Illumination Recovery and Appearance Sampling for Photorealistic Rendering

(写実的な画像生成のための光源環境推定と
物体表面の見えの標本化)

佐藤 いまり

Acknowledgments

I would first like to show my genuine appreciation to my advisor, Katsushi Ikeuchi, for the support and encouragement he gave me through my Ph.D. course of Interdisciplinary Information Studies at the University of Tokyo. He always cheered me up through good times and bad, and I have learned from him how to enjoy doing research in computer vision. I would also like to thank my secondary advisor, Yoichiro Kawaguchi, and my thesis committee, Hiroshi Harashima, Yasushi Yamaguchi, and Takeshi Naemura, for their valuable feedback regarding my research. I would like to show my gratitude to Takeo Kanade for giving me opportunities to study computer vision and computer graphics at the Robotics Institute in Carnegie Mellon University.

During my stay at the University of Tokyo, I was fortunate to have many wonderful people to work with. Many thanks to my colleagues at Computer Vision Laboratory and Human Computer Interaction Laboratory. My special thanks go to Masataka Kagesawa for his warm support, to Takahiro Okabe, Robby T. Tan, Takeshi Oishi, Hiroshi Kawasaki, Daisuke Miyazaki, Rei Kawakami, Dong Wang, and Morihiro Hayashida for the interesting discussions we had about research in the field of physics-based vision. I also thank Sakie Suzuki, Mariko Araki, Keiko Motoki, Yoshiko Matsuura, Yuko Saiki, Yukie Hirayama, and Mawo Kamakura for delightful tea time chatting that made me feel relaxed, and to Mark Ashdown, Kenji Oka, Yasuhiro Ono, Yoshinori Kobayashi for sharing with me the riches of their knowledge.

I am also very grateful to the late Marie Elm and Joan Knapp for taking the time to edit my papers. Their kind support enabled me to complete my thesis. Since the beginning of my research career, the late Marie Elm was always kind in giving her time to correct my writing and improve my writing skills. I wish I could show this thesis to Marie.

Finally I would like to thank my family, Shunzo, Nobuko, Yoshitaka, Kazuko, Yukari, and Toshiyuki, for always supporting me unconditionally. Especially I would like to express my deepest gratitude to my husband, Yoichi, who is my consistent encourager and my lifetime mentor, for his support that keeps me pushing my way toward a Ph.D.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

One of the important goals in computer graphics is achieving photographic realism of synthesized images. Techniques for rendering realistic computer graphics objects and seamlessly merging them into images of real scenes especially attract a great deal of attention in media industries such as film making, television broadcasting, game making, advertising, and so on.

For the purpose of synthesizing the realistic appearance of objects, conventional *model-based rendering* techniques have been intensively developed. Model-based rendering techniques synthesize the appearance of objects based on empirically or analytically given reflection models.[1] To use these models, the geometric and photometric information about the scene needs to be provided: the shapes of objects in the scene, their surface reflectance properties, and the lighting conditions of the scene where those objects are placed.

Regarding the geometric information of a scene, 3D shapes of objects are often manually produced by using well-designed CAD modelers. Also, for modeling the more complex shapes of real objects, a number of approaches have been developed, such as range image merging techniques and 3D photography techniques. In contrast, providing photometric information about a scene, such as surface reflectance properties and lighting conditions, has been a difficult task.

Surface reflectance properties greatly influence the appearance of an object. The ap-

[1] Commonly used reflection models in the field of computer vision and computer graphics include the Lambertian model, the Phong model, the Blinn-Phong model, the Torrance-Sparrow model, the Cook-Torrance model, the Ward model, and the Lafortune model.

pearance of a metallic surface is completely different from that of a matted surface even under the same lighting conditions. In addition, its appearance changes significantly under different lighting conditions. It is thus important to provide not only appropriate surface reflectance properties of objects in a scene but also appropriate illumination conditions so that the realistic appearance of the objects can be synthesized under these illumination conditions. Nevertheless, photometric information about a scene tends to be manually provided by a user.

Since it is difficult to imagine the appearance of an object directly from reflectance parameters, the input process of manually specifying its reflectance properties is normally non-intuitive and thus time-consuming. The correct appearance of a scene is difficult to achieve unless we stop relying on our instinct for adjusting reflectance parameters. As for providing lighting conditions, a scene generally includes both direct and indirect illumination distributed in a complex way, and it is difficult for a user to manually specify such complex illumination distribution.

In order to overcome these difficulties in providing photometric information about a scene, techniques for automatically providing the photometric models of a scene have been studied in the fields of both computer vision and computer graphics research. In particular, techniques that use a set of images of a scene provided under different viewing and/or lighting conditions for determining its geometric and photometric information are called *image-based modeling*.

This thesis addresses two issues of image-based modeling for synthesizing the photorealistic appearance of real objects under natural illumination conditions: capturing real-world illumination, and modeling the complex appearance of real objects for variable illumination. Regarding the first issue of capturing real-world illumination, both image-based lighting and inverse lighting approaches are studied. As for the second issue of modeling the complex appearance of real objects for variable illumination, we carefully investigate the requirement for input images in image-based rendering in order to correctly produce the appearance of an object under arbitrary illumination. We present a novel method for analytically obtaining a set of basis images from input images of the object taken under a point light source or extended light sources.

1.1 Image-based Lighting for Measuring Real-World Illumination

One of the image-based modeling techniques that measures real-world illumination conditions from photographically acquired images of a real scene is called *image-based lighting*. Image-based lighting techniques have been developed successfully with practical applications [17, 12, 7, 29]. However, two difficulties in image-based lighting still remain to be solved: how to construct a geometric model of the scene, and how to capture a wide field of view of the scene.

In Chapter 2, we confront these two difficulties and propose an efficient method for automatically measuring illumination distribution of a real scene by using a set of omni-directional images of the scene taken by a CCD camera with a fisheye lens. There are three reasons why we use omni-directional images rather than images taken by a camera with an ordinary lens. First, because of the wide field of view of a fisheye lens (180 degrees), we can easily capture illumination from all directions from a far fewer number of omni-directional images. Second, since a fisheye lens is designed so that an incoming ray from a particular direction is projected onto a particular point on an imaging plane, we do not have to concern ourselves with computing directions of incoming rays and considering a sampling frequency of the incoming rays. Third, we are able to use the directions of the incoming rays for automatically constructing a geometric model of the scene with the wide field of view of a fisheye lens.

In our method, based on an omni-directional stereo algorithm, a geometric model of the scene is first constructed from a pair of omni-directional images taken from different locations. Then radiance of the scene is computed from a sequence of omni-directional images taken with different shutter speeds and mapped onto the constructed geometric model. The construction of this geometric model with radiance, referred to as a radiance map, is necessary in order to compute the radiance distribution seen from any point in the scene. In other words, without constructing a radiance map, we can determine only the radiance distribution seen from the particular point where the omni-directional image was captured. Once a radiance map is constructed as a triangular mesh, an appropriate radiance distribution can be used for rendering a virtual object and for generating shadows cast by the virtual object onto the real scene wherever the virtual object is placed in the

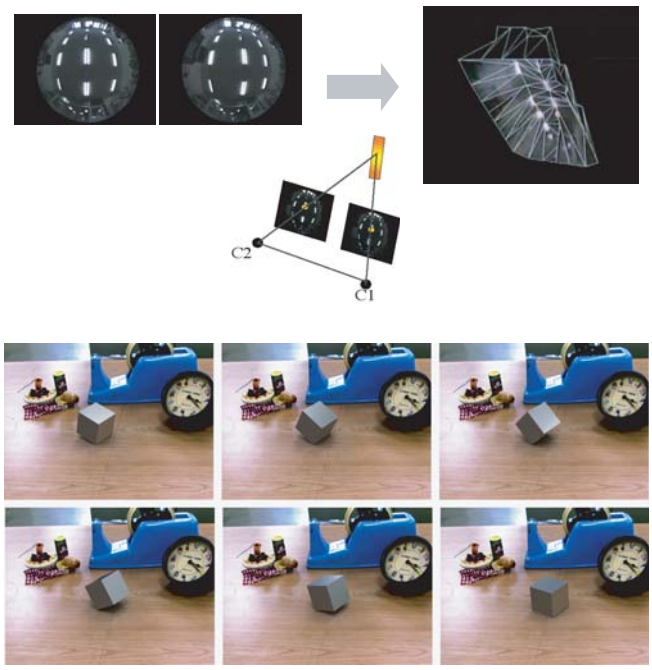


Figure 1.1: Measured illumination distribution based on the proposed omni-directional stereo algorithm and synthesized images under the captured illumination distribution.

scene.

Figure 1.1 shows the obtained triangular mesh that represents the radiance distribution as its color texture and synthesized images under the captured illumination distribution. In the images synthesized by our method, shading of the virtual object blends well into the scene. Also, the virtual object casts a shadow with a soft edge on the tabletop in the same way as do the other objects in the scene.

1.1.1 Image Synthesis under Dynamically Changing Illumination

In Chapter 3, we pursue the possibility of real-time rendering of synthetic objects with natural shading and cast shadows superimposed onto a real scene whose illumination condition is dynamically changing.

In general, high computational cost for rendering virtual objects with convincing shading and shadows, such as interreflections or soft shadows under area light sources, prohibits real-time synthesis of composite images with superimposed virtual objects. From this limitation, simple rendering algorithms supported by commonly available graphics hardware

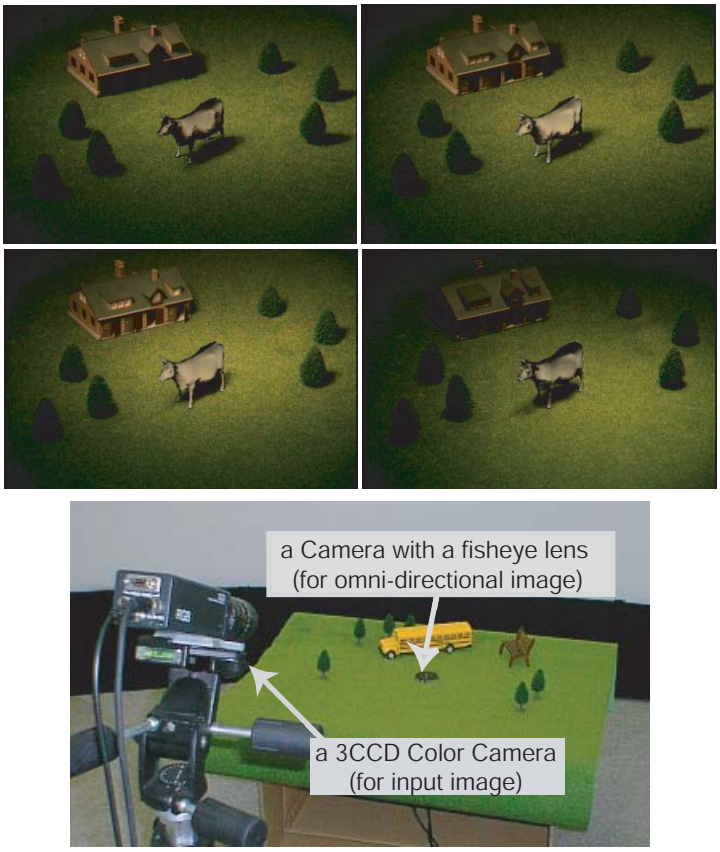


Figure 1.2: Image acquisition system and synthesized images by our method: a color camera with a fish-eye lens is used for capturing illumination of the scene.

need to be used for the applications that require real-time image synthesis. Computationally expensive rendering algorithms are not usually supported by such graphics hardware, and this leads to some restrictions on achievable image qualities.

Alternative approaches have been proposed for re-rendering a scene as a linear combination of a set of pre-rendered basis images of the scene [10, 44, 11]. These approaches are based on the linearity of scene radiance with respect to illumination intensities. Since this linearity holds for scenes with complex geometry and complex photometric properties such as interreflections between objects and cast shadows, photo-realistic appearance of a scene can be synthesized based on this simple linear combination framework.

Most of the previously proposed methods, however, have been developed for the task of interactive lighting design. Therefore, basis lights under which a set of basis images are rendered are intentionally positioned at the desired locations, so that a scene under desired

lighting configurations can be efficiently synthesized. Recently, Debevec et al. introduced a method for re-rendering a human face based on a linear combination of face images taken under densely sampled incident illumination directions in [8]. This method further considered a model of skin reflectance to estimate the appearance of the face seen from novel viewing directions and under arbitrary illumination.

In this thesis, we generalize the approach based on the linearity of scene radiance with respect to illumination radiance and present an efficient technique for superimposing synthetic objects with natural shadings and cast shadows onto a real scene whose illumination is dynamically changing.[2] The main advantage of the proposed method is that image quality is not affected by the requirement for real-time processing, since reference images are rendered off-line. This enables us to employ computationally expensive algorithms for providing reference images, and this results in achieving high quality in the final composite images of the scene.

The image acquisition system and several examples of synthesized images are shown in Figure 1.2. In those images, the virtual cow casts a shadow with a soft edge on the grass in the same way as the other objects such as trees in the scene do.

1.2 Inverse Lighting for Estimating Real-World Illumination

There has been another approach called *inverse rendering* in image-based modeling that deals with an inverse problem of traditional model-based rendering. As described in model-based rendering, the image brightness of an object can be computed as the function of the shape of the object, its surface reflectance properties, and the illumination condition where the object is located [23, 24]. The relationship among these provides three research areas in inverse rendering:

- Shape-from-Brightness for recovering the shape of the object from its reflectance properties and the known illumination condition,
- Reflectance-from-Brightness for recovering the surface reflectance properties of the object from its shape and the known illumination condition, and

[2] We consider that a scene consists of both real objects and synthetic objects with fixed scene geometry and the scene is viewed from a fixed viewing point under dynamically changing illumination.

- Illumination-from-Brightness for recovering unknown illumination conditions of the scene based on the knowledge of the shape and the surface reflectance properties of the object.

In inverse rendering, the first two kinds of analyses, shape-from-brightness and reflectance-from-brightness, have been intensively studied using the shape from shading method [25, 27, 26, 50], as well as through reflectance analysis research [28, 45, 22, 1, 32, 37, 57, 71].

In contrast, relatively limited amounts of research have been conducted in the third area, also known as *inverse lighting* [49]. One of the main advantages of inverse lighting over the former image-based lighting is that it does not require additional images for capturing illumination of a scene, but uses the appearance of objects located in a scene instead for recovering an illumination distribution of the scene.

In general, real scenes include both direct and indirect illumination distributed in a complex way, and this makes it difficult to analyze characteristics of the illumination distribution of the scene from image brightness in inverse lighting. As a consequence, most of the previously proposed approaches were conducted under very specific illumination conditions, for example, there were several point light sources in the scene, and those approaches were difficult to extend to more natural illumination conditions [26, 28, 57, 74, 65]. Or multiple input images taken from different viewing angles were necessary [34, 47].

Pioneering work in the field of inverse lighting for recovering natural illumination conditions of real scenes was proposed by Marschner and Greenberg [39]. This work proposed to approximate the entire illumination with a set of basis lights located in a scene and estimated their radiance values from shadings of objects observed in that scene. Although this method had an advantage over the previous methods of not requiring knowledge about the light locations of the scene, the estimation relies on the changes in appearance observed on an object surface assumed to be Lambertian, and therefore some restrictions were imposed on the shape of the object, for example, the object must have a large amount of curvature.

Later, Ramamoorthi and Hanrahan defined the conditions under which condition inverse rendering could be done robustly, based on their proposed signal-processing framework that described the reflected light field as a convolution of the lighting and the bidirectional reflectance distribution function (BRDF) [52]. It was shown through their analysis that changes in appearance observed on Lambertian surfaces were not necessarily suitable for estimating high frequency components of illumination distribution of a scene.

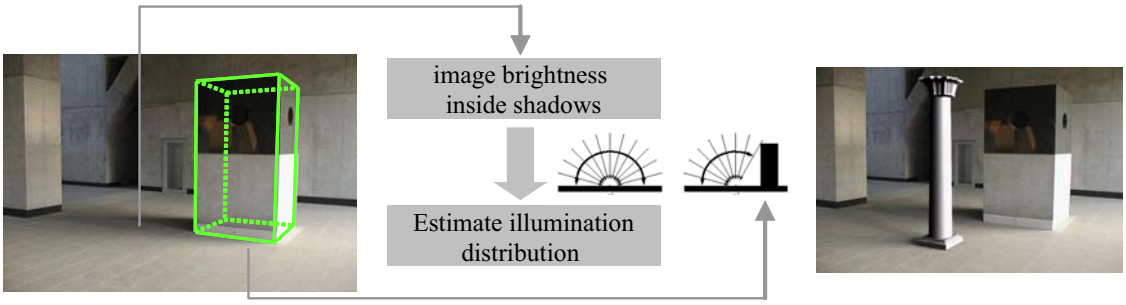


Figure 1.3: An inverse lighting approach for recovering an illumination distribution of a scene from image brightness inside shadows cast by an object of known shape in the scene.

In Chapter 4 of this thesis, we demonstrate the effectiveness of using occluding information of incoming light in estimating an illumination distribution of a scene. Shadows in a scene are caused by the occlusion of incoming light, and thus contain various pieces of information about the illumination of the scene. Nevertheless, shadows have been used for determining the 3D shape and orientation of an object that casts shadows onto the scene [38, 33, 60, 4], while very few studies have focused on the illuminant information that shadows could provide. In our method, image brightness inside shadows is effectively used for providing distinct clues to estimate an illumination distribution.

Chapter 4 further addresses two issues in inverse lighting. First, the method combines the illumination analysis with an estimation of the reflectance properties of a shadow surface. This makes the method applicable to cases where reflectance properties of a surface are not known a priori and enlarges the variety of images applicable to the method. Second, an adaptive sampling framework for efficient estimation of illumination distribution is introduced. Using this framework, we are able to avoid unnecessarily dense sampling of the illumination and can estimate the entire illumination distribution more efficiently with a smaller number of sampling directions of the illumination distribution.

Later in Chapter 5, we also discuss the amount of the information about the illumination distribution of the scene obtainable from a given image of a scene. In general, the amount of information obtainable from an image is determined depending on how much of the shadow surface is blocked by objects in a scene and how much of the scene is covered by the field of view of the camera taking the image of the scene. In particular, two main factors that control the stability of the illumination estimation from shadows are analyzed: blocked view of shadows and limited sampling resolution for radiance distribution inside

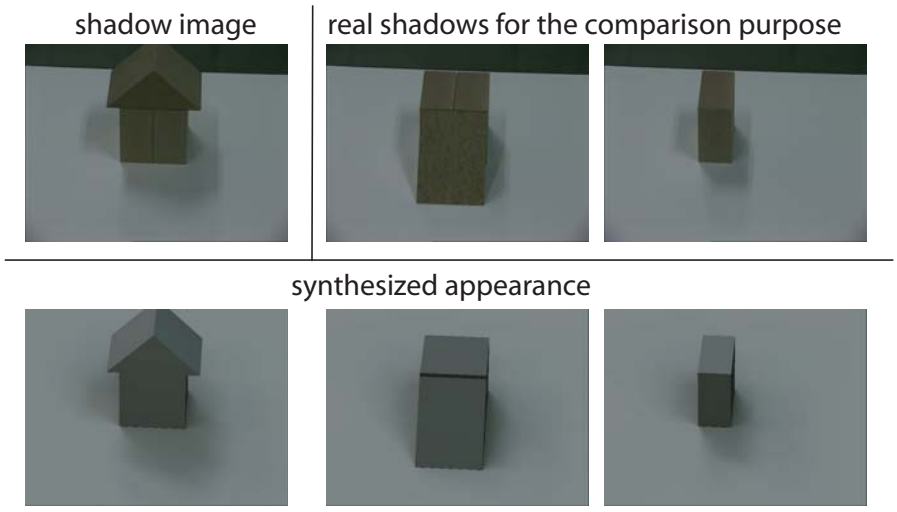


Figure 1.4: Synthesized appearance using the estimated illumination distribution.

shadows.

Then, based on the analysis, a robust method is presented. For estimating the illumination distribution of a scene reliably by taking stability issues into considerations, we propose to change the sampling density of the illumination distribution depending on the amount of the information obtainable from a shadow image for a particular direction of the illumination distribution. For using radiance distribution inside the penumbra of the shadows correctly, we introduce a super-sampling scheme for examining the occlusion of incoming light from each light source. We also explain the optimal sampling of image pixels and the selection of illumination distribution samplings for more stable computation.

All of these extensions contribute to improve stability and accuracy of illumination estimation from shadows, and illumination distribution can be estimated in a reliable manner with these proposed improvements regardless of types of input images such as the shape of an occluding object or a camera position.

In the bottom row of Figure 1.4, several synthetic objects were superimposed onto the surface using the illumination distribution estimated from the input shadow image shown in this figure. It is worth noting that in this example, a relatively large area of the shadow surface is occluded by the occluding object, and it is often difficult to provide a correct estimate of the illumination distribution in such a case.

Even in this challenging case, our proposed approach could reliably estimate the illumination distribution of the scene by taking stability issues into consideration. Shadows cast by those synthetic objects closely resemble those cast by the real objects, and this shows that the estimated illumination distribution gives a good presentation of that of the real scene.

Recently, based on the signal-processing framework proposed by Ramamoorthi and Hanrahan, it was shown that high frequency components of the appearance of an object surface could retain significant energy by taking the occlusion of incoming light as well as its bidirectional reflectance distribution function (BRDF) into account in [48]. This indicates that the use of shadows for illumination estimation has the significant advantage of providing more clues to the high frequency components of illumination distribution of a scene.

1.2.1 Obtaining Shading Models of Oil Paintings

As an application for the proposed inverse lighting approach we also present a new technique for superimposing synthetic objects onto oil paintings with artistic shadings that are consistent with those originally painted by the artists in Appendix. In a colored medium such as oil painting, artists often use color shift techniques for adding some artistic tones to their paintings as well as for enlarging their dynamic ranges.

(a) (b)

Figure 1.5: Superimposing a synthetic chair into the painting by Rembrandt.

In this thesis, we attempt to determine the mechanisms for color shifts performed by artists and to automate their processes so that we can superimpose onto paintings synthetic

objects that have consistent shadings. We first study characteristics of shadows observed both in real scenes and in paintings to discover how intrinsic color shifts were performed by artists. In particular, we analyze brightness distribution inside shadows observed in a painting. Then, we adapt the acquired mechanisms so that we can superimpose synthetic objects with consistent shadings onto oil paintings (Figure 1.5).

1.3 Image-Based Rendering under Novel Lighting Conditions

Inverse rendering carries out the opposite procedures of model-based rendering to provide object and illumination models of a real scene from photographically available information of the scene. Once models of a scene are acquired, new images of the scene under novel lighting and/or viewing conditions can be synthesized by using conventional model-based rendering techniques.

On the other hand, the approach called *image-based rendering* directly uses the original set of input images of a scene for producing new images of the scene under novel conditions [61]. Depending on which scene conditions should be modified, image-based rendering techniques are classified into three categories: image-based rendering under novel viewing conditions, image-based rendering under novel lighting conditions, and image-based rendering under novel viewing and novel lighting conditions. In this thesis, we consider the second category, image-based rendering under novel lighting conditions.

In contrast with model-based rendering techniques, image-based rendering techniques do not require full radiometric computation to synthesize the photo-realistic appearance of objects in a scene. This makes the cost to produce new images of the scene independent of the scene complexity. Also image-based rendering techniques normally do not require geometric and photometric models of a scene.[3] Image-based rendering, however, has a tendency to require many input images of a scene to synthesize reasonably realistic appearance of the scene. This results in the requirement for a large amount of both computer memory and data storage.

While there may seem to be a large variety of possible appearances for a given object,

[3] Some image-based rendering techniques make use of geometric models of a scene for better compression of its appearance.

it has been demonstrated in previous research that the changes in appearance of an object for varying illumination can be represented with a linear subspace spanned by a set of basis images of the object. For instance, in the case of a convex Lambertian object, its appearance seen under distant illumination without attached and cast shadows can be described with a 3-D linear subspace spanned from three input images of the object taken under linearly independent lighting conditions [42, 58, 73]. Even taking into account attached shadows, most of the image variation of a human face or other object under varying illumination was shown to be adequately represented by a low-dimensional linear subspace slightly higher than 3-D [21, 14, 72]. A similar observation was utilized for object recognition in [18, 19].

A set of basis images spanning such a linear subspace is often provided by applying principal-component analysis to the input images of an object taken under different lighting conditions. Since little is known about how to sample the appearance of an object in order to obtain its basis images correctly, a large number of input images taken by moving a point light source along a sphere surrounding the object are generally provided.

Recent investigations in frequency-space analysis of reflection have shown that the appearance of an object under varying complex illumination conditions can be well represented with a linear subspace spanned by basis images of the object, called *harmonic images*, each of which corresponds to an image of the object illuminated under *harmonic lights* whose distributions are specified in terms of spherical harmonics [52, 53, 2].[4]

Hence if harmonic lights can be physically constructed in a real setting, harmonic images of a real object can be obtained simply as images of the object seen under these light sources. However, harmonic lights are complex diffuse light sources comprising both negative and positive radiance and are thus difficult to physically construct in a real setting. Therefore, most of the previously proposed techniques synthetically compute harmonic images from the knowledge of an object's 3-D shape and reflectance properties.

[4] Harmonic images have been also used for the purpose of efficient rendering of an object under complex illumination [54, 62].

1.3.1 Appearance Sampling for Obtaining a Set of Basis Images for Variable Illumination

This motivated us to develop a method for analytically obtaining a set of basis images for varying illumination from input images of the object taken under a set of realizable point light sources. The main contribution of our work is that we show that a set of lighting directions can be determined for sampling images of an object depending on the spectrum of the object's BRDF in the angular frequency domain such that a set of harmonic images can be obtained analytically based on the sampling theorem on spherical harmonics [13].

Using those sampling directions determined from the sampling theory, we are able to obtain harmonic images by using a significantly smaller number of input images than other techniques that do not take into account a relationship between a spectrum of BRDFs and a sampling density of illumination directions.

In addition, unlike other methods based on spherical harmonics, our method does not require the shape and reflectance model of an object used for rendering harmonics images of the object synthetically. Thus, our method can be easily applied to determine a set of basis images for representing the appearance change of a real object under varying illumination conditions.

An overview of our hardware setup[5] used for obtaining the input images of the objects is shown in Figure 1.6; an array of light sources is mounted on a turntable. These light sources are equally spaced in elevation, and the set of light sources is rotated around the objects in azimuthal. By using this hardware setup, input images of an object are taken under a point light source positioned at equiangular grid points and used for obtaining a set of basis images of the object: 36 input images were used for the sheep and the Venus examples, and 144 input images were used for the fish example.

To demonstrate how well the recovered harmonic images represent the appearance of those objects under varying illumination, we synthesize their appearance from the recovered harmonic images under several natural illumination conditions provided by high-dynamic range light probe measurements by [7]. In Figure 1.6, synthesized appearance changes dynamically depending on characteristics of the illumination, and one can say from this that the proposed method succeeded in providing a set of basis images representing

[5] *Surface Reflectance Sampler*, TechnoDream21 corporation

Figure 1.6: Hardware setup and synthesized images of objects under natural illumination. The first row shows illumination maps. The second, third and forth rows show synthesized appearance of objects under the corresponding illumination map.

appearance of those objects under varying illumination.

1.3.2 Anti-aliasing Framework in Appearance Sampling

Our proposed appearance sampling approach does not require object models such as the 3D shape and the surface reflectance properties of an object, but the information about the highest frequency that the appearance of an object contains needs to be provided in order to determine the number of point light sources required to capture its appearance. The sampling theorem states that the higher the frequency content of an object's appearance, the more input images are required to obtain a correct set of basis images. The number of input images required may become extremely large in the case of highly specular surfaces containing a large quantity of high frequency components in their reflection.

In this case, insufficient sampling of an object appearance will result in aliasing in the basis images, and this will lead to undesirable artifacts in the synthesized appearance. Since the number of input images provided for modeling an object's appearance is usually limited, an anti-aliasing framework for obtaining a set of correct basis images from an insufficient number of object input images is needed. However, this aliasing problem has not been carefully considered in previous studies.

In Chapter 7, we carefully study this issue of aliasing and extend the method based

on the sampling theorem further for reducing the artifacts due to aliasing, by substituting extended light sources (*ELS*) for a point light source to sample the reflection kernel of a real object. The use of ELS for modeling the shape and reflectance of an object was originally introduced in [46]. We extend their analysis further in the angular frequency domain so that the harmonic images of an object of arbitrary surface materials can be obtained without suffering from aliasing caused by insufficient sampling of its appearance.

The use of ELS has the following advantages. ELS have a radiance distribution that is similar to that of the Gaussian function, and this enables extended sources to function as a low-pass filter when the appearance of an object is sampled under them. This enables us to obtain a set of basis images of an object for varying illumination without suffering from aliasing caused by insufficient sampling of its appearance. In addition, ELS can reduce high contrast in image intensities due to specular and diffuse reflection components. This helps avoid saturation so that we are able to observe both specular and diffuse reflection components in the same image taken with a single shutter speed.

Once a set of basis images of an object is obtained, its appearance under natural illumination conditions can be synthesized simply as a linear combination of these basis images whose linear coefficients are computed from the given lighting conditions, and these lighting conditions can be modeled by our proposed image-based or inverse lighting methods. We describe more details about each method in the following chapters and present concluding remarks in Chapter 8.

Chapter 2

Acquiring Illumination Based on Omni-Directional Stereo Algorithm

Image-based lighting techniques measure real-world illumination conditions from photographycally acquired images of a real scene. In this chapter, we propose an efficient method for automatically measuring illumination distribution of a real scene by using a set of omni-directional images of the scene taken by a CCD camera with a fisheye lens.

2.1 Introduction and Related Works

The seamless integration of virtual objects with an image of a real scene is an important step toward the long term goal of achieving photographic realism of synthesized images. Techniques for merging virtual objects with a real scene attract a great deal of attention in the fields of both computer graphics and computer vision research.

The synthesized world called augmented reality allows us to see the real scene with virtual objects superimposed onto the scene. In this augmented reality, we are able to handle phenomena not only in the real world but also in a virtual world, while virtual reality technologies immerse a user in a fully computer-generated scene.

To enhance the quality of synthesized images in augmented reality, three aspects have to be taken into account: *geometry*, *illumination*, and *time*. More specifically, the virtual object has to be positioned at a desired location in the real scene, and the object must

appear at the correct location in the image (*consistency of geometry*). Also, shading of the virtual object has to match that of other objects in the scene, and the virtual object must cast a correct shadow, i.e., a shadow whose characteristics are consistent with those of shadows in the real scene (*consistency of illumination*). Lastly, motions of virtual objects and real objects have to be correctly coordinated (*consistency of time*).

Pioneering work in this field was proposed by Fournier et al. [17]. Fournier et al.'s method takes into account not only direct illumination but also indirect illumination by using the radiosity algorithm, which is commonly used for rendering diffuse interreflection [5]. This method is effective for modeling subtle indirect illumination from nearby objects. However, this method requires a user to specify the 3D shapes of all objects in the scene. This object selection process could be tedious and difficult if a scene were full of objects. Also, since this method computes global illumination using pixel values of an input image, it is required that the image have a reasonably wide field of view. Even so, this method cannot model direct illumination from outside of the input image correctly unless a user specifies the positions of all lights.

Drettakis et al. [12] extended Fourier et al.'s work. Drettakis et al.'s method made the creation of the 3D model much easier using computer vision techniques. They also introduced the use of a panoramic image built by image mosaicing to enlarge the field-of-view of the input image, and the use of hierarchical radiosity for efficient computation of global illumination. However, this method still requires a user to define the vertices and topology of all objects in the scene, and it is often the case that the achieved field-of-view is not wide enough to cover all surfaces in the scene. This causes the same limitation on direct illumination outside the input image as in Fournier et al.'s method.

Later, Debevec [7] introduced a framework of constructing a light-based model of a real scene and using it for superimposing virtual objects into the scene with consistent shadings. A light-based model is a radiometric representation of a scene that is constructed by mapping reflections on a spherical mirror placed in the scene onto a geometric model of the scene.[1]

Although this method succeeded in superimposing virtual objects onto an image of a real scene with convincing shadings, it still requires a user's efforts to construct a light-base model: specifying a geometric model of the distant scene, selecting viewing points for

[1] State et al. previously introduced the use of a steel ball to capture the reflections at a single point [63].

observing the mirror so that the reflections on the mirror can cover the entire geometric model of the scene. Also, a sampling frequency of the reflected light differs depending on the 3D geometry of the surface points on the mirror and the camera's viewing directions. This might cause poor sampling of the reflected light in some portions of the constructed light-base model.

To summarise, the previously proposed methods suffered from two difficulties: how to construct a geometric model of the scene, and how to capture a wide field of view of the scene. With regard to the first difficulty, no simple solution has yet been proposed, and its construction still requires user's efforts. With regard to the second difficulty, special equipment such as a spherical mirror or a panoramic image built by mosaicing was proposed. However, multiple images taken from different viewing angles are still necessary to capture the radiance of the entire scene, and therefore the image registration process is also required.

2.2 Overview of Proposed Method

As a solution to these problems, an efficient method for automatically measuring a radiance distribution of a real scene and using it for superimposing virtual objects appropriately onto a real scene is proposed in this chapter.

Our proposed method automatically measures a radiance distribution of a real scene by using a set of omni-directional images taken by a CCD camera with a fisheye lens. There are three reasons why we use omni-directional images rather than images taken by a camera with an ordinary lens. First, because of fisheye lens' wide field of view, e.g., 180 degrees, we can easily capture illumination from all directions from far less number of omni-directional images. Second, since a fisheye lens is designed so that an incoming ray from a particular direction is projected onto a particular point on an imaging plane, we do not have to concern ourselves with computing directions of incoming rays and considering a sampling frequency of the incoming rays. Third, we are also able to use the directions of the incoming rays for automatically constructing a geometric model of the scene with fisheye lens' wide field of view. By using omni-directional images, we are able to overcome the two difficulties that the previously proposed method encountered

Using an omni-directional stereo algorithm, the proposed method first constructs a

geometric model of the scene from a pair of omni-directional images taken from different locations. Then radiance of the scene is computed from a sequence of omni-directional images taken with different shutter speeds and mapped onto the constructed geometric model. We refer to this geometric model with the radiance as a radiance map. The construction of a radiance map is necessary in order to compute a radiance distribution seen from any point in the scene. In other words, without constructing a radiance map, we can determine only the radiance distribution seen from the particular point where the omni-directional image was captured. To overcome this limitation, our method measures the radiance distribution of the scene as a triangular mesh. Once a radiance map is constructed as a triangular mesh, an appropriate radiance distribution can be used for rendering a virtual object and for generating shadows cast by the virtual object onto the real scene wherever the virtual object is placed in the scene.

In this work, we are not concerned with the consistency of time, and only static scenes are considered. We assume that there is no moving object in an image of the real scene, and therefore we do not have to coordinate virtual object motions and real object motions.

The rest of this chapter is organized as follows. Section 2.3 explains how to determine the transformation between the world coordinate system and the image coordinate system. Section 2.4 describes how to measure a radiance distribution of the real scene by using a pair of omni-directional images. Section 2.5 explains how to superimpose virtual objects onto the real scene by using the world-to-image transformation and the measured radiance distribution. Section 2.6 shows experimental results of the proposed method applied to real images of both indoor and outdoor environments. Section 2.7 presents concluding remarks.

2.3 Consistency of Geometry

In this section, we describe how to define the world coordinate system in the real scene, and how to determine the transformation between the 3D coordinate system of the real scene and the 2D coordinate system of an image onto which virtual objects are superimposed.

2.3.1 Definition of World Coordinate System

We take an image of the real scene by using a color CCD camera. Virtual objects are later superimposed onto the input image. Then we place a calibration board with regularly spaced dots, e.g., 81 dots in our experiments, in the scene. Without changing the camera setting, we take another image of the scene so that all dots on the calibration board appear in the calibration image. (For instance, see Figure 2.5 (b).)

Using the calibration board, the world coordinate system is defined in the real scene such that a) the calibration board becomes a plane of $z = 0$, b) the center dot on the calibration board becomes the origin of the world coordinate system, and c) two edges of the calibration board are parallel to the x-axis and the y-axis of the world coordinate system. Once the world coordinate system is defined, we can place virtual objects at arbitrary locations.

2.3.2 Transformation between the World Coordinate System and the Image Coordinate System

After defining the world coordinate system in the real scene, we estimate the transformation between the world coordinate system and the image coordinate system. For this estimation, we use the camera calibration algorithm proposed by Tsai [68], which is known to be able to estimate camera parameters reliably by taking into account various effects causing image distortion, e.g., radial distortion, displacement of the image center, and mismatching between camera and frame-grabber scanning rate.

Tsai's camera model gives a transformation between a 3D world coordinate system and a 2D image coordinate system, e.g., the projection of a 3D point in the scene onto the input image, generating a 3D ray extending from the camera projection center through an image pixel.

Using the transformation between the world coordinate system and the image coordinate system, we can compute where a virtual object appears in the input image once the object is placed in the real scene.

2.4 Consistency of Illumination

In the following subsections, we explain how to measure a radiance distribution of a real scene; this distribution will then be used for rendering virtual objects superimposed onto the real scene.

In Section 2.4.1, we describe an omni-directional stereo algorithm for determining the 3D location of distinct features in the scene such as a fluorescent lamp on a ceiling. After the 3D locations of distinct features are obtained, the whole shape of the scene is approximated as a 3D triangular mesh whose vertices are the obtained features (Section 2.4.2). In Section 2.4.3, we describe how to estimate a scene radiance from an irradiance measured as the brightness of the omni-directional image. Then the brightness of the omni-directional image is used for determining the radiance distribution of the scene whose shape is obtained as a triangular mesh.

2.4.1 Locations of Distinct Features from Omni-directional Stereo

A CCD camera with a fisheye lens[2] is used to take omni-directional images of the real scene. The camera is placed at two known locations in the scene to capture two omni-directional images from different locations.

The imaging system used in our method is illustrated in Figure 2.1. **C1** and **C2** are the camera projection centers at each of the two locations. These two locations, **C1** and **C2**, are known a priori as a user places the camera at these locations. In the following, we denote the omni-directional image taken at **C1** as $FEV1$, and the other one taken at **C2** as $FEV2$.

A real scene contains a very wide range of radiance. Therefore, due to the limited dynamic range of a CCD camera, pixel values of an image taken with one shutter speed cannot measure radiance in the scene accurately. To avoid this problem, multiple images taken with different shutter speeds are combined to produce each omni-directional image with a virtually extended dynamic range [9].

The fisheye lens used in the image acquisition system is designed so that an incoming

[2] 3CCD color camera (Victor KYF-57) and fisheye lens (Fit Corporation FI-19 with field of view of 180 degrees)

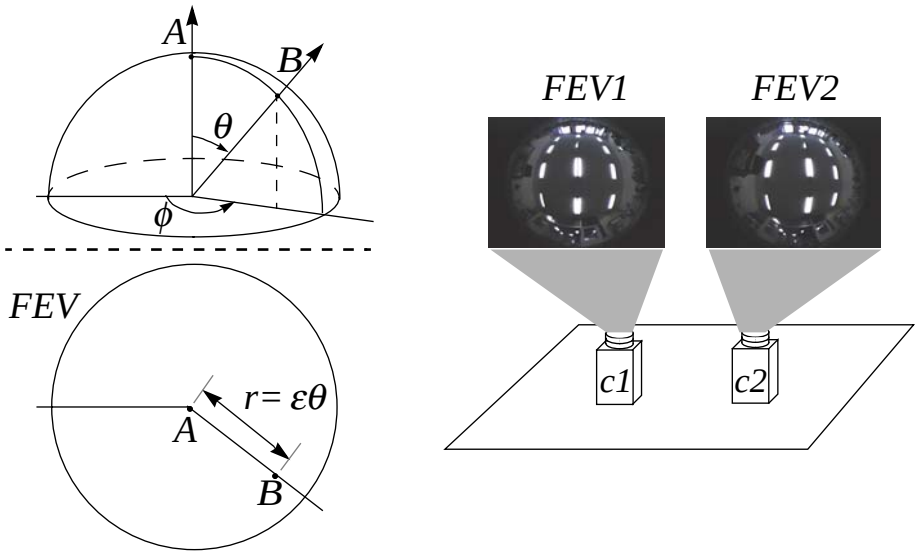


Figure 2.1: Omni-directional image acquisition system: The fisheye lens used in our system is designed so that an incoming ray to the lens is projected onto a particular location in an imaging plane.

ray to the lens is projected onto an imaging plane as

$$r = f\theta \tag{2.1}$$

where f is the focal length of the lens, r is the distance between the image center and the projection of the ray, and θ is an incident angle of the ray (Figure 2.1).

Using this projection model, the incident angle of the ray is given as $\theta = r/f$, where r is determined from the image coordinate of the point corresponding to the ray. For instance, if a direct light source appears as a point in the omni-directional image, the direction from the camera projection center to the light source is determined from the image coordinate of the point in the image.

Most of the incoming light energy in a real scene comes from direct light sources such as a fluorescent lamp or a window to the outside, while the rest of the incoming light energy comes from indirect illumination such as reflection from a wall. For this reason, it is important to know the accurate locations of direct light sources to represent an illumination distribution of a real scene.

Fortunately, direct light sources usually appear as significantly bright points in an omni-directional image. Therefore, it should be relatively easy to identify direct light

sources in the image. In our method, we extract feature points with high contrast in the two omni-directional images $FEV1$ and $FEV2$ by using the feature extraction algorithm proposed by Tomasi and Kanade [66]. In the algorithm, an image pixel with high gradient values in two orthogonal directions, e.g., a corner point, is extracted as a feature point.

After feature points are extracted in $FEV1$ and $FEV2$, 3D coordinates of points in the real scene corresponding to the extracted feature points are determined by using a stereo algorithm [15].

Consider a 3D line that extends from the camera projection center through the extracted feature point on the image plane. The 3D line is given as

$$\mathbf{L1}_i(t) = \mathbf{C_1} + t_i\mathbf{T}_i, \tag{2.2}$$

$$\mathbf{L2}_j(s) = \mathbf{C_2} + s_j\mathbf{S}_j, \qquad (i, j = 1, 2, \cdots, N) \tag{2.3}$$

where N is the number of the extracted feature points, **C1** and **C2** are the camera projection centers, $\mathbf{T_i}$ and $\mathbf{S_j}$ are the directional vectors of the 3D line, and t_i and s_j are scalar values of the line for $FEV1$ and $FEV2$, respectively.

As described in the beginning of this section, **C1** and **C2** are known a priori, and $\mathbf{T_i}$ and $\mathbf{S_j}$ are given as

$$\begin{aligned} \mathbf{T}_i &= [\sin\theta_{Ti}\cos\phi_{Ti}, \sin\theta_{Ti}\sin\phi_{Ti}, \cos\theta_{Ti}], \\ \mathbf{S}_j &= [\sin\theta_{Sj}\cos\phi_{Sj}, \sin\theta_{Sj}\sin\phi_{Sj}, \cos\theta_{Sj}]. \end{aligned} \tag{2.4}$$

Hence, the two lines $\mathbf{L1_i}(t)$ and $\mathbf{L2_j}(s)$ are defined uniquely for the extracted feature points.

Once we have obtained 3D lines corresponding to the extracted feature points in $FEV1$ and $FEV2$, we establish correspondence of the 3D lines between $FEV1$ and $FEV2$. If a feature point i in $FEV1$ and a feature point j in $FEV2$ correspond to the same 3D point in the scene, the 3D lines obtained from the two feature points $L1(t_i)$ and $L2(s_j)$ must intersect. Then a coordinate of the 3D point is determined from the intersection of the two lines.

However, due to various kinds of errors such as an error in $C1$ and $C2$, and an error in $\mathbf{T_i}$ and $\mathbf{S_j}$, the two lines may not intersect. Therefore, we consider that the two lines intersect if the distance between the two lines is sufficiently short. The distance between

the two 3D lines is given by substituting t_i and s_j from the following equations into (2.2) and (2.3) [20].

$$t_i = \frac{\det\{(\mathbf{C_2} - \mathbf{C_1}), \mathbf{S_j}, \mathbf{T_i} \times \mathbf{S_j}\}}{|\mathbf{T_i} \times \mathbf{S_j}|^2} \tag{2.5}$$

$$s_j = \frac{\det\{(\mathbf{C_2} - \mathbf{C_1}), \mathbf{T_i}, \mathbf{T_i} \times \mathbf{S_j}\}}{|\mathbf{T_i} \times \mathbf{S_j}|^2} \tag{2.6}$$

If a 3D line from one omni-directional image, e.g., $FEV1$, has multiple candidates for a matching line from the other omni-directional image, e.g., $FEV2$, we select a matching line in $FEV2$, so that the feature point in $FEV2$ has a pixel coordinate closest to that of the feature point in $FEV1$.

2.4.2 The Entire Shape of the Scene

In the previous section, we described how to determine 3D coordinates of distinct points in the real scene such as a fluorescent lamp on a ceiling and a window to the outside. However, 3D coordinates of the remaining part of the real scene cannot be determined in the same manner. This part includes a wall, a ceiling, and other object surfaces that act as indirect light sources and therefore do not appear as distinct points in omni-directional images.

In our method, 3D coordinates for that part are approximated by using the 3D coordinates of distinct features in the scene. In particular, we generate a 3D triangular mesh by using the distinct feature points. First, we construct a 2D triangular mesh by applying 2D Delaunay triangulation to the extracted feature points in $FEV1$. That determines the connectivity of a 3D triangular mesh whose vertices are the 3D points corresponding to the feature points in $FEV1$. Then, using the connectivity, a 3D triangular mesh is created from the 3D feature points. The obtained triangular mesh approximates an entire shape of the real scene, e.g., the ceiling and walls of a room, which act as direct or indirect light sources.

Note that the fisheye lens we used in our experiments has a field of view of 180 degrees, and therefore we can obtain only the upper half of the scene model from a pair of the omni-directional images. Thus, a horizontal ground plane is assumed and added to the model as a lower half of the model in our experiments. However, if it is necessary, we are able

to obtain the lower half of the model by using another omni-directional image pair which captures the left part of the scene.

2.4.3 Radiance Distribution of the Scene

After the shape of the real scene is obtained as a triangular mesh, the radiance of the scene is estimated by using the brightness of the omni-directional images.

First, we explain the relationship between the radiance at a point on an object and the irradiance at the corresponding point in an omni-directional image. The radiance L of a small patch in the real scene (δO in Figure 2.2) can be computed from the irradiance E_c at a small patch on the imaging plane δI. This irradiance on the imaging plane is measured as a pixel brightness in $FEV1$ or $FEV2$.

As shown in Figure 2.2, consider a small patch on a real object surface δO that acts as a light source, and a small patch δI on an imaging plane. In Figure 2.2, d is the diameter of the fisheye lens, f is the distance from δI to the lens, z is the distance from the lens to δO, α is the angle between the surface normal at δO and the ray from δO to the lens center, θ is the angle between the lens optical axis and the ray from δO to the lens center, and β is the angle between the lens optical axis and the ray from δI to the lens center.

The apparent area of the patch δI as seen from the center of the lens is $\delta I \cos\beta$, while the distance of this patch from the center of the lens is $f/\cos\beta$. Thus, the solid angle subtended by the patch δI is $\delta I \cos\beta/(f/\cos\beta)^2$.

Similarly, the solid angle of the patch δO as seen from the center of the lens is $\delta O \cos\alpha/(z/\cos\theta)^2$.

Since the ratio of the two solid angles is $(\frac{\mathrm{d}\beta}{\mathrm{d}\theta})^2$, we must have

$$\frac{\delta I \cos\beta}{(f/\cos\beta)^2} = (\frac{\mathrm{d}\beta}{\mathrm{d}\theta})^2 \frac{\delta O \cos\alpha}{(z/\cos\theta)^2} \tag{2.7}$$

From the projection model of the fisheye lens, we have $\beta = \tan^{-1}\theta$ and $\frac{\mathrm{d}\beta}{\mathrm{d}\theta} = \frac{\mathrm{d}\tan^{-1}(\theta)}{\mathrm{d}\theta} = \frac{1}{1+\theta^2}$. (2.7) becomes

$$\frac{\delta O}{\delta I} = \frac{\cos^3(\tan^{-1}\theta)}{\cos\alpha\cos^2\theta}(\frac{z}{f})^2(1+\theta^2)^2. \tag{2.8}$$

Next, we need to determine how much of the light emitted by the surface makes its way through the lens. The solid angle subtended by the lens, as seen from the patch δO, is

$$\Omega = \frac{\pi}{4}\frac{d^2\cos\theta}{(z/\cos\theta)^2} = \frac{\pi}{4}(\frac{d}{z})^2\cos^3\theta. \tag{2.9}$$

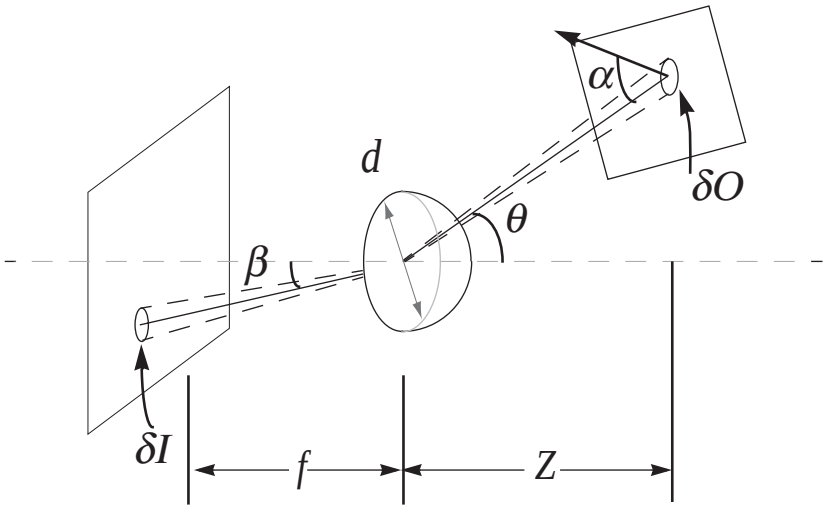


Figure 2.2: Scene radiance and image irradiance.

Thus, the power of the light originating on the patch and passing through the lens is

$$\delta P = L\delta O \cos\alpha\Omega = L\delta O \cos\alpha \frac{\pi}{4}(\frac{d}{z})^2 \cos^3\theta, \tag{2.10}$$

where L is the radiance of the surface in the direction toward the lens.

Since no light from other areas in the scene reaches the patch δI, we have

$$E_c = \frac{\delta P}{\delta I} = L\frac{\delta O}{\delta I}\frac{\pi}{4}(\frac{d}{z})^2 \cos^3\theta\cos\alpha. \tag{2.11}$$

Substituting (2.8) into (2.11), we finally obtain the equation that represents the relationship between the irradiance on the imaging plane and the radiance on the object surface in the scene:

$$E_c = L\frac{\pi}{4}(\frac{d}{f})^2(1+\theta^2)^2\cos^3(\tan^{-1}\theta)\cos\theta. \tag{2.12}$$

However, this equation cannot be used to measure irradiance values in the unit of Wm^{-2} because the equation does not take into account other factors such as D/A and A/D conversions in a CCD camera and a frame grabber. By compensating for $(1+\theta^2)^2\cos^3(\tan^{-1}\theta)\cos\theta$, we can measure values proportional to real irradiance. The scaling factor can be determined by measuring brightness of the same light source, using both a CCD camera and a photometer, a task which was not performed in this work.

Next, we project the omni-directional image $FEV1$ onto the 3D triangular mesh to assign color texture to the mesh. The assigned color texture at a point on the triangular mesh represents the irradiance of the corresponding point in $FEV1$. Therefore, by converting values of the assigned color texture into the radiance L using (2.12), we finally obtain the radiance distribution of the scene.

2.5 Superimposing Virtual Objects onto a Real Scene

In the previous sections, we described how to determine the transformation between the world coordinate system and the image coordinate system (consistency of geometry in Section 2.3), and how to measure a radiance distribution of the real scene (consistency of illumination in Section 2.4). In this section, we explain how to superimpose virtual objects onto the real scene by using the transformation and the measured radiance distribution.

First, we describe how to use the radiance map to compute the total irradiance at a point on a virtual object surface or a point on a real object surface. Then, we explain how to compute the color of a virtual object surface and how to generate the shadow cast by a virtual object using the computed irradiance.

It should be noted that superimposing virtual objects without knowing the shapes and reflectance properties of all nearby objects results in the fact that we cannot model the interreflections between the real objects and the virtual objects. To take the interreflections into account, we need to compute the global illumination in the same manner as in Fournier et al.'s method [17]. However, it is required to define or estimate shapes and reflectance properties of all nearby objects in order to do so. Rather, in this work, we make the rendering process simple, and we consider only the emitted light from the radiance map.

2.5.1 Total Irradiance from Real Illumination

For rendering a surface of a virtual object and for generating the shadow cast by a virtual object, a total irradiance at the surface from the radiance map has to be obtained.

Consider an infinitesimal patch of the extended light source, of size $\delta\theta_i$ in polar angle and $\delta\phi_i$ in azimuth (Figure 2.3). Seen from the center point A, this patch subtends a solid angle $\delta\omega = \sin\theta_i\delta\theta_i\delta\phi_i$. If we let $L(\theta_i, \phi_i)$ be the radiance per unit solid angle coming from the direction (θ_i, ϕ_i), then the radiance from the patch under consideration

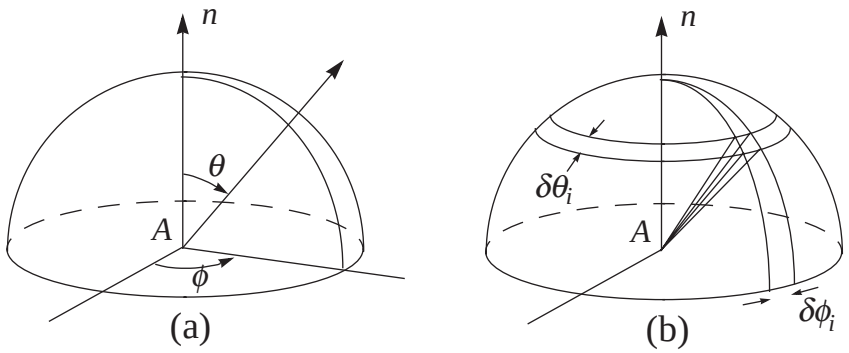


Figure 2.3: (a) the direction of incident and emitted light rays (b) infinitesimal patch of an extended light source

is $L(\theta_i, \phi_i) \sin\theta_i \delta\theta_i \delta\phi_i$, and the total irradiance of the surface is

$$E = \int_{-\pi}^{\pi} \int_{0}^{\frac{\pi}{2}} L(\theta_i, \phi_i) \cos\theta_i \sin\theta_i \mathrm{d}\theta_\mathrm{i} \mathrm{d}\phi_\mathrm{i} \tag{2.13}$$

where the radiance L is measured as shown in (2.12) [23].

To compute the irradiance E, the double integral in (2.13) needs to be approximated by discrete sampling over the entire hemisphere. In our method, nodes of a geodesic dome [3] are used for the discrete sampling. For each of the nodes, the radiance of a corresponding point in the real scene is used as the radiance of the node. Consider a ray from a point on a virtual object surface to the node. A color texture at an intersection of the ray and the 3D triangular mesh described in Section 2.4.2 is obtained as the radiance.

Nodes of a geodesic dome are uniformly distributed over the surface of a sphere. Therefore, by using N nodes of a geodesic dome in a northern hemisphere as a sampling direction, the double integral in (2.13) can be approximated as a sampling at equal solid angle $\delta\omega = 2\pi/\mathrm{N}$. The number of the nodes N can be adjusted by changing the sampling frequency of a geodesic dome.

Using the discrete sampling, (2.13) can be approximated as

$$E = \sum_{i=0}^{N} \frac{2\pi}{N} L(\theta_i, \phi_i) \cos\theta_i. \tag{2.14}$$

Note that a radiance does not depend on the distance between a viewpoint and a light source. Therefore, the distance from the point on the virtual object to the real scene does

not affect the radiance. Also, we assume that the real scene reflects or emits light as a perfect Lambertian plane. That is to say, at each surface point in the scene, light energy is emitted equally in all directions. Due to this assumption, we do not consider directional light sources such as spotlights in our method. However, by combining multiple omni-directional images taken from different locations, our method could be extended to model changes of a radiance, depending on viewing directions.

2.5.2 Ray Casting for Superimposing Virtual Objects

For superimposing virtual objects onto an input image of a real scene, the ray casting algorithm is used as follows.

1. For each pixel in the input image of the real scene, a ray extending from the camera projection center through the pixel is generated by using the transformation between the world coordinate system and the image coordinate system. Then it is determined whether a ray intersects a virtual object or a real object in the scene, e.g., a tabletop.

2. If the ray intersects a virtual object, we consider that the pixel corresponds to a point on the virtual object surface. Then we compute a color to be observed at the surface point under the measured radiance distribution of the real scene using a reflection model. The computed color is stored in the pixel as the surface color of the virtual object at the pixel.

3. Otherwise, we consider that the pixel corresponds to a point on a real object surface. Then we modify an observed color at the point on the real object surface, so that a shadow cast by the virtual object onto the real object is generated.

We are not concerned here with the problem of occlusion because it is beyond the scope of this study to measure accurate 3D shapes of real objects.[3] We assume that a virtual object always exists between the camera projection center and real objects in the scene. In other words, seen from the camera projection center, a virtual object exists only in front of real objects, and a real object never occludes the view of the virtual object. If a more dense 3D depth map is obtainable from the process of constructing a radiance map, we can make full use of the depth map for the problem of occlusion.

[3] Techniques for determining correct occlusion between virtual and real objects using the shapes of the real objects are called Z-key. For instance, see [30].

2.5.3 Rendering Virtual Object Surface

As described in the previous section, if a ray through an image pixel intersects a virtual object, the color at the intersection under the measured radiance distribution is computed, and the color is stored at the image pixel as the color of the virtual object surface.

For computing the color on the virtual object surface, a simplified Torrance-Sparrow reflection model [45], [67] was used in our method. The reflection model for the radiance distribution shown in (2.14) is represented as

$$\begin{aligned} I_m = \quad & K_{d,m} \quad \sum_{i=0}^{N} L_m(\theta_i, \phi_i) S(\theta_i, \phi_i) \cos\theta_i \\ + \quad & K_{s,m} \quad \sum_{i=0}^{N} L_m(\theta_i, \phi_i) S(\theta_i, \phi_i) \frac{1}{\cos\theta_r} e^{\frac{-\gamma(\theta_i,\phi_i)^2}{2\sigma^2}} \end{aligned} \tag{2.15}$$

m = R, G, B

where θ_i is the angle between the surface normal and the direction to each point light source, θ_r is the angle between the surface normal and the viewing direction, $\gamma(\theta_i, \phi_i)$ is the angle between the surface normal and the bisector of the light source direction and the viewing direction, $S(\theta_i, \phi_i)$ are shadow coefficients where $S(\theta_i, \phi_i) = 0$ if other surface point on the virtual object occludes $L(\theta_i, \phi_i)$, and $S(\theta_i, \phi_i) = 1$ otherwise. $K_{d,m}$ and $K_{s,m}$ are constants for the diffuse and specular reflection components, and σ is the standard deviation of a facet slope of the Torrance-Sparrow reflection model. $K_{d,m}$ and $K_{s,m}$ are simply given for virtual objects, or they can be determined for real objects by using Sato et al.'s method [57]. Also, the constant 2π/N in (2.14) is included in the constants $K_{d,m}$ and $K_{s,m}$ in (2.15).

2.5.4 Soft Shadow Cast by a Virtual Object

If a ray through an image pixel does not intersect with a virtual object, the image pixel corresponds to a point on a real object surface. The virtual object may occlude some of incoming light to the point on the object surface. Thus, the color of the image pixel needs to be modified, so that a shadow cast by the virtual object is created on the real object surface.

A shadow cast by a virtual object is created as follows:

1. Obtain a 3D coordinate of a point on a real object surface where a ray through an image pixel intersects the real object. We know the shape of the plane onto which a

virtual object is placed, because a plane of $z = 0$ is defined on the plane as explained in Section 2.3.1. Hence, we compute an intersection between the ray and the plane of $z = 0$, and generate shadows cast on the plane. In this work, we do not deal with shadows cast by virtual objects onto real objects other than the plane. However, our method can be easily extended to handle other real objects if their shapes are obtained from the process of constructing a radiance map or by other means.

2. Compute a total irradiance E_1 at the surface point from the radiance distribution of the real scene. In this case, a virtual object does not occlude incoming light at the surface point (Figure 2.4.a). As described in Section 2.5.1, the total irradiance E_1 can be computed using (2.14).

3. With a virtual object placed in the real scene, compute a total irradiance E_2 at the surface point that is not occluded by the virtual object. In this case, the virtual object occludes some of the irradiance from the real scene (Figure 2.4.b) . As a result, the total irradiance E_2 becomes smaller than the total irradiance E_1. The total irradiance E_2 can be obtained as

$$E_{2,m} = \sum_{i=0}^{N} \frac{2\pi}{N} S(\theta_i, \phi_i) L_m(\theta_i, \phi_i) \cos\theta_i, \qquad m = R, G, B, \tag{2.16}$$

where $S(\theta_i, \phi_i) = 0$ if the virtual object occludes $L(\theta_i, \phi_i)$, and $S(\theta_i, \phi_i) = 1$ otherwise.[4]

4. Compute the ratio of the total radiance E_2 to the total radiance E_1. Then multiply the ratio E_2/E_1 to the color at the intersection between the ray and the plane of $z = 0$. The ratio represents how much of the irradiance at the intersection would still be preserved if the virtual object were placed in the scene. By multiplying the ratio E_2/E_1 to the observed color of the image pixel I_m, we obtain the color I'_m that would be the color of the image pixel if there were a virtual object:

$$I'_m = I_m \frac{E_{2,m}}{E_{1,m}}, \qquad m = R, G, B. \tag{2.17}$$

A similar discussion can be found also in [17].

[4] Alternatively, instead of just blocking the radiance $L(\theta_i, \phi_i)$, the scene radiance of the blocking surface point could be used as a secondary light source with it's own radiance.

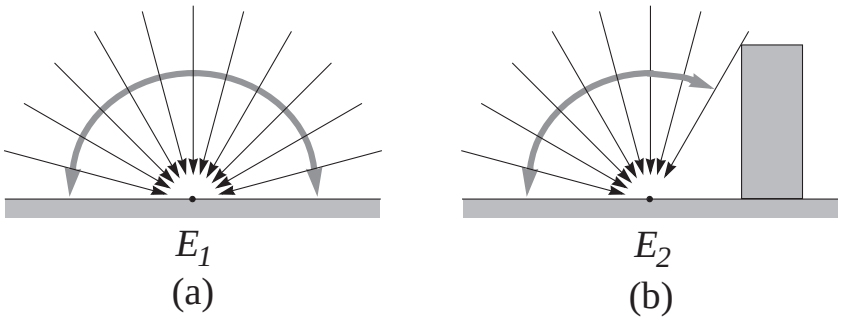


Figure 2.4: Total irradiance (a)without virtual objects (b)with virtual objects

2.6 Experimental Results

We have tested the proposed method by using real images taken in both indoor and outdoor environments. First, we described experimental results for an indoor environment in Section 2.6.1, and then in Section 2.6.2 we presented experimental results for an outdoor environment.

2.6.1 Experimental Results for an Indoor Scene

An image of a tabletop and miscellaneous objects on the tabletop in our laboratory was taken. From the same camera position, another image of the tabletop with a calibration board was taken. The input image and the calibration image are shown in Figure 2.5 (a) and (b).

First, regularly spaced dots on the calibration board were extracted in the calibration image to determine their 2D image coordinates. From pairs of the 2D image coordinates and the 3D world coordinates that were given a priori, the transformation between the world coordinate system and the image coordinate system was estimated by using the camera calibration algorithm as described in Section 2.3.

By using the imaging system illustrated in Figure 2.1, two omni-directional images of the scene, e.g., the ceiling of the laboratory in this experiment, were taken. Figure 2.5 (c) shows the two omni-directional images.

First, feature points were extracted from each of the omni-directional images as described in Section 2.4.1. Then pairs of corresponding feature points in the two omni-directional images were found to determine 3D coordinates of the feature points. Finally,

a triangular mesh was constructed from the feature points to represent the radiance distribution of the scene as explained in Section 2.4.2. However, due to the lack of distinct feature points along the floor in the scene, the obtained triangular mesh did not include vertices along the floor. Therefore, vertices representing the floor were automatically added afterward based on the assumption that the size of the floor was approximately the same as the size of the reconstructed ceiling.

Figure 2.6 shows the obtained triangular mesh that represents the radiance distribution as its color texture, called a radiance map. The reconstructed scene model shows the size of the room within an error of 8 percent. Using the world-to-image transformation and the radiance distribution of the scene, a virtual object was superimposed onto the input image of the scene as explained in Section 2.5. To generate the images in Figure 2.7, 1048 nodes of a geodesic dome were used for sampling the radiance distribution from the scene.[5]

In the images synthesized by our method, shading of the virtual object blends well into the scene. Also, the virtual object casts a shadow with a soft edge on the tabletop in the same way as do the other objects in the scene.

To demonstrate the effectiveness of constructing a radiance map as a geometric model of the scene with its radiance value, we superimposed a metallic hemisphere at three different locations on the tabletop. The synthesized results are shown in Figure 2.8. Using the geometric information of the radiance map, an appropriate radiance distribution is computed at each location of the object and is used for superimposing the object, along with correct shadings, onto the scene . As a consequence, appearance changes are observed on the object surface as it moves on the tabletop.

2.6.2 Experimental Results for an Outdoor Scene

We also applied our method to real images taken in an outdoor environment. The input images used in this experiment are shown in Figure 2.9.

In the same way as with the indoor images, the transformation between the world coordinate system and the image coordinate system was estimated by using the calibration image. Then, the two omni-directional images of the outdoor scene were used to compute the radiance distribution of the scene. The obtained radiance distribution is shown as a color-textured triangular mesh in Figure 2.10.

[5] The model of the mug used in the images was created by using the method proposed by Sato et al.[57].

From this figure, we can see that the shape of the outdoor scene was captured reasonably well. In the radiance distribution as a color-textured mesh, the sky is represented as a planar region extending between the tops of the surrounding buildings because the sky did not contain any distinct feature points to be used as vertices of the triangular mesh.

Using the radiance distribution of the outdoor scene, several virtual objects were superimposed onto the input image of the scene. Figure 2.11 shows the superimposed images. The virtual objects that we used in this experiment are a shiny metallic sphere, a rough metallic dodecahedron, and a yellow matte cube, all of which are floating at the same location in the outdoor scene. We can see that shadings and shadows are represented appropriately in the synthesized images by using our method. In particular. the reflection of the surrounding buildings appears appropriately on the metallic sphere. Unless the real radiance distribution of the scene is used as in our method, a convincing reflection on a virtual object as shown in this figure cannot be created.

2.7 Conclusions

In this chapter we proposed a new method for superimposing virtual objects onto an image of a real scene by taking into account the radiance distribution of the scene. In our method, a camera calibration algorithm is used for matching geometry between virtual objects and the real scene. For matching illumination, the radiance distribution of the real scene is measured by using two omni-directional images of the scene.

Unlike the previously proposed methods, our method can automatically measure an entire radiance distribution of the scene by using multiple omni-directional images. As a result, our method can superimpose virtual objects with convincing shadings and shadows onto the real scene. In addition, we obtain the radiance distribution as a triangular mesh representing the approximate shape of the scene. Therefore, a correct radiance distribution can be used for rendering virtual objects and generating shadows cast by the virtual objects wherever the objects are placed in the real scene.

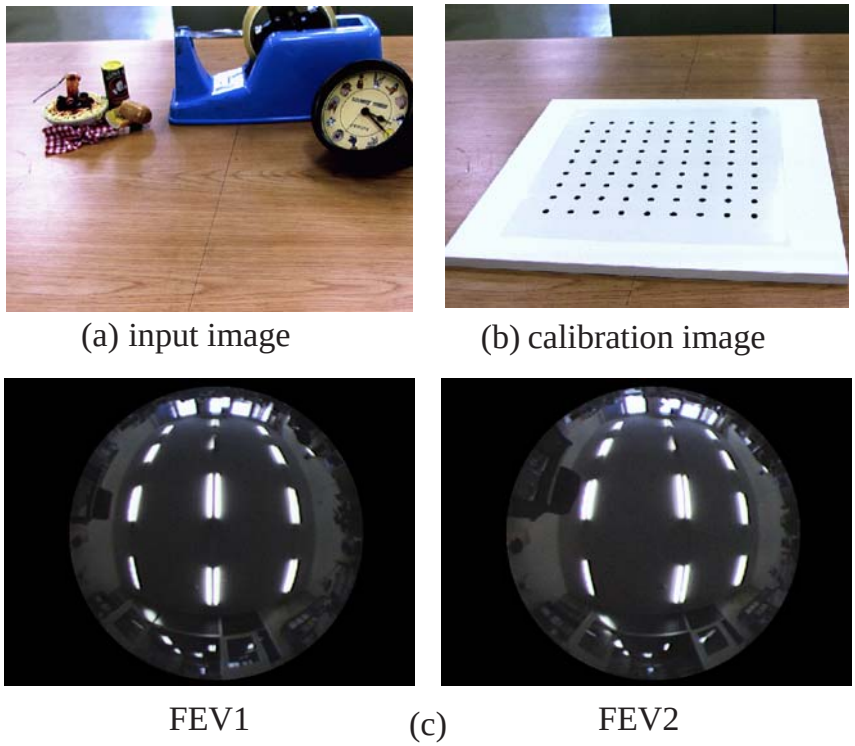


Figure 2.5: (a) input image (b) calibration image (c) omni-directional images

Figure 2.6: Measured radiance distribution.

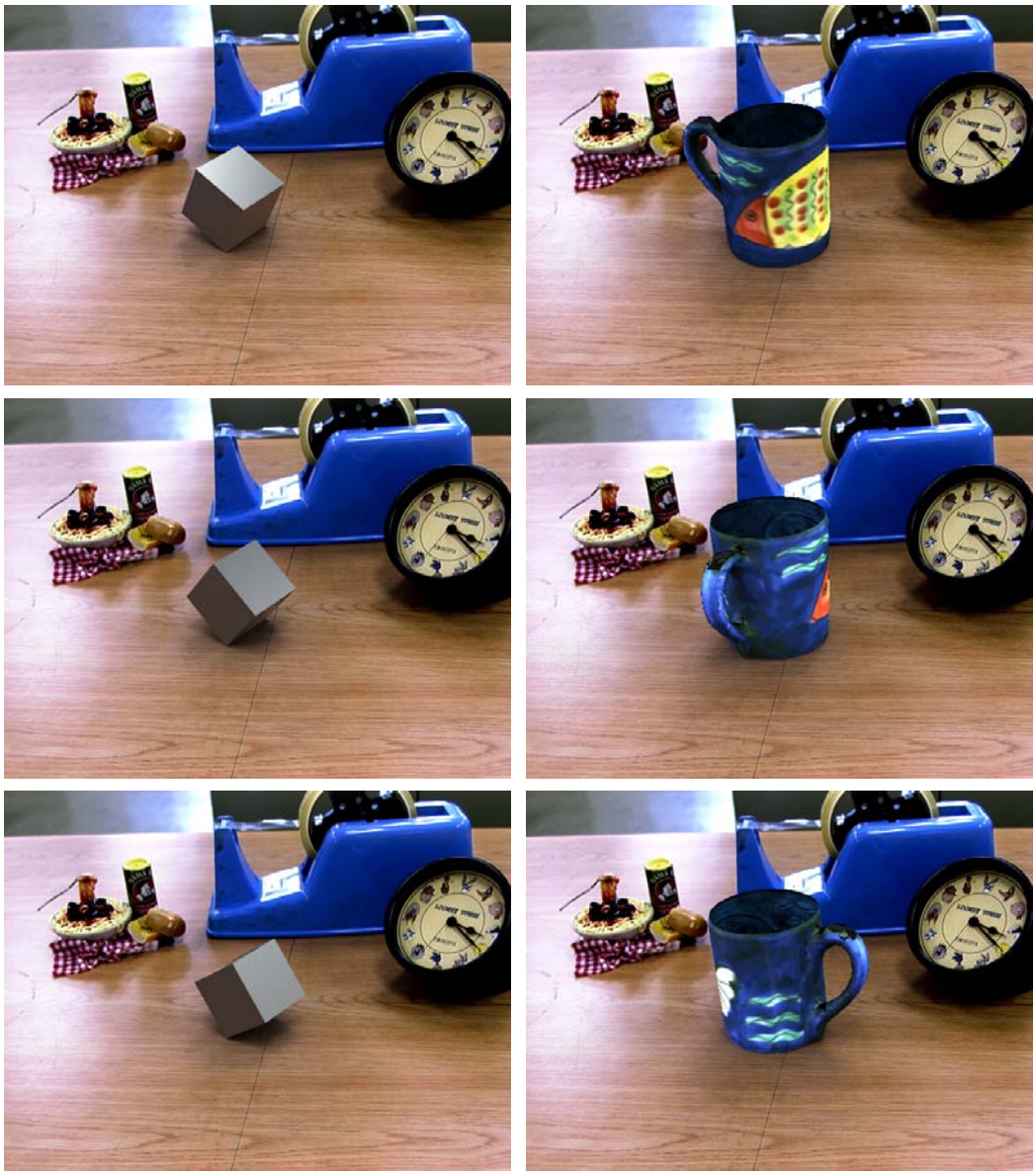


Figure 2.7: Images synthesized with our method.

Figure 2.8: Images synthesized with our method: appearance changes observed on a metallic hemisphere.

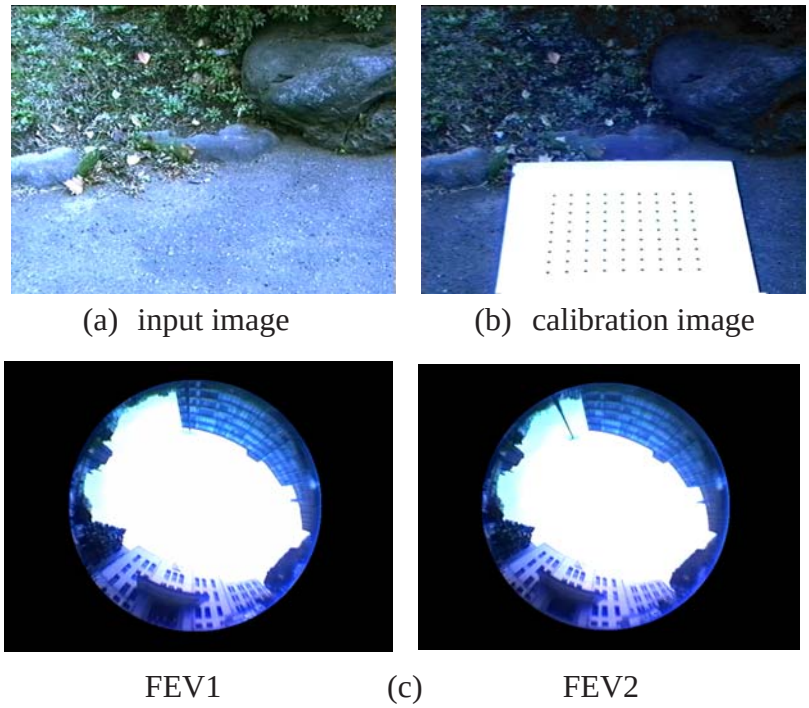


Figure 2.9: (a) input image (b) calibration image (c) omni-directional images

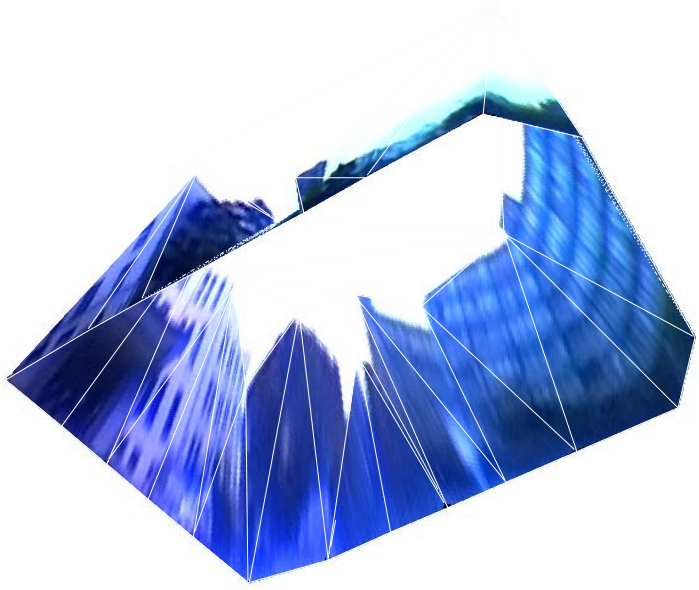


Figure 2.10: Measured radiance distribution.

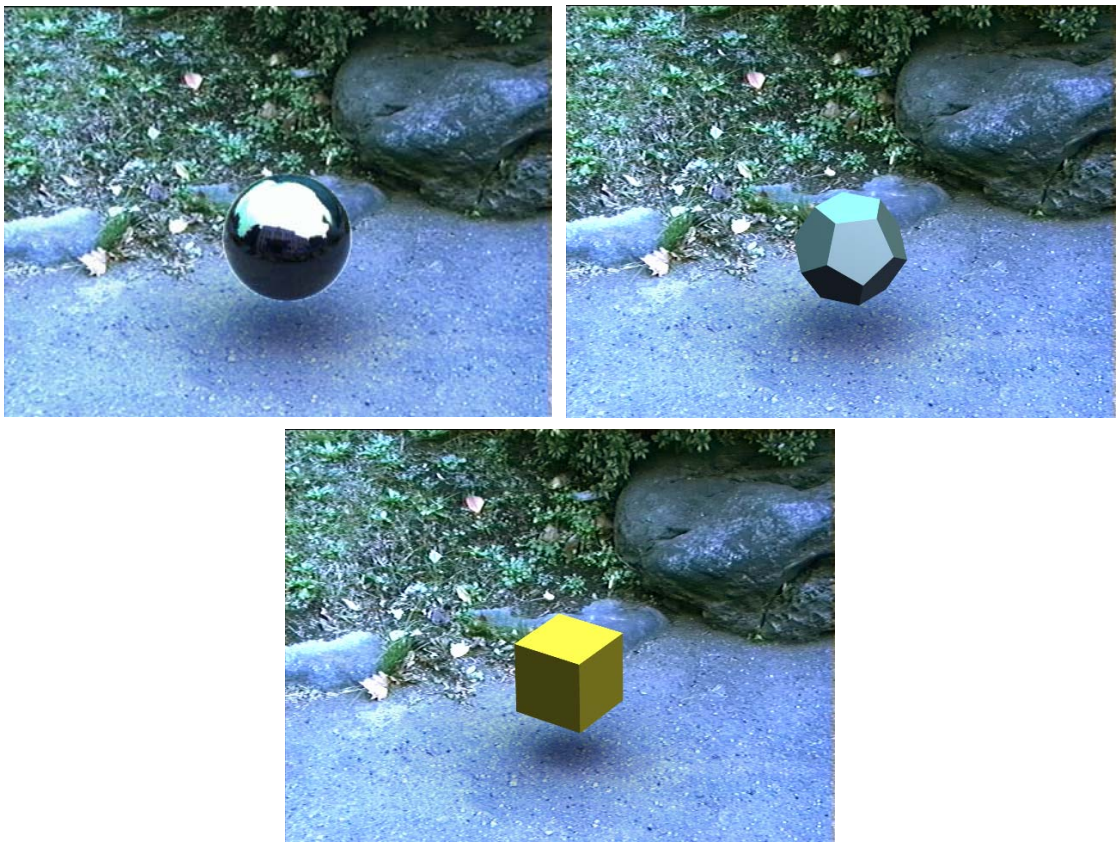


Figure 2.11: Images synthesized with our method.

Chapter 3

Image Synthesis under Dynamically Changing Illumination

In this chapter, we pursue the possibility of real-time image synthesis of virtual objects with natural shading and cast shadows superimposed onto a real scene whose illumination condition is dynamically changing. In general, high computational cost for rendering virtual objects with convincing shading and shadows, such as interreflections or soft shadows under area light sources, prohibits real-time integration of virtual objects into real images.

In order to efficiently re-render a synthetic scene when some aspects of the scene, such as scene geometry and illumination of the scene, is modified, several incremental techniques have been proposed in the field of computer graphics. However, depending on the complexity of a scene, these techniques are still not fast enough to be used for the applications that require real-time image synthesis. As a result, simple rendering algorithms supported by commonly available graphics hardware need to be used for re-rendering the scene in this case, while high qualities in the synthesized images are hardly to be expected.

Alternative approaches have been proposed for re-rendering a scene as a linear combination of a set of pre-rendered basis images of the scene [10, 44, 11]. These approaches are based on the linearity of scene radiance with respect to illumination intensities. Since this linearity holds for scenes with complex geometry and complex photometric properties

such as interreflections between objects and cast shadows, photo-realistic appearance of a scene can be synthesized based on this simple linear combination operation.

Most of the previously proposed methods, however, have been developed for the task of interactive lighting design. Therefore, basis lights under which a set of basis images are rendered are intentionally positioned at the desired locations, so that a scene under desired lighting configurations can be efficiently synthesized.

We generalize the approach based on the linearity of scene radiance with respect to illumination radiance and present an efficient technique for superimposing synthetic objects with natural shadings and cast shadows onto a real scene whose illumination is dynamically changing: we consider that a scene consists of both real objects and synthetic objects with fixed scene geometry, and we synthesize the image of a scene viewed from a fixed viewing point under dynamically changing illumination.

Although additional steps for measuring real-world illumination and taking images of a real scene is required in our case, our method succeeded in superimposing synthetic objects with natural shadings rendered under dynamically changing real-world illumination into a real scene at the frame rate of approximately 3 to 4 frames per second.

3.1 Linearity Between Scene Radiance and Illumination Irradiance

Let us first describe the linearity of the relationship between change of brightness observed on an object surface and change of illumination radiance in a scene in this section.

Consider one case where a light source L_1 illuminates an object (Figure 3.1. (a)) and another case where another light source L_2 placed at a different location illuminates the same object (Figure 3.1. (b)). If we let I_1 be the image observed in the first case and I_2 be the image observed in the second case from the same view point, a novel image I, for which both light sources L_1 and L_2 illuminate the object, can then be synthesized as the sum of I_1 and I_2:

$$I = I_1 + I_2. \tag{3.1}$$

Similarly, we can obtain an image I', which should be observed when the radiance values of the light sources L_1 and L_2 change by considering a linear combination of I_1 and I_2 as

$$I' = r_1 I_1 + r_2 I_2 \tag{3.2}$$

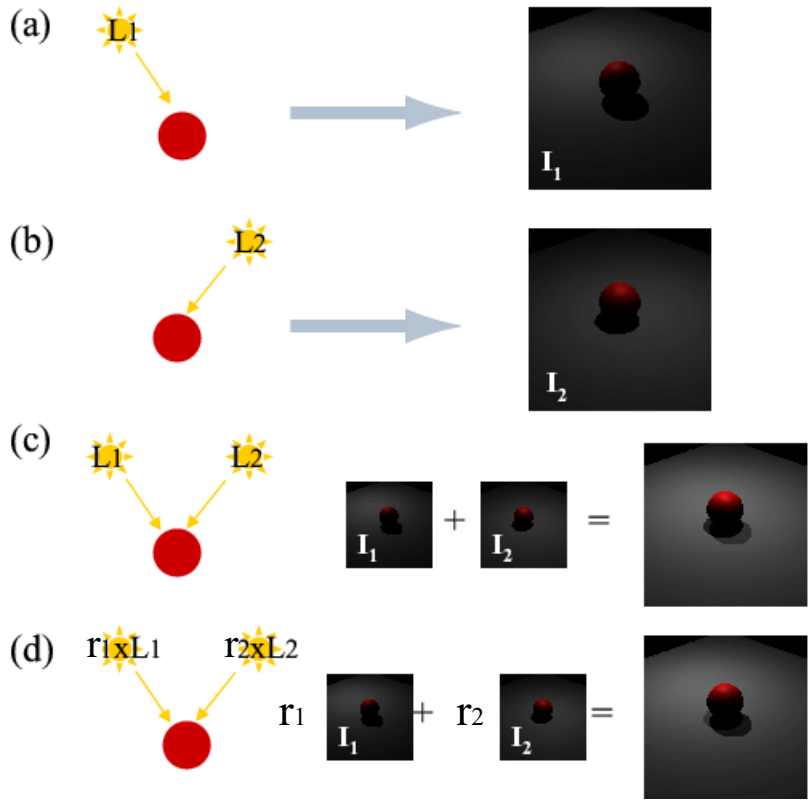


Figure 3.1: Linearity between illumination radiance and scene radiance.

where r_1 and r_2 are scaling factors representing the changes of the radiance values of L_1 and L_2, respectively. For instance, if the radiance of L_1 decreases by half, then r_1 is set to 0.5. It should be noted that this linearity holds not only for direct illumination from lights but also for indirect illumination such as interreflections between or among objects.

Here distant illumination of the scene is assumed; light sources in the scene are sufficiently distant from the objects, and thus all light sources project parallel rays onto the object surface. Namely, the distances from the objects to the light sources are not considered.

3.2 Overview of Proposed Approach

Taking advantage of this linear relationship between brightness observed on an object surface and radiance values of light sources in a scene, the proposed method synthesizes a new image for novel lighting conditions as described in the following steps.

Step1: The entire illumination of a scene is approximated as a collection of area sources $L_i(i = 1, 2, \cdots, n)$ which are equally distributed in the scene. In our current implementation of the proposed method, a set of triangular patches of a geodesic dome are used as a collection of area light sources that approximate the entire illumination distribution (Figure 3.2.1). A set of patches of a geodesic dome are uniformly distributed over the spherical surface, and the number of the patches n can be controlled by changing the sampling frequency of that geodesic dome.

Step2: Two images which are referred to as reference images are rendered under each area light source: one with a virtual object superimposed onto the scene O_i, and the other without the object S_i (Figure 3.2.2).

Step3: Scaling factors of the light source radiance values r_i $(i = 1, 2, \cdots, n)$ are measured by using an omni-directional image of the scene taken by a camera with a fisheye lens (Figure 3.2.3).

Step4: New images $I_o{}'$ and $I_s{}'$, which should be observed under the current illumination condition, are obtained as a linear combination of O_i's and S_i's with the measured scaling factors r_i's, respectively (Figure 3.2.4).

Step5: Using $I_o{}'$ and $I_s{}'$, the virtual object is superimposed onto the image of the scene along with natural shading and shadows that are consistent with those of real objects (Figure 3.2.5). The ray casing algorithm is imposed here; if an image pixel corresponds to the virtual object surface, the color of the corresponding pixels in $I_o{}'$ is assigned as the value of the pixel. Otherwise, the effects on the real objects caused by the virtual object, i.e., shadows and secondary reflection, are added by multiplying the pixel value by the ratio of $I_o{}'$ to $I_s{}'$(Figure 3.2.5).

The main advantage of the proposed method is that composite images are synthesized by simple linear operations based on reference images pre-rendered as an off-line process. From this the quality of output images is not affected by the real-time processing requirement on-the-fly. In the following, section 3.3 explains how to prepare the reference images, and section 3.4 describes a method for measuring an illumination distribution of

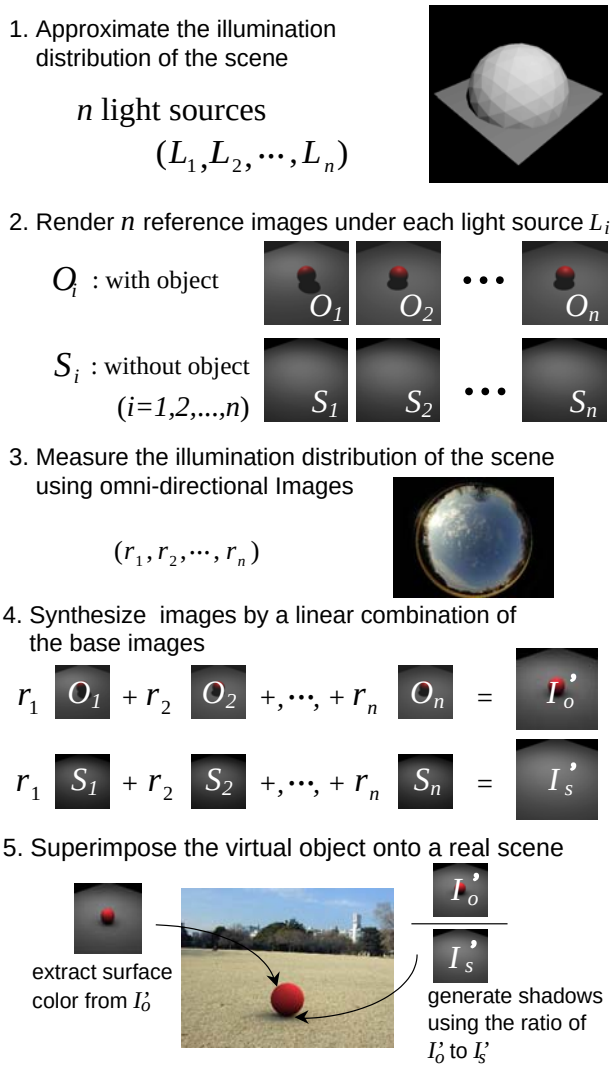


Figure 3.2: Basic steps of the proposed method.

a real scene to determine the radiance scaling factor r_i's. Then Section 3.5 explains the entire process of superimposing a virtual object onto the real scene by using our proposed method.

3.3 Rendering Reference Images

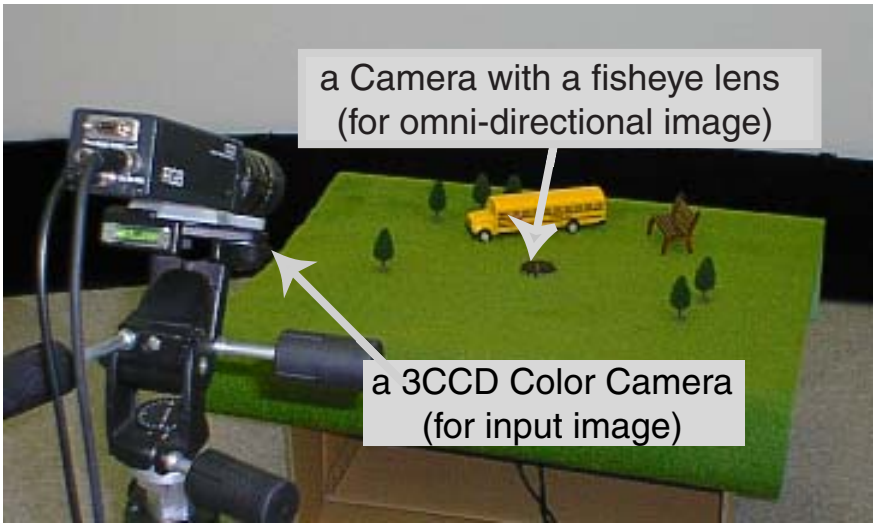


Figure 3.3: Image acquisition system: A 3CCD color camera is used for taking an image of the scene, and another color camera with a fish-eye lens is used for capturing illumination of the scene.

The image acquisition system used in our experiment is shown in Figure 3.3.[1] A 3CCD color camera (*Camera A*) is used for taking an image of a real scene onto which a virtual object is superimposed. In addition, another color camera with a fish-eye lens (*Camera B*) is used for taking an onmi-directional image of the scene which is then used for determining the illumination distribution of the scene.

First, we estimate the camera parameters of *Camera A* by using Tsai's camera calibration method [68] in the same manner described in Section 2.3.2. This camera calibration is necessary for making virtual objects appear at a desired position in final composite images, i.e., consistency of geometry.

Then reference images are rendered using the rendering software called *Radiance*, which can accurately simulate reflections of lights based on physical models of lights [76]. Using Radiance, the following two reference images are rendered off-line for each area light source L_i.

[1] A similar image acquisition set-up was used in [35] to superimpose synthetic objects onto a real scene under real-world illumination.

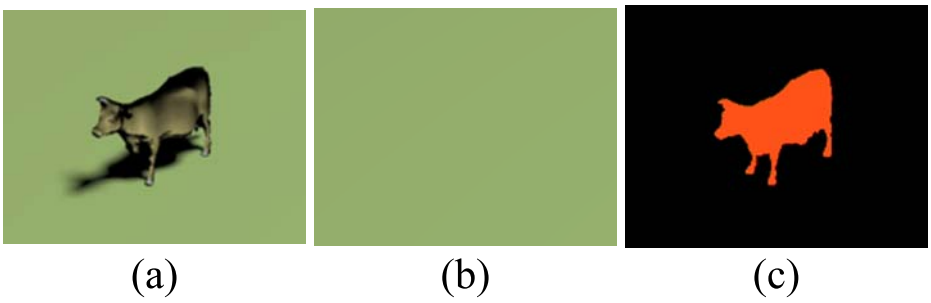


Figure 3.4: Reference image: (a)*shadow image*, (b) *surface image*, (c)*mask image*

1. *Object image* O_i are images of a scene that contains a virtual object (Figure 3.4.a).
2. *Surface image* S_i are images of the same scene without the virtual object (Figure 3.4.b).

For rendering those reference images, we need to know geometric and photometric models of the real scene, i.e., shape and reflectance properties of real objects near a virtual object. The recovery of geometric and photometric model of a real object is not an easy task, and it has been intensively studied in the past. Since geometric and photometric modeling of a real object is not the main focus of our study, those models used for our proposed method are manually specified by a user.

In addition to the *object image* and the *surface image*, another image which we call the *mask image* is prepared. The *mask image* is used for determining whether a ray extending from the camera projection center *Camera A* through each image pixel intersects a virtual object or a real object in the scene.

3.4 Measuring Illumination Distribution

Next, we describe how to measure radiance scaling factors r_i which represent an illumination distribution of a scene by using an omni-directional image taken by a camera with a fisheye lens shown in Figure 3.3. The same camera as used in Section 3.4 (3 CCD color camera Victor KYF-57 and fisheye lens (Fit Corporation FI-19 with field of view of 180 degree) is reused here for measuring scaling factors r_i.

The fisheye lens used in our system is designed in such a way that an incoming ray to the lens is projected onto a certain location in an omni-directional image according to a

desired projection model. Therefore, for each area light source, projection of three vertices of the area light source patch can be determined based on this projection model of the lens. Then the radiance scaling factor for the area light source is computed by using the average pixel values inside the projection of the area light source in the omni-directional image.

Note that we also need to determine an additional scaling factor between a unit radiance value used for rendering the reference images and radiance scaling factor r_i that are measured from the omni-directional image as described above. In our experiment in Section 3.6, this scaling factor is manually specified so that surface color of a virtual object which is rendered using measured radiance scaling factors becomes a certain value.

In order to automatically determine this scaling factor, it is necessary to perform a photometric calibration between unit radiance value used for rendering reference images and illumination radiance measured from an omni-directional image.

3.5 Superimposing Virtual Objects onto a Real Scene

After the radiance scaling factors r_i are determined, the virtual object is superimposed onto the real scene with consistent shading under the captured illumination condition. Note that each of the three color bands (R, G, B) of a composite image is separately synthesized using the corresponding band of the reference images and the scaling factor r_i's. For simplicity of discussion, we do not describe color bands explicitly in this section.

First, a new *object image* and a new *surface image* that should be observed in the measured illumination condition are synthesized as a linear combination of the corresponding reference images O_i's and S_i's, respectively. Let the new *object image* be $I_o{}'$ and the new *surface image* be $I_s{}'$; then those images are obtained as

$$I_o{}' = \sum_{i=1}^{n} r_i O_i \tag{3.3}$$

$$I_s{}' = \sum_{i=1}^{n} r_i S_i. \tag{3.4}$$

Then, for each image pixel in the composite image, we determine whether the pixel corresponds to the virtual object or to a real object in the scene by using the previously prepared *mask image*. If a pixel corresponds to a point on the virtual object surface, a

color to be observed at this point is extracted from the same pixel coordinate of the $I_o{}'$ and stored in the composite image.

Otherwise, we modify an observed color at the point on the real object surface, so that effects caused by the virtual object onto the real object are represented. For this purpose, the ratio of $I_o{}'$ to $I_s{}'$ is first computed. The ratio represents how much of the color at the real object surface should be modified if the virtual object was placed in the scene. Let P the observed color of the image pixel of *camera A*; then the color P' that should be observed if there were a virtual object in the scene is computed by using the ratio as

$$P' = P\frac{I_o{}'}{I_s{}'}. \tag{3.5}$$

3.6 Experimental Results

We have tested the proposed method by using real images of an indoor environment as shown in Figure 3.3. In this experiment, SGI Onyx2 with 6 CPUs was used for capturing a sequence of input images and for synthesizing final composite images. For efficient generation of the composite images, the following two strategies were employed.

1. As described in Section3.5, each pixel in the composite image requires a different type of operation, depending on which object surface the pixel corresponds to. Computational cost of synthesizing each image pixel in the composite image is taken into account for distributing computation evenly among all of the 6 CPUs.

2. Because effects due to the light sources with low radiance values are negligible in the final composite image, the reference images rendered under light sources whose radiance values are lower than a prefixed threshold value are omitted when a linear combination of the reference images is computed. In this way, we can reduce the number of reference images used for synthesizing final composite images, and achieve required processing time by adjusting the threshold value.

Several examples of synthesized composite images are shown in Figure 3.5. In those composite images, appearance of the virtual object blends well into the scene, and the virtual object casts a shadow with a soft edge on the grass in the same way as the other objects such as trees in the scene do. Table 3.1 shows changes in processing time due to

Figure 3.5: Synthesized results.

the number of reference images. [2]

Since the composite images are synthesized not by rendering the whole scene over and over but rather based on the linear combination of pre-rendered reference images, the processing time for synthesizing the composite images never affects the achievable image quality.

In the examples shown in Figure 3.5, we approximated the real illumination distribution by using 400 light sources and selected those light sources according to a threshold value. In this case, in spite of the highly realistic shading and shadows achieved in the composite image, the virtual object can be superimposed onto the scene at the frame rate of approximately 3 to 4 frames per second by using our current implementation of the proposed method.

[2] As described before, the number of reference images is equal to the number of area light sources that approximate the entire illumination distribution of the scene.

Table 3.1: Processing time due to the number of reference images.

num of reference images	processing time (sec)
6	0.08
25	0.2
90	0.33
190	0.55
400	1.2

In the experiment shown in Figure 3.6, the 3D shapes and reflectance properties of the surrounding real objects were modeled by using a photomodeling tool called *Canoma*; the white hemisphere placed in the center of these images is a synthetic object, and the other blocks are real objects. In the images synthesized by our proposed technique, the brightness of the virtual object changes in a similar manner as the other blocks due to changes in the real-world illumination distribution. In the resulting image on the right, the shadows of the real objects fall naturally on the virtual object. In the synthetic image on the left, interreflections are observed between the virtual object and the cube-shaped block.

3.7 Discussion

There is one limitation in the approaches based on the linearity of scene radiance with respect to illumination radiance. This limitation is that an image of the scene under novel illumination conditions cannot be synthesized accurately if the illumination condition lies in the linear subspace spanned by the basis lights under which basis images are rendered. Therefore, it is essential to provide basis lights and their associated weights properly so that images under desirable illumination conditions can be synthesized.

Previous studies investigated how to choose a set of basis lights so that these lights could efficiently represent specific lighting configurations. For instance, Nimeoff et al. introduced a set of steerable area light sources as basis lights to approximate the illumination effect of daylight [44]. Dobashi et al. defined a set of basis lights to represent

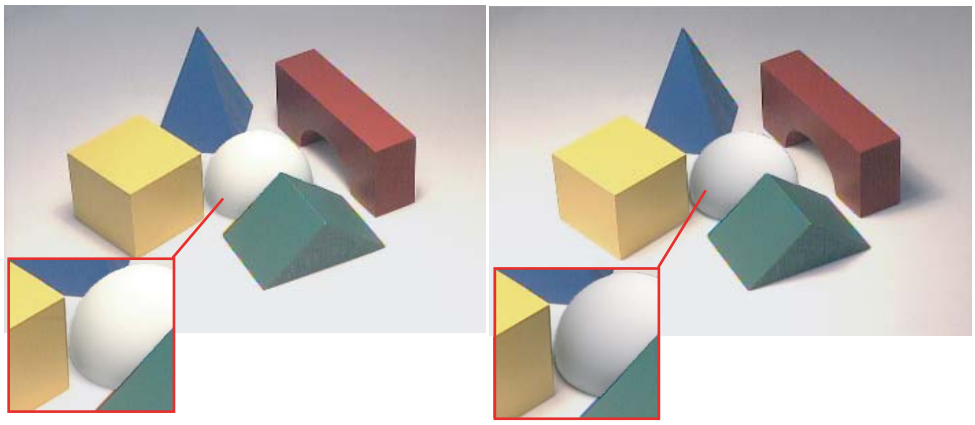


Figure 3.6: The 3D shape and reflection properties of the real objects were model and used for rendering basis images. The shadows of the real objects fall naturally on the virtual object, and the interreflections are observed between the virtual object and the real cube-shaped block.

directionality of spotlights based on spherical harmonics [10]. Later, Teo et al. introduced a hybrid method for synthesizing illumination effects from both area sources and directional spotlight sources [64]. In addition, this method adopted a strategy for reducing the number of basis images based on principal components analysis.

In terms of generating cast shadows with distinct boundaries, Naemura et al. introduced a way of adjusting interpolated weights of predefined basis point sources [43]. Also in the context of synthesizing a photorealistic appearance of human skin, Debevec et al. introduced a method for re-rendering a human face based on a linear combination of face images taken under densely sampled incident illumination directions in [8]. This method further considered a model of skin reflectance to estimate the appearance of the face seen from novel viewing directions and under arbitrary illumination.

Recently, the effect of defining a set of basis lights in the frequency domain based on spherical harmonics was demonstrated in [52, 53, 2]. Later in Chapter 6, we will reconsider this issue of efficiently synthesizing images of a scene under arbitrary illumination conditions and present a novel method for obtaining a set of basis images of real objects.

3.8 Conclusions

In this chapter, we proposed a new method for superimposing virtual objects with natural shading and cast shadows onto a real scene whose illumination condition was dynamically changing. Based on linearity of the relationship between change of brightness observed on an object surface and change of illumination radiance values in a scene, final composite images are synthesized as a linear combination of pre-rendered reference images of virtual objects placed in the scene in our proposed method. The main advantage of our proposed method is that rendering algorithms that are computationally too expensive for real-time processing can be used for rendering reference images since the rendering process is done off-line. Thus the image quality is not affected by the requirement for real-time processing, and the proposed method is able to satisfy both high image quality and real-time processing.

Chapter 4

Illumination from Shadows

There has been another approach called *inverse rendering* in image-based modeling that deals with a inverse problem of traditional model-based rendering. In inverse rendering, techniques that recover unknown illumination conditions of a scene based on the knowledge of the shape and the surface reflectance properties of objects located in that scene is called *inverse lighting*. One of the main advantages of inverse lighting over the former image-based lighting is that it does not require additional images for capturing illumination of a scene, but uses the appearance of objects located in a scene instead for recovering an illumination distribution of the scene.

In a natural illumination condition, a scene includes both direct and indirect illumination distributed in a complex way, and it is often difficult to recover an illumination distribution from image brightness observed on an object surface. The main reason for this difficulty is that there is usually not adequate variation in the image brightness observed on the object surface to reflect the subtle characteristics of the entire illumination.

In this caper, we demonstrate the effectiveness of using occluding information of incoming light in estimating an illumination distribution of a scene. Shadows in a real scene are caused by the occlusion of incoming light, and thus analyzing the relationships between the image brightness and the occlusions of incoming light enables us to reliably estimate an illumination distribution of a scene even in a complex illumination environment.

4.1 Introduction

The image brightness of a three-dimensional object is the function of the following three components: the distribution of light sources, the shape of a real object; and the reflectance of that real object surface [24, 23]. The relationship among them has provided three major research areas in physics-based vision:

- shape-from-brightness (with a known reflectance and illumination)
- reflectance-from-brightness (with a known shape and illumination)
- illumination-from-brightness (with a known shape and reflectance).

The first two kinds of analyses, shape-from-brightness and reflectance-from-brightness, have been intensively studied using the shape from shading method [25, 27, 26, 50], as well as through reflectance analysis research [28, 45, 22, 1, 32, 37, 57, 71].

In contrast, relatively limited amounts of research have been conducted in the third area, illumination-from-brightness. This is because real scenes usually include both direct and indirect illumination distributed in a complex way and it is difficult to analyze characteristics of the illumination distribution of the scene from image brightness. Most of the previously proposed approaches were conducted under very specific illumination conditions, e.g., there were several point light sources in the scene, and those approaches were difficult to be extended for more natural illumination conditions [26, 28, 57, 65, 74], or multiple input images taken from different viewing angles were necessary [34, 47].

In this study, we present a method for recovering an illumination distribution of a scene from image brightness observed on an object surface in that scene. To compensate for the insufficient information for the estimation, we propose to use occluding information of incoming light caused by an object in that scene as well as the observed image brightness on an object surface. More specifically, the proposed method recovers the illumination distribution of the scene from image brightness inside shadows cast by an object of known shape in the scene.

4.2 Overview of Proposed Approach

Shadows in a real scene are caused by the occlusion of incoming light, and thus contain various pieces of information about the illumination of the scene. However, shadows have

been used for determining the 3D shapes and orientations of an object which casts shadows onto the scene [38, 33, 60, 4], while very few studies have focused on the the illuminant information which shadows could provide. In our method, image brightness inside shadows is effectively used for providing distinct clues to estimate an illumination distribution.

The proposed approach estimates an illumination distribution of a scene using the following procedures. The illumination distribution of a scene is first approximated by discrete sampling of an extended light source; whole distribution is represented as a set of point sources equally distributed in the scene. Then this approximation leads each image pixel inside shadows to provide a linear equation with radiance of those light sources as unknowns. Finally, the unknown radiance of each light source is solved from the obtained set of equations. In this paper, we refer to the image with shadows as the *shadow image*, to the object which casts shadows onto the scene as the *occluding object*, and to the surface onto which the occluding object casts shadows as the *shadow surface*.

The assumptions that we made for the proposed approach are as follows.

- Known geometry: the 3D shapes of both the occluding object and the *occluding surface* are known as well as their relative poses and locations.

- Distant illumination: we assume that light sources in the scene are sufficiently distant from the objects, and thus all light sources project parallel rays onto the object surface. Namely, the distances from the objects to the light sources are not considered in the proposed approach.

- No interreflection: the proposed method does not take into account interreflection between the shadow surface and the occluding object on the assumption that there is no severe interreflection between them and no multiple scattering of the interreflected rays from them to the scene. As a consequence, objects with darker color and weaker specularity are preferred as *occluding objects* since they do not significantly act as secondary light sources, and the method cannot handle a scene consisting of shiny occluding objects.

This study further concerns the following two issues that need to be considered when we tackle this illumination-from-brightness problem. The first issue is how to provide reflectance properties of the shadow surface in cases where they are not given a priori. Since it is a common situation that reflectance properties of a surface are not known, a

solution to this problem is required to enlarge the variety of images to which the proposed method can be applied. In our approach, instead of assuming any particular reflectance properties of the surface inside shadows, both the illumination distribution of the scene and the reflectance properties of the surface are estimated simultaneously, based on an iterative optimization framework.

The second issue is how to efficiently estimate an illumination distribution of a scene without losing the accuracy of the estimated illumination distribution. To capture the subtle characteristics of the illumination distribution of a scene, the entire distribution must be discretized densely, and this makes the solution exceedingly expensive in terms of processing time and storage requirements. To satisfy both the accuracy and the efficiency claims, we introduce an adaptive sampling framework to sample the illumination distribution of a scene. Using this adaptive sampling framework, we are able to avoid unnecessarily dense sampling of the illumination and estimate overall illumination more efficiently by using fewer sampling directions.

The rest of this chapter is organized as follows. We describe how we first obtain a formula which relates an illumination distribution of a scene with the image irradiance of the shadow image in Section 4.3 and Section 4.4. The formula will later be used as a basis for estimating both the illumination distribution of a real scene and the reflectance properties of the shadow surface. Using this formula, we explain how to estimate an illumination radiance distribution from the observed image irradiance of a shadow image in Section 4.5. In this section, we consider the following two cases separately: (1) where the reflectance properties of the shadow surface are known, (2) where those properties are not known. We also introduce an adaptive sampling framework for efficient approximation of the entire illumination, and show experimental results of the proposed method applied to images taken in natural illumination conditions in Section 4.6. Finally, we discuss some of the related work proposed in the field of computer graphics in Section 4.7, and present concluding remarks and future research directions in Section 4.8.

4.3 Relating Illumination Radiance with Image Irradiance

In this section, we describe how we obtain a formula which relates an illumination distribution of a real scene with the image irradiance of a shadow image. The formula will

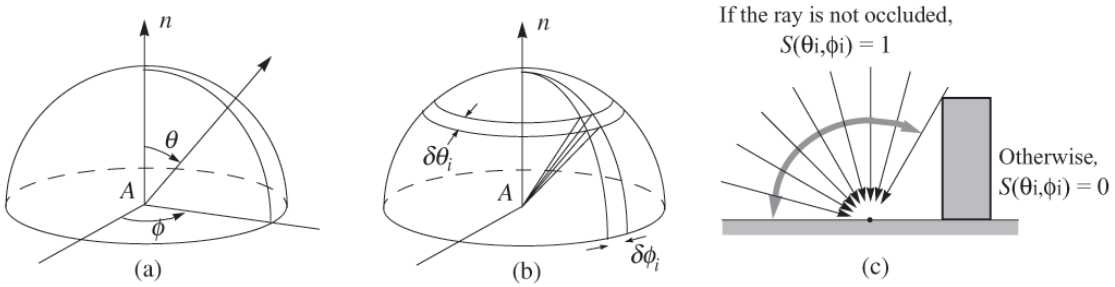


Figure 4.1: (a) the direction of incident and emitted light rays (b) infinitesimal patch of an extended light source, (c) occlusion of incoming light

later be used as a basis for estimating the illumination distribution of a scene and the reflectance properties of a shadow surface.

First, we find a relationship between the illumination distribution of a real scene and the irradiance at a surface point in the scene.[1] To take illumination from all directions into account, let us consider an infinitesimal patch of the extended light source, of a size $\delta\theta_i$ in polar angle and $\delta\phi_i$ in azimuth as shown in Figure 4.1.

Seen from the center point A, this patch subtends a solid angle $\delta\omega = sin\theta_i\delta\theta_i\delta\phi_i$. Let $L_0(\theta_i, \phi_i)$ be the illumination radiance per unit solid angle coming from the direction (θ_i, ϕ_i); then the radiance from the patch is $L_0(\theta_i, \phi_i)sin\theta_i\delta\theta_i\delta\phi_i$, and the total irradiance of the surface point A is [23]

$$E = \int_{-\pi}^{\pi}\int_{0}^{\frac{\pi}{2}} L_0(\theta_i, \phi_i)\cos\theta_i \sin\theta_i \mathrm{d}\theta_\mathrm{i}\mathrm{d}\phi_\mathrm{i}. \tag{4.1}$$

Then occlusion of the incoming light by the *occluding object* is considered as

$$E = \int_{-\pi}^{\pi}\int_{0}^{\frac{\pi}{2}} L_0(\theta_i, \phi_i)S(\theta_i, \phi_i)\cos\theta_i \sin\theta_i \mathrm{d}\theta_\mathrm{i}\mathrm{d}\phi_\mathrm{i}, \tag{4.2}$$

where $S(\theta_i, \phi_i)$ are occlusion coefficients; $S(\theta_i, \phi_i) = 0$ if $L_0(\theta_i, \phi_i)$ is occluded by the *occluding object*; Otherwise $S(\theta_i, \phi_i) = 1$ (Figure 4.1 (c)).

Some of the incoming light at point A is reflected toward the image plane. As a result, point A becomes a secondary light source with scene radiance.

[1] For a good reference of the radiometric properties of light in a space, see [41].

The bidirectional reflectance distribution function (BRDF) $f(\theta_i, \phi_i; \theta_e, \phi_e)$ is defined as a ratio of the radiance of a surface as viewed from the direction (θ_e, ϕ_e) to the irradiance resulting from illumination from the direction (θ_i, ϕ_i). Thus, by integrating the product of the BRDF and the illumination radiance over the entire hemisphere, the scene radiance $Rs(\theta_e, \phi_e)$ viewed from the direction (θ_e, ϕ_e) is computed as

$$Rs(\theta_e, \phi_e) = \int_{-\pi}^{\pi} \int_{0}^{\frac{\pi}{2}} f(\theta_i, \phi_i; \theta_e, \phi_e) L_0(\theta_i, \phi_i) S(\theta_i, \phi_i) \cos\theta_i \sin\theta_i \mathrm{d}\theta_\mathrm{i} \mathrm{d}\phi_\mathrm{i}. \tag{4.3}$$

Finally, the illumination radiance of the scene is related with image irradiance on the image plane. Since what we actually observe is not image irradiance on the image plane, but rather a recorded pixel value in a shadow image, it is necessary to consider the conversion of the image irradiance into a pixel value of a corresponding point in the image. This conversion includes several factors such as D/A and A/D conversions in a CCD camera and a frame grabber.

Other studies concluded that image irradiance was proportional to scene radiance [23]. In our work, we calibrate a linearity of the CCD camera by using a Macbeth color chart with known reflectivity so that the recorded pixel values also become proportional to the scene radiance of the surface. From (4.3) the pixel value of the shadow image $P(\theta_e, \phi_e)$ is thus computed as

$$P(\theta_e, \phi_e) = k \int_{-\pi}^{\pi} \int_{0}^{\frac{\pi}{2}} f(\theta_i, \phi_i; \theta_e, \phi_e) L_0(\theta_i, \phi_i) S(\theta_i, \phi_i) \cos\theta_i \sin\theta_i \mathrm{d}\theta_\mathrm{i} \mathrm{d}\phi_\mathrm{i}, \tag{4.4}$$

where k is a scaling factor between scene radiance and a pixel value. Due to the scaling factor k, we are able to estimate unknown $L_0(\theta_i, \phi_i)(i = 1, 2, \cdots, n)$ up to scale. To obtain the scale factor k, we need to perform photometric calibration between pixel intensity and the physical unit (watt/m^2) for the irradiance.

4.4 Approximation of Illumination Distribution with Discrete Sampling

In our implementation of the proposed method, to solve for the unknown radiance $L_0(\theta_i, \phi_i)$ which is continuously distributed on the surface of the extended light source from the recorded pixel values of the shadow surface, the illumination distribution is approximated by discrete sampling of radiance over the entire surface of the extended light source. This

can be considered as representing the illumination distribution of the scene by using a collection of light sources with an equal solid angle. As a result, the double integral in (4.4) is approximated as

$$P(\theta_e, \phi_e) = \sum_{i=1}^{n} f(\theta_i, \phi_i; \theta_e, \phi_e) L(\theta_i, \phi_i) S(\theta_i, \phi_i) \cos\theta_i \omega_i, \tag{4.5}$$

where n is the number of sampling directions, $L(\theta_i, \phi_i)$ is the illumination radiance per unit solid angle coming from the direction (θ_i, ϕ_i), which also includes the scaling factor k between scene radiance and a pixel value, and ω_i is a solid angle for the sampling direction (θ_i, ϕ_i).

For instance, node directions of a geodesic dome can be used for uniform sampling of the illumination distribution. By using n nodes of a geodesic dome in a northern hemisphere as a sampling direction, the illumination distribution of the scene is approximated as a collection of directional light sources distributed with an equal solid angle $\omega = 2\pi/n$.

4.5 Estimation of Radiance Distribution based on Reflectance Properties of Shadow Surface

After obtaining the formula that relates the illumination radiance of the scene with the pixel values of the shadow image, illumination radiance is estimated based on the recorded pixel values of the shadow image.

In (4.5), the recorded pixel value $P(\theta_e, \phi_e)$ is computed as a function of the illumination radiance $L(\theta_i, \phi_i)$ and the BRDF $f(\theta_i, \phi_i; \theta_e, \phi_e)$. Accordingly, in the following sections, we take different approaches, depending on whether BRDF of the surface is given.

4.5.1 Known Reflectance Properties: Lambertian

Let us start with the simplest case where the shadow surface is a Lambertian surface whose reflectance properties are given.

BRDF $f(\theta_i, \phi_i; \theta_e, \phi_e)$ for a Lambertian surface is known to be a constant. From (4.5), an equation for a Lambertian surface is obtained as

$$P(\theta_e, \phi_e) = \sum_{i=1}^{n} K_d L(\theta_i, \phi_i) S(\theta_i, \phi_i) \cos\theta_i \omega_i \tag{4.6}$$

where K_d is a diffuse reflection parameter of the surface.

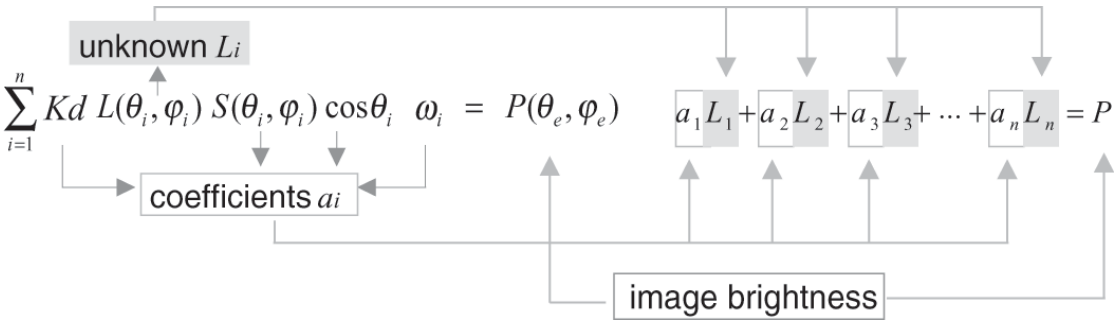


Figure 4.2: Each pixel provides a linear equation.

From (4.6), the recorded pixel value P for an image pixel is given as

$$P = \sum_{i=1}^{n} a_i L_i, \tag{4.7}$$

where L_i $(i = 1, 2, \cdots, n)$ are n unknown illumination radiance specified by n node directions of a geodesic dome. As shown in Figure 4.2, the coefficients $a_i (i = 1, 2, \cdots, n)$ represent $K_d S(\theta_i, \phi_i) \cos\theta_i \omega_i$ in (4.6); we can compute these coefficients from the 3D geometry of a surface point, the occluding object and the illuminant direction. In our examples, we use a modeling tool called the 3D Builder from 3D Construction Company [75] to recover the shape of an occluding object and also the camera parameters from a shadow image. At the same time, the plane of $z = 0$ is defined on the shadow surface.

It is worth noting that, as long as the 3D shape of the shadow surface is provided in some manner, e.g., by using a range sensor, the shadow surface need not be a planar surface. It is rather reasonable to think that a curved shadow surface has the potential for providing more variation in a_{ji} in (4.8) than an ordinary planar surface would since $\cos\theta_i$ term of a_i in (4.7) also slightly changes from image pixel to pixel. Unless severe interreflection occurs within the surface, our method can be applied to a curved shadow surface.

Here if we select a number of pixels, say m pixels, a set of linear equations is obtained as

$$P_j = \sum_{i=1}^{n} a_{ji} L_i \qquad (j = 1, 2, \cdots, m) \tag{4.8}$$

Therefore, by selecting a sufficiently large number of image pixels, we are able to

solve for a solution set of unknown L_i's.[2] For our current implementation, we solve the problem by using the linear least square algorithm with non-negativity constraints (using a standard MATLAB function) to obtain an optimal solution with no negative components.

It should be noted that each pixel $P(\theta_e, \phi_e)$ consists of 3 color bands (R, G, and B) and likewise the diffuse parameter K_d consists of 3 parameters ($K_{d,R}$, $K_{d,G}$, $K_{d,B}$). Each color band of $L(\theta_i, \phi_i)$ is thus estimated separately from the corresponding band of the image pixel and that of the diffuse parameter K_d. For the sake of simplicity in our discussion, we explain the proposed method by using $L(\theta_i, \phi_i)$, $P(\theta_e, \phi_e)$, K_d and do not refer to each of their color bands in the following sections.

We must also draw attention to the importance of shadow coefficients for recovering illumination distribution of a scene from the observed brightness on an object surface. For instance, consider the case described above where the given shadow surface is a Lambertian surface. If we do not take into account the occlusion of including light, the variation of a_{ji} in (4.8) is caused only by the cosine factor of incoming light direction and the surface normal direction of the corresponding point on the shadow surface ($\cos\theta_{ji}$). As a result, the estimation of illumination distribution by solving (4.8) for a solution set of unknown L_i's tends to become numerically unstable.

In contrast, taking into account the occlusion of incoming light tells us that the variation of a_{ji} becomes much larger since the shadow coefficient $S(\theta_i, \phi_i)$ in (4.6) changes significantly, i.e., either 0 or 1, from point to point on the shadow surface depending on the spatial relationships among the location of the point, the occluding object, and the sampling direction of the illumination distribution, as is well illustrated in Lambert's work described in [16]. This characteristics of shadows enables us to reliably recover the illumination distribution of the scene from brightness changes inside shadows.

4.5.2 Known Reflectance Properties: Non-Lambertian surface

The proposed approach may be extended to other reflection models as well. The only condition for a model to satisfy is that it enables us to analytically solve for a solution set of unknown illumination radiance from image brightness.

Take a simplified Torrance-Sparrow reflection model [45, 67] for example; the pixel

[2] More details of the pixel selection are found in Chapter 5.

value of a shadow image $P(\theta_e, \phi_e)$ is computed as

$$\begin{aligned} P(\theta_e, \phi_e) = & \sum_{i=1}^{n} K_d L(\theta_i, \phi_i) S(\theta_i, \phi_i) \cos\theta_i \omega_i + \\ & \sum_{i=1}^{n} K_s L(\theta_i, \phi_i) S(\theta_i, \phi_i) \omega_i \frac{1}{\cos\theta_e} e^{\frac{-\gamma(\theta_i,\phi_i)^2}{2\sigma^2}} \\ = & \sum_{i=1}^{n} (K_d \cos\theta_i + K_s \frac{1}{\cos\theta_e} e^{\frac{-\gamma(\theta_i,\phi_i)^2}{2\sigma^2}}) \\ & S(\theta_i, \phi_i) \omega_i L(\theta_i, \phi_i), \end{aligned} \tag{4.9}$$

where $\gamma(\theta_i, \phi_i)$ is the angle between the surface normal and the bisector of the light source direction and the viewing direction, K_d and K_s are constants for the diffuse and specular reflection components, and σ is the standard deviation of a facet slope of the Torrance-Sparrow reflection model.

From (4.9), we obtain a linear equation for each image pixel where $L(\theta_i, \phi_i)(i = 1, 2, \cdots, n)$ are unknown illumination radiance, and
$(K_d \cos\theta_i + K_s \frac{1}{\cos\theta_e} e^{\frac{-\gamma(\theta_i,\phi_i)^2}{2\sigma^2}}) S(\theta_i, \phi_i)\omega_i$ $(i = 1, 2, \cdots, n)$ are known coefficients.

Again, if we use a sufficiently large number of pixels for the estimation, we are able to solve for a solution set of unknown illumination radiance $L(\theta_i, \phi_i)(i = 1, 2, \cdots, n)$ in this case.

4.5.3 Experimental Results for Known Lambertian Surface

We have tested the proposed approach using an image with an occluding object, i.e., shadow image, taken under usual illumination environmental conditions, including direct light sources such as fluorescent lamps, as well as indirect illumination such as reflections from a ceiling and a wall (Figure 4.3 (b)).

First, an illumination distribution of the scene was estimated using the image irradiance inside shadows in the shadow image. Then a synthetic object with the same shape as that of the occluding object was superimposed onto an image of the scene taken without the occluding object, which is referred as a *surface image*, using the rendering method described in Chapter 2. Note that our algorithm does not require this surface image in the process of illumination estimation in the case where the reflectance properties of the shadow surface are given. This surface image is used here as a background image for superimposing the synthetic occluding object onto the scene.

Synthesized results are shown in Figure 4.4 (a), (b), and (c). Also, we superimposed another synthetic object of a different shape onto the scene in Figure 4.4 (d). The number

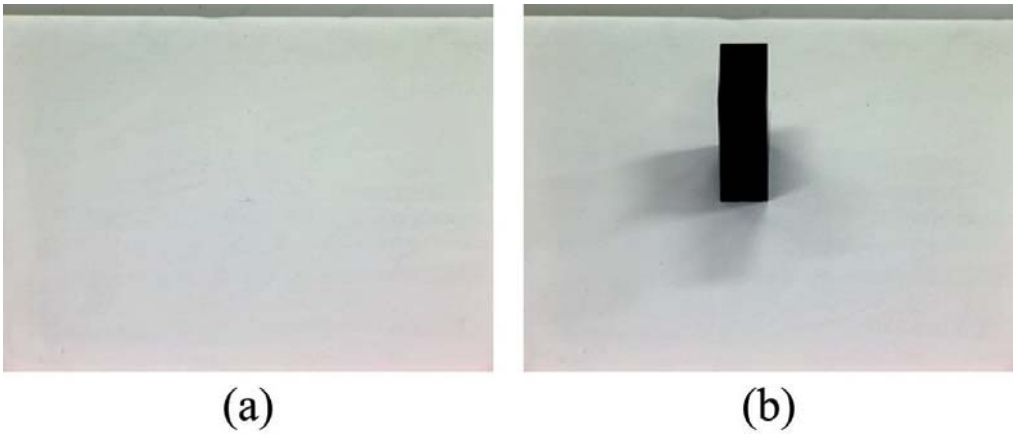


Figure 4.3: Input images : (a) *surface image* (b) shadow image

of nodes of a geodesic dome used for the estimation is shown under the resulting image.

As we see in Figure 4.4, the larger the number of nodes we used, the more the shadows of the synthetic object resembled those of the occluding object in the shadow image. Especially in the case of 520 nodes, the shadows of the synthetic object are nearly indistinguishable from those of the occluding object in the shadow image; this shows that the estimated illumination distribution gives a good presentation of that of the real scene.

Here we see some "stair-casting" artifacts inside the synthesized shadows that result from our approximation of continuous illumination distribution by discrete sampling of its radiance. One approach to reduce this artifact is to increase the number of point light sources used for rendering, whose radiance is given by interpolating the estimated radiance L_i of the neighboring light sources.

Figure 4.5 numerically shows the improvement of the accuracy obtained by increasing the number of samplings. Here, the estimated illumination distribution was evaluated in a different scene condition where the occluding object was rotated 45 degrees on its z axis. The vertical axis represents the average error in pixel values (0 to 255) inside the shadow regions in the synthesized images compared with those in the shadow image. The horizontal axis represents the number of nodes of a geodesic dome used for the estimation. The small pictures right next to the plot show error distributions inside shadow regions in the synthesized images. Darker color represents larger error in a pixel value. The average error at the zero sampling of illumination radiance represents, for reference, the average

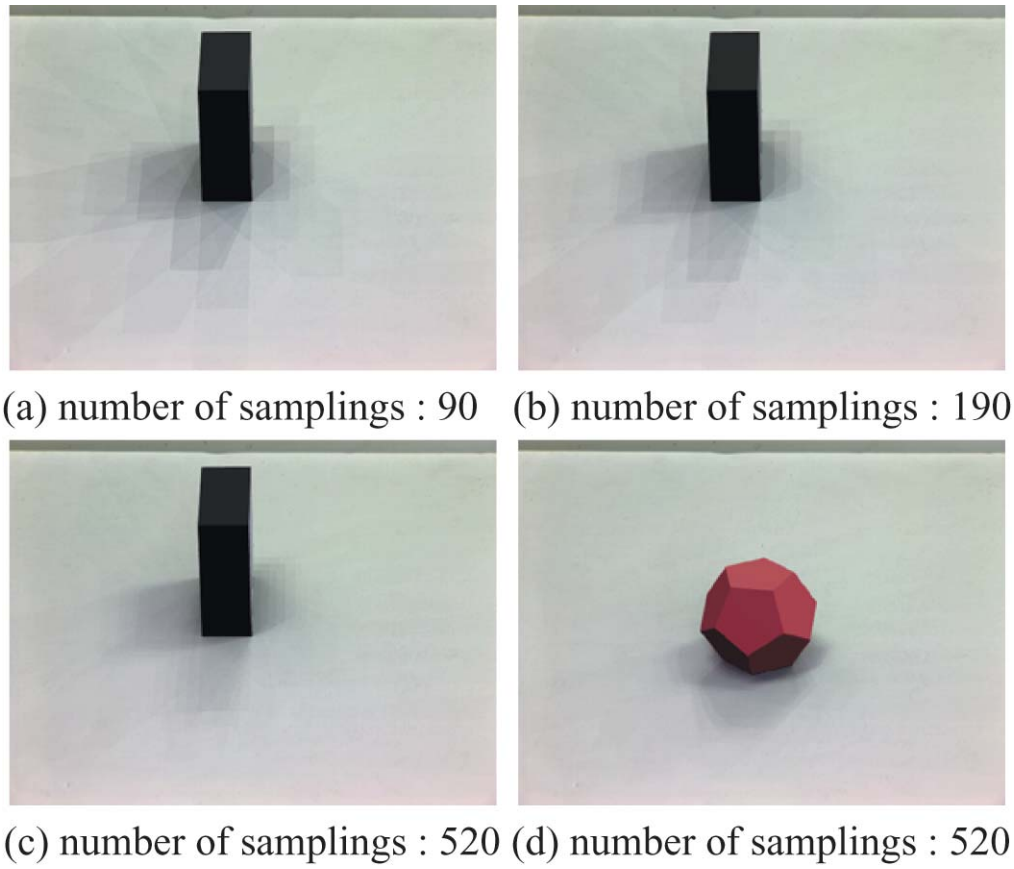


Figure 4.4: Synthesized images: known reflectance property

pixel value inside the shadow regions of the shadow image.

In this plot, we see that flattening of the error curve before convergence to the exact solutions occurs. This is likely caused by the following factors. The first of these is the errors in the measurements of the 3D shapes of the occluding object and the shadow surface. The second factor is the errors in the given reflectance properties of the shadow surface; there is some possibility that the shadow surface is not a perfect Lambertian surface. The third factor is related to our lighting model that approximates the illumination distribution of a scene by discrete sampling of an extended light source on the assumption that light sources in the scene are sufficiently distant from the objects. The input image used in this experiment was taken in an indoor environment, this assumption might not perfectly hold. The last factor concerns our assumption that there is no interreflection between the occluding object and the shadow surface.

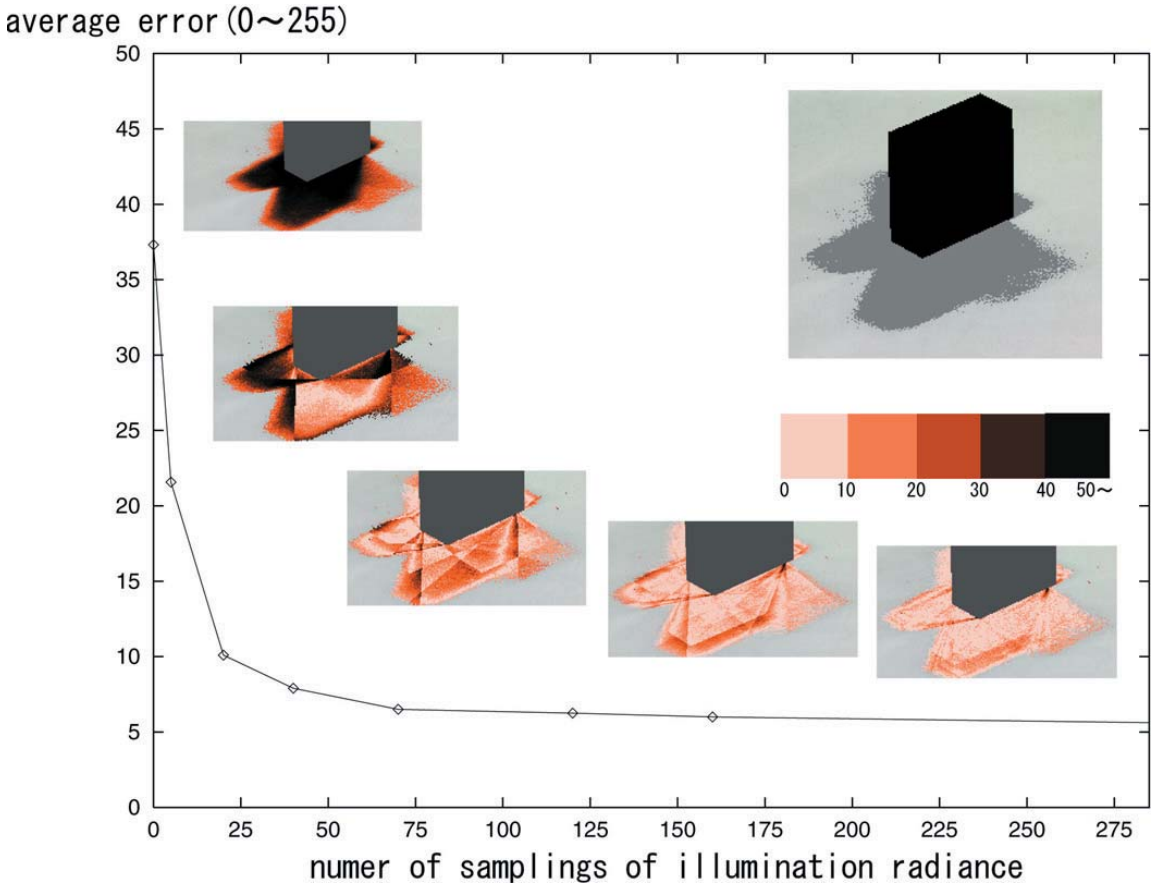


Figure 4.5: Error Analysis: The vertical axis represents the average error in pixel values (0 to 255) inside the shadow regions in the synthesized images compared with those in the shadow image. The horizontal axis represents the number of nodes of a geodesic dome used for the estimation. The small pictures right next to the plot show error distributions inside shadow regions in the synthesized images. Darker color represents larger error in a pixel value.

4.5.4 Unknown Reflectance Properties: Uniform Surface Material

There certainly may be a situation where reflectance properties of a surface are not known a priori, and we need to somehow provide those properties in advance. To cope with this situation, we combine the illumination analysis with an estimation of the reflectance properties of the shadow surface. More specifically, we estimate both the illumination distribution of the scene and the reflectance properties of the surface simultaneously, based on an iterative optimization framework. We will later explain details of this iterative optimization framework in conjunction with the adaptive sampling method described in Section4.6. This approach is based on the assumption that the shadow surface has uniform reflectance properties over the entire surface.

4.5.5 Unknown Reflectance Properties: Non-Uniform Surface Material

The last case is where the BRDF is not given, and the shadow surface does not have uniform reflectance properties, e.g., surfaces with some textures. Even in this case, we are still able to estimate an illumination distribution of a scene from shadows if it is conceivable that the shadow surface is a Lambertian surface.

The question we have to consider here is how to cancel the additional unknown number K_d in (4.6). An additional image of the scene taken without the occluding object, called the *surface image*, is used for this purpose. The image irradiance of a surface image represents the surface color in the case where none of the incoming light is occluded. Accordingly, in the case of the surface image, the shadow coefficients $S(\theta_i, \phi_i)$ always become $S(\theta_i, \phi_i) = 1$. Therefore, the image irradiance $P'(\theta_e, \phi_e)$ of the surface image is computed from (4.6) as

$$P'(\theta_e, \phi_e) = K_d \sum_{j=1}^{n} L(\theta_j, \phi_j) \cos\theta_j \omega_j. \tag{4.10}$$

From (4.6) and (4.10), the unknown K_d is canceled.

$$\begin{aligned} \frac{P(\theta_e, \phi_e)}{P'(\theta_e, \phi_e)} &= \frac{K_d \sum_{i=1}^{n} L(\theta_i, \phi_i) \cos\theta_i S(\theta_i, \phi_i) \omega_i}{K_d \sum_{j=1}^{n} L(\theta_j, \phi_j) \cos\theta_j \omega_j} \\ &= \sum_{i=1}^{n} \frac{L(\theta_i, \phi_i)}{\sum_{j=1}^{n} L(\theta_j, \phi_j) \cos\theta_j \omega_j} \cos\theta_i S(\theta_i, \phi_i) \omega_i \end{aligned} \tag{4.11}$$

Finally, we obtain a linear equation for each image pixel where $\frac{L(\theta_i, \phi_i)}{\sum_{j=1}^{n} L(\theta_j, \phi_j) \cos\theta_j \omega_j}$ are unknowns, $\cos\theta_i S(\theta_i, \phi_i)\omega_i$ are computable coefficients, and $\frac{P(\theta_e, \phi_e)}{P'(\theta_e, \phi_e)}$ is a right-hand side

quantity. Again, if we use a sufficiently large number of pixels for the estimation, we are able to solve the set of linear equations for a solution set of unknown $\frac{L(\theta_i,\phi_i)}{\sum_{j=1}^{n} L(\theta_j,\phi_j)\cos\theta_j}$ $(i = 1, 2, \cdots, n)$.

We should point out that the estimated radiance from these equations is a ratio of the illumination radiance in one direction $L(\theta_i, \phi_i)$ to scene irradiance at the surface point $\sum_{j=1}^{n} L(\theta_j, \phi_j)\cos\theta_j\omega_j$. As a result, we cannot relate the estimated radiance over the color bands unless the ratio of the scene irradiance among color bands is provided. It is, however, not so difficult to obtain this ratio. For instance, if there is a surface with a white color in the scene, the recorded color of the surface directly indicates the ratio of the scene irradiance among color bands. Otherwise, an assumption regarding total radiance is required.

4.5.6 Experimental Results for Non-Uniform Surface Material

The input images used in this experiment are shown in Figure 4.6. The illumination distribution of the scene was estimated using the image irradiance of both the shadow image and the surface image, and a synthetic occluding object was superimposed onto the surface image using the estimated illumination distribution. Synthesized results are shown in Figure 4.7. Again, in the case of 520 nodes, the shadows in the resulting image resemble those of the occluding object in the shadow image. This shows that the estimated illumination distribution gives a good representation of the characteristics of the real scene.

Figure 4.8 numerically shows the improvement of the accuracy by increasing the number of samplings. The estimated illumination distribution was evaluated using a new occluding object with a different shape. From the plot in the figure, we can clearly see that the accuracy improves as we use more sampling directions of the illumination distribution. It is likely that the flattening of the error curve before convergence to the exact solutions is caused by the same factors described in Section 4.5.3.

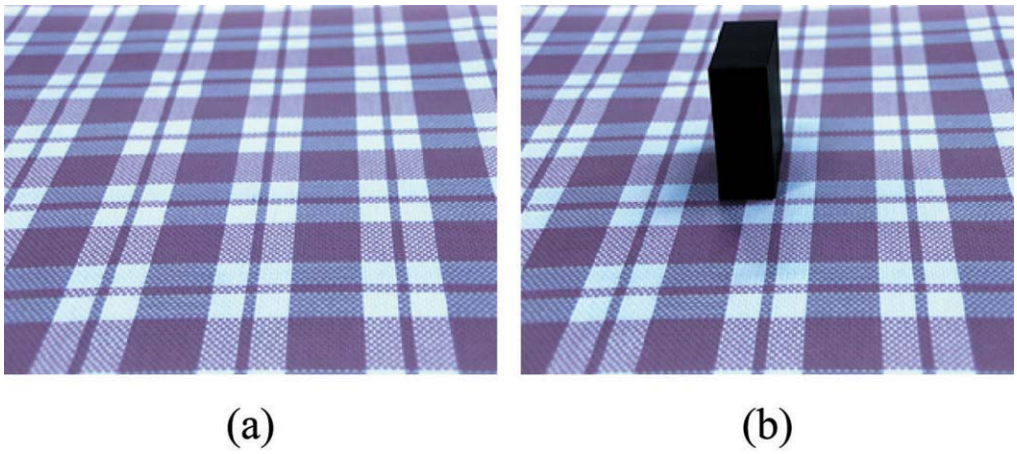


Figure 4.6: Input images : (a) surface image (b) shadow image

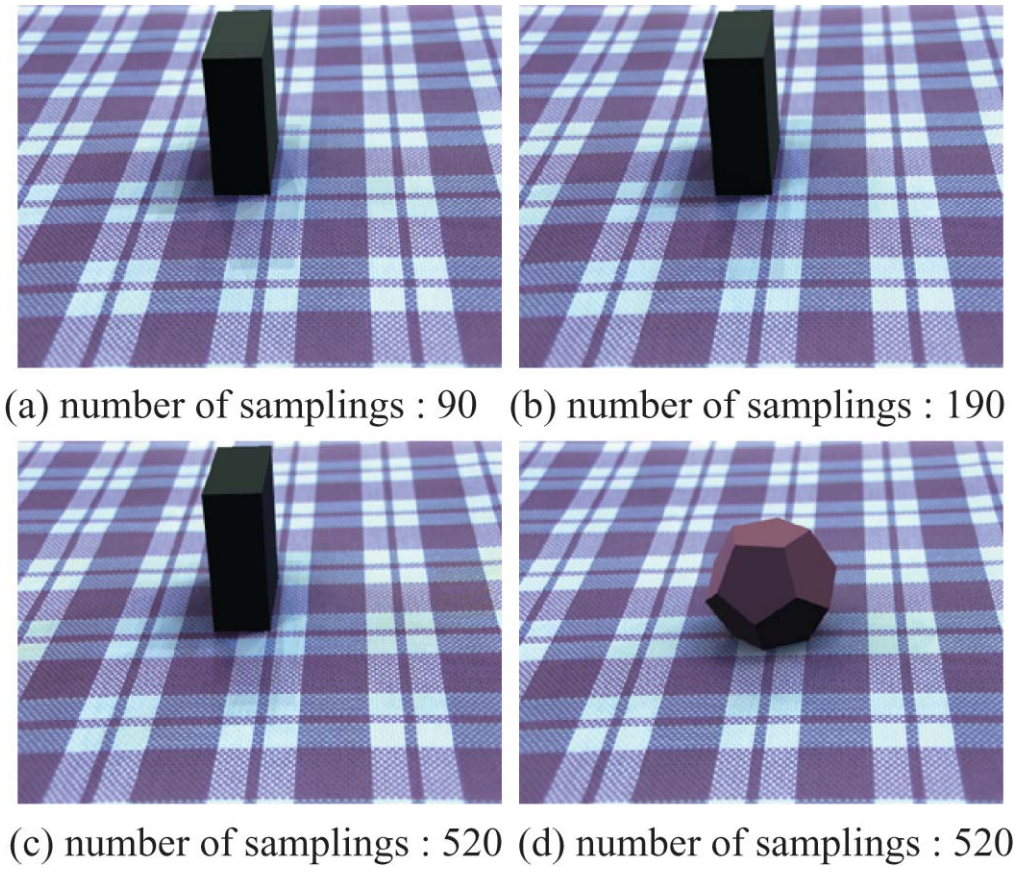


Figure 4.7: Synthesized images: unknown reflectance property

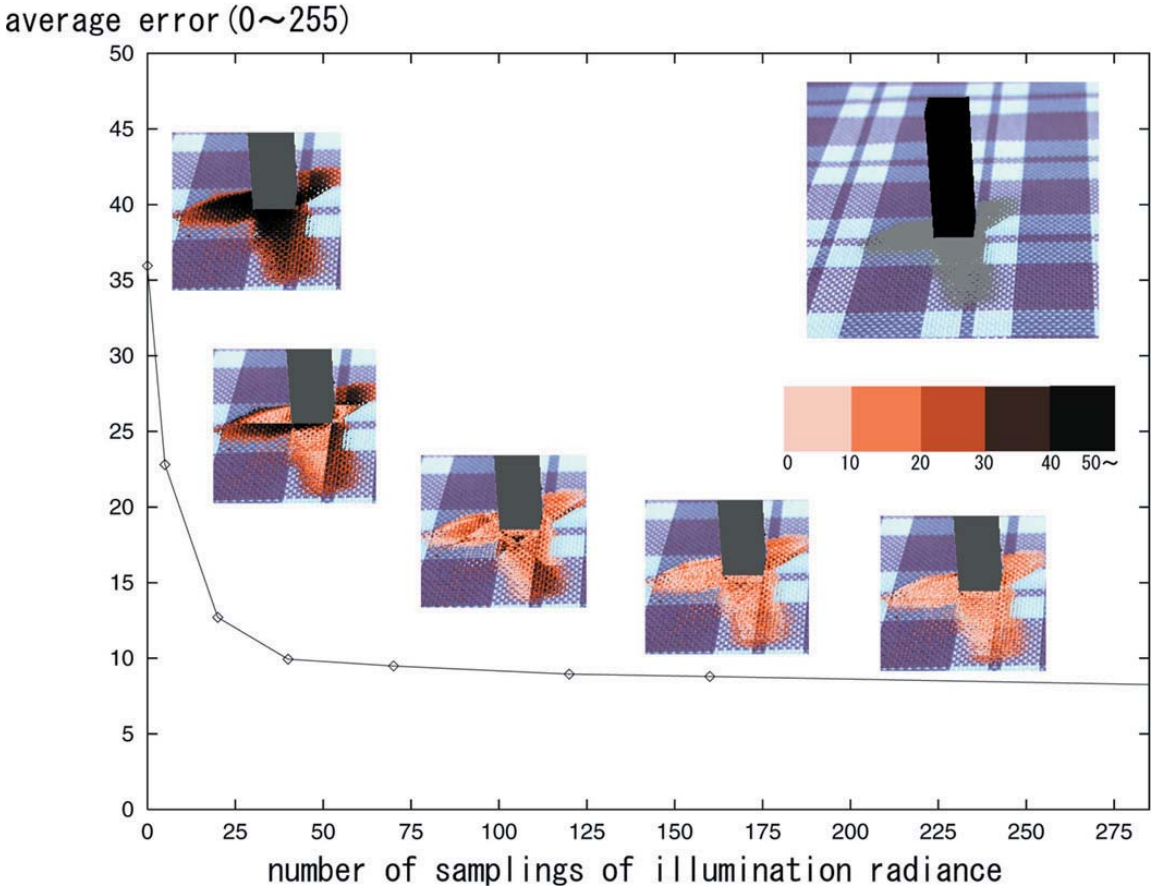


Figure 4.8: Error Analysis: the vertical axis represents the average error in pixel values (0 to 255) inside the shadow regions in the synthesized images compared with those in the shadow image. The horizontal axis represents the number of nodes of a geodesic dome used for the estimation. The small pictures right next to the plot show error distributions inside shadow regions in the synthesized images. Darker color represents larger error in a pixel value.

4.6 Adaptive Estimation of Illumination Distribution with Unknown Reflectance Properties of Shadow Surface

This section further addresses two issues in inverse lighting. The first issue is how to provide reflectance properties of a surface inside shadows in cases where they are not given a priori, and the shadow surface is not conceivable as a Lambertian surface.

The second issue is how to efficiently estimate an illumination distribution of a scene without losing the accuracy of the estimated illumination distribution. In Section 4.5, we have estimated an illumination distribution of a real scene by sampling the distribution at an equal solid angle given as node directions of a geodesic dome. For more efficiently estimating the illumination distribution of a scene using fewer sampling directions of illumination radiance, we propose to increase sampling directions adaptively based on the estimation at the previous iteration, rather than by using a uniform discretization of the overall illumination distribution in this section.

4.6.1 Basic Steps of the Proposed Approach

To take both diffuse and specular reflections of the shadow surface into consideration, a simplified Torrance-Sparrow reflection model (4.9) described in Section 4.5.2 is reused. Based on this reflection model, both the illumination distribution of the scene and the reflectance properties of the shadow surface are estimated from image brightness inside shadows as described in the following steps.

1. Initialize the reflectance parameters of the shadow surface. Typically, we assume the shadow surface to be Lambertian, and the diffuse parameter K_d is set to be the pixel value of the brightest point on the shadow surface. The specular parameters are set to be zero ($K_s = 0, \sigma = 0$). Note that the initial value of K_d is not so important since there is a scaling factor between the reflectance parameters and illumination radiance values in any case. To fix the scaling factor, we need to perform photometric calibration of our imaging system with a calibration target whose reflectance is given a priori.

2. Select image pixels whose brightness is used for estimating both the illumination distribution of a scene and the reflectance properties of the shadow surface. This selection is done by examining the coefficients a_i in (4.7).

3. Estimate illumination radiance $L(\theta_i, \phi_i)$. Using the reflectance parameters (K_d, K_s, σ) and image brightness inside shadows in the shadow image, the radiance distribution $L(\theta_i, \phi_i)$ is estimated as described in Section 4.5.2.

4. Estimate the reflectance parameters of the shadow surface (K_d, K_s, σ) from the obtained radiance distribution of the scene $L(\theta_i, \phi_i)$ by using an optimization technique. (Section 4.6.2)

5. Proceed to the next step if there is no significant change in the estimated values $L(\theta_i, \phi_i), K_d, K_s$, and σ. Otherwise, go back to Step 3. By estimating both the radiance distribution of the scene and the reflectance parameters of the shadow surface iteratively, we can obtain the best estimation of those values for a given set of sampling directions of the illumination radiance distribution of the scene.

6. Terminate the estimation process if the obtained illumination radiance distribution approximates the real radiance distribution with sufficient accuracy. Otherwise, proceed to the next step.

7. Increase the sampling directions of the illumination distribution adaptively based on the obtained illumination radiance distribution $L(\theta_i, \phi_i)$. (Section 4.6.3)
   Then go back to Step 2.

In the following sections, each step of the proposed approach will be explained in more detail.

4.6.2 Estimation of Reflectance Parameters of Shadow Surface based on Radiance Distribution

In this section, we describe how to estimate the reflectance parameters of the shadow surface (K_d, K_s, σ) by using the estimated radiance distribution of the scene $L(\theta_i, \phi_i)$.

Unlike the estimation of the radiance distribution of the scene $L(\theta_i, \phi_i)$, which can be done by solving a set of linear equations, we estimate the reflectance parameters of the shadow surface by minimizing the sum of the squared difference between the observed pixel intensities in the shadow image and the pixel values computed for the corresponding surface points. Hence, the function to be minimized is defined as

$$f = \sum_{j=1}^{m} (P_j' - P_j)^2 \tag{4.12}$$

where $P_j{}'$ is the observed pixel intensity in shadows cast by the occluding object, P_j is the pixel value of the corresponding surface points computed by using the given radiance distribution of the scene $L(\theta_i, \phi_i)$ in (4.9), m is the number of pixels used for minimization. In our method, the error function in (4.12) is minimized with respect to the reflectance parameters K_d, K_s, and σ by the Powell method to obtain the best estimation of those reflectance parameters [51]. As has been noted, this approach is based on the assumption that the shadow surface has uniform reflectance properties over the entire surface.

4.6.3 Adaptive Sampling of Radiance Distribution

If the estimated radiance distribution for a set of sampling directions does not approximate the illumination distribution of the scene with sufficient accuracy, we increase the sampling directions adaptively based on the current estimation of the illumination radiance distribution.

Radiance distribution changes very rapidly around a direct light source such as a fluorescent light. Therefore, the radiance distribution has to be approximated by using a large number of samplings so that the rapid change of radiance distribution around the direct light source is captured. Also, to correctly reproduce soft shadows cast by extended light sources, radiance distribution inside a direct light source has to be sampled densely.

On the other hand, coarse sampling of radiance distribution is enough for an indirect light source such as a wall whose amount of radiance remains small. As a result, the number of sampling directions required for accurately estimating an illumination distribution of a real scene becomes exceedingly large.

To overcome this problem, we increase sampling directions adaptively based on the estimation at the previous iteration, rather than by using a uniform discretization of the overall illumination distribution. In particular, we increase sampling directions around and within direct light sources.[3]

Based on the estimated radiance distribution $L(\theta_i, \phi_i)$ for the sampling directions at the previous step, additional sampling directions are determined as follows.

Suppose three sampling directions with radiance values L_1, L_2, and L_3 are placed to form a triangle M_1 as illustrated in Figure 4.9. To determine whether a new sampling

[3] A similar sampling method has been employed in radiosity computation to efficiently simulate the brightness distribution of a room [6].

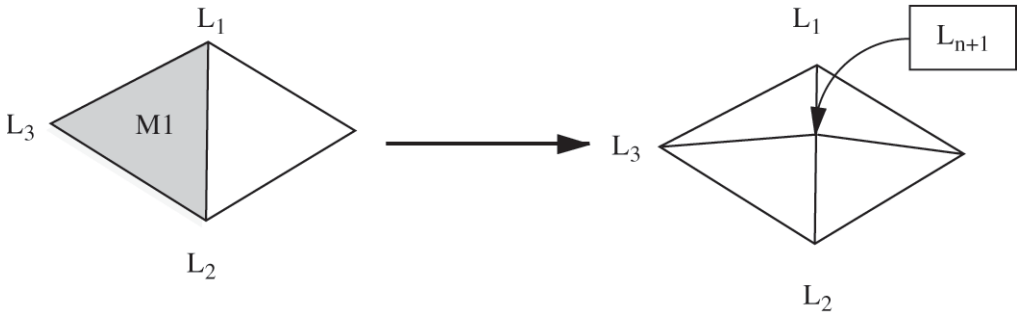


Figure 4.9: Subdivision of sampling directions

direction needs to be added between L_1 and L_2, we consider the following cost function

$$U(L_1, L_2) = \mathit{diff}(L_1, L_2) + \alpha \mathit{min}(L_1, L_2)\mathit{angle}(L_1, L_2). \tag{4.13}$$

where $\mathit{diff}(L_1, L_2)$ is the radiance difference between L_1 and L_2, $\mathit{min}(L_1, L_2)$ gives the smaller radiance of L_1 and L_2, $\mathit{angle}(L_1, L_2)$ is the angle between directions to L_1 and L_2, and α is the manually specified parameter which determines the relative weights of those two factors. The first term is required to capture the rapid change of radiance distribution around direct light sources, while the second term leads to fine sampling of the radiance distribution inside direct light sources. The additional term $\mathit{angle}(L_1, L_2)$ is used for avoiding unnecessarily dense sampling inside direct light sources. In our experiments, α is set to 0.5.

If the cost U is large, a new sampling direction is added between L_1 and L_2. In our experiments, we computed the cost function values U for all pairs of neighboring sampling directions, then added additional sampling directions for the first 50% of all the pairs in the order of the cost function values U.

4.6.4 Experimental Results for Adaptive Sampling Method

We have tested the proposed adaptive sampling method by using real images taken in both indoor and outdoor environments. In the following experiment, the adaptive sampling technique was applied to a uniform unknown reflectance properties case. The shadow image shown in Figure 4.10 was used here as an input of a indoor scene.

Starting from a small number of uniform sampling directions of the illumination distribution, the estimation of the radiance distribution of the scene was iteratively refined. At the same time, the reflectance parameters (K_d, K_s, and σ) of the shadow surface were estimated as explained in Section 4.6.2. Then an appearance of the shadow surface around the occluding object was synthesized using the estimated radiance distribution of the scene and the estimated reflectance parameters of the shadow surface.

The region inside the red rectangle in Figure 4.10 (b) was replaced with the synthesized appearances in the left column in Figure 4.11. The number of sampling directions of the radiance distribution used for the estimation is shown under the resulting images. Here, the resulting shadows of the synthetic occluding object resemble more closely those of the occluding object in the shadow image as we increase the number of sampling directions based on the proposed adaptive sampling technique. Finally, in the case of 140 sampling directions, the synthesized shadows of the occluding object blend into the input shadow image well, and few distinct boundaries in the shadows are seen in the resulting composite image.

To see how well the adaptive sampling of radiance distribution works in this example, we took an omni-directional image of the office scene as a ground truth. The middle column of Figure 4.11 shows the omni-directional image of the scene taken by placing a camera with a fisheye lens looking upward on the shadow surface in Figure 4.10 (a). The omni-directional image shows both direct light sources, i.e., the fluorescent lamps in our office, and indirect light sources such as the ceiling and walls. The right column of Figure 4.11 shows the estimated radiance values visualized for comparison with the ground truth. In those images in Figure 4.11 (b) and (c), we can see that sampling directions of the radiance distribution were nicely added only around the direct light sources at each step by the proposed adaptive sampling framework, starting from the coarse sampling directions at the top row.

Figure 4.12 numerically shows the improvement of the accuracy by adaptive refinement of sampling directions and the estimation of reflectance properties of the shadow surface. From the plot in the figure, we can clearly see that the accuracy improves rapidly as we adaptively increase the sampling directions of the radiance distribution.

To confirm the merit of the adaptive sampling framework, we also estimated the illumination radiance distribution with uniform sampling. In that case, even 300 uniformly

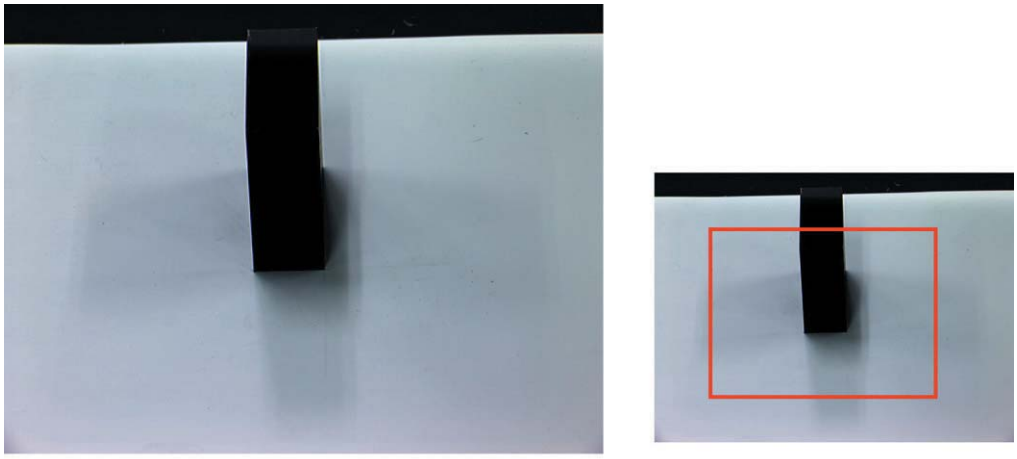


Figure 4.10: Input image : (a) shadow image taken of an indoor scene (b) the region which synthesized images with the estimated radiance distribution and reflectance parameters are superimposed.

sampled directions could not achieve the same level of accuracy as the estimation result obtained by 80 sampling directions with the proposed framework.

Figure 4.13 (a) shows another example image taken outside the entrance lobby of our building in the late afternoon. In this image, we used the rectangular pole with two colors as an *occluding object* casting shadows. In the same way as the previous example, an appearance of the shadow surface around the *occluding object*, illustrated with a red rectangle in Figure 4.13 (b), was synthesized by using the estimated radiance distribution of the scene and the estimated reflectance parameters of the shadow surface. Figure 4.14 shows the resulting images obtained by the use of our method. Although the grid pattern on the shadow surface is missing in those synthesized images due to the assumption of uniform reflectance on the shadow image, the appearance of the shadow around the occluding objects is virtually indistinguishable in the case of 140 sampling directions. This shows that the estimated illumination distribution well represents the characteristics of the real scene.

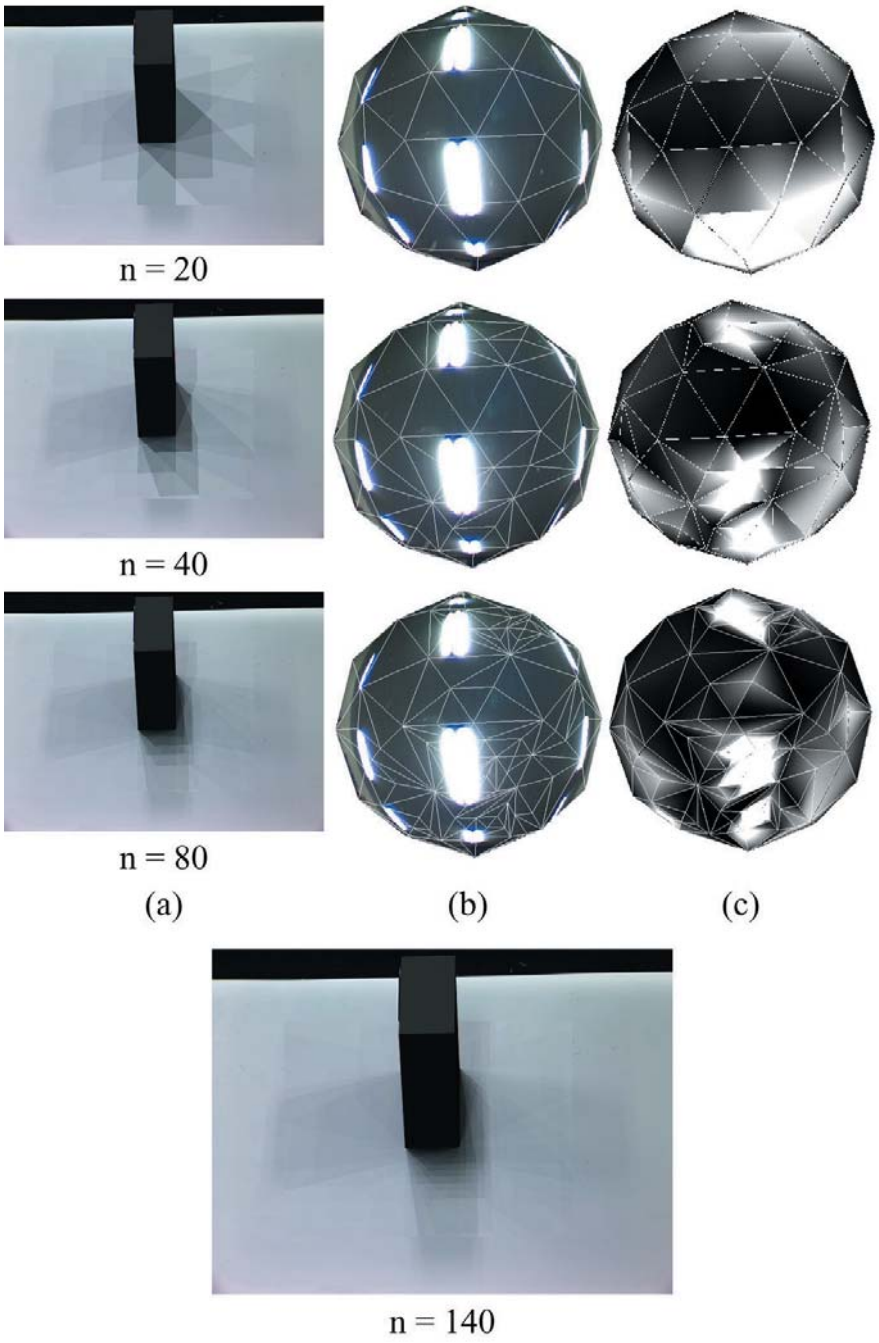


Figure 4.11: Adaptive refinement of illumination distribution estimation: (a) synthesized images with the estimated radiance distribution and reflectance parameters (b) adaptive refinement of sampling directions with a ground truth of an omni-directional image of the scene (c) the estimated radiance values visualized for comparison with the ground truth

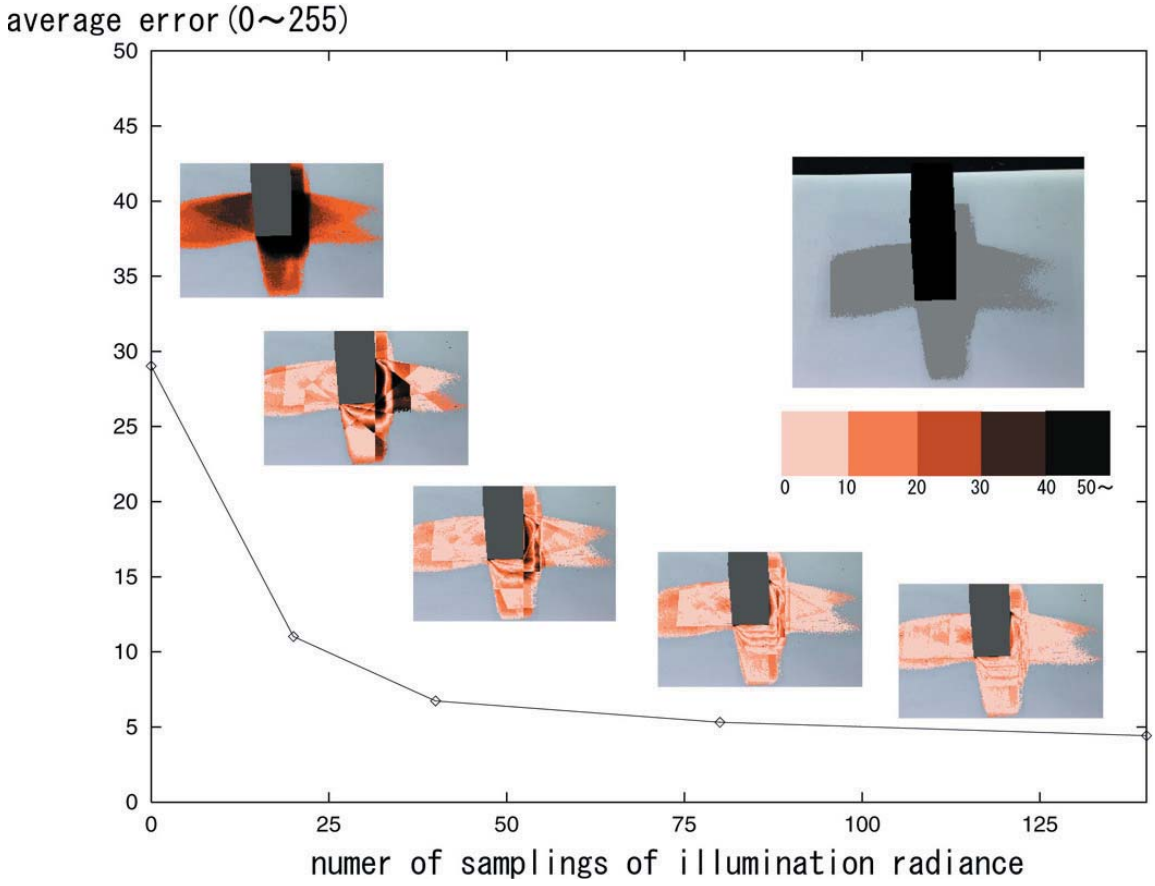


Figure 4.12: Error Analysis

4.7 Discussion

In the field of computer graphics, there have also been several methods proposed for recovering an illumination distribution of a real scene from appearance changes observed on object surfaces in the scene. Most of them have been conducted in fully computer-generated environments where a designer specified the geometry and reflectance of whole scenes and the locations of all light sources [59, 31]; those methods were mainly developed for assisting designers in achieving desirable effects such as highlights and shadows in the synthesized image.

Some researchers extended similar approaches for real images and estimated radiance of real light sources in the scene. Fournier et al. [17] estimated radiance of the light sources with users' defined fix locations on the assumption that all surfaces were Lambertian surfaces. Later, Marschner and Greenberg [39] introduced to approximate the entire illumination with a set of basis lights located in the scene. Although this method has an advantage over the previous methods of not requiring knowledge about the light locations of the scene, the estimation depends entirely on the appearance changes observed on an object surface assumed to be Lambertian, and therefore some restrictions are imposed on the shape of the object, e.g., the object must have a large amount of curvature.

Later, Ramamoorthi and Hanrahan [52] introduced a signal-processing framework that described the reflected light field as a convolution of the lighting and BRDF, and showed under which condition lighting and BRDF recovery could be done robustly. Simultaneously, Basri and Jacobs [2] reported that the images of a convex Lambertian object obtained under a wide variety of lighting conditions could be approximated with a low-dimensional linear subspace by using a similar signal-processing approach.

Recently, based on the signal-processing framework proposed by Ramamoorthi and Hanrahan, it was shown that high frequency components of the appearance of an object surface could retain significant energy by taking the occlusion of incoming light as well as its bidirectional reflectance distribution function (BRDF) into account in [48]. This indicates that the use of shadows for illumination estimation has the significant advantage of providing more clues to the high frequency components of illumination distribution of a scene.

4.8 Conclusions

In this chapter, we have presented a method for recovering an illumination distribution of a scene from image brightness observed on a real object surface. In a natural illumination condition, it is hard to recover an illumination distribution from image brightness observed on a real object surface. The main reason for this difficulty is that there is usually not adequate variation in the image brightness observed on the surface to reflect the subtle characteristics of the entire illumination.

One of the main contributions of our work is to demonstrate the effectiveness of using occluding information of incoming light in estimating the illumination distribution of a scene. Analyzing the relationships between the image brightness and the occlusions of incoming light enabled us to reliably estimate the illumination distribution of a scene even in a complex illumination environment.

In addition, the question of how to provide the reflectance properties of a surface to be used for the estimation still remains from the previously proposed methods. Since it is a common situation that reflectance properties of a surface are not known, solutions to this problem are required. Another contribution of our work is that we have combined the illumination analysis with an estimation of the reflectance properties of a surface in the scene. This makes the method applicable to the case where reflectance properties of a surface are not known, and it enlarges the variety of images to which the method can be applied.

We also introduced an adaptive sampling of the illumination distribution of a scene as a solution to the question of how we could efficiently estimate the entire illumination with a smaller number of sampling directions of the entire distribution in this paper. Using the proposed adaptive sampling framework, we were able to avoid unnecessarily dense sampling of the illumination and to estimate overall illumination more efficiently by using fewer sampling directions.

While the effectiveness of the proposed method in estimating illumination distributions of usual scenes was demonstrated in this paper, the estimation was based on the assumptions that light sources in the scene were sufficiently distant from the objects, and that there was no severe interreflection between a shadow region and an occluding object. The future directions of this study include extending the method for: (1) considering the interreflection between a shadow region and an occluding object; and (2) taking distances

(a) (b)

Figure 4.13: Input image : (a) shadow image taken in an outdoor scene (b) the region where synthesized images with the estimated radiance distribution and reflectance parameters are superimposed in Figure 4.14

from the occluding object to light sources into account.

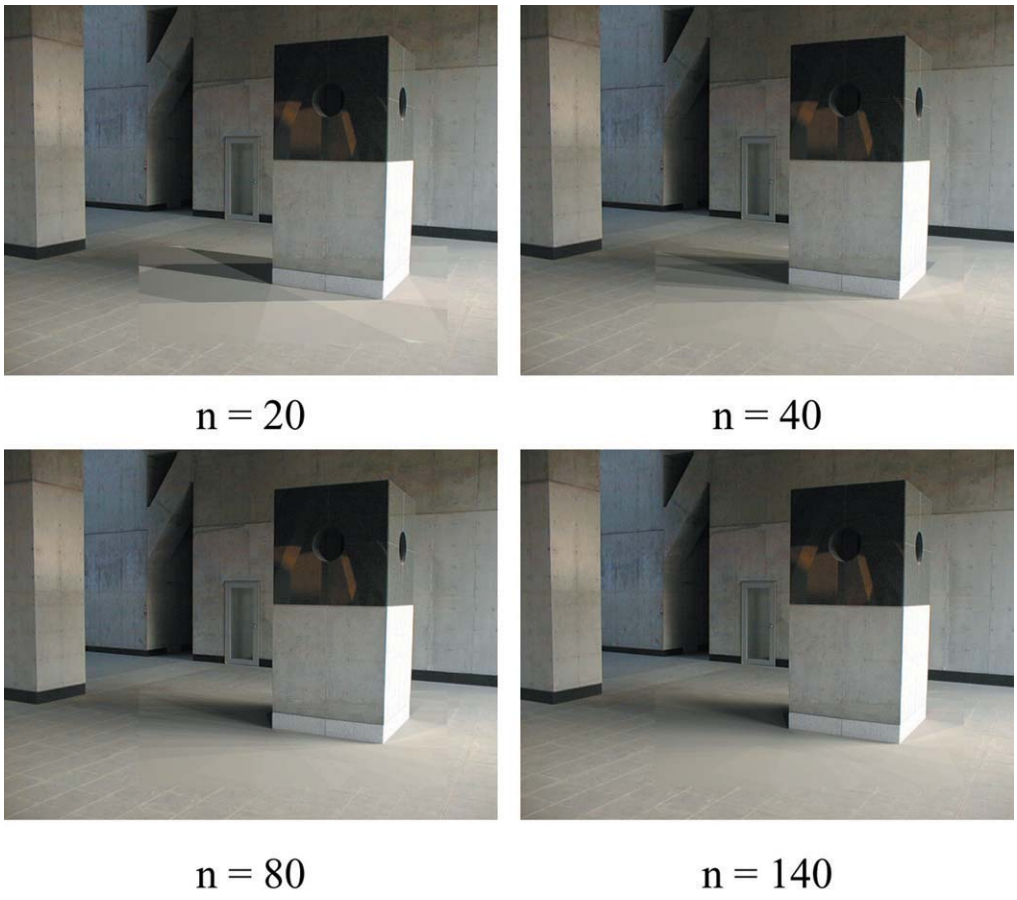


Figure 4.14: Adaptive refinement of illumination distribution estimation: synthesized images with the estimated radiance distribution and reflectance parameters

Chapter 5

Stability Issues in Illumination Estimation

In this chapter, we discuss the amount of the information about the illumination distribution of the scene obtainable from a given image of a scene. It is worth noting that brightness distributions inside shadows vary as a function of the lighting environment of the scene and the shape of an occluding object. As a consequence, the amount of the information contained in a shadow image about the illumination distribution of a scene changes from image to image, depending on how much of the shadow surfaces are blocked by the occluding object and how much are covered by the field of view of the camera taking the shadow image. Hence it is essential to evaluate an shadow image first in terms of how much information about the illumination distribution of the scene the image can provide. We explain two main factors that control the stability of the illumination estimation from shadows: blocked view of shadows; limited sampling resolution for radiance distribution inside shadows. Then we introduce solutions for overcoming these two problems.

5.1 Illumination Distribution Solvable From a Given Shadow Image

Let us first consider the ideal case where radiance values of light source L_i's can be estimated without suffering from any instability problems. As shown in Figure 4.2, a set

of linear equations is derived from a given shadow image. Recall (4.8)

$$P_j = \sum_{i=1}^{n} a_{ji} L_i \qquad (j = 1, 2, \cdots, m),$$

and rewrite it in a matrix form as

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ & \cdots & \cdots & & \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ L_3 \\ \vdots \\ L_n \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ \vdots \\ P_m \end{bmatrix}$$

$$A \cdot L = P. \tag{5.1}$$

Here the coefficients matrix A of the system $A \cdot L = P$ is consists of n column vectors $\mathbf{a_i} = (a_{1i}, a_{2i}, \ldots, a_{mi})^{\mathrm{T}}$ where n is the number of sampling directions, and m is the number of the image pixels. Here, if the coefficients matrix A consist of n linearly independent column vectors $\mathbf{a_i}$, we numerically obtain enough constraints for solving A for unknown radiance of L_i's.

The amount of the information that shadows in a shadow image provide is obviously limited. Under some circumstances, we might not be able to derive n linear independent vectors $\mathbf{a_i}$ from a given shadow image. As shown in Figure 4.2, the coefficients a_{ij} in the coefficients matrix A represent $f(\theta_i, \phi_i; \theta_e, \phi_e) S(\theta_i, \phi_i) \cos\theta_i$, and the most powerful clue to the illumination distribution of a scene among the components of a_i is $S(\theta_i, \phi_i)$ (either 0 or 1) since the other components do not change so much from image pixel to pixel.

As a consequence, it makes the estimation more stable if we observe the difference between radiance of two shadow regions for each light source: one illuminated and the other not illuminated by the light source. In other words, the more variations in $S(\theta_i, \phi_i)$'s in shadow regions we have, the more likely it is that we will obtain n linearly independent column vectors, thereby making it possible to robustly estimate radiance values of the light sources approximating illumination distribution of the scene.

In the following sections, we describe the main sources that prevent us from observing both shadow and non-shadow regions for a light source in a particular direction. This tends to makes the matrix A ill-conditioned and leads to the instability problem of the proposed method.

5.2 Blocked View of Shadows

The visible portion of a shadow surface changes from image to image depending on how wide the field of view of a camera taking the shadow image is, and what size portions of the view of the camera are occluded by the occluding object. Only part of the shadows cast by the occluding object can be usually observed in the shadow image, with the exception of special cases such as a camera aimed directly toward a floor onto which shadows are cast.

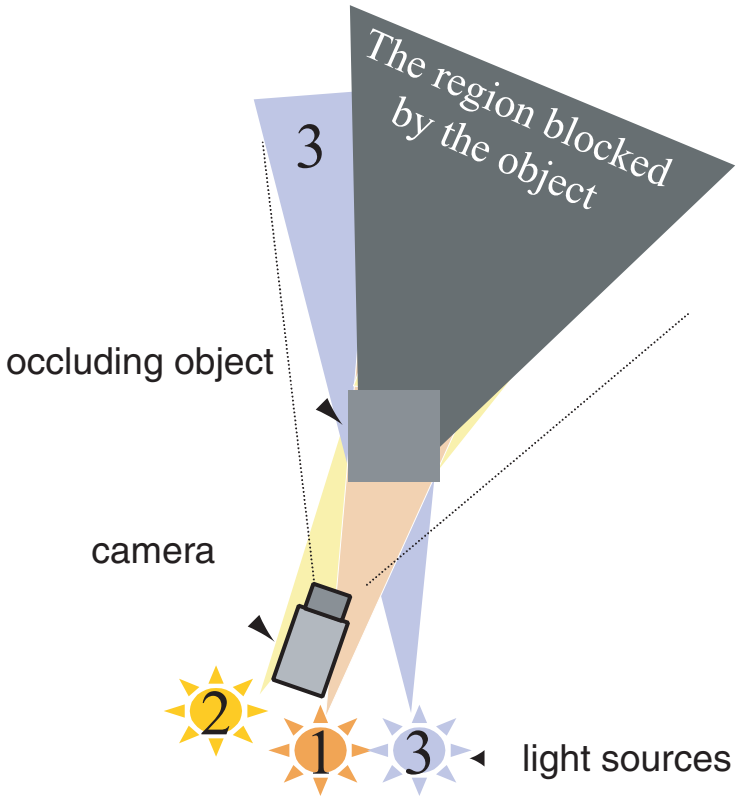


Figure 5.1: Blocked view of shadows.

Consider the case shown in Figure 5.1 as an example. In this case, the region behind the occluding object cannot be seen by the camera. More precisely, only the shadow caused by the light source 3 can be observed in the shadow image, and shadows cast by the light sources 1 and 2 are not seen.

Let $\mathbf{a_1}$ and $\mathbf{a_2}$ as column vectors and L_1 and L_2 as radiance values for the light source 1 and 2 respectively, then in the case shown in this example, $\mathbf{a_1}$ and $\mathbf{a_2}$ become close to

singular. In other words, the occlusion coefficients $S(\theta_i, \phi_i)$ of $\mathbf{a_1}$ and those of $\mathbf{a_2}$ are the same for all shadow pixels, and only differences for distinguishing the light source 1 from the light source 2 are in $f(\theta_i, \phi_i; \theta_e, \phi_e) \cos\theta_i$. As a result, the estimation becomes too sensitive to the image noise, and radiance values of the light source 1 and 2 can not be reliably estimated.

5.3 Limited Sampling Resolution For Radiance Distribution inside Shadows

Second, the number of image pixels is fixed, and therefore radiance distribution inside shadow regions can be measured only up to a certain resolution.

In the basic formulation of the previously proposed method described in Chapter 4.5, it was simply assumed that we could obtain a large enough number of image pixels inside shadow regions to provide sufficient constraints for solving a set of linear equation (5.1) for unknown radiance values of point light sources. However, this is true only when we can measure radiance values inside shadows at any resolution.

To illustrate this point, consider two cases shown in Figure 5.2. In the first case in (1), illumination distribution of a scene is represented by coarsely distributed point light sources 1, 2, and 3. On the other hand, more densely distributed point light sources are used in the other case shown in (2). Shadows are partitioned into smaller regions depending on which light source is occluded by the occluding object. For instance, the shadow in the left case is partitioned into the following regions in order from left to right: 3, 1&3, 1, 1&2&3, 1&2, and 2.

As we can see in the case in Figure 5.2 (2), the shadow does not contain a region where only the point light source 1 is occluded. Also, partitioned regions such as 3 or 2 are significantly smaller than those in the case in Figure 5.2 (1). As a result, we obtain less of a variety of partitioned regions inside shadows, thus making the estimation less reliable, and the estimation further becomes more sensitive to image noise because each partitioned region contains fewer image pixels.

Since the number of image pixels is limited, radiance distribution inside shadows can be measured only up to a certain resolution. As a result, sampling resolution of illumination distribution of a real scene is also bound to a certain limit. For instance, if we use more

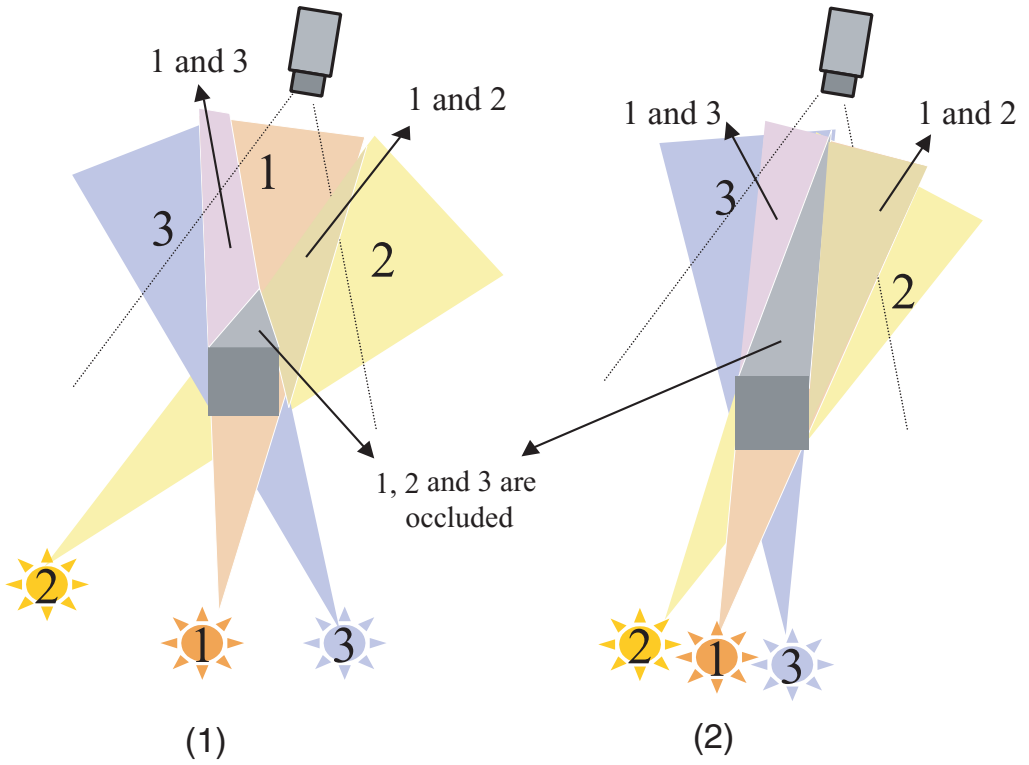


Figure 5.2: Difference in visible combinations of occlusions of lights: (1) coarsely distributed light sources and (2) densely distributed light sources.

densely distributed point light sources for the estimation, we might not even observe any image pixels of some shadow regions.

It follows from the discussions above that there is a chance that the estimation becomes too unstable to provide a correct estimate of the illumination distribution of the scene for a given set of point light sources for the estimation. As a solution to this problem, a new procedure to evaluate an input shadow image in terms of the information about the illumination distribution of a scene obtainable from the image is considered in the next chapter. Using the outputs from this procedure, a set of point light sources is defined so that radiance values of those light sources can be estimated reliably from the given *shadow images*.

5.4 Techniques For Robust Estimation

Recall the case shown in Figure 5.1 where two column vectors $\mathbf{a_1}$ and $\mathbf{a_2}$ of the light sources 1 and 2 respectively are close to singular. Here, if the light sources 1 and 2 are combined to form a larger light source for more coarse sampling of illumination distribution, we are able avoid facing the instability problem caused by the light source 1 and the light source 2. Similarly, the sampling resolution problem described in Section 5.3 can be avoided if we use the more coarse sampling instead of dense sampling of illumination radiance whenever it is decided to be necessary in terms of the stability of computation for solving the set of linear equations obtained from a given shadow image.

In this chapter, we will adopt a strategy for changing the sampling density of the illumination distribution depending on the amount of the information obtainable from a shadow image for a particular direction of the illumination distribution.

5.5 Selection of Illumination Distribution Samplings

The derived coefficients matrix A from a shadow image is first examined in terms of the stability of computation for solving the set of linear equations obtained from a given shadow image.

Stability of a given system is measured by the condition number. Some systems are sensitive to errors and others are not. For a positive matrix A, the condition number c is computed (again using a standard MATLAB function) as

$$c = \sigma_{\max}/\sigma_{\min}, \tag{5.2}$$

where $\sigma_{\max}$ is the maximum singular value of the matrix A, and $\sigma_{\min}$ is the minimum singular value.

If this condition number c is sufficiently small, (5.1) can be solved without numerical instability. Therefore, the given set of light sources representing real illumination distribution is appropriate for the given shadow image. On the other hand, if the condition number c is large, the problem of illumination estimation for the given set of light sources is close to ill-conditioned. Therefore, we need to reduce light sources representing illumination distribution by combining several light sources into one with a larger solid angle.

Then, our question is how to select those light sources to be combined for sampling

the illumination distribution more coarsely. This is done by examining the column vector $\mathbf{a_i} = (a_{1i}, a_{2i}, \ldots, a_{mi})^{\mathrm{T}}$ corresponding to a light source L_i in (5.1). If two vectors $\mathbf{a_i}$ and $\mathbf{a_j}$ for two light sources L_i and L_j are about the same, (5.1) becomes ill-conditioned and brightness distribution inside shadows does not provide sufficient information for determining radiance of those two light sources. Therefore, those two light sources need to be combined to form a larger light source for more coarse sampling of illumination distribution. We evaluate the similarity of two vectors $\mathbf{a_i}$ and $\mathbf{a_j}$ with their dot product $\mathbf{a_i} \cdot \mathbf{a_j}$.

Since the entire stability computation depends on the coefficients matrix A and the radiance values L_i are estimated from these coefficients as well, it is important to compute a_i of the matrix A as accurately as possible. For this purpose, we further introduce the following two techniques. One is a technique for computing more accurate shadow coefficients $S(\theta_i, \phi_i)$ using the more appropriate light model for L_i. The other is a technique for sampling pixels from a shadow image in such a way as to maximize the information obtainable from the image.

5.6 Occlusion Test of Incoming Lights

In the proposed approach, the entire illumination distribution of a scene is represented as a set of point sources equally distributed in the scene whose solid angle is defined by a sampling density n. Nevertheless, the occlusion tests are previously performed simply by examining occlusions of the center points of those light sources.[1]

The important point to note is that all objects casting shadows from extended light sources produce three types of regions: completely illuminated, completely occluded (*umbra*), and partially occluded (*penumbra*). As illustrated in Figure 5.3, a penumbra surrounds an umbra and there is always a gradual change in intensity from a penumbra and an umbra. The previous approach can treat only hard edged shadows correctly, and the approximation with point light sources introduces inaccuracy in illumination estimation in the presence of soft edged shadows, i.e., shadows with umbra and penumbra. Mixing up the umbra and the penumbra area not only prevents us from computing accurate coefficients a_i in (4.7) but also reduces the varieties of occlusions observed in a shadow

[1] As described in Section 5.1, occlusion coefficients $S(\theta_i, \phi_i)$ becomes 0 if the center point of L_i is occluded by the occluding object, and $S(\theta_i, \phi_i) = 1$ otherwise.

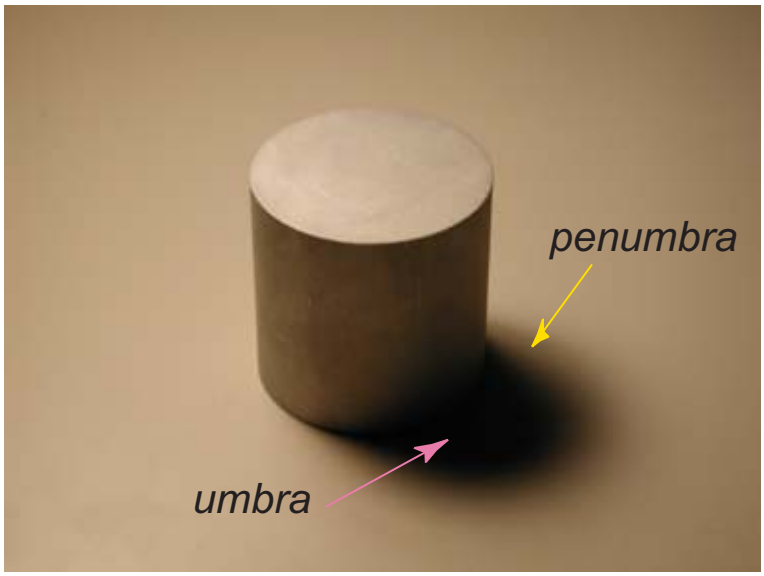


Figure 5.3: Umbra and penumbra.

image.

One approach to reduce the side-effect from this approximation is to increase the number of point light sources used for approximating illumination distribution and to compute occlusion coefficient $S(\theta_i, \phi_i)$ more accurately between 0 and 1, cf., 0 or 1.[2]

In our actual implementation of this super-sampling scheme, a hexagonally packed grid is defined inside each light source and used for computing the occlusion coefficients. The size of a hexagonally packed grid for a light source is determined from the solid angle of the light source. Figure 5.4 shows a hexagonal grid, made up of 19 hexagons, used in our experiments. For each point inside the hexagonal grid, we determine whether a light ray coming from the point is occluded by an occluding object. Then, occlusion coefficient $S(\theta_i, \phi_i)$ for the light source is given as a ratio of unoccluded sampling points inside the hexagonal grid.

[2] In theory, any subtle change in illumination distribution can be approximated with sufficient accuracy if an infinite number of point light sources is used. However, the computational cost would increase prohibitively if too many light sources were to be used. Moreover, the use of too many point light sources causes the sampling resolution problem described in Section 5.3.

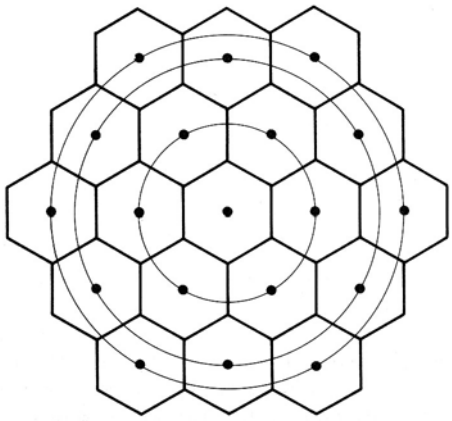


Figure 5.4: Hexagonal grids used for supersampling inside one light source.

5.7 Pixel Selection

First, the visible part of shadow surface is partitioned into clusters based on the combinations of occlusion coefficients $S(\theta_i, \phi_i)$. In other words, pixels that have the same combination of coefficients $S(\theta_i, \phi_i)$ are clustered into the same group. Conceptually, those clusters are similar to partitioned regions in shadows in Figure 5.2. At the same time, the number of pixels in one group is examined so that we can avoid selecting a pixel from a group that contains fewer image pixels.

Figure 5.5 shows several examples of partitioning based on the occlusion coefficients $S(\theta_i, \phi_i)$ with pseudo colors. Here each color represents an individual class with a different combination of occlusions of the light sources, and the block region corresponds to the occluding objects. From these examples of partitioning, we also see that patterns of partitioning differ from image to image depending on factors such as the shape of an occluding object and the camera viewpoint.

After the shadow surface is partitioned into clusters, one pixel is selected from each of those clusters. By selecting image pixels in this way, we can maximize variation of patterns of occlusion of light sources by an occluding object, and therefore, we are able to evaluate the input shadow image in an appropriate manner. In addition, we are able to avoid selecting redundant pixels, i.e., pixels that provide the same information about the illumination of the scene as other pixels.

For evaluating an input shadow image correctly, it is essential to select image pixels from a shadow image to maximize variation of patterns of occlusion of light sources by an occluding object.

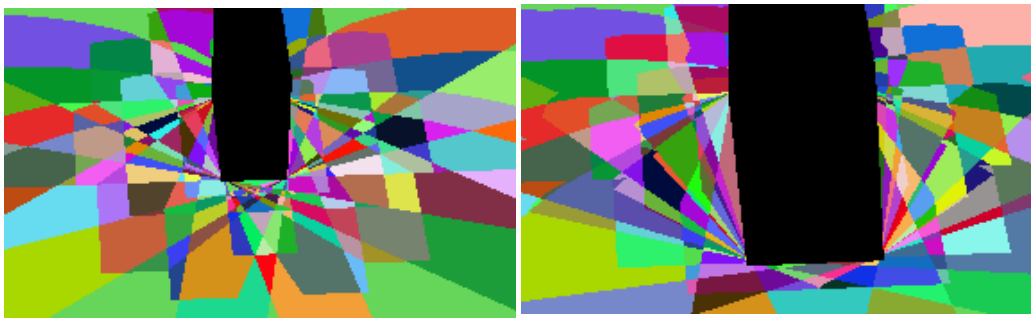


Figure 5.5: Clustering results.

5.8 Experimental Results

We have tested those proposed techniques for the robust estimation by using real images taken in both a laboratory and an office. First, we have tested our proposed method by using images taken under a relatively simple illumination environment so that we can examine the performance of our proposed method carefully. Figure 5.6 shows our experimental setup with three light sources with different colors. The input images taken by using this setup with different occluding objects are shown in Figure 5.7 (1).

The results of illumination estimation with coarse samplings are shown in Figure 5.7 (2). Here, the estimated illumination radiance is visualized by mapping their values onto a hemisphere shown in Figure 4.1. As we can see in these results, illumination distribution was correctly estimated for this set of light source samplings. However, after we increased the number of light sources for sampling illumination distribution, the estimation became unstable. Figure 5.7 (3) shows densely distributed light sources and in fact, we can see that the results of estimation shown in this column are erroneous especially around the blue and green light sources.

On the other hand, Figure 5.7 (4) shows the result of illumination estimation with our proposed method. Their stability measure is represented in the left of Figure 5.7 (4). Here, brighter light sources represent more reliable sampling regions which are not required to

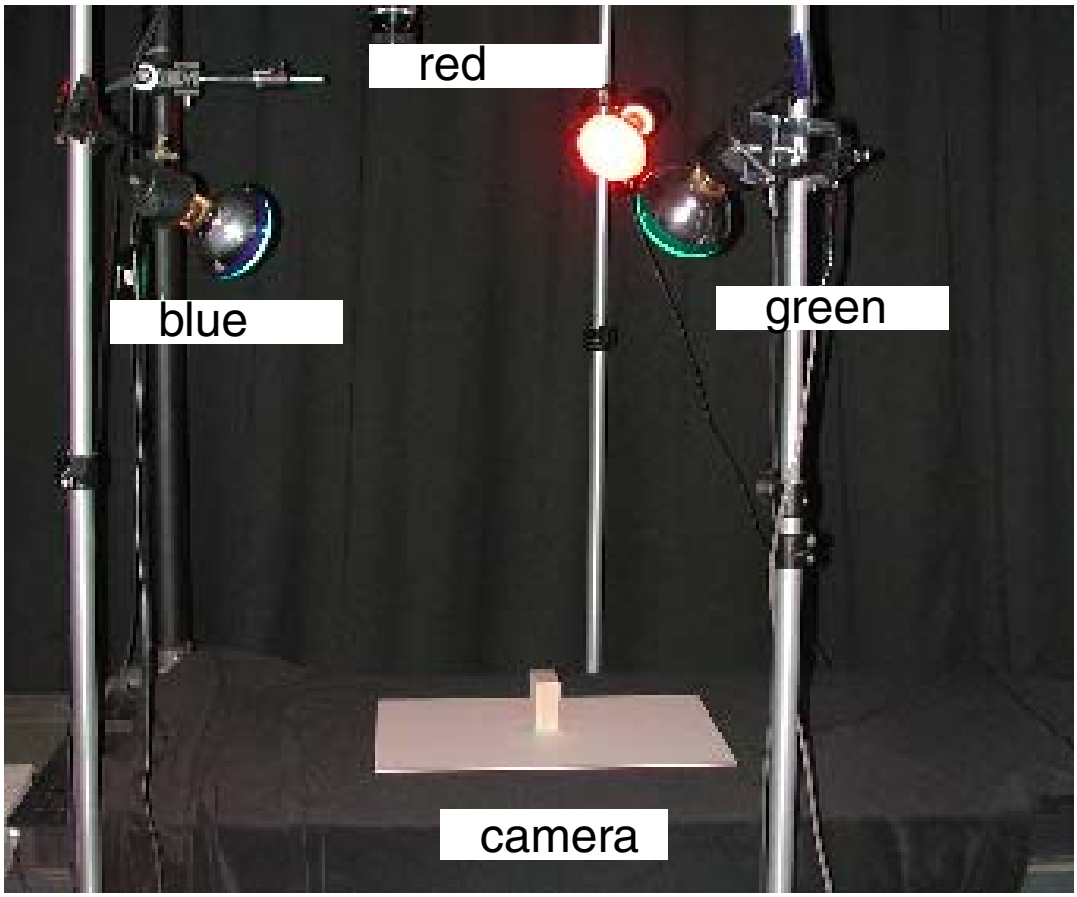


Figure 5.6: Experimental setup with three area light sources with different colors.

be merged to form a larger light source for more coarse sampling. Unlike the estimation results in Figure 5.7 (3), illumination distribution due to three light sources with different colors was correctly estimated in the right of Figure 5.7 (4).

Based on the estimated illumination distribution from the input shadow image shown Figure 5.8 (1), several synthetic objects were superimposed onto a synthetic surface in the bottom row of Figure 5.8 (1). Here, real objects with the same shape as that of the synthetic objects and shadows cast by those objects are shown in the top row for the comparison purpose.

In the proposed approach, since the entire illumination distribution of a scene is represented as a set of area light sources whose solid angle is adjusted depending on its stability, the distribution of shadows is a little different from those of real shadows cast by the real objects. However, it is found through our experiments that if we instead used a

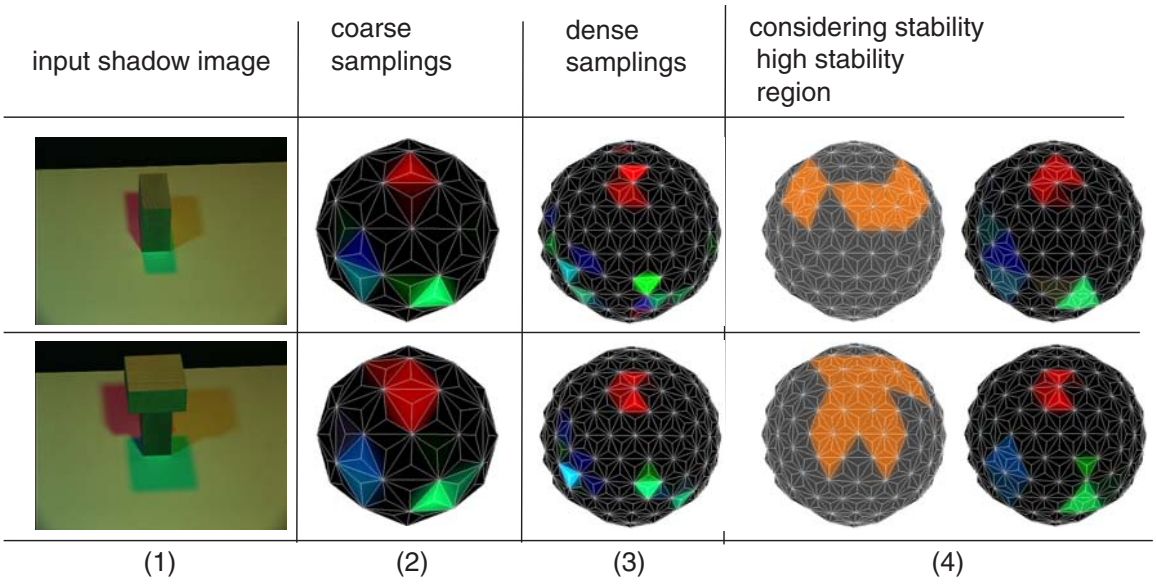


Figure 5.7: Illumination estimation without and with considering stability.

set of point light sources and examined only occlusions of the center points of those light sources, the entire estimation became unstable unless the method happened to find the correct locations for those three light sources with coarse sampling.

We have also tested our proposed method by using real images taken in an ordinary office environment. In the bottom row of Figure 5.8 (2), several synthetic objects were also superimposed onto the surface using the illumination distribution estimated from the input shadow image shown in this figure. It is worth noting that in this example, a relatively large area of the shadow surface is occluded by the occluding object, and it is often difficult to provide a correct estimate of the illumination distribution in such case. Even in this challenging case, our proposed approach could reliably estimate the illumination distribution of the scene by taking stability issues into considerations. Shadows cast by those synthetic objects resemble well those cast by the real objects, and this shows that the estimated illumination distribution gives a good presentation of that of the real scene.

5.9 Conclusions

In this chapter, an improved approach for estimating illumination distribution of a real scene from shadows in a given image in a reliable manner has been presented. First, the source of instability of the previously described *illumination-from-shadows* method was carefully studied in several aspects: invisible regions of shadows; limited sampling resolution for radiance distribution inside shadows; and approximation of illumination distribution as a collection of point light sources.

Then, based on that analysis, a robust method to overcome those problems has been presented. The proposed method consists of three components. For estimating the illumination distribution of a scene reliably by taking stability issues into considerations, we proposed to change the sampling density of the illumination distribution depending on the amount of the information obtainable from a shadow image for a particular direction of the illumination distribution. For using radiance distribution inside penumbra of shadows correctly, we introduced a super-sampling scheme for examining occlusion of incoming light from each light source. We also explained the optimal sampling of image pixels and the selection of illumination distribution samplings for more stable computation.

All of these extensions contribute to improve stability and accuracy of illumination estimation from shadows, and illumination distribution can be estimated in a reliable manner with these proposed improvements regardless of types of input images such as the shape of an occluding object or a camera position.

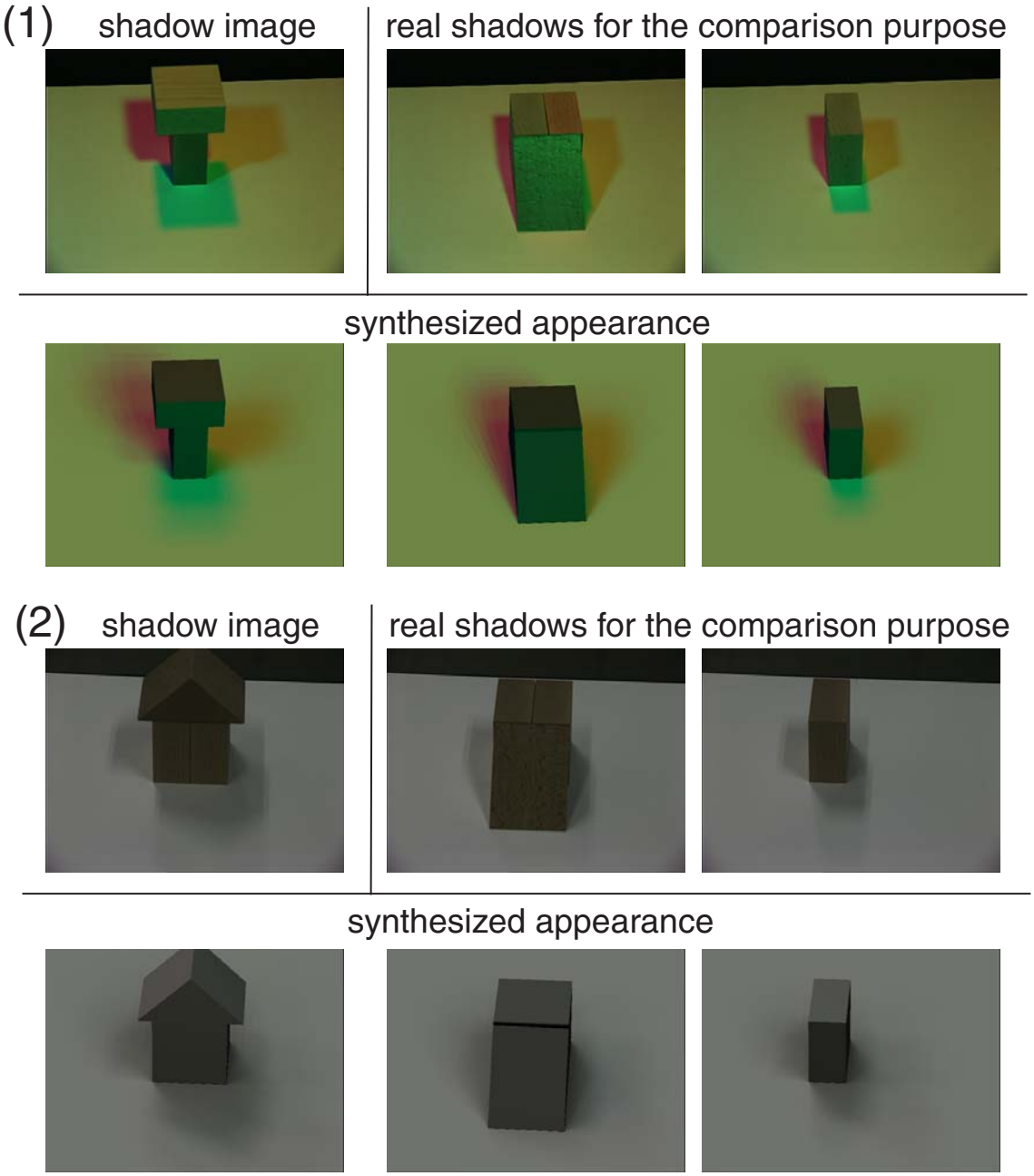


Figure 5.8: Synthesized appearance using the estimated illumination distribution.

Chapter 6

Modeling Appearance of Objects for Variable Illumination

Inverse rendering carries out the opposite procedures of model-based rendering to provide object and illumination models of a real scene from photographically available information of the scene. Once models of a scene are acquired, new images of the scene under novel lighting and/or viewing conditions can be synthesized by using conventional model-based rendering techniques.

On the other hand, the approach called *image-based rendering* directly uses the original set of input images of a scene for producing new images of the scene under novel conditions [61]. Depending on which scene conditions should be modified, image-based rendering techniques are classified into three categories: image-based rendering under novel viewing conditions, image-based rendering under novel lighting conditions, and image-based rendering under novel viewing and novel lighting conditions. In this thesis, we consider the second category, image-based rendering under novel lighting conditions.

In contrast with model-based rendering techniques, image-based rendering techniques do not require full radiometric computation to synthesize the photo-realistic appearance of objects in a scene. This makes the cost to produce new images of the scene independent of the scene complexity. Also image-based rendering techniques normally do not require geometric and photometric models of a scene.[1] Image-based rendering, however, has a tendency to require many input images of a scene to synthesize reasonably realistic

[1] Some image-based rendering techniques make use of geometric models of a scene for better compression of its appearance.

appearance of the scene. This results in the requirement for a large amount of both computer memory and data storage.

The main goal of this chapter is determining a set of lighting directions for sampling images of an object so that a set of basis images for representing the appearance change of the object under varying illumination conditions can be obtained from those images.

6.1 Related Works

It is well known that the appearance of an object changes significantly under different illumination conditions. For instance, the appearance change of someone's face often becomes much larger than the difference of two different faces under the same lighting. Thus, for the task of object recognition and image synthesis, it is very important to be able to predict the variation of objects' appearance under varying illumination conditions.

While there may seem to be a large variety of possible appearances for a given object, it has been demonstrated in previous research that the changes in appearance of an object for varying illumination can be represented with a linear subspace spanned by a set of basis images of the object. For instance, in the most simplistic case of a convex Lambertian object under distant illumination without attached and cast shadows, the appearance of the object can be completely described with a 3-D linear subspace defined with three input images taken under linearly independent lighting conditions [42, 58, 73]. However, the assumption of this model would be too restrictive to be used for object recognition in more realistic settings.

Other researchers have reported empirical studies for representing image variation due to varying illumination for human faces and other objects [21, 14, 72]. Interestingly enough, most of the image variation caused by varying illumination can be explained with a low-dimensional linear subspace slightly higher than 3D even when images contain a significant amount of shadows. For instance, Hallinan reported a) that a 5-D subspace would suffice to represent most of the image variation due to illumination change including extreme cases where a face is lit from its sides, and b) that a 3-D subspace would be sufficient when a face is lit mainly from its front [21]. Georghiades et al. used a similar observation more specifically for object recognition under varying lighting conditions [18, 19].

A set of basis images spanning such a linear subspace can be obtained by applying the

principal component analysis to a large number of images of an object taken under different lighting, e.g., by moving a point light source along a sphere surrounding the object. One might ask whether a certain set of input images of an object would be sufficient to fully span the subspace of the object for arbitrary illumination conditions. Previous empirical studies do not necessarily provide enough insight to this important question, and thus it has been a common practice to use as many input images as possible to ensure that the set of input images span a subspace entirely.

Another interesting observation of Hallinan's early work is that a subspace obtained by PCA does not vary widely between sparsely and densely sampled sets of lighting directions in the case of faces [21]. However, it was not known how many images would be sufficient in order to obtain the basis images correctly.

Recently, it was shown through the frequency-space analysis of reflection that appearance of a convex Lambertian object can be well represented with a 9-D linear subspace spanned by basis images of the object, called *harmonic images*, each of which corresponds to an image of the object illuminated under *harmonic lights* whose distributions are specified in terms of spherical harmonics[52, 53, 2]. Basri and Jacobs successfully used this 9-D subspace defined with harmonic images for face recognition under varying illumination [2]. Other researchers have also used a set of harmonic images for the purpose of efficient rendering of an object under complex illumination [54, 62]. More recently, Ramamoorthi has shown theoretically that, under certain conditions of lighting and object shape and reflectance, a set of basis images obtained from PCA on input images taken under varying lighting coincides with a set of harmonic images [55].

While harmonic images are specified analytically, it is difficult to obtain harmonic images for various kinds of real objects because harmonic lights consist of both negative and positive values distributed in a complex way and thus do not exist as real lighting. Therefore, the previously proposed techniques for object recognition and image synthesis based on harmonic images require the model of an object about its 3-D shape and reflectance properties such as albedo so that harmonic images of the object under harmonic light can be rendered synthetically. This is not a problem for the case of synthetic objects because the model of a synthetic object is given a priori by definition. On the other hand, acquiring shape and reflectance properties of real objects is not an easy problem and therefore, most of the previous studies treated real objects as Lambertian objects and estimated only

albedos by using vision techniques such as photometric stereo.

This motivated us to develop a method for determining a set of harmonic images of a real object by using only input images of the object under simple lighting such as a point light source. Lee et al. recently proposed an interesting method to determine a configuration of 9 light source directions such that input images taken under those light source directions approximate a 9-D subspace spanned by harmonic images [36]. They have reported that such a configuration of 9 light source directions does not vary much for similar objects, e.g., different faces. However, Lee et al.'s method still needs a set of harmonic images that are rendered synthetically by using an object's model in order to determine a set of 9 lighting directions. Moreover, Lee et al.'s method chooses 9 light directions from a large number of candidates and therefore, a large number of input images are required for each new object or new class of objects.

6.2 Proposed Appearance Sampling Approach

In this chapter, we present a novel method for analytically obtaining a set of basis images of a convex object for arbitrary illumination from input images of the object taken under a point light source. The main contribution of our work is that we show that a set of lighting directions can be determined for sampling images of an object depending on the spectrum of the object's BRDF in the angular frequency domain such that a set of harmonic images can be obtained analytically based on the sampling theorem on spherical harmonics [13].

Using those sampling directions determined from the sampling theory, we are able to obtain harmonic images by using a significantly smaller number of input images than other techniques which do not take into account a relationship between a spectrum of BRDFs and a sampling density of illumination directions. In addition, unlike other methods based on spherical harmonics, our method does not require the shape and reflectance model of an object used for rendering harmonics images of the object synthetically. Thus, our method can be easily applied for determining a set of basis images for representing the appearance change of a real object under varying illumination conditions.

The rest of the chapter is organized as follows. We briefly review the spherical harmonics transformation of a function over the unit sphere and harmonics images based on spherical harmonics in Section 6.3. We show details of the sampling theorem used in our

method for obtaining harmonic images of real objects in Section 6.4, and consider issues in obtaining harmonic images of real objects based on this sampling theorem in Section 6.5. Finally, we show experimental results of the proposed method applied to images of real objects in Section 7.4.2, and present concluding remarks in Section 6.7.

6.3 A Set of Basis Images for Variable Illumination

6.3.1 Review of Spherical Harmonics

Spherical Harmonics define an orthonormal basis over the unit sphere. Consider the unit sphere in R^3, a unit vector on the sphere can be described by the polar coordinate system $\theta, (0 \leq \theta \leq \pi)$ in elevation and $\phi, (0 \leq \phi < 2\pi)$ in azimuth as shown in Figure 6.1. In this coordinate system, spherical harmonics $Y_l^m(\theta, \phi), (l \geq 0, -l \leq m \leq l)$ are defined as

$$Y_l^m(\theta, \phi) = N_l^m P_l^m(\cos\theta) e^{Im\phi}, \tag{6.1}$$

where N_l^m are the normalized constants, and $P_l^m(\cdot)$ are the associated Legendre functions of degree l and order m.

A function $f(\theta, \phi)$ defined over the unit sphere is expanded as a linear combination of spherical harmonics as

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_l^m Y_l^m(\theta, \phi), \tag{6.2}$$

and f_l^m denote coefficients in its spherical harmonic expansion computed as [2]

$$f_l^m = \int_0^{2\pi} \int_0^{\pi} f(\theta, \phi) Y_l^m(\theta, \phi) \sin\theta \mathrm{d}\theta \mathrm{d}\phi. \tag{6.3}$$

6.3.2 Harmonic Image Representation

The reflectance property of an object is characterized by a bidirectional reflectance distribution function (BRDF) $\rho(\theta_i', \phi_i', \theta_o', \phi_o')$, where (θ_i', ϕ_i') and (θ_o', ϕ_o') are incident and reflection directions with respect to the surface normal of the object surface as shown in Figure 6.2 whose local coordinate is denoted by using $'$ Namely, the north pole direction $(0', 0')$ in the local coordinates corresponds to the surface normal of the object surface.

[2] In this study, we consider spherical harmonics in a real form.

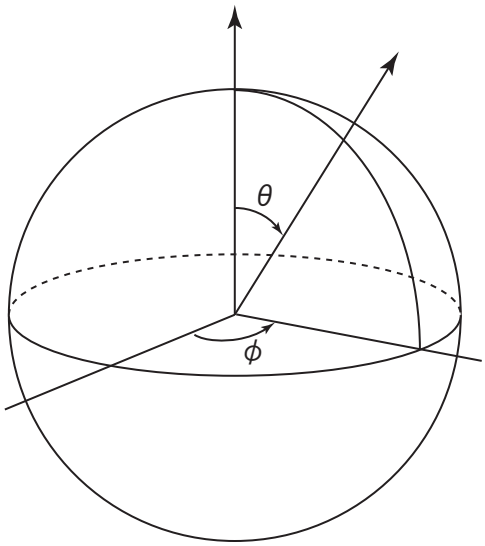


Figure 6.1: Polar coordinate system

Then, the brightness E of the object surface with a surface normal at the direction (θ_n, ϕ_n) is represented by the global coordinate defined on the unit sphere as

$$E = \int_0^{\pi} \int_0^{2\pi} L(\theta, \phi) \rho(M_{\theta_n}^{\phi_n}(\theta, \phi), \theta'_o, \phi'_o) \cos\theta \sin\theta \mathrm{d}\theta \mathrm{d}\phi, \tag{6.4}$$

where $L(\theta, \phi)$ is light source distribution, and $M_{\theta_n}^{\phi_n}(\cdot)$ represents a rotation operator that rotates (θ, ϕ) into the local coordinate defined with respect to the surface normal of the object surface.

In this study, we consider the appearance of an object under variable illumination seen from a fixed viewpoint and therefore, we represent $\rho(M_{\theta_n}^{\phi_n}(\theta, \phi), \theta'_o, \phi'_o) \cos\theta$ by using the global coordinate as $R(\theta, \phi)$, and refer to it as the *reflection kernel*.

Since both the light source distribution and the reflection kernel are functions defined on the unit sphere, we can represent them as

$$L(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} L_l^m Y_l^m(\theta, \phi), \tag{6.5}$$

$$R(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} R_l^m Y_l^m(\theta, \phi). \tag{6.6}$$

where L_l^m and R_l^m are coefficients in their spherical harmonic expansion from (6.3). Assuming spherical light sources with $Y_l^m(\theta, \phi)$ radiance in its (θ, ϕ) direction, R_l^m represents the brightness seen under these spherical light sources called harmonic lights in (6.3).

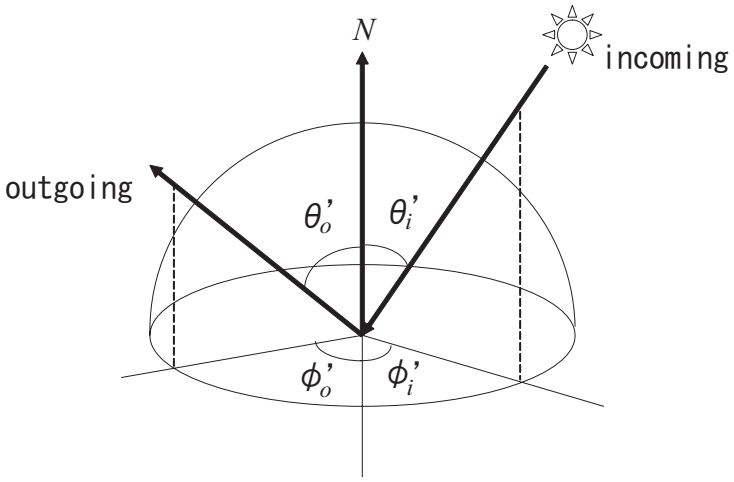


Figure 6.2: BRDF defined in the polar coordinate system

From (6.5), (6.6), and the orthonormality of spherical harmonics, the surface brightness E in (6.4) is represented as

$$E = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} L_l^m R_l^m. \tag{6.7}$$

Here, if we prepare R_l^m for all pixels, that is, for all corresponding points on the object surface, images under variable illumination are represented from (6.7). Images containing R_l^m are called harmonic images [2].

6.4 Methods for Obtaining Harmonic Image

There are several approaches for obtaining harmonic images of an object. One approach is to provide a reflection kernel $R(\theta, \phi)$ of the object from the knowledge of its 3D shape and reflectance properties. Since $Y_l^m(\theta, \phi)$ are predefined functions, R_l^m are computed from (6.3). Most of the previously proposed methods employed this approach to compute harmonic images synthetically.

One might think of observing an object under physically constructed harmonic lights. In this case, each pixel value directly corresponds to coefficients R_l^m, and thus the 3D shapes and reflectance properties of the object need not be given a priori. However, as pointed out by other researchers [2, 36], harmonic lights are complex diffuse lighting consisting of both negative and positive radiance. Thus, it is difficult to physically construct

such lighting in a real setting.

In this work, we take an alternative approach of observing a reflection kernel $R(\theta, \phi)$ for each surface point on an object directly by taking an image of the object under a point light source located at the direction (θ, ϕ). This approach is based on the assumption that a point light source used in this method is sufficiently distant from the objects and thus, this light source projects parallel rays onto the object surface.

The reflectance kernel $R(\theta, \phi)$ represents radiance of reflected light due to incoming light with unit radiance from the direction (θ, ϕ). Thus if we suppose that this point light source has unit radiance, the value of $R(\theta, \phi)$ can be obtained simply as an observed brightness of the surface point when the object is illuminated under a light source from the direction (θ, ϕ).[3] Once we determine the values of $R(\theta, \phi)$ for all directions, we can compute pixel values of harmonic images as R_l^m from (6.3). In this way, we do not need to synthetically compute the reflection kernel $R(\theta, \phi)$ of the object, nor are the 3D shapes and reflectance properties of the object required.

Since the function $R(\theta, \phi)$ are continuously distributed on the unit sphere in (6.3), we first need to approximate its distribution by a discrete set of the function $R(\theta, \phi)$ so that we can sample $R(\theta, \phi)$ using a point light source physically located at (θ, ϕ) direction. Then, the question we have to ask is how densely $R(\theta, \phi)$ need be sampled in order to correctly compute coefficients R_l^m from them. In other words, we want to know how densely a point light source needs to be positioned around an object to obtain harmonics images of the object correctly. In the next section, we will consider this issue in terms of the characteristics of spherical harmonics $Y_l^m(\theta, \phi)$.

6.4.1 Sampling Theorem on Spherical Harmonics Transformation

There have been several methods proposed in the field of applied mathematics to efficiently compute coefficients of a function $R(\theta, \phi)$ in its spherical harmonic expansion using fewer samplings of the function $R(\theta, \phi)$. We have adapted one of their theories to enable us to compute harmonic images using fewer input images of objects taken by moving a point light source to particular locations.

It is common knowledge that the sampling theorem on the 1D line tells us that a

[3] Note that we can determine $R(\theta, \phi)$ only up to some unknown scaling factor, so it is reasonable to treat a point light source to have unit radiance.

band-limited function can be reconstructed accurately from properly sampled discrete data. Namely, Fourier coefficients of the function can be determined by weighted sums of the function sampled at even intervals. In the case of a function defined on the sphere, the similar theorem, the sampling theorem on the sphere, has been proved [13]. In this section, we outline the theorem.

Let us assume that the reflection kernel $R(\theta, \phi)$ is *band-limited* with *bandwidth* N, that is, $R_l^m = 0$ $(l \geq N)$. Then, consider the "comb" function $s(\theta, \phi)$ with equiangular grid $(\Delta\theta, \Delta\phi) = (\pi/2N, 2\pi/2N)$

$$s(\theta, \phi) = \frac{\sqrt{2\pi}}{2N} \sum_{j=0}^{2N-1} \sum_{k=0}^{2N-1} w_j \delta(\theta - \theta_j)\delta(\phi - \phi_k), \tag{6.8}$$

where $(\theta_j, \phi_k) = (\pi(j + 1/2)/2N, 2\pi k/2N)$ are sampling points on the sphere, and w_j are weight of the points. We can compute the coefficients s_l^m of the comb function from definition (6.3). For degree $l < 2N$, we obtain

$$s_l^m = \sum_{j=0}^{2N-1} w_j \sqrt{\frac{2l+1}{2}} P_l^0(\cos\theta_j)\delta_{m0}, \tag{6.9}$$

where the Kronecker delta $\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$ if $i \neq j$. Here, we can uniquely choose the weight w_j so that $\sum_{j=0}^{2N-1} w_j P_l^0(\cos\theta_j) = \sqrt{2}\delta_{l0}$. Thus, the coefficients are described simply as $s_l^m = \delta_{l0}\delta_{m0}$ for degree $l < 2N$. Equivalently, the comb function is represented by the addition of $Y_0^0(\theta, \phi) = const.$ and higher-degree terms as

$$s(\theta, \phi) = 1 + \sum_{j \geq 2N} \sum_{|k| \leq j} s_j^k Y_j^k(\theta, \phi). \tag{6.10}$$

Then, from (6.6) and (6.10), the product of the reflection kernel and the comb function is written as

$$\begin{aligned} R(\theta, \phi) \cdot s(\theta, \phi) &= R(\theta, \phi) \\ &+ \sum_{l<N} \sum_{|m| \leq l} \sum_{j \geq 2N} \sum_{|k| \leq j} R_l^m s_j^k Y_l^m(\theta, \phi) Y_j^k(\theta, \phi), \end{aligned} \tag{6.11}$$

where degree $l < N$ because $R(\theta, \phi)$ is band-limited. The second term is known as *aliasing* introduced by discrete sampling. However, it is known that the product of the spherical harmonics $Y_l^m(\theta, \phi)Y_j^k(\theta, \phi)$ is represented as a linear combination of spherical harmonics with a degree greater than or equal to $|l - j|$. Accordingly, aliasing appears to be a

degree greater than or equal to $|N - 2N| = N$ in this case. Therefore, for degree $l < N$, $R(\theta, \phi) \cdot s(\theta, \phi)$ is equal to $R(\theta, \phi)$, that is,

$$(R(\theta, \phi) \cdot s(\theta, \phi))_l^m = R_l^m. \tag{6.12}$$

Hence, the coefficients of the reflection kernel can be computed accurately by the Fourier transform of $R(\theta, \phi) \cdot s(\theta, \phi)$ as

$$R_l^m = \frac{\sqrt{2\pi}}{2N} \sum_{j=0}^{2N-1} \sum_{k=0}^{2N-1} w_j R(\theta_j, \phi_k) Y_l^m(\theta_j, \phi_k), \tag{6.13}$$

where the weight w_j are analytically given by

$$w_j = \frac{2\sqrt{2}}{2N} \sin\theta_j \sum_{n=0}^{N-1} \frac{1}{2n+1} \sin[(2n+1)\theta_j]. \tag{6.14}$$

6.5 Appearance Sampling of Real Objects based on Sampling Theorem

The sampling theorem described in the previous section tells us the minimum number of sampling $2N \times 2N = 4N^2$ to compute spherical harmonics transformation of a band-limited function with bandwidth N. In this section, we consider issues in defining bandwidth of an object. We will later discuss what kind of artifacts we should expect when the function is not band-limited within N in Chapter 7.

6.5.1 Convex Lambertian Surface

Let us start with the simplest case of convex Lambertian objects. It has been shown in the previous studies that the first nine spherical harmonics with the order $l = 2$ are sufficient to capture more than 99% of the reflection energy of a convex Lambertian surface [2, 52, 53].

Accordingly, we can consider that the function $R(\theta, \phi)$ is band-limited with bandwidth $N = 3$, and this results in $4N^2 = 36$ samplings of $R(\theta, \phi)$ necessary for computing R_l^m correctly. In other words, the coefficients R_l^m are given as a finite weighted sums of the function $R(\theta, \phi)$ sampled at equiangular grid: $\theta_j = \frac{\pi(j+1/2)}{6}$ $(j = 0, \ldots, 5)$, $\phi_k = \frac{2\pi k}{6}$ $(k = 0, \ldots, 5)$. Namely, 36 input images of an object taken by moving a point light source to the directions specified with (θ_j, ϕ_k) on a sphere around the object are required to compute harmonic images of the object.

Note that what needs to be satisfied is not the number of samplings ($4N^2$) of the function but rather, the sampling intervals ($\Delta\theta = \pi/2N$, $\Delta\phi = 2\pi/2N$) that this sampling can provide. For instance, even when a large number of input images of an object taken under a point light source are available, there is no guarantee that this set of images can produce correct harmonic images that can span a low-dimensional linear subspace representing the appearance of the object under arbitrary illumination unless those intervals are satisfied.

6.5.2 Complex Appearance beyond Lambertian

Ramamoorthi and Hanrahan analytically derived the bandwidth of reflection kernels of objects that have more complex appearance than convex Lambertian surfaces such as the Phong reflection model and the Microfacet BRDF [52].[4] For instance, this study shows that the bandwidth of the Microfacet BRDF is approximately $N \approx \sigma^{-1}$. Thus, if the surface roughness of an object is predicted even roughly, it should help us to find the bandwidth of the reflection kernel of the object.

There are a large number of previous studies on BRDF measurements and BRDF databases which show reflectance parameters of various kinds of surface materials. This knowledge should be useful for estimating bandwidths of reflection kernels of various objects based on the analysis presented in [52].

6.6 Experimental Results

We have tested the proposed method using real images of several objects taken by moving a point light source to equiangular grid points defined by the sampling theorem. Figure 6.3 shows an overview of our hardware set-up[5] used for obtaining the input images of the objects; an array of light sources is mounted on a turntable. These light sources are equally spaced in elevation, and the set of light sources is rotated around the objects in azimuthal.

In the case of the sheep and the Venus examples, since those objects have an appearance similar to that of a Lambertian surface, 36 input images of them are taken under a point

[4] [56] is also a good reference to get insights into the bandwidth of a reflection kernel of various types of objects in CURET database. This work discusses appropriate sampling resolutions to model appearance of those objects based on their experiments as well.

[5] *Surface Reflectance Sampler*, TechnoDream21 corporation

Figure 6.3: Image acquisition set-up:an array of light sources is mounted on a turntable. These light sources are equally spaced in elevation, and the set of light sources is rotated around the objects in azimuthal.

light source positioned at equiangular grid points: $\theta_j = \frac{\pi(j+1/2)}{6}$, $\phi_k = \frac{2\pi k}{6}(j,k = 0,\ldots,5)$. For the fish example, since it has a more complex appearance, 144 input images are taken under a point light source at equiangular grid points : $\theta_j = \frac{\pi(j+1/2)}{12}$, $\phi_k = \frac{2\pi k}{12}(j,k = 0,\ldots,11)$. Based on the sampling theorem, coefficients R_l^m are computed up to the degree $l = 2$ from the 36 images, and up to the degree $l = 5$ from the 144 images.

The first nine harmonic images obtained from those input images are shown in Figure 6.4. In spite of given discrete sets of the appearance of the objects, the obtained harmonic images have complex and smooth shading that reflects the distribution of harmonic lights in these examples.

To evaluate the accuracy of the recovered harmonic images, we took images of those

objects under normal lighting conditions in our laboratory, including direct light sources such as fluorescent lamps. Coefficients L_l^m in (6.5) representing the illumination condition are computed from an omni-directional image of the scene taken by a camera with a fish-eye lens in the same manner as described in [54]. Then appearance of the objects under this illumination condition is synthesized from (6.7).

In Figure 6.5, the left column shows the real images of the objects and the right column shows the synthesized appearance. The synthesized appearance of the objects resembles that of the objects in real images, and this shows that the recovered harmonic images provide a good representation of the appearance of the objects. The shoulder of the plaster figure in the middle right image appears darker than that in the real image. This is due to severe cast shadows observed in the input images that do not follow our method's assumption for convex objects.

To demonstrate how well the recovered harmonic images represent the appearance of those objects under varying illumination, we synthesize their appearance from the recovered harmonic images under several natural illumination conditions provided by high-dynamic range light probe measurements by [7]. Figure 6.6 shows the results. In this figure, synthesized appearance changes dynamically depending on characteristics of the illumination, and one can say from this that the proposed method succeeded in providing a set of basis images representing appearance of those objects under varying illumination.

6.7 Conclusions

In this study, we have presented a novel method for analytically obtaining a set of basis images of an object for arbitrary illumination from input images of the object taken under a point light source. The main contribution of our work is that we have shown that a set of lighting directions can be determined for sampling images of an object depending on the spectrum of the object's BRDF in the angular frequency domain such that a set of harmonic images can be obtained analytically based on the sampling theorem on spherical harmonics.

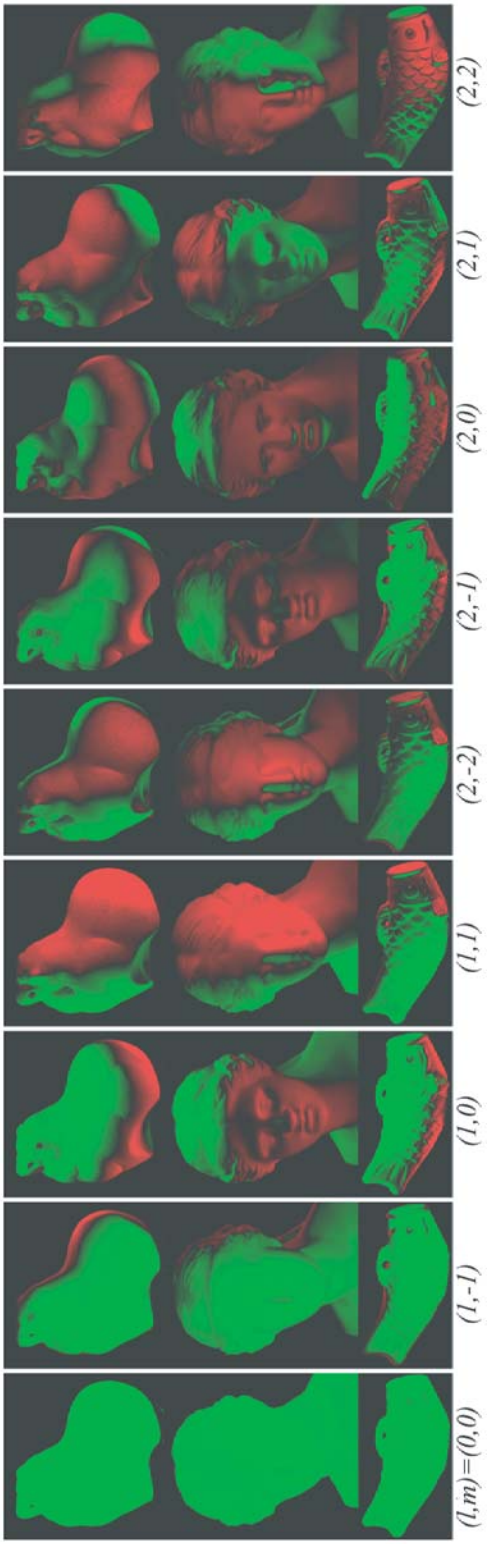


Figure 6.4: Obtained harmonic images: positive values are shown in green, and negative values are shown in red.

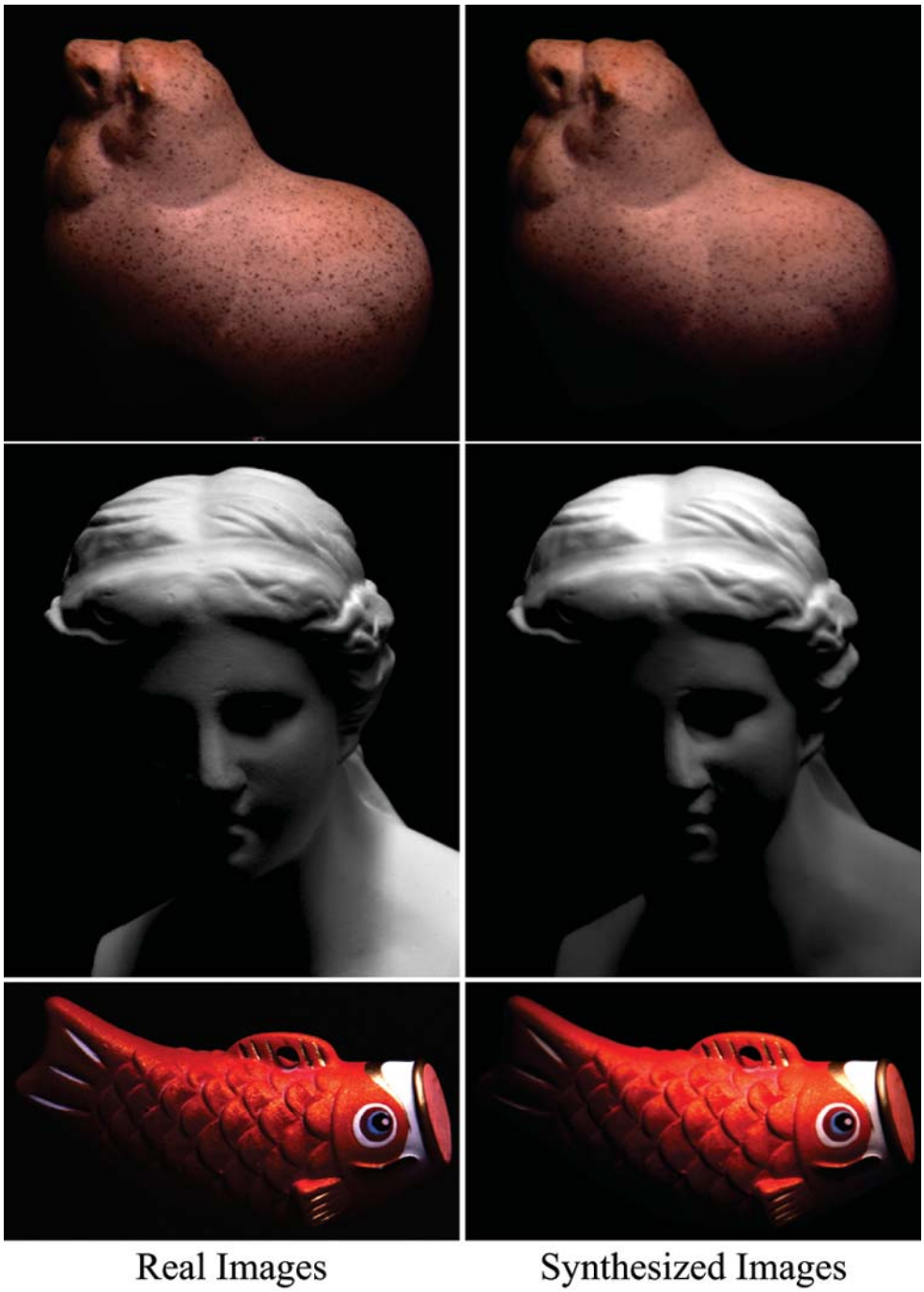


Figure 6.5: Comparison between real images and synthesized images under complex illumination

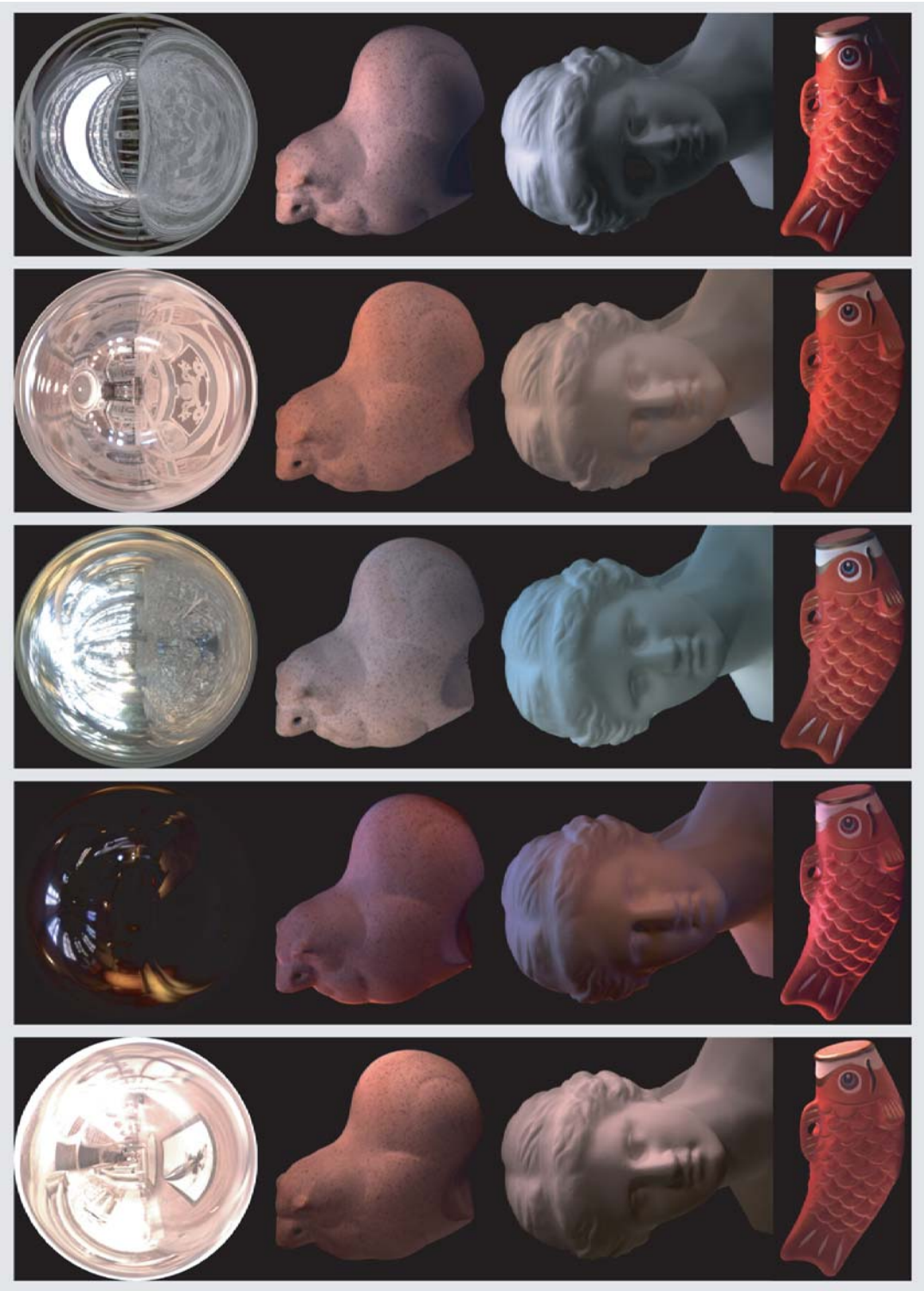


Figure 6.6: Synthesized images of objects under natural illumination. The first row shows illumination maps. The second, third and forth rows show synthesized appearance of objects under the corresponding illumination map.

Chapter 7

Anti-aliasing Method for Appearance Sampling of Real Objects

There certainly is a situation where bandwidth of reflection kernels of an object is difficult to predict, or where only a limited number of samplings of the function are obtainable depending on a particular hardware set-up used for acquiring input images. In this chapter, we carefully study the issues of aliasing in modeling appearance of real objects and extend the method based on the sampling theorem described in Chapter 6 further for reducing the artifacts due to aliasing, by substituting extended light sources for a point light source to sample the reflection kernel of a real object.

7.1 Aliasing Caused by Insufficient Sampling

First, let us consider what kinds of artifacts we should expect when the function $R(\theta, \phi)$ has a lower or a higher bandwidth than N determined from the number of samplings $2N \times 2N$ based on the sampling theorem.

Let B denote actual bandwidth of the function. Driscoll and Healy proved that the error in spherical harmonics transformation, generally known as aliasing, is confined to coefficients R_l^m of degree greater than or equal to $|B - 2N|$ in [13]. This can be confirmed by replacing the bandwidth N of the reflection function with B in (6.11).

From this, in the case where the function has bandwidth B lower than N ($B < N$),

there is no aliasing for degree less than $N+k$ for $B = N-k$. Since we compute coefficients R_l^m for $l < N$, those coefficients are correctly computed. In this case, all of the coefficients R_l^m for $B \leq l < N$ become 0.

In the case where the function $R(\theta, \phi)$ has a bandwidth B higher than N ($B > N$), error due to insufficient sampling of the function is confined to the coefficients of degree $l \geq N - k$ for $B = N + k$. In this case, the coefficients R_l^m for $l < N - k$ degree are correctly computed. Therefore, closer to B we select N, smaller error in R_l^m we expect.

Next we show what happens when we have sufficient or insufficient samplings of the function $R(\theta, \phi)$ by using synthetic data. The function $R(\theta, \phi)$ used here is a reflection kernel for the Torrance-Sparrow reflection model with known reflection parameters ($K_d = \{.62, .71, .62\}, K_s = 0.29, \sigma = 0.1$).[1] The surface normal and viewing direction for this reflection kernel are set at the direction ($\theta = 0, \phi = 0$).

We computed the coefficients R_l^m up to the degree $B = \sigma^{-1} \approx 10$ for this reflection kernel from discrete samplings of $R(\theta, \phi)$ at different sampling intervals: $N = 200$ (160000 sampling), $N = 11$ (484 samplings), $N = 5$ (100 samplings), $N = 3$ (36 samplings). The computed coefficients are shown in Figure 7.1a)∼d). Here spherical harmonic coefficient for given degree l and order m are represented using a single index $r = l^2 + l + m$. The horizontal axis represents the index r, and the vertical axis represents the computed coefficients R_l^m. The number of samplings is shown for each graph, and a red arrow represents the upper bound N determined from the number of samplings.

Figure 7.1e)∼h) are visualization of the reflection kernel $R(\theta, \phi)$ reconstructed from the computed coefficients R_l^m in (6.6) up to the degree $l = 10$ for the cases of $N = 200$, $N = 11$, $N = 5$, and $N = 3$, respectively. Figure 7.1i) shows a reflection kernel reconstructed from the coefficients up to the degree $l = 2$ for the cases of $N = 3$. In these figures, the upper half of $R(\theta, \phi)$ is visualized in a polar coordinate system with radius indicating $\{\theta | 0 \leq \theta \leq \pi/2\}$, and angle indicating $\{\phi | 0 \leq \phi < 2\pi\}$ as illustrated in Figure 7.1j).

From Figure 7.1a) and b), we can see that the coefficients R_l^m obtained from $4N^2 = 484$ samplings are almost the same as those obtained from $4N^2 = 160000$ samplings, and the $R(\theta, \phi)$ is reconstructed correctly in both cases. This shows that a set of samplings for $N = 11$ are sufficient to capture the appearance of a surface with this reflection kernel

[1] K_d and K_s are constants for the diffuse and specular reflection components, and σ is the standard deviation of a facet slope of a simplified Torrance-Sparrow reflection model.

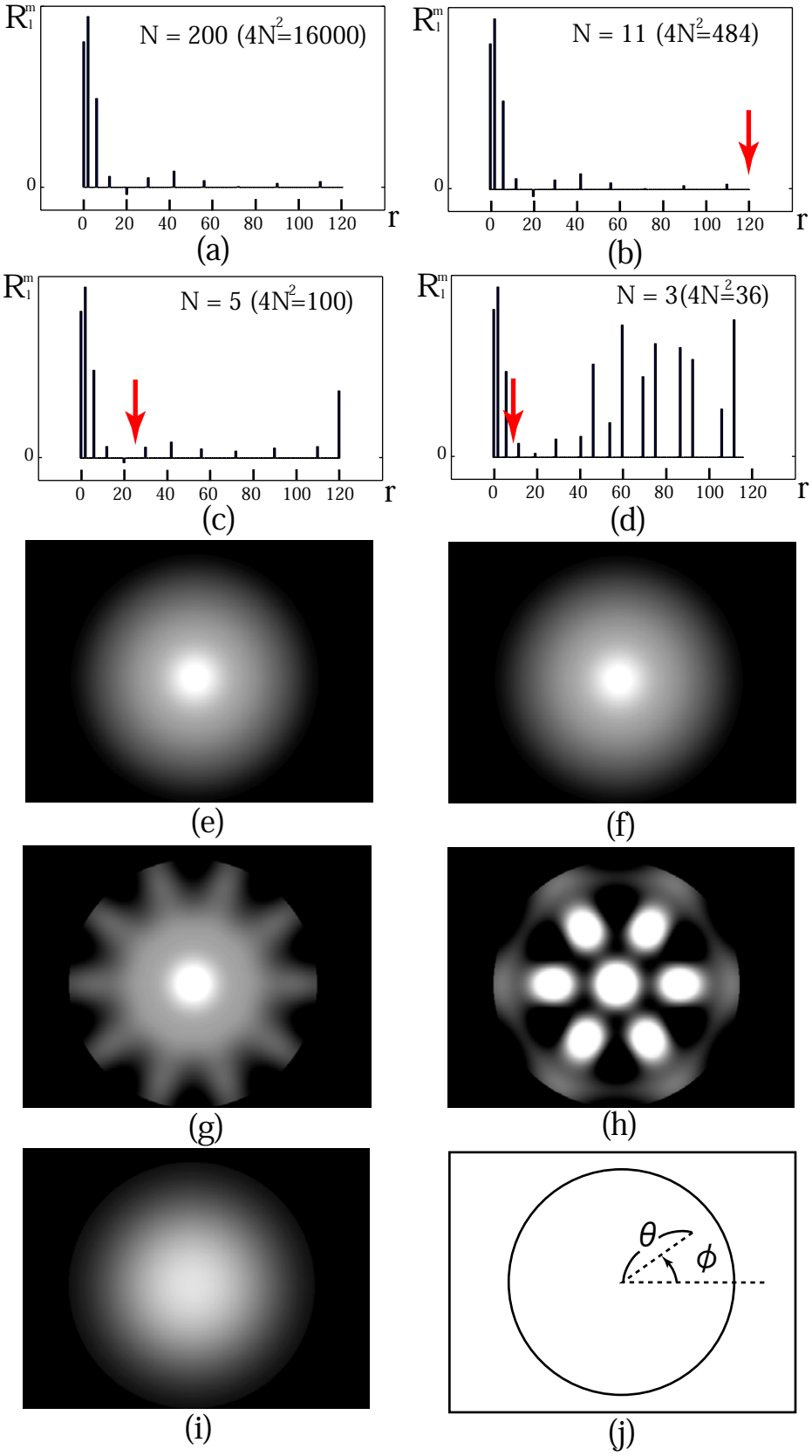


Figure 7.1: Experimental results using synthetic data: coefficients are computed from a discrete set of $f(\theta, \phi)$ sampled at different sampling intervals.

under varying illumination. This follows what the sampling theorem tells us.

On the other hand, in the cases of $N = 5$ and $N = 3$ where N is smaller than $B = 10$, the computed coefficients suffer from aliasing. It is important to note that the coefficients R_l^m for $l < N$ (the left side of the red arrow in Figure 7.1c) and d)) closely match the coefficients obtained from sufficient samplings (Figure 7.1a)) while the coefficients R_l^m for $N \leq l \leq B$ (the right side of the red arrow) differ. As a result, distinct effects are evident in the reflection kernel reconstructed from the coefficients up to the degree $l = 10$ (Figure 7.1g) and h)).

On the other hand, the reflection kernel reconstructed from the coefficients up to the degree $l = 2$ for $N = 3$ shown in Figure 7.1i) does not show such effects while high frequency components such as specularity are missing at the center. It follows from these examples that it is appropriate to compute R_l^m for $l < N$ only when $4N^2$ samplings of the function are available. In this way, we can avoid annoying effects caused by R_l^m for $N \leq l$ in the reconstructed function $R(\theta, \phi)$.

7.2 Use of Extended Light Sources

The sampling theorem states that the higher the frequency content of an object's appearance, the more input images are required to obtain a correct set of basis images. The number of input images required may become extremely large in the case of highly specular surfaces containing a large quantity of high frequency components in their reflection.

In this case, insufficient sampling of an object appearance will result in aliasing in the basis images, and this will lead to undesirable artifacts in the synthesized appearance. Since the number of input images provided for modeling an object's appearance is usually limited, an anti-aliasing framework for obtaining a set of correct basis images from an insufficient number of object input images is needed. However, this aliasing problem has not been carefully considered in previous studies.

In this section, we propose a novel approach for sampling the appearance of an object under appropriately provided Extended Light Sources *(ELS)*. The use of ELS for modeling the shape and reflectance of an object was originally introduced in [46]. We extend their analysis further in the angular frequency domain so that the harmonic images of an object of arbitrary surface materials can be obtained without suffering from aliasing caused by

insufficient sampling of its appearance.

Extended light sources (*ELS*) may be constructed by illuminating a spherical diffuser with a point light source as illustrated in Figure 7.2. The use of ELS brings the following three advantages:

Functioning as a low-pass filter: ELS have a radiance distribution that is similar to that of the Gaussian function. This enables extended sources to function as a low-pass filter when the reflection kernel of an object is sampled under ELS. From this, we are able to model its appearance without suffering from aliasing.

Adjustable bandwidth: The distribution of ELS can be adjusted by changing the distance from the diffuser to a point light source. This enables us to adjust the bandwidth of the ELS in the frequency domain.

Reducing saturation problem: When an object is illuminated by a point light source, image intensities from its specular reflection components often become much greater than those from its diffuse reflection components. Under ELS, the gap between image intensities due to specular and diffuse reflection components are narrowed [46]. This helps avoid saturation, so that we are able to observe both specular and diffuse reflection components in the same image.

In the following, we consider the reflection kernel of an object sampled under ELS.

7.2.1 Radiance Distribution

First, the radiance distribution of ELS needs to be provided. The radiance distribution of ELS is symmetric with respect to the point source direction and therefore has no azimuth dependence around the point source direction. Accordingly, the radiance of the inner surface of the diffuser E may be represented as a function of the elevation angle θ'_e, that is $E(\theta'_e, \phi'_e) = E(\theta'_e)$. Here θ'_e represents the elevation angle defined with respect to the direction of the point light source as shown in Figure 7.2.

Then $E(\theta'_e)$ are computed from the analytic formula derived in [46] as

$$E(\theta'_e) = \frac{CP[(S+H)\cos\theta'_e - S]}{[(S+H-S\cos\theta'_e)^2 + (S\sin\theta'_e)^2]^{\frac{3}{2}}}, \tag{7.1}$$

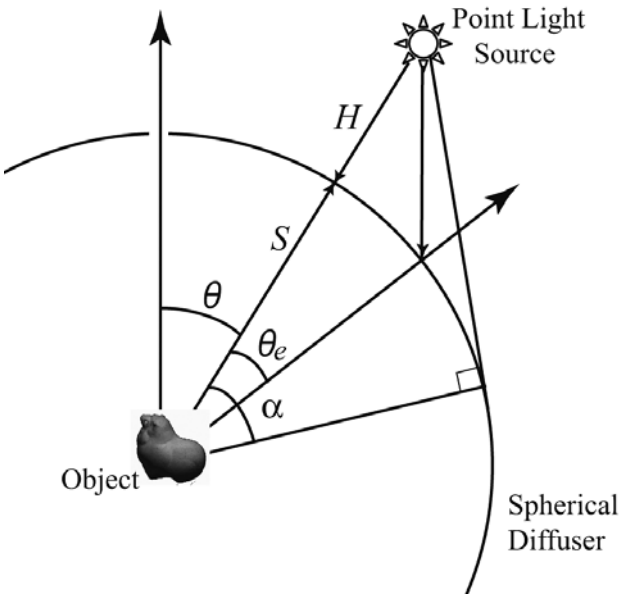


Figure 7.2: ELS are constructed by a spherical diffuser and a point light source in the same manner as described in [46].

where P denotes the radiance of the point source and C is a constant representing the proportionality between the irradiance of the outer surface of the diffuser to the radiance of its inner surfaces. As illustrated in Figure 7.2, S is the radius of the spherical diffuser, H is the distance from the the diffuser's surface to the point light source.[2]

Here the surface points on the diffuser within the range of $\theta'_e < \alpha$ can receive energy from the point light source, and the effective range α is determined by the relationships between S and H as

$$\alpha = \cos^{-1}(\frac{S}{S+H}). \tag{7.2}$$

7.2.2 Modified Reflection Kernel

Next, let us consider a reflection kernel $R(\theta, \phi)$ sampled as an observed brightness of an object surface when it is illuminated by ELS whose center is located at the direction (θ, ϕ).

Since radiance of the ELS are distributed over the spherical surface of the diffuser, the modified reflection kernel seen under ELS, denoted as $R'(\theta, \phi)$ is determined by computing

[2] This formula is based on the condition that the spherical diffuser is ideal and thus incident energy is scattered equally in all directions.

the integral of the scene radiance resulting from illumination of the entire surface of the extended source as

$$R'(\theta,\phi) = \int_0^{\pi}\int_0^{2\pi} R(M_\theta^\phi(\theta'_e,\phi'_e))E(\theta'_e)\sin\theta'_e \mathrm{d}\theta'_\mathrm{e}\mathrm{d}\phi'_\mathrm{e}, \tag{7.3}$$

where $M_\theta^\phi(\cdot)$ is a rotation operator that rotates the distribution of the ELS $E(\theta'_e)$ so that its center ($\theta'_e = 0$) is located at the direction (θ,ϕ). In other words, the modified reflection kernel $R'(\theta,\phi)$ is determined by convolving the original reflection kernel $R(\theta,\phi)$ with the radiance distribution of the ELS $E(\theta'_e)$ as illustrated in Figure 7.3.

$R(M_\theta^\phi(\theta'_e,\phi'_e))$ and $E(\theta'_e)$ in (7.3) can be expanded as a linear combination of spherical harmonics:

$$E(\theta'_e) = \sum_{n=0}^{\infty} E_n Y_n^0(\theta'_e) \tag{7.4}$$

$$R(M_\theta^\phi(\theta'_e,\phi'_e)) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} R_l^m Y_l^m(M_\theta^\phi(\theta'_e,\phi'_e)), \tag{7.5}$$

where R_l^m and E_n denote coefficients in their spherical harmonic expansion. Then a rotation formula for spherical harmonics is given in [52] as

$$Y_l^m(M_\theta^\phi(\theta'_e,\phi'_e)) = \sum_{m'=-l}^{l} D_{m,m'}^l(\theta)e^{Im\phi}Y_l^{m'}(\theta'_e,\phi'_e), \tag{7.6}$$

where the term $e^{Im\phi}$ considers the rotation about ϕ, and the matrix D^l tells us how to compute a rotated spherical harmonic as a linear combination of all the spherical harmonics of the same order l.

Substituting (7.4) and (7.5), (7.3) becomes

$$R'(\theta,\phi) = \sum_{n=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\sum_{m'=-l}^{l} R_l^m E_n D_{m,m'}^l(\theta)e^{Im\phi} \int_0^{\pi}\int_0^{2\pi} Y_l^{m'}(\theta'_e,\phi'_e)Y_n^0(\theta'_e)\sin\theta'_e \mathrm{d}\theta'_\mathrm{e}\mathrm{d}\phi'_\mathrm{e}. \tag{7.7}$$

Then the orthonormality of the spherical harmonics tells us

$$\int_0^{\pi}\int_0^{2\pi} Y_l^{m'}(\theta'_e,\phi'_e)Y_n^0(\theta'_e)\sin\theta'_e \mathrm{d}\theta'_\mathrm{e}\mathrm{d}\phi'_\mathrm{e} = \delta_{\mathrm{ln}}\delta_{\mathrm{m'0}},$$

and therefore

$$R'(\theta,\phi) = \sum_{n=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\sum_{m'=-l}^{l} R_l^m E_n D_{m,m'}^l(\theta)e^{Im\phi}\delta_{ln}\delta_{m'0}. \tag{7.8}$$

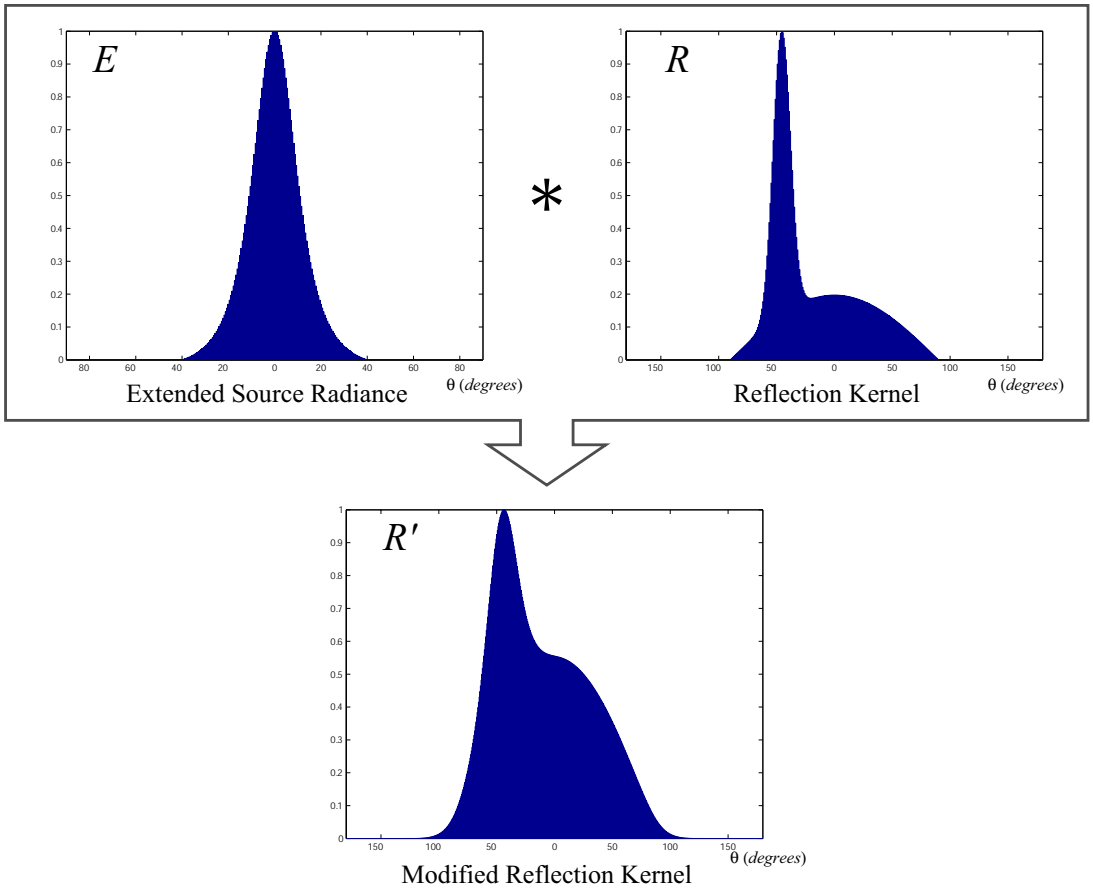


Figure 7.3: Modified Reflection Kernel obtained by convolving the original reflection kernel $R(\theta, \phi)$ with the distribution of the extended source $E(\theta'_e)$.

Finally, from the characteristics of Kronecker delta ($\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$ if $i \neq j$), the modified reflection kernel is obtained

$$R'(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} R_l^m E_l D_{m,0}^l(\theta) e^{Im\phi}. \tag{7.9}$$

The important point to note in (7.9) is that the modified reflection kernel $R'(\theta, \phi)$ is computed as a product of the coefficient R_l^m and E_l. This indicates that the modified reflection kernel $R'(\theta, \phi)$ is band-limited with the lower bandwidth between $R(\theta, \phi)$ and $E(\theta'_e)$. Fortunately, it is not difficult to adjust the bandwidth of the ELS.

7.2.3 Adjusting Bandwidth of Extended Light Sources

As has been noted above, ELS used in our approach have a distribution similar to that of the Gaussian function, and their distribution and effective range are determined by the radius S of a spherical diffuser and the distance H from the diffuser surface to a point light source as shown in (7.1) and (7.2). In other words, the closer the point light source is positioned to the diffuser, the narrower its range α becomes. This property of Extended Light Sources enables us to specify their bandwidth.

The Fourier transform of the Gaussian function is known to lead to another Gaussian function in the frequency domain. In addition, the standard deviation σ_f of the Gaussian function in the frequency domain is known to become inversely proportional to the standard deviation σ_s of the Gaussian function in the spatial domain. That is $\sigma_f = \frac{1}{\sigma_s}$. Similarly, we can safely say that the Gaussian function defined on the unit sphere also results in a half-Gaussian distribution of the coefficients, that is a Gaussian distribution with a half range, in its spherical harmonic expansion.

To see how bandwidth of ELS changes depending on their α values, we synthetically provide several radiance distributions of ELS with different α values from (7.1). The coefficients E_l and the corresponding α of those extended sources are shown in Figure 7.4. In this figure, spherical harmonic coefficients for given degree l and order m are represented using a single index $r = l^2 + l + m$. The horizontal axis represents the index r, and the vertical axis represents the computed coefficients E_l.

In each graph in this figure, the left side of a red arrow corresponds to the coefficients up to the degree $l = 8$. The energy captured by spherical harmonics up to the degree $l = 8$, denoted as P_8 are also computed by sums of the squares of their respective coefficients divided by the total squared energy of the transformed function. In this figure, we clearly see that P_8 becomes closer to 100% as their α values increase from 30 to 50 degrees. Especially in the case of $\alpha = 50$ degrees, more than 99% of the total energy is captured by the spherical harmonics up to the degree $l = 8$, so it is reasonable to assume that the extended source generated with $\alpha = 50$ degrees is band limited with bandwidth $l = 8$. As shown in this example, the bandwidth of ELS can be set by adjusting their effective range α.

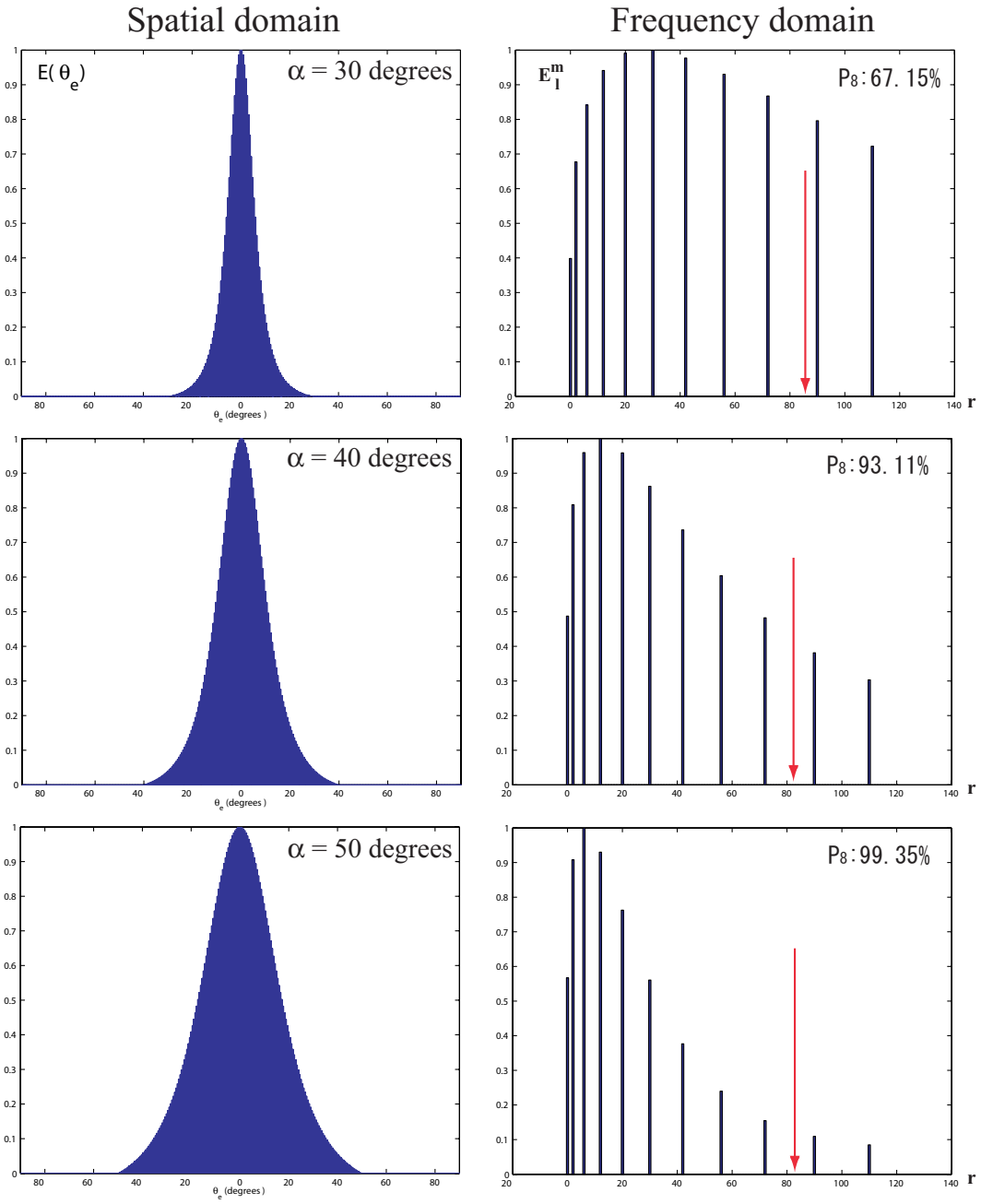


Figure 7.4: Distribution of Extended Light Sources (left: spatial domain, right: frequency domain).

7.3 Modeling Appearance by using Extended Light Sources

Based on the analysis provided in the previous section, we propose a novel scheme for sampling the reflection kernel of an object by ELS with a properly adjusted range α. Here we employ the efficient sampling theorem in spherical harmonics proposed in [40].

This sampling theorem tells us that spherical harmonics transformation of a band-limited function with bandwidth B ($R_l^m = 0$ $(l \geq B)$), can be computed by weighted sums of $2B^2 - B$ sampled function values based on Gaussian quadrature [51] as

$$F_l^m = \frac{2\pi}{2B-1} \sum_{j=0}^{B-1} \sum_{k=0}^{2B-2} w_j f(\theta_j, \phi_k) Y_l^m(\theta_j, \phi_k), \tag{7.10}$$

where the weights w_j are weights of Gaussian quadrature, θ_j are Gaussian nodes in $\cos\theta$, and $\phi_k = \frac{2\pi_k}{(2B-1)}$ are equally sampled in azimuth [40, 48].

In order to avoid aliasing caused by insufficient sampling of the appearance of an object, we observe its reflection kernel under ELS whose bandwidth is properly adjusted to B. By doing so, the modified reflection kernel $R'(\theta, \phi)$ of the object sampled under the ELS also becomes band-limited with the same bandwidth B as the ELS, and this enables us to model the appearance of the object without suffering from aliasing caused due to insufficient sampling of the original reflection kernel $R(\theta, \phi)$.

7.4 Experimental Results

We have tested our proposed approach based on ELS by using both synthetic and real data.

7.4.1 Synthetic Data

The reflection kernel $R(\theta, \phi)$ of several surface materials are synthetically provided based on the Ward isotropic reflection model [69] with known reflection parameters as shown in Table 7.1. Based on the Ward reflection model, reflection kernel $R(\theta, \phi)$ is computed as

$$R(\theta, \phi) = \frac{K_d}{\pi} + K_s \frac{1}{\sqrt{\cos\theta \cos\theta_o}} \frac{\exp(-\tan^2\gamma(\theta, \phi)/\sigma^2)}{4\pi\sigma^2}, \tag{7.11}$$

where K_d and K_s are constants for the diffuse and specular reflection components, and σ is the standard deviation of the surface slope. Here the surface normal and viewing direction

for the reflection kernel are set at the direction $(\theta_n = 0, \phi_n = 0)$ and $(\theta_o = 45, \phi_o = 0)$ respectively.

The ELS used in this experiment are generated with $\alpha = 50$ degrees whose bandwidth is equivalent to $B = 9$. This makes the modified reflection kernel $R'(\theta, \phi)$ band-limited with $B = 9$. Accordingly, the spherical harmonic coefficients $R_l'^m$ of the band-limited function $R'(\theta, \phi)$ were obtained from a properly sampled discrete set of the function values (153 samplings for $B = 9$ in (7.10)).

The right columns in Figure 7.5, and 7.6 show the obtained reflection kernel for glossy gray paper and lightly brushed aluminum surfaces respectively. b) shows computed coefficients $R_l'^m$, and d) and f) show the distribution of $R'(\theta, \phi)$ reconstructed from the obtained $R_l'^m$ in (6.2) up to the degree $l = 8$.[3]

For reference, the left columns of these figures show the coefficients and distribution of the original reflection kernel $R(\theta, \phi)$. a) shows its coefficients computed from 100000 samplings of $R(\theta, \phi)$, and c) and e) show its distribution computed by substituting the reflection parameters into the Ward reflection model.

Comparing a) with b), while the magnitude of the coefficients is different because of the multiplication with the coefficients of the ELS whose distribution is similar to that of a half Gaussian function, the recovered $R_l'^m$ show almost the same distribution as that of R_l^m. In Figure 7.7 a) and b), we also provide the low-frequency appearance of the original reflection kernel for reference by substituting the coefficients R_l^m up to degree $l = 8$ into (6.2). In other words, all coefficients with the index $l > 8$ of R_l^m are truncated in this case.

Comparing f) with the low-frequency appearance of the original reflection kernel in Figure 7.7 a), the modified reflection kernel reconstructed by our method provides a good

Table 7.1: Reflection parameters for the Ward model.

Material	k_d	k_s	σ
glossy gray paper	.29	.083	.082
lightly brushed aluminum	.15	.19	.088

[3] d) shows $R'(\theta, \phi)$ scanned in the line of $\{\theta | 0 \leq \theta \leq \pi/2\}$, $\{\phi | 0 \text{ or } \pi\}$, and f) shows the upper half of the reflection kernel visualized in a polar coordinate system with radius indicating $\{\theta | 0 \leq \theta \leq \pi/2\}$, and angle indicating $\{\phi | 0 \leq \phi < 2\pi\}$.

representation of the low-frequency appearance of the original reflection kernel, although high-frequency appearance such as the specular peak in e) is blurred due to the approximation of the original reflection kernel up to the degree $l = 8$.

It is worth noting that almost no undesirable artifacts appear in the reconstructed reflection kernel $R'(\theta, \phi)$ in f). This shows that the proposed method succeeded in computing the coefficients $R_l'^m$ of the modified reflection kernel from the discrete 153 samplings of $R'(\theta, \phi)$ without suffering from aliasing due to insufficient sampling of its appearance. In contrast, as Westin et al. pointed out in [70], simply truncating all coefficients R_l^m with $l > 8$ resulted in undesirable artifacts such as ringing in the reconstructed reflection. This ringing is called the Gibbs phenomenon and is shown in Figures 7.7 a) and b).

In order to avoid the Gibbs phenomenon, Westin et al. progressively reduced the magnitude of the coefficients according to a half Gaussian distribution of an empirically determined width. Fortunately, the coefficients of the ELS used in our method have a distribution that is similar to that of a half Gaussian distribution, as shown in Figure 7.4. It follows from this that the use of ELS is desirable not only as a method of modifying the original reflection kernel to be band-limited, but also for reducing the number of artifacts caused by the truncation of the coefficients with index $l \geq B$ of the original reflection kernel. [4]

[4] $R'(\theta, \phi)$ is computed as a product of R_l^m, E_l, and D^l in (7.9). Since E_l and D^l are computable numbers, we are able to recover R_l^m of the original reflection kernel from the set of modified reflection kernel $R'(\theta, \phi)$ if necessary.

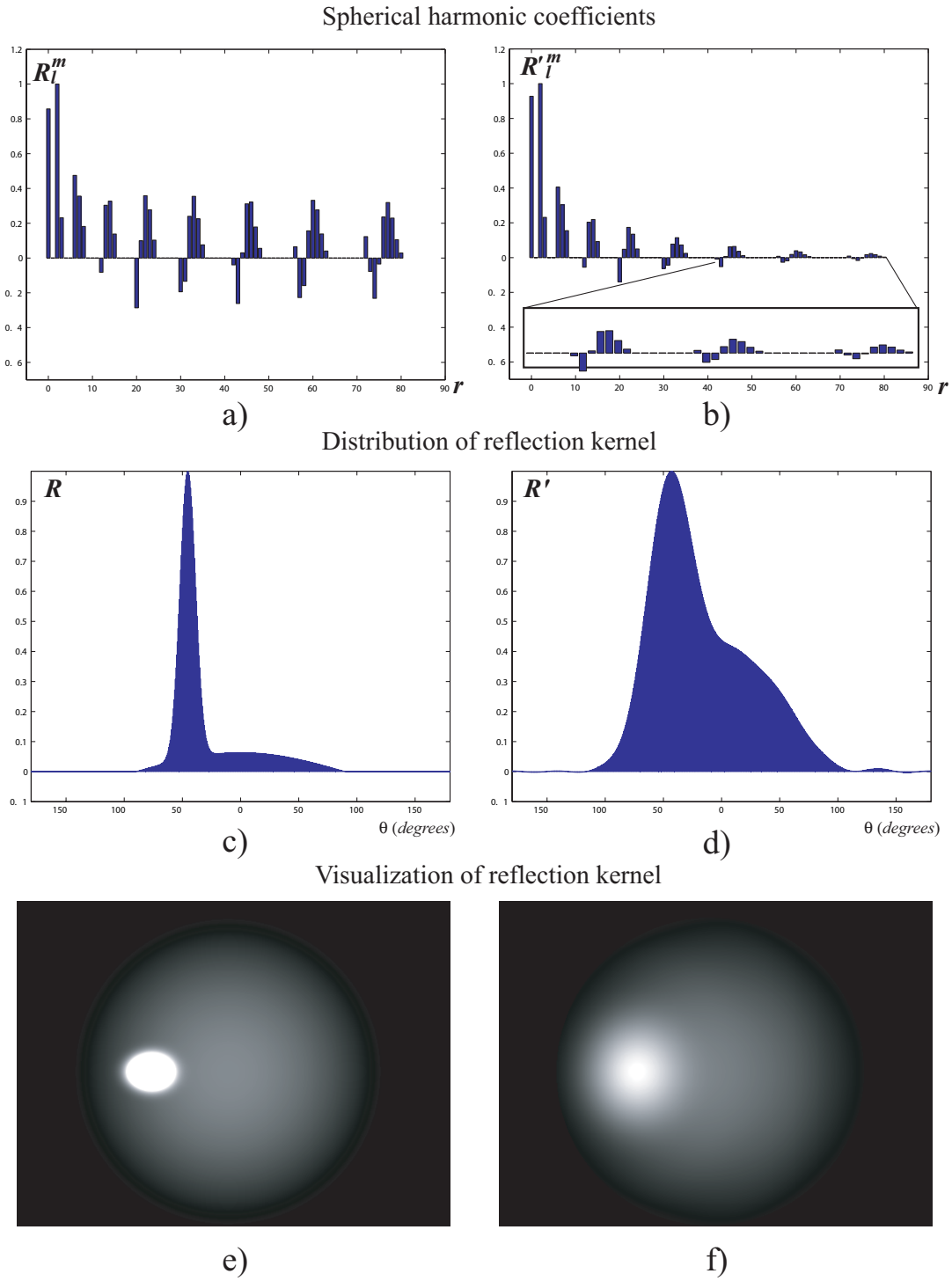


Figure 7.5: Appearance sampling of glossy gray paper : original reflection kernel (left), recovered modified reflection kernel (right)

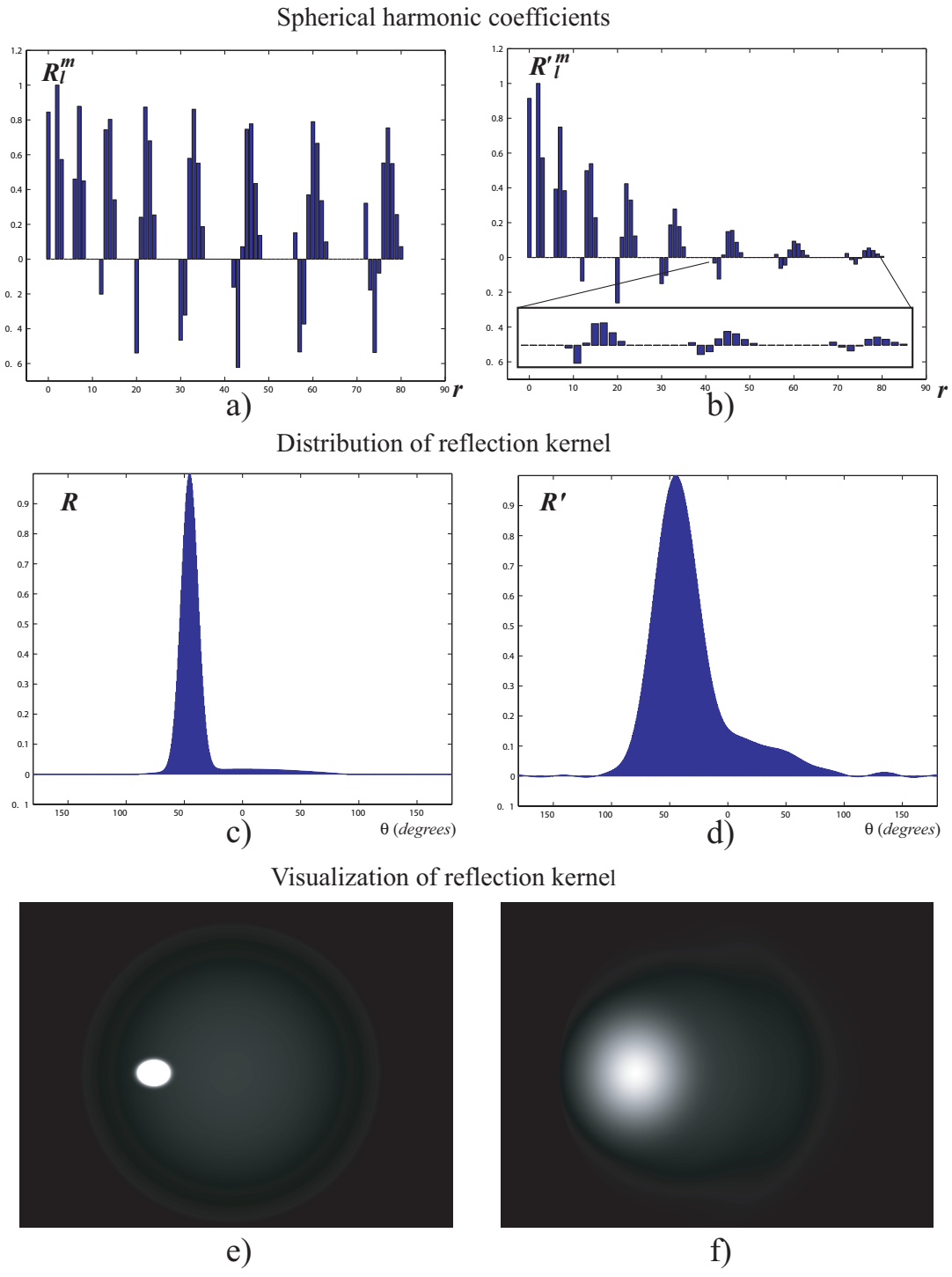


Figure 7.6: Appearance sampling of lightly brushed aluminum : original reflection kernel (left), recovered modified reflection kernel (right)

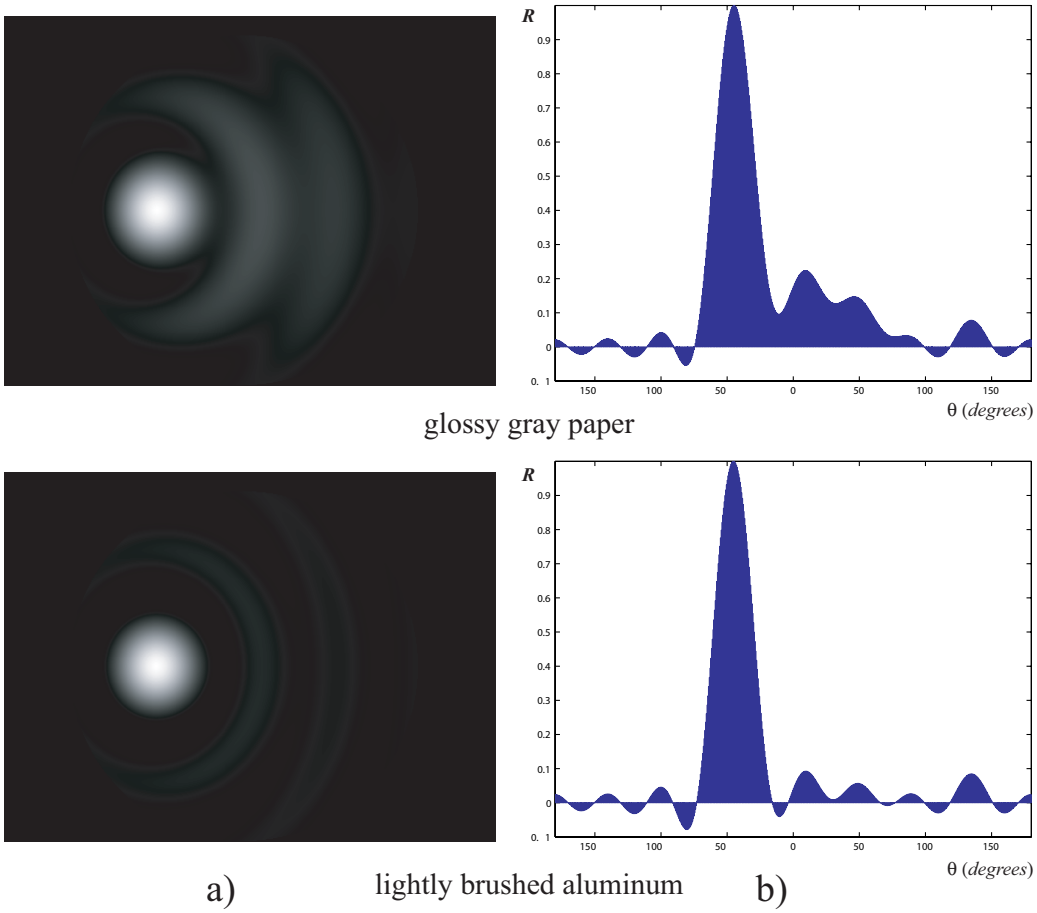


Figure 7.7: Reconstructed reflection kernel from the coefficients R_l^m of the original reflection kernel up to the degree $l = 8$

7.4.2 Real Data

Real images of abalone shellfish were taken under the physically constructed ELS apparatus shown in Figure 7.8. Abalone shellfish is famous for its interesting structural colors which alter greatly depending on viewing direction and illumination conditions. In this set-up, 9 halogen lamps were positioned at Gaussian nodes in elevation, and an array of these light sources were mounted on a turntable and rotated around the spherical diffuser (an acrylic globe with a diameter of 35cm) by $\phi_k = \frac{2\pi_k}{(18-1)}, (k = 0, \ldots, 16)$ degrees in azimuth. Here the number of point light sources in elevation indicates that a modified reflection kernel has to be band-limited with $B = 9$. Accordingly the distance between the diffuser and the point light sources was adjusted so that the bandwidth of the constructed ELS was set to $B = 9$.

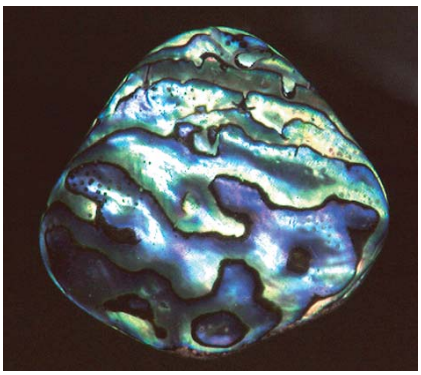
In total, 153 input images of the abalone shellfish were taken to sample its reflection kernel at each grid point $(\theta_j,\ \phi_k)$. Then coefficients $R_l^{\prime m}$ of this reflection kernel were computed up to the degree $l = 8$ by substituting the observed reflection kernel $R'(\theta_j, \phi_k)$ into (7.10). The first nine harmonic images obtained by our method are shown in Figure 7.9 (a).

Figure 7.9 (b) shows the appearance of the abalone shellfish synthesized from (6.7) under natural illumination conditions provided as high-dynamic range light probe measurements by [7]. In this figure, the synthesized appearance of the abalone shellfish significantly changes depending on the characteristics of the given illumination distributions, and this shows that the complex appearance of its structural colors are well represented by a set of basis images obtained by our method.

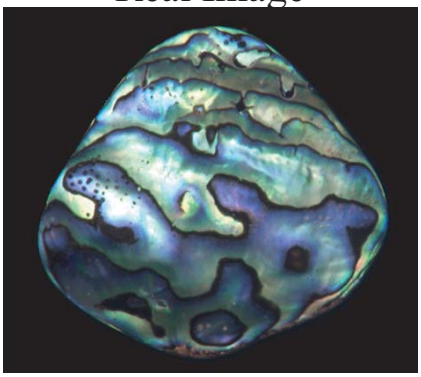
In addition, the synthesized appearance of the shellfish is compared with its real appearance as seen under a normal lighting condition in our laboratory in Figure 7.8.[5] In this figure, the synthesized appearance highly resembles the real image.

As has been noted before, in the case of a highly specular surface, image intensities from specular reflection components tend to be much greater than those from diffuse reflection components. This is the reason why image intensities in the real image become saturated in some specular regions. In general, the limited dynamic range of an image taken with one shutter speed makes it difficult to model the appearance of highly specular

[5] Coefficients L_l^m of this lighting condition are computed from an omni-directional image of the scene taken by a camera with a fish-eye lens.



Real Image



Synthesized Image

Figure 7.8: Comparison between real image and synthesized image under complex illumination. The appearance of the shellfish is modeled by using extended light sources.

surfaces. In contrast, the use of ELS contributes to reduce the high contrast between image intensities from both specular and diffuse components. This helps us to observe and model the appearance of highly specular surfaces from images taken at a single shutter speed.

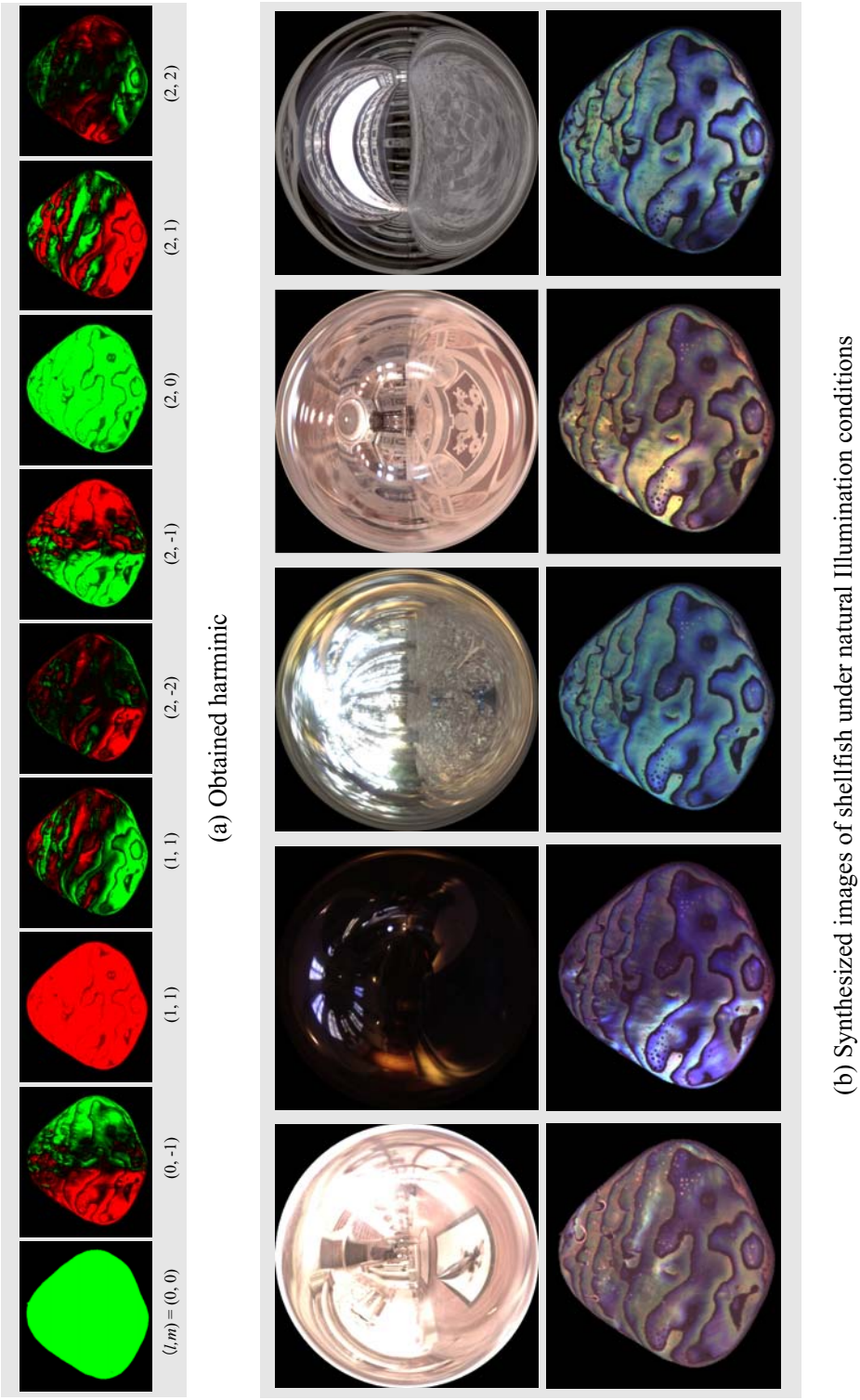


Figure 7.9: (a) Obtained harmonic images: positive values are shown in green, and negative values are shown in red. (b) Synthesized images of objects under natural illumination. The first row shows illumination maps, and the second row shows synthesized appearance under the corresponding illumination map.

7.5 Conclusions

In this chapter we have demonstrated the effectiveness of using Extended Light Sources (ELS) for modeling the appearance of an object under varying illumination. The use of ELS for modeling the appearance of objects has the following advantages: (1) ELS have the ability to function as a low-pass filter for sampling objects' appearance, so basis images of an object can be obtained without suffering from aliasing caused by insufficient sampling of its reflection kernel, (2) ELS can reduce the high contrast between image intensities from both specular and diffuse components. This helps us observe and model the appearance from images taken at a single shutter speed, and (3) ELS can minimize undesirable artifacts resulting from the truncation of spherical harmonic coefficients of higher degrees. These advantages help provide basis images of an object from a limited number of samplings of its appearance.

Our method can adequately synthesize the appearance of an object up to a certain frequency. However, objects of high bandwidth such as a mirror illuminated by a point light source are difficult to be adequately modeled. A future research direction of this work would be to integrate other modeling techniques with the objective of recovery of an object original reflection kernel.

Chapter 8

Summary and Conclusions

Conventional model-based rendering techniques have been intensively developed for synthesizing the realistic appearance of computer graphics objects. Model-based rendering techniques synthesize the appearance of objects based on empirically or analytically given reflection models. To use these models, the geometric and photometric information about the scene needs to be provided: the shapes of objects in the scene, their surface reflectance properties, and the lighting condition of the scene where those objects are placed.

Surface reflectance properties greatly influence the appearance of an object. The appearance of a metallic surface is completely different from that of a matted surface even under the same lighting condition. In addition, the appearance of an object changes significantly under different lighting conditions. It is thus important to provide not only appropriate surface reflectance properties of objects in a scene but also appropriate illumination conditions so that the realistic appearance of the objects can be synthesized under these illumination conditions. Nevertheless, the photometric information about a scene tends to be manually provided by a user.

Since it is difficult to imagine the appearance of an object directly from reflectance parameters, this input process of manually specifying its reflectance properties is normally non-intuitive and thus time-consuming. The correct appearance of a scene is difficult to achieve unless we stop relying on our instinct for adjusting reflectance parameters. As for providing lighting conditions, a scene generally includes both direct and indirect illumination distributed in a complex way, and it is difficult for a user to manually specify such complex illumination distribution.

In order to overcome these difficulties in providing photometric information about a

scene, techniques for automatically providing the photometric models of a scene have been studied in the fields of both computer vision and computer graphics research. In particular, techniques that use a set of images of a scene provided under different viewing and/or lighting conditions for determining its geometric and photometric information are called image-based modeling.

This thesis addressed two issues of image-based modeling for synthesizing the photorealistic appearance of an object under natural illumination conditions: capturing real-world illumination, and modeling complex appearance of real objects for variable illumination. Regarding the first issue of capturing and modeling real-world illumination, both of an image-based approach and an inverse lighting approaches were studied.

8.1 Modeling Real-World Illumination

The technique that measures real-world illumination conditions from photographically acquired images of the scene is called image-based lighting. While image-based lighting techniques have been developed successfully with practical applications, two difficulties still remained to be solved: how to construct a geometric model of the scene, and how to capture a wide field of view of the scene.

In Chapter 2, we confronted these two difficulties and proposed an efficient method for automatically measuring illumination distribution of a real scene by using a pair of omni-directional images taken by a CCD camera with a fisheye lens based on an omni-directional stereo algorithm. In Chapter 3, we pursued the possibility of real-time rendering of synthetic objects with natural shading and cast shadows onto a real scene whose illumination condition was dynamically changing.

There has been another approach called inverse lighting in image-based modeling that deals with an inverse problem of traditional model-based rendering. One of the main advantages of inverse lighting over the former image-based lighting is that it does not require additional images for capturing illumination of a scene, but uses the appearance of objects located in a scene instead for recovering an illumination distribution of the scene. Real scenes normally include both direct and indirect illumination distributed in a complicated way, and this makes it difficult to analyze characteristics of the illumination conditions of the scene from image brightness in inverse lighting.

In Chapter 4, we demonstrated the effectiveness of using occluding information of incoming light in estimating an illumination distribution of a scene. Shadows in a scene are caused by the occlusion of incoming light, and thus contain various pieces of information about the illumination of the scene. In our method, image brightness inside shadows was effectively used for providing distinct clues to estimate an illumination distribution. Chapter 4 further addressed two issues in inverse lighting. First, the method combined the illumination analysis with an estimation of the reflectance properties of a shadow surface. This makes the method applicable to cases where reflectance properties of a surface were not known a priori. Second, an adaptive sampling framework for efficient estimation of illumination distribution was introduced.

Later in Chapter 5, the amount of the information about the illumination distribution of scene obtainable from a given image of a scene was discussed. In particular, two main factors that controlled the stability of the illumination estimation from shadows were analyzed: blocked view of shadows and limited sampling resolution for radiance distribution inside shadows. Based on this analysis, a robust method was presented.

8.2 Modeling Appearances of Objects for Variable Illumination

Inverse rendering carries out the opposite procedures of model-based rendering to provide object and illumination models of a real scene from photographically available information of the scene. Once models of a scene are acquired, new image of the scene under novel lighting and/or viewing conditions can be synthesized by using conventional model-based rendering techniques.

On the other hand, the approach called image-based rendering directly uses the original set of input images of a scene for producing new images of the scene under novel conditions. Depending on which scene conditions should be modified, image-based rendering techniques are classified into three categories: image-based render under novel viewing conditions, image-based rendering under novel lighting conditions, and image-based rendering under novel viewing and novel lighting conditions. In this thesis, the second category, image-based rendering under novel lighting conditions was considered.

In contrast with model-based rendering techniques, image-based rendering techniques

do not require full radiometric computation to synthesize the photorealistic appearance of a scene. This makes the cost to produce new images of the scene independent of the scene complexity. Image-based rendering, however, has a tendency to require many input images of a scene to synthesize reasonably realistic appearance of objects in a scene. This results in the requirement for a large amount of both computer memory and data storage.

While there may seem to be a large variety of possible appearances for a given object, it has been demonstrated in previous research that changes in appearance of an object for varying illumination can be represented with a linear subspace spanned by a set of basis images of the object. A set of basis images spanning such a linear subspace is often provided by applying principal-component analysis to the input images of an object taken under different lighting conditions. Since little is known about how to sample the appearance of an object in order to obtain its basis images correctly, a large number of input images taken by moving a point light source along a sphere surrounding the object are generally provided.

Recent investigations in frequency-space analysis of reflection have shown that the appearance of an object under varying complex illumination conditions can be well represented with a linear subspace spanned by basis images of the object, called harmonics images, each of which corresponds to an images of the object illumination u under harmonics lights whose distribution are specified in terms of spherical harmonics.

In Chapter 6, we presented a method for analytically obtaining a set of harmonics images for varying illumination from input images of the object taken under a set of point light sources. The main contribution of our work is that we show that a set of lighting directions can be determined for sampling images of an object depending on the spectrum of the object's BRDF in the angular frequency domain such that a set of harmonic images can be obtained analytically based on the sampling theorem on spherical harmonics.

Using those sampling directions determined from the sampling theory, we are able to obtain harmonic images by using a significantly smaller number of input images than other techniques that do not take into account a relationship between a spectrum of BRDFs and a sampling density of illumination directions. In addition, unlike other methods based on spherical harmonics, our method does not require the shape and reflectance model of an object used for rendering harmonics images of the object synthetically. Thus, our method can be easily applied to determine a set of basis images for representing the appearance

change of a real object under varying illumination conditions.

The sampling theorem, however, states that the higher the frequency content of an object's appearance, the more input images are required to obtain a correct set of basis images. The number of input images required may become extremely large in the case of highly specular surfaces containing a large quantity of high frequency components in their reflection.

In this case, insufficient sampling of an object appearance will result in aliasing in the basis images, and this will lead to undesirable artifacts in the synthesized appearance. Since the number of input images provided for modeling an object's appearance is usually limited, an anti-aliasing framework for obtaining a set of correct basis images from an insufficient number of object input images is needed. Nevertheless, this aliasing problem has not been investigated in previous studies.

In Chapter 7 we carefully studied this issue of aliasing and extend the method based on the sampling theorem further for reducing the artifacts due to aliasing, by substituting extended light sources (*ELS*) for a point light source to sample the reflection kernel of a real object. The use of ELS for modeling the shape and reflectance of an object was originally introduced in [46]. We extended their analysis further in the angular frequency domain so that the harmonic images of an object of arbitrary surface materials can be obtained without suffering from aliasing caused by insufficient sampling of its appearance.

The use of ELS has the following advantages. ELS have a radiance distribution that is similar to that of the Gaussian function, and this enables extended sources to function as a low-pass filter when the appearance of an object is sampled under them. This enables us to obtain a set of basis images of an object for varying illumination without suffering from aliasing caused by insufficient sampling of its appearance. In addition, ELS can reduce high contrast in image intensities due to specular and diffuse reflection components. This helps avoid saturation so that we are able to observe both specular and diffuse reflection components in the same image taken with a single shutter speed.

Once a set of basis images of an object is obtained, its appearance under natural illumination conditions can be synthesized simply as a linear combination of these basis images whose linear coefficients are computed from the given lighting conditions, and these lighting conditions can be modeled by our proposed image-based or inverse lighting methods.

Our method can adequately synthesize the appearance of an object up to a certain frequency. However, objects of high bandwidth such as a mirror illuminated by a point light source are difficult to be adequately modeled. A future research direction of this work would be to integrate other modeling techniques with the objective of recovery of an object original reflection kernel and to model appearances of objects seen from arbitrary viewing directions.

Bibliography

[1] R. Baribeau, M. Rioux, and G. Godin, "Color Reflectance Modeling Using a Polychromatic Laser Range Sensor," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 263-269, 1992.

[2] R. Basri and D. Jacobs, "Lambertian Reflectance and Linear Subspaces," *Proc. IEEE Int. Conf. Computer Vision '01*, pp.383-389, Dec. 2001.

[3] D. H. Ballard & C. M. Brown, *Computer Vision*, Prentice-Hall, Englewood Cliffs, NJ., 1982.

[4] J. Bouguet and P. Perona, "3D Photography on Your Desk," *Proc. IEEE Int. Conf. Computer Vision '98*, pp.43-50, 1998.

[5] M.F. Cohen, S.E. Chen, J.R. Wallace, D.P. Greenberg, "A Progressive Refinement Approach to Fast Radiosity Image Generation," *Proc. SIGGRAPH '88*. pp. 75-84, 1988.

[6] M. F. Cohen, D. P. Greenberg, D. S. Immel, and P. J. Brock, "An Efficient Radiosity Approach for Realistic Image Synthesis," *J. IEEE Computer Graphics and Applications*, Vol. 6, No.3, , pp.26-35, Mar. 1986.

[7] P. E. Debevec, "Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-based Graphics with Global Illumination and High Dynamic Range Photography," *Proc. SIGGRAPH '98*, pp. 189-198, July, 1998.

[8] P. Debevec, T. Hawkins, C. Tchou, H. Duiker, W. Sarokin, and M. Sagar, "Acquiring the Reflectance Field of a Human Face", *Proc. SIGGRAPH '00*, pp. 145-156, 2000.

[9] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," *Proc. SIGGRAPH '97*, pp. 369-378, August, 1997.

[10] Y. Dobashi, K. Kanade, H. Naktani, and H. Ymashita, "A quick rendering method using basis function for interactive lighting design," *Proc. Eurographics '95*, pp.229-240, 1995.

[11] J. Dorsey, J.Arvo, and D. Greenberg, "Interactive Design of Complex Time-Dependent Lighting", *J. IEEE Computer Graphics and Applications*, vol. 15, no. 2, pp.26-36, 1995.

[12] G. Drettakis, L. Robert, S. Bougnoux, "Interactive Common Illumination for Computer Augmented Reality" *Proc. 8th Eurographics Workshop on Rendering*, pp. 45-57, 1997.

[13] J. Driscoll and D. Healy, Jr, "Computing Fourier transforms and convolutions on the 2-sphere," *J. Advanced in Applied Mathematics*, Vol.15, pp. 202-250, 1994.

[14] R. Epstein, P. Hallinan, and A. Yuille, "5+/-2 eigenimages suffice: An empirical investigation of low-dimensional lighting models," *Proc. IEEE Workshop on Physics-Based Modeling in Computer Vision*, pp. 108-116, 1995.

[15] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge, MA., 1993.

[16] D. A. Forsyth and J. Ponce, *Computer Vision A Modern Approach*, Prentice Hall, Upper Saddle River, NJ., 2002.

[17] A. Fournier, A. Gunawan and C. Romanzin, "Common Illumination between Real and Computer Generated Scenes," *Proc. Graphics Interface '93*, pp.254-262, 1993.

[18] A. Georghiades, D. Kriegman, and P. Belhumeur, "Illumination cones for recognition under variable lighting: Faces," *Proc. IEEE Conf. Computer Vision and Pattern Recognition '98*, pp. 52-59, 1998.

[19] A. Georghiades, D. Kriegman, and P. Belhumeur, "From few to many: Generative models for recognition under variable pose and illumination," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643-660, 2001.

[20] A. S. Glassner, *Graphics Gems*, AP Professional, Cambridge, MA., 1990.

[21] P. Hallinan, "A low-dimensional representation of human faces for arbitrary lighting conditions," *Proc. IEEE Conf. Computer Vision and Pattern Recognition '94*, pp. 995-999, 1994.

[22] G. E. Healey, S. A. Shafer, and L. B. Wolff, *Physics-Based Vision Principles and Practice, Color*, Jones and Bartlett Publishers, Boston, MA., 1992.

[23] B. K. P. Horn, *Robot Vision*, The MIT Press, Cambridge, MA., 1986.

[24] B. K. P. Horn,"Understanding Image Intensities," *Artificial Intelligence*, Vol. 8, No.2, pp.201-231, 1997.

[25] B. K. P. Horn, "Obtaining Shape from Shading Information," *The psychology of Computer Vision*, McGraw-Hill Book Co., New York, N.Y., 1975.

[26] B. K. P. Horn and M. J. Brooks, "The Variational Approach to Shape from Shading," *Computer Vision, Graphics, and Image Processing*, Vol.33. No.2, pp.174-208, 1986.

[27] K. Ikeuchi and B. K. P. Horn, "Numerical Shape from Shading and Occluding Boundaries," *Artificial Intelligence*, Vol. 17, No.1-3, pp.141-184, 1981.

[28] K. Ikeuchi and K. Sato, "Determining Reflectance using Range and Brightness Images," *Proc. IEEE Intl. Conf. Computer Vision '90*, pp.12-20, 1990.

[29] K. Jacobs and C. Loscos, " Classification of Illumination Methods for Mixed Reality, " *Eurographics State of the Art Reports*, pp. 95 -118, 2004.

[30] T. Kanade, A. Yoshida, K. Oda, H. Kano, and M. Tanaka, "A Video-Rate Stereo Machine and Its New Applications," *Proc. IEEE Conf. Computer Vision and Pattern Recognition '96*, pp.196-202, 1996.

[31] J. Kawai, J. Painter, and M. Cohen, "Radioptimization - Boal Bawsed Rendering"*Proc. ACM SIGGRAPH '93*, pp. 147-154, 1993.

[32] G. Kay and T. Caelli, "Estimating the Parameters of an Illumination Model using Photometric Stereo,"*Graphical Models and Image Processing*, vol. 57, no. 5, pp. 365-388, 1995.

[33] J. R. Kender and E. M. Smith, "Shape from Darkness: Deriving Surface Information from Dynamic Shadows," *Proc. IEEE Intl. Conf. Computer Vision '87*, pp.539-546, 1987.

[34] T. Kim, Y. Seo, and K. Hong, "Improving AR using Shadows Arising from Natural Illumination Distribution in Video Sequence," *Proc. IEEE Intl. Conf. Computer Vision '01*, pp. 329-334, 2001.

[35] M. Koudelka, P. Belhumeur, S. Magda, and D. Kriegman, " Image-based Modeling and Rendering of Surfaces with Arbitrary BRDFs, " *Proc. IEEE Conf. Computer Vision and Pattern '01*, vol. 1, pp. 568-575, 2001.

[36] K. C. Lee, J. Ho, and D. Kriegman, "Nine points of light: Acquiring subspaces for face recognition under variable lighting," *Proc. IEEE Conf. Computer Vision and Pattern Recognition '01*, pp.519-526, Dec. 2001.

[37] J. Lu and J. Little, "Reflectance Function Estimation and Shape Recovery from Image Sequence of a Rotating Object," *Proc. IEEE Intl. Conf. Computer Vision '95*, pp. 80-86, 1995.

[38] A. K. Mackworth, *On the Interpretation of Drawings as Three-Dimensional Scenes*, doctorial dissertation, University of Sussex, 1974.

[39] S. R. Marschner and D. P. Greenberg, "Inverse Lighting for Photography," *Proc. IS&T/SID Fifth Color Imaging Conference*, pp.262-265, 1997.

[40] M. J. Mohlenkamp, "A Fast Transform for Spherical Harmonics," *J. Fourier Analysis and Applications*, vol. 5(2/3), pp.159-184, 1999.

[41] P. Moon and D. E. Spencer, *The Photic Field*, The MIT Press, Cambridge, MA., 1981.

[42] H. Murase and S. Nayar, "Visual learning and recognition of 3-D objects from appearance," *Int. J. Computer Vision*, vol. 14, no. 1, pp. 5-24, 1995.

[43] 苗村，金子，原島, "光線記述に基づく仮想光源画像の合成，"映像情報メディア学会誌, vol. 52, no. 9, pp. 1328-1335, 1998.

[44] J. Nameroff, E. Simoncelli, and J. Dorsey, " Efficient re-rendering of naturally illuminated environments," *Proc. 5th Eugograohics Workshop on Rendering*, 1994.

[45] S. K. Nayar, K. Ikeuchi, and T. Kanade, "Surface Reflection: physical and geometrical perspectives," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, no. 7, pp. 611-634, 1991.

[46] S. K. Nayar, K. Ikeuchi, and T. Kanade, "Determining shape and reflectance of hybrid surfaces by photometric sampling," *IEEE Trans. Robotics and Automation*, vol. 6, no. 4, pp. 418-43, 1990.

[47] K. Nishino, Z. Zhang, and K. Ikeuchi, "Determining Reflectance Parameters and Illumination Distribution from Sparse Set of Images for View-dependent Image Synthesis," *Proc. IEEE Intl. Conf. Computer Vision '01*, pp. 599-606, Jul. 2001.

[48] T. Okabe, I. Sato, and Y. Sato, "Spherical Harmonics vs. Haar Wavelets: Basis for Recovering Illumination from Cast Shadows," *Proc. IEEE Conf. Computer Vision and Pattern Recognition '04*, pp.I-50-57, 2004.

[49] G. Patow and X. Pueyo, " A survey of inverse rendering problems, " *Computer Graphics forum*, 22(4), pp. 663-687, 2003.

[50] A. P. Pentland, "Linear Shape From Shading," *Intl. J. Computer Vision*, Vol.4, No.2, pp.153-162, 1990.

[51] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, Cambridge, 1988.

[52] R. Ramamoorthi and P. Hanrahan, "A signal-procession framework for inverse rendering", *Proc. SIGGRAPH '01*, pp.117-128, 2001.

[53] R. Ramamoorthi and P. Hanrahan, "On the relationship between radiance and irradiance: Determining the illumination from images of a convex lambertian object," *J. Optical Society of America A*, vol. 18, no. 10, pp. 2448-2459, 2001.

[54] R. Ramamoorthi and P. Hanrahan, "An efficient representation for irradiance environment maps," *Proc. SIGGRAPH '01*, pp. 497-500, 2001.

[55] R. Ramamoorthi, "Analytic PCA construction for theoretical analysis of lighting variability in images of a lambertian object"' *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 10, pp. 1322-1333, 2002.

[56] R. Ramamoorthi and P. Hanrahan, "Frequency space environment map rendering"' *Proc. SIGGRAPH '02*, pp. 517-526, 2002.

[57] Y. Sato, M. D. Wheeler, and K. Ikeuchi, "Object shape and reflectance modeling from observation," *Proc. of SIGGRAPH '97*, pp. 379-387, Aug. 1997.

[58] A. Shashua, "On photometric issues in 3D visual recognition from a single image," *Int. J. Computer Vision*, vol. 21, pp. 99-122, 1997.

[59] C. Schoeneman, J. Dorsey, B. Smits, J. Arvo, and D. Greenburg, "Painting with Light,"*Proc. SIGGRAPH '93*, pp. 143-146, 1993.

[60] S. A. Shafer and T. Kanade, "Using Shadows in Finding Surface Orientations," *Computer Vision, Graphics, and Image Processing*, Vol.22, No.1, pp. 145-176, 1983.

[61] H. Y. Shum and S. B. Kang, "A Review of Image-based Rendering Techniques," *Proc. IEEE/SPIE Visual Communications and Image (VCIP) 2000*, pp. 2-13, Perth, Jun. 2000.

[62] P.-P. Sloan, J. Kautz, and J. Snyder, "Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments," *Proc. SIGGRAPH '02*, pp. 527-536, 2002.

[63] A. State, G. Hirota, D. T. Chen, W. F. Garrett, and M. A. Livingston, "Superior augmented-reality registration by integrating landmark tracking and magnetic tracking," *Proc. SIGGRAPH '96*, pp. 429-438, Aug, 1996.

[64] P.C. Teo, E.P. Simoncelli and D.J. Heeger. "Efficient Linear Re-rendering for Interactive Lighting Design," Report No. STAN-CS-TN-97-60. Dept. Computer Science. Stanford University. Oct. 1997.

[65] S. Tominaga and N. Tanaka, "Estimating Reflectance Parameters from a Single Color Image," *J. IEEE Computer Graphics & Applications*, vol.20, No. 5, 2000, pp.58-66.

[66] C.Tomasi and T. Kanade, "Shape and Motion from Image Streams under Orthography: a Factorization Method," *Int. J. Computer Vision*, vol. 9, no. 2, pp. 137-154, 1992.

[67] K. E. Torrance and E. M. Sparrow, "Theory for off-specular reflection from roughened surface," *J. Optical Society of America*, vol.57, pp.1105-1114, 1967.

[68] R. Tsai, "A Versatile Camera Calibration Technique for High Accuracy Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses," *IEEE J. Robotics and Automation*, vol. 3, no. 4, pp. 323-344, 1987.

[69] G. J. Ward, "Measuring and modeling anisotropic reflection," *Proc. SIGGRAPH '92*, pp. 265-272. 1992.

[70] S. Westin, J. Arvo, K. Torrance, "Predicting Reflectance Functions from Complex Surfaces, " *Proc. SIGGRAPH '92*, pp. 255-264, 1992.

[71] Y. Yu, P. Devebec, J. Malik, and T. Hawkins, "Inverse Global Illumination: Recovering reflectance Models of Real Scenes from Photographs", *Proc. SIGGRAPH '99*, Aug. 1999, pp.215-224.

[72] A. Yuille, D. Snow, R. Epstein, and P. Belhumeur, "Determining generative models of objects under varying illumination: Shape and albedo from multiple images using SVD and integrability," *Int. J. Computer Vision*, vol. 35, no. 3, pp. 203-222, 1999.

[73] L. Zhao and Y. Yang, "Theoretical analysis of illumination in PCA-based vision systems," *Pattern Recognition*, vol. 32, pp. 547-564, 1999.

[74] Y. Zang and Y. Yang, "Illuminant Direction Determination for Multiple Light Sources," *Proc. IEEE Conf. Computer Vision and Pattern Recognition '00*, pp. 269-276, Jun. 2000.

[75] 3D Construction Company, http://www.3dcomstruction.com

[76] Radiance is Unix freeware for lighting design and rendering developed by the U.S. Department of Energy and the Swiss federal government. http://hobbes.lbl.gov/radiance/

Chapter 9

Appendix: Non-Photorealistic Shading Model For Artistic Shadings

As an application of our proposed inverse lighting method described in Chapter 4, we present a new technique for superimposing synthetic objects onto oil paintings with artistic shadings that are consistent with those originally painted by the artists.

In a colored medium such as oil painting, artists often use color shift techniques for adding some artistic tones to their paintings as well as for enlarging their dynamic ranges. For instance, in the painting by Vincent van Gogh shown in Figure 9.2, a non-photorealistic blue to white color transition is observed inside shadows cast by the book placed on the wooden table. Those kinds of shadows cannot be synthesized by using traditional rendering techniques.

In this chapter, we attempt to determine the mechanisms for color shifts performed by artists and to automate their processes so that we can superimpose onto paintings synthetic objects that have consistent shadings. To discover how intrinsic color shifts were performed by artists, we first study characteristics of shadows observed both in real scenes and in paintings. In particular, we analyze HSV (hue, saturation, value) color distributions inside shadows observed in paintings. We then adapt the acquired mechanisms to superimpose synthetic objects with consistent shadings onto oil paintings.

9.1 Shadows in Real Scenes

From (4.5), the influence of an object onto another object surface in a real scene can be considered as follows. First, scene radiance $r(\theta_e, \phi_e)$ of a surface point viewed from the direction (θ_e, ϕ_e) is computed in the case where there is no object that occludes any incoming light at the surface point (Figure 9.1 (a)).

$$r(\theta_e, \phi_e) = \sum_{i=1}^{n} f(\theta_i, \phi_i; \theta_e, \phi_e) L(\theta_i, \phi_i) \cos\theta_i \tag{9.1}$$

where n is the number of sampling directions of illumination radiance, and $L(\theta_i, \phi_i)$ is the illumination radiance per unit solid angle $\delta\omega = 2\pi/n$ coming from the direction (θ_i, ϕ_i).

Second, scene radiance $r'(\theta_e, \phi_e)$ at the surface point in the case where an object in the scene occludes some of the incoming light is computed as (Figure 9.1 (b)).

$$r'(\theta_e, \phi_e) = \sum_{i=1}^{n} f(\theta_i, \phi_i; \theta_e, \phi_e) L(\theta_i, \phi_i) S(\theta_i, \phi_i) \cos\theta_i \tag{9.2}$$

where $S(\theta_i, \phi_i)$ represents the occlusion of incoming light $L(\theta_i, \phi_i)$ by the object.

Then, the ratio of r' to r is computed. This ratio represents how much of the radiance at the surface point would still be preserved if the object were to be placed in the scene. We refer to $\frac{r'}{r}$ as the *shadow ratio*.

Let I be a color of each surface point we observe in an image of the scene in case of r. Then by multiplying the *shadow ratio* $\frac{r'}{r}$ to I, the color I' that would be the color of the surface if there were an object placed in the scene can be computed as

$$I' = I\frac{r'}{r} \tag{9.3}$$

Note that I' always becomes proportional to $\frac{r'}{r}$, and $\frac{I'}{I}$ is always equal to $\frac{r'}{r}$.

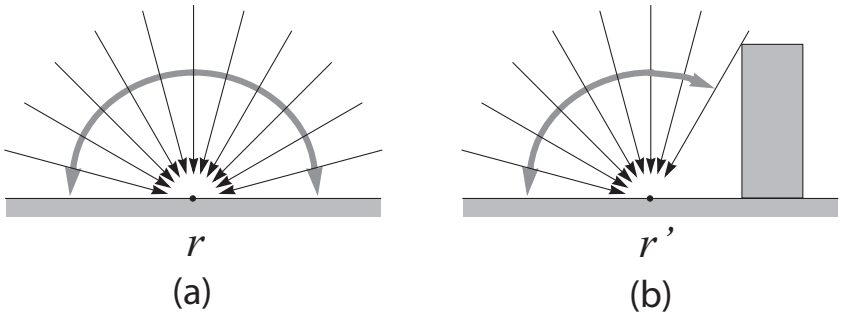


Figure 9.1: Scene radiance (a)without objects (b)with objects.

9.2 Shadows in Oil Paintings

There are some rules in conventional paintings, such as perspective, anatomy, and modeling with a scene with darker shadows. The impressionists attempted to break some of these rules. In particular, they changed the color of shadows such as gray or black to complementary hues. Therefore, image brightness observed inside shadows in their paintings changes differently from that observed in real scenes. We clearly see this in the painting by Vincent van Gogh shown in Figure 9.2, as a non-photorealistic blue to white color transition observed inside shadows cast by the book placed on the wooden table.

If we superimpose a synthetic object onto an oil painting without considering the color modifications performed by an artist, the rendered appearances and shadows of the synthetic object result in something quite different from those originally painted by the artist. Thus it is essential to first determine the intrinsic mechanisms for color modifications performed by an artist and then to use them for rendering a synthetic object superimposed onto the input painting.

Figure 9.2: "Still Life: Drawing Board, Pipe, Onions and Sealing-Wax" by Vincent van Gogh.

9.2.1 Color Modification Function

It is known that artists sometimes prefer to use the HSV (hue, saturation, value) color model than other models such as RGB (red, green, blue) or $CMYK$ (cyan, magenta, yellow, black), because of the similarity of the HSV color model to the way humans perceive color. From this, we analyze color modifications made by artists using the HSV color model.

Figure 9.3 shows the HSV color of the painting by Vincent van Gogh. Images containing each component of HSV are referred to as *Himage*, *Simage*, and *Vimage* respectively. In the *Himage* and the *Simage* of the painting by Vincent, we see that the hue and saturation of the painting greatly change inside the shadows cast by the objects on the table. From this, it seems reasonable to suppose that the changes in hue and saturation are related to the *shadow ratio* of corresponding surface points on the shadow surface.

In this study, we define a modification function $F_h(\frac{r'}{r})$ that represents the change in hue for a particular shadow ratio $\frac{r'}{r}$ as

$$F_h(\frac{r'}{r}) = h' - h, \tag{9.4}$$

where h represents the hue observed for a surface point when its scene radiance is r, and h' represents that observed for a surface point when its scene radiance is r'. In other words, the function $F_h(\frac{r'}{r})$ shows how the hue of the surface is shifted when the scene radiance of the shadow surface changes from r to r'.

Similarly, the changes in saturation and value are defined as

$$F_s(\frac{r'}{r}) = \frac{s'}{s}, \tag{9.5}$$

$$F_v(\frac{r'}{r}) = \frac{v'}{v}. \tag{9.6}$$

Here s and v represent the saturation and value for a surface point with its scene radiance r, and s' and v' represent those for a surface point with its scene radiance r'.

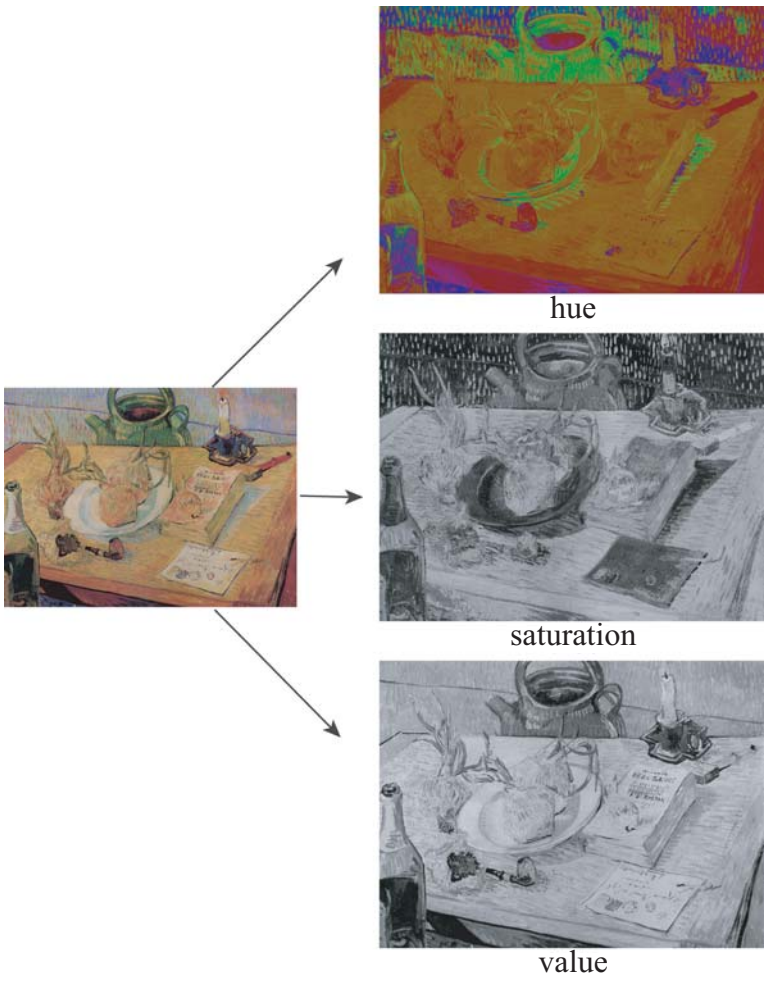


Figure 9.3: Color conversion from RGB (red, green, blue) to HSV (hue, saturation, value).

9.3 Obtaining Color Modification Function for a Painting

We analyze mechanisms for the color modification performed by an artist as described in the following steps.

1. We select an occluding object and a shadow surface in an input painting. Then the 3D shapes of the occluding object and the shadow surface are provided by using the same modeling tool as in Chapter 4 for this purpose (Figure 9.4).

2. The shadows cast by the occluding object in the *Vimage* are used for recovering an illumination distribution of the scene by using the method described in Chapter 5. Assuming that there are only white light sources in the scene, what we need to estimate is the radiance of those light sources.[1] We also assume the shadow surface to be Lambertian, and the diffuse parameter is set to be the pixel value of the brightest point on the shadow surface.

3. Cast shadows of the occluding object are simulated by using the estimated illumination distribution. Assuming uniform Lambertian reflectance on the shadow surface, the hue of the shadow surface, h in (9.4) is set to be the pixel value of a surface point in the *Himage* where its scene radiance is r.[2] Similarly, the saturation and value of the shadow surface, s and v in (9.5) and (9.6), are set to the pixel values of this surface point in the *Simage* and the *Vimage* respectively.

4. For other points on the shadow surface, their shadow ratios $\frac{r'}{r}$ are computed based on the scene geometry and the estimated illumination distribution of the scene. Then their h', s', and v' that are observed values in *Himage, Simage*, and *Vimage* respectively, are substituted into (9.4), (9.5), (9.6), and the function values $F_h(\frac{r'}{r}), F_s(\frac{r'}{r})$, and $F_v(\frac{r'}{r})$ are stored in the computer 's memory for the given shadow ratio $\frac{r'}{r}$.

5. Using all points in the painting, we are able to obtain a set of function values for different shadow ratios. This enables us to analyze how the color modifications were made by the artist as the shadow ratio changes. Finally, for each of $F_h(\frac{r'}{r}), F_s(\frac{r'}{r})$,

[1] Since no knowledge of how ratio conversions work in the painting is available at first, an assumption regarding the color of light sources is required.

[2] In other words, h shows the hue of a surface point on the shadow surface where none of the objects in the scene occludes any incoming light at this surface point.

and $F_v(\frac{r'}{r})$, we compute a parametric function that enables us to perform the color modifications made by the artist using a small number of parameters (Section 9.3.2).

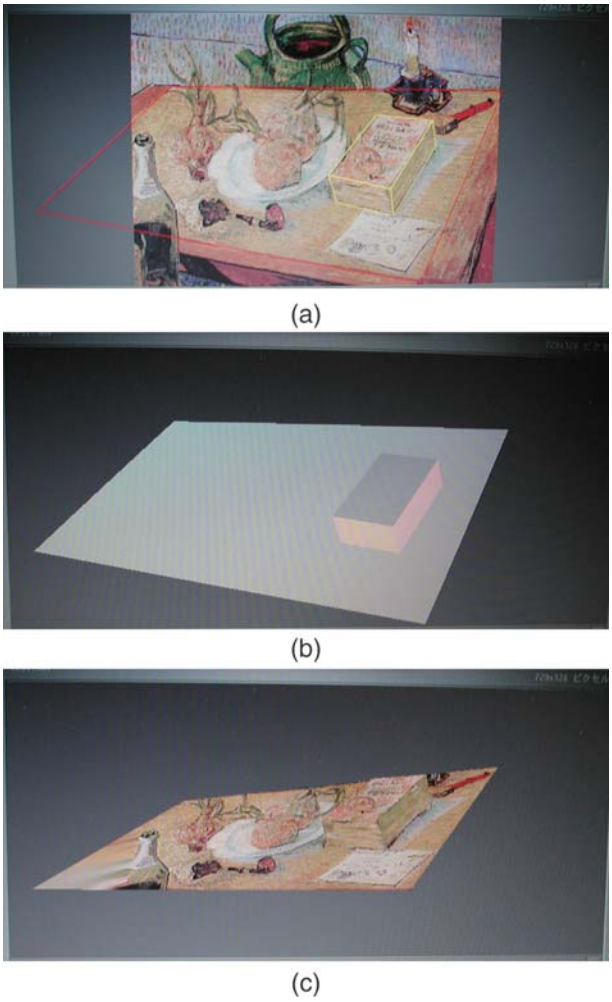


Figure 9.4: Reconstruction of the scene: (a) original photograph with marked edges indicated, (b) recovered scene geometry, (c) textured-mapped view from a different viewing position.

9.3.1 Obtained Color Modification Functions

Figure 9.5 numerically shows the color modifications made by Vincent van Gogh (left) and by Rembrandt (right). In the plots shown in this figure, the horizontal axis represents the shadow ratio $\frac{r'}{r}$ within the range of 0.0 to 1.0, which represents how the surface radiance at the surface point would be changed if we place the occluding object in the scene. The vertical axis represents function values $F_h(\frac{r'}{r}), F_s(\frac{r'}{r})$, and $F_v(\frac{r'}{r})$ for the corresponding shadow ratio $\frac{r'}{r}$ respectively.

In the plot of $F_h(\frac{r'}{r})$ obtained from the painting by Vincent van Gogh, the hue of the painted shadows changes gradually from the color of the wooden desk (yellow) to its complementary hue (blue) as the shadow ratio gets closer to 0.0. Also, its saturation changes non-linearly as we seen in the plot of $F_s(\frac{r'}{r})$ when the shadow ratio is around 0.6. These are reasons why we see a gradual blue to white color transition inside shadow regions in the painting by Vincent van Gogh.

In contrast, in the case of the painting by Rembrandt that gives with a realistic description, the hue and saturation of the shadow surface do not change much, and the plot of $F_v(\frac{r'}{r})$ shows a linear transition. A Lambertian surface normally shows a similar plot in a real image.

9.3.2 Approximating Color Conversion Functions

We derive functions that simulate the modifications performed by an artist so that we can automatically superimpose synthetic objects onto his or her painting with consistent shadings.

The function to be used for approximating artists' color modifications is defined as

$$f_c(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + a_5 x^5 \tag{9.7}$$

$$c = h, s, v$$

where x corresponds to a shadow ratio.

From the plots $F_h(\frac{r'}{r}), F_s(\frac{r'}{r})$, and $F_v(\frac{r'}{r})$ obtained above, the coefficients $a_0, a_1, a_2, a_3, a_4, a_5$ are estimated by minimizing the sum of the squared difference between the observed values $F_c(\frac{r'}{r})$ and the computed function values $f_c(\frac{r'}{r})$ in (9.7) for a set of shadow ratio $\frac{r'}{r}$. The

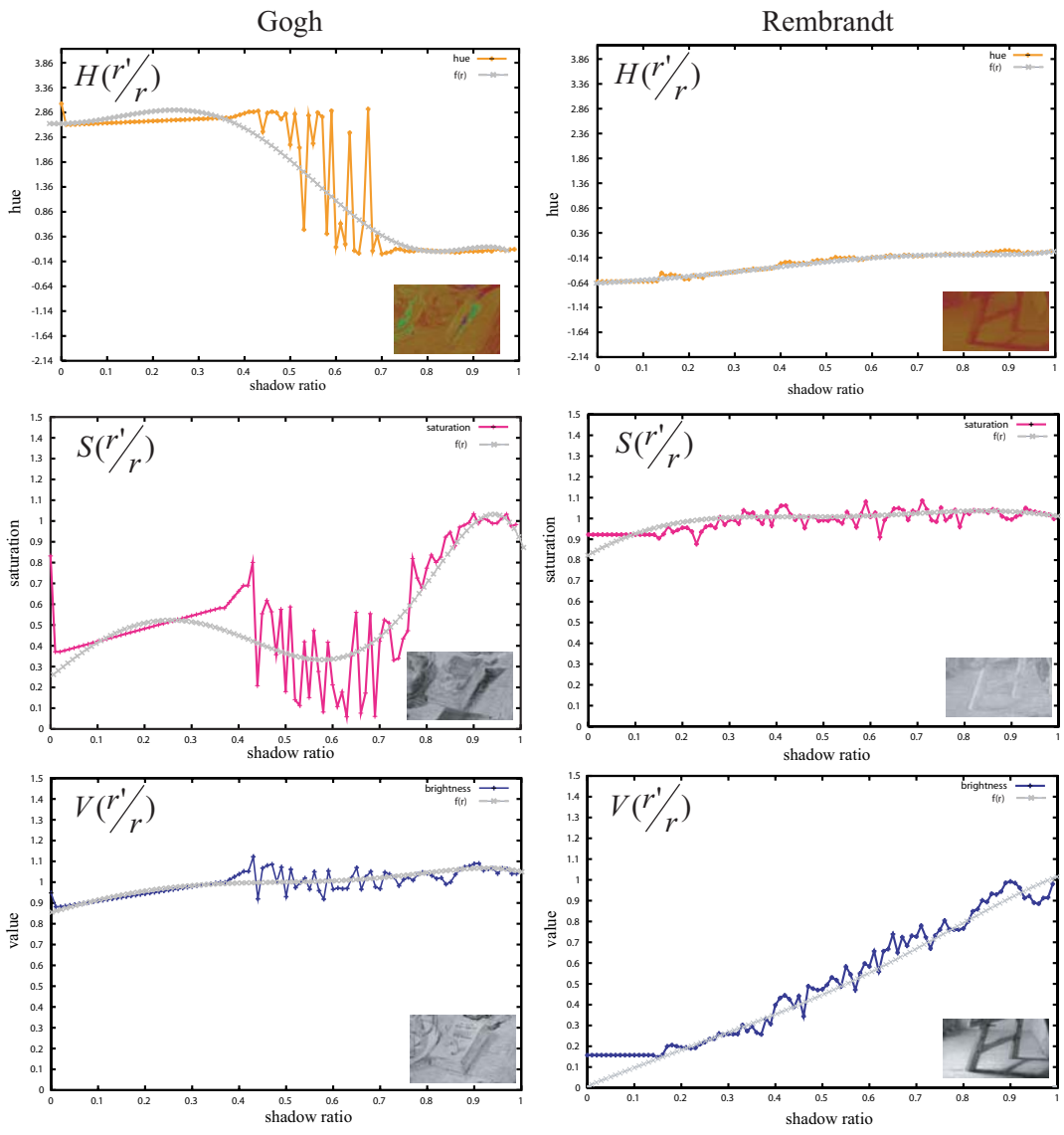


Figure 9.5: The color modifications made by Vincent van Gogh and Rembrandt: $f(\frac{r'}{r})$ show the estimated functions that approximate the original plots.

objective function E_c to be minimized is thus defined as

$$E_c = \sum_{i=0}^{n} (F_c(x_i) - f_c(x_i))^2 \tag{9.8}$$

$$c = h, s, v$$

where n is the number of samplings of shadow ratios.

In our experiments, this objective function E_c is minimized with respect to the coefficients $a_0, a_1, a_2, a_3, a_4, a_5$ by using a standard MATLAB function to obtain the best estimation of those coefficients.

The gray lines in Figure 9.5 show plots computed from the estimated parametric functions of the paintings by Vincent van Gogh and Rembrandt. In these figures, we can see that the estimated functions approximate the original plots fairly well.

9.4 Superimposing Synthetic Objects onto Paintings

Based on the modification functions we obtained so far, a synthetic object is superimposed onto a painting by using the ray casting algorithm as follows:

1. For each image pixel in the painting, a ray is extended from the projection center of the painting through the pixel using the camera parameters obtained at the modeling step of the scene geometry.[3]

2. If the ray intersects a synthetic object, we consider that the pixel corresponds to a point on the synthetic object surface and compute a color to be observed at the surface point under the estimated illumination distribution of the scene. Then the computed color is stored in the pixel as the surface color of the synthetic object at the pixel (Section 9.4.1).

3. Otherwise, we consider that the pixel corresponds to a point on a object surface originally painted by the artist referred to as the *painted surface*, and modify an observed color at the point on the painted surface so that shadows cast by the synthetic object onto the painted surface are generated (Section 9.4.2).

[3] At the modeling step of the scene geometry, the camera parameters of the painting are also recovered by using a photo-modeling tool interactively.

9.4.1 Rendering the Synthetic Object Surface

For computing the color on the synthetic object surface, we first compute surface radiance r and r' at the surface point under the estimated illumination distribution from (9.1) and (9.2), respectively, to obtain a shadow ratio $\frac{r'}{r}$ at this surface point.[4]

Then the function values $F_c(\frac{r'}{r})(c = h, s, v)$ are computed by assigning the obtained shadow ratio $\frac{r'}{r}$ to the modification function $f_c(\frac{r'}{r})$ in (9.7). This color modification is necessary to take self-occlusions of the synthetic object into consideration. In addition, h, s, and v of this surface point are computed for the case where the irradiance of this surface point is r.

Finally, we obtain h', s', and v' that would be the color of the surface point in the case where its surface radiance becomes r' and the color modifications were made by the artist according to this shadow ratio as

$$h' = h + F_h(\frac{r'}{r}) \tag{9.9}$$

$$s' = sF_s(\frac{r'}{r}) \tag{9.10}$$

$$v' = vF_v(\frac{r'}{r}) \tag{9.11}$$

Then the RGB color of this surface point is computed from h', s', v' and stored in the pixel as the surface color of the synthetic object at the pixel.

9.4.2 Shadows Cast by a Synthetic Object

In the case where the ray through an image pixel does not intersect with the synthetic object, the pixel corresponds to a point on the painted surface, and the color of the pixel needs to be modified so that shadows cast by the synthetic object are created on the painted surface.

As in the previous section, we first compute surface radiance r and r' to obtain a shadow ratio $\frac{r'}{r}$ at this surface point.[5] Then, the values of $F_h(\frac{r'}{r}), F_s(\frac{r'}{r})$, and $F_v(\frac{r'}{r})$ are computed by assigning the obtained shadow ratio $\frac{r'}{r}$ to the modification function $f(\frac{r'}{r})$ in (9.7).

---

[4] Here the occlusion coefficients $S(\theta_i, \phi_i)$ in (9.2) represent self-occlusions by the object.

[5] The occlusion coefficients $S(\theta_i, \phi_i)$ in (9.2) represent occlusions of incoming light by the synthetic object at the painted surface.

Finally, given h, s, and v of the painted surface, we obtain the color h', s' and v' that would be the color of the painted surface if the synthetic object were placed in the scene from (9.9), (9.10), and (9.11) respectively.

9.5 Synthesized Results

Synthesized results are shown in Figure 9.6. In Figure 9.6 (b), a synthetic object with the same shape as that of the occluding object was rendered on a synthetic surface with the same shape and reflectance as the shadow surface. In this result, the shadows cast by the synthetic object strongly resemble those of the occluding object in the original painting. This shows not only that the estimated illumination distribution gives a good representation of the characteristics of the scene but also that the computed ratio conversion functions work well.

We also superimposed a synthetic object of a different shape onto the scene in Figure 9.6 (c). In the images synthesized by our method, the synthetic object casts artistic shadows on the wooden table that are similar to those of the other objects originally painted by the artist, and this shows that the color modifications made by the artist are well approximated by the color modification functions obtained by our method.

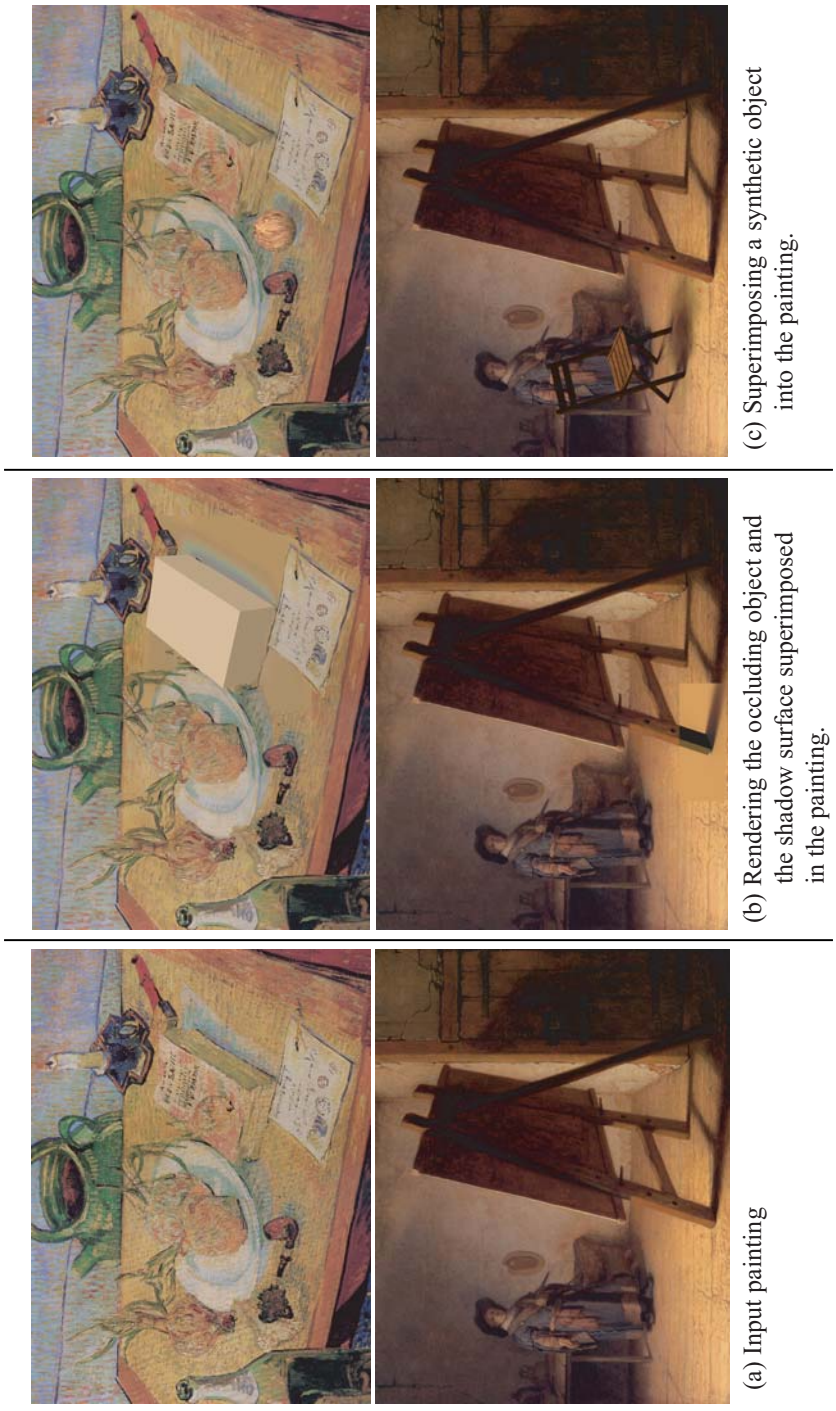


Figure 9.6: Results: (a) input paintings, (b) Synthesized occluding objects and shadow surfaces, and (c) Synthesized new objects superimposed into the painting.

List of Publications

論文誌

- I. Sato, Y. Sato, and K. Ikeuchi, "Acquiring a Radiance Distribution to Superimpose Virtual Objects onto a Real Scene," *IEEE Trans. Visualization and Computer Graphics*, Vol. 5, No. 1, pp. 1-12, January-March 1999.

- I. Sato, Y. Sato, and K. Ikeuchi, "Illumination from shadows," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 25, No. 3, pp. 290-300, March 2003.

- K. Ikeuchi,, M. Sakauchi, H. Kawasaki, and I. Sato, "Constructing Virtual Cities by Using Panoramic Images," *Int'l J. Computer Vision*, Vol. 53, No. 3, pp. 237-247, March 2004.

- 佐藤いまり, 佐藤洋一, 池内克史, "全方位ステレオによる実光源環境の計測とそれにもとづく仮想物体の実画像への重ね込み, "電子情報通信学会論文誌 D-II, Vol. J81-D-II, No. 5, pp. 861-871, May 1998.

- 池内克史, 佐藤洋一, 西野恒, 佐藤いまり, "複合現実感における光学的整合性の実現, "日本バーチャルリアリティ学会論文誌「複合現実感」特集号, Vol.4, No. 4, pp. 623-630, December 1999.

- 佐藤いまり，佐藤洋一，池内克史，"物体の陰影に基づく光源環境の推定，" 情報処理学会論文誌: コンピュータビジョンとイメージメディア，Vol. 41，No. SIG 10, pp. 31-40,December 2000.

- 佐藤いまり，林田守広，甲斐郁代，佐藤洋一，池内克史，" 実光源環境下での画像生成：基礎画像の線形和による高速レンダリング手法，"電子情報通信学会論文誌 D-II, Vol. J84-D-II, No. 8, pp. 1864-1872, Aug 2001.

- 池内克史，坂内正夫，川崎洋，高橋拓二，村尾真洋，佐藤いまり，甲斐郁代,"全方位画像による仮想都市空間の生成，" 情報処理学会論文誌: コンピュータビジョンとイメージメディア，Vol. 42，No. SIG 14, pp. 49-58, December 2001.

- 佐藤いまり, 佐藤洋一, 池内克史，"絵画の陰影特徴解析に基づく仮想物体の絵画への重ね込み," 情報処理学会論文誌 コンピュータビジョンとイメージメディア, Vol.

44, No. SIG9, pp. 132-141, July 2003.

国際会議

- I. Sato, Y. Sato, and K. Ikeuchi, "A Method for Estimating Illumination Distribution of a Real Scene Based on Soft Shadows," *Proc. Int'l Conf. Advanced Multimedia Content Processing*, pp. 47-61, November 1998.

- I. Sato, Y. Sato, and K. Ikeuchi, "Acquiring a Radiance Distribution to Superimpose Virtual Objects onto a Real Scene," *Proc. IAPR Workshop on Machine Vision Applications*, pp. 19-22, November 1998.

- I. Sato, Y. Sato, and K. Ikeuchi, "Ilumination Distribution from Shadows," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 306-312, June 1999.

- I. Sato, Y. Sato, and K. Ikeuchi, "Illumination Distribution from Brightness in Shadows: Adaptive Estimation of Illumination Distribution with Unknown Reflectance Properties in Shadow Regions," *Proc. IEEE Int'l Conf. Computer Vision*, pp. 875-882, September 1999.

- T. Kim, K. Hong, I. Sato, and K. Ikeuchi, "Enhanced Augmented Reality with Shadows in Naturally Illuminated Environments", *Proc. IAPR Workshop on Machine Vision Applications*, pp. 27-30, November 2000.

- I. Sato, Y. Sato, and K. Ikeuchi, "Stability Issues in Recovering Illumination Distribution from Brightness in Shadows," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. II-400-407, December 2001.

- I. Sato, T. Okabe, Y. Sato, and K. Ikeuchi, "Appearance sampling for obtaining a set of basis images for variable illumination," *Proc. IEEE Int'l Conf. Computer Vision*, pp. 800-807, October 2003.

国内シンポジウム

- 佐藤いまり, 佐藤洋一, 池内克史,"光学的整合性を考慮した実画像への仮想物体の重ね込み、"第 3 回知能情報メディアシンポジウム, pp. 23-32, December 1997.

- 佐藤いまり, 佐藤洋一, 池内克史,”実世界型媒介空間のための光源環境の構築：ソフトシャドウからの光源分布の推定”情報処理学会 画像の認識・理解シンポジウム, pp. 287-294, vol.1, July 1998.

- 佐藤いまり, 林田守広, 甲斐郁代, 佐藤洋一, 池内克史,”複合現実感における光学的整合性の実現：基礎画像の線形和による高速レンダリング手法, ”画像の認識・理解シンポジウム, pp. I-107-122, July 2000.

- 佐藤いまり, 佐藤洋一, 池内克史, ”絵画における陰影特徴の解析とその画像合成への応用”, 画像の認識と理解シンポジウム, pp. II-421-428, July 2002.

- 佐藤いまり, 岡部孝弘, 佐藤洋一, 池内克史, ”任意光源下における画像生成のための基底画像の獲得”, 画像の認識と理解シンポジウム, pp. II-157-162, July 2004.