A PORTABLE AND INEXPENSIVE LASER-BASED 3D-CAMERA

Thesis by

Askold Veceslav Strat

In Partial Fulfillment of the Requirements

for the Degree of

Master of Science in Computer Science

State University of New York

at Stony Brook

July 2002

State University of New York
at Stony Brook

The Graduate School

Askold Veceslav Strat

We, the thesis committee for the above candidate for the

Master of Science degree,
Hereby recommend acceptance of this thesis.

_____
Professor Manuel Menezes de Oliveira Neto, Advisor

_____
Leading Professor Arie Kaufman

_____
Professor Klaus Mueller

This thesis is accepted by the Graduate School

_____
Dean of the Graduate School

Acknowledgements

The initiative to build a 3D-camera belongs exclusively to Professor Manuel M. Oliveira, who has been a wonderful advisor, a talented teacher, and a close friend to me for almost one year. Professor Olivera has managed to turn my skepticism about this project into enthusiasm and patience and taught me the importance of documenting every step of the process: "Write, write, write!" The late night discussions on the various topics of the project were invaluable for the rapid and consistent progress. His methodic approach to solving problems and his persistence gave me the confidence I needed to complete this work. Most of all, I want to thank Professor Oliveira for trusting in me.

I am grateful to Fred Wood for always being willing to share his solid practical knowledge. The few brainstorms we carried, positively influenced the construction of the 3D-camera prototype. The laser raster generator could not have been built without some of the essential parts that Fred saved for me.

I would like to thank Dr. Joseph Katz for taking some of his time to investigate the possible application of a 3D camera in the context of Symbol's business, and also for allowing me to pursue this work on my own, while using some of Symbol's resources.

I also like to thank Raj Bridgelall for the idea of using the Laser Projection Display, a product under development at Symbol, as a raster generator and to Bruce Willins for the idea of using the 3D-camera as a interactive length measuring device.

Most of all, I am thankful to my wife Daniella, for working hard to free up my time for the past several monts, while working on this project and for showing her confidence in the success of my work; I am thankful to my children, Adrian and Dwana for understanding my choice of not spending as much time with them as I would have liked to.

Abstract

Shape acquisition of real world objects and scenes has become an essential component of modern computer graphics systems. Presently, there are numerous methods and technologies available to perform this task. While each them has proven very appropriate for a set of applications, there is no single one capable of providing satisfactory results for all situations. Additionally, most of the existing devices have in common some shortcomings such as high cost, complex usage, and lack of portability due to large sizes, weight or power requirements, which have prevented them from becoming consumer-grade products.

This thesis describes the design and implementation of a portable, and inexpensive 3D camera suitable for shape acquisition of smooth surfaces. Its operation is very similar to the operation of a regular camera, requiring a single shot for shape recovery. The 3D camera consists of a laser pattern generator, an ordinary digital camera and a software module. The pattern generator uses two scanning mirrors and a laser diode to project laser stripes onto the target object as the camera acquires the image of the scene. The 3D models are reconstructed from a set of spatial coordinates recovered from the imaged laser stripes using triangulation. These models can be manipulated in 3D and visualized as sets of points or polygonal meshes, with or without texture. Textures are extracted from the original images and filtered for removal of the laser lines.

The effectiveness of the proposed design has been demonstrated by recovering the 3D shapes of several real-world objects, such as human faces. As all the other laser-based approaches, this one is also subject to several physical limitations due to the optical material properties, ambient illumination, occlusions, and speckle noise. While the design decision of recovering 3D shape from a single image makes the system easier to use, it makes the task significantly more challenging. Given all these constraints, the results are satisfactory, with all major features of the original surfaces being preserved. With the appropriate improvements, such a prototype may evolve into a commodity item, opening the doors to consumer 3D photography.

Contents

1 Introduction 1
1.1 Motivation 1
1.2 A Construction of a 3D-camera 2
1.3 Outline 4
2 Active Shape Acquisition Techniques 5
2.1 Contact-based Shape Acquisition 5
2.2 Non-contact-based Shape Acquisition 6
2.2.1 Transmissive 6
2.2.2 Reflective 6
2.3 Summary 9
3 Background 11
3.1 Speckle Effect 11
3.2 Laser Beam Deflection 12
3.3 Lissajous Patterns 16
3.4 Camera-Laser-Raster Geometry 17
3.5 Summary 19
4 Hardware 21
4.1 System Description and Use 21
4.2 Digital Camera 22
4.2.1 The Camera Optical System 23
4.2.2 The Sensor Array 24
4.2.3 Electronic Control and Digital Storage 25
4.3 Raster Generator 25
4.3.1 Functional description 26
4.3.2 Constructive Description 27
4.3.3 Electronic Control Circuit 33
4.4 System Calibration 37
4.5 Discussion 39
4.6 Summary 39
5 Software 41
5.1 The Image Processing Software 41
5.1.1 Extracting Geometric Information 41
5.1.2 Model Rendering 48
5.2 Camera Control Software 49
5.3 Raster Generator Driver Software 50
5.4 Summary 51
6 Results 52
7 Conclusions and Further Work 58
7.1 Absolute Reference for Reconstruction 60
7.2 Ambient Light Shading Information 60
7.3 Laser Lines Shading Information 61
7.4 Time-intensity Matrix 61
7.5 Measuring Tool 62

Figures

Figure 1-1: Laser lines following the objects' contours 3
Figure 2-1: Active Shape Acquisition Taxonomy (adapted from [Curless 97]) 5
Figure 3-1: Laser speckle patterns 12
Figure 3-2: Laser beam deflection 13
Figure 3-3: Mirror's deflection 14
Figure 3-4: Laser's spot speed 15
Figure 3-5: Wide scan angle effect (A=22deg.) 16
Figure 3-6: Lissajous Patterns 17
Figure 3-7: Triangulation 20
Figure 4-1: 3D-camera prototype 21
Figure 4-2: Connectors 22
Figure 4-3: Raster projected on a flat surface 26
Figure 4-4: Raster generator diagram 26
Figure 4-5: The Raster Generator 27
Figure 4-6: Laser module 28
Figure 4-7: Laser module focusing 29
Figure 4-8: High frequency scanning element 31
Figure 4-9: Low frequency scanning element 32
Figure 4-10: Low frequency scanning element deflection 33
Figure 4-11: Electronic control circuit board 34
Figure 4-12: Calibration setup 38
Figure 4-13: Electronic control circuit 40
Figure 5-1: Speckle effect in laser lines: before (a) and after (b) low pass filtering 42
Figure 5-2: Extracted Laser Lines after intensity gradient filtering 44
Figure 5-3: Lines ambiguity 45
Figure 5-4: Angle assignment 47
Figure 5-5: Camera control software interface 50
Figure 5-6: Waveform generator interface 51
Figure 6-1: Face Profile 53

Figure 6-2: Error propagation 54
Figure 6-3: Incorrect angle assignment 55
Figure 6-4: The left ear not rendered 55
Figure 6-5: Effect of y-mirror nonlinearly 56
Figure 6-6: Original image (left) and 3D model with texture (right) 57

1 Introduction

The rapid evolution of computers over the past decade enabled powerful hardware and sophisticated specialized software to render virtual 3D environments in real time. Modeling such environments, on the other hand, remains a very time consuming task, in which human intervention is essential. Recently, various technologies and implementations have become available for scanning real-world objects, enabling rapid and accurate shape acquisition. Unfortunately, in most cases, such solutions have high cost and lack portability. This thesis describes the design and implementation of an inexpensive, lightweight, fast and easy to use 3D-camera, intended for reconstruction of relatively smooth surfaces. The constructed prototype can acquire shapes of smooth surfaces up to about one cubic foot in size, which is sufficient for digitization of human faces. Conceptually, the system can be scaled either up or down, maintaining the same design.

Our approach is capable of extracting depth information from a single image. Compared with other laser-based triangulation devices, which acquire a plurality of images, each containing one laser line only, this approach faces the challenge of disambiguating among dozens of lines, each projected at different angles from a raster generator module. In this case, each of the lines appearing in an image has to be associated with the correct angle in order to provide the accurate shape reconstruction.

An important feature of our 3D-camera is its ability to operate on batteries, without a need of high voltage conversion, as in the case of a camera flash. The rather crude implementation presented here, although sufficient for the proof of concept, can be significantly improved by a more careful construction and a better design of the driving electronics.

This chapter begins with short evolutionary overview of shape and color acquisition, followed by a brief description of the constructed prototype. The last describes the organization of this thesis.

1.1 Motivation

3D-shape acquisition is a natural extension to photography, which evolved over the years from monochrome to color and from projection to 3D-shape acquisition. The recent emergence of

consumer grade CCD and CMOS based cameras added more capabilities to traditional film-based image capture. One can see the acquired image immediately after it is taken and decide if it is representative of what he or she wanted to accomplish. Even with such good feedback and a number of images taken from different angles, it may still be difficult to evaluate the shape of an object from set of pictures, especially when the illumination is not appropriate. Shadows, color saturation and reflections can produce optical illusions and distortions. The availability of devices capable of directly acquiring 3D shape is desirable and can find numerous practical applications such as reverse engineering, inspection, computer games, medicine, rare items and work of art [Curless 97]. For the purpose of comparison, not many people realized how dry and artificial monaural music sounds, until stereophonic music became ubiquitous. A new frontier beyond color photography is now to be conquered with 3D-shape acquisition. Adding such capability to a consumer-grade digital camera brings digital photography to a new dimension.

Existing 3D acquisition devices lack important requirements for becoming commodity items. Probably, the most prohibitive is the price, followed by lack of portability (due to the size, weight, and power requirements), and last but not least, the difficulty to use. This situation results from the need to employ high precision elements in the device construction, the need for calibration, and specialized software. Such cameras are often being utilized for industrial applications as reverse engineering and inspection, where accuracy is critical.

1.2 A Construction of a 3D-camera

It is possible to construct a low-end version of a 3D-camera using inexpensive digital hardware such as a laser raster generator, which can be build with components currently costing under $50 and a digital camera. Although such a device may not be suitable for accurate measurements, it should be appropriate for some graphics applications where precision is not essential and some errors can be tolerated. This thesis describes the design and implementation of such a device and discusses the results obtained with our prototype. In addition, it presents the theoretical background that the reader should consider for further developing this approach.

For an intuitive understanding of the principle exploited in this work, **Figure 1-1** shows a stripe light pattern deformed by the contours of an egg on top of a rectangular box. Behind the two, there is a flat wall, which illustrates the uniform characteristic of the pattern. In this example, the pattern was projected from below the camera that took the picture; the shadows created by the objects are visible on the wall as dark regions.

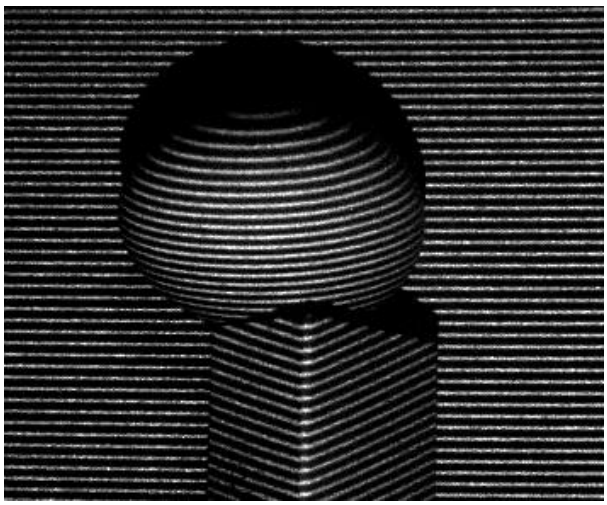


Figure 1-1: Laser lines following the objects' contours

It is possible to reconstruct the shapes of the objects in **Figure 1-1**, given that the characteristics of the pattern projector are well defined and the lines deformed by the objects' contours are distinguishable in the image. In this particular arrangement, all the lines in the image are projected from the same point, which is at fixed known distance from the camera that took the picture. Since the stripped pattern is projected from one point, each of the lines can be imagined as being the intersection between a plane of light fanning out from that point and the surfaces of the objects in the image. Each of these planes makes a well-defined angle with the optical axis of the camera. Knowing these angles, the object can be reconstructed from the image, recovering the actual dimensions and shapes. Evidently, surfaces that are not visible in the image cannot be reconstructed; in those cases, more than one image would be needed to cover the entire surface of the object.

The accuracy of the reconstruction is much dependent on the "quality" of the projected pattern. The process assumes that the projected lines are straight and the angles between the

light planes are equal. In other words, all imaginary laser planes have to intersect in one line, which passes through the point from which the laser raster is generated. That line should be parallel with the horizontal direction of the camera (for the standard camera coordinate system) and the angles between successive planes should be equal. Furthermore, the lines have to be bright enough to be clearly visible in the image. Part of the challenge in building this type of 3D-camera is constructing such stripe projector in a small package.

1.3 Outline

This thesis is divided into seven chapters. This first chapter presents the motivation for creating such a device, and describes the general concept. Chapter 2, **Active Shape Acquisition Techniques**, provides a review of the various technologies used for 3D shape acquisition, discussing their associated advantages and shortcomings. In Chapter3, **Background**, the reader is provided with the fundamental concepts that are necessary to understand the details of our design. The 3D-camera, which was built as part of this project, is described in detail in Chapter 4, **Hardware**. The chapter focuses on the constructive aspects, while explaining the design decisions made along the project. The software responsible for the operation of the 3D-camera together with the image processing software is presented in Chapter 5. Chapter 6, entitled **Results**, discusses and illustrates some of the results obtained with our 3D-camera prototype. Conclusions and directions for future work make the subject of Chapter 7.

2 Active Shape Acquisition Techniques

There are numerous technologies for active shape acquisition (**Figure 2-1**). Some purely mechanical, others are based on electromagnetic or sound waves, the most common being the laser-based ones. The reason why so many different methods have been developed and are currently being used, is the existence of shortcomings for each of them. No single method can outperform the others in all the applications; on the other hand, each method has certain advantages given a specific application. This chapter gives an overview of these methods.

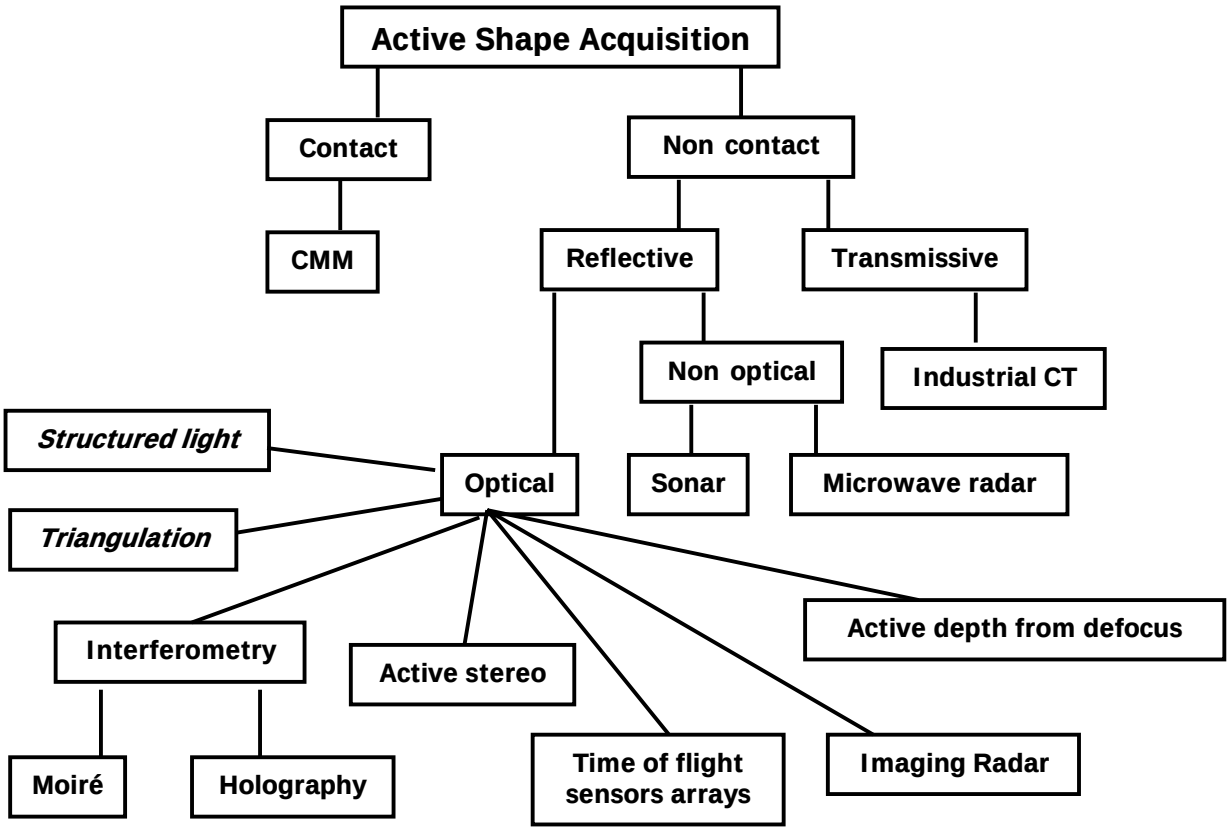


Figure 2-1: Active Shape Acquisition Taxonomy (adapted from [Curless 97])

2.1 Contact-based Shape Acquisition

Contact-based methods rely on a mechanical contact between a touch probe and the surface of the object to be digitized. The probe needs to cover the entire surface of the object in order

to guarantee a complete measurement. The machines used for this purpose, called Coordinate Measuring Machines (CMM), are extremely precise, however slow, require human operator, and are usually costly. Due to the required physical contact, they cannot be used on very fragile or soft materials as well as on rare or precious objects, such as artwork. A reputable manufacturer of CMMs, from which several models are available, is "Brown and Sharpe" [B&S]

2.2 Non-contact-based Shape Acquisition

These methods use some form of wave energy, either light, X-rays, or ultrasounds. The wave energy is recorded and analyzed after it has interacted in some way with the object to be characterized.

2.2.1 Transmissive

These methods are based on radiant energy attenuation as a function of material density and thickness. They are used for homogeneous materials, usually to detect internal cavities and imperfections, as well as in medicine for anatomic investigations [Reichl]. An example of this technology is industrial computer tomography (CT), which uses X-rays to bombard an object and measures the radiation passing through the object along various lines of sight. This information is then used to reconstruct the object's shape using back projection.

The particularity of this method is the capability to detect internal cavities. Among disadvantages are the high cost, susceptibility to material density, and potential hazard because of the radioactivity.

2.2.2 Reflective

The reflective methods are the most numerous and diverse. They measure the wave radiation that bounces back from the surface of an object. This radiation can be either light energy (I.R., visible, or UV) or some other form of radiation (ultra-sounds, microwaves).

2.2.2.1 Non-optical methods

A non-optical method can use either electromagnetic or mechanical wave energy. In either case, it is necessary that the wave is directional and the wavelength needs to be shorter than the objects details that need to be resolved.

Sonar is an ultrasonic technique that measures the time that takes to a pulse to travel to and from the surface whose range is to be determined. Similarly, microwaves radars measure the time for a pulse of electromagnetic radiation. Usually, these methods are used for long-range measurements and approximate shape estimation. Example of such device can be found at OmniTech AS [OmniTech].

2.2.2.2 Optical methods

Optical methods provide very good results in most of the situations. They rely on the relative optical uniformity of the material's surface. An object that is transparent, specular, very porous, or which diffuses the light into the surface may not be suited for any of the optical methods.

Imaging radar is similar to the ***sonar*** or ***microwave radar*** in the sense that it measures the time of flight of a laser beam. For large objects, it has proven very accurate; for objects under one meter it requires extremely high speed circuitry, since it needs to measure time differences in order of femptoseconds. 3rdTech, Inc. [3rdTech] makes a relatively compact version of this product.

Time of flight sensor arrays are similar to the ***Imaging radar***, but employ a sensor array similar to a CCD; each of the pixels is capable to measure the time at which the light pulse arrives. If the entire object is illuminated at once with a pulse of light, the reflected light will arrive sooner from the surface regions that are closer. This is a very recent technology, proprietary to Canesta Inc. [Canesta], whose potential is still being investigated.

Interferometry (Moire or Holography) uses lasers and evaluates an interference pattern between a reference source and the reflected light, resulting in constructive and destructive interference, which manifests itself as intensity variations. Both are suited for microscopic objects, of the order of the laser wavelength. Moire interferometry can exhibit phase discrimination problems when dealing with non-smooth surfaces. Holographic methods have accuracy of a fraction of wavelength. Photomechnics Inc. [Photomechanics] produces a few models which are based on Moire interferometry, while Nanofocus Inc. [Nanofocus] is a manufacturer of Holgraphy based 3D shape acquisition devices.

Active depth from defocus estimates the image blur to determine the distance, assuming that the blur is proportional to the distance from the focal plane. This method requires either a surface texture (passive) or a pattern to be projected on the object (active). For the passive approach, the quality of the shape extraction is dependent on the sharpness of the texture. For the active method the accuracy is typically moderate.

Active stereo uses two or more cameras looking at the scene while the depth is calculated based on epipolar geometry. This method is used in robotics, emulating the human vision [Fangeras 93].

Triangulation uses a camera and a focused light source (*e.g.,* laser), whose position is well defined, and which is projected on the object. The distance is calculated based on the position of the projected laser trace in the image. This approach results into highly accurate measurements, typically in the sub-pixel range, given that very precise calibration is performed. Optical triangulation is currently the most popular optical range finding method. The following is a non-exhaustive list of triangulation based systems manufacturers:

Laser Design Inc.:	http://www.laserdesign.com/
Digibotics:	http://www.digibotics.com/
3D Digital Corp.:	http://www.3ddigitalcorp.com/
Arius 3D Inc.:	http://www.arius3d.com
Cyberware Inc.:	http://www.cyberware.com/
Minolta:	http://www.minoltausa.com/vivid/default.asp

A special case of triangulation are the structured light methods, in which some light pattern, either color or shape coded, is projected on the object to be digitized. The object is reconstructed after finding the elements of the structured light in the 2D image. This approach provides quick acquisition and, in principle, utilizes hardware that can be made lightweight and cost effective. Eyetronics and 3DMetrics Inc. have products based on structured light. Both use a xenon flash lamp to project a pattern onto the object to be digitized.

Eyetronics *ShapeCam* [Eyetronics] projects a fine grid pattern, which is recorded with a digital camera, and the model is reconstructed using their proprietary software. The device can acquire shape of objects ranging between 4 inches to 8 feet in height, according to the product's specifications. The texture can be recovered by filtering the grid from the first image,

or by taking a second image, without the grid. This device consists of a SLR-size camera and a small pattern projector, mounted on a relatively large frame (about 2ft x 1ft). The distance between the camera and the projector is adjustable, in order to accommodate various object sizes.

A similar concept is used by 3D Metrics [3D Metrics], however this uses a color-coded pattern, instead of a monochrome grid. *3DFlash!cam*, which is the name of the product, incorporates one camera for the pattern acquisition, one camera for texture, and the pattern projector in one enclosure, having a volume of about one cubic foot. The acquisition time can be as low as 5ms. Objects ranging between 3 inches to 3 feet across can be digitized with high accuracy, according to the product's specifications.

In November 1999, Minolta introduced the *3D 1500* model [Minolta], a truly portable 3D acquisition color camera. Weighting only 600g and measuring 240 x 77 x 76mm, this device could acquire 200 3D images using eight AA alkaline batteries. As the other tow 3D cameras mentioned earlier, the Minolta model uses a proprietary flash-device for pattern projection. The operator requires some training and the images could only be taken only in low ("subdue") ambient illumination. Furthermore, after acquisition, it was necessary to manually "clean" the images. Priced at $4000, it did not enjoy much popularity, being now discontinued from production.

2.3 Summary

In this chapter, various methods for active shape acquisition were mentioned, with references to available industrial products. For each of the methods, the most important advantages and limitations were enumerated. There are contact and non-contact based methods. The non-contact-based ones can be divided into transmissive and reflective. The later can be further subdivided in optical and non-optical methods. Optical methods are the most numerous due to the many applications for which they are fit. Triangulation methods have very good accuracy for ranges between few inches to tens of feet.

This page was left blank intentionally.

3 Background

The purpose of this chapter is to familiarize the reader with some basic concepts that form the foundation of this work. Some of the notions described here are fundamental Physics concepts while others are more technology related. The intent is not to define these terms in general, but to describe them in the context of this work.

3.1 Speckle Effect

Before introducing this notion, the reader should be aware that the terms "specular" and "speckle" refer to different phenomena. While the former applies to shiny, mirror-like surfaces, the later is characteristic to coherent light bouncing of an irregular (at the wavelength scale) surface. Coherent light is produced by lasers and can be defined in simple terms as monochromatic light (single wavelength) whose rays that form the beam are in phase. By contrast, natural ambient light is incoherent.

Further, to understand the speckle effect, one needs to be familiar with the notions of constructive and destructive interference. This phenomenon refers to spatial addition of time varying signals, as electromagnetic or mechanical waves, such as light and sound respectively. The constructive interference occurs when two waves having the same wavelength add up in phase resulting in a wave, which has larger amplitude than the original ones. Destructive interference takes place when the two waves have opposite phase resulting in a wave that has smaller amplitude than any of the two. When this phenomenon occurs in the case of coherent light, it manifests itself in light intensity variation: brighter spots correspond to constructive interference, while dark ones correspond to destructive interference.

When the coherent beam of light from a laser encounters an irregular surface, the light is scattered in various directions, due to the surface's microscopic irregularities. At various points in space, the rays interfere creating a large number of bright and dark spots (**Figure 3-1**). When visible laser is used, the speckle pattern is visible due to the scattered rays that arrive at the retina. For an imaging system such as an electronic camera, the speckle is formed on the sensor's plane.

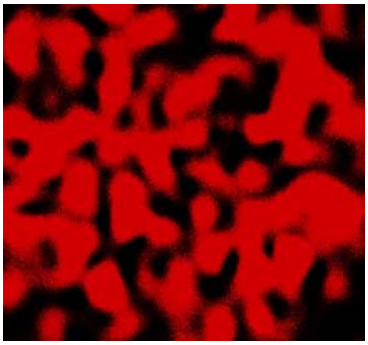


Figure 3-1: Laser speckle patterns

The speckle pattern is an undesired effect for laser-based triangulation, since it affects accuracy. Normally, a focused laser spot is characterized by a Gaussian energy distribution, meaning that starting from the outer edge of a spot and moving diametrically across it, the light intensity will follow a Gaussian profile. Considering this model, the center of the spot can be found by looking for the highest intensity area. Due to the speckle effect, the apparent intensity will be randomly distributed. As the spot moves over the surface of the object, the speckle pattern varies according to the irregularities.

In high-precision laser-based triangulation systems, speckle is the limiting factor for accuracy. Although no solution exists to this problem (this is a fundamental Physical phenomenon), there are ways to reduce it. For instance, by making the camera's aperture larger (reducing the F-number), the amount of collected rays that arrive on an elementary area of the photo sensor (pixel) is larger and because of averaging, the speckle size becomes smaller. When the speckle is several times smaller than the pixel size, it is integrated, resulting in a more uniform pattern. Some lasers have lower coherence than others; for instance, lasers produced by semiconductors are generally less coherent than gas-generated ones. There are laser diodes, however, that can produce highly coherent lasers.

3.2 Laser Beam Deflection

A laser beam can be deflected by a mirror, as any other beam of light, being subject to the law of reflection, which states that the angle between an incident beam of light and the normal to the mirror is equal to the angle between the reflected beam and the normal to the mirror. If the mirror is rotated around an axis by an angle α, then the laser beam is deflected by 2α,

relative to the angle before rotation, as illustrated in **Figure 3-2.** When the mirror oscillates around that axis, between +α and –α, the laser beam is deflected between +2α and –2α respectively. If projected onto a planar surface, it produces a spot, which moves along a linear trajectory. Furthermore, if the frequency of the oscillation is above 15Hz or so, the moving spot cannot be distinguished by the human eye, but its locus appears as a continuous line, with some flicker. As the frequency increases, the flicker becomes less noticeable until it becomes unperceivable. The light receptors of the retina are insensitive to very fast variations in light intensity, although they are capable to average out in time the photon energy. Similarly, the sensor arrays used in digital cameras average out the light over some integration time, such that a moving laser spot appears as a line. If the laser spot moves rapidly with respect to the integration time of the camera, it may be captured by the camera as line patterns.

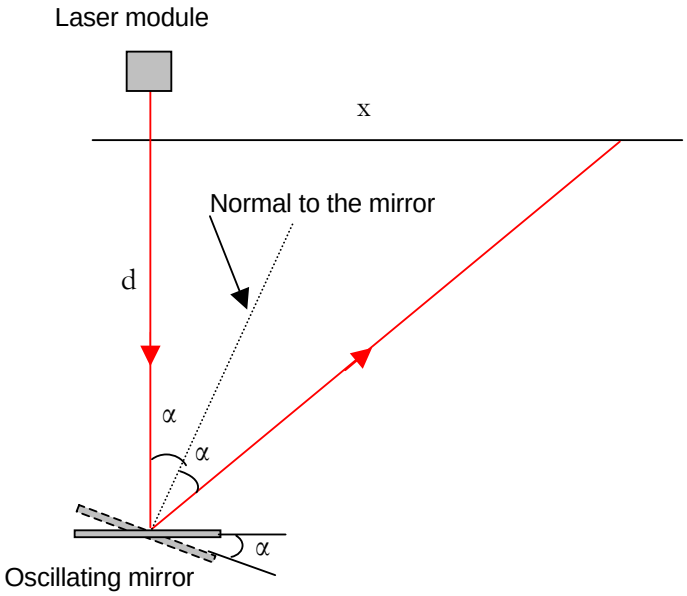


Figure 3-2: Laser beam deflection

The beam produced by an oscillating mirror can be deflected onto a second mirror, having its axis of oscillation perpendicular to the first one. The motion of the laser beam, deflected by two oscillating mirrors, having their axes mutually orthogonal results in Lissajous patterns [Lissajous], and will be described further in this chapter.

The velocity of the laser beam along each of the two perpendicular directions is a function of frequency, amplitude, and the waveform used for driving each of the two scan elements. To produce a raster-like pattern with horizontal lines, it is necessary to have a high ratio between the frequency of the horizontal scanning mirror and that of the vertical-scanning mirror. This ratio is 400:1 in our prototype, the actual values being 2020Hz and 6Hz, for practical reasons. For the low frequency mirror, one half period of the oscillation has to be similar to the integration time of the camera (30ms maximum), since one frame has to be completed in this time interval. For the high frequency element, the constrain is the power consumption needed to move the oscillating element; the highest efficiency is achieved when it is driven at a frequency equal to its natural mechanical resonance. This statement is fundamentally true for any resonating system. When driven at resonance the mirror follows a sinusoidal motion. The laser spot, being the projection of the laser beam onto a surface, has an instantaneous velocity that depends on the angle at which the beam projects onto that surface.

Considering the case of one mirror only, its sinusoidal motion can be described as:

$$\alpha = \mathrm{A}\sin\omega t$$
$$\omega = 2\pi\nu \qquad \textbf{(3-1)}$$

where α is the instantaneous angle of the mirror, A is the maximum rotation angle, ν is the oscillation frequency, and ω is the angular velocity.

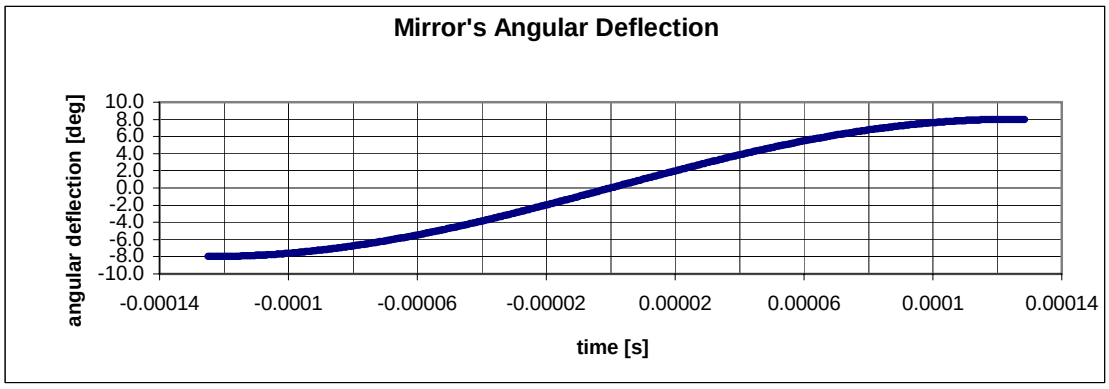


Figure 3-3: Mirror's deflection

Figure 3-3 illustrates the angular motion of the high frequency mirror for a half period (one sweep from left to right), for A = 8° and ν = 2KHz. The speed of the laser spot projected onto a flat surface placed at 0.25 m from the mirror and normal to the laser beam, when is at rest, is shown in **Figure 3-4.**

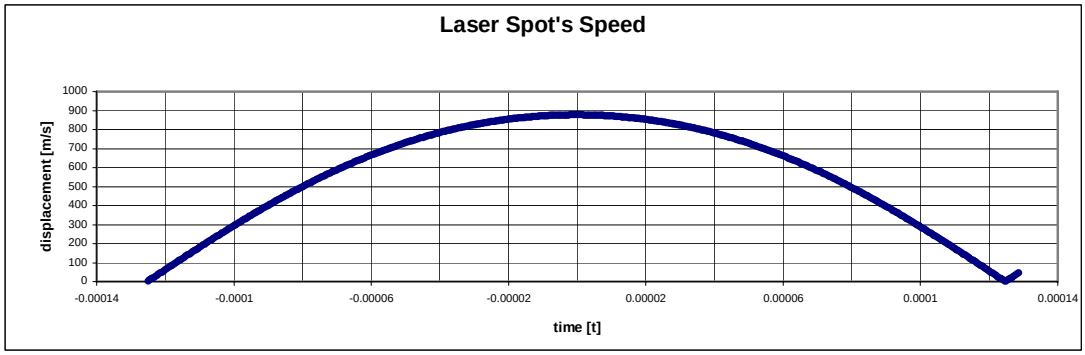


Figure 3-4: Laser's spot speed

The spot speed, as illustrated in **Figure 3-4**, was calculated based on the geometry shown in **Figure 3-2**. The equation describing the position of the laser spot onto a flat surface as described above is:

$$x = d\tan 2\alpha$$
$$x = d\tan(2A\sin\omega t) \qquad \textbf{(3-2)}$$

According to **Figure 3-4**, if a scanning mirror projects a laser beam onto a flat surface and a camera takes a snapshot as the spot travels during one half period, the image will show a line exhibiting uneven intensity. The brightness recorded by the camera is proportional to the amount of time it takes to the image of the spot to travel over one pixel. In the middle of the line, where the spot speed is higher, the line is dimmer than near the extremities, where the spot slows down. It is important to note that the speed of the spot is not fully characterized by the sinusoidal motion of the mirror, but is also dependent on the geometry of the surface onto which it is projected. As it can be seen in **Figure 3-5**, for a spot projected onto a flat surface, assuming a wide scan angle (A=22°), the speed of the spot becomes almost constant in the middle, due to the tangent factor (Equation 3-2).

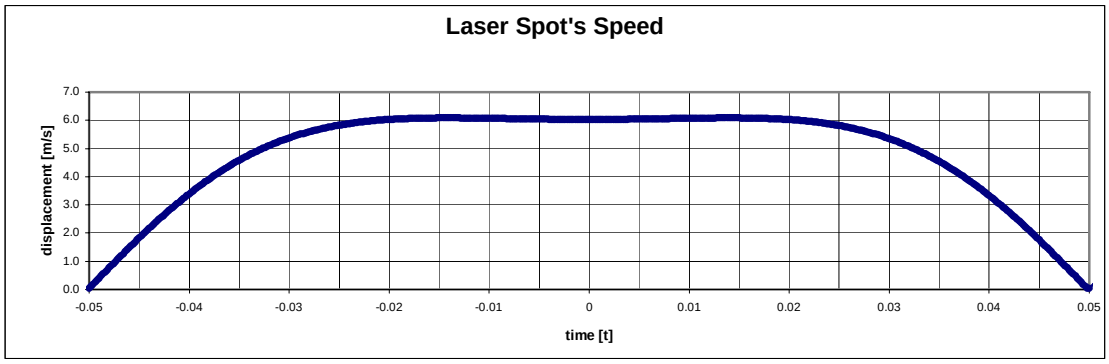


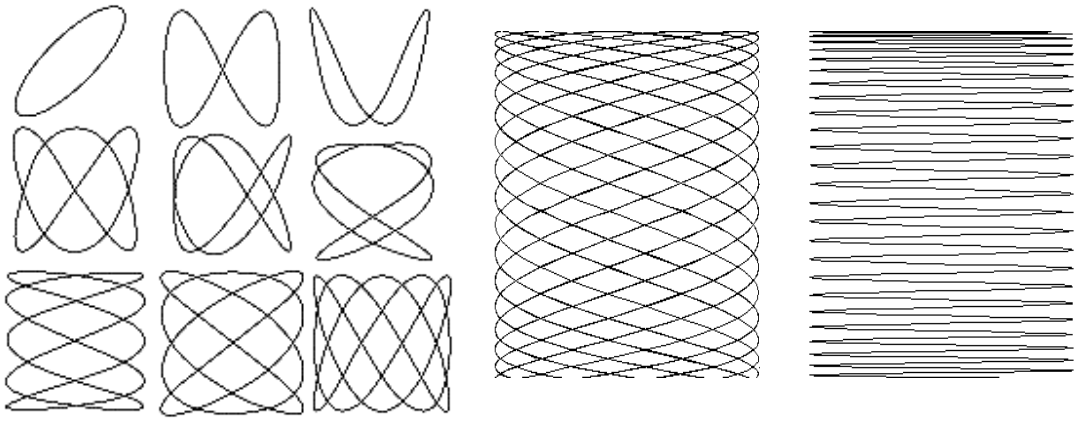
Figure 3-5: Wide scan angle effect (A=22deg.)

3.3 Lissajous Patterns

Lissajous Patterns are trajectory lines produced by the composition of two mutually orthogonal harmonic oscillatory motions. These patterns can be viewed on an oscilloscope screen by applying one sinusoidal signal to each of the vertical and horizontal channels.

Depending on the frequency and amplitude ratios and on the phase difference of the two oscillations, the patterns can have different looks. When the two frequencies are equal and the phase is arbitrary the pattern is elliptical; it becomes a circle when the amplitudes are equal while the two oscillations are 90° out of phase and is a diagonal line (with a slope as a function of amplitude) when the two oscillations are in phase. When the frequencies ratio is very high the pattern looks like a zigzagged line. **Figure 3-6** illustrates various Lissajous patterns, obtained synthetically on the computer. The jagged look of the lines are artifacts; these lines are normally smooth.

The patterns can vary significantly depending on the parameters mentioned above as shown in **Figure 3-6** (a). As the ratio of the two frequencies increases, the pattern will appear eventually as a zigzagged line (**Figure 3-6** (c) with angles between every other lines approaching zero. If every other line is removed from a Lissajous pattern with a high frequencies ratio, one can imagine that in the middle portion along vertical direction, it will approximate a raster with parallel lines.



(a) various frequency ratios **(b) 4:1 ratio** **(c) 60:1 ratio**

Figure 3-6: Lissajous Patterns

The locus of a point, which follows the equations for two orthogonal oscillatory motions, can be analytically described as:

$$\begin{aligned} x &= A_1 \sin \omega_1 t \qquad \omega_1 = 2\pi\nu_1 \\ y &= A_2 \sin \omega_2 t \qquad \omega_2 = 2\pi\nu_2 \end{aligned} \qquad \textbf{(3-3)}$$

where *(x,y)* are the Cartesian coordinates of the point at moment *t*, A_1 and A_2 the amplitudes of the motions, and ω_1 and ω_2 the two angular velocities. The two equations on the right (Equation 3-3) show the relation between frequency (ν) and angular velocity (ω).

The Lissajous figures can be produced using a laser and two mirrors oscillating around normal axes. If a laser beam bounces from one mirror to the other and then to a flat surface, the laser spot on the surface will describe such patterns.

3.4 Camera-Laser-Raster Geometry

Triangulation can be used to recover the depth information from a single image. **Figure 3-7** illustrates the underlying geometry for a single point, produced by the intersection of a laser beam with the object's surface. The relative position between the raster generator and the camera shown in the figure are consistent with the actual construction of our prototype. The figure shows the simplified geometry, for an ideal pinhole camera, since this model is

sufficiently accurate with respect to our construction. The aperture of the actual camera corresponds to the position of the pinhole of the ideal camera. The line normal to the sensor plane, passing through the pinhole represents the optical axis. The principal point is the intersection of optical axis with imaging plane and ideally should be in the center of the image. In fact, due to the construction tolerances, the principal point is usually slightly off-center. The actual location of the principal point can be determined by camera calibration, which is a standard procedure and does not make the subject of this thesis. For our camera, however, we assume that the principal point is close enough to the center of the image.

The raster generator sweeps the laser in a plane perpendicular to the plane of the page. A number of sweeps, each at a different angle θ, is needed to cover the entire surface of the object with laser lines. The distance between the point C (the position of the laser spot where the laser beam intercepts the surface of the object), and the optical axis of the camera (the normal to the image plane passing through the pinhole aperture) is X. The laser spot is visible to the camera and is imaged onto the sensor element at position x with respect to the principal point. The focal distance of the camera lens being *f*, perspective projection gives:

$$\frac{x}{X} = \frac{f}{Z} \qquad \textbf{(3-4)}$$

From the ABC triangle in **Figure 3-7,** the following relation can be written:

$$\tan\theta = \frac{d + X}{Z + a} \qquad \textbf{(3-5)}$$

After eliminating X and solving for Z, using equations (3-4) and (3-5), the formula for depth of point C with respect to the camera is given by:

$$Z = \frac{a\tan\theta - d}{\frac{x}{f} - \tan\theta} \qquad \textbf{(3-6)}$$

Since *a, d,* and *f* are intrinsic parameters of the system (fixed by construction), the depth Z can be uniquely determined from any given *(θ, x)* pair.

3.5 Summary

After reading this chapter, one should have the basic notions required to understand the work presented in this thesis. The laser speckle effect, which is a phenomenon associated with coherent, *i.e.*, laser light, affects the accuracy of the reconstruction, and requires the image processing software to be designed accordingly. The mechanical and optical aspects of the oscillating mirror were covered, explaining how the laser beam manifests itself as it bounces off the mirrors. The desired raster pattern is a particular case of the Lissajous patterns. A simplified geometry of the camera-laser pattern generator was presented and a formula for depth calculation was derived.

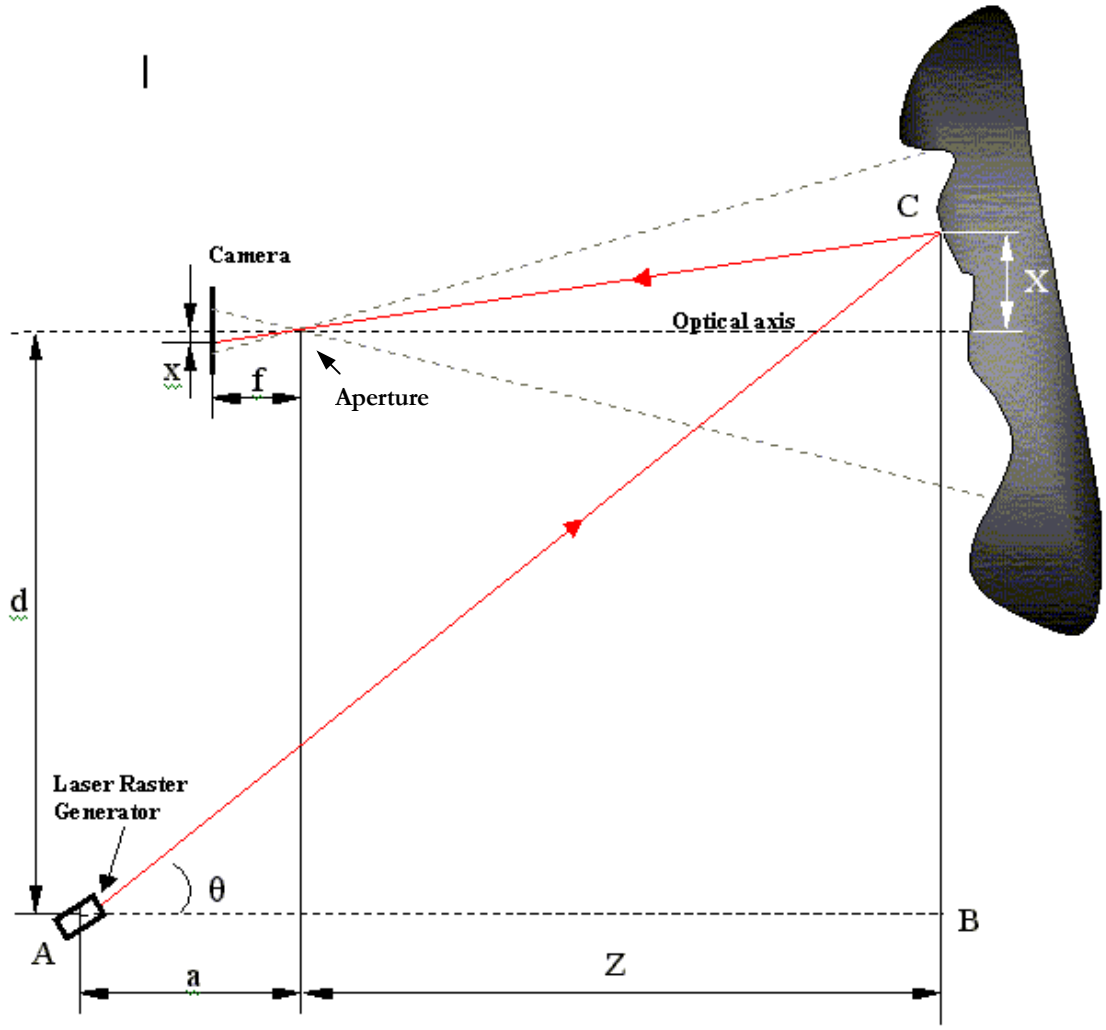


Figure 3-7: Triangulation

4 Hardware

This chapter elaborates on the equipment that was constructed to produce the images needed for 3D reconstruction. There are two major components, with distinct functions, which are synchronized for proper image capture: the camera and the raster generator. Each will be described in detail in the following paragraphs. The last section of this chapter presents a simple method used to calibrate the system.

4.1 System Description and Use

The electronic camera utilized is a CCD-based monochrome VGA (640X480), capable to produce 256 levels of gray. The raster generator module constructed with modified scan elements used in some Symbol products The electronic circuit for the raster generator was built with off the shelf components, some purchased at Radio Shack. **Figure 4-1** shows the completed 3D-camera prototype.

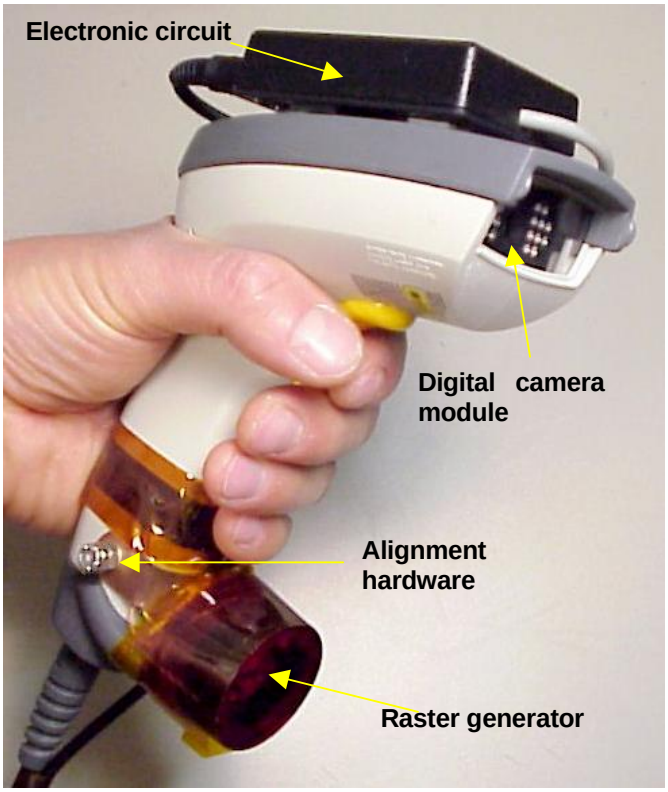


Figure 4-1: 3D-camera prototype

The CCD module is protected by a clear plastic window, inside the housing which extends with a handle and a cable used for power and data communication. This cable is terminated with a DB9 connector which plugs into a computer serial port. The DB9 connector has incorporated a small two contact standard connector in which a 5V power supply should be plugged. The images acquired by the camera are compressed as jpeg pictures and sent over the serial communication to the computer. The stereo jack connects to the headphone output of the PC sound card.

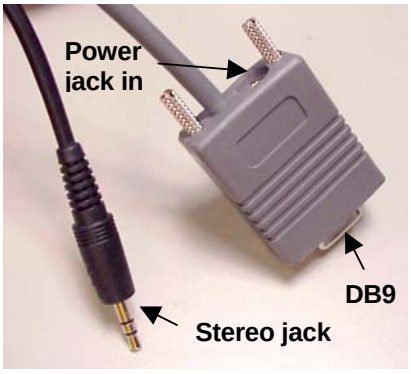


Figure 4-2: Connectors

The cylindrical module, located at the bottom of the handle is the raster generator, comprised of two scan elements and a laser diode, the whole assembly being protected by a red plastic window. The electronics which control the laser module are located in the black box, on top of the camera. This circuitry could be easily miniaturized and incorporated inside the barcode reader housing. The black cable running from the black box connects to the PC sound card output (headphones).

4.2 Digital Camera

The camera is the device that acquires and stores the images that are being utilized for laser lines extraction and texture capture, if more than the wire frame is to be reconstructed. The captured images are uploaded to a notebook computer for processing. The quality of the images has a direct impact on the accuracy of the reconstruction.

The 3D-camera prototype described in this thesis was built using Symbol Technologies Inc. VS4000 CCD-based barcode reader. Any commercial –grade camera of equal or better quality could be utilized, given that the integration (*i.e.,* exposure) time can be controlled and start of integration can be triggered electronically. It is important that the beginning of the integration is synchronized with the raster generation for the proper imaging of the projected laser pattern. The camera should start the integration when the low frequency (vertical) mirror has traveled ¼ down from the extreme top position. This way, the portion where the mirror accelerates, until it reaches its nominal velocity, does not become part of the image; otherwise, the spacing between the horizontal laser lines would not be uniform, resulting in distortions of the reconstructed 3D shape.

Constructively, the camera consists of an optical system, a sensor array, an electronic control circuit, and a digital storage block. These elements are typically integral part of any electronic camera. The following few paragraphs describe their functionality in the context of the 3D-camera.

4.2.1 The Camera Optical System

The optical system projects the image of a scene onto the sensor array. Ideally, as in the case of a pinhole camera, the image should be an accurate representation of the scene, at a scaling factor and limited by the field of view. Such image is governed by the geometry of perspective projection, which is well defined and can be easily modeled. In practice, however, a pinhole camera cannot be implemented in most cases, since it requires very high light intensity to produce a bright enough image. Instead, various lens-based collimators are being utilized, which have the advantage that can capture a large amount of light rays to produce images. Such device, known as objective, or improperly lens, consists of one or more lenses and an aperture, arranged in a precise configuration. A complex design is often used to eliminate aberrations, which are exceptions from the "ideal" optical geometry and occur due to the wide range of visible light wavelengths. Since our system uses a red laser with a well-defined wavelength, a simple lens system is appropriate. Obviously, a bad quality lens can introduce distortions; however, due to the large number of mid to low cost consumer-grade cameras available on the market today, reasonable quality lenses are common.

When a collimating lens is utilized in an optical system, two terms need to be defined along: focal distance and depth of focus. The focal distance is a measure characteristic to a lens, defining how far from the lens a beam of parallel rays converges to a point. This is a rather simplistic definition, but should be adequate for the purpose of this section. Typically for a film camera, the focal length of the lens is in the neighborhood of a few centimeters, the most common being the 35mm lens. The camera utilized in our prototype has a focal length of 8mm, which allowed miniaturization.

To form an image of an object placed at distance significantly larger than the focal length, the focal plane (the plane on which the image is formed) has to be slightly further from the lens than the focal length. If the sensor array (CCD element) is placed at any other position than the focal plane, the image formed will be blurred. Reciprocally, if the distance between the lens and the sensor array is fixed to form the image of an object at a given distance, when the distance from the object is changed its image becomes blurred. The range that the object can be moved without significantly blurring the image is referred to as the depth of focus.

For a fixed lens camera, the depth of focus determines the working range of the camera. An aperture used together with a lens increases the camera's depth of focus, however reducing the brightness of the image: the smaller the aperture, the larger depth of focus and the less bright the image.

4.2.2 The Sensor Array

The sensor array converts light into electricity. The image, which is formed by the optical system, is projected on the surface of the sensor, which is constructively a 2D array of light sensitive cells. For a monochrome system, as the one utilized in our design, all cells are identical. The color versions use red, green, and blue filters on top of the cells, alternating in an orderly fashion. Each of the cells converts the light received into an electrical charge, proportional with the amount of light received. The charge is then converted to a voltage level, which is output to the associated electronic circuit.

The light energy is integrated over time by each of the cells; therefore, the longer the sensor is exposed to light the higher the charge built-up, to the point of saturation. Saturation occurs when the maximum charge that a cell can hold has been reached. In the case of CCD arrays, if

the light continues to exist over the cell after the maximum charge capacity has been reached, the charge can "spill over" into neighboring cells, creating what is known as *bloom*. These conditions will result into distorted images, with loss of information.

When using a laser to illuminate an specular reflection may have a substantially higher intensity than the rest of the image created by scattered light and can produce saturation, causing what has been describe earlier. To avoid this, objects with non-reflective surfaces (mate) are preferably used in digitization procedures.

4.2.3 Electronic Control and Digital Storage

The electronic control is responsible for generating various timing and conditioning signals necessary for the proper operation of the sensor element. The integration (exposure) time and the gain (amplification) need control as well, either manual or automatic. These controls are accessible through the software control interface, an application that runs on a portable computer, which controls the camera in the current prototype. Section **5.2** describes this software.

The voltage levels, corresponding to the light intensity received by the sensor array, are converted further into digital levels. The camera utilized here is capable of producing 256 distinct values or shades of gray. Each pixel in the image is assigned a number form 0 to 255 corresponding to the light intensity level. This information is stored in the camera's memory and represents the digital image.

4.3 Raster Generator

A raster is a pattern of horizontal scanning lines, as shown in **Figure 4-3** and explained in sections **3.2** and **3.3**. The raster generator is responsible for producing these laser traces, which are used for the 3D reconstruction of the scanned objects. The quality of the raster and its consistent behavior are very important factors for the accuracy of the reconstruction. This section provides a detailed functional description and analyzes all the constructive elements of the raster generator. It further discusses some essential aspects for a robust opto-mechanical and electronic design.

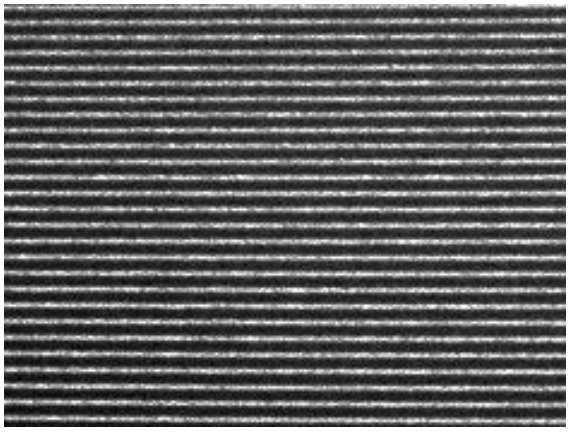


Figure 4-3: Raster projected on a flat surface

4.3.1 Functional description

The raster is produced while a narrow laser beam bounces onto two mirrors, which oscillate around axes perpendicular to each other (**Figure 4-4**). The laser bounces off a mirror with vertical oscillation axis (x-mirror), which oscillates at around 2KHz, then it encounters the second mirror (y-mirror), which oscillates at 6Hz and having its axis perpendicular on the first one. The motions of the two mirrors are quasi-sinusoidal, therefore producing Lissajous patterns. In particular, this type of motion composition results in a zigzagged line (**Figure 3-6 (c)**, since the high frequency oscillation is much faster than the low frequency one.

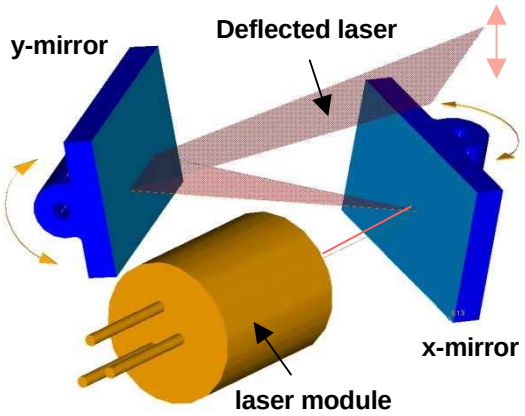


Figure 4-4: Raster generator diagram

Figure 4-3 shows a raster pattern projected onto a flat surface. Ideally, the raster should look like a dense ensemble of parallel lines, when projected onto a plane normal to the raster generator's optical axis (the direction of the laser beam when the two mirrors are in the rest positions). Unfortunately, the lines generated are slightly curved near the extremities, as the x-mirror decelerates, stops and accelerates in the opposite direction. For the lines to be straight, it is necessary that the ratio of the two velocities be constant. In practice, this is not possible during the entire motion, for a few reasons discussed in section 4.2.2. The proposed design presents a solution that attempts to produce a raster that approximates the ideal case.

The laser beam in its motion describes a series of surfaces that intersect at the point where the laser is generated. These surfaces can be approximated by planes fanning out from an origin at equal angles. This approximation significantly simplifies the computations, when compared to the case in which the surfaces are not planar and have variable angles between them.

4.3.2 Constructive Description

The laser raster generator consists of three main components mounted in a particular configuration illustrated in **Figure 4-5**. They are:

- Laser module
- High frequency oscillating element (x-mirror)
- Low frequency oscillating element (y-mirror)

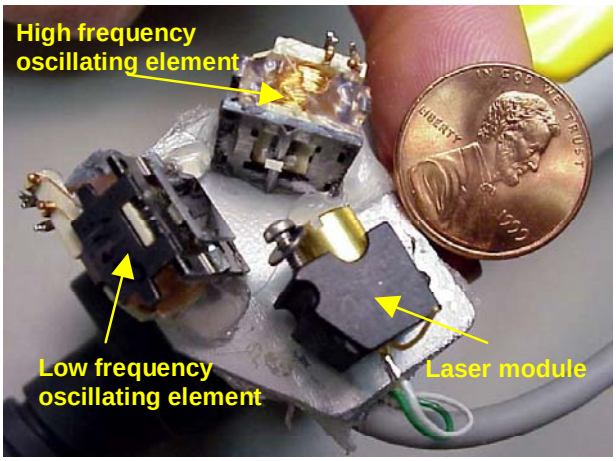


Figure 4-5: The Raster Generator

Next, I describe each of these elements, explaining the important aspects associated with the design of the 3D-camera.

4.3.2.1 The Laser Module

This section discusses the laser module utilized in our construction, in practical terms. There is much theory behind laser diodes dealing with optical, mechanical, and electrical aspects. The interested reader can find a good coverage of this topic at "Sam's Laser FAQ" [Sam's FAQ].

Our laser beam is produced by a laser module together with a laser control circuit. The raster's definition is strongly affected by the focusing and brightness of the laser spot. The laser module (**Figure 4-6**) consists of a laser diode and a focusing assembly, which is needed to collimate the light into a narrow beam.

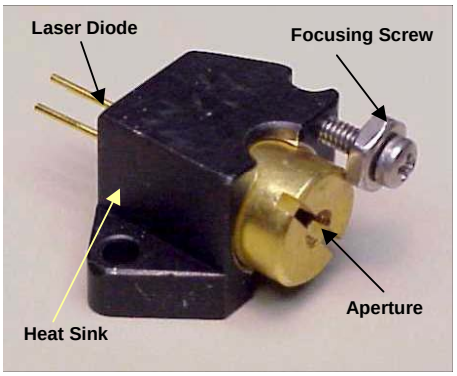


Figure 4-6: Laser module

Figure 4-7 illustrates the focusing characteristics of the laser module. Normally, the light emitted by a raw laser diode diverges in an elliptical cone at angles typically 10 and 30 degrees for the two axes. The light intensity along a cross-section of the pattern has a Gaussian distribution, fading out away from the center. This characteristic is maintained after the light is collimated by a plano-convex lens. Along with a lens, an aperture is used to further form the beam and modify its wave front shape and consequently, its propagation characteristics. When an aperture is utilized, the laser beam diameter is reduced overall, but due to diffraction, the beam waist becomes larger. It is important mentioning that the laser spot cannot be made

arbitrarily small over any distance, because of diffraction, which affects the laser beam propagation.

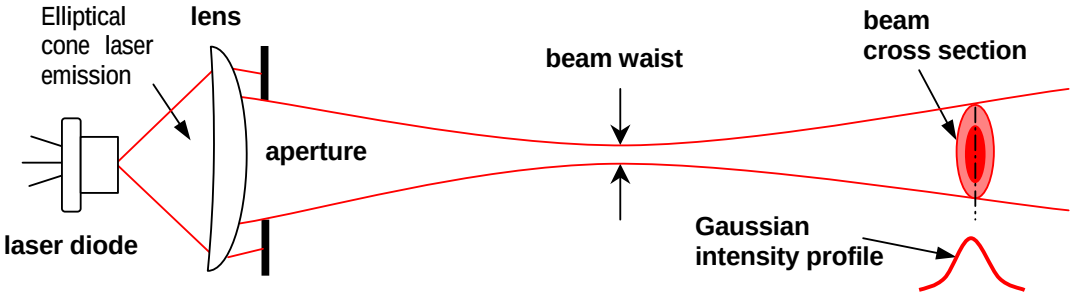


Figure 4-7: Laser module focusing

It is desired for the lines of the raster, when imaged, to be one pixel wide and bright enough to be distinguished from illumination produced by the ambient light. The fact that the laser spot intensity has a Gaussian distribution is used by the image processing software (**section 5.1**) to discern the laser lines from other bright elements in the image. Ideally, the laser line width should span over three pixels including the tails of the bell-shaped intensity profile. The laser can be focused into a small spot using a plano-convex lens with the focal length of 6 mm and an aperture of 1mm diameter. Experimentally, I found that the laser throughput power should be at least 10mW to produce bright enough lines.

The laser light power is a function of the electrical power applied to the laser diode. The efficiency is typically below 10%, decreasing as the temperature increases. Therefore, to maintain a constant light output, it may be necessary to increase the current as the diode heats up. To avoid thermal run-away, a heat sink is mounted on the laser diode. Thermal ran-away is a phenomenon characteristic to semiconductors, when a constant voltage is applied. The current caused by the potential difference results in a thermal energy, which raises the temperature of the semiconductor. The conductivity of a semiconductor increases as the temperature raises, and more current passes through, further increasing the temperature, in a self-sustaining process. The high temperature build-up results eventually in the damage of the semiconductor device. By using a radiator mounted on the semiconductor component, the thermal energy is dissipated in the ambient air, and the component reaches a thermal equilibrium (the heat energy generated equals the dissipated heat).

To produce 20mW of raw optical power (generated by the photons' energy from the laser), which results in about 10mW throughput after the beam exits the aperture, the laser diode utilized requires about 2.7V at 80mA. If a good heat sink is utilized, the diode can be driven in constant light output mode, where the electrical current is automatically adjusted by an electronic circuit. A photodiode internal to the laser diode package monitors the light output providing a feedback for the circuit that controls the current accordingly.

Without proper heat sinking it is necessary to limit the current. In this case, the light output will slightly decrease as the laser diode's temperature increases until it reaches thermal equilibrium. Our prototype was build for experimentation mainly, so the laser had to be capable to stay on for extended time intervals. In normal applications, however, the laser need not be turned on for more than the duration of an image acquisition, which is less than 30ms. Considering a case when an image is taken every second, the duty cycle would be below 2% (the laser is on only for every other line), resulting in 2.7Vx0.08Ax2%=0.004W average power. For such small power consumption, no heat sink is necessary, but a current limitation is recommended, to protect the laser diode when the ambient temperature is high. Thus, a constant current source was implemented with an adjustable voltage regulator integrated circuit.

In order to obtain the desirable raster pattern, the driver circuit should be able to turn off the laser for every other line in the raster. This is necessary, to eliminate the retrace and obtain a pattern with quasi-parallel lines. To accomplish that, a transistor is used to shunt the laser diode. The transistor is commanded by a circuit synchronized with the high frequency oscillating element driver.

4.3.2.2 High Frequency Scanning Element

The high frequency element shown in **Figure 4-8** is a mechanical resonating device actuated by an electromagnetic field. Constructively, it consists of a coil with an iron core, an elastic steel taut band, a magnet, and a mirror. The mirror and the magnet are attached at the middle of the taut band, which is mounted together with the coil. When an alternative current is applied to the coil, the electromagnetic field acts upon the magnet, forcing the band together with the mirror to twist accordingly. The direction of the spin is determined by the

polarization of the magnet and the direction of the current through the coil. As the current changes direction over 2000 times a second, the mirror-magnet assembly follows accordingly.

In order to achieve a frequency in the order of 2KHz, with the total amplitude around 15 degrees, it is necessary to keep the masses of the mirror and magnet small and make the band stiff. Also, it is most efficient to drive the oscillating element at resonance, *i.e.*, the natural resonating frequency of the mirror-magnet-band assembly should be equal to the frequency at which we wish to operate. In this mode, the element produces a sinusoidal oscillation. A total of 15 degrees of mechanical rotation of the mirror produces twice as much (30 degrees) deflection of the laser beam. This angle has been chosen to be similar to the horizontal field of view angle of the camera.

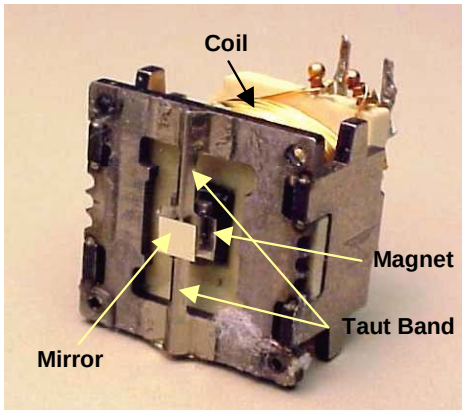


Figure 4-8: High frequency scanning element

The apparent brightness of the laser lines is a function of laser's intensity and laser's spot velocity. The 2KHz frequency was picked to produce lines that are bright enough in the images based on the laser used (~10mW throughput) and the camera's sensitivity. Higher frequency would result in dimmer laser lines.

4.3.2.3 Low Frequency Scanning Element

The low frequency element (**Figure 4-9**) is similar in construction to the high frequency one, but has a much larger mirror. This is necessary since the first element, deflects the laser beam in a fan-like pattern, which needs to be accommodated by the second element's mirror (see illustration in **Figure 4-4**).

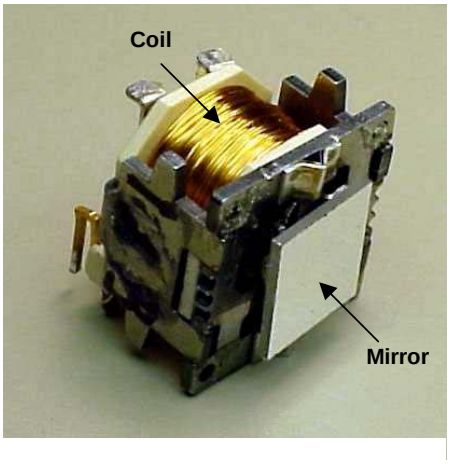


Figure 4-9: Low frequency scanning element

Unlike the fast element, the low frequency one should not operate at resonance, since at resonance the motion would be sinusoidal. Instead, we desire to produce a linear motion; therefore, the element needs to be driven with a triangular waveform. Experimentally, I found that the mirror's angular deflection is proportional to the current through the coil, for angles up to about 15 degrees one way. This is illustrated in the graph shown in **Figure 4-10**. This mechanical angle is sufficient to produce an optical deflection of 60 degrees total (see section **3.2 Laser Beam Deflection**). Given that our camera has a vertical field of view of 20 degrees, the maximum amplitude of the low frequency mirror is more than enough.

The low frequency element was constructed to have the natural mechanical resonance at around 300 Hz, although the element is driven at 6 Hz. The high intrinsic resonating frequency equates into a stiff spring with respect to the oscillating mass, making it less susceptible to mechanical shocks and externally induced vibrations. The masses of the mirror and magnet can be significantly larger than for the high frequency element. Therefore, a more powerful magnet is utilized to produce more torque. Consequently, the electromagnetic field generated by the coil does not need to be as strong as the field for the high frequency element to actuate the device, resulting in less electrical power consumption. The element can be damped to further reduce the susceptibility to unwanted vibrations that may affect the uniformity of the laser lines.

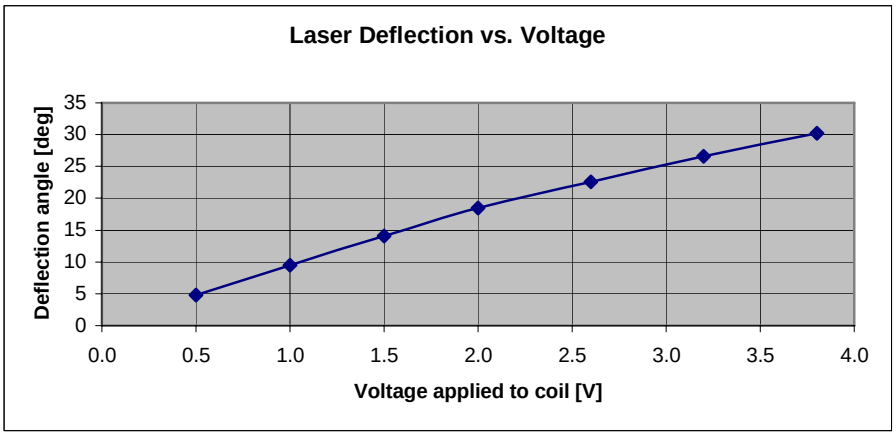


Figure 4-10: Low frequency scanning element deflection

To produce a raster with equally spaced lines (in fact, equal angles between the laser planes) the angular velocity of the mirror has to be constant throughout the entire range of motion. This is not possible since the mirror and the magnet make a large mass, which needs to be decelerated/accelerated when the mirror changes direction. It is possible however, to linearise the motion away from the ends of the travel (*e.g.* in the middle portion of the motion), where the velocity is uniform (**Figure 4-10**). To obtain such result, the camera's shutter should be opened only during the linear portion of the motion.

4.3.3 Electronic Control Circuit

The control circuit has the following functions:

- High frequency element driver
- Low frequency element driver
- Camera shutter synchronization
- Laser control

Constructively, the electronic circuit was realized using discrete surface mount components, which were soldered and connected by wires on a perforated circuit board as seen on **Figure 4-11**.

Figure 4-11: Electronic control circuit board

The following paragraphs address the functional description of the electronics, without further reference to the construction. **Figure 4-13** represents the electronic schematic of the circuit; the highlighted portions are functional blocks, which have distinct functional roles. Before beginning the description of these blocks, a brief description of the overall function of the circuit is will be given.

The high and low frequencies oscillating elements are being actuated by applying periodic waveforms of 2KHz and 6Hz to their respective coils. The laser is turned off every other line to eliminate half of the lines from the zigzagged pattern and generate a raster with parallel lines. When the position of the low frequency element is ¼ from the top position, traveling down, the camera shutter is opened and maintained so, until the element is at ¾ of angular position along its motion (about 30ms). This is necessary to eliminate the top and bottom portions where the y-mirror accelerates and decelerates; otherwise, the pattern would have unequal angularly spaced lines in these regions.

The waveforms necessary to drive the two elements are produced by a (notebook) PC sound card. The left and right channels of the audio output are fed respectively to the 2KHz and 6Hz electronic elements' drivers. These will boost the voltage and current, since the PC sound card cannot generate enough power to drive the two elements.

4.3.3.1 High Frequency Element Driver

The high frequency driver consists of a band-pass filter that has the 3dB cutoff points at 1.5KHz and 2.5KHz respectively. This band pass filter helps reduce the DC offset, which may be present from the sound card and eliminates the noise that may be induced in the circuit. The amplifier used is a TPA711, which is a Texas Instruments differential audio amplifier designed for portable music players.

The circuit is driven with a sinusoidal waveform from the sound card, chosen to match the mechanical resonating frequency of the high frequency element. Driving the element at resonance is necessary to maintain the power consumption relatively low. For this arrangement, the power consumption is about 0.5W RMS.

Because the velocity of the laser spot is quasi-sinusoidal, the apparent brightness in the laser pattern is not uniform, the lines appearing brighter towards the extremities. This, however, does not have any detrimental effect over the captured image, since the algorithm does not analyze the intensity variation for depth computations.

4.3.3.2 Laser Control Circuit

Following the high frequency motor driver, a synchronized circuit turns the laser on and off, depending on the mirror's movement direction. When a scan element is driven at resonance, the electrical drive signal has a 90 degrees phase shift with respect to the oscillatory motion of the element. Consequently, the zero crossings of the electrical sine wave occur when the mirror is at one the end positions (full deflection). The laser needs to be turned off when the mirror changes direction one way and then turned back on, when the mirror changes direction the other way. This eliminates the unwanted retrace lines. The circuit produces a square wave, which is synchronized with the drive sine wave. The square wave is used to gate (turn on and off) the laser.

When the frequency of the drive signal is not exactly matched to the resonance of the mechanical element, the phase delay is not 90 degrees. Therefore, the timing of the laser control is shifted with respect to the zero crossing of the sine wave drive signal. To control the delay, the drive signal is fed into the inputs of a comparator with a delay that is determined by an RC constant. By adjusting this constant, the square wave produced can be synchronized to

the stop positions of the mirror motion, regardless of the electrical-mechanical phase relation. This signal controls the base of a PNP transistor that shunts the laser diode during half cycle of the high frequency element.

In series with the laser diode is a constant current source circuit. This was designed using a LP2951 low dropout voltage regulator. The components were chosen to supply 80mA, which is necessary to produce about 20mW of raw optical power from the laser diode. A few capacitors were employed to prevent the current from overshooting during power-up.

4.3.3.3 Low Frequency Element Driver

The low frequency element driver consists of a series of components that produce a triangular waveform and the same type of amplifier as for the high frequency element. As mentioned earlier, the triangle wave is needed to produce a uniform motion of the mirror and therefore, a uniform raster.

The sound card of the PC is not capable of producing a triangular wave at 6Hz, but it can generate a sine wave. Additional circuitry is needed to transform this into a triangular signal. The first stage consists of a unity-gain buffer, used to de-reference the sine wave. This is followed by a comparator, which produces a square wave in phase with the original signal. The next stage is an integrator, which converts the square wave into a triangle wave. Finally, the triangle wave is fed into the amplifier and filtered. The 3dB cutoff points of the filters are 3Hz and 20Hz, respectively. The 3Hz cutoff eliminates the possible DC component, while the 20Hz cutoff rounds off the triangle wave at the vertices, producing a smoother acceleration/deceleration of the mirror.

4.3.3.4 Camera Shutter Synchronization

The drive signal from the low frequency element is fed into the inputs of a comparator, which produces a square wave with the edges corresponding to the zero crossing. An RC filter removes the high frequency noise, while the two diodes clamp the signal to maintain a low voltage swing at the inputs of the comparator.

A passive high pass filter feeds the rising edge of the square wave signal to a 555 timer IC, configured as a one-shot. This stage is capable of producing a delay adjustable between 10ms

and 100ms. This delay is adjusted to the time between the zero crossing and the moment when the camera's shutter has to be opened. At the end of the delay, as the output of the 555 timer goes low, a comparator is configured to produce a positive pulse. Finally, this pulse is used to trigger the camera's shutter. The exposure time of the camera is controlled independently by the camera software, running on the notebook computer.

4.4 System Calibration

This section describes the process of positioning and mounting the raster generator together with the camera. The accuracy of the 3D reconstruction depends on the calibration precision.

The raster generator module is mounted at about 5.5 inches from the camera, on its handle, using a bracket and a few bolts (**Figure 4-1**). This bracket allows the raster generator to swivel in the vertical direction and to spin on an axis perpendicular to the first one. To perform the calibration, the camera-raster generator assembly is mounted in a vice, placed on a table next to a flat wall (**Figure 4-12**). The camera was poisoned at 18 inches from the wall, with the optical axis approximately perpendicular to the wall. To verify that the optical axis was normal to the wall, a square was painted and several images were taken, while adjusting the position of the camera, until the all four sides of the square had equal lengths in the image. The even lengths of the square's sides were checked counting the numbers of pixels in the image (the CCD sensor used has square pixels).

Next, the raster generator was turned on and while swiveled in the vertical direction, several images were taken, until the top line of the raster was visible near the top of an image. The mounting hardware for this degree of freedom was then secured. Similarly, the raster generator was rotated until the lines became horizontal in the image. Since the lines were not exactly parallel, due to constructive imperfections, the adjustment was performed with respect to a middle line. The mounting hardware was then secured. At this point, additional images were taken to confirm that the alignment is sufficiently precise.

Once the relative position between the camera and the raster generator were fixed, several measurements were necessary to determine the exact distance between the two, the angle of the top line (laser plane) with respect to the camera's optical axis, and the angular distance between the laser planes (that generate the laser lines). The distance between the camera and

the raster generator (shown as **d** in **Figure 3-7**) was measured as the difference between the distance from the center of the camera lens to the table and the distance from the middle of the y-mirror to the table. The distance **a (Figure 3-7)** was measured by hanging a small plumb line near the camera lens and recording the distance between the middle y-mirror to the line.

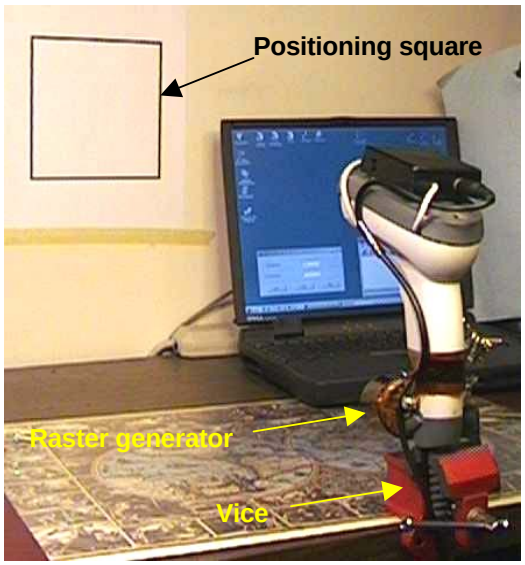


Figure 4-12: Calibration setup

The angle for the top laser line (θ) was measured as:

$$\theta = \arctan(CB - (a + Z))$$

were CB is the difference between the distance of the laser line to the table and the distance of the y-mirror to the table (previously recorded), Z = 18 inches and a = 1 (measured earlier). The angle for the bottom line was measured in a similar way. The angle increment between consecutive lines was calculated as the ratio of the difference between the top and bottom angles and the number of lines minus one. For our current prototype, this angle is 0.5 degrees.

These parameters, together with the intrinsic parameter of the camera (which, if not available, they could be obtained separately by a standard camera calibration procedures) are being used by the software for 3D reconstruction.

4.5 Discussion

It is assumed that the frequencies necessary to drive the two scan elements are produced by a computer's sound card. This is very convenient in the current configuration, because the computer is needed to control the digital camera used by our prototype, so it is readily available. In principle, however, the camera can be self-contained, with the frequencies generator constructed as part of the electronic control circuit. In fact, for the high frequency element, this approach is desirable, since a closed loop circuit can adjust the driving frequency, in order to track the possible small drift of the mechanical resonance of the element, making it more stable over temperature change.

4.6 Summary

This chapter has presented the core of the thesis, in which the raster generator is considered the essential component. Each of the elements of the system has been described in sufficient detail to expose the advantages and limitations of the design. By now, the reader should be familiar with the construction of the laser module and the two scanning mirror elements, as well as with the mode in which they are being controlled. For those with an electronics background, the schematic of the control circuit, along with its description should be straightforward. The digital camera is described as well, although it does not have any significant particularities or adaptations for this application. The system calibration section provides a simple method for aligning the raster generator module with the camera.

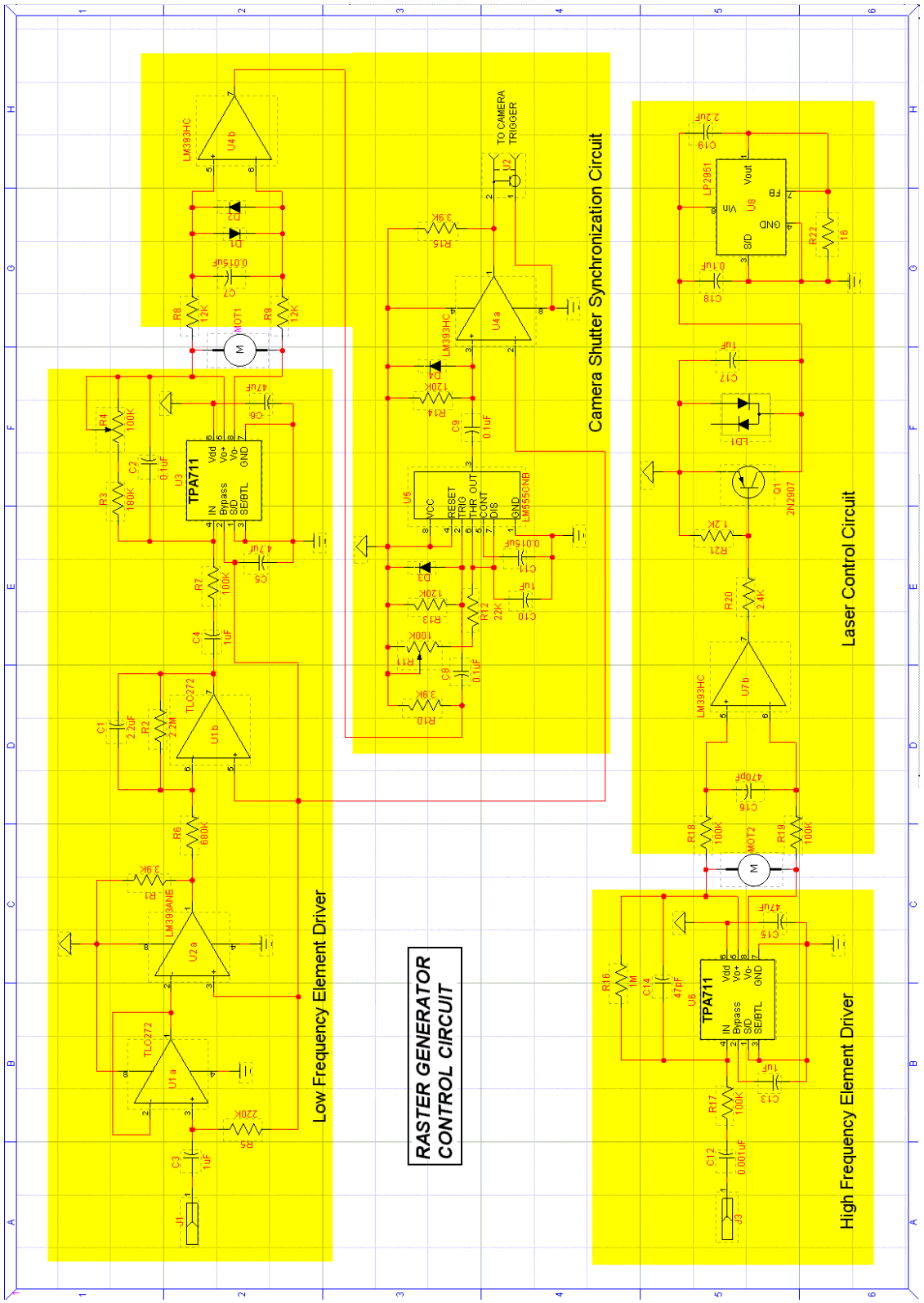


Figure 4-13: Electronic control circuit

5 Software

The software part of the 3D-camera project can be divided into three categories, according to their functions: image processing, camera control, and raster generator control. The following sections explain their functions.

5.1 The Image Processing Software

The image processing software has the role to read in the image containing the projected laser stripped pattern and create a 3D model, which can be in manipulated, in turn, via a graphical user interface. The output model can be visualized as a 3D mesh (with or without texture), as lines, or as points. To support the model reconstruction, various image-processing algorithms are being applied to the raw image. This section will go over the various steps necessary to extract the geometric information from the original picture.

The laser lines from the original images are used for extracting depth information. The fidelity of the reconstruction depends on how accurate the laser traces are being recovered. Various factors such as object texture, color, and reflectivity make the laser lines appear with various intensities, sometimes being barely visible. Even in an image where the surfaces of the objects are smooth and uniform (**Figure 4-3**), the lines do not have even intensity due to the laser speckle effect. The software should be capable to identify the laser lines and deal with the intensity variation and discontinuities in such a way that it does not significantly distort the shapes of the objects to be recovered.

5.1.1 Extracting Geometric Information

An image of an object, containing the projected lines produced by the laser, is presented to the software, which is expected, after a number of iterations to recover a 3D data from the laser lines. Various steps are necessary, each meant to remove from the original image more of what is "background" and retain what are "laser lines".

5.1.1.1 Low Pass Filter

The first step consists in blurring the image. Its main purpose is to reduce the speckle, which characterizes the laser lines in the image (**Figure 5-1 (a)**). The abrupt intensity variation along the laser lines can be regarded as high frequency noise; thus, a Gaussian filter is used to

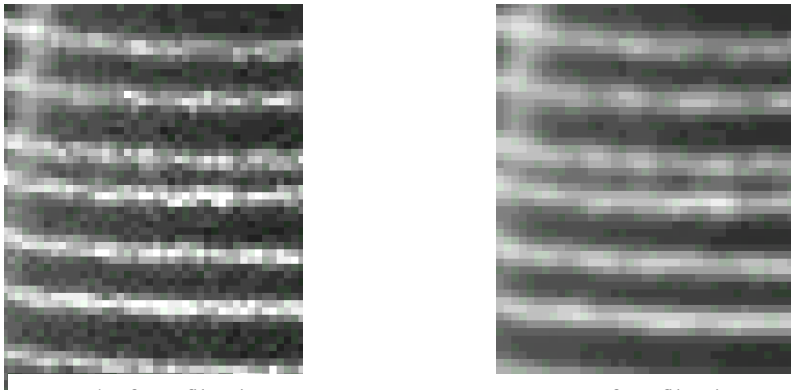
smooth the lines. This filter was chosen since it naturally characterizes the laser's intensity profile; therefore, convolving the image with a 2D Gaussian kernel should enhance the laser lines and reduce sporadic high frequency noise in the image. The size of the kernel was chosen according to the laser's spot size. Experimentally, I found that a kernel of size three smoothens the line intensity, without blurring the lines too much. The horizontal and vertical 1D Gaussian filters are given by:

$$I_{i,j} = (hI_{i,j-1} + I_{i,j} + hI_{i,j+1})/3 \qquad 0 < h < 1$$

$$I_{i,j} = (kI_{i-1,j} + I_{i,j} + kI_{i+1,j})/3 \qquad 0 < k < 1$$

respectively, where h and k are related to the variance. In our implementation, I have experimented with values of k between 0.4 and 0.6, which gave acceptable results for most of the images used. In general, these numbers have to be chosen as a function of the laser spot size. For a larger spot, a Gaussian kernel of size 5 would be preferable.

Due to the separability of Gaussian filters [Gomes 97], the 1D filters are applied in both horizontal and vertical directions independently, which is equivalent to applying a 2D kernel to the entire image. The result is an image with softer intensity variations (**Figure 5-1 (b)**). After the low pass filter is applied, the laser lines become about three pixels wide. For accurately calculating the depth, the lines in the images should be made one pixel wide.



(a) before filtering **(b)** after filtering

Figure 5-1: Speckle effect in laser lines: before (a) and after (b) low pass filtering

5.1.1.2 Intensity Gradient

The intensity profile of a laser line obtained from the previous step is more or less Gaussian in the direction perpendicular to that line. The intensity peak corresponds to the center of the laser spot, which produced the line as it moved across the surface of the object. The loci of the laser spot's center are the ideal lines that need to be extracted from the image. For practical purposes, we will produce lines that are one pixel wide and are obtained from the blurred lines produced by the filtering process. We qualify a pixel in the image as belonging to a laser line if its intensity is above a predefined threshold and satisfy one of the following:

- The top and bottom neighboring pixels have intensity slopes of opposite signs above a predefined threshold.
- The diagonal top-left and bottom-right neighboring pixels have intensity slopes of opposite signs above a predefined threshold.
- The diagonal top-right and bottom-left neighboring pixels have intensity slopes of opposite signs above a predefined threshold.

The line extraction function begins by calculating the intensity gradient for each pixel along vertical and the two diagonal directions:

$$
\begin{aligned}
Dy_{i,j} &= (-I_{i-1,j} + I_{i+1,j}) \\
Dxy1_{i,j} &= (-I_{i-1,j-1} + I_{i+1,j+1}) \\
Dxy2_{i,j} &= (-I_{i-1,j+1} + I_{i+1,j-1})
\end{aligned}
$$

As a result, a new image is generated, so that the intensity of all pixels that do not satisfy the conditions above, are set to zero:

$$
I_{i,j} = \begin{cases} I_{i,j} & \min\left(Dy_{i-1,j}, -Dy_{i+1,j}\right) > t \\ & or \quad \min\left(Dxy1_{i-1,j-1}, -Dxy1_{i+1,j+1}\right) > t \\ & or \quad \min\left(Dxy2_{i-1,j+1}, -Dxy2_{i+1,j-1}\right) > t \\ 0 & else \end{cases}
$$

where t is the predefined threshold. This value should be higher if the laser lines in the image are bright and it should be lower if the lines are dim. Ideally, some preprocessing of the image should estimate this value and assign it adaptively. For the images analyzed in this project, values between 8 and 11 were most appropriate.

This selection criterion extracts horizontal and diagonal lines rejecting the vertical ones, because a smooth surface could not produce laser lines with high slope. Near vertical lines would be produced by a surface which has an abrupt change in depth; such objects are not considered for this implementation. The algorithm assigns the original gray level to any pixel considered part of line, while all other pixels are assigned black. The resulting image consists of numerous line segments of various lengths, mostly one pixel wide, on a black background, as can be seen in **Figure 5-2**.

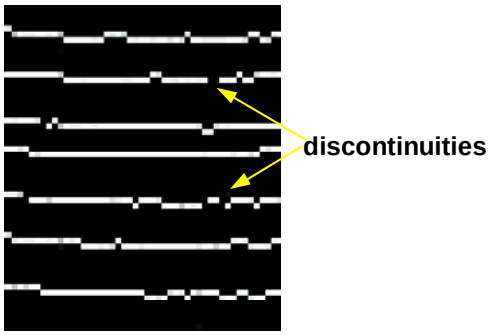


Figure 5-2: Extracted Laser Lines after intensity gradient filtering

5.1.1.3 Line Cleaning and "Chaining"

In some cases the lines are wider than one pixel, generating ambiguities. In this case, only the pixel with the highest intensity is retained by a following step:

$$I_{i,j} = \begin{cases} I_{i,j} & I_{i-1,j} < I_{i,j} > I_{i+1,j} \\ 0 & else \end{cases}$$

The lines so obtained have numerous small discontinuities (**Figure 5-2**). These small gaps are often caused by surface irregularities. Each segment that is made of adjacent pixels is assigned a *segment ID* number. The line segments separated by small gaps should be linked together since they most likely belong to the same laser line, *i.e.*, they have been projected at the same vertical angle from the raster generator. The program simply fills the one-pixel gaps by

interpolating the neighboring non-black pixels. For gaps larger than one pixel, we opted for a conservative approach and the missing portions of the lines are not reconstructed.

The next step attempts to link, without connecting, the segments that appear to be part of the same laser line. This assumption is based on the proximity of segments' ends, which should be only a few pixels apart. Each segment formed by adjacent pixels has, at this point, a *segment ID* number assigned; all the segments that are separated by only a few pixels (three pixels for our prototype) are now assigned the same *line ID* number. The *segment ID* numbers are independent from the *line ID* numbers. If more than three pixels separate line segments, no assumption is made that they belong to the same line. Scattered pixels and the short segments that are not part of long line chains (<5 pixels) are discarded.

Figure 5-3 shows a situation when segments cannot be consistently assigned to lines.

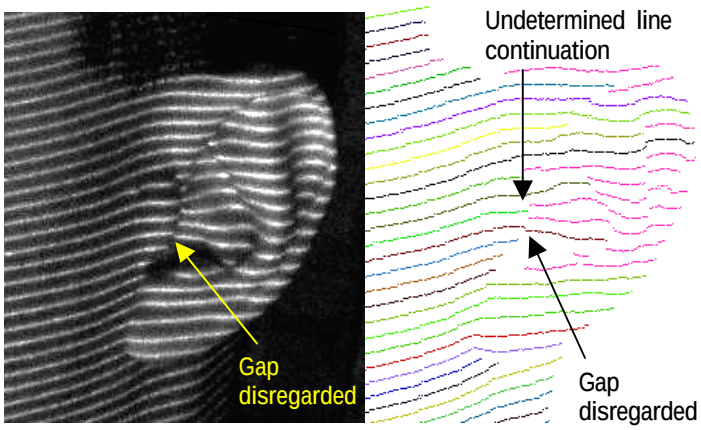


Figure 5-3: Lines ambiguity

On the left side of the figure, an image of an ear is shown, where due to the depth discontinuities, the laser lines are separated into segments. On the right, the recovered lines are shown with different colors to make the illustration clear. The lines, with the exception of the pink ones, have been assigned ID numbers by the first pass of the algorithm over the image. The line indicated by the arrows pointing upwards has been assigned one ID, although the two constituent segments should belong to different laser lines. Since the two segments appear continuous, the software treats them as one line, therefore assigning a wrong ID to one of them. In the right side image, the light green line, pointed to by the top arrow, could be

continued by any of the two magenta color lines on the right. This ambiguity situation, may occur when depth discontinuities are present. Solutions to this case of arbitrary geometry are difficult. Automatic procedures are likely to fail to correctly assign IDs.

5.1.1.4 Assigning Angles to Lines

The output of the previous steps consists of segments chains, which are used for reconstruction. Each of these segments chains is assumed to have been generated by the raster generator at a particular angle, but some of the segments chains, which were assigned the different line ID number, could have been generated by the same laser planes. It is necessary to establish a succession of the lines, in the vertical direction, which reflects their actual relative position. For a smooth surface, the order of the laser strips, for any vertical column of pixels, is in monotonic relation to the angle θ, made by the laser plane with the optical axis (**Figure 3-7**), *i.e.*, a traverse of the image in the vertical direction, from top to bottom, intersects the laser lines in an incremental order (but not necessarily consecutive, because of possible gaps between segments).

The image processing software performs a pass through the entire image, determining the pixels column (vertical line) that has the highest number of laser lines crossings. This is the starting point for determining the order of the lines. **Figure 5-4** exemplifies the algorithm used for the entire process of assigning angles to the laser lines. For the column that contains the vertical line with the largest number of crossings, the laser lines are assigned position numbers in incremental order, starting from the top. The positional numbers assigned by this step do not account for the lines that are not intersected because of the gaps, but the monotonicity for the lines intersected is preserved. The laser lines not intersected by the vertical line do not have positions assigned yet (shown in magenta).

In the step following, the program attempts to assign position numbers to the lines that have not been assigned previously. As in the previous step, the assumption was made that along any vertical column in the image the lines are monotonically ordered. Therefore, any line that has not been assigned a position number and is located below one that has been assigned should have a larger position number. The zoomed portion in **Figure 5-4** illustrates this situation for the line between line positions 31 and 32. The magenta line should be assigned number 32 and all the position numbers below have to be incremented by one in this case. If multiple lines

were present in between the assigned lines, they would have been assigned position numbers incrementally, while the lines below would have been incremented by the number of lines in between.

The algorithm does just that for the entire image, starting with the left top corner. This is repeated, covering the entire image and every time a non-numbered line is encountered, it is assigned a number consecutive to the previous line crossed, and the position numbers for the lines following, is increased by one. Once the process has been applied for the entire image, it is repeated from left to right and then from the bottom up, to account for the top lines which have not been assigned in the step (*e.g.,* the first magenta line in **Figure 5-4**). This approach allows the propagation of information along the horizontal direction due to the interleaving of the segments.

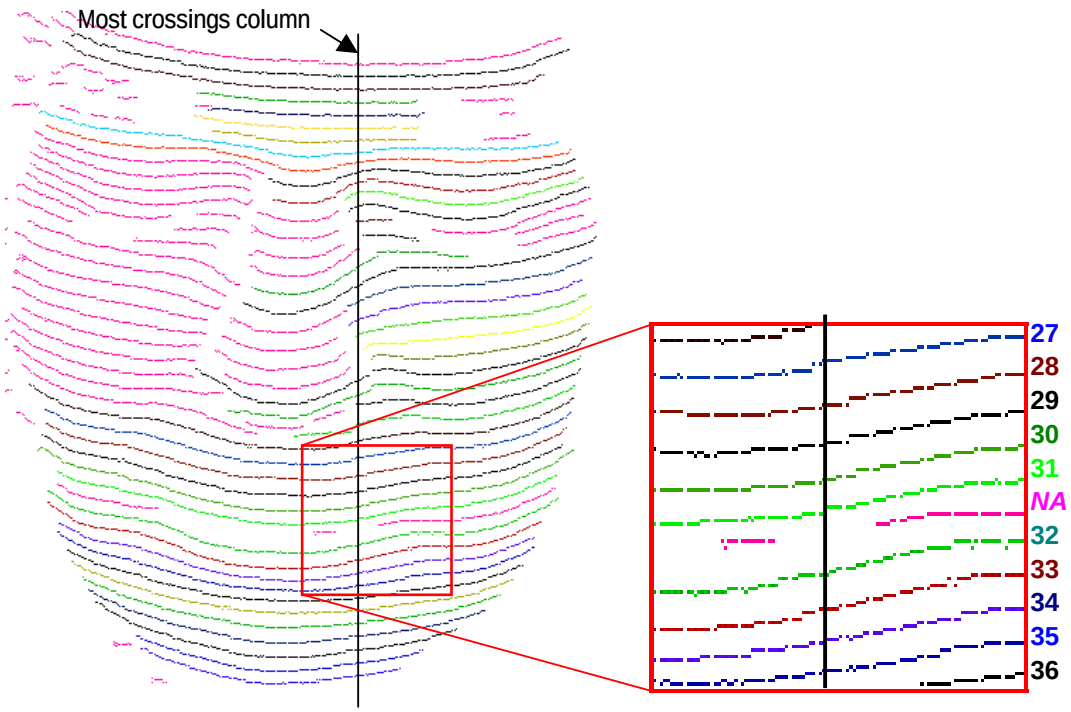


Figure 5-4: Angle assignment

The angles are assigned to the lines, based on the assumption that no projected line was completely obstructed, which is reasonable for smooth surfaces. Knowing the angle between

consecutive laser planes, as an intrinsic design parameter (0.5 degrees), once the angle for one of the lines is known the others can be assigned in consecutive order. In the current implementation it is assumed that all the lines projected by the laser generator are visible, at least partially, in the image. Assuming known the angle at which the top line was projected, the other lines are assigned angles in consecutive order. In an unconstrained situation, it is not possible to guarantee that the top line is always visible, since that depends on the position of the object with respect to the 3D-camera

Although not implemented in this construction, it is possible to take two consecutive images, one with all the lines and the other with only one line projected at a known angle. The projection angle of one line only can be controlled by the raster generator electronics. Using the angle information from the second image into the first, all the lines can be assigned angles based on their relative position.

The depth Z is assigned to each pixel on the line, based on the vertical coordinate y and angle of the line θ, (assigned in the previous steps) according to equation **(3-6)**.

5.1.2 Model Rendering

The X and Y world coordinates are calculated based on perspective projection equation **(3-4)**. Every pixel on the recovered lines is assigned (X, Y, Z) world coordinates; therefore, the absolute position with respect to the camera is known. OpenGL is used to produce the visualizations of the model, which can be arbitrarily translated and rotated in 3D.

The initial points recovered are used to create a wire frame and a mesh on which skin and texture for the model can be applied. The most appropriate is to use the texture of the object from the initial image without the laser lines. The laser lines in the original image have Gaussian intensity distributions in the directions normal to the slope of each line, therefore the laser light spreads over three or four pixels across. It is necessary to interpolate based on pixels that are at a distance of two or three pixels away from the middle of the line (normal to the local slope). Depending on the amount of ambient light that contributes to the brightness of the object, the light scattered by the laser lines due to the surface diffusion of the object may or may not be dominant. If two little ambient light is present, after removing the laser lines the

texture becomes dark, so it may be preferable in such cases to use only the mesh or the wire frame (laser lines).

Our prototype supports the rendering of the reconstructed model as triangle meshes, laser lines and points. In all modes, the model can be rendered with texture extracted form the original picture, with the laser lines removed.

5.2 Camera Control Software

The images acquired by the camera are downloaded to a PC via the serial port, in order to be processed. These images need to be taken at a precise time relative to the raster generation; the beginning of the image acquisition has to be synchronized with the raster generator such that it takes place at the beginning of a scan line and after the mirror of the low frequency element has traveled ¼ of the full scan angle. Both, the length of the integration and the gain, which controls the brightness of the image, need to be adjustable. It is important to stress the fact that the brightness of the laser lines is not a function of the integration time (camera exposure time), while the rest of the image, illuminated by the ambient light is. Consequently, the higher the integration time, the lower the contrast between the laser lines and the image illuminated by the ambient light. The integration time determines how many laser lines are visible in the image for a given frequency of the high frequency element. A frequency of 2000Hz will produce 60 lines in 30ms, which is the maximum integration time for our camera. At 30 ms, however, the environment has to be quite dark in order for the laser lines, produced with the 10mW laser throughput, to produce distinguishable lines. For typical indoor illumination, it is necessary to reduce the integration time below 20ms, resulting in less than 40 lines, which is currently done in our prototype. The shorter integration time helps also reduce the possible blurring of the image, due to the operator's hand jitter. The contrast can be improved by using a higher throughput laser, possibly infrared, or by placing a narrow-band optical filter, tuned to the laser's wavelength, in the path of the light that enters the camera. Both modifications are considered for future development.

The camera control software runs on a PC and is capable of adjusting various parameters as integration time, gain, image compression, and communication type. Since the camera is connected to the computer's serial port, these parameters are uploaded to the camera when the

user decides to do so. The software has a graphical user interface as shown in **Figure 5-5**. The windows for changing these parameters can be opened by clicking the Scanner menu and selecting options In order to acquire an image, the user has to click on the button showing the light bulb. The camera produces a short beep indicating the acquisition of the frame, followed by a second beep, which indicates that the image has been stored into the memory and is ready to acquire a new frame.

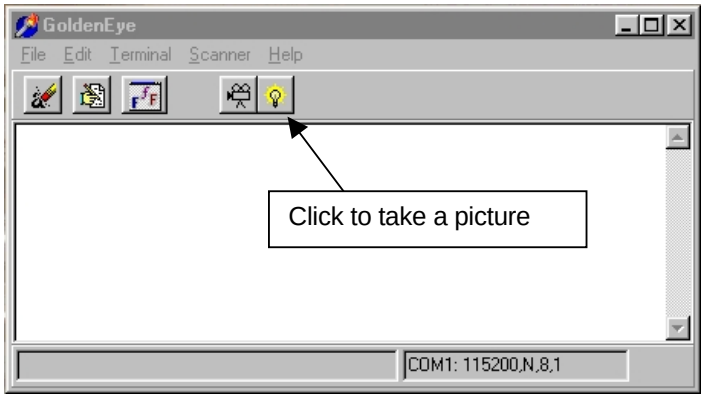


Figure 5-5: Camera control software interface

The integration time is adjustable between 1ms and 30ms and the gain can be adjusted over a range of more than 20dB. The acquired image can be stored in bitmap or jpeg format, in this case with adjustable compression. Although it is preferable to use high compression to reduce the image download time, this in not recommended since it creates artifacts, which affect the line detection algorithms. The maximum communication speed between the computer and the camera used in the current prototype is 115 Kb/s. The rate could be increased by several orders of magnitude if the camera could communicate over a USB or a Firewire port.

5.3 Raster Generator Driver Software

The raster generator needs constant frequency sine waves for the low and high frequency scan elements. The two waveforms are produced by a notebook PC sound card through the left and right audio channels, respectively. In order to generate the desired frequencies, a simple freeware application from the Physics department of Rutgers University

(http://duncan.rutgers.edu/physicsfreewares.htm), named "Beat in brain" is currently being used. This allows the generation of independent frequencies for each of the sound card channels, adjustable in increments of 1Hz. The software graphical user interface is shown in **Figure 5-6**.

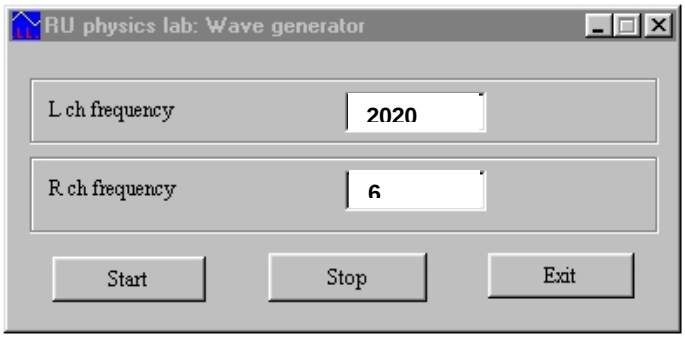


Figure 5-6: Waveform generator interface

The frequency for the left channel should be set to 2020 Hz (which corresponds to the mechanical resonance of the high frequency element) and the frequency for the right channel should be 6 Hz. Once the frequencies are entered and the application is started, the sound card produces the two sine waves continuously.

5.4 Summary

Three different software applications are used by our 3D-camera prototype. Most of the chapter explains the algorithm for shape reconstruction, systematically, providing the reader with an understanding of the reasons for the numerous image processing steps. The chapter presents our approach for extracting the laser lines from the image and determining their order. The success of the reconstruction is based on the laser lines extraction and on how the angular assignments are made. The OpenGL 3D viewer is not described here, since it is has little relevance for the purpose of this thesis. The software for adjusting the camera parameters (exposure time, gain, and image compression) and the software for creating the mirrors driving waveforms were briefly discussed.

This page was left blank intentionally.

6 Results

This chapter presents the results and discusses the limitation of our 3D-camera. These are either due to fundamental limitations or to the quality of the current prototype's hardware and software. Various images are shown, as acquired by the camera, next to wire frame representations of the reconstructed objects. The objects used were chosen to have relatively smooth and non-shiny surfaces, uniform textures, and not too many discontinuities. The human face satisfies all the requirements above and was found to be an appropriate and challenging test case of practical interest. Other objects are shown ass well, to illustrate the behavior of the camera.

Figure 6-1 left shows the image of a person's face profile, as taken with the camera; on the right, one can see the 3D reconstruction from a different viewpoint. The lines that fall onto the smooth surface of the skin are clearly visible and unambiguous, especially where the normal to the surface does not make a large angle with the optical axis of the camera. These surfaces are easily and accurately reconstructed.

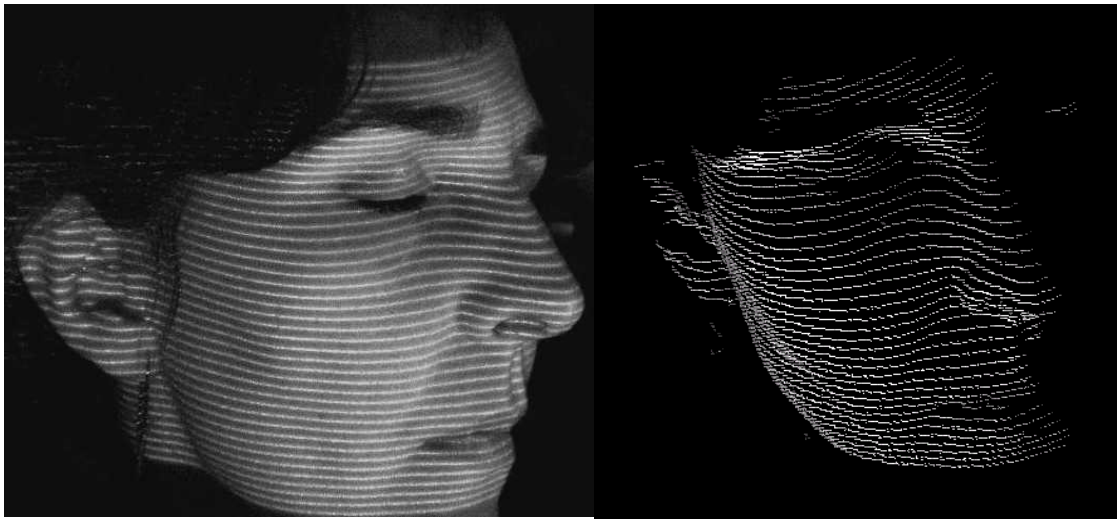


Figure 6-1: Face Profile: Original image (left) and reconstructed 3D shape (right)

By contrast, the laser lines produced over the hair, or over the regions where the surface has sharp curvatures, cannot provide accurate contour representations. When the surface is not

reflective enough the lines cannot be distinguished with enough confidence from edges or other line-like objects in the image. The software rejects such line-like structures, as can be seen in the case of the hair.

When the surfaces are not smooth, the lines ambiguity can create artifacts. Such artifact is visible in **Figure 6-1** right, in the region of the ear. Some of the line segments, shown over the ear in the left image of the figure, were not assigned the angles correctly, creating a step in depth of the object (along the direction of the optical axis).

Sometimes, the incorrect angle assignment, due to a small portion of the image, can propagate along an entire line. The reconstruction in **Figure 6-2** preserved the overall shape of the face, with the exception of one line, which is clearly discrepant. This line has been assigned the same angle as the one above it, because the lines' ambiguity around the left nostril. Notice that the entire line is shifted forward, but has a similar profile to its top and bottom neighboring lines. . The software could be constrained to reject this line, since although possible, it is unlikely for such surface to have been scanned.

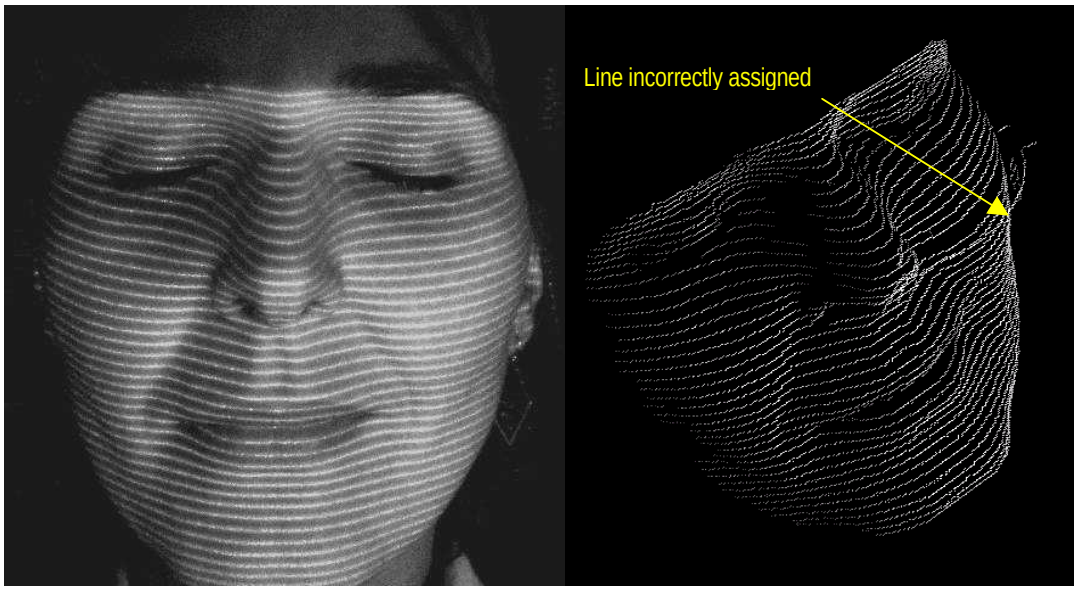


Figure 6-2: Error propagation

The same problem is shown in **Figure 6-3**, for the case of a hand, but with a more pronounced effect. In this case the line is isolated from the rest and one can easily observe that it cannot be part of the object. Since the line is isolated from the rest of the model, the

software may be designed to disregard it. The current software version, however, does not address this issue, although relatively straightforward to implement.

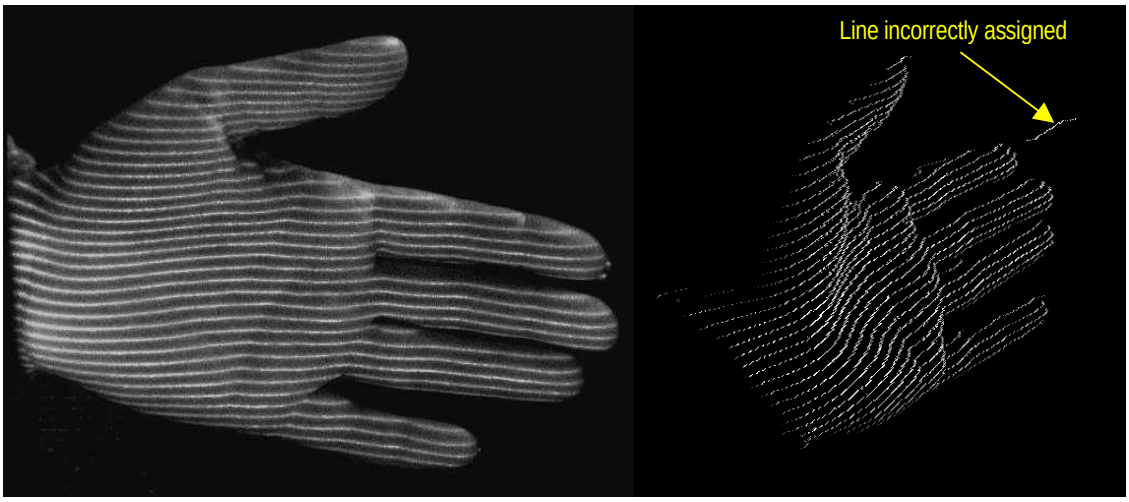


Figure 6-3: Incorrect angle assignment

When a group of lines that does not contain any line with an absolute angle reference is reconstructed, the position of this entity cannot be determined in space. This segmentation takes place when a discontinuity is present; the current software discards such portions. **Figure 6-4** shows on the left a reconstructed face in which the left ear does not appear. The relative position of the ear, with respect to the face could not be accurately determined.

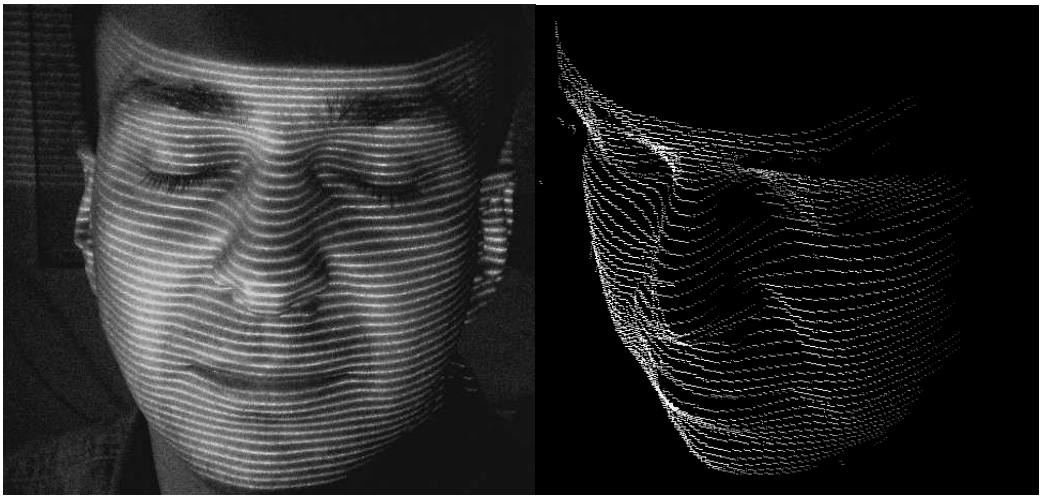


Figure 6-4: The left ear not rendered

One aspect that affects the accuracy of the reconstruction is the uniformity of the angles between consecutive laser planes, which are assumed to be equal, as the angular velocity of the mirror is assumed constant. In reality, for this particular construction, this is not true, due to some constructive problems with the low frequency element electronic circuit driver, which can be corrected in a future implementation.

In the current prototype the mirror slightly accelerates as it approaches the middle position and it decelerates as it moves away, therefore the angles between the laser planes are larger near the center of the raster and smaller toward the ends. This effect can be observed when the raster is projected onto a flat surface (**Figure 6-5**).

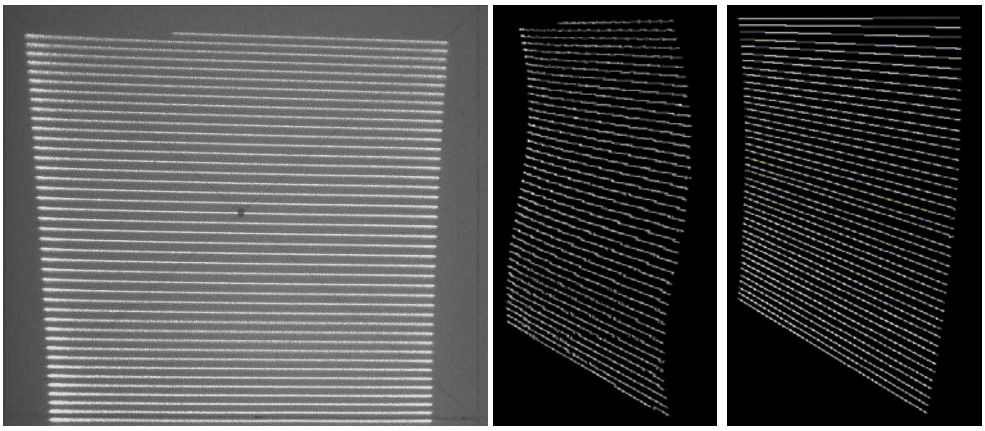
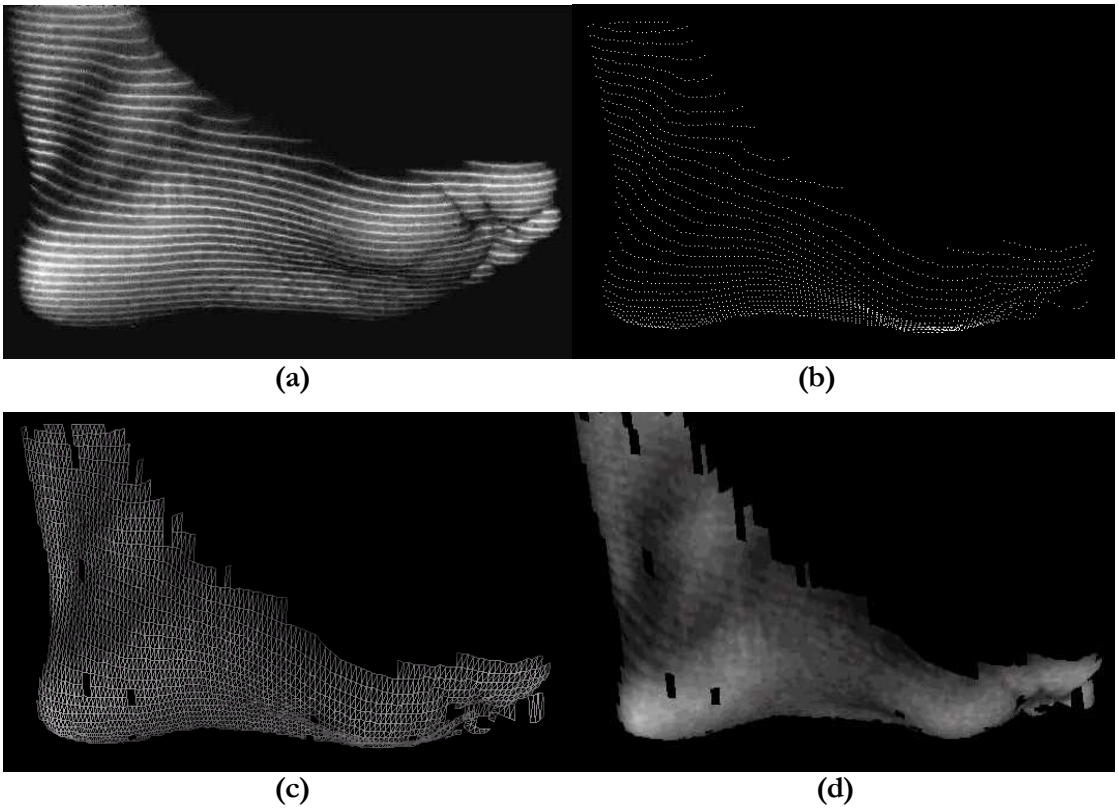


Figure 6-5: Effect of y-mirror nonlinearly

The left image is the picture of the raster projected onto a wall; note that the trapezoidal shape is due to the perspective projection of the pattern on the flat surface, which is normal to the optical axis of the camera. The middle image is the 3D reconstruction of the left image. The rightmost image is the reconstruction of a synthesized jpeg file containing parallel and equidistant lines. In reality the lines projected on a plane perpendicular to the camera's optical axis would not be equidistant, but would get denser from top to bottom, due to the projection angle of the raster generator. The curvature of the reconstruction of the synthesized image is present because the lines in the synthesized image are equidistant.

The OpenGL interface is gives the user the option to render the models not only as wire-frame, but also as points, surface mesh, and skin (**Figure 6-6**). The user has the option to select the texture mode, in which case, the original texture with the lines removed is rendered.

The surface rendering mode, however, is less forgiving to errors in reconstruction, producing visibly unpleasant artifacts. Missing points in the mesh result in rectangular gaps in the skin surface (**Figure 6-6 (d)**), while incorrectly assigned points result in various artifacts of the geometry (not shown). Further work is necessary to improve quality of the acquisition and the image processing software.



(a) **(b)**

(c) **(d)**

Figure 6-6: Original image (a) and 3D models: points (b), mesh (c), and textured surface (d)

This page was left blank intentionally.

7 Conclusions and Further Work

This thesis presented the design and implementation of a portable and inexpensive 3D-camera. A detailed description of its hardware and software components has been provided. The effectiveness of the proposed design has been demonstrated by recovering the 3D shapes of several real-world objects, such as human faces. As all the other laser-based approaches, this one is also subject to several physical limitations due to the optical material properties, ambient illumination, occlusions, and speckle noise. While the design decision of recovering 3D shape from a single image makes the system easier to use, it makes the task significantly more challenging. Some artifacts still noticeable in the reconstructed models as small holes result from non-reflected laser light or by some limitations in the current software, such as improper angular assignment for the lines and incomplete rendering. While the artifacts resulting from the first class of problems cannot be overcome in general, significant improvements can be obtained with the implementation of more careful rendering procedures. With the appropriate improvements, such a prototype may evolve into a commodity item, opening the doors to consumer 3D photography.

The prototype and the ideas presented in this thesis are the result of several months of both theoretical and experimental work, which includes the design and the construction of the most essential mechanical and electronic subassemblies. The prototype described here was intended as a proof of concept rather than a pre-production unit; the greatest value of its construction was the learning experience. During the design end experimentation process, only few of the many ideas we had were put in practice; most were discarded, but some, which were more difficult to implement at that stage, are worth mentioning. It was apparent from the beginning of the work that some factors, such as line ambiguity, surface discontinuity, deep cavities, material reflectivity, and others would introduce fundamental limitations, narrowing the range of applications for our prototype. However, it was equally exciting putting together a shape acquisition device using "home-made" elements. No mechanical stress analysis, temperature, humidity, electrostatic discharge, and other engineering tests were required to be considered before seeing the device working.

Aside from bringing the prototype closer to a product, several features can be added, either to extend its range of applications or to improve performance at a fundamental level. The following sections discuss briefly some of the ideas that can be considered for future work.

7.1 Absolute Reference for Reconstruction

In the last paragraph of section **5.1.1.4 Assigning Angles to Lines,** it has been mentioned that the angles are being assigned to the laser lines successively, in the vertical direction, starting from the top line. This assumes that the top line produced by the raster generator is visible in the image and its angle is known from construction. There are cases, however, when the operator positions the 3D-camera with respect to the object to be digitized, such that the top line does not appear in the acquired image. In this case, as suggested before, a second image, containing only the middle line of the raster, may be taken to be used as a reference for the other lines.

An alternative is to mount a laser pointer on the 3D-camera such that it projects a laser beam, at a known angle in the vertical plane with the optical axis of the CCD camera. The position of the spot in the image can be used as absolute reference. This approach will allow the reconstruction of the model from one image only.

7.2 Ambient Light Shading Information

If the ambient light is significant compared to the laser lines, it can be used to extract more information from the image. Once the shape of the object is reconstructed, the shading effect of the ambient light can be analyzed as well. The portions of the image between the laser lines show the object illuminated by the ambient light. After 3D reconstruction, the recovered normals to the surface can be used to estimate the position of the non-diffuse light sources. The specular reflections of the surface, which appear in the image as particularly bright areas, may be good hints in determining the position of the ambient light sources. The model obtained would have a more realistic look when translated and rotated, if the original illumination is created in the modeling software.

7.3 Laser Lines Shading Information

One particularity of the current implementation is the fact that all the laser lines are projected from one point and the intensity of the laser is constant during the integration time of the camera. Due to the sinusoidal motion of the high frequency scanning mirror, the laser lines are dimmer in the middle portion. It is possible to control the laser intensity such that is constant for the entire angular motion, for any given elementary angle $\Delta\alpha$. The raster generator can then be approximated for a point light source, which emits light uniformly along planes, fanning out at equal angles. For the image acquired, this is equivalent to having a point source projecting a stripped pattern. For a surface that has a uniform texture, shape from shading may be used to provide additional information for the 3D reconstruction.

7.4 Time-intensity Matrix

The most difficult problem for this "multiple lines in one image" approach is to determine at what angle each of the lines was projected onto the surface of the object. The intensity of a laser line varies as a function of the spot's velocity and the surface reflectance, as well as the angle of incidence and the material's texture. Although each point of a line can be associated with a unique time at which it has been produced, the entire image is integrated over a relatively long time interval, without preserving the absolute time information. In relative terms, however, since the lines are produced in an orderly fashion, a succession can be established. For example, when traversing any part of an image (containing the projected laser pattern) along the vertical direction, the intersected lines are in monotonic order, but not necessarily consecutive. This is because in most images some portions of the lines are not visible, either due to obstruction or due to the existence of textures onto which the laser lines do not appear bright enough. When these discontinuities are present, it may be very difficult to assign the visible segments their correct angles. If, however, each line has a time stamp associated with each of its points, then the time-intensity information can be used to disambiguate among segments. Such information can be obtained by recording a matrix of line intensities versus time. A photo sensor may be used to acquire the intensity information as the laser scans across the object. The time-intensity matrix obtained this way may be correlated with lines intensity from the image, providing the means to select between line segments.

7.5 Measuring Tool

Often in a 2D image is very difficult to estimate lengths and evaluate distances. Sometimes photographers use this to create interesting optical illusions. For example, in advertisement, pictures of the products look more appealing than the actual objects due to well-chosen illumination and viewing angle. If instead, a 3D model is used, that would allow the user not only to visualize the object from various angles, but it could measure lengths and distances. A point and click interface, showing the spatial distances between the points that the user desires to evaluate, could be added to the 3D-visualization software. This feature would be especially useful in the case of e-commerce when purchasing products such as personal items.

These ideas deserve further investigation and the current prototype provides a good framework for further development.

References

[3D Metrics] 3D Metrics, Inc. http://www.3dmetrics.com (July, 2002)

[3rdTech] 3rdTech, Inc. http://www.3rdtech.com/DeltaSphere.htm (July, 2002)

[B&S] Brown and Sharpe. http://www.brownandsharpe.com (July, 2002)

[Besl 89] Besl, Paul J. "Advances in Machine Vision: Active Optical Range Imaging Sensors." New York: Spinger-Verlag, 1989, pp. 1-63.

[Borghese 98] Borghese, Nunzio A, et al. "Autoscan: A Flexible and Portable 3D Scanner." IEEEComputer Graphics and Applications, May/June 1998 pp. 38-41.

[Bouguet 99] Bouguet, Jean-Yves. "Visual Methods for Three-dimensional Modeling." Ph.D. thesis. California Institute of Technology, 1999.

[Canesta] Canesta, Inc. http://www.canesta.com (July, 2002)

[Clark 97] Clark, James, Emanuele Trucco, and Lawrence B. Wolf. "Using Light Polarization in Laser Scanning." Image and Vision Computing Journal, Vol. 15, 1997, pp. 107-117.

[Curless 97] Curless, Brian L. "New Methods for Surface Reconstruction from Range Images." Ph.D. thesis. Stanford University, June 1997.

[Eyetronics] Eyetronics USA. http://www.eyetronics.com (July, 2002)

[Fangeras 93] Fangeras, Olivier. "Three-Dimensional Computer Vision: A Geometric Viewpoint." MIT Press, 1993.

[Fofi 01] Fofi, David, J. Salvi, and E. Mouaddib. "Uncalibrated Vision Based on Structured Light." IEEE International Conference on Robotics and Automation, Seoul, Korea, May 2001.

[Gomes 97] Gomes, Jonas, and Luis Velho. "Image Processing for Computer Graphics." Springer-Verlag, New York, 1997

[Herbert] Herbert, P., and M. Rioux. "Toward a Hand-held Laser Range Scanner: Integrating Observations-based Motion Compensation." Visual Information Technology Group, National Research Council of Canada.

[Lissajous] Exploratorium Lissajous Patterns. http://www.exploratorium.edu/xref/exhibits/lissajous_patterns.html (July, 2002)

[McCallum 96] McCallum, BC, W. R. Fright, and N. B. Price. "A Feasibility Study of Hand-held Laser Surface Scanning." Image and Vision Computing, New Zealand, August 1996, pp.103-108.

[McCallum 98] McCallum, BC, et al. "Hand-held Laser Scanning in Practice ." Proceedings, Image and Vision Computing, New Zealand, Auckland, October 1998, pp. 17-22.

[Minolta] Minolta Co., Ltd. http://www.minoltaeurope.com/minolta/products/3_d/3d_1500.html#s (July, 2002)

[Nanofocus] Nanofocus, Inc. http://www.nanofocus-inc.com (July, 2002)

[OmniTech] OmniTech AS. http://www.reichl-electronics.com (July, 2002)

[Petrov 98] Petrov, Michael, et al. "Optical 3D Digitizers: Bringing Life to the Virtual World" IEEE Computer Graphics and Applications, May/June 1998, pp. 28-37.

[Photomechanics] Photomechnaics, Inc. http://www.photomechanics.com (July, 2002)

[Proesmans 98] Proesmans, Mark, Luc Van Gool, and Filip Defoort. "Reading between the lines– a method for extracting dynamic 3D with texture." Proceedings of ICCV '98, Bombay, 1998, pp. 1081-1086.

[Reichl] Reichl Electronics. http://www.reichl-electronics.com (July, 2002)

[Sam's FAQ] Sam's Laser FAQ. http://www.repairfaq.org/sam/laserdio.htm (July, 2002)