Use of Digital Cameras as an Adjunct to LODOX

BY

Wayne Joseph Nin

Submitted to the Faculty of Health Sciences at
the University of Cape Town
in partial fulfillment of the requirements for the degree of
Master of Science in Medicine in Biomedical Engineering.

Cape Town, South Africa
November, 2002

Declaration

Use of Digital Cameras as an Adjunct to LODOX:

I, WAYNE JOSEPH NIN, hereby

(a) declare that:

  (i) the above thesis is my own unaided work, both in concept and execution, and that apart from the normal guidance from my supervisor, I have received no assistance except as stated below:

  (ii) except as stated below, neither the substance nor any part of the above thesis has been submitted in the past, or is being, or is to be submitted for a degree in the University or any other university.

The thesis has been presented by me for examination for the degree of Master of Science in Medicine in Biomedical Engineering.

Signed by candidate

Signature

Signed by candidate

Date

Acknowledgements

My supervisor, Prof. C.L. Vaughan, Head of Human Biology Department at the University of Cape Town for his help and guidance throughout this project.

Ben Wright and Fernando Martinez for helping me with my MS Visual C++ coding.

Harry Hall and John Markis for constructing components needed to mount the cameras on the LODOX C-arm.

Funding by African Medical Imaging (Pty) Ltd and the National Research Foundation (NRF) to make this thesis possible.

Thanks to the LODOX lab and African Medical Imaging for all their ideas in helping me put together this thesis.

Special thanks to my family as well as Sue-Ann and her family, who have supported me throughout my postgraduate studies.

Abstract

The thesis project involved the use of digital cameras to produce a three-dimensional representation of the surface anatomy of an object or person. The system was designed to be implemented as part of the LODOX digital X-ray scanner where it used its scanning motion to simultaneously capture the surface topology and corresponding radiograph of the object or person being scanned. With the ability to capture a surface with an X-ray, the LODOX system's functionality increased and therefore a higher duty cycle was achievable. This 3D reconstruction system allows physicians or surgeons to visualise underlying structures with reference to surface landmarks. This can be particularly helpful in the field of reconstructive surgery, body volume measurements or even as a communication tool for teaching or treatment planning. A review of relevant scanning techniques and scanning systems was undertaken to establish whether this technique of surfaces and X-rays are commonly used and found to be rare.

The main problems included choosing the appropriate scanning technique, and what configuration is needed to allow surface reconstruction to occur without interfering with the normal scanning operation of the LODOX machine. The system consists of two Sony DFW-VL500 digital cameras, focused onto the LODOX scanning scene, with each camera providing a separate reconstruction. This reconstruction approach was based on standard photogrammetric principles, however mathematical analysis, in the form of B-dual space first developed and tested by Bouguet as part of his Phd thesis in 1999, was used. The chosen algorithm extracted 3D data by tracking the intersection points of a moving light/laser plane through the person or object. A simple merging technique was also developed to merge the two 3D surface images to form a fuller more complete reconstruction. This final 3D surface was then overlaid onto its corresponding X-ray image and outputted as a VRML (Virtual Reality Modelling Language) version 2.0 file which is viewable with a standard VRML viewer.

A simple, low cost 3D reconstruction system which has been integrated onto the LODOX system, has been presented. The system requires very little equipment and can easily be operated. The system measures 3D surfaces with an accuracy of ≈5 mm with about 50% of the surfaces in the scene being reconstructed. The low surface reconstruction was mainly due to the limitations encountered when using a single laser fan-beam. However, it was shown that with the incorporation of an additional laser fan-beam the reconstruction of surfaces can be greatly improved. Typical reconstruction times for 3D surfaces are about 6 minutes for small objects while around 13 minutes are required for a full body.

Comparing the accuracies achieved by other 3D reconstruction systems (≈1 mm), it was found that there is still room for improvement in the proposed system. Therefore a number of recommendations are given to help improve the accuracy, reconstruction and efficient running of the system. These are: incorporating all the separate processing steps into a single software package to optimise performance, improving the edge detection algorithm to detect points more accurately and using two laser fan-beams instead of one to improve the surface reconstruction.

Table Of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The techniques of three-dimensional reconstruction from two-dimensional images have been around for many years and are well established in the industry (Motavalli, 1998). The main field is in industrial design engineering where the technique is used to re-capture and refine a prototype's dimensions. The context of this project will therefore investigate and demonstrate how one of these three-dimensional reconstruction methods can be applied or adapted to the LODOX system. The LODOX (LOw DOse X-ray) system is a newly developed linear digital X-ray imaging system. The goal was therefore to present a digitised three-dimensional surface/object with its corresponding X-ray.

The initial phase of the project involved using a reconstruction method, first implemented and tested in the PhD thesis of Bouguet (1999), as an exercise to become familiar with the theory, methods and difficulties in the field of reconstruction. This algorithm consisted of using a moving shadow, instead of the more commonly used laser, to acquire the shape of a scanned object. The use of the shadow technique was primarily due to cost savings, however the technique is relatively standard and is a good base for understanding reconstruction algorithms.

From experience gained through analyst Bouguet's methodology, a three-dimensional acquisition system was developed for the LODOX machine. The system operates by capturing a set of two-dimensional images during the scanning action of the LODOX machine and storing the images to file. A program then automatically extracts the three-dimensional geometry and surface properties of the scene using the information in the set of images. The 3D image is then generated in a 3D-image format and made available for viewing.

However, before we continue, it is necessary to define a few terms to prevent confusion when referring to the various types of images. We will define a **frame** or **image** as a single captured 2D image as produced by the digital camera, secondly the radiographic produced by the LODOX machine will be referred to as an **X-ray image** and lastly a **3D image** defines a surface in a three-dimensional coordinate system.

1.1 Motivation for the Project

The main motivational drive for this project was to give the LODOX machine more functionality. This will allow more departments to make use of the machine when it is idle, therefore improving its usefulness. Although the aim of this project is to have a 3D surface with its corresponding X-ray, the system is not limited to this configuration as the 3D surface can be captured independently and thereby eliminating any harmful effects associated with X-rays. Three-dimensional reconstruction was chosen since many fields make use of this technology.

Medically there are a large number of disciplines that could benefit by this non-invasive, tissue, limb or body measurement tool. Radiotherapy is an example in which measurements for radiation field shaping can be obtained in relation to a tumour. Here the coordinates of the patient could be used to determine the geometrical coordinates of radiation beam modifying devices (compensation filters) designed to tailor the radiation dose distribution to the individual patient (Andrew and Aldrich, 1989; Mcgivney and Cooney, 2000). It can also be used for the purpose of absorbed dose measurements ($J.kg^{-1}$) where volume can be used to estimate mass (Wilks, 1993).

The benefits of 3D surface geometry is also evident when considering the use of stereolithography in medicine. Stereolithography, also known as 3D layering or 3D printing, allows you to create solid, plastic, three-dimensional objects from a data set in a matter of hours. This has been shown in a survey conducted by Erben *et al.* (2002) where the use of rapid prototyping techniques was useful in helping with communication between medical doctors, students and patient, anatomical orientation and simulation of intervention (treatment planning).

The fields which might use prototyping techniques are in: craniomaxillofacial surgery where length and shape of grafts could be estimated, and plates can be prefabricated to hold the future bone graft (Kermer *et al.*, 1998); prosthetic developments such as complex moulding of facial prostheses (Bibb *et al.*, 2000) and; body volume measurement (Owens and Sadlier, 2000). Body volume measurement can also be used to detect underlying posture abnormalities, as in the case of scoliosis (Turner-Smith *et al.*, 1988) or evan in studies to accurately measure or visualise external landmarks in relation to underlying bones and structures in patients.

An additional benefit could be in tele-diagnosis as it could help physicians visualise how the patient was positioned during the scan to determine if particular features present on the X-ray image are from incorrect positioning or not (*e.g.* for assessing postural asymmetry). The 3D reconstruction may also help in case management (*i.e.* fitting a face to the X-rays).

Aside from the functional aspects of 3D scanning, one of the motivational factors in choosing this project was cost. A major drawback of standard active lighting systems is the cost of a motorised transport system for the object and/or active lighting of the scene (laser, LCD projector). Without this system the reconstruction accuracy is reduced, although it is considerably less expensive. Fortunately, the LODOX system provides us with an advantage in that it already has a motorised transport system built in (C-arm) together with a laser source. This therefore allows the reconstruction accuracy to be kept high with less added costs.

1.2 Description of the Problem

This thesis involves the integration of 3D surface reconstruction technology into the LODOX system. The main obstacle for this project is choosing which 3D surface reconstruction techniques to consider for optimum results, since there is a limited amount of movement and angles of view that can be used when scanning. An introduction of which parameters are needed for the reconstruction and how they are used will be reviewed. These parameters include both internal and external parameters, where internal parameters refer to camera lens distortion, optical centres, etc. and are typically constant for the specific cameras. External parameters involve the positions of the camera and light source with respect to the scene. The system is designed to maximise *line of sight* and minimise *shadows* (shadows cast by object surface) since reconstruction depends on what the cameras can see and how well the scene is illuminated by the light source. These parameters are the most variable and will be discussed in more detail later.

The above mentioned parameters will be limited by the type of scanning method and the actual mechanical structure and movements of the LODOX system. Therefore the following considerations should be addressed:

- Position of the imaging devices and light sources to maximise overall reconstruction.
- Adaptation of the reconstruction algorithm to the movement of the LODOX system.
- Reconstruction to be done without disrupting the X-ray scanning ability of the system.
- Scanning of non-reflective surfaces.
- Registration of 3D surfaces.
- Alignment of digital radiograph with the 3D surface.

1.3 Objectives of the Study

From the considerations mentioned in the previous section, a systematic plan was formulated to complete the project. The following objectives were therefore set:

1. Review current and past reconstruction algorithms to see which is the best method to use.
2. Review current and past 3D scanning systems to determine pros and cons of certain configurations.
3. Select, implement and test a reconstruction system as a stand-alone system to develop a basic system with a 3D reconstruction capability.

4. Adapt the reconstruction procedure to the LODOX system.
5. Experiment with various configurations (camera and lighting) to improve 3D reconstruction.

Chapter 2

Literature Review

As previously mentioned, techniques for three-dimensional reconstruction from two-dimensional images have been around for many years and are well established in the industry. The following paragraphs will give some background on existing research in the field of surface imaging, currently available scanning systems, and common 3D-image formats.

The literature review was prepared by processing the following literature sources:

- Journal Articles
- World Wide Web (WWW)
- Patent sites
- Theses
- Books

2.1 3D Surface Reconstruction

Fundamentally there are mainly two lighting techniques used for 3D reconstruction, namely passive and active. Passive lighting refers to the use of only the environment's natural light to illuminate the scene. Therefore by using a single camera to capture a sequence of images at different positions around the object, it is possible to reconstruct the object's 3D shape by passive cues (*e.g.* texture, shadows and other surface discontinuities). This is typically what the human brain does with images seen by our eyes. Because humans have two eyes, stereoscopic disparity can also be used to infer shape and structure. This passive cue is also known as stereometry. Currently, a substantial amount of computing power is required to accomplish this and therefore the use of this method in industry is limited (plain stereometry). The more common method used is active lighting. Here a projector or laser is used to project light patterns onto the object and then the shape of the object is extracted from these patterns by applying the principles of triangulation. A brief introduction to the two methods will be given in the following sections, but let us first consider some basic reconstruction theory.

2.1.1 Basic Reconstruction Theory

The most common sensor used in surface reconstruction is the video camera and since cameras are used in this project, the related theory will be discussed. A simple definition of a camera can be that of a device that transforms 3D scenes into 2D images. To find this transformation in terms of the camera's parameters, calibration of the cameras is required.

Camera Calibration

The most frequent camera calibration used, originally presented by Tsai (1987), is based on the pinhole camera model. However, modern day cameras use lenses to capture images and as a result, there may not be a mathematically precise principal or focal point for the real lens, as illustrated in Figure 2.1. However Debevec (1996), mentions that the locus of convergence is almost always small enough to be treated as a point although not in the case with extreme wide-angle lenses.

Typically, before we can generate any 3D data, we need to find the intrinsic and extrinsic camera parameters. There are three main intrinsic camera parameters that need to be found: focal length f, image centre c and radial distortion k_c.

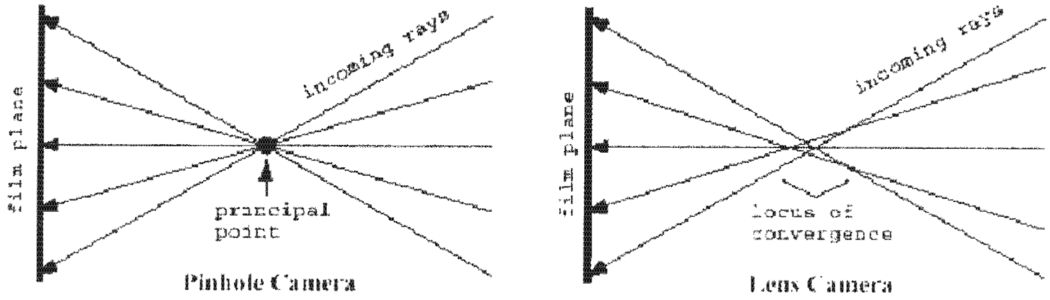


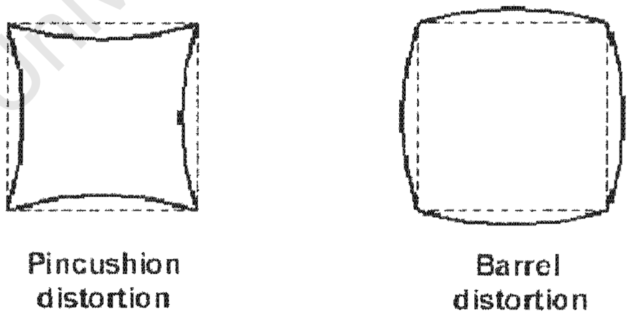
Figure 2.1: Difference between a pinhole and lens camera (Debevec, 1996).

The focal length f (sometimes called the principal axis distance) is a parameter of the lens and is defined as the distance between the lens (focal point) and the point where the image is in focus.

The image centre c is the centre of radial lens distortion and the piercing point of the camera coordinate frame's Z axis with the camera's sensor plane.

Radial distortion that causes the image to bulge toward the centre is called barrel distortion, and distortion that causes the image to shrink toward the centre is called pincushion distortion (see Figure 2.2).

These parameters can be extracted by imaging a calibration pattern with known geometry(*e.g.* checker board) on a two or three dimensional plane using the set of equations 2.1 (Tsai, 1987).


Pincushion distortion

Barrel distortion

Figure 2.2: Two types of radial distortion.

$$
\begin{aligned}
\bar{a} &= \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \\
\bar{b} &= \begin{bmatrix} b_x \\ b_y \end{bmatrix} = \bar{x}(1 + k_c\|\bar{a}\|^2) \\
\bar{p} &= \begin{bmatrix} p_x \\ p_y \end{bmatrix} = \begin{bmatrix} f_x b_x + c_x \\ f_y b_y + c_y \end{bmatrix}
\end{aligned}
\tag{2.1}
$$

where f_x and f_y are derived from the focal length divided by the image sensor's dimensions in the X and Y axes.

Combining the the set of equations 2.1 gives:

$$
\begin{bmatrix} sp_x \\ sp_y \\ s \end{bmatrix} = \begin{bmatrix} f_x r & 0 & c_x & 0 \\ 0 & f_x r & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}
\tag{2.2}
$$

where $r = (1 + k_c\|\bar{a}\|^2)$

The 3 x 4 matrix in Equation 2.2 is usually referred to as the camera intrinsic matrix $[INT]$ because it depends completely on the optical and geometric characteristics of the camera.

In practice, it is no trivial matter to measure the 3D coordinates of a point relative to the camera coordinate system so often another global coordinate system, to which the object can be easily referred is used. This shift to the world coordinate system is done through the extrinsic camera parameters. The extrinsic parameters are defined by R and T as seen in Equation 2.3. The rotation matrix R expresses three elementary rotations of the co-ordinate axes, rotations along the X, Y, and Z axes are termed pan, tilt, and roll, respectively. The translation vector T gives three elements of the translation of the origin of the world coordinate system with respect to the camera coordinate system (Figure 2.3).

$$
\mathbf{X}_{\text{world}} = \mathbf{R}.\mathbf{X}_{\text{camera}} + \mathbf{T}
\tag{2.3}
$$

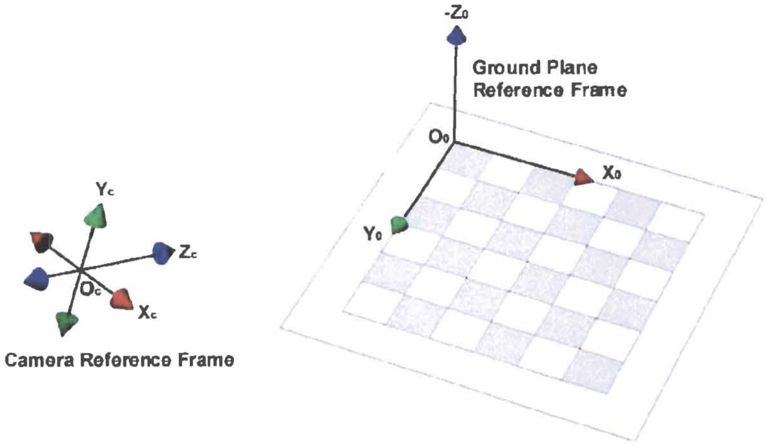


Figure 2.3: Diagram showing the camera and ground reference frame.

or in the matrix form of

$$\begin{bmatrix} X_{\text{camera}} \\ Y_{\text{camera}} \\ Z_{\text{camera}} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_{\text{world}} \\ Y_{\text{world}} \\ Z_{\text{world}} \\ 1 \end{bmatrix} \tag{2.4}$$

The 4×4 matrix in Equation 2.4 is often referred to as the camera extrinsic matrix $[EXT]$ because it depends completely of the relative position and orientation of the camera with respect to the world coordinate system.

By combining Equations 2.2 and 2.4, we get the following equation:

$$\begin{bmatrix} sp_x \\ sp_y \\ s \end{bmatrix} = \mathrm{M}_{3\times4} \begin{bmatrix} X_{\text{world}} \\ Y_{\text{world}} \\ Z_{\text{world}} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathrm{M}_{11} & \mathrm{M}_{12} & \mathrm{M}_{13} & \mathrm{M}_{14} \\ \mathrm{M}_{21} & \mathrm{M}_{22} & \mathrm{M}_{23} & \mathrm{M}_{24} \\ \mathrm{M}_{31} & \mathrm{M}_{32} & \mathrm{M}_{33} & \mathrm{M}_{34} \end{bmatrix} \begin{bmatrix} X_{\text{world}} \\ Y_{\text{world}} \\ Z_{\text{world}} \\ 1 \end{bmatrix} \tag{2.5}$$

where the perspective transformation matrix $\mathbf{M}_{3\times4} = [INT][EXT]$.

Substituting for s we get a set of linear equations having the form

$$Aa = b \tag{2.6}$$

and by using multiple points one can thus solve the perspective transformation matrix $\mathbf{M}_{3\times4}$ using the least squares method given by:

$$a = (A^T A)^{-1} A^T b \tag{2.7}$$

Therefore by knowing the intrinsic and extrinsic camera parameters one can determine the 3D position of a point on a ray projection (from the optical centre of the camera) given the 2D position of the point (pixel) on the image plane. It should be noted that the above equations were derived using projective space $\mathfrak{P}$. Points in the projective space are expressed in homogeneous (also projective) coordinates with the one-to-one mapping from Euclidian space ($\mathfrak{R}$) into Projective space ($\mathfrak{P}$) (Triggs and Mohr, 1996) given as follows:

$$\mathfrak{R} \in \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \rightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{bmatrix} \in \mathfrak{P} \tag{2.8}$$

The introduction of projective space and homogeneous coordinate vectors made the expressions simpler.

Because of the wide use of cameras in 3D reconstruction there is some free calibration software that can be downloaded from the Internet, which will do the calibration. Most are in the form of a toolbox for Matlab (Calibration, 2002) but there are some standard Windows packages like Microsoft Easy Camera Calibration Tool.

2.1.2 Passive Lighting Scanning Methods

An example of a passive lighting technique is plain stereometry (*i.e* only cameras are used in the setup). Stereometry is a technique in which one can deduce the 3D shape of an object from a stereoscopic image pair by means of triangulation. Suppose we

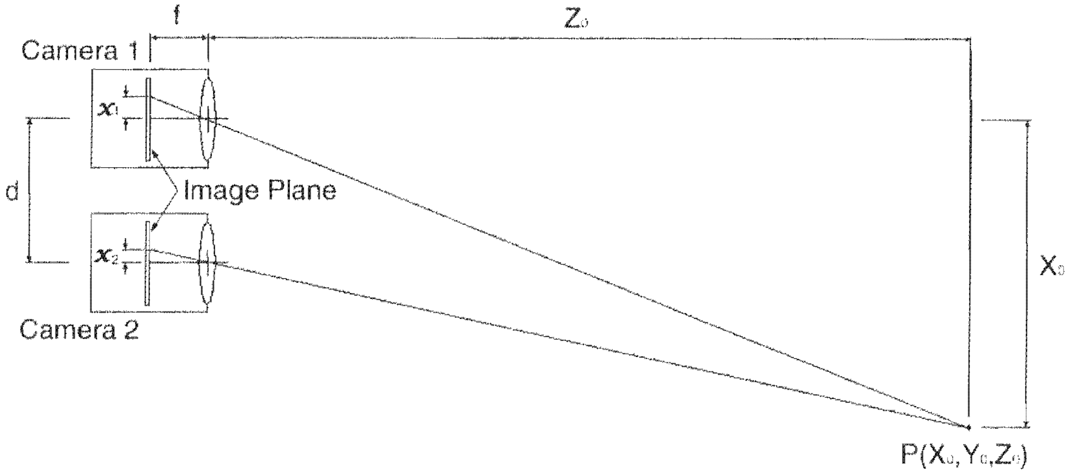


Figure 2.4: Cameras in stereoscopic configuration (Gao, 2001).

have two cameras configured as shown in Figure 2.4 with both cameras having a focal length f and separated by distance d (considering only the $X - Z$ plane). Due to the reflection of natural light from point P, the point is projected through the lens of the cameras onto the image planes represented by points x_1 and x_2. To determine the 3D coordinate of point P, a common reference system needs to be used. This reference system is chosen such that its origin is at the centre of camera 1's lens.

It can easily be shown that:

$$X_0 = x_1 \frac{Z_0}{f} = x_2 \frac{Z_0}{f} - d \tag{2.9}$$

Similarly we can determine the Y-coordinate using Equation 2.10, where y_1 and y_2 are the projections onto the image plane in the Y-axis (not shown).

$$Y_0 = \frac{Z_0}{f} y_1 = \frac{Z_0}{f} y_2 \tag{2.10}$$

and solving for Z_0 in Equation 2.9 gives:

$$Z_0 = \frac{fd}{x_2 - x_1} = \frac{fd}{y_2 - y_1} \tag{2.11}$$

With Z_0 solved, the X and Y coordinates of point P are determined by substituting Z_0 into Equations 2.9 and 2.10. To get the 3D surface of the whole scene we use the above equations for each pixel on the image plane. This method of reconstruction is known as the decomposition method.

Another method for stereoscopic reconstruction, called the direct method, is to use the perspective transformation matrix $\mathrm{M}_{3\times4}$ (Equation 2.5) for each camera. Knowing a common point in each camera, the following equation can be derived:

$$\begin{bmatrix} \mathrm{M}_{11} - \mathrm{M}_{31}p_x & \mathrm{M}_{12} - \mathrm{M}_{32}p_x & \mathrm{M}_{13} - \mathrm{M}_{33}p_x \\ \mathrm{M}_{21} - \mathrm{M}_{31}p_y & \mathrm{M}_{22} - \mathrm{M}_{32}p_y & \mathrm{M}_{13} - \mathrm{M}_{33}p_y \\ \mathrm{M'}_{11} - \mathrm{M'}_{31}p'_x & \mathrm{M'}_{12} - \mathrm{M'}_{32}p'_x & \mathrm{M'}_{13} - \mathrm{M'}_{33}p'_x \\ \mathrm{M'}_{21} - \mathrm{M'}_{31}p'_y & \mathrm{M'}_{22} - \mathrm{M'}_{32}p'_y & \mathrm{M'}_{13} - \mathrm{M'}_{33}p'_y \end{bmatrix} \begin{bmatrix} X_{\text{world}} \\ Y_{\text{world}} \\ Z_{\text{world}} \end{bmatrix} = \begin{bmatrix} p_x - \mathrm{M}_{14} \\ p_y - \mathrm{M}_{24} \\ p'_x - \mathrm{M'}_{14} \\ p'_y - \mathrm{M'}_{24} \end{bmatrix} \tag{2.12}$$

where ' denotes the parameters for the second camera.

The real world coordinates can then be solved using the least-squares method described in Equation 2.7. The advantages of this method over the decomposition method is that it's simple, both cameras can be positioned anywhere rather than on the same horizontal plane and that it yields better accuracies (Cheng, 1997).

Passive techniques are widely used in aerial photogrammetry because the stereo images usually have many features that can be easily matched (*e.g.* trees, buildings, shadows, etc.). The main problem arises when scanning human figures which do not have a lot of texture, as in this project. If a computer cannot easily determine large amounts of corresponding points in both images when doing the triangulation, reconstructions will be poor or inaccurate. Therefore purely passive techniques for 3D reconstruction for humans are generally rare. From the literature search conducted only Morency *et al.* (2002) has attempted a purely passive system for facial reconstruction. Their system used a 'Brightness Change Constraint Equation' to attempt to match points.

More commonly, researchers have used a number of active lighting techniques to make it easier to pinpoint similar points in each camera image to overcome this problem. This leads us to the next lighting method where the primary reconstruction technique requires the use of an external light source.

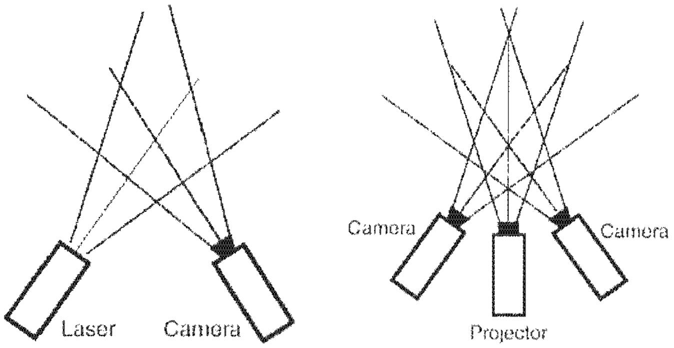


Figure 2.5: Laser-camera and cameras-projector configurations (Siebert and Marshall, 2000).

2.1.3 Active Lighting Scanning Methods

Over the years, numerous techniques have evolved to infer surface topology or surface contour by means of structuring a light system and measuring the corresponding distortion in the light structure. Light structures have been based on numerous mechanisms such as a spot, light stripe (Chu *et al.*, 2001; Addleman and Addleman, 1985; Curless, 1999; Wilks, 1993; Turner-Smith *et al.*, 1988), multi-stripes (Velkley and Oliver, 1979; TriForm, 2001; Daley and Hassebrook, 1998), grid projection (Barros *et al.*, 2002), gray-scale encoding (Naftel and Mao, 2002; Siebert and Marshall, 2000), colour encoding (Koschan *et al.*, 1996; Zhang *et al.*, 2002), time of flight (Fenster *et al.*, 1995), and interferometry (Creath and Wyant, 1992; Paakkari, 1998). An introduction to the methodologies used in some of these techniques is given below. The majority of these approaches will usually employ a laser-camera or camera-projector configuration (Figure 2.5).

Stereo Techniques with Structured Lighting

Light Source Manipulation: Crossing the boundary from passive to active lighting we get a system like that of Debevec *et al.* (2000), they use 'Reflectance Fields' to accomplish stereo matching. This system works by manipulating the light source, either in its position or its intensity to generate additional passive cues (shadows and surface discontinuities) therefore improving the quality of the reconstructions (Figure 2.6).

Figure 2.6: Images taken with the light source in five different positions (Debevec *et al.*, 2000).

Light Stripe: The more common approach when using two stereo images is to project a light stripe on the object so to help with stereo matching (Abi-Rached *et al.*, 1998). Using a standard stereo geometry configuration the epipolar lines match the image line, the matching process can be simplified as corresponding pixels can only be found in the same row in both images. If a sequence of stereo images are captured with a light stripe moving over the surface, common points can be matched by tracking the light stripe on each horizontal scan line in each image, thereby producing a reconstruction of the surface (Figure 2.7). The advantage of this system is that if the cameras have a high capture rate and/or the stripe is moved slowly over the scene, a high density range map can be achieved. The disadvantage however is that this method is time consuming in both capturing ($\approx$ 10s) and processing (*i.e.* more than one image needs to be processed) as well as errors introduced by patient movement.

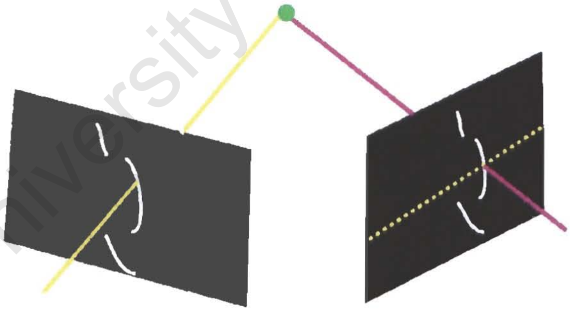


Figure 2.7: Range through triangulation using one stripe (Abi-Rached *et al.*, 1998).

Multiple Light Stripes: Structure light with multiple light stripes have drawn attraction because light stripe pattern is simple for image processing and it does not require scanning or moving mechanisms (Figure 2.8). However, since multiple light stripes are projected simultaneously, the index of a light stripe on which a surface point lies has to be identified in order to establish the association between 2D image and 3D

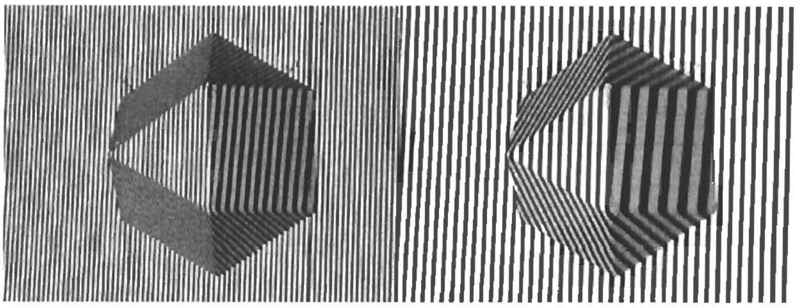


Figure 2.8: Two multi-stripe patterns with different spatial frequencies (Daley and Hassebrook, 1998).

coordinates. For complex shapes this might become difficult due to line breaks but various methods are used to overcome this. One is to use reference stripes (wider stripe than the rest) spaced equally within the stripe sequence (TriForm, 2001). Other option is to use colour coded stripes as done by Zhang *et al.* (2002) (See Figure 2.9). The advantage of this is that the acquisition speed is very fast (<1s) but the resolution of the range map is dependant on the spatial frequency of the stripe pattern. If using colour coded stripes, resolution will also depend on the ability of the cameras to separate the colours. Inexpensive cameras tend to have lower accuracies when separating colours.

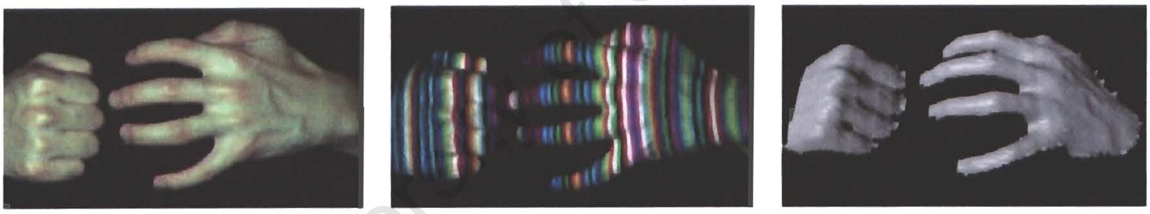


Figure 2.9: Using colour stripes to identify stripe index (Zhang *et al.*, 2002).

Speckle Texture Projection: Another form of structured light is called speckle texture or speckle texture projection. It makes use of a random pattern to allow extraction of the exact points in two images. The stereo matching is done by comparing a pattern in a small image patch around a pixel in one image and matching it to the other image patch. Prior to matching by a least squares method, one of the image patches is modified by applying a transformation (translation, rotation, sheering and scaling) (D'Apuzzo, 1998). The 3D map is then computed by triangulation (space intersection). The method has the advantage of fast acquisition time and dense range data but the processing time required is quite lengthy. However, numerous algorithms have

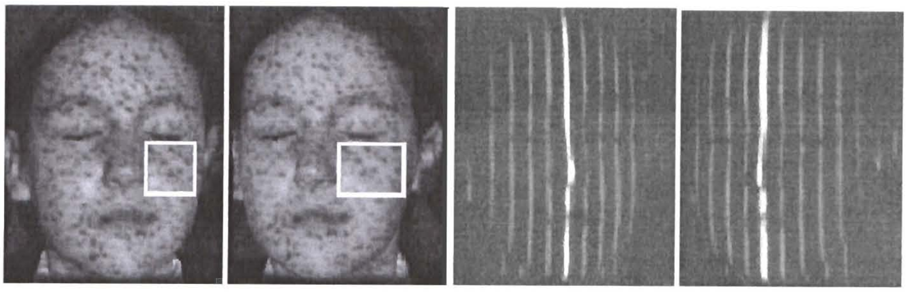


Figure 2.10: Stereo pairs acquired with texture projection and multi-stripe projection (Naftel and Mao, 2002).

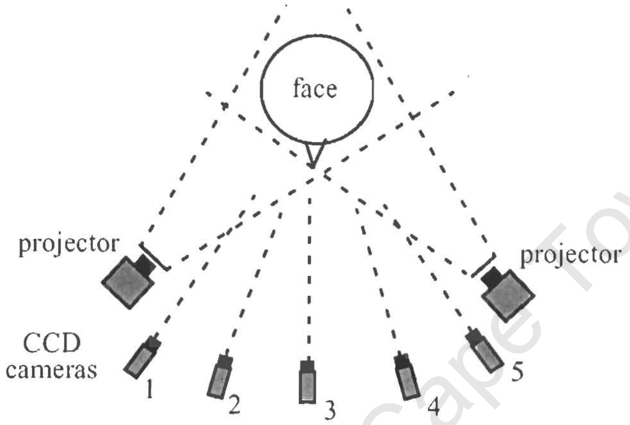


Figure 2.11: Multi-camera setup for improved coverage (D'Apuzzo, 1998).

been developed to improve on the processing times. One is to use a combination of texture projection and multi-striping to narrow search areas (Naftel and Mao, 2002) as shown in Figure 2.10.

A common disadvantage of using any of the stereo photogrammetry techniques mentioned above, is that it can only reconstruct points which can be seen in both images. To improve on this one can increase the number of cameras used as shown in Figure 2.11, therefore having better coverage of the object.

Single Camera with Structured Lighting

Laser Planes: Laser lighting techniques use projective geometry to solve for 3D coordinates. Instead of using stereo techniques (triangulation between two rays), one makes use of triangulation between rays and planes in space. This technique only

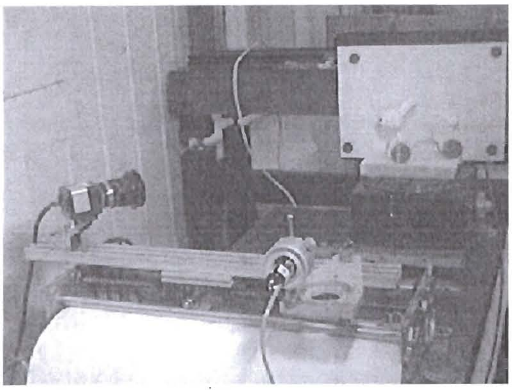


Figure 2.12: Example of a laser plane rangefinder, the camera and laser can be seen mounted on a movable track (Sonka *et al.*, 1998).

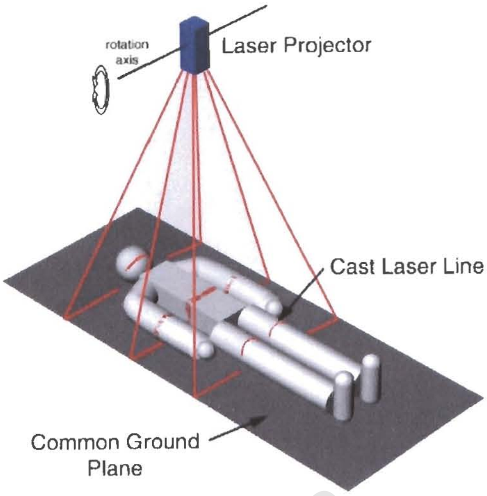


Figure 2.13: Rotating plane of light cast by laser.

requires one camera but the parameters of the plane needs to be known. Solving for the Z component of the point is achieved using the fundamental triangulation equation between a ray and a plane in space given below:

$$Z = \frac{1}{\bar{w}.\bar{x}} \tag{2.13}$$

where $\bar{w}$ represents the plane's coordinate vector in space and $\bar{x}$ the ray as a homogeneous coordinate vector.

Typically a laser projects a sheet of light onto the scene, producing an intersection line between the laser plane and the object (shown as the 'Cast Laser Line' in Figure 2.13. Using the intersection line to define the ray parameters and knowing the plane's parameters, using Equation 2.13, the depth can be computed. Substituting back into the homogenous coordinate vectors of the ray (Equation 2.10), the X and Y coordinates can be calculated.

To obtain a complete range map of the object the plane must be moved through the object. For relatively small objects, a turntable is used to rotate the object through the light plane whereas for larger objects, the camera and laser unit move together over the surface (Figure 2.12). Both these methods usually require some sort of accurate

Figure 2.14: Technique for shadow reconstruction (Bouguet, 1999).

movement control system to track the position of the light plane relative to the object in time (*i.e.* particular angle of rotation for turntable or the positional shift for laser-camera unit). A more common technique is to rotate the light source along one of its axes thereby allowing the light plane to sweep through the object (Figure 2.13). This can be accomplished by accurately controlling the movement of the light-source or a deflector (*e.g.* mirror). If however the light plane cannot be accurately controlled, the problem of a varying light plane can be overcome by knowing the centre of the light source $\bar{X_S}$, a common plane $\bar{w_g}$ (ground plane in Figure 2.13) to which the light plane $\bar{w_l}$ intersects and the line produced at this intersection $\bar{\lambda g}$. To infer the light plane parameters we use the following equation (Bouguet, 1999):

$$\bar{w_l} = \bar{w_g} + \alpha_g \bar{\lambda g} \tag{2.14}$$

where

$$\alpha_g = \frac{1 - \bar{w_g}.\bar{X_S}}{\bar{\lambda g}.\bar{X_S}} \tag{2.15}$$

Typically the common plane is set to be the ground plane to simplify the calculations. This has already been shown in Section 2.1.1 where a new coordinate system, other than the camera coordinate system, is defined.

Shadow Planes: Alternatively, weak structured lighting or shadow reconstruction (Bouguet, 1999) can be used. This method is similar to the laser system but involves the tracking of a shadow plane instead of a light plane (Figure 2.14). This is substantially cheaper than a laser system but the accuracy of the system is reduced due to the 'sharpness' of the shadow edge. Additionally, one could consider using the in-

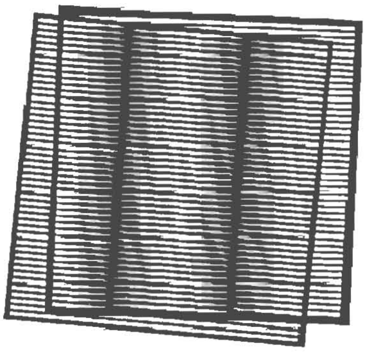


Figure 2.15: Two line patterns superimposed.

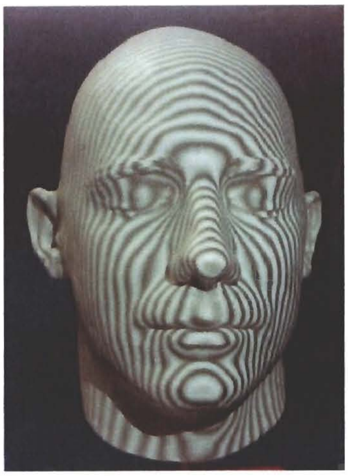


Figure 2.16: Moiré fringe contouring superimposed over a model head (3D-MATIN, 2002).

verse, where a white light plane be tracked in a dark environment instead of a shadow plane in a brightly lit environment (Abi-Rached *et al.*, 1998).

Multi-Planes: The single laser plane technique can also be expanded to a multi-plane technique, where multiple planes could be projected onto the surface, similar to the multi-stripe system mentioned previously. This will involve calculating each plane's parameters from the scan images or in a calibration step. The advantage of using multi-planes is capture times are quick but spacial resolution issues arise as mentioned earlier with the multi-stripe pattern.

Moiré Fringe Contouring: Another method that uses striped pattern projection is called Moiré Fringe Contouring (Creath and Wyant, 1992). Moiré fringes are formed when one line pattern is superimposed upon a similar line pattern as shown in Figure 2.15. These fringes represent contours, which can be used to extract 3D surface data as seen in Figure 2.16.

Moiré topography is used to measure and display the 3D form of an object. The formation of depth contours by arranging the light source and viewing point at the same distance from the grid is a key feature of Moiré topography. This concept was formulated independently and concurrently by Meadows *et al.* (1970) and Takasaki (1970).

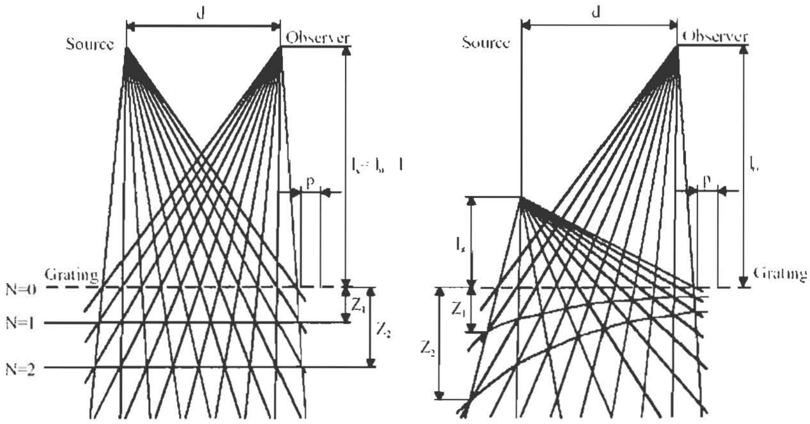


Figure 2.17: Schematic showing the variation of the contours levels with even and uneven light-observer positions (Paakkari, 1998).

The shadow and projection methods are the commonly used Moiré techniques for 3D shape measurements (Pekelsky and van Wijk, 1989).

In the case of the shadow Moiré technique one grid is used in combination with a light source and recording equipment, usually a CCD camera. The Moiré fringe is generated from the shadows cast on the object and the actual grid obstructing the observers view. The grid needs to be installed close to the object to be measured and should be at least as large as the object. Figure 2.17 shows how the light rays are projected and the light received from the observer for two different setups. As can be seen, when the light source and the observer are in line, the contours Z_i are parallel to the grid, whereas not so in the other case. Z_i is calculated using the following equation for parallel contours since contours are not equally spaced:

$$Z_N = \frac{Npl}{d - Np} \tag{2.16}$$

where N is the contour order, p is the pitch of the grating, d is the distance between the observer and source and l the distance from the grating to observer. The need for a large grid and the limitations on height variations in the object due to the proximity of the grid are sometimes treated as disadvantages of the shadow Moiré technique.

The optical projection Moiré technique however is implemented using physical gratings in both the projection and observation systems (Figure 2.18). The observation system is usually a twin of the projection system. The Moiré contours captured by

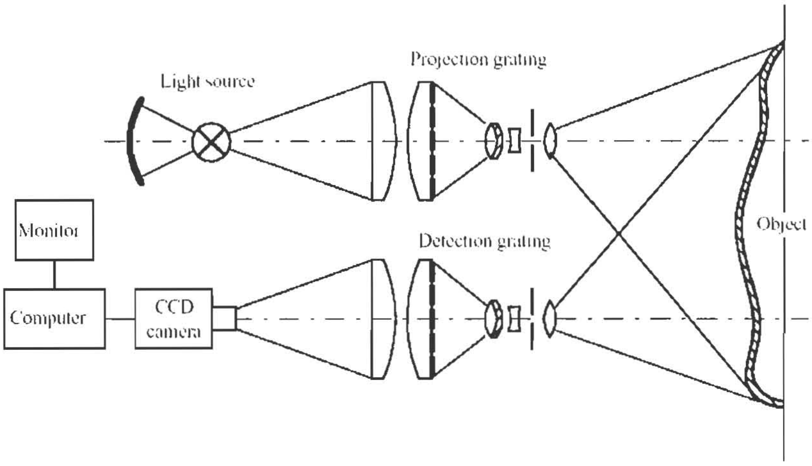


Figure 2.18: A typical projection Moiré system setup.

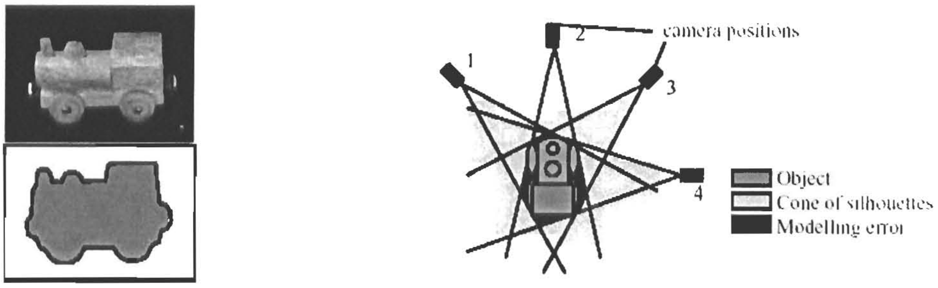
the camera have properties analogous to those formed by the shadow technique and the geometrical interpretation of contour formation remains valid. The Nth contour is however given by Takeda (1982) as follows:

$$Z_N = \frac{Npl_{rp}(l_{rp} - f)}{df - Np(l_{rp} - f)} \tag{2.17}$$

where N is the contour order, p is the pitch of the grating, d is the distance between the optical centre of the lenses and l_{rp} the distance from the centre of the lenses to a reference plane and f the focal length of the lenses.

The temporal phase shifting technique (TPS), which analyses three or more interferograms acquired serially over time, forms the basis for most modern interferometers. TPS produces the highest spatial resolution and highest accuracy in the case of static wave fronts, so unless some good reason exists to use an alternative analysis technique, TPS should be used (Paakkari, 1998).

Volume Carving: This technique uses the most simplest of structured lighting, namely back-lighting. An approach for the construction of volume models using multiple views was first described by Martin and Aggarwal (1983). They introduced the method of 'occluding contours' which uses the object silhouettes and the related cam-



(a) (b)

Figure 2.19: (a) Extraction of silhouette from image. (b) Modelling error with limited views. (Niem, 1994)

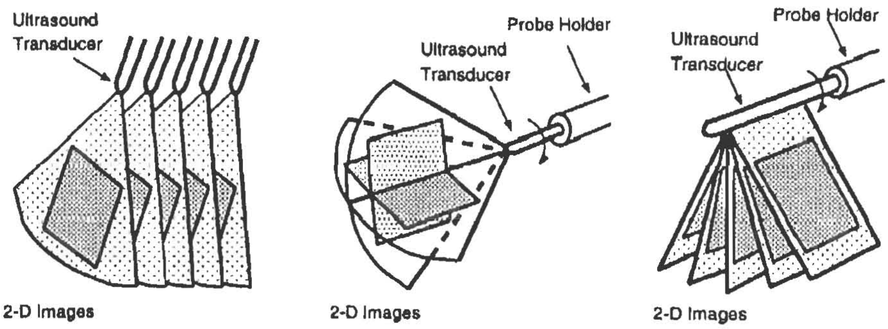
era parameters to construct a volume model of the real object. The principle of this method can be divided into the following steps (Niem, 1999; Lyness *et al.*, 1999):

1. Define a volume representation (e.g. octree - a tree of recursively subdivided cubes down to the finest resolution)
2. Extract the silhouette from an input image
3. Project the silhouette onto the volume representation in the 3D world
4. Discard or 'carve off' the volume that falls outside the silhouette from the volume representation
5. Repeat for images taken from multiple angles

As more of the silhouettes are used the volume becomes a closer approximation of the object. Applying a surface to the volume set the object can be visualised. The limitations of such a silhouette based approach are primarily the inability of the algorithm to detect object concavities (like the hollow in a cup) and with limited views, models may not be correctly reconstructed (See Figure 2.19b).

Non Optical Methods

Time of Flight: Time of Flight systems make use of radio waves to detect distance (e.g. radar, ultrasound, sonar, etc.). These systems measure the time it takes for a wave to travel to and from the surface and by knowing the velocity at which the wave travels, distance can be calculated using the following equation:



a) b) c)

Figure 2.20: Schematic diagram of three possible methods for acquiring 2D ultrasound images necessary for 3D reconstruction. a)parallel scanning; b) rotation of the probe to produce a sector of planes; and, c) rotation of the probe to produce an intersecting set of planes. (Fenster *et al.*, 1995).

$$\text{distance to surface} = \frac{\text{time} \times \text{velocity of wave}}{2} \tag{2.18}$$

One of the more interesting fields is that of 3D ultrasound imaging. 3D ultrasound data can be obtained using various methods as shown in Figure 2.20. Ultrasound may be used for external surface measurements but is more commonly used for internal measurement because greater information can be acquired from the reflections from small changes in the mediums densities (*i.e.* only one surface would be captured with an air to tissue interface). The problem with these methods is that the physical apparatus can be expensive and some deflection mechanism is required, which makes them sensitive to mechanical influences. They are also usually slow. These techniques are beyond the scope of this project so will not be discussed further.

2.2 Examples of Scanning Systems

This section will give an overview of some of the scanning systems that have been developed for scientific experiments as well as ones that are commercially available. Although the surface reconstruction algorithms can be used for most camera positions, the accuracy depends on the viewing angle. This section is therefore to show

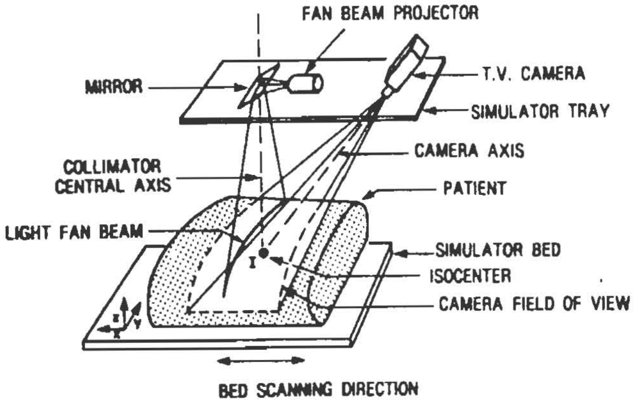


Figure 2.21: Schematic diagram of the contour acquisition system showing the try mounted camera and fan beam generator and patient lying on the simulator bed (Andrew and Aldrich, 1989).

the various system configurations used for the various surface measurement methods.

2.2.1 Experimental Systems

An early contour acquisition system is shown schematically in Figure 2.21. This system was developed by Andrew and Aldrich (1989) to generate a template for making a radiation compensator from sheet lead. As shown the system uses a single camera imaging a defined laser plane, in this case perpendicular to the ground plane. Surface reconstruction is achieved by moving the patient through the laser plane by shifting the bed. The angle of the camera is critical in obtaining a good depth perspective, for this system, the angle of the camera was 30° to the horizontal and yielded accuracies of ±2 mm. In addition, the system can be easily converted to a sweeping fan beam by controlling the angle of the mirror, which is what is done in other systems.

Figure 2.22 demonstrates the typical configuration of a stereo system. Where two cameras, separated by a distance, are used with a projector projecting a colour multi-stripe pattern onto the scene. A research laboratory, 3D-MATIC, of the Department of Computing Science at the University of Glasgow has also developed a similar system. Their system, called C3D, uses random texture and least squares matching to locate stereo point (Figure 2.23). This system produces 3D models with an accuracy of around 5 mm (Siebert and Marshall, 2000).

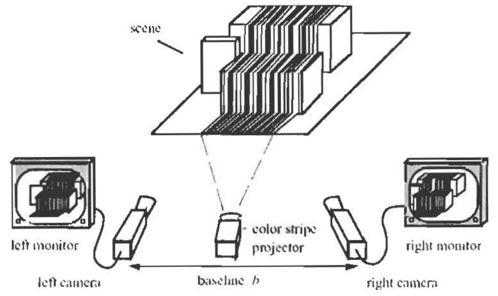


Figure 2.22: Principle of the stereo system using active colour illumination (Koschan *et al.*, 1996).

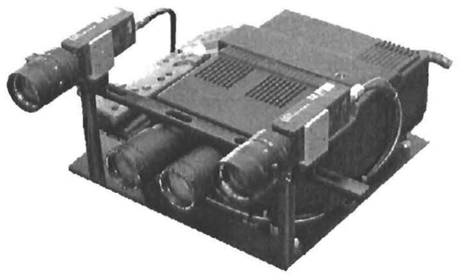


Figure 2.23: C3D modular 'pod' developed by 3D-MATIC (3D-MATIN, 2002).

2.2.2 Commercial Systems

From the literature review on the scanning methodologies it was found that Cyberware is one of the preferred scanning systems employed for doing anthropometric measurements (D'Apuzzo, 1998; Lemchen, 2000). This company designs whole systems that include both the reconstruction and motorised transport systems. As can be seen from the examples in Figure 2.24, these systems make use of the laser-camera configuration. All employ a laser plane to be moved through the object whether it be translation, rotation or plane sweeping or a combination (Section 2.1.3). These systems have accuracies of about 1mm (Nebel *et al.*, 2001). Some reconstructions from these systems are shown in Figure 2.25.

Similarly, a company called Wicks and Wilson has developed a system for digitising the surfaces of real objects. Their system called TriForm^{TM}, makes use of mirrors to get two viewing angles and a multi-plane structured lighting technique (Section 2.1.3), to reconstruct objects. As can be seen in Figure 2.26, the use of reference stripes helps to index the stripes to help with plane indexing [1]. Accuracies of ±2 mm for a single view is expected while for a full body image, ±8 mm are expected when all images are aligned (*i.e.* multiple views). Some reconstructions from these systems are shown in Figure 2.27.

Another product called Lightscribe, reconstructs objects by extracting the silhouette of the object as it rotates on a turntable. The silhouette is enhanced by using a back light

---

[1] Note how this configuration could easily be used as a stereo technique.

Figure 2.24: Cyberware 3D scanning systems (Cyberware, 2001).

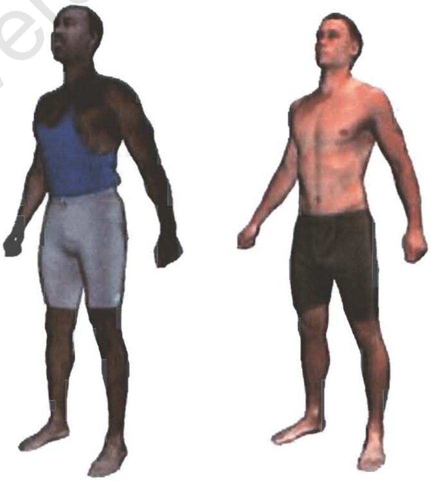


Figure 2.25: Cyberware 3D scanning systems results (Cyberware, 2001).

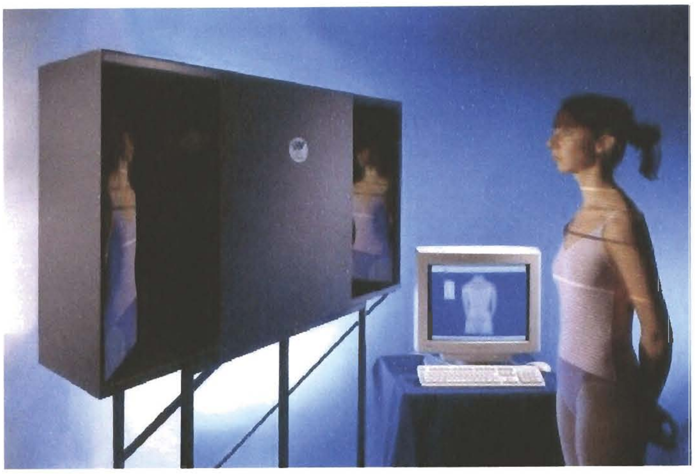


Figure 2.26: TriForm non-contact 3D image capture system (TriForm, 2001).

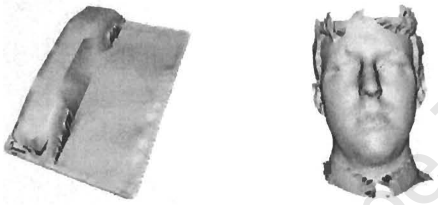


Figure 2.27: Triform 3D scanning systems results (TriForm, 2001)

as shown in Figure 2.28. However, because this method cannot reconstruct concave surfaces on the volume, as mentioned previously in Section 2.1.3, another surface extraction technique is needed to help refine the reconstruction. In this case further refinement is done by painting the object with a stripe of laser light. The company claims to provide sub-millimetre accuracies from the system.

2.3 X-Rays and 3D Surfaces

The system or method of incorporating correlated three-dimensional surface data and an X-ray image has never really been exploited. The first documented use of this technique was in mid-1998 where it was used in the orthodontic diagnostic and planning industry (Lemchen, 2000). This particular system makes use of surface scanning technology from Cyberware. A more common application where X-rays and 3D surfaces have been used is in maxillofacial surgical planning (3D-MATIN, 2002; Gross, 1998). An example is a system developed by Ayoub *et al.* (1996), which

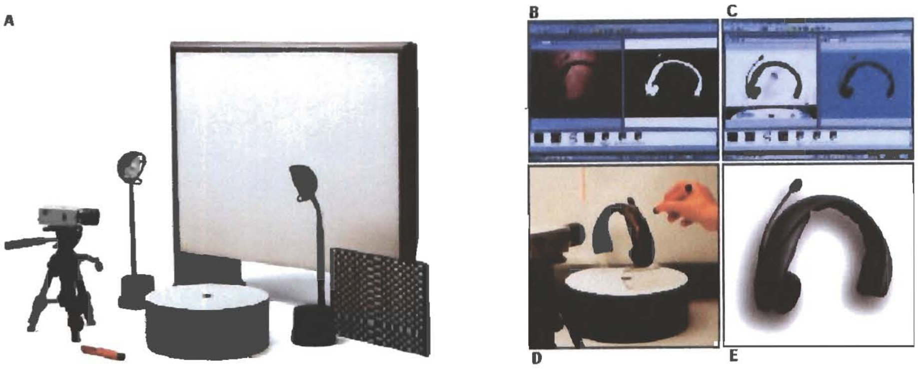


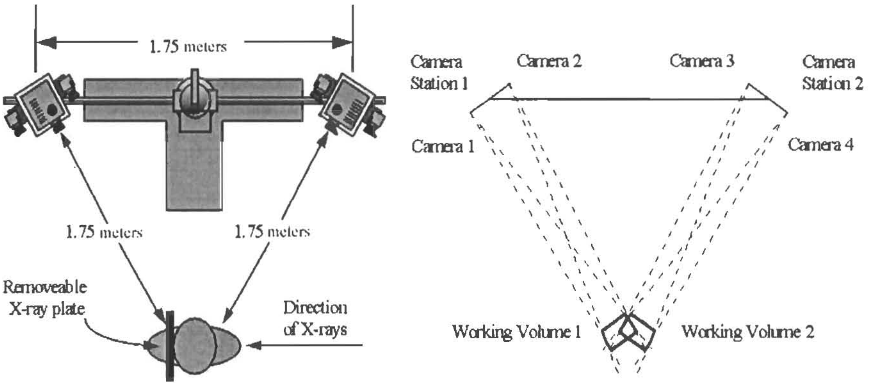
Figure 2.28: Main system components of LightScribe 3D with screenshots and example (LightScribe, 2001).

used two C3D systems, developed by the Turing Institute, to capture two stereo pair images from two sides of the patient together with an X-ray, Figure 2.29 (a). By then superimposing the soft and hard tissue (see Figure 2.29 (b)) an image is formed that can help in surgical treatment planning. Other than these mentioned cases, no other system has yet been found.

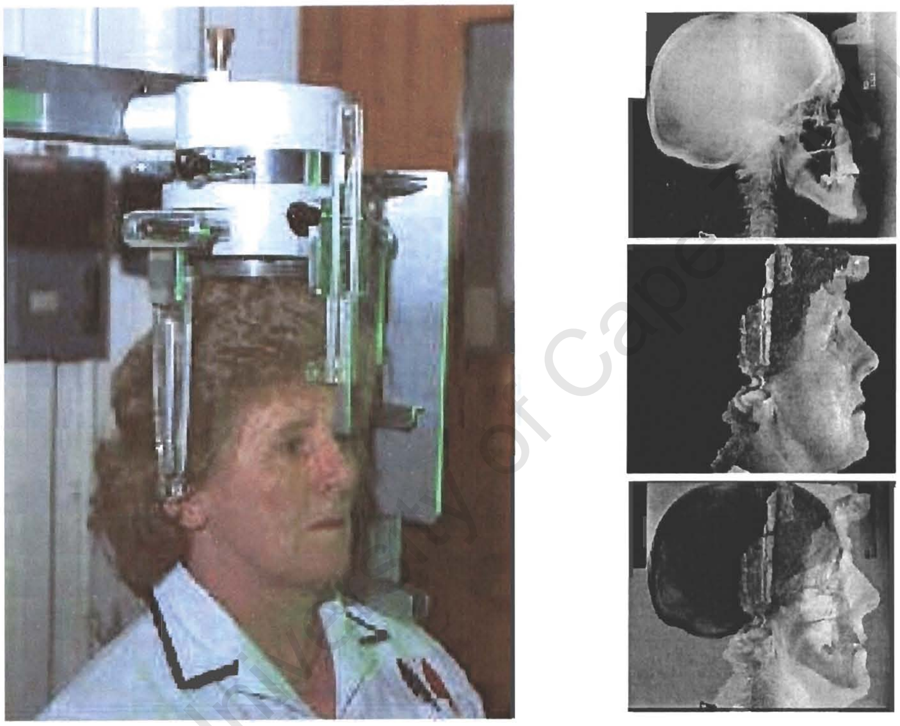
Since we are working with digital X-rays in this project, an investigation into how to access medical images stored on digital imaging modalities is performed. Currently, there are a number of formats available because each individual manufacturer produces their own patient file formats for their particular digital X-ray system. However, there is now a big effort in the medical imaging community to define certain standards for patient imaging files. Some of the more common medical image format standards are (Clunie, 2001):

- Digital Imaging and Communications in Medicine (DICOM)
- American College of Radiologists (ACR) and the National Electrical Manufacturers Association (NEMA)
- Papyrus
- Interfile V3.3

For more information on the image format standards refer to Clunie (2001). A short introduction to DICOM is given in Appendix A.2 since this is the format used by the LODOX system.



(a) System configuration.



(b) Patient setup.

(c) X-ray and 3D surface with superimposed image.

Figure 2.29: System showing the use of 3D surfaces and X-rays (3D-MATIN, 2002).

2.4 Meshing Algorithms

The process of reverse engineering usually results in the generation of precise 3D data points or point cloud from an algorithm. However, visualising a complex arrangement of points may seem complicated. To make visualisation easier the point cloud is converted to a geometrical model of vertices and surfaces. A surface can

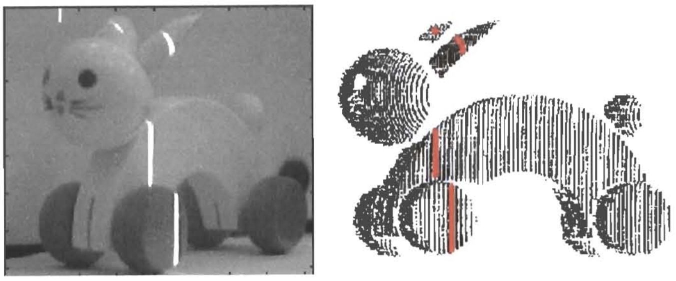


Figure 2.30: Single frame (left) from a image sequence with the extracted points from this image (in red) shown in complete point cloud image (right) (Sonka *et al.*, 1998).

be represented by means of a number of small planar triangles (polygons) each connecting 3 data points, but a meshing algorithm needs to be used to choose these points. The choice of algorithm is highly dependent on how the points are arranged. If the points are structured, *i.e.* the points are organised into a rectangular grid, the meshing is relatively straightforward. An example of such a system is one utilising a photosensitive array (*e.g.* a 2D-charge couple device or CCD) with structured light to generate an image/grid. However when we have unstructured points (*i.e.* no ordering of points), like in the case of a touch probe system which outputs only a point cloud, meshing becomes complicated and time consuming.

2.4.1 Structured Data

The structured data format is of particular interest to this research, since many surface measurement systems use CCDs to obtain range maps composed of a regular grid of data points. A polygonal surface can be created in a straightforward manner by linking data points in a neighbourhood to form the mesh (Niem, 1999; Sun *et al.*, 2001). If an object has been digitised through the acquisition of multiple range maps, then an appropriate registration and alignment technique must be implemented to merge the adjoining polygonal domains.

Simple Meshing Algorithm

Systems which produce structured data are techniques that use multi-stripe patterns or image sequences. Multi-stripe patterns structure the data with the stripe index of the pattern while image sequences structure data by time (*i.e.* speed of the camera, fps)(Figure 2.30). Assuming an image size with a height of 100 pixels, each image therefore generates 100 data points per frame. Storing this data in a matrix that represents the number of data pixels as the rows, and the frame number as the column, we get a 2D matrix of point positions. Considering a single entry in this matrix, there are usually 8 neighbouring points to which 8 polygons can be fitted. Polygons are only assigned if they follow certain rules, for example (Niem, 1999; Sun *et al.*, 2001):

Maximum point-to-point distance The edge length of the polygons cannot exceed a specified threshold.

Minimum polygon angle The smallest angle within any polygon must be larger than the minimum angle parameter, typically 12° : This prevents nearly co-linear vertices from forming long, narrow triangular facets, which because of their surface normal, may not accurately match the object's local surface curvature.

Polygon assignment Each polygon can only be shared by a maximum of two other adjacent polygon edges.

Applying these rule to every point in the matrix results in a polygonal surface mesh that covers the point cloud. The larger the rule set, the more accurate the surface is represented as more stray or false points are eliminated, however strict rules may also deteriorate the surface mesh.

Merging of 3D Images

Merging of 3D images, more generally known as registration, is normally necessary when we need to align 3D data sets acquired from different view points. This problem arises because reconstruction can only be done on data that appears on the camera image(s), and therefore to generate a more complete scene description multiple cameras or camera views are required. The final image is therefore the result of many 3D images 'stitched' together to form one. This process consists of two parts: first the 3D surfaces are aligned on top of each other and secondly the surface is combined to form one complete surface.

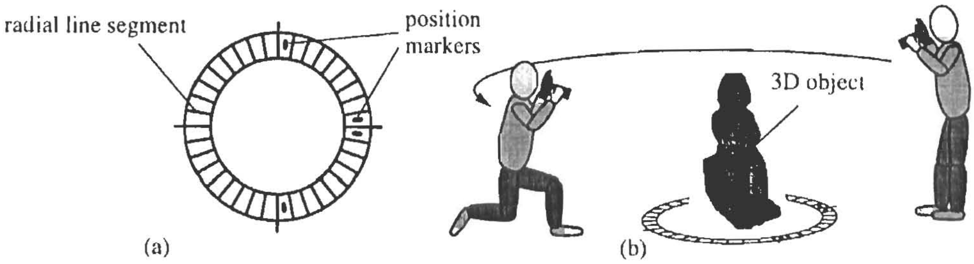


Figure 2.31: Calibration pattern (a) and its usage within the measurement environment (b) (Niem, 1999).

Surface Alignment

Many algorithms for aligning 3D surfaces have been developed in the past (Kneöaurek *et al.*, 2000; Johnson and Hebert, 1997; Chua and Jarvis, 1996; Bergevin *et al.*, 1995) with some of the more common ones being:

Control-point based By defining control points in the environment (*e.g.* markers, landmarks) one can find the transformation that maps the control points in the one image onto the ones in the other image as long as a sufficient amount of control points can be seen in each multiple view. An example is shown in Figure 2.31 where the pattern on the ground serves as the control points.

Edge based This method is usually used to register images where neat contours exist. An example would be to use edge extraction in the orthographic views and attempt a match or otherwise matching of principal curvatures. Principal curvatures can be selected by a three point based representation, where on the one mesh, three points are selected with the requirement that (1) their curvature values can be reliably computed; (2) they are not directly connected; and (3) the points are separated as much as possible. Then, under the curvature, distance, and direction constraints, different sets of three points on the model surface are found to correspond to the three point set on the other mesh (Chua and Jarvis, 1996).

Optimisation The goal here is not the extraction of features but rather transforming one of the images and measuring the similarity between images until the best fit is achieved, *i.e.* iterative closest point algorithm (Besl and McKay, 1992). Given the algorithm's simplicity, it performs quite well but there are two major drawbacks. Firstly, the one surface must be a subset of the other. The presence of points in each set that are not in the other leads to incorrect correspondences. Secondly, since there are six axes of movement, this can involve large processing times. Attempts to solve these two problems have led to several variants of the original algorithm (Eggert *et al.*, 1998).

Surface Stitching

Once the meshes are aligned the two surface meshes are combined to form one file. However, the surface still consists of two parts with areas of overlap. To merge the two surfaces one may use a number of techniques. A common technique discussed in the literature is one developed by Turk and Levoy (1994), called 'zippering', which simultaneously '*eats*' back the boundaries of each mesh that lie directly on top of each other until there is no more overlap. Next, the triangles of one mesh are clipped to the boundary of the other mesh and the vertices on the boundary are shared (the meshes are *zipped* together). The drawback of this approach is where the two surfaces overlap, there may be regions that have been reconstructed in the one mesh and not the other and as explained, redundancy data is lost. However, since 'stitching' methods are quite complex on an implementation level, it might be worth obtaining the combined surface by aligning the surfaces, discarding the mesh structuring (*i.e.* left with a point cloud) and then applying unstructured data meshing techniques.

2.4.2 Unstructured Data

Unstructured data usually occurs when a complex structure is being reconstructed. Typically, we have a cluster of points where points would not be distinguishable to be in front or behind other points when viewed. Although by rotating the object, human beings are usually able to visualise the 3D shape, but it is not a simple task to implement this on a computer. There are however a number of algorithms that can accomplish the meshing of unstructured data. A few of these are mentioned below:

Advancing Front With an advancing front method, the algorithm begins with a set of points connected by lines, defined as the initial front (usually the edges of the point cloud). Starting on two adjacent points on the initial front, create a set of nearby points, prioritise and choose a vertex from this set. Delete the starting face from the front and add any new point. Repeating this process triangles are constructed from the front segments and grow towards the interior of the domain, '*advancing*' the front as it proceeds (Tristano *et al.*, 1998).

Alpha-shapes A common explanation to this algorithm is given by Perzl (1997) and Bajaj *et al.* (1997) who both define α-shapes as:

> Given a finite point set S in real space. Let the real space be filled with styrofoam and the points of S made of a more solid material, such as tiny rocks of granite. Now imagine a spherical eraser with radius α. It is omnipresent in the sense that it carves out styrofoam at all positions where it does not enclose any of the sprinkled rocks, that is, points of S. The resulting object will be called the α-hull.

A variation to this algorithm is the Ball-Pivoting Algorithm (Bernardini *et al.*, 1999). Their method also defines a working sphere with radius α, which cannot pass through the surface without touching sample points. Placing the sphere in contact with three initial points, the sphere is '*pivoted*' around two of the initial points until it touches another. Triplets of points that the sphere contacts form new polygons. The set of polygons formed while the ball '*walks*' on the surface constitutes the interpolating mesh.

Clustering Clustering works by grouping adjacent points that are nearly planar to form sets of tiles. Projecting these point onto a 2D plane, the 2D Delaunay triangulation (faster than the 3D method) is applied to generate the polygon connective data that represents a surface/tile. What remains is to apply polygons to the '*gaps*' between the tiles which results in a complete surface model. This results in a rough structuring of the points and can be later refined to produce a more accurate final reconstruction (Heckel *et al.*, 1999).

3D Delaunay Triangulation The result of this algorithm gives the territorialisation of a set of points P in such a way that each tetrahedra can be built such that the circumscribing sphere of each tetrahedron does not contain any other point of P than its vertices. The result is therefore a set tetrahedra consisting of four polygons connecting points. Applying rules as mentioned in Section 2.4.1 polygons are discarded resulting in a triangular surface mesh covering the point cloud (Bajaj *et al.*, 1997).

Chapter 3

Materials

In the following sections the components of the 3D scanning system will be introduced.

3.1 Hardware

The equipment used in this project consists of three main components, namely:

- Digital Cameras
- LODOX Medical Protocol system
- Personal Computer

3.1.1 Digital Camera

The Sony DFW-VL500 (Digital Fire Wire Video Lens) cameras (Figure 3.1) are capable of capturing full motion picture at a standard VGA resolution (640 X 480) at 30 frames per second (fps). The cameras are fully digital and use the Sony Wfine CCD^{TM} image sensor which provides high colour accuracy and high-resolution images. With the ability to directly digitise images to a 24-bit bitmap image and a cap-

Figure 3.1: Sony DFW-VL500 digital camera

ture rate of 30 fps, a very fast data transport system is required. It therefore incorporates one of the most versatile interconnect technologies available today – IEEE1394 (FireWire) – a high-speed serial bus capable of 400Mbps. An added feature of the camera is that it does not need external power since power is provided through the FireWire cable. A PC interface card capable of supporting the transfer speed over inexpensive twisted-pair wire is also required. The cost of the cameras amounted to R14 000 each, making them the most expensive components of the system.

3.1.2 LODOX Limited Angle Computer Tomography (LACT)

The LODOX machine, shown in Figure 3.2, is a linear digital X-ray imaging system primarily designed for use in the trauma environment as a quick diagnostic tool. It can be used in this environment because it uses low dose X-rays to acquire X-ray images. The scanner acquires a radiograph by means of the horizontal movement of the C-arm. This is necessary because the system capture the X-rays with a linear camera. A linear camera captures single lines of data and stacks the data together to form a two-dimensional image. Additionally, the system uses a laser to indicate the collimator width. As the laser is used in a way that reconstruction methods use them, it was possible to utilise the same laser to do reconstructions without any modifications.

The speed at which the C-arm moves is an important factor for the project as it will affect the resolution of the 3D image in the scanning direction. The C-arm currently can move at three speeds, see Table 3.1. Although these speeds are set for X-

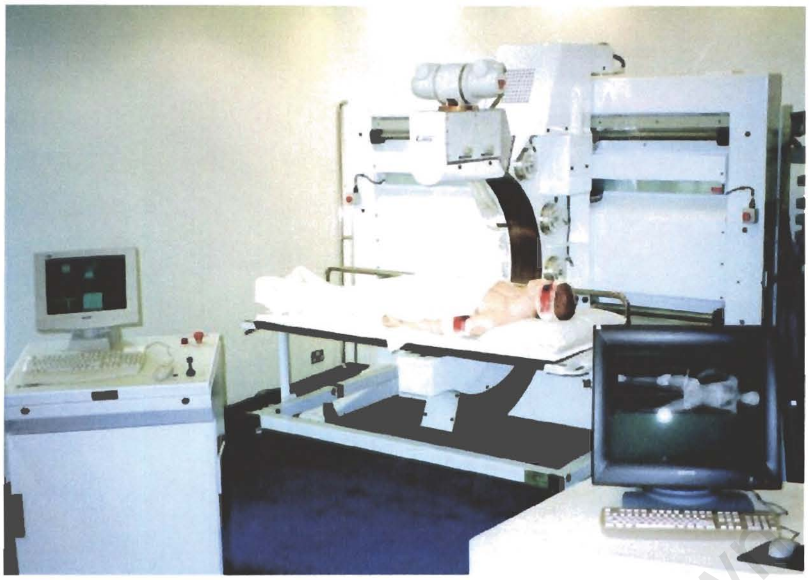


Figure 3.2: LODOX LACT X-ray machine

ray scanning the actual speed can be varied to any speed at the Programmable Logic Controller (PLC) level. However, we will only consider these three speeds since the goal is to simultaneously capture the 3D surface together with the X-ray image. The implications involved with the limited scanning speeds will be discussed later in Section 6.2.

Table 3.1: Speed classification of C-arm movement in scanning direction

Speed Class	Speed (mm/s)	Change (mm)/frame at 30fps
Normal	140	4.7
Half	70	2.3
Quarter	35	1.2

Currently, the LODOX system stores its X-ray images in two parts: the image data in a Khoros visualisation image file (.viff) and a separate file for patient and configuration data. Although this is not a recognised medical image format the LODOX system has the ability to export the file in the DICOM standard.

3.1.3 Personal Computer

There are a number of requirements when choosing the computer to do the processing for this project. Typically a fast computer would be preferable to be able to do the computations faster so that time needed for displaying the 3D image is kept to a minimum. The two most important components will be the IEEE 1394 interface, as mentioned in Section 3.1.1, and the hard drive. Due to the current frame capture algorithm the images from the cameras are stored directly onto the hard drive, and therefore to maintain the frame rate of the cameras, a hard drive with a high data writing capability will be required. An evaluation of the hard drives tested for this project is given in Table 3.2.

Table 3.2: Comparison of hard drive capture rates.

Resolution	**Hard Drives**	
	IBM	**Fujitsu**
640x480 YUV 4:2:2	8fps	16fps
640x480 YUV 4:1:1	16fps	18fps
320x240 YUV 4:2:2	30fps	30fps
160x120 YUV 4:4:4	30fps	30fps

The computer that captured the images from the camera consisted of a 700 MHz AMD Duron processor with 256MB of RAM, a Fujitsu hard drive and a 400Mbits FireBoard^{TM} IEEE1394 interface card.

3.1.4 Calibration Tools

The tools needed for calibrating the cameras consist of a simple planar checkerboard pattern and an object with a calibrated vertical edge. The checkerboard is used for calibrating the internal parameters of the camera while the vertical edge is used to extract the position of the light source. The checkerboard consisted of 2.765cm squares while a edge height of 19.1cm was used.

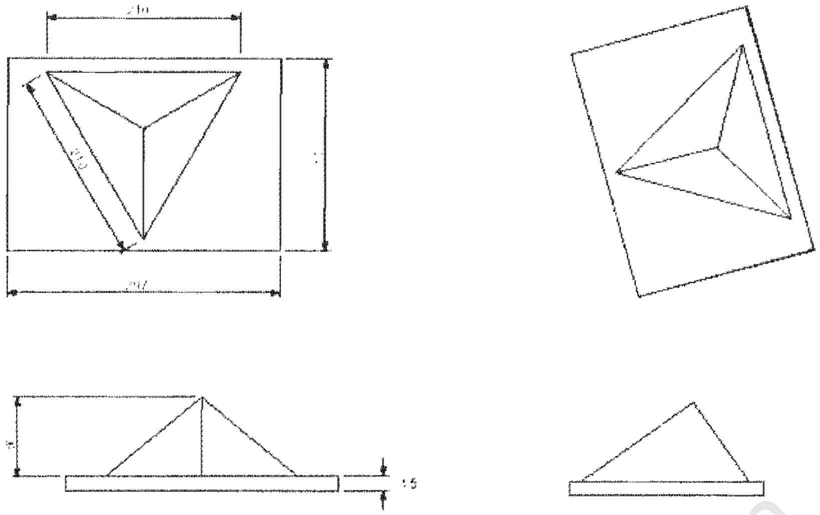


Figure 3.3: Schematic of calibration object used for merging to 3D images.

When using the weak structured lighting scanning method, a bright light source is required to cast a distinct shadow. In the experiments conducted, an overhead projector was used as the light source because it provided good lighting and was able to approximate a good point light source (due to the lens). The better the approximation to a point light source the less shadow outlines are seen and a more distinct edge is obtainable.

Another calibration tool needed was an object that had distinct control points so to obtain the rigid transformation that will align the two camera views. This transformation would then be used to merge two 3D images to obtain a more complete reconstruction. Although the camera's world coordinate systems is calibrated using the same checkerboard pattern (*i.e.* the board is placed in both camera's field of view). The extraction of the origin in the world coordinate system may not lie on the same corner position of the pattern (drawback of the chosen calibration program). Therefore a simple calibration rig was constructed out of cardboard to check and if need be, be used to correct for the error. A pyramid was chosen to provide a pivot point (*i.e.* highest peak) for rotation in the $X - Y$ plane while the elevated base was chosen to give a number of control points that would be used for orientation. The shape and dimensions of the calibration rig is shown in Figure 3.3.

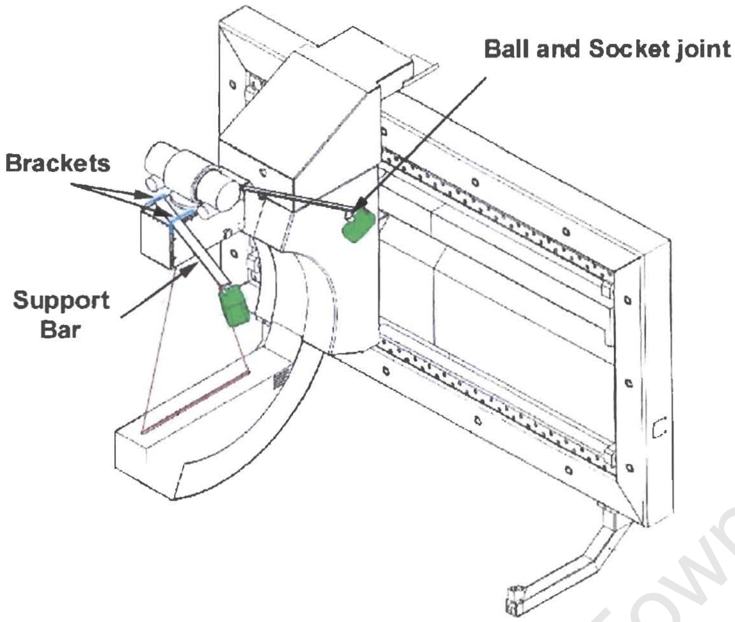


Figure 3.4: Camera mounting on LODOX.

3.1.5 Camera Mounting

The positioning of the camera on the LODOX C-arm was achieved by connecting a camera to one end of a support bar, by a standard ball and socket joint, and clamping the other end of the bar on the top of the C-arm via the brackets (see Figure 3.4). The brackets were designed to allow for some degree of flexibility with regard to the camera position so different positions could be evaluated to see how positioning would affect the 3D reconstruction. There were two main configurations used, one as seen in Figure 3.4 where the cameras are approximately 60° apart and one where they are 180° apart (*i.e.* mirror the one camera position so that its on the other side of the C-arm). Although the positions of the camera(s) are dependent on the type of scanning method chosen, the reason for choosing this configuration will be discussed later in Section 6.1.

3.1.6 Radiation Protection for Camera

Due to the position of the cameras on the C-arm, the cameras are vulnerable to radiation damage. Although typical CCDs can directly image X-rays, damage does

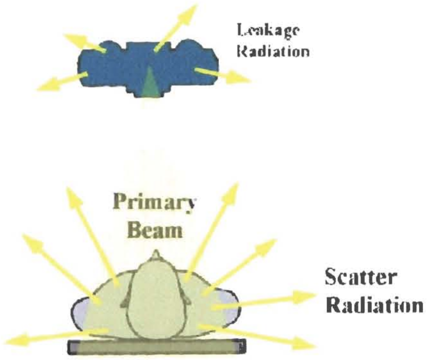


Figure 3.5: Leakage and scattered radiation (Fluoroscopy, 2001).

occur by increasing the dark current (charge at zero intensity) within the exposed pixels (Eikenberry *et al.*, 2001). However, more importantly, care should be taken to shield non-imaging parts of the CCD, such as the output amplifier, since radiation can adversely affect chip performance even at lower radiation doses. It is therefore important to investigate what protective measures need to be taken to ensure proper camera operations. There are two types of radiation that are of concern when mounting the cameras near the tube, namely leakage and scattered radiation (also known as secondary radiation). See Figure 3.5.

Leakage Radiation

Leakage radiation is radiation that escapes from the tube itself and radiates in all directions mainly from the spot where electrons strike the anode (Johns and Cunningham, 1983). It is widely documented that the leakage radiation from every X-ray tube must be contained within a protective housing that reduce the leakage radiation to less than 100mR/hr at a distance of 1m from the housing (Radiation, 2001; Johns and Cunningham, 1983). The current tube specifications for LODOX are 65mR/hr at 150KV, 20mA, at 1m.

Scatter Radiation

The main source of scattered radiation is the patient. When X-rays enter the patient, some are absorbed totally by certain tissues, some are partially absorbed and change

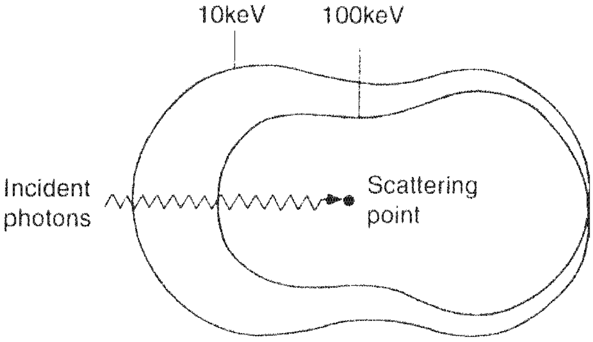


Figure 3.6: Polar diagrams showing the spatial distribution of scattered X-rays around a free electron at two different energies (Dendy and Heaton, 1999).

direction, and others penetrate the patient. The X-rays that change direction and exit all sides of the patient, including back toward the X-ray tube, are scattered X-rays. However, exposure from scatter is highly angular dependent, see Figure 3.6. Thus, the classic inverse square law may not necessarily apply to scatter radiation. Some factors affecting scatter levels are high kV and mA and wide open collimators, as an increase in dose to the patient results in increased scatter levels. Concurrently, wide open collimators will increase exposure levels since the amount of scatter originating from the patient is directly related to the area of the beam (Limacher *et al.*, 1998).

Protective Measures

To determine the protective barriers needed for the camera we make use of formulas from shielding guides for diagnostic X-ray installations. However, due to the manufacturers of the camera not supplying the radiation hardness of the camera, we will compare values obtained from the literature.

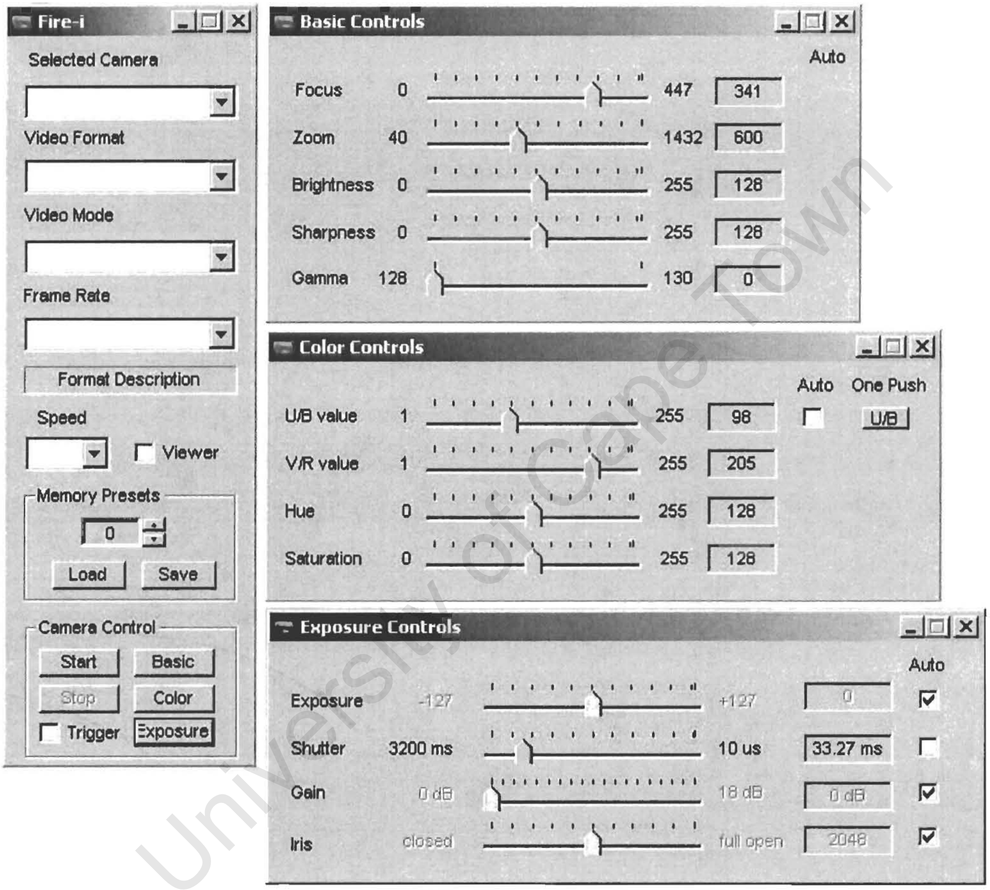
Assuming a whole body scan (13 s) is done every half an hour. The radiation the camera will receive is: using equation B.1, the estimated leakage radiation dose is 0.063 Gy/year and; using equation B.2, the estimated scatter radiation dose is 0.9842 Gy/year (using a conversion factor of 1 R $\approx$ 1 rad = 10 mGy). Therefore the total dose the cameras receive equates to roughly 1.05 Gy/year. Reviewing data from radiation experiments on commercial off-the-shelf electronic equipment. It was found that, CCD chips only start to degrade progressively at a total dose of 25 Gy (number of white spots on image increased continuously)(Rausch and Tavlet, 1999; Adams, 2001). However, since only the CCD is considered and not the individual camera components, we need to estimate what the minimum dose failure threshold might be for the whole camera. As shown in Table B.3 the most vulnerable component is the NMOS device family which experience failures at about 6 Gy. It is therefore clear that the cameras should operate for six years without showing signs of failure. If longer camera expectancy is required, protection from radiation must be considered or maintenance plans implemented to replace the cameras after six years. Protection from leakage radiation can be achieved by wrapping the camera in lead sheeting while some lead lined glass can be used to protect the lens side of the camera from the scattered radiation. Due to the duty cycle of the LODOX machine at the time of project, no radiation protection was needed.

3.2 Software

The algorithms and graphical user interface has been written mostly in Microsoft Visual C++ version 6 with camera calibration and some post 3D-image manipulation done in Matlab version 6.1. The reconstruction and calibration algorithms have been adapted from work done by Bouguet (1999). The final 3D image is written as a VRML (Virtual Reality Modelling Language) file which requires third party 3D viewing software. The only specialised software required was that for the cameras.

3.2.1 Camera Control Software

In order to control and customise the camera to our particular needs, it was necessary to purchase a Software Development Kit (SDK) for the camera. The SDK called Fire-i^{TM} and developed by Unibrain, provided the basic control functions needed to operate the camera. The camera control software has a standard interface (Figure 3.7) which allows the user to select which camera to use, the mode/resolution, frame rate as well as the colour, exposure and zoom/focus. It should be noted that the standard functions available did not allow for continuous recording of images, and therefore a customised function had to be developed. The solution to this was to loop the single image capture function while changing the filename.



☐ Viewer

0

☐ Trigger

Auto

Focus 0 447 341

Zoom 40 1432 600

Brightness 0 255 128

Sharpness 0 255 128

Gamma 128 130 0

Auto One Push

U/B value 1 255 98 ☐

V/R value 1 255 205

Hue 0 255 128

Saturation 0 255 128

Auto

Exposure -127 +127 0 ☑

Shutter 3200 ms 10 us 33.27 ms ☐

Gain 0 dB 18 dB 0 dB ☑

Iris closed full open 2048 ☑

Figure 3.7: Layout of camera control interface

Chapter 4

Methods

The methods described in this chapter are centred around the laser scanning technique as this is the technique chosen for this project. These methods can be divided into three main sections:

1. Pre-Scan
2. Scan
3. Image Processing

An overview of the steps involved from the scan to the final 3D image is represented as a flow chart in Figure 4.1.

4.1 Pre-Scanning Procedure

The pre-scan procedure is only required if the system is being adjusted or installed for the first time since it involves calibrating the cameras in their mounted positions as well as extracting all the relevant configuration information needed for a 3D reconstruction. The cameras are mounted as previously discussed in Section 3.1.5.

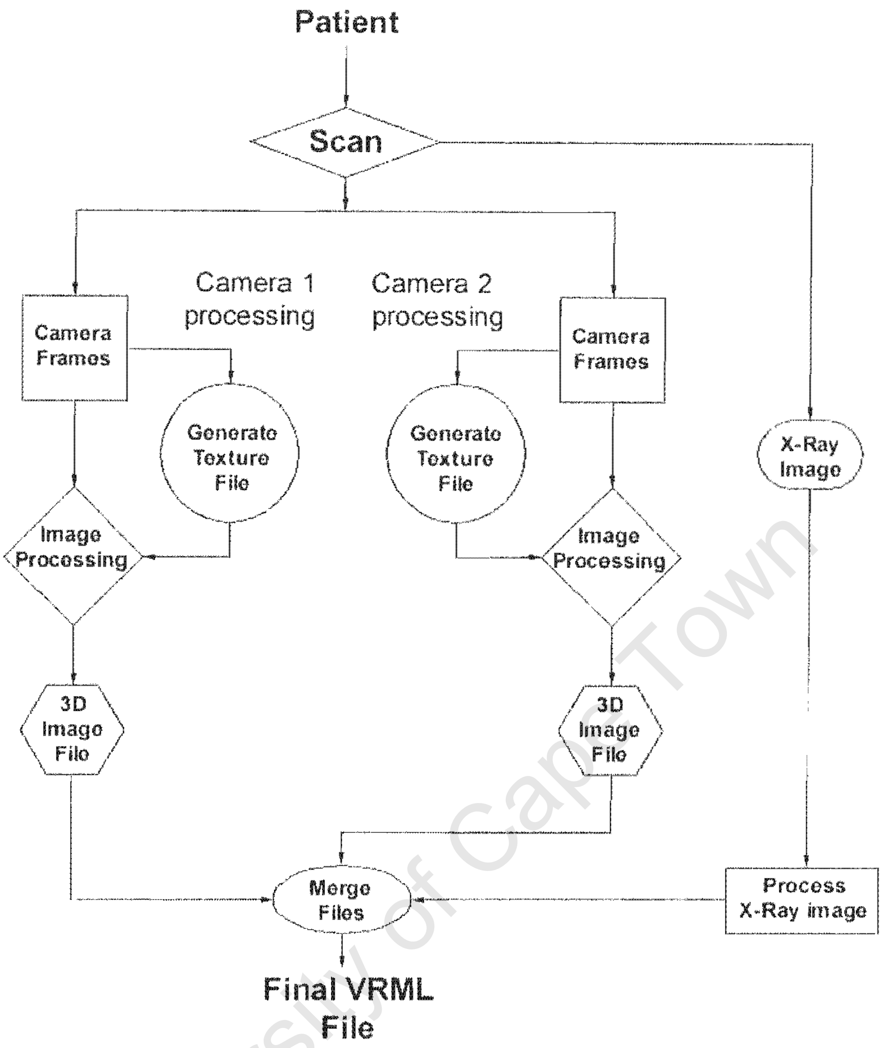


Figure 4.1: Flow chart showing the steps needed to achieve the final 3D file.

4.1.1 Calibration

There are three calibration steps that need to be completed. They provide parameters that are used in the reconstruction steps and are described below:

Camera Parameters: Determines the internal and external parameters for the camera. This needs to be done for both cameras.

Light Source: Obtains the position of the light source in each camera's coordinate system.

Camera Position: Determines the position of the cameras in a common coordinate reference system so the 3D images can be merged.

The camera parameters are obtained using a calibration method previously mentioned in Section 2.1.1. It involves capturing about three to five images of the checkerboard in various orientations and processing them using the Tsai camera calibration software to obtain the camera's parameters. With the camera correctly calibrated the position, 3D data can now be computed. To estimate the light source's 3D coordinate extrapolation methods are used, described by the following steps:

1. Image an object with a known vertical edge on the ground plane so that a shadow is seen.
2. Calculate the 3D coordinate at the tip of the shadow and the base of the edge.
3. Add the edge height to the base coordinate to get the position at the top of the edge.
4. Fit a line through the tip of the shadow and the tip of the edge.
5. Repeat until there is an acceptable set of lines, all passing through the light source position.

By doing this the 3D coordinate of the light source is estimated. As for camera position calibration between the two cameras, this will be discussed in Section 4.3.5 where we explain the merging of two 3D images.

4.2 Scan

In the scan phase the 3D scanning system should capture a sequence of frames from each camera while the LODOX system captures the radiograph. However, due to the complexity of integrating the capturing of images from both cameras simultaneously, it was decided to rather capture data from one camera at a time, therefore requiring the need to do the scan twice. As for the generation of a sequence of frames, the camera SDK was used to modify the camera software to capture a set of indexed frames and store them on the hard drive. The frames, 24-bit colour images, are stored with a filename of '*frame_XXX.bmp*' where *XXX* refers to a three digit number. With the scan complete the images are processed by the C++ program and Matlab functions to obtain data relevant to the reconstruction (*i.e.* 3D coordinates and texture file). Furthermore, these frames need to be adjusted to conform to the correct orientation and format before being processed by the C++ program.

4.2.1 Converting Frames to Greyscale

A Matlab function was written to convert the 24-bit colour frames to greyscale (8-bit) since the scanning algorithm was only written to process 8-bit images. 8-bit images are also used to limit the memory requirements for the processing computer as colour images take up three times the amount of memory. However, instead of just discarding the colour information, we make use of it by extracting only the red band of the colour image so to get a more distinct laser edge or signal-to-noise ratio. Additionally, because the cameras are mounted upside down, we need to rotate all the frames 180° to orientate them correctly. This is required because the test program was already written for right-side-up orientated frames.

4.3 Image Processing

With the setup of the system established and the set of images captured we can now look at the image processing algorithms.

4.3.1 Description of the Scanning Algorithms

From the camera configuration and the moving action of the LODOX machine, it can be seen that the the system closely resembles one developed by Andrew and Aldrich (1989). However, instead of one camera we use two and the camera and light beam unit move over the object instead of the object moving through the beam. Using triangulation principles between rays and planes (Section 2.1.3), only one camera needs to be considered as each camera is just a sub-system of the whole system.

To implement the photogrammetric principle, mathematical algorithms were adapted from Bouguet's shadow/weak structured lighting technique. Bouguet used B-Dual-Space geometry to represent planes as points in space, which make the calculations easier (see Section A.3). A brief explanation of the image processing algorithm is however given in Euclidean space.

Considering the processing , we have a laser edge that is deforming with time, or alternatively per frame or distance. Since we know the direction and speed the C-arm (Table 3.1) and the camera capture rate (Table 3.2), the distance between frames can be calculated. Therefore by extracting the laser edge from every image, in the sequence, and adding an off-set in the scanning direction a point cloud can be generated that represents the surface. Applying a mesh to the structured coordinates hence builds a surface that can be visualised. The following section will now describe the edge localisation and coordinate calculation for each frame (see Figure 4.2).

1. The light plane parameters $\prod$ is calculated by extracting the line produced when two planes intersect (in this case the light and ground plane - λ_h on image plane) together with the light source position S. The line is defined by tracking the edges at the upper and lower limits of the image, which, together with S provides three points needed for plane calculation. This step is only done for the first frame of the sequence.
2. Points along the laser edge in the frame are detected using edge detection techniques, in this case detecting a particular gradient across five points above a particular threshold value. Scanning for the edge is done from left to right and top to bottom. The result is the pixel position p for these points
3. A ray $\bar{x}_c$ is then calculated using points p and the camera centre O_c. Triangulating the intersection of this ray with the light plane $\prod$, the Z_c coordinate is of point P is calculated using Equation 4.1. The X_c and Y_c coordinate is obtained, by substituting Z_c into the ray $\bar{x}_c$.

$$Z_c = \frac{1}{\bar{w}_c.\bar{x}_c} \tag{4.1}$$

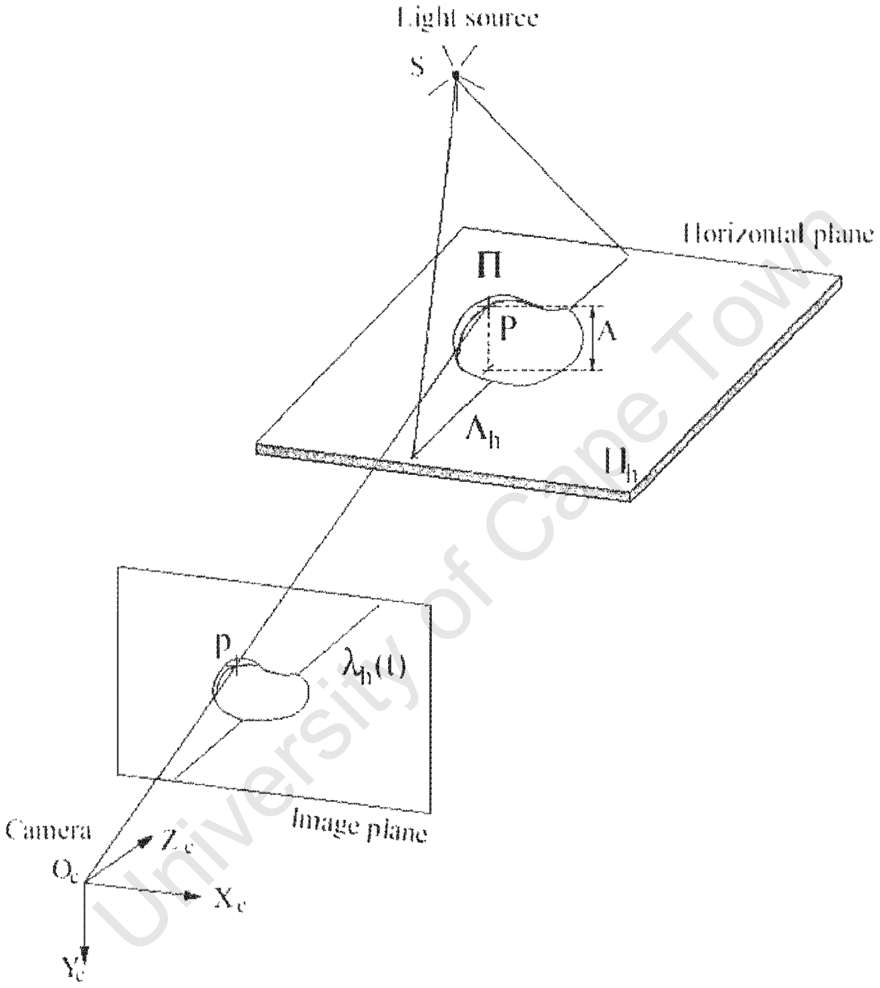


Figure 4.2: Schematic showing scanning technique (adaptation from Bouguet (1999)).

where: $\bar{w}_c$ = Homogeneous coordinate vector of shadow/laser plane.
$\bar{x}_c$ = Coordinate of image pixel.

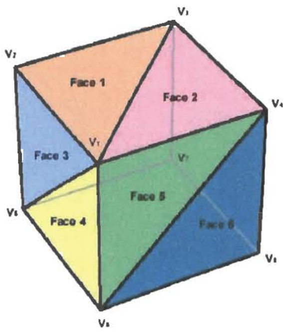
4.3.2 Mesh Generation

From the type of scanning method the reconstruction data generated is in a structured form, therefore allowing us to employ a simple meshing algorithm. The structuring of the point cloud is in the form of a matrix where each column represents a slice of the object obtained from a frame, while the row represents the vertical resolution of the image. (Example: capturing N frames at 320×240 makes a matrix size of $240 \times N$). Each entry therefore represents a 3D coordinate of a slice. The meshing is done by comparing adjacent entries on the matrix and if the XYZ coordinates are within a certain range, the points are connected to form one side of a polygon, the minimum polygon angle was also applied to prevent meshing anomalies. However, because triangles can only span one entry length, wherever a 3D coordinate was omitted a hole is formed. We attempt to patch small holes by interpolating between points and adding points where some have been left out, however this is only done where single points are missing so that structure representation is not compromised (*i.e.* deliberate holes are not filled).

4.3.3 VRML Format

The VRML format is a standard file that is written with two or three parts or shapes. Two of the shapes represent the two 3D surfaces while the other represents the X-ray image. This allows us to customise the properties of the shapes. For example, making the shapes transparent so the underlying structure/surface can be seen or making one shape a different colour so as to distinguish between the two surfaces.

As mentioned in Section A.1.1, 3D images can be written in various ways. In this project the 3D reconstruction is written as a number of Indexed Face Sets. This is a set of vertices together with a list of three indices for each of these vertices, representing a planar triangle as seen in Figure 4.3. By storing the data in this format considerable space is saved.



Vertices	X	Y	Z
V_1	0	0	1
V_2	0	1	1
V_3	1	1	1
V_4	1	1	1
V_5	0	0	0
V_6	0	1	0
V_7	1	1	0
V_8	1	0	0

Face Set			
$Face_1$	V_1	V_2	V_3
$Face_2$	V_1	V_3	V_4
$Face_3$	V_1	V_6	V_2
$Face_4$	V_1	V_5	V_6
$Face_5$	V_1	V_4	V_5
$Face_6$	V_5	V_4	V_8
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Figure 4.3: Index face set data structure

Texturing of VRML Files

Texturing is a technique used to give a triangulated surface mesh a realistic look by applying the colour of the surface at the time of scan onto the triangulated surface mesh. This is done be mapping parts of a JPEG colour image file onto the planar triangles of the surface. For a detailed explanation on this procedure see the VRML Standards (VRML-1, 1997).

The texture file for the generated surface is extracted from the colour frames of the scan. The method of texturing is accomplished by a Matlab program that finds the edge of the laser and captures the pixel intensities at that edge. However this is technically impossible because the laser edge obscures the true colour of the surface. This can be overcome by extracting the pixel intensity a few pixels before or after the edge. This is possible due to the linear scanning motion of the C-arm since the pixel at time t is equal to the pixel at time $t \pm d$, where d represents a time delay. To correct for the introduction of this delay, the final image is shifted by either adding or subtracting a pixel step equivalent to time d to the correct side of the image, which therefore aligns the texture image with the surface.

The texturing feature also allows us to display the X-ray image together with the surface structure. When this is done the surface is made semi-transparent to allow one to see the X-ray image through the surface. However, the X-ray image first needs to be converted to an image format recognised by the VRML code. The VRML code can only recognise images with 256 colours/grey levels, and therefore the X-ray image was converted to the correct format using the appropriate Window Width and Window Centre as mentioned in Section A.2. Additionally, the X-ray image needs to be cropped before it can be textured because the image width is not stored according to the X-rays detected but according to which cameras have detected X-rays. Therefore X-ray image widths from the LODOX machine are in multiples of *(camera pixel width)/(binning)* where binning is a term used to represent grouping of surrounding pixels (*e.g.* 3×3 binning equates to 9 pixels grouped to form 1). The cropping is necessary for alignment purposes as will be explained in the next section.

4.3.4 Alignment of 3D Image with X-ray Image

When a scan is done, the operator controls the width of the scan by means of adjusting the collimator, which in turn affects the laser width. Since reconstruction is only done where an edge can be detected, the laser can also be used to align the corresponding X-ray image, as the path of the laser dictates the path of the X-ray beam. Although the widths of the two images (See Figure 4.4) vary according to the distance between the ground plane (table top) and the detectors. This can be corrected for if the X-ray plane is defined at the four corners of the reconstructed ground plane. Therefore when the X-ray image is textured onto this plane we have effectively aligned the surface image with the X-ray image and by doing this we have also corrected for any distortions that have been produced by the fan-beam properties of the X-ray beam.

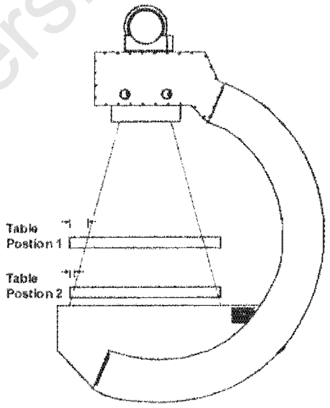


Figure 4.4: Diagram showing difference between widths of X-ray image and 3D image.

4.3.5 Merging of Two 3D Data Sets

By using two cameras to capture 3D data a method to merge the two 3D images is necessary. The merge consists of mainly aligning the two point clouds so that the points represent the surface of object being reconstructed. No surface stitching is required due to the ability of modern desktop computers to handle and process large amounts of data. This also has the advantage of redundancy and each surface property can be controlled individually (*e.g.* each surface and be represented in its own colour). Imaging the calibration object in Figure 3.3, two point clouds are gathered from the cameras.

A crude algorithm was written in Matlab that showed the one 3D point cloud from one camera superimposed over the point cloud of the other camera. The user is then prompted to input certain parameters to move and rotate the one image so that a good match is achieved. Once the user is satisfied that there is good correlation between the two surfaces, the translation and rotations are then stored as calibration data for use in all subsequent scans. The calibration object is shown merged in Figure 5.4 in the next chapter.

The merged images are then created by reading in the two 3D image files, transforming the coordinates in one of the images with the calibration data and then adding the new coordinates, as a new shape (see Appendix A.1.1), to the first image file. Unfortunately this results in a very large 3D image file because we are storing redundant data. As mentioned earlier, modern computers should be able to process these large files, but if storage space or older computers are an issue, a simplification step can be applied using the '*reducepatch*' function in Matlab. This function reduces the number of faces in the structure while trying to preserve the overall shape of the structure.

The algorithm is designed to be easily portable to a completely integrated program so merging of the images can be preformed automatically, but this can only be achieved when the cameras are synchronised to capture together. The experiments that were conducted all required some manual processing because the two image sequences were not synchronised. Image synchronisation was achieved by deleting all stationary images in the two sequences as the scanning procedure on the LODOX machine was repeatable.

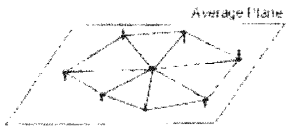


Figure 4.5: Point averaged to itself and to its surrounding points.

4.3.6 Smoothing of 3D Images

The problem with noise in an image may introduce small errors on the surface, where the surface may seem rough while in actual fact it is smooth. This is seen in Figure 5.4. Therefore to reduce this noise we apply some smoothing to the surface. Smoothing is accomplished by taking a point and modifying that point to the average of all other points connected to it, see Figure 4.5. This is done by defining a square matrix $C_{N \times N}$, with length and width equal to the number of points representing the surface, N. The length and width both represent the index to the surface point set. Therefore stepping through the connectivity data of the polygons, the matrix $C_{N \times N}$ is built up of 1's and 0's (the properties of *sparse matrices* are used here since $C_{N \times N}$ is large matrix but contains relatively little data) , where a 1 represents a connection between points (*e.g.* if the 5th point connects to the 40th point in the point set, then place a 1 in position $C(5, 40)$ and $C(40, 5)$ in the matrix $C_{N \times N}$). With this matrix, determining the points connected to a single point can easily be done by looking up where 1's occur in a particular row or column, since each row or column represents a single point in the point set.

Chapter 5

Results

This chapter will show some results of the 3D scanning system and gives a brief discussion on some of the features of the result.

5.1 Shadow Reconstruction

The first experiments were done to get the original shadow algorithm working with our hardware. After coding the algorithms the first tests were conducted using a standard desk lamp with a 100 Watt globe. The results where unsatisfactory due to the contrast of the shadow against the scene (Figure 5.1) but it was found that by using a brighter light source (overhead projector) we were able to get a more distinct shadow and hence a better reconstruction (Figures 5.2 and 5.3). In these examples some of the main drawbacks such as '*line of sight*' reconstruction, shadows and very low reflective areas (text on mouse) can be seen. High reflective areas can also introduce errors as can be seen in the box example in Figure 5.3, where a false elevation was reconstructed on the ground plane near the diagonal side of the box. This is due to the calculation of the edge threshold of the algorithm. Because the area is brighter than expected (due to reflections), the algorithm's expected threshold value has been offset and this therefore resulted in the reconstruction of the elevation. The properties of the reconstructions are listed in Table 5.1.

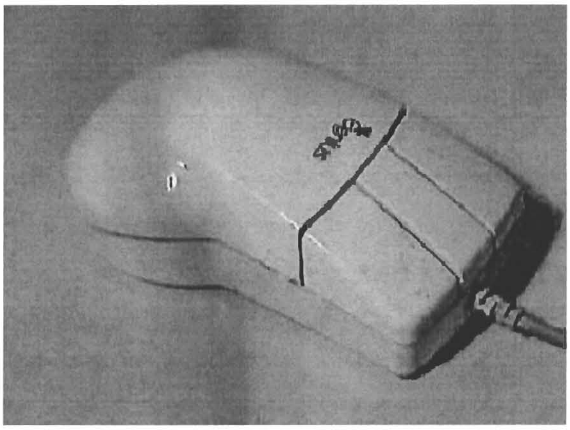
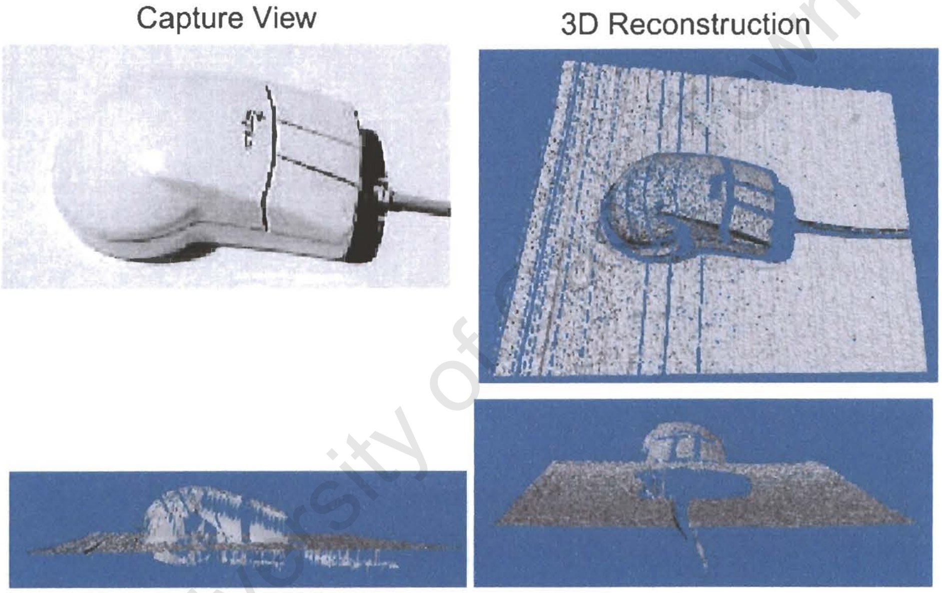


Figure 5.1: Image showing the poor contrast between the shadow and mouse.



Capture View

3D Reconstruction

Figure 5.2: Reconstruction of a computer mouse.

Table 5.1: Statistics on box and mouse reconstructions.

	Mouse	Box
No. Frames	496	340
No. of Vertices	59321	56790
No. of Triangles	110831	110868
Size of File (KB)	4,103	4,027
Capture View Size	320×240	320×240

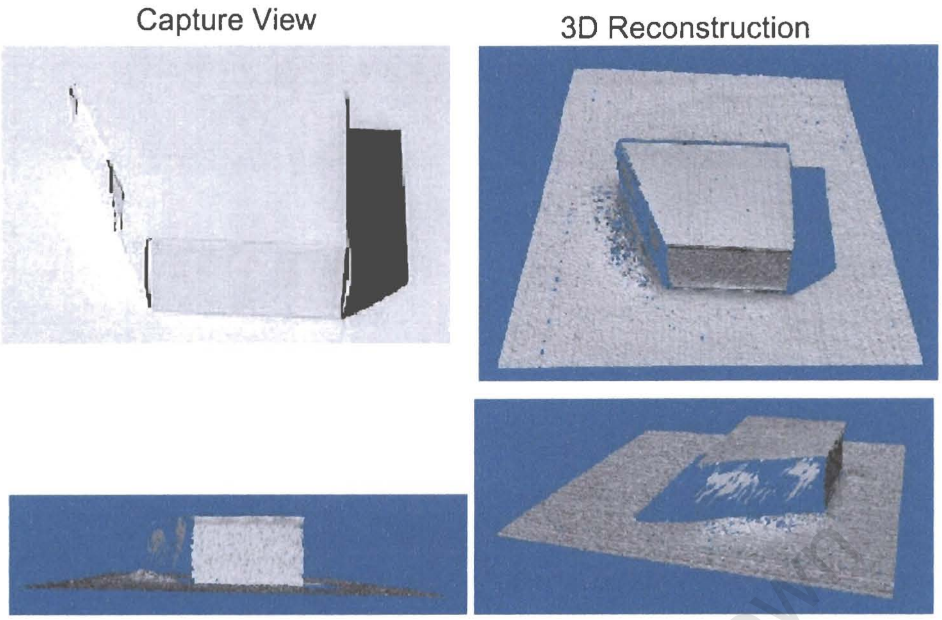


Figure 5.3: Reconstruction of a box.

5.1.1 Shadow Reconstruction Accuracy

To determine the accuracy of the reconstructions a comparison of a number of measurements were made on the original object and compared to the reconstruction result. The object used for the measurements was the box seen in Figure 5.3. There are four parameters that are used to determine the accuracy of the surface, surface noise and errors in the XYZ axes. The results of these measurements can be seen in Table 5.1.1. The surface noise was calculated by taking a neighbourhood of about 100 points on a particular plane and finding the standard deviation. The XYZ axis errors are the differences of the true value and the reconstructed value. To do the surface measurements a 3D image viewer called Actify 3D View was used. This 3D-viewing software has the ability to conduct a number of measurements on the reconstructed data points. The reconstruction therefore gives an error of about 4 mm.

5.2 Laser Reconstruction

To do the laser reconstruction a two camera setup was used, the majority of tests being conducted with the cameras 60° apart as shown in Figure 3.4. All the experi-

Table 5.2: Accuracy results for shadow reconstruction.

Box			
	True Value (mm)	Reconstruction (mm)	Relative Error (%)
Surface Noise (std)		0.011	
X axis	83.9	78.5	6.5
Y axis	73.4	70.6	3.9
Z axis	41.8	44.1	5.5

ments described in the following sections were conducted on the normal-speed setting of 140 mm/s unless otherwise stated and with individual camera characteristics. The camera characteristics are the colour, exposure and zoom/focus as mentioned in Section 3.2.1. Colour and exposure settings were all set to automatic whereas zoom and focus was chosen to give the best reconstruction resolution since the higher the zoom the better the reconstruction. Therefore for each experiment the cameras were set to maximise the size of the object in the field of view of the camera. Additionally, the ground plane needed to be covered with either white paper or a white sheet (full body) to improve the reflectiveness of the laser. Where the sheet was used a perfectly smooth plane could not be achieved due to creases on the sheet. Furthermore, the LODOX system does have the ability to move the C-arm without taking an X-ray so human subjects can be reconstructed without exposing the subject to X-rays.

5.2.1 Experiment 1: Pyramid

In this experiment a reconstruction of a pyramid was done using the half-speed setting of the LODOX machine. The reconstruction and properties of both cameras and the merge can be seen in Figure 5.4 and Table 5.2.1 respectively. The merged result was simplified and smoothed. As can be seen, each camera has reconstructed about 80% of its view. Although this seems to be high it is considered to be quite poor since most of the lost data is actually the ground plane, which should be reconstructed easily. The reason for this is due to the misalignment of the laser plane and a narrow collimator width, which slightly obstructs the laser on the one side. This produces a lighter edge, which therefore falls out of the threshold during edge detection and

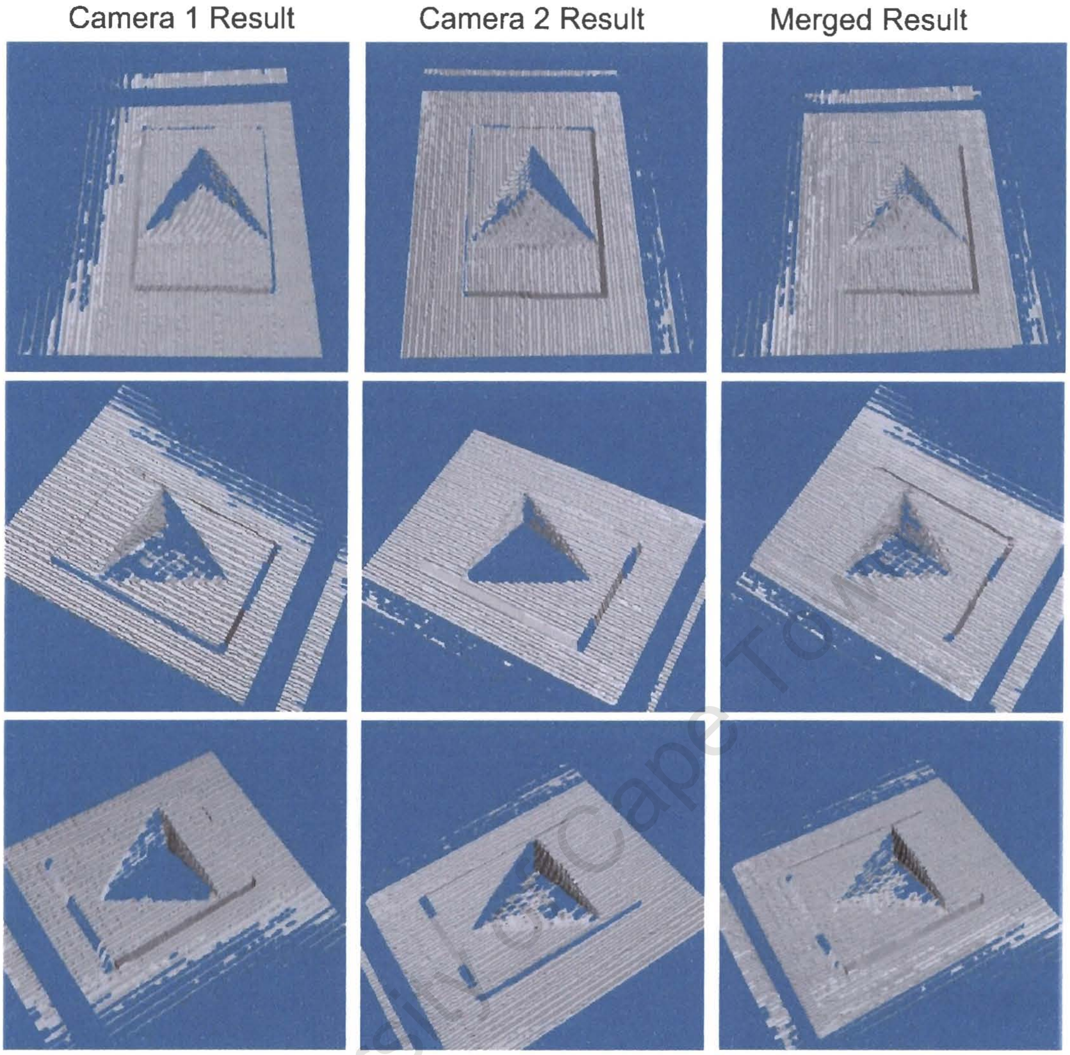


Figure 5.4: Result of the merged calibration object.

hence resulting in loss of data. The loss of reconstruction where the misalignment does not come into play (far faces of the pyramid) can be explained by the fact that the faces/planes are close to parallel to the camera's view. When this occurs the laser edge is not seen so distinctly whereas with planes more perpendicular to the camera view (near face of pyramid) are very distinct. Another factor that contributes to the loss of reconstruction is the meshing algorithm, in that if certain points are not close enough together, a triangle cannot be fitted which produces holes in the surface. The other noticeable feature in the reconstructions is the absence of a band across the 3D image. This is due to a function of the LODOX system in that the laser is actually switching off while travelling in the scanning direction. This function, set to switch the laser on and off when either in the start or end set points, is used as an indicator to show these position for the scan.

Table 5.3: Statistics on pyramid reconstructions.

	Camera 1	Camera 2	Merged
No. Frames	118	116	-
No. of Vertices	19937	20870	21190
No. of Triangles	37141	39639	38389
Size of File (KB)	1,763	1,867	1,687
Capture View Size	320x240	320x240	-

5.2.2 Experiment 2: Human Face

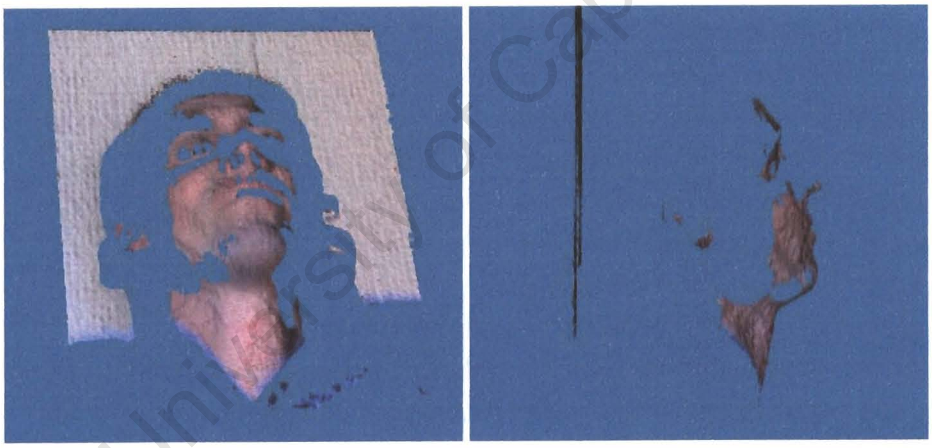


Figure 5.5: Example of reconstruction of face with texturing.

This example demonstrates the texturing of the surface to give it a more realistic look. A human subject was scanned from head to shoulders at the half-speed setting using only one camera. The result, shown in Figure 5.5, was done on the half-speed setting and had been smoothed three times using the smoothing algorithm described in Section 4.3.6. Texturing is an option available after the reconstruction and requires the generation of the corresponding JPEG texture file. This texture file is generated from the colour images recorded and relates to the natural colour of each of the laser lines in the set of images. Texturing a surface will usually cause the quality of the surface to drop because of the introduction of more colour (particularly red) in the images. The increase in red makes the edge detection harder because the edge may become hidden in the image therefore making it less distinct. A particular trade-off therefore needs to be made to retain good surfaces and good texturing. This experiment was done in medium lighting so that a reasonable texture file could be generated. Because the texture file is separate, image enhancement programs can be used to obtain a better texture (*e.g.* increase contrast and brightness). However, if one requires a more realistic look one should consider work done by Debevec *et al.* (2000) who have accounted for the reflectance characteristics of the human face.

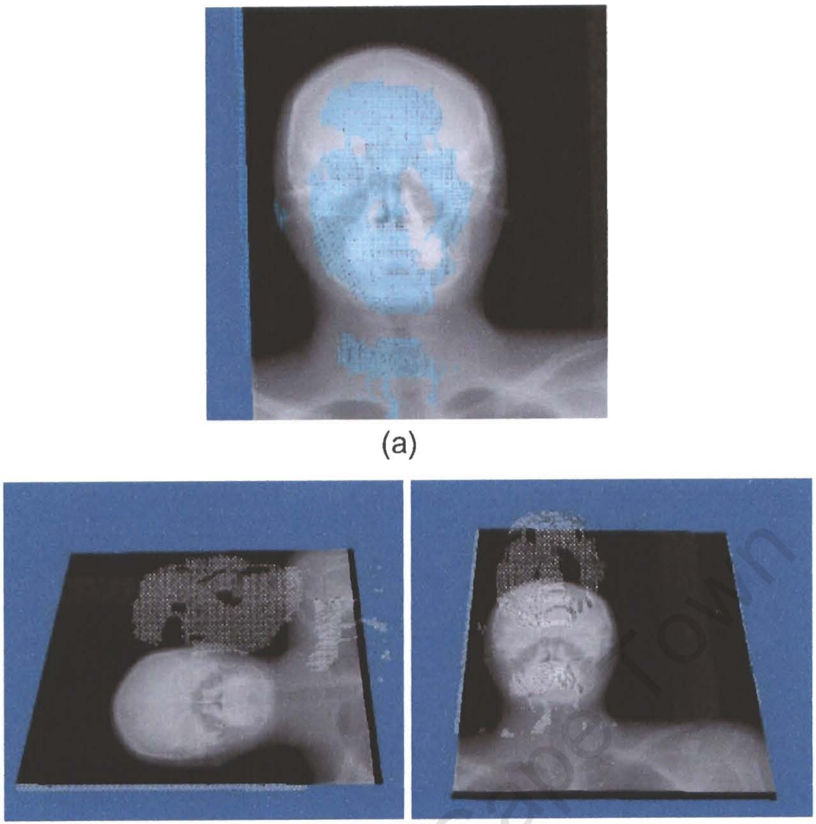
5.2.3 Experiment 3: X-Ray and 3D Surface

This example shows the incorporation of the radiograph in the 3D image. Here we demonstrate how the surface is made transparent to allow one to see through to the X-ray image below. The result shown has been done with only one camera, as seen in Figure 5.6.

The alignment is not exactly correct because the two systems (camera and LODOX) are not synchronised *i.e.* we presently do not know exactly were the X-ray starts and ends. This lack of synchronisation results in a small offset in the scanning direction when aligning the images (Figure 5.6(b)), although the alignment is correct perpendicular to the scanning direction (Figure 5.6(c)). We are still able to get a fairly good alignment even though we do not exactly know where the X-rays start, because the start position is indicated by the flashing of the laser light.

5.2.4 Experiment 4: Full Body

Because the LODOX system is a full body scanner an example of a full body is included here. Although the result shown in Figure 5.7 is not the complete full body this was because the subject was sitting lower on the table than normal and therefore the feet were out of the C-arm's range. This reconstruction is done using the dual camera setup where the grey surface represents one camera and the blue represents the other camera. The blue section does not represent the complete surface obtained



(a)

(b) (c)

Figure 5.6: Example of head reconstruction with X-ray plane shown.

by the camera but only the parts that are not captured from the first one. From the reconstruction it was seen that the added value of the second camera does not increase reconstruction coverage as expected. This is probably the same reason as mentioned before in that due to the laser line falling on oblique surfaces and not being reflected enough, together with a ridge edge detection algorithm, the edge is often not detected during the scan. This led us to experiment with a different setup configuration in the hope of improving either the reflection of the laser edge or achieving a better "line of sight". This will be demonstrated in the next experiment.

5.2.5 Experiment 5: Different Camera Setup

This experiment was performed to see if we could improve the reconstruction by having the cameras separated by approximately 180° instead of the 60°. As the results show this configuration yields better 3D surfaces than the previous setup (Figure 5.8).

Figure 5.7: Example of full body reconstruction.

The increase in quality can probably be explained by the fact that the faces/planes are more perpendicular to the camera's view and therefore reflect more of the laser light. It should be noted however that this example was done with the half-speed setting so the quality of the reconstruction can be misleading if comparing to examples done at normal speed.

5.2.6 Experiment 6: 90° Plane Reconstruction with Mirror

One of the major drawbacks of the current scanning system is the use of a fan-beam laser plane, since it limits the reconstruction of vertical and near vertical surfaces. It was therefore decided to conduct an experiment to investigate how the algorithm

Figure 5.8: Reconstruction using a 180° camera separation.

would perform without this limitation. This was accomplished by placing a mirror as shown in Figure 5.9 to reflect a portion of the laser light to illuminate vertical edges. This configuration was set up to simulate a laser illuminating from the side.

As Figure 5.10 shows, the results are encouraging since in the reconstruction, one can almost make out the word "Rooibos" on the reconstruction where the non-reflective regions (text) did not yield many points. However, by using this method we can also see that this will only improve reconstruction on vertical planes parallel to the scanning direction.

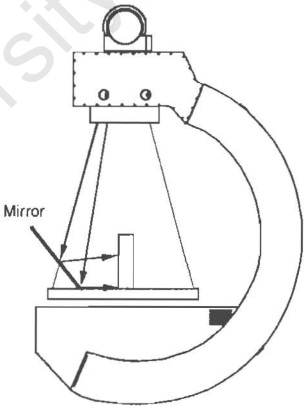


Figure 5.9: Using a mirror to illuminate vertical edges.

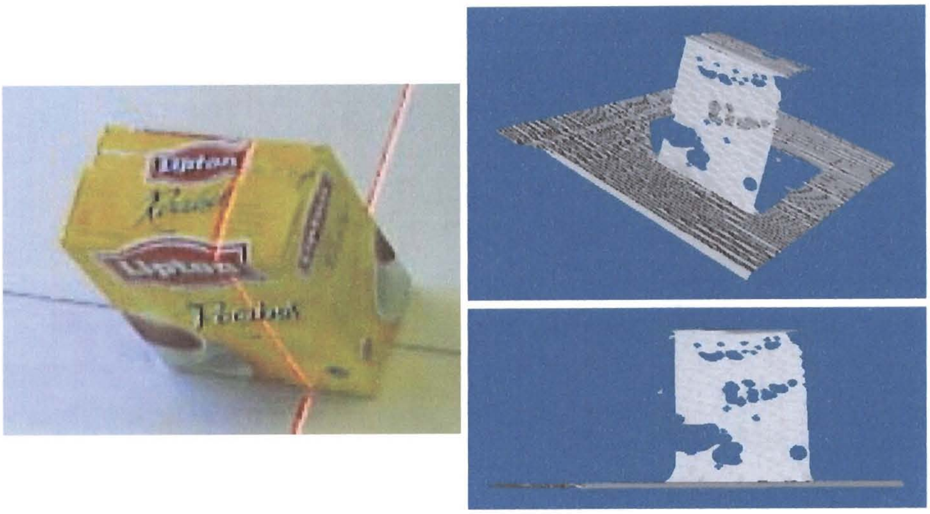


Figure 5.10: Rooibos tea box reconstruction using a mirror.

5.2.7 Laser Reconstruction Accuracy

To determine the accuracy of the reconstructions we used the same criteria as that for the shadow reconstruction accuracies. However, since we expect to use two cameras for reconstruction, the accuracy for both is shown in Table 5.2.7. The accuracy of the laser reconstruction system is therefore within about 5 mm.

5.3 Processing Time

Processing time is important if the system is to be commercialised because one has a finite time to keep the operator interested. If one has to wait half an hour for an image to be generated or displayed users will tend to lose interest in the actual result. An analysis of the expected processing times has therefore been done to see where the bottlenecks are and where methods should be optimised. The times for each step are broken down into section as previously shown in the flow chart in Figure 4.1. It should be noted that processing times are highly dependent on the speed of the processing computer. The processing times mentioned in the following sections were achieved using a 700 MHz AMD Duron processor with 256 MB of RAM.

Table 5.4: Accuracy of laser reconstruction.

Pyramid Camera 1			
	True Value (mm)	Reconstruction (mm)	Relative Error (%)
Surface Noise (std)		0.1085	
X axis	210.0	206.9	1.5
Y axis	297.0	294.0	1.0
Z axis	86.0	81.0	5.8

Pyramid Camera 2			
	True Value (mm)	Reconstruction (mm)	Relative Error (%)
Surface Noise (std)		0.0953	
X axis	210.0	206.4	1.7
Y axis	297.0	294.0	1.0
Z axis	86.0	80.9	5.9

5.3.1 Scan

The scanning of a patient can vary depending on what is being imaged. It can range from a small image like the head or hand right through to a full body image (largest). Due to the limitations of the high voltage generator for the X-ray tube, the maximum allowed time for any scan is 13 seconds. Therefore the expected frames generated from one camera is 390 frames and hence 780 frames when using two. Since these frames are 320×240 bitmap images, we need to rotate and convert these images as mentioned in Section 4.2.1. This is done via Matlab and takes roughly 0.5 seconds per frame.

5.3.2 Image Processing

Here the images for each camera are processed separately to get the 3D coordinates. This program initially requires the camera and laser light calibration parameters. These parameters are obtained from the pre-scan procedure and therefore do not affect the outputting of the 3D file (typical calibration times are around 5min per camera). Once the parameters are read in, the processing of each image can begin. To process one image takes about 0.084 seconds for a fairly good reconstruction. This value is dependent on how well the edges are detected since less edges mean less processing.

There are three versions of 3D image files that can be generated. They are:

Plain 3D surface This depends on how many points are detected and how well the mesh is constructed. This is estimated to take about 0.02 seconds per frame.

Surface with Texture Due to the structure of the VRML file adding a texture increases the amount of code 1.5 times therefore the time per frame would instead would take 0.03 seconds.

Surface with Texture and X-ray image Adding the X-ray image to the file adds very little code to the VRML file and can be considered to be negligible.

This step only generates the actual coordinates and triangles for the VRML file and not the actual colour for the triangles. This step needs to be done separately and is discussed in the next section.

5.3.3 Texture and X-ray File

The texture file is generated in Matlab where the colour images need to be processed to extract the corresponding colour. This method has already been explained previously in Section 4.3.3. This process takes about 0.1 seconds per frame. As for generating the correct X-ray image this can only be done once the X-ray image is available on the LODOX system which is typically about 20 seconds after the end of the scan. The X-ray image from the LODOX system then needs to be converted to the DICOM format and cropped to only X-ray data so that alignment with the 3D image will be correct. This is highly dependent on the size of the scan since large images take time to be read into Matlab. A 865$\times$698 image will be cropped in $\approx$ 5 seconds.

5.3.4 Merging two 3D Files

As mentioned in a previous section this step will become automated once we integrate capturing of frames simultaneously with both cameras. If done this step should take about 10secs once the rotation and translation of the two cameras are know. However, since this is not the case we need to manually merge the two images together as is done when finding the rotation and translation of the two cameras in the calibration process. This usually takes about five minutes to complete.

With the estimated processing times defined we can estimate what the time till viewing will be as seen in Table 5.3.4.

Table 5.5: Table of showing time till output.

	s/frame	**Set time (s)**	**Full Body**	**Head**
Number of Cameras			2	2
Number of Frames			780	200
Conversion and Rotation	0.5		390	100
Plain Surface with X-ray	0.02		15.6	4
X-ray File generation		5	5	5
			405.6	104
Texturing	0.01		7.8	2
Texture File generation	0.1		78	20
			491.4	126
Automatic Merging		10		
Manual Merging		300	300	300
Time till viewing (s)			791.4	426
Time till viewing (min)			13.19	7.1

It can therefore be seen that, to generate 3D image with X-ray and texturing, a full body scan will take approximately 792 seconds or 13.2 minutes while a head scan will take 426 seconds or 7.1 minutes. The processing time of the X-ray is not included in the table because it is a parallel process and will only become a factor if the reconstruction processing times are less than 30 seconds.

Chapter 6

Discussion

A discussion of the scanning system, its methods, results and limitations will follow in the following sections together with a number of recommendations to improve the system.

6.1 3D Scanning Setup

The decision on which was the most appropriate 3D scanning setup to use was based on the following two questions:

1. What is the best suited scanning methodology?
2. Knowing the scanning methodology what are the best camera positions?

6.1.1 Chosen Scanning Method

To choose the scanning method we need to apply some constraints. There are three constraints that were judged to be important, and they were:

- Minimise the cost of the scanning system.
- Maximise 3D image reconstruction (coverage) and quality (accuracy).

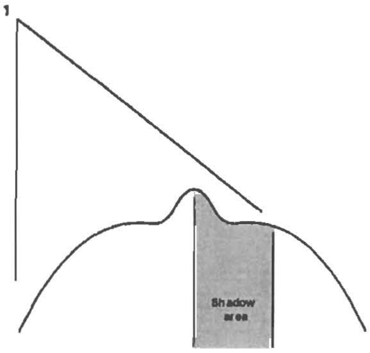


Figure 6.1: Line of sight limitation (Bibb *et al.*, 2000).

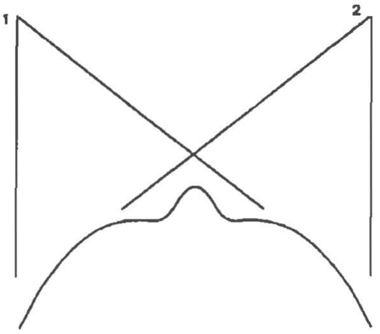


Figure 6.2: Solution to line of sight problem.

- Do not interfere with the primary operation of the machine.

Considering the first constraint, we note the cameras are the most expensive component of the system and that by using passive techniques or stereometry, a minimum of two cameras are required to do a reconstruction whereas the laser-camera needs only one (see Section 2.1). Additionally, a laser already exists on the LODOX system which can be used to do the reconstruction and therefore does not add to the cost of the system. However, using one camera we are limited to a small reconstruction coverage and shadows may result in a poor reconstruction (Figure 6.1). A two camera system would therefore prove to be a better choice. Since two cameras are being used let us reconsider the use of stereometry. This technique, from the literature review, should give a good quality 3D-image but will limit reconstruction coverage since stereo-photogrammetry techniques can only reconstruct pixels seen in both stereo image pairs. Additionally, stereo techniques require the cameras to remain stationery when using laser scanning (Section 2.1.3 and therefore the ability to obtain large surfaces with a high resolution is limited. Opting for a higher level of reconstruction coverage than accuracy, we therefore chose to use a two camera laser system because we maximise coverage of the object (Figure 6.2). Using the laser method also has the added advantage of being able to scan a larger area.

Therefore the appropriate solution was to use a two camera, laser-camera reconstruction system.

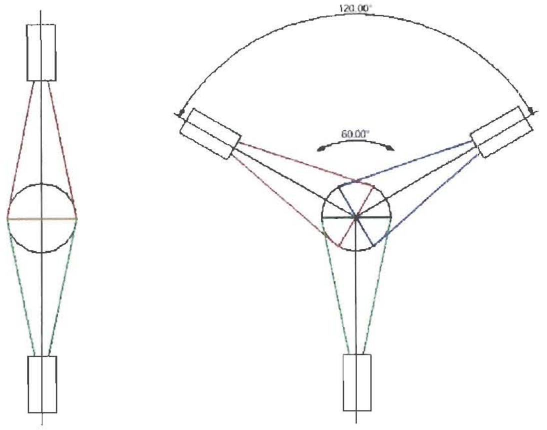


Figure 6.3: Two and three camera setup providing maximum coverage.

6.1.2 Camera Positioning

Firstly, the cameras will be limited to be within the plane of the top of the C-arm so as not to interfere with the X-ray scanning operations of the LODOX machine. Secondly, the cameras need to be positioned such that good depth perception can be achieved. Therefore the position of the cameras needs to be fairly far from the laser light source. A good angle to view the light plane from would be around 45° from the vertical. However, this would require long support beams for the camera which, depending on the stiffness of the beam, would make the cameras susceptible to bouncing. Therefore to prevent large vibrations a 30° angle was used instead.

The angle between the two cameras is also important since for a good reconstruction we need to maximise object coverage. Consider the 2D problem: the absolute minimum number of cameras needed is two (cameras 180° apart), although, depending on the shape of the object, there might be parts of the object that are inadequately imaged. A better solution would be a three camera setup (cameras 120° apart) as it will greatly increase the overall reconstruction since the whole structure is imaged with some redundancy. The three camera setup also improves the merging of the 3D images since there is a 60° overlap. The two camera setup won't be as easy (see Figure 6.3). However, as discussed above, we want to minimise cost so an extra camera would not be feasible reasonable reconstructions can be done with two.

If we now consider the 3D problem we first note that the bottom face is impossible to image (due to the table) and therefore only the top face needs to be incorporated.

This has already been solved because as mentioned earlier the cameras are limiting to the plane of the top of the C-arm. Additionally, the problem of merging is solved because now both images have a common surface to which merging can be achieved. Two setups were tested, one with the camera separation to be 60° and the other 180° and from the results, it can be seen that the 180° camera separation gives a better reconstruction due to more of the surface being imaged by the two cameras.

6.2 Accuracy

Before we discuss the accuracy of the system let us first give a brief introduction to the terminology used. We will designate the coordinate system as follows: the trolley top will be the $X - Y$ plane with Y being in the direction of travel of the C-arm and Z being perpendicular to the trolley top. The accuracy of the shadow system is presented briefly to compare how well the algorithms were implemented but the main discussion will be focused on our system.

6.2.1 Shadow Scanning

The experimental accuracy of the shadow scanning algorithm is given by Bouguet (1999) as follows: surface noise of approximately 0.7mm in standard deviation over a 9x9 pixel neighbourhood (81 pixels) and relative errors in the X and Z axis of 1.3% and 3% respectively. Although these values differ quite a lot from the ones obtained for our experiment (surface noise 0.011 mm, X-axis 6.5% and Z-axis 5.5%), it should be noted that the configuration setup was very different. Firstly, Bouguet's accuracy results were done on a large scene whereas our experiment used a smaller scene, although this should have actually made our system more accurate since the resolution is higher. An explanation might be that the difference in hardware may have caused the discrepancy. Additionally, Bouguet may have used extra optimisation techniques not used in our code. However, the results do show some correlation in that the errors in the Z direction are much higher than those in the X direction. What is significantly different is the surface noise.

Our calculations for surface noise was much less than what Bouguet recorded, in orders of magnitude. This might be also due to the differences in the configuration and hardware of the two systems, but could also be in the conduction of the experiments. A factor that definitely affects the reconstruction quality is the speed at which the shadow is moved over the scene. If this is not similar in the experiments, comparisons may be invalid. Another issue might be from the calibration of the cameras, as the accuracy can vary depending on the number of calibration frames and the orientation of the checkerboard in these frames.

6.2.2 Laser Scanning

One of the differences between the shadow and laser scanning accuracies are in resolution of the surface. This is because the algorithm takes into account the movement of the camera. The result is that the resolution is constant (depends on image resolution) in the X-axis and Z-axis while being variable in the Y-axis. This is due to the various speed settings of the C-arm in the scanning direction as mentioned in Section 3.1.2. Theoretically the resolution in the Y-axis can be calculated using Equation 6.1 and Table 3.1.

$$Resolution = \frac{Speed\ of\ Travel\ [\ mm.s^{-1}\ ]}{Frame\ Rate\ [\ s^{-1}\ ]} \tag{6.1}$$

Some possible resolutions are represented in Table 6.1

Table 6.1: Resolution of 3D image in the Y-axis using a Fujitsu hard drive.

LODOX speeds [mm.s^{-1}]	140	70	35	140	70	35
Frame Rate [s^{-1}]	30	30	30	16	16	16
Resolution [mm]	4.7	2.4	1.2	8.8	4.4	2.3

Although the accuracies measurements were conducted separately, and on different objects, they are compared to see if any correlation can be found. Comparing the shadow and laser system accuracies, it is seen that the laser system improved in the X-axis (error 5.4 to 3.6 mm), was consistent in the Y-axis (error 2.8 to 3.0 mm) while deteriorating in the Z-axis (error 2.3 to 5.1 mm). The reason for the improved accuracy in the X direction is probably due to the change in edge detection although the edge detection criteria is much more simpler than the one used in the shadow algorithm. The consistent accuracy in the Y-axis indicates that the Y step modification is valid and working correctly. The deterioration in the Z-axis may have been due to the camera having different depth perceptions (camera configuration). This is probably the highest contributing factor to the differences encountered, since the laser reconstruction was done on a larger scene than the shadow experiment.

The greatest difference from the shadow results and the laser results was the surface noise. The laser system had a surface noise 10 times that of the shadow system. This was in fact not noise but an artifact of some sort which produced saw-tooth surfaces in Y-axis. This was found to be a resolution problem, since the edge was not sufficiently represented. This will be discussed in more detail in Section 6.2.4.

One other feature that was noted was that the accuracy of the reconstruction was better for objects closer to the camera. This was found to be the case in both the shadow and laser reconstruction methods and is probably due to the camera lens, in that objects further away have a lower resolution due to the divergent properties of the optical rays.

If we compare the accuracy of the current system we see that the accuracy is not nearly as accurate as systems that are commercially available. These systems, mentioned in Section 2.2, have accuracies of about 1mm (Nebel *et al.*, 2001).

From the above mentioned accuracies it is possible that this may not be the correct scanning technique to use. Though a number of modifications will be needed to be implemented to improve the accuracy of the system, one may want to consider using stereometry techniques. Stereometry techniques will result in a similar reconstruction coverage but with higher accuracies and will not require any modifications to the LODOX configuration. However, this system may still suffer from the same resolution problems encountered with the laser system (discussed later in Section 6.4.2).

6.2.3 Error Analysis

An analysis of where errors could have been introduced and what can be done to minimise these errors was conducted. As with most systems, calibration is the most important factor when trying to improve reconstruction quality. Therefore to be able to achieve good accuracy we need to have good calibration. It would therefore be beneficial to redo the calibration or use a different calibration method to obtain better results. One of the most important calibrations that needs to be accurate is the light source position since the definition of the laser plane is calculated from this. If this is inaccurate all the data is affected.

Alternatively the light source calibration can be done by:

1. Taking one image of the laser line on the ground plane.
2. Finding the centre of this line, point c (assuming the laser plane has been collimated symmetrically).
3. Projecting a line from point c perpendicular to the ground plane with length equal to the distance between the ground plane and the focal spot on the generator.
4. Finding the coordinate at the end of this perpendicular line.

The coordinate will therefore be an estimate of the laser light source.

Other sources of errors may have arisen from the actual validation method of the system, as when measuring the surface points/edges, the exact point for measuring may not have been reconstructed and therefore had to be estimated. An object with greater surface definition might improve the measurement and hence the accuracy.

External sources of error may have been introduced from camera vibration during the scan. Though this errors should be small because of the stiffness of the supporting bar. These errors however are usually at the start of the scan and usually stabilises quickly and should not effect the rest of the reconstruction. A more important factor that might affect the accuracy of the reconstruction is patient movement, particularly since a 'double' scan needs to be done. However this may only affect the merging accuracy. This won't be a problem once the final product is developed since the cameras will be synchronised and therefore only patient movement during one scan

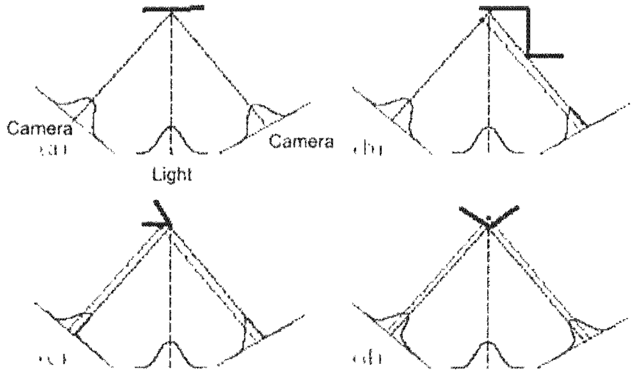


Figure 6.4: Error sources when tracking the centre of a light strip from two camera positions (Abi-Rached *et al.*, 1998).

will pose a problem. Scanning times will however not be more than 14 seconds which for patients is achievable. Another system which required the patient to remain still (Debevec *et al.*, 2000), reported that capture times of about one minute are also achievable.

6.2.4 Edge Detection

The use of a simple edge detection algorithm may have been the cause of some of the errors in reconstructing. The rigidity of the algorithm (gradient method with thresholding) prevents edges that are unclear to be detected. For example, a distinct peak (representing the edge) may appear on the row but because of a high threshold the edge is not detected. A more flexible approach would be to have a floating threshold in that the algorithm checks through a range of thresholds instead of just one. This way a fuller reconstruction can be achieved. Alternatively, finding a particular characteristic on the row might also be worth considering (*e.g.* spike), however this may require much more processing.

Sources of error may also occur as a result of light falling on the object at a grazing angle. When the projected light falls on a portion of the object that is nearly parallel to the light's path, the sensor sees a dim and stretched-out version of the pattern. This adds uncertainty to the position of the range points as the gradient change is less steep. Tracking the edge by a gradient method may prove inefficient if objects have many surfaces parallel to the light's path. An alternative might be to try tracking

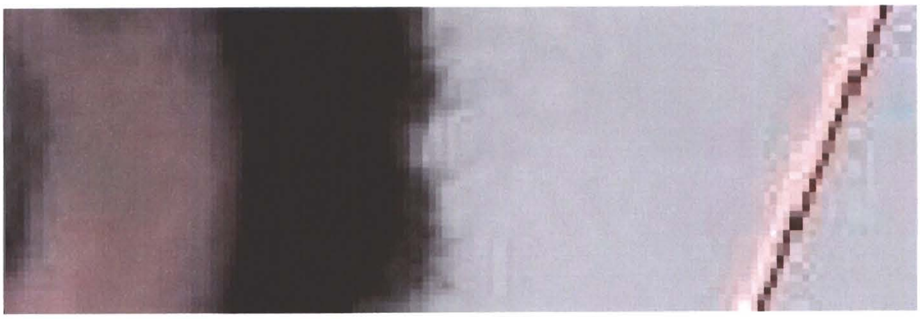


Figure 6.5: Zooming in view of the laser line in a frame.

the centre of light beam as shown in Figure 6.4(a). However, this method also suffers from uncertainties. Some of these are self occlusion (Figure 6.4(b)), sharp edges (Figure 6.4(c)) and edge creases (Figure 6.4(d)). Although these errors are more common for wide light stripes where the intensity distribution is spread over many pixels, thin stripes will not produce large errors but the drawback is that they have lower intensities thereby having a lower signal-to-noise ratio.

What was also discovered was that because of the finite resolution of the CCD, a slightly diagonal line is represented as a number of shorter vertical lines offset by a pixel (depending on the angle), see Figure 6.5. The effect of this digitisation is shown in Figure 6.6, where a step pattern artefact is generated on the surface due to the edges being detected in the same column of the image. Therefore to get rid of this artefact, the use of sub-pixels needs to be considered. Here one would use interpolation techniques to determine the edge position to a sub-pixel accuracy, before the triangulation of the world point is done. A few attempts were made to implement the code but due to time constraints, the code was eventually commented out. Additionally, the noise or artefact could be reduced by using a higher resolution setting for the cameras (*i.e.* 640×480), but this would only be feasible when the capture rate can be kept high, either by a newer capture program or faster hard drive. If using the higher resolution, 3D surface resolution in the scanning direction will diminish to about 10 mm (fps drop from 30 to 16). This can be minimised if a slower scanning speed is selected, but then radiation dose increases. Another trade-off in increasing the resolution is also that processing times become longer since four times the amount of data needs to be processed in each frame.

Another source of error is when parts of the surface have a non-reflective property as in the case of text or hair on the surface. Since the laser light does not reflect well

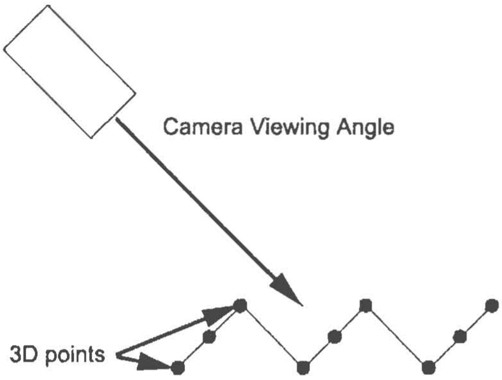


Figure 6.6: Schematic showing the reason for texture on 3D surface.

of these areas, the localisation (if any) of the true point may fail. An example of how text may cause errors is shown in Figure 6.7 where text caused protrusions in the surface where there shouldn't be. Some solutions to this problem have be developed like space-time analysis (Curless and Levoy, 1995) and line shift processing (Gühring *et al.*, 2000). To improve areas where hair is a problem, a simple technique of dusting the area with a white power may be used to improve edge detection (Bibb *et al.*, 2000).

6.2.5 Meshing

The simple meshing algorithm also has a negative effect on the reconstruction coverage, since it cannot reconstruct large surface gradients. This was seen in examples where the slope of the surface was very steep. Although the edge was detected the

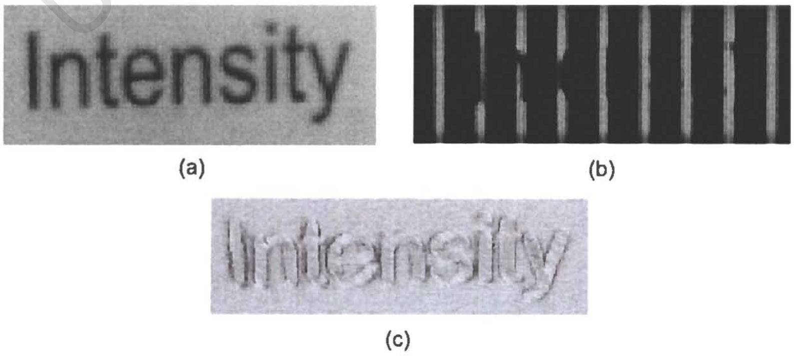


Figure 6.7: (a) Fully illuminated image. (b) Reflectance changes due to text. (c) Rendered view of the 3D surface.

distance between the points exceeds the mesh threshold and therefore the reconstruction is lost in this area. To prevent this, it would be better to use one of the unstructured meshing algorithms mentioned in Section 2.4.2 where a surface would be fitted to the complete set of points.

6.2.6 Merging

The merging of the two 3D images has not been optimised since the result contains a lot of redundant information. This results in large files making storage, transmission, computation, and visualisation more difficult, but with the speed and processing power of modern computers these problems are slowly becoming negligible. One might consider a cropping method, where a common reference plane be defined in both 3D images, the points to the left of this plane be cropped in the one image while, points on the right be cropped in the other. By joining these cropped surfaces along this plane the data structure of the reconstruction will be much less. However by using this method the redundant property of the current method is lost and holes may appear on the surface. This is a similar technique proposed by Turk and Levoy (1994) that was called 'zippering'. One might even consider using unstructured meshing techniques to do the meshing and see if better results can be achieved.

6.3 Processing Time

The processing time at present is quite lengthy and will need to be reduced if it is to be used directly after a scan. One of the drawbacks of the current method is the need to jump from one program to the other. A considerable amount of time is lost by doing this since when transferring data from program to program, data must be outputted so that the receiving program can read in the relevant data. From the breakdown of the processing time line, we can see that most of the time is spent on image conversion and merging.

We can do away with the image conversion step by getting the image processing program to do the conversion itself. Additionally, time can be saved by storing frames in memory instead of on a hard drive, though this may require large memory capacity we can instead extract all relevant information from a frame and then discard it. Therefore by implementing these steps we would have to speed up the processing time considerably since we now only need to load a frame once. An added feature is that we can start the image processing as soon as a frame is available therefore eliminating the delay in waiting for the scan to finish *i.e.* processing can run parallel with the scanning.

As for merging, considering that the current files are manually merged, which takes about 5 minutes ($\approx$ 83% of the time for reconstructing a head), the synchronisation of the cameras will play a major role in improving processing times. Additionally, the use of dual processing may improve on turn around time.

6.4 Limitations

There are a number of limitations found with the laser system being implemented on the LODOX system. Some are from typical scanning conditions while others are from the LODOX machine itself. These will be discussed in the next sections.

6.4.1 Scanning Conditions

The problem with low reflective surfaces is one of the major factors when doing the reconstruction since if one cannot detect an edge, a coordinate cannot be assigned to that pixel/point. This is seen in most of the reconstructions conducted, since dark hair and text on the objects were not reconstructed. The edge is also dependant on the environment lighting, since if the lighting in the room is bright then the edge can become hidden in the colour of the image, which will reduce the ability of the system to reconstruct surfaces. Also to be able to texture the surface the environment lighting needs to be bright enough for the cameras to detect colour. A method of solving this dilemma would be to do the scan in a dark environment but have another light plane (white light) offset from the laser light and then calibrate the texturing algorithm to extract the colour in this plane.

Camera position becomes an issue here too since when the C-arm rotates to the lateral position, only one camera can be used to reconstruct surfaces because the other camera is occluded by the table. A reconstruction is still does not guarantee because when the object is scanned in the anterior-posterior view the cameras are facing not on the centre of the C-arm but off-centre. Therefore when the C-arm is in the lateral position the cameras may not have the object in its field of view. This limits the number of angular positions where reconstruction are possible. Additionally, it should be noted that a complete 3D reconstruction can be achieve since surface against the tabletop can ever be reconstructed (*e.g.* if reconstructing a cube, only 5 sides can be reconstructed during one scan).

6.4.2 LODOX Machine

A drawback with the current configuration of the LODOX machine is that the light source (laser) produces a fan-beam so, depending on the shape being scanned, not all points in the light plane will be illuminated (*i.e.* a shadow effect). This is most noticeable when an object, directly underneath the light source, has a plane perpendicular to the ground plane as Figure 6.8 illustrates. However the reconstruction algorithm is not limited by this property since a simple test was performed to show that if these planes were illuminated the reconstruction could be realised. Another issue regarding the laser was that the light plane's intensity falls off when using thin collimator widths thereby lowering the signal-to-noise ratio and hence degrading the reconstruction. Collimator misalignment issues were also experienced with the machine where the light plane was not uniform (*i.e.* width of the plane was wider on the one side). However this problem can easily be fixed manually.

To correct for this disadvantage it is suggested that a different laser system be used during scanning. Similar to how camera positions are chosen to obtain maximum coverage and eliminate shadows, the same technique can be used but with 2 lasers instead of cameras, see Figure 6.2. This will lead to better illumination of all sides of the object and hence a better reconstruction. Although this improves the reconstruction, this will be limited to only vertical planes parallel to the scanning direction. To be able to implement the system to reconstruct $X - Z$ planes seems to be a complex task and will be left for future work. However, a way to overcome this without any

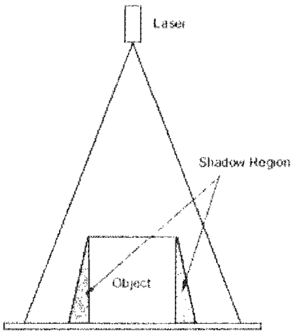


Figure 6.8: Shadow region created from using a fan-beam approach.

modifications would be to orientate the object so that planes/sufaces that are perpendicular to the scanning direction are kept to a minimal therefore allowing the laser light to illuminate most of the surfaces.

The other limitation is the fixed speeds of the C-arm. The fixed speed limits the resolution of the 3D images. Two options are available here, first the resolution can be improved by using cameras with faster capture rates or secondly, by modifying the LODOX system to travel at a slower speed. Both these changes have its advantages and disadvantages in that increasing the frame rate will require more expensive cameras whereas slowing down the C-arm speed will increase radiation dose to the patient if an X-ray is required with the 3D image. The system will therefore need to be tailored to the particular needs of the user.

6.5 Final Comments

Since radiologists can mentally assimilate two-dimensional data and interpret it in a three-dimensional manner, the application of viewing a 3D surface with a radiograph may not prove to be beneficial at all. However, this might not be the case for patients or students as they may not have the experience, so 3D surfaces may be beneficial as a teaching tool. On the other hand, medical physicists may however find it useful in having 3D surface data for automatic manufacturing of compensation filters by computerised milling machines. Additionally, where the 3D data might also become useful is in adding surface data to CT images once this capability is available on the LODOX machine.

Chapter 7

Conclusions and Future Work

7.1 Conclusions and Recommendations

We have presented a simple, low cost 3D reconstruction system that has been integrated onto the LODOX system. The system requires very little equipment and can be easily operated. A two-camera laser system was chosen and by using dual-space geometry mathematical formula to do the triangulation, a digital map of an object's topology was extracted, the result being a VRML file with a textured surface geometry with its corresponding radiograph.

The reconstruction coverage of the scanning system is not great because of the limitations encountered with the laser fan-beam while the accuracy was not very good either ($\approx$ 5 mm). The accuracy which was expected to be similar to the original reconstruction algorithm may have been degraded by the modifications made to the algorithm. Typical reconstruction times are about 6 minutes for small objects while around 13 minutes for a full body. By putting more work into better calibration techniques, slight modifications to the LODOX laser system and a new edge detection algorithm, an improvement to the scanning system should definitely be seen.

Some of the more important recommendations will now be mentioned. Currently the 3D scanning system operates independently from the LODOX system *i.e.* no direct communication is possible between the two systems and requires the usage of various software programs to complete the reconstruction. Therefore, work should be done to form a more integrated system where all the software be written in one package and the scanning and LODOX system be synchronised to perform faster reconstructions. Additionally, implementation of parallel processing, to reduce time to visualisation should be considered as processing times are currently lengthy.

However, if one prefers to leave the LODOX configured as is, one should maybe consider using stereometry techniques, as a similar reconstruction coverage with higher accuracies might be achieved.

In conclusion, the implementation of the scanning system determined that 3D reconstruction is possible using the current LODOX system configuration but is limited by properties of the system, particularly the fan-beam laser plane, which reduces the overall reconstruction coverage. However, a thorough analysis of where possible errors may have occurred was done and a number of recommendations have been made to help improve the reconstruction ability.

7.2 Prospects for Future Developments

With the invention of new scanning methodologies, future work may involve the replacement of the current scanning algorithm with a more accurate one. Additionally, real time processing may allow for adjustment of LODOX scanning parameters *e.g.* changing the width of the collimator during a scan.

Furthermore, to increase the functionality of the scanning system, work should done on methods to improve resolution of the system. This can be done by using the zoom capability of the cameras to filling the field of view of the camera. For this to be feasible, a database of calibration parameters will need to be established so that the 3D reconstruction system will automatically select the correct parameters needed for the algorithm depending on the zoom. By incorporating this into the system the resolution of the reconstructions can be improved.

One limitation that was ignored in this project was the reconstruction of vertical planes perpendicular to the scanning direction. A method of reconstructing these planes may be in using another laser to produce a light plane. Viewing the LODOX system in the $Y-Z$ plane we have the original laser plane at 90° to the tabletop. If we use another laser plane at 45° together with the 90° plane we could reconstruct vertical planes. However, two of these would be needed on both sides of the camera if one wanted to reconstruct vertical planes on both sides of the object.

Appendix A

Image Formats

A.1 3D Image Formats and Display Tools

Typically the most common field for the use of 3D images lies in computer graphics in the form of animation and Virtual Reality (VR) environments. This is because being able to display 3D objects on a computer becomes a very useful tool in allowing for modelling, simulating and visualising complex objects (Kalawsky, 1996). For the purpose of this project we generate our own virtual environment consisting simply of the scanned object and its corresponding X-ray. However, before the 3D image files can be generated, an investigation into the different types and structures of these files needs to be done. A brief description to a number of 3D image files and their viewers/browsers will be discussed in the following paragraphs.

The structure of 3D image files consists mainly of a tree type structure, where there is an environment that represents the 'trunk' and a number of 'branches' connected to the truck called objects. Each object can consist of either a specific shape with each shape having particular properties. Properties include the type of shape (sphere, box, polygons, etc.) and appearance (transparency, colour, etc.). The final 3D image can therefore consist of a number of different objects (branches) in a single environment (trunk). Although this structure is typical for most 3D image files, each vendor prefers to store the data in various formats. This results in a large number of file types being encountered today, for example: Apple 3D Metafile (.3dmf, .3dm), Open Inventor (.iv), VRML (.wrl), 3D Studio (.3ds), Wavefront (.obj) and AutoCAD (.dxt) (3D Format, 2002; Costello and Bee, 1997). Because we have chosen to use the VRML format, a more detailed explanation of its structure will be given.

A.1.1 VRML 1.0 and 2.0/95

The Virtual Reality Modelling Language (VRML) enables world builders to create and explore 3D worlds on the World Wide Web. VRML is a platform-independent language for virtual reality scene design based on the Open Inventor format. Currently, the latest version of VRML (VRML 97) and has been approved by the International Standards Organisation (ISO) and the International Electrotechnical Commission (IEC) as an international standard. VRML files consist of 4 main parts:

Header Identifies the file as a VRML file and specifies the encoding type as well as additional semantic information.

Scene graph Contains nodes which describe the objects and their properties. Nodes are abstractions of various real-world objects and concepts. Examples include spheres, polygons, lights, and material descriptions like texturing and appearance.

Prototype (Optional) Allows the user to extend node types by creating new nodes from already defined (built-in) node types. Once defined, prototyped node types may be instantiated in the scene graph exactly like the built-in node types.

Event routing (Optional) Allows the scene to respond to environmental changes or user interactions. This process can change the state of the node, generate additional events, or change the structure of the scene graph. An example would be to trigger change in movement from collisions (VRML-1, 1997).

For more information on VRML visit the VRML Repository at *www.vrml.org*.

For this project we chose to use the VRML format because of its wide spread use and the fact that it does not required specialised software. A VRML "plug-in" is all that is require for an internet browser to be able to explore a virtual environment (*e.g.* CosmoPlayer, Liquid Reality, Community Place browsers). The choice of browser is highly dependent on the user's preference for visualisation and navigation. The preferred view for this project is called 3Space Assistant from TGS Inc. (2002), a vendor for interactive graphics software.

A.2 DICOM 3

The DICOM standard was developed to aid the distribution and viewing of medical images. The keys to the success of DICOM are the use of standard network facilities for interconnection (Transmission Control Protocol/Internet Protocol (TCP/IP) and ISO-Open Systems Interconnection (OSI) reference model) and the use of the concept of Information Object Definition (IOD). An example of an IOD is shown in Figure A.1. As can be seen, the IODs can be broken down into information entities, modules and attributes, which represent all information related to the image. DICOM image files are stored with a ".dcm" extension.

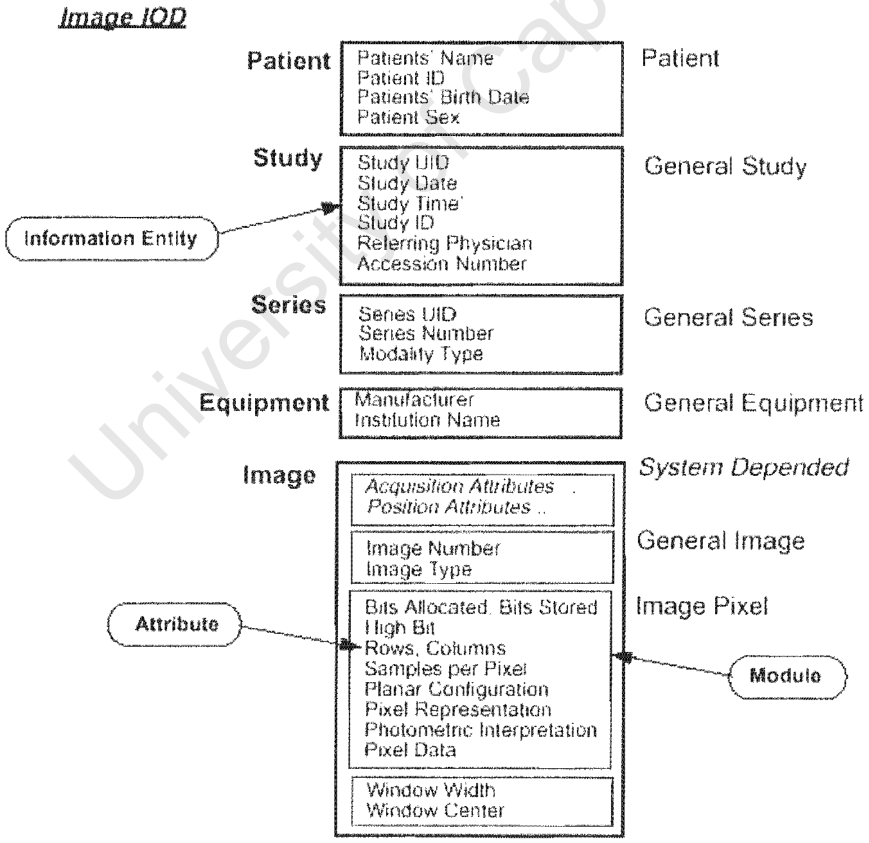


Figure A.1: DICOM information object example (Revet, 1997)

Two of the entries in the information object that might not be clear are the *Window Centre* and *Window Width*. These parameters are necessary because the human eye cannot distinguish more than 256 colours or shades of grey. In essence the scanning system can capture more information than the eye can see (this is only achievable on digital images). Therefore to be able to see all the information we adjust the image to only 256 levels at a time. This is where the two parameters come in, the *Window Centre* states where the 256 window is centred and the *Window Width* states the viewing window of grey levels. A typical LODOX X-ray image is stored as a 14-bit image, which equates to 16384 grey levels. For example, a LODOX image which has a Window Width of 16384 (its maximum) and the Window Centre of 8492 (half the maximum) needs to then convert 16384 grey levels to 256 grey levels (resulting in 64 original grey levels equaling one new grey level). Therefore in this image all the data is viewed but there is a low distinction between various structures in the image. The value of the Window Width and Window Centre represents the current view of the image when it is exported.

A.3 Duality Principle

The Duality Principle is given as follows (Triggs and Mohr, 1996):

> For any projective result established using points and hyperplanes, a symmetrical result holds in which the roles of hyperplanes and points are interchanged: points become planes, the points in a plane become the planes through a point, etc.

Example: A plane is defined as:

$$\bar{n}.\bar{X} = d \tag{A.1}$$

where $\bar{n}$ is the normal to the plane, $\bar{X}$ is a point on the plane and d is the perpendicular distance from the origin to the plane. With $\|\bar{n}\| = 1$ and $d \neq 0$ we can write:

$$\bar{w}.\bar{X} = 1 \tag{A.2}$$

where the plane vector is:

$$\bar{w} = \frac{\bar{n}}{d} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \tag{A.3}$$

where $\bar{w} \neq 0$

Appendix B

Shielding Guides for Diagnostic X-ray Installations

B.1 Barrier Against Leakage Radiation

$$B = \frac{600IPd2}{WT} \tag{B.1}$$

where: P = maximum permissible exposure rate, expressed in R/week. For controlled areas P = 0.04 R/week. For uncontrolled areas, P = 0.002 R/week.

B = Transmission factor.

d = distance in metres from the target to the barrier.

I = tube current in milliamperes.

W = workload in mA-min/week.

T = occupancy factor.

B.1.1 Calculations

Assuming the maximum tube potential of 200mA and a work load of a full body scan (13s) every half an hour ($(200mA\times13s\times2\times24hrs\times7days)/60s = 14560mA.min/week$). We can solve for P the exposure rate at a full occupancy factor ($T = 1$) with no

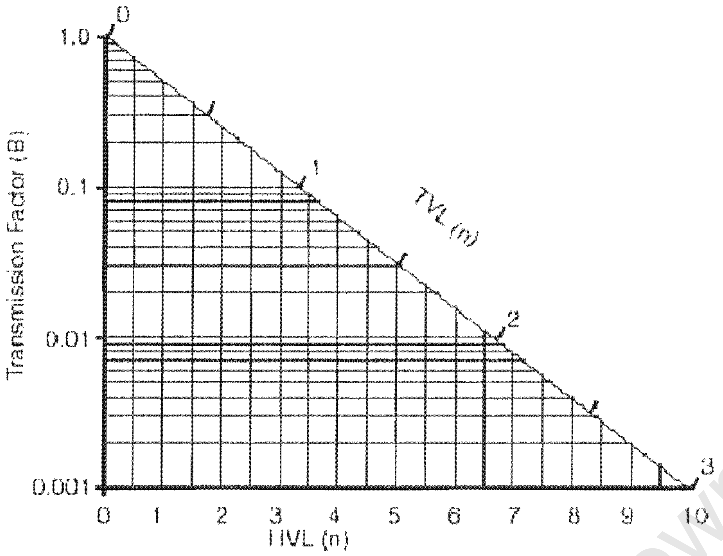


Figure B.1: Relation between the transmission Factor B and the number of half-value layers, N, or tenth-value layers, n (Radiation-1, 1999).

protection ($B = 1$).

$$
\begin{aligned}
P &= \frac{WTB}{600Id2} \\
&= \frac{(14560)(1)(1)}{600(200)(0.5)2} \\
&= 0.1213R/week \\
&= 6.3076R/year
\end{aligned}
$$

B.2 Barrier Against Scatter Radiation

$$K = \frac{400Pd^2D^2}{aWTF} \tag{B.2}$$

Table B.1: Half-value layers and tenth-value layers for heavily filtered X-radiation under broad-beam conditions

Tube Potential	**Attenuation Material for Lead (mm)**	
KVp	HVL	TVL
50	0.06	0.17
70	0.17	0.52
85	0.22	0.73
100	0.27	0.88
125	0.28	0.93
150	0.30	0.99
200	0.52	1.70
250	0.88	2.90
300	1.47	4.80

where: K = exposure per unit workload at 1 metre, expressed in R per mA-min at 1 m.
P = maximum permissible exposure rate, expressed in R/week. For controlled areas P = 0.04 R/week. For uncontrolled areas, P = 0.002 R/week.
d = distance in metres from the target to the scatterer.
D = distance in metres from scatterer to barrier.
a = ratio of scattered to incident exposure (Tabulated in Table B.2).
W = workload in mA-min/week.
T = occupancy factor.
F = field area in cm^2.

B.2.1 Calculations

Assuming the maximum tube potential of 200mA, a work load of a full body scan (13s) every half an hour ($(200mA \times 13s \times 2 \times 24hrs \times 7days)/60s = 14560mA.min/week$)

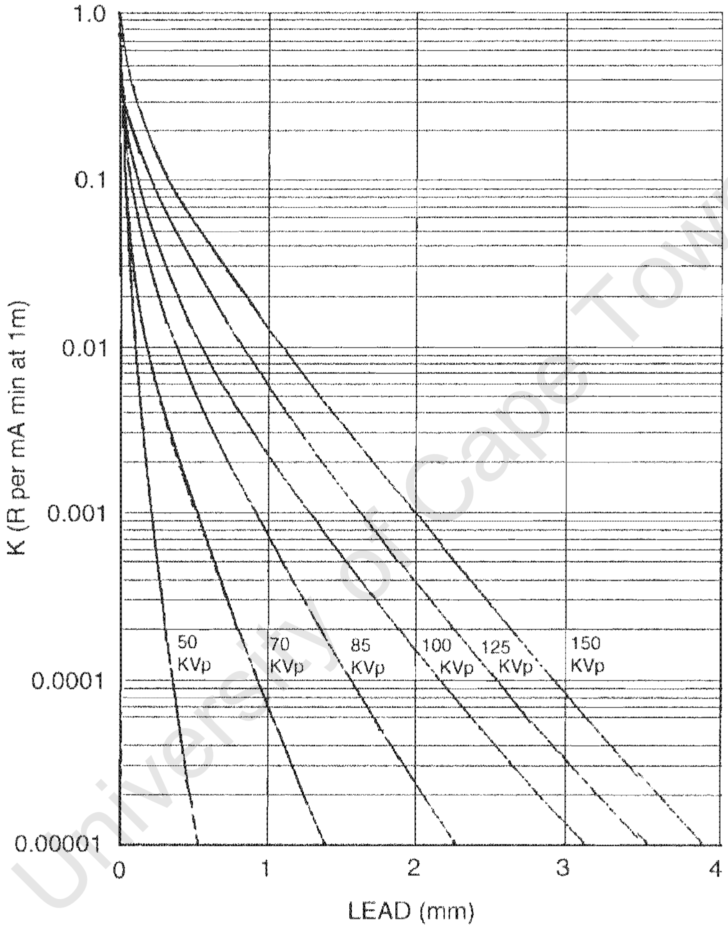


Figure B.2: Attenuation in lead of X-rays generated at 50 to 300 KVp (Radiation-1, 1999)

Table B.2: Ratio, a, of scattered to incident exposure

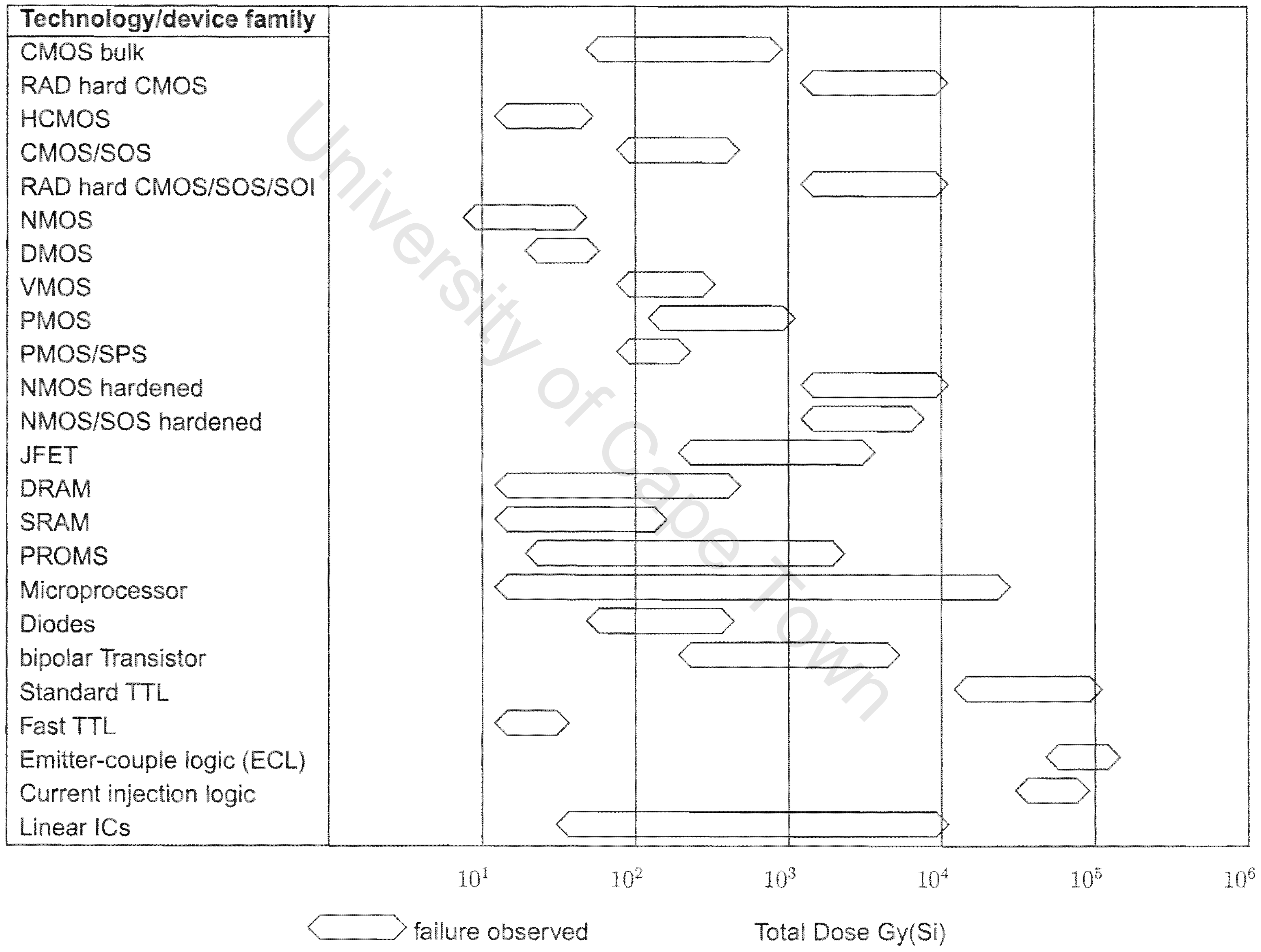
Tube Potential	**Scattering Angle (from Central Axis of Beam)**					
KVp	30°	45°	60°	90°	120°	135°
50	0.0005	0.0002	0.00025	0.00035	0.0008	0.0010
70	0.00065	0.00035	0.00035	0.0005	0.0010	0.0013
85	0.0012	0.0007	0.0007	0.0009	0.0015	0.0017
100	0.0015	0.0012	0.0012	0.0013	0.0020	0.0022
125	0.0018	0.0015	0.0015	0.0015	0.0023	0.0025
150	0.0020	0.0016	0.0016	0.0016	0.0024	0.0026
200	0.0024	0.0020	0.0019	0.0019	0.0027	0.0028
250	0.0025	0.0021	0.0019	0.0019	0.0027	0.0028
300	0.0026	0.0022	0.0020	0.0019	0.0026	0.0028

and a field size of $20cm^2$. We can solve for P the exposure rate at a full occupancy factor ($T = 1$) with no protection ($K = 1$).

$$\begin{aligned} P &= \frac{aWTF}{400Kd^2D^2} \\ &= \frac{(0.0026)(14560)(1)(20)}{(400)(1)(1)^2(1)^2} \\ &= 1.8928R/week \\ &= 98.4256R/year \end{aligned}$$

B.3 Radiation Effects on Semi-conductors

Table B.3: Summary of ionisation and atomic displacement effects (Braäunigand and Wulf, 1994).



Appendix C

Orthographic Projections for Mounting Bracket

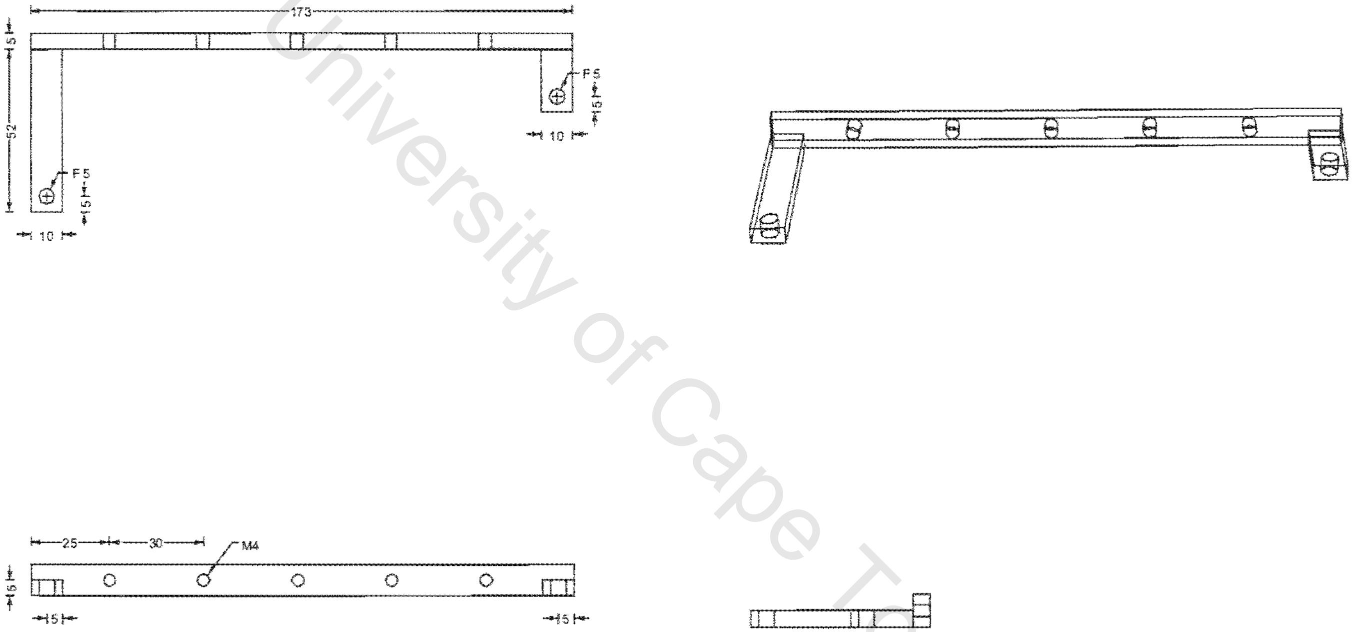


Figure C.1: Dimensions of outer bracket

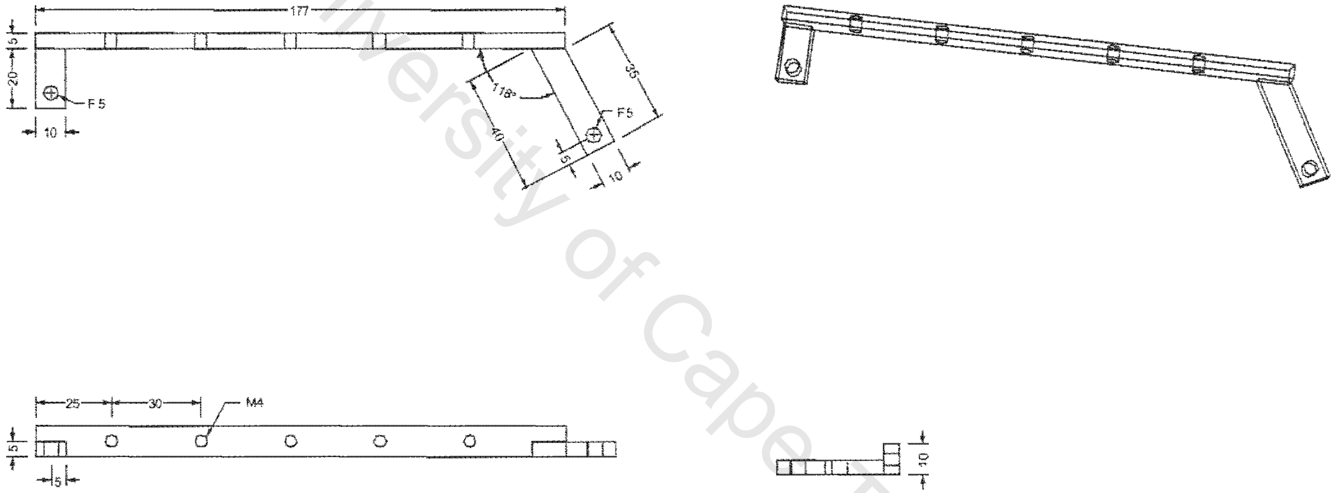


Figure C.2: Dimensions of inner bracket

Appendix D

Hardware Specifications

The following pages contain the specification sheets for the:

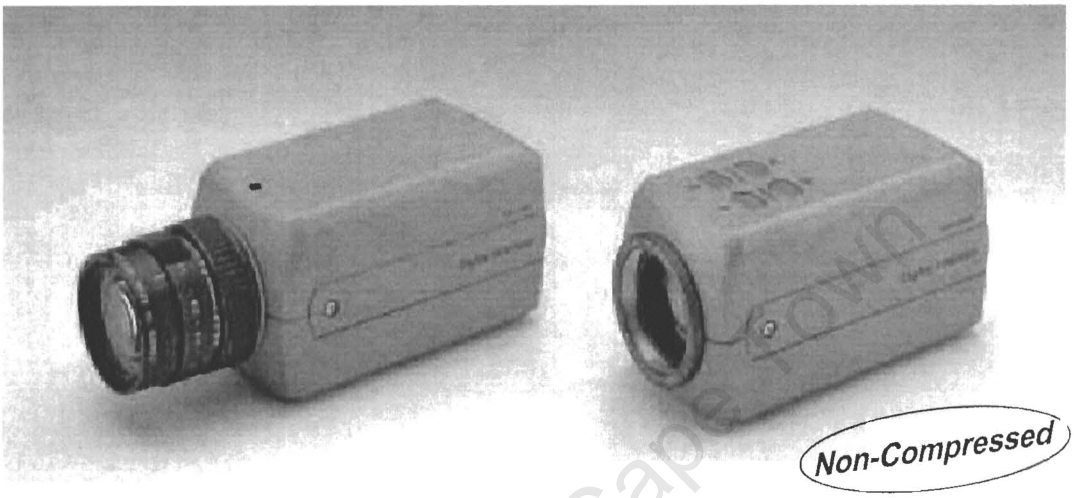
1. Sony DFW-VL500 Digital Camera.
2. Unibrain FireBoard^{TM}400 IEEE1394 interface card.
3. LODOX Digital X-ray Scanner.

SONY

DIGITAL CAMERA MODULE

DFW-V500
DFW-VL500

Component / OEM



OUTLINE

The DFW-V500/VL500 is a fully digital camera which adopts the IEEE1394-1995 standard. Both cameras incorporate the latest 400 Mbps chip set and feature Sony's Wfine CCD™ which integrates a primary color filter ensuring high color accuracy square pixels and Progressive Scan technology which provides sharp, high resolution images, even of fast moving objects. Also, both models include an external trigger mode for asynchronous trigger operation that provides jitter-free pictures since the camera acquisition can be synchronized to full random events.

Through the IEEE1394-1995 high performance serial bus, the DFW-V500/VL500 presents 30 fps in VGA (640×480) resolution format.

The DFW-V500 has a C type lens mount while the DFW-VL500 includes a 12× zoom lens with motorized zoom, iris and focus.

A latching 6-pin IEEE1394 connector is used to output the digital image, to power the camera and to control all functions of the camera through a computer.

*"Wfine CCD™" is a trademark of Sony Corporation.

- External Trigger
- Wfine CCD™

HIGHLIGHTS

- IEEE1394
- External Trigger
- Wfine CCD™
    - Primary Color Filter
    - Progressive Scan
    - Square Pixel
- VGA (640×480), Non-Compressed YUV(4:2:2) Digital Output
- 30 fps Full Motion Picture
- 400 Mbps, High Speed Data Transfers
- Aluminum Diecasting Chassis
- Supplied 6pin Cable with Latch Connector

IEEE1394

Digital Makes
The Difference

IEEE1394-1995

■Standard for High Performance Serial Bus

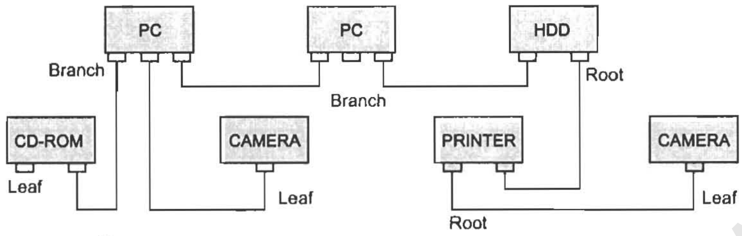
High Speed Data Transferees: 100, 200 and 400 Mbps. The DFW-V500/VL500 adopts 400 Mbps
DS-Link Encoding for highly efficinency Data Transfer

■Plug and Play

No Terminator Placement
Automatic Address Selection

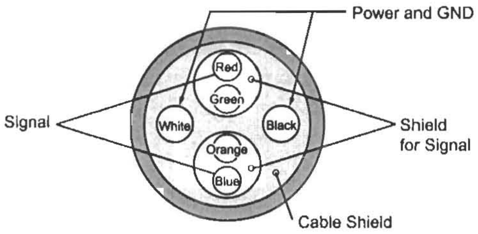
■Flexible Topology

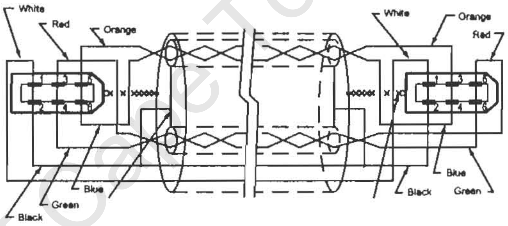
Allowes for Branching and Daisy-chaining: Max of 63 Nodes



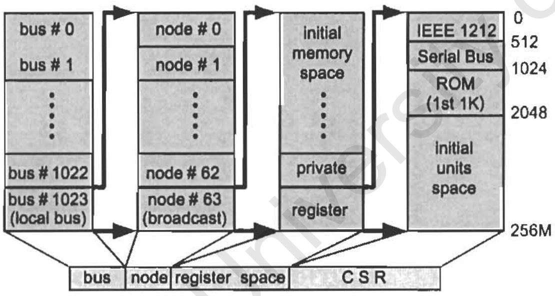
■Small Connector and Cable

6 Conductors Including Power





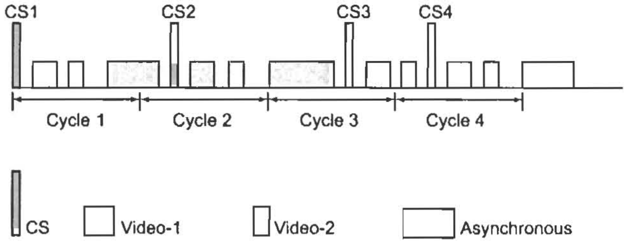
■Addressing



- ● Standardized Address from IEEE1212 Architecture
- ● Configuration ROM from IEEE1212 Architecture
- ● Direct 64 bit Addressing (48 bits per node)

■Protocol

Two Subsections : Asynchronous
Isochronous

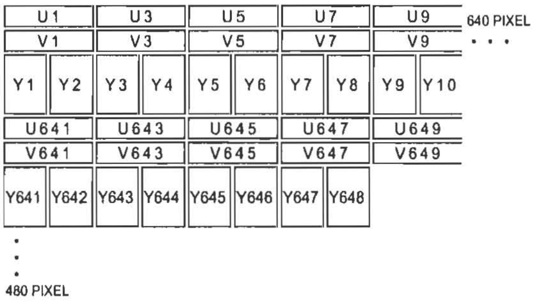


Isochronous

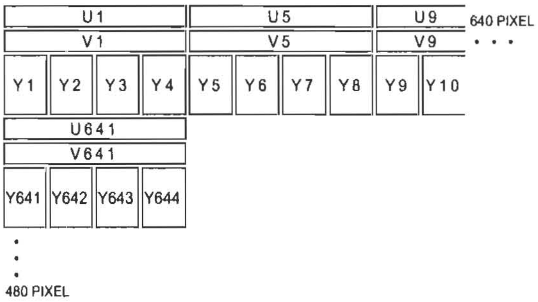
- ● Isochronous prior to Asynchronous
  →Realized Real Time Transfers
- ● Broadcast
- ● Cycle Start appears every 125μ sec.

IMAGE FORMAT

Mode_3 640×480(4:2:2)

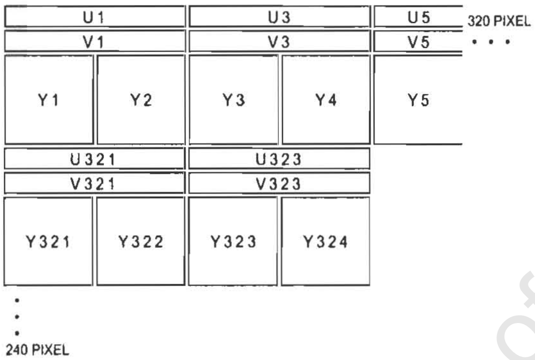


Mode_2 640×480(4:1:1)

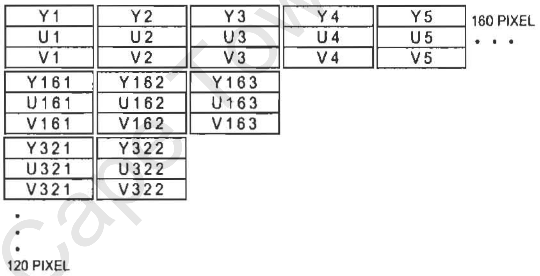


Also, the DFW-V500/VL500 features Mode_1 320×240 YUV(4:2:2) and Mode_0 160×120 YUV(4:4:4)

Mode_1 320×240(4:2:2)



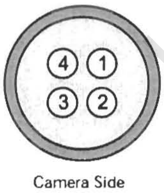
Mode_0 160×120(4:4:4)



EXTERNAL TRIGGER

The DFW-V500/VL500 features on "External Trigger", which is specified by "IEEE1394 based Digital Camera Specification (ver.1.20)". The timing of the image data is controlled by the external trigger. The integration time is defined by the camera itself. (trigger Mode_0)

4pin

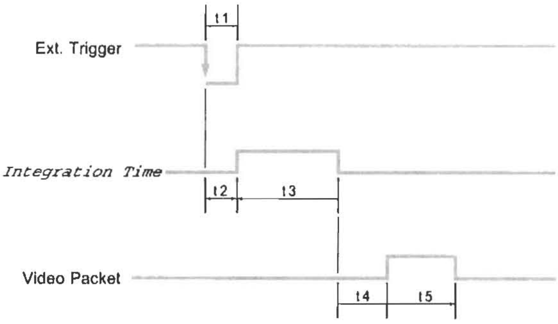


Camera Side

Pin Assignment

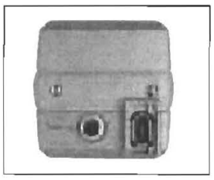
1	NC
2	GND
3	TRIG IN
4	NC

Trigger vs. Packet

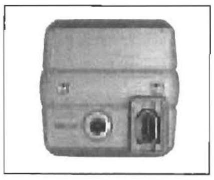


t1 : min. 1 msec
t2 : typ. 1.9 µsec
t3 : integration time
t4 : 1 to 34 msec
t5 : typ. 30 msec (240 packets) [Mode_3, 30 fps]

Rear Panel



DFW-V500



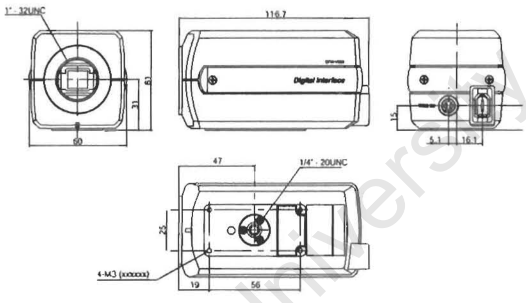
DFW-VL500

SPECIFICATIONS

	DFW-V500	DFW-VL500
Interface Format	IEEE1394-1995	
Data Format	640×480 YUV(4:2:2), YUV 8bit each 640×480 YUV(4:1:1), YUV 8bit each	320×240YUV(4:2:2), YUV 8bit each 160×120YUV(4:4:4), YUV 8bit each
Frame Rate	3.75, 7.5, 15.0, 30 fps and one shot	
Image Device	1/3" Wfine CCD™	
Mini. Senstivity	6 lx (F1.2)	14 lx (F1.8)
White Balance	ATW, One Push , 3200°k, 5600°k, Manual	
Shutter Speed	5~1/15, 1/30~1/100000 sec.	
Gamma	ON/OFF1/OFF2	
Sharpness	Adjustable	
Hue	Adjustable	
Saturation	Adjustable	
Brightness	Adjustable	
Gain	Automatic and Manual Control	
Lens Control	N/A, C mount	12x Zoom Lens, Automatic and Manual
Power	Supplied through IEEE1394-1995 Cable (8 to 30vdc), 4W	
Ope. Temp./Sto.Temp.	-10 to +50°C/-20 to +60°C	
Weight	305g	335g
Supplied Accessory	IEEE1394-1995 Cable with Latch Connector, (6pin), 4pin Connector	

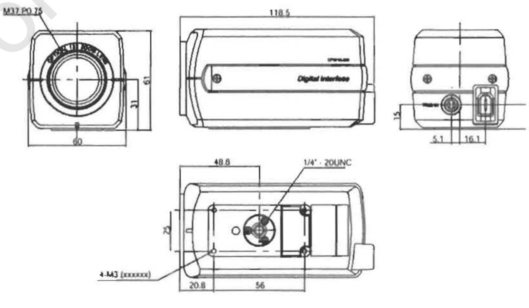
DIMENSIONS

DFW-V500



unit : mm

DFW-VL500



unit : mm

NOTICE

- **If the PC does not have an IEEE1394 connector, a Host Adapter Card is needed.**
- **The DFW-V500/VL500 does not come with any application software or drivers.**
- **The use of the software with customer developed application software may damage hardware, application program and camera. Sony Corporation is not liable for any damages under these conditions.**

● **Sony Electronics Inc. (USA) HQ**	1 Sony Drive Park Ridge, NJ 07656	(TEL:+1-201-930-7451) http://www.sony.com/professional	(FAX:+1-201-358-4401)
● **Sony of Canada Ltd. (CANADA)**	411 Gordon Baker Road, Willowdale, Ontario M2H 2S6	(TEL:+1-416-499-1414)	(FAX:+1-416-497-1774)
● **Sony Broadcast & Professional Europe HQ**	15, rue Floreal 75831 Paris Cedex 17, France	(TEL:+33-1-40-87-35-11) http://www.bpe.sony-europe.com	(FAX:+33-1-40-87-35-17)
Germany	Hugo-Eckener-Str. 20, 50829 Koln	(TEL:+49-221-5966-322)	(FAX:+49-221-5966-491)
France	15, rue Floreal 75831 Paris Cedex 17	(TEL:+33-1-49-45-41-62)	(FAX:+33-1-47-31-13-57)
UK	The Heights, Brooklands, Weybridge, Surrey KT13 0XW	(TEL:+44-990-331122)	(FAX:+44-1932-817011)
Nordic	Per Albin Hanssons vag 20 S-214 32 Malmo Sweden	(TEL:+46-40-190-800)	(FAX:+46-40-190-450)
Italy	Via Galileo Galilei 40 I-20092 Cinisello Balsamo, Milano	(TEL:+39-2-618-38-431)	(FAX:+39-2-618-38-402)
● **Sony Corp. B&P Systems Co. ISP Dpt. (JAPAN)**	4-16-1 Okata, Atsugi-shi, Kanagawa-ken, 243-0021	(TEL:+81-462-27-2345)	(FAX:+81-462-27-2347)

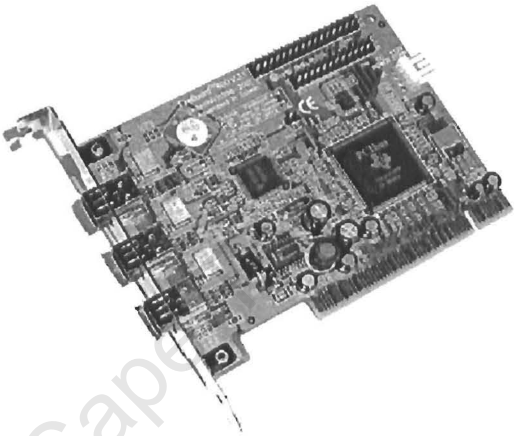
Sony online **http://www.sony.co.jp/ISP/**

98C

unibrain
The 1394 Innovators

FireBoard™400

400 Mbps IEEE-1394 (firewire, i-Link) PCI Host Adapter Board



Features

Physical Layer

- Supports Provisions of IEEE-1394 Standard and the 1394a-2000 Supplement for High Performance Serial Bus.
- Provides Three Fully Compliant Cable Ports at 100/200/400 Megabits per Second.
- Cable Ports Monitor Line Conditions for Active Connection to Remote Node.
- 25.576 MHz Crystal Oscillator and PLL Provide Transmit/Receive Data at 100/200/400 Megabits per Second, and Link-Layer Controller Clock at 50 MHz.
- Node Power-Class Information Signaling for System Power Management.
- Cable Power Presence Monitoring.

Link Layer and PCI Interface

- Performs the Function of all IEEE-1394 Cycle Master.
- Supports IEEE-1394 Transfer Rates of 100, 200 and 400 Megabits per Second.
- Provides PCI Bus Master Function for Supporting DMA Operations.
- Compliant with PCI Specification 2.1.
- Supports Distributed DMA Transfers between 1394 and Local Bus RAM, ROM, AUX, or Zoomed Video
- **Extension Headers** for the Local Bus
- **Autoboot** mode capability that allows data-moving systems to be designed to operate on the PCI bus without the need for a host CPU.

Extension Capabilities

FireBoard™ 400 can be extended in functionality and operability. This can be achieved combining the Autoboot mode capability and the Extension Headers of the Local Bus. You can design your custom system based on the PCILynx local bus as a biggy-back board attached to the FireBoard™ 400.

Unibrain's FireBoard™400 Host Adapter brings high-performance IEEE-1394 (firewire, i-Link) technology to PCI systems. **FireBoard™ 400 Design Issues**

Logic Design
The entire design of Unibrain's FireBoard™400 is based on two main IEEE-1394 chip-sets:

- the Texas Instruments **TSB41LAB3** IEEE-1394 400Mbps PHY layer transceiver and
- the Texas Instruments **TSB12LV21B (PCILynx-2)** IEEE-1394 link controller to PCI bus.

PCB Design
The FireBoard™400 card is a four-layer printed circuit PCI board. Two of the layers provide the signal interconnections (signal layers) while the remaining two layers provide the required power and ground interconnections (plane layers). ***External Power Supply***
The FireBoard™400 has an has an optional power connector in order to receive the cable power not only from the +12V of the PCI connector but also from an external power connector provided by the host PC power supply.

LODOX DIGITAL X-RAY SCANNER
TECHNICAL SPECIFICATION

Applicability This specification is applicable to the LODOX PPM (Pre-Production Model Digital X-ray Scanner, Model LDXPPM.

Notice These specifications were correct at the time of writing. Where parameters are considered critical or marginal, these parameters should confirmed with the original manufacturer.

Image Quality	
Contrast Resolution	>16000 grey levels
Spatial resolution:	
Fundamental pixel size.	60 ?m (1x1 binning)
Normal resolution mode	1.67 lp/mm at 5x5 pixel binning Maximum size = full body (2500 x 6000 pixels)
High resolution mode	2.78 lp/mm at 3x3 pixel binning Maximum size = full body (4000 x 10000 pixels)
Ultra high resolution mode	4.17 lp/mm at 2x2 pixel binning Maximum size = half of full body size
Test mode	8.34 lp/mm at 1x1 pixel binning
Image flexibility	
Imaging angles	0 to 90 degrees. The supine or prone patient can be x-rayed in any radial angle from A-P (or P-A) to the supine lateral positions.
	The trolley(gurney) top height can be adjusted vertically by 0 to 340mm, by adjusting only one side of the trolley gives Trendelenburg angles up to 10 degrees.
Scanner throughput	
Linear scanning rate (3 settings)	Variable from 34.5mm/s, 69mm/s, and 138mm/s
Instantaneous frame rate	22 - 88 milliseconds
Time to complete a full field scan	10 seconds
Time from "end-of-scan" until a diagnostic image becomes available on the DVS screen	15 seconds
Best case time between two successive x-rays on the same patient.	32 seconds, can optionally be decreased
Diagnostic viewing station	
Type	Barco 21" High Luminance (120Cd/m²) Medical-quality monitor, Type MGD 321
Resolution	1600 x 1200 pixels
Image Viewing	
Image recall	Images viewed by Previous, Next, Latest, Patient name, Referring doctor, Hospital number etc
Database	Images in the database belonging to a study or patient can be easily retrieved.
Zoom Functions	Various zoom functions exist for viewing the image. The image can also be rotated in different directions
Image Manipulation	Unsharp masking "flattens" the image, so that the entire image can be viewed without grey scale adjustments.
Palette Function	The palette function allows the user to independently select the range of grey scales that can be displayed
Invert Function	Changes the view from a "black bone" to a "white bone" image.
Image size	Variable dependant on exposure area, as well as resolution E.G. Full-body, normal resolution, approximately 30Mbyte. Equates to 300 images on a 9Gb disk
Input	
Input mains Supply	380 volt Three Phase AC (50 Hz). Permanent Connection to supply. The LODOX PPM is designed to operate at the indicated voltages. Operation at US supply voltages is possible with modifications (to ANSI C84.1 1995). Maximum peak demand 160 amp

Output	
Radiation Type	This equipment emits ionized X-radiation in a narrow fan-beam across the patient in order to produce an image.
Direct Absorbed dose	Maximum direct dose absorption is 1 mGy
Leakage Radiation	The maximum leakage radiation measured 1 meter from the X-ray focal spot in any direction is less than 20 uGy per hour @ 130kV 25mA
X-ray Tube	Rotating Anode type. CE approved
X-ray Generator	Peak Power 40 KW Voltage Range 40kVp - 135 kVp Voltage Rise Time 2mS (10 to 90%) Voltage Stability <4% Voltage Ripple 0.5% Current Range 25mA to 400mA Exposure Time 0.6s to 13 s Operating frequency 100kHz CE Approved
Environment	
Temperature	+10 to + 35° Celcius.
Humidity	20% to 80% non condensing
Atmospheric Pressure	540 hPa to 1035 hPa.
Image Information	
Image size	680mm x 1800mm, full body. Adjustable from 100mm X 100mm to full size in any position on the patient
Detector System	
Detector	12 CCD cameras
X-ray to light conversion	Rarex Green Fast ("Gadox" ($GdOS_2$:Tb))
Dimensions	
Scanner	
Height	3000 mm
Width	2500 mm
Length	2900 mm
Electrical Cabinet	
Height	1800 mm
Width	950 mm
Depth	430 mm
Operator Console	
Height	1450 mm
Width	790 mm
Depth	670 mm

Appendix E

Photos of Cameras Mounted on C-arm

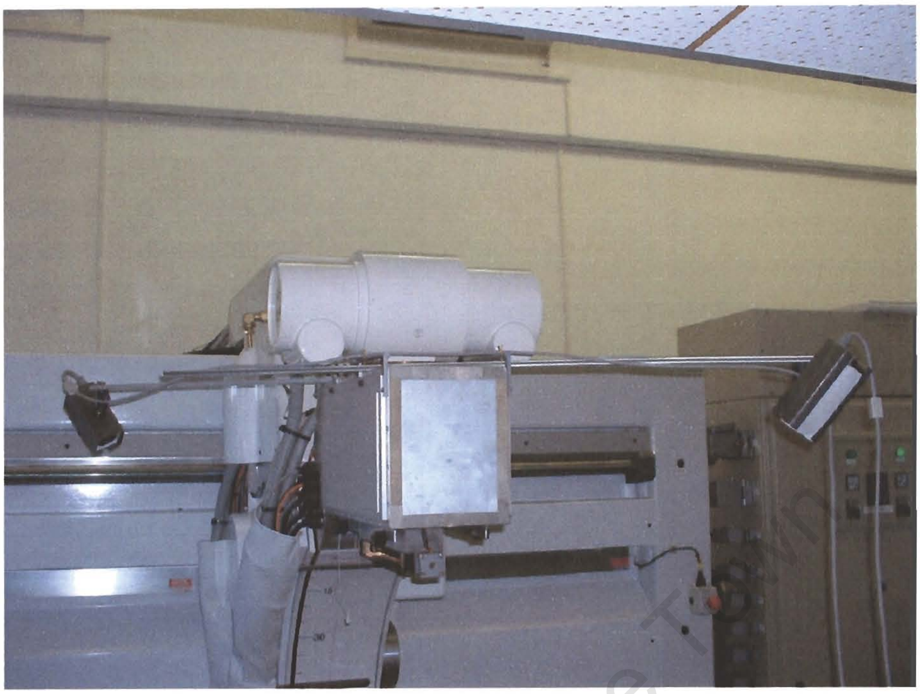


Figure E.1: Cameras positioned 180° apart.

Figure E.2: Cameras positioned 60° apart.

Appendix F

Operation Manual for Programs

This document describes the operating procedures used to do 3D reconstruction on the LODOX system.

F.1 Pre-Scanning Phase

The pre-scanning phase consists of:

1. Camera Position
2. Calibration of camera and light source.

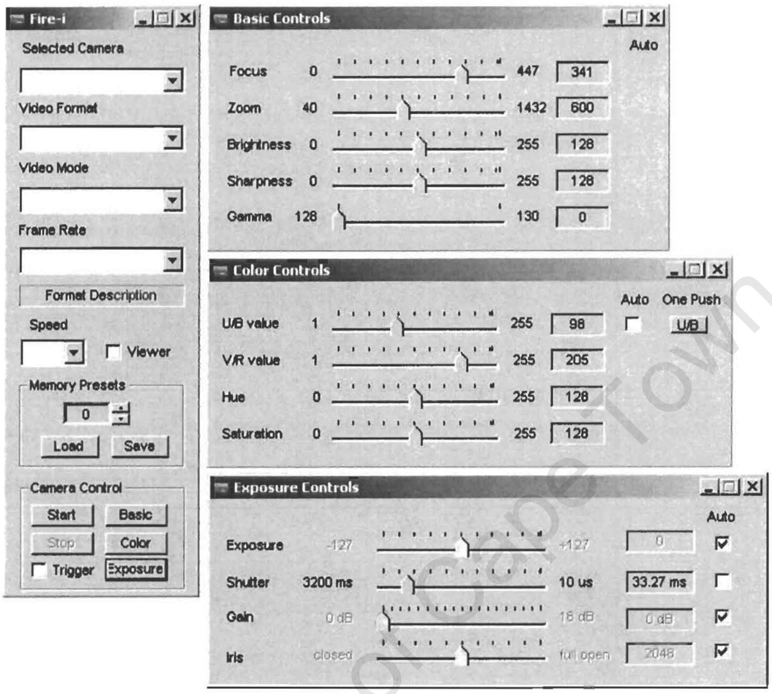
F.1.1 Camera Position

Requires program: Fire-i.exe

Firstly mount the cameras in the desired position and connect the cameras to the computer. Start the Fire-i program to see what the cameras are capturing. Secondly, position the cameras so that it captures an area where the laser line can be seen on the tabletop. To view the camera output (images) you need to initialise the cameras:

1. Start the camera control software called Fire-i.

2. Select the camera, video format and mode, frame rate and speed from the drop-down list boxes. See camera specifications for explanation of the various modes.
3. Click **Start**.



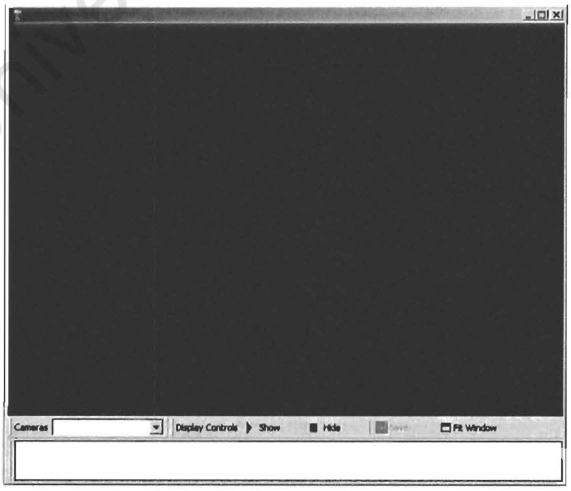
☐ Viewer
Memory Presets 0
☐ Trigger

Focus 0 447 341
Zoom 40 1432 600
Brightness 0 255 128
Sharpness 0 255 128
Gamma 128 130 0

U/B value 1 255 98 ☐
V/R value 1 255 205
Hue 0 255 128
Saturation 0 255 128

Exposure -127 +127 0 ☑
Shutter 3200 ms 10 us 33.27 ms ☐
Gain 0 dB 18 dB 0 dB ☑
Iris closed full open 2048 ☑

The camera is now initialised and running. To view the camera output check the **Viewer** tick-box, this will start the following program:



1. Check that that correct camera is selected in the **Camera** list box.

2. Click **Show** to display the camera images on the screen.
3. The cameras output can be controlled by clicking **Basic**, **Color** and **Exposure** on the Fire-i dialog box.
4. To save a single bitmap image click **Save**. The file will be saved to the current directory.

With the ability to see what the camera is capturing position the laser line from the top to bottom of the view (does not have to be vertical). This is necessary because the reconstruction algorithm detects edges from left to right. The cameras zoom and focus settings can be set to any value but cannot be changed after calibration.

F.1.2 Calibration

Requires program:	Fire-i.exe Viewer (captures 1 cam_cal_).exe Viewer (captures 1 light_cal_).exe Matlab
Requires files:	calib_tool.m (main file - sub-files see Calibration Tools directory) convert_laser.m
Requires hardware:	Checkerboard Vertical object with straight edge

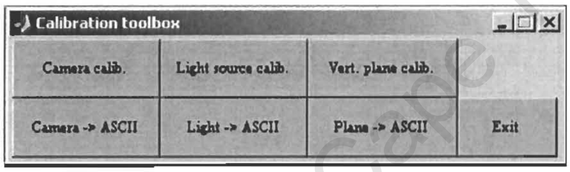
Each camera must be calibrated using the checkerboard to extract the camera's internal and external parameters as well as to obtain the position of the laser light source with respect to the cameras position.

For each camera:

1. Initialise the camera as mentioned above if not already done.
2. Take about 5 images of the checkerboard using the **Viewer(captures 1 cam_cal_).exe** program instead of checking the viewer checkbox. This program saves indexed frames with a filename of 'cam_cal_X.bmp', where X is a number. The order of the frames should be: the first image with the check-board flat on the tabletop/ground-plane and then in other various positions.

3. Next take about 5 images of the vertical edge object so that the base of the vertical edge can be seen together with the shadow that the edge casts. Use the **Viewer(captures 1 light_cal_).exe** program to save with a 'light_cal_X.bmp' filename.
4. The images now have to be converted to a compatible format for the calibration program using the **conver_laser.m** file. This file converts the camera images to greyscale while light images are filtered so only the red band of the image remains. Both the camera and light images are also rotated to take into account the upsidedown position of the camera.

With the correct files, run the calibration program to obtain the parameter files needed for the reconstruction program. Run the **calib_tool.m** file to get a dialog box as follows:



1. Click **Camera Calib** to start the camera calibration. Follow the instructions for the calibration.
2. Click **Camera -> ASCII** to write the camera calibration file, 'camera_file'.
3. Click **Light source calib** to start the camera calibration. Follow the instructions for the calibration.
4. Click **Light -> ASCII** to write the light calibration file, 'light_file'.

The 'camera_file' and 'light_file' contain the camera's internal parameters and position of the light source with respect to the camera, respectively. These will be needed for the reconstruction process. The **Vert. plane calib.** and **Plane -> ASCII** are functions for the shadow scanner and are not used.

F.2 Scanning Phase

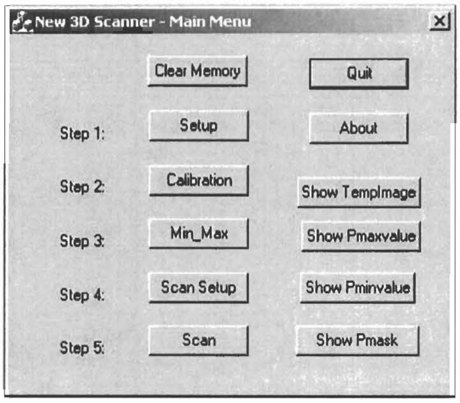
Requires program:	Fire-i.exe Viewer (captures 300 frame).exe Matlab
Requires files:	convert_laser.m

1. Initialise the camera as mentioned above if not already done.
2. Take a sequence of images while moving the laser across the scene with the **Viewer (captures 300 frame).exe** program. This program saves indexed frames with a filename of 'frame_XXX.bmp', where XXX is a three digit number.
3. Convert the images as mentioned before using the **conver_laser.m** file.

If a 3D image is required with an X-ray. This scanning phase must be done simultaneously when the X-ray is captured. If no X-ray is required use the Local Positioning Console (LPC) to move the C-arm.

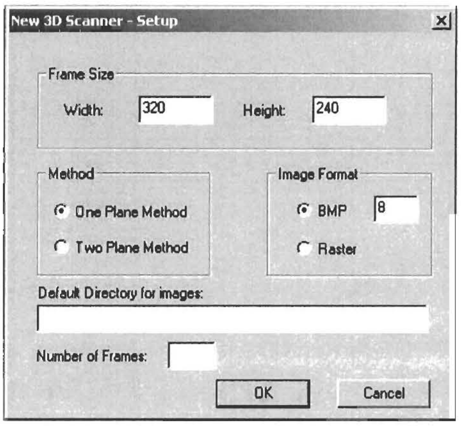
F.3 Post-Scanning Phase

Requires program:	Laser Scanner.exe Matlab
Requires files:	run_cal_merge_laser.m run_join_2_vrmls.m



1. Start the **Laser Scanner.exe** program. The initial dialog box will appear.
2. Click the **Setup** button to display the setup dialog box where you need to input parameters for the scan. These will depend on where the

frames from the scan are stored, how many there are, what format and size each frame is. Click **OK** when done.


New 3D Scanner - Setup

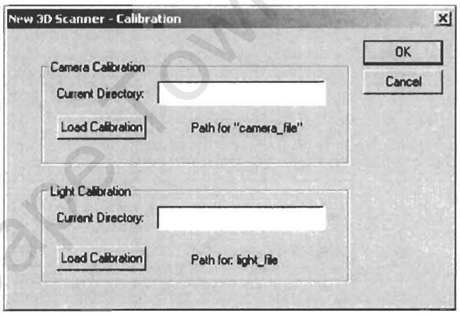
Frame Size
Width: 320
Height: 240

Method
● One Plane Method
○ Two Plane Method

Image Format
● BMP 8
○ Raster

Default Directory for images:
Number of Frames:

3. Click **Calibration** button to display the calibration dialog box:


New 3D Scanner - Calibration

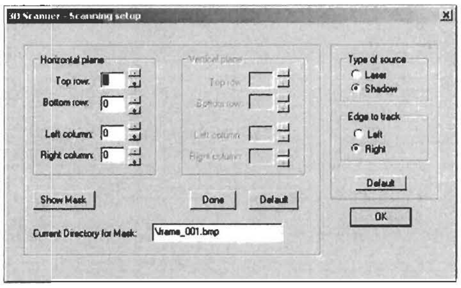
Camera Calibration
Current Directory:
Path for "camera_file"

Light Calibration
Current Directory:
Path for: light_file

4. Type in the path of the calibration files mentioned above (*i.e.* 'camera_file' and 'light_file'). Click **Load Calibration** for each file. Click **OK** when done.

5. The **Min_Max** button is not used in the Laser scanner program

6. Click **Scan Setup** to define an area to look for the laser edge.


3D Scanner - Scanning setup

Horizontal plane
Top row:
Bottom row: 0
Left column: 0
Right column: 0

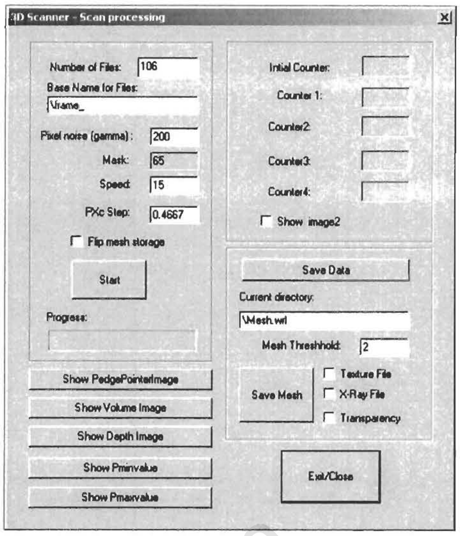
Vertical plane
Top row:
Bottom row:
Left column:
Right column:

Type of source
○ Laser
● Shadow

Edge to track
○ Left
● Right

Current Directory for Mask: \frame_001.bmp

7. Adjust the bounding box and click **Show Mask** to display the bounding box on the first image. Click **OK** when done.



Number of Files: 106
Base Name for Files: \frame_
Pixel noise (gamma): 200
Mask: 65
Speed: 15
PXc Step: 0.4667
☐ Flip mesh storage
Current directory: \Mesh.wrl
Mesh Threshhold: 2
☐ Texture File
☐ X-Ray File
☐ Transparency

8. Click **Scan**
9. Check that the parameters for the scan is set correctly (*i.e.* Paths, PXc step - related to scan speed, Gamma - the edge threshold value, Speed - no used).
10. Click **Start** to process the frames.
11. To save the 3D coordinate points as text files click **Save Data**.
12. Set mesh parameters (*i.e.* Mesh Threshold, Texture, X-ray plane, Transparency).
13. Click **Make Mesh** to generate the VRML file.
14. Click **Exit/Close** to exit we done.

The above process needs to be completed to generate a 3D image file for each camera used. After the two file are generated, they need to be merged. Merging is currently manual and is first done be running **run_cal_merge_laser.m** which will prompt you to crop the image and step the one image to merge with the other. When you are satisfied with the merge, the steps will be outputted to a **final_moves.mat** file. Using **run_join_2_vrmls.m** the two complete 3D image files will be merged into a new 3D image file using the previously generated **final_moves.mat** file.

Appendix G

Source Code for Project

G.1 Visual C++

On the attached CD-ROM the source code and pre-build executables for the following programs can be found:

1. Laser Scanner
2. Shadow Scanner
3. Viewer

These files were written and compiled on a Windows 2000 operating system using Visual C++ 6.0

G.2 Matlab

A description of the Matlab m-files written is shown in Table G.1 while the actual source files can be found on the attached CD-ROM.

File Name	Type	Description
cal_merge1.m	function	This function lets you merge two 3D data sets.
check_index.m	function	Makes sure that the coords do not exceed the index values.
check_line.m	batch job	This m-file checks plots the points one at a time to see how straight a reconstructed line is.
clean_up.m	function	This function lets you crop the points from views from the 3 axes.
Clip.m	function	This function lets you crop the points from views from the 3 axes.
convert.m	batch job	Converts calibration frames and captured frames from RGB to Gray Scale
convert_bmp2jpg.m	batch job	Converts captured frames from RGB to JPG
convert_laser.m	batch job	Converts calibration frames and captured frames from RGB to Gray Scale together with 180 degree rotation
dicom_to_grayscale.m	function	This function reads in a DICOM file and outputs the greyscale image using the WindowWidth and WindowCenter.
drop_view.m	batch job	Makes the minimum heights in View1, View2 and View3 equal to zero.
filter_index.m	function	This function makes sure that there are no large triangles in the mesh i.e. makes sure that triangles consist of only close points.
filter_index1.m	function	This function makes sure that there are no large triangles in the mesh i.e. makes sure that triangles consist of only close points.
find_average_change.m	batch job	This batch job is a template for running jobs.
find_triangles.m	batch job	This m-file eliminates all zeros in the index matrix.
gen_texture_image.m	batch job	This m-file prompts the user for info regarding the xray image and generates the JPG texture file.

File Name	Type	Description
gen_texture_image2.m	batch job	This m-file prompts the user for info regarding the xray image and generates the JPG texture file.
gen_texture_image3.m	batch job	This m-file prompts the user for info regarding the xray image and generates the JPG texture file.
Im24_to_Im8_temp.m	batch job	This m-file converts files "pic_XX.bmp" to "frame_XXX.jpg".
join_index.m	function	This functions combines the index matrices of 2 views
Line_calib.m	function	Determines the change that needs to be applied to make line straight
load_coords.m	batch job	Loads the coordinates for the mesh and its coordinate index.
make_coords.m	batch job	Loads the raw coords for each pixel - PXc, PYc and PZc and generates the coord file.
make_index.m	batch job	Joins the point to make triangles
make_simp_coord_matrix.m	batch job	Create the simplified matrix with coords of laser in a straight vert line instead of a skew one as stored in PXc etc. Is is already done in C++ code.
make_texture_coords.m	batch job	This file works out the texture coords for the VRML file (old gen_texture_image file)
merge.m	function	This function lets you merge 2 3D data sets.
mesh2vrml_new.m	function	Generate a VRML file from a set of triangulated points I.e. write VRML file
read_coords.m	function	Loads the coordinates for the mesh file and its coordinate index and translation between camera and world reference frame.
readpgm.m	function	READPGM Read a raw pgm file as a matrix , Matthew Dailey, 1997
reindex.m	function	Re-index the index when coords are removed from the coord matrix.

File Name	Type	Description
run_cal_merge_laser.m	batch job	This file loads the coords for the two views and cleans up the views for calibrating. The calibrating finds the translation and rotation with respect to the two views or cameras.
run_job.m	batch job	This m-file is to merge the three views of a pyramid in the shadow reconstructions.
run_join_2_vrmls.m	batch job	This m-file is used to write a different version of a VRML file, in that two different indexed face sets are written instead of one.
run_line_mesh.m	batch job	This m-file shows which points are being connected together in the mesh.
run_make_xray_jpg.m	batch job	This m-file finds a solid back line (with in a threshold) and works out the edge of the x-ray image and outputs a new image with just the xrays and not full camera view.
run_make_xray_jpg_new.m	batch job	This m-file finds a solid back line (with in a threshold) and works out the edge of the x-ray image and outputs a new image with just the xrays and not full camera view.
run_merge.m	batch job	This m-file was used to merge the 3 shadow views of the pyramid, uses the same process as cal_merge.
run_merge_laser.m	batch job	Reads in the coords for the two views as well as the final_moves for merging (rotation and translation) I.e. merges files automatically and prompts if simplification must be done.
run_simplify_1_vrml.m	batch job	This m-file reads in coords and index and prompts user if simplification must be done.
show_merge.m	function	A 3D view of two objects...ViewA is plotted in red, ViewB is plotted in blue
show_plot_row_xray.m	batch job	The m-file plots each row of the DICOM(greyscale image).

File Name	Type	Description
show_plot_rows.m	batch job	This m-file shows the colour profile of image (one frame)
show_view3d.m	function	A few 3D view of one object...
simp_struc.m	function	This function simplifies the structure
smooth_mesh_new.m	function	This function does the smoothing of a mesh
smooth_run.m	batch job	This file prompts the user if the coords and index need to be read into memory and asks how many smoothing operations the user wants done on the 3D image. The result is written as 'test.wrl' in the current directory.
test_delunay_method.m	batch job	M-file testing the use of delunay meshing
writepgm.m	function	Write a matrix as a raw pgm file, Matthew Dailey 1999
writevrml.m	function	Generate a VRML file from two sets of triangulated points
zero_at_tip.m	function	This function makes the xy origin to be at the tip of pyramid.

Table G.1: Description of Matlab source code m-files.

Appendix H

More Reconstruction Examples

A list of all the reconstruction examples conducted and their reconstruction quality is listed in Table H.1. The VRML files can be found in the '\Experiment' directory on the attached CD-ROM.

Directory	Cam 4 /Base	Cam 7 /Out	Colour Frames	Result Quality	Merged Quality	Comments
Shadow Experiments						
\Experiment 10 May (lecture room)		75		good		Mouse
\Experiment 18 May (Dark Room)		430		bad		Mouse
\Experiment 24 July (Student Room)		240 & 100		bad		Pyramid, two light source positions
\Experiment 28 June (Dark Room) - ken		250 & 100		none		No reconstruction due to no xtop or xbot therefore cannot extract shadow plane.
\Experiment 28 June (Dark Room) - Nose		225 & 155		none		No reconstruction due to no xtop or xbot therefore cannot extract shadow plane.
\Experiment 31 May (Dark Room)		400		bad		Glue-stick and Smint box
\Experiment 4 July (Room)		213		none		plastic Memory holder
\Experiment 6 June (Dark Room)		487		bad		Toilet roll and Smint box
\ExperimentA 26 July (Student Room)		40		none		3D triangle, surface not reflective enough
\First Experiment						
\Paper Cylinder Experiment						
\Second Experiment						

Directory	Cam 4 /Base	Cam 7 /Out	Colour Frames	Result Quality	Merged Quality	Comments
\Experiment 20 June (Dark Room)		340		good		
\Experiment 25 May (Dark Room)		496		medium		
Laser Experiments						
\2001-10-30 (blocks - 84)		84		bad		
\2001-10-30 (pyramid - 66)		66		good		
\2001-10-30 (triangle - 50)		50		good		mirror
\2001-11-07 (skull)		106	Yes	bad		
\2001-11-07 (tea box)		78	Yes	medium		mirror
\2001-11-09 (tea box half-speed)		155	Yes	good		
\2001-11-19 (hand half-speed)		148		good		
\2001-11-19 (Mark half-speed)		200	Yes	good		
\2001-11-19 (mask half-speed)		181		good		
\2001-12-05 (calibration)	81	76		good		
\2001-12-05 (Ken)	113	108		medium	medium	
\2001-12-05 (Radiation Test)		77				
\2001-12-07 (Mark)	188	196		bad		
\2001-12-07 (pyramid)	118	116		medium	good	Thin collimator width
\2001-12-10 (mouse_toolbox)	73	75	Yes	good	good	X-ray, alignment not correct
\2001-12-12 (toolbox)		55	Yes	good		X-ray, alignment better

Directory	Cam 4 /Base	Cam 7 /Out	Colour Frames	Result Quality	Merged Quality	Comments
\2002-02-09 (pyramid - dark - 1)	132	128		good	good	Done at night - dark environment
\2002-02-09 (pyramid 1)	134	129		good		
\2002-02-11 (box mirror)		74		good		mirror
\2002-02-11 (cup)\7		75		medium		
\2002-02-11 (cup)\7 and mirror		50		medium		mirror
\2002-02-11 (mouse xray)	87	95		good		
\2002-02-12 (Wayne dual - 1)	360	344	Yes	good	good	
\2002-02-13 (mouse)		50		good		X-ray, alignment still bit out
\2002-02-15 (mark half)	180	151		good		Cameras 180 degrees apart
\2002-02-15 (mark2 half)	201	168		good	good	Cameras 180 degrees apart
\2002-02-18 (Wayne)		87		medium		X-ray, alignment okay

Table H.1: List of all experiments conducted.

References

3D Format, 2002. The Graphics File Formats Page. Last accessed: 2002/01/15.
URL `http://www.dcs.ed.ac.uk/home/mxr/gfx/3d-hi.html`

3D-MATIN, 2002. 3D Face Imaging for Maxillofacial Surgical Planning. Last accessed: 2002/02/25.
URL `http://www.faraday.gla.ac.uk/medical.htm`

Abi-Rached, H., Duchamp, T., Shapiro, L. G., Stuetzle, W., 1998. Acquisition and visualization of colored 3d objects. In: *Proceedings of ICPR 98*, pp. 11–15, pp. 11–15.

Adams, L., 2001. Guidelines for the use of electronic components in the space radiation environment. Prepared by: SPUR Electron Ltd.

Addleman, D., Addleman, L., 1985. Rapid 3D digitising: an innovative technique holds promise for a variety of applications. *Comp Graph World* 11, 41–44.

Andrew, J., Aldrich, J., 1989. A video-based patient contour acquisition system for the design of radiotherapy compensators. *Medical Physics* 16 (3), 425–430.

Ayoub, A., Wray, D., Moos, K., Siebert, P., Jin, J., Niblett, T., Urquhart, C., Mowforth, R., 1996. Three-dimensional modeling for modern diagnosis and planning in maxillofacial surgery. *The International Journal of Adult Orthodontics and Orthognathic Surgery* 11 (3), 225–233.

Bajaj, C., Bernardini, F., Xu, G., 1997. Reconstructing surfaces and functions on surfaces from unorganized 3D data. *Algorithmica* 19, 243–261.

Barros, R., Figueroa, P., Paterniani, P., Brenzikofer, R., Limo Filho, E., 2002. Three-dimensional reconstruction and analysis of human back surface. In: *3-D Analysis of Human Movement*, Seventh International Symposium, pp. 17–20, pp. 17–20.

Bergevin, R., Laurendeau, D., Poussart, D., 1995. Registering range views of multipart objects. *Computer and Vision Image Understanding* 61 (1), 1–16.

Bernardini, F., Mittleman, J., Rushmeier, H., Silva, C., Taubin, G., 1999. The ball-pivoting algorithm for surface reconstruction. *IEEE Transactions on Visualization and Computer Graphics* 5 (4), 1–10.

Besl, P., McKay, N., 1992. A method of registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14 (2), 239–256.

Bibb, R., Freeman, P., Brown, R., Sugar, A., Evans, P., Bocca, A., 2000. An investigation of three-dimensional scanning of human body surfaces and its use in the design and manufacture of prostheses. *Proceedings of Institution of Mechanical Engineers - H Engineering in Medicine* 214, 589–594.

Bouguet, J., 1999. *Visual methods for three-dimensional modeling.*. Phd thesis, California Institute of Technology, Pasadena, California.

Braäunigand, D., Wulf, F., 1994. Atomic displacement and total ionizing dose damage in semiconductors. *Radiation Physics and Chemistry* 43 (1), 105–127.

Calibration, 2002. Camera Calibration Toolbox for Matlab. Last accessed: 2001/06/05.
URL `http://www.vision.caltech.edu/bouguetj/calib_doc/links.html`

Cheng, M., 1997. Stereoscopic motion tracking in biomedical visualisation. The University of Western Australia. Graduate Diploma.

Chu, C. W., Hwang, S., Jung, S. K., 2001. Calibration-free approach to 3d reconstruction using light stripe projections on a cube frame. IEEE, Third International Conference on 3D Digital Imaging and Modeling, Quebec City, Canada.

Chua, C., Jarvis, R., 1996. 3-d free-form surface registration and object recognition. *International Journal of Computer Vision* 17, 77–99.

Clunie, D. A., 2001. Medical Image Format FAQ. Last accessed: 2001/10/13.
URL `http://www.dclunie.com/medical-image-faq/`

Costello, P., Bee, S., 1997. *3D Model Databases: The Availability of 3D Models on the World Wide Web.*. Tech. rep., Advanced VR Research Centre, Dept. of Human Sciences, Loughborough University of Technology.

Creath, K., Wyant, J., 1992. *Optical Shop Testing*, Wiley, New York, chap. Moir and Fringe Projection Techniques. 2nd ed., p. 657.

Curless, B., 1999. From range scans to 3d models. *Computer Graphics* 33 (4).

Curless, B., Levoy, M., 1995. Better optical triangulation through space-time analysis. In: *Proceedings in IEEE Conference on Computer Vision (ICCV)*, pp. 987–994, pp. 987–994.

Cyberware, 2001. Cyberware home page. Last accessed: 2001/02/20.
URL `http://www.cyberware.com/products/wbInfo.html`

Daley, R., Hassebrook, L., 1998. Channel capacity model of binary encoded structured light-stripe illumination. *APPLIED OPTICS* 37 (17), 3689–3696.

D'Apuzzo, N., 1998. Human body 3D imaging by speckle texture projection photogrammetry. *International Archives of Photogrammetry and Remote Sensing, Hakodate, Japan* 32 (B5), 402–407.

Debevec, P., Hawkins, T., Tchou, C., Duiker, H., Sarokin, W., Sagar, M., 2000. Acquiring the Reflectance Field of a Human Face. In: *SIGGRAPH 2000 Conference Proceedings*, University of California at Berkeley.

Debevec, P. E., 1996. *Modeling and Rendering Architecture from Photographs*. Phd thesis, University of California at Berkeley, Computer Science Division, Berkeley CA.

Dendy, P. P., Heaton, B., 1999. *Physics for Diagnostic Radiology*. Institute of Physics Publishing, Bristol and Philadelphia, 2nd ed.

Eggert, D. W., Fitzgibbon, A. W., Fisher, R. B., 1998. Simultaneous registration of multiple range views for use in reverse engineering of CAD models. *Computer Vision and Image Understanding: CVIU* 69 (3), 253–272.
URL `citeseer.nj.nec.com/eggert96simultaneous.html`

Eikenberry, E., Tate, M., Gruner, S., 2001. Ccd detectors. In: M. Rossmann, E. Arnold (Eds.), *International Tables of Crystallography. Volume F: Macromolecular Crystallography*.

Erben, C., Vitt, K., Wulf, J., 2002. The phidias validation study of stereolithographic models (preliminary report). *Phidias Newsletter: Rapid Prototyping in Medicine* 8, 15–16.

Fenster, A., Tong, S., Sherebrin, S., Downey, D., Rankin, R., 1995. Three-dimensional ultrasound imaging. In: R. Van Metter (Ed.), *Proceedings of SPIE, Medical Imaging*, vol. 2432.

Fluoroscopy, 2001. Saint Luke's Hospital of Kansas City - Fluoroscopy User Manual. Lasted accessed: 2001/11/22.
URL `http://www.planetkc.com/scottkar/fluoro/fluoro.htm`

Gao, J., 2001. *2D and 3D Surface Image Processing Algorithms and their Applications*. Ph.D. thesis, New Jersey Institute of Technology.

Gross, M. H., 1998. Computer Graphics in Medicine: From Visualization to Surgery Simulation. *SIGGRAPH Computer Graphics* 32 (2).

Gühring, J., Brenner, C., Böhm, J., Fritsch, D., 2000. Data processing and calibration of a cross-pattern stripe projector. In: *International Society for Photogrammetry and Remote Sensing Congress 2000, IAPRS 33(5)*, Amsterdam, Netherlands.

Heckel, B., Uva, A. E., Hamann, B., 1999. *Highly Efficient Generation of Hierarchical Surface Models*. Tech. rep., Center for Image Processing and Integrated Computing (CIPIC), Department of Computer Science, University of California.

Johns, H., Cunningham, J., 1983. *The Physics of Radiogy*. Charles C Thomas, Springfield, Illinois, USA, 4th ed.

Johnson, A., Hebert, M., 1997. Surface registration by matching oriented points. Recent Advances in 3-D Digital Imaging and Modeling.

Kalawsky, R., 1996. *Exploiting Virtual Reality Techniques in Education and Training: Technological Issues.*. Tech. rep., SIMA Report Series.

Kermer, C., Lindner, A., Friede, I., Wagner, A., Millesi, W., 1998. Preoperative stereolithographic model planning for primary reconstruction in craniomaxillofacial trauma surgery. *Journal of Craniomaxillofacial Surgery* 26, 136–139.

Kneöaurek, K., Ivanovic, M., Machac, J., Weber, D. A., 2000. Medical Image Registration. *Europhysics News* 31 (4).

Koschan, A., Rodehorst, V., Spiller, K., 1996. Color stereo vision using hierarchical block matching and active color illumination. In: *Proceedings 13th International Conference on Pattern Recognition ICPR96*, pp. 835–839, pp. 835–839.

Lemchen, M., 2000. Scanning device or methodology to produce an image incorporating correlated superficial, three-dimensional surface and x-ray images and measurements of an object. *United States Patent* 6081739.

LightScribe, 2001. Immersion 3D technologies. Lasted accessed: 2001/02/10.
URL `http://www.immersion.com/3dtech.html`

Limacher, M. C., Douglas, P. S., Germano, G., Laskey, W. K., Lindsay, B. D., Mcketty, M. H., Moore, M. E., Park, J. K., Prigent, F. M., Walsh, M. N., 1998. Radiation Safety in the Practice of Cardiology. *American College of Cardiology* 31 (4), 892913.

Lyness, C., Marte, O., Wong, B., Marais, P., 1999. Low-Cost Model Reconstruction from Image Sequences. *Collaborative Visual Computing Laboratory, University of Cape Town* .

Martin, W., Aggarwal, J., 1983. Volumetric descriptions of objects from multiple views. *Transactions on Pattern Analysis and Machine Intelligence* PAMI5, 150159.

Mcgivney, J., Cooney, P., 2000. Radiation therapy and technology. *Journal of the Institution of Engineers of Ireland* .

Meadows, D., Johnson, W., Allen, T., 1970. Generation of surface contours by moiré patterns. *Applied optics* 9 (4), 942–947.

Morency, L., Rahimi, A., Darrell, T., 2002. Fast 3d model acquisition from stereo images. *3D PVT* .

Motavalli, S., 1998. Review of reverse engineering approaches. *Computers and industrial engineering* 35 (1-2), 25–28.

Naftel, A., Mao, Z., 2002. Acquiring 3-d facial models using structured-light assisted stereo correspondence. In: *International Conference on Computer Vision and Graphics*, ICCVG 2002, Zakopane, Poland.

Nebel, J.-C., Rodriguez-Miguel, F., Cockshott, P., 2001. Stroboscopic stereo rangefinder. 3D-MATIC Research Laboratory.

Niem, W., 1994. Robust and fast modelling of 3d natural objects from multiple views. In: *SPIE Proceedings: Image and Video Processing II*, vol. 2182, pp. 388–397, pp. 388–397.

Niem, W., 1999. Automatic reconstruction of 3D objects using a mobile camera. *Image and Computing* 17, 125–134.

Owens, R., Sadlier, R., 2000. Biological shape measurement. Last accessed: 2001/06/18.
URL `http://www.cs.uwa.edu.au/r̃obyn/bioshape.html`

Paakkari, J., 1998. *On-line flatness measurement of large steel plates using Moiré topography*. Ph.D. thesis, University of Oulu.

Pekelsky, J., van Wijk, M., 1989. *Non-Topographic Photogrammetry 2nd ed.*, VA: American Society for Photogrammetry and Remote Sensing, Falls Church, chap. Moire topography: Systems and applications.

Perzl, M. A., 1997. *Development of 3D models for the simulation of fluid dynamics and particle transport in realistic airway geometries*. Ph.D. thesis, Munich University of Technology, Munich University of Technology.

Radiation, L., 2001. Diagnostic X-ray Safety Manual - Fundamentals of Health Physics Principles Last accessed: 2001/10/27.
URL `http://ehs.ucdavis.edu/hp/axs/fhpp.html`

Radiation-1, 1999. *X-ray Equipment in Medical Diagnosis Part A: Recommended Safety Procedures for Installation and Use*. Canada Communication Group.

Rausch, R., Tavlet, M., 1999. On-line radiation test facility for industrial equipment needed for the large hadron collider at cern. In: *International Conference on Accelerator and Large Experimental Physics Control Systems*, Trieste, Italy.

Revet, B., 1997. *DICOM Cook Book for Implementations in Modalities*. Tech. rep., Philips Medical Systems Nederland B.V.

Siebert, J., Marshall, S., 2000. Human body 3D imaging by speckle texture projection photogrammetry. *Sensor Review* 20 (3), 218–226.

Sonka, M., Hlavac, V., Boyle, R., 1998. *Image Processing, Analysis, and Machine Vision*, PWS - an Imprint of Brooks and Cole Publishing, chap. 9: 3D Vision (Part I). 2nd ed.

Sun, W., Bradley, C., Zhang, Y., Loh, H., 2001. Cloud data modelling employing a unied, non-redundant triangular mesh. *Computer-Aided Design* 33, 183–193.

Takasaki, H., 1970. Moiré topography. *Applied optics* 9 (6), 1467–1472.

Takeda, M., 1982. Fringe formula for projection type moiré topography. *Optics and Lasers in Engineering* 3 (1), 45–52.

TGS Inc., 2002. TGS Inc. homepage. Last accessed: 2002/03/12.
URL `http://www.tgs.com/`

TriForm, 2001. Non-contact 3D image capture system from Wicks and Wilson. Last accessed: 2002/02/15.
URL `http://www.wwl.co.uk/triform.htm`

Triggs, B., Mohr, R., 1996. Projective geometry for image analysis. A Tutorial given at ISPRS.
URL `http://www.dai.ed.ac.uk/CVonline/LOCAL_COPIES/MOHR_TRIGGS/isprs96.html`

Tristano, J. R., Owen, S. J., Canann, S. A., 1998. Advancing Front Surface Mesh Generation in Parametric Space using a Riemannian Surface Definition. *Proceedings 7th International Meshing Roundtable* .

Tsai, R., 1987. A Versatile Camera Calibration Technique for High-accuracy 3D Machine Vision Metrology using off-the-shelf TV Cameras and Lenses. *IEEE Journal of Robotics and Automation* RA-3 (4), 323–344.

Turk, G., Levoy, M., 1994. Zippered polygonal meshes from range images. In: *SIGGRAPH Conference Proceedings*, vol. 2, pp. 311–318, pp. 311–318.

Turner-Smith, A., Harris, J., Houghton, G. R., Jefferson, R. J., 1988. A method for analysis of back shape in scoliosis. *Journal of Biomechanics* 21 (6), 497–509.

Velkley, D., Oliver, G., 1979. Stereo-photgrammetry for the determination of patient surface geometry. *Medical Physics* 6 (2), 100–104.

VRML-1, 1997. The Virtual Reality Modeling Language. *International Standard ISO/IEC* 14772 (1).

Wilks, R., 1993. An optical system for measuring surface shapes for radiotherapy planning. *The British Journal of Radiology* 66, 351–359.

Zhang, L., Curless, B., Seitz, S. M., 2002. Rapid shape acquisition using color structured light and multi-pass dynamic programming. 1st international symposium on 3D data processing, visualization, and transmission, Padova, Italy.