Doctoral Dissertation

Physics-based Image Decompositions for Obtaining a Shape and Visual Textures

Tsuyoshi Takatani

August 2, 2019

Graduate School of Information Science
Nara Institute of Science and Technology

A Doctoral Dissertation
submitted to Graduate School of Information Science,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of ENGINEERING

Tsuyoshi Takatani

Thesis Committee:

Professor Yasuhiro Mukaigawa	(Supervisor)
Professor Hirokazu Kato	(Co-supervisor)
Associate Professor Takuya Funatomi	(Co-supervisor)
Assistant Professor Kenichiro Tanaka	(Co-supervisor)
Assistant Professor Hiroyuki Kubo	(Co-supervisor)

Physics-based Image Decompositions for Obtaining a Shape and Visual Textures*

Tsuyoshi Takatani

Abstract

Measuring a shape and visual textures of an object is an important task in computer vision and graphics. Most methods aiming at measuring them often employ a physics-based approach because it recovers an accurate shape and the visual textures, *e.g.*, color and translucency, are derived from optical phenomena. For example, photometric stereo is a well-studied technique to obtain detailed surface orientations of an object. However, since photometric stereo assumes diffuse reflection on the surface, it is difficult to be applied for a translucent object in which complex optical phenomena, *e.g.*, subsurface scattering, occur. Such assumptions in physics-based measurements have limited opportunities of their practical usages. A possible solution is an image decomposition to extract a specific component. If a physics-based image assumed by a measurement method can be extracted, the method is expected to work properly on any materials. Additionally, a physics-based image decomposition enables to obtain visual textures which cannot be obtained by the existing measurement methods.

This thesis provides physics-based image decomposition techniques toward obtaining a shape and visual textures of an object. In general, a detectors observes a radiant intensity from the object and the observed intensity essentially includes a wide variety of physics-based components, such as optical and thermal phenomena. That is why we aim at decomposing the observed intensity into physics-based components which are assumed by methods to measure a shape and visual textures. Since the radiant intensity is a sum of various light rays, the

*Doctoral Dissertation, Graduate School of Information Science, Nara Institute of Science and Technology, August 2, 2019.

physics-base image decomposition can be conducted in spatial, temporal, and spectral perspectives.

In the spatial perspective, a unified decomposition method, multiple weighted measurements, is proposed to combine different decomposition approaches. Experimental results show that the proposed method can decompose observed images into physics-based images, such as diffuse and specular reflection and single and multiple scattering. The decomposition result is applied to enable a novel image segmentation in which a segment is based on visual textures, *e.g.*, translucency and opacity. The translucency of a translucent material is mostly derived from subsurface scattering. Since subsurface scattering spatially blurs light in an object, an incident spot light is blurred and its outgoing light becomes a distribution, called point spread function (PSF). A conventional camera observes a sum of PSFs at all surface points but the PSF at each surface point describes the translucency at the point. Therefore, the PSFs are decomposed into one at each point using projector-camera (pro-cam) system. Measured PSFs are applied for physically reproducing the translucency of a material by a UV inkjet printer. In addition to the image plane axes, a decomposition along the optical principle axis is proposed to observe inside of an object. a combination of spatio-sequentially modulated illumination with direct conversion enables to extract a continuous series of layer images in an object. Experimental results show that the proposed method decomposes observations into a series of layer images in a semi-transparent object.

In the temporal perspective, far infrared (FIR) light transport decomposition is proposed based on optical and thermal phenomena. Light absorbed by matters in an object transforms to heat and then conducts inside the object. This thermal phenomena is so slow that it can be observed by a conventional thermal camera at a video framerate, while optical phenomena, *e.g.*, reflection of light, is too fast to be observed at the video framerate. Thus, a temporal analysis brings a decomposition of optical and thermal phenomena under controlled illumination. Absorption is also related to both a surface normal and a light direction as well as diffuse reflection. Thus, photometric stereo can be applied with the decomposed thermal component. More importantly, since most materials absorb and radiate FIR light, it can be used for various materials, such as translucent glass,

translucent plastic, and black object.

In the spectral perspective, a novel one-shot hyperspectral imaging using faced reflectors is proposed to obtain spectral reflectance. A coupled mirror on which a color filter is sticked is mounted in front of a conventional camera. Then, the camera observes multi-bounce images whose spectra are different due to the color filter at one-shot. At last, a hyperspectral image is reconstructed from the multi-bounce images. Optical phenomena vary with respect to wavelength of light. For example, light of longer wavelengths penetrates deeper inside of an object than one of shorter wavelengths. Thus, analyzing spectral images reduces the effect of subsurface scattering in photometric stereo. Experimental results show that the use of spectral images makes photometric stereo more robust to translucent materials and multi-textured objects.

Throughout this thesis, physics-based image decompositions are proposed in spatial, temporal, and spectral perspectives toward obtaining a shape and visual textures of an object. The proposed methods enable to obtain surface orientations of an object with various materials, such as multi-textured, translucent, transparent, and black objects, and various visual textures, such as optical feature, translucency, inner structure, and spectral reflectance.

Keywords:

image decomposition, optical phenomena, spectral analysis, translucency, photometric stereo

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Light interacts with matters and it causes a wide variety of complex optical phenomena. As light hits a ripe apple, light of longer wavelengths in the visible spectrum is strongly reflected off surface of the apple and thus we see that it is a red apple. Small droplets of water in the atmosphere scatter light of all wavelengths in the visible spectrum, which makes clouds look white in the sky. Refraction through raindrops disperses sunlight and it produces a rainbow. These complex optical phenomena are preferred for photographic arts due to their beauty. However, the complex optical phenomena often disturb photometric measurements in computer vision.

The photometric measurement is a physics-based technique to obtain a variety of information about a target object, *e.g.*, shape and visual textures. In general, it recovers an accurate shape comparing with geometric methods, *e.g.*, multi-view stereo [5]. That is why the photometric measurement is practical and applicable for an industrial use. But the photometric measurements, in general, assume simple optical phenomena in a scene, which limits opportunities of its practical use. Thus, the photometric measurements in a scene of the complex optical phenomena are not necessarily guaranteed a proper performance. For example, photometric stereo [178], a well-known method to reconstruct surface orientations, assumes diffuse reflection on the surface. It is difficult to be applied for glossy and translucent materials because unexpected optical phenomena occur in/on those materials, *e.g.*, specular reflection and subsurface scattering. In order to overcome such a problem, a lot of reflectance models for photometric stereo have

been proposed to take account of the complex optical phenomena [155, 15, 44, 52]. Although the reflectance models could be applied not only for photometric stereo but also for other photometric measurements, *e.g.*, measurement of bi-directional reflectance distribution function (BRDF), it is required to re-formulate them for a specific case, and this procedure is not a trivial task. Another effective approach is an image decomposition or separation [105, 124, 97]. Since an observed intensity includes various components, extracting only an expected component from the observed intensity arrows a photometric measurement to work properly without the re-formulation. Intrinsic image decomposition [14] has been studied in computer vision to separate a shading image and a reflectance image. In general, Lambertian reflectance, *i.e.*, ideal diffuse reflection, is assumed in the intrinsic image decomposition problem and the separated images are not based on physical phenomena. However, the photometric measurements are based on physical phenomena. Therefore, physics-based image decompositions are addressed in this thesis. Additionally, physics-based components can explain and obtain visual textures, *e.g.*, color and translucency.

The main topic in this thesis is physics-based decompositions. Target physics-based components can be derived from optical, thermal, mechanistic, electric or magnetic phenomena. In this thesis, we target on the optical and thermal phenomena because our final application is photometric measurements and optical absorption is related to thermal radiation. In order to observe the phenomena as an image, an optical camera and a thermal camera are easily used, while it is difficult to observe the mechanistic, electric, and magnetic phenomena. A detector observes a radiant intensity in a scene and the intensity essentially includes a wide variety of physics-based components, *e.g.*, the complex optical phenomena. Since the radiant intensity is a sum of light rays, physics-based image decompositions can be performed in spatial, temporal, spectral, and directional perspectives due to the plenoptic function. In this thesis, we target on the spatial, temporal, and spectral perspectives.

Once physics-based image decompositions are done, it is possible to use a single physics-based component for the photometric measurements. In this thesis, we apply the decompositions for measuring a shape and visual textures of an object. Photometric stereo reconstructs surface orientations based on diffuse

reflection, as mentioned above, and the surface orientations can be regarded as the shape. As visual textures, spectral reflectance, *i.e.*, color, translucency, and glossiness are obtained via physics-based image decompositions. An inner layers of the object can be a visual texture and also be shapes of matters in the object. Obtaining the inner layers is important for inspections of human bodies and foods. Recognizing materials is also important for object recognition problem because a computer has to understand materials of an object when a robot picks it up. Otherwise, the robot might crush the object with its strong power.

This thesis provides physics-based image decompositions toward obtaining a shape and visual textures of an object. Target physics-based components are optical and thermal phenomena in the spatial, temporal, and spectral perspectives. The contributions of this thesis are achieving the ability to;

- spatially decompose observations into components based on optical phenomena, such as diffuse and specular reflection and single and multiple scattering;
- spatially decompose observations into subsurface scattering components and then physically reproduce the translucency of a material by a UV printer;
- spatially decompose observations into inner layer images of a semi-transparent object along the optical principle axis;
- temporally decompose thermal observations into diffuse and global far infrared radiation components and then innovate thermal photometric stereo; and
- observe a hyperspectral image at one-shot with a readily available system and then improve the robustness of photometric stereo.

1.1. Physics-based components

When illuminating an object, light interacts with matters which consist of the object. The interaction is basically categorized to five optical phenomena; reflection, transmission, scattering, absorption, and fluorescence. Some light reached to a

boundary between the air and the object is reflected off the boundary. Reflection of light is often categorized to two phenomena; diffuse and specular reflection. Diffuse reflection is a phenomenon in which light is reflected to all directions regardless of an incident angle. Specular reflection is one in which light is reflected to the regular direction to the incident angle. On surfaces of a concave object, reflected light is again reflected off the boundary, which is called interreflections. The remaining of light penetrates into the object. If a material of the object is transparent, most of it transmits through the other-side boundary. The light transporting inside of the object extincts along its light path because of scattering and absorption. Scattering of light is caused by collisions with particles in the object. Light which is scattered once with a particle is categorized to single scattering and light of multiple collisions is called multiple scattering. Especially, scattering beneath the boundary is called subsurface scattering. Light energy absorbed by the matters excites molecules in the object and then some of it transforms and emits light again but its wavelength becomes longer. This phenomenon is called fluorescence. Other of the absorbed light transforms to heat via vibration of molecules, called the photothermal effect, and then heat conducts in the object. All thermal objects radiate far infrared (FIR) light with respect to its temperature. Thus, the brightness of FIR light radiated from the object temporally vary with respect to the heat conduction.

Additionally, the photoelectric and photoacoustic effects occur depending on the matters. The photoelectric effect is the emission of electrons or to produce a current when light interacts with the matters. This effect has been applied for a digital camera and solar cells. The photoacoustic effect is to produce sounds because of the vibrations of matters when light is absorbed. This effect is applied for imaging inner layers deeply, *e.g.*, an observation of arteries.

1.2. Photometric stereo

Photometric stereo [178, 149] is a technique to estimate surface orientations from multiple images taken under different lighting conditions. A key factor in this process is the surface reflectance, which describes how the shading at each surface point changes in relations to the lighting direction and the surface normal. By

varying the light direction, the corresponding changes in the shading are used to infer the surface normals according to the reflectance model.

Surface reflectance is normally assumed to follow Lambert's law, whereby the intensity I of the reflected light is proportional to the inner product of the lighting vector $\boldsymbol{l}$ and the surface normal vector $\boldsymbol{n}$:

$$I(x) = \rho(x)\boldsymbol{n}^{\mathrm{T}}\boldsymbol{l}, \tag{1.1}$$

where $\rho(x)$ denotes the albedo, or intrinsic color, of the surface at point x. When changing the light direction, multiple observations can be obtained that can be superposed in a matrix form as

$$\boldsymbol{I} = \rho(x)\mathbf{L}\boldsymbol{n}, \tag{1.2}$$

where $\boldsymbol{I}$ and $\mathbf{L}$ are the superposed observation vector and lighting matrix, respectively. If the lighting matrix is full-rank, then the surface normal vector can be obtained as

$$\boldsymbol{n} = \frac{\mathbf{L}^{\dagger}\boldsymbol{I}}{\|\mathbf{L}^{\dagger}\boldsymbol{I}\|_2}, \tag{1.3}$$

where $\mathbf{L}^{\dagger}$ is the pseudo-inverse matrix of $\mathbf{L}$. This Lambertian model of reflectance is frequently used in photometric stereo because of its simplicity and convenience. Since photometric stereo can obtain micro-asperity of the surface, it is often applied for inspections on a factory production line, archiving oil paintings, and scanning human skin.

1.3. Visual textures

When we see an object we can feel how soft, hard, warm, cold, smooth, or rough the object is from visual textures of the object. According to the recent study about textures [183], factors for our visual feeling are shape, color, position, surface reflectance, glossiness, translucency, micro-surface structure, and viscosity. The color of the object is described as a spectrum in this thesis. The optical phenomena explained in Section 1.1 affect some of the visual textures. Diffuse and specular reflection derives the surface reflectance and glossiness. Especially, the glossiness is strongly related to specular reflection. Subsurface scattering

including single and multiple scattering produces the translucency, which is obviously important to render a photorealistic graphics. The micro-surface structure, *e.g.* fur and asperity, causes the directional difference in reflectance. The viscosity is a dynamic visual texture and difficult to be taken account of in the shape and material measurement using static images. In this thesis, the spectral reflectance, glossiness, and translucency are focused on as important visual textures.

1.4. Overview

The goal of this thesis is to propose physics-based decompositions in the spatial, temporal, and spectral perspectives toward obtaining a shape and visual textures. A key question is what to be decomposed and how to use the components when using a photometric measurement. In Chapter 3, a unified decomposition method is proposed to spatially decompose observations into components based on the optical phenomena. In experiments, the proposed method decomposes into four different optical components; diffuse and specular reflection and single and multiple scattering. Since the intensity ratio of the components varies with respect to a material, a decomposition result is applicable to classify different materials in an image. Chapter 4 establishes a system to physically reproduce the translucency of an object. In the measurement step, PSF at each surface point is obtained via a spatial decomposition using modulated illumination. In the fabrication step, a translucency similar to the measured one is physically reproduced by UV printing, finding the best print parameters in a lookup table between the parameters and the printed translucency. In Chapter 5, a novel method to observe inside an object is proposed. Spatio-sequential modulated illumination enables to decompose observations into inner layer images. It allows us to see the inner structure of the object, which affects the visual texture. In Chapter 6, far infrared spectrum, much longer wavelengths than the visible spectrum, is focused on. A novel method to decompose far infrared light transport is proposed including the photo-thermal effect. A fact is found out, that a radiation component is available for photometric stereo. Since almost all materials can be observed within the far infrared spectrum, it is possible to apply photometric stereo for various materials, such as transparent, translucent, and black objects. Chapter 7

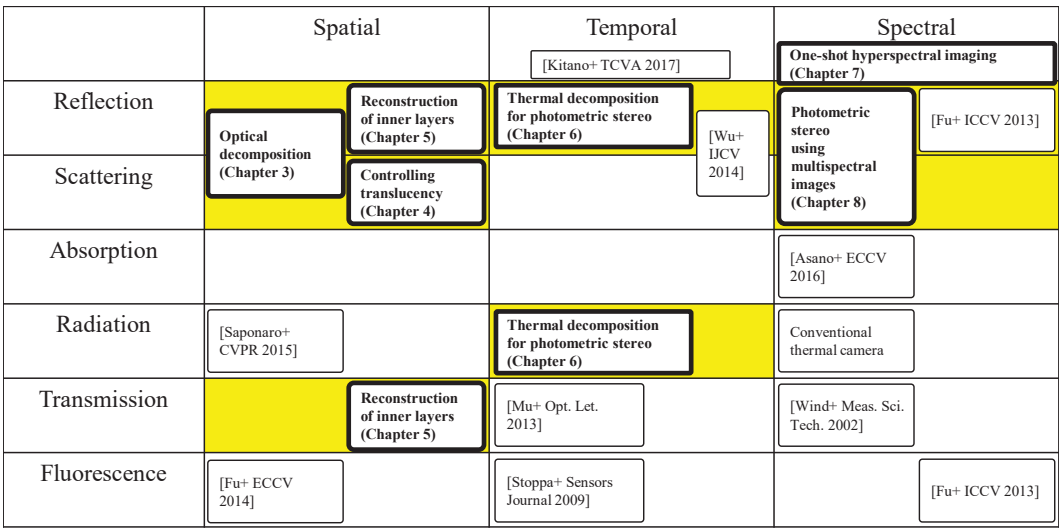


Figure 1.1: The relationship among methods proposed in this thesis.

leads spectral imaging. Decomposing into each wavelength image enables to analyze materials and optical phenomena. Although the existing system are costly and time-consuming, a novel hyperspectral imaging technique is proposed, which can be implemented with readily available mirrors and a color filer and enables to obtain a hyperspectral image at one-shot. Chapter 8 introduces a novel technique to enhance photometric stereo using multispectral images. As explained above, subsurface scattering depends on wavelength of light. That is, some wavelength can be less affected by subsurface scattering in a material. Therefore, spectral analysis enables to reduce the effect of subsurface scattering and improve the accuracy of photometric stereo.

The relationship among methods proposed in this thesis is explained in Fig. 1.1. The physics-based decompositions are categorized into the spatial, temporal, and spectral perspectives, which is the horizontal axis on a table in the figure. The vertical axis describes physical phenomena to be decomposed. Yellow cells in the table are targets for the proposed methods.

Chapter 2

Related work

In this chapter, work related to this thesis is introduced. The main topic of this thesis is image decomposition and photometric measurements of a shape and visual textures.

2.1. Spatial decomposition

Researchers in computer vision and graphics have studied to separate images to extract or remove some physics-based components from observed images. The word expressions are different but they essentially mean the same as decomposition. Target components are different with respect to an application.

The first interest is in reflection of light. Shafer [146] has proposed the dichromatic reflectance model, in which the color of specular reflection depends on the color of a light source while the color of diffuse reflection depends on the color of an object. And then, a lot of work employed the dichromatic reflectance model to separate the diffuse and specular reflection components [89, 138, 139, 156, 83, 127, 182, 6, 136]. Another effective technique to separate the reflection components is based on polarization. Wolff and Boult [176] utilized linear polarization to remove the specular reflection component of the observed image. Many researchers also used linear or circular polarization [115, 25, 102, 46]. Both of the methods based on color and polarization can be combined since both are in complementary relationship [121, 98, 82, 168, 172]. Moreover, other clues are used to separate the diffuse and specular reflection components. Ikeuchi and Sato [72] used both a

range and brightness images. Nishino *et al.* [128] assumed a known geometry of an object to separate view-independent, as diffuse, and view-dependent, as specular, components. Mukaigawa *et al.* [113] analyzed the reflection components based on photometric linearization. Mallick *et al.* [104] formulated a decomposition model based on locally spatial and spatio-temporal interactions. Tao *et al.* [161] used line consistency based on relationship between light field data and the dichromatic model. Interreflections are often targeted as other reflection component, which is a phenomenon of multiple reflections within a scene. Seitz *et al.* [145] proposed a theory of inverse light transport to separate interreflections into each bounce component. Bai *et al.* [11] developed a duality theory of forward and inverse light transports and then separated interreflections. Fu *et al.* [38] proposed a separation method of interreflections from a single image using fluorescence. A lot of decomposition work for optical phenomena has been introduced above. Another approach is intrinsic image decomposition, which is to separate an image into two intrinsic components; a reflectance image and a shading image [14]. The reflectance image contains intrinsic colors of surface points which are independent of an illumination, while the shading image includes various effects of lighting, such as shadows, specular highlights, and shading. This topic is related to that of this thesis but the intrinsic image is different from the physics-based image, which is based on physical phenomenon, such as reflection and scattering of light.

Light scattering is often regarded as a component to be removed because it disturbs the measurement methods. Gilbert and Pernicka [48] removed the single scattering component in water by using circular polarization. Many polarization-based methods were proposed for removing scattering component toward clear appearance in hazy atmosphere [142] and muddy water [165]. Ghosh *et al.* [47] modeled the layered facial reflectance consisting of specular reflection, single scattering, and shallow and deep subsurface scattering components by using a polarization-based method to achieve high quality rendering. Kim *et al.* [85] fused the polarization technique with a light field camera to decompose specular reflection, single scattering, and scattering in different layer components. Narasimhan and Nayar [119, 120] analytically modeled light scattering in atmosphere and then proposed a method to remove scattering components of fog and haze. Wu and Tang [181] decomposed the diffuse and specular reflection and subsurface

scattering components based on the model proposed by Lin and Lee [99]. Nayar *et al.* [122] proposed an effective method to fast separate direct and global illumination components, called high frequency illumination. Gupta *et al.* [55] combined high frequency illumination with the polarization technique to remove scattering components. Mukaigawa *et al.* [114] extended high frequency illumination to separate the single and multiple scattering components. Decomposition also plays an important role for understanding optical phenomena. Fuchs *et al.* [40] employed confocal imaging for descattering. Kim *et al.* [84] removed the scattering components by analyzing light field data. Various motivations exist for decomposition methods. Light scattering causes an unclear image in atmosphere and water. Many methods proposed to remove the effect of scattering and obtain a clear image [48, 119, 120, 142, 77]. The removed scattering component can be use for reconstructing a depth map [31]. The effect of scattering depends on the distance. Thus, the distance can be estimated once the scattering component is extracted. The scattering component plays an important role in such a method, while that component is often regarded as an obstacle.

2.2. Temporal decomposition

Analyzing physics-based images in the time domain is a relatively new topic in computer vision but important to directly understand interreflections and scattering of light. Wu *et al.* [179] developed a method to decompose diffuse reflection, inter-reflection, and subsurface scattering based on the temporal response of a femtosecond-pulsed laser. Interferometer [49] and holography [80] are also used for time-resolved decomposition of images. As we can imagine, light reflected once in a scene after emission from a system, which has a light source and a detector, is arriving at the detector faster than light reflected twice. Resolving the multi-path problem in the time-of-flight camera is an active research topic and it has been studied by assuming the two-bounce or simplified reflection models [41, 30, 51, 76], K-sparsity [18, 37, 134], parametric model [64, 87], consistency between ToF and stereo [94], simplified indirect reflections [117], and large-scale multi-path [78]. Direct and indirect light transport can also be decomposed by the time-resolved approach [130, 56]. The temporal decomposition approach can

be used to recover the shape of transparent and translucent objects [148, 157]. Decomposing interreflection components reveals a response signal from a scene which cannot directly be seen by a camera, which is applied for non-line-of-sight imaging [170, 64, 78, 167]. Saponaro *et al.* [137] used the FIR radiation component to classify materials. When heating or cooling a material, a spatio-temporal observation of the radiation characterizes the material property. Mu *et al.* [112] employed a temporal measurement using single-photon avalanche diode (SPAD) to extract directly transmitted light. Stoppa *et al.* [152] also used SPADs for time-resolved fluorescence measurements.

2.3. Spectral decomposition

Traditional hyperspectral imaging employs a number of narrow band filters [162], a tunable narrow band filter [42, 109], and diffractive media [67, 32]. In general, methods based on a narrow band filter are time-consuming because it is necessary to capture a scene multiple times with a number of different filters. Methods based on diffractive media such as a grating and a prism also requires long time to capture a scene because its hyperspectral image consists of multiple columns obtained by push-broom imaging. Most commercial hyperspectral cameras are based on those traditional approaches that are time-consuming and costly. To deal with those problems, there are many alternative approaches.

There have been a number of methods for multispectral/hyperspectral imaging in computational photography. Most of the methods rely on active illumination. D'Zmura [33] recovered spectral reflectance through estimating coefficients in a linear model using a set of illumination patterns whose spectra are independent from each other. Park *et al.* [133] reconstructed a multispectral video by capturing a scene under multiplexed illumination which is combinations of different LEDs. Chi *et al.* [22] selected an optimized set of wide band filters to estimate spectral reflectance. They put the set of filters in front of a light source instead of a camera. Han *et al.* [60] proposed a method to fast recover the spectral reflectance by using a DLP projector and a high-speed camera. Instead of active illumination, Oh *et al.* [129] proposed a framework for reconstructing hyperspectral images by using multiple consumer-level digital cameras. They employed

small differences of spectral sensitivities along multiple cameras for reconstructing a hyperspectral image. However multiple cameras are necessary to implement this method. Although those methods are very effective to make hyperspectral imaging more accurate and easier to be used, they still require expensive and specialized equipments.

Several methods for one-shot hyperspectral imaging have been proposed. Morovic and Finlayson [111] proposed a method to estimate spectral reflectance from a single RGB image. It is impossible to establish a unique correspondence between an RGB vector and spectral reflectance because the dimension of reflectance is higher than that of an RGB vector. Therefore, they made strong assumptions that reflectance follows a normal probability distribution and is smooth, and then trained a model under some conditions to which reflectance must adhere. Abed *et al.* [2] proposed a linear interpolation method using lookup tables. In a case of a scene under the same illumination, the reflectance of a polytope that encloses an RGB point in the scene is used for interpolation. And Nguyen *et al.* [126] introduced a non-linear mapping strategy for modeling the mapping between an RGB value and a spectra. Those one-shot hyperspectral imaging techniques are effective to some extent but their accuracy highly rely on the dataset for training. Others for one-shot hyperspectral imaging are computed tomography image spectrometers [58], which obtains diffracted signals by slicing a hyperspectral image as a 3D data through a diffraction grating, and coded aperture snapshot spectral imagers [9], which employs compressive sensing by dispersive elements and a coded aperture. Both of the methods can estimate the 3D data from the diffracted signals but still require expensive and specialized equipments. Manakov *et al.* [106] proposed a camera add-on using a kaleidoscope for high dynamic range, multispectral, polarization, and light-field imaging. Their work looks similar to our idea but they copied the input image onto 3×3 images, and then used 9 selected color filters. On the other hand, our basic idea is to use only a color filter and put the filter on the reflectors, that is totally different from the work by Manakov *et al.*.

There are several methods using spectral imaging to separate some components. Fu *et al.* [39] separated the reflection and fluorescence components using high frequency illumination in the spectral domain. Asano *et al.* [10] extracted the

absorption component from a subtraction between images at different wavelengths and then used it to recover the shape of an object in water. Wind and Szymanski [175] separated the scattering and transmission components using based on Beer-Lambert law.

2.4. Photometric measurements of shape

In general, there are two types of methodologies to obtain a shape of an object. One is a passive method, such as stereo camera [141] and multi-view stereo [144], which is preferred for obtaining a large-size object, *e.g.* buildings. Another is an active method, such as structured light [140] and time-of-flight sensing [150, 35], which is practical for obtaining a detailed shape, *e.g.* industrial products and human face. In this thesis, the active method is focused on because of its practicality. There have been many active methods for obtaining depths as a shape clue, as given above, but reconstructing surface orientations is also important to obtain a detailed shape.

The photometric stereo has been a broad interest in the computer vision field. The Lambertian photometric stereo [177] is a standard way to recover the surface normal by assuming Lambert reflection, no optical effect such as shadow and scattering, orthogonal projection, and parallel lights. Since the traditional photometric stereo assumes the diffuse reflection, it does not work well for glossy surface. A solution is to separate the specular reflection component based on the dichromatic reflectance model in preprocessing [143, 138, 139]. To apply the Lambertian photometric stereo for a non-Lambert surface, other optical components need to be separated by pattern projection [124], polarization [123], and fluorescence [164]. Inoshita *et al.* [73] improves the photometric stereo for translucent objects using surface normal deconvolution, Ngo *et al.* [125] use a polarization cue to recover a smooth surface, and Murez *et al.* [116] develop a photometric stereo in a scattering media that consider the blur depending on the distance. While these methods jointly compensate for the global light transport in their solutions, we aim to separate the far infrared light transport.

2.5. Photometric measurements of visual textures

Color is an important factor for visual textures and it can be obtained through spectral analysis, as mentioned in Section 2.3. Decompositions can be used to measure a slice of BRDF [118], Other important factors are translucency and glossiness. Translucency of a material is derived from scattering and absorption parameters. Thus, methods to measure the parameters [54, 103] are regarded as ones to obtain the translucency of a material. Obtaining and reproducing the translucency has been studied in computer vision and graphics [12, 28, 50]. Glossiness of a material is often studied as directional reflectance. In fact, the glossiness is more affected by micro-surface structure than its material. Thus, measurement of bi-directional reflectance distribution functions (BRDFs) can be regarded as obtaining the glossiness [34, 36, 101].

Currently, the main application of obtained visual textures is to virtually render photorealistic graphics via physically-based rendering. However, a relatively new application, physically reproduction of the visual textures, is challenged in this thesis. To control color, a factor of the appearance, there are some commercial 3D printers that can build a colorful object. Brunton *et al.* [20] increased the number of controllable colors in 3D printing by utilizing natural error diffusion. Weyrich *et al.* [174] controlled glossiness, a factor of the appearance, by using a CNC router. Hašan *et al.* [62] and Dong *et al.* [27] controlled translucency, also a factor of the appearance, by using a 3D printer. Papas *et al.* [132] replicated both of the color and translucency by mixing different pigments.

Chapter 3

Optical decomposition using multiple-weighted measurements

An observed image is composed of multiple components based on optical phenomena, such as light reflection and scattering. However, most scene analysis methods in computer vision assume only simple optical phenomena as explained in Chapter 1. Thus, decomposition methods are important for various computer vision tasks because unexpected optical components in the observed image could disturb the scene analysis methods.

Various optical components have been targeted for decomposition so that only an expected component is extracted because the expected component is different with respect to the scene analysis methods. For example, a polarization-based method [176] is expected as removing specular reflection component. An active method using a projector-camera system [122] supposes to separate direct and indirect illumination components. However, the polarization-based method also removes single scattering component and the direct illumination component still has various components such as diffuse and specular reflection. Combining those different methods could enable to decompose into more detailed components but no general approach to combine them exists.

In this chapter, we propose a general approach to combine different decomposition methods in a linear algebraic manner, called multiple weighted measurements. With a novel perspective, a decomposition method can be regarded as a weighted measurement, which weakens some of all components with some

weights derived from the method. A weighted measurement is formulated in a linear algebra, which makes it possible to combine different kinds of decomposition methods.

Experimental results show that the proposed approach decomposes observed images into four optical components based on diffuse and specular reflection and single and multiple scattering. The decomposed components are applied to material segmentation as an application.

3.1. Multiple weighted measurements

An intensity on an image observed by a conventional camera is mixture of signals derived from various optical phenomena, such as light reflection and scattering. Assuming m components in the mixture, the observed image $\boldsymbol{s} \in \mathbb{R}^{\mathrm{P}}$ is written as below;

$$\boldsymbol{s} = \sum_{i=1}^{m} \boldsymbol{c}_i, \tag{3.1}$$

where P is the number of pixels in an image and $\boldsymbol{c}_i \in \mathbb{R}^{\mathrm{P}} (1 \leq i \leq m)$ is a component image. The purpose of this chapter is to obtain each component image $\boldsymbol{c}_i$ from multiple observations. The component images can be defined in various manners, *e.g.*, diffuse and specular reflection, single and multiple scattering, or direct and global illumination components. If a method which individually measures each of the components can exist, then no decomposition method is required. However, such an individual measurement does not exist and that is why there are a lot of decomposition methods. Even so, the decomposition methods do not still provide the individual measurement. For example, a decomposition method using polarization is expected to separate the specular reflection component from others but the separated specular reflection component by polarization still includes the single scattering component. Thus, we regard a decomposition method as extraction of a part of the mixture, named a weighted measurement. The decomposition method weakens some components with a weight vector $\boldsymbol{w} \in \mathbb{R}^m$;

$$\boldsymbol{w} = \left[w_1 w_2 \cdots w_m \right]^\top. \tag{3.2}$$

The observed image $\boldsymbol{s}$ can be expressed by using the weight vector $\boldsymbol{w}$, as follow;

$$\boldsymbol{s} = \sum_{i=1}^{m} w_i \boldsymbol{c}_i. \tag{3.3}$$

The weighted measurement is formulated in matrix form, as follow;

$$\boldsymbol{s} = \boldsymbol{C}\boldsymbol{w}, \tag{3.4}$$

where $\boldsymbol{C} = [\boldsymbol{c}_1 \boldsymbol{c}_2 \cdots \boldsymbol{c}_m] \in \mathbb{R}^{\mathrm{P}\times m}$ as a component matrix.

Given $n(\geq m)$ different weighted measurements, an observed image $\boldsymbol{s}^j (1 \leq j \leq n)$ by each of the measurements with a weight vector $\boldsymbol{w}^j$ is formulated as below;

$$\boldsymbol{s}^j = \boldsymbol{C}\boldsymbol{w}^j. \tag{3.5}$$

All the measurements can be expressed in matrix form, as below;

$$\boldsymbol{S} = \boldsymbol{C}\boldsymbol{W}, \tag{3.6}$$

where $\boldsymbol{S} = [\boldsymbol{s}^1 \boldsymbol{s}^2 \cdots \boldsymbol{s}^n] \in \mathbb{R}^{\mathrm{P}\times n}$, an observation matrix, and $\boldsymbol{W} = [\boldsymbol{w}^1 \boldsymbol{w}^2 \cdots \boldsymbol{w}^n] \in \mathbb{R}^{m\times n}$, a weight matrix, that is called multiple weighted measurements.

Decomposition which this chapter aims at is to obtain the component matrix $\boldsymbol{C}$. When the shape of the weight matrix is square, $n = m$, and the rank of the matrix is full, $\mathrm{rank}(\boldsymbol{W}) = m$, then the component matrix $\boldsymbol{C}$ can be computed by

$$\boldsymbol{C} = \boldsymbol{S}\boldsymbol{W}^{-1}. \tag{3.7}$$

When the shape is horizontally-long rectangle, $n > m$, and $\mathrm{rank}(\boldsymbol{W}) = m$, then the component matrix $\hat{\boldsymbol{C}}$ is estimated in a least squares manner, as follow;

$$\hat{\boldsymbol{C}} = \boldsymbol{S}\boldsymbol{W}^{+} = \boldsymbol{S}(\boldsymbol{W}^{\top}\boldsymbol{W})^{-1}\boldsymbol{W}^{\top}, \tag{3.8}$$

where $\boldsymbol{W}^{+}$ is the pseudo inverse matrix of $\boldsymbol{W}$. Finally, the decomposition is performed in a linear algebraic manner, given a set of weighted measurements.

Additionally, the rank of the weight matrix reveals feasibility of the decomposition in advance, before performing measurements. The decomposition is feasible only if the rank is full, $\mathrm{rank}(\boldsymbol{W}) = m$. Otherwise, other measurement methods are required so that the rank is full. According to the nature of least squares, the larger number of combinations is, the more stable solution is estimated, even if the rank is full.

3.2. Decomposition of reflection and scattering components

In the previous section, we explained the theory of multiple weighted measurements. A key of the proposed approach is to design the weight matrix $\boldsymbol{W}$ so that the decomposition becomes feasible. However, we cannot arbitrarily design the weight matrix because a weight vector is derived from a measurement method. This section describes how to build the weight matrix as an implementation.

3.2.1 Light reflection and scattering components

An observed intensity at a point is mixture of various optical components. Figure 3.1(a) illustrates light reflection and scattering phenomena at the point. Light reflection is often classified into two components; diffuse and specular reflection. Diffuse reflection arises because of a microfacet structure on object surface. On the other hand, specular reflection arises at an interface between the air and the object surface (Fig. 3.1(b)).

Light scattering is also classified into two components; single and multiple scattering, according to researches in computer vision [21, 114] and physics [71, 108]. Single scattering is caused by one-bounce collision with a particle, or particle aggregation, inside an object, which is often seen in optically thin media (Fig. 3.1(c)). A well-known nature of single scattering is that an intensity of single scattering exponentially decays along its light path. On the other hand, multiple scattering is a phenomenon of multi-bounce collisions, which is often seen in optically thick media (Fig. 3.1(c)).

In this chapter, we aim at decomposing observed images into the above four optical components; diffuse and specular reflection and single and multiple scattering. Interreflections are not explicitly modeled in this implementation. Since interreflections and multiple scattering phenomena are similarly based on multi-bounce collisions with surfaces and inside particles, respectively (Fig. 3.1(d)), both of the components are included in the multiple scattering component.

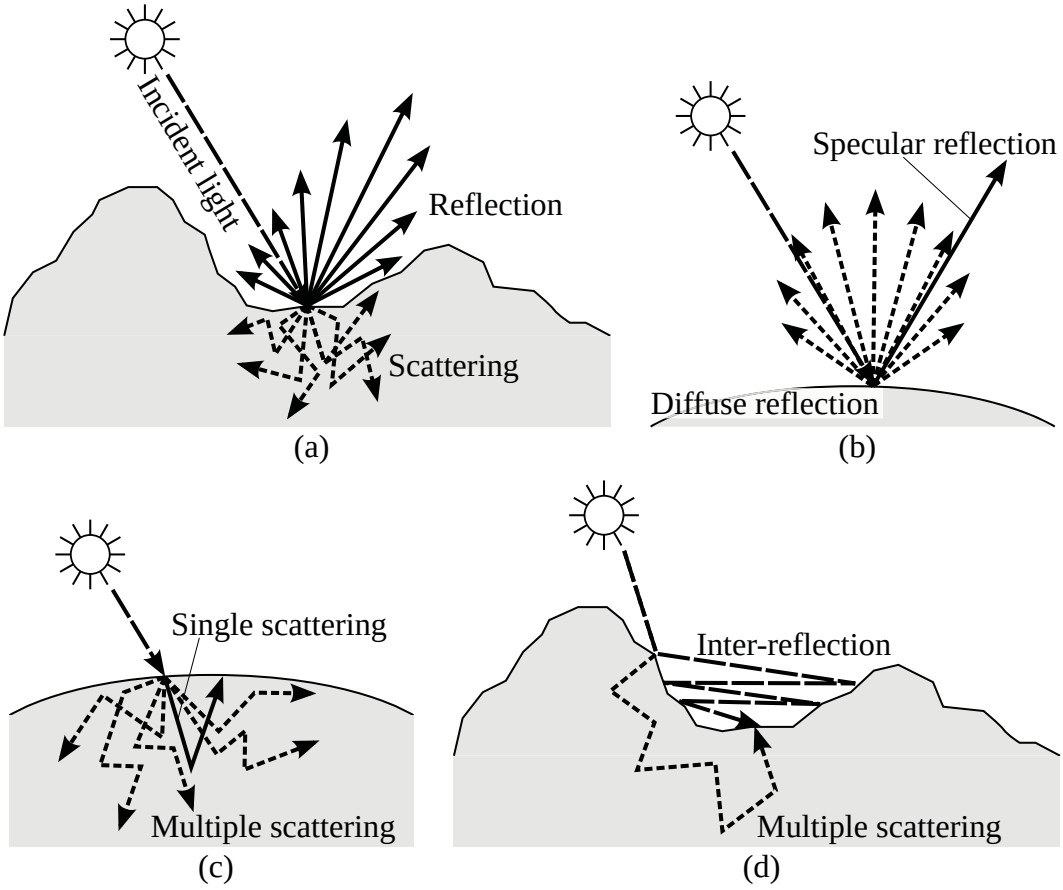


Figure 3.1: Light reflection and scattering. (a) Reflection and scattering phenomena at a point. (b) Reflection is classified into two components; diffuse and specular reflection. (c) Scattering is also classified into two components; single and multiple scattering. (d) Interreflections and multiple scattering phenomena are similarly based on multi-bounce collisions.

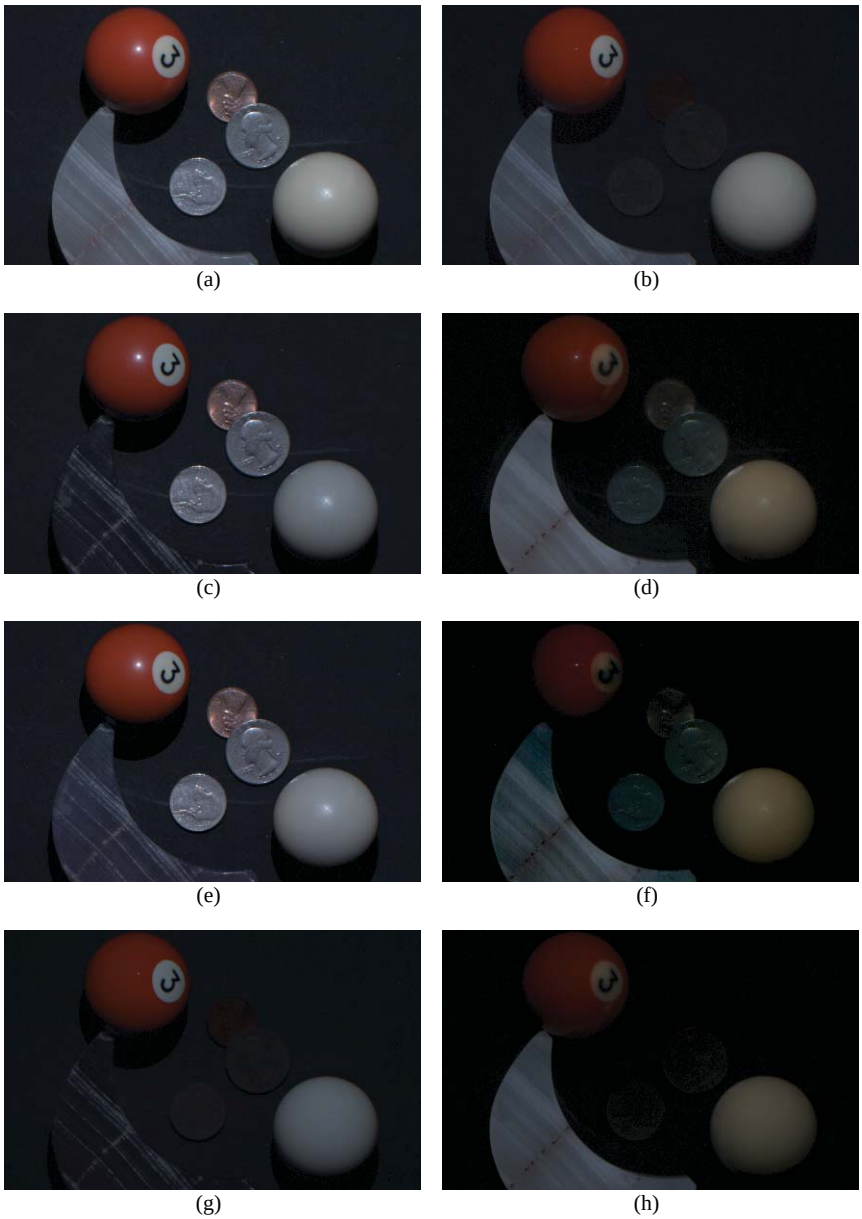


Figure 3.2: Observed images by several separation methods. (a) under an ordinary illumination, (b) circular polarization, (c-d) direct and global components by high frequency illumination, (e-f) direct and global components by sweeping high frequency illumination, (g-h) direct and global components by high frequency illumination with circular polarization.

3.2.2 Definition of measurement weights

Now we define weight vectors $\boldsymbol{w}^j$ to the four components as previously described. We consider four distinct weight elements which correspond to the four components, *i.e.*, diffuse reflection w_{DR}, specular reflection w_{SR}, single scattering w_{SS}, and multiple scattering w_{MS}. By definition, each weight is in a range of $0 \leq w_i \leq 1$. Therefore, a weight vector $\boldsymbol{w} \in \mathbb{R}^4$ is explained as

$$\boldsymbol{w} = [w_{\mathrm{DR}}\ w_{\mathrm{SR}}\ w_{\mathrm{SS}}\ w_{\mathrm{MS}}]^\top. \tag{3.9}$$

In the following, we describe separation methods and their corresponding weight vectors. Note that the weight vectors are theoretically determined from the methodology, instruments, and experimental setup.

Normal observation

An image taken under an ordinary illumination, *e.g.*, the uniform white illumination, condition contains all of the four components. We treat this observation as the one that contains all the components equally without reduction of anything. Therefore, the weight vector is defined as

$$\boldsymbol{w}^{\mathrm{NML}} = [1\ 1\ 1\ 1]^\top. \tag{3.10}$$

Figure 3.2(a) shows an image taken under white illumination projected by a projector. In the scene, there are a marble stone, two billiard balls, and three coins.

Circular polarization

Techniques based on circular polarization can separate specular reflection [176, 46] and single scattering [48, 165] from other components. The nature of circular polarization is that right-handed (or left-handed) circularly polarized light cannot transmit through a left-handed (or right-handed) circular polarizer. Since one-bounce collision reverses the handedness of the polarized light, specular reflection and single scattering, which are derived from one-bounce collision with a surface and an inside particle, respectively, change the handedness of the polarized incident light. On the other hand, multi-bounce collisions, such as diffuse reflection

and multiple scattering, turn polarized light into unpolarized one. Therefore, putting a same-handed circular polarizer in front of both a light source and a camera can remove specular reflection and single scattering.

However, in practice, a polarizer does not have a perfect capability for light transmission and shielding but the single transmittance t_s and the crossed transmittance t_c. The single transmittance t_s is the ratio of the power of light passed through the polarizer to that of the incident unpolarized light. The crossed transmittance t_c is the ratio of the power of light passed through the one-handed polarizer to that of the incident opposite-handed polarized light. Thus, the weight vector is defined as

$$\boldsymbol{w}^{\mathrm{CP}} = [t_s^2 \ t_s t_c \ t_s t_c \ t_s^2]^\top. \tag{3.11}$$

Figure 3.2(b) shows an observed image by using a circular polarization technique in the same scene. We simply put a circular polarizer in front of the projector and the same-handed circular polarizer in front of the camera. As we can see, the coins cannot almost be seen and the highlights on the balls was removed.

High frequency illumination

High frequency illumination, proposed by [122], can separate direct and global illumination components in a scene from observed images under spatially high frequent pattern illuminations, such as a checkerboard pattern. The direct illumination component includes directly reflected light on surfaces in the scene and the global one includes others, such as in-directly reflected light, scattered light, and transmitted light. For the details we refer the reader to [122]. In this instance, the direct and global illumination components correspond to reflection and scattering ones, respectively. Thus, the weight vectors for the direct and global illumination components are defined as

$$\begin{cases} \boldsymbol{w}_{\mathrm{D}}^{\mathrm{HFI}} = [1 \ 1 \ 0 \ 0]^\top, \\ \boldsymbol{w}_{\mathrm{G}}^{\mathrm{HFI}} = [0 \ 0 \ 1 \ 1]^\top. \end{cases} \tag{3.12}$$

Separated direct and global illumination components in the scene are shown in Fig. 3.2(c) and (d), respectively. Since the marble stone is a translucent object, the intensity on the marble stone region is mostly included in the global component. The billiard balls are also translucent to some extend, so the texture

on the ball, *e.g.*the number 3, is blurred in the global component while that is clear in the direct one. We can see specular interreflections on the white ball, which is reflected on the ball again after being reflected on the coins.

Sweeping high frequency illumination

Mukaigawa *et al.* [114] have proposed sweeping high frequency illumination, which can separate single and multiple scattering components in a scene by projecting spatially high frequent stripe patterns, inspired by high frequency illumination [122]. The separated direct component includes not only light reflection but also single scattering, while the global one includes the others, such as multiple scattering and intereflections. Therefore, the weight vectors for the direct and global components are defined as

$$\begin{cases} \boldsymbol{w}_{\mathrm{D}}^{\mathrm{SHFI}} = [1\ 1\ 1\ 0]^{\top}, \\ \boldsymbol{w}_{\mathrm{G}}^{\mathrm{SHFI}} = [0\ 0\ 0\ 1]^{\top}. \end{cases} \tag{3.13}$$

Separated direct and global components in the scene are shown in Fig. 3.2(e) and (f), respectively. Comparing with that in the direct component of high frequency illumination (c), the marble stone region in the direct component (e) is brighter because the single scattering component is included.

New combination: high frequency illumination with circular polarization

A combination of several separation methods let us define another weight vector. For example, we combine the high frequency illumination technique with the circular polarization technique. It is easily implemented with the projector-camera system, which is used to implement the high frequency illumination, and a pair of the same-handed circular polarizers. The combination can separate direct and global components, similar to the results of high frequency illumination, but the specular reflection and single scattering components are removed in both of the components. In this instance, each element of a new weight vector is the product of corresponding elements of the weight vectors; Eqs. (3.11) and (3.12). Thus,

the weight vectors are defined as

$$\begin{cases} \boldsymbol{w}_{\mathrm{D}}^{\mathrm{HFICP}} = \boldsymbol{w}_{\mathrm{D}}^{\mathrm{HFI}} \circ \boldsymbol{w}^{\mathrm{CP}} = [t_s^2 \; t_s t_c \; 0 \; 0]^\top, \\ \boldsymbol{w}_{\mathrm{G}}^{\mathrm{HFICP}} = \boldsymbol{w}_{\mathrm{G}}^{\mathrm{HFI}} \circ \boldsymbol{w}^{\mathrm{CP}} = [0 \; 0 \; t_s t_c \; t_s^2]^\top, \end{cases} \tag{3.14}$$

where $\circ$ is the Hadamard product operator. Note that the new weight vectors, Eq. (3.14), are linearly independent of those of the circular polarization, Eq. (3.11), and the high frequency illumination, Eq. (3.12). This is a way that we can obtain a new weighted measurement by simply combining several separation methods.

Figure 3.2(g) and (h) shows separated direct and global components, respectively. As we can see, the specular reflection component is removed in the direct component (g). Actually, there exists specular interreflections in the global component of high frequency illumination (d), *e.g.*, coins. However, in the global component (h), those are perfectly removed thanks to the effect of circular polarization.

3.2.3 Weight matrix

We employ the five weighted measurements, as described above, to implement the multiple weighted measurements for decomposition into four components; diffuse and specular reflection and single and multiple scattering. In this instance, the weight matrix $\boldsymbol{W} \in \mathbb{R}^{4\times 8}$ consists of the eight weight vectors, as

$$\begin{aligned} \boldsymbol{W} = [& s^{\mathrm{NML}}\boldsymbol{w}^{\mathrm{NML}} \quad s^{\mathrm{CP}}\boldsymbol{w}^{\mathrm{CP}} \quad s^{\mathrm{HFI}}\boldsymbol{w}_{\mathrm{D}}^{\mathrm{HFI}} \quad s^{\mathrm{HFI}}\boldsymbol{w}_{\mathrm{G}}^{\mathrm{HFI}} \\ & s^{\mathrm{SHFI}}\boldsymbol{w}_{\mathrm{D}}^{\mathrm{SHFI}} \quad s^{\mathrm{SHFI}}\boldsymbol{w}_{\mathrm{G}}^{\mathrm{SHFI}} \quad s^{\mathrm{HFICP}}\boldsymbol{w}_{\mathrm{D}}^{\mathrm{HFICP}} \quad s^{\mathrm{HFICP}}\boldsymbol{w}_{\mathrm{G}}^{\mathrm{HFICP}}] , \end{aligned} \tag{3.15}$$

where s^j is the global scales for each weighted measurement. The scales are decided by an experimental setup. In practice, the scales are normalized to one because the experimental setup is not changed while performing all of the weighted measurements. The eight weight vectors do not have to be linearly independent to each other as long as the wights matrix $\boldsymbol{W}$ has a full-rank. For example, $[\boldsymbol{w}^{\mathrm{NML}} \; \boldsymbol{w}_{\mathrm{D}}^{\mathrm{HFI}} \; \boldsymbol{w}_{\mathrm{G}}^{\mathrm{HFI}}]$ consists of linearly dependent columns because $\boldsymbol{w}^{\mathrm{NML}} = \boldsymbol{w}_{\mathrm{D}}^{\mathrm{HFI}} + \boldsymbol{w}_{\mathrm{G}}^{\mathrm{HFI}}$. However, all of them can be combined together in the weight matrix $\boldsymbol{W}$ for a stable computation. Designing a weight matrix can be

done ahead before measuring and computing, that is, the rank analysis of the designed weight matrix let us know whether decomposition is feasible, or not, in advance. In this instance, the weight matrix $\boldsymbol{W}$ has full-rank because $t_s \gg t_c$ for a general polarizer. Therefore, the decomposition is a well-posed problem.

In fact, the rank of the weight matrix $\boldsymbol{W}$ can be systematically analyzed in this case. Let us consider the product $\boldsymbol{W}\boldsymbol{W}^\top \in \mathbb{R}^{4\times4}$. Its determinant has a closed-form expression as

$$\det(\boldsymbol{W}\boldsymbol{W}^\top) = t_s^2(t_s - t_c)^2(15t_s^4 + 24t_s^2t_c(t_c - t_s) + 16). \tag{3.16}$$

Therefore, with the condition $t_s \gg t_c > 0$, the determinant becomes positive; the weight matrix $\boldsymbol{W}$ is full-rank. In practice, the conditioning of the weight matrix $\boldsymbol{W}$ is more important for the stability of the pseudo inverse $\boldsymbol{W}^+$. One of the ways to evaluate the conditioning is to assess the ratio between the largest and smallest singular values, $\sigma_{\max}$ and $\sigma_{\min}$, of the weight matrix $\boldsymbol{W}$, which can be numerically computed as

$$\kappa(\boldsymbol{W}) = \frac{\sigma_{\max}}{\sigma_{\min}}. \tag{3.17}$$

$\kappa(\boldsymbol{W})$ is often called the condition number of $\boldsymbol{W}$.

3.3. Experiments

First, we verify a result of the decomposition by the proposed approach. In the verification, we use a simple scene where there are some typical materials in Section 3.3.1. Second, we analyze the repeatability of the decomposition and the effect of each of the weighted measurements in Section 3.3.2. Finally, we perform the decomposition in various complex scenes and discuss the decomposition results in Section 3.3.3.

We begin at describing the experimental setup in this section. In all of the experiments performed in this chapter, we use a 3M MPro160 projector as a light source and a Point Grey Research Chameleon color camera as a recording device, as shown in Fig. 6.6(a). To employ the circular polarization technique, we used two circular polarizers, Kenko SQ Circular-PL with $t_s = 0.399$ and $t_c = 0.0005$ as the product-specific values. In measurements of the polarization approach, we put them in front of the projector and the camera, as shown in Fig. 6.6(b). For

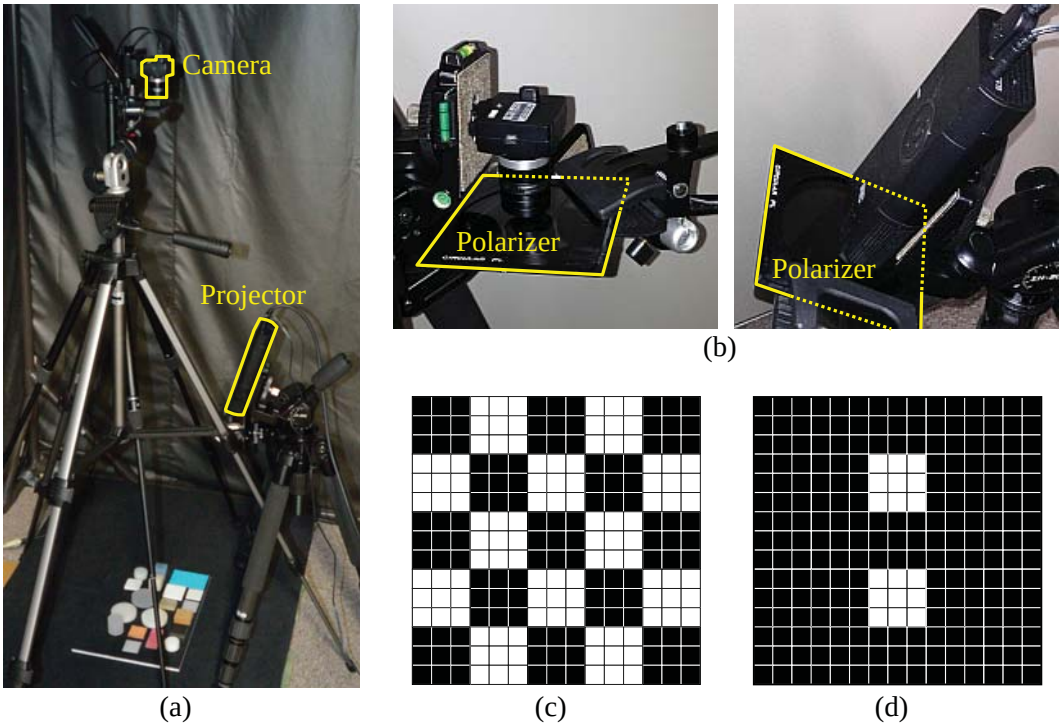


Figure 3.3: Experimental setup. (a) Locations of the camera and the projector. (b) Circular polarizers placed in front of both of the camera and the projector. (c) a part of a checkerboard pattern used for high frequency illumination. (d) a part of a dotted line pattern used for sweeping high frequency illumination.

the high frequency illumination, we project several checkerboard patterns whose block size is a 3×3 pixels square, as shown in Fig. 6.6(c). Figure 6.6(d) illustrates a dotted line pattern for the sweeping high frequency illumination, which also consists of only vertically, or horizontally, repeated 3×3 pixels squares.

In the experiments in this chapter, we employ all of the weighted measurements, described in Section 3.2.2, to obtain the observation matrix $\boldsymbol{S}$. Since the weight matrix $\boldsymbol{W}$ is defined as Eq. (3.15), we can compute the component matrix $\hat{\boldsymbol{C}}$ by Eq. (3.8), that is, we can obtain the decomposition into diffuse and specular reflection and single and multiple scattering components. Note that all of the weighted measurements are done under the same experimental setup. Thus, we assume all of the global scales in Eq. (3.15) have been normalized.

3.3.1 Verification

To verify the decomposition by the proposed approach, we use a simple and well-designed scene, as shown in Fig. 3.4(a). The target scene consists of four typical materials; a ceramic board, a duralumin plate, a block of milky epoxy resin, and a cylinder of polyoxymethylene (POM) resin. On a surface of matte ceramics, light tends to be evenly diffused for all angles because of its microstructure. Duralumin, a type of aluminum alloys, strongly reflects light on its surface, therefore a specularity becomes dominant. Both of the resins are translucent media but they have different translucencies, as shown in Fig. 3.4(b). The block of milky epoxy resin consists of an optically thin medium, thus we can observe a light ray in the medium, which is a feature of single scattering and depends on the incident light angle. On the other hand, in a optically thick medium, such as the cylinder of POM resin, the observed light does not depend on the incident light angle but evenly spreads because of multiple scattering.

We observed the scene with the five weighted measurements and then decomposed them into the four optical components by computing Eq. (3.8). The decomposed result is shown in Fig. 3.4(c-f), which are (c) diffuse reflection, (d) specular reflection, (e) single scattering, and (f) multiple scattering components. To analyze the result, we computed the averages of intensities in each material region on each optical component image and summarized the proportion of the averages in Fig. 3.4(g). As similar to our expectation, the dominant optical components varied across the materials; diffuse reflection became dominant in the ceramic board (78.7%), specular reflection in the duralumin plate (68.1%), single scattering in the block of milky epoxy resin (41.0%), and multiple scattering in the cylinder of POM (50.6%). Consequently, the verification shows that the decomposition by the proposed approach leads a significant decomposition of observations into the four optical components; diffuse and specular reflection and single and multiple scattering, while it is difficult to quantitatively analyze its performance. Note that the decomposition cannot be achieved by applying any of the existing separation methods.

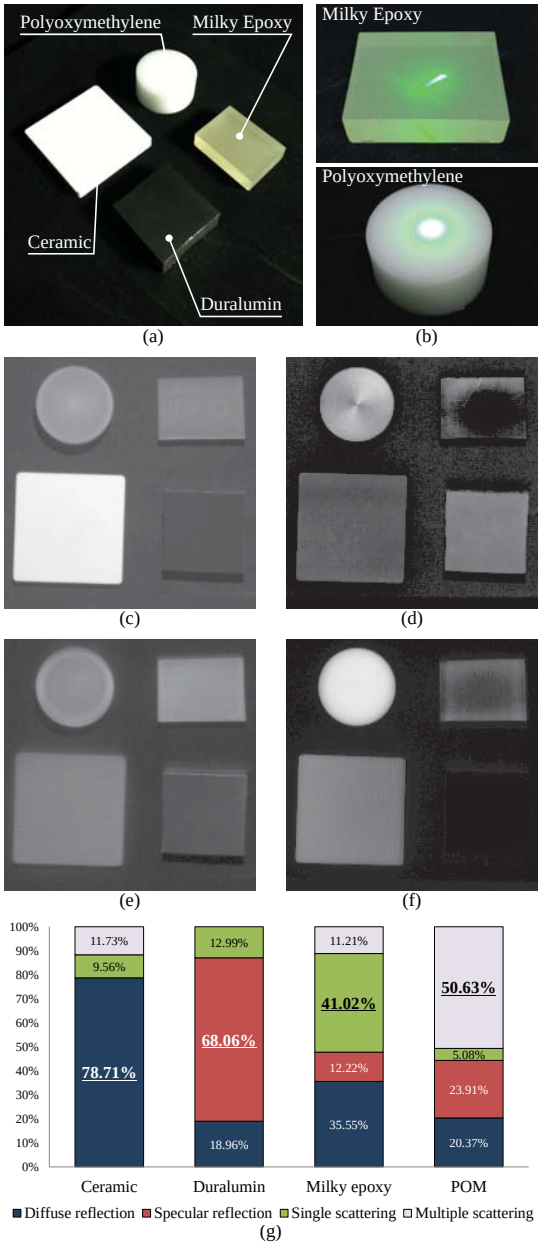


Figure 3.4: Verification. (a) The target scene consists of four materials; a ceramic board, a duralumin plate, a block of milky epoxy resin, and a cylinder of polyoxymethylen (POM). (b) The resins have different scattering properties. The observations are decomposed into four components; (c) diffuse and (d) specular reflection and (e) single and (f) multiple scattering components. (g) The proportion of the averaged intensities in each material region.

Table 3.1: Evaluation of the effect of each of the weighted measurements. We performed the decomposition without one of the weighted measurements and compared the result with that with all of them in PSNR[dB].

Removed measurement	Diffuse reflection	Specular reflection	Single scattering	Multiple scattering
Normal observation	25.5	26.7	24.1	26.8
Circular Polarization	25.2	26.4	24.1	26.6
High Frequency Illumination	26.3	27.6	25.2	26.1
Sweeping Hight Frequency Illumination	25.3	26.4	24.6	24.4
High Frequency Illumination with Circular Polarization	26.8	26.9	24.1	26.5

3.3.2 Analysis of decomposition results

We analyze the decomposition with two different perspectives. First, we show the repeatability of the decomposition. We take each weighted measurement five times under the same experimental setup, and then compare decomposition results. Each of the decomposition results is evaluated in peak signal-to-noise ratio (PSNR) between the others. The comparison resulted in 42.1[dB] in PSNR on average with the standard deviation of 1.36[dB]. The average (and the standard deviation) of PSNRs for diffuse and specular reflection and single and multiple scattering components are 42.1(1.58), 42.0(1.42), 42.2(1.15), and 42.0(1.24)[dB], respectively. Consequently, it shows that the repeatability of the decomposition by the proposed approach is quite high.

Second, we evaluate the effect of each of the weighted measurements by comparing the decomposition result with all of them and that without one of them. As shown in Table 3.1, the comparison results say that disusing one of the weighted measurements leads to a large change in a decomposition result. This is because the total number of weight vectors is few, thus $\boldsymbol{W}^{+}$ is significantly changed. For example, when the circular polarization is disused, the PSNRs for diffuse and specular reflection and single scattering components become the lowest. That is, the circular polarization is important for the decomposition. On the other hand,

the PSNRs when disusing the high frequency illumination are relatively high. This is because of the redundancy of the multiple weighted measurements.

3.3.3 Decomposition in complex scenes

We apply the decomposition to more realistic and complex scenes, where there are various everyday objects, as shown in Fig. 3.5. The scene (a) consists of plastic cards, coins, wax candles, and a plastic cup of soap water, the scene (b) consists of a mechanical pencil, a leather pen case, an eraser, an aluminum ruler, and a sticky-paper, and the scene (c) consists of coins, phenolic billiard balls, and a marble stone. Figure 3.5(c) shows the decomposed results; diffuse and specular reflection and single and multiple scattering components, respectively, from left to right.

In the diffuse reflection, single scattering, and multiple scattering components, an intensity is observed to some extent on all the materials except for metals, such as the coins and the ruler. This is because of subsurface scattering, as mentioned in [173, 92]. Almost all real-world materials are translucent to some extent except for metals. In the specular reflection component, an intensity is observed not only on metal materials but also on other materials because specular reflection arises on a smooth surface, such as the surface of the billiard balls. The scattering media, such as the wax candles, the eraser, and the marble stone, show strong intensities in the single and multiple scattering components. Optically thin media, such as the soap water, the eraser, and the marble stone, show relatively stronger intensities than the other materials in the single scattering component. Moreover, the intensity in the single scattering component seems to depend on the shape of an object, *e.g.*, the edges of the wax candles have stronger intensities than other parts. Note that we do not distinguish interreflections from multiple scattering in this chapter, as mentioned in Section 3.2.1, so that interreflections in the scenes are included in the multiple scattering component.

3.4. Application: Raw material segmentation

The decomposition enables a scene analysis in detail. In this chapter, a raw material means unpainted and individually consisting of a single material. The

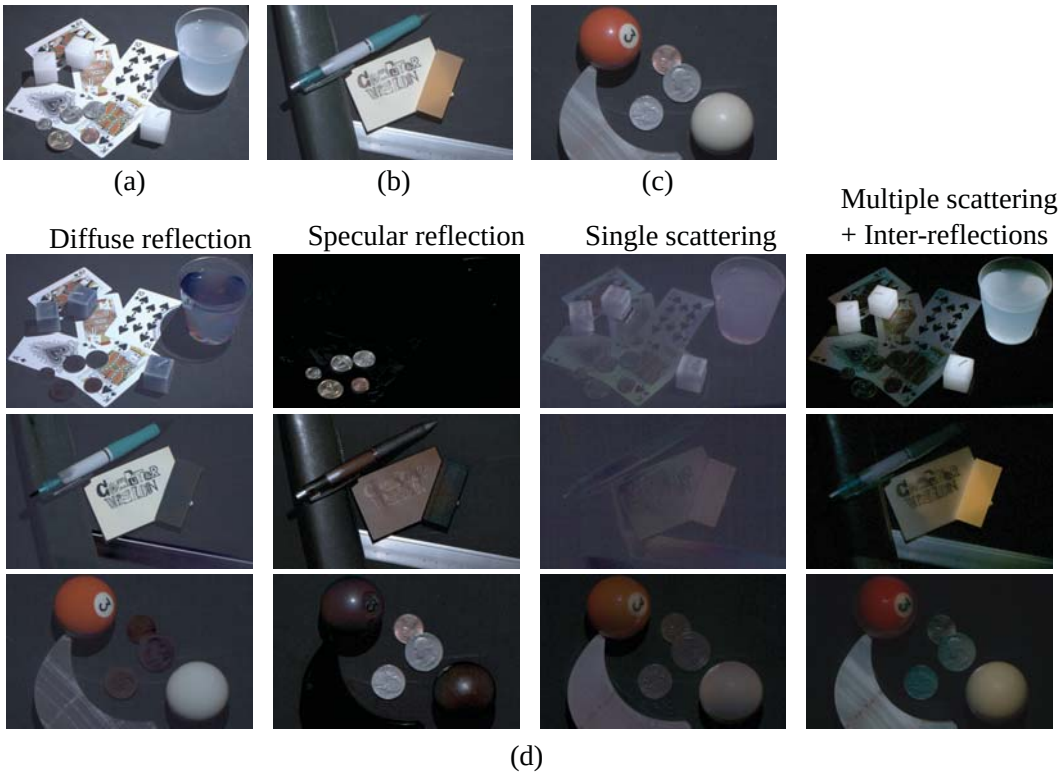


Figure 3.5: Decomposition results in complex scenes. (a-c) the target scenes. (d) the observations in each of the scenes are decomposed into diffuse and specular reflection and single and multiple scattering components.

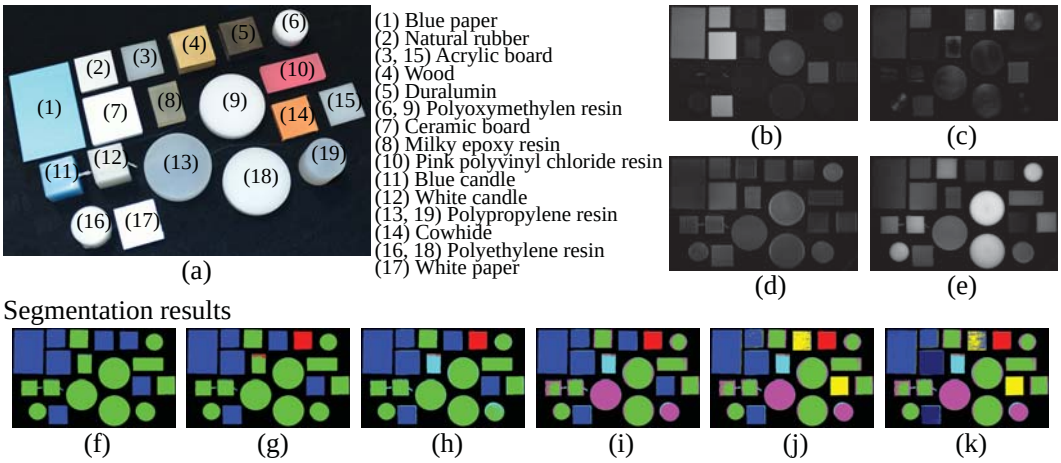


Figure 3.6: Raw material segmentation. (a) the target scene has 19 objects with 13 different raw materials. (b-e) decomposed components by the proposed method; diffuse and specular reflection and single and multiple reflection, respectively. (f) segmented result by k-means clustering with different k values from $k = 2$ to 13

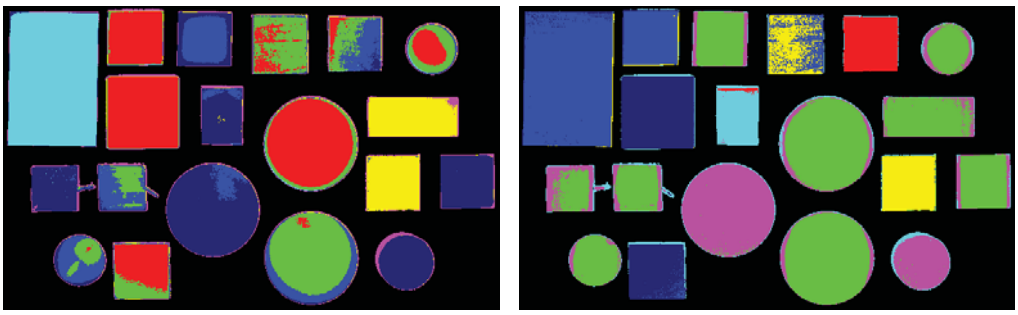
goal of the raw material segmentation, similar to [100], where discriminative illuminations are used for classifying materials, is to classify materials in an image based on the opacity and translucency. Since the proportion of optical components carries significant information about the material property as we have seen in Fig. 3.4(g). To show the potential of the decomposition, we perform the decomposition in a scene, as shown in Fig. 3.6(a), where there are 19 objects with 13 different materials, and then apply a segmentation based on its decomposition result.

We show the decomposition of observations into the four optical components in Fig. 3.6(b-e), which are diffuse and specular reflection and single and multiple scattering components, respectively. From the decomposition result, we form a normalized 4-D feature vector, consisting of the four components, pixel by pixel. And then we simply perform a conventional k-means clustering method as segmentation to assess the effectiveness of the decomposition. The segmentation results are shown in Fig. 3.6(f) with a varying parameter $k(2 \leq k \leq 13)$. As a visualization, the same color regions belong to the same segment.

When $k = 2$, the segmentation result clearly shows a distinction between

opaque and translucent materials; the blue and green regions correspond to opaque and translucent materials, respectively. When $k = 3$, *Material 5* (duralumin) is segmented as another isolated region because of its unique material property, *i.e.*, specular reflection is strongly seen on *Material 5* because duralmin is a type of aluminum alloys. When $k = 4$, *Material 8* (milky epoxy resin) is segmented as a blue-sky region because of its strong single scattering component. When $k = 5$, *Material 13* and *19* (polypropylene resins, PP) are segmented as a different region. A PP resin is a translucent medium with an optically thinner property than the other translucent media except for the milky epoxy resin. In the segmentation result with $k = 6$, *Material 4* (wood) and *Material 14* (cowhide) are separately segmented because both of them are opaquer than the other materials segmented as the blue regions. When $k = 7$, *Material 7* (ceramic) and *17* (paper) are segmented as a new isolated region because of the fact that those materials show stronger scattering components comparing with the other opaque materials. When $k = 8$, *Material 7* (ceramic) is separated as another region. When $k = 9$, *Material 3*, *15* (Acrylic) and *10* (polyvinyl chloride resin) are mainly separated. However, *Material 16*, *18* (polyethylene resins, PE), *Material 11*, and *12* (candles) are partially separated even though they consist of one material. This is because of the colors and the angle of illumination. When $k = 10$, *Material 2* (rubber) is separated from *Material 1* (paper). When $k = 11, 12$, some regions on the same materials are separated because of the angle of illumination. The result at $k = 13$ has only 12 segments which are the same as $k = 12$. Consequently, the translucent materials are classified into six types and the opaque materials into six. This application shows that it is reasonable to classify various opaque and translucent materials based on the decomposition by the proposed approach.

Additionally, we compare the segmentation result with a conventional baseline one. Assuming that only RGB channels are available for segmentation, we performed k-means clustering with $k = 7$ in the RGB space, which resulted in Fig. 3.7(a). Apparently, it is difficult to separate segments based on material properties by using a color-based segmentation approach. Comparing with that, the segmentation result by our approach shows a segmentation based on material properties, as shown in Fig. 3.7(b).



(a) (b)

Figure 3.7: Comparison of segmentation results. (a) segmented result in the RGB space. (b) segmented result based on the decomposition.

3.5. Discussion

In this chapter, we proposed the general approach, called Multiple Weighted Measurements, which enables to uniformly combine any kind of separation methods, such as color-based, polarization-based, and active projection-based, to finely decompose observations. As an implementation, we defined the weight vector of five different weighted measurements and combined them in the proposed approach to decompose observations into four optical component; diffuse and specular reflection and single and multiple scattering components. The experimental verification showed that the decomposition was reasonable because the proportions of decomposed components were similar to the expectation based on physical property for each material region on the image. In the experiments we performed the decomposition in the various complex scenes. We also showed the possibility of its application for raw material segmentation. The decomposition enables a novel segmentation based on the opacity and translucency of materials unlike a conventional segmentation based on the colors.

There are a few limitations in the proposed approach. First, a shadow is not explicitly handled in the linear formulation (Eq. (3.1)). This may yield an unmodeled error in the shadow region as computing the decomposition. Second, unmodeled components are erroneously included in some of the four components. There exist other optical phenomena, such as refraction and fluorescence, although only the four components has been introduced in the chapter. For example, refracted

light on the plastic cup of soap water in the target scene (a) in Fig. 3.5 can be seen in the diffuse component. Third, since it is based on a combination of multiple separation methods, a scene has to be static and the total processing time is a summation of ones for which individual separation methods take. The first limitation is a challenging problem to be solved but worthy to be considered in order to expand the applicability of the decomposition. In order to resolve the second limitation, a method which can separate the other components must be added to the proposed approach. The third limitation cannot essentially be resolved but the total processing time can be reduced if the target components are confined. As shown in Table 3.1, the implemented combination has a redundancy for the decomposition. That is, there must exist the optical combination corresponding to a target component. If the number of combined measurements is reduced, the total processing time also reduces.

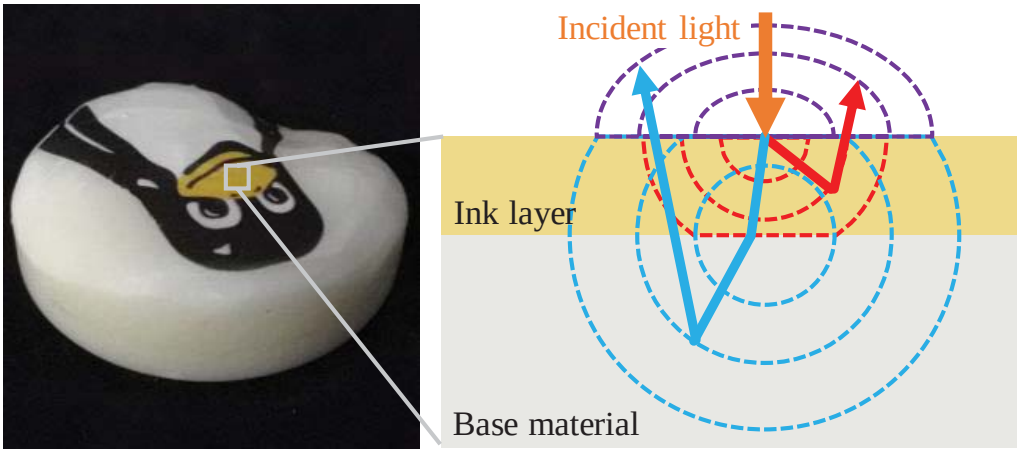
Chapter 4

Controlling translucency by UV printing on a translucent object

Digital fabrication tools, such as 3D printer and CNC router, allow us to easily build a complex-shaped object, and hence are often applied to make a custom-built object, *e.g.*, personal prosthetic limb. Since not only shape but also appearance is an important factor for our visions, techniques to control the appearance in digital fabrication are required.

To control color, a factor of the appearance, there are some commercial 3D printers that can build a colorful object. Brunton *et al.* [20] increased the number of controllable colors in 3D printing by utilizing natural error diffusion. Weyrich *et al.* [174] controlled glossiness, a factor of the appearance, by using a CNC router. Hašan *et al.* [62] and Dong *et al.* [27] controlled translucency, also a factor of the appearance, by using a 3D printer. Papas *et al.* [132] replicated both of the color and translucency by mixing different pigments. The above methods for controlling the translucency modify inside of the object. In contrast, our purpose is also to control the translucency but by modifying outside of the object.

In this chapter, we propose a method to control the translucency by using a UV printer, which can print 2D patterns on 3D objects consisting of various materials, such as plastic and metal. When putting an ink on a base material, the translucency of the printed object depends on both the translucency of the material and ink because of subsurface scattering, as shown in Fig. 4.1. The



(a) Fabricated object (b) Subsurface scattering

Figure 4.1: The translucency of a fabricated object. (a) The fabricated object consists of a UV ink layer on a white wax. (b) Incident light is scattered inside of both of the ink layer (red) and base material (blue), called subsurface scattering, and hence the radiated light (purple) is mixture of them.

translucency can be modified by a combination of factors, such as base material, color of an ink, and the number of printed layers. The proposed method controls the translucency by UV printing through changing such a combination.

Three contributions of this chapter are to propose a novel approach using a UV printer for controlling the translucency, fuse two different manners for resolving tradeoffs, and implement the working system.

4.1. Controlling translucency by UV printing

In general, a UV printer prints the bottom layer with a matte white ink to remove the translucency of base materials. On the other hand, we rather utilize the translucency without printing the white layer. The translucency of the printed object depends on both of the translucency of the base materials and inks, and can be changed by a combination of the factors. If the translucency of the printed objects in all combinations can be measured, it is easy to control the translucency in an example-based manner. However, it is almost impossible to print and

measure in all the combinations. Therefore, we fuse such a manner with a different manner based on physics model.

The proposed method consists of measurement and fabrication steps, as shown in Fig. 4.2. In the measurement step, the individual translucencies of the base materials and inks are measured. The translucency of the printed object in a combination is rendered based on a physics model. The translucency in all the combinations can be rendered by simulation with few measurements. Therefore, it enables to build a lookup table between the combinations and translucency, like the example-based manner. In the fabrication step, given a query about translucency, either measured or manually designed, a combination can be found in the lookup table so that the translucency of the printed object is the most similar to the query. The rest of this section explains the rendering method, measuring the translucency, and building the lookup table and finding the best combination.

4.1.1 Rendering the translucency of a layered object

The translucency of the printed object depends on both of the translucency of the base material and ink, as mentioned above. A key feature is that the printed object has a layered structure, as shown in Fig. 4.1(b). Thus, we apply Kubelka's layer model [91] to render the translucency. The original model formulates scalar reflectance r and transmittance t of a two-layered object, as follow;

$$\begin{cases} r = r_1 + t_1^2 r_2(1 + r_1 r_2 + \cdots) = r_1 + \dfrac{t_1^2 r_2}{1 - r_1 r_2}, & (4.1) \\ t = t_1 t_2(1 + r_1 r_2 + \cdots) = \dfrac{t_1 t_2}{1 - r_1 r_2}, & (4.2) \end{cases}$$

where r_1, t_1 are the reflectance and transmittance of the top layer and r_2, t_2 those of the bottom one, respectively. Now, let us take subsurface scattering into consideration because it is a main cause why an object looks translucent. Since subsurface scattering diffusely spreads light, it is modeled by point spread function (PSF). The PSF can also be separated into reflective and transmissive PSFs. The reflective and transmissive PSFs of the top layer are defined as $R_1(x), T_1(x)$, respectively. As well, those of the bottom one are defined as $R_2(x), T_2(x)$. The

reflective PSF of the two-layered object is, therefore, written as

$$\begin{aligned} R(x) =& R_1(x) + ((T_1 * R_2) * T_1)(x) \\ &+ ((((T_1 * R_2) * R_1) * R_2) * T_1)(x) + \cdots , \end{aligned} \tag{4.3}$$

where '$*$' is the convolution operator. The transmissive PSF can also be written as well but it is omitted here for saving the space. By Fourier transform, Eq. (4.3) is transformed into

$$\begin{aligned} \mathcal{F}[R] &= \mathcal{F}[R_1] + \mathcal{F}[T_1]\mathcal{F}[R_2]\mathcal{F}[T_1] + \cdots \\ &= \mathcal{F}[R_1] + \frac{\mathcal{F}[T_1]^2\mathcal{F}[R_2]}{1 - \mathcal{F}[R_1]\mathcal{F}[R_2]}, \end{aligned} \tag{4.4}$$

where $\mathcal{F}[\cdot]$ means Fourier transform and the argument x is omitted. If the PSFs are isotropic, the imaginary parts in the frequency domain become zero. Therefore, $\mathcal{F}[R]$ can be regarded as the reflective modulation transfer function (MTF), which is defined as $\hat{R}(f_x)$, where f_x is the spatial frequency. Finally, the reflective and transmissive MTFs of the two-layered object are written, as follow;

$$\begin{cases} \hat{R}(f_x) = \hat{R}_1(f_x) + \dfrac{\hat{T_1}^2(f_x)\hat{R}_2(f_x)}{1 - \hat{R}_1(f_x)\hat{R}_2(f_x)}, & (4.5) \\ \hat{T}(f_x) = \dfrac{\hat{T}_1(f_x)\hat{T}_2(f_x)}{1 - \hat{R}_1(f_x)\hat{R}_2(f_x)}, & (4.6) \end{cases}$$

where $\hat{A}(f_x)$ means the corresponding MTF to a PSF $A(x)$.

Once the individual reflective and transmissive MTFs of the base materials and inks are given, those of the printed object can be rendered by using Eqs. (4.5) and (4.6). Recursively applying Eqs. (4.5) and (4.6), it is also possible to render the MTFs of a multi-layered object.

4.1.2 Measuring the modulation transfer functions

The MTFs are employed for explaining the translucency, as mentioned above. In this chapter, to measure the MTFs, we apply the modulated imaging, proposed by Cuccia *et al.* [24] for measuring the quantitative scattering and absorption coefficients of a medium. The modulated imaging is based on measuring MTFs

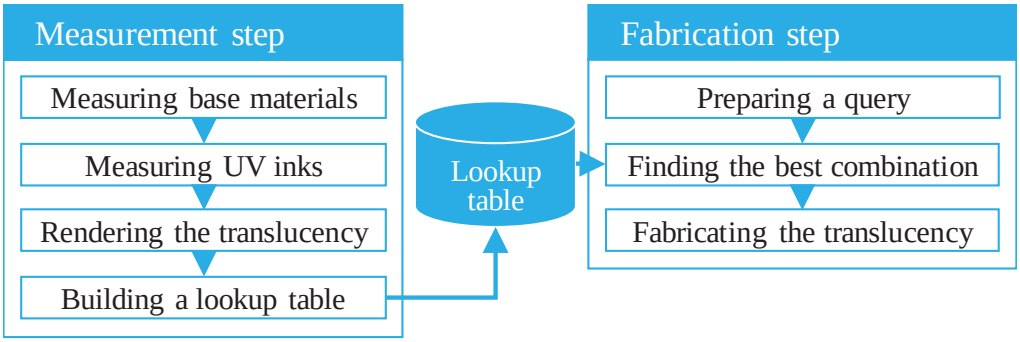


Figure 4.2: The whole system of the proposed method. That consists of measurement and fabrication steps.

in a Pro-Cam system. Projecting a sinusoidal pattern onto the medium, light scattering in the medium blurs the pattern, and hence the amplitude of the measured sinusoidal pattern is attenuated in comparison with that of the projected one. The MTF consists of the rates of attenuation at different frequencies. We refer the readers to [24] for the details.

We can directly measure the MTFs of the base materials in the Pro-Cam system. However, it is impossible to directly measure the MTFs of the inks because the ink layer has to be printed on an object. Therefore, we estimate those by utilizing Eq. (4.5). As printing an ink whose reflective and transmissive MTFs are $\hat{R}_1(f_x)$ and $\hat{T}_1(f_x)$, respectively, on a mirror, the reflective MTF of the printed object $\hat{M}_1(f_x)$ is, as below;

$$\hat{M}_1(f_x) = \hat{R}_1(f_x) + \frac{\hat{T}_1{}^2(f_x)}{1 - \hat{R}_1(f_x)}, \tag{4.7}$$

where we assume $\hat{R}_2(f_x) = 1$ in Eq. (4.5) because of only specular reflection on mirror. Also, as printing the same ink twice, the reflective MTF $\hat{M}_2(f_x)$ is, as below;

$$\hat{M}_2(f_x) = \hat{R}_1(f_x) + \frac{\hat{T}_1{}^2(f_x)\hat{M}_1(f_X)}{1 - \hat{R}_1(f_x)\hat{M}_1(f_X)}. \tag{4.8}$$

Now, the reflective and transmissive MTFs of the ink, $\hat{R}_1(f_x)$ and $\hat{T}_1(f_x)$, can be estimated from Eqs. (4.7) and (4.8) at least because both of $\hat{M}_1$ and $\hat{M}_2$ can be

measured. It is also possible to increase the number of the same ink layers and then use it for a stable estimation in a least-squares method.

4.1.3 The lookup table

Finally, it is possible to build the lookup table by using the rendering method with the measurements. The base materials should be translucent, such as rubber and wax, not transparent and opaque. The lookup table is built by rendering in all the combinations. Building the lookup table is a time-consuming process but it is required only once.

In the fabrication step, given a query about translucency, a combination is searched in the lookup table so that the translucency of the printed object is the most similar to the query. Here, it is required to define distance for representing how much a MTF is close to another. Thus, we employ a root-mean-square error (RMSE) for that. The distance E_{AB} between MTFs $\hat{A}(f_x)$ and $\hat{B}(f_x)$ is defined as

$$E_{AB} = \sqrt{\frac{1}{|\mathbb{F}|} \sum_{f_x \in \mathbb{F}} \left\{ \hat{A}(f_x) - \hat{B}(f_x) \right\}^2}, \tag{4.9}$$

where $\mathbb{F}$ is a set of discrete frequencies to be used for calculating the distance.

4.2. Experiments

We make experiments for evaluating the rendering method and the whole system. For the measurements, we constructed a Pro-Cam system with a projector (Vivitek QUMI Q8) and an RGB camera (FLIR Grasshopper3 Color). In the experiments, we use three base materials; a plastic eraser, a wax candle, and a piece of milky acrylic board.

4.2.1 Quantitative evaluation of the rendering method

The rendering method plays an important role in the proposed method because the lookup table is based on that. Hence, firstly, we quantitatively evaluate the rendering method. In this evaluation, the reflective MTFs of the printed objects in several combinations are rendered and then compared with those which are

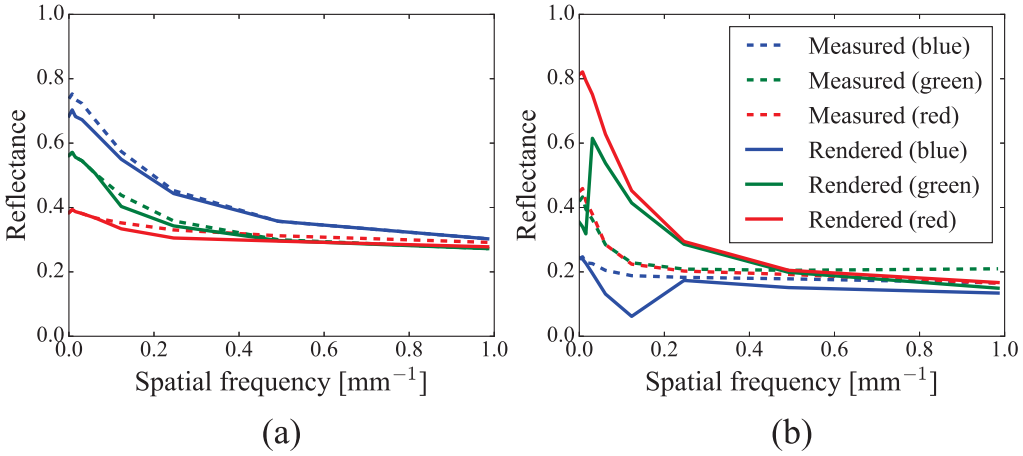


Figure 4.3: The rendered MTFs in comparison with the measured ones. (a) is a case of the cyan on the plastic eraser and (b) the yellow on the acrylic board.

Table 4.1: The RMSEs of the rendered MTFs in comparison with the measured ones.

	Base materials			
Inks	Plastic eraser	Wax candle	Milky acrylic	Average
Cyan	0.0242	0.0838	0.0548	0.0543
Magenta	0.0676	0.0726	0.0918	0.0773
Yellow	0.0741	0.1127	0.1861	0.1243
Average	0.0553	0.0897	0.1109	0.0853

measured through being fabricated. Here, we use three inks; cyan, magenta, and yellow. As examples of the results, the rendered MTFs, as solid lines, and the measured ones, as dashed lines, in three channels are shown in Fig. 4.3, in which (a) is a case of the cyan on the eraser and (b) the yellow on the acrylic. In Fig. 4.3(a), the rendered MTFs are very similar to the measured ones, whose RMSE is on average 0.0242. On the other hand, in Fig. 4.3(b), both of them are less similar, whose RMSE is on average 0.1861. The other RMSEs are witten in Table 4.1 and the totally averaged RMSE is 0.0853. Since the range of reflectance is in 0 to 1, it can be said that the error is on average 8.53%. We believe that is reasonable for building the lookup table.

4.2.2 Controlling the translucency for replication

We make experiments for controlling the translucency of the printed object. The lookup table is built with the three base materials and fifteen inks. Though only the inks of 100% were used in the previous experiment, the same inks of 20, 40, 60, and 80% are also used for building the lookup table here. Setting the maximum number of printed layers five, it results in $2,440,845$ elements in the lookup table. A query about translucency can be manually designed but it is easier to measure a real-world object. We, therefore, measure the reflective MTFs of a slice of salmon and kiwi, as shown in Fig. 4.4, and a small region of those is used as a query. As a result of finding the best combination to the salmon, it was to print in the order of the magenta of 40%, yellow of 100%, magenta of 100%, and magenta of 100% on the wax candle, whose RMSE was 0.0812. As well, the best combination to the kiwi was to print in the order of the cyan of 100%, magenta of 40%, magenta of 40%, magenta of 40%, and yellow of 60% on the wax candle, whose RMSE was 0.0714. The results are shown in Fig. 4.4.

Moreover, we make an experiment for replicating the reflective MTFs of the whole salmon pixel by pixel. According to the previous experiment, the wax candle suits to the MTF of the salmon, thus a part of the lookup table where the base material is the wax candle is used for finding the best combination for each pixel. As the result, an image of the printed object is shown in Fig. 4.5(a), which consists of four layers of inks on the wax candle. The layer images and the error map are shown in Fig. 4.5(b-e) and (f), respectively.

4.3. Discussion

In this chapter, we proposed a novel method to control the translucency by UV printing on a translucent material. We applied Kubelka's layer model with few measurements to render the translucency of the printed object, and then built the lookup table. Given a query about translucency, it is possible to find a combination of the factors in print, such as base materials, inks, and the number of printed layers, in the lookup table so that the translucency by the combination is the most similar to the query. That is, we can control the translucency of the printed object.

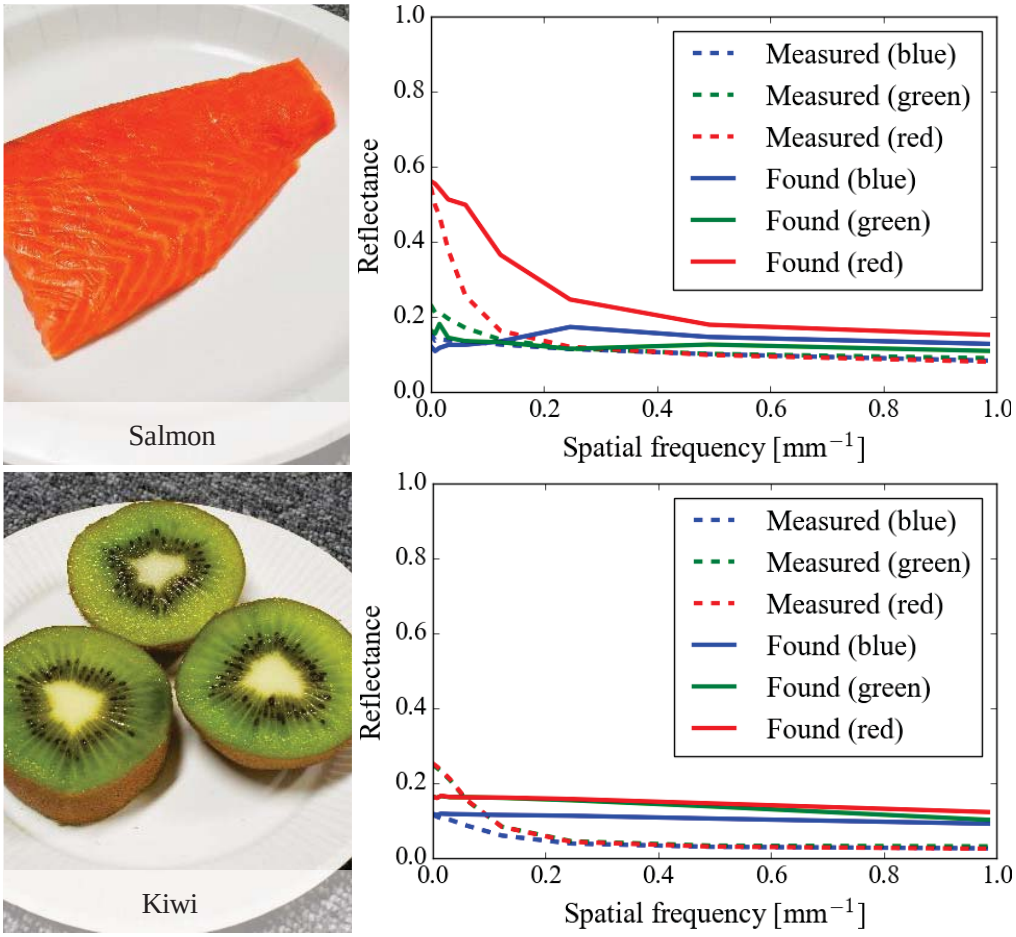


Figure 4.4: Controlling the reflective MTFs for the queries based on real-world objects; (a) salmon and (b) kiwi.

There are some limitations. Currently, we assume the PSFs are isotropic but some real-world objects has anisotropic PSFs. Since the print-head of a conventional UV printer moves within a 2D space, it is difficult to print on a complex-shaped object but a transfer sheet for UV printing could resolve that problem.

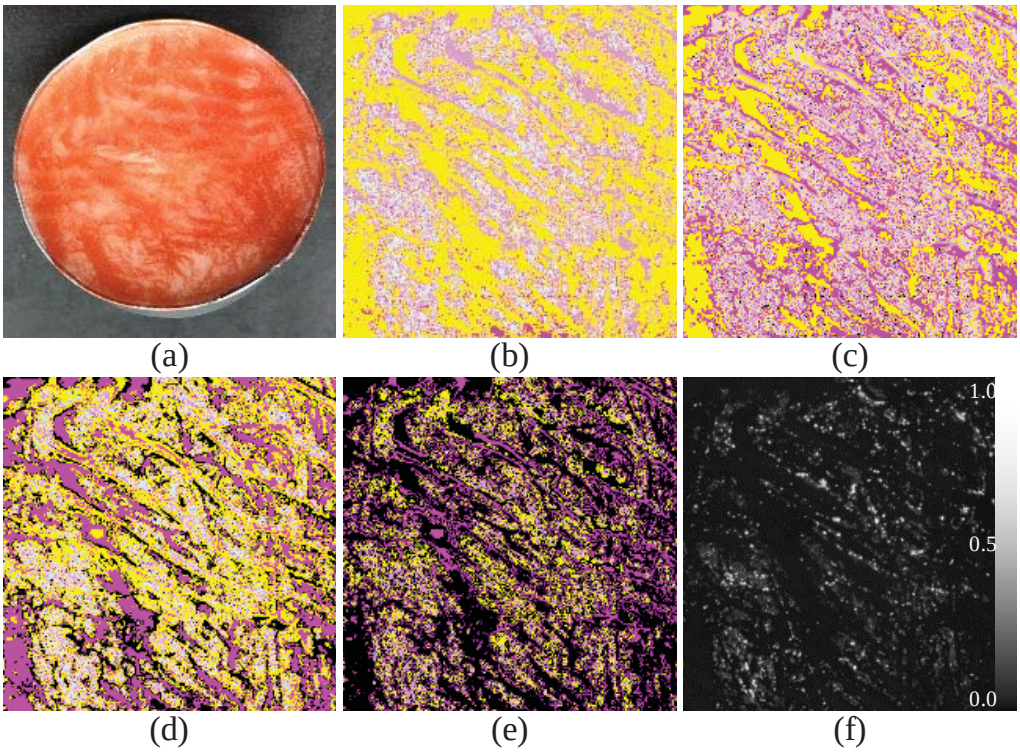


Figure 4.5: Replicating the reflective MTFs of the salmon. (a) The printed object. (b-e) The four layer images from the bottom to the top. (f) The error map.

Chapter 5

Reconstruction of inner layers using frequency correlation imaging

Observing inside an object is still a difficult task but also important for a wide variety of applications, such as bio-imaging, medical imaging, and industrial inspection. Therefore, various imaging techniques have been developed to achieve a clear observation inside an object, such as X-ray fluorescence technique [26], infrared reflectography [43], and terahertz imaging technique [69]. Researchers in computational photography have used many kinds of coded projections as spatial modulation. Levoy *et al.* [95] proposed a combination of a synthetic aperture technique [96] and a confocal imaging technique [86] to reconstruct images behind occlusions. Fuchs *et al.* [40] used the confocal imaging for recovering a solid target in scattering media. Those methods utilize a pair of a projector and a camera, called Pro-Cam system, to implement confocal imaging, for which both the projector and camera need to focus on the same depth plane. Nayer *et al.* [122] proposed high frequency illumination to separate observations into two components based on direct and global illuminations by projecting spatially high frequency patterns, which is typically spatial modulation. Achar and Narasimhan [4] extended the high frequency illumination for a multi-focus projector to recover scene shape and global illumination, simultaneously. This method is based on a similar idea on which physically changing the projector focus is used

for a spatio-temporal modulation. Instead, the proposed method in this chapter just changes the frequency of a projection pattern to realize a spatio-temporal modulation. Gupta *et al.* [57] obtained depths with direct and global separation. Tanaka *et al.* [158] proposed multi-frequency illumination to recover inner slices of a translucent object. This approach is also similar to ours because of changing the frequency of a projection pattern. However, it does not modulate a signal, so that it is required to use an empirical optimization.

A temporal modulation is another way to analyze a scene. Heide *et al.* [65] swept a modulated frequency and phase of a time-of-flight (ToF) camera to recover the light propagation inside scattering medium. Kadambi *et al.* [79] built a coded-illumination ToF camera with a deconvolution technique. Tadano *et al.* [154] proposed a coded ToF camera which is capable to select a target depth. These temporal modulation systems could be useful to reconstruct inner layer reflectance. However, a ToF method requires a temporally severe synchronization, while our Pro-Cam system just modulate a projection pattern.

In this chapter, we propose a novel technique to reconstruct reflectance maps in inner layers of an object by using a Pro-Cam system. An easy implementation consists of a perspective projector and an orthographic camera. When projecting a sinusoidal pattern at a spatial frequency, a spatial frequency observed by the camera varies with respect to a depth. However, it is not enough to reconstruct reflectance in an inner layer because of its heterogeneity. Thus, we change the frequency of the projection pattern multiple times. It enables to modulate a temporal frequency depending on the position of a 3-D point in the scene. Since an observed intensity includes response signals at various 3-D points, the response signal at a certain point has to be extracted from observations to reconstruct the reflectance. To do that, we employ a heterodyning technique, which we briefly explain in Section 5.1. We introduce the proposed method, *frequency correlation imaging*, in Section 5.2, perform simulations and analyses in Section 5.3, and make experiments on real data in Section 7.3. Finally, we discuss the proposed method and conclude this chapter in Section 5.5.

5.1. The basics of heterodyning

Heterodyning is a traditional signal processing technique, invented by Fessenden [135], to generate new frequencies by combining two or more frequencies. The technique is often used for shifting a range of frequency into a new different range, as shown in Fig. 5.1. Now, we briefly introduce the algorithm of heterodyning. Let us think about a case where we want to obtain an oscillated signal at a frequency f_1 in an observation $I(t)$. The observation is assumed as

$$I(t) = D + A(t)\sin(\omega_1 t + \phi(t)), \tag{5.1}$$

where D is the offset, $A(t)$ the amplitude, $\omega_1 \coloneqq 2\pi f_1$ the angular frequency, and $\phi(t)$ the phase of the signal. It often happens that the frequency f_1 is too high to detect directly. Here, when a sinusoidal signal at a frequency f_2 is multiplied, it becomes

$$\begin{aligned} & I(t)\sin(\omega_2 t) && (5.2)\\ & = D\sin(\omega_2 t) + A(t)\sin(\omega_2 t)\sin(\omega_1 t + \phi(t)) && (5.3)\\ & = D\sin(\omega_2 t) - \frac{A(t)}{2}\cos\{(\omega_1+\omega_2)t + \phi(t)\} && (5.4)\\ & \quad + \frac{A(t)}{2}\cos\{(\omega_1-\omega_2)t + \phi(t)\}, && (5.5) \end{aligned}$$

where $\omega_2 \coloneqq 2\pi f_2$. As a result, new frequencies $\omega_1+\omega_2$ and $\omega_1-\omega_2$ are generated. By using a low pass filter, only the third term remains, as below;

$$\langle I(t)\sin(\omega_2 t)\rangle_{\text{low}} = \frac{A(t)}{2}\cos\{(\omega_1-\omega_2)t + \phi(t)\}. \tag{5.6}$$

It becomes a low frequency signal, so that an analysis of the signal gets easier than the observation. This procedure is called heterodyne detection.

Moreover, when $f_2 = f_1$, it is independent to f_1, as below;

$$\langle I(t)\sin(\omega_2 t)\rangle_{\text{low}} = \frac{A(t)}{2}\cos\phi(t). \tag{5.7}$$

However, it is still impossible to separate the amplitude $A(t)$ and the phase $\phi(t)$.

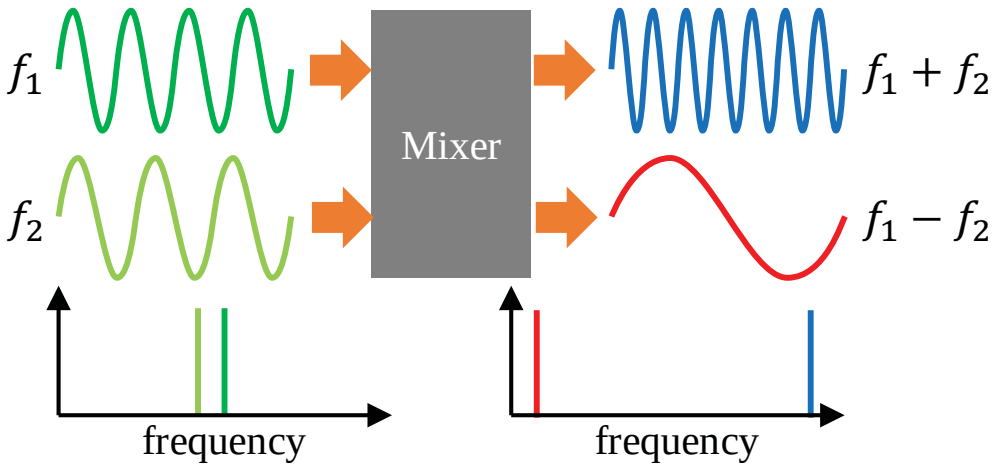


Figure 5.1: The basics of heterodyning, a signal processing technique to shift a range of frequency.

Thus, multiplying Eq. (5.1) by $\cos(\omega_2 t)$, as well, it becomes

$$I(t)\cos(\omega_2 t) \tag{5.8}$$
$$= D\cos(\omega_2 t) + A(t)\cos(\omega_2 t)\sin(\omega_1 t + \phi(t)) \tag{5.9}$$
$$= D\cos(\omega_2 t) + \frac{A(t)}{2}\sin\{(\omega_1 + \omega_2)t + \phi(t)\} \tag{5.10}$$
$$+ \frac{A(t)}{2}\sin\{(\omega_1 - \omega_2)t + \phi(t)\}. \tag{5.11}$$

As well, by using a low pass filter, the third term remains, as follow;

$$\langle I(t)\cos(\omega_2 t)\rangle_{\text{low}} = \frac{A(t)}{2}\sin\phi(t). \tag{5.12}$$

Finally, it is possible to reconstruct $A(t)$ and $\phi(t)$ from Eqs. (5.7) (5.12). This procedure is called direct-conversion.

5.1.1 Imaging using heterodyning

In fact, heterodyning has been applied for imaging techniques, such as synthetic array heterodyne [153], interferometry [53], and optical coherence tomography (OCT) [70]. Because those optical heterodyning are based on modulations of the phase and/or frequency, that is, wavelength, of light, a severe alignment

is required to work properly. Especially, since interferometry and OCT require coherent light, it cannot be applied to reconstruct a deep inner layer. Also, indirect time-of-flight imaging employs heterodyning based on temporal modulation of the amplitude [29], which is basically used for measuring a distance, not for reconstructing reflectance in an inner layer.

In this chapter, we propose a novel spatio-temporal modulation method in which the temporal frequency of illumination is modulated depending on the position of a 3-D point. A response signal is mixed with other signals until being observed. Therefore, we employ the direct-conversion technique to extract the response signal for reconstructing the reflectance.

5.2. Frequency correlation imaging

Now, we begin with introducing the proposed method to reconstruct reflectance maps in inner layers of an object. Firstly, we formulate a novel spatio-temporal modulation technique, secondly build the observation model in a Pro-Cam system, and then, finally, explain a reconstruction of reflectance at each point in inner layers, based on the heterodyning technique. We call the proposed method *frequency correlation imaging*.

Now, we assume that a Pro-Cam system consists of a perspective projector and an orthographic camera, as shown in Fig. 5.2. When projecting a sinusoidal pattern onto a reflectance standard located at a depth z_1, let f_1 a frequency of the pattern on the image. According to a nature of an orthographic camera, the size of an object on the image is always the same regardless of a distance between the object and the camera. On the other hand, a projection pattern from the projector spatially broadens with respect to a depth. Therefore, when the standard is moved to another depth $z_2 > z_1$, a frequency f_2 on the image becomes lower than f_1, that is, $f_2 < f_1$. This relationship can be held in any case between two depths. Thus, we formulate such a depth-dependent frequency as a spatial modulation in Section 5.2.1.

Theoretically, as long as the depth-dependent frequency is available, various setups can be built up. For example, it is possible to pair a perspective projector with a perspective camera, whose focal length is different from projector's one,

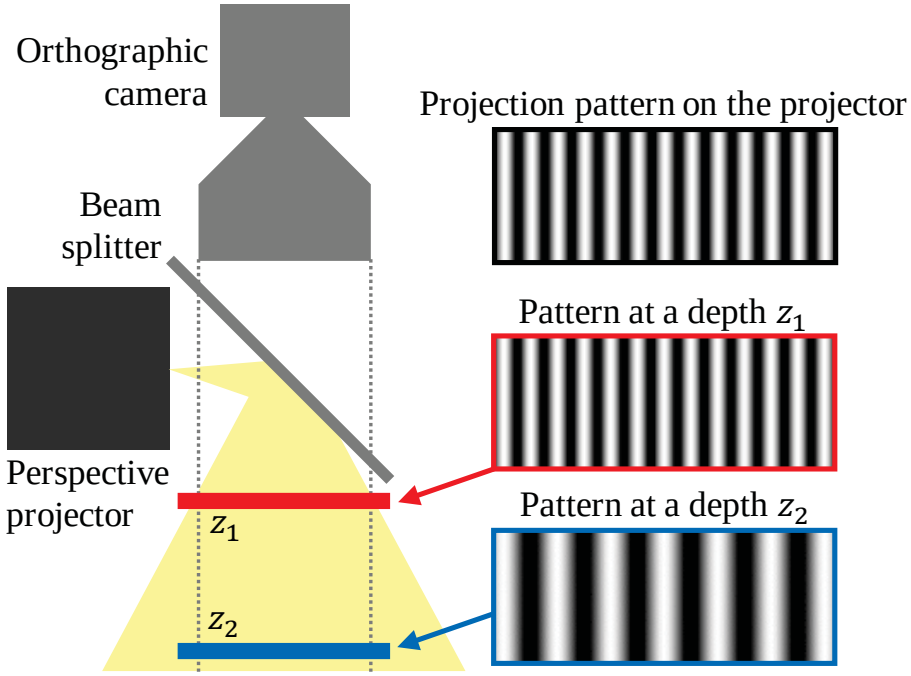


Figure 5.2: An example of a Pro-Cam system consisting of a perspective projector and an orthographic camera.

such as a coaxial setup and a parallel planes setup, as shown in Fig. 5.3. Moreover, it is also implementable that a projector is rotated to a camera and a projector is located even behind an object to observe transmission. Hereafter, we assume an easy setup consisting of a perspective projector and an orthographic camera to explain the spatial modulation, the observation model, and the reconstruction method.

5.2.1 Spatial modulation for depth-dependent frequency

We assume the projector is located as its projection plane is parallel to the image plane and both spatial axes on the projection plane are also parallel to those on the image plane, as shown in Fig. 5.2. The origin of the image plane is regarded as the origin of the Pro-Cam system. Also, the projector is assumed to be available for focus-free projection, such as a laser projector. Now, we think a case where it projects a 1-D sinusoidal pattern along x axis. The projection pattern L_1 at a

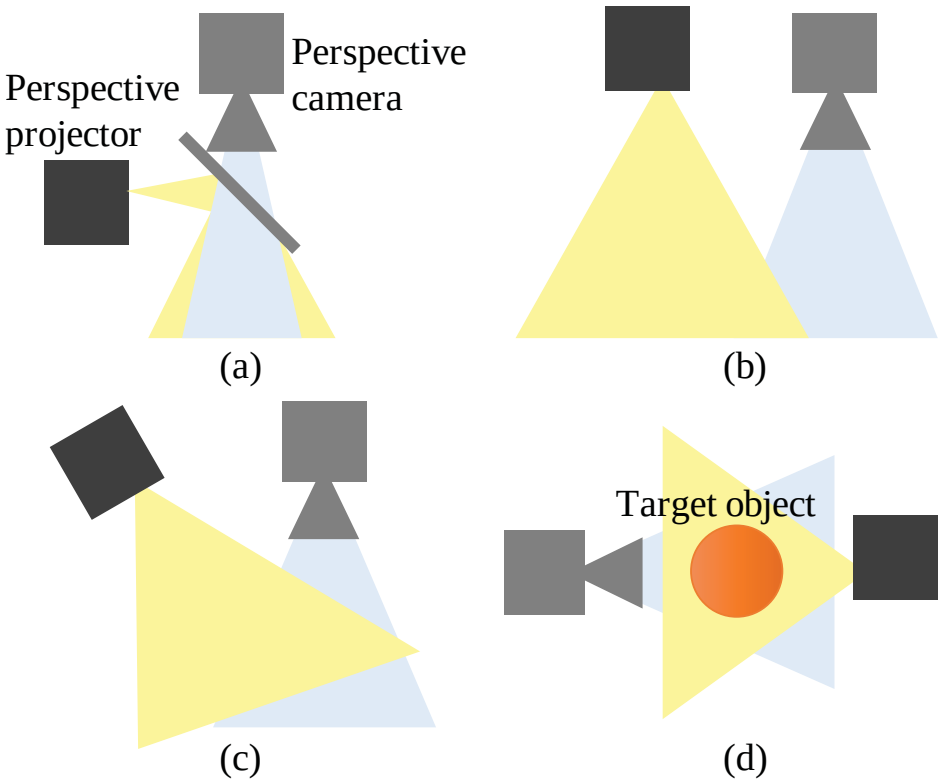


Figure 5.3: Various Pro-Cam setups for frequency correlation imaging. Theoretically, as long as the depth-dependent frequency is observed in the system, a modulation model can be formulated. (a) a coaxial setup with different focal lengths, (b) a parallel planes setup with the same focal lengths, (c) a rotated projector setup, and (d) a transmission setup.

depth z_1 is written as below;

$$L_1(x, y) = D_1 + A_1 \sin(\omega_1 x + \phi), \tag{5.13}$$

where D_1 is the offset, A_1 the amplitude, ω_1 the angular frequency, and ϕ the phase of the sinusoidal pattern. (x, y) is a position on the image plane and the x axis is the modulation axis. Since a projection pattern from the perspective projector is spatially broadened with respect to a depth, the pattern at a different depth $z_2(= z_1 + \delta z)$ is represented as

$$L_2(x, y) = D_2(x, y) + A_2(x, y) \sin(\omega_2 x + \phi), \tag{5.14}$$

where $D_2(x, y), A_2(x, y)$ and ω_2 are the offset, amplitude, and angular frequency of the sinusoidal pattern at the depth z_2, respectively. By using the angular frequency at the depth z_1, that at the depth z_2 can be formulated as below;

$$\omega_2 = \frac{z_1}{z_2}\omega_1. \tag{5.15}$$

Note that the phase of the pattern does not change for each of the depths. Here, light to illuminate the depth z_2 can be regarded as one transmitted by a plane at the depth z_1. If transmittance at the depth z_1 is spatially homogeneous, the offset D_2 and amplitude A_2 at the depth z_2 are spatially invariant. For a general scene, the transmittance spatially varies, so that the offset and amplitude become position-dependent.

In general, it can be formulated for a discrete depth $z_i(= z_1 + (i-1)\delta z)$, as below;

$$L_i(x, y) = D_i(x, y) + A_i(x, y) \sin(\omega_i x + \phi), \tag{5.16}$$

where $L_i(x, y)$ is the projection pattern at a depth z_i and $D_i(x, y), A_i(x, y)$ and ω_i are the offset, amplitude, and angular frequency of the pattern. $D_i(x, y) \leq D_j(x, y)$ and $A_i(x, y) \leq A_j(x, y)$ hold for any $i > j$. Based on the first depth z_1, the angular frequency at i-th depth z_i can be written as

$$\omega_i = \frac{z_1}{z_1 + (i-1)\delta z}\omega_1. \tag{5.17}$$

That is, the angular frequency of a projected sinusoidal pattern is reduced relative to a depth.

5.2.2 Observation model in the Pro-Cam system

The image plane is parallel to the projection one but the optical axes are not necessary to be coaxial. Now, the difference between the origins of the image and projection planes is defined as (x_p, y_p). In fact, the size of a pixel in the projection plane is not necessarily the same as that in the image one. Thus, we define the size of a projected pixel at the first depth z_1 relative to that in the image plane as the scale factor s. Finally, when the origin of the projection plane is projected at (x_p, y_p) in the image plane, Eq. (5.16) is rewritten as

$$L_i(x, y) = D_i(\hat{x}, \hat{y}) + A_i(\hat{x}, \hat{y}) \sin(\omega_i \hat{x} + \phi), \tag{5.18}$$

where $\hat{x}(x) \coloneqq s(x - x_p)$ and $\hat{y}(y) \coloneqq s(y - y_p)$. Here, we omit the arguments of them to improve the readability, as long as obvious.

An observed intensity $I(x, y)$ on the image plane is represented as a summation of all intensities from different depths, as below;

$$I(x, y) = \sum_i I_i(x, y), \tag{5.19}$$

where $I_i(x, y)$ is an intensity at a depth z_i. When a scene is projected by the projector explained above, a discrete scene point $p_i(x, y)$ at i-th depth z_i is illuminated as $L_i(x, y)$. Assuming reflectance of the scene point is $R_i(x, y)$, an intensity $I_i(x, y)$ can be represented as

$$I_i(x, y) = L_i(x, y) R_i(x, y) E_i(x, y), \tag{5.20}$$

where $E_i(x, y)$ is an attenuation rate at which the light reaches to the camera after being reflected at the scene point and then passing through the scene. Assuming light not reflected on the plane at a depth $z_{i-1} (i \geq 2)$ is all transmitted to the next depth z_i, the offset and amplitude at the depth z_i can be written as

$$D_i(x, y) = T_{i-1}(r_i x, r_i y) D_{i-1}(\hat{x}(r_i x), \hat{y}(r_i y)), \tag{5.21}$$

$$A_i(x, y) = T_{i-1}(r_i x, r_i y) A_{i-1}(\hat{x}(r_i x), \hat{y}(r_i y)), \tag{5.22}$$

where $r_i \coloneqq \frac{z_1 + (i-2)\delta z}{z_1 + (i-1)\delta z}$ and $T_i(x, y) \coloneqq 1 - R_i(x, y)$. Because the light reflected at i-th depth z_i attenuates at all of the first to the $(i-1)$-th depths, the attenuation

rate can be written as

$$E_i(x,y) = \prod_{j=1}^{i-1} T_j(x,y). \tag{5.23}$$

Finally, the observed intensity $I(x,y)$ on the image plane is written as below;

$$\begin{aligned} &I(x,y) \\ &= \sum_i L_i(x,y) R_i(x,y) E_i(x,y) &(5.24)\\ &= \sum_i \left\{ D_i^*(x,y) + A_i^*(x,y) \sin(\omega_i \hat{x} + \phi) \right\}, &(5.25)\\ &\text{where,} &(5.26)\\ &\quad D_i^*(x,y) \coloneqq R_i(x,y) E_i(x,y) D_i(x,y), &(5.27)\\ &\quad A_i^*(x,y) \coloneqq R_i(x,y) E_i(x,y) A_i(x,y). &(5.28) \end{aligned}$$

5.2.3 Reconstruction of reflectance maps in layers

Let us consider a case where the angular frequency of the sinusoidal pattern temporally varies. An observed intensity $I(x,y,t)$ at a time t under projecting a sinusoidal pattern with a time-varying angular frequency $\omega_i(t)$ can be represented as

$$I(x,y,t) = \sum_i \left\{ D_i^*(x,y) + A_i^*(x,y) \sin(\omega_i(t) \hat{x} + \phi) \right\}. \tag{5.29}$$

In practice, changing the angular frequency ω_1 at the depth z_1 leads the time-varying angular frequency $\omega_i(t)$ based on Eq. (5.17). Here, we utilize the direct-conversion. As multiplying the both sides of Eq. (5.29) by $\sin(\omega_k(t)\hat{x})$, it becomes

$$\begin{aligned} &I(x,y,t) \sin(\omega_k(t)\hat{x}) \\ &= \sum_i \Bigg\{ D_i^*(x,y) \sin(\omega_k(t)\hat{x}) \\ &\quad - \frac{A_i^*(x,y)}{2} \cos((\omega_i(t) + \omega_k(t))\hat{x} + \phi) \\ &\quad + \frac{A_i^*(x,y)}{2} \cos((\omega_i(t) - \omega_k(t))\hat{x} + \phi) \Bigg\}. &(5.30) \end{aligned}$$

Focusing on the right side of Eq. (5.30), only when $i = k$, the time-dependent part, $\omega_i(t) - \omega_k(t)$, in the third term becomes time-independent. That is, the third

term becomes the DC component in the time-series data $I(x, y, t)\sin(\omega_k(t)\hat{x})$, only when $i = k$. Therefore, defining an extraction of the DC component of time-series data $\tau(t)$ through Fourier transform as $\langle\mathcal{F}[\tau(t)]\rangle_{\mathrm{DC}}$, the DC component can be extracted as below;

$$\langle\mathcal{F}[I(x, y, t)\sin(\omega_k(t)\hat{x})]\rangle_{\mathrm{DC}} = \frac{A_k^*(x, y)}{2}\cos\phi. \tag{5.31}$$

As well, by multiplying the both sides of Eq. (5.29) by $\cos(\omega_k(t)\hat{x})$, we can get

$$\langle\mathcal{F}[I(x, y, t)\cos(\omega_k(t)\hat{x})]\rangle_{\mathrm{DC}} = \frac{A_k^*(x, y)}{2}\sin\phi. \tag{5.32}$$

Because of $\cos^2\phi + \sin^2\phi = 1$, a summation of squares of Eqs. (5.31)(5.32) becomes

$$\left\{\frac{A_k^*(x, y)}{2}\cos\phi\right\}^2 + \left\{\frac{A_k^*(x, y)}{2}\sin\phi\right\}^2 = \frac{A_k^*(x, y)^2}{2}. \tag{5.33}$$

Thus, we can obtain $A_k^*(x, y)$ for any k-th depth z_k.

Reconstruction begins at the first depth to deeper depths. According to Eq. (5.28), $A_1^*(x, y)$ for the first depth z_1 is $A_1^*(x, y) = R_1(x, y)E_1(x, y)A_1(x, y)$. Because the first depth is assumed to be the most closest plane to the camera, the light reflected at the first depth does not attenuate, that is $E_1(x, y) = 1$. Therefore, once the amplitude $A_1(x, y)$ at the first depth is measured in advance, the reflectance $R_1(x, y)$ can be reconstructed, as below;

$$R_1(x, y) = \frac{A_1^*(x, y)}{A_1(x, y)}. \tag{5.34}$$

The amplitude $A_2(x, y)$ and attenuation rate $E_2(x, y)$ at the second depth z_2 can be computed by Eq. (5.22) and Eq. (5.23), respectively. Thus, the reflectance $R_2(x, y)$ can also be reconstructed, as well. For a general depth z_i, the reflectance $R_i(x, y)$ can be reconstructed as follow;

$$R_i(x, y) = \frac{A_i^*(x, y)}{E_i(x, y)A_i(x, y)}. \tag{5.35}$$

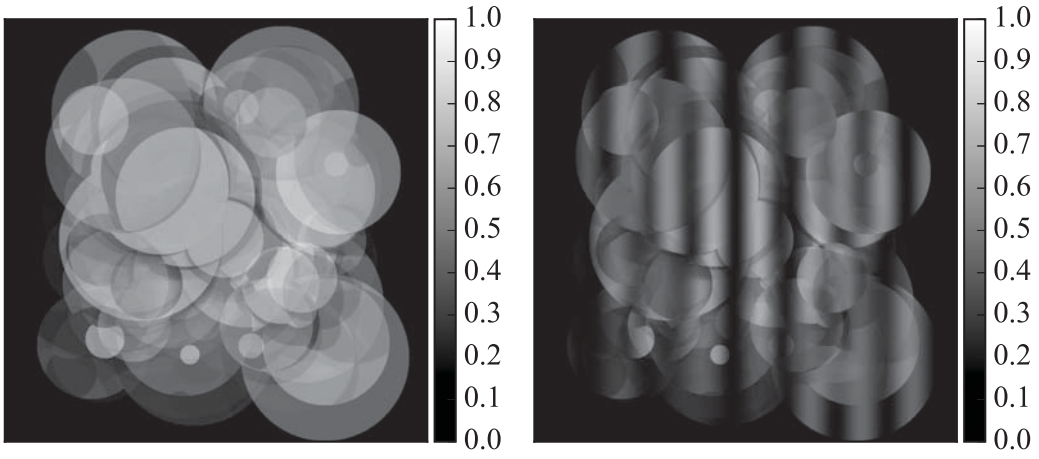


Figure 5.4: Rendered images of the ten layers object. On the left, the object is illuminated by a whole white pattern, while a sinusoidal pattern on the right. The bars on the right of images indicate reflectance.

5.3. Simulations and analyses

5.3.1 Validation

We evaluate the proposed method on synthetic data. To render the synthetic data, we simulate the same setup as explained in Section 5.2, consisting of a perspective projector and an orthographic camera. In this simulation, we set the origin of the projection plane to $(x_p, y_p) = (0, 0)$, the scale factor to $s = 1$, and the first depth to $z_1 = 100$mm. The size of an image is 1000×1000 pixels, as well as the size of a projection pattern. For projection, the offset and amplitude at the first depth z_1 is set to $D_1 = 0.5$ and $A_1 = 0.5$, as the maximum brightness of the projector is 1. The frequency f_1 of a sinusoidal pattern at the first depth z_1 is set to every 0.001Hz in the range of $0.001 \leq f_1 < 0.5$Hz, where the angular frequency $\omega_1 = 2\pi f_1$.

Using an object consisting of multiple layers as a target, we evaluate the accuracy of reconstruction. The object consists of ten layers whose gap is set to $\delta z = 10$mm. There are circles on the layers, where are filled with different reflectances. In Fig. 5.4, rendered images when projecting a whole white pattern and a sinusoidal pattern are shown. Figure 5.5 shows the reference reflectance

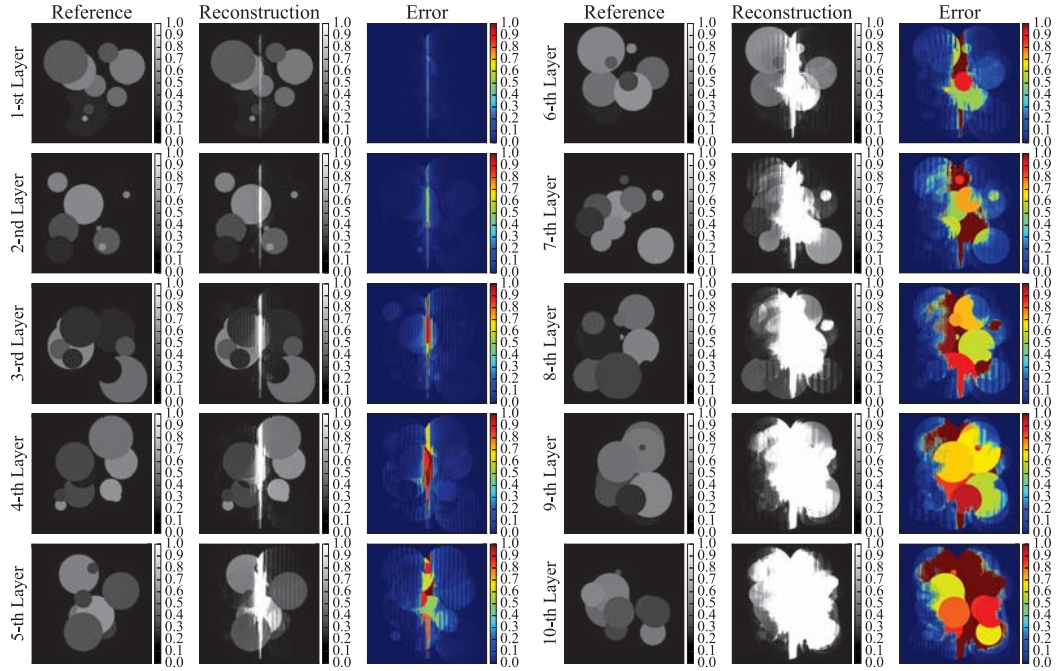


Figure 5.5: Evaluation of reconstructing reflectance maps in the ten layers. The PSNRs in the reconstruction are 28.89, 22.91, 17.53, 15.53, 14.70, 11.97, 9.17, 7.71, 6.83, and 4.95dB for the top and bottom layers, respectively.

map, the reconstructed one, and the reconstruction error map of each of the inner layers. For the first and second layers, the reconstructed maps seem quite similar to the reference ones. In fact, peak signal-to-noise ratios (PSNRs) between those are 28.89 and 22.91dB for the first and second layers, respectively. As we can see, the errors around the central region are large. This is because light rays around the origin of the projection plane do not broaden for the orthographic camera. Therefore, all the three terms in Eq. (5.30) are mostly regarded as the DC components, even though the frequency ω_i varies, which means the frequency is not modulated. It results in that heterodyning does not work properly around the region. The remaining PSNRs for the third to tenth layers are 17.53, 15.53, 14.70, 11.97, 9.17, 7.71, 6.83, and 4.95dB, respectively. As easily imagined, the deeper a layer is, the low the accuracy of reconstruction is. In a case with a number of layers, it is difficult for light to reach deeper layers and also to come

back to the surface of an object. That is a main cause why the accuracy of reconstruction gets worse as the number of layers increases.

5.3.2 Analysis of the depth resolution

We analyze the depth resolution of the proposed method. For validation, we only set the first depth and the gap to $z_1 = 100$mm and $\delta z = 10$mm for simplicity. Here, we evaluate the accuracy of reconstruction in PSNR when the first depth and the gap are changed. We change the first depth in a range of $1 \leq z_1 \leq 500$mm and the gap in a range of $0.1 \leq \delta z \leq 30$mm, and then render two layers object, which are extracted from the ten layers. The frequencies for projection are also the same as that on the two layers object. The PNSRs while changing the first depth and the gap are shown in Fig. 5.6. Looking around a range of smaller gaps, cases where the first depth is smaller resulted in accurate reconstructions, relatively. In simulation, the projector is assumed to project the whole scene, so that, in those cases, the angle of projection becomes quite large. It results in that the amount of change in the depth-dependent frequency is made large. That is why those cases has a high resolution in depth. On the other hand, because the frequency fast decreases for a deeper depth, the accuracy when the gap is set to a large value decreases. In practice, it is almost impossible to set the first depth to a small value such as $z_1 = 1, 5, 10$mm. In a practical region of the first depth, the accuracy increases as well as the gap increases.

5.3.3 Analysis of the effect of the amount of change in frequency

For the temporal modulation in frequency domain, it is important how to change the frequency. Here, we analyze the effect of the amount of change in frequency. In simulation, we use the two layers object again. The first depth and the gap are set to $z_1 = 50$mm and $\delta z = 15$mm because it seems to lead an accurate reconstruction as a possibly implementable setup, according to Fig. 5.6. Now, we change the frequency of the projection pattern in Eq. (5.29). The frequency is changed by

$$f_1(t) = f_1^0 + \delta f t, \tag{5.36}$$

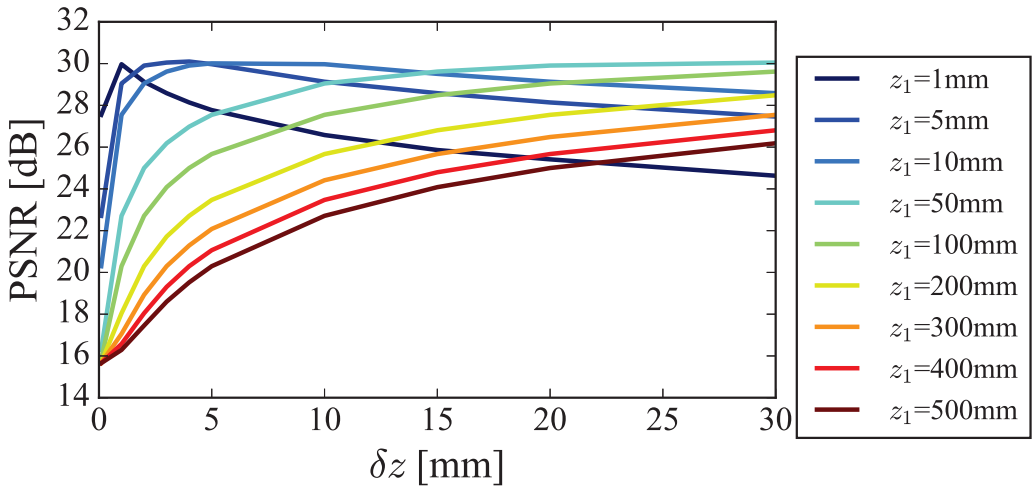


Figure 5.6: Analysis of the depth resolution.

where f_1^0 is the offset frequency, now set to 0.001Hz because the spatial resolution of the projector is 1000 pixels in the modulation axis. The time factor t is currently just a frame number. When changing the amount of change δf in a range of $0.001 \leq \delta f \leq 0.12$Hz, we evaluate the PNSR in reconstruction. As a result, shown in Fig. 5.7, the accuracy of reconstruction gradually decreases as the amount of change in frequency δf increases. Thus, it can be said that the amount of change in frequency for the temporal modulation should be set to a small value.

5.4. Experiments on real data

Finally, we make experiments on real data. A setup consists of a perspective laser projector (SK Telecom, LB-UH6CB, 1280×720) and an RGB camera (FLIR, Grasshopper3, 1920×1440) with a telecentric lens (Edmund Optics, #55-348), as shown in Fig. 5.8. Note that we use linear polarizers to remove specular reflection which makes it difficult to observe projected patterns. A projection brightness can be set within the range of 0 to 255, so that the offset and amplitude of a sinusoidal pattern at the first depth z_1 is set to $D_1 = 127$ and $A_1 = 127$, respectively. The frequency f_1 of the pattern is set to every 7.8125×10^{-4}Hz

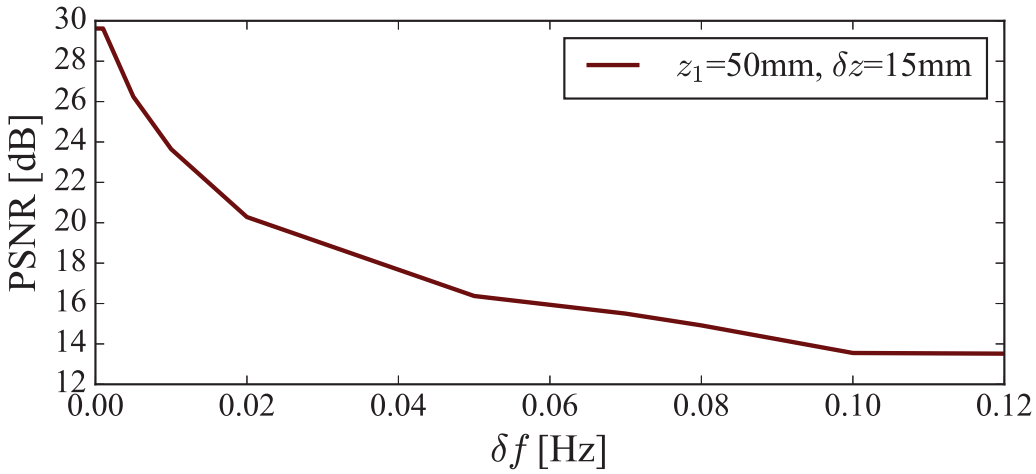


Figure 5.7: Analysis of the effect of the amount of change in frequency δf.

in the range of $7.8125 \times 10^{-4} \leq f_1 < 0.5$Hz. In practice, the projector is not perfectly focus-free despite a laser projector because it has a range of the working distance. According to our measurement, the projector requires at least 250mm as the working distance to focus. The telecentric lens also has practical limitations, which are a shallow depth of field (DoF) and a limited working distance. In the setup, the DoF of the capture system is 60mm and the closest working distance is 130mm based on our measurement.

5.4.1 Calibration

The first depth z_1, the scale factor s, and the origin of the projection plane (x_p, y_p) are required for the reconstruction. To achieve a better reconstruction, a spatial resolution of the projection should be higher in the image plane. However, in the setup, it is limited because of the working distance of the projector and the DoF and working distance of the telecentric lens. Therefore, we put a reflectance standard as close to the Pro-Cam system as a projected chessboard pattern is in focus, and then, obtain the factors. Firstly, to obtain the origin x_p, the projector projects a vertical line pattern whose central position is $x = 0$. The thickness of the line can be changed up to an observed intensity. The origin y_p is, as well, obtained by projecting a horizontal line pattern. Secondly, by projecting a thick

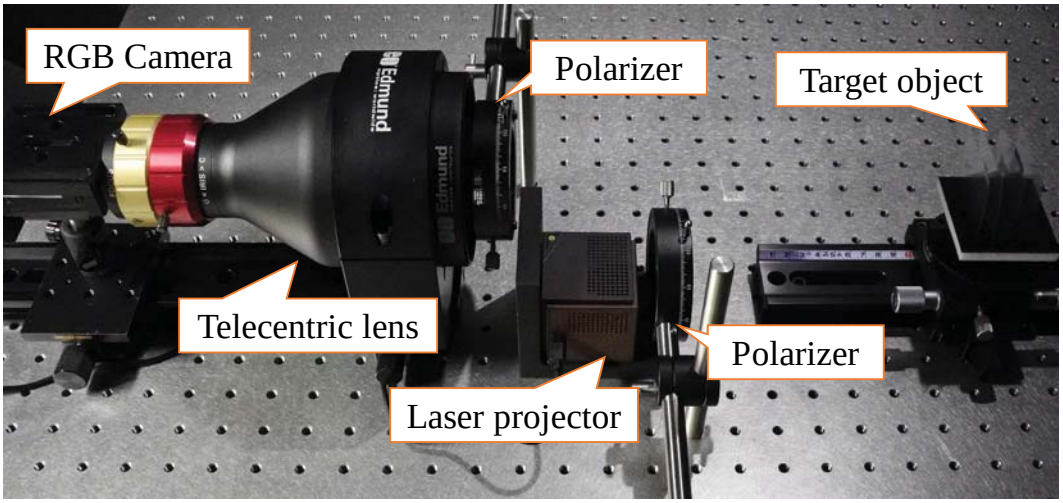


Figure 5.8: Experimental setup.

vertical line pattern, it is possible to obtain the scale factor s through a simple rate between the projected and an observed thicknesses in pixel. Finally, the first depth z_1 can be obtained from observations when moving the standard to another depth while projecting a thick line pattern. Let an observed thickness of the line at the first depth z_1 and one at another depth $z_1 + \delta z$ be w_1 and w_2, respectively. From the similarity relation, as well as Eq. (5.17), the following equation holds;

$$w_2 = \frac{z_1 + \delta z}{z_1} w_1. \tag{5.37}$$

Now, since the parameters expect for z_1 are known, the first depth z_1 can be obtained as

$$z_1 = \frac{w_1 \delta z}{w_2 - w_1}. \tag{5.38}$$

As a result, we obtained $x_p = 656.08$, $y_p = 27.00$, $s = 0.299$, and $z_1 = 252.83$mm in the setup.

5.4.2 Reconstruction of reflectance maps

Firstly, to evaluate the proposed method in the real world, we make an experiment on a simple object which consists of two overhead projector (OHP) sheets. On each of the OHP sheets, a different colored circle is printed, as shown in

Figure 5.9: Target objects used in the experiments on real data. Those consist of OHP sheets. (a) A circle, whose color and size are different, are printed on each. (b) CVPR logos in 2017 and 2018 are printed. (c) A, B, C, and D are printed, respectively.

Fig. 5.9(a). The gap between the OHP sheets is set to $\delta z = 10$mm. Figure 5.10 shows the reference reflectance map, the reconstructed one, and the reconstruction error map on each of the layers. Here, the reference reflectance maps are taken by observing each of the sheet under a white pattern illumination without collapsing its alignment. The circle on the top layer is clearer in the reconstructed top layer than the reconstructed bottom layer. The circle on the bottom layer can be seen through the circle on the top layer in the reconstructed bottom layer. As a result, the PSNRs in reconstruction are 34.32 and 30.76dB for the top and bottom layers, respectively.

The second target object consists of two OHP sheets on which the CVPR logos were printed, as shown in Fig. 5.9(b). The gap between the layers is set to $\delta z = 10$mm, as well. Figure 5.11 shows the reconstructed reflectance maps on the layers. The estimated top layer (a) seems only the logo in 2018 and we can see texts of "CVPR 2018" and "Salt Lake City" and the Wasatch Mountains. On the estimated bottom layer (b), it can be seen that a text of "CVPR 2017" appears on the Diamond Head, instead of the Wasatch Mountains, and even a small text of "21-26", which is the dates when the CVPR 2017 was held, can be

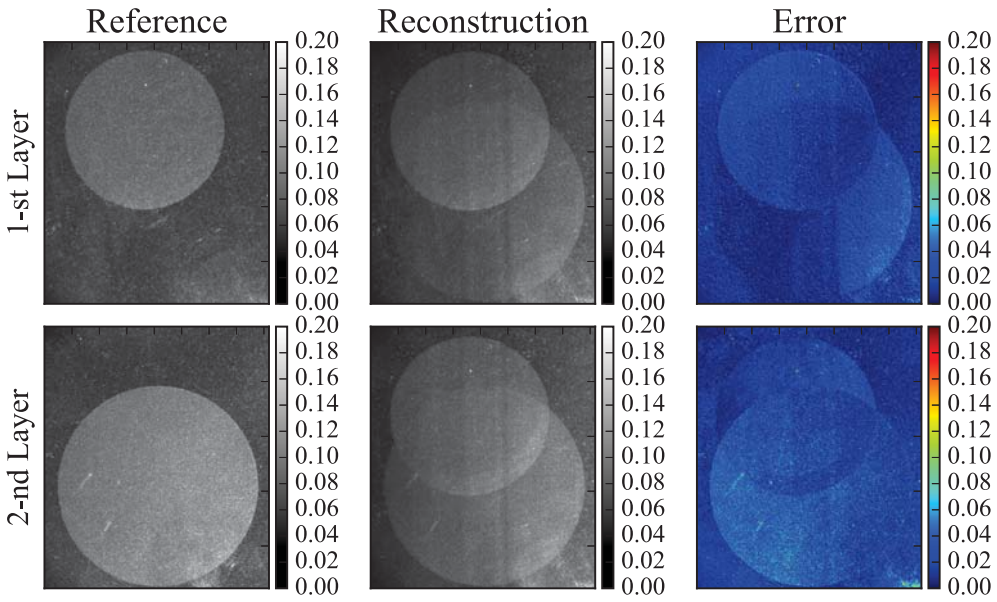


Figure 5.10: Experimental result on the real data. The target object consists of two OHP sheets on which a circle is printed. The gap between the sheets is set to 10mm.

seen. However it is strongly affected by the top layer due to its shadows.

The third target object also consists of four OHP sheets on which alphabets A, B, C, and D are printed from the first to fourth layers, respectively. The each gap is set to $\delta z = 10$mm, as well. Figure 5.12 shows the reconstructed reflectance maps on the layers. The alphabet A in Fig. 5.12(a) is quite clear comparing with the others. The alphabets B, C, and D are relatively clear in Fig. 5.12(b-d), respectively. However, it can be seen that a top layer affects to a bottom layers. In Fig. 5.12(d), all the alphabets can be seen to some extent. Finally, the reconstructed result of layers as the gap is set to $\delta z = 1$mm is shown in Fig. 5.13. The range of reconstructed depths is from 0 to 34mm and the target depths are writen on the left-top in each image. The depth resolusion of the current setup is not so high that the reconstructed images between the OHP sheets have some intensity, which are ideally supposed to be black images.

(a) (b)

Figure 5.11: Reconstruction of reflectance maps in two layers. The target object consists of two OHP sheets on which CVPR logos are printed.

5.5. Discussions

We proposed a novel technique to reconstruct reflectance maps in multiple inner layers of an object by using a Pro-Cam system. The key idea is spatio-temporally modulated illumination, which modulates a temporal frequency depending on the position of a 3-D point in a scene. A heterodyning technique leads extracting a response signal from observations under the proposed illumination. Finally, we reconstruct reflectance maps in inner layers based on the extracted response signals. In the simulations, we evaluated the accuracy of reconstruction and analyzed the depth resolution and the effect of the amount of change in frequency for the temporal modulation. The quantitative evaluation on the real data was also performed, which shows the proposed method can reconstruct reflectance maps in layers.

There are still some limitations in the proposed method. It assumes there is no refraction on a boundary but it always occurs in the real world. Refraction will also change the frequency, so it is a next challenge to model the phenomenon. In a general object, light is scattered and absorbed to some extent. Especially, light scattering is a well known phenomenon working as a low pass filtering, so that it is a significant issue for the proposed method. The scattering phenomenon is often formulated as a point spread function and the phenomenon can be written as a convolution operation [158]. This formulation could be installed in our model.

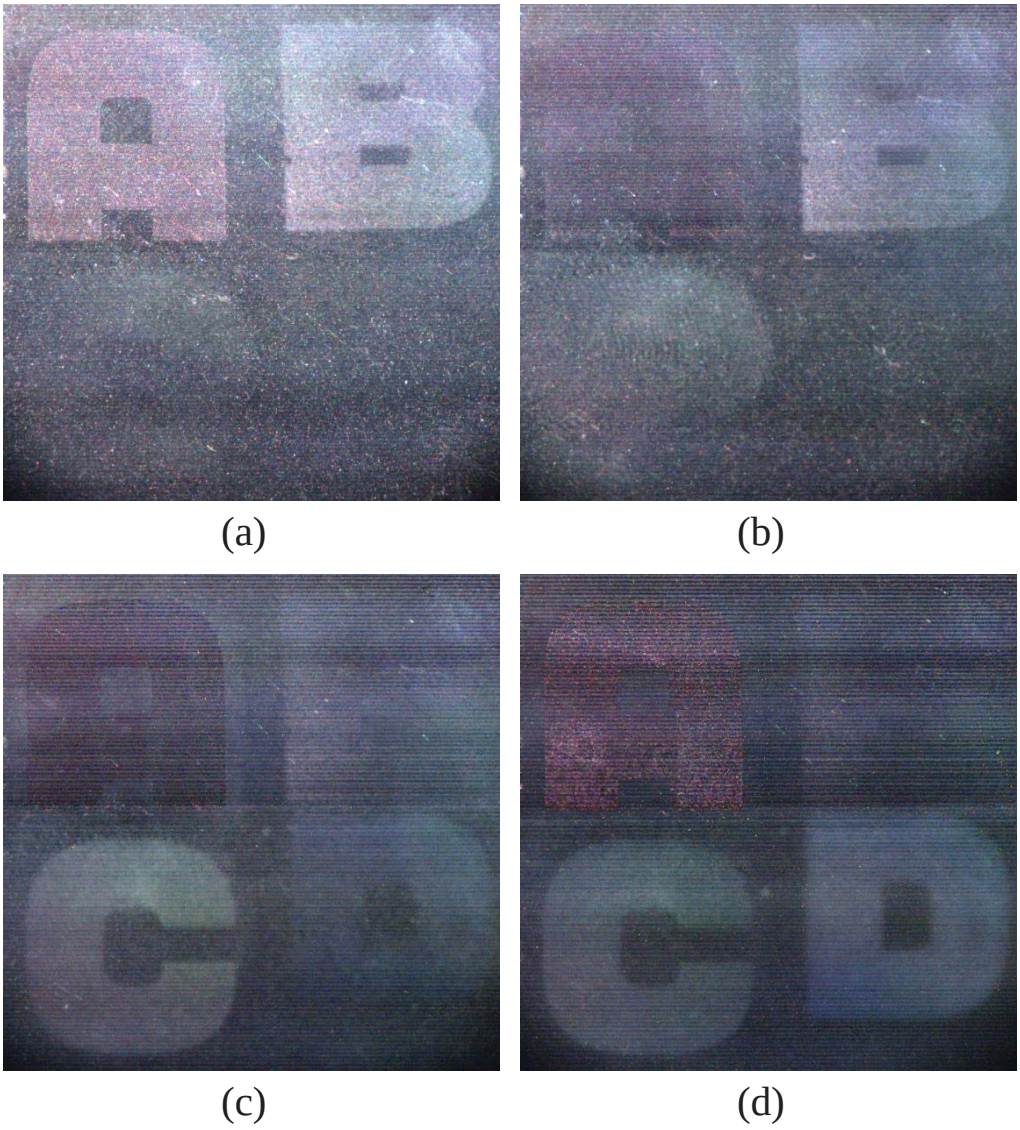


Figure 5.12: Reconstruction of reflectance maps in four layers. The target object consists of four OHP sheets on which alphabets A, B, C, and D are printed.

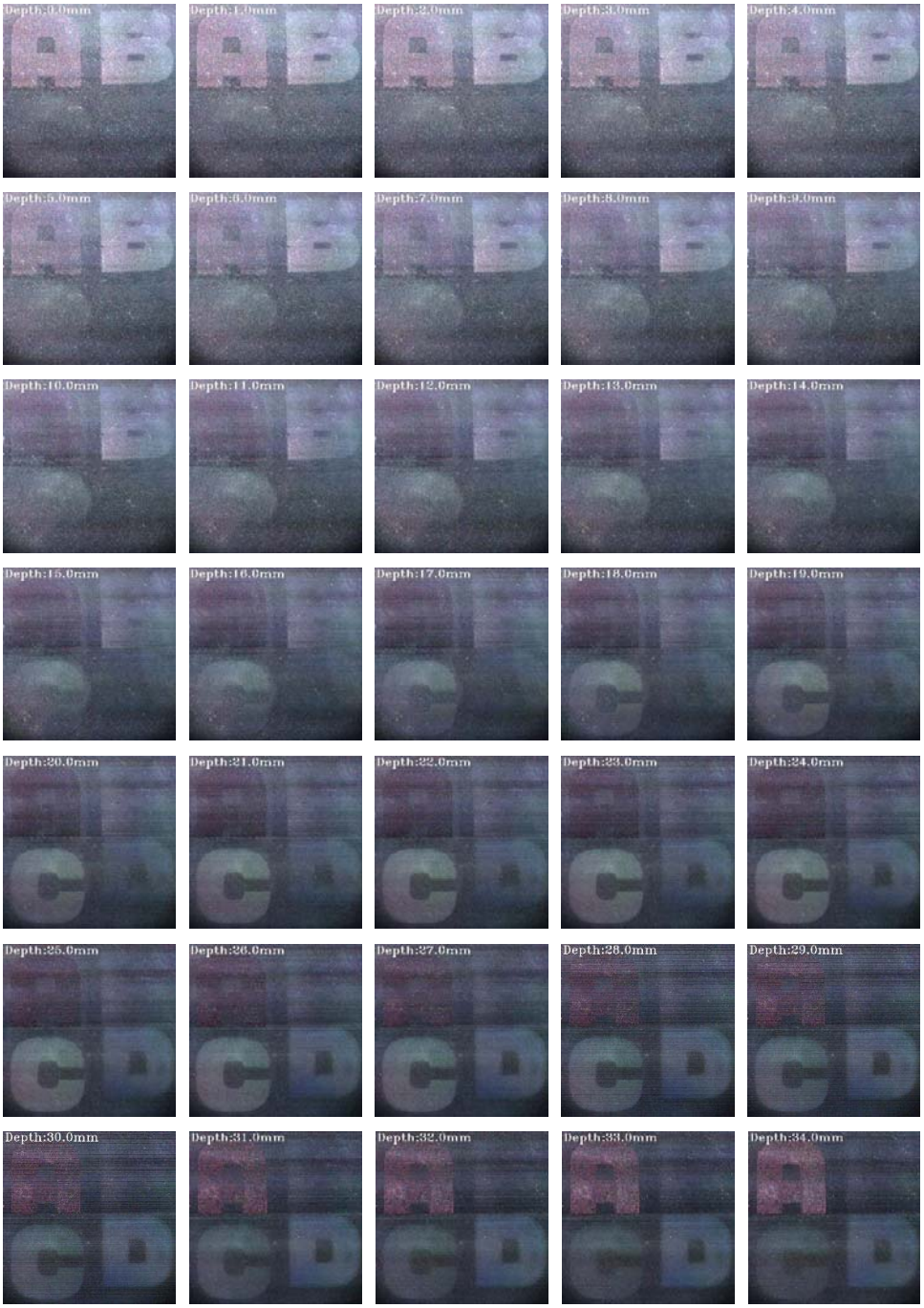


Figure 5.13: Analysis of the reconstruction in four layers. The gap for reconstruction is set to 1mm.

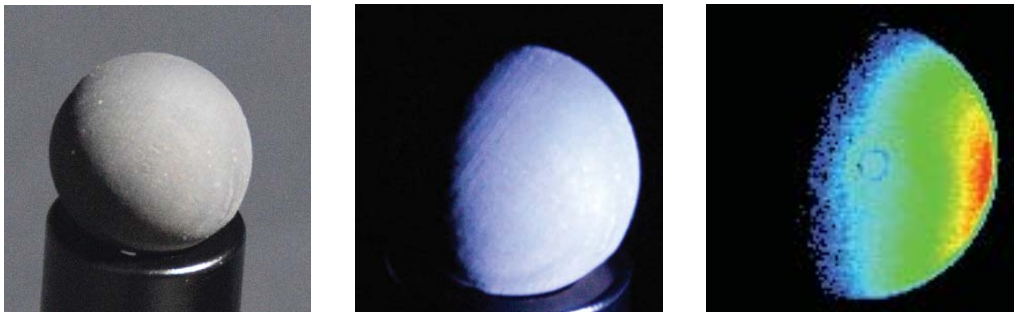
Chapter 6

Thermal photometric stereo

Light transport decomposition has attracted broad interest in the computer vision and computer graphics fields. This is because many computer vision techniques implicitly or explicitly assume only diffuse reflection, which simplifies observation models. An image in the real world is, however, composed of many optical components, such as specular reflection, inter-reflection, and subsurface scattering. Light transport decomposition plays an important role to bridge the gap between the real world and the models.

Prior works have decomposed light transport using color [147], polarization [166], and active illumination [124]. The time-resolved approach has emerged because the each optical components has different transient properties on the tens of pico seconds order [179]. There are multiple time-resolved approaches, for example, with the use of a femto-pulsed laser, interferometer [49], time-of-flight camera modifications [63, 88], and single-photon sensor [131]. Inspired by the temporal decomposition of light transport, we develop a novel time-resolved decomposition technique for far infrared light transport. A key observation is that the speed of heat propagation is extremely slow compared with the speed of light propagation. Using thermal imaging, the time-resolved decomposition is feasible at a video frame rate.

To date, thermal imaging has been treated as being different from visible light imaging: The thermal image represents the temperature of the object, while the visible light image reflects the visual information. We show, however, similar images can be obtained when the observation environment is appropriately con-



(a) A ball. (b) Color image. (c) Thermal image.

Figure 6.1: A ball captured by a conventional color camera and a thermal camera. (a) The target object. (b) Reflection image using a conventional camera. (c) Thermal image of the same object. When the object is carefully illuminated, shading of both images is the same, which implies conventional computer vision techniques can be applied to the thermal images.

trolled, because thermal imaging makes up a part of far infrared light imaging. Figure 6.1 shows an image captured by a color camera and a thermal image, where a ball is illuminated by a point light source. Both the color image and the thermal image exhibit the same shading. This observation implies that computer vision techniques can also be applied for thermal images.

In this chapter, we show that far infrared light transport can be regarded as a composition of multiple optical and thermal effects similar to the visible light transport. We define ambient, specular, diffuse, and global components in the thermal observation, and show the transient property of each component. Based on this model, a time-resolved decomposition of the far infrared light transport is proposed. Moreover, we show that the surface normal can be estimated based on the Lambertian photometric stereo, because the diffuse component, which follows the cosine law, of the far infrared light is separated. The proposed thermal photometric stereo can be applied for any objects that absorb light and convert it into heat, including black body, transparent, and translucent objects. It has a wide applicability compared with the photometric stereo using visible light.

The chief contributions of this study are threefold. First, we extend the visible light transport model to the far infrared light transport. We show that the

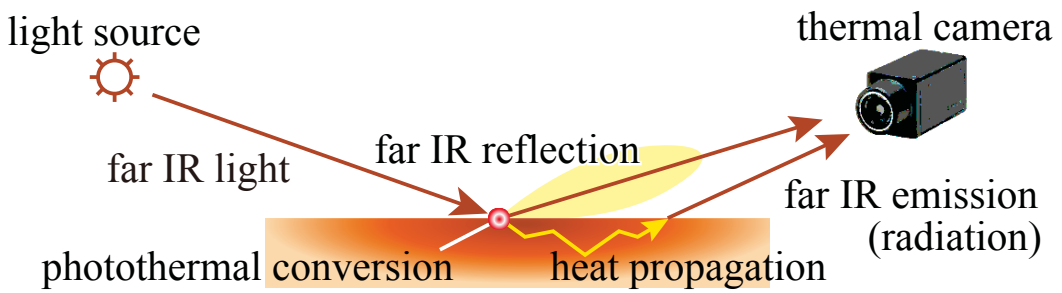


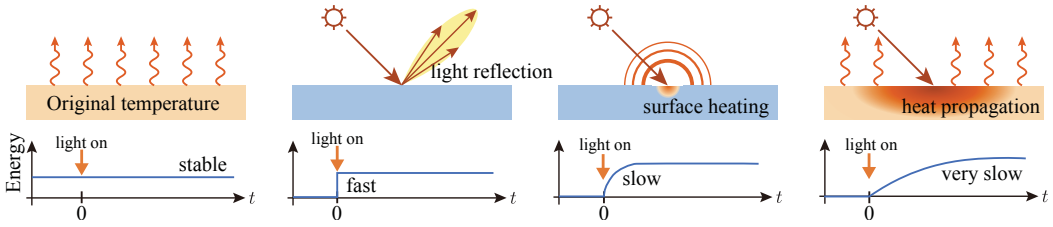
Figure 6.2: Far infrared light transport. While far infrared light can partially be reflected on the surface, the rest of the light is converted to heat energy, propagates inside the object, and is then converted to far infrared light corresponding to the temperature. The composition of all the components are captured by a camera. The observation system is closed in the far infrared light domain.

thermal image is a composition of ambient, specular, diffuse, and global components, which is similar to the visible light transport. Second, a novel approach for time-resolved light transport decomposition is provided based on the difference of the transient property of the far infrared light transport. Finally, we show that ordinary computer vision techniques can be straightforwardly applied to thermal images. As a proof of the concept, we propose a method to recover the surface normal using a photometric stereo after decomposing the far infrared light transport. The surface normal of challenging objects that have complicated optical effects can be recovered.

6.1. Far infrared light transport

We start with a brief review of thermal and far infrared light imaging. A typical thermal camera observes the temperature of the object by measuring the intensity of far infrared light because all objects emit far infrared light according to its temperature. When the object is a black body, the temperature and the intensity of far infrared light are governed by the Stefan-Boltzmann law [68], which represents a one-to-one correspondence between temperature and intensity:

$$E = \sigma T^4, \tag{6.1}$$



(a) Ambient (b) Specular reflection (c) Diffuse radiation (d) Global radiation

Figure 6.3: Far infrared light and heat transport components. Similar to the visible light transport, far infrared light transport consists of (a) ambient, which is the original temperature, (b) specular reflection as light, (c) diffuse radiation, and (d) global radiation caused by heat propagation. Because the speed of heat is slower than that of light, every components has distinctive transient properties hence they are separable.

where E is the intensity of the radiated far infrared light, σ is the Stefan-Boltzmann constant, and T is the thermodynamic temperature. We can handle the intensity of the far infrared light as the temperature, and vice versa.

We assume that the scene is illuminated by a stable parallel light source of far infrared light and the object is captured by a thermal camera as shown in Fig. 6.2. When the object is not a black body, a part of the far infrared light reflects on the surface, while the rest of the light is absorbed and converted to the heat energy, the temperature increases, and far infrared light is emitted corresponding to its temperature. The observation is the sum of these effects and we term this total energy transport as *far infrared light transport* because the observation system is closed in the far infrared light domain.

An image captured by a normal camera is the composition of multiple light transport effects, *e.g.*, specular and diffuse reflections, inter-reflection, and subsurface scattering. Similarly, the thermal image is a sum of the multiple far infrared light transports as shown in Fig. 6.3. A main difference between visible and far infrared light transport is that the temporal transient properties are significantly different among the light transport components. The transient state of visible light transport is not observable at a video frame rate because the speed of light

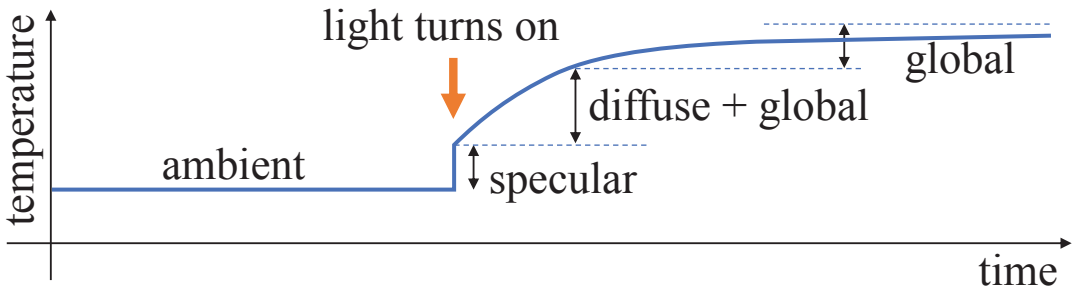


Figure 6.4: Transient properties of far infrared light transport. Because the temporal responses of the components are significantly different, they can be separated from the thermal video frames.

is extremely fast, while that of the far infrared light transport is easily observable because the heat conversion and propagation are relatively slow. Figure 6.4 illustrates a concept of the temperature transition of the far infrared light transport components. Before the light source is turned on, the observation consists of only the ambient component. The specular reflection appears immediately after the light source is turned on, and diffuse and global radiation slowly appear as the temperature increases. Then, the diffuse radiation reaches the steady state faster than the global radiation.

The observed thermal image $I(t)$ at a video frame t can be modeled as

$$I(t) = A(t) + S(t) + D(t) + G(t), \tag{6.2}$$

where A, S, D, G are the ambient, specular reflection, diffuse radiation, and global radiation components, respectively. We omit the camera pixel c because this observation is pixel-wise, and we assume the light source is turned on at $t = 0$ without the loss of generality. We review the detailed properties of the far infrared light transport components below.

Ambient The ambient component is the original temperature of the object. This corresponds to the intensity of far infrared light coming from the object when the heat source is turned off. The ambient component is assumed to be constant over time, and the effect of the heat source can be extracted by subtracting the initial state of the heat radiation. In the context of optical measurements, the ambient component corresponds to the ambient light or the dark current. The

ambient component $A(t)$ is expressed as

$$A(t) = \tau, \tag{6.3}$$

where τ is the original temperature of the object.

Specular reflection The specular component is an effect of far infrared light itself and not related to the heat propagation or temperature of the object. In the temperature measurement context, the specular component disturbs the observed temperature. Because it is the behavior of light, the specular component has the same properties as the visible light transport. The specular component $S(t)$ is only observed when the light source is turned on as

$$S(t) = \begin{cases} L_0 r_S & (t > 0) \\ 0 & (t \leq 0) \end{cases}, \tag{6.4}$$

where r_S is the specular reflectance, and L_0 is the intensity of the ideal light source. Because the speed of light is very fast, the specular component is stable and no transient state is observed on the video frame scale. Note that we ignore other optical effects such as diffuse reflection and subsurface scattering because they can be negligible for many materials owing to the long wavelength. They have however the same transient property, hence they can be safely regarded as a part of the specular reflection.

Diffuse and global radiation Diffuse radiation is defined as the surface heating. The energy is absorbed on the surface and the temperature of the surface is raised by photothermal conversion. Corresponding to the irradiance of light, the absorption energy of the incident far infrared light follows the cosine law [68]. The emission energy is linear to the absorbed energy, which is known as Kirchhoff's law of thermal radiation [68], which is given as

$$\alpha E = \varepsilon E, \tag{6.5}$$

where α and ε are the coefficients of absorption and emission, respectively. The raised temperature can be observed from any camera position; hence it corresponds to the Lambertian reflection of visible light transport. Because diffuse radiation is the effect on a single point, it is nearly stable but there is a small

transient state when the energy is absorbed, the temperature is raised, and far infrared light is emitted. Compared with the temporal spread of diffuse reflection [169], which is about tens of picoseconds, the temporal scale of diffuse radiation is much slower, and can be captured at the video frame rate.

When the object is heated over a sufficiently long time, the temperature is propagated in all directions, and we term this heat propagation as *global radiation*. Heat propagation is very slow, where it takes a few seconds to minutes, which is much slower than the diffuse radiation. Because global radiation spreads spatially, it corresponds to the subsurface scattering in the visible light observation.

Because the intensity of the far infrared light corresponding to the inner temperature follows the Fresnel law [110], the global radiation varies with respect to the viewing angle. However, it works as a scale factor with our setting, hence the effect of Fresnel refraction can be safely ignored.

It has been reported that the diffuse reflection and subsurface scattering can be regarded as the same physical phenomenon [74, 61, 159]; the light scatters on or beneath the surface and eventually bounces off of the material in random directions. Diffuse reflection represents the total intensity of light close to the incident point on the surface, and the subsurface scattering represents the light at a distance away from the incident point on the surface. The same thing can be said of the radiation; diffuse radiation is the heat energy whose heating point is local and global radiation is the heat that is propagated in all directions. Separating these components is a heuristic problem and we adopt exponential fitting to separate them.

We model the transient state of radiation using the exponential functions as

$$\begin{cases} D(t) &= R_\infty(1 - e^{-\sigma_d t})d_\infty \\ G(t) &= R_\infty(1 - e^{-\sigma_g t})g_\infty, \end{cases} \tag{6.6}$$

where σ_d and σ_g ($\sigma_g \ll \sigma_d$) represent the coefficient of the transient speed of diffuse and global radiations, respectively, and d_∞ and g_∞ represents the ratio of diffuse and the global radiation components at the steady state to the total radiation, respectively.

6.2. Thermal Photometric Stereo

Based on the difference in the transient properties of the far infrared light transport components, we develop a decomposition method and thermal photometric stereo.

6.2.1 Decomposition

Based on the observation of the transient properties, we decompose these components. First, the ambient component is observed before the light source is turned on. The light source is turned on at $t = 0$, and the specular component is the increased observation immediately after the light source is turned on. The radiation component is the transient state of increasing temperature and it is observed until the temperature becomes steady. Finally, the radiation is separated into diffuse and global radiations based on the speed to reach the steady state. We will next explain the details of the separation.

Separating ambient component The ambient component is the observation before the light source is turned on, and is determined as

$$A = I(0). \tag{6.7}$$

The transient observation $T_r(t)$ is the rest of the observation, given as

$$T_r(t) = I(t) - A. \tag{6.8}$$

Separating specular reflection and radiation The specular component is the reflection of light and has no transient state; hence it can be obtained as the increase immediately after the light source is turned on. The specular component S is obtained as

$$S = T_r(\epsilon), \tag{6.9}$$

where ϵ is an infinitesimal time duration.

The rest is the radiation, which has a temporal transient state. The radiation $R(t)$ can be obtained as

$$R(t) = T_r(t) - S. \tag{6.10}$$

Separating diffuse and global radiation

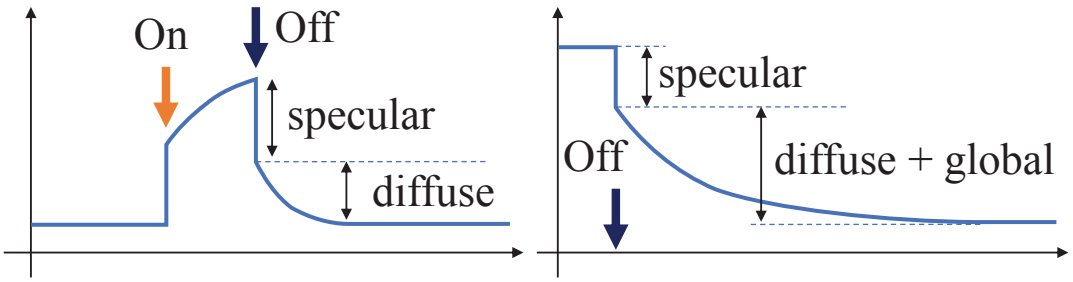
We fit the radiation components $R(t)$ to the model defined in Eq. 6.6 as

$$\begin{aligned}
\hat{\sigma_d}, \hat{d_\infty}, \hat{\sigma_g}, \hat{g_\infty} = & \underset{\sigma_d, d_\infty, \sigma_g, g_\infty}{\arg\min} \ \|R(t) - D(t) - G(t)\|_2^2 \\
\text{s.t.} \qquad & \min_t \frac{-\log\left(R_\infty - R(t)\right)}{t} \leq \sigma_g \ll \sigma_d \\
& 0 \leq d_\infty \leq 1 \\
& 0 \leq g_\infty \leq 1 \\
& d_\infty + g_\infty = 1, \qquad (6.11)
\end{aligned}$$

where $R_\infty = R(\infty)$ is the steady state of the radiation components. The first constraint represents that the time duration to the steady state of each component is smaller than the time for the observation to reach the steady state. Because the diffuse radiation is faster than the global radiation, σ_g is less than σ_d. The second and third constraints represent that the intensity of the diffuse and global radiations are smaller than the total radiation. The last constraint represents that the total radiation is a sum of diffuse and specular reflection, which reduces one degree of freedom. Fitting these parameters is not a convex problem so we use a grid search to find the global optimum. This does not involve a large computational cost because there are only three variables and the boundaries of the parameters can be predicted by the radiation profile $R(t)$.

Other options Another viable approach is to use the decrease in temperature after the light source is turned off. By switching on and off the light source over a short duration, the specular reflection and diffuse radiation can be directly obtained, as shown in Fig. 6.5(a), because the effect of heat propagation is negligible over a very short time. However, the diffuse radiation does not reach the steady state, hence it may suffer from extremely low SNR. To extend the heating time could improve the SNR, however, the global radiation cannot be ignored. To determine the suitable heating duration is another heuristic problem.

The cooling process is also useful to analyze far infrared light transport as shown in Fig. 6.5(b). Because heating and cooling are the reverse phenomena, light transport decomposition can be achieved in a very similar way. Because this takes twice as long time, we chose to analyze only the heating process.



(a) Heating in a short duration. (b) Transient state of cooling.

Figure 6.5: Other viable approaches. (a) By turning on and off the light source in a sufficiently short time, the specular reflection and diffuse radiation can be directly obtained. (b) Transient state after the light source is turned off contains similar information.

6.2.2 Surface normal estimation

When the object is heated by a narrow beam, the point absorbs the energy and radiates far infrared light according to the increased temperature. The absorbed energy follows the cosine law [68] as is observed for the light irradiance. Therefore, the diffuse component at the stable state can be represented as

$$D(\infty) = R_\infty d_\infty = R_\infty \rho \boldsymbol{i}^\top \boldsymbol{n}, \tag{6.12}$$

where ρ is the albedo of heat energy, and $\boldsymbol{i} \in \mathbb{R}^3$ and $\boldsymbol{n} \in \mathbb{R}^3$ represent the light direction and surface normal, respectively.

Because the diffuse radiation and diffuse reflection follow the same cosine law, the ordinary photometric stereo can be applied for diffuse radiation. The ordinary photometric stereo is not applicable for black body, transparent objects, and translucent objects that does not have diffuse reflection or are governed by other light transports. However, the diffuse radiation is a phenomenon of energy absorption and emission, so the surface normal of much more objects can be uniformly obtained using diffuse radiation. We propose a photometric stereo approach to the diffuse radiation, which we call the *thermal photometric stereo*.

As shown in Eq. 6.12, the decomposed diffuse radiation follows the cosine law hence it can be directly used for the Lambertian photometric stereo. The

estimated diffuse radiation component $\hat{d_\infty}$ can be simply represented as

$$\hat{d_\infty} = \rho \boldsymbol{i}^\top \boldsymbol{n}. \tag{6.13}$$

When multiple light sources are placed at different positions, multiple observations can be obtained that can be superposed in a matrix form as

$$\boldsymbol{d} = \rho \boldsymbol{I} \boldsymbol{n}, \tag{6.14}$$

where $\boldsymbol{d}$ and $\boldsymbol{I}$ are the superposed diffuse and light source direction matrices, respectively. When the light direction matrix is a full-rank matrix, the surface normal can be obtained as

$$\boldsymbol{n} = \frac{\boldsymbol{I}^\dagger \boldsymbol{d}}{\left\| \boldsymbol{I}^\dagger \boldsymbol{d} \right\|_2}, \tag{6.15}$$

where $\boldsymbol{I}^\dagger$ is a pseudo-inverse matrix of $\boldsymbol{I}$.

6.3. Experiments

The experimental setup is shown in Fig. 6.6. The target object is illuminated by far infrared spot lights (Exo Terra Heat-Glo 100W) and measured by a thermal camera (InfRec R500). The ambient component is observed before the light source is turned on. Then, the light source is turned on and the change of temperature is captured as a video.

The real light bulb is not stable immediately after turning on and requires a warm-up period in practice. In our experiments, the bulb is warmed up outside the experiment room and brought in under a cover. Removal of the cover is the actual meaning of the light being turned on. The wall of the room is heated over the experiment time and it could become a heat source. To avoid this effect, we place the object far from the wall and the room is actively cooled using an air-conditioner.

Decomposition result A black painted wooden sphere as shown in Fig. 6.7(a) is measured. A frame of the measured thermal video is shown in Fig. 6.7(b). Figure 6.7(c) shows the transition of the measured temperature at the black circular point shown in Fig. 6.7(b). The ambient component is the measured

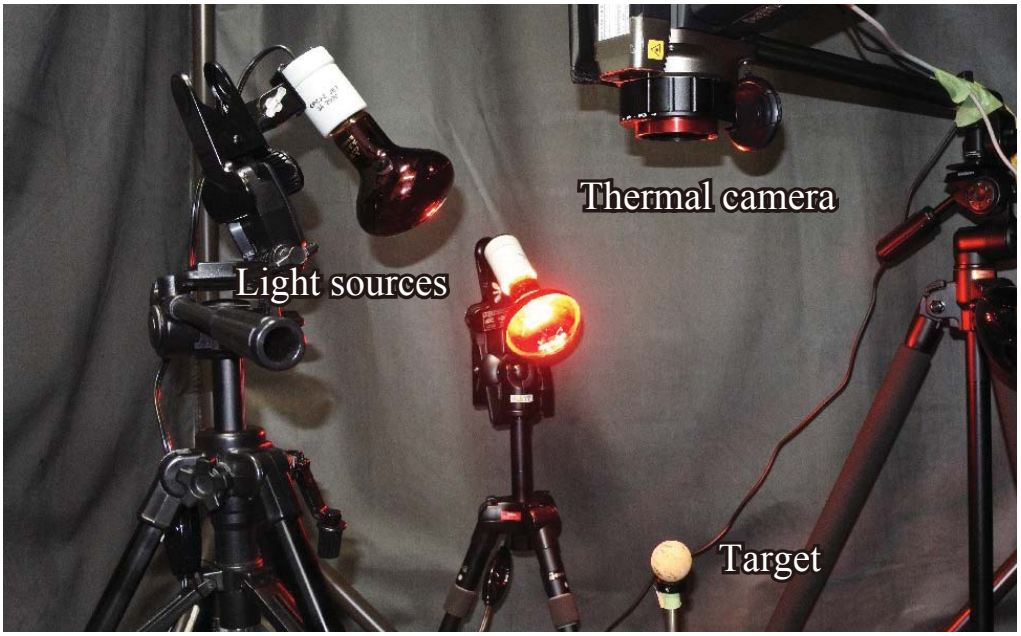


Figure 6.6: Experimental setup. The object is illuminated by far infrared light and captured by a thermal camera.

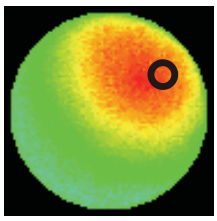
temperature before turning on the light source, and specular component is the increased intensity immediately after the light source is turned on. The radiation components are the rest, which is shown in Fig. 6.7(d). The radiation components are not fitted well by a single exponential curve because this is a sum of the diffuse and global radiations. Figure 6.7(e) shows the decomposed diffuse and global radiations. The sum of these fit well to the observation.

This procedure is applied for all the pixels, and the decomposed images are shown in Figs. 6.7(f) - (h). The specular component represents the reflection of the light source on the surface, the diffuse radiation represents the reasonable shading, and the global radiation represents the warming of the entire object.

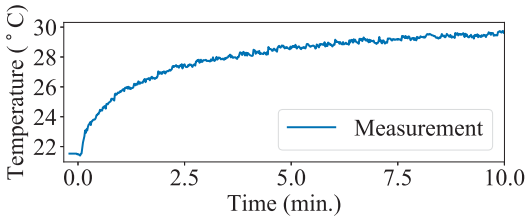
Surface normal estimation

By using multiple light source positions and separating each diffuse radiation, we can apply the Lambertian photometric stereo. Figure 6.8 shows the result of the thermal photometric stereo for the same object as shown in Fig. 6.7. A normal of the sphere is obtained as shown in Fig. 6.8(d). The result is compared

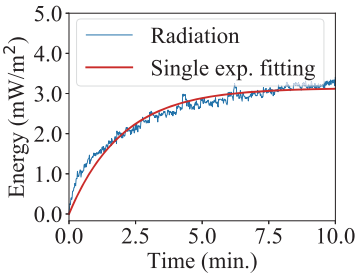
(a) The target.



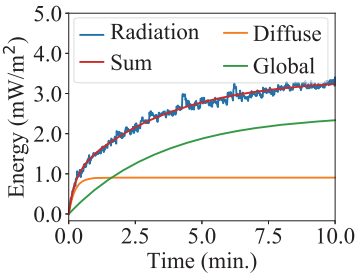
(b) Example frame.



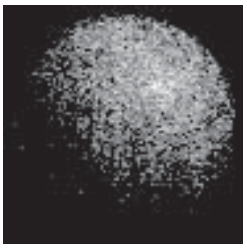
(c) Temporal profile of the temperature.



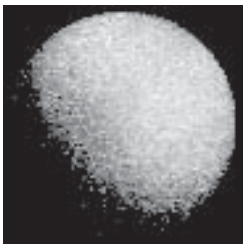
(d) Radiation profile.



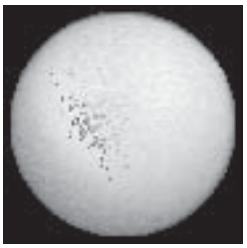
(e) Decomposed radiations.



(f) Specular comp.

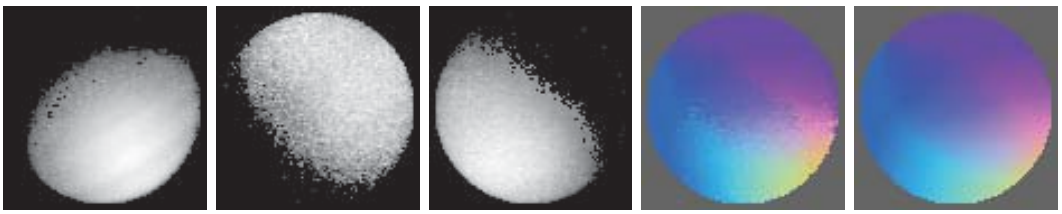


(g) Diffuse radiation.



(h) Global radiation.

Figure 6.7: Decomposition result for a black painted wooden ball. (a) The scene. (b) One of thermal video frames. Transient profiles of a point, indicated by the black circle, are shown. (c) Measured temperature transition. (d) Radiation profile. Ambient and specular reflection are subtracted from (c). (e) Decomposed diffuse and global radiations. (f-h) Decomposed images of specular reflection, diffuse, and global radiation, respectively.



(a) (b) (c) (d) (e)

Figure 6.8: Results of the thermal photometric stereo. (a - c) Decomposed diffuse radiation at different light positions. (d) Estimated surface normal. (e) The ground-truth normal.

with the result without light transport decomposition (composition of specular, diffuse, and global) and radiation (composition of diffuse and global) as shown in Fig. 6.9. As the temperature is not raised around $t = 0$, the compared results are noisy. The error increases owing to the global radiation at a longer time. As the best result, the angular errors of the result without decomposition and that of radiation is 7.71 and 6.50 degrees, respectively, while our method achieves a better result and the angular error is 5.85 degrees. This result shows the effectiveness of the separation of diffuse radiation.

We apply our method to other materials, including crystal glass, translucent plastic, and translucent marble. The decomposed diffuse component and estimated surface normal are shown in Fig. 6.10. Because our method is based on the diffuse radiation, materials that are difficult to measure with the ordinary vision techniques, *e.g.*, transparent and translucent objects, can be measured in the same way. A plastic ornament is also measured, and the result shows the feasibility of our method to a complex shaped objects.

Our method does is not suitable for some objects that does not absorb the far infrared light. Metallic materials are such objects and the thermal observation of a metallic ball is shown in Fig. 6.11. The metallic ball reflects all the incident light and behaves like a mirror. There is no transient state as shown in Fig. 6.11(c) which shows there are no radiation components. This problem is identical to the visible light observation.

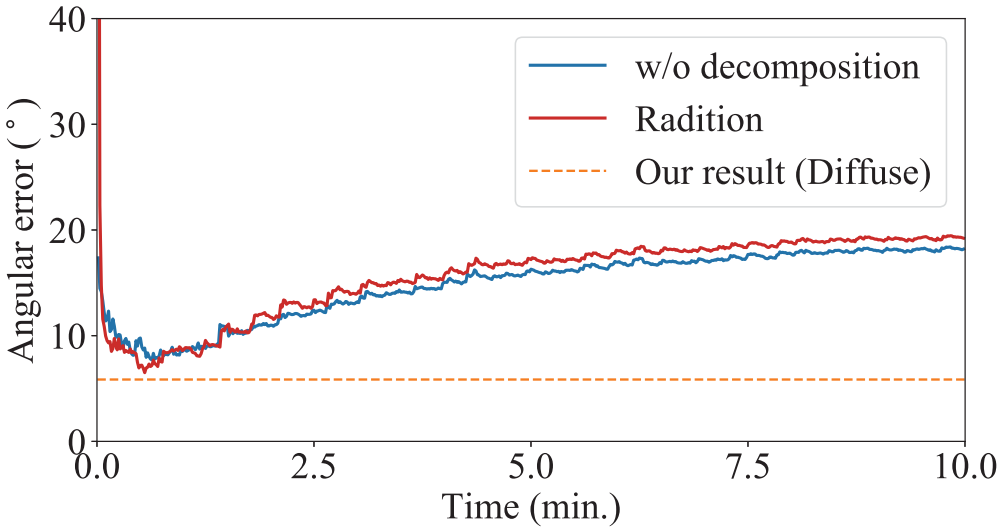


Figure 6.9: The effectiveness of decomposition. Photometric stereo result without decomposition, result using radiation components, and comparison with our method. Our method is time invariant and the accuracy is shown as a dotted line. The angular error of our method is 5.85 degrees, which shows that our decomposition is effective for the separation of diffuse radiation.

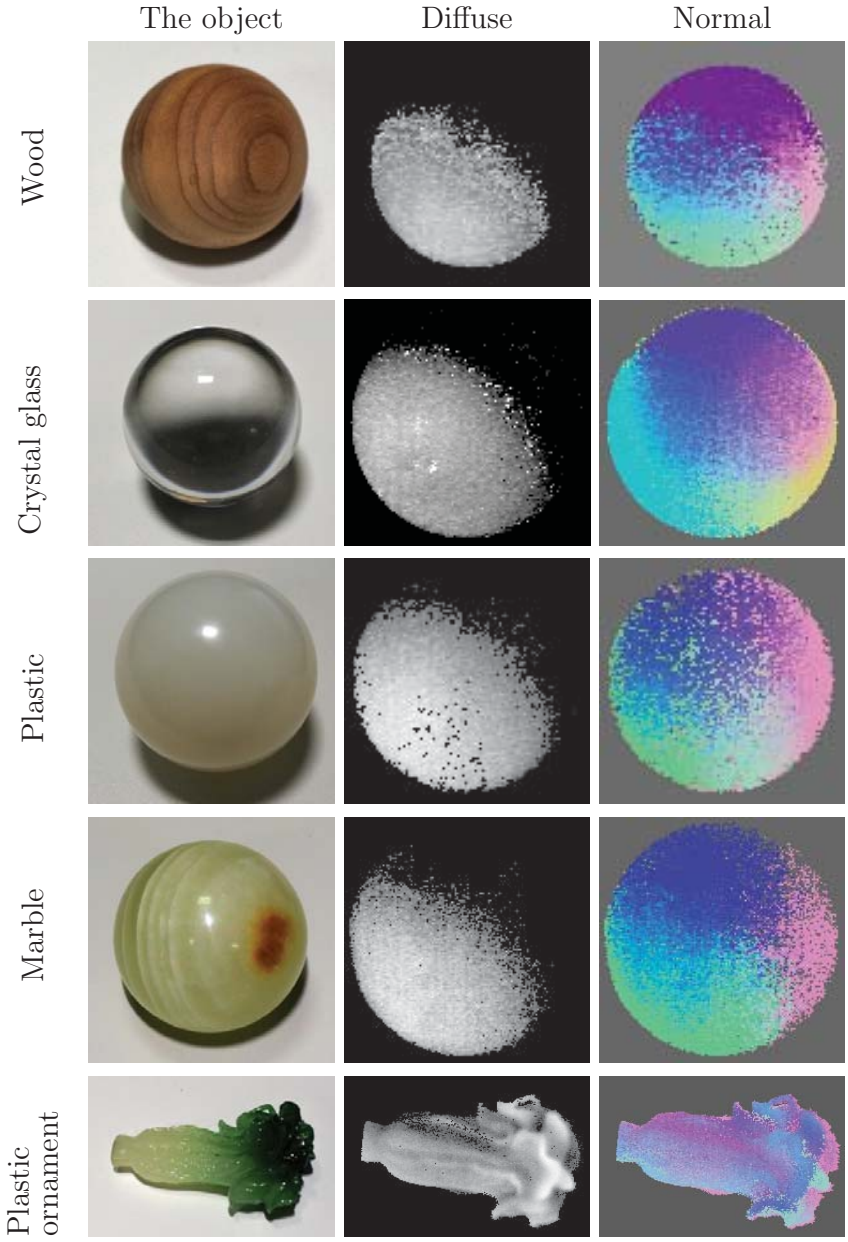
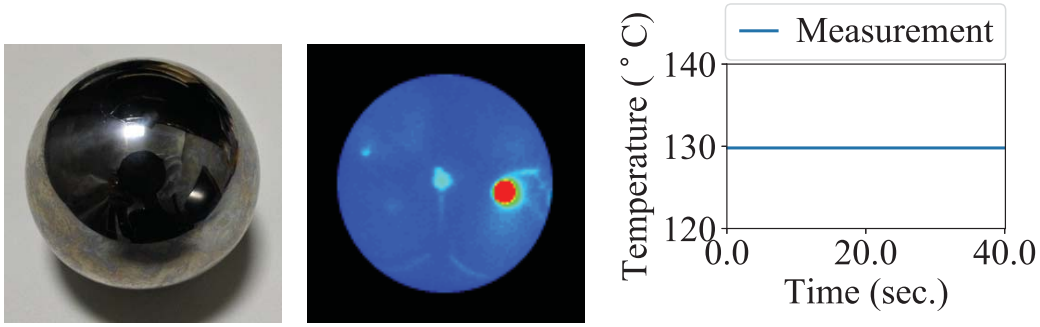


Figure 6.10: Results on various materials. Spheres made by wood, crystal grass, plastic, and marble are measured, which are challenging objects for ordinary computer vision techniques. Our method uniformly recovers the surface normal for many materials. A complex shape is also measured, and our method recovers the normal appropriately.



(a) Metallic ball. (b) Observation. (c) Temperature profile.

Figure 6.11: A failure case. The absorption rate is too small hence the metallic ball reflects all the incident light and shows the spherical thermal map of the room. The plot shows the temporal response of the highlighted point, where no transient components are observed. In such a case, only the ambient and specular components can be obtained and the photometric stereo does not work well.

6.4. Discussion

This chapter presents a novel technique for the time-resolved decomposition of far infrared light transport. We describe the far infrared light transport model, its transient properties, and that the ordinary vision techniques can be applied to decomposed thermal images. We propose a surface normal reconstruction using a photometric stereo after the diffuse component in a thermal image is separated. Our method recovers the surface normal of any objects that absorbs the incident light, including transparent, translucent, and black objects as well as matte objects.

While the effectiveness of our method is shown by some real-world experiments, some limitations are also encountered. First, the result is noisy owing to the low SNR observations and pixel-wise calculation. Because far infrared light cannot be measured by silicon sensors, the quality of the imaging sensor is not well developed. Naturally, this will be improved in the future, and it will directly improve our results. A global optimization that considers smoothness or simply using a smoothing filter are other options to improve the results.

Another limitation is that some materials, such as metals, do not exhibit much

diffuse radiation. In such a case, the ambient and specular components can be separated; however, the photometric stereo is not applicable. This problem is the same as that encountered with visible light observation, *e.g.*, photometric stereo suffers from mirror surface objects. In contrast, the absorption of many objects, including glass, is high, hence the potential applicability of our method is relatively higher than visible light observation techniques.

Chapter 7

One-shot hyperspectral imaging using faced reflectors

A hyperspectral image includes valuable information which is useful for various computer vision tasks such as material recognition [163], color consistency [3], and anomaly detection [151]. However, conventional hyperspectral imaging requires an expensive and professional system that is not easily available. Most commercial hyperspectral cameras capture a scene multiple times with a large number of narrow band filters or a tunable narrow band filter. Since a narrow band filter allows the spectral light at a certain wavelength to pass through, it is possible to capture an image at the wavelength by attaching the filter in front of the camera. Multiple capturing with different narrow band filters enables us to obtain a hyperspectral image. Another commercial system employs dispersing the light into individual wavelengths by using a prism or a diffraction grating. Since the camera captures only the dispersed light on a 2D slice in the scene at a time, sweeping in the whole scene is necessary to obtain a hyperspectral image.

In this chapter, we propose a novel technique of one-shot hyperspectral imaging using faced reflectors on which color filters are attached, as shown in Fig. 7.1. The light emitted from the light source is diffusely reflected on the target scene and then observed by the camera as an image. Moreover, the light after being reflected on the scene is also reflected on the reflectors before being observed by the camera. Thus, the spectrum of that observation is different from that without reflecting on the reflectors. That is, we can observe another image filtered by a

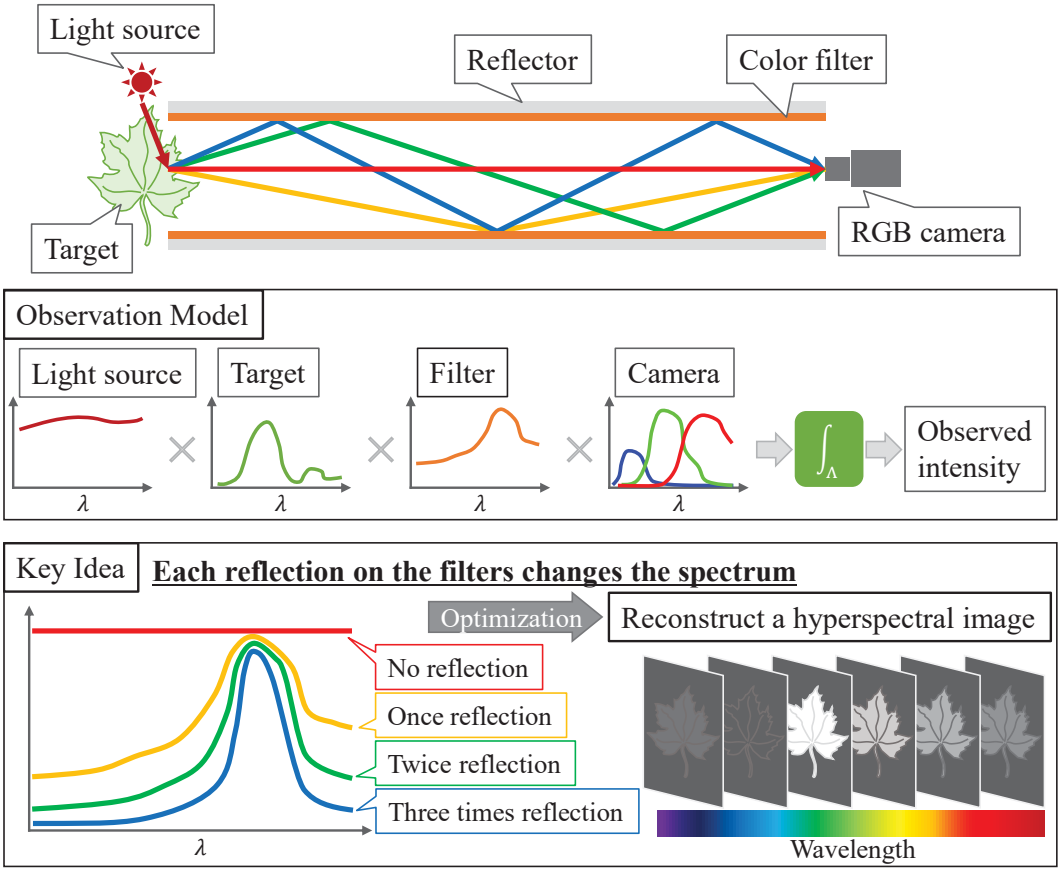


Figure 7.1: Overview of the proposed technique of one-shot hyperspectral imaging using faced reflectors. This simple setup employs a coupled mirror on which color filters are attached. Since the spectrum of the light varies at each reflection, we can obtain multiple observations from a single image. The observations allow us to reconstruct a hyperspectral image.

different spectrum. On the faced reflectors, where the light is reflected multiple times, the light at each reflection is regarded as filtered by a different spectrum. Therefore, we can obtain multiple observations filtered by different spectra at once.

The technique can be implemented by various faced reflectors setup, such as a coupled mirror and a kaleidoscope geometry, as will be explained in Section 7.2. Our experimental results show that even a coupled mirror setup with only one kind of a color filter is capable of capturing a hyperspectral image whose reconstruction error is on average 14% on real data, as shown in Section 7.3.

Contributions Our contributions are summarized as:

- *One-shot hyperspectral imaging:* Our technique is capable of capturing a hyperspectral image with one-shot by using a coupled mirror with only one kind of color filter, and its quantitative error is on average 14%.
- *Extremely low cost measurement system:* Implementing our technique only requires a pair of planar mirrors and a color filter, which are readily available for most users, and all of them totally cost less than $100.

7.1. One-shot hyperspectral imaging technique

7.1.1 Appearance model of faced reflectors

We first introduce the model to take an image by an RGB camera. We assume that isotropic spectral reflectance on the surface under a uniform illumination for a whole scene. An observed intensity y_k in the k-th channel of an image can be expressed as

$$y_k = \int_\Lambda l(\lambda)s(\lambda)c_k(\lambda)d\lambda, \tag{7.1}$$

where λ is the wavelength, $l(\lambda)$ is the spectrum of the illumination, $s(\lambda)$ is the spectral reflectance, $c_k(\lambda)$ is the spectral sensitivity of the k-th channel on the camera, and Λ is the range of wavelength, for example, from 400 to 700nm if the visible light is assumed. When a pair of reflectors is placed between the scene and the camera as shown in Fig. 7.1, the light reflected in the scene is specularly

reflected on the reflectors before being observed by the camera. Suppose that a color filter whose spectrum is $f(\lambda)$ is attached on both reflectors and the spectral reflectance of the reflector is flat, the observed intensity of the light once reflected on the reflector, see the yellow line in Fig. 7.1, can be expressed as

$$y_{k,1} = \int_{\Lambda} l(\lambda)s(\lambda)f(\lambda)c_k(\lambda)d\lambda. \tag{7.2}$$

When the light is twice reflected, shown as the green line in Fig. 7.1, it is multiplied by one more $f(\lambda)$. Regarding the light without reflection on the reflector, the red line in Fig. 7.1, as the case of $f^0(\lambda)$, the observed intensity $y_{k,i}$ to the i-bounce light reflected on the reflectors can be formulated as

$$y_{k,i} = \int_{\Lambda} l(\lambda)s(\lambda)f^i(\lambda)c_k(\lambda)d\lambda, \tag{7.3}$$

where $0 \leq i \leq N$ and N is the maximum number of bounces. N is an important factor to stably reconstruct the spectral reflectance and so we will discuss it in Section 7.1.3.

7.1.2 Problem formulation

The problem is here to estimate the spectral reflectance $s(\lambda)$ when all of the spectra of the light source $l(\lambda)$ and the color filter $f(\lambda)$ and the spectral sensitivities of the camera $c_k(\lambda)$ are known. To solve that, we transform Eq. 7.3 into a discrete formulation. The discrete formulation to Eq. 7.3 is

$$\begin{aligned} y_{k,i} &= \sum_{\lambda_b \leq \lambda \leq \lambda_e} l_\lambda s_\lambda f_\lambda^i c_{k,\lambda} \\ &= \sum_{\lambda_b \leq \lambda \leq \lambda_e} a_{k,i,\lambda} s_\lambda, \end{aligned} \tag{7.4}$$

where $a_{k,i,\lambda} \triangleq l_\lambda f_\lambda^i c_{k,\lambda}$, λ_b and λ_e are the minimum and maximum wavelengths in the range, respectively. Eq. 7.4 can be written in a matrix format as follows:

$$y_{k,i} = \boldsymbol{a}_{k,i}^{\mathrm{T}} \boldsymbol{s}, \tag{7.5}$$

where $\boldsymbol{a}_{k,i} = (a_{k,i,\lambda_b}, a_{k,i,\lambda_b+d\lambda}, \cdots, a_{k,i,\lambda_e})^{\mathrm{T}}$, $\boldsymbol{s} = (s_{\lambda_b}, s_{\lambda_b+d\lambda}, \cdots, s_{\lambda_e})^{\mathrm{T}}$, and $d\lambda$ is the granularity of wavelength in the discrete formulation. The resolution of

wavelength is defined as $N_\lambda = \frac{\lambda_e - \lambda_b}{d\lambda}$. When all intensities of 3 channels and N bounces are measured, a simultaneous equation can be composed as

$$\boldsymbol{y} = \mathbf{A}\boldsymbol{s}, \tag{7.6}$$

where

$$\boldsymbol{y} = (y_{\mathrm{r},0}, y_{\mathrm{g},0}, y_{\mathrm{b},0}, y_{\mathrm{r},1}, \cdots, y_{\mathrm{b},N})^{\mathrm{T}} \in \mathbb{R}^{3N}, \tag{7.7}$$

$$\mathbf{A} = [\boldsymbol{a}_{\mathrm{r},0}, \boldsymbol{a}_{\mathrm{g},0}, \boldsymbol{a}_{\mathrm{b},0}, \boldsymbol{a}_{\mathrm{r},1}, \cdots, \boldsymbol{a}_{\mathrm{b},N}]^{\mathrm{T}} \in \mathbb{R}^{3N \times N_\lambda}. \tag{7.8}$$

Here, when the intensity vector $\boldsymbol{y}$ is observable and the coefficient matrix $\mathbf{A}$ is known, then the spectral reflectance $\boldsymbol{s}$ can be reconstructed as follows:

$$\hat{\boldsymbol{s}} = \arg\min_{\boldsymbol{s}} \|\mathbf{A}\boldsymbol{s} - \boldsymbol{y}\|^2. \tag{7.9}$$

When the rank of the coefficient matrix is sufficient, Eq. 7.9 can easily be solved by a conventional least squares technique.

7.1.3 The nature of the coefficient matrix A

The reconstruction of the spectral reflectance can be solved by a conventional least squares technique as mentioned in Section 7.1.2. However, the stability of its computation depends on the coefficient matrix $\mathbf{A}$. To obtain a stable solution, the rank of the coefficient matrix $\mathbf{A}$ has to be sufficient.

Furthermore, it is well known that the condition number of a problem is an important factor to stably solve the problem in the field of numerical analysis. A problem with a low condition number can stably be solved. In our formulation, the condition number of the problem is decided by the coefficient matrix as follows:

$$\kappa(\mathbf{A}) = \frac{\sigma_{\max}}{\sigma_{\min}}, \tag{7.10}$$

where $\sigma_{\max}, \sigma_{\min}$ are the maximum and minimum of singular values of $\mathbf{A}$, respectively. We define $\sigma_{\max} \triangleq \sigma_1$ and $\sigma_{\min} \triangleq \sigma_{\mathrm{rank}(()\mathbf{A})}$.

Because we have assumed that the spectrum of the illumination l_λ, the spectral sensitivities of the camera $c_{k,\lambda}$, and the spectrum f_λ are all known, then the coefficient matrix $\mathbf{A}$ can be evaluated in advance. Therefore, when we have

multiple color filters, it is possible to select the optimal color filter which make the rank sufficient and the condition number the lowest. We verify this nature in Section 7.3.2 by an experiment on synthetic data.

7.1.4 Constrained optimization

Although, as mentioned in Section 7.1.2, Eq. 7.9 can be solved by a least squares method, it often gets unstable because of the nature of the coefficient matrix $\mathbf{A}$. To deal with that, we have explained how to construct the optimal setup by selecting the color filter to be used in Section 7.1.3. Moreover, we employ a convex optimization technique to make the computation more stable.

The spectral reflectance physically can neither be a negative value nor over 1.0. This fact can be used as a strong box constraint. We can adopt a smoothness constraint on the spectral reflectance since it is often measured in the real world. Thus, we can rewrite Eq. 7.9 as

$$\begin{aligned} \hat{\boldsymbol{s}} = \arg\min_{\boldsymbol{s}} \left\{ \|\mathbf{A}\boldsymbol{s} - \boldsymbol{y}\|^2 + \alpha \int_{\Lambda} \left(\frac{\partial^2 s(\lambda)}{\partial \lambda^2} \right)^2 d\lambda \right\}, \\ \text{s.t.} \quad 0 \leq s(\lambda) \leq 1 \quad (\lambda \in \Lambda), \end{aligned} \tag{7.11}$$

where α is the coefficient for the smoothness term. Then, Eq. 7.11 can be expressed in a matrix format as

$$\begin{aligned} \hat{\boldsymbol{s}} = \arg\min_{\boldsymbol{s}} \left\{ \|\mathbf{A}\boldsymbol{s} - \boldsymbol{y}\|^2 + \alpha \|\mathbf{D}\boldsymbol{s}\|^2 \right\}, \\ \text{s.t.} \quad 0 \leq s_\lambda \leq 1 \quad (\lambda_b \leq \lambda \leq \lambda_e), \end{aligned} \tag{7.12}$$

where $\mathbf{D}$ is the second-order difference matrix. The objective function in Eq. 7.12 can be expressed in a quadratic program format as

$$\begin{aligned} &\|\mathbf{A}\boldsymbol{s} - \boldsymbol{y}\|^2 + \alpha \|\mathbf{D}\boldsymbol{s}\|^2 \\ &= \boldsymbol{s}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{A} + \alpha \mathbf{D}^{\mathrm{T}}\mathbf{D})\boldsymbol{s} - 2\boldsymbol{y}^{\mathrm{T}}\mathbf{A}\boldsymbol{s} + \boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}. \end{aligned} \tag{7.13}$$

Since the third term $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}$ is constant, then Eq. 7.12 is equivalent to

$$\begin{aligned} \hat{\boldsymbol{s}} = \arg\min_{\boldsymbol{s}} \left\{ \boldsymbol{s}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{A} + \alpha \mathbf{D}^{\mathrm{T}}\mathbf{D})\boldsymbol{s} - 2\boldsymbol{y}^{\mathrm{T}}\mathbf{A}\boldsymbol{s} \right\}, \\ \text{s.t.} \quad 0 \leq s_\lambda \leq 1 \quad (\lambda_b \leq \lambda \leq \lambda_e). \end{aligned} \tag{7.14}$$

Figure 7.2: The setup of a coupled mirror.

We solve Eq. 7.14 by using the quadratic cone programming technique [19]. To implement that optimization program, we utilize the python optimization library `cvxopt` [1].

7.2. Various setups as implementation

In Section 7.1, we have explained our technique of one-shot hyperspectral imaging by using faced mirrors. In order to implement the technique, various setups are available. In this section, we introduce two types of feasible setups by using a coupled mirror and a kaleidoscope.

7.2.1 Coupled mirror geometry

The simplest setup uses a pair of planar reflectors, faced to each other, on which color filters are attached. This setup is generally called a coupled mirror. First, we assume that the same filters are attached on both reflectors. Figure 7.2 illustrates the setup of a coupled mirror whose length is z and intervening distance is d. The angle of view of the camera is defined as θ. In this setup, the number of bounces N is geometrically limited to

$$N \leq \left\lfloor \frac{z}{d} \tan \frac{\theta}{2} \right\rfloor . \tag{7.15}$$

Note that, in the actual setup, the number of bounces is also limited by the energy of the irradiance even if the right-hand side in Eq. 7.15 is a large number.

For instance, when an RGB camera whose angle of view is 120° and a coupled mirror whose length is 300mm and intervening distance is 10mm are used, $N \leq 51$. The resolution of wavelength N_λ has to be less than or equal to $3N$ because the number of channels is three and the coefficient matrix $\mathbf{A}$ has to be a vertically long or square matrix for well-posedness. Therefore, the granularity of wavelength is limited to $d\lambda \geq \frac{\lambda_e - \lambda_b}{3N}$; that is, for example, it is theoretically possible to perform hyperspectral imaging at each 1.96nm wavelength, if the range is the visible light.

Secondly, we assume that two different color filters whose spectra are $f(\lambda)$ and $g(\lambda)$ are attached on the reflectors, respectively. In this case, Eq. 7.3 can be rewritten as

$$y_{k,i,j} = \int_\Lambda l(\lambda)s(\lambda)f^i(\lambda)g^j(\lambda)c_k(\lambda)d\lambda, \tag{7.16}$$

where $0 \leq (i+j) \leq N$. Since, geometrically, the reflection alternately happens on each reflector, $|i-j| \leq 1$. Therefore, if $1 \leq i \leq \frac{N}{2} - 1$ then the patterns of j for each i are three; $j \in \{i-1, i, i+1\}$, and if $i \in \{0, \frac{N}{2}\}$ then the patterns are two. Totally, the number of combinations of (i, j) is $(\frac{3}{2}N - 2)$. In the case of using the same filter, the number of bounces represents the number of row vectors in the coefficient matrix. On the other hand, if the number of bounces is more than four, using two different filters makes the number of row vectors larger, and thus, the computation becomes more stable.

7.2.2 Kaleidoscope geometry

In computer vision, the kaleidoscope geometry has often been used to obtain many information by one-shot imaging for measuring the bi-directional texture function [59] and the shape [93] of an object. This approach can as well be useful for our technique of one-shot hyperspectral imaging. There are many kinds of kaleidoscope such as triangular pole, quadrangular pole, and hexagonal pole. Since our basic technique is based on a pair of reflectors, a setup consisting of multiple pairs of reflectors is a good match. Here, we introduce the quadrangular pole geometry as an implementation.

The quadrangular pole, shown in Fig. 7.3, consists of two pairs of reflectors. In principle, it is possible to use four different color filters but we only use two different filters whose spectra are $f(\lambda)$ and $g(\lambda)$. One kind of filter is attached

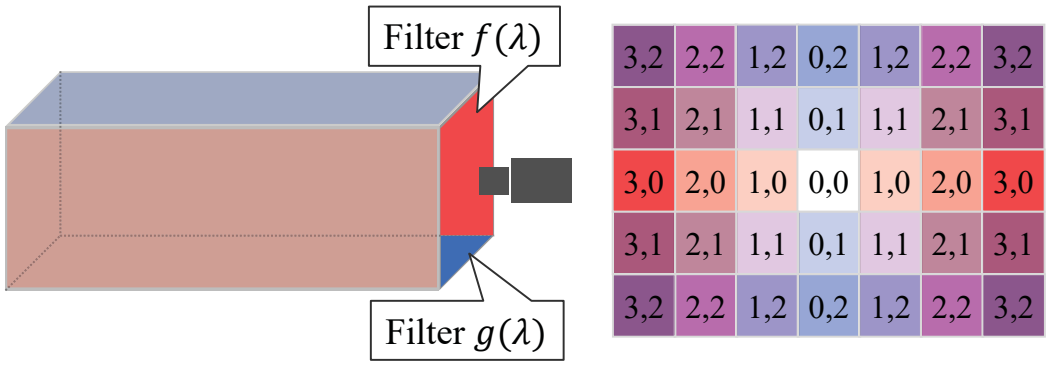


Figure 7.3: The setup of a kaleidoscope. Left: a quadrangular pole setup when two different color filters, whose spectra are $f(\lambda)$ and $g(\lambda)$, are used. Right: the combinations of (i, j) in Eq. 7.17 at a taken image.

on one of the pairs and another is attached on another of the pairs. Then, an observed intensity can be expressed as

$$y_{k,i,j} = \int_{\Lambda} l(\lambda)s(\lambda)f^{i}(\lambda)g^{j}(\lambda)c_{k}(\lambda)d\lambda, \tag{7.17}$$

where $0 \leq i \leq N$ and $0 \leq j \leq N$. The number of combinations of (i, j) is naturally N^2, if the camera has the same horizontal and vertical angles of view. Generally, the vertical angle is less than the horizontal one, so the number of combinations can be lower but still is larger than that of a coupled mirror geometry configuration.

7.3. Experiments

7.3.1 Spectral dataset

We first explain the spectral dataset to be used in our experiments. In our experiments, we define the range of wavelength as the visible light, from 400 to 700nm. To generate synthetic data of spectral reflectance, we use a set of 1995 spectral reflectance including 350 surfaces Krinov dataset [90], 120 Dupont paint chips, 170 natural objects [171], 24 Macbeth color checker patches, 1269 Munsell

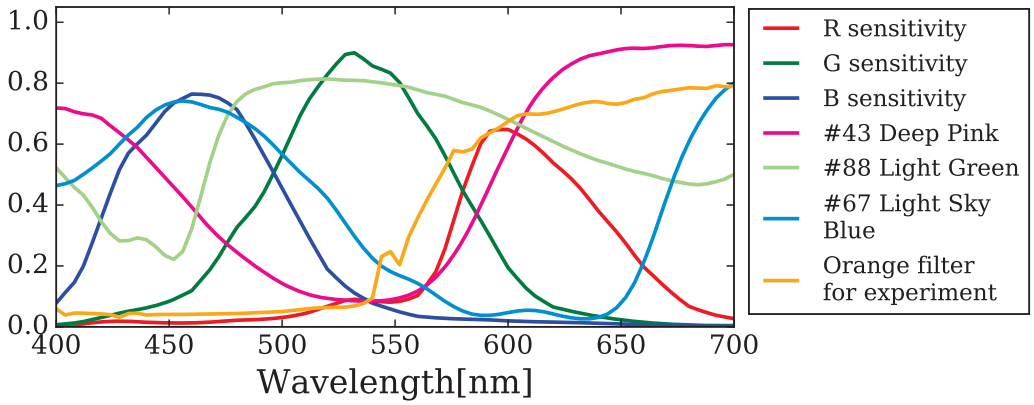


Figure 7.4: The spectral dataset.

chips, and 62 surfaces dataset [13]. We then measured the spectra of 253 Roscolux color filters and a thin clear orange filter for real data by using a spectrometer, OceanOptics Maya2000Pro. The spectral sensitivity of the camera, Nikon D5100, was estimated by the PCA-based method with data from [75, 81], as explained in [129]. Figure 7.4 shows the spectral sensitivity of the camera and the spectra of several of the color filters.

7.3.2 Synthetic data

We first perform experiments on synthetic data to validate our technique. Assuming a coupled mirror case with only one type of color filter, we simulated all observations for each bounce on each channel of the camera when using each of the 253 Roscolux filters and the thin clear orange filter. We repeated this simulation for observing each of the 1995 materials by Eq. 7.3. Then, we added zero-mean random Gaussian noise with a standard deviation of 0.1% onto the simulated intensities.

We reconstruct each spectral reflectance of the materials when using each of the filters by using Eq. 7.14. Then, we calculate their root-mean-square errors (RMSEs) to the ground truth for quantitative evaluation. The reconstructed results for the spectral reflectance of the Macbeth color checker are shown in

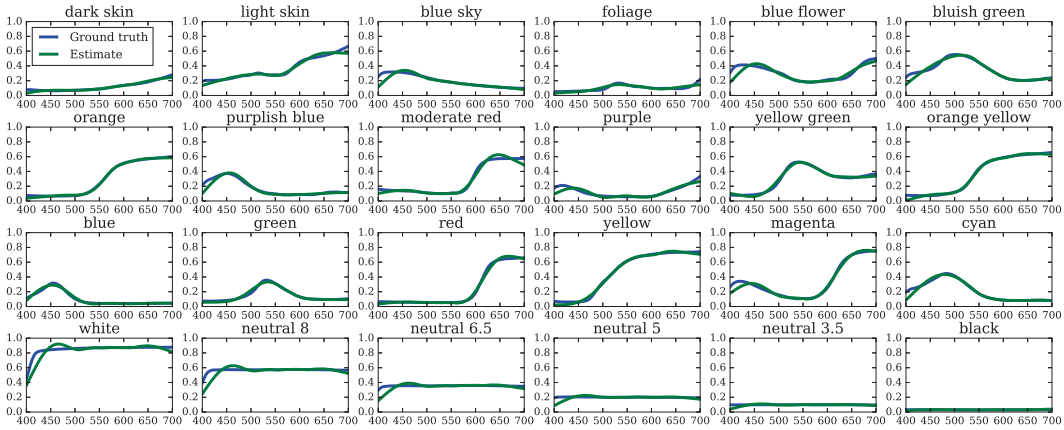


Figure 7.5: Experimental results on synthetic data for reconstructing the spectral reflectance of the Macbeth color chart in the spectral dataset when using the thin clear orange filter to be used in experiments on real data. The horizontal and vertical axes indicate wavelengths[nm] and reflectance, respectively.

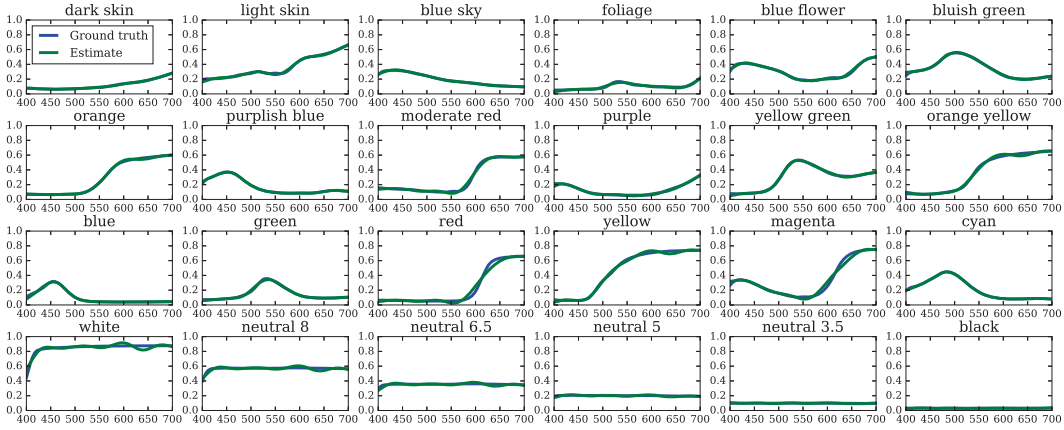


Figure 7.6: Experimental results on synthetic data for reconstructing the spectral reflectance of the Macbeth color chart in the spectral dataset when using the #67 Light Sky Blue filter in the Roscolux color filters. The horizontal and vertical axes indicate wavelengths[nm] and reflectance, respectively.

Figs. 7.5 and 7.6 as examples, in which we used the thin clear orange filter and the #67 Light Sky Blue filter. As can be seen, the reconstruction results are mostly similar to the ground truth but the error at shorter wavelength is relatively larger. The condition numbers when using the thin clear orange filter and the #67 Light Sky Blue filter are 1.29×10^5 and 0.26×10^5, respectively. The #67 Light Sky Blue filter has the smallest condition number of the Roscolux color filters. The averages of RMSEs for the Macbeth color checker by using the thin clear orange filter and the #67 Light Sky Blue filter were 0.03 and 0.01, respectively. Since the maximum value of reflectance can be 1.00, the average reconstruction errors is 3% when using the thin clear orange filter. The camera simulated in the experiments, shown in Figs. 7.5 and 7.6, can be regarded as an ideal float-value camera. However, a practical camera is an integer-value camera with 16 or 8 bits, usually. Thus, we evaluate RMSEs when using 16-bit and 8-bit cameras. The reconstructed results when using 16-bit and 8-bit cameras of the Macbeth color checker are shown in Figs. 7.7 and 7.8, respectively. The averages of RMSEs for the Macbeth color checker when using the 16-bit and 8-bit cameras were 0.02 and 0.05, respectively. It is shown that the performance of camera affects to the reconstruction accuracy. We will discuss this relationship between the condition number and the reconstruction error later. Then, we analyze the reconstruction errors from both perspectives of color filters and materials.

Figure 7.9 shows the average RMSE for reconstructing all materials when using each of the filters. After sorting in ascending order, we pick up 6 filters and then plot their spectra. This analysis explains what kind of filter is better in terms of smaller reconstruction error. As a result, a filter which allows a wide range of wavelengths and has higher gradients along wavelength is better as compared to a filter whose spectrum is very low or flat and has only one hump. Figure 7.10 shows the average RMSE for all filters when reconstructing each of the materials. We select 8 materials and then plot their spectral reflectance. This analysis shows what kind of material is difficult to be reconstructed. Results show that when the spectrum is smoother, the RMSE is smaller, but when if has a gap along wavelength, like the red line in Fig. 7.10, it is hard to be reconstructed. This is because of the smoothness constraint in Eq. 7.11.

Secondly, we analyze the relationship between condition numbers and recon-

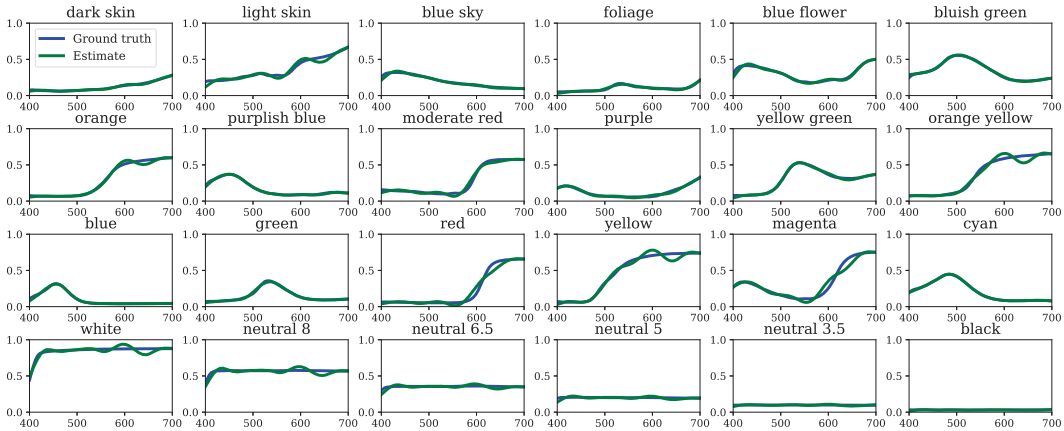


Figure 7.7: Experimental results on synthetic data for reconstructing the spectral reflectance of the Macbeth color chart in the spectral dataset when using a 16-bit camera.

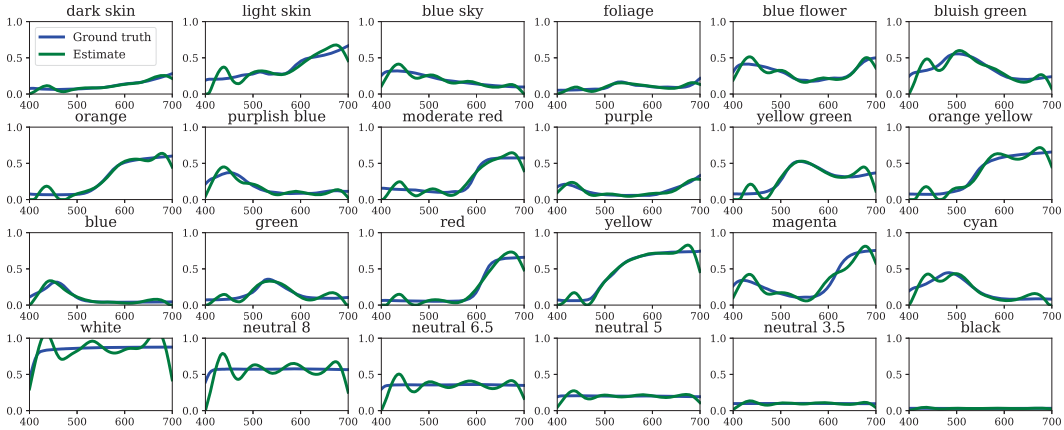


Figure 7.8: Experimental results on synthetic data for reconstructing the spectral reflectance of the Macbeth color chart in the spectral dataset when using an 8-bit camera.

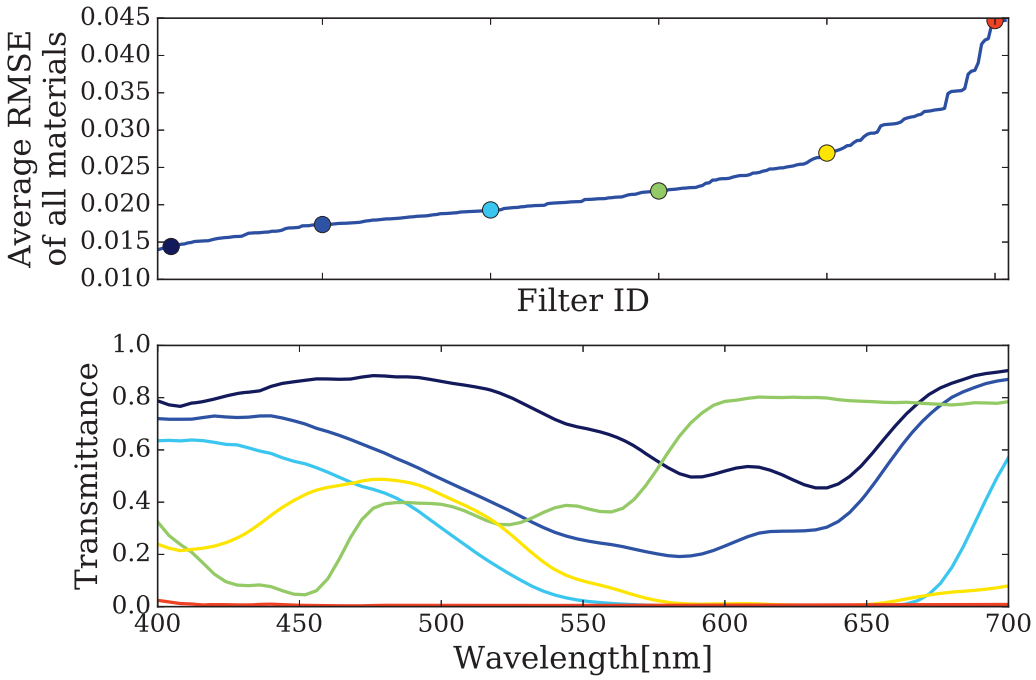


Figure 7.9: Analysis of the reconstruction errors for each filter. Top: the average RMSE for reconstructing all materials when using each of the 253 filters, sorted in ascending order. Bottom: the spectra of 6 filters chosen as color points on the top.

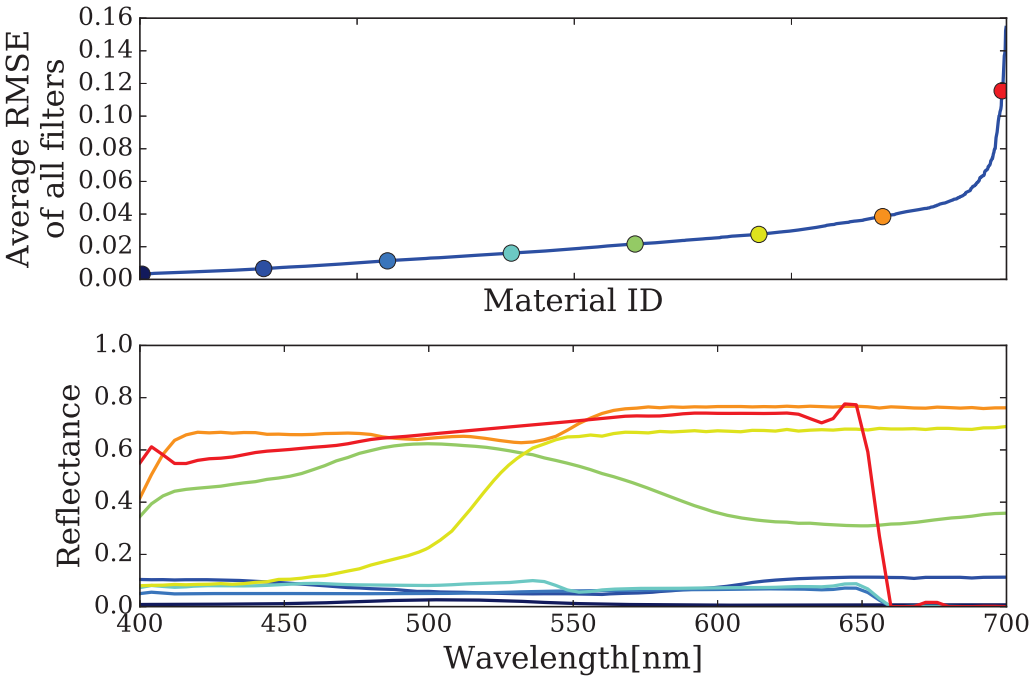


Figure 7.10: Analysis of the reconstruction errors for each material. Top: the average RMSE for all filters when reconstructing each of the 1995 materials, sorted in ascending order. Bottom: the spectral reflectance of 8 materials chosen as color points on the top.

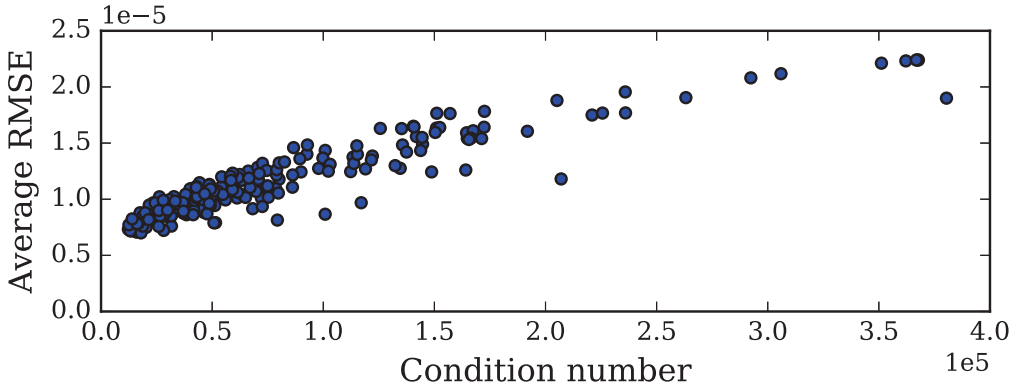


Figure 7.11: The relationship between the condition number of the coefficient matrix **A** and the reconstruction error. This shows the condition number is correlated to the reconstruction error.

struction errors to verify our assertion mentioned in Section 7.1.3. We compute the condition number of coefficient matrix **A** when using each of the filters, and then compare that with the average RMSE for the filter. Figure 7.11 shows the relationship between the condition numbers and the average RMSEs for all of the filters. It can be seen that the condition number is correlated with the average RMSE which supports our claim. This correlative relationship is helpful in selecting an optimal color filter.

7.3.3 Real data

After testing the proposed technique on synthetic data, we perform experiments on real data. We first implemented a simple setup which consists of two planar front surface reflectors and the thin clear orange filter, as shown in Fig. 7.12(a). All real data are taken by the camera Nikon D5100 under a white light source D65. Since the length of the reflectors is 100mm, the intervening distance is 20mm, and the angle of view is 120°, up to 8-th bounce reflections can be observed according to Eq. 7.15. Experimentally, it is only possible to observe up to 4-th bounce, as shown in Fig. 7.12(b), because of a physical interference between the camera and

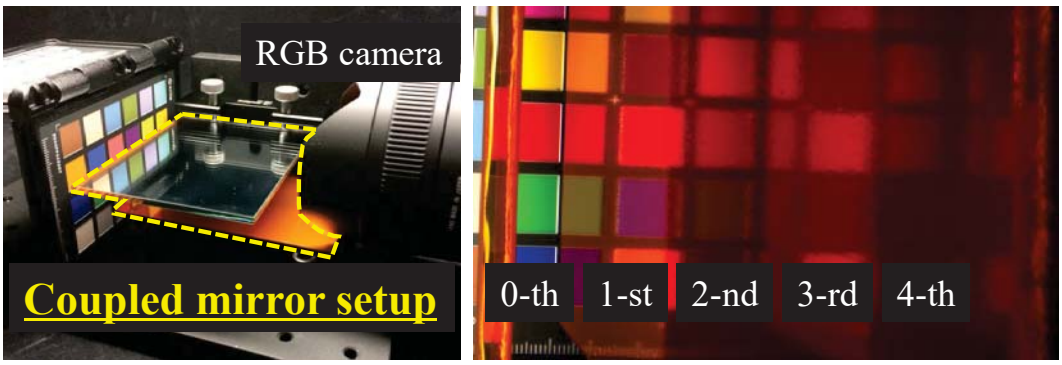


Figure 7.12: Real setup. (a) coupled mirror setup consisting of two surface reflectors and the thin clear orange filter on the both reflectors. (b) a captured image by the camera where up to 4-th bounce can be seen.

Figure 7.13: Experimental target: a green leaf. Left: a captured image by the camera with the real setup, where up to 2-nd bounce can be seen. Right: the area inside blue dash line is a target for reconstructing a hyperspectral image.

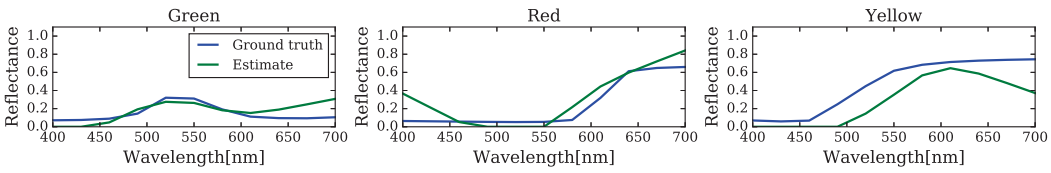


Figure 7.14: Experimental results on real data for reconstructing the spectral reflectance of green, red, and yellow chips on the Macbeth color chart.

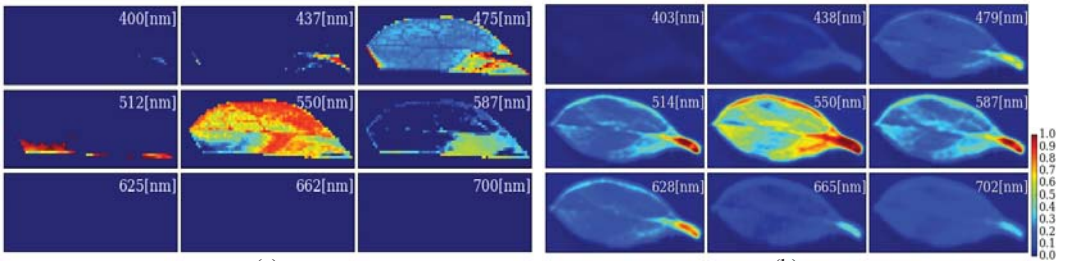


Figure 7.15: Experimental result on real data for reconstructing a hyperspectral image when capturing a green leaf. (a) the hyperspectral image reconstructed by our technique. (b) the ground truth hyperspectral image captured by SurfaceOptics SOC710. Reflectance is shown as pseudo-color. Note that the indicated wavelengths are different from each other because the resolutions of wavelength are different.

the setup.

To evaluate the accuracy of reconstructing spectral reflectance, we again use several chips on the Macbeth color chart, whose reflectance is known. As mentioned above, because only up to 4-th bounce can be observed, the number of total observations is 15, from 3 channels and 0 to 4-th bounces. This means that the granularity of wavelength is limited to 20nm. We reconstruct spectral reflectance in 30nm for stable computation. On each of yellow, red, and green chips, the pixel values are averaged to be used. Then, we reconstruct the spectral reflectance of those color chips. Figure 7.14 shows the spectral reflectance reconstructed by our technique and the ground truths. As a result, it can be seen that the reconstruction results roughly fit to the ground truths for all color chips. The RMSEs for green, red, and yellow chips are 0.09, 0.13, and 0.20, respectively.

We finally perform a practical experiment for reconstructing a hyperspectral image. A target object is a green leaf, as shown in Fig. 7.13. When using the simple real setup, only up to 2-nd bounce can be observed because of the size of the leaf, so the number of total observations is nine. Furthermore, we can observe only a region inside blue bash line in Fig. 7.13 because of physical interference. Since it is necessary to perform geometric registration along all observed images,

we put four markers in the scene and then compute homography transformation of the images. Figure 7.15 shows the hyperspectral image reconstructed by our technique and the ground truth captured by the hyperspectral camera SurfaceOptics SOC710. Note that the wavelengths indicated in Fig. 7.15 are not exactly same because the resolutions of wavelength are different between our method and the hyperspectral camera. Although only nine observations are available, our result has similar distributions to that of the ground truth, especially at 550 and 587nm. On the other hand, at shorter and longer wavelengths, the reconstructed images almost have no signal. This is because of the lack of observations and low signal-to-noise ratio.

7.4. Discussion

Our experimental results on synthetic data have shown that it is possible to reconstruct the accurate spectral reflectance by using a coupled mirror setup if the number of observations is large enough. A comparison between Figs. 7.5 and 7.6 shows that it is important to select the optimal color filter when the camera and the light source have been decided. In fact, it is difficult to observe many bounces in a real setup, as shown in the experiments on real data in Section 7.3.3. This is mainly because of two reasons: a) attenuation in light intensity and b) physical interference between the camera and the setup. Another problem comes from the color filter which is attached on the reflectors. Although we used a very thin filter, 2.5μm, the light was diffusely reflected a little and so the image after few bounces became unclear. This issue can be resolved by using thin color glass with a metallic sheet. While it is still difficult to obtain an accurate hyperspectral image as much as one by a commercial hyperspectral camera, the proposed technique can be used for capturing a very detailed spectral image around a certain wavelength, not whole wavelengths.

In this chapter, we have proposed a novel technique of one-shot hyperspectral imaging using faced reflectors, such as a coupled mirror and a kaleidoscope, on which color filters are attached. The key idea is based on that each of multiple reflections on the filters has a different spectrum, which allows us to observe multiple intensities through different spectra. We formulated an appearance model

for faced reflectors. Then, we analyzed the nature of the coefficient matrix, which is one of our contribution in this chapter, and showed that the condition number was correlated with the reconstruction error. Also, we analyzed the reconstruction errors to reveal what kind of filter should be used for a real implementation. In our experiments on real data, we showed the error of reconstructing spectral reflectance was on average 14% and performed hyperspectral imaging. The accuracy is still low but it supports our claim that one-shot hyperspectral imaging is feasible using faced reflectors with appropriate color filters.

7.4.1 Trade-off and limitations

There is a trade-off between spectral and spatial resolutions. The trade-off is derived from the granularity of wavelength, $d\lambda \geq \frac{\lambda_e - \lambda_b}{3N}$. To obtain high spectral resolution, the number of bounces N needs to be enough large. However, it leads to a fact that the size of each bounce region in a image, which is equal to the spatial resolution, reduces. A way to make N larger is to shorten the intervening distance but it leads to a narrow field of view since the distance is equal to the field of view of the proposed setup.

Currently, the proposed setup has two limitations; low SNR after a number of bounces on the reflectors and only flat-surface objects as the target. Because of the low SNR, the variations of observations caused by a small number of observable bounces are limited but it can be resolved by using the kaleidoscope geometry with different color filters. Another way is to capture an image by a high dynamic range technique. Targeting a object with complex geometry is a challenging work because shape reconstruction is required to find corresponding between different bounce images. Those will be our future work.

Chapter 8

Enhanced photometric stereo using multispectral images

Photometric stereo, a technique originally proposed by Woodham [178] and Silver [149], is a traditional method for estimating surface orientations from multiple images taken under different lighting conditions. A key factor in this process is the surface reflectance, which describes how the shading at each surface point changes in relation to the lighting direction and the surface normal. By varying the light direction, the corresponding changes in the shading are used to infer the surface normals according to the reflectance model.

Surface reflectance is normally assumed to follow Lambert's law, whereby the intensity I of the reflected light is proportional to the inner product of the lighting vector $\mathbf{l}$ and the surface normal vector $\mathbf{n}$:

$$I(x) = \rho(x)\mathbf{n}^{\mathrm{T}}\mathbf{l}, \tag{8.1}$$

where $\rho(x)$ denotes the albedo, or intrinsic color, of the surface at point x. This Lambertian model of reflectance is frequently used in photometric stereo because of its simplicity and convenience. However, the small deviations from the Lambertian model displayed by many materials produces inaccuracies in shape reconstruction.

Various methods have been proposed for dealing with non-Lambertian reflectance. Non-Lambertian highlights are treated as outliers within a Lambertian photometric stereo framework [16], or removed from images [105]. Some instances

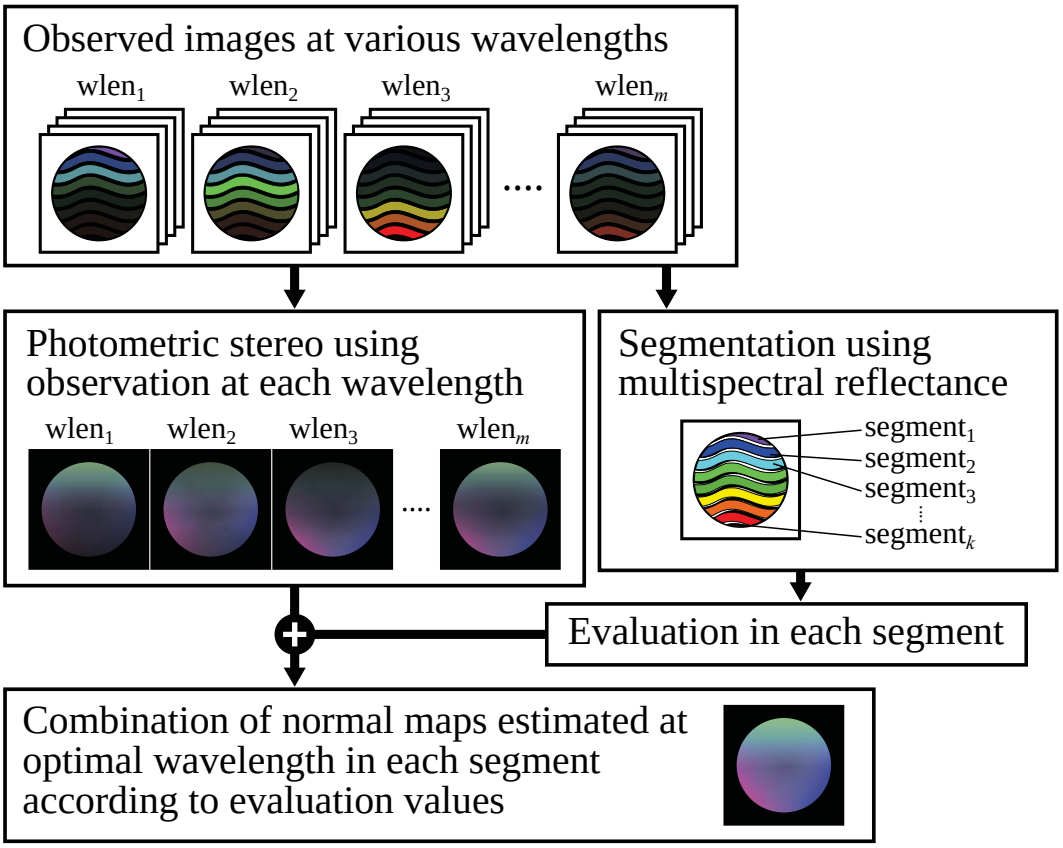


Figure 8.1: Overview of our algorithm for enhancing the performance of photometric stereo using multispectral images.

of photometric stereo are formulated for specific parametric reflectance models, such as the Torrance-Sparrow model [45], a combination of isotropic Ward models [52], and m-lobed reflectance maps [155]. Several work have been presented for an arbitrary form for the reflectance [7, 66]. Some methods analyze color for photometric stereo, but mainly to minimize image acquisition with differently colored illuminations [8], to improve robustness by incorporating additional measurements [23], or to aid in removing non-Lambertian reflectance components [105, 97].

Our algorithm is outlined in Figure 8.1. Our input is a set of images captured at various wavelengths and for different lighting directions. For each wavelength,

we apply Lambertian photometric stereo. In addition, the multispectral image constructed from the set of wavelengths is segmented into different multispectral regions. In each of these, we evaluate the photometric stereo results for each wavelength, using matrix rank analysis to identify which wavelength is most consistent with the Lambertian model. The surface normal map computed for each wavelength are then merged by selecting the optimal wavelength for each region. We thus obtain the final normal map. This approach is validated with simulated images. Experiments on real images support this multispectral approach, showing visible improvements in surface normal estimation over conventional photometric stereo based on brightness images.

8.1. Photometric Stereo using Multispectral Images

8.1.1 Wavelength-dependent Reflectance

We first examine the wavelength-dependence of reflectance and its effects on photometric stereo reconstruction. When accounting for a continuous range of wavelengths, the Lambertian reflectance model in Eq. (8.1) can be expressed as

$$I(\lambda, x) = \sigma(\lambda)\rho(\lambda, x)e(\lambda)\mathbf{n}(x)^{\mathrm{T}}\mathbf{l}, \tag{8.2}$$

where λ ranges over the measurable wavelengths of the sensor, σ is the spectral sensitivity of the camera, ρ is the spectral reflectance at a given surface point, and e is the spectrum of the light source.

The diffuse reflection that is conventionally used for photometric stereo arises from a physical process in which light penetrates an object surface, scatters off material particles, and is emitted out of the surface. According to the Lambertian model, this outgoing light is distributed uniformly in all directions, with an intensity given by Eq. (8.2).

In reality, the directional distribution of the outgoing light is determined by how the light is scattered within the material. This subsurface scattering is physically determined with respect to the optical properties of the material, virtually all of which are wavelength-dependent. For example, the wavelength can affect

the number of scattering events. For details on the wavelength-dependence of optical properties in materials, we refer the reader to [17].

This wavelength-dependence of optical properties suggests that surface reflectance is also spectrum-sensitive, thereby affecting the reconstruction quality of photometric stereo. We demonstrate this in Figure 8.2 with a simulated sphere rendered using real reflectance data for red plastic, obtained from the MERL BRDF dataset [107]. Forty images with different lighting directions were used for photometric stereo, and the color spectrum was coarsely sampled in red, green, and blue channels. The angular error maps in the reconstructions reveal different degrees of accuracy in the different color channels.

8.1.2 Region-based Identification of Optimal Wavelengths

Given a set of multispectral images captured with a stationary camera and different lighting directions, our algorithm first computes Lambertian photometric stereo, performs a region segmentation, and identifies the optimal wavelength for each region, which is then used to merge the normal maps computed with different wavelengths. The details of the process are as follows.

Region segmentation As mentioned previously, the optimal wavelength for Lambertian photometric stereo may vary across the image because of varying BRDFs. In particular, our investigations have shown that regions of different color often have different optimal wavelengths. To improve the reconstruction accuracy, we therefore segment the scene by normalized multispectral information, and determine the optimal wavelength for each of the regions. We apply k-means clustering for the segmentation, using an image where the object is fully illuminated (*i.e.*, no shadows).

Optimal wavelength identification in each region To find the optimal wavelength in each segmented region, we evaluate the data in each region. For a region with N pixels and M lighting conditions, we can express Eq. (8.2) in matrix form as

$$\mathbf{I} = g(\lambda)\mathbf{N}\mathbf{L}, \tag{8.3}$$

where $g(\lambda) = \sigma(\lambda)\rho(\lambda)e(\lambda)$, $\mathbf{I} \in \mathbb{R}^{N\times M}$ is a matrix of image intensities, $\mathbf{N} \in \mathbb{R}^{N\times 3}$

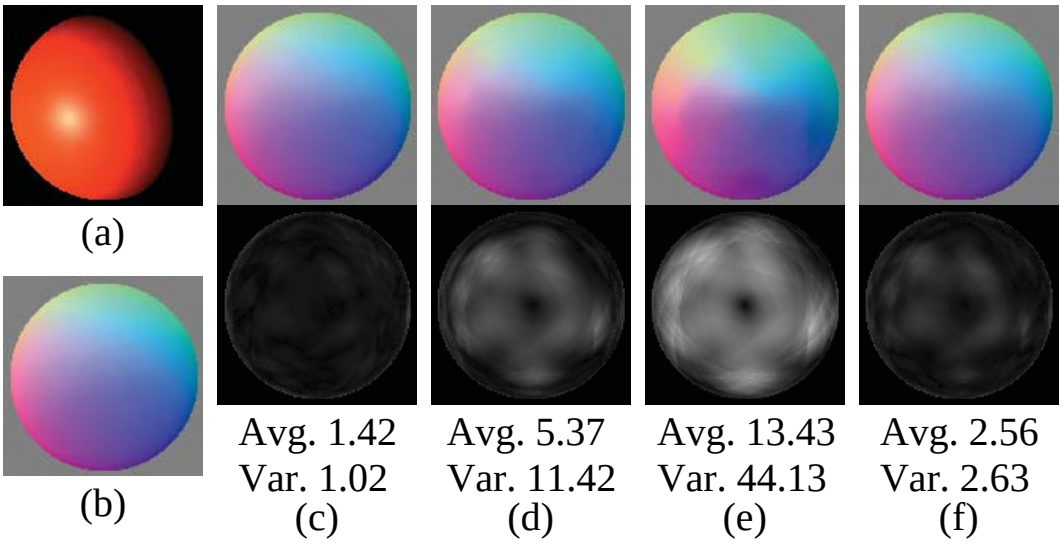


Figure 8.2: Effect of different wavelengths on photometric stereo. (a) One of 40 images rendered with different lighting directions; (b) ground truth of the normal map; normal maps and angular error maps using the (c) red, (d) green, and (e) blue color channels for Lambertian photometric stereo; and (f) results with brightness images.

is a matrix of surface normals, and $\mathbf{L} \in \mathbb{R}^{3\times M}$ is a matrix of lighting vectors.

As seen in Eq. (8.3), the Lambertian model predicts that the rank of the observation matrix $\mathbf{I}$ must equal that of the normal vector matrix $\mathbf{N}$, given that the sampled lighting directions are not co-planar (*i.e.*, rank(()$\mathbf{L}$) = 3). As rank(()$\mathbf{I}$) deviates further from rank(()$\mathbf{N}$), the data becomes more poorly fitted by the Lambertian model and then Lambertian photometric stereo should produce less accurate results as discussed in [180].

In view of this effect, we evaluate the wavelengths by determining how close the rank of its observation matrix is to rank(()$\mathbf{N}$). The evaluation function used in this work is

$$E = \frac{\sigma_{k+1}}{\sigma_k}, \tag{8.4}$$

where k = rank(()$\mathbf{N}$) and σ_i denotes the i-th eigenvalue of $\mathbf{I}$. This function becomes large when the rank of $\mathbf{I}$ falls below k (*i.e.*, σ_k close to 0) or exceeds k (*i.e.*, σ_{k+1} becomes large). Thus, the wavelength that minimizes E is that

for which the rank of $\mathbf{I}$ is closest to rank(()$\mathbf{N}$). A small additive term may be included in the denominator to avoid division by zero, but in practice we have found σ_k never to be zero.

Assuming the rank of $\mathbf{N}$ to be three, Eq. (8.4) is evaluated with $k = 3$. Corresponding values of k should be used in degenerate cases of lower rank. For each segmented region, the optimal wavelength minimizes E, and to enhance Lambertian photometric stereo, we estimate the final normal map by combining the normal maps of each region that correspond to its optimal wavelength.

8.2. Experimental Results

8.2.1 Simulation images

We first validate our method on a simulated scene for which the ground truth surface normals are known. Our scene consists of an egg-shaped surface rendered with the reflectances of seven different materials from the MERL BRDF dataset [107]. Forty images under different lighting directions are generated. One of these is shown in Figure 8.3(a). The ground truth is exhibited in Figure 8.3(b). We note that the MERL BRDF dataset provides reflectance data in terms of only red, green, and blue, so our multispectral technique is performed with three color channels in this instance.

Photometric stereo is first computed for each of the color channels separately, and also for the brightness images obtained by summing the three color channels. The estimated normal maps, angular error maps, and error statistics are shown in Figure 8.3(c-f). Overall, the brightness images yield better results than the images of individual color channels, probably because of a higher signal-to-noise ratio (SNR) resulting from summing multiple measurements. However, in some regions, using an individual color channel yields a better estimate than the brightness images (see the red rectangles in Figure 8.3 for example).

The result of combining normal map regions from different color channels is shown in Figure 8.3(g). Each region from the three color channel images is evaluated using Eq. (8.4), giving the values shown in Table 8.1. Pixels with sharp highlights detected by intensity thresholding were discarded from the evaluation

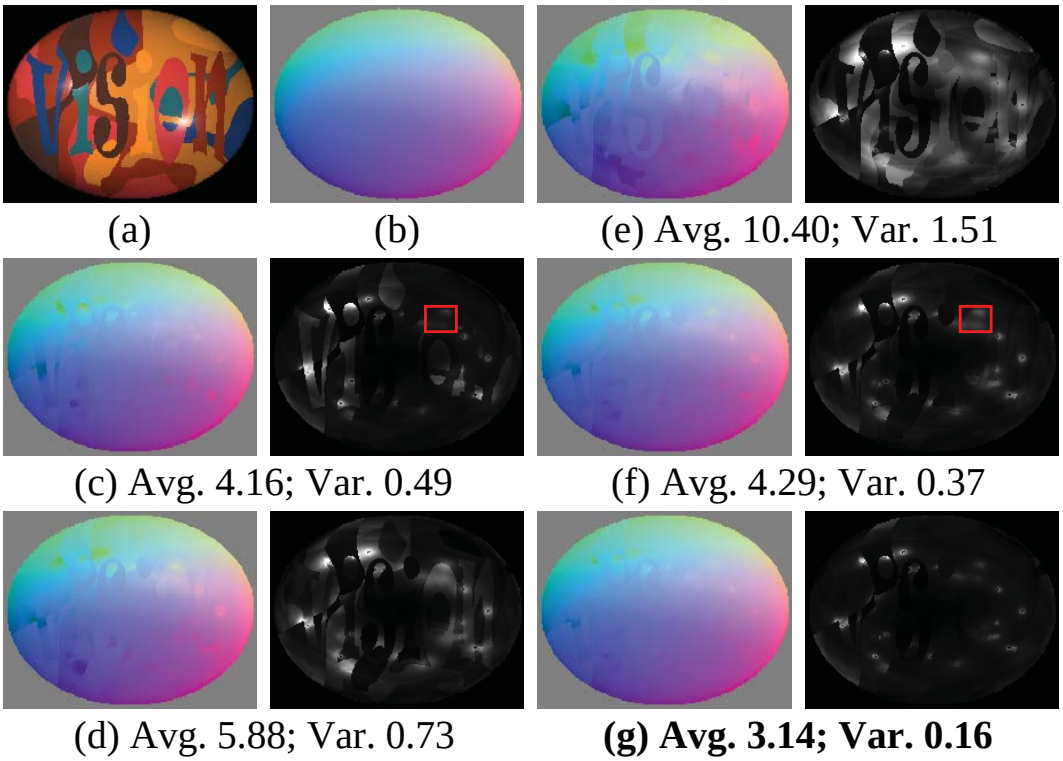


Figure 8.3: Normal map estimation for simulation images. (a) One of 40 images rendered under different light directions; (b) ground truth of the normal map; normal maps and angular error maps using the (c) red, (d) green, and (e) blue color channels; (f) with brightness images; (g) with our method.

by removing their rows from the observation matrix. With this multispectral technique, we achieve reconstructions that surpass those derived from brightness images, despite our technique not using summed measurements to improve the SNR.

8.2.2 Real images

We conducted an experiment with a real scene captured using a fixed camera (Point Grey Research Grasshopper), a halogen lamp (OSRAM), and nine narrow band filters at 450, 488, 580, 650, 694, 730, 780, 880 and 940 nm (Edmond Optics). A translation stage (Sigma) was used to switch the narrow band filters

Table 8.1: Evaluation function values.

Material name	Red	Green	Blue
blue-acrylic	0.960	0.985	**0.642**
green-plastic	0.840	**0.831**	0.885
light-brown-fabric	**0.687**	0.852	0.968
orange-paint	**0.767**	0.812	0.897
purple-paint	**0.348**	0.837	0.847
red-phenolic	**0.403**	0.932	0.900
yellow-matte-plastic	**0.434**	0.604	0.898

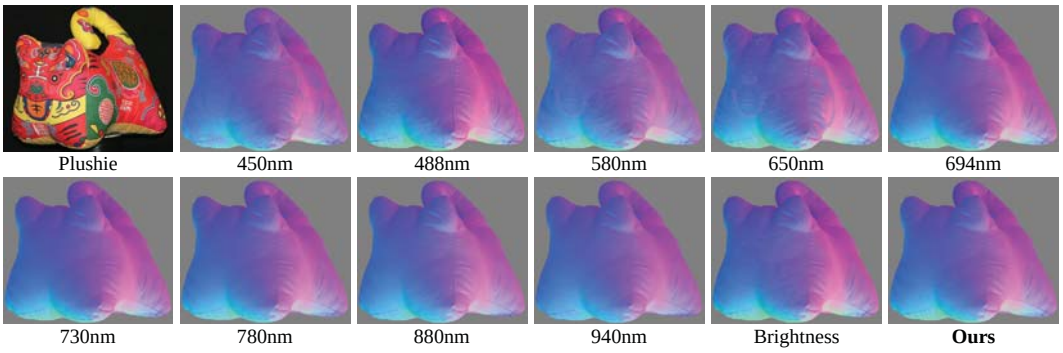


Figure 8.4: Picture and normal maps estimated using different wavelengths, brightness images, by applying our method.

automatically. Images were recorded under twelve different lighting directions. We applied photometric stereo to the images under perspective projection [160] to estimate the surface normals.

The target object in this experiment is a plushie shown in Figure 8.4, made of fabric of different colors.

Normal estimation using each wavelength The normal maps of the plushie estimated from different wavelengths are shown in Figure 8.4. Subtle but visible differences are apparent in the results obtained from the different wavelengths. Long wavelengths tend to yield smoother reconstruction results, possibly because

they penetrate deeper into the surface. By contrast, shorter wavelengths often yield sharper surface details. At certain wavelengths, such as 650 nm, surface normal discontinuities may appear, coinciding with sharp color gradients on the surface. These effects may indicate problems in fitting the Lambertian model to the data at these wavelengths.

From the surface normal reconstructions for the different wavelengths, we next determine which wavelengths are most consistent within the Lambertian model with our segmentation and evaluation scheme.

Segmentation and region evaluation Segmentation is performed using the nine color channels of our multispectral images. The segmentation result of the target object is shown in Figure 8.5(a). The plushie is divided into nine different region types. Disconnected areas on the image with the same region index are treated collectively as one region.

After segmentation, Eq. (8.4) is used to determine the optimal wavelength for each region. The values of the evaluation function for the different wavelengths are shown in Figure 8.5(c) for each region. The evaluation curves tend to be even over the color spectrum, possibly because the fabric material is less transparent, and hence more similar degrees of light penetration into the surface at different regions and for different wavelengths.

Combination of normal maps using evaluation function values The evaluation function minima in Figure 8.5(c) yield the optimal wavelength for each region, displayed in the label map in Figure 8.5(b). Long wavelengths in the visible range are often selected.

A reasonable supposition is that the optimal wavelength corresponds to the surface color of a region, since these wavelengths produce the brightest image and a high SNR. However, our results suggest otherwise. For example, the optimal wavelength is in the red spectral range even for a green surface, because of the subsurface scattering properties of this region.

The normal map computed by our method from the label map and the normal maps at different wavelengths is shown in Figure 8.4. Though the ground truth is unavailable for this object, our normal map appears qualitatively more correct than the normal map computed from brightness images. With our method, the

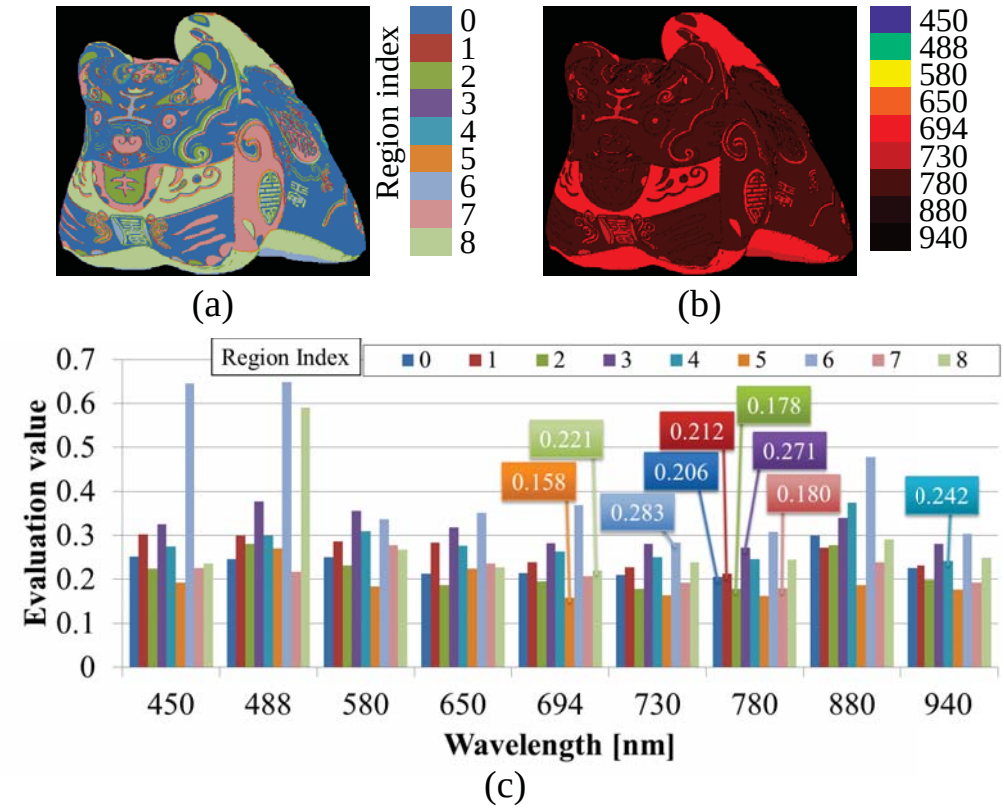


Figure 8.5: Our algorithm illustrated. (a) Nine regions segmented by k-means clustering in multispectral space. (b) Optimal wavelength for each region identified by Eq. (8.4). (c) Evaluation function values versus wavelength for each region.

normal map is less sensitive to texture.

8.3. Discussion

In this chapter, we proposed a new technique based on multispectral images to improve Lambertian photometric stereo. The resulting normal map is obtained by combining different normal maps computed at an optimal wavelength determined in each region of the image. Optimal wavelength determination does not require ground truth geometry, but is based instead on matrix rank analysis.

Our experiments demonstrated significant wavelength-dependence of various materials in photometric stereo reconstruction. Our method demonstrably shows improvements over traditional photometric stereo based on brightness images.

Some materials have reflectance properties that differ significantly from the Lambertian model for any wavelength. To broaden the applicability of our technique, future work will extend our approach to other parametric reflectance models used in photometric stereo. In addition, we intend to investigate the wavelength-dependent reflectance of various objects for the purpose of material recognition.

Chapter 9

Conclusion

Throughout this thesis, physics-based image decompositions were proposed toward obtaining a shape and visual textures of an object. As the shape, surface orientations are reconstructed by photometric stereo using the decompositions. As the visual textures, spectral reflectance, translucency, and inner structure of the object are obtained. In the spatial perspective, we proposed the image decomposition based on optical phenomena, *e.g.*, diffuse and specular reflection, single and multiple scattering, and subsurface scattering as PSFs. Also, the decomposition of inner layers was proposed to observe the inside of the object. In the temporal perspective, FIR light transport was decomposed based on optical and thermal phenomena, *e.g.*, reflection of light and diffuse and global radiation. In the spectral perspective, the one-shot hyperspectal imaging was proposed to obtain the spectral reflectance image. Finally, the achievement of this thesis are enabling to;

- reconstruct surface orientations of an object which consists of various materials, *e.g.*, multi-textured, translucent, transparent, and black objects, with a less effect of subsurface scattering;
- obtain a hyperspectral image at one-shot;
- obtain PSFs at each surface point as the translucency of a material; and
- observe inner layers of an object whose materials are relatively semi-transparent.

In Chapter 3, we proposed a general approach to uniformly combine any kind of separation methods for achieving a detailed decomposition. The experimental results showed that the proposed approach can decompose observations into four component images based on diffuse and specular reflection and single and multiple scattering. The decomposed result enables a novel image segmentation based on the opacity and translucency of materials.

In Chapter 4, we established a system to measure translucency of an object and physically reproduce it by UV printing. We defined a modulated transfer function as the translucency. The system obtains the modulated transfer function of a translucent object using a pro-cam system and prints multiple layers with different colors using a UV printer. The experiment result showed that the proposed system can reproduce the appearance of translucent materials, such as a slice of salmon and beef.

In Chapter 5, we proposed a technique to reconstruct reflectance maps of inner layers in an object using a pro-cam system. Frequency correlation imaging using spatio-sequentially modulated illumination and homodyne detection enables to extract only a signal from a specific depth. The experimental result showed that the proposed technique can reconstruct four different layers of an object whose depths can be manually assigned.

In Chapter 6, we discovered that radiation of far infrared light from an object when illuminating is also available for photometric stereo. A far infrared light transport including light and heat is temporally decomposed to extract the radiation component. Since almost all materials except metals radiate far infrared light, we achieved to reconstruct surface orientations of various materials, such as transparent, translucent, and black objects.

In Chapter 7, techniques of spectral imaging were proposed. First, we proposed a novel method to obtain a hyperspectral image at one-shot using faced reflectors. A capture system can easily be implemented by setting a coupled mirror on which a color filter is stuck in front of a conventional camera. The experimental result showed that the averaged RMSE of a spectrum reconstructed by the proposed method is 14%. Second, in Chapter 8, we utilized multispectral images to improve photometric stereo. Since the effect of subsurface scattering is different depending on a wavelength of light, the optimal wavelength for a mate-

rial is found out in photometric stereo. The experimental result showed that the proposed technique can improve the robustness of photometric stereo comparing with the traditional one using grayscale images.

9.1. Limitations and future work

While this thesis proposes radiometric decompositions for obtaining a shape, an inner structure, and visual textures of an object with various materials, there are still some limitations. The proposed photometric stereo technique using far infrared light is applicable for various materials except for metals. The emissivity of metals is so low that its radiation component cannot be extracted. Obtaining a shape of metals is important for industrial applications but also a challenging task. Thus, it will be our future work.

Currently, frequency correlation imaging for observing an inner structure cannot target a scattering media, *e.g.* human skin. For wide and practical applications, scattering in an object should be considered. Even if scattering occurs in the object, the spatial frequency of a sinusoidal projection pattern is not changed and thus it is possible to extend the technique.

The proposed system for measuring and reproducing translucency of an object cannot reproduce other visual textures, such as glossiness and micro-surface structure. However, those are important for making appearance of a fabricated object look realistic. Polarization imaging might be helpful to obtain those visual textures and can be future work.

Acknowledgements

Throughout the writing of this thesis I have received a great deal of support and assistance. First and foremost, I am forever indebted to my supervisor, Professor Yasuhiro Mukaigawa, for his guidance and support throughout my time at Osaka University and Nara Institute of Science and Technology (NAIST). His attitude to research has affected me deeply. He has always given me a lot of positive and constructive advice. I could not have asked for a better teacher or mentor.

I also deeply thank Professor Hirokazu Kato for being on my thesis committee. I would like to thank all of my advisors, Professor Takuya Funatomi, Professor Hiroyuki Kubo, and Professor Kenichiro Tanaka, throughout my time at NAIST. I am grateful to them for a lot of constructive discussions and contributions to our collaborative work. The work would have never been possible without their ingoing comments and help for writing.

I would like to thank my another supervisor, Professor Yasushi Yagi, for his support and advice throughout my time at Osaka University for my bachelor's and master's degrees. I appreciate his advice for my plan, point of view, attitude, and career. Professor Yasuyuki Matsushita has been my respectful advisor and mentor during my time at Osaka University and my internship at Microsoft Research Asia (MSRA). I have always respected his intelligence and conviction, which influenced me a lot.

The work in this thesis is also the achievements of several productive collaborations. Dr. Takahito Aoto has been one of my best colleagues in my professional and private lives. His cleaver ideas have always inspired me. Dr. Stephen Lin helped me to work on the spectral analysis during the internship at MSRA. Thanks to his experiential advice, we could archive successful results. I thank Mr. Nobuhiro Ikeya, Mr. Kazusa Tsubota, and Mr. Koki Fujita for their collab-

orative work with me during my time at NAIST. They told me how efficient and progressive collaborative work is.

I appreciate all supports from Ms. Toshie Nobori for accounting, office work, and management of my grants. I want to acknowledge all students and friends at OMI laboratory who enriched my doctoral life: Dr. Takafumi Iwaguchi, Mr. Takahiro Kushida, Mr. Kazuya Kitano, Mr. Arnaud Delmotte, Mr. Yuta Inoue, Mr. Tomoki Ueda, Mr. Yukiko Nakanishi, Mr. Takeshi Muraji, Mr. Kohei Yamamoto, Mr. Hayato Ikezawa, Mr. Daichi Ishii, Mr. Takehiro Ogawa, Mr. Masaki Kaga, Mr. Hiromu Kitajima, Mr. Tetsuro Mikami, Mr. Ryohei Morishima, Mr. Akira Yamada, Mr. Fu Xingjian, Ms. Camille Duquennoy and far too many others to name here. I would like to thank all students, postdocs, and professors at Yagi laboratory who have supported me: Professor Yasushi Makihara, Professor Ikuhisa Mitsugami, Professor Daigo Muramatsu, Dr. Seiichi Tagawa, Dr. Chika Inoshita, Dr. Kazuhiro Sakashita, Mr. Masahiro Suzuki, Mr. Yoshihiro Nagase, Mr. Ryo Kawai, Mr. Kohei Shiraga, Mr. Hozuma Nakajima and far too many others to name here.

Thanks to Dr. Ryuzo Okada, Dr. Tsuyoshi Tasaki and Ms. Mayu Okumura for their tutoring on industrial research at Toshiba Research & Development Center, I have learned a lot of skills to write official documents such as research proposal and report.

A part of this work is supported by JSPS KAKENHI Grant-in-Aid for JSPS Fellows, Grant Number JP17J05602, from the Japan Society for the Promotion of Science and ACT-I, Japan Science and Technology Agency.

I would like to thank all of my friends who shared our time for having a lot of fun: Xian, Peki, Dianne, Chi, Nikko, Paul, Felix, Rodo, Ayumi, Satoshi and far too many others to name here. Finally, I would like to thank my parents, Risa, Ryuki and Alyeska for their understanding and supporting my way throughout my life time.

References

[1] Cvxopt: Python software for convex optimization. `http://cvxopt.org/`. Accessed: 2019-07-29.

[2] Farhad Moghareh Abed, Seyed Hossein Amirshahi, and Mohammad Reza Moghareh Abed. Reconstruction of reflectance data using an interpolation technique. *Journal of the Optical Society of America A (JOSA A)*, 26(3):613–624, 2009.

[3] Andrea Abrardo, Luciano Alparone, I Cappellini, and Andrea Prosperi. Color constancy from multispectral images. In *Proc. of International Conference on Image Processing (ICIP)*, volume 3, pages 570–574. IEEE, 1999.

[4] Supreeth Achar and Srinivasa G Narasimhan. Multi focus structured light for recovering scene shape and global illumination. In *Proc. of European Conference on Computer Vision (ECCV)*, pages 205–219. Springer, 2014.

[5] Sameer Agarwal, Yasutaka Furukawa, Noah Snavely, Ian Simon, Brian Curless, Steven M. Seitz, and Richard Szeliski. Building rome in a day. *Communications of the ACM*, 54(10):105–112, October 2011.

[6] Yasushi Akashi and Takayuki Okatani. Separation of reflection components by sparse non-negative matrix factorization. *Computer Vision and Image Understanding (CVIU)*, 146:77–85, 2016.

[7] N. Alldrin and D. Kriegman. Photometric stereo with non-parametric and spatially-varying reflectance. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008.

[8] Robert Anderson, Bjorn Stenger, and Roberto Cipolla. Color photometric stereo for multicolored surfaces. In *Proc. of International Conference on Computer Vision (ICCV)*, 2011.

[9] Gonzalo R Arce, David J Brady, Lawrence Carin, Henry Arguello, and David S Kittle. Compressive coded aperture spectral imaging: An introduction. *IEEE Signal Processing Magazine*, 31(1):105–115, 2014.

[10] Yuta Asano, Yinqiang Zheng, Ko Nishino, and Imari Sato. Shape from water: Bispectral light absorption for depth recovery. In Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, editors, *Proc. of European Conference on Computer Vision (ECCV)*, pages 635–649, 2016.

[11] J. Bai, M. Chandraker, T.-T. Ng, and R. Ramamoorthi. A dual theory of inverse and forward light transport. In *Proc. of European Conference on Computer Vision (ECCV)*, pages 294–307, 2010.

[12] Gladimir V. G. Baranoski and Aravind Krishnaswamy. Light interaction with human skin: From believable images to predictable models. In *Proc. of SIGGRAPH Classes*, pages 10:1–10:80, New York, NY, USA, 2008. ACM.

[13] Kobus Barnard, Lindsay Martin, Brian Funt, and Adam Coath. A data set for colour research. *Color Research and Application*, 27(3):147–151, 2002.

[14] Harry G. Barrow and Jay M. Tenenbaum. Recovering intrinsic scene characteristics from images. In *Computer Vision System*, pages 3–26. 1978.

[15] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 25(10):1239–1252, 2003.

[16] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 25(10):1239–1252, 2003.

[17] Michael Bass, editor. *Handbook of Optics: Optical properties of materials, nonlinear optics, quantum optics*. Optical Society of America, 2009.

[18] Ayush Bhandari, Micha Feigin, Shahram Izadi, Christoph Rhemann, Mirko Schmidt, and Ramesh Raskar. Resolving Multipath Interference in Kinect: an Inverse Problem Approach. In *IEEE SENSORS*, pages 614–617. IEEE, 2014.

[19] Stephen Boyd and Lieven Vandenberghe, editors. *Convex Optimization*. Cambridge University Press, 2004.

[20] Alan Brunton, Can Ates Arikan, and Philipp Urban. Pushing the limits of 3d color printing: Error diffusion with translucent materials. *ACM Trans. on Graphics (ToG)*, 35(1):4:1–4:13, December 2015.

[21] Tongbo Chen, Hendrik PA Lensch, Christian Fuchs, and Hans-Peter Seidel. Polarization and phase-shifting for 3d scanning of translucent objects. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2007.

[22] Cui Chi, Hyunjin Yoo, and Moshe Ben-Ezra. Multi-spectral imaging by optimized wide band illumination. *International Journal of Computer Vision (IJCV)*, 86(2-3):140–151, 2010.

[23] Per H. Christensen and Linda G. Shapiro. Three-dimensional shape from color photometric stereo. *International Journal of Computer Vision (IJCV)*, 13(2):213–227, 1994.

[24] David J. Cuccia, Frederic Bevilacqua, Anthony J. Durkin, and Bruce J. Tromberg. Modulated imaging: quantitative analysis and tomography of turbid media in the spatial-frequency domain. *Optics Letters*, 30(11):1354–1356, Jun 2005.

[25] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. Acquiring the reflectance field of a human face. In *Proc. of ACM SIGGRAPH*, pages 145–156, 2000.

[26] Joris Dik, Koen Janssens, Geert Van Der Snickt, Luuk van der Loeff, Karen Rickers, and Marine Cotte. Visualization of a lost painting by vincent van gogh using synchrotron radiation based x-ray fluorescence elemental mapping. *Analytical Chemistry*, 80(16):6436–6442, 2008.

[27] Yue Dong, Jiaping Wang, Fabio Pellacini, Xin Tong, and Baining Guo. Fabricating spatially-varying subsurface scattering. *ACM Trans. on Graphics (ToG)*, 29(4):62:1–62:10, July 2010.

[28] Craig Donner and Henrik Wann Jensen. Rendering translucent materials using photon diffusion. In *Proc. of SIGGRAPH Classes*, pages 4:1–4:9, 2008.

[29] A A Dorrington, M J Cree, A D Payne, R M Conroy, and D A Carnegie. Achieving sub-millimetre precision with a solid-state full-field heterodyning range imaging camera. *Measurement Science and Technology*, 18(9):2809, 2007.

[30] A. A. Dorrington, J. P. Godbaz, M. J. Cree, A. D. Payne, and L. V. Streeter. Separating True Range Measurements from Multi-Path and Scattering Interference in Commercial Range Cameras. In *SPIE 7864, Three-Dimensional Imaging, Interaction, and Measurement*, page 786404, 2011.

[31] P. L. J. Drews, E. R. Nascimento, S. S. C. Botelho, and M. F. M. Campos. Underwater depth estimation and image restoration based on single images. *IEEE Computer Graphics and Applications*, 36(2):24–35, Mar 2016.

[32] Hao Du, Xin Tong, Xun Cao, and Stephen Lin. A prism-based system for multispectral video acquisition. In *Proc. of International Conference on Computer Vision (ICCV)*, pages 175–182. IEEE, 2009.

[33] Michael D'Zmura. Color constancy: surface color from changing illumination. *Journal of the Optical Society of America A (JOSA A)*, 9(3):490–493, 1992.

[34] Kenneth K. Ellis. Polarimetric bidirectional reflectance distribution function of glossy coatings. *Journal of the Optical Society of America A (JOSA A)*, 13(8):1758–1762, Aug 1996.

[35] S. Foix, G. Alenya, and C. Torras. Lock-in time-of-flight (tof) cameras: A survey. *IEEE Sensors Journal*, 11(9):1917–1926, Sep. 2011.

[36] Y. Francken, T. Cuypers, T. Mertens, J. Gielis, and P. Bekaert. High quality mesostructure acquisition using specularities. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–7, June 2008.

[37] Daniel Freedman, Eyal Krupka, Yoni Smolin, Ido Leichter, and Mirko Schmidt. SRA: Fast Removal of General Multipath for ToF Sensors. In *Proc. of European Conference on Computer Vision (ECCV)*, pages 234–249, 2014.

[38] Ying Fu, Antony Lam, Yasuyuki Matsushita, Imari Sato, and Yoichi Sato. Interreflection removal using fluorescence. In David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, *Proc. of European Conference on Computer Vision (ECCV)*, pages 203–217, 2014.

[39] Ying Fu, Antony Lam, Imari Sato, Takahiro Okabe, and Yoichi Sato. Separating reflective and fluorescent components using high frequency illumination in the spectral domain. In *Proc. of International Conference on Computer Vision (ICCV)*, December 2013.

[40] C. Fuchs, M. Heinz, M. Levoy, H. P. Seidel, and H. P. A. Lensch. Combining confocal imaging and descattering. *Computer Graphics Forum*, 27(4):1245–1253, 2008.

[41] Stefan Fuchs. Multipath Interference Compensation in Time-of-Flight Camera Images. In *Proc. of International Conference on Pattern Recognition (ICPR)*, pages 3583–3586. IEEE, 2010.

[42] Nahum Gat. Imaging spectroscopy using tunable filters: a review. In *Proc. SPIE Wavelet Applications VII*, pages 50–64, 2000.

[43] D Gavrilov, C Ibarra-Castanedo, E Maeva, O Grube, X Maldague, and R Maev. Infrared methods in noninvasive inspection of artwork. In *Proc. of 9th international conference on NDT of Art*, pages 25–30, 2008.

[44] A. Georghiades. Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo. In *Proc. of International Conference on Computer Vision (ICCV)*, pages 816–823, 2003.

[45] A. Georghiades. Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo. In *Proc. of International Conference on Computer Vision (ICCV)*, pages 816–823. IEEE, 2003.

[46] Abhijeet Ghosh, Tongbo Chen, Pieter Peers, Cyrus A Wilson, and Paul Debevec. Circularly polarized spherical illumination reflectometry. *ACM Trans. on Graphics (ToG)*, 29(6):162–172, 2010.

[47] Abhijeet Ghosh, Tim Hawkins, Pieter Peers, Sune Frederiksen, and Paul Debevec. Practical modeling and acquisition of layered facial reflectance. In *ACM Trans. on Graphics (ToG)*, volume 27, page 139, 2008.

[48] G. D. Gilbert and J. C. Pernicka. Improvement of underwater visibility by reduction of backscatter with a circular polarization technique. *Applied Optics*, 6(4):741–746, 1967.

[49] Ioannis Gkioulekas, Anat Levin, Frédo Durand, and Todd Zickler. Micron-Scale Light Transport Decomposition using Interferometry. *ACM Trans. on Graphics (ToG)*, 34(4):37:1–37:14, 2015.

[50] Ioannis Gkioulekas, Shuang Zhao, Kavita Bala, Todd Zickler, and Anat Levin. Inverse volume rendering with material dictionaries. *ACM Trans. on Graphics (ToG)*, 32(6):162:1–162:13, Nov 2013.

[51] John P. Godbaz, Michael J. Cree, and Adrian A. Dorrington. Closed-Form Inverses for the Mixed Pixel/Multipath Interference Problem in AMCW Lider. In *SPIE 8296, Computational Imaging X*, page 909778, 2012.

[52] Dan B. Goldman, Brian Curless, and Aaron Hertzmann Steven M. Seitz. Shape and spatially-varying brdfs from photometric stereo. In *Proc. of International Conference on Computer Vision (ICCV)*, volume 1, pages 341–348. IEEE, 2005.

[53] J. E. Greivenkamp. Generalized data reduction for heterodyne interferometry. *Optical Engineering*, 23:23 – 23 – 3, 1984.

[54] R. A. J. Groenhuis, H. A. Ferwerda, and J. J. Ten Bosch. Scattering and absorption of turbid materials determined from reflection measurements. 1: Theory. *Applied Optics*, 22(16):2456–2462, Aug 1983.

[55] M. Gupta, S. G. Narasimhan, and Y. Y. Schechner. On controlling light transport in poor visibility environments. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2008.

[56] Mohit Gupta, Shree K. Nayar, Matthias B. Hullin, and Jaime Martin. Phasor Imaging: a Generalization of Correlation-Based Time-of-Flight Imaging. *ACM Trans. on Graphics (ToG)*, 34(5):156:1–156:18, 2015.

[57] Mohit Gupta, Yuandong Tian, Srinivasa G Narasimhan, and Li Zhang. A combined theory of defocused illumination and global light transport. *International Journal of Computer Vision (IJCV)*, 98(2):146–167, 2012.

[58] Ralf Habel, Michael Kudenov, and Michael Wimmer. Practical spectral photography. In *Computer Graphics Forum*, volume 31, pages 449–458. Wiley Online Library, 2012.

[59] Jefferson Y Han and Ken Perlin. Measuring bidirectional texture reflectance with a kaleidoscope. *ACM Trans. on Graphics (ToG)*, 22(3):741–748, 2003.

[60] Shuai Han, Imari Sato, Takahiro Okabe, and Yoichi Sato. Fast spectral reflectance recovery using dlp projector. *International Journal of Computer Vision (IJCV)*, 110(2):172–184, 2014.

[61] Pat Hanrahan and Wolfgang Krueger. Reflection from Layered Surfaces Due to Subsurface Scattering. In *Proc. of ACM SIGGRAPH*, pages 165–174. ACM Press, 1993.

[62] Miloš Hašan, Martin Fuchs, Wojciech Matusik, Hanspeter Pfister, and Szymon Rusinkiewicz. Physical reproduction of materials with specified subsurface scattering. *ACM Trans. on Graphics (ToG)*, 29(4):61:1–61:1, August 2010.

[63] Felix Heide, Matthias B. Hullin, James Gregson, and Wolfgang Heidrich. Low-Budget Transient Imaging using Photonic Mixer Devices. *ACM Trans. on Graphics (ToG)*, 32(4):45:1–45:10, 2013.

[64] Felix Heide, Lei Xiao, Andreas Kolb, Matthias B Hullin, and Wolfgang Heidrich. Imaging in Scattering Media using Correlation Image Sensors and Sparse Convolutional Coding. *Optics express*, 22(21):26338–26350, 2014.

[65] Felix Heide, Lei Xiao, Andreas Kolb, Matthias B Hullin, and Wolfgang Heidrich. Imaging in scattering media using correlation image sensors and sparse convolutional coding. *Optics Express*, 22(21):26338–26350, 2014.

[66] T. Higo, Y. Matsushita, and K. Ikeuchi. Consensus photometric stereo. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1157–1164, 2010.

[67] Michele Hinnrichs and Mark A. Massie. New approach to imaging spectroscopy using diffractive optics. In *Optical Science, Engineering and Instrumentation '97*, volume 3118, pages 194–205, 1997.

[68] John R. Howell, M. Pinar Mengüc, and Robert Siegel. *Thermal Radiation Heat Transfer, Sixth Edition*. CRC Press, 2015.

[69] BB Hu and MC Nuss. Imaging with terahertz waves. *Optics Letters*, 20(16):1716–1718, 1995.

[70] David Huang, Eric A Swanson, Charles P Lin, Joel S Schuman, William G Stinson, Warren Chang, Michael R Hee, Thomas Flotte, Kenton Gregory, Carmen A Puliafito, et al. Optical coherence tomography. *Science (New York, NY)*, 254(5035):1178, 1991.

[71] Hendrik Christoffel Hulst and Hendrik C. van de Hulst, editors. *Light scattering: by small particles*. Courier Dover Publications, New York, 1957.

[72] K. Ikeuchi and K. Sato. Determining reflectance properties of an object using range and brightness images. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 13(11):1139–1153, 1991.

[73] Chika Inoshita, Yasuhiro Mukaigawa, Yasuyuki Matsushita, and Yasushi Yagi. Surface Normal Decomposition: Photometric Stereo for Optically Thick Translucent Objects. In *Proc. of European Conference on Computer Vision (ECCV)*, pages 346–359, 2014.

[74] Henrik Wann Jensen, Stephen R. Marschner, Marc Levoy, and Pat Hanrahan. A Practical Model for Subsurface Light Transport. In *Proc. of ACM SIGGRAPH*, pages 511–518. ACM Press, 2001.

[75] Jun Jiang, Dengyu Liu, Jinwei Gu, and Sabine Süsstrunk. What is the space of spectral sensitivity functions for digital color cameras? In *IEEE Workshop on Applications of Computer Vision (WACV)*, pages 168–179. IEEE, 2013.

[76] D. Jimenez, D. Pizarro, M. Mazo, and S. Palazuelos. Modelling and Correction of Multipath Interference in Time of Flight Cameras. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 893–900. IEEE, 2012.

[77] Mingye Ju, Dengyin Zhang, and Xuemei Wang. Single image dehazing via an improved atmospheric scattering model. *The Visual Computer*, pages 1–13, 2016.

[78] Achuta Kadambi, Jamie Schiel, and Ramesh Raskar. Macroscopic Interferometry: Rethinking Depth Estimation with Frequency-Domain Time-of-Flight. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 893–902, 2016.

[79] Achuta Kadambi, Refael Whyte, Ayush Bhandari, Lee Streeter, Christopher Barsi, Adrian Dorrington, and Ramesh Raskar. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM Trans. on Graphics (ToG)*, 32(6):167, 2013.

[80] Takashi Kakue, Kazuya Tosa, Junpei Yuasa, Tatsuki Tahara, Yasuhiro Awatsuji, Kenzo Nishio, Shogo Ura, and Toshihiro Kubota. Digital Light-in-Flight Recording by Holography by Use of a Femtosecond Pulsed Laser.

IEEE Journal of Selected Topics in Quantum Electronics, 18(1):479–485, 2012.

[81] Rei Kawakami, Yasuyuki Matsushita, John Wright, Moshe Ben-Ezra, Yu-Wing Tai, and Katsushi Ikeuchi. High-resolution hyperspectral imaging via matrix factorization. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2329–2336. IEEE, 2011.

[82] Dae Woong Kim, Stephen Lin, Ki-Sang Hong, and Heung-Yeung Shum. Variational specular separation using color and polarization. In *Proc. of IAPR Workshop on Machine Vision Applications*, pages 176–179, 2002.

[83] Hyeongwoo Kim, Hailin Jin, Sunil Hadap, and Inso Kweon. Specular reflection separation using dark channel prior. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1460–1467, 2013.

[84] J. Kim, D. Lanman, Y. Mukaigawa, and R. Raskar. Descattering transmission via angular filtering. In *Proc. of European Conference on Computer Vision (ECCV)*, pages 86–99, 2010.

[85] Jaewon Kim, Shahram Izadi, and Abhijeet Ghosh. Single-shot layered reflectance separation using a polarized light field camera. In *Proc. of Eurographics Symposium on Rendering (ESR)*, 2016.

[86] Gordon S Kino and Timothy R Corle. *Confocal scanning optical microscopy and related imaging systems*. Academic Press, 1996.

[87] Ahmed Kirmani, Arrigo Benedetti, and Philip A. Chou. SPUMIC: Simultaneous Phase Unwrapping and Multipath Interference Cancellation in Time-of-Flight Cameras using Spectral Methods. In *IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6. IEEE, 2013.

[88] Kazuya Kitano, Takanori Okamoto, Kenichiro Tanaka, Takahito Aoto, Hiroyuki Kubo, Takuya Funatomi, and Yasuhiro Mukaigawa. Recovering Temporal PSF using ToF Camera with Delayed Light Emission. *IPSJ Trans. on Computer Vision and Applications*, 9(15), June 2017.

[89] G. Klinker, S. Shafer, and T. Kanade. The measurement of highlights in color images. *International Journal of Computer Vision (IJCV)*, 2(1):7–32, 1988.

[90] E. L. Krinov. Spectral reflectance properties of natural formations. Technical report, National Research Council of Canada, 1947.

[91] Paul Kubelka. New contributions to the optics of intensely light-scattering materials. part ii: Nonhomogeneous layers∗. *Journal of the Optical Society of America A (JOSA A)*, 44(4):330–335, Apr 1954.

[92] Noriko Kurachi. *The Magic of Computer Graphics*. CRC Press, Florida, 2011.

[93] Douglas Lanman, Daniel Crispell, and Gabriel Taubin. Surround structured lighting for full object scanning. In *Proc. International Conference on 3-D Digital Imaging and Modeling*, pages 107–116. IEEE, 2007.

[94] Seungkyu Lee and Hyunjung Shim. Skewed Stereo Time-of-Flight Camera for Translucent Object Imaging. *Image and Vision Computing*, 43(C):27–38, 2015.

[95] Marc Levoy, Billy Chen, Vaibhav Vaish, Mark Horowitz, Ian McDowall, and Mark Bolas. Synthetic aperture confocal imaging. In *ACM Trans. on Graphics (ToG)*, volume 23, pages 825–834. ACM, 2004.

[96] Marc Levoy and Pat Hanrahan. Light field rendering. In *Proc. of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, pages 31–42. ACM, 1996.

[97] Miao Liao, Xinyu Huang, and Ruigang Yang. Interreflection removal for photometric stereo by using spectrum-dependent albedo. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 689–696. IEEE, 2011.

[98] S. Lin and S.-W. Lee. Detection of specularity using stereo in color and polarization space. *Computer Vision and Image Understanding (CVIU)*, 65(2):336–346, 1997.

[99] S. Lin and S.-W. Lee. An appearance representation for multiple reflection components. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 105–110, 2000.

[100] Chao Liu and Jinwei Gu. Discriminative illumination: Per-pixel classification of raw materials based on optimal projections of spectral brdf. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 797–804, 2012.

[101] Joakim Löw, Joel Kronander, Anders Ynnerman, and Jonas Unger. Brdf models for accurate and efficient rendering of glossy surfaces. *ACM Trans. on Graphics (ToG)*, 31(1):9:1–9:14, Feb 2012.

[102] Wan-Chun Ma, Tim Hawkins, Pieter Peers, Charles-Felix Chabert, Malte Weiss, and Paul Debevec. Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination. In *Proc. of Eurographics Symposium on Rendering (ESR)*, pages 183–194, 2007.

[103] Robert A. Maffione and David R. Dana. Instruments and methods for measuring the backward-scattering coefficient of ocean waters. *Applied Optics*, 36(24):6057–6067, Aug 1997.

[104] S. Mallick, T. Zickler, P. Belhumeur, and D. Kriegman. Specularity removal in images and videos: A pde approach. In *Proc. of European Conference on Computer Vision (ECCV)*, pages 550–563, 2006.

[105] Satya P. Mallick, Todd E. Zickler, David J. Kriegman, and Peter N. Belhumeur. Beyond lambert: Reconstructing specular surfaces using color. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 619–626. IEEE, 2005.

[106] Alkhazur Manakov, John Restrepo, Oliver Klehm, Ramon Hegedus, Elmar Eisemann, Hans-Peter Seidel, and Ivo Ihrke. A reconfigurable camera add-on for high dynamic range, multispectral, polarization, and light-field imaging. *ACM Trans. on Graphics (ToG)*, 32(4):47–1, 2013.

[107] Wojciech Matusik, Hanspeter Pfister, Matt Brand, and Leonard McMillan. A data-driven reflectance model. *ACM Trans. on Graphics (ToG)*, 22(3):759–769, 2003.

[108] William V Meyer, David S Cannell, Anthony E Smart, Thomas W Taylor, and Padetha Tin. Multiple-scattering suppression by cross correlation. *Applied optics*, 36(30):7551–7558, 1997.

[109] Peter J Miller and Clifford C Hoyt. Multispectral imaging with a liquid crystal tunable filter. In *Photonics for Industrial Applications*, pages 354–365. International Society for Optics and Photonics, 1995.

[110] Daisuke Miyazaki, Megumi Saito, Yoichi Sato, and Katsushi Ikeuchi. Determining surface orientations of transparent objects based on polarization degrees in visible and infrared wavelengths. *Journal of the Optical Society of America A (JOSA A)*, 19(4):687–694, 2002.

[111] Peter Morovic and Graham D Finlayson. Metamer-set-based approach to estimating surface reflectance from camera rgb. *Journal of the Optical Society of America A (JOSA A)*, 23(8):1814–1822, 2006.

[112] Ying Mu, Niksa Valim, and Mark J. Niedre. Evaluation of a fast single-photon avalanche photodiode for measurement of early transmitted photons through diffusive media. *Optics Letters*, 38 12:2098–100, 2013.

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

[114] Y. Mukaigawa, R. Raskar, and Y. Yagi. Analysis of scattering light transport in translucent media. *IPSJ Trans. on Computer Vision and Applications*, 3:122–133, 2011.

[115] Volker Müller. Elimination of specular surface-reflectance using polarized and unpolarized light. In *Proc. of European Conference on Computer Vision (ECCV)*, pages 625–635, 1996.

[116] Z. Murez, T. Treibitz, R. Ramamoorthi, and D. J. Kriegman. Photometric Stereo in a Scattering Medium. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(9):1880–1891, Sept 2017.

[117] Nikhil Naik, Achuta Kadambi, Christoph Rhemann, Shahram Izadi, Ramesh Raskar, and Sing Bing Kang. A Light Transport Model for Mitigating Multipath Interference in Time-of-Flight Sensors. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 73–81, 2015.

[118] Nikhil Naik, Shuang Zhao, Andreas Velten, Ramesh Raskar, and Kavita Bala. Single View Reflectance Capture using Multiplexed Scattering and Time-of-Flight Imaging. *ACM Trans. on Graphics (ToG)*, 30(6):171:1–171:10, 2011.

[119] S. G. Narasimhan and S. K. Nayar. Vision and the atmosphere. *International Journal of Computer Vision (IJCV)*, 48(3):233–254, 2002.

[120] S. G. Narasimhan and S. K. Nayar. Contrast restoration of weather degraded images. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 25(6):713–724, 2003.

[121] S. K. Nayar, X. S. Fang, and T. Boult. Separation of reflection components using color and polarization. *International Journal of Computer Vision (IJCV)*, 21(3):163–186, 1997.

[122] S. K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar. Fast separation of direct and global components of a scene using high frequency illumination. In *Proc. of ACM SIGGRAPH*, pages 935–944, 2006.

[123] Shree K. Nayar, Xi-Sheng Fang, and Terrance Boult. Separation of Reflection Components Using Color and Polarization. *International Journal of Computer Vision (IJCV)*, 21(3):163–186, 1997.

[124] Shree K. Nayar, Gurunandan Krishnan, Michael D. Grossberg, and Ramesh Raskar. Fast Separation of Direct and Global Components of a Scene using High Frequency Illumination. *ACM Trans. on Graphics (ToG)*, 25(3):935–944, 2006.

[125] Trung Thanh Ngo, Hajime Nagahara, and Rinichiro Taniguchi. Shape and Light Directions from Shading and Polarization. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[126] Rang MH Nguyen, Dilip K Prasad, and Michael S Brown. Training-based spectral reconstruction from a single rgb image. In *Proc. of European Conference on Computer Vision (ECCV)*, pages 186–201. Springer, 2014.

[127] Tam Nguyen, Quang Nhat Vo, Hyung-Jeong Yang, Soo-Hyung Kim, and Guee-Sang Lee. Separation of specular and diffuse components using tensor voting in color images. *Applied optics*, 53(33):7924–7936, 2014.

[128] K. Nishino, Z. Zhang, and K. Ikeuchi. Determining reflectance parameters and illumination distribution from a sparse set of images for view-dependent image synthesis. In *Proc. of International Conference on Computer Vision (ICCV)*, pages 599–606, 2001.

[129] Seoung Wug Oh, Michael S. Brown, Marc Pollefeys, and Seon Joo Kim. Do it yourself hyperspectral imaging with everyday digital cameras. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2461–2469, 2016.

[130] Matthew O'Toole, Felix Heide, Lei Xiao, Matthias B. Hullin, Wolfgang Heidrich, and Kiriakos N. Kutulakos. Temporal Frequency Probing for 5D Transient Analysis of Global Light Transport. *ACM Trans. on Graphics (ToG)*, 33(4):87:1–87:11, 2014.

[131] M. O' Toole, F. Heide, D. Lindell, K. Zang, S. Diamond, and G. Wetzstein. Reconstructing Transient Images from Single-Photon Sensors. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[132] Marios Papas, Christian Regg, Wojciech Jarosz, Bernd Bickel, Philip Jackson, Wojciech Matusik, Steve Marschner, and Markus Gross. Fabricating translucent materials using continuous pigment mixtures. *ACM Trans. on Graphics (ToG)*, 32(4):146:1–146:12, July 2013.

[133] Jong-Il Park, Moon-Hyun Lee, Michael D Grossberg, and Shree K Nayar. Multispectral imaging using multiplexed illumination. In *Proc. of International Conference on Computer Vision (ICCV)*, pages 1–8. IEEE, 2007.

[134] Hui Qiao, Jingyu Lin, Yebin Liu, Matthias B Hullin, and Qionghai Dai. Resolving Transient Time Profile in ToF Imaging via Log-Sum Sparse Regularization. *Optics letters*, 40(6):918–21, 2015.

[135] Fessenden Reginald. Electric signaling apparatus, January 14 1913. US Patent 1050441.

[136] Weihong Ren, Jiandong Tian, and Yandong Tang. Specular reflection separation with color-lines constraint. *IEEE Trans. on Image Processing*, 26(5):2327–2337, 2017.

[137] Philip Saponaro, Scott Sorensen, Abhishek Kolagunda, and Chandra Kambhamettu. Material classification with thermal imagery. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[138] Y. Sato and K. Ikeuchi. Temporal-color space analysis of reflection. *Journal of the Optical Society of America A (JOSA A)*, 11(7):2990–3002, 1994.

[139] Y. Sato, M. Wheeler, and K. Ikeuch. Object shape and reflectance modeling from observation. In *Proc. of ACM SIGGRAPH*, pages 379–387, 1997.

[140] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, June 2003.

[141] Daniel Scharstein and Richard Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision (IJCV)*, 47(1):7–42, Apr 2002.

[142] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar. Polarization-based vision through haze. *Applied Optics*, 42(3):511–525, 2003.

[143] Karsten Schlüns and Oliver Wittig. Photometric stereo for non-lambertian surfaces using color information. In *Proc. of International Conference on Image Analysis and Processing*, 1993.

[144] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 519–528, June 2006.

[145] S. M. Seitz, Y. Matsushita, and K. N. Kutulakos. A theory of inverse light transport. In *Proc. of International Conference on Computer Vision (ICCV)*, pages 1440–1447, 2005.

[146] S. A. Shafer. Using color to separate reflection components. *Color Research & Application*, 10(4):210–218, 1985.

[147] S. A. Shafer. Using Color to Separate Reflection Components. *Color Research & Application*, 10(4):210–218, 1985.

[148] Hyunjung Shim and Seungkyu Lee. Recovering Translucent Object using a Single Time-of-Flight Depth Camera. *IEEE Transactions on Circuits and Systems for Video Technology*, 26(5):841–854, 2015.

[149] W. Silver. Determining shape and reflectance using multiple images. Technical report, Master thesis, MIT, 1980.

[150] J. D. Spinhirne. Micro pulse lidar. *IEEE Trans. on Geoscience and Remote Sensing*, 31(1):48–55, Jan 1993.

[151] David WJ Stein, Scott G Beaven, Lawrence E Hoff, Edwin M Winter, Alan P Schaum, and Alan D Stocker. Anomaly detection from hyperspectral imagery. *IEEE Signal Processing Magazine*, 19(1):58–69, 2002.

[152] D. Stoppa, D. Mosconi, L. Pancheri, and L. Gonzo. Single-photon avalanche diode cmos sensor for time-resolved fluorescence measurements. *IEEE Sensors Journal*, 9(9):1084–1090, Sep. 2009.

[153] Charlie E. M. Strauss. Synthetic-array heterodyne detection: a single-element detector acts as an array. *Optics Letters*, 19(20):1609–1611, Oct 1994.

[154] Ryuichi Tadano, Adithya Kumar Pediredla, and Ashok Veeraraghavan. Depth selective camera: A direct, on-chip, programmable technique for depth selectivity in photography. In *Proc. of International Conference on Computer Vision (ICCV)*, pages 3595–3603, 2015.

[155] H. Tagare and R. de Figueiredo. A theory of photometric stereo for a class of diffuse non-lambertian surfaces. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 13(2):133–152, 1991.

[156] R. T. Tan and K. Ikeuch. Separating reflection components of textured surfaces using a single image. In *Proc. of International Conference on Computer Vision (ICCV)*, pages 870–877, 2003.

[157] Kenichiro Tanaka, Yasuhiro Mukaigawa, Hiroyuki Kubo, Yasuyuki Matsushita, and Yasushi Yagi. Recovering Transparent Shape from Time-of-Flight Distortion. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4387–4395, 2016.

[158] Kenichiro Tanaka, Yasuhiro Mukaigawa, Hiroyuki Kubo, Yasuyuki Matsushita, and Yasushi Yagi. Recovering inner slices of layered translucent objects by multi-frequency illumination. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 39(4):746–757, 2017.

[159] Kenichiro Tanaka, Yasuhiro Mukaigawa, Hiroyuki Kubo, Yasuyuki Matsushita, and Yasushi Yagi. Recovering Inner Slices of Layered Translucent Objects by Multi-frequency Illumination. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 39(4):746–757, 4 2017.

[160] Ariel Tankus and Nahum Kiryati. Photometric stereo under perspective projection. In *Proc. of International Conference on Computer Vision (ICCV)*, volume 1, pages 611–616. IEEE, 2005.

[161] Michael W Tao, Jong-Chyi Su, Ting-Chun Wang, Jitendra Malik, and Ravi Ramamoorthi. Depth estimation and specular removal for glossy surfaces using point and line consistency with light-field cameras. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 38(6):1155–1169, 2016.

[162] Shoji Tominaga. Multichannel vision system for estimating surface and illumination functions. *Journal of the Optical Society of America A (JOSA A)*, 13(11):2163–2173, 1996.

[163] Shoji Tominaga and Ryousuke Okajima. Object recognition by multispectral imaging with a liquid crystal filter. In *Proc. of International Conference on Pattern Recognition (ICPR)*, volume 1, pages 708–711. IEEE, 2000.

[164] T. Treibitz, Z. Murez, B. G. Mitchell, and D. Kreigman. Shape from Fluorescence. In *Proc. of European Conference on Computer Vision (ECCV)*, 2012.

[165] T. Treibitz and Y. Y. Schechner. Active polarization descattering. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 31(3):385–399, 2009.

[166] T. Treibitz and Yoav. Y. Schechner. Active Polarization Descattering. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 31(3):385–399, 2009.

[167] Chia-Yin Tsai, Kiriakos N Kutulakos, Srinivasa G Narasimhan, and Aswin C Sankaranarayanan. The geometry of first-returning photons for non-line-of-sight imaging. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7216–7224, 2017.

[168] Shinji Umeyama and Guy Godin. Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 26(5):639–647, 2004.

[169] Andreas Velten, Ramesh Raskar, Di Wu, Adrian Jarabo, Belen Masia, Christopher Barsi, Chinmaya Joshi, Everett Lawson, Moungi Bawendi, and Diego Gutierrez. Femto-Photography: Capturing and Visualizing the Propagation of Light. *ACM Trans. on Graphics (ToG)*, 32(4):44:1–44:8, 2013.

[170] Andreas Velten, Thomas Willwacher, Otkrist Gupta, Ashok Veeraraghavan, Moungi G Bawendi, and Ramesh Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature communications*, 3:745, 2012.

[171] Michael J. Vrhel, Ron Gershon, and Lawrence S. Iwan. Measurement and analysis of object reflectance spectra. *Color Research and Application*, 19(1):4–9, 1994.

[172] Fan Wang, Samia Ainouz, Caroline Petitjean, and Abdelaziz Bensrhair. Specularity removal: A global energy minimization approach based on polarization imaging. *Computer Vision and Image Understanding (CVIU)*, 158:31–39, 2017.

[173] Rui Wang, Ewen Cheslack-Postava, Rui Wang, David Luebke, Qianyong Chen, Wei Hua, Qunsheng Peng, and Hujun Bao. Real-time editing and relighting of homogeneous translucent materials. *The Visual Computer*, 24(7):565–575, 2008.

[174] Tim Weyrich, Pieter Peers, Wojciech Matusik, and Szymon Rusinkiewicz. Fabricating microgeometry for custom surface reflectance. *ACM Trans. on Graphics (ToG)*, 28(3):32:1–32:6, July 2009.

[175] L Wind and W W Szymanski. Quantification of scattering corrections to the beer-lambert law for transmittance measurements in turbid media. *Measurement Science and Technology*, 13(3):270–275, Feb 2002.

[176] L. B. Wolff and T. E. Boult. Constraining object features using a polarization reflectance model. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 13(7):635–657, 1991.

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

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

[179] Di Wu, Andreas Velten, Matthew O'Toole, Belen Masia, Amit Agrawal, Qionghai Dai, and Ramesh Raskar. Decomposing Global Light Transport using Time of Flight Imaging. *International Journal of Computer Vision (IJCV)*, 107(2):123–138, 2014.

[180] Lun Wu, Arvind Ganesh, Boxin Shi, Yasuyuki Matsushita, Yongtian Wang, and Yi Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *Proc. of Asian Conference on Computer Vision (ACCV)*, 2010.

[181] T.-P. Wu and C.-K. Tang. Separating specular, diffuse, and subsurface scattering reflectances from photometric images. In *Proc. of European Conference on Computer Vision (ECCV)*, pages 419–433, 2004.

[182] Qingxiong Yang, Jinhui Tang, and Narendra Ahuja. Efficient and robust specular highlight removal. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 37(6):1304–1311, 2015.

[183] 小松 英彦. 質感の科学 : 知覚・認知メカニズムと分析・表現の技術. 朝倉書店, 2016. (in Japanese).

List of Publications

Award

1. MIRU フロンティア賞, 第 21 回 画像の認識・理解シンポジウム（MIRU2018），2018 年 8 月 8 日

2. Best Poster Award, The 7th International Workshop on Robust Computer Vision (IWRCV), Jan. 5, 2013

3. 卒論セッション最優秀賞，情報処理学会 コンピュータビジョンとイメージメディア（CVIM）研究会，2011 年 5 月 20 日

Invited Talk

1. Tsuyoshi Takatani, Kenichiro Tanaka, Hiroyuki Kubo, Takuya Funatomi, Yasuhiro Mukaigawa, Reproducing Translucency by Computational Fabrication, Imaging Conference JAPAN 2019, Chiba, Japan, Jul. 2019

2. Kenichiro Tanaka, Nobuhiro Ikeya, Tsuyoshi Takatani, Takuya Funatomi, Hiroyuki Kubo, Yasuhiro Mukaigawa, Time-resolved Light Transport Decomposition for Thermal Photometric Stereo, 第 21 回 画像の認識・理解シンポジウム（MIRU2018），北海道札幌市，2018 年 8 月

3. Tsuyoshi Takatani, Takahito Aoto, Yasuhiro Mukaigawa, One-shot Hyperspectral Imaging using Faced Reflectors, The 6th IEEE International Workshop on Computational Cameras and Displays (CCD), Honolulu, Hawaii, Jul, 2017

Journal Paper

1. Tsuyoshi Takatani, Yasuhiro Mukaigawa, Yasuyuki Matsushita, Yasushi Yagi, Decomposition of Reflection and Scattering by Multiple-Weighted Measurements, IPSJ Transactions on Computer Vision and Applications, 10:13, pp. 1–13, Oct. 2018

2. Tsuyoshi Takatani, Koki Fujita, Kenichiro Tanaka, Takuya Funatomi, Yasuhiro Mukaigawa, Controlling Translucency by UV Printing on a Translucent Object, IPSJ Transactions on Computer Vision and Applications, 10:7, pp. 1–6, Jun. 2018

International Conference (Reviewed)

1. Tsuyoshi Takatani, Takahito Aoto, Kenichiro Tanaka, Takuya Funatomi, Yasuhiro Mukaigawa, Reconstruction of Volumetric Reflectance using Spatio-Sequential Frequency Correlation Imaging, ACM SIGGRAPH Asia Technical Briefs, pp. 1–4, Tokyo, Japan, Dec. 2018

2. Kenichiro Tanaka, Nobuhiro Ikeya, Tsuyoshi Takatani, Takuya Funatomi, Hiroyuki Kubo, and Yasuhiro Mukaigawa, Time-resolved Light Transport Decomposition for Thermal Photometric Stereo, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4804–4813, Salt Lake City, Utah, Jun. 2018

3. Kazusa Tsubota, Tsuyoshi Takatani, Takahito Aoto, Kenichiro Tanaka, Hiroyuki Kubo, Takuya Funatomi, and Yasuhiro Mukaigawa, Examining Single Scattering Region in Concentration, Depth, and Wavelength on Diluted Media, OSA-OSJ Joint Symposia on Nanophotonics and Digital Photonics, pp. 101–102, Tokyo, Japan, Oct. 2017

4. Tsuyoshi Takatani, Takahito Aoto, and Yasuhiro Mukaigawa, One-shot Hyperspectral Imaging using Faced Reflectors, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2692-2700, Honolulu, Hawaii, Jul. 2017

5. Tsuyoshi Takatani, Yasuyuki Matsushita, Stephen Lin, Yasuhiro Mukaigawa, and Yasushi Yagi, Enhanced Photometric Stereo with Multispectral Images, IAPR tInternational Conference on Machine Vision Applications (MVA), pp. 343–346, Kyoto, Japan, May 2013

International Workshop

1. Kenichiro Tanaka, Nobuhiro Ikeya, Tsuyoshi Takatani, Takuya Funatomi, Hiroyuki Kubo, and Yasuhiro Mukaigawa, Time-resolved Light Transport Decomposition for Thermal Photometric Stereo, IEEE International Conference on Computational Photography (ICCP), Pittsburgh, Pennsylvania, May 2018

2. Tsuyoshi Takatani, Takahito Aoto, and Yasuhiro Mukaigawa, One-shot Hyperspectral Imaging using Faced Reflectors, The 12th International Workshop on Robust Computer Vision (IWRCV), Nara, Japan, Jan. 2018

3. Tsuyoshi Takatani, Yasuhiro Mukaigawa, Yasuyuki Matsushita, and Yasushi Yagi, Multiple Weighted Measurements for Decomposition of Photometric Components, The 7th International Workshop on Robust Computer Vision (IWRCV), Osaka, Japan, Jan. 2013.

Domestic Conference (Reviewed)

1. 高谷剛志，向川康博，松下康之，八木康史，多重重み付け計測による反射・散乱光の分解，第 14 回 画像の認識・理解シンポジウム（MIRU2011），石川県金沢市，2011 年 7 月

Domestic Workshop

1. Masaki Kaga, Takahiro Kushida, Tsuyoshi Takatani, Kenichiro Tanaka, Takuya Funatomi, Yasuhiro Mukaigawa, Thermal Non-Line-of-Sight Imaging from Specular and Diffuse Reflections, 第 22 回 画像の認識・理解シンポジウム（MIRU2019），大阪府大阪市，2019 年 7 月

2. Tsuyoshi Takatani, Koki Fujita, Kenichiro Tanaka, Takuya Funatomi, Yasuhiro Mukaigawa, Controlling Translucency by UV Printing on a Translucent Object, 第 21 回 画像の認識・理解シンポジウム（MIRU2018），北海道札幌市，2018 年 8 月

3. 高谷剛志，藤田紘樹，田中賢一郎，舩冨卓哉，向川康博，半透明材質への UV 印刷による半透明感の制御，情報処理学会 コンピュータグラフィクスとビジュアル情報学（CGVI）研究会，東京都文京区，2018 年 3 月

4. 池谷信寛，田中賢一郎，高谷剛志，久保尋之，舩冨卓哉，向川康博，熱伝導の経時特性に基づく成分分解を用いた遠赤外照度差ステレオ法，情報処理学会 コンピュータビジョンとイメージメディア（CVIM）研究会，大阪府堺市，2018 年 1 月

5. 高谷剛志，松下康之，Stephen Lin，向川康博，八木康史，多波長画像を用いた照度差ステレオ法の精度向上，情報処理学会 コンピュータビジョンとイメージメディア（CVIM）研究会，東京都小金井市，2012 年 9 月

6. 高谷剛志，向川康博，松下康之，八木康史，多重重み付け計測による反射光と散乱光の分解，情報処理学会 コンピュータビジョンとイメージメディア（CVIM）研究会，大阪府茨木市，2011 年 5 月