

# Frequency Estimation of the Plenoptic Function Using the Autocorrelation Theorem

Changjian Zhu<sup>1</sup>, Li Yu<sup>2</sup>, *Member, IEEE*, Zengqiang Yan, and Sen Xiang, *Student Member, IEEE*

**Abstract**—The frequency estimation of the plenoptic function (POF) is an important research topic in spectral analysis for determining the minimum sampling rate of image-based rendering. In this paper, we mathematically derive a frequency estimation function of the POF using the autocorrelation theory. The autocorrelation function (ACF) of the POF is studied along both the spatial and image plane frequency axes. The influence of the scene's complexity and depth on the ACF of the POF is analyzed. Furthermore, we study the frequency estimation error to analyze the performance of the method. Existing techniques typically use Fourier transformation to determine the frequency of the POF. The technique presented herein simply starts from measurements of the light field in a finite number of positions, and the information from the light field is analyzed in the spatial domain. Finally, experimental results are presented to demonstrate that the proposed method can effectively estimate the frequency of the POF. The results are shown to be in good agreement with traditional methods.

**Index Terms**—autocorrelation function (ACF), frequency estimation, image-based rendering, plenoptic function (POF), sampling.

## I. INTRODUCTION

THE technique of image-based rendering (IBR) has been extensively applied to three-dimensional television (3DTV) [1], free viewpoint TV [2], and immersive communication systems [3]. Generally, this technique consists of image capture (i.e., the sampling of the IBR) and novel view synthesis. The sampling problem is one of the most basic and fascinating topics in IBR. To investigate the sampling problem, we can determine the minimum number of capturing cameras that are needed while achieving an anti-aliased IBR [4].

Manuscript received November 24, 2015; revised May 23, 2016 and November 15, 2016; accepted February 9, 2017. Date of publication February 17, 2017; date of current version November 6, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61231010, in part by the 863 High-Tech Research and Development Program under Grant 2015AA015901 and Grant 2015AA015903, and in part by the science and technology research projects of Guangxi universities under Grant KY2015YB367. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Neel Joshi.

C. Zhu and L. Yu are with the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: changjianzhu@hust.edu.cn, hustlyu@hust.edu.cn).

Z. Yan is with the Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong (e-mail: zyanad@connect.ust.hk).

S. Xiang is with the School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China (e-mail: xiangsen@wust.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCI.2017.2671450

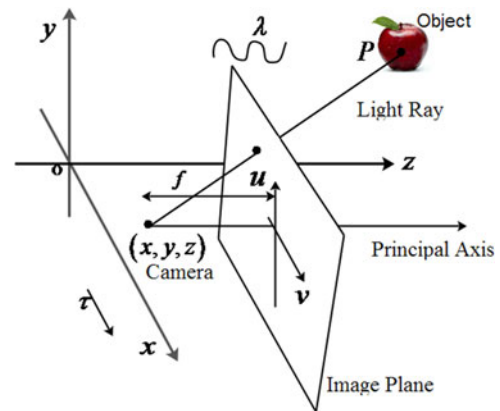


Fig. 1. Diagram of the 7-D POF and pinhole camera model [8], [9]. The camera position is represented using  $(x, y, z)$ , and the image plane is  $(u, v)$ . The distance that separates the camera center and the image plane is known as the focal length  $f$ . A light ray, whose wavelength is  $\lambda$ , from point  $P$  intersects the image plane and the camera center at a time  $\tau$ .

The answers to the sampling problem of the IBR and related questions are addressed by multi-view acquisition and rendering [5], [6]. The acquired view information is dependent on the viewing position  $(x, y, z)$ , viewing direction  $(u, v)$ , specific wavelength  $\lambda$  and time  $\tau$ . Generally, the plenoptic function (POF), which was introduced by Adelson and Bergen [7], is applied to describe the multi-view scheme in three-dimensional (3-D) space. A more precise diagram of the POF is illustrated in Fig. 1 [7], [8]. All parameters of the acquired view are gathered into a seven-dimensional (7-D) function  $\Phi(x, y, z, u, v, \lambda, \tau)$ . Therefore, the minimum sampling rate of the IBR can be determined using the results of the POF sampling [9]–[12].

The most well-known approach uses a spectral analysis of the POF to determine the minimum sampling rate [5], [6]. In this approach, an important issue is estimating the frequency of the POF [9], [12]. The concept of frequency estimation is initially mentioned in communication signal processing [13]–[18]. Two principal categories of methods are employed for frequency estimation, i.e., frequency domain methods [13] and time domain methods [14]–[18]. Generally, frequency domain methods can be applied to effectively estimate the frequency of the POF, such as studies that employ the spectrum of the POF [9], [10] and the bandwidth of the POF [11], [12]. However, a mathematical framework has not been fully developed for studying the sampling problems in the POF. Additionally, time domain methods can be applied if a rapid and accurate estimation is required as a high priority in the presence of noise interference. Therefore,

to acquire a new approach to accurately determine the spectrum support of the POF for complex situations (e.g., existing occlusion) and, inspired by the works of [15]–[18], we try to study using the time domain methods to estimate the frequency of the POF.

In the time domain methods, one of the most effective approaches to estimate the frequency of signals is the autocorrelation function (ACF) [15], [16]. This approach is employed because the ACF of a periodic signal is a periodic function, and its cycle is the same as the periodic signal based on the Wiener-Khinchin theorem [19], [20]. The advantages of the autocorrelation theorem are its low degree of complexity and high-accuracy. Considering that the POF is a periodic function [9], we can derive the ACF of the POF to estimate the maximum frequency of the POF. The frequency estimation error can be evaluated to analyze the frequency estimation performance of the ACF. Our analysis reveals that the ACF will be more complicated if the variations in the scene geometry become more complex or if the depth of the scene increases. Given this characterization, the spectral support of the POF is bounded by the maximum frequency. Based on the spectral support, an essential bandwidth of the plenoptic spectrum is determined and applied to determine the minimum sampling rate of the POF. Our framework is based on Chai and Gilliam *et al.* [9]–[12] and extends it in several ways. These theoretical results can be verified in experiments.

Our proposed framework has the following new insights:

1) In contrast to the current state-of-the-art methodologies, we are the first researchers to address the sampling problem of the POF in the spatial domain (for the ACF of the POF; using the spatial domain instead of the time domain due to the plenoptic sampling along the spatial axis).

2) We simplify the complexity of the frequency estimation which is the most important research result from the spectral analysis of the POF. For example, the calculation of the ACF is performed in the spatial domain, and the frequency of the POF is directly obtained by the phase of the ACF.

3) Our proposed method is more accurate than the existing methods in the literature. For the complex scene, such as an existing occlusion, the estimated accuracy is higher when the autocorrelation theorem is applied to perform the frequency estimation of the POF.

## II. RELATED WORK

The sampling problem of the IBR was first discussed in [21]. Additionally, this problem was analyzed by Shum *et al.* [22], Gortler *et al.* [23] and Levoy *et al.* [24]. The first analysis of plenoptic sampling for use in spectral analysis was published by Chai *et al.* [9]. The minimum sampling rate of the uniform plenoptic sampling for the IBR was determined using spectral analysis. Based on an assumption of a truncated/piecewise constant depth, the authors noted that if the scene surface was Lambertian without occlusions, the spectral support of the POF was bounded by the minimum depth and the maximum depth. The minimum sampling rate of the IBR was obtained using the characteristics of the spectral support of the POF. In [10], the effect

of the non-Lambertian and occluded scenes on the spectral support of the plenoptic signal were investigated. Zhang *et al.* noted that the depth variations and occlusions, as well as the spectrum of the POF, were broadened when the scene surface included non-Lambertian reflections. In [11], N. Do *et al.* investigated the bandwidth of the POF. The results of the characterization of the POF bandwidth were obtained using a representation of the light field by its spatial spectrum. The results indicated that the continuous plenoptic spectrum was not bandlimited unless the scene surface was flat. Additionally, the spectral support was related to the maximum surface slope and the maximum frequency of the painted texture signal. Similar to the bandlimited nature of the POF, an exact closed-form expression of the plenoptic spectrum was presented in [12]. The results of the plenoptic spectrum of a slanted plane with a finite width were obtained. Additionally, the reconstruction of the POF aspects, such as the reconstruction filter, was discussed. Based on the spectrum of the POF, in [25], Stewart *et al.* presented a linear, and spatially invariant reconstruction filter that reduces ghosting artifacts of under-sampled light fields. Using this reconstruction filter, the number of necessary camera-plane samples for light field rendering can be substantially reduced, and aliasing in the light field reconstruction can be avoided.

Unstructured light field [26] aspects were discussed. Recently, Chen and Schonfeld in [27] investigated the plenoptic sampling theory of unstructured camera systems, under which cameras can be located at arbitrary locations and orientations. The authors extended the conclusion of plenoptic sampling [9] to study unstructured cameras with a limited rotation and field of view (FOV). Additionally, spherical four-dimensional (4-D) light field sampling was presented in [28]. Using this sampling scheme, a spherical light field without aliasing can be reconstructed.

In contrast to the previously discussed methods, other methods have taken different approaches to plenoptic sampling. Based on the causality requirement in scale-space theory, Lin *et al.* investigated the lower bound for the number of samples required in light field/Lumigraph rendering in [29]. Then, a closed-form solution of the lower bounds that avoided horizontal and vertical double images was determined. In [30], Liang *et al.* presented a unified mathematical framework to represent the light fields. This framework consists of three different parameterizations (e.g., plane-sphere, two-plane, and in-camera). Based on this framework, they modeled the image formation process in the frequency domain using the spectral analysis. The results obtained by this framework quantitatively match the results of the traditional models. The term “layer-based modeling” was originated in [31]. Some studies [32] employed the concept of layered depth to analyze video view interpolation. In [33], Pearson *et al.* used plenoptic layer-based modeling to analyze the minimum sampling rate problem of the IBR. Using the layer-based method, a plenoptic sampling analysis can be reduced to the necessary geometric information of a 3-D scene. Similarly, depending on the depth of the scene, Chaurasia *et al.* presented a new approach in [34] that used introductory silhouette aware variation warping to compensate for incorrect or incomplete geometric information. Based on this approach, the

TABLE I  
NOTATIONS AND DEFINITIONS

$t, r$	The camera position on the camera plane
$u, v$	The pixel location on the image plane
$\theta$	The angle between a light ray and principal axis of the camera
$f$	The focal length of the camera
$s$	The curvilinear surface coordinate
$S$	The scene surface function
$\Psi(s)$	The function of $x$ with respect to $s$
$v_m$	The maximum value of a FOV
$\Delta t$	The spacing between the cameras along the spatial axis
$\Delta v$	The spacing between the pixels on the image plane
$\omega_s$	The maximum frequency of a texture signal
$\omega_t$	The indexing of the spatial frequency in the $t$ -axis
$\omega_v$	The indexing of the image plane frequency in the $v$ -axis
$\Omega_t$	The maximum of $\omega_t$
$\Omega_v$	The maximum of $\omega_v$
$\hat{\Omega}_t$	The estimation of $\Omega_t$
$\hat{\Omega}_v$	The estimation of $\Omega_v$
$D_t$	The estimation variance of $\Omega_t$
$D_v$	The estimation variance of $\Omega_v$
$T$	The maximum of $x$
$B$	The maximum of $s$
$N_a$	The number of cameras
$M_a$	The number of pixels
$p(t, v)$	The EPI function, i.e., the 2-D light field
$R(\cdot)$	The ACF of the EPI function $p(t, v)$
$l(s, \theta)$	The intensity of a light ray emitted from $s$ on the scene surface in direction $\theta$

number of images that are necessary for alias-free rendering can be reduced in the IBR.

### III. PARAMETERIZATION

For clarity, we list important notations and associated definitions that are utilized throughout this paper in Table I.

#### A. Problem Statement

The high dimensionality POF is illustrated in Fig. 1. It can be observed that it is important to address the seven dimensions of the POF. To overcome this obstacle, we can use certain solutions to reduce the dimensionality of the POF. For example, by assuming that the scene is static and the wavelength is fixed, McMillan *et al.* [35] used a set of discrete samples to define a continuous five-dimensional (5-D) POF. Based on the 5-D POF and an assumption that the radiance of a light ray is constant along its path, Gortler *et al.* [23] and Levoy *et al.* [24] used a 4-D light field/Lumigraph to represent each light ray. Specifically, each light ray is parameterized by two intersections with two parallel planes. One of the planes is termed the camera plane, and its coordinates are  $(t, r)$ . The other plane is termed the image plane, and its coordinates are  $(u, v)$ . The separation of the two parallel planes is equivalent to the focal length  $f$ . Considering these definitions of the light field, the function  $p(t, r, u, v)$  is defined to represent the intensity of a light ray at a camera location  $(t, r)$  and the pixel location  $(u, v)$ .

For the simplicity of exposition, as in [9]–[12], a two-dimensional (2-D) version of the light field,  $p(t, v)$ , was considered by keeping  $u$  and  $r$  constant. After determining the sampling results for the  $(t, v)$  dimensions, the results can be

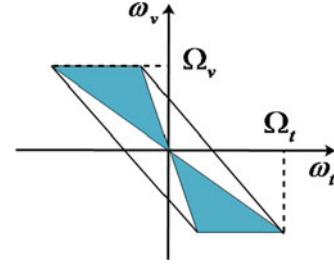


Fig. 2. Diagram that illustrates the parametric essential bandwidth of the POF in reference to [9] and [12].

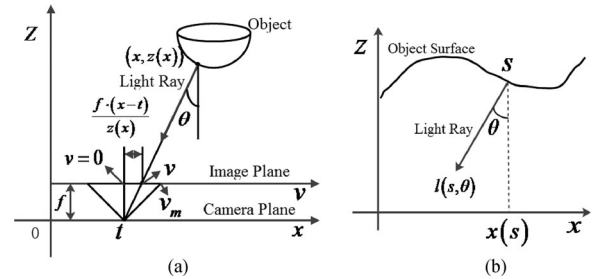


Fig. 3. (a) An illustration of the 2-D light field  $p(t, v)$ , in which the intensity of a light ray is defined by its captured position of the camera at  $t$  and the viewing direction  $v$  on the image plane. In addition,  $p(t, v)$  also describes the relationship between the camera and the scene's geometry  $z(x)$ . Diagram (b) illustrates the light ray representation using the SPF.

easily extended to the  $(r, u)$  dimensions. An illustration of the 2-D light field  $(t, v)$  is provided in Fig. 3(a).

As previously mentioned for the spectral analysis of the POF, the spectral support is bounded by the maximum frequency in all of the frequency axes. Based on the light field parameterizations, a diagram of the spectral support of the POF is shown in Fig. 2, where  $\omega_t$  is the spatial frequency in the  $t$ -axis, and  $\Omega_t$  is its maximum value.  $\omega_v$  denotes the image plane frequency in the  $v$ -axis, and  $\Omega_v$  is its maximum value. The spectral support is primarily bounded by the maximum spatial frequency  $\Omega_t$  and the maximum image plane frequency  $\Omega_v$ . Therefore, we are interested in estimating  $\Omega_t$  and  $\Omega_v$ .

Now, the frequency estimation problem of the POF is stated as follows.

**Definition 1:** For the given 2-D light field  $p(t, v)$ , its maximum frequencies (i.e.,  $\Omega_t$  and  $\Omega_v$  in Fig. 2) in the  $\omega_t$ -axis and the  $\omega_v$ -axis can be calculated as

$$\Omega_t = \Phi_t(p(t, v)), \quad (1)$$

$$\Omega_v = \Phi_v(p(t, v)), \quad (2)$$

where  $\Phi_t(\cdot)$  and  $\Phi_v(\cdot)$  are the frequency calculation functions, which are dependent on the approach of the frequency estimation.

In the subsequent sections, we demonstrate that this frequency estimation causes angle variations in the POF, which can be efficiently solved using a modified ACF algorithm [15], [16].

### B. Standard Parameterization and Notation

To construct the ACF of the POF, the 2-D light field  $p(t, v)$  is generally represented by the concept of the epipolar plane image (EPI) [36]. Using the EPI, a visual representation of the 2-D light field can be constructed if we plot its intensity at all possible  $(t, v)$  coordinates. Here, we can also use  $p(t, v)$  to denote the EPI function. Then, we can define the mapping expression of the EPI function  $p(t, v)$ .

**Definition 2:** Given the camera position  $t$ , the image plane coordinate  $v$ , and the depth  $z(x)$  of a scene with respect to  $x$  as shown in Fig. 3(a), the relationships between the camera and the image plane and the scene object for the EPI of a light ray along the  $x$ -axis can be expressed as

$$v = f \tan(\theta), \quad (3)$$

$$t = x - z(x) \frac{v}{f}. \quad (4)$$

In (3),  $\theta$  is the angle between the light ray and the principal axis of the camera, which is also known as the light ray direction.

These formulas for light ray parameterization are also represented by the surface plenoptic function (SPF) [10]–[12], [37], [38]. In this report, we employ the representation of the SPF in [12]. The standard light ray representation of the SPF involves the origination point of the light ray emission on the scene surface  $S$ , and the direction of the light ray. Generally, we use the curvilinear surface coordinate  $(s)$  to parameterize the scene surface in 2-D space. The variables  $[x, z]^T$  are defined to represent the Cartesian coordinates of a point [39]. Then, we use the function  $S(s) = [x(s), z(s)]^T$  to represent the position of a point on the scene surface in the 2-D space. Therefore, all parameters of the SPF are gathered into a single function  $l(s, \theta)$ . Using this function, the intensity of a light ray emitted from a point  $(s)$  on a 2-D surface at the viewing direction  $\theta$  can be parameterized. In Fig. 3(b), we demonstrate an example of the SPF in 2-D space. Using the SPF, the scene surface is described as

$$S = \begin{cases} z(x) = g(x), x \in [0, T], \\ x(s) = \Psi(s), s \in [0, B], \end{cases} \quad (5)$$

where  $\Psi(s)$  is a function of  $x$  with respect to  $s$  and is dependent on the scene geometry, and  $g(x)$  is a function of  $z$  with respect to  $x$ .  $T$  is the maximum value of  $x$  for the scene surface coordinates, and  $B$  is the maximum value of  $s$ . Then, we substitute (5) into (4) to obtain a new mapping expression for the EPI as

$$t = \Psi(s) - z(\Psi(s)) \frac{v}{f} = \Psi(s) - z(\Psi(s)) \tan(\theta). \quad (6)$$

Additionally, we assume that the object surface has no self-occlusions in the FOV of the cameras. In this case, each light ray  $(t, v)$  within the FOV can intersect with maximum of one point on the object surface  $z(x)$  [11], which indicates that  $t$  in (3), (4) and (6) is a strictly monotonic function of  $x$ . Thus, the scene object function  $z(x)$  is constrained by

$$|z'(x)| < \frac{f}{v_m}, \quad (7)$$

where  $v_m$  is the maximum FOV. Here, we assume that the FOV of the camera is finite (i.e.,  $-v_m \leq v \leq v_m$ ) and that the scene object is completely covered by the camera.  $z'(x)$  is the derivative of the scene surface  $z$  with respect to coordinate  $x$ . Using (5) and (6), the EPI expression is also represented by  $l(s, \theta)$  as

$$l(s, \theta) = p(\Psi(s) - z(\Psi(s)) \tan(\theta), f \tan(\theta)). \quad (8)$$

Furthermore, let  $\theta_m$  be the maximum  $\theta$ , which can be represented by  $v_m$  as  $\theta_m = \arctan(v_m/f)$ .

Let  $\tau_1$  be the indexing of autocorrelation along the camera position's  $t$ -axis, and  $\tau_2$  be the indexing of autocorrelation along the image pixel's  $v$ -axis. The exception is any function that carries the superscript “\*”, which denotes a conjugate operation.

### IV. ACF DEFINITION OF THE POF

To estimate the frequency of the POF, the ACF must be defined using the EPI. To study the properties of the ACF, the ACF of the POF must be analyzed from every possible camera position in the 3-D space. For simplicity and based on the 2-D light field, we only show that the ACF is a 2-D function of the POF. Based on the ACF definition in [15]–[18], we define the ACF of the 2-D light field as the following definition.

**Definition 3:** Given a 2-D light field  $p(t, v)$ , its ACF can be defined as

$$\begin{aligned} R(\tau_1, \tau_2) &= E[p(t, v) \cdot p^*(t + \tau_1, v + \tau_2)] \\ &= \int_{-\infty}^{\infty} \int_{-v_m}^{v_m} p(t, v) \cdot p^*(t + \tau_1, v + \tau_2) dt dv, \end{aligned} \quad (9)$$

where  $E[\cdot]$  represents the expectation operation.

As shown in (9), calculating the ACF is dependent on the expression of the EPI. In particular, the expression for the EPI is associated with the scene complexity (e.g., geometrical and texture information), capturing camera, and output resolution [9]. Therefore, the scene complexity can be simplified to expediently analyze the corresponding characteristic of the ACF.

Because the relationship between  $l(s, \theta)$  and  $p(t, v)$  is described by (8), the integral variables in (9) are substituted using (8). First, it is necessary to explain the relationship between  $p^*(t + \tau_1, v + \tau_2)$  in (9) and the SPF using (8). According to (3) and (6),  $v + \tau_2$  and  $t + \tau_1$  can be represented by (10) and (11), respectively, as

$$v + \tau_2 = f \tan(\theta + \tau_\theta), \quad (10)$$

$$t + \tau_1 = \Psi(s + \tau_s) - z(\Psi(s + \tau_s)) \tan(\theta + \tau_\theta), \quad (11)$$

where  $\tau_s$  is the indexing for the corresponding  $\tau_1$  along the  $s$ -axis, and  $\tau_\theta$  is the indexing for the corresponding  $\tau_2$  along the  $\theta$ -axis. Substituting (8), (10) and (11) into (9), we obtain a new ACF expression for the 2-D light field. Consequently, the



expression can be written as

$$R(\tau_1, \tau_2) = R(\tau_s, \tau_\theta) = f \int_{-\infty}^{\infty} \int_{-\theta_m}^{\theta_m} l(s, \theta) \cdot l^*(s + \tau_s, \theta + \tau_\theta) \Psi'(s) (1 - z'(\Psi(s)) \tan(\theta)) \sec^2 \theta ds, \quad (12)$$

where  $z'(\Psi(s))$  is the derivative of  $z$  with respect to  $\Psi(s)$ , and  $\Psi'(s)$  is the derivative of  $\Psi$  with respect to  $s$ . Note that the ACF of the POF is dependent on the  $z(\Psi(s))$  and  $\Psi(s)$ , which are the depth of the scene and the geometry information of the scene, respectively.

## V. AUTOCORRELATION ANALYSIS WITH SURFACE MODEL

In the previous section, we presented the ACF derivation of the POF. This derivation required the scene surface to be restricted by a no-occlusion constraint. However, the ACF expression of the POF can not be completely obtained because the EPI expression is not given in the Lambertian surface of the scene and the geometry information is unknown. Additionally, we determine that calculating the autocorrelation in (12) can be simplified in the following manner by making certain assumptions.

### A. ACF With the Scene Surface

To study the ACF with the scene surface, the scene can be assumed to be a Lambertian surface to simplify the EPI expression. In the Lambertian reflectance, the surface's luminance is isotropic, and the luminous intensity obeys Lambert's cosine law. Therefore, we restrict our analysis to the Lambertian surface, i.e.,  $l(s, \theta) = l(s)$ . In this case, the ACF of the POF along the  $\tau_2$ -axis is a constant, i.e.,  $R(\tau_s, \tau_\theta) = R(\tau_s)$ . Thus, the ACF expression for (12) can be simplified as

$$R(\tau_s) = f \int_{-\infty}^{\infty} \int_{-\theta_m}^{\theta_m} l(s) \cdot l^*(s + \tau_s) \Psi'(s) (1 - z'(\Psi(s)) \tan(\theta)) \sec^2 \theta ds d\theta. \quad (13)$$

To describe the relationship between the ACF  $R(\tau_s)$  and the scene's information, an intermediate variable  $H(s)$  can be defined to represent the ACF as

$$\begin{aligned} H(s) &= \Psi'(s) l(s) \cdot l^*(s + \tau_s) \\ &\int_{-\theta_m}^{\theta_m} (1 - z'(\Psi(s)) \tan(\theta)) \sec^2 \theta d\theta \\ &= 2\Psi'(s) l(s) \cdot l^*(s + \tau_s) \tan(\theta_m). \end{aligned} \quad (14)$$

By applying (14), the following general equation for the ACF of the POF, which is independent of the scene's information, can be obtained as

$$\begin{aligned} R(\tau_s) &= f \int_{-\infty}^{\infty} H(s) ds \\ &= 2f \tan(\theta_m) \int_{-\infty}^{\infty} \Psi'(s) l(s) \cdot l^*(s + \tau_s) ds. \end{aligned} \quad (15)$$

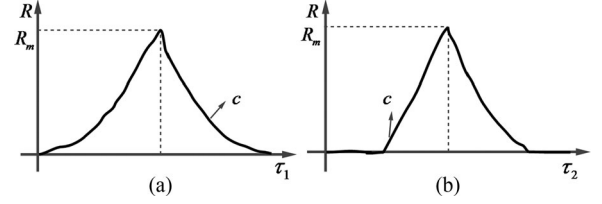


Fig. 4. The ACF support of a 2-D light field. (a) An illustration of ACF along the  $\tau_1$ -axis. Diagram (b) illustrates the ACF along the  $\tau_2$ -axis.

Equation (15) is the ACF expression when the scene surfaces are Lambertian and do not include self-occlusions. Note that  $s$  is not calculated using (15); because  $l(s)$  is unknown. Here,  $l(s)$  is related to the texture information of the scene's surface.

### B. ACF With Texture Information

To analyze the influence of the texture information on the ACF of the POF, we can define an expression for the texture information of the scene surface. Here, the expression of the texture signal is defined in reference to [10], [11]. If the scene geometry information is given, then the texture signal pasted to its surface can be considered as a sinusoid signal, e.g.,  $l(s) = \sin(\omega_s s)$ , where  $\omega_s$  is the frequency of the texture signal. By substituting the texture signal into (15), we can obtain the following expression for the ACF of the POF as

$$R(\tau_s) = 2f \tan(\theta_m) \int_{-\infty}^{\infty} \Psi'(s) \sin(\omega_s s) \cdot \cos(\omega_s (s + \tau_s)) ds. \quad (16)$$

As shown in (16),  $R(\tau_s)$  contains  $s$ ,  $\theta_m$ , and the texture frequency  $\omega_s$ . Additionally,  $s$  relates to the surface geometry shape (e.g., curvatures)  $\Psi(s)$ , and  $\theta_m$  is related to the FOV of the camera. Therefore, the ACF of the POF is also dependent on the surface geometry shape and the FOV of the camera. The surface geometry shape is generally unknown. However, if the surface geometry has been given, the ACF expression can be completely calculated using (16). The influence of the geometry information on the plenoptic sampling is discussed in [11] and [33]. We will not detail its derivation in this paper. In (16), the phase of  $R(\tau_s)$  contains accurate frequency information about the POF, such as  $\omega_s s$ ,  $\omega_s (s + \tau_s)$  and  $\tan(\theta_m)$ . Therefore, we attempt to study using the ACF phase to estimate the frequency of the POF.

### C. The ACF Support Analysis

From the previous derivation, the ACF support is bounded by the maximum value  $R_m$  of the ACF and the curvature  $c$  as shown in Fig. 4. In particular,  $R_m$  is primarily determined by the minimum depth and the maximum depth and  $\Psi(s)$  and  $\theta_m$ . The curvature  $c$  is primarily determined by  $\omega_s$  and  $\Psi(s)$ . Note that,  $s$  relates to the scene geometry shape and width (i.e.,  $B$  in (5)). Therefore, the scene width affects the  $R_m$  and the curvature  $c$ . Based on the ACF support and (16), we use the ACF phase

to estimate the frequency of the POF and give its calculation function in the following theorem.

*Theorem 1:* The frequencies of the POF can be respectively calculated by its ACF phase

$$\hat{\Omega}_t = \hat{\Phi}_t (\arg \{R(\tau_1, \cdot)\}), \quad (17)$$

$$\hat{\Omega}_v = \hat{\Phi}_v (\arg \{R(\cdot, \tau_2)\}), \quad (18)$$

where  $\hat{\Phi}_t(\cdot)$  and  $\hat{\Phi}_v(\cdot)$  are the frequency calculation functions with respect to the phase of  $R(\tau_1, \tau_2)$ . Additionally,  $\hat{\Omega}_t$  is the estimation of  $\Omega_t$ , and  $\hat{\Omega}_v$  is the estimation of  $\Omega_v$ .

*Proof:* Refer to Appendix A. ■

## VI. FREQUENCY ESTIMATION THEOREM OF THE POF

### A. Frequency Estimation Theorem of the 2-D Light Field

As mentioned in theorem 1, the frequency of the POF can be determined using the ACF phase. Consider that the ACF of the POF can be given by (9) at the particular spatial frequency  $\omega_t$  and the particular image plane frequency  $\omega_v$ . To analyze the ACF of the POF, the ACF is typically calculated using a finite number of discrete samples [15]. We present the frequency estimator of the POF from a set of discrete samples in the following theorem.

*Theorem 2:* Given the discrete 2-D light field signal  $p(n, m)$ , its spatial frequency  $\hat{\Omega}_t$  and image plane frequency  $\hat{\Omega}_v$  can be respectively calculated as

$$\hat{\Omega}_t = \arg \max \left| \sum_{n=1}^{N_a} p(n, m) e^{-j\Omega_t n} \right|^2, \quad (19)$$

$$\hat{\Omega}_v = \arg \max \left| \sum_{m=1}^{M_a} p(n, m) e^{-j\Omega_v m} \right|^2, \quad (20)$$

where  $N_a$  denotes the number of cameras, and  $M_a$  denotes the number of pixels.

*Proof:* The full proof is presented in Appendix B. A brief outline of the proof is given as follows. First, the continuous  $p(t, v)$  is sampled using the camera spacing  $\Delta t$  along the  $t$ -axis and the pixel spacing  $\Delta v$  along the  $v$ -axis. Using a rectangular sampling lattice [9], the discrete 2-D light field  $p(n, m)$  is represented by

$$p(n, m) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} p(t, v) \delta(t - n\Delta t) \delta(v - m\Delta v). \quad (21)$$

As assumed earlier for a finite number of cameras and a finite camera FOV, we let  $1 \leq n \leq N_a$  and  $1 \leq m \leq M_a$ . By the method of maximum likelihood estimation (MLE) discussed in [15] with the frequency estimation of a sinusoid, the spatial frequency  $\hat{\Omega}_t$  and the image plane frequency  $\hat{\Omega}_v$  of a 2-D light field signal can be solved as given in (19) and (20). ■

From (19) and (20), both  $\hat{\Omega}_t$  and  $\hat{\Omega}_v$  are related to the 2-D light field  $p(n, m)$ . Therefore, we can employ the properties of the 2-D light field to estimate the frequency of the POF. The frequency estimator of the POF in this paper is obtained from examining the likelihood equation of the POF. In the discrete 2-D light field  $p(n, m)$ , the expression for  $R(\tau_1, \tau_2)$  can be

modified as

$$R(i, k) = \sum_{n=i+1}^{N_a} \sum_{m=k+1}^{M_a} p(n, m) \cdot p^*(n - i, m - k), \quad (22)$$

where  $1 \leq i \leq N_a$  and  $1 \leq k \leq M_a$ .

We present the frequency estimator of the POF by examining the likelihood equation in the following theorem.

*Theorem 3:* The maximum spatial frequency  $\hat{\Omega}_t$  and the maximum image plane frequency  $\hat{\Omega}_v$  can be estimated using the ACF of the 2-D light field as

$$\hat{\Omega}_t \approx \sum_{i=1}^{N_a} \frac{i}{\pi(N_a + 1)} \arg \{R(i, k)\}, \quad (23)$$

$$\hat{\Omega}_v \approx \sum_{k=1}^{M_a} \frac{2k}{M_a + 1} \arg \{R(i, k)\}. \quad (24)$$

*Proof:* We use the maximum periodogram of the POF to solve the frequency estimation problem over the likelihood equation of the POF. First, let the derivatives of the POF periodogram equal zero. After appropriately grouping terms, the likelihood equations of the POF can be given as

$$\text{Im} \left( \sum_{i=1}^{N_a} i \cdot R(i, k) e^{-j\Omega_t i} \right) = 0, \quad (25)$$

$$\text{Im} \left( \sum_{k=1}^{M_a} k \cdot R(i, k) e^{-j\Omega_v k} \right) = 0. \quad (26)$$

By (25) and (26), the likelihood equations of the POF are primarily dependent on the ACF  $R(i, k)$ . A practical estimator is produced by approximating the likelihood equations (25) and (26). Additionally, we consider that  $\arg(E(R(i, \cdot))) = \Omega_t i$  and  $\arg(E(R(\cdot, k))) = \Omega_v k$ . For large  $N_a$  or  $M_a$ , the following functions are obtained

$$\begin{aligned} \text{Im}(R(i, k) e^{-\Omega_t i}) &\approx A \sin(\arg(R(i, k)) - \Omega_t i) \\ &\approx A(\arg(R(i, k)) - \Omega_t i), \end{aligned} \quad (27)$$

$$\begin{aligned} \text{Im}(R(i, k) e^{-\Omega_v k}) &\approx A \sin(\arg(R(i, k)) - \Omega_v k) \\ &\approx A(\arg(R(i, k)) - \Omega_v k). \end{aligned} \quad (28)$$

Then, (25) is heuristically approximated for large  $N_a$  as

$$\sum_{i=1}^{N_a} i \cdot (\arg(R(i, k)) - i \cdot \hat{\Omega}_t) \approx 0. \quad (29)$$

Similarly, (26) is heuristically approximated for large  $M_a$  as

$$\sum_{k=1}^{M_a} k \cdot (\arg(R(i, k)) - k \cdot \hat{\Omega}_v) \approx 0. \quad (30)$$

Algebra reduces (29) and (30), and then we can solve the POF's frequency estimation problem as given in (23) and (24). This frequency estimator weights the argument of the sample autocorrelation function by the lag,  $(i, k)$ , as suggested in (25) and (26). ■

From theorem 3, the frequency estimator weights primarily contain the argument of the ACF of the POF by the lag,  $R(i, k)$ ,  $N_a$  and  $M_a$ , as suggested in (23) and (24). In (23) and (24), we observe that they can be alternately viewed as the least-squares fit of the estimated phase of the ACF versus the lag to the true autocorrelation phase. In particular,  $\omega_v$  will be a constant based on (13) when the object surface is assumed to have a Lambertian reflectance. Thus,  $\Omega_v$  is also a constant. Based on this analysis, we discover that the frequency of the POF can be determined using the ACF of the discrete POF.

### B. Performance Analysis of the Frequency Estimation

In this sub-section, we analyze the frequency estimation performance. We use the concept of variance in [18] to describe the performance of the frequency estimation method for the POF. Let  $D_t$  be the variance of  $\Omega_t$  and let  $D_v$  be the variance of  $\Omega_v$ . We define the variance of the frequency estimation of the POF.

**Definition 4:** For a given  $\hat{\Omega}_t$  and  $\hat{\Omega}_v$ , their variances for the frequency estimation can be defined by (31) and (32), respectively, as

$$D_t = E \left[ \left( \hat{\Omega}_t - \Omega_t \right) \right]^2, \quad (31)$$

$$D_v = E \left[ \left( \hat{\Omega}_v - \Omega_v \right) \right]^2, \quad (32)$$

where  $E[\cdot]^2$  represents the calculation of the variance.

For the actual frequencies, note that both  $\Omega_t$  and  $\Omega_v$  are difficult to obtain. Therefore,  $D_t$  and  $D_v$  cannot be calculated using (30) and (31), respectively. However, from the expression of  $R(i, k)$ , we can determine that the phase of  $R(i, k)$  in the  $i$  direction can be written as

$$\arg \{R(i, k)\} = \tan^{-1} \left\{ \frac{v_Q(i, k)}{2N_a - i + v_I(i, k)} \right\}, \quad (33)$$

where

$$v_Q(i, k) = \text{Re} \{R(i, k)\}, \quad (34)$$

$$v_I(i, k) = \text{Im} \{R(i, k)\}. \quad (35)$$

Beginning with the ACF variance that is used to analyze the performance of the frequency estimation, the phase variance of the ACF in the  $\tau_1$  direction can be calculated as

$$V_t(i) = \text{var} \left\{ \frac{v_Q(i, k)}{2N_a - i + v_I(i, k)} \right\}, \quad (36)$$

where  $\text{var} \{\cdot\}$  is the variance operation. Similarly, the phase variance of the ACF in the  $\tau_2$  direction can be given as follows

$$V_v(k) = \text{var} \left\{ \frac{v_Q(i, k)}{2M_a - k + v_I(i, k)} \right\}. \quad (37)$$

Note that the phase variance of the ACF is primarily dependent on the maximum error of the ACF phase. Particularly, the phase variance of the ACF is  $\text{var} \{\arg [R(i, k)]\}$ . However, the ACF is not independent; it is related to the change in  $(i, k)$ . Therefore, the variances of  $\Omega_t$  and  $\Omega_v$  cannot be directly calculated using the variance of  $\arg [R(i, k)]$ . By combining the relationship between the frequency and the ACF of the POF, the

expressions for the calculations of  $D_t$  and  $D_v$  can be written as follows

$$D_t = \text{var} \left( \hat{\Omega}_t \right) = \left[ \frac{1}{\pi(N_a + 1)} \right]^2 \times \sum_{i=1}^{N_a} 2V_t(i), \quad (38)$$

$$D_v = \text{var} \left( \hat{\Omega}_v \right) = \left[ \frac{2}{M_a + 1} \right]^2 \times \sum_{k=1}^{M_a} 2V_v(k). \quad (39)$$

We can easily demonstrate that (38) and (39) achieve their minimum values for a minimum phase variance of the ACF. In the last iteration, the estimation errors are combined to form the original value and then they are applied to the frequency estimation.

### C. Analysis of the Complexity of the Frequency Estimation

To reflect the advantages of the ACF algorithm, the computational complexity is analyzed using the comparison with the method of fast Fourier transform (FFT) of the discrete POF according to the expression for the plenoptic spectrum in [9]–[12]. The discrete Fourier transform of the POF is given as

$$X(i, k) = \sum_{n=1}^{N_a} \sum_{m=1}^{M_a} p(n, m) e^{-j \frac{2\pi}{N_a} i \cdot n} e^{-j \frac{2\pi}{M_a} k \cdot m}, \quad (40)$$

$$1 \leq i \leq N_a, 1 \leq k \leq M_a.$$

Combining (22)–(24) and (40), the computational complexity can be obtained as shown in Table II. By comparison, the amount of computation in the ACF algorithm is less than the amount of computation in the FFT algorithm for the frequency estimation of the POF. For example, the addition/subtraction of the ACF algorithm is  $N_a^2 + M_a^2 - 2$  times, and the addition/subtraction of the FFT algorithm is  $2(N_a \log_2 N_a + M_a \log_2 M_a)$  times. Additionally, the  $\cos(\cdot)/\sin(\cdot)$  of the FFT is  $N_a \log_2 N_a + M_a \log_2 M_a$  times, and the  $\cos(\cdot)/\sin(\cdot)$  of the ACF is 0 times. This reason can be explained by considering that the ACF algorithm is not dependent on the rotation factor. Additionally, the frequency estimation is directly performed in the time domain, and does not need to transform the frequency domain.

### D. Sampling Theorem of the POF

After the frequency of the POF has been estimated using (23) and (24), the effect of sampling on the spatial axis of the POF are presented in this subsection. This representation considers the results of the frequency estimation in the previous section. Sampling along the spatial axis is performed by considering a margin that uses the properties of the frequency estimation of the POF by the ACF. The sampling of the POF is detailed in Fig. 5. In this figure, a group of cameras are uniformly placed along the camera plane  $(t, r)$  [9]. The variable  $\Delta t$  is the camera spacing along the spatial axis. In this paper, we discuss the sampling of the POF along the  $t$ -axis. After the sampling rate along the  $t$ -axis is determined, the results can be easily extended to the  $r$ -axis.

The schematic top view of the spectrum of the POF is shown in Fig. 6(a). If the POF is sampled along the spatial axis with

TABLE II  
COMPARISON OF COMPUTATIONAL COMPLEXITY FOR THE FREQUENCY ESTIMATORS OF THE POF

Methods	Add/Subtract	Multiplication	$\cos()$ or $\sin()$
FFT	$2(N_a \log_2 N_a + M_a \log_2 M_a)$	$N_a \log_2 N_a + M_a \log_2 M_a$	$N_a \log_2 N_a + M_a \log_2 M_a$
ACF	$N_a^2 + M_a^2 - 2$	$N_a^2 + M_a^2 + N_a + M_a$	0

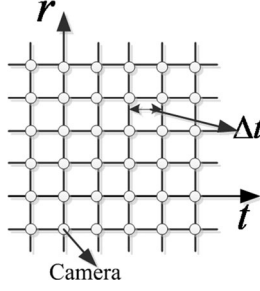


Fig. 5. Rectangular sampling grid for plenoptic sampling. The cameras are uniformly placed on the plane  $(t, r)$ .

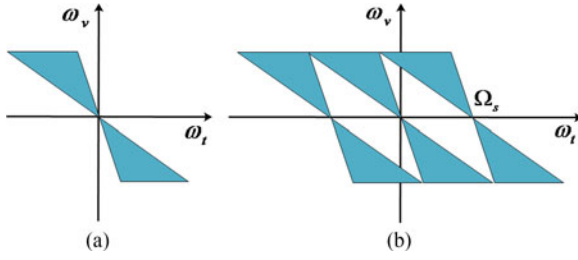


Fig. 6. Spectrum of the POF. (a) Top view of the plenoptic spectrum; and (b) top view of the plenoptic spectrum with its repetitions due to the spatial sampling.

a spatial sampling frequency of  $\Omega_s$  that is too small, then a further spectrum aliasing will occur, as indicated in Fig. 6(b). According to Shannon's uniform sampling theorem [41], the relationship between the sampling angular frequency  $\Omega_s$  and the camera spacing can be given as  $\Omega_s = 2\pi/\Delta t$ . To avoid aliasing in the reconstruction of the POF, the camera spacing  $\Delta t$  (i.e., the sampling rate of the IBR) along the spatial axis is given by [12]

$$\Delta t = \frac{2\pi z_a}{\hat{\Omega}_t (z_{\max} - z_{\min})} \quad (41)$$

where  $z_{\max}$  is the maximum depth of the scene,  $z_{\min}$  is the minimum depth of the scene, and  $z_a$  is the average depth of the scene. Additionally,  $\hat{\Omega}_t$  is estimated using (23). Therefore, the sampling rate of the POF can be obtained from the above analysis if the frequency of the POF by the ACF is known.

### E. Reconstruction by Interpolation

Finally, after determining the sampling rate, the POF can be reconstructed using the image samples. Generally, to reconstruct the POF at any location, the captured images at each point of the sampling grid must be known, and typical interpolation

techniques [26], [42] are applied. The interpolation filter to be used is dependent on the sampling grid along the spatial and image plane axes. When the samples have been obtained by rectangular sampling, the interpolation filters are two low-pass filters with support  $[-\frac{\pi}{\Delta v}, \frac{\pi}{\Delta v}]$ ,  $[-\Omega_s/2, \Omega_s/2]$ . Furthermore, the reconstruction filter is created by setting a skewed filter [11], [12]. The ideal filters can be calculated in closed forms; however, these ideal interpolation filters are not achievable in practice. Therefore, the spatial sampling frequency must be slightly increased to consider the change in the real filter. The standard techniques in [11] and [12] can be used for their design.

## VII. EXPERIMENTAL RESULTS

In this section, the experimental results for the frequency estimation of the POF using the ACF are presented. These results are compared with other frequency estimation methods for the POF sampling. In the experiments, all the scenes are rendered using 3ds Max. A wall, cube-annulus, cone, teapot, bunny and dragon [43], as illustrated in Fig. 7(a)–(f), are employed to perform the experiments. The surface complexity of the different scenes is different. All scenes are fixed at the point  $(0, -400, 4000)$  mm. In all cases, each scene is captured by a set of cameras, which are uniformly placed along the line  $L = \{(x, 0, 0)\}$  and point in the same direction. Let  $x \in [-1000, 1000]$  mm, and the images are captured every 10 mm along the straight line. The camera resolution is  $240 \times 320$  pixels.

### A. Simulation Results of the ACF

First, the simulation results of the ACF of the POF are presented to intuitively describe the characteristics of the ACF. Based on the captured images, the EPI volumes can be obtained as indicated in Fig. 7(g)–(l). The results of the ACF operation are shown in Fig. 8. The ACF will increase as  $i$  increases, and the ACF of the POF will attain its maximum value when  $i$  is  $\frac{N_a}{2}$ . Subsequently, the ACF will decrease as  $i$  increases. For the  $k$ -axis, the ACF exhibits the same phenomenon. The ACF of the POF will be at its maximum when  $k$  is  $\frac{M_a}{2}$ . As depicted in Fig. 8(c), when  $i = 100$ , the ACF is at its maximum. As shown in Fig. 8(d), when  $k = 160$ , the ACF is at its maximum. Therefore, when  $i = \frac{N_a}{2}$  and  $k = \frac{M_a}{2}$ , the POF is completely autocorrelated, as indicated in Fig. 8(a) and (b). Without a loss in generality, we select the maximum ACF of the POF to calculate the ACF phase and obtain the frequency of the POF.

The surface in Fig. 8(b) is rougher than the surface in Fig. 8(a) because the surface geometry of the dragon is more complicated than that of the bunny. Moreover, Fig. 8(c) and (d) show that the ACFs of the cube-annulus, cone, and teapot scenes are rougher, whereas the ACF of the wall is smoother. Therefore, if the



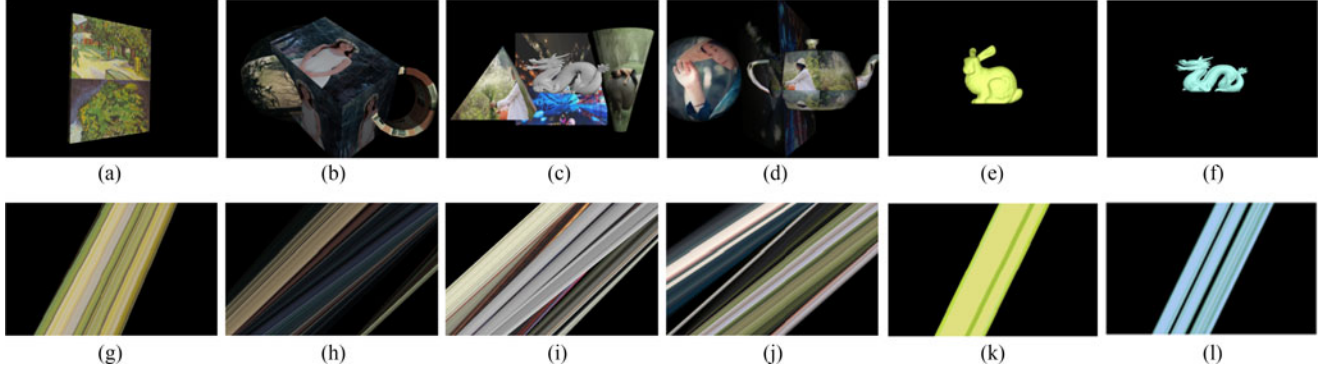


Fig. 7. Diagrams of the six scenes and their corresponding EPI volumes. Diagrams (a)–(f) illustrate the wall, cube-annulus, cone, teapot, bunny and dragon scenes. Diagrams (g)–(l) depict the EPI volumes for the corresponding scenes.

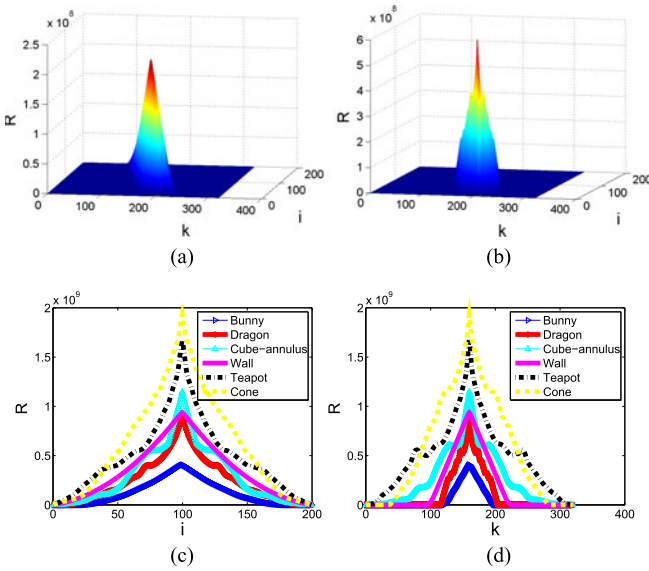


Fig. 8. ACF of the 2-D light field. Diagram (a) illustrates the ACF of the bunny scene in 3-D plots. Diagram (b) illustrates the ACF of the dragon scene in 3-D plots. Diagram (c) illustrates the ACF of the six scenes with the  $i$ -axis in 2-D plots. Diagram (d) illustrates the ACF of the six scenes with the  $k$ -axis in the 2-D plots.

surface geometry becomes complicated, the ACF of the POF will also become complicated. Another phenomenon is that the ACF of the bunny is smaller than the ACF of the dragon. Additionally, the ACF of the cube-annulus is smaller than the ACFs of the cone and the teapot because the scene complexity (e.g., geometrical and textural information) of the cube-annulus is less than that of the cone and the teapot scenes. Moreover, we determine that the ACF of the cone is at a maximum because its width is the maximum width. Thus, the ACF of the POF also depends on the scene width. Finally, if the minimum and maximum depths of the scene are fixed and the same between the different scenes, the ACF will increase as either the scene complexity increases or as the scene width increases.

### B. Experimental Results of the Frequency Estimation

The frequency estimation of the POF is simulated using the bunny and dragon scenes depicted in Fig. 7(e) and (f), respectively. For simulation purposes, a dense set of frequency

TABLE III  
THE ESTIMATED FREQUENCY AND VARIANCES BY OUR PROPOSED METHOD FOR THE BUNNY AND DRAGON SCENES

Scenes	$\hat{\Omega}_t$ (Hz)	$\hat{\Omega}_v$ (Hz)	$D_t$ (Hz)	$D_v$ (Hz)
Bunny	177.42	912.97	0.05	1.81e-005
Dragon	181.45	750.84	0.23	8.24e-005

estimations of a scene can be derived, and the proposed frequency estimation results for the ACF of the POF can be evaluated. First, based on the ACF of the POF, the frequency of the POF is estimated using (23) and (24). In addition, the estimation error is also calculated using (38) and (39). Then, by comparing the results of the reconstruction experiments with that of the theoretical calculations, the correctness of the frequencies estimated by the ACF of the POF can be confirmed.

Now, the estimation results and frequency variances are provided in Table III. It can be observed that the spatial frequency  $\hat{\Omega}_t$  of the bunny scene is 177.42 Hz, whereas the spatial frequency  $\hat{\Omega}_t$  of the dragon scene is 181.45 Hz. Then, the maximum spacing for the bunny scene, which is calculated by (41) is 5.56 mm. Furthermore, the maximum spacing for the dragon is 5.42 mm. To verify the correctness of the above results, we simultaneously reconstruct an EPI volume using a set of captured images. We analyze the relationship between the quality of the reconstruction EPI-volumes and the number of the captured images. Additionally, the relationship can be applied to determine the maximum spacing between the cameras.

For the reconstruction experiments, each EPI volume consists of 700 reconstructed images and is rendered using different numbers of the captured images. Different numbers of captured images in the range of [10, 610] are tested, and the number is increased by increments of 50 to analyze the relationship between the rendering quality and the number of captured images. The reconstruction of the EPI volumes is depicted in Fig. 9. It should be note that using 10 cameras causes significant aliasing in the reconstruction of the EPI volumes of the bunny and dragon scenes, as shown in Fig. 9(a) and (b), respectively. The aliasing will be reduced as the number of captured images increases.

For the bunny scene, when the number of captured images is 260, minimal aliasing occurs. For the dragon scene, when the

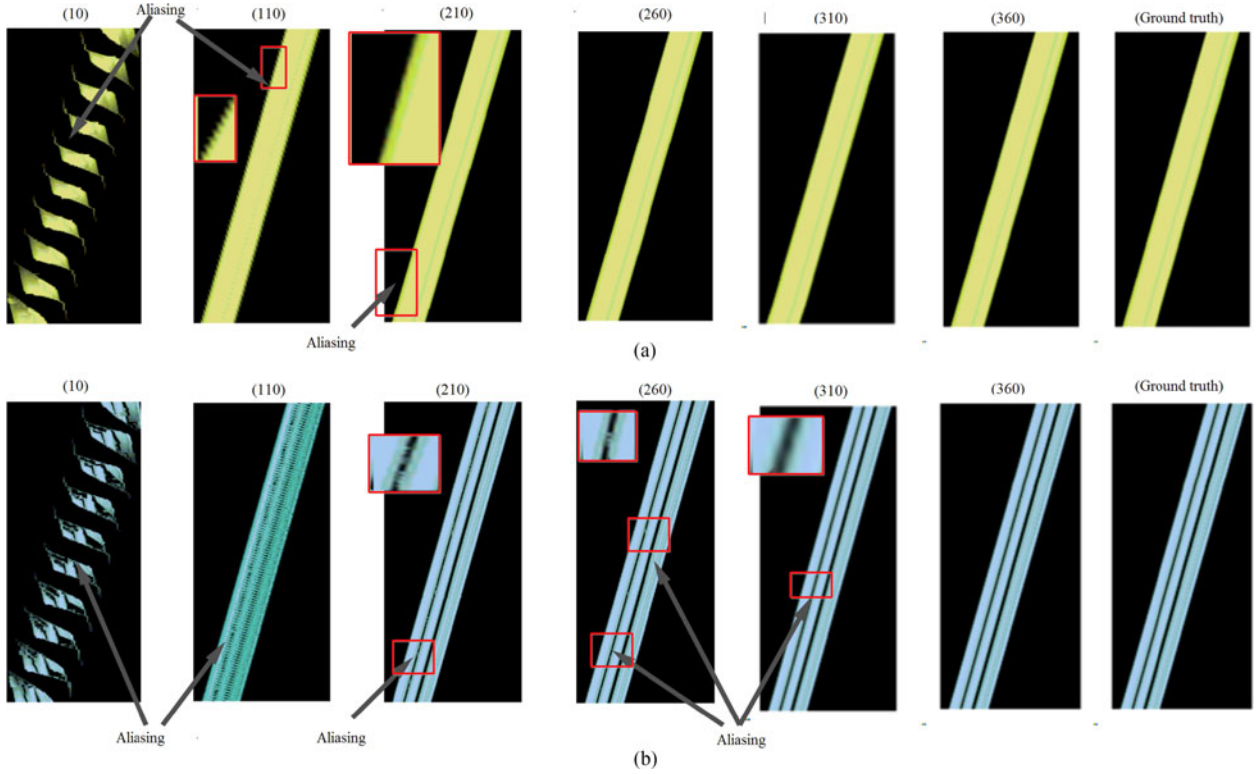


Fig. 9. Diagrams that depict the reconstruction of the synthetic EPI volumes for the bunny and the dragon. When the number of cameras is too small, aliasing will occur. The aliasing will be reduced as the number of cameras increases. Diagram (a) illustrates the reconstruction of the EPI volumes for the bunny using a different number of cameras, such as 10, 110, 210, 310 and 360. Diagram (b) illustrates the reconstruction of the EPI volumes for the dragon using a different number of cameras.

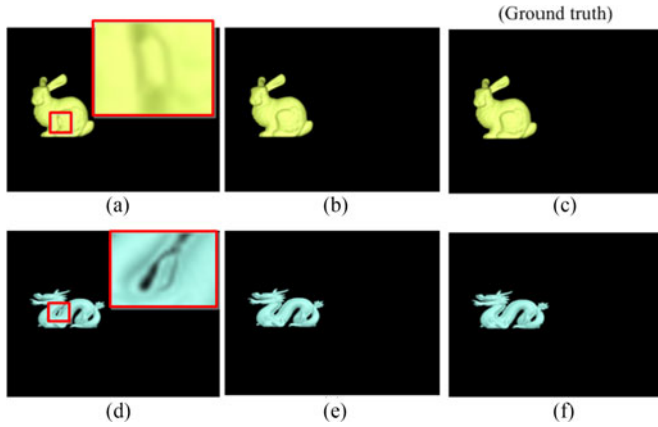


Fig. 10. Virtual views to be rendered for the bunny and dragon scenes. Diagrams (a) and (d) depict one of the 700 rendered virtual views, which are rendered using 310 captured images for the bunny and dragon, respectively. The red patch represents aliasing. Diagrams (b) and (e) depict the corresponding rendered virtual views, when 700 virtual views are rendered using 360 captured images. Diagrams (c) and (f) depict the ground truth.

number of captured images is 310, aliasing is nearly nonexistent. Using the rendered virtual views, aliasing occurs when 310 captured images are employed, as shown in Fig. 10(a) and (d). The red patch in the figures denotes the area of the aliasing. No aliasing is observed when the corresponding rendered virtual view is rendered using 360 captured images, as shown in Fig. 10(b) and (e). To expediently analyze the relationship between the quality of the reconstruction of the EPI volumes and the number of

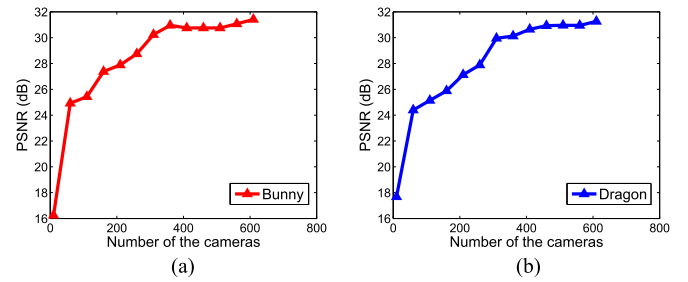


Fig. 11. Diagram (a) illustrates the PSNR of the reconstruction of the EPI volumes for the bunny using a different number of cameras. Diagram (b) illustrates the PSNR of the reconstruction of the EPI volumes for the dragon using a different number of cameras.

captured images, we calculate the PSNR for the reconstruction of the EPI volumes. Fig. 11 indicates that the PSNR will increase as the number of captured images increases. For both the bunny scenes and the dragon scenes, when more than 360-400 captured images are employed, the PSNR will be stable and will not be significantly affected by variations in the number of captured images. Additionally, the reconstruction of the EPI volume in Fig. 9(a) does not exhibit aliasing when the number of captured images is approximately 360. In this case, the maximum spacing between the cameras is approximately 5.56 mm for the bunny scene. Similarly, considering Figs. 9(b), 10(e) and 11(b), the reconstruction of the EPI volume does not exhibit aliasing when the number of captured images is approximately 370. Thus, the maximum spacing between the cameras is approximately

5.42 mm for the dragon scene. Based on this analysis, the maximum spacing between the cameras, which is calculated by (41), is equivalent to the spacing derived from the reconstruction experiments for the two scenes. Therefore, we can consider that the results of our proposed method are satisfactory. The ACF of the POF can also be applied to determine the sampling rate of the POF.

As shown in Table III, the difference in the estimated frequencies between the two scenes is 4.03 Hz. This difference is caused by the minimum and maximum depths and the difference between the scene complexity of the bunny scene and the scene complexity of the dragon scene. In addition, the dragon scene is more complicated than the bunny scene, and the  $\hat{\Omega}_t$  of the dragon scene is larger than the  $\hat{\Omega}_t$  of the bunny scene. The  $\hat{\Omega}_v$  of the bunny scene is 912.97 Hz, whereas that of the dragon scene is 750.84 Hz. This result is due to the different scene complexities of the bunny and dragon scenes, which can be observed in Fig. 7(e) and (f), respectively. Therefore, the resolution ratio of imaging for the dragon scene is less than that for the bunny scene. We can consider that  $\hat{\Omega}_t$  will be larger and that  $\hat{\Omega}_v$  will be smaller as the scene complexity or the maximum depth of the scene increases. In particular, the difference between the two scenes is also influenced by  $D_t$  and  $D_v$ . It should be noted that both  $D_t$  and  $D_v$  differ between the two scenes. Furthermore,  $D_t$  and  $D_v$  will increase as the scene complexity increases because the frequency of the POF is related to the scene depths and the scene complexity. These conclusions will be demonstrated in the following sub-sections.

### C. Comparison of the Experimental Results

To evaluate our results, the frequency estimation method using the ACF is compared with other methods. The frequency estimation method presented by Gilliam *et al.* [12] is selected for comparison purposes. This method is termed the closed-form expression of the plenoptic spectrum (CEPS), which is based on the assumption that the scene surface consists of a set of slanted planes. The authors derived a spectral expression of the POF for a slanted plane. Based on the spectral analysis of a slanted plane, the expression of the frequency estimation for  $\hat{\Omega}_t$  is calculated as

$$\hat{\Omega}_t = \frac{1}{\cos(\varphi) - \bar{v}_m |\sin(\varphi)|} \left( \omega_s + \frac{2\pi}{L_s} \right), \quad (42)$$

where  $\varphi$  is the angle between the plane and the  $x$ -axis,  $L_s$  is the length of the plane, and  $\bar{v}_m = v_m / f$ . Then, the expression of  $\hat{\Omega}_v$  is given by [12] as

$$\hat{\Omega}_v = \hat{\Omega}_t \frac{z_{max}}{f} + \frac{3\cos^2(\varphi) + 3.5(\bar{v}_m \sin(\varphi))^2}{3\cos^2(\varphi) + (\bar{v}_m \sin(\varphi))^2} \frac{\pi}{v_m}. \quad (43)$$

Then, the spectral support of the slanted plane is analyzed. This result can be used to effectively solve the frequency estimation problem. Therefore, to expediently compare their algorithm with our proposed method, the wall in Fig. 7(a) is selected for comparison. For this wall,  $\omega_s = 125$  rad/m. The width of the wall is 1270 mm, whereas the height is 1860 mm. These two parameters

TABLE IV  
THE ESTIMATED FREQUENCY BY OUR PROPOSED METHOD AND CEPS

Frequency estimation method	$\hat{\Omega}_t$ (Hz)	$\hat{\Omega}_v$ (Hz)
Using ACF	198.01	842.45
CEPS	188.29	884.13

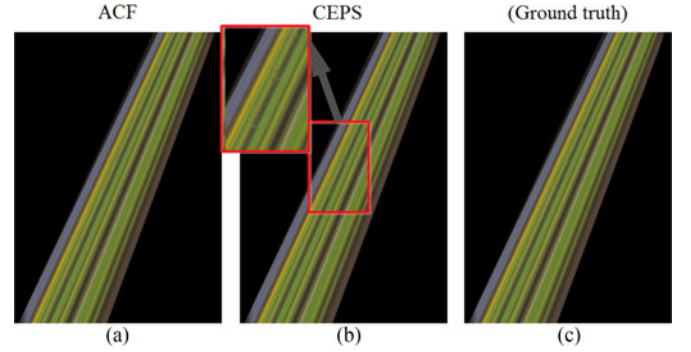


Fig. 12. In (a), the EPI-volume is rendered by a set of captured images when the maximum spacing between the cameras is calculated using (23). In (b) the EPI-volume is rendered when the maximum spacing is calculated using (42). In (c) the ground truth.

are randomly chosen and do not influence our results. The angle between the plane and the  $x$ -axis is  $\frac{\pi}{4}$  rad.

The estimated results are presented in Table IV. The values of  $\hat{\Omega}_t$  and  $\hat{\Omega}_v$  for the wall are determined to be 198.01 Hz and 842.45 Hz, respectively, using our proposed method. Using the CEPS, the  $\hat{\Omega}_t$  of the wall is 188.29 Hz, whereas  $\hat{\Omega}_v$  is 884.13 Hz. The frequency estimation using our proposed method is approximately the same as that using the CEPS. However, it should be noted that there are a few differences in the estimated frequency of the wall. The difference is approximately 9.72 Hz for  $\hat{\Omega}_t$ , and approximately 41.68 Hz for  $\hat{\Omega}_v$ . Conversely, the maximum spacing between the cameras for the wall is 5.35 mm using (23), and the corresponding PSNR of the reconstructing EPI of the wall is 34.52 dB as indicated in Fig. 12(a). However, the maximum spacing is 5.67 mm when (42) is applied in [12], and the corresponding PSNR of the reconstructing EPI is 34.10 dB as indicated in Fig. 12(b). The results indicate that these differences are attributed to the different accuracies of the two methods. The method of CEPS is based on an assumption that the scene surface contains Lambertian reflections. In practice, the scene surface does not contain non-Lambertian reflections. Our proposed method is not limited in this regard. In addition, we cannot obtain the Lambertian surface of the wall, and the actual scene also includes certain differences from the ideal scene. Therefore, a few differences in the estimation results are observed between the two methods.

To compare the performance of the two estimated methods, we reconstruct an EPI volume of the wall scene using different amounts of captured images. In the reconstruction experiments, each EPI volume consists of 450 reconstructed images and is rendered using a set of captured images. Different amounts of captured images in the range of [10, 420] are tested; the number of images is increased by increments of ten to analyze the



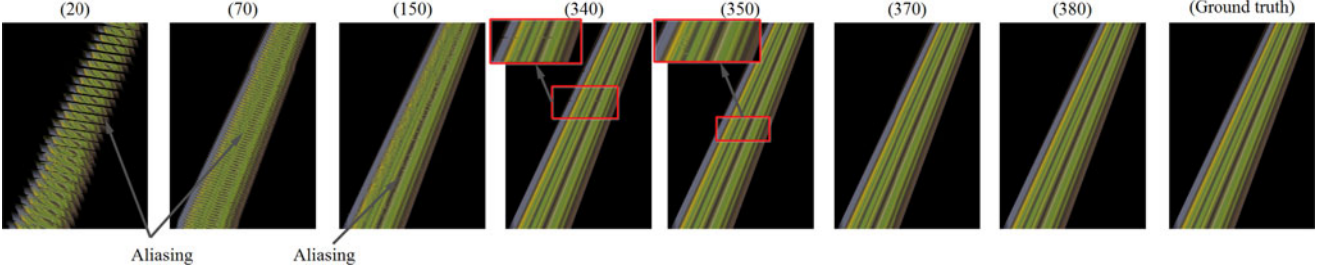


Fig. 13. The reconstruction of synthetic EPI volumes using different numbers of captured images for the wall scene.



Fig. 14. (a) Bracket-plant scene. (b) Plaster scene. (c) Chair scene. (d) Data acquisition set-up. A camera is moved along a track.

relationship between the rendering quality and the number of captured images. As shown in Fig. 13, the use of 20 cameras causes significant aliasing in the reconstruction of the EPI volumes. The aliasing will be reduced as the number of captured images increases. For example, when the number of captured images is 350, minimal aliasing occurs; when the number of captured images is 370, the aliasing is nearly nonexistent. No aliasing is observed when the corresponding rendered virtual view is rendered using 380 captured images. Thus, the maximum spacing between the cameras is approximately 5.25 mm for the wall scene.

Using this analysis, we determine the frequency estimation using our proposed method is closer to the ground truth of the POF frequency and obtain more accurate results than the results obtained using the CEPS.

#### D. Actual Scenes

To evaluate the ACF method, the frequency estimation experiments conducted using three actual scenes, as depicted in Fig. 14(a), (b), and (c). The three scenes are captured using the set-up depicted in Fig. 14(d). For each scene, a camera is uniformly moved every 5 mm along a straight line track, and then the images are captured. Let  $x \in [-500, 500]$  mm. Using

TABLE V  
THE ESTIMATED FREQUENCY AND VARIANCES BY OUR PROPOSED METHOD FOR THE ACTUAL SCENES

Methods	Scenes	$\hat{\Omega}_t$ (Hz)	$\hat{\Omega}_v$ (Hz)	$D_t$ (Hz)	$D_v$ (Hz)
ACF	Bracket-plant	184.89	821.23	0.40	6.47e-005
	Plaster	181.62	853.90	0.30	5.31e-005
	Chair	179.07	905.73	0.07	4.01e-005
CEPS	Bracket-plant	181.00	847.64		
	Plaster	170.93	874.78		
	Chair	168.07	919.07		

the captured images, the frequencies of the actual scenes are estimated, as illustrated in Table V.

Similar to Section VII-B, to verify the correctness of the results in Table V, an EPI volume consists of 200 reconstructed images and is rendered using different numbers of the captured images for each actual scene. Different numbers of captured images in the range of [10, 200] are tested, and the number is increased by increments of ten. The rendering quality of the reconstruction EPI volumes with respect to the number of the captured images is analyzed to determine the maximum spacing between the cameras. For the rendering results, as observed in Fig. 15(a1) and (a2), the use of 70 or 100 cameras causes significant aliasing in the reconstruction of the EPI volumes of the bracket-plant scene. The aliasing will be reduced as the number of captured images increases. When the number of captured images is 160, minimal aliasing occurs. When the number of captured images is 170, aliasing is nearly nonexistent. The reconstruction of the EPI volume does not exhibit aliasing when the number of captured images is approximately 180, as shown in Fig. 15(a5). The same phenomenon is evident using the rendered virtual views. In Fig. 16(a1), aliasing occurs when 160 captured images are employed. No aliasing is observed when the corresponding rendered virtual view is rendered using 170 captured images, as shown in Fig. 16(a2). Additionally, Fig. 17 indicates that the PSNR will increase as the number of captured images increases. For the bracket-plant scene, when more than 170 captured images are employed, the PSNR will be stable and will not be significantly affected by variations in the number of captured images. In this case, the maximum spacing between the cameras is approximately 5.88 mm for the bracket-plant scene. By Table V, the maximum spacing between the cameras, which is calculated by (41), is 5.87 mm. This spacing is equivalent to the spacing that is derived from



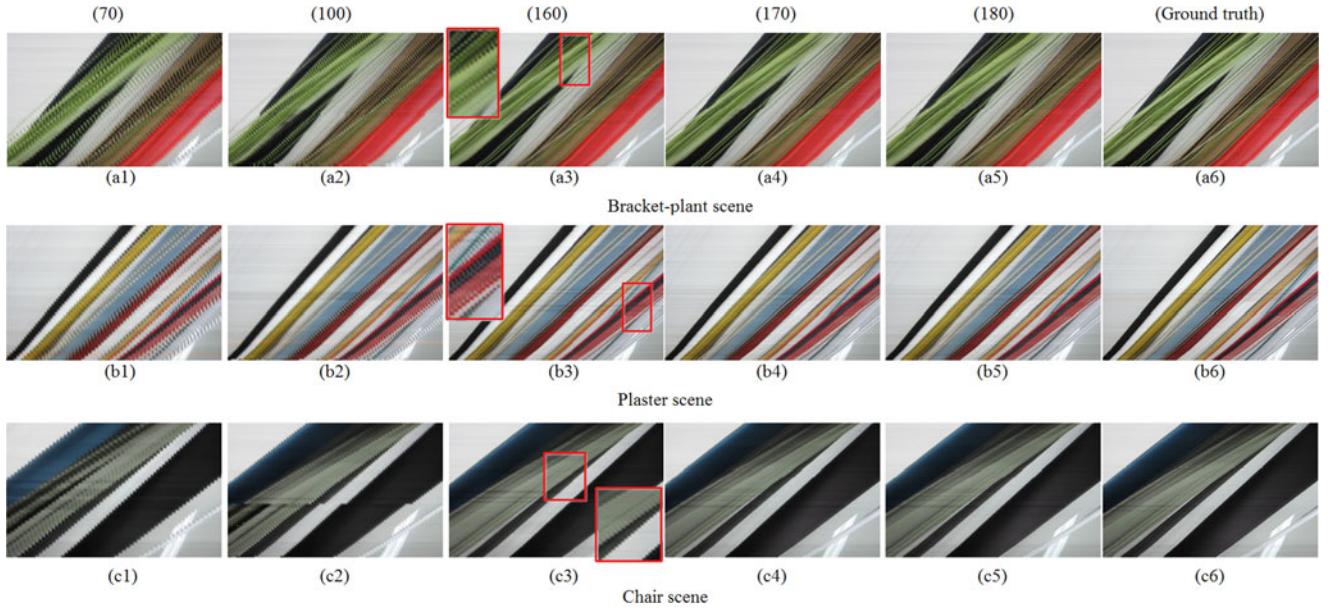


Fig. 15. The reconstruction of synthetic EPI volumes for the bracket-plant, plaster and chair scenes. When the number of cameras is too small, aliasing will occur. The aliasing will be reduced as the number of cameras increases. Diagrams (a1)–(a5) are the reconstruction EPI volumes for the bracket-plant scene using a different number of cameras, such as 70, 100, 160, 170, and 180; diagrams (b1)–(b5) depict the results for plaster scene using a different number of cameras; diagrams (c1)–(c5) display the results for chair scene using a different number of cameras. and diagrams (a6), (b6), (c6) depict the ground truth.

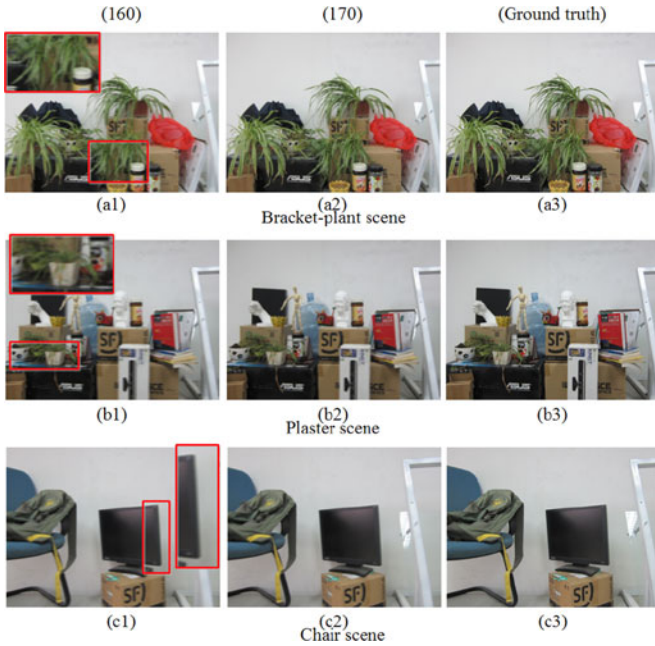


Fig. 16. Rendered images of the virtual views for the bracket-plant, plaster and chair scenes. Diagrams (a1), (b1), (c1) depict one of the 200 rendered virtual views, which are rendered using 160 captured images for the three scenes. Diagrams (a2), (b2), (c2) depict the corresponding rendered virtual view, when 200 virtual views are rendered using 170 captured images. Diagrams (a3), (b3), (c3) display the ground truth.

the reconstruction experiments for the bracket-plant scene. Similarly, the same results are obtained for the plaster and chair scenes, as shown in the Figs. 15(b1)–(b6) and 15(c1)–(c6), 16(b1)–(b3) and 16(c1)–(c3), and 17. Apparently, the maximum

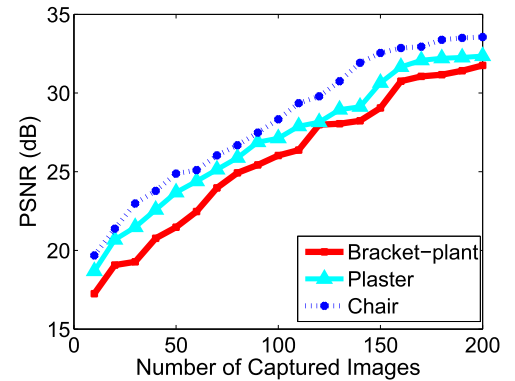


Fig. 17. The PSNR of the rendered EPI volumes with respect to the number of captured images for the actual scenes.

spacing between the cameras is approximately 6.06 mm for the plaster scene and 6.25 mm for the chair scene. When using (41) calculation, the maximum spacing is 6.19 mm for the plaster scene and 6.34 mm for the chair scene by Table V. To compare the performance of the ACF method, the CEPS method is also applied to estimate the three actual scenes. The estimated frequencies for the actual scenes are presented in Table V. For the bracket-plant scene, the maximum spacing between two cameras is 5.97 mm when (42) is applied. Similarly, the maximum spacing is 6.32 mm for the plaster scene. The maximum spacing is 6.57 mm for the chair scene. Compare to the results by the two methods, the frequency estimation using our proposed method is closer to the ground truth of the POF frequency and obtain more accurate results than the results obtained using the CEPS for the actual scenes. These

TABLE VI  
THE ESTIMATED FREQUENCY AND VARIANCES BY OUR PROPOSED METHOD  
FOR FOUR DIFFERENT SCENES

Scene	$\hat{\Omega}_t$ (Hz)	$\hat{\Omega}_v$ (Hz)	$D_t$ (Hz)	$D_v$ (Hz)
Wall	198.21	$1.20e+003$	1.55	$0.06e-004$
Cube-annulus	206.01	$1.18e+003$	1.80	$9.41e-004$
Cone	218.01	952.67	1.94	$10.0e-004$
Teapot	213.51	$1.15e+003$	1.91	$9.91e-004$

results demonstrate that the method of ACF can be applied to determine the sampling rate of the POF in the actual scenes.

### E. Frequency Estimation With the Scene Complexity

To analyze the influence of the scene complexity on the frequency estimation of the POF, two experiments are performed in this sub-section. The two experiments use four scenes (i.e., the wall, cube-annulus, cone and teapot in Fig. 7). The results are demonstrated below.

1) *Frequency Estimation With the Scene Surface Variation:* For this experimental setting, the surface complexity (e.g., curvatures) of the different scenes is different, and the minimum and maximum depths of these scenes are the same. The estimated frequencies of the scenes are shown in Table VI. We observe that  $\hat{\Omega}_t$  of the wall is 198.21 Hz, and  $\hat{\Omega}_v$  is  $1.20e+003$  Hz. Additionally,  $\hat{\Omega}_t$  of the cube-annulus is 206.01 Hz, and  $\hat{\Omega}_v$  is  $1.18e+003$  Hz. For the cone,  $\hat{\Omega}_t$  is 218.01 Hz, and  $\hat{\Omega}_v$  is 952.67 Hz. For the teapot,  $\hat{\Omega}_t$  is 213.51 Hz, and  $\hat{\Omega}_v$  is  $1.15e+003$  Hz. Therefore,  $\hat{\Omega}_t$  will be larger and  $\hat{\Omega}_v$  will be smaller as the surface geometry shape becomes more complicated. In contrast, the estimation error will increase as the surface geometry becomes more complicated. For example,  $D_t$  of the cone is 1.94, and  $D_v$  is  $10.0e-004$ ; they are both larger than the  $D_t$  and  $D_v$  values of the wall, cube-annulus and teapot.

2) *Frequency Estimation With the Scene Width:* To analyze the influence of the scene width on the frequency estimation of the POF, each scene width varies from 500 to 2000 mm. Furthermore, the height of the scenes is constant. Specifically, we change the size of  $B$  in (5).

The ACF of the various widths of the scenes is applied to estimate the frequency of the scene and calculate the frequency estimation error. The results for the frequency estimation with various widths are shown in Fig. 18. Fig. 18(a) and (b) illustrate that the frequency of the scene will increase as the width of the scene increases. For example, when the width of the wall is 500 mm,  $\hat{\Omega}_t$  of the wall is 100.24 Hz, while  $\hat{\Omega}_v$  of the wall is 364.48 Hz. When the width of the wall is 1600 mm,  $\hat{\Omega}_t$  is 190.64 Hz, and  $\hat{\Omega}_v$  is 896.27 Hz. Furthermore, Fig. 18(c) and (d) illustrate that the estimation error of the frequency will increase as the width of the scene increases. For example, when the width of the wall is 500 mm,  $D_t$  of the wall is 0.12 Hz, and  $D_v$  of the wall is  $0.21e-004$  Hz. When the width of the wall is 1600 mm,  $D_t$  is 1.71 Hz, and  $D_v$  is  $7.91e-004$  Hz because the width of the wall varies. Based on these simulation results, we determine that the frequency of the POF will increase if the scene complexity increases.

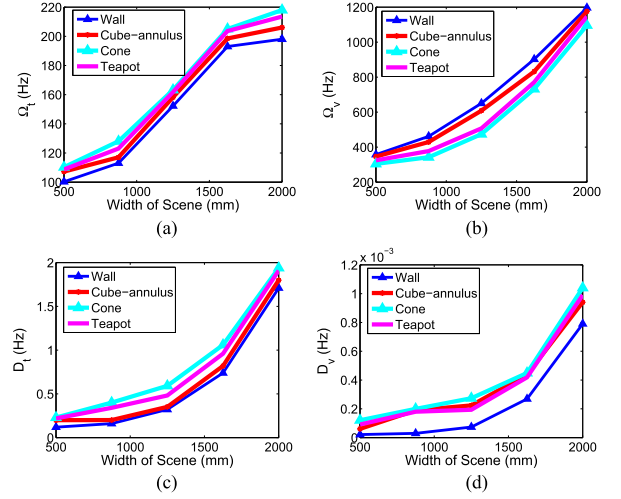


Fig. 18. Diagram of the frequency estimation and the estimation error with respect to the scene width. Diagram (a) illustrates the influence of  $\hat{\Omega}_t$  on the scene width; diagram (b) illustrates the influence of  $\hat{\Omega}_v$  on the scene width; diagram (c) illustrates  $D_t$  with respect to the scene width; and diagram (d) illustrates  $D_v$  with respect to the scene width.

### F. Frequency Estimation With the Depth

The bunny and dragon scenes are also selected for this experiment. Thus, the influence of the scene depth on the frequency estimation can be derived. The experimental results are obtained using a completely static space. First, the ACFs of the POF are measured at different spatial positions. Specifically, the bunny and dragon scenes are moved to a position of 500 mm along a line  $L_z = \{(0, 0, z)\}$ , where  $z$  denotes the variation in the bunny and dragon scene positions,  $z \in [3000, 6000]$  mm. In other words, we can change the distance between the bunny or dragon and the  $x$ -axis (the camera plane) to change the depths of the scenes. Similar to the technique presented in Section VII-A, a set of images are captured with a camera spacing of 10 mm along a line  $L = \{(x, 0, 0)\}$  for each depth setting.

The results for the frequency estimation with varying depths are provided in Fig. 19. Fig. 19(a) illustrates that the frequency of the scenes along the  $\hat{\Omega}_t$ -axis will increase as the depth of the scenes increases. For example, when the depth of the scenes is 3000 mm,  $\hat{\Omega}_t$  of the bunny is 177.42 Hz, and  $\hat{\Omega}_t$  of the dragon is 181.45 Hz. When the depth of the bunny is 6000 mm,  $\hat{\Omega}_t$  of the bunny is 198.01 Hz, and  $\hat{\Omega}_t$  of the dragon is 184.32 Hz. It should be noted that various differences exist between the bunny and the dragon scenes. The differences are caused by the differences in the size and the scene complexity between the bunny and the dragon scenes.

The relationship between  $\hat{\Omega}_v$  and the depth of the scene is different from  $\hat{\Omega}_t$  in Fig. 19(b). This figure illustrates that the frequency of the scene along the  $\hat{\Omega}_v$ -axis will decrease as the depth of the scene increases because the resolution of the captured image will decrease as the depth of the scene increases. However, the estimation error will decrease if the depth of the scene increases, as indicated in Fig. 19(c) and (d). It should be noted that because the surface complexity of the dragon is greater than that of the bunny, the estimation error of the dragon

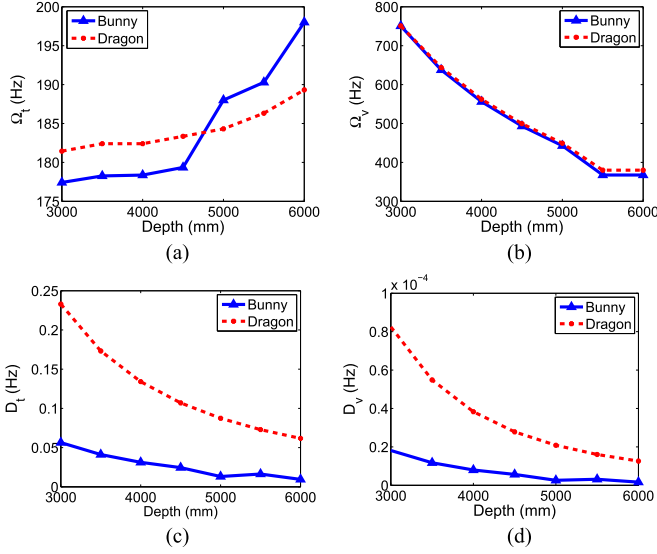


Fig. 19. Frequency estimation and the estimation error with respect to the scene depth. The scenes are for the bunny and the dragon. Diagram (a) illustrates the influence of  $\hat{\Omega}_t$  on the scene depth; diagram (b) illustrates the influence of  $\hat{\Omega}_v$  on the scene depth; diagram (c) illustrates  $D_t$  with respect to the scene depth; and diagram (d) illustrates  $D_v$  with respect to the scene depth.

scene is larger than that of the bunny scene. Furthermore, it can be noted that  $D_t$  is larger than  $D_v$  for both the dragon and bunny scenes because the influence of the plenoptic sampling on the camera position along the  $t$ -axis is larger than the image plane along the  $v$ -axis. Moreover, the differences between the scenes influence  $D_t$  and  $D_v$ . It should be noted that both  $D_t$  and  $D_v$  are different for all of the scenes because the frequency of the POF is related to the scene depth and the scene complexity.

## VIII. CONCLUSION

In this paper, we introduced and examined the frequency estimation of the POF to determine the minimum sampling rate. The POF characterizes a scene that is static at any point in space. This function and its linear ACF with respect to the object surface are investigated. The influence of scene depth and scene geometry on the ACF and both the spatial frequency and the image plane frequency axes are investigated. Based on the ACF support of the POF, the frequency of the POF is determined. The frequency of the POF increases as the scene complexity increases. The frequency on the  $\Omega_t$ -axis will increase as the depth of the scene increases and  $\Omega_v$  will decrease. The number and spacing between the cameras that are needed to reconstruct the continuous light field to a certain spatial frequency are determined. The experimental results are presented and compared with the theoretical results.

## APPENDIX A

According to the results of Chai and Gilliam *et al.* [9]–[12], the  $p(t, v)$  is a 2-D periodic function. Its frequencies are  $\Omega_t = F_t \{p(t, v)\}$  and  $\Omega_v = F_v \{p(t, v)\}$ , where  $F_t \{\cdot\}$  and  $F_v \{\cdot\}$  are the Fourier transform. In particular, we consider the

$p(t, v)$  to consist of a single complex 2-D sinusoid in complex scenes. Denoting the phases of  $p(t, v)$  by  $\alpha_t$  and  $\beta_v$ , and they expressions are accordingly given as

$$\alpha_t = \arg(p(t, \cdot)) = \Omega_t t, \quad (44)$$

$$\beta_v = \arg(p(\cdot, v)) = \Omega_v v. \quad (45)$$

Assuming that we wish to estimate only the frequencies (i.e.,  $\Omega_t$  and  $\Omega_v$ ), by (44) and (45), we can avoid phase unwrapping by considering the differenced phase data

$$\Delta_t = \alpha_{t+1} - \alpha_t = \Omega_t, \quad (46)$$

$$\Delta_v = \beta_{v+1} - \beta_v = \Omega_v. \quad (47)$$

First, by the linear model of (46), the MLE of  $\Omega_t$  which is equivalent to the minimum variance unbiased estimator is found by minimizing [17]

$$J = (\mathbf{\Delta} - \mathbf{\Omega}_t)^H \mathbf{C}^{-1} (\mathbf{\Delta} - \mathbf{\Omega}_t), \quad (48)$$

where  $H$  is the transpose operation,  $\mathbf{\Delta} = [\Delta_0, \Delta_1, \dots, \Delta_{T-1}]^H$ , and  $\mathbf{C}$  is the  $(T-1) \times (T-1)$  covariance matrix of  $\Delta_t$ . The solution to this problem can be derived as

$$\hat{\Omega}_t = \frac{\mathbf{C}^{-1} \mathbf{\Delta}}{\mathbf{C}^{-1}}. \quad (49)$$

Because the complexity of  $\mathbf{C}$  and the expression of  $p(t, v)$  is unknown, we can't evaluate the characteristic of  $\mathbf{C}^{-1}$  to calculate  $\hat{\Omega}_t$ . However, according to (49), we can describe  $\hat{\Omega}_t$  using a function with respect to  $\Delta_t$  as

$$\hat{\Omega}_t = \hat{\Phi}_t(\Delta_t). \quad (50)$$

Because  $\hat{\Omega}_t$  is an unbiased estimator, the frequency estimator may further be written by using the equivalence

$$\Delta_t = \alpha_{t+1} - \alpha_t = \alpha_t^* \alpha_{t+1} = \arg\{R(\tau_1, \cdot)\}. \quad (51)$$

Substituting (51) to (50), the estimated expression of  $\hat{\Omega}_t$  with respect to the ACF phase is written as

$$\hat{\Omega}_t = \hat{\Phi}_t(\arg\{R(\tau_1, \cdot)\}). \quad (52)$$

After that, similarly to the derivation of (52),  $\hat{\Omega}_v$  also can be written as a function with respect to the phase of  $R(\cdot, \tau_2)$

$$\hat{\Omega}_v = \hat{\Phi}_v(\arg\{R(\cdot, \tau_2)\}). \quad (53)$$

## APPENDIX B

In this appendix, we derive the frequency estimator of the POF from a set of discrete samples. The frequency estimator presented in this paper results from examining the MLE of the light field. First, the continuous  $p(t, v)$  is sampled using a camera spacing  $\Delta t$  along the  $t$ -axis and a pixel spacing  $\Delta v$  along the  $v$ -axis. Using a rectangular sampling lattice [9], the discrete 2-D light field  $p(n, m)$  is represented by

$$p(n, m) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} p(t, v) \delta(t - n\Delta t) \delta(v - m\Delta v). \quad (54)$$



Because the assumption for a finite number of cameras and a finite camera FOV, we can let  $1 \leq n \leq N_a$  and  $1 \leq m \leq M_a$ .

After that, we derive a frequency estimator of the discrete 2-D light field using the ACF in reference to [15], [16]. Specifically, the proposed frequency estimation algorithm is an approximation of the MLE discussed in [15]. On the other hand, Fitz demonstrated in [15] that the MLE of the frequency of a sinusoid could be obtained by an estimator of the form as

$$\hat{\Omega} = \arg \max \left| \sum_{k=1}^N x_k e^{-j\Omega k} \right|^2, \quad (55)$$

where  $x_k$  is a discrete signal,  $\Omega$  is its frequency, and  $N$  is the length of the signal. Similar to (55), the MLE of the spatial frequency of a 2-D light field in the model  $p(n, m)$  is the maximum of a periodogram [40], which is given as

$$\hat{\Omega}_t = \arg \max \left| \sum_{n=1}^{N_a} p(n, m) e^{-j\Omega_t n} \right|^2. \quad (56)$$

In (56),  $\hat{\Omega}_t$  is calculated using the periodogram of  $p(n, m)$  along the  $t$ -axis. Similarly,  $\hat{\Omega}_v$  is also calculated using the periodogram of  $p(n, m)$  along the  $v$ -axis as

$$\hat{\Omega}_v = \arg \max \left| \sum_{m=1}^{M_a} p(n, m) e^{-j\Omega_v m} \right|^2. \quad (57)$$

## REFERENCES

- [1] A. Kubota, A. Smolic, M. Magnor, M. Tanimoto, T. Chen, and C. Zhang, "Multiview imaging and 3DTV," *IEEE Signal Process. Mag.*, vol. 24, no. 6, pp. 10–21, Nov. 2007.
- [2] M. Tanimoto, M. P. Tehrani, T. Fujii, and T. Yendo, "Free-viewpoint TV," *IEEE Signal Process. Mag.*, vol. 28, no. 1, pp. 67–76, Jan. 2011.
- [3] M. N. Do, Q. H. Nguyen, H. T. Nguyen, D. Kubacki, and S. J. Patel, "Immersive visual communication," *IEEE Signal Process. Mag.*, vol. 28, no. 1, pp. 58–66, Jan. 2011.
- [4] H.-Y. Shum, S.-C. Chan, and S. B. Kang, *Image-Based Rendering*. Berlin, Germany: Springer, 2007.
- [5] H.-Y. Shum, S. Kang, and S.-C. Chan, "Survey of image-based representations and compression techniques," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 13, no. 11, pp. 1020–1037, Nov. 2003.
- [6] C. Zhang and T. Chen, "A survey on image-based rendering-representation, sampling and compression," *EURASIP Signal Proc.: Image Commun.*, vol. 19, no. 1, pp. 1–28, Jan. 2004.
- [7] E. H. Adelson and J. R. Bergen, "The plenoptic function and the elements of early vision," in *Proc. Comput. Models Vis. Process.*, 1991, pp. 3–20.
- [8] J. Berent and P. L. Dragotti, "Plenoptic manifolds," *IEEE Signal Process. Mag.*, vol. 24, no. 6, pp. 34–44, Nov. 2007.
- [9] J.-X. Chai, X. Tong, S.-C. Chan, and H.-Y. Shum, "Plenoptic sampling," in *Proc. SIGGRAPH*, 2000, pp. 307–318.
- [10] C. Zhang and T. Chen, "Spectral analysis for sampling image-based rendering data," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 11, pp. 1038–1050, Nov. 2003.
- [11] M. N. Do, D. Marchand-Maillet, and M. Vetterli, "On the bandwidth of the plenoptic function," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 708–717, Feb. 2012.
- [12] C. Gilliam, P. Dragotti, and M. Brookes, "On the spectrum of the plenoptic function," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 502–516, Feb. 2014.
- [13] L. Palmer, "Coarse frequency estimation using the discrete Fourier transform (Corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 1, pp. 104–109, 1974.
- [14] S. Tretter, "Estimating the frequency of a noisy sinusoid by linear regression (Corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 6, pp. 832–835, 1985.
- [15] M. P. Fitz, "Further results in the fast estimation of a single frequency," *IEEE Trans. Commun.*, vol. 42, no. 234, pp. 862–864, Feb–Apr. 1994.
- [16] U. Mengali, and M. Morelli, "Data-aided frequency estimation for burst digital transmission," *IEEE Trans. Commun.*, vol. 45, no. 1, pp. 23–25, Jan. 1997.
- [17] S. Kay, "A fast and accurate single frequency estimator," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 37, no. 12, pp. 1987–1990, Dec. 1989.
- [18] T. Brown and M. M. Wang, "An iterative algorithm for single-frequency estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2671–2682, Nov. 2002.
- [19] N. Wiener, "Generalized harmonic analysis," *Acta Mathematica*, vol. 55, no. 1, pp. 117–258, 1930.
- [20] A. Khintchine, "Korrelationstheorie der stationären stochastischen Prozesse," *Mathematische Annalen*, vol. 109, no. 1, pp. 604–615, 1934.
- [21] E. Camahort, A. Lieros, and D. Fussell, *Uniformly Sampled Light Fields (Rendering Techniques '98)*. Vienna: Austria: Springer, 1998, pp. 117–130.
- [22] H.-Y. Shum and L.-W. He, "Rendering with concentric mosaics," in *Proc. Comput. Graph.*, Aug. 1999, pp. 299–306.
- [23] S. Gortler, R. Grzeszczuk, R. Szeliski, and M. Cohen, "The lumigraph," in *Proc. SIGGRAPH*, 1996, pp. 43–54.
- [24] M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. SIGGRAPH*, New Orleans, LA, USA, 1996, pp. 31–40.
- [25] J. Stewart, J. Yu, S. J. Gortler, and L. McMillan, "A new reconstruction filter for undersampled light fields," in *ACM Int. Conf. Proc. Ser.* vol. 44, Jun. 2003, pp. 150–156.
- [26] C. Buehler, M. Bosse, L. McMillan, S. J. Gortler, and M. F. Cohen, "Unstructured lumigraph rendering," in *Proc. SIGGRAPH*, 2001, pp. 425–432.
- [27] C. Chen, and D. Schonfeld, "Geometrical plenoptic sampling," in *Proc. 16th IEEE Int. Conf. Image Process.*, 2009, pp. 3769–3772.
- [28] L. Bagnato, P. Frossard, and P. Vanderghyest, "Plenoptic spherical sampling," in *Proc. 19th IEEE ICIP*, Sep./Oct. 2012, pp. 357–360.
- [29] Z. Lin and H.-Y. Shum, "On the number of samples needed in light field rendering with constant-depth assumption," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2000, vol. 1, pp. 588–595.
- [30] C. K. Liang, Y. C. Shih, and H. H. Chen, "Light field analysis for modeling image formation," *IEEE Trans. Image Process.*, vol. 20, no. 2, pp. 446–460, Feb. 2011.
- [31] J. Shade, S. Gortler, L. W. He, and R. Szeliski, "Layered depth images," in *Proc. SIGGRAPH*, Orlando, FL, USA, 1998, pp. 231–242.
- [32] C. L. Zitnick *et al.*, "High-quality video view interpolation using a layered representation," *ACM Trans. Graph.*, vol. 23, no. 3, 2004, pp. 600–608.
- [33] J. Pearson, M. Brookes, and P. L. Dragotti, "Plenoptic layer-based modeling for image based rendering," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3405–3419, Sep. 2013.
- [34] G. Chaurasia, O. Sorkine-Hornung, and G. Drettakis, "Silhouette-aware warping for image-based rendering," *Proc. Comput. Graph. Forum*, vol. 30, no. 4, pp. 1223–1232, 2011.
- [35] L. McMillan and G. Bishop, "Plenoptic modeling: An image-based rendering system," in *Proc. Comput. Graph.*, Aug. 1995, pp. 39–46.
- [36] R. Bolles, H. Baker, and D. Marimont, "Epipolar-plane image analysis: An approach to determining structure from motion," *Int. J. Comput. Vis.*, vol. 1, no. 1, pp. 7–55, 1987.
- [37] G. S. P. Miller, S. Rubin, and D. Ponceleon, "Lazy decompression of surface light fields for precomputed global illumination," in *Rendering Techniques'98 (Proceedings of the Eurographics Workshop on Rendering)*. New York, NY, USA: Springer, 1998, pp. 281–292.
- [38] D. N. Wood *et al.*, "Surface light fields for 3D photography," in *Proc. Comput. Graph.*, Jul. 2000, pp. 287–296.
- [39] H. T. Nguyen and M. N. Do, "Error analysis for image-based rendering with depth information," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 703–716, Apr. 2009.
- [40] A. J. Viterbi, *Principles of Coherent Communication*. New York, NY, USA: McGraw-Hill, 1966.
- [41] C. E. Shannon, "Communications in the presence of noise," *Proc. IREE*, vol. 37, no. 1, pp. 10–21, Jan. 1949.
- [42] P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1992.
- [43] *Stanford 3D scanning*. Aug. 2014. [Online]. Available: <https://graphics.stanford.edu/data/3Dscanrep/>

Authors' photographs and biographies not available at the time of publication.