Spectral Analysis for Sampling Image-Based Rendering Data

Cha Zhang and Tsuhan Chen
Electrical & Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213
Email: {czhang, tsuhan}@andrew.cmu.edu

Abstract

Image-based rendering (IBR) has become a very active research area in recent years. The spectral analysis problem for IBR has not been completely solved. In this paper, we present a new method to parameterize the problem, which is applicable for general-purpose IBR spectral analysis. We notice that any plenoptic function is generated by light ray emitted/reflected/refracted from the object surface. We introduce the surface plenoptic function (SPF), which represents the light rays starting from the object surface. Given that radiance along a light ray does not change unless the light ray is blocked, SPF reduces the dimension of the original plenoptic function to 6D. We are then able to map or transform the SPF to IBR representations captured along any camera trajectory. Assuming some properties on the SPF, we can analyze the properties of IBR for generic scenes such as scenes with Lambertian or non-Lambertian surfaces and scenes with or without occlusions, and for different sampling strategies such as lightfield/concentric mosaic. We find that in most cases, even though the SPF may be band-limited, the frequency spectrum of IBR is not band-limited. We show that non-Lambertian reflections, depth variations and occlusions can all broaden the spectrum, with the latter two being more significant.

SPF is defined for scenes with known geometry. When the geometry is unknown, spectral analysis is still possible. We show that with the "truncating windows" analysis and some conclusions obtained with SPF, the spectrum expansion caused by non-Lambertian reflections and occlusions can be quantatively estimated, even when the scene geometry is not explicitly known.

Given the spectrum of IBR, we also study how to sample IBR data more efficiently. Our analysis is based on the generalized periodic sampling theory with arbitrary geometry. We show that the sampling density can be up to twice of that when we use rectangular sampling. The advantages and disadvantages of generalized periodic sampling for IBR are also discussed.

Keywords: plenoptic function, image-based rendering, lightfield, concentric mosaic, spectrum analysis.

. Contact Author: Prof. Tsuhan Chen, Department of Electrical and Computer Engineering, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213. Tel: (412) 268-7536, Fax: (412) 268-3890, Email: tsuhan@cmu.edu
. Work supported in part by NSF CAREER Award 9984858 and Industrial Technology Research Institute.
. Submitted to *IEEE Trans. on Circuits and Systems for Video Technology: Special Issue on Image-based Modeling, Rendering and Animation*, Sep. 2001. Revised and resubmitted in Sep. 2002. The final draft is submitted in Apr. 2003.

I. Introduction

Image-based rendering (IBR) has attracted a lot of attention recently. Proposed by Adelson and Bergen [1] as a 7D plenoptic function, IBR models a 3D dynamic scene by recording the light rays at every space location, towards every possible direction, over any range of wave-lengths and at any time. By removing the time and the wavelength dimensions, McMillan and Bishop [2] defined plenoptic modeling as generating a continuous 5D plenoptic function from a set of discrete samples. The concepts of Lightfield by Levoy and Hanrahan [3] and Lumigraph by Gortler *et al.* [4] effectively parameterized the plenoptic function into a 4D function under the constraint that the object lies in free space. Shum and He [5] proposed concentric mosaic, where the number of dimensions of the plenoptic function is decreased further to three by restricting the viewer's movement inside a circle.

The principle of the Lightfield [3] can be briefly addressed as follows. In free space, radiance does not change along a line unless the light ray is blocked. The light rays are thus parameterized by their intersections with two planes in arbitrary position. One of the planes is indexed with coordinate (u,v) and the other with coordinate (s,t). In Figure 1, we show an example where the two planes, the camera plane and the focal plane, are parallel. This is the most widely used setup. An example light ray is shown and indexed as (u_0,v_0,s_0,t_0). The coordinates are then discretized so that a finite number of light rays are recorded. In practice, we put a 2D array of cameras on the camera plane, which share the same focal plane. The images captured form a 2D image array, as is shown in Figure 2. To create a new view of the object, we just split the view into its light rays, which are then calculated by interpolating existing nearby light rays in the image array. The new view is then generated by reassembling the split rays together.

Concentric mosaic [5] further restricts that both the cameras and the viewer are on a plane. The scene is captured by mounting a camera at the end of a level beam, and shooting images at regular intervals as the beam rotates, as is shown in Figure 3. The light rays are then indexed by the camera position or the beam rotation angle α, and the pixel locations (u,v). This parameterization is equivalent to having many slit cameras rotating around a common center and taking images along the tangent direction, thus the name concentric mosaic. The concentric mosaic can be rendered with light ray interpolation, too.

IBR has many advantages over the traditional model-based rendering. When enough images are captured, IBR requires very little or even no geometrical information about the scene to realistically render it. This is very attractive due to the difficulty to construct geometrical models for objects in a real-world scene. The rendering speed of IBR is linearly proportional to the image size, but not the scene complexity, which makes IBR very suitable for rendering realistic scenes with complex geometry. Nevertheless, often a huge number of images have to be taken in order to avoid aliasing or ghosting during the rendering. It is therefore extremely important to know how many images are enough to capture a scene, based on the spectrum analysis of the plenoptic function.

Recently Chai et al. [6] proposed a method to study the frequency spectrum of lightfield. Assuming Lambertian surface and no occlusion, they showed that any image taken at position (s,t) can be mapped to the image captured

at position $(0,0)$. A key conclusion is that the spectral support of a lightfield signal is bounded by the minimum and maximum depths of objects in the scene only, no matter how complicated the scene is. They claim that by knowing the bound of the support, the maximum sampling density can be achieved by compacting copies of the spectral support in the frequency domain. They further proposed the "Sampling Curve", which gives the relationship between the number of images to be taken and the knowledge about the scene geometry (number of bits that represents the depth field). Their analysis provides a fairly good first-order approximation of the spectrum analysis of IBR. However, the dependency on mapping images captured at arbitrary position to one position (as in their case from (s,t) to $(0,0)$) prevents it from being applied to more complicated scenes such as non-Lambertian surface, scenes with occlusions and other IBR methods such as concentric mosaic.

Due to the complexity of the problem, few ideas have been proposed to enhance the work in [6]. In [7], Chan and Shum proposed to view the plenoptic sampling problem as a multidimensional sampling problem. They showed that depth variation causes the plenoptic function to spread in the frequency domain. The same image mapping in [6] is applied in their paper and thus it bears the same disadvantages.

Marchand-Maillet and Vetterli [8] extended the work for scenes with functional surfaces. They fixed the light ray direction and tried to find a one-to-one mapping from the object surface to the camera plane. This one-to-one mapping is valid when no occlusion occurs. They showed that even when there is no occlusion, a band-limited signal pasted on a functional surface will not result in a band-limited lightfield spectrum. The limitation of their analysis is that they did not consider general scenes with occlusions or non-Lambertian surfaces. Lightfield was the only IBR method that was considered. They also proposed a geometrical analysis for ideal sampling of lightfield. Occlusion was discussed, yet no significant and practical results were drawn from their analysis.

Another interesting work was done by Lin and Shum [9]. They performed sampling analysis on both lightfield and concentric mosaic with scale-space theory under constant-depth assumption. The bounds are derived from the aspect of geometry and based on the goal that no "spurious detail" should be generated during the rendering (referred as the causality requirement). Although the viewpoint of their analysis is very interesting, the constant-depth assumption is often too strong to describe the real world.

In this paper, we propose a new method to parameterize the plenoptic function for IBR spectrum analysis. Similar to [6], we also work in the Fourier transform domain. However, instead of mapping images captured at arbitrary position to one position, we map the IBR representation onto the so-called surface plenoptic function (SPF, as defined in Section II). This allows us to solve for occluded, non-Lambertian scenes, as well as other IBR capturing methods such as concentric mosaic. For non-Lambertian surface, we show that when the surface emission/reflection changes slowly along different directions, the Fourier transform of the scene is also expanded by a limited amount. We then show that when the scene has depth variations, the spectrum will be broadened due to the nonlinearity between the SPF and the IBR representation. We also show that occlusion causes a modulation effect between layers in the scene, which generates additional frequency components. Among the three sources we find that depth variations and occlusions are often much stronger in spectrum spreading. For analyzing the spectrums of scenes with unknown geometry, we apply the "truncating windows" analysis [6] and some of the

above conclusions. We are able to quantatively estimate the spectrum spread due to non-Lambertian surface and occlusions.

Given the spectrum, we apply generalized periodic sampling theory with arbitrary geometry to IBR. We show that non-rectangular or quincunx/hexagonal sampling can reduce the sampling rate by a factor of 2 theoretically. The price is the increase of complexity for both data capturing and rendering. Experiments and comments are given at the end of the paper.

The paper is organized as follows. In Section II, we present the new method of parameterization. Section III gives examples on different scenes sampled with lightfield. We analyze the lightfield spectrums of scenes with unknown geometry in Section IV. Section V briefly shows the analysis for concentric mosaic. We describe the non-rectangular sampling for IBR in Section VI. Conclusions and future work are given in Section VII.

II. Parameterization Method

Image-based rendering, as defined by McMillan and Bishop in [2], is to generate a continuous representation of the plenoptic function given a set of discrete samples (complete or incomplete) from that function. To analyze the spectrum for IBR, it is important to have a close look at how plenoptic function is generated. Obviously, any light ray in the free space has a source. It can be either emitted from some light source (e.g., the Sun), or reflected by some object surface. If the object is transparent, refraction is also involved. Let the entire surface of all the light sources and objects be **S**. We can always trace a light ray in the free space back to a point on **S**. Assume the radiance does not change along a line unless the light ray is blocked, the 7D plenoptic function can be reparameterized to 6D including time (1D), wavelength (1D), point on the surface **S** (2D), and azimuth and elevation angles (2D) the light ray is emitted. We name this 6D function as the *surface plenoptic function* (SPF).

SPF bears close relationship to the surface light field proposed in [10][11]. A minor difference is that SPF is more general than surface light field because it includes the time and wavelength information. Wood *et al.* [11] focused on the representation, rendering and compression of surface light field, while we use SPF for sampling analysis. Wong's plenoptic illumination function [12] and Lin's reflected irradiance field [13] are also similar to SPF in the sense that they tried to model different light ray radiances from same surface points. However, their work studied scenes under different illuminations (fixed viewpoint), while we consider scenes under different viewpoints (fixed illumination). The sampling of reflected irradiance field [13] is independent of the scene geometry, because the viewpoint is fixed and the correspondence between light rays from the same surface point is known (always at the same pixel position). In contrast, for IBR sampling analysis we need some geometry information in order to know the light ray correspondences, such as the constant depth assumption made in [6].

The assumption that the radiance does not change along a line unless the light ray is blocked is widely adopted in many IBR representations [3][5]. Under this assumption, the light rays coming out from the convex hull of the scene can also be parameterized based on the camera surface (the surface where the cameras locate), which is the essence of IBR representations. Since light rays start from the object surface and end at capturing cameras, there exists an onto mapping from the SPF to the IBR representations. Such a mapping depends on both the scene geometry and the camera surface, but not the surface property such as BRDF. If we have some knowledge about

the scene, in other words, if we know some property about the SPF, related property can be derived for the IBR representation. More importantly, the mapping does not require the scene to have no occlusions or Lambertian surface, which is very attractive for IBR spectral analysis.

Without loss of generality, we use the 2D world as example throughout this paper for conciseness. The conclusions drawn in this paper are easy to extend to the 3D world. In the 2D world, surface of objects/light sources is described with curves. Ignoring time and wavelength, the SPF is 2D: one dimension for describing a point on a curve, the other for illustrating the direction of the light ray emitted/reflected/refracted. An example scene is shown in Figure 4. The surface can be represented by either $S_i(x,y)=0$ or $\begin{cases} x = x_i(s) \\ y = y_i(s) \end{cases}$, where *s* can be the arc length, *i* be the index for different objects. For a certain object *i*, define its SPF as:

$$l_i(s,\theta) \text{ on the curve } \begin{cases} x = x_i(s) \\ y = y_i(s) \end{cases} \tag{1}$$

where $0 \le \theta < 2\pi$ is the direction of the light ray. *l* is the radiance of the light ray that can be traced back to the surface point determined by *s*. Notice that the above function does not appear to be related with what people often use for calculating lightings, such as surface normal, BRDF etc. We intend to do so because it is often too complicated if we try to model how the light rays are generated, in addition to the fact that such a model does not always exist for real scenes. Therefore, we only consider the resultant lighting effects in (1), and assume that we have some knowledge about the SPF. If we are able to model the lighting very well, the following analysis can still apply after calculating the lighting effects based on the known model.

Lambertian is the first assumption we could make for the scene, since it has been exclusively used in the literature. In terms of SPF, Lambertian property gives the following relationship:

$$l_i(s,\theta) = l_{is}(s) \tag{2}$$

and its Fourier transform is: $L_i(\Omega_s,\Omega_\theta) = L_{is}(\Omega_s)\delta(\Omega_\theta)$, where $L_i(\Omega_s,\Omega_\theta)$ is the Fourier transform of $l_i(s,\theta)$, $L_{is}(\Omega_s)$ is the Fourier transform of $l_{is}(s)$[1].

In the real world, pure Lambertian objects are rare. Highly reflective surface (like a mirror) is infrequent, too. In most cases, light rays from the same point on the object surface tend to be similar and have slow changes with respect to their angles. It is therefore reasonable to assume that $L_i(\Omega_s,\Omega_\theta)$ is can be approximated by a band-limited signal. That is:

$$L_i(\Omega_s,\Omega_\theta) \approx L_i(\Omega_s,\Omega_\theta) I_{B_{\theta i}}(\Omega_\theta) \tag{3}$$

where $I_{B_{\theta i}}(\Omega_\theta)$ is the indicator function, which is defined as:

$$I_{B_{\theta i}}(\Omega_\theta) = \begin{cases} 1, & \text{if } |\Omega_\theta| < B_{\theta i}, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

[1] Strictly speaking, at a certain point on the surface, $l_i(s,\theta)$ need to be truncated on θ, because we can only observe light rays that come out of the object. Therefore, along Ω_θ $L_i(\Omega_s,\Omega_\theta)$ cannot be a delta function, nor can it be band-limited. However, we assume that this windowing artifact is negligible.

and $B_{\theta i}$ defines the bandwidth. Our band-limitness assumption can be connected to the band-limitness of the surface BRDF with the signal-processing framework for inverse rendering recently presented by Ramamoorthi and Hanrahan [14]. For points on a reflective surface, the outgoing light can be described as the convolution of the incoming light and the surface BRDF. If the incoming light is far (thus the incoming light can be considered as a delta function with respect to the angle), whenever the BRDF is band-limited, the outgoing light will be band-limited.

In order to capture the plenoptic function or surface plenoptic function, existing IBR approaches align cameras on a path/surface and take images for the scene. For example, cameras are placed on a plane in lightfield, and on a circle in concentric mosaic. In the 2D world, 2D lightfield has cameras on a line, while 2D concentric mosaic has cameras on a circle. In general, the cameras can be put along an arbitrary curve, as is shown in Figure 4. Let the camera path be $S_c(x,y)=0$ or $\begin{cases} x = x_c(t) \\ y = y_c(t) \end{cases}$, where t is the arc length. Assume that the camera path curve is differentiable, and the optical axes of our cameras are identical to the normals of the camera path. That is, at arc length t, the optical axis has direction $(-y_c'(t), x_c'(t))$, where $x_c'(t)$ and $y_c'(t)$ are the first order derivatives. Denote the direction of the optical axis with angle β, then $\tan(\beta) = -\frac{x_c'(t)}{y_c'(t)}$. The image pixels can be indexed by the angle between the captured light ray and the optical axis, as is represented by α in Figure 4. Denote the radiance of the light ray captured at arc length t, angle α as $l_c(t,\alpha)$. The goal of spectral analysis for this specific IBR representation is to find the Fourier transform of $l_c(t,\alpha)$, denoted as $L_c(\Omega_t,\Omega_\alpha)$, so that we can determine the minimum number of images we have to capture. Given the knowledge we have on the SPF $l_i(s,\theta)$ and its Fourier transform $L_i(\Omega_s,\Omega_\theta)$, the strategy is to associate $l_c(t,\alpha)$ with $l_i(s,\theta)$ and hope that we can represent $L_c(\Omega_t,\Omega_\alpha)$ in terms of $L_i(\Omega_s,\Omega_\theta)$. In the next two sections, we will show examples where we can apply the above strategy to lightfield and concentric mosaic.

Before moving on to the next section, it is necessary to compare our work to the work by Marchand-Maillet and Vetterli in [8]. Being independently developed, they share some similar ideas. For example, both of them try to find a mapping between the captured lightfield and the light rays from the object surface. However, our approach is more general than that in [8]. We employ the concept of surface plenoptic function, which allows us to perform a better analysis on general IBR representations, as long as we can represent $L_c(\Omega_t,\Omega_\alpha)$ in terms of $L_i(\Omega_s,\Omega_\theta)$. This includes lightfield and concentric mosaic. In addition, occlusions and non-Lambertian surface effects can be better analyzed within our framework. Another work that might be related is shape from texture (SFT) [15]. Both SFT and our approach are about transforming some spectral property from a certain object surface to the captured images. However, IBR spectral analysis is significantly different from shape-from texture. First, the goal of SFT is to reconstruct the scene geometry from texture information, while we try to find the Fourier spectrum of the IBR representation given the geometry. Second, SFT often recovers geometry from a single image (although it can be extended to multiple views), while we deal with many images. Third, SFT only considers texture spectrum (e.g.,

isotropy or homogeneity), while we consider both texture and light ray radiance change along different directions, including non-Lambertian surface and occlusions.

III. Analysis of Lightfield

As shown in Figure 5, 2D lightfield is parameterized by two parallel lines, indexed by *t* and *v*, respectively. The *t* line is the camera line, while the *v* line is the focal line. The distance between the two lines is *f*, which is the focal length of the cameras. It is easy to show that a light ray indexed by pair (t,v) satisfies the following algebraic equation:

$$fx - vy - ft = 0 \tag{5}$$

Notice that the focal line is indexed locally with respect to where the camera is.

The relationship between lightfield $l_c(t,v)$ and the SPF $l_i(s,\theta)$ is as follows. For the same light ray emitted/reflected/refracted from a surface point, it must be captured at the corresponding angle. That is:

$$\tan(\theta) = \frac{f}{v} \text{ or } \theta = \frac{3\pi}{2} - \tan^{-1}\left(\frac{v}{f}\right) \tag{6}$$

where $-v_0 \le v \le v_0$ and $2\tan^{-1}\left(\frac{v_0}{f}\right)$ tells the field of view. The above equation is actually the mapping between a pixel's angular position θ and its image coordinate *v*. Such a mapping can be linearized as:

$$\theta \approx \frac{3\pi}{2} - \frac{v}{f} \tag{7}$$

The above linearization will introduce 4.3% maximum error if the FOV of the camera is 40º. In practice, the approximation in Equation (7) can be replaced by a simple pixel re-arrangement.

Another constraint is that the light ray (t,v) can be traced back to a cross point on the object surface, whose arc length *s* can be obtained through solving:

$$\begin{cases} x = x_i(s) \\ y = y_i(s) \\ fx - vy - ft = 0 \end{cases} \tag{8}$$

When multiple objects exist in the scene or some objects occlude themselves, Equation (8) may have multiple answers. We have to figure out which cross point is the closest to the cameras. The closest point will occlude all the others. This may make scenes with occlusions hard to analyze. However, for simple scenes this is still doable and we will show examples later in this section.

A. Scene at a constant depth

The simplest scene we can have for lightfield is one at a constant depth. The surface can be described by:

$$\begin{cases} x = x_0(s) = s \\ y = y_0(s) = d_0 \end{cases} \tag{9}$$

We can solve Equation (8) without concerning about occlusion:

$$fs - vd_0 - ft = 0 \Rightarrow s = \frac{vd_0}{f} + t \tag{10}$$

The lightfield spectrum can be derived as:

$$\begin{aligned} L_c(\Omega_t, \Omega_v) &= \iint l_c(t,v) e^{-j\Omega_t t - j\Omega_v v} dtdv \\ &= \iint l_0\left(\frac{vd_0}{f} + t, \frac{3\pi}{2} - \frac{v}{f}\right) e^{-j\Omega_t t - j\Omega_v v} dtdv \\ &= \iint l_0(s,\theta) e^{-j\Omega_t\left[s - \left(\frac{3\pi}{2} - \theta\right)d_0\right] - j\Omega_v\left(\frac{3\pi}{2} - \theta\right)f} fdsd\theta \\ &= fL_0(\Omega_t, d_0\Omega_t - f\Omega_v) e^{j\frac{3\pi}{2}(d_0\Omega_t - f\Omega_v)} \end{aligned} \tag{11}$$

We can see that the spectrum of the lightfield at constant depth is a rotated version of the SPF spectrum, with some constant factor in magnitude and some shift in phase. The rotation angle is determined by the scene depth d_0 and the focal length *f*.

If the object surface is Lambertian, we have $L_0(\Omega_s, \Omega_\theta) = L_{0s}(\Omega_s)\delta(\Omega_\theta)$ as in Equation (2). Therefore,

$$\begin{aligned} L_c(\Omega_t, \Omega_v) &= fL_0(\Omega_t, d_0\Omega_t - f\Omega_v) e^{j\frac{3\pi}{2}(d_0\Omega_t - f\Omega_v)} \\ &= fL_{0s}(\Omega_t)\delta(d_0\Omega_t - f\Omega_v) e^{j\frac{3\pi}{2}(d_0\Omega_t - f\Omega_v)} \end{aligned} \tag{12}$$

This is a tilted line in the (Ω_t, Ω_v) space, which is the same conclusion as that in [6].

When the object surface is non-Lambertian but the SPF is band-limited, we have $L_i(\Omega_s, \Omega_\theta) = L_i(\Omega_s, \Omega_\theta) I_{\Omega_{i0}}(\Omega_\theta)$ as in Equation (3). Consequently,

$$\begin{aligned} L_c(\Omega_t, \Omega_v) &= fL_0(\Omega_t, d_0\Omega_t - f\Omega_v) e^{j\frac{3\pi}{2}(d_0\Omega_t - f\Omega_v)} \\ &= fL_0(\Omega_t, d_0\Omega_t - f\Omega_v) e^{j\frac{3\pi}{2}(d_0\Omega_t - f\Omega_v)} I_{B_{\theta 0}}(d_0\Omega_t - f\Omega_v) \end{aligned} \tag{13}$$

The spectrum is also tilted, but this time it has a finite width $\frac{2B_{\theta 0}}{\sqrt{d_0^{\,2} + f^2}}$ perpendicular to the tilted spectrum (or $\frac{2B_{\theta 0}}{d_0}$ horizontally) because of the indicator function.

The above analysis is illustrated in Figure 6. A scene at constant depth has two sinusoids (different frequency) pasted on it as texture, as shown in Figure 6 (a). Figure 6 (b) is the epipolar image (EPI) when the scene is Lambertian. Figure 6 (c) is its Fourier transform. The spectrum has several peaks because the texture on the scene object is pure sinusoids. It basically lies on a tilted line with some small horizontal and vertical windowing artifact that is due to the truncation of the range of *s* and θ. We ignored the windowing artifacts in our analysis for simplicity. Figure 6 (d) shows the EPI for a non-Lambertian case at the same depth. It can be seen that because of the non-Lambertian property, its Fourier transform in Figure 6 (e) is expanded towards the normal direction of the tilted line.

B. Scene on a tilted line

We next study a scene on a tilted line. We write the surface equation as:

$$\begin{cases} x = x_0(s) = s\cos\varphi + x_0 \\ y = y_0(s) = s\sin\varphi + y_0 \end{cases} \tag{14}$$

where $0 \le \varphi < \pi$ is an angle that is known. Assuming no occlusion, we can solve Equation (8) as:

$$f(s\cos\varphi + x_0) - v(s\sin\varphi + y_0) - ft = 0 \Rightarrow s = \frac{ft - fx_0 + vy_0}{f\cos\varphi - v\sin\varphi} \tag{15}$$

To guarantee that within each captured image the parameter s is continuous, φ cannot be arbitrary. If $-v_0 \le v \le v_0$, we can have constraint that $f\cos\varphi - v_0|\sin\varphi| > 0$.

The lightfield spectrum is:

$$\begin{aligned} L_c(\Omega_t, \Omega_v) &= \iint l_c(t,v) e^{-j\Omega_t t - j\Omega_v v} dtdv \\ &= \iint l_0\left(\frac{ft - fx_0 + vy_0}{f\cos\varphi - v\sin\varphi}, \frac{3\pi}{2} - \frac{v}{f}\right) e^{-j\Omega_t t - j\Omega_v v} dtdv \\ &= \iint l_0(s,\theta) e^{-j\Omega_t\left[s\left(\cos\varphi - \left(\frac{3\pi}{2}-\theta\right)\sin\varphi\right) + \left(x_0 - \left(\frac{3\pi}{2}-\theta\right)y_0\right)\right] - j\Omega_v\left(\frac{3\pi}{2}-\theta\right)f} \left(\cos\varphi - \left(\frac{3\pi}{2}-\theta\right)\sin\varphi\right) fdsd\theta \end{aligned} \tag{16}$$

Unfortunately, even for a scene as simple as a line, Equation (16) is too complex to solve. Here we consider a case where Equation (16) can be further simplified. Let the scene be Lambertian, i.e., $l_0(s,\theta) = l_{0s}(s)$ as in Equation (2). We have:

$$\begin{aligned} L_c(\Omega_t, \Omega_v) &= \iint l_0(s,\theta) e^{-j\Omega_t\left[s\left(\cos\varphi - \left(\frac{3\pi}{2}-\theta\right)\sin\varphi\right) + \left(x_0 - \left(\frac{3\pi}{2}-\theta\right)y_0\right)\right] - j\Omega_v\left(\frac{3\pi}{2}-\theta\right)f} \left(\cos\varphi - \left(\frac{3\pi}{2}-\theta\right)\sin\varphi\right) fdsd\theta \\ &= \iint l_{0s}(s) e^{-j\Omega_t\left[s\left(\cos\varphi - \left(\frac{3\pi}{2}-\theta\right)\sin\varphi\right) + \left(x_0 - \left(\frac{3\pi}{2}-\theta\right)y_0\right)\right] - j\Omega_v\left(\frac{3\pi}{2}-\theta\right)f} \left(\cos\varphi - \left(\frac{3\pi}{2}-\theta\right)\sin\varphi\right) fdsd\theta \\ &= \int L_{0s}\left[\Omega_t\left(\cos\varphi - \left(\frac{3\pi}{2}-\theta\right)\sin\varphi\right)\right] e^{-j\Omega_t\left(x_0 - \left(\frac{3\pi}{2}-\theta\right)y_0\right) - j\Omega_v\left(\frac{3\pi}{2}-\theta\right)f} \left(\cos\varphi - \left(\frac{3\pi}{2}-\theta\right)\sin\varphi\right) fd\theta \end{aligned} \tag{17}$$

If we have a sinusoid pasted as the texture, e.g., $l_{0s}(s) = \sin(\Omega_0 s)$. Its Fourier transform is:

$$L_{0s}(\Omega_s) = \frac{1}{2j}\left[\delta(\Omega_s - \Omega_0) - \delta(\Omega_s + \Omega_0)\right] \tag{18}$$

Due to symmetry, let us consider the magnitude of $L_c(\Omega_t, \Omega_v)$ when $\Omega_t > 0$ and $\Omega_s > 0$. Let $\gamma = \frac{3\pi}{2} - \theta$ and $-\frac{v_0}{f} \le \gamma \le \frac{v_0}{f}$ due to the limited FOV. Plug into Equation (17) and we have:

$$\begin{aligned} |L_c(\Omega_t, \Omega_v)| &= \frac{1}{2}\left|\int_{-\frac{v_0}{f}}^{\frac{v_0}{f}} \delta[\Omega_t(\cos\varphi - \gamma\sin\varphi) - \Omega_0] e^{-j\Omega_t(x_0 - \gamma y_0) - j\Omega_v \gamma f} (\cos\varphi - \gamma\sin\varphi) fd\gamma\right| \\ &= \begin{cases} \dfrac{\Omega_0 f}{2\Omega_t^{\,2}|\sin\varphi|}, & \text{when } \dfrac{\Omega_0 f}{f\cos\varphi + v_0|\sin\varphi|} \le \Omega_t \le \dfrac{\Omega_0 f}{f\cos\varphi - v_0|\sin\varphi|} \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{19}$$

Figure 7 (a) shows a scene on a line with $\varphi = \pi/6$. A single sinusoid is pasted on the line as texture. Figure 7 (b) is the EPI of the lightfield, and Figure 7 (c) is its Fourier transform. Notice that our analysis matches very well with the example. For example, the spectrum is non-zero only between certain Ω_t thresholds. In the non-zero region, the spectrum decays when $|\Omega_t|$ increases. The spectrum is bounded for Ω_v due to the minimum and maximum depth of the scene, as was described in Section III-A. One thing we may notice in this example is that compared with the original SPF (which is just a sinusoid), a scene as simple as a tilted line can cause the IBR spectrum to spread a lot. Considering the fact that light rays from a certain surface point tend to change slowly in the real world, we may conclude that the spectrum spreading in real world scenes is mainly caused by its irregular shape, including the occluding effects which we will discuss in the next subsection.

There are two practical considerations. First, the geometry of the scene may be very difficult to obtain. When the geometry is unknown, we may approximate the scene geometry with piecewise constant depths, as discussed in Section IV. Following the "truncating windows" analysis there, we may obtain first-order approximations of the spectrum, even if the scene is non-Lambertian. Second, given the scene geometry, sometimes an analytical expression associating the spectrum of SPF and that of IBR representation may not exist. To solve for such scenes, numerical methods may be applied. For example, by replacing $l_0(s,\theta)$ with the inverse Fourier transform of $L_0(\Omega_s,\Omega_\theta)$, Equation (16) could be solved numerically.

C. Occlusion between objects

When occlusion happens, IBR spectral analysis becomes very difficult due to the discontinuity at the occlusion boundary. In fact, any occlusion will simply cause the IBR spectrum to be band-unlimited. In this subsection, we analyze occluded scenes under certain assumptions, and give some insights on the formation of the spectrum of such scenes.

With SPF, solving for occluded scenes means finding the closest cross point among multiple solutions for Equation (8). We first assume that objects do not occlude themselves. Marchand-Maillet and Vetterli [8] once derived the no-occlusion condition for functional surfaces. Similarly, in our notation, no occlusion requires:

$$\max_s \left| \frac{y_i'(s)}{x_i'(s)} \right| < \frac{f}{v_0} \tag{20}$$

where $y_i'(s)$ and $x_i'(s)$ are first order derivatives. Equation (20) simply means that the slope of the surface should not be greater than that of any possible light rays captured. This assumption is, however, hard to justify. We treat this as an example where occlusion can be solved under our parameterization. In the meantime, in practice mutual occlusions are often more significant than self-occlusions[2], and self occluded objects can sometimes be decomposed into smaller objects so that the occlusions become mutual.

When objects do no occlude themselves, their spectrums can be obtained through previously mentioned methods, as there are no occlusions. Let the number of objects in the scene be *N*. Let $l_{ci}(t,v), 0 \le i \le N-1$ be the *N*

[2] We claim mutual occlusion is more significant than self-occlusion because the former often cause a sharp boundary/edge in the EPI, while self-occlusion often does not if the surface normal changes slowly.

lightfields and $L_{ci}(\Omega_t, \Omega_v), 0 \le i \le N-1$ be their Fourier transforms. We also define a silhouette lightfield for each object *i* as:

$$sil_{ci}(t,v) = \begin{cases} 1, & \text{when light ray } (t,v) \text{ can be traced back to object } i \\ 0, & \text{otherwise} \end{cases}, 0 \le i \le N-1 \qquad (21)$$

Their Fourier transforms are denoted as $SIL_{ci}(\Omega_t, \Omega_v), 0 \le i \le N-1$. Notice that the silhouette's spectrum can be obtained by setting $l_i(s,\theta) = 1$ on object *i*'s surface.

We are now ready to find the spectrum of the occluded scene. Since Fourier transform is linear, the overall spectrum is simply the sum of individual objects' spectrums. If an object is not occluded, we can just add its spectrum to the overall one. Otherwise, let object *i* be occluded by *M* other objects ($M < N-1$). Denote the *M* occluding objects' silhouette lightfields as $sil_{ci_j}(t,v), 0 \le j \le M-1$. The contribution object *i* has to the overall lightfield can be written as:

$$l_{ci}^{occ}(t,v) = l_{ci}(t,v) \prod_{j=0}^{M-1} \left[1 - sil_{ci_j}(t,v)\right] \qquad (22)$$

where $l_{ci}^{occ}(t,v)$ is the occluded lightfield of object *i*. Its Fourier transform is:

$$L_{ci}^{occ}(\Omega_t, \Omega_v) = L_{ci}(\Omega_t, \Omega_v) * \left[\delta(\Omega_t, \Omega_v) - SIL_{ci_0}(\Omega_t, \Omega_v)\right] * \cdots * \left[\delta(\Omega_t, \Omega_v) - SIL_{ci_{M-1}}(\Omega_t, \Omega_v)\right] \qquad (23)$$

where $*$ stands for convolution, $\delta(\Omega_t, \Omega_v)$ is the Fourier transform of constant 1. From Equation (23) we can see that the spectrum of an occluded object is its unoccluded spectrum modulated by all the occluding objects' silhouette spectrum. This modulation will bring some additional components to the overall spectrum.

Figure 8 (a) shows an example Lambertian scene that has three objects. Each object is at a constant depth and has two sinusoids pasted as texture. Therefore, from what we had in subsection A, if no occlusion is considered, for $0 \le i \le 2$, $L_{ci}(\Omega_t, \Omega_v)$ and $SIL_{ci}(\Omega_t, \Omega_v)$ both lie on a tilted line whose slope is $\frac{d_i}{f}$, where d_i is the depth of the object. Since objects at greater depth are occluded by closer objects, we will notice additional modulated components along larger slope lines. Moreover, the additional modulated components will be along smaller slopes corresponding to the closer object. This is clearly shown in Figure 8 (c).

IV. Analysis of Scenes with Unknown Geometry

The surface plenoptic function is defined on the scene's geometry. It may cause the misunderstanding that the above analysis is not applicable when the scene geometry in unknown. In this section, we show how such scenes can be analyzed with the "truncating windows" analysis [6] and the conclusions obtained above.

We know that it is possible to approximate a complex scene with multiple constant depth layers. In [6], Chai *et al.* adopted such assumption and gave a "truncating windows" analysis for non-Lambertian scenes. They showed that if there is no occlusion, the spectrum support of each layer looks like a narrow ellipse. The overall spectrum is simply the sum of the spectrums for all the layers. As a first order approximation, the overall spectrum will be

between two slopes – one determined by the minimum depth $d_{\min}$ and the other determined by the maximum depth $d_{\max}$. This is drawn in Figure 9 (a). Without loss of generality we assume in this paper that the spectrum is mainly bounded by the resolution of the camera, and the unit of the *u* axis is 1, so that $-\pi < \Omega_v < \pi$. If rectangular sampling is applied, at minimum sampling rate the spectrum is replicated as in Figure 9 (b). The corresponding maximum sampling distance is $\Delta t_{\max} = \frac{2}{f\left(1/d_{\min} - 1/d_{\max}\right)}$, where *f* represents the focal length. The authors of [6] also proposed to use $d_{opt} = \frac{2}{1/d_{\min} + 1/d_{\max}}$ as the best depth to render the scene, which can be easily obtained through Figure 9 (b).

If the scene is non-Lambertian, it is clear that the "truncating windows" analysis is still valid except that for each layer, the spectrum support becomes a "fatter" ellipse. This causes the original fan-like spectrum to be expanded. If the SPF is band-limited along Ω_θ, the lightfield bandwidth expansion is also limited, as is shown in Figure 10 (a). The dotted region stands for the extra spectrum support caused by the non-Lambertian surface. A conservative estimation of the amount of expansion is $\frac{B_{\theta 0}}{d_{\min}}$ on each side along the horizontal axis Ω_t, based on the discussion after Equation (13). Therefore, with rectangular sampling, the best way to compact the spectrum is as Figure 10 (b), and the sampling rate of non-Lambertian scenes has to be increased. Nevertheless, the effect of non-Lambertian property is often difficult to observe in their spectrum because typically the light rays from the same surface point change radiance very slowly. We next show an example where we have to take more images in order to render correctly because of the non-Lambertian property. We render two spheres (Figure 11 (a) and (b)) - one is Lambertian and the other is a purely reflective surface. Meanwhile, we assume that we have a 6-bit depth map of the scene. According to the sampling curve proposed in [6], the number of images required for rendering the scene is reduced if we have an accurate depth map. We found that at a certain sampling rate, we are able to reconstruct the Lambertian scene very well as in Figure 11 (c). However, if we use the same sampling rate for the reflective surface, the reconstruction is very bad (Figure 11 (d)). This means that we need to sample more images to reconstruct the latter surface.

The "truncating windows" analysis can be applied to more complex scenes if occlusions have to be considered. When the scene is heavily occluded, far layers will be blocked by close layers, which causes a spectrum modulation as in Equation (23). In worst case, the major occluding objects are at $d_{\min}$, which generates additional modulated components along the slope $\frac{d_{\min}}{f}$, where *f* is the focal length. The corresponding spectrum is given in Figure 12 (a). The dotted support is the additional modulated components. If we still apply the rectangular sampling strategy, the spectrum can be compacted as in Figure 12 (b). Notice that we have to increase the minimum sampling rate by a factor of 2. Or, the maximum sampling distance is $\Delta t_{\max} = \frac{1}{f\left(1/d_{\min} - 1/d_{\max}\right)}$. Meanwhile, from Figure 12 (b) we see that the optimal rendering depth becomes $d_{\min}$ instead of d_{opt}. In general,

when the major occluding objects are at distance d_{occ}, the spectrum is shown in Figure 12 (c). After the most compact rectangular sampling, the spectrum is shown in Figure 12 (d). The minimum sampling rate and the optimal rendering depth both depend on d_{occ}.

We give an example to verify the above analysis. In Figure 13 (a) we show an OpenGL scene composed of 9×9×9 cubes. The cubes stay on a 3D regular grid and have a lot of occlusions. The scene parameters are: focal length $f = 128$, minimum depth $d_{\min} = 6$ and maximum depth $d_{\max} = 18$. All the renderings are done through depth-corrected bilinear interpolation. We first sample the scene with the approach in [6], i.e., $\Delta t_{\max} = \frac{2}{f(1/d_{\min} - 1/d_{\max})} = 0.1406$. Figure 13 (b) and (c) show the rendered scene with rendering depth d_{opt} and $d_{\min}$, respectively. We found that the scene rendered with d_{opt} appears better than the other with $d_{\min}$. However, both images appear blurred. In Figure 13 (d) and (e) we increase the sampling rate by a factor of 2. Figure 13 (d) is rendered at d_{opt}, while Figure 13 (e) is rendered at $d_{\min}$. It is obvious that (e) is much better than (d), and (e) is also better than (b). The fact that (e) is much better than (d) proves that the additional modulated components have to be considered during the sampling and rendering. If the additional modulated components could be ignored, rendering at d_{opt} would have at least generated as good results as (e). That is, from Figure 12 (b) if there was no dotted area due to the occlusions, d_{opt} would be one of the best rendering depths.

V. Analysis on Concentric Mosaic

2D concentric mosaic is captured by rotating a camera along a circle, as is shown in Figure 14. Let the circle radius be *R*. We index concentric mosaic by two angles, α and β. $-\alpha_0 \le \alpha \le \alpha_0$ represents the angle between the light ray and the camera optical axis; β denotes where the camera is on the circle. For concentric mosaic, we use polar coordinates instead of rectangular coordinates. The object surfaces are represented as $S_i(r, \varphi) = 0$ or $\begin{cases} r = r_i(s) \\ \varphi = \varphi_i(s) \end{cases}$.

The light ray line indexed by pair (β, α) is:

$$r - \frac{R \sin\alpha}{\sin(\alpha + \beta - \varphi)} = 0 \tag{24}$$

Equation (24) is fairly complex if we want to solve for α and β from *r* and φ. Provided that in practice the FOV of the cameras is often limited[3], we let $\sin\alpha \approx \alpha$. Because of the setup of concentric mosaic, the objects in the scene must be outside the capturing circle, we have $|\alpha + \beta - \varphi| < |\alpha|$. Therefore, we can also let $\sin(\alpha + \beta - \varphi) \approx \alpha + \beta - \varphi$. Equation (24) is thus simplified as:

$$r(\alpha + \beta - \varphi) - R\alpha = 0 \tag{25}$$

[3] When the FOV is about 40 degrees, i.e., $\alpha_0 \approx \pi/9$. Making $\sin\alpha \approx \alpha$ has 2% maximum error. Notice that in concentric mosaic we have the same linearization error as in lightfield (Equation (7)).

We also have simple relationship:

$$\alpha + \beta + \pi = \theta \tag{26}$$

Similar to the analysis in lightfield, light ray (α, β) should be traced back to a cross point on the object surface, whose arc length *s* can be obtained through solving:

$$\begin{cases} r = r_i(s) \\ \varphi = \varphi_i(s) \\ r(\alpha + \beta - \varphi) - R\alpha = 0 \end{cases} \tag{27}$$

Equation (27) is in great similarity to Equation (8). Therefore, discussions on lightfield can be borrowed directly. In this section, we only give a simple example by considering a scene at constant depth.

A constant depth concentric mosaic scene is described as:

$$\begin{cases} r = r_i(s) = r_0 \\ \varphi = \varphi_i(s) = \dfrac{s}{r_0} \end{cases} \tag{28}$$

where r_0 is the depth of the scene. We can easily solve Equation (27):

$$r_0\left(\alpha + \beta - \frac{s}{r_0}\right) - R\alpha = 0 \Rightarrow s = r_0(\alpha + \beta) - R\alpha \tag{29}$$

The concentric mosaic spectrum can be derived as:

$$\begin{aligned} L_c(\Omega_\beta, \Omega_\alpha) &= \iint l_c(\beta, \alpha) e^{-j\Omega_\beta \beta - j\Omega_\alpha \alpha} d\beta d\alpha \\ &= \iint l_0(r_0(\alpha + \beta) - R\alpha, \alpha + \beta + \pi) e^{-j\Omega_\beta \beta - j\Omega_\alpha \alpha} d\beta d\alpha \\ &= \iint l_0(s, \theta) e^{-j\Omega_\beta \frac{s + (r_0 - R)(\pi - \theta)}{R} - j\Omega_\alpha \frac{r_0\theta - s - \pi r_0}{R}} \frac{1}{R} ds d\theta \\ &= \frac{1}{R} L_0\left(\frac{\Omega_\beta - \Omega_\alpha}{R}, \frac{r_0\Omega_\alpha - (r_0 - R)\Omega_\beta}{R}\right) e^{-j\Omega_\beta \frac{(r_0 - R)\pi}{R} + j\Omega_\alpha \frac{\pi r_0}{R}} \end{aligned} \tag{30}$$

Again, the spectrum is a rotated version of that of SPF, plus some magnitude change and phase shift. If the scene is Lambertian, or $L_0(\Omega_s, \Omega_\theta) = L_{0s}(\Omega_s)\delta(\Omega_\theta)$, we further have:

$$\begin{aligned} L_c(\Omega_\beta, \Omega_\alpha) &= \frac{1}{R} L_0\left(\frac{\Omega_\beta - \Omega_\alpha}{R}, \frac{r_0\Omega_\alpha - (r_0 - R)\Omega_\beta}{R}\right) e^{-j\Omega_\beta \frac{(r_0 - R)\pi}{R} + j\Omega_\alpha \frac{\pi r_0}{R}} \\ &= \frac{1}{R} L_{0s}\left(\frac{\Omega_\beta - \Omega_\alpha}{R}\right) \delta\left(\frac{r_0\Omega_\alpha - (r_0 - R)\Omega_\beta}{R}\right) e^{-j\Omega_\beta \frac{(r_0 - R)\pi}{R} + j\Omega_\alpha \frac{\pi r_0}{R}} \end{aligned} \tag{31}$$

The spectrum lies on a line that has slope $\dfrac{\Omega_\alpha}{\Omega_\beta} = \dfrac{r_0 - R}{r_0}$. Since $r_0 > R$, the slope is always less than 45 degree.

Figure 15 (a) shows an example scene at constant depth with two sinusoids pasted as texture. In Figure 15 (b) the coordinates α and β are at different scale so the slope does not map to the true value. However, the spectrum in Figure 15 (c) is obviously along a line. If the scale between α and β is the same, the line will achieve maximum slope when the depth of the scene becomes infinite.

VI. Non-rectangular Sampling for IBR

Chan and Shum were the first authors who realized that the plenoptic sampling problem is basically a multidimensional sampling problem [7]. It is well known that the rectangular sampling strategy is not necessary optimal for multidimensional signals [17]. Instead, the generalized periodic sampling theory with arbitrary geometry should be applied. We refer the details of the theory to [17][18]. In [7], the authors mentioned that it is also possible to determine the minimum sampling densities for the quincunx and hexagonal sampling lattices. We have performed some experiments on non-rectangular sampling for IBR. We will briefly go over our experiments and give some comments.

Multidimensional signals can be sampled on a lattice in the spatial domain, which leads to the replication of the Fourier transform on another lattice in the frequency domain. For the lightfield spectrum, take non-Lambertian and no occlusions scenes as an example. The spectrum can be compacted as in Figure 16 (a), which implies quincunx/hexagonal sampling. With this sampling strategy, the sampling efficiency can be improved by a factor of 2 compared to rectangular sampling in Figure 9 (b), which means we only need 50% of the samples. Obviously, the reconstruction filter is a tilted fan-like filter. In theory, we may be able to reconstruct the original IBR signal without losing any information. In practice, however, we cannot design an ideal reconstruction filter without transition band. We choose to reduce the sampling density for the ease of reconstruction. To give a fair comparison between the quincunx/hexagonal sampling (QS) and the rectangular sampling (RS), we let them have the same sampling density. The RS spectrum is shown in Figure 9 (b), while the QS spectrum is shown in Figure 16 (b). We design reconstruction filters for QS with the eigenfilter approach [19][20]. For RS, we use the bilinear interpolation. Please refer to [16] for detailed description on the QS and reconstruction filter design.

We show some experimental results on the two different sampling approaches. In order to have full control on the scenes and the cameras, we choose two scenes rendered from 3D models with texture. These scenes are shown in Figure 17, where scene (a) and (b) are named *Duck* and *Containers*, respectively. Since we focus on 2D lightfield in this paper, we take the center horizontal line to construct the EPIs. Figure 17 (a1) and (b1) are snapshots of the scenes; (a2) and (b2) are their EPIs; (a3) and (b3) are the Fourier transform of the EPIs. Although occlusions can be observed in the scenes, we ignore them in our analysis since the additional modulated components caused by occlusions are not significant in these two examples.

From the Fourier transform of the scenes in Figure 17, we can find the corresponding sampling rate to compact the spectrum to Figure 9 (b) and Figure 16 (b). One thousand random images are then rendered for each scene with the reconstruction filters. These images are also synthesized through the 3D model rendering engine. The difference between the synthesized images and the rendered images is used to measure the quality of the sampling process. In our experiments, PSNR is used to measure such differences. The results are shown in Table I . We observe that for smooth scene as *Duck*, the image quality of RS and bilinear interpolation is actually better than that of QS and its corresponding reconstruction filter. When scenes have more high frequency components as in *Containers*, the latter approach is slightly better than the former one.

The above statement is generally valid and can be justified by looking at how RS and QS replicate the spectrum. In Figure 9 (b) for RS, two neighboring replications have a huge empty space in between when the frequency is low,

but they almost touch when the frequency is high. As a comparison, in Figure 16 (b) for QS, the space in between is kept as a constant. Therefore, RS is more suitable for scenes that have expanded low frequency components and do not have too much high frequency components. This is often true for real world scenes. Even for the *Containers* scene, the improvement by using QS is minor, which cannot justify the increased complexity. Since QS is inconsistent with how the images are taken, the required re-sampling may introduce error. The rendering speed is another concern. Detailed in [16], the reconstruction filter for QS has to do four times more bilinear interpolations than that for RS. Therefore, we conclude that rectangular sampling is preferable to non-rectangular sampling for image-based rendering.

VII. Conclusions and Future Work

In this paper, we presented a new approach to analyzing the IBR spectrum. We first defined the surface plenoptic function, then showed that any IBR representation can be studied if we can find the light ray mapping from the SPF to the IBR representation. We showed examples where the mapping can be done for lightfield and concentric mosaic. Non-Lambertian property, scene depth variations and occlusions fit naturally into our approach and we were able to analyze the effects of them as long as certain assumptions are satisfied. We believe that this approach greatly improves the understanding of IBR spectral analysis.

As we already showed in the paper, the spectrum of the IBR representation is very likely to be band-unlimited. The same result was also drawn in [8]. This may lead to the conclusion that IBR sampling is impossible. On the other hand, effects such as non-Lambertian property, depth variations and occlusions all bring finite spreading of the spectrum after throwing away very small sidebands, and in many cases the amount of spread is predictable. In practice, the first order approximation (spectrum is bounded by the slopes determined by the minimum depth and the maximum depth) in [6] does give good results. When the scenes are highly reflective or the occlusions are very heavy, we showed in section III that the minimum sampling rate has to be increased up to a certain amount.

SPF will be useful in the further understanding of a number of IBR research problems. For example, since the mapping between SPF and the IBR representation is determined by both the scene geometry and the camera path, is there an “optimal” path that we should use for capturing a certain scene? If the camera path is predetermined, and if we know that the SPF has good properties such as band-limitness, it makes sense to capture images non-uniformly such that the effective sampling on the SPF is close to uniform. Work is in progress looking into these opportunities.

Acknowledgement

This work is supported in part by NSF Career Award 9984858 and Industrial Technology Research Institute. We would also like to thank the reviewers who gave us invaluable comments and helped us to improve the quality of the paper.

References

[1] E. H. Adelson, and J. R. Bergen, "The plenoptic function and the elements of early vision", *Computational Models of Visual Processing*, Chapter 1, Edited by Michael Landy and J. Anthony Movshon. The MIT Press, Cambridge, Mass. 1991.

[2] L. McMillan and G. Bishop, "Plenoptic modeling: an image-based rendering system", *Computer Graphics (SIGGRAPH'95)*, pp. 39-46, Aug. 1995.

[3] M. Levoy and P. Hanrahan, "Light field rendering", *Computer Graphics (SIGGRAPH'96)*, pp. 31, Aug. 1996.

[4] S. J. Gortler, R. Grzeszczuk, R. Szeliski and M. F. Cohen, "The Lumigraph", *Computer Graphics (SIGGRAPH'96)*, pp. 43-54, Aug. 1996.

[5] H.Y. Shum and L.-W. He. "Rendering with concentric mosaics", *Computer Graphics (SIGGRAPH'99)*, pp.299-306, Aug. 1999.

[6] J.X. Chai, X. Tong, S.C. Chan and H. Y. Shum, "Plenoptic sampling", *Computer Graphics (SIGGRAPH'00)*, pp.307-318, July 2000.

[7] S. C. Chan and H. Y. Shum, "A Spectral Analysis for Light Field Rendering", *ICIP 2000*.

[8] D. Marchand-Maillet and M. Vetterli, "Sampling Theory for Image-Based Rendering", Master thesis, EPFL, Apr. 2001.

[9] Z. C. Lin and H. Y. Shum, "On the Number of Samples Needed in Light Field Rendering with Constant-Depth Assumption", *Proc. CVPR 2000*.

[10] G. S. P. Miller, S. Rubin and D. Ponceleon, "Lazy Decompression of Surface Light Fields for Precomputed Global Illumination", *Rendering Techniques'98 (Proc. Eurographics Workshop on Rendering)*, pp. 281-292. Springer, 1998.

[11] D. N. Wood, D. I. Azuma, K. Aldinger, B. Curless, T. Duchamp, D. H. Salesin and W. Stuetzle, "Surface Light Fields for 3D Photography", *SIGGRAPH'00*, pp. 287-296, July 2000.

[12] T.-T. Wong, C. W. Fu, P.-A. Heng and C.-S. Leung, "The Plenoptic Illumination Function", *IEEE Trans. on Multimedia*, pp. 361-371, Vol. 4, No. 3, Sep. 2002.

[13] Z.-C. Lin, T.-T. Wong and H.-Y. Shum, "Relighting with the Reflected Irradiance Field: Representation, Sampling and Reconstruction", *International Journal of Computer Vision*, pp. 229-246, Vol. 49, No. 2-3, Sep.-Oct. 2002.

[14] R. Ramamoorthi, P. Hanrahan, "A Signal-Processing Framework for Inverse Rendering", *SIGGRAPH '01*, pp.117-128, Aug. 2001.

[15] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*, Prentice Hall. 1st edition (August 14, 2002.

[16] C. Zhang and T. Chen, "Generalized Plenoptic Sampling", *Carnegie Mellon University Technical Report, AMP01-06*.

[17] D.E. Dudgeon and R.M. Mersereau, *Multidimensional Digital Signal Processing*, Prentice-hall signal processing series, 1984.

[18] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-hall signal processing series, 1993.

[19] P. P. Vaidyanathan and T. Q. Nguyen, "Eigenfilters: A New Approach to Least-Squares FIR Filter Design and Applications Including Nyquist Filters", *IEEE Trans. on CAS*, pp. 11-23, Jan. 1987.

[20] T. Chen, "Unified Eigenfilter Approach: with Applications to Spectral/Spatial Filtering", *IEEE ISCAS*, Chicago, May 1993.

Tables

Table I Rendered images qualities for different sampling methods and reconstruction filters

	Duck	Containers
RS, bilinear interpolation	36.10 dB	21.67 dB
QS, reconstruction filter designed by the eigenfilter approach	36.06 dB	22.21 dB

Figures

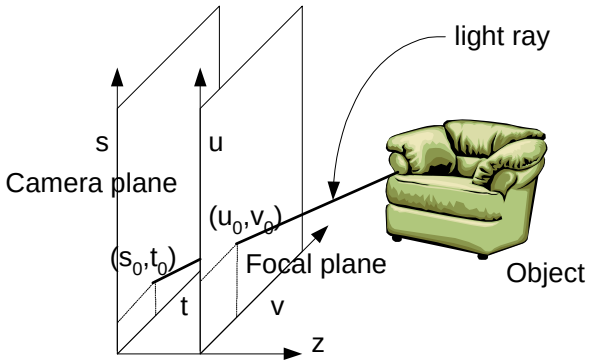

Figure 1 One parameterization of the Lightfield.

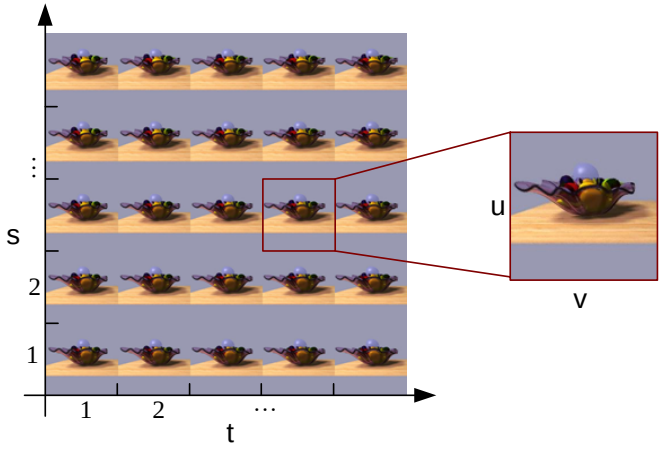

Figure 2 A sample Lightfield image array: fruit plate.

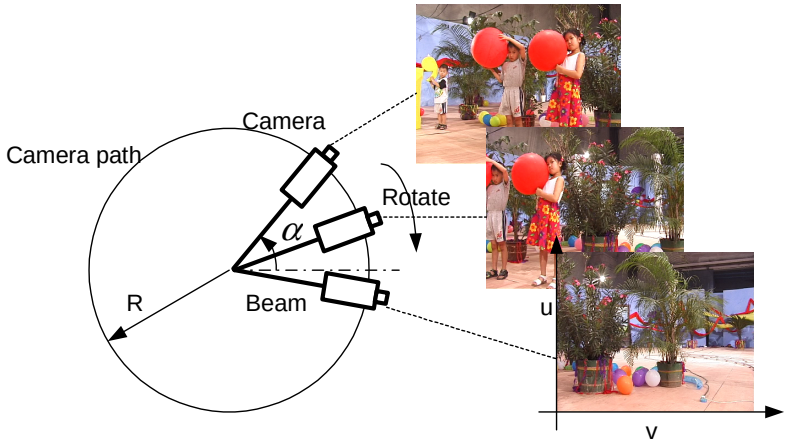

Figure 3 Concentric mosaic capturing.

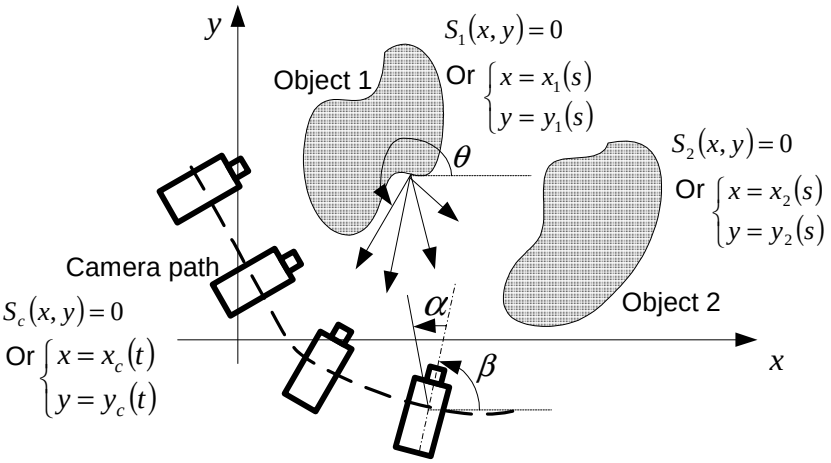

Figure 4 2D surface plenoptic function and general IBR capturing.

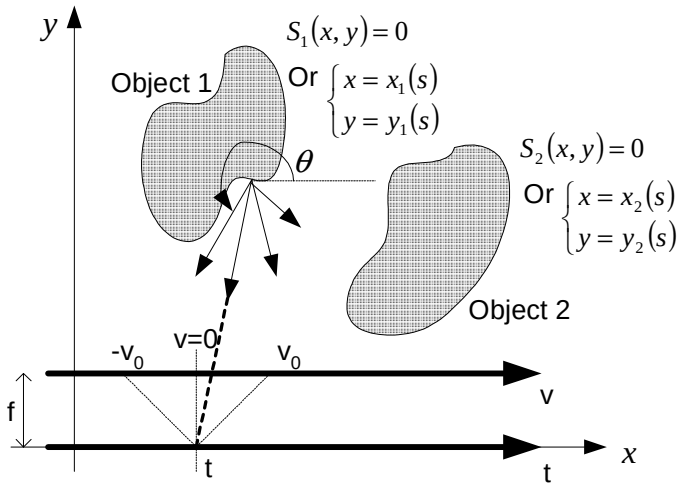

Figure 5 The lightfield parameterization.

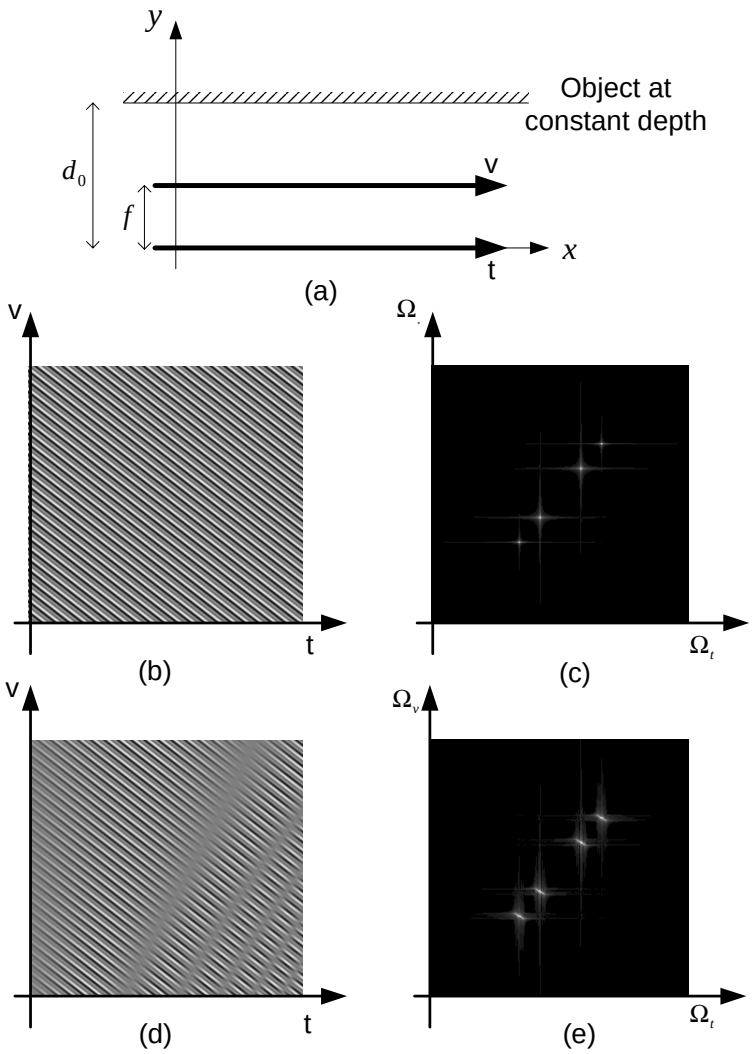

Figure 6 Lambertian and non-Lambertian scenes at constant depth. (a) A scene at constant depth. (b) The EPI of the lightfield when the scene is Lambertian. (c) The Fourier transform of (b). (d) The EPI of the lightfield when the scene is non-Lambertian. (e) The Fourier transform of (d).

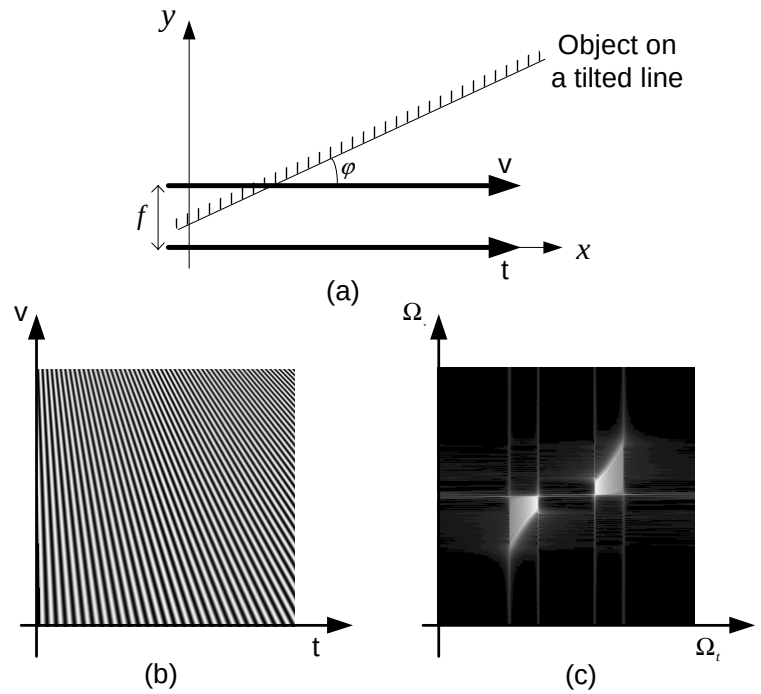

Figure 7 A tilted planar scene with a sinusoid pasted as texture. (a) The scene on a tilted line. (b) The EPI of the lightfield. (c) The Fourier transform of (b).

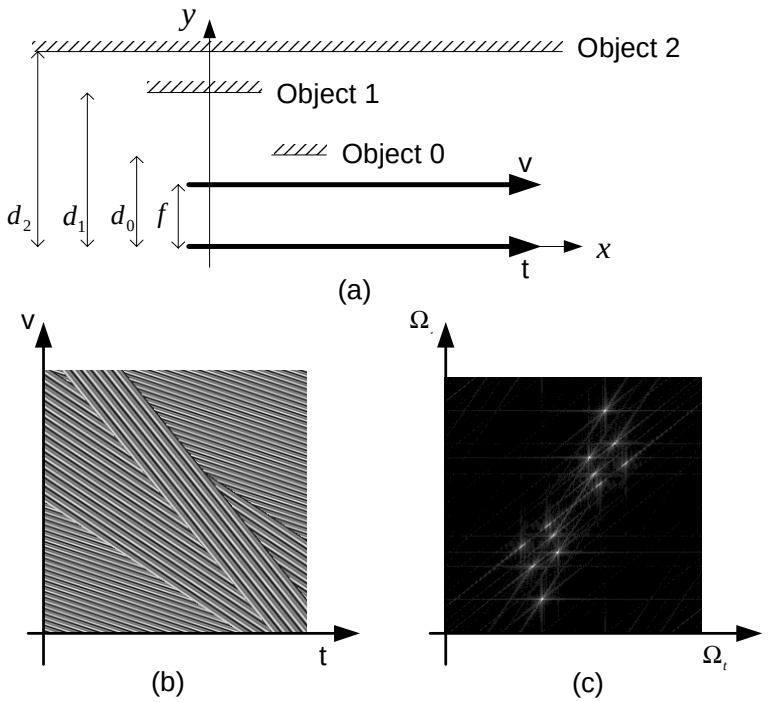

Figure 8 A three-layer scene with sinusoids pasted as texture. (a) Three objects on different constant depths. (b) The EPI of the lightfield. (c) The Fourier transform of (b).

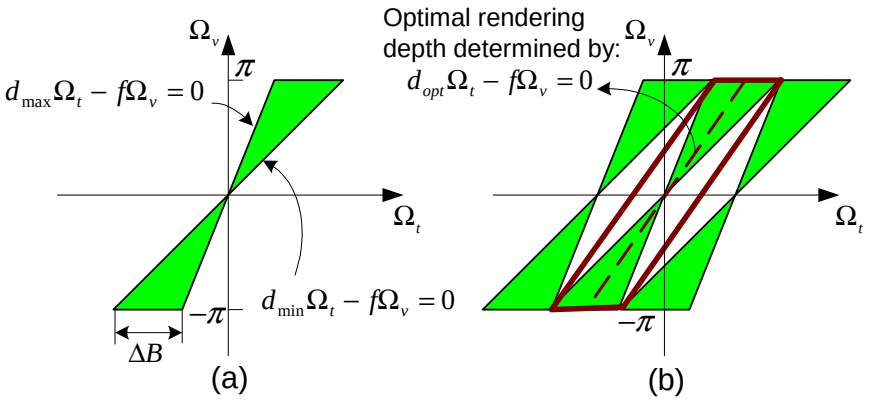

Figure 9 Frequency response of a no occlusion EPI and the "optimal" sampling in [6]. (a) Frequency support of a scene with no occlusion and Lambertian model. (b) The "Optimal" compacting and the reconstruction filter in [6].

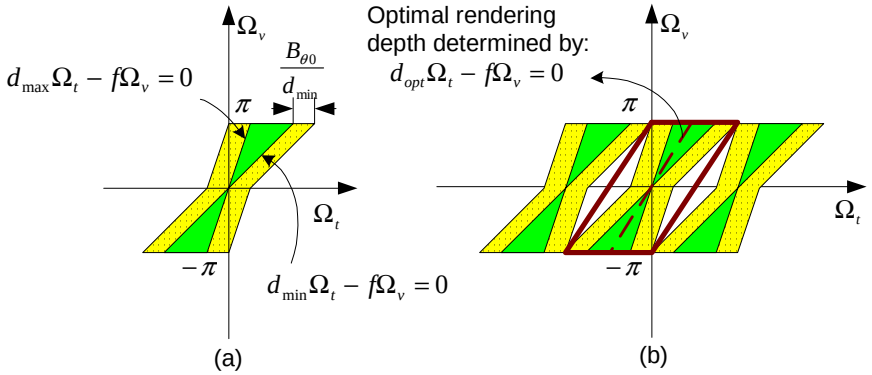

Figure 10 Optimal sampling when the scene is non-Lambertian. (a) The spectrum is expanded from the Lambertian case. (b) Optimal compacting of the spectrum for rectangular sampling.

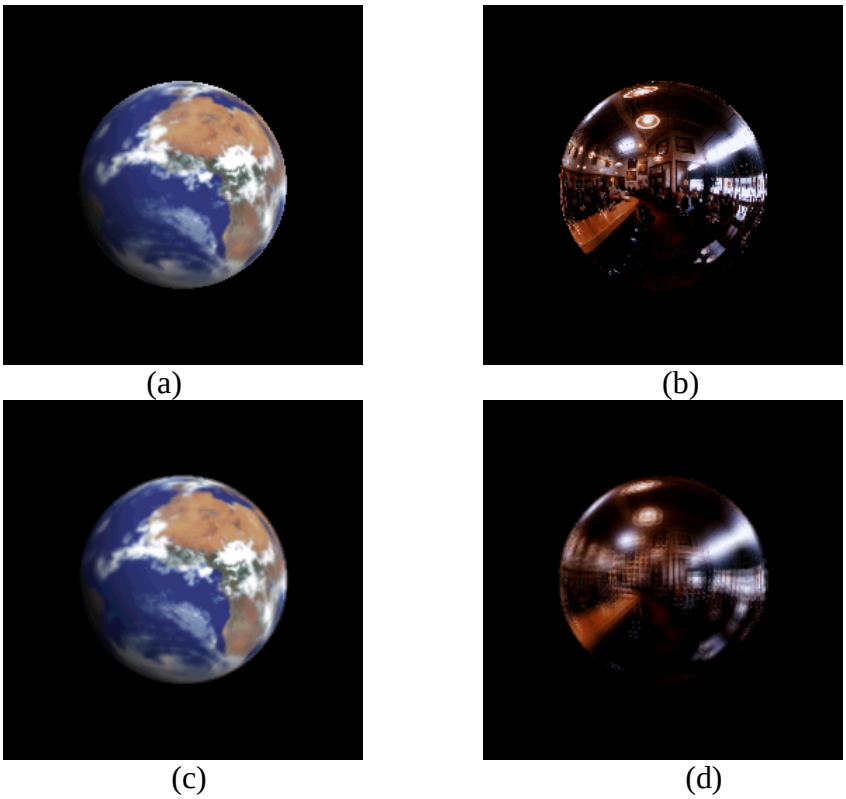

Figure 11 An example to show the minimum sampling rate for non-Lambertian surface. (a) A Lambertian sphere. (b) A purely reflective sphere. (c) At a certain sampling rate, we are able to reconstruct the scene very well for the Lambertian sphere. (d) At the same sampling rate, the reconstruction is very bad for the reflective surface.

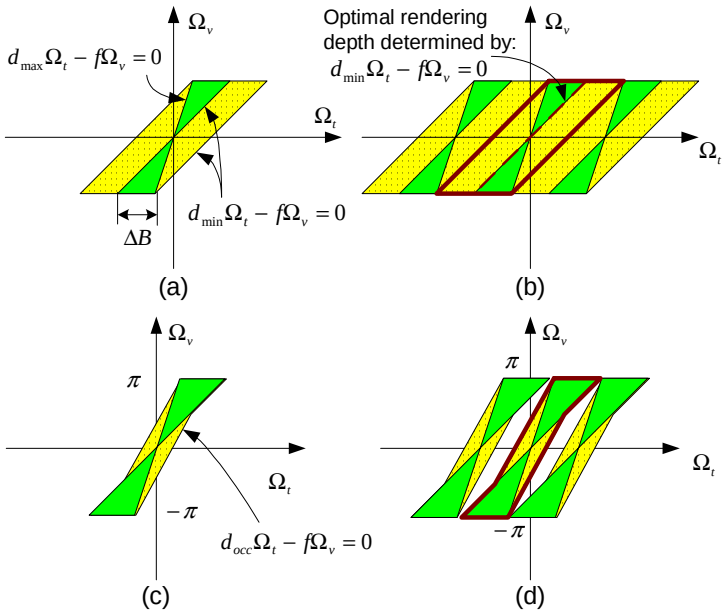

Figure 12 Optimal sampling when there are occlusions. (a) The most conservative estimation of the spectrum when the object at $d_{\min}$ causes the major occlusions (b) The optimal rectangular sampling strategy for compacting (a). (c) The spectrum given that the object at d_{occ} is the main source of occlusions. (d) The optimal rectangular sampling strategy for compacting (c).

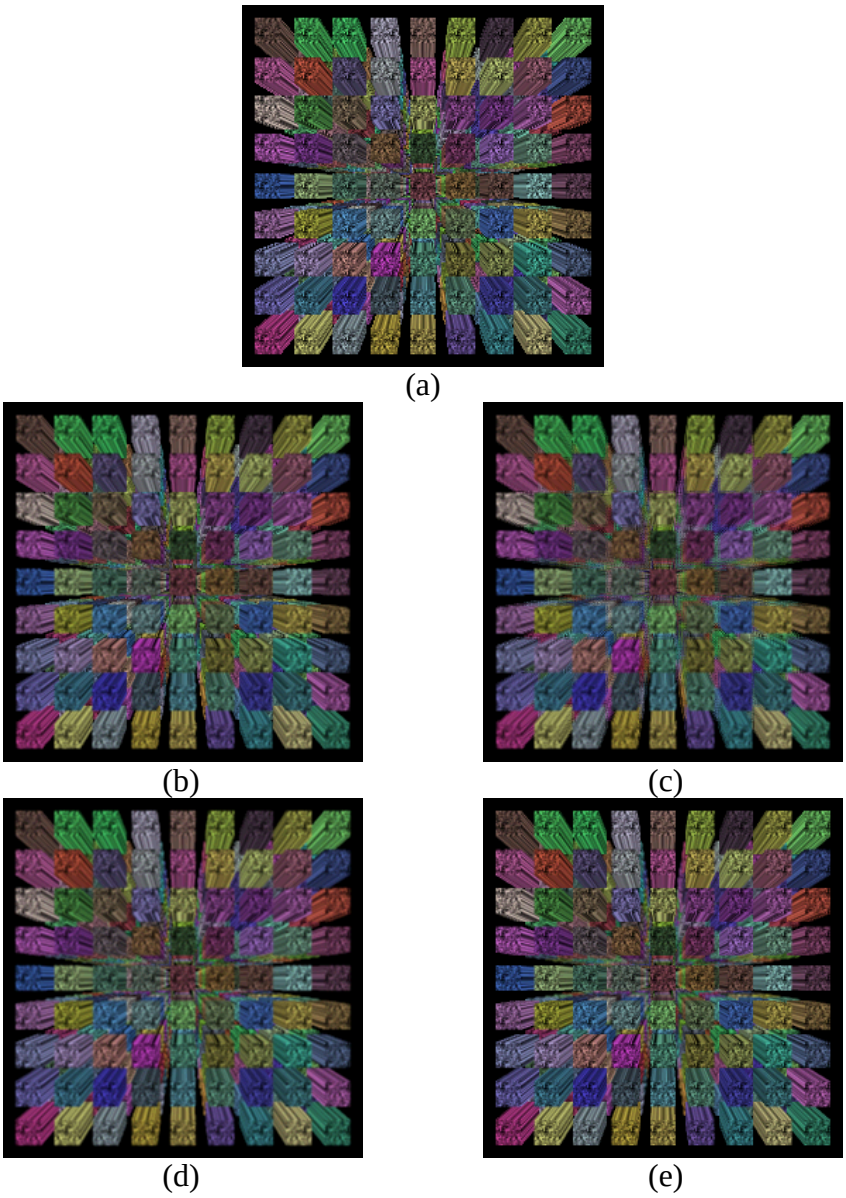

Figure 13 An example to show the minimum sampling rate under heavy occlusion.

(a) The heavily occluded scene. (b) Use the minimum sampling rate in [6], rendered with d_{opt}. (c) Same as (b) but rendered with $d_{\min}$. (d) Use twice the sampling rate in (b), rendered with d_{opt}. (e) Same as (d) but rendered with $d_{\min}$.

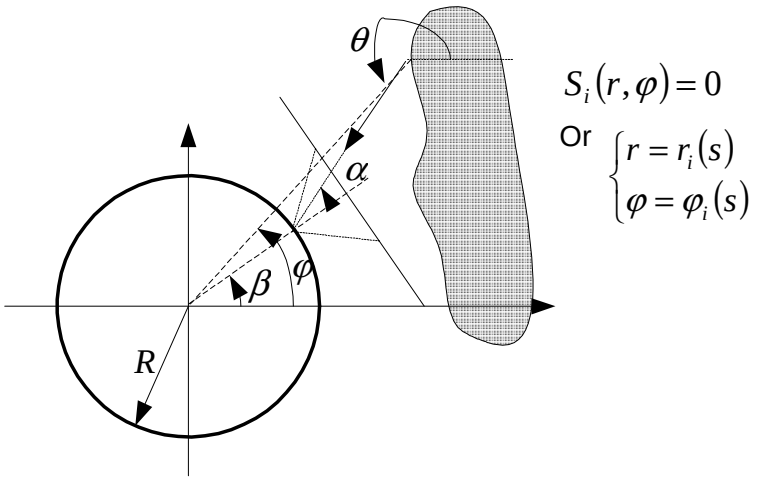

Figure 14 The concentric mosaic parameterization.

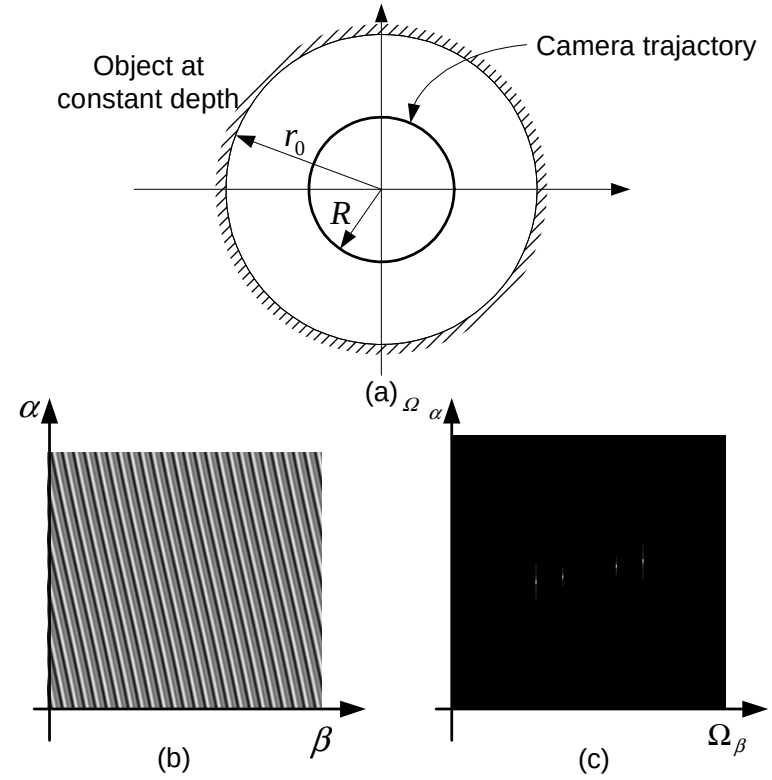

Figure 15 A constant depth concentric mosaic scene with sinusoids pasted as texture. (a) A scene at constant depth. (b) The EPI of the concentric mosaic when the scene is Lambertian. (c) The Fourier transform of (b).

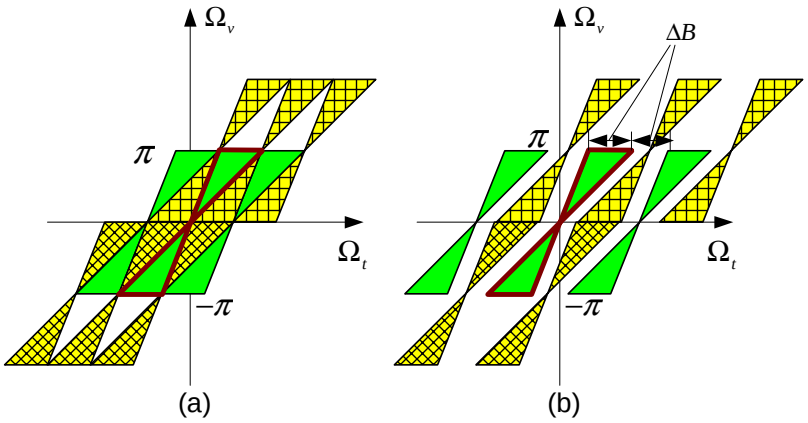

Figure 16 Quincunx/hexagonal sampling for IBR. (a) The most compacting parallelepiped sampling. (b) The quincunx/hexagonal sampling that has the same sampling rate as the optimal rectangular sampling.

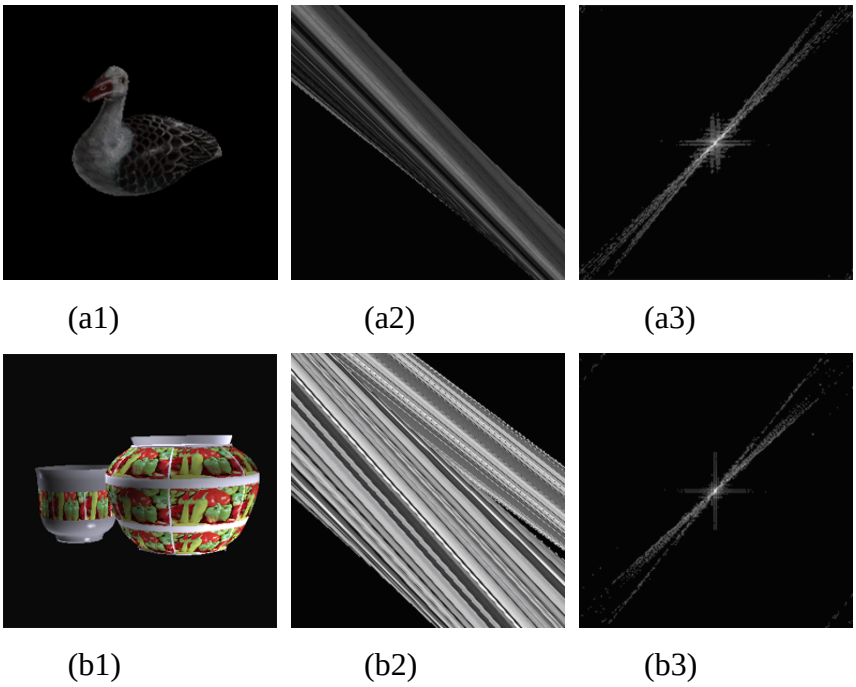

Figure 17 The testing scene EPIs and their Fourier transforms. (a) The scene *Duck*. (b) The scene *Containers*.