Abstract

The availability of rapid prototyping enhances a designer's creativity and speed, enabling quicker development of new products. However, because this process relies heavily on computer-aided design (CAD) models it can often be time costly and inefficient when a component is needed urgently in the field. This paper proposes a method to seamlessly integrate the digitization of existing objects with the rapid prototyping process. Our technique makes use of multiple structured-light techniques in conjunction with photogrammetry to build a more efficient means of product development. This combination of methods allows our developed application to rapidly scan an entire object using inexpensive hardware. Single views obtained by projecting binary and sinusoidal patterns are combined using photogrammetric feature tracking to create a computer model of the subject.

We present also the results of the application of these concepts, as applied to several familiar objects–these objects have been scanned, modified, and sent to a rapid prototyping machine to demonstrate the power of this tool. This technique is useful in a wide range of engineering applications, both in the field and in the lab. Future projects may improve the accuracy of the scans through better calibration and meshing, and test the accuracy of the digitized models more thoroughly.

Contents

List of Figures

List of Tables

Nomenclature

DMD digital micro-mirror device

FPGA field-programmable gate array

Chapter 1

Introduction

Chapter 2

Background

2.1 3-Dimensional Optical Shape Measurement Techniques

There are many different optical techniques for 3-Dimensional shape measurement. The different techniques each have their advantages and disadvantages, and are utilized in different professional fields. This section will give a brief overview of some of the many techniques available for 3D shape measurement.

2.1.1 Photogrammetry

Photogrammetry is a technique to construct a 3D image from several 2D photos. In order to achieve this, photogrammetry utilizes feature, pattern, or color matching. Algorithms frequently use reflectivity, shading, and focus to recover shape information.[16] The advantage of this system is that it can create 3D reconstructions without knowledge of the location of the cameras. However this method has much lower accuracy than most other methods, and therefore is generally not used in engineering or medical fields.

2.1.2 Time of Flight

This method directly measures the time of flight of a laser or light source. The amount of time between the light being emitted, reflected off the object, and then received by the sensor is used to calculate the distance to the object.[3] Time of flight techniques have an advantage of being longer ranged than most other shape measurement techniques, but they are also lower resolution. This makes it useful for surveying, and other long range purposes.

2.1.3 Triangulation

Optical Triangulation techniques utilize the geometry of the system to calculate the distance from the camera to the object being measured. In most cases a projector and a camera are positioned a known distance apart. The central axis of the camera is angled by a known amount relative to the central axis of the projector. This angle is known as the triangulation angle. Triangulation techniques use these known values, along with a measured value extracted from the image data, to compute the 3D data. The most common measured values are displacement and optical phase.

2.1.3.1 Laser Scanning

Laser scanning techniques work by projecting one or more laser lines on the object to be measured. The line(s) are scanned across the object while a camera captures images of the object. The camera is positioned in a known triangulation geometry with the projector. The distance to the object at each point along the line is calculated to generate a profile of the object illuminated by the line. The profiles of the line at every location as it scans across the object are combined to create a full image.[3] In order to have a resolution greater than the thickness of the laser line, an algorithm to find the center of the line can be used.

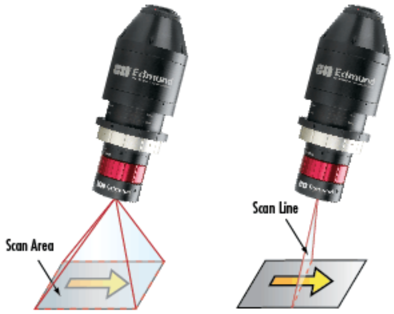

Figure 2.1: Laser Scanner[12]

2.1.3.2 Structured Light

Structured Light is a category of optical imaging techniques that use a coded pattern in projected light in conjunction with a camera to perform triangulation. There are two main categories of structured light techniques, which are continuous coding and discrete coding.

Continuous Coding Continuous coding is a term for any structured light technique that projects a continuous pattern in order to code the shape data into an image. Most continuous coding techniques utilize a sinusoidal pattern, but there are some that use other forms of continuous information. [14]

Sinusoidal Fringe Projection Sinusoidal Fringe Projection is the most commonly used form of continuous coding. This method utilizes a Digital Light Projector (DLP) to project vertical fringes that vary sinusoidally in intensity. To achieve this sinusoidal pattern the projector and CCD camera must be synched so that the fringes start at their maximum size and shrink to a single pixel wide during a single exposure of the camera. The intensity at each point is averaged over the exposure time so that a sinusoid is created. Four frames of data are captured, each phase shifted by a quarter of the wavelength of the sinusoid. The multiple phase shifted images are used to calculate the wrapped phase map of the object. An unwrapping algorithm gives the unwrapped phase map, and the actual size data is calculated using the triangulation geometry. [18]

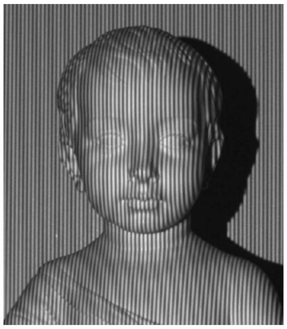

Figure 2.2: Sinosoidal Fringe Projection [18]

Binary Fringe Projection Instead of a sinusoidal pattern, a binary fringe pattern is projected. This pattern is defocused to approximate a sinusoidal pattern. The phase can then be found through phase shifting in the same manner as standard sinusoidal fringe projection. The advantage of this method is that near sinusoidal fringes can be created without long exposure times and dynamic projectors. This allows for faster applications such as single frame acquisition, but reduces the accuracy of the system by introducing higher order noise.[10]

Fourier Transform Profilometry Fourier Transform Profilometry is a method for calculating the wrapped phase map of the object in a single frame. The method takes the Fourier transform of the intensity and isolates the shape containing phase information. When the Fourier transform is performed, 3 distinct peaks result in the Fourier domain. The central peak is the brightness information and can be masked out. The two remaining peaks are symmetric about the origin and contain the shape information. One of these peaks is masked out, and the remaining one is shifted by the carrier frequency so it is located on the origin.[17] The inverse FFT is calculated and the phase data is separated from the contrast by taking the arc tangent of the imaginary components over the real components. [18]

Color Coded Fringe Projection Instead of a single fringe pattern being phase shifted between several pictures, three phase shifted patterns are projected simultaneously for single frame acquisition. These three patterns are different colors, typically RGB. The three patterns are separated and used to generate a wrapped phase map. The main limitation of this method is that it is sensitive to the transmittance, reflectivity, and absorption of the object being measured. To compensate for this the exact wavelengths used can be chosen based on the color and material of the object, or some form of coating can be applied to the object to improve the conditions. [7]

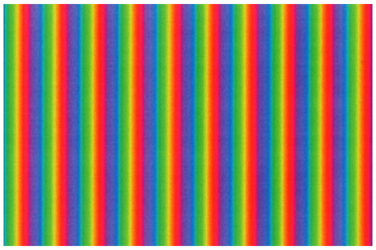

Figure 2.3: Color Coded Fringes[7]

Continuous Spatial Grading A continuous grayscale or color scale is projected onto the object. Every X coordinate in the undistorted projection has a unique intensity value, allowing for triangulation similar to a line scanner. This method is extremely sensitive to the color of the target object and shadowing on the object. [14]

Discrete Coding Discrete coding consists of projecting non-continuous patterns onto an object. These patterns are designed such that every part of the image is uniquely identified by the pattern. This identification is referred to as the "codeword" for that location. The locations identified by codewords can either be lines or pixels, depending on whether the patter is 1D or 2D. Since the location of each "codeword" is known in the projected image the displacement of the "codeword" when the pattern is projected on the object can be measured. This displacement can be used to triangulate the distance to the object for each location, thus giving the 3-dimensional shape. The two methods of discrete coding are spatial multiplexing and time multiplexing [14]

Spatial Multiplexing Spatial multiplexing methods project only a single pattern. In order to identify the “codewords” for that pattern, the surroundings are used. For a 1-D pattern this means the sequence of lines two either side of any given line are unique and thus identify that line. In a 2-D case it the surroundings in all directions within the plane are taken into account. [15]

De Bruijn Coding A pattern is constructed using a pseudorandom sequence known as a De Bruijn sequence. The properties of a De Bruijn sequence ensure that any projected line can be identified by the bordering lines, allowing for triangulation. This pattern can be binary (similar to bar code), grayscale, or color. [15]

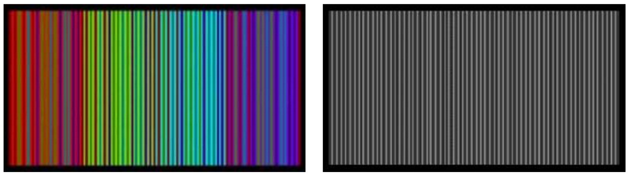

Figure 2.4: De Bruijn Patterns[14]

M-Arrays M-arrays are a 2D equivalent to the 1-D De Bruijn patterns. An array of pseudorandom dots is projected onto the target object. Any dot can be identified by the adjacent dots, allowing for triangulation. This method can utilize both binary, color or grayscale. The “codeword” of the dot can be identified not only by the type of dots around it, but by the relative density of the dots as well. [14]

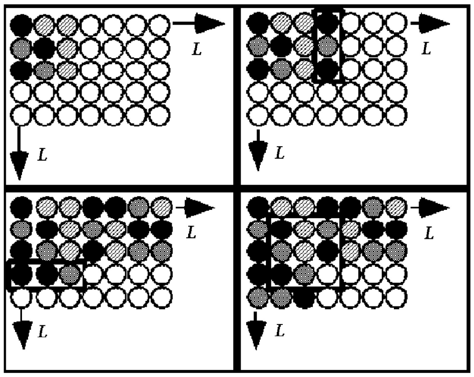

Figure 2.5: M-Array[14]

Non-Formal Coding Non-formal coding is a term used to categorize any number of spatial multiplexing methods that use unique patterns for specific purposes. These patterns do not necessarily directly uniquely identify the line or pixel as the previous methods do. Instead non-formal coding usually serves a more specific purpose such as calibration patterns. [15]

Time Multiplexing Time multiplexing captures successive images with different patterns in order to generate the necessary "codeword" for each location. The patterns a generated such that each location has a unique sequence of values throughout the series of images. Since multiple frames are needed to generate a 3D image, this method is not viable for high speed applications that require single frame acquisition. [14]

Binary Codes Binary codes function by projecting a series of binary patterns. These patterns are typically vertical of varying thicknesses or densities, similar to a bar code. A single pattern alone does nothing, but by taking into account the binary value of each line or pixel over the entire series of projected patterns, uniqueness is established. [15]

N-Array Codes N-Array codes utilize the same basic concept as binary codes, except they are not restricted to binary patterns. They can utilize color or grayscale patterns to greatly reduce the number of frames necessary to uniquely identify each location in the image. [15]

Hybrid Coding Hybrid coding consists of a combination of spatial and time multiplexing. Several spatial multiplexing patterns are displayed in series so as to create a time multiplex with them. This method achieves the high accuracy of time multiplexing, while greatly reducing the number of patterns necessary. [14]

2.1.4 Interferometry

Interferometry utilizes beam splitter to separate a single beam into two beams. One of the beams, the sample beam, is reflected off the target object and then into sensor. This beam then meets with the other beam, the reference beam, in an interferometer. The interference between these two beams gives the phase difference of the lasers. The phases from all the points on the object are generated into a wrapped phase map image of the object. [3]The phase maps are unwrapped, giving the shape of the object.

There are many advanced imaging techniques that use interferometry as a basis to generate absolute 3D measurements. One such technique is called laser speckle pattern sectioning. This method projects a speckle pattern on the target object which is measured using a CCD array. The pattern is scanned through a range of wavelengths. Each wavelength corresponds to a 2D slice of the 3D object. By adding these slices together into a 3D data array, and then performing a 3D Fourier transform, the 3D shape can be found. [3]

Interferometry has higher resolution and accuracy than many of the other techniques, and can be performed on a large range of object sizes depending on the setup. For this reason it can be used in a large variety fields, making it a versatile technique.[4]

2.2 Existing Commercial Products

There are many different commercial scanners currently on the market. Most of them cater to different professional fields and have specifications that those fields find desirable. A few of these products will be discussed in this section.

2.2.0.1 Kinect

The Kinect is a 3D imaging device made by Microsoft for use with their Xbox 360 gaming console. The Kinect works based on a fixed pseudorandom array of dots projected on an infrared wavelength. The dots are formed by an array of small micro lenses, each with a slightly different focal length. The included infrared camera picks up the projection of these dots on their environment. Groups of dots are then compared against an image taken on a reference plane. Due to the pseudorandom nature of the dot array, each group is unique enough to allow identification of a particular dot based on the relative positions of neighboring dots. Furthermore, due to the different focal lengths of the micro lenses, the pattern itself will vary based on the distance between the camera and the object [6].

Microsoft has kept its specific algorithms for calculating the depth proprietary. However, the open-source community has had some success in reverse-engineering the Kinect. In its operating range between 0.8 and 3.5 meters, the Kinect can resolve depth with about 10 mm accuracy along the optical axis, and position to about 3 mm perpendicular to the optical axis [1].

2.2.0.2 Next Engine

Next engine is a device that projects multiple laser lines onto the target object. To construct a 3D image of an object it performs line scanning in both the vertical and horizontal directions. It takes about two minutes to create a single 3D point cloud of the object. The software that is bundled with this product has the capacity to stitch

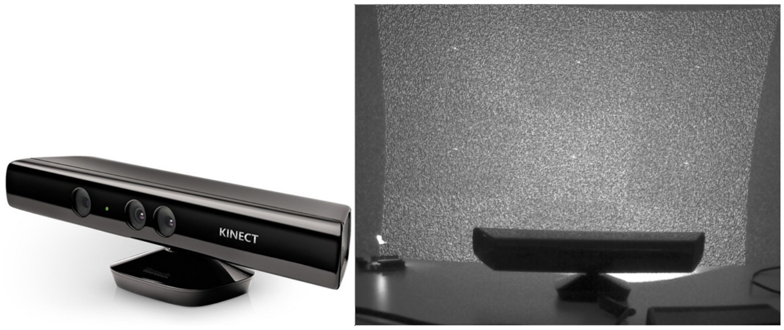

Figure 2.6: a) Microsoft Kinect[11] b) Projected pattern as seen through an infrared camera[2]

together multiple views to create full 3D images. The 3D images are in full color and can be output to several common CAD formats. The Next Engine Scanner is marketed for use in design, manufacturing, CGI, art, and medical applications. This system boasts accuracy to 0.005inches in macro mode and to 0.015inches in wide mode. [8] (image)

2.2.0.3 David 3D Laser Scanner

David 3D Laser Scanners come in two types. The first is a line scanning method that uses a line laser pointer and a digital camera. The laser pointer is scanned across the object by hand while the camera captures the image data. The other scanner is a sinusoidal fringe projection system. This scanner comes with calibration patterns and a software program capable of creating and stitching 3D point clouds. The scanner has a object size range from 10mm-600mm with a accuracy up to 2% of the object size. It takes 2-4 seconds per scan and generates grayscale images. [9] (image)

2.2.0.4 Handy Scan 3D

The Handy Scan 3D scanner is a portable line scanner. It boasts an accuracy of up to 40 microns. The Handy Scan projects a cross hair onto the target object and scans in both x and y simultaneously. The device has a camera built into it so the triangulation geometry remains constant as the laser is scanned along the object. The technology requires several sensors to be placed on the object. These sensors are randomly placed on the object and are triangulated by two cameras on the scanner. This allows the scanner to know its location relative to the object, making the freehand scanning possible. This product is marketed for reverse engineering, design, and part inspection. [5] (image)

Chapter 3

Methods

3.1 Encoding

Scan Studio combines two methods to obtain high-quality results. First, the scanner collects low-frequency data via Gray code, also known as reflected binary. Then, the system projects and decodes sinusoidal fringes to collect high-frequency data. This gives the system subpixel accuracy with inexpensive equipment.

The software uses a method called *direct coding*, in which every pixel of a camera image is encoded with the coordinates of the projector pixel that originated that ray of light. This is a very flexible approach that can be applied to any setup geometry. The patterns are *temporally multiplexed*, meaning that a single frame is not enough for a measurement; a series of frames is used to compute the result.

3.1.1 Reflected binary encoding

The easiest method of temporal multiplexing to understand is binary encoding. First, each projector-pixel index is broken down into the ones and zeros that represent it in software. Next, each bit of the resulting number is projected twice, once as a positive and again as a negative. Finally, for each bit, the two frames are subtracted. A positive

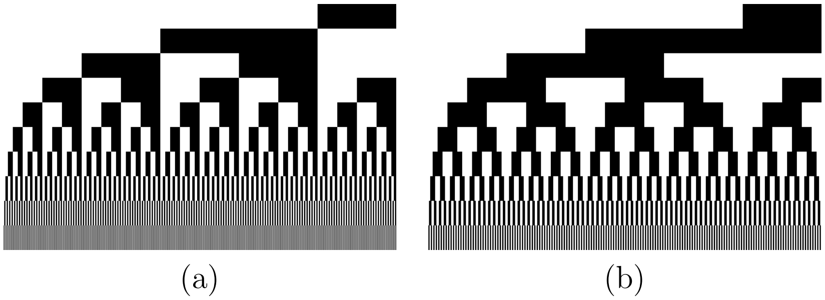

Figure 3.1: Progression of (a) a binary sequence and (b) a reflected-binary sequence, both of width 640. Each row represents a projected frame; time advances as you move down the figure. Notice that in the ordinary binary sequence, a pixel on a bit boundary in a given frame will change between every subsequent frame. The reflected binary sequence does not have this property.

difference indicates a binary zero; a negative difference indicates a binary one. These bits are summed for all frames to recover the encoded frame.

A problem of binary encoding is that pixels on the edge of higher bits are susceptible to noise. This noise can cause the decoded pixel index to vary far from its actual value. In addition, projecting the lowest bit results in a pattern of alternating black-and-white stripes 1 pixel wide. This can cause Moiré issues if the camera and projector are mismatched.

The use of reflected binary solves these issues. A pixel on the edge of a wider binary band, if resolved incorrectly, will only differ from its ideal value by one. This prevents catastrophic failure during decoding. In addition, the lowest bit of the pattern is a series of stripes 2 pixels wide, reducing the effects of Moiré. These effects may still be present, though, requiring the application of another technique for resolution of finer details.

3.1.2 Phase-shifted fringes

Another method of encoding a pixel's value is through phase shifting. This projects a pattern whose intensity varies sinusoidally from left to right. This pattern is projected multiple times, shifting to the right each time.

Decoding this requires a least-squares approach.

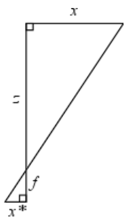

Figure 3.2: Converting camera coordinates to real-world coordinates.

One issue with phase-shifting is that it is susceptible to noise. Adding more shift steps can reduce the impact of noise but increases the number of patterns required. In addition, phase shifting requires phase unwrapping.

3.1.3 Hybrid approach

Our method uses a hybrid binary-encoding/phase-shifting approach. *Scan Studio* uses binary encoding to quickly capture high-order data. Then, it uses phase shifting to enhance this binary data down to subpixel levels.

3.2 Triangulation

Once an encoded image has been captured, this encoding must be translated from (u_c, v_c, u_p) to (x, y, z) coordinates.

3.2.1 Relating image coordinates to world coordinates

Solving the system requires knowledge of θ_C, the angle the incoming ray creates with the camera's optical axis. Figure 3.2 demonstrates the geometry involved in this operation.

By itself, the camera can directly measure neither x_C nor z_C. (If it could, this project would be rather pointless.) However, figure 3.2 demonstrates that

$$\frac{x_C}{z_C} = \frac{x_C^\star}{f_C} \tag{3.1}$$

f_C is the distance between the focal point and the camera sensor. The sensor is composed of many tiny pixel sensors. These sensors give the camera sensor a resolution R_C, measured in $^{\mathrm{mm}}/_{\mathrm{px}}$. These sensors convert the image into pixel coordinates with the following relationship:

$$u'_C = \frac{x^\star_C}{R_C} \tag{3.2}$$

We can express a normalized focal length $\bar{f}_C$ in terms of pixels with

$$\bar{f}_C \equiv \frac{f_C}{R_C} \tag{3.3}$$

Combining (3.1) with (3.2) and 3.3, we get

$$u'_C = \bar{f}_C \frac{x_C}{z_C} \tag{3.4}$$

This means that, given the normalized focal length, we can easily calculate θ_C:

$$\begin{aligned} \tan\theta_C &= \frac{x_C}{z_C} && (3.5) \\ \tan\theta_C &= \frac{u'_C}{\bar{f}_C} && (3.6) \end{aligned}$$

Similarly,

$$\begin{aligned} \tan\phi_C &= \frac{v'_C}{\bar{f}_C} && (3.7) \\ \tan\theta_P &= \frac{u'_P}{\bar{f}_P} && (3.8) \\ \tan\phi_P &= \frac{v'_P}{\bar{f}_P} && (3.9) \end{aligned}$$

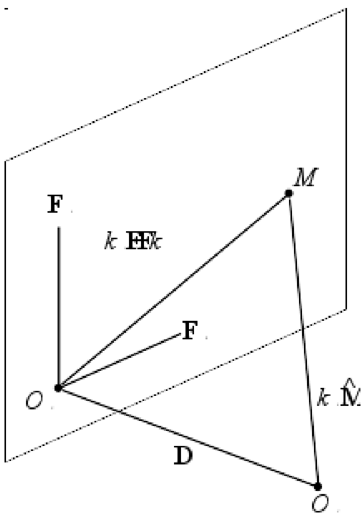

Figure 3.3: Triangulation geometry.

3.2.2 Recovering 3D information

All 3D work in this section is performed using the camera's coordinate system. We start by assuming a distance vector measured from the camera to the projector.

$$\mathbf{D} = \begin{pmatrix} D_x \\ D_y \\ D_z \end{pmatrix}$$

$\hat{\mathbf{M}}$ is the unit vector pointing from the camera to the measured point $\mathbf{M}$.

$$1 = \hat{M}_x^2 + \hat{M}_y^2 + \hat{M}_z^2 \tag{3.10}$$

$$\hat{M}_z = \left[\tan^2\theta_C + \tan^2\phi_C + 1\right]^{-\frac{1}{2}} \tag{3.11}$$

$$\hat{\mathbf{M}} = \begin{pmatrix} \hat{M}_z \tan\theta_C \\ \hat{M}_z \tan\phi_C \\ \hat{M}_z \end{pmatrix} \tag{3.12}$$

Even though the fringe appears as a contour where it illuminates a surface, it may be conceptualized as a plane extending from the projector through every possible point of illumination. Any point on the fringe plane may be identified uniquely and parametrically as a linear combination of two vectors $\mathbf{F_1}$ and $\mathbf{F_2}$.

$$\mathbf{F}\left(t_1, t_2\right) = t_1\mathbf{F_1} + t_2\mathbf{F_2} \tag{3.13}$$

Vertical fringes will always pass through the y-axis, meaning

$$\mathbf{F_1} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

To define the other vector, we can use an angle ψ, defined as

$$\psi \equiv \pi + \theta_P - \beta$$

The physical significance of ψ is that it is the angle formed in the $y = 0$ plane by the fringe plane's trace and the camera's optical axis ($x = y = 0$).

We can now relate $\mathbf{F_2}$ to the projector's image coordinates:

$$\mathbf{F_2} = \begin{pmatrix} \sin\psi \\ 0 \\ \cos\psi \end{pmatrix}$$

Now we have three vectors which sum to $\mathbf{D}$ in some linear combination. k_1 and k_2 are not necessarily positive.

$$k_0\hat{\mathbf{M}} + k_1\mathbf{F_1} + k_2\mathbf{F_2} = \mathbf{D}$$

This may be solved by using the equation

$$\begin{pmatrix} \hat{\mathbf{M}} & \mathbf{F_1} & \mathbf{F_2} \end{pmatrix} \begin{pmatrix} k_0 \\ k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} \mathbf{D} \end{pmatrix}$$

$$\begin{pmatrix} \hat{\mathbf{M}} & \mathbf{F_1} & \mathbf{F_2} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{D} \end{pmatrix} = \begin{pmatrix} k_0 \\ k_1 \\ k_2 \end{pmatrix}$$

Now that we have our constants, the point $\mathbf{M}$ is simply

$$\mathbf{M} = k_0 \hat{\mathbf{M}} \tag{3.14}$$

in the camera's coordinate space.

3.2.3 Converting to surfels

Applying the above techniques to an input image results in a two-dimensional sparse matrix of three-dimensional points. This group of points may be immediately exported as a point cloud for quick verification and visualization of results. However, to be most useful in a CAD package, these points must be merged into a mesh. This merging is best done in the scanning software, as the input image implicitly describes relationships between adjacent points.

Scan Studio uses surfels to represent the surface. A surfel, short for *surf*ace *el*ement, represents a small patch of the object's surface in much the same way that a pixel represents a small patch of a computer screen. A surfel stores location, normal, color, and size.

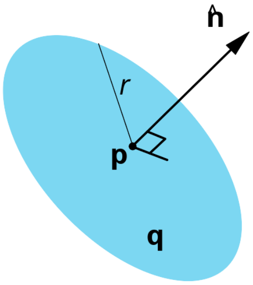

Figure 3.4: A surfel, with location **p**, normal **n̂**, color **q**, and radius r.

3.3 Hardware Selection

3.3.1 Projector

The projector was selected to be low cost, compact, and have a high resolution. The spatial resolution of the scanning system is limited by the resolution of the projector. Additionally the spatial resolution is affected by the distance to the target object, so projectors with a short minimum throw distance were favored. In order to ensure a large field of view at the close range, the projector also needed to have a large throw ratio. Taking all of these parameters into account, the projector that was selected is the Optoma PK201 Pico Pocket Projector. See Appendix A for technical specifications of the projector.

Figure 3.5: Optoma PK201 Pico Pocket Projector

3.3.2 Cameras

The cameras for the scanning system were selected to be low cost and high resolution. The resolution of the cameras needs to be equal to or greater than the resolution of the projector. Since cost was a main factor webcams were chosen for the cameras. The specific model selected is the Logitech C905 webcam. Full specifications on the cameras can be found in Appendix A.

Figure 3.6: Logitech C905 Webcam

3.3.3 Mounting System

In order to secure the cameras and projector in triangulation geometry a mounting system was created. It was important to be able to adjust the distance between the projector and cameras and the triangulation angle of the system. Additionally it was critical that the cameras and projector be able to be locked in place so that they would not move relative to each other after calibration.

The mounting system utilized an L-beam as a mounting rail which the cameras and projector could slide along. A custom mount was designed to hold the projector, and the mount was fixed in place on the rail by a hole drilled through the L-beam. The projector mount had a countersunk screw hole for holding the projector in place, and another one for attaching the assembly to a tripod.

Next the existing camera casings were removed so that only the lenses and circuitry remained. New casings were designed that had holes to mount from the bottom. Two mounts were designed to slide on to the L-beam and clamp into place with a screw hole.

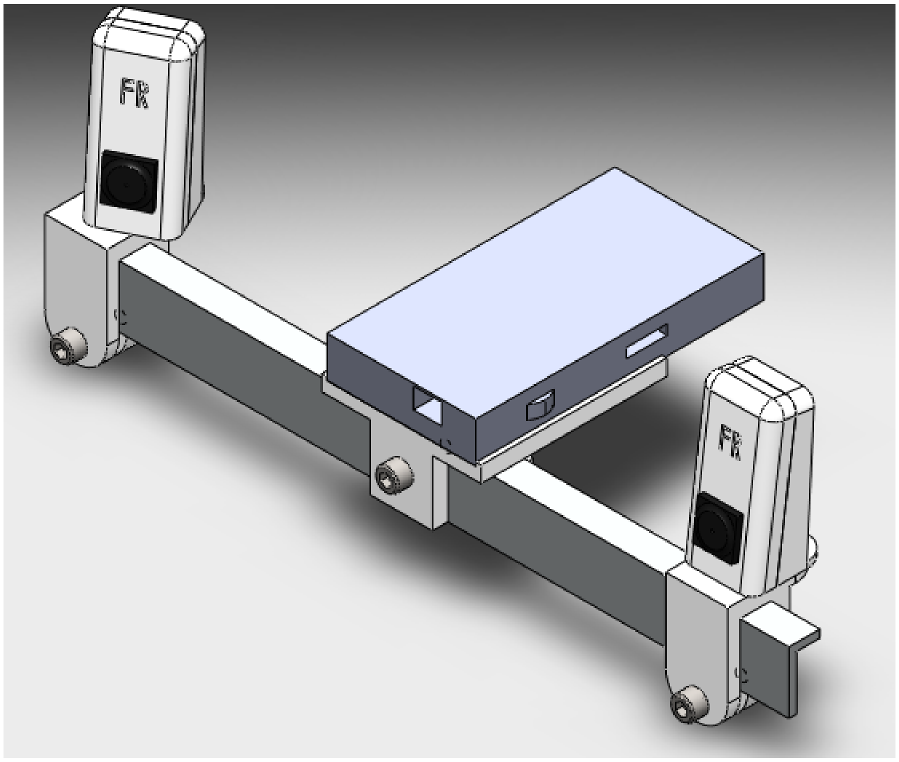

Figure 3.7: CAD Mounting Assembly

The mounts had a hole on top for the new camera casings, which could pivot about the hole until tightened into place. The custom mounts and caseings were then rapid prototyped.

3.4 Overview of Scanning Process

The process of scanning a real world object and using it to make a new object can be broken into 4 distinct steps. These steps are digitization, conversion to mesh, manipulating in CAD software, and manufacturing. Each of these steps will be explained in further detail below.

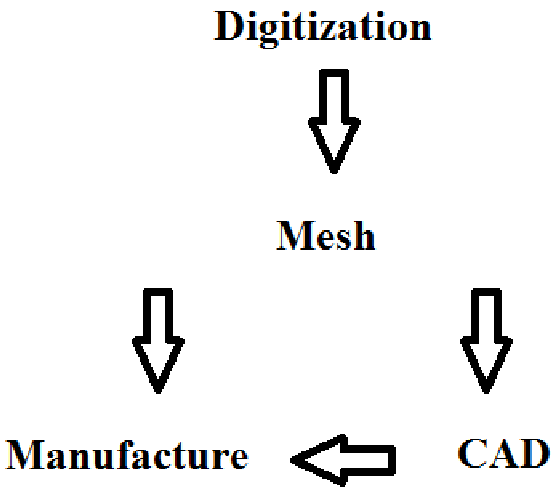

Figure 3.8: Scanning Process Overview

3.4.1 Digitization

Digitization refers to the process of converting the dimensions of the 3-dimensional object into data to be used by computer software. This process can be broken into three distinct sections; setup, data capture, and repositioning.

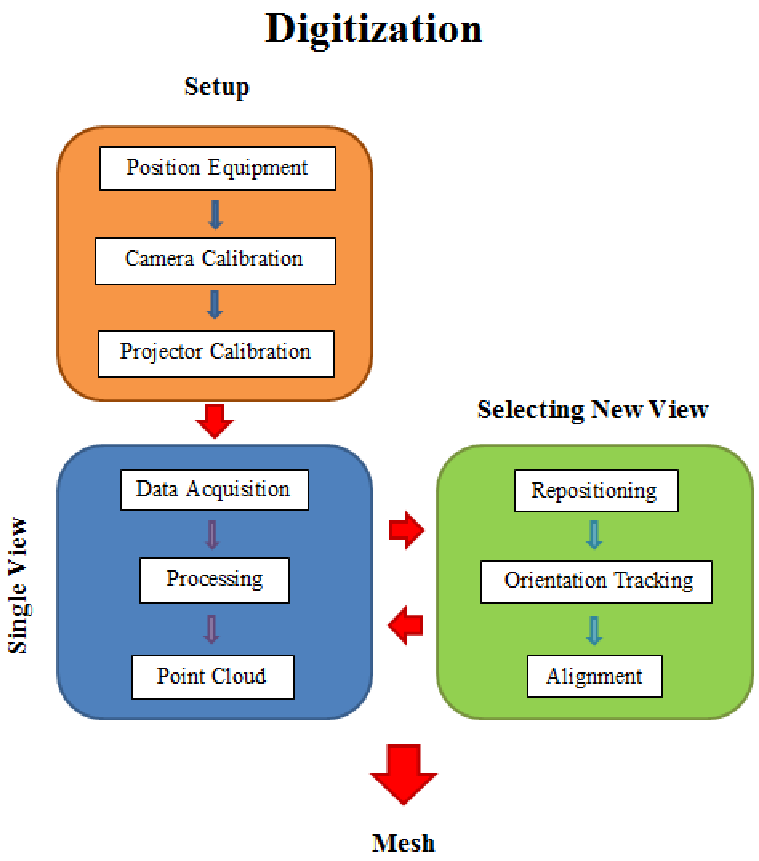

Figure 3.9: Digitization Process Overview

3.4.1.1 Setup

The setup process for scanning is the preparation of the equipment. For this scanning process the necessary equipment are an active projector and two or more digital cameras. The quality of these pieces of equipment are determined by the quality of the results required. The cameras and projector must be placed in a triangulation geometry. Optimal positioning will improve results and is determined by size of the object and the desired accuracy and precision.

Once the equipment is positioned it is necessary to calibrate it. Calibration is used to determine both the intrinsic and extrinsic parameters of the cameras and projector. This means that the calibration serves to both remove the distortions from the lenses of the devices, but also to establish the relative locations and orientations of the different pieces of equipment relative to each other. First the cameras are calibrated simultaneously, and then the projector is calibrated. Detailed descriptions of the calibration procedure can be found _________

3.4.1.2 Data Capture

Data capture can be broken into three distinct steps; data aquisition, data processing, and point cloud generation. The data aquisition is achieved by taking images with the two cameras. In the case of this method 20 frames of data are needed for each data capture view. Once the data is aquired the images are correct for distortions using the calibration data and are then processed using the triangulation algorythm. Once each pixel is processed a point cloud is constructed using the x, y, and z information obtained from the calibration and processing. Each pixel in the picture equates to one point in the point cloud. Background points can be filtered out either before or after the processing into point cloud form.

3.4.1.3 Repositioning

After each data capture view is complete it is necessary to reposition the target object to capture another view. Photogrammetry tracking using ORB matching algorythms is employed to automatically rotate the point cloud to match the object rotation. In this way the next point cloud can be added to the existing in proper alignment. The process of data capture and repositioning is repeated until the point cloud is completed with no holes from shadowing or lack of views.

3.4.2 Meshing

The point cloud obtained during digitization needs to be converted into a mesh in order to be used by other software and for manufacturing.

3.4.3 CAD software

Once the STL mesh is created it is possible to create a CAD model from it. Commercially available CAD packages such as SolidWorks have the capability of creating a CAD file from an STL automatically.(Cite) Depending on the CAD packages algorithm and the complexity of the part there will be a varying level of "intelligence" to the part. This means that the CAD software may have recognized editable features such as holes or extrudes in the model. Once a CAD file is created, the CAD software can be used to heal the geometry of the part in any regions that lacked data, such as the insides of holes. Additionally changes can be made to specific aspects of the part, and new features can be added. Finally analysis software such as FEA can be used with the CAD model.

3.4.4 Manufacture

Once all changes have been made to the 3D model or mesh it can be used to manufacture new parts. The two simplest ways of manufacturing parts from a 3D scan are using rapid prototyping machines, or CNC machining. CNC machining requires additional creating of tool paths in CAM software before the part to be manufactured. Rapid Prototyping machines on the other hand are capable of creating parts from a mesh file, most commonly an STL(cite).

3.5 Calibration

The techniques outlined in the previous sections rely on calibration for accurate reconstruction of a single view. Calibration allows the system to measure both intrinsic parameters (such as focal length) and extrinsic parameters (such as relative position and orientation) of the cameras and projector.

3.5.1 Camera Calibration

3.5.1.1 Calibration Process

Camera calibration consists of determining the intrinsic and extrinsic properties of the cameras. The intrinsic properties of the camera are the focal length and lens distortions. The focal length is a very important parameter to triangulation, so even if the manufacturer gives this value it still must be calculated during calibration in order to ensure the accuracy of the given value. Correcting for lens distortions is critical for obtaining accurate input, and doing so removes distortions from the resulting mesh, vastly improving results. The extrinsic parameters are the location and orientition of the cameras. For simplicity the coordinate system is always defined such that the theoretical pinhole of one of the cameras is at the origin, with the optical axis of that camera aligned with the z-axis of the system. Extrinsic calibration of the cameras allows for stereo vision, which adds to the accuracy of the measurments being taken.

Both intrinsic and extrinsic calibration are carried out simultaneously. To do so it is necessary to have a calibration board, which is a flat board or panel with a calibration pattern on it. The calibration pattern we utilize is a simple checkerboard pattern with a known square size. This square size is input into the software as the only parameter for calibration. Then pictures of the pattern are taken at several different orientations. It is important to use many orientations, and to fill every portion of the image with the pattern at least once if possible. These calibration images are then batched processed by the camera calibration algorithm to generate the camera calibration parameters.

3.5.1.2 Calibration Algorithms

3.5.2 Projector Calibration

3.5.2.1 Calibration Process

After the camera calibration is complete, the projector calibration can be carried out. Just like with camera calibration, projector calibration consists of determining both intrinsic and extrinsic parameters. The intrinsic parameters are once again focal length and distortion, while the extrinsic parameters are the position and orientation of the projector relative to the coordinate system established during camera calibration.

In order to perform projector calibration the calibration board needs to be in a fixed position where it can be seen by both cameras. Then a series of binary patterns are projected onto the calibration board with the stereo camera system taking an image of each pattern. The patterns consist of first vertical stripes and then horizontal stripes. These images are then batch processed by the projector calibration algorithm in order to generate the projector calibration parameters. These parameters are used in conjunction with the camera calibration parameters when processing any images taken with the sytem into meshes.

3.5.2.2 Calibration Algorithms

3.6 Orientation tracking

3.7 Surface reconstruction

Chapter 4

Results

4.1 Scan Studio

Scan studio is the custom software created for this project. Scan studio was programmed in visual C++ and has a graphical user interface (GUI) that is built in the QT framework. Scan studio can be used to run simulations or perform scans.

4.1.1 Calibration

Scan Studio has a built in camera and projector calibration process that utilizes the calibration algorithms detailed in the calibration section. While calibrating the cameras the real time images from both cameras is displayed in the program. Once calibration is complete the calibration results are automatically stored and used in any subsequent scans until they are either cleared or overwritten with a new calibration.

4.1.2 Scanning

Scan Studio creates a single view scan of an object with a single button click. Once the scan button is pressed the software automatically projects the binary patterns and saves images from the cameras of each pattern. Additionally it is possible to chose how many images are taken and averaged together for each pattern. Once the images are all taken they are automatically processed by the triangulation algorithm and a point cloud is generated.

In addition to generating point clouds directly from a scan, Scan Studio is capable of generating point clouds from two other sources. The first is from a series of image files captured externally. The image files must contain the same reflected binary patterns, and the filenames must be formated correctly for Scan Studio to process them. The other source Scan Studio can process images from is simulations. Scan Studio is capable of interfacing with the POV-ray simulation software to generate images of an object with the same projected patterns. These images are then processed in the same way as real scan data.

4.1.3 Object View

Scan studio renders point cloud data in a virtual 3D environment. The point cloud can be viewed from any angle as the coordinate system is rotated. Additionally wireframe models of the cameras and projector are displayed in this view and are located at the positions indicated by the extrinsic calibration. Object view utilizes OpenGL, which is an open source library for rendering 2D and 3D graphics.

Object view is capable of displaying multiple point clouds simultaneously. Additionally it updates the poistion of the existing point clouds based on the movement of the real object through the use of photogrametric tracking.

4.1.4 Meshing

The point cloud data generated by Scan Studio can be exported as an .stl mesh file. The program generates the mesh using algorithms from Point Cloud Library (PCL). PCL is an open source library for processing point clouds.

4.2 Simulation Results

4.3 Scanning Results

Chapter 5

Conclusion

Appendix A

Equipment used

Camera The camera is a Logitech C905 Webcam. Specifications are listed in table A.1.

Resolution	1600 x 1200
Frame rate	Up to 30 fps
Bit depth	Up to 16 bits/pixel
Min Focal Length	50 mm
Sensor cell size	2.8µm x 2.8µm

Table A.1: Camera properties.

Projector The projector is an Optoma PK201 Pico Pocket Projector. Specifications are listed in table A.2.

Projection Type	DLP
Native Resolution	854 x 480
Max Resolution	1280 x 800
Brightness	20 ANSI Lumens
Contrast Ratio	2000:1
Light Source	LED
Throw Ratio	2.2:1
Projection Distance	10” -126”
Weight	0.35lb
Dimensions	2.4”x0.7”x4.6”

Table A.2: Projector properties.[13]

Appendix B

Open Source Libraries

B.1 OpenCV

OpenCV is an open source computer vision library created by Intel in 1999 for the purpose of advancing the computer vision field(cite). OpenCV has a BSD software license, which is a minimally restricting license. OpenCV source code is free to use for academic and commercial purposes so long as the proper acknowledgments and copyrights are used. The OpenCV library contains a vast amount of source code for various computer vision application, and is supported by a large open source community.

OpenCV code was used in this project as part of the scanning program. The main use of OpenCV is in the calibration software. The camera calibration code for locating the corners on a checkerboard pattern was created using OpenCV source code. The other optional calibration patterns supported by the software also use OpenCV code. Additionally the solvePNP code used to match features for stereo calibration also utilizes OpenCV. The use of existing OpenCV source code to perform these complicated and vital tasks allowed for time to be focused on other areas of the development. This of course freed up vital time and allowed for other areas of the software to be more robust. The OpenCV stereo calibration code was also modified and re-purposed for the projector calibration. In this way the OpenCV library is responsible for most of the calibration code. Additionally, OpenCV code is used for the photogrammetry tracking used to track the orientation of the object being measured. In this way the OpenCV libraries form a crucial backbone for our software, working in support of our triangulation algorithm and binary projection code.

B.2 OpenGL

B.3 PCL

Index

Bibliography

[1] Imaging information.

[2] Arne Bernin. Kinect chess board meta pattern. web page, November 2010.

[3] Frank Chen, Gordon M. Brown, and Mumin Song. Overview of three-dimensional shape measurement using optical methods. *Optical Engineering*, 39(1):10–22, 2000.

[4] Z Corperation. 3d scanners: Selection criteria for common applications. Technical report, T. A. Grimm and Associates Inc., 2012.

[5] Creaform. Handy scan 3d. Online, 2012.

[6] Barak Freedman, Alexander Shpunt, Meir Machline, and Yoel Arieli. Depth mapping using projected patterns, 2009.

[7] P.S. Huang, Q. Hu, F. Jin, and F.P. Chiang. Color-encoded digital fringe projection technique for high-speed three-dimensional surface contouring. *Optical Engineering*, 38(6):1065–1071, 1999.

[8] NextEngine Inc. Next engine. Online, 2012.

[9] David LaserScanner. David laserscanner. Online, 2012.

[10] S. Lei and S. Zhang. Digital sinusoidal fringe pattern generation: defocusing binary patterns vs focusing sinusoidal patterns. *Optics and Lasers in Engineering*, 48(5):561–569, 2010.

[11] Microsoft. Kinect for windows. web page, 2012.

[12] Edmund Optics. Laser scanner. website.

[13] Optoma. *Optoma Pico Pocket Projector PK201 Product Sheet.* Projector Centeral.

[14] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado. A state of the art in structured light patterns for surface profilometry. *Pattern recognition*, 43(8):2666–2680, 2010.

[15] J. Salvi, J. Pages, and J. Batlle. Pattern codification strategies in structured light systems. *Pattern Recognition*, 37(4):827–849, 2004.

[16] K. N. Snavely. *Scene reconstruction and visualization from internet photo collections.* PhD thesis, University of Washington, 2008.

[17] Mitsuo Takeda and Kazuhiro Mutoh. Fourier transform profilometry for the automatic measurement of 3-d object shape. *Appl. Opt.*, 22(24):3977–3982, Dec 1983.

[18] Michael J. Zervas. Development of a high-speed, robust system for full field-of-view 3d shape measurements. Master's thesis, Worcester Polytechnic Institute, August 2011.