

---

# DEVELOPEMENT AND ACCURACY ANALYSIS OF CODED PHASE SHIFT 3D SCANNER

---

Mr. Pranav Kant Gaur  
ENGG01201101001

BHABHA ATOMIC RESEARCH CENTRE, MUMBAI

*A dissertation submitted to the Board of Studies in Engineering Sciences  
In partial fulfillment of requirements for the Degree of  
Master of Technology*

*of*

HOMI BHABHA NATIONAL INSTITUTE



January 2, 2013



# **Declaration**

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

Pranav Kant Gaur  
ENGG01201101001

# Acknowledgments

Author would like to acknowledge his Technical advisor Mr.D.M.Sarode for guidance throughout this work in pointing out the literature that can be of help in solving a particular problem.Guide Dr.P.K.Pal has provided time-to-time feedback on the desirable course that this work should take.Further author would like to thank Mr.S.K.Bose,Head,Computer graphics & Visualization section,Computer Division, for providing valuable guidance on technical problems as well as encouragement to follow the problem until it is solved.And last but not the least Computer Division administration and staff for providing freedom to the author to complete this work.

# Abstract

In this work, a 3D scanner based on a non-contact active technique is developed which can be used as a metrology equipment. Further, the factors which affect measurement accuracy of the system is analyzed by studying the sources of systematic errors and an attempt has been made to quantize them.

This work was initiated to address the growing need to acquire expertise in surface metrology required in medical applications and to explore alternatives to conventional contact based metrology devices such as CMM.

Specifically, *Coded phase shift technique* is used in this work, which being a combination of both phase-shift and binary coded technique provides high resolution 3D reconstruction as of phase shift technique and high robustness against noise as of binary coded technique.

Accuracy and repeatability of calibration algorithms used in this work are quantified. For assessing accuracy of stereo-correspondence an approach is proposed to quantify correspondence error and has been demonstrated with real experiments. Further, measurement accuracy and precision of 3D data acquired from the developed system has been studied and compared with that of popularly used sensor Microsoft Kinect. 3D data acquisition rate of kinect is considerably higher than the current version of developed system. Future versions of developed system are expected to reduce this gap.

Open-source development of this system can provide a low cost, portable and flexible alternative to the commercially available 3D scanners. Further it provides a flexible, research-oriented system whose basic parameters can be experimented with. This will provide us further insights into science of 3D metrology.

# Contents

<b>1</b>	<b>Introduction</b>	<b>10</b>
1.1	Motivations for development and accuracy analysis of 3D scanner . . . . .	11
1.2	Scope of this work . . . . .	12
1.3	Structure of thesis . . . . .	13
1.4	Principle of operation and graphical layout of the developed system . . . . .	14
<b>2</b>	<b>System calibration</b>	<b>17</b>
2.1	Camera model . . . . .	18
2.1.1	Graphical abstraction . . . . .	18
2.1.2	Mathematical abstraction . . . . .	19
2.2	Current approaches:A review . . . . .	24
2.3	Approach followed in this work . . . . .	28
2.4	Working of developed system calibration module . . . . .	30
2.4.1	Camera calibration . . . . .	31
2.4.2	Projector calibration . . . . .	34
2.4.3	Camera-projector relative extrinsic calibration . . . . .	37
2.5	Discussion . . . . .	41
2.6	Summary . . . . .	41
<b>3</b>	<b>Stereo correspondence</b>	<b>42</b>
3.1	Overview of correspondence problem . . . . .	42
3.2	State-of-art approaches . . . . .	42
3.3	Approach followed in this work . . . . .	52

3.4	Working of developed stereo-correspondence modules . . . . .	53
3.4.1	Pattern generation . . . . .	53
3.4.2	Pattern projection and capture . . . . .	55
3.4.3	Phase wrapping . . . . .	57
3.4.4	Pixel decoding and phase unwrapping . . . . .	59
3.4.5	Correspondence computation using absolute phase . . . . .	60
3.5	Discussion . . . . .	61
3.6	Summary . . . . .	62
<b>4</b>	<b>Triangulation</b>	<b>63</b>
4.1	Current techniques for triangulation . . . . .	63
4.2	Approach followed in this work . . . . .	66
4.3	Working of developed triangulation module . . . . .	67
4.4	Procedure . . . . .	69
4.5	Visualization of Point clouds . . . . .	69
4.6	Discussion . . . . .	70
4.7	Summary . . . . .	70
<b>5</b>	<b>Accuracy evaluation and analysis</b>	<b>72</b>
5.1	Evaluation of accuracy and precision of developed system . . . . .	73
5.1.1	Precision . . . . .	74
5.1.2	Measurement accuracy . . . . .	74
5.2	Comparative evaluation with <i>Microsoft Kinect</i> . . . . .	76
5.2.1	Kinect:Principle of operation . . . . .	77
5.2.2	Experiments . . . . .	79
5.2.3	Observations and conclusions . . . . .	80
5.3	System calibration . . . . .	82
5.3.1	Experiments and results . . . . .	82
5.3.2	Limitations of the used testing method . . . . .	87
5.4	Stereo correspondence . . . . .	87
5.4.1	Experiments and results . . . . .	88

5.4.2	Limitations of the used testing method . . . . .	91
5.5	Triangulation . . . . .	91
5.6	Summary . . . . .	92
<b>6</b>	<b>Conclusions and future works</b>	<b>93</b>
6.1	Conclusions and proposed future works . . . . .	93
6.2	Contributions of this work . . . . .	96

# List of Figures

1-1	Stereo correspondence and triangulation . . . . .	15
1-2	Architecture of the developed system . . . . .	16
2-1	Transformation of a real world 3D point from world coordinate system to computer-buffer coordinate system . . . . .	20
2-2	Camera model pipeline . . . . .	23
2-3	Radial alignment constraint used in Tsai's algorithm . . . . .	25
2-4	System setup . . . . .	31
2-5	Some views used for camera calibration . . . . .	32
2-6	Visualization of camera calibration results:View along X-axis . . . . .	33
2-7	Visualization of camera calibration results:View along Y-axis . . . . .	34
2-8	Some view used for projector calibration . . . . .	36
2-9	Visualization of projector calibration views along Y-axis . . . . .	36
2-10	Visualization of projector calibration views along X-axis . . . . .	37
2-11	Actual setup for extrinsic calibration . . . . .	40
2-12	Estimated system extrinsic geometry . . . . .	40
3-1	A typical structured light system . . . . .	44
3-2	Binary coded pattern:Design . . . . .	46
3-3	Binary coded pattern:Projection . . . . .	46
3-4	Gray coded patterns . . . . .	47
3-5	Combination of Phase-shift and gray code techniques . . . . .	49
3-6	De-bruijn sequence patterns . . . . .	50
3-7	M-array based pattern . . . . .	51

3-8	Direct coding method:Rainbow patterns . . . . .	52
3-9	Vertical phase shifted patterns . . . . .	54
3-10	Horizontal phase shifted patterns . . . . .	54
3-11	Vertical binary coded patterns . . . . .	55
3-12	Horizontal binary coded patterns . . . . .	55
3-13	Some captured vertical fringe and binary coded patterns . . . . .	56
3-14	Some captured horizontal fringe and binary coded patterns . . . . .	56
3-15	Computed wrapped phase . . . . .	58
3-16	Computed unwrapped phase . . . . .	60
3-17	Stereo correspondence between camera and projector: <i>green</i> spot in (b) corresponds to selected point(cursor) in (a) . . . . .	61
4-1	Some 3D reconstruction results . . . . .	71
5-1	Experiment setup . . . . .	75
5-2	Orientations and positions used for experiments . . . . .	76
5-3	Kinect 3D reconstruction process . . . . .	78
5-4	Distance versus precision along X,Y,Z axes . . . . .	80
5-5	Distance versus measurement accuracy with indication of uncertainty(vertical bars) . . . . .	81
5-6	Points A,B,C,D were used for measuring camera calibration accuracy	84
5-7	Corners A,B,C,D were used for assessing accuracy of projector calibra- tion parameters . . . . .	86
5-8	Effect of gamma on 3D reconstruction of a plane . . . . .	88
5-9	Effect of increasing number of phase shifted patterns on <i>waviness</i> in 3D reconstruction . . . . .	89

# List of Tables

2.1	Estimated Camera intrinsic and distortion model parameters . . . . .	34
2.2	Estimated Projector intrinsic model parameters . . . . .	36
2.3	Estimated camera extrinsic parameters . . . . .	39
2.4	Estimated projector extrinsic parameters . . . . .	39
2.5	Estimated projector-camera relative parameters . . . . .	39
5.1	CPSS:Precision along X,Y,Z axes . . . . .	74
5.2	CPSS:Dependence of measurement accuracy on distance . . . . .	77
5.3	CPSS:Dependence of measurement uncertainty on distance . . . . .	77
5.4	Kinect:Precision along X,Y,Z axes . . . . .	79
5.5	Kinect:Dependence of measurement accuracy on distance . . . . .	79
5.6	Kinect:Dependence of measurement uncertainty on distance . . . . .	81
5.7	Percentage average absolute deviation of camera calibration parameters	83
5.8	Camera calibration:Radial error of projection . . . . .	84
5.9	Percentage average absolute deviation of projector calibration parameters	85
5.10	Projector calibration:Radial error of projection . . . . .	86
5.11	Stereo-correspondence error for various gamma values for 3 phase-shifted method . . . . .	90
5.12	Stereo correspondence error for 3,4,5 phase shifted patterns . . . . .	90

# Chapter 1

## Introduction

A 3D scanner is a combination of hardware and software components used for acquisition of three dimensional information from a scene. Recent applications[1] of 3D scanners seems to suggest that they are not only used for metric reconstructions but also for acquiring information regarding relative placement of objects in 3D space which do not essentially requires metric information, examples are *affine* and *projective* reconstructions[2],[3], example application includes robotic navigation where 3D coordinates in terms of physical units are not required but only a non-metric 3D information suffices[4], whereas the *euclidean* reconstruction comes under category of metric reconstruction and is also the subject of this work.

3D scanners find applications in robotic controls for example a laser scanner can act like an eye for robot, in site modeling and lay-outing, establishing benchmark for preexisting shape/state for studying structural changes in response to extreme environmental conditions, to create Geometric information systems, it is also used for sub-surface scanning in mines. Of course in entertainment where currently Microsoft Kinect is a leading example. It is also used in reverse-engineering a mechanical component which requires a precise digital model of object. In cultural heritage preservation, where Michelangelo, Monticello, Kasubi-tombs etc are prime examples. It is also used in Medical CAD/CAM for example to produce orthosis, prosthesis and dental implants. In industrial quality inspection and meteorology it is used for measuring

geometric dimension accuracy, automation of quality assurance process for example in assembling a car it must be ensured that geometry of individual components is accurate so that it can fit together for reliable operation.

Although recent work on Monocular 3D[5] has lead to development of methods which can extract approximate three dimensional information even from single view using cues like perspective distortion, boundary shades, shadow, occlusion etc, exact metric reconstruction still requires at least 2 views. In simplest case with 2 views, there is a need to recover a relationship between these views so as to compute depth of a point common to both views with respect to a global coordinate system. But since depth is to be interpreted and utilized for real world practical applications(as is the case with this work) it has to be expressed in terms of physical units like ‘mm’, ‘cm’, ‘feet’ etc.

So clearly above discussion point towards three major problems associated with 3D metric reconstruction:Correspondence between views, Calibration of viewing elements to determine the relation between abstract world of these elements(in our case these elements are projector and camera) and our physical world, and computation of depth given correspondence and calibration information. Specifically for 3D reconstruction, first system is calibrated, then correspondence between camera and projector is calculated, and then using the calibration and correspondence information triangulation is performed to compute 3D coordinates of point in real world with respect to a reference coordinate system.

## 1.1 Motivations for development and accuracy analysis of 3D scanner

Motivation for this work comes from the fact that as per our current knowledge there is no development expertise in the field of 3D scanners based on structured-light

technique in our country. This approach provides high accuracy, portability, low-cost solution to dimension measurement and reverse-engineering applications. Although measurement accuracy of these systems is still lower than CMM machines, it can be expected to be achievable with high quality camera and projector optics and improvements in measurement algorithm. Furthermore current commercial scanners using structured-light technique are considerably more costlier for example[6] where even cost of 3D scanner software is  $>1.5$  Lakhs which excludes cost of hardware components. Hence our development work can provide a low-cost alternative to these systems.

Specifically, in CDM,BARC we have a CMM machine [Zeiss UPMC-850] for dimensional inspection of fabricated job which is reported to have accuracy of  $4\mu m$  within measurement volume of 850mm X 1200mm X 600 mm. But dimension measurement time for complex jobs, strict temperature requirements to maintain acceptable sphere-form of Ruby balls used in probes, lack of portability of measuring device, higher power consumption for maintaining near frictionless air-bearings and a certain operating temperature and larger space requirements are motivating factors for this work. In addition, in B.A.R.C Hospital we have a common requirement of producing artificial teeth for implantation but an expertise is needed to deal with its technical details. Furthermore Computer Division is working on CFD simulation problem in collaboration with IIT-Guwahati where a digitized model of a physical object will be more insightful to use instead of using an artificial object with ideal/perfect dimensions, further development of our work can provide 360 degree 3D scans of an object which can be used in such simulations.

## 1.2 Scope of this work

Scope of the work is limited to development of a 3D dimension measurement system based on Coded-phase shift technique. Further, tests are devised to assess accuracy of individual components of the system developed. Specifically, the effect of

non-linearity of output response of projector(characterized by *gamma value*) on accuracy of stereo-correspondence is studied. In addition, the effect of increasing the number of phase-shifted patterns on accuracy of stereo-correspondence has also been studied as it has been both theoretically and practically demonstrated in the literature[7] that increasing the number of phase shifted patterns reduces the effect of non-linearity of projector output response and hence increases accuracy of estimated stereo-correspondence. Here i have tested these claims for 3, 4 and 5 phase-shifted sinusoidal fringe patterns. Similarly, tests to assess accuracy of system calibration parameters are performed. To evaluate accuracy of complete 3D scanning pipeline, a comparative evaluation of measurement accuracy and precision of developed system with nowadays commonly available 3D sensor *Microsoft Kinect* is performed. However during development and performance assessments of developed system, shortcomings of techniques used for both development and accuracy assessments were observed which are reported in chapter-6 of this thesis which may pave the way for our future work.

### 1.3 Structure of thesis

Chapter-2 describes the currently most popularly used approaches for solving system calibration problem. It further describes the development and working of system calibration module. Chapter-3 reviews the current state of approaches to solve stereo-correspondence problem and further describes the development and working of Coded phase-shift scanning technique. Chapter-4 first reviews the current approaches to solve the triangulation problem and then describes the approach used in this work. Chapter-5 explains the experiments performed to assess accuracy of modules within the scope of this work namely stereo-correspondence and system calibration modules. It further describes the work done for comparative evaluation of measurement accuracy and precision of Coded-phase-shift scanner and Microsoft Kinect. In chapter-6 conclusions emerging from this work are described and directions for future work are proposed. Finally, Bibliography concludes the thesis.

## 1.4 Principle of operation and graphical layout of the developed system

Solving stereo-correspondence problem amounts to providing unique relation between points of camera and projector such that related camera and projector points will be 2D projections of same point in real 3D world. Coded phase-shift technique being a structured light technique solves this problem by assigning unique code to each point(theoretically) in the 3D scene. Specifically, this technique involves projection of sinusoidal phase-shifted patterns(which assigns unique phase value to each projector pixel within a sinusoidal period) followed by binary-coded patterns (which assign unique period number to each sinusoidal fringe period) onto object to assign unique phase-value to each point in the scene. Camera captures these patterns and computes the phase of incident signal at each pixel(which indirectly assigns phase to the pixel). Correspondence problem is then solved by determining the camera-projector pixel pair having same phase value.

Sinusoidal and binary-coded patterns are generated by ***Pattern generator module***. Camera captures the projected patterns through ***Pattern projection and capture module*** and for each pixel, phase of the original incident signal is computed through ***Phase wrapper and Phase unwrapper modules***. This phase value at each pixel relates a camera pixel to the corresponding projector pixel which are seeing a common point in the real 3D world through ***Absolute phase computation module***. Further the calibration parameters for camera, projector and relative orientation and position of camera with respect to projector are estimated by ***System calibration module***. Once the system calibration parameters and stereo-correspondence information is known, triangulation can be applied through ***Triangulation module***. Figure 1-1 shows the optical triangulation process.

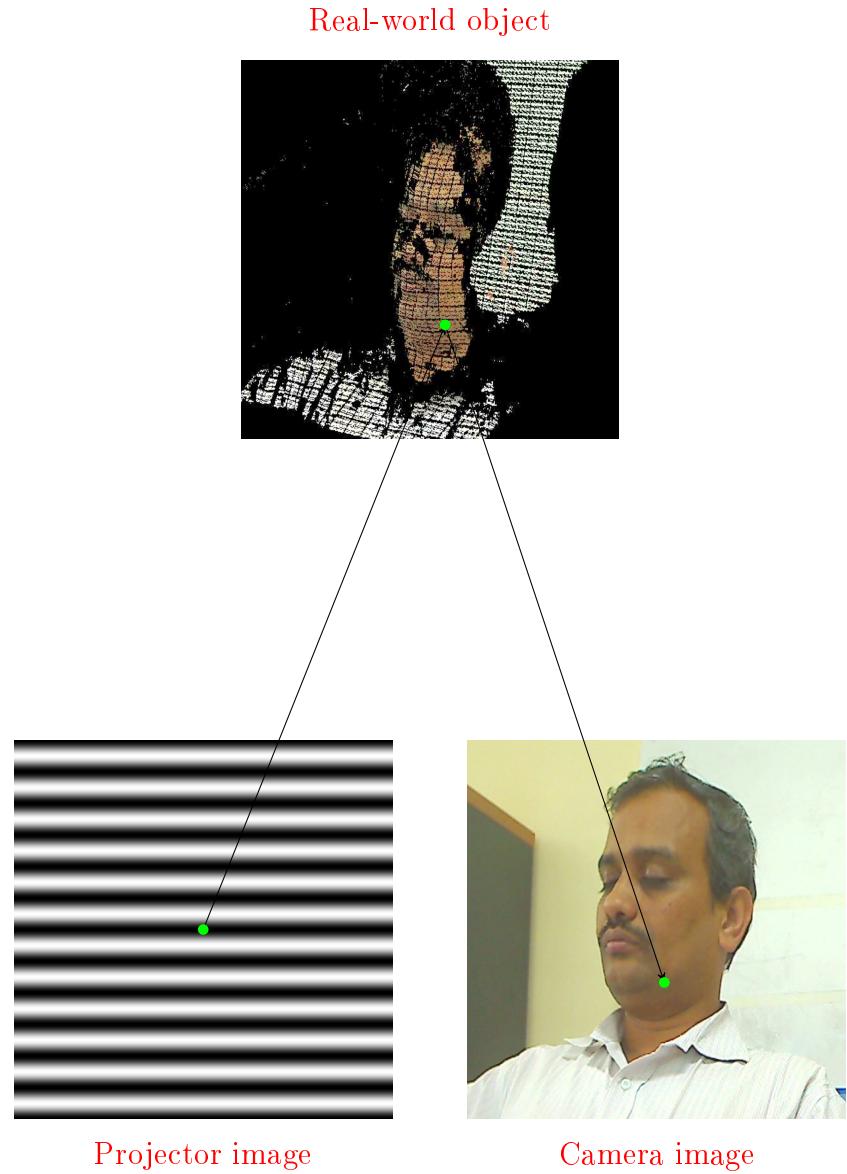


Figure 1-1: Stereo correspondence and triangulation

In figure 1-1, a point in *Projector image* is related to a point in *Camera image* because they are 2D projections of a common point in the scene using the estimated stereo-correspondence. Once system is calibrated, ray-ray intersection between optical ray of camera and that of corresponding projector pixel gives the (X,Y,Z) coordinates of real-world point in *physical unit* like ‘mm’, ‘m’ etc.

Figure 1-2 graphically represents the layout of our 3D scanner system, which can

be referred during thesis. Working of ***System calibration module*** is documented in chapter-2. ***Pattern generation, Pattern projection and capture, Phase wrapper, Phase unwrapper, Absolute phase computation modules*** are explained in chapter-3. ***Triangulation module*** has been described in chapter-4.

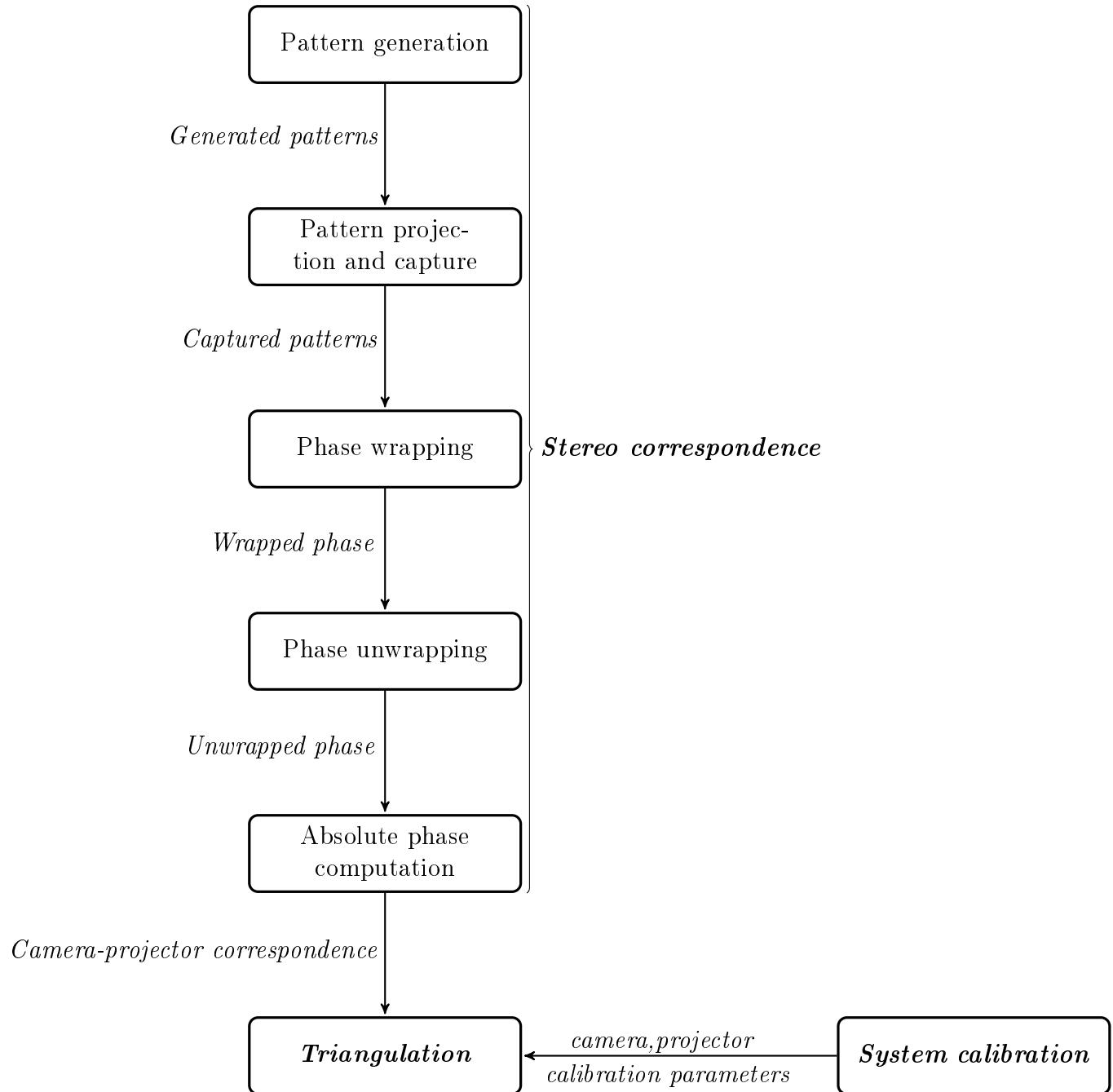


Figure 1-2: Architecture of the developed system

# Chapter 2

## System calibration

For performing non-metric reconstruction only point-to-point correspondences between camera and projector are needed to perform triangulation[3]. But to relate the depth computed using triangulation with physical metrics such as 'mm', 'cm' etc there is need to relate the camera and projector geometry with physical measurement units, this task is accomplished by process of System calibration. Calibration is a process of determining intrinsic and extrinsic characteristics of the imaging devices(camera and projector in our case). Intrinsic characteristics include focal length, coordinates of principal points of lens used in camera and projectors, image axis skew coefficients and lens distortion coefficients. Extrinsic characteristics include rigid-body transformations that can transform a point represented in World-coordinate system to individual coordinates systems of camera and projector.

Above discussion clearly implies the need to establish defining equations which relate a real-world three dimensional point(with coordinates expressed in *physical* units) with pixel coordinates in camera(or projector). For this purpose a projective model of camera which applies to projector as well is introduced in the next section which clearly describes the problem of calibration, following which the current approaches and the approach used in this work to solve this problem will be reviewed. In the end, a note on *relative extrinsic calibration* of projector-camera system is mentioned, which is a process of determining relative rotation and translation of camera and

projector coordinate systems which in simple words means to determine the relative orientation and position of camera with respect to projector or vice-versa.

## 2.1 Camera model

Camera model is a concise mathematical abstraction of the internal(i.e., its optics and sensor arrangement) and external(i.e., its position relative to a scene) geometry of camera which is responsible for the imaging process. In this section, these geometric parameters will be described and hence the imaging process. Camera model described here is referred from seminal paper by [8]. Other calibration techniques also follow variants of this formulation. Relatively recent work in [9] has shown projector to be *inverse-camera* hence further discussion on camera calibration holds equally to projector as well. Section 2.4.2 will describe how projector was calibrated similar to camera with a real example. In next subsection the coordinate-systems which are relevant for geometric modeling of imaging process under consideration will be described.

### 2.1.1 Graphical abstraction

In this subsection the graphical model of camera depicting coordinate systems involved in transforming a real world 3D point to camera 2D coordinate representation will be described.

**World-coordinate system.** World-coordinate system defines the reference frame in the real world that will be used for defining the positions of object in scene in physical units. Typically any location in the real scene can be used as origin for world coordinate system axes. For example in this work world coordinate system origin was assumed to be present at a corner on a wall.

**Camera-coordinate system.** This is a 3D reference coordinate-system internal to camera, it is similar to world coordinate system since it also works in physical units.

Its origin is assumed to be at center of projection of camera. It is used to generate 3D interpretation of the scene from camera viewpoint.

**Image-coordinate system.** This is a 2D reference coordinate-system for indexing points in the real-image. Note here *real* means that it can take real values as coordinates. It is used to index the points in perspective projection of 3D world.

**Computer-frame coordinate system.** This is a 2D reference coordinate-system which indexes pixels in the computer-frame. It can only have integral coordinates. Images with which we normally interact are indexed with respect to this coordinate system.

Figure 2-1<sup>1</sup> illustrates the coordinate systems except computer-frame coordinate system since it is not geometrically connected with other coordinate systems:

Here points  $O_w, O_c$  and  $O_i$  represent origins of world,camera and real-image coordinate systems respectively. Further, axes  $X_w, Y_w$  and  $Z_w$  represent the world-coordinate system, axes  $X_c, Y_c$  and  $Z_c$  denote the camera coordinate system, similarly axes  $X_i$  and  $Y_i$  show the 2D real-image coordinate system,  $P$  is a real-world point,  $P_u$  is the ideal projection of  $P$  on real-image plane(i.e., image-coordinate system),  $P_d$  is corresponding distorted projection which differ from  $P_u$  due to practical imperfections in camera system. Equations relating these coordinate systems are described in the following paragraphs of this section.

### 2.1.2 Mathematical abstraction

In this subsection, the actual mathematical form of camera imaging process which is used for transforming a real-world 3D point to a computer-buffer pixel will be sequentially described. This involves a series of coordinate transformations which are described here.

---

<sup>1</sup>Figure source: Refer [8]

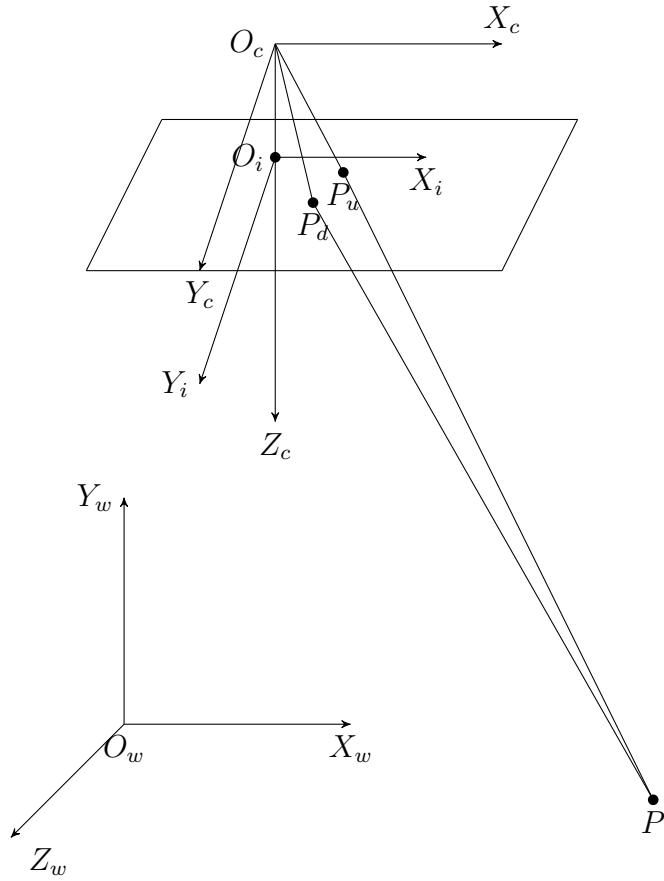


Figure 2-1: Transformation of a real world 3D point from world coordinate system to computer-buffer coordinate system

**1.Rigid body transformations from object world-coordinate system to camera coordinate system.** Points in the world-coordinate coordinates system ( $x_w, y_w, z_w$ ) are transformed to the camera-coordinate system ( $x_c, y_c, z_c$ ). To model this process, rotation( $R$ ) and translation( $T$ ) matrices are used as:

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = R * \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + T \quad (2.1)$$

$$\text{where } R = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} \\ r_{2,1} & r_{2,2} & r_{2,3} \\ r_{3,1} & r_{3,2} & r_{3,3} \end{bmatrix}, T = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

Rotation and translation parameters are called *extrinsic parameters* of camera/projector. It should be noted here that rotation followed by translation is used since it ensures unique transformations, whereas translation followed by rotation can result in multiple possible transformations[8].

**2.Perspective transformation from camera-coordinates to real-image coordinates.** In this step, 3D scene as viewed from camera-coordinate origin is perspective projected to a plane. Hence it is a 3D to 2D transformation resulting in loss of Z-coordinate(i.e., depth) since multiple points are projected onto same point in the image-plane. Following equations describe this process:

$$\begin{bmatrix} x_u \\ y_u \end{bmatrix} = \left( \frac{f}{z_c} \right) * \begin{bmatrix} x_c \\ y_c \end{bmatrix} \quad (2.2)$$

where  $f$  represents the focal length of camera/projector lens.  $(x_u, y_u)$  is the perspective projection of point  $(x_c, y_c, z_c)$  on real-image plane.

**3.Radial and tangential lens distortion.** In this step, real-image coordinates are distorted due to radially imperfections in practical lens which is characterized by *radial distortion* and oblique alignment of principal plane of lens and sensor-plane which is characterized by *tangential distortion*. Although Tsai's work only accounted for radial distortions subsequent works have accounted for tangential[10] and prism distortions[11] also in an attempt to increase the accuracy of estimated calibration parameters further. Following equations describes the effect of radial and tangential distortion on *ideal* image coordinates:

$$\begin{aligned} x_d &= x_u(1 + k_1r^2 + k_2r^4 + k_3r^6 + ...) + 2p_1x_u y_u + p_2(r^2 + 2x_u) \\ y_d &= y_u(1 + k_1r^2 + k_2r^4 + k_3r^6 + ...) + 2p_2x_u y_u + p_1(r^2 + 2y_u) \end{aligned} \quad (2.3)$$

where,

$(x_d, y_d)$ : distorted image coordinates

$(k_1, k_2, k_3)$ : radial distortion coefficients

$(p_1, p_2)$ : tangential distortion coefficients

$$r = \sqrt[2]{X_u^2 + Y_u^2}$$

**4.Transforming real-image coordinates to computer-frame coordinates.** Finally to map 3D world-coordinates to frame coordinates of camera, i.e., the pixels that we see, following equations are assumed:

$$\begin{bmatrix} x_f \\ y_f \end{bmatrix} = \begin{bmatrix} s_x * (d'_x)^{-1} * x_d \\ d_y^{-1} * y_d \end{bmatrix} + \begin{bmatrix} C_x \\ C_y \end{bmatrix} \quad (2.4)$$

where,

$(x_f, y_f)$ :computer frame-buffer coordinates for point  $(x_d, y_d)$

$s_x$ :uncertainty image-scale factor

$d_x$ :center-to-center distance between sensor cells along X-direction

$N_{cx}$ :Number of sensors along X-direction

$N_{fx}$ :Number of pixels along X-line sampled by computer

$$d'_x = d_x * \left( \frac{N_{cx}}{N_{fx}} \right)$$

It should be noted that at full camera resolution there is one-to-one relation between number of sensor elements and number of pixels in computer-frame buffer hence  $N_{cx} = N_{fx}$  and hence  $d'_x = d_x$  but in general  $d'_x >= d_x$ . [8] attributed  $s_x$  to be due to possibility of mismatch between camera image acquisition hardware and camera scanning hardware. Figure 2-2<sup>2</sup> illustrates the complete imaging process.

So based on above discussion we can deduce that to define the geometric imaging process of camera(or projector) we have to estimate following parameters:

1. Internal parameters:

- (a) Camera lens focal length
- (b) Position of principal point on the real-image plane
- (c) Lens distortion parameters

---

<sup>2</sup>Figure source: Refer [8]

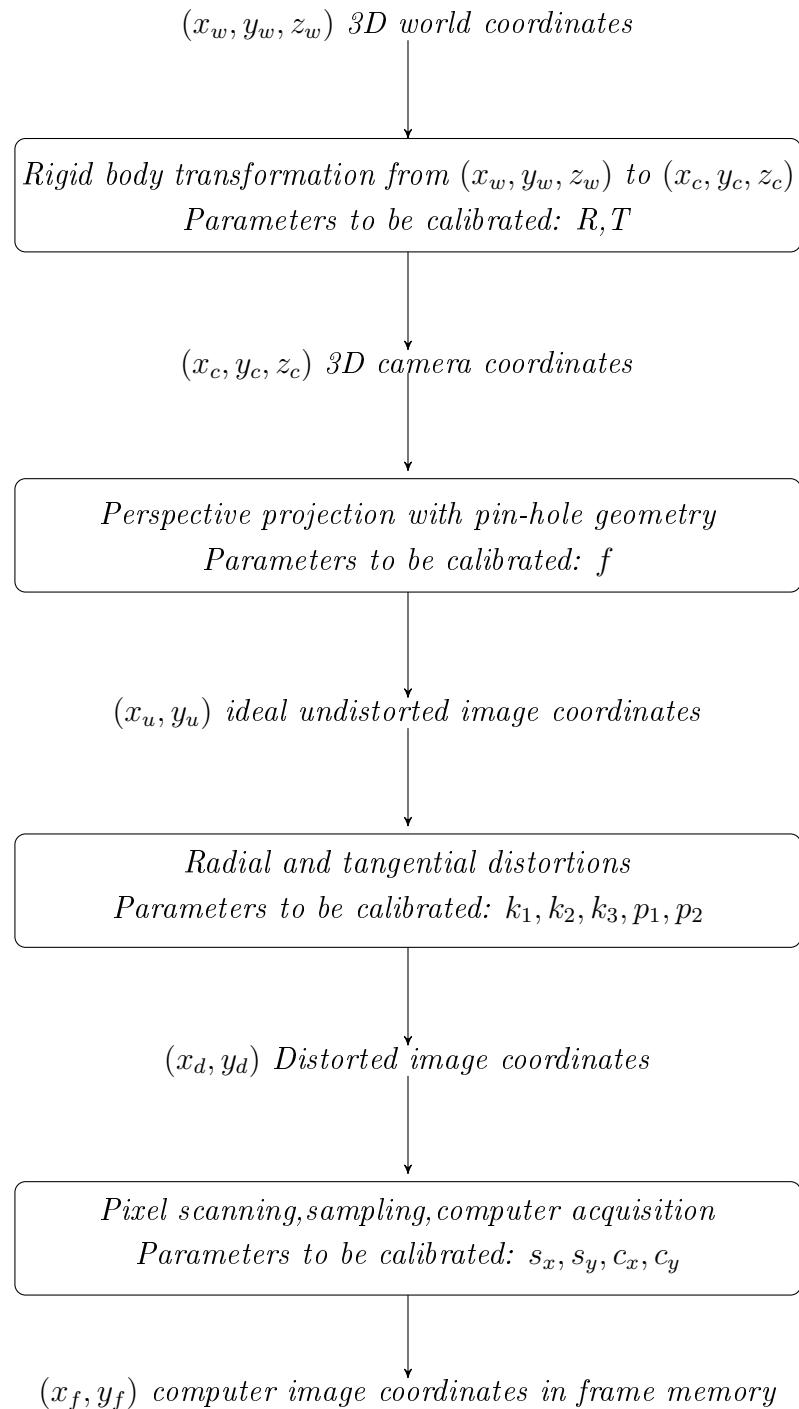


Figure 2-2: Camera model pipeline

2. External parameters:
  - (a) Rotation between camera-coordinate system and world-coordinate system
  - (b) Translation between camera-coordinate system and world-coordinate system

## 2.2 Current approaches:A review

A taxonomy was proposed by [10] for camera calibration approaches based on their requirement of additional calibration objects as:

- a) **Photogrammetric calibration.** These techniques perform camera calibration by observing a calibration object. Calibration object can be any object with known geometry i.e., we know world-coordinates of feature-points on the object that we want to use for calibration. In this category many techniques for both 3D and 2D calibration objects were proposed. For example use of 3D calibration object with three orthogonal planes or planar calibration rigs with known translation across its views have been used for calibration[8] although requiring elaborate setup for calibration and often time-consuming.
- b) **Self-calibration.** These techniques are more flexible than photogrammetric techniques in that they do not require an explicit calibration rig for calibration, but by moving the camera across a static rigid scene they determine constraints on internal parameters of camera that can be used for calibration[12][13]. Other than these techniques there are techniques that do not fall into these categories such as camera calibration using vanishing points for orthogonal directions[14], calibration from pure rotation[15]. In this work, a photogrammetric method for camera and projector calibration was used. Following paragraphs will briefly describe the details of some highly cited and practically used algorithms in chronological order.

**R.Y.Tsai's algorithm.** Tsai's algorithm restricts the lens-distortion effects to only radial distortion and do not account for skewness or lack of orthogonality of projection. Calibration was divided in 2 separate stages. In first stage all extrinsic parameters except for  $t_z$  are computed using the *Radial alignment constraint(RAC)* shown in figure 2-3<sup>3</sup>. If assumed lens distortion is only *radial*, then the direction of vector  $O_iP_d$  remains unchanged and is parallel to vector  $P_{oz}P_w$ . Tsai showed that RAC is independent of magnitude of effective focal length, radial distortion coefficients and position of point  $P_w$  along z-axis. He further proved that this constraint is sufficient to estimate 3D rotation and translation of  $P_w$  along X and Y axes.

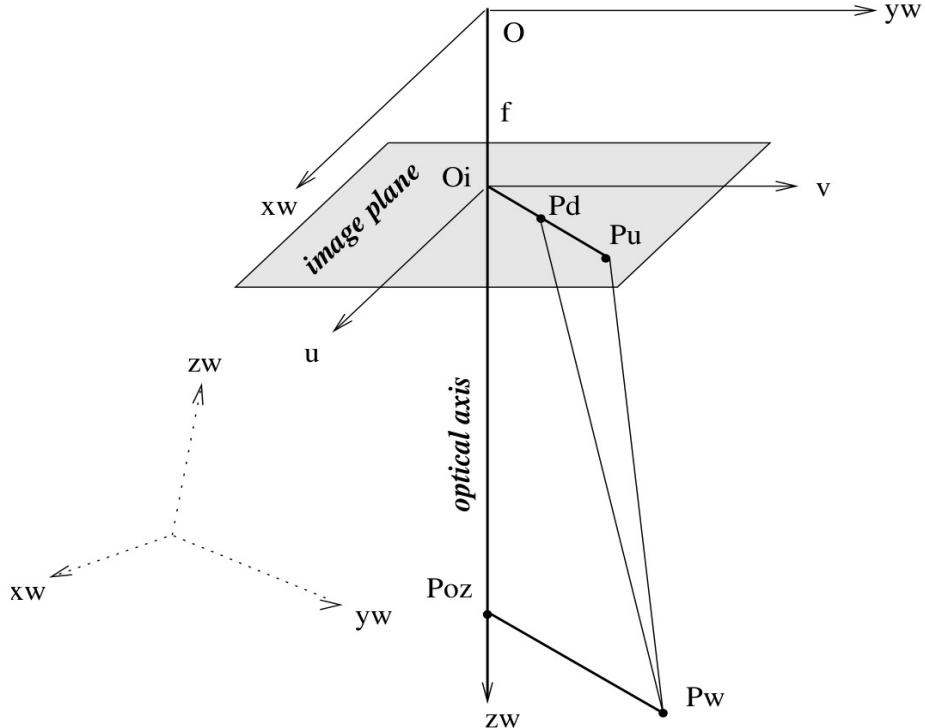


Figure 2-3: Radial alignment constraint used in Tsai's algorithm

In second stage effective focal length, distortion coefficients and  $t_z$  are estimated using non-linear optimization.

---

<sup>3</sup>Figure source: Refer [16]

**Z.Zhang's algorithm.** Z.Zhang proposed a plane based camera calibration algorithm[10]. It first acquires real-world coordinate to camera image coordinate correspondences using planar calibration board. Then it estimates all intrinsic and extrinsic parameters using a closed-form solution. This closed form solution assumes zero lens distortion and is based on two constraints on intrinsic parameters derived from the *orthonormality* of rotation vector expressed in equations 2.5,2.6.

$$h_1^T A^{-T} A^{-1} h_2 = 0 \quad (2.5)$$

$$h_1^T A^{-T} A^{-1} h_1 = h_2^T A^{-T} A^{-1} h_2 \quad (2.6)$$

where,

$h_i$ :  $i^{th}$  column vector of Homography

$A$ : Camera projective matrix

Further, the coefficients of radial distortion  $k_1, k_2$  are estimated by solving linear least squares expressed in equation 2.7.

$$\begin{bmatrix} (u - u_0)(x^2 + y^2) & (u - u_0)(x^2 + y^2)^2 \\ (v - v_0)(x^2 + y^2) & (v - v_0)(x^2 + y^2)^2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} u' - u \\ v' - v \end{bmatrix} \quad (2.7)$$

where,

$(u, v)$ : ideal pixel image coordinates

$(u_0, v_0)$ : coordinates of principal point

$(x, y)$ : ideal normalized image coordinates

Estimated values of all parameters are considered as initial input for non-linear least squares minimization which minimizes expression in equation 2.8.

$$\sum_{i=1}^m \sum_{j=1}^n \|m_{ij} - m'(A, k_1, k_2, R_i, t_i, M_j)\|^2 \quad (2.8)$$

where,

$m$ : Total number of features on the calibration board

$n$ : Total number views of planar calibration board

$M_j$ : 3D coordinates of  $j^{th}$  feature point on calibration board

$m_{ij}$ : True 2D coordinates of projection of  $M_j$  in camera image

$m'$ : Estimated 2D coordinates for  $M_j$  using calibration parameters

$R_i$ : Rotation parameters for  $i^{th}$  view used for calibration

$t_i$ : Translation parameters for  $i^{th}$  view used for calibration

**J.Heikkila et al.'s algorithm.** This algorithm assumes a more elaborate camera model which includes both radial and tangential components as lens distortion coefficients. It first linearizes the camera model by ignoring lens distortion coefficients and performs linear parameter estimation using Direct Linear Transformation(DLT)[17]. Equation 2-9 shows the linear model solved using DLT method.

$$La = 0 \quad (2.9)$$

where,

$$a = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \end{bmatrix}$$

$$L = \begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -X_1u_1 & -Y_1u_1 & -Z_1u_1 & -u_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -X_1v_1 & -Y_1v_1 & -Z_1v_1 & -v_1 \\ \vdots & \vdots \\ X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -X_iu_i & -Y_iu_i & -Z_iu_i & -u_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -X_iv_i & -Y_iv_i & -Z_iv_i & -v_i \\ \vdots & \vdots \\ X_n & Y_n & Z_n & 1 & 0 & 0 & 0 & 0 & -X_nu_n & -Y_nu_n & -Z_nu_n & -u_n \\ 0 & 0 & 0 & 0 & X_n & Y_n & Z_n & 1 & -X_nv_n & -Y_nv_n & -Z_nv_n & -v_n \end{bmatrix}$$

Here,  $(X_i, Y_i, Z_i)$  denotes 3D coordinates of  $i^{th}$  feature point used in calibration,  $(u_i, v_i)$  denotes the corresponding 2D image coordinates,  $a$  denotes the projective matrix mapping  $(X_i, Y_i, Z_i)$  to  $(u_i, v_i)$ . In second step, Gaussian noise model is assumed and a non-linear least square minimization is performed to minimize equation 2-10.

$$F = \sum_{i=1}^N [(U_i - u_i)^2 + (V_i - v_i)^2] \quad (2.10)$$

where,  $N$  is the number of observations,  $(U_i, V_i)$  are the observed 2D coordinates of  $i^{th}$  feature point,  $(u_i, v_i)$  are corresponding 2D coordinates estimated using calibration parameters.

## 2.3 Approach followed in this work

For camera and projector calibration an extension of [10] method was used which was first described and developed in [18] and later ported to OpenCV. Unfortunately, author in [18] has not published the details of algorithm except describing the approach to be based on camera model described in [17] and implementation as a variant of [10] hence OpenCV documentation other than reading the source code itself is the only source of theoretical information on this technique. During this project work, an effort is initiated to fully understand the calibration algorithms and effect of each calibration parameters on accuracy of 3D reconstruction by actually implementing an algorithm so that every input to the algorithm can be changed and its effect seen. For which technique described in [17] was selected and partially implemented. Here a brief outline of camera/projector calibration algorithm implemented in OpenCV/Matlab[18][19] is presented:

1. A planar calibration board is selected with known world-coordinates of reference points on it.
2. Calibration board is shown for multiple views(at least 2) in front of camera by

varying its rotation and translation(amount need not be known) with respect to camera and for every pose feature-points are detected in camera.

3. After all such views are captured and feature-points detected, homography for each view is calculated.
4. Closed form solution of intrinsic parameters from homographies using orthogonality of vanishing points is computed but unlike [10] no initial estimates of distortion coefficients are computed.
5. Initial estimates of calibration parameters computed earlier are given as input to Maximum likelihood estimator Levenberg-Marquardt algorithm[20] to refine the parameters. Output of this process is a set of calibration parameters which optimally satisfy the maximum likelihood criteria.

**Algorithm for extrinsic calibration of camera and projector.** Extrinsic parameters relate camera-coordinate system and projector-coordinate system to world-coordinate system. For extrinsic calibration, OpenCV[21] performs minimization of re-projection error between observed point projections given as input(i.e., image points) and projection computed by applying the estimated model to the given object points. Extrinsic parameters that minimize this criteria are considered as the optimal pose(Rotation and translation) parameters.

**A note on Relative calibration of projector and camera.** After camera and projector are calibrated extrinsically, there is need to relate their coordinate-systems with each other so that a point in projector coordinate-system can also be expressed in terms of camera-coordinate system. This is required because of an implicit assumption in triangulation is that it requires intersection of optical rays from camera and projector, which is possible only if we can express both rays in a common coordinate-system either in that of camera or projector. This is achieved using our world-coordinate system as a common link between projector and camera. Following

equations relate camera-coordinate system to projector-coordinate system via world-coordinate system[22]:

$$\begin{bmatrix} P_c \\ P_p \end{bmatrix} = \begin{bmatrix} R_{wc} * P_w \\ R_{wp} * P_w \end{bmatrix} + \begin{bmatrix} T_{wc} \\ T_{wp} \end{bmatrix} \quad (2.11)$$

hence,

$$P_c = \underbrace{(R_c * R_p^{-1}) * P_p}_{R_{pc}} + \underbrace{(T_c - R_c * R_p^{-1} * P_p)}_{T_{pc}}$$

where,

$P_w$ :A point in world coordinate system

$P_c$ :Representation of  $P_w$  with respect to camera coordinate system

$P_p$ :Representation of  $P_w$  with respect to projector coordinate system

$R_{wc}$ :Rotation transformation from world-to-camera coordinate system

$R_{wp}$ :Rotation transformation from world-to-projector coordinate system

$T_{wc}$ :Translation transformation from world-to-camera coordinate system

$T_{wp}$ :Translation transformation from world-to-projector coordinate system

$R_{pc}$ :Rotation transformation from projector-to-camera coordinate system

$T_{pc}$ :Translation transformation from projector-to-camera coordinate system

Here  $R_{pc}$  and  $T_{pc}$  are the final outcomes which allows us to map a point in projector coordinate system to camera coordinate system which is an requirement for performing stereo-triangulation.

## 2.4 Working of developed system calibration module

System calibration is composed of camera calibration, projector calibration and relative extrinsic calibration of camera with respect to projector. A module for system calibration was developed which performs above mentioned functions using OpenCV camera calibration and extrinsic calibration algorithms. This section first describes the system setup and then explains the practical procedure for camera, projector and extrinsic calibration of projector with respect to camera.

**System setup.** In this work, Logitech Quickcam sphere AF web-cam at 1600X1200 pixels resolution was used as a capture device, Sharp PG-F200X projector at 1024X768 pixels resolution was used for pattern projection, and 4 markers (shown in figure 2-4 as ‘black’ points) as world coordinate references forming a rectangular region of 1900mm X 1700mm and a physical checkerboard with square size 25.6mm arranged in a 10X8 square grid. Figure 2-4 shows setup of our 3D scanner system:



Figure 2-4: System setup

### 2.4.1 Camera calibration

In this work, camera calibration was performed using OpenCV camera calibration function whose theoretical foundations were laid by the well known algorithms by [10] and [17]. Basic overview of algorithm was presented in Section-2.3. OpenCV estimates 9 intrinsic parameters to define internal geometry of camera and 6 extrinsic parameters to define external geometry of camera with respect to world coordinate system.

#### **Intrinsic and distortion parameters**

$f_x$ : focal length of camera lens expressed in number of pixels along X - axis

$f_y$ : focal length of camera lens expressed in number of pixels along Y - axis

$(c_x, c_y)$ : Pixel coordinates of principal point

$(k_1, k_2, k_3)$ : Radial distortion coefficients

$(p_1, p_2)$ : Tangential distortion coefficients

### Extrinsic parameters

$(r_x, r_y, r_z)$ : Rotation vector between camera and world coordinate system

$(t_x, t_y, t_z)$ : Translation vector between camera and world coordinate system

**Procedure:** Aim of this process is to acquire 3D-2D correspondences between world coordinate system and camera image coordinate system which are then used to estimate intrinsic and extrinsic parameters of camera. World coordinate system origin is assumed to be at top-left corner of the physical checkerboard although it can be at any arbitrary position. Please note that this coordinate system origin is different from the world coordinate origin used for measurement(i.e., a corner on wall) described earlier.

Multiple views of the checkerboard are captured by camera and inner corners are detected with sub-pixel accuracy. Figure 2-5 shows some such views used.

Each such view provides a set of 3D-2D correspondence and adds a separate set

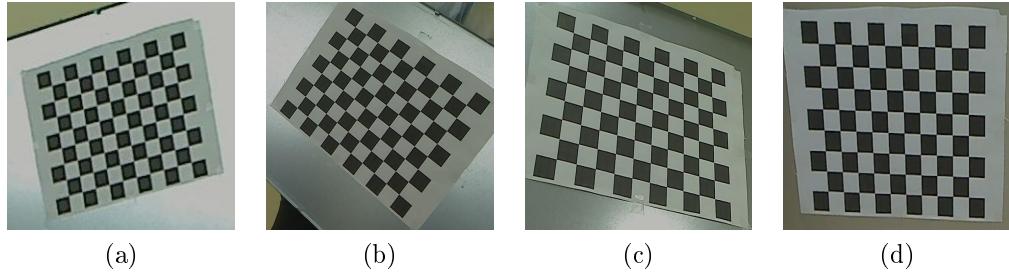


Figure 2-5: Some views used for camera calibration

of rotation and translation parameters to the list of unknowns. Since there are 15 unknowns per view hence at least 15 equations needs to be solved for the unknowns. This requires at least 2 views since a single view can provide only 8 constraints(i.e., 4 points are required to completely define a homography). Practically, we use more than 2 views(typically greater than 10) with more than 4 points

per view to account for effect of noise in corner-detection. Acquired 3D-2D point correspondences are used to calibrate camera as explained in Section-2.3.

**Results and visualization.** For visualization of camera calibration process a program to plot the views used for calibration was developed. This visualization is similar to [18] but it was not available for projector calibration.

Figures 2-6, 2-7 show snapshots of 3D plot of all views of checkerboard used for camera calibration. Here, figure 2-6 shows that our calibration range was 50cm to 250cm along Z-axis, -25cm to +45cm along Y-axis. Figure 2-7 shows that calibration range

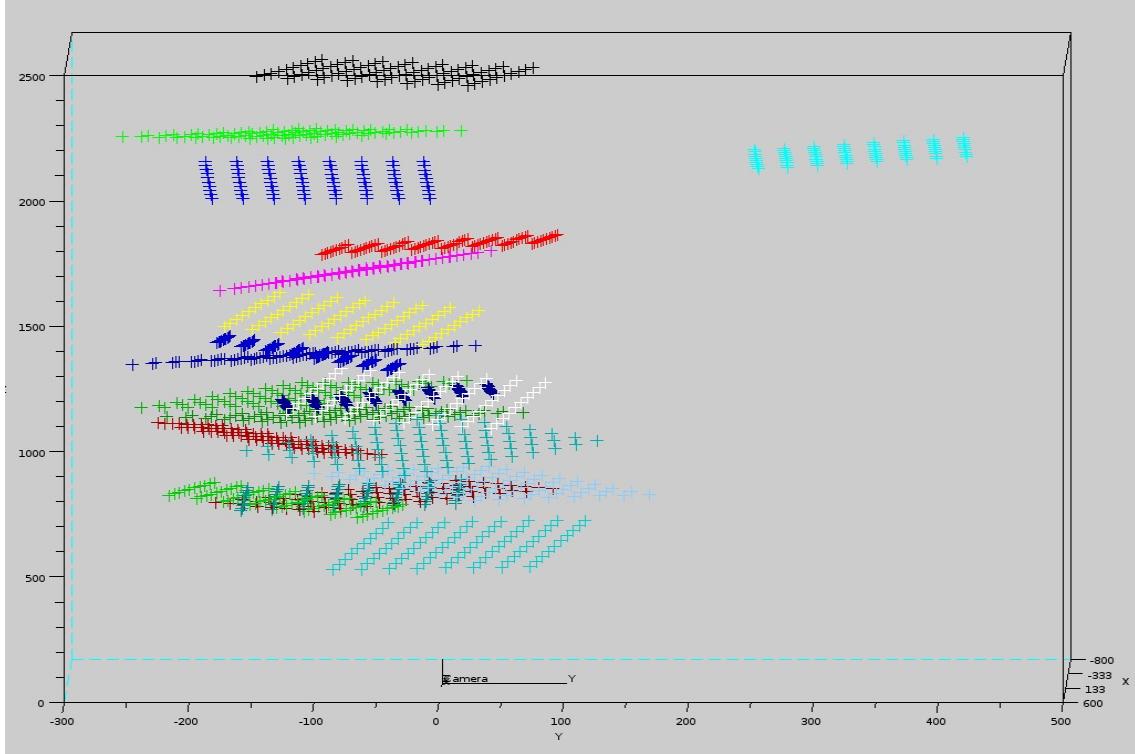


Figure 2-6: Visualization of camera calibration results:View along X-axis

along X-axis was -50cm just above +40cm. All above mentioned distances are with respect to camera coordinates system, which is at (0,0,0) in both figures. So the camera calibration volume was approximately 90cm X 70cm X 200cm. Camera calibration parameters estimated through experiments for Logitech Quickcam sphere AF webcam at native resolution are tabulated in table 2-1. *Reprojection error* mentioned in the table can be defined as a measure of deviation of fitted parameters from the

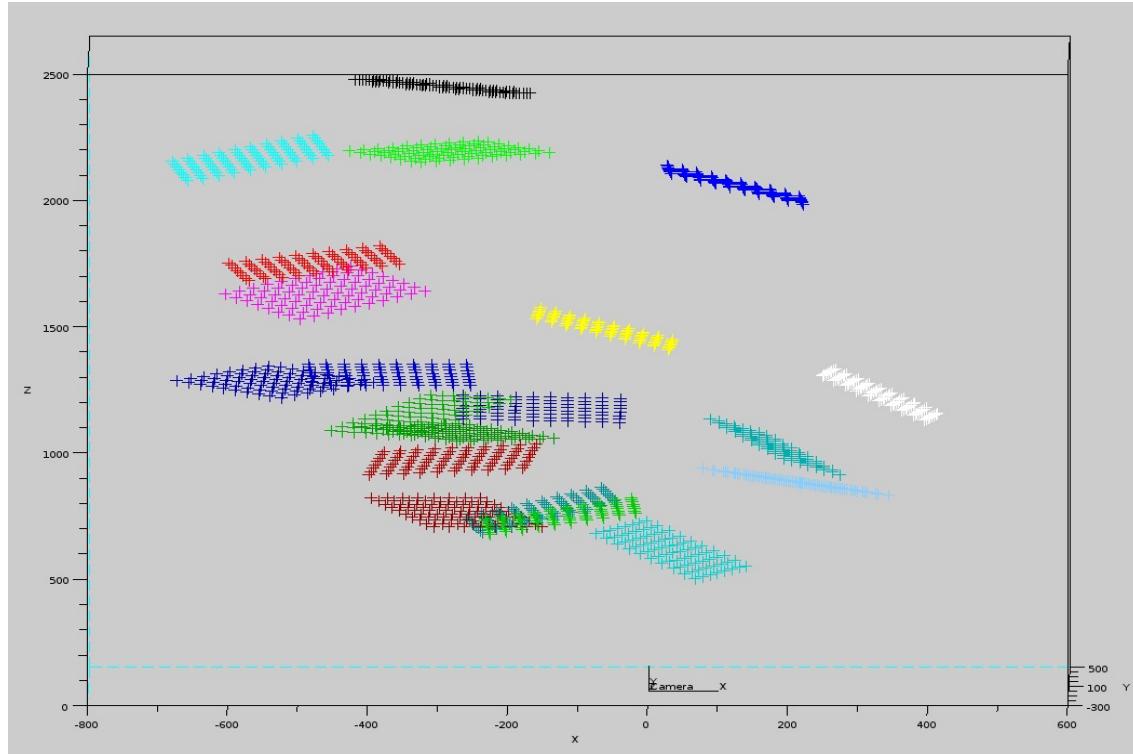


Figure 2-7: Visualization of camera calibration results:View along Y-axis

observed world-camera point correspondences.

Parameter	Estimated value
$f_x$	1362.2
$f_y$	1372.2
$c_x$	803.9
$c_y$	590.1
$k_1$	0.07
$k_2$	-0.14
Reprojection error	0.20

Table 2.1: Estimated Camera intrinsic and distortion model parameters

## 2.4.2 Projector calibration

As already mentioned in Section-2.1, projector calibration problem can be regarded as an inverse-camera calibration problem. Consequently, same procedure as of camera calibration has been used to calibrate the projector. Basically, the idea is to use

camera as a feedback device which will provide the 3D coordinates for the 2D features projected by the projector, hence providing the 3D-2D point correspondence between world-coordinate system and projector-coordinate system needed for projector calibration similar to camera calibration process.

**Procedure:**

1. Camera-to-world homography is computed once.
2. For each view,
  - (a) Projector is physically displaced allowing for rotations and translation with respect to world coordinate system and a checkerboard pattern is projected.
  - (b) Camera captures the projected features and detects their coordinates.
  - (c) Using camera-to-world homography, corresponding 3D coordinates are computed.
  - (d) These 3D coordinates along with the 2D coordinates of projected features(in projector buffer) forms the required 3D-2D correspondence for this view.
3. After sufficient number of such views(typically greater than 10) are captured projector calibration is performed using same algorithm as for camera calibration.

Figure 2-8 shows few views used for projector calibration, projected checkerboard pattern has board dimension 11 squares X 9 squares and has each square of size of 75 pixels X 75 pixels.

**Results and visualization.** Calibration parameters for Sharp PG F200X projector computed by performing above mentioned procedure are tabulated in table 2.2. Figures 2-9, 2-10 show plots of views used for projector calibration. Here, figure 2-9 shows that calibration range for along X-axis is -100cm to +150cm along Z-axis it

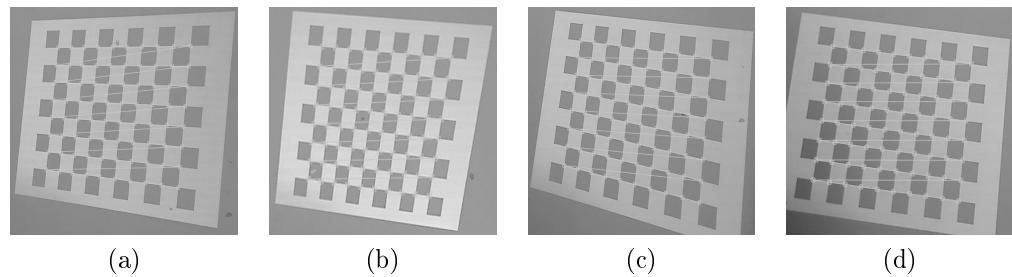


Figure 2-8: Some view used for projector calibration

Parameter	Estimated value
$f_x$	2261.7
$f_y$	2262.8
$c_x$	522.7
$c_y$	713.8

Table 2.2: Estimated Projector intrinsic model parameters

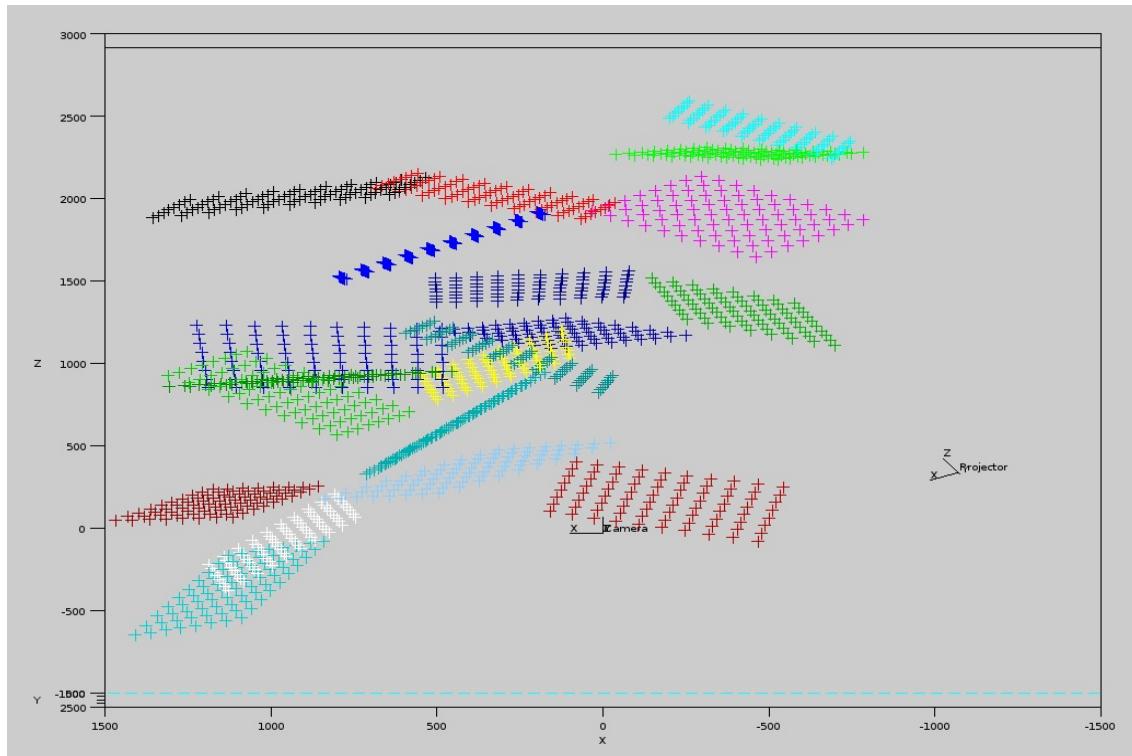


Figure 2-9: Visualization of projector calibration views along Y-axis

is -50cm to +250cm. Figure 2-10 shows calibration range along Y-axis to be from -125cm to +225cm. All measurements are with respect to camera-coordinate system which is shown at (0,0,0) in the figures. Hence calibration volume for projector was 250cm X 400cm X 300cm.

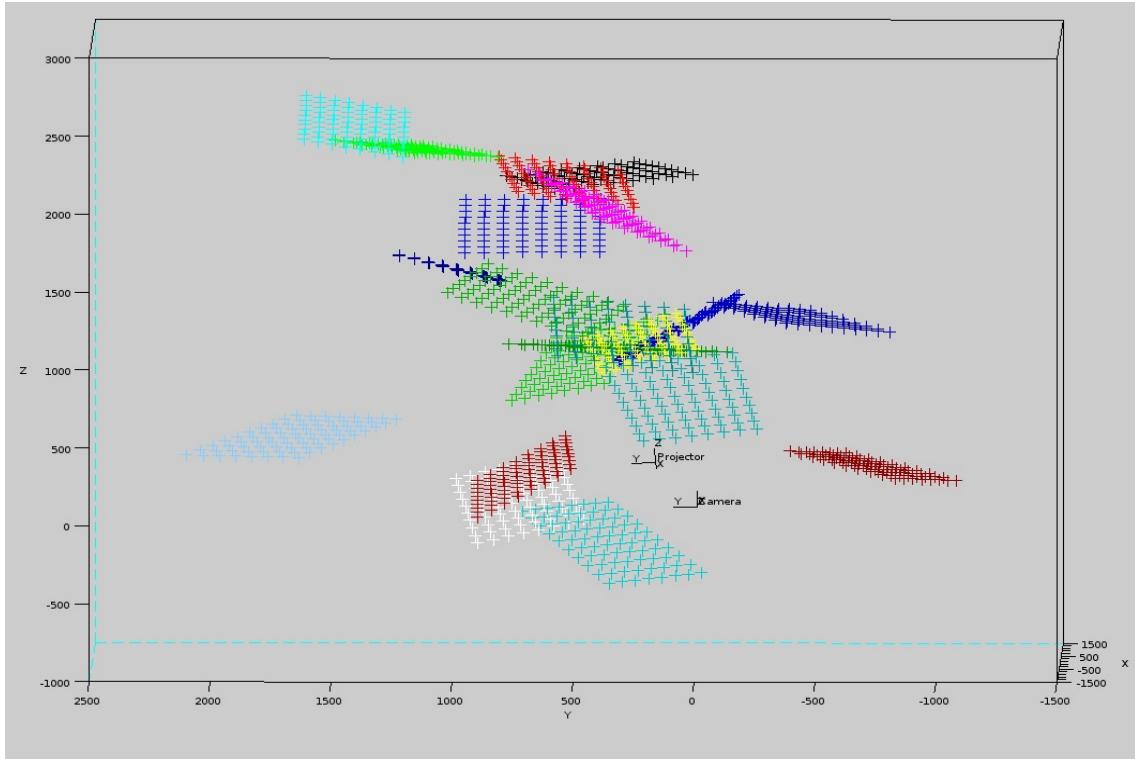


Figure 2-10: Visualization of projector calibration views along X-axis

#### 2.4.3 Camera-projector relative extrinsic calibration

As mentioned in an end note in section 2.3, relative calibration of projector with respect to camera is needed to bring both projector and camera to a common coordinate system which is an implicit assumption in triangulation. In this work, extrinsic calibration function available in OpenCV to determine the extrinsic parameters(i.e., rotation translation of projector coordinate system with respect to camera coordinate system) was used. Rigid body transformations of camera and projector coordinate systems with respect to world coordinate system allow for determining the transformations between them(refer Note on relative calibration in section-2.3 for underlying

equations)

**Procedure:**

1. Extrinsic calibration of camera with respect to world coordinate system:
  - (a) A known reference plane (whose dimensions are known) containing world coordinate system is captured and corners are detected.
  - (b) These detected features are used for computing camera-to-world homography.
  - (c) Using the 3D-2D correspondences determined above, extrinsic calibration of camera is performed.
2. Extrinsic calibration of projector with respect to world coordinate system:
  - (a) Projector projects a checkerboard pattern, which is captured by the camera.
  - (b) Using the camera-to-world homography computed above, 3D coordinates for the projected feature points can be computed.
  - (c) These 3D-2D correspondences (3D coordinates of projected feature points and corresponding 2D coordinates in projector image) are used for computing relative rotation and translation of projector coordinate system with respect to world coordinate system.
3. Relative rotation and translation of projector coordinate system with respect to camera coordinate system is computed (refer Note on relative calibration in section 2.3 for underlying equations)

**Results and visualization.** Estimated extrinsic calibration parameters for a sample camera-projector geometry are tabulated in table 2.3, 2.4, 2.5. It should be noted however that these results are for one specific system configuration and do not represent any intrinsic characteristic of system. Hence they are called extrinsic parameters.

In tables 2.3. 2.4, 2.5,  $(r_1^{wc}, r_2^{wc}, r_3^{wc})$  is rotation transformation from world coordinate system to camera coordinate system  $(t_1^{wc}, t_2^{wc}, t_3^{wc})$  is translation transformation from world coordinate system to camera coordinate system  $(r_1^{wp}, r_2^{wp}, r_3^{wp})$  is rotation transformation from world coordinate system to projector coordinate system  $(t_1^{wp}, t_2^{wp}, t_3^{wp})$  is translation transformation from world coordinate system to projector coordinate system  $(r_1^{pc}, r_2^{pc}, r_3^{pc})$  is rotation transformation from projector coordinate system to camera coordinate system  $(t_1^{pc}, t_2^{pc}, t_3^{pc})$  is translation transformation from projector coordinate system to camera coordinates system. Rotation vectors are expressed in radians in axis-angle form, translation is expressed in ‘mm’.

Parameter	Estimated value
$r_1^{wc}$	0.128
$r_2^{wc}$	0.239
$r_3^{wc}$	0.035
$t_x^{wc}$	-920.7
$t_y^{wc}$	-834.6
$t_z^{wc}$	2960.6

Table 2.3: Estimated camera extrinsic parameters

Parameter	Estimated value
$r_1^{wp}$	-0.047
$r_2^{wp}$	-0.246
$r_3^{wp}$	0.008
$t_x^{wp}$	-1054.0
$t_y^{wp}$	-1235.0
$t_z^{wp}$	2281.0

Table 2.4: Estimated projector extrinsic parameters

Parameter	Estimated value
$r_1^{pc}$	-0.086
$r_2^{pc}$	0.484
$r_3^{pc}$	0.006
$t_x^{pc}$	-1080.3
$t_y^{pc}$	188.9
$t_z^{pc}$	359.4

Table 2.5: Estimated projector-camera relative parameters

View used for extrinsic calibration of projector camera system is depicted in Figure 2-11. Figure 2-12 shows the plan-view of estimated projector-camera-world configuration. In figure 2-12 ‘green’ object is the top-view of projected checkerboard. Estimated positions of world-coordinate system, projector-coordinate system and camera-coordinate system are also shown.



Figure 2-11: Actual setup for extrinsic calibration

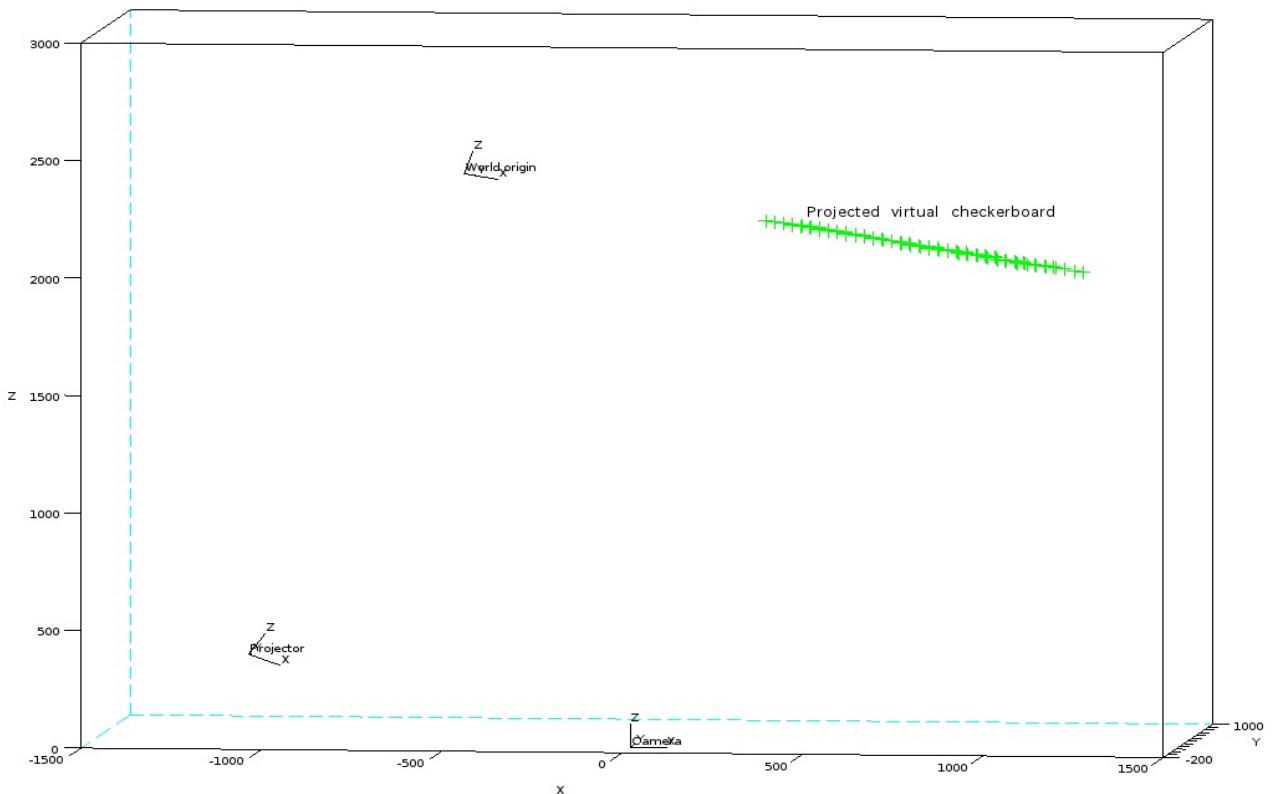


Figure 2-12: Estimated system extrinsic geometry

## 2.5 Discussion

During camera calibration it was observed that for larger distances ( $\sim 2m$ ) of calibration board from camera, sub-pixel corner detection algorithm gives inaccurate results. Further, there has been no study on optimal *window size* and *termination criteria* for sub-pixel corner detection algorithm in the literature due to which these parameters were decided experimentally.

Projector calibration using OpenCV camera calibration algorithm gives non-repeatable estimated parameters of which reason is still unknown and deserves further study.

It was observed that there are certain camera-projector geometric configurations for example when base-line between optical centers of camera and projector is small, for which OpenCV pose-estimation gives degenerate results, again this problem needs further study.

## 2.6 Summary

In this chapter, camera model was described as a series of transformations on a real world 3D point. In review, currently most popularly used calibration algorithms were briefly described. Further, working of the developed system calibration module was described with individual explanations of used procedure for camera, projector and relative extrinsic calibration. To make calibration results more comprehensive, individual plots for camera, projector and relative extrinsic calibration process were presented.

# Chapter 3

## Stereo correspondence

### 3.1 Overview of correspondence problem

As described earlier, correspondence problem involves determining point-to-point relationship between 2 views but since these views come from projective elements like cameras, projectors, a pixel-to-pixel relation between this pair (called binocular-pair or stereo-pair) must be determined. In nutshell, establishing correspondence (technically called stereo-correspondence) means to know for a point in 3D world, which pixel in camera & in projector (or another camera) will see it. Following section gives introduction to the current state-of-art in field of stereo-correspondence.

### 3.2 State-of-art approaches

To solve shape-acquisition problem following classification has been reported in literature [23][24]:

1. Contact techniques:

    1.1 Non-destructive techniques:

        1.1.1 Coordinate measurement machines

        1.1.2 Jointed arms

    1.2 Destructive techniques:

        1.2.1 Slicing

2 Non-contact techniques:

2.1 Transmissive techniques:

2.1.1 Industrial CT

2.2 Reflective techniques:

2.2.1 Optical techniques:

2.2.1.1 Active techniques:

2.2.1.1.1 Active stereo

2.2.1.1.2 Imaging radar

2.2.1.1.3 Active depth from defocus

2.2.1.2 Passive techniques:

2.2.1.2.1 Passive stereo

2.2.1.2.2 Depth from defocus

2.2.1.2.3 Shape from shading

2.2.1.2.4 Shape from silhouette

2.2.2 Non-optical methods:

2.2.2.1 Microwave radar

2.2.2.2 SONAR

In this work focus is on active stereo approach using structured light technique. So present survey will describe current state-of-art in structured-light techniques. Structured light systems deploy off-the-shelf video cameras & projectors as their binocular elements. Here a camera pair of passive stereo methods is replaced by a camera-projector pair. It provides solution to the problem of correspondence for texture-less surfaces experienced by passive stereo camera pair because of lack of texture marks which act as important cues for establishing correspondence. Projection of explicit patterns which create an artificial texture on object surface, greatly simplifies the problem of correspondence. In these techniques, projector projects a pattern(s) which encode the information regarding the projector pixel which has illuminated a point in the scene. Camera captures these patterns and by decoding the pattern which is functionally equivalent to solving the correspondence problem in passive stereo vi-

sion, it extracts the unique ID of projector pixel that illuminated the same point in 3D space as is captured by camera. Figure 3-1<sup>1</sup> provides conceptual representation of structured-light systems.

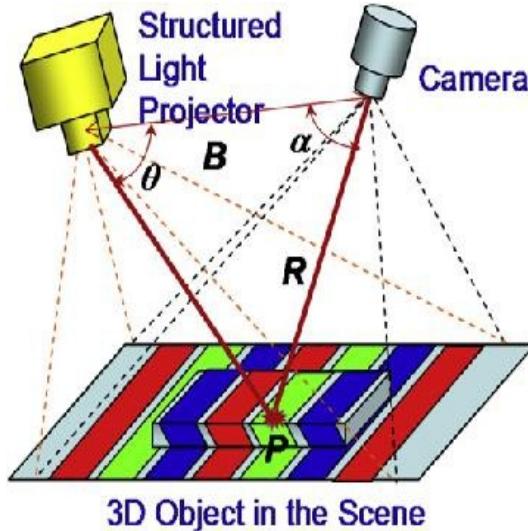


Figure 3-1: A typical structured light system

In figure 3-1,  $\theta$  denotes the angle between a projector optical-ray & baseline between camera and projector optical center. Similarly  $\alpha$  denotes the angle between corresponding camera optical-ray & baseline between camera and projector optical center. Projected pattern in the figure is just a representative of structured-light technique.

Structured light techniques have been classified in 4 major classes[23][24][26] based on how correspondence information is encoded in the projected patterns:

**Time multiplexing.** Every point is encoded by series of intensities across multiple patterns.

**Spatial encoding.** Every pixel is encoded based on intensities of its neighborhood pixels.

---

<sup>1</sup>Figure source: Refer [25]

**Direct encoding.** Every pixel is encoded by its unique intensity and does not depend upon any spatial or temporal correlations among pixels.

**Hybrid encoding.** It encodes pixels both spatially and temporally to project lesser number of patterns than conventional time-multiplexing techniques and to have more robustness as compared to spatial methods.

It should be noted however that this classification do not enforce unique code for each projected pixel i.e., there may be a set of pixels with same code. This pattern design decision depends upon required resolution in the reconstructed scene.

## 1. Time multiplexing

Time multiplexing techniques encode unique ID's for pixel(s) of projector over multiple patterns projected sequentially over a period of time. Camera captures all these pattern in an order to determine correct ID for each pixel(or pixel group). Depending upon type of correspondence required i.e. camera-pixel to projector-pixel or camera pixel-plane to projector pixel-number of axes of codification can be chosen to be both vertical and horizontal or one of them respectively. To be more clear, axis of codification means axis of projected pattern along which the codes are embedded other axis will simply have same copy of intensity for each code in each pattern and hence will not contribute to pixel codification. Examples of time-multiplexing based structured light patterns are binary-coded patterns, N-ary codes, gray code+phase shift method. Here i will discuss these methods as representatives for time multiplexing approach.

### (a) Binary coded patterns

In these patterns[27], any pixel in camera and projector is identified by the sequence of binary intensities it receives or projects respectively. Specifically, for a  $N * M$  projector resolution, if pixels are to be divided into K groups along say Horizontal axis then a total of  $\frac{N}{K}$  code-words will be used which will requires projection of  $\log_2 \lceil \frac{N}{K} \rceil$  patterns sequentially. A pattern

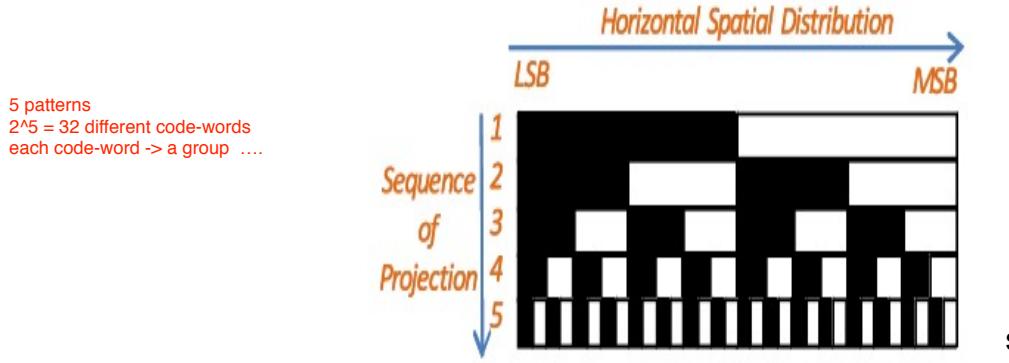


Figure 3-2: Binary coded pattern:Design

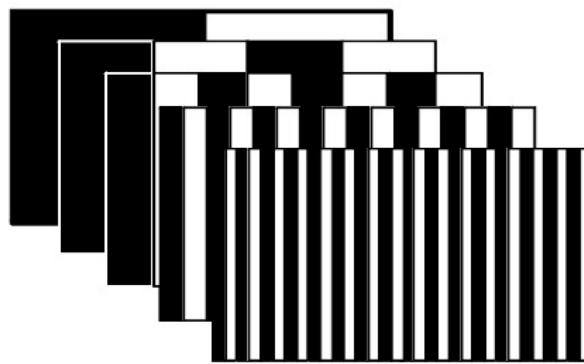


Figure 3-3: Binary coded pattern:Projection

say  $j^{th}$  will represent intensities corresponding to the  $j^{th}$  bit of all code-words across its horizontal axis. Hence every pattern is called a *bit-plane*. Figure 3-2<sup>2</sup> shows design of code-words, figure 3-3<sup>3</sup> shows projection of simple binary-coded patterns scheme. After all patterns are captured by camera, a code at every pixel is computed based on intensities incident on that pixel during pattern projection. A single pattern determines single bit in code-word of all pixels in the camera image.

### (b) N-ary coded patterns

In case of binary-coded patterns it can be observed that for high resolution more code-words will be required which directly increases the number of patterns to be projected. For applications requiring faster acquisition rates,

---

<sup>2</sup>Figure source: Refer [25]

<sup>3</sup>Figure source: Refer [25]

number of patterns can be reduced yet meeting the resolution requirement if we increase number of intensity levels that a pixel can have[25][28][29]. This approach can allow for dynamic scene shape recovery at the cost of decreased robustness against environmental noise. As it can be deduced, color-coded patterns are special cases of this technique. Figure 3-4<sup>4</sup> shows the gray-level coding scheme, here multiple allowed gray-level can be observed.

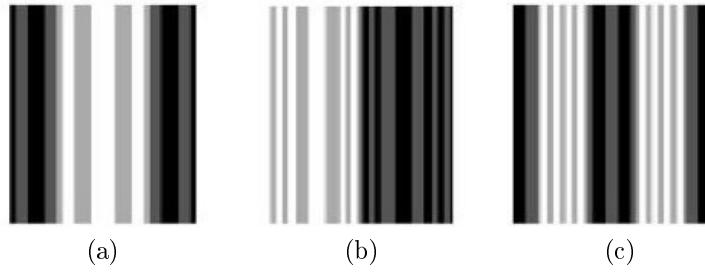


Figure 3-4: Gray coded patterns

### (c) Combination of gray-coded patterns and phase-shift techniques

Gray-coded patterns were proposed to increase robustness of code recovery process against possibility of false-code extraction as compared to simple binary-coded patterns, since successive gray code-words differ at only one bit position thereby reducing chances of assigning same code to the pixels belonging to neighboring code-word groups. It has a straightforward analogy with the considerations in code designing for noisy communication channel. For a treatment of structured-light patterns in way similar to communication code designing over a noisy channel, refer[30].

In phase-shift technique, phase shifted sinusoidal interference fringes are projected on object surface. In camera, unique code to each pixel is assigned in the form of phase of the signal incident on it. Similarly in projector image a phase value is assigned to each pixel which represents spatial position of that pixel in the projected sinusoidal periodic waveform. It can

---

<sup>4</sup>Figure source: Refer [25]

be deduced that this phase value indirectly encodes the position of projector pixel that projected that signal and hence the correspondence information. Figure 3-5<sup>5</sup> shows the spatially aligned combination of phase-shifted and gray coded patterns. Further details of this technique are described in section-3.4 with its actual working.

Gray-coded patterns provide a way to determine a region of pixels uniquely by assigning unique code to it, whereas phase-shift technique can provide phase variation up to pixel-level within that region. Although phase-shift technique is potent to be used alone to provide pixel-level resolution but computational complexity of accurate phase value computation constraints the image resolution and scene geometry.[31] describes details of problems in accurate phase estimation called *phase-unwrapping*. Hence to alleviate problem of low resolution of gray-coded patterns and computationally expensive phase-unwrapping problem, combination of both techniques is used. Here period number of a sinusoidal fringe is determined with the help of gray-coded pattern and pixel-level correspondence within that period is computed using phase-shift technique. Although it comes at a cost of increased number of patterns(typically greater than 10) that must be projected as compared to the original phase shift approach which typically requires 3 or 4 patterns. In pattern design, width of a code-word(in pixels) is equal to width of a period in phase-shifted pattern to assign a unique code-word to each fringe period and hence perform phase unwrapping. A variant of this approach is the subject of this work.

## 2. Spatial encoding methods

As already mentioned these methods extract more information per pattern by encoding unique identity of pixels in their neighborhood(called *window*) also. Although it can be clearly argued that this approach will have less robustness

---

<sup>5</sup>Figure source: Refer [25]

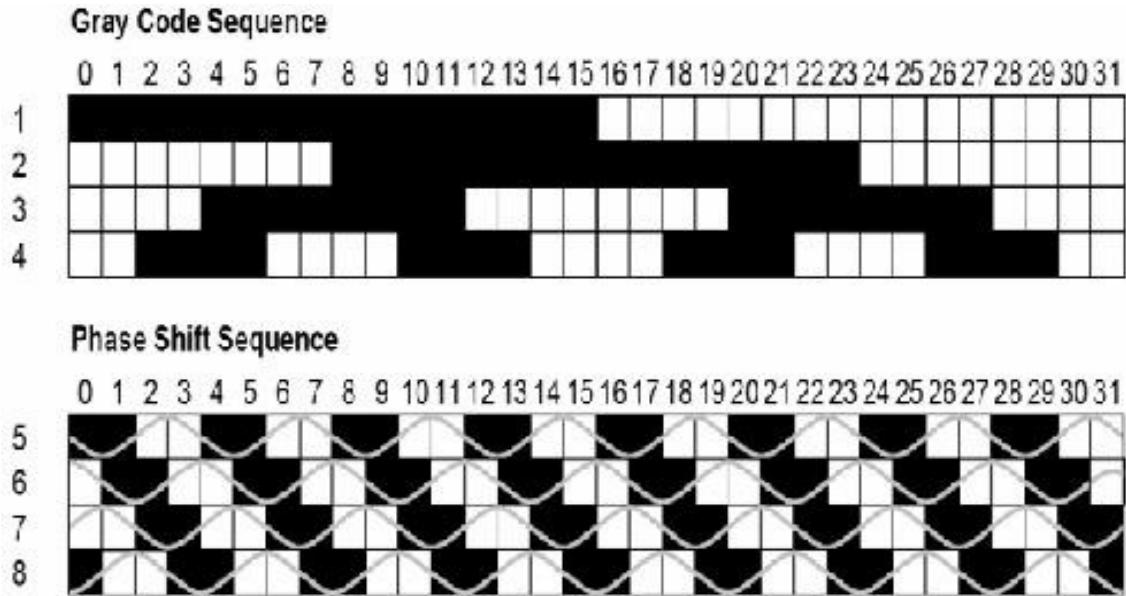


Figure 3-5: Combination of Phase-shift and gray code techniques

as compared to time-multiplexing methods since error in a decoding of one pixel will affect decoding of all dependent pixels. Here size of the window is directly proportional to the number of pixels encoded and inversely proportional to the number of intensity levels used for encoding. Basically, aim of these methods is to obtain correspondence from single captured image which facilitates them to be used for dynamic scenes (like moving objects). As can be noticed, increased number of allowed intensity-levels increases chances of error in decoding because natural scene illumination can potentially interact with these projected patterns making accurate pattern decoding more challenging than binary-coded patterns. De Bruijn patterns and M-arrays are discussed here as representatives of this technique.

#### (a) De Bruijn sequence based pattern

De Bruijn sequence[32] of order 'm' over an alphabet of 'n' symbols is a pseudo-random sequence of length  $n^m$  where every possible combination of length 'm' occurs only once. These sequences have been adapted for pattern generation by considering number of allowed intensity-level as 'n' and window length as 'm'. In terminology of graph theory it can be interpreted

as a graph of  $n^m$  vertices and problem of generating debruijn sequence is to determine a Hamiltonian path in that graph, this problem is also called Traveling salesmen problem. Specifically, a pattern encodes a debruijn sequence. Pattern is divided into slits where each slit has width equal to the order of the pattern. Intensities of sub-slits are generated using individual code-words occurring in debruijn sequence. Total number of intensity levels used in the pattern are limited by the alphabet size of used de-bruijn sequence. To decode a pixel it is only required to determine the code-word or window to which it belongs which will depend upon the intensities of its neighborhood pixels. Figure 3-6<sup>6</sup> shows patterns generated using debruijn sequence of various alphabets and order.

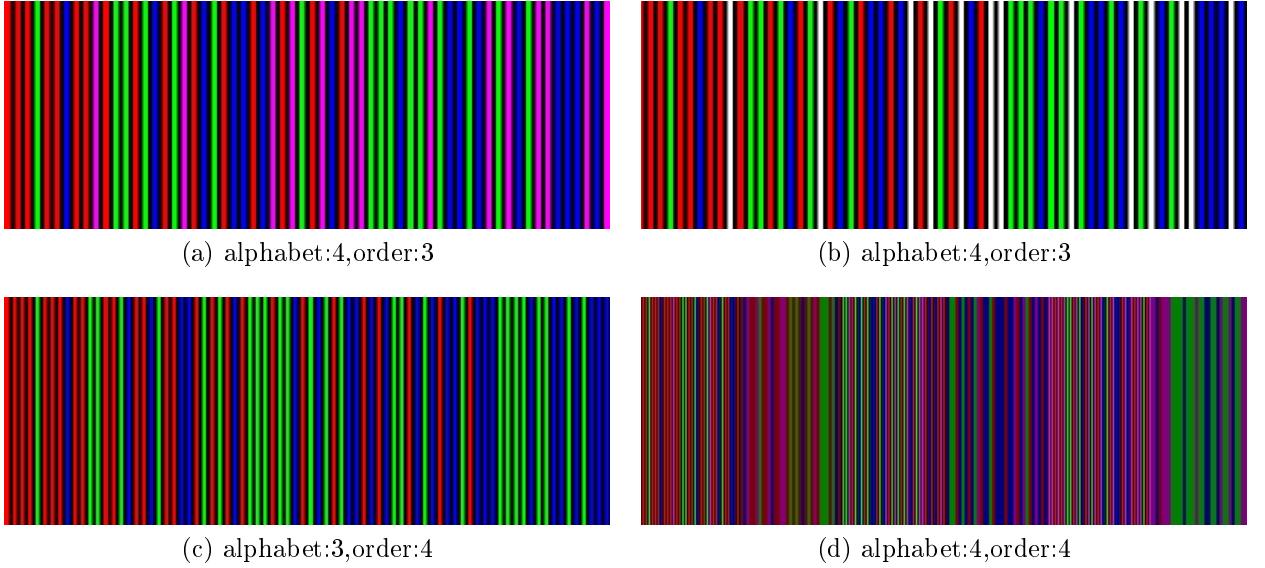


Figure 3-6: De-bruijn sequence patterns

### (b) M-array based pattern

M-arrays[34][35] are bi-dimensional extensions of debruijn sequences. Hence a M-array of  $N * M$  dimension will have all 2D code-words(in window) of dimension  $W * H$  that will occur only once in complete sequence/array. One code-word will be present in one window, so it precisely means that

---

<sup>6</sup>Figure source: [33]

one 2D-window will appear only once in the whole array. Figure 3-7 shows a M-array based pattern.



Figure 3-7: M-array based pattern

### 3. Direct coding techniques

In these techniques, every encoded pixel is identified by its own color/intensity. Although theoretically high resolution for dynamic scenes can be attained but practically these approach are highly susceptible to noise because of large number of intensity-levels used. It also implies that camera with very large pixel depth have to be used for detecting changes in intensity across the scene to extract code from it. Efforts in this direction include grey patterns[36][37][38] for example intensity ratio based grey pattern by [36] which fades away from black to white. Colored patterns[39][40] for example [39] proposed projection of rainbow pattern whereas [40] proposed relatively robust patterns which effectively cancels out the object color by projection of three shifted patterns but at the cost of increasing acquisition time. This approach can be classified as hybrid approach also since it includes combination of direct and temporal techniques. Figure 3-8<sup>7</sup> shows the 3 shifted rainbow pattern as representatives of colored patterns used for direct coding techniques.

---

<sup>7</sup>Figure source: Refer [26]

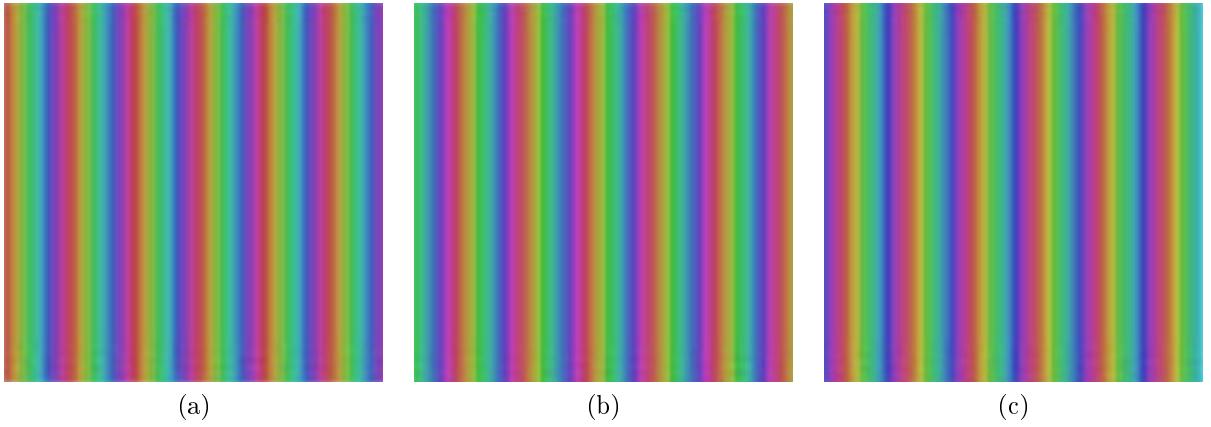


Figure 3-8: Direct coding method:Rainbow patterns

#### 4. Hybrid coding techniques

In order to attain benefits of both spatial and temporal techniques, hybrid patterns[41] were also proposed where patterns are projected temporally but their encoding also depends on neighborhood thereby allowing more information to be encoded per pattern and increased robustness from temporal technique since multiple frames contain the encoding information hence dependence on neighborhood pixels is less as compared to purely spatial neighborhood approach. This approach allowed for slow motion in the scene which is exceptional from point of view temporal-encoding schemes.

### 3.3 Approach followed in this work

In this work, i have attempted to investigate the accuracy achievable by Phase-shift technique for 3D scanning since currently it is the most popular approach[42][43] for its claimed capability to achieve high accuracy and resolution in the structured-light category. It was decided to develop it also rather than just accuracy analysis because there is no *open-source* implementation available for this technique except the work of [44] but it does not have documentation, is still under development and it is only for windows platform whereas we are working mainly on Linux systems. Work by [45] also attempted for three phase scanning but at the time of development had no

means for system calibration and had only naive phase unwrapping algorithm which was far from being usable as a reference for accuracy analysis.

Here, a combination of phase-shift technique with binary coded patterns instead of purely using phase-shifting based approach is used, reason for which is described in section-3.5. This approach gives higher spatial resolution of phase-shift technique and high noise robustness and considerable reduction in correspondence computation complexity of binary-coded patterns. Implementation details of this approach will be described in following section.

## 3.4 Working of developed stereo-correspondence modules

As already shown in Figure 1-2, stereo-correspondence problem is solved using a group of modules which will be described in the next subsections.

### 3.4.1 Pattern generation

This module allows for generation of sinusoidal interference fringe patterns of arbitrary phase shift, fringe width and binary coded patterns of arbitrary number of codewords. Binary coded patterns are generated keeping width of a sinusoid period in mind because we want to resolve among individual sinusoid periods across a pattern using binary coded patterns, so bit-width of a binary coded pattern must be equal to the period of a sinusoid. In our implementation, target was to perform ray-ray triangulation(computing intersection of corresponding optical rays from projector and camera pixels) hence we need to project horizontal and vertical versions of both fringe and binary coded patterns. Figures 3-9 and 3-10 show vertical and horizontal fringe patterns whereas figures 3-11 and 3-12 show vertical and horizontal binary coded patterns respectively. Binary coded patterns were designed using following relations:

$$N_v^c = \frac{W_p}{w} \quad (3.1)$$

$$N_h^c = \frac{H_p}{w} \quad (3.2)$$

$$N_v^p = \log_2(N_v^c) \quad (3.3)$$

$$N_h^p = \log_2(N_h^c) \quad (3.4)$$

where,

$N_v^c$ : Number of vertical binary code-words

$N_h^c$ : Number of horizontal binary code-words

$N_v^p$ : Number of vertical binary coded patterns

$N_h^p$ : Number of horizontal binary coded patterns

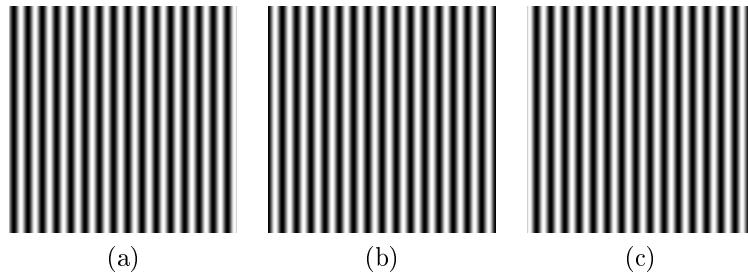


Figure 3-9: Vertical phase shifted patterns

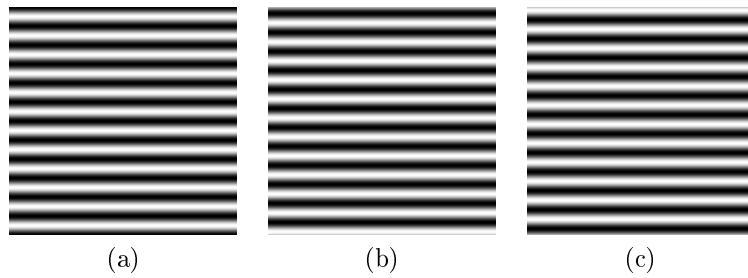


Figure 3-10: Horizontal phase shifted patterns

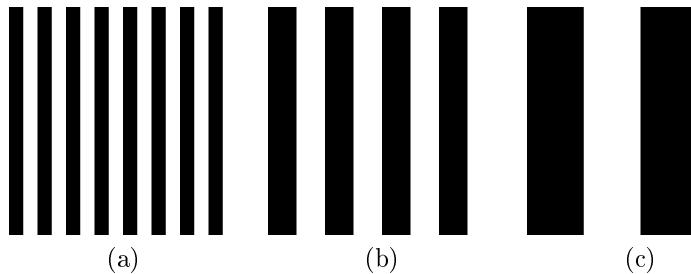


Figure 3-11: Vertical binary coded patterns

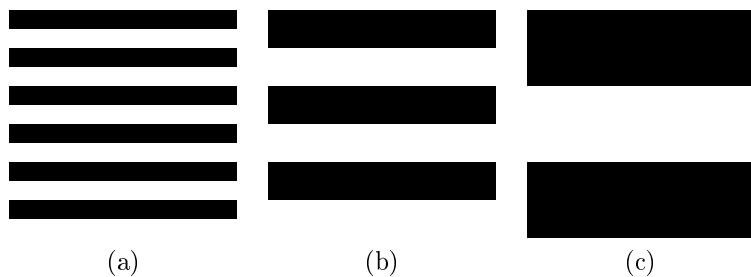


Figure 3-12: Horizontal binary coded patterns

### 3.4.2 Pattern projection and capture

This module is developed for projection of binary coded and fringe patterns and to capture the patterns in sequential order. Currently it has only interactive control to project and capture patterns, but future version will include a synchronizing function to determine a common frequency of pattern projection and capture. In addition this module allows to control brightness of projector and camera. It allows to manipulate shutter-speed of camera hence facilitating to adapt according to varying illumination conditions. All projected and captured images are undistorted(i.e. camera/projector lens distortion effects are reduced) using the projector and camera intrinsic calibration parameters respectively. Figure 3-13 shows the captured vertical fringe and binary coded patterns. Figure 3-14 shows some captured horizontal fringe and binary coded patterns.

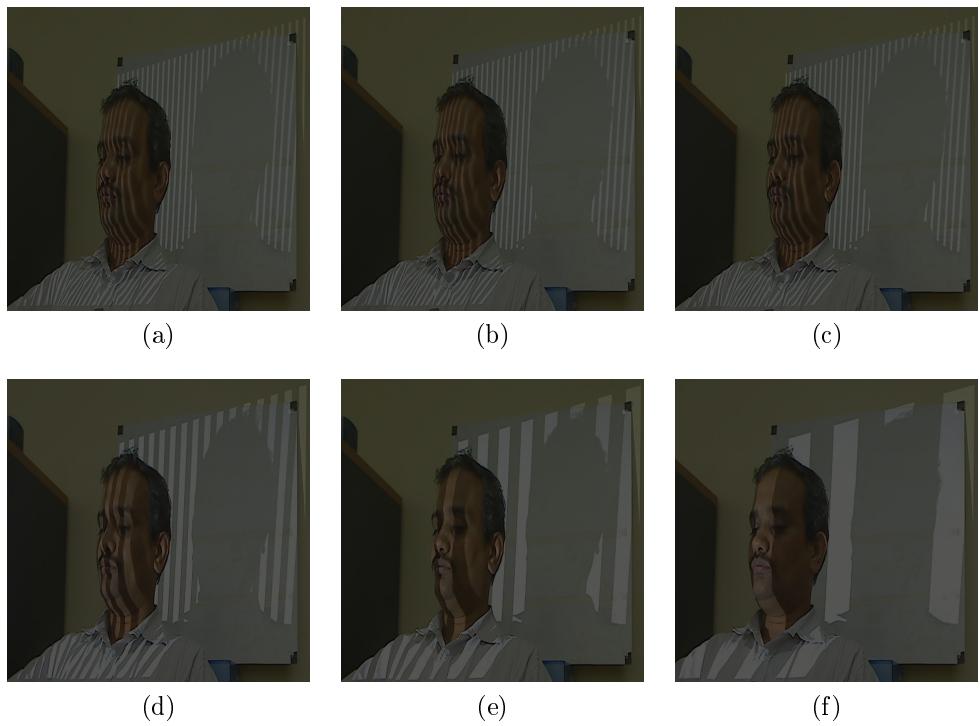


Figure 3-13: Some captured vertical fringe and binary coded patterns

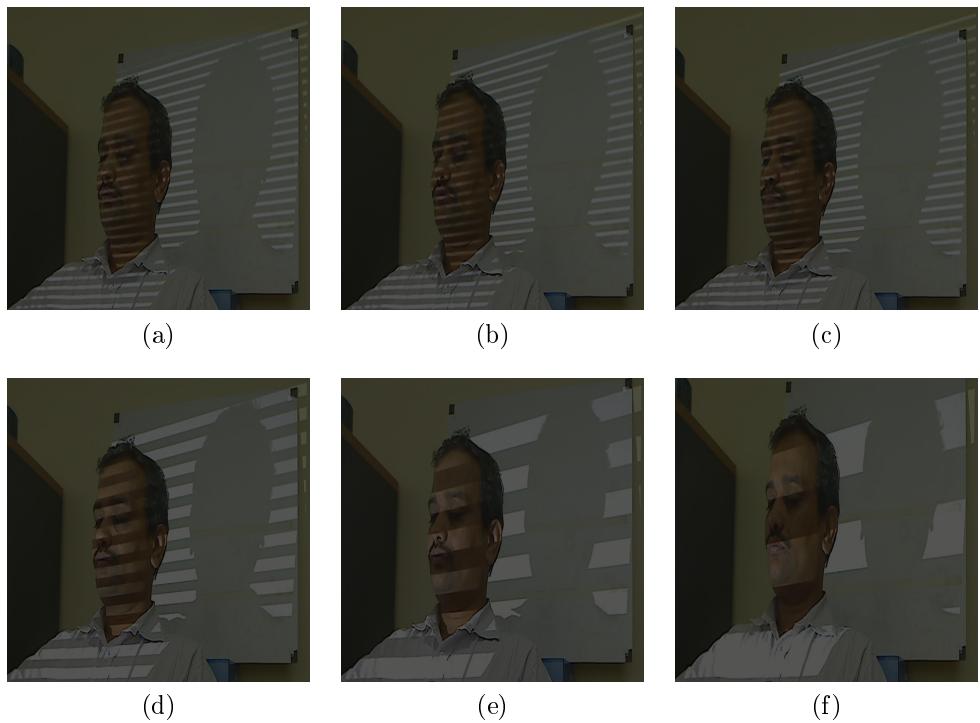


Figure 3-14: Some captured horizontal fringe and binary coded patterns

### 3.4.3 Phase wrapping

Projected sinusoidal fringe pattern is periodic in nature. It implies that there exists ambiguity in determining the phase of any point across the pattern. This ambiguity is manifested in the illumination model used in Phase-shift technique. Following equations show the assumed illumination model at every camera pixel for three phase-shift technique[9] for both horizontal and vertical fringe patterns:

$$\begin{aligned} I_1 &= I_{dc} + I_{mod} * \cos(\phi - \psi) \\ I_2 &= I_{dc} + I_{mod} * \cos(\phi) \\ I_3 &= I_{dc} + I_{mod} * \cos(\phi + \psi) \end{aligned} \quad (3.5)$$

where,

$I_i$ : Captured intensity when  $i^{th}$  pattern is projected

$\phi$ : Wrapped phase

$\psi$ : Constant phase shift between successive patterns

$I_{dc}$ : Background intensity

$I_{mod}$ : Intensity modulation

Assuming  $\psi = \frac{2\pi}{3}$ , using this system of equations  $\phi$  can be solved at every pixel as:

$$\phi = \tan^{-1} \left[ \frac{\sqrt[3]{3}(I_1 - I_3)}{2I_2 - I_1 - I_3} \right] \quad (3.6)$$

where,  $-\pi \leq \phi \leq \pi$

Since  $\phi$  lies from  $-\pi$  to  $\pi$ , information regarding actual phase at every point is mapped to this range hence it is called wrapped phase. It cannot be resolved for correct corresponding projector pixel unless we compute the corresponding period number. This task is performed by module described in next subsection. Figure 3-15 shows the wrapped phase computed from the captured phase shifted fringe images.

**Extension to four and five phase shifted fringe patterns.** In order to assess the effect of increasing the number of phase shifted patterns on accuracy of estimated

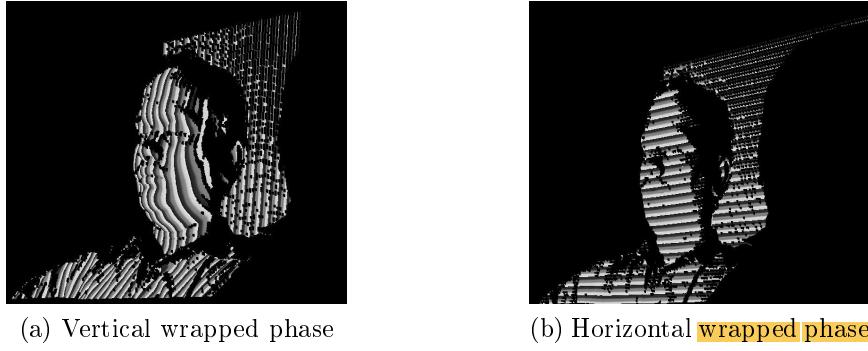


Figure 3-15: Computed wrapped phase

stereo-correspondence, the basic 3D scanner with 3 phase shifted patterns was extended to be able to project and process four and five phase shifted patterns also. It should be noted however that only pattern generator and phase wrapper modules need to be extended for this purpose as all remaining modules are independent of number of patterns used for 3D scanning. Further it should be pointed out that *three* is the minimum number of phase shifted patterns that can be used assuming  $I_{dc}$ ,  $I_{mod}$  and  $\phi$  as unknowns which requires at least 3 equations and hence 3 patterns. Equation 3-7 describes assumed illumination model for 4 phase-shifted pattern technique, equation 3-8 expresses the corresponding wrapped phase. Whereas equation 3-9 expresses the illumination model for 5 phase shifted pattern technique and equation 3-10 represents the corresponding wrapped phase. Unknowns  $I_{dc}$ ,  $I_{mod}$ ,  $\phi$  have same meaning as in equation 3-5,3-6.

For four sinusoidal fringe patterns phase-shifted by  $\pi/2$ ,

$$\begin{aligned}
 I_0 &= I_{dc} + I_{mod} * \cos(\phi) \\
 I_1 &= I_{dc} + I_{mod} * \cos(\phi + \pi/2) \\
 I_2 &= I_{dc} + I_{mod} * \cos(\phi + \pi) \\
 I_3 &= I_{dc} + I_{mod} * \cos(\phi + 3\pi/2)
 \end{aligned} \tag{3.7}$$

Simultaneously solving equations 3-7 gives wrapped phase,  $\phi$  as:

$$\phi = \tan^{-1} \left[ \frac{I_3 - I_1}{I_0 - I_2} \right] \tag{3.8}$$

Similarly for five sinusoidal fringe patterns phase-shifted by  $2\pi/5$ ,

$$\begin{aligned}
 I_0 &= I_{dc} + I_{mod} * \cos(\phi - 4\pi/5) \\
 I_1 &= I_{dc} + I_{mod} * \cos(\phi - 2\pi/5) \\
 I_2 &= I_{dc} + I_{mod} * \cos(\phi) \\
 I_3 &= I_{dc} + I_{mod} * \cos(\phi + 2\pi/5) \\
 I_4 &= I_{dc} + I_{mod} * \cos(\phi + 4\pi/5)
 \end{aligned} \tag{3.9}$$

Solving equations 3-9 gives wrapped phase  $\phi$  as:

$$\phi = \tan^{-1} \left[ 2 \sin \alpha * \left( \frac{I_1 - I_3}{2I_2 - I_4 - I_0} \right) \right] \tag{3.10}$$

### 3.4.4 Pixel decoding and phase unwrapping

Periodic nature of sinusoidal interference fringes creates ambiguous condition in which multiple pixels in camera image can have same wrapped phase hence there is need to assign distinct period numbers to such candidates thereby eliminating the ambiguity. This step is called *phase unwrapping*. Therefore a module to perform phase unwrapping was developed which uses information from binary code projected onto the object to recover this period number. It should be noted here that wrapped and hence unwrapped phase is computed in both vertical and horizontal directions. Hence we get a pair of unwrapped phase associated with each valid camera pixel. Here a camera pixel is called *valid* if it corresponds to the region on which pattern is projected and has a Signal/Noise ratio greater than a threshold(this threshold is environment illumination dependent in current implementation).Equation 3.11 describes the criteria used as a measure of Signal/Noise ratio for 3 phase-shift approach.

$$\eta = \frac{I_{mod}}{I_{dc}} = \frac{\sqrt[3]{3(I_1 - I_3)^2 + (2I_2 - I_1 - I_3)^2}}{I_1 + I_2 + I_3} \tag{3.11}$$

Here, $\eta$  represents the Signal/Noise ratio,hence,pixel(x,y) is *valid* if  $\eta(x, y) > T$ ,where  $T$  is assumed threshold.

Unwrapped phase is calculated using information from both fringe projection and binary coded techniques as. Following equation represent represents this problem:

$$\oint(x, y) = \phi(x, y) + 2\pi * C(x, y) \quad (3.12)$$

where,

$\oint(x, y)$ : Unwrapped phase at pixel (x,y)

$C(x, y)$ : Decoded binary code at pixel (x,y)

This equation is applied for both horizontal and vertical phase unwrapping. Figure 3-16 represents the unwrapped phase corresponding to the wrapped phase shown in figure 3-15.

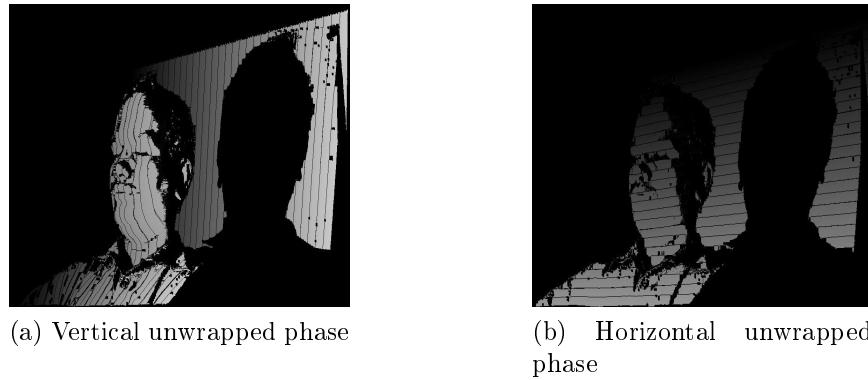


Figure 3-16: Computed unwrapped phase

### 3.4.5 Correspondence computation using absolute phase

Stereo-correspondence between binocular elements like a projector-camera pair relates the points in projector and camera image-space which are observing a common 3D point in the scene.

Once the unwrapped phase is computed,camera-projector pixel-pixel correspondences can be computed directly by searching for pixels in projector with same vertical and horizontal unwrapped phase values as that of corresponding valid camera pixel.

It is computed as:

$$\begin{aligned} X_p &= \lfloor w_{fringe} * \left( \frac{\phi_v(X_c, Y_c)}{2\pi} \right) \rfloor \\ Y_p &= \lfloor w_{fringe} * \left( \frac{\phi_h(X_c, Y_c)}{2\pi} \right) \rfloor \end{aligned} \quad (3.13)$$

where,

$(X_p, Y_p)$ : Projector coordinates corresponding to camera coordinates  $(X_c, Y_c)$

$w_{fringe}$ : Width of projected fringe in pixels.

Figure 3-17 shows the one result of correspondence computation using the developed system.

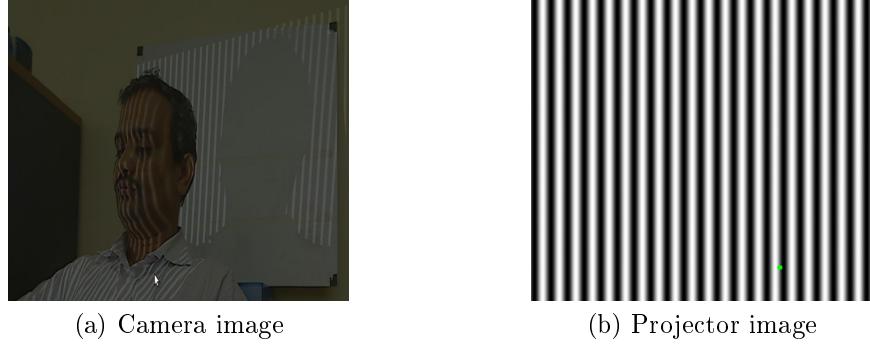


Figure 3-17: Stereo correspondence between camera and projector:*green* spot in (b) corresponds to selected point(cursor) in (a)

### 3.5 Discussion

It was observed that phase-shift technique do not work well on surface with very high(or very low) surface reflectance, surface interreflections, subsurface scattering and in cases when projector or camera is/are defocused. These problems are still a limitation of structured light technique. However some developments for this problem have been reported recently[46].

Further, an ambiguity in decoding the binary coded patterns was observed at the edges of strips which due to projector or camera blur becomes wider *grey* region which cannot be reliably judged black or white in general. Because of this problem, points in the strip region were discarded while evaluating accuracy and precision of

the system.

Although original plan was to fully develop the system purely based on phase-shift technique but our phase-unwrapping implementation[47] was not converging to a solution and was running for >3hrs without solution(on Core2Duo processor E8400 at 3.0GHZ and 2 GB RAM) for an image resolution of 960X720 even after several attempts to check correctness of implementation hence keeping project deadline and time-efficiency of the solution to be developed under consideration it was decided to shift to a combination phase-shift and binary-coded technique to reduce the computation time requirement in the phase-unwrapping process

### 3.6 Summary

In this chapter, the taxonomy of state-of-art approaches for 3D metrology was presented. Since this work is concerned with structured-light based techniques, a brief review of currently used approaches in this field was presented. All stages of coded phase-shift technique for estimating stereo-correspondence were described along with formal description in form of equations and real examples of output from each stage. Section 3.5 described the issues faced during development of this module.

# Chapter 4

## Triangulation

Triangulation is a technique of computing distance of a point with respect to a known reference coordinate system, given calibration parameters, relative pose and correspondence information between imaging elements by performing optical ray-ray or optical ray-plane intersection between corresponding pixels(or plane of pixels) of imaging elements. When noise is present the optical rays may not intersect, hence typically an approximate intersection point is estimated under some assumed noise model.

### 4.1 Current techniques for triangulation

In this section, the current techniques used for triangulation will be briefly reviewed. This section is largely based on [4]. In this work only metric reconstruction (specifically, euclidean reconstruction) is attempted but readers interested in more detailed exposition of different types of 3D reconstructions possible given point correspondences between multiple views of the scene, can refer [3][48][49].

#### 1. Mid-point method

This method proposes to use mid-point of common perpendicular between optical rays corresponding to matching points as the intersection point.

## 2. Quasi-Euclidean approach

In this method a approximation to the correct euclidean frame is selected and mid-point method is performed in it. [4] describes it as a sub-optimal method.

## 3. Linear triangulation

Let  $u_c = P * U_w$ , where  $P$  is projective transformation between a point  $u_c$  in camera image plane and a point  $U_w$  in world coordinate system. Here  $u_c = w * (x, y, 1)^T$  where 'w' is an unknown scale factor and  $u_c$  is in homogenous coordinate representation. Let us denote  $p_i^T$  as  $i^{th}$  row of  $P$ . This relation gives us three equations for one view:

$$\begin{aligned} w * x &= p_1^T * U_w \\ w * y &= p_2^T * U_w \\ w &= p_3^T * U_w \end{aligned} \tag{4.1}$$

Eliminating  $w$ , we get two constraints on  $U_w$ :

$$\begin{aligned} x * (p_3^T * U_w) &= p_1^T * U_w \\ y * (p_3^T * U_w) &= p_2^T * U_w \end{aligned} \tag{4.2}$$

Hence using two views we get 4 equations in three unknowns which can be represented in form of system of Homogenous linear equations:

$$A * U = 0 \tag{4.3}$$

Linear triangulation methods[50] are based on solving system of linear equations mentioned above.

### (a) Linear-Eigen method

Minimizes  $\|AU\|$  under the constraint  $\|U\| = 1$ . Solution is a unit eigenvector, corresponding to smallest eigenvalue of  $A^T * A$ . This problem may be solved using Singular-value-decomposition or Jacobi's method for finding

eigenvalues of symmetric matrix.

### (b) Linear-LS method

This method assumed there can be no point on plane at infinity, thereby setting  $w=1$  in  $U_w$ . This constraint allows system of four linear equations for 3 unknowns whose least-squares solution can be computed using for instance Singular value decomposition method.

## 4. Iterative linear methods

[4] points out the reason for failure of Linear method is due to lack of geometric interpretation of  $\|AU\|$  and they do not minimize the optimal cost criteria proposed by them. To account for this an equivalent iterative version of methods is proposed in literature called *iterative-LS* and *iterative-Eigen*. Although authors in [4] claim iterative methods to be more insensitive to projective transformations than corresponding non-iterative versions. But they also point out that non-convergence is an issue with these methods especially in the case when matching points are near epipoles, hence they suggest these methods to be used with some other supportive method which can be used in case the non-convergence is detected.

## 5. Minimizing the sum-of-magnitudes of distances

In this method, sum magnitude of Euclidean distance between observed noisy correspondence  $(u_r, u_l)$  and possible candidate  $(u'_r, u'_l)$  is minimized, ideal solution is the one which will exactly satisfy:

$$u_r^T * F = 0 \quad (4.4)$$

where,

$u_r$ : position of a point in right view

$u_l$ : position of a point in left view

F: Fundamental matrix[40]

Hence the cost function becomes:

$$c = d(u_r, u'_r) + d(u_l, u'_l) \quad (4.5)$$

where,

$d(u_i, u'_i)$ : represents the 2D euclidean distance between image points  $u_i$  and  $u'_i$

## 6. Polynomial method

In this approach, Gaussian error model is assumed in detected features in image. It minimizes sum of squares of euclidean distances between corresponding ideal and measured points in the individual images as:

$$c = d(u_r, u'_r)^2 + d(u_l, u'_l)^2 \quad (4.6)$$

It is proved to be optimal provided the assumed Gaussian noise-model holds. [4] argues that problem of accurate triangulation is more concerned with accurate determination of point-correspondences. This approach assumes that camera matrices (and hence fundamental matrix) is known more accurately than point-correspondences and assumption that correct matching pair will be most closer to the measured matching pair.

## 4.2 Approach followed in this work

In this work, I have actually implemented [51], Mid-point method and Linear-LS for triangulation but it was observed that only the Linear-LS approach seems to give relatively accurate results. But it has not been analyzed in this work why other approaches do not work since our initial target was to develop a system that can perform 3D reconstruction.

### 4.3 Working of developed triangulation module

For triangulation a module was developed for solving system of linear equations described in section 4.1 as Linear-LS. It used setup since it was assumed that the world coordinate system origin is at a reference plane board corner, all measurements(3D coordinates) are with respect to it. It should be noted that position of world coordinate system is not disturbed after relative extrinsic calibration of camera and projector system. Here equations described in [52] has been rephrased since these are relevant for 3D reconstruction: Let  $(u_i^c, v_i^c)$  and  $(u_i^p, v_i^p)$  be the corresponding camera and projector pixels which will see  $(x_i^w, y_i^w, z_i^w)$  in world. Let  $A_c, A_p$  be the projective transformations from world-coordinate system to camera and projector coordinate system respectively. In homogeneous coordinate representation projective relations can be expressed as:

For camera:

$$\begin{bmatrix} w_c * u_i^c \\ w_c * v_i^c \\ w_c \end{bmatrix} = \begin{bmatrix} a_{1,1}^c & a_{1,2}^c & a_{1,3}^c & a_{1,4}^c \\ a_{2,1}^c & a_{2,2}^c & a_{2,3}^c & a_{2,4}^c \\ a_{3,1}^c & a_{3,2}^c & a_{3,3}^c & a_{3,4}^c \end{bmatrix} \begin{bmatrix} x_i^w \\ y_i^w \\ z_i^w \\ 1 \end{bmatrix} \quad (4.7)$$

Similarly, for projector:

$$\begin{bmatrix} w_c * u_i^p \\ w_c * v_i^p \\ w_p \end{bmatrix} = \begin{bmatrix} a_{1,1}^p & a_{1,2}^p & a_{1,3}^p & a_{1,4}^p \\ a_{2,1}^p & a_{2,2}^p & a_{2,3}^p & a_{2,4}^p \\ a_{3,1}^p & a_{3,2}^p & a_{3,3}^p & a_{3,4}^p \end{bmatrix} \begin{bmatrix} x_i^w \\ y_i^w \\ z_i^w \\ 1 \end{bmatrix} \quad (4.8)$$

In non-matrix form we have,

For camera,

$$\begin{aligned} w_c * u_i^c &= a_{1,1}^c * x_i^w + a_{1,2}^c * y_i^w + a_{1,3}^c * z_i^w + a_{1,4}^c \\ w_c * v_i^c &= a_{2,1}^c * x_i^w + a_{2,2}^c * y_i^w + a_{2,3}^c * z_i^w + a_{2,4}^c \\ w_c &= a_{3,1}^c * x_i^w + a_{3,2}^c * y_i^w + a_{3,3}^c * z_i^w + 1 \end{aligned} \quad (4.9)$$

For projector,

$$\begin{aligned}
w_p * u_i^p &= a_{1,1}^p * x_i^w + a_{1,2}^p * y_i^w + a_{1,3}^p * z_i^w + a_{1,4}^p \\
w_p * v_i^c &= a_{2,1}^p * x_i^w + a_{2,2}^p * y_i^w + a_{2,3}^p * z_i^w + a_{2,4}^p \\
w_p &= a_{3,1}^p * x_i^w + a_{3,2}^p * y_i^w + a_{3,3}^p * z_i^w + 1
\end{aligned} \tag{4.10}$$

Eliminating  $w_p$  and  $w_c$  we get,

For camera,

$$\begin{aligned}
u_i^c - a_{1,4}^c &= (a_{1,1}^c - u_i^c * a_{3,1}^c) * x_i^w + (a_{1,2}^c - u_i^c * a_{3,2}^c) * y_i^w + (a_{1,3}^c - u_i^c * a_{3,3}^c) * z_i^w \\
v_i^c - a_{2,4}^c &= (a_{2,1}^c - v_i^c * a_{3,1}^c) * x_i^w + (a_{2,2}^c - v_i^c * a_{3,2}^c) * y_i^w + (a_{2,3}^c - v_i^c * a_{3,3}^c) * z_i^w
\end{aligned} \tag{4.11}$$

For projector,

$$\begin{aligned}
u_i^p - a_{1,4}^p &= (a_{1,1}^p - u_i^p * a_{3,1}^p) * x_i^w + (a_{1,2}^p - u_i^p * a_{3,2}^p) * y_i^w + (a_{1,3}^p - u_i^p * a_{3,3}^p) * z_i^w \\
v_i^p - a_{2,4}^p &= (a_{2,1}^p - v_i^p * a_{3,1}^p) * x_i^w + (a_{2,2}^p - v_i^p * a_{3,2}^p) * y_i^w + (a_{2,3}^p - v_i^p * a_{3,3}^p) * z_i^w
\end{aligned} \tag{4.12}$$

System of equations 4.11,4.12 can be combined and written in the matrix form as:

$$PV = F \tag{4.13}$$

where,

$$P = \begin{bmatrix} (a_{1,1}^c - u_i^c * a_{3,1}^c) & (a_{1,2}^c - u_i^c * a_{3,2}^c) & (a_{1,3}^c - u_i^c * a_{3,3}^c) \\ (a_{2,1}^c - v_i^c * a_{3,1}^c) & (a_{2,2}^c - v_i^c * a_{3,2}^c) & (a_{2,3}^c - v_i^c * a_{3,3}^c) \\ (a_{1,1}^p - u_i^p * a_{3,1}^p) & (a_{1,2}^p - u_i^p * a_{3,2}^p) & (a_{1,3}^p - u_i^p * a_{3,3}^p) \\ (a_{2,1}^p - v_i^p * a_{3,1}^p) & (a_{2,2}^p - v_i^p * a_{3,2}^p) & (a_{2,3}^p - v_i^p * a_{3,3}^p) \end{bmatrix}$$

$$V = \begin{bmatrix} x_i^w \\ y_i^w \\ z_i^w \end{bmatrix}$$

$$F = \begin{bmatrix} a_{3,4}^c * u_i^c - a_{1,4}^c \\ a_{3,4}^c * v_i^c - a_{2,4}^c \\ a_{3,4}^p * u_i^p - a_{1,4}^p \\ a_{3,4}^p * v_i^p - a_{2,4}^p \end{bmatrix}$$

Hence at each pixel  $V = (P^T * P)^{-1} * P^T * F$  holds, which represents the 3D coordinate corresponding to  $(u_i^c, v_i^c)$  and  $(u_i^p, v_i^p)$  in camera and projector respectively.

## 4.4 Procedure

Here correspondence computed from absolute phase computation step(please refer equation 3.12) is used along with the camera and projector intrinsic and extrinsic calibration parameters to compute depth.

## 4.5 Visualization of Point clouds

To visualize, filter, smooth-out, down-sample and triangulate(mesh-generation from 3D data-set) the results, *Point cloud library(PCL)*[53] was used. Specifically, *PCL visualizer* was used for point cloud visualization. *Statistical outlier detection algorithm* was used for filtering point cloud data. *Voxel-based down-sampling method* was used for reducing the number of points on rendered point cloud so as to provide relatively faster 3D rendering rate. *Moving least squares(MLS)* algorithm was used for smoothing the point cloud data. *Greedy triangulation algorithm* was used to generate mesh output. 3D point-clouds were also visualized in the ingeniously developed visualization tool *ANUVI-v1.0.2*. Figure 4-1 shows snapshots of some sample 3D reconstructions using our system.

## 4.6 Discussion

During 3D reconstruction of our reference plane(in our case it was a wall) an angle between *true* reference plane(i.e., Z=0) and reconstructed plane was observed which is suspected to be because of incorrect system calibration parameters specifically extrinsic relative calibration of projector with respect to camera. This issue needs further study and experimentation to reveal the actual source of the problem.

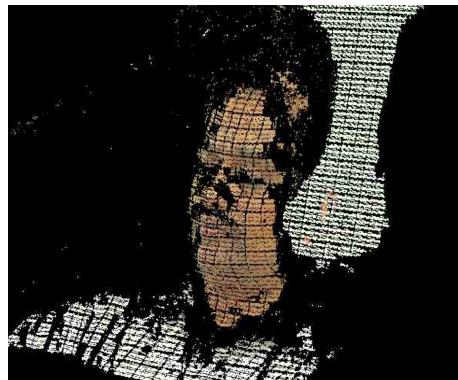
Further, evaluation of measurement accuracy(described in more detail in chapter-5) revealed average absolute relative error of  $\sim 1.018\%$  in distance range of 1.3m to 2.5m from the developed system.

## 4.7 Summary

In this chapter, a brief description of the currently used optical triangulation techniques used for metric reconstruction was presented. Further the approach used in this work for triangulation was mathematically described along with some 3D reconstruction examples to show the working of triangulation module.



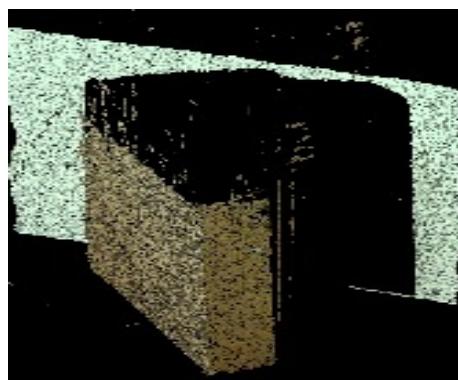
(a) 2D face image



(b) 3D reconstruction of face



(c) 2D image of box in front of wall



(d) 3D reconstruction of box in front of wall



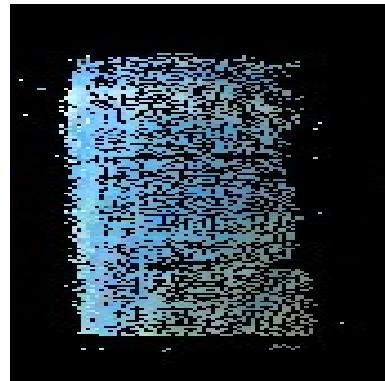
(e) 2D image of a chair with background



(f) 3D reconstruction of chair with background



(g) A cup



(h) 3D reconstruction of cup

# Chapter 5

## Accuracy evaluation and analysis

Measurement accuracy is among the most crucial benchmarks to judge worth of a metrology equipment. To assess measurement accuracy of the developed system the problem can be viewed as having three components as even the structure of this thesis suggests: System calibration, stereo-correspondence computation and triangulation. Hence it was decided to perform assessments of accuracy of each of these individual components. However scope of this work is limited to accuracy evaluation of stereo-correspondence and system calibration only because these components decide the accuracy of optical triangulation as can be implied from equations 4.7,4.8. Hence the accuracy of optical triangulation method cannot be studied completely isolated from that of stereo-correspondence and system calibration unless both correspondence and calibration estimates are kept fixed.

This chapter is logically divided into two parts. First part covered by sections 5.1 and 5.2 will describe the experiments performed to assess the measurement accuracy and precision of the system developed and further compares it with the nowadays popular 3D sensor *Microsoft Kinect*. This comes under *accuracy evaluation*. Second part covered in sections 5-3 and 5-4 further delves deeper into quantitative assessment of *independent* individual components of the developed system namely system calibration and stereo-correspondence module. This comes under *accuracy analysis*. Coded phase-shift scanner will be abbreviated by *CPSS* in this chapter.

## 5.1 Evaluation of accuracy and precision of developed system

For evaluating the measurement accuracy and precision 3D reconstruction of a planar surface was used. For visualization of point clouds Point cloud library[53] was used. Before describing experimental procedures and results each term will be described for clarity in equations 5-1 and 5-2. Same definition was used for comparative evaluation of developed system with Kinect in section 5.2 and hence will not be repeated there.

**Precision.** It refers to the percentage deviation of measurements from their mean value however it should be noted that precision does not imply accuracy[54]. Higher the percentage deviation lower the measurement precision. It is a measure of *uncertainty* in the measurement. The definition used in this work is:-

$$Precision = \frac{\sum_{p=1}^{vp} \left[ \frac{\sum_{i=1}^{vs_p} \left[ \frac{mean_p - sample_i}{mean_p + 1} \right]}{vs_p} \right]}{vp} \quad (5.1)$$

where  $p$  is the index of pixel number and  $vp$  denotes total *valid* pixels <sup>1</sup> in an image,  $mean_p$ <sup>2</sup> is the mean X/Y/Z value of 3D point corresponding to pixel  $p$ .  $sample_i$  is the corresponding X/Y/Z value of sample  $i$ ,  $vs_p$  is the total valid samples for pixel  $p$ . To avoid divide-by-zero  $mean+1$  has been used.

**Measurement accuracy.** It refers to the percentage deviation of a measurement from its actual value. Higher the deviation lower the accuracy. The definition used in this work is:

---

<sup>1</sup>**CPSS**:Generally thresholding at the edge of strips in the captured binary coded patterns is ambiguous due to a grey region at such regions instead of unambiguous black or white hence these pixels are considered as *invalid*.

**Kinect**:Pixels with NaN's in any of corresponding estimated X, Y or Z coordinates are *invalid*.

<sup>2</sup> $mean_p = \frac{\sum_{i=1}^{vs_p} [sample_i]}{vs_p}$

$$Accuracy = \frac{\sum_{i=1}^N \left[ \frac{Actual_i - measured_i}{Actual_i} \right]}{N} \quad (5.2)$$

where  $N$  is total number of measured lengths,  $Actual_i$  is the actual value for  $i^{th}$  measured length, similarly  $measured_i$  is the corresponding average measured length.

### 5.1.1 Precision

To assess precision(or uncertainty), a planar surface was scammed 10 times for 5 depth levels. For each depth, the percentage average absolute deviation was computed for the 3D-coordinates returned by system corresponding to each pixel with respect to mean of measurements at that pixel, and averaged it over the complete image to get the percentage average absolute deviation at that depth. Table 5-1 shows the variation of % average absolute deviation of X,Y,Z component measurements with depth.

Distance	X%	Y%	Z%
1.3	0.006	0.010	0.007
1.6	0.015	0.019	0.037
1.9	0.010	0.024	0.050
2.2	0.014	0.029	0.094
2.5	0.010	0.025	0.141

Table 5.1: CPSS:Precision along X,Y,Z axes

### 5.1.2 Measurement accuracy

To assess the dependence of measurement accuracy on distance of object from sensor it was decided to determine error in length measurements. Specifically, a planar-face of a box was used with a checkerboard attached to it. Figure 5-1 shows the overall setup for measurement. Using checkerboard detection[55] the image coordinates of 4 outermost corners were determined and 6 3D distances (4 edges and 2 diagonals) were measured using corresponding 3D coordinates in the point cloud. Object was scanned

the 10 times for each depth and the average measured lengths from these scans were calculated. It should be pointed out that the depth range was limited by the depth of field of camera-projector system. Although phase shift technique suffers from non-linearity of projector and camera system and needs to have gamma correction[7], measurements were performed on raw point cloud without any correction/enhancement to get upper bounds on accuracy of the CPSS implementation. Figure 5-2 shows the orientations and distances used in measurement experiments. This figure shows position of each view of measurement object used for evaluating measurement accuracy. Individual views are indexed as  $\#id$  where  $id$  is the number of view. Further, relative position and orientation of developed system and Kinect(described in next section) are shown. The apparent curved path for the positions used for measurement(see figure 5-2) ensured that measurement object remains in common field-of-view of both camera and projector of CPSS system.



Figure 5-1: Experiment setup

Table 5.2 summarizes the observed dependence of measurement accuracy on distance. Whereas in Table 5.3 measurement uncertainty is shown which is the average % relative deviation of 10 measurements with respect to their mean value. It should be noted that measurement uncertainty is a measure of random error as it is calculated

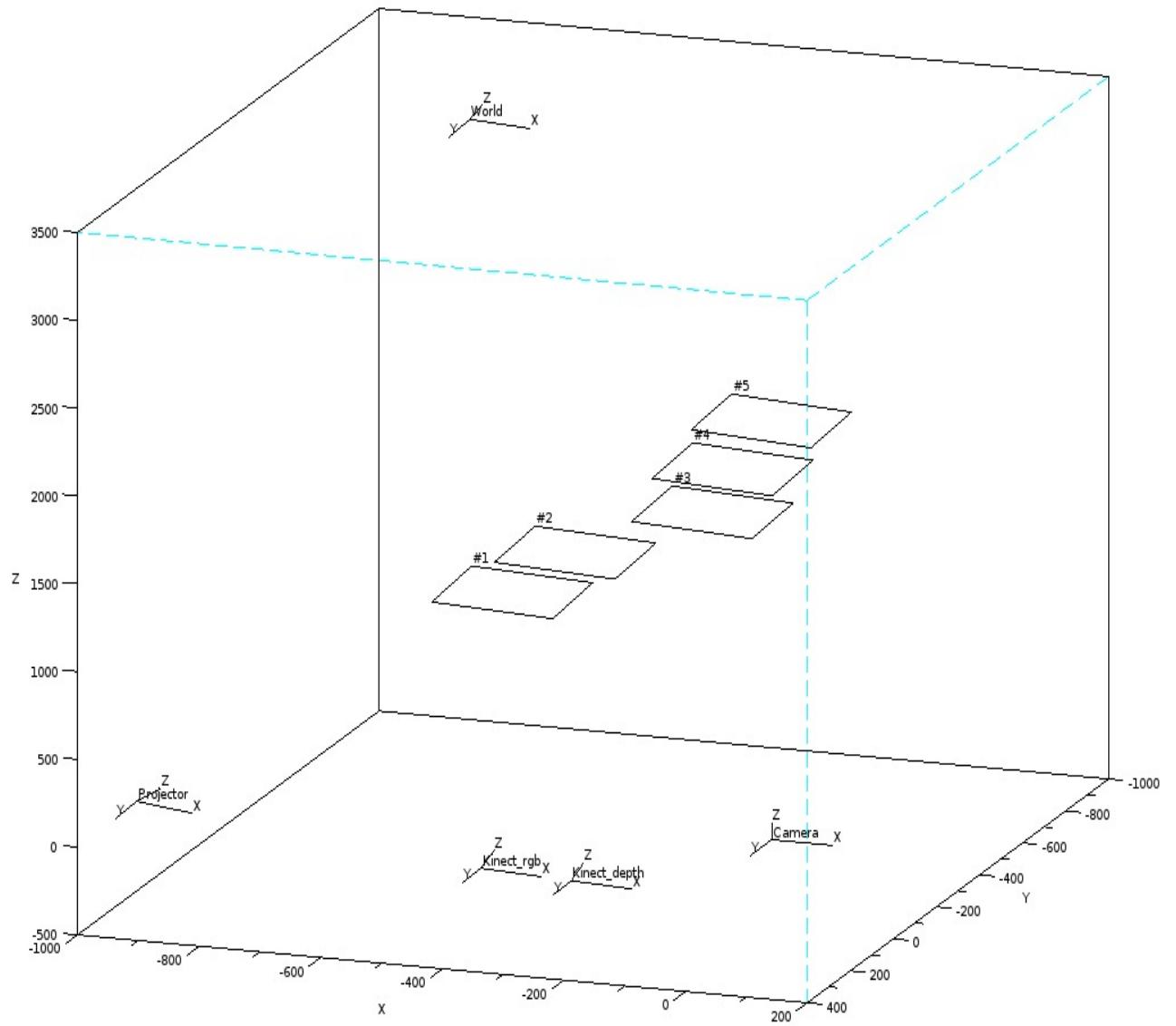


Figure 5-2: Orientations and positions used for experiments

from individual measurements rather than from an averaged dataset.

## 5.2 Comparative evaluation with *Microsoft Kinect*

Phase shift technique has been used in conjunction with binary(or gray) coded technique[56] to take advantage of both phase shift technique in terms of its capability to provide high resolution correspondence information and binary coded technique for providing a robust approach towards establishing correspondence between camera

Distance	CPSS(% error)
1.3	1.153
1.6	0.744
1.9	1.393
2.2	1.040
2.5	0.758

Table 5.2: CPSS:Dependence of measurement accuracy on distance

Distance	CPSS(% uncertainty)
1.3	0.095
1.6	0.152
1.9	0.433
2.2	0.204
2.5	0.134

Table 5.3: CPSS:Dependence of measurement uncertainty on distance

and projector pixels. Kinect on the other hand was initially designed for providing natural user interface[57] but its capability to acquire 3D frames at rates of approximately 30FPS and its object tracking algorithms have gained attention of researchers, enthusiasts from different domains, result has been the tremendous growth of its developer community[58][59] and many applications[60][61] have been built on top of opensource drivers[58][59][62]. Some domains like robotic planning, Rapid prototyping for CFD simulations, dental modelling, medical surgery assistance requires high accuracy of 3D details, to use kinect for such purposes an accuracy analysis should be in place. There has been its comparison with commercial laser scanners[63] but in our knowledge there has been no accuracy comparison of 3D reconstruction using kinect and with the structured light scanners. In the next subsection we will describe brief details of hardware specification and principle of operation of Kinect sensor.

### 5.2.1 Kinect:Principle of operation

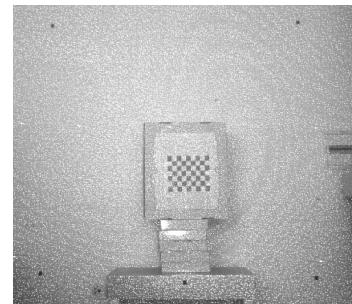
Kinect has an IR projector which is basically an IR LED with diffuser, transparency[64], a RGB camera, an IR camera. RGB camera has native resolution of 1280X1024 [63][65][66], IR camera has a resolution of 1280X1024[67]. Depth(or disparity) im-

age is generated from IR image. Due to bandwidth limitations and to maintain approximately 30FPS 3D datarate IR and RGB images are resized to 640X480 pixel resolution.

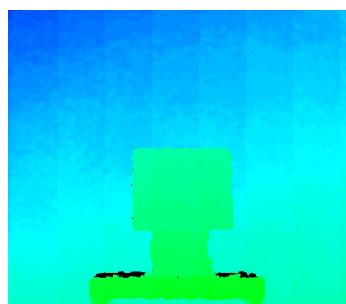
It involves projection of a speckle pattern which is pseudorandom along horizontal axis allowing unique identification of a speckle along every row[64] since IR camera and projector have no relative rotation but translation along X-axis hence deformation of projected pattern due to object surface variations displaces the speckle only in horizontal direction along baseline with respect to a reference image captured with pattern projected at a certain known depth. Hence triangulation involves computing disparity between observed position of a speckle along X-axis and its corresponding position(along X-axis) in reference image. Figure 5-3 shows the images used by kinect for 3D reconstruction.



(a) RGB image



(b) IR image



(c) Disparity map



(d) Point cloud

Figure 5-3: Kinect 3D reconstruction process

### 5.2.2 Experiments

In this section only the results and observations are reported as the experiment procedure and criteria for measurement accuracy and precision are same as described in section 5-1. In table 5.4, precision of the device along X, Y and Z axes is reported.

Distance	X%	Y%	Z%
1.3	0.033	0.011	0.059
1.6	0.018	0.018	0.070
1.9	0.021	0.030	0.095
2.2	0.035	0.046	0.136
2.5	0.050	0.075	0.195

Table 5.4: Kinect:Precision along X,Y,Z axes

For kinect, considerable difference between precision of Z-coordinate as compared to X and Y was observed which might be due to disparity quantization. Figure 5-4 summarizes the results indicating decrease in % deviation with reduction in distance from the sensors, at the same time % deviation in CPSS is lower than that in kinect. Furthermore, it can be observed that precision along Z-axis varies linearly within 1.3m to 2.5 m range and is in agreement with [68].

Table 5-5 shows the dependence of measurement accuracy on distance for Kinect. Table 5-6 shows the dependence of measurement uncertainty on distance of measurement object from kinect sensor.

Distance	Kinect(% error)
1.3	1.194
1.6	1.827
1.9	1.911
2.2	1.737
2.5	1.057

Table 5.5: Kinect:Dependence of measurement accuracy on distance

Although relative error in CPSS and Kinect doesn't give a monotonous variation within the considered depth range but measurement uncertainty in kinect showed

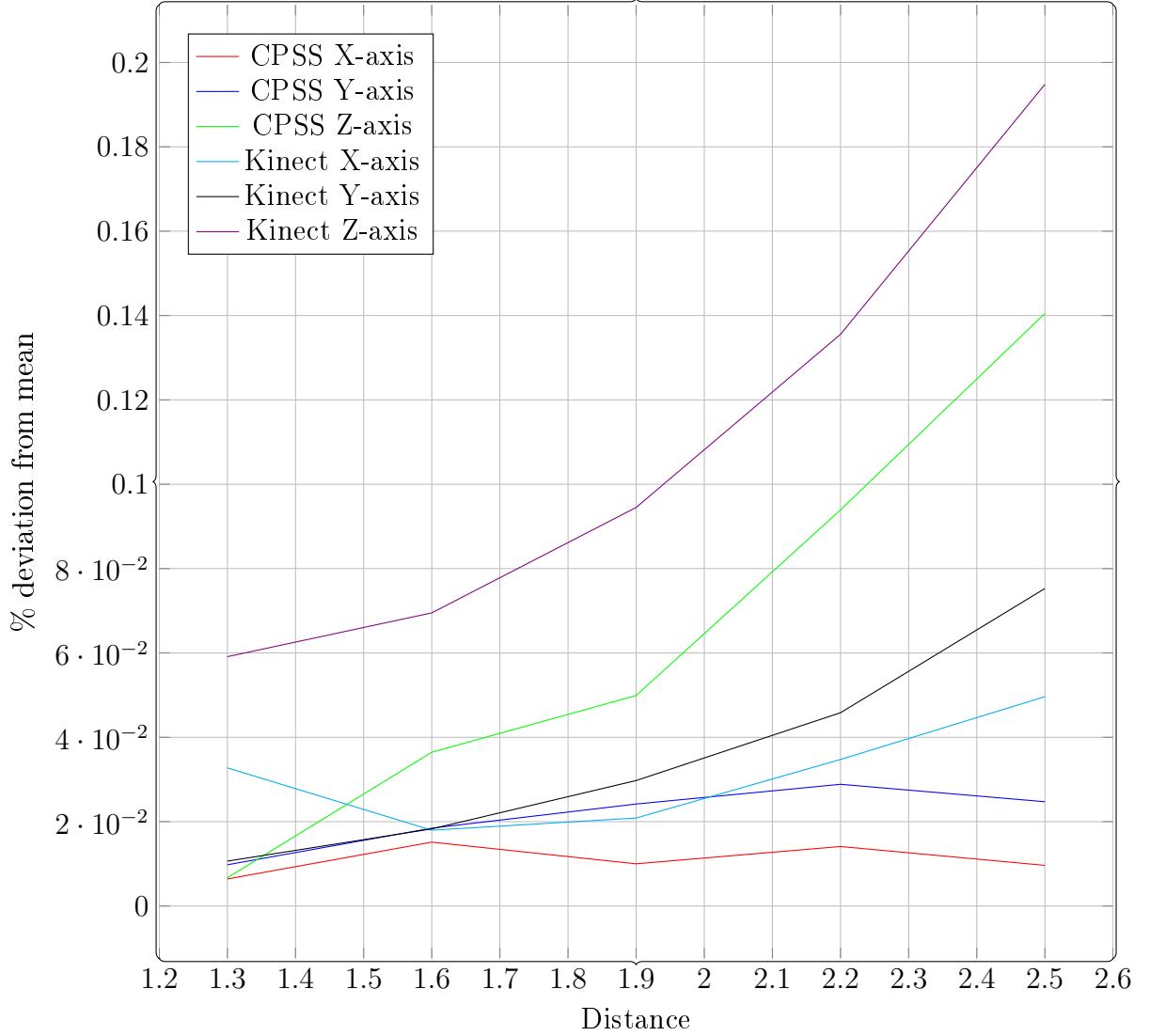


Figure 5-4: Distance versus precision along X,Y,Z axes

strict rise with increasing distance from sensor. We can summarize these observations as average percentage relative error for CPSS was  $\sim 1.018\%$  while that for kinect was  $\sim 1.545\%$  within range of 1.3m-2.5m. Figure 5-5 represents the observed variation of measurement accuracy with distance, along with measurement uncertainty.

### 5.2.3 Observations and conclusions

Kinect accuracy can be improved using system calibration[58][65][67][69]. In our work we have used OpenNI default calibration parameters. Error-prone nature of both

Distance	Kinect(% uncertainty)
1.3	0.069
1.6	0.191
1.9	0.245
2.2	0.261
2.5	0.289

Table 5.6: Kinect:Dependence of measurement uncertainty on distance

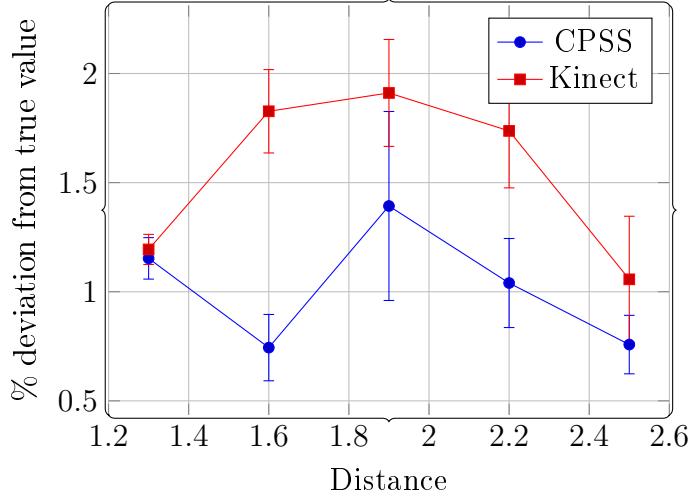


Figure 5-5: Distance versus measurement accuracy with indication of uncertainty(vertical bars)

kinect and CPSS against surfaces with high reflectivity and very low reflectivity was observed. Further it was observed that due to disparity quantization multiple planes are visible in the scan of a single planar object and separation between these planes increases with increase in distance between object and sensor which was also reported by [63]. Here an approach is introduced for accuracy evaluation using 3D length measurements however deviations reported here may not be general statistics as [70] have observed different measurement accuracy for different units of kinect. Depth uncertainty in kinect is considerably higher than CPSS as shown in Tables 1 and 2. Although [71] reports relatively small improvement due to averaging over multiple measurements in range less than 3m, but it was found that the measurement uncertainty over 10 measurements was considerable(as shown in table 5.6) and should not be ignored, hence it suggests that averaging the measurements using multiple scans

is important even within 2.5m.

## 5.3 System calibration

System calibration results in estimation of parameters which allow 3D sensor to map real-world 3D points defined in *physical* units to 2D points defined in terms of *pixels*. In this section, the tests performed to assess repeatability and accuracy of calibration algorithm are described. To assess repeatability, calibration was performed multiple times to determine % average absolute deviation of estimated parameters from their mean values. For assessing accuracy, some *reference* 3D points were selected whose 3D coordinates were known. These points were projected to camera/projector pixel coordinate space and the radial distance of projection from the *observed* projection was considered as *error*. A similar approach is described in [72]. The definition of *system calibration error*  $\varepsilon$  used in this work is described by the following equation:

$$\varepsilon = \frac{\sum_{i=1}^N \sqrt{(X_t^i - X_e^i)^2 + (Y_t^i - Y_e^i)^2}}{N} \quad (5.3)$$

$N$ : Total 3D reference points used for testing system calibration parameters

$(X_t^i, Y_t^i)$ : True coordinates of 2D projection of  $i^{th}$  reference point

$(X_e^i, Y_e^i)$ : 2D coordinates of  $i^{th}$  reference point computed using system calibration parameters

### 5.3.1 Experiments and results

Experiments to evaluate repeatability and accuracy of camera and projector intrinsic and extrinsic calibration parameters were performed. Procedure is already described in the parent section. For sake of brevity, equation 5.2 rephrases the relation used for projecting the  $i^{th}$  reference point from real world 3D coordinates to 2D image coordinates in experiment for measuring accuracy of calibration parameters. Please

refer section 2.1.2 for detailed explanation of this equation.

$$\begin{aligned} \begin{bmatrix} U_e^i \\ V_e^i \end{bmatrix} &= A_c[R|T] \begin{bmatrix} X_w^i \\ Y_w^i \\ Z_w^i \end{bmatrix} \\ \begin{bmatrix} X_e^i \\ Y_e^i \end{bmatrix} &= f(U_e^i, V_e^i) \end{aligned} \quad (5.4)$$

where,

$(X_w^i, Y_w^i, Z_w^i)$ : 3D coordinates of  $i^{th}$  reference point

$(U_e^i, V_e^i)$ : 2D undistorted coordinates of reference point in camera

$A_c$ : Camera intrinsic parameter matrix

$[R|T]$ : Camera extrinsic parameter matrix

$f$ : Camera distortion function estimated by calibration

$(X_e^i, Y_e^i)$ : 2D estimated coordinates of projection of reference point in camera image coordinate system

## 1. Camera calibration

In order to quantify the repeatability of the used OpenCV camera calibration algorithm which is an important measure of stability of any algorithm, we performed camera calibration 5 times and determined the % average absolute deviation of its parameters. Table 5-7 shows the results. Results show consider-

Parameter	% deviation
$f_x$	1.105
$f_y$	1.075
$c_x$	0.707
$c_y$	0.894
$k_1$	7.357
$k_2$	8.825

Table 5.7: Percentage average absolute deviation of camera calibration parameters

ably high uncertainty in estimating  $k_1$  and  $k_2$ . This was observed for projector

calibration as well(see table 5.9). So based on this test we can conclude that algorithm is repeatable within  $\sim 1\%$  for all parameters except distortion coefficients. Although more number of such tests can put more tighter and reliable limits.

Figure 5-6 shows the test scene with reference points which were used for estimating accuracy of calibration parameters. Coordinates of reference points were physically measured and used as *true values*. Here true values were defined using corner-detection to detect image coordinates of reference points. Table 5-8 shows the result of projection of 3D points to camera image-space and their radial distance from corresponding *observed* value. This test was repeatedly performed 5 times to reduce effect of random error in corner detection.



Figure 5-6: Points A,B,C,D were used for measuring camera calibration accuracy

Point	Radial distance from true projection
A	2.553
B	3.453
C	4.080
D	4.180

Table 5.8: Camera calibration:Radial error of projection

In table 5-8, radial error  $\varepsilon$  points towards a systematic error since in all 5 measurements we observed that relation:  $\varepsilon_A < \varepsilon_B < \varepsilon_C < \varepsilon_D$  holds, which demands further study.

## 2. Projector calibration

During development it was observed that calibration parameters returned by OpenCV calibration algorithm for projector were not repeatable hence it was decided to use VPCLib[73] which gives considerably higher repeatability as shown in table 5-9.

VPCLib assumes no distortion parameters in projector model. It estimates projector to 3D-plane homographies for multiple positions of projector with respect to plane. It further uses these homographies to compute points in 3D plane corresponding projector image points. Once these point correspondences are known it uses plane based calibration method for estimating the calibration parameters. Further the estimated parameters are refined using *Bundle adjustment* or non-linear minimization[73].

Parameter	OpenCV method	VPCLib method
$f_x$	7.625	2.166
$f_y$	7.445	2.047
$c_x$	5.559	1.795
$c_y$	4.381	1.861
$k_1$	38.196	Not estimated
$k_2$	96.227	Not estimated

Table 5.9: Percentage average absolute deviation of projector calibration parameters

It can be clearly observed that VPCLib method gives much more repeatable results than OpenCV algorithm for projector calibration. Further it can be observed that estimated  $k_1$  and  $k_2$  exhibit unacceptably high % deviation. Hence VPCLib method was used for further analysis.

As mentioned above, VPCLib method do not assume distortion parameters in projector model since projectors practically have negligible radial distortion

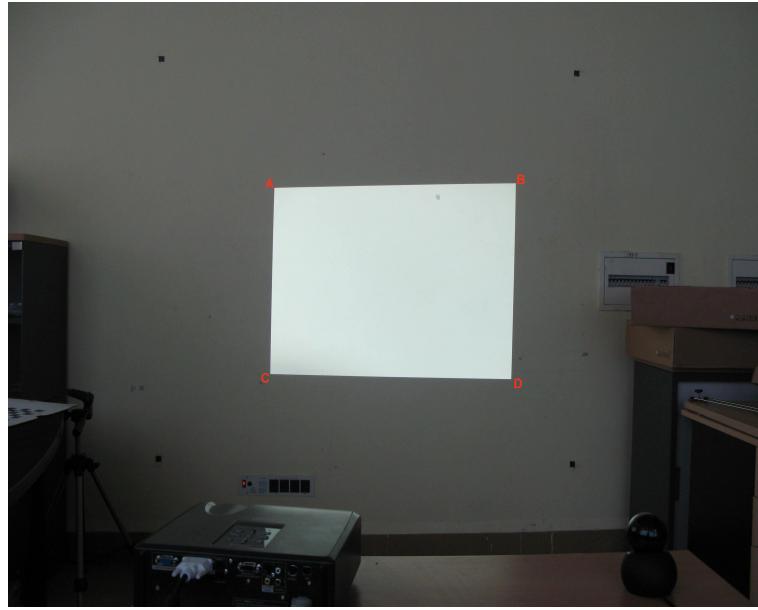


Figure 5-7: Corners A,B,C,D were used for assessing accuracy of projector calibration parameters

hence  $k_1$  and  $k_2$  were not estimated and were assumed to be 0.

To assess accuracy of projector intrinsic and extrinsic parameters estimated by VPCLib, 3D coordinates of corners of projection on a plane were determined through physical measurement and then they were reprojected back to projector image. Figure 5-7 shows the projected blank image whose corners were used for accuracy assessment. Radial distance between *true* image coordinates and coordinates estimated through calibration parameters was considered as *error*. Table 5-10 shows the deviation between observed and estimated 2D projector coordinates. Here again we observe a pattern in radial error in which error is monotonically increasing from point A to point D which is similar to that was observed in case of camera calibration in table 5-8.

Point	Radial distance from true projection
A	6.776
B	8.608
C	11.728
D	18.248

Table 5.10: Projector calibration:Radial error of projection

### 5.3.2 Limitations of the used testing method

This approach requires us to define *true* 3D reference coordinates which is not practical at larger distances from reference origin without using a CMM machine or any other highly accurate metrology equipment. Hence this approach do not allow to practically cover entire *calibration volume* to define accuracy of calibration parameters in general. Hence there is a need of an *automated* method which can efficiently cover entire calibration volume without any requirement of manual intervention to define true values for 3D reference points. Even for exhaustive testing at distance where true 3D coordinates can be conveniently measured it is very time consuming process. Further to define  $(X_t^i, Y_t^i)$  we need to use a *feature detector* which in itself will add corner detection errors.

## 5.4 Stereo correspondence

Stereo-correspondence relates pixels in camera and projector image which corresponds to a common point in real world. To assess accuracy of stereo-correspondence achieved using coded phase shift technique a *known* checkerboard pattern was projected. Corners of checkerboard were detected and corresponding projector coordinates were determined using the stereo-correspondence estimated using coded phase shift technique. Accuracy is defined to be the *radial* distance of estimated projector coordinates for detected checkerboard corners from the true coordinates of the checkerboard corners in projector image. The definition of *stereo-correspondence error* used in this work is described by the following equation:

$$\varepsilon = \frac{\sum_{i=1}^N \sqrt{(X_t^i - X_e^i)^2 + (Y_t^i - Y_e^i)^2}}{N} \quad (5.5)$$

where,

$N$ : Total number of checkerboard corners used for testing stereo-correspondence

$(X_t^i, Y_t^i)$ : True coordinates of  $i^{th}$  checkerboard corner

$(X_e^i, Y_e^i)$ : Estimated coordinates of  $i^{th}$  checkerboard corner

### 5.4.1 Experiments and results

Stereo-correspondence for a phase-shift technique is mainly affected by non-linear output response of projection and capture devices to the applied input[7][42]. This non-linearity can be compensated by a factor called *gamma*. Projector Gamma allows manipulation of contrast of projected image such that non-linearity of input-to-output behavior of projector can be compensated. Furthermore a projected sinusoidal pattern can *appear* non-sinusoidal if contrast of projected image has not been controlled to suppress non-linearity of device. Although same holds for camera also but in this work we have limited our scope to study the effects of projector gamma.

Incorrect value of gamma leads to apparent *waviness* in the 3D reconstructions as shown in figure 5-8 which shows the zoomed view of reconstruction of a planar object for different values of gamma. Here  $\gamma = 2.041$  seems to be more accurate gamma value since planarity of reconstructed surface is higher than that for  $\gamma = 1.0$ . Clearly

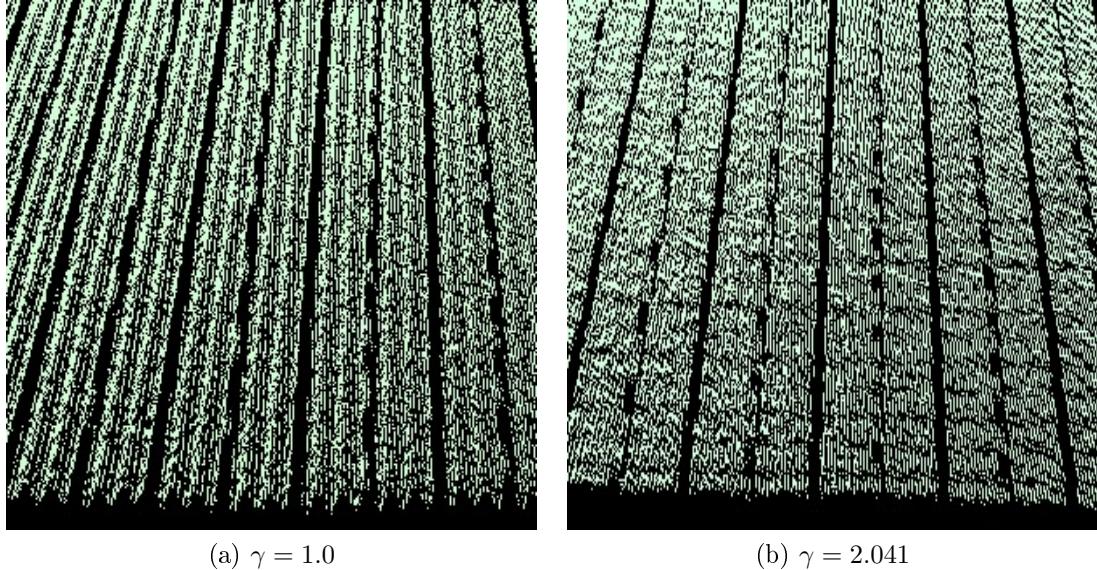
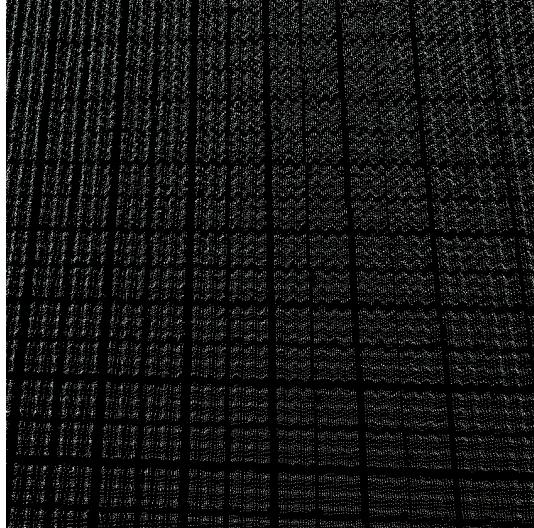


Figure 5-8: Effect of gamma on 3D reconstruction of a plane

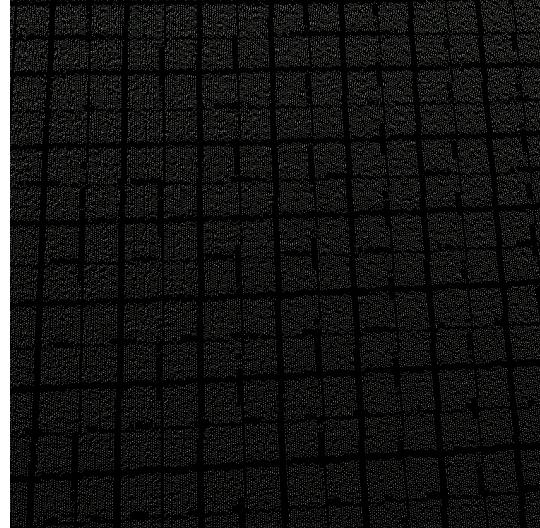
this will affect the measurement accuracy of the system. So the target in this work was to study the effect of projector gamma on stereo-correspondence error.

Further, in literature it has been reported that effect of non-linearity reduces with increasing number of phase shifted patterns used to acquire stereo-correspondence[7].

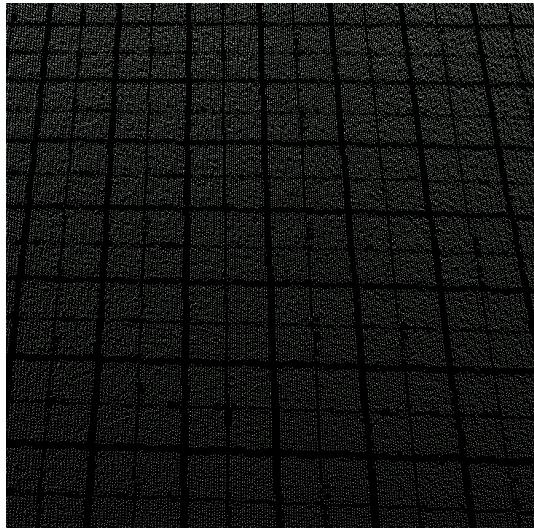
This can be observed in figure 5-9 which shows reduction in *waviness* in zoomed view of reconstruction of a planar object using 4 and 5 phase shifted patterns as compared to 3 phase shifted patterns. Hence, it was further attempted to quantify the effect of number of phase shifted patterns on stereo-correspondence error.



(a) 3 phase



(b) 4 phase



(c) 5 phase

Figure 5-9: Effect of increasing number of phase shifted patterns on *waviness* in 3D reconstruction

Tests as already described in the parent section were performed for both cases i.e., varying gamma for 3 phase shifted pattern approach and by varying the number of

projected phase-shifted fringe patterns under a constant gamma( $\gamma = 1$ ). Table 5.11 shows the effect of projector gamma on stereo-correspondence error. Table 5.12 shows the effect of increasing number of phase-shifted patterns on stereo-correspondence error.

Gamma( $\gamma$ )	Stereo-correspondence error
1.0	1.890
1.524	1.682
1.912	1.730
2.041	1.588
2.559	2.073
2.947	2.394

Table 5.11: Stereo-correspondence error for various gamma values for 3 phase-shifted method

Number of patterns	Stereo-correspondence error
3	1.890
4	2.100
5	2.454

Table 5.12: Stereo correspondence error for 3,4,5 phase shifted patterns

In table 5.11 it can be observed that there is reduction in stereo-correspondence error from  $\gamma = 1.0$  to  $\gamma = 2.041$  followed by increase monotonous increase in error from  $\gamma = 2.559$ . But in table 5.12 we got contradictory results to our expectation. To further confirm this observation we performed experiment 5 times but same pattern of monotonous increase in correspondence error from 3 to 5 phase shifted patterns was observed. This highlights the shortcomings of used criteria for stereo-correspondence error. Although criteria is correct in its formulation because if increasing number of phase-shifted patterns reduces effect of system non-linearity as has been observed in figure 5-9 then it must result in reduction in stereo-correspondence error as defined here. It encourages for further study on how to apply this criteria correctly.

### 5.4.2 Limitations of the used testing method

This approach works as long as projected features can be uniquely recognized so that explicit correspondence estimation for the reference pattern features itself is not required. This can be conveniently used for assessing stereo-correspondence errors in case of a planar scene with no occlusions, since correspondence between camera and projector for feature points in the reference pattern can be easily established. But for complex surfaces with surface discontinuities, curvature, where correspondence between camera and projector for reference pattern cannot be implicitly assumed we either have to use any other technique to define correspondence for the features or an altogether different approach to test stereo-correspondence. Further, accuracy of feature detection will also affect the computed correspondence error.

## 5.5 Triangulation

Triangulation as described in chapter-4 computes 3D coordinates of a real-world point given stereo-correspondence between camera-projector and system calibration parameters. Although it is already mentioned that accuracy evaluation of optical triangulation cannot be done reliably unless we have exact quantitative error bounds for calibration and correspondence for any test case, it is possible to compare different triangulation methods with respect to the accuracy of estimated 3D coordinates of any feature in the scene without changing correspondence and system calibration parameters in the process. This approach do not require us to have knowledge about *exact* error bounds of correspondence and calibration. Hence it is proposed to define *triangulation error* as the deviation of estimated 3D coordinates of a feature in the scene with respect to that reported by CMM machine. Here coordinate measurement is proposed rather than conventional dimension measurement since coordinate measurement indirectly reflects accuracy of dimensions.

## 5.6 Summary

In this chapter, the approaches used for assessing accuracy of stereo-correspondence and system calibration parameters was described along with the limitations of such approaches. Furthermore, a test was proposed to evaluate accuracy of optical triangulation which was not in the scope of work due to time limitations.

In addition, comparison of measurement accuracy and precision of developed system with Microsoft Kinect has been described. Here we examined reconstruction accuracy in static scenes and observed that within depth range of 1.3m-2.5m(based on combine depth-of-field of camera projector system) kinect reconstructs with 0.5 percent higher relative error as compared to CPSS, further the precision of 3D data generated by both techniques using reconstruction of a planar object at various distances was analyzed, and it was observed that for both kinect and CPSS, there is a strong relationship between distance of object from sensor and uncertainty in measurement. Overall CPSS exhibits relatively lower uncertainty in X,Y,Z, coordinates. Furthermore it was observed that CPSS has almost equal uncertainties in X,Y,Z coordinates but kinect exhibits considerably higher uncertainty along Z-axis(depth-uncertainty) as compared to that along X and Y axes.

# Chapter 6

## Conclusions and future works

In this chapter, conclusions from the work performed, original contributions of this work and possible future directions emerging from this work will be discussed. Section 6-1 describes the conclusions and propose future directions from the work done on system calibration and stereo-correspondence. Section 6-2 describes the original contributions of this work.

### 6.1 Conclusions and proposed future works

#### 1. System calibration

Development and accuracy evaluation of system calibration module offered following insights:

- (a) As mentioned in table 5.7 repeatability of camera calibration parameters using OpenCV method was within  $\sim 1\%$  except distortion coefficients. While this may be sufficiently reliable for some applications but for domains like dental transplantations, CFD, inspection of fabricated job where very high accuracy and repeatability of metrology device is expected this range of repeatability may not be sufficient. Recently [11] have proposed an algorithm which they claim to provide appreciably higher accuracy than commonly used OpenCV method. Hence our future work includes

experimenting with this algorithm to evaluate its accuracy to assess its applicability in the domains mentioned above.

- (b) Furthermore, higher values of % deviation for  $k_1$  and  $k_2$  possibly suggest instability of the calibration algorithm in estimating  $k_1$  and  $k_2$ . In this work high quality lenses were used, hence it may be due to *overfitting* leading to estimating coefficients for radial distortion even though lens may not have significant distortion.
- (c) For projector calibration table-5.9 clearly showed that OpenCV based calibration method cannot be reliably used due to unacceptably high values of % deviations for estimated parameters. Instead a study should be done to determine the reason why this algorithm do not reliably works for projector although it works for camera calibration. This question is significant because projector is modeled as *inverse camera*.
- (d) To make the developed system more flexible and mobile, there is need to account for world coordinate system internally or by precalibrating system for physical units like as done in kinect. This will remove the constraint on the system to position world coordinate system within measurement volume which can be inconvenient and cumbersome at times.
- (e) During calibration, subpixel corner detection algorithm was used. It needs multiple initial parameters decision of which considerably changes the position of the detected corner and hence the estimated calibration parameters. But in our knowledge, there has been no thorough study of optimal parameters for certain level of desired accuracy of subpixel corner coordinates or a study relating optimal choice of parameters with structure of image.

## 2. Stereo-correspondence

Work on stereo-correspondence provided practical insight into coded phase-shift technique:

- (a) In practice it was observed that coded phase shift technique is ineffective in highly specular or reflective environments. Furthermore it is not suitable

to be used outdoors. It is primarily because the currently used illumination model for phase shift technique does not take such factors into account.

- (b) Pattern acquisition time is the main hurdle in allowing one to do real-time 3D scanning using this technique. Therefore, there is need to reduce number of projected pattern. This will be part of our future work.
- (c) It was observed that non-sinusoidal nature of captured patterns depends upon camera gamma, environmental lighting, camera resolution, perspective of camera with respect to region in the scene where pattern is projected etc., in addition to projector gamma. Hence there is a need of a study to account for all these factors besides projector gamma.
- (d) Further a problem of locating edge of projected binary coded patterns in camera image was encountered. Strip edge region is ambiguous for decoding binary codewords since it typically has grey intensity hence it cannot be classified as black or white only. In this work, pixels corresponding to such regions were discarded and were considered as invalid. This has raised the need for an algorithm which can atleast approximately determine the position of strip edge.
- (e) It was initially attempted to develop system solely based on phase-shift method without any supporting technique like binary coded patterns but unacceptably poor time efficiency of the algorithm used for phase-unwrapping was observed hence it was decided to changed it to combination of binary coded and phase shift methods. It reduces computational complexity of phase-unwrapping but increases pattern acquisition time. To reduce this time and problems of strip edge detection with binary coded patterns, it is proposed to develop a time efficient phase unwrapping algorithm which will allow us to solve stereo-correspondence using only phase-shift approach. This will considerably reduce the required number of patterns.
- (f) During experiments with stereo-calibration module it was observed that the proposed *stereo-correspondence error* criterion gives different error es-

timates under different lighting conditions. Hence it has been proposed to study the effect of various radiometric conditions on the behavior of this criterion. And it is further speculated that contradictory results obtained in table 5.12 require accommodation of radiometric conditions into account.

### 3. 3D reconstruction and triangulation

It is possible to improve the time efficiency of system by replacing ray-ray triangulation which requires projection of both horizontal and vertical patterns as described in chapter-3 to ray-plane triangulation which calculates 3D coordinates of a point using intersection of an optical ray from camera and corresponding plane from projector or vice versa. Ray-plane triangulation will require projection of only vertical or horizontal patterns thereby reducing the required number of patterns by half.

Furthermore, for our application of dental imaging, CFD and inspection of fabricated jobs it is required to have complete 3 dimensional view of the object. For this purpose it is proposed to extend the developed system such that it allows user to take complete 360 degree view of the object.

## 6.2 Contributions of this work

This work is focused on development and accuracy analysis of a 3D scanner, where in both development and accuracy evaluation some novelties are observed:

1. Developed system is first *in-house* 3D measurement system based on Coded phase-shift technique which is currently among the most accurate 3D metrology approaches based on active vision.
2. The measurement accuracy and precision of developed system was experimentally studied and compared with Microsoft Kinect, this work can be of immediate

ate use to researchers and practitioners planning to use kinect for accuracy and reliability critical tasks.

3. An approach to assess *measurement accuracy* was proposed and experimentally demonstrated.
4. A *Stereo-correspondence error* criteria has been proposed and experimentally demonstrated which will allow us to evaluate accuracy of stereo-correspondence.

# Bibliography

- [1] Microsoft Photosynth.  
[photosynth.net](http://photosynth.net)  
Accessed 3 December 2012
- [2] Projective geometry tutorial.  
<http://www.cse.iitd.ernet.in/~suban/vision/tutorial/node1.html>  
Accessed 4 December 2012
- [3] Faugeras,Olivier.(1995).Stratification of 3-D vision: projective, affine, and metric representations,Journal of the Optical Society of America A,Vol-12,46548-4
- [4] Hartley,Richard I.,Sturm,Peter(1997).Triangulation, Computer Vision and Image Understanding,Volume 68, Issue 2, November 1997,Pages 146-157
- [5] Saxena,Ashutosh,Sun,Min,Ng,Andrew Y.(2007).3-D Reconstruction from Sparse Views using Monocular Vision,ICCV workshop on Virtual Representations and Modeling of Large-scale environments (VRML),2007
- [6] Flexiscan3D scanner software  
<http://www.3d3solutions.com/products/flexscan3d/>  
Accessed:3 December 2012
- [7] Pan,B., Kemaq,Q.,Huang,L.,and Asundi,A.(2009).Phase error analysis and compensation for nonsinusoidal waveforms in phase-shifting digital fringe projection profilometry, Opt. Lett. 34, 416-418.
- [8] Tsai,Roger Y. (1987) :A Versatile Camera Calibration Technique for High- Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses,IEEE Journal of Robotics and Automation, Vol. RA-3, No. 4, August 1987, pp. 323-344.
- [9] Zhang,Song(2005).High-resolution, Real-time 3-D Shape Measurement,Ph.D. dissertation, Stony Brook University, Stony Brook, NY, 2005.

- [10] Zhang,Z.(2000).A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000
- [11] Vo,Minh,Wang,Zhaoyang,Pan,Bing,Pan,Tongyan(2012).Hyper-accurate flexible calibration technique for fringe-projection-based three-dimensional imaging,Opt. Express 20, 16926-16941 (2012)
- [12] Hartley,R.I.(1994),An algorithm for self calibration from several viewes.In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition,pages 908-912,IEEE Computer Society Press
- [13] Luong,Q. T,Faugeras,O.(1997).Self-calibration of a moving camera from point correspondences and fundamental matrices.The International Journal of Computer Vision,22(3):261-289,1997
- [14] Caprile,B,Torre,V.(1990).Using Vanishing Points for Camera Calibration.The International Journal Of Computer Vision,4(2):127-140,Mar.1990
- [15] Hartley,R.(1994).Self-calibration from multiple views with a rotating camera.In Proceedings of 3rd ECCV,volume 800-801 of 'Lecture notes in Computer Science' pages 471-478,Stockholm,Sweden,May 1994.Springer-Verlag
- [16] Zollner, Helmut and Sablatnig,Robert(2004).Comparison of Methods for Geometric Camera Calibration using Planar Calibration Targets,Proceedings of the 28th Workshop of the Austrian Association of Pattern Recognition, AAPR 2004,pp.237–244,OCG Schriftenreihe, Österreichische Arbeitsgemeinschaft für Mustererkennung
- [17] Heikkilä,J.,Silven,O.(1997).A Four-Step Camera Calibration Procedure with Implicit Image Correction. In Proc. of IEEE Computer Vision and Pattern Recognition, pp. 1106-1112, 1997
- [18] Camera Calibration Toolbox For Matlab  
[http://www.vision.caltech.edu/bouguetj/calib\\_doc/index.html](http://www.vision.caltech.edu/bouguetj/calib_doc/index.html)  
Accessed:3 December 2012
- [19] OpenCV documentation on Camera calibration  
<http://docs.opencv.org/modules/calib3d/doc/calib3d.html>  
Accessed 5 December 2012
- [20] Levenberg Marquardt algorithm.  
[http://en.wikipedia.org/wiki/Levenberg%E2%80%93Marquardt\\_algorithm](http://en.wikipedia.org/wiki/Levenberg%E2%80%93Marquardt_algorithm)  
Accessed 4 December 2012.

- [21] Extrinsic parameters estimation in OpenCV  
[http://opencv.willowgarage.com/documentation/camera\\_calibration\\_and\\_3d\\_reconstruction.html#findextrinsiccameraparams2](http://opencv.willowgarage.com/documentation/camera_calibration_and_3d_reconstruction.html#findextrinsiccameraparams2)  
Accessed 5 December 2012
- [22] Bradski,Gary and Kaehler,Adrian(2008).Learning OpenCV,Computer Vision with the OpenCV Library.O'Reilly Media,September 2008
- [23] Asla Medeiros e Sa,Esdras Soares de Medeiros Filho,Paulo Cezar Carvalho,Luiz Velho.Coded Structured Light for 3D-photography:an Overview.  
[www.visgraf.impa.br/Data/RefBib/PS\\_PDF/rita-survey/survey.pdf](http://www.visgraf.impa.br/Data/RefBib/PS_PDF/rita-survey/survey.pdf)  
Accessed 5 December 2012
- [24] Salvi,J.,Pagès,J,Tutorial on Coded Light Projection techniques  
[http://eia.udg.es/~qsalvi/Tutorial\\_Coded\\_Light\\_Projection\\_Techniques\\_arxivos/v3\\_document.html](http://eia.udg.es/~qsalvi/Tutorial_Coded_Light_Projection_Techniques_arxivos/v3_document.html)  
Accessed 4 December 2012
- [25] Geng,J.(2011)."Structured-light 3D surface imaging: a tutorial," Adv. Opt. Photon.3, 128-160 .
- [26] Pagès,Jordi,Salvi,Joaquim,Garcia,Rafael,Matabosch,Carles(2003).Overview of coded light projection techniques for automatic 3D profiling
- [27] Posdamer,J. L.,Altschuler,M. D.(1982). Surface measurement by space-encoded projected beam systems, Comput. Graph. Image Processing 18(1), 1-17 (1982).
- [28] Inokuchi,S.,Sato,K.,Matsuda,F.(1984).Range-imaging for 3-D object recognition, in International Conference on Pattern Recognition (International Association for Pattern Recognition, 1984), pp. 806-808.
- [29] Caspi,D.,Kiryati,N.,Shamir,J.(1998),Range imaging with adaptive color structured light, IEEE Trans. Pattern Anal. Mach. Intell. 20(5), 470-480 (May 1998).
- [30] Horn,Eli,Kiryati,Nahum(1997).Toward Optimal Structured Light Patterns,Image and Vision Computing,Vol. 17,pp.87-97
- [31] Ghiglia,Dennis C.,Pritt,Mark D..Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software.Wiley Publications,ISBN: 978-0-471-24935-1
- [32] MacWilliams,F.J.,Sloane,N.J.A.(1976).Pseudorandom sequences and arrays, Proc. IEEE 64(12), 1715-1729.

- [33] De Bruijn sequences  
[http://feed-back.be/nick/?page\\_id=310](http://feed-back.be/nick/?page_id=310)  
Accessed 12 December 2012
- [34] Moigne,J.Le,Waxman,A. M.(1985), Multi-resolution grid patterns for building range maps,in Vision-85, Applied Machine Vision Conference (ASME) (Society of Manufacturing Engineers, 1985), pp. 22-39.
- [35] Ulusoy,A. Osman,Calakli,F.,Taubin,G.(2009), One-shot scanning using De Bruijn spaced grids, in 2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops) (IEEE, 2009), pp. 1786-1792.
- [36] Carrihill,B.,Hummel,R.(1985). Experiments with the intensity ratio depth sensor. In Computer Vision, Graphics and Image Processing, volume 32, pages 337–358. Academic Press, 1985
- [37] Chazan,G.,Kiryati,N.(1995). Pyramidal intensity ratio depth sensor. Technical report 121, Center for Communication and Information Technologies, Department of Electrical Engineering, Technion, Haifa, Israel, October 1995.
- [38] Hung,D.C.D.(1993). 3d scene modelling by sinusoid encoded illumination. Image and Vision Computing, 11:251–256, 1993.
- [39] Tajima,J.,Iwakawa,M.(1990). 3D data acquisition by rainbow range finder. In International Conference on Pattern Recognition, pages 309–313, 1990
- [40] Sato,T.(1999). Multispectral pattern projection range finder. In Proceedings of the Conference on ThreeDimensional Image Capture and Applications II, volume 3640, pages 28–37, San Jose, California, January 1999. SPIE.
- [41] Stripe boundary codes for real-time structured-light range scanning of moving objects. In The 8th IEEE International Conference on Computer Vision, pages II: 359-366, 2001.
- [42] Zhang,Song(2010).Recent progresses on real-time 3D shape measurement using digital fringe projection techniques.Optics and Lasers in Engineering, Volume 48, Issue 2, February 2010, Pages 149-158
- [43] Gorthi,S.S.,Rastogi,P(2009).Fringe projection technique:Whither we are?.Optics and Lasers in Engineering, 2009
- [44] ‘open-light’-Google code  
<http://code.google.com/p/open-light/>  
Accessed 5 December 2012

- [45] 'Structured-light'-Goole code  
<http://code.google.com/p/structured-light/>  
Accessed 5 December 2012
- [46] Gupta,Mohit and Agrawal, Amit and Veeraraghavan, Ashok and Narasimhan, Srinivasa,G.(2012).A Practical Approach to 3D Scanning in the Presence of Interreflections,Subsurface Scattering and Defocus.International Journal of Computer Vision,pp.1-23,Springer
- [47] Ghiglia,Dennis C.,Romero,Louis A.(1996)."Minimum L<sub>p</sub>-norm two-dimensional phase unwrapping," J. Opt. Soc. Am. A 13, 1999-2013 (1996)
- [48] Mundy,J.,Zisserman,A.(1992).Appendix:Projective geometry for machine vision-Geometric invariances in computer vision,MIT Press,1992
- [49] R.I.,Hartley,Zisserman,A.(2004).Multiple View Geometry in Computer Vision,Second edition,Cambridge University Press, ISBN:0521540518
- [50] Hartley,R.,Gupta,R.,Chang,T.(1992).Stereo from uncalibrated cameras, in Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, 1992, pp. 761-764.
- [51] Triangulation  
<http://en.wikipedia.org/wiki/Triangulation>  
Accessed 5 December 2012
- [52] Brett R.,Jones(2010).AUGMENTING COMPLEX SURFACES WITH PROJECTOR-CAMERA SYSTEMS,M.S.Thesis,University of Illinois at Urbana-Champaign
- [53] Point cloud library.  
[pointclouds.org/](http://pointclouds.org/).  
Accessed 24 November 2012
- [54] Accuracy & precision.  
[http://en.wikipedia.org/wiki/Accuracy\\_and\\_precision](http://en.wikipedia.org/wiki/Accuracy_and_precision)  
Accessed 24 November 2012
- [55] OpenCV function for checkerboard corner detection.  
[http://opencv.willowgarage.com/documentation/camera\\_calibration\\_and\\_3d\\_reconstruction.html#findchessboardcorners](http://opencv.willowgarage.com/documentation/camera_calibration_and_3d_reconstruction.html#findchessboardcorners)  
Accessed 24 November 2012
- [56] Sansoni,G., Carocci,M., and Rodella,R.(1999). "Three-Dimensional Vision Based on a Combination of Gray-Code and Phase-Shift Light Projection: Analysis and Compensation of the Systematic Errors," Appl. Opt. 38, 6565-6573 .

- [57] Microsoft Kinect.  
<http://www.xbox.com/en-US/KINECT>.  
Accessed 24 November 2012
- [58] OpenNi home page.  
[openni.org/](http://openni.org/).Accessed 24 November 2012
- [59] OpenKinect home page.  
[openkinect.org/](http://openkinect.org/).Accessed 24 November 2012
- [60] RGBDemo program for accessing kinect data.  
[labs.manctl.com/rgbdemo/](http://labs.manctl.com/rgbdemo/)  
Accessed 24 November 2012
- [61] Skanect for Windows & Mac-plateform.  
<http://manctl.com/products.html>.  
Accessed 24 November 2012
- [62] Code laboratory.  
<http://codelaboratories.com/nui/>.  
Accessed 24 November 2012
- [63] Khoshelham, K., Elberink, S.O.(2012). Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications. *Sensors*, 12, 1437-1454.
- [64] Freedman, B., Shpunt, A., Machline, M., & Arieli, Y. (2010). Patent No. US2010/0018123 A1. United States of America.
- [65] Smisek, J., Jancosek, M., Pajdla, T., (2011), "3D with Kinect," Computer Vision Workshops (ICCV Workshops),pp.1154-1160, 6-13 Nov. 2011 doi: 10.1109/ICCVW.2011.6130380
- [66] ifixit Kinect teardown.  
<http://www.ifixit.com/Teardown/Microsoft+Kinect+Teardown/4066/1>.  
Accessed 24 November 2012
- [67] ROS kinect calibration page.  
[http://www.ros.org/wiki/openni\\_launch/Tutorials/IntrinsicCalibration?action=show&redirect=openni\\_camera%2Fcalibration](http://www.ros.org/wiki/openni_launch/Tutorials/IntrinsicCalibration?action=show&redirect=openni_camera%2Fcalibration).  
Accessed 24 November 2012
- [68] Molnàr,B., Toth,C.K., and Detreköi,A.(2012). ACCURACY TEST OF MICROSOFT KINECT FOR HUMAN MORPHOLOGIC MEASUREMENTS, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXIX-B3, 543-547, doi:10.5194/isprsarchives-XXXIX-B3-543-2012
- [69] Herrera C.,D.,Kannala, J.,Heikkilä,J.(2012).Joint depth and color camera calibration with distortion correction, *TPAMI*,

- [70] Boehm,J.(2012).NATURAL USER INTERFACE SENSORS FOR HUMAN BODY MEASUREMENT, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XXXIX-B3, 531-536, doi:10.5194/isprsarchives-XXXIX-B3-531-2012
- [71] Chow,J.C.K.,Ang,K.D.,Lichti,D.D.,and Teskey,W.F.(2012). PERFORMANCE ANALYSIS OF A LOW-COST TRIANGULATION-BASED 3D CAMERA: MICROSOFT KINECT SYSTEM International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXIX-B5 XXII ISPRS Congress
- [72] Salvi,Joaquim,Armanguè,Xavier,Batlle,Joan(2002).A comparative review of camera calibrating methods with accuracy evaluation, Pattern Recognition, Volume 35, Issue 7, July 2002, Pages 1617-1635, ISSN 0031-3203, 10.1016/S0031-3203(01)00126-1.
- [73] Drarèni, J. and Roy, S. and Sturm, P.(2012).Methods for Geometrical Video Projector Calibration,Machine Vision and Applications