博 士 論 文

Photometric Stereo with
Auto Radiometric Calibration

(カメラ応答関数の自動校正を伴う照度差ステレオ)

モンコンマーン　ウィーンナート

Abstract

Scene structure recovery from images is one of the fundamental problems that has been studied since the early day of computer vision. It is well-known that scene appearance in images does not only rely on scene geometric structure but also surface properties, illumination, and the properties of camera that is used to capture the images. Then, the scene geometric structure can be estimated from the appearance of the images under some assumption regarding surface reflection properties and surrounding environment such as illumination, hence the name photometric-based scene recovery. However, many photometric-based scene recovery studies emphasize only in the relationship of scene appearance, surface properties, and illumination, while take for the granted that the camera properties are calibrated or measured beforehand.

In this work, we focus on improving the photometric-based surface recovery method to account for the camera properties such as radiometric response function so we can avoid time-consuming and cumbersome camera calibration with no additional images are required. The key idea behind this work is to make use of inherent properties that lie inside the input images such as the consistency between the irradiance converted from the camera's image formation model and the irradiance estimated from a reflection model. In other words, we use the physical clues that reside in the images to perform camera calibration while simultaneously estimate the scene structure.

First, we present photometric stereo techniques that can estimate surface orientations from a sequence of object images taken under different lighting directions with a radiometrically uncalibrated camera. The original photometric stereo assumes the images are captured by a camera with a linear response function. However, cameras often have a non-linear response function, and thus, the radiometric calibration is required to cancel the effect of nonlinear response function before taking images which are later used for physics-based analysis of the scene. Unfortunately, the radiometric calibration is a time-consuming pre-process that requires either many additional images or a calibration target.

Here we use the consistency between the irradiance converted from the inverse response function and the irradiance estimated from Lambertian reflection model to formulate a linear optimization problem to estimate the surface normals of a Lambertian surface and the response function simultaneously. We empirically show that our proposed method can produce surface orientation from images accurately even though the images were captured by radiometrically uncalibrated cameras.

Then, we extend the proposed method for the surfaces that do not follow Lambertian reflection model. Many real-world objects contain reflections that do not follow Lambertian reflection property such as specular highlights or weak specular lobe. Our simultaneous estimation model does not account for these kinds of reflection so it produces distorted surfaces when specular highlights are observed in the input images.

However, many non-Lambertian surfaces exhibit reflection similar to Lambertian surfaces where no specular highlight is not observed. Therefore, we can treat specular highlights pixels as outliers to Lambertian reflection model and use a robust estimation technique such as RANSAC to determine in-

verse response function and surface normals that maximize the number of diffuse reflection. The experiments on synthetic images and real images illustrate that our proposed method can compensate the nonlinearity of response function even though there exist corrupted measurements such as specular highlights in the scene.

Finally, we refocus ourselves into a more practical ground of photometric stereo by pursuing a radiometric calibration method for uncalibrated photometric stereo where light source directions and intensities are unknown. Without light source directions, there exist multiple solutions of surface normals, light source directions, and response functions that satisfy the given photometric stereo images. Assume that the target object has Lambertian reflection property, we can express radiometrically calibrated images of the object as an image matrix with the rank of three that exhibits linear dependency property under linear response function. However, this assumption is no longer true under nonlinear response functions as the nonlinearity of the response function variates the images so that each element of the image matrix deviates from the linear model we assume. As the result, the image matrix loses its rank-3 structure and becomes full rank.

Here we propose a radiometric calibration method for uncalibrated photometric stereo. Given the response function is monotonicity increasing, there exists an inverse function that can restore the rank-3 property of the image matrix. Therefore, we formulate a singular value minimization problem with a rank constraint to find an inverse response function that produces the best rank-3 image matrix from the full-rank input matrix.

With this method, ones can compensate the nonlinearity of the response function even when the light source directions and intensities are not known beforehand. Therefore, our method allows existing uncalibrated photometric

stereo techniques to estimate surface normals without neither the calibration of light sources and the calibration of the camera response function. Moreover, we do not assume that the input images must be color and our method works well with grayscale objects. That is, our method allows one to estimate response function for each color channel separately.

We conducted experiments on both synthetic and real images to validate the inverse response function calibration of our method. We used existing uncalibrated photometric stereo techniques to estimate surfaces from images calibrated with our method and compared that to those estimated from radiometrically calibrated and uncalibrated images. The experimental results illustrated that our proposed method can compensate the nonlinearity of the response function used for capturing the scene even when light source directions and intensities are unknown so that the surfaces estimated from images calibrated with our method were similar to ones estimated from radiometrically calibrated images.

Acknowledgements

Earning the doctorate degree is a long and difficult process. I certainly could not finish it without the help and support from following people.

First of all, I owe my deepest gratitude to my advisor, Professor Yoichi Sato for his continuous support throughout my graduate study and his kindness encouragement. Discussion with him is always engaging and incredibly helpful. I would treasure this experience and I could not imagine having a better advisor.

Besides my advisor, I am grateful to Professor Takahiro Okabe who always give me insightful advice and helped me through my research countless times. His guidance from him during the beginning of my graduate study was the essential in this work.

I am grateful to Panasonic Scholarships who gave me the opportunity to pursue my degree in Japan and Sato Yo Scholarship Foundation who gave me support during my doctorate course. Moreover, the scholarships staffs always took a good care of me and gave helpful advice regarding living and studying in Japan.

I would like to thank all wonderful members of Sato-lab for their friendship. I am fortunate to have a chance to know and work with these talents. I also would like to thank the secretaries, Sakie Suzuki, Yoko Imagawa and Chio Usui for their kindness. They helped me from the start even before I

come to Japan. My gratitude also goes to Game Cloud Team at Line Corp. to allow me to have time to work on this thesis after my office hours. Also, I wish to sincerely thank all of my Thai friends in Tokyo for helping me get through many problems.

Finally, I would like to express my gratitude to my family: my parents, my sister, as well as my late grandmother for continued understanding and support. I could never have completed this thesis without them.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

3D scene reconstruction from 2D images is one of the fundamental problems in computer vision that has been studied extensively for a long time. Various methods have been developed to accurately recover the surface and geometry of a scene from given one or multiple images. Those surface recover methods can be categorized into two groups: geometry-based and photometry-based surface recovery methods.

Geometry-based methods make use of the relationship of geometry and its 2D projection on images to subsume the location and geometry of the surface. Well-known examples of these methods are including multi-view stereo vision, shape from structured light and time-of-flight camera. In contrast, photometry-based methods use of the relationship between the surface's reflection properties, scene illumination, light sources, and the light that is reflected into the camera to recover the surface orientation.

Photometric stereo is one of the photometry-based methods that recover surface orientations from multiple images of a stationary object taken un-

der varying light sources. The classic Woodham's photometric stereo[60] assumes that the object follows the Lambertian reflection model and that the directions and intensities of directional light sources are known a priori. These assumptions are difficult to achieve outside the laboratory and prevent photometric stereo from being used outside the computer vision community. Thus, current research trends in photometric stereo emphasize relaxing such assumptions; generalization of materials and light sources.

Another assumption that is usually taken for granted is that images are captured using a camera with a linear radiometric response function; *i.e.*, pixel values in the images are linearly proportional to irradiance values. Unfortunately, this assumption is not always true because response functions in consumer cameras are often nonlinear to enhance image quality visually, *e.g.*, compensating for a nonlinear mapping of a display, and simulating traditional films response. Moreover, these nonlinear relationships vary among camera models and manufacturers. Despite its benefit in aesthetics and displaying system, this nonlinearity may lead to unsatisfactory corrupted results in many computer vision methods which assume a linear camera response function. Therefore, cumbersome and time-consuming radiometric calibration is required to cancel the effect of the nonlinear relationships so that irradiance values can be subsumed from pixel values.

In this thesis, we discuss the relationship between radiometric response function calibration and surfaces reflection property. Concretely, we propose a framework in which allow surface recovery using photometric stereo from radiometrically uncalibrated images. Then, we extend the framework to handle corrupted measurements that do not follow Lambertian reflection. In a later chapter, we sought for more practical use of auto radiometric calibration by continuing our research onto uncalibrated photometric stereo

which the light source directions and intensities are unknown.

1.2 Overview of the Thesis

This thesis focuses on developing photometric-based surface recovery methods that invariant to nonlinear radiometric response function. The contributions of this thesis can be separated as follows:

1.2.1 Photometric Stereo with Auto Radiometric Calibration

First, we address the problem of Lambertian surface recovery from a set of radiometric uncalibrated images. Specifically, we propose a photometric stereo method to determine normals of Lambertian surfaces regardless nonlinearity of response function. The key idea is the consistency of irradiance images calculated by inverse response function and irradiance estimated from reflection model. We then formulate a linear optimization problem to simultaneous estimate surface normals and response function without additional images for radiometric calibration. With this method, one can recover surfaces of objects from uncalibrated images without aberration caused by nonlinear response function. We empirically show that our method can estimate surfaces normal from both synthetic images and real images were taken by a camera with a nonlinear response function.

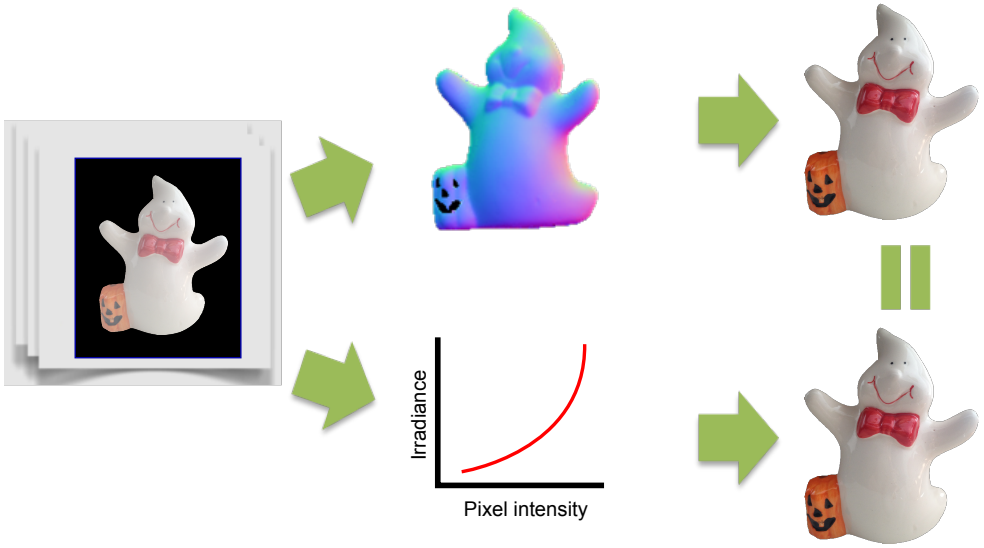


Figure 1.1: Photometric stereo with auto radiometric calibration using irradiance consistency. The irradiance estimated from surface property (top) must equal to the irradiance estimated from the inverse response function (bottom).

1.2.2 Non-Lambertian Surface Recovery from Radiometrically Uncalibrated Images

Second, we focus on the simultaneous estimation of non-Lambertian surfaces and radiometric response function from a set of radiometric uncalibrated images. In this work, we extend our method to handle the specular highlights, noise, and other corrupted measurements. We regard those pixel values as outliers to a reflection model, specifically Lambertian reflection model, and propose a framework to determine inliers, and surfaces at the same time. Our experimental results show that our method can remove distortion due to specularity while compensating the effect of nonlinearity of response function that resides in the input images.

1.2.3 Auto Radiometric Calibration in Uncalibrated Photometric Stereo

Last but not least, we study about auto radiometric calibration method using uncalibrated photometric stereo input images. We propose an auto radiometric calibration method which makes use of photometric properties to determine an inverse response function even when we do not have information about light source directions and intensities. We observe that pixel intensities form a matrix with the rank of 3 under a linear response function but the matrix is altered to be a full-rank when the response function is nonlinear. Given the response function is monotonically increasing, we can find an inverse response function that returns the matrix back into the rank of 3. We then use the rank constraint to formulate an optimization problem to determine the inverse response function with neither additional images and a priori knowledge about light source directions. The experimental results demonstrate that the proposed method can compensate the nonlinearity of the matrix so that the estimated surfaces were similar to the one estimated with calibrated images.

Chapter 2

Preliminaries

In order to establish a concrete foundation for the main subject of this dissertation, this chapter briefly explains fundamental knowledge of reflectance analysis and radiometric calibration.

2.1 Reflection

Reflection of a ray of light that fall onto an object surface can be classified into two categories: *diffuse reflection* and *specular reflection*. Consider a light ray fall onto a surface with rough micro-structure. The light may multiple-scatters inside the microstructure and sub-surface before scatter back to outside in broad directions. This kind of reflection is called diffuse reflection. On the other hand, the ray of light fall may directly reflect off the surface is known as specular reflection. The specular reflection mainly exhibit mirror-like reflection in which the incident light and the reflected light have the same angle respect to the reflected surface.[1]

The brightness of the reflected light depends on its microstructure, sur-

[1]We assume that the surface point where the light incident and the light ray emitted are sufficiently close.

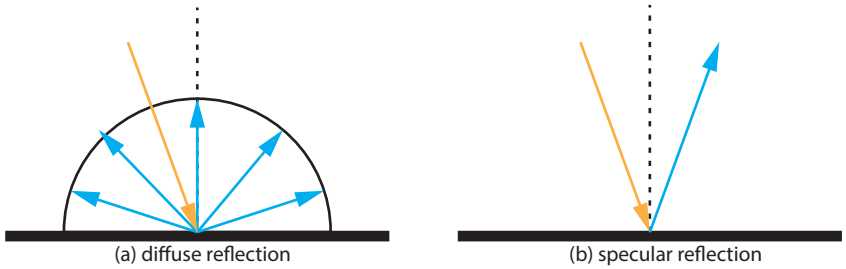


Figure 2.1: Reflections on surfaces (left) diffuse reflection (right) specular reflection

face orientation, and state of polarization of the incident illumination [60]. Moreover, the brightness perceived by an observer also depends on the direction that the reflected light emitted to. Therefore, we can model the reflectance with a function of the incident angle (θ_{in}, ϕ_{in}) and the reflected angle $(\theta_{out}, \phi_{out})$ called *Bi-directional Reflectance Distribution Function* (BRDF). A BRDF $\rho(\theta_{in}, \phi_{in}, \theta_{out}, \phi_{out})$ can be expressed as,

$$\rho(\theta_{in}, \phi_{in}, \theta_{out}, \phi_{out}) = \frac{E(\theta_{out}, \phi_{out})}{I(\theta_{in}, \phi_{in})}, \tag{2.1}$$

which $I(\theta_{in}, \phi_{in})$ is the irradiance, the light that incidents on the surface in (θ_{in}, ϕ_{in}) direction, and $R(\theta_{out}, \phi_{out}$ is the radiance emitted by the light reflected from the incidented surface into $(\theta_{out}, \phi_{out})$ direction.

2.1.1 Diffuse Reflection

For rough surfaces like chalk and plaster, their micro-structures are very rough so the light reflected off the surfaces to every direction equally, *i.e.*, constant regardless observing directions. The diffuse reflection component of the surfaces can be approximately expressed by so called *Lambertian model*. The BRDF of such surfaces then constant respect to $(\theta_{in}, \phi_{in}, \theta_{out}, \phi_{out})$. The

brightness of the surface then solely depends on the radiance that falls on to the surface which can be given by:

$$E = max(0, L\rho(\cos\theta)) \tag{2.2}$$

where L is the light source intensity and the $cos\theta$ is the angle difference between the light source respect to the surface normal. When $L\rho(\cos\theta) < 0$, it means the light direction is on the opposite side of the surface normal so there is no incident light from that light source onto the surface.

Conventionally, we can write 2.2 in vector form as an inner product of incident light direction scaled by its intensity $\boldsymbol{l}$ and surface normal scaled by its albedo $\boldsymbol{n}$:

$$E = \boldsymbol{l} \cdot \boldsymbol{n}. \tag{2.3}$$

2.1.2 Specular Reflection

An object with very smooth surfaces behaves like a mirror so it precisely reflects the incident light into the same angle respect to its surface normal. This mirror-like reflection is known as specular reflection.

When the surface is not perfectly smooth, the incident light may not be reflected into a particular angle and this result blurred specular highlights instead of sharp edges. This blurred highlights can be explained by the existence of microfacets[57]. Microfacets are tiny facets which each of them acts as a perfect specular reflector. Therefore, the shininess of an object can be expressed as a distribution of the angle of microfacets normals in an area.

Typically, we can assume that specular highlight can be observed within small angles and the reflection consist of the additive combination of diffuse

component and specular component. Therefore we can model a reflection model with specular highlight as following:

$$E_{total} = k_d E_{diffuse} + k_s E_{specular}. \tag{2.4}$$

Here, k_d and k_s are diffuse and specular component contants. They represent the ratio of reflection of diffuse components and specular component to the incoming light, respectively.

Phong[51] proposed a model to determine the specular component at a point given an observed direction $\hat{\boldsymbol{v}}$ and the reflected light direction scaled by the light source intensity $\hat{\boldsymbol{r}}$ as,

$$E_{specular} = (\boldsymbol{r} \cdot \boldsymbol{v})^{\alpha}. \tag{2.5}$$

Here, α represents the shininess or distribution of microfacets of the surface at that point. Subsitute 2.5 in 2.6 and assume the diffuse component follow Lambertian reflection model, we can write Phong reflection model as,

$$E_{Phong} = k_d (\boldsymbol{l} \cdot \boldsymbol{n}) + k_s (\boldsymbol{r} \cdot \boldsymbol{v})^{\alpha}. \tag{2.6}$$

There are numerous reflection models that explain the reflection closer to physical reality, *e.g.*, Cook and Torrance[10] developed a model that accounts for microfacets distribution, material's refraction index, and geometry attenuation function to represent the ratio of light that remains after the reflection. Ward BRDF model[58] uses asymmetric distribution to model the reflection of anisotropic surfaces which the surface property depends on the observed direction.

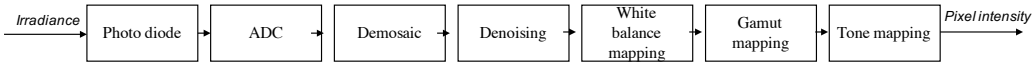


Figure 2.2: An image acquisition pipeline

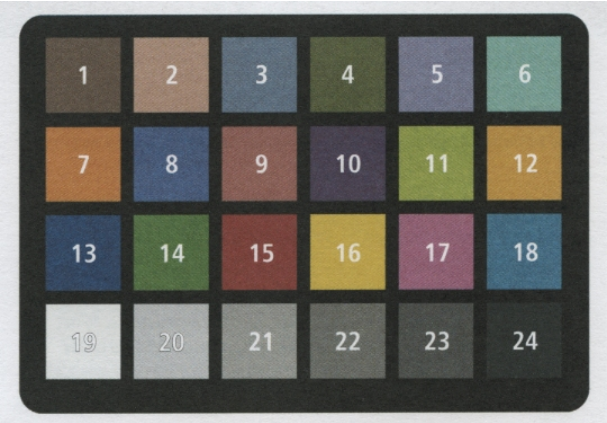
2.2 Radiometric Calibration

This section illustrates an overview of radiometric calibration and radiometric response functions in cameras.

Let us consider a typical digital image processing pipeline[53]. As depicted in Fig. 2.2, after light incidents on a camera sensor, the sensor measures the light energy and converts it into electrical energy. Then, the camera performs various image signal processing such as demosaicking, noise reduction, white balancing, and other post-processing to enhance the image, mainly for aesthetic reasons. Eventually, the processed signal is transformed into a brightness level in form of pixel intensity. Here, we can treat aforementioned multi-step signal processing as a so-called camera radiometric response function that transforms the measured incident light brightness into pixel intensities. Let E an irradiance level detected by a camera sensor and its corresponding pixel intensity I on the recorded image. Response functions f can be expressed as

$$I = f(E). \tag{2.7}$$

Generally, physics analysis of scenes expects light brightness levels in form of image irradiance. If the scenes are captured by a linear response function, we can directly use pixel intensities instead of image irradiance due to their linear relationship. However, this is not necessarily always true so we need to cancel out the effect of nonlinear response function by radiometric calibration. Since the radiometric response function is a monotonically increasing



No.	Number	sRGB R	sRGB G	sRGB B	CIE L*a*b* L*	CIE L*a*b* a*	CIE L*a*b* b*	Munsell Notation Hue Value / Chroma
1.	dark skin	115	82	68	37.986	13.555	14.059	3 YR 3.7 / 3.2
2.	light skin	194	150	130	65.711	18.13	17.81	2.2 YR 6.47 / 4.1
3.	blue sky	98	122	157	49.927	-4.88	-21.925	4.3 PB 4.95 / 5.5
4.	foliage	87	108	67	43.139	-13.095	21.905	6.7 GY 4.2 / 4.1
5.	blue flower	133	128	177	55.112	8.844	-25.399	9.7 PB 5.47 / 6.7
6.	bluish green	103	189	170	70.719	-33.397	-0.199	2.5 BG 7 / 6
7.	orange	214	126	44	62.661	36.067	57.096	5 YR 6 / 11
8.	purplish blue	80	91	166	40.02	10.41	-45.964	7.5 PB 4 / 10.7
9.	moderate red	193	90	99	51.124	48.239	16.248	2.5 R 5 / 10
10.	purple	94	60	108	30.325	22.976	-21.587	5 P 3 / 7
11.	yellow green	157	188	64	72.532	-23.709	57.255	5 GY 7.1 / 9.1
12.	orange yellow	224	163	46	71.941	19.363	67.857	10 YR 7 / 10.5
13.	blue	56	61	150	28.778	14.179	-50.297	7.5 PB 2.9 / 12.7
14.	green	70	148	73	55.261	-38.342	31.37	0.25 G 5.4 / 8.65
15.	red	175	54	60	42.101	53.378	28.19	5 R 4 / 12
16.	yellow	231	199	31	81.733	4.039	79.819	5 Y 8 / 11.1
17.	magenta	187	86	149	51.935	49.986	-14.574	2.5 RP 5 / 12
18.	cyan	8	133	161	51.038	-28.631	-28.638	5 B 5 / 8
19.	white (.05*)	243	243	242	96.539	-0.425	1.186	N 9.5 /
20.	neutral 8 (.23*)	200	200	200	81.257	-0.638	-0.335	N 8 /
21.	neutral 6.5 (.44*)	160	160	160	66.766	-0.734	-0.504	N 6.5 /
22.	neutral 5 (.70*)	122	122	121	50.867	-0.153	-0.27	N 5 /
23.	neutral 3.5 (1.05*)	85	85	85	35.656	-0.421	-1.231	N 3.5 /
24.	black (1.50*)	52	52	52	20.461	-0.079	-0.973	N 2 /

Figure 2.3: (left) X-Rite ColorChecker Color Redition Chart. (right) Colorimetric values for each patch in the charts.[63]

function, there is a unique inverse function g that maps an I to an E, *i.e.*,

$$g(I) = f^{-1}(I) = E. \tag{2.8}$$

so we can apply g to the captured images to cancel the response function's effect.

A conventional method to determine response function makes use of a calibration target with known reflectance and color such as a color chart (Fig.2.3). We can estimate the response function by measuring the deviation of patches of the color chart in the captured images to their reference value.

One of the research directions in radiometric calibration focuses on relaxing the requirement of the calibration target, *i.e.*, radiometric calibration without a calibration target. The most well-known approach uses multiple photos of a static scene taken under varying exposure times [35, 11, 40, 17]. If the response function is constant during the capturing of the scene, it can be estimated with the relationship between the exposure times and the recorded pixel intensities. Manning and Picard [35] models this nonlinear response functions with a gamma curve and used a regression method to estimate the

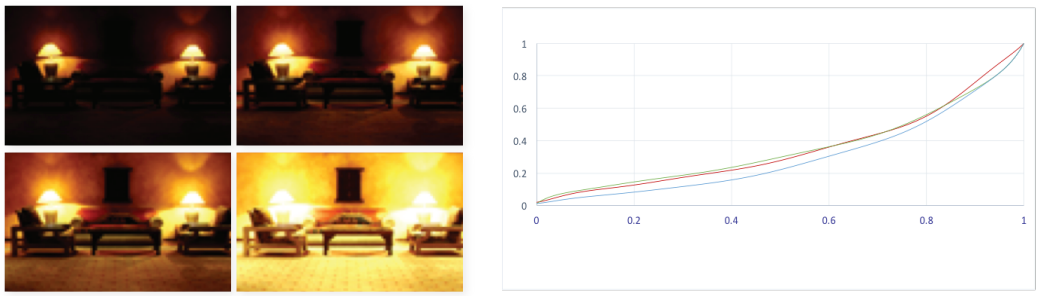


Figure 2.4: Radiometric calibration using multiple images (a) examples of input images (b) the estimated response functions for RGB channels. Images from Mitsunaga and Nayar [40]

curve function parameters. Devebec and Malik [11] used a non-parametric model with a smoothness term to approximate a response function in order to produce HDR images.

Mitsunaga and Nayar [40] proposed a method that represents a response function with a linear combination of polynomials. This method can accurately estimate a response function even when only rough estimates of exposure times are available. Grossberg and Nayar [17] studies the space of response function of camera films and digital sensors. Then they developed an empirical model that can represent most of the response functions with only a few basis functions.

Recent works in radiometric calibration focus on relaxing the assumptions to allow fewer restrictions on input images, *e.g.*, requiring only a single input image, or allowing camera movement. Lin *et al.* [30] proposed a method for estimating a radiometric response function from a single color image by using color mixture near edges of regions. The color gradation near edge region should form a straight line in the RGB space. However, a nonlinear response function bends this line into a nonlinear curve so the IRF can be estimated as a function that bends the nonlinear curve back into a straight line. For

grayscale images, intensity mixtures along edge boundaries exhibit a similar property as the color mixtures so can use a distribution of intensity mixtures instead of using color mixtures [31]. And instead of spatial irradiance mixture, Wilburn *et al.* [59] made use of temporal irradiance mixtures and developed a method for radiometric calibration from motion blur in a single image. Kim and Pollefeys [25] estimated a radiometric response function from an image sequence captured using a moving camera. Because irradiance values observed at a certain point should not change when the camera moves, the changes in pixel values can be used to compute the changes in exposure times and the response function. Matsushita and Lin [36] suggested that noise distributions in uncalibrated images are not symmetry due to nonlinear response functions. They developed a calibration method to estimate an IRF that makes the noise distribution symmetry. Recently, Lee *et al.* [27] have formulated the radiometric calibration from multiple images with different exposure times as a rank minimization problem. This method is more robust to noise and can estimate response functions more accurately than previous methods.

Chapter 3

Photometric Stereo with Auto Radiometric Calibration

3.1 Background

Photometric stereo is a method for estimating the surface orientations of a static object from a set of images taken from a fixed viewpoint but under different lighting environment. The classic Woodham's photometric stereo method[60] assumes that surfaces follow the Lambertian reflection model, and the directions and intensities of directional light sources are known a priori. These assumptions are difficult to achieve outside the laboratory and prevent photometric stereo from being used outside the computer vision community. Thus, current research trends in photometric stereo emphasize relaxing such assumptions; *i.e.*, generalizing materials and light sources.

Another assumption that is usually taken for granted is that images are captured using a camera with a linear radiometric response function; *i.e.*, pixel values in the images are linearly proportional to irradiance values. Unfortunately, this assumption is not always true because response functions in

consumer cameras are often nonlinear to enhance image quality visually, *e.g.*, compensating for a nonlinear mapping of a display, and simulating traditional films response. Moreover, these nonlinear relationships vary among camera models and manufacturers. Therefore, cumbersome and time-consuming radiometric calibration is required to cancel the effect of the nonlinear relationships so that irradiance values can be subsumed from pixel values.

In this thesis, we propose an auto radiometric calibration framework for photometric stereo techniques to estimate surface orientations from a sequence of radiometrically uncalibrated images. The main idea is to use the consistency of irradiance; the irradiance values converted from pixel values by using the inverse response function (IRF) must be consistent with the irradiance values computed by the corresponding surface normals and reflection property, specifically the Lambertian reflection model. In other words, we exploit the cues inherent in the image formation process to estimate the IRF as well as surface normals. Then, we derive a linear least-squares problem with linear constraints to simultaneously determine surface normals and response functions. Our experiments show that our method allows surface normals to be estimated from a set of radiometrically uncalibrated images regardless of nonlinearity of a response function.

The contribution of this research is an auto radiometric calibration framework for photometric stereo methods. This framework requires no additional images for radiometric calibration. Therefore, it allows one to estimate surface orientations from images taken using consumer cameras without additional effort to calibrate nonlinearity of radiometric response functions.

This chapter is organized as follows: Section 3.2 gives a brief overview of previous studies on radiometric calibration, photometric stereo methods, and relationships to our method. Section 3.3 presents our photometric stereo with

auto radiometric calibration method for Lambertian surfaces. Section 3.4 verifies the surface normal and response function estimation accuracy with experiments on both synthetic images and images taken from a real camera. Finally, we conclude this Chapter in Section 3.5.

3.2 Related Work

In this section, we briefly review previous works on radiometric camera calibration and how they are related to our present work.

Originally, one can use an image of a calibration target with known reflectances [63]. The response function can be estimated from the relationship between the referenced reflectances of the calibration target and their corresponding pixel values in the image.

An alternative approach uses multiple images of a static scene taken with varying exposure times [35, 11, 40, 17]. Manning and Picard [35] used a regression method to estimate parameters of a response function represented by a gamma correction curve. Devebec and Malik [11] developed a method for HDR images processing. Their method employs a non-parametric model with a smoothness term to approximate a response function in order to produce HDR images. Mitsunaga and Nayar [40] proposed a radiometric calibration method which can accurately estimate a response function even when only rough estimates of exposure times are available. They also suggested that a response function can be estimated with a linear combination of polynomials. Grossberg and Nayar [17] studies the space of response function and proposed an empirical model that can capture most of the response functions with a few basis functions.

Recent works in radiometric calibration have fewer restrictions on input

images, *e.g.*, requiring only a single input image, or allowing camera movement. Lin *et al.* [30] proposed a method for estimating a radiometric response function from a single color image. Color mixtures in edge regions lie on a straight line in the RGB space. A nonlinear response function bends this line into a nonlinear curve. The IRF can be estimated as a function that bends the nonlinear curve back into a straight line. For grayscale images, one can use a distribution of intensity mixtures along edge boundaries instead of using color mixtures [31]. Wilburn *et al.* [59] conducted radiometric calibration on the basis of motion blur in a single image. They made use of temporal irradiance mixtures instead of spatial ones. Grossberg and Nayar [17] argued that radiometric calibration can be done using the pixel value histograms of two image frames without exact registration. Matsushita and Lin [36] suggested that the noise distributions in uncalibrated images of a scene are symmetry under a linear response function but the noise distributions become asymmetric due to nonlinear response functions. So they developed a calibration method to estimate an IRF that returns the noise distribution back into symmetric one.

Although we can conduct radiometric calibration by using these existing methods separately, they are often cumbersome and require additional images used for radiometric calibration. Unlike these methods, our proposed method does not require additional images. More importantly, our method is based on a different cue; we take advantage of the image formation process for radiometric calibration.

Previously, radiometric calibration and photometric stereo have been studied separately. As far as we know, Shi *et al.* [55] was the first to use photometric stereo images for radiometric calibration. They calibrate input images for photometric stereo by using color profile linearity. A color profile

is the RGB values of a surface point under varying illumination conditions. They observed that a color profile of a Lambertian point forms a straight line under the RGB space if the response function is linear. In contrast, color profiles become nonlinear curves if the response function is nonlinear. Therefore, the IRF can be estimated as a function that makes the nonlinear curve into a straight line. [27] used a similar idea to cast the radiometric calibration in photometric stereo image sets into a low-rank matrix completion problem. They unroll images into matrices in which each consists of three-row vectors for R, G, and B channels. These matrices give lowest rank if response function is linear and the rank is increased under nonlinear response function. Therefore, the IRF can be determined by finding a function that minimizes the rank.

Although their methods do not require additional images, the radiometric calibration step can be considered as a preprocessing and requires nonlinear optimization. More importantly, their methods cannot handle a certain class of radiometric response functions such as gamma correction curves as well as gray objects whose color profiles remain in straight lines even when the response function is nonlinear.

3.3 Recovering Lambertian Surfaces

3.3.1 Photometric Stereo

We briefly summarize the classic photometric stereo method [60] in which radiometrically calibrated input images of Lambert surfaces illuminated by known light sources are assumed. Let us consider a surface observed under orthographic projection and light sources far from the surface, then the viewing direction and light source directions are constant across the surface.

Let us denote the irradiance value at the p-th pixel ($p = 1, 2, \ldots, P$) under the d-th light source ($d = 1, 2, 3, \ldots, D$) by $E_{p,d}$. Assuming that the surface is Lambertian, the irradiance value $E_{p,d}$ can be described as

$$E_{p,d} = \mathbf{n}_p^T \mathbf{l}_d, \tag{3.1}$$

where $\mathbf{n}_p$ is the surface normal at the p-th pixel scaled by its albedo and $\mathbf{l}_d$ is the direction of the d-th light source scaled by its intensity.

The classic photometric stereo method estimates the scaled surface normal $\mathbf{n}_p$ from the irradiance value $E_{p,d}$ with known light sources $\mathbf{l}_d$. Because a scaled surface normal $\mathbf{n}_p$ has three degrees of freedom, *i.e.*, two for a normal with unit length and one for an albedo, the surface normal can be estimated from at least three images.

Conventionally, the irradiance value $E_{p,d}$ in (3.1) is described in a matrix form:

$$\begin{aligned} \begin{pmatrix} E_{p,1} \\ \vdots \\ E_{p,D} \end{pmatrix} &= \begin{pmatrix} l_{1,x} & l_{1,y} & l_{1,z} \\ \vdots & \vdots & \vdots \\ l_{D,x} & l_{D,y} & l_{D,z} \end{pmatrix} \begin{pmatrix} n_{p,x} \\ n_{p,y} \\ n_{p,z} \end{pmatrix}, \\ \boldsymbol{E}_p &= \boldsymbol{L}\mathbf{n}_p, \end{aligned} \tag{3.2}$$

where $\mathbf{l}_d = (l_{d,x}, l_{d,y}, l_{d,z})^T$ and $\mathbf{n}_p = (n_{p,x}, n_{p,y}, n_{p,z})^T$. The estimate of the scaled surface normal $\hat{\mathbf{n}}_p$ is given by the least-square method:

$$\hat{\mathbf{n}}_p = \left(\boldsymbol{L}^T \boldsymbol{L}\right)^{-1} \boldsymbol{L}^T \boldsymbol{E}_p. \tag{3.3}$$

This is equivalent to

$$\hat{\mathbf{n}}_p = \arg\min_{\mathbf{n}_p} \sum_{d=1}^{D} \left(E_{p,d} - \mathbf{n}_p^T \mathbf{l}_d \right)^2 . \tag{3.4}$$

The surface normal and albedo are computed from the estimated scaled surface normal $\hat{\mathbf{n}}_p$ as $\hat{\mathbf{n}}_p / |\hat{\mathbf{n}}_p|$ and $|\hat{\mathbf{n}}_p|$ respectively.

3.3.2 Radiometric Response Function

Suppose a radiometric response function f maps an irradiance value E into a pixel value I, *i.e.*, $I = f(E)$. Since the radiometric response function is a monotonically increasing function, there is a unique inverse function $g = f^{-1}$ that maps an I to an E, *i.e.*, $g(I) = f^{-1}(I) = E$. Hereafter, we normalize the ranges of pixel values and irradiance values to $[0, 1]$ without loss of generality.

Assume an IRF g can be approximated as a linear combination of basis functions such as polynomials [40], or eigenvectors (or eigenfunctions in this context) of response curve data [17]. In this work, we use the K-parameters EMoR approximation[17] in which the basis functions are derived from Principle Component Analysis (PCA) on the real world response function database. The approximated IRF is in the form:

$$g(I) = g_0(I) + \sum_{k=1}^{K} c_k g_k(I), \tag{3.5}$$

subjects to the boundary conditions, $g(0) = 0$ and $g(1) = 1$. Here g_0 is the mean curve or the 0-th basis function, and g_k is the k-th basis functions with their coefficients c_k.

3.3.3 Simultaneous Estimation

We propose a technique that uses irradiance consistency to estimate both the surface normals $\mathbf{n}_p$ $(p = 1, 2, ..., P)$ and coefficients of the IRF c_k $(k = 1, 2, ..., K)$ at the same time.

With calibrated images, we directly substitute the irradiance $E_{p,d}$ in (3.4) with the corresponding pixel value $I_{p,d}$. For uncalibrated images, we can compensate the nonlinearity of the response function by substituting $E_{p,d}$ with its irradiance approximated by the IRF shown in (3.5) Then the relationship between Lambertian reflection property and pixel intensity can be given by,

$$\begin{aligned} E_{p,d} &= g(I_{p,d}) \\ \mathbf{n}_p^T \mathbf{l}_d &= g_0(I_{p,d}) + \sum_{k=1}^{K} c_k g_k(I_{p,d}). \end{aligned} \tag{3.6}$$

Using this relationship, we can simultaneously estimate surface normals $\hat{\mathbf{n}}_p$ and IRF $\hat{g}$ minimizing the difference in irradiance values estimated from surface normals and IRF. More specifically, we combine the PCA approximation of irradiance in (3.6) into (3.4) to obtain the estimates of the surface normals and coefficients of the IRF $\{\hat{c}_k\}$:

$$\{\hat{\mathbf{n}}_p, \hat{c}_k\} = \arg \min_{\{\mathbf{n}_p, c_k\}} \sum_{p=1}^{P} \sum_{d=1}^{D} \left[g_0 \left(I_{p,d}\right) + \sum_{k=1}^{K} c_k g_k \left(I_{p,d}\right) - \mathbf{n}_p^T \mathbf{l}_d \right]^2 . \tag{3.7}$$

As response functions and their inverse are monotonicity non-decreasing function, eq.(3.7) subjects to the constraints $\partial g / \partial i \geq 0$, which can be given as discrete derivatives of g as $g\left(i_{s+1}\right) - g\left(i_s\right) \geq 0$. The monotonicity constraints

can be derived as follows:

$$
\begin{aligned}
& g\left(i_{s+1}\right) - g\left(i_s\right) \geq 0, \\
& \left(g_0\left(i_{s+1}\right) + \sum_{k=1}^{K} c_k g_k\left(i_{s+1}\right)\right) - \left(g_0\left(i_s\right) + \sum_{k=1}^{K} c_k g_k\left(i_s\right)\right) \geq 0, \\
& \sum_{k=1}^{K} \left(c_k g_k(i_{s+1}) - c_k g_k(i_s)\right) \geq \left(g_0\left(i_s\right) - g_0\left(i_{s+1}\right)\right), \qquad (3.8)
\end{aligned}
$$

$\forall i_s \in I_S$, where $I_S = \{i_0, \ldots, i_S\}$ is the dense sampling of pixel intensities in the range $[0, 1]$ in monotonically increasing order, *e.g.*, $I_S = \{0, \frac{1}{255}, \frac{2}{255}, \ldots, \frac{255}{255}\}$ for 8-bit images.

Thus, the simultaneous estimation of the surface normals and the IRF results in the linear least-square problem in (3.7) with the linear constraints in (3.8) given that the input images are well-exposed, *i.e.*, the input pixel values distributed uniformly to cover the whole range of pixel intensity $[0, 1]$.

When the radiometric response function is linear, we can estimate a surface normal at each surface point independently, as shown in (3.2). On the other hand, when the radiometric response function is nonlinear, we can no longer estimate each surface point independently as all pixels subject to a common response function. Therefore, the naive optimization of (3.7) subject to the constraints of (3.8) is computationally expensive when the number of pixels increases. Given all pixel values in an image are affected with the same response function and the number of foreground pixels is greatly larger than the number of basis functions ($PD \gg K$), we could reduce the computational cost by estimating the IRF (and the surface normals) with a small number of randomly selected pixels. Using the estimated IRF, all the pixel values now can be converted to irradiance values, so we can finally estimate the surface normals of all foreground points using (3.4).

Note that our simultaneous estimation also has degenerate cases; it fails to estimate an IRF if pixel values are not well distributed, *e.g.*, estimating an IRF from images of a plane which is illuminated by light sources rotated around its perpendicular axis so the pixel values remain constant across all images. In such scenes, the nonlinear relationship between irradiance and pixel values cannot be observed from the images; thus the IRF cannot be estimated.

3.4 Experiments

3.4.1 Experiments on Synthetic Images

We validated the surfaces and IRF estimation of our method with the experiments on synthetic images of a Lambertian sphere. The images were illuminated with 20 directional light sources whose directions were uniformly selected from a hemisphere. The uncalibrated images were obtained by applying nonlinear response functions from DoRF. We detected shadows in all images with thresholding, *i.e.*, a pixel value is considered in shadow if its intensity is less than $\frac{5}{255}$. Then, we discarded all pixels that consisted of only one or two non-shadow pixel values from the estimation.

We implemented the optimization of (3.7) and its constraint (3.8) by using a MATLAB's built-in function *fmincon* to illustrate that our method can improve surface estimation accuracy over previous methods despite its sensitivity to outliers. The number of basis functions was fixed to $K = 4$. We used first four functions derived from DoRF database with PCA technique as our basis functions. To make the estimation tractable, we first estimated an IRF from 50 randomly selected pixels to calibrate the images then estimated surface normals with (3.4). The surface normals and IRF estimation took

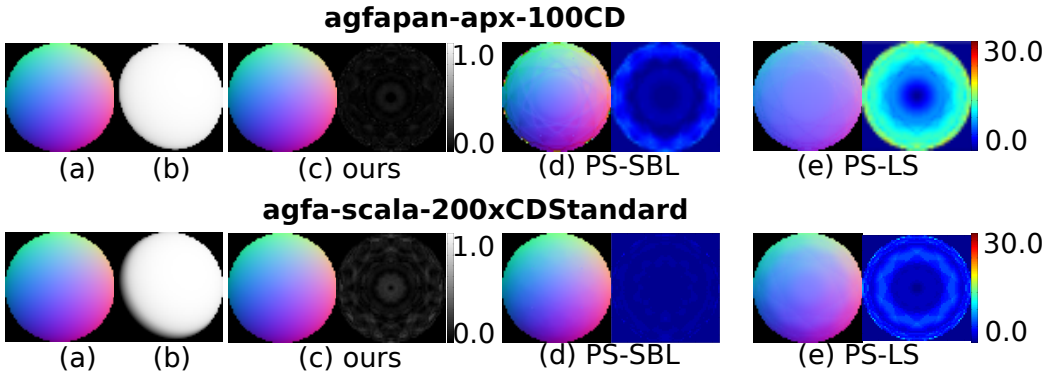


Figure 3.1: Samples of surface estimated from synthetic data sets and the comparison to the ground truth. The response function used to generate the images are specified over the results. (a) ground truth (b) examples of input images (c)-(e) Estimated surface normals and difference to the ground truth in degrees.

Table 3.1: Comparison of average of mean angular error and their SD of the surface normals estimated from the synthetic data sets using our proposed method (ours), photometric stereo with sparse Bayesian regression (PS-SBL), and the classic photometric stereo (PS-LS).

Method	Mean angular error (degrees)		Elapsed Time (secs.)
	Uncalib. images	Calib. images	
ours	$\mathbf{0.68 \pm 0.71}$	0.53 ± 0.23	2.18 ± 1.35
PS-SBL	4.42 ± 4.203	0.002 ± 0.003	27.83 ± 0.15
PS-LS	12.09 ± 3.11	$\mathbf{0.0002 \pm 0.004}$	$\mathbf{0.39 \pm 0.01}$

about 6 seconds on an Intel Core i7-2600 3.4GHz machine for each set of images.

To assess surface estimation accuracy, we compared the surfaces estimated using our method (ours) to the surfaces estimated from radiometrically uncalibrated images with the classic photometric stereo (PS-LS) [60], and photometric stereo with sparse Bayesian learning (PS-SBL) [24] to illustrate the advantage over the state-of-the-art method. The parameters for PS-SBL were configured as suggested in the original paper, *i.e.*, $p = 3$, $\lambda_{sbl} = 10^{-2}$, $\sigma_a^2 = 10^{-2}$, and $\sigma_n^2 = 10^6$.

The mean angular error of the estimated surfaces are summarized in Table 3.1 and few examples of the estimated surfaces and their difference in degrees to the ground truth are shown in Fig. 3.1 as RGB color-coded surface normals along with their difference to the ground truth. It is clear that our auto calibration framework compensates nonlinear response functions so the estimated surfaces are more similar to the ground truth than one of those estimated using PS-LS and PS-SBL. Meanwhile, the nonlinear reflectance model used in PS-SBL compensates the nonlinearity of response functions, therefore the mean of angular errors of the estimated surfaces was significantly lower than those estimated by PS-LS.

We performed the experiments with several numbers of images with a different combination of images in each iteration. Results summarized in Fig. 3.2 shows our method can accurately estimate surfaces even when we used only four images and the accuracy increased when we used more images.

We also conducted experiments on radiometrically calibrated images. As now the surfaces certainly follow Lambertian reflection model, it is unnecessary to address nonlinear reflectance function so we fixed the number of piecewise reflectance function $p = 1$ for PS-SBL. The experimental results show that the difference of the surfaces estimated by the proposed method from calibrated images is comparable to (or slightly smaller than) that for uncalibrated images. However, the surfaces estimated by PS-SBL and PS-LS have less error than those estimated from our proposed method. This is because the basis functions we used cannot represent a linear function well, so the estimated IRFs had a subtle deviation from the linear IRF. Those small errors eventually propagated to the subtle errors on every point of the estimated surfaces. In contrast, PS-LS and PS-SBL assume calibrated images and the input images were images of Lambertian surfaces without noise

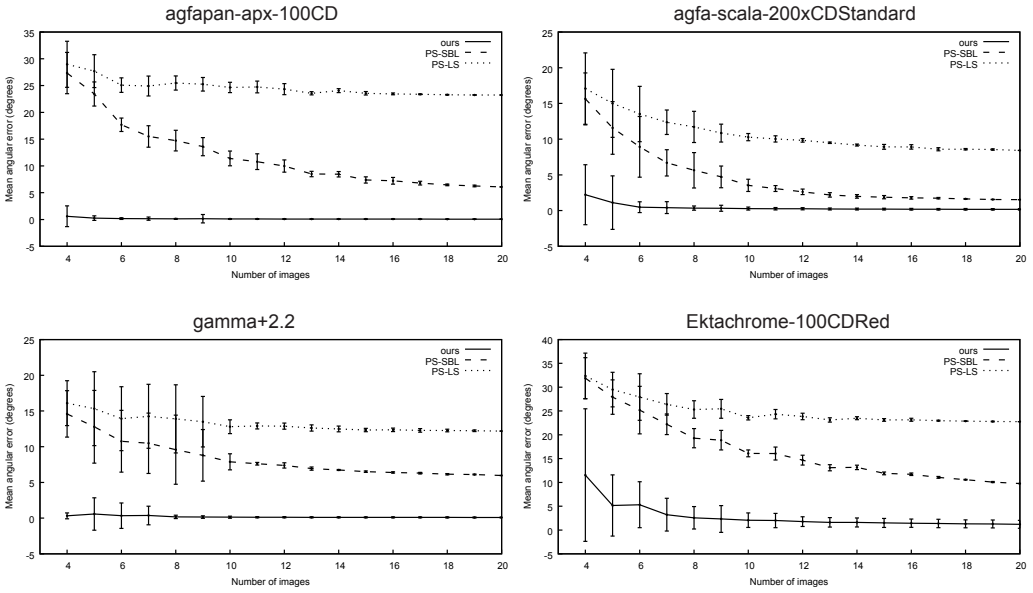


Figure 3.2: Comparison of average of mean angular error in degrees of the surface normals estimated from the synthetic data sets with various number of input images. The error bars show the standard deviation of the angular errors.

so that the only precision error become the source of error in the estimated surfaces.

Also, we evaluated the IRF estimation accuracy of our method to [40] with EMoR representation [17]. Note that we used images of a static scene with varying exposure times to estimate IRF with [40], hence, indirect comparison. Here, we show some examples of the estimated IRFs fitted to their corresponding ground truth in Fig. 3.3. The root mean square error (RMSE) of the fitted IRFs are also shown in the figure. The average of RMSE for all synthetic images sets is 0.0134 while the RMSE of the IRFs estimated with [40] is 0.0098.

However, it is worth to mention that our method can estimate the IRFs accurately up to scale. When selected irradiance values do not well cover the entire range of irradiance values, there is little information to constrain

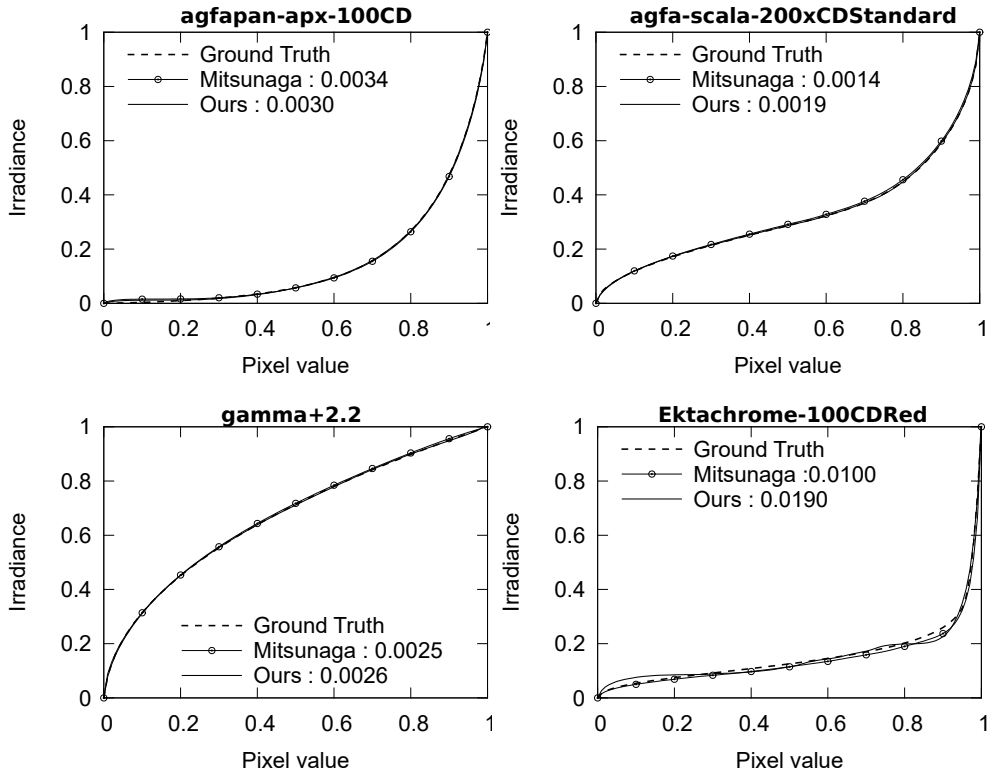


Figure 3.3: The IRFs estimated by *ours* from synthetic data sets. The dotted lines show ground truth and the solid lines show the estimated functions. The RMSE for each IRF are depicted over the graph.

the estimated IRF. Therefore it is possible to obtain multiple IRFs with different scales that satisfy the selected irradiance values. In our experiments on synthetic images, this scaling ambiguity did not affect the estimated IRF but affects the overall scale of the estimated albedos.

3.4.2 Experiments on Real Images

We evaluated the accuracy of our method through the experiments on images of two real-world objects: *sphere*, and *statue*. We used 20 light sources placed randomly over the objects and calculate the light source direction from images of a chrome sphere. Then we used images of a Lambertian

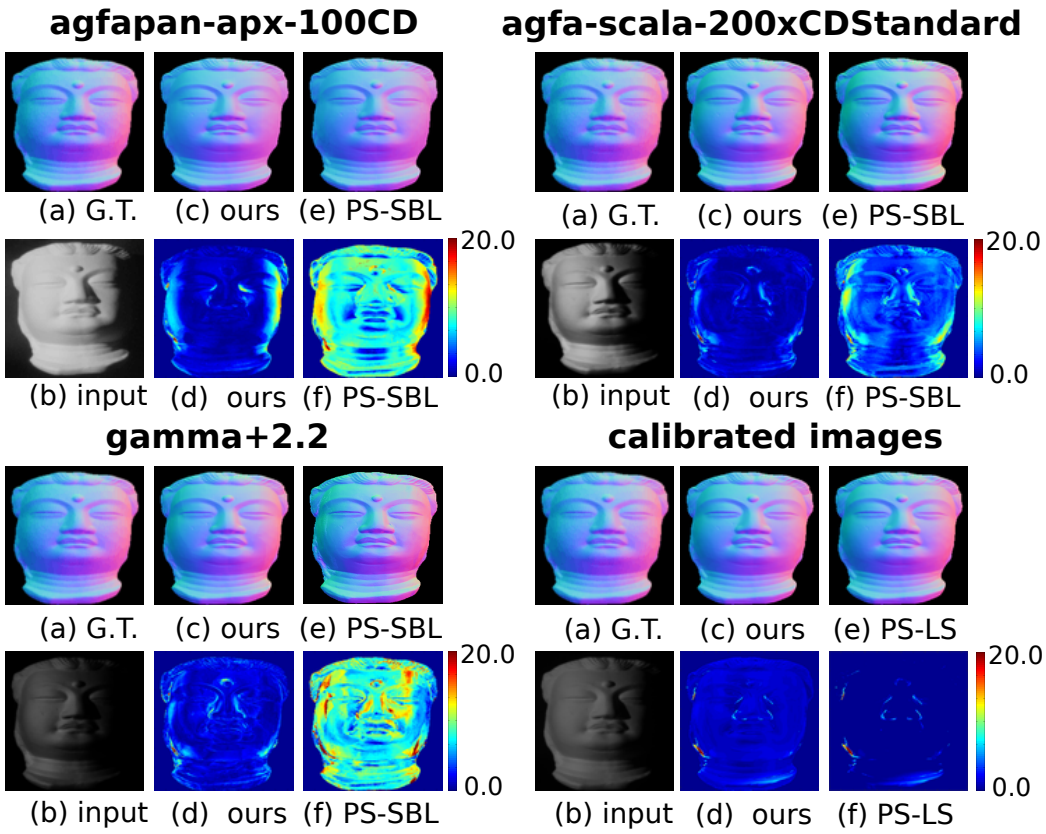


Figure 3.4: The estimated surface normals *statue*. The name of response functions are specified in the top row and the method used to estimated the surfaces are specified in the bottom. (a) ground truth (b) sample input image (c)-(f) estimated surfaces and difference to the ground truth in degrees.

sphere taken with a calibrated camera to calculate the light source intensities with (3.1). Instead of using calibrated images, ones can use a luminometer to directly measure the light source intensities.

We captured images of the objects with a Point Grey Flea 2 camera. Although the camera provides linear measurements of light intensity, we can configure its intensity lookup table so it acts like a nonlinear camera. We configured the lookup table with the measured agfapan-apx-100CD, agfa-scala-200xCDStandard, and gamma+2.2 nonlinear response functions from DoRF database [17]. These functions represent three common shapes of

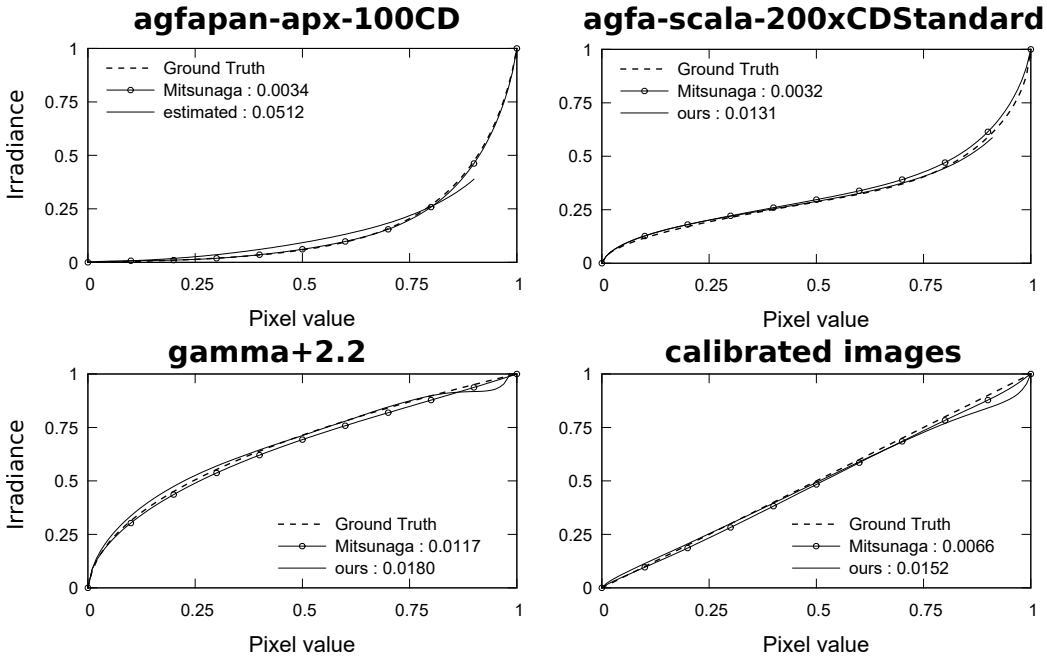


Figure 3.5: The IRFs estimated from *statue* with [40] and ours compare to the ground truth. The RMSEs of the estimated IRFs are included in the legend.

nonlinear functions in the database: concave, convex, and sigmoid. Since the ground truth of *sphere*, and *statue* were not available, we used the surfaces estimated from the radiometrically calibrated input images with PS-SBL [24] as ground truth.

The mean of angular error of the estimated surfaces are presented in Fig. 3.6. Similar to the experiments with the synthetic images, our auto radiometric calibration compensated the nonlinearity of response functions so the estimated surfaces were more similar to the ground truth than those estimated by PS-LS and PS-SBL from the uncalibrated images. However, it is possible that shadow pixel values were raised by nonlinear response functions and became non-shadow in pixel value space. Eventually, those false pixel values were included in the surface estimation and deviated the surfaces. This can be observed in the side parts of the estimated *statue (agfapan-apx-*

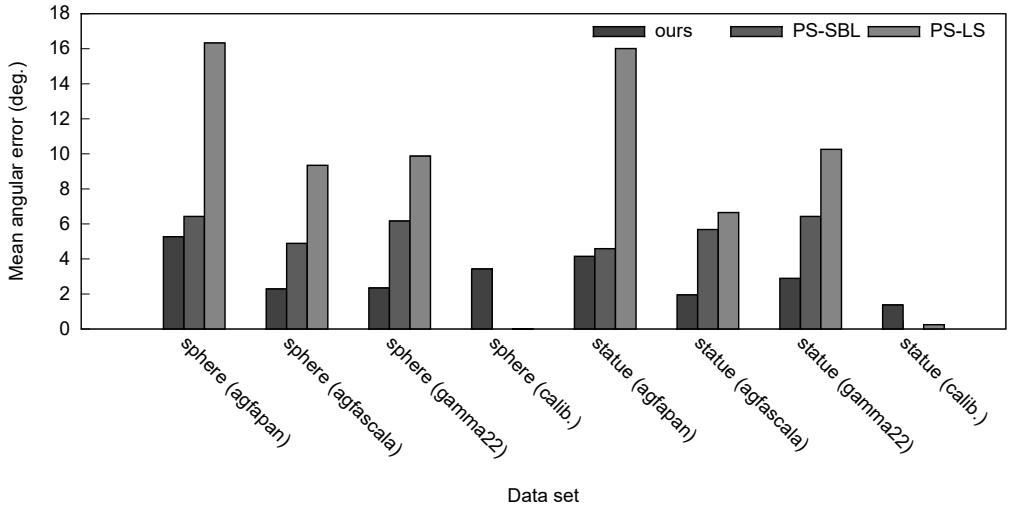


Figure 3.6: Mean angular error of the surfaces estimated from real images with ours, PS-SBL, and PS-LS methods. Note that the error for *sphere (calib.)* is near zero and surfaces estimated from PS-SBL are groud truth.

100CD) in Fig. 3.4. Moreover, noises in the input images also caused the estimated IRF to deviate from the ground truth IRF. The images calibrated with such IRFs were marginally different to the ground truth images so that the surfaces estimated from those images had small angular error over the surfaces.

We also compared the accuracy of the estimated IRFs to the ones estimated from the images of a static scene using radiometric calibration technique[40]. Instead of images with varying light sources, we captured the images of a static scene with different exposure times under the same response functions used for capturing the input images for our method. The IRFs estimated from both [40] and ours are shown in Fig. 3.5. The RMSE of the estimated IRFs shows that our method can estimate IRFs accurately without any additional image for radiometric calibration.

3.5 Conclusion

We presented a method for joint estimating surface normals and a radiometric response function of a camera. Our proposed method takes advantage of the consistency between the irradiance values from an inverse radiometric response function and those estimated from surface normals. This method requires neither cumbersome radiometric calibration preprocessing nor additional images. We demonstrate experimentally that our method can estimate surface normals accurately even when images are captured by using cameras with nonlinear radiometric response functions.

Chapter 4

Non-Lambertian Surface Recovery from Uncalibrated Images

This chapter introduces a RANSAC-based framework to integrate photometric stereo for non-Lambertian surfaces methods into the simultaneous estimation of surfaces and IRF.

In chapter 3, we assume that objects exhibit Lambertian reflection model and simultaneously estimate surface normals and inverse response function. However, there is a very limited number of materials that follow Lambertian reflection model. For example, specular highlights on surfaces of many real-world objects are usually observed as saturated pixel values which do not follow Lambertian reflection. In such kind of surfaces, areas with specular lobe also behave differently to the pixels consisted of diffuse component solely. Therefore, including such non-Lambertian pixel values in the simultaneous estimation leads to warped surfaces and incorrect response functions.

Fortunately, many non-Lambertian surfaces where specular highlights do

not exist behave similarly to Lambertian surfaces. Assuming that specular highlights can be observed only within limited angles, we can treat highlights as outliers that deviate from the Lambertian reflection model. Consequently, it would be possible to integrate a robust estimation technique based on RANSAC [13] into our framework to estimate the surfaces of a non-Lambertian object and the response function of a camera.

The contribution of this work is to propose a RANSAC auto radiometric calibration framework for photometric stereo technique. One can use our framework with a photometric stereo technique to estimate surface normals and inverse response function simultaneously. Our experiments show that our framework can be integrated with photometric stereo for non-Lambertian surfaces to avoid outliers such as specular highlights, noisy pixels, and shadow in the IRF estimation from objects with uniform and non-uniform materials.

The organization of this chapter is as following: Section 4.1 discusses about previous studies in photometric stero on non-Lambertian surfaces. Section 4.2 explains the RANSAC-based framework in detail. Section4.3 describes the experiment detail and evaluation of our framework. Lastly, we concluded our method in Section 4.4.

4.1 Related Work

The problem of estimation of surfaces with specular highlights has been studied by many researchers which can be categorized as the modeling approach and the statistical approach.

The modeling approach estimates surface normals by assuming a specific reflectance property. Georghiades' [14] used Torrance-Sparrow model to solve the photometric stereo problem with unknown light sources. Goldman *et*

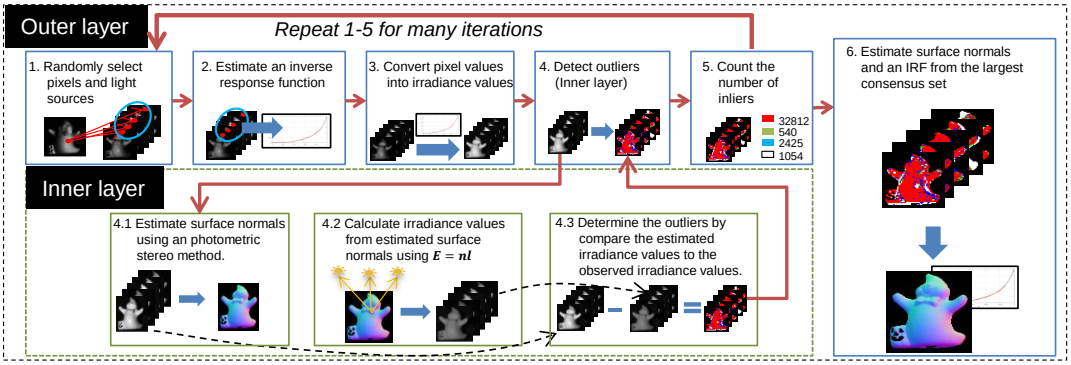


Figure 4.1: Overview of our framework for photometric stereo with auto radiometric calibration. The outer layer is used for IRF estimation and the inner layer is used for surface estimation.

al. [15] used the isotropic Ward reflectance model for their example-based photometric stereo.

The statistical approach regards specular pixels, shadows, as well as other corrupted measurements as outliers to the Lambertian reflectance model, so this approach detects and excludes these corrupted measurements from surface estimation. One of the pioneering works in the statistical approach is the four-sources photometric stereo technique by Coleman and Jain [9]. They took this approach and proposed a photometric stereo technique for non-Lambertian surface using four light sources. They estimated four candidate albedos at a certain location from four possible triplets selected from four light sources. If all of the candidate albedos do not coincide, it is regarded due to specularity, and the smallest albedo is used for surface estimation. Rushmiere *et al.* [54] use similar idea to design a system to estimate surfaces of non-Lambertian surfaces under five light sources. However, they use only the second, third, and fourth brightest pixel intensities out of five to avoid shadows and specular. Barsky and Petrou [4] modified this method to detect both highlights and shadows by using temporal pixel intensity and

linear dependency. Contrast to[54], this technique does not presume that the brightest pixel and the darkest pixels must be specular highlights and shadows. Argyriou *et al.*[3] generalized [4] to handle arbitrary number of light sources. With an assumption that specular highlights at a point are observed at most once, they developed a greedy algorithm to iteratively maximize the number of inliers that match input images. Wu *et al.* [62] show that corrupted measurements increase the rank of observation matrix. So they cast the photometric stereo problem as a rank-minimization problem. Ikehata *et al.* [23] used a hierarchical Bayesian model derived from a sparse Bayesian learning framework instead of principal component analysis used in [62]. Unfortunately, all of these existing photometric stereo techniques for non-Lambertian surfaces assume input images are captured by a camera with a linear response function.

Our method is similar in spirit to one proposed by Mukaigawa *et al.*[44] to use RANSAC-like robust estimation [13] to determine specular region. Unlike the previous work, our proposed method avoids specular highlight in the estimation by using RANSAC framework to determine specular region and estimate response function at the same time. RANSAC has been used by several works to handle specular highlights in face recognition task [47] and to remove specularity from non-Lambertian surfaces in photometric stereo [21, 19].

4.2 Framework for Simultaneous Surface Normals and IRF Estimation

In this section, we present our RANSAC-based framework to integrate photometric stereo for non-Lambertian surfaces methods into the simultaneous

estimation of surfaces and IRF.

Our framework consisted of two layers for surface and IRF estimation as shown in Fig. 4.1. The outer layer of the framework repeatedly estimates IRFs. With the estimated IRFs, the inner layer estimates surface normals using an existing photometric stereo method, and then we determine inliers which are diffuse pixel values that follow the Lambertian reflection model with respect to each estimated IRF. The inlier set with the maximum number of supporting inliers is considered as the consensus diffuse pixel values and we can use these pixel values to estimate the IRF and surfaces without being affected by specularity.

Concretely, we use a RANSAC loop in the outer layer to estimate candidate IRFs using randomly sampled pixels. Using the estimated IRF, we generate candidate calibrated image sets for each candidate IRF. Once the images have been calibrated, we now can use an existing photometric stereo method to estimate surface normals from the calibrated images for each set of candidate IRF. Here, we assume that the photometric stereo method we use can handle corrupted observations such as shadows, and specular highlights properly so the estimated surfaces have no distortion. Then we determine inlier pixels in the images, which are diffuse pixel values that follow the Lambertian reflection model with respect to each estimated inverse response function. As we assume that diffuse pixels in the input be dominant and specular can be observed within a small angle, the inlier set with the maximum number of supporting inliers is considered as the consensus diffuse pixel values and we can use these pixel values to estimate the IRF and surfaces without being affected by specularity.

The detailed algorithm for our framework is explained in the rest of this section as follows:

1) Randomly Selecting Pixels and Light Sources: The first step of our method begins by randomly selecting pixel values for the IRF estimation using RANSAC technique. To avoid corrupted observation, we randomly select fewest pixel values that still can estimate an IRF so we first randomly select one foreground pixel.

For a foreground pixel, there are 3 unknowns because a scaled surface normal has 3 degrees of freedom: two for direction, and one for the scale of the surface normal. In addition, there are K unknowns for K coefficients in inverse response functions if we approximate an IRF with K basis functions. Therefore, we select $(3+K)$ light sources and their corresponding pixel values for the simultaneous estimation.

$$t_{min} = \left\lceil \frac{3+K-1}{s} \right\rceil .t_{min} = (3+K-1)\,. \tag{4.1}$$

That is, we select t_{min} light sources and their corresponding pixel values of the selected pixel of the selected pixel for the simultaneous estimation.

2) Estimating IRF from Pixel Values of Selected Pixels Under Selected Light Sources: After the pixel values have been selected, we estimate an IRF that satisfies the pixel values with (3.7) and (3.8).

Although one might argue that one of the selected pixel values is possibly a specular highlight, we assume that all selected pixel values are diffuse and follow Lambertian reflection model in this step and the IRF is estimated with no distortion from specularity. The goodness of the estimated IRF will be determined in a later step.

3) Converting All Pixel Values Into Irradiance Values: In this step, we convert the pixel values in the input images into irradiance values by using the estimated IRF.

4) Detecting Outliers such as Specular Pixels: In this step, we de-

termine outlier irradiance values that violate Lambertian reflection property.

We first estimate surface normals from the calibrated input images by using a photometric stereo method such as [44], or [24]. Assuming the images are correctly calibrated, the estimated surface normals should have no distortion from both specularity and nonlinear response function. Then, we estimate irradiance values from the estimated surface normals with (3.1) and compare them to the observed irradiance values. An observed irradiance value is a supporting inlier of the estimated surface normal if it is equal to the corresponding estimated irradiance value.

In practice, the estimates might contain small errors due to noise in the input images. We relax the equality constraint by introducing a threshold for the error. An observed irradiance value is considered as an inlier if the difference between the observed and estimated irradiance values is less than this threshold. Let $\tilde{\mathbf{n}}_p$ denote the estimated surface normal scaled by its albedo. Assuming that input images are contaminated with photon shot noise[1], the variance of noise is proportional to the irradiance value [8]. An observed irradiance value $E_{p,d}$ supports the estimated normal $\tilde{\mathbf{n}}_p$ if

$$\left(E_{p,d} - \tilde{\mathbf{n}}_p^\top \mathbf{l}_d\right)^2 \leq \tau_s E_{p,d}, \tag{4.2}$$

where τ_s is a specified parameter.

5) Counting Number of Inliers: We then count the number of inliers to evaluate the goodness of the estimated IRFs. If the IRF is estimated without outlier, it must be consistent with all diffuse irradiance values in the images. In contrast, if the IRF is distorted, it will be consistent with only a few irradiance values. Therefore, the best IRF should maximize the number of supporting diffuse irradiance values.

[1]Poisson noise due to random arrivals of photons

6) Estimating Surface Normals and IRF from Largest Consensus Set: Up to now, we use only a few randomly selected pixel values to estimate an IRF. However, it is possible that one of the selected pixel values is a specular highlight. Therefore, we repeat the whole process multiple times to produce a set of candidate IRFs. Given the number of iterations is sufficiently large, there probably exists an iteration in which all selected pixel values are diffuse. Without distortion from specularity, the IRF should have largest supporting inlier set. The inliers in the largest inlier set are then determined as the consensus-diffuse pixel values.

Finally, we estimate the final IRF from the maximum consensus-diffuse pixel values by using (3.7) and (3.8). Because the pixel values are all diffuse, the final IRF have no distortion from specularity and surface normals can be estimated using a photometric stereo method.

When input images are under-exposed, foreground pixel values in the images probably do not cover the pixel intensity levels. The IRF estimated from (3.7) might become unrealistic, *e.g.*, rising sharply in the range with no foreground pixel values. We can use additional constraints such as smoothness, and integrability in order to avoid such unrealistic IRFs as suggested in previous studies [30, 36]. In this work, we propose to use prior knowledge derived from an existing response function database[17]. Let $P(\mathbf{c})$ a prior model of IRF coefficients constructed by fitting the probabilistic distribution of coefficients of IRF in DoRF to a multivariate Gaussian mixture model:

$$P\left(\mathbf{c}\right) = \sum_{m=1}^{M} \alpha_m \mathcal{N}\left(\mathbf{c}; \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m\right), \tag{4.3}$$

where $\mathcal{N}$ is a Gaussian distribution with mean $\boldsymbol{\mu}_m$, and covariance matrix $\boldsymbol{\Sigma}_m$. α_m is the weighting factor for the m-th distribution, and $\mathbf{c} =$

$(c_1, \cdots, c_k)^T$ are the coefficients of IRF. The model parameters $\boldsymbol{\mu}_m$, $\boldsymbol{\Sigma}_m$ and α_m can be obtained using EM algorithm or cross-entropy method [7]. Then, the prior term can be added to (3.7) as following:

$$\{\hat{\mathbf{n}}_p, \hat{\mathbf{c}}\} = \arg\min_{\{\mathbf{n}_p, \mathbf{c}\}} \{\frac{1}{2PD} \sum_{p=1}^{P} \sum_{d=1}^{D} w\left(I_{p,d} | \mathbf{n}_p, \mathbf{c}\right) - \lambda \log\left(P\left(\mathbf{c}\right)\right)\}, \tag{4.4}$$

where $w\left(I_{p,d} | \mathbf{n}_p, \mathbf{c}\right) = \left(g_0\left(I_{p,d}\right) + \sum_{k=1}^{K} c_k g_k\left(I_{p,d}\right) - \mathbf{n}_p^T \mathbf{l}_d\right)^2$, and λ is a regularization factor for the prior model term. Although an arbitrary number of normal distributions can be used, we assume the number of normal distributions to $M = 5$ similar to the previous studies [30, 36] and use Matlab's *fmincon* for this optimization problem.

We remove saturated foreground pixel values and shadows from the IRF estimation. Those pixel values violate Lambertian reflection model and shadows pixel values might contain excessive noise. We determine all pixel values with $I_{pd} = 1$ as saturated and detect shadow pixels by thresholding.

4.3 Experiments on Specular Objects

4.3.1 Experiments on Synthetic Images

We conducted experiments on both synthetic and real images to validate the estimation accuracy of our framework on non-Lambertian surfaces.

In our experiments, we used 20 synthesized images of a sphere under the assumptions of Torrance-Sparrow model [57] and directional light sources with the same intensities. The light directions were uniformly distributed over a hemisphere. The images were applied with response functions from DoRF to obtain uncalibrated images.

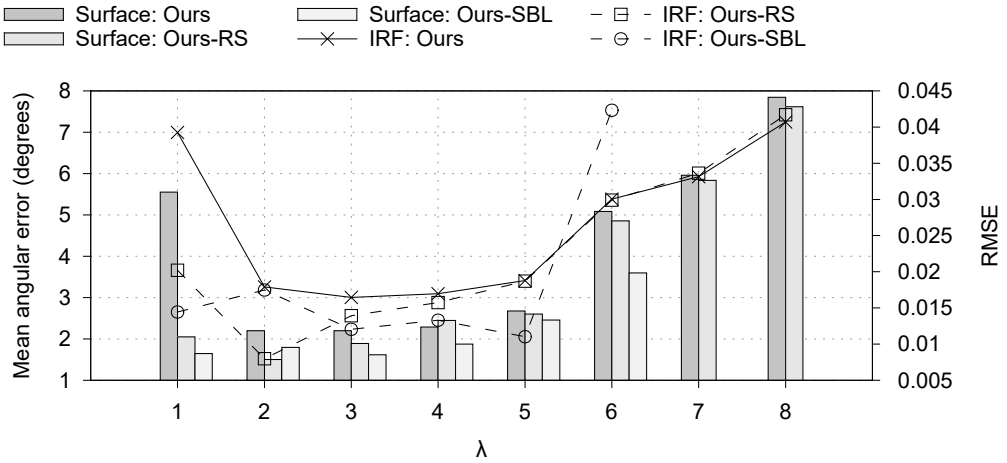


Figure 4.2: Experimental results with respect to the value of λ.

We implemented two variants of the inner layer of our framework with two photometric stereo methods: a straightforward RANSAC-based method (ours-RS), and a sparse Bayesian learning-based method (ours-SBL) based on [24]. Unlike [44], we directly applied RANSAC to Lambertian photometric stereo to estimate surface normals for each foreground pixel. More specifically, we select three light directions and their corresponding pixel values to estimate surface normals using (3.4). Then the supporting inliers are determined as the observed irradiance values that match ones predicted with the estimated surface normals. This process is repeated for many iterations and we regard the support inliers of the iteration that yields a maximum number of inliers as consensus inliers. Then the final surface normals are estimated using only consensus inliers.

We use the following equation to determine the number of iterations for RANSAC processes[13],

$$\text{number of iterations} = \frac{\log 1 - p}{\log 1 - w^n}, \tag{4.5}$$

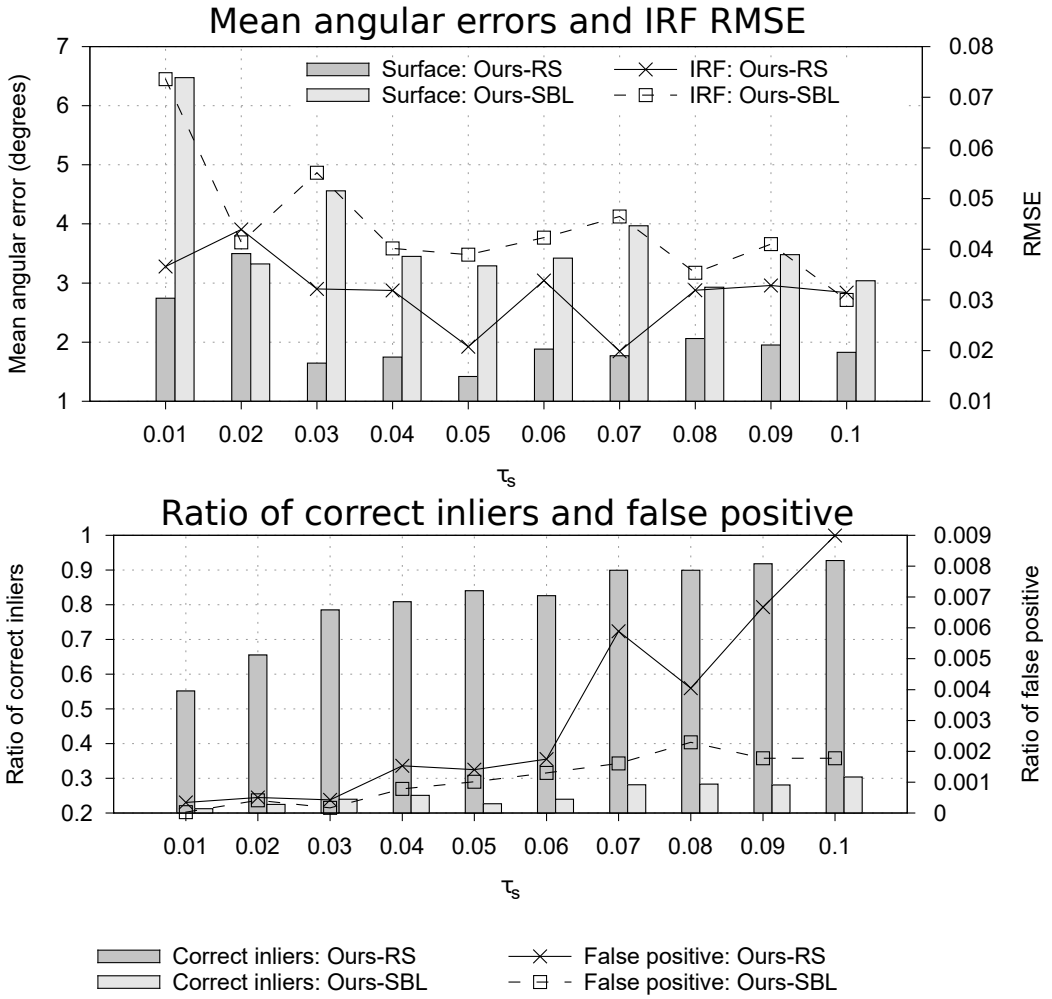


Figure 4.3: Experimental results with respect to the value of τ_s.

where p is the probability that there is an iteration that the RANSAC algorithm selects only inliers, w is the ratio of inliers and all samples, and n is a number of samples. Since we assume that specular highlights can be seen in limited angles, we aggressively set $p = 0.95$ and $w = 0.8$. With these settings, it took about 53 seconds for Ours-RS and 414 seconds for Ours-SBL to estimate an IRF and surface normals from a set of 20 images with 3,228 foreground pixels.

To calculate the RMSE of the IRFs, we discarded the top ten percent of

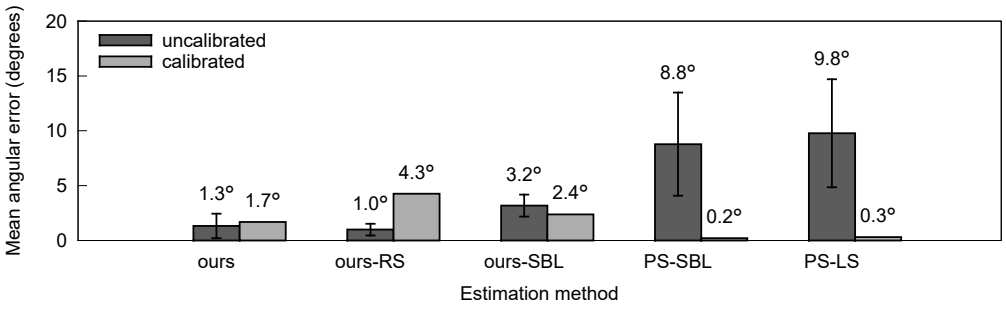


Figure 4.4: Comparison of mean angular error in degrees of the surfaces estimated from synthetic images using the specified methods. The error bar shows SD of the errors.

the brightest selected pixel values and used only the rest. This is because the number of bright diffuse pixel values was small due to specularity. Therefore, the estimated IRF where the pixel value was near 1 could not be constrained well; thus, not accurate.

First, we observed the effect of λ in the prior term to the estimated surfaces and IRF. Here we picked a small positive value to factor the prior term to match the magnitude of the surfaces estimation term in (4.4). More specifically, we used various λ to perform surfaces and IRFs estimation from the synthetic images generated with a selected set of response functions, then we selected the most appropriate value based on the estimation results.

Fig.4.2 shows that the value of $\lambda = 10^{-5}$ gave the best balance of the IRF and surface estimation accuracy. We found that the IRFs is overfitted to the prior term when $\lambda >= 10^{-2}$. As the pixel values no longer follow the Lambertian model when the IRFs is overfitted, ours-SBL misclassified all foreground pixel values, *i.e.*, no result when $\lambda >= 10^{-2}$. In contrast, the sampling process in ours-RS guarantees that there are at least four pixel values have been determined as inliers. Therefore, the surfaces could be estimated even when the IRFs were overfitted.

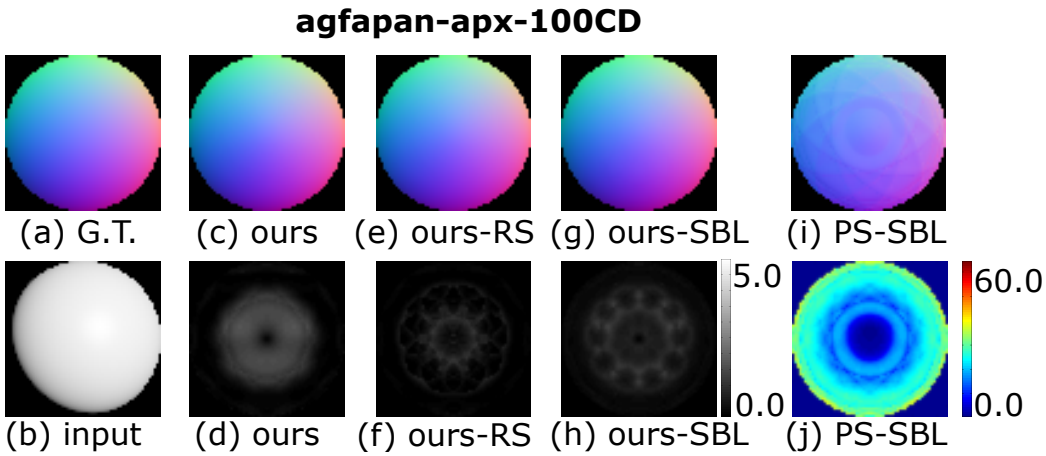


Figure 4.5: Qualitative results from the synthetic specular sphere. (a) ground truth, (b) example of an input image from a synthetic images set, (c)-(j) estimated surfaces estimated using the specified methods. The bottom row shows angular differences to the ground truth in degrees.

The effect of the threshold τ_s on the surface and IRF estimation performance is shown in Fig.4.3. A larger τ_s leads to more false positive, *i.e.*, more outliers are misclassified as inliers. In contrast, a smaller τ_s causes less pixel values to be classified as inliers. Here we varied the value of τ_s and performed experiments on the synthetic datasets.

We found that $\tau_s = 0.05$ gave the best balance between the classification accuracy, angular error, and RMSE of the estimated IRFS. Therefore, we fixed the threshold to $\tau_s = 0.05$ for the rest of the experiments. With the given λ and τ_s, the average RMSE of the IRFs estimated from ours, ours-RS, and ours-SBL for all synthetic data sets were 0.0128, 0.0156, and 0.0208 respectively.

We observed that ours-SBL produced surfaces and IRFs with larger errors than the ones estimated by ours without outlier detection. This was due to scaling ambiguity that distorted the overall shape of the estimated IRF. When the outlier detection removes bright specular highlights, the range of

pixel values is narrower so there is less information to constrain the IRF that are outside of the range of inliers values. Eventually, our method further minimized (4.4) by minimizing the scale of IRF while still satisfying both input pixel values and the boundary conditions. Without enough pixel intensities to constrain the whole range of IRF, the IRF near the end of the range of inliers values are therefore incorrectly estimated. Since PS-SBL uses diffuse components in the pixel with specular highlight to achieve better results, Ours-SBL took the disrupted bright pixel values into account and it led to more angular errors. In contrast, Ours-RS was more capable to handle this kind of distortions since it aggressively discarded outliers such as the incorrect calibrated pixel values from the surface estimation, therefore specular highlights were removed more properly.

We compared the surface normals estimated using our proposed framework to those estimated from PS-SBL, and the simultaneous estimation without outlier detection (ours). Figure 4.5 shows an example of the estimated surfaces. The bright areas shown in the error maps correspond to the specular highlight areas in the input images. One can see that ours-RS can remove the specular pixel values from the surface estimation so it reduces distortion caused by specular highlights on the estimated surfaces. We summarized the mean angular errors of the surfaces estimated from all datasets and their standard deviation in Fig. 4.4. The mean angular error of the surfaces estimated from ours, ours-RS, and ours-SBL are 1.32°, 0.99°, and 3.19° respectively. We performed statistical significant with paired samples t-test and found that ours-RS can produce surfaces lower mean angular errors at significance level of 0.05.

4.3.2 Experiments on Real Images

We conducted experiments on images of four real objects made from four different materials; matte ceramic with painted smooth areas *seal*, polished ceramic *ghost*, opaque plastic *tomato*, and glossy-painted ceramic *fish*. The images were captured under the same lighting environment and nonlinear response functions used in the experiments with diffuse objects, *i.e.*, agfapan-apx-100CD, agfa-scala-200xCDStandard, and gamma+2.2. The ground truth was obtained as the surface normals estimated from the radiometrically calibrated images by using PS-SBL [24] with a number of piecewise functions $p = 3$ as the ground truth.

The qualitative results from the *ghost* data set are shown in Fig. 4.6. The error maps indicate that the methods with auto radiometric calibration have less angular errors. This is because the nonlinearity has been compensated. And the angular errors from specular highlights in the results estimated from ours-SBL and ours-RS are reduced. However, we still observed large angular error along the concave regions due to inter-reflection and cast shadows on the surfaces.

Also, our framework still has difficulty in recovering the regions with reflective materials as observed in the results from the *fish* data set. The edges of the scale on the object is a non-convex shape and it is made of reflective glossy paint. Therefore, the pixel values in those areas are greatly influenced by specular reflection component, hence exhibit a strong effect of interreflection. Moreover, we observed the distortion in the bottom left part of the surfaces estimated from *fish (gamma 2.2)* by Ours-SBL. We performed statistical significant with paired sample t-test and found that ours-RS did not have lower mean angular error than ours at significance level of 0.05 with the specified RANSAC parameters of $p = 0.95$ and $w = 0.8$ but it could

produce surfaces with lower mean angular error than ours at significance level of 0.05 when we used more conservative parameters of $p = 0.95$ and $w = 0.7$.

We show the quantitative results of the experiments in Fig. 4.8. One can see that our RANSAC-based method gives less distortion from specularity so the angular error decreased in most datasets. However, we observed that there are small errors in the estimated surfaces caused by overly removed pixel intensities that exceed the threshold in (4.2). These small errors could be observed over the estimated surfaces: therefore, mean angular error remained relatively high, even though specular highlights were removed.

4.4 Conclusion

We proposed a framework for photometric stereo techniques to recover surface normals from images captured using a camera with an unknown non-linear response function. Our framework can be integrated with an existing photometric stereo method to handle outliers such as specular pixels. The experiments show that our method can estimate surface normals of the non-Lambertian surfaces more accurate than the existing methods when images are radiometrically uncalibrated.

There are two limitations with our proposed method. First, it can estimate response functions and surface reflectance up to scale. Since our method has no constraint regarding neither scale of albedos nor scale of response function, the scale of albedo can propagate to the scale of response function and vice versa. Therefore, it cannot determine the correct scale of estimated IRF and albedos without additional cues. Second, our shadow thresholding works incorrectly if noisy pixels are greatly modified by the response function. From

the experimental results, noisy pixels in shadow areas were greatly amplified by the agfapan-apx-100CD response function, so the pixel values exceed the shadow threshold and became non-shadow in pixel-value space. Those non-shadow pixels were eventually included in the simultaneous estimation and caused distortion in such areas. Similarly, non-shadow pixels were darkened by gamma+2.2 function so they become shadow pixels.

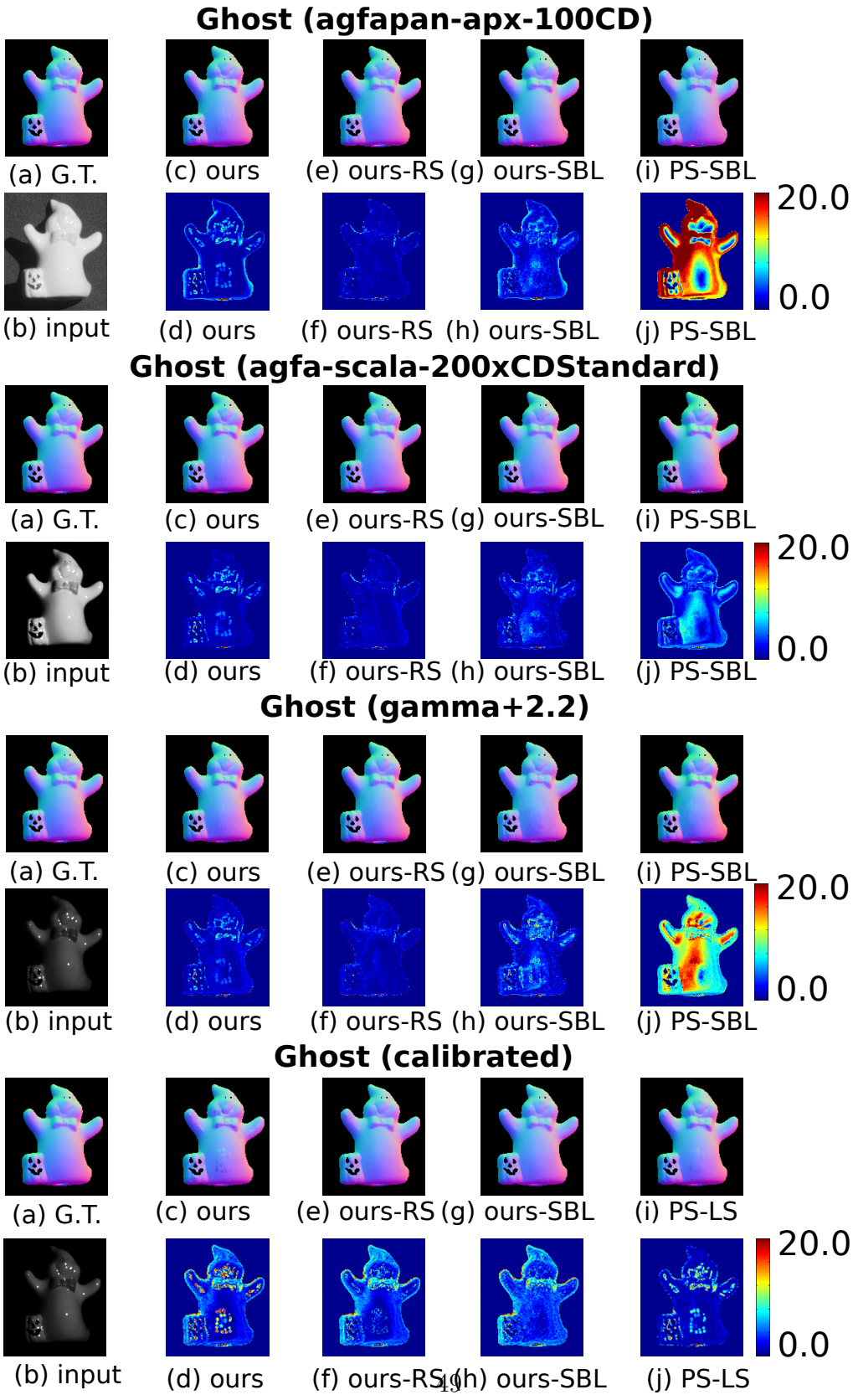


Figure 4.6: Qualitative results from the data set *ghost*. The response functions used for capturing the images are specified in parenthesis. (a) ground truth (b) sample of input images (c)-(j) the top row shows the normal maps estimated from the specified method and the bottom row shows the error maps.

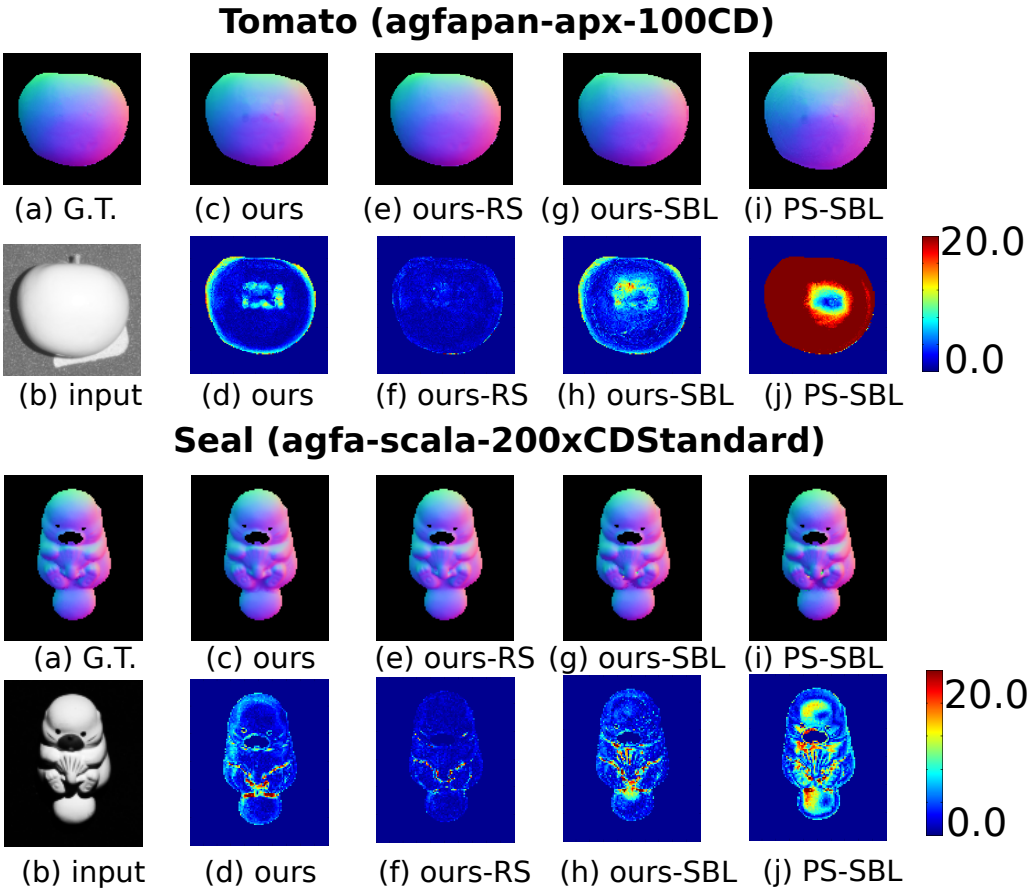


Figure 4.7: Samples of qualitative results from the data set *seal* and *tomato*. The response functions used for capturing the images are specified in parenthesis. (a) ground truth (b) sample of input images (c)-(j) the top row shows the normal maps estimated from the specified method and the bottom row shows the error maps.

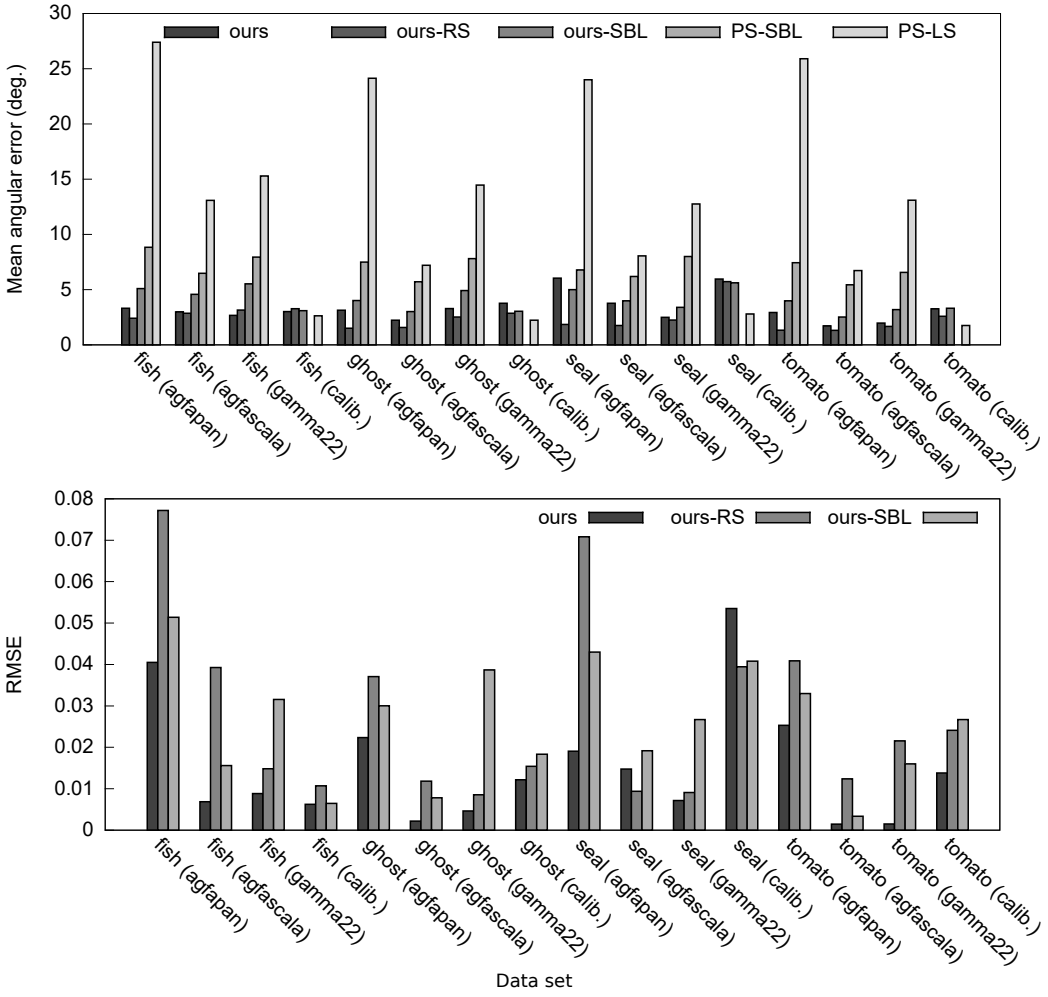


Figure 4.8: The quantitative results from the experiments with real images. (top) Mean angular error in degrees of surface normals estimated with ours, ours-RS, ours-SBL, PS-SBL, and PS-LS. (bottom) RMSE of agfapan-apx-100CD (agfapan), agfa-scala-200xCDStandard (agfascala), and gamma+2.2 response functions estimated from fish, ghost, seal, and tomato images sets. The RMSE of the IRF estimated by [40] are 0.0034, 0.0032, 0.0117, and 0.0066 respectively.

Chapter 5

Auto Radiometric Calibration in Uncalibrated Photometric Stereo

5.1 Background

This chapter presents a method to perform auto-radiometric calibration to compensate nonlinearity of the response function that resides in input images of uncalibrated photometric stereo. In this context, directions of the light sources used for capturing the scene are unknown.

The classic photometric stereo method[60] assumes that the directions and intensities of the varying light sources are known or calibrated beforehand. It then uses the relationship between the varying light sources and the changes of irradiances according to Lambertian reflection property to estimate the object's surface orientations. This condition is hard to achieve in an uncontrolled environment where light sources are generally unknown. Therefore this limits the practical use of the photometric stereo.

There are a lot of studies try to relax this assumption by developing photometric stereo methods which do not require the light sources to be known in advance. They extend the photometric stereo to estimate surface normals from images that are captured under complex lightings such as unknown directional light sources[18, 6, 55, 37, 33], unknown point light sources[39, 64, 48, 29], or even under natural illumination[5, 22, 20]. In the context of unknown directional light sources, the photometric stereo problem is much more difficult than the one with known light source direction especially when target surfaces are assumed to follow Lambertian reflection property as it suffers from an ambiguity in a class of transformation call Generalized Bas-Relief transformation (GBR)[6]. Previous studies make use of additional clues to resolve the GBR ambiguity such as material isotopy[56], inter-reflection[45], low-dimension feature space[5], prior knowledge about albedo[2], diffuse maxima[48], variation of depth map[52], pixel color profiles[55], etc. However, most of those methods take it for granted that the input images must be radiometrically calibrated.

In this chapter, we propose a method to perform radiometric calibration in photometric stereo input images which light source directions and intensities are not known. Our method makes use of the rank-3 property of a matrix constructed by organizing the input images to exhibit linear dependency. Nonlinear response functions do not preserve this rank property and turn this image matrix into a full rank one. Therefore, the inverse response function then can be determined as a function that turns the full rank image matrix back into the one with rank-3 structure.

The main contribution of this study is an auto radiometric calibration method for uncalibrated photometric stereo. With this method, ones can estimate surfaces using uncalibrated photometric stereo methods even when

light sources directions and intensities are unknown.

The rest of this chapter is organized as follows: Section 5.2 give a brief explanation regarding previous studies on uncalibrated photometric stereo and radiometric calibration. Section 5.3 explains about image formation and rank-3 structure in the image matrix. Section 5.4 explains our proposed method. Section 5.5 shows experimental results. Lastly, Section 5.6 concludes this chapter.

5.2 Related Work

This section introduces previous works on uncalibrated photometric stereo methods and discusses the relationship to our proposed method.

Conventional photometric stereo[60] assumes a motionless Lambertian surface placed under varying directional light sources with known directions and intensities. Photometric stereo that does not assume known light source directions was first pioneered by Hayakawa[18]. He proposed a method to estimate surfaces from images without using neither light source directions nor intensities. However, there exists an ambiguity in the estimated surfaces which can be resolved if there are at least 6 known surface normals in the images or 6 known light source directions. Belhumeur *et al.*[6] used integrability to reduce the ambiguity to a class of transformation, namely, Generalized Bas-Relief transformations (GBR). More concretely, given a pair of light source directions and Lambertian surfaces, we can find a countless number of transformations that result in the same images when viewing orthographically regardless the transformed surfaces and light source directions. This study suggested that we can only estimate surface normals up to a general bas-relief transformation without additional clues.

The trend of research in uncalibrated photometric stereo then was refocused on disambiguating GBR ambiguity. There are numerous studies proposed variety of cues that can be used to disambiguate GBR transformations including specularities[12, 14], symmetry in isotropy for non-Lambertian materials[34], attached shadow coding[46], and inter-reflections[45]. In context of Lambertian surfaces, many researches suggested that albedo distribution[2], local diffuse maxima[48], image structure[44, 43, 5], grouping of color and appearance [55] are feasible for disambiguating the GBR transformation. Almost all of these methods except [55] take it for granted that the input images are already radiometrically calibrated.

Although radiometric calibration is a necessary step before photometric scene analysis can be carried out, the research on radiometric calibration itself is usually done separately. Primitively, one can estimate an inverse response function by using a calibration chart with known reflectance. The difference of the known reflectance and the recorded pixel intensities can be used for estimating the inverse response function. To relax the requirement of using the calibration target, the pioneering works suggested that multiple photos of a static scene taken under varying exposure times provide enough information for radiometric calibration [35, 11] along with an empirical [17] and parametric model[40] that can approximate inverse response function accurately. However, those methods still need images purposely captured for radiometric calibration. Later studies aimed toward relaxing this restriction by using additional clues that lie inside original images or video sequences such as the distribution of noises[36], distribution of pixel intensities at near edge [30, 31], vignette [26], low-rank structure [27, 41, 28], reflectance properties of skin pigments [28], and temporal changes of irradiance [59, 25].

In the context of surface estimation, Abrams *et al.*[1] incorporate non-

linear response function into his method for estimate surface normals of an outdoor scene from time-lapse images taken by a web camera. Despite lacking light direction, they look up the sun direction using information from GPS and timestamp. The most well-known method for performing radiometric calibration in the context of uncalibrated photometric stereo was proposed by Shi *et al.*[55]. They illustrate that color profile in a pixel across input images should be linear, but it becomes a nonlinear curve under nonlinear response function. Therefore, they can formulate an optimization problem to determine an inverse response function as a function that can bend the curves back to a straight line. However, this method assumes color images and fails to estimate inverse response functions from gray objects which the color profiles remain straight lines regardless nonlinear response function. Unlike these methods, our method assumes only directional light sources without information regarding their light source direction and intensities. At its core, our method makes use of linear dependency similar to that used by Lee et al. [27]. However, we do not assume color input images and ours is able to estimate the inverse response function from gray images, hence, ours can estimate inverse response function for each image channel separately.

5.3 Rank-3 Structure in Radiometrically Calibrated Images

In this section, we explain photometric stereo input image representation in this work and its rank-3 structure in which we can use for the inverse response function estimation.

Let $\mathbf{e}_d$ denotes the d-th image as a column vector, we can stack all of D

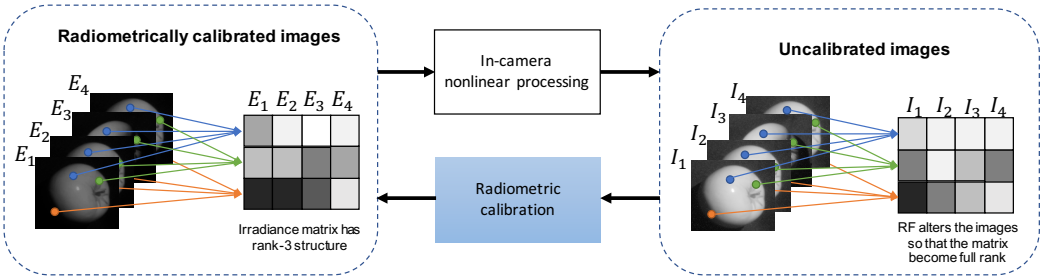


Figure 5.1: Irradiance matrix of a Lambertian surface has rank-3 structure. In-camera nonlinear response functions convert irradiance into a full-rank pixel intensity matrix. Then, the inverse response functions can be determined as the one that converts the full-rank intensity matrix back into the matrix with the rank of 3.

images together to form an irradiance matrix.

$$\begin{aligned} \mathbf{E} &= [\mathbf{e}_1, \ldots, \mathbf{e}_D], \\ &= \begin{bmatrix} E_{1,1} & E_{1,2} & \cdots & E_{1,D} \\ E_{2,1} & E_{2,2} & \cdots & E_{2,D} \\ \vdots & \vdots & \ddots & \vdots \\ E_{P,1} & E_{P,2} & \cdots & E_{P,D} \end{bmatrix} \end{aligned} \tag{5.1}$$

Let $\boldsymbol{l}_d = [l_x, l_y, l_z]^T$ denotes a d-th light source vector with a scaled light intensity, and $\boldsymbol{n}_p = [n_x, n_y, n_z]$ is a surface normal scaled with an albedo at the point p of a Lambertian object. Irradiance $E_{p,d}$ of the point p illuminated under $\boldsymbol{l}_d$ then can be calculated as a dot product of the light source and the surface normal,

$$E_{p,d} = \boldsymbol{n}_p \cdot \boldsymbol{l}_d. \tag{5.2}$$

Let $\boldsymbol{N}$ is a surface normal matrix where each row is a surface normal, i.e.,$\boldsymbol{N} = [\boldsymbol{n}_1, \boldsymbol{n}_1, \ldots, \boldsymbol{n}_P]^T$, and $\boldsymbol{L}$ is a light source matrix where each column

represent the light source direction. We can rewrite the image matrix $\mathbf{E}$ as:

$$\begin{aligned}\mathbf{E} &= \begin{bmatrix} \boldsymbol{n}_1 \cdot \boldsymbol{l}_1 & \boldsymbol{n}_1 \cdot \boldsymbol{l}_2 & \cdots & \boldsymbol{n}_1 \cdot \boldsymbol{l}_D \\ \boldsymbol{n}_2 \cdot \boldsymbol{l}_1 & \boldsymbol{n}_2 \cdot \boldsymbol{l}_2 & \cdots & \boldsymbol{n}_2 \cdot \boldsymbol{l}_D \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{n}_P \cdot \boldsymbol{l}_1 & \boldsymbol{n}_P \cdot \boldsymbol{l}_2 & \cdots & \boldsymbol{n}_P \cdot \boldsymbol{l}_D \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{n}_1 \\ \boldsymbol{n}_2 \\ \vdots \\ \boldsymbol{n}_P \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{l}_1, \boldsymbol{l}_2, \ldots, \boldsymbol{l}_D \end{bmatrix} \\ &= \mathbf{N} \cdot \mathbf{L}. \end{aligned} \tag{5.3}$$

Eq.(5.3) suggests that the matrix $\boldsymbol{E}$ has the rank of 3.

Under linear response function, a pixel intensity $I_{p,d}$ of the point p on d-th image are linearly proportional to the irradiance received by the image sensor so we can easily substitute pixel intensities with irradiances, *i.e.*, $I_{p,d} = E_{p,d}$. The rank of the result image matrix then remains to be the rank of 3. However, the effect of nonlinear response functions is nonlinearly proportional to input irradiance of so the pixel intensity matrix is no longer have linear dependency and its rank becomes full rank. Note that we normalize the irradiances and the intensities into the range of $[0, 1]$ without loss of generality.

5.4 Proposed Method

In this section, we proposed a method to estimate an inverse response function from the pixel intensity matrix constructed from photometric stereo

images without using light source information.

5.4.1 Inverse Response Function Estimation with Ratio of Singular Value Minimization

We first begin by explaining the image structure under nonlinear response functions. Let $\mathbf{I}_d$ is a column vector that represents pixel intensities of the d-th image captured under dth light source. We can construct an image matrix $\mathbf{I}$ by stacking image column vectors $[\mathbf{I}_1, \ldots, \mathbf{I}_D]$. Here we take for grant that the images are aligned so that the pixels from the same coordinate are on the same row.

The image matrix $\mathbf{I}$ is equivalent to that converted from the irradiance matrix with a response function, *i.e.*, $\mathbf{I} = f(\mathbf{E})$ where f is a monotonic increasing function that operates on the matrix $\mathbf{E}$ element-wise.

The nonlinearity of response function alters all elements in $\mathbf{I}$ so that the linear dependency property does not hold and $\mathbf{I}$ becomes full-rank. Since f is a monotonic function, there exists an inverse response function g such that $E = g(I) = f^{-1}(I)$. So we propose a radiometric calibration method that estimates an inverse response function g as a function that turns the full-rank $\mathbf{I}$ into a rank-3 matrix.

We then formulate the calibration method by using a ratio of singular values. Let $\mathbf{s}^{\boldsymbol{I}} = [s_1, s_2, \ldots]$ denotes singular values of the matrix $\mathbf{I}$. The rank-3 property suggests that the three largest singular values s_1, s_2, and s_3 should be non-zero while the rest of singular values are near zero. However, the effect of nonlinear response function causes variation in pixel intensities so the singular values become nonzero everywhere, hence, full-rank structure. Therefore we can cast the radiometric calibration problem to a singular value minimization problem to determine the inverse response function $\hat{g}$ that min-

imizes the rest of singular values.

$$\begin{aligned} \hat{g} &= \arg\min_{g} \sum_{i=4}^{D} s_i^A, \\ &\text{subjects to } \mathbf{A} = g(\mathbf{I}), \frac{\partial g(x)}{\partial x} > 0 \end{aligned} \tag{5.4}$$

where x is a dense sampling of pixel intensities in that range of $[0, 1]$, *e.g.*, $x \in [0, 0.001, 0.002, \ldots, 1]$. Note that we normalized irradiance and pixel intensities into the range of $[0, 1]$ without losing its generality so we can enforce boundary constraints $g(0) = 0$ and $g(1) = 1$.

This optimization setting is closely similar to rank minimization problem which can be efficiently approximated by a nuclear norm minimization[61]. However, in our case, both nuclear norm minimization and directly minimization of the $\sum_{i=4}^{D} s_i^A$ are more likely to produce a degenerate solution which causes all singular values to be zero and turns the image matrix into a rank-1 matrix. We avoid this issue by adding a rank constraint to (5.4) to enforce rank-3 structure. Intuitively, we want to prevent the third singular value from becoming zero while minimizing the summation of the singular values, therefore we make a modification by adding s_3^A as:

$$\begin{aligned} \hat{g} &= \arg\min_{g} \sum_{i=4}^{D} \frac{s_i^A}{s_3^A}, \\ &\text{subjects to } \mathbf{A} = g(\mathbf{I}), \frac{\partial g(x)}{\partial x} > 0. \end{aligned} \tag{5.5}$$

With this method we can estimate the inverse response function from 4 input images of uncalibrated photometric stereo images. Although the cost function in (5.4) can estimate an inverse response function accurately from noise-free images or when there is a small amount of noise, we found that

the IRF estimation performance degrades rapidly when noise increases.

We took the same approach as suggested by Lee *et al.*[27] to handle the noise. Since the noise variates all singular values evenly, the smaller singular values are therefore more severely affected by the noise. Instead, we can conveniently use only the fourth singular in the IRF estimation to avoid the variation from noise. The optimization problem then can be given by:

$$\begin{aligned} \hat{g} &= \arg\min_{g} \frac{s_4^A}{s_3^A}, \\ &\text{subjects to } \mathbf{A} = g(\mathbf{I}), \frac{\partial g(x)}{\partial x} > 0. \end{aligned} \tag{5.6}$$

Since all of the pixels share the same response functions, and given that number of foreground pixels are larger than a number of unknowns, we can speed up the estimation by using a subset of the images for the inverse response function estimation instead of the whole images.

5.4.2 Response Function Representations

Thus far, we have not mentioned about the response function representation of our method. Our method estimates inverse response functions that are approximated by a linear combination of either a parametric model such as polynomials[40] or an empirical model[17]. The polynomial representation is one of the most widely used representations for radiometric calibration researches[41, 27, 55]. It approximates a response function with K degree polynomials as:

$$g(I) = I + \sum_{k=1}^{K} c_k I^k, \tag{5.7}$$

subjects to the boundary conditions, $g(0) = 0$ and $g(1) = 1$, and monotonicity constraint $\frac{\partial g(x)}{\partial x} > 0$. This representation has been widely used in

recent prior works[55, 27] because it can well approximate most of the typical response functions with only fifth-degree polynomials.

The empirical model of response function (EMoR) is a model derived from real-world response functions and it has been used in various previous works[30, 36, 32, 28]. The K-parameters EMoR approximation of an inverse response function is in the form:

$$g(I) = g_0(I) + \sum_{k=1}^{K} c_k g_k(I), \tag{5.8}$$

subjects to the boundary conditions and monotonicity constraints. Here, g_0 is the mean curve or the 0-th basis function, g_k is the k-th basis functions derived from the response functions in the database with Principle Component Analysis (PCA), and c_k is the coefficient of the k-th basis function.

By substituting (5.7) or (5.8]) into (5.6), the optimization problem for response function then becomes the problem of determining the coefficients of the basis functions, *e.g.*, for EMoR representation,

$$\begin{aligned} \{\hat{c}_k\} &= \arg\min_{\{c_k\}} \frac{s_4^A}{s_3^A}, \\ &\text{subjects to } \mathbf{A} = g_0(\mathbf{I}) + \sum_{k=1}^{K} c_k g_k(\mathbf{I}), \frac{\partial g(x)}{\partial x} > 0. \end{aligned} \tag{5.9}$$

Despite reported in a prior study[27] that the EMoR representation of response function is unsuitable for gradient-based convex optimization, we found that our implementation of the EMoR model works well in our experiments.

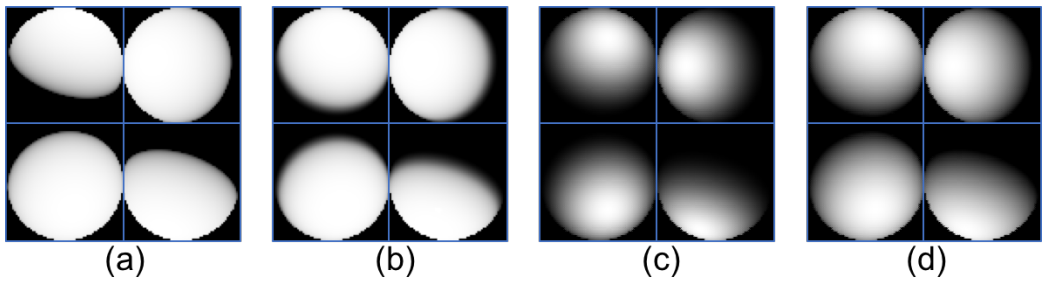


Figure 5.2: Samples of input images compared to the radiometric calibrated images. (a) agfapan-apx-100CD, (b) agfa-scala-200xCDStandard (agfascala), (c)gamma+2.2 (d) radiometric calibrated images.

5.5 Experiments

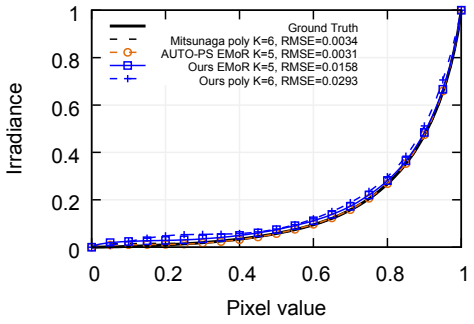
We evaluate the estimation accuracy of our proposed method on both synthetic and real images.

5.5.1 Experiments on Synthetic Images

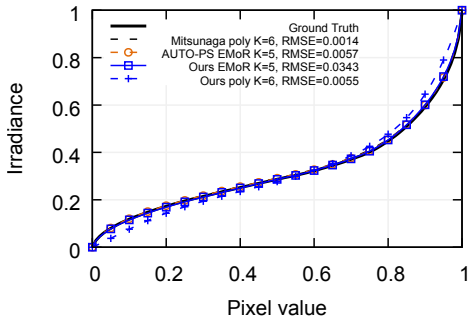
We performed the experiments to observe the inverse response function estimation accuracy with synthetic images. We generated the synthetic images of a Lambertian sphere, illuminated under 20 light source with a uniform intensity. The light source was distributed so that the light source directions were uniformly over a hemisphere in which the sphere was placed at the hemisphere's center. We manually masked out the background and used only 3228 foreground pixels in the estimation.

To obtain uncalibrated images, we applied nonlinear response functions picked from DoRF[16] database to the calibrated images. The pixel intensities that have intensity values less than $^{5}/_{255}$ were considered shadows and were excluded from the inverse response function estimation.

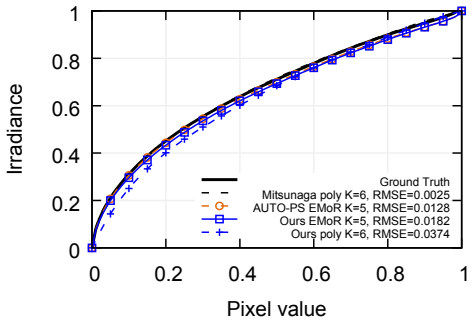
We implemented two variants of our method with polynomial representation (Ours-poly) and EMoR representation (Ours-EMoR). Both variants



(a) agfapan-apx-100CD



(b) agfa-scala-200xCDStandard



(c) gamma+2.2

Figure 5.3: Samples of the inverse response function estimated using our proposed method compared to the ones estimated from previous methods. The RMSE of the estimated functions are denoted in the legend. (left) agfapan-apx-100CD (middle) agfa-scala-200xCDStandard (right) gamma+2.2

were implemented by using Matlab's *fmincon*. They took about 5 seconds for the EMoR variant and 1.3 seconds for the polynomial variant on an Intel Core i5 2.7GHz machine to estimate an inverse response function.

We estimated the IRF using images prepared with response functions the EMoR database. The number of basis functions for EMoR is $K = 5$ and we use 6-th degree polynomials ($K = 6$) for the polynomial variant. In this experiment, we empirically used the number of random pixels at $P = 400$. To demonstrate the estimation accuracy, we compared the root mean square error (RMSE) of the estimated IRF to those of estimated by auto radiometric calibration photometric stereo (Auto-PS) with $K = 5$ in Figure

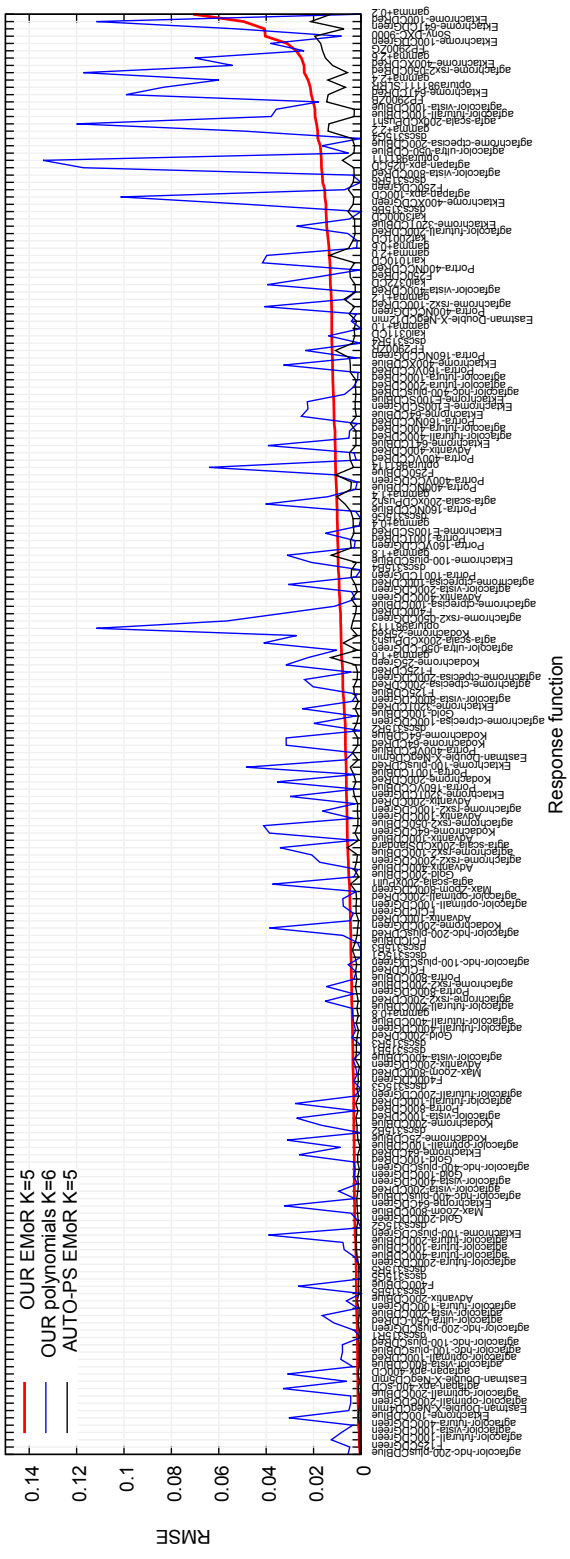


Figure 5.4: RMSE of the estimated inverse response functions using our proposed method (OUR) with polynomials and EMoR representation compared to those estimated of a previous method (AUTO-PS)[42]. Note that AUTO-PS used light source directions for the estimation.

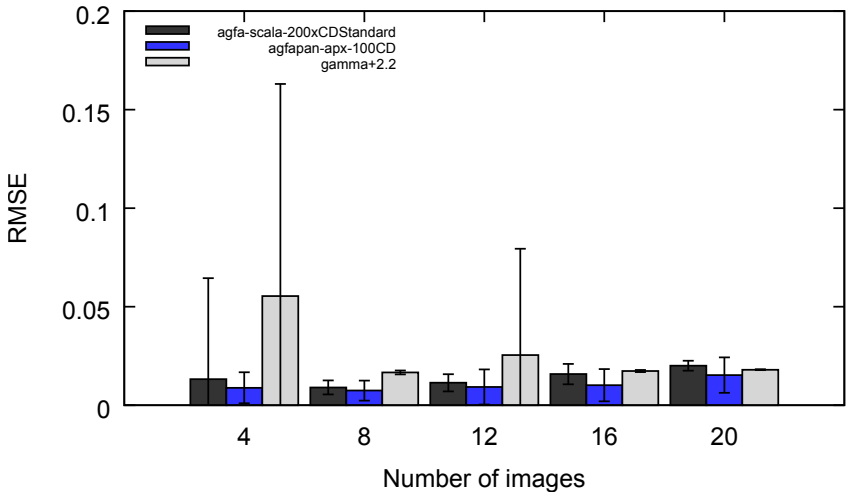


Figure 5.5: Average RMSE of the estimated inverse response functions using various number of images. The error bars indicate standard deviation of the RMSE.

5.4. The mean and standard deviation of RMSE of the inverse response function estimated by Ours-EMor, Ours-poly, and Auto-PS were 0.0088 (Sd. 0.0086), 0.0182 (Sd. 0.0256), and 0.0035 (Sd. 0.0038) respectively. Note that this was an indirect comparison because Auto-PS requires known light source directions.

The experimental results suggested that Ours-EMoR can estimate the inverse response function relatively well for most of the response functions in the database except for a few ones that cannot be represented well with EMoR representation. The number of the estimated inverse response functions that have the RMSE less than 0.02 was 185 functions. For Ours-poly, it can accurately estimate smaller subsets of the response functions and there were 135 inverse response functions with the RMSE less than 0.02. We show examples of the estimated functions using our proposed method compared with the previous methods in Figure 5.3.

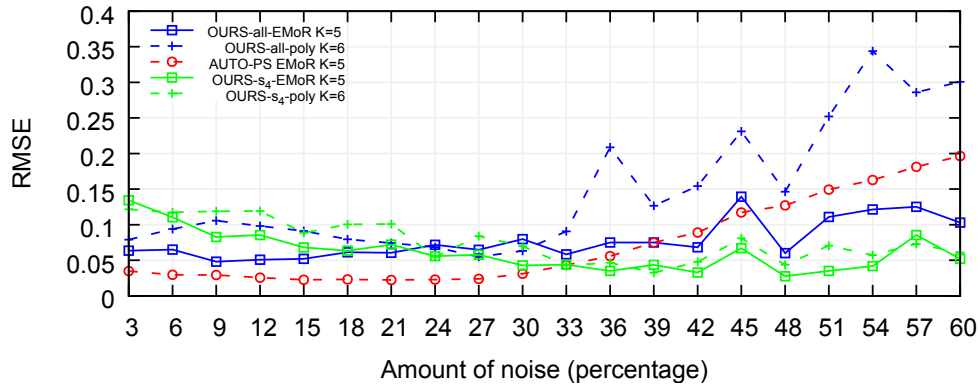
We then observed the effect of the number of images on the estimated

function. In this experiments, we prepared images with three response functions, namely, agfapan-apx-100CD, agfa-scala-200xCDStandard, gamma+2.2. Instead of using the whole set of input images, we selected only a small subset of images for the inverse response function estimation. We repeated this for several times with various numbers of images. We show the RMSE of the estimated images with a different number of images in the Figure 5.5.

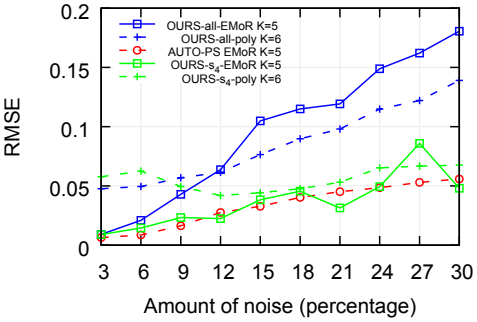
If a subset of the selected images were relatively similar to each other, or the light sources used for capturing those images lie close to each other in the same plane, the images exhibited linear dependency in the image matrix. When the number of images is small, our assumption that the irradiance matrix should have the rank of three might no longer hold. Therefore, in this case, our algorithm produced degenerate solutions which caused a very large RMSE. On the other hand, when the number of images was sufficiently large, the probability that the selected images can cause linear dependency so that is lower, our algorithm could produce appropriate solutions even though there might have some similar images selected in the estimation. Therefore the standard deviation became smaller.

The experimental results suggested that the average RMSE increased when the number of images was increased. This was because our shadow thresholding aggressively removed pixel coordinates that overlap with shadow regions. Therefore, it was more likely that the pixels near shadow regions were also removed so that the inverse response function could not be constrained well.

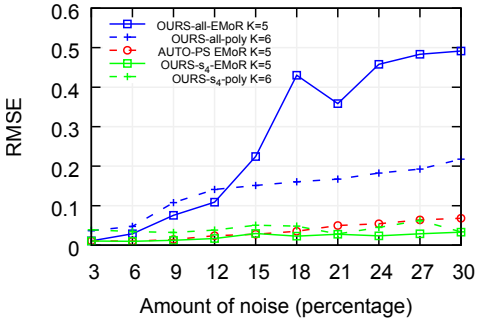
We assessed the inverse response function estimation accuracy of our method from noisy images. In these experiments, we implemented two cost functions given by (5.5) and (5.6) with two response function representation. We denote the implementation of (5.5) with EMoR and polynomials represen-



(a) agfapan-apx-100CD



(b) agfa-scala-200xCDStandard



(c) gamma+2.2

Figure 5.6: The mean of RMSE of the inverse response function estimated from images with various noise levels. (top) agfapan-apx-100CD (bottom-left) agfa-scala-200xCDStandard (bottom-right) gamma+2.2

tation as *Ours-all-EMoR* and *Ours-all-poly*, and denote the implementation of (5.6) with EMoR and polynomials representation as Ours-s_4-EMoR and Ours-s_4-poly.

We prepared the images for the experiments by using additive Gaussian noise to the calibrated images and then applied the nonlinear response functions to obtain the uncalibrated images.

We compared the inverse response functions estimated with all variants in Figure 5.7. It is clear that the variants with (5.6) were more robust to noise as they could produce inverse response functions with smaller RMSE

Table 5.1: The mean of angular difference in degrees of the surfaces estimated from synthetic images calibrated with our method, uncalibrated images, and calibrated images compared to the ground truth.

Response functions	Ours-s_4 PF14	Ours-all PF14	PF14 (uncal.)	PF14 (cal.)
agfapan-apx-100CD	5.15	5.16	9.40	5.226
agfa-scala-200xCDStandard	5.19	5.18	6.260	5.226
gamma+2.2	5.18	5.18	6.913	5.226
Response functions	Ours-s_4 QLD15	Ours-all QLD15	QLD15 (uncal.)	QLD15 (cal.)
agfapan-apx-100CD	5.14	5.15	13.18	5.101
agfa-scala-200xCDStandard	5.14	5.58	5.18	5.101
gamma+2.2	5.65	5.57	7.32	5.101

when noise increased. However, we found that Ours-all-EMoR and Ours-all-poly produced more accurate results than Ours-s_4-EMoR and Ours-s_4-poly when there is a small amount of noise. This was because nonlinear response functions were the major sources of the singular values variation in the image matrix when the noise was minimal. The Ours-s_4 variants, therefore, could not capture this variation, hence, produced sub-optimal solutions. In contrast, minimizing all of the singular values leads to the more accurate inverse response functions in this case.

On the other hand, when the noise level increased, the effect of noise on the singular value variation became larger everywhere. Ours-s_4-EMoR and Ours-s_4-poly excluded small singular values from the inverse response function estimation, therefore, they were more robust to noise.

Lastly, we assessed the accuracy in the surface estimation by applying our inverse response function estimation to two existing uncalibrated photometric stereo methods, namely, PF14[48] and QLD15[52]. Both methods assume directional light sources and require about 10-20 images to estimate surfaces with uncalibrated photometric stereo. In these experiments, we used the

Table 5.2: The RMSE of the estimated inverse response functions estimated from real images. Note that Auto-PS requires light source directions for the estimation.

input images	Ours-s_4	Ours-all	Auto-PS
sphere agfapan-apx-100CD	0.019	0.057	0.003
sphere agfa-scala-200xCDStandard	0.016	0.016	0.006
sphere gamma+2.2	0.018	0.018	0.066

synthetic images prepared with the response functions *agfapan-apx-100CD*, *agfa-scala-200xCDStandard*, and *gamma+2.2*.

We compared the estimated surfaces to those estimated with radiometrically uncalibrated and calibrated images. The qualitative results are shown in the Figure 5.8. The error maps indicated that our proposed method can compensate the nonlinearity of response functions so the estimated surfaces were more similar to the ones estimated from calibrated images regardless the inaccurate estimated inverse response functions. In contrast, the surfaces that were estimated directly from the uncalibrated images suffered from the nonlinearity so they differed from the ground truth.

The mean of angular error of the estimated surfaces was summarized in Table 5.1. The angular errors of the surfaces estimated from the calibrated images and from the uncalibrated images with our method were relatively the same. As we used noise-free images in these experiments, the difference of the mean angular error of the surfaces estimated by Ours-all and Ours-s_4 variants were insignificant.

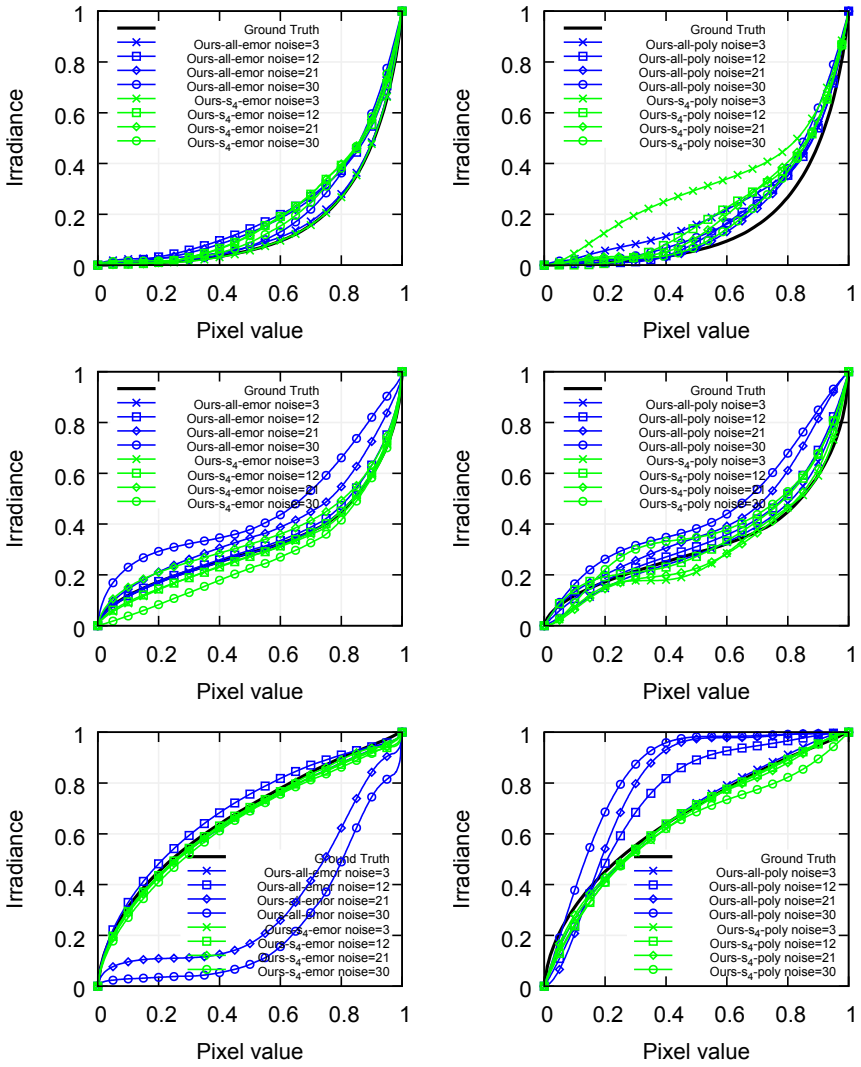
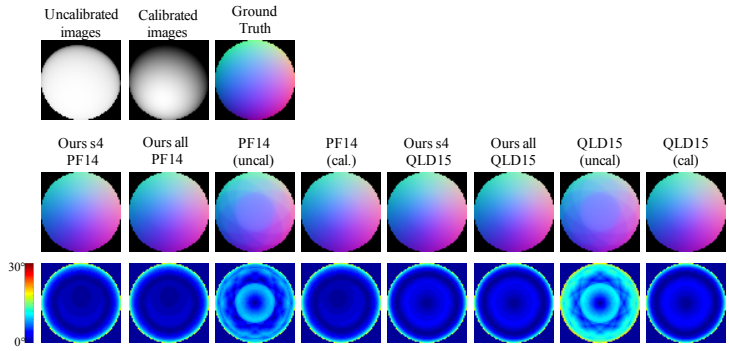
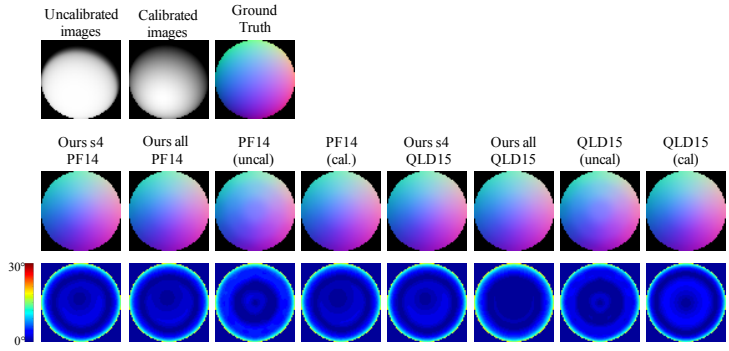


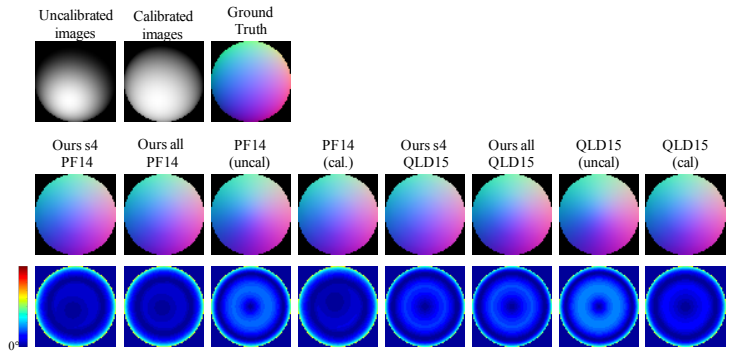
Figure 5.7: Samples of the inverse response function estimated from images with different noise level (top) agfapan-apx-100CD (middle) agfa-scala-200xCDStandard (bottom) gamma+2.2



(a) agfapan-apx-100CD



(b) agfa-scala-200xCD Standard



(c) gamma+2.2

Figure 5.8: Surface estimation results from synthetic images. (top) sample of inputs (middle) estimated surface normals (bottom) angular difference to the ground truth

5.5.2 Experiments on Real Images

We evaluated the estimation accuracy of our method with the experiments on images of real objects, *sphere*, and *statue*. *sphere* is a plaster sphere with broad specular surfaces and *statue* is a plaster sculpture. We placed the objects of interests under 20 light sources placed randomly over the objects then captured the images with a Point Grey Flea 2 camera. Instead of linear measurements of light intensity, we configured the camera's intensity lookup table so acted like a nonlinear camera in this setting. We configured the camera with the measured agfapan-apx-100CD, agfa-scala-200xCDStandard, and gamma+2.2 nonlinear response functions to represent three common shapes of nonlinear functions in the response function database[17]: concave, convex, and sigmoid.

Since there is no ground truth available for the target objects, we considered the surface normals estimated from the radiometrically calibrated input images with PS-SBL [24] as ground truth. Both have 28900 and 82800 foreground pixels respectively.

We performed the inverse response function estimation using two versions of our proposed method. Both versions used the EMoR representation but used different cost functions for the inverse response function estimation. More specifically, we implemented the optimization given in (5.5) and (5.6) and denoted them with Our-all and Our-s_4 respectively.

We then compared the estimated surfaces to the ones estimated from both radiometrically calibrated and uncalibrated with PF14 and QLD15.

Figure 5.9 and 5.11 show the qualitative experimental results. The error maps show the angular difference of the estimated surfaces compared to the ground truth. It is obvious to see that the surfaces estimated from the uncalibrated method were more different to the ground truth than ones estimated

Table 5.3: The mean of angular difference in degrees of the surfaces estimated from real images calibrated with our method, uncalibrated images, and calibrated images compared to the ground truth.

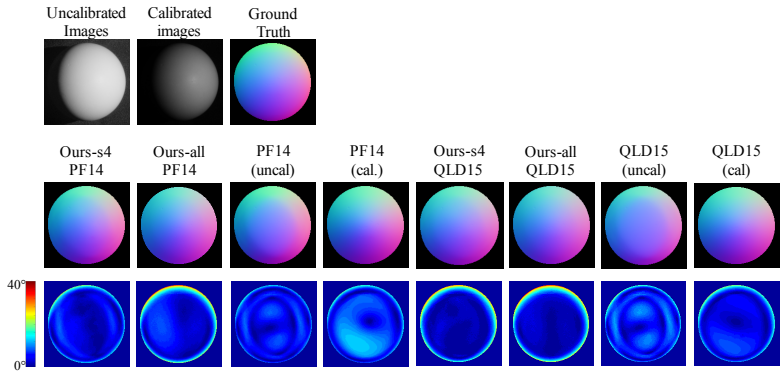
input images	Ours-s_4 PF14	Ours-all PF14	PF14 (uncal.)	PF14 (cal.)
sphere agfapan-apx-100CD	4.69	6.55	5.15	7.79
sphere agfa-scala-200xCDStandard	6.38	6.27	13.06	7.79
sphere gamma+2.2	9.48	8.18	30.10	7.79
statue agfapan-apx-100CD	4.87	9.35	7.49	7.49
statue agfa-scala-200xCDStandard	7.23	6.27	8.90	7.49
statue gamma+2.2	7.12	7.00	10.31	7.49
input images	Ours-s_4 QLD15	Ours-all QLD15	QLD15 (uncal.)	QLD15 (cal.)
sphere agfapan-apx-100CD	5.49	5.59	11.44	5.49
sphere agfa-scala-200xCDStandard	3.93	3.95	11.03	5.49
sphere gamma+2.2	6.49	5.83	20.90	5.49
statue agfapan-apx-100CD	9.05	13.04	10.74	10.5
statue agfa-scala-200xCDStandard	9.21	9.64	10.04	10.5
statue gamma+2.2	8.34	8.69	13.39	10.5

from the calibrated images. On the other hand, the inverse response function estimated from our proposed method compensated the nonlinearity of the response functions so that the estimated surfaces were relatively similar to the ones estimated from calibrated images.

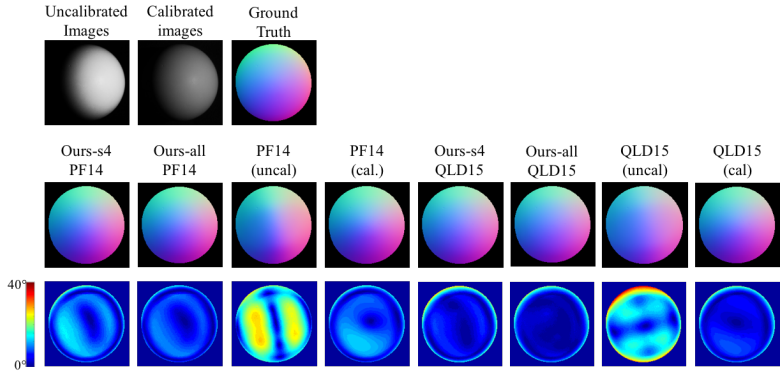
The quantitative results were summarized in Table 5.3. We can see that the angular error of the surfaces estimated with the calibrated images and the images calibrated with our methods had comparable angular errors. Although there were some image sets that the angular error did not improve the angular error, namely, *statue agfapan-apx-100CD* and *statue agfa-scala-200xCDStandard* estimated by QLD15, we can see that the images calibrated with our method produced more similar surfaces to the ground truth.

Table 5.4: The RMSE of the estimated inverse response functions estimated from real images. Note that Auto-PS requires light source directions for the estimation.

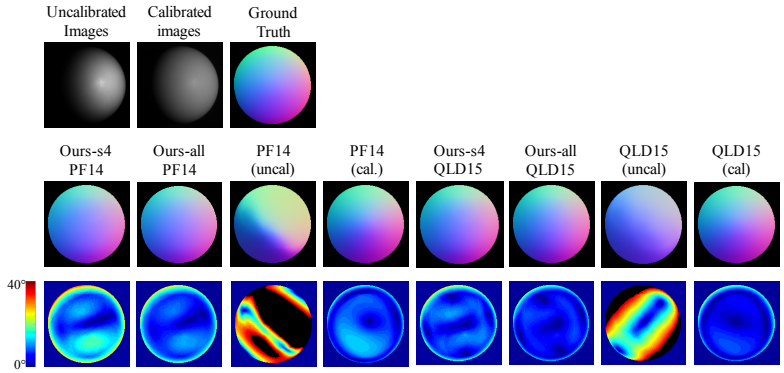
input images	Ours-s_4	Ours-all	Auto-PS
sphere agfapan-apx-100CD	0.143	0.168	0.061
sphere agfa-scala-200xCDStandard	0.117	0.104	0.038
sphere gamma+2.2	0.063	0.120	0.106
statue agfapan-apx-100CD	0.017	0.067	0.051
statue agfa-scala-200xCDStandard	0.079	0.187	0.020
statue gamma+2.2	0.058	0.126	0.115



(a) agfapan-apx-100CD

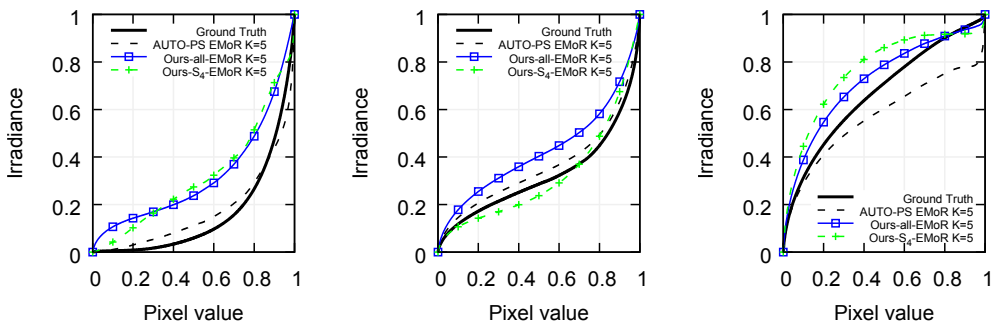


(b) agfa-scala-200xCD Standard



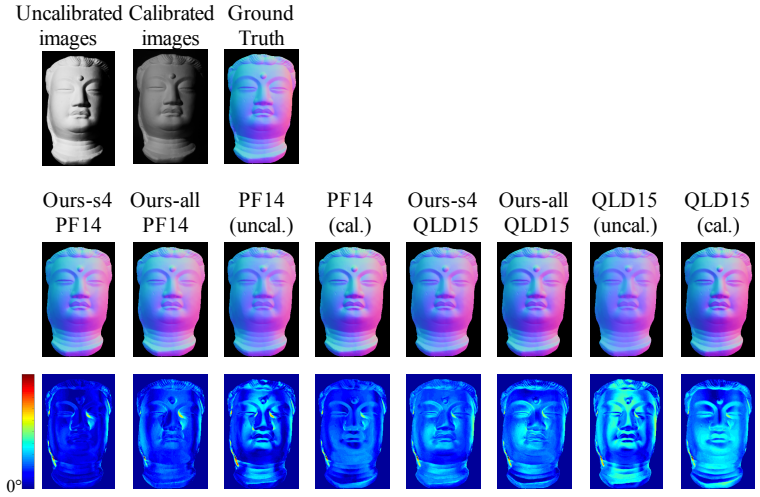
(c) gamma+2.2

Figure 5.9: Surface estimation results from real images *sphere*. (top) sample of inputs (middle) estimated surface normals (bottom) angular difference to the ground truth

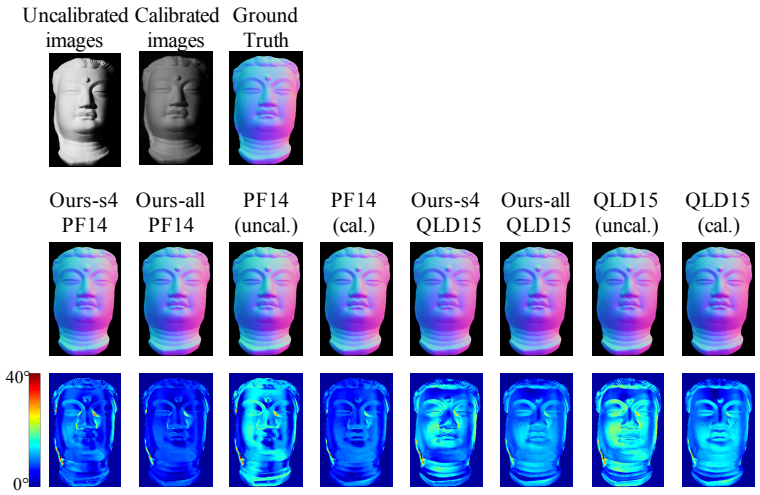


(a) (b) (c)

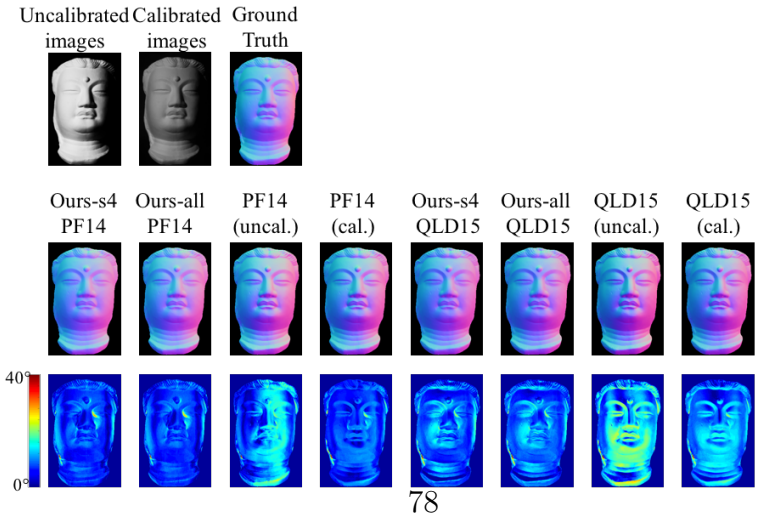
Figure 5.10: Inverse response functions estimated from the real images *Sphere*. (a) agfapan-apx-100CD (b) agfa-scala-200xCD Standard (c) gamma+2.2



(a) agfapan-apx-100CD

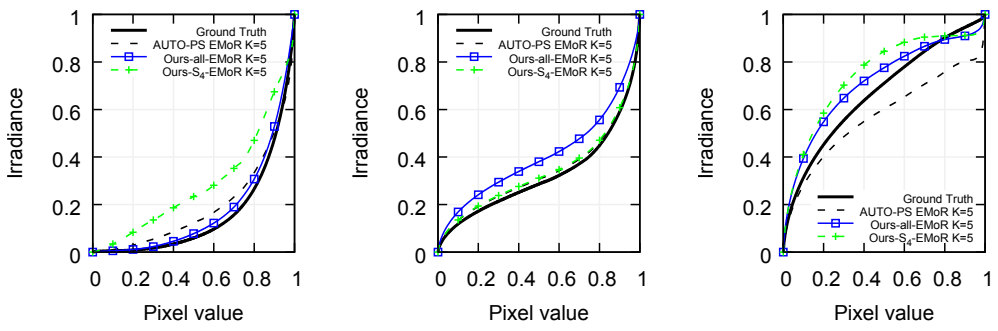


(b) agfa-scala-200xCD Standard



(c) gamma+2.2

Figure 5.11: Surface estimation results from real images *statue*. (top) sample of inputs (middle) estimated surface normals (bottom) angular difference to the ground truth



(a) (b) (c)

Figure 5.12: Inverse response functions estimated from the real images *Statue*. (a) agfapan-apx-100CD (b) agfa-scala-200xCD Standard (c) gamma+2.2

5.6 Conclusion

In this chapter, we present an auto radiometric calibration method for photometric stereo images. It makes use of linear dependency and prior knowledge about the rank of the to determine an inverse response function that restores the best rank-3 structure from the input images. This method does not assume known light source directions and intensities. Therefore we can apply this method to the datasets for uncalibrated photometric stereo. We empirically show that our proposed method can compensate the nonlinearity of the response function so estimate the inverse response function relatively well even though the light source direction is not known beforehand.

Our proposed method has a major limitation. It clearly cannot determine response functions of a data set which the images are on the same plane or relatively similar. Moreover, we aggressively exclude pixel coordinates from the inverse response function estimation even if there exists a shadow pixel on that coordinate. Therefore, there will be fewer pixels left for the inverse response function estimation if the input images have large regions of shadows in the foreground.

Chapter 6

Conclusion

We have presented a photometric stereo technique that accounts for the effect of nonlinearity of response function that resides in the input images that were captured by a camera with an unknown non-linear response function. The proposed method makes use of the consistency of the irradiance calculated from a reflection model and the irradiance calculated from an inverse response function. By expressing the inverse response function as a linear combination of basis functions, we can formulate a linear least squares problem to estimate surface normals and coefficients that best match the input images.

Then, we have also presented an extension to handle non-Lambertian surfaces. In this work, we developed a unified framework that incorporates arbitrary photometric stereo technique into our inverse response function estimation method. Our key assumption in this work is that pixels with diffusion reflection are dominant and specular highlights can be observed within limited angles. Then the corrupted measurements such as noisy pixels and specular highlights can be regarded as outliers to the Lambertian reflection model. We have done this by using RANSAC algorithm to propose the simultaneous surface normals and inverse response function estimation framework

that maximizes the number of diffuse pixels.

We refocused our attention to pursue more practical ground in auto radiometric calibration by continuing our research into the uncalibrated photometric stereo. We have presented an auto radiometric calibration method that takes uncalibrated photometric stereo input images and gives inverse response functions to cancel out the effect of the response function used for capturing the scenes as an output. We used a priori knowledge regarding the reflectance of the objects to construct an image matrix that exhibits linear dependency. Then we constrain the rank property to solve for the coefficients of the inverse response functions.

As a whole, we have proposed surface estimation techniques that not only accounts for surface properties but also camera properties. They make use of inherent cues that reside in the images to allow ones to recover scene structure from images with less cumbersome calibration step.

6.1 Future Work

There are a number of open questions that still left to be addressed. We have not explored how to incorporate non-Lambertian diffuse reflection and more complex materials such as BRDF into our model. Since we assume that non-Lambertian model exhibit Lambertian reflection where the specular highlight is not observed, it fails to handle many real-world surfaces that are better explained by BRDF. Therefore, it is interesting to directly extend the consistency of the irradiance for such models to capture more real-world reflections.

Another possible research would be incorporating the auto radiometric calibration under more complex illumination such as a near-point light

source. It is known that light fall off is an important clue for constraining GBR ambiguity [38, 49, 64] and radiometric calibration [50]. So it would be interesting whether it can be applied to auto radiometric calibration as well.

Last but not least, there are some things we have yet to try in our experiments, *e.g.*, replacing our fix threshold for shadow thresholding with an adaptive threshold so it can capture more non-shadow pixel intensities that are altered to be shadow more effectively.

Bibliography

[1] A. Abrams, C. Hawley, and R. Pless. Heliometric stereo: shape from sun position. In *Computer Vision–ECCV 2012*, pages 357–370. Springer, 2012.

[2] N. Alldrin, S. Mallick, and D. Kriegman. Resolving the generalized bas-relief ambiguity by entropy minimization. In *Proc. of Computer Vision and Pattern Recognition*, pages 1 –7, Jun. 2007.

[3] V. Argyriou, M. Petrou, and S. Barsky. Photometric stereo with an arbitrary number of illuminants. *Computer Vision and Image Understanding*, 114(8):887–900, 2010.

[4] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(10):1239 – 1252, Oct. 2003.

[5] R. Basri, D. Jacobs, and I. Kemelmacher". "photometric stereo with general, unknown lighting.". *"International Journal of Computer Vision"*, "72"("3"):"239–257", "2007".

[6] P. Belhumeur, D. Kriegman, and A. Yuille. The bas-relief ambiguity. In *Proc. of Computer Vision and Pattern Recognition*, pages 1060 –1066, Jun. 1997.

[7] Z. Botev and D. Kroese. Global likelihood optimization via the cross-entropy method with an application to mixture models. In *Simulation Conference, 2004. Proceedings of the 2004 Winter*, volume 1, pages –535, Dec 2004.

[8] N. Campbell. Discontinuities in light emission. *Proceedings of the Cambridge Philosophical Society, Mathematical and physical sciences*, 15:310–328, 1909.

[9] E. N. Coleman Jr. and R. Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Computer Graphics and Image Processing*, 18(4):309 – 328, 1982.

[10] R. L. Cook and K. E. Torrance. A reflectance model for computer graphics. *ACM Transactions on Graphics (TOG)*, 1(1):7–24, 1982.

[11] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *Proc. of Computer Graphics and Interactive Techniques*, pages 369–378, 1997.

[12] O. Drbohlav and R. Sára. Specularities reduce ambiguity of uncalibrated photometric stereo. In *Proc. of European Conference on Computer Vision-Part II*, pages 46–62, May 2002.

[13] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, Jun. 1981.

[14] A. Georghiades. Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo. In *Proc. of Int'l Conf. on Computer Vision*, volume 2, pages 816–823, Oct. 2003.

[15] D. Goldman, B. Curless, A. Hertzmann, and S. Seitz. Shape and spatially-varying brdfs from photometric stereo. In *Proc. of Int'l Conf. on Computer Vision*, volume 1, pages 341–348, Oct. 2005.

[16] M. Grossberg and S. Nayar. Determining the camera response from images: What is knowable? *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(11):1455–1467, Nov. 2003.

[17] M. Grossberg and S. Nayar. What is the Space of Camera Response Functions? In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume II, pages 602–609, Jun 2003.

[18] H. Hayakawa. Photometric stereo under a light source with arbitrary motion. *Journal of the Optical Society of America*, 11(11):3079–3089, Nov. 1994.

[19] T. Higo, D. Miyazaki, and K. Ikeuchi. Multi-view photometric stereo using rough shape data application for estimation of reflection parameters. In *Proc. of Meeting on Image Recognition and Understanding*, pages 1280–1287, Jul. 2008.

[20] Y. Hold-Geoffroy, J. Zhang, P. F. U. Gotardo, and J. F. Lalonde. x-hour outdoor photometric stereo. In *2015 International Conference on 3D Vision*, pages 28–36, Oct 2015.

[21] M. Holroyd, J. Lawrence, G. Humphreys, and T. Zickler. A photometric approach for estimating normals and tangents. *ACM Trans. Graph.*, 27(5):133:1–133:9, Dec. 2008.

[22] C.-H. Hung, T.-P. Wu, Y. Matsushita, L. Xu, J. Jia, and C.-K. Tang. Photometric stereo in the wild. In *Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on*, pages 302–309. IEEE, 2015.

[23] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa. Robust photometric stereo using sparse regression. In *Proc. of Computer Vision and Pattern Recognition*, pages 318–325, Jun. 2012.

[24] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa. Photometric stereo using sparse bayesian regression for general diffuse surfaces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 36(9):1816–1831, Sept 2014.

[25] S. J. Kim and M. Pollefeys. Radiometric alignment of image sequences. In *Proc. of Computer Vision and Pattern Recognition*, volume 1, pages I–645–I–651, 2004.

[26] S. J. Kim and M. Pollefeys. Robust radiometric calibration and vignetting correction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(4):562–576, April 2008.

[27] J. Y. Lee, Y. Matsushita, B. Shi, I. S. Kweon, and K. Ikeuchi. Radiometric calibration by rank minimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):144–156, Jan 2013.

[28] C. Li, S. Lin, K. Zhou, and K. Ikeuchi. Radiometric calibration from faces in images. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[29] J. Liao, B. Buchholz, J.-M. Thiery, P. Bauszat, and E. Eisemann. Indoor scene reconstruction using near-light photometric stereo. *IEEE Transactions on Image Processing*, 26(3):1089–1101, 2017.

[30] S. Lin, J. Gu, S. Yamazaki, and H.-Y. Shum. Radiometric calibration from a single image. In *Proc. of Computer Vision and Pattern Recognition*, volume 2, pages II–938–II–945, Jun.-Jul. 2004.

[31] S. Lin and L. Zhang. Determining the radiometric response function from a single grayscale image. In *Proc. of Computer Vision and Pattern Recognition*, volume 2, pages 66–73, Jun. 2005.

[32] S. Lin and L. Zhang. Determining the radiometric response function from a single grayscale image. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, pages 66–73 vol. 2, June 2005.

[33] F. Lu, X. Chen, I. Sato, and Y. Sato. Symps: Brdf symmetry guided photometric stereo for shape and light source estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PP(99):1–1, 2017.

[34] F. Lu, I. Sato, and Y. Sato. Uncalibrated photometric stereo based on elevation angle recovery from brdf symmetry of isotropic materials. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 168–176, June 2015.

[35] S. Mann and R. W. Picard. On being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures. In *Proc. of IS&T's 48th Annual Conference*, pages 442–448, 1995.

[36] Y. Matsushita and S. Lin. Radiometric calibration from noise distributions. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.

[37] K. Midorikawa, T. Yamasaki, and K. Aizawa. Uncalibrated photometric stereo by stepwise optimization using principal components of isotropic brdfs. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4350–4358, June 2016.

[38] T. Migita, S. Ogino, and T. Shakunaga. *Direct Bundle Estimation for Recovery of Shape, Reflectance Property and Light Position*, pages 412–

425. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[39] T. Migita, K. Sogawa, and T. Shakunaga. *Specular-Free Residual Minimization for Photometric Stereo with Unknown Light Sources*, pages 178–189. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[40] T. Mitsunaga and S. Nayar. Radiometric self calibration. In *Proc. of Computer Vision and Pattern Recognition*, volume 1, pages 374–380, 1999.

[41] Z. Mo, B. Shi, S.-K. Yeung, and Y. Matsushita. Radiometric calibration for internet photo collections. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4738–4746, 2017.

[42] W. Mongkulmann, T. Okabe, and Y. Sato. Photometric stereo with auto-radiometric calibration. In *Proc. of IEEE Workshop on Color and Photometry in Computer Vision (CPCV2011)*, pages 753–758, Nov. 2011.

[43] T. Mori, R. Taketa, S. Hiura, and K. Sato. *Photometric Linearization by Robust PCA for Shadow and Specular Removal*, pages 211–224. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[44] Y. Mukaigawa, Y. Ishii, and T. Shakunaga. Analysis of photometric factors based on photometric linearization. *Journal of the Optical Society of America*, 24(10):3326–3334, Oct. 2007.

[45] S. Nayar, K. Ikeuchi, and T. Kanade. Determining Shape and Reflectance of Hybrid Surfaces by Photometric Sampling. *IEEE Trans. Robot. Autom.*, 6(4):418–431, Aug. 1990.

[46] T. Okabe, I. Sato, and Y. Sato. Attached shadow coding: Estimating surface normals from shadows under unknown reflectance and lighting

conditions. *2009 IEEE 12th International Conference on Computer Vision*, pages 1693–1700, 2009.

[47] T. Okabe and Y. Sato. Object recognition based on photometric alignment using ransac. In *Proc. of Computer Vision and Pattern Recognition*, volume 1, pages I–221–I–228, Jun. 2003.

[48] T. Papadhimitri and P. Favaro. A closed-form, consistent and robust solution to uncalibrated photometric stereo via local diffuse reflectance maxima. *International Journal of Computer Vision*, 107(2):139–154, Apr 2014.

[49] T. Papadhimitri and P. Favaro. Uncalibrated near-light photometric stereo. In *Proceedings of the British Machine Vision Conference*. BMVA Press, 2014.

[50] J. Park, S. N. Sinha, Y. Matsushita, Y. W. Tai, and I. S. Kweon. Calibrating a non-isotropic near point light source using a plane. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2267–2274, June 2014.

[51] B. T. Phong. Illumination for computer generated pictures. *Communications of the ACM*, 18(6):311–317, 1975.

[52] Y. Quéau, F. Lauze, and J.-D. Durou. Solving uncalibrated photometric stereo using total variation. *Journal of Mathematical Imaging and Vision*, 52(1):87–107, May 2015.

[53] R. Ramanath, W. Snyder, Y. Yoo, and M. Drew. Color image processing pipeline. *Signal Processing Magazine, IEEE*, 22(1):34–43, Jan 2005.

[54] H. Rushmeier, G. Taubin, and A. Guéziec. Applying shape from lighting variation to bump map capture. In J. Dorsey and P. Slusallek, editors,

Rendering Techniques '97, Eurographics, pages 35–44. Springer Vienna, 1997.

[55] B. Shi, Y. Matsushita, Y. Wei, C. Xu, and P. Tan. Self-calibrating photometric stereo. In *Proc. of Computer Vision and Pattern Recognition*, pages 1118–1125, Jun. 2010.

[56] P. Tan, S. P. Mallick, L. Quan, D. J. Kriegman, and T. Zickler. Isotropy, reciprocity and the generalized bas-relief ambiguity. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.

[57] K. E. Torrance and E. M. Sparrow. Theory for off-specular reflection from roughened surfaces. *JOSA*, 57(9):1105–1112, 1967.

[58] G. J. Ward. Measuring and modeling anisotropic reflection. *ACM SIGGRAPH Computer Graphics*, 26(2):265–272, 1992.

[59] B. Wilburn, H. Xu, and Y. Matsushita. Radiometric calibration using temporal irradiance mixtures. In *Proc. of Computer Vision and Pattern Recognition*, pages 1–7, Jun. 2008.

[60] R. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.

[61] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma. Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. In *Advances in neural information processing systems*, pages 2080–2088, 2009.

[62] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *Proc. of Asian Conference on Computer vision*, pages 703–717. Springer-Verlag, 2011.

[63] +Colorimetric values for ColorChecker Family of Targets, 2010. http://xritephoto.com/ph_product_overview.aspx?ID=824&Action=Support&SupportID=5159.

[64] W. Xie, C. Dai, and C. C. L. Wang. Photometric stereo with near point lighting: A solution by mesh deformation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4585–4593, June 2015.

Publications

1. Wiennat Mongkulmann, Takahiro Okabe, Yoichi Sato, "Automatic Radiometric Calibration in Photometric Stereo", IEICE Trans., on Information and Systems. vol. 99, no. 12, pp.3154-3164, 2016.

2. モンコンマーン　ウィーンナート, 岡部　隆弘, 佐藤　洋一, "カメラ応答関数の自動校正を伴う照度差ステレオ：非ランバート物体への拡張". 2013-CVIM-185(3), pp.1-8, Jan. 2013.

3. Wiennat Mongkulmann, Takahiro Okabe, Yoichi Sato, "Automatic Radiometric Calibration in Photometric Stereo by Using Irradiance Consistency". 2011-CVIM-179(21), pp.1-7, Nov. 2011.

4. Wiennat Mongkulmann, Takahiro Okabe, and Yoichi Sato, "Photometric stereo with auto-radiometric calibration", in ICCV Workshops, pp. 753-758, Nov. 2011.