

Projector Calibration via Overlapping Point Cloud Distance Minimization

by

Pranav Kumar Ayee Goundar Venkatesan

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2025

© Pranav Kumar Ayee Goundar Venkatesan 2025

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis includes a paper for which I am the first author. I was actively involved in every stage of its development, including formulating the research problem, designing the solution, collecting data, conducting experiments, analyzing results, and writing the paper.

[68] Ayee Goundar Venkatesan, P. K., Moradi, S., Fieguth, P., & Lamm, M. (2025). *Evaluating the Impact of Stereo Overlap on Calibration Parameters in Camera and Projector Systems*. *Journal of Computational Vision and Imaging Systems*, **10**(1), 30–33. <https://doi.org/10.15353/jcvis.v10i1.10005>

Excerpts from this paper appear in Chapter 3 Problem Formulation and Chapter 4 Stereo Overlap Between Optical Devices.

Contributor	Concept Design	Writing and Editing
Ayee Goundar Venkatesan,P (Candidate)	75%	80%
Moradi, S.	10%	8%
Fieguth, P.	10%	10%
Lamm, M.	5%	2%

Abstract

Projection mapping is a technique that transforms any 3D surface into an interactive display by projecting visuals that conform to the surface’s shape. Camera and projector calibration is a fundamental prerequisite for accurate spatial measurement and perception in projection mapping. The calibration of projectors is necessary to ensure that they correctly map the surface they are projecting onto. An accurately calibrated projector will have the image projected perfectly in line with the intended surface, without any misalignment. Such alignment is important for a number of applications, one of which is projection mapping, in which multiple projectors are used to create immersive visual displays on complex surfaces such as buildings, stages, or statues.

Since projector calibration involves highly nonlinear relationships between the projector’s parameters, a nonlinear optimization is required. Typically such optimizations include reprojection error as the objective functions, which is also one of the most commonly used objective function to calibrate projectors. However, in scenarios where there is a lack of ground truth reprojection error fails to accurately align the multiple overlapping projector point clouds, resulting in visible gaps between them instead of forming a continuous surface.

To address this issue, this thesis proposes a multi-step optimization with a novel global objective function. To analyze its robustness, the proposed optimization is tested both on simulated and on multiple real world configurations. Experimental results show that the proposed approach achieves higher calibration accuracy compared to existing methods, while also maintaining low runtime.

The proposed multi-step optimization process parameterizes the calibration problem as a function of the degree of stereo overlap to improve accuracy. The stereo overlap plays a key role in projection mapping, influencing calibration accuracy and system cost. Understanding the relationship between stereo overlap and calibration error allows for reducing overlap while maintaining acceptable accuracy, thus reducing the number of cameras needed and cutting costs.

In summary, this thesis introduces a novel global objective function that minimizes the distance between overlapping projector point clouds, rather than relying solely on reprojection error. It also provides insight into the required overlap between devices, including both cameras and projectors, helping to achieve higher accuracy while efficiently covering an entire projection surface that may not be uniform.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to Prof. Paul Fieguth for his patience and unwavering support throughout my master's journey. As a beginner in research, the initial stages were overwhelming, but his guidance and encouragement shaped my thinking and helped me grow as a researcher. I remain deeply grateful for his mentorship.

I would also like to thank Prof. John Zelek, Prof. Olga Veksler, and Dr. Saed Moradi for sharing their expertise and taking the time to review my thesis.

A special note of appreciation to Dr. Saed Moradi for his continuous encouragement and steadfast support throughout the research process.

Finally, I am thankful to Mark Lamm and the engineering team at Christie Digital Systems. As my second mentor, Mark generously shared his knowledge and helped me gain a strong understanding of projector systems. His support in facilitating various experiments and guiding me through practical challenges was truly invaluable.

Dedication

I would like to dedicate this work to my grandmother Pavayee and my mother Devi, the two most important women in my life, whose unwavering love helped me stay focused throughout this journey. I also extend this dedication to my entire family, whose constant love and support made this research possible.

Finally, I am grateful to all my friends at the University of Waterloo, who became my second family in a new country and continuously motivated me throughout the completion of my thesis.

Table of Contents

Author’s Declaration	ii
Statement of Contributions	iii
Abstract	iv
Acknowledgements	v
Dedication	vi
List of Figures	x
List of Tables	xi
1 Introduction	1
1.1 Motivation	2
1.2 Objectives	6
1.3 Thesis Organization	7
2 Background	9
2.1 Camera Model and its Mathematical Modeling	9
2.1.1 Pinhole Camera model	9
2.1.2 Lens Distortion	12

2.2	Two-View Geometry	13
2.2.1	Epipolar Geometry and Fundamental Matrix	14
2.2.2	Essential Matrix	15
2.2.3	Triangulation	17
2.2.4	Reprojection Error	17
2.3	Camera Calibration	18
2.3.1	Manual Calibration	18
2.3.2	Self-Calibration	20
2.4	Projector Calibration	21
2.4.1	Adapting the Camera model for Projector Calibration	21
2.4.2	Structured Light	22
2.4.3	Calibration Techniques	23
3	Problem Formulation	27
3.1	Stereo Overlaps and their Significance	28
3.2	Calibration of Projectors	28
3.2.1	Need for Auto-Calibration of Projectors	28
3.2.2	Challenges	29
3.3	Proposed Strategy	30
3.3.1	Overlapping Distance Error	30
3.3.2	Cyclical Effect	31
3.3.3	Multi-Step Optimization Strategy	32
4	Stereo Overlap Between Optical Devices	33
4.1	Synthetic Data	34
4.2	Methodology and Experimental Setup	35
4.2.1	Impact of the Degree of Stereo Overlap on Camera Calibration Accuracy	35

4.2.2	Impact of Camera Overlap on Projector Calibration Accuracy . . .	36
4.2.3	Calibration Accuracy in Sequentially Overlapping Projectors	38
4.3	Experimental Procedure	38
4.4	Results	39
5	Multi-Step Calibration of Projectors	45
5.1	Methodology	45
5.1.1	Point Correspondence Collection	46
5.1.2	Camera Calibration	46
5.1.3	Initial Estimate of Projector Parameters with Optimization 1 . . .	47
5.1.4	Optimization 2	49
5.2	Experiment Data and Setup	52
5.3	Experimentation Results	54
5.3.1	Synthetic data	54
5.3.2	Real data	55
5.3.3	Limited Points	56
6	Conclusion	60
6.1	Future Directions	61
	References	63

List of Figures

1.1	Sample display of Projection Mapping	3
1.2	Basic setup with a dome as the target surface.	4
1.3	Point cloud with visible gaps	6
2.1	Pinhole Camera Model	10
2.2	Two-view Epipolar Geometry	16
2.3	Binary Coded Structured Light	24
3.1	Illustration of cyclical error across overlapping projectors	31
4.1	Illustration of field of view overlap between two devices.	34
4.2	Synthetic Data Setup for Stereo Overlap Analysis	37
4.3	Analysis of the Effect of Stereo Overlap on Camera Calibration Accuracy .	40
4.4	Analysis of the Effect of Stereo Overlap on Projector Calibration Accuracy	41
4.5	Impact of Calibration Accuracy Across Sequential Projectors with Varying Overlap	42
4.6	Impact on Calibration Accuracy Across Sequential Projectors with Fixed Overlap	43
5.1	Point cloud reconstruction using only reprojection error	50
5.2	Synthetic and Real Point Clouds Used in Optimization	53
5.3	Point Cloud Reconstruction Before and After Optimization	56
5.4	Overlapping distance error vs Point Count in Optimizations 1 and 2	57
5.5	Runtime vs Point Count in Optimizations 1 and 2	58

List of Tables

5.1	Results of Multi-step Optimization on Synthetic Data	54
5.2	Results of Multi-step Optimization on Real Data	55

Chapter 1

Introduction

Traditional flat-screen projection involves projecting visuals onto flat, white surfaces such as walls or screens, where the image alignment is straightforward and unaffected by surface geometry. In contrast, projection mapping is a technique that transforms any 3D surface into an interactive display by projecting video and animations onto it [23, 29]. The target 3D surface can be any non-flat or non-white surface. Unlike traditional flat-screen projections, this technique adapts the visuals to the contours of 3D surfaces, resulting in immersive and dynamic experiences. This technology merges art and science to enhance storytelling and spatial perception. Projection mapping is widely used across various domains, including 3D reconstruction and artistic displays for educational, entertainment, and cultural applications [6, 7, 9, 29]. An example is shown in Figure 1.1, where projection mapping is used on a building to illustrate the origin and evolution of life.

In a typical projection mapping setup, three fundamental components are required: a projector, projection animation, and a target surface [6]. The accuracy of this process depends on the geometric information of the projector and its spatial relationship to the target surface. To determine this geometric relationship, the projector must be calibrated by estimating its internal characteristics and position in space relative to the scene. This step is critical to ensure that the projected image aligns accurately with the target 3D surface. High-resolution cameras are used to calibrate the projectors by measuring their spatial relationship to the projection surface [44, 70]. Since the cameras are spatially offset from the projectors, some form of calibration is performed to estimate the relative transformation between the camera and projector coordinate frames.

For a planar target surface, performing projection mapping is relatively straightforward, as there is typically no rotation involved and the process primarily depends on the internal

parameters of the projectors [29]. While it may seem feasible to first calibrate projectors using a flat white surface and then move them to a different location, this approach is generally not practical. The actual 3D surface often requires repositioning the projectors at different orientations and distances, which can alter both intrinsic and extrinsic parameters. As a result, the projectors must be directly calibrated on the 3D surface itself to ensure accurate alignment. This becomes particularly challenging for complex surfaces (such as a building, illustrated in Figure 1.1), where identifying both the pose of each projector and its internal parameters is significantly more difficult.

Due to its challenging nature, conventional geometric calibration of projectors is often infeasible and prone to significant error [45]. As a result, non-linear optimization methods [32, 46] have been widely adopted to achieve accurate projector calibration. Among these methods, reprojection error is the most commonly used objective function. However, in the absence of ground truth, reprojection error becomes ineffective when dealing with large projection surfaces and a minimal number of cameras, as discussed in detail in Section 1.1.

In this thesis, we introduce a novel objective function to improve projector calibration. Using this objective function, we develop a multi-step optimization algorithm that is rigorously tested on both synthetic and real-world data to evaluate its robustness. All analyses in this work are performed on a setup consisting of at least one projector and a minimum of two cameras (stereo pair), ensuring sufficient overlap for reliable calibration. A simplified version of this setup is illustrated in Figure 1.2. Further experiments are conducted in this thesis to better understand the relationship between overlap and calibration accuracy.

1.1 Motivation

Projection mapping is a rapidly growing industry, projected to reach a market value of \$4.43 billion USD by 2025, with a compound annual growth rate of 18.3% [52]. Given the high economic stakes and intense competition, accurate projector calibration is essential for maintaining a competitive position in the market. Higher calibration accuracy leads to improved visual quality, which enhances the consumer experience. In addition, greater accuracy reduces the need for a large number of cameras, thereby lowering the significant capital costs associated with projection mapping setups.

Over the past two decades, extensive research has been conducted to improve calibration accuracy [3, 34, 44, 51]. Although many techniques have achieved accurate projector calibration, they often require manual intervention. In projector calibration, manual calibration typically involves using printed calibration patterns such as checkerboards, physical



Figure 1.1: Projection mapping displayed on the Archbishop's Palace in Prague, Czech Republic, depicting the evolution of life. Image obtained from Christie Digital Systems.

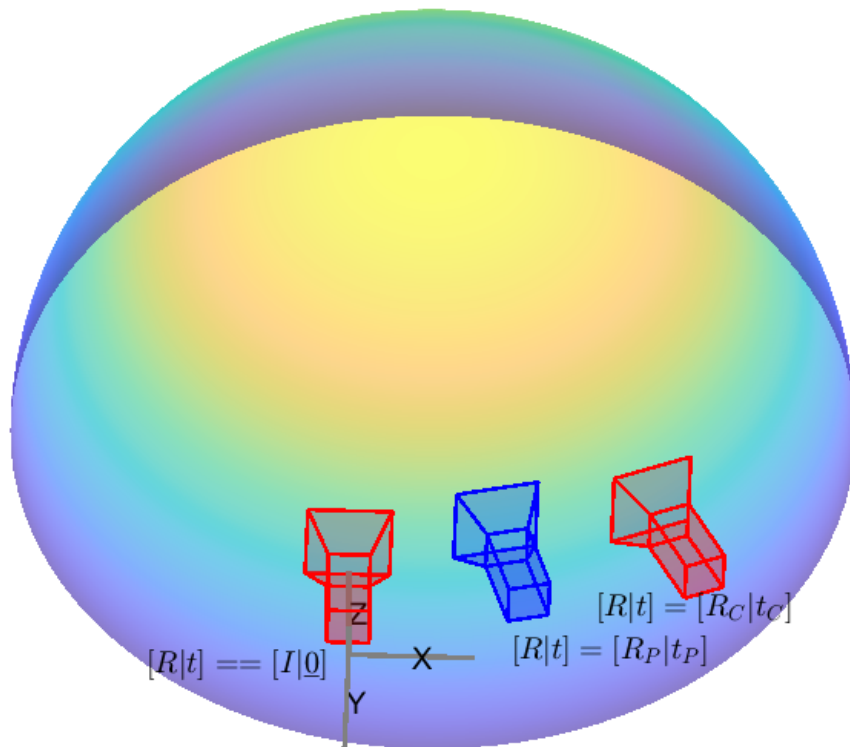


Figure 1.2: A simple setup with two cameras and one projector projecting onto a 3D dome surface, with red representing cameras and blue representing the projector.

markers, or embedded sensors placed on the target surface to enhance accuracy. These sensors are typically light-emitting or reflective elements that assist optical devices in detecting and localizing calibration targets more precisely. However, these approaches present several limitations. For example, accuracy decreases when projecting onto distant surfaces because the physical sensors may appear too small to be detected effectively. Furthermore, manual alignment with calibration targets is time-consuming and requires human involvement at each step.

These limitations highlight the need for self-calibration methods in camera and projector systems that can eliminate manual procedures while preserving or improving calibration accuracy. Researchers have proposed several algorithms to enable effective self-calibration of projectors [62, 69]. These methods introduced robust techniques for calibrating projectors without requiring manual intervention or physical patterns.

Although effective, these approaches still rely on sufficient overlap between the projector and the common stereo field of view of both cameras. For example, Tehrani et al. [62] assumed that at least 20% of the projected area by each projector must be consistently observed by a minimum of two cameras. When projection mapping is applied to a large surface, such as one involving a dozen projectors, this assumption leads to two possible outcomes:

- Increase in the number of cameras: Since the projector calibration algorithm [62] requires that at least 20% of the projected area be seen by two cameras, more cameras must be installed in the physical setup. This substantially increases capital costs, which may be unacceptable. However, reducing the number of cameras compromises the algorithm’s accuracy, resulting in a poor visual display.
- High field of view coverage between cameras: To reduce the number of cameras while maintaining coverage, each projector must have at least 20% overlap with the stereo field of view of the camera pair, as required by the projector calibration algorithm. On large target surfaces, this often necessitates placing the cameras farther from the surface to expand their field of view. At greater distances, however, each pixel in the camera image corresponds to a larger area on the projection surface. For instance, Martynov et al. [40] reported that doubling the camera-to-surface distance (from 60 cm to 120 cm) led to a 25% increase in reprojection error, even when the viewing angle remained constant. This degradation occurs because projected features occupy fewer pixels in the image, thereby reducing detection precision.

From the above observations, it can be concluded that current self-calibration algorithms for projectors create a trade-off between accuracy and cost when aiming to produce a

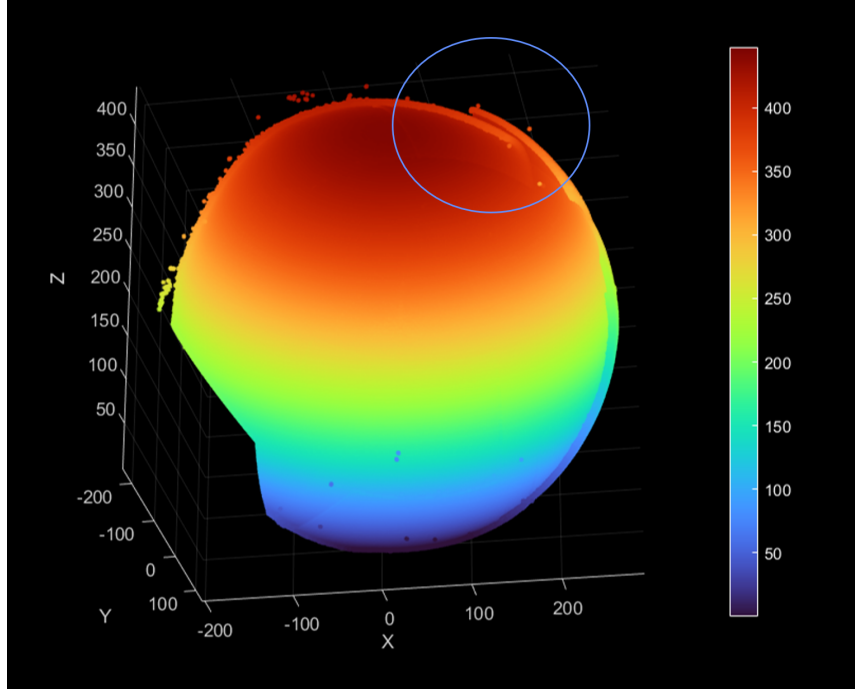


Figure 1.3: Point cloud of projectors calibrated using existing methods with less than 20% stereo overlap between each projector and the stereo camera. The visible gaps, particularly in the top-right region, highlight misalignments in the reconstructed structure.

visually appealing display. Reducing the overlaps often leads to visible gaps between projector point clouds in the reconstructed structure, as shown in the top-right corner of Figure 1.3. Therefore, there is a clear need for a robust algorithm that enables accurate projector calibration even when projecting onto distant surfaces, while also minimizing the number of cameras required. Achieving this would significantly reduce the overall capital cost of projection mapping setups without compromising accuracy.

1.2 Objectives

This research has three primary objectives:

1. To analyze various stereo overlaps between cameras and projectors in order to gain insights into how to optimally position the projections, ensuring full surface coverage while maintaining accuracy.

2. To develop a novel objective function that accurately calibrates projectors with minimal cameras. The minimal setup consists of two cameras, sufficient to ensure stereo coverage across the target surface.
3. To devise a quick optimization strategy that completes within 10 minutes, based on the proposed objective function and conduct rigorous testing to validate its effectiveness.

By addressing these objectives, this thesis aims to address the following challenges:

- The need for physical calibration targets.
- The high capital cost associated with deploying numerous cameras.
- The visible gaps between overlapping projector regions in the reconstructed structure, enabling a seamless and continuous surface.

1.3 Thesis Organization

This thesis is organised as follows:

- Chapter 2: Background. This chapter introduces the key concepts and important equations used throughout the thesis. It also provides a detailed overview of existing calibration techniques for projectors and cameras.
- Chapter 3: Problem Formulation. This chapter outlines the core objectives of the thesis. It expands on the motivation introduced in Chapter 1, discusses the challenges associated with current methodologies, and highlights the necessity of developing a new approach.
- Chapter 4: Stereo Overlap Between Optical Devices. This chapter presents experiments and a detailed analysis of the different types of stereo overlap that can occur between optical devices.
- Chapter 5: Multi-Step Calibration of Projectors. In this chapter, the proposed method is applied to both synthetic and real-world data. The chapter explains the methodology in detail, presents the optimization process, and analyzes the resulting performance.

- Chapter 6: Conclusion. The final chapter summarizes the key findings and contributions of the thesis and concludes by suggesting potential directions for future research.

Chapter 2

Background

2.1 Camera Model and its Mathematical Modeling

Cameras are devices that project a three-dimensional (3D) scene onto a two-dimensional (2D) image plane. The relationship between this 3D scene to 2D image is modelled by the camera model [25, 59]. Understanding this model is critical in computer vision tasks such as 3D reconstruction, pose estimation, and augmented reality. The most basic and widely used camera model is the pinhole camera model [12].

2.1.1 Pinhole Camera model

Pinhole cameras are closed boxes with a tiny hole which allows the passage of light through it to form an image. Pinhole camera models represent the projection from a 3D scene to a 2D image, as captured by pinhole cameras. It serves as the foundational model in computer vision. Visual representation of pinhole camera model is shown in Figure 2.1.

Mathematically, the projection of a 3D point onto a 2D image plane in the pinhole camera model is represented using the homogeneous coordinates [16, 25, 59]. Homogeneous coordinates are an extension of Cartesian coordinates that allow representation of points at infinity, such as the apparent intersection point of parallel lines, and enable the use of linear transformations in projective geometry.

Using homogeneous coordinates, the relationship between a 3D point in world coordi-

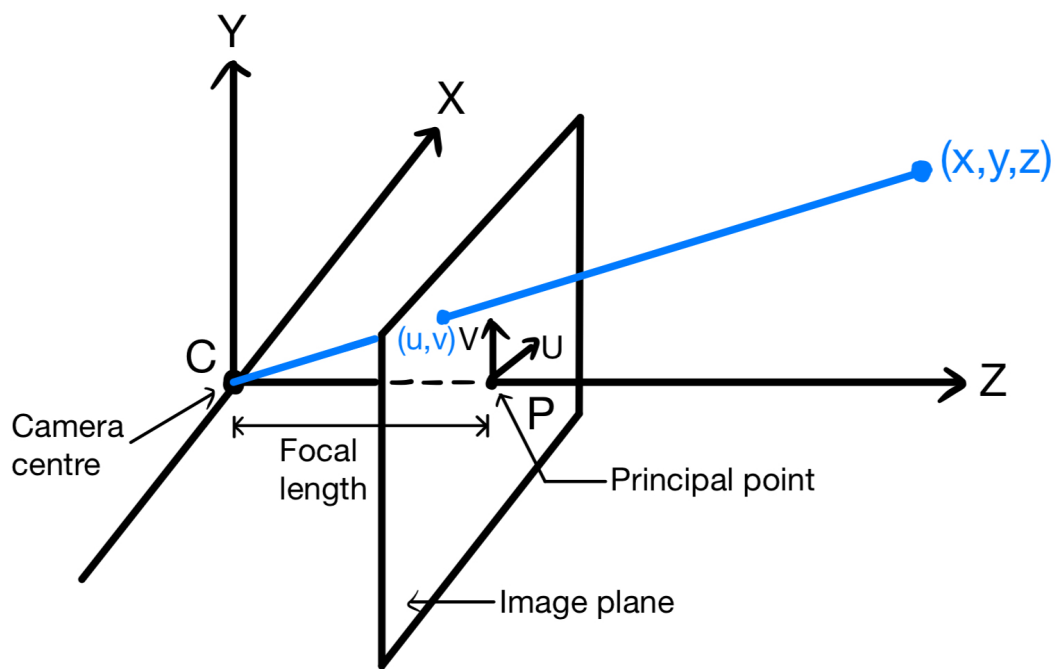


Figure 2.1: Pinhole camera models describe the projection of 3D world coordinates (x, y, z) onto 2D image coordinates (u, v) .

nates and its corresponding 2D point in image coordinates is formulated as:

$$\alpha \begin{bmatrix} u_C \\ v_C \\ 1 \end{bmatrix} = P_C \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.1)$$

Here, $(x, y, z, 1)^T$ represents a point in the 3D world (in homogeneous coordinates), and $(u_C, v_C, 1)^T$ represents the corresponding pixel location in the camera image. The matrix P_C is a 3×4 projection matrix that encodes both the intrinsic and extrinsic parameters of the camera [16, 25]. The scalar α reflects the scale ambiguity inherent in homogeneous coordinates, since image points are only defined up to scale in projective geometry. The Projection matrix can be decomposed into:

$$P_C = K_C \cdot [R_C \mid t_C] \quad (2.2)$$

Here,

- K_C is a 3×3 intrinsic matrix containing internal camera parameters such as focal length and principal point. These parameters describe how 3D points in the camera's coordinate system are projected onto the 2D image plane.
- $[R_C \mid t_C]$ is the extrinsic matrix, where R_C is a 3×3 rotation matrix and t_C is a 3×1 translation vector. These parameters describe the camera's orientation and position relative to the world coordinate system.

For a reference camera, we often assume one of the cameras in the setup is located at the origin (chosen arbitrarily) of the world coordinate system. In this case:

$$[R_{C1} \mid t_{C1}] = [I \mid 0] \quad (2.3)$$

Here, R_{C1} , t_{C1} represent the rotation matrix and translation vector of camera 1. Other cameras or projectors that are used may have different extrinsic parameters, such as R_{C2} , t_{C2} for a second camera, or R_P , t_P for a projector, which represent their pose relative to the world coordinate frame. The intrinsic matrix K_C is typically defined as [16, 25]:

$$K_C = \begin{bmatrix} f_x & s & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2.4)$$

where:

- f_x, f_y : focal lengths along the x and y axes. These represent the scaling factors that convert 3D camera coordinates to 2D image coordinates along each axis
- p_x, p_y : principal point coordinates (usually near the center of the image). It indicates where the optical axis intersects the image plane.
- s : skew parameter (typically zero), representing the non-orthogonality between the x and y image axes [17, 48].

Although the pinhole camera model is the most commonly used model in many computer vision tasks, it is a simplified representation that does not fully capture the behavior of real-world cameras. Even in the absence of imperfections, real lenses deviate from the ideal pinhole model due to their optical design. This deviation is particularly evident in cameras with wide-angle lenses, such as fisheye cameras, which exhibit significant lens distortion. Therefore, an extended version of the pinhole camera model that accounts for distortion is used to represent the cameras in this thesis.

2.1.2 Lens Distortion

To more accurately model real-world camera behavior, lens distortion must be taken into account [16]. Lens distortion is modeled using distortion parameters and can be broadly categorized into two types: radial distortion and tangential distortion [61, 76].

Radial Distortion

Radial distortion causes straight lines to appear curved, particularly near the image edges, due to the spherical shape of lenses. This effect is especially significant in wide-angle or fisheye lenses. To model radial distortion, we first express image coordinates in a normalized form, which removes the influence of the camera’s intrinsic parameters. This normalization is achieved by centering the coordinates at the principal point and scaling them by the focal lengths:

$$\hat{u} = \frac{u - p_x}{f_x}, \quad \hat{v} = \frac{v - p_y}{f_y} \quad (2.5)$$

Here, (u, v) are the pixel coordinates, (p_x, p_y) denote the principal point coordinates, and f_x, f_y are the focal lengths along the x and y axes, respectively. The normalized coordinates

$\hat{u}$, $\hat{v}$ represent the position of a point in the image plane relative to the optical center, scaled by the focal length. The radial distance from the optical center is then computed as:

$$r = \sqrt{\hat{u}^2 + \hat{v}^2} \quad (2.6)$$

Radial distortion is modeled by scaling the normalized coordinates based on their distance from the center:

$$\begin{aligned} \hat{u}_{\text{rad}} &= \hat{u} (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \\ \hat{v}_{\text{rad}} &= \hat{v} (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \end{aligned} \quad (2.7)$$

In this model, k_1, k_2, k_3 are the radial distortion coefficients. Typically, using k_1 alone is sufficient for accurate calibration in most practical application [18]. However, if the camera exhibits significant distortion, the other coefficients (k_2, k_3) may also be included.

By applying this model, we can correct for radial distortion, ensuring that straight lines in the real world are represented as straight lines in the image, which is crucial for accurate image analysis and 3D reconstruction.

Tangential Distortion

Tangential distortion [61] arises when the lens and the image sensor are not perfectly parallel. It is modeled as:

$$\begin{aligned} \hat{u}_{\text{tan}} &= \hat{u} + 2p_1 \hat{u}\hat{v} + p_2(r^2 + 2\hat{u}^2) \\ \hat{v}_{\text{tan}} &= \hat{v} + p_1(r^2 + 2\hat{v}^2) + 2p_2 \hat{u}\hat{v} \end{aligned} \quad (2.8)$$

Here, p_1, p_2 describe the tangential distortion coefficients and $\hat{u}_{\text{tan}}, \hat{v}_{\text{tan}}$ denote the tangential distorted normalized image coordinates.

In practice, both types of distortion are applied together to correct simulate lens behavior. These parameters are crucial for accurate camera calibration and undistortion of captured images.

2.2 Two-View Geometry

Following the mathematical formulation of the camera model and its projection behavior in Section 2.1.1, it is important to understand how multiple views of a scene interact geometrically. This leads to the concept of two-view geometry, which forms the basis for stereo vision and 3D reconstruction.

Two-view geometry [16, 25] describes the geometric relationship between two different views of the same 3D scene. These views may come from the same camera at different positions or from two separate cameras. Let I_l and I_r denote the left and right views, respectively.

Let a homogeneous 3D point be $X = (x, y, z, 1)^T$, and its projections onto the left and right views be $I_l = (u_l, v_l, 1)^T$ and $I_r = (u_r, v_r, 1)^T$. Let the projection matrices of the left and right cameras be P_l and P_r . Using Equation (2.1), the relationship between the 3D point in world space and the 2D image point for both cameras can be written as:

$$I_l = P_l X \quad (2.9)$$

$$I_r = P_r X \quad (2.10)$$

The relationship between I_l and I_r is governed by epipolar geometry, which is discussed in the following subsection.

2.2.1 Epipolar Geometry and Fundamental Matrix

Epipolar geometry [16, 25, 77] captures the relationship between two views of the same scene, independent of the scene's structure. A visual representation of this relationship is shown in Figure 2.2, which involves the epipolar plane, epipoles, and epipolar lines.

- **Epipolar Plane:** The plane formed by the 3D point X and the optical centers of both cameras (C_l and C_r).
- **Epipoles:** The points of intersection between the line connecting C_l and C_r with each image plane, denoted as e_l and e_r .
- **Epipolar Lines:** The intersections of the epipolar plane with each image plane.

The fundamental matrix (F) [25, 77] is a 3×3 rank-2 matrix that encodes the relationship between the epipolar geometry of two uncalibrated images. The relationship between F and the corresponding views is defined as:

$$I_l^T F I_r = 0 \quad (2.11)$$

This equation, known as the epipolar constraint, ensures that X , C_l , C_r , and the image points I_l and I_r lie on the same epipolar plane. As a result, when searching for a

corresponding point in the second image, it is sufficient to search along the epipolar line associated with the first image, rather than the entire image. This significantly reduces the search space.

The fundamental matrix is commonly estimated using the eight-point algorithm [80]. This algorithm requires at least eight pairs of corresponding points to solve a linear system derived from the epipolar constraint. Using F for scene recovery is known as weak calibration because it reconstructs the scene up to a projective transformation, leading to projective ambiguity [25].

Let H be any non-singular 4×4 matrix representing a projective transformation in 3D. The new projection matrices and the 3D points after applying H to the scene are defined as:

$$\tilde{P}_l = P_l H^{-1}, \quad \tilde{P}_r = P_r H^{-1}, \quad \tilde{X} = HX \quad (2.12)$$

These transformed matrices yield the same image points:

$$I_l = \tilde{P}_l \tilde{X}, \quad I_r = \tilde{P}_r \tilde{X} \quad (2.13)$$

This implies that the same fundamental matrix F is preserved, thereby introducing projective ambiguity. Without additional constraints, the true Euclidean structure of the scene cannot be recovered. To overcome the limitations of projective ambiguity, the concept of the essential matrix is introduced when the camera intrinsics are known.

2.2.2 Essential Matrix

The essential matrix (E) [25, 81] is a special case of the fundamental matrix used when the intrinsic parameters of both cameras are known. It is defined as:

$$E = K_l^T F K_r \quad (2.14)$$

where K_l and K_r are the intrinsic matrices of the left and right cameras, respectively. By incorporating the intrinsic parameters of both cameras, the essential matrix allows a strong calibration because it facilitates a more precise reconstruction and eliminates the projective ambiguity associated with the fundamental matrix, as described in Section 2.2.1. It is also defined as:

$$E = [t]_{\times} R \quad (2.15)$$

where $[t]_{\times}$ is the skew-symmetric matrix form of the translation vector t , and R is the rotation matrix of the right view with respect to the left view. Decomposing E using Singular Value Decomposition (SVD) [71] yields R and t up to a scale factor. This enables accurate scene reconstruction through a process known as triangulation.

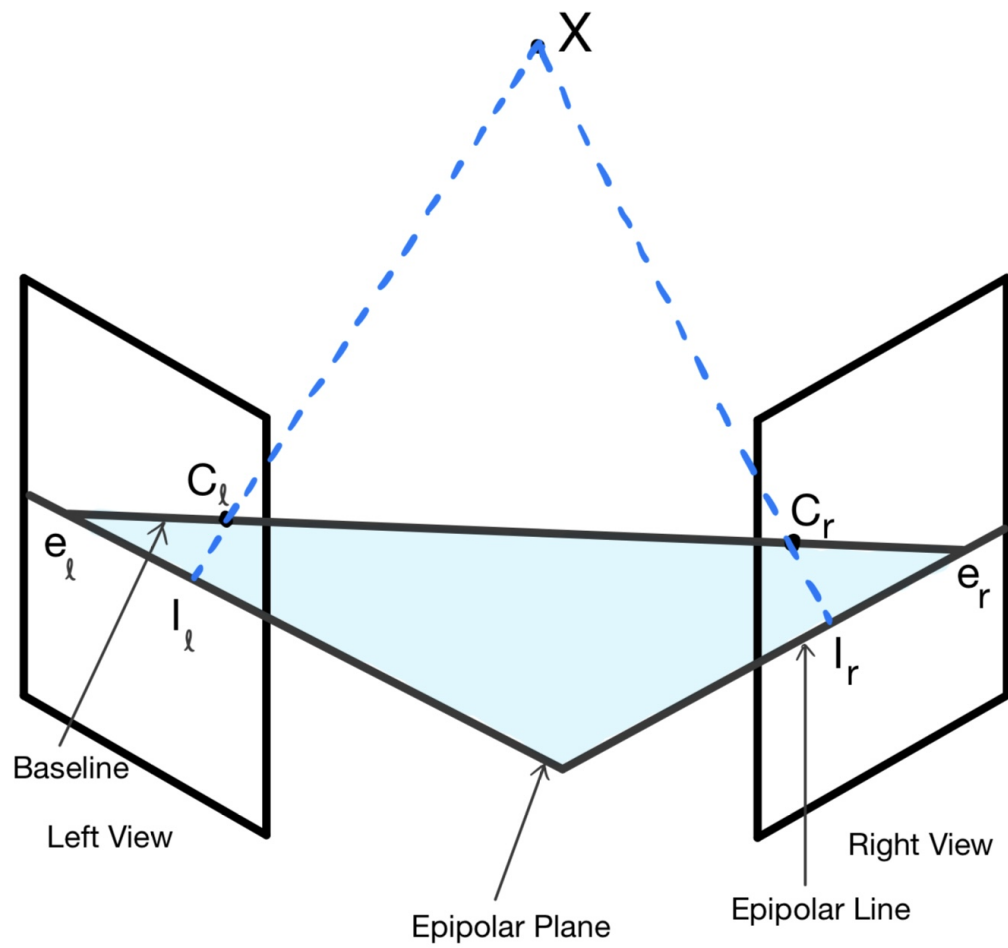


Figure 2.2: Two-view geometry showing epipolar constraints between corresponding image points. These constraints enable triangulation of the 3D point from its image projections.

2.2.3 Triangulation

Given P_l , P_r , and the corresponding image points I_l and I_r , the 3D point X can be estimated through a process called triangulation [16, 25]. It is a widely used method for depth estimation and 3D reconstruction in stereo vision [14, 42, 72].

In ideal conditions, the back projected rays from I_l and I_r would intersect precisely at X . However, due to measurement noise and calibration errors, these rays often do not intersect accurately. Triangulation estimates the 3D point X by finding the point in space that best satisfies the projection constraints from both views. This discrepancy between observed and ideal projections is also known as the reprojection error.

2.2.4 Reprojection Error

Reprojection error [16, 25] measures the difference between the observed image points and the corresponding world points projected back into the same image. The reprojection error (E_R) is defined as:

$$E_R = \sum_m \|I_{im} - \pi(P_i, X_m)\|^2 \quad (2.16)$$

where

- X_m : 3D coordinates of point m in the world coordinate system.
- I_{im} : 2D coordinates of the projection of point m onto projector i .
- P_i : Projection matrix for projector i , containing both intrinsic and extrinsic parameters.
- $\pi(P_i, X_m)$: The projection of X_m onto projector i using P_i .

In the process of calibrating projectors, both manual and auto-calibration methods (discussed in detail in Section 2.4.3) apply optimization techniques that aim to minimize the reprojection error. As a result, it is commonly used as the objective function in projector calibration [27, 38, 44, 69, 75].

2.3 Camera Calibration

Following the discussion on two-view geometry in Section 2.2 and its role in establishing spatial relationships between views, the next critical step in any stereo vision or 3D reconstruction system is camera calibration.

Geometric Camera Calibration [54, 79], which is also known as camera resectioning, is a fundamental computer vision process that establishes the relationship between 3D world coordinates and 2D image projections. Camera calibration is the process of determining the intrinsic and extrinsic camera parameters. Once the camera parameters are obtained, they are used to estimate the intrinsic and extrinsic parameters of the projector and also the estimation of the 3D scene.

Extensive research has been conducted to develop precise calibration techniques for optical devices [4, 50, 54, 55, 78]. Camera calibration techniques can be categorized into manual and self-calibration methods [58]. Manual calibration often involves using a known pattern such as a checkerboard to capture images from different angles and estimate the camera's parameters. Whereas self-calibration doesn't require any physical calibration object. Instead, it relies on the analysis of image sequences or motion to determine the camera's parameters.

2.3.1 Manual Calibration

Manual calibration [54, 55, 79] is a camera calibration technique that utilizes a physical calibration target with known dimensions typically utilizing a checkerboard grid or fiducial markers. Research on camera calibration dates back to the early 20th century with a major breakthrough in 1919 when Conrady introduced a model that classified lens distortion into Radial and Tangential [11], as described in detail in Section 2.1.2. This model was further refined in 1966 by Brown which led to the development of the Brown-Conrady model [13]. Later in 1971 Brown used his model and presented his plumb line method to calibrate the distortions [47]. Also in this method Brown defines a formula that accurately models radial distortion at a range of varying focal lengths. The advantage with this method is that it achieves high accuracy at a close range. However, this method has limitations. It includes manual determination of the points required for calibration which is time consuming and susceptible to human error.

To overcome the reliance on fiducial markers, Abdel-Aziz and Karara proposed the Direct Linear Transformation (DLT) method [33]. It is a widely employed method for directly deducing the projection matrix from image and world points without the need

of iterations. This method mandates a minimum of six pixel correspondences, aligning observed points in an image with their respective 3D world coordinates, and requires the surface to be non-planar. However, the accuracy of this method is greatly affected by noise. Hence, to improve the performance of calibration non-linear optimization is performed [35, 49].

To mitigate the drawbacks of Brown’s model, in 1987 Tsai proposed a two-step calibration method [55] which minimizes the dependency on manual input. Tsai’s method requires at least eight feature points per image, and it also solves the calibration problem using a set of linear equations based on the radial alignment constraint. It employs a second-order radial distortion model but does not consider the tangential distortion terms. Additionally, this method assumes that some parameters related to the camera would be provided by the manufacturer. By incorporating the manufacturer provided camera parameters, this method improves the accuracy and speed of calibration. The two-step technique can be applied to a single image or multiple images of a 3-D grid or plane calibration grid if the coordinates of the grid points are known. The advantage with this method is its reduced error in calibration, which is primarily because of the reduced manual point selection. However, Tsai’s method is affected by feature point noise and precise grid placement to some degree.

Although Tsai’s method was a huge improvement, it was difficult to incorporate into general purpose camera due to its dependency on the manufacturer data. All these demerits were eliminated by Zhang in 2000, when he developed his simplified planar checkerboard method [78, 79]. Instead of relying on manual 3D points, Zhang’s method is a flexible method which only requires a planar checkerboard grid at different orientations (more than two) in front of the camera for the calibration. Zhang’s algorithm extracts corner points from the checkerboard pattern and computes a projective transformation between the image points of multiple views, up to a scale factor. This method computes the camera parameters using a closed-form solution. Additionally, the third- and fifth-order radial distortion terms are estimated through a linear least-squares solution, followed by a final non-linear optimization of the reprojection error using the Levenberg-Marquardt method [19]. Zhang’s method is robust to noise and allows calibration in diverse setting without the need for specialized equipments, relying solely on images of the planar checkerboard captured from multiple orientation.

In summary, methods for the manual calibration of cameras have evolved significantly from the early distortion models by Conrady to a more refined method by Zhang. While Brown’s approach laid the foundation, Tsai’s method introduced a computationally efficient calibration model which is more practical. Finally, Zhang’s method further improved the process by utilizing multiple images and a non-linear optimization to achieve higher

accuracy with minimal manual intervention. Although numerous researchers around the world have proposed many methods for camera calibration, the most adopted methods are those of the DLT, Tsai and Zhang method. In this thesis, Zhang’s method is used for the ground truth calibration of cameras.

2.3.2 Self-Calibration

Unlike manual calibration, which requires calibration patterns or known point correspondences, self-calibration methods [43, 59, 79] rely solely on mathematical constraints to estimate the intrinsic parameters of cameras. These methods are especially useful when the use of calibration patterns is impractical.

The Kruppa equations [26] represent one of the earliest self-calibration approaches. They establish a direct relationship between the fundamental matrix and the intrinsic camera parameters. However, this method suffers from several limitations, including high sensitivity to noise in the fundamental matrix estimation and a reliance on known principal points for numerical stability.

To improve robustness and simplicity, the Mendonça–Cipolla method was introduced [41]. This approach uses the essential matrix, imposing the constraint that it must have two identical non-zero singular values and one singular value equal to zero. The method supports variations in focal lengths and principal points and delivers accuracy comparable to more complex algorithms. However, the specific constraints on the essential matrix can introduce challenges, such as sensitivity to initial parameters and is vulnerable to local minima and convergence issues when estimating multiple parameters.

Recently, camera self-calibration has increasingly shifted toward learning-based algorithms. Fang et al. [15] proposed a self-supervised approach that estimates camera intrinsics directly from video sequences without requiring calibration patterns. By relying on frame-to-frame consistency, their method achieves competitive accuracy and is adaptable to various camera models. Subsequently, Jeong et al. [31] introduced a technique that jointly optimizes camera parameters and scene geometry using photometric cues. This method, known as Self-Calibrating Neural Radiance Fields, is particularly effective in scenarios lacking a predefined scene layout, although it demands substantial computational resources.

Liao et al. [37] conducted a comprehensive survey of deep learning-based calibration methods, highlighting both their potential and associated challenges. These recent advancements reflect the growing interest in using learning-based techniques for camera calibration, marking a shift away from traditional manual approaches.

Despite significant advancements in learning-based self-calibration methods [15, 31, 37], camera self-calibration remains challenging in practical applications. A major limitation is its sensitivity to noise in feature matching and fundamental matrix estimation. In addition, non-linear optimization often suffers from convergence issues. These limitations make manual calibration approaches preferable in applications where high precision is required.

2.4 Projector Calibration

Following the calibration of cameras in Section 2.3, the next essential step in a projection mapping system is the accurate calibration of projectors [44, 70]. In projection mapping, multiple projectors are used simultaneously with overlapping content. To ensure proper alignment and create a seamless composite image across all projectors, it is essential to determine both the internal parameters and the pose of each projector. Therefore, once the cameras are calibrated, they are used in conjunction with the projectors to estimate the intrinsic matrix K_P , the rotation matrix R_P , and the translation vector t_P for each projector in the setup.

2.4.1 Adapting the Camera model for Projector Calibration

By definition, a camera is a device that captures light rays from the environment to form an image, while a projector emits light rays to display images onto the surface. Due to this relationship, in computer vision and projection mapping, a projector is modeled as the inverse of a camera [40]. The inverse relationship between the camera and projectors allows the adoption of camera calibration models for the purpose of projector calibration. Therefore, the projector model can be derived from Equation (2.1), Equation (2.2), and Equation (2.4), and is expressed as follows:

$$\alpha \begin{bmatrix} u_P \\ v_P \\ 1 \end{bmatrix} = P_P \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.17)$$

$$P_P = K_P \cdot [R_P \mid t_P] \quad (2.18)$$

Here, $(u_P, v_P, 1)^T$ represents the corresponding pixel location in the projector image plane and the matrix P_P is a 3×4 projection matrix that encodes both the intrinsic and extrinsic

parameters of the projector. R_P , t_P represent the rotation matrix and translation vector of projector. K_P can be decomposed into:

$$K_P = \begin{bmatrix} f_x & s & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2.19)$$

Similar to cameras, projectors also exhibit distortion parameters such as radial and tangential distortion. As the projectors used in our setup exhibited minimal distortion, this thesis assumes projectors with negligible distortion. Therefore, distortion parameters are not considered in the calibration process.

2.4.2 Structured Light

To accurately calibrate projector-camera systems and reconstruct 3D surfaces, structured light techniques are commonly employed [39, 70]. Structured light is one of the most effective approaches for achieving dense pixel correspondence between cameras and projectors. Although the process can be time-consuming, it offers a level of precision and pixel density that other methods generally lack.

Structured light involves projecting known pattern of light onto an object and analyzing the deformation of these patterns to obtain the depth information and the surface geometry. In a typical camera-projector system, projectors emit specific light patterns onto a target surface, which will be captured by one or more cameras viewing the patterns at different angles. By comparing the deformed patterns captured by the cameras with the actual patterns emitted by the projector, pixel correspondence can be established between the cameras and also between the projectors and the cameras. The pixel correspondence can be expressed mathematically as a mapping function:

$$f : (u_c, v_c) \mapsto (u_p, v_p)$$

where (u_c, v_c) represents the coordinates of a pixel in the camera image, and (u_p, v_p) represents the coordinates of the corresponding pixel in the projector image. The projector-camera system can utilize these pixel correspondences to triangulate and obtain the dense 3D surface reconstruction [16, 24]. The accuracy of 3D reconstruction depends on the precision of the pixel correspondence. Hence, the structured light patterns are designed to maximize the uniqueness and detectability of projected features across the entire field of view of the cameras [5].

Structured light methods can be classified into two categories, namely multiple-shot and single-shot techniques [5, 20]. Multiple shot involves projecting a series of images onto the target, to obtain an accurate and dense reconstruction. However, this method is limited to static surfaces. Examples of multiple-shot techniques include binary coding, Gray code, and phase shifting [20]. On the other hand, single-shot technique projects only a single pattern onto the target to obtain the pixel correspondence. This makes it suitable for dynamic scenes and it offers faster acquisition but suffers from lower accuracy compared to multiple-shot methods. Single-shot techniques use spatial multiplexing, where the projected pattern contains unique features that are easy to decode like the color coded patterns [20].

Since our work involves a static physical surface and does not impose constraints on pattern capture time, we adopt a binary coded multiple-shot technique [5, 20] for generating dense pixel correspondences. An example of such a structured light pattern is illustrated in Figure 2.3.

2.4.3 Calibration Techniques

With the camera–projector correspondences obtained using structured light, the next step is to calibrate the projectors. The most widely adopted approach for this is an adaptation of Zhang’s method [78], previously discussed in Section 2.3.1, which is typically used for camera calibration. In projector calibration, Zhang’s method involves projecting patterns onto a surface and capturing the projection using calibrated cameras. These captured patterns are used to determine 3D-to-2D correspondences, which are then solved to estimate the projector parameters.

Projector calibration techniques can be broadly classified into two categories: manual calibration [44, 51] and auto-calibration [69, 70]. Similar to manual camera calibration, manual projector calibration relies on physical patterns or additional apparatus.

Early approaches treated projectors as inverse cameras, utilizing planar checkerboards and homography estimation. Raskar et al. [51] projected checkerboard patterns onto flat surfaces and computed homographies based on camera-captured correspondences to correct keystone distortion. Moreno and Taubin [44] enhanced this method by integrating Gray code patterns and OpenCV-based sub-pixel corner detection, reducing reprojection errors to below 0.1 pixels through bundle adjustment [30, 65, 73] and local homography estimation.

Recent advancements in manual calibration of projectors involve utilizing phase-shifting sinusoidal patterns to improve projector calibration accuracy [36], where structured light is used to create dense correspondences between the projector and camera views. Another

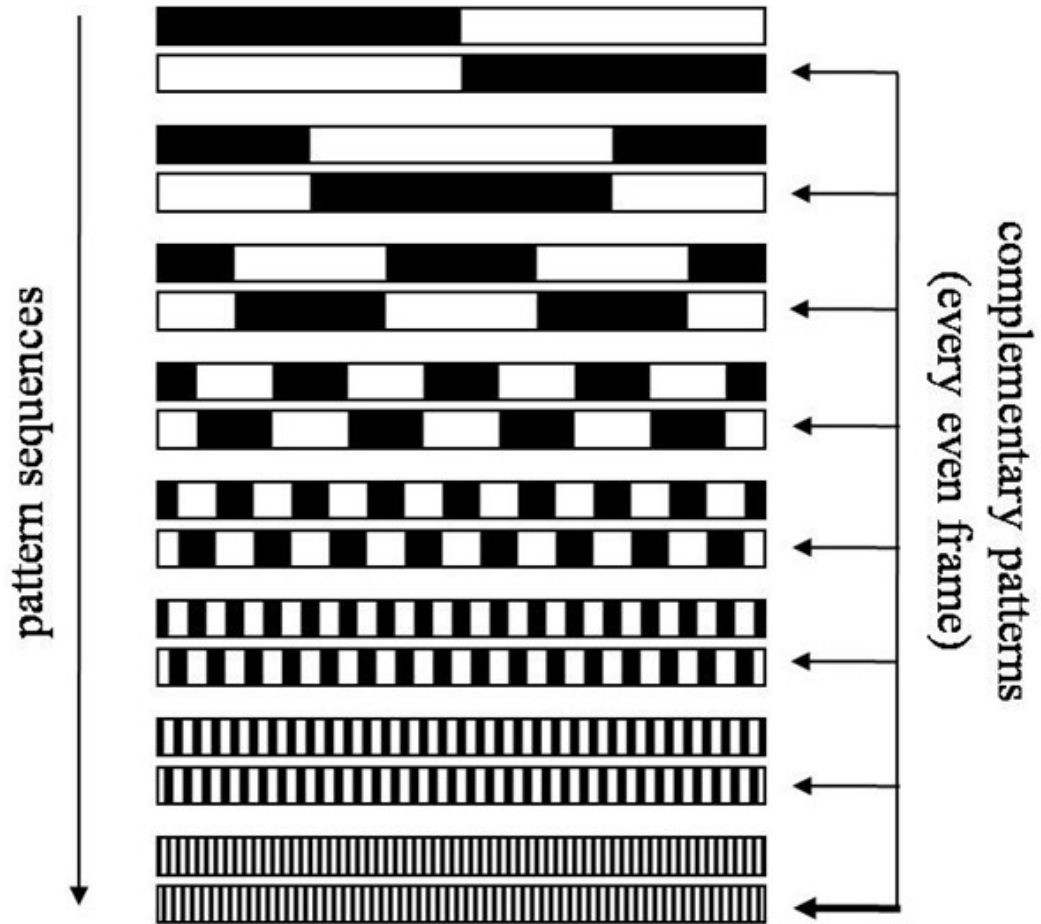


Figure 2.3: Illustration of binary coded structured light pattern. Image copied from [\[56\]](#)

approach involves using a planar mirror in complex scenarios to calibrate projector–camera systems with non-overlapping fields of view [67]. This technique uses indirect correspondence estimation by reflecting patterns, which enables calibration even when the fields of view of the camera and projector do not overlap.

Although effective and capable of achieving high accuracy, these methods are often difficult to apply in real-world scenarios, as the manual placement of calibration targets is labor-intensive and prone to errors caused by lens distortion when projectors are defocused. Furthermore, embedding markers or sensors on all surfaces is not feasible. In cases where projection is performed on distant target surfaces, sensors may not be sufficiently visible, rendering the calibration process ineffective.

In contrast, auto-calibration techniques eliminate the need for printed patterns, embedded sensors, or manual adjustments. Instead, they rely on structured light to establish dense correspondences between the projector and the camera. These correspondences enable the system to compute homographies and distortion parameters automatically, facilitating the calibration of both intrinsic and extrinsic projector properties.

One of the earliest methods in auto-calibration was proposed by Yamazaki et al. [70], utilizing structured light projection for the simultaneous calibration of a projector and a camera. This technique estimates the intrinsic parameters of both devices and their relative pose without requiring any calibration boards. The method works by decomposing the radial fundamental matrix into intrinsic and extrinsic parameters. To overcome sensitivity issues in the decomposition of the fundamental matrix, several steps are taken. These include placing the principal point at the center of lens distortions, which is not always the case and performing a final optimization to attain accurate results.

Willi and Grundhöfer [69] proposed a robust auto-calibration method that does not require any calibration board or prior geometric knowledge of the projection surface. Similar to the previous method, this approach uses structured light to obtain dense correspondences and later employs iterative optimization to calibrate the cameras and projectors. This assists in calibration on non-planar surfaces. However, the calibration of cameras assumes the principal point to be at the center of the image, which is not always the case and can affect the calibration accuracy of projectors.

Further developments in self-calibration of projectors involve decomposing structured light into directional components to estimate geometry without using a physical calibration target [60]. This method analyzes the deformation of light patterns on the surface and infers the projector parameters automatically. Recently, a quasi-calibration technique [57] introduced the use of an auxiliary camera to estimate correspondences indirectly. However, this method requires an additional camera in order to calibrate the projector.

Although these established algorithms work well, they impose constraints on the parameters during calibration, which, if not satisfied, can significantly impact the accuracy of projector calibration. Additionally, for accurate calibration of projectors, researchers often employ optimization along with initial estimates of the parameters to best estimate the projector’s characteristics. However, in certain cases when the overlap is minimal, it results in poor calibration of the projectors.

This thesis builds on these advancements and employs an auto-calibration approach by projecting structured light patterns onto the target surface to obtain dense correspondences, which are later used to calibrate the projectors. An additional contribution is the introduction of a new objective function that optimizes the overall calibration on top of existing methods, aiding in the accurate calibration of projectors even when there is limited stereo overlap.

In summary, this chapter introduced foundational concepts such as camera modeling, two-view geometry, reprojection error, and existing calibration techniques for both cameras and projectors. These form the basis for understanding how cameras and projectors interact within a 3D environment, and provide the necessary background to evaluate and apply calibration methods. A solid understanding of these concepts is essential for identifying the limitations in existing approaches and for understanding the theoretical foundations of the proposed solutions.

Chapter 3

Problem Formulation

As mentioned in Section 2.3, the calibration of projectors begins with the calibration of cameras, which can be performed using the Direct Linear Transformation (DLT) method or self-calibration methods such as the Mendonça–Cipolla method or the learning based approaches. Once the cameras are calibrated, they are utilized in the calibration of the projectors.

To calibrate the projectors using cameras, typically in a setup involving one or more projectors and at least two calibrated cameras, point correspondences between the cameras and the projectors are established through structured light. After the structured light patterns are properly projected and captured, traditional methods are employed to obtain an initial estimate of the projector parameters, as mentioned in Section 2.4.3. Subsequently, optimization algorithms can be utilized to refine these parameters for improved accuracy.

However, projectors are subject to certain constraints to achieve accurate calibration through these optimization processes. One such constraint is the requirement to maintain sufficient spatial overlap between each projector’s projection area and the field of view of both stereo cameras, as discussed in the motivation section of Chapter 1. These constraints increase the overall system cost compared to simpler setups, and adding more devices also makes the calibration process take longer.

In this chapter, we will analyse the existing problems in the current methodology of projector calibration and develop a novel solution to address it.

3.1 Stereo Overlaps and their Significance

From the motivation in Chapter 1, it is clear that the calibration of projectors depends on the amount and type of overlap between optical devices, such as between projectors and between projectors and cameras. Understanding the effects of different types of overlap helps to better identify the sources of error in projector calibration, which in turn enables the development of more accurate calibration methods.

There are three possible types of overlaps that can occur during projection mapping:

1. Overlap between the Cameras
2. Overlap between the Camera and the Projector
3. Overlap between the Projectors

Typically, for projection mapping to function correctly, at least one camera must observe each region of the surface with a small stereo overlap between cameras. However, for precise calibration in industrial settings, two cameras are usually assigned to each surface region to obtain accurate depth measurements. This requirement increases the overall cost of the setup.

Reducing the three overlaps listed above allows for broader surface coverage while lowering setup costs. However, minimizing overlap also increases projector calibration error. Therefore, analyzing and understanding these overlaps is essential for producing high-quality visual results at a reduced cost.

Comprehending these overlaps facilitates more effective use of optimization techniques and contributes to the accurate calibration of all devices involved. In this work, we consider a minimal setup consisting of two cameras and one or more projectors, which ensures sufficient stereo overlap while keeping the setup cost-effective. A detailed analysis of the types of overlaps is presented in Chapter 4.

3.2 Calibration of Projectors

3.2.1 Need for Auto-Calibration of Projectors

As explained in Section 2.4.3, auto-calibration of projectors involves calibrating projectors without the use of any physical markers or checkerboards. The auto-calibration approach addresses several limitations associated with manual calibration methods, such as:

1. **Manual Intervention:** Manual calibration requires the placement and alignment of calibration targets, which is labor-intensive and prone to human error.
2. **Scalability:** In large projection mapping setups, deploying physical calibration targets or embedding physical markers becomes impractical.
3. **Cost:** Manual calibration increases the total equipment required for projection mapping, thereby elevating the overall cost of the setup.

To address these challenges, auto-calibration methods for projectors have been introduced, as discussed in detail in Section 2.4.3. However, they present certain limitations that need to be addressed to ensure accuracy and cost-effectiveness.

3.2.2 Challenges

Despite the advantages of auto-calibration over manual methods, it has certain limitations. The key challenges include dependency on sufficient stereo overlap between projectors and cameras and the limitations of using reprojection error as the sole optimization criterion.

Dependence on Stereo Overlap

As discussed in Section 1.1, auto-calibration of projectors requires sufficient stereo overlap between the cameras and each projector’s projection area to ensure enough valid correspondences for accurate calibration. For instance, the algorithm developed by Tehrani et al. [62] requires that at least 20% of the area projected by each projector be consistently observed by a minimum of two cameras to obtain sufficient point correspondence data for calibration. Due to this requirement, large-scale projection mapping faces two primary issues:

1. **Increased number of cameras:** To maintain the necessary stereo overlap, more cameras must be deployed compared to setups without overlap constraints, which in turn increases the overall system’s capital cost.
2. **Reduced accuracy with Fewer Cameras:** To manage costs by limiting the number of cameras, the calibration accuracy for projectors may be compromised, adversely affecting the quality of the projected display.

Limitations of E_R as an Optimization Criterion

During the optimization of projector parameters, most algorithms use the reprojection error (E_R) as their objective function, as discussed in Section 2.2.4. While minimizing E_R is intended to achieve accurate projection matrix (P_i) for each projector i , it may lead to suboptimal solutions due to the presence of multiple local minima specific to the camera-projector calibration setup [25, 64]. These issues are due to the complexity of the projection geometry and limited visibility across devices [78]. Additionally, reprojection error does not inherently account for outliers in the point correspondences, making it sensitive to noise and bad correspondences [25].

In many practical situations, the actual point cloud of the target surface remains unknown. Since, obtaining the point cloud of the target surface requires laser scanning of the target surface. However, in large scale projection mapping, performing laser scan on the surface, increases the overall capital cost by a huge factor [2]. In such scenarios, relying solely on E_R can result in misalignments between overlapping projector regions, leading to visible gaps and discontinuities on the target surface.

3.3 Proposed Strategy

3.3.1 Overlapping Distance Error

As discussed in the previous section on the challenges of auto-calibrating projectors, utilizing the reprojection error (E_R) without access to the ground truth point cloud necessitates a higher degree of overlap between the projector’s projection area and the field of view of the stereo cameras. Insufficient overlap in this context can reduce the calibration accuracy of the projectors, leading to misalignments in regions where projector content overlaps on the target surface and resulting in visible gaps. This overlapping distance error can be numerically defined as:

$$\sum_{m \in \mathcal{M}_{ij}} \|X_m^i - X_m^j\|^2 \quad (3.1)$$

where X_m^i and X_m^j are the 3D coordinates of point m in the world coordinate system, reconstructed independently by projectors i and j , respectively. The set $\mathcal{M}_{ij}$ contains the world points commonly observed and reconstructed by both projectors.

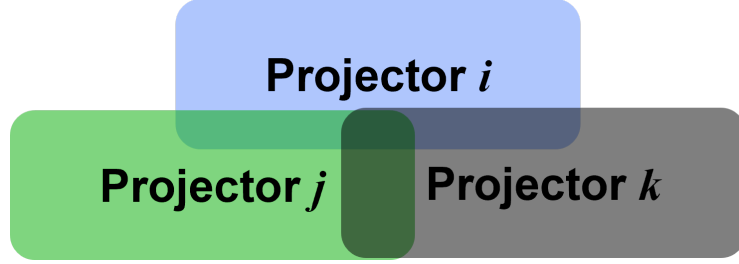


Figure 3.1: Illustration of the cyclical effect in pairwise projector optimization. While errors between adjacent projectors (e.g., i - j and j - k) are minimized, the lack of a direct constraint between i and k can lead to misalignment.

3.3.2 Cyclical Effect

Addressing the overlapping distance error individually for each projector introduces a cyclical effect. For instance, consider a setup with three projectors: i , j , and k , each overlapping with the others, as illustrated in Figure 3.1. Minimizing distance error between projectors i and j , and then between j and k , may inadvertently increase the error between projectors i and k . This occurs because the optimization is performed only on pairwise overlaps, without considering the system as a whole. As a result, calibration errors may accumulate, preventing consistent alignment across all projectors. In practical applications, large setups can involve more than ten projectors, complicating the problem and making it difficult to reach a configuration that minimizes errors across the entire system.

To address this issue, we introduce a global distance-based objective function E_{OD} , designed to simultaneously optimize all projectors. This optimization problem is mathematically formulated as:

$$\min_{\{P_i\}} E_{OD}, \quad E_{OD} = \sum_i \sum_{j>i} \sum_{m \in \mathcal{M}_{ij}} \|X_m^i - X_m^j\|^2 \quad (3.2)$$

The set $\{P_i\}$ includes all projector parameters being optimized. This objective function aims to minimize the error between overlapping content across all projectors simultaneously. Here, $\mathcal{M}_{ij}$ denotes the set of 3D points commonly observed and reconstructed by both projectors i and j , and X_m^i and X_m^j are the reconstructed coordinates of point m from projectors i and j respectively.

3.3.3 Multi-Step Optimization Strategy

The goal of this thesis is to accurately calibrate the projectors using a minimal number of cameras, as described in Section 1.2. Reducing the number of cameras results in some projectors having limited or no overlap between their projection areas and the stereo cameras' fields of view, making it difficult to establish direct camera-projector correspondences. Additionally, there may be projectors that do not overlap with the stereo cameras at all, but only with other projectors.

In such scenarios, using the proposed objective function (Equation 3.2) alone and placing it on top of traditional methods could lead to divergence in the optimization results or failure to calibrate certain projectors. However, extensive research involving the calibration of projectors using optimization based on E_R has shown that it is a robust approach and can optimize parameters to a reasonable degree [27, 38, 44, 75]. Based on this, we propose a multi-step optimization strategy, detailed below:

1. **Initial Estimate:** Obtain an initial estimate of projector parameters using traditional calibration methods.
2. **Optimization 1:** Refine the parameters by minimizing the reprojection error (E_R , see Equation 2.16).
3. **Optimization 2:** Further refine the parameters using the proposed overlapping distance error (E_{OD} , see Equation 3.2) to align overlapping regions across projectors and form a unified, joint point cloud of the target surface.

The addition of a second optimization stage on top of E_R improves the robustness of the algorithm and ensures convergence, as Optimization 1 provides a reliable initial estimate.

In summary, this chapter identified the key challenges in projector calibration, particularly under minimal stereo overlap. It also highlighted the importance of analyzing stereo overlaps and emphasized the need for auto-calibration of projectors. The proposed strategy introduced in this chapter incorporates a novel objective function and a multi-step optimization process, aiming to improve projector calibration performance when using the minimum number of cameras.

In the subsequent chapters, this thesis will analyze and explore the different types of stereo overlap and will ultimately present the implementation results of the novel objective function introduced in this chapter.

Chapter 4

Stereo Overlap Between Optical Devices

Overlap between optical devices plays a crucial role in the overall calibration of the system. As described in Section 3.1, multiple types of overlap are involved throughout the projector calibration process, namely the overlap between camera views, between the camera and projector views, and between projector projections. An illustration of the general concept of field-of-view overlap between two devices is shown in Figure 4.1. Each device may represent either a camera or a projector, and the highlighted region indicates the shared area observed or covered by both devices. In this chapter, based on the overlaps mentioned above, we will analyze three different case studies:

- Case Study 1: Impact of the Degree of Stereo Overlap on Camera Calibration Accuracy
- Case Study 2: Impact of Camera Overlap on Projector Calibration Accuracy
- Case Study 3: Calibration Accuracy in Sequentially Overlapping Projectors

The following sections will present the synthetic data used in the analysis, outline the detailed methodology for each case study, and finally analyze and examine their impact on calibration accuracy.

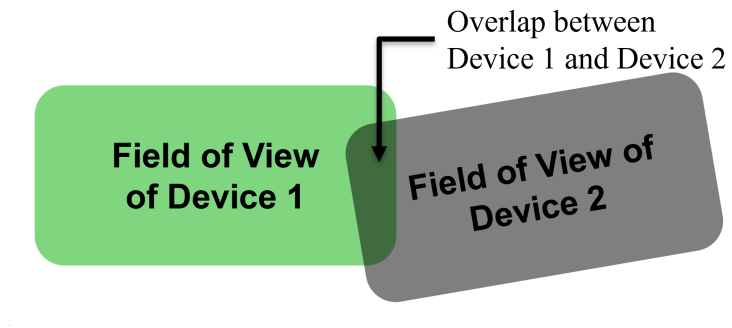


Figure 4.1: Illustration of field of view overlap between two devices. The highlighted region represents the shared area observed or projected by both Device 1 and Device 2.

4.1 Synthetic Data

Synthetic data is generated for this analysis. A dome structure and a wave structure (as shown in Figure 4.2) are used as the unknown 3D surfaces observed by the cameras and displayed onto by the projectors. The dome structure is used in Case Studies 1 and 2, while the wave structure is employed as the unknown 3D surface in Case Study 3. Initially, the correspondences between the camera image points and the world points are identified by multiplying the projection matrix with the world points. This transformation is represented by

$$\underline{u}_c = P_C \cdot \underline{u}_w \quad P_C = K_C \cdot [R_C \mid t_C] \quad (4.1)$$

where $\underline{u}_c$ represents the image points in the camera plane, $\underline{u}_w$ represents the world points, and P_C , R_C , t_C represent the projection matrix, rotation matrix and the translational vector of the camera. Two cameras are used for the analysis, however the number of projectors varies depending on the case study. Subsequently, correspondences between Cameras are identified by matching the world points projected onto the respective image planes, and designated as $(\underline{u}_{c1}, \underline{u}_{c2})$.

Additionally, the transformation described by Equation 4.2 is applied to the projector to establish correspondences between the world points and the image points on the projector's image plane. Finally, the correspondences between the projector and the cameras are established by matching the world points in the projector and camera image planes, and designated as $(\underline{u}_{c1}, \underline{u}_p)$ and $(\underline{u}_{c2}, \underline{u}_p)$.

$$\underline{u}_p = P_P \cdot \underline{u}_w \quad P_P = K_P \cdot A_p \quad (4.2)$$

where $\underline{u}_p$ represents the image points in the projector plane, and P_P is the projection matrix of the projector. To enhance the realism of the synthetic data, Gaussian noise (ϵ) is

added to the correspondences. ϵ is a random variable which follows a normal distribution with a mean of zero and a standard deviation of σ and is expressed as:

$$\epsilon \simeq \mathcal{N}(0, \sigma^2) \quad (4.3)$$

where the values of σ used for these case studies are 0.25, 0.5, 1. These values represent increasing levels of noise added to the image data and help evaluate how well the calibration process performs under different conditions. Higher values were not considered, as they would introduce extreme distortions which is unlikely to occur in practice, while lower values would not provide enough variation to reveal meaningful trends.

For all case studies, K_C and K_P are assigned standard values at the outset. Additionally, K_{C1} and K_{C2} are defined to be identical, as are all projectors:

$$K_C = K_{C1} = K_{C2}, \quad K_P = K_{P1} = \dots = K_{Pn} \quad (4.4)$$

where n is the number of projectors used in that particular case study and K_{Pn} is the intrinsic matrix of Projector n . Camera 1 is always considered as the origin, as described in Section 2.1.1. Hence, the extrinsic parameters of Camera 1 are defined as:

$$[R_{C1} \mid t_{C1}] = [I \quad 0] \quad (4.5)$$

The subsequent sections will outline the methodology and provide a detailed analysis of each case study, to understand the influence of different types of overlap on the accuracy of projector calibration.

4.2 Methodology and Experimental Setup

4.2.1 Impact of the Degree of Stereo Overlap on Camera Calibration Accuracy

Stereo Overlap between the cameras refers to the shared region between the field of view (FOV) of two or more cameras. In projection mapping, the purpose of camera is to capture the physical surface onto which projections are made. However, stereo overlap between the cameras helps in determining the depth of the physical surface, which plays a crucial role in accurate projector calibration. Since, in the absence of the ground truth point cloud of the target surface and the physical embedding, this stereo surface acts as the reference onto which rest of the projector point cloud is built to calibrate it.

To analyze the accuracy of camera calibration in relation to stereo overlap (α) between the cameras, two cameras are used in the setup. Where t_{C2} is iteratively increased such that, camera 2 is moved away from the camera in both the negative and positive directions, thereby varying the stereo overlap from 100 to 0. The value of α is calculated as the total area in the camera plane that is viewed by both cameras divided by the combined image plane area of both cameras. With the variation of α , values of R_{C2} , and t_{C2} are determined and compared against ground truth values.

For this analysis, K_C , R_{C1} , t_{C1} and $(\underline{u}_{c1}, \underline{u}_{c2})$ are known, while R_{C2} , t_{C2} are estimated. To identify the extrinsic parameters, the essential matrix (E) [25] is first computed using the normalized point correspondences [25] $(\hat{\underline{u}}_{c1}, \hat{\underline{u}}_{c2})$, as shown in Eq. 4.6. This matrix is then decomposed using Singular Value Decomposition (SVD) [71], allowing for the extraction of R_{C2} and t_{C2} .

$$\hat{\underline{u}}_{c2}^T \cdot E \cdot \hat{\underline{u}}_{c1} = 0 \quad (4.6)$$

For each value of stereo overlap, the output values are computed 15 times, and the resulting errors are averaged to reduce fluctuations caused because of the gaussian noise.

4.2.2 Impact of Camera Overlap on Projector Calibration Accuracy

This case study utilizes two cameras and one projector to evaluate the precision of projector calibration (K_P , R_P and t_P) as a function of stereo overlap (β) between the cameras and projector. The value of β is defined as the area of the projector region visible to both cameras, divided by the total area of the projector in the projector image plane. Initially the cameras are placed such that β is equal to 100%. Later, t_{C2} is iteratively increased such that camera 2 moved away from the camera 1 in both the negative and positive directions, thereby varying β from 0 to 100. As β is varied, the values of K_P , R_P and t_P are determined and compared with ground truth values.

For this analysis, the camera calibration (K_C , R_C , t_C) and $(\underline{u}_c, \underline{u}_p)$ are assumed to be known, while the projector calibration (K_P , R_P and t_P) needs to be estimated. Initially, triangulation [10, 25], as described in Section 2.2.3, is performed with cameras to identify the world points ($\underline{u}_w$) in the stereo region of the cameras and is seen by the projector.

Next, using the identified $(\underline{u}_w)$ and the known $\underline{u}_p$, the projector's projection matrix P_P is determined which is later decomposed to obtain the projector parameters. For each stereo overlap, P_P is calculated 15 times to avoid the fluctuations caused due to Gaussian noise. The setup for Study 1 and 2 is shown in Fig. 4.2a.

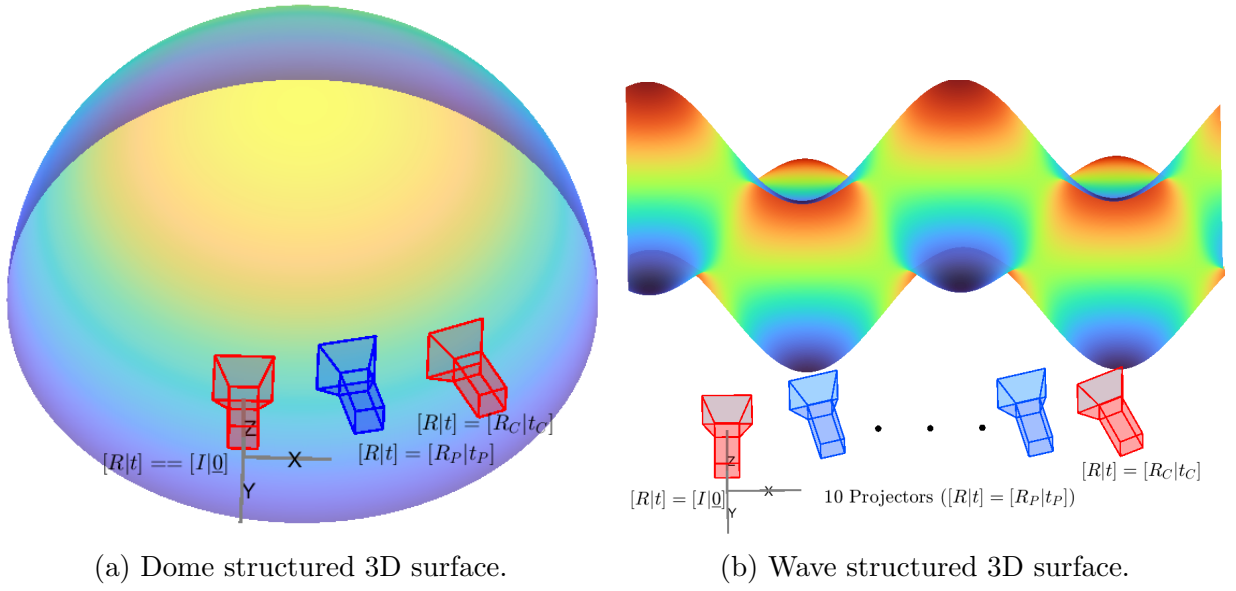


Figure 4.2: Synthetic data setup for different case studies with red representing cameras and blue representing projectors, all projecting onto a 3D surface. The devices are oriented such that their optical axes point toward the surface, and each device's field of view (FoV) covers a region centered along its viewing direction.

4.2.3 Calibration Accuracy in Sequentially Overlapping Projectors

This case study is performed to understand the flow of sequential error in the projector parameters. It involves the use of two cameras and ten projectors, where each projector is placed adjacent to the next with a percentage overlap of γ . Additionally, the cameras are positioned such that both cameras fully overlap with Projector 1. Since this case study is specifically designed to investigate the sequential error propagation and not the overlap between cameras and projectors, it is important to isolate the effect of stereo cameras. Therefore, both cameras are configured to fully view Projector 1. However, only one of the cameras observes the remaining projectors.

To analyze the accuracy of the calibrated projector parameters (K_P , R_P , and t_P) sequentially across all projectors, the variable γ is increased from 5% to 95%. The upper limit of 95% is imposed to prevent physical collision of projectors within the setup, while the lower limit of 5% ensures that some minimum amount of overlap is maintained between devices. The synthetic data setup used for this analysis is illustrated in Figure 4.2b.

As γ increases, the calibration values K_P , R_P , and t_P are determined for each projector. The estimation of projector parameters is performed in the same manner as described in Section 4.2.2. This process begins with the estimation of parameters for the projector that fully overlaps with the stereo cameras and continues sequentially toward the final projector. This approach enables a systematic assessment of the variation and accumulation of calibration errors in the projector parameters.

4.3 Experimental Procedure

To analyze the performance and accuracy of the calibration results, the reprojection error (E_R), as defined in Section 2.2.4, is used as the evaluation metric. Since the stereo overlap tests are conducted entirely using synthetic data, the true parameters of both the projector and the camera are known. The computation of E_R proceeds as follows:

- The actual image coordinates are computed using the ground truth parameters.
- The estimated image coordinates are then obtained using the newly estimated parameters.
- The E_R is calculated as the Euclidean distance between the true and estimated image coordinates in the image plane of the corresponding optical device.

The image plane varies depending on the device for which the error is being evaluated: if E_R is computed for a projector, then $\underline{u}_p$ is used; if it is computed for a camera, then $\underline{u}_c$ is used. Additionally for each value of stereo overlap (denoted by α , β , and γ), the output values are computed 15 times, and the resulting errors are averaged to reduce fluctuations caused because of the Gaussian noise.

4.4 Results

To understand the variation of E_R with changes in stereo overlap (denoted by α , β , and γ), experiments were conducted using three different noise levels (σ values corresponding to Gaussian noise): 0.25, 0.5, and 1. The results indicate that E_R increases as stereo overlap decreases. In the case of camera calibration (see Figure 4.3), a sharp rise in error is observed when stereo overlap falls below 15%. Additionally, a spike in error is also noted near 100% overlap, where the cameras are positioned extremely close to one another. This spike occurs due to numerical instability, which happens when the cameras are placed too close together. In such cases, their views become too similar, making triangulation difficult and leading to unreliable calibration results. It is important to note that E_R is calculated close to, but not exactly at, zero camera separation because at extremely small baselines, the calibration problem becomes ill-posed and numerically unstable.

In the case of projector calibration (see Figure 4.4), a similar pattern is observed. The reprojection error increases as stereo overlap decreases. In this scenario, the dotted segments in the plot represent instances where the E_R exceeds 3 pixels. While an E_R of less than 1 pixel is widely accepted as a standard threshold for accurate calibration in projector systems [45, 74], errors exceeding 3 pixels indicate significant misalignment and are unreliable for practical applications such as projection mapping.

A comparison of E_R across different noise levels ($\sigma = 0.25, 0.5, 0.75, 1.0$) highlights the sensitivity of calibration accuracy to increasing noise. In particular, the ratio of E_R between $\sigma = 0.5$ and $\sigma = 0.25$, and between $\sigma = 1.0$ and $\sigma = 0.5$, shows that doubling the noise level leads to an almost fourfold increase in E_R within the linear region. This amplification of error suggests that even moderate increases in noise can cause large degradations in calibration accuracy, highlighting the importance of minimizing noise during data acquisition.

As discussed in Section 4.2.3, Case Study 3 aims to analyze the sequential propagation of E_R across projectors. To investigate this behavior, three different values of noise σ are considered, namely 0.25, 0.5, and 1. For each noise level, γ between projectors is varied from 5% to 95%, and E_R is computed for all ten projectors.

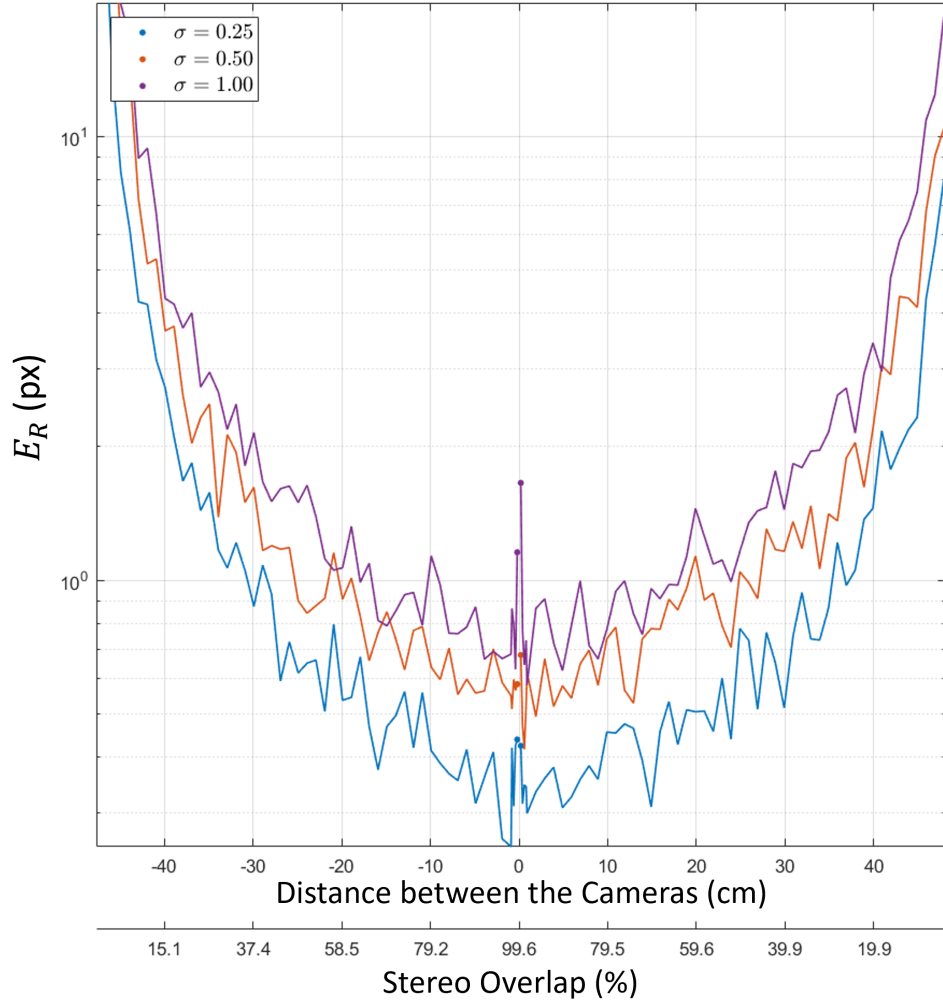


Figure 4.3: E_R vs α for Case Study 1. As stereo overlap decreases, E_R generally increases. At nearly 100% overlap (i.e., when cameras are placed very close), numerical instability causes a spike in error.

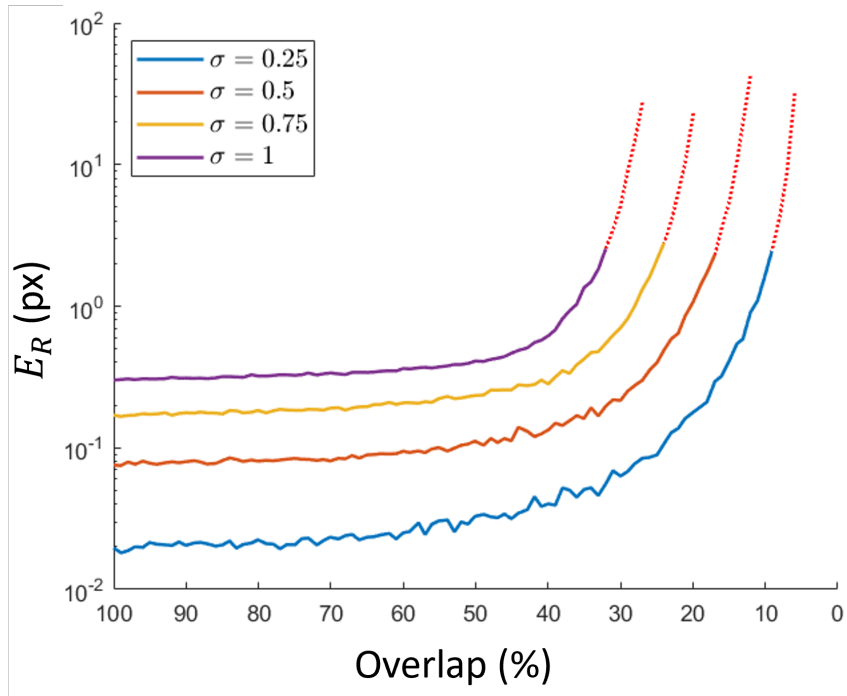
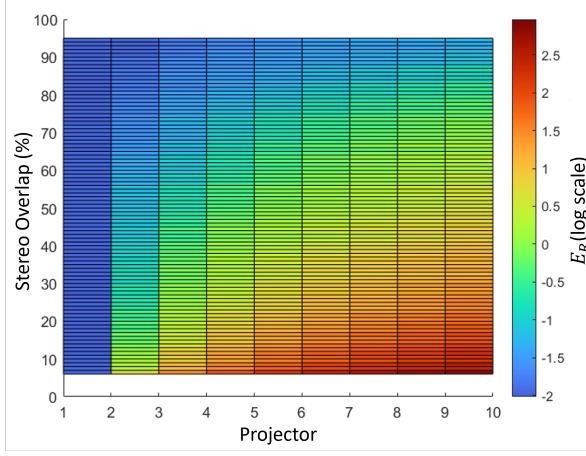
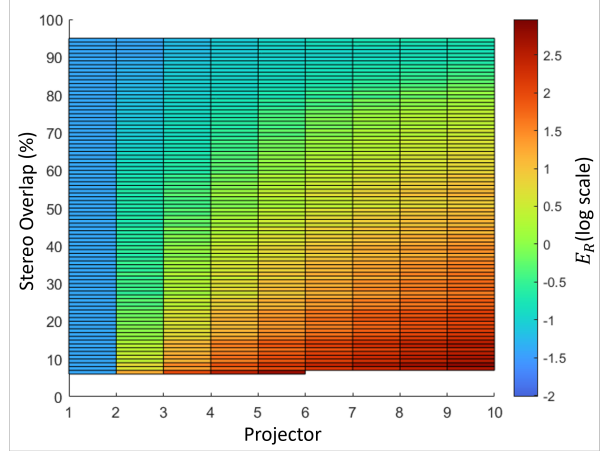


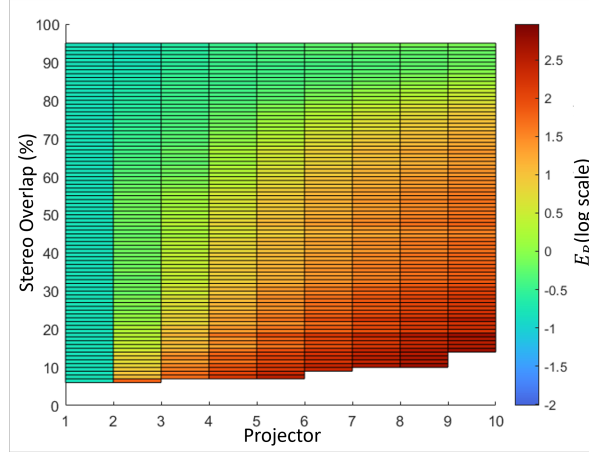
Figure 4.4: E_R vs β for Case Study 2. The error increased almost fourfold when the noise level was doubled, highlighting the sensitivity of calibration to noise. Red dotted lines marking E_R exceeding 3px, where calibration accuracy becomes unreliable.



(a) $\sigma = 0.25$

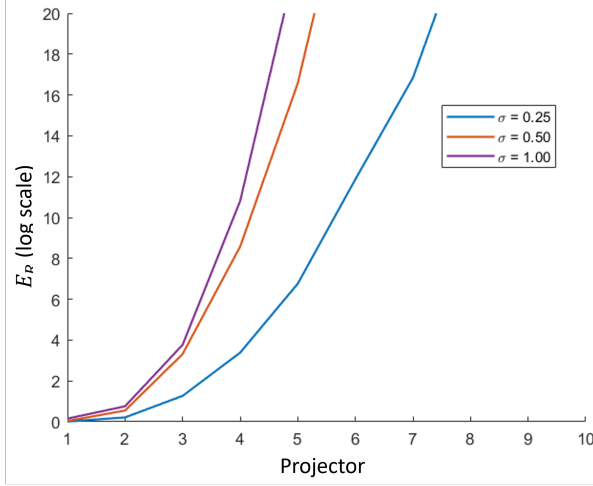


(b) $\sigma = 0.50$

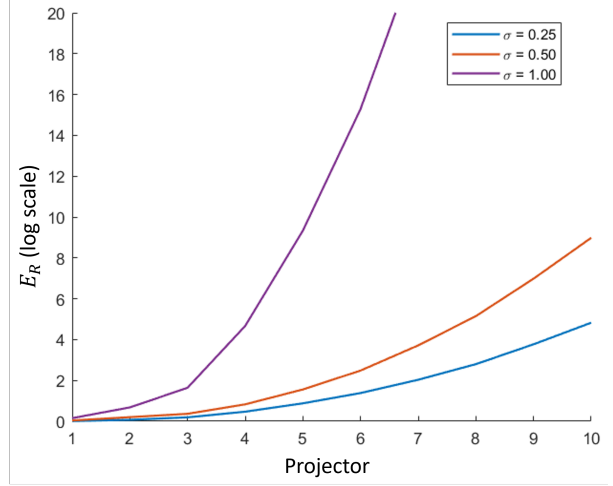


(c) $\sigma = 1.00$

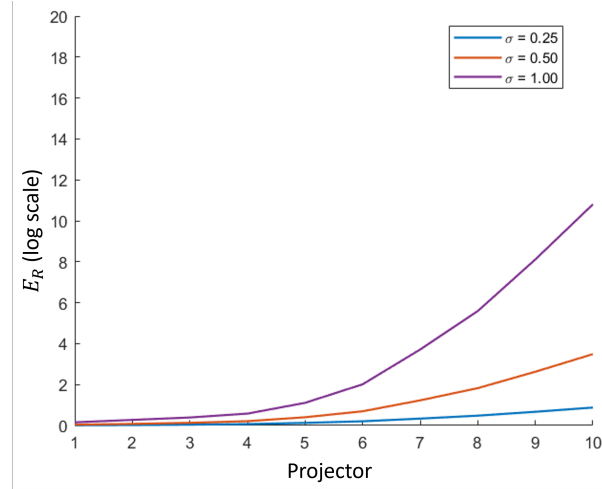
Figure 4.5: E_R as a function of varying γ and projector indices for different values of σ . As σ increases, E_R becomes more sensitive to both overlap and projector depth. For higher values of σ (subfigures b and c), several projectors are left uncalibrated due to insufficient correspondences, which is reflected in the blank regions of the plots.



(a) 25% Stereo Overlap.



(b) 50% Stereo Overlap.



(c) 75% Stereo Overlap.

Figure 4.6: E_R plotted as a function of projector index, for different σ , under fixed γ between projectors. As the sequence progresses, E_R increases more rapidly, particularly at higher noise levels. This trend suggests that calibration errors from earlier projectors influence the accuracy of later ones, leading to a gradual buildup of error across the sequence.

As illustrated in Figure 4.5, E_R increases with projector index even when stereo overlap is fixed, indicating that calibration accuracy worsens as the sequence progresses. The projector index refers to the position of a projector in the sequence, starting from Projector 1 (fully visible to both cameras) and increasing for each subsequent projector that overlaps only with its immediate neighbour. This increase in error indicates that calibration errors in the initial projectors carry over to each subsequent projector in the sequence, reducing their accuracy as well.

Additionally, for any fixed configuration, reducing stereo overlap further increases E_R , highlighting its role in maintaining consistency. This means that lower overlap between adjacent projectors leads to less reliable calibration results. This trend is even more visible in Figures 4.5b and 4.5c, where higher noise levels ($\sigma = 0.5$ and 1.0) leads to a noticeable rise in E_R values and cause some projectors to fail calibration entirely due to insufficient reliable correspondences.

Figure 4.6 presents a detailed analysis of how E_R grows as calibration proceeds sequentially through the projector array. At higher stereo overlaps, a large number of high-quality correspondences are available, resulting in a nearly linear increase in error. However, as the stereo overlap decreases, the error growth becomes more rapid and nonlinear.

These observations highlight how calibration accuracy is influenced by the type and extent of overlap between optical devices. By thoroughly understanding the propagation of calibration error through different overlap scenarios, it becomes possible to optimize the arrangement of projectors in a projection mapping system and apply appropriate calibration strategies that account for error accumulation.

To conclude, the results presented in this chapter demonstrate that calibration accuracy is strongly influenced by the degree of stereo overlap and the level of image noise. Reduced overlap and increased noise both lead to substantial increases in reprojection error, with the effect compounding across sequentially calibrated projectors. Building on these insights, the next chapter focuses on calibrating projectors using a minimal number of cameras and evaluates the performance of the proposed strategy introduced in Chapter 3 using both synthetic and real-world data.

Chapter 5

Multi-Step Calibration of Projectors

Section 2.4.3 of this thesis described the various techniques employed for projector calibration. Both manual and auto-calibration approaches have been explored by researchers and engineers. Due to the significant human intervention required in manual calibration, auto-calibration methods have gained increased attention and adoption.

Subsequently, Section 3.2 presented the challenges associated with auto-calibration. It was shown that existing algorithms often depend on either a large number of cameras or a minimum overlap between projectors and stereo camera surfaces to ensure accurate calibration.

To address these limitations, a multi-step optimization strategy has been proposed to enhance the accuracy of projector parameter calibration. This chapter outlines the proposed approach in detail and presents the results obtained from its application.

5.1 Methodology

The auto-calibration of projectors involves several preprocessing stages, ultimately leading to the estimation of accurate calibration parameters. The methodology is developed under the following assumptions:

1. The intrinsic parameters of the cameras are known. This helps isolate the effect of stereo overlap on projector calibration by removing the influence of camera intrinsics.

2. The scale (s), defined as the distance between the cameras, is known. This provides a real-world unit of measurement, allowing the calibration to be interpreted at a meaningful scale.

These assumptions facilitate the effective calibration of the system components.

5.1.1 Point Correspondence Collection

The Gray Code structured light technique is utilized to extract point correspondences between devices. A comprehensive explanation of this technique is provided in Section 2.4.2. In this framework, two cameras capture structured light images projected by the projectors. From these images, decoding is performed to generate the following data:

1. Correspondence between image points in two camera views: $(\underline{u}_{c_i}, \underline{u}_{c_j})$
2. Correspondence between camera and projector image points: $(\underline{u}_c, \underline{u}_p)$
3. Correspondence between image points in two projector views: $(\underline{u}_{p_i}, \underline{u}_{p_j})$

Here, $\underline{u}_c$ and $\underline{u}_p$ denote the image coordinates in the camera and projector image planes, respectively, for optical devices i and j .

Each correspondence type contributes uniquely to the accuracy of projector calibration. Initially, $(\underline{u}_{c_i}, \underline{u}_{c_j})$ is employed for camera calibration. Since the accuracy of projector calibration is highly dependent on the precision of camera parameters, this step is crucial. The correspondences $(\underline{u}_c, \underline{u}_p)$ are subsequently used to estimate initial projector parameters and are further refined during the first stage of optimization. Finally, $(\underline{u}_{p_i}, \underline{u}_{p_j})$ is incorporated into the second optimization phase, as it reflects the known overlap between projector image planes.

5.1.2 Camera Calibration

With known intrinsic parameters, the calibration of cameras focuses on estimating their extrinsic parameters. One camera is designated as the origin reference frame, and its extrinsic parameters are set as:

$$[R_{C1} \mid t_{C1}] = [I \quad 0] \quad (5.1)$$

This assumption simplifies the problem to the estimation of R_{C2} and t_{C2} for the second camera.

Using the known intrinsic matrix K_C and the correspondences $(\underline{u}_{c_i}, \underline{u}_{c_j})$, the essential matrix E is computed. As described in Section 2.2.2, this matrix is decomposed into a rotation matrix R_{C2} and a unit translation vector $\hat{t}_{C2}$. The actual translation vector t_{C2} is then recovered by scaling $\hat{t}_{C2}$ with the known distance s .

5.1.3 Initial Estimate of Projector Parameters with Optimization 1

An initial estimate plays a crucial role in any optimization algorithm [32]. Since real-world optimization landscapes often contain numerous local minima, a good initial estimate brings the solution closer to the global minimum and avoids convergence to local minima [32].

Initial Estimate

To compute an initial estimate of the projector parameters, the correspondence between image points and world points $(\underline{u}_p, \underline{u}_w)$ is utilized. As this correspondence is not directly available, it is derived from existing known variables.

First, triangulation is applied using camera parameters and the camera correspondences $(\underline{u}_{c_i}, \underline{u}_{c_j})$, resulting in a mapping between camera image points and world points $(\underline{u}_c, \underline{u}_w)$. Additionally, structured light decoding provides the correspondence $(\underline{u}_c, \underline{u}_p)$. Combining these correspondences yields the required $(\underline{u}_p, \underline{u}_w)$ mapping.

Using the obtained $(\underline{u}_p, \underline{u}_w)$ correspondences, the Direct Linear Transformation (DLT) algorithm [33], as described in Section 2.3.1, is employed to estimate the intrinsic matrix K_P , rotation matrix R_P , and translation vector t_P . The DLT method constructs a homogeneous linear system from image-world correspondences and solves it via generalized eigenvalue decomposition. At least six point correspondences are required for each projector to generate an initial estimate. Each projector is calibrated independently rather than simultaneously.

Optimization 1

Following the initial estimation, a robust optimization technique is applied to refine each projector's parameters individually. This thesis uses reprojection error, discussed in detail

in Section 2.2.4, as the objective function and employs Sequential Quadratic Programming (SQP) [22] as the optimization algorithm.

SQP is particularly suitable for constrained, nonlinear optimization, offering a balance between speed and accuracy while handling constraints effectively [22]. The optimization is bound-constrained, with the lower bounds for focal length and principal point set to 0.1. This avoids unrealistic parameter values, as zero is not a plausible value for either focal length or principal point.

Combined Iterative Approach

Instead of applying DLT across all projectors followed by Optimization 1, a combined step is performed: DLT is followed immediately by Optimization 1 for each projector before proceeding to the next.

This sequence helps mitigate poor estimates in projectors that have little or no overlap with the stereo camera pair. Applying Optimization 1 after DLT improves the $(\underline{u}_p, \underline{u}_w)$ correspondences, yielding a more accurate initial estimate for subsequent steps.

The calibration order is determined by the degree of overlap between each projector and the stereo camera setup. Projectors are calibrated in descending order of this overlap, beginning with the projector that has the highest overlap.

Reduction of Runtime

As noted in Section 1.2, the calibration process aims not only for accuracy but also for efficiency. Real-time applications demand faster algorithms. Structured light techniques typically yield around 100,000 point correspondences per projector, making it computationally expensive to process all of them.

While six correspondences suffice for DLT, relying solely on them often results in inaccurate parameter estimation due to noise. In this work, 100 $(\underline{u}_c, \underline{u}_p)$ correspondences are selected per projector to strike a balance between computational speed and calibration accuracy, thereby reducing the overall runtime.

To ensure these selected points are evenly distributed across the projector plane, random sampling is avoided. Instead, correspondences are selected based on uniform spacing to ensure broad surface coverage. Experiments comparing different sampling strategies in terms of accuracy and runtime are presented in Section 5.3.3.

Outlier Removal

Structured light decoding often introduces erroneous correspondences. Common causes include:

- Overexposed or underexposed projected patterns, leading to decoding errors.
- Surface reflectivity or texture, which can cause scattering and code loss.
- Ambient light interference, resulting in false detections.

To mitigate these effects, two strategies are employed during this optimization stage:

1. **Nearest Neighbour Filtering:** Given that structured light yields dense point correspondences, a threshold radius r_o is defined for each triangulated world coordinate X_i . If no neighbouring point lies within this radius, X_i is considered an outlier and removed. The parameter r_o is a tunable hyperparameter based on the projection setup.
2. **Reprojection Error Filtering:** Using the parameters estimated from the combined approach, the reprojection error for each $(\underline{u}_p, \underline{u}_w)$ pair is computed. These errors are clamped between 1.0 and 100.0 to stabilize extremes. A dynamic threshold $t_{adapt} = \text{mean} + 2 \times \text{stddev}$ is applied to remove high-error correspondences. The remaining points are refined through Bundle Adjustment (BA) [73], after which the reprojection error filter is reapplied. This iterative process continues until no more points are eliminated and the average reprojection error stabilizes, showing less than 10% change between iterations. This method is inspired by techniques developed by Disney Research [69].

These filtering strategies, although not directly improving the accuracy of the current projector’s calibration, significantly enhance the quality of input for subsequent projector calibration stages. Given that later projectors often lack sufficient overlap with the stereo camera setup, such filtering contributes to improved overall system accuracy.

5.1.4 Optimization 2

As illustrated in Figure 5.1, Optimization 1 alone is insufficient, as using only reprojection error creates visible gaps between overlapping projector content. This limitation

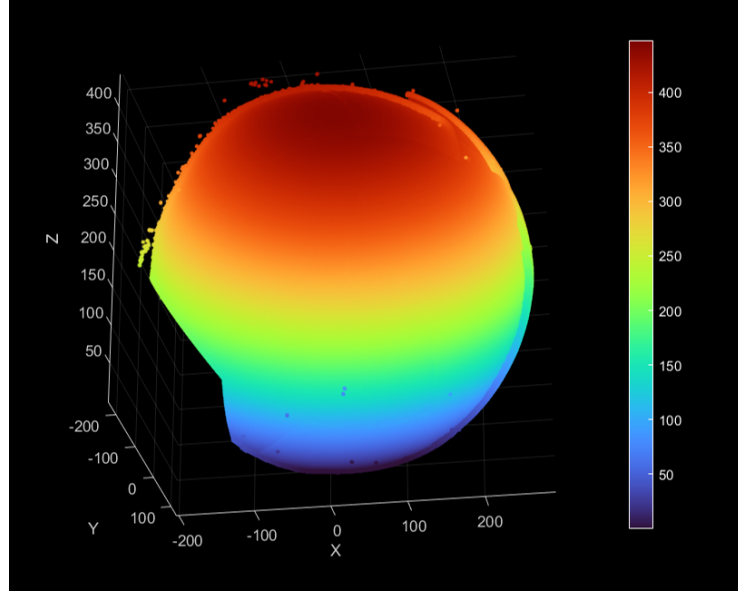


Figure 5.1: Point cloud reconstruction using only reprojection error. This approach creates visible misalignment between the projected content after Optimization 1.

necessitates an additional optimization step to further refine the estimation of projector parameters.

This thesis introduces and implements a novel objective function designed to minimize the distance between overlapping regions in order to achieve a more accurate calibration of projector parameters. In contrast to Optimization 1, which calibrates each projector independently, this second optimization step simultaneously optimizes the parameters of all projectors to be mutually consistent. A detailed description of this algorithm is provided in Section 3.3.

The distance-based objective function accepts the output of Optimization 1 as its initial parameter estimates and employs the interior-point algorithm for optimization. The choice of interior-point over Sequential Quadratic Programming (SQP) stems from the computational demands of the problem. As this stage involves a significantly larger number of parameters, the interior-point method [8] is more suitable due to its scalability and stability in large-scale optimization scenarios. In contrast, SQP becomes computationally expensive and may lose stability as the problem size increases [46].

During this optimization, the overlapping projector correspondences $(\underline{u}_{p_i}, \underline{u}_{p_j})$ are utilized. The optimization process minimizes the distances between these overlapping points

to refine the projector parameters.

In addition to the overlapping points between projectors, the dataset for Optimization 2 also includes a small number of points that overlap with the camera stereo surface. In other words, some projector overlap regions also intersect with areas observed by both cameras. In these regions, the 3D points reconstructed from all projector overlaps and the camera stereo overlap should ideally be identical, with zero distance between them, as they represent the same physical point in space. During Optimization 2, the algorithm minimizes the total distance between all overlapping points. However, since the overlap between the projector regions and the camera stereo surface is relatively limited, the overall process is generally treated as a minimization over projector correspondences. Nevertheless, the inclusion of these camera stereo points introduces a constraint that helps guide the optimization and prevents unrealistic or unstable transformations.

Outlier Removal

In contrast to Optimization 1, outlier removal in this step is performed prior to the optimization process. This is because, in Optimization 1, a reliable estimate is not available initially, and outlier removal would be ineffective without an initial calibration. However, by the time Optimization 2 begins, a reasonably good estimate has already been obtained from the previous step, making early outlier removal both feasible and beneficial for refining the calibration across all projectors. Using the initial parameter estimates obtained from Optimization 1, the distances between overlapping points are computed.

The Interquartile Range (IQR) method [53] is then applied to determine the upper bound for acceptable distances. This procedure is carried out across all projectors to derive a global upper bound (u_o). Any point with a distance error E_{OD} exceeding u_o is excluded from the optimization.

Reduction of Runtime

In Optimization 1, a fixed number of point correspondences per projector were selected, as the entire projector plane was uniformly covered, making sampling straightforward. However, in the case of overlapping regions between projectors, the extent of overlap varies significantly, ranging from very small to very large, making it challenging to determine a fixed number of points that fairly represent all overlaps.

Selecting a fixed number of correspondences could result in biased optimization, where a disproportionate number of points represent smaller overlaps between projectors, while

larger projector overlaps receive insufficient coverage. To address this, a proportional sampling strategy is adopted. In this thesis, a proportion of 0.1% of the available overlapping points $(\underline{u}_{p_i}, \underline{u}_{p_j})$ is uniformly sampled across the surface for each projector pair.

This approach ensures that all overlaps, regardless of their size, are uniformly represented during optimization, enabling a balanced and scalable calibration process. A detailed experimental evaluation of the sampling proportion, including its impact on calibration accuracy and runtime performance, is presented in Section 5.3.3.

5.2 Experiment Data and Setup

For a robust analysis of the multi-step optimization approach, two types of data are used: synthetic data and real-world data. The purpose of using synthetic data is to introduce controlled noise levels and compare the true calibration results across all projectors, leveraging the availability of ground truth parameters. On the other hand, real-world testing is conducted to validate the practical applicability of the proposed model. Unlike the simulation environment, obtaining ground truth projector parameters in real-world scenarios is challenging. Therefore, in the analysis of real-world data, performance is evaluated using metrics such as runtime and the distance between overlapping content (E_{OD}).

The synthetic and real data used for the experimentation are shown in Figure 5.2. The simulation data was designed to closely resemble the real-world setup, as the real data includes eleven projectors with multiple overlapping regions, which are central to the optimization process. Therefore, the overlap configurations in the synthetic data were aligned as closely as possible with those of the real setup. Additionally, the synthetic surface is not purely hemispherical; a flat region covering approximately 10% of the top surface was introduced to better replicate the actual structure. Gaussian noise σ was also added to the point correspondences in the synthetic data to simulate real-world measurement variations.

In both the synthetic and real setups, projectors and cameras are positioned to face the dome surface, with their optical axes directed toward it. The field of view (FoV) of each device enables partial overlap with neighboring devices to support calibration. Specifically, projectors oriented toward the center of the dome have over 20% stereo overlap with the camera pair, while the remaining projectors exhibit minimal or no overlap, typically less than 20%. This configuration reflects practical challenges in large-scale setups and is mirrored in the synthetic data to assess the impact of limited stereo visibility on calibration accuracy.

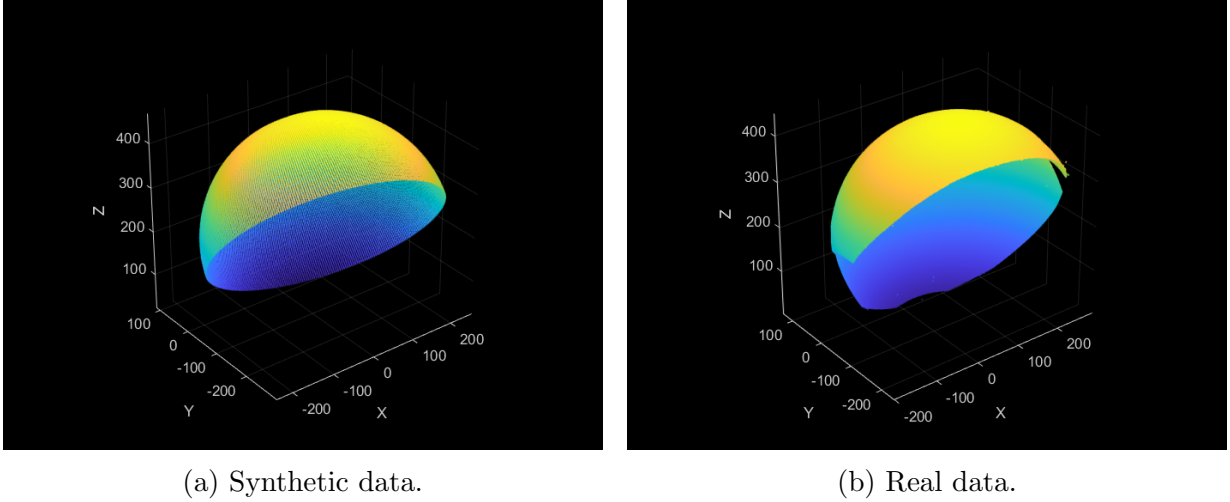


Figure 5.2: Synthetic and real point clouds of the dome. The synthetic data was manually created in MATLAB, while the real data was obtained via triangulation of structured light correspondences. Voids in the real data are caused by regions that were not visible to the cameras during capture.

While both dome-shaped and wave-shaped synthetic surfaces were used in earlier analysis, only the dome surface is considered here, as it closely matches the real test setup. This allows for more direct comparison and evaluation of the proposed method under similar conditions.

In the real dataset, some regions appear void or missing. This occurs because those areas were not visible to any of the cameras during capture. Triangulation to generate the full point cloud requires that each region be visible from at least one camera. In contrast, the synthetic data does not exhibit such missing regions, as it was generated entirely within MATLAB and does not rely on image capture or visibility constraints.

To obtain the point correspondences between the cameras, as well as between the cameras and projectors for the synthetic dome, the same procedure described in Section 4.1 is followed. Using this method, all necessary point correspondences are generated and subsequently used for the calibration and optimization of projector parameters.

The algorithm was implemented on an Alienware m15 R7 laptop running Microsoft Windows (x64-based PC) with a 12th Gen Intel Core i7-12700H processor (2.30 GHz, 14 cores/20 logical processors) and 32 GB of RAM.

σ	Optimization	Median Error (px)	Mean Error (px)
0.25	Opt 1	0.64	0.68
	Opt 2	0.61	0.60
0.5	Opt 1	0.96	1.15
	Opt 2	0.86	0.96
0.75	Opt 1	1.82	2.08
	Opt 2	1.58	1.80

Table 5.1: E_R for Optimizations 1 and 2 across different values of σ on the synthetic dataset.

5.3 Experimentation Results

5.3.1 Synthetic data

The E_R is used as the primary metric for evaluating the calibration accuracy. Since the synthetic data provides access to ground truth parameters, the same error computation procedure described in Section 4.3 is applied to the synthetic dome setup. The synthetic dataset consists of 11 projectors. For each value of σ , the error is computed 15 times and the results are averaged across all projectors. From these averaged values, the mean and median E_R are calculated to summarize performance. Both mean and median errors are calculated to reflect overall performance and robustness. The mean captures the average reprojection error, while the median helps assess consistency by reducing the influence of outliers.

Three different values of σ are considered in this analysis: 0.25, 0.5, and 0.75 pixels. These values represent increasing levels of noise added to the image data and help evaluate how well the calibration process performs under different conditions. Higher values were not considered, as they would introduce extreme distortions which are unlikely to occur in practice, while lower values would not provide enough variation to reveal meaningful trends.

For each noise level, both steps of the multi-step optimization procedure are applied. First, Optimization 1 is performed to obtain initial estimates, followed by Optimization 2, which refines the parameters using the distance-based objective function. Table 5.1 summarizes the E_R values across different noise levels and optimization stages.

The results clearly indicate that for all values of σ , the proposed multi-step optimization consistently reduces the overall E_R . Both the mean and median reprojection errors decrease

Optimization	Median Error (cm)	Mean Error (cm)
Opt 1	0.70	1.49
Opt 2	0.50	0.70

Table 5.2: E_{OD} for Optimizations 1 and 2 on the actual dataset with 100 points in Optimization 1 and 0.1% of overlapping points in Optimization 2.

following Optimization 2. The presence of projectors with minimal or no overlap with the stereo camera setup contributes to calibration difficulties, leading to higher mean errors compared to the median. This discrepancy highlights the skewed nature of the error distribution across projectors.

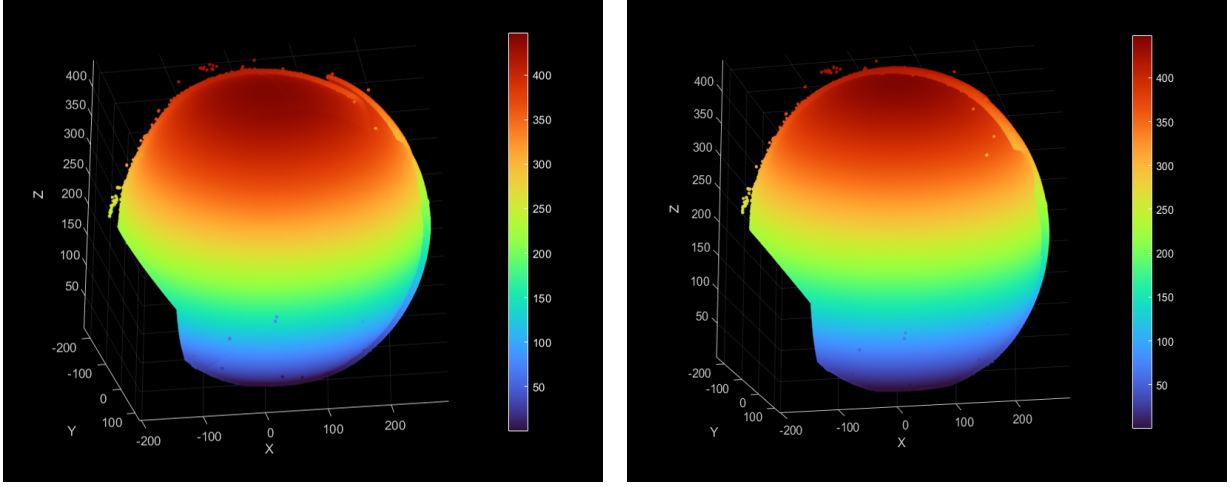
5.3.2 Real data

Unlike the simulation dataset, the real-world setup lacks access to ground truth projector parameters, making direct comparison against true values infeasible. Therefore, performance evaluation is conducted using the distance-based error metric E_{OD} , as defined in Equation 3.2.

Table 5.2 compares the E_{OD} values for Optimizations 1 and 2 on the real dataset. To reduce runtime, a fixed number of points were selected from the overall point correspondences, as detailed in Sections 5.1.3 and 5.1.4. Specifically, 100 points were used in Optimization 1, and 0.1% of the overlapping data points were used in Optimization 2. The reason for this choice of parameters will be explained in detail in Section 5.3.3. Even with this limited point set, the algorithm remained effective, resulting in a noticeable reduction in both the mean and median E_{OD} . Similar to the synthetic case, the presence of skewed error distribution results in the mean error being greater than the median.

For visual validation, Figure 5.3 illustrates the point cloud reconstructions before and after Optimization 2. The colormap encodes the depth of each point from the origin. The left image reveals visible misalignment between overlapping point clouds, particularly in the top region, which is effectively corrected in the right image after Optimization 2. It is important to note that although the visible gap appears significant, it is localized to regions between projectors with minimal overlap with the stereo camera surface. Consequently, the overall mean and median E_{OD} remain comparatively low despite the apparent visual error.

With this choice of limited points, the runtime for Optimization 1 and Optimization 2 was approximately 2 minutes and 5 minutes 40 seconds, respectively. Thus, the total



(a) Before Optimization 2.

(b) After Optimization 2.

Figure 5.3: Comparison of point cloud reconstructions. Optimization 2 effectively reduces the misalignment present in the upper region after Optimization 1.

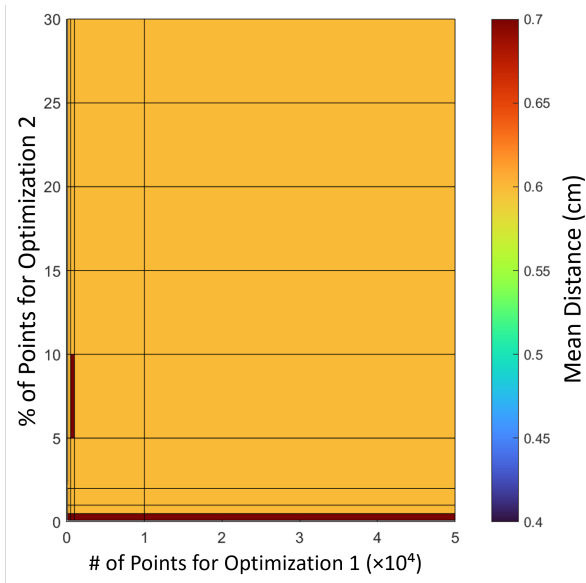
runtime for both optimization steps was around 7 minutes and 40 seconds.

5.3.3 Limited Points

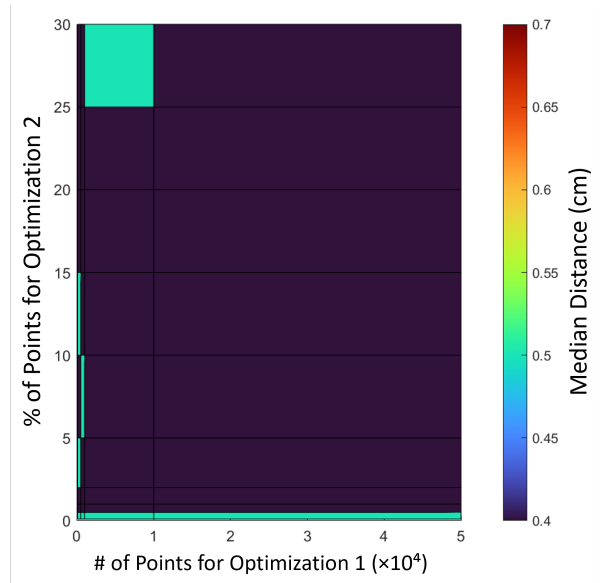
In real-world applications, achieving accurate calibration is important, but it is equally essential to minimize runtime. Since structured light images typically provide close to one hundred thousand pixel correspondences per projector, a series of experiments was conducted to analyze how calibration error and runtime vary with different selections of point correspondences.

Figure 5.4 presents the mean and median distance errors, denoted as E_{OD} , between overlapping projector regions. For Optimization 1, the number of points selected per projector included 100, 500, 1,000, 10,000, and 50,000. In Optimization 2, different proportions of overlapping points were used, ranging from 0.1% to 30%. The figure demonstrates that the distance error remains largely consistent even when the number of point correspondences is significantly reduced. Only at the lowest values of point count and proportion does the error begin to rise noticeably.

A similar trend is observed in Figure 5.5, which shows the runtime results for both optimization stages. The runtime increases approximately linearly with the number of

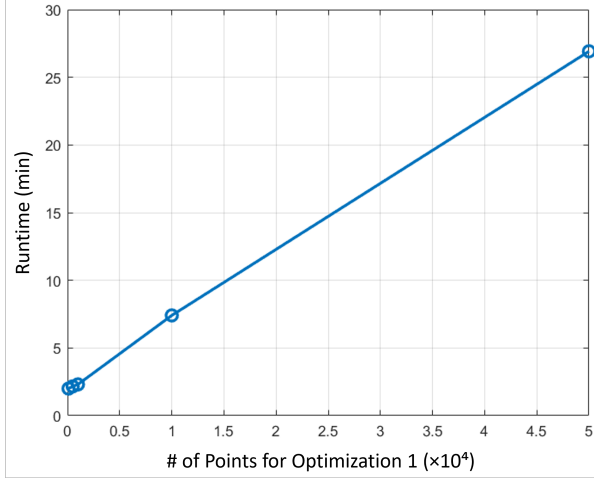


(a) Mean distance error.

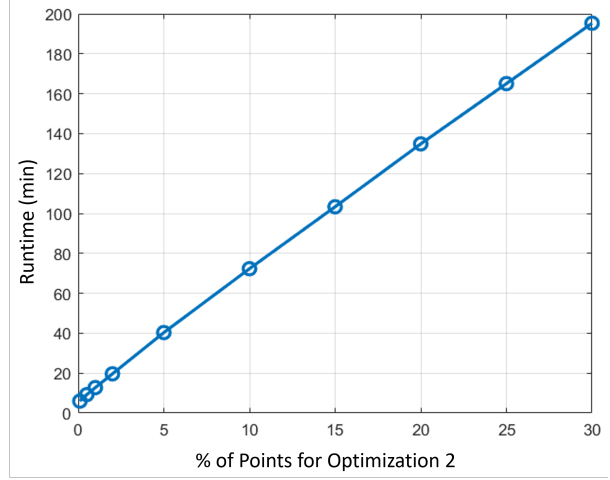


(b) Median distance error.

Figure 5.4: Comparison of E_{OD} for varying point counts in Optimization 1 and 2. The error remains largely stable despite significant reductions in point count, with only slight increases observed at extreme values.



(a) Runtime variation for Optimization 1.



(b) Runtime variation for Optimization 2.

Figure 5.5: Comparison of runtime for varying point counts in Optimization 1 and 2. The runtime increases approximately linearly with the number of points, highlighting the computational cost associated with larger input sizes.

points, indicating the computational cost associated with larger inputs. This confirms the importance of limiting the number of points to achieve faster execution.

Because the error remains nearly unchanged while the runtime decreases substantially, it is advantageous to use a reduced set of point correspondences after outlier removal. This leads to quicker and more efficient calibration without sacrificing accuracy.

When only 0.1% of the overlapping data was used in Optimization 2, the dome dataset still maintained at least three thousand overlapping point correspondences between projectors. However, this value acts as a hyperparameter that may vary depending on the density of the structured light data. For dense datasets, the current selection is sufficient. In contrast, for sparse datasets, selecting a higher proportion of overlapping points in Optimization 2 is recommended to preserve calibration quality.

Overall, this chapter displayed the effectiveness of the proposed multi-step calibration strategy using both synthetic and real-world datasets. The results confirmed that the second optimization stage improves calibration accuracy, even when using a limited set of point correspondences. Across various noise levels and data types, both mean and median errors consistently decreased after Optimization 2, validating the robustness of the distance-based objective. Furthermore, the analysis on reduced point sets revealed that

substantial reductions in runtime can be achieved without compromising accuracy, making the approach practical for real-world projection systems.

Chapter 6

Conclusion

This thesis presents a multi step optimization strategy incorporating a novel objective function to accurately calibrate projectors. Prior to the calibration of projectors, point correspondences between devices are obtained using structured light. These correspondences are then used to calibrate the extrinsic parameters of the cameras, assuming the intrinsic parameters are already known. Following this, the Direct Linear Transformation algorithm is applied to obtain an initial estimate of the projector parameters in an iterative manner. This estimate is subsequently refined through an optimization process using reprojection error as the objective function. In the final stage, the results are further improved by minimizing the distance between overlapping projector content using a distance-based objective function.

The results presented in Table 5.1 and Table 5.2 highlight the improvements achieved before and after applying the distance-based optimization. These tables clearly demonstrate the effectiveness and impact of the proposed method. Additionally, Figure 5.3 offers a visual comparison where a noticeable gap observed after the initial optimization step is corrected following the application of the final optimization. These outcomes confirm the effectiveness of the proposed multi step optimization in reducing the dependence on the degree of overlap between the projector and the stereo camera surface.

In addition to developing the optimization strategy, this thesis conducts a thorough analysis of the various types of overlaps that occur between optical devices. This includes the overlap between camera views, between stereo cameras and projectors, and between multiple projectors. Gaussian noise is introduced into the data to evaluate the robustness of the calibration under different conditions. Several notable observations emerge from this analysis. In the scenario involving stereo camera and projector overlap, it is found

that doubling the noise level causes the reprojection error to increase nearly four times. Furthermore, during the investigation of error propagation across sequentially calibrated projectors, it is observed that decreasing stereo overlap leads to a transition from gradual to more rapid error increase.

These findings contribute to a deeper understanding of the relationship between overlaps and calibration accuracy. This knowledge is valuable when designing projection mapping systems or similar configurations that require precise placement of projectors. A clear understanding of how different types of overlaps influence calibration performance allows for more informed decisions regarding device arrangement, ultimately leading to more effective usage and improved outcomes from the optimization process.

6.1 Future Directions

Several improvements can be made to enhance the overall effectiveness and efficiency of the proposed approach.

At present, the method relies on structured light to generate dense point correspondences between optical devices. While this technique offers the advantage of producing high-density correspondence data, it is also time-consuming. This is due to the sequential projection and capture of multiple structured patterns by both cameras. Replacing this with a faster and equally reliable method, such as deep learning based feature matching, could significantly reduce the overall processing time and improve scalability.

The current implementation assumes that the intrinsic parameters of the cameras are known beforehand. However, in many real-world applications, calibration of both cameras and projectors is required. Extending this approach to include the self-calibration of cameras, alongside the introduced auto-calibration of projectors, would enable a fully automated calibration pipeline without any prior parameter knowledge.

Additionally, the projectors used in both the synthetic and real-world experiments had negligible distortion parameters, and thus the optimization did not need to account for distortion. However, different projector models may exhibit significant distortion depending on their optics. Incorporating distortion parameters into the optimization framework would make the method more versatile and applicable to a broader range of setups.

In projection scenarios where the overlap between projector content is limited, the current method may face challenges, as it depends on those overlapping points for optimization. Enhancing the algorithm by combining a global reprojection error minimization with the existing distance-based objective could provide a more balanced solution.

This dual-objective strategy would simultaneously minimize reprojection error and reduce visible gaps between projector outputs, improving calibration robustness in low-overlap conditions.

Finally, the evaluation presented in this work is limited to configurations with two cameras. There is no constraint on the number of projectors, which can be one or more, depending on the scenario, and is tested across both synthetic simulations and real-world data. Scaling the method to handle multiple cameras would be beneficial for larger projection mapping installations or any complex setup that requires accurate calibration across multiple viewpoints.

References

- [1] Devesh Adlakha, Adlane Habet, Fabio Morbidi, Cédric Demonceaux, and Michel de Mathelin. Stratified autocalibration of cameras with euclidean image plane. In *31st British Machine Vision Virtual Conference (BMVC 2020)*. BMVA Press, 2020.
- [2] G&P Associates. A quick guide to understanding the costs of 3d laser scanning. <https://www.gpassociates.org/a-quick-guide-to-understanding-the-costs-of-3d-laser-scanning-bg.html>, 2024. Accessed: 2025-05-16.
- [3] Samuel Audet and Masatoshi Okutomi. A user-friendly method to geometrically calibrate projector-camera systems. In *2009 IEEE computer society conference on computer vision and pattern recognition workshops*, pages 47–54. IEEE, 2009.
- [4] Hakan Bacakoglu and Mohamed S Kamel. A three-step camera calibration method. *IEEE Transactions on Instrumentation and Measurement*, 46(5):1165–1172, 1997.
- [5] Tyler Bell, Beiwen Li, and Song Zhang. Structured light techniques and applications. *Wiley encyclopedia of electrical and electronics engineering*, pages 1–24, 1999.
- [6] Oliver Bimber and Ramesh Raskar. *Spatial augmented reality: merging real and virtual worlds*. CRC press, 2005.
- [7] Fabio Bruno, Stefano Bruno, Giovanna De Sensi, Maria-Laura Luchi, Stefania Mancuso, and Maurizio Muzzupappa. From 3d reconstruction to virtual reality: A complete methodology for digital archaeological exhibition. *Journal of Cultural Heritage*, 11(1):42–49, 2010.
- [8] Richard H Byrd, Jean Charles Gilbert, and Jorge Nocedal. A trust region method based on interior point techniques for nonlinear programming. *Mathematical programming*, 89:149–185, 2000.

- [9] Xaviéra Calixte and Pierre Leclercq. The interactive projection mapping as a spatial augmented reality to help collaborative design: Case study in architectural design. In *International Conference on Cooperative Design, Visualization and Engineering*, pages 143–152. Springer, 2017.
- [10] Jia Chen, Dongli Wu, Peng Song, Fuqin Deng, Ying He, and Shiyan Pang. Multi-view triangulation: Systematic comparison and an improved method. *Ieee Access*, 8:21017–21027, 2020.
- [11] Timothy A Clarke and John G Fryer. The development of camera calibration methods and models. *The Photogrammetric Record*, 16(91):51–66, 1998.
- [12] Kenneth M Dawson-Howe and David Vernon. Simple pinhole camera calibration. *International Journal of Imaging Systems and Technology*, 5(1):1–6, 1994.
- [13] C Brown Duane. Close-range camera calibration. *Photogramm. Eng*, 37(8):855–866, 1971.
- [14] Lutz Falkenhagen and M Strintzis. 3d object-based depth estimation from stereoscopic image sequences. In *Proc. Internat. Workshop on Stereoscopic and 3D Imaging*, volume 95, pages 81–86. Citeseer, 1995.
- [15] Jiading Fang, Igor Vasiljevic, Vitor Guizilini, Rares Ambrus, Greg Shakhnarovich, Adrien Gaidon, and Matthew R Walter. Self-supervised camera self-calibration from video. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 8468–8475. IEEE, 2022.
- [16] O. Faugeras, Q. Luong, and T. Papadopoulou. *The Geometry of Multiple Images: The Laws That Govern The Formation of Images of A Scene and Some of Their Applications*. MIT Press, Cambridge, MA, USA, 2001.
- [17] William Ford. *Numerical Linear Algebra with Applications: Using MATLAB and Octave*. Academic Press, 2014.
- [18] Clive S Fraser and Salaheddin Al-Ajlouni. Zoom-dependent camera calibration in digital close-range photogrammetry. *Photogrammetric Engineering & Remote Sensing*, 72(9):1017–1026, 2006.
- [19] Henri P Gavin. The levenberg-marquardt algorithm for nonlinear least squares curve-fitting problems. *Department of Civil and Environmental Engineering Duke University August*, 3:1–23, 2019.

- [20] Jason Geng. Structured-light 3d surface imaging: a tutorial. *Advances in optics and photonics*, 3(2):128–160, 2011.
- [21] Sepehr Ghavam. Projector self-calibration using the dual absolute quadric. Master’s thesis, University of Waterloo, 2023.
- [22] Philip E. Gill, Michael A. Saunders, and Elizabeth Wong. On the performance of sqp methods for nonlinear optimization. In Boris Defourny and Tamás Terlaky, editors, *Modeling and Optimization: Theory and Applications*, pages 95–123, Cham, 2015. Springer International Publishing.
- [23] Anselm Grundhöfer and Daisuke Iwai. Recent advances in projection mapping algorithms, hardware and applications. In *Computer graphics forum*, volume 37, pages 653–675. Wiley Online Library, 2018.
- [24] Rostam Affendi Hamzah, A Fauzan Kadmin, M Saad Hamid, S Fakhra A Ghani, and Haidi Ibrahim. Improvement of stereo matching algorithm for 3d surface reconstruction. *Signal Processing: Image Communication*, 65:165–172, 2018.
- [25] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.
- [26] SD Hppisley-Cox and John Porrill. Auto-calibration-kruppa’s equations and the intrinsic parameters of a camera. In *BMVC*, pages 1–10, 1994.
- [27] Bingyao Huang, Samed Ozdemir, Ying Tang, Chunyuan Liao, and Haibin Ling. A single-shot-per-pose camera-projector calibration system for imperfect planar targets. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 15–20. IEEE, 2018.
- [28] Junhui Huang, Zhao Wang, Qi Xue, and Jianmin Gao. Calibration of a camera–projector measurement system and error impact analysis. *Measurement Science and Technology*, 23(12):125402, 2012.
- [29] Daisuke Iwai. Projection mapping technologies: A review of current trends and future directions. *Proceedings of the Japan Academy, Series B*, 100(3):234–251, 2024.
- [30] Yekeun Jeong, David Nister, Drew Steedly, Richard Szeliski, and In-So Kweon. Pushing the envelope of modern methods for bundle adjustment. *IEEE transactions on pattern analysis and machine intelligence*, 34(8):1605–1617, 2011.

- [31] Yoonwoo Jeong, Seokjun Ahn, Christopher Choy, Anima Anandkumar, Minsu Cho, and Jaesik Park. Self-calibrating neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5846–5854, 2021.
- [32] Hubertus Th Jongen, Klaus Meer, and Eberhard Triesch. *Optimization theory*. Springer Science & Business Media, 2007.
- [33] HM Karara et al. Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. *Photogrammetric Engineering & Remote Sensing*, 81(2), 2015.
- [34] Johnny C Lee, Paul H Dietz, Dan Maynes-Aminzade, Ramesh Raskar, and Scott E Hudson. Automatic projector calibration with embedded light sensors. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 123–126, 2004.
- [35] Shenglin Li and Hwan-Sik Yoon. Enhancing camera calibration for traffic surveillance with an integrated approach of genetic algorithm and particle swarm optimization. *Sensors*, 24(5):1456, 2024.
- [36] Wenguo Li, Yongpeng Zhong, Tao Liu, and Manli Tai. Efficient and accurate projector calibration method based on phase-shifting method with sinusoidal structured light. *Heliyon*, 9(4), 2023.
- [37] Kang Liao, Lang Nie, Shujuan Huang, Chunyu Lin, Jing Zhang, Yao Zhao, Moncef Gabbouj, and Dacheng Tao. Deep learning for camera calibration and beyond: A survey. *arXiv preprint arXiv:2303.10559*, 2023.
- [38] Jinbo Liu, Xianguo Yu, Kai Yang, Xinxin Zhu, and Yue Wu. Automatic calibration method for the full parameter of a camera-projector system. *Optical Engineering*, 58(8):084105–084105, 2019.
- [39] Huafen Luo, Jing Xu, Nguyen Hoa Binh, Shuntao Liu, Chi Zhang, and Ken Chen. A simple calibration procedure for structured light system. *Optics and Lasers in Engineering*, 57:6–12, 2014.
- [40] Ivan Martynov, Joni-Kristian Kamarainen, and Lasse Lensu. Projector calibration by “inverse camera calibration”. In *Image Analysis: 17th Scandinavian Conference, SCIA 2011, Ystad, Sweden, May 2011. Proceedings 17*, pages 536–544. Springer, 2011.

- [41] Paulo RS Mendonça and Roberto Cipolla. A simple technique for self-calibration. In *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, volume 1, pages 500–505. IEEE, 1999.
- [42] Zelin Meng, Xiangbo Kong, Lin Meng, and Hiroyuki Tomiyama. Stereo vision-based depth estimation. In *Advances in Artificial Intelligence and Data Engineering: Select Proceedings of AIDE 2019*, pages 1209–1216. Springer, 2021.
- [43] Mostafa Merras, Soulaïman El Hazzat, Aziz Bouazi, Idriss Chana, Nabil El Akkad, and Khalid Satori. Camera self-calibration with varying parameters based on planes basis using particle swarm optimization. *The Visual Computer*, 39(7):3109–3122, 2023.
- [44] Daniel Moreno and Gabriel Taubin. Simple, accurate, and robust projector-camera calibration. In *2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*, pages 464–471. IEEE, 2012.
- [45] Lei Nie, Yuping Ye, and Zhan Song. Method for calibration accuracy improvement of projector-camera-based structured light system. *Optical Engineering*, 56(7):074101–074101, 2017.
- [46] Jorge Nocedal and Stephen J Wright. *Numerical optimization*. Springer, 1999.
- [47] Artur Nowakowski and Władysław Skarbek. Analysis of brown camera distortion model. In *Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2013*, volume 8903, pages 248–257. SPIE, 2013.
- [48] Peter J Olver, Chehrzad Shakiban, and Chehrzad Shakiban. *Applied linear algebra*, volume 1. Springer, 2006.
- [49] Bronislav Příbyl, Pavel Zemčík, and Martin Čadík. Absolute pose estimation from line correspondences using direct linear transformation. *Computer Vision and Image Understanding*, 161:130–144, 2017.
- [50] Srikumar Ramalingam and Peter Sturm. A unifying model for camera calibration. *IEEE transactions on pattern analysis and machine intelligence*, 39(7):1309–1319, 2016.
- [51] Ramesh Raskar and Paul Beardsley. A self-correcting projector. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 2, pages II–II. IEEE, 2001.

- [52] Research and Markets. Projection mapping market report, 2025. Accessed on April 16, 2025. URL: <https://www.researchandmarkets.com/reports/5767575/projection-mapping-market-report>.
- [53] Peter J Rousseeuw and Mia Hubert. Robust statistics for outlier detection. *Wiley interdisciplinary reviews: Data mining and knowledge discovery*, 1(1):73–79, 2011.
- [54] Joaquim Salvi, Xavier Armangué, and Joan Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern recognition*, 35(7):1617–1635, 2002.
- [55] David Samper, Jorge Santolaria, Francisco Javier Brosed, Ana Cristina Majarena, and Juan José Aguilar. Analysis of tsai calibration method using two-and three-dimensional calibration objects. *Machine vision and applications*, 24:117–131, 2013.
- [56] B-K Seo, M-H Lee, H Park, Jong-Il Park, and Young Soo Kim. Direct-projected augmented reality for robot-assisted surgery. In *39th International Symposium on Robotics, ISR 2008*, pages 598–603, 2008.
- [57] Seung-Jae Son, Yatong An, and Jae-Sang Hyun. Quasi-calibration method for structured light system with auxiliary camera. *Optics & Laser Technology*, 179:111369, 2024.
- [58] Liming Song, Wenfu Wu, Junrong Guo, and Xiuhua Li. Survey on camera calibration technique. In *2013 5th International conference on intelligent human-machine systems and cybernetics*, volume 2, pages 389–392. IEEE, 2013.
- [59] Peter Sturm, Srikumar Ramalingam, Jean-Philippe Tardif, Simone Gasparini, Joao Barreto, et al. Camera models and fundamental concepts used in geometric computer vision. *Foundations and Trends® in Computer Graphics and Vision*, 6(1–2):1–183, 2011.
- [60] Masatoki Sugimoto, Daisuke Iwai, Koki Ishida, Parinya Punpongsanon, and Kosuke Sato. Directionally decomposing structured light for projector calibration. *IEEE Transactions on Visualization and Computer Graphics*, 27(11):4161–4170, 2021.
- [61] Zhongwei Tang, Rafael Grompone von Gioi, Pascal Monasse, and Jean-Michel Morel. A precision analysis of camera distortion models. *IEEE Transactions on Image Processing*, 26(6):2694–2704, 2017. doi:10.1109/TIP.2017.2686001.

- [62] Mahdi Abbaspour Tehrani, M Gopi, and Aditi Majumder. Automated geometric registration for multi-projector displays on arbitrary 3d shapes using uncalibrated devices. *IEEE transactions on visualization and computer graphics*, 27(4):2265–2279, 2019.
- [63] Bill Triggs. Autocalibration and the absolute quadric. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 609–614. IEEE, 1997.
- [64] Bill Triggs, Philip F McLauchlan, Richard I Hartley, and Andrew W Fitzgibbon. Bundle adjustment—a modern synthesis. In *International workshop on vision algorithms*, pages 298–372. Springer, 1999.
- [65] Bill Triggs, Philip F McLauchlan, Richard I Hartley, and Andrew W Fitzgibbon. Bundle adjustment—a modern synthesis. In *Vision Algorithms: Theory and Practice: International Workshop on Vision Algorithms Corfu, Greece, September 21–22, 1999 Proceedings*, pages 298–372. Springer, 2000.
- [66] Antonio Valdés, José Ignacio Ronda, and Guillermo Gallego. The absolute line quadric and camera autocalibration. *International Journal of Computer Vision*, 66(3):283–303, 2006.
- [67] Joni Vanherck, Lode Jorissen, Brent Zoomers, and Nick Michiels. Projector-camera calibration with non-overlapping fields of view using a planar mirror. *Virtual Reality*, 29(1):1–13, 2025.
- [68] Pranav Kumar Ayee Goundar Venkatesan, Saed Moradi, Paul Fieguth, and Mark Lamm. Evaluating the impact of stereo overlap on calibration parameters in camera and projector systems. *Journal of Computational Vision and Imaging Systems*, 10(1):30–33, 2024.
- [69] Simon Willi and Anselm Grundhöfer. Robust geometric self-calibration of generic multi-projector camera systems. In *2017 IEEE International symposium on mixed and augmented reality (ISMAR)*, pages 42–51. IEEE, 2017.
- [70] Shuntaro Yamazaki, Masaaki Mochimaru, and Takeo Kanade. Simultaneous self-calibration of a projector and a camera using structured light. In *CVPR 2011 WORKSHOPS*, pages 60–67. IEEE, 2011.

- [71] Haruo Yanai, Kei Takeuchi, and Yoshio Takane. Singular value decomposition (svd). In *Projection Matrices, Generalized Inverse Matrices, and Singular Value Decomposition*, pages 125–149. Springer, 2011.
- [72] Seongwook Yoon, Taehyeon Choi, and Sanghoon Sull. Depth estimation from stereo cameras through a curved transparent medium. *Pattern Recognition Letters*, 129:101–107, 2020.
- [73] Christopher Zach. Robust bundle adjustment revisited. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V 13*, pages 772–787. Springer, 2014.
- [74] Jiarui Zhang, Yingjie Zhang, and Bo Chen. Out-of-focus projector calibration method with distortion correction on the projection plane in the structured light three-dimensional measurement system. *Sensors*, 17(12):2963, 2017.
- [75] Xu Zhang and Limin Zhu. Projector calibration from the camera image point of view. *Optical Engineering*, 48(11):117208–117208, 2009.
- [76] Zhengyou Zhang. On the epipolar geometry between two images with lens distortion. In *Proceedings of 13th International Conference on Pattern Recognition*, volume 1, pages 407–411. IEEE, 1996.
- [77] Zhengyou Zhang. Determining the epipolar geometry and its uncertainty: A review. *International journal of computer vision*, 27:161–195, 1998.
- [78] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11):1330–1334, 2002.
- [79] Zhengyou Zhang. Camera calibration. In *Computer vision: a reference guide*, pages 130–131. Springer, 2021.
- [80] Zhengyou Zhang. Eight-point algorithm. In *Computer Vision: A Reference Guide*, pages 370–371. Springer, 2021.
- [81] Zhengyou Zhang. Essential matrix. In *Computer Vision: A Reference Guide*, pages 400–401. Springer, 2021.