Development of a calibration pipeline for a monocular-view structured illumination 3D sensor utilizing an array projector

Dissertation

For the acquisition of the academic title

Doctor rerum naturalium (Dr. rer. nat.)

Submitted to the Council of the Faculty of Physics and Astronomy of the Friedrich-Schiller-Universität Jena

FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

Written by

M.Sc. Eugene Wong

Gutachter:

1. Prof. Dr. rer. nat. Richard Kowarschik
   Institut für Angewandte Optik und Biophysik
   Friedrich-Schiller-Universität Jena

2. Prof. Dr. -Ing. Rainer Tutsch
   Institut für Produktionsmesstechnik
   Technische Universität Braunschweig

3. Prof. Dr. -Ing. Peter Lehmann
   Fachgebiet Messtechnik
   Universität Kassel

Tag der Disputation: 2 December, 2021

List of Acronyms

2D/3D	Two/Three-Dimensional
AD	Absolute Deviation
AOD	Acousto-Optical Deflector
BLP	Band-Limited Pattern
BRDF	Bidirectional Reflectance Distribution Function
CAD/CAM	Computer-Aided-Design/Manufacturing
CCD	Charge-Coupled Device
CMOS	Complementary Metal–Oxide–Semiconductor
CNN	Convolutional Neural Network
DLP	Digital Light Processing
DMD	Digital Mirror Device
EAD	Ellipsoidal Absolute Deviation
FFD	Fiber Fixed to Diffuser
FPP	Fringe Projection Profilometry
FTP	Fourier Transform Profilometry
GD&T	General Dimensioning and Tolerancing
GOBO	Goes Before Optics
GPU	Graphics Processing Unit
GM	Galvanometric Mirror
HMI	Human-Machine Interaction
LED	Light-Emitting-Diode
LCoS	Liquid Crystal on Silicon
MAAP	Multi-Aperture-Array Projector
MLP	Multilayer Perceptron

MV	Measurement Volume
PSP	Phase Shifting Profilometry
PBR	Physically-Based Rendering
SfM	Structure from Motion
SVD	Singular Value Decomposition
TFT	Thin-Film-Transistor
ZC	Zhang Calibration

List of Symbols

a	Vertical intensity shift of sine pattern
$a(X,Y)$	Background object image intensity at spatial coordinate (X,Y) in Fringe Projection Profilometry
$\vec{a}, \vec{a}_P$	Arbitrary vector from origin to arbitrary 3D point on a plane in the camera coordinate system and plane coordinate system respectively
$\mathcal{A}_n$	Fourier coefficient of n^{th} harmonic in Fourier Transform Profilometry
A_n	n^{th} angle of pyramidal frustum gauge
b	Intensity amplitude of sine pattern
$b(X,Y)$, $b_0(X,Y)$	Image intensity modulation due to non-uniform reflectivity of object surface and reference plane at spatial coordinate (X,Y) in Fringe Projection Profilometry respectively
B	Baseline distance
c	Wavenumber of sine pattern
C_j, C'_j	Gray-value intensity of camera pixels in stereo setup for each pattern j
$\bar{C}$, $\bar{C'}$	Arithmetic mean of gray-value intensity sequence of camera pixels in a stereo setup
C^*	Arbitrary measurement camera in proposed intrinsic 're-calibration'
d_S	Distance of each reconstructed point of sphere to best-fit center
D	Corresponding test plane increment
$\overrightarrow{D_P}$	Vector from first test plane increment to corresponding test plane increment in the plane coordinate system

f	Focal length
f_0	Fundamental carrier frequency of sine pattern projected in Fourier Transform Profilometry
F	Fundamental matrix
H	Distance between camera pinhole and reference plane in Fourier Projection Profilometry
i	Imaginary number
I	Digitized intensity bit-value of pixel in image or projection chip
$\tilde{I}$	Image intensity value due to sine pattern projected onto the object, filtering only for the spectrum about f_0 in Fourier Transform Profilometry
$\widetilde{I_0}$	Image intensity value due to sine pattern projected onto the reference plane, filtering only for the spectrum about f_0 in Fourier Transform Profilometry
k_n	n^{th} radial distortion coefficient
l	Epipolar line
L_n	n^{th} line segment of pyramidal frustum gauge
M	3×3 intrinsic camera parameter matrix
$\vec{n}, \vec{n}_P$	Normal vector of a plane in the camera/world coordinate system and plane coordinate system respectively
o	Origin of a coordinate system
O_j	Gray-value intensity of an image pixel when imaging object
$\bar{O}$	Arithmetic mean of gray-value intensity sequence of image pixel when imaging object
p_n	n^{th} tangential distortion coefficient

P_j	Gray-value intensity of an image pixel when imaging test plane increment
$\bar{P}$	Arithmetic mean of gray-value intensity sequence of image pixel when imaging test plane increment
r	Radial distance
R	3×3 rotation matrix
$\mathbb{R}+$	Set of positive real numbers
$S_{C,W}$, S_P	Camera/world coordinate system and plane coordinate system respectively
S_n	n^{th} planar surface of pyramidal frustum gauge
$\vec{t}$	Translation vector
$\overrightarrow{t_{0P}}$	Translation vector from origin to first test plane increment in plane coordinate system
$\vec{v}$	Directional vector of light ray
$\hat{x}, x_{\perp}, x$	Unknown ground truth image plane projection, corrected image plane projection, and observed image plane projection respectively
$Z(X,Y)$	Height of point on object surface at spatial coordinate (X,Y) in Fringe Projection Profilometry
(α,β), (α_d,β_d)	Image point in pixel units in the image plane coordinate system and geometrically distorted image point respectively
(u,v)	Principal point of image plane
(x_{im}, y_{im})	Image point in meter units in the image plane coordinate system
(x^r_{im}, y^r_{im}), (x^t_{im}, y^t_{im})	Radially and tangentially distorted component of image point in meter units in the image plane coordinate system respectively

(x_n, y_n)	Normalized image coordinates
(X,Y,Z)	3D spatial coordinates
Δl	Length error absolute difference between measured length and ground truth/reference
Δp	Pixel pitch
Δ°	Angle error absolute difference between measured length and ground truth or reference point cloud
δ	Distance error from Epipolar line
ɜ	Contains the set
ϵ	Belongs to the set
γ	Scalar multiplied with directional vector of light ray $\vec{v}$ to obtain 3D point on a plane
Λ, Λ_{im}	Length in 3D space and its projection onto the image plane
η	Skew coefficient of image coordinate system axis
φ(X,Y), $\varphi_0(X,Y)$, *Δφ(X,Y)*	Spatially modulated phase of sine pattern incident on object due to surface height distribution, on reference plane, and phase difference between them at spatial coordinate (X,Y) in Fringe Projection Profilometry respectively
Π	Epipolar plane
Φ	Phase shift of sine pattern in Phase-Shifted Profilometry
ρ	Cross-correlation coefficient
σ	Spread of pointwise deviation from best-fit or reference
χ,ψ,Ω	Ellipsoid distances along each ellipsoid axis to the ellipsoid center

$x_a = (x_{a1}, x_{a2}, \dots, x_{a\mathbb{D}})$	Position of a^{th} particle in the $\mathbb{D}$ dimensional search domain in the particle swarm optimization method
$v_a = (v_{a1}, v_{a2}, \dots, v_{a\mathbb{D}})$	Velocity of a^{th} particle in the $\mathbb{D}$ dimensional search domain in the particle swarm optimization method
$p_a^i = (p_{a1}^i, p_{a2}^i, \dots, p_{a\mathbb{D}}^i)$	For an individual a^{th} particle, the position in the $\mathbb{D}$ dimensional search domain that resulted in the best objective function loss achieved so far during the iterative process out of the whole group of particles in the particle swarm optimization method
$p_{a^*d}^g = (p_{a^*1}^g, p_{a^*2}^g, \dots, p_{a^*\mathbb{D}}^g)$	For the group of particles, the position in the $\mathbb{D}$ dimensional search domain that resulted in the best objective function loss achieved so far during the iterative process, where a^* is the specific particle that achieved the best objective function loss so far in the particle swarm optimization method
$r_{1_{ad}}, r_{2_{ad}}$	Random numbers drawn from a uniform distribution between [0,1] for the particle swarm optimization method.
c_1, c_2, w	Constants affecting convergence of the particle swarm optimization method

Abstract

Commercial off-the-shelf digital projection systems are commonly used in active structured illumination photogrammetry of macro-scale surfaces due to their relatively low cost, accessibility, and ease of use. They can be described as inverse pinhole modelled. The calibration pipeline of a 3D sensor utilizing pinhole devices in a projector-camera setup configuration is already well-established. Recently, there have been advances in creating projection systems offering projection speeds greater than that available from conventional off-the-shelf digital projectors. However, they cannot be calibrated using well established techniques based on the pinole assumption. They are chip-less and without projection lens. This work is based on the utilization of unconventional projection systems known as “array projectors” which contain not one but multiple projection channels that project a temporal sequence of illumination patterns. None of the channels implement a digital projection chip or a projection lens. To workaround the calibration problem, previous realizations of a 3D sensor based on an array projector required a stereo-camera setup. Triangulation took place between the two pinhole modelled cameras instead. However, a monocular setup is desired as a single camera configuration results in decreased cost, weight, and form-factor. This study presents a novel calibration pipeline that realizes a single camera setup. A generalized intrinsic calibration process without model assumptions was developed that directly samples the illumination frustum of each array projection channel. An extrinsic calibration process was then created that determines the pose of the single camera through a downhill simplex optimization initialized by particle swarm. Lastly, a method to store the intrinsic calibration with the aid of an easily realizable calibration jig was developed for re-use in arbitrary measurement camera positions so that intrinsic calibration does not have to be repeated. The only caveats of the calibration pipeline are that the projected patterns must be temporally repeatable and that the pose between the array projector and calibration jig remains fixed. The success of the developed calibration pipeline is demonstrated experimentally. 3D reconstructions using only a single camera are accomplished of white scattering geometric gauge surfaces as well as complex shapes, and the 3D measurement performance compared to reconstructions generated by a stereo-camera setup.

Zusammenfassung

Kommerzielle, handelsübliche digitale Projektionssysteme werden aufgrund ihrer relativ geringen Kosten, ihrer Verfügbarkeit und ihrer einfachen Handhabung häufig in der aktiven Photogrammetrie mit strukturierter Beleuchtung von makroskopischen Oberflächen eingesetzt. Sie können dazu durch das inverse Lochkameramodell beschrieben werden. Die Kalibrierungspipeline eines 3D-Sensors, der Lochkamerakomponenten in einer Projektor-Kamera-Konfiguration verwendet, ist bereits gut bekannt. In jüngster Zeit gab es Fortschritte bei der Entwicklung von Projektionssystemen, die eine höhere Projektionsgeschwindigkeit als herkömmliche digitale Projektoren bieten. Sie können jedoch nicht mit etablierten Techniken kalibriert werden, die auf dem Lochkameramodell basieren, da sie weder einen (Mikro-)Chip noch eine Projektionslinse besitzen. Diese Arbeit basiert auf der Verwendung von unkonventionellen Projektionssystemen, die als "Array-Projektoren" bezeichnet werden und nicht nur einen, sondern mehrere Projektionskanäle enthalten, die eine zeitliche Abfolge von Beleuchtungsmustern projizieren. Keiner der Kanäle enthält einen digitalen Projektions-Chip oder eine einzelne Projektionslinse. Um das Kalibrierungsproblem zu umgehen, erforderten frühere Realisierungen eines 3D-Sensor, der auf einem Array-Projektor basiert, eine Stereokamera-Anordnung. Die Triangulation fand dabei zwischen den beiden Kameras auf der Basis eines Lochkameramodells statt. Ein monokularer Aufbau ist jedoch erwünscht, da eine Ein-Kamera-Konfiguration zu geringeren Kosten, Gewicht und Baugröße führt. In dieser Dissertation wird eine neuartige Kalibrierungspipeline vorgestellt, die eine Ein-Kamera-Konfiguration realisiert. Es wurde ein verallgemeinerter intrinsischer Kalibrierungsprozess ohne Modellannahmen entwickelt, der direkt den Kegelstumpf eines jeden Array-Projektionskanals abtastet. Anschließend wurde ein extrinsischer Kalibrierungsprozess erstellt, der die Position der Einzelkamera durch ein abwärtsgerichtetes Simplex-Verfahren bestimmt, das durch einen Partikelschwarm initialisiert wird. Schließlich wurde eine Methode zur Speicherung der intrinsischen Kalibrierung mit Hilfe einer einfach zu realisierenden Kalibrierungsvorrichtung entwickelt, die in beliebigen Kamerapositionen wiederverwendet werden kann, sodass die intrinsische Kalibrierung nicht wiederholt werden muss. Die einzigen Einschränkungen der Kalibrierungspipeline sind, dass die projizierten Muster zeitlich wiederholbar sein müssen und dass die Position von Array-Projektor und

Kalibrierungsvorrichtung zueinander fest bleibt. Die erfolgreiche Anwendung der entwickelten Kalibrierungspipeline wird experimentell demonstriert. Es wurden 3D-Rekonstruktionen mit nur einer Kamera von weißen streuenden Kalibrierkörpern sowie von komplexen Objekten erstellt und die 3D-Messergebnisse mit Rekonstruktionen verglichen, die mit einem Stereokamera-Setup erzeugt wurden.

Table of Contents

Chapter 1.

Introduction

The desire to obtain optical photogrammetric non-tactile three-dimensional (3D) measurement of objects or scenes is increasing in a variety of industries and applications. The application of this type of 3D sensors in art and culture are of great relevance due to the topological sensitivity of historical objects, making 3D archaeological or paleontological shape measurement incredibly valuable [1]. 3D sensors have also seen use in the film and video game industries where 3D information from the motion capture of realistic facial and body movements is recorded as it is difficult for artists to manually reproduce them [2]. Non-tactile 3D measurement has also been seeing use in security applications offering enhanced biometric measurements as extra spatial and texture information is provided compared to conventional digital 2D images [3]. Non-contact 3D sensors will play a decisive role in providing solutions in the field of Human-Machine Interaction (HMI). Intensive research is being done on implementing 3D sensors for autonomous vehicle vision to realize self-driving automobiles [4]. In the same vein, these sensors are also highly desirable in the field of drones and robotics [5].

Naturally, demand exists to implement such solutions in industrial manufacturing for quality assurance and General Dimensioning and Tolerancing (GD&T) of sophisticated manufactured parts [6, 7]. Such 3D sensors also greatly improve any workflow process as users can nearly instantly obtain results. There has also been interest in the medical and life sciences towards the use of computer-aided design and computer-aided-manufacturing (CAD/CAM) of orthopedic and orthodontic devices [8]. In addition, these sensors could also see application in the diagnosis and treatment of musculoskeletal diseases such as cystic fibrosis as a means of measuring 3D chest wall volume [9]. Sensors with high quality 3D measurement performance are of importance considering dimensional error tolerances can have medical consequences.

While there are a variety of non-tactile optical photogrammetric 3D sensors, it is not intended to exhaustively review all possible technologies but focus on active structured illumination systems. Even with the focus on active structured illumination,

no specific methodology can be considered as an all-encompassing 3D measurement solution for every application. In this work, the focus is to operate in the macro-scale; delivering 3D measurements for approximately meter-scale working distances with cubic meter-scale measurement volumes. The camera is considered to have negligible distortion. While many types of surface materials exist, this work shall be limited to nearly-Lambertian surfaces such as evenly matte textures.

Active structured illumination is a highly active methodology in both research and industry where the test object is illuminated by a pattern of spatially varying intensity. This facilitates the identification of correspondences between the illuminating device and the imager such as a Charged-Coupled Device (CCD) or Complementary Metal Oxide Semiconductor (CMOS) digital camera. This enables 3D coordinates of surface points to be obtained commonly through calculations based on the triangulation of light rays between the illuminator and imager. To carry out this calculation, the geometric directionality of rays therefore needs to be determined within the same reference frame. Determining the directionality of rays of an optical device in its own coordinate system is known as intrinsic calibration. To bring both devices into the same reference frame, the pose relationship between them needs to be known and is obtained through extrinsic calibration.

With the wide accessibility, low-cost, and ease of use of conventional off-the-shelf digital projection devices such as Digital Light Projection (DLP) or Liquid Crystal on Silicon (LCoS), these devices as illuminators have been implemented in both research and industry. Since these devices operate using a digital projection chip and projection lens, the optical function is described as inverse perspective projection (or inverse pinhole modelled) as it operates similarly as a camera, but projects light rays into space instead of receiving them. Hence, the same well-known pinhole camera intrinsic parameters can be applied to model the outgoing geometric directionality of light rays. Thus, digital projectors can also be intrinsically calibrated using the well-established Zhang calibration method [10]. Once it has been calibrated, the intrinsic parameters can be saved for later use in alternate setup configurations without having to perform the intrinsic calibration again. As both the digital projector and camera are modelled as pinhole devices, the pose between them can be determined using the well-known extrinsic calibration method of the eight-point algorithm by Hartley [11].

The foundational technique of Fringe Projection Profilometry (FPP) has been a common point of comparison where the projected pattern takes the form of vertical

sinusoidal fringes. Whole-field 3D reconstructions can be made with high measurement accuracy and high measurement precision [12], particularly when dealing with diffusely reflecting surfaces [13] in cubic-meter scale measurement volumes at meter scale working distances. Two main varieties of FPP exist: Fourier Transform Profilometry (FTP) and Phase Shifting Profilometry (PSP). The techniques differ by the number of patterns projected to generate a single 3D reconstruction. Usage of FTP or PSP each come with their own trade-offs. FTP requires only a single fringe pattern to be projected to generate a reconstruction but is only capable of 3D measurement of smooth surfaces. PSP on the other hand does not have the smoothness limitation but requires at least three fringe patterns shifted by a constant phase to produce a single 3D measurement.

Phase unwrapping is a crucial step in FPP to determine phase-correspondences between the illuminator and imager to disambiguate phase and resolve phase discontinuities in the image. Phase unwrapping can be grouped into two classes: spatial unwrapping and temporal unwrapping. Each comes with their own trade-offs. Spatial unwrapping is restricted to spatially continuous surfaces but only requires one pattern as the fringe order of each pixel is tracked in relation with its neighbors. Temporal unwrapping does not have the continuity restriction but requires additional patterns to be projected to induce another code that keeps track of fringe order.

To work around these issues, stochastic pattern modalities have been developed for use such as laser speckle [14], Band-Limited Patterns (BLP) [15], and aperiodic sinusoidal fringes [16]. These stochastic pattern modalities could use spatial-based methods such as digital image correlation [17] to determine correspondences for single pattern projection, or temporal correlation-based methods for multiple pattern projection. Again, spatial or temporal based methods come with their own trade-offs. Spatial methods are dependent on the pixel window size. Correspondence identification in small windows are susceptible to imaging noise. Larger windows, while more robust to imaging noise, are susceptible to errors at depth discontinuities due to view occlusion [18]. While temporal methods are not susceptible to depth discontinuity errors, decreasing susceptibility to imaging noise requires an increasing number of projected patterns to generate a single 3D measurement.

Multi-pattern projection methods, while more favorable due to robustness against spatial scene discontinuities and high surface gradient measurement errors,

are therefore at a disadvantage when attempting to obtain 3D measurements at high-speed. This use-case has been seeing growing interest especially due to dynamic scenes. An example of dynamic scenes benefitting from whole-field 3D measurement would be airbag quality control and inspection during car crash testing [19]. The disadvantage is particularly evident when using off-the-shelf digital projectors as they are generally capable of projection rates up to only ~100-200 Hz. If one were to perform a conventional N = 3 phase-shifting method, then the hypothetical 3D measurement rate achievable is ~33 3D measurements/s assuming a sequence of independent patterns is utilized for a single 3D reconstruction. However, to measure such dynamic scenes, “ultrafast” measurement rates of 300 3D measurements/s may be required [20]. This would still be greater than what could be possible with conventional off-the-shelf digital projectors, even if considering a sliding sequence of patterns per 3D reconstruction.

This limitation in multi-pattern projection techniques has led to investigations on the hardware side toward improving the projection speed of the illuminating device. Unconventional high-speed projection devices have been created that do not depend on pattern formation using the digital chip-projection lens optical configuration. This avoids the projection rate limitations associated with conventional off-the-shelf digital projection. Examples of such devices have been proposed that involve a rotating diffuser disc or an Acousto-Optical Deflector (AOD) projecting laser speckles [21], a wobbling mirror that quickly spatially shifts a single BLP [15], or rotating transmission masks in Goes Before Optics (GOBO) projection [19].

A class of unconventional high-speed projection devices known as “array projectors” shall be studied here. One such device called a Multi-Aperture-Array-Projector (MAAP), created by Heist et al. at the Fraunhofer Institute of Optics and Fine Mechanics, was both utilized and emulated in this study [16]. The MAAP was demonstrated to be capable of 3D measurement rates up to 0.33 kHz with a shown projection rate of 3 kHz. It is capable of such high projection rates due to it containing six distinct projection channels that can be electrically switched rapidly in succession. Each projection channel projects a distinct aperiodic sinusoidal fringe pattern whose template is stochastically generated.

As unconventional projectors such as the MAAP are chip-less and do not utilize a projection lens, they cannot be pinhole modelled and therefore cannot be intrinsically calibrated using the Zhang calibration method. Any 3D measurement system that

utilizes them also cannot be extrinsically calibrated using the eight-point algorithm as it too assumes pinhole modelled optical devices. To work around this, a stereo-camera setup is implemented such that correspondence detection and triangulation is carried out between the two cameras to perform 3D measurement. This brings the total number of optical devices from two to three. With a greater number of devices required, this increases the total cost, weight, and form-factor of a 3D measurement system.

For systems of a chip-based pinhole modelled digital projector and a single camera setup, a well-established calibration pipeline already exists that must be performed to obtain 3D measurements. The three main steps of the calibration pipeline are to:

- Intrinsically calibrate the projector and the camera (or in the stereo case; the two cameras) using the Zhang method.
- Extrinsically calibrate to determine the pose between the projector and camera (or in the stereo case; between the two cameras) using the eight-point algorithm.
- Store the results of the intrinsic calibration for potential re-use in alternate setup configurations.

If one implements a 3D measurement system using a MAAP in a single camera configuration, no such calibration pipeline exists. The MAAP cannot be pinhole modelled. Hence, no 3D reconstructions would be possible. Therefore, the purpose of this work is to develop a comprehensive calibration pipeline so that a 3D measurement system using a MAAP and only a single camera can be implemented. Not only is this an academically challenging problem, but overcoming this monocular challenge also has practical benefits of decreased cost, weight, and form-factor when implementing a MAAP 3D sensor.

A series of proposals need to be assessed of their suitability if the successful development of a comprehensive calibration pipeline enabling a MAAP-single camera 3D measurement system is to be achieved.

- It shall be experimentally investigated using a physical setup if a proposed intensity sequence-to-depth mapping that empirically samples the projection frustum of each distinct projection channel of the MAAP is capable of intrinsically calibrating it, utilizing only a single camera as the optical capturing device. The

appropriateness of the proposed method is evaluated by comparing the surface precision and length accuracy of the generated 3D reconstruction with respect to geometrically calibrated gauges.

- It will be experimentally tested if the proposed intrinsic calibration method is still capable of enabling 3D reconstructions of scenes with spatial discontinuities and complex surface gradients, where each camera pixel independently reconstructs a 3D point.
- The intrinsic calibration method will be experimentally investigated to observe if it enables greater measurement volume depths compared to depths reported by a similar state-of-the-art phase-to-depth mapping method that utilized purely sinusoidal periodic fringes.
- An experimental demonstration shall be carried out to determine if the intrinsic calibration method is still capable if two-dimensionally varying intensity patterns are projected using a created “repeatable speckle” projector, rather than just being limited to the calibration of one-dimensionally varying intensity patterns of aperiodic sinusoidal fringes.
- A proposed downhill simplex optimization method initialized by a particle-swarm will be experimentally assessed using a MAAP emulator to determine if it is capable of extrinsically calibrating the physical MAAP-single camera 3D measurement setup. The proposed method is evaluated by comparing the surface precision, length accuracy, completeness, and consistency of the generated 3D reconstruction with respect to reference point clouds of a complex surface and with the measurements produced using a stereo-camera setup.
- An appropriate objective function to minimize in the proposed optimization method is searched for by carrying out a brute-force investigation to observe the convexity of the error surface of several different objective functions.
- The proposed optimization method implemented with the appropriate objective function is tested over a large sample of randomized separate optimization runs to observe if extrinsic calibration without *a priori* knowledge of an appropriate initial guess is consistent. This shall be compared to using solely the downhill simplex method with randomized initial guess.

- An experimental investigation using a MAAP emulator will be performed to ascertain if a proposed method dubbed intrinsic ‘re-calibration’ that artificially synthesizes intrinsic calibration images from previously stored images enables 3D measurements to be performed without having to repeat the entire intrinsic calibration process again for alternate arbitrary measurement camera views.
- The 3D measurements generated when using synthesized intrinsic calibration images is compared with 3D measurements generated from ground truth intrinsic calibration images. This shall be done by evaluating the surface precision, length accuracy, completeness, and consistency of the generated 3D reconstructions with respect to reference point clouds of a complex surface, and with the measurements produced using a stereo-camera setup.

1.1. Structure of thesis

This work is divided such that Chapters 2-5 provide in-depth the background necessary to understand the history and motivation behind the series of contributions this work demonstrates in Chapters 6-8. The actual breakdown of the chapters are as follows:

Chapter 2 gives an overview of fringe projection profilometry. This technique is considered the forerunner of triangulation-based active structured illumination methods in which 3D sensor operation is commonly compared with.

Chapter 3 gives an overview of the essential aspects of photogrammetry. Photogrammetry is the overall technique of obtaining 3D measurements through photographic images. In this chapter, the concepts required to understand the intrinsic calibration of digital cameras and projectors are summarized. The concepts required to understand extrinsic calibration of a triangulation-based active structured illumination 3D sensor is then briefly described.

Chapter 4 summarizes the composition of the MAAP studied in this work and how it is capable of high projection speeds. The chapter also gives an overview of the pattern modality it projects and how a conventional stereo-camera system performs triangulation with the array projector.

Chapter 5 gives an overview of the direction of research of recent literature and state-of-the-art on various active structured illumination techniques.

Chapter 6 describes the first step of the novel calibration pipeline; development of a process that intrinsically calibrates the MAAP. The proposed method is also shown to work for patterns whose intensity varies spatially two-dimensionally.

Chapter 7 describes the development of an extrinsic system calibration process for a single camera setup.

Chapter 8 describes the development of a view synthesis method that enables posterior intrinsic ‘re-calibration’. Without this, the intrinsic calibration process must be repeated every time an alternate measurement camera view is desired.

Chapter 9 details some possibilities for improvement beyond the scope of this work.

Chapter 10 provides a summary and conclusion.

Chapter 2.

Basics of fringe projection profilometry

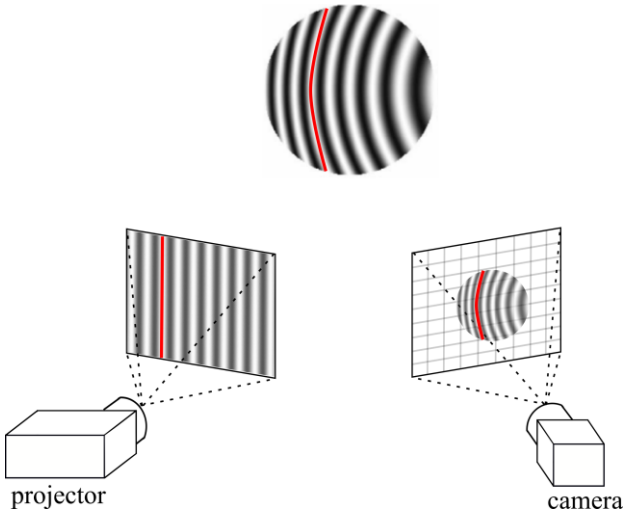

Fig. 2.1 *Digital fringe projection profilometry setup with red line depicting corresponding phase of sinusoidal fringe pattern.*

Fringe Projection Profilometry (FPP) is one of the classic forerunners to modern active structured illumination techniques for macroscopic and microscopic 3D optical metrology. A macroscopic digital FPP setup can be seen in Fig. 2.1. Often in many studies, a digital projection device projects structures in the form of sinusoidal fringes and images are recorded using a single digital camera. FPP can be generally described as:

- a projection device projects a sinusoidal fringe pattern onto the object surface
- an imaging device captures the phase modulation of fringes resulting from the surface height distribution
- analysis of the image to demodulate phase
- phase unwrapping to obtain a continuous phase distribution

Key differences between different FPP methods arise in how to phase modulate the surface, how to perform phase demodulation, and how to perform phase unwrapping. The following sections provide a brief overview of the techniques.

2.1. Phase modulation

2.1.1. Fourier Transform Profilometry

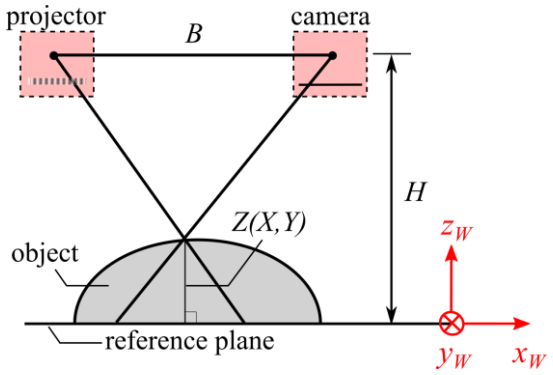

Fig. 2.2 *Setup geometry to obtain height distribution Z(X,Y) of object. The world coordinate system denoted with subscript W, B the baseline distance between projection centers, and H the height to the reference plane.*

Takeda et al. presented the basic principle on Fourier Transform Profilometry (FTP) [22]. Other similar variations exist using different transforms [23]. With a single sinusoidal fringe projected onto the object surface, the surface height distribution *Z(X,Y)* is compared to a reference plane as seen in Fig. 2.2, where *(X,Y)* is the spatial coordinate in the world coordinate system defined on the reference plane.

The captured image from a single digital camera can be represented as a Fourier series, such that

$$I(X,Y) = a(X,Y) + b(X,Y)\cos[2\pi f_0 X + \varphi(X,Y)]$$

$$I(X,Y) = a(X,Y) + b(X,Y)\sum_{n=-\infty}^{\infty} \mathcal{A}_n \exp i[2\pi f_0 nX + n\varphi(X,Y)] \tag{2.1}$$

where *I(X,Y)* is the image intensity, *a(X,Y)* is the background image intensity, *b(X,Y)* is the intensity modulation resulting from the non-uniform reflectivity of the object surface, $\varphi(X,Y)$ is the spatially modulated phase due to the surface height distribution, $\mathcal{A}_n$ is the Fourier coefficient for the n^{th} harmonic, f_0 is the fundamental carrier frequency of the underlying sinusoidal pattern, and *i* is the imaginary number. Assuming that *a(X,Y)*, *b(X,Y)*, and $\varphi(X,Y)$ vary much more slowly than f_0 of the sinusoidal pattern, performing the Fourier transform on Eqn. 2.1 and filtering only for the spectrum about f_0, the filtered image in the spatial domain can be described as

$$\tilde{I}(X,Y) = b(X,Y)\mathcal{A}_1 \exp i[2\pi f_0 X + \varphi(X,Y)] \tag{2.2}$$

Similarly, the filtered image of the pattern projected onto a reference plane in which the object surface height distribution is with respect to can be written as

$$\widetilde{I_0}(X,Y) = b_0(X,Y)\mathcal{A}_1 \exp i[2\pi f_0 X + \varphi_0(X,Y)] \tag{2.3}$$

where $b_0(X,Y)$ is the intensity modulation due to non-uniform reflectivity of the reference plane, and $\varphi_0(X,Y)$ is the spatially modulated phase due to the height distribution of the reference plane. Therefore, the phase difference $\Delta\varphi(X,Y) = \varphi_0(X,Y) - \varphi(X,Y)$ between the phase modulation of the object and the reference plane is obtained by taking the imaginary part, where i is the imaginary number, of

$$\ln\left[\tilde{I}(X,Y)\widetilde{I_0^*}(X,Y)\right] = \ln[b(X,Y)b_0(X,Y)|\mathcal{A}_1|^2] + i\Delta\varphi(X,Y) \tag{2.4}$$

Assuming that the direction of rays of both the camera and the fringe projector operate on the perspective projection model, the height Z of the object point seen at (X,Y) can be determined using the system geometry from Fig. 2.2 resulting in a phase-to-height equation

$$Z(X,Y) = \frac{H\Delta\varphi(X,Y)}{\Delta\varphi(X,Y) - 2\pi f_0 B} \tag{2.5}$$

where H is the distance between the camera projection center and the reference plane and B is the baseline distance between the projector and camera projection centers. To be able to separate the f_0 spectrum from the DC component and higher order harmonics as seen in Fig. 2.3, there must be no significant overlap between spectrums. Thus, the variation in the phase induced by the object surface needs to be

$$\left|\frac{d\varphi}{dX}\right|_{\max} < \frac{1}{3} \cdot 2\pi f_0 \tag{2.6}$$

This leads to a surface slope restriction such that

$$\left|\frac{dZ(X,Y)}{dX}\right|_{\max} < \frac{1}{3} \cdot \frac{H}{B} \tag{2.7}$$

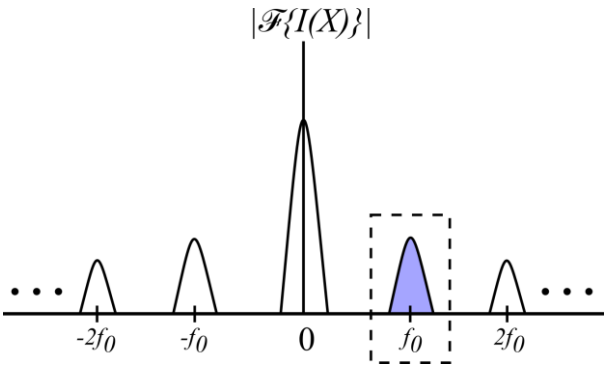

Fig. 2.3 *Separation of the fundamental carrier frequency f_0 spectrum (blue) of the underlying sinusoidal pattern from higher order harmonics and the DC component.*

where only smooth surfaces without strong surface complexity can be measured in typical FTP.

2.1.2. Phase-Shifted Profilometry

An alternative method to phase modulate the surface and perform demodulation is known as Phase-Shifted Profilometry (PSP). The basic principle involves the projection of a sequence of multiple phase-shifted patterns. Classically, an $N = 3$ pattern sequence is utilized. Other variations exist involving different sequence lengths [24–26]. Typically, a larger number of patterns result in improved and more robust 3D measurement performance [26]. For an $N = 3$ sequence projected onto the object surface, the consecutive images from a digital camera can be described as

$$I_1(X,Y) = a(X,Y) + b(X,Y)\cos[2\pi fX + \varphi(X,Y) - \phi]$$

$$I_2(X,Y) = a(X,Y) + b(X,Y)\cos[2\pi fX + \varphi(X,Y)] \tag{2.8}$$

$$I_3(X,Y) = a(X,Y) + b(X,Y)\cos[2\pi fX + \varphi(X,Y) + \phi]$$

for a shift in phase of the pattern by constant Φ. A common phase shift for an $N = 3$ sequence is $\Phi = 2\pi/3$. The three unknowns of background image intensity $a(X,Y)$, intensity modulation due to non-uniform reflectivity of the object surface $b(X,Y)$, and the spatially modulated phase due to the surface height distribution $\varphi(X,Y)$ can be solved for using the three equations of Eqn. 2.8. Solving for the spatially modulated phase obtains

$$\varphi(X,Y) = \tan^{-1}\left[\frac{\sqrt{3}[I_1(X,Y) - I_3(X,Y)]}{2I_2(X,Y) - I_1(X,Y) - I_3(X,Y)}\right] \tag{2.9}$$

The phase modulation on the reference plane $\varphi_0(X,Y)$ can be obtained similarly, allowing one to obtain the difference $\Delta\varphi(X,Y)$ so that the object height distribution can be determined using Eqn. 2.5. Contrasted with FTP, there are no measurement surface restrictions introduced by the modulation process. On the other hand, as three phase-shifted patterns must be projected, the acquisition time to obtain a single 3D measurement increases.

2.2. Phase unwrapping

Phase unwrapping is a critical step after illumination has been performed. With sinusoidal fringes, only phase values [-π, π] are detected in an image. This results in phase ambiguities across the object surface. Phase discontinuities also arise when a total phase of 2π has been traversed as the phase jumps from one end of the interval to the other, -π or π. Thus, phase values need to be made continuous. Generally, phase unwrapping methods can be grouped into two classes: spatial and temporal [26].

2.2.1. Spatial unwrapping

Spatial phase unwrapping methods requires all phase values measured relative to a designated reference point and can be performed using a single image. As seen in Fig. 2.4, the fringe order is tracked by adding or subtracting multiples of 2π to the phase according to the reference's neighbours [26]. However, this can only be done for objects having a connected surface. Obtaining a continuous phase map fails for surface discontinuities such as holes and spatially separated objects.

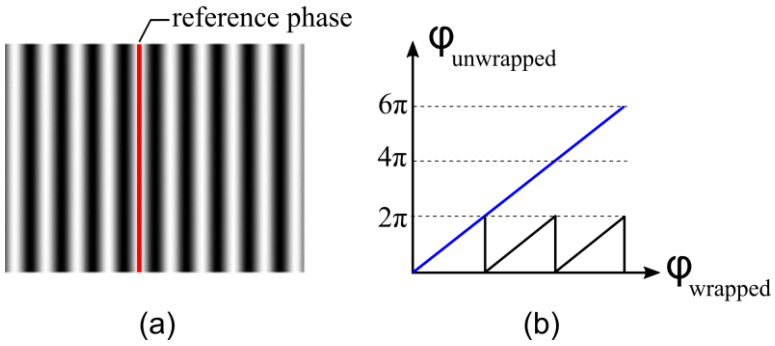

Fig. 2.4 *(a-b) The phase φ is unwrapped in spatial methods by tracking the fringe order of neighboring image pixels with respect to a designated reference phase.*

2.2.2. Temporal unwrapping

Temporal unwrapping attempts to alleviate the discontinuity issue by providing additional information in the time domain by projecting supplemental patterns [27]. Each pixel receives a complementary code that clarifies the fringe order of its phase value. Variations in supplementary patterns have been extensively studied. Examples include binary code (as seen in Fig. 2.5), gray code, De Bruijn code, and different phase maps with different fringe periods [24, 27, 28]. While surface discontinuities and spatial scene separations can be measured using temporal phase unwrapping, it increases the total number of patterns projected. The acquisition time to obtain a single 3D measurement therefore increases.

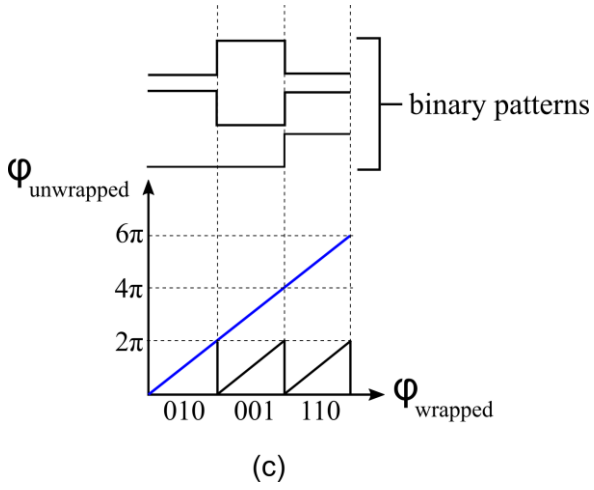

Fig. 2.5 *The phase is unwrapped in temporal methods by projecting an additional pattern sequence. In this example, an N = 3 binary pattern sequence keeps track of fringe order with an extra binary code.*

Chapter 3.

Basics of 3D photogrammetry

3.1. Perspective projection and intrinsic calibration

Understanding how light scattered from an object travels through 3D space and is imaged onto the 2D camera image plane is vital in photogrammetry. The most common model is known as perspective projection or 'pinhole' model. As seen in Fig. 3.1, the projection mapping of a 3D point onto a 2D plane can be geometrically modelled such that a single ray scattered from any particular point in 3D space is constrained to pass through a single point in the camera in order to reach the imaging plane. This single point, known as the projection center or pinhole, is located a distance away from the image plane equal to the focal length of the camera objective lens. In Fig. 3.1 the image plane coordinate system and the camera coordinate system are conventionally defined.

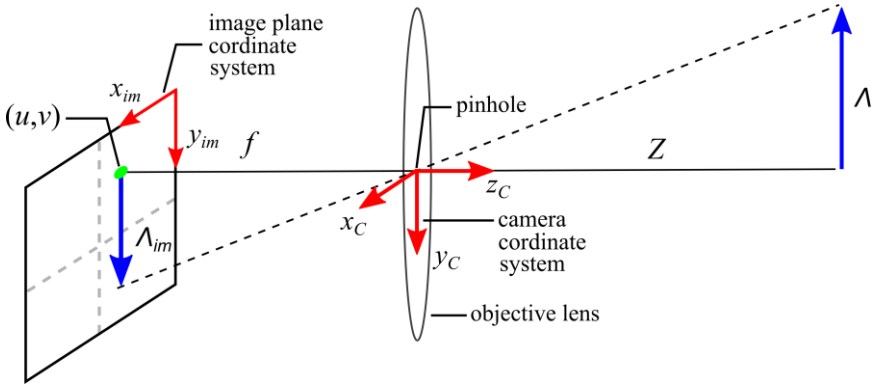

Fig. 3.1 *Perspective projection of the length Λ in 3D space onto a 2D image plane of a camera denoted as the projected length Λ_{im}, where all rays from space are constrained to pass through the camera pinhole. The point of intersection of the optical axis with the image plane is the principal point denoted as (u,v). Ideal alignment between the image plane and optical axis would place (u,v) at the center of the image plane.*

In order to more easily convey mathematically the 3D-2D projection mapping, let us first assume a mathematical abstraction that is functionally similar where the image plane is in front of the pinhole such that images are not inverted on the image plane in a format known as the 'retinal' convention [29]. Observing Fig. 3.1, considering the basic geometry principle of similar triangles in the reference frame of the camera coordinate system, it can be seen that

$$\frac{\Lambda_{im}}{f} = \frac{\Lambda}{Z} \tag{3.1}$$

where the projection of the length Λ onto the camera image plane is the length Λ_{im}, Z is the distance from the pinhole, and f is the focal length.

The point at which the optical axis intersects the imaging plane is known as the principal point. Ideally, the principal point is located at the center of the image plane. However, due to the camera manufacturing process, there will always be some offset of the image chip from the ideal principal point location. Let the offset of the principal point (u,v) be relative to the image plane coordinate system in meter units. Let the length from the optical axis to a 3D point (X,Y,Z) in space in meter units be relative to the camera coordinate system. This length is projected onto the camera image plane. The projected length from the principal point to a 2D point (x_{im},y_{im}) in meter units in the image plane coordinate system is therefore

$$\begin{pmatrix} x_{im} \\ y_{im} \end{pmatrix} = \begin{pmatrix} f_x\left(\frac{X}{Z}\right) + u \\ f_y\left(\frac{Y}{Z}\right) + v \end{pmatrix} \tag{3.2}$$

Knowing the pixel pitch Δp, the projected length on the image plane in pixel units can be expressed in matrix form considering homogenous coordinates such that

$$\begin{pmatrix} \alpha \\ \beta \\ 1 \end{pmatrix} = \frac{1}{\Delta p}\begin{pmatrix} f_x & \eta & u \\ 0 & f_y & v \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{3.3}$$

where $\eta = 0$ is the skew coefficient for perpendicular image axes. A non-zero skew-coefficient is atypical. The mapping matrix containing the camera parameters is known as the intrinsic matrix.

The Zernike polynomials are used in optics theory to describe imaging aberrations arising from wavefront deformations. Spherical Aberration, Coma, Astigmatism, Curvature of Field, and Distortion are five primary types. The main source of deviation of concern from the pinhole model of camera imaging is distortion as the geometric principle of similar triangles, and thus perspective projection, would no longer hold. Instead of using the wavefront description of distortion, Brown

modelled image distortion as non-linear deviations of each image point from its ideal non-distorted position [30]. Let the normalized coordinates be (x_n, y_n). The distorted point in pixel coordinates is thus

$$\begin{pmatrix} x^r_{im} \\ y^r_{im} \end{pmatrix} = \begin{pmatrix} x_n[1 + k_1 r^2 + k_2 r^4 + \cdots] \\ y_n[1 + k_1 r^2 + k_2 r^4 + \cdots] \end{pmatrix} \tag{3.4}$$

$$\begin{pmatrix} x^t_{im} \\ y^t_{im} \end{pmatrix} = \begin{pmatrix} p_2(r^2 + 2x_n) + 2p_1 x_n y_n \\ p_1(r^2 + 2y_n) + 2p_2 x_n y_n \end{pmatrix} \tag{3.5}$$

$$\begin{pmatrix} \alpha_d \\ \beta_d \end{pmatrix} = \frac{1}{\Delta p} \begin{pmatrix} f_x(x_r + x_t) + u \\ f_y(y_r + y_t) + v \end{pmatrix} \tag{3.6}$$

where (α_d, β_d) are the distorted pixel coordinates, (x^r_{im}, y^r_{im}) is known as the radial distortion, (x^t_{im}, y^t_{im}) is known as the tangential distortion, k_n is the n^{th} radial distortion coefficient, p_n is the n^{th} tangential distortion coefficient, and the radial distance is $r^2 = (x_n - u)^2 + (y_n - v)^2$. Tangential distortion arises due to tilt of the lens with respect to the image plane while radial distortion arises due to non-uniform magnification variation with respect to distance from the lens' optical axis.

As there will be some error from reported specifications due to unavoidable errors arising from the camera manufacturing process, the actual intrinsic parameters need to be determined. The seminal work that presented a comprehensive procedure to obtain the camera intrinsic matrix is known as the Zhang method and is considered well-established [10]. It requires the camera to image a planar calibration target containing identifiable feature points so that correspondences between the image plane and the calibrating plane are identified. Appendix A can also be referred to for further detail. As seen in Fig. 3.2, the calibrating plane must be imaged for $N \geq 3$ distinct positions. Each plane position must have different orientation from one another. Most commonly, a chessboard is used as the corners can be easily located and used as feature points. As all rays are constrained to pass through the pinhole, the directional vector of each ray is therefore dictated by the intrinsic camera parameters. Determining the directionality of light rays is known as intrinsic calibration. As long as the camera condition does not change, one can save and re-use the same intrinsic parameters.

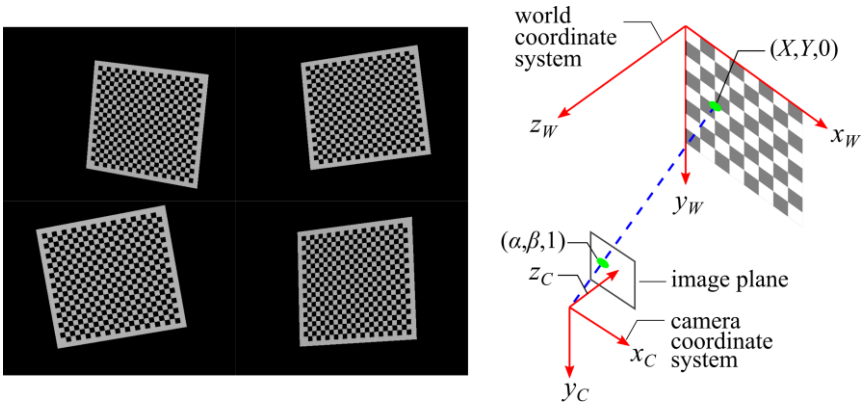

Fig. 3.2 *(a) Typical planar calibration target of a chessboard used in the Zhang method of intrinsic calibration. (b) The projection of a corner with 3D coordinates (X,Y,0) on the planar target used as a feature point and projected onto the image plane with homogeneous image coordinates (α,β,1). All 3D points are with respect to the planar coordinate system which is used as the world coordinate system.*

3.2. Epipolar geometry, extrinsic calibration, and triangulation

Beyond the FPP phase-to-height equation, the use of triangulating stereo-setups similar to binocular human vision are also used. As seen in the typical triangulation configurations of Fig. 3.3, to triangulate an object point *(X,Y,Z)* in 3D space, identifying the homologous points *(α,β)* and *(α,β)'* in each image plane is required and is known as the correspondence problem. The structured illumination in active photogrammetry aids in determining these corresponding points through the encoding of additional information. Triangulation can then be carried out similarly either between two

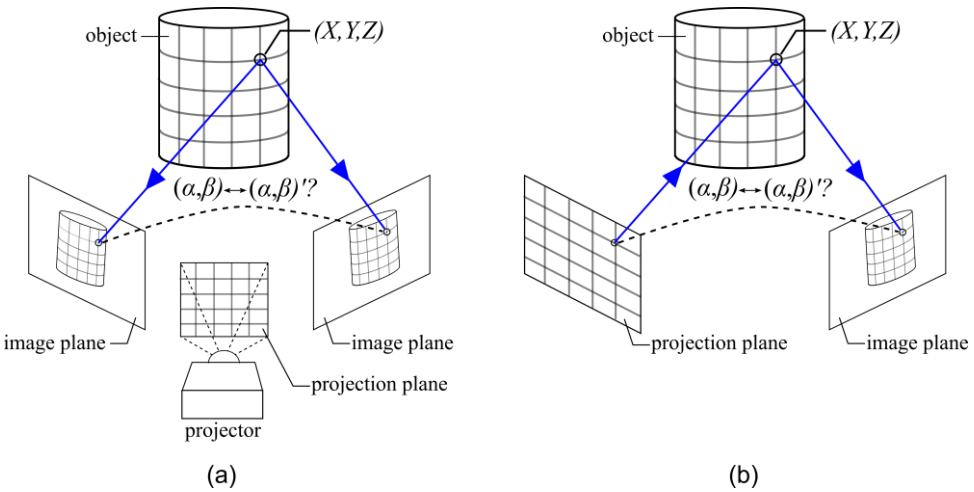

Fig. 3.3 *Typical triangulation setup configurations: (a) between two cameras or (b) between a digital projector and camera. Regardless of configuration, triangulation of (X,Y,Z) requires the detection of corresponding points (α,β)-(α,β)' between the two triangulating devices.*

cameras in a stereo configuration or between a camera and a projector as long as the two triangulating components are intrinsically calibrated. Conventional off-the-shelf digital projectors such as Digital Light Processing (DLP) systems utilizing Digital Mirror Devices (DMD) or other digital projectors using Liquid Crystal on Silicon (LCoS) are typically used due to their accessibility and low cost. They can be intrinsically calibrated in a similar fashion to a camera as they can also be pinhole modelled but operating inversely [31–33]. Instead of its optical function being to receive light rays, they project light rays into space.

The relational pose aligning two pinhole modelled devices so that all 3D coordinates are in a common world reference frame is known as the Epipolar geometry. Determining the Epipolar geometry is called extrinsic system calibration. The seminal works on extrinsic calibration and the Epipolar geometry were developed by Luong [29] and a popular well-established method was created by Hartley known as the eight-point algorithm [11]. Appendix B can be referred to for further detail. Performing the eight-point algorithm requires $N \geq 8$ corresponding points from the view of a non-planar scene [34–37]. From the eight-point algorithm, the algebraic representation of the Epipolar geometry known as the 3×3 Fundamental matrix F is obtained. F is independent of the scene.

From the Epipolar geometry, a constraint arises for a stereo setup. As seen in Fig. 3.4, let the symbol P represent any point in 3D space with spatial coordinates of (X,Y,Z). For a homologous corresponding pair (P_L,P_R) of the observed object point P, the pinholes of each device on the left and right-hand side (P^o_L, P^o_R) all lie on a single plane in 3D space known as the Epipolar plane Π. Π contains the baseline distance B connecting the pinholes of each camera. The points where the baseline intersects each image plane are known as the Epipoles P^e_L and P^e_R. The intersection of Π with each image plane is known as the Epipolar line l_L *and* l_R. P_L and P_R are constrained such that they lie along l_L *and* l_R in their respective image planes. If the projections in pixel coordinates of P_L and P_R are written in homogeneous form as $(\alpha,\beta,1)_L$ and $(\alpha,\beta,1)_R$ respectively, the search for all other corresponding points in pixel coordinates can therefore be simplified by exploiting the Epipolar constraint

$$(\alpha \;\; \beta \;\; 1)_L \, F \begin{pmatrix} \alpha \\ \beta \\ 1 \end{pmatrix}_R = 0 \tag{3.7}$$

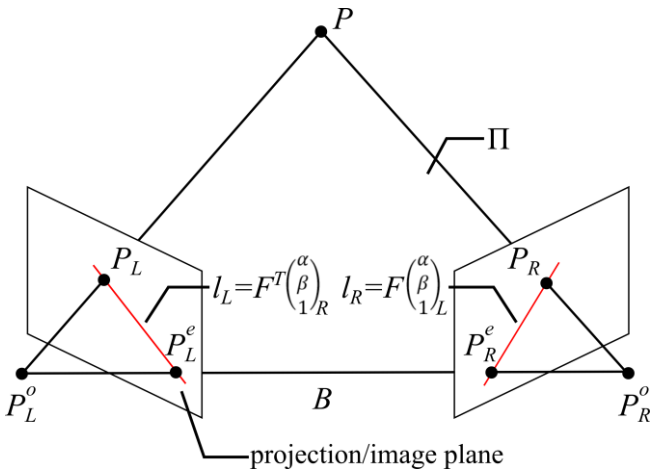

Fig. 3.4 *Epipolar geometry of triangulating pinhole modelled devices.*

From the Epipolar constraint, the Epipolar lines in each image are derived such that

$$l_R = F\begin{pmatrix}\alpha\\\beta\\1\end{pmatrix}_L = 0 \qquad l_L = F^T\begin{pmatrix}\alpha\\\beta\\1\end{pmatrix}_R = 0 \tag{3.8}$$

where the lines can be formulated as the line equation in standard form.

The Fundamental matrix of the stereo-view geometry can also be decomposed into a form more familiar as the rotation and translation of one camera pinhole with respect to the other camera pinhole

$$[t_c]_\times R_c = M^T_L \, F \, M_R \tag{3.9}$$

where $[t_c]_\times R_c$ is the matrix cross product of the translation vector and the rotation matrix between the two camera coordinate systems, and M_L and M_R are the intrinsic camera parameter matrices of the left and right cameras respectively.

The Fundamental matrix *F* however is not unique in terms of scale as utilizing *F* would result in a 3D reconstruction with correct surface topography but with incorrect scale compared to the ground truth. Thus, extrinsic calibration also requires a gauge whose ground truth geometric dimensions are known in order to determine the correct scale to apply to the 3D reconstruction. As a result, to achieve a final 3D reconstruction with correct scale, a calibration pipeline must be executed as seen in the flowchart in Fig. 3.5.

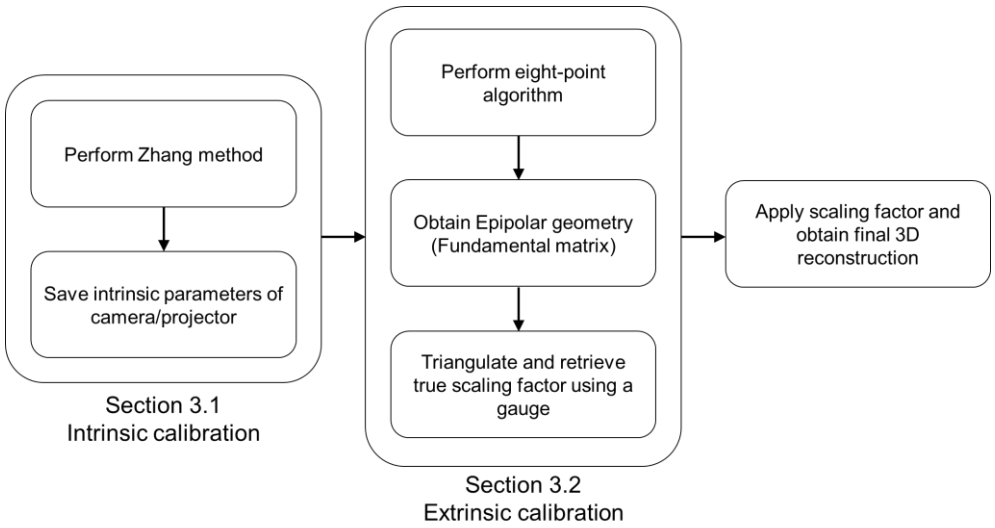

Fig. 3.5 *Flowchart depicting calibration pipeline for 3D sensors relying on triangulation between pinhole modelled devices.*

Chapter 4.

Multi-Aperture Array Projection

In this work, the opportunity to utilize a projection device known as a Multi-Aperture Array Projector (MAAP) was presented. Before detailing the contributions of this study, a brief overview of the mechanism of projection and aperiodic sinusoidal fringes will be introduced. The operating principle of the MAAP has been covered extensively in previous literature [16].

The MAAP was created with the intent of achieving high projection speeds. Conventional off-the-shelf chip-based projection systems such as Digital Light Projection (DLP) using digital mirror devices (DMD) or projectors utilizing Liquid Crystal on Silicon (LCoS) chips usually have max projection rates of ~0.1 - 0.2 kHz [14]. In a typical $N = 3$ phase-shifting method, then the hypothetical 3D measurement rate achievable is ~33 3D measurements per second (assuming an independent image sequence). With the ability to project at drastically higher rates, this makes multi-pattern projection methods much more attractive. It opens the ability to measure highly dynamic scenes as well as decreasing the process times to generate a 3D reconstruction.

The MAAP is capable of high projection rates with potential up to 100 kHz. 3D measurement rates up to 0.33 kHz using a stereo-camera system was demonstrated. This is made possible as the MAAP contains a linear array of six independent projection channels that can be rapidly electrically switched in succession. Each channel projects a distinct aperiodic sinusoidal fringe pattern, for a sequence of six total patterns projected. Each independent projection channel utilizes a cluster of high intensity LEDs acting as a light source. Formation of a single image in each projection channel is produced by a cluster of light-modulating micro-projection optics rather than a single digital projection chip.

4.1. Projection Mechanism

A projection channel of the MAAP comprises of a cluster of identical Köhler-illumination micro-projection units hexagonally packed and realized through

lithographic etching onto a glass substrate. As seen in Fig. 4.1, each individual micro-projector includes a condenser lens that focuses incoherent collimated light into the projection lens aperture. A chromium-based transmission structure is deposited between each condenser-projector lens pair whose transmission profile is of the target output image. By illuminating the cluster with collimated light, the projections from each micro-projection superimpose to form a single desired output image with sufficient intensity. The superposition of each micro-projector's output illumination is achieved by adjusting the pitch between the transmission structure and the projector lens and by adjusting the tilt of each micro-projector [38]. A cluster is compact, occupying an area ~10×10 mm^2. Due to the small size of the projection apertures, a high depth-of field is present. In comparison with a typical off-the-shelf DMD-based projector, the patterns projected by the MAAP are fixed due to the etched transmission structure. As off-the-shelf projectors have only a single projection channel, a sequence of patterns are projected along the same propagation path as seen in Fig. 4.2a. As each projection channel of the MAAP is laterally displaced from one another, a lateral shift between patterns arises as seen in Fig. 4.2b. The size of the lateral shift depends on the depth from the MAAP.

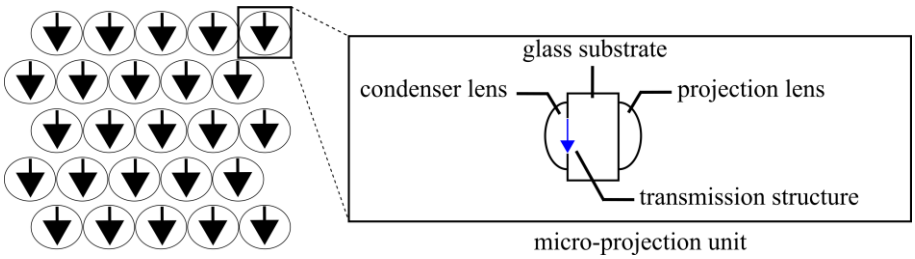

Fig. 4.1 *Cluster of micro-projection units composing a single pattern projecting aperture [39].*

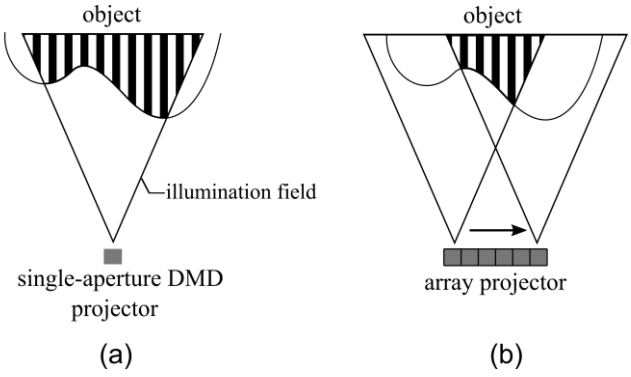

Fig. 4.2 *Illumination field projected by a (a) single-aperture DMD based projector compared with a (b) multi-aperture array projector where there is a shift in each projected pattern due to laterally displaced apertures [39].*

4.2. Temporal coding with aperiodic sinusoidal fringes

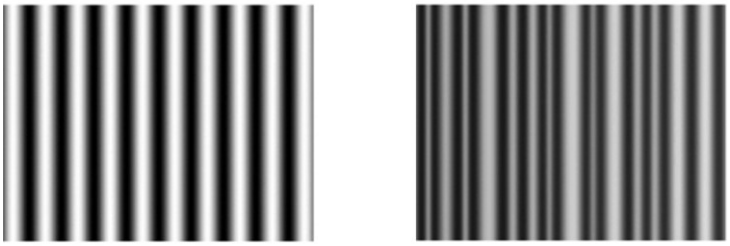

Fig. 4.3 *Example of (a) sinusoidal periodic fringes and (b) sinusoidal aperiodic fringes.*

In traditional phase-shifting profilometry (PSP), an $N \geq 3$ sequence of phase-shifted sinusoidal fringe patterns as demonstrated in Fig. 4.3a is projected. The spatial variation in intensity profile $I(\alpha,\beta)$ for pixel coordinates (α,β) of the projection chip of the j^{th} sinusoidal fringe pattern projected can be described as

$$I_j(\alpha,\beta) = a + b\big(\sin(c\alpha + \phi_j)\big), \qquad j \in \{1 \dots N\} \quad (4.1)$$

where a is the vertical intensity shift, b is the intensity amplitude, c is the wavenumber, $2\pi/c$ is the period length, and Φ_j is some constant phase shift. A common design in traditional PSP is to have the phase shift such that $\Phi_j = 2\pi(j-1)/N$.

Due to the size of the lateral shift of patterns being dependent on depth, conventional phase-shifting methods discussed in Sec. 2.1.2 (pg. 12) to phase modulate the surface cannot be utilized. The phase-shift would not be constant throughout the measurement volume. Rather than project sinusoidal fringes, the MAAP projects a sequence of aperiodic sinusoidal fringe patterns instead. Each aperiodic fringe pattern j has the intensity profile for N total patterns in the sequence such that

$$I_j(\alpha,\beta) = a_j(\alpha) + b_j(\alpha)\big[\sin(c_j(\alpha)\alpha + \phi_j(\alpha))\big], \qquad j \in \{1 \dots N\} \quad (4.2)$$

with varying vertical intensity shift $a_j(\alpha)$, varying intensity amplitude $b_j(\alpha)$, varying wavenumber $c_j(\alpha)$, varying period length $2\pi/c_j(\alpha)$, and varying phase $\Phi_j(\alpha)$. In Fig. 4.3b, the visual difference between a sinusoidal fringe pattern and an aperiodic fringe pattern can be seen. Due to the different pattern modality, single camera techniques derived from phase-based methods cannot be used to perform 3D reconstruction. A stereo configuration implementing two grayscale cameras is utilized instead.

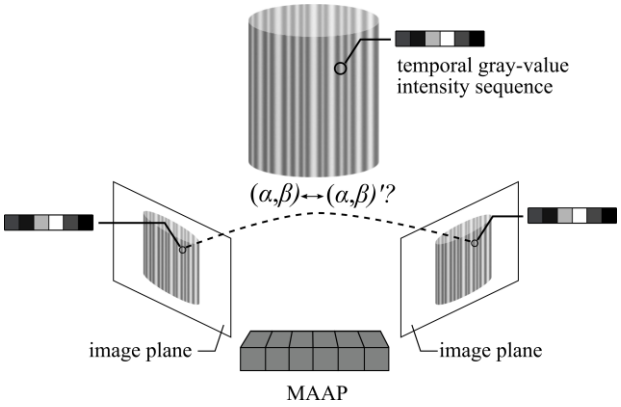

Fig. 4.4 *In temporal structured illumination, correspondences between triangulating pinhole devices are identified through normalized cross-correlation of gray-value intensity sequences.*

Each camera in the stereo setup synchronously images every projected pattern such that every image pixel of the observed object is encoded by six gray-value intensities in the time domain as shown in Fig. 4.4. Every pixel therefore has the capability to independently generate a 3D point without reliance on its neighbors.

Finding corresponding points between two views is solved through normalized cross-correlation between their temporal gray-value intensity sequences to determine their similarity such that

$$\rho(C, C') = \frac{\sum_{j=1}^{N}(C_j - \bar{C}) \cdot (C'_j - \bar{C'})}{\sqrt{\sum_{j=1}^{N}(C_j - \bar{C})^2}\sqrt{\sum_{j=1}^{N}(C'_j - \bar{C'})^2}} \tag{4.3}$$

where for each pattern j, C_j and C'_j is the j^{th} gray-value of the C and C' camera pixel, and $\bar{C}$ and $\bar{C'}$ is the arithmetic mean of the gray-value sequence of the C and C' camera pixel over N = 6 images. The cross-correlation coefficient ρ varies between [-1,1]. A value approaching 1 means the two sequences are similar. A value approaching -1 means they are dissimilar. The quality of correspondence matching improves for increasing number of patterns N in the temporal sequence [14], but at the sacrifice of the acquisition speed to generate a single 3D measurement. No temporal code is present in the presence of spatial discontinuities as a gap scatters no illumination back to the cameras. This results in a low cross-correlation coefficient between two views, meaning no correspondence is identified. Thus, no additional step of phase unwrapping after initial illumination is needed to reconstruct discontinuous scenes as experienced with phase-based methods. Due to the fundamentally different nature compared to FTP, surface smoothness restrictions are not present.

Chapter 5.

Current literature and state-of-the-art: a balance of tradeoffs

Unlike traditional fringe projection profilometry (FPP) methods, the ideal measurement technique should not introduce spatial limitations on what scene can be measured. For example, it was seen from Sec. 2.1.1 (pg. 10), that demodulation of phase in Fourier Transform Profilometry (FTP) results in being limited to measure only smooth surfaces. Another limitation arises when resolving phase ambiguities using spatial unwrapping techniques as described in Sec. 2.2.1 (pg. 13). The ability to measure a surface becomes limited further to only continuous scenes as the phase tracking with pixel neighbors is disrupted for spatial disjointedness. The issues seen in FTP and spatial phase unwrapping can be mitigated by using Phase-Shifting Profilometry (PSP) methods with temporal phase unwrapping. But these require additional projected patterns and increase the time to acquire a single 3D reconstruction of the surface[1].

The ideal structured illumination system should have each camera image pixel independently capable of generating a 3D point without requiring information from its neighbors. This capability should be present regardless of the continuity or gradients of the surface distribution. Not only would an ideal system be capable of fast 3D measurements but would also be implemented with fewer cameras to the limit of only a monocular view. In this way, utilizing a setup is enabled with lower cost, weight, and smaller form-factor as another optical device is no longer needed.

The limitations experienced in classical FPP have led the research community to propose various methods to overcome them and attempt to achieve the ideal structured illumination system. The following are selected examples that indicate the current direction of research.

[1] Obtaining a 3D measurement from active structured illumination can be considered in two phases: data acquisition (all necessary projected patterns) and data processing (calculation of the 3D coordinates). While obtaining 3D measurements can be considered sped up by decreasing the computation time of data processing, data acquisition is the limiting factor. ‘Fast 3D measurements’ is considered in the scope of how quickly data can be acquired.

5.1. Spatial coding with random codeword or statistical patterns

Spatial coding can involve a pseudo-random codeword grid and a 3D measurement acquired using only a single projected pattern. Some recent studies have involved filling the grid with either binary or colored rhombus elements where each rhombus is embedded with a random feature [40]. Considering a window around a rhombic grid point, each embedded feature within the window forms a unique codeword. Through corner detection, every grid point can be located, and the correspondence problem solved by matching codewords between a camera and projector. Other possibilities involve binary mini-patterns or pseudo-random dots [41, 42]. Due to the codeword window, the codeword itself occupies a finite amount of space of the surface. This results in low density point clouds as pixel-by-pixel reconstruction is not achieved. Another form of spatial coding is through statistical pattern projection such as laser speckle [17]. Instead of a codeword window, intensity windows are implemented so that digital image correlation-based or optical flow-based methods can be used. While a single projected pattern is only needed, a stereo-camera setup is involved. However, intensity windows typically result in foreground fattening errors at depth discontinuities due to occlusion between views [18, 43, 44].

5.2. Hybrid FTP-PSP

The release of the DLP Discovery 4100 development kit by Texas Instruments has made advanced hardware control of high-speed projection highly accessible to researchers [45]. The kit is capable of binary pattern projection rates greater than 32 kHz and gray-scale pattern projection rates up to nearly 2 kHz. The ability to project at drastically higher rates makes multi-pattern projection methods with digital projectors much more desirable than before.

Proposed by Zuo et al., the kit was utilized such that a projection rate of 20 kHz was accomplished, and 3D measurements achieved with a rate of 10 kHz [46, 47]. Three phase-shifted binarized sinusoids defocused by the projector lens and three flat patterns between sinusoids are projected to exploit the 32 kHz rate. Assuming the DC component of background ambient light can be neglected, the phase and surface reflectivity can be obtained using only a single sinusoid pattern and a single flat

pattern. To measure discontinuous complex surfaces, a hybrid spatio-temporal phase unwrapping method was proposed. However, errors arise in the presence of spatio-temporal discontinuities occurring when an object moves faster than consecutive phase-shifting sequences can be recorded.

5.3. Linear LED array phase shifting

An approach proposed by Fujigaki et al. surpasses the limits of conventional off-the-shelf DLP projection rates. It presented a Ronchi ruling illuminated by a linear LED array, resulting in a laterally shifted sequence of illumination similar in functionality to the MAAP [48–50]. It was shown that stable projection rates up to 100 kHz was possible and 3D measurement rates up to 4 kHz was demonstrated. Due to this lateral shift of illumination, the phase-shift of the sinusoidal pattern is dependent on depth. Thus, conventional PSP and temporal phase unwrapping cannot be performed. Instead, a test plane is scanned depth-wise through the measurement volume and an image of the projection at each depth increment z is taken by a single camera. By using the three-step phase-shifting method for each individual camera pixel, the corresponding phase φ_0, φ_1, φ_2, … at every depth z_0, z_1, z_2 … can be determined. Thus, along a light ray, the change of phase with respect to depth is realized. No phase unwrapping method was proposed to resolve the depth-wise phase ambiguity along a camera ray. This limited the possible measurement volume depth allowing only low surface profiles to be measured in order to ensure depth-wise phase uniqueness.

5.4. Multi-wavelength projection

With the availability of multispectral cameras, there have been proposed methods recently of projecting in the near infrared spectrum [51]. The surface is encoded by a single spectrally multiplexed pattern produced by a spectral array projector. Each image pixel therefore contains a spectral intensity sequence. Using a synchronized stereo-camera setup, the Epipolar geometry can be exploited to solve the correspondence problem. The normalized cross-correlation of spectral intensity sequences is used to find which camera pixel corresponds to which pixel in the other. In comparison to non-spectral imagers of the same imaging resolution, multispectral imagers are effectively spatially under-sampled, leading to a decrease in the point

cloud density. This is due to the greater distance present between filtering elements of a specific wavelength. This issue can be worked around by choosing a multispectral camera with greater imaging resolution than a comparable non-spectral one.

5.5. Temporal projection of laser speckle

To project quickly a temporally varying set of objective speckle patterns, coherent illumination of a laser beam is focused and incident on a rotating diffusing medium [14, 52]. Alternatively, an Acousto-Optical Deflector (AOD) rapidly deflecting the angle of laser beam incidence was also demonstrated [21]. Using a synchronized stereo-camera setup, the correspondence problem between two views is solved through the temporal normalized cross-correlation approach. The AOD method demonstrated that the speckle pattern modality is capable of projection rates up to 205 kHz and 0.33 kHz 3D measurement rate using 15 independent images was demonstrated. No additional surface restrictions are imparted by the technique as each camera pixel is independently encoded by a gray-value intensity sequence and therefore each observed object point can be independently triangulated.

In multi-pattern projection methods operating with high projection rates, cameras capable of high frame rates are also needed. However, the cost of a camera generally increases for increasing frame-rate capability. Thus, stereo-camera configurations in multi-pattern methods such as this compound the problem of 3D measurement system cost.

5.6. Spatial shift of single pattern

Proposed by Große et al., a single randomly generated Band-Limited Pattern (BLP) is projected onto a rotating wobbling mirror and re-projected into space [15]. As a result, the re-projected BLP is spatially shifted in time. Using a synchronized stereo-camera setup, temporal cross-correlation is implemented to solve the correspondence problem. The spatial shift generates a sufficiently altered image for unique pixel intensities at different time steps. Using 15 independent images, a 3D measurement rate of 0.70 kHz was demonstrated. Similarly in work by Heist et al., a Goes Before Optics (GOBO) projector was created by implementing a rotating transmission mask in front of a high radiant flux source [19]. The transmission pattern projects an

aperiodic sinusoidal fringe pattern, resulting in a demonstrated 3D measurement rate of 1.30 kHz using nine independent images. In both examples, no additional surface restrictions are introduced due to pixel independent intensity sequence coding, but two cameras are needed. Recently, Liu et al. was able to accomplish a method utilizing a rotating mask to project purely sinusoidal fringes [53]. The mask is synchronized with the camera through optocoupling to image at the correct time to obtain the correct phase shift. As the GOBO principle was implemented, a projection center still exists. Thus, the phase-to-height equation of Eqn. 2.5 (pg. 11) could still be utilized to perform single camera 3D measurement. While able to reconstruct surface holes/gaps, the proposal is still susceptible to spatial separations where no continuous phase path exists between neighbouring pixels.

5.7. Monocular-view challenge

	Highest demonstrated measurement rate (independent image sequence)	Pixel-by-pixel independent without introducing spatial measurability limitations?	Cameras
5.1 Spatial coding	Single-shot	No	1/2
5.2 Hybrid FTP-PSP	10 kHz	No	1
5.3 Linear LED array	4 kHz	No	1
5.4 Multi-wavelength	Single-shot	No	2
5.5 Temporal laser speckle	0.33 kHz	Yes	2
5.6 Spatial shift of single pattern	3 kHz	Yes/No	1/2

Table 5.1 *Summary of recent literature and state-of-the-art showing the tradeoffs made between the qualities of measurement speed, surface spatial measurability, and the number of camera views.*

Shown in Table 5.1 is a summary of the recent literature reviewed in this chapter. It summarizes the operating characteristics of each method in comparison to an ideal structured illumination 3D measurement system. It was seen that other proposals not utilizing phase-based methods implementing unconventional projectors and temporal pattern projection have been shown to be capable of both fast acquisition speed and independent pixel-by-pixel reconstruction without introducing spatial

measurability limitations. However, a stereo-camera setup is still required. There has not been significant progress demonstrated in reducing to a monocular view.

A monocular configuration utilizing unconventional array projectors such as the MAAP is therefore a particularly challenging obstacle. Not only is achieving a single camera setup an academically interesting problem, but it has also directly applicable benefits of decreased setup complications as well as decreasing cost, weight, and form factor of such a metrology system.

Chapter 6.

Intrinsic calibration of a Multi-Aperture Array Projector

6.1. State-of-the-art

Since a MAAP uses a different pattern modality, phase-to-height equations as used in FPP are not applicable. In addition, since a MAAP does not utilize a typical digital projection chip or projection optics, the established pinhole modelled means of intrinsic calibration cannot be implemented to achieve 3D reconstruction through triangulation. Applying the pinhole model to the MAAP to describe the optical projection function would not be possible as it does not describe the MAAP's actual operating function. No geometric parameters physically exist such as focal length or principal point of a projection chip to implement the intrinsic matrix. No inherent projection center is responsible for illumination. Besides this, if one were to successfully geometrically model a single MAAP projection channel, a challenging extrinsic calibration is still needed to determine the rotation and translation between each aperture in order to transform them all into a single world coordinate system.

As seen in the flowchart of Fig. 6.1 first introduced in Sec. 3.2 (pg. 18), the pipeline to generate 3D reconstructions requires intrinsic and extrinsic calibration. As the MAAP cannot be intrinsically calibrated by existing methods that rely on the pinhole model, a novel intrinsic calibration process needs to be developed.

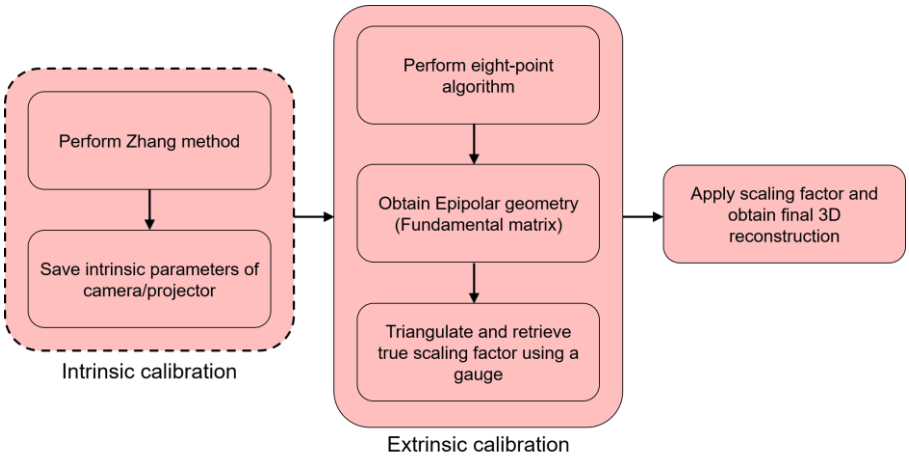

Fig. 6.1 *Flowchart depicting missing components of the calibration pipeline. The processes calibrating triangulating pinhole devices that cannot be applied here are in red. This chapter's focus is depicted in the dashed line.*

Alternate analytical models without utilizing explicitly the intrinsic matrix parameters have been proposed in traditional monocular view fringe projection profilometry. The most representative work on this was presented by Du et al. [54] and Zhang et al. [55] where a direct non-linear least squares approach was proposed. An alternate mathematically parameterized equation was developed that directly obtains the depth of an object point with respect to a reference plane using its observed phase and pixel coordinates in the camera image plane. Using at least two block gauges of different but known uniform height, twelve parameter constants are determined through minimization of the squared error with the block gauge heights through non-linear least squares optimization. While not explicitly using conventional geometric parameters of the pinhole model in the equation, the twelve parameters of the direct least squares method are still derived from the assumption that both the camera and projector are pinhole modelled. This direct approach therefore cannot be adapted to the MAAP. Also, while twelve parameter constants are required when using a single aperture chip-based projector, the number of these constants could become computationally prohibitive for considering every distinct projection channel in a MAAP.

Other methods to perform intrinsic calibration independent of any analytic geometric model have also been developed in traditional fringe projection profilometry. Empirical methods known as phase-to-height/phase-to-depth look-up-table mapping have been previously proposed while utilizing conventional digital projectors incapable of high projection rates. A test plane is incremented depth-wise through the measurement volume. At each test plane increment, the camera images the projected pattern incident on the test plane. In many of the proposals, the test plane must contain a grid structure to create a coordinate system in the *x*-*y* plane [34, 56–58]. To obtain phase mappings along an incoming light ray without a depth-wise 2π discontinuity, the aforementioned studies employ temporal phase unwrapping methods. Since the phase mappings along a ray belong to each individual pixel independently, the problems of measuring surface discontinuities or high surface gradients do not arise.

Utilizing the empirical phase-to-depth look-up-table mapping method, Fujigaki et al. and Morimoto et al. calibrated a high-speed LED-based array projector in which a Ronchi ruling acts as a transmission mask to construct laterally displaced purely sinusoidal fringes [48–50, 58]. This is a concept similar to the MAAP. It too was capable of high projection rates. However, because of the lateral displacement of

patterns, the phase-shift depends on depth. Thus, no established phase unwrapping method could be applied to resolve depth-wise phase ambiguity. Consequently, it was limited to 3D measurement with limited measurement depth. The largest measurement volume range of depth experimentally demonstrated was 15 mm [58]. To largen the depth range, Guo et al. proposed a two frequency phase-based approach [59]. A low frequency such that no phase ambiguity is present through the defined measurement volume acts as a complementary phase value to a high frequency. This also enabled spatial discontinuities to be measurable. While the measurement volume depth was widened to 70 mm, a conventional digital projector was utilized. With a projection center, this enabled phase-to-height/phase-to-depth calculations as per Eqn. 2.5 (pg. 11). Thus, this method is unable to be implemented with the MAAP.

In this chapter, a method is presented that intrinsically calibrates the MAAP without depending on a geometrically parameterized model. This would enable 3D measurements to be generated using only a single camera. Using the aperiodic fringe pattern modality, surface discontinuities and high gradients are still measurable. Each image pixel shall be capable of generating a 3D point without dependency on its neighbors. The proposed method does not require an additional step after illumination as with phase unwrapping. The proposed method does not limit measurement depth in comparison to the phase shifting method presented by Fujigaki et al. [48–50].

6.2. Mapping intensity sequence to depth

Due to the inconveniences present in the phase-shifting approaches, an alternate pattern modality should be used for depth mapping.

As stated in Chapter 4 (pg. 22) the MAAP projects instead a series of six aperiodic sinusoidal fringes. Rather than relying on phase-shifting methods to encode each depth along a light ray with a phase value, it is proposed in this chapter that the series of projected aperiodic fringe patterns encode each depth along a ray with a temporal gray-value intensity sequence when imaging with a grayscale camera. Temporal intensity sequence encoding provides greater possibilities to uniquely code each depth in comparison to the phase approach where duplicate phase values arise depth-wise.

Like phase-to-depth look-up-table mapping methods, this study also utilizes a test plane that is incremented depth-wise through the measurement volume. Depth is

defined as the orthogonal distance between the camera pinhole and the test plane surface. At each increment position, the test plane is imaged for each aperiodic fringe pattern projected that is incident upon it. One can think of this process as taking discretized slices of the illumination frustum produced by each projection channel of the MAAP. As seen in Fig. 6.2a, assuming a pinhole modelled camera, the measurement volume becomes discretized such that each intersection point between a test plane position and a camera ray is encoded by a gray-value intensity sequence. This creates a volumetric 3D grid of data such that

$$(X,Y,Z)_P \ni \left\{ I_j(X,Y,Z) \;\middle|\; j \in \{1 \dots N\} \right\} \tag{6.1}$$

where $(X,Y,Z)_P$ is the intersection point with spatial coordinates (X,Y,Z) between a test plane increment and a light ray, and I is the set of gray-value intensities at (X,Y,Z) for each pattern j in the sequence. The number of patterns in this study is fixed to $N = 6$, equal to the number of projection channels. With this process, a 3D gray-value intensity sequence map of the illumination propagating through the measurement volume has been created.

As seen in Fig. 6.2b, an object placed in the measurement volume is therefore subject to the same illumination sequence. Thus, the object intersects the intensity map such that each surface point has a certain depth in the discretized measurement volume. The 3D coordinates of the surface point can therefore be found by calculating the intersection between the point's corresponding depth in the intensity map and the

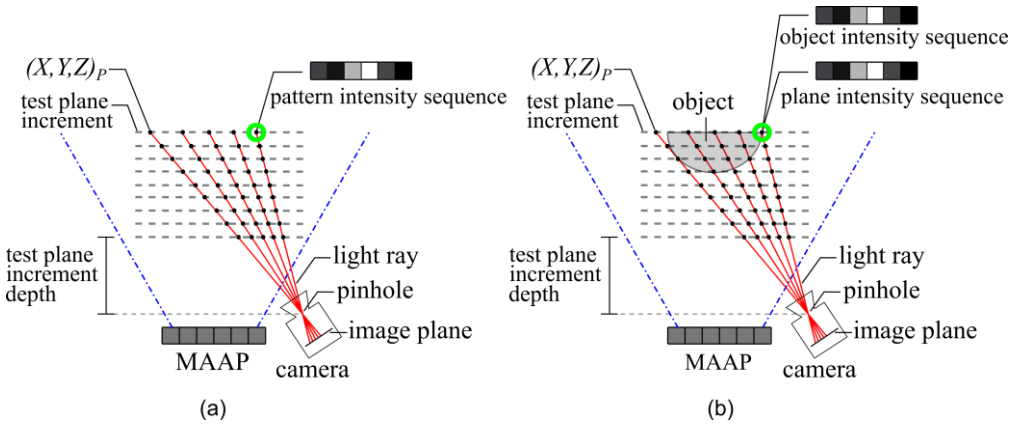

Fig. 6.2 *(a) Volumetric 3D grid of temporal gray-value intensity sequence data generated by imaging each projected pattern incident on a test plane at discrete depth slices. Each data point in the 3D grid $(X,Y,Z)_P$ is the intersection between a depth slice and light ray. (b) An object intersecting the volumetric 3D grid [39].*

directional vector of a light ray incoming to the camera.

6.3. Depth correspondence identification

If an object point and a test plane increment point intersect the same position in the measurement volume, they must therefore share the same gray-value intensity sequence. Thus, a similarity measure is required to identify which depth sequence is the same as the object point sequence. The normalized cross-correlation is used to compare the similarity between two gray-value intensity sequences

$$\rho(O,P) = \frac{\sum_{j=1}^{N}(O_j - \bar{O}) \cdot (P_j - \bar{P})}{\sqrt{\sum_{j=1}^{N}\left(O_j - \bar{O}\right)^2}\sqrt{\sum_{j=1}^{N}\left(P_j - \bar{P}\right)^2}} \tag{6.2}$$

where O_j is the gray-value intensity of a pixel in image j of the object, P_j is the gray-value intensity of a pixel in image j of a test plane increment, $\bar{O}$ is the mean intensity of the pixel over N = 6 object images, and $\bar{P}$ is the mean intensity of the pixel over N = 6 test plane increment images. A cross-correlation coefficient ρ = 1 indicates a perfect match between the two sequences while ρ = -1 indicates anti-correlated behavior.

For all 3D points in space lying along the same light ray, their 2D image projections onto the imaging chip are all incident on the same pixel (α,β). Let O be the intensity sequence at (α,β) in the object image and P be the intensity sequence at (α,β) in a test plane increment image. The search for the corresponding test plane increment D can then be expressed as

$$D := \underset{d:d\in\{0,\ldots,N\}}{\operatorname{argmax}} \rho(O,P_d) \tag{6.3}$$

where P_d is P for the d^{th} test plane increment over N total increments of the measurement volume. Thus, as seen in Fig. 6.3, to find the corresponding test plane increment D of an object point $(X,Y,Z)_O$, all test plane points $(X,Y,Z)_P$ that lie along a light ray must be searched to find D maximizing ρ.

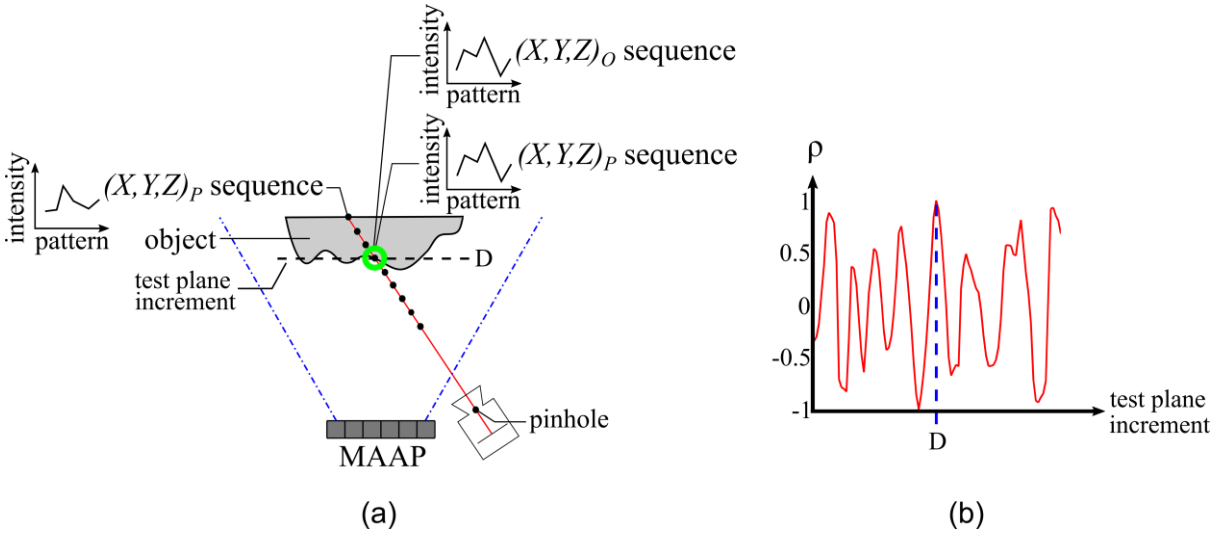

Fig. 6.3 *Determining the corresponding test plane increment D of an object point $(X,Y,Z)_O$ by searching each test plane point $(X,Y,Z)_P$ along a light ray, (b) finding which D maximizes the normalized cross-correlation coefficient with the object point's intensity sequence [39].*

6.4. Obtaining 3D coordinates

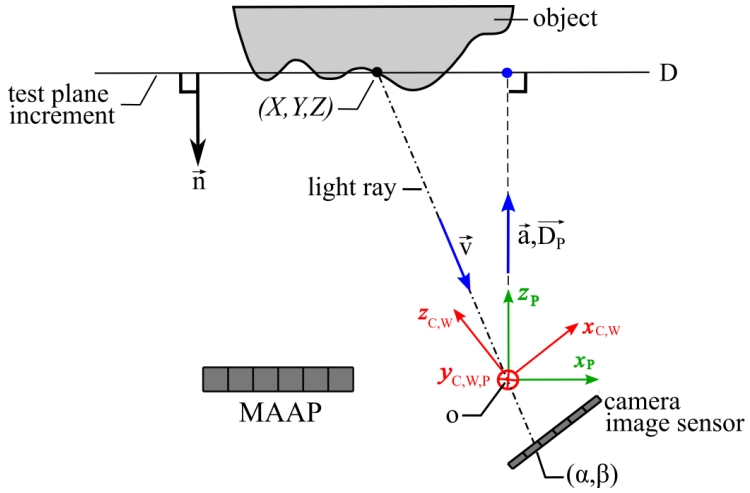

Fig. 6.4 *Determining the 3D coordinates (X,Y,Z) of a point on the object point's corresponding test plane increment D by calculating the intersection between the ray's direction vector $\vec{v}$ and D. (α,β) is the pixel the light ray of $\vec{v}$ is incident. The coordinate system of the plane is denoted with subscript P and the coordinate system of the camera/world is denoted with subscript C,W. The origin o of the two coordinates systems is shared at the camera pinhole. $\vec{a}$ is a vector in the camera/world coordinate system from the origin to an arbitrary point on the test plane at D whose direction can be arbitrarily set the same as the vector $\overrightarrow{D_P}$ from the origin to D in the plane coordinate system [39].*

A world coordinate system in which all measurements are made with respect to needs to be defined. Assuming the camera operates under the pinhole model, the camera's pinhole is chosen as the origin. The world coordinate system has the same orientation as the camera reference frame, which shall be called $S_{C,W}$. As seen in Fig. 6.4, let us

define that the test plane itself has its own coordinate system S_P. S_P is defined such that its *z*-axis shall be antiparallel to the test plane surface normal so that $\overrightarrow{n_P} = \overrightarrow{(0,0,-1)}$.

First assume that by using a linear translation stage, the increment distance between test plane increments is constant. There is also the assumption that the test plane orientation is fixed such that the difference in orientation between each test plane increment is negligible. Thus, the depth of any corresponding test plane increment $\overrightarrow{\boldsymbol{D_P}}$ in the plane coordinate system can be obtained such that

$$\overrightarrow{D_P} = \overrightarrow{t_{0_P}} + k\overrightarrow{(0,0,D)} \tag{6.4}$$

where $\overrightarrow{t_{0_P}} = \overrightarrow{(0,0,t_{0_P})}$ is the orthogonal translation vector from the pinhole to the surface of the first test plane increment along the *z*-axis of S_P, and $k\overrightarrow{(0,0,D)}$ is also an orthogonal vector along the *z*-axis of S_P from the first test plane increment to the corresponding test plane increment *D* such that *k* is the constant increment distance.

With the discretization of the measurement volume intensity map due to the test plane being incremented at discrete depth slices, the increment *D* that maximizes ρ of Eqn. 6.3 could lie between the discrete depth samples. From Fig. 6.3, first say that the coarse estimate of *D* is initially identified. A fine estimate of *D* is then obtained by using the nearest depth neighbors: *D-1, D, D+1, D+2*. Using the gray-value intensities *I* of each pattern *j* of the nearest depth neighbors $I_j(D-1)$, $I_j(D)$, $I_j(D+1)$, $I_j(D+2)$ along a light ray, the gray-value intensities of each pattern between the nearest depth neighbors are obtained through cubic interpolation. The search for the fine estimate of *D* that maximizes ρ is then carried out again along the light ray, searching only between the nearest neighbor depths using the interpolated gray-value intensities. D_P is then calculated using Eqn. 6.4.

From Fig. 6.4, where *o* is the origin of $S_{C,W}$, the 3D coordinates *(X,Y,Z)* of the object point with corresponding depth D_P can be calculated by

$$\gamma = \frac{\vec{n} \cdot \vec{a}}{\vec{n} \cdot \vec{v}}$$
$$(X, Y, Z) = \gamma \vec{v} \tag{6.5}$$

where γ is a scalar, $\vec{n}$ is the test plane surface normal 3D vector in the world coordinate system $S_{C,W}$, $\vec{a}$ is an arbitrary vector from o to an arbitrary 3D point on the test plane in $S_{C,W}$. $\vec{v}$ is the 3D directional vector of the light ray incoming to the camera which is already in $S_{C,W}$.

Obtaining $\vec{n}$ only requires a rotational transformation such that $\vec{n} = R\overrightarrow{n_P}$. R is the 3×3 rotation matrix that aligns S_P to $S_{C,W}$. If $\overrightarrow{a_P}$ is the arbitrary vector in the reference frame of S_P, considering an infinite plane, a convenient choice for the endpoint on the test plane at corresponding depth D_P would be $(0,0,D_P)$. To convert the arbitrary vector into $S_{C,W}$, the same transformation is applied such that $\vec{a} = R\overrightarrow{a_p}$. The directional vector $\vec{v}$ is determined from the camera's intrinsic parameters of the pinhole model assumption and already in $S_{C,W}$.

To determine the rotation R, a calibration grid is manually placed in the same orientation as the first test plane increment and is imaged by the camera. The calibration grid coordinate system has the same rotational orientation as the test plane reference frame. The calibration grid provides feature points in which a perspective-n-point algorithm can be used to determine the pose aligning the reference frame of the control points on the calibration grid and the reference frame of their corresponding projections on the image plane [60]. To determine the endpoint $(0,0,D_P)$ of $\overrightarrow{a_p}$, the orthogonal distance between the first test plane increment and the camera pinhole is manually measured to obtain the orthogonal translation vector $\overrightarrow{t_{0_P}}$ and inputted into Eqn. 6.4. $\overrightarrow{a_p}$ is then converted to $S_{C,W}$ using the obtained R. In summary, substituting R and $\overrightarrow{t_{0_P}}$ into Eqn. 6.5, gives the equation

$$\gamma = \frac{\vec{n} \cdot \vec{a}}{\vec{n} \cdot \vec{v}} = \frac{R\overrightarrow{n_P} \cdot R\overrightarrow{(0,0,t_{0_P} + kD)}}{R\overrightarrow{n_P} \cdot \vec{v}} \tag{6.6}$$

As the set of intensities of each pixel and calculation of 3D coordinates is independent of one another, this makes the process ideal for fast computation in heterogeneous computing platforms such as Graphics Processing Units (GPU) as the computation is 'embarrassingly parallel'. Each pixel therefore has a GPU thread dedicated to generating a 3D point independently in parallel with all other pixels.

6.5. Experimental results

6.5.1. Setup

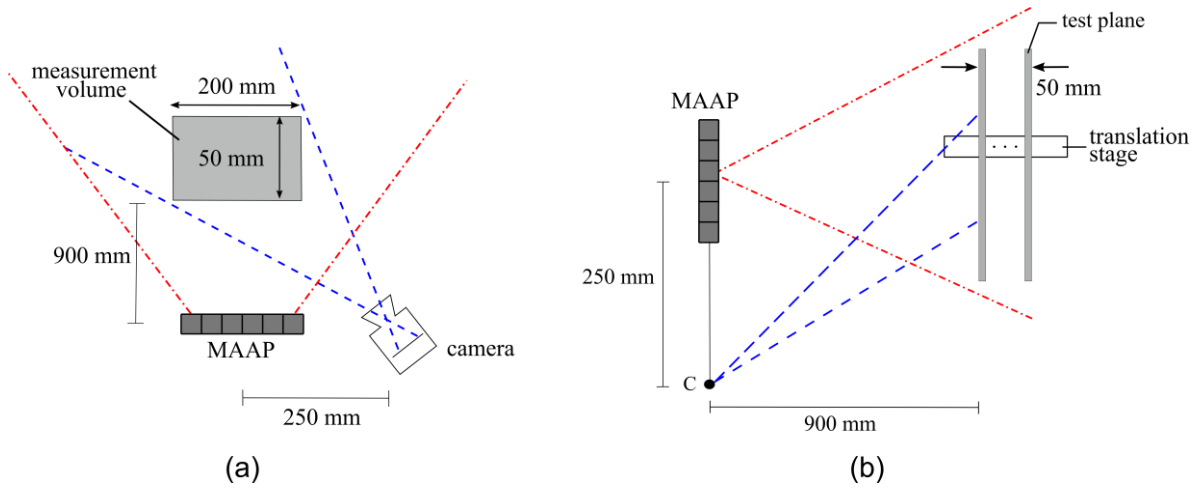

Fig. 6.5 *(a) Experimental setup configuration, (b) translation stage scanning the test plane through the measurement volume [39].*

To test the proposed MAAP calibration method to obtain 3D measurements using a single camera, the setup shown in Fig. 6.5 was created. The camera was placed 250 mm away from the MAAP along a common baseline connecting the center of the MAAP and the camera pinhole. The measurement volume was defined to begin 900 mm away from the common baseline and had dimensions 200×200×50 (H×W×D\mm). The imaging camera used was an AVT Pike F-032 with objective lens of focal length f = 25 mm, a quadratic pixel pitch of Δp = 7.4 µm, and a resolution of 640×480 px. imaging in 16-bit grayscale mode. The camera was assumed to be pinhole modelled and intrinsically calibrated using the Zhang method [10]. It was found that the distortion parameters were minimal, and any nonlinear geometric distortions were therefore omitted. The optical axis of the camera was aligned with the center of the measurement volume. The test plane was placed in an orientation perpendicular to the direction of illumination. From the perspective-n-point algorithm, the unitless 3×3 rotation matrix R aligning the plane coordinate system and the camera/world coordinate system was found to be

$$R = \begin{pmatrix} 0.9611 & -0.0167 & 0.2756 \\ 0.0181 & 0.9998 & -0.0023 \\ -0.2755 & 0.0072 & 0.9612 \end{pmatrix} \tag{6.7}$$

The test plane filled the entire field of view of the camera and was incremented using a linear translation stage in 0.5 mm increments resulting in a total of 101 test plane increments. The linear translation stage has a positioning precision of ±0.020 mm. It was found that the deviation of test plane surface normals was < 0.04° with respect to the first increment normal. To obtain the fine depth estimate, the interpolation of gray-value intensities is done in 0.1 mm interpolating sub-steps.

6.5.2. 3D measurement performance

A surface discontinuity is a gap that is not encoded with a gray-value intensity sequence. This should result in a low cross-correlation coefficient between the pixels of the object scene and the pixels of each test plane increment due to the presence of only imaging noise. By calculating 3D coordinates only for pixels that achieve a $\rho \geq 0.99$ threshold, no corresponding depth should be identified and 3D points should not be generated where discontinuities are present.

The ability to 3D reconstruct a discontinuous scene with isolated surfaces is successfully demonstrated in Fig. 6.6. A white scattering object in the shape of a bust of German poet Friedrich Schiller that occupies a volume of ~150×100×100 (H×W×D/mm) along with two separate white scattering spheres with radii of 15 mm were placed in the measurement volume. The missing hair features in the reconstructed point cloud of Fig. 6.6b are due to these features being outside the measurement volume. Missing points on the right side of the nose are due to occlusion of the illumination such that the nose casts a shadow. It can also be seen that there are missing points occurring along columns. As the fringes are imaged nearly parallel

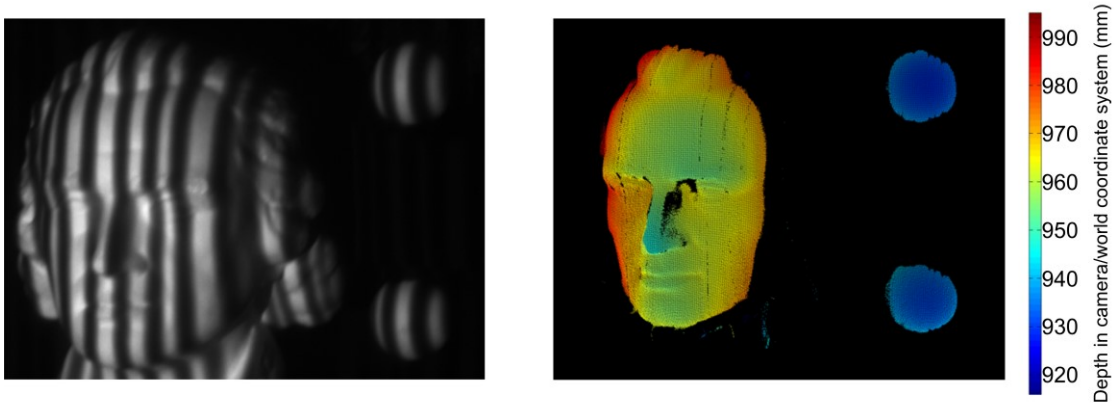

Fig. 6.6 *Test object of bust and two separate spheres: (a) snapshot during projection of aperiodic fringes, (b) resulting 3D reconstruction of the discontinuous surfaces [39].*

to the pixel columns of the imaging chip, pixel columns have similar gray-value intensity sequences. These missing points occur possibly because of low variation in the gray-value intensity sequence, leaving them susceptible to corruption by imaging noise and not reaching the cross-correlation threshold. This corruption would also cause the identification of false corresponding depths, resulting in 3D point outliers.

To evaluate 3D measurement performance quantitatively, a white ceramic scattering pyramidal frustum shaped gauge containing five distinct planes was placed into the measurement volume as seen in Fig. 6.7. The planar surfaces S_n had a certified peak-to-valley surface deviation of < 7 µm. The certified line segments L_n and certified angles A_n are also shown. For each 3D reconstructed point cloud of S_n, a plane is fitted and the orthogonal deviation of each point to the best-fit plane is determined [61]. Points with deviation > 1 mm from a best-fit plane are filtered. The number of points filtered was minimal as > 99% of points remained. The deviation σ_{plane} of remaining points to the best-fit plane is a measure of 3D measurement surface precision.

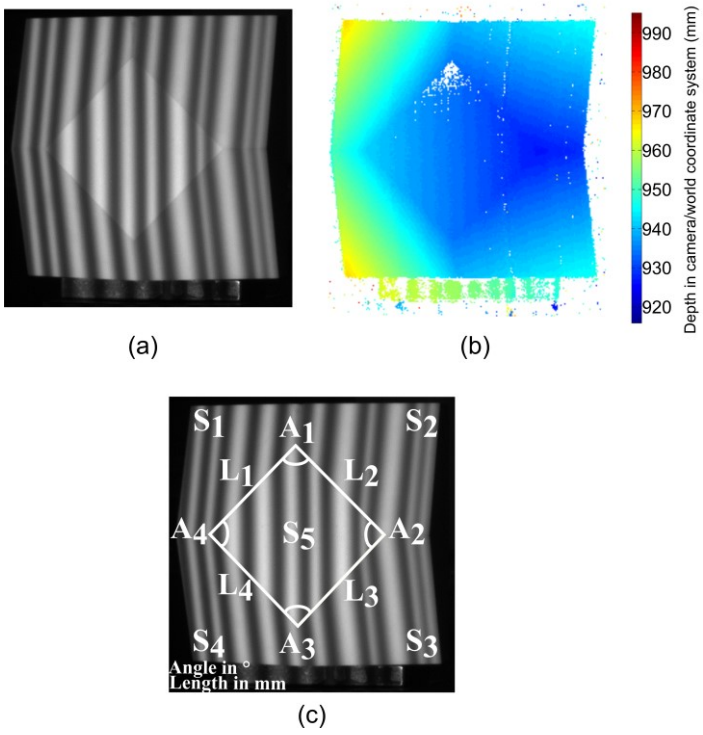

Fig. 6.7 *Pyramidal frustum gauge: (a) snapshot during projection of aperiodic fringes, (b) resulting 3D reconstruction, (c) layout of certified ground truth quantities of surface S, angles A, and lengths L [39].*

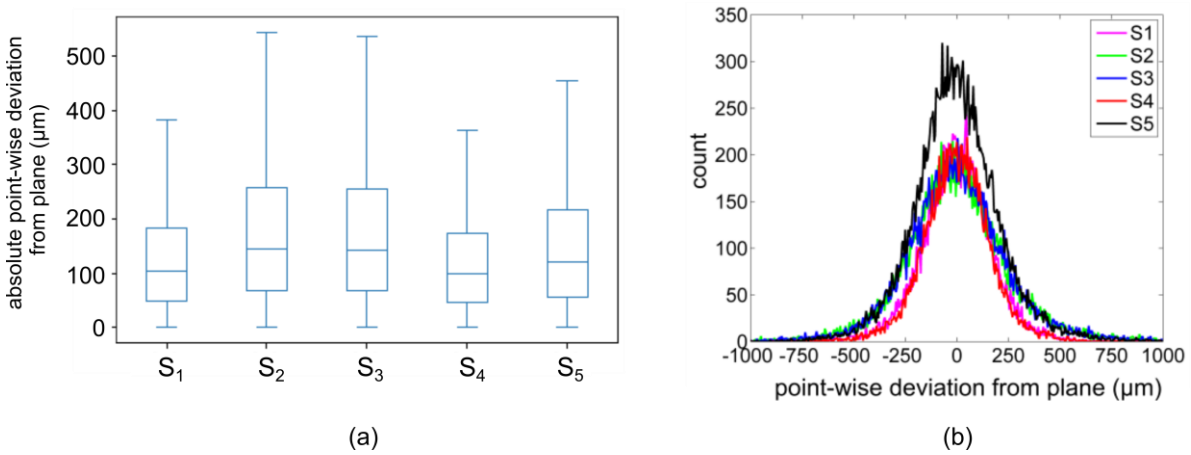

Fig. 6.8 *(a) Quantiles of the absolute deviations of reconstructed point-wise deviations from the best-fit planes of the pyramidal frustum gauge, (b) distribution of the deviations*

Seen in Table 6.1 are the quantiles of the absolute deviations for each plane S_n and plotted in Fig. 6.8a. They define the percentile of reconstructed points whose absolute deviation from the best-fit plane falls below. In Fig. 6.8b is the distribution of the deviations from the best-fit plane. It can be seen from Table 6.1 and Fig. 6.8 that the distribution of σ_{plane} for each plane nearly overlap, thus demonstrating the consistency of 3D measurement surface precision of the proposed method.

Plane	25th percentile σ_{plane} (µm)	Median σ_{plane} (µm)	75th percentile σ_{plane} (µm)
S_1	49	104	183
S_2	68	145	258
S_3	67	144	255
S_4	47	100	173
S_5	57	122	217

Table 6.1 *Quantiles of reconstructed absolute point-wise deviations from best-fit plane*.

Shown in Table 6.2 are measures of the 3D measurement accuracy. The length error Δl and angle error $\Delta°$ are calculated as the absolute difference between the measured value and the certified value. As the angles and lengths are defined by intersecting best-fit planes, Δl and $\Delta°$ arise due to the propagation of error of the determined normal of best-fit planes from their ground truth. It can be observed that angle A_2 and length L_2 have the greatest deviations from their ground truth values. This is possibly due to these geometries being defined by planes S_2, S_3, and S_5, whose σ_{plane} quantiles are greater than the other surfaces. Their reduced surface precision

Angle	Certified nominal ground truth (°)	$\Delta°$	Line	Certified nominal ground truth (mm)	Δl (µm)
A_1	90.00	0.07	L_1	43.955	325
A_2	89.58	0.79	L_2	43.939	886
A_3	90.01	0.29	L_3	43.960	397
A_4	89.58	0.25	L_4	43.946	534

Table 6.2 *Absolute angle deviation $\Delta°$ and absolute length error Δl from the angles and lengths of the pyramidal frustum gauge.*

leads to the increased likelihood that the determined best-fit plane normal differs more considerably from the ground truth normal.

As seen in Fig. 6.9, to obtain measurement performance for a non-planar surface, a white ceramic scattering sphere with ground truth radius r_{true} = 50 mm was placed in the measurement volume. Similarly, a sphere is fitted to the 3D reconstructed point cloud and the difference between r_{true} and each point's distance from the best-fit center is taken [61]. Points with difference > 1 mm are filtered. The number of points filtered was minimal as > 99% of points remained. The difference of all remaining points is called the spherical surface deviation σ_{sphere}. Seen in Table 6.3 are the quantiles of σ_{sphere} while in Fig. 6.10 is the distribution of the deviations from the best-fit sphere. Another measurement accuracy assessment is the best-fit sphere radius in comparison to r_{true}. The absolute deviation between the two, $AD = |r_{true} - r_{measured}|$, was found to be AD = 333 µm. Observing the results of the sphere and the planes show that there is comparable 3D measurement performance.

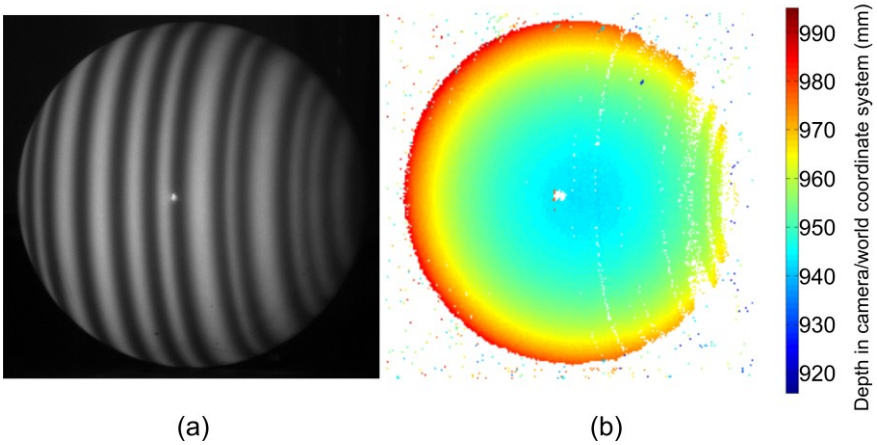

Fig. 6.9 *Sphere gauge: (a) snapshot during projection of aperiodic fringes with (b) resulting 3D reconstruction [39].*

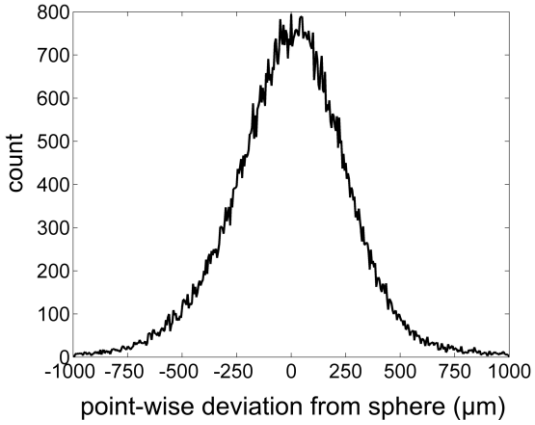

Fig. 6.10 *Distribution of point-wise deviations of the 3D reconstruction from the best-fit sphere surface of the sphere gauge shown in Fig. 6.9a.*

25th percentile σ_{sphere} (µm)	Median σ_{sphere} (µm)	75th percentile σ_{sphere} (µm)
77	166	291

Table 6.3 *Quantiles of point-wise deviations from best-fit sphere*.

6.6. Temporally repeatable speckle projector: a proof-of-concept

Due to the proposed single camera method being model-less, the method should be able to be utilized with any pattern modality. It is demonstrated only as a proof-of-concept that the proposed method also works on the projection of objective speckle patterns whose intensity distribution varies spatially in two dimensions under the condition that the speckle patterns are temporally repeatable.

As mentioned previously in Sec. 5.5 (pg. 29), objective speckles were introduced as a pattern modality also capable of high-speed 3D measurement where a coherent light source of a laser beam is incident on a rotating diffusing medium. An alternate method rapidly deflects the angle of incidence of the beam on a static diffusing medium. As with the MAAP, this type of projection system is chip-less and cannot be conventionally intrinsically calibrated as the pinhole model cannot be assumed. Thus, the use of such a speckle projector in previous studies was implemented with a synchronized triangulating stereo-camera setup where the correspondence problem was solved between the two views [14, 21].

However, temporally random pattern sequences cannot be used for the proposed single camera method as the gray-value intensity distribution propagated through the measurement volume would not be consistent. For a temporally random

projection set, a sequence of gray-value intensities recorded for a duration of time T would be I_1, I_2, I_3, …, I_N. At another time duration T' the gray-value intensities recorded would be I_1', I_2', I_3',…, I_N'. As a result, determining depth correspondence with the proposed single camera method would not be possible for an object point since there would be no matching intensity sequences as I_1, I_2, I_3,…,$I_N \neq I_1'$, I_2', I_3',…, I_N'. For the proposed single camera method to be utilized with objective speckles, a temporally repeating sequence of speckle patterns needs to be generated such that I_1, I_2, I_3,…,$I_N = I_1'$, I_2', I_3',…, I_N'. Temporal synchronization of the camera to repeatedly image at the precise location in which the laser beam is incident on the diffuser would be highly challenging due to the risk of unstable or imprecise camera hardware timing. Alternatively, a precisely rotating actuator may be possible, but it is currently unknown the level of rotation precision needed to be sustained.

As seen in Fig. 6.11, a projection system designed in collaboration with Stark and Weigel and fabricated by Weigel was developed [62]. The system is capable of producing a temporally repeating set of speckle patterns by directly attaching several single-mode fibers to a static diffusing medium. The use of single-mode fibers coupled with a diffusing medium offers stable output speckle intensity distributions under fiber bending or twisting. Although objective speckles can be generated using multi-mode fibers alone without having the laser beam propagate through a diffusing medium, the output intensity distribution is highly sensitive to external perturbations due to mode interference under changing fiber conditions [63].

As seen in Fig. 6.12a, a galvanometric mirror system was used to focus a 532-nm laser into each individual fiber. Different objective speckle patterns are projected

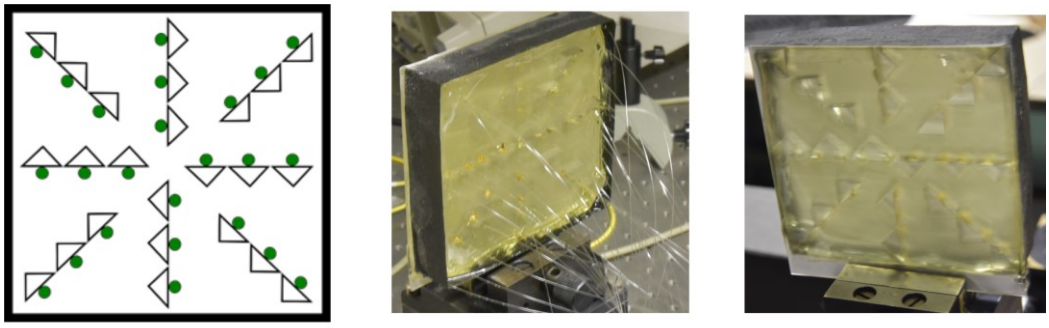

Fig. 6.11 *(a) Schematic of fibers (green) and arrangement of glass block position stabilizers constituting the repeatable speckle projector, (b) image of backside, and (c) image of front [62].*

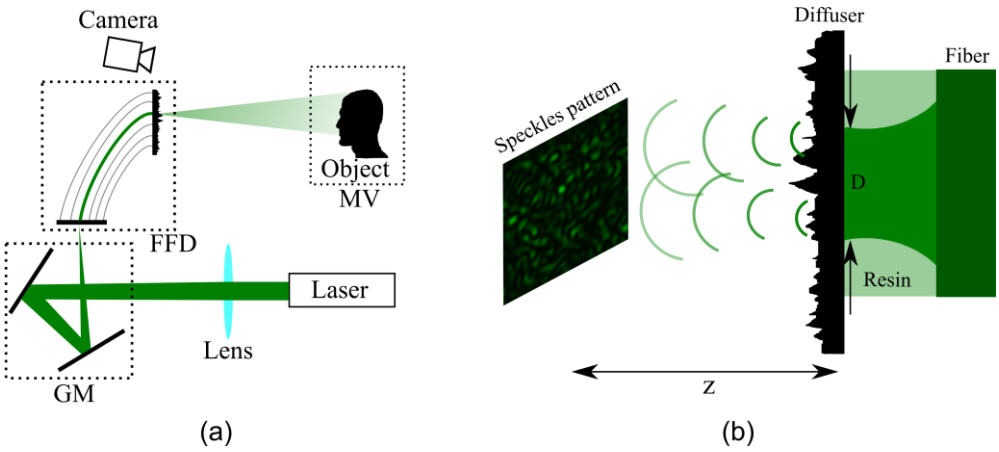

Fig. 6.12 *(a) Scheme of the setup with measurement volume (MV), galvanometric mirror (GM), and fibers fixed to diffuser (FFD), (b) production of a repeatable speckle pattern with a fiber affixed to diffuser plate. D is the diameter of the beam spot incident on the diffuser and z is the propagation distance [62].*

depending on which fiber the mirror system addresses. By addressing each fiber in a specific order, a temporally repeating sequence of objective speckles is generated. As seen in Fig. 6.12b, single-mode fibers were affixed to a diffusing medium by applying an epoxy resin layer. More than two dozen fibers were attached with the aid of glass blocks to stabilize the fiber position so that each fiber end face is roughly parallel to the diffusing medium surface. As every fiber end is fixed, the beam exiting from each fiber is incident on the same diffuser location. Therefore, the resulting patterns are unlikely to change due to this mechanical fixation.

Shown in Fig. 6.13a is an object in the shape of a bird with dimensions of ~300×100×100 (H×W×D\mm). In Fig. 6.13b is the object illuminated. The generated 3D reconstruction utilizing six objective speckle patterns is shown in Fig. 6.13c. In comparison to the MAAP which uses the same number of patterns in a temporally repeatable sequence, the repeatable speckle projector measurement performance is worse due to the presence of subjective speckle noise. The surface deviation of the 3D reconstruction overcomes the finer spatial details such that the wing feathers and eye of the bird are not well discernible. Subjective speckle noise is a phenomenon that arises simply from observing coherent illumination incident on a surface. To reduce subjective speckle noise, a larger camera aperture can be utilized. While more sophisticated means of subjective speckle suppression exist, it is outside the scope of this work and will not be discussed in this work but can be reviewed in the literature [64].

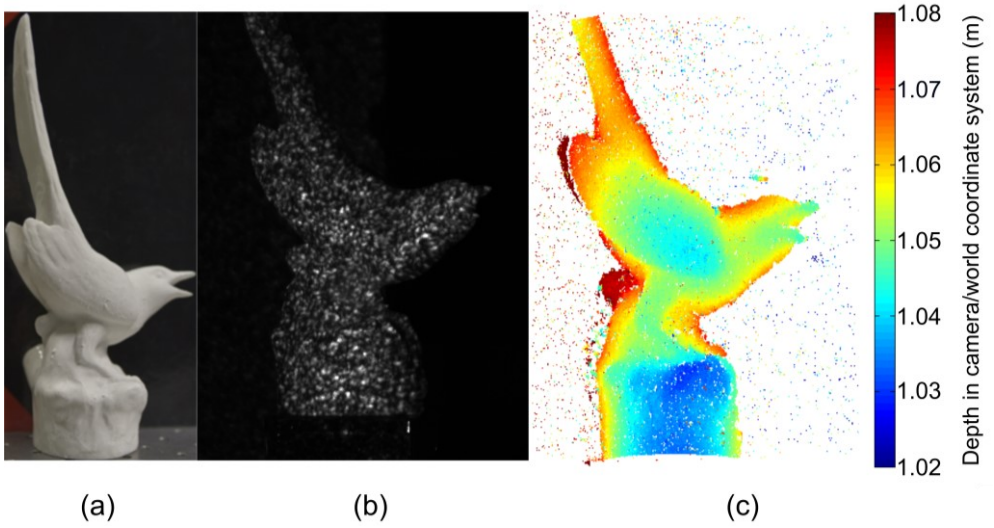

Fig. 6.13 *(a) Image of bird test object with (b) bird under speckle illumination using proposed repeatable speckle projector, and (c) single-camera reconstructed 3D point cloud [62].*

6.7. Conclusion

In this chapter, a method was proposed and experimentally demonstrated to intrinsically calibrate the MAAP without a geometrically parameterized model, thus enabling 3D reconstructions to be made using only a single camera. The method requires depth sampling of each projection channel using a test plane incremented through the measurement volume. This is as if discrete slices are imaged of each illumination frustum. This creates a volumetric grid that stores the temporal gray-value intensity sequence to depth mapping through the measurement volume. For every object point encoded by a temporal intensity sequence, its corresponding depth is determined through a search along the light ray scattered from the object point to determine which discrete depth slice contains the temporal intensity sequence maximizing the normalized cross-correlation coefficient. The object point's 3D coordinates are then calculated by determining the intersection between the determined corresponding depth and the directional vector of the light ray.

In comparison with other single camera methods, the proposed method does not introduce scene measurability limitations. Each individual pixel is capable of generating a 3D point. As the proposal does not use a phase-based method, it does not inherit any of the issues when spatial discontinuities are present. No additional work is required after illumination as involved with phase unwrapping. There would be no problems either with high surface gradients present in complex surfaces as experienced by Fourier Transform Profilometry. Since each individual pixel contains

an intensity sequence independent of its neighbors, surface measurability would not be spatially limited compared to spatial codeword based techniques.

Compared with past studies using similar array-projection devices projecting purely sinusoidal fringes [48, 49, 58], a greater measurement volume depth of ~3 times was experimentally demonstrated. Nevertheless, the method itself does not limit the measurement volume depth. Limits on measurement volume depth is what the user desires it to be in terms of how many test plane increments shall be taken, as well as GPU memory to store the volumetric 3D grid of intensity sequence data for computation.

This chapter also successfully showed that the proposed single camera method is generalized and can generate 3D reconstructions for different pattern modalities such as objective speckle patterns, whose intensity distribution varies spatially in two dimensions. However, the projected illumination must be temporally repeating. From this caveat, a proof-of-concept of a novel repeatable speckle projection system was created, where only temporally random speckle generation was exhibited in prior studies using a stereo-camera setup [14, 21].

Chapter 7.

Extrinsic calibration of a Multi-Aperture Array Projector-single camera system

7.1. State-of-the-art

From the previous chapter, the proposed intrinsic calibration process of the MAAP was described. This process required a test plane incremented through the measurement volume. At each increment position, the camera images the pattern projected by each projection channel as if imaging discrete depth slices of each illumination frustum. To obtain the 3D coordinates of an object in the measurement volume, the corresponding test plane increment depth of every object point needs to be identified in order to calculate its intersection with the scattered light ray. However, the test plane needs to be transformed to the camera/world coordinate system. Therefore, the purpose of extrinsic calibration is to determine the rotation and translation of the test plane with respect to the camera. From Sec. 6.4 (pg. 37), the rotation was determined by manually placing a calibration grid in an orientation similar to the test plane. The calibration grid provides feature points in which a perspective-n-point algorithm was used to determine the pose aligning the reference frame of the control points on the grid and the reference frame of their corresponding projections on the image plane [60]. Translation was manually measured as the orthogonal distance between the test plane surface and the camera. Due to manual measurement of these parameters, additional errors can be introduced to the 3D measurement. Hence, alternative means of extrinsic calibration of a single camera 3D sensor utilizing a MAAP is needed without hands-on intervention.

The need for an extrinsic calibration method in the pipeline to generate 3D reconstructions is reflected in Fig. 7.1. Fig. 7.1 was first introduced in Sec. 3.2 (pg. 18). Without the propagation of error due to manual extrinsic calibration, the result of the calibration pipeline is a final 3D reconstruction that should agree better with the ground truth surface.

Methods of extrinsic calibration have been thoroughly explored and applied to chip-based digital projection systems in monocular view Fringe Projection Profilometry. Rather than using the phase-to-depth equation with respect to a

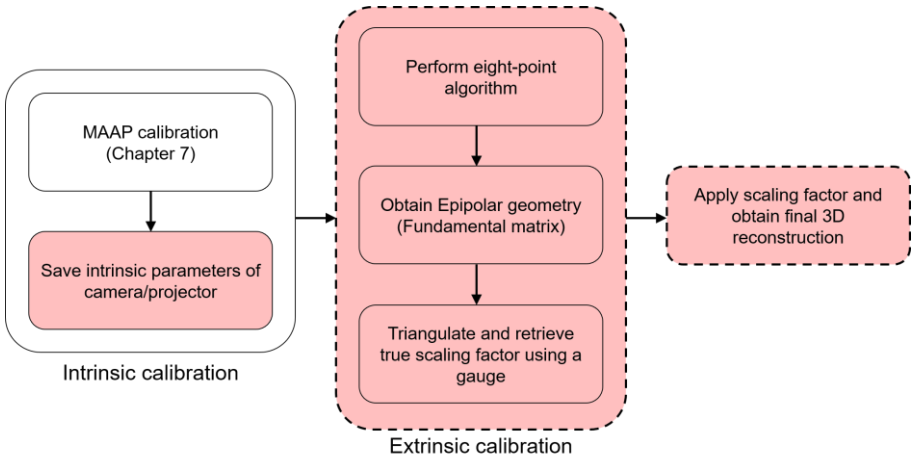

Fig. 7.1 *Flowchart depicting missing components of the calibration pipeline. The processes calibrating triangulating pinhole devices that cannot be applied here are in red. This chapter's focus is depicted in the dashed line.*

reference plane as in Eqn. 2.5 (pg. 11), by treating the geometric optical function of the projector as an inverse camera with its own focal length and principal point [31–33], an approach is to project onto a planar grid a set of phase-shifted sinusoidal fringe patterns to determine the phase values projected onto each feature point of the grid [26, 33, 65, 66, 67]. In this way, corresponding points between the camera image plane and projector projection plane can be determined with respect to a common world coordinate system defined on the planar grid. The extrinsic calibration is typically accomplished through nonlinear optimization of the rotation and translation between the projector and camera. The purpose is to minimize the re-projection error of the grid feature points in the camera and projector image planes. With the rotation and translation, triangulation can then be performed via the projector and camera rays. This method of extrinsic calibration however cannot be adapted for the MAAP as it does not utilize a digital projection chip with discretized pixels. Correspondences with the camera image plane are not possible and cannot be obtained.

Instead of determining projector-camera correspondences, a direct non-linear least squares approach has been studied in single camera Fringe Projection Profilometry. A mathematically parameterized formula was proposed that directly determines object point depth with respect to a reference plane using its observed phase and pixel coordinates in the camera image plane [54, 68, 69]. Twelve parameter constants need to be determined in the equation. Using at least two block gauges of different known uniform heights with respect to a reference plane, these twelve constants can be obtained through non-linear least squares optimization of the

squared error with the block gauge heights. While not explicitly using conventional geometric parameters of the pinhole model in the equation, the twelve parameters of the direct least squares method are still derived from the assumption that both the camera and projector are pinhole modelled. This direct approach therefore cannot be adapted to the MAAP. Also, while twelve parameter constants are required when using a single aperture chip-based digital projector, the number of these constants could become computationally prohibitive for considering every distinct projection channel in a MAAP.

Morimoto et al. and Fujigaki et al. developed a high-speed LED-based array projector projecting laterally displaced purely sinusoidal fringes using a Ronchi ruling as a transmission mask [48–50, 58]. This mode of operation is similar to the operating principal of the MAAP. It was intrinsically calibrated by using a gridded test plane incremented through the measurement volume. Each projected phase-shifted pattern incident on the gridded test plane was imaged so that direct phase-to-depth look-up-tables were obtained. The gridded test plane implements a world coordinate system in which all 3D measurements are made with respect to. Not only can depth be mapped, but the phase for every *x,y* coordinate on the grid can also be acquired. Although this method can be easily adapted to the MAAP, 3D measurements become limited to the *x,y* sampling period of the grid. Not only must depth between increments be interpolated but also areas between *x,y* sampling points.

In this chapter, a comprehensive method is proposed to extrinsically calibrate the 3D measurement system based on a MAAP and single camera without manual pose measurement of a grid-less test plane. Without assuming a geometric projection model of the MAAP, the method presents a meta-heuristic derivative-free optimization-based approach to obtain the pose using a specially designed objective function as no model is available for gradient-based optimization solutions.

7.2. Influence of incorrect rotation and translation

As a thought experiment, the influence of incorrect rotation and translation between the camera and test plane on the geometry of a 3D reconstruction is investigated. Recall from Sec. 6.4 (pg. 37) the following assumptions: the test plane orientation is fixed such that the difference in orientation between each test plane increment is negligible, and the increment distance between test plane increments is constant.

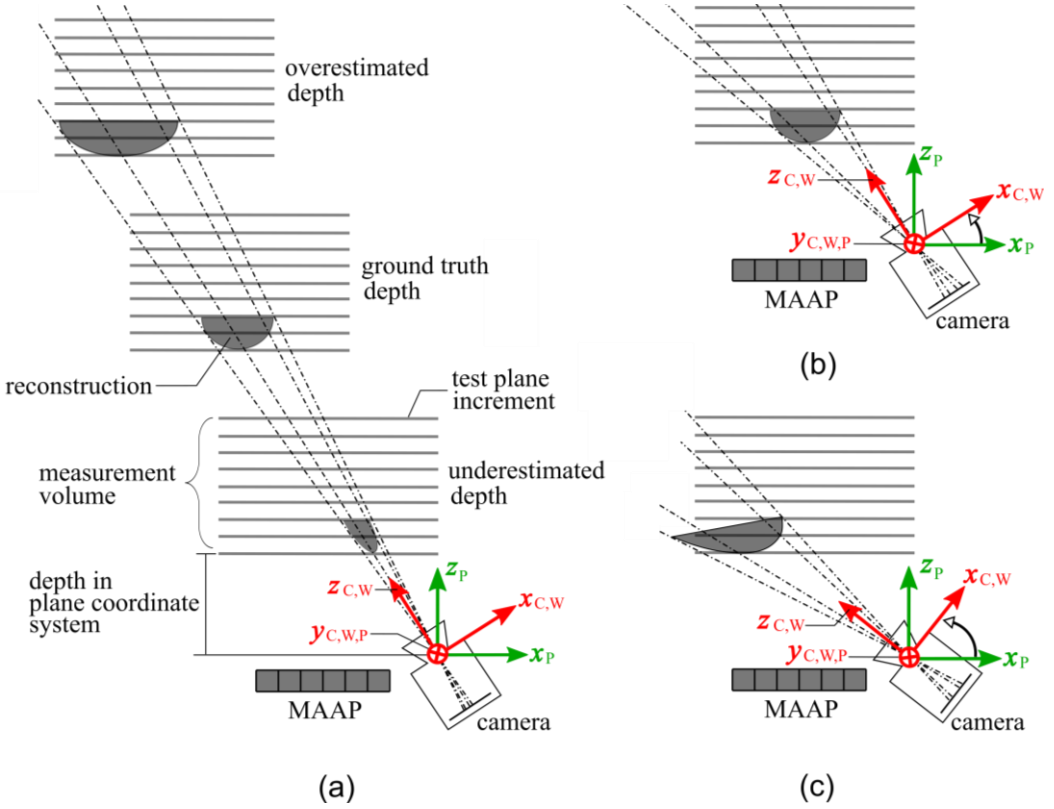

Fig. 7.2 *The effect on the reconstructed point cloud of an ideal sphere when estimating the location and orientation of the measurement volume. The plane coordinate system is denoted with subscript P and the camera/world coordinate system denoted with subscript C,W. (a) reconstruction resulting from depth estimation, (b) reconstruction resulting from correctly estimated rotation, and (c) reconstruction resulting from incorrectly estimated rotation [70].*

Thus, only the rotation and depth translation of the first test plane increment with respect to the camera needs to be determined as was shown in Eqn. 6.6 (pg. 39). This essentially defines the orientation and location of the measurement volume. The depth translation is specified as the orthogonal distance between the pinhole and the first test plane increment surface along the *z*-axis in the reference frame of the plane coordinate system. The rotation is specified as the orientation that aligns the plane coordinate system to the camera/world coordinate system. As demonstrated in Fig. 7.2a, by measuring a hypothetical ideal sphere, the effect of underestimating the depth of the first test plane increment to the camera would result in a prolate 3D reconstruction in comparison to the ground truth. Conversely, overestimating the depth would result in an oblate 3D reconstruction in comparison to the ground truth. Now assume that the camera and test plane are fixed about their *y*-axes as they would be mounted horizontally on the surface of an optical table. Comparing Fig. 7.2b and c, incorrectly estimating the rotation between the first test plane increment and camera would cause a shear transformation of the 3D reconstruction with respect to the

ground truth. From the thought experiment, by obtaining accurately the rotation and depth between the camera and the beginning of the measurement volume, good 3D measurement performance would be obtained.

7.3. Brute-force estimation of rotation and translation

Considering the presented thought experiment, if one has a geometric gauge object, then the 3D reconstruction will be highly similar to the gauge geometry when an accurate estimate of rotation and depth translation is obtained. With this ansatz, one can search for the rotation and translation until the error between the 3D reconstructed point cloud and the ground truth gauge geometry is minimized. Utilizing a geometric gauge object during extrinsic calibration would not be out of the norm. As was summarized in Sec. 3.2 (pg. 18), in a conventional triangulation-based setup, the final 3D reconstruction must be corrected for scale using a gauge of known ground truth geometry as the Fundamental matrix is not unique in terms of scale.

Considering a sphere gauge, two typical measures of comparison could be used as possible objective functions: the absolute deviation *AD* and the sphere surface deviation *σ*. A white scattering ceramic sphere gauge is used with a certified radius of r_{true} = 15.001 mm ± 1 µm and a peak-to-valley surface deviation < 3µm. *AD* is defined as the absolute difference between r_{true} and the best-fit radius of the 3D reconstruction $r_{measured}$. Therefore, one possible objective to minimize is

$$\min(AD) := \min\left(|r_{true} - r_{measured}|\right) \tag{7.1}$$

Now considering the distance d_S of each point of the reconstructed sphere (X_S, Y_S, Z_S) from the best-fit center (X_0, Y_0, Z_0)

$$d_S = \left|\sqrt{(X_S - X_0)^2 + (Y_S - Y_0)^2 + (Z_S - Z_0)^2} - r_{true}\right| \tag{7.2}$$

σ is then defined as the standard deviation of all d_S. Therefore, another possible objective to minimize is

$$\min(\sigma) := \min\left(\sqrt{\frac{1}{N-1}\sum_{j=1}^{N}|d_{S_j} - \mu|}\right) \tag{7.3}$$

where μ is the mean of all d_S, and N is the number of points for the j^{th} point in the reconstruction.

The search for the minimizing combination of rotation and translation can be done through brute-force. However, to accomplish a brute-force search, the computation time needs to be considered. The rotation consists of three degrees of freedom; rotation of the test plane coordinate system about each of its *x,y,z* axes to align to the camera coordinate system. Depth translation adds an additional degree of freedom for a total of four degrees of freedom. Since each additional degree of freedom increases the search time quadratically, a reduction in the considered number of degrees of freedom is made. Assuming that the camera and test plane are fixed about their *y*-axes as they are mounted along a horizontal optical table, there are now only two degrees of freedom to search through. To further simplify the search domain, knowing that the perpendicular distance between the test plane and camera is ~ 900 mm in the setup configuration, the domain is limited in terms of translation to search only between [800, 1000] mm. In addition, because the camera must be viewing the measurement volume, the search domain is restricted such that only counter-clockwise rotation between [0, 50]° is considered. 0° indicates that the camera view is perpendicular to the surface of the test plane. 90° indicates the camera view is parallel to the surface of the test plane. However, approaching 90° means that the plane is not in the camera view. The 50° upper limit is considered as for rotations > 50° in the setup configuration, parts of the test plane are already not in the camera view.

Shown in Fig. 7.3 is the brute-force search for the *y*-axis rotation aligning the test plane coordinate system to the camera coordinate system and the depth translation that minimizes the two possible objectives *AD* and σ. The search is performed in 1° and 10 mm increments. Observing Fig. 7.3a, it can be seen that multiple combinations of *y*-axis rotation and depth translation achieve an AD < 100 µm. Similarly, observing Fig. 7.3b, it can be seen that at a *y*-axis rotation of 17°, a wide range of depths achieve σ < 60 µm. If one were to attempt to use alone either the *AD* or σ as minimizing objectives, there will be an uncertainty in deciding which combination of rotation and

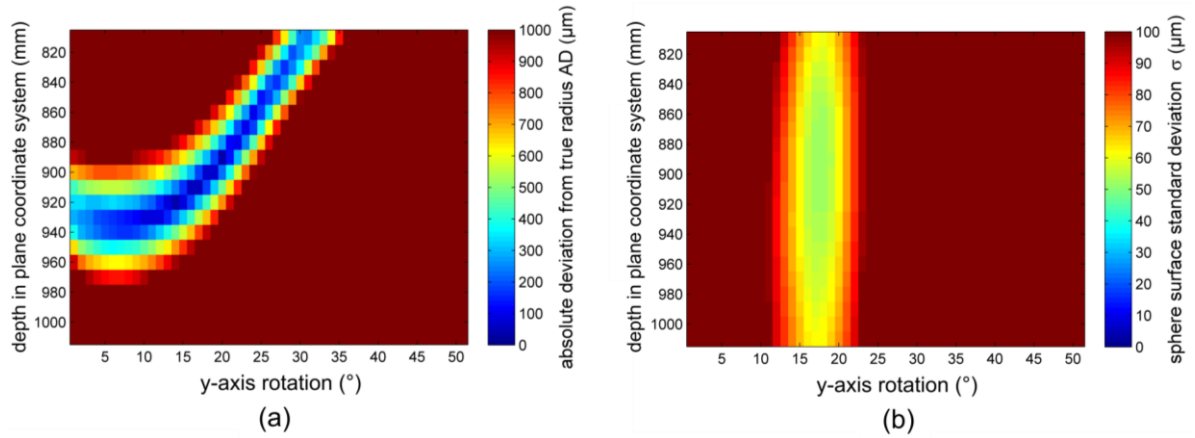

Fig. 7.3 *Search for the y-axis rotation and depth between the first test plane increment and camera that minimizes (a) the deviation of the best-fit sphere radius from the ground truth radius AD and (b) the sphere surface deviation σ [70].*

translation minimizes the geometric error with the ground truth sphere. One must use caution as some of the combinations of rotation and translation in Fig. 7.3 minimizing *AD* are not physically meaningful. The least-squares regression best-fitting process [61] attempts to fit a sphere regardless if the 3D reconstruction is more ellipsoidal in geometry. Similarly, caution is required in minimizing σ as it minimizes the surface deviation of a fitted sphere regardless of the fitted sphere's size. Since using *AD* or σ alone is insufficient, they could be used in conjunction. Cross-referencing Fig. 7.3a and b along the 17° *y*-axis rotation, both the *AD* and σ are minimized for a depth of 900 mm.

7.4. Ellipsoidal deviation

Although using both the *AD* and σ objectives in conjunction is possible, it would be more suitable and convenient if there was a single comprehensive objective. In contrast, this objective should have a single easily identifiable minimum such that there is no decision uncertainty in its location. A convex-shaped objective would satisfy this condition. It would also enable the use of so-called 'direct-search' optimization techniques rather than relying on the time-consuming brute-force method.

Fig. 7.3 indicated multiple potential combinations of rotation and translation minimizing *AD* or σ. However, some of these combinations are invalid as the least squares sphere fit procedure does not consider if the 3D reconstruction is more ellipsoidal in nature. To overcome this limitation, a potential comprehensive objective should therefore consider the ellipsoidal geometry of the 3D reconstruction. An

ellipsoid in a Cartesian coordinate system is defined as

$$\frac{x^2}{\chi^2}+\frac{y^2}{\psi^2}+\frac{z^2}{\Omega^2}=1 \tag{7.4}$$

Where $\chi,\psi,\Omega \in \mathbb{R}+$ such that χ,ψ,Ω are the distances from the center of the ellipsoid in each spatial axis direction *x,y,z*. For the case where $\chi = \psi = \Omega$, the ellipsoid is therefore a sphere [71]. From this requirement, another possible objective is introduced that can be obtained where the *AD* between each axis distance and the ground truth radius r_{true} is minimized such that

$$\min(EAD) := \min\left(\sum_{j=\chi,\psi,\Omega} |j - r_{true}|\right) \tag{7.5}$$

This sum of *AD*s shall be called the Ellipsoidal Absolute Deviation (*EAD*). Shown in Fig. 7.4 is the same brute-force search utilizing the *EAD* as the objective. A convex-shaped profile occurs with a single easily identifiable minimum at a *y*-axis rotation of 17° and a depth of 900 mm. The brute-force search of Fig. 7.4 was only performed for one sphere gauge position.

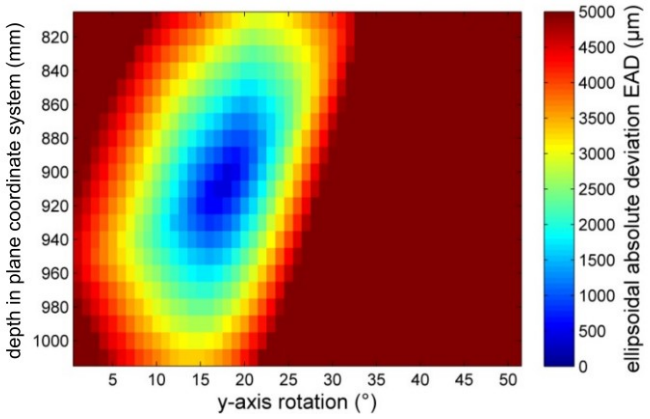

Fig. 7.4 *Search for the y-axis rotation and depth translation between the first test plane increment and camera that minimizes the EAD [70].*

If *EAD* is a truly comprehensive objective, then the convex-shaped profile should occur regardless of sphere gauge position. As seen in Fig. 7.5a and b, the sphere gauge was placed in several different positions in the measurement volume to investigate if the convex-shape of the *EAD* still holds. It is observed in Fig. 7.5c that there is no change in the convex-shape of the *EAD* when using different sphere gauge

positions. To investigate further, a finer brute-force search was then carried out in 1° and 1 mm increments. From Table 7.1, it can be seen that no substantial deviation occurs between the locations of the *EAD* minima.

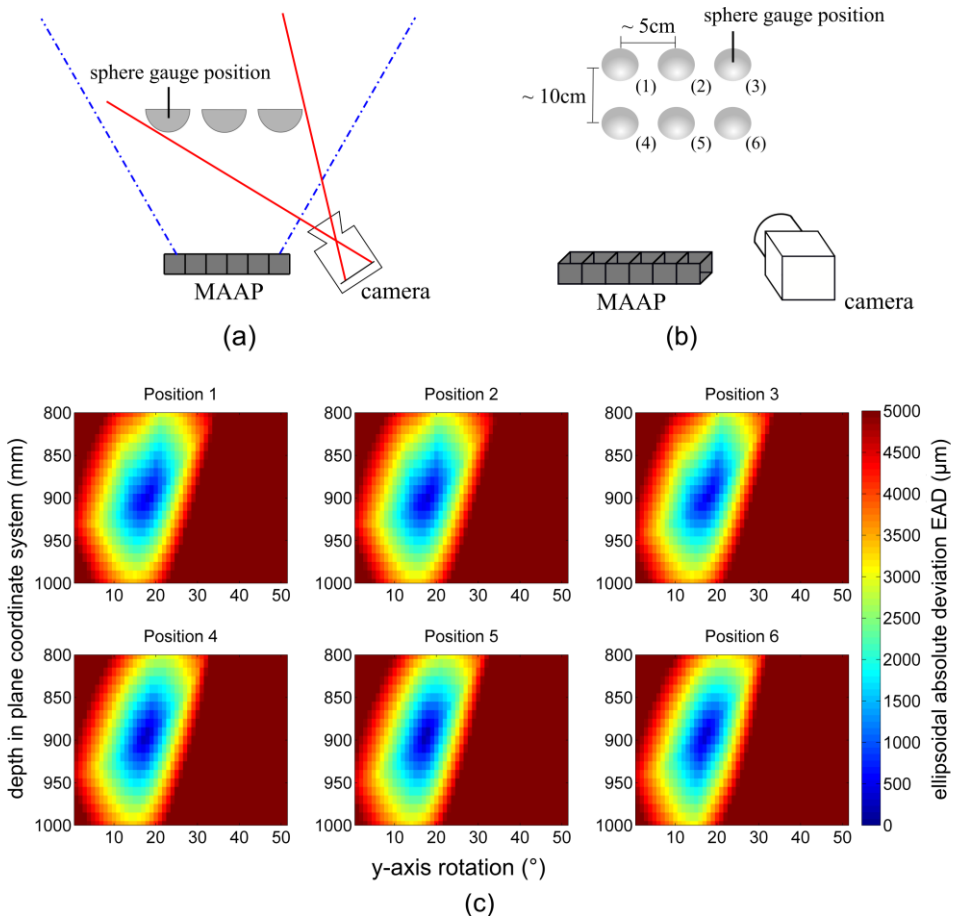

Fig. 7.5 *(a) Top-view diagram indicating sphere gauge placement in the camera view, (b) front-view, and (c) brute-force search for y-axis rotation and depth translation between the first test plane increment and camera that minimizes the EAD for different sphere gauge positions [70].*

Sphere position	Depth (mm)	*y*-axis rotation (°)	*EAD* (µm)
1	902	17	470
2	900	17	410
3	900	17	493
4	895	17	264
5	899	16	224
6	895	16	303

Table 7.1 *EAD minimum for each sphere gauge position found through brute-force search.*

7.5. Brute-force estimation of rotation and translation for all degrees of freedom

As the brute-force searches were only performed for two degrees of freedom; *y*-axis rotation and depth translation, it needs to be investigated if the convex-shape of the *EAD* still holds for all degrees of freedom; rotation of the test plane coordinate system about each of its *x,y,z* axis to align with the camera coordinate system, and depth for a total of four degrees of freedom.

However, rotation about the *z*-axis can be eliminated considering that the normal vector of the plane in the plane coordinate system is always $\overrightarrow{n_P} = (0,0,-1)$. It is defined to always be anti-parallel to the plane coordinate system *z*-axis. Thus, any rotation of the plane about its *z*-axis will not affect the direction of its normal vector. With this rotational invariance, the total number of degrees of freedom is permanently reduced to three; rotation about the *x,y* axes and depth.

For the *EAD* to be considered as a truly comprehensive objective to minimize, the convex shape of the *EAD* should hold when all degrees of freedom are considered. Shown in Fig. 7.6 is a volumetric plot of the *EAD* for the total three degrees of freedom. The search domain is restricted for an *x*-axis rotation of ±10° as outside this range, sections of the test plane are no longer in the camera view. The same restriction for the *y*-axis rotation of [0, 50]° as in the previous section is applied. The search was performed in increments of 1 mm and 1°. It can be seen that the convex shape of the *EAD* objective still holds over the search domain.

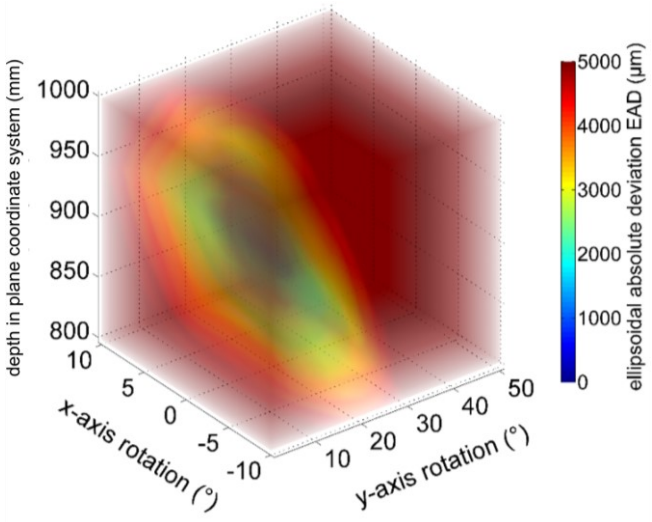

Fig. 7.6 *Search for the x and y-axis rotation and depth translation between the first test plane increment and camera that minimizes EAD [70].*

7.6. Meta-heuristic optimization-based estimation of rotation and translation for all degrees of freedom

With the introduction of the *EAD*, there is now a single comprehensive convex-shaped objective. Because the *EAD* is not a modelling function, no derivative is available for gradient-based optimization techniques. However, so-called 'direct-search' optimization methods can therefore be used rather than the time-consuming brute-force search to locate the minimizing combination of rotation and depth translation in the multidimensional domain. A classic direct-search optimization technique is the downhill simplex method [72]. Other than being time-efficient, the downhill simplex method is also not constrained to only search in discrete steps in contrast with the brute-force method. The downhill simplex method however requires an initial guess as a starting point. From Fig. 7.6, it is possible that it appears deceptively smooth. If the *EAD* is plotted in finer search increments, there could be unseen local minima in which the simplex gets stuck in in later simplex contraction stages, leading to a local search away from the global minimum. This emphasizes the need for an appropriate initial guess. However, to create a truly comprehensive method, an initial guess needs to be made without *a priori* knowledge.

Here, it is proposed to use the meta-heuristic particle swarm optimization method to obtain an initial guess without *a priori* knowledge. The particle swarm method itself first places multiple "particles" representing different candidate solutions at regular intervals or randomly throughout the search domain. The properties of each a^{th} particle can be defined as composed of two multicomponent arrays: its position in the $\mathbb{D}$ dimensional search domain $x_a = (x_{a1}, x_{a2}, \dots, x_{a\mathbb{D}})$ and its velocity $v_a = (v_{a1}, v_{a2}, v_{a3}, \dots, v_{a\mathbb{D}})$. The position and velocity in a specific *d* dimension for each k^{th} iteration is updated by

$$v_{ad}^{k+1} = w v_{ad}^{k} + r_{1_{ad}} c_1 \left(p_{ad}^{i} - x_{ad}^{k}\right) + r_{2_{ad}} c_2 (p_{a^* d}^{g} - x_{ad}^{k}) \tag{7.6}$$

$$x_{ad}^{k+1} = x_{ad}^{k} + v_{ad}^{k} \tag{7.7}$$

For each a^{th} particle, the position that resulted in the best *EAD* achieved so far during the iterative process is $p_a^i = (p_{a1}^i, p_{a2}^i, \dots, p_{a\mathbb{D}}^i)$. The position that resulted in the best *EAD* achieved so far during the iterative process out of the whole group of particles is

$p^g_{a^*d} = (p^g_{a^*1}, p^g_{a^*2}, \ldots, p^g_{a^*\mathbb{D}})$, where a^* is the specific particle that achieved the best *EAD* so far out of the whole group. $r_{1_{ad}}$ and $r_{2_{ad}}$ are random numbers drawn from a uniform distribution between [0,1]. Observing Eqn. 7.6, c_1 is a constant that influences each particle to move towards its own best position while c_2 is a constant that influences a particle to move towards the best position achieved by the entire group of particles. *w* is a constant which weights the influence of a particle's previous velocity. Thus, the constants c_1, c_2, and *w* affect convergence of the particle swarm method.

The selection of c_1, c_2, and *w* in conventional particle swarm optimization is problem dependent and convergence cannot be guaranteed [73]. A complex balance exists between each of the three parameters. Individualistic particle movement or movement as a group can be emphasized. In addition, too large *w* can cause overshooting of the region containing the global minimum, while too small *w* can affect its tendency to become stuck in a region of a local minimum [74]. Due to the possible range of c_1, c_2, and *w* values and combinations, the particle swarm method was chosen only as a means to obtain an initial guess for the downhill simplex optimization rather than using the particle swarm alone. In addition, the ability to start the particle swarm optimization with multiple particles spread throughout the entire search domain enables initial exploration of different regions of the search domain rather than being limited to explore the area around a single region.

In this study, the constant parameters are set to c_1 = 0.2, c_2 = 0.1, and *w* = 0.5. A group of 20 particles are first randomly positioned according to a uniform probability distribution throughout the search domain. The velocity of each of the particles is randomly initialized in each dimension of the search domain between [-1,1] according to a uniform probability distribution. The particle swarm optimization runs for 10 iterations and takes ~30 s to compute on an Intel i5-4590 CPU. After completion of the 10 iterations, the rotation and depth translation corresponding to $p^g_{a^*d}$ is used as the initial guess to initialize the downhill simplex optimization. The downhill simplex optimization also completes in ~30 s. To demonstrate that the proposed process is also a more comprehensive solution than the brute-force approach, an expanded search domain is utilized for depth of [0, 2] m, *y*-axis rotation [0, 90]°, and *x*-axis rotation [-90, 90]°.

7.7. Results of optimization

Due to the particle swarm and downhill simplex methods being heuristic in nature, the proposed process is performed for 1000 different runs. If large deviations between runs exist, this would indicate that the *EAD* has local minima present that are large enough to trap the contracting simplex in later iterations. The particle swarm initialization is compared to initialization with a uniformly random initial guess of depth, *y*-axis rotation, and *x*-axis rotation.

Shown in Fig. 7.7 is the count histogram of depth, *y*-axis rotation, and *x*-axis rotation obtained from the 1000 runs. A single run is completed when the optimization has reached an *EAD* tolerance ≤ 100 µm and step tolerance ≤ 1e-4. In Fig. 7.7a-c, it can be seen that the region about a global minimum is reached by a majority of the runs with the uniform random initial guess. However, as seen as the red colored bins, there are instances where an ellipsoid cannot be fitted to the point cloud during downhill simplex optimization. Observing Fig. 7.7d-f, there are also some instances where the simplex becomes trapped in a local minimum. In Fig. 7.7g-i, all 1000 runs reach a region about the global minimum using the proposed particle swarm initialization.

A more detailed breakdown of the histogram counts can be seen in Fig. 7.7j-l, which shows no major deviation between separate optimization runs. Thus, the convergence of the proposed process of initializing the downhill simplex optimization using the particle swarm method to optimize the *EAD* was demonstrated to be highly consistent. Over the 1000 separate runs, the average depth translation was (895.11 ± 0.27) mm, the *y*-axis rotation was (16.56 ± 0.01)°, and the average *x*-axis rotation was (−0.66 ± 0.01)°.

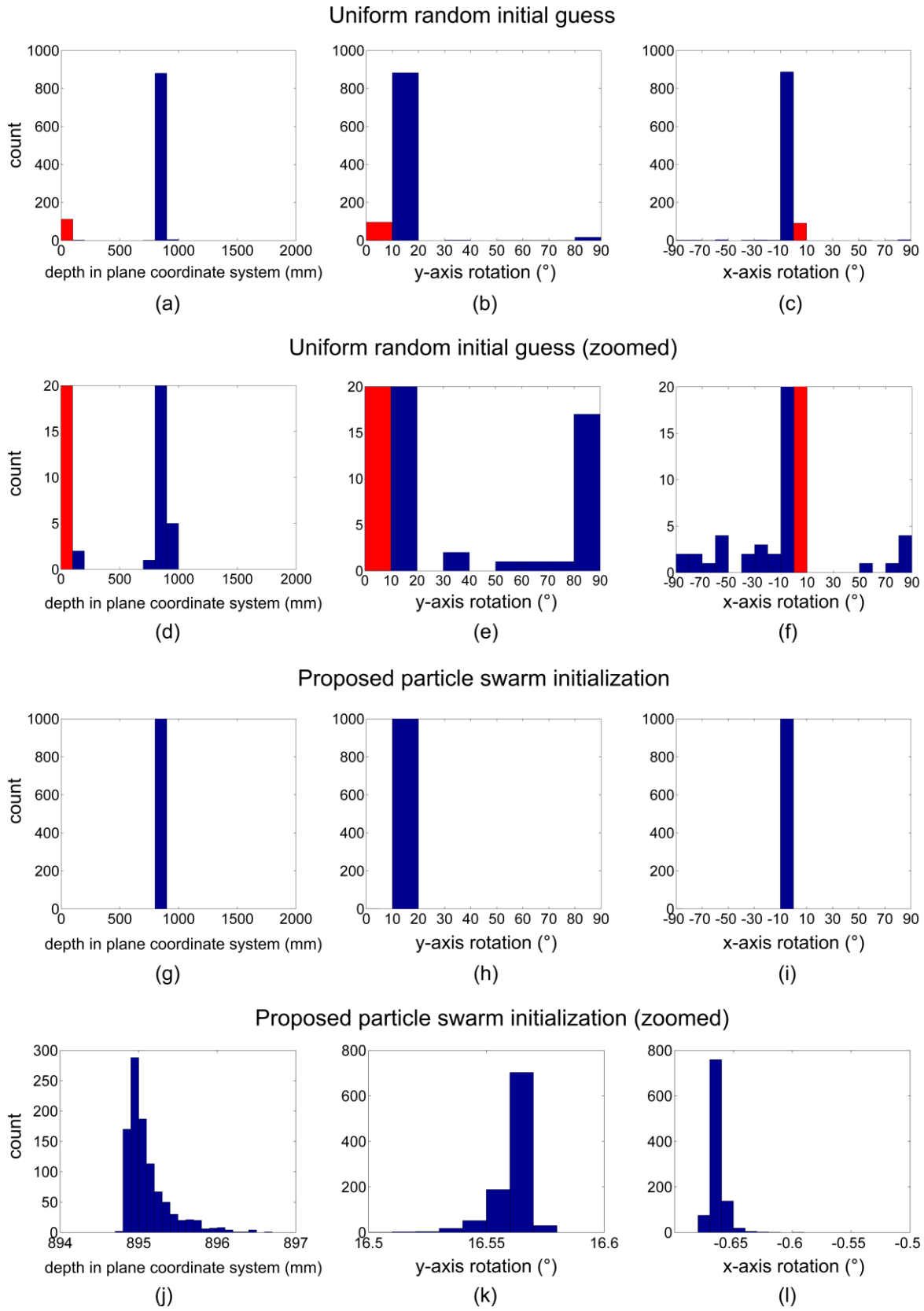

Fig. 7.7 *Histogram of rotation and depth translation of first test plane increment with respect to camera obtained after 1000 separate optimization runs. (a-c) Using the randomly initialized downhill simplex, (d-f) zoomed graphs of (a-c) with red bars indicating instances where the ellipsoid fit could not be initialized, (g-i) using the particle swarm initialization, (j-l) zoomed graphs of (g-i).*

7.8. 3D measurement performance

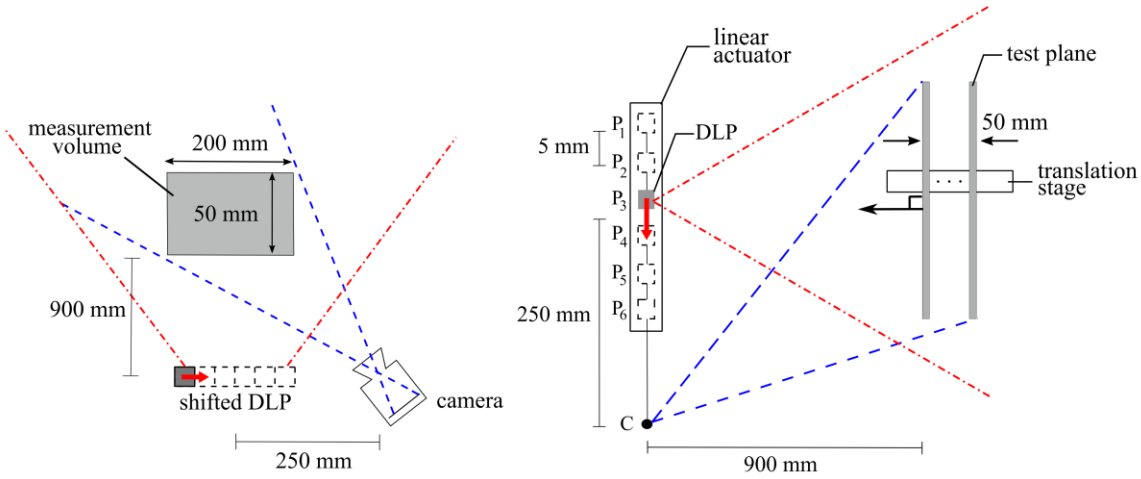

Fig. 7.8 *Experimental test-bed setup of linearly shifted DLP emulating the function of the MAAP with camera C and emulated projection channels P [70].*

Using the average rotation and translation of the 1000 separate runs determined in the previous section, the resulting 3D measurement performance is experimentally evaluated using a test-bed setup as shown in Fig. 7.8. Recall from Sec. 6.5.1 (pg. 40) that the configuration was set up such that the first test-plane position was manually placed about 900 mm from the projector-camera baseline connecting the pinhole and physical center of distinct projection channels. The test-plane was placed orthogonal to the illumination direction. For a measurement volume of 200×200×50 (H×W×D/mm), the plane is incremented in steps of 0.5 mm using a linear translation stage. A grayscale AVT Pike F-032B camera with focal length f = 25 mm, quadratic pixel pitch of Δp = 7.4 µm, and resolution of 640×480 px. imaging in 16-bit grayscale mode, was placed 250 mm away from the baseline.

The test-bed setup acts as a MAAP emulator as it implements a linearly shifted single aperture DLP projector. At each shift position, a distinct aperiodic fringe pattern of the pattern sequence is projected on the same test plane increment and imaged. This linear shift imitates each distinct projection channel of the MAAP. The benefit of using the test-bed is that different pattern sequences can be utilized unlike the fixed pattern projection of the actual MAAP device. Thus, the consistency of 3D measurement results can be evaluated over different pattern sequences. Aperiodic fringe patterns were implemented such that the projected fringe bandwidth was between [7,10] mm on the test plane. Since a sphere gauge object was utilized as a

reference object to minimize the *EAD* such that the 3D error of a sphere is purposefully minimized, it would be inappropriate to evaluate 3D measurement performance with respect to a sphere. It would also be interesting to investigate the measurement capability of a complex surface distribution. Instead, the measurement performance with respect to a non-planar object in the shape of German poet Johann Wolfgang von Goethe was evaluated to determine measurement performance of geometrically complex shapes. The dimensions of the bust are ~80×50×40 (H×W×D/mm).

Shown in Fig. 7.9a is the bust of Fig. 7.9b illuminated by aperiodic fringes for a sequence of six patterns. Shown in Fig. 7.9c is the bust 3D reconstructed by a separate stereo-photogrammetric system. The separate stereo-photogrammetric system utilizes the projection of thirty band-limited patterns [75, 76] and is capable of measurement surface standard deviation of σ = 26.7 µm from the ground truth without any outlier filtering [15]. The 3D reconstruction generated by the separate stereo-photogrammetric system is the surface topology of the bust as close to the ground truth as could be known and is used as a reference point cloud.

Fig. 7.9d shows the 3D reconstructed point cloud generated utilizing the proposed optimization process to extrinsically calibrate the single camera system. Fig. 7.9e shows the point cloud generated through manual extrinsic calibration. Fig. 7.9f shows the point cloud generated with a stereo-camera system using the same aperiodic fringes and the same cross-correlation threshold of $\rho \geq 0.99$ to determine correspondences. The absolute point-wise deviation σ_{bust} of these point clouds from the reference point cloud can be seen in Fig. 7.9g-i.

Fig. 7.9b also indicates vertical and horizontal lengths obtained from the generated point clouds, where the length error Δl is the absolute difference between the measured lengths and the reference such that $\Delta l = |l_{measured} - l_{ref}|$. Another important factor is the completeness of the point cloud

$$\text{completeness} = \frac{\text{number of generated points in point cloud}}{\text{number of possible points}} \tag{7.8}$$

since significant missing points result in a loss of spatial details. Comparing point clouds directly generated by the stereo and monocular-view setups for completeness is difficult. Occlusion occurring between two camera views results in missing reconstructed points that would appear in the single camera setup and vice versa. As

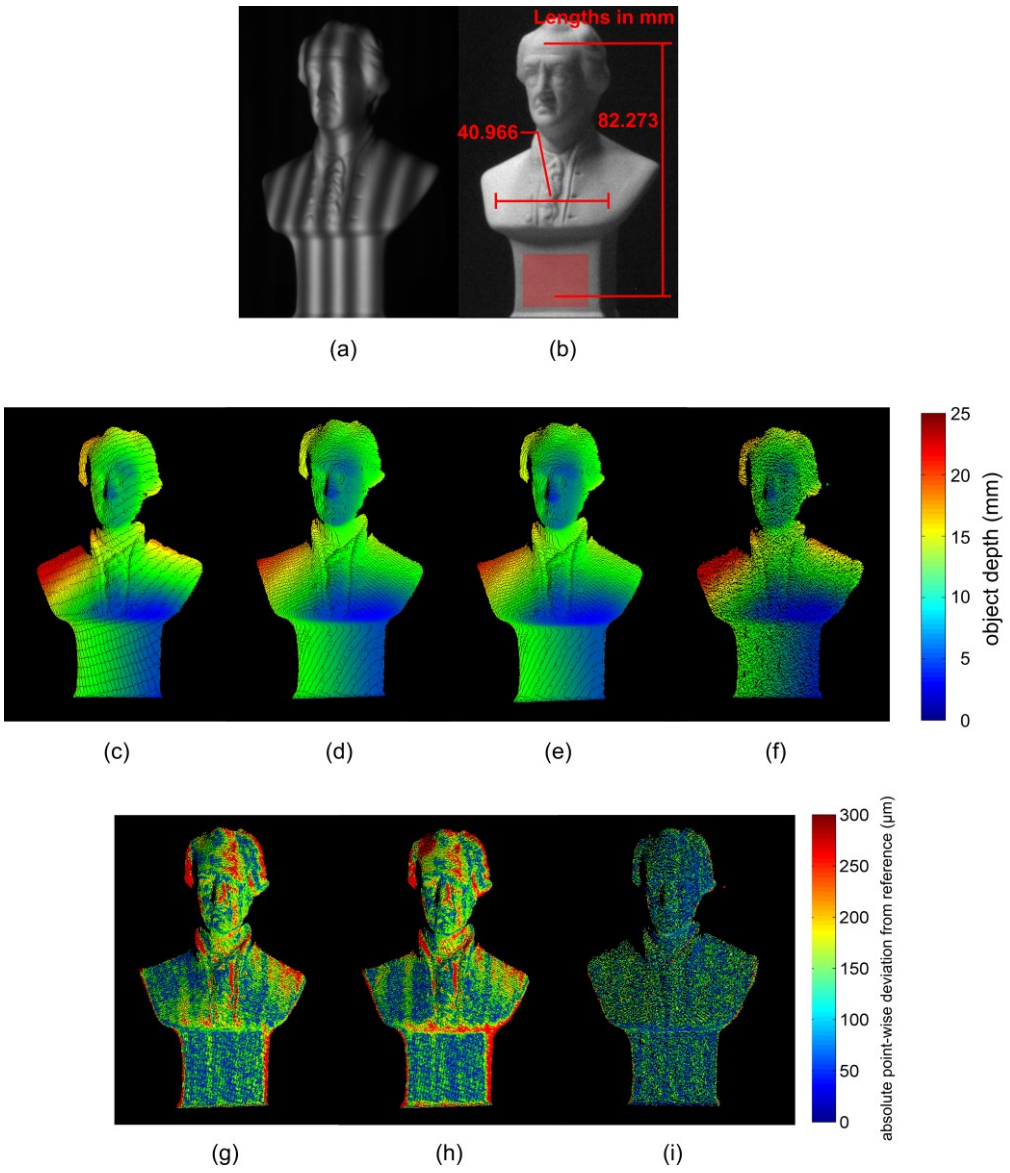

Fig. 7.9 *Bust object in the shape of Johann Wolfgang von Goethe (a) illuminated by an aperiodic fringe pattern, (b) bust object indicating vertical and horizontal reference lengths for measurement comparison and the selected area for completeness comparison, (c) reference point cloud, (d) single-camera 3D reconstruction generated using proposed extrinsic calibration optimization method, (e) single-camera 3D reconstruction generated through manual extrinsic calibration, (f) stereo-camera 3D reconstruction generated using same pattern sequence as Fig.d-e, (g) point-wise deviation of Fig.d from reference, (h) point-wise deviation of Fig.e from reference, (i) point-wise deviation of Fig.f from reference [70].*

seen in Fig. 7.9b, the area common between the setups is highlighted to show where the completeness was determined.

The point clouds for a total of ten different aperiodic fringe sequences were generated and the aggregated absolute point-wise deviation σ_{bust} is shown in Fig. 7.10 and Table 7.2. Outlier points of each point cloud with deviation > 1 mm from the

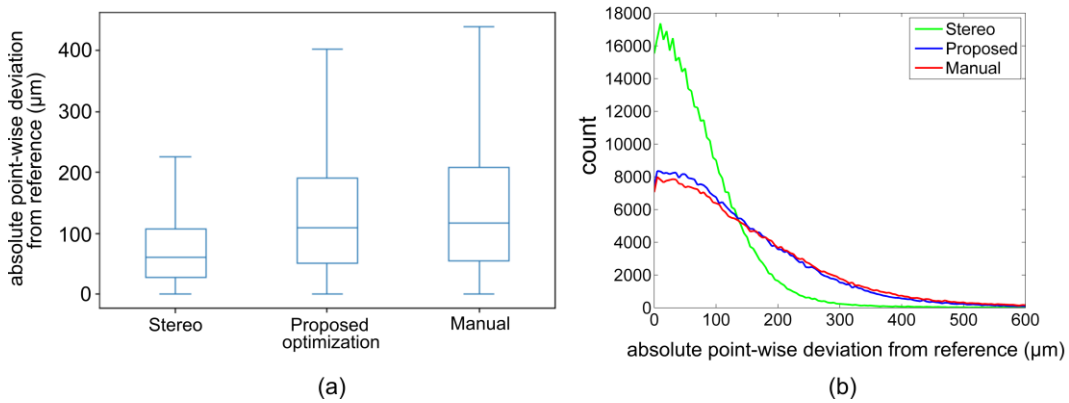

Fig. 7.10 *Comparing the absolute point-wise deviations generated by different extrinsically calibrated methods between the 3D reconstructed bust and reference. Graphs aggregate all points from ten reconstruction samples: (a) Quantile boxplot of deviations and (b) distribution of deviations.*

	25th percentile σ_{bust} (µm)	Median σ_{bust} (µm)	75th percentile σ_{bust} (µm)
Stereo	29	62	108
Proposed	53	112	198
Manual	56	120	215

Table 7.2 *Quantiles of absolute point-wise deviations and completeness for different extrinsically calibrated methods.*

	Outliers removed (%)	Completeness (%)
Stereo	0.5	96.7
Proposed	2.7	99.8
Manual	2.7	99.8

Table 7.3 *Percentage of point cloud outliers removed and completeness for different extrinsically calibrated methods.*

reference are filtered. Table 7.3 shows the average percentage of outliers removed and completeness of the ten point clouds. The quantiles of the length error Δl of the ten point clouds are shown in Fig. 7.11 and Table 7.4.

It can be seen from Fig. 7.10 and Fig. 7.11 that the stereo setup generates 3D reconstructions with better measurement performance. The point-wise deviation σ_{bust}, a measure of reconstructed surface precision, is considerably lower. The 50th percentile deviation of the stereo setup is nearly equal to the 25th percentile of the single camera setups. The stereo setup also offers better measurement performance of length error, a measure of reconstruction accuracy. The 75th percentile of measured lengths in the stereo setup is less than the 25th percentile of the single camera setups.

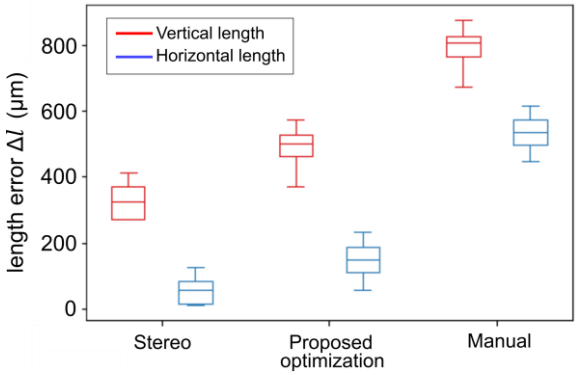

Fig. 7.11 *Comparing the length errors generated by different extrinsically calibrated methods between the 3D reconstructed bust and the reference.*

	Vertical length		
	25th percentile Δl	Median Δl	75th percentile Δl
Stereo	232	327	370
Proposed	453	502	532
Manual	755	806	834
	Horizontal length		
	25th percentile Δl	Median Δl	75th percentile Δl
Stereo	13	57	89
Proposed	105	150	192
Manual	490	534	576

Table 7.4 *Quantiles of length errors for different extrinsically calibrated methods.*

This better general 3D measurement performance is expected due to the stereo setup using the same Epipolar geometry as the setup generating the reference. The point-wise surface deviation and length error from the reference is due to using only six patterns in the temporal pattern sequence. The correspondence match quality is degraded as a shorter temporal gray-value intensity sequence is more easily susceptible to image noise corruption, causing the corresponding points to deviate from their ground truth locations.

Comparing the point-wise deviations in Fig. 7.10 between the proposed optimization-based and manual extrinsic calibration methods for a single-camera setup shows similar performance. This is not surprising as the exact same pattern sequence was used. The differences between the two extrinsic calibration methods is more apparent in the length error observing Fig. 7.11. It can be seen that the proposed optimization-based method better estimates the test plane pose as the length errors

are substantially lower than the manual method. No matter how meticulous, manual measurement of the test plane pose is naturally predisposed to deviate more from the ground truth pose causing prolate or oblate reconstruction geometry errors. Also seen is that the vertical lengths are systematically larger than the horizontal lengths. This is possibly due to residual error in the estimation of rotation and translation of test plane. The systematic difference could have possibly arisen due to the test plane depth still being slightly underestimated compared to the ground truth, a small unaccounted for rotation about the *x*-axis, or some combination of the two.

In Table 7.3, it can be observed that the completeness of point clouds generated using the single camera setup using either the proposed or manual extrinsic calibration are equal. This is expected as the same pattern sequence results in the same number of points regardless of its geometry. The single camera methods outperform the point cloud completeness of the stereo-setup. This indicates that the identification of corresponding depth while scanning along a light ray in the single camera setup is more easily detectable for $N = 6$ patterns compared to the identification of stereo-correspondences scanning along an Epipolar line. Scanning along light rays in the single camera setup yields more points reaching the cross-correlation threshold of $\rho \geq 0.99$. However, this increased detectability of correspondence comes at the cost of susceptibility to sequence corruption by imaging noise. This results in false correspondences, causing a greater number of outliers with deviation > 1 mm that need to be filtered.

7.9. Conclusion

In this chapter, a method was proposed and demonstrated to extrinsically calibrate a 3D measurement system composed of a MAAP and single camera and is summarized in Table 7.5. It implements a derivative-free optimization-based approach that minimizes the geometry error between the ground truth geometry of a sphere gauge and its initial 3D reconstruction. The optimization solution is a downhill simplex optimizer coupled with a particle swarm optimizer whose function is to provide an initial guess for the downhill simplex method without *a priori* knowledge. The aim is to minimize a specially designed objective function that considers the geometry error along each major axis of the initially 3D reconstructed ellipsoid and the ground truth sphere gauge. It was observed that the optimization converged to the minimum

consistently. With the presented extrinsic calibration method, 3D measurements were made that agree well with the reference surface. The presented extrinsic calibration method is comprehensive without manual intervention and produces measurement results better than manual extrinsic calibration. The measurement performance was observed to be worse compared to when using a stereo setup but achieved greater point cloud completeness.

Particle swarm initialized downhill simplex extrinsic calibration
1. Intrinsically calibrate the MAAP according to Chapter 6.
2. Image the projected aperiodic fringes onto a sphere gauge.
3. Perform particle swarm optimization as described in Sec. 7.6 to search for the rotation and depth of the test plane that minimizes the reconstructed point cloud's Ellipsoidal Absolute Deviation (*EAD*) of Eqn. 7.5.
4. Use the found rotation and depth from step 3 as an initial guess for a downhill simplex optimization as described in Sec. 7.6. The search parameters and optimizing objective are the same as in step 3.
5. Completing step 4 obtains the optimal rotation and depth parameters that extrinsically calibrate the system.

Table 7.5 *Summary of the extrinsic calibration method proposed in this chapter.*

Chapter 8.

Intrinsic ‘re-calibration’ of a Multi-Aperture Array Projector for arbitrary monocular measurement views through artificial view synthesis

8.1. State-of-the-art

Intrinsic calibration of the optical devices in a stereo-view setup is a straightforward process using the Zhang method assuming they can be pinhole modelled [10]. Unless the camera or projector experiences some mis-adjustment of the lens, the intrinsic parameters obtained from calibration remain the same regardless of device position. Therefore, one can save the intrinsic parameters for alternate setup configurations such that only an extrinsic calibration is needed to determine the pose relationship between the triangulating devices.

A setup using the MAAP however does not have this capability. From Chapter 6 (pg. 32), this study proposed to intrinsically calibrate the MAAP where a test plane is incremented depth-wise through the measurement volume. For each increment position, the camera images the pattern projected by each projection channel as if discrete depth slices are imaged of each illumination frustum. To obtain the 3D coordinates of an object in the measurement volume, the corresponding depth of each object point needs to be determined. If the camera position has changed, then the entire intrinsic calibration process must be performed again as the intensities through the measurement volume have been encoded only for a specific camera view. Consequently, while the intrinsic calibration of a pinhole modelled DLP or other types of digital projectors is solely dependent on the device itself, the intrinsic calibration of the MAAP is coupled with the camera pose.

As seen in the flowchart of Fig. 8.1 first introduced in Sec. 3.2 (pg. 18), the pipeline to generate 3D reconstructions is nearly complete with newly developed intrinsic and extrinsic calibration processes proposed in Chapter 6 (pg. 32) and Chapter 7 (pg. 50) respectively. However, a method to save the results of the intrinsic calibration for re-use in alternate setup configurations is not yet possible.

The stipulation that the camera device must be fixed to maintain the MAAP intrinsic calibration is inconvenient as it restricts setup flexibility and does not lend well

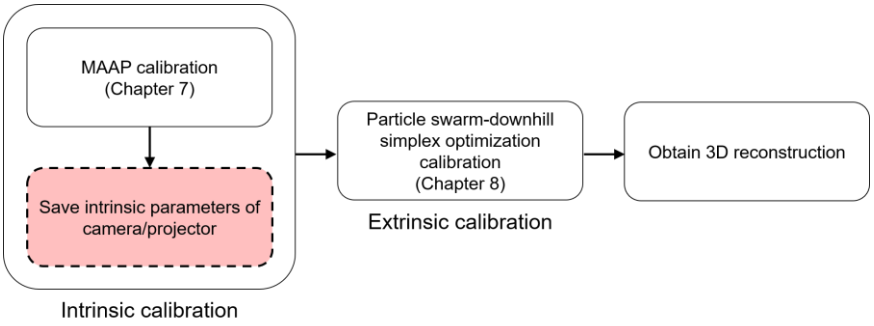

Fig. 8.1 *Flowchart depicting missing components of the calibration pipeline. The processes calibrating triangulating pinhole devices that cannot be applied here are in red. This chapter’s focus is depicted in the dashed line.*

in case of accidental perturbation of camera position. It would be laborious and time-consuming to repeat the entire test plane incrementation process all over again to obtain the volumetric 3D grid of temporal gray-value intensity sequence data.

This raises the question: is intrinsic MAAP calibration possible for arbitrary camera views without having to take images of each aperture’s illumination frustum again? For accessibility, this should be achieved only with off-the-shelf devices available to a typical machine vision facility and without requiring additional aiding bespoke hardware, tailor-made only for this application.

A possible solution arising from the posed question is to artificially synthesize the intrinsic calibration images as if they were directly taken from the arbitrary camera view. View morphing or artificial view synthesis has been mainly investigated in the computer vision field. The most prominent first work was investigated by Seitz et al. where artificial views are interpolated between two cameras [77]. However, this limits view synthesis to only along the baseline distance connecting the two views. Artificial view synthesis is a highly active field of study in Machine Learning with the application of Deep Learning approaches [78–81]. But as of this study, the synthesized views are not produced with error minimal enough compared to the ground truth for this application, takes large amounts of training data of nearly 100 sparse images of a scene, and takes 1-2 days of training time for a scene using an NVIDIA V100 Tensor Core GPU [79].

In this chapter, a posterior method is proposed to intrinsically ‘re-calibrate’ the MAAP by initially saving the volumetric 3D grid of temporal gray-value intensity sequence data for later use such that subsequent arbitrary monocular measurement views of the MAAP illumination frustums can be artificially synthesized. Hence, the arbitrary monocular view that actually performs 3D measurement does not need to

carry out explicitly the entire MAAP calibration process. The entire time-consuming process of taking images of multiple test plane increments does not have to be performed again to calibrate the MAAP for the arbitrary measurement camera view. The proposed method therefore decouples the MAAP intrinsic calibration from camera position and orientation. As a result, 3D measurements shall be possible without the camera being fixed to a specific calibrated position.

8.2. Proposed view synthesis principle

For an arbitrary monocular measurement position *C**, the process to intrinsically calibrate the MAAP must be carried out again as the measurement volume would not be encoded from this unknown view. Here, a slight modification from the proposed intrinsic MAAP calibration procedure of Chapter 6 (pg. 32) is presented in order to enable artificial view synthesis.

Prior to performing any 3D measurement from the arbitrary monocular measurement position *C**, an initial ‘aiding’ triangulation configuration is set up utilizing Structure from Motion (SfM). The cameras are placed at ‘aiding’ positions denoted as C_1 and C_2. For any point on a test plane increment, its 3D coordinates *(X,Y,Z)* can therefore be initially triangulated utilizing C_1 and C_2 with respect to either aiding camera position. Let the 3D coordinates be calculated with respect to the reference frame of aiding position C_2 and denoted as $(X,Y,Z)_{C_2}$. Thus, every point contains an intensity sequence $I_j(X,Y,Z)$ for every pattern *j* projected such that

$$(X,Y,Z)_{C_2} \ni \left\{ I_j(X,Y,Z) \;\middle|\; j \in \{1 \dots N\} \right\}$$
$$(\alpha,\beta)_{C_2} \ni \left\{ I_j(X,Y,Z) \;\middle|\; j \in \{1 \dots N\} \right\} \tag{8.1}$$

With the pinhole model assumption, recall that every 3D point $(X,Y,Z)_{C_2}$ is projected onto a pixel of the imaging chip at position C_2, denoted as $(\alpha,\beta)_{C_2}$. Hence, every $(\alpha,\beta)_{C_2}$ is the intersection between the light ray scattered from $(X,Y,Z)_{C_2}$ and the image chip at position C_2.

If the relational pose geometry aligning the reference frames of the aiding position C_2 and any arbitrary monocular measurement position *C** could be

determined, every 3D point $(X,Y,Z)_{C_2}$ could be transformed to be made with respect to the *C** reference frame to obtain $(X,Y,Z)_{C^*}$ such that

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{C^*} = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix}_{C_2 \to C^*} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{C_2} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}_{C_2 \to C^*} \tag{8.2}$$

where the unitless 3×3 rotation matrix $R_{C_2 \to C^*}$ and translation vector $\vec{t}_{C_2 \to C^*}$ align the C_2 reference frame to the *C** reference frame.

As seen in Fig. 8.2, for a light ray scattered from a 3D point on a test plane increment in the reference frame of *C** $(X,Y,Z)^P_{C^*}$, its intersection at the *C** image plane can be calculated such that

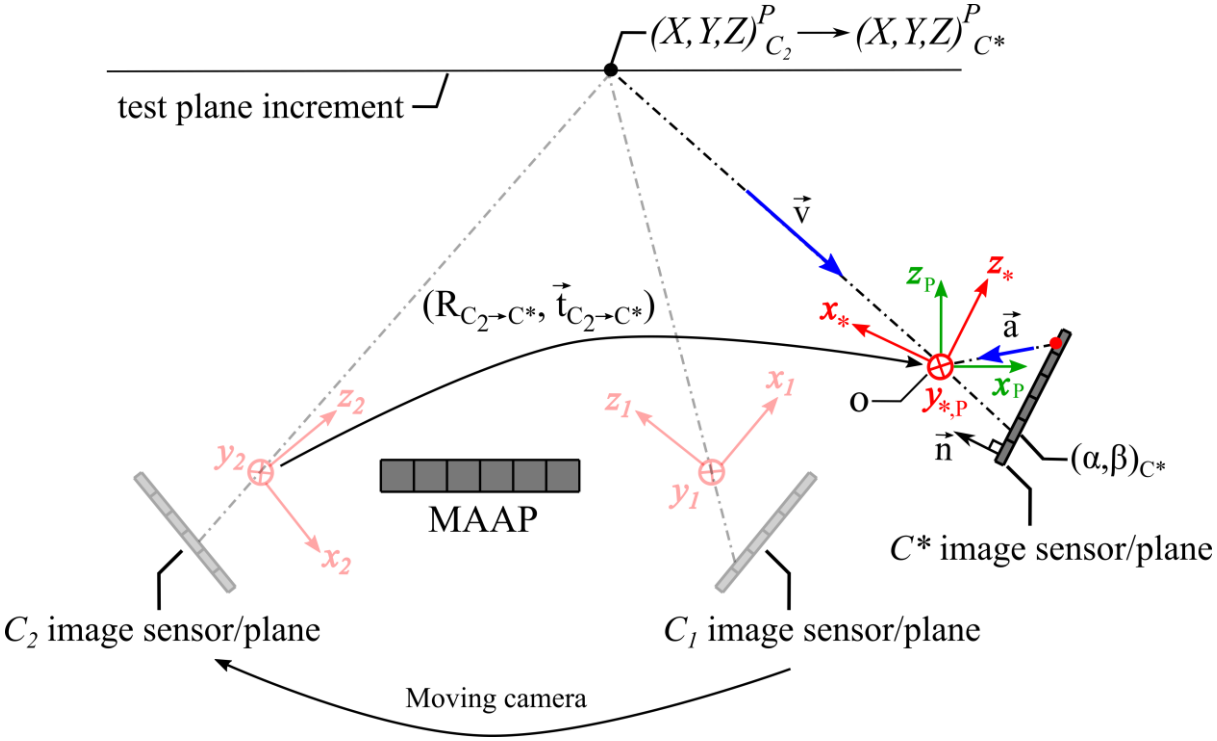

Fig. 8.2 *Projecting the triangulated 3D point on a test plane increment $(X,Y,Z)^P_{C_2}$ made in the reference frame of the aiding camera position C_2 to the reference frame of the monocular arbitrary measurement position C* requires the pose relationship between $C_2 \to C^*$. C_2 and C_1 are grayed out as they do not have to be physically present when imaging with C*. The pixel coordinates $(\alpha,\beta)_{C^*}$ are determined by calculating the intersection between the ray direction vector $\vec{v}$ and the C* image plane. The coordinate system of the test plane is denoted with subscript P and the coordinate system of C* is denoted with subscript *. The origin o of the two coordinates systems is shared. $\vec{a}$ is an arbitrary vector from an arbitrary point on the C* image plane to the origin o [82].*

$$\gamma = \frac{\vec{n} \cdot \vec{a}}{\vec{n} \cdot \vec{v}}$$
$$(X,Y,Z)^{I}_{C^*} = \gamma \vec{v} \tag{8.3}$$

where γ is a scalar, the normal vector of the image plane in the *C** reference frame is $\vec{n} = (0,0,1)$, and $\vec{a}$ is an arbitrary vector from an arbitrary 3D start point on the image plane in the *C** reference frame to the origin of the *C** reference frame $o = (0,0,0)$. The start point of $\vec{a}$ can be conveniently taken as $(0,0,f)$ with objective lens focal length *f*. $(X,Y,Z)^{I}_{C^*}$ is the intersection point with the image plane. Since the light ray is scattered from $(X,Y,Z)^{P}_{C^*}$ and travels through the *C** pinhole, the directional vector is therefore $\vec{v} = o - (X,Y,Z)^{P}_{C^*}$. With these values, Eqn. 8.3 reduces to

$$\gamma = \frac{f}{Z^{P}_{C^*}}$$
$$(X,Y,Z)^{I}_{C^*} = \frac{f}{Z^{P}_{C^*}} \vec{v} = \left(\frac{fX^{P}_{C^*}}{Z^{P}_{C^*}}, \frac{fY^{P}_{C^*}}{Z^{P}_{C^*}}, f \right) \tag{8.4}$$

The pixel coordinates $(\alpha,\beta)_{C^*}$ of the intersection $(X,Y,Z)^{I}_{C^*}$ can be calculated using the pixel pitch of the camera *Δp*, and the principal point in pixel units *(u,v)* in which the camera’s optical axis runs through

$$(\alpha,\beta)_{C^*} = \left(\frac{1}{\Delta p} \frac{fX^{P}_{C^*}}{Z^{P}_{C^*}} + u, \frac{1}{\Delta p} \frac{fY^{P}_{C^*}}{Z^{P}_{C^*}} + v \right) \tag{8.5}$$

Therefore, once the rotation $R_{C_2 \to C^*}$ and the translation vector $\vec{t}_{C_2 \to C^*}$ aligning the pose from aiding position C_2 to arbitrary monocular measurement position *C** are obtained, any 3D point on a test plane increment initially in the C_2 reference frame $(X,Y,Z)^{P}_{C_2}$ can be converted into the *C** reference frame to obtain $(X,Y,Z)^{P}_{C^*}$ using Eqn. 8.2. With $(X,Y,Z)^{P}_{C^*}$, the projection on the *C** image chip $(\alpha,\beta)_{C^*}$ using Eqn. 8.5 can then be obtained. Therefore, for every $(\alpha,\beta)_{C_2}$ the ‘apparent’ corresponding $(\alpha,\beta)_{C^*}$ can be determined without direct imaging by *C**. By Eqn. 8.1, this ‘apparent’ correspondence means the gray-value intensity sequence of $(\alpha,\beta)_{C*}$ is known such

that $(\alpha, \beta)_{C^*} \ni I_j(X, Y, Z)$. Thus, the images of every test plane increment from any arbitrary monocular measurement position *C** can be artificially synthesized using only the images captured by the initial aiding triangulation SfM setup.

By storing the images, the MAAP can be intrinsically ‘re-calibrated’ posteriorly for alternate arbitrary measurement positions in the future, similar to how one can save the intrinsic parameters of a pinhole modelled projector for later use. This effectively stores the volumetric 3D grid of temporal gray-value intensity sequence data for later use.

The initial aiding triangulation setup does not have to be present when imaging with the arbitrary monocular measurement position *C**. Thus, for a completely arbitrary single measurement camera position and orientation, the time-consuming process of having to image every test plane position does not have to be repeated.

8.3. Pose determination with background fixed calibration jig

What remains to be determined is the relational pose between the arbitrary measurement camera position and the aiding camera position with which test plane 3D coordinates are with respect to. Thus, assuming that the test plane 3D coordinates were made with respect to aiding position C_2 of the initial aiding triangulation SfM setup, the Epipolar geometry between the two reference frames, $C_2 \rightarrow C^*$, needs to be calculated. It can be done using the eight-point algorithm where at least eight homologous corresponding points between the two views are required [11]. The eight corresponding points must not be located on a single dominant planar scene [34, 35]. $C_2 \rightarrow C^*$ is obtained by using a calibration jig composed of two Thin-Film-Transistor (TFT) panels angled with respect to one another as seen in Fig. 8.3. Recall from Eqn. 3.9 (pg. 20) that the Fundamental matrix of the Epipolar geometry can then be decomposed into the rotation and translation aligning the two camera reference frames. As the scale of the translation vector would be incorrect due to scale ambiguity of the Fundamental matrix, the correct scaling factor can be obtained using a polygon of known ground truth geometry defined on the TFT panel itself as the TFT pixel pitch is specified by the manufacturer.

The TFTs allow flexibility to display any desired pattern design with feature points so that homologous corresponding points between the two views can be obtained. By using TFT panels, creation of such a calibration jig should be easily

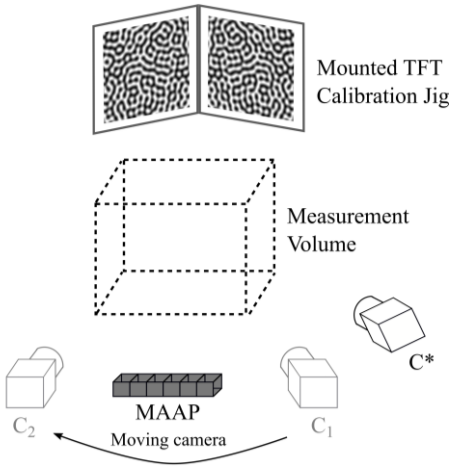

Fig. 8.3 *Permanently mounted calibration jig in the background behind the intended measurement volume comprising two TFT panels angled with respect to one another. The TFT displays a pattern full of features that enable correspondence detection between aiding camera positions C_1 and C_2, and monocular arbitrary measurement position C*. When performing 3D measurement with C*, C_1 and C_2 are grayed out to denote that they do not have to be present when measuring using C* [82].*

accessible by any machine vision facility. By permanently mounting the calibration jig behind the desired measurement volume on a background wall for example, the scene remains constant. If the monocular arbitrary measurement position *C** views the same scene of the fixed calibration jig, the pose between $C_2 \rightarrow C^*$ can also be determined. Even if the initial aiding triangulation SfM setup is removed and not physically present, $C_2 \rightarrow C^*$ can always be obtained as long as the aiding position C_2 images of the fixed calibration jig are stored. The caveats to this are that since the projected MAAP patterns are initially taken by the aiding triangulation SfM setup from a certain MAAP position, then the orientation between the MAAP and the calibration jig must also be fixed. In addition, there must be a shared set of feature points on the calibration jig seen by the aiding positions C_1, C_2 and the arbitrary monocular measurement position *C**.

8.4. Optimal projection

Discussion has been focused on the test plane 3D coordinates generated by the aiding triangulation SfM setup made in the coordinate system of an aiding camera whose placement is on the opposite side of the MAAP (i.e. C_2 in Fig. 8.2 and Fig. 8.3) with respect to the arbitrary monocular measurement position *C**. Naturally, the question arises if one can use the test plane 3D coordinates from an aiding position on the

same side of the MAAP as *C** (i.e. C_1 in Fig. 8.2 and Fig. 8.3). While this is possible, there is a risk that *C** is placed coincident or near coincident with C_1. C_1 coincident with *C** is a degenerate configuration and the Fundamental matrix cannot be obtained [37]. Near coincidence is also not desirable as computational determination of the Fundamental matrix becomes numerically ill-conditioned [36]. To avoid this coincidence risk and maximize flexibility in positioning *C**, one has the option to use either of the aiding positions of the initial aiding triangulation SfM setup, one on each side of the MAAP such that the images of *C** are synthesized using the aiding position on the opposing side.

Since either aiding position C_1 or C_2 can be utilized, let us now denote them generally as aiding position *C*. Even by using a proper distance between the aiding position *C* and arbitrary monocular measurement position *C**, there will still be some unavoidable error of the measured *C** projection compared to the ground truth. This is due to imaging noise that is assumed to be Gaussian. Thus, projecting from a triangulated 3D point made in the reference frame of *C* to the *C** image plane will have some deviation from the Epipolar constraint. This leads to the *C** point not lying along the Epipolar line. To overcome this, the Sampson distance is utilized [37]. Shown in Fig. 8.4, for the unknown ground truth projection $\hat{x}$, it lies somewhere along the Epipolar line l while the measured projection x is located some distance δ away from $\hat{x}$. The distance δ is the error induced by imaging noise and can never be determined perfectly such that $\delta = 0$. Thus, the operation $x - \delta$ cannot be performed to restore $\hat{x}$. Instead, the minimal distance from $\hat{x}$ that still lies along l is the point $x_\perp$ that lies at the foot of the perpendicular. Therefore, to minimize error from the *C** ground truth image, for each and every measured projection x, its perpendicular point $x_\perp$ is used to artificially synthesize views of *C**.

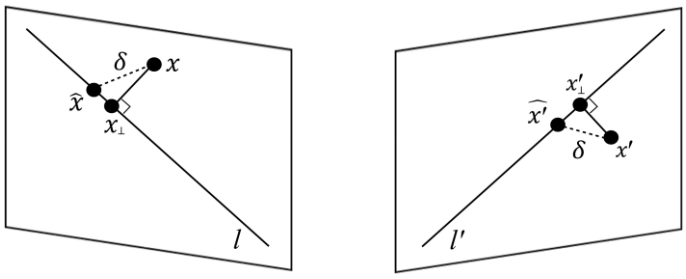

Fig. 8.4 *Sampson distance - the optimal correction of observed correspondences $x_\perp$ and $x'_\perp$ lie at the foot of the perpendicular to the Epipolar lines l and l' [82].*

8.5. View synthesis in practice

Using only six patterns to encode each point with a temporal gray-value intensity sequence leads to decreased surface precision as well as a greater number of outliers compared to using 20-30 patterns [14]. This is due to the reduction of temporal data in the intensity sequence such that the sequence can be more easily corrupted by imaging noise. This leads to false test plane depth correspondences, resulting in point cloud outliers. Projecting from outliers to the arbitrary monocular measurement position C^* would result in an incorrectly synthesized image. These outliers are removed using a statistical filter algorithm provided by CloudCompare [83]. For every 3D point in the reconstructed point cloud of a test plane increment, 200 of its nearest neighbours are taken and the sample mean and sample standard deviation distance from the 3D point is determined. Points outside a distance of 0.15σ are removed with the statistical outlier filter. The remaining 3D points are then projected to the arbitrary monocular measurement position C^* image plane. Since a projection is determined by finding the intersection of a light ray with the image plane as described by Eqn. 8.3 - 8.5 (pg. 75), then it is likely that a projection falls between image pixels. To artificially create the C^* view through digital image synthesis, then the gray-value intensities for integer-value pixels are needed. The intensities at integer pixels are obtained through bi-cubic interpolation of the scattered grid of intensities belonging to the projected real-valued pixels.

Shown in Fig. 8.5a is a synthesized image of an aperiodic fringe pattern incident on the test plane without statistical filtering to remove outliers. In Fig. 8.5b is a synthesized image with statistical filtering. Fig. 8.5c shows the real ground truth image of the test plane actually taken at the arbitrary monocular measurement position C^*. In practice, the entire view cannot be completely shared between the aiding position C and C^*. However, in Fig. 8.5a, artefacts arise as seen on the left side of the synthesized image where no gray-value intensities should be present. These artefacts are due to some outliers whose C^* projection happens to fall into the unshared view area. Due to the sparseness of these outlier projections, the gray-value intensity is interpolated from this sparsely scattered data, resulting in the ‘smearing’ artefact effect. Contrast this with Fig. 8.5b where statistical filtering removes these outliers resulting in the absence of artefacts in the unshared view area.

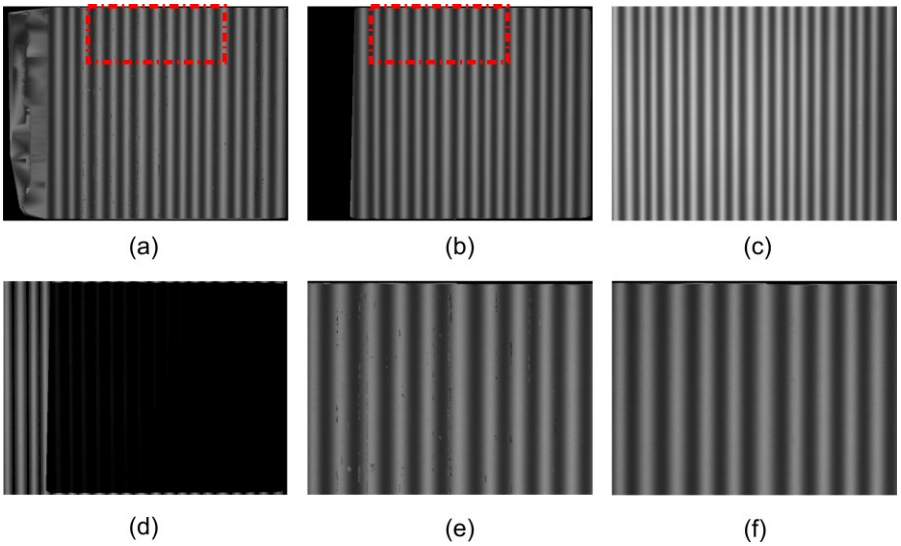

Fig. 8.5 *(a) Artificially synthesized image of a projected aperiodic fringe pattern incident on a test plane increment generated without implementing statistical outlier filter, (b) artificially synthesized image generated implementing statistical outlier filter, (c) real ground truth image taken by camera, (d) absolute difference of gray-scale intensities between histogram matched images of Fig.b and Fig.c, (e) zoomed image of the area denoted in Fig.a showing artefacts, and (e) zoomed image of area denoted in Fig.b showing no artefacts [82].*

After matching the histograms of the synthesized and real ground truth images, the absolute difference of gray-scale intensities between them can be seen in Fig. 8.5d. The synthesized image and real ground truth image resemble well one another due to the low absolute difference observed between the two. Histogram matching was required due to the different contrast between the two images because of the aperture size of the camera at the aiding position C and the single camera at the arbitrary measurement position C^* cannot be perfectly matched. Shown in Fig. 8.5e is a zoom of Fig. 8.5a where one can see holes in the synthesized image. However, these are not holes but areas where outliers were projected whose lower gray-value intensity belonged to a completely different part of the image, thus not matching the higher gray-value intensity of its surroundings. Conversely, the same effect can also be observed in the same synthesized image where higher gray-value intensities can be seen that do not belong to the lower gray-value intensity of its surroundings. Shown in Fig. 8.5f is the zoom of Fig. 8.5b. It can be observed that the ‘hole’ artefacts are no longer present due to the statistical filtering of outliers.

For every test plane increment through the measurement volume, the normalized cross-correlation between each pixel of the artificially synthesized image with statistical outlier filtering and each pixel of the real ground truth image is performed.

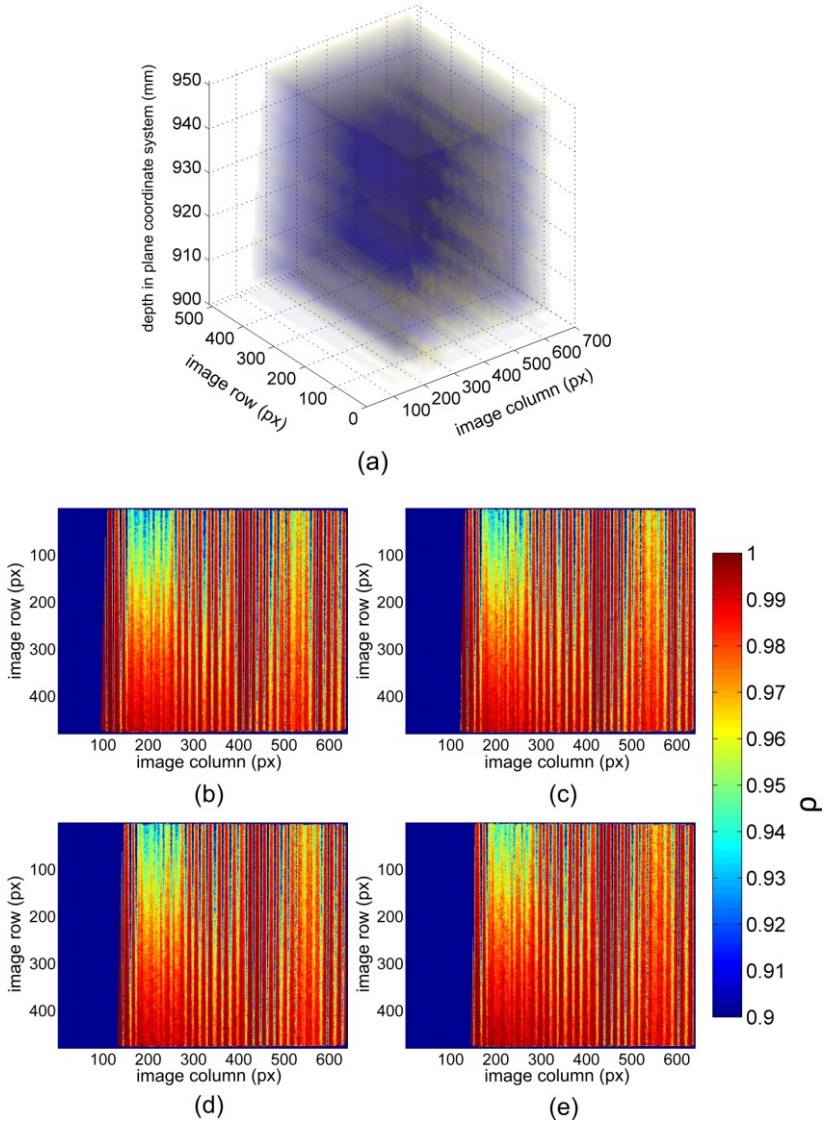

Fig. 8.6 *(a) Cross-correlation between real images and artificially synthesized images for each test plane increment through the measurement volume. Yellow points indicate ρ < 0.93 and blue points indicate ρ ≥ 0.93, (b) cross-correlation at depth slice 900.0 mm away from the camera in the test plane coordinate system, (c) 912.5 mm, (d) 925.0 mm, (e) 937.5 mm [82].*

This compares the similarity between synthesized gray-value intensity sequences to its ground truth. This similarity is shown in the volumetric plot of an aperiodic fringe pattern sequence seen in Fig. 8.6a. Synthesized pixels with a cross-correlation of $\rho < 0.93$ are displayed in yellow while those with $\rho \geq 0.93$ are displayed in blue. The remaining points with $\rho < 0.93$ can be seen in Fig. 8.6a to occur in columns. This is due to fringes occurring nearly parallel to the pixel columns of the imaging chip, resulting in pixels along image columns to share the same intensity sequence. Repeating the volumetric calculation for a total of ten different aperiodic pattern sequences, it was found that an average of 90% of the gray-value intensity sequences have a cross-correlation coefficient of $\rho \geq 0.93$, indicating that the artificially

synthesized images resemble closely the real ground truth images. Shown in Fig. 8.6b-e is the cross-correlation coefficient at depth slices through the volume for distances 900.0, 912.5, 925.0, and 937.5 mm.

8.6. 3D measurement performance

Using the same experimental MAAP emulator test-bed setup introduced in Sec. 7.8 (pg. 64) to be able to test for measurement consistency over different pattern sequences, the camera positions C_1 and C_2 of the initial aiding triangulation SfM setup that image the test plane were positioned 250 mm on both sides of the MAAP along a common baseline such that the baseline distance between them was 500 mm. 3D points belonging to each test plane increment are then reconstructed through triangulation between C_1 and C_2. This initial aiding triangulation SfM setup was then removed. A camera was then arbitrarily placed such that there is no knowledge of its true position. This arbitrary monocular measurement position C^* for which the view of each test plane increment is artificially synthesized utilizes the triangulated 3D points made in the coordinate system of the aiding position C_2. Thus, C^* was on the opposite side of the MAAP with respect to the position of C_2 to avoid the risk of coincident/near-coincident camera positions.

The configuration was set up such that the first test-plane increment was manually placed 900 mm from the baseline. The calibration jig was composed of two Samsung SA200 monitors and setup about 2 m from the baseline. The test-plane surface was oriented orthogonal to the illumination direction. For the 200×200×50 (H×W×D/mm) measurement volume, the plane is incremented in steps of 0.5 mm using a linear translation stage. The camera at aiding positions C_1, C_2 along with the arbitrary monocular measurement position C^* used the same grayscale AVT Pike F-032B with focal length f = 25 mm, quadratic pixel pitch of Δp = 7.4 µm, and resolution of 640×480 px. imaging in 16-bit gray-scale mode.

To evaluate the 3D measurement performance when measuring a complex surface, the white scattering ceramic bust in the shape of German poet Johann Wolfgang von Goethe was again used as a test object. It had dimensions of ~80×50×40 (H×W×D/mm). Shown in Fig. 8.7a is the bust of Fig. 8.7b illuminated by aperiodic fringes for a sequence of six patterns. Shown in Fig. 8.7c is the bust 3D

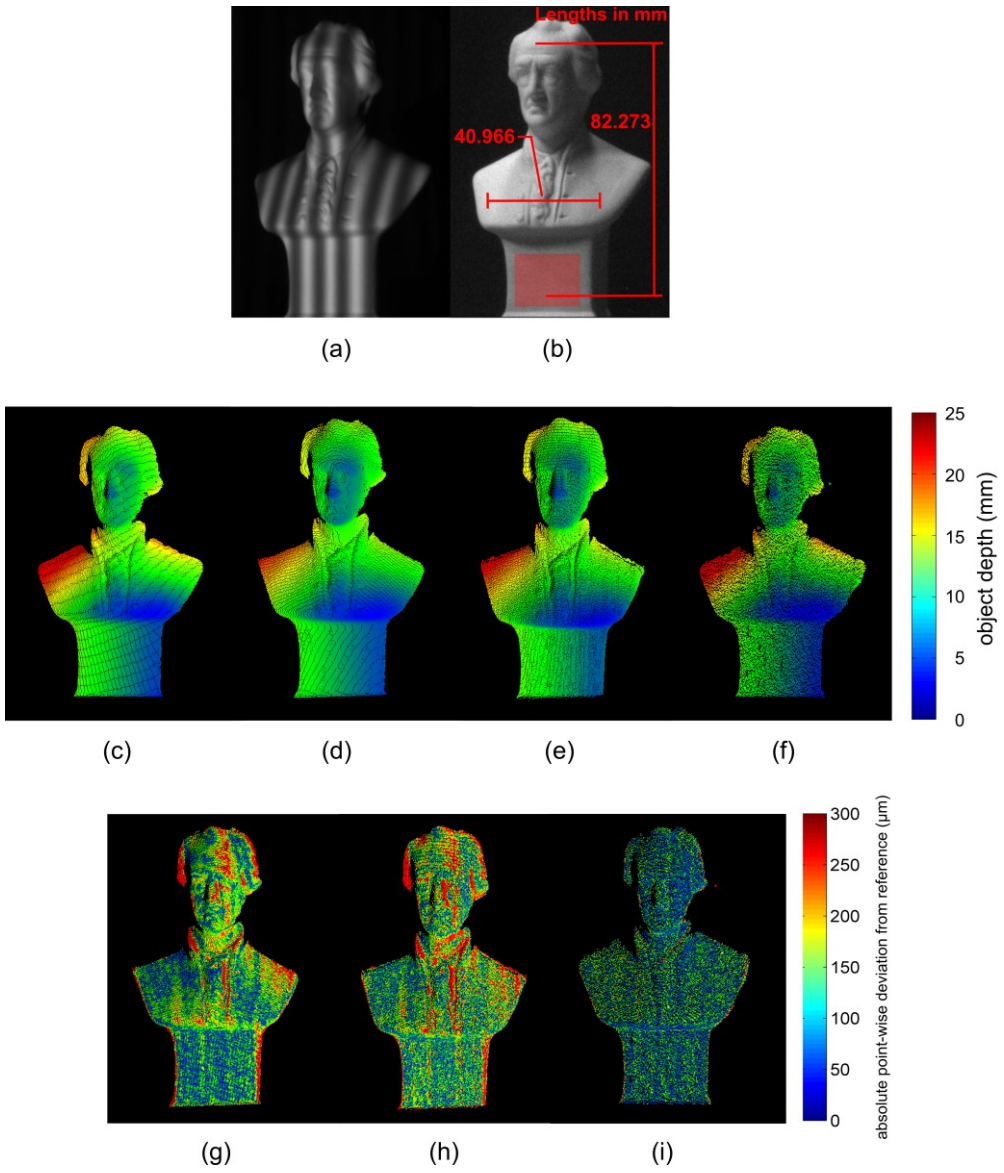

Fig. 8.7 *Bust object in the shape of Johann Wolfgang von Goethe (a) illuminated by an aperiodic fringe pattern, (b) bust object indicating vertical and horizontal reference lengths for measurement comparison and the selected area for completeness comparison, (c) reference point cloud, (d) single-camera 3D reconstruction generated using true calibration images, (e) single-camera 3D reconstruction generated using synthesized calibration images, (f) stereo-camera 3D reconstruction generated using same pattern sequence as (d-e), (g) point-wise deviation of Fig.d from reference, (h) point-wise deviation of Fig.e from reference, (i) point-wise deviation of Fig.f from reference [82].*

reconstructed by a separate stereo-photogrammetric system. As was previously done in Sec. 7.8 (pg. 64), this separate system projects thirty Band-Limited Patterns (BLP) capable of measurement surface standard deviation of σ = 26.7 µm from the ground truth without outlier filtering [15]. It is used to obtain the surface topology of the bust as close to ground truth as possible. The 3D reconstruction generated is utilized as a reference point cloud. Fig. 8.7d-e show the point clouds generated implementing real

calibration images and artificially synthesized calibration images, respectively. Fig. 8.7f shows the point cloud generated with a stereo-camera system employing the same aperiodic fringe sequence and the same cross-correlation threshold of $\rho \geq 0.99$ to determine correspondences. The absolute point-wise deviation σ_{bust} of these point clouds from the reference point cloud can be seen in Fig. 8.7g-i. Fig. 8.8b also indicates vertical and horizontal lengths obtained from the generated point clouds, where the length error Δl is the absolute difference between the measured lengths and the reference such that $\Delta l = |l_{measured} - l_{ref}|$. In addition, the area where the completeness is determined is also highlighted. The point clouds for a total of ten different aperiodic fringe sequences were generated and the aggregated absolute point-wise deviation σ_{bust} obtained in Fig. 8.8 and Table 8.1. Outlier points of each point cloud with deviation > 1 mm from the reference are filtered. Table 8.2 shows the average percentage of outliers removed and the completeness of the ten point clouds. The quantiles of the length error Δl of the ten point clouds are shown in Fig. 8.9 and Table 8.3.

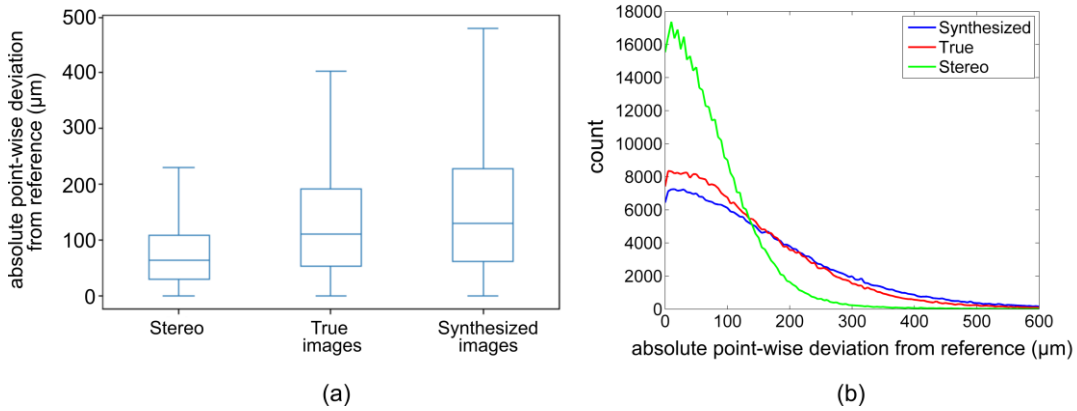

Fig. 8.8 *Comparing the absolute point-wise deviations generated between the 3D reconstructed bust and reference when using true calibration images, synthesized calibration images, and a stereo-camera setup. Graphs aggregate all points from ten reconstruction samples: (a) Quantile boxplot of deviations and (b) distribution of deviations [82].*

	25th percentile σ_{bust} (µm)	Median σ_{bust} (µm)	75th percentile σ_{bust} (µm)
Stereo	29	62	108
True	53	112	198
Synthesized	60	129	228

Table 8.1 *Quantiles of absolute point-wise deviations when using true calibration images, synthesized calibration images, and a stereo-camera setup.*

	Outliers removed (%)	Completeness (%)
Stereo	0.5	96.7
True	2.7	99.8
Synthesized	3.2	99.7

Table 8.2 *Percentage of outliers removed and completeness when using true calibration images, synthesized calibration images, and a stereo-camera setup.*

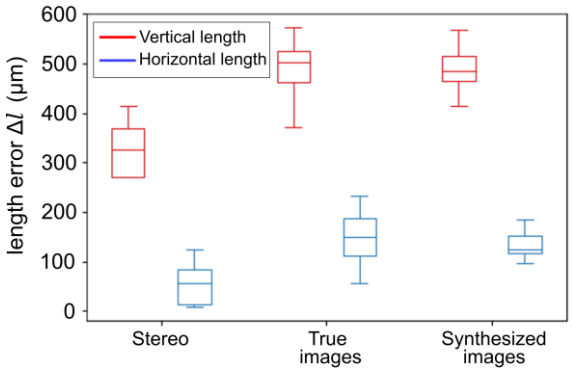

Fig. 8.9 *Comparing the length errors generated when using true calibration images, synthesized calibration images, and a stereo-camera setup between the 3D reconstructed bust and reference [82].*

	Vertical length		
	25th percentile Δl	Median Δl	75th percentile Δl
Stereo	232	327	370
True	453	502	532
Synthesized	463	485	527
	Horizontal length		
	25th percentile Δl	Median Δl	75th percentile Δl
Stereo	13	57	89
True	105	150	192
Synthesized	112	125	160

Table 8.3 *Quantiles of length errors Δl when using true calibration images, synthesized calibration images, and a stereo-camera setup.*

Similar to Sec. 7.8 (pg. 64), it can be seen from Fig. 8.8 and Fig. 8.9 that the stereo setup generates 3D reconstructions with better point-wise deviation and length error. This better general 3D measurement performance is expected due to the stereo setup using the same Epipolar geometry as the setup generating the reference. The stereo setup again showed decreased completeness and fewer outliers. This again

indicates that the identification of corresponding depth while scanning along a light ray in the single camera setup is more easily detectable for N = 6 patterns compared to the identification of stereo-correspondences scanning along an Epipolar line.

Comparing the point-wise deviations in Fig. 8.8 between using true calibration images and synthesized calibration images in the single-camera setup, it can be observed that the measurement performances are nearly similar when using either one as the point-wise deviation distribution and percentiles nearly overlap. However, there is some degradation in performance when using synthesized calibration images as there is a greater number of outliers removed. This is due to some areas in the measurement volume where the synthesized calibration images were not accurately reproduced in comparison to the real calibration images. This results in greater likelihood that a false corresponding depth is identified rather than the ground truth corresponding depth. Observing Fig. 8.9, the difference in length errors between the two methods show similar results as their percentiles overlap. This is as expected as they share the same pose between the camera and the test plane, meaning that the geometry of any 3D reconstruction will be the same between the two. The high similarity in the measurement performances shows that the synthesized images resemble well the ground truth images and their usage is still appropriate for 3D measurement with a monocular view.

8.7. Conclusion

A posterior method was proposed and demonstrated to intrinsically ‘re-calibrate’ the MAAP such that it decouples intrinsic MAAP calibration from the camera view. Any arbitrary camera placement for 3D measurement becomes possible without having to repeat the entire MAAP intrinsic calibration process. This results in significant savings in time and labor. The proposed method requires an initial aiding Structure from Motion (SfM) setup to image the projected patterns onto each test plane increment and triangulate them. With the aid of a simple calibration jig permanently mounted in the background, the projected patterns on each test plane increment can be artificially synthesized from the view of any subsequent monocular arbitrary measurement position through re-projection from the triangulated test plane. The artificial view can be synthesized even if the initial aiding triangulation SfM setup is not currently present. Only the images of the aiding camera at each aiding SfM position need to be stored

for later use. This is similar to how the intrinsic parameters of a pinhole modelled device can be stored for future alternate setup configurations. The only caveat is that the orientation between the MAAP and calibration jig is fixed, and that some points on the jig are shared by both the aiding SfM setup view and the arbitrary measurement camera view.

The developed method as summarized in Table 8.4 is readily accessible. It requires minimal additional devices available to any machine vision facility and does not need bespoke hardware customized solely for this application.

Intrinsic MAAP ‘re-calibration’ for arbitrary measurement camera views through artificial view-synthesis	
1.	Image each test plane increment using an initial aiding triangulation Structure from Motion (SfM) setup. Store the images of the projected patterns from each camera aiding position.
2.	Have the camera at each aiding position of the initial aiding triangulation SfM setup image a background fixed calibration jig whose setup is described in Sec. 8.3. Determine the Epipolar geometry between SfM aiding positions. Store the images.
3.	Triangulate each test plane increment and statistically filter 3D reconstructed outliers according to Sec. 8.5.
4.	Remove the initial aiding triangulation SfM setup.
5.	Install a camera at an arbitrary desired measurement position to perform monocular 3D measurement, taking care to avoid degenerate positions as described in Sec. 8.4.
6.	Image the calibration jig from the monocular arbitrary measurement position.
7.	Determine the Epipolar geometry between an aiding camera position C and monocular arbitrary measurement camera position C^* using the well-established eight-point algorithm.
8.	Re-project the test plane increment 3D points of step 3 to the image plane of C^* according to Sec. 8.4 to artificially synthesize the MAAP calibration images as if they were taken from the view of C^*.
9.	Extrinsically calibrate monocular arbitrary measurement camera position C^* according to Chapter 7.
10.	Completing steps 8 and 9 enable single camera 3D reconstructions with the arbitrary position without having to perform the entire intrinsic calibration process of Chapter 6 again.

Table 8.4 *Summary of the intrinsic MAAP ‘re-calibration’ through view synthesis method proposed in this chapter.*

With the presented view synthesis method, it was shown that the artificial calibration images resemble well the actual ground truth images. As expected, 3D measurements generated using artificial calibration images were slightly degraded from 3D measurements generated using real calibration images but still similar. Due to their measurement similarities, they share the same characteristic that they both perform worse compared to a stereo setup but achieve greater point cloud completeness.

Chapter 9.

Future work

The following list short-term and long-term prospects for improving this work in the future. Short-term prospects are ones that can be more readily studied in comparison to more involved work required with long-term prospects.

9.1. Short-term prospects

9.1.1. Optimizing pattern design

In active structured-illumination photogrammetry, the accurate assignment of correspondences is an important factor in 3D measurement performance [84]. The importance of accurate correspondence assignment also applies to this study as an accurate corresponding depth is needed to be mapped to each object point. Since correspondence assignment is dependent on cross-correlation matching of temporal gray-value intensity sequences, then the uniqueness of sequences throughout the measurement volume plays a key factor. Hence, the spatial bandwidth of aperiodic fringes projected needs to be optimized to enhance 3D measurement performance. The bit depth of captured images also plays a role. Throughout this study, a single spatial bandwidth of aperiodic fringes was used such that fringe widths were between 7 and 10 mm throughout the measurement volume for 16-bit depth images. To further investigate the affect of pattern design and image bit depth, bandwidth and bit depth optimization can be systematically studied through Physically Based Rendering (PBR) ray-tracing simulation to generate realistic synthetic pattern projection and image capture of the simulated measurement setup [84, 85]. The study can then be applied experimentally to validate the synthetic results in terms of measurement performance and reconstruction completeness.

9.1.2. Distortion compensation

This work implemented a camera with objective lens with a low degree of geometric image distortion and distortion compensation models were not needed to be applied.

However, possible future work involves including distortion modelling such as the well accepted Brown-Conrady model [30] into the artificial view synthesis process as it will generalize the method to be robust against stronger geometric optical distortions. This results in 3D measurement performance robustness in the presence of optical distortions as the directional vectors of light rays would not deviate greatly from their ground truth directions. Other possibilities include implementing different camera intrinsic calibration methods where 3D measurement performance is more robust against strong geometric distortions. Such methods include ray-based camera calibration [86] where distortions are not parametrically modelled, but the directional vectors of all incoming light rays to the camera are directly obtained.

9.2. Long-term prospects

9.2.1. Improved data processing

The speed of 3D measurement can be considered in two stages: data acquisition and data processing. The data acquisition side pertains to the speed at which all the necessary patterns are captured to perform 3D reconstruction. The data processing side pertains to how fast calculations can be made to create a single 3D reconstruction.

In temporal cross-correlation based methods, work has been proposed to speed up the temporal cross-correlation based correspondence search itself. The cross-correlation calculation is computationally expensive as it requires multiple square-root operations [87]. A fast coarse search has been proposed where the Hamming distance is matched between hashed binarized intensity sequences [88]. The hash function binarizes based on if the intensity is greater or less than the mean of the temporal intensity sequence. The coarse correspondence is then refined using the temporal cross-correlation between the original intensity sequences, as opposed to performing the entire search using the temporal cross-correlation. Implementing the proposed BiCOS algorithm could also yield computational speed improvements in the temporal cross-correlation correspondence depth search process of this work.

Another possibility is the implementation of Machine Learning techniques. A study by Wong et al. has shown that even simple Multilayer-Perceptron neural networks (MLPs) can directly obtain the 3D coordinates of a surface simply by

inputting the stereo-correspondences [89]. One-to-one mappings between pre-determined stereo-correspondences and their 3D coordinates can be used as training data for the MLP to approximate a model relating stereo-correspondences and 3D coordinates directly. Therefore, not only is the triangulation calculation no longer needed to produce a 3D reconstruction, but intrinsic camera calibration and extrinsic system calibration also do not need to be performed either.

Extending from this, Van der Jeught et al. implemented a Feed-Forward Convolutional Neural Network (CNN) [90]. First, a large set of synthetic surface depth maps was generated and the phase modulation of a single projected sinusoidal fringe pattern was simulated as a 128×128 px. image. The set of phase modulated images and corresponding depth maps are used as training data to approximate a mapping between the depth distribution and grayscale intensities resulting from phase modulation of the surface. It was recently shown experimentally by Nguyen et al. [91] that the trained CNN can produce a single 3D reconstruction in a single shot without intermediate processing steps of phase demodulation and phase unwrapping. However, additional care needs to be taken as the one-to-one mapping would no longer hold if phase ambiguities exist along depth for a single image pixel. The opportunity to utilize the aperiodic fringe pattern modality arises to train the CNN to approximate a mapping between grayscale intensity sequences and the surface depth distribution. Thus, the cross-correlation search is no longer needed to determine the corresponding depth of points on the object. However, additional investigation needs to be performed to determine if fringe aperiodicity provides sufficient intensity sequence uniqueness in the lateral and depth direction for the one-to-one mapping to hold.

9.2.2. Uncooperative surfaces

The 3D measurement of uncooperative, non-Lambertian (i.e. non-diffuse scattering) surfaces is still an active area of research as they still pose significant challenges due to the variety of possible materials. This challenge is exacerbated with the MAAP's multiple projection channels.

Due to the changing angle of illumination incidence, the Bidirectional Reflectance Distribution Function (BRDF) of the surface causes a scaling factor K to arise that is dependent on the angle of illumination incidence. Assuming that the

calibrating test plane is nearly Lambertian, for non-Lambertian object surfaces, the normalized cross-correlation becomes

$$\rho(KO,P)=\frac{\sum_{i=1}^{N}(K_iO_i-\overline{K_iO})\cdot(P_i-\bar{P})}{\sqrt{\sum_{i=1}^{N}(K_iO_i-\overline{K_iO})^2}\sqrt{\sum_{i=1}^{N}(P_i-\bar{P})^2}}\neq\rho(O,P) \tag{10.1}$$

where K_iO_i are the angle-dependent gray-value intensities when imaging the object. Therefore, a decrease in the cross-correlation coefficient occurs when a measured object has different material properties (hence different BRDF) than the test plane. This can cause a greater chance in a false corresponding depth being determined for an object point, resulting in a decrease in 3D measurement performance.

To overcome this such that *K* is no longer angle-dependent, a possible solution is inspired from a study by Heist et al. where a stereo-view configuration with a MAAP was simulated to overcome the challenges of measuring isotropic reflective surfaces [92]. However, a significant challenge still remains where the extrinsic calibration to determine the rotation and translation of each projection channel with respect to a world coordinate system has not yet been shown experimentally.

A possible solution is to first generate a preliminary point cloud of the object so that every surface normal $\vec{n}$ of each observed object point can be determined. A ray-based calibration on each projection channel of the MAAP to obtain the directional vectors of projected rays. Thus, the illumination angle $\vec{\imath}$ and camera viewing direction $\vec{v}$ with respect to $\vec{n}$ for every object point can be determined. Utilizing a BRDF (Ashikhmin-Shirley in the isotropic reflective case [93]) that models the material's radiometric behaviour, the reflected radiance $f(\vec{\imath},\vec{v})$ in the direction of $\vec{v}$ can be determined using $\vec{\imath}$, $\vec{v}$, and $\vec{n}$, relative to a Lambertian surface where *f* is invariant to viewing direction. Thus, a correction factor $f^{-1}(\vec{\imath},\vec{v})$ is multiplied with the gray-value intensity of each pixel resulting in an image of the object as if it were a Lambertian surface and is done for every image in the illumination sequence. This correction procedure could possibly be similarly adapted to different uncooperative materials depending on an appropriate radiometric model.

Chapter 10.

Summary and conclusion

Active structured illumination optical photogrammetric non-tactile 3D sensors based on triangulation utilizing conventional pinhole modeled digital cameras and projectors have a well-established calibration pipeline. Commercial off-the-shelf digital cameras such as CCD or CMOS based devices, and digital projectors based on DMD or LCoS devices are intrinsically calibrated using the well-known Zhang method and the intrinsic parameters stored for re-use in alternate setup configurations [10]. Extrinsic calibration of the Epipolar geometry between the triangulating devices is performed using the eight-point algorithm popularized by Hartley [11].

However, novel chip-less projection systems designed to achieve high projection rates, and therefore high 3D measurement speed, cannot be pinhole modelled. Without a digital chip or projection lens, this class of projectors does not physically have the traditional intrinsic parameters of focal length and principal point of the digital chip. Attempting to apply the Zhang method and the eight-point algorithm would therefore not be possible.

This work investigated a six-pattern aperiodic sinusoidal fringe projecting array projector known as a Multi-Aperture-Array Projector (MAAP) that belongs to this class of chip-less projection systems. The purpose was to measure matte-like near-Lambertian surfaces for cubic meter scale measurement volumes at meter-scale working distances. The camera is considered to have negligible distortion. As the MAAP could not be intrinsically calibrated, a previous iteration of 3D sensor utilizing this projection device required a stereo-camera setup calibrated through the Zhang-Hatley pipeline. However, no calibration pipeline exists that can be applied to enable a functional monocular-view setup.

This work has developed that needed calibration pipeline as seen in the flowchart of Fig. 10.1. Achieving the calibration pipeline required a series of accomplishments that were demonstrated in this work.

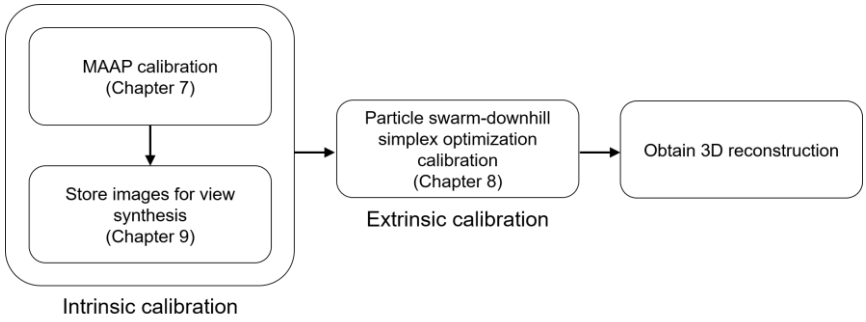

Fig. 10.1 *Flowchart depicting the components of the calibration pipeline developed in this work.*

In Chapter 6, the following series of accomplishments were demonstrated towards achieving intrinsic calibration of the MAAP:

- A generalized model-less method to intrinsically calibrate the MAAP using a single recording camera was proposed and experimentally demonstrated. An intensity sequence-to-depth mapping forms a volumetric 3D grid of temporal gray-value intensity sequence data. This grid is therefore coupled to the specific camera view. This is achieved by imaging in discrete depth slices the illumination of each frustum incident on a scanned test plane. For an object illuminated by the same pattern sequence, the corresponding depth of each object point was determined through normalized temporal cross-correlation matching of point's gray-value intensity sequence with the volumetric 3D grid. It was observed that the surface precision and length accuracy of the generated 3D reconstructions agreed well with plane and sphere gauges. The caveat to the proposed intrinsic calibration method is that the projected patterns must be temporally repeatable.

- 3D reconstructions were still capable of being generated of spatially disconnected scenes and of bust figures whose complex surface contained highly varying gradients. Each image pixel independently produced a 3D point in the reconstruction.

- It was experimentally demonstrated that the proposed intrinsic calibration method was capable of ~3 times greater measurement volume depth compared to the previous state-of-the-art phase-to-depth look-up-table mapping methods that utilized purely sinusoidal periodic fringes.

- The generalization of the intrinsic calibration method facilitated the creation of a novel repeatable speckle projector that projects two-dimensional spatially varying speckle patterns. It was experimentally demonstrated that 3D reconstructions using a single camera could be generated. Hence, the proposed calibration method is not exclusive to aperiodic sinusoidal fringe patterns. Previous state-of-the-art speckle projectors generated temporally random patterns that required a stereo-camera setup.

In Chapter 7, the following series of accomplishments were demonstrated towards achieving extrinsic calibration of a monocular MAAP 3D measurement system:

- A downhill-simplex optimization method initialized by particle-swarm was developed and experimentally demonstrated using a MAAP emulator. It extrinsically calibrated the monocular system by determining the orientation and position between the camera and scanned test plane. The proposed method does not require *a priori* knowledge of an appropriate initial guess as with the conventional downhill-simplex method alone. It was experimentally observed that the surface precision and length accuracy of the generated 3D reconstructions agreed well with reference point clouds of a complex surface bust figure with high reconstruction completeness. The resulting 3D measurement performance for the same projected pattern sequence was seen to be degraded from the stereo-camera case but with greater point cloud completeness.
- Several objective functions in the proposed optimization method were investigated through brute-force search of the parameter space. The parameter space is the rotation and depth translation between the test plane and camera. A minimizing objective function was designed and dubbed the ‘Ellipsoidal Absolute Deviation’ (*EAD*). The *EAD* is the sum of absolute deviations from the ground truth radius along each major axis of the 3D reconstruction of a sphere gauge. The *EAD* visually displayed a convex error surface with an unambiguous global minimum.
- To ensure that minimizing the *EAD* delivers consistently the rotation and depth translation that best extrinsically calibrate the monocular MAAP system, a large trial of random parameter initializations was carried out. It was observed that the *EAD* delivers consistently the rotation and depth translation over all trials

regardless of the initialization. This was compared to the conventional downhill-simplex method where it was observed that it could end up trapped in incorrect local minimums if the initial guess was not appropriate.

In Chapter 8, the following series of accomplishments were demonstrated towards achieving a method to store an intrinsic calibration for later use in arbitrary measurement camera views in a setup. Thus, the intrinsic calibration does not have to be performed again:

- A posterior method was developed and experimentally demonstrated using a MAAP emulator that intrinsically ‘re-calibrates’ the single camera setup for any arbitrary measurement camera view. The method requires a simple calibration jig installed permanently in the background of the measurement volume and an initial aiding triangulation Structure from Motion (SfM) setup. It was shown that the volumetric 3D grid of temporal gray-value intensity sequence data can be artificially synthesized for arbitrary single measurement camera views through re-projection from previously triangulated test plane 3D points. As long as the initial aiding SfM setup images are stored, it does not have to be physically present during 3D measurement with the arbitrary single camera view. The only stipulations are that the orientation between the MAAP and calibration jig must remain fixed and a shared set of viewed points on the jig are viewed by the SfM setup and the arbitrary single measurement camera.
- It was observed experimentally that the synthesized calibration images resemble closely the ground truth calibration images. The 3D reconstructions generated utilizing either image was seen to be similar as the surface precision, length accuracy, and reconstruction completeness of a bust figure were nearly the same. Due to this similarity, the resulting 3D measurement performance for the same projected pattern sequence in the synthesized calibration case was also seen to be degraded from the stereo-camera case but with greater point cloud completeness.

In Chapter 9, short-term and long-term improvements were discussed. Short-term improvements that can be more readily performed involve the experimental investigation to optimize fringe spatial bandwidth of the aperiodic pattern design in

order to enhance 3D measurement performance and reconstruction completeness. Implementation of distortion compensation models or alternate camera calibration techniques more robust to distortion could also be done in the future. Long-term improvements include the implementation of faster correspondence search algorithms and reconstruction algorithms, and the experimental execution of compensation techniques to overcome 3D measurement challenges arising from uncooperative non-Lambertian surfaces.

In conclusion, this work accomplished a series of achievements culminating in the overall goal of producing a novel calibration pipeline that realizes monocular 3D sensors utilizing an aperiodic sinusoidal fringe projecting MAAP. Overcoming this challenge therefore brings about potential decreases in cost, weight, and form-factor of such 3D sensors utilizing such an array projector in photogrammetric optical metrology applications. Besides the immediate benefits, the presented pipeline could one day provide a potentially new avenue to create a portable 3D sensor that is better ergonomically suited. A single, freely-moving, handheld camera only needs to be carried to perform 3D measurement if artificial view synthesis could be implemented in real-time.

Appendix A

Zhang method

Intrinsic calibration of pinhole cameras is typically performed through Zhang Calibration (ZC) where a planar pattern must be imaged. This planar pattern can be of any design as long as it contains identifiable fiduciary control points such that each point on the plane can be designated with 3D coordinates in the world reference frame *(X,Y,Z)*, and their projections assigned homogeneous pixel coordinates in the image plane *(α,β,1)*. The origin of the world reference frame can be defined arbitrarily on the plane. A chessboard is commonly used as each corner can be easily detected using the Harris detector [94]. Since a plane is used, the 3D coordinates of all planar points can be reduced such that $Z = 0$.

Let K be the 3×3 intrinsic matrix and [*r1 r2 r3* $\vec{t}$] be the columns of the 3×4 extrinsic matrix relating the pose between the world reference frame and the camera reference frame based at the pinhole. [*r1 r2 r3*] are the columns of the rotation matrix R and t is the translation vector. While an analytic closed-form solution to determine K exists, estimation of K has been found to be handier through non-linear minimization of the sum of the absolute square error

$$\min\left(\sum_{i=1}^{n}\sum_{j=1}^{m}\left|m_{ij}-\widehat{m}\left(K,k,p,R_i,\overrightarrow{t_\iota},M_j\right)\right|^2\right) \tag{A1}$$

where n is the number of plane positions, m is the number of identifiable planar points, R_i is the rotation between coordinate systems of the camera and the i^{th} plane position, $\overrightarrow{t_\iota}$ is the translation between coordinate systems of the camera and the i^{th} plane position, m_{ij} is the observed image point, M_j is the plane point (X Y 0), k are the radial distortion coefficients depending on the order of the radius considered [*k1 k2 k3* …], and p are the tangential distortion coefficients depending on the order of the radius considered [*p1 p2 p3* …]. Eqn. A1 can be minimized using the Levenberg-Marquardt algorithm. Thus, the goal of the optimization is to minimize the deviation between the observed image point and the re-projection of M_j onto the image plane, $\widehat{m}\left(K,k,p,R_i,\overrightarrow{t_\iota},M_j\right)$, where K, k, p, R_i, $\overrightarrow{t_\iota}$ are the iterated parameters.

Appendix B

Eight-point algorithm

The Epipolar constraint where *p′ = (x′ y′ 1)* and *p = (x y 1)* are corresponding stereo-pairs expressed in homogeneous image coordinates is

$$p'^{T} F p = 0 \tag{B1}$$

To determine the elements of the 3x3 Fundamental matrix *F*, at least eight Epipolar constraint equations arising from eight corresponding pairs are required. With $n \geq 8$ equations, the expanded form of Eqn. B2 is obtained where A is the stack of corresponding pairs and *f* is the stack of fundamental matrix elements.

$$\begin{bmatrix} xx' & xy' & x & yx' & yy' & y & x' & y' & 1 \\ & & & \vdots & & & & & \end{bmatrix}^{T} \begin{bmatrix} F_{11} & F_{12} & F_{13} & F_{21} & F_{22} & F_{23} & F_{31} & F_{32} & F_{32} \\ & & & \vdots & & & & & \end{bmatrix} = 0 \tag{B2}$$

$$Af = 0$$

Eqn. B2 is a problem that can be solved using Singular Value Decomposition (SVD). Care must be taken to normalize the coordinates of each point, such that the average distance of all points to the center of the image is equal to $\sqrt{2}$. The obtained *F* matrix is then denormalized as this form satisfies the Epipolar constraint between corresponding pairs using homogeneous pixel coordinates.

References

1. ERRICKSON, David. *Shedding Light on Skeletal Remains: The Use of Structured Light Scanning for 3D Archiving*. Academic Press, 2017. DOI: 10.1016/B978-0-12-804602-9.00008-4

2. FYFFE, Graham, HAWKINS, Tim, WATTS, Chris, MA, Wan Chun and DEBEVEC, Paul. Comprehensive facial performance capture. In : *Computer Graphics Forum*. 2011. p. 425–434. DOI: 10.1111/j.1467-8659.2011.01888.x

3. BUSCH, Christoph, NOUAK, Alexander, ZHOU, Xuebing, DERAVI, Farzin, VAN VEEN, Michiel Der and SUCHIER, Jean Marc. Towards unattended and privacy protected border control. In : *2007 Biometrics Symposium*. Baltimore, 2007. DOI: 10.1109/BCC.2007.4430536

4. GEIGER, A., LENZ, P. and URTASUN, R. Are we ready for autonomous driving? The KITTI vision benchmark suite. In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Providence, 2012. p. 3354–3361. DOI: 10.1109/CVPR.2012.6248074

5. THRUN, Sebastian, BURGARD, Wolfram and FOX, Dieter. Real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping. In : *Proceedings - IEEE International Conference on Robotics and Automation*. San Francisco, 2000. p. 321–328. DOI: 10.1109/robot.2000.844077

6. GAO, J., GINDY, N. and CHEN, X. An automated GD&T inspection system based on non-contact 3D digitization. *International Journal of Production Research*. 2006. DOI: 10.1080/09638280500219737.

7. SU, Cheng, JIANG, Xin, HUO, Guanying, ZOU, Qiang, ZHENG, Zhiming and FENG, Hsi Yung. Accurate model construction of deformed aero-engine blades for remanufacturing. *International Journal of Advanced Manufacturing Technology*. 2020. DOI: 10.1007/s00170-019-04688-w.

8. HAJEER, M Y, MILLETT, D T, AYOUB, a F and SIEBERT, J P. Applications of

3D imaging in orthodontics: part II. *Journal of orthodontics*. 2004. Vol. 31, no. 1, p. 62–70. DOI: 10.1179/146531204225020472.

9. HARTE, James M., GOLBY, Christopher K., ACOSTA, Johanna, NASH, Edward F., KIRACI, Ercihan, WILLIAMS, Mark A., ARVANITIS, Theodoros N. and NAIDU, Babu. Chest wall motion analysis in healthy volunteers and adults with cystic fibrosis using a novel Kinect-based motion tracking system. *Medical & Biological Engineering & Computing*. 2016. Vol. 54, no. 11, p. 1631–1640. DOI: 10.1007/s11517-015-1433-1.

10. ZHANG, Zhangyou. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2000. Vol. 22, no. 11, p. 1330–1334. DOI: 10.1109/34.888718.

11. HARTLEY, Richard I. In defense of the eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1997. Vol. 19, no. 6, p. 580–593. DOI: 10.1109/34.601246.

12. GORTHI, Sai Siva and RASTOGI, Pramod. Fringe projection techniques: Whither we are? *Optics and Lasers in Engineering*. 2010. Vol. 48, no. 2, p. 133–140. DOI: 10.1016/j.optlaseng.2009.09.001.

13. REN, Mingjun, REN, Jieji, WANG, Xi, GAO, Feng, ZHU, Limin and YAO, Zhenqiang. Multi-scale measurement of high-reflective surfaces by integrating near-field photometric stereo with touch trigger probe. *CIRP Annals*. 2020. DOI: 10.1016/j.cirp.2020.04.016.

14. SCHAFFER, Martin, GROSSE, Marcus and KOWARSCHIK, Richard. High-speed pattern projection for three-dimensional shape measurement using laser speckles. *Applied Optics*. 2010. Vol. 49, no. 18, p. 3622–3629. DOI: 10.1364/AO.49.003622.

15. GROSSE, Marcus, SCHAFFER, Martin, HARENDT, Bastian. and KOWARSCHIK, Richard. Fast data acquisition for three-dimensional shape measurement using fixed-pattern projection and temporal coding. *Optical Engineering*. 2011. Vol. 50, no. 11, p. 100503. DOI: 10.1117/1.3646100.

16. HEIST, Stefan, MANN, Andreas, KÜHMSTEDT, Peter, SCHREIBER, Peter and NOTNI, Gunther. Array projection of aperiodic sinusoidal fringes for high-speed three-dimensional shape measurement. *Optical Engineering*. 2014. Vol. 53, no. 11, p. 112208. DOI: 10.1117/1.OE.53.11.112208.

17. KHAN, Danish, SHIRAZI, Muhammad Ayaz and KIM, Min Young. Single shot laser speckle based 3D acquisition system for medical applications. *Optics and Lasers in Engineering*. 2018. Vol. 105, p. 43–53. DOI: 10.1016/j.optlaseng.2018.01.001.

18. HIRSCHMÜLLER, Heiko, INNOCENT, Peter R. and GARIBALDI, Jon. Real-time correlation-based stereo vision with reduced border errors. *International Journal of Computer Vision*. 2002. Vol. 47, no. 1–3, p. 229–246. DOI: 10.1023/A:1014554110407.

19. HEIST, Stefan, LUTZKE, Peter, SCHMIDT, Ingo, DIETRICH, Patrick, KÜHMSTEDT, Peter, TÜNNERMANN, Andreas and NOTNI, Gunther. High-speed three-dimensional shape measurement using GOBO projection. *Optics and Lasers in Engineering*. 2016. Vol. 87, p. 90–96. DOI: 10.1016/j.optlaseng.2016.02.017.

20. ZHANG, Song. Recent progresses on real-time 3D shape measurement using digital fringe projection techniques. *Optics and Lasers in Engineering*. 2010. Vol. 48, no. 2, p. 149–158. DOI: 10.1016/j.optlaseng.2009.03.008.

21. SCHAFFER, Martin, GROSSE, Marcus, HARENDT, Bastian and KOWARSCHIK, Richard. High-speed three-dimensional shape measurements of objects with laser speckles and acousto-optical deflection. *Optics Letters*. 2011. Vol. 36, no. 16, p. 3097. DOI: 10.1364/OL.36.003097.

22. TAKEDA, Mitsuo and MUTOH, Kazuhiro. Fourier transform profilometry for the automatic measurement of 3-D object shapes. *Applied Optics*. 1983. Vol. 22, no. 24, p. 3977. DOI: 10.1364/AO.22.003977.

23. HUANG, Lei, KEMAO, Qian, PAN, Bing and ASUNDI, Anand Krishna. Comparison of Fourier transform, windowed Fourier transform, and wavelet

transform methods for phase extraction from a single fringe pattern in fringe projection profilometry. *Optics and Lasers in Engineering*. 2010. Vol. 48, no. 2, p. 141–148. DOI: 10.1016/j.optlaseng.2009.04.003.

24. ZUO, Chao, HUANG, Lei, ZHANG, Minliang, CHEN, Qian and ASUNDI, Anand. Temporal phase unwrapping algorithms for fringe projection profilometry: A comparative review. *Optics and Lasers in Engineering*. 2016. Vol. 85, p. 84–103. DOI: 10.1016/j.optlaseng.2016.04.022.

25. ZUO, Chao, FENG, Shijie, HUANG, Lei, TAO, Tianyang, YIN, Wei and CHEN, Qian. Phase shifting algorithms for fringe projection profilometry: A review. *Optics and Lasers in Engineering*. 2018. Vol. 109, p. 23–59. DOI: 10.1016/j.optlaseng.2018.04.019.

26. ZHANG, Song. High-speed 3D shape measurement with structured light methods: A review. *Optics and Lasers in Engineering*. 2018. Vol. 106, p. 119–131. DOI: 10.1016/j.optlaseng.2018.02.017.

27. ZHANG, Qican, SU, Xianyu, XIANG, Liqun and SUN, Xuezhen. 3-D shape measurement based on complementary Gray-code light. *Optics and Lasers in Engineering*. 2012. Vol. 50, no. 4, p. 574–579. DOI: 10.1016/j.optlaseng.2011.06.024.

28. SALVI, Joaquim, PAGÈS, Jordi and BATLLE, Joan. Pattern codification strategies in structured light systems. *Pattern Recognition*. 2004. Vol. 37, no. 4, p. 827–849. DOI: 10.1016/j.patcog.2003.10.002.

29. LUONG, Quang-Tuan and FAUGERAS, Olivier. Self-Calibration of a Moving Camera from Point Correspondences and Fundamental Matrices. *International Journal of Computer Vision*. 1997. Vol. 22, no. 3, p. 261–289. DOI: 10.1023/A:1007982716991.

30. BROWN, D.C. Close- range camera calibration. *Photogramm Eng*. 1971. Vol. 37, no. 8, p. 855–866.

31. HUANG, Peisen S. Novel method for structured light system calibration. *Optical Engineering*. 2006. Vol. 45, no. 8, p. 083601. DOI: 10.1117/1.2336196.

32. MORENO, Daniel and TAUBIN, Gabriel. Simple, accurate, and robust projector-camera calibration. In : *Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*. Zurich, 2012. p. 464–471. DOI: 10.1109/3DIMPVT.2012.77

33. LI, Zhongwei. Accurate calibration method for a structured light system. *Optical Engineering*. 2008. Vol. 47, no. 5, p. 053604. DOI: 10.1117/1.2931517.

34. SITNIK, Robert. Digital fringe projection system for large-volume 360-deg shape measurement. *Optical Engineering*. 2002. Vol. 41, no. 2, p. 443–449. DOI: 10.1117/1.1430422.

35. DECKER, Peter, PAULUS, Dietrich and FELDMANN, Tobias. Dealing with degeneracy in essential matrix estimation. In : *15th IEEE International Conference on Image Processing*. San Diego, 2008. p. 1964–1967. DOI: 10.1109/ICIP.2008.4712167

36. FAUGERAS, Olivier and LUONG, Quang-Tuan. *The Geometry of Multiple Images: The Laws that Govern the Formation of Multiple Images of a Scene and Some of Their Applications*. MIT Press, 2018. DOI: 10.7551/mitpress/3259.001.0001

37. HARTLEY, Richard and ZISSERMAN, Andrew. *Multiple View Geometry in Computer Vision*. Cambridge : Cambridge University Press, 2004. DOI: 10.1017/CBO9780511811685

38. SIELER, M., FISCHER, S., SCHREIBER, P., DANNBERG, P. and BRÄUER, A. Microoptical array projectors for free-form screen applications. *Optics Express*. 2013. Vol. 21, no. 23, p. 28702–28709. DOI: 10.1364/oe.21.028702.

39. WONG, Eugene, HEIST, Stefan, BRÄUER-BURCHARDT, Christian, BABOVSKY, Holger and KOWARSCHIK, Richard. Calibration of an array projector used for high-speed three-dimensional shape measurements using a single camera. *Applied Optics*. 2018. Vol. 57, no. 26, p. 7570–7578. DOI: 10.1364/AO.57.007570.

40. TANG, Suming, ZHANG, Xu, SONG, Zhan, JIANG, Hualie and NIE, Lei.

Three-dimensional surface reconstruction via a robust binary shape-coded structured light method. *Optical Engineering*. 2017. Vol. 56, no. 1, p. 2017. DOI: 10.1117/1.oe.56.1.014102.

41. GENG, Jason. Structured-light 3D surface imaging: a tutorial. *Advances in Optics and Photonics*. 2011. Vol. 3, no. 2, p. 128–160. DOI: 10.1364/aop.3.000128.

42. GU, Feifei, CAO, Huazhao, SONG, Zhan, XIE, Pengju, ZHAO, Juan and LIU, Jing. Dot-coded structured light for accurate and robust 3D reconstruction. *Applied Optics*. 2020. Vol. 59, no. 33, p. 10574–10583. DOI: 10.1364/ao.403624.

43. KANADE, Takeo and OKUTOMI, Masatoshi. A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1994. DOI: 10.1109/34.310690.

44. HERNANDEZ-BELTRAN, Jose Enrique, DIAZ-RAMIREZ, Victor H. and JUAREZ-SALAZAR, Rigoberto. Stereo matching using adaptive windows and correlation filtering. In : *Proc. SPIE 10751, Optics and Photonics for Information Processing XII*, Vol. 1075102. San Diego, 2018. DOI: 10.1117/12.2322043

45. DLP Discovery™ D4100 Development Kit. *Texas Instruments*. 2020. Available from: https://dlinnovations.com/products/dlp-discovery-d4100/

46. ZUO, Chao, TAO, Tianyang, FENG, Shijie, HUANG, Lei, ASUNDI, Anand and CHEN, Qian. Micro Fourier Transform Profilometry (μFTP): 3D shape measurement at 10,000 frames per second. *Optics and Lasers in Engineering*. 2018. Vol. 102, p. 70–91. DOI: 10.1016/j.optlaseng.2017.10.013.

47. QIAN, Jiaming, TAO, Tianyang, FENG, Shijie, CHEN, Qian and ZUO, Chao. Motion-artifact-free dynamic 3D shape measurement with hybrid Fourier-transform phase-shifting profilometry. *Optics Express*. 2019. Vol. 27, no. 3, p. 2713–2731. DOI: 10.1364/oe.27.002713.

48. FUJIGAKI, Motoharu, SAKAGUCHI, Toshimasa and MURATA, Yorinobu.

Development of a compact 3D shape measurement unit using the light-source-stepping method. *Optics and Lasers in Engineering*. 2016. Vol. 85, p. 9–17. DOI: 10.1016/j.optlaseng.2016.04.016.

49. FUJIGAKI, Motoharu, OURA, Yohei, ASAI, Daisuke and MURATA, Yorinobu. High-speed height measurement by a light-source-stepping method using a linear LED array. *Optics Express*. 2013. Vol. 21, no. 20, p. 23169. DOI: 10.1364/OE.21.023169.

50. FUJIGAKI, Motoharu, KUSUNOKI, Yoshiyuki, HARA, Takuya and TANAKA, Hideyuki. 3D shape measurement with an uncalibrated camera using the feature quantity type whole-space tabulation method. *OSA Continuum*. 2020. Vol. 3, no. 8, p. 2277–2288. DOI: 10.1364/osac.401639.

51. ZHANG, Chen, BRAHM, Anika, BREITBARTH, Andreas, ROSENBERGER, Maik and NOTNI, Gunther. Single-shot 3D shape reconstruction using multi-wavelength pattern projection. In : *Dimensional Optical Metrology and Inspection for Practical Applications VIII*. Baltimore, 2019. p. 109910H. DOI: 10.1117/12.2518420

52. STARK, Andreas, WONG, Eugene, BABOVSKY, Holger and KOWARSCHIK, Richard. Subjective speckle suppression for 3D measurement using one-dimensional numerical filtering. *Applied Optics*. 2019. Vol. 58, no. 34, p. 9473–9483 DOI: 10.1364/ao.58.009473.

53. LIU, Yihang, ZHANG, Qican, LIU, Yuankun, YU, Xin, HOU, Yanli and CHEN, Wenjing. High-speed 3D shape measurement using a rotary mechanical projector. *Optics Express*. 2021. Vol. 29, no. 5, p. 7885–7903. DOI: 10.1364/oe.419206.

54. DU, Hua and WANG, Zhaoyang. Three-dimensional shape measurement with an arbitrarily arranged fringe projection profilometry system. *Optics Letters*. 2007. Vol. 32, no. 16, p. 2439–2440. DOI: 10.1364/ol.32.002438.

55. ZHANG, Xiaoling, LIN, Yuchi, ZHAO, Meirong, NIU, Xiaobing and HUANG, Yinguo. Calibration of a fringe projection profilometry system using virtual

phase calibrating model planes. *Journal of Optics A: Pure and Applied Optics*. 2005. Vol. 7, no. 4, p. 192. DOI: 10.1088/1464-4258/7/4/007.

56. ZHOU, W. S. and SU, X. Y. A direct mapping algorithm for phase-measuring profilometry. *Journal of Modern Optics*. 1994. Vol. 41, no. 1, p. 89–94. DOI: 10.1080/09500349414550101.

57. TAVARES, Paulo J. and VAZ, Mário A. Linear calibration procedure for the phase-to-height relationship in phase measurement profilometry. *Optics Communications*. 2007. Vol. 274, no. 2, p. 307–314. DOI: 10.1016/j.optcom.2007.02.038.

58. MORIMOTO, Yoshiharu, MASAYA, Akihiro, FUJIGAKI, Motoharu and ASAI, Daisuke. Shape Measurement by Phase-Stepping Method Using Multi-Line LEDs. *Applied Measurement Systems*. 2012. Vol. 102, p. 70–91. DOI: 10.5772/36127.

59. GUO, Wenbo, WU, Zhoujie, XU, Renchao, ZHANG, Qican and FUJIGAKI, Motoharu. A fast reconstruction method for three-dimensional shape measurement using dual-frequency grating projection and phase-to-height lookup table. *Optics and Laser Technology*. 2019. Vol. 112, p. 269–2777. DOI: 10.1016/j.optlastec.2018.11.009.

60. FISCHLER, Martin A. and BOLLES, Robert C. Random sample consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*. 1981. Vol. 24, no. 6, p. 381–395. DOI: 10.1145/358669.358692.

61. SHAKARJI, C.M. Least-squares fitting algorithms of the NIST algorithm testing system. *Journal of Research of the National Institute of Standards and Technology*. November 1998. Vol. 103, no. 6, p. 633. DOI: 10.6028/jres.103.043.

62. STARK, Andreas, WONG, Eugene, WEIGEL, Daniel, BABOVSKY, Holger and KOWARSCHIK, Richard. Repeatable speckle projector for single-camera three-dimensional measurement. *Optical Engineering*. 2018. Vol. 57, no. 12,

p. 120501. DOI: 10.1117/1.oe.57.12.120501.

63. ZHONG, Tianting, YU, Zhipeng, LI, Huanhao, LI, Zihao, LI, Haohong and LAI, Puxiang. Active wavefront shaping for controlling and improving multimode fiber sensor. *Journal of Innovative Optical Health Sciences*. 2019. Vol. 12, no. 4, p. 1942007. DOI: 10.1142/S1793545819420070.

64. STARK, Andreas, WONG, Eugene, WEIGEL, Daniel, BABOVSKY, Holger, SCHOTT, Thomas and KOWARSCHIK, Richard. Subjective speckle suppression in laser-based stereo photogrammetry. *Optical Engineering*. 2016. Vol. 55, no. 12, p. 121713. DOI: 10.1117/1.oe.55.12.121713.

65. CHEN, Xiaobo, XI, Juntong, JIN, Ye and SUN, Jin. Accurate calibration for a camera–projector measurement system based on structured light projection. *Optics and Lasers in Engineering*. 2009. Vol. 47, no. 3–4, p. 310–319. DOI: 10.1016/j.optlaseng.2007.12.001.

66. LEGARDA-SA´ENZ, Ricardo. Accurate procedure for the calibration of a structured light system. *Optical Engineering*. 2004. Vol. 43, no. 2, p. 464. DOI: 10.1117/1.1635373.

67. HUANG, Bingyao, TANG, Ying, OZDEMIR, Samed and LING, Haibin. A Fast and Flexible Projector-Camera Calibration System. *IEEE Transactions on Automation Science and Engineering*. 2020. p. 1–15. DOI: 10.1109/tase.2020.2994223.

68. HUANG, Lei, CHUA, Patrick S.K. and ASUNDI, A. Least-squares calibration method for fringe projection profilometry considering camera lens distortion. *Applied Optics*. 2010. Vol. 49, no. 9, p. 1539–1548. DOI: 10.1364/AO.49.001539.

69. HUANG, Tingting, FU, Xiao, LI, Xiuming, ZHANG, Cong, DUAN, Fajie, JIANG, Jiajia and TAN, Haishu. Phase-to-depth calibration in fringe projection profilometry based on blockwise surface fitting. *Optics and Laser Technology*. 2020. Vol. 122, p.105844. DOI: 10.1016/j.optlastec.2019.105844.

70. WONG, Eugene, HEIST, Stefan, BRÄUER-BURCHARDT, Christian, STARK,

Andreas, BABOVSKY, Holger and KOWARSCHIK, Richard. Optimization-based extrinsic calibration of a three-dimensional sensor composed of an array projector and a single camera. *Optical Engineering*. 2019. Vol. 58, no. 10, p. 104109. DOI: 10.1117/1.oe.58.10.104109.

71. HARRIS, J.W and STÖCKER, H. *Handbook of Mathematics and Computational Science*. New York : Springer-Verlag New York, 1998.

72. NELDER, J. A. and MEAD, R. A Simplex Method for Function Minimization. *Mathematical and Computer Modelling*. 1965. Vol. 7, no. 4, p. 308–313. DOI: 10.1093/comjnl/7.4.308.

73. SHI, Yuhui and EBERHART, Russell C. Parameter selection in particle swarm optimization. In : *Evolutionary Programming VII International Conference on Evolutionary Programming*. San Diego, 1998. p. 591–600. DOI: 10.1007/bfb0040810

74. NICKABADI, Ahmad, EBADZADEH, Mohammad Mehdi and SAFABAKHSH, Reza. A novel particle swarm optimization algorithm with adaptive inertia weight. *Applied Soft Computing Journal*. 2011. Vol. 11, no. 4, p. 3658–3670. DOI: 10.1016/j.asoc.2011.01.037.

75. WIEGMANN, A., WAGNER, H. and KOWARSCHIK, R. Human face measurement by projecting bandlimited random patterns. *Optics Express*. 2006. Vol. 14, no. 17, p. 7692. DOI: 10.1364/OE.14.007692.

76. GROSSE, Marcus, SCHAFFER, Martin, HARENDT, Bastian and KOWARSCHIK, Richard. Camera calibration using time-coded planar patterns. *Optical Engineering*. 2012. Vol. 51, no. 8, p. 083604. DOI: 10.1117/1.OE.51.8.083604.

77. SEITZ, Steven M. and DYER, Charles R. View morphing. In : *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1996*. Anaheim, 1996. p. 21–30. DOI: 10.1145/237170.237196

78. JI, Dinghuang, KWON, Junghyun, MCFARLAND, Max and SAVARESE, Silvio.

Deep view morphing. In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, 2017. p. 2155–2163. DOI: 10.1109/CVPR.2017.750

79. MILDENHALL, Ben, PRATUL P., Srinivasan, MATTHEW, Tancik, JONATHAN T., Barron, RAVI, Ramamoorthi and REN, Ng. Representing Scenes as Neural Radiance Fields for View Synthesis. *arXiv preprint*. 2020. Vol. 2003.08934.

80. SUN, Shao Hua, HUH, Minyoung, LIAO, Yuan Hong, ZHANG, Ning and LIM, Joseph J. Multi-view to Novel View: Synthesizing Novel Views With Self-learned Confidence. In : *Proceedings of the European Conference on Computer Vision*. Munich, 2018. p. 155–171. DOI: 10.1007/978-3-030-01219-9_10

81. PARK, Eunbyung, YANG, Jimei, YUMER, Ersin, CEYLAN, Duygu and BERG, Alexander C. Transformation-grounded image generation network for novel 3D view synthesis. In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, 2017. p. 3500–3509. DOI: 10.1109/CVPR.2017.82

82. WONG, Eugene, HEIST, Stefan, BRÄUER-BURCHARDT, Christian, STARK, Andreas, BABOVSKY, Holger and KOWARSCHIK, Richard. View-synthesized ‘re-calibration’ of an array projector for 3D measurement from an arbitrary monocular view. *Optics and Lasers in Engineering*. 2021. Vol. 141, p. 106559. DOI: 10.1016/j.optlaseng.2021.106559.

83. CloudCompare, Version 2.9.1, GPL Software. . 2017. Available from: https://www.cloudcompare.org/

84. GROSSE, Marcus, SCHAFFER, Martin, HARENDT, Bastian and KOWARSCHIK, Richard. Influence of the structured illumination frequency content on the correspondence assignment precision in stereophotogrammetry. In : *Optical Measurement Systems for Industrial Inspection VIII*. Munich, 2013. DOI: 10.1117/12.2020448

85. HEIST, Stefan, KÜHMSTEDT, Peter, TÜNNERMANN, Andreas and NOTNI,

Gunther. Theoretical considerations on aperiodic sinusoidal fringes in comparison to phase-shifted sinusoidal fringes for high-speed three-dimensional shape measurement. *Applied Optics*. 2015. Vol. 54, no. 35, p. 10541. DOI: 10.1364/AO.54.010541.

86. WONG, Eugene, GROSSE, Marcus, BABOVSKY, Holger and KOWARSCHIK, Richard. 3D Reconstruction through the Ray Calibration of an Optically Distorted Stereo-photogrammetric System. In : *116th DGaO Proceedings*. Brno, 2015.

87. MOROZ, Leonid V., WALCZYK, Cezary J., HRYNCHYSHYN, Andriy, HOLIMATH, Vijay and CIEŚLIŃSKI, Jan L. Fast calculation of inverse square root with the use of magic constant – analytical approach. *Applied Mathematics and Computation*. 2018. Vol. 316, p. 245–255. DOI: 10.1016/j.amc.2017.08.025.

88. DIETRICH, Patrick, HEIST, Stefan, LANDMANN, Martin, KÜHMSTEDT, Peter and NOTNI, Gunther. BICOS-an algorithm for fast real-time Correspondence Search for statistical pattern projection-based active stereo sensors. *Applied Sciences*. 2019. Vol. 9, no. 16, p. 3330. DOI: 10.3390/app9163330.

89. WONG, Eugene, STARK, Andreas Walter, BABOVSKY, Holger and KOWARSCHIK, Richard. High-density three-dimensional measurements through multilayer perceptron calibration and statistical band-limited patterns. *Optical Engineering*. 2018. Vol. 57, no. 07, p. 1. DOI: 10.1117/1.oe.57.7.070501.

90. VAN DER JEUGHT, Sam and DIRCKX, Joris J. J. Deep neural networks for single shot structured light profilometry. *Optics Express*. 2019. Vol. 27, no. 12, p. 17091–17101. DOI: 10.1364/oe.27.017091.

91. NGUYEN, Hieu, WANG, Yuzeng and WANG, Zhaoyang. Single-shot 3d shape reconstruction using structured light and deep convolutional neural networks. *Sensors*. 2020. Vol. 20, no. 13, p. 3718. DOI: 10.3390/s20133718.

92. HEIST, Stefan, KÜHMSTEDT, Peter, TÜNNERMANN, Andreas and NOTNI,

Gunther. BRDF-dependent accuracy of arrayprojection-based 3D sensors. *Applied optics*. 2017. Vol. 56, no. 8, p. 2162–2170. DOI: 10.1364/AO.56.002162.

93. ASHIKHMIN, Michael and SHIRLEY, Peter. An Anisotropic Phong BRDF Model. *Journal of Graphics Tools*. 2000. DOI: 10.1080/10867651.2000.10487522.

94. HARRIS, Chris and STEPHENS, Mike. A combined edge and corner detector. In : *Proc 4th Alvey Vision Conference*. Manchester, 1988. p. 10–5244.

Acknowledgments

I must express great gratitude and appreciation to my colleagues from the Institute of Applied Optics and Biophysics. First and foremost, to Prof. Dr. Richard Kowarschik for his guidance, scientific insight, and wisdom in his role as supervisor. Much appreciation also must be given to Andreas Stark and Dr. Holger Babovsky for fruitful discussions and their unbiased constructive scientific criticisms. Gratefulness must also be expressed to both Daniel Weigel and Andreas Stark for our collaboration on the design of the temporally repeatable speckle projector, especially for Daniel Weigel's hard work on its fabrication. To all my colleagues at the Institute of Applied Optics and Biophysics, I am also incredibly grateful for their guidance in traversing through the intricacies of the German academic system, for which I was not so adept at negotiating through.

Thanks must also go to my colleagues at the Fraunhofer Institute of Optics and Fine Mechanics. I must express my gratefulness to Dr. Stefan Heist for our fruitful collaboration and granting the use of the MAAP for experiments. I am also grateful to Dr. Christrian Bräuer-Burchardt for spearheading this joint project between the two institutes to the project alliance 3Dsensation. I would also like to express thanks to the project alliance 3Dsensation and the Bundesministerium für Bildung und Forschung (BMBF) under grant numbers 03ZZ0425H, 03ZZ0469, and 03ZZ0409 for this project's financing. I would like to acknowledge the 3Dsensation graduate school in creating a stimulating environment in promoting interdisciplinary research towards the advancement of achieving Industry 4.0. I also express thanks to the Abbe School of Photonics and the city of Jena for being at the forefront of promoting international students such as myself to experience such an exciting environment for research and study. It was fun and stimulating to feel as part of a greater research community.

Apart from my much appreciated work colleagues, I would like to thank my family and friends for their encouragement outside of work. Their support, especially during the time where this study was written during the COVID-19 pandemic, was priceless.

List of Publications

Peer-reviewed publications contributed as part of this work:

1. **WONG, Eugene***, HEIST, Stefan, BRÄUER-BURCHARDT, Christian, BABOVSKY, Holger and KOWARSCHIK, Richard. Calibration of an array projector used for high-speed three-dimensional shape measurements using a single camera. *Applied Optics*. 2018. Vol. 57, no. 26, p. 7570–7578. DOI: 10.1364/AO.57.007570.

2. STARK, Andreas*, **WONG, Eugene***, WEIGEL, Daniel, BABOVSKY, Holger and KOWARSCHIK, Richard. Repeatable speckle projector for single-camera three-dimensional measurement. *Optical Engineering*. 2018. Vol. 57, no. 12, p. 120501. DOI: 10.1117/1.oe.57.12.120501.

3. **WONG, Eugene***, HEIST, Stefan, BRÄUER-BURCHARDT, Christian, STARK, Andreas, BABOVSKY, Holger and KOWARSCHIK, Richard. Optimization-based extrinsic calibration of a three-dimensional sensor composed of an array projector and a single camera. *Optical Engineering*. 2019. Vol. 58, no. 10, p. 104109. DOI: 10.1117/1.oe.58.10.104109.

4. **WONG, Eugene***, HEIST, Stefan, BRÄUER-BURCHARDT, Christian, STARK, Andreas, BABOVSKY, Holger and KOWARSCHIK, Richard. View-synthesized ‘re-calibration’ of an array projector for 3D measurement from an arbitrary monocular view. *Optics and Lasers in Engineering*. 2021. Vol. 141, p. 106559. DOI: 10.1016/j.optlaseng.2021.106559.

Peer-reviewed publications contributed not part of this work:

1. **WONG, Eugene***, STARK, Andreas, BABOVSKY, Holger and KOWARSCHIK, Richard. High-density three-dimensional measurements through multilayer perceptron calibration and statistical band-limited patterns. *Optical Engineering*. 2018. Vol. 57, no. 07, p. 1. DOI: 10.1117/1.oe.57.7.070501.

Patent applications contributed as part of this work:

1. Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V, Friedrich-Schiller-Universität Jena. *Verfahren und Vorrichtung zum berührungslosen Vermessen dreidimensionaler Oberflächenkonturen*. Inventors: BRÄUER-BURCHARDT, Christian , HEIST, Stefan, **WONG, Eugene**, and KOWARSCHIK, Richard. *Verfahren zur strukturierten Beleuchtung*. Registration date: 20.11.2017. DE, Offenlegungsschrift DE 10 2017 220 720 A1. Publication date: 23.05.2019.

2. Friedrich-Schiller-Universität Jena, *Verfahren zur strukturierten Beleuchtung*. Inventors: WEIGEL, Daniel, **WONG, Eugene**, STARK, Andreas, BABOVSKY, Holger and KOWARSCHIK, Richard. Registration date: 22.05.2018. DE, Offenlegungsschrift DE 10 2018 004 078 A1. Publication date: 28.11.2019.

Conference proceedings contributed as part of this work:

1. **WONG, Eugene***, HEIST, Stefan, BRÄUER-BURCHARDT, Christian, STARK, Andreas, BABOVSKY, Holger and KOWARSCHIK, Richard. Single camera photogrammetry utilizing non-DMD/LCoS based projection systems. In : *119th DGaO Proceedings*. Aalen, 2018.

2. **WONG, Eugene***, HEIST, Stefan, BRÄUER-BURCHARDT, Christian, STARK, Andreas, BABOVSKY, Holger and KOWARSCHIK, Richard. New View Post-Calibration of an Array Projector. In : *120th DGaO Proceedings*. Darmstadt, 2019.

3. **WONG, Eugene***, HEIST, Stefan, BRÄUER-BURCHARDT, Christian, STARK, Andreas, BABOVSKY, Holger and KOWARSCHIK, Richard. Extrinsic calibration of a 3D sensor based on an array projector and a single camera. In : *Optical Measurement Systems for Industrial Inspection XI*. Munich, 2019. Vol. 11056, p. 1105617. DOI: 10.1117/1.OE.58.10.104109

Ehrenwörtliche Erklärung

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit selbständig, ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen.

Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem In-halt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die geltende Promotionsordnung der Physikalisch-Astronomischen Fakultät ist mir bekannt.

Ich versichere ehrenwörtlich, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Jena, 02/12/2021 Eugene Wong