

論文 / 著書情報
Article / Book Information

題目(和文)	シングルキャプチャによる照度差ステレオ法の開発
Title(English)	Development of single capture photometric stereo methods
著者(和文)	小澤圭右
Author(English)	keisuke ozawa
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第10922号, 授与年月日:2018年6月30日, 学位の種別:課程博士, 審査員:佐藤 いまり,山口 雅浩,熊澤 逸夫,金子 寛彦,小尾 高史
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第10922号, Conferred date:2018/6/30, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Development of single capture photometric stereo methods (シングルキャプチャによる照度差ステレオ法の開発)

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Engineering
at
Tokyo Institute of Technology
Department of Information Processing

Keisuke Ozawa

Supervised by: Imari Sato and Masahiro Yamaguchi

May 23th, 2018

Contents

1	Introduction	1
1.1	3D geometry from 2D image	1
1.2	Photometric stereo and issue	2
1.3	Outline	4
2	Basics of Photometric stereo	6
2.1	Light transport and image formation	6
2.2	Linear photometric stereo problem for Lambertian surfaces	8
2.3	Extension to nonlinear problems	9
2.4	Color photometric stereo	10
2.5	Related work in color photometric stereo	11
3	Hyperspectral photmoetric stereo for a single capture	15
3.1	Method	16
3.1.1	Overview	16
3.1.2	Illumination design	18
3.1.3	Reflectance estimation	20
3.1.4	Surface normal estimation	21
3.2	Simulation	22
3.2.1	Comb illumination	22
3.2.2	Setup	22
3.2.3	Results and discussion	23
3.2.4	Filtered illumination	24
3.2.5	Setup	24
3.2.6	Results and discussion	28
3.3	Real scene	29
3.3.1	Experimental setup	30
3.3.2	Hyperspectral image denoising	31
3.3.3	Results and discussion	32
3.3.4	Effect of notched-band on spectral estimation	33
3.4	Conclusion	35
4	Single color image photometric stereo for multi-colored surfaces	36
4.1	Method	37
4.1.1	Observation under three light sources	37
4.1.2	On unique solution for color reflectance	39
4.1.3	Surface-color feature on multi-colored surfaces	40
4.1.4	Surface classification via reflectance norm distribution	42
4.1.5	Surface color classification algorithm	45
4.1.6	Resolving classification ambiguity	47
4.2	Simulations	48
4.2.1	Evaluation for variety of textures	50
4.2.2	Comparison against standard photometric stereo	51
4.2.3	Complicated surface geometry	52

4.3	Experiment with real objects	52
4.3.1	Experimental setup	52
4.3.2	Results and discussion	53
4.4	Conclusion	58
5	General consideration	59
5.1	Multiband surface color classification in photometric stereo	59
5.2	Multiband spectral regression in photometric stereo	61
5.3	Conclusion	64

1 Introduction

1.1 3D geometry from 2D image

The surface geometry of an object provides important visual information. Once a surface geometry is measured and registered in memory, it can then be used for various applications such as the digital archive of historical heritages, the surface analysis of artifacts, computer graphics and 3D printer models, and medical analysis, depending on for what the surface is measured. Unlike that the human eyes can sense depth in what we see, cameras except for RGB-D cameras that capture RGB images as well as depth, while having been invented to reproduce and store human vision, do not have depth information themselves. With deep-learning technology it has recently been attempted to estimate depth in a 2D image [1], and even a 2D image alone is now regarded informative for depth recognition once trained adaptively as human perception is recognized to be in its unconscious stratum. On the other hand, such approaches have been developed for rough depth acquisition in automated control or human assistance. Contrary, we focus on fine-structural reconstruction from 2D observations.

The technical foundation established in the history of surface measurement has achieved increasing accuracy and stability from geometrical measurement methods that calculate the depth maps or 3D coordinate sets, from the 2D measurements of scenes. Presumably a class of the most widely used techniques incorporate electromagnetic fields, which interact with the object surfaces electrically and without mechanical contact. Such optical measurement techniques are in general preferable for applications in non-intrusive analyses and scenes that we cannot physically contact with. Especially in computer vision, radiometry (photometry in jargon of the human eye system) provides a central cue for surface measurement. Such 3D-from-2D reconstruction techniques, while having been widely applied in industries, have unsettled problems and remain popular research interests today.

Photometric stereo is a technique that recovers surface geometry from 2D images. It observes radiometric responses rather than geometric cues found in multiple images that are often used in stereopsis. A specific advantage of photometric stereo comes from the tangent-recovering nature, and the recovered geometries generally exhibit high spatial-frequency details.

1.2 Photometric stereo and issue

We have depth sensation from geometric cues such as the relative sizes of objects near and far from us. It is endowed by nature as an indispensable part of human recognition system. Binocular stereo or stereopsis, and multi-view stereo [2] might be recognized as a computational realization of that by giving point-to-point correspondence between multiple views, and has been evolving for better accuracy and larger scale. However, this technique and its derivatives require multiple cameras and at least a pair of corresponding points in two images from different camera poses, and often the results suffer from inaccurate reconstruction against textureless surfaces.

Another important visual cue is the brightness and shading defined in radiometry. Photometric stereo recovers the surface normals of a surface from multiple shading images. Let us see a simplest sketch of how this method works. Consider a diffuse hemisphere lit by a light source, and we can roughly guess the curvature of the surface from the shading. Unlike geometric measurement techniques, the relationship between a light direction and surface normals is embedded in a single shading image. A constraint on surface normals is determined from those shading images captured under known lighting directions. Shape from shading technique [3] aims to reconstruct surface geometry from a shading image of a surface of a uniform reflectance. In photometric stereo [4], surface normals are determined uniquely from at least three shading images from different lighting directions.

Recovering surface normals means a fine surface reconstruction to its high-frequency details. A surface geometry is accompanied by its surface normal field, each vector at a point of which is equivalent with the tangent space at that point. Therefore, photometric stereo problems are in essence partial differential problems and require a boundary condition on the surface, which is not given as a prior knowledge in most photometric stereo problems. However, this is not problematic in many applications where the surface geometry itself or relative distances are important rather than the depth values in the external world. Except for some approaches that estimate accurate heights, the 3D shape is represented by a height field with arbitrary zero height. The capability of recovering spatially fine, high-frequency surface details and single-view setups are rather preferable to other surface reconstruction methods based on geometric cues. One might see the conceptual difference between radiometry-based photometric stereo and distance-based methods in Figure 1 that also sums up the procedural concept of photometric stereo that tracks the solid arrows, while the half-dot lines indicates a direct depth measurement procedure in the other optical methods. Note that these two approaches using shading

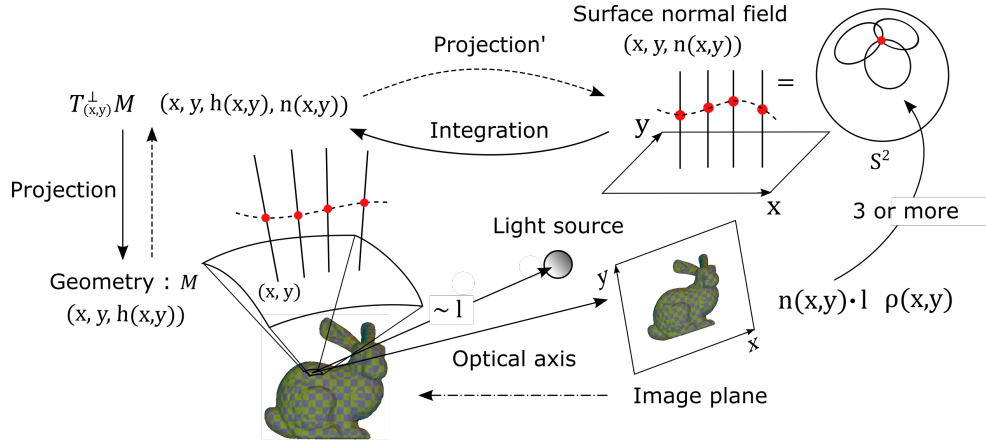


Figure 1: Diagram of surface geometry measurement in photometric stereo.

and geometric cue have been combined e.g. in [5] and many others.

An important issue in photometric stereo research is development for dynamic scenes. Since basic approaches in photometric stereo assume surfaces to be stationary while capturing images, it cannot be directly applied to dynamic scenes where the camera-surface-light source positions change over frames. This is a bottleneck for photometric stereo to become a practical surface measurement technique in a wider kind of scenes. Other important and widely studied issues must be in mind that extend for surfaces of complicated reflectance properties and considerations on uncalibrated light sources. The first conventional photometric stereo [4] was developed for Lambertian surfaces, and has been extended to non-Lambertian surfaces. Throughout this dissertation that focuses on how to estimate surface geometry with a single capture, we target Lambertian surfaces under calibrated light source directions as any other proposed methods, which we review in Section 2.5.

To this end, a folk of photometric stereo that lit multiple light sources of different emission spectra, called color photometric stereo, has been studied mainly for this reason toward dynamic scenes. The conventional photometric stereo method with multiple gray-scale images discard the color information, while in color photometric stereo methods, the color of surface and illumination is facilitated. All current color photometric stereo methods use multiple cameras or frames, or even with a single frame assuming relatively hard geometric or chromatic constraints on surfaces. As will be reviewed in Chapter 2.5, these methods have in fact achieved dynamic scene reconstruction under some geometric or chromatic limitations. On the other hand, it still remains an issue to develop single-capture-based methods in the color photometric stereo framework against colored surfaces.

1.3 Outline

This dissertation is aimed at developing photometric stereo toward single capture techniques. The problem is that what kind of surface color set is permissible of which surface geometry can be reconstructed from a single capture. To mitigate hard assumptions in the existing methods, we develop two single-image-based photometric stereo methods that require a single camera and frame, targeting multi-colored Lambertian surfaces with moderate assumptions on the colors.

We firstly summarize the basics of photometric stereo and color photometric stereo in Chapter 2. We consider the full-spectral formulation of light transport and the photometric stereo linear problem as its first-order approximation. Some related work on color photometric stereo are listed and compared against our assumptions and methods. We also briefly review efficient extensions for non-Lambertian surfaces of photometric stereo. Although not being relevant to our methods, such nonlinear extensions of photometric stereo have quite importance for the realization of practical photometric stereo methods. We do not refer to uncalibrated photometric stereo methods rather than suggesting that our second method in Chapter 4 assumes calibrated light directions but uncalibrated light source intensities for estimating only surface normals without knowing the material-intrinsic reflectances. The latter calibration often appears as the more difficult prerequisite, from which our method is free.

In Chapter 3, we present a single capture photometric stereo method that use a single hyperspectral image. Light transport phenomena are the propagation of light and its interaction with objects. Hyperspectral imaging techniques enable the full-wavelength imaging over a certain wavelength region, and convey informative spectral signals emitted or reflected from object surface. We design illumination spectra and spatial coordination specifically for obtaining reflectance spectra of target surfaces and multiple shading images in a single hyperspectral image. This method in principle enables a point-wise estimation of surface normals and the full reflectance spectra on a target surface. The spectral design of the illumination has some options in its notched bands, the optimum choice of which is also discussed.

Although hyperspectral imaging techniques have been recently developing themselves in spatial resolution, signal-to-noise (SNR) ratio, and capture speed, their use are still caught a dilemma between the current ability of hyperspectral cameras and required accuracy and cost not only for speed-SNR trade off but also for expensiveness. In Chapter 4, we present another single capture method that use a single RGB color image. It is an ill-posed problem to simultaneously determine the RGB color and

surface normal at a surface point from an RGB color observation. On the other hand, it is known that the global color observation is informative in determining surface normals within a region of an unknown, but constant color. We also facilitate global color observation and introduce a surface color feature that enables to classify colors that are ambiguously observed under illumination of three different emission spectra, and then estimate surface normals. From evaluations using synthetic images and real scenes, we show the accuracy and effectiveness of our method.

In chapter 5, we extend our presented methods to general multiband cases. Hyperspectral images and RGB images provide an upper bound and the lower bound for estimating surface normals from shading images. In each method, we facilitated global spectral information or global geometric/chromatic information to estimate reflectance spectra or color classification. In this final chapter, we discuss some possible augmentations of these approaches that estimate surface normals and multispectral reflectances from a single multiband image.

2 Basics of Photometric stereo

In this section, we provide the basics of photometric stereo. First, we start with the conventional monochromatic photometric stereo, and then color photometric stereo, a spectral augmentation of which and our interest. We also review some existing color photometric stereo methods, where unsettled problems are discussed in relation to our purpose.

2.1 Light transport and image formation

Surface geometry is accompanied by its tangent bundle, which is uniquely determined by the surface normal field. A surface geometry as represented by a height map is obtained from the surface normal integration. Photometric stereo relies on radiometry on the tangent space of a surface, and the light transport equation describes the radiometric phenomena [6].

When the wavelengths of light that stimulates a surface electrically are sufficiently small compared to the spatial dielectric structure of the surface, the ray optics takes the place of the electric field propagation. Microscopic responses of electric substances are now averaged to macroscopic response functions on the tangent space called the bidirectional reflectance distribution functions (BRDFs). We do not discuss the intermediate scales that have provided some analytical reflectance models, and consider isotropic surfaces. Since such macroscopic optical responses depend on the light position, the tangent space, and the observer position, a BRDF has three scalar arguments of the inner products between the light direction and surface normal, the surface normal and observer direction, and the observer direction and light direction.

Using the BRDF $f(\mathbf{x}'' \rightarrow \mathbf{x}' \rightarrow \mathbf{x})$ from a light source position $\mathbf{x}''$ through a surface point $\mathbf{x}'$ to the observer position $\mathbf{x}$, the observed radiance L at $\mathbf{x}'$ is described by the following light transport equation that takes all the pass integrated over the surface area Ω :

$$L(\mathbf{x}' \rightarrow \mathbf{x}) = E(\mathbf{x}' \rightarrow \mathbf{x}) + \int_{\mathbf{x}'' \in \Omega} f(\mathbf{x}'' \rightarrow \mathbf{x}' \rightarrow \mathbf{x}) L(\mathbf{x}' \rightarrow \mathbf{x}) V(\mathbf{x}'' \leftrightarrow \mathbf{x}') \cos \theta_{\text{obs}} d\omega(\mathbf{x}'' \rightarrow \mathbf{x}'), \quad (1)$$

where E is an light emission and V is the mutual visibility function for two points, and θ_{obs} is the angle between the lien segment $\overline{\mathbf{x}'\mathbf{x}}$ and the surface normal at $\mathbf{x}$, or the direction of the observation plane. Note that the BRDF is phenomenological macroscopic representation of surface response, while the

visibility function is introduced in purely macroscopic view. Since the solid angle is

$$d\omega(\mathbf{x}'' \rightarrow \mathbf{x}') = \frac{\cos \theta_{\mathbf{x}'' \rightarrow \mathbf{x}'}}{\|\mathbf{x}'' - \mathbf{x}'\|^2} d\mu'' \quad (2)$$

using the light incidence angle $\theta_{\mathbf{x}'' \rightarrow \mathbf{x}'}$, the light transport equation becomes

$$L(\mathbf{x}' \rightarrow \mathbf{x}) = E(\mathbf{x}' \rightarrow \mathbf{x}) + \int_{\Omega} f(\mathbf{x}'' \rightarrow \mathbf{x}' \rightarrow \mathbf{x}) L(\mathbf{x}' \rightarrow \mathbf{x}) G(\mathbf{x}'' \leftrightarrow \mathbf{x}') d\mu'', \quad (3)$$

where

$$G(\mathbf{x}'' \leftrightarrow \mathbf{x}') = V(\mathbf{x}'' \leftrightarrow \mathbf{x}') \frac{\cos \theta_{\mathbf{x}'' \rightarrow \mathbf{x}'} \cos \theta_{\text{obs}}}{\|\mathbf{x}'' - \mathbf{x}'\|^2} \quad (4)$$

is called the geometric term.

This self-consistent equation is expanded as

$$L(\mathbf{x}' \rightarrow \mathbf{x}) = E(\mathbf{x}' \rightarrow \mathbf{x}) + \int_{\Omega} f(\mathbf{x}'' \rightarrow \mathbf{x}' \rightarrow \mathbf{x}) E(\mathbf{x}'' \rightarrow \mathbf{x}') G(\mathbf{x}'' \leftrightarrow \mathbf{x}') d\mu'' + \mathcal{O}(f^2), \quad (5)$$

where the first component is the emission of a source, and the second is the primary term of transport, or the direct reflection component, and the higher order terms represent interreflections.

Light emissions are not observed in our measurement system: $E(\mathbf{x}' \rightarrow \mathbf{x}) = 0$. In photometric stereo often be assumed full-convex surfaces and thus no interreflections, which also can be as treated as errors or outliers, and the surface is assumed to be located at the origin so that $\|\mathbf{x}'' - \mathbf{x}'\|$ is constant over the system. By omitting the scalar terms of $\cos \theta_{\text{obs}}$ and $\|\mathbf{x}'' - \mathbf{x}'\|^2$ emerged into the measurement system and rewriting with the observation point $\mathbf{x}_c$, observed surface position $\mathbf{x}$, and light source positions on $\mathbf{x}_s \in \Omega_s$, the radiance is simply

$$L(\mathbf{x}) = \int_{\Omega_s} f(\mathbf{x}, \alpha, \beta, \gamma) E(\mathbf{x}_s \rightarrow \mathbf{x}) \cos \theta_{\mathbf{x}_s \rightarrow \mathbf{x}} d\mu_s \quad (6)$$

$$\approx f(\mathbf{x}, \alpha, \beta, \gamma) E \mathbf{l} \cdot \mathbf{n}(\mathbf{x}), \quad (7)$$

where the light sources are assumed to be point light sources at $\mathbf{x}_{s,0}$: $E(\mathbf{x}_s \rightarrow \mathbf{x}) = E\delta(\mathbf{x}_s - \mathbf{x}_{s,0})$ and the light direction is $\mathbf{l} \equiv \mathbf{x}_{s,0}/\|\mathbf{x}_{s,0}\|$ in the approximation: throughout this dissertation, it is assumed that the surface is enclosed within a ball of a radius sufficiently smaller than the distance to the light sources. Such a point light source assumption can be alleviated by expressing it with parameterized

distribution functions [7]. We have also used the fact that BRDFs are functions of the inner products between the surface normal, source direction, and viewing direction: $\alpha = \alpha(\mathbf{x}) = \mathbf{n}(\mathbf{x}) \cdot \mathbf{l}$, $\beta = \mathbf{l} \cdot \mathbf{v}$, and $\gamma = \gamma(\mathbf{x}) = \mathbf{v} \cdot \mathbf{n}(\mathbf{x})$, since BRDFs are scalars independent of the coordinates and each vector of three is normalized. We have so far discussed in a monochromatic manner, assuming no wavelength dependency. Since surface materials have dispersion relationships of wavelength and spectral characteristics for light sources, the observer, and digital cameras in our interest, being incorporated the camera response functions of wavelength and having tone-curves, which we assume constant rather than scales in our actual measurement setups throughout this dissertation, the pixel value at a channel $k \in \{1, 2, \dots, K\}$ is

$$v^k(\mathbf{x}) = \int_{\Lambda} f(\mathbf{x}, \alpha, \beta, \gamma; \lambda) E(\lambda) S^k(\lambda) \mathbf{l} \cdot \mathbf{n} d\lambda, \quad (8)$$

where Λ is a wavelength region in interest and S^k is the camera response function of k -th channel. In this equation, we factorized the shading term $\mathbf{l} \cdot \mathbf{n}$ instead of α since the linearized photometric stereo problems prefer this expression.

2.2 Linear photometric stereo problem for Lambertian surfaces

A surface is called Lambertian when $f(\mathbf{x}, \alpha, \beta, \gamma; \lambda) = \rho(\mathbf{x}, \lambda)$, that is the reflectance is independent of local geometry and depends on the material at $\mathbf{x}$ and the wavelength. Then a linear relationship at a certain channel is

$$v(\mathbf{x}) = A \mathbf{l} \cdot \mathbf{n}, \quad (9)$$

where

$$A = \int_{\Lambda} \rho(\lambda) E(\lambda) S(\lambda) d\lambda \quad (10)$$

is the effective reflectance encompassing the light source spectrum and camera response function.

Photometric stereo methods take multiple images as input captured under various light directions $\mathbf{l}_j (j = 1, 2, \dots, J)$. Concatenating the input images and converting into a column vector, we obtain a linear equation

$$\mathbf{v} = A \hat{L} \mathbf{n}, \quad (11)$$

where $\hat{L} = (\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_J)^T$. The surface normal is readily obtained as

$$\mathbf{n} = A^{-1} \hat{L}^\dagger \mathbf{v} \quad (12)$$

minimizing the measurement error or outliers, or as

$$\mathbf{n} = \frac{\hat{L}^\dagger \mathbf{n}}{\|\hat{L}^\dagger \mathbf{n}\|} \quad (13)$$

when reflectance is unknown in advance.

2.3 Extension to nonlinear problems

Many real world objects have non-Lambertian reflectance properties due to their material compositions. Metals are shiny and have specular lobes at points the surface normals of around the half angle between the light and viewing directions. Potteries look either Lambertian or specular depending on the light and viewing directions. Light is directly reflected back into air or refracted into the enamel layer of a pottery and reflected on the sintered powders back into the enamel layer again, and then finally enter our eyes. Since the enamel layer is thin and transparent, the appearance of potteries is the composite of specularities and its diffuse colors.

Extensions to surfaces of more complicated reflectance properties have also been investigated. We only refer to methods that are based on noise/specular-robust approach and so-called photometric invariants [8]. There are still a variety of approaches incorporating explicit BRDF models that represent the reflectance properties of some class of materials, and an extensive review has been recently provided [9].

Specularities are often dealt with as outliers in observations, as opposed to inlier diffuse components. Wu et al. [10] applied low-rank approximation to an observation matrix the columns of which are vectorized shading images captured under different lighting directions. The fact that this matrix has rank-3 is used for eliminating specular lobes, shadows, and noise as outliers, and the resulting rank-3 matrix corresponds to the diffuse components. Ikehata et al. [11] also facilitated this fact, but their formulation is motivated to estimate the diffuse components and outliers found in pixel-wise sparsity that is difficult to be detect in preprocessed low-rank approximation. They solved their sparse regression problem as l1-norm approximation and Bayesian sparse regression. Their method has been shown to be the state-of-the-art of outliers-robust photometric stereo against some class of materials.

Another important approach finds a solvable systems in a quantity called a photometric invariant. The material-intrinsic reflectance value is factored out while the geometric-dependent term only remains in the photometric invariants [8]. Such approaches are independent of the shape of BRDFs and

provides a universal projection from surface geometry onto the surface normal field. The photometric-invariant-based approaches turns out to extract the partial differential equations for surface normal fields from photometric observations. A state-of-the-art method against a assumes a single-lobe reflectance and discusses a solution for a material-independent reflectances [12].

As briefly outlined above, photometric stereo methods have been studied not only for surface reconstruction of Lambertian surfaces but also extended to surfaces of complicated surface reflectance properties including specularities, shadows, noise, or developed photometric-invariant approaches.

Another important drawback of photometric stereo methods is its requiring multiple input images for surface reconstruction: even for the simplest Lambertian surfaces, it requires three or more number of images. This causes estimation error when photometric stereo is applied to deforming objects, since surface normals to be estimated fluctuate while capturing the required number of images. To tackle this problem, time-multiplexing techniques and multi-view augmentation techniques have been proposed. Another approach is to reduce the input image number to one. In his pioneering work, Woodham [4] pointed out that red, green, and blue light sources generate three shading images at once, and that surface normals are determined for gray surfaces. Rigorously, it means that the reflectances at three color channels are the same (or known values). In section 2.5, we review the details of some color photometric stereo techniques.

2.4 Color photometric stereo

Conventional photometric stereo methods use gray-scale images. Color cameras have three or more channels, and it is expected that three or more number of input images under different lighting directions can be obtained in a single color image (Figure 2) as suggested by Woodham [4]. A conventional color photometric stereo problem can be written as

$$\mathbf{v} = \hat{A}\hat{L}\mathbf{n}, \quad (14)$$

where the effective reflectances are now a matrix, the element of which is

$$A_j^k = \int_{\Lambda} \rho(\lambda) E_j(\lambda) S^k(\lambda) d\lambda \quad (15)$$

for $k = 1, 2, \dots, K$ over all the channels and $j = 1, 2, \dots, J$ over all the light sources.

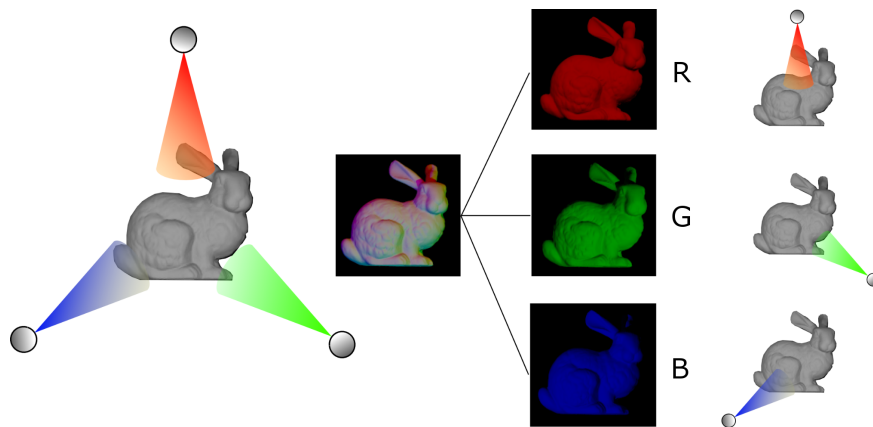


Figure 2: Color photometric stereo.

However, a challenge in turn comes in by having alternated multiple gray-scale inputs with a single color image: the effective reflectance is no longer a scalar but a matrix with KJ -degrees of freedom with 2-degrees of freedom from the surface normal. This gives an ill-conditioned linear problem with K constraints against $KJ + 2$ unknowns.

To modify this ill-conditioned problem to a more approachable problem, some constraints have been considered for the color photometric stereo problem. We give a brief review in the next section.

2.5 Related work in color photometric stereo

Ikeuchi [13] augmented photometric stereo by using multiview stereo, and more recently, Anderson et al. [14] developed a real-time photometric stereo method for objects with a piecewise uniform texture. They first coarsely estimated heights by using binocular stereo and then retrieved fine structures radiometrically. Although their assumption gives a practical scenario, there might be situations in which two cameras cannot be placed, e.g., when observing an object under an optical microscope or when observers need to synchronize movement of devices.

Some other approaches have introduced time-multiplexing in reconstructing surface geometries. Wenger [15] temporally varied illumination and approximated the surface normals to be constant over frames on each point tracked using optical flow. Janko et al. [16] also developed a time-multiplexing method, further adding a regularization term that explains the smoothness of the surface. More recently, Kim et al. [17] have succeeded in reconstructing the fine geometries of moving objects. They elaborately selected LED spectra to be turned on temporally, generating a sequence of images: one for

surface normal estimation, and the others for effective reflectance coefficients. Although their results were quite reasonable and the instrumental setups are feasible to implement, these methods would be difficult to use when the optical flow estimation fails and cannot take into account situations where the color changes over time. Kurihara et al. [18] uniquely applied their correlation image sensor [19] to a single-shot photometric stereo with a specific light configuration. The three degrees of freedom for surface normal and reflectance were retrieved from three temporally correlated outputs, each integrated over a frame period. Although their method is robust due to its proven theory as well as its accurate control and measurement of the amplitude and phase of illumination, the integration over time tends to average the surface normals of a moving object.

Although these photometric stereo methods augmented by multiview stereo, time-multiplexing, or coded exposure have succeeded in reconstructing the surface geometries of moving objects, a common problem is that they are not based on a single capture from a single point of view or point of time. On the other hand, some researchers have tried to reconstruct surface geometry from a single shot by limiting surface color.

In a single-shot method [20] that used structured light composed of different wave vectors of a visible scale, an image was decomposed into images lit from different directions and then applied to photometric stereo with spatial resolution being sacrificed.

Smith and Smith [21] developed a single-shot photometric stereo method active in the near infrared region. They focused on materials such as paints that transmit well near infrared spectra, the characteristic of which they facilitated for treating a textured surface as that of constant reflectances within the spectral region. Moreover, the use of near infrared illumination for photometric stereo has been investigated and suggested to be preferable for human skin [22].

Fyffe et al. [23] presented a method to simultaneously estimate color and surface normals by using a multispectral camera. They regarded the problem as a bilinear optimization on a unit sphere and a subspace spanned by a basis set in the spectral domain. The accuracy seems insufficient when there are few channels and needs to be calibrated with the ground truth data, but as Fyffe et al. suggested, the accuracy would increase as the number of channels increases. While our proposed method resembles their approach in terms of the spectral augmentation, we do not restrict reflectance spectra in a spectral subspace but enable them to be stably estimated by directly measuring a wide part of the spectra. Our first method in Chapter 3 is partly similar to this approach in terms of spectral augmentation.

A long tailed consideration given centrally by Nikolaev, Petrov, Kontsevich, Brill, and Drew has

been suggested that the global geometry, that is some subset of the tangent bundle space, provides a sufficient information on the color and surface normals of a uniformly colored surface region.

Nikolaev [24] presented a classification algorithm based on the concept of the image-field rank. Brill [25] later surveyed Nikolaev’s theory and gave extensions. Note that Nikolaev studied the problem before color photometric stereo was developed. Petrov [26] gave a general classification of surface shape and color but without discussing global color classification capability and the latent classification ambiguities on multi-colored surfaces, which we discuss in Section 4.1. Later, Drew [27] demonstrated through simulations that the surface normals can be retrieved by partly using the information of the known global property. Kontsevich et al. [28] generalized Nikolaev’s theory and applied it to the shape-reconstruction problem (they did not term it as photometric stereo), where they considered the integrability in surface reconstruction to resolve the rotation ambiguity. Instead, Drew and Kontsevich [29] developed a closed-form method to resolve the rotation-transformation ambiguity. In the same period, Petrov and Kontsevich [30] categorized surfaces and their mapped color spaces based on Nikolaev’s theory. Color identification on uniform-color surfaces was succeeded in previous studies [28, 30], but has not been applied to multi-colored surfaces.

More recently, color photometric stereo methods for multi-colored surfaces have been presented. Drew [31] extended them to multi-colored surfaces and retrieved the surface geometry through simulation but relied on a given surface classification. Chakrabarti and Sunkavalli [32] showed that the reflectance of a uniformly colored region is uniquely determined when a reflected light from a light source is observed exclusively in a channel, as we discuss in a different manner, and jointly estimated the surface normals and reflectances of a surface. Despite their elaborate formulation, due to the unknown surface classification, they needed to process each patch by combining polynomial functions [33] that approximate surface geometries and an optimization in the overlaps between the patches. The results seem insufficient in terms of accuracy compared with the standard photometric stereo method.

We found Nikolaev’s image field concept and the numerous discussions in the above studies to be invaluable for surface classification and reconstruction with unknown colors, but neither a basic principle nor a practical method for surface classification and reconstruction for multi-colored surfaces has been provided because current strategies are aimed mainly at uniform color and require searches over every patch with an assumption that the color should be the same over the designated patches: it is necessary to identify at once the whole union of the surface regions of the same color with any

spatial distribution. This is the main difficulty in applying color photometric stereo to multi-colored surfaces. Our second method presented in Chapter 4 is closely related to these succeeding studies and gives a practical surface classification algorithm that also enables the estimation of surface normals from a single color image.

3 Hyperspectral photometric stereo for a single capture

In this chapter, we present a single-capture photometric stereo method that uses a single hyperspectral image [34]. The spectral observation of a surface point is insufficient to determine both of the surface normal and reflectance spectra or color at that point irrespective of the number of channels, since the degree of freedom of the spectral reflectance also increases. Therefore, the reflectances must be estimated under certain geometric and/or chromatic constraints. We assume a chromatic constraint against reflectance spectra in this method.

Most of surface reflectance spectra of real world materials are smooth functions without high-frequency oscillations or discontinuities in the wavelength domain. From this property, it is expected that reflectance values at the wavelengths within a narrow band can be estimated from the functional behavior of reflectance spectra near out of the band. Since the ill-posedness of single-color-image photometric stereo problem comes from the ambiguity between color reflectances and surface normals, separating the processes of reflectance estimation and shading acquisition is effective for resolving the reflectance-shading ambiguity in a single spectral image. This approach is enabled by using a hyperspectral camera instead of an RGB camera due to the abundant spectral information.

Hyperspectral cameras capture 2D images with full-spectral information at every point in the images. Especially, hyperspectral cameras have been developed mainly in remote sensing and artificial satellite, then spread into agriculture, food sanitation inspection, medical imaging, and so on. Spectral imaging beyond RGB color's is advantageous or even necessary in e.g. material classification and medical imaging, and its introduction to image capturing system is becoming an important step toward new technologies. Although there are latent negative aspects for hyperspectral imaging: the trade-off between spatial and spectral resolution and between signal-to-noise ratios (SNRs) and capture speed, improvements and novel capture methodologies and signal processing algorithms have been intensively studied, and high spectral and spatial resolution, capture speed and good SNRs have been achieved and even developing.

Our setup is composed of a hyperspectral camera and a spectrally and spatially designed illumination. The illumination is designed to have a twofold role: a reflectance probe and a resource for independent shading images. The first is to measure and estimate reflectance. Our illumination has wide spectral regions for measuring reflectance spectra, and the remainder narrow bands are used for the second role of obtaining shading images generated under different lighting directions. The lack

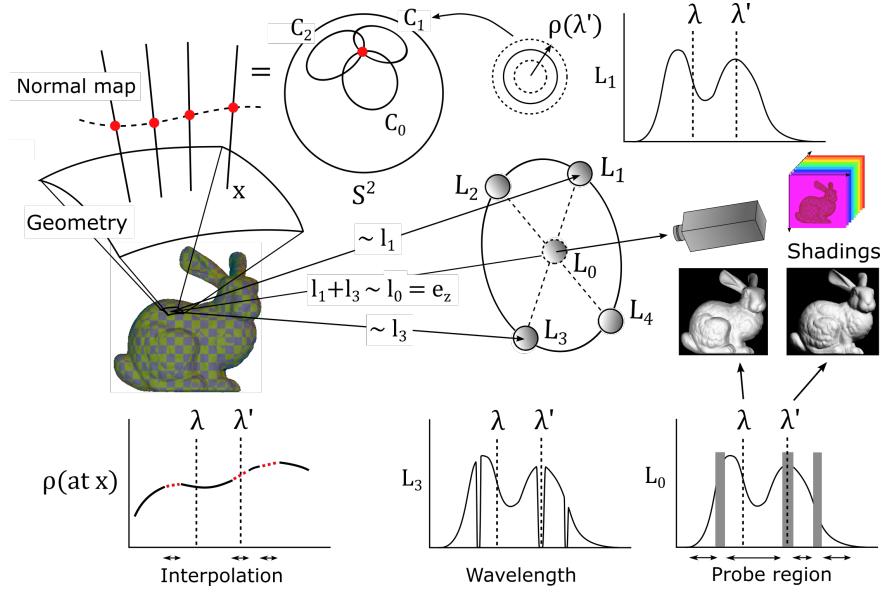


Figure 3: Conceptual schematic of single hyperspectral image photometric stereo (with permission from [34], Journal of the Optical Society of America A).

of reflectance information within a narrow band could be estimated from the functional behavior of reflectance spectra near the band. The multiple shading images are provided by spatially distributed light sources, each of which is identified in one of the narrow bands that corresponds to the lighting from one of the spectrally tuned light sources. We first depict the conceptual schematic in the next section, and then describe the detailed setup and the calculation procedure.

This chapter fully covers the published materials in [34]. The figures and script are included under the permission from the Journal of the Optical Society of America A.

3.1 Method

3.1.1 Overview

Figure 3 shows a conceptual schematic of the method. Let a diffuse surface be in the view of a hyperspectral camera from the e_z direction. we need to obtain three shading images and the reflectances at the wavelengths. To obtain the requisites from a single hyperspectral image, we provide light sources that do not overlap in the wavelength range designated for shading images, and that can work equivalent to a light source as a reflectance probe when restricted in the other wavelength regions.

To provide such an illumination, we prepare four light sources positioned on a circle. Consider a circular board, the center of which is on the optical axis and the normal vector is along this axis. The

light sources L_1 and L_3 are placed centrosymmetrically on the circle, and their spectra are the same in a wide wavelength region, which we call probe region (Figure 3). In a narrow band the spectrum of the light source L_3 is optically notched around a center wavelength λ' , and then only the light source L_1 generates shading on the surface. In the same way, we place two other light sources L_2 and L_4 with a notched band around λ'' , where only the light source L_2 generates shading as the light source L_1 does. Since $\mathbf{l}_1 + \mathbf{l}_3 = \mathbf{l}_2 + \mathbf{l}_4 \parallel \mathbf{e}_z$, the whole light sources works as a probe light source L_0 that measures reflectance spectra in wide spectral regions, virtually positioned in $\mathbf{l}_0 (= \mathbf{e}_z)$ direction. The reflectance spectra can be measured within the nonzero support of the spectrum of this probe light source. The reflectance $\rho(\lambda')$ and $\rho(\lambda'')$ are estimated from curve interpolation of the measured data by assuming that the reflectance spectra have C^1 -smoothness. Figure 3 bottom-left shows the schematic of the interpolation of a reflectance spectrum. Using the reflectance value $\rho(\lambda)$ at a wavelength selected from the nonzero support of L_0 as well as $\rho(\lambda')$ and $\rho(\lambda'')$, three shadings at λ, λ' and λ'' can be applied to shape reconstruction from photometric stereo.

There are other possibility for illumination designs that provide multiple shading information in a single spectral image. In principle, as described above, two pairs of light sources tuned in some wide probe regions and optically notched or added intensities in narrow bands are sufficient to achieve the purpose. A multi-notch-filtered light source and two bandpass-filtered light sources, each center wavelength of which falls exclusively within one of the notched bands, is another option. It is also possible to place a light source as a reflectance probe and bandpass-filtered light sources as shading resources. In this case, however, we should take into account overlaps in the spectral domain between the probing light source and the filtered light sources. The overlaps reduce the slant angles between effective light vectors that are the combinations of the overlapping light vectors, and then cause gaps between the actual light coordination and expected coordinates. It is known that the reduction of the slant angles between the light vectors increase estimation error of surface normals [4, 35, 36]. The proposed illumination design can retain the light vectors as coordinated as the experimenter's discretion by notching narrow bands at different wavelengths against the pair of L_1 and L_3 and that of L_2 and L_4 , while satisfying the equally separated tilt angles between light vectors suggested to be the best by a numerical analysis [36]. For providing the minimum required number of shading resources, we notch a band against each one light source of each light source pair in our setup.

Note that our method is in principle restricted to materials that have nonzero reflectances at least at every two candidate wavelengths from both Λ_{probe} and $\Lambda \setminus \Lambda_{\text{probe}}$. We can select two channels

that have nonzero reflectances, provided that the illumination spectra have multiple notched bands. Although some materials such as red and blue paints have almost zero reflectances in shorter and longer wavelength regions, many materials or chemical compounds often have nonzero reflectances over wide spectral regions. For example, our target includes human skin, the spectrum of which varies among racial groups and health conditions, which requires an accurate measurement of the reflectance spectrum for diagnosis. Regarding application of photometric stereo to skin or medical imaging, some researchers have shown advantages of photometric method against skin surface reconstruction using infrared light [41], in which Sohaib et al. analyzed human skin and found that human skin exhibit diffuse reflection especially for red-to-near infrared light, and endoscopic image analysis [42].

To relieve the limitation on reflectance spectra, we also present an ideal illumination design in Section 3.3.1 and evaluated its effectiveness in simulations in Section 3.3.2-3.

3.1.2 Illumination design

We describe the detailed illumination design and the linear equation from shading with considering illumination spectra.

We define four (normalized) light vectors $\mathbf{l}_1 = (\bar{r}, \phi_1, \bar{z})$, $\mathbf{l}_2 = (\bar{r}, \phi_2, \bar{z})$, $\mathbf{l}_3 = (\bar{r}, \phi_1 + \pi, \bar{z})$, and $\mathbf{l}_4 = (\bar{r}, \phi_2 + \pi, \bar{z})$ for any $\bar{r}$, ϕ_1 , ϕ_2 , and $\bar{z}$ in cylindrical coordinates, which are directed from the object center to the light positions. z -axis is along the optical axis of the camera. Provided that we observe over wavelength region within Λ , $E_j(\lambda)$ ($j = 1, \dots, 4$) represent the spectra of the illuminations that satisfy

$$\begin{cases} E_1(\lambda) = E_3(\lambda), & E_2(\lambda) = E_4(\lambda) & (\lambda \in \Lambda_{\text{probe}}), \\ E_1(\lambda) \neq E_3(\lambda), & E_2(\lambda) \neq E_4(\lambda) & (\lambda \in \Lambda \setminus \Lambda_{\text{probe}}). \end{cases} \quad (16)$$

Figure 4 shows the spatially and spectrally designed illumination in our setup. The top row of Figure 4 illustrates a single pair of light sources for simplicity, and the bottom row shows the full version with two pairs of light sources used in practice. The light source with the spectrum E_1 generates a shading image at the wavelength λ' (blue color), since the light source E_3 is optically notched around the wavelength and no light is emitted from this direction. On the other hand, in Λ_{probe} , both pairs of illuminations; E_1 and E_3 , E_2 and E_4 work together, being equivalent to an illumination on the z -axis unless any of them suffer from occlusions. Although E_1 and E_2 can be determined arbitrarily, we assume that they have nonzero and zero intensities with no overlaps. For an actual setup, this

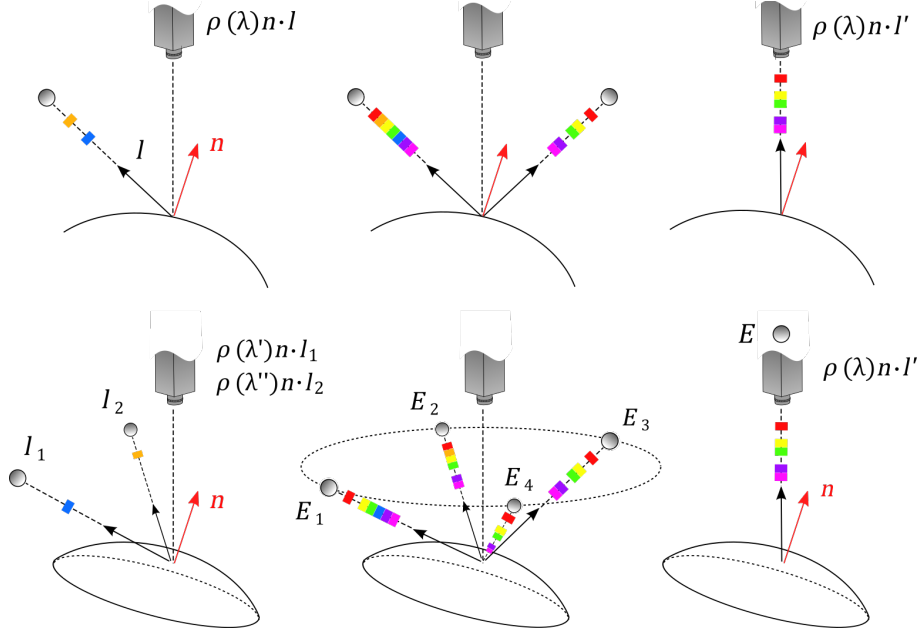


Figure 4: Illumination design (top row: single pair of light sources; bottom row: full version with two pairs of light sources) for the simultaneous acquisition of multiple shadings (first column) and reflectance spectra within the probe region (third column).

facilitates tuning of illumination spectra.

Since we consider diffuse surfaces, the observed intensity at a wavelength λ is

$$v(\mathbf{x}; \lambda) = \Phi_{\lambda} \left[\sum_j \mathbf{n}(\mathbf{x}) \cdot \mathbf{l}^j \rho(\mathbf{x}; \lambda) E_j(\lambda) \right], \quad (17)$$

where $\mathbf{n}(\mathbf{x})$ is the surface normal vector, $\rho(\mathbf{x}; \lambda)$ is the reflectance spectrum at the surface point and $\Phi[\bullet]$ is a nonlinear tone reproduction function at each wavelength that relates the incident number of photons to the pixel value, in accordance with the exposure time. We confirmed that Φ is linear and omitting the constant coefficient, by taking spectral data with a white reference over a range of exposure time; then Equation (17) becomes

$$v(\mathbf{x}; \lambda) = \rho(\mathbf{x}; \lambda) \sum_j \sigma^j(\mathbf{x}) E_j(\lambda). \quad (18)$$

$\sigma^j(\mathbf{x}) = \mathbf{n}(\mathbf{x}) \cdot \mathbf{l}^j$ is the shading, a geometric term that depends only on the surface geometry and the lighting directions.

From the light source positions, giving $\mathbf{l}_1 + \mathbf{l}_3 = \mathbf{l}_2 + \mathbf{l}_4$, $E_1(\lambda) = E_3(\lambda)$ and $E_2(\lambda) = E_4(\lambda)$ ($\lambda \in$

Λ_{probe}), Equation (18) becomes

$$v(\mathbf{x}; \lambda) = \rho(\mathbf{x}; \lambda) \sum_{j=1,2} \tilde{\sigma}^\bullet(\mathbf{x}) E_j(\lambda) = \rho(\mathbf{x}; \lambda) \tilde{\sigma}^\bullet(\mathbf{x}) E(\lambda) \quad (\lambda \in \Lambda_{\text{probe}}), \quad (19)$$

where $\tilde{\sigma}^\bullet \equiv \tilde{\sigma}^j$ ($j = 1, \dots, 4$) and $\tilde{\sigma}^j \equiv \sigma^j + \sigma^{j'}$ ($\{j, j'\} = \{1, 3\}, \{2, 4\}$) are independent to the light position labeled by j , and $E(\lambda) = E_1(\lambda) + E_2(\lambda) = E_3(\lambda) + E_4(\lambda)$.

Surface reflectance spectra can be directly obtained within the probe region Λ_{probe} under this illumination that is designed to have an equivalent single light source within the wavelength domain.

3.1.3 Reflectance estimation

The reflectance spectra within Λ_{probe} can be obtained by simply dividing the observed intensities by light source intensities, since the geometrical factor $\tilde{\sigma}^\bullet$ is independent of the light positions. The reflectance spectrum at point $\mathbf{x}$ within the probe region Λ_{probe} is

$$\rho(\mathbf{x}; \lambda) = v(\mathbf{x}; \lambda) / (\tilde{\sigma}^\bullet(\mathbf{x}) E(\lambda)) \quad (\lambda \in \Lambda_{\text{probe}}). \quad (20)$$

$v(\mathbf{x}; \lambda)$ and $E(\lambda)$ are known in advance, while $\tilde{\sigma}^\bullet$ is an unknown constant value. Here, we can rescale the reflectance to one at $\lambda_{ch(1)}$; then, the geometrical factor disappears and the reflectance spectrum at $\mathbf{x}$ is given by

$$\rho(\mathbf{x}; \lambda) = v(\mathbf{x}; \lambda) / E(\lambda) \quad (\lambda \in \Lambda_{\text{probe}}). \quad (21)$$

We next estimate complementary spectral curves within $\Lambda \setminus \Lambda_{\text{probe}}$ by curve interpolation. Provided that reflectance spectra are C^1 -smooth functions without rapid fluctuations, as the visible spectra of many materials are [37–39], each full spectrum is expected to be estimated by interpolation. Although the short number of observations less than the degree of freedom remains the problem still intractable, the full-spectral information has provided a reasonable restriction on the functional space where reflectance spectra exist, and we approximately rely on C^1 space. The accuracy will be increased as the filtered bands in $\Lambda \setminus \Lambda_{\text{probe}}$ become narrower. Throughout the simulations and the experiment, we recovered full spectra by piecewise cubic Hermite polynomial interpolation in Matlab.

Another option to recover the full spectra is to restrict the reflectance spectra within a spectral subspace spanned by a reflectance basis set [37–39]. This approach is widely used and shown ef-

fective to represent reflectance spectra with a small number of basis sets, having been attempted for photometric stereo [23, 40] as well as the other computer vision fields. We will also discuss this issue on this approximation when used in surface reconstruction, our specific interest, in Chapter 5.

3.1.4 Surface normal estimation

After estimating the reflectance spectrum at each surface point, we next reconstruct the surface normals. From Equation (18), we obtain

$$\mathbf{v}(\mathbf{x}) = \hat{\rho}(\mathbf{x})\hat{E}\boldsymbol{\sigma}(\mathbf{x}), \quad (22)$$

where $(\mathbf{v})_k = v(\lambda_k)$ is the observation at k -th wavelength, $(\boldsymbol{\sigma})^j = \sigma^j$ is the unknown parameter that includes the information about the surface normal vector, $\hat{E}$ is the illumination matrix of each element being $(\hat{E})_j^k = E_j(\lambda_k)$, and $\hat{\rho} = \text{diag}(\rho(\lambda_1), \dots, \rho(\lambda_K))$ is the reflectance matrix with the size of the number of acquired wavelengths K . The geometrical factor is solved as

$$\boldsymbol{\sigma}(\mathbf{x}) = \hat{\rho}(\mathbf{x})^{-1}\hat{E}^\dagger\mathbf{v}(\mathbf{x}), \quad (23)$$

from which the surface normal vector is given by

$$\mathbf{n}(\mathbf{x}) = \frac{\hat{L}^\dagger\boldsymbol{\sigma}(\mathbf{x})}{\|\hat{L}^\dagger\boldsymbol{\sigma}(\mathbf{x})\|}, \quad (24)$$

where $\hat{L} \equiv (\mathbf{l}_1, \dots, \mathbf{l}_4)^\text{T}$ is the light vector matrix and $\hat{L}^\dagger$ is the psuedo-inverse matrix. The scale ambiguity of reflectance spectra is removed by the normalization. In practice, four images are sufficient, selected at two different wavelengths with nonzero illumination intensities each from Λ_{probe} and $\Lambda \setminus \Lambda_{\text{probe}}$. Then from $(\hat{E})_j^q = E_j(\lambda_{ch(q)})$ ($q = 1, \dots, 4$), Equation (23) becomes $\boldsymbol{\sigma}(\mathbf{x}) = \hat{\rho}(\mathbf{x})^{-1}\hat{E}^{-1}\mathbf{v}(\mathbf{x})$. Note that although three images are also sufficient if we replace $\mathbf{l}_3$ and $\mathbf{l}_4$ with $\mathbf{e}_z$, we consider all the four light sources explicitly as a more general expression (e.g. the four light sources generate independent shadings when four different bands are filtered).

3.2 Simulation

We evaluate the estimation accuracy of our method using noise-free synthetic hyperspectral image. We consider two types of illumination: comb spectrum and filtered spectrum, the former can measure reflectance spectra that have zero reflectance bands. Such an illumination can be realized by using a spectral synthesizer [43]. The latter illumination design can be achieved only by using optical filters, and is easily implemented. Although this illumination design cannot be applied to locally nonzero reflectance spectra, it can be applied to a variety of materials and recover smooth, widely nonzero reflectance spectra.

We rendered hyperspectral images with the rendering software pbrt-v2 [44]. For the specific purpose of spectral rendering, we slightly modified the software so that it can output hyperspectral images. Throughout the simulations, the spectra of light sources were calibrated with a white reference that has a constant reflectance over a visible spectrum. Note that the unit of the coordinates is arbitrary since we do not consider the effects of diffraction and other electric scale effects.

3.2.1 Comb illumination

3.2.2 Setup

To quantitatively evaluate the accuracy of our method, we target a colored sphere of a spatially uniform color by varying the surface reflectance spectrum. For this purpose, we prepared a set of uniformly colored spheres with different reflectance spectra and tested them one by one. The sphere was set sufficiently small (0.75) so that fitted within the image window, and rays emitted from a point light source were regarded as parallel. The surface reflectance spectra were chosen from the Macbeth Color Checker given in a pbrt file.

We designed a set of comb-shaped illumination spectra: (L_1) a base spectrum with the period of 24 nm, (L_2) a spectrum shifted by a half period without overlaps against the base spectrum (the term "base" is relativistic), (L_3) a partially notched base spectrum and (L_4) another shifted and notched spectrum (Figure 5 (a)). We set the degree of notch to 0.2, which was optional if it was small enough, against the normalized intensity. To be able to deal with the localized reflectance spectra, we selected the notch wavelengths to fall around blue, green and red visible regions. No two notched bands in L_1 and L_2 or L_3 and L_4 are next to each other, otherwise interpolations of reflectance spectra would be less reliable. Depending on the narrow peaks of the illumination spectra, we set the spectral resolution

of a camera to 1.2 nm, within which the signals were averaged. Reflectance spectra were measured every 12 nm from 400 nm to 700 nm, and then the gaps in the notched bands were interpolated over 24 nm (Figure 5 (b-2)).

We describe the scene by a Cartesian coordinate system with the origin at the position of the camera. Every colored sphere was placed at $(0, 0, -200)^T$. The light sources were placed at $(100, 100, 0)^T$, $(-100, -100, 0)^T$, $(100, -100, 0)^T$ and $(-100, 100, 0)^T$ for L_1 , L_2 , L_3 , and L_4 , respectively (Figure 5 (b-1)). The camera was set to orthographic projection, and the image size was set to 128×128 .

3.2.3 Results and discussion

Figure 5 shows (c-1) the rendered input image in the RGB format, (c-2) the estimated surface normal map, (c-3) the error map against the ground truth, and (c-4) the reconstructed surface geometry with the ground truth, each for a sphere of Macbeth-1 color (Macbeth- i is referred to as the i -th index in the color checker). The surface geometries were reconstructed by using the Frankot-Chellappa algorithm [45]. The results against the other reflectance spectra are almost the same except for Macbeth-15. The increasing error of surface normals in the neighborhood of the rim of the sphere is due to occlusions from the farthest illuminations in our setup. This is a limitation of photometric stereo, possibly being relaxed in specific situations by taking a large number of images under widely distributed light sources or by increasing the number of light sources in our case. Our method is directed to be simply able to reconstruct surface geometries that are not occluded from any light source. In Figure 5 (c-4), the geometry reconstructed from the surface normals within the unoccluded area corresponded well with the ground truth.

The Comb column in Table 1 lists the mean and maximum errors in surface normals and reflectance over the wavelength region of interest. The error in reflectance at λ was evaluated by

$$100 \frac{|\rho(\lambda) - \rho^*(\lambda)|}{\rho^*(\lambda)}, \quad (25)$$

where $\rho(\lambda)$ and $\rho^*(\lambda)$ are the estimated reflectance and the ground truth at λ , respectively. Estimation was accurate enough except for a sphere of Macbeth-15 color. As shown in Figure 6, all the reflectance spectra except for the Macbeth-15 spectrum were estimated well, including within the notched bands. The reflectance spectra were evaluated at the top of the spheres. The results suggest that the notched bands distributed with moderate gaps enable appropriate channels to be adaptively

Table 1: Mean and maximum errors in surface normals and reflectances against sphere of reflectance spectra from Macbeth Color Checker (No.1-24), for comb illumination and filtered illumination (with permission from [34], Journal of the Optical Society of America A).

No.	Comb				Filter			
	Normal (°)		Reflectance (%)		Normal (°)		Reflectance (%)	
	Mean	Max	Mean	Max	Mean	Max	Mean	Max
1	.993	3.32	1.01	1.93	.601	3.01	.286	1.53
2	.992	3.24	.573	3.66	2.35	5.07	.336	2.17
3	1.26	3.63	.351	1.06	.930	3.41	.274	.920
4	1.49	4.60	1.17	9.26	2.16	5.22	.611	4.00
5	1.58	4.15	.546	1.66	3.49	7.20	.730	5.37
6	1.60	4.00	.525	1.57	3.03	5.89	.567	3.59
7	.950	3.21	.920	3.17	1.34	3.69	.198	1.49
8	1.06	3.59	.760	2.00	1.68	4.58	.651	3.43
9	2.31	5.41	1.29	8.55	3.07	6.34	.827	5.53
10	.930	3.22	.766	2.13	2.45	5.05	.891	4.80
11	1.47	3.91	.913	4.95	1.41	4.23	.590	3.67
12	.848	3.11	.767	3.10	.660	2.70	.322	1.76
13	2.06	4.58	1.42	2.93	5.37	10.1	1.52	9.51
14	.960	3.28	1.22	6.13	3.40	6.25	.870	4.38
15	4.48	8.21	1.35	9.88	4.01	8.16	1.04	6.69
16	.903	3.19	.730	4.43	8.57	15.2	1.40	14.4
17	1.74	4.32	.964	4.40	2.39	5.10	.710	2.22
18	.990	3.26	.933	2.14	2.31	4.87	.335	1.81
19	.962	3.28	.369	1.46	1.41	3.82	.679	1.20
20	1.07	3.43	.180	1.74	.863	3.05	.287	.446
21	.902	3.19	.146	.991	.978	3.20	.133	.320
22	1.01	3.30	.119	.621	.679	2.91	.184	.942
23	1.25	3.55	.257	1.02	.723	2.86	.239	1.29
24	.979	3.30	.261	1.91	2.16	4.84	.255	2.55

selected. For Macbeth-15 as shown in Figure 8 (a), the slope over 600 nm is steep and the reflectances were small around the wavelength, resulting in the large relative error of estimated reflectances and the surface normals.

3.2.4 Filtered illumination

3.2.5 Setup

To evaluate the potential accuracy of a feasible preliminary experimental setup, we examine another illumination design similarly prepared in the actual experiment. Figure 7 (a) and (b) show a schematic of the simulation setup and the illumination spectra.

Each light source has a broad spectrum and one or two filters: (F1) a shortpass filter, (F2) a

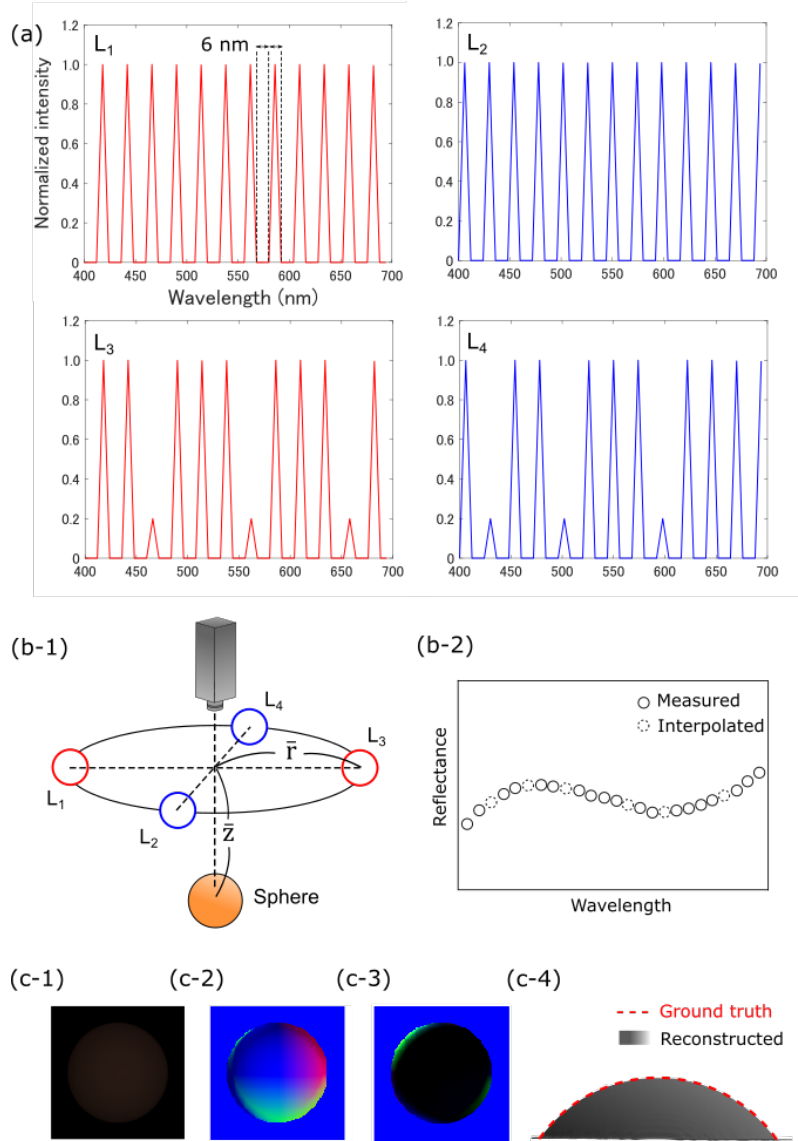


Figure 5: (a) Comb illumination spectra. (b-1) Schematic of setup, and (b-2) probe region (solid circles) and interpolated region (dotted circles). (c-1) Rendered image, (c-2) normal map, (c-3) residual error map, and (c-4) section of reconstructed shape (gray area) against ground truth (red dotted line), each for a sphere of uniform color of Macbeth-1 reflectance spectrum (with permission from [34], Journal of the Optical Society of America A).

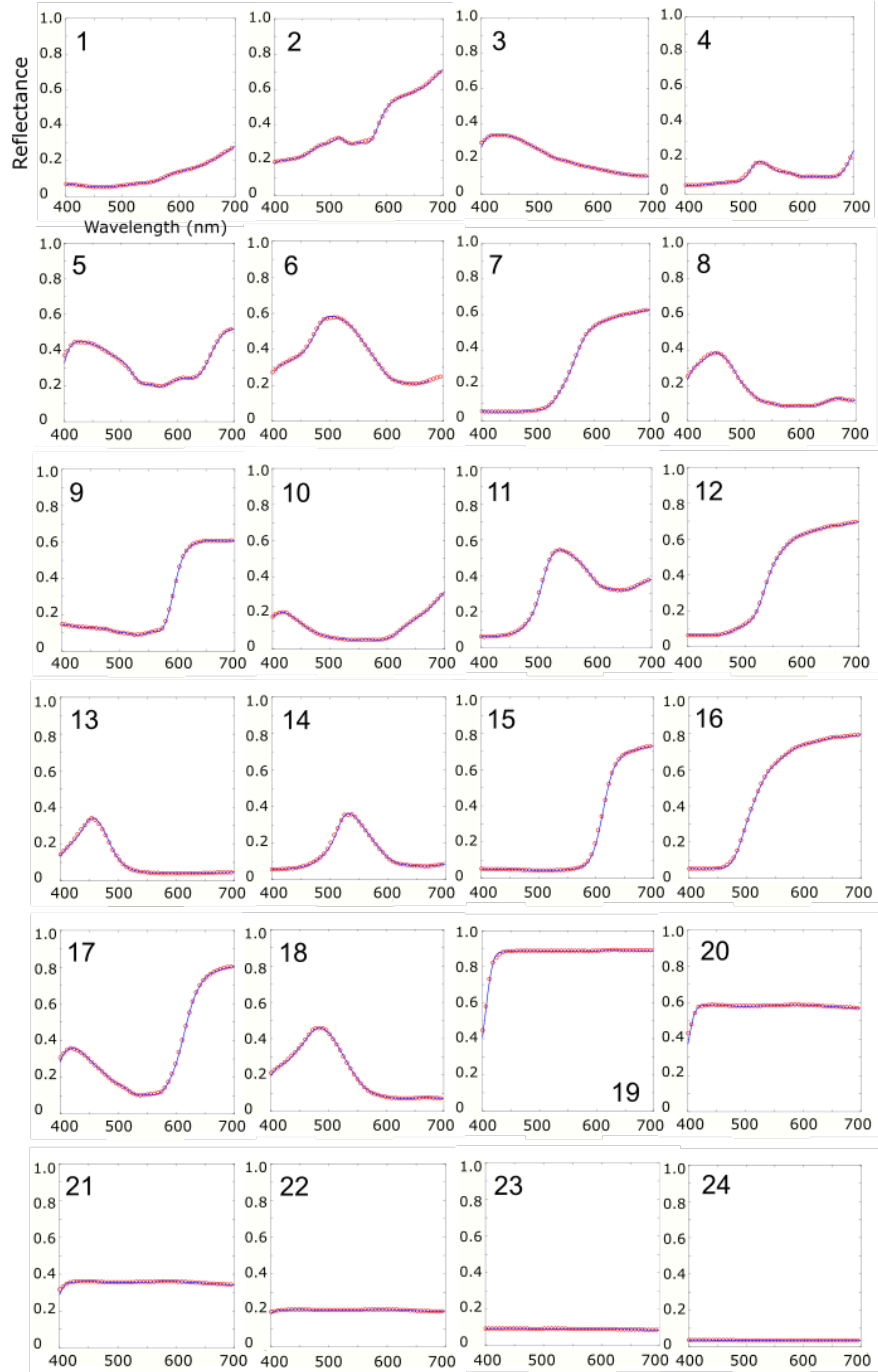


Figure 6: Estimated reflectance spectra at top of spheres (red circles) and ground truths (blue solid lines) (with permission from [34], Journal of the Optical Society of America A).

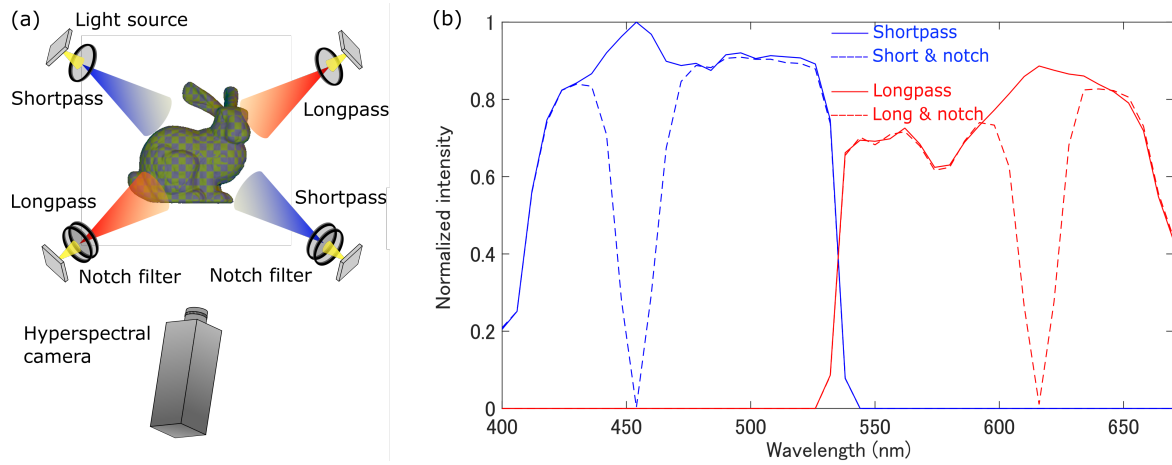


Figure 7: (a) Schematic of setup and (b) illumination spectra: shortpass-filtered (blue solid), shortpass/notch-filtered (blue dot), longpass filtered (red solid) and longpass/notch-filtered (red dot) spectra (with permission from [34], Journal of the Optical Society of America A).

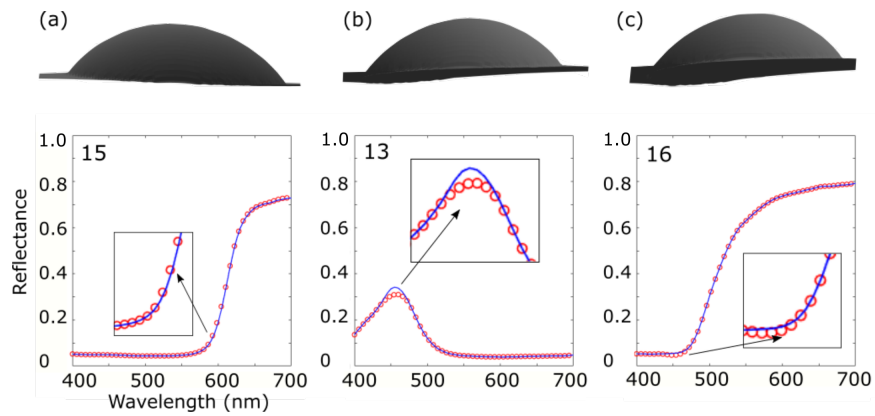


Figure 8: Reconstructed shapes in top row and estimated reflectance spectra (red circles) against ground truths (blue-solid line) in bottom row for spheres of Macbeth-15 (using comb illumination), 13 and 16 (using filtered illumination) (with permission from [34], Journal of the Optical Society of America A).

longpass filter, (F3) a shortpass and a notch filter, and (F4) a longpass and another notch filter. The shortpass and longpass filters divide the broad light spectrum into "bluish" and "reddish" spectra (see the schematic in Figure 7(a) and the photographs in Figure 10 (d,e)), and the two notch filters cut narrow bands off within the spectral regions. We set the daylight spectrum in a pbrt file as the light source and additionally designed a longpass filter, a shortpass filter, and two notch filters. The light sources are placed at $(100, 100, 0)^T$, $(-100, -100, 0)^T$, $(100, -100, 0)^T$, and $(-100, 100, 0)^T$ for shortpass, shortpass/notch, longpass, and longpass/notch-filtered ones, respectively. The setup of the camera was the same as for the comb illumination.

The shortpass-filtered and shortpass/notch-filtered spectra have the same intensities outside of the notched spectral region, where we selected a channel $\lambda_{ch(2)} = 502$ nm (see 3.1.4). The shortpass-filtered spectrum has a nonzero intensity at $\lambda_{ch(1)} = 454$ nm, while the shortpass/notch-filtered one has almost zero intensity at the wavelength. The longpass-filtered and the longpass/notch-filtered spectra also have the same intensities outside of the other notched spectral region, where we also selected $\lambda_{ch(3)} = 574$ nm. At $\lambda_{ch(4)} = 616$ nm, the longpass-filtered spectrum has a nonzero intensity, while the longpass/notch-filtered spectrum has almost zero intensity.

As in the case of the comb illumination, we evaluated the mean and maximum errors in surface normals and reflectances for spheres of a spatially uniform color by varying the reflectance spectrum.

Moreover, to emphasize that our method achieves a point-wise estimation, we performed another simulation, where a Stanford bunny with a bicolor texture was placed at $(0, 0, -200)^T$ in the scene. The reflectance spectra of the texture were those of Macbeth-5 and Macbeth-11. The camera was set to orthographic projection and the image resolution was set to 256×256 .

3.2.6 Results and discussion

The Filter column in Table 1 lists mean and maximum errors in surface normals and reflectances over the wavelength region of interest. The no.*i* represents Macbeth-*i*. Other than Macbeth-5, 9, 13, 15, and 16, the mean normal errors and maximum reflectance errors were small. Also, the surface geometries were retrieved well for the well-estimated reflectance spectra. The reconstructed spheres for Macbeth-5, 9, and 15 were distorted to as small an extent as in the comb illumination. For Macbeth-13 and 16, the reconstructed spheres were more distorted (the top row in Figure 8 (b, c)). The surface geometries of the spheres of all the other reflectance spectra were retrieved well. The large errors in the maximum

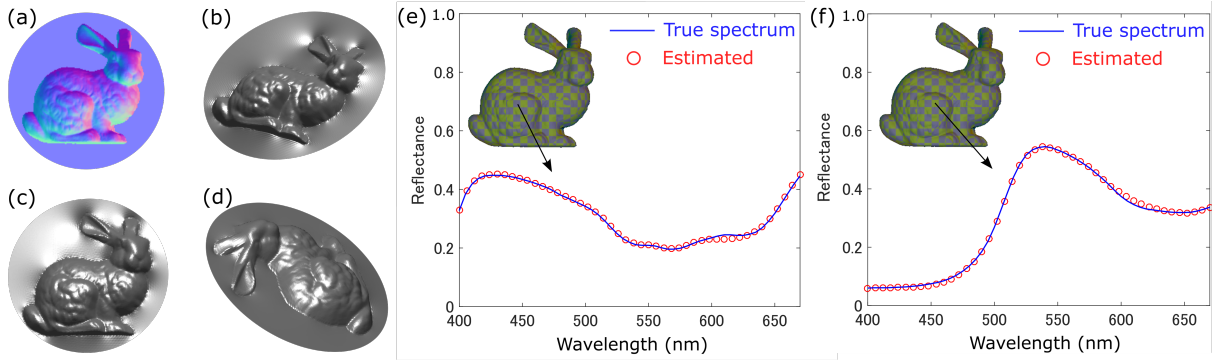


Figure 9: (a) Estimated normal map, (b-d) reconstructed geometries from new views, and (e, f) estimated reflectance spectra with ground truths of Macbeth-5 and 11 (with permission from [34], Journal of the Optical Society of America A).

reflectance errors for Macbeth-16 can be attributed to the steep change in the reflectance spectrum around the wavelength with a small reflectance. As shown in the bottom row in Figure 8, there were larger relative differences between the interpolated spectra and the ground truths.

The results for the scene of the textured bunny are shown in Figure 9: (a) the estimated surface normal map, (c) the reconstructed geometry seen from the camera position, (b, d) geometries from new points of view, and (e, f) the reflectance spectra of the two color patches, and (e) Macbeth-5 and (f) Macbeth-11 patches. The surface normal map, the fine geometrical structures, and the reflectance spectra were retrieved well from a single hyperspectral image. Like in most photometric stereo methods, the effect of interreflections between locally convex structures impairs the reconstruction. With the quantitative analyses against spheres, these results indicate that our method works correctly with the filtered illumination. Note that the RGB image shown in Figure 9 (e, f) retains the sensuously natural appearance of the texture. An advantage of our method is that the lighting environment does not affect the appearance of the eyes, unlike time-multiplexing methods, especially when used in medical imaging analyses, where in situ human perceptivity and intuition become important.

3.3 Real scene

In this section, we demonstrate the effectiveness of the proposed method using a real scene by generating a hyperspectral image of an object under a spectrally designed illumination. In our setup, hyperspectral images have serious noise due to a low illumination intensity, and this makes image denoising difficult. We therefore targeted an object of a uniform color in this study, which facilitates

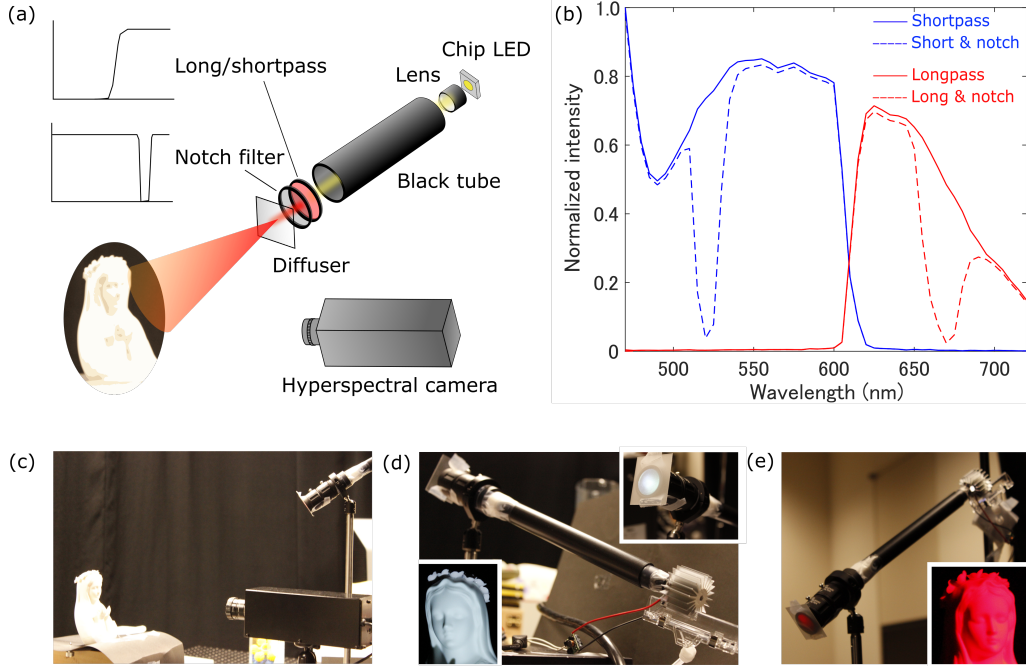


Figure 10: (a) Schematic of experimental setup, (b) measured illumination spectra, and (c) photograph of experimental environment. Illumination equipment and lit up images under (d) shortpass-filtered illumination and (e) longpass-filtered illumination (with permission from [34], Journal of the Optical Society of America A).

image denoising. Experiments against textured objects will be conducted in our future work.

3.3.1 Experimental setup

Figure 10 (a) and (c) show a schematic and a photograph of our system. We placed a statue of Maria at the origin of a Cartesian coordinate system, a hyperspectral camera NH7 (EBA Japan Co.Ltd.) on the z-axis, a longpass-filtered illumination at $(20, 20, 40)^T$, a longpass/notch-filtered illumination at $(-20, -20, 40)^T$, a shortpass-filtered illumination at $(-20, 20, 40)^T$, and a shortpass/notch-filtered illumination at $(20, -20, 40)^T$, in centimeters. The spectral resolution of the camera was 5 nm.

In the actual procedure, we generated a single hyperspectral image by adding four hyperspectral images captured under each illumination lit one by one because not all the light sources could be fixed on our work table at the same time. Each position of the light sources was determined by the position of a supporting pole shown in Figure 10 (c-e), adjusted so that the light vector was directed from Maria's face to the light source.

We used an NS3W183H1 LED chip (Nichia Corporation) as a light source with a broad spectrum over the visible region (Figure 10 (b)), lit with a constant current of 340 mA and a forward voltage of

12 V. Emitted light was collimated using a convex lens and a black tube guiding to the clear aperture of the filters with near-perpendicular incidence. The collimated light was spread for a spatially uniform lighting. For this purpose, we used a microlens array optical diffuser with a 20 degree spread (OSC Japan). Additionally, difficulty can also be reduced by using the diffuser that the spectrum of the collimated light is not spatially uniform due to the distribution of the fluorescent paints in the chips.

Figure 10 (b) shows the illumination spectra measured using a white reference board and the camera. To achieve these illuminations, we used a shortpass filter and a longpass filter, the band edges of which cross over between 605 nm and 625 nm, and two notch filters, the center wavelengths of which are around 520 nm and 670 nm with the notched bands of 35 nm and 45 nm wide, respectively (Thorlabs Inc.). These notch wavelengths are suitable for measuring the reflectance spectra e.g. of human skin and blood vessels that exhibit moderate spectral changes within the notched bands.

3.3.2 Hyperspectral image denoising

Removing noise from hyperspectral images is important in hyperspectral image processing. Due to recent advances in sparse modeling [46], we can resort to an approach based on the linear mixing model: in real-world scenes, constituent materials are pure or mixed chemical compounds, which have a unique or linear combinations of specific reflectance spectra. A hyperspectral image can be treated as a matrix of columns and rows that represent vectorized spatial and spectral data, respectively. When interreflections and multi-scattering effects can be ignored, the rank of the matrix has an upper bound number of the reflectance spectra specific to the constituent materials, and thus is often much lower than the number of rows of the matrix.

We recovered a low rank matrix using the SSGoDec algorithm [47], as in a recent study [48], except that the GoDec algorithm was applied in [48] because the SSGodec algorithm performed better in our tests. Unlike [48], we did not divide pixels into pixel blocks.

Given a matrix X , in which each i -th column represents a vectorized image of the i -th band, we solve

$$\min_{L,S} \|X - L - S\|_F^2 + \mu \|S\|_1 \quad \text{s.t.} \quad \text{rank}(L) \leq r, \quad (26)$$

where L and S are the low rank matrix and the sparse matrix, and $\|\cdot\|_F$ and $\|\cdot\|_1$ denote the Frobenius norm and the ℓ_1 -norm, respectively. The ℓ_1 -norm is a convex relaxation of the ℓ_0 -norm, the cardinality of nonzero elements. The algorithm alternatively minimizes the norm functions by bilateral random

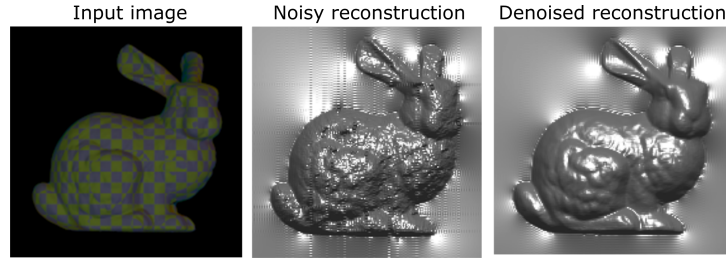


Figure 11: Surface reconstruction with and without denoising.

projections [49] augmented by power iterations [50] for L and a soft-thresholding algorithm for S [51]. A bilateral projection of a rank r gives a rank- r approximation of X [49]. We obtain a denoised hyperspectral image by rearranging L in a sequence of monotone images.

In practice, since each image contains too many spike lines that arise from the low illumination intensity for recovering a low rank matrix when they are added into an image, we recovered low rank matrices of four input images taken under different lighting directions and then took a sum. Note that this problem would not occur when an image under all light sources is captured at once. We set the rank as $r = 1$ and the coefficient $\mu = 1$ and $power = 3$ for the power iteration for each matrix as the best values found in our tests. The computation finished within a minute for each image.

Figure 11 shows the surface reconstruction from the synthetic bunny image with and without denoising. In contrast with the results using the noisy image, the surface was well reconstructed after image denoising.

3.3.3 Results and discussion

The first four columns in Figure 12 (a) show the raw input images (top row) and the denoised images (second row) for the selected channels. Although the raw images were heavily corrupted by noise, the denoised images were clear enough as inputs for photometric stereo. Note that, at the notched wavelengths $\lambda_{ch(1,4)}$, the shadings shown in the second row in Figure 12 (a) are the same as those lit by the top left light source or the top right light source. On the other hand, at $\lambda_{ch(2,3)}$ as well as all the other wavelengths (not shown), the shadings are equivalent to ones lit in front.

Figure 12 (b-1) shows the estimated reflectance spectrum at a point of the object. Compared with the ground truth spectrum, which was measured using the same camera and a white reference and arranged the same at 600 nm, the reflectance spectrum was estimated accurately below 600 nm, while it exhibited almost a scale difference above the wavelength. We conducted the same experiment

twice under the same conditions and obtained almost the same results. The measurement errors are probably attributed to the experimental procedure and the camera response, where we took images individually, especially below and above around 600 nm. To correct the spectrum, we compared the reflectances obtained below and above 600 nm and then rescaled them at 600 nm. The corrected reflectance spectrum is shown in Figure 12 (b-2). This problem will be resolved by capturing a single hyperspectral image by turning on all the light sources at once. To realize this condition, we need two sets of two white LEDs of almost the same illumination spectrum. Fortunately, LED chips of broad spectra that can be tuned precisely have been developed for commercial use, including the LED chips used in this study. Fiber optic illuminators can also be used.

The two rightmost columns in Figure 12 (a) show the surface normal map in a pseudo color and the reconstructed shape for the raw input data (top row) and for the denoised data (second row). Although the normal map and the surface geometry estimated from the raw data were corrupted by noise, a clear normal map and the surface geometry were retrieved well from the denoised data. (c-1) to (c-4) show the surface geometry from new points of view. The details like the hollows around the eyes, the nose, the lips and the chin on the Maria's face and the hair and the wimple on the head were clearly retrieved. The geometries around the neck and the rim of the face were depressed due to the occlusion from some light sources.

3.3.4 Effect of notched-band on spectral estimation

In the simulations and experiment we fixed the notched bands in our illumination designs, but the accuracy of full-spectral estimation by interpolation must vary depending on the notched bands. We investigated the effect of the notched bands on the estimation accuracy in simulations. We tested on 60 water paints and 91 oil paints, the reflectance spectra of which were measured from 380 nm to 730 nm with 10 nm steps, and a band of tough situated 40 nm width was notched at every interpolation test with varying the center wavelength by 20 nm.

Table 2 shows the resulting mean and maximum errors in the interpolated spectra at each center wavelengths (CWs). The interpolated spectra retrieved the original ones better in the longer wavelength region, especially for the water paints. The oil paints are also characterized by a similar behavior, but less accurate were the interpolation results. This is what Smith and Smith [21] have noticed the well-known property of materials used for colorants that the reflectance spectra become less sensitive

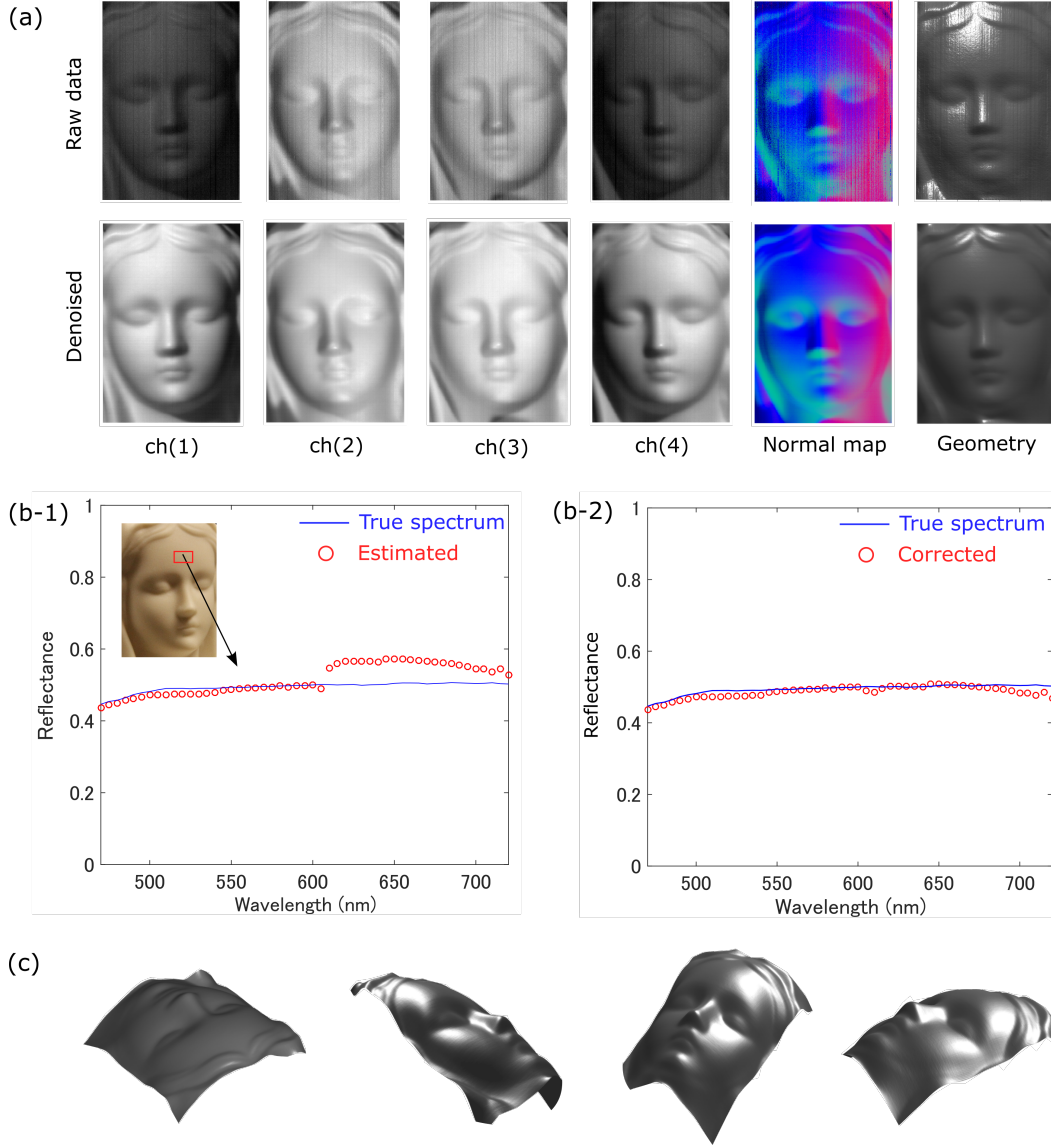


Figure 12: (a) From left to right, images at selected wavelengths: $ch(1) = 520$ nm, $ch(2) = 570$ nm, $ch(3) = 640$ nm, and $ch(4) = 670$ nm, estimated normal map and reconstructed geometry: raw input data in first row and denoised data in second row. (b-1) Estimated reflectance spectrum against ground truth, (b-2) corrected spectrum, and (c) reconstructed geometry from new views (with permission from [34], Journal of the Optical Society of America A).

to wavelength in longer wavelength regions. The results suggest that longer wavelengths are better to be notched in the illumination design. If available, near-infrared regions are also suitable for notched bands.

Despite this result, full-spectral estimation using interpolation is insufficient in the worse case that we have no expectation for the spectral behavior. We will consider possible solutions to alleviate the estimation failure of reflectance spectra in Chapter 5.

Table 2: **Interpolation test results with a notched-band variation.**

CW (nm).	Water paints		Oil paints	
	Reflectance (%)		Reflectance (%)	
	Mean	Max	Mean	Max
400	6.82	49.8	13.9	56.3
420	1.83	12.5	4.95	72.2
440	1.72	8.25	5.16	60.8
460	8.71	101	9.08	68.6
480	9.48	86.1	14.0	180
500	7.08	55.0	11.3	152
520	3.55	28.4	8.63	34.4
540	3.54	31.8	8.30	75.0
560	5.23	65.9	12.4	118
580	4.71	36.7	10.5	109
600	3.66	7.76	7.76	72.1
620	1.90	15.0	3.07	21.9
640	1.14	5.69	2.88	25.2
660	0.878	3.19	2.86	35.0
680	0.311	1.26	2.81	41.0
700	0.811	4.88	3.02	24.0

3.4 Conclusion

We have presented a single-capture photometric stereo against arbitrary surface colors by using a hyperspectral camera. The spectrally and spatially designed illumination worked as a reflectance probe and shading resources. Full reflectance spectra were estimated by interpolation with a smooth reflectance assumption, and we analyzed the accuracy of our method in simulations and conducted an experiment against a real scene. We also investigated the optimum notch wavelengths with real data and simulated interpolation.

4 Single color image photometric stereo for multi-colored surfaces

We have presented a method that estimate the reflectance spectra and surface normals of a colored Lambertian surface from a single hyperspectral image. Despite its effectiveness against arbitrary colors with moderate conditions, there still remain practical challenges on capturing a hyperspectral image: slow capture speed by e.g. line-scans, relatively high SNRs, and high cost.

Most of conventional color photometric stereo methods use RGB cameras with elaborating illumination, using multiple camera setups or inter-frame constraints. Since RGB cameras are easily available and generally have low SNRs and high FPSs, developing a single-color-image based method solves the practical difficulties on the usage of hyperspectral cameras and makes photometric stereo a more effective tool in suitable applications. However, unlike there being assumed interchannel correlations in a hyperspectral data, an RGB image has less correlations between different channels: provided that red and blue reflectance values were known, the green reflectance cannot be estimated in general. Thus, we focus not just on the color at a surface point but on the global color distribution on the surface, which is expected to have information of the surface normals and colors, embedded in the color space. We consider conditions on color and geometry, under which the true color classification of a surface and accurate surface normals are obtained. There has been a small number of single color-image based photometric stereo method that recover sufficiently accurate geometries against multi-colored surfaces of unknown reflectances.

In this chapter, we present a photometric stereo method [52] that requires a single color image, targeting Lambertian surfaces with arbitrary textures composed of a finite number of color reflectances (RGB reflectances) that have nonzero reflectance values for all color channels. Unlike conventional color photometric stereo methods, target surfaces are allowed to have a countable number of color reflectances, but color gradation is not allowed as in fact one cannot count the number of color reflectances on such surfaces. This assumption on surface color is equivalent to state-of-the-art single capture color photometric stereo methods [14,32], but our method uses only a single capture, while we must aware that using two or more frames or viewpoints always provides certain benefits like depth accuracy as well as local concave accuracy by using multi-view stereo augmentation. Going back and move to our next assumption, we also assume a moderate condition on surface geometry to be sufficiently curved. Under these assumptions, we show that surface normals are determined almost everywhere in the range of the Gaussian map as well as the corresponding effective color reflectances

that subsumes illumination intensities. This in turn suggests a benefit of omitting the calibration process of the illumination intensities, which is often difficult compared to that of light source positions. Based on the observation that the norms of surface normals fluctuate around one if they are calculated under any wrong reflectance, it is shown that the surface classification into constant reflectance regions is obtained with allowing multiple solutions on some surface points. We also show that the best candidate solution is estimated under an assumption that the surface prefers a small curvature. Unlike the existing single-color-image-based photometric stereo methods, our method encompasses multi-colored surfaces, and the surface classification is instant in that the whole surface region of the same color is detected at once whether it is connected or disconnected.

This chapter fully covers the published materials in [52]. The figures and script are included under the permission from the journal Computer Vision and Image Understanding.

4.1 Method

4.1.1 Observation under three light sources

Although we have already shown the general color photometric stereo problem in Equation (14) and (15), we revisit them again. We consider a Lambertian surface lit by three light sources with different spectra. The coordinate system describes the relative positions of the surface, light sources, and an RGB camera, and it is assumed that every point on the surface is approximately at the origin. We denote the surface geometry projected onto the image plane as $\mathcal{M}$. The j -th channel pixel value at $x \in \mathcal{M}$ is

$$v^j(x) = \sum_k \mathbf{n}(x) \cdot \mathbf{l}^k \int_{\lambda \in \text{supp}(S^j)} \rho(x, \lambda) S^j(\lambda) E_k(\lambda) d\lambda, \quad (27)$$

where $\mathbf{n}$ and ρ are the surface normal and reflectance spectrum, respectively, $\mathbf{l}^k$ and E_k ($k = 1, 2, 3$) are the k -th light vector and its radiance spectrum, and S^j is the camera-sensitivity function of the j -th band. We assume that only the light vectors $\mathbf{l}^k$ (only the orientations) are known; the intensities can be unknown. In Equation (27), the surface normals are described as vectors on a unit sphere embedded in $\mathbb{R}^3$ with a normalized orthogonal basis set $\mathbf{e}_i$ ($i = 1, 2, 3$), and its dual basis set is defined for light vectors. Provided that all channels have nonzero intensities, the map from the surface normal space to the observation color space is isomorphic, and we can take a basis set $\mathbf{e}'_i$ ($i = 1, 2, 3$) for the color space that has coordinates $(v^1, v^2, v^3)^T$ on it. Hereafter, matrices (denoted with uppercase letters with

hats) and vectors (bold lowercase letters) are assumed to be representations in those sets of coordinates, and the combined surface-reflectance property and illumination geometry gives a map from surface normals to observed color observations, which is represented by a matrix. The assumption that all color channels observe nonzero intensities are not always satisfied for real materials. However, such counterexamples are moderate when compared with nonzero-intensity assumption within narrow bands, since most RGB cameras have wide nonzero supports of sensitivity functions and each of them likely overlap with reflectance spectra. For application purpose, to avoid the lack of observation due to zero reflectance, we can optionally prepare appropriate cameras for the category of target materials.

Using a color reflectance matrix $\hat{A}(x)$ that is effective with the sensitivity function of the camera, the elements of which are

$$A_k^j(x) = \int_{\lambda \in \text{supp}(S^j)} \rho(x, \lambda) S^j(\lambda) E_k(\lambda) d\lambda, \quad (28)$$

and a light vector $L = (\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3)^T$, Equation (1) becomes

$$\mathbf{v}(x) = \hat{B}(x)\mathbf{n}(x), \quad (29)$$

where we define another matrix $\hat{B}(x) \equiv \hat{A}(x)\hat{L}$. Matrices $\hat{A}$ and $\hat{B}$ are supposed to be full-rank, and this is true when all the three channels have nonzero intensities, and the three light sources and the origin are not in the same plane. Note that, our method does not need the light intensities to be calibrated, since the effective color reflectance subsumes them.

Since the color reflectance and surface normals are unknown, we have to retrieve $\mathbf{n}$ and $\hat{A}$ from the underdetermined linear equations. In general, it is known that under a certain condition $\hat{B}$ can be determined only up to a subset of the general linear group, while a specific form of $\hat{B}$ can be determined uniquely. We discuss this fact in the following subsection. Our discussion is largely based on the following work by Nikolaev [24], Petrov [26]; Woodham et al. [53], who investigated an uncalibrated three-source photometric stereo problem but we can see a close relevance and an essentially the same formulation as the other studies on color photometric stereo; Drew [27], Kontsevich et al. [28], Petrov and Kontsevich [30], Drew and Brill [31], all of which have shown that $\hat{B}$ can be determined up to orthogonal group actions; and the important work by Chakrabarti and Sunkavalli [32], who first showed that $\hat{B}$ can be determined uniquely when $\hat{A}$ is diagonal. In Section 4.1.2, we summarize these

results by first discussing the surface normal ambiguity for an arbitrary reflectance matrix and then its uniqueness for a diagonal matrix.

4.1.2 On unique solution for color reflectance

First, we consider a uniformly colored surface and thus the reflectance matrix $\hat{A}$ and $\hat{B}$ are constant. The surface normal at $x \in \mathcal{M}$ is

$$\mathbf{n}(x) = \hat{B}^{-1}\mathbf{v}(x) \quad (30)$$

with an unknown matrix $\hat{B}^{-1}$. Using the fact that its square norm is one, a quadratic form is obtained:

$$\mathbf{v}^T \hat{C} \mathbf{v} = 1, \quad (31)$$

where $\hat{C} \equiv (\hat{B}^{-1})^T \hat{B}^{-1}$ is a symmetric matrix. Since $\hat{C}$ has six independent parameters from its symmetry, we need six independent quadratic forms to determine $\hat{C}$. Using the coordinates:

$$\sum_{i \leq j} \tilde{C}_i^j v^i v_j = 1, \quad (32)$$

where $\tilde{C}_i^i = C_i^i$ and $\tilde{C}_i^j = 2C_i^j$ ($i < j$), six independent linear combinations of direct products $\{(v^i \mathbf{e}'_i, v^j \mathbf{e}'_j) \mid i \leq j\}$ are necessary and sufficient. Six independent linear combinations of $\{(n^i \mathbf{e}_i, n^j \mathbf{e}_j) \mid i \leq j\}$ are necessary and sufficient to determine the matrix $\hat{C}$ in the surface normal space that is isomorphic to the observation space. For example, we can see that a set of (normalized) surface normals, with a coordinates system in which the x -axis is aligned to the optical axis, $\{(1, 0, 0)^T, (1, \delta, 0)^T, (1, \delta', 0)^T, (1, 0, \varepsilon)^T, (1, 0, \varepsilon')^T, (1, \delta'', \varepsilon'')^T\}$ with $\delta \neq \delta' \neq \delta'', \varepsilon \neq \varepsilon' \neq \varepsilon''$ is a case that determines $\hat{C}$ uniquely. A smooth surface that has non-degenerate points on it satisfies the condition, which is called the rank-3 shape [26].

Even when $\hat{C}$ is determined, $\hat{B}$ is identified only up to $\text{GL}(3)/\text{O}(3)$, where $\text{GL}(3)$ and $\text{O}(3)$ denote the general linear group and the orthogonal group on $\mathbb{R}^3$, respectively. Provided that the correct $\hat{B}$ does not violate Equation (31), the members of the quotient set $[\hat{B}] \in \text{GL}(3)/\text{O}(3)$ do not violate it either, since $\text{O}(3)$ is isometric and it holds that $(\hat{B}^{-1})^T (\hat{R}^{-1})^T \hat{R}^{-1} \hat{B}^{-1} = (\hat{B}^{-1})^T \hat{B}^{-1}$ for any $\hat{R} \in \text{O}(3)$. Thus, there exists $\mathbf{n}$ that satisfies Equation (29) without violating Equation (31) for $\hat{B}\hat{R}$ as well as $\hat{B}$. Constraining $\hat{A}$ to be a positive matrix or constraining the visibility of points for light

sources are not considered in our discussion.

This ambiguity is resolved by the fact that the set of diagonal matrices is not transitive under any left actions of $O(3)$ (in any representation): consider $\hat{B} = \hat{A}\hat{L}$, where $\hat{A}$ is a diagonal matrix. If there exists another diagonal matrix $\hat{A}'$ such that $\hat{B}\hat{R} = \hat{A}'\hat{L}$ under $\hat{R} \in O(3)$, then $\hat{B}$ has orthogonal ambiguity. However, such a diagonal matrix does not exist, since $\hat{A}'\hat{L} = \hat{B}\hat{R} = \hat{A}\hat{L}\hat{R}$ and $\hat{A}' = \hat{A}\hat{L}\hat{R}\hat{L}^{-1}$, which is not a diagonal matrix even if $\hat{A}$ is diagonal. Thus, $\hat{B}$ can be determined uniquely, if $\hat{A}$ is diagonal. In this case, three independent surface normals are required and sufficient to determine $\hat{A}$, since $\hat{L}$ is assumed to be known by calibration and the elements of $\hat{A}$ are determined from Equation (31): $\mathbf{v}^T(\hat{A}^{-1})^T(\hat{L}^{-1})^T(\hat{L}^{-1})\hat{A}^{-1}\mathbf{v} = 1$, given three independent linear combinations of $\{(v^i\mathbf{e}'_i, v^i\mathbf{e}'_i) \mid i = 1, 2, 3\}$, each of which extracts just one of the diagonal elements. As in the case of an arbitrary reflectance matrices, $\{(1, 0, 0)^T, (1, \delta, 0)^T, (1, 0, \varepsilon)^T\}$ with $\delta \neq 0$ and $\varepsilon \neq 0$ is possible. In practice, a sufficiently curved surface satisfies this condition as already mentioned. This condition is achieved when the radiance spectra and camera sensitivities satisfy $\text{supp}(E_k) \subset \text{supp}(S^k) - (\cup_{j \neq k} \text{supp}(S^j))$; e.g., reflected blue light is observed only in the blue channel. Hereafter, $\hat{A}$ denotes a diagonal matrix with three unknown parameters, each of which corresponds to the color observed by the camera.

4.1.3 Surface-color feature on multi-colored surfaces

Since surface-color classification is not generally given in advance, we must consider another criterion that can classify a surface into regions of the same color. Such a criterion can be found in the following discussion of what set of surface normals and a fixed reflectance matrix fit the observations and vice versa, and what simple geometric and color features can determine it.

To find such a feature, we compare the surface normals that are calculated with the correct reflectance matrix and with an incorrect one by using Equation (29). With the correct reflectance matrix $\hat{A}$ (and $\hat{B} = \hat{A}\hat{L}$), we have $\mathbf{v}(x) = \hat{B}\mathbf{n}(x)$, and with an incorrect reflectance matrix $\hat{A}'$ ($\hat{B}' = \hat{A}'\hat{L}$), we have $\mathbf{v}(x) = \hat{B}'\mathbf{n}'(x)$. Comparing them, we find

$$\hat{B}'^{-1}\hat{B}\mathbf{n}(x) = \mathbf{n}'(x). \quad (33)$$

Using the normalization constraint on $\mathbf{n}'$, the incorrect surface normals appear on the intersection

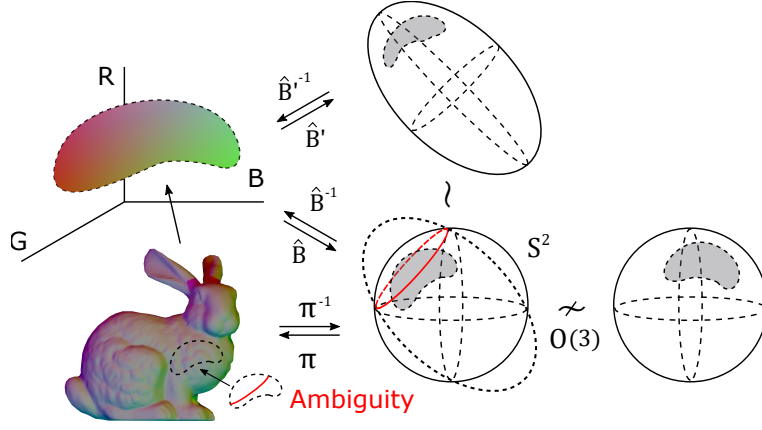


Figure 13: Uniqueness of surface normals on the same colored region.

curve between a quadratic surface

$$\mathbf{n}^T (\hat{B}'^{-1} \hat{B})^T (\hat{B}'^{-1} \hat{B}) \mathbf{n} = 1 \quad (34)$$

and the Gaussian sphere

$$\|\mathbf{n}\|^2 = 1, \quad (35)$$

which are mapped from the true surface geometry. The induced measure of this intersection curve on the two-dimensional surfaces is zero unless $\hat{A}' = \hat{A}$, since, as discussed in section 4.1.2, there is no $\hat{R} \in O(3)$ such that $\hat{B}' = \hat{B}\hat{R}$ thereby reducing the ellipsoid in Equation (34) to the unit sphere. Figure 13 shows the surface normal uniqueness on the same colored region. Since the reflectance matrix is diagonal, the set of surface normals (gray shaded) mapped from the surface geometry using the correct reflectance matrix $\hat{B}$ resides on the unit sphere, while it cannot be equivalent to another region of the unit sphere with satisfying the observed color. Although the surface normals solved using a incorrect (diagonal) reflectance matrix $\hat{B}'$ are mapped on an ellipsoid, this is wrong. This simple fact that evaluated from the normalization of surface normals reflects the global geometric property of same brightness as well as color composition.

Conversely, very few surface directions solved for with an incorrect reflectance matrix reside on the Gaussian sphere. That is, $\|\hat{B}^{-1} \mathbf{v}\|$ is one everywhere on the surface only if the correct reflectance matrix is given, while it deviates from one almost everywhere if given an incorrect reflectance matrix. Let us express the above idea in a representation that can be used to reduce the dimension of the color

reflectance in a more manageable representation: provided that the norm of the reflectance matrix is normalized, then $\|\hat{\tilde{A}}\|_F = 1$ and $\|\hat{L}^{-1}\hat{\tilde{A}}^{-1}\mathbf{v}\| = r$ for some constant r if given the normalized one of the correct reflectance matrix, while $\|\hat{L}^{-1}\hat{\tilde{A}}^{-1}\mathbf{v}\|$ varies on the surface and deviates from that value almost everywhere if an incorrect reflectance matrix is given. In the case of the correct reflectance matrix, this means that $r = \|\hat{A}\|_F$ and the surface normals have unit lengths.

This observation shows that the distribution of the norms of color-reflectance matrices is expected to be a global feature of surface regions of the same color, and this surface color feature depends on the global surface geometry and global color distribution. Note that, once the surface classification and the color reflectance are obtained, the geometry is retrieved locally, as the standard photometric stereo method benefits from shading information purely described locally in the tangent spaces on the surface.

Although the intersection curves where ambiguities exist would be wider due to image noise and color unevenness in real scenes even for a surface colored by the same materials, this classifier should still work if it is allowed to have classification margins, as we will introduce in Section 3.5. In this sense, the discussion about the intersection curves must be modified so that the classification ambiguities appear in small areas on the Gaussian sphere.

4.1.4 Surface classification via reflectance norm distribution

We consider a multi-colored surface and introduce the reflectance norm as a surface color feature. Suppose that $\mathcal{M}$ has Q regions of constant colors: $\mathcal{M} = \cup_{q=1,\dots,Q} \Omega_q$, the reflectance matrix $\hat{A}_{\Omega_q}^*(x)$ is a constant matrix within the same colored region Ω_q ($q = 1, \dots, Q$), where we indicate by asterisks the correct reflectance matrices and the spherical coordinates in Equation (37). Omitting the class label for simplicity, the surface normal at $x \in \Omega$ for a region is

$$\mathbf{n}(x) = \hat{L}^{-1} \hat{A}_{\Omega}^{*-1} \mathbf{v}(x). \quad (36)$$

Let us represent the constant reflectance matrix using the spherical coordinate system as

$$\hat{A}_{\Omega} = \begin{pmatrix} r \cos \theta \sin \phi & 0 & 0 \\ 0 & r \sin \theta \sin \phi & 0 \\ 0 & 0 & r \cos \phi \end{pmatrix}, \quad (37)$$

where $r \in (0, \infty)$ is the norm of the reflectance matrix, $\|\hat{A}\|_F$, and $\theta, \phi \in (0, \pi/2)$. Within the region Ω , surface directions solved for with an incorrect reflectance matrix exist apart from the unit sphere. On the other hand, surface norms reside correctly on the unit sphere when solved for with the correct reflectance matrix $\hat{A}_\Omega^*(r^*, \theta^*, \phi^*)$. In the correct case, it makes sense that we can rewrite Equation (36) using (normalized) surface normals as

$$r^* \mathbf{n}(x) = \hat{L}^{-1} \bar{\hat{A}}_\Omega^{*-1} \mathbf{v}(x), \quad (38)$$

where $\hat{A}_\Omega^* = r^* \bar{\hat{A}}_\Omega^*(\theta^*, \phi^*)$. Equation (38) can be extended to the whole range of (r, θ, ϕ) , which determines a map

$$r(x; \theta, \phi) = \|\hat{L}^{-1} \bar{\hat{A}}_\Omega^{-1}(\theta, \phi) \mathbf{v}(x)\| \quad (39)$$

of $r(\cdot; \theta, \phi) : \mathcal{M} \rightarrow (0, \infty)$ for each fixed (θ, ϕ) . From the discussion in Section 4.1.3, r takes the same value in each constant reflectance region only if the correct (θ^*, ϕ^*) is given. It can also be seen from Section 4.1.2 that the pair (θ^*, ϕ^*) is uniquely determined almost everywhere on the mapped Gaussian sphere so that r maps points in the colored region to the same value, when the geometrical condition discussed in Section 4.1.2 is satisfied for each constant reflectance region. Therefore, each region of the same color is characterized uniquely by r when each surface region is sufficiently curved. Although $r(x; \theta, \phi)$ is no longer the norm of such a reflectance matrix in the case of an incorrect reflectance matrix, we call $r(x; \theta, \phi)$ the reflectance norm. The reflectance norm subsumes implicitly the deviations of the normalized reflectance matrix, the norm of the reflectance matrix expressed in Equation (37), and the resulting surface directions. The reflectance norm is constant over x in a uniformly colored region only when one of the correct reflectance matrices is given.

We will summarize our discussion by giving a simple example. Figure 14 schematically shows the estimation process of our method. The reflectance matrices are diagonal, the space spanned by which is isomorphic to $\mathbb{R}_{>}^3$, where the subscript $>$ denotes that all the coordinates are strictly positive. Consider a curved surface composed of two regions with reflectances $\hat{A}_1^*$ and $\hat{A}_2^*$, lit by blue, green, and red light sources. We equivalently transform the sets of reflectance matrices and surface normals, $(\hat{A}, \mathbf{n}) \in \mathbb{R}_{>}^3 \times \mathcal{S}^2$, to the sets of the normalized reflectance matrices and (unnormalized) surface directions, $(\bar{\hat{A}}, r\mathbf{n}) \in \mathcal{S}_{>}^2 \times \mathbb{R}^3$, where $r = \|\hat{A}\|_F$ is the reflectance norm and $\mathcal{S}^2$ is the unit sphere in $\mathbb{R}^3$. Let the surface projected on the image plane be classified as $\mathcal{M} = \Omega(\hat{A}_1^*) \cup \Omega(\hat{A}_2^*)$.

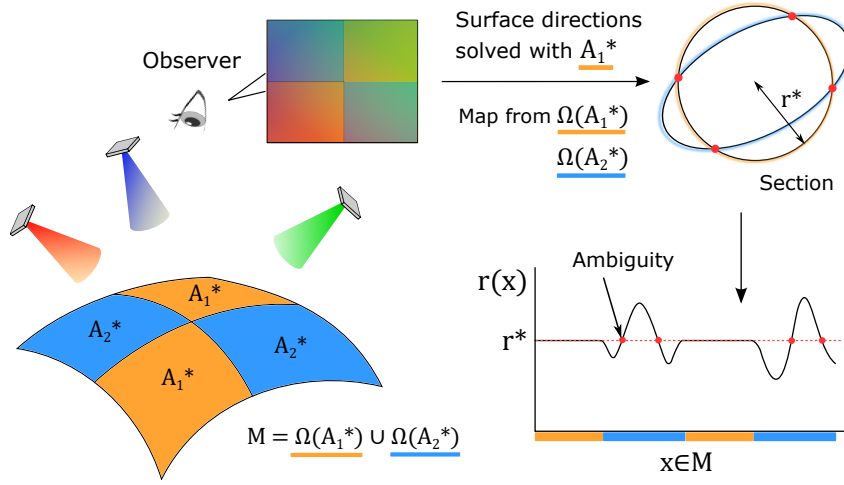


Figure 14: Schematic of single color image photometric stereo (with permission from [52], Computer Vision and Image Understanding).

By focusing on the region $\Omega(\hat{A}_1^*)$ and provided a reflectance $\bar{\hat{A}}_1^*$, the obtained surface directions are mapped from $\Omega(\hat{A}_1^*)$ onto a sphere with a diameter of r_1^* . This gives a dense reflectance norm distribution, which is correct, while almost everywhere in $\Omega(\hat{A}_2^*)$ cannot be mapped onto a sphere of any diameter, which is wrong. It is only when given one of the normalized correct reflectance matrices that the calculated surface directions on a region of the same color are mapped onto a sphere. As a result, $\Omega(\hat{A}_1^*)$ is found in the range of the inverse map from r_1^* . Ambiguities can occur only at points, the calculated reflectance norms of which are occasionally the same. We tackle this ambiguity with a surface smoothness constraint in Section 4.1.6.

Intuitively, the possible variation in the reflectance norm over surface points in Equation (39) can be interpreted in two ways. First, the supposed reflectance matrix was wrong; thus, there is no value of r with which $\bar{\hat{A}}$ returns correctly normalized surface normals on any constant reflectance region. Second, there is a variety of r 's on the surface with the same (θ, ϕ) ; in other words, the brightness change on the surface is caused by the spatially fluctuating distribution of the reflectance norm rather than by shading. The first interpretation is our concern, since we cannot apply our method to the second case. Our scope is that shadings are the cause of the color variation rather than spatially fluctuating reflectance norms.

4.1.5 Surface color classification algorithm

We now present a practical algorithm for surface classification and estimation of effective color reflectances. The algorithm is composed of three steps, the first and second of which are described in this section, and the third in section 4.1.6.

In the first step, we identify all the normalized reflectance matrices on the surface. Since the reflectance norms from a same colored region calculated with the correct normalized reflectance matrix concentrate on a value r , we can determine the correct normalized reflectance matrix from the histogram of the reflectance norms. Once the set of parameters (θ, ϕ) is fixed, we can calculate the reflectance norm at each point $x \in \mathcal{M}$ from Equation (39). We count the frequencies of reflectance norms from the histogram sampled over all the points $x \in \mathcal{M}$, and then find the most frequent reflectance norm at the (θ, ϕ) . The number of bins in the histogram is set to $(\max(r) - \min(r))/(\delta \cdot \tilde{r})$, where δ is a parameter that determines the bin width and $\tilde{r}$ is the median value of the reflectance norms, balancing the size of the histogram. By searching for (θ, ϕ) over $(0, \pi/2) \times (0, \pi/2)$ every 1 degree and comparing the most frequent reflectance norms between all the (θ, ϕ) 's in $(0, \pi/2) \times (0, \pi/2)$, we can determine the reflectance norm r_1^* and (θ_1^*, ϕ_1^*) for the first color on the surface. Next, we find the second color by applying the above procedure to all the points $x \in \mathcal{M} - r^{-1}(\mathcal{U}_\varepsilon(r_1^*); \theta_1^*, \phi_1^*)$ while removing the first candidate colored region, where $r^{-1}(\cdot; \theta, \phi) : (0, \infty) \rightarrow \mathcal{M}$ is the inverse map of Equation (39) for (θ, ϕ) , and $\mathcal{U}_\varepsilon(r) = (r - \varepsilon, r + \varepsilon)$ allows some margin for small fluctuations in reflectance norms due to color unevenness, noise, and the coarseness of the discretization of reflectance values. $r^{-1}(\mathcal{U}_\varepsilon(r^*); \theta^*, \phi^*)$ returns the pixel sets, where the calculated reflectance norms fall within the norm interval $\mathcal{U}_\varepsilon(r^*)$. We tuned this parameter as $\varepsilon(r) = \varepsilon r$ with a constant value ε . In general for finding the q -th color on the surface, we perform the same procedure on all the points $x \in \mathcal{M} - \cup_{p=1, \dots, (q-1)} r^{-1}(\mathcal{U}_\varepsilon(r_p^*); \theta_p^*, \phi_p^*)$, until all the pixels are classified at the final Q -th iteration.

Figure 15 shows an example of the histogram of r values for a synthetic textured sphere. The most frequent reflectance norm while changing (θ, ϕ) over $(0, \pi/2) \times (0, \pi/2)$ occurs when the correct normalized reflectance matrix (left) is given. On the other hand, the reflectance norms fluctuate over surface points when given a wrong one (right).

It is important for our method that every q -th classified region must satisfy the geometrical condition in Equation (32) in Section 4.1.2: for the reflectance norm histogram to work in practice, sufficiently curved surface regions of the same color are required to exhibit strong peaks in the histogram.

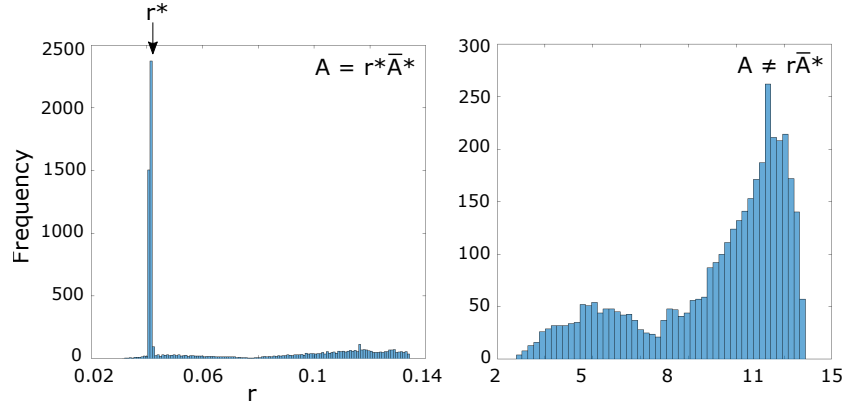


Figure 15: Histogram of reflectance norms for sphere for the normalized correct reflectance matrix (left) and an incorrect one (right) (with permission from [52], Computer Vision and Image Understanding).

Unfortunate cases are that the inverse map of $r(\cdot; \theta, \phi) : \mathcal{M} \rightarrow (0, \infty)$ from a reflectance norm on the intersection between Equation (34) and (35) has too large an area to be disambiguated with the smoothness constraint described in Section 4.1.6.

One can see a closely related idea that introduces a reflectance histogram in a study of uncalibrated photometric stereo by Alldrin et al. [54]. In fact, after the surface classification is obtained, the entropy of the reflectance norm distribution should be minimum. Note, that method [54] relies less on the surface color distribution than our method does.

In the second step, we update the color classification so that each classified region fully covers each same colored region. Although we have found the colors that exist on the surface and their effective reflectance matrices, some of the classified regions do not cover the whole same colored regions. This is because in the first step the ambiguous regions labeled formerly have been removed in the following search for the remaining colors, and thus have not been classified to the correct region. Thus, in the second step, for each fixed color (θ_q^*, ϕ_q^*) ($q = 1, \dots, Q$), we calculate the reflectance norms over *all* the points on $\mathcal{M}$ from Equation (39) and find the most frequent reflectance norm from the histogram. It holds that r_q^* determined in the first step and the q -th reflectance norm found in this step are the same in most cases. On the other hand, we need to take into account another cause of misclassification in which an ambiguous region of some class q covered in the second step has a larger area than a region labeled by $q' (> q)$ in the first step. In this case, the colored region with a smaller area is missed. To avoid such a lack in the latent classes, we define each classified region Ω_q ($q = 1, \dots, Q$) as a union of the two q -th classified regions in the first and second steps. Finally, we obtain a sequence of sets of

constant reflectance regions and reflectance matrices, $\{(\Omega_1, \hat{A}_{\Omega_1}^*), \dots, (\Omega_Q, \hat{A}_{\Omega_Q}^*)\}$. By substituting each reflectance matrix $\hat{A}_{\Omega_q}^*$ into Equation (36) over Ω_q , the surface-normal field is obtained. Note that there remain classification ambiguities $\Omega_p \cap \Omega_q \neq \emptyset$ for some p and q ; this is discussed in the next section.

4.1.6 Resolving classification ambiguity

Since the following third step relies only on the surface smoothness assumption and is independent of the reflectance norm, it can be treated separately from the first and second steps.

Points may exist that are attributed to two or more classified regions. Such classification ambiguity is included in the reflectance sets obtained above. To identify the correct surface normal from the candidate surface normals at each point, we introduce an energy function in an intuitive manner. Surface regions with two or more candidate surface normals are of zero-measure in the ideal case, noiseless and composed of perfectly finite number of reflectances, but even in real scenes, regions with classification ambiguities would have small areas.

Since the angles between the surface normals at neighborhood points tend to be small for a smooth surface, we determine the best candidate class at x by minimizing the angles between the surface normals to be determined and the ones that have already been determined uniquely. Provided that each point $x' \in \Omega_{q'}$ near the point x has been classified uniquely to some color class q' , the surface normal $\mathbf{n}(x')$ there has also been determined. Using these surface normals, we determine the best color class q^* at the ambiguous point x by

$$q^*(x) = \arg \min_{q \in \{1, \dots, Q\}} \sum_{\substack{x' \neq x, \\ x' \in \Omega_{q'}}} -\alpha(x, x') \mathbf{n}_q(x) \cdot \mathbf{n}(x'), \quad (40)$$

where $\mathbf{n}_q$ is the surface normal obtained using the q -th reflectance matrix, and the summation is over surface normals that have been uniquely obtained. The $\alpha(x, x')$ is a positive function that attenuates with x' far from x , and we define it as $\alpha(x, x') = e^{-\|x - x'\|}$. For efficient calculations and meaningful surface normal similarities, the summation in Equation (40) is restricted to be within $\|x - x'\| \leq \Delta_{\text{smooth}}$ with a parameter Δ_{smooth} , with which surface smoothness is assumed. Note that the metric depends on the pixel count.

When the area of an ambiguous region $\Omega_p \cap \Omega_q$ for some classes p and q is large enough to exclude

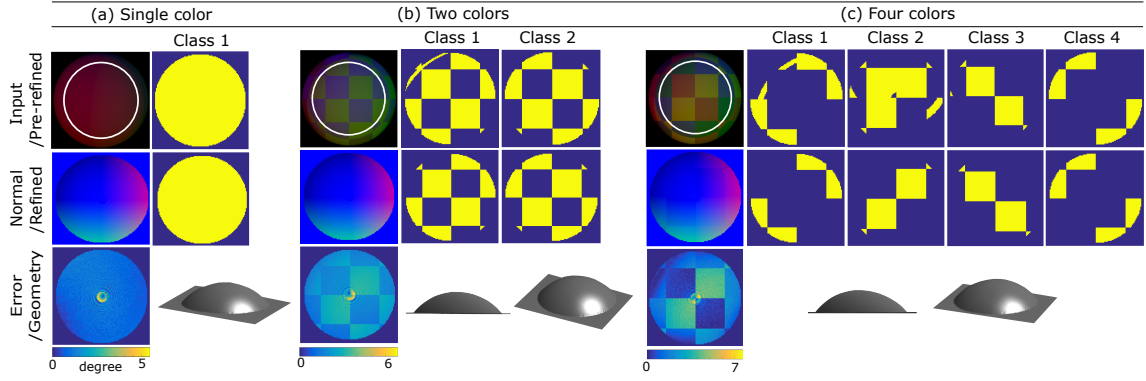


Figure 16: Simulation results for textured spheres: (a) single-color texture of Macbeth-15, (b) two-color texture of Macbeth-5 and 11, and (c) four-color texture of Macbeth-3, 7, 14, and 16. Pre-refined results during surface-classification process (top row), refined classes after resolving ambiguities (second), surface-normal fields and their errors (third), and retrieved surface geometries (bottom) (with permission from [52], Computer Vision and Image Understanding).

all neighborhood points within Δ_{smooth} , the smoothing constraint is unreasonable and the energies in Equation (40) are set the same for these classes. In this case, we select the formerly labeled class as the final output. Since $\mu(\Omega_p) - \mu(\Omega_p \cap \Omega_q) > \mu(\Omega_q) - \mu(\Omega_p \cap \Omega_q)$ for these classes, we consider that the region Ω_p is less likely than Ω_q to include misclassified regions.

4.2 Simulations

We evaluated the accuracy of our method using synthetic images. In the simulations, we used the rendering software pbrt-v2 [44] for making the images and MatlabR2016a for the evaluation. The height map was calculated from estimated surface normals by using the Frankot-Chellappa algorithm [45]. The spectral samples were taken every 1.2 nm within a wavelength range of 400 and 700 nm, and then summed over each support of the camera sensitivity functions. The light-source spectra and the RGB camera-sensitivity functions were designed so that the reflectance matrices would be diagonal. The camera was assumed to be orthographic, but the surface-classification algorithm in our method is independent of this assumption. Throughout the simulations, the target surfaces were placed at the origin of the Cartesian coordinate system, the camera at (0, 0, 200) in centimeters with the optical axis along the z-axis.

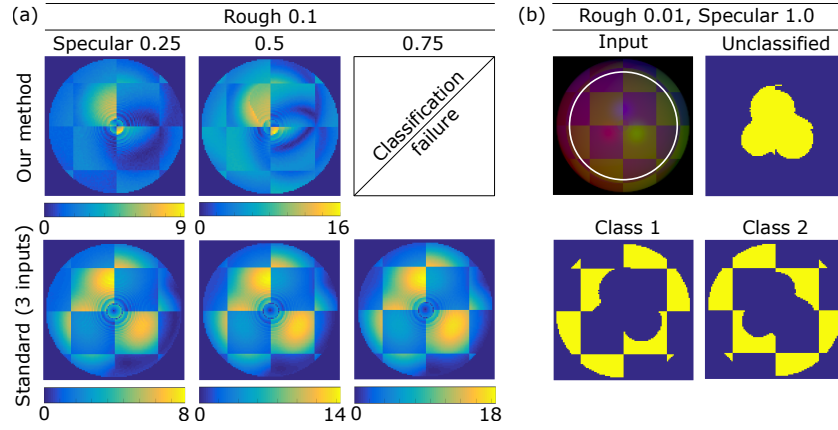


Figure 17: Simulation results for plastic materials.

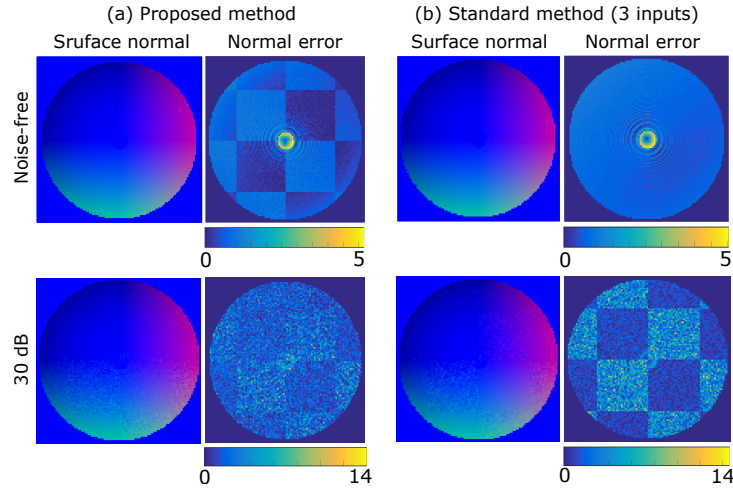


Figure 18: Comparison of (a) our method against (b) standard method: noise-free input (top) and input with 30 dB of added noise (bottom) (with permission from [52], Computer Vision and Image Understanding).

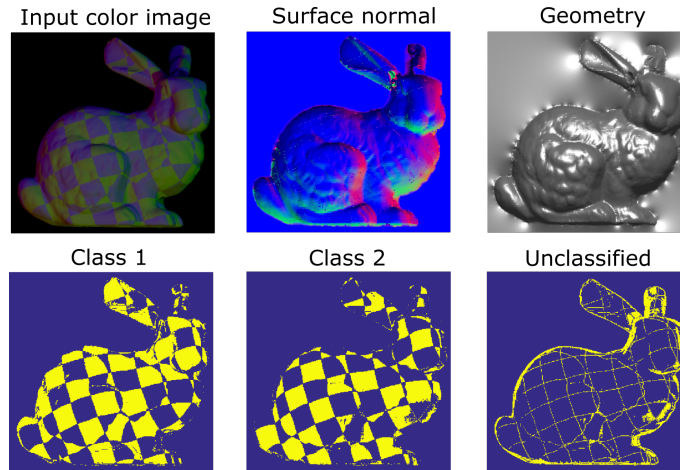


Figure 19: Simulation result (with permission from [52], Computer Vision and Image Understanding).

4.2.1 Evaluation for variety of textures

A textured Lambertian spherical surface with a diameter of 5 cm was simulated to evaluate the effectiveness of our method against scenes that satisfy the assumptions that a surface has a finite number of colors and is sufficiently curved as discussed in Section 4.1.2. The surface had a convex geometry without inter-reflection and was composed of a finite number of colors. The light sources were placed at $(0, -100, 200)^T$, $(86.6, 50.0, 200)^T$, and $(-86.6, 50.0, 200)^T$ in centimeters on a circle perpendicular to the optical axis, with even separation angles [36]. (θ, ϕ) for the reflectance search were discretized into 1 degree intervals, and we empirically determined the parameters as $\varepsilon = 0.075$, $\delta = 0.02$, and $\Delta_{\text{smooth}} = 5$. Note that in the surface classification tests, we found that $\delta \in (0.01, 0.03)$ and $\varepsilon \in (0.05, 0.1)$ were allowable, but selected values that were most robust to 40 dB of noise.

Three different textures were designed for the evaluation: (a) a single-color texture with Macbeth-15 reflectance spectra, (b) two-color texture with Macbeth-5 and 11, and (c) four-color texture with Macbeth-3, 7, 14, and 16, respectively.

Figure 16 shows the results. The (1,1) block (one separate square) in (a) is the input image of the sphere with the single-color texture, where the white circle (as in the other two input images) designates the region in which none of the light sources are occluded. Pre-refined and refined classification results before and after resolving the classification ambiguity due to the occasional coincidence of reflectance norms are shown in the (1,2) and (2,2) blocks in (a). The estimated surface normal field and error against the ground truth are shown in the (2, 1) and (3, 1) blocks, and the reconstructed surface geometry from a new view is at the bottom. The average normal error and standard deviation were 0.95 and 0.38 degrees, respectively. The large errors around the top of the sphere were due to the aliasing of the rendered image. This aliasing was the same for the other texture models below.

The (1,3) block is the input image for the two-color texture, and the results are shown in Figure 3 (b). A small area of classification ambiguity remains in the pre-refined results in the first row, whereas this ambiguity was resolved and refined classifications were correctly obtained in the second row. The average normal error of the estimation was 2.2 degrees and the standard deviation was 0.69 degrees.

The (1,6) block is the input for the four-color texture. As in the two-color case, some classification ambiguity remains in the pre-refined results. The reason for the relatively large area of classification ambiguity in the (1, 8) block is considered to be that the yellow- and orange-like colors are similar; thus the overlaps between the two quadratic surfaces of Equation (34) and (35), with an allowable margin ε

for the reflectance norm, are larger than those for the other reflectance combinations. Nevertheless, the refined classification results were correctly obtained with the smoothness constraint, and the surface normals were accurately estimated with an average normal error of 1.9 degrees and standard deviation of 1.5 degrees. Reconstructed surface geometry from new views are shown at the bottom.

To evaluate undesirable effects from specular reflection, we simulated spheres with two-colored plastic textures by varying the surface roughness and specularity in pbrt-v2 [44]. Figure 17 shows the results. For comparison using the same textures, we also show the surface normals estimated by the standard 3-inputs photometric stereo methods at the bottom row. The surface was correctly classified when the specular reflection is weak, while the classification and surface normals were corrupted for textures with strong specular highlights (Figure 17 (b)).

4.2.2 Comparison against standard photometric stereo

To compare the accuracies of our method and the standard photometric stereo method [4] that uses three input images, we added 30 dB of noise to the input images of the textured sphere with Macbeth-5 and 11 colors and compared the results of running each method on the images with and without noise.

For comparison, the degree range of the reflectance search was discretized into 0.5-degree intervals. We set the parameters as $\varepsilon = 0.15$ and $\Delta_{\text{smooth}} = 5$. The estimated surface normal fields and errors are shown in Figure 18. The top row shows results for images without added noise, and the bottom row shows results for images with 30 dB added noise. This finer discretization as 0.5 degree did not dramatically affect the resulting accuracy compared with the results from 1 degree discretization. In the noise-free case, the average normal error and standard deviation for our method were 0.65 and 0.45 degrees, while they were 0.77 and 0.35 degrees for the standard method. In the 30 dB added noise case, the average normal error and standard deviation for our method were 2.2 and 1.5 degrees, while they were 2.6 and 2.1 degrees for the standard method. The large error around the top of the sphere in both methods is due to aliasing from the texture, as described in Section 4.2.1. Our method correctly classified the colored regions and the estimated surface normals were comparable with those of the standard photometric stereo method that uses three images. Note that the estimation errors in our method were slightly smaller than those of the standard method, which we consider is due to the finite size of the setup. Since the object surface has a finite size, unit vectors from the surface points to the lights deviate from the light vectors that we fixed under an assumption that all the surface points reside

at the origin. Thus, the surface normals estimated using the standard method also have small error in general, but we should note that in principle the standard method can retrieve more accurate surface normals when the light sources are placed farther from the object. In this simulation example, we consider that our method compensated the deviation from the light vectors by finding the reflectance matrix that best fitted the observation. A similar effect can be found in the entropy minimization of the uncalibrated photometric stereo method proposed by Alldrin et al. [54].

4.2.3 Complicated surface geometry

Another simulation was conducted using the Stanford bunny geometry model with Macbeth-5 and 11 colors. This complicated geometry exhibits weak inter-reflections at concave regions, but nevertheless, the surface classification results were correctly obtained in the most of the surface regions. Figure 19 shows the input image, classification results, estimated surface normals, and the reconstructed surface geometry. The unclassified surface points are located at the boundaries between the two color regions, or occluded from one or more light sources. The surface normals at those unclassified points were estimated by interpolating the surface-normal field by using an interpolation method for regions lacking data [55].

4.3 Experiment with real objects

We have shown through simulations that our method works in principle, but it must be considered that real capturing processes inevitably include camera noise, imperfect light calibration, and color unevenness even for a surface colored with a countable number of materials. To demonstrate the practicability of our method, therefore, we captured images of real objects with various surface colors with an RGB digital camera and applied our method to them.

4.3.1 Experimental setup

We placed an RGB digital camera (FLOVEL FD-420M) and red, green, and blue light sources in front of the objects (Figure 20). The light sources were distributed around the optical axis so that the light vectors were linearly independent in the camera coordinate system. We confirmed that the light-source spectra and camera-sensitivity functions satisfy the assumption that the light-source spectra and the supports of the sensitivity functions have one-to-one correspondence. The gamma value of the camera

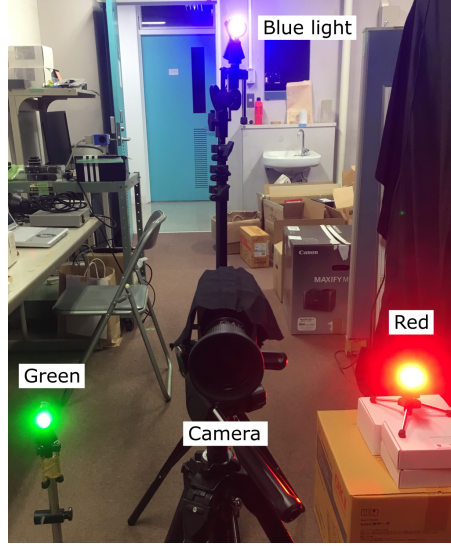


Figure 20: Camera and light sources in the experimental setup (with permission from [52], Computer Vision and Image Understanding).

was set to one so that the pixel intensities would be proportional to the irradiances. The light directions were calibrated with a polished steel sphere before capturing the objects.

4.3.2 Results and discussion

Figure 21 shows the results. From left to right, the columns show the photos of target objects, input color images, classified surface reflectance regions, estimated surface normals, and various views of rendered geometries calculated with the algorithm by Frankot and Chellappa [45]. For each object, the same colored regions were accurately classified, and reasonable surface normals and geometries were retrieved. We empirically set the parameters to $\delta = 0.02$, $\varepsilon \in [0.1, 0.2]$, and $\Delta_{\text{smooth}} \in [1, 10]$ and selected the best values for each object. Due to outlier pixels, e.g., color boundaries, dark channels, and background reflections that disturb the end of the search, we instead manually set the upper limit number of the classes. We allowed an excess to the upper limit on the number of classes, because the whole surface region is classified and excluded from the search region after completing the loop at the correct number of colors. Although we empirically determined the parameters in this experiment, a possible way to automatically determine the optimal parameters is to minimize the area of the ambiguous regions by decreasing the parameters starting from some large values while the surface is covered with the given upper number of colors.

Objects with two, three, and four colors are shown in the top four, middle four, and bottom two

rows in Figure 21, respectively. It took for example about 140 seconds for Robot (174×101 pixels) on a quad-core 2.60-GHz CPU with 16-GB memory. Figure 22 shows, for example, the reflectance norm histogram and the distribution of reflectance norms over the surface points when color class 1 of Bear was found. The peak in the histogram was broader than that of the simulation result, while the surface regions of the same reflectance were accurately classified. The broad spike might have been due to a large quantization noise and nonlinearity of the camera response, but most likely it was from the unevenness of the reflectance values even of the same materials. Although misclassifications existed around the hair of Matryoshka, the surrounding region was painted with two materials and a smoother surface was selected by our ambiguity resolution. The misclassification in the bottom-right region was certainly a failure, but the resulting surface normals and geometry were reasonably obtained. The concave areas of the objects were often misclassified because of inter-reflections or occlusions. The test objects had diffuse surfaces, and we did not consider specular effects in the surface normal estimation. The reflectance norm is affected by outliers like specularities, shadows, and strong interreflections. To reduce specular reflection effects, we can optionally place polarizers in front of the camera and light sources to reduce specular reflections as Fyffe et al. [23] did.

Although the classification results, surface normals, and reconstructed geometry were well estimated within each classified region, the inaccurately estimated surface normals at the color boundaries resulted in geometrical artifacts, as can be seen in, e.g., Robot in Figure 21. These artifacts are noticeable partly because of the low-resolution images used in the experiment (100-200 pixels for rows and columns). To reduce them, we shrank each classified region by a pixel and interpolated as in the simulation [55]. The resulting surface geometry is rendered from various views in the first column of Figure 23 (a). The second column in Figure 23 (a) shows the reconstruction using the standard three-input photometric stereo method from the red channel, where the light directions were the same but the light intensity was not constant at each light position, and we calibrated them by facilitating prior knowledge of the surface color. Since at least one vector of the surface normals on the class of the largest area was directed to each light source, the largest pixel value must have been proportional to the light intensity. Our reconstructed result matched the three-input result. The main cause of the local differences between these results is likely the color unevenness in the same color regions.

We also compared our method with a state-of-the-art method for single-image RGB photometric stereo, CS-2016 [32]. We ran their Matlab software on a GeForce GTX Titan X with 24-GB memory. For the Robot object, we determined the best parameters for CS-2016 by evaluating the surface normal

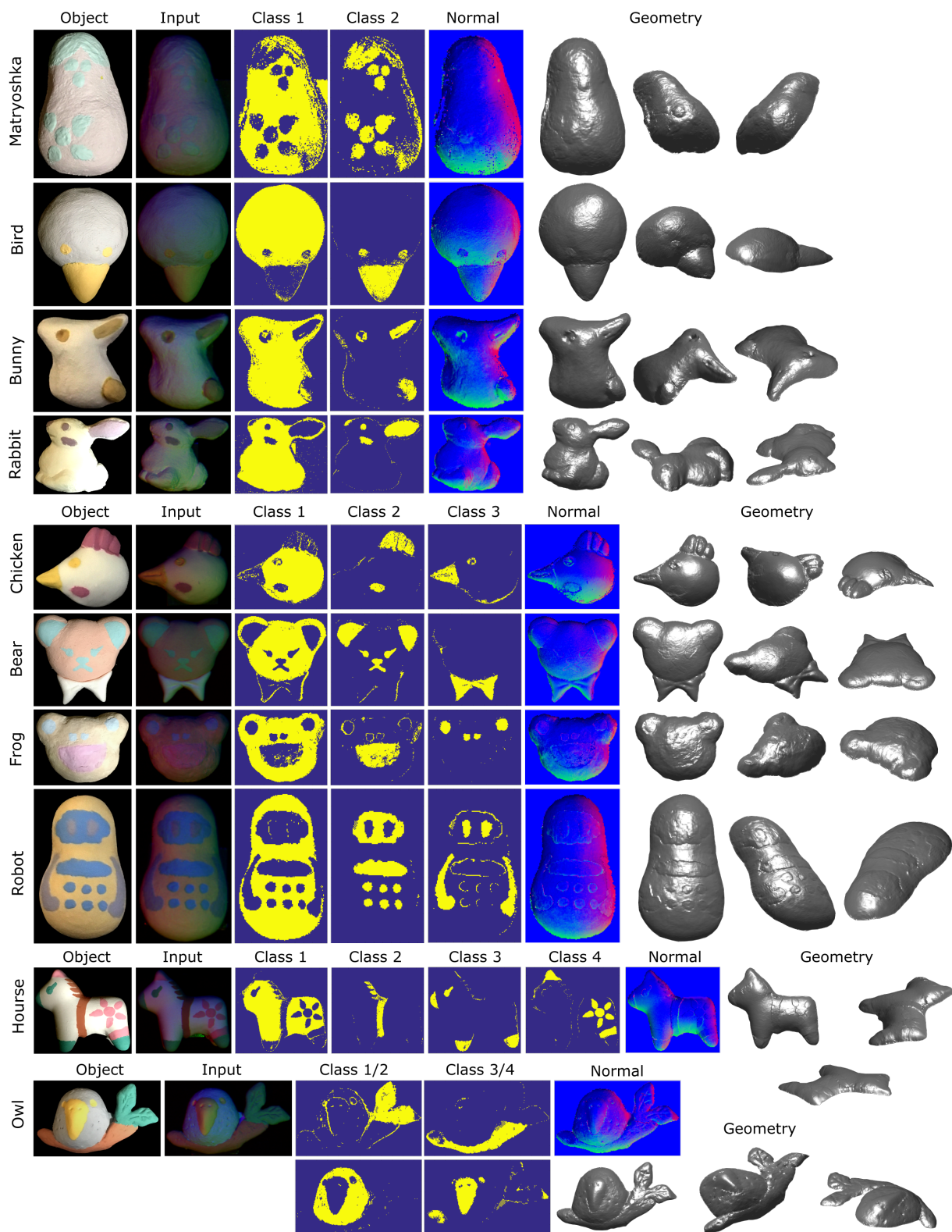


Figure 21: Experimental results for real objects: from left to right, columns show input images, classified surface regions, estimated surface normals, and various views of rendered geometries (with permission from [52], Computer Vision and Image Understanding).

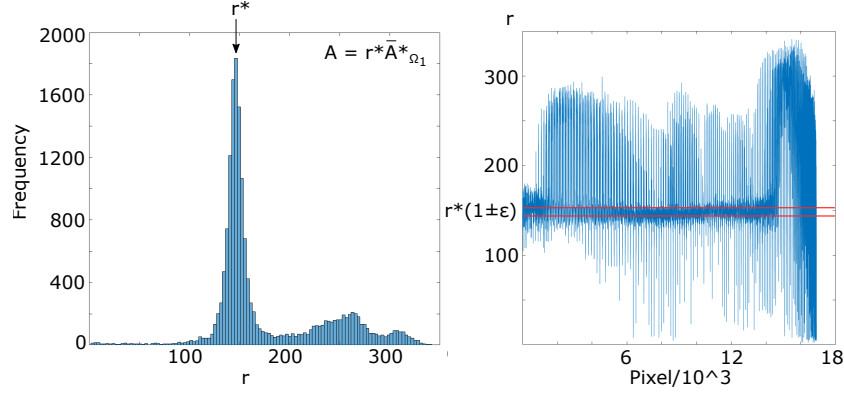


Figure 22: Histogram of reflectance norms for Bear for the correct normalized reflectance matrix (left) and distribution of reflectance norms over the surface points (right) (with permission from [52], Computer Vision and Image Understanding).

error against the ground truth from the 3-input standard method. We especially searched for trade-off parameters in CS-2016: from 2 to 10 in steps of 2 for the patch size, from 1 to 10 steps of 1 for the outlier threshold, and from 10^{-3} to 10^{-1} every 2×10^{-3} for the histogram error threshold. Other than setting azimuth, elevation, and luminance for reflectance quantization to $90 \times 90 \times 10000$, as well as setting 300 for the maximum luminance, the other parameters were set to the default values: 5 for the polynomial degree for the shape model, 100 for the size of the global reflectance set, and the value of the internal-varied parameter λ of Equation (10) in their paper. As a result, the best parameters were a patch size of 8, outlier threshold of 1, and histogram error threshold of 0.075. The surface normal estimation error did not dramatically change with different parameters, and their method was stable.

The estimated surface normals and geometry of the Robot are shown in the third column in Figure 23 (a). Our method retrieved more accurate surface normals and geometry compared against the ground truth obtained by the three-input standard method (second column) than the CS-2016 did (third column). Note also that the CS-2016 method approximately estimates surface geometries with color gradation, cases of which our method cannot address.

We also applied our method to the Rabbit image provided by Chakrabarti and Sunkavalli [32], using the same parameter for the Robot object. The image size was large for our method and the calculation slowed down significantly, because the calculation time of residual point sets when removing a classified region from the surface increases as the image becomes bigger. CS-2016 does not suffer from the image size problem; it calculates by patch in parallel. Thus, we trimmed three representative regions of 256×256 from the image affected by the multicolor problem: the belly and legs have two

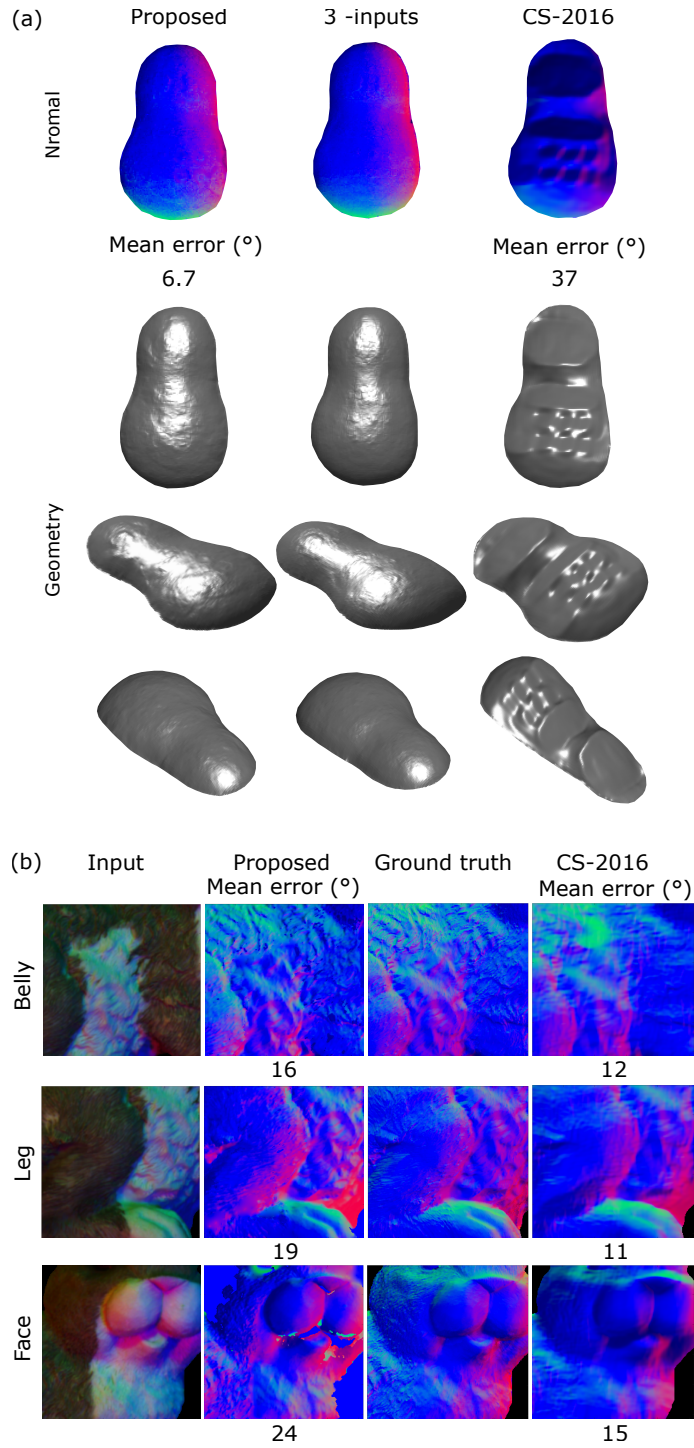


Figure 23: Comparison of three-input photometric stereo method, our method, and method of Chakrabarti and Sunkavalli (2016) (CS-2016) for (a) Robot and (b) Rabbit (provided by Chakrabarti and Sunkavalli) (with permission from [52], Computer Vision and Image Understanding).

colors, and the face seems to have three or more colors with color gradation. Figure 23 (b) shows the result; CS-2016 performed better than our method in terms of the mean error against the ground truth normals.

It is considered that the estimation error of CS-2016 for the Robot object was due to the concurrence of two or more colors in a local patch, which violates the assumption of Chakrabarti and Sunkavalli [32]. Another possible reason for the error is that the object surface is smooth and thus most local patches have less variation in surface direction, which might cause error in CS-2016. Our method handles the whole surface and thus all surface directions at once. On the other hand, the error of our method for the Rabbit object is considered to come from color unevenness, gradation, and shadows; our method suffered more from such irregularities in the reflectance estimation. In such a case, the geometric representation using a polynomial model in CS-2016 seems to stably approximate surface geometry by mitigating such outlier effects, while our method gives inaccurate surface normals. In the future, we should develop a more efficient algorithm to deal with the whole surface in a bigger image.

We consider that our method is suitable for objects like our targets that have smooth surfaces and have even disconnected colored regions. However, the above comparison points to the need to make it more robust to shadows, interreflections, color unevenness, and gradations.

4.4 Conclusion

We have presented a photometric stereo method for multi-colored surfaces that uses a single RGB image. It is an ill-posed problem to determine the surface normal and surface color at a point from a single color image with any color photometric stereo setup. In the proposed method, Under the assumption that the surface is colored with a finite number of materials of the same reflectance and the union of the surface regions of the same reflectance is curved, the difficulty of the surface reflectances being unknown is resolved by facilitating the global geometric information of reflectance norm distribution. Our method is comparable in accuracy to the three-input photometric stereo method, and it worked well in an experiment on real objects.

5 General consideration

In this final chapter, we present a general discussion on our proposed methods. Throughout this dissertation, we developed two single-capture photometric stereo methods for colored surfaces. Existing methods have severely restricted surface color or geometry to adopt the limited number of observations. Our methods enable to apply photometric stereo reconstruction to colored surfaces of arbitrary colors from a single hyperspectral image or multi-colored, curved surfaces from a single RGB image. The first approach facilitated the global spectral information, that is densely sampled signal values, for estimating the reflectance spectra without using any geometric prior, while the second facilitated the global geometric/chromatic information for estimating color classification and the color reflectance as well as the surface normal field without using global spectral information. We consider two classes of augmentative possibility for these methods: the first is to extend to multiband case the color classification criterion introduced in Chapter 4, and the second is to design illumination spectra for a spectral regression from subband reflectance values as in Chapter 3. Finally, we give a conclusion of this dissertation.

5.1 Multiband surface color classification in photometric stereo

We can apply the surface color feature to a $K(> 3)$ -band image for countable, multi-colored, curved surfaces as discussed in Chapter 4 for the $K = 3$ case there. Although our method enables a single-shot photometric stereo against some conditioned colored surfaces, there would be an uncertainty on surface normals when the surface has so small curvature regions with color classification ambiguities that the estimated surface normals would be inaccurate. Another problem is that we would lack some required observations in channels with no reflectance. Actually, some paints like red and blue often have no reflectance at their counterpart wavelengths.

By extending our discussion to a multiband case, we see that the surface color classification ambiguity could be resolved under some condition without assuming surface smoothness, that is the correct and unique surface classification could be augmented with combinatorially available additional reflectance norm distributions. Let us use $K(> 3)$ -bands with adding multiple constraints on possible color classifications. When the whole target surface colors have nonzero reflectances other than K_0 from K channels, we will have $_{K-K_0}C_3$ combinations of observable channels that can be used to estimate color classification and surface normal field. Note that K light sources and K channels should

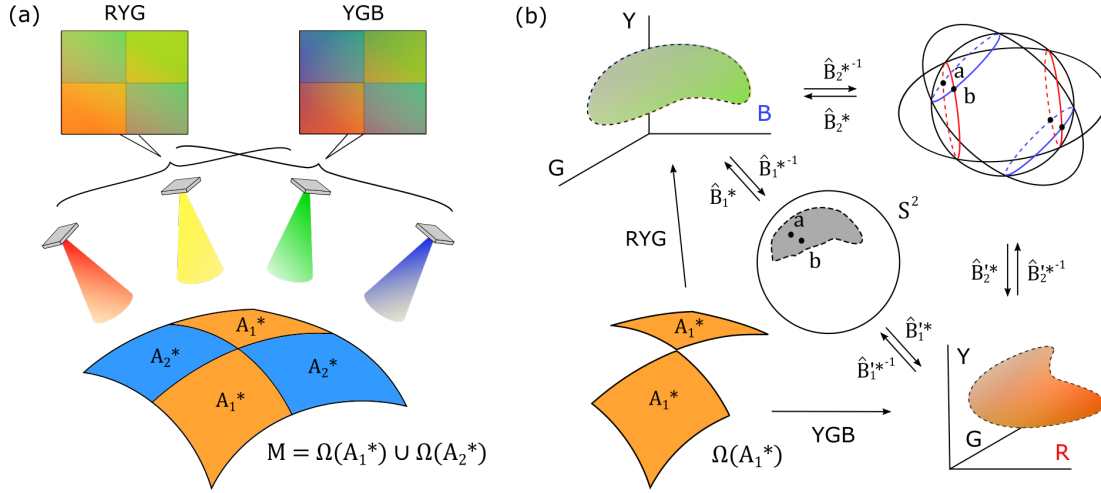


Figure 24: Schematic of surface normal ambiguities in a multiband image.

have one-to-one correspondence in our setup as already discussed. Each surface point is uniquely classified if one or more than one candidate color classes obtained under some combinations of channels uniquely classify the surface point. Figure 24 illustrates an example of a reduction of surface normal ambiguities. Provided that two three-color sets, e.g. (red(R), yellow(Y), green(G)) and (yellow, green, blue(B)), the reflectance norm distribution finds the two correct reflectance matrices $\hat{B}_1^*$ and $\hat{B}_1'^*$ for RYG and YGB channels, respectively, while the ambiguities appear at the intersections between the unit sphere and ellipsoids mapped from $\Omega(A_2^*)$. As a result, the ambiguities remain at the point a and b on the unit sphere, and the possibility for a surface point to encounter with an ambiguity is reduced. Moreover, it is also possible to characterize a color region if one of the selected channels observes no reflection from the surface, which reflects a part of surface color property. Since an effective color reflectance, which varies among materials, is never controlled only by tuning light source spectra or their emission intensities, this augmentation just increases the chance of correct classifications.

Another possibility is to use some bands for surface color classification in the first step before estimating the surface normal field under the global surface color distribution in the next step. For simplicity, consider K camera spectral sensitivity functions and the one-to-one corresponding K light sources. These two sets of K pairs can be divided into four subsets with at least one pair. We firstly classify colors from a subset, within which spectral region the surface is illuminated from one direction with being remained no ambiguity for calculating effective reflectances up to a scalar-magnitude difference, and then solve or estimate the surface normal field on each color region. Figure 25 illus-

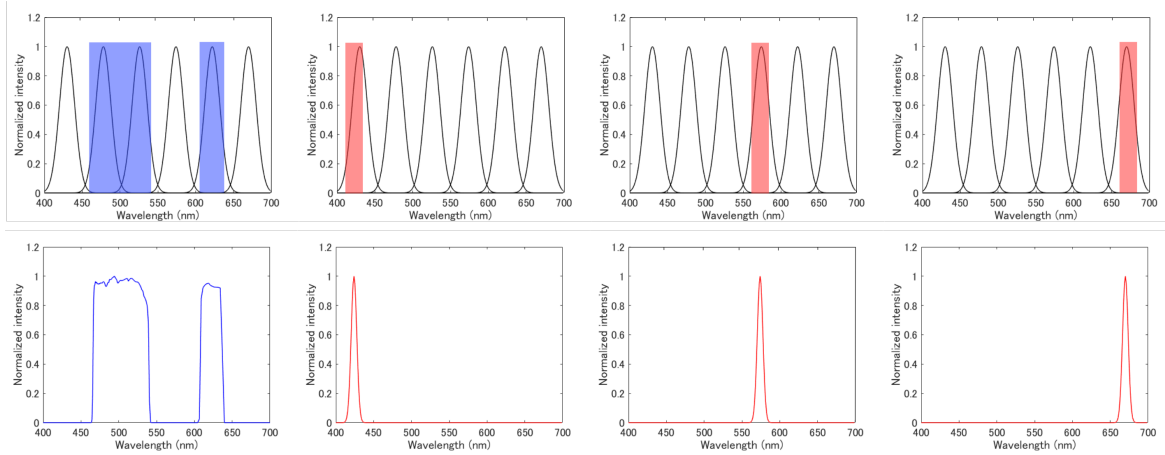


Figure 25: Simulated camera sensitivity functions and illumination spectra for 6-band photometric stereo.

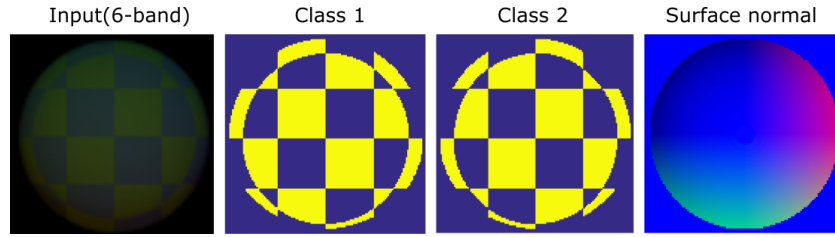


Figure 26: Results for the 6-band photometric stereo method: from left to right, input image rendered as an RGB image, color-class 1, 2, and estimated surface normal field.

trates an example of spectral sensitivity functions and light source spectra for $K = 6$ case: the first column for color classification and the others for surface normal estimation. In this case, we can use 3 from 6 channels for color classification and the four subsets can be used for estimating the surface normal field if the reflectance value within their channels are nonzero.

Figure 26 shows the color classification result and estimated surface normal field. In contrast to the case of RGB input, the surface was classified without ambiguity.

5.2 Multiband spectral regression in photometric stereo

What differentiates the two approaches in Chapter 3 and 4 is that while a certain part of a full reflectance spectrum has close interchannel correlation within itself, an RGB signal has less interchannel correlation and each channel value conveys color information quite independently: thus we introduced the color feature under the map that can be interpreted as one from normalized color reflectance to a color norm defined on naturally deduced surface normals over each same-colored surface region. This

is the lower bound case that enables to retrieve a surface normal at a point with the color being known there, and therefore required a smallest-observation oriented solution as shown in Chapter 4.

On the other hand, as in the hyperspectral case, bare color information in some bands are available without shading ambiguities if we are allowed to observe $K(> 3)$ -bands: the simplest is a direct $K - 2$ colors measurement and 3 independent shadings without interference between shadings resulting from overlapped camera sensitivities and illumination spectra. We will discuss an extension of our methods to a K -(multi)band case, especially on how we divide bands into subbands for shadings and reflectance measurement and how retrieve multiband reflectances accurately from the subbands for reflectance measurement enough to surface reconstruction. There still remain the other possibilities that both the surface normals and the colors are retrieved simultaneously, but such approaches would definitely assume geometric and chromatic conditions such as surface smoothness and clustered colors, and moreover would not prove an accurate retrieval. It must be mentioned that our approach using a hyperspectral camera would fail even if our assumptions hold, since the widths of notched bands affect the interpolation accuracy.

Despite the fact that a multiband image gives an underdetermined photometric stereo problem when simply placing somehow different illumination spectra, increasing the number of bands K gives more constraints on retrieving reflectance spectra. Provided that we have reflectance atoms or a basis set that perfectly represent as some linear combination any reflectance spectrum in interest, the original spectrum is retrieved from a sufficient number of independent observations. Even when such a spectral factorization is not given, the reflectance spectra in a scene embedded in this K -dimensional space are expected to distribute near around some points on the projective space determined by the K -dimensional color space projected at a band-axis; moreover, these points of embedded spectra are also expected to fall around a low-dimensional surface in the color space. What it requires is not to find widely studied embedding maps from the spectral space to a color space of some smaller number of bands, but more simply finding and attributing the correct multiband reflectance values onto each point of an unknown-colored surface: this results in just a simple discrete attribution problem.

We suggest two spectral estimation methods in a multiband photometric stereo problem: the first is to use a prior spectral dictionary, and the second is to use two capture but both of them are fully used for surface normal estimation; this is different from the other methods that used multiple frames with fixed surface normals assumption in that each capture of them has sufficient color information for surface normal estimation by unifying the two captures. Note that these alternative methods can be

realized with multiband cameras in principle, which makes our methods more feasible. Note that these frame-augmentation methods, since it uses two frames, deviates from our main topic as single-capture photometric stereo methods. However, these suggested methods do not assume constant surface normals over required frames. The first augmentation method provides in principle accurate spectral estimation for all N frames out of $N + 1$ frames, and the second method also does for each frame from every 2 frames provided the 2 frames. Although in addition these suggested methods require a technical complexity, they meet the essential purpose of this dissertation toward dynamic capture.

The first dictionary-augmentation requires an image captured under an illumination composed of a narrow-band light source bundle with some center wavelengths. Such an illumination that emit a spatially uniform spectral distribution over a target surface can be realized using a optical fibers or a diffuser in front of the light source bundle. This image can be used as a spectral (or multi-channel color) dictionary to refer to. For multiple shading monochromatic images embedded in a single multiband image, we choose two or more light sources of non-overlapping spectra, and locate them at different positions on the uniquely determined hemisphere with the origin at the surface and the light source bundle. Then the remained light source band works as a reflectance probe for a part of spectra with $K - 2$ or less channels, and the full reflectance spectra (or multi-channel reflectance values) can be estimated via the dictionary by e.g. a resemblance measure. It is generally difficult to represent arbitrary full-spectral information or high-dimensional spectral data in a lower dimensional space of $K - 2$ or less. But just so as to a surface we are to observe, the dictionary is expected to be used for an accurate indexing.

The second alternative method uses two frames taken under a pairwise somewhat different illumination every over-two-frame estimation of reflectances and surface normal fields. Again we assume a K -light source bundle with the support wavelength region Λ , and this time replicate $K_1 (\geq 2)$ -light source bundle with the support Λ_1 extracted from the K -light source bundle, the $k_1^{(p)}$ ($p = 1, 2, \dots, P$) of which being separately placed as $\sum_{p=1}^P k_1^{(p)} = K_1$ and another nonoverlapping $K_2 (\geq 2)$ -light source bundle with the support Λ_2 placed as the K_1 -bundle. A pairwise capture takes the first-frame image with turning on $(K - K_1)$ -light source bundle with the support $\Lambda \setminus \Lambda_1$ and the K_2 -bundle, and the second-frame image with turning on $(K - K_2)$ -light source bundle with the support $\Lambda \setminus \Lambda_2$ and the K_1 -bundle. Then, the wavelength region $\Lambda \setminus (\Lambda_1 \cup \Lambda_2)$ can be used for measuring reflectance, and the full K -spectral data are retrieved by concatenating observed data at pixels with the same $(K - K_1 - K_2)$ -multiband reflectance from the two frames. The low-dimensional

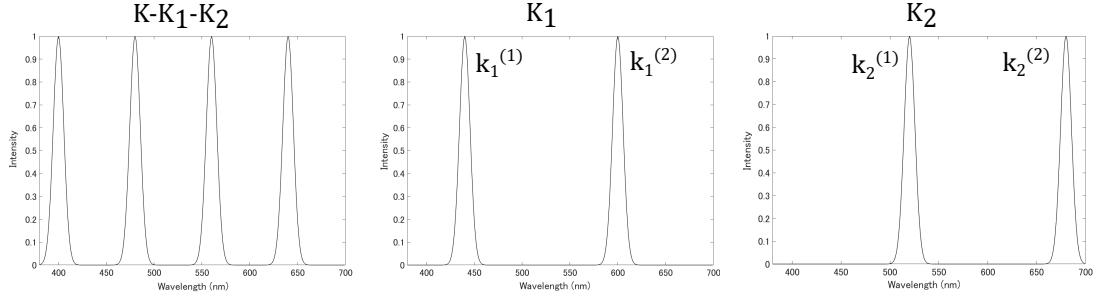


Figure 27: Illumination design for two-frame-based method.

assumption comes in with this procedure, but the spectral dictionary is generated while two sequential frames are captured. This allows a moderate temporal change in reflectance spectra and the usage of inter-pixel distances over two frames: two near pixel locations should have the common $(K - K_1 - K_2)$ -multiband reflectance. A possible multiband camera and illumination has $(K, K_1, K_2) = (8, 2, 2), (16, 2, 2), (16, 4, 4)$ and so on. The setup with $K_1, K_2 \geq 4$ enables to avoid zero reflectances by placing the $k_1^{(p)}$ and $k_2^{(p')}$ subbundles with a spectrally well-separated supports. Figure 27 shows an example of temporally tuned illumination spectra for the $(K, K_1, K_2) = (8, 2, 2)$ case.

5.3 Conclusion

In this dissertation, we developed two single-capture photometric stereo methods. Decreasing the required number of images is an important issue for applying photometric stereo to dynamic scenes. Color photometric stereo has been studied to obtain multiple shading images toward dynamic scene capture, but there have been remained theoretical and practical issues in the target surface colors as well as limited hard geometric assumptions. We alleviated these assumptions by firstly introducing a hyperspectral camera and spectrally and spatially designed illumination to estimate full-spectral reflectance spectra and surface normals from a single capture. Our second method used a single RGB digital camera and obtained color classification and surface normals simultaneously from a single RGB image. We discussed the color classification algorithm under moderate assumptions on surface geometry and surface colors. We also extended these methods to further increase the estimation accuracy or system feasibility by using a multiband camera, where surface classification accuracy was increased or the possibility that reflectance spectra can be estimated by regression.

References

- [1] D. Eigen, C. Puhrsch, and R. Fergus. Depth map prediction from a single image using a multi-scale deep network. In *Advances in Neural Information Processing Systems*, 2366-2374 (2014).
- [2] Steven M. Seitz, Brian Curless, James Diebel, Daniel Scharstein, and Richard Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *IEEE Conference on Computer Vision and Pattern Recognition*, 519-528 (2006).
- [3] R. Zhang, P. Tsai, J. Cryer, and M. Shah. Shape from shading: A survey. In *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(8), 690-706 (1999)
- [4] R. J. Woodham. Photometric Stereo: A Reflectance Map Technique for Determining Surface Orientation from Image Intensity. In *Proc. SPIE*, 155, 136-143 (1978).
- [5] F. Langguth, K. Sunkavalli, S. Hadap, and M. Goesele. Shading-aware multi-view stereo. In *European Conference on Computer Vision*, (2016).
- [6] J. T. Kajiya. The Rendering Equation. In *SIGGRAPH*, 143-150 (1986).
- [7] T. Papadhimetri and P. Favaro. Uncalibrated near-light photometric stereo. In *British Machine Vision Conference*, (2014).
- [8] S. Lee and M. Brady. Integrating stereo and photometric stereo to monitor the development of glaucoma. In *Image and Vision Computing*, 9, 39-44 (1991).
- [9] B. Shi, Z. Wu, Z. Mo, D. Duan, S.-K. Yeung, and P. Tan. A benchmark dataset and evaluation for non-Lambertian and uncalibrated photometric stereo. In *IEEE Conference on Computer Vision and Pattern Recognition*, (2016).
- [10] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *Asian Conference on Computer Vision*, (2010).
- [11] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa. Photometric stereo using sparse Bayesian regression for general diffuse surface. In *IEEE Trans. Pattern Anal. Mach. Intell.*, 36(9), 1816-1831 (2014).

- [12] R. Mecca, Y. Queau, F. Logothetis, R. Cipolla. A Single Lobe Photometric Stereo Approach for Heterogeneous Material. In SIAM Journal on Imaging Sciences, 9(4), 1858-1888 (2016).
- [13] K. Ikeuchi. Determining a depth map using a dual photometric stereo. In International Journal of Robotics Research, 6(1), 15-31 (1987).
- [14] R. Anderson, B. Stenger, and R. Cipolla. Color Photometric Stereo for Multicolored Surfaces. In IEEE International Conference on Computer Vision, 6-13 (2011).
- [15] A. Wenger, A. Gardner, C. Tchou, J. Unger, T. Hawkins, and P. Debevec. Performance Relighting and Reflectance Transformation with Time-Multiplexed Illumination, In SIGGRAPH, 756-764 (2005).
- [16] Z. Janko, A. Delaunoy, and E. Prados. Colour Dynamic Photometric Stereo for Textured Surfaces. In Asian Conference on Computer Vision, 55-66 (2010).
- [17] H. Kim, B. Wilburn, and M. Ben-Ezra. Photometric Stereo for Dynamic Surface Orientations. In European Conference on Computer Vision, 59-72 (2011).
- [18] T. Kurihara, N. Ono, S. Ando. Single-frame surface orientation imager using triangular three-phase amplitude-modulated illuminations and correlation image sensor (in Japanese). In Trans. SICE, 48(8), 505-513 (2012).
- [19] A. Ando and A. Kimachi. Correlation image sensor: Two-dimensional matched detection of amplitude modulated light. In IEEE. Trans. Electron Devices, 50-10, 2059-2066 (2003).
- [20] C. Chen, D. Vaquero, M. Turk. Illumination Demultiplexing from a Single Image. In IEEE International Conference on Computer Vision, 17-24 (2011).
- [21] M. Smith and L. Smith. Dynamic photometric stereo: A new technique for moving surface analysis. In Image and Vision Computing, 23, 841-852 (2005).
- [22] G. Atkinson, M. F. Hansen, M. Smith, and L. Smith. A efficient and practical 3D face scanner using near infrared and visible photometric stereo. In Proceedings of the International Conference and Exhibition on Biometrics Technology, 11-19 (2010).
- [23] G. Fyffe, X. Yu, P. Debevec. Single-shot photometric stereo by spectral multiplexing. In IEEE International Conference on Computational Photography, 1-6 (2011).

- [24] P. P. Nikolaev. Algorithms for recognition of surface color. In *Modeling of Learning and Behavior*, 121-151 (1975).
- [25] M. H. Brill. Image segmentation by object color: a unifying framework and connection to color constancy. In *Journal of the Optical Society of America A*, 7, 2041-2047 (1990).
- [26] A. P. Petrov. Color and Grassmann-Cayley coordinates of shape. In *Human Vision, Visual Processing, and Digital Display II*, 342 (1991).
- [27] M. S. Drew. Shape from color. In *Technical Report CSS/LCCR TR 92-07*, School of Computing Science, Simon Fraser Univ. (1992).
- [28] L. L. Kontsevich, A. P. Petrov, and I. S. Vergelskaya. Reconstruction of shape from shading in color images. In *Journal of the Optical Society of America A*, 11(3), 1047-1052 (1994).
- [29] M. S. Drew and L. L. Kontsevich. Closed-form attitude determination under spectrally varying illumination. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, 985-990 (1994).
- [30] A. P. Petrov and L. L. Kontsevich. Properties of color images of surfaces under multiple illuminants. In *Journal of the Optical Society of America A*, 11(10), 2745-2749 (1994).
- [31] M. S. Drew and M. H. Brill. Color from Shape from Color: A Simple Formalism with Known Light Sources. In *Journal of the Optical Society of America A*, 17(8), 1371-1381 (2000).
- [32] A. Chakrabarti and K. Sunkavalli. Single-image RGB Photometric Stereo With Spatially-varying Albedo. In *Proc. of the IEEE International Conference on 3D Vision* (2016).
- [33] Y. Xiong, A. Chakrabarti, R. Basri, S. J. Gortler, D. W. Jacobs, , and T. Zickler. From shading to local shape In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(1), 67-79 (2015).
- [34] K. Ozawa, I. Sato, and M. Yamaguchi. Hyperspectral photometric stereo for a single capture. In *Journal of the Optical Society of America A*, 34(3), 384-394 (2017).
- [35] A. D. Spence and M. J. Chantler. Optimal illumination for three-image photometric stereo acquisition of texture. In *Texture 2003: Proc. of the 3rd international workshop on texture analysis and synthesis*, 89-94 (2003).

- [36] O. Drbohlav and M. J. Chantler. On Optimal Light Configurations in Photometric Stereo In Proc. International Conference on Computer Vision (2005).
- [37] W. K. Pratt and C. E. Mancill. Spectral estimation techniques for the spectral calibration of a color image scanner. In *Applied Optics*, 15(1), 73-75 (1976).
- [38] J. P. S. Parkkinen, J. Hallikainen, T. Jaaskelainen. Characteristic spectra of Munsell colors. In *Journal of the Optical Society of America A*, 6(2), 318-322 (1989).
- [39] D. H. Marimont and B. A. Wandell. Linear models of surface and illuminant spectra. In *Journal of the Optical Society of America A*, 9, 1905-1913 (1992).
- [40] M. Kitahara, T. Okabe, C. Fuchs, and H. P. A. Lensch. Simultaneous Estimation of Spectral Reflectance and Normal from a Small Number of Images. In *Proc. International Conference on Computer Vision Theory and Applications*, 303-313 (2015).
- [41] A. Sohaib, A. R. Farooq, G. A. Atkinson, L. N. Smith, M. L. Smith, and R. Warr. In vivo measurement of skin microrelief using photometric stereo in the presence of interreflections. In *Journal of the Optical Society of America A*, 30(3), 278-286 (2013).
- [42] L. Gao, R. T. Kester, N. Hagen, T. S. Tkaczyk. Snapshot Image Mapping Spectrometer (IMS) with high sampling density for hyperspectral microscopy. *Optics Express*, 18(14), 14330-14344 (2010).
- [43] M. Hauta-Kasari, K. Miyazawa, S. Toyooka. Spectral vision system for measuring color images. In *Journal of the Optical Society of America A*, 16(10), 2352-2362 (1999).
- [44] Pharr, Humphreys. *Physically Based Rendering*, Second Edition (Morgan Kaufmann, 2010).
- [45] R. T. Frankot and R. Chellappa. A Method for Enforcing Integrability in Shape from Shading Algorithms. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4), 439-451 (1988).
- [46] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing* (Springer, 2010).
- [47] T. Zhou and T. Tao, gGo Decomposition (GoDec): Randomized Lowrank and Sparse Matrix Decomposition in Noisy Case, In the International Conference on Machine Learning, (2011).

- [48] H. Zhang, W. He, L. Zhang, H. Shen, Q. Yuan. Hyperspectral Image Restoration Using Low-Rank Matrix Recovery. In IEEE Transactions on Geoscience and Remote Sensing, 52(8), 4729-4743 (2014).
- [49] M. Fazel, E. Candes, B. Recht, P. Parrilo. Compressed Sensing and Robust Recovery of Low Rank Matrices In Asilomar Conference on Signals, Systems and Computers, 1043-1047 (2008).
- [50] S. Roweis. EM Algorithms for PCA and SPCA. In Advances in neural information processing systems, 626-632 (1997).
- [51] Z. Lin, M. Chen, L. Wu, and Y. Ma. The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices. In arXiv preprint arXiv: 1009.5055 (2010).
- [52] K. Ozawa, I. Sato, and M. Yamaguchi. Single color image photometric stereo for multi-colored surfaces. In Computer Vision and Image Understanding, <https://doi.org/10.1016/j.cviu.2018.04.003> (2018).
- [53] Woodham, R.J., Iwahori, Y, and Barman, R.A., “Photometric Stereo: Lambertian reflectance and light sources with unknown direction and strength,” Technical Report TR 91-18, University of British Columbia, Department of Computing Science (1991).
- [54] N. G. Alldrin, S. P. Mallick, D. J. Kriegman. Resolving the generalized bas-relief ambiguity by entropy minimization. In Computer Vision and Pattern Recognition, 1-7 (2007).
- [55] J. DfErrico. inpaint nans, MATLAB Central File Exchange. Retrieved 13 Aug, (2012).

Publication list

論文誌（査読あり）

- 1) Keisuke Ozawa, Imari Sato, and Masahiro Yamaguchi, “Hyperspectral photometric stereo for a single capture,” Journal of the Optical Society of America A, vol. 34, Issue 3, pp. 384-394 (2017).
- 2) Keisuke Ozawa, Imari Sato, and Masahiro Yamaguchi, “Single color image photometric stereo for multi-colored surfaces” Computer Vision and Image Understanding, <https://doi.org/10.1016/j.cviu.2018.04.003>, (2018).

学会発表（査読なし）

- 1) ハイパースペクトル画像を用いた照度差ステレオ法（電子情報通信学会、電子ディスプレイ研究会 2016 年秋）
- 2) 多色物体に対する単一カラー画像照度差ステレオ法（日本光学会年次学術講演会 2017 年、11 月 2 日）

本研究にあたりお世話になった方々にお礼申し上げます。

物理情報システム専攻の佐藤いまり先生には、研究の機会と貴重な機器に出会わせていただきました。山口雅治先生は投稿論文執筆にあたり丁寧にご指導下さり、カメラの準備やセッティングまで助けて下さいました。その際、山口研究室の Syukran Hakim さんと五十嵐俊亮さんにも手をお貸しいただき、皆様の親切なサポートのおかげで円滑に実験を進めることができました。また、中村祐哉先生は、比較実験にあたり GPUPC を自由にに使わせて下さったばかりか、査読の遅れで不安に過ごす私を気遣って様々なアドバイスを下さいました。参加させていただいた山口研究室でのゼミでは、私の基本的な質問に対しても皆様が丁寧に答えて下さり、勉強になるだけでなく、嬉しい時間を過ごさせていただきました。本論文の審査にあたり、審査員の小尾高史先生、金子寛彦先生、熊澤逸夫先生には貴重なご意見をいただきました。エバ・ジャパン株式会社の野呂直樹様、高良洋平様、安藤史識様には、ハイパースペクトルカメラの取り扱いに関する丁寧なサポートをいただきました。個人的にはありますが、光源準備にあたり、新光電子株式会社様と LED Ecology WebShop 様にも多大な助けをいただきました。改めて皆様にお礼申し上げます。

東京工業大学には長きにわたり大変お世話になりました。本研究との直接の関わりではありませんが、学部時代に師事させていただいた材料物理学専攻の細野秀雄先生には、私にとっての研究者の理想を見せていただきました。そして、分野を変えての博士後期課程進学に際しては、気持ちを前向きに固めるきっかけを下さいました。入学時、色々な勉強をするようにと投げかけられたお言葉を、実現できる限りで大切にしてきたつもりです。少しずつですが今は実りを得はじめ、学業、研究上での新たな出会いにも恵まれました。そして私が知るわずかばかりの、本の中だけでない光学の生きた知識は全て、材料物理学専攻（当時）の山本直紀先生から教えていただいたものです。山本研究室の皆様にはいくら感謝してもしきれないと思います。また、セラミックスとガラスのことを教えてくださった無機材料工学科の先生方にお礼申し上げます。大学での学びの思い出として、吉野太郎先生は学部数年間にわたり、私の図々しい質問に対し、毎回貴重なお時間を割いて丁寧に答えて下さり、そのたびに知的な喜びをいただきました。少しですが授業でお話しさせていただいた細谷暁夫先生には大きな励みと驚きをいただきました。蓄積された出会いの恵みを忘れずに勉強し続けたいと思います。

残るお名前を全て挙げることはできませんが、佐藤研究室、山口研究室、事務局の皆様、無機材料工学科の友人達、ピアサポート、大学図書館、人文図書館、数学相談室、大学生協、構内のあらゆるサポートをしてくださる皆様方にお礼申し上げます。

最後に、大学生生活と研究活動をあらゆる面で支えてくれた家族に感謝します。