2011

ENHANCEMENTS TO THE MODIFIED COMPOSITE PATTERN METHOD OF STRUCTURED LIGHT 3D CAPTURE

Charles Joseph Casey
University of Kentucky, ccasey0@hotmail.com

ABSTRACT OF DISSERTATION

Charles Joseph Casey

The Graduate School

University Of Kentucky

2011

ENHANCEMENTS TO THE MODIFIED COMPOSITE PATTERN METHOD OF STRUCTURED LIGHT 3D CAPTURE

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Engineering
at the University of Kentucky

By
Charles Joseph Casey

Lexington, Ky

Director: Dr. Laurence G. Hassebrook, Professor of Electrical and Computer Engineering

Lexington, Ky

2011

ABSTRACT OF DISSERTATION

ENHANCEMENTS TO THE MODIFIED COMPOSITE PATTERN METHOD OF STRUCTURED LIGHT 3D CAPTURE

The use of structured light illumination techniques for three-dimensional data acquisition is, in many cases, limited to stationary subjects due to the multiple pattern projections needed for depth analysis. Traditional Composite Pattern (CP) multiplexing utilizes sinusoidal modulation of individual projection patterns to allow numerous patterns to be combined into a single image. However, due to demodulation artifacts, it is often difficult to accurately recover the subject surface contour information. On the other hand, if one were to project an image consisting of many thin, identical stripes onto the surface, one could, by isolating each stripe center, recreate a very accurate representation of surface contour. But in this case, recovery of depth information via triangulation would be quite difficult. The method described herein, Modified Composite Pattern (MCP), is a conjunction of these two concepts. Combining a traditional Composite Pattern multiplexed projection image with a pattern of thin stripes allows for accurate surface representation combined with non-ambiguous identification of projection pattern elements. In this way, it is possible to recover surface depth characteristics using only a single structured light projection.

The technique described utilizes a binary structured light projection sequence (consisting of four unique images) modulated according to Composite Pattern methodology. A stripe pattern overlay is then applied to the pattern. Upon projection and imaging of the subject surface, the stripe pattern is isolated, and the composite pattern information demodulated and recovered, allowing for 3D surface representation.

In this research, the MCP technique is considered specifically in the context of a Hidden Markov Process Model. Updated processing methodologies explained herein make use of the Viterbi algorithm for the purpose of optimal analysis of MCP encoded images. Additionally, we techniques are introduced which, when implemented, allow fully automated processing of the Modified Composite Pattern image.

Keywords: Composite Pattern, Structured Light, 3D, image processing, scan

_____Charles Joseph Casey_____

_____May 06, 2011_____

ENHANCEMENTS TO THE MODIFIED COMPOSITE PATTERN METHOD OF STRUCTURED LIGHT 3D CAPTURE

By

Charles Joseph Casey

_____Laurence G. Hassebrook_____
(Director of Dissertation)

_____Stephen Gedney_____
(Director of Graduate Studies)

_____May 06, 2011_____

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name	Date

DISSERTATION

Charles Joseph Casey

The Graduate School

University Of Kentucky

2011

ENHANCEMENTS TO THE MODIFIED COMPOSITE PATTERN METHOD OF STRUCTURED LIGHT 3D CAPTURE

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Engineering
at the University of Kentucky

By
Charles Joseph Casey
Lexington, Ky

Director: Dr. Laurence G. Hassebrook, Professor of Electrical and Computer Engineering
Lexington, Ky

2011

ACKNOWLEDGEMENTS

I'd like to thank Dr. Laurence Hassebrook for all of his efforts as my advisor and mentor during my time as a researcher. He has endured a lot from me, and has gone through a great deal of trouble on my behalf. I will remember this.

I would also like to thank each member of my dissertation defense committee; Dr. Kevin Donohue, Dr. Daniel Lau, Dr. Brent Seales, and examiner Dr. William Murphy, for their patience and assistance. I hope that they have found my efforts satisfactory, pleasing, and worthwhile.

And I thank the staff throughout the University of Kentucky, and especially that of the ECE department and Center for Visualization, whose efforts and kindnesses have not gone unappreciated.

I'd also like to acknowledge and thank every professor who has instructed me over the years, perhaps especially the professors at Georgetown College, who were exceptionally supportive during my early, formative years of study.

Thanks go also to my family. They have made things most pleasant for me.

And most of all, thanks go to my God YHWH, the creator, the one who IS, without whom I could have accomplished, literally, nothing at all.

Thank you, everyone.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS III

LIST OF TABLES VI

LIST OF FIGURES VII

CHAPTER 1 INTRODUCTION 1

1.1 3D ACQUISITION 1
1.2 STRUCTURED LIGHT ILLUMINATION 5
1.3 SINGLE-FRAME TECHNIQUES 15
1.4 MOTIVATIONS 27
1.5 CONTRIBUTIONS 30
1.6 OUTLINE 34

CHAPTER 2 MODIFIED COMPOSITE PATTERN 36

2.1 INTRODUCTION TO MCP 36
2.2 MODULATION THEORY 40
2.3 SYSTEMATIC INTERFERENCE ANALYSIS 44

CHAPTER 3 CAMERA MODEL, TRIANGULATION, AND CALIBRATION 56

3.1 WORLD TO CAMERA 3D TRANSFORMATION 56
3.2 CAMERA 3D TO CAMERA 2D TRANSFORMATION 57
3.3 WORLD TO CAMERA 2D TRANSFORMATION 58
3.4 CALIBRATION 59
3.5 WORLD 3D RECOVERY 61
3.6 CALIBRATION PROCESS SUMMARY 63
3.7 SPECIAL MODIFIED COMPOSITE PATTERN CALIBRATION PROCEDURE 65

CHAPTER 4 MCP PROCESSING 67

4.1 PEAK ISOLATION 68
4.2 SURFACE INCLINATION SEGMENTATION 73
4.3 ANALYSIS-BY-SYNTHESIS TARGET FREQUENCY ISOLATION 77
4.4 CHANNEL ISOLATION 80
4.5 THRESHOLDING AND INITIAL ERROR ESTIMATION 82

CHAPTER 5 ERROR CORRECTION AND POST-PROCESSING 91

5.1 BACKGROUND INFORMATION 91
5.2 MCP HIDDEN MARKOV MODEL 94
5.3 THE VITERBI BEST PATH ALGORITHM 99
5.4 POST-PROCESSING OPERATIONS 101
5.5 RESULTS 105
5.6 EFFECTIVENESS OF PROCESSING 118

CHAPTER 6 CONCLUSION AND FUTURE WORK 128

6.1 FUTURE WORK 129

APPENDIX A: NOMENCLATURE, TERMINOLOGY, AND MATHEMATICAL SYMBOLS REFERENCE 131

REFERENCES 133
VITA 147

LIST OF TABLES

TABLE 2.1, 92 CYCLES/FOV - SINGLE FREQUENCY MODULATED 54
TABLE 2.2, 92 CYCLES/FOV - FOUR FREQUENCIES MODULATED 54
TABLE 2.3, 115 CYCLES/FOV - FOUR FREQUENCIES MODULATED 55
TABLE 5.1, TESTING MCP DECODING PROCESS - FLAT BOARD 119
TABLE 5.2, TESTING MCP DECODING PROCESS - SPHERE 120
TABLE 5.3, ERROR RESULTS (SPHERE-MCP METHOD) 122
TABLE 5.4, ERROR RESULTS (SPHERE-MCP METHOD-ISOLATED SURFACE ONLY) 123
TABLE 5.5, ERROR RESULTS (BOARD-MCP METHOD) 123
TABLE 5.6, ERROR RESULTS (GENERATION 1 CP VS. 4 PATTERN PMP) 126
TABLE 5.7, ERROR RESULTS (GENERATION 3 MCP VS. MULTI-FREQUENCY PMP) 126

LIST OF FIGURES

FIGURE 1.1, HIERARCHY OF 3D ACQUISITION DEVICES 2
FIGURE 1.2, EXAMPLES OF CONTACT SCANNERS [21-24] 3
FIGURE 1.3, PATTERN DISTORTION DUE TO SUBJECT SURFACE 7
FIGURE 1.4, SLI SYSTEM GEOMETRY [32] 7
FIGURE 1.5, GRAY CODE PATTERN AS USED IN SLI [38] 10
FIGURE 1.6, DE BRUIJN SEQUENCE MULTI-FRAME SCAN EXAMPLE [39] 11
FIGURE 1.7, (LEFT) EXAMPLE PMP PROJECTION CROSS SECTION AND (RIGHT) RESULTING PHASE MAP CROSS SECTION FROM IMAGE ANALYSIS [42] 12
FIGURE 1.8, PMP 3D ACQUISITION RESULT EXAMPLE [32] 13
FIGURE 1.9, AMBIGUITY IN PROJECTED PATTERN DUE TO SURFACE SHAPE [49] 16
FIGURE 1.10, (RIGHT) SHOWS DISTORTION DUE TO IMPERFECTIONS IN PROJECTOR/CAMERA COLOR FILTRATION OF (LEFT) AN IDEAL COLOR MULTIPLEXED 3 CHANNEL PROJECTION PATTERN 19
FIGURE 1.11, EXAMPLE OF ERRORS INTRODUCED TO 3 PATTERN SLI DUE TO COLOR INTERFERENCE 20
FIGURE 1.12, COMPOSITE PATTERN (LEFT) PROJECTED ONTO A HUMAN SUBJECT AND (RIGHT) ANALYZED TO CREATE A DEPTH MAP [32] 21
FIGURE 1.13, COMPOSITE PATTERN (LEFT) PROJECTED ONTO A HUMAN HAND AND (RIGHT) ANALYZED TO CREATE A DEPTH MAP [32] 21
FIGURE 1.14, (TOP) IDEAL RECONSTRUCTION OF A SMALL, FLAT DISC, (BOTTOM) COMPOSITE PATTERN SCAN RESULT AFTER A SINGLE POST-PROCESSING CORRECTION ITERATION [32] 23
FIGURE 1.15, (LEFT) EXAMPLE OF A HOLD PATTERN PROJECTION AND (RIGHT) RESULTING "SNAKES" 25
FIGURE 1.16, (LEFT) LOCK SCAN AND (RIGHT) SEQUENCE OF RECONSTRUCED FRAMES FROM HOLD IMAGES 26
FIGURE 1.17, MCP SCAN OF A LATEX BUST, (LEFT) MCP DATA POINTS WITH INTERPOLATED LINEAR FILL REGIONS, (RIGHT) RESULT AFTER SMOOTHING AND OTHER POST-PROCESSING STEPS 29
FIGURE 1.18, ADAPTIVE THRESHOLD GENERATION 31
FIGURE 1.19, EXAMPLE OF STRIPE ROTATION EFFECTS 32
FIGURE 1.20, (LEFT) ISOLATED POSITIVE INCLINATION STRIPE REGIONS, (RIGHT) ISOLATED NEGATIVE INCLINATION STRIPE REGIONS 32
FIGURE 1.21, ANALYSIS BY SYNTHESIS PROCESS, (TOP) INITIAL ARRAY A, (MIDDLE) ARRAY B, WHERE B REPRESENTS ISOLATED FREQUENCY BAND COMPONENTS, AFTER ONE ITERATION, (BOTTOM) UPDATED ARRAY C WHERE C=A-B, WHERE THE PRESENCE OF B IS INDICATED BY THE DOTTED CURVE 33
FIGURE 2.1, VISUALIZATION OF BINARY PATTERN ENCODING 37
FIGURE 2.2, AN EXAMPLE OF A MODIFIED COMPOSITE PATTERN 38
FIGURE 2.3, (A) INPUT TEST IMAGE, (B) WINDOWED INPUT, (C) WINDOWED DEMODULATED REGION, (D) INVERSE-WINDOWED DEMODULATED REGION 53
FIGURE 3.1, SIMPLIFIED CAMERA MODEL (Y DIRECTION) 57
FIGURE 3.2, CALIBRATION SOFTWARE 64
FIGURE 4.1, MCP FLOWCHART 67

FIGURE 4.2, ADAPTIVE THRESHOLD GENERATION 69
FIGURE 4.3, ENHANCED IMAGE 71
FIGURE 4.4, THRESHOLDED ISOLATED REGION IMAGE 71
FIGURE 4.5, EXAMPLE OF STRIPE ROTATION EFFECTS 73
FIGURE 4.6, (LEFT) ISOLATED POSITIVE INCLINATION STRIPE REGIONS, (RIGHT) ISOLATED NEGATIVE INCLINATION STRIPE REGIONS 76
FIGURE 4.7, FOUR QUADRANT COMPOSITE FOURIER IMAGE 78
FIGURE 4.8, VERTICAL PSD SUMMATION 79
FIGURE 4.9, ANALYSIS BY SYNTHESIS PROCESS, (TOP) INITIAL ARRAY A, (MIDDLE) ARRAY B, WHERE B REPRESENTS ISOLATED FREQUENCY BAND COMPONENTS, AFTER ONE ITERATION, (BOTTOM) UPDATED ARRAY C WHERE C=A-B, WHERE THE PRESENCE OF B IS INDICATED BY THE DOTTED CURVE 80
FIGURE 4.10, FREQUENCY ISOLATION COMPARISON OF METHODS EXAMPLE 1, UNSHIFTED (LEFT) VERSUS SHIFTED (RIGHT) 81
FIGURE 4.11, FREQUENCY ISOLATION COMPARISION OF METHODS EXAMPLE 2, UNSHIFTED (LEFT) VERSUS SHIFTED (RIGHT) 81
FIGURE 4.12, PEAK THRESHOLDING EXAMPLE 86
FIGURE 5.1, BASIC SYMBOL SEQUENCE 97
FIGURE 5.2, SYMBOL SEQUENCE WITH ERROR STATES 97
FIGURE 5.3, SYMBOL SEQUENCE WITH ERRORS AND MISSES 98
FIGURE 5.4, INPUT IMAGE 105
FIGURE 5.5, CONTRAST ENHANCED INPUT IMAGE 106
FIGURE 5.6, STRIPE REGION ISOLATED IMAGE 107
FIGURE 5.7, ISOLATED STRIPES 107
FIGURE 5.8, SPECTRAL REGION 108
FIGURE 5.9, PLOT OF FREQUENCY BANDS (MAGNITUDE VS. FREQUENCY IN CYCLES/FOV) AS USED IN ABS AUTOMATED FREQUENCY SELECTION 109
FIGURE 5.10, INDIVIDUAL ISOLATED BINARY REGION INFORMATION IMAGES 110
FIGURE 5.11, INITIAL RECOMBINED REGION/STRIPE INFORMATION ("CODE IMAGE") 111
FIGURE 5.12, CODE IMAGE AFTER INITIAL HMM PROCESSING 112
FIGURE 5.13, CODE INFORMATION RECOVERED FOR MODULATED AND UNMODULATED STRIPE REGIONS 113
FIGURE 5.14, FULLY RECOVERED CODE INFORMATION, INCLUDING PRESENCE OF 2X AND 4X STRIPES 114
FIGURE 5.15, SINGLE COLUMN PLOT OF PHASE INFORMATION 115
FIGURE 5.16, (LEFT) PMP GROUND TRUTH 3D, (RIGHT) RECOVERED MCP 3D 116
FIGURE 5.17, EDGE VIEW OF PMP BOARD 116
FIGURE 5.18, EDGE VIEW OF MCP 3D 117
FIGURE 5.19, (LEFT) MCP RESULT FOR A MOUNTED WHITE SPHERE, (RIGHT) MASKED PMP RESULTS FOR THE SAME SPHERE 121
FIGURE 5.20, (LEFT) ISOLATED SURFACE REGION OF AN MCP SCAN OF A MOUNTED WHITE SPHERE, (RIGHT), ISOLATED SURFACE REGION OF A PMP SCAN OF A MOUNTED WHITE SPHERE 123
FIGURE 5.21, PMP (LEFT) VERSUS DEMODULATED CP (RIGHT). INDIVIDUAL PATTERNS ARE SHOWN IN (A-D), RESULTING IN THE CALCULATED PHASE IMAGE (E). COLOR REPRESENTS INTENSITY 125

FIGURE 5.22, (LEFT) MCP SCAN IMAGE OF THE SCUPLTURE, (RIGHT) ALBEDO IMAGE OF THE SCULPTURE 126
FIGURE 5.23, MCP SCAN RESULTS FOR A HUMAN SCULTPURE - MULTIPLE VIEWS..... 127

Chapter 1 Introduction

Structured Light Illumination (SLI) techniques for three-dimensional (3D) depth measurement have been in development since the early twentieth century [1, 2]. With the introduction of computing machinery, SLI data acquisition has been an active area of image processing research, and a wide range of patterns and codification strategies have been developed. Today, there is widespread demand for technologies with the ability to acquire a digital representation of the three 3D form of a real object [3-16]. In the entertainment industry, special effects often demand computer graphics (CG) representations of objects. Manufacturing concerns will often make use of physical models to generate design data for parts and products or for quality control purposes [4, 5]. Archivists, artists, and historical scholars can benefit from 3D models of various objects, artifacts, or sculptures [10]. 3D imaging will often be used in the field of robotics [6-8], and even those who work in the art of photography, at both professional and consumer levels, may wish to simply extend the scope of their work by integration of real 3D data. In order to satisfy these demands, a highly varied assortment of technical solutions has been developed with the aim of fast, accurate, and reliable 3D data acquisition.

1.1 3D Acquisition

The classes of available technologies can be roughly categorized according to a simple hierarchy, as shown in Figure 1.1. The first distinction is drawn between contact scanning, which requires physically touching the object; and non-contact scanning.

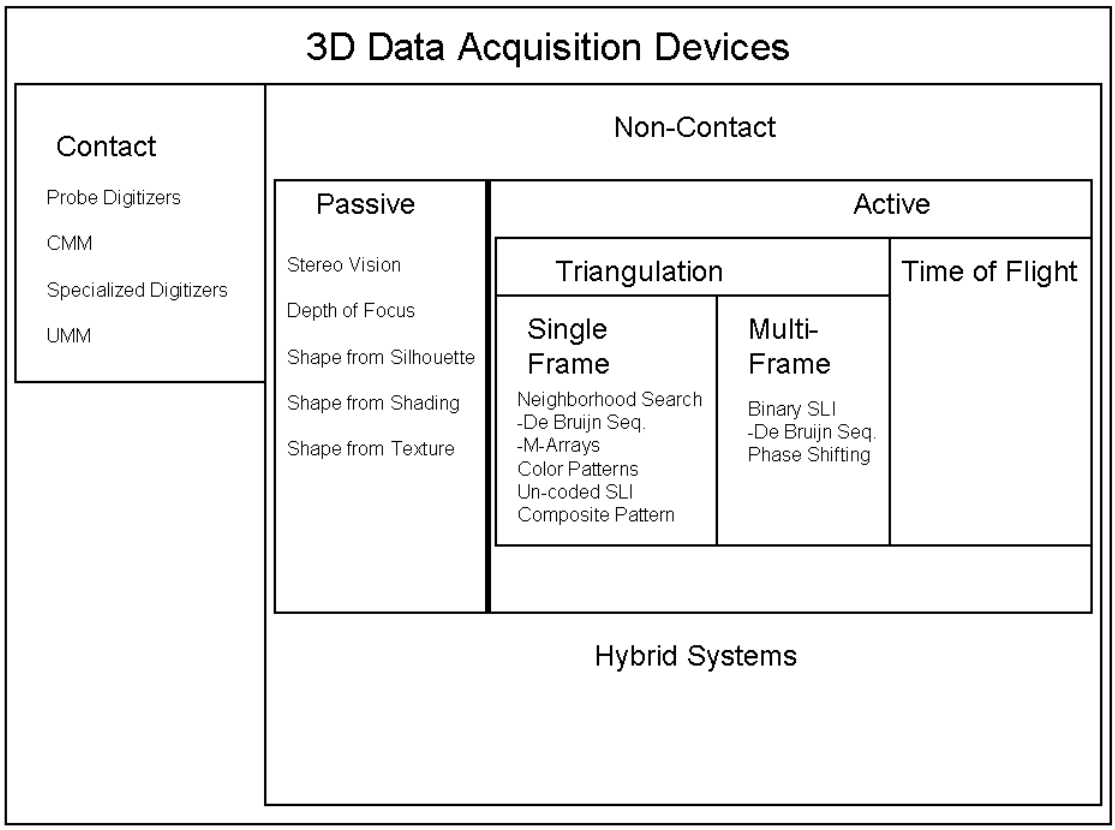

Figure 1.1, Hierarchy of 3D acquisition devices

Contact 3D scanning machines, like all scanning machines, are highly varied in capacity and characteristics. Large, industrial Coordinate Measuring Machines (CMM) and Universal Measuring Machines (UMM) allow extremely precise and dense measurements, at the cost of highly specialized operation and long scan times [17, 18]. Stand-alone measuring probes [19, 20] have been developed for use with common machine tools to allow similar capabilities on smaller scales as well, yet are subject to many of the same drawbacks as the larger devices. Hand-operated CMM probes [21, 22] allow data collection in a quick and intuitive way, but due to their manual nature, are often incapable of acquiring dense data. That is, the highest possible scan density will often be on the order of centimeters, rather than the sub-millimeter precision possible

with many other devices. Additionally, specialized scanners exist, often to acquire data regarding very specific objects, such as biological surfaces [23]. Examples of the wide variety of contact scanning devices can be seen in Figure 1.2.

Many of the weaknesses of contact scanners have been mentioned, such as the lack of measurement resolution. The slow acquisition rate of most contact scanners is due primarily to their inability to collect data for multiple points simultaneously. Additionally, a contact digitizer must actually touch the surface of the scanned object. They may even damage the object. Additionally, measurement on a soft, deformable surface (surfaces such as gel, foams, or hair) may be rendered difficult or impossible.

Figure 1.2, Examples of contact scanners [21-24]

Non-contact scanners attempt to overcome the weaknesses of contact digitizers, often with the use of cameras and illumination devices. Non-contact 3D acquisition devices are most frequently classified as either passive types, which utilize only cameras with no additional radiation source, or active types, which require a specialized illumination device to acquire relevant surface data.

Passive acquisition devices can be further categorized based on their primary method of operation. Perhaps the most common and well-known type of passive 3D acquisition is stereo-vision. These stereo-vision digitizers (largely inspired by biological stereo-vision) attempt to represent an object's 3D form based on data acquired by two (or more) cameras. These stereo-vision systems usually operate by finding the corresponding points on the subject surface in each available image, and then (utilizing the known positions of the cameras which produced these images) geometrically calculating the true position of that surface point relative to each camera pair in a process known as "triangulation." A different well-known passive technology is depth-of-focus [25-28] acquisition, in which the depth of each point on a surface is found by capturing numerous images of the subject under a series of different optical focus settings. A point will be perfectly in focus (a characteristic calculated via any number of image processing techniques) only at a particular camera lens focal length, which is well-defined for each image. This focal length corresponds to a specific distance from the camera lens, thereby allowing one to calculate the depth of each point on the object surface. Still other methods make use of different characteristics of the subject [29, 30] (such as object silhouette, texture, shading, or motion) to attempt to define the 3D surface information.

As effective as these methods are, they are often subject to many of the same weaknesses. Being, as they are, entirely image based, they are greatly subject to the surface characteristics of the target. A major problem in stereo-vision research, for example, is the problem of finding corresponding points in multiple images for surfaces with very little in the way of identifiable features.

Active scanning methods offer an alternative solution. Rather than the capture-only based approach of passive techniques, active scanners use a transmit-capture method. That is, each scan consists of data transmitted (in the form of electromagnetic radiation, and most often visible light), reflected from the object, and then received and processed. Within this there are again, two most common categories: time-of-flight devices where 3D surface data is calculated from the time-difference between transmission and reception of light (as used in LIDAR and laser rangefinder systems), and active triangulation based devices, in which a pattern is projected upon an object (by a laser, projector, or other emitter device), deformed by the subject surface, and then recovered and analyzed in an image.

1.2 Structured Light Illumination

Structured Light Illumination is one of the most accurate non-contact surface scanning methods under development today. It is a commonly used method in scientific and industrial applications [13-16] because of its high degree of accuracy and scalability to different object sizes. The concept is to project a structured pattern of light onto the target surface and extract measurements of depth from the amount of deviation that the reflected light pattern undergoes. In actual application, SLI patterns have become highly sophisticated and are commonly used in a sequence of multi-patterns or even multi-

patterns of different colors. The state-of-the-art techniques can be used in the presence of ambient light, yield non-ambiguous depth, and when processed result in a high density of accurate 3-D point measurements. SLI techniques, when compared to the passive techniques, are more easily able to overcome the fundamental ambiguities which generally occur with low texture targets.

Structured light illumination (SLI) allows one to measure the depth information of a surface by measuring the deformation in a projected light pattern [31]. A simple example would be a pattern of stripes projected onto a sphere. When viewed obliquely, the light stripes on the sphere appear curved as shown in Figure 1.3. For a given arrangement of the projector and camera, the variation in a pattern can be characterized extremely accurately, such that a precise model of the surface can be reconstructed. Most modern implementations of SLI systems make use of digital projectors to illuminate the subject and a digital camera to capture an image of the illuminated subject, though in certain cases static projection devices (slide projectors, for example) may be used.

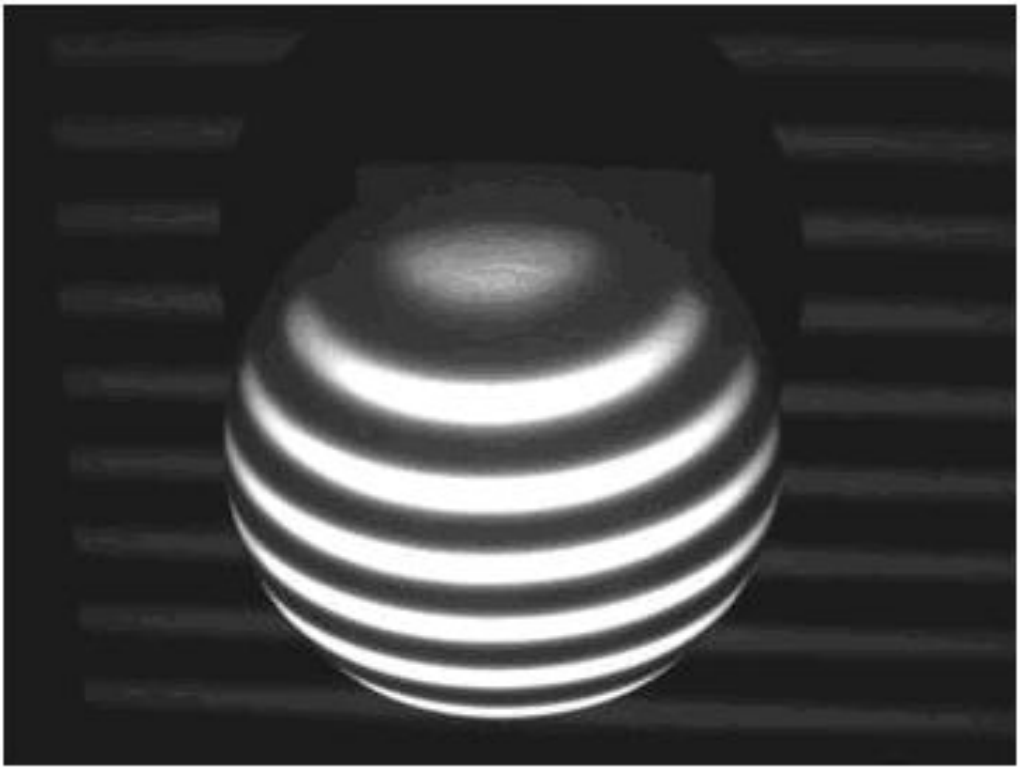

Figure 1.3, Pattern distortion due to subject surface

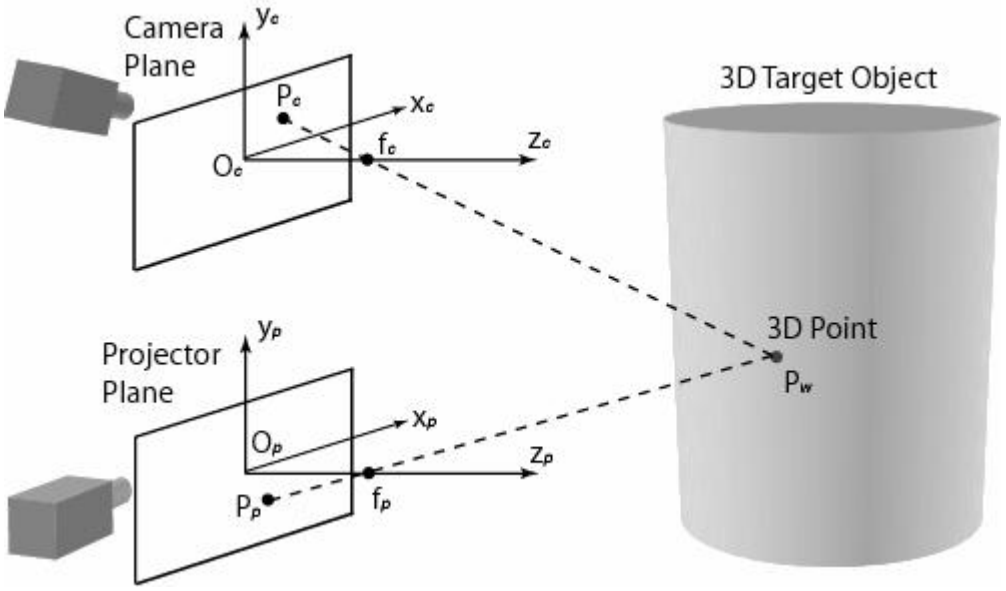

Figure 1.4, SLI system geometry [32]

Mathematically, the SLI measurement process is based on triangulation. Accurate results can be produced only when there is a well defined relationship between a single point on the projection plane and the corresponding point on the captured image, as indicated in Figure 1.4. It is to establish this relationship that projection patterns are utilized. A projection pattern is designed such that each pixel (or row or column, depending on the specific implementation) of the projection image is uniquely characterized, either by some characteristic intensity value sequence or by some other identifiable property such as a local pattern of shapes or colors. When projected onto a subject, the captured image can be analyzed to locate these identifiable projection pattern points. Given a fixed placement of camera and projector, the location of any given pattern point on the subject creates a unique triangle with dimensions defined by the depth of the subject surface.

Structured Light systems are further classified into single-frame techniques, which require only one image capture to calculate surface characteristics, and multi-frame (or "time multiplexed") techniques, which require multiple projection/capture instances in order to acquire enough information for surface reconstruction.

As they require numerous images captured sequentially, multi-frame techniques are often unsuited for subjects that are likely to move, such as living things. Movement during the capture process will normally introduce error, and often very significant measurement error can be introduced by only a small amount of subject motion. However, multi-frame techniques also offer extremely high measurement resolution (measurement precision is often on the order of a fraction of a millimeter) and relatively fast scans when compared to contact measurement methods (a multi-frame SLI scan can

take anywhere from minutes for a laser-based system, to three seconds or less for a digital light projector (DLP) structured light system).

Binary SLI

Multi-frame systems are further classified according to the pattern projection and analysis methodology used to acquire depth information. Proposed in the late twentieth century [33], binary acquisition techniques utilize multiple projections of simple black and white stripe patterns. In each captured image, any point on the subject will either be fully illuminated or fully without illumination. Over the course of a sequence of projections, the sequence of illumination states for each point correspond to a binary "code word" which can be associated with a specific projector region. This association allows for triangulation and thus, depth measurement. In order to identify N regions, one must project log2 (N) patterns.

Practical limitations on the lateral resolution of such a system are, however, imposed due to projector and camera pixel size as well as the resolution capabilities of the analysis software. Thus it may be impractical to uniquely encode each individual projector column or row with a binary pattern.

Based on the standard binary pattern concept, Gray Code patterns utilize a different set of code words such that each pattern region differs from the adjacent region by only one bit [34-37] Using this prior knowledge, analysis becomes more robust against noise and region identification is therefore more accurate.

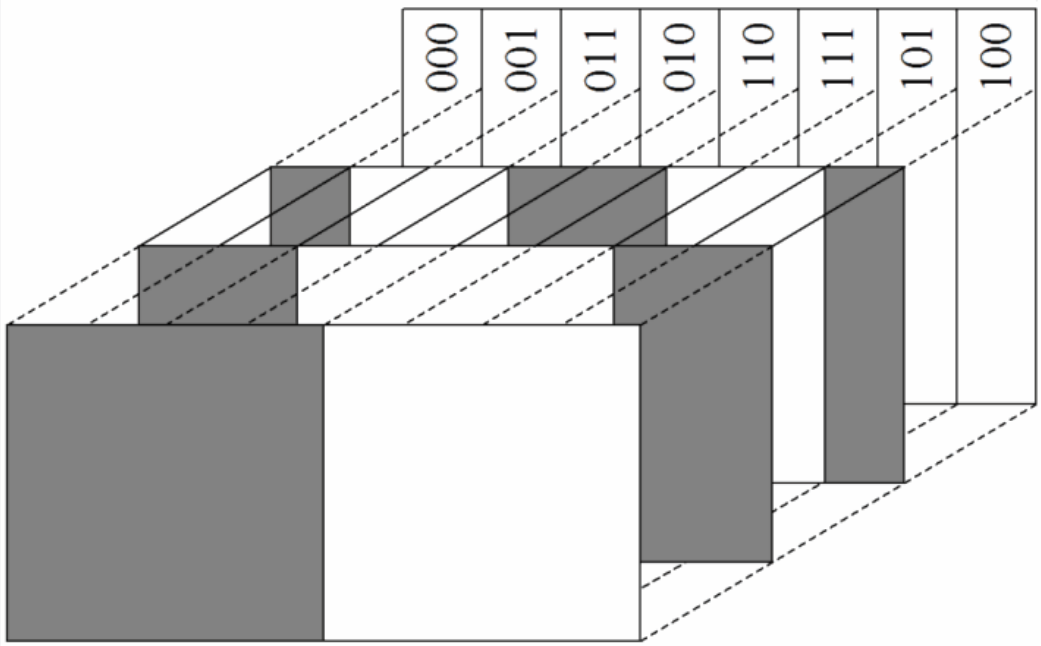

Figure 1.5, Gray code pattern as used in SLI [38]

An alternative binary system introduces De Bruijn sequence encoding, that is, binary sequences wherein a subsequence of a given length occurs only once. They are often employed in the form of stripe boundary codes [39, 40], that is, they define information only at the boundary of the binary stripes (unlike standard binary encoding which use the centers of the stripes, for example). Stripe boundaries inherently carry two bits of information (the high/low value of both adjacent regions). De Bruijn sequences allow one to properly define the binary stripe sequences necessary to unambiguously define unique codewords at each stripe boundary, allowing greater scan density with fewer pattern projections when compared to standard binary SLI.

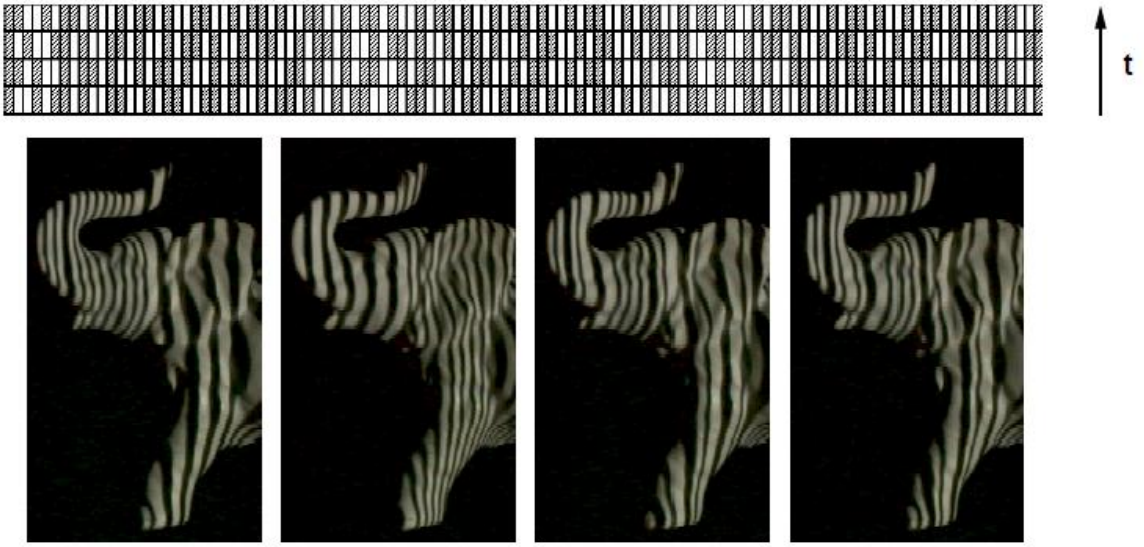

Figure 1.6, De Bruijn sequence multi-frame scan example [39]

Phase Shifting Techniques

Phase shifting techniques utilize patterns which are shifted spatially by a certain known amount in each frame of projection. By utilizing the series of images captured during each frame of this shifting process, it is possible to calculate a unique identity (often represented as a "phase" measurement, corresponding to values normalized between 0 and 2pi) for each pixel in the projected image, and thereby, achieve the necessary pixel-wise correspondence between projection and capture necessary to complete triangulation.

Phase Measuring Profilometry (PMP) [41], is exemplary of the phase shifting paradigm. In this case, the projection pattern consists of a gray level field which varies sinusoidally. In each subsequent projection, the field is spatially shifted in the direction of variation for one sinusoidal period. Due to the continuously varying nature of the

pattern, each projector row or column may be encoded uniquely and identified according to the equation

$$I_n(x^p, y^p) = A + B\cos\left(2\pi f\, y^p - 2\pi n/N\right) \tag{1.1}$$

where I represents the intensity value of any projector pixel at location (x^p, y^p), A and B are constants, N represents the total number of phase shifts to be performed, n is an index defining the frame in a projection sequence, and f indicates the sine wave frequency (in cycles per field-of-view). Notice that projector intensity varies only in the y direction.

Given a set of N patterns, a set of N captures can be analyzed according to

$$\Phi(x^c, y^c) = \arctan\left[\frac{\sum_{n=1}^{N} I_n(x^c, y^c)\sin 2\pi n/N}{\sum_{n=1}^{N} I_n(x^c, y^c)\cos 2\pi n/N}\right] \tag{1.2}$$

where Φ denotes a phase value between 0 and 2π, $In(x^c, y^c)$ represents the intensity of capture n at location (x^c, y^c) in camera space, and is related to the projector y^p coordinate according to

$$y^p = \Phi(x^c, y^c)/(2\pi f) \tag{1.3}$$

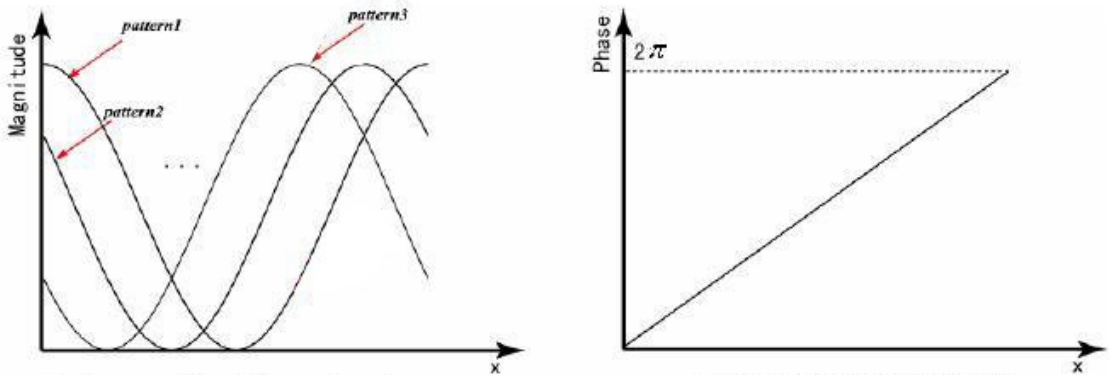

Figure 1.7, (left) Example PMP projection cross section and (right) resulting phase map cross section from image analysis [42]

The technique requires a minimum of three shifted patterns, and depth measurement accuracy increases with the number of shifts. More recently [43, 44] PMP has been extended to incorporate multiple sinusoidal frequencies. This multi-pattern PMP has the advantage of higher accuracy for a given number of frames and, equivalently, fewer necessary frames for any desired scan accuracy with respect to standard PMP.

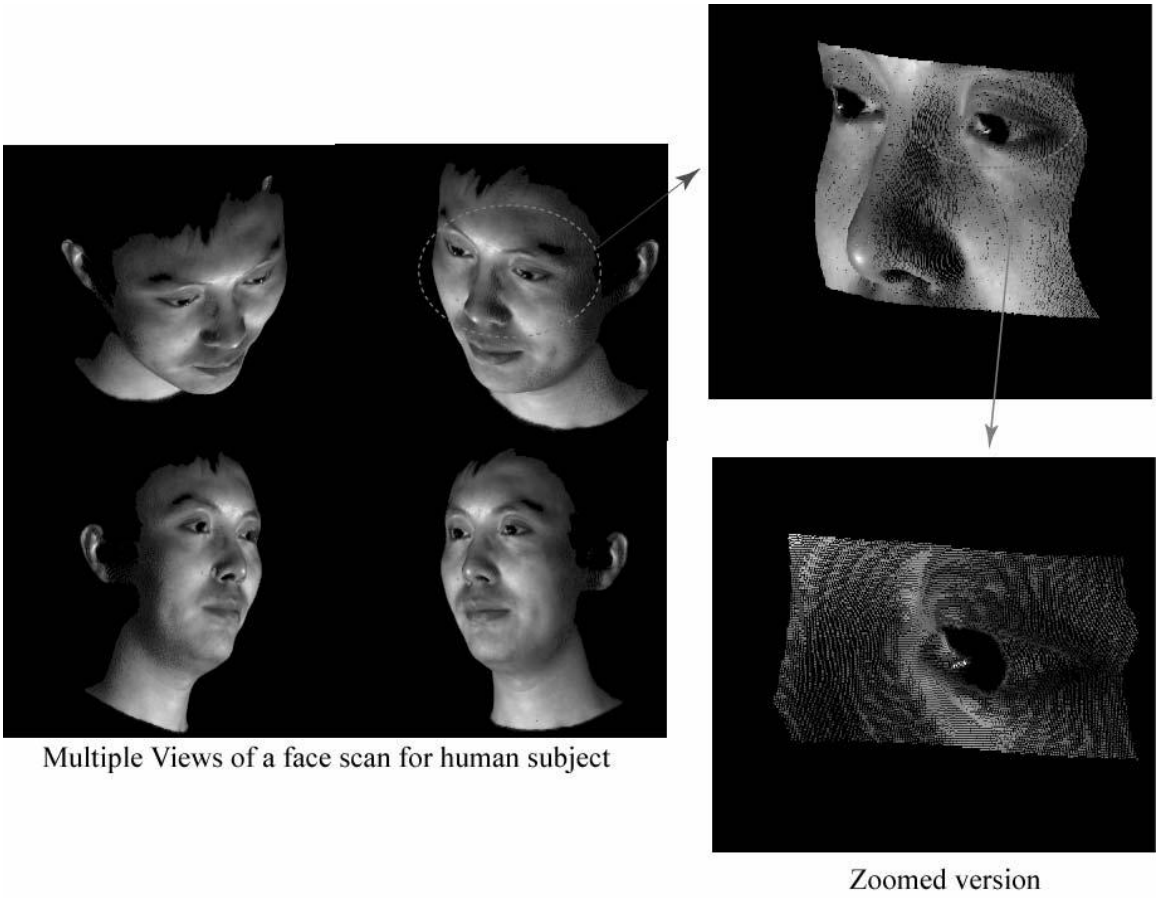

Figure 1.8, PMP 3D acquisition result example [32]

Hybrid Systems

Hybrid SLI systems make use of the methodology of a passive system (commonly stereo-vision) but with enhancements implemented with structured light [45-47]. As noted in the previous section, passive systems are often limited by the characteristics of the subject. Calculating disparity measurements necessary for stereo-vision reconstruction, for example, is quite difficult for smooth, featureless subjects (or areas of subjects). It is quite common to use a projected light pattern to overcome these inherent ambiguities. By projecting one or more patterns onto a subject, the systems introduce local variations which can be used to calculate disparity measures etc. as needed for depth calculation. These systems are sometimes multi-frame systems, and sometimes single frame systems, depending on the particular implementation of the technology.

Active Real-Time Systems

Active Real-Time SLI systems are not a specific subclass of multi-frame systems. Rather, they may be considered the next stage of advancement in multi-frame active depth acquisition systems. Using a variety of multi-frame pattern techniques [39, 48], these systems allow real-time acquisition of highly dense surface data (often similar to the results shown in Figure 1.8) through the utilization of specialized high-speed hardware and software. The concept is really quite simple: multi-frame techniques work quite well as long as the subject doesn't move during the scanning period. Active Real-Time systems work according to the same principles as standard SLI systems, but acquire data so fast (performing entire multi-frame scans in mere hundredths of a second) that a subject is unable to move enough during a scan to significantly impact the results.

The data that the systems acquire is usually of a very high visual and metric quality. Nevertheless, these systems are frequently limited by data throughput limitations. That is, it is often difficult or impossible to actually store the massive amount of 3D data that these systems generate in real time. Additionally, the hardware requirements often make these systems significantly more expensive than competing 3D imaging technologies.

1.3 Single-frame Techniques

While multi-frame techniques are often quite effective, it is nevertheless true that for many scanning applications it would be ideal to capture the required surface information in only a single projection/image instance. Fortunately, there are options for single frame SLI capture, which offer not only reduced scan time, but also allow one to scan a moving subject.

Un-coded Methods

One of the most significant problems that any structured light system must overcome is that of ambiguity. In order for 3D data to be properly acquired, there must be an absolutely certain correspondence between a pixel on the projection pattern and a point on the subject illuminated by that pixel. However, consider the situation shown in Figure 1.7. You can see in the left image an apparently continuous field of stripes. Nevertheless, as seen in the right image, this is an incorrect assessment of the true situation. In actuality, the stripes that appear to be connected on either side of the obstruction are most certainly not connected. It is therefore likely that the

correspondence between a projected stripe and a captured stripe would not represent the true situation, and therefore, calculations based on this correspondence would render incorrect depth measurements.

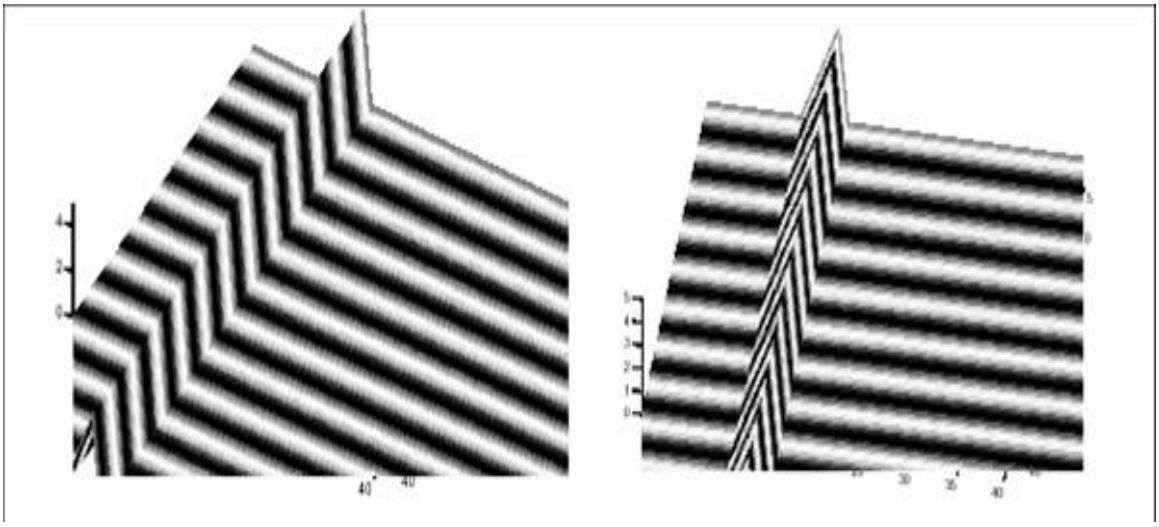

Figure 1.9, Ambiguity in projected pattern due to surface shape [49]

In order to avoid problems such as these, most structured light systems use some method by which ambiguity can be avoided. Often this is the motivation behind the development of new pattern methodologies, both single pattern as well as multi-pattern techniques. Nevertheless, there are methods that simply utilize these un-coded patterns regardless of ambiguity. Shapes such as stripe fields or grids are popular choices as un-coded pattern types [50-52].

Since there is little in the way of information embedded within the pattern itself, un-coded SLI is based entirely on creative analysis of the captured image. Often, statistical approximations, artificial intelligence, and algorithms that make use of prior knowledge of expected pattern distortions are used in an attempt to reconstruct the

surface. As such, the quality of results acquired using these systems can be unpredictable, though in general they tend to work best for subjects with minimal surface depth variation.

Neighborhood Search Methods

Neighborhood-search methods [53-58] take a somewhat different approach. These techniques utilize a pattern (usually binary in nature, that is, black and white colors only) in which subsections of the pattern can be uniquely identified in some way. Specific implementations may utilize patterns of noise or streaks [56] in which a point can be identified according to the known local statistical characteristics of the pattern, or deterministic sub-patterns defined by De Bruijn binary sequences [57] or M-Arrays (effectively a 2D pattern based on the same principle as De Bruijn sequences, that is, an M-Array is a matrix in which any sub-matrix of a given size is unique)[54] wherein the identity of one point can be determined by the information contained in nearby points. Like color-multiplexed systems, the accuracy of neighborhood-search based techniques may be strongly dependent on subject surface characteristics. If pieces of the pattern can be obscured by subject features or distorted too much by local gradients, correct identification of the pattern points may be impossible. In addition, the primarily binary nature of the patterns can limit the resolution possible. Thus, only a small portion of surface points may be measured.

Neighborhood search methods may also make use of color patterns in order to further increase the amount of information recovered or the reliability of the scan (that is,

ensuring that data is recovered accurately). Nevertheless, such problems still suffer the same weaknesses common to all color-based patterns.

Color Patterns

Color patterns are very often employed in structured light technology [59-69], either considered as unique standalone pattern configurations, or in the form of color-multiplexed patterns. Color-multiplexing [64-69] simply combines individual patterns from some multi-pattern technique into a single pattern by coloring each differently. A three pattern PMP sequence, for example, can easily be combined into a single pattern by coloring each of the three patterns red, green, or blue. In this way, each pattern can be isolated independently of the others by considering only the R, G, or B channel of the captured image. Each channel image is effectively identical to a single frame of the corresponding multi-pattern PMP scan process.

The use of colored De Bruijn sequence patterns [68, 70] is effectively another instance of color multiplexing. By using color elements, the effective length of unique De Bruijn subsequences is increased, and thereby, the density of recovered information is increased far beyond what would be possible in a single binary projection.

While the concept is simple, the number of patterns that can be combined in this way is usually relatively limited and analysis is plagued by non-idealities [65, 67]. Additionally, all color patterns introduce a strong dependence on subject coloration and luminance properties. If a subject is strongly colored blue, for example, there may be insufficient information in the R and G image channels to properly reconstruct the surface.

Yet another significant problem is the unfortunate reality of color-cross-talk. Due to imperfections in the color isolation components of both cameras and projection devices, it is often difficult or impossible to totally isolate an individual pattern. Significant research [67] has gone into solving this problem.

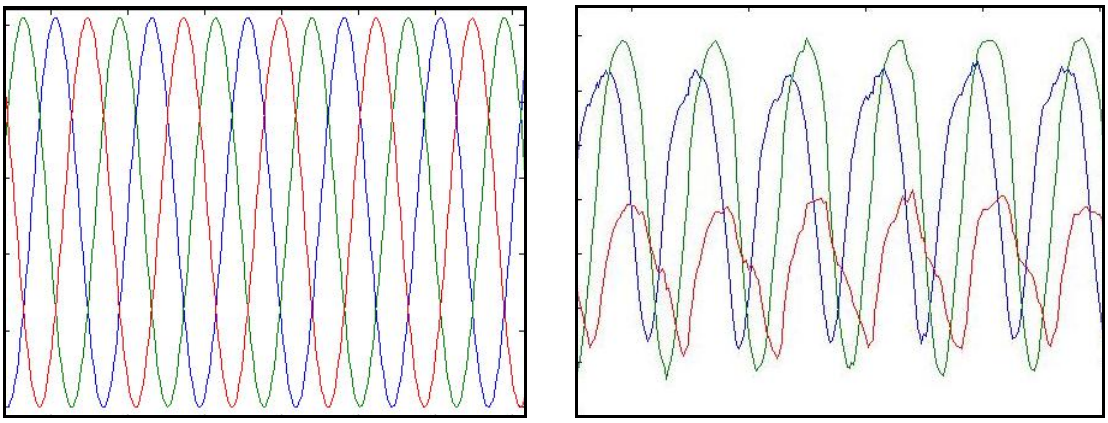

Figure 1.10, (right) shows distortion due to imperfections in projector/camera color filtration of (left) an ideal color multiplexed 3 channel projection pattern

Figure 1.10 shows just how significant hardware distortions of a color pattern can be. On the left is the conjunction of an ideal three sine wave pattern according to the color multiplexing methodology (one sine wave in each of the RGB color channels). To the right is the unaltered shape of the recovered sine waves which resulted from a single projection/capture event upon the subject of a flat, white sheet of paper. Notice that not only have the relative magnitudes of the projected sine waves been altered (a distortion known as color imbalance), but also that the shape of the sine waves themselves are greatly distorted due to cross talk between each color channel.

The effects of these distortions can be seen in Figure 1.11. A small cropped (nose) section of a color-multiplexed scan of a statue is shown. Note the strong, regular

wavelike distortions in the recovered surface. It can be clearly seen that the weaknesses inherent to color patterns are not to be ignored.

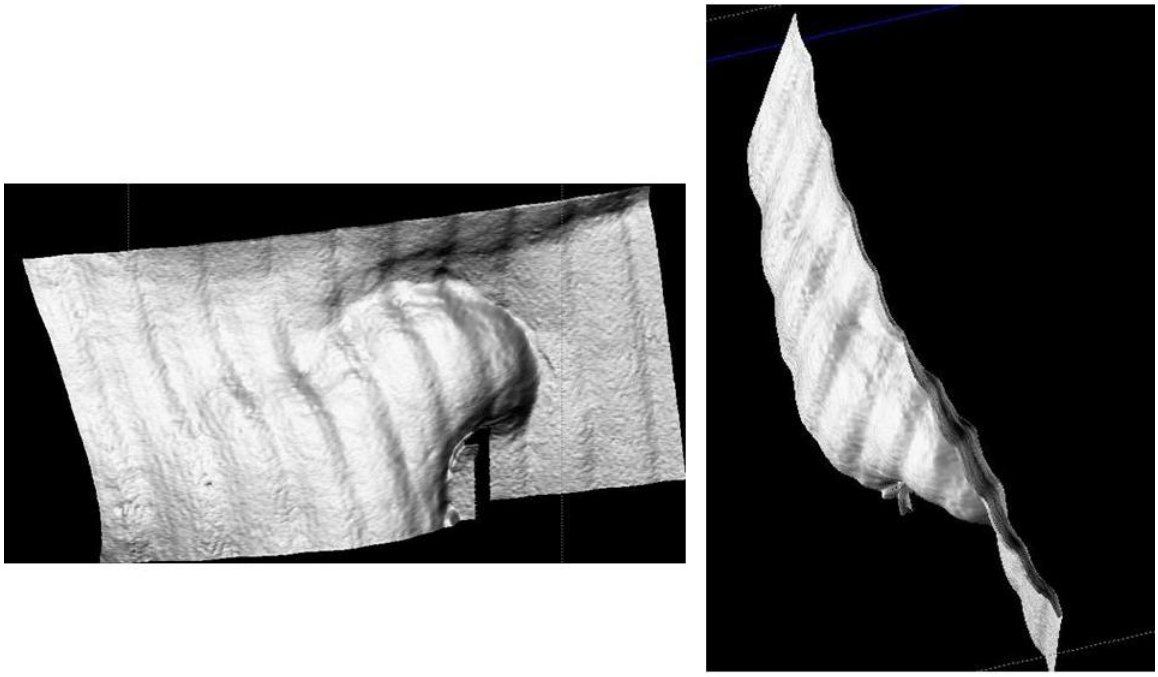

Figure 1.11, Example of errors introduced to 3 pattern SLI due to color interference

Composite Pattern

Composite Pattern (CP) [32, 42, 49, 71] is not a standalone pattern type, but is instead a way to combine multiple projection patterns into a single projection. Due to the method of triangulation that is used in many SLI techniques, only one dimension of the projected pattern is normally modified for the purposes of depth calculation. This is known as the "phase direction", a term derived from phase measuring profilometry. The perpendicular dimension is known as the "orthogonal direction." A CP projection pattern consists of multiple single projection patterns multiplied in the orthogonal direction by different (constant) frequency sine wave variation functions, essentially a form of spatial amplitude modulation. These individual modulated patterns are then added together to

form a single pattern. In a captured image of a subject illuminated by such a pattern, one can independently isolate the illumination effects of each modulated pattern by isolating the modulated signal envelope in the Fourier domain, thereby allowing one to acquire as much information in a single image as one would normally acquire in an entire SLI pattern sequence.

Figure 1.12, Composite pattern (left) projected onto a human subject and (right) analyzed to create a depth map [32]

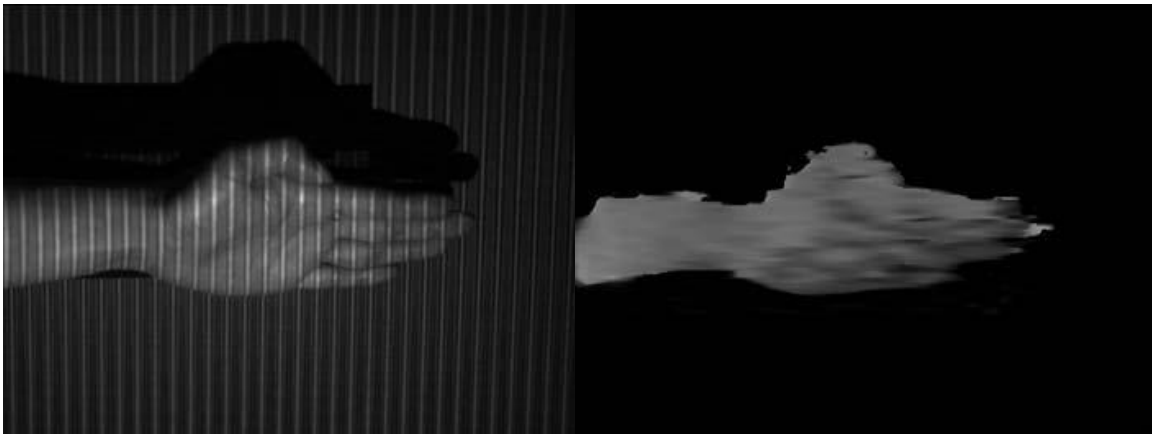

Figure 1.13, Composite pattern (left) projected onto a human hand and (right) analyzed to create a depth map [32]

Figures 1.12 and 1.13 show some early experimental applications of composite pattern multiplexing for a four pattern PMP sequence. Notice that in comparison to standard PMP results (Figure 1.7) and even a color-multiplexed result (Figure 1.8) this

early composite pattern was quite prone to surface measurement error. It can be seen that there is often significant depth error near surface edges especially due to CP component pattern isolation (which requires frequency domain filtering).

During early stages of development, the most common method of correction [32, 42] utilized a multiple-iteration dynamic programming approach. The initial CP projection pattern undergoes epipolar correction to further orthogonalize the phase and orthogonal depth distortion. Once a scan image is collected and depth measurements are made, these measurements are used to create a simulated capture image by warping and transforming an image of the projected pattern. Using correspondence matching algorithms developed for stereo-vision problems, the two images were compared and point-by-point disparity is calculated. One was able to correct the depth measurements by minimizing the disparity between the real and simulated images. Early simulations suggested the viability of the method. However, when put into experimental practice, even with post-processing the resulting reconstruction had considerable depth noise, as shown in Figure 1.14. It was noted that the algorithm was still potentially sensitive to albedo variations.

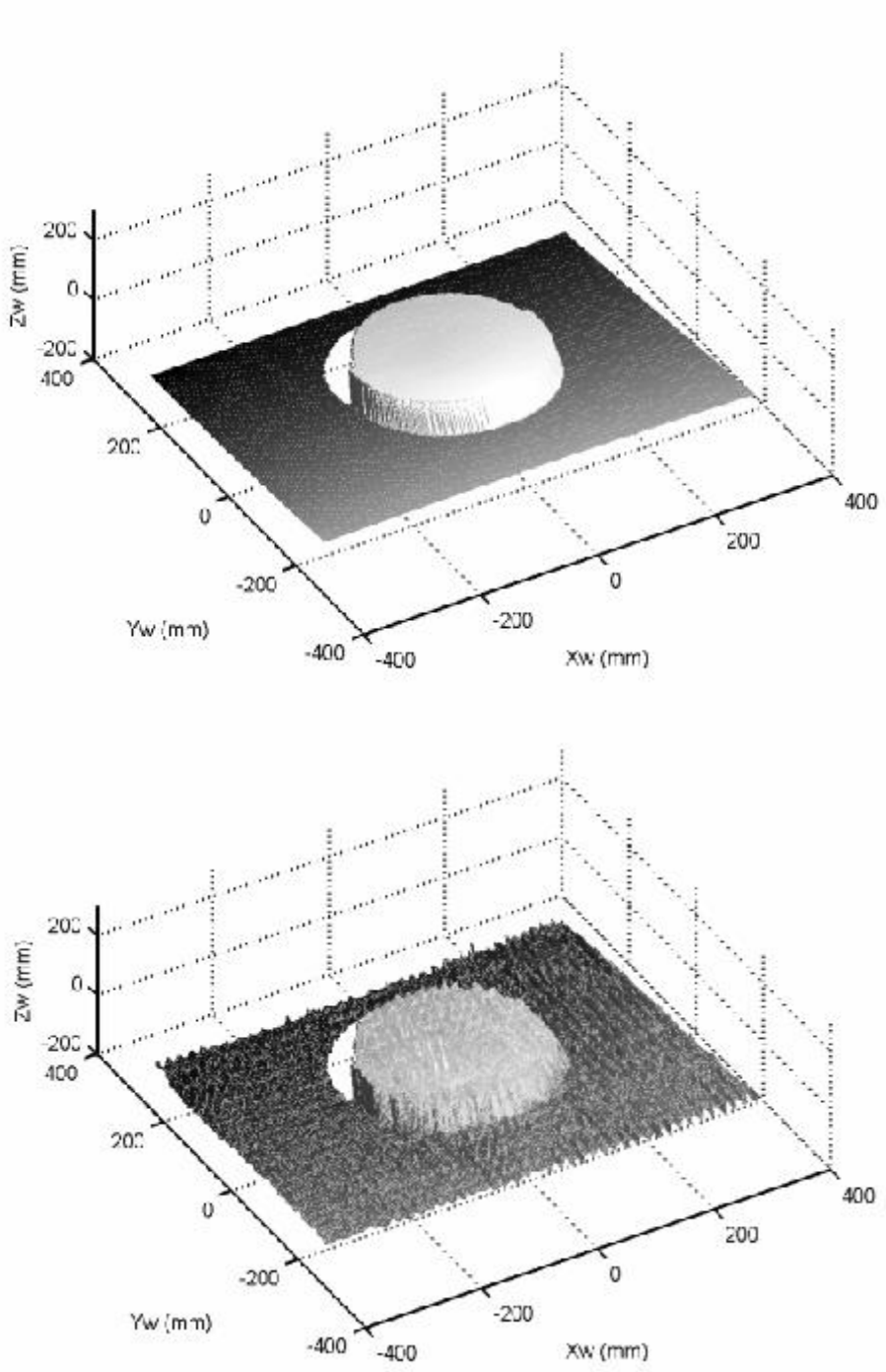

Figure 1.14, (top) Ideal reconstruction of a small, flat disc, (bottom) Composite Pattern scan result after a single post-processing correction iteration [32]

Lock and Hold Structured Light Illumination

Like the Composite Pattern technique, Lock and Hold [72] motion capture was an idea inspired by communications theory [73]. The idea is that, as in the operation of a phase-lock loop, if one can "lock on" to an unknown signal, then the changes in that signal can be easily tracked by compensating for the small incremental changes that occur through time.

Lock and Hold motion capture blurs the line between single frame and multi-frame SLI. It uses an un-coded structured light pattern (usually a pattern of stripes with triangular or sinusoidal cross sections) to capture the depth data of a moving surface. Changes in this "Hold pattern" are traced through the multiple frames of the capture video sequence in order to acquire a continually updated accurate depth map of the subject. Unlike similar systems that utilize un-coded SLI [50, 74] the system avoids difficulties involved with pattern ambiguity by the use of the "Lock sequence," a preliminary 3D scan taken before the subject is allowed to move. Since an un-coded pattern has numerous identical elements, it can't generally be used to measure absolute depth in the same way as a coded pattern method (such as PMP or even Composite Pattern) since a projected pattern point may correspond to any number of pattern points on the captured image (a problematic called simply "ambiguity" in common practice). By performing a preliminary 3D scan using PMP, the relationship between an identified point on the Hold pattern projection and Hold pattern capture can be unambiguously defined.

A simple explanation of the Lock and Hold process is as follows: to begin, a standard 3D scan of a subject is taken using a method such as PMP. Immediately following this, the Hold pattern projection begins and the subject is allowed to move. The

Lock scan creates an unambiguous “phase map” which relates each point of the projection pattern to a single point that it illuminates on the subject image. If a Hold pattern is immediately projected, the first frame of the Hold capture sequence is directly related to the phase map. In other words, each isolated feature of the Hold pattern maps to a single phase value from the PMP scan. In this way, the depth of each isolated Hold pattern feature (i.e., “snake” shown in Figure 1.15) can be calculated using triangulation techniques.

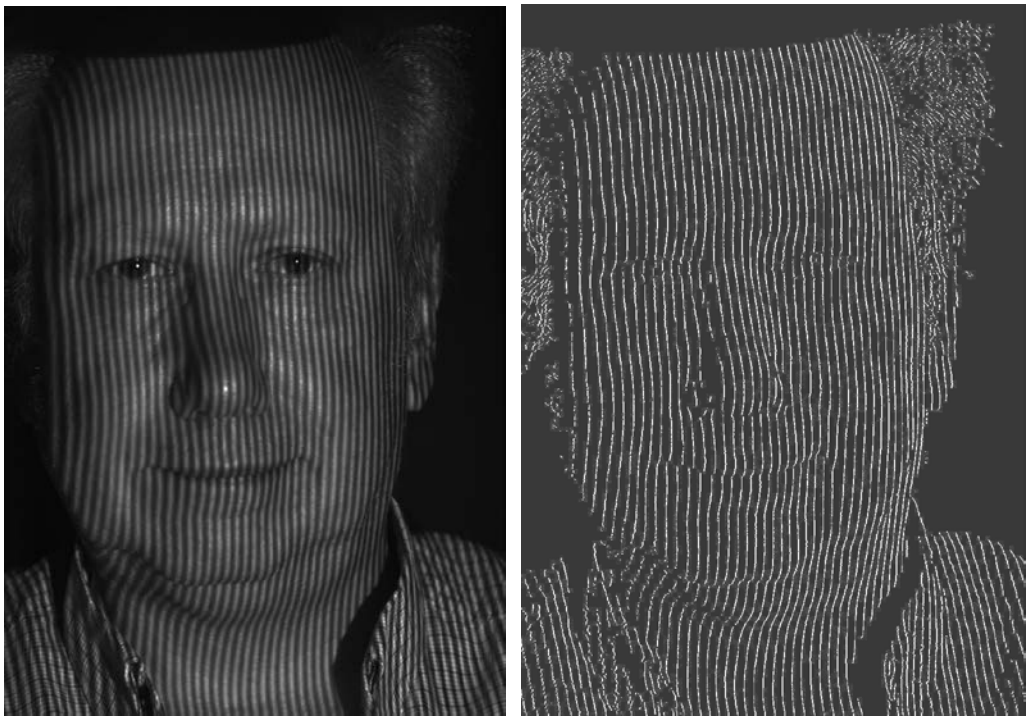

Figure 1.15, (left) Example of a Hold pattern projection and (right) resulting "snakes"

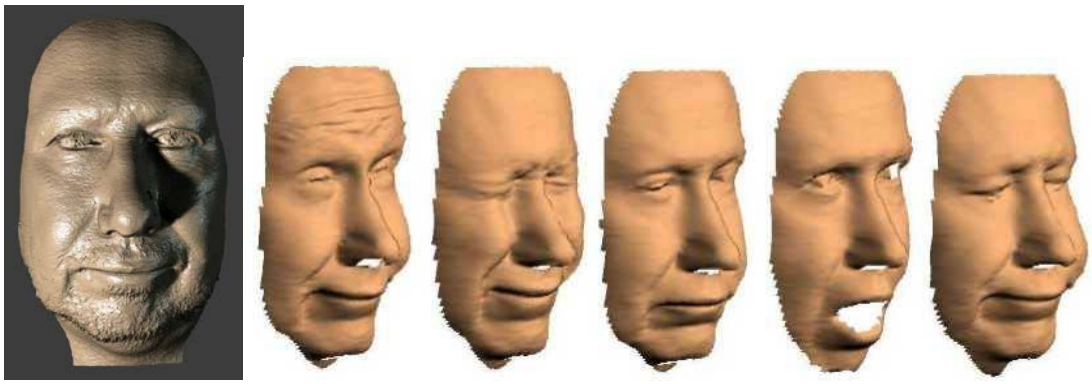

Figure 1.16, (left) Lock scan and (right) sequence of reconstruced frames from Hold images

Once each feature in the first Hold frame is unambiguously identified, features in the next frame are isolated. Then, at each identified point in the first frame, a search is performed in a window around the corresponding position in the next frame. If a suitable feature is found in that frame, it is assigned the appropriate identity. In this way features can be traced through the numerous frames of the Hold sequence, and a depth map for each frame can be calculated. A Lock scan and five samples of Hold scans are shown in Figure 1.16. Practical implementations of the process require additional steps, of course. Depending on the shape of the subject and the speed of its movements (relative to the capture rate of the camera), the initial tracking process may identify some features incorrectly. Thus, techniques for error prevention or correction are normally required for optimal results. However, for our purposes, a detailed description of this process is not necessary.

1.4 Motivations

Let us consider the ideal structured light 3D data acquisition system. Such a system should be first of all accurate, able to create a data model of a subject that is actually representative of the physical object within a small acceptable margin of measurement error. If a system cannot image the 3D surface accurately, then the data acquired is often simply unusable for any practical purpose. It should also be robust, able to capture a good scan with minimal regard to subject or environment characteristics, and thereby lending itself for use in many different situations and many different subjects. Multi-frame techniques, such as PMP, are often quite close to ideal in these respects.

A perfect system would also be fast, able to collect 3D data instantaneously. This would allow it to capture subjects which are unpredictable or mobile (biological subjects for example). It would also be more convenient for the user, who would often not wish to wait for their data to be collected, and would prefer to avoid the need to take multiple attempts at data acquisition due to scan failures thanks to a moving subject. Single-frame and real-time active systems excel in this regard.

Finally, the system would be simple and user friendly. It should be inexpensive, easy to use and hard to use incorrectly. In these areas, single-frame systems have a distinct advantage, as they can often be used with static projection systems (unlike most other structured light systems, which require active electronic illumination devices such as digital projectors). The system should require minimal knowledge and skill on the part of the user.

Composite Pattern was developed as an attempt to create, or at least approach, this ideal system. It was intended to be a method with the accuracy and robustness of a multi-frame system, but with the speed and versatility of a single frame system; a new

form of pattern multiplexing which could avoid the problematic drawbacks inherent to either a color-multiplexed single pattern method (which, once again, can be clearly seen in Figures 1.10 and 1.11) or any passive methods, such as stereo-vision (passive methods, as noted in previous sections, often have difficulty resolving the correspondence problem for certain subjects, such as those with limited surface features).

The development was successful, to a certain extent. Using CP, one could indeed capture 3D depth information in only a single frame, while avoiding the majority of the drawbacks present in a color or stereo-vision system. However, as seen in Figures 1.12 - 1.14, CP came with its own set of flaws. This was to be expected, perhaps, as it was an unprecedented new system, but still unacceptable for many applications. While first generation CP could be useful in areas such as machine vision or user interface control (as demonstrated specifically in reference [71]) it was considered too inaccurate for most matters involving human subjective evaluation (it was not an acceptable imaging solution for 3D graphics, for example) and likely for most forms of record-keeping or measurement as well.

Thus a second generation of development was undertaken, leading to the creation of the Modified Composite Pattern (MCP) method. By introducing significant changes to the CP pattern design as well as additional processing steps, MCP was able to greatly improve the performance of Composite Pattern, creating results that were both more accurate as well as visually pleasing (as seen in Figure 1.17) than anything possible with first generation techniques.

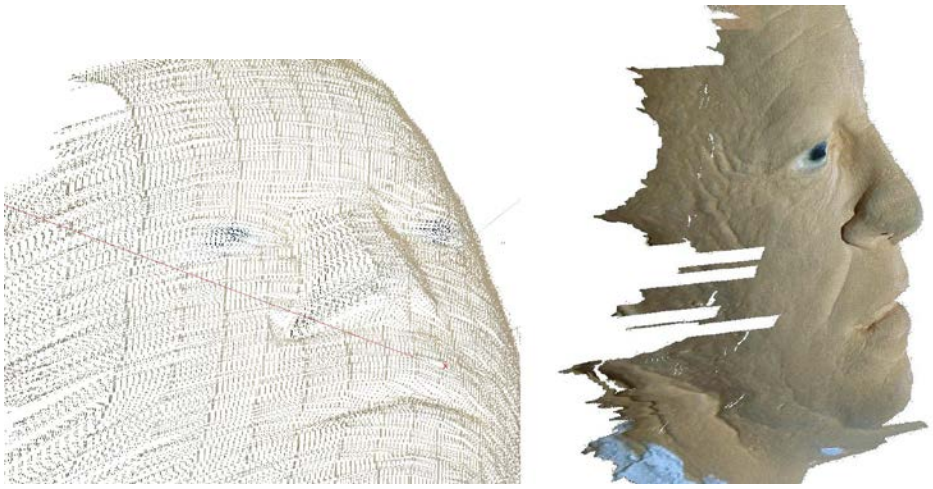

Figure 1.17, MCP scan of a latex bust, (left) MCP data points with interpolated linear fill regions, (right) result after smoothing and other post-processing steps

Nevertheless, problems still remained. The system relied heavily on user input parameters. An incorrect choice for any one of the many necessary processing parameters would lead to error, which was often very significant, in the output result. Additionally, the system, while effective, often showed a lack of robustness. Result quality was often highly dependent on individual subject characteristics. Being, as it was, fairly new and developed as an experimental offshoot of an already existing technique, the method was highly algorithmic in nature and lacked a consistent theoretical basis. It was the goal of this research to overcome these systematic weaknesses.

The research in this dissertation constitutes a third generation of Composite Pattern development, with the intention of further advancing the CP paradigm toward the final goal of creating an ideal structured light 3D acquisition technology.

1.5 Contributions

Second generation Modified Composite Pattern is frequently implemented as a binary SLI pattern, multiplexed according to Composite Pattern methodology. This single pattern is then further altered with the addition of a stripe pattern (usually sinusoidal or triangular in cross-section, thereby defining a unique peak location), allowing improvement in the quality of contour results. However, the process by which these stripes were traditionally identified was found to be problematic.

The stripe identification algorithm (effectively a form of local peak detection) was based on the original technique used in Lock and Hold structured light processing. It calculated, at each pixel in the input image, a peak-to-sidelobe ratio, defined as the image intensity value at that pixel divided by the maximum intensity value at some specific lateral distance away (ideally defining the ratio between the highest and lowest points on the stripe cross section). This process required the user to define a specific lateral distance measure, in pixels, at which to calculate this ratio. It was observed that the most appropriate value for this parameter would often vary among images, and even among different stripes in the same image, leading to inconsistencies that would often degrade the quality of a recovered surface scan.

Additionally, the user was again forced to select an appropriate threshold measurement for this peak-to-sidelobe ratio measurement in order to define stripe regions (which could then be locally searched for the actual peak present). Again, this ideal measurement was often variable, even within a single scan image. Even small errors in these selected parameters could render a scan highly prone to error, potentially rendering the scan data unusable.

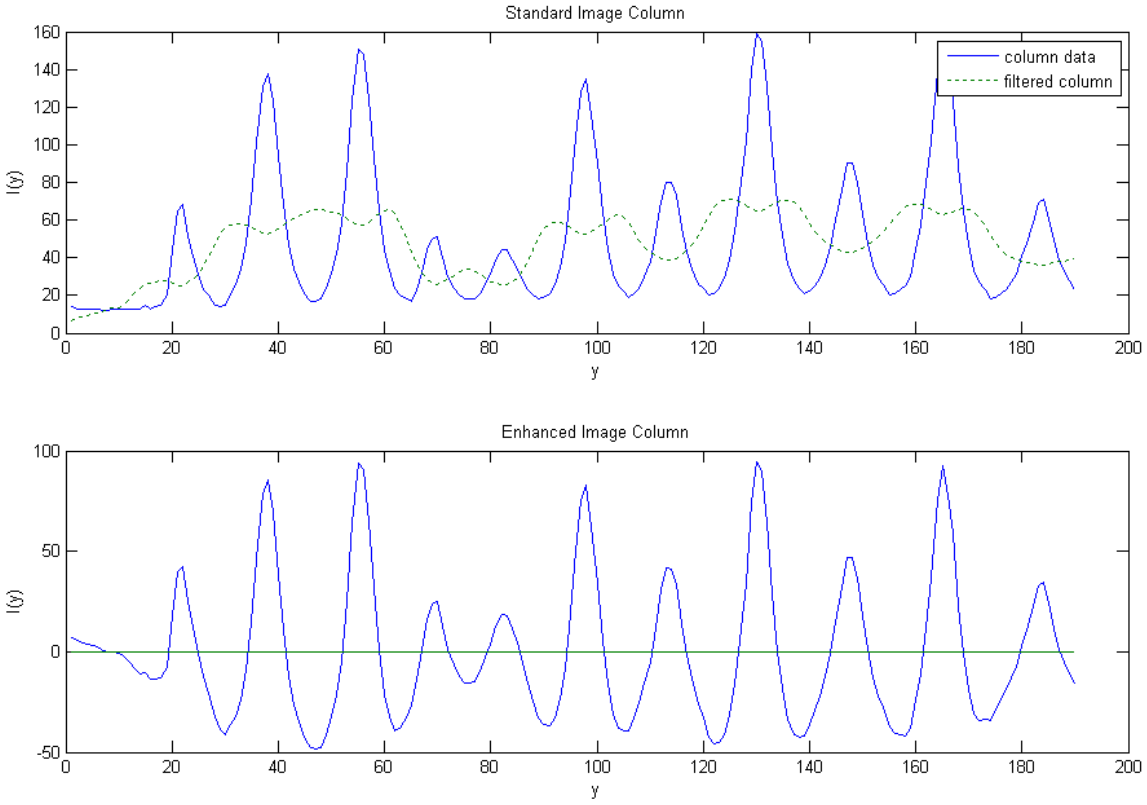

Figure 1.18, Adaptive threshold generation

The solution implemented in this research makes use of a linear-filter based locally adaptive threshold method (illustrated in Figure 1.18) which effectively eliminates the need for these earlier user-defined parameters. This specifically removes a major source of procedural error, thereby increasing the system's overall robustness. The method is also significantly less impacted by local intensity variations, and is able to define usable stripe regions even in noisy or low-intensity regions, which earlier systems were often unable to do.

Early generation CP systems also required specification of detection frequencies by the user. That is, the user was to observe the frequency spectrum of the CP input image and effectively pick out which bands constituted modulated information. However, in many cases this would prove difficult or effectively impossible due to object

features (as seen in Figure 1.19). Incorrectly identifying a modulation frequency, even by as little as a few cycles per field of view, could introduce significant measurement error as well. These strict requirements, including the necessity for a highly expert user, created severe limitations on the system, effectively rendering it unusable for practical or commercial scenarios.

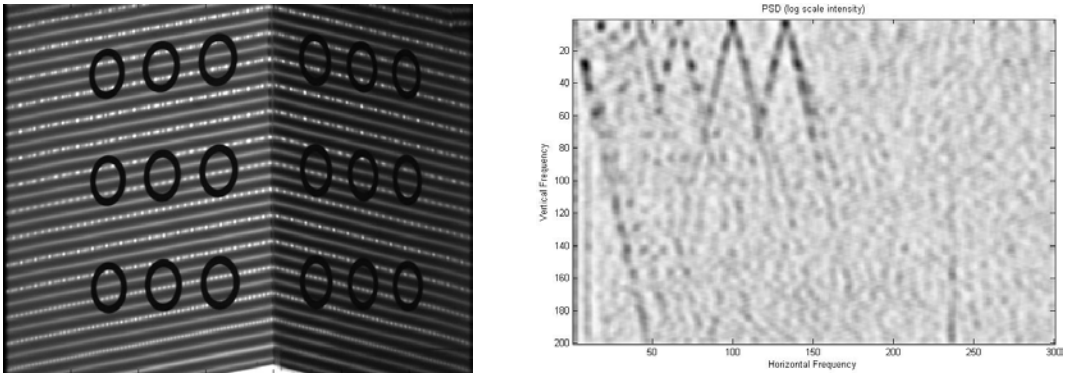

Figure 1.19, Example of stripe rotation effects

The research performed during the development of this dissertation effectively eliminated these problems with the inclusion of a method by which problematic subjects may be segmented and corrected (as displayed in Figure 1.20).

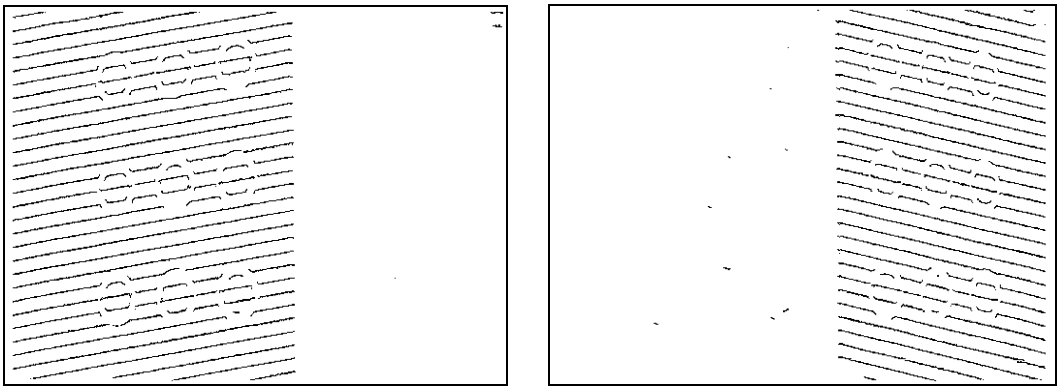

Figure 1.20, (left) Isolated positive inclination stripe regions, (right) Isolated negative inclination stripe regions

Detection of modulated frequency bands was also automated using an analysis-by-synthesis envelope detection algorithm (illustrated in Figure 1.21), generating an ideal definition of each modulation frequency, and allowing for consistent and effective CP component isolation and analysis.

With these advances, the need for user interaction in the CP demodulation process can be effectively eliminated, rendering the system much more robust and easy to use, while simultaneously facilitating ease of further advancement.

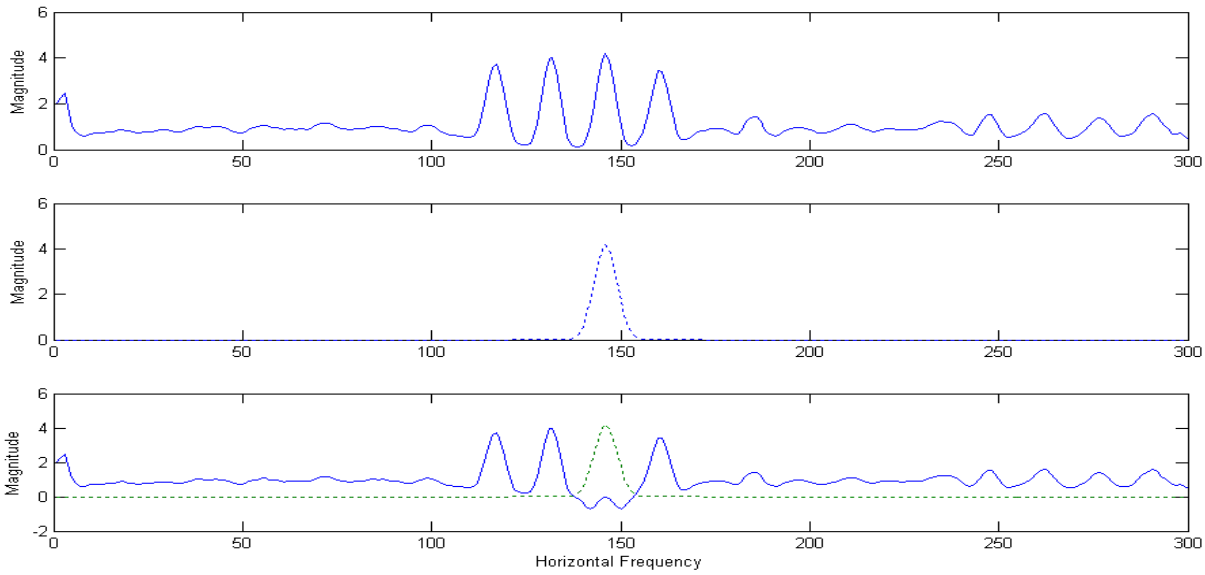

Figure 1.21, Analysis By Synthesis process, (top) initial array A, (middle) array B, where B represents isolated frequency band components, after one iteration, (bottom) updated array C where C=A-B, where the presence of B is indicated by the dotted curve

Early generation Composite Pattern research treated the method as a typical structured light system. Corrective measures such as epipolar correction were implemented, but were only partially effective, and largely unable to overcome the weaknesses inherent to the system. However, the advent of MCP, with its unique characteristics, moved the CP paradigm into new and unknown territory. As such,

second generation systems lacked a meaningful and consistent theoretical basis for development.

Normal SLI systems are usually concerned with overcoming ambiguity. However, MCP introduces a SLI methodology which effectively eliminates the problem of ambiguity entirely. Each stripe in an MCP scan is uniquely identifiable, and all identities will occur in a known and predictable order upon the subject surface. During the course of this research, it was determined that a Hidden Markov Model would most appropriately characterize the situation encountered in an MCP scan. It was determined that optimal analysis via Viterbi sequential decoding would be an appropriate and effective measure for MCP. This characterization is heretofore unknown in the context of SLI research. It has proven effective in the course of this particular research and, by providing a new and reasonable context for theoretical development, promotes future progress in Composite Pattern methodology.

In summation, with the research performed during this third generation of Modified Composite Pattern development, the technique has effectively been established as a viable competitor with all available 3D imaging techniques and has been made suitable for future use and refinement toward the eventual goal that CP was created to achieve.

1.6 Outline

The body of this dissertation is organized as follows: the initial chapter introduces the most general information regarding the research performed. Chapter 2 discusses, in great depth, the MCP technique. The method is first explained, and then characterized

via specific mathematical representations. Consideration is also made for various sources of systematic error which one is likely to encounter. Once the MCP pattern and processing methodology has been adequately explained, chapter 3 goes into detail regarding the actual 3D reconstruction that will be performed. Also included is information regarding the specific calibration technique utilized for the current generation MCP technology. Chapter 4 examines the practical stages in MCP processing. The chapter also includes specific information regarding automated analysis techniques implemented in the course of the research. Chapter 5 finishes the discussion of the MCP technique by introducing advanced post-processing and optimization techniques utilized to further improve depth recovery, along with a collection of scan processing results and tests for depth recovery accuracy. Finally, chapter 6 concludes the dissertation and discusses potential avenues for further research and summary improvement of the MCP methodology.

Chapter 2 Modified Composite Pattern

Modified Composite Pattern is frequently implemented as a binary SLI pattern, multiplexed according to Composite Pattern methodology. This single pattern is then further altered with the addition of a stripe pattern, allowing improvement in the quality of contour results. The inclusion of this additional pattern also adds complexity in processing, versus traditional CP.

2.1 Introduction to MCP

A binary code structured light pattern allows 3D reconstruction of a subject by encoding each point in an image of the subject with a unique value. This value corresponds to a (ideally unique) point in a projector, and as such allows 3D reconstruction by triangulation. Each point is encoded by projecting a sequence of binary images of various forms, such that for any given projection, a point may be either illuminated (thus being encoded with a binary 1 for that frame) or not illuminated (in which it is encoded with a 0 for that frame) depending on the shape of each particular pattern in the sequence. The values collected during the sequence create multi-bit representation of a number in base 2, identified as the point's binary code-word or code value. The binary code form of Composite Pattern (CP) SLI allows a single frame equivalent of this process. A visualization of a binary pattern is shown in Figure 2.1. Notice that the combination of binary information from each of the 4 unique pattern frames creates 16 uniquely valued regions (indicated as unique color combinations).

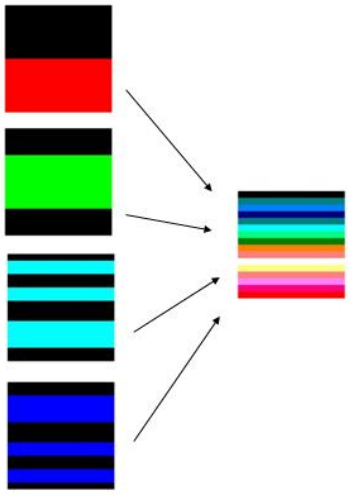

Figure 2.1, Visualization of binary pattern encoding

The pattern modification that we use consists of a field of stripes which are most frequently sinusoidal or triangular, without loss of generality. A modification pattern could hypothetically consist of any recognizable form, such as circles, dots, or a grid pattern, and the selection of stripes was simply an intuitive choice. In a binary code Modified Composite Pattern (MCP) a single modulated stripe will be associated with a specific binary code value. Additionally, the particular implementation utilized in this research places an unmodified stripe between any two binary code stripes. Using Composite Pattern analysis techniques, each modulated stripe can be identified; but in addition to this, each unmodified stripe can be used for reconstruction as well, since any unmodified stripe positioned between two sequential binary-encoded stripes is uniquely identifiable. The presence of unmodified stripes effectively doubles the amount of useable surface contour information. An example of an MCP pattern is shown in Figure 2.2.

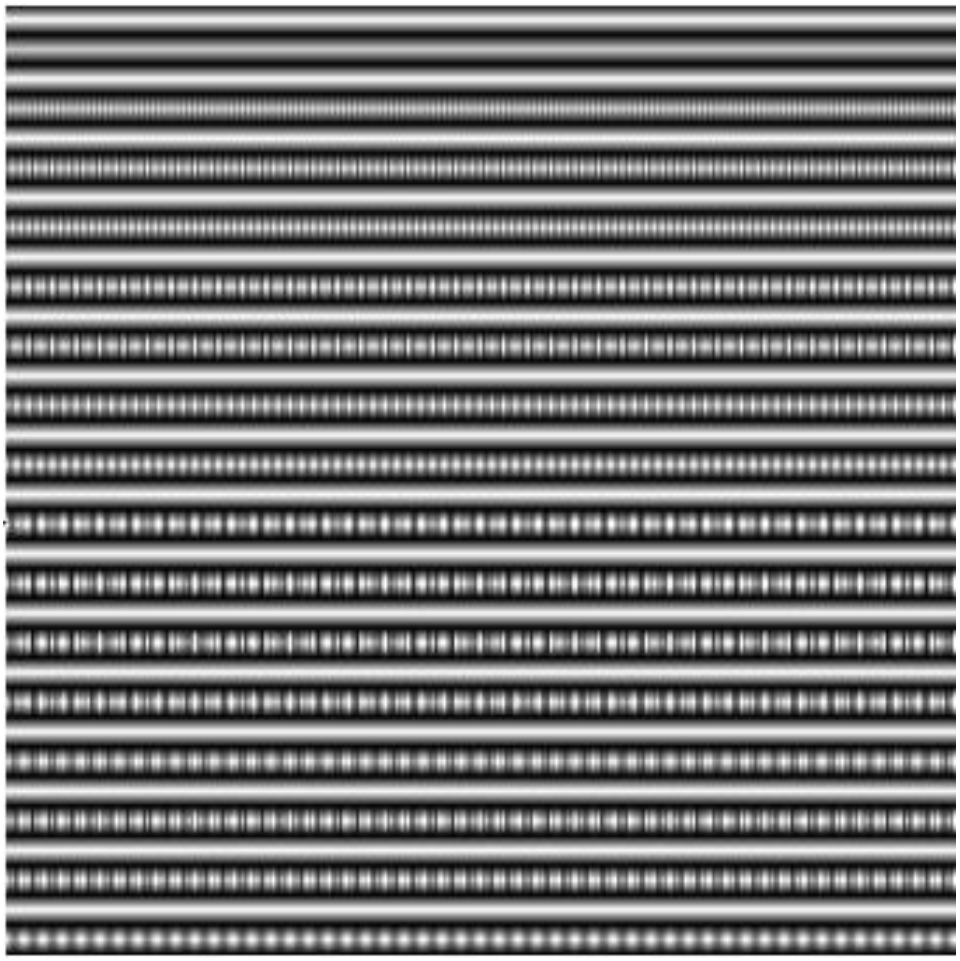

Figure 2.2, An example of a Modified Composite Pattern

In addition to these stripes, each minimum falling between two stripes (known as “negative stripes” or “2x stripes” due to the fact that they further double the amount of available information) can also be identified based on the unique combination of binary-encoded stripe and unmodified stripe that it falls between. It is also possible to track the intensity midpoints of the stripes (between each peak and trough) for a further increase in the number of identifiable stripes per image (and these zero-crossing stripes are known as “4x stripes”). While a standard four image Binary code scan may be limited to 15 unique values (the ‘0000’ code word is unusable due to the presence of unmodified stripes), equivalent to a mere 15 points of resolution in the direction of code variation, a positive,

negative, and unmodified stripe pattern with 15 unique binary code values can be evaluated for 30, 60, or 120 points of resolution.

Consider the binary structured light pattern $p_i(x,y)$, where i=0…N-1 for a binary structured light sequence of length N. Suppose that each individual pattern p_i is element-wise multiplied by an image $q_i(x,y)$ where q_i is constant in the y dimension and varies only in the x dimension according to

$$q_i(x) = \frac{1}{2}\left(\cos(2\pi f_i x) + 1\right) \tag{3.1}$$

resulting in

$$p_i'(x,y) = q_i(x,y)\, p_i(x,y) \tag{3.2}$$

A composite pattern can be represented by

$$p(x,y) = \sum_{i=0}^{N-1} M q_i(x,y)\, p_i(x,y) \tag{3.3}$$

where M is simply a normalization constant.

Modified Composite Pattern further alters this pattern with the addition of a spatial stripe pattern, $g(x,y)$, and an additional unmodified stripe pattern $f(x,y)$, according to

$$I(x,y) = g(x,y)\, p(x,y) + f(x,y) \tag{3.4}$$

where both $g(x,y)$ and $f(x,y)$ are constant in the x dimension, varying in the y dimension according to

$$g(y) = M_a\left[\sum_{i=0}^{2^N-1} \delta(2ih - \tfrac{h}{2}) * \Lambda(\tfrac{2y}{h})\right] \tag{3.5}$$

and

$$f(y) = M_a\left[\sum_{i=0}^{2^N-1} \delta((2i+1)h - \tfrac{h}{2}) * \Lambda(\tfrac{2y}{h})\right] \tag{3.6}$$

where $h = \frac{Y}{2^{N+1}}$, and Y is the y dimension maximum of the pattern space, and M_a is an intensity multiplier. Here Λ represents the triangle function, and defines the shape of each spatial stripe. Note that the stripe cross section can be altered by replacing instances of the triangle function with whichever function one desires, such as saw tooth or sinusoidal functions.

2.2 Modulation Theory

Consider the binary structured light pattern $p_i(x,y)$, where i=0…N-1 for a binary structured light sequence of length N. Each such pattern contains inherent spectral information $P_i(f_x,f_y)$, where P_i is the Fourier domain representation of p_i. Upon projection onto a target, each pattern undergoes alteration due to the characteristics of the surface. The resulting captured image can be approximated as

$$c_i(x,y) = A(x,y) + B(x,y)d_i(x,y) \tag{3.7}$$

where

$$d_i(x,y) = S\big(p_i(x,y)\big) \tag{3.8}$$

Here, $A(x,y)$ represents information included due to ambient light, $B(x,y)$ represents a modification field due to target surface reflectance characteristics, and $S(\,)$ is an operator representing the element-wise transformation of each pattern pixel due to surface depth and world-camera coordinate transformation. Allow $S(\,)$ to be defined such that it has the property

$$\Im\{S(\,)\} = S_f \tag{3.9}$$

Where S_f is the equivalent spectral space operator for $S(\,)$.

It is the recovery of the quantity $d_i(x,y)$ for each pattern/image i in the sequence that allows one to recover the desired surface depth information according to structured light illumination methodology.

Composite Pattern Modulation

Let us now consider how the MCP capture process actually functions in a mathematical sense. Suppose that each individual binary pattern p_i is element-wise multiplied by an image $q_i(x,y)$ where q_i is constant in the y dimension and varies only in the x dimension according to

$$q_i(x) = \frac{1}{2}\left(\cos(2\pi f_i x) + 1\right) \tag{3.10}$$

resulting in

$$p_i'(x,y) = q_i(x,y)\, p_i(x,y) \tag{3.11}$$

The effect of this, in the spectral domain, is an effective duplication and translation in the f_x dimension of the baseband information $P_i(f_x,f_y)$, such that

$$P_i'(f_x, f_y) = \frac{P_i(f_x, f_y)}{2} + \frac{P_i(f_x - f_i, f_y)}{4} + \frac{P_i(f_x + f_i, f_y)}{4} \tag{3.12}$$

In other words, the pattern has been amplitude-modulated.

Consider now the (composite) pattern $p(x,y)$ where

$$p(x,y) = \sum_{i=0}^{N-1} Mq_i(x,y)\, p_i(x,y) \tag{3.13}$$

where M is simply a normalization constant. In this case, the spectral domain representation can be shown to be

$$P(f_x, f_y) = \sum_{i=0}^{N-1} M\left(\frac{P_i(f_x, f_y)}{2} + \frac{P_i(f_x - f_i, f_y)}{4} + \frac{P_i(f_x + f_i, f_y)}{4} \right) \tag{3.14}$$

Composite Pattern Capture

Now consider the effect of projecting this pattern upon a target surface. As in eqn. 3.7, the recovered information can be seen to be

$$c(x,y) = A(x,y) + B(x,y)d(x,y) \tag{3.15}$$

where

$$d(x,y) = S\big(p(x,y)\big) \tag{3.16}$$

and the Fourier domain representation of d is

$$D(f_x, f_y) = S_f \sum_{i=0}^{N-1} M\left(\frac{P_i(f_x, f_y)}{2} + \frac{P_i(f_x - f_i, f_y)}{4} + \frac{P_i(f_x + f_i, f_y)}{4} \right) \tag{3.17}$$

where S_f is the equivalent Fourier space representation of the effects of S.

It can therefore be seen that

$$\begin{aligned} & C(f_x, f_y) \\ & = \Im(c(x,y)) \\ & = \Im(A(x,y)) + \Im(B(x,y)) * S_f \sum_{i=0}^{N-1} M\left(\frac{P_i(f_x, f_y)}{2} + \frac{P_i(f_x - f_i, f_y)}{4} + \frac{P_i(f_x + f_i, f_y)}{4} \right) \end{aligned} \tag{3.18}$$

Note the effects of B in the equation above. For now, let us assume that B consists of a constant value, such that

$$\Im(B(x,y)) = \delta(f_x, f_y) \tag{3.19}$$

Let us also assume that the intensity of ambient light is so low as to make the effects of $A(x,y)$ negligible. In this case, it can be seen that

$$c(x,y) = d(x,y) \tag{3.20}$$

And thus, eqn. 3.18 reduces to

$$C(f_x, f_y) = \Im(c(x,y)) = S_f \sum_{i=0}^{N-1} M\left(\frac{P_i(f_x, f_y)}{2} + \frac{P_i(f_x - f_i, f_y)}{4} + \frac{P_i(f_x + f_i, f_y)}{4}\right) \tag{3.21}$$

For clarity, let us rearrange 3.21 such that

$$C(f_x, f_y) = \sum_{i=0}^{N-1} S_f \frac{M}{2} P_i(f_x, f_y) + \sum_{i=0}^{N-1} S_f \frac{M}{2}\left(\frac{P_i(f_x - f_i, f_y)}{2} + \frac{P_i(f_x + f_i, f_y)}{2}\right) \tag{3.22}$$

Composite Pattern Demodulation

Now, let $h_i(x,y)$ represent the impulse response of an ideal bandpass filter applied to the collected image, such that

$$\Im(c(x,y) * h_i(x,y)) = S_f \frac{M}{2}\left(\frac{P_i(f_x - f_i, f_y)}{2} + \frac{P_i(f_x + f_i, f_y)}{2}\right) \tag{3.23}$$

In other words, the filter has allowed isolation of a single modulated band component. Consider now the effects of multiplying the filtered image by a sinusoidal field (an image varying in the x dimension and constant in the y dimension), such that

$$
\begin{aligned}
&C_i'(f_x, f_y) \\
&= \Im\left(\left[c(x,y) * h_i(x,y)\right]\cos(2\pi f_i x)\right) \\
&= \left[S_f \frac{M}{2}\left(\frac{P_i(f_x - f_i, f_y)}{2} + \frac{P_i(f_x + f_i, f_y)}{2}\right)\right] * \left[\frac{\delta(f_x - f_i, f_y)}{2} + \frac{\delta(f_x + f_i, f_y)}{2}\right] \\
&= S_f \frac{M}{2}\left(\frac{P_i(f_x, f_y)}{2}\right) + S_f \frac{M}{2}\left(\frac{P_i(f_x - 2f_i, f_y)}{4} + \frac{P(f_x + 2f_i, f_y)}{4}\right)
\end{aligned}
\tag{3.24}
$$

This multiplication is effectively demodulating the isolated pattern information. Applying a simple low-pass filter allows one to isolate the key result

$$
C_i''(f_x, f_y) = \frac{M}{4} S_f P_i(f_x, f_y) \tag{3.25}
$$

It can be seen that

$$
\begin{aligned}
&c_i''(x,y) = \Im^{-1}(C_i''(f_x, f_y)) \\
&= \frac{M}{4} S\left(p_i(x,y)\right) \\
&= \frac{M}{4} d_i(x,y)
\end{aligned}
\tag{3.26}
$$

And thus, it is shown that by performing the prescribed series of operations, one may indeed recover the individual components of depth-recovery information $d_i(x,y)$ from a captured image $c(x,y)$ of a target subject illuminated by a composite pattern $p(x,y)$.

2.3 Systematic Interference Analysis

One inevitable difficulty encountered in MCP processing is the necessity of dealing with undesirable interference at the modulating frequencies. The primary causes of this are deviations from theoretical ideality, projection image gamma distortion effects, and also the introduction of spurious frequency modulation due to image intensity saturation.

Deviation From Theory

Given the theoretical representation of composite pattern processing from section 2.2, what conclusions may be drawn regarding a real composite pattern scan?

The first and most obvious fact is that the operator $S(\ )$ will, in practice, not be isolated separately from the demodulated pattern information. That is, the goal of CP processing is to isolate $d_i(x,y)$, and from this information, recover the depth of the surface. It will not necessarily be desirable to isolate $p_i(x,y)$ directly, and as such, the effects of $S(\ )$ will be considered coupled to the pattern p_i information throughout the process. The greatest practical effect that this will have is that of altering the apparent modulation frequencies. In other words, the frequencies f_i used to generate the pattern are not necessarily the same as the frequencies needed to demodulate the pattern.

Additionally, it was assumed that ambient light effects A and surface reflectance characteristics B had negligible influence on the captured image. In a realistic scenario, this can't always be ensured. From eqn. 3.18, it can be seen that A will add a certain amount of extraneous information into each isolated information band. Additionally, the theoretical consideration of B was such that it has influence only at $f_x=0$ and $f_y=0$ (that is, "DC"). In reality, B is likely to have significant energy at DC (thus allowing recovery of pattern information according to the theoretical model), but is also likely to have energy at all other possible frequencies as well. Since the spatial effect of B is multiplicative, frequency information contained therein results in further modulation of the pattern information. Due to the use of bandpass filters in the demodulation process, most of

these effects will be eliminated. However, there will still be two effects of B which may be potentially significant.

The modulation from B may create undesirable cross talk between patterns. Consider the energy in the surface reflectance pattern B at $f_x=f_i-f_{i-1}$, $f_y=0$. Due to the convolution operation seen in eqn. 3.18, this energy will modulate pattern information i according to

$$b\delta(f_i - f_{i-1},0) * \left[S_f M\left(\frac{P_i(f_x, f_y)}{2} + \frac{P_i(f_x - f_i, f_y)}{4} + \frac{P(f_x + f_i, f_y)}{4} \right) \right]$$
$$= bS_f M\left(\frac{P_i(f_x - (f_i - f_{i-1}), f_y)}{2} + \frac{P_i(f_x - f_i - (f_i - f_{i-1}), f_y)}{4} + \frac{P_i(f_x + f_i - (f_i - f_{i-1}), f_y)}{4} \right)$$
$$= bS_f M\left(\frac{P_i(f_x - (f_i - f_{i-1}), f_y)}{2} + \frac{P_i(f_x - 2f_i + f_{i-1}, f_y)}{4} + \frac{P_i(f_x + f_{i-1}, f_y)}{4} \right)$$
(3.27)

where b is some constant. The last term is the most significant. Note here how an image of pattern p_i is being placed directly within the bandwidth of pattern p_{i-1}. The effect is to create instances of "false positive" detection within each isolated binary pattern image.

Similarly, energy in B at low frequencies will create aliased images of each pattern which fall within its own isolated bandwidth. Thus, it may cause spatial intensity distortion in the isolated pattern images, the severity of which is related to the amount of low-frequency energy in B itself. The primary effect here will be the introduction of "misses," that is, failed detections, within each isolated binary pattern image due to intensity modulation.

Gamma Interference

As shown in reference [75], gamma distortion causes significant bleeding of signal energy into nearby harmonic bands. In the context of pattern modulation, gamma related effects are only likely to occur when using lower modulation frequencies (that is, a frequency low enough that modulation at a harmonic multiple is even possible for a specific device implementation). Nevertheless, in a worst case scenario, the pattern modulation frequencies may fall on exact harmonic multiples of each other. Kai et. al. have derived a simple algebraic formula to calculate the precise ratio of energy actually contained in these gamma induced harmonic sidebands, given as

$$\frac{B_{k+1}}{B_k} = \frac{\gamma - k}{\gamma + k + 1} \tag{3.28}$$

where B_k indicates the signal energy at the kth harmonic, and γ is the measure of projector gamma.

Note the significance of the gamma induced interference; equation (3.5) shows that for a gamma value of 2 (a realistic estimate of a gamma distortion encountered in a digital projector) there will be an influence of 25% of the signal energy on the first harmonic.

Original examples of composite pattern data used low frequency modulation, and were therefore highly subject to this particular type of interference. Recent implementations make use of much higher modulating frequencies, thereby reducing the influence of gamma distorted harmonics.

Saturation Interference

Due to unpredictable surface reflectivity and lighting conditions, it is sometimes possible to encounter saturation or "clipping" effects in modulated pattern regions. By flattening the modulated pattern signal, spurious modulation images are created, potentially interfering with other channel information and introducing possible errors in pattern analysis.

Without loss of generality, the situation can be approximately modeled in a 1D case. Consider the signal $S(t)$, where

$$S(t) = \sum_{i=1}^{N} s_i(t) \tag{3.28}$$

and each of s_i is a pattern signal modulated by an in-phase cosine with frequency f_i. The distorted signal can be approximately modeled as

$$S'(t) = S(t) + -R(t)S(t) + bR(t) \tag{3.29}$$

where

$$0 \le b \le \max(S(t)) \tag{3.30}$$

and

$$R(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_1) * \prod\left(\frac{t}{T_2}\right) \tag{3.31}$$

Here, **Π** represents the rectangle function and $*$ indicates convolution.

The values T_1 and T_2 are defined as

$$T_1 = \frac{1}{f_{\text{gcd}}} \text{ where } f_{\text{gcd}} = \gcd(\{f_1, f_2 \cdots f_N\}) \tag{3.32}$$

and

$$T_2 = \frac{1}{2 f_m} \text{ where } f_m = \max(\{f_1, f_2 \cdots f_N\}) \tag{3.33}$$

Now, T_1 represents the period of occurrence of clipping events, under the assumption that saturation will occur at the maximum values of the summation. It can be shown that, under summation of two in-phase sinusoidal functions, these maximum values (regions of constructive interference) occur with a frequency equal to the greatest common devisor of the additive frequencies. The value T_2 is approximated as half of the smallest modulating component period owing to the fact that it is the highest frequency component present that will define the primary visible pattern of peaks in the summation, and thus will approximately define an upper bound on the width of the saturated region. Note that, in the event that the image intensity is such that saturation occurs in regions other than these peak areas, the model may no longer be considered accurate.

The significance of saturation interference can be seen under transformation to the Fourier domain. In this case

$$\Im\{S'(t)\} = \Im\{S(t)\} + \Im\{-R(t)\} * \Im\{S(t)\} + b\Im\{R(t)\} \tag{3.34}$$

where

$$\Im\{R(t)\} = \sum_{k=-\infty}^{\infty} T_2 \delta(f - \frac{k}{T_1}) Sa(\pi T_2 f) \tag{3.35}$$

and

$$\Im\{S(t)\} = \sum_{i=1}^{N} \Im\{s_i(t)\} * \left(\frac{\delta(f - f_i) + \delta(f + f_i)}{2} \right) \tag{3.36}$$

Thus, a certain amount of signal energy is added at subharmonic frequency multiples of kf_{gcd} and also, a certain amount is lost at each modulating frequency and surrounding $\pm kf_{\text{gcd}}$ frequencies. The first effect creates the appearance of, sometimes strong, low

frequency subharmonic peaks, while the second effect can cause direct interference with nearby pattern channels.

Isolated Regions

Consider the following situation: a modified composite pattern is projected upon a surface with significant, and very abrupt, surface variation, such as a pile of wooden blocks. It is very likely, in a realistic scenario such as this, that one will encounter small, isolated segments of each MCP stripe. It is desirable to know, therefore, what is the minimum amount of stripe information necessary to make an accurate assessment of the stripe's code-word identity. Would a stripe segment that is only a few pixels wide be properly recognized, for example? And if not, then what is the smallest region that will?

As in equation eqn. 3.7 and 3.15, a recovered MCP image can be modeled as

$$c(x, y) = A(x, y) + B(x, y)d(x, y) \tag{3.37}$$

where

$$d(x, y) = S\big(p(x, y)\big) \tag{3.38}$$

It may be possible that *c* effectively contains an isolated segment of pattern *p*, due either to the effects of *B*, or the effects of *S*. In either case, the result may be modeled by replacing $p(x,y)$ with $p'(x,y)$ where *p'* is a summation of windowed patterns, that is

$$p'(x, y) = \sum_{i=1}^{m} w_i \, p(x, y) \tag{3.39}$$

Where each w_i (for i=1…m) represents a window function that effectively isolates, within pattern *p*, an individual pattern section isolated in the captured image *c*.

Consider, for a moment, the effect of isolating any one of these windowed segments, that is, let

$$p'_q(x,y) = w_q\, p(x,y) \tag{3.40}$$

where 1≤q≤m. In this case, we have

$$\begin{aligned} & c(x,y) \\ & = A(x,y) + B(x,y)d(x,y) \\ & = A(x,y) + B(x,y)Sp'_q(x,y) \\ & = A(x,y) + B(x,y)Sw_q\, p(x,y) \end{aligned} \tag{3.41}$$

Under the assumption of negligible A and constant, unit-value B, this becomes

$$c(x,y) = S\left(w_q\, p(x,y)\right) \tag{3.42}$$

Now, based on eqn. 3.21, the spectral representation is given by

$$C(f_x, f_y) = \Im(c(x,y)) = \Im(w_q) * S_f \sum_{i=0}^{N-1} M\left(\frac{P_i(f_x, f_y)}{2} + \frac{P_i(f_x - f_i, f_y)}{4} + \frac{P_i(f_x + f_i, f_y)}{4}\right) \tag{3.43}$$

Which, rearranged, is eqn. 3.22

$$C(f_x, f_y) = \Im(w_q) * \sum_{i=0}^{N-1} S_f\, \frac{M}{2} P_i(f_x, f_y) + \Im(w_q) * \sum_{i=0}^{N-1} S_f\, \frac{M}{2}\left(\frac{P_i(f_x - f_i, f_y)}{2} + \frac{P_i(f_x + f_i, f_y)}{2}\right) \tag{3.44}$$

Note that, due to distributivity of the convolution operation, this may be written

$$\begin{aligned} C(f_x, f_y) = & \sum_{i=0}^{N-1} S_f\, \frac{M}{2} \Im(w_q) * P_i(f_x, f_y) \\ & + \sum_{i=0}^{N-1} S_f\, \frac{M}{2}\left(\frac{\Im(w_q) * P_i(f_x - f_i, f_y)}{2} + \frac{\Im(w_q) * P_i(f_x + f_i, f_y)}{2}\right) \end{aligned} \tag{3.45}$$

Thus, the result of isolation of any given pattern segment can be shown to be convolution of the 2D Fourier transform of the window function with the original modulated pattern information.

In addition, a similar convolution takes place for all un-modulated components (including any DC within modulated stripes) introducing additional energy into numerous frequency bands.

Due to this fact, certain assumptions can be made regarding the effect of shape or coloration on the recovery of pattern information. For example, a very thin isolated section of the target surface would, according to the well-documented characteristics of spectral transformation, have a very wide-bandwidth Fourier representation. Thus, when the inherently narrow bandpass filter is applied to isolate each modulation band, it will result in a loss of information in the windowing function. This will cause familiar effects, such as ringing or "smearing" around step-edges. On the other hand, very wide isolated segments would have a narrow-bandwidth Fourier representation, potentially resulting in much less distortion.

In order to test the practical accuracy of this model, an experiment was performed. The input image (Figure 2.3a) is first collected by projecting a composite pattern onto a flat, white surface. A window was applied to the image, mimicking the effects of a region isolated due to subject surface variation (Figure 2.3b) and was processed for frequency selection and stripe peak isolation according to the MCP methodology. Upon demodulation, the window is again applied to the result image, isolating a demodulated region as seen in (Figure 2.3c). Isolated pixels that are within this windowed region "AND" within the isolated modulated stripe regions are considered correct, while those which fall outside (Figure 2.3d) are considered false positives. Using the isolated stripe mask, the fraction of correctly identified snake pixels in the windowed region was recorded, as was the average length (in pixels) of the false positive regions

that extend beyond the window edges. An additional false positive measure (False II) measures the length of positives which occur inside the window, but on stripe peaks specifically isolated as unmodulated regions (which, by definition, should not contain positive information in any modulation band). False II regions are due exclusively to the distortion energy introduced by the windowing function on unmodulated stripes.

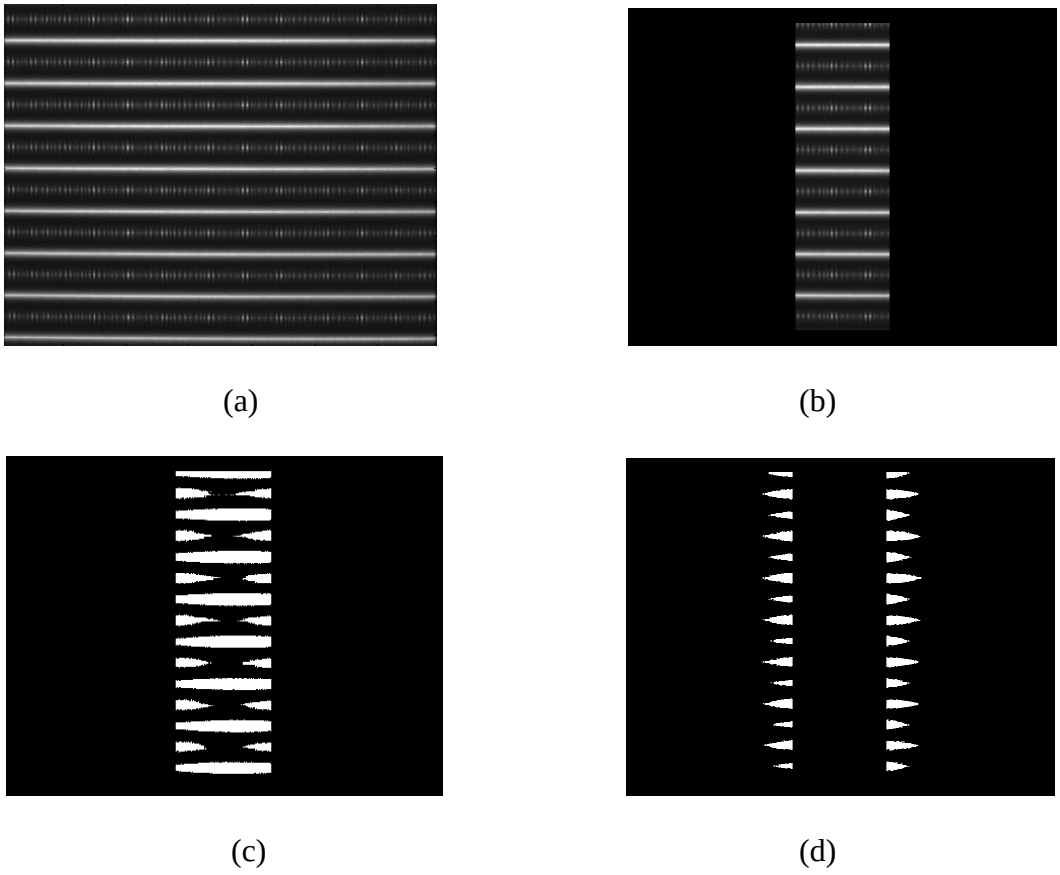

Figure 2.3, (a) input test image, (b) windowed input, (c) windowed demodulated region, (d) inverse-windowed demodulated region

The experiment was performed on two separate input images; each collected using a pattern type developed specifically for testing purposes. The first input was collected using a pattern that contains only a single modulation frequency (but modulating a fixed number of stripes; 16). The second input was collected using a

pattern containing four unique modulation frequencies, as is the current standard for MCP, but used in an 'all on' manner. That is, it is equivalent to a binary pattern in which the only encoded value is '1111'. From this, a single indicated frequency band was demodulated, producing results similar to those seen in Figure 2.3(b) and (c).

Utilizing two pattern types in this way allows better isolation of effects, that is, one may measure separately the actual influence of the isolating window versus additional interference which may be caused by the influence of multiple modulation frequencies. An ideally demodulated image (with all identifications correct and false positive region lengths of zero) may be considered the experimental control.

Table 2.1, 92 cycles/FOV - single frequency modulated

Window width	Fraction correct	False II region width	False region length
301	1	19	47
201	1	30	65
101	1	25	83
51	1	13	102
25	1	6	92
15	1	4	119
7	1	2	114
3	1	1	116

Table 2.2, 92 cycles/FOV - four frequencies modulated

Window width	Fraction correct	False II region width	False region length
301	1	42	77
201	1	40	85
101	1	20	91
51	1	10	94
25	1	3	69
15	1	3	128
7	1	2	107
3	1	1	110

Table 2.3, 115 cycles/FOV - four frequencies modulated

Window width	Fraction correct	False II region width	False region length
301	1	47	73
201	1	40	88
101	1	20	98
51	1	10	103
25	1	4	123
15	1	3	88
7	1	2	109
3	1	1	111

Results indicate that distortion introduced due to segment isolation is much, much more likely to introduce false positive results than to cause failed detections. Note that the results also act in accordance to predictions, in that smaller windows generally result in proportionally larger distortion effects. Also of note, inclusion of multiple frequency bands does seem to create larger distortions as well; possibly due to cross-talk (the convolution of the windowing function may introduce signal energy from one modulated pattern into nearby energy bands, depending on its shape).

Chapter 3 Camera Model, Triangulation, and Calibration

In order to truly understand the structured light illumination 3D recovery process, one must be familiar with the underlying mathematical operations being performed. Thus, the following section presents the explanation, from reasonably basic principles, of exactly how a structured light scan allows recovery of 3D surface information, and an explanation of the necessary calibration processes involved.

3.1 World to Camera 3D Transformation

The transformation from an arbitrary 3D world coordinate system (X^w, Y^w, Z^w) to a camera-specific 3D coordinate system (X^c, Y^c, Z^c) is based on the following simple linear tranformation

$$\begin{bmatrix} X^c \\ Y^c \\ Z^c \end{bmatrix} = R \begin{bmatrix} X^w \\ Y^w \\ Z^w \end{bmatrix} + T \tag{4.1}$$

where

$$R = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \tag{4.2}$$

representing a rotation operation, and

$$T = \begin{bmatrix} X^0 \\ Y^0 \\ Z^0 \end{bmatrix} \tag{4.3}$$

representing a simple translation of the world coordinate system origin.

Alternatively, this may be written

$$\begin{bmatrix} X^c \\ Y^c \\ Z^c \end{bmatrix} = \begin{bmatrix} r_1 & r_2 & r_3 & X^0 \\ r_4 & r_5 & r_6 & Y^0 \\ r_7 & r_8 & r_9 & Z^0 \end{bmatrix} \begin{bmatrix} X^w \\ Y^w \\ Z^w \\ 1 \end{bmatrix} \tag{4.4}$$

3.2 Camera 3D to Camera 2D Transformation

This research utilized a pinhole camera model, a simple representation of which can be seen in Figure 3.1.

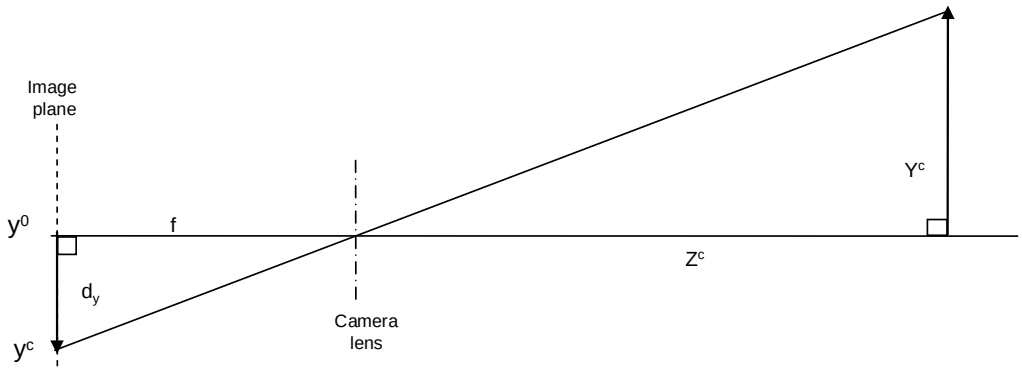

Figure 3.1, Simplified camera model (Y direction)

The relationship of similar triangles reveals

$$\frac{X^c}{Z^c} = \frac{d_x}{f} = \frac{(x^c - x^0)}{a^x f} \tag{4.5}$$

$$\frac{Y^c}{Z^c} = \frac{d_y}{f} = \frac{(y^c - y^0)}{a^y f} \tag{4.6}$$

Here, a^x and a^y represent scalar transformation factors (often taken as simply -1), f represents the camera focal length, x^c and y^c are defined as camera 2D space position coordinates, and x^0 and y^0 represent the axial origin positions for this 2D space. Let

$$s^x = a^x f \tag{4.7}$$

$$s^y = a^y f \tag{4.8}$$

By substituting eqns. 4.7 and 4.8 into eqns. 4.5 and 4.6, and rearranging, one finds

$$X^c s^x + Z^c x^0 = Z^c x^c \tag{4.9}$$

$$Y^c s^y + Z^c y^0 = Z^c y^c \tag{4.10}$$

Notice how Z^c acts as a scalar multiplier affecting both x^c and y^c. Thus, let us define a scale factor s such that

$$s = Z^c \tag{4.11}$$

In this case, the transformation may be represented by

$$\begin{bmatrix} sx^c \\ sy^c \\ s \end{bmatrix} = \begin{bmatrix} s^x & 0 & x^0 \\ 0 & s^y & y^0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X^c \\ Y^c \\ Z^c \end{bmatrix} \tag{4.12}$$

3.3 World to Camera 2D Transformation

Combining eqn. 4.12 and eqn. 4.4, one finds a system representation defined by

$$\begin{bmatrix} sx^c \\ sy^c \\ s \end{bmatrix} = \begin{bmatrix} s^x & 0 & x^0 \\ 0 & s^y & y^0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_1 & r_2 & r_3 & X^0 \\ r_4 & r_5 & r_6 & Y^0 \\ r_7 & r_8 & r_9 & Z^0 \end{bmatrix} \begin{bmatrix} X^w \\ Y^w \\ Z^w \\ 1 \end{bmatrix} \tag{4.13}$$

Or, simplifying

$$\begin{bmatrix} sx^c \\ sy^c \\ s \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X^w \\ Y^w \\ Z^w \\ 1 \end{bmatrix} \tag{4.14}$$

creating a basic representation commonly known as the Hall model [76]. Note that camera lens distortion is not accounted for in the model. However, as necessary, it may be accounted for by a corrective process at some other point in the calibration or imaging procedures [77-79].

3.4 Calibration

From 4.14, we can see that

$$x^c = \frac{sx^c}{s} = \frac{m_{11}X^w + m_{12}Y^w + m_{13}Z^w + m_{14}}{m_{31}X^w + m_{32}Y^w + m_{33}Z^w + m_{34}} \tag{4.15}$$

$$y^c = \frac{sy^c}{s} = \frac{m_{21}X^w + m_{22}Y^w + m_{23}Z^w + m_{24}}{m_{31}X^w + m_{32}Y^w + m_{33}Z^w + m_{34}} \tag{4.16}$$

Rearranging, these may be expressed as

$$x^c(m_{31}X^w + m_{32}Y^w + m_{33}Z^w + m_{34}) = m_{11}X^w + m_{12}Y^w + m_{13}Z^w + m_{14} \tag{4.17}$$

and

$$y^c(m_{31}X^w + m_{32}Y^w + m_{33}Z^w + m_{34}) = m_{21}X^w + m_{22}Y^w + m_{23}Z^w + m_{24} \tag{4.18}$$

that is

$$m_{34}x^c = m_{11}X^w + m_{12}Y^w + m_{13}Z^w + m_{14} - m_{31}x^cX^w - m_{32}x^cY^w - m_{33}x^cZ^w \tag{4.19}$$

$$m_{34}y^c = m_{21}X^w + m_{22}Y^w + m_{23}Z^w + m_{24} - m_{31}y^cX^w - m_{32}y^cY^w - m_{33}y^cZ^w \tag{4.20}$$

The significance of these representations will become clear shortly.

Now, define

$$m_c = \left[m_{11}\ m_{12}\ m_{13}\ m_{14}\ m_{21} \ldots m_{33}\right]^T \tag{4.21}$$

And let

$$m_{34} = 1 \tag{4.22}$$

Given i predefined points where X^w_i, Y^w_i, Z^w_i, x^c_i, y^c_i are known for all i (these points normally being supplied by the image of a calibration target), one can define

$$Am_c = B \tag{4.23}$$

by repeated vertical concatenation of vectors defined by 4.19 and 4.20, that is

$$A_{2i-1} = \begin{bmatrix} X_i^w & Y_i^w & Z_i^w & 1 & 0 & 0 & 0 & 0 & -x_i^c X_i^w & -x_i^c Y_i^w & -x_i^c Z_i^w \end{bmatrix} \tag{4.24}$$

$$B_{2i-1} = \begin{bmatrix} x_i^c \end{bmatrix} \tag{4.25}$$

and

$$A_{2i} = \begin{bmatrix} 0 & 0 & 0 & 0 & X_i^w & Y_i^w & Z_i^w & 1 & -y_i^c X_i^w & -y_i^c Y_i^w & -y_i^c Z_i^w \end{bmatrix} \tag{4.26}$$

$$B_{2i} = \begin{bmatrix} y_i^c \end{bmatrix} \tag{4.27}$$

The pseudo-inverse method least squares solution estimate for m_c can be found as

$$m_c = \left(A^T A\right)^{-1} A^T B \tag{4.28}$$

In a similar way, one may define a relationship between a point in world-space and the projector 2D pixel which illuminates it.

$$\begin{bmatrix} s_p x^p \\ s_p y^p \\ s_p \end{bmatrix} = \begin{bmatrix} m^p{}_{11} & m^p{}_{12} & m^p{}_{13} & m^p{}_{14} \\ m^p{}_{21} & m^p{}_{22} & m^p{}_{23} & m^p{}_{24} \\ m^p{}_{31} & m^p{}_{32} & m^p{}_{33} & m^p{}_{34} \end{bmatrix} \begin{bmatrix} X^w \\ Y^w \\ Z^w \\ 1 \end{bmatrix} \tag{4.29}$$

In this case, recovery of x^p and y^p is performed according to the structured light illumination method being used (that is, these values will be defined by specific structured light pattern code values, and must be obtained by performing a scan). Thus

$$x^p = \frac{s_p x^p}{s_p} = \frac{m^p{}_{11} X^w + m^p{}_{12} Y^w + m^p{}_{13} Z^w + m^p{}_{14}}{m^p{}_{31} X^w + m^p{}_{32} Y^w + m^p{}_{33} Z^w + m^p{}_{34}} \tag{4.30}$$

$$y^p = \frac{s_p y^p}{s_p} = \frac{m^p{}_{21} X^w + m^p{}_{22} Y^w + m^p{}_{23} Z^w + m^p{}_{24}}{m^p{}_{31} X^w + m^p{}_{32} Y^w + m^p{}_{33} Z^w + m^p{}_{34}} \tag{4.31}$$

Since the model for the projector is effectively identical to that of the camera, one may again utilize a least squares approximation to recover a calibration vector m_p, allowing transformation from world 3D coordinates to equivalent projector space 2D.

Defining

$$m_p = \left[m^p{}_{11} \; m^p{}_{12} \; m^p{}_{13} \; m^p{}_{14} \; m^p{}_{21} \ldots m^p{}_{33} \right]^T \tag{4.32}$$

and letting

$$m^p{}_{34} = 1 \tag{4.33}$$

and utilizing a scan of calibration target to define a series of i calibration points for which $X^w{}_i$, $Y^w{}_i$, $Z^w{}_i$, $x^p{}_i$, and $y^p{}_i$ may be found, one may then form appropriate matrices A_p and B_p (similar to 4.24-4.27) and summarily recover m_p according to

$$m_p = \left(A_p{}^T A_p \right)^{-1} A_p{}^T B_p \tag{4.34}$$

3.5 World 3D Recovery

Note that by properly aligning the camera and projector, it is possible to effectively remove one of the projector dimensions from consideration. Thus, the following scenario assumes that the structured light method used will allow recovery of

only y^p. Upon successful calibration, the m_c and m_p transformation vectors have been defined. Now, for any given point in a scan result image, one may collect values for x^c, y^c, (defined as simple camera space coordinates for the point) and y^p (defined by the structured light pattern projection information recovered at the point).

From 4.19 and 4.20, one may again create another useful representation

$$(m_{11} - m_{31}x^c)X^w + (m_{12} - m_{32}x^c)Y^w + (m_{13} - m_{33}x^c)Z^w = m_{34}x^c - m_{14} \qquad (4.35)$$

and similarly, one may define

$$(m_{21} - m_{31}y^c)X^w + (m_{22} - m_{32}y^c)Y^w + (m_{23} - m_{33}y^c)Z^w = m_{34}y^c - m_{24} \qquad (4.36)$$

and

$$(m^p{}_{21} - m^p{}_{31}y^p)X^w + (m^p{}_{22} - m^p{}_{32}y^p)Y^w + (m^p{}_{23} - m^p{}_{33}y^p)Z^w = m^p{}_{34}y^p - m^p{}_{24} \qquad (4.37)$$

Thus, we have three independent equations with three unknowns (X^w, Y^w, and Z^w). From eqns. 4.35, 4.36 and 4.37 we may form the following representation

$$\begin{bmatrix} m_{11} - m_{31}x^c & m_{12} - mx_{32}{}^c & m_{13} - m_{33}x^c \\ m_{21} - m_{31}y^c & m_{22} - m_{32}y^c & m_{23} - m_{33}y^c \\ m^p{}_{21} - m^p{}_{31}y^p & m^p{}_{22} - m^p{}_{32}y^p & m^p{}_{23} - m^p{}_{33}y^p \end{bmatrix} \begin{bmatrix} X^w \\ Y^w \\ Z^w \end{bmatrix} = \begin{bmatrix} m_{34}x^c - m_{14} \\ m_{34}y^c - m_{24} \\ m^p{}_{34}y^p - m^p{}_{24} \end{bmatrix} \qquad (4.38)$$

so, let

$$C = \begin{bmatrix} m_{11} - m_{31}x^c & m_{12} - mx_{32}{}^c & m_{13} - m_{33}x^c \\ m_{21} - m_{31}y^c & m_{22} - m_{32}y^c & m_{23} - m_{33}y^c \\ m^p{}_{21} - m^p{}_{31}y^p & m^p{}_{22} - m^p{}_{32}y^p & m^p{}_{23} - m^p{}_{33}y^p \end{bmatrix} \qquad (4.39)$$

and

$$D = \begin{bmatrix} m_{34}x^c - m_{14} \\ m_{34}y^c - m_{24} \\ m^p{}_{34}y^p - m^p{}_{24} \end{bmatrix} \qquad (4.40)$$

Then, it can be seen that

$$C\begin{bmatrix} X^w \\ Y^w \\ Z^w \end{bmatrix} = D \tag{4.41}$$

and so, any world coordinate 3D point can be calculated as

$$\begin{bmatrix} X^w \\ Y^w \\ Z^w \end{bmatrix} = C^{-1}D \tag{4.42}$$

3.6 Calibration Process Summary

A summary of the basic SLI calibration process is as follows: first, a calibration target is created, consisting of i=12 or more calibration target points with X^w_i, Y^w_i, and Z^w_i coordinates known (as defined in some known coordinate system; for example, millimeter units as measured from an origin point defined upon the calibration object itself).

Next, a SLI scan of this calibration object is acquired, creating two particular images: a "phase map" in which each point in the image space is marked with the corresponding projector y^p encoded value (this is the image which may actually be processed to collect a 3D representation of the object), and an 'albedo image', which is normally a basic visual representation of the object (that is, a largely unaltered digital photograph).

Using the albedo image, one locates the calibration target points (either automatically or manually) and thus acquires x^c_i, and y^c_i for each of the i calibration target points. The y^p_i encoded value of the equivalent position in the phase map is collected and recorded.

Finally, the necessary matrices (A, B, A_p and B_p, as in eqns. 4.28 and 4.34) are created using the collected calibration target point data, and calculation to acquire m_c and m_p is performed.

In the specific context of research performed, the SLI method used for calibration purposes was multifrequency phase modulation profilometry (PMP, as described in chapter 1) [43]. The calibration object was a wedge shaped grid with 18 calibration target points, marked by circles. During the calibration process, data regarding each target point or circle (including target position) was collected using a semi-automatic procedure implemented by a specifically designed software application shown in Figure 3.2. This software application performed all necessary calculations and summarily recorded the relevant calibration data in appropriate data files for later use.

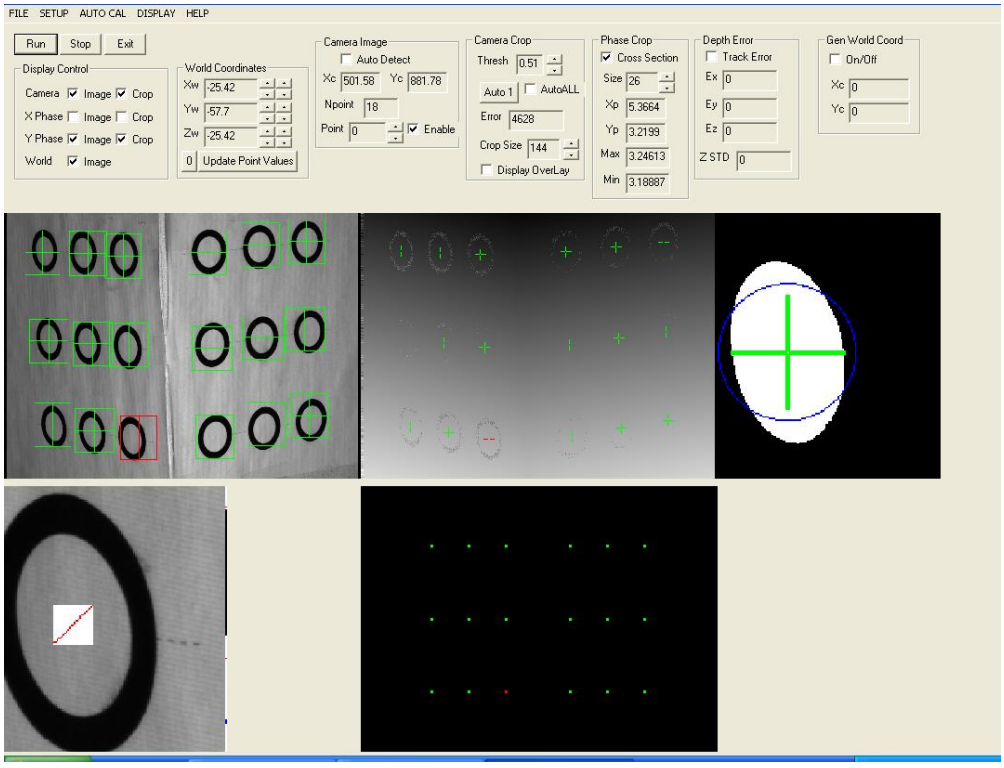

Figure 3.2, Calibration software

In the figure, one can see displays of the following data images (clockwise from the top-leftmost image): the albedo image of a calibration target, the phase image created from the PMP scan of the target, an isolated localized region representing the inner portion of a target circle, a small grid of dots representing the corresponding world-coordinate position of the selected target circle, and finally the albedo image subsection containing the selected target circle.

3.7 Special Modified Composite Pattern Calibration Procedure

Calibrating for PMP scanning can be performed directly using the software and procedure described in the previous section. However, due to the nature of MCP, this calibration method is unsuitable. MCP does not create a dense phase image as PMP does, and so, finding the camera-projector correspondence for an arbitrary point on a calibration target is difficult and, if attempted, would require potentially detrimental interpolation. Additionally, the presence of calibration targets, such as circles, in an MCP image will create areas of local interference during the demodulation process, potentially corrupting the recovered encoding information at the very spot where correct decoding is the most necessary. Also, as will be discussed in a later chapter, the strongly angled surface presented by a calibration object can also introduce difficulties into the demodulation process.

For these reasons, a slightly altered method was implemented in order to calibrate MCP scans. Note that, since MCP encoded stripe peaks correspond to a single row in projector space, and a single PMP phase value corresponds to a single row as well, it is, in fact, possible to find a correspondence between MCP peak code values and PMP phase

values. If one could, for a given system setup, acquire PMP calibration information as well as the mathematical correspondence between MCP encoding and PMP phase values, this would be all the information necessary to reconstruct a depth model from an MCP scan.

Thus, the following procedure is performed to acquire this effective "MCP calibration": first, the camera-projector system is set up with appropriate focus and alignment, etc. and a PMP calibration procedure is performed (thus, the system is now enabled for PMP scans and data reconstruction). Next, a PMP scan of a matte-white flat surface is acquired, followed immediately by a MCP scan of the surface. The MCP scan is then processed, yielding an image of stripes bearing base-10 encoding values. Finally, a least-squares linear fit is found between each encoding value and the phase value at the same (x^c, y^c) point in the PMP phase image. The linear fit coefficients are then recorded.

Using this information, it is possible to transform MCP code values to PMP phase values and then, using the PMP calibration data for the system, calculate world-coordinate data of the surface. It should be noted that the process performed this way yields 3D data which is, at best, only as accurate as the PMP scan used in initial calibration. However, any errors present in the MCP result (due to incorrect decoding) will adversely affect the linear fit calibration process, and so it is especially important that the MCP result data used in this calibration be as correct as possible, potentially necessitating additional post-processing correction procedures.

Chapter 4 MCP Processing

The updated Modified Composite Pattern (MCP) processing is a fairly complex affair. The process for early stage processing (effectively the CP demodulation and initial recovery of binary SLI data) is summarized in Figure 4.1.

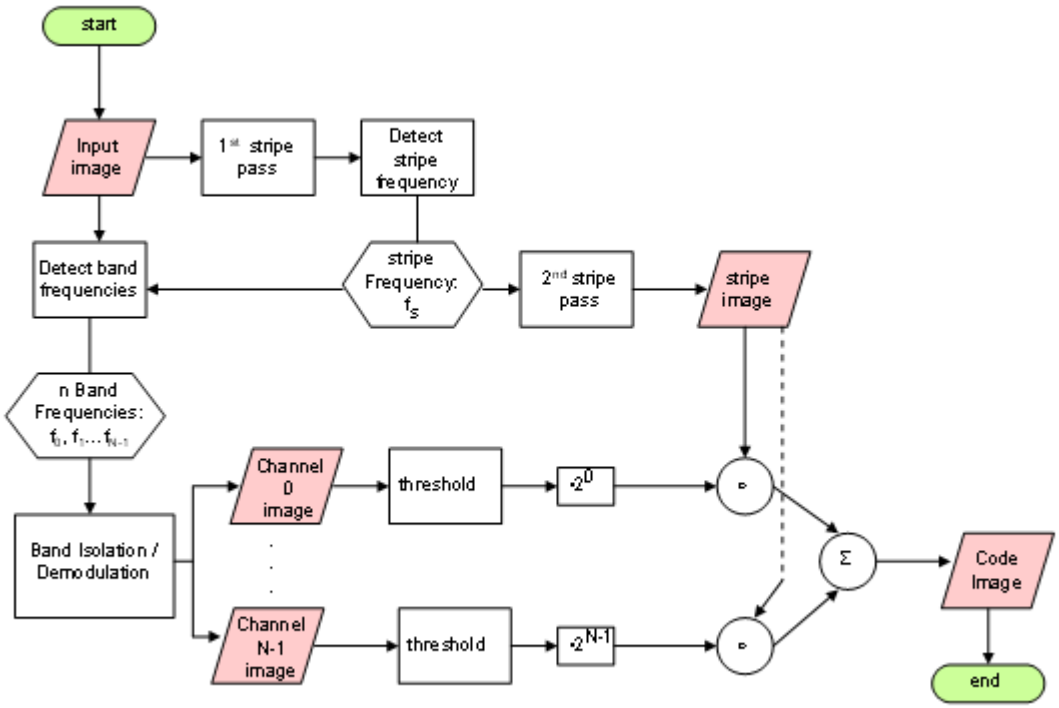

Figure 4.1, MCP flowchart

Parallelograms represent data images, while hexagons represent collected data variables to be used in specific processes. Rectangles represent an image processing step. The dots in the operation circles represent element-wise multiplication. The procedure can be effectively divided into two, largely parallel branches; a stripe detection branch, and a decoding branch.

Stripe detection utilizes an automated detection algorithm, followed by a frequency detection process to isolate the "stripe frequency." This is then used to perform an updated detection process, resulting in the final, accurate "stripe image."

Decoding begins, as seen in Fig. 5.1, with the automated detection of N modulation band frequencies (which may be enhanced by the use of prior knowledge of stripe frequency data, collected separately). These band frequencies are used to direct a filtration process which isolates N individual channel images. These images are then thresholded and element-wise multiplied by a constant value. The resulting images are then element-wise multiplied by the binary stripe image (effectively using the stripe image as a binary mask), and all images thus created are finally summed together.

The output of the process is a single image in which each isolated peak has been identified by its appropriate binary sequence code-word. This "code image" is then used as input to a second processing stage which makes use of various operations (such as morphological operations and Hidden Markov Model / Viterbi best path analysis) to allow even more accurate recovery of 3D surface data.

4.1 Peak Isolation

The hallmark of MCP is its use of spatial domain information to enhance the standard Composite Pattern depth measurement techniques. Utilizing characteristic information in the pattern itself, it is possible to reduce or prevent CP demodulation errors and simultaneously eliminate the influence of systematic errors on the recovered CP depth data.

The triangular cross section of the MCP was utilized in order to allow easy and accurate definition and recognition of pattern regions. Each peak runs the full length of the pattern, and will be distorted as it is projected onto a target surface. Isolation of the stripes on the surface can allow for recovery of 3D information in accordance with standard structured light triangulation methods. Pixel-accurate isolations of the stripe peaks can be automatically obtained using a linear filtering process illustrated in Figure 4.2. Inspired by the theory of digital unsharp masking, and similar to the ring filter technique used in reference [80], this filtering process was deemed most suited to this particular application due to the limited necessity for user interaction.

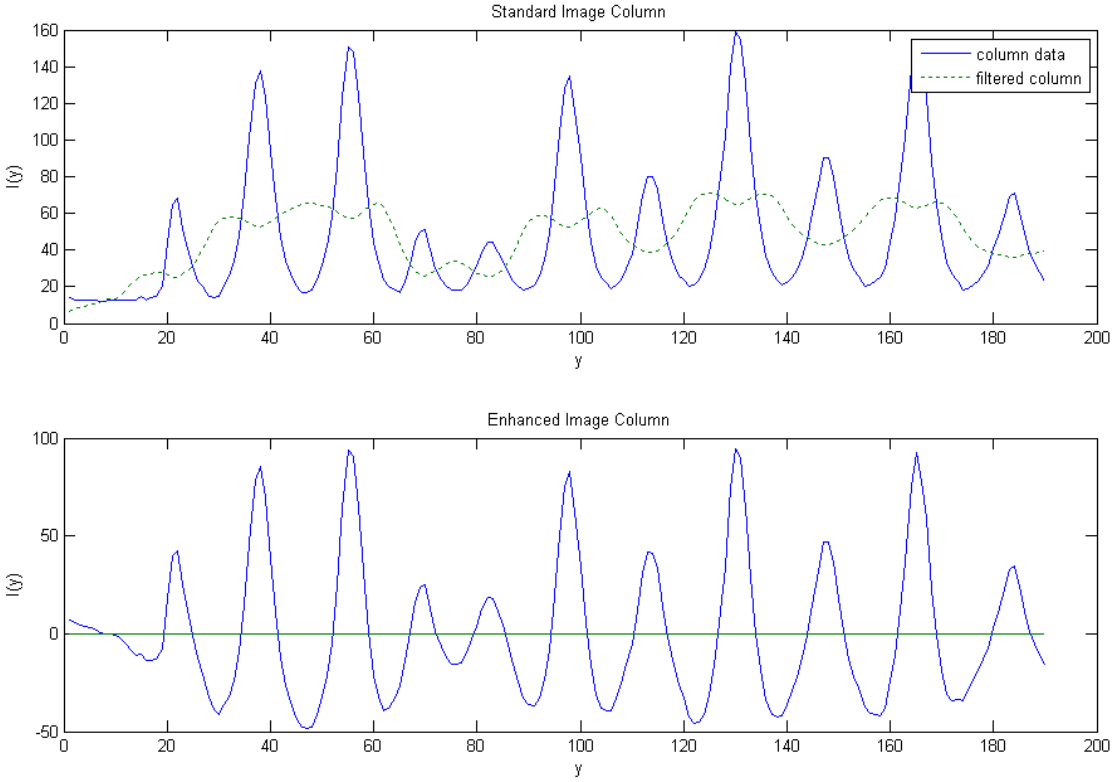

Figure 4.2, Adaptive threshold generation

The peak isolation process is quite simple. The image is first strongly smoothed with a simple linear filter; a single-column-wide rectangle with a length equal to the approximate peak wavelength, with a zero-valued section, possibly as small as one pixel, in the center (effectively an averaging filter with a "hole"). This smoothed image is then subtracted from the original, resulting in an approximate shifting of the mean of all stripes to zero while also enhancing the local contrast (advantageous, as it may allow one to reduce the loss of peak data during aggressive intensity thresholding).

The filter essentially defines an adaptive threshold that has the additional effect of enhancing locally symmetric peak areas. A simple threshold at zero operation on the filtered image will isolate the positive stripe regions from the negative. The actual peaks can then be isolated by a local search bounded by the edges of the positive stripe regions, performed either on the mean-shifted filtered image or the original, unfiltered image (notice that the filter has an additional undesirable effect of skewing asymmetric regions, which may create error in identifying peak locations. As such, it may be advantageous to perform final peak localization processes using the unfiltered image instead).

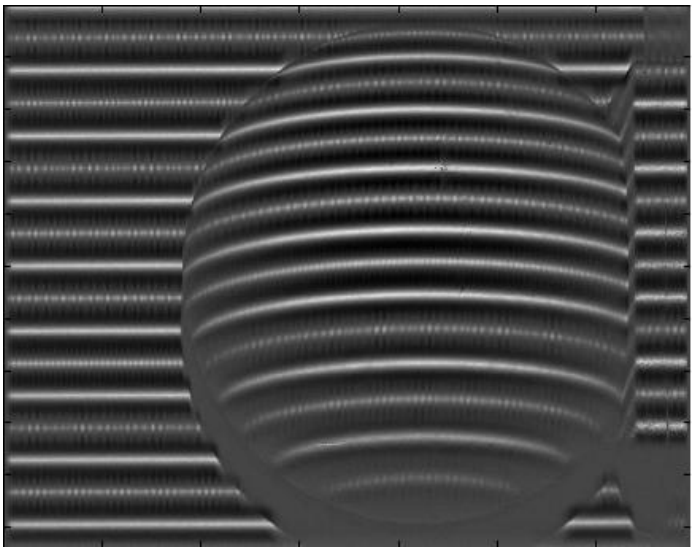

Figure 4.3, Enhanced image

Figure 4.4, Thresholded isolated region image

The process can actually be modeled by a single filter, defined by the following

$$F(y) = \delta(y) - \frac{1}{T_1 - T_2}\left(\Pi\left(\frac{y}{T_1}\right) - \Pi\left(\frac{y}{T_2}\right)\right) \tag{5.1}$$

Where

$$T_1 > T_2 \geq 1 \tag{5.2}$$

This can be easily shown to have a frequency response of

$$\Im\{F(y)\} = 1 - \frac{1}{T_1 - T_2}\left(\frac{T_1 \sin(\pi f T_1)}{\pi f T_1} - \frac{T_2 \sin(\pi f T_2)}{\pi f T_2}\right) \tag{5.3}$$

Of course, the situation could be considered in yet another slightly different way as well. Rather than a two step process (creating a filtered image, then threshold), one could simply utilize the adaptive threshold directly, reducing this to a single step process. However, it was deemed potentially advantageous to have the filtered image data available, rather than just the thresholding result, as it allows potential utilization of multiple threshold values, and may also be useful during the peak isolation process.

The peak isolation process is most effective when the size of the smoothing filter is roughly the size of an individual stripe wavelength. Nevertheless, even a filter with poorly chosen size parameters can still be used to adequate affect. Taking advantage of this fact, the implemented MCP analysis process utilizes a two step automated procedure for stripe detection. An initial pass performs stripe region isolation and thresholding using an initial stripe wavelength estimate. This region image is smoothed with a default sized moving average filter (to suppress harmonics) and transformed to Fourier space and analyzed to isolate the primary vertical frequency; that is, the effective stripe frequency.

This is accomplished very easily by considering the first quadrant of the Fourier space representation of the thresholded region image (after removing the DC component of the image), and summarily finding the location of maximum value along the first column. This value is then recorded as the stripe frequency and used to form a new estimate for stripe wavelength. The new estimated value is used for the second filtration pass, allowing for optimally isolated stripe peaks.

The final step in stripe isolation is, as indicated previously, the application of an intensity threshold. It was found, during early stages of the research, that utilization of an intensity threshold could often serve to demarcate certain object boundaries, and thus, could be useful in ensuring correct stripe identification.

4.2 Surface Inclination Segmentation

Composite Pattern Fourier domain information, which must be analyzed and demodulated in order to recover subject depth information, is unfortunately sensitive to local rotations due to surface shape. Consider the image in Figure 4.5 (left) and a cropped image of the corresponding Fourier domain power spectral density Figure 4.5 (right).

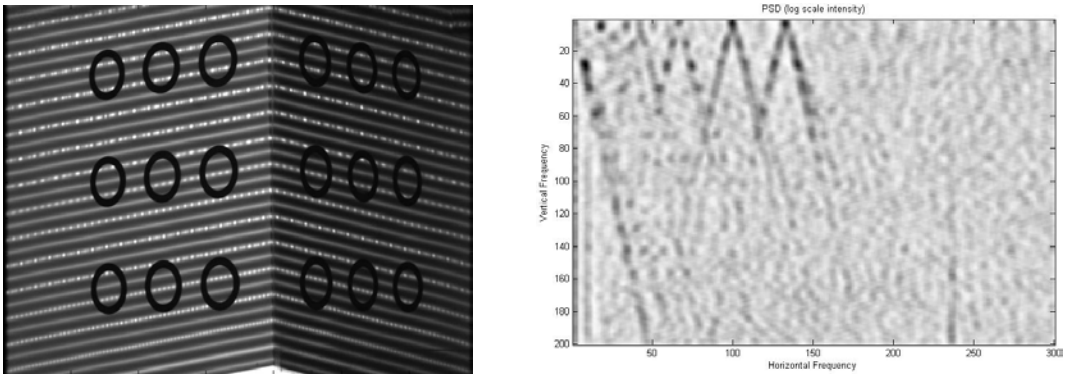

Figure 4.5, Example of stripe rotation effects

Note how the modulated image bands in Fourier space suffer noticeable rotation. If uncorrected, this information will create significant difficulties in demodulation and analysis, including difficulty in determining the appropriate modulating frequencies and in isolating the modulated information (note how the “crossing” of the bands will create

large amounts of inter-channel interference and will be detrimental to proper stripe identification).

Thus, in order to avoid this undesirable complication, it is necessary to perform correction of the image in the spatial domain before significant Fourier domain analysis can be performed. It is a simple task to segment the image into regions defined, for example, by the average local linear inclination identified in each. By then processing each region individually, the maximum amount of correct encoded pattern information can be obtained without corruption due to frequency shifting (similar to the technique in reference [81]).

By isolating the regions and dividing the original image into three or more inclination segmented images (one for highly positive inclination, one for highly negative inclination, and one for negligible inclination), appropriate rotations can be utilized to compensate, allowing processing of each to be performed according to the standard methodology. The resulting decoded stripe identity images may then be rotated back, once more creating a single image that contains all necessary information for depth calculation.

Isolation can be performed in the following manner: within the binary isolated stripe peak image I_p, at any non-zero point (x,y), consider the set of non-zero points in a local windowed region. Let

$$W(x, y) = \left[w_1, w_2, \cdots w_n\right] \tag{5.4}$$

represent the collection of such points, and for all i

$$w_i = (w_{i,x}, w_{i,y}) \tag{5.5}$$

represents the image space coordinates of each non-zero pixel.

Define a matrix X, such that for each row i

$$X[i,1] = w_{i,x} \tag{5.6}$$

and

$$X[i,2] = 1 \tag{5.7}$$

Next, define a column vector Y, such that for each row i,

$$Y[i] = w_{i,y} \tag{5.8}$$

One may then define the linear parameterization for this local window using the least-squares estimation, given by

$$Z = (X'X)^{-1} X'Y \tag{5.9}$$

In which case element Z[1,1] represents the pixel-wise slope for the best fit linear model of all non-zero pixels within the window.

Define $z(x,y)=Z[1,1]$ thus acquired from $W(x,y)$. We may now define a new image matrix L, such that

$$L[y,x] = \begin{cases} z(x,y) & for\ I_p[y,x] > 0 \\ 0 & for\ I_p[y,x] = 0 \end{cases} \tag{5.10}$$

Consider now three copies of image matrix I_p, appropriately named I_{p1}, I_{p2}, and I_{p3}. Given a positive inclination threshold t_p and negative inclination threshold t_n, one may form three segmented binary images in which

$$\begin{aligned} I_{p1}[y,x] &= 1 \qquad iff\ \ L[y,x] >= t_p \\ I_{p2}[y,x] &= 1 \qquad iff\ \ L[y,x] > t_n, \quad L[y,x] < t_p \\ I_{p3}[y,x] &= 1 \qquad iff\ \ L[y,x] <= t_n \end{aligned} \tag{5.11}$$

One may now define rotation compensation values r_p (for image I_{p1}) and r_n (for image I_{p3}) according to

$$r_p = -\tan^{-1}\left(\frac{\sum_{i=0}^{N_1} L_1[i]}{N_1}\right) \tag{5.12}$$

and

$$r_n = -\tan^{-1}\left(\frac{\sum_{i=0}^{N_3} L_3[i]}{N_3}\right) \tag{5.13}$$

where L_n represents the set of non-zero elements in the matrix defined by Hadamard multiplication of matrix L and I_{pn}. The value N_n represents the number of elements in each set L_n.

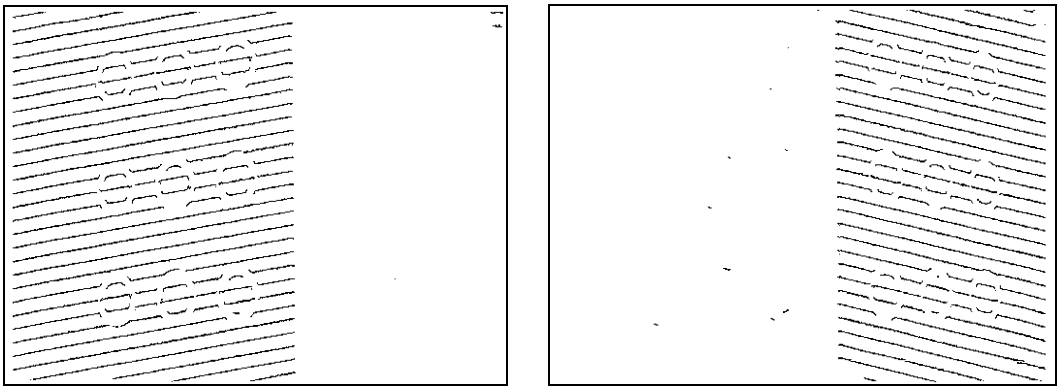

Figure 4.6, (left) Isolated positive inclination stripe regions, (right) Isolated negative inclination stripe regions

Figure 4.6 displays the results of inclination filtering, using a simple threshold at zero to isolate regions of positive inclination versus regions of negative inclination (note that only two regions, rather than three, were identified simply for demonstration purposes).

Performing morphological dilation (or closing) using a large structuring element on each I_p image allows the creation of three separate segmentation masks which may be applied to duplicates of the input image. Each image, thus segmented, can be corrected using the appropriate rotation compensation factor, and summarily processed according to the standard MCP processing methodology.

4.3 Analysis-By-Synthesis Target Frequency Isolation

A significant portion of Composite Pattern processing is performed in the spectral domain. The goal is to isolate the binary pattern information as accurately as possible in order to properly identify each stripe codeword identifier. With each stripe identity known, the surface depth can be calculated via triangulation (as in most structured light implementations). A secondary goal during project development is to design a method to complete the analysis process with minimal user input or observation. Early versions of the process required significant user manipulation, and were thus difficult to use and prone to error.

In order to recover CP depth information accurately, it is necessary to recover each individual pattern channel with as little error as possible. The fundamental aspect of this is the accurate identification of the apparent pattern modulation frequencies. Though the projected pattern modulation frequencies are known, the effects of the camera and projector optics will generally alter the frequencies significantly in the captured image.

To begin, a Fast Fourier Transform is applied to the input image. Each element is appropriately multiplied by its own complex conjugate in order to create a two dimensional power spectral density image. All quadrants of the PSD image are then

averaged (utilizing appropriate cropping and element-wise ordering in order to add corresponding positive and negative frequency components) to create a single image as shown in Figure 4.7.

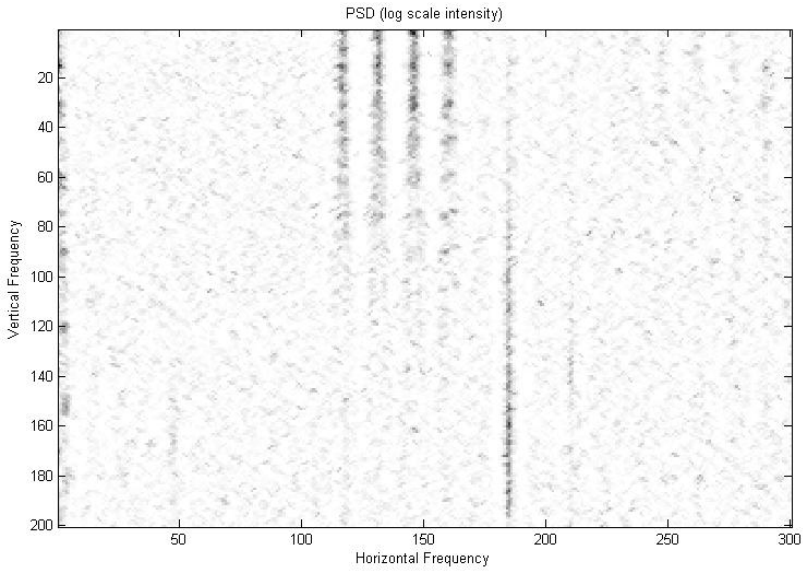

Figure 4.7, Four quadrant composite Fourier image

A simple column-wise summation process is employed (limited to vertical frequencies between zero and the calculated stripe frequency, though this empirically defined limit may need to be altered depending on circumstantial implementation characteristics), creating a data array similar to that shown in Figure 4.8.

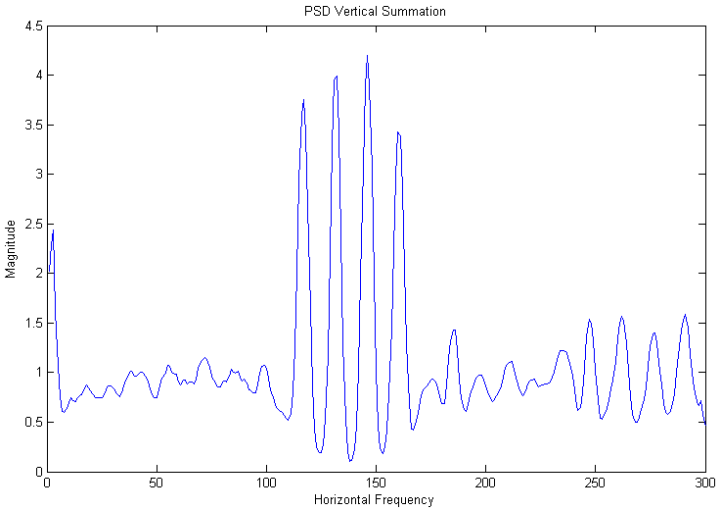

Figure 4.8, Vertical PSD summation

Note that the data averaging and summation processes can be expected to reduce the influence of uncorrelated noise on the underlying modulated data. A second array is then created by filtering this data with a 2nd derivative filter. Using these, the individual frequency elements are isolated with an analysis-by-synthesis method. The algorithm is summarized as follows:

- Isolate data array A
- Construct 2nd derivative (discrete Laplace) filtered array D
- Construct copy C of array A (as shown in Fig. 5.9 (top))
- Construct a blank data array B
- Begin loop - end after a pre-determined number of iterations
    - Find and record the location X of the maximum value M in C
    - Find nearby upper and lower zero values Z_+ and Z_- in D
    - Using M, X, Z_+ , and Z_- , define a discrete Gaussian curve, and add this curve to array B (shown in Fig. 5.9 (middle))
    - Redefine C according to C=A-B (shown in Fig 5.9 (bottom))
- End loop

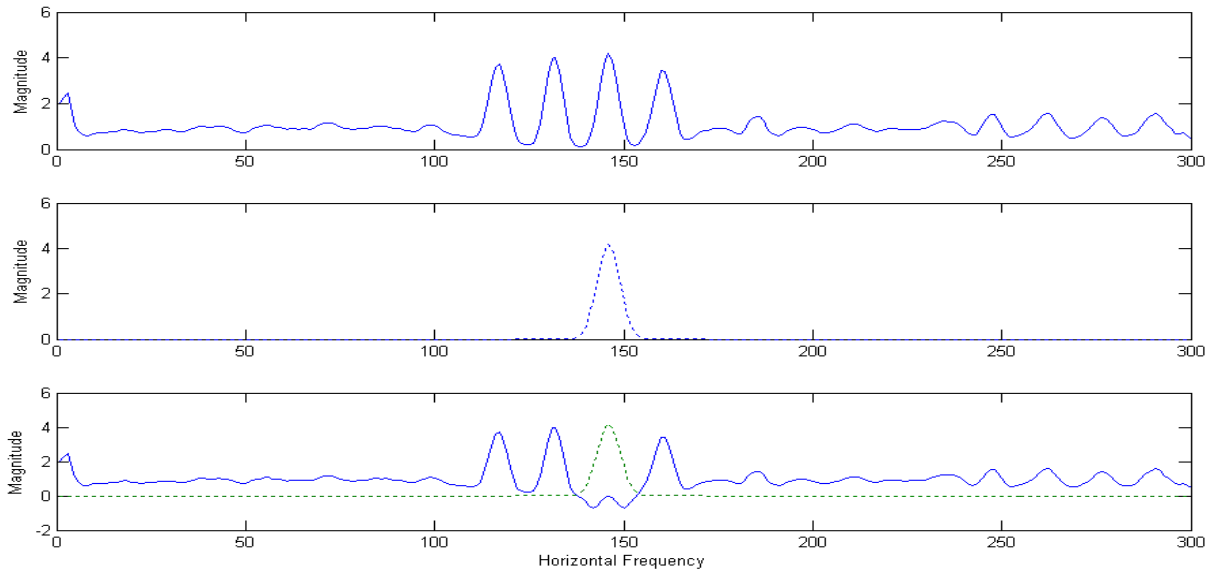

Figure 4.9, Analysis By Synthesis process, (top) initial array A, (middle) array B, where B represents isolated frequency band components, after one iteration, (bottom) updated array C where C=A-B, where the presence of B is indicated by the dotted curve

This process will effectively synthesize, in B, the information contained in A, while also collecting relevant data (thus it is "analysis by synthesis" (ABS)). The desired number of isolated frequencies, collected as recorded *X* values, determines the number of loop iterations to be performed. As part of the process, obviously incorrect frequencies (such as near-zero frequencies) can be eliminated from consideration.

4.4 Channel Isolation

Once the target modulating frequencies have been isolated, the desired pattern data may be isolated and demodulated. The basic process of CP demodulation is explained in-depth in chapter 2.

In the specific implantation of this project, the modulation bands are isolated using a Gaussian shaped bandpass filter which mimics the shape of the modulated energy, according to information gathered in the ABS procedure. This choice greatly reduces the likelihood of inter-channel interference at the cost of reduced signal energy.

The result is then inverse-transformed so that further processing may be performed in the spatial domain.

Note that, in contrast to the theoretical implementation, the isolated signals are not demodulated according to the classical methodology, and so, the resulting inverse-transformed spatial domain images are effectively still modulated by high-frequency sinusoids. This result is actually observably better than the alternative.

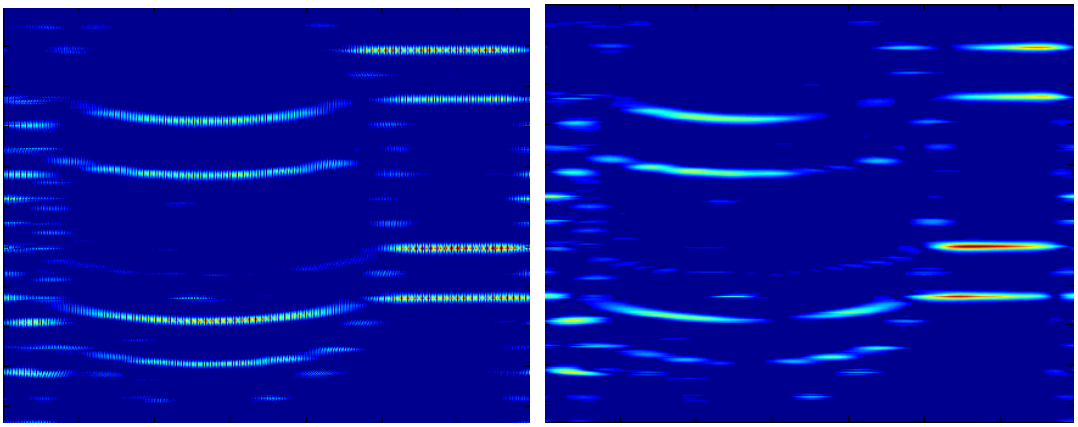

Figure 4.10, Frequency isolation comparison of methods example 1, unshifted (left) versus shifted (right)

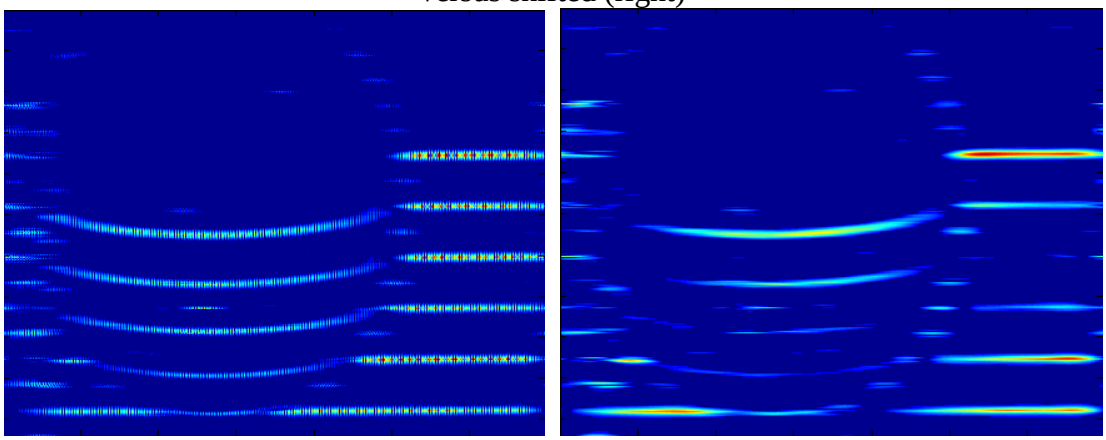

Figure 4.11, Frequency isolation comparision of methods example 2, unshifted (left) versus shifted (right)

As seen in Figure 4.10 and Figure 4.11, the act of demodulating (that is, shifting each band to zero frequency) the isolated frequency bands results in a significant loss of

information. This is likely due to optical distortion introducing small local deviations in frequency within each modulated stripes.

The "isolated-but-not-frequency-shifted" spatial images contain sinusoidal modulation, creating the presence of numerous small, empty sections (at the sinusoidal zero crossings). These are, however, easily dealt with. The isolated images need only be thresholded and summarily dilated with a very small morphological structuring element. This very minor correction allows easy recovery of the full stripe information.

4.5 Thresholding and Initial Error Estimation

The final stages of processing for a MCP image take place once again in the spatial domain. Once the isolated pattern data images have been created, they must be thresholded to isolate the stripe identity code-word values. These identity values may then be applied to the stripe peak mask. This image may then be used for 3D data recovery.

Modified Composite Pattern encoding is binary in nature, while in practice, illuminating a surface will produce an obviously non-binary image. This is further complicated due to pattern modulation / demodulation. As such, a binarizing threshold must be applied to the demodulated pattern images.

At the same time, attempting to define an intensity decision boundary for thresholding is complicated due to the fact that key aspects of threshold definition (average intensity values, interference and noise levels etc.) are highly variable, even in different areas of a single image, due to specifics of subject and system setup. For this reason, it is necessary to apply a reasonable, data specific statistical analysis and

thresholding method that takes into account known statistical characteristics of the image being processed. These statistics can then be carried forward into the specifications of the HMM very easily. Using the HMM, a more realistic image analysis can be performed (if necessary) using a multiple pass Viterbi analysis methodology, as will be explained in later sections.

For the purposes of experimentation, the threshold level was chosen to be two standard deviations from the statistical mean value of the isolated spatial intensity image of each demodulated data channel. This choice was made as this threshold produced satisfactory results during early stage testing, in addition to providing a statistically consistent threshold methodology, suitable for further study. As such, a characterization of this threshold follows, in order to assess the suitability of the method in practical scenarios.

Threshold analysis

Any demodulated MCP pattern image can be considered as consisting of two parts: the triangular encoded "active regions", and low-intensity "inactive" regions which occur between the peaks. Let us consider a one-dimensional image *B*, consisting of *N* total regions, a number *a* of which are active and (*N*-*a*) of which are inactive. The expectation value for the region *B* can be calculated as the sum of the expectation values of the active regions and the inactive regions plus the expectation value of an additive noise field, that is

$$E\{B\} = E\{B_a\} + E\{B_i\} + E\{noise\} \tag{5.14}$$

Given that the probability of a region being active is P_a, and that the mean value of a given active region is M_a, it is clear that

$$E\{B_a\} = P_a M_a \tag{5.15}$$

And similarly for inactive regions

$$E\{B_i\} = P_i M_i \tag{5.16}$$

And for noise, let

$$E\{noise\} = \mu_n \tag{5.17}$$

Owing to the fact that an active region consists of a single symmetric triangular waveform, it can be shown that the mean active region value can be easily calculated as

$$M_a = \frac{E\{H\}}{2} \tag{5.18}$$

Where

$$E\{H\} = \frac{1}{a}\sum_{i=1}^{a} H_i \tag{5.19}$$

which describes the mean height of all active peaks.

From the description of the image, it can be seen that

$$P_a = \frac{a}{N} \tag{5.20}$$

And also that

$$P_i = \frac{(N-a)}{N} \tag{5.21}$$

and

$$M_i = 0 \tag{5.22}$$

And thus

$$E\{B\} = \frac{aE\{H\}}{2N} + \mu_n \tag{5.23}$$

The variance of the image can be shown to be given by

$$\sigma^2 = E\{(B - E\{B\})^2\} = E\{B^2\} - E\{B\}^2 \tag{5.24}$$

Where

$$E\{B^2\} = E\{B_a{}^2\} + E\{B_i{}^2\} + E\{noise^2\} \tag{5.25}$$

Under the same assumptions described for the previous equations, it can be shown that

$$E\{B^2\} = \frac{aE\{H^2\}}{3N} + \sigma_n^2 \tag{5.26}$$

And thus

$$\sigma^2 = \frac{aE\{H^2\}}{3N} + \sigma_n^2 - \left(\frac{aE\{H\}}{2N} + \mu_n\right)^2 \tag{5.27}$$

or

$$\sigma^2 = \frac{aE\{H^2\}}{3N} + \sigma_n^2 - \left(\frac{aE\{H\}}{2N}\right)^2 - \frac{\mu_n aE\{H\}}{N} - \mu_n^2 \tag{5.28}$$

It can be shown that the situation of largest variance occurs in the event in which half of the peaks have a large height value, while the rest have a much lower height (assuming an even number of active regions). In this situation of maximum variance, the question becomes: what is the smallest possible value of peak that will not be thresholded away under an image-wide threshold of two standard deviations above the mean of the image? For the sake of convenience, let the situation be defined such that the active region peak height mean is normalized to a value of 1.

Define values x (small peak height) and H_b (large peak height) such that

$$\frac{H_b \frac{a}{2}}{a} + \frac{x \frac{a}{2}}{a} = \frac{H_b}{2} + \frac{x}{2} = E\{H\} = 1 \tag{5.29}$$

as described by the situation of maximum variance. The situation is illustrated in Figure 4.12 below.

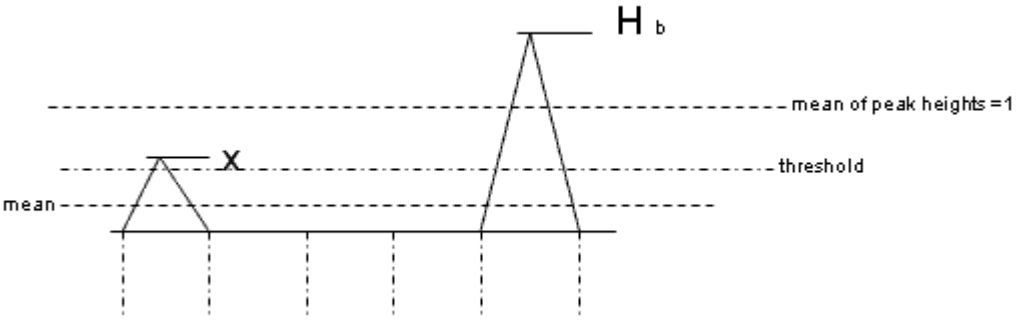

Figure 4.12, Peak thresholding example

Similarly,

$$\frac{H_b{}^2}{2} + \frac{x^2}{2} = E\{H^2\} \tag{5.30}$$

Using equation 5.29, one may solve for the expectation of squared heights in terms of x only, as

$$E\{H^2\} = x^2 - 2x + 2 \tag{5.31}$$

According to the proposed threshold, let

$$x = 2\sigma + \frac{aE\{H\}}{2N} + \mu_n \tag{5.32}$$

$$0 < x \leq 1 \tag{5.33}$$

wherein peaks of height x will fall exactly on the threshold value.

And since $E\{H\}$=1,

$$x = 2\sigma + \frac{a}{2N} + \mu_n \tag{5.34}$$

and also

$$\sigma^2 = \frac{aE\{H^2\}}{3N} + \sigma_n^2 - \left(\frac{a}{2N}\right)^2 - \frac{\mu_n a}{N} - \mu_n^2 \tag{5.35}$$

Now, define

$$n = \sigma_n^2 - \frac{\mu_n a}{N} - \mu_n^2 \tag{5.36}$$

which can be defined as a noise contribution term (not to be confused with the noise subscript n, as in μ_n)

Then

$$\sigma^2 = \frac{aE\{H^2\}}{3N} - \left(\frac{a}{2N}\right)^2 + n \tag{5.37}$$

From equation 5.31, this can be written

$$\sigma^2 = \frac{a[x^2 - 2x + 2]}{3N} - \left(\frac{a}{2N}\right)^2 + n \tag{5.38}$$

Or alternatively

$$\sigma^2 = \frac{a}{N}\left(\frac{1}{3}\left[x^2 - 2x + 2\right] - \frac{a}{4N}\right) + n \tag{5.39}$$

Which simplifies to

$$\sigma^2 = \frac{a}{3N}x^2 - 2\frac{a}{3N}x + \left[\frac{2a}{3N} - \frac{a^2}{4N^2} + n\right] \tag{5.40}$$

Now, defining some more useful variables, let

$$u = \frac{a}{3N} \tag{5.41}$$

$$w = \frac{a}{2N} + \mu_n \tag{5.42}$$

$$v = \frac{2a}{3N} - \frac{a^2}{4N^2} + n \tag{5.43}$$

One may write 5.40 as

$$\sigma^2 = ux^2 - 2ux + v \tag{5.44}$$

And making proper substitutions for *x* from 5.34, this becomes

$$\sigma^2 = u(2\sigma + w)^2 - 2u(2\sigma + w) + v \tag{5.45}$$

Or

$$\sigma^2 = u(4\sigma^2 + 4\sigma w + w^2) - 2u(2\sigma + w) + v \tag{5.46}$$

By simplifying and rearranging, one can arrive at a crucial result

$$0 = (4u - 1)\sigma^2 + 4u(w - 1)\sigma + (uw^2 - 2uw + v) \tag{5.47}$$

Which may be solved easily using the quadratic formula, revealing

$$\sigma = \frac{-4u(w-1) \pm \sqrt{(4u(w-1))^2 - 4(4u-1)(uw^2 - 2uw + v)}}{2(4u-1)} \tag{5.48}$$

Using typical values *a*=8 and *N*=32, and assuming negligible contribution from noise (that is, μ_n=0 and *n*=0), we find that

$$\text{x} = .6776 \tag{5.49}$$

$$H_b = 1.322 \tag{5.50}$$

and so

$$\frac{H_b}{x} \approx 1.95 \tag{5.51}$$

In other words, given maximum possible variation between stripe band peak heights, the threshold of two standard deviations above the mean will function appropriately (that is, not eliminate any signal peaks) when the maximum height is less than roughly 1.95 times the lowest peak height.

This value provides an estimate of effectiveness, though in realistic scenarios, noise contribution is likely to reduce this allowable variation. Additionally, blurring and filtration effects will create peak shapes which are not truly triangular. Nevertheless, it is also true that in realistic scenarios, variation between peak heights may also be reduced significantly, thereby reducing the variance; and thus, reducing the corresponding threshold.

Error rate estimation

Once thresholding has been performed, estimations are utilized in order to assign theoretical 'false positive' and 'miss' error rates to the image. These values will then be used in the definition of the Hidden Markov Model emission matrix.

Defining the false positive error rate is based on the assumption that any noise peaks which have been incorrectly included will have intensity values that are much smaller than the intensity values of rightly included peaks. Under this assumption, and the assumption of a statistically normal process, we will thus consider all pixels with intensity below one standard deviation of the mean (of remaining thresholded pixels) to be introduction of false positives. Therefore, for all patterns, the initial false positive rate is assumed to be roughly 15.9% (that is, the proportion of remaining peaks below one

standard deviation; peaks that are considered to be noise). Due to the characteristics of these assumptions, this initial rate can be assumed to be a very crude estimate.

The miss error rate is defined in a more practical way. Given the stripe image, it is known approximately how many stripe pixels must be present in the image. Prior knowledge of the pattern encoding allows one to define how many 'high' encoded stripe pixels should appear in each channel image. Therefore, an estimate of the miss error rate is defined as one minus the ratio of actual 'high' stripe pixels versus the expected number of 'high' stripe pixels in each channel image.

Chapter 5 Error Correction and Post-Processing

The previous sections have explained the theoretical operations of the Modified Composite Pattern method. While it is, perhaps, well suited for tasks such as optical control or machine vision, the results acquired from the basic method are often considered inadequate from a subjective standpoint of the human viewer. In order to further reduce depth recovery error, and thus improve the visible quality of the output model, it was necessary to implement a number of error correcting and post-processing procedures.

5.1 Background Information

Based on observations specific to MCP, it was found early in the research process that it would be necessary to consider the MCP encoding sequence in a more formal context. Specifically, it was decided that the recovered MCP codeword sequences should be analyzed as the output of a finite state machine, specifically of the Hidden Markov Model subclass.

Finite State Machines

A finite state machine (FSM, also known as a finite automaton, FA, or finite state automaton FSA) is theoretical system model. It consists of a set of system states, a set of known transitions between states, a known starting state, a set of end states, and (where relevant) alphabets of input and output symbols. The physical manifestations of these abstractions depend on the particular system being modeled. For example, a computer program may take typed data characters as inputs. Recognition of the input characters

may cause transition between states represented by changes in some numerical flag value, and cause the output of a pre-determined set of graphical color changes. Physical systems can be modeled this way as well. A FSA representing a door may take "open the door" and "close the door" actions as inputs, transitioning between an opened state and a closed state.

From the initial starting state, an input from the input symbol alphabet will cause a transition to one of any number of next states, from which the process of taking in inputs and transitioning between states may be continued until an end state is reached. FSAs may be classified as deterministic (in which a single input symbol is related to only a single possible transition from any given state) or non-deterministic (in which a single input symbol may be related to any number of possible transitions from a given state), as well as acceptors (also known as recognizers or sequence detectors, in which only end states generate an associated output) or transducers (in which all states, or even state transitions, may generate an associated output).

Hidden Markov Models

The Hidden Markov model (HMM) is a class of statistical signal models frequently utilized in signal processing applications. A Hidden Markov model is constructed under the assumption that an observed symbol sequence is generated, in a statistically well-defined manner, by an underlying hidden process. The underlying process is modeled by a Markov chain, effectively characterized as a transducer-type finite state automaton with a known number of states *N*. A Markov chain is a discrete stochastic process defined by the Markov assumption, that is, the probability of transition

from some state A to some other state B depends only on state A, and is independent of the path by which state A was reached. Thus, for each state pair, there is a single transition probability value. These probabilities define the HMM "transition probability distribution." Each state (or state transition as the case may be) is assumed to have a certain probability of generating (that is, "emitting") one of *M* observation symbols. These probabilities define the HMM "emission probability distribution". Additionally, there is a certain probability that any individual sequence will begin with any given state, that is, each state has an associated probability that it is the initial state in any sequence. These probabilities define the HMM "initial state probability distribution". Together, the transition, emission, and initial state probability distributions (usually considered in the form of matrices), along with associated values *N* and *M*, define in full the HMM proper [82].

Hidden Markov models were first introduced in their purely mathematical form in the mid twentieth century. Early utilizations for speech recognition followed soon after [82]. HMM implementations lend themselves well to sequential symbol recognition problems such as speech and text recognition, and have been used in varied fields such as musicology [82, 83], bioinformatics [84], and image processing.

A particularly attractive characteristic of HMMs is their utility as a mechanism of machine learning. It is possible to actively create or modify a given HMM to reflect training data introduced to the system. This has been taken advantage of in the development of speech recognition systems [85], pattern classification and computer process modeling [86].

There are three problems of particular interest when utilizing HMMs. The first is posed: given a particular sequence of observations, what is the probability of the sequence being produced by the system? The second is similar: given a particular sequence of observations, what is the particular sequence of system states that is most likely to have produced the sequence? The third can be summarized: given one or more sequences of observations and a hypothetical hidden system FSA, what are the most likely associated transition and emission probability distributions? Note that the solution to this third problem is particularly important in the utilization of HMMs for machine learning systems.

5.2 MCP Hidden Markov Model

Once each channel image is thresholded, multiplied by the appropriate constant, and combined (the stripe image acting as a binary mask), the resulting encoded image (the "Code Image" in the flowchart in Figure 4.1) must then be processed using a Hidden Markov Model decoder.

Consider a single column in the encoded image. Ideally, all pixels encountered should correspond to the correct encoding values in the proper order. However, in practice, a column will frequently contain errors such as incorrect values, missing code values, and additional extraneous pixels, all due to data loss from thresholding or demodulation, or due to undesirable noise peaks.

However, due to prior knowledge of the encoding scheme, the proper order of encoded values is well defined. Additionally, as noted in the previous section, it is possible to estimate the likelihood of error for any given code value. These

characteristics suggest that a HMM modeling scheme may be very useful. In this case, the HMM will not be used due specifically to its learning ability, but rather due to the applicability of the Veterbi algorithm in identifying the most likely underlying code structure given a sequence of (potentially corrupted) observations.

A proper Hidden Markov Model requires an emission matrix, state transition matrix, and starting state matrix, in addition to the set of observations upon which it will operate. In this situation, processing will be performed in a column-wise fashion. Observations will consist of the set of nonzero pixel values in an encoded image column.

Emission Matrix

The system emission matrix defines the probability of observation of a given value for each underlying true encoded value (in terms of a specific example, the emission matrix value at coordinate (8,5) is the probability that signal value 5 will be observed, even though the actual transmitted signal value is 8). These probabilities can be defined very simply from the probability approximations discussed in previous sections.

Assuming statistical independence between each modulation channel (interference analysis can allow an estimate of just how accurate this assumption actually is) the probability of interpretation of each codeword can be calculated as the product of probabilities of interpretation for each codeword symbol. If Y is the set of received symbols for each channel, and X is the set of symbols sent, then the probability of each codeword is defined as

$$P(Y_1 Y_2 \cdots Y_N) = P(Y_1 \mid X_1) P(Y_2 \mid X_2) \cdots P(Y_N \mid X_N) = \prod_{n=1}^{N} P(Y_n \mid X_n) \tag{6.5}$$

For an example, let Pf_x represent the false positive probability for channel x. Similarly, Pm_x represents the missed symbol probability for channel x. Consider that the N=4 codeword sequence {1,1,0,1} has been encoded on a given stripe. The probability of receiving, on the corresponding stripe in the received image, the codeword sequence {0,0,1,1} is given by

$$P(0_1 0_2 1_3 1_3) = P(0_1 \mid 1_1)P(0_2 \mid 1_2)P(1_3 \mid 0_3)P(1_4 \mid 1_4) \tag{6.6}$$

that is

$$P(0_1 0_2 1_3 1_3) = P_{1} \cdot Pm_2 \cdot Pm_3 \cdot (1 - Pf_4) \tag{6.7}$$

State Transition Matrix

According to the definition of a Markov chain, the probability of occurrence of symbol X_{n+1}, given the occurrence of the previous symbol X_n, is a unique value. That is

$$P(X_{n+1} \mid X_n) = p_{n,n+1} \tag{6.8}$$

These values are collected as elements of the Hidden Markov Model state transition matrix. Practically, it means this: given that encoded pattern value *Xi* has been found, the probability that the next encoded pattern value is *Xj* is defined by the transition matrix element $T(i,j)$.

The state transition matrix for MCP is defined by the encoding pattern used. Ideally, given a sequence of encoded symbols, transitions will occur only from one symbol to the immediate next symbol in the sequence, as in Figure 5.1.

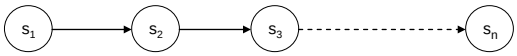

Figure 5.1, Basic symbol sequence

However, in practice, it is possible to encounter extraneous stripe pixels, interrupting the ideal sequence with an error value. In order to compensate for this, one may consider the system to be defined by the following representation, Figure 5.2.

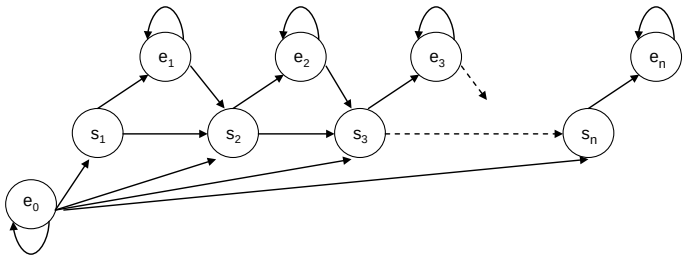

Figure 5.2, Symbol sequence with error states

The symbols e_0 through e_n represent potential error states, that is, extraneous pixels. Note that it is allowed that any number of error pixels may be encountered between any two proper symbols.

This interpretation of the system is more representative to the actual pattern likely to be encountered, and can therefore be expected to give superior results to the first

model. However, it doesn't take into account a second source of error, the possibility of a missed symbol. Therefore, an even more accurate representation is shown in Figure 5.3.

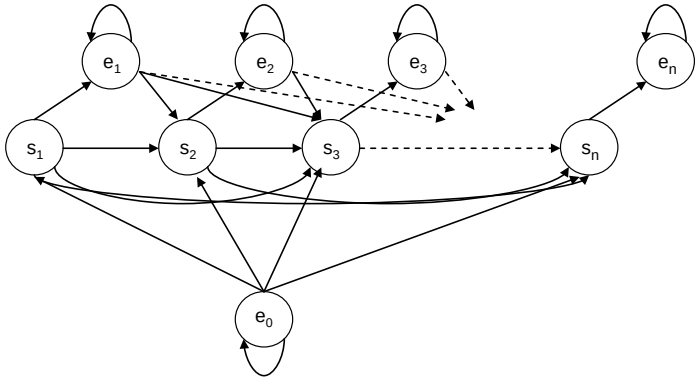

Figure 5.3, Symbol sequence with errors and misses

As you can see, this is much more complicated than the initial ideal system. Note that it is possible for each symbol state to transition to any other later symbol state or to an error state, but never to a previous state. As such, the system can be classified as a left-right HMM.

The transition probabilities in the system are not necessarily easy to define. The sequence of symbols encountered in any column may well be different from the sequence encountered in any other column due to subject and image capture characteristics. In order to find an approximation for each necessary transition, one can make use of a few helpful assumptions.

First, it may be assumed that it is quite unlikely to encounter an error pixel at all, due to the stripe isolation process characteristics. Therefore, any transition to an error state must be very small.

Second, one may assume that the likelihood of a single missed symbol is significantly greater than the likelihood of multiple missed symbols, and that the most likely outcome is that a symbol is not missed at all. That is, transition probability is related to symbol proximity. Therefore, one can reduce the transition probabilities between symbols using a scheduling method, defining transition probabilities according to

$$p_{i\ j} = \frac{(1-p_e)}{C} \cdot \frac{1}{(j-i)^k} \qquad j > i \tag{6.9}$$

Where k is a constant, C is a normalization constant, and p_e is the approximate transition probability to an error state.

Start Matrix

The start matrix, the matrix of probabilities that the sequence begins at any particular symbol, is actually quite easy to define. Due to the fact that no assumptions can actually be made regarding which symbol will be the first encountered in any given column, all non-error symbols (and only e_0) can be considered to have equal probability.

5.3 The Viterbi Best Path Algorithm

The Viterbi algorithm was developed in the late 1960s by Andrew Viterbi for use in decoding error-correcting convolution codes [87]. It is a "path finding algorithm". In the context of HMMs, it is used to determine the most likely sequence of system transitions responsible for an observed sequence of emitted observable signals. This most likely path (also known as the Viterbi path) is given as

$$\text{argmax}(P(O|X)) \tag{6.1}$$

where O is a sequence of observation signals and X is a sequence of underlying hidden states. Due to the Markov assumption, the probability of any observation sequence is given by

$$P(O|X)=[P(o_0|x_0)P(x_0)][P(o_1|x_1)P(x_1|x_0)][P(o_2|x_2)P(x_2|x_1)]... \tag{6.2}$$

Given that the set of X and O probabilities are known as given parameters of the HMM, it is possible to find the Viterbi path simply by calculating the probability of a set of observations under all possible permutations of hidden states and summarily finding which state sequence produces the greatest observation probability. However, the Viterbi algorithm allows one to avoid the computational complexity required by this method, instead utilizing a much more efficient recursive process.

Let $X(t)=\{x_0,x_1...x_t\}$ define a path of t transitions. Let $S(t)=\{x_0,x_1...x_n\}$ define the set of hidden states which may be reached in t transitions. Given t transitions, for the observation o_t, for each state x in S(t), there exists one best path X(t-1) which maximizes

$$P(x|t)=P(X(t))=P(o_t|x)P(x|x_{t-1})P(X(t-1)) \tag{6.3}$$

Paths calculated thus are called “partial best paths”. The definition of X(0) makes use of starting probabilities according to

$$P(X(0))=P(o_0|x_0)P(x_0) \tag{6.4}$$

Notice that the calculation of P(X(t)) makes use of the Viterbi path probability to intermediate state x_{t-1}, that is P(X(t-1)). Thus, the calculation of the overall Viterbi path may be performed recursively beginning with t=0.

For each t, for each state in S(t), one must record both the partial best path probability P(X(t)), as well as a pointer value to the state x_{t-1}. This pointer is recorded so

that when the final Viterbi path probability is calculated, the actual sequence of states may be easily traced from the final state to the initial state while eliminating the need to store an entire X(t-1) state sequence for each intermediate hidden state. [88]

It was decided that MCP be analyzed using the Viterbi algorithm due to the nature of any MCP scan: the scan contains a subset of a known sequence of symbols (the stripe code-words). This sequence is constant and well-defined. However, due to decoding errors and the nature of the photographed subject, it is possible that many of these symbols will be incorrectly identified or absent in the decoded image. The goal of MCP processing is to recover, as well as possible, the information represented in the true sequence regardless of these errors. And so, modeling the MCP scan recovery process in terms of HMM, and using Viterbi optimal path decoding, was deemed to be the ideal method of analysis.

5.4 Post-Processing Operations

Despite the optimality of the Viterbi decoding process, intractable errors are likely to persist in MCP images due, if nothing else, to the unpredictable nature of visual data. As such, it has been necessary to implement a series of corrective functions, most commonly to be used as post processing. While effective, these functions are, at present, largely non-automated, sometimes leaning more heavily than other functions on user defined parameters.

Modal Filtration

Modal filtration is, as the name implies, a nonlinear filtration process which simply replaces all elements within a windowed region of the code image with the modal value of that region (implemented to ignore background values of zero). Based on the assumption that, in a small region containing multiple code-word values, the correct value will occur more frequently than error values, it has the helpful effect of correcting very small regions of incorrect coding that occur in an otherwise correctly decoded region. Section 5.5 further shows how this apparently small correction can have a measurably significant effect.

Morphological Blob Elimination

In the morphological blob elimination stage, the image is effectively divided into an array of binary component images. Each contains a binary representation of only a single code-word value from the image (that is, the third image in the array would contain a positive value only where the pixel has a corresponding value of 3 in the code image). Each component image is then processed with a morphological open operation. This has the effect of eliminating isolated regions smaller than the user-defined (and usually, fairly large) size of the structuring element. Finally, all binary result images are once more combined into a new code image via simple multiplication by an appropriate base-10 code-word value followed by addition.

It must be noted that, due to the nature of the morphological open operation, it is possible to lose sizable regions of data, and the quality of the results may be influenced by the shape of the target subject or the incidence of intensity thresholding (which has often been employed as a simple method of error reduction in low-intensity areas, as

noted in section 4.1). Nevertheless, when employed with a suitably sized structuring element, such as one derived from empirical data of the sort available in section 2.3, it is able to reliably eliminate decoding error due to step-edges (which notably occur frequently at the borders of MCP scan images). Also note that, unlike modal filtration, blob elimination does not correct data, but merely eliminates it. Thus, in order to recover the maximum amount of data available, it is necessary to implement an additional function to fill in empty stripe region areas.

Region-based Growth

Upon completion of blob elimination, it is often the case that some pieces of the MCP image are left with no decoded value at all. Yet an ideal depth map would make use of as much of the given data as possible. As such, it is desirable to fill in these empty regions with the appropriate code-word stripe information.

A process of expansion of existing coded stripe regions was present in earlier implementations of MCP. Unfortunately, this growth process relied heavily on the correctness of user input for proper operation. However, with the implementation of stripe localization via adaptive thresholding, it was found that these binary stripe regions provided an ideal path for the expansion of stripes.

The method operates as a two-pass process. Given three element matrices: the stripe image, which contains all pixels which should contain any codeword value; the code image, which, after blob elimination and modal processing, is expected to contain only the most accurately identified encoded pixels; and the binary region image, containing the most basic, unfiltered identification of stripe regions, the method proceeds

as follows: first, the stripe image is dilated by a single-column-wide structuring element (assuming stripes run horizontally, as in all previous examples). This dilated stripe image is subtracted from the region image, and the result is then combined with the original stripe image to create an enhanced growth mask image. This procedure effectively allows the region image to "fill in" gaps in the stripe image (occurring due to thresholding, for example) in a way that still preserves the consistency of the initial stripe results (that is, code values will only be extended into pixels that were, at the most basic level, still part of the same stripe).

Next, a copy of the code image is created as an initial starting point for the output image (identified as the "blob image" as it contains all remaining blobs of encoded pixels). Each pixel in the input image is then processed. If pixels adjacent to the corresponding point are found to be empty in the blob matrix, but not empty in the growth mask, a growth loop initiates, propagating the value (in the blob matrix) along the path provided in the growth mask. This continues until all of the pixels in the original image have been processed.

Now, consider that some regions may be such that two different code identity stripes could be extended into it. The result at regions such as this is entirely dependent on the order in which the input pixels were processed. Thus, in order to prevent misidentification in these potentially problematic regions, the entire process is repeated with a second blob matrix, this time searching through the input pixels in the reverse order. The two resulting blob matrices are then compared, and only matching pixels are passed as the result. Finally, this result is multiplied by the true initial stripe mask to remove any extraneous pixels. The result of the growth process is most frequently the

final result in the decoding phase, and may be processed to acquire the MCP depth image.

5.5 Results

The following results were collected for an image of a flat, white poster presentation board at varying stages of processing.

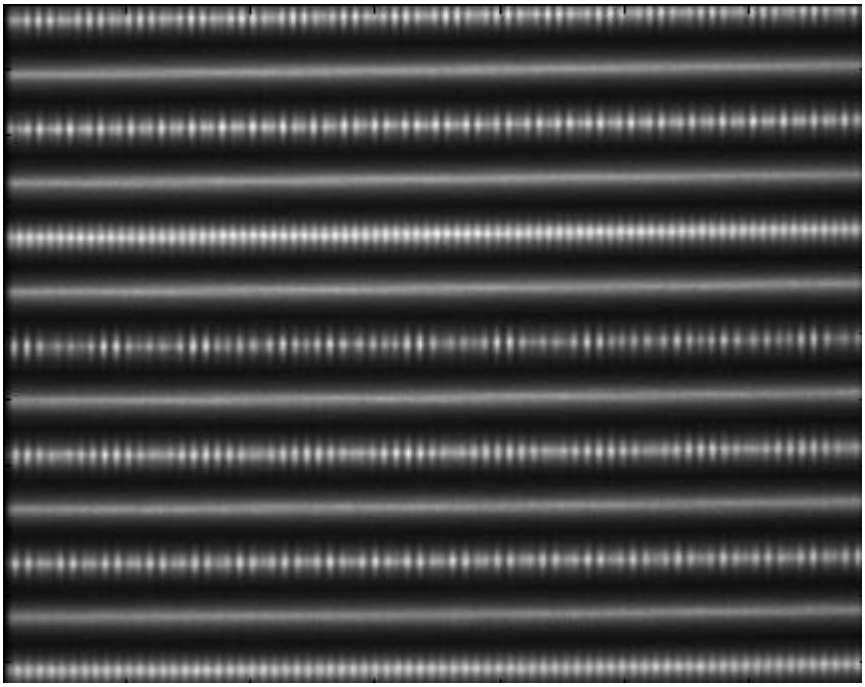

Figure 5.4, Input image

Figure 5.4 shows the initial unprocessed input image.

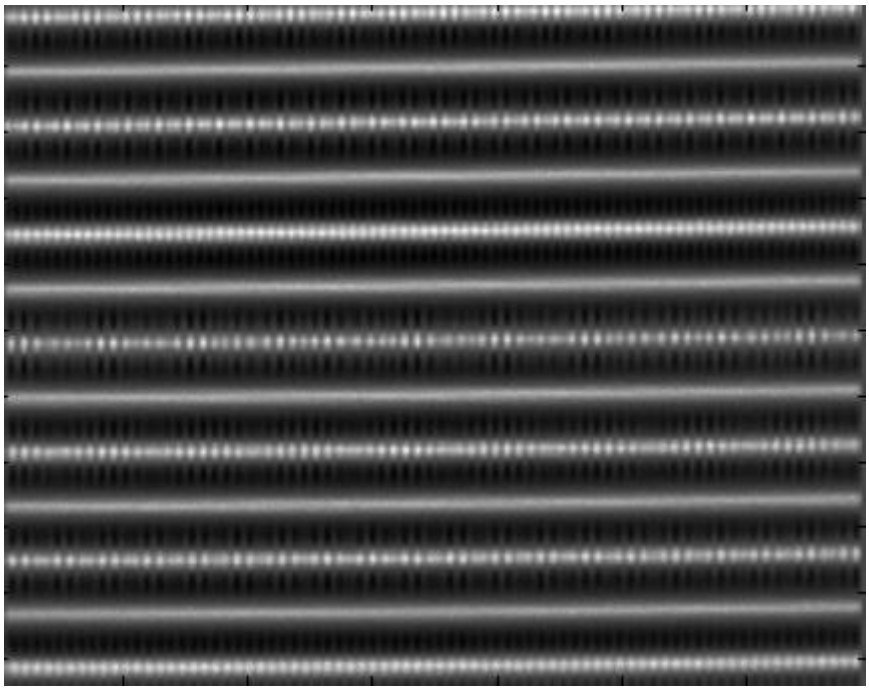

Figure 5.5, Contrast enhanced input image

Figure 5.5 shows the input image, processed with the contrast enhancing filter (discussed in chapter 4) in preparation for thresholding and peak identification.

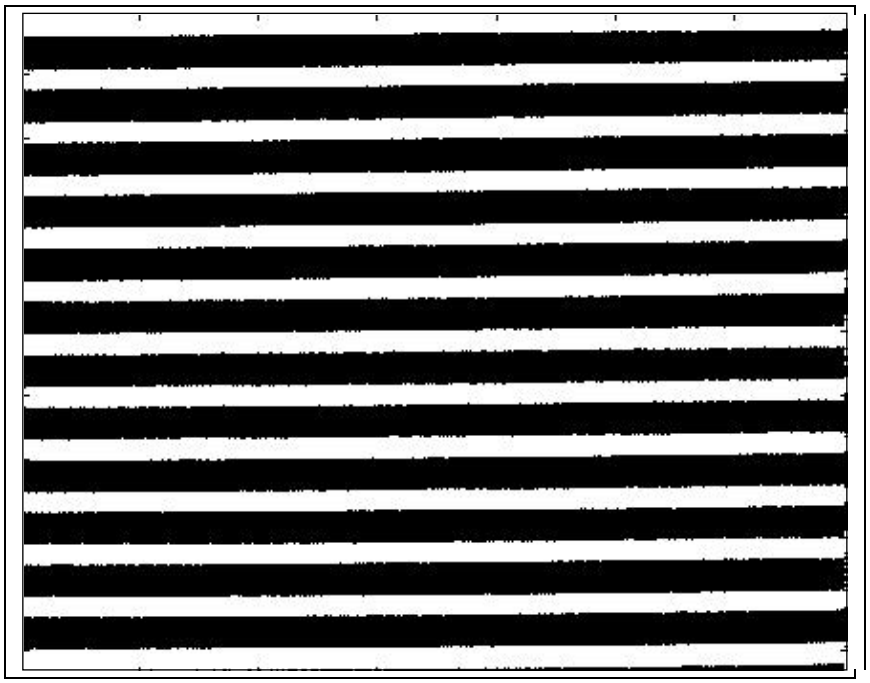

Figure 5.6, Stripe region isolated image

Figure 5.6 shows the results of local adaptive thresholding, which isolates the regions containing stripe peaks.

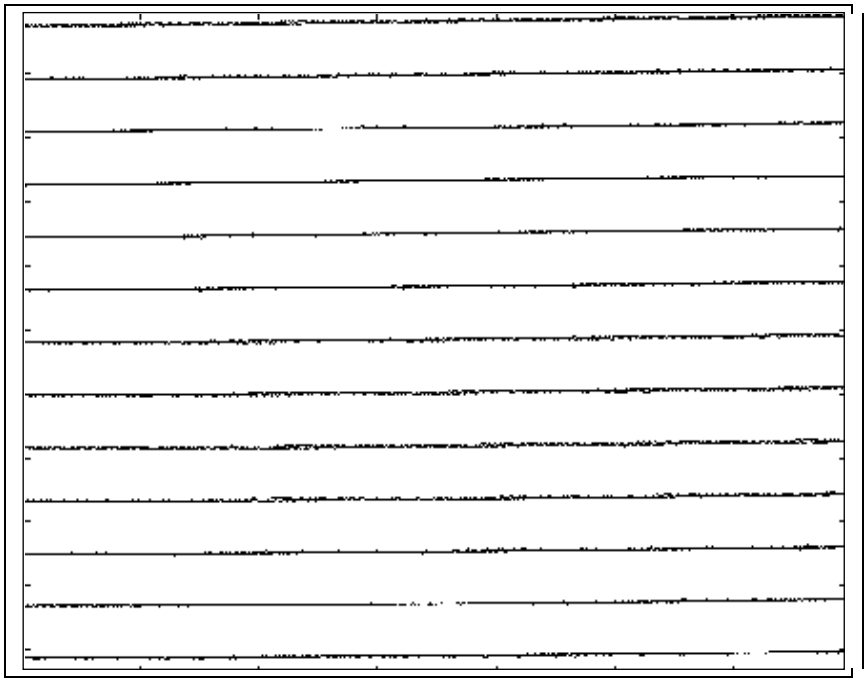

Figure 5.7, Isolated stripes

Figure 5.7 shows the pattern of isolated stripe peaks identified in the image.

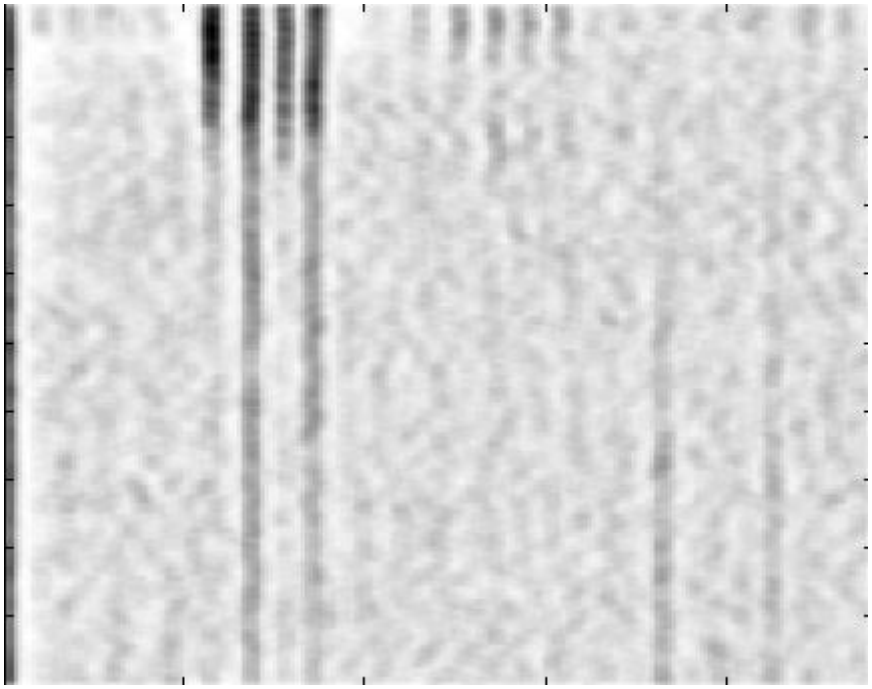

Figure 5.8, Spectral region

Figure 5.8 shows the prepared spectral representation of the input image. This sub-region (collected as described in chapter 4) is used in the collection of the individual modulation frequencies. Note the obvious presence of strong modulated signal bands.

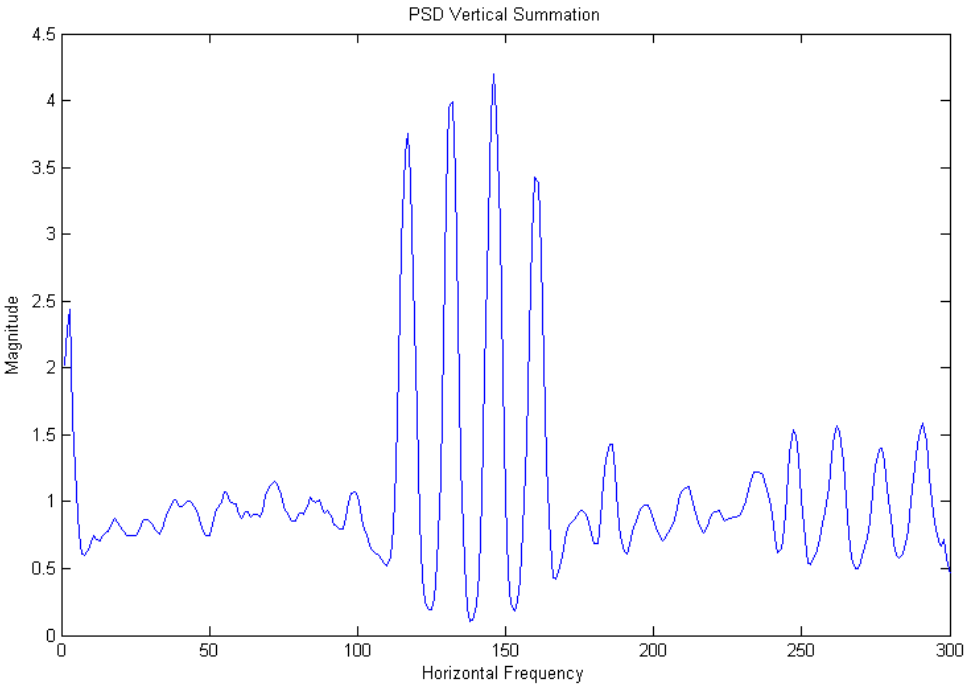

Figure 5.9, Plot of frequency bands (magnitude vs. frequency in cycles/fov) as used in ABS automated frequency selection

Figure 5.9 shows the data array to be used in the analysis-by-synthesis frequency detection algorithm. This plot is of particular importance, as it clearly displays some of the most intrusive problems encountered in Composite Pattern processing.

Notice that, in addition to the strong modulated pattern bands (located between 50 and 100 cycles per field of view) there is also some remaining baseband energy, located near DC. This energy is particularly troublesome due to its large magnitude, and must be accounted for, either in the frequency detection process, or by the use of a high-pass filter.

One can also see images of the modulation bands, located between approximately 120 and 160 cycles per field of view. In this case, these image bands lack sufficient energy to cause any disruption in the frequency detection process. However, this may not

always be true. Thus, one can see the importance of maximizing modulation energy and reducing interference.

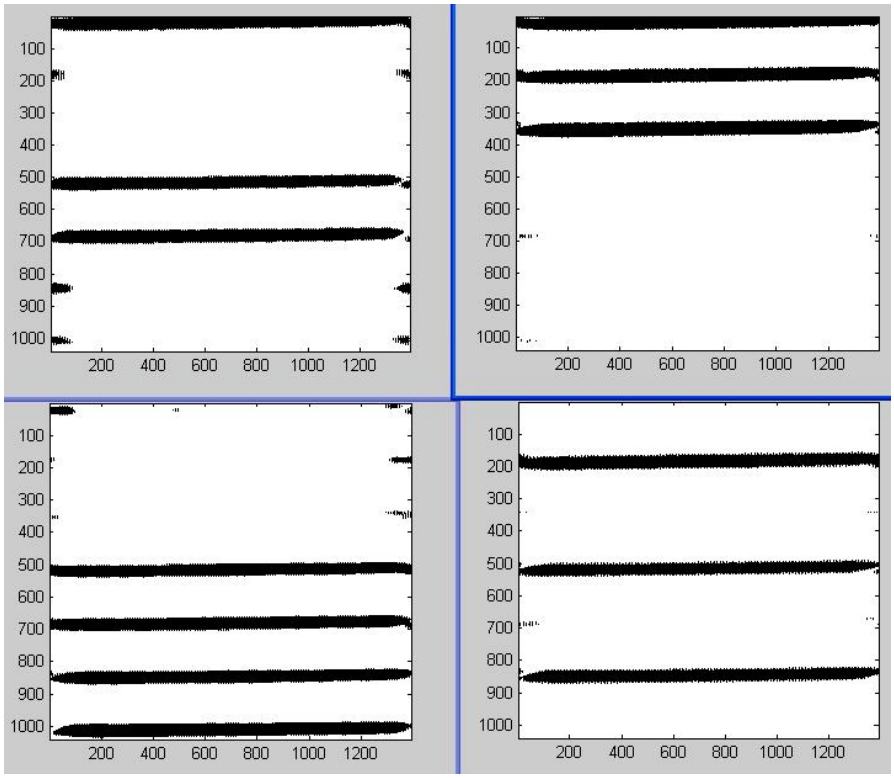

Figure 5.10, Individual isolated binary region information images
(axes simply represent pixel coordinates)

Figure 5.10 shows the set of structured light pattern information collected in the demodulation process, thresholded to binary values (where black indicates a high value). Take note of the distortion and incorrect information present around the image edges. As discussed in chapter 2, step edges may lead to the introduction of artifacts in demodulated bands, and this fact must be accounted for in processing.

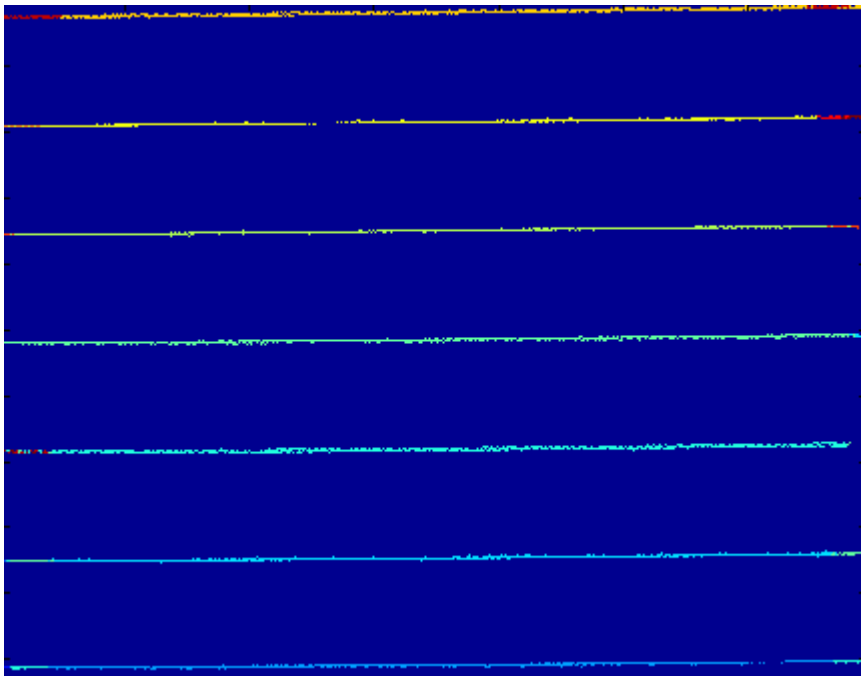

Figure 5.11, Initial recombined region/stripe information ("Code image")

Figure 5.11 shows the results of combining the four binary regions into base 10 code-words, and combining the result with the binary stripe peak mask (each color indicates a different code value). This image was mentioned in chapter 4, specifically as the “code image” mentioned in Figure 4.1. Additionally, the image was further processed with a basic modal filter (as discussed in section 5.4). It was found that modal filtering could often improve results of further processing steps.

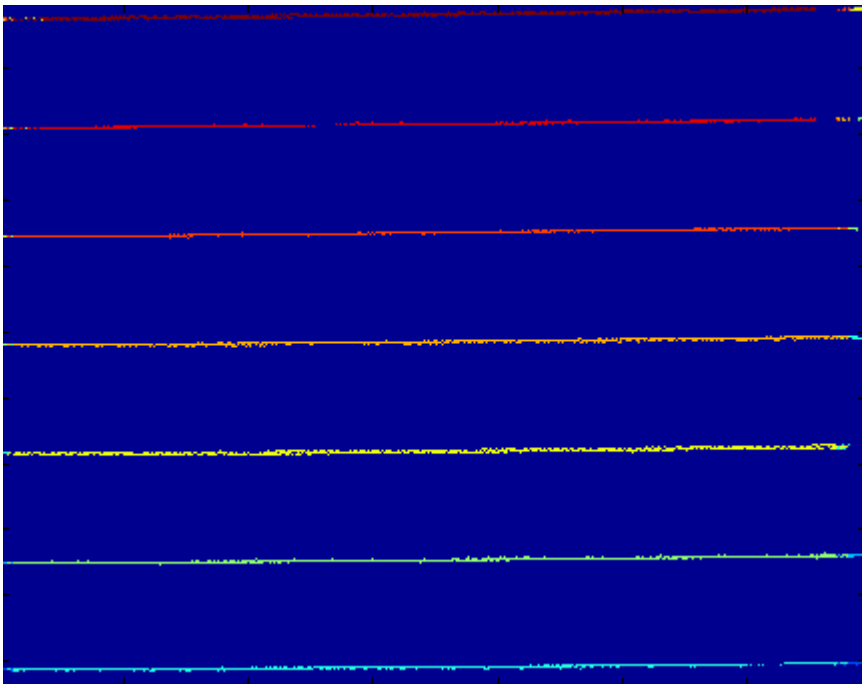

Figure 5.12, Code image after initial HMM processing

Figure 5.12 shows the code image after HMM processing. Although the improvements may not be obvious due to the already high quality of the image, they can be seen in certain areas. Note that, due to the extreme nature of the interference near the edges of the images, it was often impossible to correct the encoding errors.

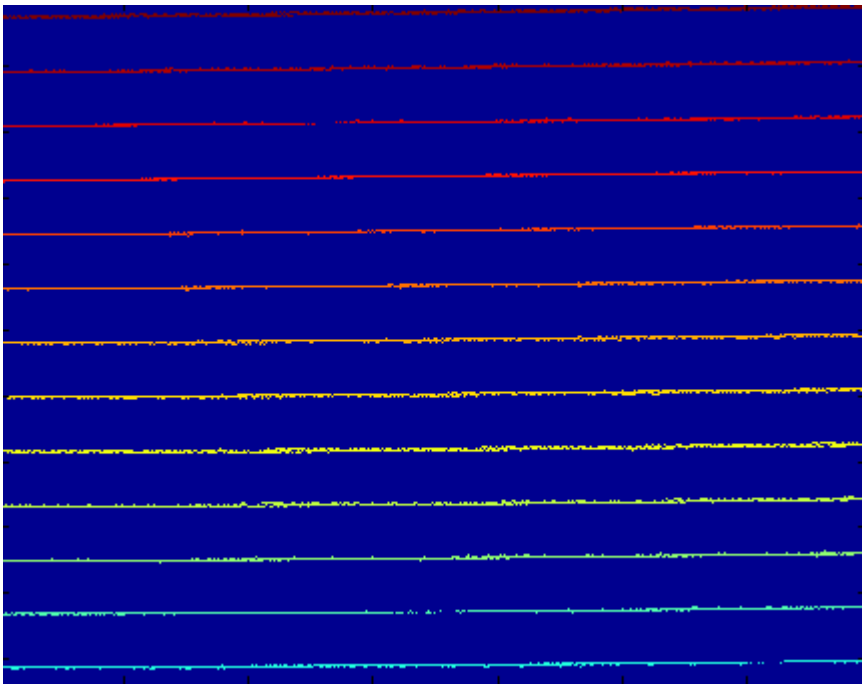

Figure 5.13, Code information recovered for modulated and unmodulated stripe regions

Figure 5.13 shows the result after additional post processing (modal filtration, blob elimination, and region-based growth process) and combination with the identified unmodulated stripe information. Notice that stripe code-word identification is almost entirely free of errors after post-processing is complete.

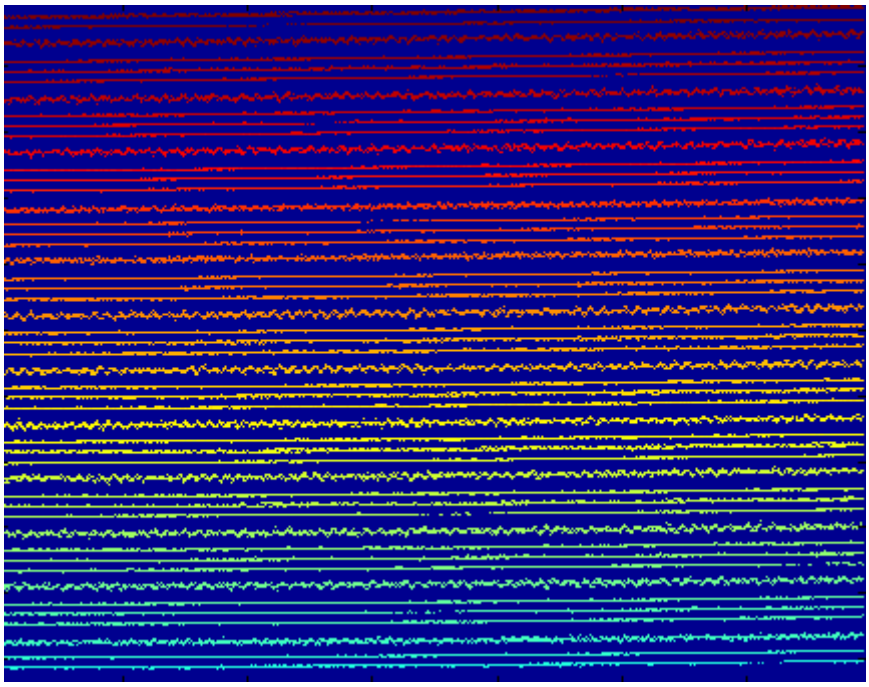

Figure 5.14, Fully recovered code information, including presence of 2x and 4x stripes

Figure 5.14 shows the final MCP decoding result, combined with interpolated 2x and 4x stripe regions (as discussed in chapter 2). As expected, the density of the encoded data available is greatly increased.

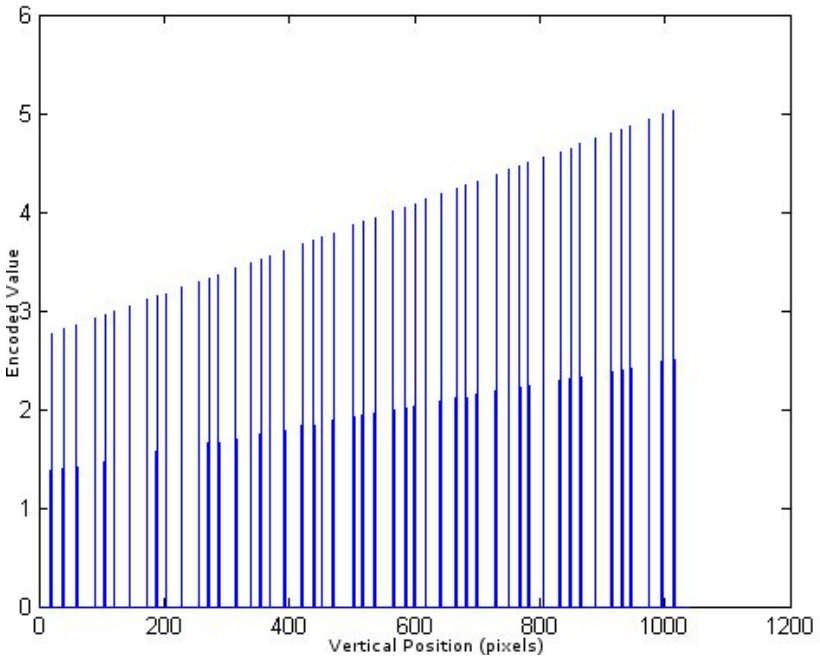

Figure 5.15, Single column plot of phase information

Figure 5.15 displays a single column of the code image, transformed (via appropriate scale factors, as discussed in chapter 3) into a “phase space” encoding representation necessary for 3D data recovery. Note the marked linearity, as would be appropriate for a flat surface.

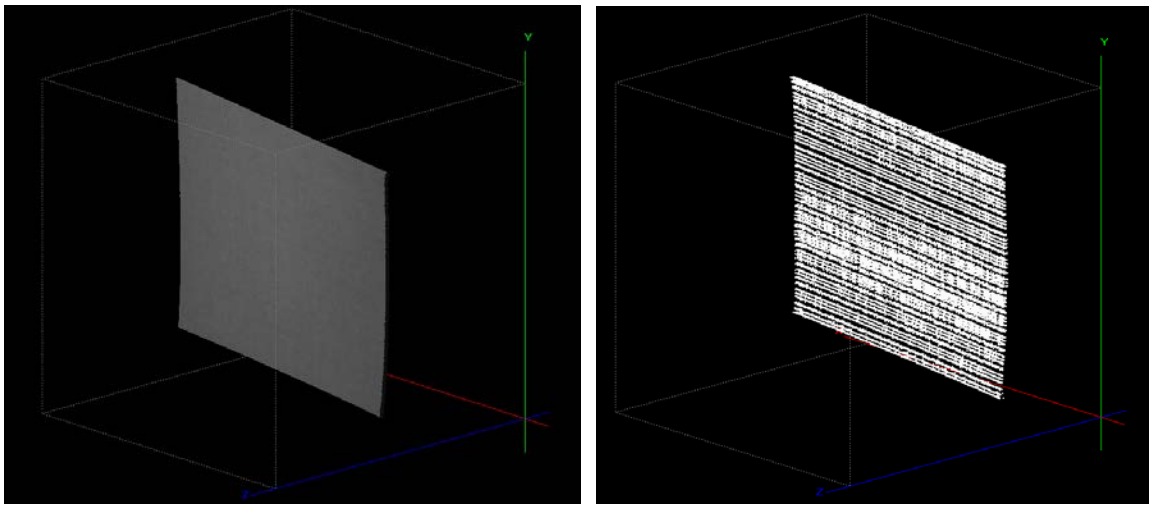

Figure 5.16, (left) PMP ground truth 3D, (right) recovered MCP 3D

In Figure 5.16, one can see the actual 3D models acquired from the scan, the left being the PMP “ground truth” scan of the board, and the right being the 3D data recovered from MCP processing.

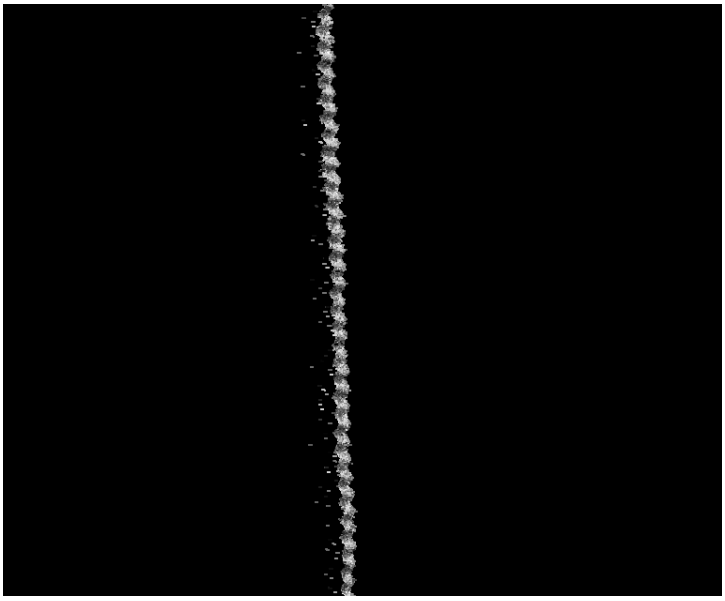

Figure 5.17, Edge view of PMP board

Figure 5.17 shows the edge view of a section of the PMP board scan. The apparent undulation is likely due to a combination of depth recovery error and actual surface roughness (the board, being made of corrugated cardboard, was not, itself, perfectly flat).

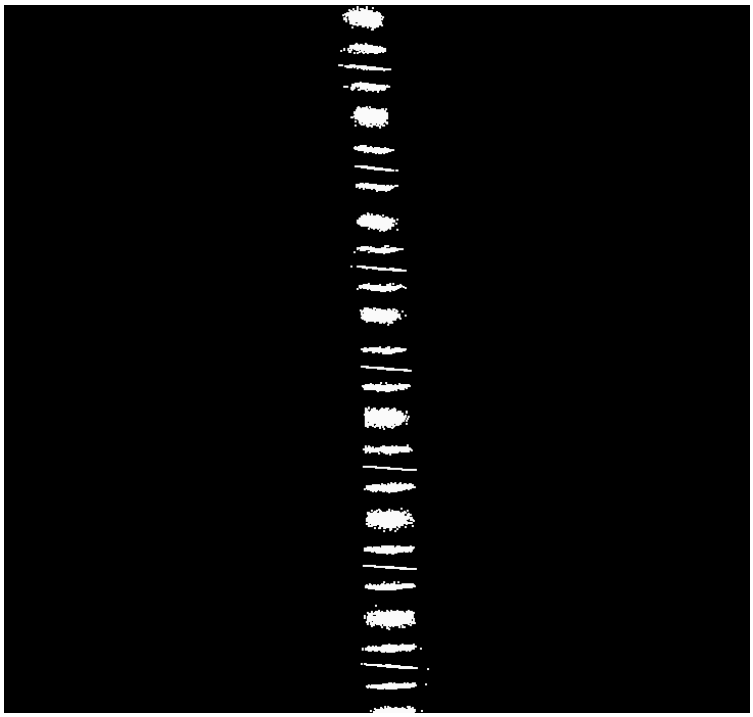

Figure 5.18, Edge view of MCP 3D

Figure 5.18 shows the edge view of the MCP scan 3D. Notice that, in contrast to the PMP scan data, there is a marked "spread" in the model's Z direction. This effect is due to the pixel-accurate nature of the current generation MCP processing. That is, there is a very slight systematic error in localization of stripe peak position, and thus, a small amount of systematic deviation of each stripe from the ideal depth. The imperfection may be corrected using sub-pixel localization methods, or more simply, via basic interpolation and smoothing which would likely occur in a realistic use of the method.

5.6 Effectiveness of Processing

Some results have been shown, but nevertheless, the question remains: just how effective are the proposed processing techniques in a quantitative sense? To this end, an experiment was performed.

Two objects were tested: a flat board (the same discussed in the previous section. Scan results used for the following tests are seen in Figure 5.16) and a mounted, white sphere. For each, a PMP scan was performed, followed immediately by a MCP scan.

MCP pattern modulation frequencies were isolated, in the captured image, to 78, 92, 104, 115 cycles / fov.

MCP decoding

The first test concerned only the effectiveness of decoding procedures for Modified Composite Pattern. For each object, the final MCP decoding result was compared to results at various stages of processing. Under the assumption that the final result is completely correct (which, as shown by results depicted in section 5.5, is not an unreasonable assumption), this allows one to identify the relative effectiveness of each processing step in defining a correctly decoded result.

The results collected for the flat board are shown in Table 5.1.

Table 5.1, Testing MCP decoding process - Flat board

Code #	Original	Modal	HMM	HMM / Modal
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0.9143	0.9158	0.9259	0.9259
5	0.9143	0.9165	0.9259	0.9245
6	0.9702	0.9724	0.9266	0.923
7	0.9964	1	0.9266	0.923
8	0.9688	0.9739	0.9253	0.9216
9	0.9448	0.9441	0.9253	0.9253
10	0.9005	0.8998	0.9259	0.9259
11	0.931	0.9303	0.9259	0.9259
12	0	0	0	0
13	0	0	0	0
14	0	0	0	0
15	0	0	0	0

The column “Code #” represents a base 10 code-word present in the decoded image. Note that, due to the capture process, not all available code-words are present. The other values shown represent what fraction (in comparison to the final result) of pixels is correctly identified at each stage.

The column “Original” denotes the initial decoded result, before any further processing is applied. The column “Modal” denotes results acquired by processing the original decoded result with a simple modal filter. The “HMM” column represents the results of processing the modal image with the Hidden Markov Model best path algorithm. The last column, “HMM / Modal” represents results of using a modal filter on the HMM processing result image. Each result value is given with the implication that, after final post processing, all codeword pixels have been identified with 100% accuracy (that is, one can consider each table to have an unstated column of “Final Results” in which each value equals exactly 1).

As you can see from Table 5.1, results are mixed. The initial decoding result is, in fact, very accurate. In some cases, processing seemed to cause an overall reduction in correctness.

Table 5.2, Testing MCP decoding process - Sphere

Code #	Original	Modal	HMM	HMM / Modal
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0.8928	0.9002	0.9177	0.9177
6	0.8499	0.8594	0.8602	0.8728
7	0.8847	0.895	0.8612	0.8855
8	0.8112	0.8179	0.856	0.859
9	0.7285	0.7506	0.8709	0.8892
10	0.7448	0.762	0.8533	0.8742
11	0.7117	0.7117	0.8906	0.9002
12	0.3842	0.3863	0.8016	0.8185
13	0.9474	0.995	0.9073	0.9198
14	0	0	0	0
15	0	0	0	0

Table 5.2 shows results from a more complicated target; a mounted sphere (shown in Figure 5.19). Here, the data more clearly suggests the effectiveness of the applied techniques. There is, in all but one case, an increase in correct identification. In some cases, the increase is highly significant (code words 9, 10, and 11 all show greater than 10% increase in coding correctness).

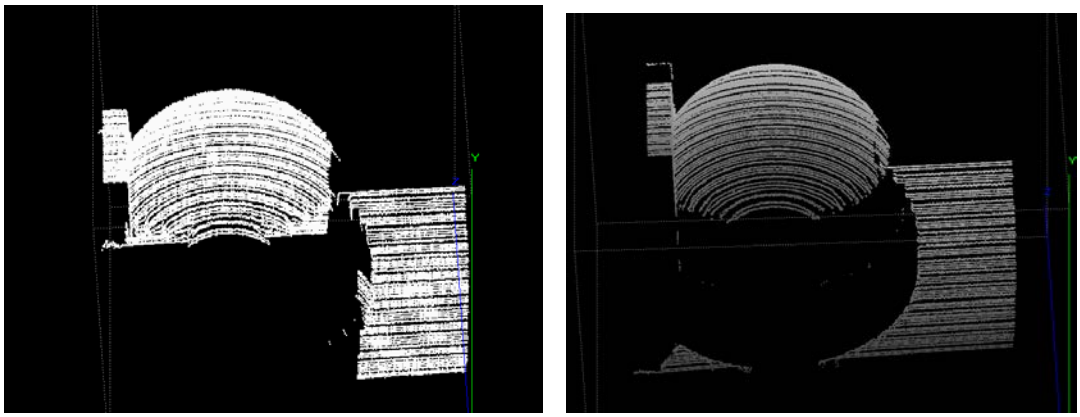

Figure 5.19, (left) MCP result for a mounted white sphere, (right) masked PMP results for the same sphere

Take note that data reflected for code-word 12 is, in fact, incorrect. In the constructed "ground truth" processed image, growth errors created a large and observable incorrectly coded region of code-word 12 (the effects of which can be seen at the edge of the sphere in Figure 5.19(left)). The initial estimate is, in this case, a more accurate representation of the true position. This all goes to show the significant difficulties that can occur due to subject shape.

While the final results of processing, as shown in Figure 5.19, are very promising, they are still not perfect. It is suggested that future work be focused on further improvement of decoding correctness, whether via systematic improvements to the signal processing component (as detailed in previous chapters), improvements to the Hidden Markov Model implementation, or via some clever application of spatial image processing techniques or post-processing.

Accuracy of Recovered Depth

The second area of concern to be considered is that of measurement accuracy itself. That is, assuming that decoding has been performed, how accurate can one expect the results to be? How much additional spatial measurement error should one expect in an MCP scan reconstruction, even when decoding is performed with no incorrect identification whatsoever?

As stated earlier, both the flat board and sphere were scanned twice. For each, a PMP scan was performed, followed immediately by a MCP scan. Each scan was processed to collect the necessary MCP coding information and 3D data according to standard processing procedure, but for one small difference: the PMP phase data for each object was multiplied by a binary stripe mask representing all points present in the MCP image of the object. This allows a 1-to-1 correspondence in points between each type of scan, and thus allows direct comparison between them.

Consider first the sphere. Let us define error terms as being the root-mean-square of the difference between the MCP world coordinate result and the PMP result (for all pixels in which an MCP result is defined). In this case, the results shown in Table 5.3 were collected.

Table 5.3, Error results (Sphere-MCP method)

Subject (MCP)	X error (mm)	Y error (mm)	Z error (mm)
Sphere	1.0341	0.3388	13.0667

As you can see from the results, significant measurement error was found in the Z dimension. This error calculation is due (nearly exclusively, as it will be seen) to misidentification errors and edge effects which may be corrected in various ways.

When the spherical surface itself is isolated as in Figure 5.20 (thereby discounting edge effects and potential stripe misidentification) the actual mean squared error is greatly reduced, as seen in Table 5.4.

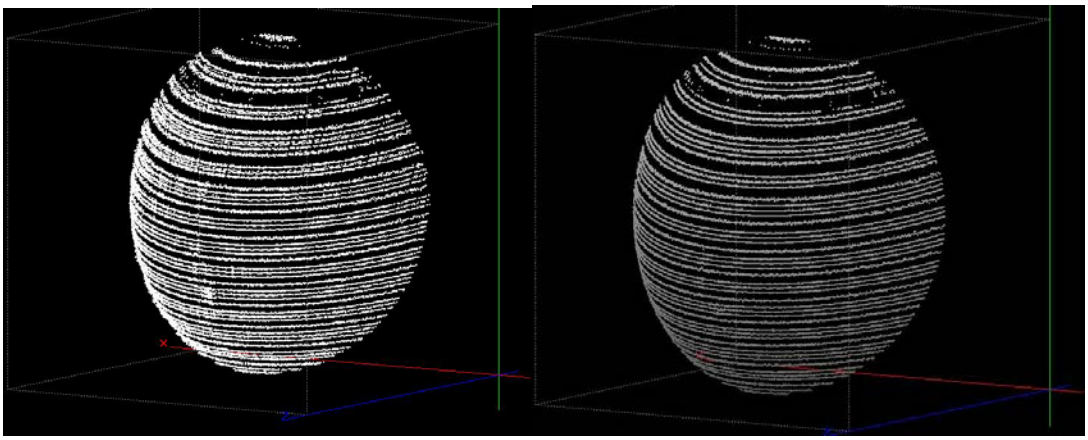

Figure 5.20, (left) isolated surface region of an MCP scan of a mounted white sphere, (right), isolated surface region of a PMP scan of a mounted white sphere

Table 5.4, Error results (Sphere-MCP method-isolated surface only)

Subject (MCP)	X error (mm)	Y error (mm)	Z error (mm)
Sphere (surface only)	0.0849	0.1446	1.7628

Now let us consider the flat board scan, in which there were no readily apparent misidentification errors present (as can be clearly seen in Figure 5.16). Here, the data in Table 5.5 was observed.

Table 5.5, Error results (Board-MCP method)

Subject (MCP)	X error (mm)	Y error (mm)	Z error (mm)
Board	0.0448	0.0401	0.5647

These results are indicative of the measurement errors encountered in an ideal MCP scan. This particular camera/projector arrangement utilized a calibration volume with a depth of 84 millimeters. Thus, the depth error introduced was roughly .67% of the maximum calibrated realizable depth measurement.

Comparison To First Generation Composite Pattern

So how does this new MCP method compare to the standard CP technique? To answer this question, the two techniques were directly compared in the following experiment.

A four pattern, unit-frequency PMP sequence was modulated according to the first generation composite pattern methodology (using frequencies 50, 100, 150, and 200 cycles per field of view, as in reference [32]). Similarly, a third generation MCP pattern was generated for a four bit binary sequence (using modulating carrier frequencies 136, 182, 232, and 253 cycles per field of view). Both images had dimensions of 1040 (vertical pixels) by 1392 (horizontal pixels) and had a gamma correction of 2.3 applied before projection.

The projector used was an Infocus LP120 DLP with resolution 864x1152 and a field of view of 47.3 cm by 63.8 cm. The camera was a Pulnix TM-1402CL with resolution 1040x1392 with a field of view 41.3 cm by 55.5 cm. The camera and projector were offset (vertically) by 23.9cm, and positioned 134.6 cm from the target (that is, the flat board used in MCP calibration, as described in Chapter 3).

The first generation pattern was projected upon a flat, matte white board target. Following this, each modulated sinusoidal pattern was projected sequentially upon the same target according to the traditional PMP phase shifting method, thereby creating an effective "ground truth" for a four pattern unit-frequency PMP scan. Mimicking as closely as possible the methodology implemented in reference [32], the captured image under generation one CP illumination was analyzed to generate a world coordinate matrix. These world coordinates were then compared directly to the world coordinate

results from traditional, non-modulated PMP pattern images. The individual sinusoidal pattern images and resulting phase maps from each can be seen in Figure 5.21.

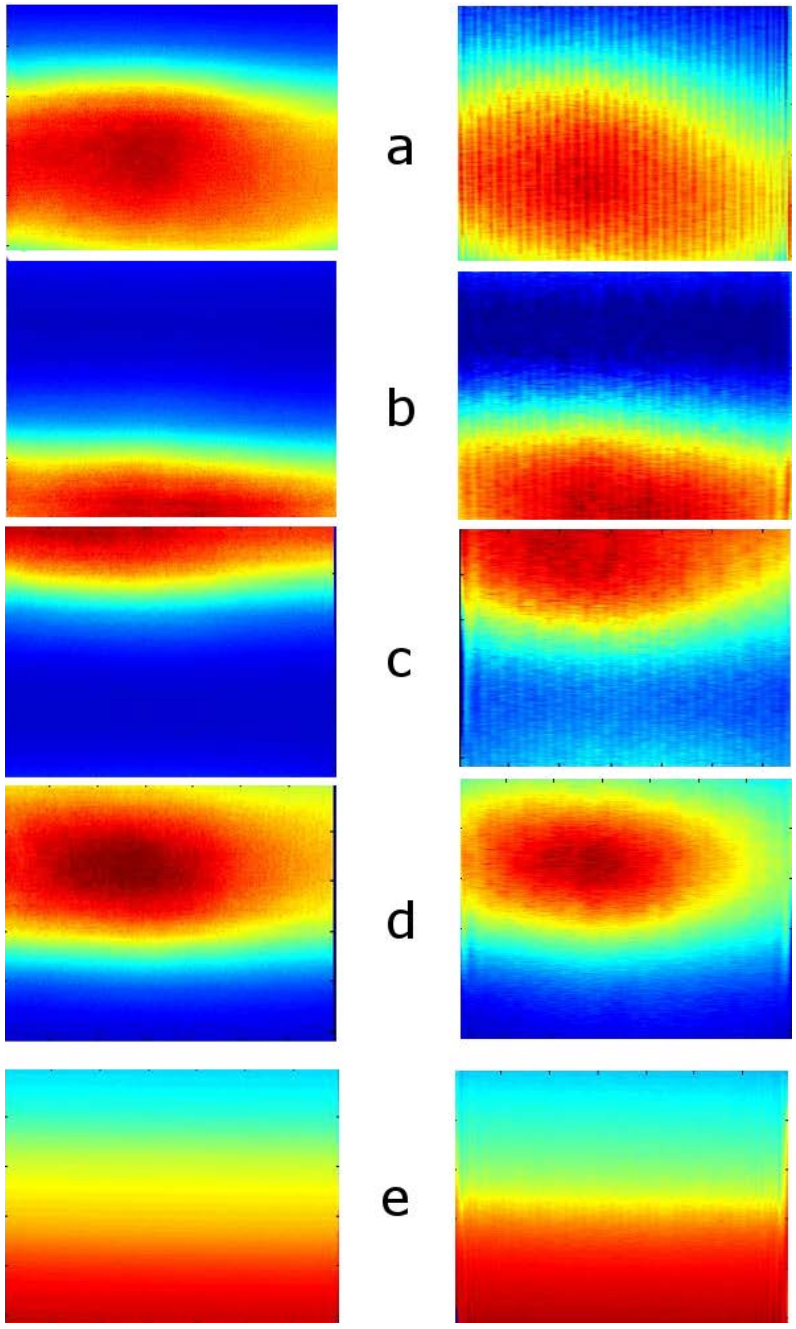

Figure 5.21, PMP (left) versus demodulated CP (right). Individual patterns are shown in (a-d), resulting in the calculated phase image (e). Color represents intensity.

The root mean squared coordinate error for the compared scans can be seen in Table 5.6.

Table 5.6, Error results (Generation 1 CP vs. 4 pattern PMP)

Subject (CP vs. 4 pattern PMP)	X error (mm)	Y error (mm)	Z error (mm)
Board	37.3001	26.3979	72.0078

Similarly, the same board was then scanned with the new third generation MCP pattern, followed by a full multi-frequency PMP scan to create a ground truth depth image (similar to the experiment documented in the previous section, see Figure 5.16 and Table 5.5). The rms error of the comparison of each can be seen in Table 5.7 below.

Table 5.7, Error results (Generation 3 MCP vs. Multi-frequency PMP)

Subject (MCP vs. Multi-frequency PMP)	X error (mm)	Y error (mm)	Z error (mm)
Board	0.7362	0.4826	2.4707

As a final test of the methodology, third generation MCP was used to scan a complex subject representative of a realistic subject, specifically, a latex human face sculpture (as shown in Figure 5.22). An albedo image was acquired separately and used merely for surface coloration.

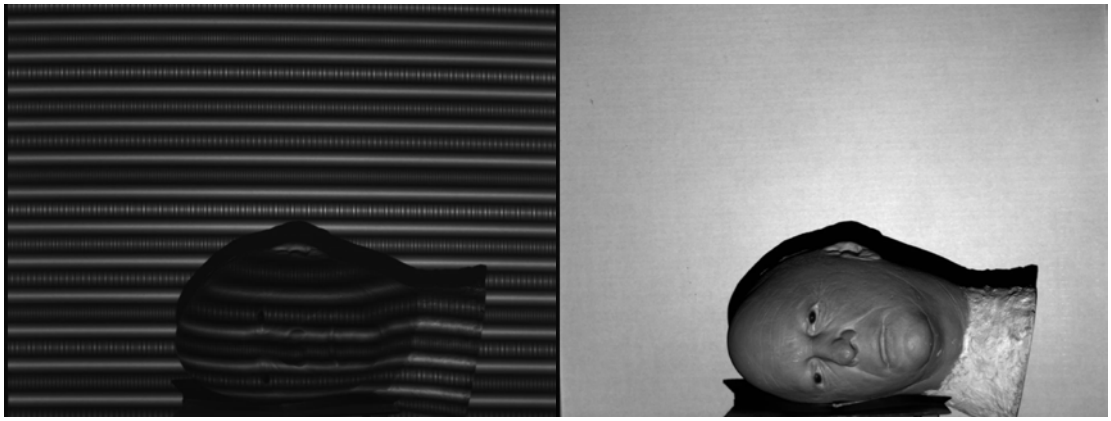

Figure 5.22, (left) MCP scan image of the scuplture, (right) albedo image of the sculpture

The image was processed to recover depth, and a linear interpolation and smoothing process was applied to create a continuous surface. The acquired depth model results are shown in Figure 5.23

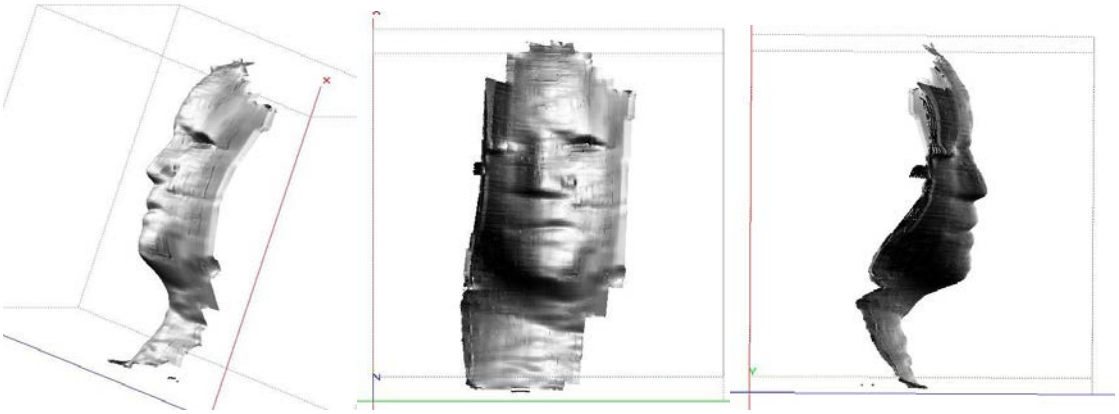

Figure 5.23, MCP scan results for a human scultpure - multiple views

As you can see, even complex subject contours can be captured very effectively using this new MCP methodology. Thus, MCP can be seen, both subjectively and quantitatively, to be a meaningful improvement on the original CP technique.

Chapter 6 Conclusion and Future Work

This dissertation introduced the Modified Composite Pattern technique for structured light depth recovery. A mathematical model of the pattern was developed, in addition to models of potential sources of systematic interference. These were used in the research to guide improvements to the pattern, allowing for potential improvements in decoding, specifically in the reduction of gamma based pattern cross-talk. The results of step-edge isolation of sub-regions of the projection pattern (an effect which is quite likely to occur in realistic scanning situations) were investigated, modeled, and experimentally quantified. These quantified results lend themselves to the direction of various post-processing operations. The Hall camera model and basic calibration methodology used in the research were derived, and the new MCP specific calibration procedure was described. An adaptive thresholding method for the isolation of MCP stripe pattern elements was developed. This method also proved useful when utilized in a new stripe extension post-processing technique. A statistical threshold was suggested for MCP images, with justifications following. An analysis-by-synthesis envelope detection process was introduced, allowing fully automated recovery and demodulation of composite pattern scan component images; and a technique for segmentation of MCP images by apparent surface inclination was developed which may allow one to successfully recover the depth of objects that display significant local Z direction inclination (which is effectively impossible with only basic CP operations). The MCP decoding process was considered in the context of a Hidden Markov Model, and a Viterbi best-path decoder was implemented. Finally, results at each stage of MCP processing

were displayed, and final results of the processing were shown, and the error involved in the decoding and recovery of the 3D model were measured.

6.1 Future Work

Despite the success of the research already performed, the Modified Composite Pattern method, while apparently an improvement over standard Composite Pattern methods, may yet be further developed.

Improvements to the MCP Hidden Markov Model

The accuracy of decoding achieved by the Viterbi decoder is limited by the quality of the HMM representation of the system in question. In this case, an MCP scan was described using a model derived by only the most basic observations available. It is likely that further improvement of the accuracy of the model will allow the Viterbi process to be even more effective in ensuring the quality of the MCP model.

One option for improvement of the model is the use of traditional HMM training sets, allowing the model to be derived in an orderly way from real results. To accomplish this, one would have to collect useful results for a wide variety of subjects and somehow ensure their quality as training images. As the processing method is initially imperfect (thus producing imperfect training images), this suggests an iterative process, gradually improving the model by using results of earlier, less accurate models (in conjunction with useful post-processing methods).

A second option, along the same line, could be to iteratively improve the model by making additional statistical observations of output results (performing essentially the same sorts of observations described in section 4.5) and integrating these new observations into the model according to some form of scheduling method. This option may advantageously be performed on any MCP scan individually without the need for a large training set, but at the same time, can potentially carry errors as well as improvements into each new iteration, potentially degrading the model rather than improving it.

Object Segmentation

One extremely useful area of improvement would be the integration of object segmentation methods into the MCP paradigm. Certainly, isolating individual objects in a scan image can lead to detrimental effects (as shown in section 3.3). Nevertheless, if these step-edge effects can be eliminated from consideration (using Viterbi decoding, or perhaps blob-isolation or other post-processing methods), the resulting object would no longer be subject to the effects of ambiguous stripe ambiguity. If the sphere in Figure 5.19, for example, were isolated from the surrounding background before post-processing operations were performed, it is unlikely that it would display the misidentification errors present, as the stripe growth post-processing algorithm would no longer incorrectly extend the sphere or background stripes.

Appendix A: Nomenclature, Terminology, and Mathematical Symbols Reference

ABS- Analysis By Synthesis

CG- Computer Graphics

CMM- Coordinate Measuring Machine

CP- Composite Pattern

DC- Direct Current - Defined as the zero frequency component of any spectrum

DLP- Digital Light Projector

FA- Finite Automaton

FSA- Finite State Automaton

FFT- Fast Fourier Transform

FSM- Finite State Machine

HMM- Hidden Markov Model

LIDAR- Light Detection And Ranging

MCP- Modified Composite Pattern

PMP- Phase Modulation Profilometry

PSD- Power Spectral Density

SLI- Structured Light Illumination

UMM- Universal Measuring Machine

albedo- An image representing surface reflectance

ambiguity- Uncertainty in camera/projector image correspondence information

gamma- Intensity nonlinearity in a display device

Hadamard multiplication- Multiplication of corresponding elements in two matrices

orthogonal direction- Direction in a structured light scan or pattern which is orthogonal to the phase direction

phase- From PMP methodology of SLI, an encoding of camera/projector correspondence with a range of 2π.

phase map- An image wherein each pixel value represents a phase value

phase direction- From PMP methodology of SLI, the direction in which camera/projector correspondence values vary in a given SLI scanner arrangement.

Λ- Canonical triangle function

$\mathfrak{I}(\ )$- Fourier transform of

$\sum_{A}^{B}$ - Summation of, within the range defined by lower limit A and upper limit B

$\mathbf{\Pi}$- Canonical rectangle function

δ - Delta function, defined as the Dirac delta function in continuous contexts, and Kronecker delta function in discrete contexts

$E\{\ \}$- Expectation of

μ- Variable representing mean value

σ- Variable representing standard deviation

argmax($f(x)$)- The result of this operation is the argument value of x which maximizes the function $f(x)$.

References

[1] Jarvis, R. A. (1983), A Perspective on Range Finding Techniques for Computer Vision, *IEEE Trans. on Pattern Analysis and Machine Intelligence* **5**(2), 122-139, 1983.

[2] T. Strand, Optical Three-Dimensional sensing for Machine Vision, *Optical Engineering*, **24**, 33-40 1985.

[3] T.S. Newman, "A survey of automated visual inspection," Image Understanding, 61, 231-262, 1995.

[4] P. Graebling, A. Lallement, D. Zhou and E. Hirsch, "Optical high-precision three dimensional vision-based quality control of manufactured parts by use of synthetic images and knowledge for image-data evaluation and interpretation," Applied Optics, 41(14), 2627-2643, 2002.

[5] G. Bonser and G. A. Parker, "Robotic gas metal arc welding of small diameter saddle joints using multistripe structured light," Opt. Eng., 38(11), 1943-1949, 1999.

[6] A. Broggi, "Vision-based driving assistance in vehicles of the future," IEEE Intell. Systems, 13(6), 22-23, 1998.

[7] L. Charbronnier and A. Fournier, “Heading guidance and obstacles localization for an indoor mobile robot,” IEEE International Conference on Advanced Robotics, 507-513, 1995.

[8] H. Moravec and M. Martin, “Robot Navigation by 3D Spatial Evidence Grids. Mobile Robot Laboratory,” Robotics Institute, Carnegie Mellon University, 1994.

[9] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin and H. Fuchs, “The office of the future: a unified approach to image-based modeling and spatially immersive displays,” SIGGRAPH 98, Orlando, Florida, July 19-24, 1998.

[10] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Gintzton, J. Davis, J. Ginsberg, J. Shade and D. Fulk, “The digital Michelangelo project: 3D scanning of large statues,” ACM SIGGRAPH Proceedings, 131-144, 2000.

[11] O. D. Faugeras and M. Hebert, “The representation, recognition, and locating of 3-d objects,” International Journal of Robotic Research, 5(3) 27-52, 1986.

[12] P. J. Besl and N. D. McKay, “A method for registration of 3-d shapes,” IEEE Trans. on Pattern Anal. Mach. Intell. 14(2) 239-256, 1992.

[13] De Piero, F. and M. Trivedi, 3D Computer Vision using Structured Light: Design, Calibration, and Implementation Issues, *Advanced in Computers*, **43**, pp. 243-278,1996.

[14] M. Tistarelli, E. Grosso, G. Sandini, Dynamic Stereo in Visual Navigation, Proceedings of the IEEE CVPR, 186-193, 1991

[15] D. Ashbrook T. Starner, J. Auxier and M. Gandy, The gesture pendant: A selfilluminating, wearable, infrared computer vision system for home automation control and medical monitoring, Fourth International Symposium on Wearable Computers (ISWC'00), Atlanta, Georgia, 2000.

[16] J. C. Carr, W. R. Fright, A. H. Gee, R. W. Prager and K. J. Dalton, "3D Shape Reconstruction using Volume Intersection Techniques". Proceedings of the IEEE Intl. Conf. Comp. Vision, Bombay, India, pp1095-1110, January 1998.

[17] A. Weckenmann, G. Peggs and J. Hoffman, "Probing systems for dimensional micro-and nano- metrology," Meas. Sci. Technol., 17, 504509, 2006.

[18] M. Bauza, R. Hocken, S. Smith, and S. Woody, "The development of a virtual probe tip with application to high aspect ratio microscale features," Rev. Sci Instrutum, 76(9) 095112, 2005.

[19]http://www.larkencnc.com/probe.htm

[20]http://www.micro-machine-shop.com/3D_digitizing.htm#3D_Digitizing_Probe

[21]http://paleo-tech.com/microscribeg2.aspx

[22]http://www.boulderinnovators.com/PDFs/indCatalog.pdf

[23]http://www.amfit.com/products/contact-digitizer

[24]http://en.wikipedia.org/wiki/File:9.12.17_Coordinate_measuring_machine.png

[25] S. K. Nayar and Y. Nakagawa, Shape from focus: an effective approach for rough surfaces, *IEEE Trans. On Pattern Analysis and Machine Intelligence*, **16**(8): 824-831, August 1994.

[26] T. Darrell and K. Wohn, Pyramid based depth from focus, Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, pp. 504-509, June 1988.

[27] J. Ens and P. Lawrence, A matrix based method for determining depth from focus, Proc. of IEEE conf. on Computer Vision and Pattern Recognition, pp. 600-609, June 1991.180

[28] A. Pentland, A new sense for depth of field, *IEEE Trans. On Pattern Analysis and Machine Intelligence*, **9**(4): 523-531, July 1987

[29] Ruo Zhang, Ping-Sing Tsai, James Edwin Cryer, and Mubarak Shah. Analysis of shape from shading techniques. In Proceedings, Conference on Computer Vision and Pattern Recognition, pages 377--384, Seattle, Washington, June 21--23 1994.

[30] B. K P. Horn, Shape from shading: A method for obtaining the shape of a smooth opaque object from one view, M.I.T. Project MAC Rep. TR-79, 1970

[31] G. Schmaltz of Schmaltz Brothers Laboratories, "A method for presenting the profile curves of rough surfaces," *Naturwiss* 18, 315–316 (1932)

[32] Chun Guan, "Composite Pattern for Single Frame 3D Acquisition," PhD Dissertation, University of Kentucky, Lexington, Ky, USA, 2004.

[33] J. Potsdamer and M. Altschuler “Surface measurement by space-encoded projected beam system,” *Computer Graphics Image Processing* 18:1–17 (1982)

[34] F. Gray. *Pulse code communication*, March 17, 1953 (filed Nov. 1947)

[35] K. Sato and S. Inokuchi, “Three-dimensional Surface Measurement by Space Encoding Range Imaging,” *Journal of Robotic Systems*, 2(1), 27-39, (Spring 1985)

[36] S. Inokuchi, K. Sato, F. Matsuda, Range imaging system for 3D object recognition, Proceedings of the international Conference on Pattern Recognition, pp. 806-808,

1984.

[37] M. Minou, T. Kanade, and T. Sakai, A method of timecoded parallel planes of light for depth measurement, *Transactions of the IECE of Japan*, 64:521--528, 8 1981.

[38]http://upload.wikimedia.org/wikipedia/en/a/ad/13-stripes-s.png

[39] O. Hall-Holt and S. Rusinkiewicz, "Stripe Boundary Codes for Real-Time Structured-Light Range Scanning of Moving Objects," *Proc. Int'l Conf. Computer Vision*, 359-366, (2001)

[40] S. Rusinkiewicz, O. Hall-Holt, and M. Levoy, "Real-time 3d model acquisition," ACM Trans. Graphics Proc: ACM SIGGRAPH 02, 21(3), 2002.

[41] V. Srinivasan, H.C. Liu, M. Halioua, "Automated phase-measuring profilometry of 3-D diffuse objects," *Applied Optics*, 23(18), 3105-8, (1984)

[42] Veeraganesh Yalla, "Optimal Phase Measuring Profilometry Techniques for Static and Dynamic 3D Data Acquisition," PhD Dissertation, University of Kentucky, Lexington, Ky, USA, 2006

[43] Jielin Li, Laurence G. Hassebrook, and Chun Guan, "Optimized two-frequency phase-measuring profilometry light-sensor temporal-noise sensitivity," *J. Opt. Soc. Am. A*, 20(1), (2003)

[44] Jie-lin Li, Hong-jun Su and Xian-yu Su, Two-frequency grating used in phasemeasuring profilometry, *Applied Optics*, vol. **36**, January 10, pp. 277-280, 1997.

[45] Daniel Scharstein and Richard Szeliski "High Accuracy Stereo Depth Maps Using Structured Light," In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2003), volume 1, pages 195-202, Madison, WI, June 2003.

[46] Lavoie, P.; Ionescu, D.; Petriu, E.M.; , "3D object model recovery from 2D images using structured light," Instrumentation and Measurement, IEEE Transactions on , vol.53, no.2, pp. 437- 443, April 2004

[47] Qingcang Yu, Xiaojun Jia, Jian Tao, Yun Zhao. "An Encoded Mini-grid Structured Light Pattern for Dynamic Scenes." In Proceedings of ICIC (1)'2005. pp.126~135

[48] Song Zhang and Peisen S. Huang, "High-resolution Real-time 3-D Shape Measurement," *Opt. Eng.*, Vol. 45, No 12, 123601 2006.

[49] Pratibha Gupta, "Gray Code Composite Pattern Structured Light Illumination," MS Thesis, University of Kentucky, Lexington, Ky, USA, 2007.

[50] A. Robinson, L. Alboul, and M. Rodrigues. “Methods for indexing stripes in uncoded structured light scanning systems,” *Journal of WSCG*, 12(3):371-378, February 2004.

[51] M. Proesmans, L. Van Gool, F. Defoort, Reading between the lines - a method for extracting dynamic 3D with texture, Sixth international conference on computer vision, pp. 1081-1086, January 4-7, 1998.

[52] W. Brink, A. Robinson, M. Rodrigues: Indexing Uncoded Stripe Patterns in Structured Light Systems by Maximum Spanning Trees, British Machine Vision Conference BMVC 2008, Leeds, UK, 1-4 Sep 2008

[53] P. Vuylsteke and V. Oosterlinck, “Range image acquisition with a single binary-encoded light pattern,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(2):148-163, (1990)

[54] H. Morita, K. Yajima, S. Sakata, “Reconstruction of surfaces of 3-d objects by *m*-array pattern projection method,” *IEEE International Conference on Computer Vision*, 468–473 (1988)

[55] F. J. MacWilliams and N. J. A. Sloane, “Pseudorandom sequences and arrays,” *Proceedings of the IEEE*, 64(12):1715-1729, (1976)

[56] M. Maruyama and S. Abe, “Range sensing by projecting multiple slits with random cuts,” *IEEE Trans. Pattern. Anal. Mach. Intell*. 15, 647–651 (1993)

[57] O. Hall-Holt and S. Rusinkiewicz, “Stripe Boundary Codes for Real-Time Structured-Light Range Scanning of Moving Objects,” *Proc. Int'l Conf. Computer Vision*, 359-366, (2001)

[58] J. Pagès, J. Salvi and C. Matabosch, “Implementation of a robust coded structured light technique for dynamic 3D measurements,” *IEEE International Conference on Image Processing Barcelona, Spain* 1073–1076 (September 2003)

[59] Weiyi Liu, Zhaoqi Wang, Guoguang Mu, and Zhiliang Fang,"Color-coded projection grating method for shape measurement with a single exposure," *Applied Optics*, 39(20):3504-8, July 2000.

[60] Changsoo Je, Sang Wook Lee, and Rae-Hong Park, "High-Contrast Color-Stripe Pattern for Rapid Structured-Light Range Imaging," *ECCV 2004, LNCS 3021*, pp. 95–107, 2004.

[61] Hao Li, Raphael Straub, Hartmut Prautzsch, “Fast Subpixel Accurate Reconstruction Using Colored Structured Light,” *Proceedings of the Fourth IASTED International Conference on Visualization, Imaging and Image Processi 2004*, 396-401, 09/2004

[62] Z. J. Geng, "Rainbow three-dimensional camera: new concept of high-speed three-dimensional vision systems," Opt. Eng. 35(2), 376–383 (1996)

[63] D. Caspi, N. Kiryati, and J. Shamir, Range imaging with adaptive color structured light, *Pattern analysis and machine intelligence*, **20**(5):470--480, May 1998.

[64] Wei-Hung Su, “Color-encoded fringe projection for 3D shape measurements,” *Opt. Express* 15, 13167-13181 (2007)

[65] Jaihui Pan, Peisen S. Huang, and Fu-Pen Chiang, “Color-phase shifting technique for three-dimensional shape measurement,” *Optical Engineering* -- January 2006 -- Volume 45, Issue 1, 013602

[66] J. Pan, P. S. Huang, and F. -P. Chiang, "Color-coded binary fringe projection technique for 3-D shape measurement," Opt. Eng. 44(2), 623606 (2005)

[67] P. S. Huang, Q. Hu, F. Jin, and F. P. Chiang, "Color-encoded digital fringe projection technique for high-speed three-dimensional surface contouring," Opt. Eng. 38(6), 1065–1071 (1999)

[68] Li Zhang, Brian Curless, and Steven M. Seitz, "Rapid Shape Acquisition Using Color Structured Light and Multi-pass Dynamic Programming," *Proceedings of the 1st*

International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT), Padova, Italy, June 19-21, 2002, pp. 24-36.

[69] Oleksandr A. Skydan, Michael J. Lalor, and David R. Burton, "Technique for phase measurement and surface reconstruction by use of colored structured light," *Applied Optics*, 41(29):6104-17, 2002.

[70] Hao Li, Raphael Straub, Hartmut Prautzsch “FAST SUBPIXEL ACCURATE RECONSTRUCTION USING COLOR STRUCTURED LIGHT” Proceedings of the Fourth IASTED International Conference on Visualization, Imaging and Image Processing 2004, 396-401, 09/2004

[71] C Guan, LG Hassebrook, DL Lau, "Composite structured light pattern for three-dimensional video," *Optics Express*,11(5): 406–17 (2003)

[72] C. J. Casey, L.G. Hassebrook and D. L. Lau, “Structured Light Illumination Methods for Continuous Motion Hand and Face-Computer Interaction,” *Human-Computer Interaction, New Developments, International Journal of Advanced Robotic System*, edited by Kikuo Asai, published by In-Teh, Croation branch of I-Tech Education and Publishing KG, Vienna, Austria, pp 297-308, (copyright 2008) ISBN 978-953-7619-14-5.

[73] L. G. Hassebrook, Ray C. Daley and William Chimitt, "Application of Communication Theory to High Speed Structured Light Illumination," Edited by Harding and Svetkoff, *SPIE Proceedings*, **3204**(15), 102-113 (October 1997).

[74] M.A. Rodrigues, A. Robinson, W. Brink, "Issues in Fast 3D Reconstruction from Video Sequences", Lecture Notes in Signal Science, Internet and Education, Proceedings of 7th WSEAS International Conference on MULTIMEDIA, INTERNET & VIDEO TECHNOLOGIES (MIV '07), Beijing, China, September 15-17, 2007, pp 213-218.

[75] Kai Liu, Yongchang Wang, Daniel L. Lau, Qi Hao, and Laurence G. Hassebrook, "Gamma model and its analysis for phase measuring profilometry," *J. Opt. Soc. Am. A* 27, 553-562 (2010)

[76] E.L. Hall, J.B.K. Tio, C.A. McPherson, F.A. Sadjadi, Measuring curved surfaces for robot vision, *Comput. J.* **15**, 42–54, 1982.

[77] J. Salvi, X. Armangué and J. Batlle. A Comparative Review of Camera Calibrating Methods with Accuracy Evaluation. *Pattern Recognition*, pp. 1617-1635, vol. **35**, no. 7, 2002

[78] J. Weng, P. Cohen, and M. Herniou, Camera calibration with distortion models and accuracy evaluation, *IEEE Trans. Pattern Analysis Machine Intel*., **14**:965–980, 1992.

[79] R.Y. Tsai, A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses, *IEEE J. Robotics and Automation*, vol. **3**, no. 4, pp. 323--344, August 1987

[80] Laurence G. Hassebrook, "Composite correlation filter for O-ring detection in stationary colored noise," Invited paper, Optical Pattern Recognition XX, SPIE Defense and Security Symposium, edited by David P. Casasent; Tien-Hsin Chao, Orlando, Florida. Vol. 7340, pp 734007-1 to 734007-8 (April 2009)

[81] P. M. Will and K. S. Pennington. 1971. Grid coding: a preprocessing technique for robot and machine vision. In *Proceedings of the 2nd international joint conference on Artificial intelligence* (IJCAI'71). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 66-70.

[82] L. R.Rabiner. “A tutorial on hidden markov models and selected apllications in speech recognition,” In A. Waibel and K. F. Lee, editors, *Readings in Speech Recognition*, pages 267-296. Kaufmann, San Mateo, CA, 1990

[83] B. Pardo and W. Birmingham, “Modeling Form for On-line Following of Musical Performances,” *AAAI-05 Proc.*, July 2005

[84] Lee A Newberg. “Error statistics of hidden Markov model and hidden Boltzmann model results” BMC Bioinformatics 2009, 10:212 doi:10.1186/1471-2105-10-212

[85] Ara V. Nefian, Luhong Liang, Xiaobo Pi, Xiaoxing Liu, Crusoe Mao Kevin Murphy, “A coupled HMM for audio-visual speech recognition”, *IEEE International Conference on Acoustic Speech and Signal Processing* 2002.

[86] Wei Wang, Xiaohong Guan, Xiangliang Zhang, "Modeling Program Behaviors by Hidden Markov Models for Intrusion Detection". *The Third IEEE International Conference on Machine Learning and Cybernetics (ICMLC2004),* Shanghai, China, pp. 2830-2835. Aug 2004

[87] A. J. Viterbi, “Error bounds for convolutional codes and an asymptotically optimum decoding algorithm,” IEEE Trans. Inform. Theory, vol. IT–13, pp. 260–269, April 1967

[88] Boyle, R. D. "Hidden Markov Models." School of Computing - Faculty of Engineering - University of Leeds. Web. 13 Jan. 2011. <http://www.comp.leeds.ac.uk/roger/HiddenMarkovModels/html_dev/main.html> l>

VITA

Charles Joseph Casey

Born: Concord, Mass. USA
March 18, 1984

Education

- M.S.E.E. University of Kentucky College of Engineering , 2008
- Thesis: Structured Light Motion Capture
- B.S. Georgetown College, 2006
- Majors: Physics / Mathematics
- Minors: Computer Science / Chemistry

Professional Experience

- Project "Real-Time, 3-D Finger and Palm-Print Scanner for Entry and Access Portal Security," Department of Homeland Security administered by National Institute of Hometown Security 2006-2007-Research Assistant

- Project "Rolled Equivalent, 3-D Fingerprint/Palm Print Scanner." Department of Homeland Security and the National Institute of Hometown Security, 2010 - Research Assistant

- Project "Active and Passive Range Sensor Fusion for Automated Surveillance and Face Recognition," Department of Homeland Security administered by National Institute of Hometown Security 2006-2008-Research Assistant

- Project "High Resolution Composite Pattern Proof of Concept" 3D Imaging LLC. 2007-Research Assistant

- Project "Facial Expression 3-D Scanner" Gentle Giant Studios, Inc. 2006- 2007 - Research Assistant.

- Project "3-Dimensional Surround Scanning and Surface Feature Tracking," Kentucky Science and Technology Corporation, 2008-Research Assistant

- Inventor -UKRF IP Disclosure *Structured Light Time Division Multiplexing or Pattern Interleaving.*

- Inventor -UKRF IP Disclosure No. 1428 *Lock and Hold Structured Light Illumination.*

- Expert declarant regarding novel techniques introduced in Patent 10/444,033 *System and Technique for Retrieving Depth Information about a Surface by Projecting a Composite Image of Modulated Light Patterns.*

Publications

- **Charles Casey**, Laurence G. Hassebrook, Eli Crane, Aaron Davidson, "Multi-Feature Distortion Insensitive Constellation Detection," - Applied Optics - Accepted March 2011

- **Charles Casey**, Laurence G. Hassebrook, "Multifeature Constellation Correlation Filters" Accepted Optical Pattern Recognition XXII, *SPIE Defense and Security Symposium* (2011*)*

- **Charles Casey**, Laurence G. Hassebrook, "Automated Modified Composite Pattern Single Image Depth Acquisition " Accepted Optical Pattern Recognition XXII, *SPIE Defense and Security Symposium (*2011)

- **C. J. Casey**, L.G. Hassebrook and D. L. Lau, "Structured Light Illumination Methods for Continuous Motion Hand and Face-Computer Interaction," *Human-Computer Interaction, New Developments, International Journal of Advanced Robotic System*, edited by Kikuo Asai, published by In-Teh, Croation branch of I-Tech Education and Publishing KG, Vienna, Austria, pp 297-308, (copyright 2008) ISBN 978-953-7619-14-5.

- Chun Guan, Laurence G. Hassebrook, MEMBER SPIE; Daniel L. Lau, Veeraganesh G. Yalla, and **Charles J. Casey** "Improved composite-pattern structured-light profilometry by means of postprocessing" Opt. Eng., Vol. 47, 097203 (2008);

- E. R. Crane, L.G. Hassebrook, C.T. Begley, W. F. Lundby and **C. J. Casey**, "Methodology and Technology for Rapid Three-Dimensional Scanning of In Situ Archaeological Materials in Remote Areas," (presented and under preparation) Edited by Bernard Breuckmann. Interpretation and Evaluation of High Definition 3D Surface Data III. *Computer Applications and Quantitative Methods in Archaeology*, Granada, Spain (April 2010).

- Laurence G. Hassebrook, **Charles J. Casey** and Walter Lundby, "Non-Contact Fiducial Based 3-Dimensional Patch Merging Methodogy and Performance," Three-Dimensional Surface Recording, Analysis, and Interpretation in Archaeology and Anthropology, *Computer Applications and Quantitative Methods in Archaeology*, Williamsburg, Virginia, in press, number 346 (May 2009).

- **Charles Casey**, Laurence G. Hassebrook and Priyanka Chaudhary, "Correlation based swarm trackers for 3-dimensional manifold mesh formation," Optical Pattern Recognition XX, *SPIE Defense and Security Symposium*, edited by David P. Casasent; Tien-Hsin Chao, Orlando, Florida. Vol. 7340, pp 73400G-1 to 73400G-7 (April 2009).

- Laurence G. Hassebrook, Akshay Pethe, Veeraganesh Yalla, **Charles Casey** and Daniel L. Lau, "Super Resolution Structured Light Illumination," Sensors and Systems for Space Applications, *SPIE Defense and Security Symposium*, edited by Richard T. Howard; Robert D. Richards, Orlando, Florida. Vol. 6555, (April 2007).

- Technical Report - "3-Dimensional Scanner Prototype for Scanning Archeological Artifacts in Remote Areas" L.G. Hassebrook, E. Crane, C. Casey, W. Lundby and C. Begley (2009).

- Invention Disclosure - L.G. Hassebrook, **Charles J. Casey**, Eli R. Crane, Walter F. Lundby, Yongchang Wang, Kai Liu and Daniel L. Lau, "Rotate and Hold and Scan Structured Light Illumination Pattern Encoding and Decoding," Submitted to University of Kentucky Intellectual Property Development Office, 11/30/2009. INV09/1706.

- Invention Disclosure - **C. J. Casey**, L. G. Hassebrook and D. L. Lau, "Structured Light Time Division Multiplexing, aka: Pattern Interleaving" submitted to University of Kentucky Intellectual Property Development Office, April 14, 2008

- Provisional Patent: L.G. Hassebrook, **Charles J. Casey**, Eli R. Crane, Walter F. Lundby, Kai Liu, Yongchang Wang, and Daniel L. Lau, "Rotate and Hold and Scan (RAHAS) Structured Light Illumination Pattern Encoding and Decoding," Filed 6/24/2010. Application No. 61/358,397, UKRF-136P.

- Patent Pending: L. G. Hassebrook, Daniel L. Lau and **Charles J. Casey**, "Lock and Hold Structured Light Illumination," Patent Application No. 12/284,253. **MBH** Attorney Docket UKRF-133P, University of Kentucky Intellectual Property Development, University of Kentucky, (September 18, 2008).

- Provisional Patent: L. G. Hassebrook, Daniel L. Lau and **Charles J. Casey**, "Lock and Hold Structured Light Illumination," Provisional Patent Application No. 60/994,181. MBH Attorney Docket UKRF-133P, University of Kentucky Intellectual Property Development, University of Kentucky, (September 18, 2007).

- Veeraganesh Yalla, **Charles Casey**, Pratibha Gupta, Wei Su, Meng, L.G. Hassebrook and D. L. Lau, Poster Presentation "College of Engineering Academic Fair, 3D Facial Visualization," University of Kentucky, October 13 and 14, 2006.