Research Survey:

Laser scanners for three-dimensional image digitisation

Josep Forest-Collado

Director: Dr. Joaquim Salvi Mas

Programa de Doctorat en Enginyeria Informàtica Industrial /

Tecnologies Avançades de Control del

Departament d'Electrònica, Informàtica i Automàtica

Universitat de Girona

Contents

List of Figures

List of Tables

To Maite, with love...

Chapter 1

Introduction

A three-dimensional digitiser is often called *range finder* too, stating its capability for acquiring **range** or *depth* information. These devices use to grab "range images" or "$2\frac{1}{2}$D images", which are dense arrays of values related to the distance of the scene to a known point or plane.

Many laser scanning three-dimensional digitiser systems use one or more standard CCD cameras and scanning laser slits, and triangulation techniques are mostly used in the 3D reconstruction of the scenes. Some of these systems operate grabbing one image per each sweep step, so it is a time consuming process, yielding a big amount of images. However, using high speed vision modules, rates of up to 1.6 complete scans per second have been achieved, as is the case of the *MINOLTA vivid 700*.

Speeds of up to 3 $2\frac{1}{2}$D images per second have been achieved using standard CCD cameras, though. The operation of such systems (i.e. CubicScope) rely on the fast switching of the light source and the CCD shutter operation in order to get space-encoded images.

A lot of manufacturers of scanning three-dimensional digitisers exist: *Intecu*, *Cubiscope*, *3DScanners*, *Philips Industrial Vision*, *DigiBotics*, *Steinbichler Optotechnik*, and more. A complete list of manufacturers with links to their own web pages can be found in *http://www.3dlinks.com/hardware_scanners.cfm*.

In real time applications range information is required continuously, without gaps in

the data. This limits dramatically the use of commercial three-dimensional digitisers in such fields.

Some research has been done by [13, 41] toward the development of *Smart sensors* capable of delivering up to 1000 $2\frac{1}{2}$D images per second, but only laboratory prototypes were developed.

This paper is intended to give an overview of laser slit scanning three dimensional digitiser systems paying a special attention to their calibration methods, reporting a number of references for the interested reader. In addition the different surveyed authors and their corresponding systems or methods are classified in table 2.1 in order to provide the audience with a reference tool. The next section discusses the use of triangulation-based reconstruction briefly, specially focusing on laser scanning systems. In section 2.3 an illustrative classification table is proposed. Section 2.2 expands the proposed classification showing a more in-depth description of some of the systems. Conclusions are presented in section 4.

Chapter 2

Laser scanners: a state-of-the-art

2.1 The principle of triangulation

This section introduces the principle of triangulation focusing its use in laser slit scanning. The reader is pointed to [17] for a very thorough survey on triangulation techniques.

Figure 2.1 shows how a point P in 3D space is *seen* by two different cameras. If the spatial relationship between both cameras is accurately known, the position of P can be found by its two projections (p' and p'') on the corresponding image planes, computing the intersection of lines $P - p'$ and $P - p''$.

The spatial relationship between both cameras can be easily found by applying a standard camera calibration procedure to be chosen among the multiple existing ones, depending on how distorted are the captured images. Different camera calibration methods with and without distortion consideration can be found in [15, 12, 39] and[40]. In addition a survey has been published recently by [29] comparing the performance of different methods in terms of their accuracy.

But, unfortunately the biggest problem here is: how can P be uniquely identified in both image planes ? i.e. how sure can be stated that both p' and p'' correspond to the projection of P ? The answer to these questions is not unique and a lot of research is being done in this direction. This problem is widely known as *the correspondence*

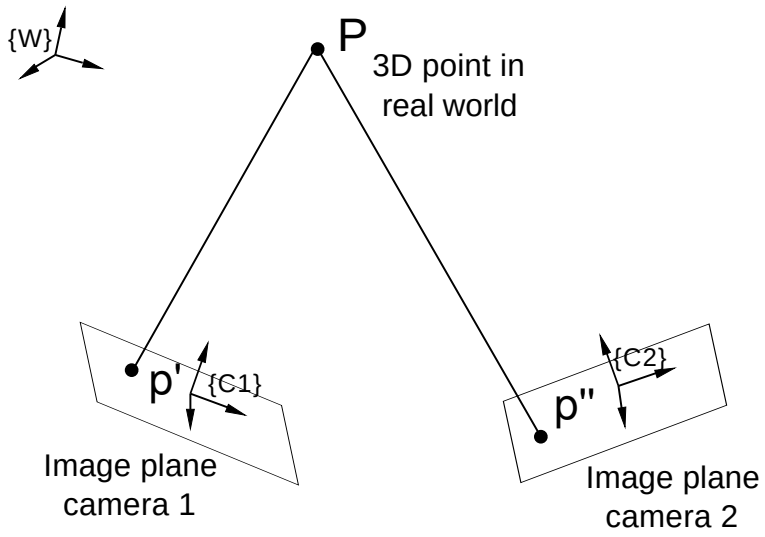


Figure 2.1: Triangulation principle using 2 cameras.

problem.

One of the techniques that works very well for identifying points in 3D space is the use of *Active Vision*. The term *active* means that some well known light pattern (typically laser or coded/uncoded structured light) in different geometries (points, planes, crosses, parallel bars, etc.) is projected onto the scene, so a laser illuminated point in the 3D space is well *seen* by both cameras, and hence, the 3D coordinates of P are easily computed by triangulation. An excellent survey on active range finders can be found in [5]. Moreover, a thorough survey comparing the use of different coded structured light schemes toward the solution to the correspondence problem has been published by [4].

In addition, if the light source geometry, i.e. its pose and sweep movement, is well known, one of the cameras can be replaced by the light source and 3D information can be obtained by triangulation as well. This procedure is widely used by commercial and research prototype range finders.

Figure 2.2 shows how camera 2 in figure 2.1 has been replaced by a sweeping laser light source. Using this technique the camera model has to be computed using a calibration procedure, and the line or plane laser equation must be known in order to compute its

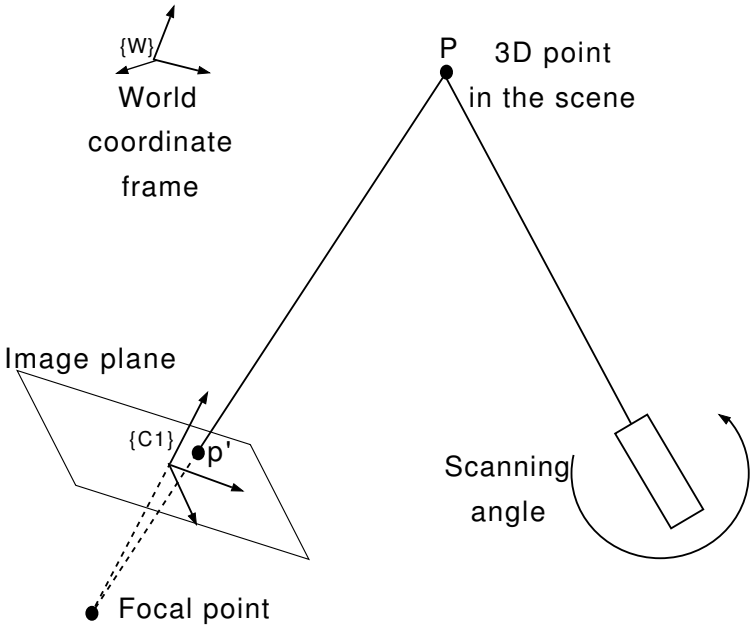


Figure 2.2: Laser scanner system with scanning light source.

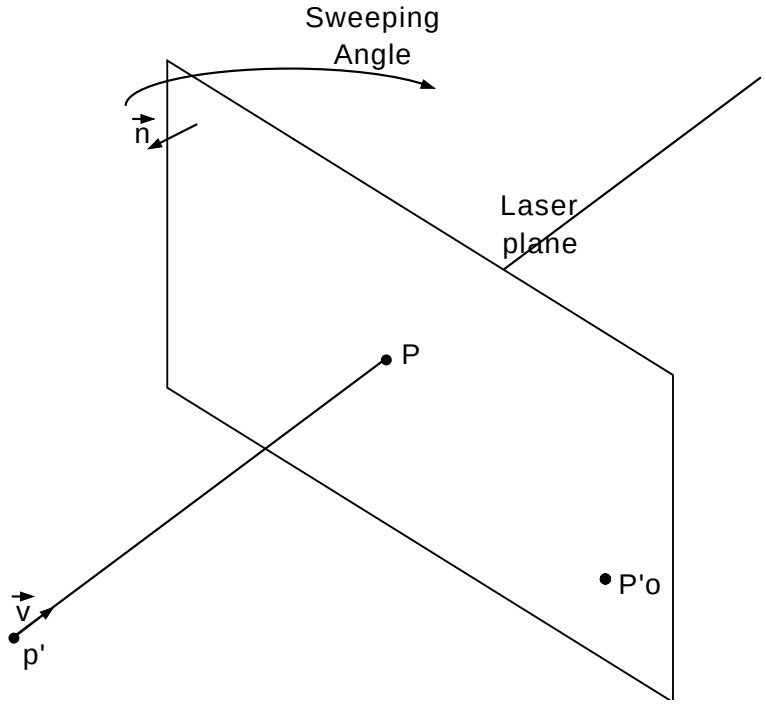


Figure 2.3: 3D reconstruction by triangulation using a laser plane and a camera.

intersection with the line $P - p'$.

In order to know the laser equation, the sweeping angle must be measured by means of some external mechanical or optical rotation sensor (encoders or potentiometers).

Equations 2.1 show how a 3D point can be computed, knowing one point of the sweeping laser plane (P'_0), the $P-p'$ line equation (with direction vector $\overrightarrow{v}$) and the plane's normal vector $\overrightarrow{n}$ (which depends exclusively on the sweeping angle). Figure 2.3 shows the laser plane and camera arrangement where these equations can be applied.

$$\begin{cases} P = P_0 + \lambda \overrightarrow{v} \\ (P - P'_0) \cdot \overrightarrow{n} = 0 \end{cases} \tag{2.1}$$

2.2 Systems and methods for shape acquisition

Shape measurement is intended to be as much accurate as possible. The calibration procedure together with the stripe detection method play an important role in this sense. At the end, accuracy is dependent on both the accurate measurement of the relative pose between the imager and the light projector, and in how certain is the stripe detection in the imager itself. The imager and slit projector pose may be measured using very accurate mechanical means, although it is not applicable to the systems which are subject to a modification of their *default* mechanical arrangement.
In addition, sub-pixel accuracy stripe detection methods may be applied in order to minimise the uncertainty. A comparison between five different such methods can be found in [37].
This section discusses the methods surveyed in section 2.3, addressing the general description of each system according to the dissertation of the authors. It follows the same structure as table 2.1. A special attention is made on calibration and reconstruction issues.

2.2.1 Raw acquisition

In this work, the term *raw acquisition* is used as a reference including those methods in which no encoding techniques have been applied to the process of acquiring a scene range map. That is, one image is captured by a frame grabber whenever a slit is projected onto the scene for a new position increment of the scanning mechanics. Hence, assuming the laser slit is vertical with respect to the image plane, $\boldsymbol{n}$ images are needed in order to get an $\boldsymbol{n}$ *column* range map.
Shirai & Suwa [36] and Agin & Binford [1] used a video camera together with a plane-shaped laser light source projecting a slit onto a scene. The slit scanned the scene using a mirror attached to a stepper motor, and the camera was used for the acquisition of the stripe image.
Chen & Kak [9] developed a similar three-dimensional digitiser, but the camera-projector set was arranged so that it formed a rigid structure, keeping their relative orientations invariant to each other. The whole system was mounted on the end effector of a robot

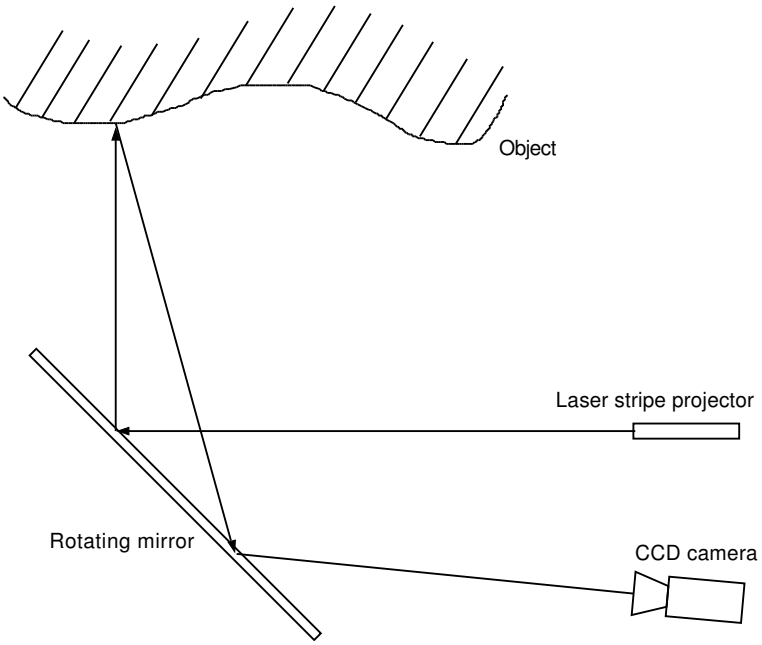


Figure 2.4: Oxford/NEL range-finder scheme.

arm in order to scan the scene under consideration.

Reid [28] used the same idea providing the Oxford AGV with object recognition and acquisition capabilities (see figure 2.4). In this system, instead of making the device scan the scene itself, it is mounted on a static platform and a mirror is used in order to reflect both the laser stripe and the image. Experimental data is reported showing a 1% of error relative to the measured depths.

Other approaches have been reported using different scanning techniques: Sato et al. [34] described a system consisting on two laser slit projectors and a CCD camera, but the object was placed on a turntable, being the laser-camera set static. The two laser sources were used in order to minimise the occlusions. Champleboux et al. [7] and [8] described a range imaging sensor where the laser was linearly displaced using a linear actuator. It was used together with an object localisation algorithm for the matching of human face surfaces.

Calibration

Shirai & Suwa [36] calibrated their system by measuring the relative pose between the camera and the projector. The use of this method is strongly discouraged if an accurate mechanical positioning system is not available. Agin & Binford [1] used a calibration procedure consisting on a former camera calibration, followed by the estimation of the light projector position by grabbing a series of stripe images at known angular positions of the projector. A relaxation algorithm was used in order to minimise the position error. The accuracy in scene reconstruction is mostly dependent on the right choice of the camera calibration method.

Chen & Kak proposed a new calibration method based on the application of projective geometry (see figure 2.6). They stated that the points on the light plane could be mapped onto the image plane by finding the planar transformation (homography) which relates both planes, as shown in figure 2.5. The reader is pointed to [12] or [16] to get deeper into projective geometry. The planes **s** and **r** correspond to the laser and image planes respectively. The calibration procedure is carried out with respect to a reference coordinate frame. A different coordinate frame F_s is associated to the laser plane which describes its pose with respect to the reference frame. A bi-dimensional coordinate frame F_{2s} is associated with the laser plane, where its axis x and y coincide with the axis x and y of F_s. This is an artifact used by the authors of [9] in order to easily change from the bi-dimensional coordinates of points on **s** to their corresponding three-dimensional representation.

$$\rho \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \cdot \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} \tag{2.2}$$

Using the property of the cross-ratio invariance in projective space, one can define a 3x3 planar transformation mapping the points laying on a plane **s** into another plane **r**. Then, equation 2.2 is obtained, where ρ is a scale factor. The left hand side of this equation is the vector coordinate of a point laying on **s** (X_s), while the right hand side column vector describes the coordinates of a point laying on **r** (X_r). Considering the

transformation F_s, the transformation from F_{2s} to F_s and equation 2.2, equation 2.3 is obtained (where $x'_1 = u$, $x'_2 = v$ and $x'_3 = 1$), which maps a point on the image plane into its corresponding three dimensional point. The matrix T_{cb} in equation 2.3 is called the *conversion matrix*.

From equation 2.3, the expressions for x, y and z in terms of the t_{ij} components of the conversion matrix can be obtained. Hence, a set of three linear equations is set for every image point. Since ρ is a scale factor, the 4x3 matrix T_{cb} can be simplified making $t_{43} = 1$ without loss of generality. Then four points are enough in order to solve for the 11 unknowns (t_{11} to t_{42}). A system of 12 equations with 11 unknowns is raised, but since the points must be chosen such that they are coplanar, they must satisfy equation 2.4 too, which is *de facto* implicit in equation 2.3, meaning that the 12th equation is determined by the other 11. In practise, a set of more than 4 points is used in order to set an overdetermined system of equations obtaining a solution that best fits some minimisation criterion like a least square optimisation. Experimental results made on a simple cubic object show that the relative accuracy of the system is dependent on the distance from the object to be scanned, obtaining errors of <0.05 inch at a distance of 20 inches and <0.04 inch at 14 inches, for the measure of the width, which is the side *perpendicular* to the optical axis of the camera. In addition, when the accuracy is evaluated measuring the height, which is the side *parallel* to the optical axis of the camera, the errors are one order of magnitude bigger, obtaining <0.30 inch at 20 inch and <0.14 inch at 14 inch. Again the accurate stripe detection is of great importance for an improved measurement accuracy.

A similar calibration procedure was used by Reid [28].

Sato et al. [34] used a light target together with a known object in order to calibrate their system. A point of light was attached to the turntable at known radius and heights and the turntable was operated, grabbing several images of it. The accurate measurement of the laser projector location was finally used in order to obtain three-dimensional data by triangulation. The reconstruction accuracy is dependent on how accurate is the location of the maximum light power of the target.

Champleboux et al. used an accurate positioning mechanical system in order to identify

the light plane equation. Then, the system calibration was done by identifying the transformation $G : R^3 \rightarrow R^3$ defined by the relation (X,Y,Z)=G(u,v,d), where (u,v) are the pixel coordinates of point (X,Y,Z) and $\boldsymbol{d}$ is the linear displacement of the laser slit.

$$\rho \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = F_s \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \\ t_{41} & t_{42} & t_{43} \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = T_{cb} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{2.3}$$

$$\det[X_b^1 X_b^2 X_b^3 X_b^4] = 0 \tag{2.4}$$

2.2.2 Space encoding. Switching the laser slit

Space-encoding stands for the projection of successive binary patterns onto the scene to be acquired (see figure 2.7).

The number of bits used in the codification is directly related to the number of patterns projected. Usually, the bright regions are assigned the logic level '1', and the dark regions are assigned a '0'. Finally, if each pattern is assigned a different weight, every pixel is associated to a unique binary code, which can either be natural, grey or any other suitable one. Space-encoding techniques are being applied using pattern projection ([27, 22, 21, 30]). In these cases, the projector is modelled as a *reversed camera*.

Sato et al. applied this technique in the *CubicScope* three-dimensional digitiser [35], using a scanning laser slit mechanism. The technique takes advantage of the shutter period of a standard CCD camera: within the time period in which the shutter is open, the laser slit must scan the scene in such a way that illuminated (bright) and shadowed (dark) regions appear. This technique requires the laser to be rapidly switched on

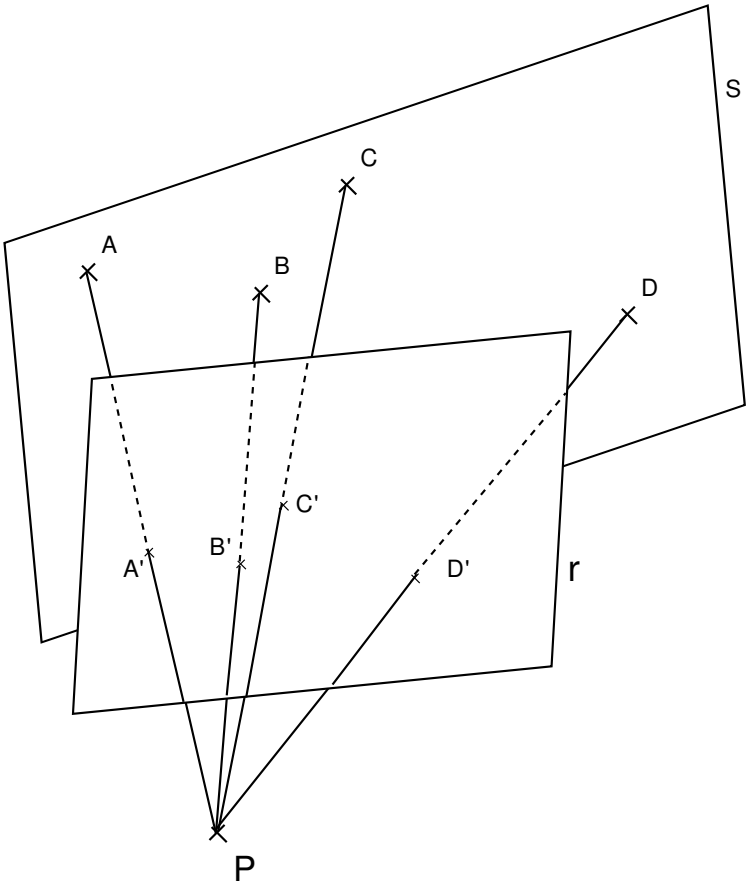


Figure 2.5: Elements of two dimensional projectivity.

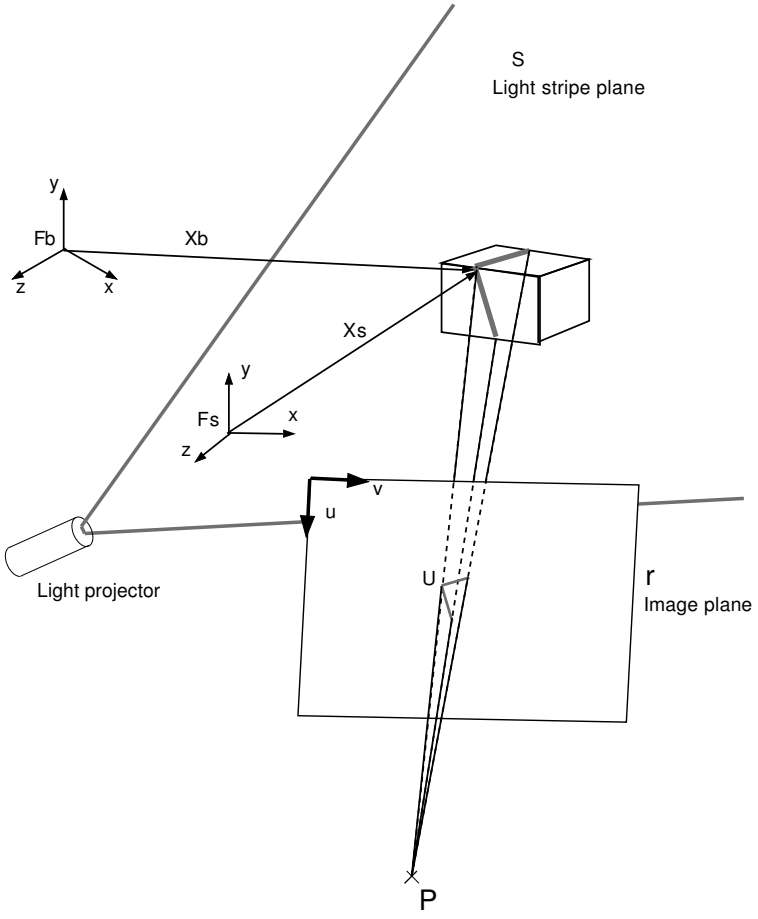


Figure 2.6: The system arrangement of Chen & Kak.

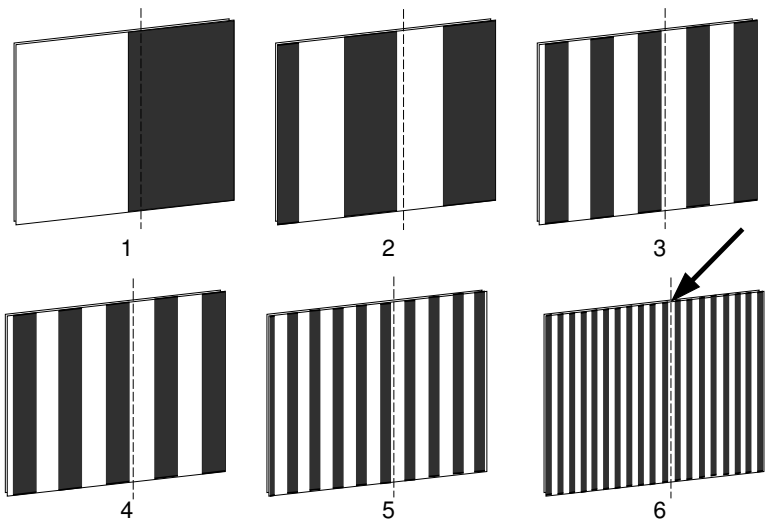


Figure 2.7: Space-encoding method.

and off. Due to the light integration given on the CCD, it stores a whole image as though a slide pattern was projected onto the scene. Hence, for each shutter period, a different pattern must be generated using the switching technique in order to get the proper patterned image. Generally, N encoded images produce the same range map than $2^N - 1$ non-encoded images, taking one image shot per each laser scan step.

As an example of performance, consider the operation of a standard NTSC video camera which operates at 30 frames per second. Then, in order to achieve a rate of 1 $2\frac{1}{2}$D image in 0.3 seconds, a series of 9 patterns must be projected. This implies that a $2^9 - 1 = 511$ *columns* range map is obtained with 1% of accuracy, according to [35].

Yu et al. [43] used the space-encoding method in their laser scanning device, whose mechanical arrangement is shown in figure 2.8. Expressions for X,Y and Z are obtained using trigonometrical analysis, as equations 2.5 show, where $2\theta_1$ and $2\beta_1$ are the vertical and horizontal view angles of the camera, respectively. The number of horizontal and vertical pixels are 2N and 2M respectively, while $\boldsymbol{n}$ and $\boldsymbol{m}$ are the corresponding pixel indexes.

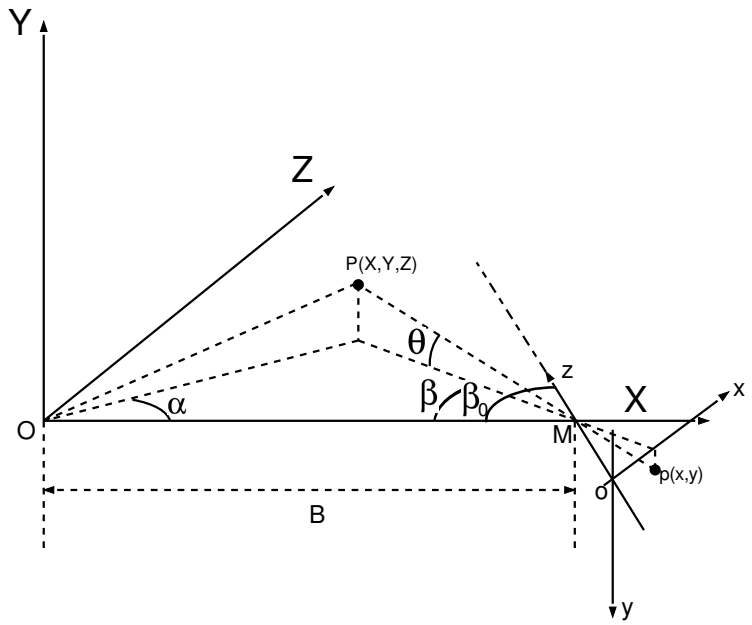


Figure 2.8: Yu et al. digitiser geometric parameters.

$$
\begin{aligned}
Z &= \frac{B}{\cot\alpha + \frac{\cot\beta_0 + n\cdot(\tan\beta_1/N)}{1 - n\cdot\cot\beta_0\cdot(\tan(\beta_1/N)}} \\
Y &= \frac{m\cdot Z\cdot(\tan\theta_1/M)}{\sin\beta_0 - n\cdot\cos\beta_0\cdot(\tan\beta_1/N)} \\
X &= Z\cdot\cot\alpha
\end{aligned}
\tag{2.5}
$$

The three-dimensional errors introduced by the error in the pixel detection (Δx and Δy) as well as the angle error ($\Delta\alpha$) are studied obtaining the expressions for ΔX, ΔY and ΔZ, by partially deriving X, Y and Z with respect to $\boldsymbol{x}$, $\boldsymbol{y}$ and α. A study about the error introduced by the geometry of the polygonal mirror was reported (see table 2.1). In such arrangement, the rotating mirror has several different reflecting sides in order to achieve the optimal number of scans per motor turn. This makes the reflected slit to *effectively be projected* from somewhere away from the rotational axis, as shown in figure 2.9, and this introduces an error in the calculation of the baseline B (ΔB_0). Equation 2.6 shows an expression for ΔB_0, deduced by [43]. A very thorough work on the estimation of calibration errors in laser scanning devices using space-encoding can be found in [42]. This work was done from the results published in [43].

$$
\Delta B_0 = 2R\cdot(\cos a - \cos a_0)\cdot\frac{\sin(b-a)}{\cos(b-2a)} \tag{2.6}
$$

Calibration

Cubicscope [35] is calibrated in two steps: first, the CCD camera is calibrated using Tsai's method [38]. Second, regarding the scanning nature of the laser slit projection, each of the coded regions is assigned to a plane equation, hence, the second step of the calibration process is in charge of the accurate identification of every equation. In this sense, a mechanical arrangement consisting on an accurately positioned plane at different ranges from the digitiser is set up, then, on each of the planes a whole set of patterns is projected and the images are stored. Every laser slit plane is identified by the equally coded regions imaged on each of the calibration planes, hence one laser slit plane equation is yield for every coded region. In [19], a new approach using pattern shifting is tested with the aim of improving the accuracy at the expense of acquisition speed. The pattern shifting method relies on the accurate detection of the regions' edges

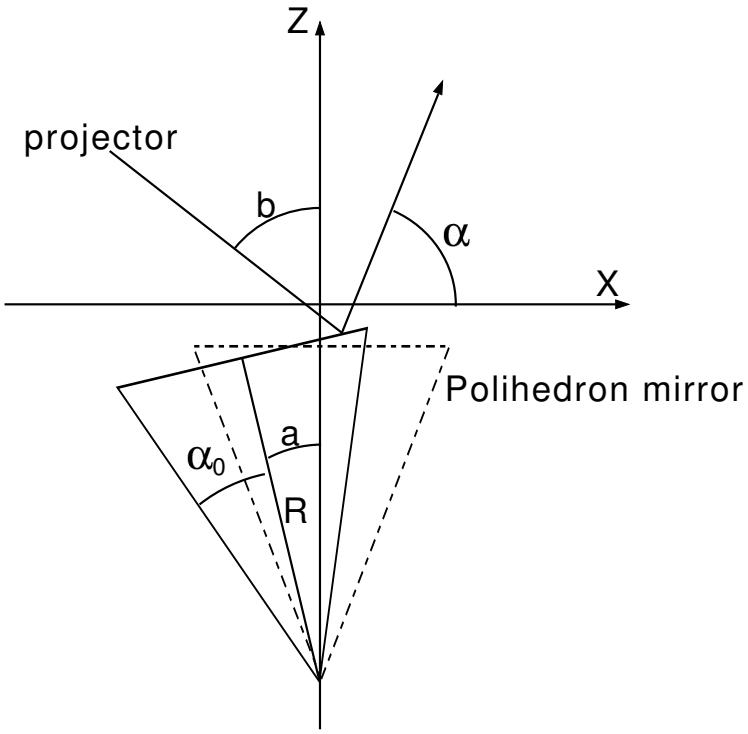


Figure 2.9: Change on the slit ray projection axis.

using the CCD charge storing effect.
In [42], Yu's digitiser is calibrated using a different method. The mechanical arrangement consists of a standardised plane, placed in front of the lenses, within the view of the system, such that it is parallel to the xoy plane and perpendicular to the xoz. Using the switching technique for pattern generation, two concentric, different size squares are projected onto the calibration plane. Then, the plane is moved along the z axis until the biggest square fills the whole image plane, the calibration plane position is recorded, and the process is repeated for the smallest square. Using the known geometrical relation between the two squares, the different parameters are obtained. The reconstruction is done using triangulation and an accuracy of 1.5% is achieved.
In space-encoding, the *width* of each coded region determines both the resolution and the reconstruction accuracy. Hence, the number of bits involved in the pattern generation must be carefully chosen according to the application specs. In addition, a unique criterion must be systematically applied in order to determine which of the several subsets of points laying on a coded region is assigned to each plane equation with the aim of minimising the stripe detection uncertainty as well as the reconstruction inaccuracies.

2.2.3 Smart sensors. On-chip range computation

Under the assumption of constant rotation speed, it can be stated that the angular position is proportional to the rotation time ($\theta = \omega \cdot t$), where ω, the rotation speed, is the constant of proportionality. If the rotation speed of a scanning mirror can be accurately controlled so it can be considered constant, and if the rotation time can be accurately measured, the angular position can be obtained algebraically eliminating the need for an angular position sensor. In addition, since no position control is required, only a simple speed controller is needed in order to comply with the assumption of constant rotation speed.
Popplestone [26], used a range finder that relied on the operation of a *vidicon* TV camera, measuring the time that a TV line scan takes until it achieves the image of the laser slit. The apparatus returned a number of clock pulses per each TV line, using a fast clock signal.

Araki et al. [2] proposed a new method for measuring the *time* at which the slit image projects onto a photosensitive cell of a specifically designed discrete array. Each of the photosensitive cells has a timer register associated which is incremented until the image of the laser slit illuminates its associated photosensitive cell. This kind of operating procedure allows a very fast scanning, since each of the photosensitive cells is concurrently *processing* by its own. Hence, this operating procedure is called ***cell-parallel*** processing. Two optical switches are used in order to indicate the start and the end of a scan period.

A simple threshold circuit was used in each cell in order to detect the light stripe image. The output of this circuit was used to stop the timer counting and store its value on the register itself. Once the light stripe mechanism fulfils a complete scan period, the register array is configured such that a big shift register is formed. Then the time values are transferred serially to the host computer.

A rotating mirror arrangement similar to that of [19] or [43] is used in this implementation.

Gruss and Kanade [13, 14, 23] proposed a VLSI version of a similar system. The conceptual difference with the system by [2] is that the time computation is made analogically, using capacitor-based timing circuits, and a periodic sawtooth signal. The stripe detection was accomplished by using simple threshold circuits integrated in each cell.

Analog time stamp recording is used instead of digital timers in order to avoid EMI from the switching clock signals to affect the photosensitive cells output.

Kanade et al. [23] and [14] described a prototype implementation of the system in [13].

Yokohama et al. [41] and [32] proposed a new prototype similar to that of [13], incorporating notable improvements in slit detection. They used two side by side photosensitive areas in every cell, sensing the *difference* in light intensity between the two twin photo-diodes in every cell. The cell architecture uses three clock signals in order to set a 4-phase operation for synchronously sensing and transferring the data. A notable improvement in the accuracy has been achieved, obtaining a range error of -24μm to +14μm, and is, in addition, very robust under the presence of significant levels of ambient light. Hori et al. [20] and Baba et al. [3] proposed similar systems based on

the *cell-parallel* processing concept, introduced by [2].
Recently, Brajovic and Kanade [6] have proposed a new approach to VLSI acquisition of $2\frac{1}{2}$D images, introducing the concept of *row-parallel* processing. This new idea takes advantage of the fact that laser slit scanning is inherently row-parallel, since only one cell on each row is illuminated at a particular instant of time.

Calibration

In [14] the calibration target is a planar surface out of which a triangular section has been removed. This target is mounted on an accurate 3DOF[1] positioning device so its pose is accurately known. The calibration procedure is as follows: 1) the line-of-sight rays for a few cells are measured, and 2) a pinhole-camera model is fit to the measured line-of-sight straights in order to approximate the line-of-sight equation for all sensing elements.
The target is positioned so that its surface is parallel to the reference xy plane. Then it is moved along the search path (see figure 2.11) until an occlusion (the photosensitive cell does not detect the slit) occurs. Finally, the 3D position of the target is recorded, and the process is repeated until both the bottom and top edges of the triangle are found. This procedure is made for different depths, so a set of *triangles* is recorded. The right corner of the triangle is found by calculating the bottom and top lines intersection for each of the triangles at the different depths. Finally, a line-of-sight equation is found by fitting a straight on the different corners, as shown in figure 2.10.
This calibration procedure is too slow to be applied to all the sensing elements, so in practise, a number of 25 evenly spaced of them are found. Finally, a least-squares procedure is used to fit a pinhole camera model. Experimental data for this system shows that for a range value within 0.5mm to 500mm, an accuracy of 0.1% is achieved. A transfer rate of 1000 $2\frac{1}{2}$D images per second is achieved using this system. Although the improvement in the reconstruction accuracy is very significant, it could even be improved if a real *peak detector* was used instead of a simple threshold circuit. In addition, threshold time stamp recording is strongly dependent on the object reflectivity

[1] DOF:Degrees Of Freedom

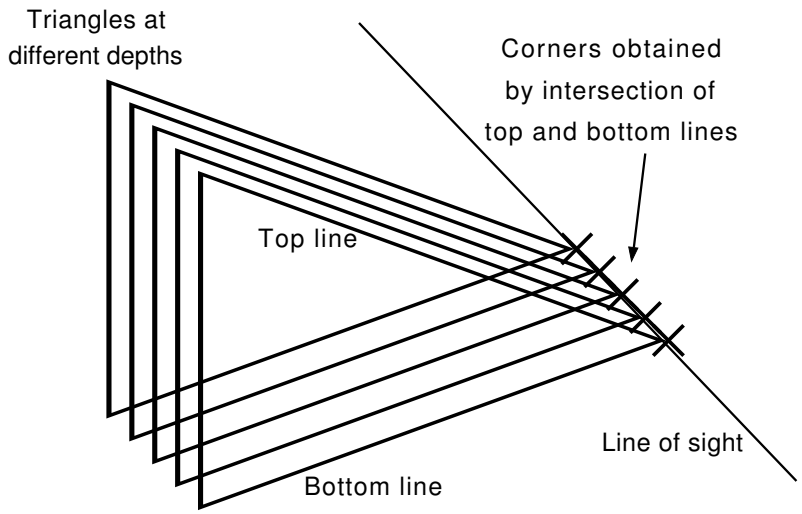


Figure 2.10: Line of sight computation in Gruss et al.

properties, hence different time stamps are obtained for equally shaped object with different reflectivities.

Yokohama et al. used a plane object for calibration. The procedure consists of 4 scans over the same plane, using 4 different depths. In each scan, the relationship between the scanning angle and the plane depth is recorded for each photo-sensitive cell. At the end of the process, a lookup table is generated in order to allow for the real-time three-dimensional reconstruction. The fact that stripe detection is made by comparison between two side-by-side photosensitive areas in each cell yields a significant improvement in the accuracy. Several considerations must be taken, since the signal-to-noise ratio of each electric signal must be maximised in order to get the highest reconstruction accuracy. To this aim, a compromise must be achieved among scanning speed, photosensitive area and optical laser power in order to successfully comply with the application specs.

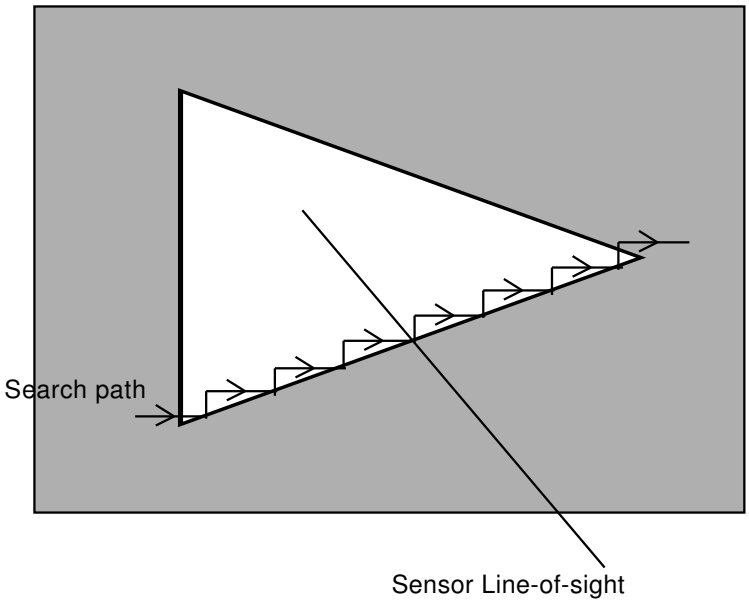


Figure 2.11: Calibration target used in Gruss et al.

2.3 Proposed classification

Table 2.1 shows the reviewed three-dimensional digitisers, reporting the main features they exhibit. The classification is structured in three main categories: *Raw acquisition*, *Space Encoding* and *Smart Sensors*. For each of them, several authors are referenced, reporting the main features of the systems they have proposed, and including an illustrative sketch. This table is intended as a fast reference for the readers.

Table 2.1: Classification of three-dimensional digitiser systems.

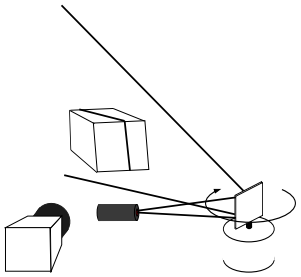
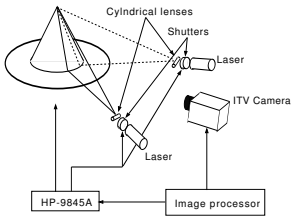
Characteristic	Features
	*Auth:*Shirai et al.[36], Agin et al.[1], Popplestone[26] *Calib:* [36, 26] Accurate pose measurement of system devices [1] (A)-Camera calibration (B)-Estimation of laser slit position by image measurements *Features:* • Single slit • Rotating mirror reflects only the laser slit
	*Auth:* Sato[34] *Calib:* Point of light attached to the turntable at known radius and heights. *Features:* • Two slit projection • **Turntable operation** • Static laser and camera set • Minimisation of occlusions
Raw Acquisition	

Table 2.1: (continued)

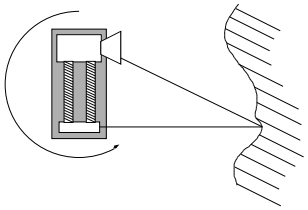
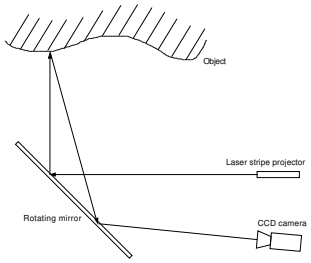
Characteristic	Features
	*Auth:*Chen & Kak [9] *Calib*:Projective geometry *Features:* • Single slit • **Fixed geometry, rotates the whole system**
	*Auth:*Reid [28] *Calib*:Projective geometry *Features:* • Single slit • Fixed geometry • **Rotating mirror reflects both the image and laser slit**
Raw Acquisition	

Table 2.1: (continued)

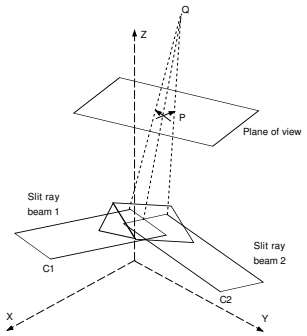
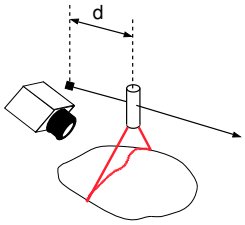
Characteristic	Features
	*Auth:* Nakano[25] *Calib:* Accurate pose measurement of system devices *Features:* • Two slit projection • Laser angular scanning • **Eliminates reconstruction errors with lambertian surfaces**
	*Auth:*Champleboux[7][8] *Calib:* (A) - Camera calibration using NPBS method. (B) - Accurate mechanical measurement and $R^3 \rightarrow R^3$transformation identification *Features:* • Single slit • **Linear scanning**
Raw Acquisition	

Table 2.1: (continued)

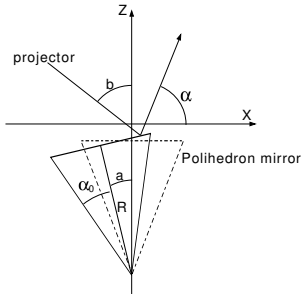
Characteristic	Features
	*Auth:*Yu[43][42] *Calib:* Square pattern projection. System layout estimation using geometry analysis *Features:* • Single slit • Rotating mirror • **Baseline error compensation** • No special code spedified, although the system is versatile enough for adopting any codification • Compromise between acquisition speed and accuracy
Space Encoding	

Table 2.1: (continued)

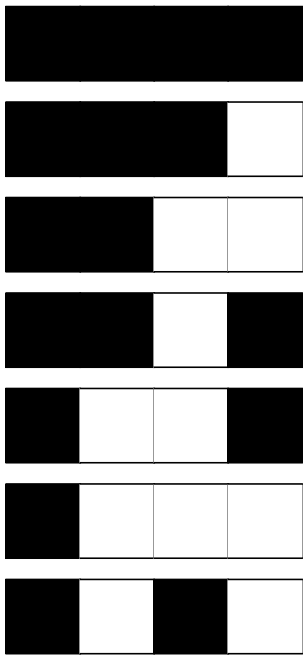
Characteristic	Features
	*Auth:*Sato[18][33][19] *Calib:* (A) - Camera calibration (Tsai [38]). (B) - Plane equation identification according to the space coding. *Features:* • Single slit • Rotating mirror • Compromise between acquisition speed and accuracy • **Use of gray code in the pattern projection** • Pattern shifting for accuracy improvement
Space Encoding	

Table 2.1: (continued)

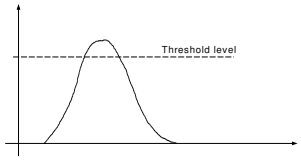
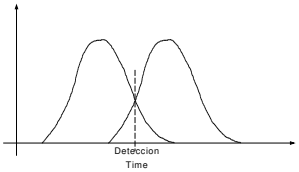
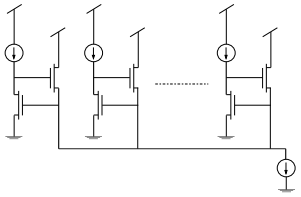
Characteristic	Features
	*Auth:* Gruss et al.[13][14], Kanade et al.[23], Hori et al.[20], Baba et al.[3] *Calib:* [13][14][23] Line-of-sight identification using a specially designed calibration target. The target is accurately positioned using a mechanical system *Features:* • Single slit • Rotating mirror • Cell-parallel architecture • Accuracy is determined by slit detection method • **Slit detection by threshold of the signal level**
Smart Sensors	

Table 2.1: (continued)

Characteristic	Features
	*Auth:* K.Sato et al.[31], Yokohama et al.[41] *Calib:* Lookup table generation by finding the relationship between the scanning *time* and known depths. A plane is used as a calibration target *Features:* • Single slit • Rotating mirror • Cell-parallel architecture • **Slit detection using comparison between two photosensitive areas in each cell**
	*Auth:* Brajovic et al. [6] *Calib:* Acurate pose measurement. *Features:* • Single slit • Rotating mirror • **One-per-row slit peak detectors (winner-take-all)** [24] • Row-parallel architecture
Smart Sensors	

Chapter 3

Performance evaluation

In order to get deeper in how a three-dimensional digitiser works, some of the methods discussed above have been simulated using *Octave*. The calibration procedures have been followed as faithfully as the level of abstraction allowed, and the performance has been evaluated making the three-dimensional reconstruction of a well known plane at four different ranges from the system (typically taking the image plane of the image sensor as a reference), and computing the maximum absolute reconstruction errors in the X,Y and Z directions.

3.1 The Chen & Kak system ([9])

Figure 3.1 shows the co-ordinate system arrangement which has been considered during the simulation. The relative pose between the camera and the laser emitter has been established such that the planes $X_L Y_L$ and $X_C Y_C$ are coincident, just for simplicity in the generation of the synthetic data. This does not affect in principle the performance of the system, nor the calibration procedure, since this method is transparent to explicit geometric considerations between those devices.
B is the separation or *base-line* between the centres of both the camera $\{C\}$ and the laser $\{L\}$ co-ordinate frames in the X direction. Note that the planes $X_L Z_L$ and $X_C Z_C$ are not coincident. **OLz** is the distance between the centres of the world $\{W\}$ and the laser

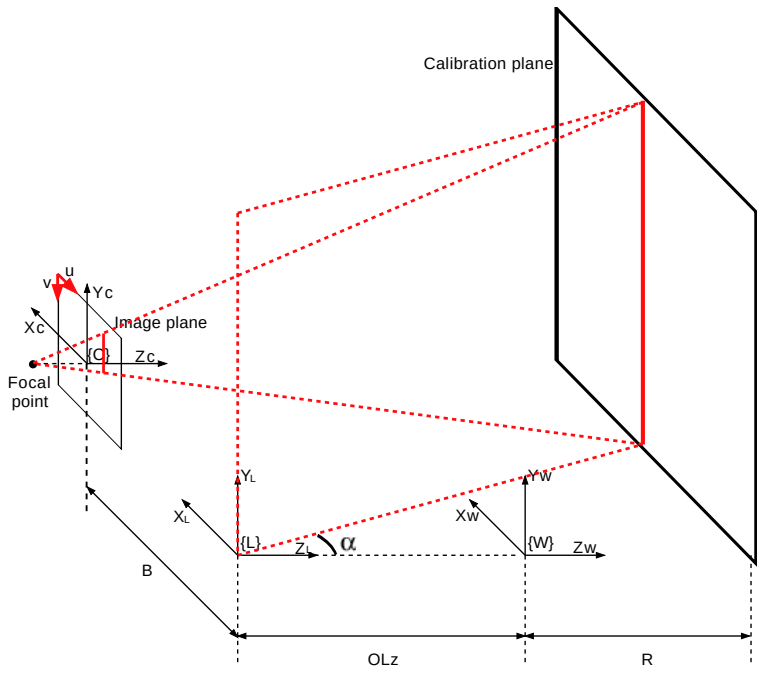


Figure 3.1: Co-ordinate frames for Chen & Kak system simulation.

$\{L\}$ co-ordinate frames. $\mathbf{R}$ is the range at which the calibration plane is located with respect to the world co-ordinate frame. $\mathbf{R}$ is known during the simulated calibration procedure, and takes the values from 0 to 110 mm in increments of 10 mm. The angle of projection (α) between the laser plane and Z_L is constant and set to 20 degrees.

According to Hartley & Zisserman ([16]), the performance evaluation of any three-dimensional system based on one or more cameras must take into consideration the existence of noise in the imaged data. This noise has been found to fit a gaussian probability distribution.

In addition, Trucco et.al. ([37]) state that subpixel accuracy for slit detection is strongly recommended for high accuracy three-dimensional reconstructions, hence it has been taken into consideration for simulation purposes. Table 3.1 summarises the results obtained after the simulation of a calibration and reconstruction of a known plane. The

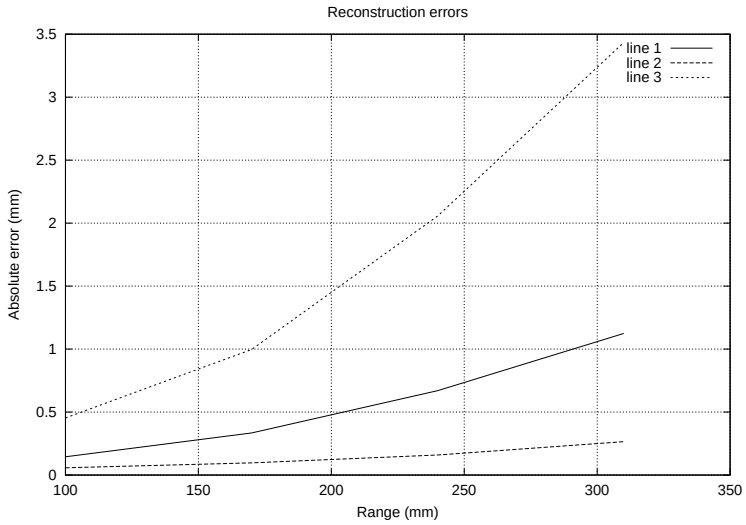


Figure 3.2: Reconstruction errors as a function of range.

four right-hand-side columns of this table correspond to the ranges at which the accuracy has been evaluated. These four distances at which the plane has been reconstructed are referred to the $\{C\}$ and $\{L\}$ co-ordinate frames, instead of $\{W\}$, in order to see more clearly what is the real distance of the object to the scanner. The added gaussian noise has been chosen such that the maximum error in slit detection is of about 7 pixels, as shown in row number 2. Figure 3.2 shows these results graphically. The upper curve (the largest error) corresponds to the error in range or Z direction, as it was expected, while the smallest error is found in the Y direction.

Table 3.1: Performance of the Chen & Kak scanner.

Range (mm)	100	170	240	310
Image error (pix.)	$6,8916$	$7,0088$	$7,0505$	$7,0395$
X error (mm)	$0,1447$	$0,3342$	$0,6703$	$1,1239$
Y error (mm)	$0,0573$	$0,0963$	$0,1594$	$0,2651$
Z error (mm)	$0,4535$	$0,9967$	$2,0569$	$3,4345$

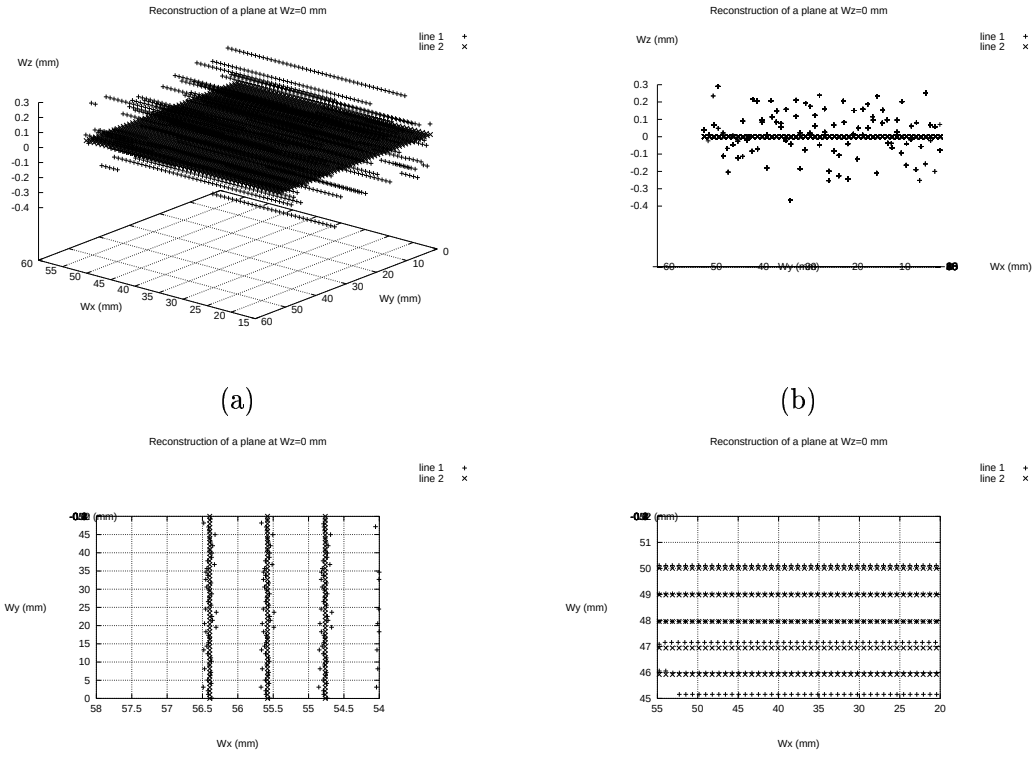


Figure 3.3: Reconstruction of a plane at Wz=0.

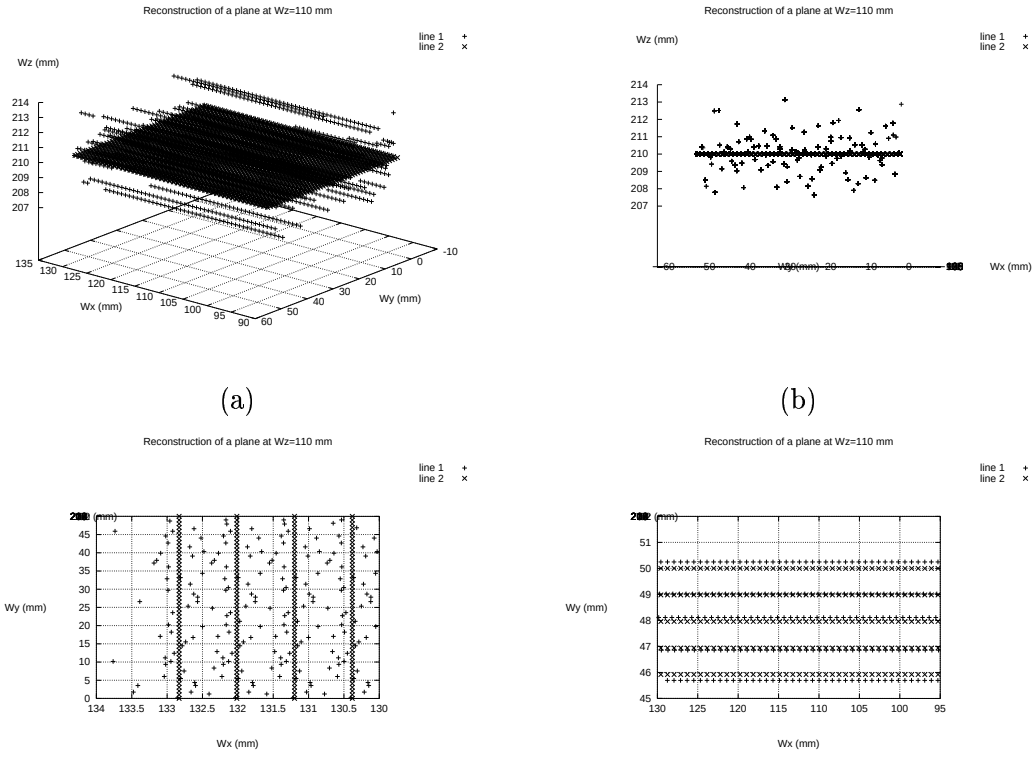


Figure 3.4: Reconstruction of a plane at Wz=110.

Figures 3.3 and 3.4 (a) to (d) show the reconstruction of a plane for Wz=0 and Wz=110 mm respectively. Parts (a) show a general view of the reconstructions, while parts (b) to (d) show different views of (a) where the errors in the Z, X and Y directions respectively can be appreciated in more detail. As can be seen, considering the error in pixels, the performance of the system is considerably good, although it is expected to decrease when it is implemented using real devices. One of the reasons is that in this simulation, the stripe detection has been assumed to be carried out with a nearly ideal sub-pixel accuracy, that is, up to the machine accuracy, which is very high. In real systems based on discrete array cameras, the quantisation error is considerable, although sub-pixel accuracy peak detection techniques have shown to perform up to 1/8 - 1/9 of a pixel. At the time of writing this work, a working prototype based on a basis similar to the system under consideration is being implemented.

3.2 Smart sensors. Yet another approach

Figure 3.5 shows the geometrical arrangement for the simulation of a hypothetical such device. At the time of writing this work, a prototype based on the S7585 sensor of Hamamatsu is under construction. This sensor is a simple 5x5 photodiode array, hence the rest of the cirtuitry, both analog and digital, is being assembled in separate PCBs. In addition, a detection scheme similar to that of Yokohama et al.([41]) has been adopted and a new calibration method has been contributed, where just a simple known plane is required as a calibration target.

3.2.1 Slit detection

Yokohama et al. used two photo-sensitive areas in each of the pixels in the array, one of them being slightly bigger in area than the other. The slit detection was achieved by comparison between the voltage outputs of both neighbour photo-diodes, hence, a detection signal was triggered at the instant in which the peak light intensity falled onto the middle of the two photo-diodes. Figure 3.6 depicts the evolution of the two voltages with time, showing the instant of detection. In our approach, the photo-diodes in the

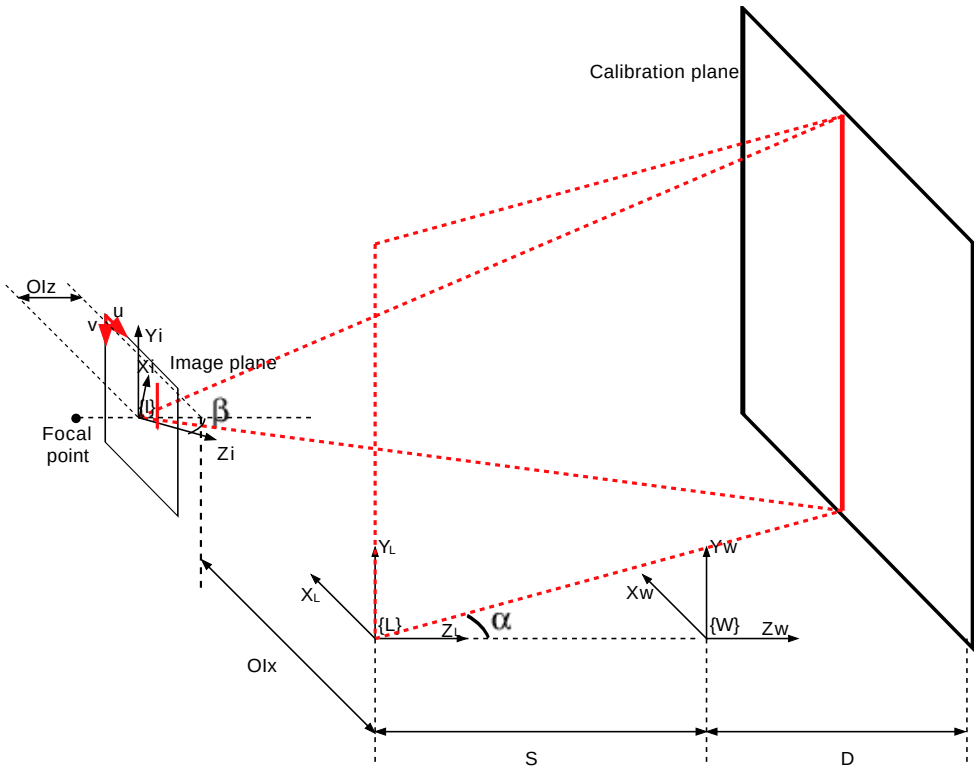


Figure 3.5: Co-ordinate frames for Smart Sensor-based system simulation.

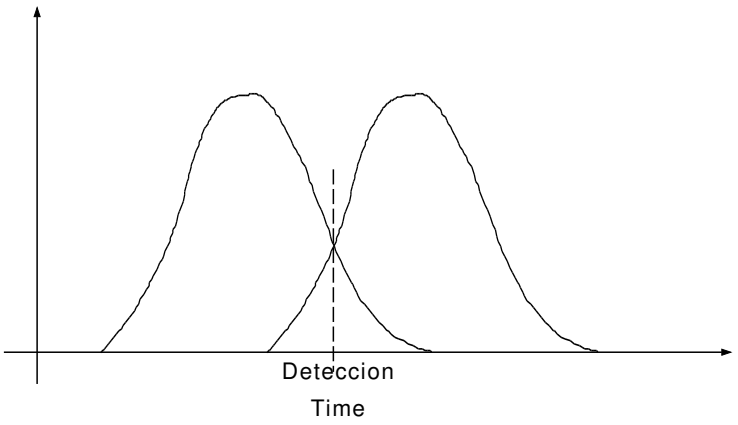


Figure 3.6: Detection of laser slit by voltage comparison.

array have been used in couples, such that four detection "points" exist in each row. Hence the two voltages to be compared have been taken from the two neighbour pixels in each couple. Figures 3.7 and 3.8 show one row of the array, in the common kathode configuration (by manufacture default) and a simple comparator circuit which provides the detection signal, respectively. In addition, an offset voltage has been added in one of the comparator inputs in order to cancel out thermal noise, avoiding false detections.

3.2.2 Calibration procecure

According to figure 3.5, the world co-ordinate frame is chosen to be placed in front of the {L} (laser) co-ordinate frame, with both Z_L and Z_W axis being coincident and the origins separated by OLz. The image sensor co-ordinate frame {I} is considered to be separated from {L} by OIx and OIz along the X_L and Z_L axis respectively. Note that the $X_C Z_C$ and $X_L Z_L$ planes are supposed to be coincident. In addition, {I} has a non-zero orientation angle (β) with respect to the Z_W axe.

Assuming a pinhole model, a set of linear equations may be used in order to describe the behaviour of a camera, as stated and proved by [12]. The pinhole model is valid if no zoom or wide angle lenses are used, since these lenses introduce radial as well as

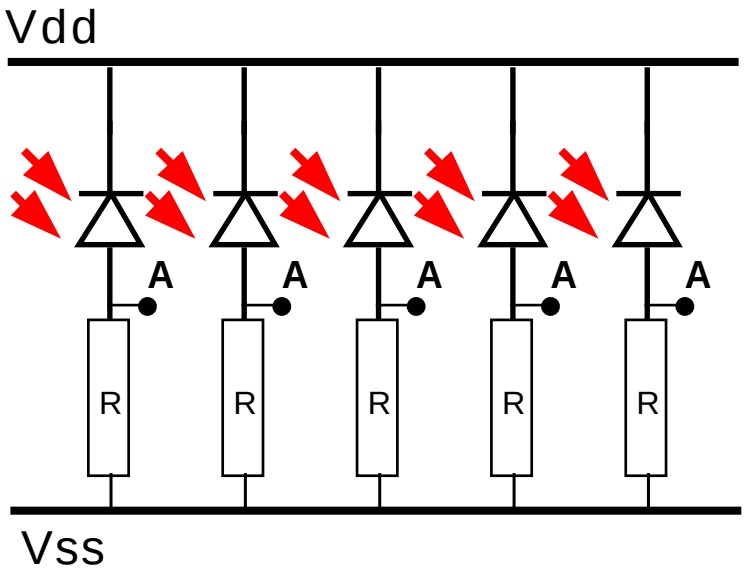


Figure 3.7: One row in common kathode.

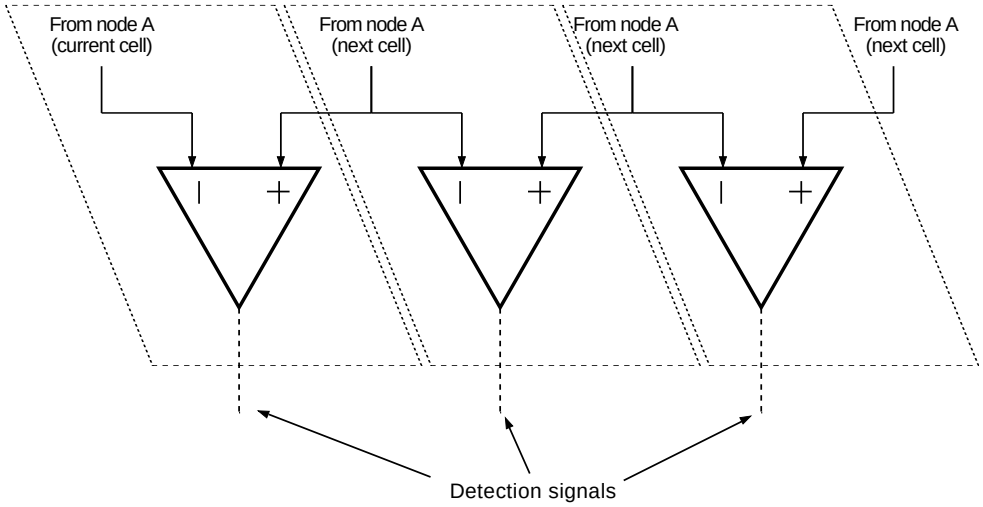


Figure 3.8: Detection circuit.

tangential distorsion in the image, both of them non-linear effects. Since we are using a 50mm lens (a reflex camera body has been used as a housing for the 5x5 array), it is reasonable to assume this linear camera model for our purposes.

Camera model: intrinsic and extrinsic parameters

Any kind of discrete array-based video camera may be modelled in two stages, consisting in how the image points or pixels (in pixel co-ordinates) are related to an arbitrarily placed metric co-ordinate frame -*Intrinsic Parameters*-, and in how this metric co-ordinate frame is located with respect to an arbitrary world co-ordinate frame {W} -*Extrinsic Parameters*-.

The intrinsic parameters of a camera, assuming a pinhole model, are the focal distance (f), the principal point pixel co-ordinates (u_0, v_0) and the horizontal and vertical pixel pitch constants of proportionality (K_u, K_v). The reader is pointed to [12] and [29] to get deeper in camera calibration issues. In our case, though, we are not dealing with an already housed camera, but it is *the* camera what is being built, hence the horizontal and vertical constants of proportionality are supplied by the sensor manufacturer and the principal point must be well aligned with the optical axe of the lens, which is a mechanical issue, and has been chosen to be $(0, 0)$. Hence only f must be estimated in the calibration procedure.

The estimation of the camera extrinsic parameters describe how the camera itself (or more concretely, the camera co-ordinate frame) is located with respect to the world co-ordinate frame. In our case, only the X (OIx) and Z (OIz) co-ordinates have been assumed to be variable (i.e. unknown a priori), since Y_C has been assumed to be parallel to both Y_L and Y_W. This assumption may be adopted if the laser plane emitter is perfectly vertical with respect to Y_L.

Calibration equations

The position and orientation of {I} with respect to {W} is described by the homogeneous transformation shown in equation 3.1, hence, the point co-ordinates of the focal point, expressed as ${}^{I}FP = [0, 0, -f, 1]^T$ with respect to {I}, is obtained with respect to {W}

making ${}^{W}FP = {}^{W}T_I \cdot {}^{I}P$, yielding the expression of equation 3.2.

$$ {}^{W}T_I = \begin{bmatrix} cos(\beta) & 0 & sin(\beta) & OIx \\ 0 & 1 & 0 & 0 \\ -sin(\beta) & 0 & cos(\beta) & OIz \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.1} $$

$$ {}^{W}FP = \begin{bmatrix} OIx - f \cdot sin(\beta) \\ 0 \\ OIz - f \cdot cos(\beta) \\ 1 \end{bmatrix} \tag{3.2} $$

Once the focal point is expressed with respect to {W}, each of the detection point locations will be used in order to get the direction vectors (equation 3.3) which, together with ${}^{W}FP$, yield the expression of equation 3.4, which is the line-of-sight equation in parametric form for each detection point, where R_X, R_Y, R_Z are the co-ordinates of the line-of-sight points, nx, ny are the pixel co-ordinates of each detection point, with respect to {I}, and dx, dy are the pixel pitch in the horizontal and vertical directions respectively. The pixel pitch is the inverse of the pixel pitch constant of proportionality (K_u, K_v). As can be seen in equation 3.4, R_Y is dependent only on the Y co-ordinate in pixels (ny) and the pixel vertical pitch (py). This is due to the fact that there is no *vertical* information provided by the laser scan, since the laser shape is a vertical plane, which projects a slit parallel to the Y_W axe. Hence, the measurement error in the Y direction is ±7.5mm at a 500mm range, independently of the scanning mechanism, if the vertical detection point is assumed to fall in the centre of the pixels. Clearly, this shows that in order to minimise the reconstruction errors in the vertical, or Y, direction it is recommended that the pixels have the vertical pitch as short as possible.

$$ V_d^W = \begin{bmatrix} nx \cdot dx \cdot cos(\beta) - f \cdot sin(\beta) \\ ny \cdot dy \\ nx \cdot dx \cdot sin(\beta) - f \cdot cos(\beta) \\ 0 \end{bmatrix} \tag{3.3} $$

$$\begin{bmatrix} R_x \\ R_y \\ R_z \\ 1 \end{bmatrix} = \lambda \cdot \begin{bmatrix} nx \cdot dx \cdot cos(\beta) - f \cdot sin(\beta) \\ ny \cdot dy \\ nx \cdot dx \cdot sin(\beta) - f \cdot cos(\beta) \\ 0 \end{bmatrix} + \begin{bmatrix} OIx - f \cdot sin(\beta) \\ 0 \\ OIz - f \cdot cos(\beta) \\ 1 \end{bmatrix} \tag{3.4}$$

As shown in chapter 2.1, in order to get the reconstruction (or estimated three-dimensional co-ordinates) of a point contained in a line-of-sight, it is necessary to obtain the point intersection of the laser plane with the line-of-sight itself. Hence, the laser plane equation must be obtained in terms of α, the scanning angle. ${}^{L}O = [0, 0, 0, 1]^T$ is the laser rotation centre point with respect to {L}. Since the orientation of {L} has chosen to be equal to the orientation of {W}, and its position is $-S$ mm away from the origin of {W} along Z_W, equation 3.5 expresses ${}^{W}T_L$ by means of a homogeneous transformation. Hence, the laser rotation point (it is a rotation *axe*, indeed, but a *point* has been chosen for notation convenience only), can be found with respect to {W} by left-multipliying ${}^{L}O$, such that ${}^{W}O = {}^{W}T_L \cdot {}^{L}O$, and its expression is shown in equation 3.6.

$${}^{W}T_L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -S \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.5}$$

$${}^{L}O = \begin{bmatrix} 0 \\ 0 \\ -S \\ 1 \end{bmatrix} \tag{3.6}$$

The normal vector which describes the laser plane orientation may be obtained by left-multipliying the normal vector expression when $\alpha = 0$, by a rotation homogeneous transformation upon the Y_L axe, as it is expressed in equation 3.7. In addition, if V_n^L is left-multiplied by ${}^{W}T_L$, the normal vector expression with respect to {W} is obtained, as shown in equation 3.8. Since the orientation of {L} and {W} are equal, there have been no changes in the normal vector expression. Once both the normal vector and a

point of the laser plane (which has been chosen to be the rotation centre for convenience) have been found, the laser plane equation in terms of α is configured, and its expression is shown in equation 3.9.

$$V_n^L = \begin{bmatrix} cos(\alpha) & 0 & sin(\alpha) & 0 \\ 0 & 1 & 0 & 0 \\ -sin(\alpha) & 0 & cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} cos(\alpha) \\ 0 \\ -sin(\alpha) \\ 0 \end{bmatrix} \tag{3.7}$$

$$V_n^W = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -S \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} cos(\alpha) \\ 0 \\ -sin(\alpha) \\ 0 \end{bmatrix} = \begin{bmatrix} cos(\alpha) \\ 0 \\ -sin(\alpha) \\ 0 \end{bmatrix} \tag{3.8}$$

$$X \cdot cos(\alpha) - Z \cdot sin(\alpha) = S \cdot sin(\alpha) \tag{3.9}$$

Once both the parametric equation of the line-of-sight and the laser plane equation have been found, a system of 4 linear equations can be set, but there is still a fifth equation which must be considered in order to obtain the calibration parameters. The calibration parameters are β, S, OIx, OIz and f, and they must be obtained from the knowledge of a calibration target object, which is, in our case, a plane located at known distances along the Z_W axe and parallel to the $X_W Y_W$ plane. Hence the fifth equation of the linear system is $Z = d_i$, where d_i is the distance of the i-th plane to the origin of {W}, as shown in equations 3.10.

$$\begin{cases} (1) X \cdot cos(\alpha) - Z \cdot sin(\alpha) = S \cdot sin(\alpha) \\ (2) X = \lambda \cdot (nx \cdot dx \cdot cos(\beta) - f \cdot sin(\beta)) + OIx - f \cdot sin(\beta) \\ (3) Y = \lambda \cdot ny \cdot dy \\ (4) Z = \lambda \cdot (-nx \cdot dx \cdot sin(\beta) - f \cdot cos(\beta)) + OIz - f \cdot cos(\beta) \\ (5) Z = d_i \end{cases} \tag{3.10}$$

If λ is isolated from the 3.10(2) and 3.10(4) equations, two expressions for λ are obtained. If these two expressions are equated, the line-of-sight equation in the plane $X_W Z_W$ is

obtained, as shown in equation 3.11. Note that it is senseless to consider the line-of-sight equation in the vertical direction, since it does not contribute any additional information for obtaining the calibration parameters.

$$\begin{aligned} &Z \cdot (nx \cdot dx \cdot cos(\beta) - f \cdot sin(\beta)) + X \cdot (nx \cdot dx \cdot sin(\beta) + f \cdot cos(\beta)) + \\ &+OIz \cdot (f \cdot sin(\beta) - nx \cdot dx \cdot cos(\beta)) - OIx \cdot (f \cdot cos(\beta) + nx \cdot dx \cdot sin(\beta)) + \\ &+f \cdot nx \cdot dx = 0 \end{aligned} \tag{3.11}$$

If 3.10(5) is substituted into 3.10(1) and 3.11, a new system of 2 linear equations is obtained, which, after isolating X in both of them and equating, equation 3.12 is obtained, where the only unknowns are the calibration parameters.

$$\begin{aligned} &OIz - OIx \cdot tan(\beta) + nx \cdot dx \cdot (OIz \cdot \tfrac{tan(\beta)}{f} + OIx \cdot \tfrac{1}{f} - sec(\beta)) - \\ &-nx \cdot dx \cdot d_i \cdot \tfrac{1}{f} + d_i \cdot tan(\beta) - nx \cdot dx \cdot d_i \cdot tan(\alpha) \cdot \tfrac{tan(\beta)}{f} - \\ &-nx \cdot dx \cdot tan(\alpha) \cdot \tfrac{tan(\beta) \cdot S}{f} - tan(\alpha) \cdot S = d_i \cdot tan(\alpha) \end{aligned} \tag{3.12}$$

Obtaining the calibration parameters

In order to obtain the calibration parameters (β, S, OIx, OIz and f), several complete scans on the calibration plane at different ranges must be made. In these scans, all the detection points **in a row** must detect the scanning stripe projected on the calibration plane. A complete scan will detect the scanning *times* in which the stripe *cuts* each of the line-of-sights, hence, in a given instant of time, when the stripe illuminates the **i-th** calibration plane, the laser light will intersect the **j-th** line-of-sight, as shown in figure 3.9. Using i and j as the plane and line-of-sight indices, equation 3.12 can be arranged as shown in equation 3.13.

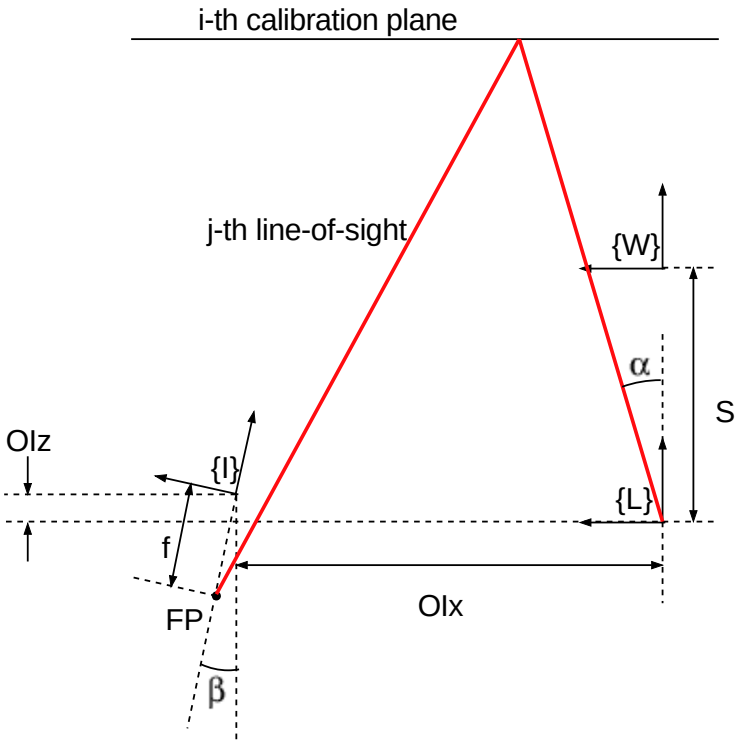


Figure 3.9: Intersection between a line-of-sight and the laser plane.

$$\begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & dx \cdot n_j & -dx \cdot d_i \cdot n_j & d_i & -dx \cdot d_i \cdot n_j \cdot tan(\alpha_{ij}) & -dx \cdot n_j \cdot tan(\alpha_{ij}) & -tan(\alpha_{ij}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \cdot$$

$$\cdot \begin{bmatrix} OIz - OIx \cdot tan(\beta) \\ OIz \cdot \frac{tan(\beta)}{f} + OIx \cdot \frac{1}{f} - \frac{1}{cos(\beta)} \\ \frac{1}{f} \\ tan(\beta) \\ \frac{tan(\beta)}{f} \\ S \cdot \frac{tan(\beta)}{f} \\ S \end{bmatrix} = \begin{bmatrix} \vdots \\ d_i tan(\alpha_{ij}) \\ \vdots \end{bmatrix} \tag{3.13}$$

In order to solve the unknowns, a minimum of 7 *detections* or points should be provided for satisfiying all the constraints, but since any real image capture system is noisy, according to [16], it is much more advisable to include many more points so an overdetermined set of equations can be solved minimising some cost function. In our simulation, the calibration plane has been scanned at 4 different ranges, obtaining 4 points per scan, which yields a total of 16 points. In addition, since it has been proved sufficient above, only the row at $Y_I = 0$ has been used, and the singular value decomposition method has been utilised in order to solve the set of linear equations.

If the unknowns column vector is called U, with components $U_1 \ldots U_7$, the expressions of the calibration parameters can be obtained from 3.13, as shown in equation 3.14. The real parameters and the estimated ones do not show to be appretiably different, and their values in our simulation are: $\beta = 15^o$, $S = 200$ mm, $OIz = 20$ mm, $OIx = 400$ mm and $f = 50$ mm. Section 3.2.3 reports a study of the noise influencing in the light detection and how it affects the reconstruction, while section 3.2.4 shows the simulation results for different ranges and noise sources. The estimated noise power has been considered

when the calibration parameters have been calculated.

$$\begin{aligned} \beta &= atan(U_4) \\ S &= U_7 \\ f &= \frac{1}{U_3} \\ OIz &= \frac{U_2 - U_1 U_3 U_4 + sec(\beta)}{U_3(1+U_4^2)} \\ IOx &= U_1 + OIz \cdot U_4 \end{aligned} \tag{3.14}$$

3.2.3 Noise evaluation

Noise plays an important role in the results evaluation of any measuring device. Since laser scanners use to include at least 2 measuring device components (a camera and an angular sensor), it is necessary to evaluate which is the error introduced by both. In standard cameras, the most influencing noise has been found to follow a gaussian probability distribution, which is a consequence of the point spread function, due to the imperfections in the lenses and the gray level digitisation. On the other hand, an angular sensor introduces electrical noise, as well as quantisation noise, either if this angular sensor is of a discrete nature or in the analog-to-digital stage conversion. In our case, the angular sensor is not used as a measure of the scanning angle itself, but it is the same image sensor, together with the additional circuitry, which provides a measure of the scanning angle. The angular sensor is used for constant speed control purposes only.

According to the above discussion, it is straightforward to spend a few paragraphs evaluating which kind of noise affects mainly our measuring device, and to which extent this noise distorts the measure. Since it is clear that it is the image sensor itself which provides the measure of the scanning angle, the noise evaluation study is restricted to the noise present in the image sensor, while the noise present in the angular sensor is assumed to be filtered out by the constant speed controller electronics. Under these assumptions, two main noise sources have been found to influence the three-dimensional measure of our device: **electrical noise** and **quantisation noise**. The quantisation noise introduces a significant error in the measure, due to the truncation which the number of bits employed in the angular measure digitisation induces. It has been proved

Table 3.2: The three main electrical noise sources under consideration.

Photon shot noise	$\#e^- = \sqrt{Q_{signal}/q}$
Dark current noise	$\#e^- = \sqrt{Q_{dark}/q}$
Johnson noise	$V_{RMS} = \sqrt{4 \cdot k \cdot T \cdot BW}$

that this error is more important for low bit numbers, but when the number of bits rises above a particular threshold, the electrical noise becomes more important in the measure distortion.

Electrical noise, which has a uniform probability distribution, has shown to be of a great importance in image formation quality. Dierickx [10] and [11] described the different noise sources which influence in image sensors in general but in CMOS-based image sensors particularly. Since we have used a CMOS photo-diode array, it has been considered of a great convenience to follow Dierickx's performance evaluation given in the CMOS image sensors-concepts and limits SPIE Photonics West 2000 Short Courses. The influencing electrical noise, in our sensor, can be decomposed into 3 categories: ***Photon shot noise***, ***Dark current noise*** and ***Thermal or Johnson noise***. Photon shot noise is due to the statistical variation on the number of absorbed/converted/collected photons (converted to electrons or charge carriers), it is considered as a fundamental noise source and it is device-independent, i.e there is no way to cancel it out, but it is inherent to the photon to charge conversion. Dark current is the apparent current present in the photo-diode under darkness conditions. The presence of dark current introduces the so called dark current noise, which is the statistical variation on the number of generated/recombined/collected charge carriers. Finally, the thermal or Johnson noise is the statistical variation of charge carriers agitation due to the temperature. Table 3.2 collects the algebraic expressions of these 3 noise sources. In order to have a thorough performance evaluation, the total signal-to-noise ratio of the sensor under normal operation must be taken into account. This total noise is the sum of the noise induced by the three noise sources independently, since the principle of superposition applies. The device specifications as well as the signal-to-noise ratio evaluation are

shown in table 3.3, where the laser power has been chosen to be of 6 mW peak, and only the case in which the laser is projected on a dark object (20% of reflectivity).

Table 3.3: Performance evaluation in terms of the S/N.

Light features		
Distance of performance evaluation	0.3	[m]
Aperture angle	85	[deg]
Laser emitter peak light power	6	[mW]
Major axis	549.798704410454	[mm]
Minor axis	1.5	[mm]
Slit area	0.0025908653560936	[m2]
Light flux (direct)	2.31582856511175	[W/m2]
Light flux (reflected from white object)	1.8526628520894	[W/m2]
Light flux (reflected from dark object)	0.463165713022351	[W/m2]
Sensor features		
Pixel size	1300	[um]
Pixel area (square pixel)	0.00000169	[m2]
Pixel power	0.000000782750055007773	[W]
Photo sensitivity	0.39	[A/W]
Pixel current	0.00000305272521453031	[A]
Integration time	0.000065	[s]
Photo charge	0.000000000019842713894447	[C]
Photo charge	124016961.840294	[e-]
Pixel capacitance	0.00000000001	[F]
Pixel voltage	1.9842713894447	[V]
Noise Calculation		
Photon shot noise	11136.2903087291	[e-]

Dark currrent	0.05	[nA]
Dark current noise	142.521928137392	[e-]
Signal-to-noise ratio		
Total noise (shot+dark current)	11278.8122368665	[e-]
S/N (Photo charge/noise electrons)	10995.5693237738	Linear
S/N (Photo charge/noise electrons)	80.82435441977	(dB)
S/N considering thermal noise		
Noise equivalent voltage (shot+dark)	0.000180460995789864	[V]
Thermal (Johnson) noise	0.000000122758299108451	[V]
Total noise voltage	0.000180583754088973	
S/N (Pixel voltage/Noise total voltage)	**10988.0946902182**	Linear
S/N (Pixel voltage/Noise total voltage)	**80.8184478661937**	(dB)

The laser stripe has been modelled as an elongated ellipse, according to the manufacturer's data. The sensor features have been picked from the photo-diode array data-sheet.The photo charge is the generated charge (or current) in the photo-diode, during the exposition (or integration) time, which is the time during which a laser stripe image illuminates one pixel. As can be observed, the S/N does not change considerably either accounting for thermal noise or not, hence, a signal-to-noise ratio of **80.82** dB has been considered in our simulations.

3.2.4 Simulation results

Table 3.4 show the results of the reconstruction of a plane for different ranges, whith no errors due to quantisation. Table 3.5 show the reconstruction errors due only to quantisation, for different quantisation bits. As can be observed, the error due to quantisation becomes smaller and smaller as the number of quantisation bits increase, but when using 16 bits or more, errors due to electrical noise become more important, hence, it does not make sense to include more than 24 bits, for the quantisation, since, as can be seen in table 3.6, the errors stabilise for 16, 24 and 32 bits.

Table 3.4: Metric errors due to electric noise.

Zw (mm)	10	110	210	310
dX (um)	+4,37E-01	+1,08E+01	+2,93E+01	+5,53E+01
dZ (um)	+4,93E-01	+1,21E+01	+3,31E+01	+6,24E+01

Table 3.5: Metric errors due to quantisation.

	Zw (mm)	10	110	210	310
8 bits	**dX (um)**	+1,73E+03	+1,19E+03	+2,47E+03	+3,74E+03
	dZ (um)	+1,88E+03	+1,95E+03	+3,05E+03	+4,04E+03
16 bits	**dX (um)**	+2,70E+00	+1,28E+01	+2,59E+00	+2,80E+01
	dZ (um)	+5,93E+00	+1,45E+01	+5,69E+00	+3,16E+01
24 bits	**dX (um)**	+1,14E-02	+3,87E-02	+6,29E-02	+4,85E-02
	dZ (um)	+1,98E-02	+4,37E-02	+7,10E-02	+5,47E-02
32 bits	**dX (um)**	+9,44E-05	+1,81E-04	+2,88E-04	+4,32E-04
	dZ (um)	+1,06E-04	+2,04E-04	+3,25E-04	+4,87E-04

Table 3.6: Metric errors due to both quantisation and electric noise.

	Zw (mm)	10	110	210	310
8 bits	**dX (um)**	+1,73E+03	+1,19E+03	+2,47E+03	+3,73E+03
	dZ (um)	+1,88E+03	+1,94E+03	+3,05E+03	+4,03E+03
16 bits	**dX (um)**	+8,53E+00	+6,96E-01	+1,42E+01	+4,76E+01
	dZ (um)	+9,62E+00	+2,91E+00	+1,60E+01	+5,36E+01
24 bits	**dX (um)**	+4,20E-01	+1,03E+01	+2,82E+01	+5,29E+01
	dZ (um)	+4,74E-01	+1,16E+01	+3,17E+01	+5,97E+01
32 bits	**dX (um)**	+4,37E-01	+1,08E+01	+2,93E+01	+5,53E+01
	dZ (um)	+4,93E-01	+1,21E+01	+3,31E+01	+6,24E+01

Chapter 4

Conclusions and further work

An exhaustive review of representative laser slit scanning three-dimensional digitiser systems has been reported. A classification table has been proposed showing the main features for each system, including references to the corresponding authors, expanding the discussion about the different calibration methods.

There are no two identical calibration methods, but a different one exists for every different digitiser. Nevertheless, according to the nature of the hardware involved in each system, accuracy is always strongly dependent on the system calibration as well as the stripe detection accuracy.

If video cameras are involved in the digitiser hardware, very accurate camera calibration methods have to be chosen according to the application specs and image distortion. In addition, the relative pose between the imager and the laser emitter must be carefully and accurately measured in order to get a proper shape measurement. Manual pose measurements may yield very significant inaccuracies in the shape measurement.

Smart sensors is a very interesting field of application, since they offer fast and accurate shape measurements, although resolution is still a challenge for these devices. Since laser slit scanning time measurement is inherent to the imager, the system calibration does not involve any camera calibration method, because both the bi-dimensional point position (on the image sensor) and angular measurement (by time stamp recording) is fulfilled at the same time on the same device. In this kind of sensors, accuracy depends on the

photosensitive area, the stripe detection method, the S/N of the sensing electronics and the laser power.

In addition, the method of Chen & Kak has been reproduced as a simulation and a working prototype is being implemented in order to obtain more thorough accuracy evaluation. The real implementation of such system will allow our research lab to include a digitisation facility for research purposes, as well as experimenting with new calibration algorithms as well as new stripe detection operators.

A new calibration technique for smart sensor-based digitisers has been reported and a working prototype is currently being implemented using a standard reflex camera housing. Although a simple 5x5 photo-diode imager sensor is being used, it is sufficient for acquisition speed and accuracy evaluation as well as for image acquisition hardware design for range or $2\frac{1}{2}$D images. In addition, it is our purpose to continue with this field of application and a new digitiser based on a similar principle of operation is being envisaged, using a linear image sensor as a first approach, extending the capabilities to a bi-dimensional smart sensor array in the future, using CMOS-based image sensor design facilities.

Bibliography

[1] Gerald J. Agin and Thomas O. Binford. Computer description of curved objects. In *International Joint Conferences on Artificial Intelligence*, pages 629–640, 1973.

[2] Kazuo Araki, Yukio Sato, and Srinivasan Parthasarathy. High speed rangefinder. In SPIE, editor, *Optics, illumination and image sensing*, volume 850, pages 184–188, 1987.

[3] Mitsuru Baba, Tadataka Konishi, and Nobuaki Kobayashi. A new fast rangefinding method based on a non-mechanical scanning mechanism and a high-speed image sensor. In *IEEE Instrumentation and Measurement Technology Conference*, pages 957–962, 1997.

[4] J Batlle, E Mouaddib, and J Salvi. A survey: Recent progress in coded structured light as a technique to solve the correspondence problem. *Pattern Recognition*, 31(7):963–928, 1998.

[5] P.J. Besl. Active, optical range imaging sensors. *Machine Vision and Applications*, 1:127–152, 1988.

[6] Vladimir Brajovic, Kenichi Mori, and Nebojsa Jankovic. 100 frames/s cmos range image sensor. In *2001 IEEE International Solid State Conference*, pages 256,257 and 453, 2001.

[7] G. Champleboux, S. Lavallée, P. Sautout, and P. Cinquin. Accurate calibration of cameras and range imaging sensors: the npbs method. In *Proceedings of the*

1992 IEEE International Conference on Robotics and Automation, pages 1552–1557. IEEE, 1992.

[8] Guillaume Champleboux, Stéphane Lavalée, Richard Szelinski, and Lionel Brunie. From accurate imaging sensor calibration to accurate model-based 3-d object localization. In *Proceedings of the IEEE Computer Vision and Pattern Recognition conference*, pages 83–89. IEEE, 1992.

[9] C.H. Chen and A.C. Kak. Modeling and calibration of a structured light scanner for 3-d robot viison. In *Proceedings of the IEEE conference on robotics and automation*, pages 807–815. IEEE, 1987.

[10] Bart Dierickx. *CMOS image sensors. Concepts.* SPIE.

[11] Bart Dierickx. *CMOS image sensors. Limits.* SPIE.

[12] O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint.* The MIT Press, 1993.

[13] Andrew Gruss, Takeo Kanade, and L. Richard Carley. *Machine Vision for Three-Dimensional Scenes*, chapter A fast lightstripe rangefinding system with smart VLSI sensor, pages 381–397. Academic Press, Inc., 1990.

[14] Andrew Gruss, Shigeyuki Tada, and Takeo Kanade. A vlsi smart sensor for fast range imaging. In *Proceedings of the 1992 IEEE/RSJ International Conference on Intelligent Robots ans Systems*, pages 349–358, 1992.

[15] E.L. Hall, C.A. McPherson J.B.K. Tio, and F.A. Sadhadi. Measuring curved surfaces for robot vision. *Computer Journal*, 1982.

[16] Richard Hartley and Andrew Zisserman. *Multiple view geometry.* Cambridge University Press, 2000.

[17] Richard I. Hartley. Triangulation. *Computer Vision and Image Understanding*, pages 146–157, 1997.

[18] Kazuyuki Hattori and Yukio Sato. Handy rangefinder for active vision. In *Proceedings of the 1995 ieee international conference on robotics and automation*, pages 1423–1428, 1995.

[19] Kazuyuki Hattori and Yukio Sato. Accurate rangefinder with laser pattern shifting. In *Proceedings of the 1996 ieee international conference on robotics and automation*, pages 849–853, 1996.

[20] Yuko Hori, Mitsuru Baba, and Tadataka Konishi. A simultaneous high-speed measuring system of multiple points. In *IMTC'94*, pages 1365–1368, 1994.

[21] E. Horn and N. Kiryati. Toward optimal structured light patterns. *Image and Vision Computing*, 17(2):87–97, February 1999.

[22] S. Inokuchi, K. Sato, and F. Matsuda. Range imaging system for 3-D object recognition. In *Proceedings of the International Conference on Pattern Recognition*, pages 806–808, 1984.

[23] Takeo Kanade, A. Gruss, and L.R. Carley. A very fast vlsi rangefinder. In *Proc. IEEE International Conference on Robotics and Automation*, volume 2, pages 1322–1329, April 1991.

[24] J. Lazzaro, S. Ryckebusch, M.A. Mahowald, and C.A. Mead. Winner-take-all networks of o(n) complexity. In D.Tourestzky, editor, *Advances in Neural Information Processing Systems*, volume 1, pages 703–711, 1988.

[25] Kouichi Nakano, Yasuo Watanabe, and Sukeyasu Kanno. Extraction and recognition of 3-dimensional information by projecting a pair of slit-ray beams. In *Proceedings of the 9th International Conference on Pattern Recognition*, pages 736–743, 1988.

[26] R.J. Popplestone, C.M. Brown, and G.F. Crawford A.P. Ambler. Forming models of plane-and-cylinder faceted bodies from light stripes. In *International Joint Conferences on Artificial Intelligence*, pages 664–668, 1975.

[27] J. L. Posdamer and M. D. Altschuler. Surface measurement by space-encoded projected beam systems. *Computer Graphics and Image Processing*, 18(1):1–17, 1982.

[28] Ian D. Reid. Projective calibration of a laser-stripe range finder. *Image and Vision Computing*, (14):659–666, 1996.

[29] Joaquim Salvi, Xavier Armangué, and Joan Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, 35(7):1617–1635, 2002.

[30] G. Sansoni, M. Carocci, and R. Rodella. Calibration and performance evaluation of a 3-D imaging sensor based on the projection of structured light. *IEEE Transactions on instrumentation and measurement*, 49(3):628–636, June 2000.

[31] Kosuke Sato, Atsushi Yokohama, and Seiji Inokuchi. Silicon range finder – a realtime range finding vlsi sensor –. In *IEEE 1994 Custom Integrated Circuits Conference*, pages 14.5.1–14.5.4, 1994.

[32] Kotsuke Sato, Atsushi Yokohama, and Seiji Inokuchi. Silicon range finder. In *Proceedings of the IEEE 1994 custom integrated circuits conference*, pages 339–342, 1994.

[33] Yukio Sato. Active rangefinding and recognition with cubicscope. In *Proceedings of the 2nd asian conference on computer vision*, pages 211–217, 1995.

[34] Yukio Sato, Hiroo Kitagawa, and Hiroichi Fujita. Shape measurement of curved objects using multiple slit-ray projections. *IEEE Transactions on Pattern Analysis and Machine Intellignece*, PAMI-4(6):641–646, 1982.

[35] Yukio Sato and Masaki Otsuki. Three-dimensional shape reconstruction by active rangefinder. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 142–147. IEEE, 1993.

[36] Yoshiaki Shirai and Motoi Suwa. Recognition of polyhedrons with a range finder. In *International Joint Conferences on Artificial Intelligence*, pages 80–87, 1971.

[37] E. Trucco, R.B. Fisher, A.W. Fitzgibbon, and D.K. Naidu. Calibration, data consistency and model acquisition with a 3-d laser striper. *International Journal of Computer Integrated Manufacturing*, 11(4):292–310, 1998.

[38] R.Y. Tsai. *Synopsis of recent progress on camera calibration for 3D machine vision*, chapter The Robotic Review, pages 147–159. The MIT Press, 1986.

[39] R.Y. Tsai. A versatile camera calibration technique for high accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE International Journal on Robotics and Automation*, RA-3:323–344, 1987.

[40] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distorsion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:965–980, 1992.

[41] A. Yokoyama, K. Sato, T. Yoshigahara, and S. Inokuchi. Realtime range imaging using adjustment-free photo-vlsi. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1751–1758, 1994.

[42] Xiaoyang Yu, Jian Zhang, Liying Wu, and Qing Lin. Calibration errors in laser scanning 3d vision measurement using the space encoding method. In *Automated optical inspection*, volume 3558, pages 298–303. SPIE, 1998.

[43] Xiaoyang Yu, Jian Zhang, Liying Wu, and Xifu Qiang. Laser scanning device used in space-encoding rangefinder. In *Automated optical inspection for industry*, pages 490–495. SPIE, 1996.