Light Field Analysis for Modeling Image Formation

Chia-Kai Liang, *Member, IEEE*, Yi-Chang Shih, and Homer H. Chen, *Fellow, IEEE*

Abstract—Image formation is traditionally described by a number of individual models, one for each specific effect in the image formation process. However, it is difficult to aggregate the effects by concatenating such individual models. In this paper, we apply light transport analysis to derive a unified image formation model that represents the radiance along a light ray as a 4-D light field signal and physical phenomena such as lens refraction and blocking as linear transformations or modulations of the light field. This unified mathematical framework allows the entire image formation process to be elegantly described by a single equation. It also allows most geometric and photometric effects of imaging, including perspective transformation, defocus blur, and vignetting, to be represented in both 4-D primal and dual domains. The result matches that of traditional models. Generalizations and applications of this theoretic framework are discussed.

Index Terms—Defocus blur, image formation, light field, light transport analysis, perspective transformation, vignetting.

I. INTRODUCTION

MODELING the appearance of a scene on an image is the most fundamental step in image processing and computer vision. It entails the use of the properties of scene (geometry, reflectance, etc.) and imaging system (location, lens, etc.) to control image formation. The scene appearance models characterize the scene properties and map the scene to the image.

Traditional image formation modeling describes each specific imaging phenomenon by an individual model. For example, the classic pinhole camera model describes the perspective projection part of an image formation process. Since the image formation process usually involves various geometric and photometric effects, such as defocus blur and vignetting, it is often difficult to describe the combined effect by concatenating the individual models, if possible at all.

In this paper, we develop a unified framework for modeling the image formation process by representing the radiance along the light ray as a high-dimensional light field. An ordinary image is a projected signal of the light field. With this framework, the transportation of the radiance from the object surface to the image plane can be simply formulated as linear transformations or modulations of the light field. Furthermore, these various light field manipulations can be combined into a single operation to account for the aggregation of photographic effects. The light field representation also enables spectral analysis of the radiance distribution, leading to mathematically more elegant analysis of the photographic effects in the frequency domain.

A unified framework of image formation is appealing. It allows all photographic effects to be modeled on the same base and facilitates a better estimation of the scene and imaging parameters. It encompasses all elements of the light transport process of an imaging system and is applicable to different kinds of imaging systems. Although light transport analysis has been applied to many image-related problems, a unified framework for image formation has not been proposed till now.

The rest of the paper is organized as follows. Section II describes the related work on image formulation and light transport analysis. Section III presents various light field operations that serve as the basic building blocks of our framework. Section IV shows how these operations can be combined into a single photographic operation and how the transportation of light from the object surface all the way to the image sensor can be represented. Section V discusses the relation of the proposed unified framework to the traditional photographic models. The generalizations, applications, and limitations of the framework are discussed in Section VI, followed by a conclusion in Section VII.

II. RELATED WORK

A. Image Formation

Image formation modeling, which describes the appearance of objects on an image, has been studied for decades, and the basic models can be found in computer vision and image processing books [1], [2].

Existing image formation models can be roughly classified into two categories. The first category deals with the geometrical aspect of image formation. For example, the projection of a 3-D scene onto a 2-D image is modeled as a perspective transformation, and the image of an out-of-focus scene is modeled as the convolution of a sharp image with a spatially variant low-pass kernel that is a function of object distance and camera parameters [3]. The second category deals with the photometric aspect of image formation. For example, the vignetting effect is usually modeled by a 2-D smooth function of view angle, surface orientation, and camera parameters [1], [4].

While having been widely applied, such models have several fundamental problems. First, as these models were separately developed to account for different photographic effects, an integrated treatment of the effects is rather difficult to do. For example, while the defocus blur and the vignetting are functions

C.-K. Liang and H. H. Chen are with the Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan, 10617 (e-mail: homer@cc.ntu.edu.tw).

Y.-C. Shih was with the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, 10617. He is now with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

of the camera parameters, they cannot be analyzed together with existing models. Second, most models are confined to the 2-D image domain, which is not an ideal domain for analysis of a number of photographic effects. For instance, an occluded object can be partially observed by a lens camera because the finite-size aperture induces a small parallax, but this effect cannot be analyzed in the image domain [6]. Third, while physically all photographic effects occur simultaneously, existing models can only be sequentially applied to the image. To our knowledge, there is not yet any study about which order these models can be applied to best describe the entire image formation process. Last but not least, many models assume that a perfect image can be used as the reference. However, this is not always true in practice.

B. Light Transport Analysis

By representing the radiance of a light ray by a plenoptic function [7] or light field [8], we can apply standard signal processing techniques, such as convolution and spectral analysis [9], to analyze the light transport. This concept was first used to analyze the bandwidth of the light field. Chai *et al.* [11] showed that the spectral support of the light field depends upon the range of the scene depth and, thus, knowledge of scene geometry helps improve the sampling and reconstruction efficient. In [12], a depth-dependent reconstruction filter is proposed to suppress the aliasing effect. Note, however, that only the propagation of light rays in free space is considered in these studies.

Ramamoorthi and Hanrahan modeled the surface reflection as a convolution [13], [14]. Typically, the surface acts as a low-pass filter on the incoming light. Therefore, the lighting and the reflectance properties (precisely, the bidirectional reflectance distribution function) of the surface can be recovered from the image by a simple deconvolution. This also implies that the surface reflection may be efficiently evaluated in the frequency domain [15], which is the basic idea of many precomputed radiance transfer algorithms [16].

The occlusion of the object in the path of a light ray may be modeled as a modulation of the light field [17]–[20]. An analytical representation of the irradiance field due to diffuse emitters and occluders is derived in [17]. The effect of modulation is exploited to predict the bandwidth of the local light field for efficient sampling [18], and the occluder is recovered from the modulated light field [19]. In this paper, we describe the effect of an aperture on the light field as modulation and discuss its relation with the defocus blur.

Light transport analysis is a technique often employed in the design of light field cameras and displays [21]–[26]. Ng showed that a photo is equivalent to an integral projection of the light field, or a slice of the light field spectrum [26]. Therefore, by properly modulating or transforming the light field such that all the information falls on the slice, a light field can be captured with a 2-D sensor array [24].

The framework presented here is significantly different from previous work in several ways. First, we present a full derivation of the light transport from the object surface all the way to the image plane, including the blocking effect of the aperture. On the contrary, previous works usually present or utilize a small part of this framework. For example, the effect of the aperture

is discussed in [21]. In [22], the light transport equation without aperture is presented, but no derivation is provided. Second, unlike previous work that only considers geometric transformation, we include photometric transformation in the light transport process as well to account for effects such as vignetting. Third, we show that the photographic effects such as perspective transformation and defocus blur can be modeled in the unified framework.

During the submission of the paper, Levin *et al.* concurrently developed a related analysis to design a lattice-focal lens for extending the depth-of-field [27]. There are a few notable differences between their work and ours. First, we model the lens refraction and aperture separately for better flexibility, while they model them together as a single 4-D convolution for a simpler representation. Second, while they focus on the analysis of defocus blur in many different imaging systems, we develop a framework to describe various photographic effects in the standard thin-lens camera.

III. LIGHT TRANSPORT OPERATIONS

In this section, we describe the representation of light field and basic light transport operations. Like the traditional models, we consider the effects of geometric optics [1], for which the object and the space of concern are much larger than the wavelength of the light, and neglect the wave effects such as diffraction and interference. We first describe how to represent the radiances of light rays as a light field. Then, we describe the light field operations corresponding to the propagation, refraction, and blocking of light rays. We also use Fourier analysis to interpret these operations in the frequency domain. Finally, we show that these operations can be combined into a single one and that the backward light transport can be easily derived. For conciseness, we only present the light transport operations needed to model the image formation process. Other light transport operations and analyses can be founded in [14] and [18].

A. Light Field Representation

In geometric optics, if a light ray can be uniquely parameterized, the representation of its radiance value can be easily determined. Generally, a light ray in 3-D space can be represented by a point (3-D) the light ray passes through and the direction (2-D) of the light ray. Hence, the radiance of a light ray is a sample of a 5-D signal, which is the *plenoptic function* [7].

However, the radiance measured along a light ray is constant if the light ray is not blocked by any occluding object. Therefore, when a light ray leaves the convex hull of a scene, the dimensionality of the plenoptic function is decreased by one. In this case, it has been shown that each light ray can be specified by its intersections with two 2-D manifolds and that the radiance of the light ray is a sample of a 4-D signal called *light field* [8] or *lumigraph* [10]. We adopt the former name throughout the paper.

An example of the light field representation is shown in Fig. 1. Two infinite planes with independent coordinate systems are defined on the right-hand side of the scene (a toy house). Each rightward light ray emitted from the scene intersects these two planes at two points: one at the coordinates (u, v) on Plane 1 and the other at (x, y) on Plane 2. Therefore, the radiance R of
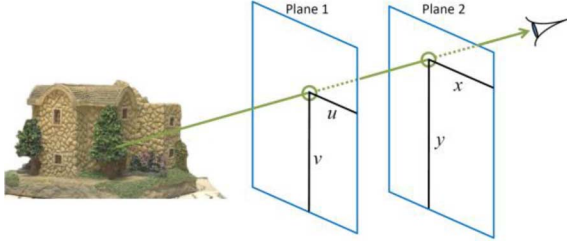
Fig. 1. 4-D representation of a light ray in 3-D space.

Fig. 2. Light field representations. (a) Plane–sphere parameterization. (b) Two-plane parameterization. (c) In-camera parameterization.

TABLE I
NOTATION

Φ	Radiant flux
x	Spatial coordinate in all parameterizations
u	Angular coordinate in two-plane parameterizations
θ	Angular coordinate in sphere-plane parameterization
$\mathbf{x}$	$[x\,u]^T$
rect(x)	Rectangular function
T	Texture function of a Lambertian surface
R	Radiance, traditional definition
I	Irradiance, traditional definition
S	Light field under the sphere-plane parameterization
L	Light field under the two-plane parameterization
$\hat{L}$	Light field under the in-camera parameterization
g	Parameterization function
j	$(-1)^{0.5}$
$\mathcal{L}$	Fourier transform of L
ω_x, ω_u	Frequencies along x and u
$\boldsymbol{\omega}$	$[\omega_x\,\omega_u]^T$
$A \otimes B$	Convolution of signals A and B

a light ray, which is a sample of the light field L, can be described by

$$R = L(u, v, x, y). \tag{1}$$

For simplicity, we derive our framework in 2-D space, where 2-D manifolds become 1-D and, thus, the light field becomes 2-D. The derivation is readily extendable to 3-D space as described in Section VI. The notation used in this paper is defined in Table I.

We use three different parameterizations to represent the light field: plane–sphere, two-plane, and in-camera. In the plane–sphere parameterization shown in Fig. 2(a), the light ray is represented by the intersection of the light ray with a reference plane X,[1] denoted by x, and the angle θ between the light ray and the normal of the reference plane. In the two-plane parameterization, the light ray is represented by its intersections with two parallel reference planes U and X as shown in Fig. 2(b), where U is at a unit distance from X. The coordinates of the intersection point u on U are defined with respect to the intersection point x on X. That is, u is defined in the local frame fixed at x. These two parameterizations are related to each other by $u = \tan \theta$. For convenience, we call the x-axis *spatial* axis and the u-axis *angular* axis since it is related to the direction of the light ray.

[1]In 2-D space, a 2-D plane is degenerated into a 1-D line, but we still call it plane for convenience.
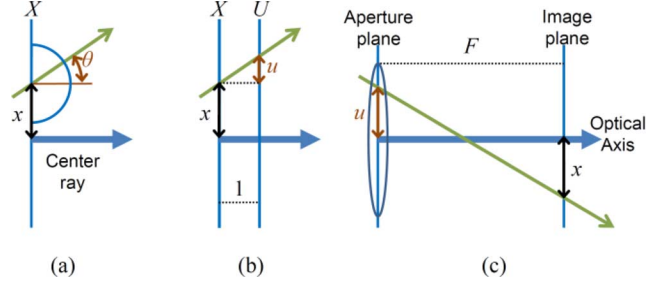
The third method, in-camera parameterization, is used to represent the light rays entering the camera, as shown in Fig. 2(c). It is similar to the two-plane parameterization but one reference plane is aligned with the aperture plane and another with the image plane. The distance between these two planes is F. The coordinate systems of these two planes are independent, but their origins are on the optical axis of the camera.

The radiances of the light rays in different parameterizations are measured in different ways. Traditionally, the radiance R along a light ray (x, θ) is measured by its power Φ per unit projected area and per unit solid angle. That is

$$R(x, \theta) = \frac{\mathrm{d}^2\Phi}{\cos\theta\,\mathrm{d}x\,\mathrm{d}\theta}. \tag{2}$$

In the plane–sphere parameterization, we remove the cosine term and define the radiance S as

$$S(x, \theta) = \frac{\mathrm{d}^2\Phi}{\mathrm{d}x\,\mathrm{d}\theta}. \tag{3}$$

In the two-plane parameterizations the radiance L is defined as

$$L([x, u]^T) = L(\mathbf{x}) = \frac{\mathrm{d}^2\Phi}{\mathrm{d}x\,\mathrm{d}u}. \tag{4}$$

We pack x and u into a vector $\mathbf{x}$ to facilitate the derivation. The conversion between these radiance representations is given in Appendix A.

Because the light field L under the two-plane parameterization is a regular 2-D signal, it can be transformed to the Fourier spectrum

$$\mathcal{L}([\omega_x, \omega_u]^T) = \mathcal{L}(\boldsymbol{\omega}) = \int_{u=-\infty}^{\infty} \int_{x=-\infty}^{\infty} L(\mathbf{x}) \exp(-2\pi j \boldsymbol{\omega}^T \mathbf{x})\,\mathrm{d}x\,\mathrm{d}u \tag{5}$$

where $\boldsymbol{\omega} = [\omega_x, \omega_u]^T$ is the vector representation of the frequency components. We use the calligraphic symbols to denote the signal spectrums. Note that Fourier analysis used here is not related to Fourier optics.

When the reference plane X is aligned with the surface of an object, the resulting light field, denoted by L_s and illustrated in Fig. 3, is called surface light field [28]. If the reflectance of the object is Lambertian, the radiances of the light rays emitted from the object are independent of the view angle. Therefore, the light field can be described by a 1-D *texture function* T

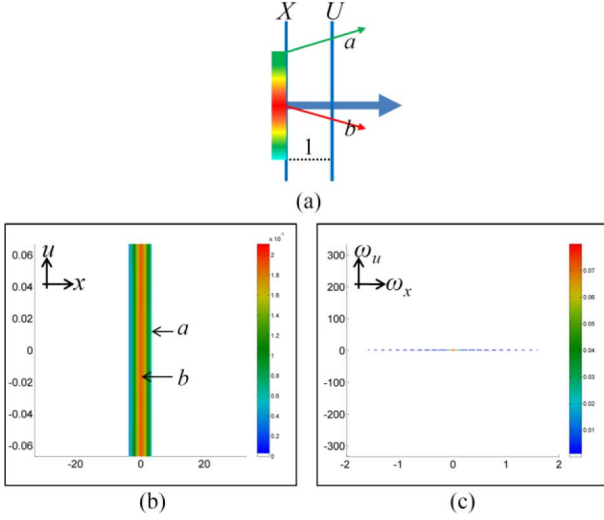$$L_s(x, u) \cong R(x, \theta) = T(x). \tag{6}$$

Fig. 3. (a) Construction of the surface light field of a Lambertian object. (b) Resulting light field. (c) Fourier spectrum of (b).

The derivation is described in Appendix B. Therefore, the surface light field is constant in the u direction. Moreover, because it is invariant along the u-axis, we have

$$\mathcal{L}_s(\boldsymbol{\omega}) = \mathcal{T}(\omega_x)\delta(\omega_u) \tag{7}$$

where δ is the Dirac delta function. That is, all energies of the spectrum fall on the plane $\omega_u = 0$ in the Fourier domain, as shown in Fig. 3(c).

If we move the location of the reference plane as the light rays traverse in space, the light field representation is changed even though the radiances of the light rays are not. In the following, we discuss the operations of light field.

B. Light Ray Propagation

When a light ray traverses in free space, the radiance along the light ray is unchanged. Consider the example shown in Fig. 4. If we displace the reference planes by d units in the center ray direction, the representation of the light ray (x_0, u_0) becomes (x_1, u_1), and $(x_0, u_0) = (x_1 - d \cdot u_1, u_1)$. We can relate the original light field L_0 defined by planes X_0 and U_0 to the new light field L_1 defined by planes X_1 and U_1 by

$$L_1([x_1, u_1]^T) = L_0([x_0, u_0]^T)$$
$$= L_0\left(\begin{bmatrix} 1 & -d \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ u_1 \end{bmatrix}\right) \triangleq L_0(\mathbf{T}_d\mathbf{x}). \tag{8}$$

That is, when the light rays traverse by d units, the new light field is a linear transformation of the original light field, where the transformation matrix $\mathbf{T}_d$ is a function of d.

Moreover, according to the Fourier linear transformation theorem [29], the spectrum of the light field $\mathcal{L}_1$ is also the result of a linear transformation of the original spectrum

$$\mathcal{L}_1(\boldsymbol{\omega}) = |\det(\mathbf{T}_d)|^{-1}\mathcal{L}_0(\mathbf{T}_d^{-T}\boldsymbol{\omega}) = \mathcal{L}_0\left(\begin{bmatrix} 1 & 0 \\ d & 1 \end{bmatrix}\boldsymbol{\omega}\right). \tag{9}$$

Consider again the example in Fig. 4, assuming the surface light field shown in Fig. 3(a) is the initial light field L_0. As the
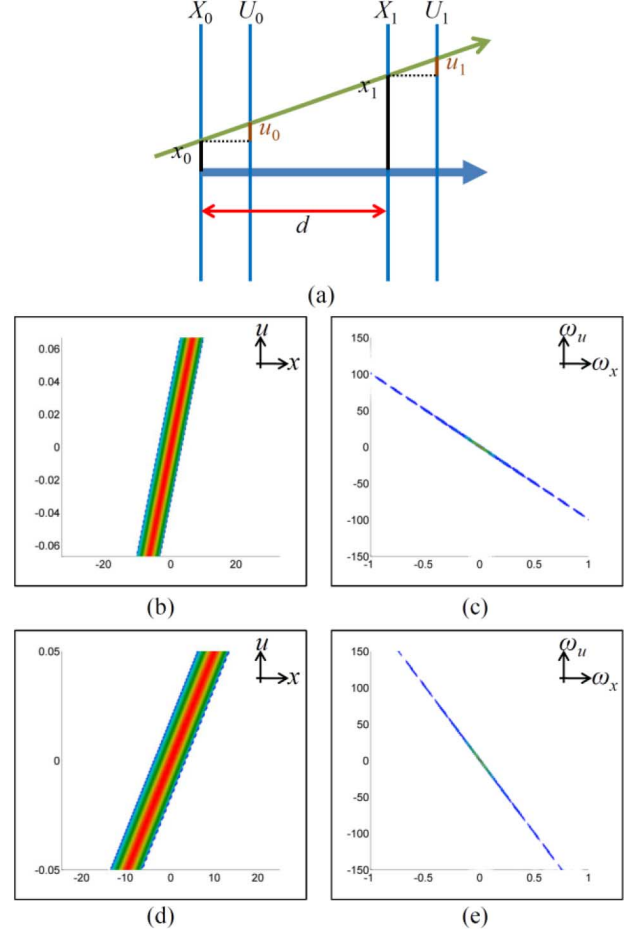

Fig. 4. Light ray propagation. (a) One light ray can be represented using two different sets of reference planes. The initial light field is the surface light field shown in Fig. 3, we can use (8) and (9) to obtain (b) the light field at $d = 100$ away from the surface, and (c) the spectrum of the transformed light field. (d) and (e) Are the light field and its spectrum at $d = 200$, respectively.

light rays travel along the central ray by, for example, 100 units, according to (8) the light field is sheared [Fig. 4(b)], so is the corresponding spectrum of the light field [Fig. 4(c)]. As we increase the distance between the surface and the reference plane, the degree of shearing increases. Although the initial light field contains only spatial variations, they become angular variations due to the linear transformation.

C. Lens Refraction

According to the Snell's law, a light ray changes its direction when passing from one medium to another that has a different refractive index. Unfortunately, this direction change cannot be described as a linear transformation in the light field representation. However, for the purpose of image formation modeling, we only have to consider lens refraction, not general refraction. Specifically, we consider the thick lens or thin lens models in Gaussian optics [5]. These models are commonly used in image processing and computer vision.

For a thin lens model, if we align the reference plane X with the thin lens, then the light fields L_0 before refraction is related to the light field L_1 after the refraction by

$$L_1(\mathbf{x}) = L_0\left(\begin{bmatrix} 1 & 0 \\ f^{-1} & 1 \end{bmatrix}\begin{bmatrix} x \\ u \end{bmatrix}\right) \triangleq L_0(\mathbf{R}_f\mathbf{x}) \tag{10}$$

where f is the focal length of the lens. We can see that the refraction of the lens can also be modeled as a linear transformation of the light field. Similar to the case described by (9), the spectrum of the refracted light field is related to that of the original light field by a linear transformation.

Note that matrices similar to $\mathbf{T}_d$ and $\mathbf{R}_f$ are used in matrix optics to describe the behavior of individual light rays [30]. Here we use them to transform the light field that represents the radiances of all light rays. An advantage of analysis in the light field space is that we can simultaneously describe both geometric and photometric changes of the light rays. On the contrary, traditional matrix optics can only describe the geometric transformations of the light rays. Another advantage of light filed analysis is that the techniques well developed in signal processing become immediately applicable.

D. Occlusion

When an occluder is present in the path of a light ray, the radiance along the light ray cannot be propagated through the occluder. This blocking effect can be described as a modulation of the light field

$$L_1(\mathbf{x}) = L_0(\mathbf{x})B(\mathbf{x}) \tag{11}$$

where $B(\,\cdot\,)$ is the blocking function, or the shield field [19], of the occluder. The value of $B(\mathbf{x})$ controls how the light ray (x, u) is blocked or attenuated by the occluder. Unlike the previous two operations that describe the geometric changes of the light rays, the blocking operation describes the photometric changes of the light rays in the light transport process.

According to the convolution theorem and the duality property, when a signal L_1 is a modulation of inputs L_0 and B, its spectrum is the convolution of the two input signals

$$\mathcal{L}_1(\boldsymbol{\omega}) = \mathcal{L}_0(\boldsymbol{\omega}) \otimes \mathcal{B}(\boldsymbol{\omega}). \tag{12}$$

Determining the blocking function of an arbitrary occluder is generally difficult. However, several occluders needed for our analysis can be derived. For example, when an occluder is infinitesimally thin and is aligned with the reference plane X, whether a light ray is blocked would solely depend upon the intersection point x with the occluder. In this case, the blocking function B and its spectrum are similar to the surface light field of a Lambertian object

$$B(\mathbf{x}) = O(x) \tag{13}$$
$$\mathcal{B}(\boldsymbol{\omega}) = O(\omega_x)\delta(\omega_u) \tag{14}$$

where O is a 1-D signal describing the transmittance of the occluder. Clearly, the energy of the blocking function concentrates on the plane $\omega_u = 0$ and, thus, (12) is effectively a convolution of a 2-D signal with a 1-D signal.

An example occlusion is illustrated in Fig. 5, where the initial light field is the one shown in Fig. 4(d) and the occluder is a thin periodic grating. Fig. 5(b) shows that the effect of occlusion is a modulation of the light field by a periodic function. Because the spectrum of the blocking function is a 1-D signal, convolving it with the initial light field results in several small replicas along
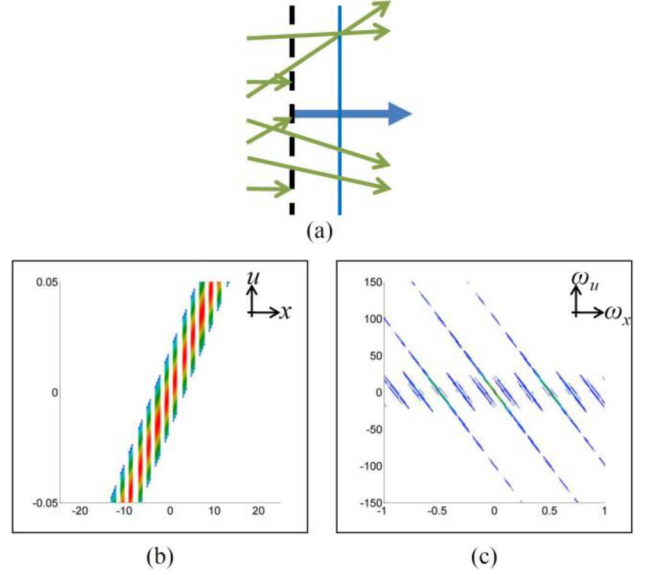

Fig. 5.   Effect of occlusion on the light field. (a) Thin periodic occluder is placed at the reference plane. Assume the initial light field is the propagated surface light field in Fig. 4(d), we can use (11) to obtain (b) the modulated light field and (12) to obtain the convolved spectrum.

the ω_x-axis. The power of each replica depends upon the power spectrum of the blocking function.

E. Reparameterization

For the light rays entering the camera body, we use the in-camera parameterization shown in Fig. 2(c) to represent the light field. By aligning the reference plane X in the original two-plane parameterization with the image plane, the light field $\hat{L}$ in the in-camera parameterization is represented as a linear transformation of the original light field L

$$\hat{L}(\mathbf{x}) = \frac{1}{F} L\left(\begin{bmatrix} 1 & 0 \\ F^{-1} & -F^{-1} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \right) \triangleq \frac{1}{F} L(\mathbf{P}_F \mathbf{x}) \tag{15}$$

where the scaling factor $1/F$ is the Jacobian of the transformation. Also, according to the Fourier linear transform theorem, the spectrum of $\hat{L}$ is a linear transformation of $\mathcal{L}$. In the following, we use the symbol ˆ to denote signals represented by the in-camera parameterization.

F. Integration of Radiances Into an Irradiance

For a sensor with a uniform angular sensitivity, the sensed irradiance value is the integration of the incoming radiances times the sensor sensitivity, ρ

$$I(x) = \int\limits_{\theta=-\pi/2}^{\pi/2} \rho R(x, \theta) \cos\theta\mathrm{d}\theta = \int\limits_{u=-\infty}^{\infty} \rho L(\mathbf{x})\mathrm{d}u \tag{16}$$

where the second equality is due to different light field representations. We can see that the 1-D irradiance is obtained by integrating the light field along the u-axis. Note that this formulation is similar to the reflected reflectance from the Lambertian object [18].

For an N-dimensional signal, integrating along $N–k$ dimensions reduces the signal to k-dimensional. This process is called *canonical projection operation*. According to the Fourier slice theorem [26], projecting the signal in the primal domain is equivalent to extracting a spectrum of a k-dimensional hyperplane in the Fourier domain, and this process is called a *slicing operation*.

For the completeness, we derive the theorem for $(N, k) = (2, 1)$, which is the case for (16). The spectrum $\mathcal{I}(\omega_x)$ is the Fourier transformation of the image $I(x)$

$$
\begin{aligned}
\mathcal{I}(\omega_x) &= \int_{x=-\infty}^{\infty} I(x) \exp(-2j\omega_x x)\mathrm{d}x \\
&= \int_{x=-\infty}^{\infty} \left(\int_{u=-\infty}^{-\infty} \rho L(\mathbf{x})\mathrm{d}u \right) \exp(-2j\omega_x x)\mathrm{d}x \\
&= \int_{x=-\infty}^{\infty} \int_{u=-\infty}^{-\infty} \rho L(\mathbf{x}) \exp(-2j(\omega_x x + \omega_u \cdot 0))\mathrm{d}u\mathrm{d}x \\
&= \rho\mathcal{L}([\omega_x, 0]^T) = [\rho\mathcal{L}(\boldsymbol{\omega})]_{\omega_u=0}.
\end{aligned}
\tag{17}
$$

In other words, $\mathcal{I}$ is equivalent to the subset $\mathcal{L}$ along the slice $\omega_u = 0$. The slicing operation plays an important role in the image formation as we show in Section IV.

G. Combined and Backward Operations

We have shown that several processes in light transport can be described as transformations or modulations of the light field. Because these operations are linear, they can be combined together. For example, given the initial light field L_0, if the light rays propagate for a units, we can obtain the new light field L_1 using (8). If the light rays propagate for b more units, the light field L_2 can be obtained in a similar way. Because matrix multiplication is associative, we have

$$
L_2(\mathbf{x}) = L_1(\mathbf{T}_b\mathbf{x}) = L_0(\mathbf{T}_a(\mathbf{T}_b\mathbf{x})) = L_0((\mathbf{T}_a\mathbf{T}_b)\mathbf{x}).
\tag{18}
$$

The combined matrix $\mathbf{T}_a\mathbf{T}_b$ describes the transformation of the light field when the light rays are propagated by $(a + b)$ units from the original reference plane. We can also obtain L_2 from L_0 directly by

$$
L_2(\mathbf{x}) = L_0(\mathbf{T}_{a+b}\mathbf{x})
\tag{19}
$$

and with simple calculation we have $\mathbf{T}_{a+b} = \mathbf{T}_a\mathbf{T}_b$. That is, the combined operation is equivalent to the direct operation.

Therefore, as the light rays are refracted or blocked during traversal, we can simply concatenate all individual operations to represent the overall light transport process without deriving it—a major benefit of the light transport analysis.

While the operations of the light transport process are discussed in a forward fashion, many of them can be described in a backward manner. For example, if a light field L_0 is linearly transformed to L_1 by $\mathbf{A}$ due to either propagation or refraction;
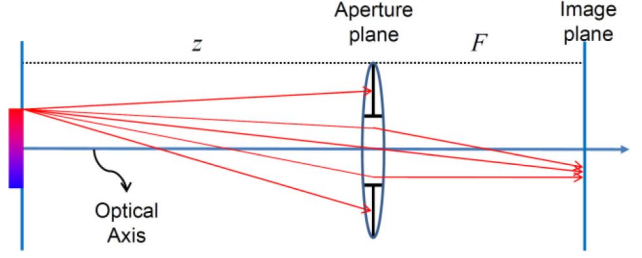

Fig. 6. Camera configuration. Traditionally, each light ray is traced independently to construct the image formation models.

we have $L_1(\mathbf{x}) = L_0(\mathbf{A}\mathbf{x})$. Then it is straightforward to show that

$$
L_0(\mathbf{x}) = L_1(\mathbf{A}^{-1}\mathbf{x})
\tag{20}
$$

$$
\mathcal{L}_0(\mathbf{x}) = |\det(\mathbf{A}^{-1})|^{-1}\mathcal{L}_1(\mathbf{A}^T\mathbf{x}).
\tag{21}
$$

Therefore, given the forward operation and a specific coordinate system in the transformed light field, we can easily obtain its coordinates in the original light field using the backward operation. Similarly, the modulation operation can be reversed by multiplying the light field by the inverse of the modulation function.

However, the projection operation, or the equivalent slicing operation in the Fourier domain, irreversibly reduces the dimensionality of the signal. Therefore, we cannot recover the light field from an irradiance signal without relying on proper prior information of the light field [36]. A special case of layered scenes is discussed in Section VI-C.

IV. PHOTOGRAPHIC OPERATION FOR IMAGE FORMATION

We have described the basic operations of light field. In the following, we use these operations to model the image formation process. Specifically, we show how the light field emitting from the object surface propagates in space before it enters the camera, and how the image is generated from the light field in the camera.

Without loss of generality, we assume the reference plane X of the initial light field L_i is z units away from the thin lens of the camera, the focal length of the lens is f, and the distance between the image plane and the aperture plane is F. The reference plane can be aligned with the object surface as shown in Fig. 6, but it is not necessary.

In traditional image formation models, one has to apply a ray-tracing to determine the end points of all light rays emitting from the object. Only those light rays that hit the image plane contribute to the final image. However, ray-tracing is difficult to express elegantly. On the contrary, we use the light transport operations to describe the transformation of all radiances together. The light rays first propagate z units to the lens and, thus, the light field is transformed by $\mathbf{T}_z$ according to (8). Then the light rays not blocked by the aperture are refracted by the lens. The blocking process causes a modulation of the light field, and then the light field is transformed by $\mathbf{R}_f$ according to (10). The refracted light rays then propagate F units to the image plane; the light field is transformed by $\mathbf{T}_F$. Finally, we change the parameterization of the light field to the in-camera parameterization
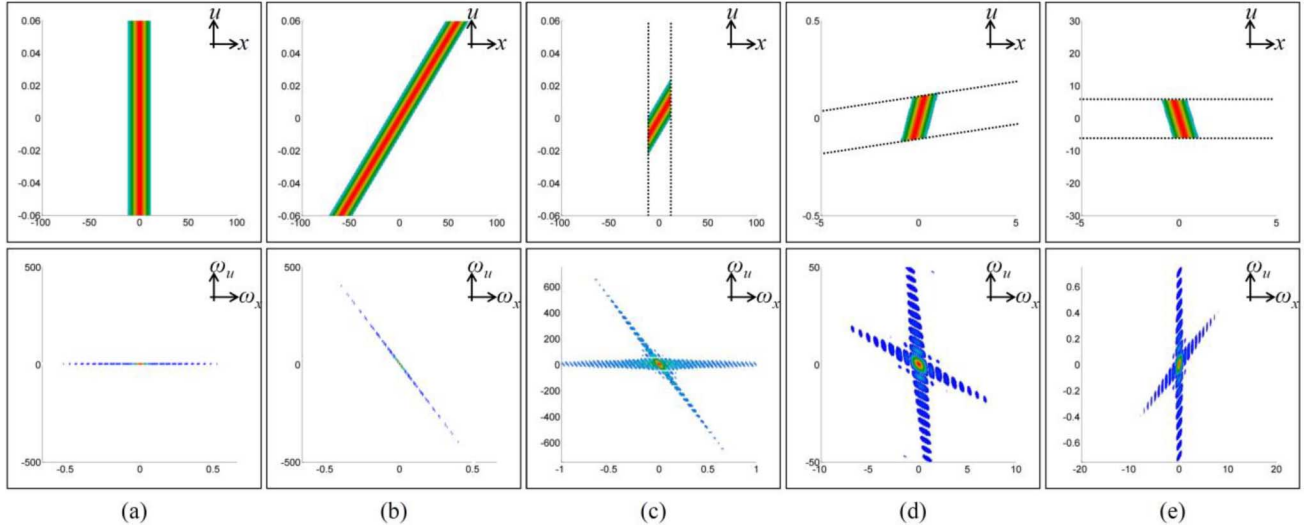
Fig. 7. Transformation of the light field and its spectrum in image formation. The first row shows the light fields at different steps in image formation, and the second row shows the corresponding spectrums. The scene and camera parameters in this case is $z = 5010$, $f = 50$, $A = 12$, and $F = (1/50 - 1/5000)^{-1}$ (i.e., the object at $z = 5000$ is in-focus). (a) The surface light field. (b) The light field observed at the lens plane that is z units away from the surface. (c) The light field modulated by the blocking function of the aperture. The dotted lines define the width of the aperture. (d) The light field observed at the image plane. (e) The light field represented by the in-camera parameterization.
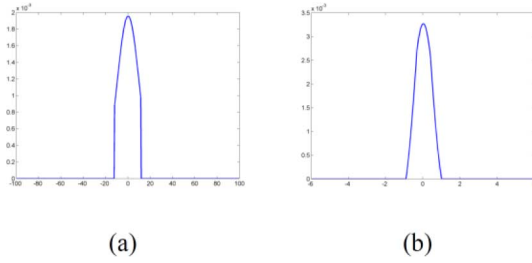

Fig. 8. (a) Texture function of the surface in the Fig. 7. (b) Resulting defocused image using the scene and camera parameters given in Fig. 7.

using (15) and have the final light field $\hat{L}_r$. One example of this process is shown in Figs. 7 and 8, which are used to assist the derivations shown in the following and discussed in the end of the section.

A. Derivation Without Aperture

We can see that $\hat{L}_r$ is related to L_i by a series of linear transformations and modulation. For now, we neglect the modulation due to the aperture and let L_p denote the *perfectly* transformed light field at the image plane before changing the parameterization. Because we can combine several operations together, we have

$$L_p(\mathbf{x}) = L_i(\mathbf{C}_{zfF}\mathbf{x}) \triangleq L_i(\mathbf{T}_z\mathbf{R}_f\mathbf{T}_F\mathbf{x})$$
$$= L_i\left(\begin{bmatrix} 1 - z/f & -Fz(1/z + 1/F - 1/f) \\ 1/f & 1 - F/f \end{bmatrix}\begin{bmatrix} x \\ u \end{bmatrix}\right). \tag{22}$$

This transformation is called *photographic operation* and $\mathbf{C}_{zfF}$ is dubbed *photographic matrix*. We let Δ denote $(1/z + 1/F -$

$1/f)$ for convenience. According to the Fourier linear transform theorem, the spectrum of L_p is

$$\mathcal{L}_p(\boldsymbol{\omega}) = \mathcal{L}_i\left(\mathbf{C}_{zfF}^{-T}\boldsymbol{\omega}\right)$$
$$= \mathcal{L}_i\left(\begin{bmatrix} 1 - F/f & -1/f \\ Fz\Delta & 1 - z/f \end{bmatrix}\begin{bmatrix} \omega_x \\ \omega_u \end{bmatrix}\right). \tag{23}$$

The sensors at the image plane integrate the radiances into the irradiances. According to (17) and remove ρ for simplicity, the spectrum of the captured image is

$$\mathcal{I}_p(\omega_x) = [\mathcal{L}_p(\boldsymbol{\omega})]_{\omega_u = 0} = \mathcal{L}_p([\omega_x, 0]^T). \tag{24}$$

Therefore, if the size of the aperture is infinite, the spectrum of the image is equivalent to a slice of $\mathcal{L}_p$ along $\omega_u = 0$.

B. Derivation With Aperture

When the size of the aperture is finite, a number of light rays are blocked from entering the camera. As described in Section III-D, this causes a modulation of the light field. If the aperture is infinitesimally thin and its size is A, when we align the reference plane with the aperture, the blocking function B_A is a 1-D rectangular function

$$B_A(\mathbf{x}) = \text{rect}(x/A) = \begin{cases} 1, & |x| < A/2 \\ 0, & \text{otherwise} \end{cases} \tag{25}$$

and according to (14) its spectrum is

$$\mathcal{B}_A(\boldsymbol{\omega}) = A\sin c(A\omega_x)\delta(\omega_u)$$
$$= (\pi\omega_x)^{-1}\sin(A\pi\omega_x)\delta(\omega_u). \tag{26}$$

After the light field is modulated with the blocking function, the light rays propagate F units to the image plane. Due to the

linearity of the operations and the distributive rule, we can transform both the input light field and the blocking function and then multiply them together. Therefore, when the aperture is present, the light field L_r at the image plane and its spectrum are

$$
\begin{aligned}
L_r(\mathbf{x}) &= L_i(\mathbf{T}_d \mathrm{R}_f \mathrm{T}_F \mathbf{x}) B_A(\mathrm{T}_F \mathbf{x}) \\
&= L_i(\mathbf{C}_{zfF} \mathbf{x}) B_A(\mathbf{T}_F \mathbf{x})
\end{aligned} \tag{27}
$$

$$
\mathcal{L}_r(\boldsymbol{\omega}) = \mathcal{L}_i\left(\mathbf{C}_{zfF}^{-T}\boldsymbol{\omega}\right) \otimes B_A\left(\mathrm{T}_F^{-T}\boldsymbol{\omega}\right). \tag{28}
$$

We can see that the blocking function is transformed due to the propagation operation. This causes a skew to the axis-aligned rectangular function and complicates our analysis [see Fig. 7(d)]. To address this issue, we change the parameterization to the in-camera parameterization shown in Fig. 2(c). Then the blocking function due to the aperture becomes

$$
\hat{B}_A(\mathbf{x}) = B_A(\mathbf{T}_F \mathbf{P}_F \mathbf{x}) = \mathrm{rect}(u/A) \tag{29}
$$

and its energies fall on the plane $\omega_x = 0$. Using (15), (27), and (29), we have the final light field $\hat{L}_r$

$$
\begin{aligned}
\hat{L}_r(\mathbf{x}) &= \hat{L}_p(\mathbf{x})\hat{B}_A(\mathbf{x}) = L_p(\mathbf{P}_F \mathbf{x})\hat{B}_A(\mathbf{x}) \\
&= L_i(\mathbf{C}_{zfF}\mathbf{P}_F\mathbf{x})\hat{B}_A(\mathbf{x}) \triangleq L_i(\hat{\mathbf{C}}_{zfF}\mathbf{x})\hat{B}_A(\mathbf{x}) \\
&= L_i\left(\begin{bmatrix} -z/F & z\Delta \\ 1/F & 1/f - 1/F \end{bmatrix}\begin{bmatrix} x \\ u \end{bmatrix}\right)\hat{B}_A(\mathbf{x}) \quad (30)
\end{aligned}
$$

and its spectrum

$$
\begin{aligned}
\hat{\mathcal{L}}_r(\boldsymbol{\omega}) &= \hat{\mathcal{L}}_p(\boldsymbol{\omega}) \otimes \hat{\mathcal{B}}_A(\boldsymbol{\omega}) = \mathcal{L}_i(\mathbf{C}_{zfF}^{-T}\omega) \otimes \hat{\mathcal{B}}_A(\boldsymbol{\omega}) \\
&= \mathcal{L}_i\left(\begin{bmatrix} 1 - F/f & 1 \\ Fz\Delta & z \end{bmatrix}\begin{bmatrix} \omega_x \\ \omega_u \end{bmatrix}\right) \otimes \hat{\mathcal{B}}_A(\boldsymbol{\omega}). \quad (31)
\end{aligned}
$$

Here we remove the constant Jacobian $1/F$ in (15), which has no effect on our analysis. Note that because the spectrum of the blocking function falls on the plane $\omega_x = 0$, the transformed light field is only convolved along the ω_u-axis.

According to (16), the captured image I_r, is

$$
I_r(x) = \int_{u=-\infty}^{\infty} \hat{L}_r(\mathbf{x})\mathrm{d}u \tag{32}
$$

and according to (24) and (26), the spectrum of the image is

$$
\begin{aligned}
\mathcal{I}_r(\omega_x) &= [\hat{\mathcal{L}}_r(\boldsymbol{\omega})]_{\omega_u=0} = \left[\mathcal{L}_i\left(\hat{\mathbf{C}}_{zfF}^{-T}\boldsymbol{\omega}\right) \otimes \hat{\mathcal{B}}_A(\boldsymbol{\omega})\right]_{\omega_u=0} \\
&= \left[\mathcal{L}_i\left(\hat{\mathbf{C}}_{zfF}^{-T}\boldsymbol{\omega}\right) \otimes (A\sin\mathrm{c}(A\omega_u)\delta(\omega_x))\right]_{\omega_u=0}.
\end{aligned} \tag{33}
$$

In summary, given the initial light field L_i representing the radiance distribution of the scene, we can obtain the light field in the camera by linearly transforming the initial light field using the photographic matrix, and modulating it with the blocking function. The modulation causes a convolution of the light field spectrum. The spectrum of the image is the result of the slicing operation on the transformed light field.

One example of the full light transport process of image formation is shown in Fig. 7. Starting from the object surface, the light rays first propagate z units to reach the lens plane, and the light field is transformed according to (8), as shown in Fig. 7(b). The light field is then modulated by the aperture function, and the spectrum is spread out due to the convolution operation, as shown in Fig. 7 (c). Finally, the light rays propagate to the image plane, and by changing the parameterization, the final light field is obtained [Fig. 7(e)]. The texture function of the surface in this example is a windowed Gaussian function as shown in Fig. 8(a). Because the surface is not in-focus, the steep boundaries of the surface are blurred [Fig. 8(b)].

We can see that in the light transport process, the light field is never filtered. This should be clear because the light rays do not interact with other, and neither do the radiances along the light rays. This is different from the traditional methods that perform analysis on the image domain, where filtering is a common operation.

Equations (30)–(33) describe the exact physical process in image formation. They present both the changes of the geometric and photometric properties of the light rays during light transport. Moreover, we can map those equations to several traditional models of photographic effects.

V. RELATION TO TRADITIONAL MODELS

In this section, we show that the equations presented in the previous section can be used to describe geometric and photometric effects in image formation. We also show that those effects can be analyzed in the frequency domain using spectral analysis.

A. Perspective Transformation

It is well known that in the pinhole model, a scene point (x, z) is projected to $-xF/z$ on the image plane. The projection is described as a perspective transformation of 3-D primitives using the homogenous coordinate system. Here we show that the perspective projection is essentially a light field transformation due to light transport.

We first discuss the case of pinhole camera, of which the aperture size is infinitesimally small. That is, in $A \to 0$ (29) and the blocking function becomes a delta function of u. Substituting it into (30) and (32), we have

$$
\begin{aligned}
I_r(x) &= \int_{u=-\infty}^{\infty} \hat{L}_r(\mathbf{x})\mathrm{d}u \\
&= \int_{u=-\infty}^{\infty} L_i(\hat{\mathbf{C}}_{zfF}\mathbf{x})\delta(u)\mathrm{d}u = L_i(\hat{\mathbf{C}}_{zfF}[x,0]^T) \quad (34)
\end{aligned}
$$

That is, the irradiance of x on the image plane is a radiance of the initial light field L_i. When the reference plane of L_i is aligned with the Lambertian surface, L_i can be described by a 1-D texture function as in (6)

$$
\begin{aligned}
I_r(x) &= L_i(\hat{\mathbf{C}}_{zfF}[x\ 0]^T) \\
&= T(-zx/F) \Rightarrow T(x) = I_r(-Fx/z). \tag{35}
\end{aligned}
$$

That is, when the object is at z units away from the lens, the radiance at x on the surface is mapped to the irradiance at $-xF/z$ on the image plane. This is exactly the equation of the perspective transformation. In other words, the pinhole camera captures a specific subset of light rays, or a slice of the light field. This slice is stretched on the image plane, and the stretch factor is a function of z and F.

Now consider the case of a lens camera with a finite aperture. Before we explain the defocus effect in the next subsection, we assume the object is in-focus. According to the lens formula, when the object at distance z is in-focus, $\Delta = 1/z + 1/F - 1/f = 0$, and the photographic matrix becomes

$$\hat{\mathbf{C}}_{zfF} = \begin{bmatrix} -z/F & 0 \\ 1/F & 1/f - 1/F \end{bmatrix} \tag{36}$$

then substituting it into (30) we have

$$\hat{L}_r(\mathbf{x}) = L_i([-zx/F, x/F + (1/f - 1/F)u]^T)\hat{B}_A(\mathbf{x}). \tag{37}$$

Again, when L_i is emitted from the Lambertian surface at z, all light rays reaching x on the image plane are from the scene point $(-zx/F, z)$, and only differ in the emitting angles. As a result, for a lens camera, the perspective transformation is still valid when the object is in-focus.

Because the light field is a signal, we also conduct our analysis in the frequency domain. Combining $\Delta = 0$, (31), and (33) we have

$$\mathcal{I}_r(\omega_x) = \mathcal{L}_i\left(\begin{bmatrix} 1 - F/f & 1 \\ 0 & z \end{bmatrix}\begin{bmatrix} \omega_x \\ 0 \end{bmatrix}\right) \tag{38}$$

where $1 - F/f = F(-1/z) = -F/z$ (ignoring the effect of the aperture for now). Applying variable substitution to (7) and (38), we have

$$I_r(-z\omega_x/F) = \mathcal{T}(\omega_x). \tag{39}$$

We can see that the image spectrum is a stretched version of the texture spectrum. Again the stretch factor only depends upon z and F. As the object moves away from the camera (z increases), the frequencies of its texture appear higher in the image. This result exactly matches that of the perspective transformation.

B. Defocus Blur

Defocused images are often modeled as the result of applying a pillbox or Gaussian filter to the in-focus images, for which the point spread function of the filter depends upon the object distance and the camera parameters. Here we show that such image filtering actually corresponds to the combination of the photographic operation and the projection/slicing operation in the light transport framework. We also show that even when the object is out of focus, the perspective transformation is still valid.

We use the backward transport operation to facilitate the analysis and neglect the effect of the aperture for now. According to (21) and (31), when $\Delta \neq 0$ we have

$$\mathcal{L}_i(\boldsymbol{\omega}) = \hat{\mathcal{L}}_p\left(\begin{bmatrix} -z/F & 1/F \\ z\Delta & 1/f - 1/F \end{bmatrix}\begin{bmatrix} \omega_x \\ \omega_u \end{bmatrix}\right). \tag{40}$$
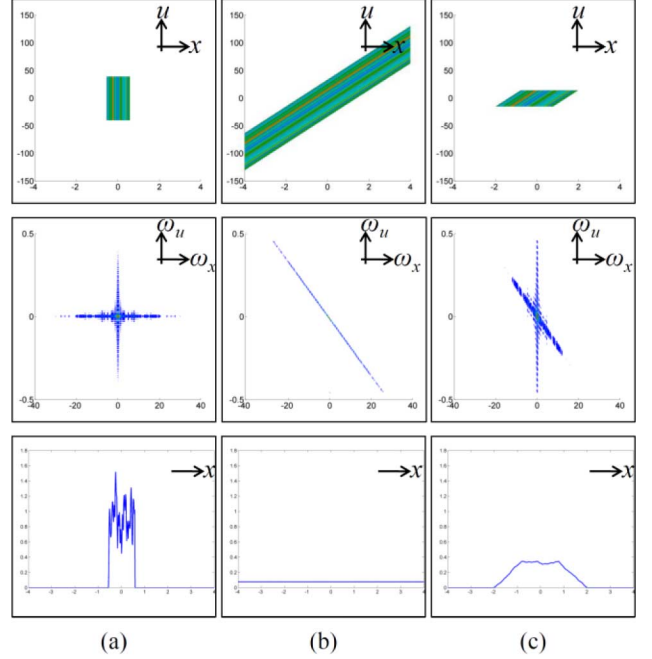

Fig. 9. Effect of the aperture size on the defocus blur. The first row shows the final light field, the second row shows the corresponding spectrums, and the third row shows the final images. (a) Object is in focus, that is, $\Delta = 0$. (b) Object is out of focus and aperture size is infinite. (c) Object is out of focus and the aperture size is finite.

Assuming that there is only one object at distance z and that the reference plane of L_i is aligned with the Lambertian surface, we immediately have that the energies concentrate on the plane $\omega_u = 0$. Then (40) becomes

$$\mathcal{L}_i([\omega_x, 0]^T) = \hat{\mathcal{L}}_p([-z\omega_x/F, z\Delta\omega_x]^T). \tag{41}$$

We can see that before the convolution with the blocking function, the energies fall on the plane $\omega_u = -F\Delta\omega_x$. Because all other frequency components are null and the blocking function only spreads the energies along the ω_u-axis, according to (33) we have

$$\begin{aligned}
\mathcal{I}(\omega_x) &= \hat{\mathcal{L}}_p([\omega_x, -F\Delta\omega_x]^T)A\sin\mathrm{c}(-AF\Delta\omega_x) \\
&= A\hat{\mathcal{L}}_i([-F\omega_x/z, 0]^T)\sin\mathrm{c}(-AF\Delta\omega_x) \\
&= A\mathcal{T}(-F\omega_x/z)\sin\mathrm{c}(-AF\Delta\omega_x).
\end{aligned} \tag{42}$$

The second and third equalities are obtained by simple variable substitution. This equation indicates that the spectrum of the image is a modulated and stretched texture function, where the modulation is dependent upon the camera parameters. According to the convolution theorem, the image is a filtered version of the stretched texture function.

An example of the defocus blur with and without aperture is shown in Fig. 9. When the object is in focus, $\Delta = 0$. According to (42), all frequency components in this case are captured without decay. Therefore, all details of the object are preserved in the image. When the object is out of focus and the aperture size is infinite, $(A \to \infty)$, $\mathcal{I}(\omega_x) = 0$ unless $\omega_x = 0$. Therefore, the image is a flat signal. Finally, when the aperture size is finite, a blurred image is captured.

It should be noted that the mapping between the texture function and the image is independent of Δ. When $\Delta = 0$, (42) is equivalent to (39) up to a scale factor. This means when the object is out of focus, the image spectrum is not shifted but only attenuated (and the attenuation is characterized by a sinc function). Therefore, the perspective transformation is still valid even when the object is out of focus. This claim is difficult to obtain from the traditional ray-based or image-based analysis.

To relate (42) to the traditional defocus model, we think of it as a linear system, where the input is the stretched texture function and the output is the image. If the texture function T is a delta function (e.g., a white point on a black surface), L_i is uniform and the image is the response function H of the system

$$\mathcal{H}(\omega_x) = A \sin c(-AF\Delta\omega_x) \tag{43}$$

$$H(x) = \frac{1}{F|1/z + 1/F - 1/f|}$$
$$\times \operatorname{rect}\left(\frac{x}{AF|1/z + 1/F - 1/f|}\right) \tag{44}$$

which is identical to the pillbox filter in the traditional model [3]. The appearance of an out-of-focus scene point is a box of width $AF|\Delta|$ with uniform irradiance distribution.

When the aperture is a pinhole $A \to 0$, the spectrum of the blocking function is uniform along ω_u. Therefore, the energy at any frequency is mapped to the slice $\mathcal{I}$ without any decay. In this case, all objects are in focus. This means that the pinhole camera is only a special case of the lens model. On the other hand, when the aperture size is infinite, the out-of-focus objects are absent from the image [see Fig. 9(b)]. Therefore, when the aperture is very large, one may see the in-focus object behind the out-of-focus one which only contributes a constant offset to the image signal, as demonstrated in [12].

Finally, one can also use our framework to derive the point spread function of defocus blur in the primal domain. Let the scene contain a single Lambertian point at $(x = 0, z)$ and the initial light field be $L_i(\mathbf{x}) = \delta(x)$. Then we obtain the transformed light field from (30) and the image from (32). It is straightforward to verify that the image of the scene point is a flat box of width $AF\Delta$, identical to the result described by (44).

C. Vignetting

The perspective transformation and the defocus blur are due to the geometric transformation of the light field. Here we turn our attention to a different kind effect on the image caused by photometric changes of the light field.

The reduction of brightness at the periphery of the image, which is a typical phenomenon encountered in practice, is in fact a combination of several effects collectively called the *vignetting* effect. The falloff due to the natural vignetting can be modeled by a cosine fourth law [1]. When the object is Lambertian with a uniform texture function $(T(x) = C)$, the image $I(x)$ is proportional to $\cos^4(\theta)$, where θ is the angle between the optical axis and the light ray pointing from the pinhole center to the image point x. In 2-D space, we have a slightly different law: The image is proportional to $\cos^3(\theta)$, as derived in Appendix C. In the following, we show that our light transport framework can also account for this effect.

First, note that we neglect the effect of the parameterization function g in the previously shown analyses. If we use (55) as the initial light field and assume that $T(x)$ is a constant C and that the object is in focus, (30) becomes

$$\hat{L}_r(\mathbf{x}) = L_i(\hat{\mathbf{C}}_{zfF}\mathbf{x})\hat{B}_A(\mathbf{x})$$
$$= \frac{C}{F[1 + (x/F + (1/f - 1/F)u)^2]^{1.5}}\operatorname{rect}(u/A). \tag{45}$$

Then, according to (16), the image $V(x)$ of the uniform Lambertian surface, which is called vignetting field, becomes

$$V(x) = \int_{u=-\infty}^{\infty} \frac{C}{F[1 + (x/F + (1/f - 1/F)u)^2]^{1.5}}$$
$$\times \hat{B}_A(\mathbf{x})\mathrm{d}u$$
$$= \int_{u=-0.5A}^{0.5A} CF^{-1}$$
$$\times [1 + (x/F + (1/f - 1/F)u)^2]^{-1.5}\mathrm{d}u. \tag{46}$$

Let $a = x/F$ and $b = 0.5A(1/f - 1/F)$, we have

$$V(x) = \frac{C}{F/f - 1}[\sin(\tan^{-1}(a + b)) - \sin(\tan^{-1}(a - b))]. \tag{47}$$

This is the vignetting field obtained from our framework. Unlike the traditional model in which the vignetting field is only a function of θ, we show that the aperture size also influences the degree of vignetting. In fact, when the aperture size is small, we obtain the traditional model

$$I(x) \cong \mathrm{ACF}^{-1}[1 + (x/F + (1/f - 1/F)0)^2]^{-1.5}$$
$$= C\frac{A}{F}\cos^3\theta. \tag{48}$$

It is clear by now that the traditional cosine falloff model is only valid when the aperture is small.

Comparing (47) with (48), we can see the main difference between the traditional model and ours is that, while the size of the aperture only causes a uniform scaling to the image in the traditional model, it causes a spatially variant effect in our model. This means, when the size of the aperture changes, the degree of the vignetting changes as well. When the aperture is very small, the vignetting field follows the simple cosine-falloff model. However, as the aperture size increases, because the angular ranges of the incident light rays are different for different sensors on the image plane, the vignetting field diverges from the cosine-falloff model.

We can also model the vignetting effect in the frequency domain. When $T(x) = C$, the initial light field spectrum is a delta function. That is, all the energies fall on the origin ($\omega_x = \omega_u = 0$). According to (33), the image spectrum only has a DC component and, hence, the image should have a constant brightness. However, if we consider the parameterization function g, the spectrum of the light field would be convoluted with a 1-D filter

in the direction $\omega_x + \omega_u = 0$.[2] Therefore, the energy of the DC component is spread out. Combining this effect and (33), the image has a few low-frequency components and matches (47).

D. Summary

We have discussed several photographic effects and shown that these effects are the result of transformation, modulation, and projection of the light field. These effects can be modeled all together by our framework.

It is straightforward to model other effects with our framework. For example, in the telecentric lens system, the aperture is placed in front of the lens so that the perspective transformation becomes invariant to the focus setting [32]. Traditionally, modeling this effect requires tedious ray-tracing. But in our framework, the new aperture is nothing but a blocking function undergoing a propagation operation. Therefore, we can easily obtain the model for the telecentric lens by slightly modifying (29). The derivation is detailed in Appendix D.

VI. DISCUSSION

In this section, we discuss how to generalize the framework to handle complex scenes, how to extend it to 3-D space, and what the possible applications and limitations are.

A. Generalization

Traditional image formation models are usually based upon the assumption that the scene is Lambertian or even just a single plane parallel to the image plane. We adopt the same assumption when we show that our framework is quantitatively equivalent to those models.

However, this assumption is not essential for our framework. We derive the relationship between the initial light field L_i and the final light field $\hat{L}_r$, but the structure or the content of L_i is not specified. When the scene is a planar Lambertian surface aligned with the image plane, L_i is defined by a 1-D texture function (6).

When there are multiple planes, we can obtain L_i by simple composition. Starting from the surface light field of the plane farthest to the camera, we propagate the light field to the next surface, modulate it with the blocking function caused by this surface, combine the modulated light field with the new surface light field, and continue to the next surface. This regular procedure does not require complex visibility testing for tracing the propagation of each individual light ray.

When the surface is not Lambertian, the surface light field has angular variation. Therefore its spectrum no longer falls on a 1-D plane in the frequency domain. However, our framework can still accurately model its propagation to the image plane. The appearance of the non-Lambertian objects in the image with defocus blur or vignetting effect can, thus, be obtained.

Finally, we can combine other light transport processes into our framework. Starting from the light source, the light field would be transformed due to propagations, convolved due to surface reflections, and modulated due to occlusions. We can combine all operations into a single one which describes the complete light transport process, from the light emitter to the image sensor.

B. Extensions to 3-D Space

While the derivation of the framework is performed in 2-D space, it is straightforward to extend it to 3-D space. In 3-D space, the reference planes become 2-D and, thus, each intersection point requires two coordinates. For example, for the two-plane parameterization [Fig. 2(b)], we can use (x, u, y, v) to represent a light ray, where (x, y) denotes the intersection point on the plane X and (u, v) denotes the intersection point on the plane U. Although the dimensionality of the light field increases, the structures of the light transport operations remain unchanged. We can easily combine 4-D propagation operation, lens refraction operation, and reparameterization to obtain the 4-D photographic operation.

However, there are two noticeable changes in the derivations. The first one happens to the parameterization function g. In 3-D space, it becomes

$$g(u, v) = (1 + u^2 + v^2)^{-2}. \tag{49}$$

Similar to the 2-D case, the parameterization function is a smooth function with a small bandwidth. Multiplying g with the surface light field causes a small blur to the spectrum. This is negligible when the vignetting effect is not considered.

The second change happens to the blocking function. The aperture is a plane with a circular hole, and therefore the blocking function is

$$\hat{B}_A(u, v) = \begin{cases} 1, & (u^2 + v^2)^{0.5} \leq A/2 \\ 0, & \text{otherwise} \end{cases} \tag{50}$$

and its spectrum is a first-order Bessel function [3], [5]. Besides these two changes, the derivation of the framework and the mapping to the traditional models can be obtained by simply following the procedure we have developed for the 2-D cases.

Due to the increase of dimensionality, the number of parameters in the light transport operations becomes larger, and new photographic effects may occur. For example, the lens refraction still results in a linear transformation of the light field

$$L_1(\mathbf{x}) = L_0(\mathbf{R}_{a,b}\mathbf{x}) = L_0\left(\begin{bmatrix} \mathbf{R}_a & 0 \\ 0 & \mathbf{R}_b \end{bmatrix} \mathbf{x}\right) \tag{51}$$

where $\mathbf{x} = [x, u, y, v]^T$, $\mathbf{R}_a$ and $\mathbf{R}_b$ are defined in (10), and a and b are the lens focal length measured in the u and v directions, respectively. When $a \neq b$, there are two in-focus planes in different directions, resulting in the so-called astigmatic effect [5]. We can easily model this effect by just adding one parameter to the light transport operation.

C. Applications

In this paper we focus on interpreting various effects in image formation, but the main result, the relationship between the light field and the image, can be used in many applications. One obvious application is light field acquisition, specifically, recovering the modulated light field $\hat{L}_r$ from I_r. Because the slicing

---

[2]The parameterization function is $g(u)$ in the two-plane parameterization. When changed to the in-camera parameterization, it becomes $g((x - u)/F)$, which is constant along the direction $x - u = 0$. Therefore, the energies fall on the plane $\omega_x + \omega_u = 0$.

operation performed by the sensors always extracts a 1-D slice of the 2-D spectrum as we have shown in Section IV, one must find a way to put all 2-D frequency components on the slice. This is the motivation of several light field cameras [19], [21], [22], [24].

Another important result of our framework is that, for Lambertian objects at different z's, their energies would fall on slices with different slopes in the spectrum of $\hat{L}_r$. Therefore, when we have a complete light field, we can synthetically refocus the image by changing the slope of the slicing operation. This is the basic concept of the Fourier slice photography [26], [34]. Moreover, by analyzing the power spectrum of the light field, we can detect the slices that possibly contain in-focus objects without estimating the per-pixel depth information [33].

The light field in the camera is modulated by the blocking function, and the spectrum is convolved. It may be interesting to remove this effect by deconvolution of the light field spectrum

$$\hat{\mathcal{L}}_p(\boldsymbol{\omega}) = \hat{\mathcal{L}}_r(\boldsymbol{\omega}) \otimes^{-1} \hat{\mathbf{B}}_A(\boldsymbol{\omega}). \tag{52}$$

However, because $\mathbf{B}_A$ is a low-pass filter, the deconvolution is an ill-posed problem. In [33] and [34], it is shown that when the number of the objects with different depths is much smaller than the spectrum resolution and the depths are known, the un-modulated spectrum is sparse and, thus, perfect deconvolution is possible. Therefore an all-focused image (i.e., an image without defocus blur) can be generated from the blurred light field spectrum.

In the images captured with different camera parameters, the change of object appearance provides a cue for estimating the scene structure. However, while all photographic effects are changed with the camera parameters, only a few of them are used in traditional algorithms. For example, in depth-from-defocus algorithms, the scene structure is estimated from images with different F's or B_A's by using the differences in the degree of defocus blur. The difference in perspective transformation is usually neglected or intentionally compensated [3], [35]. This is mainly because those two effects were modeled independently. Since our framework can model them together, it is possible to develop novel algorithms using both the perspective cue and the defocus cue to obtain more accurate results.

We have shown that the image is a filtered and stretched version of the surface light field. Therefore, the bandwidth of the image can be calculated from the bandwidth of light field. In the rendering applications, we can use this information to perform adaptive sampling. Specifically, when rendering a region with a small bandwidth, which may be due to out-of-focus or low-frequency texture functions, we can use a coarser sampling rate to reduce the computational cost. Note that the estimation of the bandwidth can be calculated analytically using (31) and (42) without performing Fourier transform. This application was first proposed in [18] where the bandwidth of the light field after propagation, reflection, and occlusion is estimated, but the effect of defocus blur was not considered.

Finally, although we derive the photographic operation according to the physic entities, we may do it in a reverse way to design new imaging systems. One can change $\hat{C}_{zfF}$ and the

blocking functions until a specific effect appears in the image, and then search for a physically feasible decomposition of $\hat{C}_{zfF}$ into a number of basic propagation and reflection operations and apertures of different shapes. Some possible transformations and the corresponding effects such as pushroom projection and afocal effect are given in [31] without considering the physical feasibility.

D. Limitations

Our framework is built upon geometric optics, and therefore the effects, including diffraction and polarization, of wave optics cannot be modeled. When the size of the element, such as sensor, in the imaging system is comparable to the wavelength, the diffraction effect must be considered. In [36], it is shown that the light field is closely related to the Wigner distributions in wave optics, but unifying these two representations for modeling the wavelength-dependent effects requires more study.

Because we use first-order approximation to model the lens refraction, several third-order effects (aberrations, coma, etc.) are not discussed. However, these effects, which are within the scope of geometric optics, result in nonlinear but invertible transformations of the light field. Finding proper methods to describe these transformations and their influences on the light field spectrum is worth pursuing in the future.

VII. CONCLUSION

In this paper, we have presented a unified framework for modeling image formation using light transport analysis. We have shown that the image formation process can be characterized as a series of transformations, modulations, and projection of the light field signal. We have shown that these operations can be combined into a single equation to describe the full process and discussed several extensions and applications of this theoretical framework.

The unified framework can be used to describe all the geometric and photometric phenomena encountered in the image formation process, including perspective transform, defocus blur, and vignetting. The results obtained by this framework quantitatively match those of the traditional models.

APPENDIX

A. Mapping Between the Radiance Representations

For a specific light ray represented by (x, θ) and according to (2), (3), and (4), we have

$$S(x, \theta)\mathrm{d}x\mathrm{d}\theta = R(x, \theta)\mathrm{d}x\mathrm{d}\theta \cos\theta \tag{53}$$

$$S(x, \theta)\mathrm{d}x\mathrm{d}\theta = L([x, u]^T)\mathrm{d}x\mathrm{d}u. \tag{54}$$

Combine the previous equations and with $u = \tan\theta$, we have

$$\begin{aligned} L([x, u]^T) &= R(x, \theta)\cos\theta\frac{\mathrm{d}\theta}{\mathrm{d}u} \\ &= R(x, \theta)\cos\theta\left(\frac{\mathrm{d}\tan\theta}{\mathrm{d}\theta}\right)^{-1} \\ &= R(x, \theta)\cos^3\theta = R(x, \theta)(1 + u^2)^{-1.5}. \end{aligned} \tag{55}$$
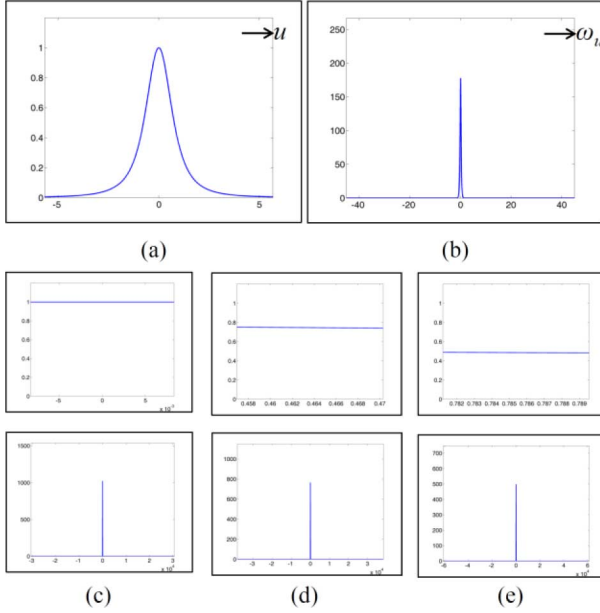
Fig. 10.   (a) Parameterization function $\}(u)$ and (b) its spectrum $\mathbf{g}(\omega_u)$. The second row shows the effective pa-rameterization function of different scene points when the aperture size is 50 units, and the third row shows the cor-re-sponding spectrums. The scene point is at (c) $(x, z) = (0, 3000), (d)(x, z) = (1500, 3000)$, and (e) $(x, z) = (3000, 3000)$.

In our example shown in Fig. 2 where the object is Lambertian, the emitted radiance R is constant regardless of the view angle. Let g denote the parameterization function $(1+u^2)^{-1.5}$. Given $R(x, \theta) = T(x)$ in (6), we have

$$L(\mathbf{x}) = T(x)g(u) \triangleq T(x)(1+u^2)^{-1.5} \qquad (56)$$

which can be approximated as $T(x)$ when the range of u is small, as shown in the next section.

B. Effect of the Parameterization Function

We repeat the parameterization function here for convenience

$$g(u) = (1+u^2)^{-1.5} \qquad (57)$$

which is plotted in Fig. 10(a) and its spectrum is shown in Fig. 10(b). Because it is a smooth function, modulating the light field with it causes a very small blur to the light field spectrum.

In this paper, our main interest is modeling the light transport in image formation of regular cameras. For normal scene configuration and camera parameters, the effective range of u of each scene point is very small. Several examples are shown in Fig. 10(c)–(e), where we can see that the effect of the parameterization function can be approximated by uniform scaling. That is, if we only consider the light rays around the direction u_0, we have

$$L(\mathbf{x}) \cong T(x)g(u_0). \qquad (58)$$

Given a specific u_0, we can embed $g(u_0)$ into the texture function and obtain the approximate equality in (6).
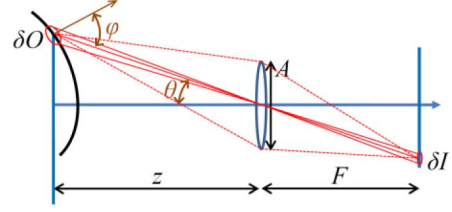

Fig. 11.   Typical scene and camera configuration for the derivation of the vignetting effect in 2-D space.

C. Cosine-Falloff in 2-D Space

For completeness we give the derivation of the vignetting effect in 2-D space using the traditional method. The derivation in 3-D space can be found in [1]. Consider the scene and camera configuration shown in Fig. 11. The lens is of diameter A at F away from the image plane. Let a patch on the Lambertian surface at distance z from the lens have area δO and the radiance value C. The image of the patch has area δI. The angle between the optical axis and the light ray from the object to the lens center is θ, and the angle between the surface normal and that light ray is φ.

Because the light ray passing through the lens center is not deflected, the solid angles formed at the lens center subtended by δO and δI are the same. That is

$$\frac{\delta I}{F/\cos\theta} = \frac{\delta O}{z/\cos\varphi}. \qquad (59)$$

If we assume the z is much larger than A, the solid angle Ω formed at the surface patch subtended by the lens is

$$\Omega \cong A\cos\theta/(z\cos\theta). \qquad (60)$$

Thus, the power δP emitting from δO and passing through the lens is

$$\delta P = (C\cos\varphi)\delta O\Omega = A(C\cos\varphi)\delta O\cos^2\theta/z \qquad (61)$$

and because no light ray from other patch reaches the patch δI, its irradiance I is

$$I = \frac{\delta P}{\delta I} = \frac{\delta O}{\delta I}\frac{A(C\cos\varphi)\cos^2\theta}{z} = C\frac{A}{F}\cos^3\theta \qquad (62)$$

which is identical to our result in (48) following the same approximation.

D. Modeling of the Telecentric Effect

We have shown that the perspective transformation and the defocus blur are both dependent upon the object depth (z) and the distance between the lens and the image plane (F). Therefore, when adjust the focus setting, the magnification factor of the object changes simultaneously. By placing an aperture/stop at the focus plane of the lens on the object side, the telecentric lens makes the magnification factor constant, no matter how F is changed.

Previously, the functionality of the telecentric lens is explained by ray tracing [32]. The systematic configuration of a
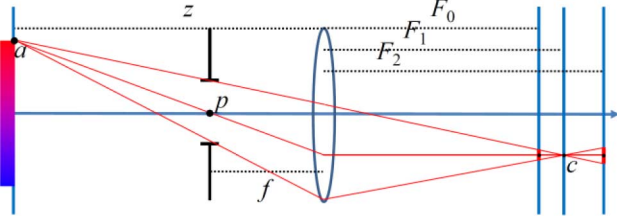
Fig. 12. Configuration of the telecentric lens. When the point a is out-of-focus, its image is always a pillbox centered at c regardless of the value of F.

telecentric lens is shown in Fig. 12. The aperture of the telecentric lens is placed in front of the lens by f units. Consider the light ray emitting from point a on the surface toward the point p at the center of the aperture. After lens refraction, this light ray is parallel to the optical axis and always hits the same point c on the image plane regardless of the changes in F. It is shown that all the light rays from a hit the image plane within a box centered at c [31]. Therefore, while the magnification factor becomes constant, the defocus effect still varies with z and F.

However, the ray tracing method is difficult to analyze. Here we show that the telecentric effect can be easily modeled using light transport analysis. According to Section III-D, the aperture of size A at distance f is mathematically a blocking function $O_A(\mathbf{x}) = \text{rect}(x/A)$.[3] Physically, the initial light field is first modulated with this blocking function before reaching the lens. However, mathematically, we can transform both the light field and the blocking function and then modulate them together. As a result, the final light field $\hat{L}_t$ is

$$\hat{L}_t(\mathbf{x}) = L_i(\hat{\mathbf{C}}_{zfF}\mathbf{x})O_A(\hat{\mathbf{C}}_{ffF}\mathbf{x}) = \hat{L}_p(\mathbf{x})\hat{O}_A(\mathbf{x})$$
$$= \hat{L}_p(\mathbf{x})\text{rect}((x - c)f/(AF)). \qquad (63)$$

We can see the transformed light field is modulated by a binary function, which is 1 if $|(x - u)| \leq (FA)/(2f)$ and 0 otherwise. For a regular camera, the blocking function is also a binary function, but whether a light ray is blocked only depends upon its entry point u on the lens.

Assume the scene is a Lambertian point at (x_0, z), we have $L_i(\mathbf{x}) = \delta(x - x_0)$, a line in 2-D space. By putting this into (63), the final light field is a line segment along $x + Fx_0/z = F\Delta u$, and the x coordinates of its two endpoints are

$$\frac{x_0f \pm (AzF\Delta/2)}{f - z}. \qquad (64)$$

Using the projection operation, the image of this segment is a pillbox centered at

$$x_0f/(f - z) \qquad (65)$$

with width

$$\left|\frac{AzF\Delta}{f - z}\right|. \qquad (66)$$

We can see that the center is invariant to the focus (F), but the width is not. This result is quantitatively identical to that by

[3]We use the notation O_A to distinguish it from the blocking function B_A of the aperture in normal lens camera.

ray tracing [32]. However, the analysis in the light field domain is still somewhat tedious. In the following, we show that the same result can be obtained more easily by spectral analysis. The Fourier transform of $\hat{O}_A$ is

$$\hat{O}_A(\boldsymbol{\omega}) = Af^{-1}\text{sinc}(0.5FAf^{-1}(\omega_x - \omega_u))\delta(\omega_x + \omega_u) \qquad (67)$$

of which all the energies fall on the plane $\omega_x + \omega_u = 0$. Therefore, the convolution effectively spreads the energies of the transformed light field along $\omega_x + \omega_u = 0$. Let the scene be the Lambertian surface with texture function T at z units from the lens, according to (41), (42), (63) and the convolution theorem, the relationship between the image spectrum and the texture spectrum is

$$\mathcal{I}(\omega_x) = A(f - z)^{-1}\text{sinc}\left(\frac{AFz\Delta\omega_x}{f - z}\right)\mathcal{T}(f\omega_x(f - z)^{-1}). \qquad (68)$$

Therefore, the image spectrum is the stretched and attenuated texture spectrum. The stretch/magnification factor is independent of F, but the attenuation factor is not. This result is quantitatively equivalent to the ones by ray tracing.

The light transport analysis makes the interpretation of the telecentric lens much easier than previous approaches. Furthermore, it is straightforward to combine the previously shown image formation with the photometric transformation. Therefore, one can derive the change of the vignetting effect in the telecentric lens system. When multiple apertures are present in the imaging system, we can modulate their blocking functions together to model their joint effect on the light field and the image.

ACKNOWLEDGMENT

REFERENCES

[1] B. K. P. Horn, *Robot Vision*. Cambridge, MA: MIT Press, 1986.
[2] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2003.
[3] S. K. Nayar, M. Watanabe, and M. Noguchi, "Real-time focus range sensor," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 12, pp. 1186–1198, Dec. 1996.
[4] S. B. Kang and R. S. Weiss, "Can we calibrate a camera using an image of a flat, textureless Lambertian surface?," in *Proc. 6th Eur. Conf. Comput. Vis.*, D. Vernon, Ed., 2000, pp. 640–653, Lecture Notes In Computer Science, vol. 1843. Springer-Verlag.
[5] M. Born and E. Wolf, *Principles of Optics*, 7th ed. Cambridge, U.K.: Cambridge Univ. Press, 1999.
[6] N. Asada, H. Fujiwara, and T. Matsuyama, "Seeing behind the scene: Analysis of photometric properties of occluding edges by the reversed projection blurring model," *IEEE Trans. Pattern. Anal. Mach. Intell.*, vol. 20, no. 2, pp. 155–67, Feb. 1998.
[7] J. R. Bergen and E. H. Adelson, "The plenoptic function and the elements of early vision," in *Computational Models of Visual Processing*. Cambridge, MA: MIT Press, 1991, pp. 3–20.
[8] M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. 23rd Annu. Conf. Comput. Graph. Interactive Tech.*, New York, 1996, pp. 31–42.
[9] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals and Systems*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.

[10] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The lumigraph," in *Proc. 23rd Annu. Conf. Comput. Graph. Interactive Tech.*, New York, NY, 1996, pp. 43–54.

[11] J. Chai, X. Tong, S. Chan, and H. Shum, "Plenoptic sampling," in *Proc. 27th Annu. Conf. Comput. Graph. Interactive Tech.*, New York, 2000, pp. 307–318.

[12] A. Isaksen, L. McMillan, and S. J. Gortler, "Dynamically reparameterized light fields," in *Proc. 27th Annu. Conf. Comput. Graph. Interactive Tech.*, New York, 2000, pp. 297–306.

[13] R. Ramamoorthi and P. Hanrahan, "A signal-processing framework for inverse rendering," in *Proc. 28th Annu. Conf. Comput. Graph. Interactive Tech.*, New York, 2001, pp. 117–128.

[14] R. Ramamoorthi and P. Hanrahan, "A signal-processing framework for reflection," *ACM Trans. Graph.*, vol. 23, no. 4, pp. 1004–1042, Oct. 2004.

[15] R. Ramamoorthi and P. Hanrahan, "Frequency space environment map rendering," in *Proc. 29th Annu. Conf. Comput. Graph. Interactive Tech.*, New York, 2002, pp. 517–526.

[16] J. Lehtinen, "A framework for precomputed and captured light transport," *ACM Trans. Graph.*, vol. 26, no. 4, pp. 13:1–13:22, Oct. 2007.

[17] J. Arvo, "The irradiance Jacobian for partially occluded polyhedral sources," in *Proc. 21st Annu. Conf. Comput. Graph. Interactive Tech.*, New York, 1994, pp. 343–350.

[18] F. Durand, N. Holzschuch, C. Soler, E. Chan, and F. X. Sillion, "A frequency analysis of light transport," *ACM Trans. Graph*, vol. 24, no. 3, pp. 1115–1126, Jul. 2005.

[19] D. Lanman, R. Raskar, A. Agrawal, and G. Taubin, "Shield fields: Modeling and capturing 3-D occluders," *ACM Trans. Graph.*, vol. 27, no. 5, pp. 131:1–131:10, Dec. 2008.

[20] R. Ramamoorthi, M. Koudelka, and P. Belhumeur, "A Fourier theory for cast shadows," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 2, pp. 288–295, Feb. 2005.

[21] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, "Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing," *ACM Trans. Graph.*, vol. 26, no. 3, pp. 69:1–69:12, Jul. 2007.

[22] C.-K. Liang, T.-H. Lin, B.-Y. Wong, C. Liu, and H. H. Chen, "Programmable aperture photography: Multiplexed light field acquisition," *ACM Trans. Graph*, vol. 27, no. 3, pp. 55:1–55:10, Aug. 2008.

[23] T. Georgiev, C. Zheng, S. K. Nayar, B. Curless, D. Salesin, and C. Intwala, "Spatio-angular resolution trade-offs in integral photography," in *Proc. Eurograph. Symp. Rendering*, 2006, pp. 263–272.

[24] T. Georgiev, C. Intwala, S. Babacan, and A. Lumsdaine, "Unified frequency domain analysis of lightfield cameras," in *Proc. Eur. Conf. Comput. Vis.*, 2008, pp. 224–237.

[25] M. Zwicker, W. Matusik, F. Durand, and H. Pfister, "Antialiasing for automultiscopic displays," in *Proc. Eurograph. Symp. Rendering*, 2006, pp. 73–82.

[26] R. Ng, "Fourier slice photography," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 735–744, Jul. 2005.

[27] A. Levin, S. W. Hasinoff, P. Green, F. Durand, and W. T. Freeman, "4-D frequency analysis of computational cameras for depth of field extension," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 97:1–97:14, Aug. 2009.

[28] D. N. Wood, D. I. Azuma, K. Aldinger, B. Curless, T. Duchamp, D. G. Salesin, and W. Stuetzle, "Surface light fields for 3-D photography," in *Proc. 27th Annu. Conf. Comput. Graph. Interactive Tech.*, New York, 2000, pp. 287–296.

[29] R. Ramamoorthi, D. Mahajan, and P. Belhumeur, "A first-order analysis of lighting, shading, and shadows," *ACM Trans. Graph.*, vol. 26, no. 1, pp. 2:1–2:21, Jan. 2007.

[30] A. Gerrard and J. M. Burch, *Introduction to Matrix Methods in Optics*. New York: Dover, 1994.

[31] A. Adams and M. Levoy, "General linear cameras with finite aperture," in *Proc. Eurograph. Symp. Rendering*, 2007, pp. 121–126.

[32] M. Watanabe and S. K. Nayar, "Telecentric optics for focus analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 12, pp. 1360–1365, Dec. 1997.

[33] Y.-H. Kao, C.-K. Liang, L.-W. Chang, and H. H. Chen, "Depth detection of light field," in *Proc. Int. Conf. Acoust., Speech Signal Process.*, 2007, vol. 1, pp. 893–896.

[34] C.-K. Liang, "Analysis, acquisition, and processing of light field for computational photography," Ph.D. dissertation, National Taiwan Univ., Taipei, Taiwan, 2008.

[35] Y. Y. Schechner and N. Kiryati, "Depth from defocus vs. stereo: How different really are they?," *Int. J. Comput. Vis.*, vol. 39, no. 2, pp. 141–162, Feb. 2000.

[36] Z. Zhang and M. Levoy, "Wigner distributions and how they relate to the light field," in *Proc. Int. Conf. Comput. Photogr.*, 2009, pp. 1–10.

[37] A. Levin, W. T. Freeman, and F. Durand, "Understanding camera trade-offs through a Bayesian analysis of light field projections," in *Proc. 10th Eur. Conf. Comput. Vis.*, 2008, vol. 5305, pp. 88–101.

Chia-Kai Liang (S'05–M'08) received the B.S. degree in electrical engineering and Ph.D. degree in communication engineering from National Taiwan University, Taipei, Taiwan, in 2004 and 2009, respectively.

From 2006 to 2007, he was a VLSI Design Engineer at Avisonic Corporation. In 2009, he was a Research Intern at Nokia Research Center, Palo Alto, CA. His research interests include computational photography, image processing, computer vision, and multimedia hardware architecture.

Dr. Liang received the CSVT Transactions Best Paper Award in 2008, IPPR Best Dissertation Award, NTU GICE Best Dissertation Award, and IICM Best Dissertation Award in 2009. He is a member of Phi Tau Phi, and a student member of ACM SIGGRAPH.

Yi-Chang Shih received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 2009.

He is currently with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA. His current research interests are quantum communication, image processing, and light field rendering.

Mr. Shih received International Physics Olympic (IPhO) Silver Medal Award in 2004.

Homer H. Chen (S'83–M'86–SM'01–F'03) received the Ph.D. degree in electrical and computer engineering from the University of Illinois, Urbana-Champaign.

Since 2003, he has been with the College of Electrical Engineering and Computer Science, National Taiwan University, Taipei, Taiwan, where he is currently an Irving T. Ho Chair Professor. Prior to that, he held various research and development management and engineering positions with U.S. companies over a period of 17 years, including AT&T Bell Labs, Rockwell Science Center, iVast, and Digital Island (acquired by Cable & Wireless). He was a U.S. delegate for ISO and ITU standards committees and contributed to the development of many new interactive multimedia technologies that are now part of the MPEG-4 and JPEG-2000 standards. His professional interests lie in the broad area of multimedia signal processing and communications.

Dr. Chen is an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY. He served as an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING from 1992 to 1994, Guest Editor for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY in 1999, and Associate Editor for *Pattern Recognition* from 1989 to 1999.