



University of Kentucky
UKnowledge

University of Kentucky Doctoral Dissertations

Graduate School

2010

Novel Approaches in Structured Light Illumination

Yongchang Wang

University of Kentucky, ychwang6@gmail.com

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

Recommended Citation

Wang, Yongchang, "Novel Approaches in Structured Light Illumination" (2010). *University of Kentucky Doctoral Dissertations*. 116.

https://uknowledge.uky.edu/gradschool_diss/116

This Dissertation is brought to you for free and open access by the Graduate School at UKnowledge. It has been accepted for inclusion in University of Kentucky Doctoral Dissertations by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

ABSTRACT OF DISSERTATION

Yongchang Wang

The Graduate School
University of Kentucky
2010

Novel Approaches in Structured Light Illumination

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Electrical and Computer Engineering at the University of Kentucky

By
Yongchang Wang
Lexington, Kentucky

Director: Dr. Daniel L. Lau, Associate Professor of Electrical and Computer Engineering
Lexington, Kentucky 2010

Copyright© Yongchang Wang 2010

ABSTRACT OF DISSERTATION

Novel Approaches in Structured Light Illumination

Among the various approaches to 3-D imaging, structured light illumination (SLI) is widely spread. SLI employs a pair of digital projector and digital camera such that the correspondences can be found based upon the projecting and capturing of a group of designed light patterns. As an active sensing method, SLI is known for its robustness and high accuracy. In this dissertation, I study the phase shifting method (PSM), which is one of the most employed strategy in SLI. And, three novel approaches in PSM have been proposed in this dissertation. First, by regarding the design of patterns as placing points in an N-dimensional space, I take the phase measuring profilometry (PMP) as an example and propose the edge-pattern strategy which achieves maximum signal to noise ratio (SNR) for the projected patterns. Second, I develop a novel period information embedded pattern strategy for fast, reliable 3-D data acquisition and reconstruction. The proposed period coded phase shifting strategy removes the depth ambiguity associated with traditional phase shifting patterns without reducing phase accuracy or increasing the number of projected patterns. Thus, it can be employed for high accuracy realtime 3-D system. Then, I propose a hybrid approach for high quality 3-D reconstructions with only a small number of illumination patterns by maximizing the use of correspondence information from the phase, texture, and modulation data derived from multi-view, PMP-based, SLI images, without rigorously synchronizing the cameras and projectors and calibrating the device gammas. Experimental results demonstrate the advantages of the proposed novel strategies for 3-D SLI systems.

KEYWORDS: Structured light illumination, 3-D reconstruction, period coded phase shifting, edge-pattern, hybrid system.

Author's signature: Yongchang Wang

Date: April 28, 2010

DISSERTATION

Yongchang Wang

The Graduate School
University of Kentucky
2010

Novel Approaches in Structured Light Illumination

DISSERTATION

A dissertation submitted in partial
fulfillment of the requirements for
the degree of Doctor of Philosophy
in the Department of Electrical and
Computer Engineering at the
University of Kentucky

By
Yongchang Wang
Lexington, Kentucky

Director: Dr. Daniel L. Lau, Associate Professor of Electrical and Computer
Engineering
Lexington, Kentucky 2010

Copyright© Yongchang Wang 2010

ACKNOWLEDGMENTS

I would like to thank all those people who have helped me for this dissertation.

Here, firstly, my sincere thanks to my advisor, Dr. Daniel Lau for giving me an opportunity to work on this project and lead me into this area. Also Dr. Lau taught me so much from the whole idea to the technique details. It is his persistence for perfection and interest in little details that have supplemented my own quest for knowledge. I would also like to extend my great thanks to Dr. L. G. Hassebrook for his constant encouragement and support during all the time. Whenever I met problems Dr. Lau and Dr. Hassebrook always helped me and taught me how to solve them.

Thank Kai Liu, especially when I first came into this area, Kai gave me so much help and worked together with me through all the problems. I am also grateful to Qi Hao, Eric Dedrick, Veer Ganesh Yalla and Abhishika Fatehpuria. Discussions with them regarding the work, were always interesting and intellectual.

This work would not have been possible without the support and love of my parents and wife. They were the ones who always encouraged and motivated me to go ahead. I also thank all my friends for always putting a smile on my face during tough times and support me.

Dedicated to my family.

TABLE OF CONTENTS

Acknowledgments	iii
Table of Contents	iv
List of Figures	vi
List of Tables	ix
Chapter 1 INTRODUCTION	1
1.1 Structured light illumination	6
1.2 Phase shifting methods	15
1.3 Problems in phase shifting methods	22
1.4 Contributions	28
Chapter 2 MAXIMIZATION SNR PATTERN STRATEGY	37
2.1 Phase measuring profilometry	38
2.2 Influence of intensity noise	43
2.3 Edge-pattern	46
2.4 Experimental results and discussion	52
Chapter 3 PERIOD CODED PHASE SHIFTING	61
3.1 Phase shifting structured light illumination	62
3.2 Codification of projected patterns	73
3.3 De-codification of received images	83
3.4 Accuracy analysis	87
3.5 Experimental results and discussion	89
Chapter 4 HYBRID FRAMEWORK	104
4.1 Stereo matching in phase	108
4.2 Phase error modeling	121
4.3 Hybrid 3-D reconstruction	129
4.4 Experimental results and discussion	137
Chapter 5 CONCLUSION AND FUTURE WORK	153
Bibliography	161
VITA	178
Yongchang Wang	178
Education	178
Professional experience	178
Invention disclosures	179

Publications	180
Presentations	182
Media news	183

LIST OF FIGURES

1.1	Ground-based LiDAR scanners.	2
1.2	Triangulation based 3-D scanning.	3
1.3	A plush gorilla.	4
1.4	SLI was used for jet pilot’s helmets.	6
1.5	Patterns designed by Zhang and Chen.	10
1.6	Patterns designed by Koninckx and Guan.	11
1.7	Cross sections of 3-pattern unit frequency PMP strategy.	16
1.8	3-D reconstruction using PMP.	17
1.9	The trapezoidal pattern strategy designed by Huang.	18
1.10	The 2 + 1 pattern strategy designed by Wizinowich.	19
1.11	The triangular pattern strategy designed by Jia.	20
1.12	The fast three step pattern strategy designed by Huang.	21
1.13	An SLI system developed by Liu and Wang.	21
1.14	3-D reconstruction of a white model, Alice, using 3-pattern PMP.	23
1.15	3-D reconstruction of a white model, Alice, using 5-pattern PMP.	24
1.16	Illustration of depth ambiguities in high-frequency patterns.	26
1.17	3-D reconstruction of a flat textured board.	29
1.18	3D reconstructions of Alice with 3, 4 and 5 pattern PMP and E-P strategies.	30
1.19	Three white boards were scanned using PCPS.	31
1.20	Schematic diagram of a multi-camera and multi-projector system.	32
1.21	3-D reconstruction of a giraffe and a space shuttle model.	33
1.22	3-D reconstruction using hybrid approach.	34
2.1	Cross sections of 3-pattern unit frequency PMP. The pattern height is 1024.	39
2.2	3-pattern PMP in three dimensional coding space.	40
2.3	Reflection process in coding space R^N	42
2.4	Relationship between computational length and created error.	44
2.5	Cross sections of the 3-pattern E-P strategy in coding space.	47
2.6	Cross sections of the 3-pattern E-P strategy.	48
2.7	The prototype system setup for E-P.	53
2.8	The scanned textured, flat, poster board. (a) The front view. (b) The top view.	54
2.9	(a) Cross sections of phase error of the scanned board from 3 pattern PMP and E-P strategies. (b) Cross sections of phase error from 4 pattern PMP and E-P strategies. (c) Cross sections of phase error from 5 pattern PMP and E-P strategies.	55
2.10	The scanned white plaster model, Alice.	57
2.11	3D reconstructions of Alice with 3, 4 and 5 pattern PMP and E-P strategies.	58

2.12	(a) Distribution of phase error of Alice from 3 pattern PMP and E-P strategies. (b) Distribution of phase error from 4 pattern PMP and E-P strategies. (c) Distribution of phase error from 5 pattern PMP and E-P strategies.	59
3.1	Illustration of four pattern, two period PMP strategy.	66
3.2	Initial unwrapping.	75
3.3	The dynamic range of linear period cue.	79
3.4	Coding functions when N is 4 and P is 4.	80
3.5	Cross sections of the PCPS patterns.	82
3.6	Temporal phase unwrapping.	85
3.7	Hybrid phase unwrapping.	86
3.8	The prototype system setup for PCPS.	90
3.9	Three white boards were scanned using PCPS.	91
3.10	Final decoded phase results.	92
3.11	The 3-D reconstruction of three boards through PCPS.	93
3.12	The 3-D reconstruction of a flat board with intensity texture.	94
3.13	The phase error variance versus the number of periods and modulation.	95
3.14	Realtime 3-D reconstruction employing proposed PCPS strategy.	96
3.15	3-D reconstructions of a stationary textured angel and a moving hand.	97
3.16	Nonlinear period cue when P is 4 and N is 3.	100
3.17	3-D reconstructions for different number of patterns strategies.	102
4.1	SLI 3D surfaces reconstructed using spot array patterns.	105
4.2	A cartoon deer and a white space plane.	109
4.3	SLI 3D surfaces reconstructed using spot array patterns.	110
4.4	Rotated PMP patterns.	113
4.5	The captured images of rotated PMP patterns.	113
4.6	Plot of PMP reconstruction errors vs. rotation angles.	114
4.7	3-D reconstruction of target.	115
4.8	Reconstruction errors vs. the rotation angles.	117
4.9	Matching errors in pixel vs. log ratio of H/L	119
4.10	3-D reconstruction of target using phase and stereo matchings.	120
4.11	Probability density function of phase error.	126
4.12	Pdf of phase error and its Gaussian approximation.	127
4.13	An epipolar line goes through the space between four image pixel centers.	130
4.14	Phase matching in low and high level noise.	131
4.15	Phase matching error.	132
4.16	Optimized weight function with respect to modulation value.	136
4.17	The prototype system setup for the hybrid 3-D reconstruction.	138
4.18	The project gamma distortion and flicker.	138
4.19	Reconstructed 3-D geometry of a textured flat board.	139
4.20	Cross sections of SLI and hybrid approaches.	142
4.21	Cross sections of sub-pixel phase matching and hybrid approaches.	144
4.22	3-D reconstruction of a giraffe and a space shuttle model.	145

4.23	3-D reconstruction of a thick and shiny hair gorilla.	146
4.24	3D visualization of Eq. (4.57).	148
4.25	RMS depth errors change with number of patterns.	150
4.26	Pyramid generalization of the hybrid 3-D reconstruction approach.	152

LIST OF TABLES

1.1	A Review of Recent Structured Light Patterns.	9
2.1	Edge order for 3 pattern strategy.	48
2.2	Four patterns E-P strategy.	48
2.3	Theoretical improvements of SNR for different number of patterns.	52
2.4	Standard deviation values of phase error of the scanned board for PMP and E-P.	56
2.5	Standard deviation values of phase error of the scanned Alice for PMP and E-P.	59
3.1	Comparison of Spatial Intensity Efficiency and Pattern Entropy.	71
3.2	Linear period cue values when N is 4 and P is 4.	80
3.3	Comparison of spatial intensity efficiency and pattern entropy.	83
3.4	Analysis of Execution Performances.	97
3.5	Comparison of Execution Performances.	99
4.1	Experimental results using the SLI method	143
4.2	Experimental results using the phase matching method	143
4.3	Experimental results using the hybrid algorithm	143
4.4	Error analysis for SLI and hybrid approaches	151

Chapter 1 INTRODUCTION

Three-dimensional (3-D) surface measurement is an area of active research [1], whose purpose is to measure or analyze a real-world object's shape and possibly its appearance (i.e. color). The collected data can then be used to construct digital models useful for a wide variety of applications [2,3]. Many different technologies can be used for 3-D measurement purposes. Generally, these technologies can be classified into contact or non-contact categories, where contact scanners probe the subject through physical touch. A coordinate measuring machine (CMM) is an example of a contact 3D scanner, which is used mostly in manufacturing [4]. By making physical contact with the scanned object, the act of scanning might modify or damage the target. Another disadvantage of CMMs is that they are relatively slow compared to the other scanning methods [5,6].

Non-contact 3-D techniques overcome the disadvantages in contact based scanning by means of either time-of-flight or triangulation. Time-of-flight technology is an active technique that finds the distance to the target's surface by measuring the round-trip time of a pulse of light. Since the speed of light is known, the round-trip time determines the travel distance of the light, which is twice the distance between the scanner and the surface. The accuracy of a time-of-flight 3-D scanner depends on how precisely it can measure the time, where 3.3 picoseconds is the approximate time taken for light to travel 1 *mm* [7,5].

A laser rangefinder only detects the distance of one point in its direction of view.



Figure 1.1: Ground-based LiDAR scanners (time-of-flight unless noted as otherwise) [8, 9, 10, 11, 12, 13, 14, 15, 16].

Thus, the scanner scans its entire field of view one point at a time by changing the range finder direction of view to scan different points. The view direction of the laser rangefinder can be changed by either rotating the range finder itself, or by using a system of rotating mirrors. The latter method is commonly used because mirrors are much lighter and can, thus, be rotated much faster and with greater accuracy. Several examples of ground-based LiDAR (single scanning laser) scanners are shown in Fig. 1.1 [8,9,10,11,12,13,14,15,16]. There are also some other techniques or products based on time-of-flight that indirectly measure the time-of-flight of a flash-like light source illuminating a large field of view and then capturing the returning light pulse

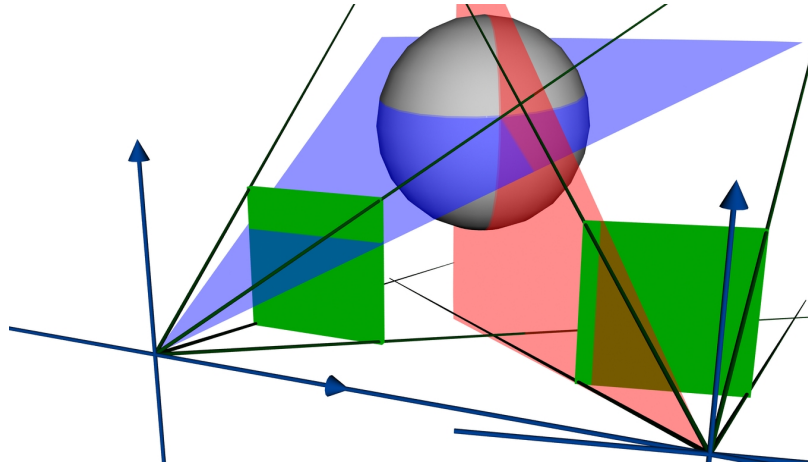


Figure 1.2: Triangulation based 3-D scanning.

using a traditional CCD camera sensor. [17, 18, 16]. Typical time-of-flight 3-D laser scanners can measure the distance of 10,000 ~ 100,000 points every second [19, 5].

There are many techniques for 3-D scanning that are based on triangulation. In triangulation based 3-D scanning, 3-D information is computed based on a pre-calibrated triangulation between two devices [20, 21, 22, 23, 24]. Figure 1.2 shows the principle of triangulation based 3-D scanning where the two devices are replaced with their pin hole models. With respect to time-of-flight 3-D scanner, triangulation scanning tries to exploit the two devices and finds the correspondences of each point between devices. For example, stereo vision techniques are employing the triangulation between two cameras. A laser scanner is based on the triangulation between a laser source and a camera.

A laser triangulation scanner shines a laser on the subject and exploits a camera to look for the location of the reflected laser spot. Depending on how far away the laser strikes a surface, the laser spot appears at different positions in the camera's field of view. After locating the spot in camera space, the camera and the laser emitter

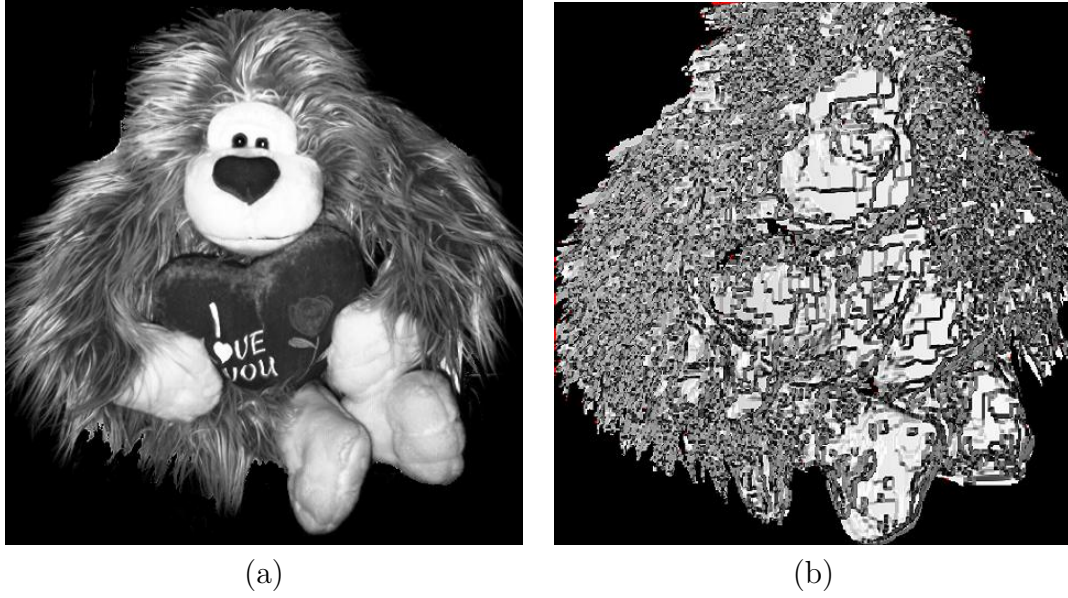


Figure 1.3: (a)A plush gorilla with shiny hair.(b)3-D reconstruction using stereo vision. Result is shown with depth rendering.

form a triangle. The length of one side of the triangle, the distance between the camera and the laser emitter is known. The angle of the laser emitter corner is also known. The angle of the camera corner can be determined by looking at the location of the laser spot in the camera's field of view. These three pieces of information fully determine the shape and size of the triangle and gives the location of the laser spot corner of the triangle.

Broadly, triangulation based techniques for 3-D scanning can be classified into active or passive approaches. Passive approaches do not emit any kind of radiation themselves, but instead rely on detecting reflected ambient radiation. Most scanners of this type detect visible light because it is a readily available ambient radiation. For example, stereo vision is widely studied in computer vision [25, 26, 27]. Two cameras take pictures of the same scene but from different views of point. A computer compares the images while shifting the two images together over top of each other to

find the parts that match. The shifted distance between the same object in between views is called the disparity, at which is used by the computer to calculate the distance from the object to the cameras [28]. Passive methods can be very cheap, because, in most cases, they do not need particular hardware. However, the accuracy, compared to active approaches, is low [29, 24]. As an example, we scanned a plush gorilla as shown in Fig. 1.3 (a). The 3-D reconstruction result using stereo vision technique (graph-cuts [30]) is shown in Fig. 1.3 (b). Due to the complicated scene of the plush gorilla, the quality of the reconstruction in Fig. 1.3 (b) is low.

Triangulation based laser scanner and structured light illumination are typically employed active triangulation approaches for 3-D surface measurement. In most cases of laser scanning, a laser stripe, instead of a single laser dot, is swept across the object to speed up the acquisition process. Thus, the scanning time is fairly long. To reduce the scanning time, a faster and more versatile method is the projection of patterns consisting of many stripes at once, or of arbitrary fringes, as this allows for the acquisition of a multitude of samples simultaneously. Seen from different viewpoints, the pattern appears geometrically distorted due to the surface shape of the object [31, 32, 33, 34]. In order to identify the stripes, the projected patterns are designed such that after capturing images, by the camera, the correspondences between the camera and the projecting device can be found [35]. And based on this idea, structured light illumination (SLI) is developed [36, 37, 38, 39].



Figure 1.4: SLI was used for jet pilot's helmets, from Dec. 1954 issue of life magazine.

1.1 Structured light illumination

SLI [40, 41] employs the projection of a group of designed light patterns onto the target's surface. By the reflected illumination of the target, the depth information can be computed. As one of the most accurate non-contact 3D surface measuring techniques, SLI has many industrial and scientific applications, including human and computer interfacing, biometrics, 3D conferencing, IC mounting on circuit board, biomedical, next-gen multimedia and entertainment, motion scanning and tracking, defect inspection and so on [32, 31, 42, 43, 35, 38, 44]. A cool application of SLI, reported from the Dec. 1954 issue of life magazine, is shown in Fig. 1.4 where SLI was used to measure the size and shape of a jet pilot's head for the purpose of

making custom fitting helmets. Compared to some passive 3D information acquisition techniques, such as stereo vision, SLI can easily overcome the correspondence problem between the two devices, which are the camera and the projector in the SLI case, via active scanning, and obtain accurate depth information through the camera-projector triangulation.

Reflection model

Generally speaking, the SLI technologies have two parts, codification and de-codification. In the codification part, all the SLI techniques employ a set, $\{I_n : n = 0, 1, \dots, N - 1\}$, of designed patterns, such that a point in the projector plane, (x^p, y^p) , is assigned with coding information Q .

After the set of illumination patterns have been projected upon a target object, an off-axis camera captures the scene such that the patterns are observed after being distorted by the surface topology under inspection. The resulting set of captured images, $\{I_n^c : n = 0, 1, \dots, N - 1\}$, can, therefore, be expressed as [45]

$$I_n^c = \alpha I_n + \alpha\beta, \tag{1.1}$$

where the two-dimensional camera coordinate (x^c, y^c) has been left out for brevity of notation for the image terms I_n^c , α , and β . In Eq. (1.1), α represents the albedo with $\alpha \in [0, 1]$ where 0 is pure black and 1 is pure white. $\alpha\beta$ represents the albedo image from ambient light with intensity β .

Thus, after projecting patterns from the projector and receiving the images $\{I_n^c\}$ in the camera, the de-codification algorithm tries to find out the coded information

Q in the original patterns $\{I_n\}$ with the present of albedo, ambient light and noise.

With the epipolar constraint which indicates that when the camera and the projector view a 3-D scene from two distinct positions, there are a number of geometric relations between the 3-D points and their projections onto the 2-D images that lead to constraints between the image points [46, 33], the coding information can be vertically or horizontally periodic depending on the system structure [47, 48]. The structured illumination patterns could be spots, stripes, or some other geometric patterns as long as the coding information can be extracted from the reflected illumination. Some of the recently studied or employed SLI pattern strategies are listed in Table 1.1.

Table 1.1: A Review of Recent Structured Light Patterns.

Technique Classification	Technique	Realtime 3-D reconstruction	Accuracy	Reliability	Implementation complexity
One-shot, spatial	De Bruijn sequences [49, 50, 51]	Y	Losing information inside stripes	Sensitive to object color	Low
One-shot, spatial	Adaptive structured light [52]	Y	Losing information inside stripes	High	Require predicting, labeling and tracking
One-shot, temporal	Composite pattern [45]	N	Reduced SNR in different channels	High	Require modulation and demodulation
One-shot, spatial	Color coded [53]	Y	Low pattern resolution	High	Low
One-shot, spatial	Laser ray matrix [54]	Y	Low pattern resolution	High	Low
One-shot, view-point-coding	Viewpoint-coding structured light [55]	Y	Increasing with camera number	High	Low
One/multi-shot, spatial	Color stripe [56, 57]	Y	Losing information inside stripes	High	Low
One/multi-pattern, temporal+spatial	Space-time stereo [58, 59]	N	High	Assumption on temporal continuity	Require good techniques in stereo matching
Multi-pattern, temporal+spatial	Motion compensation phase-shifting [60]	N	High	High	Require stereo matching and optimization
Multi-pattern, temporal+spatial	Stripe boundary coding [61, 62]	Y	High	Assumption on slow texture variation	Require segmentation, tracking and decoding
Multi-pattern, temporal	Trapezoidal phase-shifting [63]	Y	Error due to image de-focus	Introduces depth ambiguity	Low
Multi-pattern, temporal	Sinusoidal phase-shifting [29]	Y	High	Introduces depth ambiguity	Low
Multi-pattern temporal	Hierarchical structured light [64]	N	High	Introduces depth ambiguity	Low
Multi-pattern temporal	Two-step triangular-pattern [65]	N	High	Introduces depth ambiguity	Low

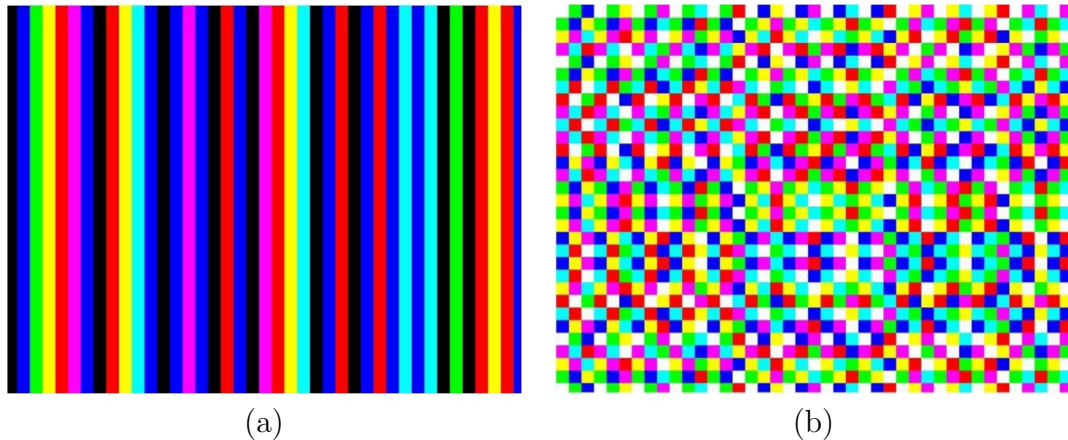


Figure 1.5: Patterns designed by Zhang and Chen.

One-shot

Generally, all the active 3-D measurement approaches can be classified into: 1) one pattern strategies, or “one-shot”, and 2) multi-pattern strategies, as we listed some of the latest developed or employed pattern strategies. Among all the “one-shot” techniques, De Bruijn sequences may be the most used [40, 53]. These spatial techniques define the neighborhood by using pseudo random sequences. But the decoding stage becomes more difficult as the spatial neighborhood cannot always be identified, and 3D errors can arise. Thus, they are sensitive to object color which is common among most one-shot strategies [40, 35]. The recent contribution from Zhang et al [49], as shown in Fig. 1.5 (a), achieves excellent performance by projecting patterns consisted of 125 vertical slits colored by using a De Bruijn sequence of third order and 8 colors (equal adjacent colors in the resulting sequence were eliminated) and inventing multi-pass dynamic programming [49]. Various other attempts have also been made to overcome the problems faced by one-shot strategies in order to present a improved robustness to the scanned scene [53, 52, 45, 54, 55, 56, 57]. Color patterns are typically

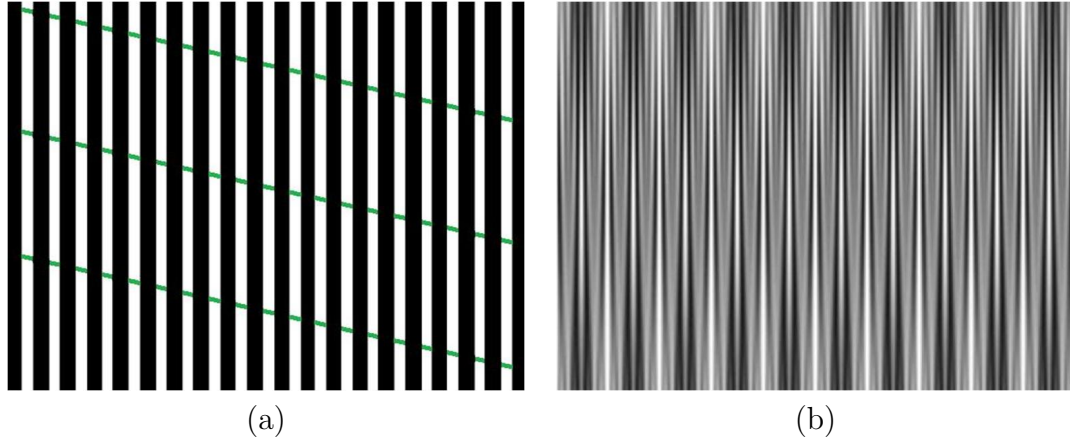


Figure 1.6: Patterns designed by Koninckx and Guan.

employed to increase the coding density [49, 53, 56, 57], e.g. Fig. 1.5 (b), but, the shape measurement result is affected, to various degrees, by the variations in the object’s surface color [66].

A real-time adaptive strategy, developed by Koninckx et al in [52], overcomes many practical difficulties. By generating better patterns in real time by taking the properties of the scene and setup into account, as shown in Fig. 1.6 (a), color coding lines are used to resolves ambiguities due to the periodic stripe pattern (the binary base stripes). The work is mainly based on binary fringe pattern, which loses information inside the binary stripes. And the system is a little complex because it requires predicting, labeling, and tracking scene features.

While most of the one-shot strategies are trying to find correspondence between camera and projector by spatially coding the neighborhood, Guan et al developed their composite pattern method, which efficiently combines multi-patterns into a single projection [45]. As shown in Fig. 1.6 (b), it successfully reduces sensitivity to ambiguity or albedo. However, the robustness is improved at a cost of accuracy due

to the reduced Signal to Noise Ratio (SNR) in each modulation channel.

An interesting approach has been proposed by Young et al [55]. By using a high frequency stripe pattern and placing multiple cameras in carefully selected locations where the epipolar projection in each camera can be made to mimic the binary encoding patterns normally projected over time. Thus, they achieve the 3-D reconstruction distinguishably from other techniques, but the accuracy of this approach is strongly related with the number of cameras.

Although the above innovations promise robust applications, limitations of these method still exist. One is the trade-off between reliability and accuracy. Since adjacent color/intensity stripes/dots should have enough spectral difference, people have to use a limited number of color/intensity stripes or apply them periodically which produces either stripe ambiguities or rough resolution [66, 53]. In general, the more patterns are used in a structured light system, the better accuracy that can be achieved. Based on the information from only one pattern, it is hard to achieve both high accuracy and reliability.

Multi-pattern

Compared to one-shot strategy, multi-pattern strategies are well known for their robustness to object color and measurement accuracy. Mostly, the multi-pattern strategies are based on temporal coding, where a set of patterns are successively projected onto the measuring surface. The codeword for a given pixel is usually formed by the sequence of illumination values for that pixel across the projected patterns. The bits of the codewords are multiplexed in time. This kinds of pattern

strategies can achieve high accuracy in their measurements due to the facts: first, as multiple patterns are projected, the codeword basis tends to be small (usually binary) and, therefore, a small set of primitives is used, which are easily distinguishable among each other. Moreover, a coarse-to-fine paradigm is followed, where the position of a pixel is encoded more precisely as the patterns are successively projected [40,35].

During the last 20 years, lots of techniques based on multi-patterns have appeared. Many of these techniques can be classified as: (i) techniques based on binary codes which project sequences of binary patterns in order to generate binary codewords [67,68], (ii) techniques based on n-ary codes which are a basis of n primitives used to generate the codewords [69], (iii) Phase shifting method (PSM) which involve projecting the same pattern, but shifting it in a certain direction in order to increase resolution [70,71,72,24]. Among these multi-patterns approaches, PSMs exploit higher spatial resolution as they project a periodic intensity pattern several times by shifting it in every projection. And PSMs achieve higher accuracy.

Typically, PSMs project a set of time-multiplexed patterns, $\{I_n^b : n = 0, 1, \dots, N - 1\}$, onto a target object such that an off-axis imaging sensor observes the scene and captures the wave patterns distorted by the surface topology under inspection. Generally, the patterns $\{I_n^b\}$ are designed as

$$I_n^b(x^j, y^j) = A + Bs(x^j, y^j), \quad (1.2)$$

where A is a temporal DC value, B is the amplitude (or projector modulation) value of a periodical signal function $s(x^j, y^j)$. The coordinate (x^j, y^j) is the corresponding coordinate in the projector. The captured images, $\{I_n^c : n = 0, 1, \dots, N - 1\}$, are then

denoted as

$$I_n^c(x^c, y^c) = \alpha(x^c, y^c)I_n^b(x^j, y^j) + \alpha(x^c, y^c)\beta(x^c, y^c), \quad (1.3)$$

where the superscript c indicates that I_n^c is now in the camera space and (x^c, y^c) is the two-dimensional camera coordinate [45]. In Eq. (1.3), $\alpha(x^c, y^c)$ represents the albedo with $\alpha(x^c, y^c) \in [0, 1]$ where 0 is pure black and 1 is pure white. The term, $\alpha(x^c, y^c)\beta(x^c, y^c)$, represents the albedo image from ambient illumination with intensity $\beta(x^c, y^c)$.

Through a decoding function and a phase unwrapping algorithm, the phase Φ that represents the coordinate of x^j or y^j , can be obtained from the “wrapped” (or coded) phase ϕ^w which is expressed as

$$\phi^w(x^c, y^c) = g \left[\frac{U(x^c, y^c)}{V(x^c, y^c)} \right], \quad (1.4)$$

where

$$U(x^c, y^c) = \sum_{n=0}^{N-1} a_n I_n^c(x^c, y^c), \quad (1.5)$$

and

$$V(x^c, y^c) = \sum_{n=0}^{N-1} b_n I_n^c(x^c, y^c). \quad (1.6)$$

The terms a_n and b_n are two coefficients in summations such that, in $U(x^c, y^c)$ and $V(x^c, y^c)$, the terms of $\beta(x^c, y^c)$ and the DC value A in images are canceled, while the division between the two summations cancels the terms of $\alpha(x^c, y^c)$. The term $g(\cdot)$ is a function that estimates the phase values $\phi^w(x^c, y^c) \in [0, 2\pi)$ out of image intensity values. As we can see, the PSMs are trying to solve the Eqs. (1.1) and

(1.4) and calculate the coding information. To do so, at least three patterns should be projected [45], since there are three unknown parameters, namely α , β and Q .

In summary, there are a variety of different designed light pattern strategies. Among these strategies, PSM, which project a pattern several times by shifting it spatially in every projection, are typically used [40, 35, 29, 70, 24]. In contrast to patterns based on binary code [67, 68] and n-array [73], PSM overcomes the discrete nature of patterns. And due to the fact that the pattern resolutions are exponentially increasing among the coarse-to-fine light projections and the fringe gap tends to 0, the resolution of PSM is greatly improved [40].

1.2 Phase shifting methods

Following, we will review and discuss some of the latest developed and used PSM strategies. In order to simplify the discussion, the patterns and techniques are presented in a typical 8-bits gray scale projector with pattern height of 1024. And all the patterns are unit-frequency.

Phase measuring profilometry

Among the many proposed SLI methods, the technique of Phase Measuring Profilometry (PMP) is one of the most widely used and precise strategies [70, 74, 75, 76, 77]. The canonical PMP technique employs a set, $\{I_n : n = 0, 1, \dots, N - 1\}$, of sinusoidal wave patterns such that at the point (x^p, y^p) , in projector space, the intensity value is assigned as:

$$I_n(x^p, y^p) = A + B \cos(\Phi(y^p) - \frac{2\pi}{N}n), \quad (1.7)$$

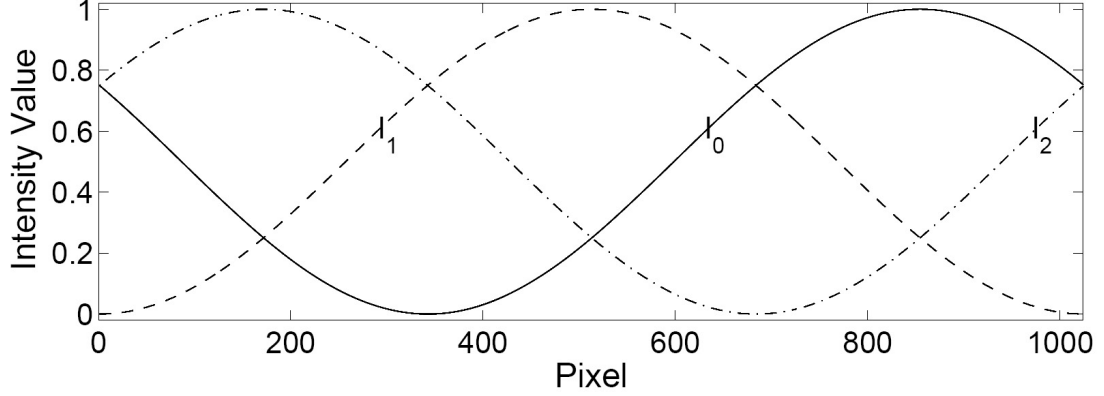


Figure 1.7: Cross sections of 3-pattern unit frequency PMP strategy.

where A is the temporal DC value, which is normally 0.5, and B is the amplitude of the temporal AC signal, which is also normally 0.5. Thus, the sinusoidal signal covers the entire dynamic range of the projector $[0, 1]$. The term $\Phi(y^p)$ is the phase information and is designed according to:

$$\Phi(y^p) = \frac{2\pi F y^p}{H}, \quad (1.8)$$

where H is the pattern height (number of points in pattern) and F is the number of periods. Note the dependence of the phase term, $\Phi(y^p)$, on y^p as this is the parameter that will be used when triangulating with the camera, which is assumed to be positioned vertically above/below the projector. Thus, we denote $I_n(x^p, y^p)$ as $I_n(y^p)$.

After projecting the patterns, the reflection process can be expressed as:

$$I_n^c(x^c, y^c) = \alpha(x^c, y^c) \left[A + B \cos\left(\Phi(y^p) - \frac{2\pi}{N}\right) + \beta(x^c, y^c) \right], \quad (1.9)$$

where (x^c, y^c) is the two-dimensional camera coordinate, $\alpha(x^c, y^c)$ represents the albedo within $[0, 1]$ where 0 is pure black and 1 is pure white, and $\alpha(x^c, y^c)\beta(x^c, y^c)$ is the albedo image from ambient illumination with intensity $\beta(x^c, y^c)$. Thus in the



Figure 1.8: 3-D reconstruction using PMP. Result is shown with depth rendering.

process of reflection, the albedo and ambient illumination effects are added into the patterns. The phase information $\Phi(y^p)$ is then obtained from a decoding function expressed as:

$$\begin{aligned}
 \phi(x^c, y^c) &= \arctan \left[\frac{\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \sin(\frac{2\pi n}{N})}{\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \cos(\frac{2\pi n}{N})} \right] \\
 &= \arctan \left[\frac{\alpha(x^c, y^c) B \sin(\Phi(y^p))}{\alpha(x^c, y^c) B \cos(\Phi(y^p))} \right].
 \end{aligned} \tag{1.10}$$

The decoding function removes the effects of albedo and ambient illumination such that the designed phase information is recovered. Generally, if a high frequency signal is employed, the obtained phase ϕ , from Eq. (1.10), needs to be further “unwrapped” into unit frequency Φ in order to find the unique correspondence over the full resolution. Once $\Phi(x^c, y^c)$ is obtained, the 3-D world coordinates of a point can be calculated from $(x^c, y^c, \Phi(x^c, y^c))$ in a pre-calibrated system.

To solve Eq. (1.10), at least 3 patterns should be projected and captured since there are $\Phi(y^p)$, $\alpha(x^c, y^c)$, and $\beta(x^c, y^c)$ unknowns in Eq. (2.3). The cross sections of

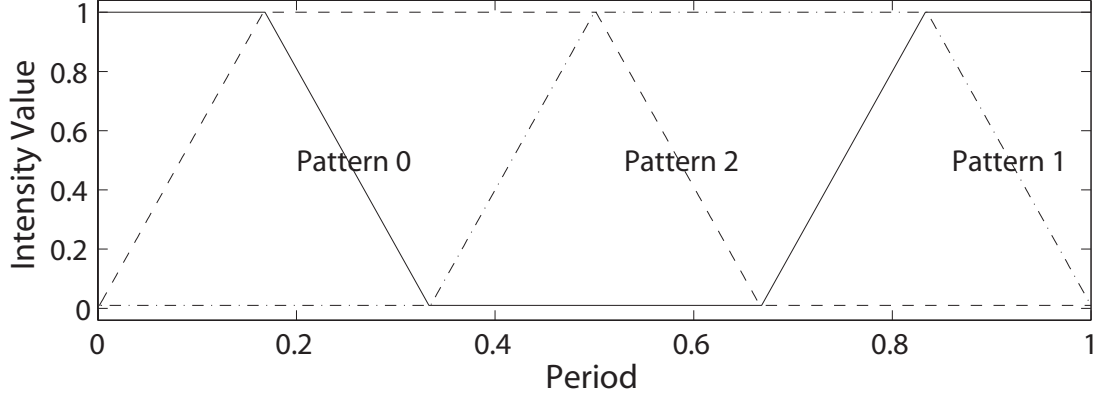


Figure 1.9: The trapezoidal pattern strategy designed by Huang.

three pattern PMP are shown in Fig. 2.1. As the most widely used pattern strategy, PMP is easy to implement and well developed for $N > 3$ patterns. The continuous smooth sinusoidal patterns are insensitive to blurring caused by an out of focus lens [78, 29]. An example of 3-D reconstruction of the plush gorilla (Fig. 1.3 (a)) is shown in Fig. 1.8. Compared to the result using stereo vision (Fig. 1.3 (b)), the 3-D reconstruction from PMP is greatly improved.

Trapezoidal pattern

The Trapezoidal Pattern (TPP) is proposed by Huang et al [63]. This method uses three patterns coded with phase shift trapezoidal shaped gray levels [63], as shown in Fig. 1.9. For each point (x^c, y^c) of captured images, the coding information Q is obtained by using the following equation:

$$Q = 2\text{round}\left(\frac{K-1}{2}\right) + (-1)^{K+1} \frac{I_{med}^c - I_{min}^c}{I_{max}^c - I_{min}^c}, \quad (1.11)$$

where I_{max}^c , I_{med}^c and I_{min}^c are respectively the maximum, median and minimum intensities of the three received images, and $K = 1, 2, \dots, 6$ is the region number

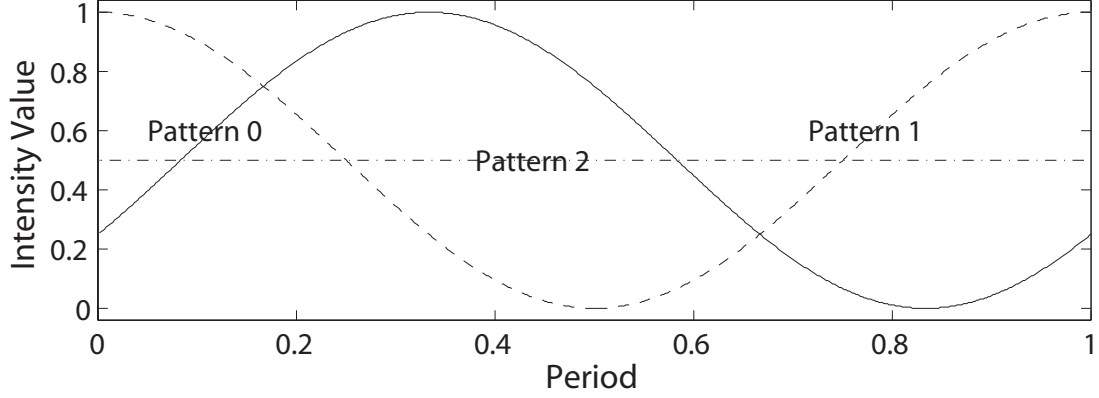


Figure 1.10: The 2 + 1 pattern strategy designed by Wizinowich.

determined by comparing the three intensities [63]. The value of Q ranges from 0 to 6. The 3-D coordinate is obtained similarly as PMP.

As claimed by Huang et al, the TPP is 6 time better than three-pattern PMP. But due to the fact that the camera devices tend to integrate over a certain area such that pixel values are affected by its neighbors [78], the the inflection points of patterns are not as sharp as theoretical results. However, because of the intensity ratio based de-codification (Eq. (1.11)), the coding information Q is not sensitive to the blurring effect.

2+1 method

The 2 + 1 phase shift patterns employ two sinusoidal patterns with a phase shift of $\frac{\pi}{2}$ and a uniform flat pattern [79], as shown in Fig. 1.10. The coding information is calculated through

$$Q = \text{atan} \left[\frac{I_0^c - I_2^c}{I_1^c - I_2^c} \right], \quad (1.12)$$

where the value of Q ranges from 0 to 2π . The 3-D coordinate is obtained similarly as PMP. This method, consisting totally three patterns, is proposed by Wizinowich

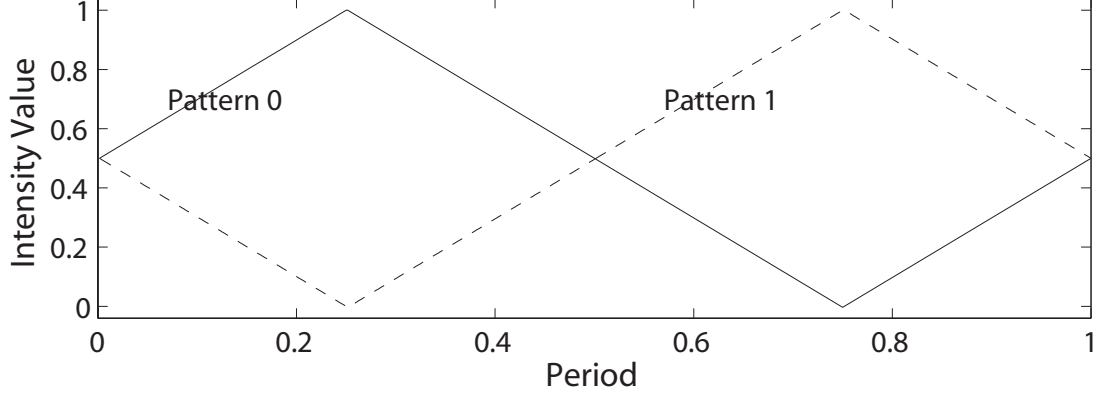


Figure 1.11: The triangular pattern strategy designed by Jia.

in [80] and lately developed by Zhang and Yau in [79]. In contrast to PMP, the method developed by Zhang and Yau is more sensitive to noise due to less number of data frames. However, it is more robust to motion and vibration, i.e. the faster geometric shape changes [79, 80].

Triangular pattern

Different from above strategies, the phase shift triangular pattern (TGP) strategy consists only two patterns [81, 65], as shown in Fig. 1.11. After receiving the patterns from camera, the Q is obtained as

$$Q = 2\text{round}\left(\frac{K-1}{2}\right) + (-1)^{K+1} \frac{|I_0^c - I_1^c|}{I_m}, \quad (1.13)$$

where Q ranges from 0 to 4, and $K = 1, 2, 3, 4$ is the region number determined by analyzing a small neighborhood of each point. I_m is the modulation intensity computed as

$$I_m = I_{max} - I_{min}, \quad (1.14)$$

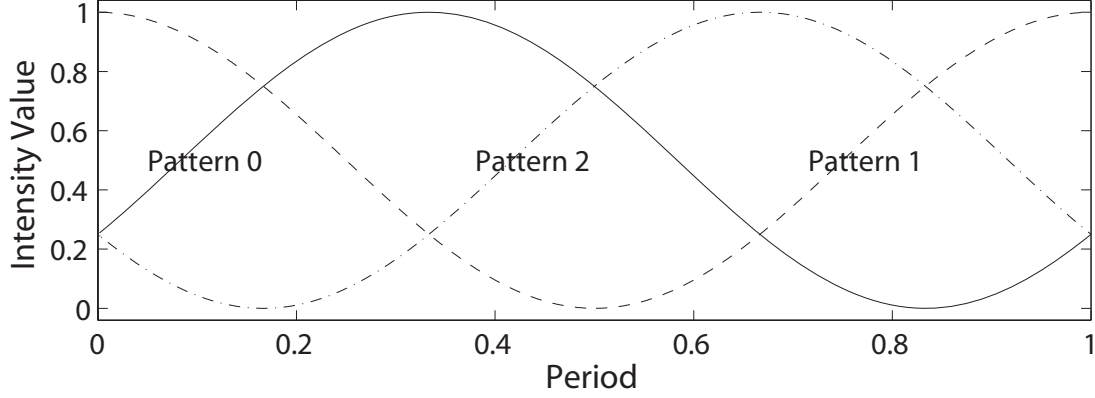


Figure 1.12: The fast three step pattern strategy designed by Huang.

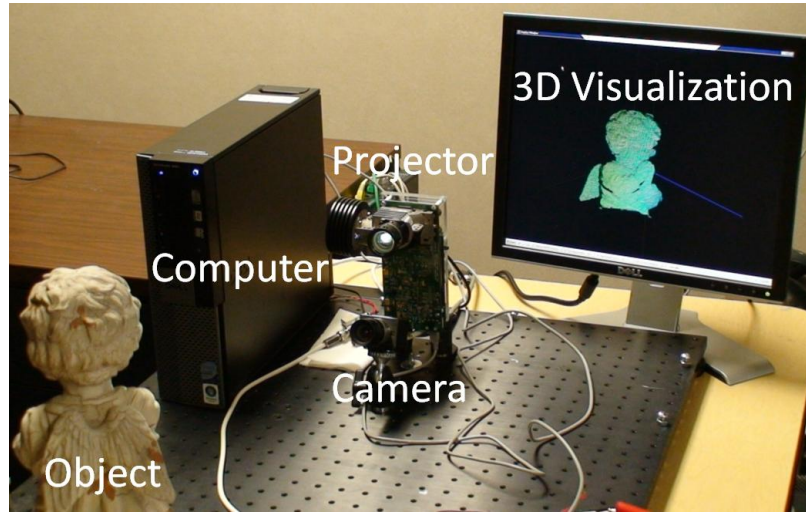


Figure 1.13: An SLI system developed by Liu and Wang.

where I_{max} and I_{min} are the maximum and minimum intensities of projected patterns I_0 and I_1 . Thus, instead of from the received images, I_m is actually obtained from the projected patterns. This makes the TGP sensitive to the albedo of scanned object, since in Eq. (1.14) $I_0^c - I_1^c$ contains the information of albedo (see in Eq. (1.1)) but I_m does not. In order to make TGP work under the presence of non-uniform albedo and ambient light, one more pattern should be projected.

There are some other PSMs which are similar as or developed from the above

strategies. For example, the fast three step method, as shown in Fig. 1.12, used by Huang et al [82] projects the sinusoidal phase shift pattern as PMP. But instead of using Eq. (1.10), it uses Eq. (1.11) for de-codification. And in [78], Gühning proposed a new method called line shifting which combines binary and sinusoidal phase shift pattern. This combination overcomes some problems in PSM, i.e. reconstructing of sharp change albedo region, but it results a large number of patterns (up to 32). An SLI system developed by Liu and Wang is shown in Fig. 1.13.

1.3 Problems in phase shifting methods

Although the PSMs are accurate and robust to various scenes. In practice, due to the system noise, the non-linear response of devices (gamma) [83], and other factors, in the PSMs, there are still problems that needs to be solved.

Pattern efficiency

In PSMs, due to the exist of noise, higher accuracy of the 3-D reconstructions can be achieved when more patterns are projected out. However, for different pattern strategies, using the same number of patterns, the achieved accuracy is not the same. In other words, the pattern efficiency is different. In general, PSMs calculate the phase, which can be regarded as a corresponding map between projector and camera, temporally for each valuable pixel in the camera, and have the advantages of less number of patterns and fast processing speeds because the calculation of the temporal intensities is rather simple. And no point matching or image enhancement is required [63, 74]. They are also resistant to target surface albedo variations and am-

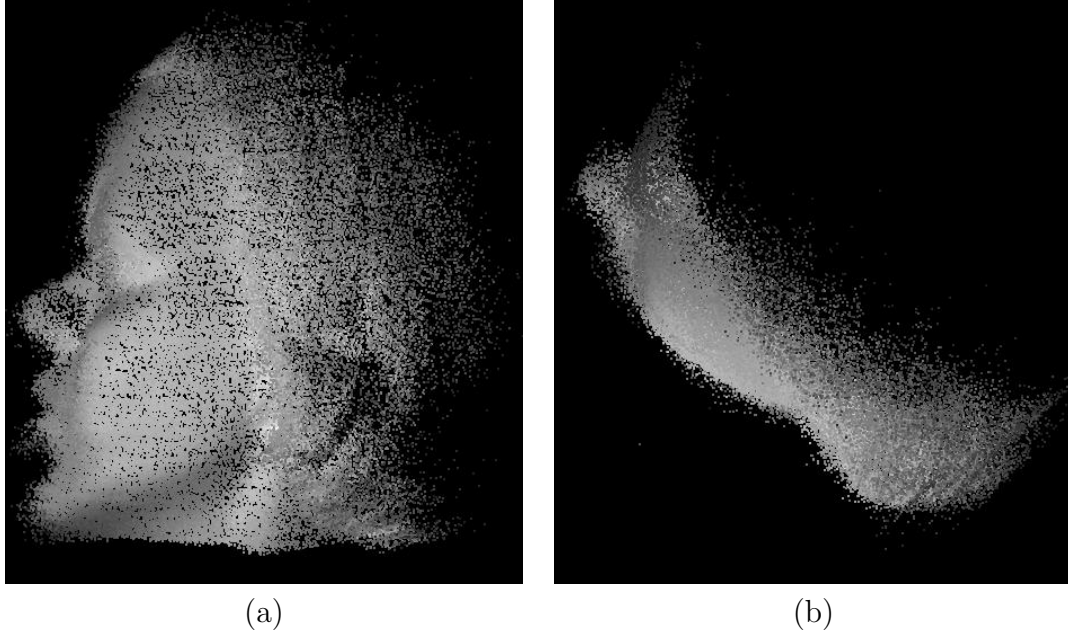


Figure 1.14: 3-D reconstruction of a white model, Alice, using 3-pattern unit frequency PMP where the error caused by noise is shown. (a) Front view of the reconstruction. (b) Top view of the reconstruction.

bient light. However in practice, error is introduced from camera noise [74], projector noise [45, 84], and ambient light noise [45] as well as other sources [85, 86, 87]. And these methods show large impact of noise [63]. On the other hand in order to scan dynamic scenes with moving objects, researchers [63, 58, 64, 60, 29] have employed high speed components to build real-time acquisition systems that attempted, as much as they could, to limit the number of projected patterns in order to minimize total scan time. Here, we first employed 3-pattern unit frequency PMP, which is the minimum number of patterns for PMP, to scan a white model, Alice. The 3-D reconstruction result is shown in Fig. 1.14 where the noisy reconstruction is due to the noise. Further, 5-pattern unit frequency PMP was employed and the result is shown in Fig. 1.15. As shown, with two more patterns, the improvement of 3-D reconstruction, which is only 22.54% reduction of the standard deviation of reconstruction error, is not obvious.

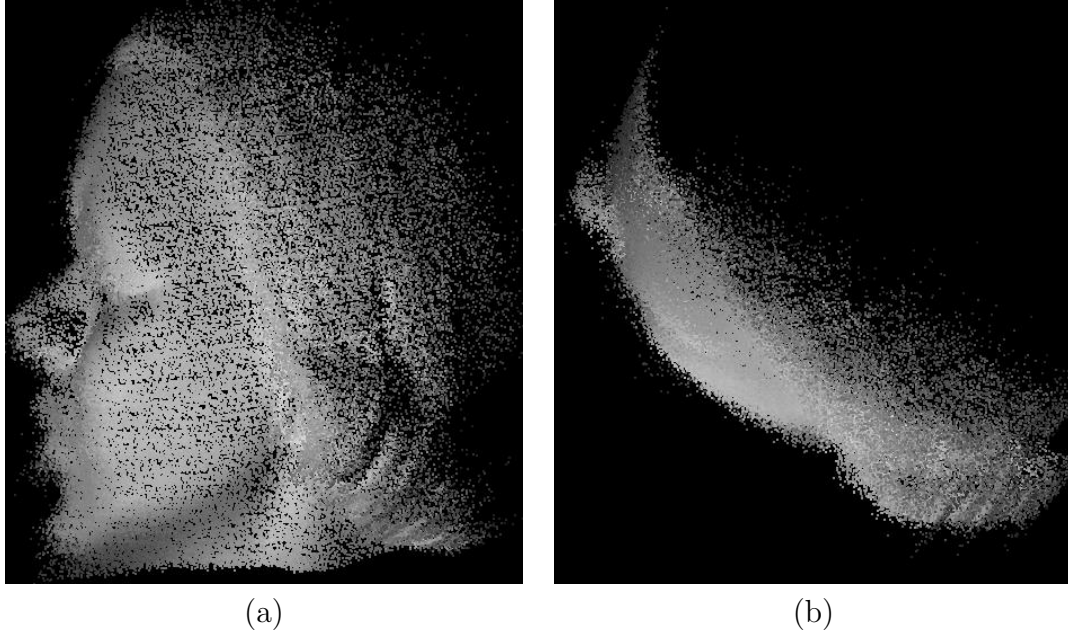


Figure 1.15: 3-D reconstruction of a white model, Alice, using 5-pattern unit frequency PMP where the error caused by noise is shown. (a) Front view of the reconstruction. (b) Top view of the reconstruction.

It is essential to study the pattern efficiency of the PSM's patterns such that the best measurement accuracy can be achieved with a certain number of patterns. With the noise in practical systems, it is desirable to define a parameter that measures the Signal to Noise Ratio (SNR) of the PSM patterns. And further by maximizing the SNR , the patterns can be optimized such that with a certain number of patterns N the noise in 3-D results can be minimized. In [69], Horn and Kiryati proposed a method for the optimal design of n-ary patterns. The design of N projection patterns with pattern height H turns out to be equivalent to the placement of H points in a N dimensional space. Thus, the authors proposed the use of space filling curves, such as Hilbert curves, for defining the codewords to maximize SNR . But, the principles and constraints of n-ary techniques are different from those of PSM such that Horn and Kiryati's optimization can not be applied to PSM, and little work in optimization

of PSM has been done by now.

Depth ambiguity

For commonly employed PSMs [40], the reconstruction is derived from the observed phase distortions caused by the wrapping of the projected stripes around an object. Because PSMs involve a series of time-multiplexed patterns, it is not commonly associated with scanning dynamic scenes with moving objects. Of the PSMs, video-rate surface reconstructions are possible using high speed imaging equipment [63, 58, 64, 60, 29, 88]. Among these high speed systems, since PSMs are sensitive to noise, high frequency pattern strategies are popular approaches in order to improve the accuracy of measurement.

For PSMs, measurement accuracy can be improved with an increase in the number of light pattern periods projected per frame of video. However in order to build a full and one-to-one correspondence map between projected patterns and captured images, captured high frequency phase data has to be unwrapped into a single period [29]. The depth ambiguities that this process entails represents a major obstacle [89] as illustrated in Fig. 1.16 where Fig. 1.16 (a) shows two isolated surfaces where no method of spatial unwrapping can uniquely determine the period numbers in the captured image. Figure 1.16 (b) shows a depth discontinuity, parallel to the stripes, resulting in a continuity artifact in phase such that two different period numbers appear as the same one in the captured image.

Various methods have been investigated to enhance the reliability of phase unwrapping through the inclusion of extraneous spatial information such as through

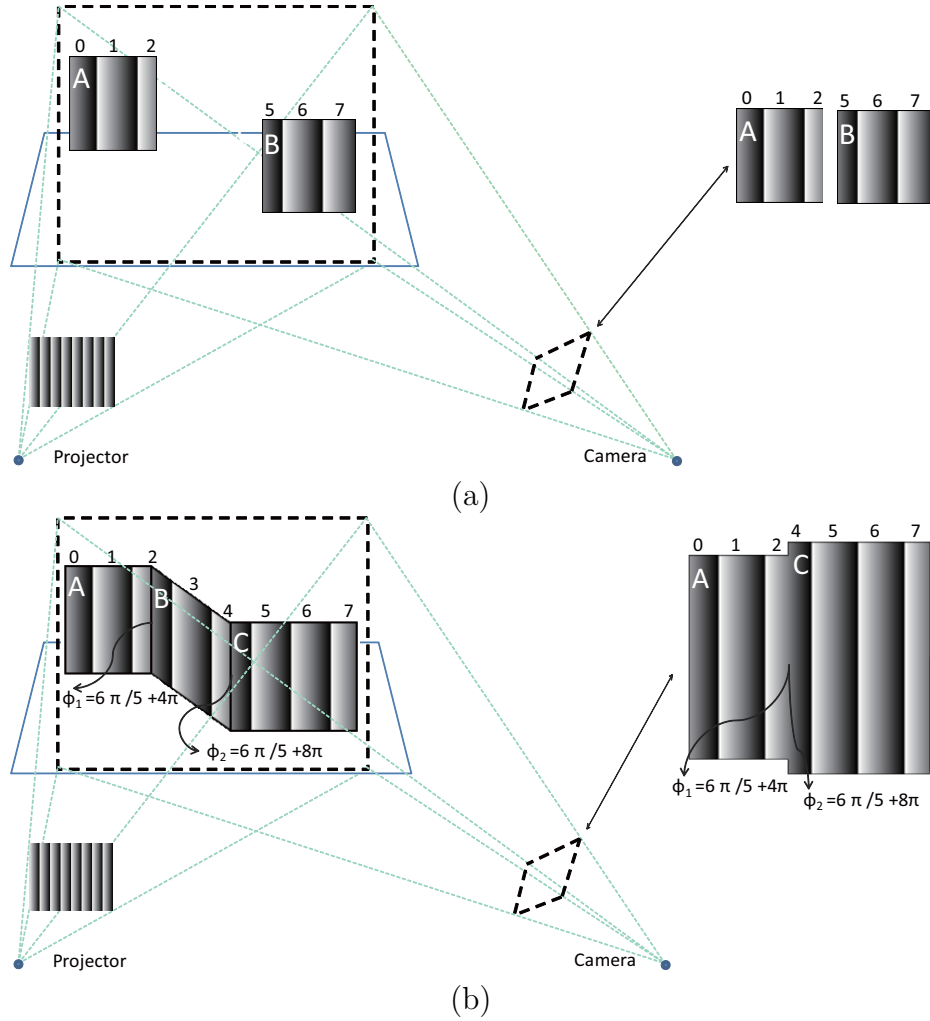


Figure 1.16: Illustration of depth ambiguities in high-frequency phase shifting SLI showing the scanned geometries and the captured images for (a) disconnected surfaces and (b) a surface step-edge.

branch cut [89], discontinuity minimization [90], agglomerative-clustering [91], or least-squares [92]. But according to Saldner and Huntley [93], these approaches alone cannot solve the difficulties presented in Fig. 1.16. On the other hand, temporal unwrapping approaches [93,94] have been proposed to achieve reliable unwrapping, but the use of extra patterns to remove the depth ambiguities is, again, unappealing for real-time acquisition.

Thus, phase information is the only signal coded into the patterns. The pixels of

the projector, having the same phase but different periods, project the same intensity values. In order to unwrap high frequency, either the algorithms need to “count” the period number from one direction to the other using a spatial approach, or extra patterns need to be projected to obtain a reference unit frequency signal (or period cue). Thus although the unwrapping problem has been studied for over 20 years [95], solutions are not satisfying in a real-time system. In fast SLI applications, spatial unwrapping methods are preferred because they use fewer patterns, but as reported by Zhang [96] for the phase unwrapping of exaggerated and difficult facial expressions, the failure rate of the scan-line algorithm was more than 86%.

Gamma, multi-light-path and others

Compared to passive methods like stereo-vision, SLI is known for its robustness and high accuracy [40, 33, 35, 34]. On the other hand, as SLI involves a scanning process with a series of time-multiplexed patterns, its application is limited to moving targets, where the standard solution is to drive the camera/projector pair at such high frame rates that the amount of object motion is small over the pattern set. Hence, there is general need to reduce the overall pattern number for these applications.

Among efforts to achieve high speed and high accuracy 3-D data acquisition, temporal multiplexed projection [29, 97] and color multiplexed [49, 53] sequences have been extensively studied to reduce the number of illumination patterns. However, the reduced number of illumination patterns tends to yield poor quality 3-D reconstructions [35, 29, 45]. Problems associated with SLI that are particular acute when using a low number of patterns include:

1. Increased sensitivity of phase data to noise and distortions in areas of high contrasted texture or depth [49, 52, 98].
2. Phase distortion in the presence of multi-light-path where the phase of one point is disturbed by the light reflected from other points on the target surface [99].
3. The sensitivity of the phase reconstruction in the presence of camera, projector and ambient light noise [74].
4. Phase distortion caused by non-linear response curves (gamma) in the projector and/or camera [84, 83].
5. Phase distortion caused by projector flicker associated with out-of-phase projector/camera pairs [45].

Without increasing the pattern number, some of these issues can be addressed through simple camera/projector calibration [84, 83] and hardware synchronization [29], but the requirement of a hardware connection may complicate the use of commodity hardware that lacks a synchronization signal interface. The fringe pattern in Fig. 1.8 is due to the gamma error. Another example of gamma distortion is shown in Fig. 1.17.

1.4 Contributions

In general, all PSMs are well known for their high resolution and accuracy when many patterns are projected. However, in order to achieve high speed, the number of patterns should be limited to as few patterns as is necessary, but limiting the number of patterns makes PSM methods especially susceptible to reconstruction errors from

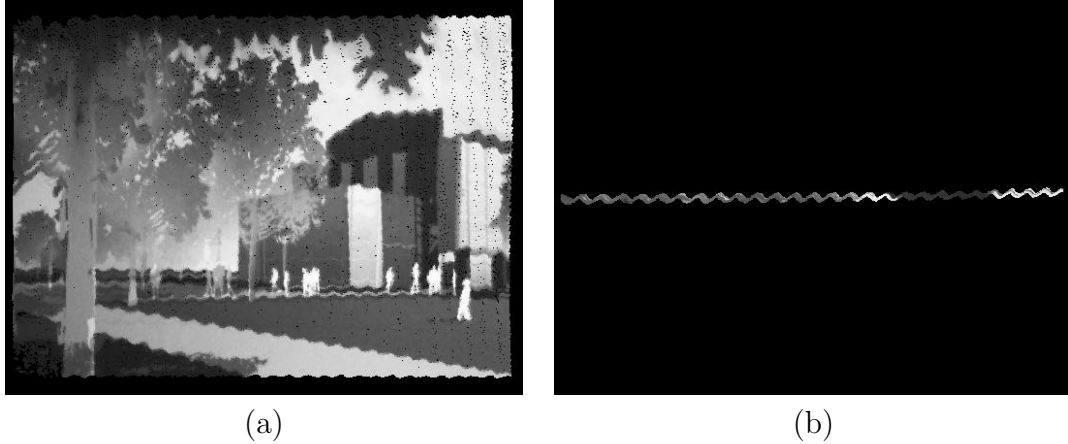


Figure 1.17: 3-D reconstruction of a flat textured board where the error caused by gamma and noise is shown. (a) Front view of the reconstruction. (b) Top view of the reconstruction.

sensor noise. In this dissertation, we study the noise in PSM systems and focus on the improvement of pattern efficiency in PSMs such that we maximize the achievable reconstruction accuracy using fewer patterns. In order to achieve this goal, in Chapter II, we take one of the most used strategy, PMP, as an example and illustrate the maximization of pattern SNR such that the pattern efficiency can be improved. By regarding the design of an N pattern strategy as placing points in an N -dimensional coding space, we define the computational length in the N -dimensional space and further mathematically derive the SNR which is related to the computational length and number of periods. Then, without introducing ambiguity of high frequency, we propose the edge-pattern strategy which maximizes the computational length and number of periods for different numbers of patterns (≥ 3).

The SNR improvements from edge-pattern to PMP increase with the number of patterns, which indicates the lower pattern efficiency of PMP with more number of patterns. Experimental results further demonstrate the high SNR of proposed

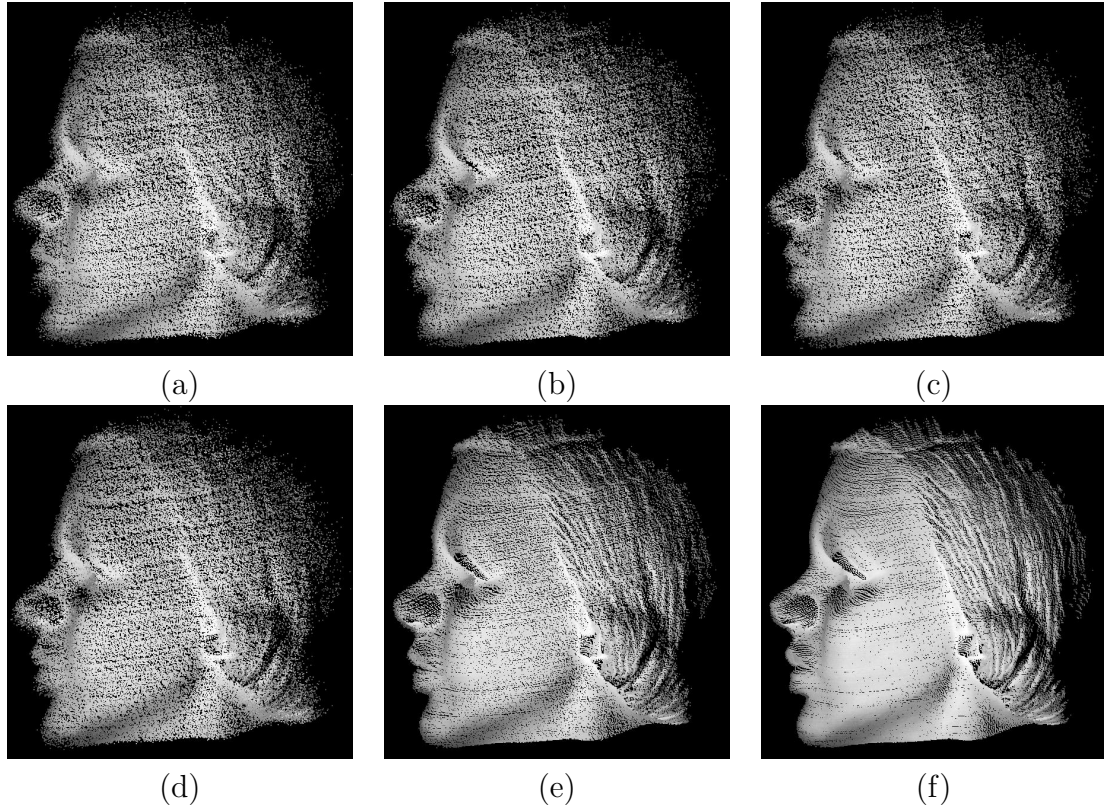


Figure 1.18: 3D reconstructions of Alice. Results are shown with depth rendering. (a) Reconstruction using traditional 3 pattern PMP. (b) Reconstruction using traditional 4 pattern PMP. (c) Reconstruction using traditional 5 pattern PMP. (d) Reconstruction using 3 pattern E-P. (e) Reconstruction using 4 pattern E-P. (f) Reconstruction using 5 pattern E-P.

pattern such that more accurate 3-D results can be achieved using less number of patterns. An example of 3-D reconstructions using the proposed edge-pattern strategy and traditional PMP with the same number of patterns is shown in Fig. 1.18. Specifically, for 3 pattern strategies, the proposed pattern strategy improves the SNR by 1.2362 times, whereas the improvement is as high as 13.6027 times for 5 pattern strategies. Detailed analysis and discussion of the edge-pattern strategy will be presented in Chapter II.

Now another way that we can reduce the effects of system noise on PSM recon-

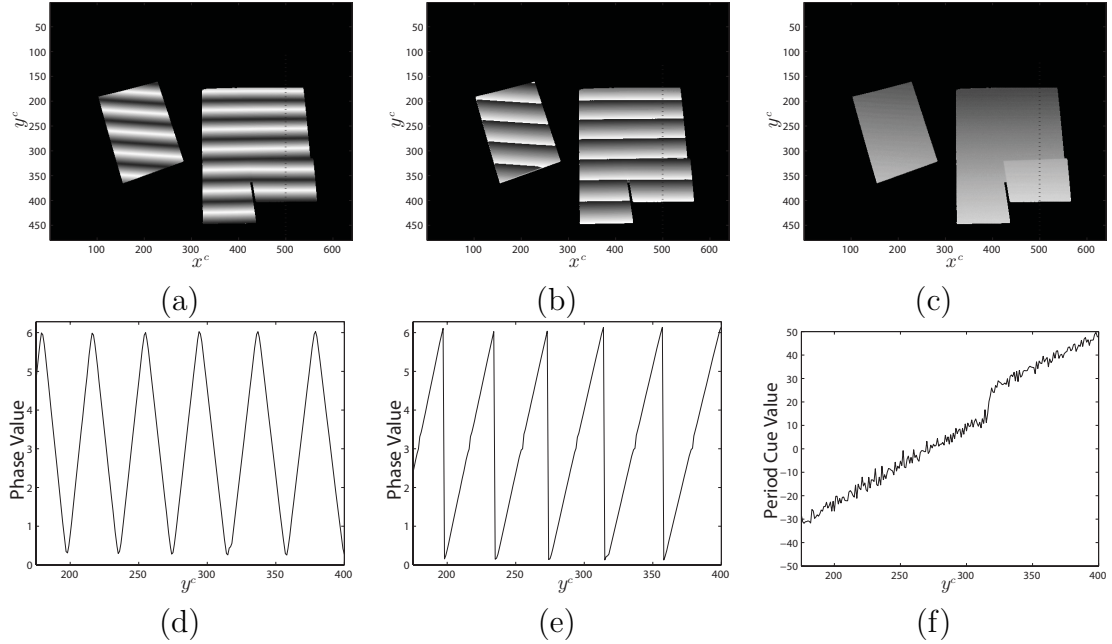


Figure 1.19: Three white boards were scanned using PCPS. (a) The wrapped phase image. (b) The phase image after initial decoding. (c) The period cue image. (d) The phase values on the 500th column of (a). (e) The phase values on the 500th column of (b). (f) The phase values on the 500th column of (c).

structions is to increase the frequency of patterns, but this introduces phase ambiguities, as illustrated in Fig. 1.16. So in order to eliminate these ambiguities, Chapter III defines a spatial intensity efficiency measure that, for those pattern sets with a spatial intensity efficiency of less than 100%, without reducing the *SNR* of original high frequency signal, we encode the period cue into the projected pattern set such that each period, of the multi-period pattern, is uniquely identifiable. This chapter further introduces pattern entropy as a separate measure from which one can optimize the embedded period cue. In combination, the proposed method, period coded phase shifting, can unwrap high frequency phase information and achieve high measurement precision without increasing the number of projected patterns and, therefore, has significant benefits when scanning moving objects. This chapter, further, introduces a

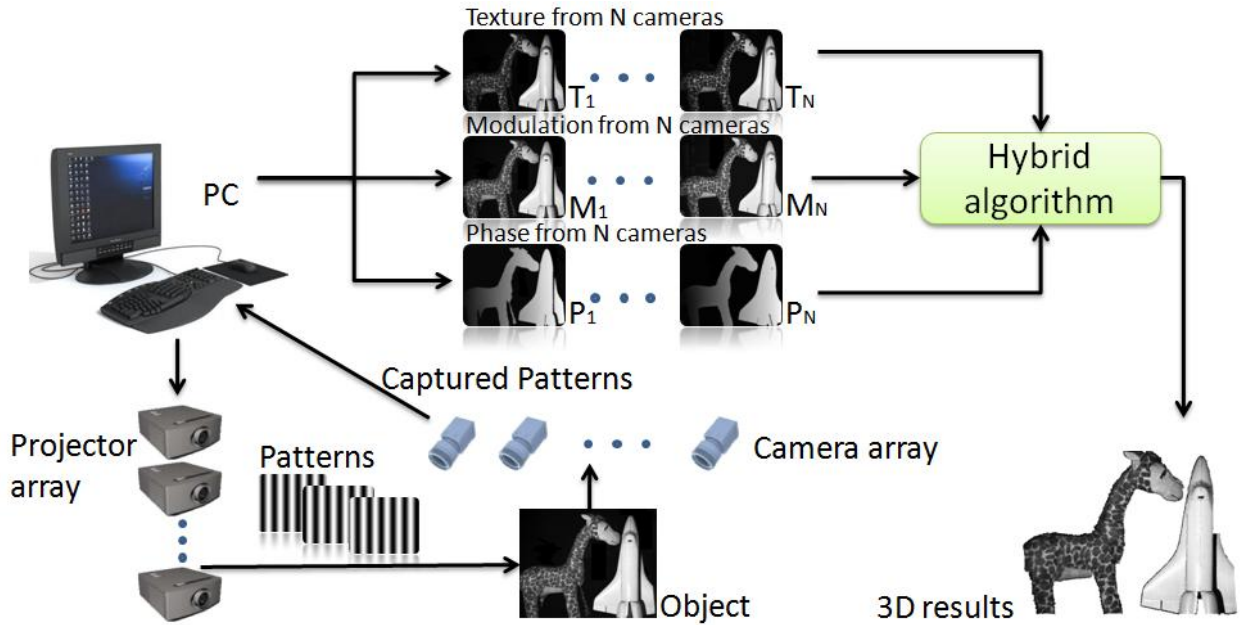


Figure 1.20: Schematic diagram of a multi-camera and multi-projector 3-D measurement system.

noise model, verified by experiment, to demonstrate the high measurement accuracy of the proposed approach.

Finally, a prototype system is demonstrated that can achieve 120 fps at 640×480 resolution for 3-D data acquisition and reconstruction. An example is shown in Fig. 1.19, where three separate white foam boards were carefully placed and scanned such that the left-most board was isolated from two boards otherwise positioned to create a phase ambiguity. Figs. 1.19 (c) and (f) show that the embedded unit frequency signal is successfully obtained such that the proposed approach in Chapter III works correctly in situations like Fig. 1.16 where spatial unwrapping fails [93] and temporal approaches only work by projecting additional patterns [96]. Detailed analysis and discussion of the period coded phase shifting method will be presented in Chapter III.

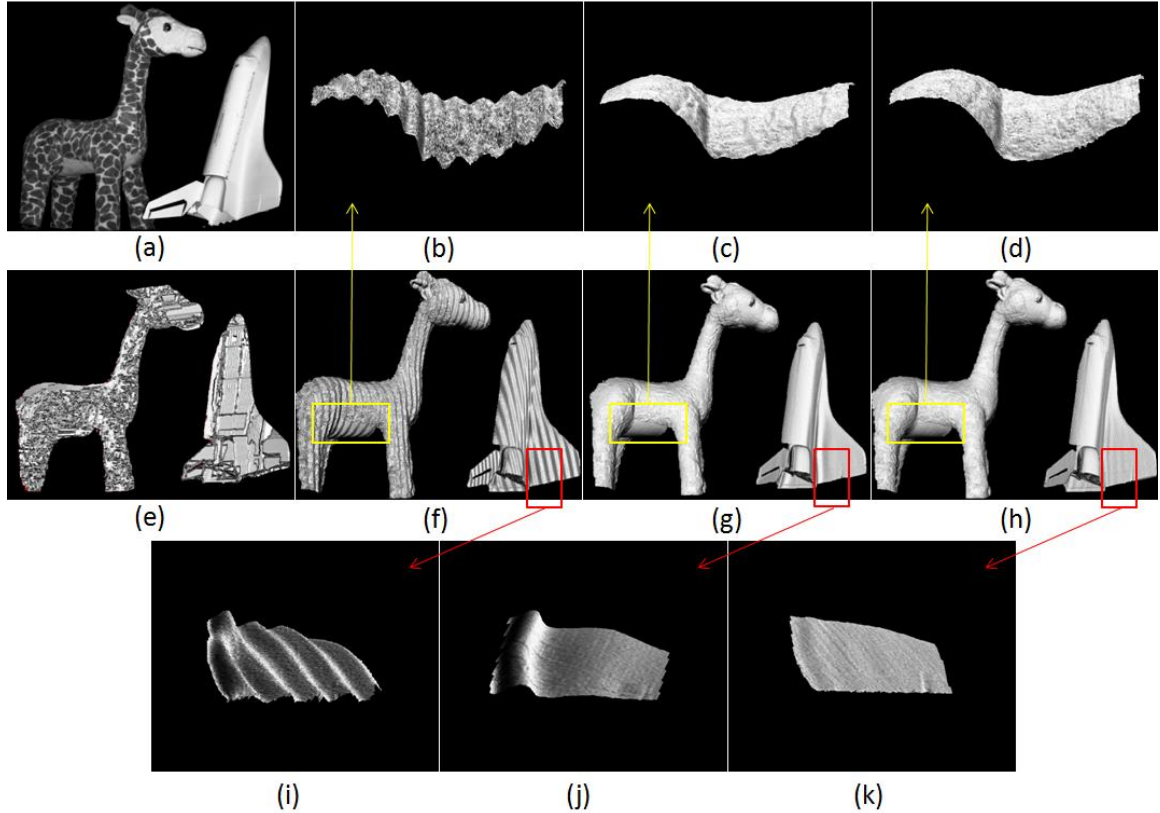


Figure 1.21: (a) A textured cartoon giraffe and a white space shuttle model. (b), (c) and (d) are zoomed-in and rotated views of (f), (g) and (h), respectively. (e) Depth rendering reconstruction by using stereo vision technique (with the graph-cut algorithm for 7 iterations). (f) Depth rendering reconstruction by using SLI technique (with 3 high frequency PMP patterns). (g) Depth rendering reconstruction by using SLI technique (with 30 high frequency PMP patterns). (h) Depth rendering reconstruction by using the hybrid approach (with 3 high frequency PMP patterns). (i), (j) and (k) are zoomed-in and rotated views of (f), (g) and (h).

In order to improve SNR and also reduce the reconstruction error cost by gamma, In Chapter IV, I proposed a hybrid approach for high accurate 3-D reconstruction which combines the SLI and stereo vision techniques. Generally, SLI is a method of range sensing involving the projection of a series of light patterns and allowing a camera to extract range data based on pattern distortions; whereas, stereo-vision reveals depth through building correspondences between textured images recorded simultaneously by two or more cameras. By employing multiple cameras, structured



Figure 1.22: 3-D reconstruction using hybrid approach. Result is shown with depth rendering.

light illumination can generate pattern phase/modulation and object texture data. The hybrid 3-D reconstruction framework achieves the 3-D reconstruction through phase-modulation-texture data fusion under temporal multiplexed illumination. The proposed scheme consists of, first, using phase data to derive initial correspondences across cameras. Second, texture data is used to eliminate correspondence ambiguities. Third, modulation data is used to estimate correspondence error ranges. Finally, Kullback-Leibler divergence refinement, based on the derived phase error models, is performed to reduce mis-registration among images. Using only a small number of light patterns, the presented approach significantly reduces measurement errors versus traditional structured-light methodologies while being insensitive to gamma distortion and projector flicker.

Experimental results demonstrate these advantages in terms of enhanced 3-D reconstruction performance and robustness against noise, distortions, and conditions

of texture and contrast. An example of 3-D reconstructions using the proposed hybrid approach is shown in Fig. 4.22 which demonstrates the improved performance of the hybrid approach in high and low textured environments and in multi-light-path regions. The plush gorilla of Fig. 1.3 (a) was also reconstructed using the hybrid approach. The result is shown in Fig. 1.22. Compared to the reconstructions using stereo vision (Fig. 1.3 (b)) and PMP (Fig. 1.8), the result from hybrid approach is greatly improved. Detailed analysis and discussion of the hybrid approach will be presented in Chapter IV.

At last in Chapter V, the conclusion and future work are given. For the edge-pattern strategy proposed in Chapter II, theoretically, the SNR is improved by 1.2381 times when using three component patterns and 15.5421 times when using five patterns. In Chapter III, the PCPS embeds the second signal such that PCPS can successfully unwrap high frequency phase information and achieve high measurement precision without increasing the number of projected patterns and, therefore, has significant benefits when scanning moving objects. The hybrid 3-D framework developed in Chapter IV combines the SLI and stereo vision techniques such that, after performing the proposed Kullback-Leibler divergence refinement, the root-mean-squared (RMS) registration error is $1/26^{th}$ that of conventional SLI. In Chapter V, I also list the future work of edge-pattern, PCPS, and Hybrid approach and propose a PSM pattern strategy with only 2 or even 1 pattern.

In final summary, minimizing the number of component patterns, without increasing the effects of noise, is a chief goal for structured light illumination system design. As such, this dissertation focused on the efficiency of phase shifting pat-

terns where traditional pattern strategies were divided into two parts: codification and de-codification. For codification, a theoretical definition of pattern efficiency was proposed where it was shown that the efficiencies of most patterns of PSMs do not reach 100%. Thus, it becomes possible to employ the remaining dynamic range of the projector to develop better pattern strategies. An example is the edge-pattern developed in chapter II where I demonstrated that by only changing the codification, could the SNR be improved. Another example to employ the remaining dynamic range for codification was proposed in chapter III where a period coded codification and its specified de-codification were designed such that two kinds of signals could be extracted from one projected group of patterns. Other functions may also be achieved through different assembly of codifications and de-codifications. An alternative approach to improve the efficiency of the pattern was proposed in chapter IV. Furthermore by introducing extra devices into an SLI system, more information becomes available, which can be employed to improve the 3-D reconstruction without projecting more patterns. In the future, the efficiency of a particular device should also be analyzed. The problem of how much information can be brought by employing extra devices and how to use this extra information should be studied.

Chapter 2 MAXIMIZATION SNR PATTERN STRATEGY

After over 20 years development, many PSMs have been explored and studied. As we reviewed in Chapter I, the most employed PSMs include: Phase Measuring Profilometry (PMP) [70], trapezoidal pattern strategy [63], 2 + 1 pattern strategy [79], and triangular pattern strategy [81]. For these PSMs, trapezoidal and 2 + 1 pattern strategies project 3 patterns, triangular pattern strategies projects 2 patterns while sacrificing robustness to object albedo. PMP strategies, alternatively, have been developed for various numbers of component patterns (≥ 3).

With the aim of defining and maximizing the SNR for PSMs, this chapter takes one of the most employed PMP strategies and illustrates the process of optimizing the patterns. Firstly, we present a short overview of the PMP strategy. By regarding the design of N patterns as placing points in N -dimensional space (coding space), we analyze the PMP patterns, the reflection process and traditional high frequency behaviors in coding space. Further, the SNR is derived mathematically

showing that, among all the coefficients in PSM pattern design, there are two that are directly related to the SNR of patterns. Then, we propose the edge-pattern strategy for various numbers of patterns ($N \geq 3$) which maximizes the two coefficients such that the pattern SNR is maximized. The SNR improvements from edge-pattern to PMP increase with the number of patterns, which indicates the lower pattern efficiency of PMP for larger N . Specifically, for 3 pattern strategies, the edge-pattern theoretically improves the SNR by 1.2381 times, whereas the improvement is as high

as 15.5421 times for 5 pattern strategies. Experimental results demonstrate that higher SNR is achieved by applying the edge-pattern compared to PMP strategy.

2.1 Phase measuring profilometry

Among the many proposed SLI methods, the technique of Phase Measuring Profilometry (PMP) is one of the most widely used and precise strategies [70, 74, 75, 76, 77]. The canonical PMP technique employs a set, $\{I_n : n = 0, 1, \dots, N - 1\}$, of sinusoidal wave patterns such that at the point (x^p, y^p) , in projector space, the intensity value is assigned as:

$$I_n(x^p, y^p) = A + B \cos(\Phi(y^p) - \frac{2\pi}{N}), \quad (2.1)$$

where A is the temporal DC value, which is normally 0.5, and B is the amplitude of the temporal AC signal, which is also normally 0.5. Thus, the sinusoidal signal covers the entire dynamic range of the projector $[0, 1]$. The term $\Phi(y^p)$ is the phase information and is designed according to:

$$\Phi(y^p) = \frac{2\pi F y^p}{H}, \quad (2.2)$$

where H is the pattern height (number of points in pattern) and F is the number of periods. Note the dependence of the phase term, $\Phi(y^p)$, on y^p as this is the parameter that will be used when triangulating with the camera, which is assumed to be positioned vertically above/below the projector. Thus, we denote $I_n(x^p, y^p)$ as $I_n(y^p)$.

After projecting the patterns, the reflection process can be expressed as:

$$I_n^c(x^c, y^c) = \alpha(x^c, y^c)[A + B \cos(\Phi(y^p) - \frac{2\pi}{N})] + \beta(x^c, y^c), \quad (2.3)$$

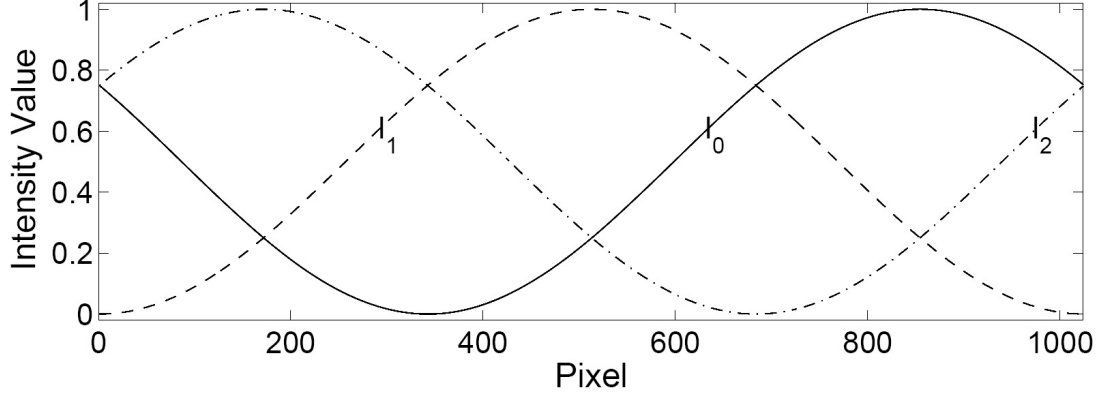


Figure 2.1: Cross sections of 3-pattern unit frequency PMP. The pattern height is 1024.

where (x^c, y^c) is the two-dimensional camera coordinate, $\alpha(x^c, y^c)$ represents the albedo within $[0, 1]$ where 0 is pure black and 1 is pure white, and $\alpha(x^c, y^c)\beta(x^c, y^c)$ is the albedo image from ambient illumination with intensity $\beta(x^c, y^c)$. Thus in the process of reflection, the albedo and ambient illumination effects are added into the patterns. The phase information $\Phi(y^p)$ is then obtained from a decoding function expressed as:

$$\begin{aligned}
 \phi(x^c, y^c) &= \arctan \left[\frac{\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \sin(\frac{2\pi n}{N})}{\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \cos(\frac{2\pi n}{N})} \right] \\
 &= \arctan \left[\frac{\alpha(x^c, y^c)B \sin(\Phi(y^p))}{\alpha(x^c, y^c)B \cos(\Phi(y^p))} \right]. \tag{2.4}
 \end{aligned}$$

The decoding function removes the effects of albedo and ambient illumination such that the designed phase information is recovered. Generally, if a high frequency signal is employed, the obtained phase ϕ , from Eq. (2.4), needs to be further “unwrapped” into unit frequency Φ in order to find the unique correspondence over the full resolution. Once $\Phi(x^c, y^c)$ is obtained, the 3-D world coordinates of a point can be calculated from $(x^c, y^c, \Phi(x^c, y^c))$ in a pre-calibrated system.

To solve Eq. (2.4), at least 3 patterns should be projected and captured since there

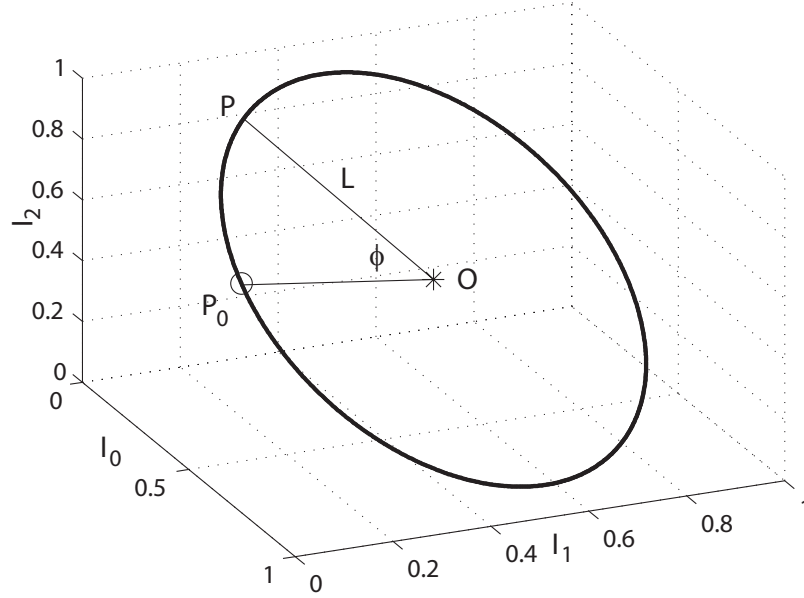


Figure 2.2: 3-pattern PMP in three dimensional coding space.

are $\Phi(y^p)$, $\alpha(x^c, y^c)$, and $\beta(x^c, y^c)$ unknowns in Eq. (2.3). The cross sections of three pattern PMP are shown in Fig. 2.1. As the most widely used pattern strategy, PMP is easy to implement and well developed for $N > 3$ patterns. The continuous smooth sinusoidal patterns are insensitive to blurring caused by an out of focus lens [78, 29].

N-dimensional coding space

In this chapter, the problem of codification for N patterns PMP with pattern height H is regarded as the placement of H points in the N -dimensional coding space R^N . Thus, the patterns $\{I_n : n = 0, 1, \dots, N-1\}$ become points $(I_0, I_1, \dots, I_{N-1})$ in R^N . For 3-pattern PMP, the visualization of the patterns in R^3 is shown in Fig. 2.2 where are total H points on the curve, corresponding to the H points in projector space. Each point P , in R^N , represents a temporal intensity sequence $(I_0, I_1, \dots, I_{N-1})$ projected from a pixel of the projector.

In order to analyze the reflection process (Eq. (2.3)) and the decoding function (Eq. (2.4)) in R^N , we propose the concepts of computational center and length where for a N -pattern PMP strategy, the *computational center* O is defined as:

$$O(o_0, o_1, \dots, o_{N-1}) = \left(\lim_{B \rightarrow 0} I_0, \lim_{B \rightarrow 0} I_1, \dots, \lim_{B \rightarrow 0} I_{N-1} \right), \quad (2.5)$$

where $\{I_n\}$ is defined in Eq. (2.1). The computational center O is the point in R^N when the temporal AC component of the patterns goes to 0. For 3-pattern PMP, the O point is (0.5, 0.5, 0.5) as shown in Fig. 2.2.

In R^N , we regard the vector from O to a point P as the computation vector \overrightarrow{OP} . The computational length, L , of \overrightarrow{OP} in R^N is then defined as:

$$L(y^p) = \left\{ \sum_{n=0}^{N-1} [I_n(y^p) - o_n]^2 \right\}^{\frac{1}{2}}, \quad (2.6)$$

where $\{y^p : 0, 1, \dots, H\}$ is the index of points in R^N , which corresponds to y^p in projector space. The computational length, L , is defined as the distance between each point P and the point O . For 3-pattern PMP, the pattern length is 0.6124, which is the same for all H points such that the curve of 3-pattern PMP in R^3 is a circle. We denote the point with $\Phi = 0$, on the circle, as P_0 . The phase Φ value of P is then the angle between $\overrightarrow{OP_0}$ and \overrightarrow{OP} , as shown in Fig. 2.2.

Reflection process and ambiguous high frequency

By projecting the patterns, the received images are affected by the albedo and ambient illumination as given in Eq. (2.3). In Fig. 2.3, a computational vector \overrightarrow{OP} is projected out, and the influence from albedo is equivalent to a “scaling” of \overrightarrow{OP} . As shown in Fig. 2.3, the vector vector \overrightarrow{OP} is scaled to $\overrightarrow{O_\alpha P_\alpha}$ due to the albedo effect, and during

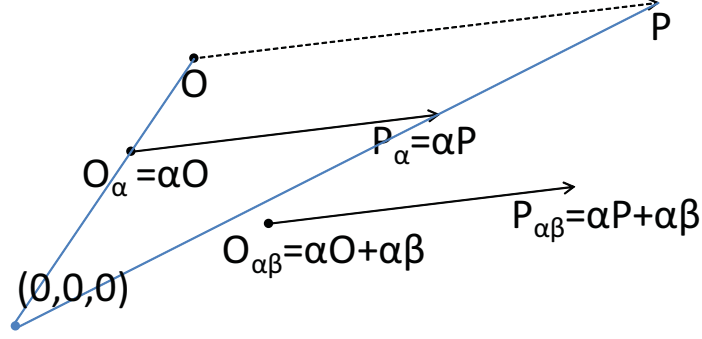


Figure 2.3: Reflection process in coding space R^N .

this linear scaling process, the length of $\overrightarrow{O_\alpha P_\alpha}$ is changed. However, the direction of $\overrightarrow{O_\alpha P_\alpha}$ remains the same.

On the other hand, the influence from ambient illumination is equivalent of a “shifting” process of $\overrightarrow{O_\alpha P_\alpha}$. As in Fig. 2.3, the vector $\overrightarrow{O_\alpha P_\alpha}$ is further shifted to $\overrightarrow{O_{\alpha\beta} P_{\alpha\beta}}$ because of the ambient illumination. In the case, the intensities of $O_{\alpha\beta}$ and $P_{\alpha\beta}$ stay within the dynamic range of the camera. During this shifting process, both the length and direction of $\overrightarrow{O_{\alpha\beta} P_{\alpha\beta}}$ are still unchanged.

After capturing the scaled and shifted vector $\overrightarrow{O_{\alpha\beta} P_{\alpha\beta}}$ in the camera, the decoding function, Eq. (2.4), calculates the angle between $\overrightarrow{O_{\alpha\beta} P_{\alpha\beta}}$ and $\overrightarrow{OP_0}$ where P_0 is the point with $\Phi = 0$. Since the reflection process keeps the direction of \overrightarrow{OP} , the obtained phase value equals to Φ in Fig. 2.2.

If a high frequency phase with $F > 1$ is encoded, the computational vector \overrightarrow{OP} revolves round the computational center O for F times. During de-codification, not only should the algorithm calculate the angle $\phi \in [0, 2\pi)$ between $\overrightarrow{O_{\alpha\beta} P_{\alpha\beta}}$ and $\overrightarrow{OP_0}$, but also should find the circle number (period number $f = [0, \dots, F - 1]$) of $\overrightarrow{O_{\alpha\beta} P_{\alpha\beta}}$ such that the unique phase value $\Phi = (2f\pi + \phi)/F$ can be obtained. However, since

using traditional high frequency phase the coordinates of points with the same ϕ but different f (periodical points) overlap in R^N , ambiguities exist making it difficult to obtain the period number f .

Indeed, traditional high frequency phase is realized by repeating the sinusoidal pattern spatially. We note this type of high frequency as ambiguous high frequency. It introduces ambiguities in phase, because periodical points have the same coordinates in R^N . The process of identifying f is well known as the unwrapping of high frequency phase and has been studied for over 20 years [95]. But the approaches are still not very satisfying. After unwrapping, the unique correspondence between projector and camera is found. Generally, the phase value Φ is further scaled back to $[0, 2\pi)$ in order to unify the 3-D reconstructions for both unit and high frequency pattern strategies.

In this chapter, we will show in Sec. 2.3 that the high frequency phase does not necessarily have to be realized by high frequency patterns. And besides the ambiguous high frequency, there is an other way to increase the frequency of phase where patterns are still in unit frequency and no ambiguity is introduced. We note this type of high frequency as non-ambiguous high frequency.

2.2 Influence of intensity noise

Although the reflection process does not change the direction of computational vector, employing such light patterns requires the projector and camera to support multiple gray levels, and noise also exists in the captured images. There are noise sources, including: camera [74], projector [45, 84], ambient light noise [45] as well as other sources [74, 85, 86, 87]. In this chapter, similar as several other researchers [74, 85, 86],

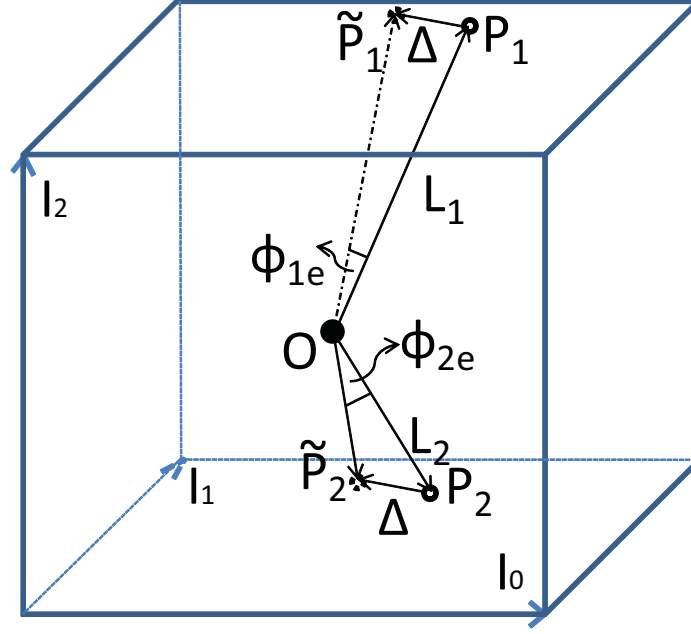


Figure 2.4: Relationship between computational length and created error. Two points P_1 and P_2 in R^3 , with $L_1 > L_2$, are affected by the same noise (w_0, w_1, w_2) such that the errors in phase Φ_{1e} and Φ_{2e} are respectively generated.

we assume that the combination of noise sources is additive, white, Gaussian noise, $w_n \sim N(0, \sigma^2)$, and by adding the noise, the captured images I_n^c can be rewritten as:

$$\tilde{I}_n^c(x^c, y^c) = \alpha(x^c, y^c)[I_n(y^p) + \beta(x^c, y^c)] + w_n(x^c, y^c), \quad (2.7)$$

where $\tilde{\cdot}$ denotes the observed variables with noise. The point $P = (I_0^c, I_1^c, \dots, I_{N-1}^c)$ is shifted to $\tilde{P} = (I_0^c + w_0, I_1^c + w_1, \dots, I_{N-1}^c + w_{N-1})$, where such shifting of P changes the direction of \overrightarrow{OP} , which will further creates the error Φ_e in phase.

In Fig. 2.4, we plot two points, P_1 and P_2 , in R^3 where $L_1 > L_2$. We assume the same noise (w_0, w_1, w_2) is added to P_1 and P_2 such that the created errors, in phase, are Φ_{1e} and Φ_{2e} respectively. Obviously with $L_1 > L_2$, Φ_{1e} is smaller than Φ_{2e} ; thus, the pattern sensitivity to noise is related to the computational length.

In order to analyze the relationship between error in phase, Φ_e , and computational

length, L , for real valued numbers a and b , the following equation always exists:

$$\arctan(a) - \arctan(b) = \arctan\left(\frac{a-b}{1+ab}\right), \quad (2.8)$$

we have,

$$\begin{aligned} \Phi_e(x^c, y^c) &= \Phi(x^c, y^c) - \tilde{\Phi}(x^c, y^c) \\ &= \arctan\left[\frac{\sum_{n=0}^{N-1} \tilde{I}_n^c(x^c, y^c) \sin(\frac{2\pi n}{N})}{\sum_{n=0}^{N-1} \tilde{I}_n^c(x^c, y^c) \cos(\frac{2\pi n}{N})}\right] - \\ &\quad \arctan\left[\frac{\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \sin(\frac{2\pi n}{N})}{\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \cos(\frac{2\pi n}{N})}\right] \\ &= \arctan\left(\frac{Y}{X}\right) \end{aligned} \quad (2.9)$$

where

$$\begin{aligned} X(x^c, y^c) &= \alpha(x^c, y^c)L(y^p) + \\ &\quad \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} w_n^c(x^c, y^c) \sin(\Phi(y^p) - \frac{2\pi n}{N}), \end{aligned} \quad (2.10)$$

and

$$Y(x^c, y^c) = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} w_n^c(x^c, y^c) \cos(\Phi(y^p) - \frac{2\pi n}{N}). \quad (2.11)$$

Because the noise, w_n , is much smaller than the computational length L , we can further approximate Eqs. (2.9) to (2.11) to

$$\Phi_e(x^c, y^c) \approx \frac{\sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} w_n^c(x^c, y^c) \cos(\Phi(y^p) - \frac{2\pi n}{N})}{\alpha(x^c, y^c)L(y^p)}. \quad (2.12)$$

By assuming $w_n \sim N(0, \sigma^2)$, where σ is regarded as the standard deviation of system noise, the standard deviation of Φ_e is

$$\sigma_{\Phi}(x^c, y^c) \approx \frac{\sigma}{\alpha(x^c, y^c)L(y^p)}. \quad (2.13)$$

When the high frequency phase with number of periods F is applied, the phase error Φ_e is divided by F after unwrapping and scaling back to $[0, 2\pi)$, such that σ_Φ is

$$\sigma_\Phi(x^c, y^c) \approx \frac{\sigma}{\alpha(x^c, y^c)FL(y^p)}, \quad (2.14)$$

From Eq. (2.14), the standard deviation of phase error decreases with the increasing of α , F , and L . The albedo, α , depends on the scene under scanning. The increasing of F will not reduce the error in high frequency phase signal, but, σ_Φ is reduced by unwrapping and scaling the high frequency phase into $[0, 2\pi)$. In this chapter, we consider the increasing of F as an approach to reduce error such that it becomes reasonable for us to define the SNR for PMP as

$$SNR(x^c, y^c) = \frac{L(y^p)F}{\sigma}. \quad (2.15)$$

As clearly shown in Eq. (2.15), the SNR value in captured images depends on the corresponding computational length L in patterns and the number of periods F . In other words, LF represents the signal strength of phase.

In this chapter, we focus on maximizing the SNR value for the phase shifting methods, which is realized by maximizing the computation length, L , and the number of periods, F , without introducing ambiguities. Here, the high frequency phase is non-ambiguous high frequency.

2.3 Edge-pattern

For a phase shifting pattern strategy, the optimization of patterns is achieved by maximizing L and F . In this subsection, we first maximize the computational length

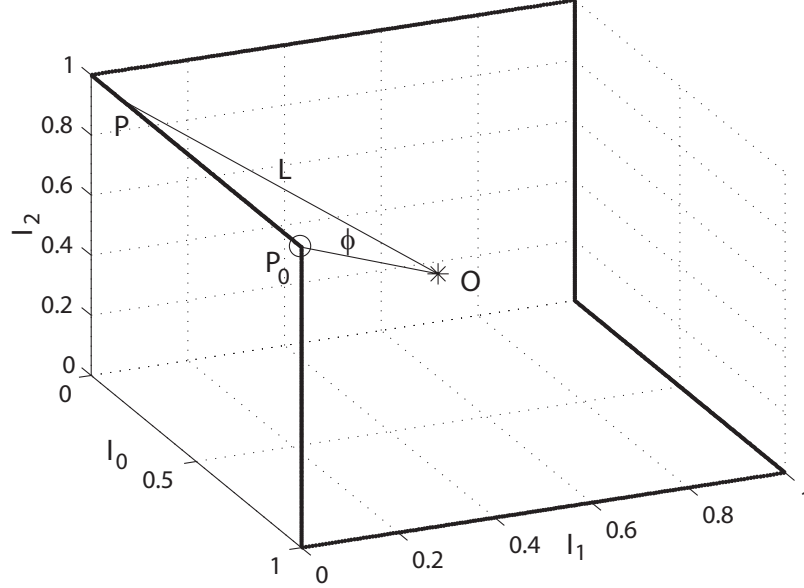


Figure 2.5: Cross sections of the 3-pattern E-P strategy in coding space.

L in the coding space R^N to achieve the highest SNR design of patterns. The maximizing of L in practice should have the following basic properties:

Property 1: The Dynamic range of the patterns: $\{I_n(x^j, y^j)\}$ should stay within the dynamic range of the projector, $[0, 1]$.

Property 2: At least three different illuminations should be projected in order to solve Eq. (2.4), since there are three unknown parameters α , β , and Φ .

Property 3: The maximization of computational length, L , should not affect the phase value at each point, in order to achieve an accurate 3D reconstruction based on phase.

For PMP, the computational center O is equal to $(0.5, 0.5, \dots, 0.5)$ in R^N ; thus, the computational length, L , defined in Eq. (2.6), can be further specified as:

$$L_{y^p} = \left[\sum_{n=0}^{N-1} (I_n(y^p) - 0.5)^2 \right]^{\frac{1}{2}}. \quad (2.16)$$

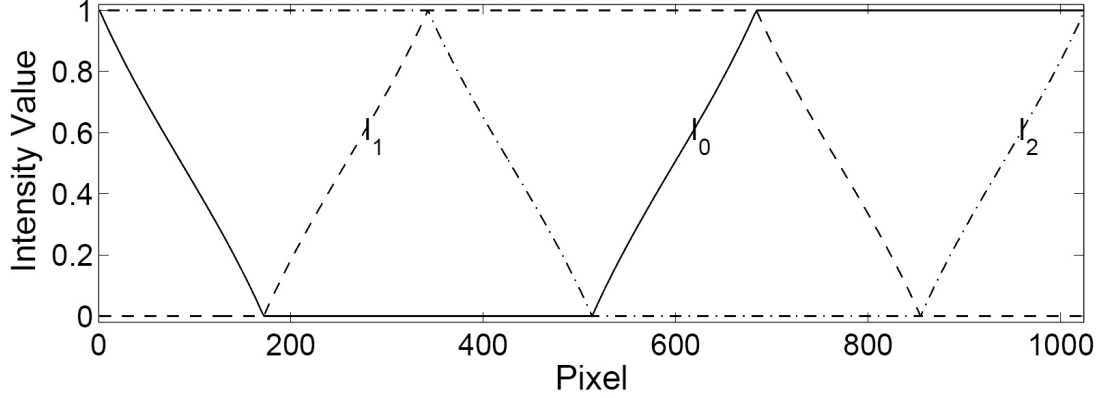


Figure 2.6: Cross sections of the 3-pattern E-P strategy.

Table 2.1: Edge order for 3 pattern strategy.

Edge number	e_0^c	e_1^c	e_2^c	e_3^c	e_4^c	e_5^c
I_0^c	med	inf	inf	med	sup	sup
I_1^c	inf	med	sup	sup	med	inf
I_2^c	sup	sup	med	inf	inf	med

Table 2.2: Four patterns E-P strategy.

Period number	f=0						f=1						f=2						f=3					
	e_0	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}	e_{11}	e_{12}	e_{13}	e_{14}	e_{15}	e_{16}	e_{17}	e_{18}	e_{19}	e_{20}	e_{21}	e_{22}	e_{23}
I_0	ie	de	ie	de	ie	de	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	1	1
I_1	0	0	0	1	1	1	de	ie	de	ie	de	ie	1	0	0	0	1	1	1	1	1	1	0	0
I_2	0	1	1	1	0	0	0	1	1	1	0	0	ie	de	ie	de	ie	de	0	1	1	1	0	0
I_3	1	1	0	0	0	1	1	1	0	0	0	1	0	0	1	1	1	0	ie	de	ie	de	ie	de

With the satisfaction of the Prop (2), the maximum L is achieved when $I_n(y^p)$ is 0 or 1.

However in order to satisfy Prop. (2), there should be a $I_k(y^p) = \lambda$ where $k \in [0, N-1]$

and $\lambda \in (0, 1)$ such that there are three different illuminations $\{i_0, i_1, i_2\} = \{0, 1, \lambda\}$

and Eq. (2.4) can be solved. And according to Prop. (3), we have

$$\arctan \left[\frac{\sum_{n=0}^2 i_n(y^p) \sin(\frac{2\pi n}{N})}{\sum_{n=0}^2 i_n(y^p) \cos(\frac{2\pi n}{N})} \right] = \Phi(y^p), \quad (2.17)$$

from where the value of λ can be obtained. In practice, we increase the span of

$\{I_0, I_1, I_2\}$ at points y^p until the span achieves 1 and then we shift $\{I_0, I_1, I_2\}$ into

$[0, 1]$.

Figure 2.5 shows the optimized 3-pattern strategy in R^3 coding space, where the patterns in projector space is shown in Fig. 2.6. Since for every point in R^3 there are illuminations 0 and 1, all the points are “pushed” onto the edges of the coding space. Thus, we denote the optimized pattern as edge-pattern (E-P). The phase value is still the angle between $\overrightarrow{OP_0}$ and \overrightarrow{OP} such that, using Eq. (2.4), Prop. (3) is satisfied. The obtained phase ranges in $[0, 2\pi)$ for unit frequency. The illumination λ is denoted as edge. For the 3-pattern strategy, there are 6 edges. If the illumination values on an edge are increasing spatially, the edge is denoted as an increasing edge (ie). Otherwise, if the values are decreasing spatially, the edge is denoted as a decreasing edge (de). The phase information is actually coded into the edges. It should be noted that the edges in projector space, as shown in Fig. 2.6, are not linear but distorted curves. The distortion ensures the linear phase obtained from edges.

After capturing the images, $\{I_n^c\}$, for each valuable point (x^c, y^c) in camera space, there are three different intensity values $\{I_0^c(x^c, y^c), I_1^c(x^c, y^c), I_2^c(x^c, y^c)\}$. We denote the infimum value as $\inf(x^c, y^c)$, the median value as $med(x^c, y^c)$, and the supremum value as $\sup(x^c, y^c)$. Thus, the six edges in the projected patterns will correspondingly have six edges in camera space $\{e_0^c, \dots, e_5^c\}$. For 3 pattern E-P, Eq.(2.4) is further specified as

$$\phi(x^c, y^c) = \arctan \left[\frac{3^{0.5}(I_1^c(x^c, y^c) - I_2^c(x^c, y^c))}{2I_0^c(x^c, y^c) - I_1^c(x^c, y^c) - I_2^c(x^c, y^c)} \right]. \quad (2.18)$$

The order of the edges in camera space should not be changed such that a linear phase ranging from 0 to 2π can be obtained from Eq. (2.18). For the 3-pattern strategy, the algorithm does not need to identify the edge number.

In the 3-pattern strategy, the number of patterns equals to the minimum requirement of Prop (2). When $N > 3$, there are extra patterns. It becomes possible to employ the illuminations with value either 0 or 1 to increase the number of periods F without introducing the ambiguity, as we will illustrate following.

For an $N > 3$ pattern strategy, we have C_N^1 choices to locate the edge where the rest of the illuminations are either 0 or 1, and at least one illumination is 0 and one illumination is 1. The rest of the illuminations should be designed such that no points in R^N overlap. Thus, the total number of edges, E , is

$$E = C_N^1(2^{N-1} - 2), \quad (2.19)$$

and every 6 edges consist of a period of phase such that the number of periods, F , is

$$F = \frac{C_N^1(2^{N-1} - 2)}{6}. \quad (2.20)$$

Thus, the non-ambiguous high frequency phase signal is embedded into the patterns. For the 4 pattern strategy. Based on the definitions of ie and de, there are 24 edges and, correspondingly, 6 periods. The E-P patterns are shown in Table 2.2.

Different from traditional high frequency phase, no points in R^N overlap such that the frequency is introduced without ambiguity as shown in Table 2.2. Each edge can be identified by referring the code created by 0 and 1 illuminations. It should be noted that it is also possible to increase the ambiguous high frequency for E-P by repeating the patterns spatially; however, similar to the traditional approaches, the unwrapping problem would be introduced. Thus, in this chapter we only focus on non-ambiguous high frequency.

When $N > 3$, for each valuable point (x^c, y^c) in camera space, the algorithm needs to first identify the median intensity $I_\lambda^c(x^c, y^c) = med(\{I^c(x^c, y^c)\})$. For intensities higher than $med(x^c, y^c)$, they are identified as supremum intensities $\{I_{sup}^c(x^c, y^c) : I_{sup}^c(x^c, y^c) > med(x^c, y^c)\}$. The intensities lower than $med(x^c, y^c)$ are identified as infimum intensities $\{I_{inf}^c(x^c, y^c) : I_{inf}^c(x^c, y^c) < med(x^c, y^c)\}$.

At each point (x^c, y^c) , there is only one median intensity which is corresponding to an ie or de in the projected patterns. However, there are several infimum and supremum intensities. In practice, the infimum and supremum values are further refined by

$$\inf(x^c, y^c) = mean(\{I_{inf}^c(x^c, y^c)\}) \quad (2.21)$$

and

$$\sup(x^c, y^c) = mean(\{I_{sup}^c(x^c, y^c)\}), \quad (2.22)$$

where the noise in $\inf(x^c, y^c)$ and $\sup(x^c, y^c)$ will be further reduced. The temporal order of infimum, median, and supremum intensities create a code, according to which the algorithm can decide the edge number at (x^c, y^c) .

The high frequency phase, ϕ , is then obtained from the three intensities $\{\inf, med, \sup\}$ using Eq. (2.18). In each period, the order of inf, med, and sup, when applying Eq. (2.18) is shown in Table 2.1. For example at point (x^c, y^c) , if the captured edge is identified as e_{20}^c , which is the third edge in period 4 according to Table 2.1, then we have $I_0^c(x^c, y^c) = \inf(x^c, y^c)$, $I_1^c(x^c, y^c) = \sup(x^c, y^c)$ and $I_2^c(x^c, y^c) = med(x^c, y^c)$ for Eq. (2.18). And the obtained high frequency phase, ϕ , is further unwrapped and

Table 2.3: Theoretical improvements of SNR for different number of patterns.

E-P/PMP	L	F	SNR
N=3	1.2381	1	1.2381
N=4	1.2834	4	5.1336
N=5	1.3318	11.67	15.5421

scaled to Φ by

$$\Phi(x^c, y^c) = \frac{\phi(x^c, y^c) + 2f(x^c, y^c)\pi}{F}. \quad (2.23)$$

Here since the edge number and period number are all obtained by the temporal order of infimum, median, and supremum intensities, no ambiguity is introduced for the high frequency phase when unwrapped and scaled to Φ .

From Eq. (2.15), the increasing of L and F improves the SNR of patterns. The theoretical average improvements in SNR for different numbers of patterns when comparing the E-P strategies and the traditional PMP strategies are shown in Table 2.3. As shown in Table 2.3, with an increasing number of patterns, the improvement in SNR also increases. For $N = 5$ strategies, the computational length is improved by 1.3318 times and the non-ambiguous high frequency is improved by 11.67 times. Thus, theoretically, the SNR of E-P is 15.5421 times higher than that of a traditional PMP strategy. Table 2.3 also indicates that the efficiency of traditional PMP pattern becomes lower with more patterns.

2.4 Experimental results and discussion

To demonstrate the advantages of the proposed pattern strategy, we developed the SL system shown in Fig. 4.17, employing an 8 bpp, monochrome, Prosilica GC640M, gigabit ethernet camera with 640×480 pixel resolution. The projector is composed of

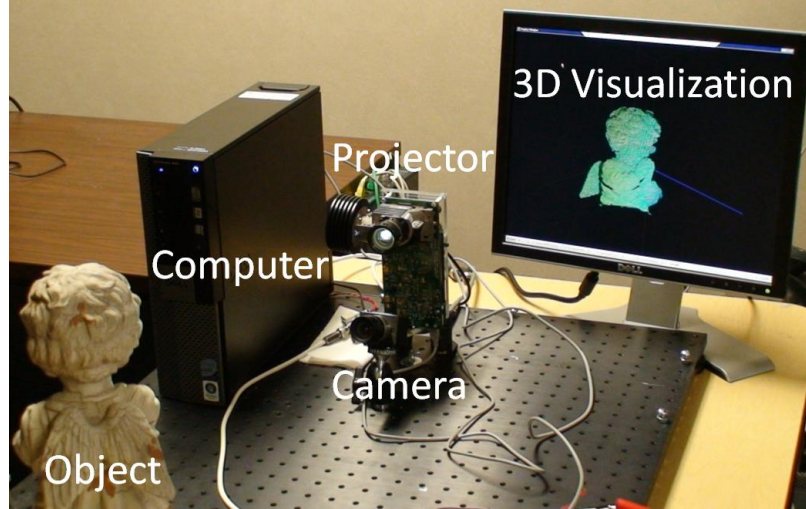


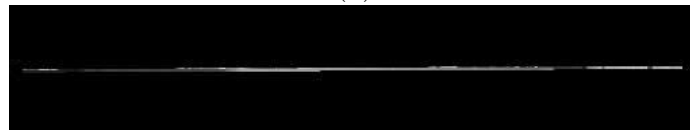
Figure 2.7: The prototype system setup for E-P.

a Texas Instrument’s Discovery 1100 board with ALP-1 controller and LED-OM with 225 ANSI lumens. The resolution of the 8 bpp, monochrome, projector is 1024×768 , with a maximum frame rate of 150 fps. The camera and projector are synchronized by an external triggering circuit with a baseline distance between camera and projector of 120 mm . For testing, the scanned object was placed around 600 mm away. Gamma correction was performed on the received images, while a lookup table was created to correct optical distortion. In the following experiments, objects were scanned with a camera exposure time of 1.6 ms . Due to the low illumination of our projector and the small exposure time, the standard deviation of system noise, σ , was 1.3365.

In the first experiment, we compare the accuracy of traditional PMP and E-P in the presence of noise. For this purpose, we performed both PMP and E-P for different N by scanning a textured, flat, poster board. The approximate size of the board is $210 \times 280 \text{ mm}$, where the front view and the top views of the board are shown in Fig. 3.12.



(a)



(b)

Figure 2.8: The scanned textured, flat, poster board. (a) The front view. (b) The top view.

In this paper, the proposed pattern strategy is to maximize the SNR of the patterns. Thus, only error caused by the uncertainty noise is concerned here. In order to obtain the “ground-truth” phase of the scanned board, we scanned the board 2000 times using the multi-frequency PMP strategy studied in [72]. And the ground-truth phase is the averaged value of the 2000 phase values. Then, the board was scanned

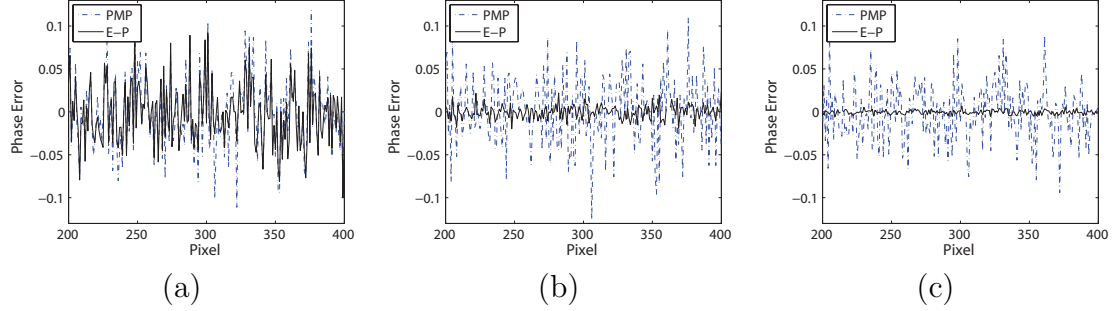


Figure 2.9: (a) Cross sections of phase error of the scanned board from 3 pattern PMP and E-P strategies. (b) Cross sections of phase error from 4 pattern PMP and E-P strategies. (c) Cross sections of phase error from 5 pattern PMP and E-P strategies.

using PMP and E-P strategies. As a means of analysis, Fig. 2.9 shows cross section comparisons of the phase error using traditional PMP and proposed E-P patterns. The phase error is the difference between the obtained phase and the ground-truth phase. As shown in Fig. 2.9 (a), for $N = 3$ pattern strategies, the phase of E-P is more accurate than that of PMP, although the difference between the two phase errors is small. With more patterns, the accuracy of E-P increases faster than that of PMP, as indicated in Figs. 2.9 (b) and (c).

Further, in order to quantify the differences between PMP and E-P, we scanned the board using PMP and E-P respectively for 2000 times. The numbers of patterns change from 3 to 5. The phase error is the difference between each individually obtained phase and the ground-truth phase. The mean values of the phase error for both PMP and E-P with different numbers of patterns are close to 0. And the standard deviation values of phase error are list in Table 2.4.

In Table 2.3, we list the theoretical improvements in SNR for different number of patterns, comparing E-P and traditional PMP strategies. In experiments, we obtain

Table 2.4: Standard deviation values of phase error of the scanned board for PMP and E-P.

Number of patterns	3	4	5
σ_{PMP}	0.0263	0.0228	0.0204
σ_{E-P}	0.0213	0.0048	0.0015
$\sigma_{PMP}/\sigma_{E-P}$	1.2362	4.7109	13.6027

the SNR improvement by measuring the reduction in standard deviation of phase error. In Table 2.4, we list the ratio between standard deviations of phase error using PMP and E-P, for $N = 3, 4,$ and 5 pattern strategies. The ratio represents the improvement. For $N = 3$, the experimental SNR improvement is 1.2362, whereas the theoretical value is 1.2381. For $N = 4$, the experimental SNR improvement is 4.7109, whereas the theoretical value is 5.1336. And for 5 pattern strategy, the experimental SNR improvement is 13.6027, whereas the theoretical value is 15.5421. The difference between the experimental and theoretical values is due to the defocus and resolution differences between projector and camera. Thus, the received patterns can be regarded as low pass filtered and the points where the projected light intensity changes from 0 to 255 become difficult to be identified. However, as shown in Table 2.4, the E-P still achieves a higher SNR than traditional PMP. And, as the number of patterns increasing, the improvement of SNR increases. This indicates that, using traditional PMP strategy, the efficiency of patterns becomes lower with a larger N .

As the second experiment, we scanned a white plaster model, Alice, as shown in Fig. 2.10, using different pattern strategies with 2000 times for each strategy. Typical 3-D reconstructions are shown in Fig. 2.11. From Fig. 2.11 (a) to (c), the results are respectively obtained using 3, 4 and 5 pattern traditional PMP strategies, whereas,



Figure 2.10: The scanned white plaster model, Alice.

from Fig. 2.11 (d) to (f), the results are created using 3, 4 and 5 pattern proposed E-P strategies. Comparing the results using the same number of PMP patterns, the proposed E-P strategy achieves better results in the present of noise. For $N = 3$, the improvement is not obvious in the 3-D results (19% reduction in the standard deviations of reconstruction error). However for $N = 4$ and $N = 5$, the improvements are visually clear. Compared to traditional PMP, the proposed E-P is more robust to the uncertainty noise.

For a quantified analysis, the “ground-truth” phase of the scanned Alice is obtained from an average of 2000 times multi-frequency PMP scanning. Thus, the phase error can be calculated. The distributions of phase errors using different strategies are shown in Fig. 2.12. The mean values of the phase error for both PMP and E-P with different numbers of patterns are close to 0. However, the standard deviation

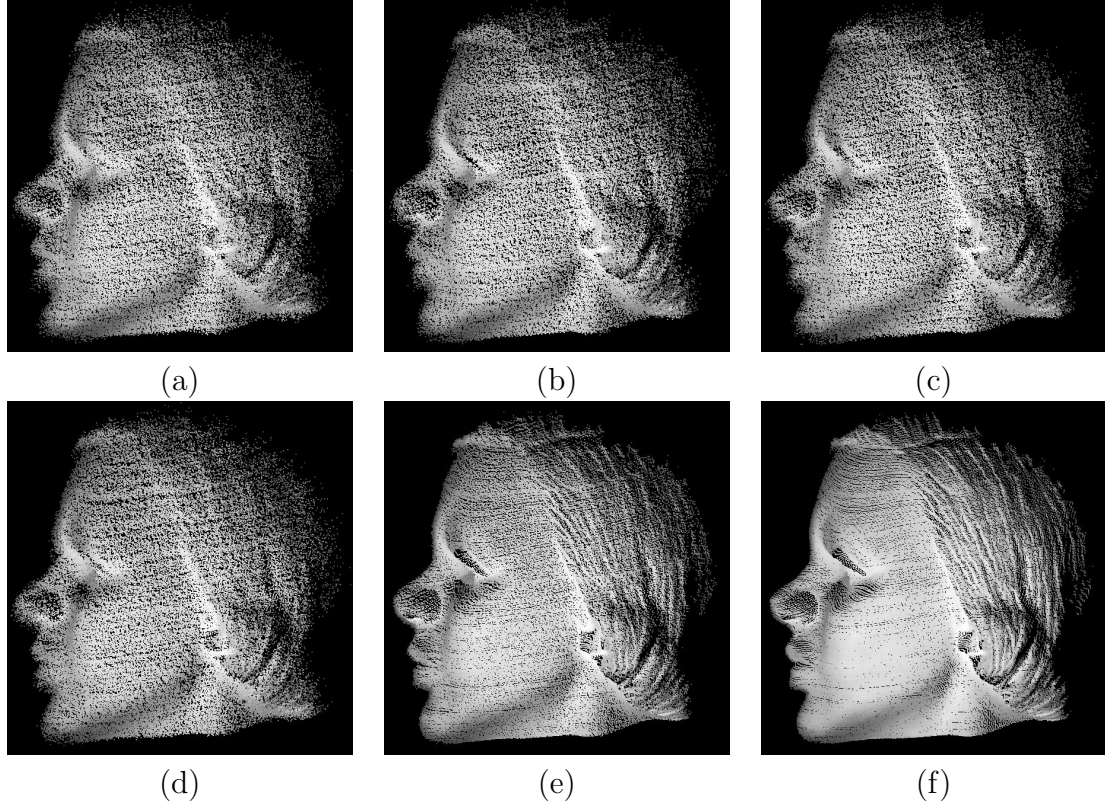


Figure 2.11: 3D reconstructions of Alice. Results are shown with depth rendering. (a) Reconstruction using traditional 3 pattern PMP. (b) Reconstruction using traditional 4 pattern PMP. (c) Reconstruction using traditional 5 pattern PMP. (d) Reconstruction using 3 pattern E-P. (e) Reconstruction using 4 pattern E-P. (f) Reconstruction using 5 pattern E-P.

values from E-P strategies are smaller than those from PMP strategies. And, from Fig. 2.12 (a) to (c), as the N varies from 3 to 5, the standard deviation values of E-P decreases faster than that of PMP.

The experimental standard deviation values of phase error from scanning Alice are listed in Table 2.5. The experimental SNR improvement for $N = 3$ is 1.2377. For $N = 4$, the experimental SNR improvement achieves 4.9729. And, it is 14.0625 for $N = 5$. The difference between the experimental and theoretical improvement values indicates that the problem caused by the de-focus and resolution differences between

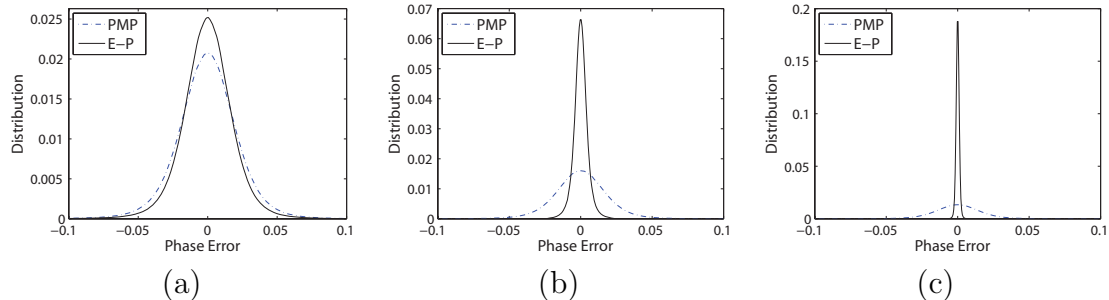


Figure 2.12: (a) Distribution of phase error of Alice from 3 pattern PMP and E-P strategies. (b) Distribution of phase error from 4 pattern PMP and E-P strategies. (c) Distribution of phase error from 5 pattern PMP and E-P strategies.

Table 2.5: Standard deviation values of phase error of the scanned Alice for PMP and E-P.

Number of patterns	3	4	5
σ_{PMP}	0.0103	0.0089	0.0080
σ_{E-P}	0.0083	0.0018	5.7×10^{-4}
$\sigma_{PMP}/\sigma_{E-P}$	1.2377	4.9729	14.0625

projector and camera is still suffered in the second experiment. However, compared to the first experiment, the results from both PMP and E-P are improved. It is due to the fact that Alice is a white plaster model which has a higher albedo value (α in Eq. (2.14)).

By regarding the design of N pattern strategy in SLI system as placing points in an N dimensional coding space, we have defined computational length and derived a mathematical SNR model for the widely used PMP pattern strategy. As indicated in the model, the SNR is strongly related to the computational length and the frequency of patterns. Thus, we proposed a scheme to increase the computational length without changing the phase value. And we proposed the concept of non-ambiguous high frequency. Different from traditional high frequency patterns where there are points overlapping in coding space, the proposed non-ambiguous high frequency patterns do not have ambiguous points such that the unwrapping of the high frequency phase

can be solved temporally without projecting extra patterns. The same as PMP, the proposed E-P strategy requires at least 3 patterns.

Experimental results demonstrate that by maximizing the computational length and employing the non-ambiguous high frequency, the E-P strategy achieves a higher SNR than the traditional PMP strategy. With more patterns, the pattern SNR improvement increases further; however, in practice, we noticed that the difference between theoretical and experimental improvements also increases with the number of patterns. We believe that it is caused by the de-focus and resolution difference of the projector and camera pair. For future research, we will further exploit the E-P strategy, especially the de-codification methods, such that we can address to the de-focus and resolution difference problems.

Chapter 3 PERIOD CODED PHASE SHIFTING

In an effort to develop a low-computation, high-frequency pattern scheme that can achieve high measurement accuracy and unwrapping robustness, this chapter presents the novel method of Period Coded Phase Shifting (PCPS) that avoids the introduction of extra patterns without losing accuracy. For this purpose, we first develop the notion of spatial intensity efficiency to measure the proportion of the projector’s dynamic range spanned by the projected patterns and, thereby, determine if the pattern intensity values are under-utilized. For pattern sets with an efficiency of less than 100%, a reference unit frequency signal (or period cue) is embedded into the pattern set such that pixels of the projector, having the same phase but different periods, no longer project the same intensity values. To maximize the difference among these same-phase-but-different-period points (periodical points), we define pattern entropy as a measure of how far apart the periodical points are. The parameters of this added information are then optimized to make maximum amplitude of the period cue.

Having laid this foundation of efficiency and entropy, this chapter proposes a hybrid unwrapping method to take advantage of both spatial and temporal unwrapping approaches where a triangular phase waveform is proposed for phase shifting methods that provides a doubled frequency compared to the traditional sawtooth phase waveform. The proposed method works for most phase shifting schemes, and in this chapter, we demonstrate the advantages of the proposed approach both mathematically and experimentally using a high-speed, real-time prototype system achieving

120 fps using a 640×480 resolution camera.

3.1 Phase shifting structured light illumination

Phase shifting methods project a set of time-multiplexed patterns, $\{I_n^b : n = 0, 1, \dots, N-1\}$, onto a target object such that an off-axis imaging sensor observes the scene and captures the wave patterns distorted by the surface topology under inspection. Generally, the patterns $\{I_n^b\}$ are designed as

$$I_n^b(x^j, y^j) = A + Bs(x^j, y^j), \quad (3.1)$$

where A is a temporal DC value, B is the amplitude (or projector modulation) value of a periodical signal function $s(x^j, y^j)$. The coordinate (x^j, y^j) is the corresponding coordinate in the projector. The captured images, $\{I_n^c : n = 0, 1, \dots, N-1\}$, are then denoted as

$$I_n^c(x^c, y^c) = \alpha(x^c, y^c)I_n^b(x^j, y^j) + \alpha(x^c, y^c)\beta(x^c, y^c), \quad (3.2)$$

where the superscript c indicates that I_n^c is now in the camera space and (x^c, y^c) is the two-dimensional camera coordinate [45]. In Eq. (3.2), $\alpha(x^c, y^c)$ represents the albedo with $\alpha(x^c, y^c) \in [0, 1]$ where 0 is pure black and 1 is pure white. The term, $\alpha(x^c, y^c)\beta(x^c, y^c)$, represents the albedo image from ambient illumination with intensity $\beta(x^c, y^c)$.

Through a decoding function and a phase unwrapping algorithm, the phase Φ that represents the coordinate of x^j or y^j , can be obtained from the “wrapped” (or

coded) phase ϕ^w which is expressed as

$$\phi^w(x^c, y^c) = g \left[\frac{U(x^c, y^c)}{V(x^c, y^c)} \right], \quad (3.3)$$

where

$$U(x^c, y^c) = \sum_{n=0}^{N-1} a_n I_n^c(x^c, y^c), \quad (3.4)$$

and

$$V(x^c, y^c) = \sum_{n=0}^{N-1} b_n I_n^c(x^c, y^c). \quad (3.5)$$

The terms a_n and b_n are two coefficients in summations such that, in $U(x^c, y^c)$ and $V(x^c, y^c)$, the terms of $\beta(x^c, y^c)$ and the DC value A in images are canceled, while the division between the two summations cancels the terms of $\alpha(x^c, y^c)$. The term $g(\cdot)$ is a function that estimates the phase values $\phi^w(x^c, y^c) \in [0, 2\pi)$ out of image intensity values. The modulation $M(x^c, y^c)$ is given by

$$M(x^c, y^c) = \gamma \sqrt{(U(x^c, y^c))^2 + (V(x^c, y^c))^2}, \quad (3.6)$$

where γ is a coefficient related with the number of patterns. Noting that the value of $\frac{M(x^c, y^c)}{\alpha(x^c, y^c)}$ represents the strength of the high frequency signal at pixel (x^c, y^c) , we define the Signal to Noise Ratio (SNR) for phase shifting methods as

$$SNR(x^c, y^c) = \lambda \frac{M(x^c, y^c)}{\alpha(x^c, y^c) \sigma}, \quad (3.7)$$

where λ is a coefficient that is a function of the number of patterns, N , and σ is the standard deviation of the system noise. It should be noted that in Eq. (3.7), the temporal DC value A in the time-multiplexed patterns $\{I_n^b\}$ can be canceled by the summations in U and V . Thus, the SNR is only related to the temporal

AC values B in $\{I_n^b\}$. The specific decoding functions, Eqs. (3.3) and (3.6), are designed for the particular pattern strategy such that, after unwrapping the wrapped phase ϕ^w in Eq. (3.3) into the final phase $\Phi \in [0, 2P\pi)$, where P is the number of periods, the correspondence problem can be solved without using any point-matching or image enhancement. The 3-D coordinates are finally recovered from pre-calibrated triangulation [100, 23].

For the specific phase shifting method of PMP, the patterns $\{I_n^b : n = 0, 1, \dots, N - 1\}$ of Eq. (3.1) are given by [70]

$$I_n^b(x^j, y^j) = A + B \sin(\Phi(x^j, y^j) - \frac{2n\pi}{N}). \quad (3.8)$$

The sinusoidal light signal can cover the whole resolution range of a projector $[0, 1]$, if both A and B are given the value of 0.5. After capturing the PMP images $\{I_n^c\}$ in a camera, the decoding function of Eq. (3.3) is given by

$$\phi^w(x^c, y^c) = \arctan \left[\frac{\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \cos(\frac{2\pi n}{N})}{\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \sin(\frac{2\pi n}{N})} \right] = \arctan \left[\frac{U(x^c, y^c)}{V(x^c, y^c)} \right], \quad (3.9)$$

where

$$U(x^c, y^c) = \frac{N}{2} \alpha(x^c, y^c) B \sin(\Phi(x^c, y^c)), \quad (3.10)$$

and

$$V(x^c, y^c) = \frac{N}{2} \alpha(x^c, y^c) B \cos(\Phi(x^c, y^c)). \quad (3.11)$$

The modulation $M(x^c, y^c) = \alpha(x^c, y^c) B$ (Eq. (3.6)) becomes

$$M(x^c, y^c) = \frac{2}{N} \sqrt{(U(x^c, y^c))^2 + (V(x^c, y^c))^2}. \quad (3.12)$$

Thus, the specific SNR of Eq. (3.7) for PMP pattern strategy is [74]:

$$SNR(x^c, y^c) = \frac{\sqrt{N}M(x^c, y^c)}{\sqrt{2}\alpha(x^c, y^c)\sigma}, \quad (3.13)$$

that is $\lambda = \sqrt{N/2}$.

Non-ambiguous phase unwrapping

In order to obtain the correspondence between camera and projector, the high frequency signal ϕ^w should be further unwrapped into $\Phi \in [0, 2P\pi)$ where P is the number of periods. To do so, a secondary unit frequency signal (period cue) can be added into the projected patterns. In this sub-section, we first propose a simple approach to add the second signal such that, in case of PMP, the projected patterns are given by

$$I_n^b(x^j, y^j) = A + B \sin(\Phi(x^j, y^j) - \frac{2n\pi}{N}) + p(x^j, y^j), \quad (3.14)$$

where $\{p(x^j, y^j) : 0, \frac{1}{2^{k-1}}, \dots, \frac{P-1}{2^{k-1}}\}$ is the period number at (x^j, y^j) when a k -bits per pixel (bpp) projector is used. After capturing the images $\{I_n^c\}$ through Eq. (3.2) and ignoring the effects of system noise and ambient light, the high frequency wrapped phase $\phi^w(x^c, y^c)$ is obtained from Eq. (3.9), while the period information $p(x^c, y^c)$ can be obtained by

$$p(x^c, y^c) = \frac{\frac{1}{N} \sum_{n=0}^{N-1} I_n^c(x^c, y^c)}{M(x^c, y^c)} B - A. \quad (3.15)$$

Thus, $\Phi(x^c, y^c) = \phi^w(x^c, y^c) + 2\pi p(x^c, y^c)$, i.e., the high frequency signal is unwrapped. Compared to the traditional PMP, this method temporally unwraps the high frequency phase information without projecting extra patterns; however in order to

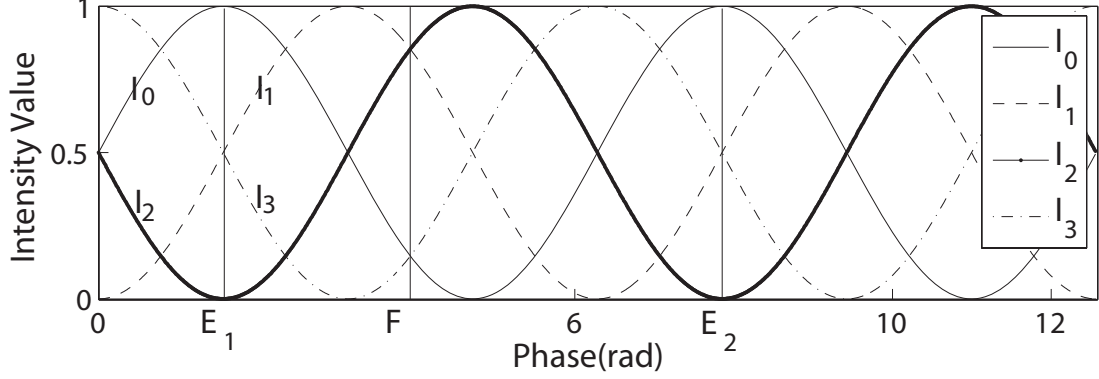


Figure 3.1: Illustration of four pattern, two period PMP strategy where the points E_1 and E_2 are shiftable, periodical points while point F is a shiftable and scalable point.

leave room within the projector's resolution range for the unit frequency signal p , the high frequency amplitude needs to be reduced from $B = 0.5$ to $0.5 - \frac{P}{2^k - 1}$. Thus, the obtained modulation M and the SNR of the high frequency signal has to be reduced, thereby, increasing the impact of noise on the phase image.

With the ability to uniquely decode the period number, it is expected that an increase in the period number of the original signal will compromise the effects of noise to a greater extent. But it would be ideal to find a low frequency period cue that uniquely decodes the period number in the demodulated phase image without reducing the SNR . To this end, we note that there are intelligent ways to design and add this second signal such that the combined signal still falls within the resolution range of a projector without reducing the high frequency signal's amplitude. Unlike the pattern strategy in Eq. (3.14), these intelligent approaches are robust to system noise and ambient illumination.

In order to develop our discussion of intelligent period cue design, we first examine the case of $N = 4$ and $P = 2$ pattern PMP (Fig. 3.1) to illustrate how a second

signal can be added without reducing the original signal's amplitude. Here, Eq. (3.9) is reduced to

$$\phi^w(x^c, y^c) = \arctan \left[\frac{I_0^c(x^c, y^c) - I_2^c(x^c, y^c)}{I_1^c(x^c, y^c) - I_3^c(x^c, y^c)} \right] \quad (3.16)$$

where high frequency signal $\phi^w(x^c, y^c)$ remains constant as long as the ratio between $(I_0^c(x^c, y^c) - I_2^c(x^c, y^c))$ and $(I_1^c(x^c, y^c) - I_3^c(x^c, y^c))$ remain the same even though the values of $\{I_0^c(x^c, y^c), \dots, I_3^c(x^c, y^c)\}$ may change. As such, we define the action of “shifting” as the process of changing the intensity values of $I_n^b(x^j, y^j)$ such that the ratios of the distances among $\{I_n^b(x^j, y^j)\}$ are changed. We further define the action of “scaling” as the process of changing the intensity values of $I_n^b(x^j, y^j)$ such that the ratios of the distances among $\{I_n^b(x^j, y^j)\}$ remain the same.

Having the notions of shifting and scaling, we can classify the points of our SLI pattern set into: (1) shiftable points, and (2) shiftable and scalable points. As shown in Fig. 3.1 point E_1 , the wrapped phase value $\phi_{E_1}^w$ is 0.5π such that $\{I_0^b = 1, I_1^b = 0.5, I_2^b = 0, I_3^b = 0.5\}$. The projected intensities of I_1^b and I_3^b can be shifted within $[0, 1]$ under the constraint that $I_1^b = I_3^b$, and the obtained $\phi_{E_1}^w$ will not change. The point E_1 is a shiftable point; however, the intensities at E_1 can not be scaled up since I_0^b achieves the supremum bound of the projector, and I_2^b achieves the infimum bound. At the same time as shown in Fig. 3.1 point F , $\phi_F^w = 1.25\pi$ and $\{I_0^b = 0.1464, I_1^b = 0.8536, I_2^b = 0.8536, I_3^b = 0.1464\}$ where, since none of the four projected intensities achieves 0 or 1, the intensities can be scaled up as long as the ratio between $(I_0^b - I_2^b)$ and $(I_1^b - I_3^b)$ stay the same. And, the intensities at F are also shiftable.

It is the operations of shifting and scaling pixels, without changing the high fre-

quency signal, that provides the possibility of encoding a second signal in the projection patterns. As stated above for the point E_1 , the shifting of $I_1^b(x^j, y^j)$ and $I_3^b(x^j, y^j)$ will not change the values of $(I_0^c(x^c, y^c) - I_2^c(x^c, y^c))$ and $(I_1^c(x^c, y^c) - I_3^c(x^c, y^c))$ under the constraint that $I_1^b(x^j, y^j) = I_3^b(x^j, y^j)$. Actually, only the A values of $I_1^b(x^j, y^j)$ and $I_3^b(x^j, y^j)$ in Eq. (3.8) are changed. Thus, the shifting of $I_1^b(x^j, y^j)$ and $I_3^b(x^j, y^j)$ does not affect the B value nor the SNR of $\phi^w(x^c, y^c)$; however, scaling the intensities causes a change in B in Eq. (3.8). By increasing the distance between intensities, the B value will be increased such that the SNR is improved. In order to embed a second signal into the pattern set without changing the SNR of the original signal, we need only be concerned with shifting intensity values, and in order to measure how much particular intensities can be shifted in the patterns, we first define the spatial intensity efficiency η .

Spatial intensity efficiency

For phase shifting patterns (in projector space), there are P points, one point for each period of the projected pattern, having the phase value $\phi^w(x^c, y^c) = \Phi(x^c, y^c) - 2p(x^c, y^c)\pi$ where $\phi^w(x^c, y^c)$ is the wrapped phase value within the range $[0, 2\pi)$. As shown in Fig. 3.1, the points E_1 and E_2 are two such periodical points having the same wrapped phase value $\phi^w = 0.5\pi$. Correspondingly, there are NP intensity values, $\{I_n(\phi^w) : n = 0, 1, \dots, NP - 1\}$, projected from these periodical points across the pattern set. For each intensity $I_n(\phi^w)$, there are 2^k available intensities that $I_n(\phi^w)$ can choose from inside a k bpp projector. We denote $\Delta_n(\phi^w)$ as the number of different intensity values that result in the same value of phase, *i.e.* the shiftable

intensity values. In Fig. 3.1, the shiftable intensity values of I_1^b at point E_1 is 2^k . The spatial intensity efficiency, η , for this phase shifting method is then defined as

$$\eta = \frac{P}{L} \sum_{\phi^w=0}^{2\pi} \left[\left(1 - \frac{\sum_{n=0}^{NP-1} \Delta_n(\phi^w)}{NP2^k} \right) \times 100\% \right], \quad (3.17)$$

where L is the pattern length or the number of pixels across which the pattern is defined. It should be noted that, for patterns with length L and periods P , there are $\frac{L}{P}$ different ϕ^w values. The spatial intensity efficiency is an average over all the wrapped phase values. If all the intensities can not be changed at all in order to obtain the same value of phase, $\Delta_n(\phi^w) = 0$ and $\eta = 100\%$. On the other hand, if no signal is coded into patterns, $\Delta_n(\phi^w) = 2^k$ and $\eta = 0\%$. For the same pattern strategy with the same numbers of patterns and periods, increasing of the amplitude of phase signal will result in a higher spatial intensity efficiency and higher SNR . Table 3.1 lists the spatial intensity efficiencies for several commonly used phase shifting methods where $L = 1, 280$ pixels on an 8 bpp projector.

Now if the spatial intensity efficiency of a phase shifting method is less than 100%, then it becomes possible to employ the remaining dynamic range of the projector for other purposes. To solve the unwrapping problem in traditional phase shifting methods, particularly in this chapter, the remaining dynamic range is employed for embedding a period cue D under four constraints (described in Sec. 3.2). Here, we introduce a generic signal such that the Eq. (3.2) becomes

$$I_n^c(x^c, y^c) = \alpha(x^c, y^c)[I_n^b(x^j, y^j) + C_n(x^j, y^j)] + \alpha(x^c, y^c)\beta(x^c, y^c), \quad (3.18)$$

where the term $\{I_n^b\}$ contains the information of high frequency phase ϕ^w , and $\{C_n\}$

contains the secondary information D . Eq. (3.3) can be re-written as

$$\phi^w(x^c, y^c) = g \left(\frac{\sum_{n=0}^{N-1} a_n [\alpha(x^c, y^c)(I_n^b(x^j, y^j) + C_n(x^j, y^j) + \beta(x^c, y^c))]}{\sum_{n=0}^{N-1} b_n [\alpha(x^c, y^c)(I_n^b(x^j, y^j) + C_n(x^j, y^j) + \beta(x^c, y^c))]} \right). \quad (3.19)$$

Similar to Eq. (3.3), the terms of $\beta(x^c, y^c)$ can be canceled by the two summations. And, it is also possible that $\{C_n(x^j, y^j)\}$ can be canceled by the two summations. For example, at the point E_1 in Fig. 3.1, the $\{C_n\}$ can be designed as $\{0, s, 0, s\}$ where $s \in [-0.5, 0.5]$ such that $\{C_n\}$ would be canceled in ϕ^w and there is still a considerably large dynamic range for the secondary generic signal since the value of s can vary in $[-0.5, 0.5]$. The details about how the secondary signal can be added without changing the amplitude of $\{I_n^b\}$ will be discussed in Sec. 3.2. Without changing the amplitude of $\{I_n^b\}$, the SNR of the original signal would also be the same after embedding the secondary information. On the other hand, the period cue is extracted according to

$$D(x^c, y^c) = F_D \left(\{ \alpha(x^c, y^c)(I_n^b(x^j, y^j) + C_n(x^j, y^j) + \beta(x^c, y^c)) \} \right), \quad (3.20)$$

where $F_D(\cdot)$ is a decoding function of period cue, as will be described in Sec. 3.3. The function $F_D(\cdot)$ may vary with the methods of coding period information as long as the signal $D(x^c, y^c)$ can be extracted. The high frequency signal can be then unwrapped by referring to the period cue.

Pattern entropy

In Eq. (3.18), the added $\{C_n\}$ will result in increased distinctions among the illuminations of periodical points $\{I_n(\phi^w) : I_n(\phi^w) = I_n^b(\phi^w) + C_n(\phi^w)\}$. For example, the projected intensities of E_1 and E_2 in Fig. 3.1 will no longer be the same, if two

Table 3.1: Comparison of Spatial Intensity Efficiency and Pattern Entropy.

Pattern Strategy	η	H	Remark
High frequency PMP [70, 74]	63.80%	14.00 <i>bit</i>	4 patterns, 4 periods
Fast three-step phase-shifting [101]	82.71%	13.87 <i>bit</i>	3 patterns, 4 periods
2 + 1 method [88, 102]	59.39%	13.81 <i>bit</i>	3 patterns, 4 periods
Three patterns trapezoidal [63]	100%	13.71 <i>bit</i>	3 patterns, 4 periods
Two-step triangular-pattern [81]	50.31%	13.67 <i>bit</i>	2 patterns, 4 periods

different sets of $\{C_n\}$ are added. Indeed, the secondary information is extracted from the differences among the illuminations of periodical points. Coding information $\{C_n\}$ can be added without losing the *SNR* of the original signal as long as $\sum_{n=0}^{N-1} a_n I_n^c(x^c, y^c)$ and $\sum_{n=0}^{N-1} b_n I_n^c(x^c, y^c)$, in Eq. (3.19), can cancel out the terms of $\{C_n\}$. Multiple solutions of a coded $\{C_n(x^c, y^c)\}$ exist, if $\Delta_n(\phi^w) > P$. So after embedding the period cue, an optimization procedure should be performed to maximize the distinctions among points with the same phase value $\{I_n(\phi^w)\}$. The maximum distinctions among these NP illumination points, in $\{I_n(\phi^w)\}$, would be achieved when $\{I_n(\phi^w)\}$ are uniformly distributed in the range of $[0, 1]$.

In order to measure and maximize the differences, we propose the definition of pattern entropy for phase shifting methods. The pattern entropy is a summation of entropies of $\{I_n(\phi^w)\}$ over $\phi^w \in [0, 2\pi)$, and it is employed to maximize the information content of the embedded secondary signal. In information theory, entropy is a measure of the uncertainty associated with a random variable [103]. In order to extend the entropy analysis to phase shifting methods, we first define the probability mass function (pmf) of the intensity values. We assume the points with different wrapped phase values are independent. For the periodical points, we first sort the projected intensity values such that $\{I_n(\phi^w) : 0 \leq I_0(\phi^w) \leq I_1(\phi^w) \dots \leq I_{NP-1}(\phi^w) \leq 1\}$. By

assuming a uniform distribution in $[\frac{I_{n-1}(\phi^w)+I_n(\phi^w)}{2}, \frac{I_n(\phi^w)+I_{n+1}(\phi^w)}{2}]$, the possibility at intensity value, $\frac{I_{n-1}(\phi^w)+I_n(\phi^w)}{2} \leq x \leq \frac{I_n(\phi^w)+I_{n+1}(\phi^w)}{2}$, is given by

$$p(x) = \sum_{I_n(\phi^w)=x} \frac{2}{NL(I_{n+1}(\phi^w) - I_{n-1}(\phi^w))}. \quad (3.21)$$

The range of intensity values is changed from $[0, 1]$ to $[-\frac{1}{2NP-2}, 1 + \frac{1}{2NP-2}]$, and the uniform distribution in $[0, 1]$ can achieve the highest pattern entropy. If $-\frac{1}{2NP-2} \leq x \leq \frac{I_0(\phi^w)+I_1(\phi^w)}{2}$, then

$$p(x) = \frac{2(NP - 1)}{NL(1 + (NP - 1)(I_0(\phi^w) + I_1(\phi^w)))}. \quad (3.22)$$

And if $\frac{I_{NP-2}(\phi^w)+I_{NP-1}(\phi^w)}{2} \leq x \leq 1 + \frac{1}{2NP-2}$, then

$$p(x) = \frac{2(NP - 1)}{NL(2NP - 1 - (NP - 1)(I_{NP-1}(\phi^w) + I_{NP-2}(\phi^w)))}. \quad (3.23)$$

So the pattern entropy, H , is defined as

$$H = - \sum_{\phi^w=0}^{2\pi} \sum_{x=-\frac{1}{2NP-2}}^{1+\frac{1}{2NP-2}} p(x) \log_2 p(x). \quad (3.24)$$

The pattern entropy is an integral over all the phase values. Among the strategies with the same number of patterns, the higher pattern entropy indicates that the intensities of the periodical points are closer to a uniform distribution. By maximizing the pattern entropy, the distinctions among pixels with the same phase value $\{I_n(\phi^w)\}$ would be maximized and, therefore, the content of the added information $\{C_n\}$ is optimized.

In Table 3.1, several commonly used phase shifting methods are listed with the values of pattern entropy, assuming a pattern length of 1,280 pixels on an 8-bpp projector. Pattern entropy measures the information in the projected patterns. In

Table 3.1, high frequency PMP achieves the highest pattern entropy since 4 patterns are used in this strategy. Not surprisingly, the two-step triangular pattern strategy has the lowest pattern entropy.

3.2 Codification of projected patterns

In this section, we consider how the signals are coded into patterns in the projector space. For PMP, the spatial intensity efficiency is 63.80%, leaving a 36.20% margin of the available intensity dynamic range for adding a period cue signal. In order to make a full use of the available dynamic range without a sacrifice in SNR , the proposed pattern strategy consists of two codification stages: (1) phase codification which introduces a triangular phase waveform, instead of saw-tooth, and encodes the high frequency phase ϕ^w , and (2) period codification which implements Eq. (3.18) and encodes the period cue D of the high frequency phase. In traditional PMP, the saw-tooth high frequency phase ϕ^w is encoded into the sinusoidal intensities $\{I_n^b\}$, as defined in Eq. (3.8). Similarly, our triangular phase, noted as ϕ^w , is extracted from intensities $\{I_n^b\}$, as defined in Eq. (3.19), whereas the period cue D is obtained from intensities (or coding functions) $\{C_n\}$, as defined in Eq. (3.20).

Triangular phase

Traditional high frequency phase obtained from Eq. (3.9) uses only temporal intensity information, and its waveform is sawtooth [70, 74]. However, the small neighboring information is actually also reliable and available in the phase image. By employing such small neighboring spatial information, the phase frequency can be further im-

proved. Thus, the reconstruction error will be reduced. In order to take advantage of the spatial information, we propose a triangular waveform for the high frequency phase such that the wrapped phase from Eq. (3.9) is first decoded into traditional sawtooth waveform by utilizing a small portion of spatial information. This procedure is for an initial coding. With a corresponding de-codification, it can provide a doubled phase frequency.

Different from the traditional sawtooth high frequency phase, we employ a set of sinusoidal wave patterns, $\{I_n^b : n = 0, 1, \dots, N - 1\}$, as the base patterns, such that each point in the projector plane, (x^j, y^j) , is given by

$$I_n^b(x^j, y^j) = A + B \sin(\Phi(y^j) - \frac{2\pi n}{N}), \quad (3.25)$$

if p is even, and

$$I_n^b(x^j, y^j) = A + B \sin(\Phi(y^j) - \frac{2\pi n}{N}), \quad (3.26)$$

if p is odd, where A and B are set to 0.5, same as traditional PMP patterns. The term, $\Phi(y^j)$, is given by

$$\Phi(y^j) = 2\pi P \frac{y^j}{L}, \quad (3.27)$$

where L is the pattern length. Note the dependence of the phase term, $\Phi(y^j)$ on y^j , which is assumed to be positioned vertically above/below the projector. When $L = 1280$ and $P = 4$, the created triangular phase is shown as the solid lines in Fig 3.2.

Given the captured images, $\{I_n^c : n = 0, 1, \dots, N - 1\}$, the received phase of the sinusoidal wave at the camera coordinate (x^c, y^c) is calculated from Eq. (3.9). To solve Eq. (3.9), at least 3 patterns are needed, and if P in Eq. (3.27) is

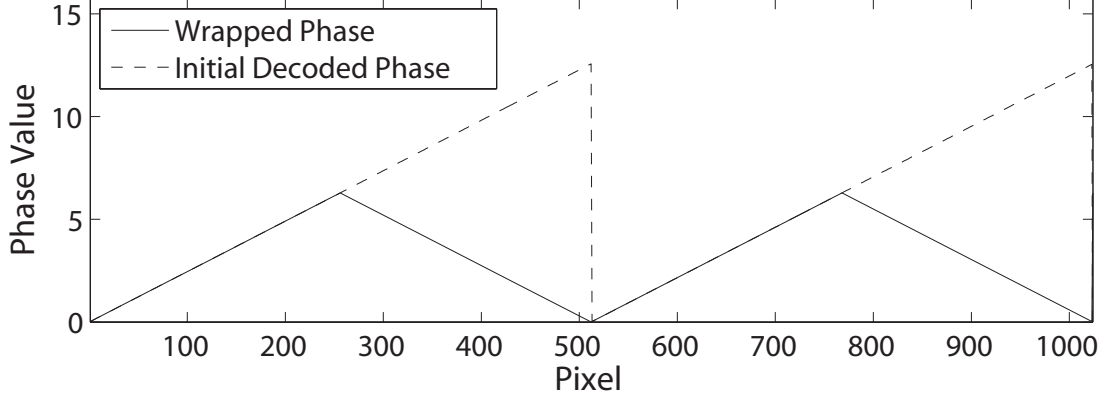


Figure 3.2: Captured phase is unwrapped into initial decoded phase by judging the location of phase value.

1 (unit frequency), then $\Phi(x^c, y^c) = \phi^w(x^c, y^c)$. The 3-D world coordinates of a point can be calculated from $(x^c, y^c, \Phi(x^c, y^c))$. However, when the number of periods, P , is greater than 1, ambiguities in phase will appear as the phase values $\Phi_{even}(x^c, y^c) = 2p(x^c, y^c)\pi + \phi^w(x^c, y^c)$ ($p(x^c, y^c) = \{1, 3, 5, \dots\}$) and $\Phi_{odd}(x^c, y^c) = 2p(x^c, y^c)\pi - \phi^w(x^c, y^c)$ ($p(x^c, y^c) = \{2, 4, 6, \dots\}$), in Eq. (4.11), all have the same value $\phi^w(x^c, y^c)$ obtained from Eq. (3.9). Figure 3.2 shows that the wrapped phase ϕ^w can be unwrapped into initial decoded phase ϕ^i by judging the position of ϕ^w .

After the initial decoding, ϕ^i is scaled to $[0, 2\pi)$ such that the error in phase is reduced during the scaling. Thus, the triangular phase ϕ^w is initially decoded and scaled to ϕ^i with traditional sawtooth waveform. By using a small portion of spatial information to determine whether the phase is increasing or decreasing, triangular phase encoding provides a doubled-frequency from the point of view of the imaging sensor, compared with the traditional sawtooth phase encoding. We denote P^i as the number of initial decoded periods, where $P^i = \frac{P}{2}$. For $P \geq 4$, $P^i \geq 2$ which means that ambiguities still exist. The high frequency ϕ^i needs to be further decoded into

unit frequency signal Φ by referring the period cue.

Period codification

As discussed in Sec. 3.1 for phase shifting methods with the spatial intensity efficiency less than 100%, we intend to embed a period cue signal into phase-shifting SLI patterns within the remaining dynamic range without any reduction in the SNR . We develop pattern codification as a means of removing ambiguities in initial decoded phase by encoding the period cue $p^i = \{0, 1, \dots, (P^i - 1)\}$ into the SLI patterns such that each point in the projector plane, (x^j, y^j) , is given by

$$I_n(x^j, y^j) = I_n^b(x^j, y^j) + C_n(y^j), \quad (3.28)$$

where $I_n^b(x^j, y^j)$ is defined in Eq. (4.11) and $\{C_n : n = 0, 1, \dots, (N - 1)\}$ represents the coding functions. Due to these added signals, the projected patterns are no longer sinusoidal wave patterns.

In an N pattern strategy, the number of coding functions is also N for each point (x^j, y^j) . We need to define a parameter, period cue $D(y^j)$, in the projector space, which is derived from the temporal values $\{C_n(y^j)\}$ and can be employed to identify the period $p^i(y^j) = \{0, 1, \dots, (P^i - 1)\}$. The decoding filter function, $F_D(\cdot)$, can be implemented in various forms as long as it translates the N temporal values $\{C_n(y^j)\}$ into a single value $D(y^j)$. For the phase values $\Phi(y^j) = 2\pi p^i(y^j) + \phi^i(y^j)$ ($p^i(y^j) = \{0, 1, \dots, (P^i - 1)\}$), $D(y^j)$ varies with respect to $p^i(y^j)$ such that the period number can be obtained from the value of $D(y^j)$.

In practice, for a pattern strategy with the spatial intensity efficiency less than

100%, the coding functions $\{C_n(y^j)\}$, in Eq. (3.28), should have the following basic properties:

Property 1: *Absence in phase*: the calculation of phase value, in the camera space, by means of Eq. (3.9) should not be affected by adding the coding functions $\{C_n(y^j)\}$, in order to achieve an accurate 3-D reconstruction based on phase.

Property 2: *Dynamic range of coding*: the illumination patterns, $\{I_n(x^j, y^j)\}$, in the projector space (Eq. (3.28)), should stay within the dynamic range of the projector, $[0, 1]$; otherwise, the patterns have to be scaled into $[0, 1]$, which will result in a reduced *SNR* of the high frequency phase.

Property 3: *Maximum index probability*: the information of period cue $D(y^j)$ should maximize the probability of correctly indexing the period number $p^i(y^j)$ in the presence of system noise. That is, the $D(y^j)$ values for different periodical points (points with the same $\phi^i(y^j)$ values) should be as much distinct as possible, to prevent false estimation of $p^i(y^j)$ under noisy conditions.

Property 4: *Maximum pattern entropy*: the illumination patterns, $\{I_n(x^j, y^j)\}$, should achieve maximum pattern entropy such that the signal strength of period cue $D(y^j)$, realized by adding the coding functions $\{C_n(y^j)\}$, is maximized.

In order to devise such a coding method that satisfies the above properties, we will develop, first, an $N = 4$ pattern strategy where, as specified by Prop. (1), the values of U and V in Eqs. (3.10) and (3.11) will remain unchanged by adding $\{C_n(y^j)\}$ if

$$\sum_{n=0}^{N-1} C_n(y^j) \cos\left(\frac{2\pi n}{N}\right) = 0, \quad (3.29)$$

and

$$\sum_{n=0}^{N-1} C_n(y^j) \sin\left(\frac{2\pi n}{N}\right) = 0. \quad (3.30)$$

For $N \geq 3$, solutions to Eqs. (3.29) and (3.30) always exist because the number of unknown parameters is less than the number of equations. Particularly since $N = 4$, $C_0(y^j) = C_2(y^j)$ and $C_1(y^j) = C_3(y^j)$ will satisfy the property of *absence in phase*. Thus, we define the period cue, $D(y^j)$, for $N = 4$ as

$$D(y^j) = \frac{C_0(y^j) + C_2(y^j) - C_1(y^j) - C_3(y^j)}{2}, \quad (3.31)$$

which is a linear function of $\{C_n(y^j)\}$.

For $N = 4$, Prop. (2) should be satisfied because: (1) the intensity values can not exceed $[0, 1]$ and (2) the dynamic range of the signals for computing high frequency phase should not be reduced. That is, B remains unchanged. Thus for a pixel $I_n(x^j, y^j)$ of illumination patterns, the coding functions $\{C_n(y^j)\}$ must satisfy

$$C_n(y^j) \geq 0 - I_n^b(x^j, y^j), \quad (3.32)$$

and

$$C_n(y^j) \leq 1 - I_n^b(x^j, y^j). \quad (3.33)$$

From Eq. (3.20), the bounds of $\{C_n(y^j)\}$ will limit the value range of period cue $D(y^j)$, which we denote as $R(y^j)$ with supremum bound $\sup\{R(y^j)\}$ and infimum bound $\inf\{R(y^j)\}$. Figure 3.3 shows the bounds of $D(y^j)$ when $N = 4$ and $P = 4$. As shown for different $\Phi(y^j)$, $R(y^j)$ is not the same. So to simplify the problem, we set our bounds to a constant value R inside the theoretical bounds $R(y^j)$.

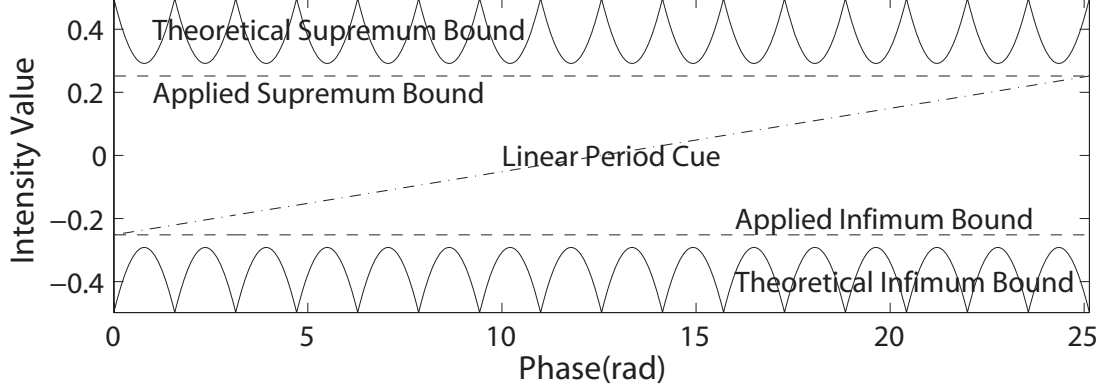


Figure 3.3: The dynamic range of linear period cue, D , when $N = 4$ and $P = 4$.

With regard to Prop. (3), the period cue pattern can be of any shape as long as the cue values can be uniquely obtained from the received patterns, and for the phase values $\Phi(y^j) = 2p^i(y^j)\pi + \phi^i(y^j)$, ($p^i(y^j) = \{0, 1, \dots, P^i - 1\}$), the period cue values must be unique. So for purposes of coding period, the cue patterns should be continuous to avoid estimation errors, especially, when the resolution of camera is higher than that of the projector. In addition, the values of $D(y^j)$ should be evenly spread in R . So, we choose a linear period cue for the coding period in the $N = 4$ pattern strategy such that

$$D(y^j) = \frac{(\sup\{R\} - \inf\{R\})y^j}{L} + \inf\{R\}. \quad (3.34)$$

A cross section of this linear period cue is shown in Fig. 3.3 while several cue values for different phase values are list in Table 3.2. As shown, the distinction in cues for all the phase values is maximized, given the limited dynamic range of cues. Despite the presence of noise, the probability of correctly identifying the period is maximized.

Through the above three properties, the values of period cues, $D(y^j)$, can be determined, but the period cues have not directly been embedded into the patterns.

Table 3.2: Linear period cue values when N is 4 and P is 4.

Period p^i	Cue values for $\Phi_1 = 2p^i\pi + 0$	Cue values for $\Phi_2 = 2p^i\pi + \frac{2\pi}{3}$	Cue values for $\Phi_3 = 2p^i\pi + \frac{4\pi}{3}$
0	-0.252	-0.084	0.084
1	0	0.168	-0.084

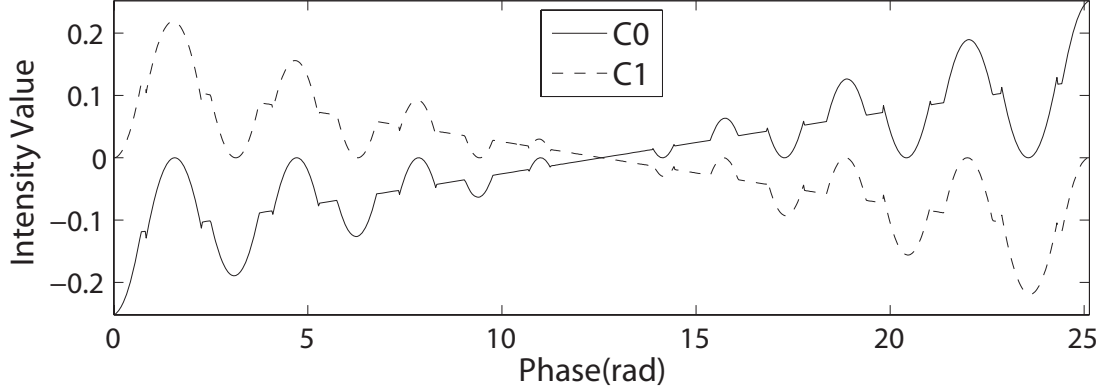


Figure 3.4: Created coding functions for an $N = 4$ and $P = 4$ strategy, where $C_0 = C_2$ and $C_1 = C_3$.

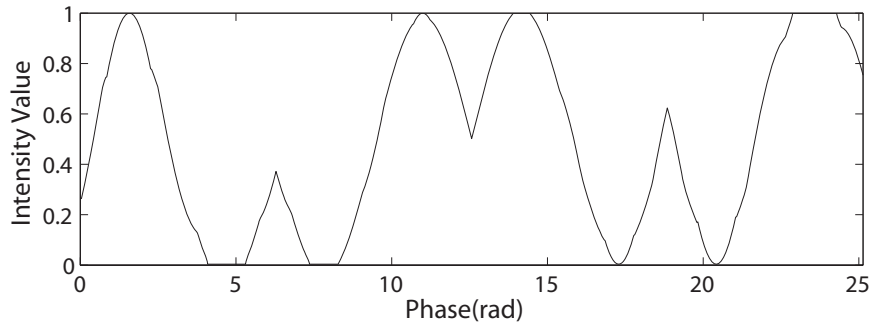
Instead, they are derived from the function $F_D(\cdot)$ in Eq. (3.20). The coding functions $\{C_n\}$ should be designed to achieve the values of D , just as ϕ^w is achieved by designing $\{I_n^b\}$. In order to increase the SNR of ϕ^w , the amplitude B , in $\{I_n^b\}$, is maximized. Similarly, the pattern entropy should be maximized to increase the SNR of D (i.e., Property (4)), but there are restrictions from Eqs. (3.29) and (3.30) for designing $\{C_n\}$. Also, the relation between D and $\{C_n\}$, given in Eq. (3.20), along with the D values given in Eq. (3.34), represent a third restriction on $\{C_n\}$.

The solution, given the above constraints, is not unique for $N \geq 4$ pattern strategies because there are more unknown parameters, $\{C_n\}$, than the number of restrictions. So in order to derive a unique solution, we employ the pattern entropy, of Eq. (3.24), to optimize the patterns such that

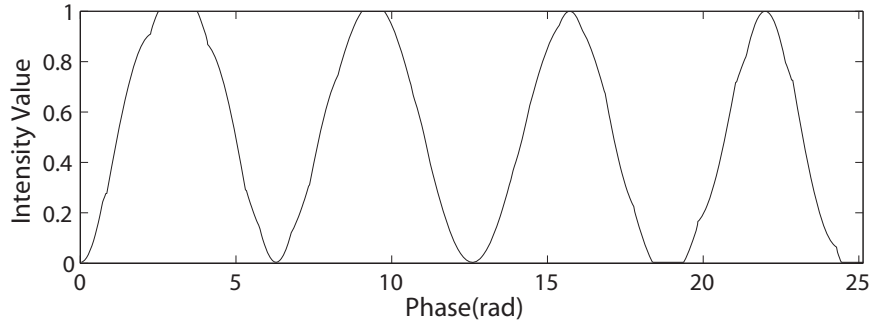
$$\{C_n\} = \arg \max_{\{C_n\}_k} H(\{I_n^b + C_n\}_k). \quad (3.35)$$

In this way, $\{C_n\}$ can be optimized to achieve maximum distinctions among the projected NP pixels with the same phase value ϕ^w . It should be noted that the Property (4) can be modified for other purposes. For example, in order to reduce the sensitivity to de-focus, the coding functions achieving the most smooth patterns (no inflection points) can be chosen. The coding functions $\{C_n\}$ for $N = 4$ pattern strategy when $P = 4$ is shown in Fig. 3.4.

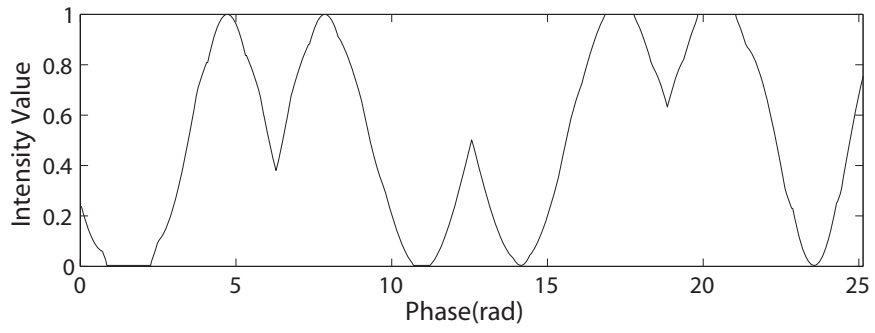
Finally, after obtaining the coding functions, SLI patterns are computed through Eq. (3.28), which take advantage of the remaining dynamic range and consist of the base patterns $\{I_n^b\}$ with the period cue D , achieved through coding functions $\{C_n\}$. For a pattern strategy where $N = 4$ and $P = 4$, the PCPS patterns are shown in Fig. 3.5, where cross-sections of the four patterns indicate that the patterns, in each period, are no longer sinusoidal waves due to the added coding functions. The spatial intensity efficiency and pattern entropy is listed in Table 3.3. Compared with $N = 4$ and $P = 4$ high frequency PMP, the spatial intensity efficiency is increased from 63.80% to 91.53%, and the pattern entropy is increased from 14.00 bits to 15.91 bits.



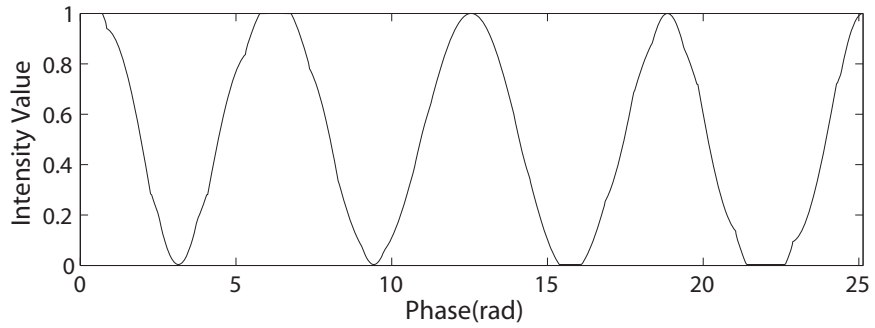
(a)



(b)



(c)



(d)

Figure 3.5: Cross sections of the PCPS patterns when $N = 4$ and $P = 4$ for (a) I_0 , (b) I_1 , (c) I_2 , and (d) I_3 .

Table 3.3: Comparison of spatial intensity efficiency and pattern entropy.

4 patterns 4 periods strategy	PMP	PCPS
Spatial intensity efficiency	63.80%	91.53%
Pattern entropy	14.00 <i>bit</i>	15.91 <i>bit</i>

3.3 De-codification of received images

Modulation and wrapped phase

In this section, we consider how the captured images are decoded in the camera space. After a set of illumination patterns described in Eq. (3.28) have been projected upon a target object, there are three types of information we need to obtain: (1) the modulation of scanned object, (2) the wrapped phase, and (3) the period cue. With regards to modulation, the parameter $M(x^c, y^c) = \alpha(x^c, y^c)B$, representing the amplitude of the observed high frequency signal reflected off the target, can be derived according to Eq. (3.12). The modulation M is used as the texture image for the 3-D reconstruction. The wrapped phase value, ϕ^w , is calculated by Eq. (3.9), where through the definition of U and V , it is clear that the coding functions $\{C_n\}$ will not affect ϕ^w and M .

In order to achieve real-time operation, there are two facts that should be noticed: (1) the computational cost of $\sin(\cdot)$, $\cos(\cdot)$, and $\arctan(\cdot)$ in Eqs. (3.9) and (3.12) is high and (2) the value range of received images inside the camera are fixed. For instance, given an 8-bit gray scale camera, illumination pattern $I_n^c(x^c, y^c)$ is represented as the integers in the range $[0, 255]$. Thus, it becomes possible for us to build look-up table (LUT) based algorithms for resolving both Eqs. (3.12) and (3.9) such that $M(x^c, y^c)$ and $\phi^w(x^c, y^c)$ are obtained with very low computational cost, and because

B in Eq. (4.11), is constants of the projector, the albedo $\alpha(x^c, y^c)$ is obtained by

$$\alpha(x^c, y^c) = \frac{M(x^c, y^c)}{B}. \quad (3.36)$$

Furthermore, the ϕ^w is initially decoded into ϕ^i as shown in Fig. 3.2.

De-codification of period cue

The purpose of decoding the period cue is to find the period $p^i(x^c, y^c)$ at the pixel (x^c, y^c) in the received images. To do so, we present two methods for this purpose:

(1) a temporal method, which decodes the period number point by point, and (2) a hybrid method, which decodes the period number with the help of spatial unwrapping approaches. The temporal period decoding attempts to identify the $p^i(x^c, y^c)$ based on the period cue value $D(x^c, y^c)$ of each point. It decodes the phase point by point with no dependence on neighboring points. The $D(x^c, y^c)$ value is calculated from Eq. (3.20), which may vary depending on the definition of $F_D(\cdot)$ in Eq. (3.20).

Particularly in a 4 pattern strategy, $D(x^c, y^c)$ is computed as

$$\begin{aligned} D(x^c, y^c) &= \frac{C_0^c(x^c, y^c) + C_2^c(x^c, y^c) - C_1^c(x^c, y^c) - C_3^c(x^c, y^c)}{2} \\ &= \frac{(I_0^c(x^c, y^c) + I_2^c(x^c, y^c)) - (I_1^c(x^c, y^c) + I_3^c(x^c, y^c))}{2\alpha(x^c, y^c)}, \end{aligned} \quad (3.37)$$

where $\alpha(x^c, y^c)$ is obtained from Eq. (3.36).

The parameter, $D(x^c, y^c)$, is only related to the coding functions $\{C_n(x^c, y^c)\}$. So once $\phi^i(x^c, y^c)$ and $D(x^c, y^c)$ are obtained as shown in Fig. 3.6, the period $p^i(x^c, y^c)$ is calculated by

$$p^i(x^c, y^c) = \text{round} \left(\frac{1}{2\pi} (2P^i\pi \frac{D(x^c, y^c) - \inf\{R\}}{\sup\{R\} - \inf\{R\}} - \phi^i(x^c, y^c)) - 0.5 \right), \quad (3.38)$$

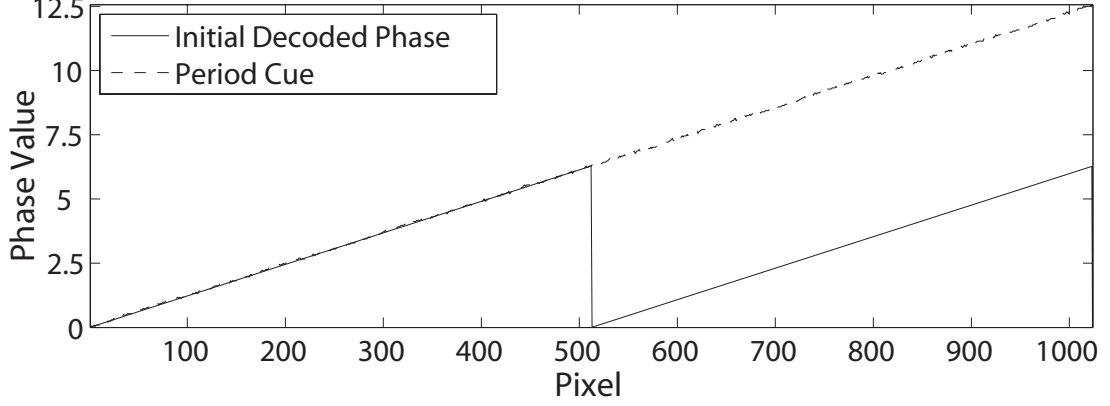


Figure 3.6: The initial unwrapped phase is finally unwrapped according to the period cue. The period cue is scaled into $[0, 2P^i\pi]$.

where \sup , \inf , and R are defined in Sec. 3.2, P^i is the constant number of periods after initial decoding and $\text{round}(\cdot)$ notes the rounding function of mapping the operand to the nearest integer value. Thus, the phase $\Phi(x^c, y^c)$ is obtained by

$$\Phi(x^c, y^c) = \phi^i(x^c, y^c) + 2\pi p^i(x^c, y^c), \quad (3.39)$$

which is used for reconstruction based on pre-calibrated triangulation. For the 4 patterns strategy, when $P = 4$, the period will be correctly decoded, in the presence of noise, as long as the absolute error in period cue, $e_D(x^c, y^c)$, is smaller than 0.126. That is when $D(x^c, y^c)$ is scaled to $[0, 4\pi]$ as shown in Fig. 3.6, the absolute error $e_D(x^c, y^c)$ is smaller than π .

We note that, when the period number P increases and since the R in Eq. (3.38) is unchanged, the points with incorrect decoded period increase. So instead of decoding the period temporally, the spatial unwrapping approaches [96,89,90] can be combined with temporal period cues for indexing the period with higher accuracy. As such, we propose hybrid de-codification where the initial decoded phase image is further processed line by line. As illustrated in Fig. 3.7 (left), the processed phase line is

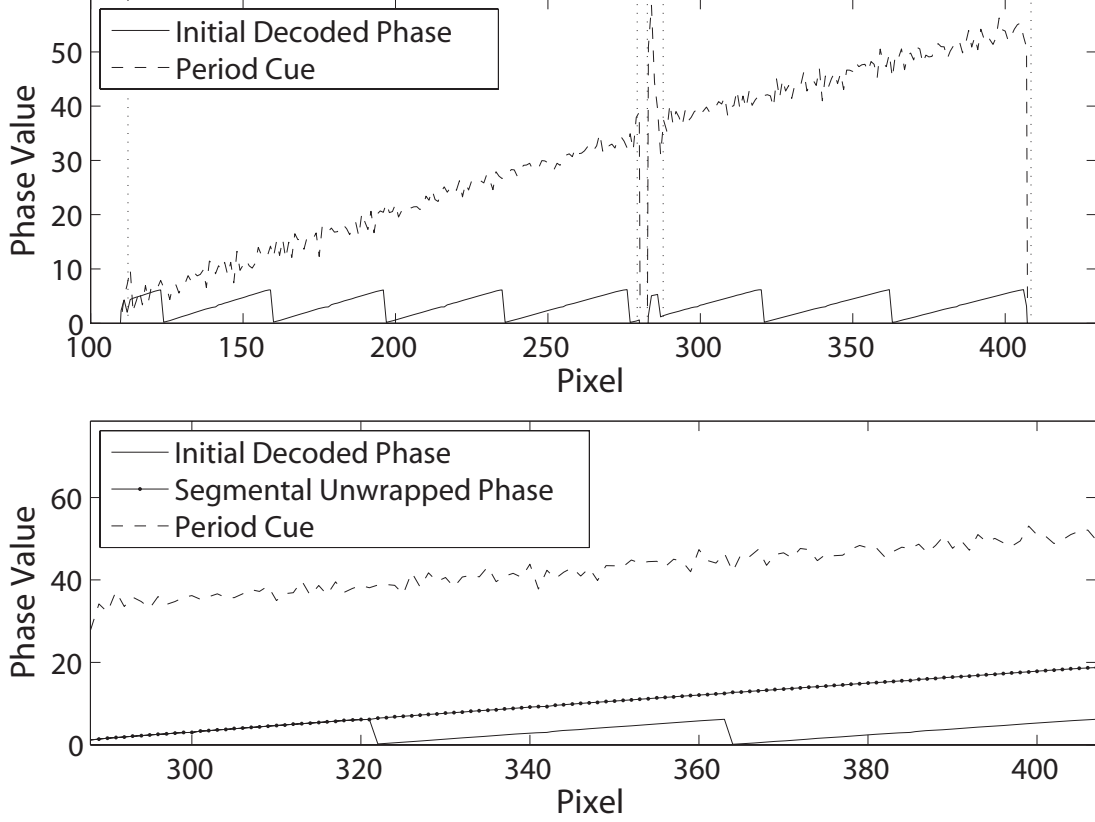


Figure 3.7: The (top) phase line is segmented according to the period cue line where period cue is scaled to $[0, 2P^i\pi)$. The (bottom) phase segment is unwrapped without knowing the start period of the segment such that distance exists between the unwrapped phase and the period cue line.

first segmented according to the corresponding period cue line. Then, the segment is unwrapped into segmental unwrapped phase, ϕ^s , as shown in Fig. 3.7 (right). Since the start period, in pixels 287 to 322, is unknown, there is a distance between the scaled period cue D and ϕ^s . So the start period p^i is computed as

$$p^i = \text{round} \left(\frac{1}{2\pi} \overline{\left(2P^i\pi \frac{D - \inf\{R\}}{\sup\{R\} - \inf\{R\}} - \phi^s \right)} - 0.5 \right), \quad (3.40)$$

which is similar to Eq. (3.38) except that $\overline{\left(2P^i\pi \frac{D - \inf\{R\}}{\sup\{R\} - \inf\{R\}} - \phi^s \right)}$ is the average distance between D and ϕ^s . This averaging reduces the degrading effect from noise and gives a correct period value p^i . Lastly, the final decoded phase, Φ , of the segment

is obtained by adding ϕ^s with $2p^i\pi$ as Eq. (3.39).

3.4 Accuracy analysis

In practice, uncertainty in $I_n^c(x^c, y^c)$ is introduced from camera noise [104] and projector noise [45], as well as other sources [87]. Since these effects degrade the quality of 3-D reconstructions, many studies have been performed on noise analysis to improve the signal to noise ratio through pattern strategy optimization [86]. In doing so, several researchers [74, 85, 86] have proposed modeling the combination of noise sources as additive, white, Gaussian noise, $w^c(x^c, y^c) \sim N(0, \sigma^2)$. Thus, the captured images, $I_n^c(x^c, y^c)$, in Eq. (3.2) can be expressed as

$$\tilde{I}_n^c(x^c, y^c) = \alpha(x^c, y^c) [I_n(x^j, y^j)] + \alpha(x^c, y^c)\beta(x^c, y^c) + w_n^c(x^c, y^c), \quad (3.41)$$

where $\tilde{\cdot}$ denotes the observed variable polluted by noise. With the same period p , the phase can be regarded as unit frequency, *i.e.* $P = 1$. The phase value with noise, $\tilde{\Phi}(x^c, y^c)$, is then calculated by

$$\begin{aligned} \tilde{\Phi}(x^c, y^c) &= \arctan \left[\frac{\sum_{n=0}^{N-1} \tilde{I}_n^c(x^c, y^c) \cos(\frac{2\pi n}{N})}{\sum_{n=0}^{N-1} \tilde{I}_n^c(x^c, y^c) \sin(\frac{2\pi n}{N})} \right] \\ &= \arctan \left[\frac{\frac{N}{2} M(x^c, y^c) \sin(\Phi(x^c, y^c)) + \sum_{n=0}^{N-1} w_n^c(x^c, y^c) \cos(\frac{2\pi n}{N})}{\frac{N}{2} M(x^c, y^c) \cos(\Phi(x^c, y^c)) + \sum_{n=0}^{N-1} w_n^c(x^c, y^c) \sin(\frac{2\pi n}{N})} \right] \end{aligned} \quad (3.42)$$

Since for real-valued numbers a and b with $ab > -1$,

$$\arctan(a) - \arctan(b) = \arctan \left(\frac{a - b}{1 + ab} \right), \quad (3.43)$$

then phase error Φ_e is obtained by

$$\Phi_e(x^c, y^c) = \Phi(x^c, y^c) - \tilde{\Phi}(x^c, y^c)$$

$$\begin{aligned}
&= \arctan \left[\frac{\sin(\Phi(x^c, y^c))}{\cos(\Phi(x^c, y^c))} \right] \\
&- \arctan \left[\frac{\frac{N}{2} M(x^c, y^c) \sin(\Phi(x^c, y^c)) + \sum_{n=0}^{N-1} w_n^c(x^c, y^c) \cos(\frac{2\pi n}{N})}{\frac{N}{2} M(x^c, y^c) \cos(\Phi(x^c, y^c)) + \sum_{n=0}^{N-1} w_n^c(x^c, y^c) \sin(\frac{2\pi n}{N})} \right] \\
&= \arctan \left(\frac{T(x^c, y^c)}{S(x^c, y^c)} \right), \tag{3.44}
\end{aligned}$$

where

$$S(x^c, y^c) = M(x^c, y^c) + \frac{2}{N} \sum_{n=0}^{N-1} w_n^c(x^c, y^c) \sin(\Phi(x^c, y^c) - \frac{2\pi n}{N}) \tag{3.45}$$

and

$$T(x^c, y^c) = \frac{2}{N} \sum_{n=0}^{N-1} w_n^c(x^c, y^c) \cos(\Phi(x^c, y^c) - \frac{2\pi n}{N}). \tag{3.46}$$

Further, it can be shown that $S(x^c, y^c)$ and $T(x^c, y^c)$ are Gaussian, and their correlation can be estimated by

$$E[S(x^c, y^c)T(x^c, y^c)] = \sigma^2 \sum_{n=0}^{N-1} \sin(2\Phi(x^c, y^c) - \frac{4\pi n}{N}) = 0. \tag{3.47}$$

Thus, $S(x^c, y^c)$ and $T(x^c, y^c)$ are independent. The joint Probability Density Function (PDF) of Φ_e and \tilde{M} is then given by

$$f_{\tilde{M}, \Phi_e}(\tilde{m}, \phi_e) = \frac{N\tilde{m}}{4\pi\sigma^2} \exp \left[-N \frac{(\tilde{m})^2 + M^2 - 2\tilde{m}(M \cos(\phi_e))}{4\sigma^2} \right], \tag{3.48}$$

as derived in [105]. Integrating Eq. (4.28) over \tilde{M} yields the marginal PDF for Φ_e given by

$$f_{\Phi_e}(\phi_e) = \frac{1}{2\pi} \exp \left[-\frac{NM^2}{4\sigma^2} \right] \left[1 + \kappa \sqrt{\pi} \exp(\kappa^2) (1 + \text{erf}(\kappa)) \right], \tag{3.49}$$

where $\text{erf}(\cdot)$ is the error function and

$$\kappa = \frac{\sqrt{N}}{2} \frac{M}{\sigma} \cos(\phi_e). \tag{3.50}$$

With sufficiently large values of $\frac{\sqrt{NM}}{\sqrt{2}\sigma}$ (>3), the error function of Eq. (4.30) will be close to 0. The term, $\kappa\sqrt{\pi}\exp(\kappa^2)(1+erf(\kappa))$, will, therefore, dominate the constant 1 such that Eq. (4.29) reduces to

$$\begin{aligned} f_{\Phi_e}(\phi_e) &\approx \frac{\sqrt{NM}\cos(\phi_e)}{2\sqrt{\pi}\sigma}\exp\left[\frac{-NM^2\sin^2(\phi_e)}{4\sigma^2}\right] \\ &\approx \frac{\sqrt{NM}}{2\sqrt{\pi}\sigma}\exp\left[\frac{-NM^2\phi_e^2}{4\sigma^2}\right], \end{aligned} \quad (3.51)$$

which can be considered a zero-mean, Gaussian distribution. Thus, the variance of Φ_e is approximated by

$$\sigma_{\Phi_e}^2(x^c, y^c) \approx \frac{2\sigma^2}{NM^2(x^c, y^c)}. \quad (3.52)$$

When using high frequency patterns, $\Phi(x^c, y^c)$ is decoded and scaled into $[0, 2\pi)$ before 3-D reconstruction. Thus, the phase error, $\Phi_e(x^c, y^c)$, is divided by P so that Eq. (4.38) becomes

$$\sigma_{\Phi_e}^2(x^c, y^c) \approx \frac{2\sigma^2}{NP^2\tilde{M}^2(x^c, y^c)}, \quad (3.53)$$

which is the same as that of the two frequency PMP strategy derived by Li [74] with only half the number of patterns. Also from Eq. (4.39), it should be noted that when P is less than $M(x^c, y^c)$, where $M(x^c, y^c)$ is generally larger than 64 on 8 bpp devices, increasing P will more efficiently reduce $\sigma_{\Phi_e}^2(x^c, y^c)$ than increasing $M(x^c, y^c)$.

3.5 Experimental results and discussion

To demonstrate the proposed pattern strategy, we developed a prototype SLI system shown in Fig. 4.17, based on an 8 bpp, monochrome, Prosilica GC640M, gigabit ethernet camera with 640×480 pixel resolution. The projector is composed of a

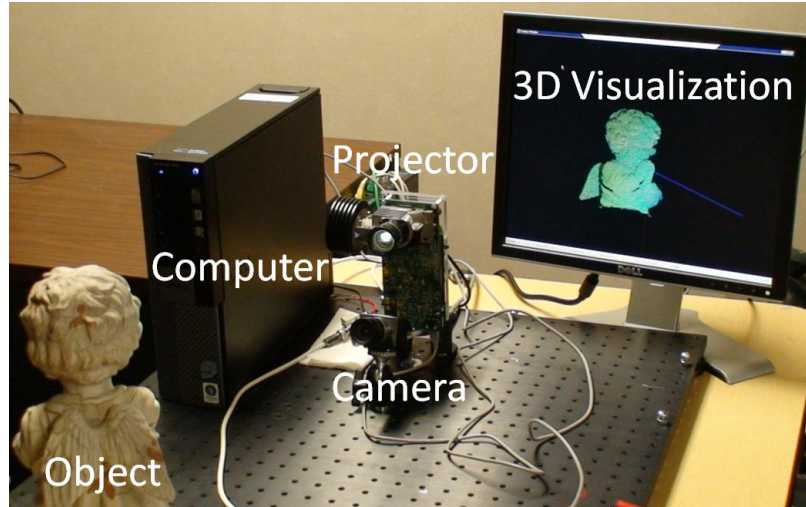


Figure 3.8: The prototype system setup for PCPS.

Texas Instrument’s Discovery 1100 board with ALP-1 controller and LED-OM with 225 ANSI lumens. The resolution of the 8 bpp, monochrome, projector is 1024×768 ($W \times L$), with a maximum frame rate of 150 fps. The camera and projector are synchronized by an external triggering circuit with a baseline distance between camera and projector of 120 mm. During the experiments, the scanned object was placed around 600 mm away. Gamma correction was performed on the received images, while a lookup table was created to correct optical distortion. We programmed the experimental system using Microsoft Visual Studio 2005 with managed C++. As our processing unit, we used a Dell Optiplex 960 with an Intel Core 2 Duo Quad Q9650 processor running at 3.0 GHz.

In the first two experiments, stationary objects were scanned with a camera exposure time of 2.4 ms. Due to the low illumination of our projector, the standard deviation of system noise, σ , was 7.2413, much higher than SLI systems using commercial projectors, *e.g.* 2.8309 in Li’s system [74]. In order to scan moving objects

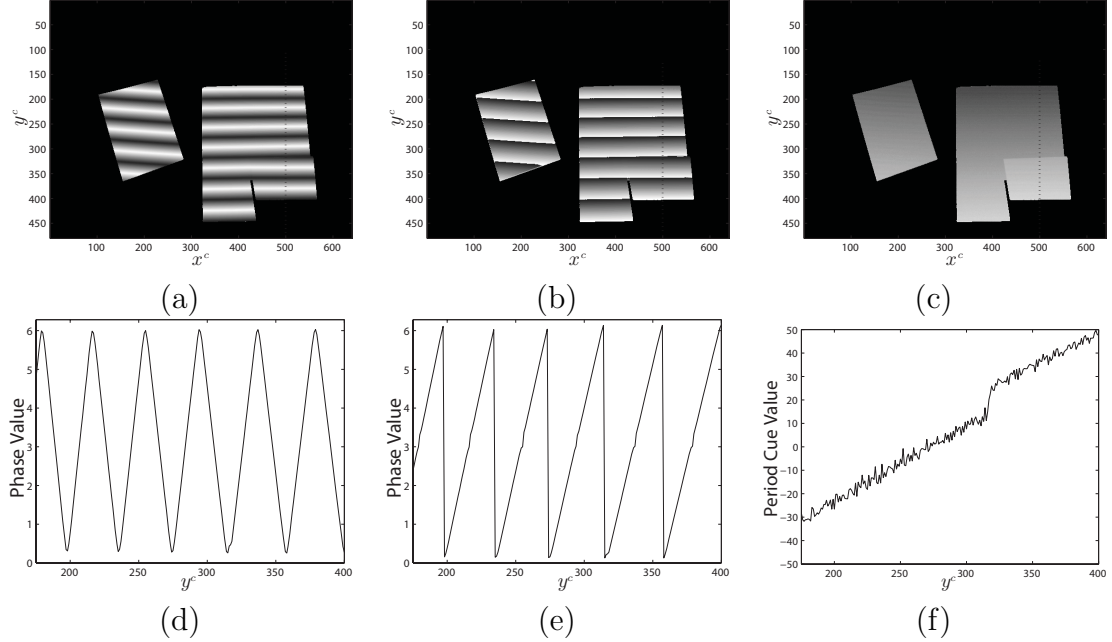


Figure 3.9: Three white boards were scanned using PCPS. (a) The wrapped phase image. (b) The phase image after initial decoding. (c) The period cue image. (d) The phase values on the 500th column of (a). (e) The phase values on the 500th column of (b). (f) The phase values on the 500th column of (c).

in our third experiment, the exposure time was reduced to 0.8 ms where σ increased to 9.9165.

Depth ambiguities

In the first experiment shown in Fig. 3.9, three separate white foam boards were carefully placed and scanned such that the left-most board was isolated from two boards otherwise positioned to create a phase ambiguity. In Fig. 3.9 (a) is the corresponding phase term, ϕ^w , while Fig. 3.9 (d) shows ϕ^w for the 500th image column, which dissects that ambiguous foam boards in half. Visually, there is no apparent discontinuity in phase to suggest that more than two distinct objects appear in the scene in either (a) or (d).

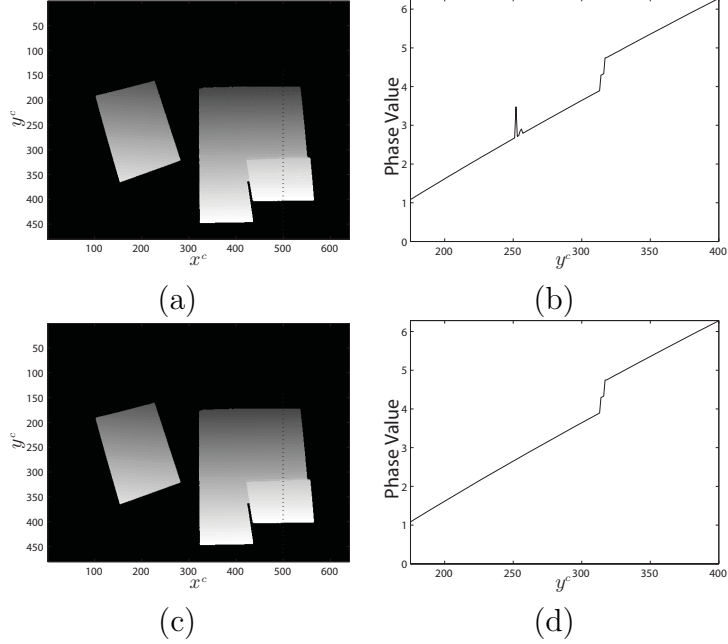


Figure 3.10: Final decoded phase results. (a) Temporal decoded phase. (b) The phase values on the 500th column of (a). (c) Hybrid decoded phase. (d) The phase values on the 500th column of (c).

After initial decoding by means of spatial phase unwrapping [93] in Fig. 3.9 (b) and (e), the phase values ϕ^i of the right two boards still appear to form a single, continuous surface; however, the depth discontinuity does show up in the period cue image of Fig. 3.9 (c) and (f), that is the D information. Thus, PCPS works to detect otherwise ambiguous phase, where it should be noted that because of the higher resolution of the projector versus that of camera and because of the de-focus problem [81], the received information should be low-pass filtered first. As a result, the inflection points of phase in Fig. 3.9 (c) and (f) are not as sharp as our theory predicts. In practice, the phases of these points can be corrected after obtaining the final decoded phase [106, 81, 63].

The final 3-D reconstruction is shown in Fig. 3.10, where because of the heavy

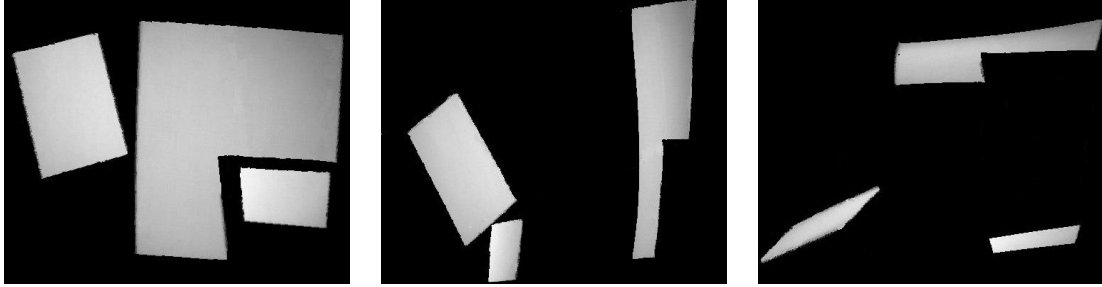


Figure 3.11: The 3-D reconstruction of three boards through PCPS using hybrid decoding method showing the (left) front, (middle) side, and (right) top views.

noise in our system ($\sigma = 7.2413$), the temporal period decoding method may produce wrong period numbers for some heavily contaminated points, *e.g.* the impulse in Fig. 3.10 (b). But even in our noisy system with $P = 16$, 99.56% points were correctly decoded through the temporal approach. By implementing the hybrid method, however, all the points were correctly decoded (Fig. 3.10 (c) and (d)). After having obtained these final phase images by means of the hybrid method, 3-D surfaces were reconstructed as illustrated in Fig. 3.11. The side and top views, of the reconstructed surfaces, demonstrate that the proposed PCPS works correctly in situations like Fig. 1.16 where spatial unwrapping fails [93] and temporal approaches only work by projecting additional patterns [96].

Accuracy

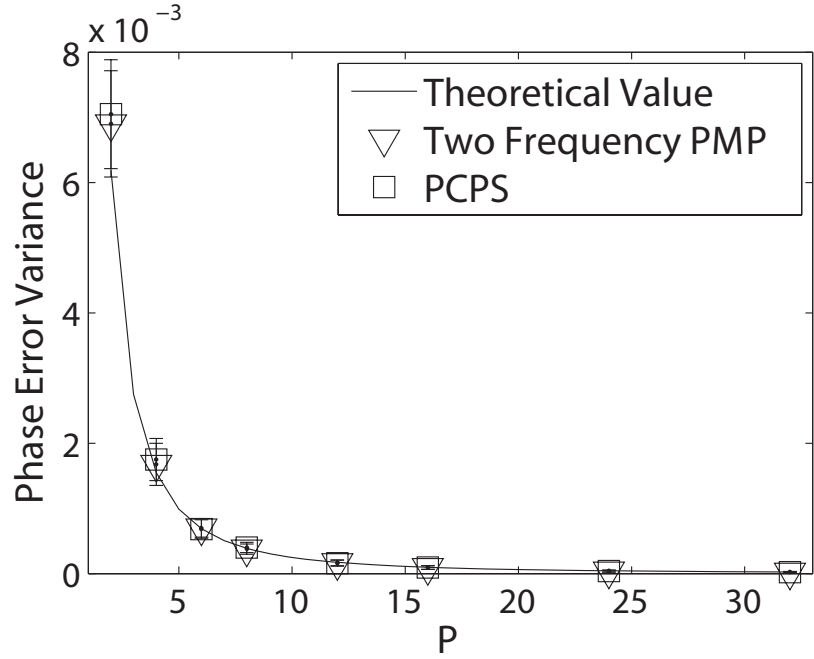
In order to demonstrate the accuracy of the PCPS in the presence of noise and object texture, we performed the second experiment by scanning a textured, flat, poster board (Fig. 3.12). The approximate size of the board is 210×175 mm (H×L). The front view and the top view of the 3-D reconstruction are shown in Fig. 3.12, where phase was obtained through 4 pattern 16 period PCPS using the hybrid decoding



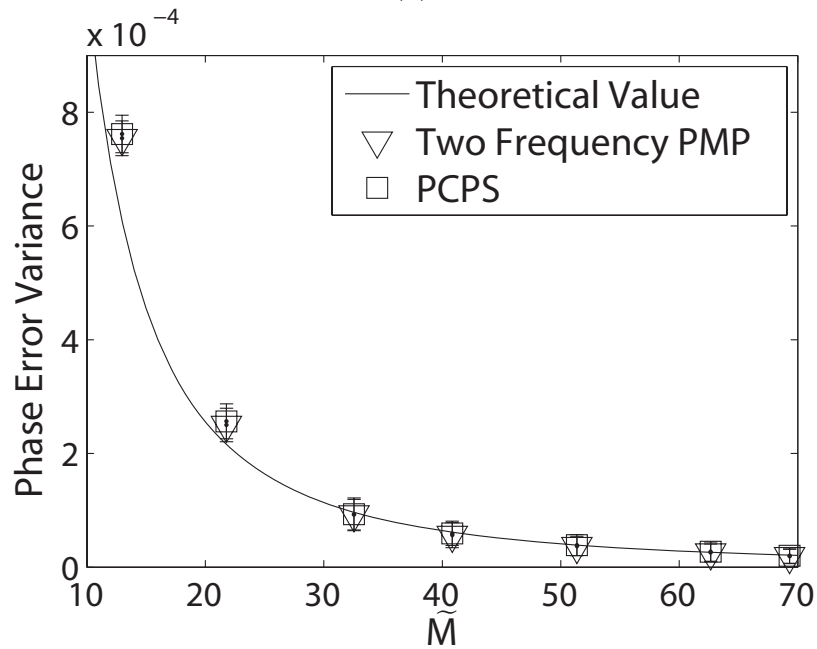
Figure 3.12: The 3-D reconstruction of a flat board with intensity texture, through 4 pattern, 16 period PCPS using hybrid decoding method showing the (top) front and (bottom) top view.

method. The texture was the modulation image obtained from Eq. 3.12.

From Eq. (4.39), the variance of phase error, $\sigma_{\phi_e}^2$, increases linearly with system noise, σ^2 , and decreases with the number of patterns N , the number of periods P , and the modulation M . This relationship is also true for PMP strategies as demonstrated in [74]. Thus, in theory the accuracy of PCPS is the same as PMP when scanning stationary objects, and to compare the accuracy of PCPS and PMP through experiments, the textured board was scanned 800 times using PCPS and two-frequency PMP. The theoretical value was obtained from Eq. (4.39), and as demonstrated, the experimental variances of phase error of PCPS and two frequency



(a)



(b)

Figure 3.13: (a) The phase error variance changes with the number of periods P . The number of patterns was 4. The standard deviation of system noise was 7.2413. The modulation was 32.55. (b) The phase error variance changes with modulation M . The number of patterns was 4. The standard deviation of system noise was 7.2413. The number of periods was 16.



Figure 3.14: Realtime 3-D reconstruction employing proposed PCPS strategy. The frame rate achieves 120 fps.

PMP are very close for all numbers of periods P (see Fig. 3.13 (left)) and modulation values \tilde{M} (see Fig. 3.13 (right)). However, it should be noted that 8 patterns were used in two frequency PMP strategy, whereas only 4 patterns were projected in PCPS. Thus, PCPS achieves the same accuracy as high frequency PMP by using only half the number of patterns.

Execution performance

In the third experiment, a stationary, textured, statue and a moving hand were scanned as shown in Fig. 3.14, and the execution performance of PCPS was analyzed. In order to achieve a high speed of scanning, the exposure time of the camera was reduced to 0.8 ms, resulting in an increased variance of 9.9165. Using lookup table (LUT) 3-D reconstruction techniques, the system realizes 120 fps for 3-D acquisition and reconstruction. Here, we projected the 4 pattern PCPS set. Since high frequency phase shifting patterns are more sensitive to motion, we employed 4 periods for motion scanning. The decoding method applied was temporal period decoding. As shown

Table 3.4: Analysis of Execution Performances.

Function	Time cost	Remark
Wrapped phase and modulation	1.72 <i>ms</i>	$\phi^w + M$
Period cue	0.59 <i>ms</i>	D
Initial decoded phase	0.96 <i>ms</i>	ϕ^i
Final decoded phase	1.24 <i>ms</i>	Temporal period decoding
Overall data processing	7.87 <i>ms</i>	Include 3-D reconstruction

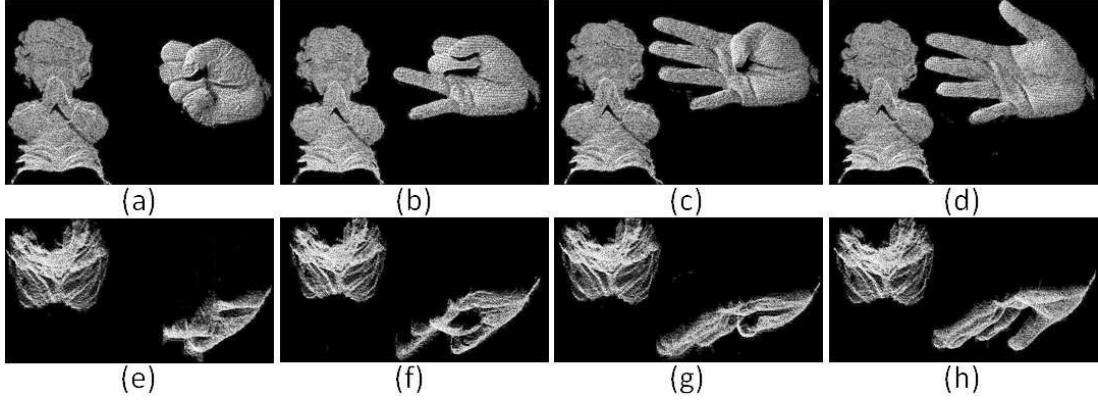


Figure 3.15: (a), (b), (c) and (d) are depth rendering 3-D reconstructions of a stationary textured angel and a moving hand. (e), (f), (g) and (h) are top views of (a), (b), (c) and (d). The standard deviation of system noise was 9.9165. No filter was applied.

in Fig. 3.15, the two separate surfaces were successfully reconstructed in real time.

The results look noisy in the high noise environment of the prototype system, as no smoothing process was performed to reduce the computational cost.

To observe the experimental time cost, we implemented the algorithm in our system by using only one CPU core. While points corresponding to low modulation strength were not displayed, our tests were based upon processing the full resolution of 640×480 pixels. The processing time each important procedure, averaged over 10,000 frames, is listed in Table 3.4. The period decoding is by means of the temporal method. The computational cost of hybrid decoding depends on the scene, since segmentation is involved. However, the one dimensional segmentation is done by

determining the period cue values. Indeed, the hybrid method will not introduce much more computational workload. The computation frame rate can be further improved by ignoring the low modulation points (un-illuminated points) and using multiple cores or GPU programming. In our case, the computation speed is fast enough for the maximum frame rate of the camera/projector pair. For comparison, we also implement Li’s algorithm [74], while Zhang et al. reported a reconstruction frame rate of 25.56 fps with a resolution of 532×500 when employing quality-guided phase unwrapping and using GPU processing on nVidia Quadro FX 3450 [102]. The comparison result is listed in Table 3.5.

In summary, the proposed PCPS strategy uses the remaining 36.20% intensity dynamic range to add the second reference signal. As shown in the first experiment, the PCPS works in situations where traditional phase shifting methods either fail or require more patterns, with high system noise standard deviation (7.2413). In the second experiment, we demonstrated Eq. (4.39), showing that the proposed PCPS method achieves the same accuracy as high frequency PMP with half the number of patterns. In the last experiment, we implemented the PCPS strategy in a real-time system that can achieve 120 fps for 3-D acquisition and reconstruction. The analysis of execution performance shows that the computational cost of PCPS is considerably low. Finally, it should be noted that for the SLI systems with lower noise level than our prototype system, the quality of both phase and period cue will be improved, leading to more accurate unwrapped phase values.

Table 3.5: Comparison of Execution Performances.

Strategy	Time cost	Remark
Two-frequency PMP [74]	86.23 <i>ms</i>	Temporal unwrapping, resolution (640×480), one core of 3.0 GHz CPU
2 + 1, quality-guided method [102]	39.12 <i>ms</i>	Spatial unwrapping, resolution (532×500), nVidia Quadro FX 3450 GPU
PCPS	7.87 <i>ms</i>	Period cue decoding, resolution (640×480), one core of 3.0 GHz CPU

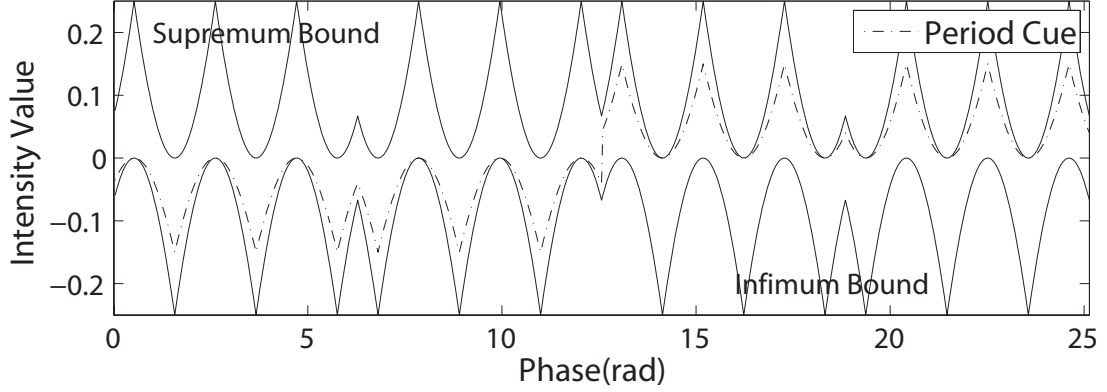


Figure 3.16: Nonlinear period cue D in 4 periods when the number of patterns is 3.

Generalization

For simplicity, we have presented the PCPS strategy and performed the experimental demonstration using 4 pattern PMP strategy. In fact, this new pattern strategy can be extended to any number of patterns as long as the pattern is obtained from Eq. (4.11) and satisfies the four coding properties in Sec. 3.2. As a phase shifting pattern strategy, the theoretical minimum number of patterns of PCPS is 3. Generally speaking, fewer patterns ($N \leq 6$) are desired to scan moving objects. With more patterns, the accuracy of phase can be increased as shown in Eq. (4.39), and it also gives a larger range of period cue D such that a more accurate period number can be decoded.

As the first example for $N = 3$, the base pattern is derived from Eq. (4.11). From Eqs. (3.29) and (3.30), the coding functions have $C_0(y^j) = C_1(y^j) = C_2(y^j)$. Thus, we define Eq. (3.20) as

$$D(y^j) = \frac{C_0(y^j) + C_1(y^j) + C_2(y^j)}{3}. \quad (3.54)$$

The period cue is designed as shown in Fig. 3.16, which is no longer linear. The

decoding process is similar to the 4 pattern strategy except that the environment illumination, $\beta(x^c, y^c)$, can not be removed from $D(x^c, y^c)$. Thus to reduce the effect of $\beta(x^c, y^c)$, the wrapped phase, $\phi^w(x^c, y^c)$, is first mapped to an initial decoded phase $\phi^i(x^c, y^c)$ where the points, with the same $\phi^i(x^c, y^c)$, are compared such that the $p^i(x^c, y^c)$ with higher $D(x^c, y^c)$ are assigned to 1 while the $p^i(x^c, y^c)$ with lower $D(x^c, y^c)$ are assigned to 0. The decoding process of 3 patterns is more complicated than that of $N > 3$ pattern strategies, and limited by the number of patterns, it becomes unreliable to decode patterns with $P > 4$ because of the environment illumination β . The 3-D reconstruction results of 3 pattern PCPS with $P = 4$ are shown in Fig. 3.17 (b).

As a second example, we present the pattern strategy when $N = 6$, where from Eqs. (3.29) and (3.30), the coding functions satisfy $C_1(y^j) + C_2(y^j) = C_4(y^j) + C_5(y^j)$ and $2C_0(y^j) + C_1(y^j) + C_5(y^j) = 2C_3(y^j) + C_2(y^j) + C_4(y^j)$, which provide even more coding freedom than 4 pattern strategy. Thus, it is possible to design more complicated period cues such that higher accuracy period decoding can be achieved. For simplicity, we define the period cue as

$$D(y^j) = \frac{2(C_0(y^j) + C_3(y^j)) - (C_1(y^j) + C_2(y^j) + C_4(y^j) + C_5(y^j))}{4}, \quad (3.55)$$

which is also a linear period cue. The decoding process is similar to the 4 pattern strategy, where because of the additional freedom, the range of $D(y^j)$ is increased. Comparison between Eqs. (3.37) and (3.55) shows that the denominator increases from 2 to 4 such that the period cue noise is reduced by half. The 3-D reconstruction results of 6 pattern PCPS with $P = 32$ is shown in Fig. 3.17 (d).

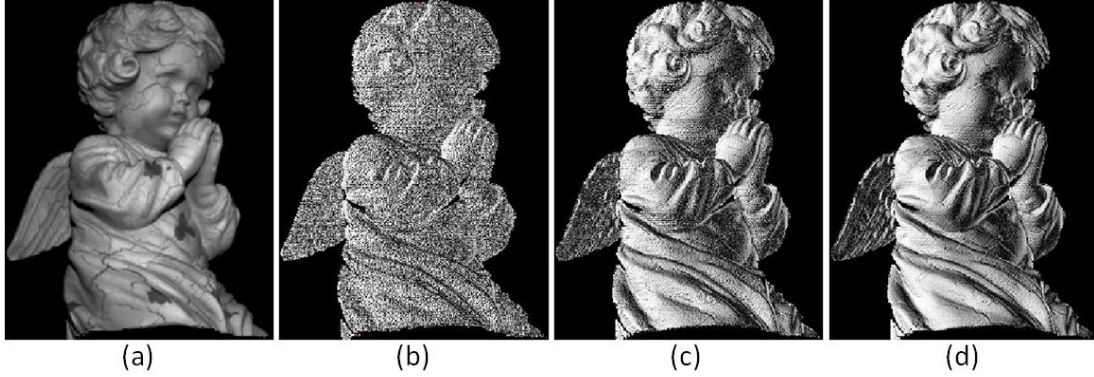


Figure 3.17: 3-D reconstructions for different number of patterns strategies. The standard deviation of system noise was 7.2413. No filter was applied. (a) Scanned Object. (b) The depth rendering result using 3 patterns 4 period PCPS. (c) The depth rendering result using 4 patterns 16 period PCPS. (d) The depth rendering result using 6 patterns 32 period PCPS.

Applying the period coded method to other phase shifting methods is similar to the procedures of PMP as long as our four stated properties are satisfied. It should be noted that, in this chapter, the SNR of the high frequency information is unchanged; however in practice, the projector amplitude value B , in Eq. (4.11), can be reduced to produce a reduction in SNR of the high frequency signal such that the M value, in Eq. (4.39), is reduced, but the smaller B value will leave more dynamic range for period codification. As a result, the P value, in Eq. (4.39), will be increased, and, based on Eq. (4.39), when the P value is less than M , the increasing of P will more efficiently reduce the phase error than increasing M . By reducing the projector modulation, the period information can be embedded into phase shifting methods even with the spatial intensity efficiency of 100%. The tradeoff between the dynamic ranges of the high frequency phase and the period cue indeed depends on the distribution of system noise.

This chapter presents two innovative concepts of spatial intensity efficiency and

pattern entropy for generalized phase shifting methods. When the spatial intensity efficiency is less than 100%, the remaining dynamic range of projection patterns can be used for embedding a secondary signal. The pattern entropy can be used to optimize the embedded secondary signal's strength. Based on these two concepts, we have developed a novel period information embedded pattern strategy (PCPS) for fast, reliable 3-D data acquisition and reconstruction. The whole strategy includes codification of projected patterns, de-codification of captured images, a real-time implementation, accuracy analysis, and execution performance analysis. A major advantage of the proposed PCPS strategy includes removing the depth ambiguity associated with traditional phase shifting patterns without reducing phase accuracy or increasing the number of projected patterns. The computational cost of PCPS is low and it can achieve 120 fps for 3-D acquisition and reconstruction. The data processing procedure for phase and modulation can be extended to other phase shifting methods. For future research, we will further exploit the remaining dynamic range such that the accuracy of period cue and the number of periods in the 3 pattern strategy can be improved. We will also work on the optimization between the dynamic ranges of high frequency phase signal and the period cue.

Chapter 4 HYBRID FRAMEWORK

Besides the pattern innovations, as we discussed in last two previous chapters, in order to improve SNR and also reduce the reconstruction error cost by gamma and multiple light pass, various other solutions have been proposed based upon post-processing methods, as well as multi-view data fusion. There are many post-processing methods existing for extracting the sub-pixel location of stripes based upon Steger's curvilinear structure detection method [107]. However, Steger's method fails in certain cases, such as in the study by Sun et al. [108], wherein many artificial lines resulted.

Various efforts have also been made at multi-channel 3-D measurement techniques [109,110,111,112,113,114] that merge measurements from multiple cameras in order to improve SNRs and, hence, 3-D reconstruction quality for a given number of projected patterns. By having multiple cameras, the probability that correspondences arise by chance is reduced, thereby resulting in better 3-D measurements [112]. However, employing such a multiple camera 3-D system brings the problem of merging data from different views. As shown in Fig. 4.3, the 3-D reconstructions from different views are mis-aligned due the reasons analyzed in [115,116]. For achieving accurate correspondences in the presence of phase distortion, phase stereo matching, or space-time stereo, has been proposed as a multi-channel 3-D measurement technique, where triangulation is performed between the component cameras without inclusion of the projector [34,111,59]. Specifically, Scharstein and Szeliski employed phase to acquire highly precise and reliable ground truth disparity measurements accurately aligned

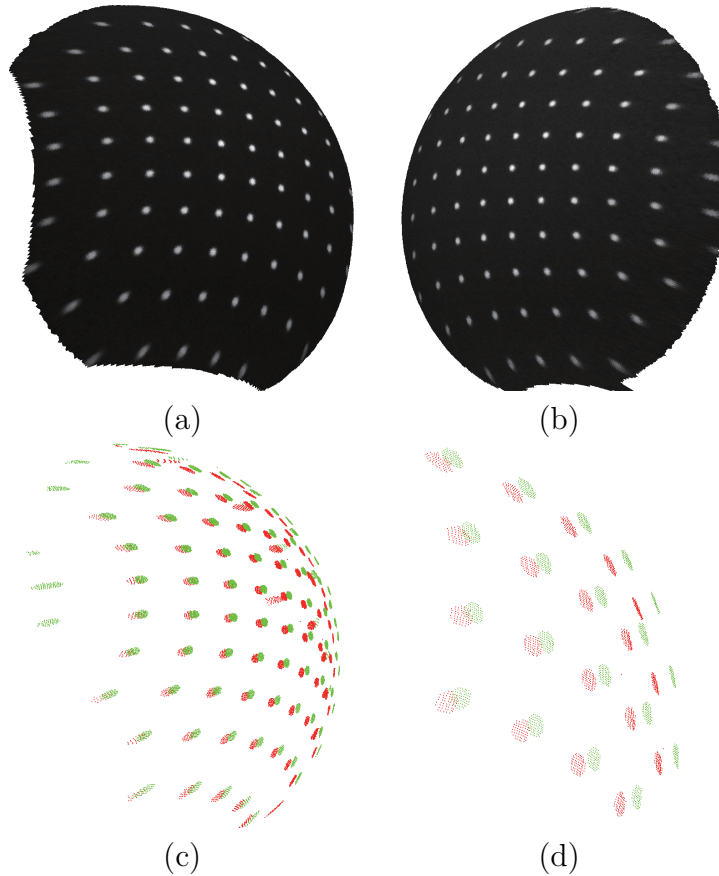


Figure 4.1: SLI 3D surfaces reconstructed using spot array patterns. Sixteen high frequency PMP patterns are used to ensure the phase data quality. (a) The 3D surface reconstructed through camera I and the projector. (b) The 3D surface reconstructed through camera II and the projector. (c) The resultant surface by merging two surfaces shown in (a) and (b). (d) An enlarged and rotated part of the surface shown in (c) to illustrate the misalignment.

with stereo image pairs [34].

The framework of space-time stereo was also developed in [111, 113, 114, 59, 58, 61, 60] where the advantages for high accuracy and motion detection were shown. In [111], Davis et al even proves that the framework would still work well by projecting unstructured light where illumination undergoes uncontrolled variation. The conclusion is due to the fact that the unstructured light eliminates the ambiguities when matching in space-time on the epipolar lines of two cameras. And by match-

ing based on space-time windows, they recovered dynamic scenes successfully. The advantage of the phase matching framework has also been presented by Zhang et al in [59, 58] who, based upon a synchronized multi-camera and multi-projector SLI system, developed an algorithm that overcomes over-fitting deficiencies while employing a novel template fitting and tracking procedure that fills in missing data and yields point correspondences across the entire sequence.

Davis' and Zhang's works demonstrate the robustness of the phase matching framework in motion detection and improving the SNR in SLI systems. However, in none of these works has the distribution of errors been analyzed nor utilized, nor was the texture or modulation data further exploited where errors exist in SLI systems [74, 84, 85, 86, 87] while decades of study on stereo-vision proves that 3-D geometry can also be revealed from texture or modulation data [33, 117] alone. More recent research shows that good quality 3-D reconstructions can be recovered especially in high texture regions [112, 30]. Thus, how to utilize these three types of information (phase, texture, and modulation) to maximize 3-D data acquisition efficiency becomes a key question to the development of advanced SLI technologies.

To take full advantage of the information provided by texture and modulation data for the purpose of improving phase-stereo matching, this chapter presents a multi-modal data fusion and hybrid 3-D reconstruction framework for a two-camera SLI system, where the pattern phase data obtained from the cameras is used to compute stereo-vision correspondences through sub-pixel phase matching with object texture used to eliminate ambiguities. Pattern modulation information will then be used to estimate a phase error model for Kullback-Leibler (KL) divergence [118] analysis as

a means of further refinement. KL divergence analysis is typically associated with measuring the “distance” between probability distributions [118]. While some recent techniques employ it to account for occlusion in stereo-vision models [119], it can also be used as an error metric between noisy values. It is particularly relevant for correspondence matching between camera views in the presence of noise and uncertainty.

Based upon PMP, we employ the theoretical minimum three patterns while achieving a significant improvement in 3-D reconstruction quality, over conventional PMP and stereo-vision techniques, that performs well in both rich and poor texture environments with complicated light paths and even in high contrast conditions. Although we require the cameras to be synchronized to each other in software, the new method is immune to the projector flicker and gamma distortion without rigorous synchronization to the projector. And in the proposed framework, a rigorous synchronization between cameras and projectors is not needed and the projector gamma distortion is removed without extra calibration and compensation.

In summary, the unique contributions of this chapter include developing a hybrid 3-D reconstruction framework for multi-camera SLI systems that utilizes the notion of stereo-vision to reduce phase matching ambiguities based on the correlation among object texture images. But first, this chapter derives a statistical phase-error model that is successfully applied, both theoretically and experimentally. As a result, the 3-D reconstructions are improved through a developed correspondence refinement mechanism to reduce mis-registration among images, based on the derived phase error model and KL divergence analysis. Modulation and phase information is fused

for this purpose.

4.1 Stereo matching in phase

Various problems arise in SLI systems when a small number of patterns are used. Especially in regions with rich texture, multi-lightpath or high contrast, phase data tends to be inaccurate and noisy. The poor quality of phase data may be caused by many factors: (i) rich texture, high contrast or multi-lightpath of the scanned object, (ii) noise from cameras, projectors and ambient light, (iii) gamma distortion, (iv) flickering problem, etc. Stereo vision, on the other hand, fails in regions with low texture or areas in short of feature points.

An illustrative example is shown in Fig. 4.2. Fig. 4.2 (a) and (b) show the stereo images of a rich-textured object and a poor-textured object. Fig. 4.2 (c) shows the phase data along epipolar lines in a rich-textured region. When only three PMP patterns are used, the phase data in such kind of regions is very noisy. Fig. 4.2 (d) shows the intensity data along the epipolar lines in a poor-textured region. Due to the lack of texture information and feature points, stereo vision techniques fail to work in such kind of regions.

Performing PMP scan with two cameras and one projector yields two 3-D surfaces. We expect by fusing these two 3-D surfaces to obtain a higher resolution, less distorted, and more wrapped version of 3-D fingerprints. The key to merge reconstructed images from different cameras lies in finding the point correspondence among cameras. Another problem associated with multi-camera SLI system is the mis-alignment of 3-D data obtained from different views. Fig. 4.3 is used to illustrate

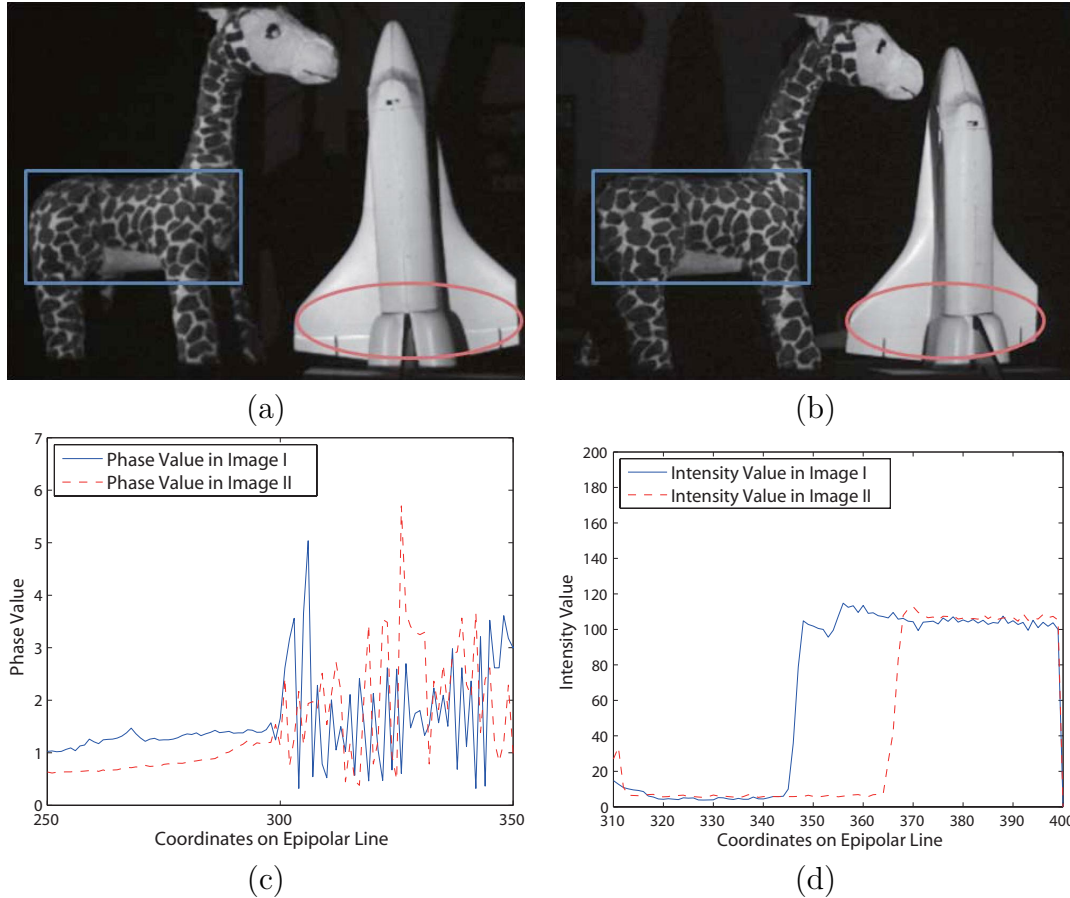


Figure 4.2: A cartoon deer with rich texture and a white space plane with poor texture. (a) The image I of objects captured by the first camera. (b) The image II of objects captured by the second camera. (c) Cross sections of phase data in the rectangular region marked in images I and II. With only three PMP patterns, phase data is of poor quality and noisy. (d) Cross sections of intensity data in the circular region marked in images I and II. Because of the poor texture information, texture-based matching does not work well.

this problem. Figs. 4.3 (a) and (b) show the 3-D surfaces reconstructed from two cameras using spot array patterns. However, when the two 3-D surfaces are positioned in the same coordinate system, it can be seen that there is a displacement between these two surfaces, as shown in Fig. 4.3 (c) and (d). Such a misalignment illustrated may be caused by many factors. Some of them are analyzed in [115, 116]. Generally speaking, in a multi-views system, data fusion should be performed at the

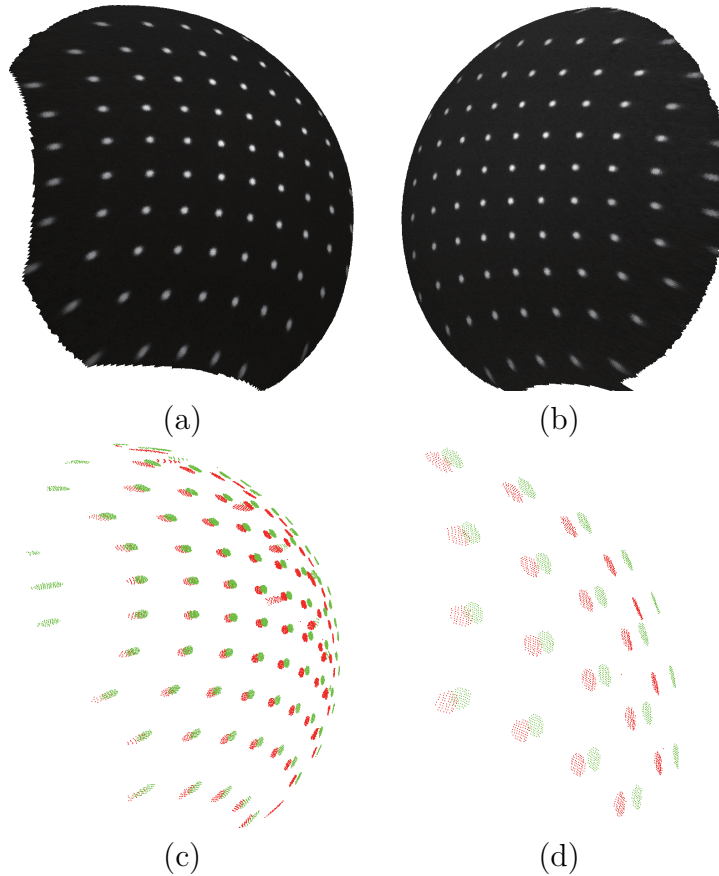


Figure 4.3: SLI 3D surfaces reconstructed using spot array patterns. Sixteen high frequency PMP patterns are used to ensure the phase data quality. (a) The 3D surface reconstructed through camera I and the projector. (b) The 3D surface reconstructed through camera II and the projector. (c) The resultant surface by merging two surfaces shown in (a) and (b). (d) An enlarged and rotated part of the surface shown in (c) to illustrate the misalignment.

pixel level instead of at the surface level.

The points mismatching, illustrated in Fig. 4.3, can be caused by many factors. Some of them are analyzed in [115, 116]. A few remedying approaches have been proposed in [116, 120, 42]. In a typical optical system, both cameras and projector will introduce distortions due to the curving lens surface and the limited aperture size. Using the PMP technique, the temporal multiplexed illumination patterns can offset the projector distortion, so the camera distortion remains the main cause of

the mismatching. To overcome the multiple 3D surfaces mismatching caused by the distortion of camera lenses, we propose a method that can achieve the precise point correspondence between two 3D surfaces based on the notion of stereo vision.

Traditional stereo matching is to find the correspondence of points from two or more planes. Stereo matching has been studied for decades in computer vision, which can be broadly classified into feature based [25, 121, 26] and point based approaches [122, 123]. Generally speaking, when come to two or more views geometry, the fundamental matrices and epipolar lines [27, 124] can be derived using an external calibration target. The corresponding points from different views must lie on the same epipolar line. That is, stereo matching greatly reduces the search for correspondences. The point correspondences deduce the stereo disparities. Through stereo triangulation, disparities can be converted into coordinates of points in 3D space.

In this section, we present a new approach for 3D surface reconstruction with multiple cameras SLI systems by incorporating PMP technique into stereo vision correspondence. We employ epipolar geometry and sub-pixel technique to find the correspondence between phase images from different cameras and reconstruct the 3D surface. The system performance is further optimized by tuning three parameters: the alignment angle of patterns, the number of patterns, and the geometry among cameras and the projector. The optimized 3D surface reconstruction by our method is also compared with the performance of the conventional correlation based method.

Rotated structured light

Traditional PMP projects either vertical or horizontal sine wave patterns onto the surfaces under scanning so that the vertical or horizontal correspondence information between the camera and the projector can be directly derived from the computed phase data.

Here, we use patterns rotated in different angles to choose the optimal pattern alignment. Denote r as the alignment angle, L^p as the image length, and H^p as the image height. If we use single frequency patterns, then

$$f = \frac{1}{L^{pl}}, \quad (4.1)$$

where L^{pl} is the length of the pattern. If the pattern alignment is horizontal, i.e. $r = 0$, L^{pl} is the height of the pattern. If the pattern alignment is vertical, i.e. $r = \frac{\pi}{2}$, L^{pl} is the length of the pattern. In between,

$$L^{pl} = \frac{L^p}{\sin(r)}, \quad (4.2)$$

for $R < r < \pi/2$, and

$$L^{pl} = \frac{H^p}{\cos(r)}, \quad (4.3)$$

for $0 < r \leq R$, where R is the $\arctan(\cdot)$ of the image length-to-height ratio. Thus, a rotated pattern with an angle r can be expressed as

$$I_n(x^p, y^p) = A^p + B^p \cos[2\pi p(\sin(r)y^p + \cos(r)x^p) - 2\pi \frac{n}{N}]. \quad (4.4)$$

Several created patterns are shown in Fig. 4.4.

From the camera point of view, the captured image is distorted by the scanned surface's topology. The reflection and data processing equations are still unchanged

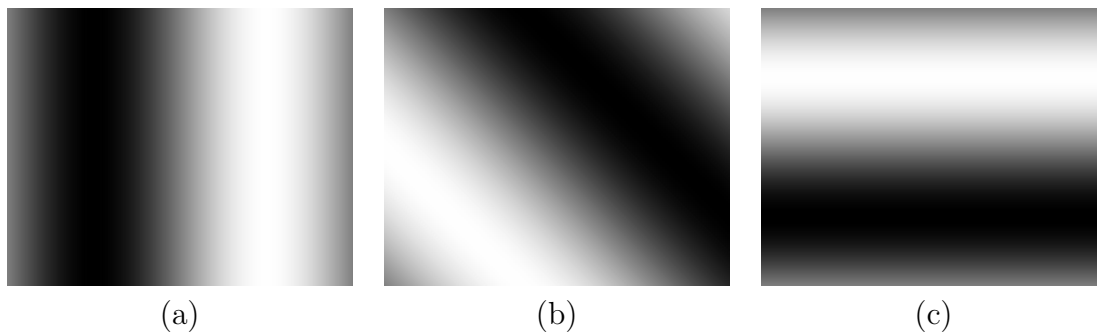


Figure 4.4: Three image patterns with different rotation angles. (a) shows the vertical pattern where $r = \pi/2$. (b), shows a pattern where $r = \pi/4$. (c) shows the horizontal pattern where $r = 0$.

as we illustrated in chapter I. After converting the phase from $[0, 2\pi)$ to $[0, 255]$, we can get the 8-bit gray level phase image, as shown in Fig. 4.5.

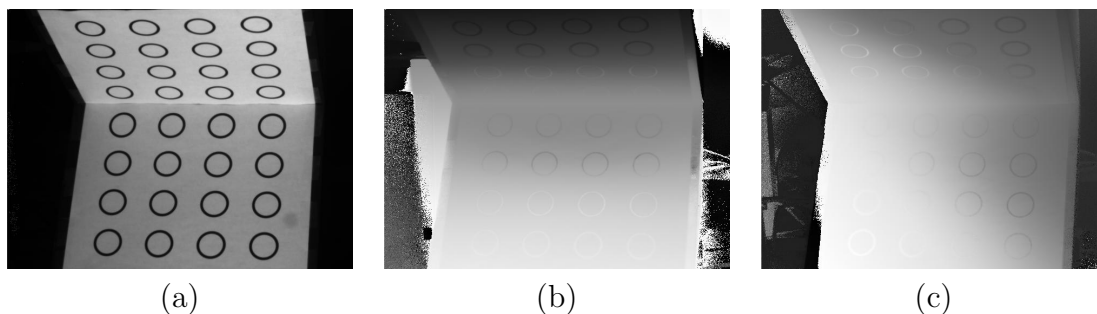


Figure 4.5: (a) The target image. (b) The phase image of the target when $r = 0$. (c) The phase image of the target when $r = \pi/4$.

The target shown in Fig. 4.5 (a) is used to evaluate the 3D reconstruction performance of the proposed rotated pattern PMP and sub-pixel stereo phase matching. Three pattern numbers in use, 25, 40 and 55, are for both traditional PMP and the rotated pattern PMP. Ten different angles are tested for each rotated pattern PMP. The results are given in Fig. 4.8. The reconstruction errors include the mean square root errors in x, y and z directions of the world coordinates in the unit of *mm*.

Figure 4.7 shows the reconstructed surface of the target using traditional PMP technique with 40 patterns. Figure 4.8 shows the 3-D reconstruction errors using

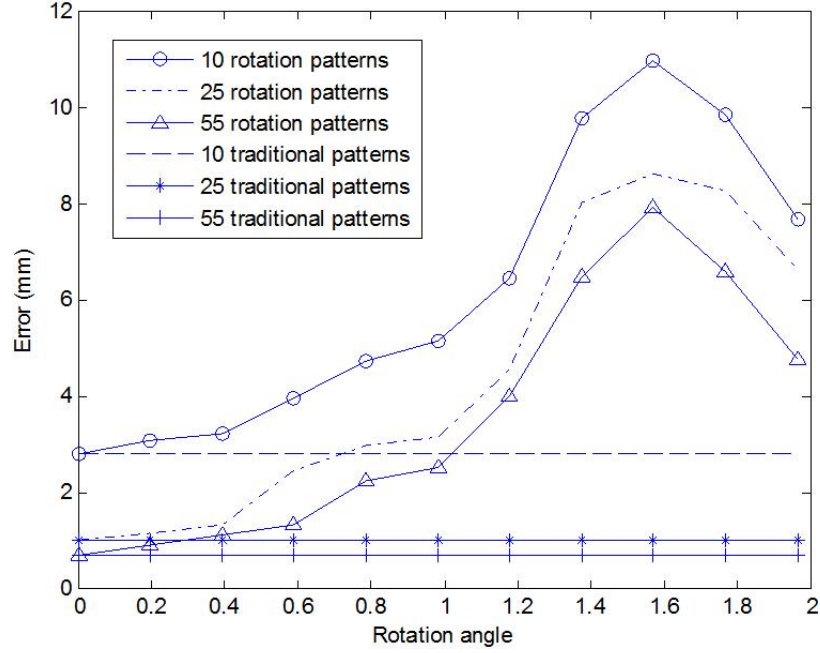


Figure 4.6: Plot of PMP reconstruction errors vs. rotation angles. The three straight lines are the errors of traditional PMP, where the pattern rotation angle is a constant.

rotated patterns in contrast to using traditional patterns, where the rotation angle is 0. With an increase of pattern number, the errors of 3-D surface reconstruction drop as expected. For the rotation pattern PMP, the errors reach their maxima as the rotation angles go near $\pi/2$. This is because in our experimental system the camera-projector line is nearly vertical. Usually the best reconstruction results could be obtained when the pattern angle is perpendicular to the camera-projector line. In theory, when the pattern alignment is parallel to the camera-projector line, little information on can be extracted, resulting in infinite reconstruction errors. In practice, as the camera and the projector are not perfectly vertically aligned, some information on still can be derived even using vertical patterns. It also can be seen that when the number of patterns is larger than 25, there is no much room to improve the reconstruction performance.

The transformation equations from real world to camera is given by

$$x^c = \frac{m_{11}^{wc}x^w + m_{12}^{wc}y^w + m_{13}^{wc}z^w + m_{14}^{wc}}{m_{31}^{wc}x^w + m_{32}^{wc}y^w + m_{33}^{wc}z^w + m_{34}^{wc}}, \quad (4.5)$$

and

$$y^c = \frac{m_{21}^{wc}x^w + m_{22}^{wc}y^w + m_{23}^{wc}z^w + m_{24}^{wc}}{m_{31}^{wc}x^w + m_{32}^{wc}y^w + m_{33}^{wc}z^w + m_{34}^{wc}}. \quad (4.6)$$

where m_{ij}^{wc} is the parameter of world-to-camera matrix M^{wc} , which is calculated from camera calibration,

$$A = \begin{bmatrix} m_{11}^{wc} & m_{12}^{wc} & m_{13}^{wc} & m_{14}^{wc} \\ m_{21}^{wc} & m_{22}^{wc} & m_{23}^{wc} & m_{24}^{wc} \\ m_{31}^{wc} & m_{32}^{wc} & m_{33}^{wc} & m_{34}^{wc} \end{bmatrix}, \quad (4.7)$$

Denote M^{wc1} as the world-to-camera matrix of the camera I, and M^{wc2} for the camera II. For a 3-D point, the relation between its world coordinates (x^w, y^w, z^w) and its projection coordinates (x^{c1}, y^{c1}) and (x^{c2}, y^{c2}) can be described as $Q[x^w, y^w, z^w]^T = R$,

where

$$Q = \begin{bmatrix} m_{31}^{wc1}x^{c1} - m_{11}^{wc1}, & m_{32}^{wc1}x^{c1} - m_{12}^{wc1}, & m_{33}^{wc1}x^{c1} - m_{13}^{wc1} \\ m_{31}^{wc1}y^{c1} - m_{21}^{wc1}, & m_{32}^{wc1}y^{c1} - m_{22}^{wc1}, & m_{33}^{wc1}y^{c1} - m_{23}^{wc1} \\ m_{31}^{wc2}x^{c2} - m_{11}^{wc2}, & m_{32}^{wc2}x^{c2} - m_{12}^{wc2}, & m_{33}^{wc2}x^{c2} - m_{13}^{wc2} \\ m_{31}^{wc2}y^{c2} - m_{21}^{wc2}, & m_{32}^{wc2}y^{c2} - m_{22}^{wc2}, & m_{33}^{wc2}y^{c2} - m_{23}^{wc2} \end{bmatrix}, \quad (4.8)$$

and

$$R = \begin{bmatrix} m_{14}^{wc1} - m_{34}^{wc1} \\ m_{24}^{wc1} - m_{34}^{wc1} \\ m_{14}^{wc2} - m_{34}^{wc2} \\ m_{24}^{wc2} - m_{34}^{wc2} \end{bmatrix}. \quad (4.9)$$

The pseudo-inverse solution can be derived as

$$[x^w, y^w, z^w]^T = (Q^T Q)^{-1} Q^T R \quad (4.10)$$

The target shown in Fig. 4.5 (a) was used again to test the 3-D reconstruction errors of the proposed sub-pixel stereo phase matching technique. We used ten different rotation angles. For each rotation angle, three different pattern numbers, 55, 25 and

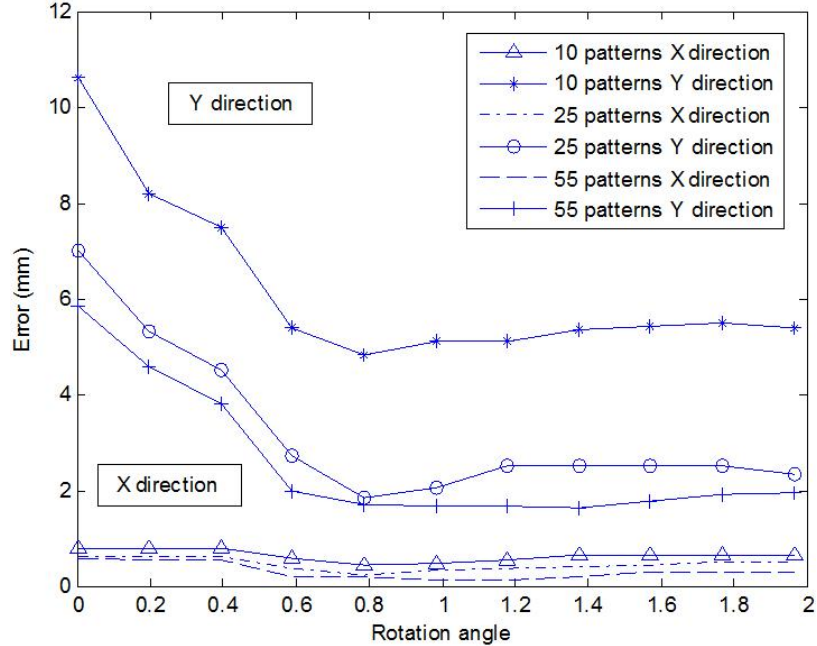


Figure 4.8: Reconstruction errors using sub-pixel stereo phase matching vs. the rotation angles. The upper three curves are errors in y direction. The lower three curves are the errors in x direction.

10, were chosen. The errors of matching point pair estimation are measured in mean square root error of x (vertical) and y (horizontal) directions of image coordinates.

The performance of sub-pixel stereo phase matching depends on both the quality of phase information and the angle between pattern alignment and epipolar lines. If the pattern alignment is parallel to one epipolar line, all the points on this line will have the same phase value; in that case, the stereo matching by phase becomes impossible. In our prototype system, the two cameras are positioned horizontally; therefore the vertical patterns should yield the best reconstruction results. However, Fig. 4.8 shows that the best results happen at the rotation angle of near 0.8, instead of $\pi/2$. That is because the reconstruction performance also relies on the quality of phase information. As Fig. 4.8 shows, the quality of phase information, equivalent

to depth errors, degrades most when the pattern alignment goes near $\pi/2$. Besides, there is an angle between the alignments of two cameras, which can not generate a set of parallel horizontal epipolar lines as expected.

It can also be seen from Fig. 4.8 that the error in x direction (vertical) is much less than those in y direction (horizontal). Since the epipolar lines are near horizontal, such a stereo vision geometry helps reduce the errors in x direction. In addition, although the phase information degrades most when the pattern rotation angle goes near $\pi/2$ as shown in Fig. 4.5 (a), the sub-pixel stereo matching reconstruction, even in y direction, does not degrade as much as conventional PMP reconstruction does. It seems that the proposed sub-pixel stereo matching is quite robust to the quality of the phase information extracted using the PMP technique.

Optimization of camera-projector geometry

Sub-pixel stereo matching in phase with vertical patterns and three pattern numbers, 15, 30, and 60, are applied in optimizing the camera-projector-camera geometry. The matching error is defined as the mean square root error between the actual matching points and the estimated matching points of the target images. To obtain better phase information, multi-frequency patterns were used [72]. Three frequencies were chosen as 1, 8, and 32.

Here, we denote the horizontal distance between two cameras as L and the vertical distance between the center of cameras and the projector as H . Figure 4.9 illustrates the point matching errors vs. $\log(H/L)$. The mean square root errors in x and y directions using multi-frequency patterns are reduced, compared to the case of using

single frequency patterns, as the quality of phase information is much improved by using multiple frequencies patterns. As shown in Fig. 4.9, the matching errors drop as the log ratio of H/L decreases. It suggests that the best position of the projector is the middle point of two cameras.

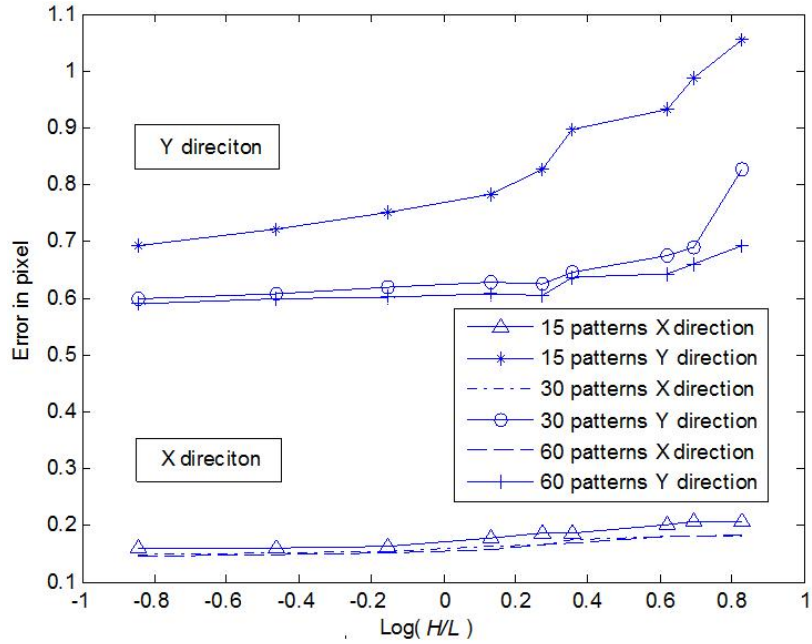


Figure 4.9: Matching errors in pixel vs. log ratio of H/L . The upper three curves are the errors in y direction. The lower three curves are the errors in x direction.

Figure 4.10 shows the reconstruction results of the target by the proposed sub-pixel phase based stereo matching and the conventional correlation based stereo matching. For the proposed method, the errors are 0.1485 pixels in x direction and 0.5833 in y direction. For the conventional method, the errors are 1.5293 pixels in x direction and 15.4825 pixels in y direction. Compared to the traditional stereo matching, the results of sub-pixel stereo phase matching are more accurate and more robust to noise. While traditional correlation based stereo matching may achieve higher quality reconstruction for objects with rich texture information, the proposed

phase based stereo matching approach can effectively reconstruct 3-D surfaces under all the texture and illumination conditions.

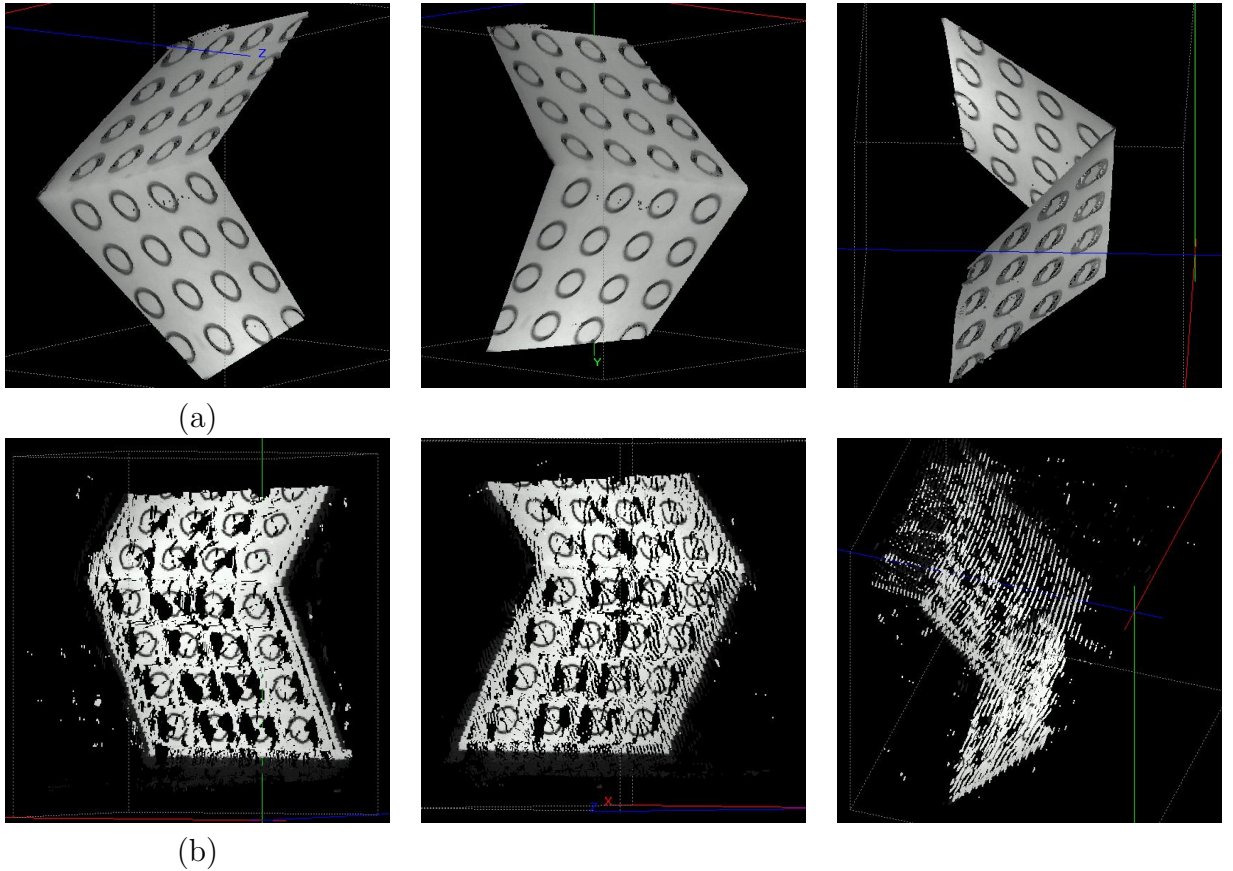


Figure 4.10: (a) Surfaces reconstructed by sub-pixel stereo matching, with 3 frequencies of 1, 8, and 32, 10 patterns for each frequency. (b) Surfaces reconstructed by traditional correlation based stereo matching.

In this section, we propose a phase information based sub-pixel stereo matching for multi-perspective structured light illumination 3D surface reconstruction. We also investigate the impact of pattern rotation angles, pattern numbers, and camera-projector-camera geometry on the reconstruction performance. The advantages of using multiple cameras in SLI scanning over the single camera SLI system include its higher robustness to the quality of phase information, as well as the measurement noise, and its capability to acquire surrounding information of objects under exami-

nation. Compared to the traditional correlation based stereo matching, the proposed method can produce high quality 3D surface reconstruction despite the texture and illumination conditions.

4.2 Phase error modeling

Of many proposed SLI methods, the technique of phase measuring profilometry (PMP), or sinusoidal phase-shifting pattern, is one of the most widely used and precise strategies [84, 75, 76, 77, 70]. The canonical PMP technique employs a set, $\{I_n^p : n = 1, 2, \dots, N\}$, of single frequency sinusoidal wave patterns, such that a point in the projector plane, (x^p, y^p) , is assigned according to:

$$I_n^p(x^p, y^p) = A^p + B^p \cos(2\pi f y^p - \frac{2\pi n}{N}) \quad (4.11)$$

where A^p and B^p are constants of the projector and f is the frequency of the sinusoidal wave [70]. Note the dependence of the phase term, $2\pi f y^p$, on the vertical coordinate inside the projector, y^p , as this is the parameter that will be used when triangulating with the camera, which is assumed to be positioned vertically above the projector. As discussed in [74], larger values of f typically result in smoother 3-D surfaces in the presence of sensor noise but also introduce ambiguities in the phase reconstruction necessary for reconstructing the 3-D surface.

After the set of illumination patterns have been projected upon a target object, an off-axis camera captures the scene such that the sinusoidal wave patterns are observed after being distorted by the surface topology under inspection. The resulting set of

captured images, $\{I_n^c : n = 1, 2, \dots, N\}$, can, therefore, be expressed as

$$I_n^c = A + B \cos\left(\Phi - \frac{2\pi n}{N}\right), \quad (4.12)$$

where the two-dimensional camera coordinate (x^c, y^c) has been left out for brevity of notation for the image terms I_n^c , A , B , and Φ . In Eq. (4.12), the parameter A is the albedo image of the captured scene, which is derived according to

$$A = \frac{1}{N} \sum_{n=1}^N I_n^c, \quad (4.13)$$

while parameter B represents the amplitude of the observed I_n^p reflected off the target [70], which is derived according to

$$B = \frac{2}{N} \sqrt{\left(\sum_{n=1}^N I_n^c \sin\left(\frac{2\pi n}{N}\right)\right)^2 + \left(\sum_{n=1}^N I_n^c \cos\left(\frac{2\pi n}{N}\right)\right)^2}. \quad (4.14)$$

Finally, the parameter Φ represents the phase value at pixel location (x^c, y^c) of the captured sinusoidal wave pattern, computed from the captured images as

$$\Phi = \text{atan} \left[\frac{\sum_{n=1}^N I_n^c \sin\left(\frac{2\pi n}{N}\right)}{\sum_{n=1}^N I_n^c \cos\left(\frac{2\pi n}{N}\right)} \right], \quad (4.15)$$

which in a noiseless system, can be calculated from $N \geq 3$ patterns [70]. And once the value of Φ is obtained, the 3-D world coordinates of a point can be calculated from (x^c, y^c, Φ) as described by Li et al [74].

Noise Analysis

In practice, error in I_n^c is introduced from camera noise [74], projector gamma, flicker and noise [45, 84], ambient light noise [45] as well as other sources [74, 85, 86, 87]. We classify the error sources to projector gamma distortion [84], outside camera noise

(including projector noise, flicker, ambient light noise and other noise sources outside the camera), and camera noise. Thus, the captured images, I_n^c , in Eq. (4.12) can be rewritten as

$$\tilde{I}_n^c = A + B\cos(\Phi - 2\pi n/N) + \rho_n + w_n^o + w_n^c, \quad (4.16)$$

where $\tilde{\cdot}$ denotes the observed variable polluted by noise while the projected pattern is of unit frequency with $\Phi \in [0, 2\pi)$. The parameter ρ_n is the error caused by projector gamma distortion while the parameters w_n^c and w_n^o represent noise sources inside and outside of the camera, respectively.

For conventional single view SLI systems, by synchronizing the projector and camera that eliminates pattern flicker, as well as performing a thorough calibration that eliminates gamma distortion [84], then Eq. (4.16) reduces to

$$\tilde{I}_n^c = A + B\cos(\Phi - 2\pi n/N) + w^s, \quad (4.17)$$

where we have the combined noise term $w^s = w_n^o + w_n^c$. This simplified form then represents that proposed by Li et al [74], Kamgar-Parsi and Kamgar-Parsi [85], and Savarese et al [86] who further model w^s as additive, white Gaussian noise, i.e., $w^s = w_n^s \sim N(0, \sigma^2)$. The phase value with noise, $\tilde{\Phi}$, is then calculated as

$$\begin{aligned} \tilde{\Phi} &= \operatorname{atan} \left[\frac{\sum_{n=1}^N \tilde{I}_n^c \sin(\frac{2\pi n}{N})}{\sum_{n=1}^N \tilde{I}_n^c \cos(\frac{2\pi n}{N})} \right] \\ &= \operatorname{atan} \left[\frac{\frac{N}{2} B \sin(\Phi) + \sum_{n=1}^N w_n^s \sin(\frac{2\pi n}{N})}{\frac{N}{2} B \cos(\Phi) + \sum_{n=1}^N w_n^s \cos(\frac{2\pi n}{N})} \right]. \end{aligned} \quad (4.18)$$

For real-valued numbers α and β , the following equation always exists,

$$\operatorname{atan}(\alpha) - \operatorname{atan}(\beta) = \operatorname{atan}\left(\frac{\alpha - \beta}{1 + \alpha\beta}\right), \quad \alpha\beta > -1 \quad (4.19)$$

by applying Eq. (4.19), Φ_e , in Eq. (4.23) is obtained:

$$\begin{aligned}
\Phi_e &= \Phi - \tilde{\Phi} \\
&= \text{atan} \left[\frac{\sin(\Phi)}{\cos(\Phi)} \right] - \text{atan} \left[\frac{\frac{N}{2} B \sin(\Phi) + \sum_{n=1}^N w_n^c \sin(\frac{2\pi n}{N})}{\frac{N}{2} B \cos(\Phi) + \sum_{n=1}^N w_n^c \cos(\frac{2\pi n}{N})} \right] \\
&= \text{atan} \left(\frac{\frac{2}{N} \sum_{n=1}^N w_n^c \sin(\Phi - \frac{2\pi n}{N})}{B + \frac{2}{N} \sum_{n=1}^N w_n^c \cos(\Phi - \frac{2\pi n}{N})} \right).
\end{aligned} \tag{4.20}$$

Further, we have

$$\begin{aligned}
\tilde{B}^2 &= (B \sin(\Phi) + \frac{2}{N} \sum_{n=1}^N w_n^c \sin(\frac{2\pi n}{N}))^2 + \\
&\quad (B \cos(\Phi) + \frac{2}{N} \sum_{n=1}^N w_n^c \cos(\frac{2\pi n}{N}))^2 \\
&= B^2 + \frac{4}{N} B \sum_{n=1}^N w_n^c \cos(\Phi - \frac{2\pi n}{N}) + \\
&\quad \frac{4}{N^2} \sum_{n=1}^N \sum_{m=1}^N w_n^c w_m^c \cos(\frac{2\pi}{N}(n-m)) \\
&= (B + \frac{2}{N} \sum_{n=1}^N w_n^c \cos(\Phi - \frac{2\pi n}{N}))^2 + \\
&\quad \frac{4}{N^2} [(\sum_{n=1}^N w_n^c \sin(\frac{2\pi n}{N}))^2 + (\sum_{n=1}^N w_n^c \cos(\frac{2\pi n}{N}))^2] - \\
&\quad \frac{4}{N^2} (\sum_{n=1}^N w_n^c \cos(\Phi - \frac{2\pi n}{N}))^2 \\
&= (B + \frac{2}{N} \sum_{n=1}^N w_n^c \cos(\Phi - \frac{2\pi n}{N}))^2 + \\
&\quad (\frac{2}{N} \sum_{n=1}^N w_n^c \sin(\Phi - \frac{2\pi n}{N}))^2.
\end{aligned} \tag{4.21}$$

Thus,

$$\begin{aligned}
\tilde{B} &= \left\{ (B + \frac{2}{N} \sum_{n=1}^N w_n^c \cos(\Phi - \frac{2\pi n}{N}))^2 + \right. \\
&\quad \left. (\frac{2}{N} \sum_{n=1}^N w_n^c \sin(\Phi - \frac{2\pi n}{N}))^2 \right\}^{\frac{1}{2}}.
\end{aligned} \tag{4.22}$$

Thereby introducing the notion of phase error, Φ_e , given by

$$\Phi_e = \Phi - \tilde{\Phi} = \text{atan} \left(\frac{Y}{X} \right). \tag{4.23}$$

Furthermore, the noisy modulation image is given by

$$\tilde{B} = \sqrt{X^2 + Y^2}, \tag{4.24}$$

where

$$X = B + \frac{2}{N} \sum_{n=1}^N w_n^s \cos(\Phi - \frac{2\pi n}{N}), \tag{4.25}$$

and

$$Y = \frac{2}{N} \sum_{n=1}^N w_n^s \sin(\Phi - \frac{2\pi n}{N}). \tag{4.26}$$

Distribution of Phase Error

Under the assumption of Gaussian noise, auxiliary variables X and Y are also Gaussian, i.e. $X \sim N(B, \frac{2}{N}\sigma^2)$ and $Y \sim N(0, \frac{2}{N}\sigma^2)$. As their correlation can be estimated by

$$E[XY] = \sigma^2 \sum_{n=1}^N \sin(2\Phi - \frac{4\pi n}{N}) = 0, \quad (4.27)$$

X and Y are independent. The joint Probability Density Function (PDF) of Φ_e and \tilde{B} is then given by

$$f_{\tilde{B}, \Phi_e}(\tilde{b}, \phi_e) = \frac{N\tilde{b}}{4\pi\sigma^2} \exp\left[-N \frac{(\tilde{b})^2 + B^2 - 2\tilde{b}(B\cos(\phi_e))}{4\sigma^2}\right], \quad (4.28)$$

derived in [105]. Integrating Eq. (4.28) over \tilde{B} yields the marginal PDF for Φ_e given by

$$f_{\Phi_e}(\phi_e) = \frac{1}{2\pi} \exp\left[-\frac{NB^2}{4\sigma^2}\right] \left[1 + \kappa\sqrt{\pi} \exp(\kappa^2)(1 + \text{erf}(\kappa))\right], \quad (4.29)$$

where

$$\kappa = \frac{\sqrt{N}B}{2\sigma} \cos(\phi_e) \quad (4.30)$$

and $\text{erf}(\cdot)$ is error function. In a given SLI system, Eq. (4.29) can be used to predict the achievable accuracy.

Defining the SNR of a SLI system using PMP patterns as $\frac{\sqrt{N}B}{\sqrt{2}\sigma}$, Fig. 4.11 shows the PDFs of phase error, Φ_e , for different SNRs. It can be seen that an increase in pattern number, N , or in modulation value, B , will reduce the phase error Φ_e . As such, the modulation image, B , is typically used in many SLI scanning systems to define a binary mask such that the pixel with sufficiently large B , ($B(x^c, y^c) > \text{threshold}$), is considered to be reliable in determining that pixel's phase, and it is only these pixels

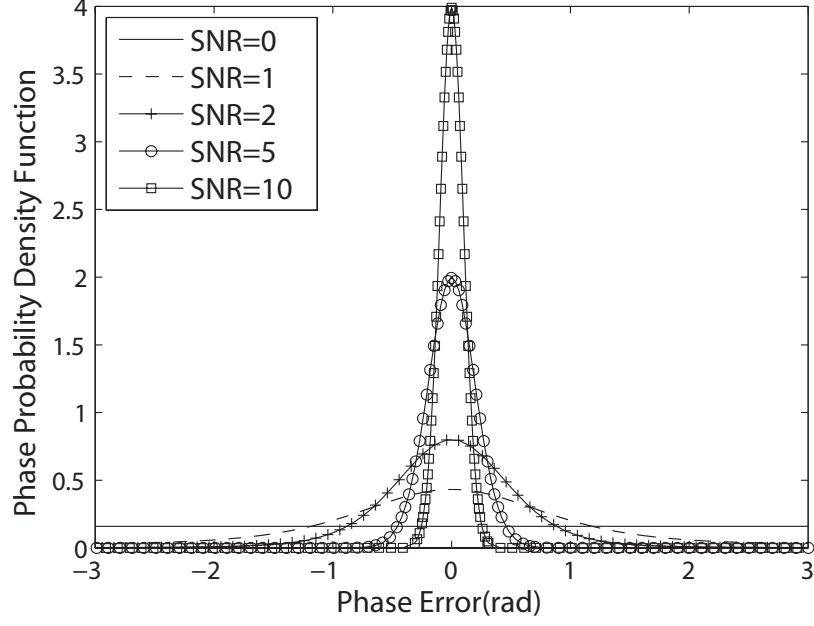


Figure 4.11: Probability density function of phase error.

that are used to reconstruct a 3-D surface. Figure 4.11 also demonstrates that the PDF of Φ_e is a uniform distribution when the SNR is 0. Theoretically, this explains the impulsive nature of noise observed in low B regions of a phase image, because $\tilde{\Phi}$ can be far from the true phase value with the same probability.

For large SNRs where \sqrt{NB} can be considered much greater than $\sqrt{2}\sigma$, the phase error will be small with the error function, of Eq. (4.30), close to 0. The second term, $\kappa\sqrt{\pi}\exp(\kappa^2)(1 + \text{erf}(\kappa))$, will then dominate the constant 1 such that Eq. (4.29) effectively reduces to

$$\begin{aligned}
 f_{\Phi_e}(\phi_e) &\approx \frac{\sqrt{NB}\cos(\phi_e)}{2\sqrt{\pi}\sigma} \exp\left[\frac{-NB^2\sin^2(\phi_e)}{4\sigma^2}\right] \\
 &\approx \frac{\sqrt{NB}}{2\sqrt{\pi}\sigma} \exp\left[\frac{-NB^2\phi_e^2}{4\sigma^2}\right], \tag{4.31}
 \end{aligned}$$

which demonstrates that the phase error distribution can be considered a zero-mean, Gaussian distribution. As shown in Fig. 4.12, the Gaussian approximation is accurate

for SNRs as small as 3.

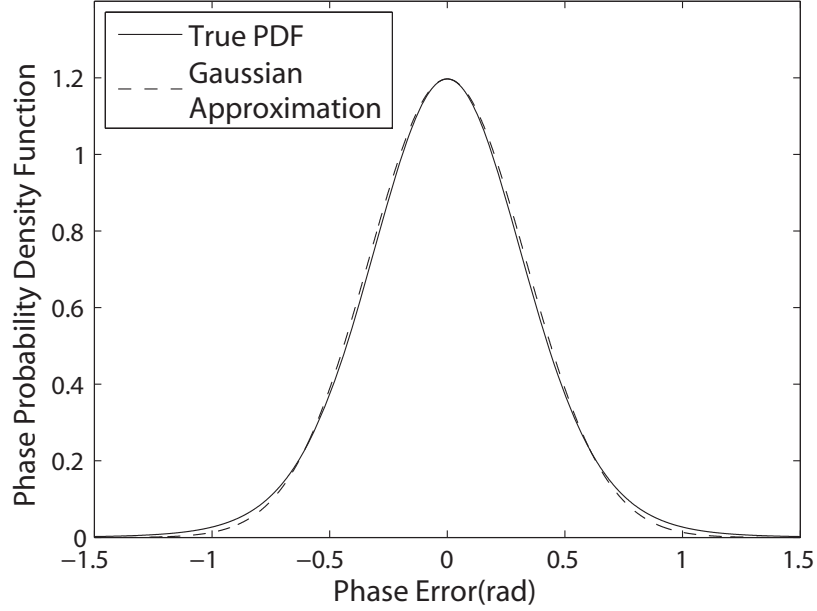


Figure 4.12: Probability density function of phase error Φ_e and its Gaussian approximation with an SNR equal to 3.0.

Variations of Phase Error and Modulation

In order to derive the variance of Φ_e as it relates to the observed \tilde{B} instead of the true B , we denote μ_X and μ_Y as the mean values of X and Y . And by integrating Eq. (4.28) over Φ_e , we obtain the marginal PDF of \tilde{B} as a Rician random variable with distribution

$$f_{\tilde{B}}(\tilde{b}) = \frac{N\tilde{b}}{2\sigma^2} \exp\left[-N\frac{(\tilde{b})^2 + B^2}{4\sigma^2}\right] I_0\left(\frac{N\tilde{b}B}{2\sigma^2}\right) \quad \tilde{b} \geq 0, \quad (4.32)$$

where $I_0(\cdot)$ is the zero-th order modified Bessel function of the first kind. As the SNR becomes bigger than 3, the PDF of modulation rapidly starts to approximate the Gaussian distribution, with the variance σ_B expressed as

$$\sigma_B^2 = \frac{2\sigma^2}{N}. \quad (4.33)$$

The second order moment of \tilde{B} is then

$$E[(\tilde{B})^2] = B^2 + \frac{4\sigma^2}{N}, \quad (4.34)$$

with an unbiased estimator of B given by:

$$B = \sqrt{\tilde{B}^2 - \frac{4\sigma^2}{N}}. \quad (4.35)$$

Replacing B in Eq. (4.23) with this estimator, Φ_e becomes:

$$\Phi_e = \text{atan} \left[\frac{\frac{2}{N} \sum_{n=1}^N w_n^s \sin(\Phi - \frac{2\pi n}{N})}{\sqrt{\tilde{B}^2 - \frac{4\sigma^2}{N} + \frac{2}{N} \sum_{n=1}^N w_n^s \cos(\Phi - \frac{2\pi n}{N})}} \right]. \quad (4.36)$$

Because the noise value is much smaller than the value of modulation data, Eq. (4.36)

can be approximated as

$$\Phi_e \approx \frac{2}{N \sqrt{\tilde{B}^2 - \frac{4\sigma^2}{N}}} \sum_{n=1}^N w_n^s \sin(\Phi - \frac{2\pi n}{N}), \quad (4.37)$$

with the variance of Φ_e approximated by

$$\sigma_{\Phi}^2 \approx \frac{2\sigma^2}{N(\tilde{B}^2 - \frac{4\sigma^2}{N})}. \quad (4.38)$$

When high frequency patterns with frequency f are employed, the phase Φ is unwrapped into $[0, 2\pi)$ before 3-D reconstruction [29, 74]. Thus, the phase error Φ_e is divided by f so that Eq. (4.38) becomes

$$\sigma_{\Phi}^2 \approx \frac{2\sigma^2}{N f^2 (\tilde{B}^2 - \frac{4\sigma^2}{N})}. \quad (4.39)$$

So in summary, Eq. (4.39) and Eq. (4.33) characterize phase error and modulation error with a Gaussian approximation, which we will later use in Sec. 4.3 in order to refine the correspondence matching between cameras of a multi-view SLI system. The PMP noise model, roughly from Eqs. (4.20) to (4.39), is derived by K. Liu in [126].

4.3 Hybrid 3-D reconstruction

For the purpose of reconstructing 3-D surfaces by means of triangulating between cameras of a multi-sensor SLI system using phase, modulation, and texture data, our proposed hybrid construction algorithm is summarized by the steps of: (1) use the phase data to derive an initial set of correspondences across cameras through sub-pixel phase matching; (2) use texture data to reduce the likelihood of ambiguous correspondences; (3) use the modulation data to characterize the phase error model of Eq. (4.39); and (4) perform KL divergence analysis, based on the derived phase error models, to reduce mis-registration among images. In the following description, we assume that we are using a two-camera, single projector system in accordance with the system used in Sec. 4.4.

Sub-pixel Phase Matching

Initially, the proposed process employs sub-pixel phase matching to match pixels across camera images where a point with phase value $\tilde{\Phi}^i$ at integer indexed location $P^i = [m^i, n^i]$, in the first camera, is matched to the interpolated point with real number indexed location $P^{ii} = (x^{ii}, y^{ii})$, in the second camera, such that

$$\tilde{\Phi}^i[m^i, n^i] = \tilde{\Phi}^{ii}(x^{ii}, y^{ii}). \quad (4.40)$$

For the purpose of sub-pixel matching, we assume that the phase value between pixels, along an epipolar line, can be computed by means of bilinear interpolation as

$$\tilde{\Phi}^{ii}(x^{ii}, y^{ii}) = s_1\tilde{\phi}_4 + s_2\tilde{\phi}_3 + s_3\tilde{\phi}_2 + s_4\tilde{\phi}_1, \quad (4.41)$$

where $\tilde{\Phi}^{ii}$ is the phase value of the point on the epipolar line, $\tilde{\phi}_1$ to $\tilde{\phi}_4$ are the phase values of the four nearest-neighboring integer indexed pixels, and s_1 to s_4 are the values of the four nearest-neighboring integer indexed pixels, and s_1 to s_4 are the areas of the sub-pixel regions as illustrated in Fig. 4.13. The localization of the matched point, $P^{ii}(x^{ii}, y^{ii})$, involves finding the phase value, $\tilde{\Phi}^{ii}$ equals to $\tilde{\Phi}^i$, along the epipolar line using Eq. (4.41).

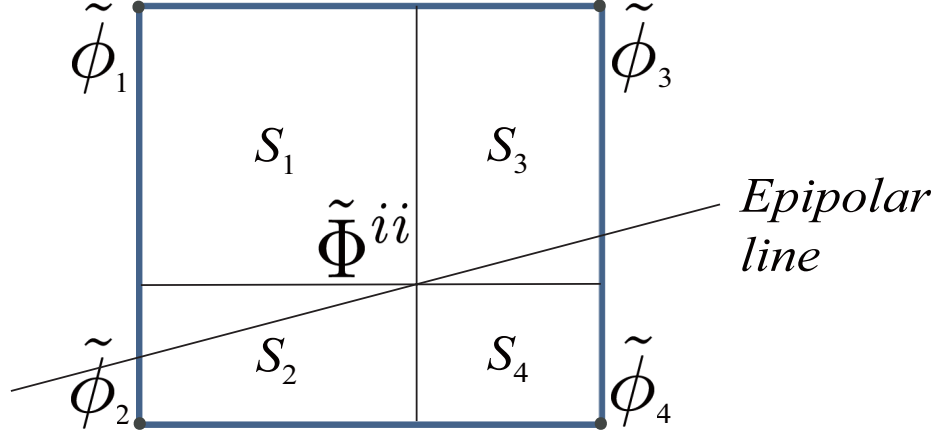


Figure 4.13: An epipolar line goes through the space between four image pixel centers. $s_1+s_2+s_3+s_4 = 1$.

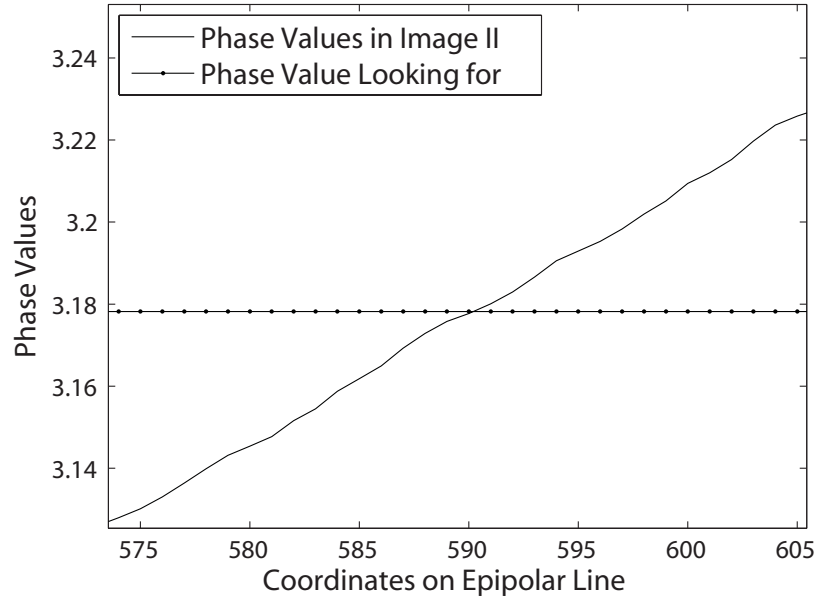
In situations like Fig. 4.14 (b) where sub-pixel phase matching results in multiple pairings such that

$$\{P_k^{ii} | \tilde{\Phi}_k^{ii} = \tilde{\Phi}^i, k \geq 2\}, \quad (4.42)$$

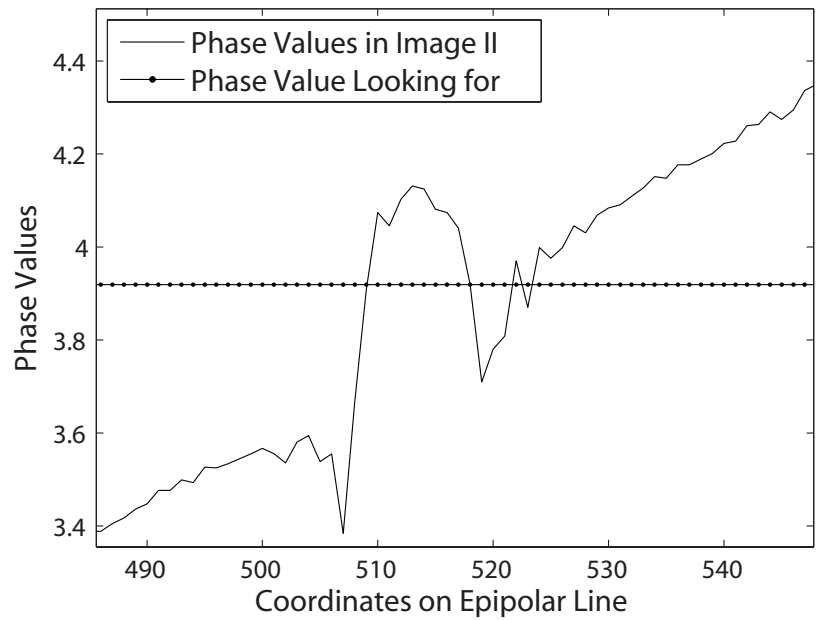
we rely on stereo-vision to identify the optimal pairing, k_{opt} , as the one that results in the highest similarity match in albedo images. For this purpose, we employ the normalized cross-correlation (NCC) defined as

$$NCC(k; i, ii) = \frac{\sum_{g,h} (A^i(x^i+g, y^i+h) - \bar{A}^i)(A^{ii}(x_k^{ii}+g, y_k^{ii}+h) - \bar{A}_k^{ii})}{\sqrt{\sum_{g,h} (A^i(x^i+g, y^i+h) - \bar{A}^i)^2 (A^{ii}(x_k^{ii}+g, y_k^{ii}+h) - \bar{A}_k^{ii})^2}}, \quad (4.43)$$

where \bar{A}^i and \bar{A}_k^{ii} are the average intensity values of two windowed regions about a particular pair given in Eq. (4.42). Therefore, k_{opt} is the pairing that maximizes Eq.



(a)



(b)

Figure 4.14: The phase value to be matched is $\tilde{\Phi}^i$. (a) In the area with low level noise, only one point is matched. (b) In the area with high level noise, five points are matched.

(4.43). The NCC can compensate for the difference in albedo for a given point in the scene as recorded by two different camera sensors [127], at high computational cost.

There, it is only used for remove the phase correspondence ambiguities.

Having eliminated phase-match ambiguities, Eq. (4.42) and Eq. (4.43) represent a means to uniquely construct a 3-D surface based upon triangulation across cameras that incorporates the principals of structured-light and stereo-vision.

Phase Matching Error

In the low SNR environments associated with smaller number of illumination patterns, registration errors exist after sub-pixel phase matching due to having different phase errors between the two cameras. For example, Fig. 4.15 shows matching results based on phase data, where the accurate matching is derived from using 200 high frequency PMP patterns (i.e. high SNR of $\frac{\sqrt{NB}}{\sqrt{2\sigma}}$) while using only 3 high frequency patterns (i.e. low SNR) leads to unreliable matching. It can be seen that registration error is 1.73 pixels.

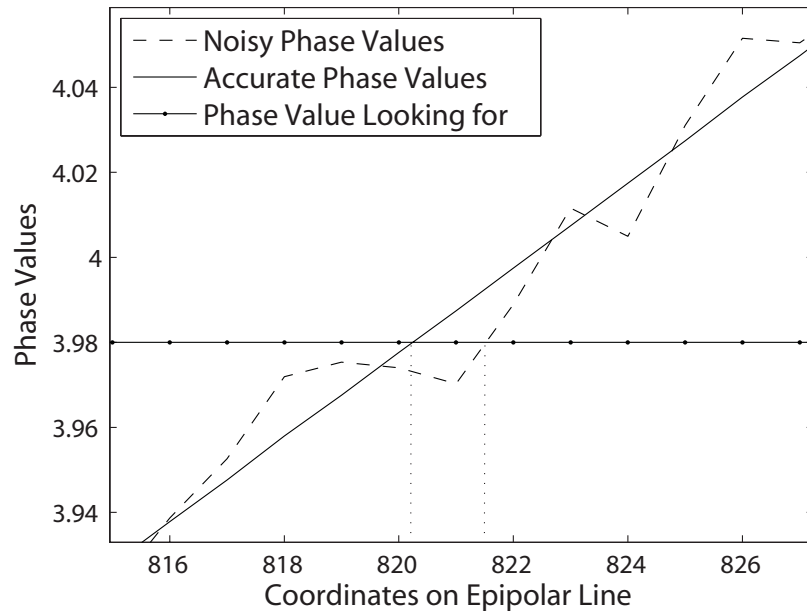


Figure 4.15: The phase value searching for is $\tilde{\Phi}^i$. The matching error is 1.27 pixels.

$$\tilde{I}_n^i = A + B\cos(\Phi - 2\pi n/N) + \rho_n + w_n^o + w_n^i \quad (4.44)$$

and

$$\tilde{I}_n^{ii} = A + B \cos(\Phi - 2\pi n/N) + \rho_n + w_n^o + w_n^{ii}, \quad (4.45)$$

where w_n^i and w_n^{ii} are the respective noise terms within the two cameras. If we assume that the cameras are rigorously synchronized and have near identical gamma characteristics, as would be the case if the cameras were of identical manufacture and model, then we expect ρ_n and w_n^o to be the same in both cameras. We, therefore, denote $\zeta_n = \rho_n + w_n^o$, and thus for the same point on the target surface, the phase difference, Φ_d , between the two cameras can be derived as

$$\begin{aligned} \Phi_d &= \tilde{\Phi}_i - \tilde{\Phi}_{ii} \\ &= \operatorname{atan} \left[\frac{\frac{N}{2} B \sin(\Phi) + \sum_{n=1}^N (w_n^i + \zeta_n) \sin(\frac{2\pi n}{N})}{\frac{N}{2} B \cos(\Phi) + \sum_{n=1}^N (w_n^i + \zeta_n) \cos(\frac{2\pi n}{N})} \right] \\ &\quad - \operatorname{atan} \left[\frac{\frac{N}{2} B \sin(\Phi) + \sum_{n=1}^N (w_n^{ii} + \zeta_n) \sin(\frac{2\pi n}{N})}{\frac{N}{2} B \cos(\Phi) + \sum_{n=1}^N (w_n^{ii} + \zeta_n) \cos(\frac{2\pi n}{N})} \right] \\ &= \operatorname{atan} \left(\frac{V}{U + E} \right), \end{aligned} \quad (4.46)$$

where

$$U = B + \frac{2}{N} \sum_{n=1}^N (w_n^i - w_n^{ii}) \cos(\Phi - \frac{2\pi n}{N}), \quad (4.47)$$

$$V = \frac{2}{N} \sum_{n=1}^N (w_n^i - w_n^{ii}) \sin(\Phi - \frac{2\pi n}{N}), \quad (4.48)$$

and

$$\begin{aligned} E &= \frac{1}{NB} \left(\sum_{n=1}^N (w_n^i + \zeta_n)(w_n^{ii} + \zeta_n) + 2 \sum_{n=1}^N \sum_{m=1, m \neq n}^N \right. \\ &\quad \left. (w_n^i + \zeta_n)(w_m^{ii} + \zeta_m) \cos(\frac{2\pi(n-m)}{N}) \right). \end{aligned} \quad (4.49)$$

Given a sufficiently large SNR (> 3), E is considerably smaller than the two parts in U such that

$$\Phi_d \approx \text{atan} \left(\frac{V}{U} \right), \quad (4.50)$$

where the error caused by outside camera sources, ζ_n , is canceled out in the phase matching result when using multiple cameras.

If we model the noise within each camera as additive, white, Gaussian noise, i.e., $w^i \sim N(0, (\sigma_c^i)^2)$ and $w^{ii} \sim N(0, (\sigma_c^{ii})^2)$ [85, 86, 128], then the PDF of Φ_d would be the same as Φ_e except that σ is replaced with $\sqrt{(\sigma_c^i)^2 + (\sigma_c^{ii})^2}$, noted as σ_c . Thus like Eq. (4.39) with high frequency f , the variance of Φ_d is

$$\sigma_{\Phi_d}^2 \approx \frac{2\sigma_c^2}{Nf^2((\tilde{B}^i)^2 - \frac{4\sigma_c^2}{N})}. \quad (4.51)$$

Similarly, the variance of B_d is

$$\sigma_{B_d}^2 = \frac{2\sigma_c^2}{N}. \quad (4.52)$$

Theoretically, comparison between Eqs. (4.39) and (4.51) proves that, in PMP-based SLI systems, the multi-view system eliminates projector gamma distortion as well as pattern flicker. Furthermore, we can expect a multi-view system, utilizing phase-stereo matching, to outperform traditional single view SLI as long as the overall, single-camera, SLI system's noise is heavier than the summation of the camera noise, e.g. $\sigma^2 > (\sigma_c^i)^2 + (\sigma_c^{ii})^2$, in the two-camera system. This notion is reasonably true as fewer noise sources are involved in multi-camera systems. In our experimental system to be introduced in Sec. (4.4), the variance of the first camera and projector pair, $(\sigma^i)^2$, is 7.2827, whereas $(\sigma_c^i)^2$ and $(\sigma_c^{ii})^2$ are respectively 1.0410 and 1.1038.

For a multi-view PMP-based SLI system, the error after sub-pixel phase matching is characterized by Eq. (4.51) that we will employ for refinement.

KL Divergence Based Refinement

To overcome the 3-D matching difficulties presented in Fig. 4.15, utilizing the noise models can improve the accuracy of registration between stereo pairs [129, 130]. The KL method can measure the divergence of the two camera’s phase distributions. As discussed in Sec. 4.2, the two phase distributions can be approximated as Gaussian with the observed values as means and variances calculated through Eq. (4.39). Thus we can employ the KL method to build correspondence between camera images using both phase and modulation information.

As originally proposed, KL divergence [118] measures the expected number of extra bits required to code samples from a random process \mathbf{P} when using a code based on random process \mathbf{Q} , rather than using a code based on \mathbf{P} itself. The process \mathbf{P} typically represents the “true” distribution of data, or a precisely calculated theoretical distribution, whereas the process \mathbf{Q} typically describes an assumed approximation of \mathbf{P} .

So given two probability distributions \mathbf{P} and \mathbf{Q} and a set of M total samples from each, the KL divergence between \mathbf{P} and \mathbf{Q} is derived as

$$D_{KL}(\mathbf{P}||\mathbf{Q}) = \sum_{i=0}^M \mathbf{P}(i) \log \frac{\mathbf{P}(i)}{\mathbf{Q}(i)}. \quad (4.53)$$

If \mathbf{P} and \mathbf{Q} are both Gaussian, then Eq. (4.58) simplified to

$$D_{KL}(\mathbf{P}||\mathbf{Q}) = \frac{1}{2} \left[2 \log \left(\frac{\sigma_{\mathbf{Q}}}{\sigma_{\mathbf{P}}} \right) + \frac{\sigma_{\mathbf{P}}^2}{\sigma_{\mathbf{Q}}^2} + \frac{(\mu_{\mathbf{P}} - \mu_{\mathbf{Q}})^2}{\sigma_{\mathbf{Q}}^2} - 1 \right]. \quad (4.54)$$

where $\mu_{\mathbf{P}}$ and $\mu_{\mathbf{Q}}$ are the mean values and $\sigma_{\mathbf{P}}$ and $\sigma_{\mathbf{Q}}$ are the variances.

For KL divergence based phase refinement, we intend to use our resulting phase values to derive probability distributions, which we will then use to refine the stereo-phase correspondences according to their KL distance apart. Specifically to measure the distance in phase between the pair of points $P^i = [m^i, n^i]$ and $P^{ii} = (x^{ii}, y^{ii})$, we note that we can derive the Gaussian probability distribution \mathbf{P} at P^i using its phase value $\tilde{\Phi}^i$ and modulation value of \tilde{B}^i , where $\mu_{\mathbf{P}} = \tilde{\Phi}^i$ and $\sigma_{\mathbf{P}}^2$ is calculated by using Eq. (4.39). Likewise, we derive the Gaussian probability distribution \mathbf{Q} at P^{ii} . Refining the phase correspondences also involves searching inside a small window around P^{ii} that minimizes its KL divergence from P^i . Specifically, we search a set of phase values within a range of $3\sigma_{\Phi_d}$ to P^{ii} along the epipolar line in camera II and the point with the minimum KL divergence value is chosen as the matched one.

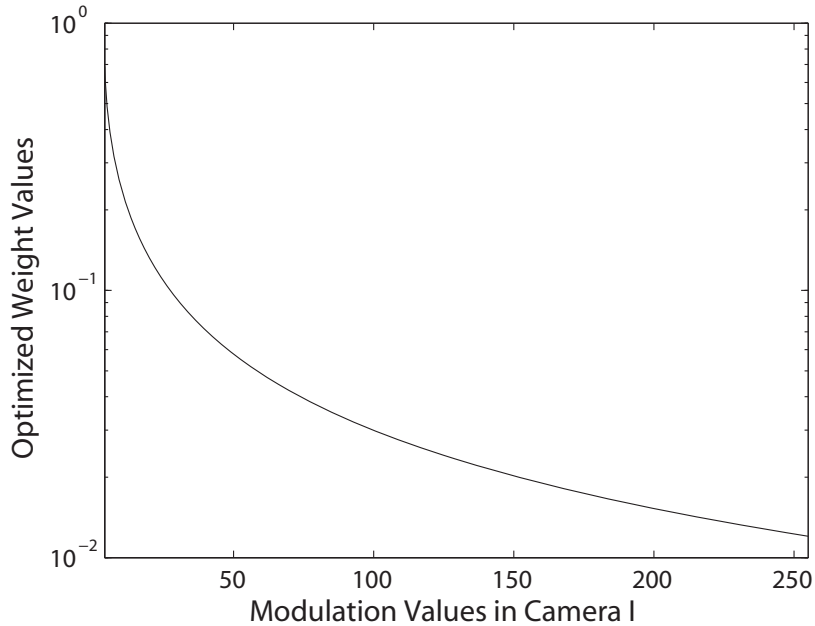


Figure 4.16: Optimized weight function α with respect to modulation value \tilde{B} .

In order to normalize $D_{KL}(\mathbf{P}||\mathbf{Q})$ to account for different units of Φ and \tilde{B} where,

for example, \tilde{B} may range from 0 to 1.0 as a floating point variable or 0 to 255 as an unsigned character or where Φ is in radians versus degrees, we introduce that parameter $\alpha(\cdot)$ into Eq. (4.54) as

$$D_{KL}(\mathbf{P}||\mathbf{Q}) = 2\log\left(\frac{\sigma_{\mathbf{Q}}}{\sigma_{\mathbf{P}}}\right) + \frac{\sigma_{\mathbf{P}}^2}{\sigma_{\mathbf{Q}}^2} + \frac{\alpha^2(\cdot)(\mu_{\mathbf{P}} - \mu_{\mathbf{Q}})^2}{\sigma_{\mathbf{Q}}^2}, \quad (4.55)$$

where $\alpha(\cdot)$ is a weighting function that provides a means of balancing the influences of phase and variance differences in \mathbf{P} and \mathbf{Q} on $D_{KL}(\mathbf{P}||\mathbf{Q})$. In particular under the assumption that

$$D_{KL}(P^i||P_1^{ii}) = D_{KL}(P^i||P_2^{ii}), \quad (4.56)$$

for two points P_1^{ii} and P_2^{ii} where $\tilde{\Phi}_1^{ii} = \tilde{\Phi}^i \pm 3\sigma_{\Phi_d}$ while $\tilde{\Phi}_2^{ii} = \tilde{\Phi}^i$ and $\tilde{B}_1^{ii} = \tilde{B}^i$ while $\tilde{B}_2^{ii} = \tilde{B}^i \pm 3\sigma_{B_d}$ such that both \tilde{B}_2^{ii} and $\tilde{\Phi}_1^{ii}$ are equally unreliable, then $\alpha(\cdot)$ can be pre-calculated as a function of \tilde{B}^i , as demonstrated in Fig. 4.16 for the system used in Sec. 4.4. Note that if we vary $\alpha(\tilde{B}^i)$, either by increasing or decreasing it versus the equality of Fig. 4.16, then we modulate the influence of Φ and \tilde{B} and, thereby, create a stereo-phase matching refinement step that favors one parameter over the other.

4.4 Experimental results and discussion

To demonstrate the proposed stereo-phase matching with KL divergence refinement, we developed a two-camera SLI system as shown in Fig. 4.17. It consists of an Infocus LP70+ DLP projector with 1280×1024 pixel resolution and 1000 ANSI lumens output coupled with two identical, 8-bit, monochrome Prosilica EC1280 cameras with 1280×1024 pixel arrays. Camera I was set up 653 mm apart from the projector

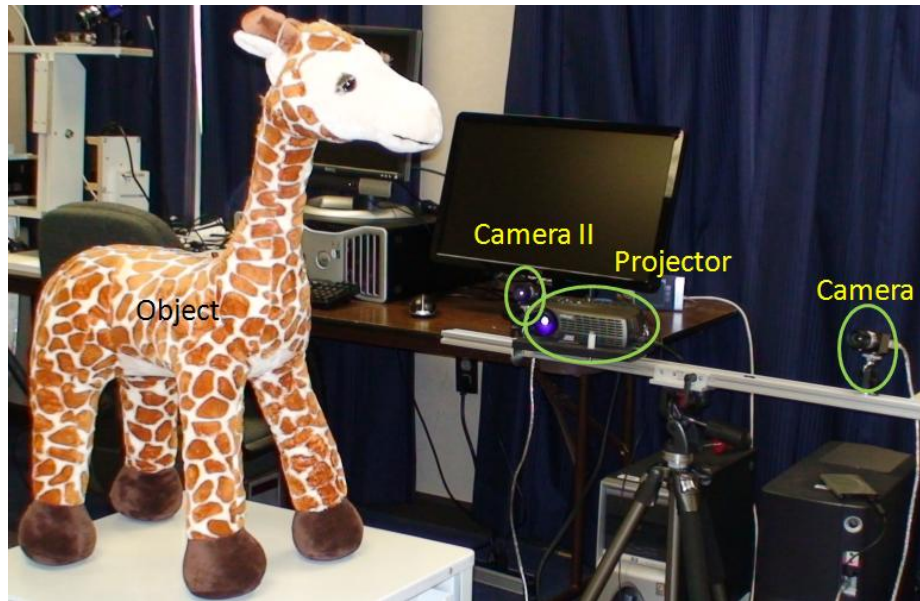


Figure 4.17: The prototype system setup for the hybrid 3-D reconstruction.

as the baseline for traditional SLI reconstruction, while camera II was placed in close proximity to the projector at a distance of only 27 mm. The stereo-phase matching is performed at a near equal triangulation angle to SLI, providing a fair basis for

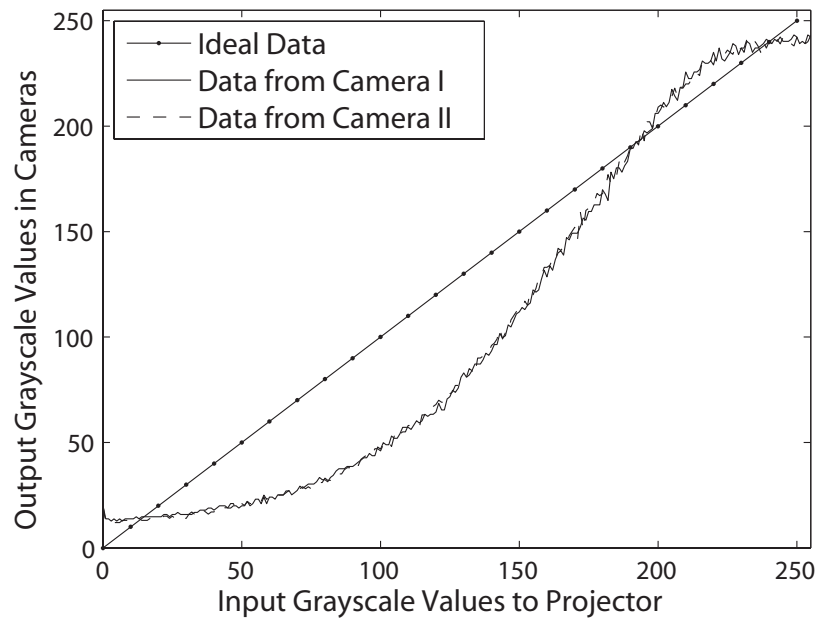


Figure 4.18: The experimental result to illustrate the project gamma distortion and flicker.

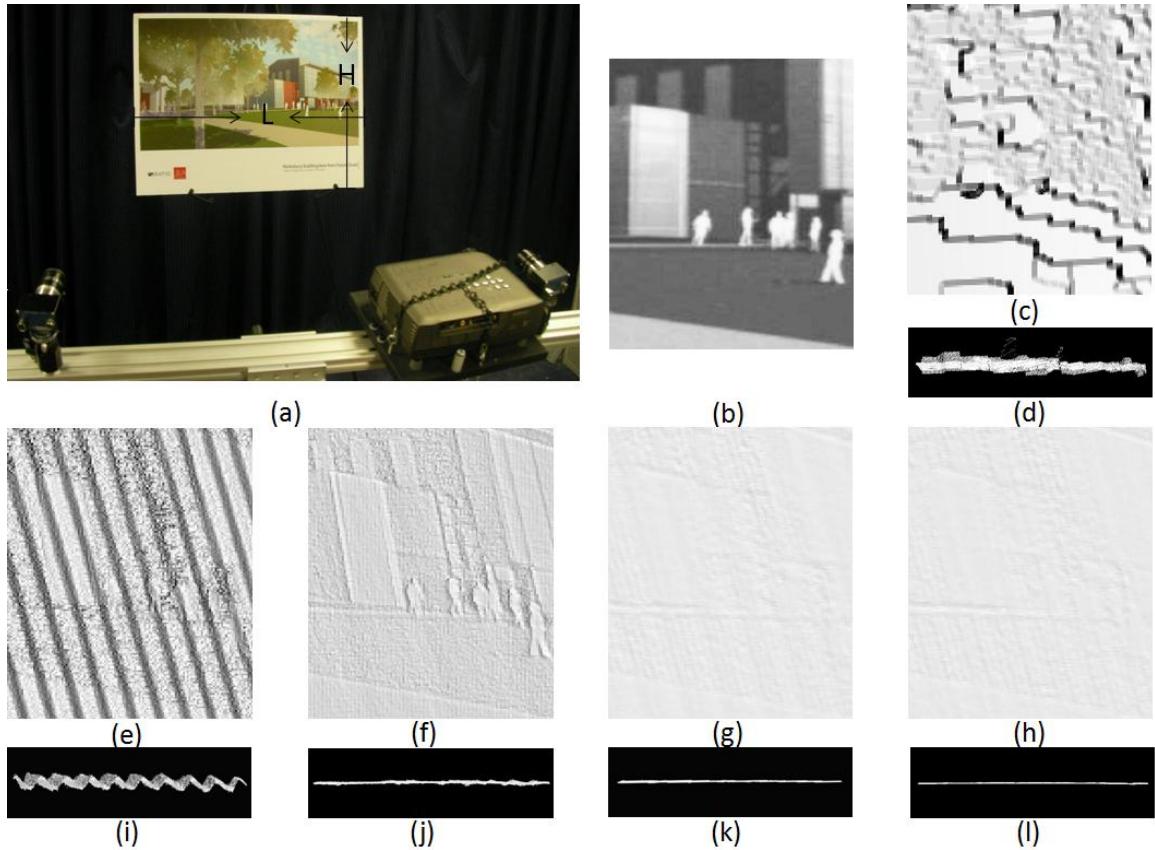


Figure 4.19: Reconstructed 3-D geometry of a textured flat board. To illustrate the problems clearly, only the reconstruction of a rectangle region is zoomed in and shown. (a) The board with rich texture under scanning. (b) The B image in the zoomed in area. Part of the building and six people figures can be seen in the region. (c) Front view of reconstruction result by using stereo-vision techniques (with the graph-cut algorithm for 6 iterations). (d) Top view of (c). (e) Front view of reconstruction result by using conventional SLI (with 3 high frequency PMP patterns). (f) Front view of reconstruction result by using conventional SLI (with 30 high frequency PMP patterns). (g) Front view of reconstruction result by using sub-pixel phase matching (with 3 high frequency PMP patterns). (h) Front view of reconstruction result by using the hybrid approach (with 3 high frequency PMP patterns). (i) Top view of (e). (j) Top view of (f). (k) Top view of (g). (l) Top view of (h). Images (c), and (e) to (h) are depth rendering images processed in the same way to show the reconstruction errors.

performance comparison among different reconstruction methods. Look-up tables were created to compensate the difference between two cameras from pixel-to-pixel and to correct the optical distortion caused by the camera lenses [131]. The system

was pre-calibrated to compute the reconstruction matrix and epipolar geometry as discussed in [75, 76].

The measured noise variance of the camera I to projector pair was 7.2827 while the variances of the two cameras were measured separately at 1.0410 and 1.1038. As discussed in Sec. (4.3), we expect phase-stereo matching between cameras I and II to outperform traditional single view SLI between camera I and the projector since their pairing noise is heavier than the summation of the two camera noises.

In order to show how the reconstruction quality using SLI is affected by the projector configuration, we measured the joint gamma and flicker distortion of the projector/camera pairs by projecting a series constant gray-scale images with consecutively increasing intensity values $\{I_g^p : g = 0, 1, 2, \dots, 255\}$. The resulting gray-level intensities of the two cameras were then measured to produce the tone reproduction curves of Fig. 4.18 where these nonlinear curves were caused by gamma distortion with the local vibration due to the noise and flicker problems. This result shows that, although heavy distortion exists in both cameras, the curves from the two cameras are near identical.

Now in order to demonstrate the advantages of the hybrid approach over conventional SLI with sub-pixel phase matching in the presence of gamma distortion, noise, projector flicker, as well as an object's texture and other effects, we performed a series of experiments beginning with scanning a textured, flat poster board (Fig. 4.19) placed at approximately 1.5 meters away, where no gamma correction was performed on the received images nor were the cameras rigorously synchronized to the projector. The cameras, themselves, were synchronized in software with each other by means of

a common fire wire bus connection to the host computer. Figure 4.19 (c) and (d) show the surfaces produced by means of stereo-vision using Kolmogorov *et al*'s graph-cut algorithm [30] applied to the B images. Although the reconstructed results differ significantly from the surfaces from Fig. 4.19 (e) to (l), they demonstrate that the B images do provide information useful for 3-D reconstruction whereas traditional stereo-vision would have matched the A images.

After having reconstructed the poster board's surface by means of traditional SLI, Fig. 4.19 (e) and (i) show the fringe pattern caused by the projector's gamma distortion. By employing 30 high frequency patterns, we can remove this fringe pattern without performing any form of gamma correction as indicated in Fig. 4.19 (f) and (j). While employing 30 patterns simultaneously reduced surface noise, there is still noticeable ringing artifacts in areas of high contrast most noticeable around step edges, which indicates the degrading effect caused by texture. And clearly in Fig. 4.19 (f), regions with low B values exhibit increased high frequency surface fluctuations due to Gaussian sensor noise.

Using sub-pixel phase matching, we achieve the reconstruction results of Fig. 4.19 (g) and (k) as well as (e) and (h) where (g) and (k) were produced without KL divergence refinement while (e) and (h) were produced with it. These results illustrate the improved smoothness that we expected by triangulating across cameras versus between camera and projector.

As a means of analysis, Fig. 4.20 shows a comparison of the disparity results using 3-pattern SLI and hybrid approaches. Since the quality of SLI depth reconstruction depends entirely on phase, the disparity errors of SLI are caused by the projector

imperfections proportional to their effect on phase. The SLI RMS registration error is, in this instance, approximately 2.6979 pixels, whereas that of the hybrid approach is only 0.1040 pixels or $1/26^{\text{th}}$ that of conventional SLI. Figure 4.21 shows a further comparison of the disparity results using sub-pixel phase matching with and without KL divergence refinement. To illustrate the effect of modulation values, B is scaled to $[1, 2]$ and also plotted in Fig. 4.21. In Area 1 where B values have small changes, the results with and without using divergence refinement contain errors of the same level; however, in Area 2 where B values are changing rapidly, the hybrid approach yields a smaller error. The overall RMS error is 0.1040 pixels with KL refinement versus 0.1464 pixels without any refinement.

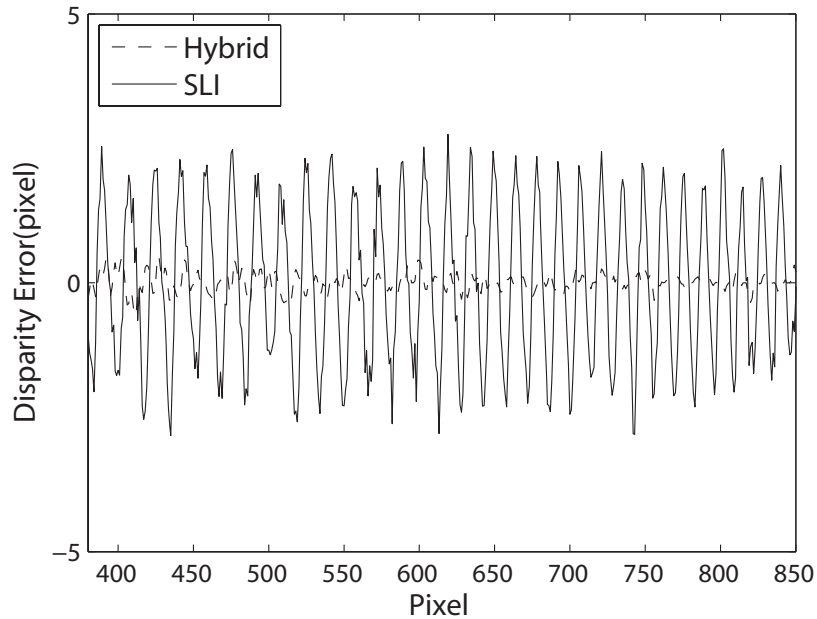


Figure 4.20: Cross sections of reconstructed plane from SLI and hybrid algorithms.

To further test the performance of the sub-pixel phase matching with KL divergence refinement versus traditional SLI, the target poster board was tilted away from the sensor array at 15° increments while still using 3 high-frequency PMP patterns.

Table 4.1: Experimental results using the SLI method

Pose	Feature	True value $v_0(mm)$	Measured value $v(mm)$	Relative error $(\Delta v/v_0, \%)$	Relative error to scene depth $(\Delta v/Z, \%)$
0	H	458	456.8500	-0.2511	-0.0320
	L	580	578.8127	-0.2047	-0.0271
1	H	458	459.6140	0.3524	0.0501
	L	580	581.5967	0.2753	0.0391
2	H	458	459.6035	0.3501	0.0576
	L	580	581.5347	0.2646	0.0435

Table 4.2: Experimental results using the phase matching method

Pose	Feature	True value $v_0(mm)$	Measured value $v(mm)$	Relative error $(\Delta v/v_0, \%)$	Relative error to scene depth $(\Delta v/Z, \%)$
0	H	458	457.9557	-0.00967	-0.0012
	L	580	579.9243	-0.01305	-0.0023
1	H	458	458.0916	0.020	0.0028
	L	580	580.1563	0.02694	0.0054
2	H	458	458.0934	0.0204	0.0034
	L	580	580.0900	0.01551	0.0037

Table 4.3: Experimental results using the hybrid algorithm

Pose	Feature	True value $v_0(mm)$	Measured value $v(mm)$	Relative error $(\Delta v/v_0, \%)$	Relative error to scene depth $(\Delta v/Z, \%)$
0	H	458	458.0341	0.00745	0.0009
	L	580	579.9455	-0.00940	-0.0017
1	H	458	458.0578	0.01261	0.0019
	L	580	580.1063	0.01832	0.0037
2	H	458	458.0758	0.01654	0.0027
	L	580	580.0684	0.01179	0.0028

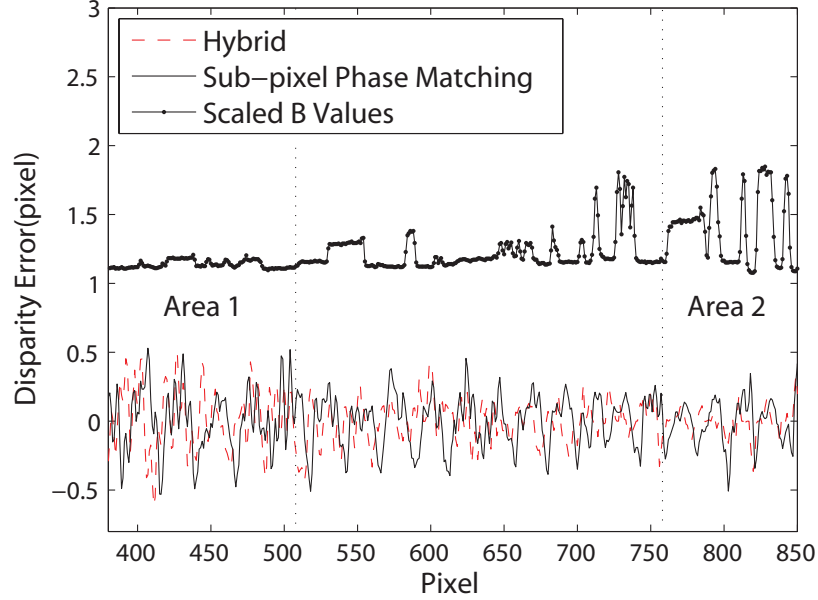


Figure 4.21: Cross sections of reconstructed plane from sub-pixel phase matching and hybrid approaches.

The height (H) and length (L) (Fig. 4.19) of the board were then measured with some typical results listed in Tables 1~3, which show that 3-D measurement errors from the SLI technique were 15 ~ 33 times larger in magnitude than those using the hybrid method, and the errors from phase matching were reduced 20% ~ 37% by employing hybrid. With only 3 patterns, the hybrid approach achieves an accuracy of 0.0009% ~ 0.0037% in scene depth measurements. Correspondingly, the errors from the hybrid reconstruction method are smaller than the state of the art approaches in [76] and [77], and very close to the errors in [75] where at least nine patterns were used.

So through the above experiments, we have demonstrated that, compared to traditional SLI, the hybrid reconstruction method is insensitive to phase errors deriving from projector flicker as well as camera/projector gamma distortion, since the new

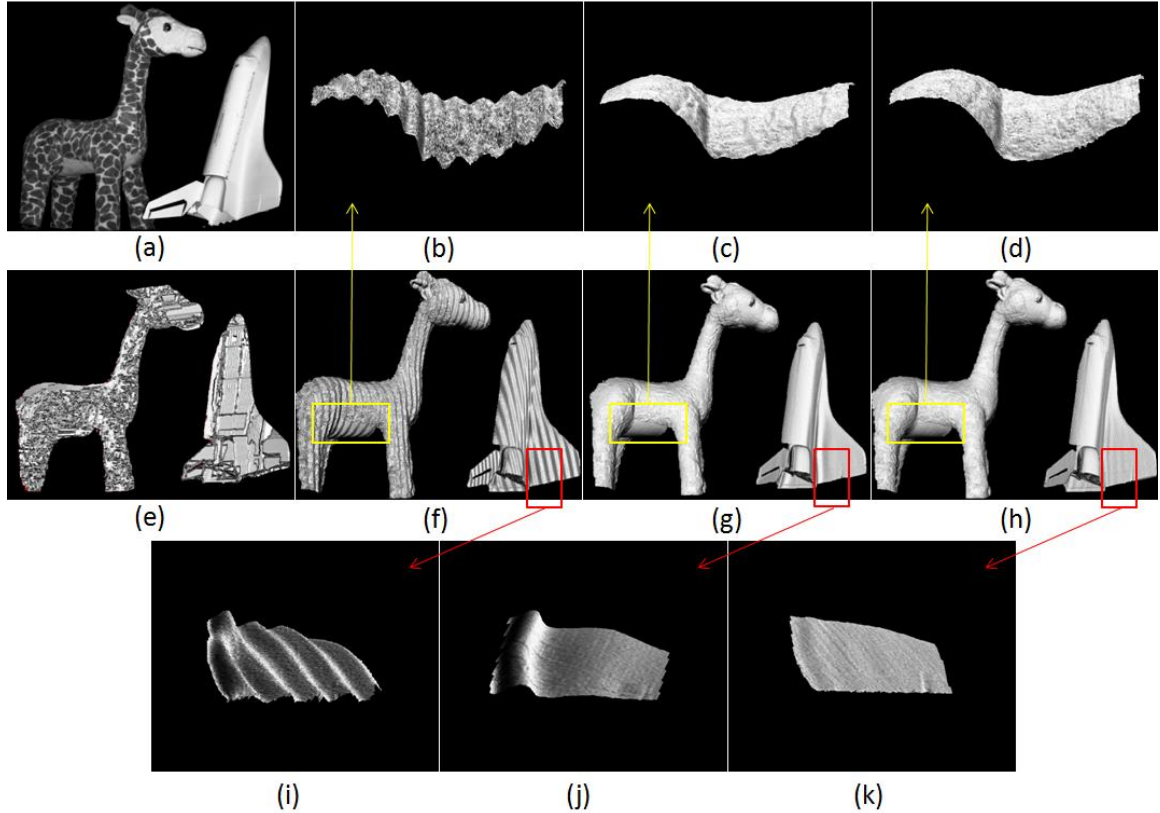


Figure 4.22: (a) A textured cartoon giraffe and a white space shuttle model. (b), (c) and (d) are zoomed-in and rotated views of (f), (g) and (h), respectively. (e) Depth rendering reconstruction by using stereo vision technique (with the graph-cut algorithm for 7 iterations). (f) Depth rendering reconstruction by using SLI technique (with 3 high frequency PMP patterns). (g) Depth rendering reconstruction by using SLI technique (with 30 high frequency PMP patterns). (h) Depth rendering reconstruction by using the hybrid approach (with 3 high frequency PMP patterns). (i), (j) and (k) are zoomed-in and rotated views of (f), (g) and (h).

method relies upon triangulation between camera sensors. Compared to sub-pixel phase matching without refinement, employing KL divergence improves the result by incorporating modulation information not previously exploited. The new approach also exhibits little impact in areas of the target surface having high-contrast texture since this texture is taken advantage of when identifying ambiguities in the phase correspondences as well as for estimating the variance parameters for KL divergence, which likewise identifies correspondences between camera images with higher preci-

sion while being resilient to noise.

Now with regards to scanning smooth surfaces with little or no texture, we reconstructed the 3-D surfaces of a rich textured cartoon giraffe along with a white space shuttle model. Here, the surface of the giraffe is coarse whereas that of the shuttle model is smooth. Again, separate 3-D reconstructions produced by means of stereo-vision, SLI, and the sub-pixel phase matching with KL divergence refinement, as illustrated shown in Figs. 4.22 (e), (f) and (h), respectively. For comparison, Fig. 4.22 (g) illustrates the reconstructed surface obtained by means of traditional SLI using 30 high frequency PMP patterns.

The giraffe's surface in Fig. 4.22 (b) is distorted by gamma and significantly noisier than by using the hybrid approach, as illustrated in Fig. 4.22 (d) and (h). By using 30 patterns, the surface in Fig. 4.22 (c) removes the gamma distortion; however, small wrinkles indicate that the reconstructed surface is degraded by the texture of the giraffe, where it can be seen that the phase quality is low. And since the hybrid

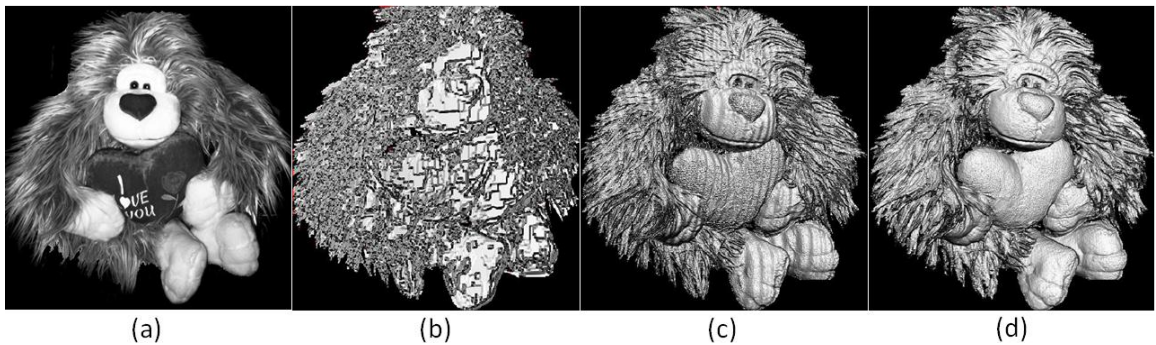


Figure 4.23: (a) A thick and shiny hair gorilla. (b) Depth rendering reconstruction by using stereo-vision technique (with the graph-cut algorithm for 9 iterations). (c) Depth rendering reconstruction by using conventional SLI technique (with 3 high frequency PMP patterns). (d) Depth rendering reconstruction by using the proposed hybrid approach (with 3 high frequency PMP patterns).

approach utilizes this texture information to complement the degraded phase data, sub-pixel phase matching with KL divergence refinement performs better.

Looking now at smooth textureless surfaces, the distortion in Fig. 4.22 (i) for traditional SLI is dominated by gamma distortion, but even after eliminating this distortion by using 30 patterns, Fig. 4.22 (j) still shows distortion caused by the reflected light from the space shuttle body, which is illustrative of the multi-light-path problem. By contrast, the hybrid approach greatly reduced this distortion in Fig. 4.22 (k), further demonstrating the improved performance of the hybrid approach in high and low textured environments in multi-light-path regions.

As the third example, we scanned a plush toy gorilla with thick and shiny hair of varying color, where the depth contrast along some edges of hair was high while the light path was complicated. The results shown in Fig. 4.23 were produced using 3, unit-frequency PMP patterns and demonstrate the reconstructions produced by stereo-vision by means of graph-cuts [30], SLI, and the hybrid approach, respectively. From visual inspection of Fig. 4.23 (c) and (d), the reconstructed result from SLI is noisy and shows a significant amount of gamma distortion. We believe that the heavy noise is mainly due to the projector’s flicker, the high contrast depth, and the multi-light-path problem.

From Fig. 4.23 (b), it can be seen that the stereo-vision reconstruction suffers little from the gamma distortion but shows significant errors caused by incorrect registration. The reconstruction produced by the hybrid approach suffers little from the gamma distortion, and has much improved registration, as shown in Fig. 4.23 (d). Using this experiment, it can be seen that the hybrid approach is less sensitive to the

projector's distortion, the condition of rich texture, multiple light paths, and high depth variation in terms of improved reconstruction performance over stereo-vision and traditional SLI techniques.

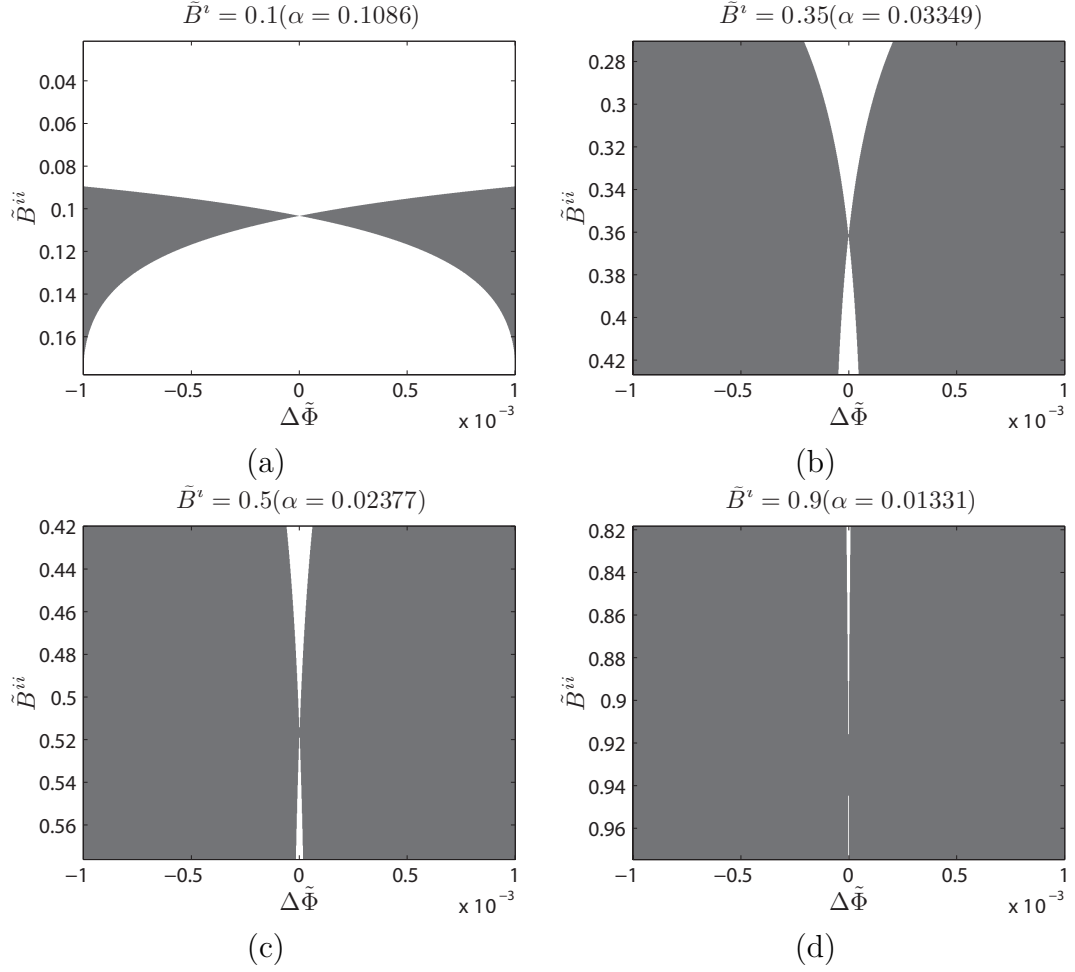


Figure 4.24: 3D visualization of Eq. (4.57) with $f = 1$, $N = 3$, $(\sigma^i)^2 = 7.2827$, and $(\sigma^i)^2 = 7.7925$. In white area, $SR > 1$. $SR < 1$ in gray area. The modulation values, B^i and B^{ii} , are scaled to $[0, 1]$. (a) The modulation values in camera I \tilde{B}^i is 0.1. (b) \tilde{B}^i is 0.35. (c) \tilde{B}^i is 0.5. (d) \tilde{B}^i is 0.9.

With the help of texture and modulation data, the hybrid approach is robust to phase data errors, yielding high quality 3-D reconstruction. Note that some high quality projectors may introduce much less gamma and noise than the one we used and that the flicker problem can be solved by rigorous synchronization between cam-

eras and projector. However, our strict adherence to using the theoretical minimum number of patterns where, regardless of the quality of hardware, the phase data quality tends to be poor. The hybrid approach improves the reconstruction performance at the cost of added computation.

Discussion

In fact, KL divergence combines information of modulation and phase data. To further study the performance of Eq. (4.55), we define the Sensitivity Ratio, SR , as

$$\begin{aligned} SR &= \left| \frac{\partial D_{KL}}{\partial \tilde{B}^{ii}} / \frac{\partial D_{KL}}{\partial \tilde{\Phi}^{ii}} \right| \\ &= \left| \frac{2}{Nf^2\alpha^2\Delta\tilde{\Phi}\tilde{B}^{ii}} \left(\frac{(\sigma^i)^2}{(\tilde{B}^i)^2} - \frac{(\sigma^{ii})^2}{(\tilde{B}^{ii})^2} \right) + \frac{\Delta\tilde{\Phi}}{\tilde{B}^{ii}} \right|, \end{aligned} \quad (4.57)$$

where $\Delta\tilde{\Phi} = \tilde{\Phi}^{ii} - \tilde{\Phi}^i$. For different modulation and phase values, Eq. (4.57) indicates whether Eq. (4.55) is more sensitive to the modulation value \tilde{B}^{ii} or the phase value $\tilde{\Phi}^{ii}$. When $SR > 1$ the KL divergence is more sensitive to \tilde{B}^{ii} than $\tilde{\Phi}^{ii}$, and when $SR < 1$ $\tilde{\Phi}^{ii}$ dominates the KL value.

As shown in Fig. 4.24, for a given \tilde{B}^i , the SR has values bigger than 1 when \tilde{B}^{ii} is small as the modulation information becomes more reliable than the phase. With increased values of \tilde{B}^{ii} , as shown in Fig. 4.24 (a), (b), (c) and (d), more regions are dominated by SR values smaller than 1, which means the KL divergence is more sensitive to phase data as the phase information becomes more reliable. Figure 4.24 also shows that when \tilde{B}^{ii} is very close to \tilde{B}^i , SR reduces rapidly and phase information dominates the KL, and when $\Delta\tilde{\Phi}$ goes to 0, SR grows fast and KL is dominated by $\tilde{\Phi}^{ii}$. Thus, Eq. (4.55) actually trades off between phase and modulation information

based on their reliability using different values of α .

In order to study the dependence of scan quality on pattern number, we measured the scan errors of the hybrid approach and traditional SLI method, using the textured poster board shown in Fig. 4.19, with different pattern numbers. The results in Fig. 4.25 show that the RMS disparity between SLI and hybrid approaches decreases as the pattern number increases. The reconstruction errors of the hybrid approach are smaller across the analyzed scale, especially for small pattern numbers, which we listed in Table 4.4. The existence of error under high pattern numbers, we believe, is mainly caused by calibration and inaccuracy of camera and projector models. As most structured light systems in practice employ only a small number of patterns, it can be concluded that the proposed hybrid approach has the great superiority in real world applications of 3D reconstruction.

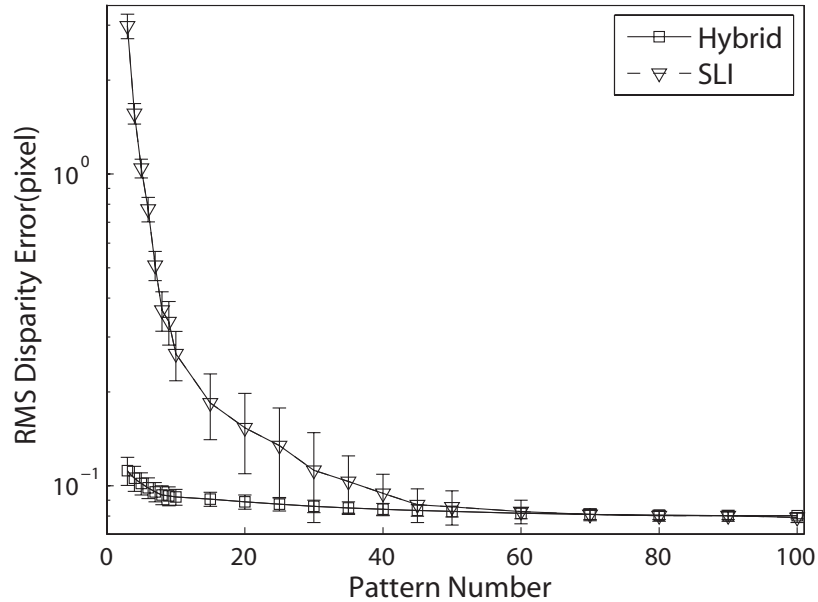


Figure 4.25: RMS depth errors change with number of patterns.

Although our experiments were performed using only two cameras and one projec-

Table 4.4: Error analysis for SLI and hybrid approaches

Pattern number	Error of SLI $E_{SLI}(pixel)$	Error of hybrid $E_{hybrid}(pixel)$
3	2.9859	0.1119
4	1.5644	0.1058
5	1.0435	0.1018
6	0.7719	0.0986

tor, the proposed stereo-phase matching approach can be extended to an N -camera, M -projector system, where the patterns in the M projectors are well aligned through pre-rectification, a process similar to camera rectification as presented in [132]. Accordingly, the N cameras would need to be positioned such that every neighboring two cameras share some common regions. The phase, modulation, and texture information obtained from these cameras would then be defined according to

$$I^1 = \{S_1^1, S_2^1, \dots, S_N^1\}, \quad (4.58)$$

where I^1 notes the information of the first level. The hybrid approach would then be applied in a bottom-up manner, as illustrated in Fig. 4.26.

At the first level, stereo-phase matching and refinement would be performed for each pair of neighboring cameras, with the results passed on to the second level I^2 where the registration result is converted to phase, modulation, and texture data, which is then fused at higher levels. The final 3-D reconstruction is then a top-down process, where 3-D surfaces exclusively obtained at the lower levels are attached to the surfaces reconstructed at the higher levels.

This chapter presents a novel hybrid 3-D reconstruction approach for multi-camera SLI systems. The major contribution of this chapter is to achieve high quality 3-D

reconstructions with only a small number of illumination patterns by maximizing the use of correspondence information from the phase, texture, and modulation data derived from multi-view, PMP-based, SLI images. The proposed approach is insensitive to gamma distortion and projector flicker that is, otherwise, present if the cameras and projectors are not rigorously synchronized nor have the device gammas been calibrated. Experimental results show that the hybrid approach can achieve superior 3-D reconstruction performances in rich/poor texture, multi-light path, and/or high/low contrast environments regardless of the phase data quality. Specifically, we achieved a $26\times$ reduction in error compared to traditional SLI.

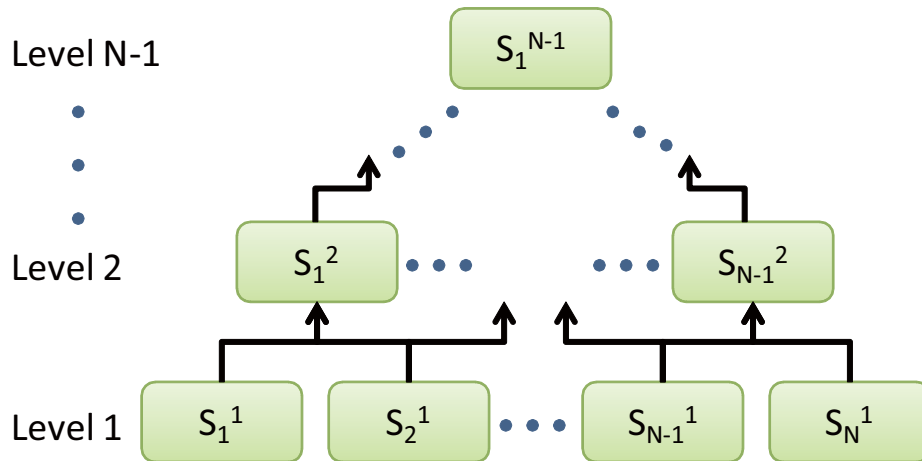


Figure 4.26: Pyramid generalization of the hybrid 3-D reconstruction approach to multi-camera and multi-projector SLI systems. S_1^2 represents the fused information from the common region of S_1^1 and S_2^1 . S_1^{N-1} is the fused information from all cameras.

Chapter 5 CONCLUSION AND FUTURE WORK

Structured light illumination (SLI) is a method of 3-D scanning that has been developing rapidly and applied in various domains. In this dissertation, we focused on the phase shifting methods (PSMs) which involve projecting the same pattern, but shifting it in a certain direction in order to increase resolution. Compared to other approaches of SLI, the PSMs overcome the discrete nature of patterns and are robust to object's texture and ambient environment light. And due to the fact that the pattern resolutions are exponentially increasing among the coarse-to-fine light projections and the fringe gap tends to 0, the resolution of PSMs is greatly improved. However in practice, error is introduced from camera noise, projector noise, and ambient light noise as well as other sources. And these methods show large measurement error because of noise, especially when only a few patterns are employed, as is the case in real-time systems.

In this research, we focused on how to improve the accuracy of PSMs. There are three major contributions of this dissertation. First, this dissertation studied the specific technique of PMP and the maximization of a pattern's SNR in Chapter II. By treating the design of an N -pattern PMP process as placing points in an N -dimensional coding space, we defined a pattern's SNR in terms of a pattern set's computational length and the number of sinusoidal periods in the projected patterns. Then, without introducing phase ambiguities, I proposed a so-called edge-pattern strategy that maximizes the computational length and number of periods. Specifi-

cally, the edge-pattern technique improves the SNR by 1.2381 times when using three component patterns and 15.5421 times when using five patterns.

Experimental results demonstrate that by maximizing the computational length and employing the non-ambiguous high frequency, the E-P strategy achieves a higher SNR than the traditional PMP strategy. With more patterns, the pattern SNR improvement increases further; however in practice, I noticed that the difference between theoretical and experimental improvements also increases with the number of patterns. This is caused, I believe, by the de-focus and resolution difference of the projector and camera pair.

For future research of edge-pattern, I will further exploit the E-P strategy such that the de-focus and resolution difference problems can be addressed. In experiments, I noticed that when the number of patterns was 3, the edge-pattern was robust to the de-focus and resolution difference. However, when the number of patterns was larger than 3, the error appeared. Thus, I believe, some edges in the non-ambiguous high frequency patterns introduces the sensitivity to de-focus and resolution difference. A possible solution to it is to reduce the non-ambiguous high frequency. Such kind of trade-off reduces the theoretical accuracy of the patterns. However, in practice, it may avoid the error caused by de-focus and resolution difference such that the theoretical and experimental accuracies would be close and the overall accuracy can be improved.

A second contribution to SLI research made in this dissertation, presented in Chapter III, was a period coded phase shifting pattern strategy where PSMs involve projecting a set of grating patterns where measurement accuracy is determined, in

part, by the number of grating periods. But by increasing the number of periods in the projected pattern set, ambiguity problems are introduced during the process of phase unwrapping. In order to eliminate these ambiguities, I defined a spatial intensity efficiency measure that, for those pattern sets with a spatial intensity efficiency of less than 100%. Without reducing the *SNR* of original high frequency signal, I encoded the period cue into the projected pattern set such that each period, of the multi-period pattern, is uniquely identifiable. This Chapter further introduced pattern entropy as a separate measure from which one can optimize the embedded period cue. In combination, the proposed method can unwrap high frequency phase information and achieve high measurement precision without increasing the number of projected patterns and, therefore, has significant benefits when scanning moving objects.

Further, in Chapter III, a noise model was introduced and verified by experiment, in order to demonstrate the high measurement accuracy of the proposed approach. Finally, a prototype system is demonstrated that can achieve 120 fps at 640×480 resolution for 3-D data acquisition and reconstruction, which demonstrates that the computational cost of the proposed PCPS is low. For future research of this chapter, I will further exploit the remaining dynamic range such that the accuracy of period cue and the number of periods in the 3 pattern strategy can be improved. I will also work on the optimization between the dynamic ranges of high frequency phase signal and the period cue. Currently, the high frequency phase signal employs the same dynamic range as the traditional PSMs. In order to achieve a high accuracy, most PSMs occupy a large portion of the dynamic range such that the left for period cue is generally small. Using such small dynamic range, the information can be

embedded is also limited. In practice, it is possible that the dynamic range of the original high frequency signal is reduced such that the range left for period cue can be increased. It would reduce the accuracy of high frequency phase before unwrapping. However, by doing so, the period cue becomes more accurate such that the frequency can be further increased. Thus, the final phase used for 3-D reconstruction after unwrapping is improved. The trade-off between the dynamic range of original high frequency phase and that of period cue should be further studied and optimized in order to achieve the highest accuracy of final unwrapped phase.

At last, a hybrid 3-D framework was developed in Chapter IV. The system combines the SLI and stereo vision techniques. By employing multiple cameras, SLI can generate pattern phase/modulation and object texture data. In Chapter IV, I first developed a new method to find the correspondence between the two cameras using both the phase information generated by the temporal multiplexed illumination patterns and stereo triangulation. I also analyzed the resulting correspondence accuracy as a function of the number of structured patterns as well as the geometric position of projector to cameras. I then presented a hybrid 3-D reconstruction framework through phase-modulation-texture data fusion under temporal multiplexed illumination. The proposed scheme consists of, first, using phase data to derive initial correspondences across cameras. Second, texture data is used to eliminate correspondence ambiguities. Third, modulation data is used to estimate correspondence error ranges. Finally, Kullback-Leibler divergence refinement, based on the derived phase error models, is performed to reduce mis-registration among images. Using only a small number of light patterns, the presented approach significantly reduces measurement errors

versus traditional structured-light methodologies while being insensitive to gamma distortion and projector flicker. Experimental results demonstrate these advantages in terms of enhanced 3-D reconstruction performance and robustness against noise, distortions, and conditions of texture and contrast. The future works for this chapter include an extension of this approach to N -camera and M -projector SLI systems and the development of single-pattern, multi-view, SLI systems. As studied by C. Guan et. al. in [71], the single pattern PSM can be achieved by composing the multiple patterns into one. However, in practice, the pattern becomes sensitive to noise. Since the hybrid technique greatly improves the accuracy and robustness to noise, it is possible that a satisfying 3-D reconstruction can be achieved by combining the composite pattern and hybrid techniques. The pattern should be designed similar to the method of composite pattern such that it would be robust to the albedo and ambient light. And the phase changing direction should be along with the epipolar geometry of N -camera system as studied in Chapter IV. Thus, although the quality of obtained phase from each camera may be low, each camera introduces both phase and albedo information and all the information can be employed through the hybrid approach such that high quality 3-D reconstruction can be still achieved.

My future research should also include the reducing of the number of patterns for PSMs. A triangulation pattern strategy has been proposed by Jia, which achieves two patterns. However, this method is sensitive to the albedo of the scene and the environment ambient light. In practice, the fact that two pattern or one pattern PSM is hard to achieve mainly due to the low-pass filter effect of the captured images, which can be caused by de-focus and resolution difference. To overcome the low pass

filter effect, the differences between pixels should be improved. By the restriction of epipolar line, it is possible for us to design a pattern, such that, in the captured images, the projected intensities in a vertical or horizontal line are the same. Thus, we can improve the differences between every neighboring vertical or horizontal lines such that we can assume that the neighboring lines are not projected at the same time but in two or even three projections. Then, the first estimated phase can be achieved by using only one or two patterns. For example, we can first assume the environment ambient light is changing slowly on the object's surface, which is generally true. Then, we can obtain the phase by using only two patterns. After obtaining the initial phase, the phase can be iteratively refined by removing the assumption and refining the albedo and ambient light values. Thus, it becomes possible for us to develop two or even one pattern's PSM which would be robust to the environment light and albedo of the scene.

3-D scanning is important to analyze a real-world object or environment. Some different technologies can be used to build these 3D scanning devices, including stereo vision, time of flight, and structured light illumination. Stereo vision usually employs two video cameras and the 3-D information is revealed based on the differences between the images seen by each camera. The system is easy to set up but the accuracy, compared with the other two, is low. The time-of-flight 3D scanner is an active scanner that uses laser light to probe the subject. Currently, the time-of-flight 3D scanners are drawing significant attention from the machine vision community because they offer real-time operation. Thus, it becomes possible for many applications, for example human and machine interface, to scan fast moving object and use the real-time

3-D information. However, compared to the SLI technique, the time of flight scanner has a lower resolution and accuracy. Among the 3-D scanning techniques, SLI has played a fundamental role. By employing high resolution cameras, SLI can achieve an ultra-high resolution and accuracy. But traditional SLI approaches require a set of patterns projected in order to achieve the high accuracy. Thus, the scanning speed is considerably low. If SLI is going to maintain a place in machine vision, it is going to have to work at or near real-time video rates. Hence, my research particularly focuses on the efficiency of patterns such that with less number of patterns the SLI systems can get the same accuracy and the scanning time is then reduced. Three different approaches are proposed in this dissertation with significant reduction of the number of patterns. As the real-time operation becomes a demand on machine vision systems, industry would have no choice to either adopt these three methods or develop new methods that attempt to repeat what I have done with focusing on pattern efficiency.

Copyright© Yongchang Wang, 2010.

Bibliography

- [1] V. Yalla, "Optimal pmp techniques for static and dynamic 3d data acquisition," Jun 2007.
- [2] O. D. Faugeras and M. Hebert, "The representation, recognition, and locating of 3-d objects," *International Journal of Robotic Research*, vol. 5, no. 3, pp. 27–52, fall 1986.
- [3] P. J. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, Feb 1992.
- [4] W. A., P. G., and H. J., "Probing systems for dimensional micro- and nano-metrology," *Meas. Sci. Technol.*, vol. 17, p. 504509, 2006.
- [5] Wikipedia, "3d scanner," http://en.wikipedia.org/wiki/3D_scanner.
- [6] 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 Instrum*, vol. 76, no. 9, p. 095112, 2005.
- [7] R. J. Cotter, "Time-of-flight mass spectrometry," *Columbus, OH: American Chemical Society, ISBN 0-8412-3474-4*, 1994.
- [8] Optech, "www.optech.ca," 2006.
- [9] Trimble, "www.trimble.com," 2006.

- [10] L. Geosystems, “www.leica-geosystems.com,” 2006.
- [11] Riegl, “www.riegl.com,” 2006.
- [12] Faro, “www.faro.com,” 2006.
- [13] Isite, “www.isite3d.com,” 2006.
- [14] Z. Frohlich, “www.zofre.de,” 2006.
- [15] InteliSum, “www.rappidmapper.net,” 2006.
- [16] C. F. L. H. Division, “Ground-based lidar, rock slope mapping and assessment,” 2008.
- [17] M. Imaging, “Sr4000 time-of-flight range imaging,” *http : //www.mesa – imaging.ch/*.
- [18] A. P. Cracknell and L. Hayes, “Introduction to remote sensing (2 ed.),” *London: Taylor and Francis, ISBN 0849392551, OCLC 70765252*, 1991.
- [19] D. Modarress, P. Svitek, K. Modarress, and D. Wilson, “Micro-optical sensors for boundary layer flow studies,” *2006 ASME Joint U.S.-European Fluids Engineering Summer Meeting*, *http : //www.measurementsci.com/papers/FEDSM0698556.pdf*, p. 095112, 2006.
- [20] R. J. Valkenburg, “Classification of camera calibration techniques,” *IVCNZ95, Lincoln, New Zealand*, pp. 43–48, Aug 1995.

- [21] R. Y. Tsai, “A versatile camera calibration technique for high accuracy 3d machine vision metrology using off-the shelf tv cameras and lenses,” *IEEE Int. J. Robot. Automat.*, vol. RA, no. 3, 1987.
- [22] H. Farid and A. C. Popescu, “Blind removal of lens distortion,” *JOSA A.*, vol. 18, no. 9, 2001.
- [23] S. Negahdaripour, H. Sekkati, and H. Pirsiavash, “Opti-acoustic stereo imaging: on system calibration and 3-d target reconstruction,” *IEEE Transactions on Image Processing*, vol. 18, no. 6, pp. 1203–1214, 2009.
- [24] K. Liu, Y. Wang, D. L. Lau, Q. Hao, and L. G. Hassebrook, “Dual-frequency pattern scheme for high-speed 3-d shape measurement,” *Optics Express*, vol. 18, no. 5, pp. 5229–5244, Mar. 2010.
- [25] U. Dhond and J. Aggarwal, “Structure from stereo - a review,” *IEEE Trans. Syst., Man and Cybern.*, vol. 19, pp. 1489–1510, 1989.
- [26] S. Pollard, J. Mayhew, and J. Frisby, “Pmf: A stereo correspondence algorithm using a disparity gradient limit,” *Perception*, vol. 14, no. 4, pp. 449–470, 1985.
- [27] R. Hartley, R. Gupta, and T. Chang, “Stereo from uncalibrated cameras,” *Proceedings of the Conference on Computer Vision and Pattern Recognition, Urbana-Champaign, Illinois*, pp. 761–764, 1992.
- [28] Wikipedia, “Stereopsis,” [http : //en.wikipedia.org/wiki/Stereopsis](http://en.wikipedia.org/wiki/Stereopsis).

- [29] S. Zhang, D. Royer, and S. T. Yau, “High-resolution, real-time-geometry video acquisition system,” *ACM SIGGRAPH*, no. 14, 2006.
- [30] V. Kolmogorov, R. Zabih, and S. Gortler, “Multi-camera scene reconstruction via graph cuts,” *In European Conference on Computer Vision*, May 2002.
- [31] G. Frankowski, M. Chen, and T. Huth, “Real-time 3d shape measurement with digital stripe projection by texas instruments micromirror devices (dmd),” *Proc. Of SPIE*, vol. 3958, p. 90106, 2000.
- [32] Wikipedia, “Structured-light 3d scanner,” *http* :
 //en.wikipedia.org/wiki/Structured-light_3d_scanner.
- [33] D. Scharstein and R. Szeliski, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithm,” *International Journal of Computer Vision*, 2002.
- [34] —, “High-accuracy stereo depth maps using structured light,” *CVPR*, vol. 1, pp. 195–202, Jul 2003.
- [35] J. Pages, J. Salvi, R. Garcia, and C. Matabosch, “Overview of coded light projection techniques for automatic 3d profiling,” *Proc. 2003 IEEE Conf. Robotics and Automation*, pp. 133–138, 2003.
- [36] J. Pages, C. Collewet, F. Chaumentte, and J. Salvi, “Optimizing plane-to-plane positioning tasks by image-based visual servoing and structured light,” *IEEE Transactions on Robotics*, vol. 22, no. 5, pp. 1000–1010, 2006.

- [37] C. Kim, J. Park, J. Yi, and M. Turk, “Structured light based depth edge detection for object shape recovery,” *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, 2005.
- [38] P. Fechteler and P. Eisert, “Adaptive color classification for structured light systems,” *Proc. of CVPR*, 2008.
- [39] T. Peng and S. Gupta, “Model and algorithms for point cloud construction using digital projection patterns,” *Journal of Computing and Information Science in Engineering*, vol. 7, no. 4, pp. 372–381, 2008.
- [40] J. Salvi, J. Pages, and J. Batlle, “Pattern codification strategies in structured light systems,” *Pattern Recognition*, vol. 37, no. 4, pp. 827–849, 2004.
- [41] P. Fechteler, P. Eisert, and J. Rurainsky, “Fast and high resolution 3d face scanning,” *Proc. of ICIP*, 2007.
- [42] W. J., C. P., and H. M., “Camera calibration with distortion models and accuracy evaluation,” *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. 14, no. 10, pp. 965–980, 1992.
- [43] Y. Wang, D. L. Lau, and L. G. Hassebrook, “Fit-sphere unwrapping and performance analysis of 3d fingerprints,” *Applied Optics*, vol. 49, no. 4, pp. 592–600, Feb. 2010.
- [44] Y. Wang, Q. Hao, A. Fatehpuria, D. L. Lau, and L. G. Hassebrook, “Data acquisition and quality analysis of 3-dimensional fingerprints,” *IEEE conference on Biometrics, Identity and Security, Tampa, Florida*, Sep. 2009.

- [45] C. Guan, L. G. Hassebrook, and D. L. Lau, “Composite structured light pattern for three-dimensional video,” *Optics Express*, vol. 11, no. 5, 2003.
- [46] O. Faugeras and Q. T. Luong, “The geometry of multiple images,” *MIT Press*, 2001.
- [47] C. Loop and Z. Zhang, “Computing rectifying homographies for stereo vision,” *CVPR*, vol. I, no. 2, pp. 125–131, 1999.
- [48] R. I. Hartley and A. Zisserman, “Multiple view geometry,” *Combridge University Press*, 2000.
- [49] L. Zhang, B. Curless, and S. M. Seitz, “Rapid shape acquisition using color structured light and multi-pass dynamic programming,” *In Int. Symp. on 3D Data Processing, Visualization, and Transmission*, 2002.
- [50] H. li, R. Straub, and H. Prautzsch, “Structured light based reconstruction under local spatial coherence assumption,” *in Proc. 3rd IEEE Int. Symp. 3D Data Processing, Visualization and Transmission, Washington DC*, pp. 575–582, Jun. 2006.
- [51] J. Pages and J. Salvi, “Optimised de bruijn patterns for one-shot shape acquisition,” *Image Vis. Comput.*, vol. 23, no. 8, pp. 707–720, 2005.
- [52] T. P. Koninckx and L. V. Gool, “Real-time range acquisition by adaptive structured light,” *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. 28, no. 3, pp. 432–445, 2006.

- [53] S. Y. Chen, Y. F. Li, and J. Zhang, "Vision processing for realtime 3-d data acquisition based on coded structured light," *IEEE Transaction on Image Processing*, vol. 17, no. 2, pp. 167–176, 2008.
- [54] A. Dipanda and S.Woo, "Towards a real-time 3d shape reconstruction using a structured light system," *Pattern Recognition*, vol. 38, p. 1632C1650, 2005.
- [55] M. Young and E. Beeson, "Viewpoint-coded structured light," *IEEE CVPR*, Jun. 2007.
- [56] C. Je, S. Lee, and R. Park, "High contrast color stripe pattern for rapid structured-light range imaging," *Proc. Eighth European Conf. Comp. Vision*, pp. 95–107, 2004.
- [57] C. Sinlapeecheewa and K. Takamasu, "3d profile measurement using color multi-line stripe pattern with one scanning," *Integr. Comput. Aided Eng.*, vol. 12, pp. 333–341, 2005.
- [58] L. Zhang, B. Curless, and S. Seitz, "Spacetime faces: High resolution capture for modeling and animation," *ACM Trans. Graphics (Proc: ACM SIGGRAPH)*, vol. 23, no. 3, pp. 548–558, 2004.
- [59] ———, "Spacetime stereo: Shape recovery for dynamic scenes," *In CVPR*, 2003.
- [60] T. Weise, B. Leibe, and L. V. G. and, "Fast 3d scanning with automatic motion compensation," *IEEE CVPR*, pp. 1–8, 2007.

- [61] S. Rusinkiewicz, O. Hall-Holt, and M. Levoy, “Real-time 3d model acquisition,” *ACM Trans. Graphics (Proc: ACM SIGGRAPH 02)*, vol. 21, no. 3, 2002.
- [62] O. Hall-Holt and S. Rusinkiewicz, “Stripe boundary codes for realtime structured-light range scanning of moving objects,” in *8th IEEE ICCV*, 2001.
- [63] P. S. Huang, S. Zhang, and F. P. Chiang, “Trapezoidal phase-shifting method for 3-d shape measurement,” *SPIE Bellingham WA*, vol. 5606, 2004.
- [64] A. K. Wong, P. Niu, and X. He, “Fast acquisition of dense depth data by a new structured light scheme,” *Computer Vision and Image Understanding*, vol. 98, no. 3, pp. 398–422, 2005.
- [65] P. Jia, J. Kofman, and C. English, “Two-step triangular-pattern phase-shifting method for three dimensional object-shape measurement,” *Optical Engineering*, vol. 46, no. 8, 2007.
- [66] S. Zhang and S. T. Yau, “High-resolution, real-time three-dimensional shape measurement,” *Optical Engineering*, vol. 45, no. 12, 2006.
- [67] C. Rocchini, P. Cignoni, C. Montani, P. Pingi, and R. Scopigno, “A low cost 3d scanner based on structured light,” In: *A. Chalmers, T.-M. Rhyne (Eds.), EG 2001 Proceedings*, vol. 20, no. 3, pp. 299–308, 2001.
- [68] D. Skocaj and A. Leonardis, “Range image acquisition of objects with non-uniform albedo using structured light range sensor,” In: *Proceedings of the 15th International Conference on Pattern Recognition*, vol. 1, pp. 778–781, 2000.

- [69] E. Horn and N. Kiryati, "Toward optimal structured light patterns," *Image Vision Comput.*, vol. 17, no. 2, pp. 87–97, 1999.
- [70] V. Srinivasan, H. Liu, and M. Halioua, "Automated phase-measuring profilometry of 3d diffuse objects," *Applied Optics*, vol. 23, no. 18, pp. 3105–3108, 1984.
- [71] C. Guan, L. G. Hassebrook, D. L. Lau, V. G. Yalla, and C. J. Casey, "Improved composite-pattern structured-light profilometry by means of postprocessing," *Optical Engineering*, vol. 47, no. 9, 2008.
- [72] V. G. Yalla and L. G. Hassebrook, "Very high resolution 3-d surface scanning using multi-frequency phase measuring profilometry," *SPIE Defense and Security, Spaceborne Sensors II, Orlando, FL*, vol. 5798-09, Mar. 28 2005.
- [73] D. Caspi, N. Kiryati, and J. Shamir, "Range imaging with adaptive color structured light," *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. 20, no. 5, pp. 470–480, 1998.
- [74] J. Li, L. G. Hassebrook, and C. Guan, "Optimized two-frequency phase-measuring-profilometry light-senso temporal-noise sensitivity," *J. Opt. Soc. Am.*, vol. 20, no. 1, 2003.
- [75] H. Liu, W. Su, K. Reichard, and S. Yin, "Calibration-based phase-shifting projected fringe profilometry for accurate absolute 3d surface profile measurement," *Opt. Commun.*, vol. 216, no. 1-3, p. 6580, 2003.

- [76] P. J. Tavares and M. A. Vaz, "Linear calibration procedure for the phase-to-height relationship in phase measurement profilometry," *Opt. Commun.*, vol. 274, no. 2, p. 307314, 2007.
- [77] R. Anchini, G. D. Leo, C. Liguori, and A. Paolillo, "A new calibration procedure for 3-d shape measurement system based on phase-shifting projected fringe profilometry," *IEEE trans. on Instrumentation and measurement*, vol. 58, no. 5, pp. 1291–1298, May 2009.
- [78] J. Ghring, "Dense 3-d surface acquisition by structured light using off-the-shelf components," in *Proc. SPIE Videometrics Optical Methods 3D Shape Measurement*, p. 220C231, 2001.
- [79] S. Zhang and S. Yau, "High-speed three-dimensional shape measurement system using a modified two-plus-one phase shifting algorithm," *Optical Engineering*, vol. 46, no. 11, pp. 113 603–1–113 603–6, 2007.
- [80] P. Wizinowich, "Phase shifting interferometry in presence of vibration: A new algorithm and system," *Appl. Opt.*, vol. 29, pp. 3271–3279, 1990.
- [81] P. Jia, J. Kofman, and C. English, "Intensity-ratio error compensation for triangular-pattern phase-shifting profilometry," *J. Opt. Soc. Am. A*, vol. 24, pp. 3150–3158, 2007.
- [82] P. S. Huang and S. Zhang, "Fast three-step phase-shifting algorithm," *Applied optics*, vol. 45, no. 21, Jul 2006.

- [83] K. Liu, Y. Wang, D. L. Lau, Q. Hao, and L. G. Hassebrook, "Gamma model and its analysis for phase measuring profilometry," *Journal of the Optical Society of America A*, vol. 27, no. 3, pp. 553–562, Mar. 2010.
- [84] S. Zhang and P. S. Huang, "Phase error compensation for a 3-d shape measurement system based on the phase-shifting method," *Optical Engineering*, vol. 46, no. 6, Jan 2007.
- [85] B. Kamgar-Parsi and B. Kamgar-Parsi, "Evaluation of quantization error in computer vision," *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. 11, pp. 929–939, 1989.
- [86] S. Savarese, J. Bouguet, and P. Perona, "3d depth recovery with grayscale structured lighting," *Caltech Technical Report*, 1999.
- [87] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur, "A theory of multiplexed illumination," *In IEEE ICCV*, vol. 2, 2003.
- [88] P. L. Wizinowich, "Phase shifting interferometry in the presence of vibration: A new algorithm and system," *Applied optics*, vol. 29, p. 3271C3279, 1990.
- [89] M. F. Salfity, P. D. Ruiz, J. M. Huntley, M. J. Graves, R. Cusack, and D. A. Beaugard, "Branch cut surface placement for unwrapping of undersampled three-dimensional phase data: application to magnetic resonance imaging arterial flow mapping," *Applied Optics*, vol. 45, pp. 2711–2722, 2006.
- [90] T. J. Flynn, "Two-dimensional phase unwrapping with minimum weighted discontinuity," *J. Opt. Soc. Am A*, vol. 14, pp. 2692–2701, 1997.

- [91] M. A. Merraez, J. G. Boticario, M. J. Labor, and D. R. Burton, “Agglomerative clustering-based approach for two-dimensional phase unwrapping,” *Applied Optics*, vol. 44, pp. 1129–1140, 2005.
- [92] J.-J. Chyou, S.-J. Chen, and Y.-K. Chen, “Two-dimensional phase unwrapping with a multichannel least-mean-square algorithm,” *Applied Optics*, vol. 43, pp. 5655–5661, 2004.
- [93] H. O. Saldner and J. M. Huntley, “Temporal phase unwrapping: application to surface profiling of discontinuous objects,” *Applied Optics*, vol. 36, pp. 2770–2775, 1997.
- [94] J. Guhring, “Dense 3-d surface acquisition by structured light using off-the-shelf components,” *Videometrics Opt. Meth. 3D Shape Meas.*, vol. 4309, pp. 220–231, 2001.
- [95] R. M. Goldstein, H. A. Zebker, and C. L. Werner, “Two dimensional phase unwrapping,” *Radio Sci.*, vol. 23, pp. 713–720, 1988.
- [96] S. Zhang, X. Li, and S. T. Yau, “Multilevel quality-guided phase unwrapping algorithm for real-time three-dimensional shape reconstruction,” *Applied Optics*, vol. 46, no. 1, Jan. 2007.
- [97] M. Ribo and M. Brandner, “State of the art on vision-based structured light system for 3d measurements,” in *Proc. IEEE Int. Workshop on Robotic Sensors: Robotic and Sensor Environments, Ottawa, CA*, vol. 22, no. 11, p. 2, Sep. 2005.

- [98] J. Guehring, “Dense 3-d surface acquisition by structured light using off-the-shelf components,” *Proc. SPIE Vol. 4309*, pp. 220–231, 2001.
- [99] H. K. George, L. T. Watkins, and R. J. Benjamin, “Method and device for reduction in noise in images from shiny parts,” *US Patent 6678057*, Jan. 2004.
- [100] G. Sansoni, M. Carocci, and R. Rodella, “Calibration and performance evaluation of a 3-d imaging sensor based on the projection of structured light,” *IEEE Transactions on Instrumentation and Measurement*, vol. 49, no. 3, pp. 628–636, 2000.
- [101] P. S. Huang and S. Zhang, “Fast three-step phase-shifting algorithm,” *Applied optics*, vol. 45, no. 21, Jul. 2006.
- [102] S. Zhang, D. Royer, and S. Yau, “Gpu-assisted high-resolution, real-time 3-d shape measurement,” *Opt. Express*, vol. 14, pp. 9120–9129, 2006.
- [103] C. E. Shannon, “A mathematical theory of communication,” *The Bell System Technical Journal*, vol. 27, no. 4, pp. 379C423, 623C656, 1948.
- [104] J. Diaz, E. Ros, R. Carrillo, and A. Prieto, “Real-time system for high-image resolution disparity estimation,” *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 280–285, 2007.
- [105] H. Gudbjartsson and S. Patz, “The rician distribution of noisy mri data,” *Magn Reson Med.*, Dec. 1995.

- [106] Y. F. Li and S. Y. Chen, "Automatic recalibration of an active structured light vision system," *IEEE trans. on Robotics and automation*, vol. 19, no. 2, pp. 259–268, Apr. 2003.
- [107] C. Steger, "An unbiased detector of curvilinear structures," *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. 20, no. 2, p. 113C125, Feb. 1998.
- [108] J. Sun, G. Zhang, Z. Wei, and F. Zhou, "Large 3d free surface measurement using a mobile coded light-based stereo vision system," *Sens. Actuators A, Phys.*, vol. 132, no. 2, p. 460C471, Nov. 2006.
- [109] R. Bergevin, M. Soucy, H. Gagnon, , and D. Laurendeau, "Towards a general multi-view registration technique," *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. 18, no. 5, p. 540C547, May 1996.
- [110] K. Pulli, "Multiview registration for large data sets."
- [111] J. Davis, D. Nehab, R. Ramamoorthi, and S. Rusinkiewicz, "Space stereo: A unifying framework for depth from triangulation," *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. 27, no. 2, pp. 296–302, Feb 2005.
- [112] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms." *In Proc. CVPR*, 2006.
- [113] S. Zhang and S. T. Yau, "Three-dimensional data merging using holoimage," *Optical Engineering*, vol. 47, no. 3, March 2008.

- [114] R. Furukawa and H. Kawasaki, “Dense3d reconstruction with an uncalibrated stereo system using coded structured light,” *In CVPR*, pp. 107–107, Jun. 2005.
- [115] Faugeras, O. D., and T. G., “Camera calibration for 3d computer vision,” *Proc. uInternational Workshop on Industrial Application of Machine Vision and Machine Intelligence, Silken, Japan*, pp. 240–247, 1987.
- [116] T. Melen, “Geometrical modeling and calibration of video cameras for underwater navigation,” *Dr. ing thesis, Norges tekniske hogaskole, Institutt for teknisk bybernetikk*.
- [117] F. Blais, “Review of 20 years of range senso development,” *J. Electron. Imag.*, vol. 13, no. 1, pp. 231–240, 2004.
- [118] S. Kullback, “Information theory and statistics,” *John Wiley and Sons, NY*, 1959.
- [119] J. Sun, N. N. Zheng, and H. Y. Shum, “Stereo matching using belief propagation,” *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. 26, no. 7, pp. 787–800, 2003.
- [120] Wei, G. Q., Ma, and S. D., “A complete two-plane camera calibration method and experimental comparisons,” *Proc. 4th International Conference on Computer Vision, Berlin, Germany*, pp. 439–446, 1993.
- [121] T. M. abd L. Van Gool, M. Proesmans, and E. Pauwels, “Affine reconstruction from perspective image pairs with a relative object-camera translation in be-

- tween,” *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. 18, no. 1, pp. 77–83, 1996.
- [122] M. Okutomi and T. Kanade, “A locally adaptive window for signal processing,” *International Journal of Computer Vision*, vol. 7, pp. 143–162, 1992.
- [123] I. Cox, S. Hingorani, and S. Rao, “A maximum likelihood stereo algorithm,” *Computer Vision and Image Understanding*, vol. 63, no. 3, May 1996.
- [124] O. Faugeras, “What can be seen in three dimensions from an uncalibrated stereo rig?” *Proceedings of the 2nd European Conference on Computer Vision, Santa Margherita Ligure, Italy*, pp. 563–578, 1992.
- [125] —, “Three-dimensional computer vision - a geometric viewpoint,” *Artificial intelligence. M.I.T. Press, Cambridge, MA*, 1993.
- [126] K. Liu, “Real-time 3-d reconstruction by means of structured light illumination,” *PhD Dissertation, Department of Electrical and Computer Engineering, University of Kentucky*, 2010.
- [127] O. Faugeras, B. Hotz, H. Mathieu, T. Viéville, Z. Zhang, P. Fua, E. Théron, L. Moll, G. Berry, J. Vuillemin, P. Bertin, and C. Proy, “Real time correlation based stereo: algorithm implementations and applications,” *Tech. Rep. 2013, INRIA*, 1993.
- [128] R. C. Gonzalez and R. E. Woods, “Digital image processing,” *Pearson Prentice Hall, ISBN 013168728X.*, 2007.

- [129] P. Leclercq and J. Morris, “Robustness to noise of stereo matching,” *Proceedings of the 12th International Conference on Image Analysis and Processing*, p. 606, 2003.
- [130] S. Lehmann, A. P. Bradley, I. V. L. Clarkson, J. Williams, and P. J. Kootsookos, “Correspondence-free determination of the affine fundamental matrix,” *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. 29, no. 1, pp. 82–97, 2007.
- [131] L. G. Hassebrook and C. Guan, “Distortion,” *Encyclopedia of Optical Engineering*, vol. DOI 10.1081/E-EOE 120009793.
- [132] C. Loop and Z. Zhang, “Computing rectifying homographies for stereo vision,” *IEEE CVPR*, vol. 1, p. 1125, 1999.

VITA

Yongchang Wang

- Birth Date: 17th Apr 1981
- Birth Place: China
- Home Page: <http://vis.uky.edu/~yongchang/>

Education

- **Master of Science**, “Performance Analysis of 3-Dimensional Fingerprint Scan System”, Dept. of Electrical and Computer Engineering, University of Kentucky, USA, 2008.
- **Bachelor of Engineering**, “Wavelet Application on Control System”, Dept. of Electrical Engineering, Zhejiang University, China, 2005.

Professional experience

- Research Assistant, University of Kentucky, USA. 2006-2009.
- Research Assistant, National University of Singapore, Singapore. 2005-2006.
- Research Assistant, Zhejiang University, China. 2003-2005.

Invention disclosures

1. **Yongchang Wang**, Kai Liu, D. L. Lau, and L.G. Hassebrook, “Maximum SNR Edge-pattern Strategy and Gamma model and error correction for Structured Light System,” INV09/1710. Approved for provisional patent application by University of Kentucky.
2. Kai Liu, **Yongchang Wang**, D. L. Lau, and L.G. Hassebrook, “Dual-frequency Phase Multiplexing (DFPM) and Period Coded Phase Measuring (PCPM) pattern strategies in 3-D structured light systems, and Lookup Table (LUT) based real-time data processing for phase measuring pattern strategies,” INV09/1686. Approved for provisional patent application.
3. **Yongchang Wang**, L. G. Hassebrook, D. L. Lau and Kai Liu, “Multi-camera Phase Measuring Profilometry for Accurate Registration and Distortion-free Fingerprint Unraveling to Aimed Definition,” INV09/1667, Sep 20, 2009. Approved for provisional patent application.
4. Laurence G. Hassebrook, C. J. Casey, E. Crane, W. Lundby, **Yongchang Wang**, and Kai Liu, “Roate and Hold and Scan Structured Light Illumination Pattern Encoding and Decoding,” INV09/1706. Approved for provisional patent application by University of Kentucky.

Publications

Journal papers

1. **Yongchang Wang**, Kai Liu, Qi Hao, Daniel L. Lau and Laurence G. Hassebrook. "Maximum SNR Pattern Strategy in Phase Shifting Methods" Submitted to Journal of the Optical Society of America, A., 2010.
2. Kai Liu, **Yongchang Wang**, Daniel L. Lau, Qi Hao and Laurence G. Hassebrook. "Gamma Model and its Analysis for Phase Measuring Profilometry," Journal of the Optical Society of America A, Vol. 27, pp. 553-562, Mar. 2010.
3. Kai Liu, **Yongchang Wang**, Daniel L. Lau, Qi Hao and Laurence G. Hassebrook. "Dual-frequency pattern scheme for high-speed 3-D shape measurement," Optics Express, Vol. 18, pp. 5229-5244, Mar. 2010.
4. **Yongchang Wang**, Daniel L. Lau and Laurence G. Hassebrook. "Fit-sphere unwrapping and performance analysis of 3D Fingerprints," Applied Optics, Vol. 49, Issue 4, pp. 592-600, Feb. 2010.
5. **Yongchang Wang**, Daniel L. Lau and Laurence G. Hassebrook. "Quality and Matching Performance Analysis of 3D Unraveled Fingerprints," Submitted to Optical Engineering, 2009.
6. **Yongchang Wang**, Kai Liu, Qi Hao, Daniel L. Lau and Laurence G. Hassebrook. "Period Coded Phase Shifting Strategy for 3-D Real-time Acquisition and Data Processing," Submitted to IEEE Transactions on Image Processing, 2009.

7. **Yongchang Wang**, Kai Liu, Qi Hao, Daniel L. Lau and Laurence G. Hassebrook. “Hybrid Structured Light Illumination and Stereo Vision 3-D Reconstruction,” Submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009.
8. **Yongchang Wang**, Laurence G. Hassebrook, and Daniel L. Lau. “Data Acquisition and Processing of 3D Fingerprints,” Submitted to IEEE Transactions on Information Forensics and Security, 2009.
9. **Yongchang Wang**, Wang Hao, Chen Liang, Zhu Shanan “Java-based inverted-pendulum Control System,” Control of China, 4, 04429, 2005.

Conference papers

1. **Yongchang Wang**, Qi Hao, Abhishika Fatehpuria, Daniel L. Lau, and Laurence G. Hassebrook “Data Acquisition and Quality Analysis of 3-Dimensional Fingerprints,” IEEE conference on Biometrics, Identity and Security, Tampa, Florida, Sep. 22-24, 2009.
2. **Yongchang Wang**, Kai Liu, Qi Hao, Daniel L. Lau, and Laurence G. Hassebrook “Multicamera Phase Measuring Profilometry For Accurate Depth Measurement,” Proceedings of SPIE, the International Society for Optical Engineering, Orlando, Florida, Vol. 6555, Apr. 9-15, 2007.

Others

1. **Yongchang Wang**, Laurence G. Hassebrook, and Daniel L. Lau “Noncontact, depth-detailed 3D fingerprinting,” SPIE Newsroom. DOI: 10.1117 / 2.1200912.002552, Dec. 2009.

Presentations

1. **Yongchang Wang**, Qi Hao, Abhishika Fatehpuria, Laurence G. Hassebrook, and Daniel L. Lau, “Data Acquisition and Quality Analysis of 3-Dimensional Fingerprints”, IEEE conference on Biometrics, Identity and Security, Tampa, Florida, Sep. 22, 2009.
2. **Yongchang Wang**, “Performance Analysis of 3-Dimensional Fingerprint Scan System”, Defence of Master Degree, University of Kentucky, May 14, 2008.
3. **Yongchang Wang**, Kai Liu, Qi Hao, Daniel L. Lau, and Laurence G. Hassebrook “Multicamera Phase Measuring Profilometry For Accurate Depth Measurement”, Proceedings of SPIE, the International Society for Optical Engineering, Orlando, Florida, Apr. 9-15, 2007.
4. Veeraganesh Yalla, Laurence G. Hassebrook, Daniel L. Lau, Pratibha Gupta, Charles Cassey and **Yongchang Wang**, “3D Facial Visualization”, University of Kentucky Engineering Career Fair, Oct 13-14, 2006.
5. **Yongchang Wang**, “Wavelet Application on Control System”, Defence of Bachelor Degree, Zhejiang University, China, Jun 14, 2005.

Media news

1. “Touchless, 3-D fingerprinting system promises speed, accuracy”, Smart Planet, Oct 2009.
2. “3-D Scanning Brings the Future of Fingerprinting”, Popular Science, Sep 2009.
3. “Touchless 3-D Fingerprinting”, Technology Review, Sep 2009.
4. “Realtime 3D system”, Prosilica Camera News, Sep 2009 (also published in the November issue of Inspect Magazine in both German and English).
5. “Machine Vision Technology Leaves Its Print on Surveillance”, Photonics Spectra, May 2008.
6. “Vision Systems Tackle Fingerprint Analysis”, Vision Systems Design, Dec 2007.