Photometric calibration and out-of-focus blur compensation for projector camera systems

Song, Peng

2006

Song, P. (2006). Photometric calibration and out-of-focus blur compensation for projector camera systems. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/2449>

<https://doi.org/10.32657/10356/2449>

Nanyang Technological University

Downloaded on 26 Apr 2023 21:20:06 SGT

Photometric Calibration and Out-of-Focus Blur Compensation for Projector Camera Systems

Song Peng

Division of Computing Systems
School of Computer Engineering

*A thesis submitted to the Nanyang Technological University
in fulfilment of the requirements for the degree of
Doctor of Philosophy*

December 2006

Abstract

Projector-camera systems are strong contenders as the enabling technology for a future environment in which all physical surfaces are potential interactive displays. Small mobile projectors and cameras can be readily placed non-obtrusively in corners of rooms, yet achieve a wide area coverage in terms of light projection and camera-based sensing. However, the physical configurations of these systems will typically encounter technical issues that result from oblique projection. In an oblique projection scenario, not only will the raw displays look geometrically distorted, but significant color balance and brightness artifacts will also be apparent. Moreover, the depth of field of projectors can easily be exceeded in an oblique projection, resulting in non-uniform out-of-focus blur in some regions of projection.

In this thesis, a review of standard technical concepts in relation to projectors and cameras is provided. Recent approaches to geometric calibration are briefly discussed, and the standard geometric calibration method used in the thesis work is described.

A novel framework for grayscale photometric calibration of overlapping projectors and cameras is presented, and extended to colorimetric calibration. The approaches are predominantly based on the analysis of isointensity curves – curves which relate different projector input intensity combinations that result in the same camera-based pixel intensities. The calibration approaches do not require a specific parametric model for the shapes of these functions, nor impose any smoothness constraints. Unlike existing approaches requiring expensive equipment, a low dynamic range camera with minimal controls is sufficient for the calibration. Results for both grayscale photometric and colorimetric calibration demonstrate the accuracy of the models mapping projector input to camera observations.

In the case of projector out-of-focus blur due to oblique projection, an image preconditioning technique is proposed to compensate for the non-uniform blur. Spatially variant point spread functions (PSF) are empirically estimated across the image and adaptively inverted through Wiener filtering, and image compositing from filtered basis images. In an alternate approach, an analytical model for the oblique PSF is derived based on an apertured thin-lens model for the projector. Methods for calibrating the different parameters of this model are described, and a PSF map estimated based on the model. Results for both the empirical and analytical approaches to non-uniform out-of-focus blur compensation are presented.

In conclusion, some future directions for extending the work in this thesis are discussed, in conjunction with potential advancements in the capabilities of projector-camera systems.

Acknowledgements

This thesis summarizes the research work done during my Ph.D candidature at the School of Computer Engineering. Over this period of time I have received considerable help from many people, but none more so than from my supervisor Dr. Cham Tat Jen. His thoughtful guidance, invaluable suggestions and constant encouragement has always been my source of inspiration in my research. What's more, I have also learned a lot about how to be a qualified researcher from his continual enthusiasm, passion and serious attitude towards research work, for which I am extremely grateful.

I would also like to thank my seniors - Mr. Chen Xi and Mr. Pong Hon Keat and my juniors - Mr. Feng Zhou and Mr. Pham Minh Tri for their insightful discussions and suggestions in relevant research. Thanks should also go to my friends in the Center for Multimedia and Network Technology (CeMNet), including Wu Jianhua, Zhou Zhi, Wang Huiqing, Ye Minghua, Zhou Chen, Liu Song, Hu Yiqun and Guo Jing for their warm help and encouragement, and my roommate Hu Pu for his daily cares and concerns.

I would like to take this opportunity to thank Nanyang Technological University for providing me the research scholarship. I would also like to thank the CeMNet director Dr. Chia Liang Tien, Clement and technicians Miss Chua Poo Hua, and Miss Oh Hwee May for their help to the convenient access to the computing facilities.

Finally, I wish to express my deepest gratitude to my parents Yiling and Zhenyu Song for their unfailing love and constant support and care.

Contents

1	Introduction	1
1.1	Motivation	1
1.1.1	Introduction of Projector Camera Systems	1
1.1.2	Advantages of using Projector Camera Systems	6
1.1.3	Research Issues	8
1.2	Problems Addressed in this Thesis	12
1.2.1	Geometric Calibration	12
1.2.2	Photometric Calibration	13
1.2.3	Out-of-Focus Blur Estimation and Image Preconditioning	13
1.3	Thesis Outline	14
2	Basic Concepts	17
2.1	Projector and Camera Characteristics	17
2.1.1	Projector Characteristics	17
2.1.2	Camera Characteristics	20
2.2	Projector Imaging Models	22
2.2.1	Pinhole Model	22
2.2.2	Thin-lens Model	24
2.3	Radiometric Image Formation	30

2.4	Image Restoration	33
3	Geometric Calibration of Projector Camera Systems	39
3.1	Homography Calibration	40
3.2	Full Geometric Calibration Approaches	42
3.2.1	Camera Geometric Calibration	42
3.2.2	Projector Geometric Calibration	48
3.3	Conclusion	53
4	Photometric Calibration of Projector Camera Systems	55
4.1	Overview	55
4.2	Grayscale Photometric Calibration	59
4.2.1	Basic Framework	59
4.2.2	Isointensity Curves	61
4.2.3	Dual-Projector Calibration	63
4.2.4	Computing Radiance Contribution from Binary Light Sources	71
4.2.5	Camera Calibration	72
4.3	Colorimetric Calibration	73
4.3.1	Expansion of Framework	73
4.3.2	Near-Degeneracies in Colorimetric Calibration	76
4.3.3	Single Projector Calibration	77
4.3.4	Dual Projector Calibration	82
4.3.5	Camera Colorimetric Calibration	84
4.4	Calibrating Reflectance Map for Multiple Pixels	84
4.5	Experiments and Results	87
4.5.1	Estimating Grayscale Intensity Transfer Functions	87

4.5.2	Estimating Color Channel Intensity Transfer Functions . . .	88
4.5.3	Prediction of Camera Output Based on Known Projector Input	92
4.5.4	Mixed and Compensating Projector Display for Camera Observations	98
4.6	Conclusion	101
5	Projector Out-of-Focus Blur Analysis and Compensation	107
5.1	Image Preconditioning	107
5.1.1	Blur Compensation Based on Image Preconditioning . . .	107
5.1.2	Limitations of Image Preconditioning	111
5.2	Image Preconditioning using an Observation-based Approach . . .	112
5.2.1	Blur Estimation	112
5.2.2	Computing the Preconditioned Image	118
5.2.3	Experiments and Results	119
5.3	Thin-lens Projection Modeling	127
5.3.1	Point Spread Function from the Thin-lens Projection Model	127
5.3.2	Calibration of the Thin-lens Projection Model	134
5.3.3	PSF Map Calculation	142
5.4	Generation of Preconditioned Images	146
5.4.1	Preconditioned Image Generation using Windowed Wiener Filtering	146
5.4.2	Experiments and Results	146
5.5	Comparison of Approaches	155
6	Conclusion and Future Work	157

6.1	Conclusion	157
6.2	Future Work	160
	Bibliography	163

List of Figures

1.1	Examples of Scientific Visualization	3
1.2	Examples of Shader Lamps	4
1.3	An Example of Immersive Environment from Metaverse Lab	5
1.4	Examples of Large Display System	6
1.5	Obliquely Positioned Projector Camera System Setup.	9
1.6	Brightness Fall-off in a Projected Imagery.	10
1.7	Projector Out-of-focus Blur	11
2.1	LCD Projectors Intensity Response	18
2.2	White Enhancement for Small DLP Projectors	19
2.3	Basic Transformations within Camera Imaging System	20
2.4	Examples of Camera Response Functions	21
2.5	Pinhole Camera Imaging	22
2.6	Pinhole Camera Model	23
2.7	Thin Lens System Imaging	25
2.8	Thin Lens Model	26
2.9	Blur in Camera Thin-lens Imaging.	27
2.10	Depth of Field Illustration	28
2.11	Projector Depth of Field Illustration	29

2.12 Radiometry Illustration	31
2.13 Radiometric Image Formation Model	32
3.1 Projector Camera Homography Calibration	40
4.1 Photometric Calibration Framework.	60
4.2 (a) Graph of camera-observed pixel intensity as a function of projector input intensities; (b) Isointensity curves.	62
4.3 The Staircase Method.	64
4.4 Recursive Bisection of Intensity Levels.	66
4.5 Gridline Intersections on Isointensity Curves	69
4.6 Color Light Path in a Projector-Camera System.	74
4.7 Gridline Meshes of Isointensity Curves for Single-projector Colorimetric Calibration (1).	78
4.8 Gridline Meshes of Isointensity Curves for Single-projector Colorimetric Calibration (2).	80
4.9 Computed Projector Camera Intensity Transfer Functions.	88
4.10 Double Projector – Single Camera Intensity Transfer Function Space.	89
4.11 Error Plot for the Grayscale Intensity Transfer Function Space	90
4.12 Estimated Projector to Camera Color Channel Intensity Transfer Functions.	91
4.13 Camera Observation Prediction over Single/Double Projectors Display (1).	94
4.14 Camera Observation Prediction over Single/Double Projectors Display (2).	95
4.15 Prediction of Camera Image (1).	96

4.16	Prediction of Camera Image (2).	97
4.17	Prediction Errors Plot.	98
4.18	Camera Observation of Double-projector Mixed Displays (1).	99
4.19	Camera Observation of Double-projector Mixed Displays (2).	100
4.20	Color Compensation Experiment (1).	102
4.21	Color Compensation Experiment (2).	103
4.22	Compensation Errors Plot.	104
5.1	Stages of our PSF estimation.	116
5.2	Estimated PSF $M_j(x, y)$.	117
5.3	Piecewise PSF Filtering.	118
5.4	Example 1	122
5.5	Example 1 Cont'd	123
5.6	Example 2	124
5.7	Example 2 Cont'd	125
5.8	Example 3	126
5.9	Light Distribution through the Thin-lens of the Projector.	129
5.10	Vector Representation of Projector Pixel Light Distribution in Thin-lens Projection.	130
5.11	Sample PSF shapes in On-axis Cases under Large and Small Apertures.	135
5.12	Sample PSF shapes in Off-axis Cases under Large and Small Apertures.	136
5.13	Projector Geometric Calibration Models.	138
5.14	Lens Aperture Radius Estimation.	141

5.15 An Obliquely Projected Image with Out-of-focus Regions and corresponding PSFs.	144
5.16 Example 1	148
5.17 Example 1 cont'd	149
5.18 Example 2	150
5.19 Example 2 Cont'd	151
5.20 Example 3	152
5.21 Example 4	153
5.22 Example 4 Cont'd	154

Chapter 1

Introduction

1.1 Motivation

1.1.1 Introduction of Projector Camera Systems

With the rapid development of digital projection and imaging technologies, projectors and cameras are becoming less expensive. With pervasive and abundant computational power and connectivity available, visualization techniques for the lay user are starting to break confines of traditional computer monitors and fixed displays. Displays in the future can be so ubiquitous that every physical interface will become an interactive display [49, 20]. Projectors and cameras can be combined together to instrument such pervasive display capabilities in an indoor environment without massive equipment resources.

Current applications of projector camera systems varies from personal computing devices, large scale data viewer, education support facilities, to family en-

tertainment equipment. A variety of fields such as research, industry, entertainment, education, are putting more efforts in the exploration of projector camera systems applications. Some typical applications of projector camera systems are:

- *Smart presentation systems in education*

Intelligent camera-projector systems allow the creation of interactive displays which would greatly enhance education. The presentations and lectures will enable gesture, laser or touch interaction with the display. Students can also participate in a lecture by interacting with the system and cooperating with lecturers on complicated problems. Researchers from the former Just Research developed such a system [56] for smart presentations.

- *Scientific visualization*

Large-scale tiled display walls can be used to visualize large scale data, flow charts or object structures. Research fields such as astronomy science, biomedicine and financial engineering crucially depend on the visualization and analysis of numerical data or physical structures simultaneously. Interactive display walls can be applied to the research of these fields, and allow many researchers to discuss and cooperate effectively. Prototype systems include the Princeton University's *Scalable Display Wall* and the University of Minnesota's *PowerWall* (see figure 1.1).

- *Manufacturing Design*

The ability to display and interact with large-scale representations of manufacturing process flow and logistic chains has always been a fundamental requirement in certain heavy industry designs, such as automotive design, airplane design, *etc.* This requirement has manifested itself in a variety of

Figure 1.1: Examples of Scientific Visualization: (a) Princeton University's "Scalable Display Wall"; (b) University of Minnesota's "PowerWall".

traditional media, such as blue-prints, tape drawings, *etc.* Large-scale interactive displays can provide convenient approaches to such designs. Buxton, et al. [10] have been carrying out research in the large displays in automotive design field.

- *Architectural Design using Spatially-augmented reality*

Although there are modern graphics systems which help architects and designers build models, there is still a tradition of constantly scaled-down physical models of buildings or products. The scaled models are however static in structure and surface characteristics. Camera-projector systems can be used in augmented reality applications without the need of head-mounted displays or goggles. Systems such as *Shader Lamps* [50, 4, 48] and *iLamps* [46] have been designed to interactively generate different effects on the object surfaces for visualization. (See figure 1.2) More recently, the metaverse lab has an augmented bookshelf using projector camera systems [13].

Figure 1.2: Examples of Shader Lamps: (a) is the physical model; (b) is a view of (a) augmented with projected imagery.

- *Immersive environment*

New interfaces can be built when multiple projectors seamlessly display on almost all the walls within an enclosed environment. Cameras and other sensors can be the interfaces to detect movements and gestures of users interacting with the system. Examples include *immersive* projection based environments [5, 14, 15]. The Metaverse Lab of the University of Kentucky has carried out research on such an interactive immersive display environment [28] (see figure 1.3).

- *Entertainment*

The traditional theater works with a single large projector, but today, accurate geometric registration and photometric calibration allow *projector arrays* to display seamless color image on planar or cylindrical surfaces in theaters and large conference rooms. A large scale display called Spitz ElectricHorizonTM [17] on a cylindrical surface is shown in figure 1.4 (a), and a multi-user display (medium scale) called Panoram PanoWallTM [42] is shown in (b). Another application is the large display system in the SGI-

Figure 1.3: An Example of Immersive Environment from Metaverse Lab, University of Kentucky.

powered *Reality Theater* in the Nanyang Technological University [59] as shown in figure 1.4 (c) and (d).

- *Pervasive computing*

In a pervasive computing future, users have access to computational resources, information and services anywhere and anytime. Present systems are based on users carrying laptops or PDAs everywhere, wearing computer graphics goggles, or accessing from distributed monitors and displays. An alternative possibility is the use of projector-camera systems to create interactive displays everywhere in an indoor environment by transforming any suitable surface into a projected interactive display screen. *The Everywhere*

Displays Projector [43] is an initial example of such an application.

(a)

(b)

(c)

(d)

Figure 1.4: Examples of Large Display System: (a)Spitz ElectricHorizonTM is a large scale display system; (b)Panoram PanoWallTM is a multi-user display system. (c) and (d) show the Reality Theater [64] in the Nanyang Technological University.

1.1.2 Advantages of using Projector Camera Systems

Projectors in the recent decade has become prevalent in business use and more and more families are buying portable projectors for home entertainment use. It has come to an age that projected display is replacing the traditional way of visualization on monitors in our daily life. Although projectors have their problems such as low visibility in bright ambient light or on inappropriate surfaces, unwanted black level, resolution loss due to keystone correction, they still have great special ad-

vantages over traditional displays and play a very important role in visualization and display nowadays:

- *Scalability.* Projected display can be largely scalable without using new equipment. However, since the traditional monitor display size can not be changed, to create a larger size of display, it is necessary to have either replacement of original monitors with larger size monitors or multiple monitors.
- *Smarter Display.* Combined with a cheap camera, projectors can “see” and react to designed commands intelligently. Corrected display can be showed even if the projectors are put in oblique positions to the screen. Other smart adaptations may be detecting and tracking humans with spot lights, interaction with laser pointer and human pose commands, *etc.*
- *Smaller size and Quick Deployment.* Modern projectors are becoming smaller because of their hardware breakthrough. Normal portable projectors are of a smaller size, and even smaller than a 15” LCD monitor. The latest series of LED projectors are small enough to be put in pockets. With the help of an off-the-shelf camera, projectors can be quickly set up for smart display.
- *Ubiquitous Displays.* With the software support, projectors can be put anywhere in any orientation to project onto any surface. Unlike traditional projectors which have to be mounted on the ceiling, the space and setup constraint of using ubiquitous displays is relaxed by the oblique front projection.

1.1.3 Research Issues

The powerful display and interactive advantages can not be achieved using existing hardware only. Together with software approaches, projectors and cameras can be used in creating smart and scalable displays. However, in creating ubiquitous displays, projectors may project onto various kinds of surfaces made of different materials. Additionally, the projection surface may not necessarily be planar. Projectors may be placed in corners to achieve wide coverage of projection, resulting in oblique projection to various surfaces. Oblique projection results in a set of technical challenges, such as geometric and photometric distortions, brightness decay across the display, color imbalance and projector out-of-focus blur. These problems are analyzed in this thesis and novel solutions proposed.

Geometry Issues

For an oblique projector camera system setup as shown in figure 1.5, images are geometrically distorted on the display surface for normal input images. Geometric relationships between the projector and the display surface need to be known in order to undistort the original input image such that the projected image observed by the camera or the user looks the same as the original image. The geometric relationships can be expressed as a simple 2D projective transformation, *i.e.* homography, or a 3D rotation and translation from the projector coordinates to the world coordinates. These relationships can be acquired through geometric calibration. Manual intervention may be needed to complete the geometric calibration if cameras are not available. With cameras, a closed loop of projection and observation can be formed and thus the calibration process becomes more automatic and accurate.

Figure 1.5: An Illustration of Obliquely Positioned Projector Camera System Setup.

Photometry Issues

Even with projectors of the same brand and model, the color balance and brightness may vary between each projector. The gradual degradation of a projector lamp may additionally result in corresponding temporal change of display color and brightness. When multiple projectors are used together to form a large scale display wall, inter-projector color and brightness differences will result in easily observable artifacts which are detrimental to display quality. When a projector is projecting obliquely onto a display surface, this problem is compounded as within-projector color and brightness is no longer uniform. Figure 1.6 gives a good example of within-projector brightness fall-off. In order to have a uniform brightness and color display, photometric calibration is needed to allow correction for the fall-off. Figure 1.6 (c) shows the prediction based on the photometric calibration based on the methods that are developed in chapter 4.

(a)

(b)

(c)

Figure 1.6: Brightness Fall-off in a Projected Imagery.(a) shows the input image to a projector; (b) shows the camera observed image of the oblique projection while (c) shows the prediction of camera image based on the photometric calibration of the projector camera system.

Projector Out-of-focus Blur

While significant advances to projector hardware have been demonstrated [40], on the whole, commodity projector hardware has not evolved to accommodate the flexibility allowed by projector-camera systems. Commodity light projectors are still designed to be used in an orthogonal (on-axis) manner with a planar display surface. While vision-based algorithms reduce these constraints and allow for more arbitrary positioning, one consequence is the lack of a single focus setting. Depth-of-field of projectors is often limited, and even slight off-axis projection can lead to blurred regions in the imagery. As shown in figure 1.7 (a), an

Figure 1.7: Projector Out-of-focus Blur. (a) Original image with blurring due to regions of the projected image being out-of-focus. (b) The same projected image with much less blur.

obliquely positioned projector may produce an image with portions that are out of focus. Unlike camera images where we can apply deblurring on the observed image to restore the original image, other ways need to be explored to solve the problem of projector lens out-of-focus. A potential way is to make changes to the

original image and after projection the image (figure 1.7 (b)) looks sharper than the unchanged ones (figure 1.7 (a)).

In all, the above problems are all due to the oblique setup of projectors and cameras. Still the current hardware are designed to be used orthogonally, vision-based algorithms can be used to break this confinement.

1.2 Problems Addressed in this Thesis

In this thesis, we will be focusing on solving the geometry, photometry and out-of-focus projection problems in projector camera systems.

1.2.1 Geometric Calibration

To solve the geometry problems due to oblique positioning of cameras and projectors, firstly, a homography-based calibration approach is introduced in chapter 2. Then standard full calibration approaches recovering camera internal, external and lens distortion parameters are reviewed. Based on the camera full calibration techniques, a projector geometric calibration procedure is proposed afterwards.

Homography calibration is the key to keystone the projection in case of oblique projection and provide basis for the future photometric calibration as well as out-of-focus blur correction, for the geometric distortion caused by the oblique projection has to be corrected in any future stage of the use of projector camera systems.

Full geometric calibration not only provides more information on the intrinsic and extrinsic parameters of the projector and camera, but it also allows for measuring the radial distortion of lens which exists in most commodity cameras. Furthermore, the out-of-focus blur problem can be solved based on the physical

thin-lens radiometric image formation model where projector intrinsic and extrinsic parameters need to be known.

1.2.2 Photometric Calibration

Photometry problems can be solved if we are able to model how the light energy transfers from the projector to the screen, is reflected and transmitted back to the camera. The brightness responses describe the relationships between the input and output of projectors and cameras. By recovering the brightness responses, it is possible to manipulate the input of projectors such that desired output is observed by the camera. In chapter 4, a grayscale photometric calibration method is firstly proposed to estimate the grayscale photometric responses of the projectors and the cameras. As a major extension, a colorimetric calibration method is designed in later parts of that chapter to estimate the colorimetric responses of projector camera systems. Photometric calibration makes it possible to correct for non-uniform colors or brightness responses.

1.2.3 Out-of-Focus Blur Estimation and Image Preconditioning

Due to the projector's limited depth of focus, out-of-focus blur may appear in some regions of the display during oblique projection. Our approach is formulated in the vein of traditional restoration algorithms focused on deblurring. Traditional approaches operate on a degraded image that has undergone some blurring process and try to approximate the original input image. Our problem is cast as the inverse – the original image is known, however, it cannot be processed once it undergoes the degradation process (*i.e.* projection). Thus, we need to pre-

condition the image such that when projected the effects from out-of-focus blur will produce a result as close as possible to the original. To accomplish this, we have developed an algorithm that estimates the blurring process as a set of spatially varying *point spread functions* (PSFs) applied over the projected imagery. A pre-conditioned image is then computed from a set of Wiener filtered images computed from the estimated PSFs. In chapter 5, two image preconditioning approaches are proposed sequentially. An empirical approach using spatially variant Gaussian blurring PSFs is proposed first. The second approach proposed recovers the PSFs from a physical lens model. The geometric calibration results can be used in the second approach for the estimation of PSFs.

1.3 Thesis Outline

In chapter 2, concepts and methods that are used in this thesis are presented. Firstly, projector and camera characteristics are discussed, followed by the description of projector imaging models. Standard image restoration techniques are also reviewed. Lastly, the radiometric image formation procedure is reviewed.

In chapter 3, the homography calibration are introduced first followed by the elaboration of full camera geometric calibration techniques. These techniques recover the camera intrinsic and extrinsic parameters as well as the lens distortion factors. The latest projector calibration techniques are reviewed and subsequently a projector full geometric calibration procedure is proposed based on the camera calibration techniques.

The theory of isointensity curves is presented first in chapter 4. Based on this theory, the basic framework for grayscale photometric calibration with three methods (staircase, recursive bisection, and gridline optimization) are described. The

grayscale framework is then extended in a color projector camera system scenario. Based on the isointensity curves theory, single-projector and dual-projector calibration approaches are proposed. Experiment results are shown to demonstrate the effectiveness of the original and extended gridline optimization methods. Further results are shown proving the performance of colorimetric calibration.

In chapter 5, an image preconditioning technique is introduced first and followed by describing how it is applied to compensate for out-of-focus blur in oblique projection. Then the limitations of the image preconditioning approach are discussed. An approach estimating spatially variant PSFs using variant Gaussian blur kernels is proposed and the composition of basis images presented. Secondly, another approach calculating variant PSFs from the physical thin lens radiometric imaging model is proposed. The calibration of the model is presented afterwards, followed by the computation of the preconditioned images. Experimental results are shown to demonstrate the apparent blur reduction in the out-of-focus region of projected display, followed by a comparison between the two approaches.

The conclusions and future research work are finally presented in chapter 6.

Chapter 2

Basic Concepts

In this chapter, some basic concepts of projector-camera systems are introduced and necessary for the understanding of the research work done in the following chapters. First of all, projector and camera characteristics are introduced, followed by geometric modeling of projectors. Radiometric image formation principles are introduced next, and finally an overview of image restoration techniques are presented.

2.1 Projector and Camera Characteristics

2.1.1 Projector Characteristics

Projectors usually exhibit a nonlinear photometric responses, *i.e.* the radiance projected on the screen has a nonlinear relationship with the intensity of pixels in a projector. Since the native responses of the LCD imaging elements are almost

linear, hardware image processing units are usually used to optimize the video display, such that the typical intensity transfer functions for most commercial LCD projectors will approximate a gamma curve, as shown in figure 2.1. Brainard [7]

Figure 2.1: LCD Projectors Intensity Responses: Intensity Transfer Functions for eight typical LCD projectors which approximate gamma curves, from M.C.Stone [54].

stated that CRT gamma model can be used successfully to characterize LCD projectors. However, this is not necessarily true due to differences in the design of different LCD projector models. Hence it is important to check each individual projector before applying the model to it. Besides the LCD imaging elements, there is also a dichronic mirror to split the outgoing light into red, green and blue separations. Different from monitors, the black level of a LCD projector can not be neglected and must be subtracted from all color measurements to indicate an additive system, which can be shown from the its CIE Lab¹ space [54].

This is different for DLP (Digital Light Processing) projectors. In both LCD and DLP technology, luminance is by default linearly proportional to the input

¹A color specification system by Commission Internationale de L'Éclairage (CIE) [1].

signal. Thus a degamma function is often applied to convert for signals that are pre-adjusted for optimal display on CRT systems. However in DLP technology, other intensity functions are applied to maximize luminance and mid-level contrast. For example, a “white” filter [30] to exploit the clear part of a DLP color wheel increases maximum white output to 145% the sum of the RGB components [54, 30]. The white filter is added in some nonuniform R,G,B values to keep a smooth luminance ramp without hue shifts, as shown in figure 2.2 (a), which plots output luminance for an 11-step gray ramp. This reforms the projectors again to show a nonlinear response, which caters to the nonlinear perceptual response to luminance of human vision system. In figure 2.2 (b), the CIE XYZ

(a)

(b)

Figure 2.2: White Enhancement for Small DLP Projectors: (a) is a plot of output luminance vs. summed RGB luminance (straight line); (b) is a scatter plot in CIE XYZ tristimulus space of data taken from a Compaq MP 1800 projector, and the extended white point is clear visible, from [54].

tristimulus space is shown. The extended white points at the clear segment are designed to enhance the luminance level.

Again DLP projectors also emit a significant amount of light while displaying

“black” pixels, which results in a black offset for the display. Consequently the black point must be included in the characterization of projectors.

Laser projectors are also used in projecting texts and mono-color line-drawings, and they are always in focus. However, they are not able to project color images as LCD and DLP projectors and are not widely used in daily life. Therefore, we are not going to discuss in detail on laser projectors.

2.1.2 Camera Characteristics

The camera imaging process is quite complex which may involve lens fall-off, spatial light distribution, sensor materials and sensitivities, *etc.* However, when measuring the relationship between a single scene radiance point and the corresponding image brightness, this process can be simplified into two basic transmission functions $O(\cdot)$ and $E(\cdot)$ regardless of the individual factors mentioned above [23]. At first, the light rays go through camera optical system to focus onto its image plane [27], and then electronic sensor or chemical films convert the irradiance into brightness values.

Figure 2.3: Basic transformations within camera imaging system are linear optical process $O(\cdot)$ that maps the scene radiance R to image irradiance I and then nonlinear electronic process $E(\cdot)$ that maps I to image brightness B .

In a digital camera imaging system, as shown in figure 2.3, the optical transmission responses from scene radiance R to image irradiance I can be modeled as $O(\cdot)$ which is normally a linear function [3]. The electronic transmission responses from image irradiance I to image brightness B can be modeled as $E(\cdot)$. This is usually a nonlinear function, as traditional films were designed to have such a nonlinear responses that a wide range of irradiance values can be compressed into a limited range of image brightness values. Now electronic cameras are also designed in this way to simulate the nonlinear response for the same purpose.

Figure 2.4: Examples from Columbia CAVE Database of 201 Camera Response Functions: The database includes photographic films, digital films, digital cameras, CCDs, and synthetic gamma curves, from [23].

In the real world, different cameras may have different response curves, but all are nonlinear. Figure 2.4 shows a wide variety of photographic films, CCD cameras, and digital cameras (detector and electronics) response curves from Columbia University's Database of Real-world Response Functions (DoRF) [23]. A good estimation of the camera response functions can help us solve the camera photometric calibration problems.

2.2 Projector Imaging Models

2.2.1 Pinhole Model

A pinhole camera is the simplest form of cameras where a tiny hole replaces the lens. Light passes through the hole and an image is formed in the camera. (See figure 2.5 for illustration.) A simple pinhole model can be illustrated mathemati-

Figure 2.5: Pinhole Camera Imaging Illustration: light rays from an object pass through a pin hole to form an image.

cally using figure 2.6 (a) where the focal length f is defined by the distance from the image plane to the pinhole. The optical axis is the line passing through the camera center (pinhole in this case) perpendicular to the image plane. It pierces the image plane at the principal point. The camera coordinates of an image point

(x, y, z) can be expressed by the following equation:

$$\begin{bmatrix} fx \\ fy \\ z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \cong \begin{bmatrix} fx/z \\ fy/z \\ 1 \end{bmatrix} \quad (2.1)$$

However for a general pinhole camera model the principal point may not lie on

Figure 2.6: Pinhole Camera Model. (a) illustrates a pinhole camera model where the principal point is on the image center; (b) illustrates a pinhole camera model when there is a shift of principal point from the image center.

the center of the image plane, but rather there is a shift P_x, P_y from the center. This is illustrated in figure 2.6 (b) and it can be expressed using homogeneous coordinates as:

$$\begin{bmatrix} fx + P_x z \\ fy + P_y z \\ z \end{bmatrix} = \begin{bmatrix} f & 0 & P_x & 0 \\ 0 & f & P_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \cong \begin{bmatrix} fx/z + P_x \\ fy/z + P_y \\ 1 \end{bmatrix} \quad (2.2)$$

Since focal length f and principal point position (P_x, P_y) are the inherent parameters of the camera, the matrix K is usually referred to as camera intrinsic or

internal matrix.

$$K = \begin{bmatrix} f & 0 & P_x \\ 0 & f & P_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2.3)$$

The world coordinates of the object point (X, Y, Z) can usually be derived through multiplication with a rotation and translation matrix.

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \cong \begin{bmatrix} f & 0 & P_x & 0 \\ 0 & f & P_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & t_1 \\ 0 & 1 & 0 & t_2 \\ 0 & 0 & 1 & t_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.4)$$

$$R = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix}, \quad T = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \quad (2.5)$$

The rotation matrix R and translation T can be called extrinsic or external matrix.

If we assume a projector is projecting through a pinhole, the pinhole camera model can be applied to it except that light rays travel reversely. In order to geometrically calibrate the projector, the internal and external matrix need to be computed.

2.2.2 Thin-lens Model

There are multiple types of cameras used in reality. Normal camera lenses are typically complex and usually comprised of multi-lenses. To simplify the complex lens analysis, we can use a theoretical single-lens model which is better than a pinhole model in mimicking the real camera lens to express the mathematical relationship, to which is referred as the thin-lens camera model [60]. The fig-

Figure 2.7: Thin Lens System Imaging. Light rays parallel to the optical axis intersect the focus after passing through the lens, while light rays intersecting the focus prior to entering the lens will become parallel to the optical axis after passing through the lens. The light rays traveling through the lens center do not change direction after passing through the lens.

Figure 2.7 is a good illustration of the imaging of a candle in a typical thin-lens camera system. While projector lens systems are typically comprised of complex and multiple lenses, the thin-lens model is used in this thesis for simplicity of analysis. A thin-lens model of a projector is illustrated as in figure 2.8.

From the similar triangles $\triangle PW_lF_l$ and $\triangle H_dOF_l$, $\triangle H_uOF_r$ and $\triangle pW_rF_r$, we get the equations:

$$\frac{Z}{Y} = \frac{f}{y} \quad (2.6)$$

$$\frac{z}{y} = \frac{f}{Y} \quad (2.7)$$

Figure 2.8: Thin Lens Model. The similar triangles $\triangle PW_lF_l$ and $\triangle H_dOF_l$, $\triangle H_uOF_r$ and $\triangle pW_rF_r$ showed the simple geometry between Z , z and f .

If we substitute with $\widehat{Z} = Z + f$, $\widehat{z} = z + f$, then we will get

$$\frac{1}{\widehat{Z}} + \frac{1}{\widehat{z}} = \frac{1}{f} \quad (2.8)$$

where $|\widehat{Z}| > |f|$, for convergent rays. For a single point on an object at a fixed distance from the lens, the imaging plane should be ideally located at a distance $\widehat{z}$ satisfying (2.8) in order for the image of the point to be in sharp focus. However, for a different object point located at a different distance, a circular spot will appear due to light rays from the object point not intersecting exactly on the plane. This circular spot is called the *circle of confusion*, which is illustrated in figure 2.9. Objects within a certain range of depths from the camera will have acceptably small circles of confusion, in which case this range is referred to as the camera *depth of field*. Aperture size and focal length are the two factors affecting

Figure 2.9: Blur in Camera Thin-lens Imaging. Point light sources P_1, P_2, P_3 are positioned at different distances away from the lens. On the camera image plane, different sizes of circular light spot (circle-of-confusion) are formed.

the camera depth of field.

Similarly, these concepts have analogous counterparts in a projector thin-lens model. When a thin-lens projector is out of focus, the light rays emitting from a single projector pixel and collected by the lens system do not converge onto a single point on the display, but are instead distributed in a circle of confusion (shown in figure 2.10).

Similarly, there is a projector depth of field which depends on the focal length

Figure 2.10: Depth of Field Illustration. The smaller the aperture size, the less light goes through the lens, the smaller the circle of confusion, and the larger the depth of field. The longer the focal length, the smaller the circle of confusion, thus the larger the depth of field.

of the projector lens and its aperture size. However, this depth of field refers to the range of depths that the *imaging plane* can be located, as opposed to the range of depths for object points in the case of a camera. This is shown in figure 2.10 that the sharpening of the image is inversely related to the aperture size. As evident from figure 2.10, a smaller aperture will result in a smaller circle of confusion, and hence a larger depth of field. Thus if the aperture size becomes infinitesimal, the lens model becomes a pin-hole model where the focal length is equivalent to the pinhole model.

However, unlike the inverse relationship between the camera depth of field and focal length, projector depth of field is positively related to its focal length, due to

Figure 2.11: Projector Depth of Field Illustration. A longer focal length ($F_2 > F_1$) results in a smaller angle of cone of light rays ($\beta < \alpha$). For the same maximum acceptable sized circle of confusion c , the smaller depth of field D_1 corresponds to the shorter focal length F_1 while D_2 corresponds to the longer focal length F_2 .

the difference in the definition for depth of field. Since light rays parallel to the optical axis will intersect the focus after passing through the lens, a longer focal length will result in a smaller angle formed by the respective convergent cone of light, shown in figure 2.11. Therefore, based on the maximum acceptable size of the circle of confusion, a lens with a longer focal length achieves a larger depth of field than lens with smaller focal length.

Out-of-Focus Blur

A blurred image is caused not just by this dispersion of light but also the additive overlap of circles-of-confusion from neighboring pixels. The blur of an image depends on both the size of the circle-of-confusion as well as the distribution profile of light within it – this distribution of light is typically called the *point-*

spread function (PSF). The PSF in turn depends on a number of factors including aperture size. Projectors and cameras typically do not have pinhole apertures and therefore have a finite depth-of-field. Projectors, in particular, are designed to have larger apertures that lead to brighter displays. Larger apertures however suffer from smaller depth-of-fields, *e.g.* in a thin-lens model the diameter of the circle-of-confusion for an out-of-focus point is directly proportional to aperture size. This is generally not a problem for projection systems as the projector is typically aligned orthogonal to a flat display surface, thereby allowing all points on the surface to be simultaneously in focus. However, in applications when the projector is significantly skewed to the display surface, or for substantially curved surfaces, there is only a small region on the display that is in sharp focus, while the other parts of the display suffer varying degrees of out-of-focus blur.

2.3 Radiometric Image Formation

Radiometry describes the relationship between surface radiance and the image irradiance. For thin lens, there is a fundamental equation determining this relationship,

$$E = L \frac{\pi}{4} \left(\frac{d}{Z'} \right)^2 \cos^4 \alpha . \quad (2.9)$$

As illustrated in figure 2.12, I is the radiance vector of the light source, L is the scene radiance at point P and E is the image irradiance at screen point p . The normal vector to the surface is denoted by $\mathbf{n}$. Passing through the optical lens, the received image irradiance E is proportional to the scene radiance L , positively related to the size of the lens aperture (the lens diameter is defined as d), and inversely related to the distance Z' from the lens to the display surface. There is a $\cos^4 \alpha$ decay of image irradiance from the center of the image to its edge, where

Figure 2.12: Radiometry Illustration. The image irradiance E at screen point p is in proportion to the scene radiance L at point P , and to the size of the optical lens aperture.

α is the angle between the surface normal of a small area on the object δO and the lens optical axis (shown in figure 2.13).

For completeness of this thesis, we provide a standard proof of this law. The image irradiance E is defined as

$$E = \frac{\delta P}{\delta I} , \quad (2.10)$$

where δP is the power received/emitted, and δI is the small area of projection on the display surface $\mathbf{n}$. The scene radiance L is defined as

$$L = \frac{\delta P}{(\delta O \cos \theta)(\Delta \Omega)} , \quad (2.11)$$

where $\delta \Omega$ is the solid angle facing the lens from δO , and θ is the angle between

any single light ray emitted from δO and the surface normal of area δO . Equating powers by invoking the law of power conservation, using equations (2.10) and (2.11),

$$E = L \Delta \Omega \cos \theta \frac{\delta O}{\delta I} . \quad (2.12)$$

This is illustrated in figure 2.13. To compute E given L , we have to find the

Figure 2.13: Radiometric Image Formation Model. Through similar solid angels $\delta \Omega_O$ and $\delta \Omega_I$, the scene radiance E is related to the image irradiance L by (2.12).

unknowns $\Delta \Omega$, and $\delta O / \delta I$. The solid angle $\Delta \Omega$ can be derived by:

$$\Delta \Omega = \frac{\pi (d/2)^2 \cos \alpha}{(Z / \cos \alpha)^2} , \quad (2.13)$$

where Z is the distance from the lens to the object. In order to find $\delta O/\delta I$, solid angles $\delta\Omega_O$ and $\delta\Omega_I$, which are from the lens center and facing δO and δI respectively, need to be computed first:

$$\delta\Omega_O = \frac{\delta O \cos\theta}{(Z/\cos\alpha)^2} \quad \delta\Omega_I = \frac{\delta I \cos\alpha}{(Z'/\cos\alpha)^2} \quad (2.14)$$

From the figure 2.13 we can see that the solid angle $\delta\Omega_O$ is equal to $\delta\Omega_I$, then

$$\frac{\delta O}{\delta I} = \frac{\cos\alpha}{\cos\theta} \left(\frac{Z}{Z'} \right)^2 \quad (2.15)$$

By combining equations (2.13) and (2.15), (2.12) is becoming

$$E = L \frac{\pi}{4} \left(\frac{d}{Z'} \right)^2 \cos^4\alpha \quad (2.16)$$

This is the equation shown in (2.9).

2.4 Image Restoration

In image processing, restoration attempts to reconstruct or recover an image that has been degraded by using a priori knowledge of the degradation phenomenon. Thus restoration techniques are oriented towards modeling the degradation first and applying the inverse process to recover the original image. Image enhancement techniques can also help to improve an image in some predefined sense, but it is a more subjective process comparing with image restoration.

According to Wallace *et al.* [65], image degradation can be caused by noise, blur, glare or scattering effects. The noise that can cause image degradation ranges widely from Gaussian noise to salt-and-pepper noise or poisson noise, *etc.* Blur degradation is non-random and usually can be modeled by a point-spread-function. Glare usually happens in microscopic images randomly to cause

degradation. Due to sampling density differences, some signals can be randomly scattered to cause degradation.

In the projector setup, the image displayed mainly is degraded by blur with noise. Thus the approaches we need to recover the image should be deblur oriented restoration.

Unsharp Masking

One simple technique to sharpen the image is through *unsharp masking*, a typical image enhancement technique. Unsharp masking [21], and the related high boost filtering, consists of performing several operations in series on the original image. First, the original image should be blurred or averaged. The second step should be to subtract the blurred image from the original image while the resulting image is called the “unsharp mask”. This operation can be express in mathematical forms:

$$f_m(x, y) = f(x, y) - \bar{f}(x, y) \quad (2.17)$$

where $f(x, y)$ is the original image, and $\bar{f}(x, y)$ is the blurred image, $f_m(x, y)$ is the processed image. To keep the original image components by subtracting the original image from the corresponding high contrast image while boosting the edges, we can increase the $f(x, y)$ factor:

$$f_{hb}(x, y) = (A - 1)f(x, y) + f_m(x, y), \quad A \geq 1 \quad (2.18)$$

Unsharp masking is a special form of deconvolution method, as its convolution kernel is of a particular form. Other deconvolution algorithms are also popular to be used for restoring the blur signal to its original position in the image. The PSF can be theoretical, empirical or estimated by blind deconvolution methods.

Inverse Filtering

If an image can be expressed as $f(x, y)$ in spatial domain, where x and y are the horizontal and vertical coordinates in the original image, the transform of its counterpart in frequency domain can be expressed by $F(u, v)$, where u and v denote the location in the frequency image. The degradation function can denoted as $H(u, v)$. In practice, an unknown noise term $N(u, v)$ has to be added after degradation. If we denote the degraded image as $g(x, y)$ in spatial domain, then its Fourier transform $G(u, v)$ can be expressed as:

$$G(u, v) = H(u, v)F(u, v) + N(u, v) \quad (2.19)$$

To compute an estimate, $\hat{F}(u, v)$ of the transform of the original image, we can apply the simple technique of inverse filtering by dividing the transform of the degraded image $G(u, v)$ by the degradation function $H(u, v)$,

$$\hat{F}(u, v) = \frac{G(u, v)}{H(u, v)} \quad (2.20)$$

Replacing $G(u, v)$ in (2.20) with (2.19), we get

$$\hat{F}(u, v) = F(u, v) + \frac{N(u, v)}{H(u, v)} \quad (2.21)$$

Therefore even if we know the degradation function, we can not recover the original undegraded image exactly because $N(u, v)$ is unknown. If the degradation has zero or very small values, then the ratio $N(u, v)/H(u, v)$ can very easily dominate the estimate $\hat{F}(u, v)$.

Wiener Filtering

To incorporate noise into restoration process, Wiener (Minimum Mean Square Error) Filtering [69] may be used. The objective is to find an estimate $\hat{f}$ of the

uncorrupted image f such that the mean squared error between them is minimized.

In frequency domain, this can be expressed as:

$$\hat{F}(u, v) = \left[\frac{H^*(u, v)}{|H(u, v)|^2 + S_\eta(u, v)/S_f(u, v)} \right] G(u, v) \quad (2.22)$$

where $H^*(u, v)$ is the complex conjugate of $H(u, v)$, $S_\eta(u, v)$ is the power spectrum of the noise, and $S_f(u, v)$ is the power spectrum of the undergraded image. Hence $S_\eta(u, v)/S_f(u, v)$ is a 2D spectrum of noise-to-signal ratios. However, this spectrum is usually simplified as a single scalar noise-to-signal ratio K . If $K = 0$, then (2.22) degenerates to (2.20), and the Wiener filter is identical to the inverse filter. However, in the presence of noise, Wiener filtering outperforms inverse filtering by incorporating the non-zero signal-to-noise ratio K , which prevents the denominator in (2.22) from becoming close to zero.

Lucy-Richardson Deconvolution

Lucy-Richardson deconvolution [51, 33] aims to recover an image by iteratively maximizing the likelihood that the recovered image leads to the observed image through convolution with the known PSF in the presence of noise. Pixel values in the observed image can be expressed as:

$$g(x, y) = \sum_u \sum_v h(u, v) f(x - u, y - v) \quad (2.23)$$

where $h(\cdot)$ is the PSF, $f(x, y)$ is the pixel value at location (x, y) in the original undegraded image, (u, v) are dummy variables and $g(x, y)$ is the observed pixel value at location (x, y) . The algorithm is to iteratively calculate the values of $f(x, y)$ based on

$$f^{t+1}(x, y) = f^t(x, y) \sum_u \sum_v \frac{h(u, v) g(x - u, y - v)}{G(x, y)} \quad (2.24)$$

$$G(x, y) = \sum_u \sum_v h(u, v) f^t(x - u, y - v) \quad (2.25)$$

Similar with Wiener Filtering, this techniques requires a known PSF as input, but it can perform multiple iterations to get the maximum likelyhood of the deconvolution. However, Wiener Filtering linearly restores the degraded image while Lucy-Richardson deconvolution provides a nonlinear approach for the restoration.

Blind Deconvolution

Blind deconvolution [62] refers to a family of convolution methods that maximizing the likelihood of the resulting image as Lucy-Richardson deconvolution does, but also can effectively restore unknown PSF. Given an initial input PSF expressed as a uniform function, the algorithm can automatically determine the PSF based on the resulting image itself by computing a variation of the Lucy-Richardson algorithm.

Chapter 3

Geometric Calibration of Projector Camera Systems

In order to display the desired images without geometric distortions, projectors and cameras need to be calibrated geometrically. Homography calibration between projector, camera and display surface can help display a geometrically undistorted image as discussed in papers [12, 29, 44]. Lens distortions can also be corrected by full calibration of projectors and cameras. The calibration of explicit parameters such as internal and external parameters of projectors and cameras [41, 45, 47] will be of further use in projector blur correction which will be discussed in section 5.3.

3.1 Homography Calibration

In an obliquely positioned projector-camera system setup, geometric relationship needs to be recovered to correctly display an image for the viewer, otherwise the image on display will be severely distorted as shown in figure 3.1.

Figure 3.1: Due to projector misalignment, the rectangular screen appears as a distorted quadrilateral(white), but can be corrected by pre-warping using the homographies above.

As the distortion is predominantly the result of perspective projection, pre-warping the image via a homography may be used to correct for the distortion. The homographies between projectors, cameras, and the display surface need to

be calibrated prior to prewarping. Sukthankar *et al.* [56, 55] proposed an automatic keystone correction method in the above setup. Assuming that the projection screen is flat and that the pinhole model may be used for projectors and cameras, the transforms between projectors, cameras, and display screen can all be modeled as 2-D planar homographies:

$$\begin{bmatrix} kx \\ ky \\ k \end{bmatrix} = \begin{bmatrix} t_1 & t_2 & t_3 \\ t_4 & t_5 & t_6 \\ t_7 & t_8 & t_9 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (3.1)$$

where (x, y) and (X, Y) are corresponding points in either two frames of reference, and $\mathbf{t} = (t_1, \dots, t_9)^T$ with $|\mathbf{t}| = 1$ are the parameters to specify the homography. Thus the homography here only has 8 degrees of freedom and can be recovered from as few as 4 pixel correspondences. If more than 4 pixel correspondences are available, more accurate solutions can be sought by using a least-squares solution. Given m point correspondences, we can have the following $2m \times 9$ matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} X_1 & Y_1 & 1 & 0 & 0 & 0 & -X_1x_1 & -Y_1x_1 & -x_1 \\ 0 & 0 & 0 & X_1 & Y_1 & 1 & -X_1y_1 & -Y_1y_1 & -y_1 \\ X_2 & Y_2 & 1 & 0 & 0 & 0 & -X_2x_2 & -Y_2x_2 & -x_2 \\ 0 & 0 & 0 & X_2 & Y_2 & 1 & -X_2y_2 & -Y_2y_2 & -y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_m & Y_m & 1 & 0 & 0 & 0 & -X_mx_m & -Y_mx_m & -x_m \\ 0 & 0 & 0 & X_m & Y_m & 1 & -X_my_m & -Y_my_m & -y_m \end{bmatrix}$$

Then the unit vector $\mathbf{t}$ needs to be found to minimize $|\mathbf{A}\mathbf{t}|$. This can be solved by finding the eigenvector corresponding to the smallest eigenvalue of $\mathbf{A}^T \mathbf{A}$.

A simple solution using 4 feature points correspondences is shown in figure 3.1. If a white rectangle against a contrast background is projected onto the

screen, the four corners of the projected rectangle in the camera image can be extracted using the Hough-transform line-fitting technique. Then with the coordinates of the corners known in projector frame, the projector-camera homography $\mathbf{T}_{c,P_i}$ can be recovered. The corners of the desired quadrilateral formed by dotted lines completely specify the camera-screen homography $\mathbf{T}_{c,s}$, given a definition of the corners in the screen coordinates, and then the homography between each projector and screen $\mathbf{T}_{P_i,s}$ can be recovered from following relationship:

$$\mathbf{T}_{P_i,s} = \mathbf{T}_{c,P_i}^{-1} \mathbf{T}_{c,s} \quad (3.2)$$

By applying the inverse transform $\mathbf{T}_{P_i,s}^{-1}$ to pre-warp the original image, a corrected image can be perceived on the display screen.

3.2 Full Geometric Calibration Approaches

3.2.1 Camera Geometric Calibration

The objective of camera geometric calibration is to estimate a group of parameters that define the internal geometric characteristics (intrinsic parameters) of a camera as well as its 3D pose (extrinsic parameters). The direct calibration of these parameters mentioned above is commonly referred to as explicit camera calibration. Implicit camera calibration, on the contrary, does not calculate the physical parameters but some implicit parameters based on different implicit camera models [67, 38], such as the computation of mapping parameters between a point on a 3D object and its 2D camera coordinates. The focus of this thesis is on explicit camera calibration methods that are needed as basis for similar projector geometric calibration.

The classic photogrammetric method [52] resolved the problem through non-linear error minimization, but has high computational cost. The DLT (Direct Linear Transformation) method was proposed by Abdel-Aziz and Karara [2] to reduce the computational cost, and later Tsai [61] proposed another closed-form approach. However, the reduction of computational costs is traded off with loss of accuracy due to simplifications of camera models. These methods do not achieve as good a result as the nonlinear minimization method. Combinations of linear and nonlinear methods (*e.g.* [68]) may be used where the initial parameters for nonlinear minimization are derived from linear methods in advance.

In our experiments, the camera geometric calibration can be carried out using popular Matlab toolboxes provided by Bouguet [6] which combines the methods in Heikkilä and Silvén [26, 25] and Zhang [71, 72].

General Camera Model for Geometric Calibration

A pinhole camera model is used assuming collinear projection as the basis for different approaches of calibration. An arbitrary object point P at location (X_i, Y_i, Z_i) in world coordinates can be transformed to camera coordinates (x_i, y_i, z_i) by a rotation and a translation:

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \quad (3.3)$$

The intrinsic camera parameters usually include the effective focal length f , scale factor s_u , and the principal point (u_0, v_0) . The origin of the image coordinate system is in the upper left corner of the image array and the coefficients D_u and D_v are used to convert units in world coordinates to pixels. Using a pinhole camera

model, the projection of the point (x_i, y_i, z_i) to the image plane can be expressed as:

$$\begin{bmatrix} \tilde{u}_i \\ \tilde{v}_i \end{bmatrix} = \frac{f}{z_i} \begin{bmatrix} x_i \\ y_i \end{bmatrix} \quad (3.4)$$

Then the corresponding image coordinates (u'_i, v'_i) in pixels can be obtained from the projection $(\tilde{u}_i, \tilde{v}_i)$ through the following equation:

$$\begin{bmatrix} u'_i \\ v'_i \end{bmatrix} = \begin{bmatrix} D_u s_u \tilde{u}_i \\ D_v \tilde{v}_i \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \quad (3.5)$$

The pinhole model is an approximation of the real camera projection. In order to deal with radial and tangential lens distortion, these are modeled using two equations. The radial distortion is expressed using:

$$\begin{bmatrix} \delta u_i^{(r)} \\ \delta v_i^{(r)} \end{bmatrix} = \begin{bmatrix} \tilde{u}_i(k_1 r_i^2 + k_2 r_i^4 + \dots) \\ \tilde{v}_i(k_1 r_i^2 + k_2 r_i^4 + \dots) \end{bmatrix} \quad (3.6)$$

where $k_1, k_2, \dots$ are the radial distortion coefficients, and $r_i = \sqrt{\tilde{u}_i^2 + \tilde{v}_i^2}$. Normally, one or two coefficients are enough to compensate for the distortion. Similarly, tangential distortion can be expressed as follows:

$$\begin{bmatrix} \delta u_i^{(t)} \\ \delta v_i^{(t)} \end{bmatrix} = \begin{bmatrix} 2p_1 \tilde{u}_i \tilde{v}_i + p_2(r_i^2 + 2\tilde{u}_i^2) \\ p_1(r_i^2 + 2\tilde{v}_i^2) + 2p_2 \tilde{u}_i \tilde{v}_i \end{bmatrix} \quad (3.7)$$

where p_1 and p_2 are the tangential distortion parameters for estimation. Thus a proper camera model can be derived using above equations combining pinhole camera model with correction for radial and tangential distortion components:

$$\begin{bmatrix} u'_i \\ v'_i \end{bmatrix} = \begin{bmatrix} D_u s_u (\tilde{u}_i + \delta u_i^{(r)} + \delta u_i^{(t)}) \\ D_v (\tilde{v}_i + \delta v_i^{(r)} + \delta v_i^{(t)}) \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \quad (3.8)$$

Heikkilä and Silvén's Method

Traditionally, the direct linear transformation (DLT) method was used by Abdel-Aziz and Karara [2] to calibrate a pinhole camera model without any lens distortion parameters. However, as an implicit method, it does not separate the intrinsic parameters from extrinsic parameters at once. To calibrate the model more accurately, Heikkilä and Silvén [26] proposed a nonlinear optimization algorithm to obtain the best estimate of the camera parameters in the presence of camera noise. Take n observations (u'_i, v'_i) , where $i = 1, \dots, n$. The objective function can be expressed as below:

$$F = \sum_{i=1}^n (u'_i - u_i)^2 + \sum_{i=1}^n (v'_i - v_i)^2 \quad (3.9)$$

Equation 3.9 can be minimized in a least square sense. The Levenberg-Marquardt optimization method [31, 37] has been shown to have the fastest convergence. To avoid the failure of trapping in the local minimum, initial parameter values can be assigned using the result of DLT method.

Zhang's Method

Heikkilä and Silvén's method [25] is based on a single view of a non-planar pattern, while Zhang's method [72] provides more flexibility by allowing the use of multiple views of a planar pattern. Starting by denoting a 3D point $\mathbf{P}$ in object coordinates as $[X, Y, Z]^T$, the corresponding 2D projection point $\mathbf{p}$ in image coordinates can be expressed as $[u, v]^T$. In homogeneous coordinates, they can be expressed as $\tilde{\mathbf{P}} = [X, Y, Z, 1]^T$ and $\mathbf{p} = [u, v, 1]^T$, and the pinhole camera model

becomes:

$$s\tilde{\mathbf{p}} = \mathbf{A}[\mathbf{R}, \mathbf{t}]\tilde{\mathbf{P}} \quad \text{with} \quad \mathbf{A} = \begin{bmatrix} \alpha & c & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.10)$$

where s is an arbitrary scalar, $\mathbf{A}$ is the camera intrinsic matrix, and the extrinsic matrix is comprised of a rotation matrix $\mathbf{R}$ and a translation vector $\mathbf{t}$. In the intrinsic matrix $\mathbf{A}$, (u_0, v_0) is the principal point, α and β are the scaling factors in image u and v axes respectively, and c is the skewness ratio factor of the two image axes.

If we assume the model plane is on $Z = 0$ of the world coordinate system, and denote the i^{th} column of the rotation matrix $\mathbf{R}$ by $\mathbf{r}_i$, we can express (3.10) using homography $\mathbf{H}$:

$$s\tilde{\mathbf{p}} = \mathbf{H}\tilde{\mathbf{P}} \quad \text{with} \quad \mathbf{H} = \mathbf{A}[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}] \quad (3.11)$$

A homography $\mathbf{H}$ can be estimated with an image of the model plane. If $\mathbf{H}$ is expressed as $[\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3]$, (3.11) can be changed to

$$[\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3] = \lambda \mathbf{A}[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}] \quad (3.12)$$

where λ is an arbitrary scale factor. Since $\mathbf{r}_1$ and $\mathbf{r}_2$ are from the rotation matrix $\mathbf{R}$ and hence orthonormal, we can have two basic constraints:

$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2 = 0 \quad (3.13)$$

$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_1 = \mathbf{h}_2^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2 \quad (3.14)$$

If matrix $\mathbf{B}$ is created as

$$\mathbf{A}^{-T} \mathbf{A}^{-1} \equiv \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} = \begin{bmatrix} \frac{1}{\alpha^2} & -\frac{c}{\alpha^2 \beta} & \frac{cv_0 - u_0 \beta}{\alpha^2 \beta} \\ -\frac{c}{\alpha^2 \beta} & \frac{c^2}{\alpha^2 \beta^2} + \frac{1}{\beta^2} & -\frac{c(cv_0 - u_0 \beta)}{\alpha^2 \beta^2} - \frac{v_0}{\beta^2} \\ \frac{cv_0 - u_0 \beta}{\alpha^2 \beta} & -\frac{c(cv_0 - u_0 \beta)}{\alpha^2 \beta^2} - \frac{v_0}{\beta^2} & \frac{(cv_0 - u_0 \beta)^2}{\alpha^2 \beta^2} + \frac{v_0^2}{\beta^2} + 1 \end{bmatrix} \quad (3.15)$$

3.2 Full Geometric Calibration Approaches

47

The matrix $\mathbf{B}$ is symmetric and can be re-expressed as a 6D vector

$$\mathbf{b} = [b_{11}, b_{12}, b_{22}, b_{13}, b_{23}, b_{33}]^T \quad (3.16)$$

If we denote $\mathbf{h}_i$ as $[h_{i1}, h_{i2}, h_{i3}]^T$, we have

$$\mathbf{h}_i^T \mathbf{B} \mathbf{h}_j = \mathbf{v}_{ij}^T \mathbf{b} \quad (3.17)$$

where $\mathbf{v}_{ij} = [h_{i1}h_{j1}, h_{i1}h_{j2}+h_{i2}h_{j1}, h_{i2}h_{j2}, h_{i3}h_{j1}+h_{i1}h_{j3}, h_{i3}h_{j2}+h_{i2}h_{j3}, h_{i3}h_{j3}]^T$.

Then the basic constraints (3.13) and (3.14) applies for any given homography and can be rewritten with $\mathbf{b}$ as

$$\begin{bmatrix} \mathbf{v}_{12}^T \\ (\mathbf{v}_{11} - \mathbf{v}_{22})^T \end{bmatrix} \mathbf{b} = 0 \quad (3.18)$$

To solve the above equation for $\mathbf{b}$, n images of the model plane need to be observed such that n equations as in (3.18) can be stacked as

$$\mathbf{V}_{2n \times 6} \mathbf{b} = 0 \quad (3.19)$$

where $\mathbf{V}_{2n \times 6}$ is $2n \times 6$ matrix stacking n matrices $(\mathbf{v}_{12}^T, (\mathbf{v}_{11} - \mathbf{v}_{22})^T)^T$ derived from the n images of the model plane. After the estimation of $\mathbf{b}$, the camera intrinsic matrix $\mathbf{A}$ can be computed analytically (for details, please refer to [71]). Once $\mathbf{A}$ is known, the extrinsic parameters can be recovered from (3.15).

To include the radial distortion parameters, we denote (u, v) as the ideal pixel coordinates, and $(\tilde{u}, \tilde{v})$ as the observed image coordinates. From the definitions in (3.5), we can get

$$\frac{u - u'_i}{u'_i - u_0} = \frac{\tilde{u} - u}{u - u_0} \quad (3.20)$$

$$\frac{v - v'_i}{v'_i - v_0} = \frac{\tilde{v} - v}{v - v_0} \quad (3.21)$$

If we define $r_i = x^2 + y^2$, where (x, y) is the ideal normalized image coordinates, (3.14) can be re-expressed as:

$$\check{u} = u + (u - u_0)[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \quad (3.22)$$

$$\check{v} = v + (v - v_0)[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \quad (3.23)$$

In order to be able to estimate k_1 and k_2 after estimating the other parameters, the above equations can be changed to

$$\begin{bmatrix} (u - u_0)(x^2 + y^2) & (u - u_0)(x^2 + y^2)^2 \\ (v - v_0)(x^2 + y^2) & (v - v_0)(x^2 + y^2)^2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} \check{u} - u \\ \check{v} - v \end{bmatrix} \quad (3.24)$$

In the observed n images, if there are m feature point correspondences for each image, the $2m \times n$ simultaneous equations can be solved by a least squares solution.

Once k_1 and k_2 are derived, since $\mathbf{A}$, $\mathbf{R}$, $\mathbf{t}$ has been computed through methods discussed above after solving (3.19), all the parameters can be put together as an initial estimate for a complete maximum likelihood estimation. By minimizing the objective function

$$\sum_{i=1}^n \sum_{j=1}^m \| m_{ij} - \check{m}(A, k_1, k_2, R_i, t_i, M_j) \|^2 \quad (3.25)$$

This nonlinear minimization problem can be quickly solved running Levenberg-Marquardt algorithm.

3.2.2 Projector Geometric Calibration

Similar to camera geometric calibration, the objective of projector geometric calibration is to estimate the intrinsic as well as extrinsic parameters of the projector. By treating projectors as pinhole projectors, we can use the same analysis as

for pinhole cameras. However, in practice, since the projector is an output device rather than a measurement device, a camera is needed to provide feedback. Raskar and Beardsley [47] proposed a method to calibrate a projector and a camera together by mounting them together as a stereo pair. This method, however, is not applicable to projectors and cameras positioned separately and randomly. Okatani and Deguchi [41] later proposed an automatic calibration of projector-camera system through the estimation of screen-to-camera homography. Their method differs from the traditional homography estimation method (Sukthankar *et al.* [56]) in that it does not need any physical markers or human interaction for the estimation. Raj and Pollefeys [45] subsequently proposed an auto-calibration approach of multi-projectors by exploring the constraints and relationships between the image of absolute conic, intrinsic matrix and homographies. Before the introduction of our calibration approach, Okatani and Deguchi's method as well as Raj and Pollefeys' method will be briefly reviewed in the following parts respectively.

Okatani and Deguchi's Method

In 2001, Okatani and Deguchi [41] proposed an autocalibration approach between a projector, a camera and a display screen. Instead of using fiducial or human interaction, they adapted virtual markers/patterns to be projected to automate the calibration. Projected markers/patterns can be viewed by the camera under different lighting conditions.

Pointing a projector to the screen, let $\mathbf{p} = [x, y, 1]^T$ represent the projector beam orientation where x, y is in projector coordinates, and $\mathbf{s} = [u, v, 1]^T$ represent the corresponding screen point where u, v is in screen coordinates. Then

these two vectors can be related by

$$\mathbf{s} \propto \mathbf{K}\mathbf{R}\mathbf{p} \quad (3.26)$$

where

$$\mathbf{K} = \begin{bmatrix} t_z & 0 & -t_x \\ 0 & t_z & -t_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3.27)$$

Here $[-t_x, -t_y, t_z]$ represent the screen coordinates of the projector position and $\mathbf{R}$ is a rotation matrix from projector frame to screen frame. This is based on the assumption that the projector has the same geometry as a pinhole camera, and a pre-calibrated projector with know intrinsic parameters is used for the auto-calibration.

If screen-to-camera homography is denoted as $\mathbf{H}_{sc}$, and projector-to-camera homography is denoted as $\mathbf{H}_{pc}$, then projector-to-screen homography can be determined by $\mathbf{H}_{ps} \propto \mathbf{H}_{sc}^{-1}\mathbf{H}_{pc}$. Since $\mathbf{H}_{ps}$ can relate $\mathbf{s}$ and $\mathbf{p}$ by

$$\mathbf{s} \propto \mathbf{H}_{ps}\mathbf{p} \quad (3.28)$$

The RQ factorization of $\mathbf{H}_{ps}$ will result in an upper triangular matrix $\mathbf{K}'$, and a rotation matrix $\mathbf{R}'$

$$\mathbf{H}_{ps} = \mathbf{K}'\mathbf{R}' \quad (3.29)$$

Given a series of $\mathbf{H}_{pc,1}, \dots, \mathbf{H}_{pc,n}$ from n projector poses, through (3.28) and (3.29), it can be derived that $\mathbf{H}_{sc}^{-1}\mathbf{H}_{pc,i} \propto \mathbf{K}_i\mathbf{R}_i$. They made a further proposition that there exists a 3×3 matrix $\mathbf{T}$ such that $\mathbf{TKR} \propto \mathbf{K}'\mathbf{R}'$. Thus the key is to find $\mathbf{T}$ if it exists and its property.

The traditional approach of camera autocalibration from an image sequence computes the factorization of $\mathbf{H} \propto \mathbf{KRT}$, while in the projector calibration $\mathbf{H} \propto$

TKR has been estimated. This difference determines that camera calibration theory can not be directly applied here. They subsequently proved the uniqueness of solution and proposed a non-iterative approach using matrix SVD to estimate the matrix $\mathbf{T}$, $\mathbf{K}$, $\mathbf{R}$ accordingly.

Raij and Pollefeys' Method

Similar to Okatani and Deguchi's calibration approach, Raij and Pollefeys[45] proposed an automatic calibration for multi-projectors without the need for fiducial markers and human interaction. However, their method additionally does not require pre-calibrated projectors, and is able to self-calibrate for the intrinsic as well as extrinsic parameters of the projectors. Their assumptions are:

- Projectors are adequately modeled as pinhole projectors;
- Projector pixels are all square;
- The principal point of a projector lies on a vertical line going through the image center;
- Principal point offsets are identical for projectors (of the same brand) with the same zoom settings.

Through observing circular points projected from all projectors by a camera, the images of the circular points must fall on the Image of Absolute Conic (IAC) in all projectors and cameras. If the IAC in projector p is defined as $\omega_p^{-1} = \mathbf{C}_p^T \mathbf{C}_p$, where $\mathbf{C}_p$ is the inverse of the p^{th} projector intrinsic matrix $\mathbf{K}_p$. Likewise, the IAC of the camera c is defined as $\omega_c^{-1} = \mathbf{C}_c^T \mathbf{C}_c$, where $\mathbf{C}_c$ is the inverse of the camera's intrinsic matrix $\mathbf{K}_c$.

If $\mathbf{H}_{cp}$ defines the homography from camera to projector, then

$$\mathbf{x}_c = \frac{1}{\sqrt{2}}(x_c + iy_c), \quad \mathbf{x}_p = \mathbf{H}_{cp}\mathbf{x}_c \quad (3.30)$$

$$\mathbf{x}'_c = \frac{1}{\sqrt{2}}(x_c - iy_c), \quad \mathbf{x}'_p = \mathbf{H}_{cp}\mathbf{x}'_c \quad (3.31)$$

where $\mathbf{x}_p$ and $\mathbf{x}'_p$ are the circular points in the projectors and $\mathbf{x}_c$ and $\mathbf{x}'_c$ are the circular points in the camera. The mathematical constraints then can be expressed as

$$\mathbf{x}_c^T \omega_c^{-1} \mathbf{x}_c = 0 \quad (3.32)$$

$$(\mathbf{H}_{cp}x_c)^T \omega_p^{-1} (\mathbf{H}_{cp}x_c) = 0 \quad \text{for } p = 1 \dots n \quad (3.33)$$

Given n projectors and a camera, this yields $2n + 2$ constraints. Based on their assumptions mentioned above, the numbers of unknowns can be eliminated from $2n + 4$ to $n + 5$, and thus can be solved by first a search for projector focal lengths and principal point offsets followed by a nonlinear minimization procedure to refine the results.

Our Calibration Procedure

The above methods discussed are considered the latest in the projector calibration literature. However, through our observation, the horizontal component of a projector's principal point often does not lie on its image center and may vary from one projector to another. As the assumptions in these two methods are too restrictive, we fall back on a traditional process for calibration.

First, through projecting distinguishing texture or features on the screen, we can derive point correspondences $\mathbf{p} = (P_x, P_y)$ in projector coordinates and $\mathbf{c} = (C_x, C_y)$ in camera coordinates which is related by projector-to-camera homography through $\mathbf{c} = \mathbf{H}_{pc}\mathbf{p}$. The screen coordinate system is defined through a

physical checkerboard pattern on the projection surface. By observing the pattern from the camera, we can estimate the camera-to-screen homography $\mathbf{H}_{cs}$ which relates the point correspondences in the camera to screen coordinates. The homography between the projector and the screen can be subsequently calculated by $\mathbf{H}_{ps} = \mathbf{H}_{pc}\mathbf{H}_{cs}$. By changing the pose of the screen with respect to the projector-camera pair, we can derive a series of n projector-to-screen homographies computed from $\mathbf{H}_{ps,i} = \mathbf{H}_{pc}\mathbf{H}_{cs,i}$ where i denotes the i^{th} pose of the screen. If an input image to the projector is defined as I_p , and the projected image on the screen is defined as I_s , the i^{th} projected pattern (*e.g.* checkerboard) can be warped using the projector-to-screen homographies $\mathbf{H}_{ps,i}$ such that the projected pattern is aligned exactly over the physical pattern. This allows for an alternative treatment of the projector as a camera, in which the warped pattern in the projector frame of reference is simply the “image” of the physical pattern. In this way, by collecting a set of warped patterns as “observed images”, the standard Matlab camera calibration toolbox [6] can directly be used to complete the calibration accurately.

Using physical checkerboard pattern on the projection surface is only possible when the projection surface is flat and the projector can be moved around, while in other cases, this method can not be applied.

3.3 Conclusion

In this chapter, projector and camera geometric calibration approaches were reviewed. Homography calibration was first introduced for easy and quick keystone correction of display. Then the two camera full calibration approaches of Heikkilä and Silvén [26] and Zhang [72] were discussed. For projector geometric calibration, Okatani and Deguchi’s method as well as Raj and Pollefeys’ method were

reviewed. Subsequently our full projector geometric calibration approach was presented.

Chapter 4

Photometric Calibration of Projector Camera Systems

4.1 Overview

Photometric calibration deals with modeling how the camera observed intensities vary with the projector input intensities, and in the case of color projectors and cameras the calibration extends to different color channels. Various methods have been proposed for camera-only photometric/colorimetric calibration [39, 16, 22, 23, 32]. The estimated camera color channel responses are then applied in reverse to infer light energy. Subsequently, either projector color channel intensity transfer functions are estimated with a calibrated camera of linear responses [24], or color maps for projector color channel nonlinearity correction are computed [34, 63]. In all of such cases, the camera has to be photometrically pre-calibrated using color charts under controlled illumination, or via the follow-

ing active methods.

Modeling of Combined Projector-to-camera Intensity Transfer Function

This may be done via a look-up table (LUT) [57]. The LUT is created by varying each projector pixel input and observing the output pixel intensity in a camera, hence forming a mapping between projector input and camera output. An alternative approach is to model the combined projector-to-camera intensity transfer function with parametric curves [29]. Jaynes *et al.* [29] proposed a parametric model for the intensity transfer function, and through nonlinear optimization, the four parameters of the model can be fitted to the measured data points (camera observations of each projector pixel input with varying intensities). While simple, these approaches are time-consuming and sensitive to noise and apply only to fixed poses of the projectors and cameras.

Use of Optical Measurement Devices

The projector-to-screen intensity transfer function may be directly measured via an optical measurement device [34, 35, 70], and the screen-to-camera transfer function deduced. Majumder *et al.* [34] first proposed this approach to calibrate the luminance and chrominance of projectors using a photometer. Per channel color maps are generated to show identical projected pixel colors from different projectors with the same input. This is achieved through mapping each input color to a new input color. In order to generate color maps, three intermediate mapping functions need to be computed. They are nonlinearity correction, luminance matching and chrominance matching. Projectors may have nonlinear intensity-to-radiance responses, similar as CRT displays, but are not simply modeled by

a gamma function. Therefore they proposed using an optical measurement device such as a spectroradiometer or photometer to record the LUT of projector intensity-to-radiance response and inversely apply it to correct its nonlinearity. Subsequently, a luminance matching step is applied to compress each projector's dynamic range into a shared range. Chrominance matching then generates chromaticity maps for each R,G,B channels of the projectors. The concatenation of maps generated from these 3 steps comprise of the per channel color map, which helps achieve the color uniformity in multi-projector displays. However, optical measurement devices such as photometers are confined to be used in a research environment, very costly, and not easily accessible.

High Dynamic Range (HDR) Imaging Approaches

These approaches involve calibrating the screen-to-camera (more specifically irradiance to intensity) transfer function through a series of images captured with different but known exposures [16, 39, 44] (either aperture size or shutter speed). Debevec and Malik first proposed this method to calibrate the digital camera transfer function from high dynamic images. They assumed the digital imaging process imitates the film imaging process in nonlinearity, and the scene is static with no lighting changes during the capturing process. The exposure can be defined as the product of the scene irradiance E_i on the imaging device and the exposure time Δt . Images with different known exposure time Δt_j are the input of their algorithm. Now the camera irradiance-to-intensity transfer function f can be related by pixel value Z_{ij} and the exposure:

$$Z_{ij} = f(E_i \Delta t_j) \quad (4.1)$$

where i is the spatial pixel index, and j is the exposure time index. f is assumed to be monotonic and thus invertible. By taking inversion and logarithm on both sides, the above equation can be changed into:

$$\ln f^{-1}(Z_{ij}) = \ln E_i + \ln \Delta t_j \quad (4.2)$$

If $f^{-1}(\cdot)$ is denoted for simplification as $g(\cdot)$, the above equation will be:

$$g(Z_{ij}) = \ln E_i + \ln \Delta t_j \quad (4.3)$$

Since the pixel value Z is of finite integrate numbers, if n denotes the number of pixels, p denotes the number of images taken, the function $g(\cdot)$ and irradiance E_i can be estimated through minimizing a quadratic objective function O in a least-squared error fashion:

$$O = \sum_{i=1}^n \sum_{j=1}^p [g(Z_{ij}) - \ln E_i - \ln \Delta t_j]^2 + \lambda \sum_{Z_{min}+1}^{Z_{max}-1} g''(z)^2 \quad (4.4)$$

where Z_{max} and Z_{min} are the largest and least pixel values respectively. The first term satisfies the equation set (4.3) in a least square sense, while the second one is the sum of squared values of the second derivative of $g(\cdot)$ to ensure its smoothness. They defined $g''(z) = g(z-1) - 2g(z) + g(z+1)$ in the discrete setting. λ is a scalable weight of the smoothness term relative to the data fitting term. Since it is quadratic in E_i 's and $g(z)$'s, the minimization of the objective function O here becomes a linear least squares problem, and can be robustly solved by Singular Value Decomposition (SVD) method.

These HDR approaches however require more expensive cameras that not only allow user variation of aperture size and/or shutter speed, but also require calibrated values for these parameters. However, in some inexpensive cameras the shutter / integration time can be qualitatively controlled, but often cannot be increased beyond 1/60 seconds; reducing the integration time below this amount

will however result in observable interference with the refresh rates of the projectors.

Our photometric calibration approach is quite different from above mentioned approaches. It exploits the overlap of projectors for calibration and is based on the principle of linear addition of light energy from multiple projectors, and is inspired by the popular photometer-free gamma correction methods for computer monitors [7]. Isointensity curves are proposed to complete the grayscale photometric calibration of intensity transfer functions of projectors and cameras.

In the second part of this chapter, colorimetric calibration will be introduced as an extension to the grayscale photometric calibration approach. The extension of colorimetric calibration are also based on the isointensity curves, but differs from the grayscale photometric calibration due to the complexity of combinations of color channels of projectors and the different color sensitivities of the camera. A direct application of grayscale photometric calibration will not be able to help calibrate a color projector or camera. More interdependence between color channels of projector and camera channels are explored to generate a robust colorimetric calibration approach.

4.2 Grayscale Photometric Calibration

4.2.1 Basic Framework

Assuming a geometrically calibrated system, there are a number of ways to express the photometric relation between a screen pixel and its corresponding projector pixels. As shown in figure 4.1, in the most general form, the camera-observed intensity Z of a particular screen pixel may be expressed as a photomet-

ric function $G(\cdot)$ that depends on the input intensities of corresponding projector pixels

$$Z = G(I_1, I_2, \dots, I_j, \dots, I_N) \quad (4.5)$$

where I_j is the corresponding source pixel intensity set in projector j .

Figure 4.1: Photometric Calibration Framework. The photometric function of projector camera system $G(\cdot)$ can be decomposed into separate projector intensity-to-radiance transfer functions $S_j(I)$, screen reflectance function $M(\cdot)$, and ambient light term A and camera irradiance-to-intensity transfer function $C(\cdot)$. Through this process, the input pixel intensity I transformed to output pixel intensity Z .

Instead of only using equation 4.5, we can also decompose $G(\cdot)$ to form a more physically-explicit photometric equation (modified from [11]) as shown:

$$Z = C(M(I_1, I_2, \dots, I_N, A)) \quad (4.6)$$

where I_j is the corresponding source pixel intensity set in projector j , A is the ambient light contribution, $C(\cdot)$ is the screen radiance to camera intensity transfer function. The additive radiance function $M(\cdot)$ is given by

$$M(I_1, I_2, \dots, I_N, A) = \sum_{j=1}^N S_j(I_j) + A \quad (4.7)$$

where $S_j(\cdot)$ is the intensity to radiance transfer function relating the pixel intensity in projector j to the resulting radiance component contributed by the projector reflecting from the screen in the direction of the camera.

There are a number of key assumptions made in this framework:

- Projectors and cameras are time-invariant devices with monotonic transfer functions of $S_j(\cdot)$ and $C(\cdot)$.
- The reflectance of the screen is invariant to time and light power (i.e. an increase in projector light power results in proportional increase in the radiance towards the camera).

In the following sections we will describe quantitatively recovering a model for the $S_j(\cdot)$ functions for each projector, as well as the $C(\cdot)$ function for the camera.

4.2.2 Isointensity Curves

In a scenario with two overlapping projectors, consider the plot of $Z = G(I_1, I_2)$ in figure 4.2 (a) – this is plot of the intensity of a particular camera-observed pixel as a function of the two input pixels of the source projector, and may be recovered directly through experimentation. The contour lines of this graph are the *isointensity curves* as shown in figure 4.2 (b). An isointensity curve relates points where different combinations of projector intensities give rise to the same camera-observed pixel intensity. Higher dimensional isointensity manifolds are the analogous case for more than two overlapping projectors; nevertheless the subsequent analysis is valid by simply considering pairs of projectors.

Suppose for the moment that Z , I_1 and I_2 are in a continuous domain, and that an isointensity curve is perfectly identified. We can measure the rate of change

Figure 4.2: (a) Graph of camera-observed pixel intensity as a function of projector input intensities; (b) Isointensity curves.

of Z with respect to I_i at *all* points along this curve. However, the rate of change may also be expressed via (4.6) and (4.7) as

$$\frac{\partial Z}{\partial I_1} = \frac{dC}{dM} \frac{\partial M}{\partial S_1} \frac{dS_1}{dI_1} = \frac{dC}{dM} \frac{dS_1}{dI_1} \quad (4.8)$$

Since we are assuming $C(\cdot)$ to be a monotonic function, and by definition it is constant along the isointensity curve, we can therefore deduce that dC/dM is also constant (but unknown) along this curve. Consequently, we can theoretically compute the shape of S_1 simply by integrating along the isointensity curve:

$$S_1 = k \int \frac{\partial Z}{\partial I_1} \frac{dI_1}{dt} dt \quad (4.9)$$

where t is an arbitrarily defined curve parameter of the isointensity curve, and k is an unknown scale factor. Note that the offset of S_1 will also be unknown in this instance. A similar process may be applied to recover S_2 .

While theoretically feasible, doing a direct discrete implementation of the above equations is highly ill-conditioned. In the following sections, we will describe more realistic approaches to estimating the $S_j(\cdot)$ functions. Nevertheless,

it is worthwhile noting that the above analysis shows that, at least in theory, *only a single isointensity curve is required to recover the projector-to-screen intensity transfer functions of two projectors*. This implies that only a very limited photosensing device is required for photometric calibration of the projectors. For practical purposes, a low dynamic range camera with minimal controls is sufficient.

4.2.3 Dual-Projector Calibration

In the following sections, we describe how pixelwise photometric calibration allows the computation of $S_j(\cdot)$ and $C(\cdot)$ to different specified degrees of accuracy.

The Staircase Method

In this method, two isointensity curves are selected. These isointensity curves define two quantum levels of radiance arriving at the camera from the screen pixel. Hence we can make the following observation:

Observation 1 *Given two isointensity curves Ψ_a and Ψ_b , a transition from any state (I_{1a}, I_{2a}) on Ψ_a in the projector input intensity plane, to any other state (I_{1b}, I_{2b}) on Ψ_b , leads to the same exact change in total radiance leaving the screen pixel for the camera. In other words, $\Delta M_{a \rightarrow b} = M(I_{1b}, I_{2b}, A) - M(I_{1a}, I_{2a}, A)$ is invariant for all such transitions. This is true because of the monotonicity assumption on $C(\cdot)$.*

This leads to the following observation:

Observation 2 *Consider all transitions from Ψ_a and Ψ_b in which there is no change in the input intensity of one projector, i.e. all $(I_{1a}, I_{2a}) \rightarrow (I_{1b}, I_{2b})$ where*

either $I_{1a} = I_{1b}$, or $I_{2a} = I_{2b}$. The changes in input intensities of the varying projector therefore **solely** cause the changes in radiance leaving the screen pixel. Moreover, all such changes in radiance are exactly the same (based on the Observation 1). For such cases, ΔM is constant implies $\Delta S_{j(a \rightarrow b)} = S_j(I_{jb}) - S_j(I_{ja})$ is the same for all such transitions.

Figure 4.3: The Staircase Method. A ‘staircase’ is bounded by two isointensity curves. Each horizontal or vertical transition results in equivalent contribution to the radiance change reaching the screen.

Figure 4.3 shows a pair of isointensity curves, and transitions selected to form a ‘staircase’ bounded by these two curves. If we consider all horizontal transitions in the graph, we can compute the $S_1(I_1)$ function up to unknown scale and offset, at discrete points corresponding to the corners of the staircase. Similarly, we can

compute $S_2(I_2)$ at discrete points corresponding to the vertical transitions.

$$S_j({}^kI_j) = k \Delta S_{j(a \rightarrow b)} + S_j({}^0I_j), \quad k \geq 1 \quad (4.10)$$

where kI_j are the discrete input intensities corresponding to the steps of the staircase, and $S_j({}^0I_j)$ is the unknown offset for the projector radiance contribution, corresponding to the start of the staircase. Recall that $\Delta S_{j(a \rightarrow b)}$ is constant but also unknown – representing ambiguity in the scale factor.

In practice, the two isointensity curves are not explicitly identified beforehand. Instead, two specific camera-observed intensities are chosen to represent the isointensity curves. The staircase path is dynamically traversed during the calibration process, in which the corner positions of the ‘steps’ in the staircase are estimated on-the-fly. The estimates are based on the similarity between the camera-observed intensities at the corner positions and the two pre-selected intensities.

There is needless to say some impact from noise and quantization errors. There is a tradeoff between selecting pairs of isointensity curves that are close together or far apart. If the selected isointensity curves are close together, the recovered $S_j(\cdot)$ ’s are densely sampled; however the noise sensitivity is poor; in particular, the estimation of kI_j is done serially and leads to large accumulative errors when used to model $S_j(\cdot)$. If the curves are far apart, the noise sensitivity improves, but $S_j(\cdot)$ is sparsely sampled.

Recursive Bisection of Intensity Levels

In order to alleviate the problems associated with the staircase method, a method based on the recursive bisection of intensity levels is devised. In this method, we aim to determine a new isointensity curve that represents a radiance that is exactly the mean of the radiances of two previously determined isointensity curves.

(a)

(b)

The bisection process is illustrated in figure 4.4. Consider two isointensity curves Ψ_a and Ψ_b , and a fixed point $(\tilde{I}_{1a}, \tilde{I}_{2a})$ that is established on Ψ_a . We then consider the family of inscribed rectangles bounded by $(\tilde{I}_{1a}, \tilde{I}_{2a})$ and an opposite corner point $(\tilde{I}_{1b}, \tilde{I}_{2b})$ lying anywhere on Ψ_b , and denote the remaining two corner points of the rectangle as (I_{1ab}, I_{2ab}) and (I'_{1ab}, I'_{2ab}) . We can then make a further observation:

Observation 3 *Suppose Ψ_a and Ψ_b have radiances that differ by ΔM . If a corner point $(\tilde{I}_{1b}, \tilde{I}_{2b})$ on Ψ_b is found such that (I_{1ab}, I_{2ab}) and (I'_{1ab}, I'_{2ab}) lie on a third isointensity curve (i.e. (I_{1ab}, I_{2ab}) and (I'_{1ab}, I'_{2ab}) map to the same camera pixel intensity), then this new isointensity curve Ψ_c has a radiance that is $(\Delta M)/2$ different from the radiances of Ψ_a and Ψ_b . This is because the transition $(\tilde{I}_{1a}, \tilde{I}_{2a}) \rightarrow (I_{1ab}, I_{2ab})$ results in the same radiance change as the transition $(I'_{1ab}, I'_{2ab}) \rightarrow (\tilde{I}_{1b}, \tilde{I}_{2b})$, and the transitions link Ψ_a to Ψ_c to Ψ_b .*

This is shown in figure 4.4(b). In the next stage, a new point $(\tilde{I}_{1c}, \tilde{I}_{2c})$ on Ψ_c is recovered such that *both* the rectangle inscribed between Ψ_a and Ψ_c , as well as the rectangle inscribed between Ψ_c and Ψ_b , satisfy the isointensity condition described in Observation 3. This leads to a further recovery of two intermediate isointensity curves (figure 4.4 (c)), such that the difference in radiance between adjacent isointensity curves are equally $(\Delta M)/4$.

This recursive procedure may be carried out further until the desired level of resolution for the radiance is reached. By considering the generated set of steps in figure 4.4 (d) (similar to the ‘staircase’ in the previous section, but in an orthogonal direction), the projector-to-screen intensity transfer functions may be recovered

$$S_j(kI_j) = k \frac{\Delta S_{j(a \rightarrow b)}}{2^B} + S_j(\tilde{I}_{ja}) \quad (4.11)$$

where B is the number of bisections carried out, and $\Delta S_{j(a \rightarrow b)} = S_j(\tilde{I}_{jb}) - S_j(\tilde{I}_{ja})$ captures the unknown scale factor.

Unlike the staircase method, there is no tradeoff between the sampling density of $S_j(\cdot)$ and noise sensitivity. The recursive bisectional estimation of ${}^k I_j$ effectively isolates errors in one part of the function from propagating to the other parts. Hence the number of bisections can be extended to the desired resolution.

Gridline Optimization

Despite the improved accuracy in the recursive bisection method, an important source of error remains in that only two points are used to validate each isointensity curve. A deeper analysis of the above approach will show that many more points can be used to validate each isointensity curve. Nevertheless, the recursive bisection approach is useful in obtaining a satisfactory solution with moderately low computational cost, or can be used as an initial guess to the next method described below.

Consider a set of variably placed horizontal and vertical gridlines in the I_1 - I_2 graph. Suppose the gridlines are placed to demarcate each of the ${}^k I_j$, where $S_j({}^{k+1} I_j) - S_j({}^k I_j)$ is constant. In other words, the gridlines are to be arranged such that the changes in radiance resulting from horizontal state transitions between any pair of adjacent vertical gridlines is exactly the same, and also equal to that resulting from vertical state transitions between horizontal gridlines. This leads to yet another observation:

Observation 4 *Consider the grid induced by the intersection of such horizontal and vertical gridlines, with the bottom-leftmost intersection point defined as the origin. For each set of points which have the same norm-1 distance from the*

origin, as measured in units of grid links, all the points in the set must lie on the same isointensity curve. This is because each upwards or rightwards transition of 1 grid link unit results in the same increase in radiance.

The isointensity set of points may also be intuitively thought of as complete chains of top-left to bottom-right diagonal neighbors, as illustrated in figure 4.5.

Figure 4.5: Gridline Intersections on Isointensity Curves. There exists a grid, of which the intersections all fall on a number of isointensity curves, among which adjacent curves have equal radiance differences.

In the presence of measurement noise, we can formulate an optimization problem, in which we minimize the variation of camera-observed intensities within

each supposed set of isointensity points, with respect to the changes in gridline placement. Suppose we have K horizontal and K vertical gridlines, we can define the set of all gridline positions as $\Gamma = \{^kI_1, ^kI_2 : 1 \leq k \leq K\}$. We can also define a set function $\Psi(m)$ on an integer distance m , such that the output is a set of camera-observed intensities $G(I_1, I_2)$ at gridline intersection points that are m norm-1 distance (in units of grid links) away from the origin. Thus

$$\Psi(m) = \{G(^kI_1, ^{m-k}I_2) : 1 \leq k \leq m\}, \quad 1 \leq m \leq 2K-1 \quad (4.12)$$

Based on Observation 4, the intensities within each output set of $\Psi(m)$ should all be the same, since their associated projector intensity states all lie on the same isointensity curve. However, due to incorrect placement of the gridlines and other noise, the intensities may differ from each other. The optimization problem may then be expressed as

$$\arg \min_{\Gamma} \left\{ \sum_{m=1}^{2K-1} \text{Var} \{ \Psi(m) \} \right\} \quad (4.13)$$

where $\text{Var}\{\cdot\}$ measures the variance of these intensities. There are $2K - 1$ unique isointensity sets that have more than one element, and the goal is to minimize the sum of variances in all the sets through adjusting kI_1 's and kI_2 's, which is intuitively thought of as “shifting” the gridlines.

Once the optimization is completed, we can define a similar projector-to-screen intensity transfer function as before:

$$S_j(^kI_j) = k \Delta S + S_j(^0I_j), \quad k \geq 1 \quad (4.14)$$

where ΔS is an arbitrary value to define the increase in radiance from either a horizontal or vertical transition between adjacent gridlines, and $S_j(^0I_j)$ corresponds to the initial but unknown projector-contributed radiance offset.

In order to minimize the energy function in (4.13), gradient-based approaches can be used. In our experiments, both the BFGS quasi-Newton method as well as the conjugate gradient descent algorithm were tested. Minimization is relatively fast, but somewhat sensitive to initialization – this is fixed by using the results from the recursive bisection method as the initial guess.

4.2.4 Computing Radiance Contribution from Binary Light Sources

Based on the earlier photometric calibration, we can compute the relative radiance contribution from binary light sources, i.e. light sources that are within our control to turn on and off for experimentation. These sources may include indoor lights, or more importantly, the relative radiance values of $S_j({}^0I_j)$ corresponding to the initial projector offsets. In the case when ${}^0I_j = 0$, then $S_j(0)$ corresponds to the *black offset*, i.e. the light emitted from the projector when input intensity is set to zero.

In order to determine the radiance contribution from a binary light source, we need to extend the previous experiments to scenarios with the source turned on and off. In particular, we need to recover projector input combinations that result in equal camera-observed intensities when the light source is on and off. For example, if we are attempting to estimate $S_1({}^0I_1)$, we need to match camera observed intensities resulting from a two-projector combination, and one with a projector 1 turned off. Based on the match, we can express the camera-incoming radiance equality via

$$S_2(I_{2\alpha}) = S_1(I_{1\beta}) + S_2(I_{2\gamma}) \quad (4.15)$$

where $I_{2\alpha}$, *etc.* express different projector intensity inputs that result in equal cam-

era intensities. Expanding via (4.14), we obtain

$$\hat{k}_\alpha \Delta S + S_2({}^0I_2) = k_\beta \Delta S + S_1({}^0I_1) + k_\gamma \Delta S + S_2({}^0I_2) \quad (4.16)$$

where k_β and k_γ are known and relate to a previously recovered projector combination on a known isointensity curve, while $\hat{k}_\alpha$ may either be known if $I_{2\alpha}$ happens to lie on a known gridline, or needs to be computed as a non-integer factor through interpolation otherwise. Hence the projector offset is simply

$$S_1({}^0I_1) = (\hat{k}_\alpha - k_\beta - k_\gamma) \Delta S \quad (4.17)$$

This establishes the projector offsets relative to the ΔS radiance change, which is still in an arbitrary scale.

For other binary light sources such as indoor lights, a very similar analysis applies.

4.2.5 Camera Calibration

In the gridline optimization method, a set of isointensity curves were determined that were equidistant apart in terms of radiance. Additionally, we can measure the actual camera observed intensities at the gridline intersection points that lie on these isointensity curves. Ideally the intensities of points within each curve should be identical, but in practice we need to take the average of these intensities. Hence we can recover the mapping

$$C^{-1}({}^m\hat{Z}) = m\Delta S + M({}^0I_1, {}^0I_2, A) \quad (4.18)$$

where ΔS is as in (4.14), $M({}^0I_1, {}^0I_2, A)$ is the radiance of the least-bright isointensity curve, and ${}^m\hat{Z}$ is the average camera observed intensity of the m -th isoin-

tensity curve (in order of brightness):

$${}^m\hat{Z} = \frac{1}{|\Psi(m)|} \sum_{Z_k \in \Psi(m)} Z_k \quad (4.19)$$

Additionally, the calibration of projector offsets and binary ambient light sources can augment (4.18) with additional radiance-to-camera-intensity mapping data points.

4.3 Colorimetric Calibration

4.3.1 Expansion of Framework

Figure 4.6 illustrates the physical process by which the light from the projector lamp interacts with the pixel value of a particular projector color channel, gets reflected from the screen, mixed with ambient light, and is subsequently observed by the camera as a pixel value in one of the camera color channels.

The spectral radiance of light from the projector lamp which passes through an arbitrary projector pixel for projection is given by $f(\lambda)$ where λ is the wavelength, and is further modulated by the spectral response $h_i(\lambda)$ of a projector channel i (one of R, G and B). The input color intensity value of the image pixel I_i controls the fraction of light emerging from the projector via the colorimetric radiance transfer function $w_i(\cdot)$, which is typically nonlinear and depends on the electronics of the projector. The spectral irradiance $P_i(\lambda)$ of the corresponding screen pixel may thus be expressed as

$$P_i(\lambda) = w_i(I_i)f(\lambda)h_i(\lambda) \quad (4.20)$$

Together with the ambient light $a(\lambda)$, the light from all 3 RGB projector channels is additively combined on the screen and reflected in the direction of the cam-

Figure 4.6: Color Light Path in a Projector-Camera System. This graph shows how light from an image pixel is modulated and transferred in the projector, subsequently projected onto the screen, reflected with added ambient light to be observed by the camera. The final observed image pixel intensity is obtained through camera modulation and transfer processes.

era with a spectral reflectance of $b(\lambda)$. The spectral irradiance received by the corresponding camera pixel, $E(\lambda)$, is therefore

$$E(\lambda) = \left[\sum_{i=R,G,B} w_i(I_i) f(\lambda) h_i(\lambda) + a(\lambda) \right] b(\lambda) \quad (4.21)$$

The incoming light is observed separately by the three RGB color channels at the camera pixel. Suppose the spectral response of a camera color channel j is denoted by $v_j(\lambda)$, the irradiant light power C_j arriving on the corresponding CCD sensor will be

$$C_j = \int \left[\sum_{i=R,G,B} w_i(I_i) f(\lambda) h_i(\lambda) + a(\lambda) \right] b(\lambda) v_j(\lambda) d\lambda \quad (4.22)$$

The light power is subsequently measured by the sensor and converted into a numerical pixel value M_j through the nonlinear intensity transfer function $c_j(\cdot)$

$$M_j = c_j(C_j) \quad (4.23)$$

For the purpose of colorimetric calibration, the intention is to obtain the mappings from (I_R, I_G, I_B) to (M_R, M_G, M_B) , which does not require the calibration of all the intermediate functions. One way to simplify the above analysis would be simply to roll all the above equations into a single function $g_j(\cdot)$

$$M_j = g_j(I_R, I_G, I_B) \quad (4.24)$$

This however does not exploit the factorization that is available to increase the efficiency and accuracy of calibration. Instead, a slightly more explicit simplification is used in this paper, where (4.23) is used in conjunction with

$$\begin{bmatrix} C_R \\ C_G \\ C_B \end{bmatrix} = \begin{bmatrix} H_{RR} & H_{RG} & H_{RB} \\ H_{GR} & H_{GG} & H_{GB} \\ H_{BR} & H_{BG} & H_{BB} \end{bmatrix} \begin{bmatrix} W_R \\ W_G \\ W_B \end{bmatrix} + \begin{bmatrix} A_R \\ A_G \\ A_B \end{bmatrix} \quad (4.25)$$

where:

$$H_{ij} = \int f(\lambda) h_i(\lambda) b(\lambda) v_j(\lambda) d\lambda, \quad (4.26)$$

$$A_j = \int a(\lambda) b(\lambda) v_j(\lambda) d\lambda, \quad (4.27)$$

$$W_i = w_i(I_i). \quad (4.28)$$

The matrix given by

$$\mathbf{H} = \begin{bmatrix} H_{RR} & H_{RG} & H_{RB} \\ H_{GR} & H_{GG} & H_{GB} \\ H_{BR} & H_{BG} & H_{BB} \end{bmatrix} \quad (4.29)$$

is the color sensitivity matrix which reflects the cross-channel sensitivity between projected and observed light power.

In this part of the chapter, colorimetric calibration refers to quantitatively recovering the $w_i(\cdot)$ functions for each projector color channel, the $c_j(\cdot)$ functions for each camera color channel, and the color sensitivity matrix $\mathbf{H}$.

The assumptions made in our framework are the same as previous photometric calibration stage except a further assumption is made here: the color channel transfer functions $w_i(\cdot)$ and $c_j(\cdot)$ are monotonically increasing functions (this is predominantly true except with minor deviation at the extreme range of intensities).

4.3.2 Near-Degeneracies in Colorimetric Calibration

Colorimetric calibration based on isointensity curves requires that the camera observations be sensitive to changes in the projector channel variations. In the theoretically general case, each camera color channel is sensitive to all projector color channels. This implies that the color sensitivity matrix $\mathbf{H}$ is a full matrix with non-zero elements.

In practice however, the per-channel spectral responses $h_i(\lambda)$ and $v_j(\lambda)$ are often highly selective in wavelength. As a typical example, the red projector channel emits only a very limited range of wavelengths that are associated with the color red, while the blue camera channel is sensitive to a small and almost mutually exclusive range of wavelengths associated with the color blue. Thus the blue camera channel is highly insensitive to changes in the red projector channel. This manifests as a near-diagonal matrix $\mathbf{H}$ where the off-diagonal terms H_{RG} , H_{RB} , etc. are very small.

In this near-degenerate condition, an isointensity plot involving say the camera G channel, the projector G channel and the projector B channel will show isointensity curves to be strongly asymmetric and aligned with a the projector G channel axis (*i.e.* will be closely parallel to horizontal axis as in figure 4.7 (b)).

From a practical standpoint, colorimetric calibration involving these two projector channels and this camera channel will be ill-conditioned, as large intensity variations in the projector B channel lead only to small, possibly negligible changes in the camera G channel.

In the following sections, we present two colorimetric calibration methods that are based on (i) using a single projector and camera if the $\mathbf{H}$ matrix is not severely degenerate, or if a real grayscale camera is used; and (ii) using two projectors where the second projector is only required to project grays, or simply a single projector and a variable dimming lamp.

4.3.3 Single Projector Calibration

In this section, a colorimetric calibration method is described which calibrates the projector and camera color channels in a pairwise manner. *Each calibration process involves calibrating **two** projector color channels, in **two** camera color channels.* An extended version of the gridline optimization method [53] is used. While the original gridline optimization method may be applied using only one camera color channel, the typical ill-conditioned color sensitivities demand that greater constraints be applied.

In a system comprised of a projector and a camera both with RGB channels, the projector to screen color intensity transfer functions are expressed in (4.25). However, since we are not interested in recovering W_i absolutely to some physical

(a)

(b)

Figure 4.7: Gridline Meshes of Isointensity Curves for Single-Projector Colorimetric Calibration (1). (a) and (b) shows gridline meshes based on camera-observed intensities in red and green channels respectively.

unit but only to a scale factor, we can absorb H_{ii} as part of the W_i . Hence (4.25) may be re-expressed as:

$$\begin{bmatrix} C_R - A_R \\ C_G - A_G \\ C_B - A_B \end{bmatrix} = \begin{bmatrix} 1 & K_{RG} & K_{RB} \\ K_{GR} & 1 & K_{GB} \\ K_{BR} & K_{BG} & 1 \end{bmatrix} \begin{bmatrix} W_{RR} \\ W_{GG} \\ W_{BB} \end{bmatrix} \quad (4.30)$$

where $K_{ij} = H_{ij}/H_{ii}$ and $W_{ii} = w_{ii}(I_i)$. Here $w_{ii}(I_i)$ is a new transfer function denoting a contribution of projector channel in the identical camera channel. If a grayscale camera is used, this $\mathbf{K}$ matrix reduces to a row vector.

Consider the calibration involving the R and G colors channels for both the projector and camera. The curves W_{GG} and $K_{RG}W_{GG}$ may be independently computed in the separate camera color channels through independent use of the gridline optimization method. However, it is clear from (4.30) that the curves are identical except that one is scaled by a factor K_{RG} . A similar analysis applies for W_{RR} and $K_{GR}W_{RR}$. This is illustrated in figure 4.8 (a, b).

An extension of the gridline optimization method is applied here. Suppose we have m_{RR} and n_{RG} gridlines for projector red and green inputs as observed in the camera red channel. The vectors of gridline positions denoted as $V_{RR} = \{^m_R I_R : 1 \leq m \leq m_{RR}\}$ and $V_{RG} = \{^n_R I_G : 1 \leq n \leq n_{RG}\}$ are shown in figure 4.7(a). The vectors of gridline positions in the camera green channel can be denoted as $V_{GR} = \{^m_G I_R : 1 \leq m \leq m_{GR}\}$ and $V_{GG} = \{^n_G I_G : 1 \leq n \leq n_{GG}\}$ where m_{GR} and n_{GG} are corresponding numbers of gridlines for projector red and green inputs as observed in the camera green channel. These gridlines are shown in figure 4.7 (b).

The extended gridline optimization method is based on the crucial observation that the sets of gridlines are *not independent*. For example, V_{RR} are essen-

(a)

(b)

Figure 4.8: Gridline Meshes of Isointensity Curves for Single-Projector Colorimetric Calibration (2). (a) and (b) are projector intensity to radiance functions in red and green channels.

tially samples of W_{RR} , while V_{GR} are samples of $K_{GR}W_{RR}$. The difficulty is that they are however not sampled at the same positions, as partly illustrated in figure 4.8. Specifically, if V_{RR} is given, the vector of spectrally-weighted radiances corresponding to the gridline positions can be written as $W_{RR} = \{i/m_{RR} : 1 \leq i \leq m_{RR}\}$ and thus the vector of radiances in camera green channel can be written as $K_{GR}W_{RR} = \{i \cdot K_{GR}/m_{RR} : 1 \leq i \leq m_{RR}\}$. However, the number of grid points in the projector red channel may be different from that in projector green channel, especially for a near-degenerate scenario. Thus we can derive a new vector of radiances corresponding to V_{GR} relating m_{RR} and n_{GG} : $K_{GR}\hat{W}_{RR} = \{j \cdot K_{GR}/n_{GG} : 1 \leq j \leq m_{GR}\}$. Through linear interpolation, we can derive the vector of gridline positions V_{GR} . In this way, V_{RR} and V_{GR} are explicitly linked. A similar analysis applies for V_{GG} and V_{RG} .

For the camera red channel, we can define a set function $\Psi_R(q)$ on an integer distance q , such that the output is a set of camera-observed intensities $g_R(I_R, I_G)$ at the intersection points of gridlines V_{RR} and V_{RG} that are q norm-1 distance (in units of grid links) and

bined cost function over $(V_{RR}, V_{GG}, K_{RG}, K_{GR})$:

$$\arg \min_{V_{RR}, V_{GG}, K_{RG}, K_{GR}} \left\{ \sum_{i,j=R,G}^{m_{jR}+n_{jG}-1} \sum_{q=1} \text{Var} \{ \Psi_i(q) \} \right\} \quad (4.32)$$

Once the optimization is completed, we can define the projector-to-screen color intensity transfer functions as before:

$$W_{ij}(^q I_j) = q \Delta W_i + W_{ij}(^0 I_j), \quad q \geq 1 \quad (4.33)$$

where $\Delta W_i \in \{1/m_{RR}, 1/n_{GG}\}$ and $i, j \in \{R, G\}$. $W_{ij}(^0 I_j)$ corresponds to the initial but unknown projector-contributed radiance offset. We can set it to zero when only a single projector is concerned. We are able to use the BFGS quasi-newton method or conjugate gradient descent to minimize the energy function form in (4.32). This optimization can be extended to RGB channels in projector and camera for a complete colorimetric calibration of the system in a very similar manner.

4.3.4 Dual Projector Calibration

In most circumstances the color sensitivity matrix will turn out to be severely degenerate, in which case it will be impossible to accurately self-calibrate a single projector as each color channel of the camera will be adequately sensitive to only one projector color channel. This means that neither the shape of the color transfer functions nor the K_{ij} terms can be computed.

In order to circumvent this problem, we introduce a second projector (or a variable dimming lamp) which has overlapping display area with the first projector.

Unless the intention is to fully calibrate the second projector as well, the calibration process may be substantially simplified by projecting only grays from the

second projector. If this is not the case, same-color channels in the projectors may be calibrated independently of other color channels, in a manner identical to a grayscale calibration. Otherwise this process is done by specifying one instead of three input values to set all RGB channels equally. Additionally, the color sensitivity matrix of the second projector may be ignored, such that (4.30) can be changed to

$$\begin{bmatrix} C_R - A_R \\ C_G - A_G \\ C_B - A_B \end{bmatrix} = \begin{bmatrix} 1 & K_{RG} & K_{RB} \\ K_{GR} & 1 & K_{GB} \\ K_{BR} & K_{BG} & 1 \end{bmatrix} \begin{bmatrix} W_{RR} \\ W_{GG} \\ W_{BB} \end{bmatrix} + \begin{bmatrix} Q_R \\ Q_G \\ Q_B \end{bmatrix} \quad (4.34)$$

where

$$\begin{bmatrix} Q_R & Q_G & Q_B \end{bmatrix}^T = \mathbf{q}(I_w) \quad (4.35)$$

in which $\mathbf{q}(I_w)$ refers to the intensity-to-radiance transfer functions for gray inputs I_w in the second projector as observed in the three separate camera color channels.

Instead of having a calibration process involving two camera color channels and two separate projector color channels, the new process also includes the gray “channel” of the second projector. Technically, (4.32) should be extended to

$$\arg \min_{V_{RR}, V_{GG}, V_{RW}, V_{GW}, K_{GR}, K_{RG}} \left\{ \sum_{i,j=R,G}^{m_{jR}+n_{jG}-1} \sum_{q=1} \text{Var} \{ \Psi'_i(q) \} \right\} \quad (4.36)$$

where V_{RW} and V_{GW} represent gridline positions for the gray inputs of the second projector as observed in the two different camera color channels, while $\Psi'_i(q)$ is a variation on $\Psi_i(q)$ in terms of involving the second projector input as well. Note that there is (rightly) no constraint that the intensity transfer functions of the second projector should appear the same shape in separate camera color channels, hence there are three sets of gridlines to optimize instead of two as in the previous section. However for practical reasons in order to reduce the optimization complexity, the gridlines are first optimized in pairs corresponding to the two

dominant projector inputs of each camera color channel – e.g. pairs of (V_{RR}, V_{RW}) and (V_{GG}, V_{GW}) are optimized separately. After the optimization, the K_{ij} entries for color sensitivity matrix of the first projector are then estimated.

4.3.5 Camera Colorimetric Calibration

Very similar to the camera photometric calibration approach proposed in section 4.2.5, calibration may be done in the same way for each color channel of the camera. Therefore, the camera output intensity and input radiance per color channel can be related by

$$c^{-1}({}^q\hat{M}_i) = q\Delta W_i + C({}^0I_j, {}^0I_i, A) \quad (4.37)$$

where ΔW_i is as in (4.33), $C({}^0I_j, {}^0I_i, A)$ is the radiance of the least-bright isointensity curve, and ${}^q\hat{M}_i$ is the average observed intensity of the q -th isointensity curve in camera channel i (in order of brightness):

$${}^q\hat{M}_i = \frac{1}{|\Psi_i(q)|} \sum_{k: M_i \in \Psi_i(q)} {}^kM_i \quad (4.38)$$

4.4 Calibrating Reflectance Map for Multiple Pixels

The previous sections describe a framework for photometric calibration on a per-pixel basis. While it is possible to do the calibration for all projector pixels separately, it is easier if we are able to use the photometric calibration result for a single pixel and apply to all pixels.

In our framework, we assume the display screen is a Lambertian reflectance surface. The Lambertian model [19] predicts that the light incident on a matte surface is scattered equally in all directions, and the total amount of light L reflected

at a specific surface point depends on its angle of incidence θ of the illumination and proportional to the $\cos$ of this angle, *i.e.* $L \propto \cos \theta$. Additionally, lens radiometry (e.g. $\cos^4$ law for thin-lens) is assumed to affect the camera irradiance as a linear scaling factor and are not separable from the simplified reflectances. These led us to the assumption that the variation of the combined projector-to-camera intensity transfer function is due to different simplified reflectances for different pixels on the screen, and the shape of the individual projector-to-screen intensity transfer function and screen-to-camera intensity transfer function is the same for all pixels. We can express this via a generalization of (4.7)

$$M_{x,y} = \sum_{j=1}^N \rho_j(x,y) S_j(I_j(x,y)) + A(x,y) \quad (4.39)$$

where $\rho_j(x,y)$ are simplified reflectances associated with each (x,y) pixel in the camera image. In this generalization, the total radiance $M_{x,y}$, projector intensities $I_j(x,y)$ and the ambient light are all dependent on pixel position.

Similarly, this generalization can be extended to color reflectance for a particular projector:

$$C_i(x,y) = \sum_{j=R,G,B} \rho_{ij}(x,y) W_{ij}(I_j(x,y)) + A_i(x,y); \quad (4.40)$$

where $\rho_{ij}(x,y)$ are the simplified reflectances associated with each (x,y) pixel in the camera image. In this generalization, the total radiance $C_i(x,y)$ and projector channel j input intensities $I_j(x,y)$ with respect to camera channel i are all dependent on pixel location.

A method for photometric calibration setups can be summarized as the following:

1. Turn off all projectors and capture a camera image which can be expressed

as $Z_0(x, y) = C(A(x, y))$ for photometric calibration setup and $M_i^0(x, y) = c(A_i(x, y))$ for colorimetric calibration setup.

2. For photometric calibration setup, turn on only one projector each time, setting a common intensity I to all input pixels, and capturing cameras images $Z_j(x, y) = \rho_j(x, y)$. For colorimetric calibration setup, turn on a projector, setting a common intensity I_j for the channel j and 0 for all other channels to all input pixels, and capturing camera images $M_i^j(x, y) = \rho_{ij}(x, y)$.
3. Using a common reference camera pixel (x_0, y_0) for which photometric calibration has been done and hence $\rho_j(x_0, y_0) = 1$ in photometric calibration setup, then

$$\rho_j(x, y) = \frac{C^{-1}(Z_j(x, y)) - C^{-1}(Z_0(x, y))}{C^{-1}(Z_j(x_0, y_0)) - C^{-1}(Z_0(x_0, y_0))} \quad (4.41)$$

and $\rho_{ij}(x_0, y_0) = 1$ in colorimetric calibration setup, then

$$\rho_{ij}(x, y) = \frac{c^{-1}(M_i^j(x, y)) - c^{-1}(M_i^0(x, y))}{c^{-1}(M_i^j(x_0, y_0)) - c^{-1}(M_i^0(x_0, y_0))} \quad (4.42)$$

These can be done for all projectors and camera-observed pixels.

As the end of this calibration, a grayscale reflectance map or a set of color-based reflectance maps are obtained from the point of view of the camera. This allows the photometric calibration for a single pixel to be done once and easily extended to all pixels.

4.5 Experiments and Results

4.5.1 Estimating Grayscale Intensity Transfer Functions

First experiments were conducted to compare our modeling of the projector-to-camera intensity transfer functions with the original raw data collected in a double projector single camera observation fashion. The modeled functions and the original raw data collected are specific to a single pixel on the screen display. With two projectors turned on, the input intensities to the projectors were varied from 0 to 255 respectively with an interval of 1 pixel intensity respectively. Camera observations were used to recover the projector-camera response space shown in figure 4.10 (a). A 64×64 grid is derived using the gridline optimization method derived in section 4.2.3. Figure 4.9 shows the computed projector and camera intensity transfer functions together with a comparison between the observed and computed projector-camera intensity transfer function for projector 1. The latter is shown in the bottom right graph of figure 4.9.

With all the intensity transfer functions computed, we were able to fully rebuild intensity transfer functions for the two projectors and the single camera as a 3D surface plot, as shown in figure 4.10 (b). Errors are plotted between the observed and computed intensity transfer functions (Figure 4.11). The mean error is 0.21027 intensity levels.

We wish to point out that the observed data does *not* represent error-free ground truth, due to the presence of temporally-varying projector and camera noise. We believe the estimated model provides factorization-based noise filtering for the transfer functions and may on average be a more accurate model than the observed data itself.

Figure 4.9: Computed Projector Camera Intensity Transfer Functions. The top two graphs show the computed projector intensity to radiance responses; the bottom left graph shows the computed camera radiance to intensity response; the bottom right graph compares the computed Projector1-Camera response (Blue) to the observed one (Green).

4.5.2 Estimating Color Channel Intensity Transfer Functions

First we conducted experiments to compare our modeling of the projector-to-camera color intensity transfer functions to the original raw data collected from a dual projector, single camera system. The modeled functions and the original raw data are specific to a single pixel on the display surface. The raw data was collected by letting each camera channel observe all the combinations of projector 1 displaying red, green and blue respectively while projector 2 displaying white

(a)

(b)

Figure 4.10: Double Projector – Single Camera Intensity Transfer Function Space. Graph (a) shows the observed projector-camera response space. Graph (b) is the computed response space using our model.

Figure 4.11: Error Plot for the Grayscale Intensity Transfer Function Space. The errors between the observed and the computed projector-camera response space is shown above with a mean of 0.21027 intensity levels.

from 0 to 255 by an increment of 3. Using an extended gridline optimization approach for the dual projector calibration, the projector and camera color channel intensity transfer functions are obtained, with linear interpolation between gridline sample points. Figure 4.12 (a-c) shows the projector channel intensity to radiance responses in the observation of each camera channel while figure 4.12 (d-f) shows the camera channel radiance to intensity responses. The matrix $\mathbf{K}$ obtained in this experiment was:

$$\mathbf{K} = \begin{bmatrix} 1 & 0.4017 & 0.0340 \\ 0.0173 & 1 & 0.0630 \\ 0.0225 & 0.0485 & 1 \end{bmatrix}$$

Table 4.1 shows the mean difference between our computed and observed camera outputs. Again we would like to emphasize that the camera-observed data may be substantially affected by the temporally-varying projector and cam-

4.5 Experiments and Results

91

Figure 4.12: (a,b,c) shows the estimated projector color intensity to radiance responses in the observation of respective camera channels; (d,e,f) shows the estimated radiance to intensity response of each camera channel.

Table 4.1: Mean Difference between Computed and Observed Camera Outputs (in pixel intensity levels)

	Proj R	Proj G	Proj B	Proj W
Cam R	0.1178	-0.7057	-0.1767	-0.2549
Cam G	1.1800	0.0695	2.8438	1.3644
Cam B	1.6458	1.6354	0.16973	1.1503

era noise and can not represent fully an error-free ground truth. In our view, the estimated model from our method provides a factorization-based noise filtering for the intensity transfer functions, and thus may be more accurate in the ground truth representation than the observed data itself.

4.5.3 Prediction of Camera Output Based on Known Projector Input

From our modeling result, we can proceed to predict the camera observation of projector displays after computing a reflectance map for each of the projectors. In our experiments, a camera image of the two projectors projecting a ramp image is shown in figure 4.13(a). Using the computed intensity transfer functions, a synthesized image in the camera view is generated and shown in figure 4.13 (b). For this image, there is a mean estimation error of 0.9589 pixel intensity levels. The pixel-by-pixel error map is shown in figure 4.13 (c). In figure 4.14, two projectors are simultaneously projecting the same image content but with different input pixel intensity levels computed from our model, such that each half of the radiance into the camera to form the observed image was from either of the two projectors. Figure 4.14 (a) shows the camera observation of this texture image

while (b) shows the predicted image based on our model.

In our setup, we calibrated the projectors and cameras using only homographies. However, there appears to be some amount of radial distortion present in the camera image resulting in small but noticeable displacement of corresponding pixels. This has led to slightly higher pixel-to-pixel errors, caused mainly by registration errors rather than errors in photometric calibration.

Per channel transfer functions for projectors and the camera are recovered using our colorimetric calibration model. Additionally color reflectance maps are computed for each projector channel in the observation of each camera channel. Color images now can be predicted from the camera's point of view based on the previous calibration results. An experiment to predict the camera image when a projector is displaying a horizontal spectrum map with large vertical variations of intensities is shown in figure 4.15 and figure 4.16. The observed and predicted images are shown in figure 4.15 (a,b) respectively. Figure 4.17 shows the absolute estimation error map across the display while the units of error are in pixel intensity levels. Furthermore, another camera observed image and corresponding predicted image are shown in figure 4.16 (a,b) respectively. Projective transformations were only used to calibrate the projectors and cameras geometrically, however, noticeable misalignment of corresponding pixels in the display occurs due to some amount of radial distortion in the camera images. This may results in slightly higher pixel-level errors when comparison are made between the observed images and the predicted images, of which the main cause lies in the geometric registration errors rather than photometric calibration errors.

(a)

(b)

(c)

Figure 4.13: Camera Observation Prediction over Single/Double Projectors Display (1). Patches of 280×210 pixels of the original images are shown here for close views. Graph (a) shows camera observation of projected images from both projectors with the same input, while Graph (b) shows the computationally synthesized images. The estimation errors for pair (a,b) were shown in (c).

(a)

(b)

Figure 4.14: Camera Observation Prediction over Single/Double Projectors Display (2). Patches of 280×210 pixels of the original images are shown here for close views. Graph (a) shows camera observation of projected images with half contribution from each projector, while Graph (b) shows the computationally synthesized images.

(a)

(b)

Figure 4.15: Prediction of Camera Image (1). Patches of size 220×210 pixels of the original 768×576 images are shown here in close-up. Graph (a) shows the camera observation of projected color ramp image and Graph (b) shows the predicted images.

4.5 Experiments and Results

97

(a)

(b)

Figure 4.16: Prediction of Camera Image (2). Patches of size 220×210 pixels of the original 768×576 images are shown here in close-up. Graph (a) shows the camera observation of projected color ramp image and Graph (b) shows the predicted images.

Figure 4.17: Prediction error maps of camera images (shown in figure 4.15(a,b)) in red, green and blue channels with finally a mean error plot over all three camera channels are shown above.

4.5.4 Mixed and Compensating Projector Display for Camera Observations

Further experiments were carried out to explore how different combinations of projector inputs can form a uniform camera image. Given a camera observation of an arbitrarily projected textured image (shown in figure 4.14 (a)), the radiance contribution per pixel can be computed from the camera transfer function. Subsequently different contributive proportions of projector radiance were assigned to be delivered by the two projectors. After computing the exact intensity level per

pixel in each projector, the resultant images were projected, and observed by the camera. Figure 4.18 shows the result of mixing displays with a radiance mixing ratio of 7:3. In a second experiment, the method was applied to allow one projector compensate for the non-uniform projection by another projector, such that a uniform display is generated. A sine wave stripe image is projected from projector 1 in figure 4.19 (a). After adding in the computed contribution from projector 2, a uniform image was obtained and observed by the camera. The mean estimation error is 0.6477 pixel levels.

Figure 4.18: Camera Observation of Double-projector Mixed Displays (1). Patches of 280×210 pixels of the original images are shown here for close views. This graph shows a 7:3 mixing of the two projector to recreate figure 4.14(a).

Based on colorimetric calibration result, a color image can be projected from one projector to mix or compensate for an image from another projector. The goal of this experiment is to create a desired color image to be observed in the

(a)

(b)

(c)

Figure 4.19: Camera Observation of Double-projector Mixed Displays (2). Patches of 280×210 pixels of the original images are shown here for close views. (a) shows a camera observation of projector 1 showing a sine wave gray image while (b) shows a uniform camera observation by adding in computed projector 2 contribution. Estimation errors are shown in (c).

camera, by modifying the output from one of the projectors to compensate for the projected image from the other projector. To modify the projector output, radiance contribution map from one projector for compensation must be computed from camera radiance-to-intensity transfer function with the reflectance map, and subsequently modified input image can be computed through projector intensity-to-radiance transfer function pixel by pixel.

This experiment involves creating a camera observed image with nearly uniform intensities. The color images projected by projector 2 is observed by the camera, as shown in figure 4.20 (a). In order to display a desired color image (figure 4.20 (c)) on the projected background, a compensation image computed from the estimated color transfer functions (figure 4.20 (b)) is projected by projector 1. Finally a camera observed image is shown in figure 4.20 (d). In figure 4.22 an error map measuring the difference between the desired image and the actual observed image with respect to each color channel and grayscale intensity of the camera. In the following experiment, a more complicated color background image (figure 4.21 (a)) is projected by projector 2, while a compensation image (figure 4.21 (b)) is projected by projector 1 to produce an image of the same content as in figure 4.16 (a,b). The desired and actual camera images are shown in figure 4.21 (c) and (d) respectively.

4.6 Conclusion

In this chapter, a principled framework for grayscale photometric calibration of multiple overlapping projectors was presented involving the analysis of iso-intensity curves. Three photometric calibration approaches were proposed: the staircase method, the recursive bisection method, and the gridline optimization

Figure 4.20: Color Compensation Experiment (1). (a) A color image projected by projector 2 alone; (b) the *compensation* image for projector 1, computed to create desired display image in (c); (d) the camera observed image based on the projection of compensation image on the 1st color image.

Figure 4.21: Color Compensation Experiment (2). (a) A color image projected by projector 2 alone; (b) the *compensation* image for projector 1, computed to create desired display image in (c); (d) the camera observed image based on the projection of compensation image on the 1st color image.

Figure 4.22: Plots of error maps between figure 4.20 (c) and (d) are shown for each camera channels and overall camera outputs.

method. An extension to the framework for colorimetric calibration of projector camera systems was presented. Two calibration methods proposed were proposed for both single-projector and dual-projector setups.

None of these methods require explicit parametric models for calibration nor impose smoothness constraints, nor do they require expensive equipment; an inexpensive camera with minimal controls and a small dynamic range may be used. Results were shown demonstrating that the recovered intensity transfer functions are able to accurately model the raw data, and can be used in two ways – predicting the camera observed image based on known projector inputs, and setting correct input intensities to compensate a textured background for a desired

camera-observed image, which successfully solves the problem of artifacts from color imbalance in overlapped displays.

Chapter 5

Projector Out-of-Focus Blur Analysis and Compensation

5.1 Image Preconditioning

5.1.1 Blur Compensation Based on Image Preconditioning

Often a camera recorded image is a degraded version of an original scene. For example, the degradation may be due to motion blur, or defects in the lens. In this chapter, we will be focusing on the degradation of an image due to out-of-focus blur. This process can be described as $I_2 = h(I_1)$, where $h(\cdot)$ is an degradation function, I_1 and I_2 are the original and degraded images respectively. Usually for camera observed degraded images, enhancement or restoration techniques can be applied to improve the degraded image or recover the original image. An enhancement/restoration function can be denoted as $g(\cdot)$, a resultant image I_3 can

be obtained through $I_3 = g(I_2) = g(h(I_1))$. In the case of out-of-focus blurring, the function is well-modeled as a convolution of the ideal image with a blurring PSF. Interestingly, the commutation characteristics of convolution allows us to change the order of input. The convolution of two signals is defined as $f_1(t) * f_2(t) = \int_{-\infty}^{\infty} f_1(\tau)f_2(t - \tau)d\tau$, if let $x = t - \tau$, then $\tau = t - x$ and $d\tau = -dx$, subsequently

$$\begin{aligned}
 & f_1(t) * f_2(t) \\
 &= \int_{-\infty}^{\infty} f_1(\tau)f_2(t - \tau)d\tau \\
 &= -(\int_{-\infty}^{\infty} f_1(t - x)f_2(x)dx) \\
 &= \int_{-\infty}^{\infty} f_2(x)f_1(t - x)dx \\
 &= f_2(t) * f_1(t)
 \end{aligned} \tag{5.1}$$

Because of the commutative property of convolution, the order of applications of $h(\cdot)$ and $g(\cdot)$ can be interchanged, *i.e.*, the operations of degradation and enhancement/restoration are commutative through convolution, the same resultant image I_3 can be obtained through $h(g(I_1))$. In this scenario, the original image vector I_1 is *preconditioned* by $g(\cdot)$ before going degradation process $h(\cdot)$. This is exploited in the research discussed in this chapter. Based on image preconditioning, two formulated methods will be proposed in the following parts.

Uniform PSFs and Wiener Filtering

We initially consider the scenario in which a projector projecting orthogonally to a flat display surface is out of focus. In this setup, the display is uniformly blurred as the PSF (on the display) is assumed to be reasonably invariant to the spatial position of the associated pixel in the image.

While the PSF depends on the lens system, it can be reasonably modeled as a

2D circular Gaussian [18] of the form

$$h_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} . \quad (5.2)$$

where σ is the standard deviation of the Gaussian Function, (x, y) denotes the pixel location in an image. The blurred image created from the overlap of the uniform PSF from different pixels can be modeled as the result of a convolution

$$\begin{aligned} i_B(x, y) &= i(x, y) \circ h(x, y) \\ &= \sum_{s=-m/2}^{m/2} \sum_{t=-n/2}^{n/2} i(x-s, y-t) h(s, t) , \end{aligned} \quad (5.3)$$

where $i(x, y)$ and $i_B(x, y)$ are the original and blurred images respectively, and (m, n) defines the size of the convolution kernel. Additionally, some additive noise may be present. In image processing, a typical problem is to recover the original but unknown image $i(x, y)$ given only the blurred image $i_B(x, y)$. If (5.3) is valid, the deblurring may also be achieved via convolution with an inverse filter $h^{-1}(x, y)$ such that

$$\begin{aligned} \hat{i}(x, y) &= i_B(x, y) \circ h^{-1}(x, y) \\ &= [i(x, y) \circ h(x, y)] \circ h^{-1}(x, y) , \end{aligned} \quad (5.4)$$

where $\hat{i}(x, y)$ is the estimated deblurred image, assuming that $h^{-1}(x, y)$ exists and the noise is small.

In our problem, the sequence of operators is different. Here the goal is to *pre-condition* the known original image such that when it is displayed via the out-of-focus projector, the output image appears similar to the original image. Because convolution operators are commutative, (5.4) may be rewritten as

$$\hat{i}(x, y) = [i(x, y) \circ h^{-1}(x, y)] \circ h(x, y) . \quad (5.5)$$

We can consider the preconditioned image to be the first term of (5.5), defined as

$$\tilde{i}(x, y) = [i(x, y) \circ h^{-1}(x, y)] . \quad (5.6)$$

Thus, the preconditioned image $\tilde{i}(x, y)$ after degradation $h(x, y)$ is an approximation of the original image $\hat{i}(x, y)$.

The challenge here is determining the optimal $h^{-1}(x, y)$, and this is easiest done in the frequency domain, where the blurring process may be dually treated as

$$I_B(u, v) = I(u, v)H(u, v) , \quad (5.7)$$

where the $I_B(\cdot)$, $I(\cdot)$ and $H(\cdot)$ functions are Fourier transforms of the $i_B(\cdot)$, $i(\cdot)$ and $h(\cdot)$ functions respectively. If the PSF is known, a classical solution that minimizes the mean squared error is Wiener filtering (see [21]), for which a simplified variation is:

$$\hat{I}(u, v) = \frac{H^*(u, v)I_B(u, v)}{|H(u, v)|^2 + 1/S} , \quad (5.8)$$

where $\hat{I}(\cdot)$ is the Fourier transform of $\hat{i}(\cdot)$, $H^*(\cdot)$ is the complex conjugate of $H(\cdot)$, and S is the estimated (or apriori) signal-to-noise ratio. Hence the preconditioning filter that is used for uniform PSF is simply given by

$$h^{-1}(x, y) = \mathcal{F}^{-1} \left\{ \frac{H^*(u, v)}{|H(u, v)|^2 + 1/S} \right\} , \quad (5.9)$$

where $\mathcal{F}^{-1}$ is simply the inverse Fourier transform.

Considering (5.6), (5.8), and (5.9), it is apparent that the preconditioned image, $\tilde{i}(x, y)$ can be obtained by applying the Wiener filtering to the original image, $i(x, y)$, with H such that:

$$\mathcal{F}^{-1}\{\tilde{I}(u, v)\} = \mathcal{F}^{-1} \left\{ \frac{H^*(u, v)I(u, v)}{|H(u, v)|^2 + 1/S} \right\} . \quad (5.10)$$

where $\tilde{I}(\cdot)$ is the Fourier transform of $\tilde{i}(\cdot)$.

Assuming that the PSF is known or can be estimated from test images (*e.g.* fiducial markers), the Wiener filter allows for the preconditioning of images for out-of-focus projectors that are projecting orthogonally to the display surface.

Non-Uniform PSFs

When the projector is skewed to the display surface or for curved surfaces, the PSF is non-uniform across the projector image. One significant consequence of this is that the convolution model no longer applies, and Wiener filtering cannot be directly used to precondition the image.

To address this problem, we estimate the spatially varying PSF profile across the projector. While we would like to estimate the PSF for each projector pixel, this is impractical in practice. As a compromise, we partition the projected image into smaller regions within which a PSF is computed. These sub-sampled PSFs are used to compute our preconditioned $\tilde{i}(x, y)$ by compositing a series of global PSF compensations described in the following section.

5.1.2 Limitations of Image Preconditioning

As with restoration of blurred images, image preconditioning takes into account the PSF and the input image. Intuitively preconditioning may be thought of as attempting to sharpen the original image. However, there are limitations when image preconditioning is applied.

The main constraint are the quantization/saturation limits of the input image. The values of image pixel array $I(x, y)$ is quantized and limited to within the range of (0-255). However, the inverse Fourier Transform of the Wiener-filtered frequency spectrum may not necessarily fall into this range and have negative

values or values above 255.

Thus to solve this problem, our assumption is, firstly, the preconditioned spatial image derived from Fourier Transform of its frequency spectrum has to be clipped if there are pixel values beyond the range of (0-255). Secondly, the preconditioned image is derived based on the discrete input pixel array for an approximation of the original continuous image. With these assumptions, we proceed to propose two approaches for projector-based display deblurring.

5.2 Image Preconditioning using an Observation-based Approach

5.2.1 Blur Estimation

Our framework begins by estimating piecewise PSFs in the projector's image. The projector displays an image of equally sized feature markers (crosses) in an off-axis manner onto a flat surface. A high-resolution camera captures an image of these projected features. Since the displayed features and their observed locations in the camera are known, we can compute the 3×3 homography matrix between the camera and display image to rectify the camera image to the original displayed image.

To derive the PSFs we would ideally compare the original image with the rectified camera image. These two images, however, are sufficiently different due to a variety of effects including the devices' imaging systems, display surface response, and properties such as the projector's lamp age and color balance settings. To align these two images together, we need to calibrate the projectors and camera geometrically and photometrically using the methods proposed in chapter 3

and chapter 4 respectively. As mentioned in chapter 4, homographic calibration may be used to rectify the camera image for comparison with the original image. However, geometric calibration is not sufficiently accurate and flexible enough to model this transformation due to reasons such as non-radial or non-tangential projector lens distortions, and not perfectly planar display surfaces, *etc.* Instead, the PSFs are derived based solely on analyzing the camera observed image. Additionally, while we are able to carry out photometric calibration as described in chapter 4, it turns out that a simple intensity normalization process can replace the photometric calibration in the blur estimation to provide the same result.

The idea of this image-based approach is to locate the most in-focus observed feature and use this rectified image as an *exemplar* for determining the PSFs of the other features. Since the camera image is rectified to the original projected image, the locations of the features are known. For each feature in the rectified camera image, a numbered sub-image (bounding box) is assigned encompassing it. These subimages are of the same size and defined as $i_{f_j}(x, y)$, where x, y denotes the pixel location within the sub-image, and j denotes the j^{th} sub-image.

Due to lighting variations within the projector and illumination fall off from off-axis projection, intensity responses across the displayed image are not uniform. Feature image's intensity need to be adjusted according to the photometric calibration result used in chapter 4. For simplification purpose in this image-based method, we manage to get away with calibration by just normalizing the features' intensities first before finding the exemplar feature. In our setup, the illuminated display surface exhibits a reasonably uniform response to the projected light. As a result, we can exploit the nature of the PSFs to perform the intensity normalization. For display surfaces with non-uniform responses, more sophisticated il-

lumination compensation approaches proposed in chapter 4 or other works such as [36] can be used.

For a geometrically rectified image, the sum of pixel intensities within one sub-image may be different from another sub-image. This is because of the brightness fall-off effect in an obliquely projected image. Based on the assumption that the out-of-focus blur has no contribution to this brightness fall-off effect, the sum of pixel intensities within each sub-image should be identical. Therefore, by normalizing the pixel intensities within each sub-image, the brightness fall-off effect can be removed. This allows us to solely model the out-of-focus blur that is brightness invariant.

The Gaussian PSF used in our blur model sums to unity and therefore does not change the overall energy of the original signal, *i.e.*, it does not change the DC component of the original $I(u, v)$. In other words:

$$I_B(0, 0) = I(0, 0)H(0, 0) = I(0, 0) ,$$

where I_B is the frequency spectrum of blurred images, and the index $(0, 0)$ represents the DC component of each I , I_B and H functions in the Fourier domain. This makes the normalization of all features possible while maintaining the sum of energy of signals as unity. First, the brightest feature i_{max} among all the sub-images can be derived by

$$\begin{aligned} i_f(j) &= \sum_{x=-m/2}^{m/2} \sum_{y=-n/2}^{n/2} i_{f_j}(x, y) \\ i_{max} &= \max i_f(j) \end{aligned} \quad (5.11)$$

where m, n denotes the number of horizontal and vertical pixels in a sub-image. Subsequently all other features $i_{f_j}(x, y)$ can be normalized as

$$i'_{f_j}(x, y) = \mathcal{F}^{-1}\{I_N(u, v)\} , \quad (5.12)$$

where

$$I_{f_j}(u, v) = \begin{cases} I_{max}(0, 0) & \text{if } u = v = 0 \\ I_N(u, v) & \text{otherwise.} \end{cases}$$

where $I_{f_j}(u, v)$ are the Fourier Transforms of $i_{f_j}(x, y)$ and $I_{max}(0, 0)$ is the DC component of the Fourier Transform of the sub-image corresponding to i_{max} . From (5.12) all the features are transformed to have the same DC component as the brightest feature.

After normalization, the sharpest feature in the image is found by computing a sharpness response in a block-wise fashion about each feature, $i'_{f_j}(x, y)$, using the Tenengrad operator [58] as follows:

$$T_j = \frac{1}{m+n} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} [s_h^2(x, y) + s_v^2(x, y)] \quad (5.13)$$

where T_j is the sharpness response for feature $i'_{f_j}(x, y)$, $s_h^2(x, y)$ and $s_v^2(x, y)$ are 5×5 horizontal and vertical Sobel filter responses applied in the spatial domain over all $(m+n)$ pixels composing the feature $i'_{f_j}(x, y)$.

Figure 5.1 shows the steps to find the exemplar feature. Figure 5.1 (a) shows the original input image captured by the camera. This image is rectified to the projected image (figure 5.1 (b)), and then normalized (figure 5.1 (c)). Sharpness responses computed using (5.13) are obtained for each block as shown in figure 5.1 (d). Our exemplar feature, $i_e(x, y)$ is taken to be the feature corresponding to $\max(T_j)$.

PSF Map Recovery

Given the exemplar template, $i_e(x, y)$, we compute a set of k blurred templates with increasing σ , such that $i_{e(\sigma_k)}(x, y) = i_e(x, y) \circ h_{\sigma_k}(x, y)$, where $h_{\sigma_k}(x, y)$

Figure 5.1: Stages of our PSF estimation. (a) Camera image of displayed features; (b) rectified image; (c) intensity normalized; (d) sharpness response for each feature.

represents the Gaussian PSF described in (5.2) with parameter σ_k . Typical values of σ_k are $0.5, 1, 1.5, \dots, 4$. These blurred $i_{e(\sigma_k)}(x, y)$ serve as templates for estimating the PSFs across the projector image. Cross correlation can be applied for each projected feature $i'_{f_j}(x, y)$ against all blurred templates, $i_{e(\sigma_k)}(x, y)$, to find most similar $i_{e(\sigma_k)}(x, y)$ for each feature. However, since the Tenengrad responses, T_j for each $i'_{f_j}(x, y)$ are already available from the exemplar search, we therefore make use of the available Tenengrad responses for each $i_{e(\sigma_k)}(x, y)$ and

Figure 5.2: Estimated PSF $M_j(x, y)$. Notice the shape of this PSF map appears as the inverse of the Tenengrad responses.

for matching PSFs.

The final result is a PSF map $M_j(x, y)$ that assigns the appropriate σ_k to each feature $i'_{f_j}(x, y)$ based on the template matching. The σ_k associated with each $M_j(x, y)$ corresponds to the PSF $h_{\sigma_k}(x, y)$ which best approximates the blurring in that region. Figure 5.2 shows the resulting $M_j(x, y)$. Not surprisingly, the shape of this map appears as the inverse of the Tenengrad responses. The σ at each pixel location (x, y) can be thus estimated from σ_k through bilinear interpolation and thus each pixel in the image can have a σ assigned. However, repeated operations will be carried out for the same σ among pixels, and the size of the image determines that the computational cost would be very high. Therefore, a faster way through image composition based on PSF map is proposed in the following sections.

5.2.2 Computing the Preconditioned Image

Basis Images via Wiener Filtering

As mentioned in section 5.1.1, because our PSFs are varying spatially within the image, Wiener filtering cannot be applied in a global manner to derive the preconditioned image $\tilde{i}(x, y)$. As a compromise, we approximate a spatially varying Wiener filter given the PSF maps $M_j(x, y)$.

Figure 5.3: Piecewise PSF Filtering: (Top) Basis images $\tilde{i}_{\sigma_k}$. (Bottom-left) Shows PSF map and four nearest neighbors to pixel (x, y) . (Bottom-middle) Shows zoomed in regions of four basis images. (Bottom-right) Final composited image.

The $M_j(x, y)$ has k distinct PSFs defined as $h_{\sigma_k}(x, y)$. Using these $h_{\sigma_k}(x, y)$, we compute a set of preconditioned *basis images*, $\tilde{i}_{\sigma_k}(x, y)$, using Wiener filtering as described in (5.10), where the filter H for (5.10) is $\mathcal{F}\{h_{\sigma_k}(x, y)\}$. Figure 5.3 (top) shows an example of these basis images.

Image Compositing

For a given pixel in the preconditioned image, $\tilde{i}(x, y)$, we compute its value using a bilinear interpolation of the basis images $\tilde{i}_{\sigma_k}(x, y)$. The appropriate basis images and weights for the interpolation are determined from the PSF $M_j(x, y)$.

Performing the appropriate coordinate scaling, we find the four closest neighbors in the PSF $M_j(x, y)$ to pixel (x, y) . These four neighbors are denoted as $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4$ and are ordered in a clockwise fashion about (x, y) . Letting $\mathbf{m}(\sigma)$ refer to the $\mathbf{m}$'s corresponding σ value, the interpolation is written as:

$$\begin{aligned} \tilde{i}(x, y) = & (1 - t)(1 - s) \tilde{i}_{\mathbf{m}_1(\sigma)}(x, y) \\ & + (1 - t)s \tilde{i}_{\mathbf{m}_2(\sigma)}(x, y) \\ & + ts \tilde{i}_{\mathbf{m}_3(\sigma)}(x, y) + t(1 - s) \tilde{i}_{\mathbf{m}_4(\sigma)}(x, y) \end{aligned} \quad (5.14)$$

where $s, 1 - s, t, 1 - t$ are the appropriate barycentric coefficients, ($s, t \in [0..1]$), in the horizontal and vertical directions between the (x, y) location and the centers of the features associated with $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4$. Performing this interpolation for each pixel we obtain the desired preconditioned image $\tilde{i}(x, y)$ needed for projection.

5.2.3 Experiments and Results

Experiments are performed using a 3M MP8749 portable LCD projector with (1024×768) resolution, an Olympus C760 digital camera with 3.2M pixels and 10x optical zoom and a IBM Intellistation M Pro. The algorithms are all implemented in unoptimized Matlab 7.0 code.

For our experiments, we project a grid of 12×16 crosses. Features are bounded by 64×64 pixels blocks. Eight PSFs are estimated using $\sigma_k = 0.5, 1, 1.5, \dots, 4$ as

discussed in Section 5.2.1. When computing the basis images, a Signal-to-Noise ratio of 0.01 is provided in the Wiener filter to estimate noise present in the degradation process. Computation time, including estimating the PSFs, constructing the basis images, and compositing the final preconditioned image is around 3-5 minutes.

We have selected test images that are sufficiently in focus and hope to demonstrate that results from our algorithm are not merely attributed to sharpening the input image. Please note that the preconditioned images will inherently appear sharper than the original, however, the original images themselves are sharp.

Figure 5.4 and figure 5.5 show an example of our approach. Figure 5.4 (top) shows the original image of a “cat” and figure 5.5 (top) shows its appearance after projection. Note the out-of-focus blur appearing in the left-bottom corner. Figure 5.4 (bottom) is the corresponding preconditioned image $\tilde{i}(x, y)$ and its result, figure 5.5 (bottom), after projection. The texture of the cat’s fur appears sharper in the preconditioned image (zoomed region). Figure 5.6 and figure 5.7 show similar results for an outdoor scene. Notice the zoomed region show the preconditioned image appearing sharper than the unprocessed image.

Figure 5.8 compares the results of our approach as an inset into the original projected image. Note that textures in the blurred regions are better preserved in the preconditioned image than the unprocessed projected image.

The above results are subjective. Given the nature of the projector-camera system it is hard to compute quantitative results. As an effort to motivate our approach the following comparison is made. In this experiment, a set of blurred basis images are first generated from an original image i using the PSF map computed in section 5.2.1. Subsequently a blurred image can be composited from the

blurred basis images using the method in section 5.2.2. This process is denoted as a synthesized blurring process $B(\cdot)$, and the original image under blur can therefore be expressed as $B(i)$. If instead of the original image, a preconditioned image $\tilde{i}$ is sent into the blurring process, the output image may be expressed as $B(\tilde{i})$. Differences in the form of residual can be derived by the subtraction of original and preconditioned images under synthesized blur from the original images respectively, in mathematical forms: $i - B(i)$ and $i - B(\tilde{i})$. Subsequently, the norms of these absolute differences are shown respectively and compared in the table below to indicate performance improvement of our method.

Figure	$\ i - B(i)\ $	$\ i - B(\tilde{i})\ $	Improvement
Colosseum (figure 5.6 and 5.7)	22204	21030	+5%
Cat (figure 5.4 and 5.5)	12217	12094	+1%
Temple (figure 5.8 left)	20621	18163	+13%
Castle (figure 5.8 right)	25806	23557	+9%

Here we focused solely on an off-axis projector to demonstrate our approach. In practice, our approach can be used with any display surface geometry. The only requirement is that the camera image of the displayed features be rectified back to the projector's coordinate frame. Several of the geometric calibration techniques (see [8]) provide methods for this rectification on non-planar surfaces.

While our Gaussian model is a reasonable approximation of projector blur, a formal investigation into image creation via projector systems and deeper analysis into projector specific PSFs would be favored. In the next section, we investigate and propose another image preconditioning approaches based on the projector optical blurring model.

Original Image

Preconditioned Image

Figure 5.4: [Example 1] (Top) Original image and (Bottom) Preconditioned image.

Projection Result of the Original Image

Projection Result of the Preconditioned Image

Figure 5.5: [Example 1 Cont'd] (Top) Projection Result of the Original Image. (Bottom) Projection Result of the Preconditioned Image. Note that textures are better preserved in the projection of preconditioned image.

Original Image

Preconditioned Image

Figure 5.6: [Example 2] (Top) Original image and (Bottom) Preconditioned image.

Projection Result of the Original Image

Projection Result of the Preconditioned Image

Figure 5.7: [Example 2 Cont'd] (Top)Projection Result of the Original Image. (Bottom) Projection Result of the Preconditioned Image. Note that textures are better preserved in the projection of preconditioned image.

Figure 5.8: [Example 3] Inset of preconditioned image together with the original image. Note the difference in texture quality in the marked regions.

5.3 Thin-lens Projection Modeling

5.3.1 Point Spread Function from the Thin-lens Projection Model

By making use of a thin-lens projection model and adopting some of the radiometric image formation principles introduced in sections 2.2.2 and 2.3, the point spread function of each projected pixel is analytically derived. Here the point spread function of a projected pixel is the distribution of light from the pixel due to the projector lens. This is different from the radiometric imaging model reviewed in section 2.3 as it only describes the light distribution for the whole image, while the light distribution per input projector pixel is desired. Therefore, it is necessary to derive such a model ourselves.

In an oblique projection scenario, due to the lighting variations within the projector and the illumination fall off from the off-axis projection, intensity responses across the displayed image are not uniform. In reality, the method of using uniform PSF for a small certain bounding area discussed in the previous section can not represent fully the nature of the light distribution from the projector. Thus a physical model of the light distribution process needs to be proposed, which can well represent uniform PSFs in the orthogonal projection and non-uniform PSFs in oblique projection.

When an image is projected by a projector on a planar display surface, a thin-lens radiometric model can be derived with the following assumptions:

- The projector lens is assumed to be a thin-lens, as a simplification of the real complex projector lens system. This is much more realistic compared to the pinhole projection model used in other existing work.
- The projector input pixel radiance has already been estimated from a known

projector intensity-to-radiance transfer function, and the preconditioned image intensity will be able to be calculated using a known projector radiance-to-intensity transfer function. The projector transfer functions are computed using photometric and colorimetric calibration methods in chapter 4.

In order to estimate the light distribution on the projection surface with respect to a specific input pixel, the lens is segmented into multiple infinitesimal patches (see figure 5.9), which results in multiple small projected patches on projection surface. By integrating multiple infinitesimal lens patches together as a whole, it would be possible to obtain the overall light distribution as a PSF by spatially accumulating the different small projected patches. Figure 5.9 shows an illustration of how the light of a specific projector pixel pass through an infinitesimal patch of lens and distribute on the projection surface.

The origin of the projector lens coordinate system is the lens center O . Pointing rightwards along the optical axis is the z axis while pointing upwards along the lens is the y axis.

To assist in the representation of this model, a series of vectors are defined for the clearness and convenience of illustration, which is shown in figure 5.10. The surface normal $\mathbf{n}$ of pixel area δO is orthogonal to the lens, *i.e.*, δO is in parallel to the projector lens. Let vector $\mathbf{x}_0$ be defined as $(x_0, y_0, z_0)^T$ and likewise for other vector in this coordinate system. Among all the vectors shown in figure 5.10, the fixed are $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_P$, while $\mathbf{x}'_0, \mathbf{x}'_1, \mathbf{x}_2$ are all dependent on the varying vector $\mathbf{x}_L$.

Using basic radiometric image formation principles, the power emitted δP_e can be represented in the following form:

$$\delta P_e = L \delta \Omega_e \delta O \left(\frac{\mathbf{x}'_0 \cdot \mathbf{n}}{\|\mathbf{x}'_0\|} \right) \quad (5.15)$$

where L is the projector pixel radiance, δO is the pixel area, $\delta \Omega_e$ is the solid

Figure 5.9: Light Distribution through the Thin-lens of the Projector. This graph shows how the light of a projector input pixel travels through an infinitesimal patch of the thin-lens and distribute on the projector surface.

angle facing δO , and $\left(\frac{\mathbf{x}'_0 \cdot \mathbf{n}}{\|\mathbf{x}'_0\|} \right)$ is the cosine of the angle between vectors $\mathbf{x}'_0$ and $\mathbf{n}$. Similarly, the power received δP_r can be represented by:

$$\delta P_r = \delta E \delta A_d \quad (5.16)$$

where δE is the irradiance of light to be distributed on the projection surface, and δA_d is the distribution area.

Equating powers emitted and received, δE can be derived from (5.15) and (5.16):

$$\delta E = L \delta O \delta \Omega_e \frac{1}{\delta A_d} \left(\frac{\mathbf{x}'_0 \cdot \mathbf{n}}{\|\mathbf{x}'_0\|} \right) \quad (5.17)$$

Since solid angle is defined in the radiometric principles as the division of

Figure 5.10: Vector Representation of Projector Pixel Light Distribution in Thin-lens Projection. This graph represents the light traveling path and lens geometry in vector forms for the convenience of derivation in mathematical forms.

orthogonal area by distance², the solid angles $\delta\Omega_e$, $\delta\Omega_l$, $\delta\Omega_d$ can be derived:

$$\delta\Omega_e = \delta A_l \left(\frac{\mathbf{x}'_0 \cdot \mathbf{n}}{\|\mathbf{x}'_0\|} \right) \frac{1}{\|\mathbf{x}'_0\|^2} \quad (5.18)$$

$$\delta\Omega_l = \delta A_l \left(\frac{\mathbf{x}'_1 \cdot \mathbf{n}}{\|\mathbf{x}'_1\|} \right) \frac{1}{\|\mathbf{x}'_1\|^2} \quad (5.19)$$

$$\delta\Omega_d = \delta A_d \left(\frac{\mathbf{x}'_1 \cdot \mathbf{x}_P}{\|\mathbf{x}'_1\| \|\mathbf{x}_P\|} \right) \frac{1}{\|\mathbf{x}'_2 - \mathbf{x}'_1\|^2} \quad (5.20)$$

where $\mathbf{x}_P$ is a vector from the origin of the lens coordinate system and orthogonally pointing to the projection surface (shown in figure 5.10). Using (5.18), (5.17) will become

$$\delta E = L \delta O \frac{\delta A_l (\mathbf{x}'_0 \cdot \mathbf{n})^2}{\delta A_d \|\mathbf{x}'_0\|^4} \quad (5.21)$$

5.3 Thin-lens Projection Modeling

131

Meanwhile, since solid angles $\delta\Omega_l$ and $\delta\Omega_d$ are equal, (5.19) and (5.20) can give out the ratio of δA_l and δA_d :

$$\frac{\delta A_l}{\delta A_d} = \frac{\mathbf{x}'_1 \cdot \mathbf{x}_P}{\mathbf{x}'_1 \cdot \mathbf{n}} \frac{\|\mathbf{x}'_1\|^2}{\|\mathbf{x}_P\| \|\mathbf{x}_2 - \mathbf{x}_1\|^2} \quad (5.22)$$

By using variables z_0 and z_1 to replace vector product $\mathbf{x}'_0 \cdot \mathbf{n}$ and $\mathbf{x}'_1 \cdot \mathbf{n}$ respectively from the vector definition, (5.21) incorporating (5.22) becomes

$$\delta E = L\delta O\delta \frac{z_0^2}{z_1} (\mathbf{x}'_1 \cdot \mathbf{x}_P) \frac{\|\mathbf{x}'_1\|^2}{\|\mathbf{x}_P\| \|\mathbf{x}'_0\|^4 \|\mathbf{x}_2 - \mathbf{x}_1\|^2}. \quad (5.23)$$

In this equation, besides the independent fixed vectors $\mathbf{x}_0$ and $\mathbf{x}_P$, there are also independent variables to be fixed: L and δO . The focal length f of the projector lens is also a fixed independent variable. The varying independent vector for plotting is $\mathbf{x}_L = (x_L, y_L, 0)^T$.

All the other dependent variables can be replaced by these independent variables. Through lens geometry principle $\frac{1}{z_0} + \frac{1}{f} = \frac{1}{z_1}$, we can derive:

$$\mathbf{x}'_1 = \mathbf{x}_1 - \mathbf{x}_L = \frac{f}{z_0 + f} \mathbf{x}_0 - \mathbf{x}_L \quad (5.24)$$

Additionally, vector $\mathbf{x}'_0$ can be replaced by $\mathbf{x}'_0 = \mathbf{x}_L - \mathbf{x}_0$, and vector $\mathbf{x}_2$ can be expressed by $\mathbf{x}_2 = \mathbf{x}_L + k\mathbf{x}'_1$, where k is a scalar determined by the plane equation $\mathbf{x}_2 \cdot \mathbf{x}_P = \|\mathbf{x}_P\|^2$. Replacing $\mathbf{x}_2$ in this plane equation and (5.24), the scalar k is determined:

$$k = \frac{\|\mathbf{x}_P\|^2 - \mathbf{x}_L \cdot \mathbf{x}_P}{\mathbf{x}'_1 \cdot \mathbf{x}_P} \quad (5.25)$$

By taking k back, we can derive $\mathbf{x}_2$ through

$$\mathbf{x}_2 = \mathbf{x}_L + \frac{\|\mathbf{x}_P\|^2 - \mathbf{x}_L \cdot \mathbf{x}_P}{(\mathbf{x}_1 - \mathbf{x}_L) \cdot \mathbf{x}_P} (\mathbf{x}_1 - \mathbf{x}_L) \quad (5.26)$$

Also, we can derive

$$\begin{aligned} \|\mathbf{x}_2 - \mathbf{x}_1\| &= \|(k-1)\mathbf{x}'_1\| \\ &= \frac{\|\mathbf{x}_P\|^2 - \mathbf{x}_L \cdot \mathbf{x}_P - (\mathbf{x}_1 - \mathbf{x}_L) \cdot \mathbf{x}_P}{(\mathbf{x}_1 - \mathbf{x}_L) \cdot \mathbf{x}_P} \|\mathbf{x}_1 - \mathbf{x}_L\| \\ &= \frac{\|\mathbf{x}_P\|^2 - \mathbf{x}_1 \cdot \mathbf{x}_P}{(\mathbf{x}_1 - \mathbf{x}_L) \cdot \mathbf{x}_P} \|\mathbf{x}_1 - \mathbf{x}_L\| \end{aligned} \quad (5.27)$$

Since $\| \mathbf{x}'_1 \| = \| \mathbf{x}_1 - \mathbf{x}_L \|$, we can replace all the dependent variables in (5.23) with all independent variables and derive the following formula.

Formula 1 (Light Distribution Formula) *The irradiance of light δE distributed in area δA_d on the display surface is expressed by*

$$\delta E = L\delta O \underbrace{\frac{z_0(z_0 + f)}{f} \frac{1}{\| \mathbf{x}_P \| (\| \mathbf{x}_P \|^2 - \mathbf{x}_1 \cdot \mathbf{x}_P)^2}}_{\text{part 1: independent of } \mathbf{x}_L} \underbrace{\frac{((\mathbf{x}_1 - \mathbf{x}_L) \cdot \mathbf{x}_P)^3}{\| \mathbf{x}_L - \mathbf{x}_0 \|^4}}_{\text{part 2: dependent on } \mathbf{x}_L}. \quad (5.28)$$

where part 1 of this equation is independent of vector $\mathbf{x}_L$ and can be fixed once the projection setup is fixed, while part 2 is dependent on $\mathbf{x}_L$ and can be plotted with varying $\mathbf{x}_L$. The constraint on the varying vector is the lens aperture radius r , where $\| \mathbf{x}_L \| \leq r$.

For orthogonal projection scenario, since $\mathbf{x}_P$ will align with the lens optical axis, the product of $(\mathbf{x}_1 - \mathbf{x}_L) \cdot \mathbf{x}_P$ will be of a constant value. Thus

$$\delta E \propto \frac{1}{\| \mathbf{x}_L - \mathbf{x}_0 \|^4}. \quad (5.29)$$

Since $\mathbf{x}'_0 = \mathbf{x}_L - \mathbf{x}_0$, it can be seen that δE is inversely related to the 4th power of the length of $\mathbf{x}'_0$.

Based on this simplification, we can make an interesting yet very important observation. In orthogonal projection, the light distribution on the projection surface shows a *more similar* shape to that of a Gaussian distribution than the shape of a pillbox distribution [66], as pillbox shaped PSFs are widely used to estimate the out-of-focus blur effect for the camera. This observation also provides the theoretical support for and justification of why we have chosen Gaussian distribution functions instead of pillbox distribution functions for approximation of the projector out-of-focus blur in the observation-based approach.

For a point on the projection surface frame of reference, it can be denoted as $\mathbf{q} = (p, q, 0)^T$. To relate it back to the original projector lens frame of reference with $\mathbf{x}_L$, consider using

$$\mathbf{x}_L = \mathbf{R}\mathbf{q} + \mathbf{x}_P \quad (5.30)$$

where $\mathbf{R}$ is a 3×3 rotation matrix with the rotation axis $\mathbf{a} = \frac{\mathbf{n} \times \mathbf{x}_P}{\|\mathbf{x}_P\|}$, such that

$$\begin{bmatrix} \mathbf{x}_P & \mathbf{a} & \mathbf{a} \times \frac{\mathbf{x}_P}{\|\mathbf{x}_P\|} \end{bmatrix} = \mathbf{R} \begin{bmatrix} 0 \\ 0 & \mathbf{a} & \mathbf{a} \times \mathbf{n} \\ \mathbf{x}_P \end{bmatrix} \quad (5.31)$$

Because of $\mathbf{R}$'s sole dependency on $\mathbf{x}_P$, it can be solved when $\mathbf{x}_P$ is calculated.

Similar as scalar k , scalar α is invented such that $\mathbf{x}_L = \mathbf{x}_1 + \alpha(\mathbf{x}_1 - \mathbf{x}_2)$. From the constraint $\mathbf{x}_L \cdot \mathbf{n} = 0$, scalar α can be derived and $\mathbf{x}_L$ becomes

$$\mathbf{x}_L = \mathbf{x}_1 - \frac{\mathbf{x}_1 \cdot \mathbf{n}}{(\mathbf{x}_1 - \mathbf{x}_2) \cdot \mathbf{n}}(\mathbf{x}_1 - \mathbf{x}_2). \quad (5.32)$$

If we write it in variable form, it becomes

$$\begin{aligned} \begin{bmatrix} x_L \\ y_L \\ 0 \end{bmatrix} &= \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} - \frac{z_1}{z_1 - z_2} \begin{bmatrix} x_1 - x_2 \\ y_1 - y_2 \\ z_1 - z_2 \end{bmatrix} \\ &= \begin{bmatrix} (z_1 x_2 - z_2 x_1)/(z_1 - z_2) \\ (z_1 y_2 - z_2 y_1)/(z_1 - z_2) \\ 0 \end{bmatrix} \end{aligned} \quad (5.33)$$

The projective transformation now can be expressed by

$$\begin{bmatrix} x_L \\ y_L \\ 1 \end{bmatrix} \propto \begin{bmatrix} z_1 & 0 & -x_1 & 0 \\ 0 & z_1 & -y_1 & 0 \\ 0 & 0 & -1 & z_1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \\ 1 \end{bmatrix} \quad (5.34)$$

With (5.30) and (5.34), given vector $\mathbf{q}$, $\mathbf{x}_L$ can be computed from the following equation:

$$\begin{bmatrix} x_L \\ y_L \\ 1 \end{bmatrix} \propto \begin{bmatrix} z_1 & 0 & -x_1 & 0 \\ 0 & z_1 & -y_1 & 0 \\ 0 & 0 & -1 & z_1 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{x}_P \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p \\ q \\ 0 \\ 1 \end{bmatrix} \quad (5.35)$$

In order to generate the PSF map for each input projector pixel, with the above framework, we also need to know the independent fixed and varying parameters for calculation. Physical parameters of the projector, such as focal length f and lens aperture radius r need to be calibrated. The size of the lens aperture has a significant effect on the shape of estimated PSFs. From (5.28), we can see that the aperture radius r only has a constraint on the length of $\mathbf{x}_L$. In a large aperture setting, the estimated PSF from our model shows a shape similar to Gaussian distribution in an on-axis projection scenario in the graph (a) of figure 5.11, while the graph (a) of figure 5.12 shows an off-axis projection scenario where the shape of the estimated PSF is asymmetric. In a small aperture setting, both the on-axis and off-axis projections shown in the graph (b) of figure 5.11 and of figure 5.11 respectively result in similar PSFs but in a severely *truncated* manner. Other parameters such as $\mathbf{x}_L$ and $\mathbf{x}_P$ can be estimated from extrinsic parameters of the projector. In the next section, a detailed calibration approach will be presented.

5.3.2 Calibration of the Thin-lens Projection Model

Difference Between Pinhole Model and Thin-lens Model

In section 2.2, we reviewed the pinhole and thin-lens camera model respectively. The primary difference is that in pinhole model, as long as the pinhole is small

Figure 5.11: Sample PSF shapes in On-axis Cases under Large and Small Apertures. Graph (a) shows a sample estimated PSF shape under a large aperture setting in an on-axis projection scenario; (b) shows a sample estimated PSF shape under a small aperture setting in the same on-axis projection scenario.

(a)

(b)

Figure 5.12: Sample PSF shapes in Off-axis Cases under Large and Small Apertures. Graph (a) shows a sample estimated PSF shape under a large aperture setting in an off-axis projection scenario; (b) shows a sample estimated PSF shape under a small aperture setting in the same off-axis projection scenario.

enough such that images can be formed sharply on the screen, the focal length is determined by the distance between the pinhole and the screen. While for thin-lens system, light has to intersect on the other side of the lens and form a sharp image in the screen is in focus. However, the depth of focus for thin-lens camera is very limited, and can be affected by its focal length and aperture changes. The larger the lens radius, the shorter the depth of focus. If the lens aperture radius contracts small enough, it will become a pinhole camera. In digital photography, there is a parameter f-number defined as the division of focal length by aperture diameter, which measures the depth of focus of a particular camera.

However, there are similarities between pinhole model and thin-lens models, with which we can make use of to assist in our thin-lens calibration avoiding calibration from scratch. In figure 5.13, intuitively, when viewing thin lens projector as a pinhole projector, the pixel distance to the lens z_0 is the same as the focal length f_p in the corresponding pinhole model, while the screen distance z_1 are identical for both models. This similarity allows the estimation of thin-lens focal length from pinhole focal length. Another similarity is that both models conforms with projective transformation, which allows the extrinsic parameters of the pinhole model to be used for the thin-lens model.

In the following parts, calibration methods for lens focal length f and aperture radius r will be proposed, followed by the calculation of the perpendicular projection surface vector $\mathbf{x}_P$ directly from the pinhole model calibration results.

Lens Focal Length Calibration

Based on the similarity of pinhole lens focal length with thin-lens pixel-to-lens distance, it allows us to calibrate the focal length f of the thin-lens model using

the following method.

Figure 5.13: Projector Geometric Calibration Models. (a) Pinhole projection model; (b) Thin-lens projection model. The same similar triangles in solid lines exist in both models. Using lens geometry together with similarities between the two model, thin-lens focal length can be recovered from pinhole focal length. See text for details.

After the focal length (intrinsic parameter) of the pinhole model has been calibrated using techniques proposed in section 3.2.2, the thin-lens projector pixel-to-lens distance z_0 is known due. Figure 5.13 shows a pinhole projection model on the left and a thin-lens projection model on the right. To recover the thin-lens projector focal length, first let us consider similar triangles in the pinhole model. It can expressed as

$$\frac{x}{f_p} = \frac{x_1}{z_1} \quad (5.36)$$

where f_p is the pinhole projector focal length. In the thin-lens model, basic lens geometry can be expressed as

$$\frac{x}{f} = \frac{x_1}{z_1 - f} \quad (5.37)$$

where f is the lens focal length. From (5.36) and (5.37), we can get the lens focal length:

$$f = \frac{z_1 f_p}{z_1 + f_p} \quad (5.38)$$

where z_1 is the depth of the point in focus. Replacing f_p with z_0 , the above equation will be solved. A detailed procedure is as follows:

1. Point projector orthogonally to the projection surface, adjust the projection distance such that sharp images are displayed;
2. Calibrate the projector intrinsic and extrinsic parameters using methods from section 3.2.2;
3. Calculate the projection distance z_1 using calibrated extrinsic parameters;
4. Use calibrated pinhole model focal length and z_1 to calculate thin-lens focal length f from (5.38).

After the focal length of the thin-lens projector has been calculated, the settings of the projector can not be adjusted, thus manual movement of the projector or the projection screen can help generate an out-of-focus blurred image on the projection surface.

During the calibration of the lens focal length, we can manually keep the display in focus so that the estimation of f can more accurate.

Lens Aperture Radius Calibration

Since the depth of focus is affected by the lens aperture, if some area on the projection screen is out-of-focus, the light will be distributed in a larger area than if the point was in focus. This allows us to observe the out-of-focus area using a camera

and subsequently compare with the computed sum of irradiance within that area. As shown in our analytic model, figure 5.11 and figure 5.12, the aperture radius r only determines how the shape of computed PSF is truncated, *i.e.* the shape of a computed PSF with infinite aperture radius will be truncated when its aperture becomes smaller. As a camera can not observe the projection of a pixel on the screen clearly enough, we can project a small patch of pixels assuming their combined PSF is an approximation of each individual PSF. Therefore an area of light distribution can be observed by the camera on the screen (shown in figure 5.14). A smaller patch δC within the projected bright area δP can be observed by the camera and estimated from our model up to a scale factor, which is equal to $L\delta O$. This is based on the assumption of an infinite aperture radius, because its truncation effect has not been spread into this small area. Subsequently, by observing a larger area δA shown in figure 5.14, the aperture radius is unknown and can be calculated based on the truncation shape of the PSF. A detailed description to compute the aperture radius r with an optimization procedure is shown below.

If the projector is projecting a black image with a white pixel at its center, an enlarged version of light distribution after projection may look like a bright area δP from the camera shown in figure 5.14. The light distribution decays from the center towards the edge. Assuming area δC is an area smaller than and within the area δP , the sum of light falling in δC can be observed using a camera and its intensity-to-radiance transfer function as $\sum_{\mathbf{x}_L \in \delta C} \delta \tilde{E}_1$. From (5.28), assuming an infinite aperture radius, *i.e.*, $\|\mathbf{x}_L\| \ll r$, the sum of light $\sum_{\mathbf{x}_L \in \delta C} \delta E_1$ can not be computed directly since L and δO can not be estimated. However, we can define a term $E_1 = \frac{\sum_{\mathbf{x}_L \in \delta C} \delta E_1}{L\delta O}$ falling within δC , which can be computed based on the previous calibration results. Since observation area δC is smaller than and

Figure 5.14: Lens Aperture Radius Estimation. A square patch of white pixels with equal brightness is projected on the projection screen. The area δP is the light distribution area, within which a smaller area δC and a much larger area δA is used as the aperture radius estimation.

Of aperture radius estimation. The area δP is the light distribution area, within which a smaller area δC and a much larger area δA is used as the aperture radius estimation.

known, E_2 is not able to be computed. In order to be able to get r , a 1-D golden section search or gradient descent optimization using least squares method can be applied over the following objective function for minimization:

$$\arg \min_r \left\| \rho E_2 - \sum_{\mathbf{x}_L \in \delta A} \delta \tilde{E}_2 \right\|^2 \quad (5.40)$$

where the radius r can be set within a range of 1 to 100 for optimization. Meanwhile, after the estimation of ρ , $L\delta O$ can also be estimated from (5.39).

Since the projection of a single white pixel onto the screen can not be properly observed due to the camera noise, a patch of 10 white pixels are used instead, and their combined PSF is computed through a summation of the per-pixel PSFs over this 10-pixel patch. For example, if the patch is projected and distributed in the area δP , the observed sum of light in δC can be expressed as $\sum_{i=1}^{10} \sum_{\mathbf{x}_L \in \delta C} \delta \tilde{E}_{1i}$, where i is the projector pixel index. This is because the 3D position of a projector pixel in the lens coordinate system which is denoted by $\mathbf{x}_0$ is a variable of the function δE in Formula 1. In this way, the lens aperture radius can be properly estimated.

Perpendicular Projection Surface Vector Calculation

The perpendicular display surface vector $\mathbf{X}_p$ shown in figure 5.10 must be calculated to complete our thin-lens modeling. While this is straightforward, from the result of geometric calibration of the pinhole model using approaches in chapter 3, $\mathbf{X}_p$ can be directly computed and hence will not be discussed in detail here.

5.3.3 PSF Map Calculation

From the light distribution of a particular projector pixel, we can generate a PSF such that the preconditioning technique proposed in section 5.2.2 can be used. The

light distribution for each projector pixel on the projection screen is presented in (5.28). Thus for each projector input pixel, a corresponding PSF will be generated. However, for an high resolution input image size of 1024×768 pixels, the time for computing all the PSFs may be up to 200+ hours. From (5.28), we can also see that the neighboring projector input pixels will generate PSFs of very similar shape but with different locations of the light distribution area on the projection surface. Thus to save computational cost, we may approximate the PSFs of neighboring input pixels by that of the center pixel in the neighborhood. By defining a neighborhood window of 128×128 points on a projection surface with 1024×768 points, a PSF map is generated with a representing PSF at the center of each neighborhood window.

If $L\delta O$ is defined as a unit irradiance, we can then define a basis PSF map as $\frac{\delta E}{L\delta O}$. The corresponding $L\delta O$ from different pixel irradiance of the input image will be multiplied into the PSF map once the basis PSF map is ready. Thus each time an input image is changed, the corresponding PSF map for this particular image can be generated from a pre-computed basis PSF map.

Our projector is a 3M M8749 portable LCD projector model with (1024×768) resolution. Some sample results computed from previous calibration steps are: projector principal point: (541.4249, -66.5086), pinhole projector focal length: 1774.9039, thin-lens projector focal length: 1362.3567, and perpendicular projection surface vector: $\mathbf{x}_P = (43030.0880, -205.3462, 40556.8573)^T$. The thin-lens projector focal length converted in millimeter is 31.7584mm which is within the range of manufacturer-reported focal length (27.3-32.8mm). Our estimated aperture radius is 9.8mm, and the f-number is thus 1.62 (focal length/aperture diameter). This value, comparing with the manufacturer-reported data (1.7-2.1), is

a bit higher.

like at different regions of an obliquely projected image. Note that the upper-left region has more more out-of-focus blur than the lower-right region, and the PSF in the upper-left region shows a broader irregular and truncated shape comparing with the more compact and Gaussian-like shape of the PSF in the lower-right region.

5.4 Generation of Preconditioned Images

5.4.1 Preconditioned Image Generation using Windowed Wiener Filtering

The analytic model can also be used for image preconditioning. When a PSF map is computed from the composite basis PSF map by multiplying with corresponding $L\delta O$ terms of each pixel, the projector input image can be now preconditioned by using windowed wiener filtering. Since the PSF computed is specific to certain input image pixels in a confined area, *i.e.* a window of 128×128 in our experiment for example, when wiener filtering is applied, it must be over that window area.

5.4.2 Experiments and Results

We are using the same system setup as in section 5.2.3. In our experiment, the projector parameters including the lens focal length and aperture radius, is calibrated using the Matlab Camera Calibration Toolbox [6] and the techniques proposed in section 5.3.2. Sharp images when projected, will have out-of-focus blur effects. These images will be subsequently preconditioned using Wiener filtering. After reprojection, the blur effects are enormously lessened and the texture of the original sharp images is better preserved in the projected preconditioned images.

Figure 5.16 shows the original and preconditioned images of a “cat” while figure 5.18 shows the results of the projection of these two “cat” images. Notice that the fur of the cat in the original image is already very sharp. In the projected images, the projection of preconditioned image much better preserved the texture of the image than the other. In figures 5.17 and 5.19 an outdoor scene of a temple

is shown to demonstrate the performance of projected preconditioned image in contrast with the projection of original image. In all the test images, we choose the original images to be sufficiently in focus so as to demonstrate the contribution of our algorithm is not merely attributed to sharpening the input image.

In figure 5.20 we compares our results as insets into the original projected images. Note the better preservation of the textures in the projection of preconditioned images.

Although our results are subjective, it is hard to give quantitative measurement of the degree of improvement of images on the projection setup, given the nature of projector camera systems. However, we have also tried preconditioning the original images by Wiener filtering with uniform blur kernels. The uniform blur kernel can be standard Gaussian blurring kernel or windowed kernels from our approximated PSF map. Four projected results of uniformed Wiener filtered images using different windowed kernels from our PSF map are shown in figure 5.21 and 5.22. From those figures we can see, the blurred parts of the projected images may be improved in some way while the in focus parts are over sharpened with apparent ringing effects and color changes. Besides Wiener Filtering, Lucy-Richardson deconvolution has also been applied in the preconditioning the original images but achieving inferior results than Wiener filtering and thus will not be shown here.

Original Image

Preconditioned Image

Figure 5.16: [Example 1] (Top) Original image and (Bottom) Preconditioned image.

Projection Result of the Original Image

Projection Result of the Preconditioned Image

Figure 5.17: [Example 1 cont'd] (Top) Projection Result of the Original Image. (Bottom) Projection Result of the Preconditioned Image. Note that textures are better preserved in the projection of preconditioned image.

Original Image

Preconditioned Image

Figure 5.18: [Example 2] (Top) Original image and (Bottom) Preconditioned image.

Projection Result of the Original Image

Projection Result of the Preconditioned Image

Figure 5.19: [Example 2 Cont'd] (Top)Projection Result of the Original Image. (Bottom) Projection Result of the Preconditioned Image. Note that textures are better preserved in the projection of preconditioned image.

Figure 5.20: [Example 3] Inset of preconditioned image together with the original image.
Note the difference in texture quality in the marked regions.

Figure 5.21: [Example 4] Projection Results of Preconditioned Image Using Uniform Blur PSFs. Top: An image is uniformly preconditioned using a PSF. The enlarged bottom-left inset is still blurred, while the enlarged upper-right inset shows some over-sharpening effects. Bottom: A broader PSF distribution is used to uniformly precondition the original image. Both of the enlarged insets show apparent over-sharpening effects.

Figure 5.22: [Example 4 Cont'd] Projection Results of Preconditioned Image Using Uniform Blur PSFs. The same original image is uniformly preconditioned using increasingly broader PSFs than the PSFs used in figure 5.21, and the observations are shown in the image above. Note the apparent over-sharpening effects shown in the insets.

5.5 Comparison of Approaches

In this chapter, we proposed two approaches lessening the out-of-focus effects of oblique projection from the projectors. Both of the methods use preconditioning techniques to lessen the blur, but differs in the estimation of PSF and the computation of preconditioned images.

The image-based empirical approach [9] using the estimation of Gaussian PSF map is intuitive and most direct. It is simple to use, but the estimation error may be big due to limitations of modeling cursorily. For example, while the LCD projectors used produce out-of-focus blur similar to Gaussian blur, out-of-focus blur of DLP projectors used do not appear similar to Gaussian blur, and can not be properly estimated using Gaussian PSF maps. Besides, the empirical approach does not have a strong mathematical foundation to support the theory of its PSF map estimation.

The physical model we proposed secondly, is a complex analytic model. It is based on a thin-lens model and is developed based on the radiometric image formation principles. While in practice projectors are comprised of complex lens, our model may be a reasonable approximation of the projector lens system and provides a sufficient mathematically supported projection model for the estimation of PSF maps. The calibration procedure may be a little bit more complex comparing with the estimation of PSF maps using empirical approach. However, the great modeling accuracy may lead to better results.

Neither of the two approaches can be measured quantitatively at the current stage. In the future, we will have a further investigation to compare the results quantitatively.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

The aim of the research done in this thesis is to provide a better display quality when projector camera systems are obliquely deployed. Geometric calibration is used to undistort the projector image and provides further basis for photometric calibration and out-of-focus blur reduction. Photometric calibration estimates the projector intensity-radiance and camera radiance-intensity transfer functions which is key to creating artifact-free uniform color/brightness display. The out-of-focus blur can be reduced through our image preconditioning approaches, and thus further improve the quality of image/video display of the projector camera system.

This thesis started with introduction of projector-camera systems and its applications in various fields. Research issues in geometry, photometry and out-of-focus projector blur were discussed next in chapter 1. Specific to each issue,

detailed problems were outlined and approaches proposed to solve them are proposed.

In chapter 2, some basic concepts were presented for the better understanding of the approaches proposed in the following chapters. The concepts are all described in mathematical forms. First, the characteristics of projectors and cameras and typical imaging models are introduced: the pin-hole model and the thin-lens model. Then standard imaging restoration techniques were reviewed. Convolution commutation was next introduced as a nature of convolution to lay a foundation for the theory of image preconditioning later in chapter 5. Finally, the radiometric image formation model was introduced for the better understanding of the formation of estimating spatially variant PSFs using physical thin-lens model.

In chapter 3, two types of geometric calibration were introduced. The first is homography calibration introducing the computation of the 2D projective transformation matrix between the projectors, cameras and screen through correspondences. Secondly, full camera geometric calibration techniques (Heikkilä and Silvén's Method and Zhang's method) were introduced. These methods can be used to calibrate the camera internal, external parameters and lens distortion factors. The latest projector geometric calibration techniques from Okatani and Deugchi, Raij and Pollefey are reviewed respectively before a projector calibration approach was proposed based on the full camera calibration methods introduced previously.

In chapter 4, an overview of the photometric and colorimetric calibration of projectors and cameras was provided first. Then a grayscale photometric calibration framework was proposed. The solution of the framework is based on the theory of isointensity curves. Three methods were given: the staircase, the recur-

sive bisection, and the gridline optimization method. In gridline optimization is robust in any situation. Results are shown demonstrating the recovered intensity transfer functions are able to accurately model the raw data, and can be used in two ways – predicting the camera observed image based on known projector inputs, and setting correct input intensities to achieve a desired camera-observed image. Subsequently a colorimetric calibration framework was presented afterwards. Single-projector and dual-projector calibration were proposed respectively. The dual-projector approach proved to be the most robust. The following experiment results fully shows that the recovered intensity transfer functions can accurately model the raw data, and further predict the camera observed images as well as compensate a projected background for a desired camera-observed image, which successfully solves the problem of artifacts from color imbalance in overlapped displays. None of the photometric and colorimetric calibration approaches require explicit parametric models nor smoothness constraints, nor do they require expensive equipment: an inexpensive camera with minimal controls and a small dynamic range may be used.

At the beginning of chapter 5, the concept of image preconditioning and its application in projected display to counter the effects of out-of-focus display were introduced. Then the limitations of image preconditioning in this sense were discussed. A camera based approach was proposed to estimate the spatially variant PSFs across the projector image. Composition of basis images and generation of the preconditioned image using wiener filtering were presented afterwards. The results shown demonstrate the successful lessening of the effects of projector blur. Another approach based on physical thin-lens imaging model was proposed next. The spatially variant PSFs can be computed from the imaging model and inde-

pendent of any camera observed images, but the model needs to be calibrated beforehand. The calibration of the projector thin-lens model and computation of the preconditioned image windowed wiener filtering were presented. Results of this approach proved the correctness of the physical model and successfully reduce the out-of-focus blur in the oblique projection scenario. The physical thin lens model based approach requires a bit more complex lens calibration than the image based approach, however, the strong mathematical foundation deducted from the nature of thin lens imaging models provide a sound explanation about the formation and reduction of the out-of-focus projector blur.

6.2 Future Work

The research until now has laid a very good foundation for the improvement of premium quality images/video on multi-projector camera systems. However, there are certainly some research methods and technologies to be developed to achieve a better display. Some potential future research work include:

- *Rapid automatic photometric calibration*

The photometric calibration method is slow mainly due to the data collection stage. Images are projected and observed by the camera, and the data collected are only with respect to a fixed point on the screen. Although it only takes several hours to get a decent dataset for the calibration, there are ways to expedite this process. One way is to research on the surface reflectance and propose a model for calibration first. In the current photometric calibration approach discussed in chapter 4, only a reflectance map is computed after the projector and camera intensity transfer functions are

derived. If a surface reflectance model can be proposed and estimated beforehand, the collection of data for calibrating intensity transfer functions would cost much less time, because instead of data collection on a single fixed screen point, multiple screen points can be used for data collection with the known surface reflectance model. This would geometrically reduce the data collection time, and even, if the display image intensities vary a lot spatially, the data collection can be completed in seconds which makes the photometric calibration become automatic and rapid.

- *Blur-free display through automatic segmented display of multi-projectors*

The approaches proposed in chapter 5 can reduce the out-of-focus to a large extent, however it can not be applied to any image for projection due to the limitations of image preconditioning technique we used. Since nowadays projectors are more prevalent in offices, homes and classrooms, *etc.* multiple overlapping projectors can be used to create large scale displays. Using both the image based approach and physical model based approach, further research can be carried out to decide the blur regions. With good geometric and photometric calibration, we can fill the out-of-focus area of one projector using another projector's in-focus area assuming these areas are overlapped. The regions to be filled can be automatically determined if we know the corresponding out-of-focus and in-focus regions.

- *Automatic ubiquitous display*

The automatic ubiquitous display is a research framework where everything in an indoor environment can be treated as a display surface which can be projected from one or more of the projectors as a display or an augmented device. The automatic system requires the projectors and cameras geometri-

cally and photometrically calibrated without or with acceptable blur. Cameras need to cover the environment to carry out real-time human, gestural detection, surface detection and advanced algorithms can be developed to choose display surface for human needs. The system can automatically self-learn through structured light on calibrated projectors and cameras.

- *Immersive interactive display*

Using the available geometric and photometric calibration techniques introduced in chapter 3 and chapter 4, immersive display environment can be created. However, to enable more interactive projection, accurate geometric registration and 3D modeling of the indoor space may be required. The interaction may be gestural or laser pointer, *etc.* and thus human gestural or laser pointer detection needs to be incorporated. Such an immersive interactive environment requires a fully integrated framework to be developed and the key research issues may be the accurate 3D modeling of the space through structured light, pose detection and estimation, *etc.*

Bibliography

- [1] Publication CIE No 17.4. *International Lighting Vocabulary*. Central Bureau of the Commission Internationale de L'Éclairage, Vienna, Austria.
- [2] Y.I. Abdel-Aziz and H.M. Karara. Direct linear transformation into object space coordinates in close-range photogrammetry. In *Proc. Symposium on Close-Range Photogrammetry.*, pages 1–18, Urbana, Illinois., 1971.
- [3] N. Asada, A. Amoano, and M. Baba. Photometric calibration of zoom lens systems. In *Proc. International Conference on Pattern Recognition, Vienna (Austria)*, volume A, pages 186–189, Vienna, Austria, 1996.
- [4] D. Bandyopadhyay, R. Raskar, and H. Fuchs. Dynamic Shader Lamps: Painting on real objects. In *Proc. of International Symposium on Augmented Reality*, pages 207–216, New York, NY, 2001.
- [5] D. Bennett. Alternate Realities Corporation, Durham, NC27703. In <http://www.virtual-reality.com/>, 1999.
- [6] J. Y. Bouguet. Camera calibration toolbox for matlab. In http://www.vision.caltech.edu/bouguetj/calib_doc/.
- [7] D. H. Brainard, D. G. Pelli, and T. Robertson. Display characterization. In

- The Encyclopedia of Imaging Science and Technology*, pages 172–188, John Wiley and Sons, New York, 2001.
- [8] M.S. Brown, A. Majumder, and R.G. Yang. Camera-based calibration techniques for seamless multi-projector displays. *IEEE Transactions Visualization and Computer Graphics*, 11(2):193–206, March 2005.
- [9] M.S. Brown, P. Song, and T.J. Cham. Image pre-conditioning for out-of-focus projector blur. In *Proc. IEEE Conference Computer Vision and Pattern Recognition, New York*, pages 1956–1963, 2006.
- [10] W. Buxton, G. Fitzmaurice, R. Balakrishnan, and G. Kurtenbach. Large displays in automotive design. *IEEE Computer Graphics and Applications*, 20(4):68–75, 2000.
- [11] T.J. Cham, J. Rehg, R. Sukthankar, and G. Sukthankar. Shadow elimination and occluder light suppression for multi-projector displays. In *Proc. IEEE Conference Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 513–520, Madison, Wisconsin, 2003.
- [12] H. Chen, R. Sukthankar, G. Wallace, and K. Li. Scalable alignment of large-format multi-projector displays using camera homography trees. In *Proc. IEEE Visualization Conference*, pages 339–346, Boston, MA, 2002.
- [13] D. Crasto, A. Kale, and C. Jaynes. The smart bookshelf: A study of camera projector scene augmentation of an everyday environment. In *Workshop on Applications of Computer Vision*, pages 218–225, January 2005.
- [14] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti. Surround-screen projection-based virtual reality: The design and implementation of the CAVE. In

- SIGGRAPH Conference Proceedings*, volume 27, pages 135–142, Anaheim, CA, 1993.
- [15] M. Czernuszenko, D. Pape, D. Sandin, T. DeFanti, L. Dawe, and M. Brown. The ImmersaDesk and InfinityWall projection-based virtual reality displays. *Computer Graphics*, pages 46–49, 1997.
- [16] P.E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH Conference Proceedings*, pages 369–378, Los Angeles, CA, 1997. [17 -2(l)]P.(c)2MRs zlitmn08-2(d)-29
- [17] Spitz ElectricHorizonTM. <http://www.spitzinc.com/>.
- [18] P. Favaro and S. Soatto. A geometric approach to shape from defocus. *IEEE Transactions Pattern Analysis and Machine Intelligence*,aaa(R)-1(A)78-6 (-)-2(4)-27.24 - Td | *IEonc.c:/s*,

- [23] M.D. Grossberg and S.K. Nayar. What is the space of camera response functions? In *Proc. IEEE Conference Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 602–609, Madison, Wisconsin, 2003.
- [24] M.D. Grossberg, H. Peri, and S.K. Nayar. Making one object look like another: controlling appearance using a projector-camera system. In *Proc. IEEE Conference Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 452 – 459, Washington, DC, 2004.
- [25] J. Heikkilä. Geometric camera calibration using circular control points. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 22(10):1066–1077, 2000.
- [26] J. Heikkilä and Silvén. A four-step camera calibration procedure with implicit image correction. In *Proc. IEEE Conference Computer Vision and Pattern Recognition, Puerto Rico*, pages 1106–1112, Washington, DC, USA, 1997.
- [27] B. Horn. *Robot Vision*. The MIT Press, 1986.
- [28] C. Jaynes, B. Seales, K. Calvert, Z. Fei, and J. Griffioen. The Metaverse - a collection of inexpensive, self-configuring, immersive environments. In *Proc. the 7th International Workshop on Immersive Projection Technology/ the 9th Eurographics Workshop on Virtual Environments*, pages 115–124, Zürich, Switzerland, 2003.
- [29] C. Jaynes, S. Webb, and R.M. Steele. Camera-based detection and removal of shadows from interactive multiprojector displays. *IEEE Transactions on Visualization and Computer Graphics*, 10(3):290–301, 2004.

- [30] W. Kunzman and G. Pettitt. White enhancement for color sequential DLP^{TM} . In *SID (Society for Information Display) International Symposium Technical Digest*, pages 121–124, 1998.
- [31] K. Levenberg. A method for the solution of certain problems in least squares. *Quarterly of Applied Mathematics*, 2(2):164–168, 1944.
- [32] S. Lin, J. Gu, S. Yamazaki, and H.Y. Shum. Radiometric calibration from a single image. In *Proc. IEEE Conference Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 938–945, Washington, DC, 2004.
- [33] L.B. Lucy. An iterative technique for the rectification of observed distributions. *Journal of Astronomics*, 79:745–754, 1974.
- [34] A. Majumder, Z. He, H. Towels, and G. Welch. Achieving color uniformity across multi-projector displays. In *Proc. of IEEE Visualization Conference*, pages 117–124, Salt Lake City, UT, 2000.
- [35] A. Majumder and R. Stevens. LAM: Luminance attenuation map for photometric uniformity in projection based displays. In *Proc. ACM symposium on Virtual Reality Software and Technology (VRST)*, pages 147–154, Hong Kong, China, 2002.
- [36] A. Majumder and R. Stevens. Color nonuniformity in projection-based displays: Analysis and solutions. *IEEE Transactions Visualization and Computer Graphics*, 10(2):177–188, 2004.
- [37] D. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal of Applied Mathematics*, 11:431–441, 1963.

- [38] J.W. McKee and J. Burgett, Sherrie. Implicit multiplane 3d camera calibration matrices for stereo image processing. *Proc. SPIE*¹, 3173:330–334, 1997.
- [39] T. Mitsunaga and S. Nayar. Radiometric self calibration. In *Proc. IEEE Conference Computer Vision and Pattern Recognition (CVPR)*, pages 374–380, Fort Collins, CO, 1999.
- [40] S.K. Nayar, V. Branzoi, and T. Boult. Programmable image using a digital micromirror array. In *Proc. IEEE Conference Computer Vision and Pattern Recognition (CVPR)*, pages 436–443, Washington, D.C., 2004.
- [41] T. Okatani and K. Deguchi. Autocalibration of projector-screen-camera system: Theory and algorithm for screen-to-camera homography estimation. In *Proc. International Conference on Computer Vision*, pages 774–781, Nice, France, 2003.
- [42] PanoWallTM. <http://www.panoramtech.com/>.
- [43] G. Pingali, C. Pinhanez, A. Levas, R. Kjeldsen, M. Podlaseck, H. Chen, and N. Sukaviriya. Steerable interfaces for pervasive computing spaces. In *Proc. IEEE International Conference on Pervasive Computing and Communications*, pages 315–322, Fort Worth, TX, 2003.
- [44] A. Raij, G. Gill, A. Majumder, H. Towles, and H. Fuchs. PixelFlex2: A comprehensive, automatic, casually-aligned multi-projector display. In *IEEE International Workshop on Projector-Camera Systems. (PROCAMS)*, Nice, France, 2003.

¹Ultrahigh- and High-Speed Photography and Image-based Motion Measurement, Andrew Davidhazy, Takeharu G. Etoh, C. Bruce Johnson, Donald R. Snyder, James S. Walton, Editors.

- [45] A. Raij and M. Pollefeys. Auto-calibration of multi-projector display walls. In *Proc. International Conference on Pattern Recognition*, volume I, pages 14–17, Cambridge, England, 2004.
- [46] R. Raskar, J. Baar, P. Beardsley, T. Willwacher, S. Rao, and C. Forlines. iLamps: Geometrically aware and self-configuring projectors. *ACM Transactions on Graphics*, 22(3):809–818, 2003.
- [47] R. Raskar and P. Beardsley. A self correcting projector. In *Proc. IEEE Conference Computer Vision and Pattern Recognition (CVPR)*, pages 504–508, Kauai, Hawaii, 2001.
- [48] R. Raskar, G. Welch, and W. Chen. Table-top spatially-augmented reality: Bringing physical models to life with projected imagery. In *Proc. the 2nd International Workshop on Augmented Reality*, pages 64–71, San Francisco, CA, 1999.
- [49] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs. The office of the future : A unified approach to image-based modeling and spatially immersive displays. In *ACM SIGGRAPH Conference Proceedings*, pages 179–188, Orlando, FL, 1998.
- [50] R. Raskar, G. Welch, K-L. Low, and D. Bandyopadhyay. Shader Lamps: Animating real objects with image-based illumination. In *Proc. the 12th Eurographics Workshop on Rendering*, pages 89–102, London, England, 2001.
- [51] W.H. Richardson. Bayesian-based iterative method of image restoration. *American Journal of Optical Society*, 62:55–59, 1972.

-
- [52] C.C. Slama. *Manual of Photogrammetry*. American Society of Photogrammetry, Falls Church, Virginia, 4th edition, 1980.
- [53] P. Song and T.J. Cham. A theory for photometric self-calibration of multiple overlapping projectors and cameras. In *IEEE International Workshop on Projector-Camera Systems. (PROCAMS)*, pages 97–104, San Diego, CA, 2005.
- [54] M.C. Stone. Color balancing experimental projection displays. In *Proc. the 9th Color Imaging Conference: Color Science, Systems, and Applications*, pages 342–347, Scottsdale, AZ, 2001.
- [55] R. Sukthankar, R. Stockton, and M. Mullin. Automatic keystone correction for camera-assisted presentation interfaces. In *Proc. International Conference on Multimodal Interfaces*, pages 607–614, Beijing, China, 2000.
- [56] R. Sukthankar, R. Stockton, and M. Mullin. Smarter Presentations: Exploiting homography in camera-projector systems. In *Proc. International Conference on Computer Vision*, pages 247–253, Vancouver, Canada, 2001.
- [57] R. Surati. *A Scalable Self-Calibrating Technology for Seamless Large-Scale Displays*. PhD thesis, Department of Electrical and Computer Engineering, Massachusetts Institute of Technology, 1999.
- [58] J.M. Tenenbaum. *Accommodation in Computer Vision*. PhD thesis, Department of Computer Science, Stanford University, 1970.
- [59] NTU Reality Theatre. <http://www.camtech.ntu.edu.sg/rt/index.html>.
- [60] E. Trucco and A. Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1998.

- [61] R.Y. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, RA-3(4):323–344, 1987.
- [62] F. Tsumuraya, N. Miura, and N. Baba. Iterative blind deconvolution method using Lucy’s algorithm. *Astronomy and Astrophysics*, 282:699–708, February 1994.
- [63] R. Ulichney and R. Sukthankar. Correcting luminance for obliquely-projected displays. In *Int. Conf. on Consumer Electronics (ICCE)*, pages 367–368, Singapore, 2005.
- [64] VR Group, Center for Advanced Media and Technology, Nanyang Technological University. OpenSG Forum Events Workshops. In http://www.inigraphics.net/press/topics/2003/issue5/5_03Events.pdf.
- [65] W. Wallace, L.H. Schaefer, and J.R. Swedlow. A workingperson’s guide to deconvolution in light microscopy. *BioTechniques*, 31:1076–1097, 2001.
- [66] M. Watanabe and S. K. Nayar. Rational filters for passive depth from defocus. *International Journal of Computer Vision*, 27(3):203–225, 1998.
- [67] G. Q. Wei and S. D. Ma. Implicit and explicit camera calibration: Theory and experiments. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 16(5):469–480, 1994.
- [68] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 14(10):965–980, 1992.

-
- [69] N. Wiener. *Extroplation, Interpolation, and Smoothing of Stationary Time Series*. Wiley, New York, USA, 1949.
- [70] R. Yang, D. Gotz, J. Hensley, H. Towles, and M. S. Brown. Pixelflex: A reconfigurable multi-projector display system. In *Proc. of IEEE Visualization Conference*, pages 167–174, San Diego, CA, 2001.
- [71] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In *Proc. International Conference on Computer Vision*, pages 666–673, Corfu, Greece, 1999.
- [72] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.