

The plenoptic spectrum analysis and reconstruction filter design for the occlusive scene

Peng Zhou, Li Yu*, Gang Zhong

School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China

ARTICLE INFO

Article history:

Received 1 February 2015

Received in revised form

23 July 2015

Accepted 23 July 2015

Keywords:

Spectral analysis

Occlusive scene

Sampling rate

Reconstruction filter

ABSTRACT

The spectral analysis of the plenoptic function is an intuitive way to acquire the sampling rate of the image-based rendering (IBR) data. In this paper, we present a new epipolar plane image (EPI) model to analyze the spectrum of the plenoptic function and derive the plenoptic spectrum of the occlusive scene. In particular, the spectrum is separated into two aspects. The main part of the spectrum is bounded by the minimum and maximum depths of the scene. The second part is the broadening area which is analyzed by modeling the EPI lines. Our analysis reveals that, the occlusion leads to the spectrum broadening, and the broadening range depends on the occluding depth difference, the size of the occluding object and the maximum frequency of the texture painted on the occluding object. Based on the spectrum analysis, a new Nyquist rate and the corresponding reconstruction filter are deduced for the occlusive scene. Experiments on both synthetic scenes and real scenes show that our sampling spacing and reconstruction filter outperform other methods.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Image-based rendering (IBR) is an attractive technique for rendering novel views of the scene by a few existing views [1,2]. It can be regarded as a sampling and reconstruction problem. Firstly, the scene is sampled with finite cameras, and then the new views are reconstructed from the samples to represent the scene. The study of sampling the plenoptic function is also termed as the plenoptic sampling [3], and it is of great importance to figure out the minimum sampling rate and the reconstruction filter to avoid aliasing.

The first plenoptic sampling research was introduced by Chai et al. [3]. By assuming the ideal scene with Lambertian reflection and no-occlusion, they analyzed the spectrum of the plenoptic function and obtained the uniform sampling rate. Zhang et al. [4] extended the frequency spectral analysis to more general cases. They introduced a new representation of surface plenoptic function (SPF) to analyze the special scenes with non-Lambertian reflection and occlusion, respectively. They concluded that the spectrum of the scenes with occlusion is broadened. Do et al. [5] investigated the bandwidth of the plenoptic function. They found that the plenoptic spectrum is band-unlimited unless the surface is flat. Furthermore, they claimed that the bandwidth of the plenoptic function depends on the maximum surface slope and

maximum frequency of painted signals [6]. Gilliam et al. [7] paid attention to the scene of the slanted plane captured by finite field-of-view (FOV) cameras. In their recent work [8], they further got a new reconstruction filter for the slanted plane scene. In our previous work [9], we analyzed the spectral broadening of the scenes with non-Lambertian reflection. By introducing the reflection function into the light field, we concluded that the spectrum is broadened and the broadening range is decided by the material property. The spectral analysis was also extended to the camera with arbitrary position and orientation. Chen et al. [10] examined the plenoptic spectrum of unstructured camera systems. Bagnato et al. [11] analyzed the plenoptic sampling when the cameras are arranged on a sphere. In summary, for an ideal scene with Lambertian reflection and no-occlusion, the spectrum of the plenoptic function is bounded by the minimum and maximum depths. As considering the non-Lambertian reflection or the occlusion, the spectrum is broadened, and the broadening range needs to be further studied.

In this paper, we examine the spectrum of the occlusive scene with finite FOV cameras. According to the early conclusion, the plenoptic spectrum of the ideal scene is only bounded by the minimum and maximum depths. As considering the complicated scene, the spectrum is broadened. For the reason that the plenoptic spectrum is *bow-tie* shaped, the broadening happens on the edge of the original spectrum. By digging the essence of the plenoptic sampling, the plenoptic spectrum could be separated into two aspects. The main part of the spectrum is bounded by the minimum and maximum depths, which conforms to the

* Corresponding author.

E-mail address: hustlyu@hust.edu.cn (L. Yu).

conclusion of paper [3]. The second part is the broadening area which is usually introduced by the non-Lambertian reflection, the occlusion or other factors. For the first time, we derive an EPI model to analyze the spectrum (especially the spectral broadening part) of the plenoptic function. Furthermore, we conclude that in the occlusive scene, the spectral broadening happens in the maximum depth area, and the broadening range depends on the occluding depth difference, the size of the occluding object and the maximum frequency of the texture painted on the occluding object. For uniform plenoptic sampling, a new Nyquist rate and the corresponding reconstruction filter for the scene with occlusion are derived. Experiments on both synthetic scenes and real scenes demonstrate the improvement of our sampling rate and reconstruction filter. Compared with the existing works, our method achieves superior quality with lower sampling rate.

In summary, the contributions of our work in this paper are twofold:

- We separate the plenoptic spectrum into two parts. The main part of the spectrum is bounded by the minimum and maximum depths. The second part is the broadening area, and the broadening range is analyzed by modeling the EPI lines on the edge of the spectrum. Compared with the existing works, a quantitative relationship of the spectrum of the occlusive scene is obtained.
- We obtain a new Nyquist rate and the corresponding reconstruction filter for the scene with occlusion. Our method achieves superior quality with lower sampling rate. The analysis of the plenoptic function of the occlusive scene presented in this paper extends the plenoptic sampling to more general scenes and benefits the design of new IBR algorithms.

The outline of the paper is organized as follows. In Section 2, we cover the state-of-the-art in uniform plenoptic sampling. In Section 3, an EPI model to analyze the plenoptic spectrum is derived, and it is validated by the application in the scene with Lambertian reflection and the constant depth scene with finite FOV cameras. Based on the EPI model, the spectral broadening in the occlusive scene is investigated in Section 4. In Section 5, we obtain the new Nyquist rate and the corresponding reconstruction filter of the occlusive scene. In addition, we analyze the sampling of the occlusive scene with the known focal plane. Our sampling method is evaluated by experiments in both synthetic and real scenes in Section 6. Finally, conclusions are given in Section 7.

2. Uniform plenoptic sampling

The uniform of plenoptic sampling is naturally analyzed within the classical Fourier framework. The sampling leads to the spectrum replication in the frequency domain and the minimum Nyquist sampling rate is needed to make sure the replicas do not overlap. Owing to this, it is essential to analyze the bandwidth of the plenoptic spectrum.

2.1. Plenoptic sampling

The 7D plenoptic function is shown in Fig. 1(a). The scene is modeled by a set of light rays captured at a space location (X, Y, Z), toward a direction (θ, ϕ), over a wavelength (λ) and at a time (τ), which forms a 7D plenoptic function [12]. To avoid its high dimensionality and tremendous amount of data, the plenoptic function is simplified by the light field representation [13,14], as shown in Fig. 1(b). The light ray in the 4D light field is described by two parallel planes, the camera plane and the focal plane, respectively. This parametrization assumes that the scene is static, and the intensity of a light ray is constant along the transmission. The light field $p(s, t, u, v)$ defines the intensity of the light ray at the camera position (s, t) and with the direction (u, v). Furthermore, by fixing the s and u , the light field is simplified to the epipolar plane image(EPI) [15]. In this way, the plenoptic function finally can be simplified to a 2D case as $p(t, v)$.

The EPI image is a volume constructed by multi-view images taken from equidistant locations along a line. In the EPI image, a point in the scene is mapped to an EPI line with a slope inverse to its depth. The line is formed by the pixels captured along the camera path, and its intensity is depended on the discrete pixels captured by the cameras. Under the condition of Lambertian reflection and no occlusion, the intensity of an EPI line maintains the same, and the intensity capture in camera position t can be mapped to the camera position 0. In paper [3], the spectrum of the 2D light field with constant depth Z is

$$\begin{aligned} P(\Omega_v, \Omega_t) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(v, t) e^{-j(\Omega_v v + \Omega_t t)} \\ dv dt &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(v, 0) e^{-j(\Omega_v v + \Omega_t t)} \\ dv dt &= \int_{-\infty}^{\infty} p(v, 0) e^{-j\Omega_v v} dv \int_{-\infty}^{\infty} e^{-j(\Omega_t + \frac{f}{Z}\Omega_v)t} dt \\ dt &= 2\pi P(\Omega_v) \delta\left(\Omega_t + \frac{f}{Z}\Omega_v\right) \end{aligned} \quad (1)$$

where $P(\Omega_v)$ is the Fourier transform of $p(v, 0)$. The spectral

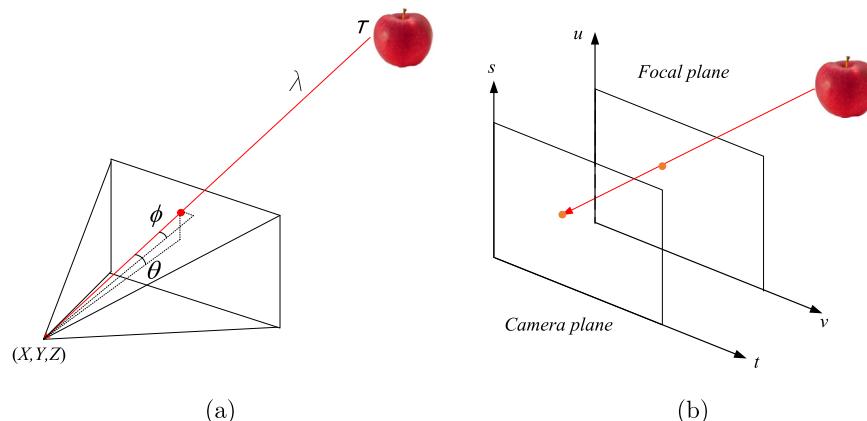


Fig. 1. (a) The 7D plenoptic function, the light rays are captured at a space location (X, Y, Z), toward a direction (θ, ϕ), over a wavelength (λ) and at a time (τ). (b) The 4D light field. The light field $p(s, t, u, v)$ defines the intensity of the light ray at the camera position (s, t) with the direction (u, v).

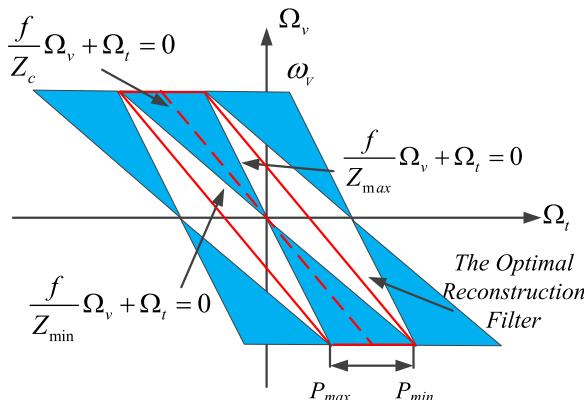


Fig. 2. The sampling rate and the optimal reconstruction filter.

support of the light field is defined by a line $\Omega_t + \frac{f}{Z} \Omega_v = 0$. As in Fig. 2, the scene is set with the maximum depth Z_{max} and the minimum depth Z_{min} , the spectral support is bounded by two lines $\Omega_t + \frac{f}{Z_{max}} \Omega_v = 0$ and $\Omega_t + \frac{f}{Z_{min}} \Omega_v = 0$. Due to the finite pixel resolution Δv , the spectrum is bandlimited in Ω_v at $\frac{\pi}{\Delta v}$. The spectrum is cut off as a bow-tie shape. The smallest interval in the spectral support is

$$|P_{max} P_{min}| = \frac{\pi f}{\Delta v} \left(\frac{1}{Z_{min}} - \frac{1}{Z_{max}} \right) 1 \quad (2)$$

The minimum sampling rate is equivalent to the maximum camera spacing Δt_c , and

$$\Delta t_c = \frac{2\pi}{|P_{max} P_{min}|} = \frac{2\Delta v Z_{max} Z_{min}}{f(Z_{max} - Z_{min})} \quad (3)$$

The corresponding reconstruction filter Z_c is calculated as

$$Z_c = \frac{2Z_{max} Z_{min}}{Z_{max} + Z_{min}} \quad (4)$$

2.2. Surface plenoptic function and spectrum of the occlusive scene

In the ideal scene, the intensity of the EPI line maintains the same because of the Lambertian reflection and no-occlusion. However, it is incorrect in more complicated scenes. Different SPF representations [4,6,9,16], are presented to solve this problem. Take the paper [4] for example, as in Fig. 3, the intensity of the plenoptic function can be related as the light intensity emitted from the object surface:

$$p(t, v) = l(s, \theta) \quad (5)$$

Zhang and Chen [4] defined a silhouette light field to analyze the spectrum of scenes with occlusion.

$$sil_i(t) = \begin{cases} 1 & \text{when ray } l(t, v) \text{ comes from the object } i \\ 0 & \text{object } i \text{ is occluded by other objects} \end{cases} \quad (6)$$

Suppose object i is occluded by M other objects, the spectrum of the occluded object i is

$$L_i^{occ}(\Omega_v, \Omega_t) = L_i(\Omega_v, \Omega_t) * [\delta(\Omega_v, \Omega_t) - SIL_0(\Omega_v, \Omega_t)] * \dots * [\delta(\Omega_v, \Omega_t) - SIL_{M-1}(\Omega_v, \Omega_t)] \quad (7)$$

where $*$ stands for convolution, $L_i(\Omega_v, \Omega_t)$ is the Fourier transform

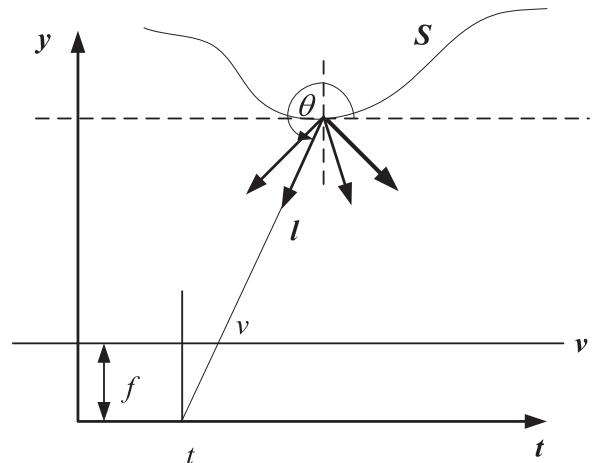


Fig. 3. The surface plenoptic function and the 2D light field.

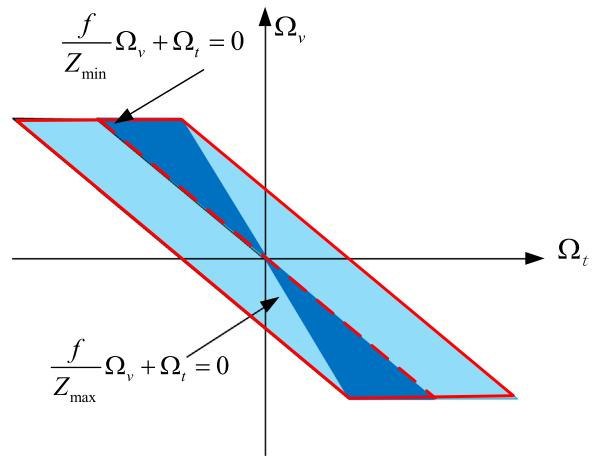


Fig. 4. The worst case of the occlusive spectrum in paper [4].

of the object i , $\delta(\Omega_v, \Omega_t)$ is the Fourier transform of constant 1, and SIL is the Fourier transform of the silhouette lightfield. The spectrum of an occluded object is modulated by the occluding object's silhouette spectrum, which broadens the final spectrum.

In their conclusion, the worst case of occlusion happens in the minimum depth and the spectrum is shown in Fig. 4. Compared with Chai's work [3], the maximum sampling rate is doubled, and the depth of the reconstruction filter becomes Z_{min} instead of Z_c in Eq. (4).

3. The EPI model

In Section 2, we covered the derivation of the state-of-the-art plenoptic sampling theory. We figured out that the plenoptic sampling theory is an extension of the Shannon sampling theory to the light field case. For the ideal scene with Lambertian reflection and no-occlusion, the signal is the 2D EPI image. When it comes to more complicated scenes with non-Lambertian reflection or occlusion, the early analysis is based on the SPF. However, the SPF is less intuitive compared with the EPI image. Considering the characteristics of the EPI, for Lambertian reflection and no-occlusion scenes, the intensity of the EPI line maintains the same. For non-Lambertian reflection or occlusion in the scenes, the intensity of EPI line changes and the spectrum broadens. The plenoptic spectrum can be analyzed based on the Fourier transform of the EPI lines. Owing to this, an EPI model is explored to analyze the plenoptic spectrum of the complicated scenes.

¹ In fact, the maximum frequency ω_v is impacted by the highest frequency of texture and the camera resolution. We assume that the highest frequency is bounded by the resolution of the capturing camera, and take $\omega_v = \frac{\pi}{\Delta v}$.

In previous works, a no-occlusion constraint [10] was defined to restricted the scene as

$$|Z'(x)| < \frac{f}{v_m} \quad (8)$$

where $Z'(x)$ is the derivative of $Z(x)$, which defines the slope of the scene. v_m is the maximum Field of View (FOV) of camera. This equation is valid when the camera has a finite FOV. However, as the previous plenoptic sampling analysis is usually with a latent assumption that the FOV is infinite, the no-occlusion constraint is only valid when the scene is a constant depth scene. Thus, without loss of generality, we will study the plenoptic spectrum under the constraint that the cameras are with finite FOV.

Before our further analysis, we first clarify that, the spectrum of the plenoptic function is analyzed from two aspects. On one hand, the main part of the spectrum is bounded by the minimum and maximum depths. On the other hand, the spectral broadening area is introduced by some factors. As the plenoptic spectrum is *bow-tie* shaped, the broadening generally happens in the maximum or minimum depth area. This claim is based on the property of the plenoptic spectrum, since we already know that the plenoptic spectrum of the ideal scene is bounded by the minimum and maximum depths and the spectrum is broadened as in the non-ideal scene. As the *bow-tie* shape of the spectrum, the broadening only happens in the minimum depth or the maximum depth. Our claim covers the spectrum form the ideal scene to the non-ideal scene. The spectral analysis of the complicated scene can be simplified to analyze the spectral broadening. In this paper, the spectral broadening is analyzed by the Fourier transform of the EPI line with the minimum or maximum depth. Specifically, the analysis is based on the modeling of the EPI lines.

The EPI model is presented to analyze the spectral broadening. As the plenoptic spectrum is analyzed by the Fourier transform of the EPI image, the information in the EPI image indicates the variation of the plenoptic spectrum. Thus, an intuitive way to analyze the spectrum is to model the EPI image.

Since Fourier transform is linear, the spectral support of the EPI image is simply the sum of the spectrum of individual EPI lines. A straight way to analyze the spectrum is to model the EPI lines. Thus, one single line of the EPI image is considered as the basic analysis unit. On the other hand, the Fourier transform is rotationally invariant, which means that the rotation of the EPI line does not change the bandwidth of the signal. Then we can consider the signal only in one dimension by defining the intensity of the EPI lines as $f(t)$. To validate the EPI model, some special scenes are analyzed as follows.

Lambertian reflection scene: In the Lambertian reflection scene, the intensity of a EPI line maintains the same, defined as $f(t) = c$, where c is a constant. The Fourier transform of this function is a delta function, which means that the bandwidth is not broadened. So in an ideal scene with Lambertian reflection, the spectrum does not broaden. This conforms to the common analysis [3].

Constant depth scene: As mentioned before, the general analysis of the plenoptic sampling assumes the capture cameras are with infinite FOV. The spectrum of the constant depth scene becomes a line and one camera with infinite FOV is enough to sample the constant depth scene. Considering the camera with finite FOV, the maximum FOV is v_m ($v_m > 0$). The EPI image of a constant depth Z is a set of parallel lines with slope f/Z and finite length $\frac{2v_m Z}{f}$. For an EPI line, the signal can be modeled as

$$f(t) = \begin{cases} c, & -\frac{v_m Z}{f} < t < \frac{v_m Z}{f} \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

The corresponding Fourier transform is

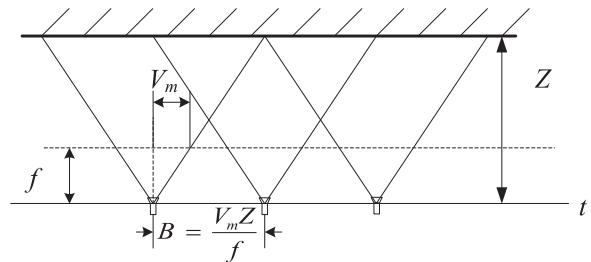


Fig. 5. The sampling spacing of the constant depth plane.

$$\mathcal{F}(Q_t) = 2c \frac{\sin \frac{v_m Z Q_t}{f}}{Q_t} \quad (10)$$

As (10) is a Sinc function, the spectrum is band-unlimited. Its essential bandwidth is approximately equal to the width of its main lobe, and the width of the main lobe is $B = \frac{2\pi f}{v_m Z}$. The camera spacing of a constant depth scene with finite FOV is

$$\Delta t = \frac{2\pi}{B} = \frac{v_m Z}{f} \quad (11)$$

As shown in Fig. 5, this camera spacing would be increased as the depth or the maximum FOV increases. However, as the cameras are fixed, there exists a maximum depth which makes sure the texture could be distinguished.

This conclusion can be extended to the common case. With a finite FOV camera, for the spectrum of no-occlusion and Lambertian reflection scene, the spectrum is constructed from two parts. The main part is bounded by the minimum and maximum depths. The broadening part is caused by the finite FOV, and its broadening range is

$$B_{FFOV} = \frac{\pi f}{v_m} \left(\frac{1}{Z_{max}} + \frac{1}{Z_{min}} \right) \quad (12)$$

This broadening happens because of the finite FOV, and the spectral broadening range is small compared with the main part of the spectrum.

In next section, we will analyze the EPI of the occlusive scene with finite FOV cameras and investigate the bandwidth of the spectrum.

4. The spectrum of the occlusive scene

The occlusion widely exists in the 3D scene, and the spectral analysis is not very clear before. Considering the EPI of the occlusive scene, an occlusive scene with two constant planes and its corresponding EPI image are shown in Fig. 6. The EPI image is constructed by the pixels on the center horizontal line. The EPI image is comprised by a group of EPI lines, and the slope of a line is inversely proportional to the depth of the corresponding point. In the occlusive scene, the points close to the camera occlude the points which are more distant. Thus, the EPI lines with steeper gradients always occlude lines with shallower gradients. So when the occlusion happens, the intensity of the EPI line at the occluded depths changes as the intensity of the occluding object, and the spectrum is broadened at the maximum depth area.

Suppose a scene is constructed by two constant planes with two depth, Z_{min} and Z_{max} , respectively. The depth difference is ΔZ , and the far plane is occluded by the near plane. This is the worst case of occlusion in paper [4]. We also define the length of the near plane is L . The EPI image is made up of two groups of parallel lines, as illustrated in Fig. 7. We assume a single sinusoid $\sin(\Omega t)$ is



(a) An occlusive scene. (b) The corresponding EPI image.

Fig. 6. An occlusive scene with two constant planes and the corresponding EPI image.

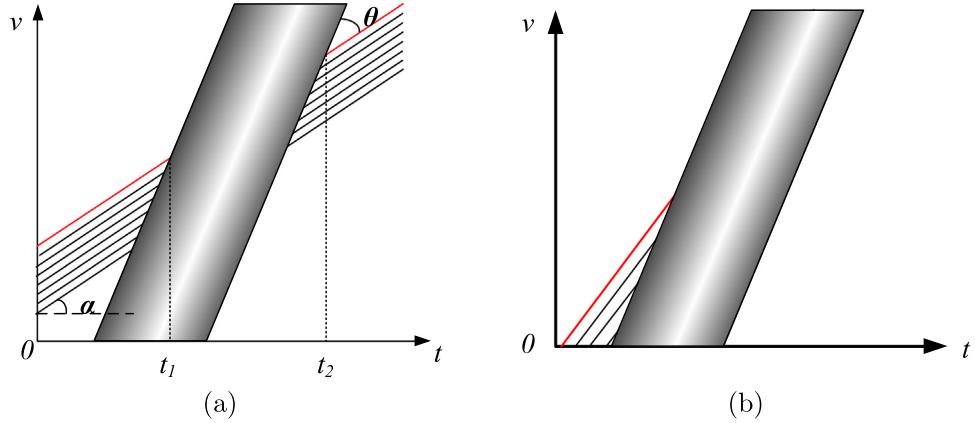


Fig. 7. The illustration of the EPI image with two cases of occlusion.

pasted on the near plane. The change of intensity in the maximum depth EPI line is modeled as

$$f(t) = c + \sin(\Omega \sin \theta t) \times \omega(t) \quad (13)$$

where $\theta (0 < \theta < 90^\circ)$ is the included angle between two depth of EPI lines. As shown in Fig. 7, we have

$$\sin \theta = \frac{\Delta Z}{\sqrt{(f^2 + Z_{min}^2)(f^2 + Z_{max}^2)}} \quad (14)$$

The space point we observe is occluded by the near plane point. There are two cases of the red line shown in Fig. 7. In the first case, the occlusion exists in a finite area. As shown in Fig. 7(a), the occlusion happens at the position between t_1 and t_2 , where $t_2 - t_1 = T$. In the second case (Fig. 7(b)), the occlusion never ends, that is, the point on the far plane is totally occluded by the near plane. The function $\omega(t)$ becomes a Heaviside step function, thus no essential bandwidth exists. This means for the camera with finite FOV, the scene cannot be represented when the occlude object has a large size. In order to represent the whole scene, a solution is to increase FOV or the distance from the object to the camera. Therefore, we will not consider this case because of the unsure bandwidth. For the first case, the occluding plane size must satisfy

$$L \leq \frac{2v_m \Delta Z}{f} \quad (15)$$

Without loss of generality, we define the function as

$$\omega(t) = \begin{cases} 1, & |t| < \frac{T}{2} \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

The length of the occluding plane is L , we have

$$T = \frac{LZ_{max}}{\Delta Z \cos \alpha} \quad (17)$$

where α is the slant angle of the far plane, and

$$\cos \alpha = \frac{Z_{max}}{\sqrt{f^2 + Z_{max}^2}} \quad (18)$$

The spectrum is band-unlimited as there is a finite function in the time domain. And in the frequency domain, the spectrum is the convolution of the Fourier transform of a sinusoid function and a rectangular window function. Then the essential bandwidth is

$$B_e = 2 \left(\Omega \sin \theta + \frac{2\pi}{T} \right) = \frac{2\Delta Z f \Omega}{\sqrt{(f^2 + Z_{max}^2)(f^2 + Z_{min}^2)}} + \frac{4\pi \Delta Z}{L \sqrt{f^2 + Z_{max}^2}} \quad (19)$$

The spectrum of the broadening range is

$$B = \frac{1}{2} B_e \frac{\sqrt{f^2 + Z_{max}^2}}{Z_{max}} = \frac{\Delta Z f \Omega}{Z_{max} \sqrt{f^2 + Z_{min}^2}} + \frac{2\pi \Delta Z}{L Z_{max}} \quad (20)$$

Ω is defined as the frequency of the texture information on the near plane. Taking account of the texture with maximum frequency B_v^s , the broadening range can be described as

$$B = \frac{\Delta Z f B_v^s}{Z_{max} \sqrt{f^2 + Z_{min}^2}} + \frac{2\pi \Delta Z}{L Z_{max}} \quad (21)$$

usually $Z_{min} \gg f$, we have

$$B = \frac{\Delta Z}{Z_{max}} \left(\frac{f B_v^s}{Z_{min}} + \frac{2\pi}{L} \right) \quad (22)$$

The broadening range will be smaller if the texture variation on the occluding object gets more uniform. One way to reduce B_v^s is to apply a low-pass filter to the light field capturing. As the size of the occluding object increases, the window size in time domain increases, and the corresponding spectrum range decreases. The broadening range is also determined by the depth difference of the occluding. The larger the focus length f , the more details the camera covers, the wider the broadening range. We find out that the broadening range is not relative with the camera resolution.

We analyze the spectrum of a occlusive scene with two planes, the spectrum is illustrated in Fig. 8. The scene set is the worst case of occlusion in paper [4]. The spectral broadening happens in the occluded depths area. Extending to the normal occlusive scene, as the *bow-tie* shape of the spectrum, the broadening is usually covered by the nearby EPI lines except the broadening area that exceeds the original spectrum. The broadening also reflects at the maximum depth area. As the cameras are determined, the parameters that impact the broadening range are the occluding depth difference, the size of the occluding object and the maximum texture frequency painted on the occluding object.

5. The sampling spacing of the occlusive scene

In this section, we will first study the sampling spacing of the occlusive scene. Then, since the camera samples the scene with the known focal plane, the corresponding reconstruction filter as well as the sampling spacing is further considered.

5.1. The plenoptic sampling spacing

From our analysis, the plenoptic spectrum of occlusive scene is band-unlimited, and the spectrum broadens in the maximum

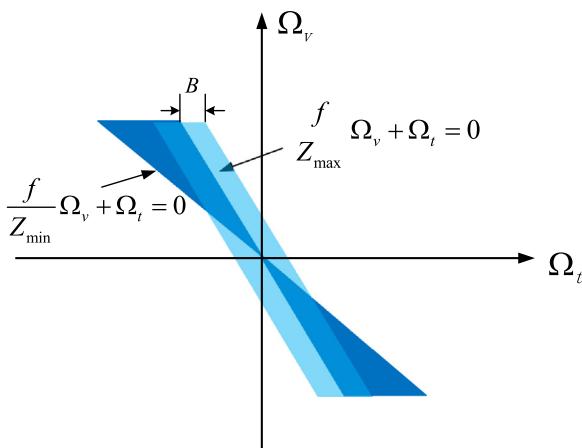


Fig. 8. An illustration of the spectrum for the scene of two occlusive planes.

depth area compared with ideal scenes. In this subsection, we obtain the maximum sampling spacing and the corresponding reconstruction filter for occlusive scenes.

The spectrum broadening is analyzed in (21), and the maximum sampling spacing is

$$\Delta t_{max} = \frac{2\pi}{|P_{max} P_{min}| + B} \quad (23)$$

Substituting (2) and (21) into (23), we have

$$\Delta t_{max} = \frac{2\pi \Delta v L Z_{min} \sqrt{f^2 + Z_{min}^2}}{\Delta Z \left(\pi f L \sqrt{f^2 + Z_{min}^2} + \Delta v f L B_v^s Z_{min} + 2\pi \Delta v Z_{min} \sqrt{f^2 + Z_{min}^2} \right)} \quad (24)$$

The corresponding depth of reconstruction filter is

$$Z_{rec} = \frac{f}{\frac{f}{Z_c} - \frac{B \Delta v}{2\pi}} \quad (25)$$

Substituting Eqs. (4) and (21) into (25), we have

$$Z_{rec} = \frac{2\pi f L Z_{max} Z_{min} \sqrt{f^2 + Z_{min}^2}}{\pi f L (Z_{max} + Z_{min}) \sqrt{f^2 + Z_{min}^2} - \Delta Z \Delta v Z_{min} (f L B_v^s + 2\pi \sqrt{f^2 + Z_{min}^2})} \quad (26)$$

Obviously, $Z_{rec} > Z_c$, for simplicity, Z_{mean} is used to approximate the depth of reconstruction filter, and

$$Z_{mean} = \frac{Z_{max} + Z_{min}}{2} \quad (27)$$

Fig. 9 is the illustration of three sampling methods. The yellow one is the classic plenoptic sampling [3] without considering the occlusion. Distinctly, it loses some high frequency information. The black one is the conclusion of paper [4]. Compared with paper [3], the black one doubles the sampling spacing and chooses the minimum depth Z_{min} as the depth of reconstruction filter. In our analysis, the spectrum broadens in the maximum depth area. Thus, the filters listed above miss part of the spectrum. The red one is our sampling method. As shown in the figure, our filter covers all the spectrum. And the sampling spacing we choose is bigger than paper [4], which means we only need less cameras to capture the scene.

5.2. The sampling for the occlusive scene with the known focal plane

In this subsection, we consider the sampling for the occlusive scene with the known focal plane. When the camera samples the scene, the focus plane needs to be determined. Usually, the optimal depth of reconstruction filter we obtain in Eq. (27) can be used as guidance for selecting the focal plane, but this is not very suitable for the occlusive scene. As in the occlusive scene, the depth of field (DOF) is large, and focus length is usually on the minimum depth plane or the object plane in ROI, which makes the reconstruction on this plane

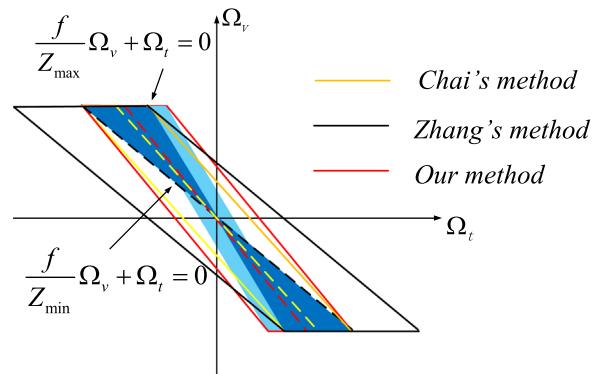


Fig. 9. Three sampling methods of the scene with occlusion.

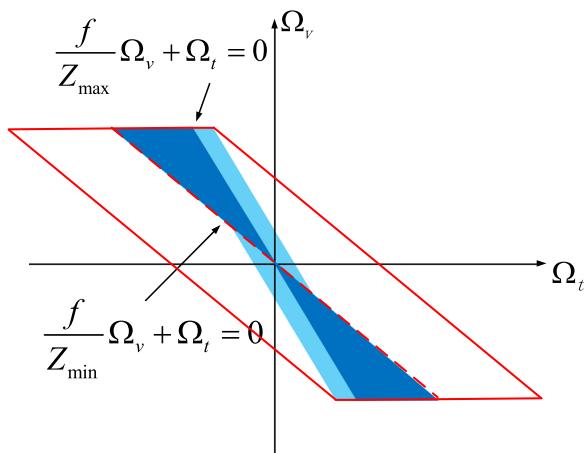


Fig. 10. The reconstruction filter with camera focus on the minimum depth plane.

is more meaningful. So, in this case, we choose the focal plane as the reconstruction filter depth. As we cannot choose the optimal reconstruction filter, the sampling is not optimal. Instead, it is a sub-optimal result depends on the spectrum and the depth of focal plane.

As in Fig. 10, an illumination of the reconstruction filter with the minimum depth and the sampling rate is shown. For the occlusive scene, the broadening happens in the maximum depths area. The sampling rate Δt_{max} is much higher than the twice of that of the ideal scene, and

$$\Delta t_{max} = \frac{\pi \Delta v L Z_{min} Z_{max} \sqrt{f^2 + Z_{min}^2}}{\Delta Z \left(\pi f L \sqrt{f^2 + Z_{min}^2} + \Delta v f L B_v^S Z_{min} + 2\pi \Delta v Z_{min} \sqrt{f^2 + Z_{min}^2} \right)} \quad (28)$$

The reconstruction filter with the constant depth will lead to blurriness due to DOF. Thus the light field is only suitable for storing scenes with an approximately constant depth [17], especially for the camera with focal plane on the minimum depth. For sampling the scene with large DOF, the effective way is to use the prefiltering of the light field [13] and extract more layers for the scene [18]. Our plenoptic spectrum analysis and reconstruction filter design for the occlusive scene can be easily extended to the sampling of complicated scenes.

6. Experiments

In this section, we evaluate our proposed sampling method for the

plenoptic function of the occlusive scene. Specifically, the camera spacing and the corresponding filter are evaluated. We perform this evaluation by sampling and reconstructing a plenoptic function generated by both synthetic and real scenes. The camera sampling rate is uniform and the reconstruction filter is a depth-corrected bilinear interpolation filter. Note that the sampling rate will be determined by the reconstruction algorithm as well [19]. We use the same kind of filter to examine the effectiveness of the sampling.

In the following analysis, the samples are captured with the camera spacing in paper [3,4] and our result in Eq. (24), and the depth of reconstruction filters are Z_c , Z_{min} and Z_{mean} , respectively. Then, the PSNRs of the synthesized views are measured.

6.1. Synthetic scenes

We set two synthetic scenes by the same cameras with $f=28$ mm and $FOV=60^\circ$. The scene is set with $Z_{min}=60$ mm, $Z_{max}=125$ mm. The scene 1 is simple, which consists of two constant depth planes. The scene 2 is a complex scene with 5 layers, each has $1, 3 \times 3, 5 \times 5, 7 \times 7$ and 9×9 cubes. The scenes are shown in Fig. 11.

The sampling and reconstruction results of the scene 1 are shown in Table 1. The Chai's sampling spacing is the biggest as not considering the occlusion. Zhang doubles the sampling rate, but the PSNR of the synthesized view is lower than Chai's. The reason is that the scene is constructed with only two planes, the depth of Zhang's reconstruction filter is Z_{min} which results in that the focus plane is on the near plane and the far plane is with a mass of defocus phenomena. Our sampling spacing is bigger than Zhang's, which means our sampling rate is smaller, and our PSNR of synthesized view is the best. The reason for this behavior is the right choice of the reconstruction filter. As in Fig. 9, the spectral support of the reconstruction filter with constant depth is determined jointly by the depth of the reconstruction filter and the sampling spacing. In particular, the spectral support comprises the band of which the center is a dash line determined by the corresponding depth, and the width of the band is depended on the sampling spacing. The experiment result verifies the effectiveness of our sampling method.

Similar results of the scene 2 are shown in Table 2. The objective quality of the synthesized views proves that our sampling and reconstruction filter is the best. Fig. 12 shows the synthesized views of the scene 2 with different sampling spacing and reconstruction filters. Obviously, our subjective quality is the best for the entire scene.

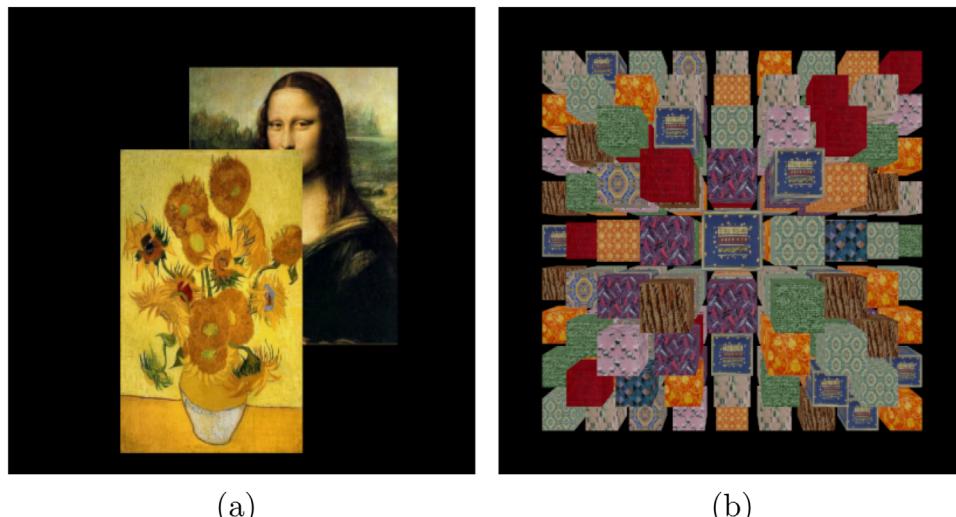


Fig. 11. Two synthetic occlusive scenes: (a) a simple scene with two constant depth planes, (b) a complex scene with 5 layers.

Table 1

The sampling and synthesis of the synthetic scene 1.

Method	Chai's [3]	Zhang's [4]	Ours
Sampling spacing (mm)	0.74	0.37	0.55
The depth of reconstruction filter (mm)	81.08	60	92.5
PSNR (dB) of the synthesis views	29.2498	29.0978	29.2722

Table 2

The sampling and synthesis of the synthetic scene 2.

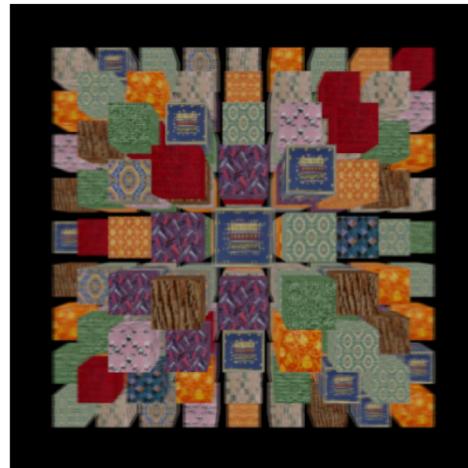
Method	Chai's [3]	Zhang's [4]	Ours
Sampling spacing (mm)	2.47	1.13	1.47
The depth of reconstruction filter (mm)	81.08	60	92.5
PSNR (dB)	25.7525	25.4520	26.5384

6.2. Real scenes

We now analyze our sampling and reconstruction filter with the real scenes. The general cameras have a fixed focus plane when capturing scenes, which impacts the choice of the reconstruction filter. In the experiment, we use a Lytro camera [20] to capture the scene and obtain the all-in-focus images. The all-in-focus image is the image where all the visible objects are focused. The scene is set



(a)



(b)



(c)



(d)

Fig. 12. The synthetic scene: (a) the ground truth of occlusion, (b) use the minimum sampling rate in paper [3], rendered with Z_c , (c) use the sampling rate in paper [4], rendered with Z_{min} , (d) use the sampling rate in Eq. (24), rendered with Z_{mean} .

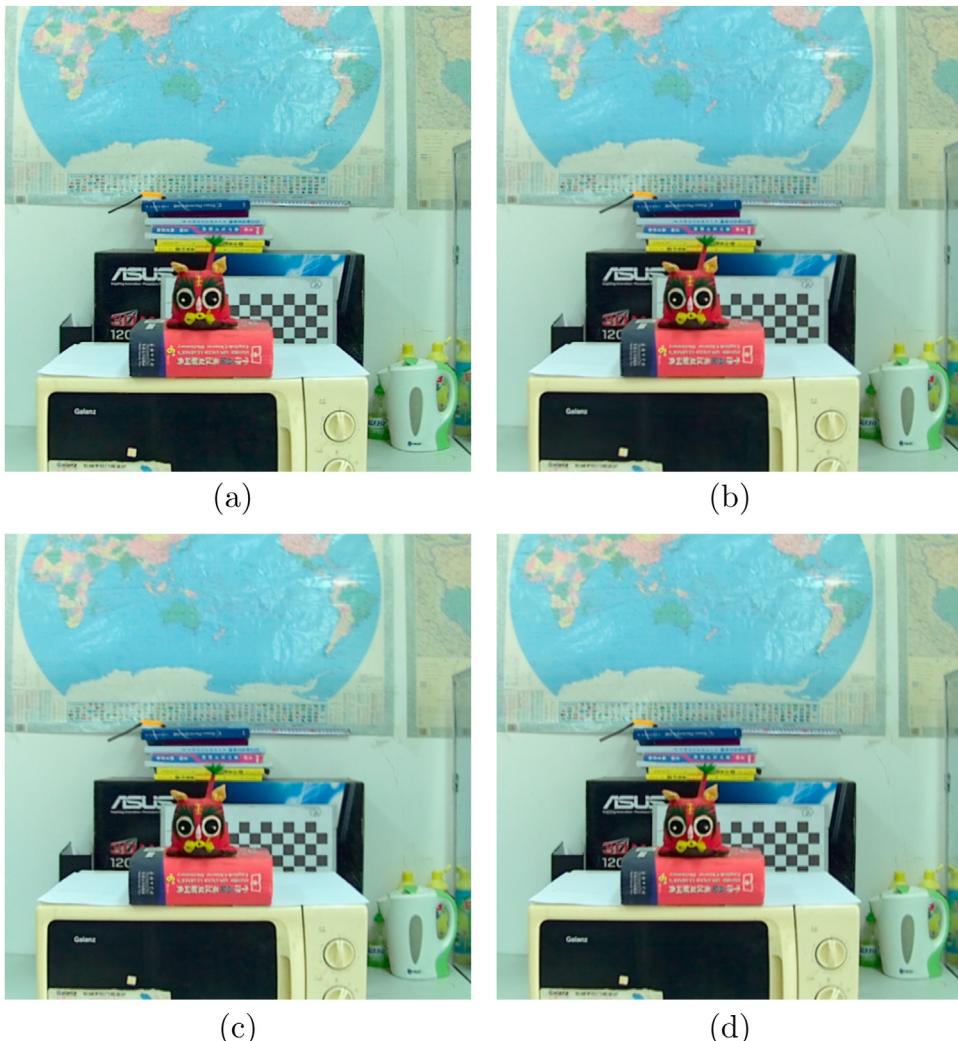


Fig. 13. The real scene: (a) the ground truth of occlusion, (b) use the minimum sampling rate in paper [3], rendered with Z_c , (c) use the sampling rate in paper [4], rendered with Z_{min} , (d) use the sampling rate in Eq. (24), rendered with Z_{mean} .

which is modeled by the EPI image with the finite field of view camera. From our analysis, the spectral broadening of the occlusion scene happens in the maximum depth area and is impacted by the occluding depth difference, the size of the occluding object and the maximum frequency of the texture painted on the occluding object. We then determined the uniform plenoptic sampling results for the occlusive scene. In particular, we derived a new reconstruction filter and a new expression for maximum camera spacing. Experiments on both synthetic and real scenes prove the validity of our method. The analysis of the spectrum of occlusive scene in this paper extends the plenoptic spectrum analysis and benefits the design of new IBR algorithms.

Acknowledgments

This work is supported in part by the National Science Foundation of China under Grant nos. 61231010, and 61202301.

References

- [1] H.-Y. Shum, S.-B. Kang, S.-C. Chan, Survey of image-based representations and compression techniques, *IEEE Trans. Circuits Syst. Video Technol.* 13 (11) (2003) 1020–1037.
- [2] C. Zhang, T. Chen, A survey on image-based rendering-representation, sampling and compression, *EURASIP Signal Process.: Image Commun.* 19 (2004) 1–28.
- [3] J.-X. Chai, X. Tong, S.-C. Chan, H.-Y. Shum, Plenoptic sampling, in: Proceedings of SIGGRAPH, New Orleans, 2000, pp. 307–318.
- [4] C. Zhang, T. Chen, Spectral analysis for sampling image-based rendering data, *IEEE Trans. Circuits Syst. Video Technol.* 13 (11) (2003) 1038–1050.
- [5] M.N. Do, D. Marchand-Maillet, M. Vetterli, On the bandlimitedness of the plenoptic function, in: Proceedings of IEEE International Conference on Image Processing, Genova, 2005.
- [6] M.N. Do, D. Marchand-Maillet, M. Vetterli, On the bandwidth of the plenoptic function, *IEEE Trans. Image Process.* 21 (2) (2012) 708–717.
- [7] C. Gilliam, P. Dragotti, M. Brookes, A closed-form expression for the bandwidth of the plenoptic function under finite field of view constraints, in: Proceedings of IEEE International Conference on Image Processing, Hong Kong, 2010, pp. 3965–3968.
- [8] C. Gilliam, P. Dragotti, M. Brookes, On the spectrum of the plenoptic function, *IEEE Trans. Image Process.* 23 (2) (2014) 502–516.
- [9] P. Zhou, L. Yu, G. Zhong, The non-Lambertian reflection in plenoptic sampling, in: Proceedings of IEEE International Conference on Image Processing, Melbourne, 2013, pp. 2154–2157.
- [10] C. Chen, D. Schonfeld, Geometrical plenoptic sampling, in: Proceedings of IEEE International Conference on Image Processing, Cairo, 2009, pp. 3769–3772.
- [11] L. Bagnato, P. Frossard, P. Vandergheynst, Plenoptic spherical sampling, in: Proceedings of IEEE International Conference on Image Processing, Orlando, 2012, pp. 375–380.
- [12] E.H. Adelson, J.R. Bergen, The plenoptic function and the elements of early vision, in: Computational Models of Visual Processing, 1991, pp. 3–20.
- [13] M. Levoy, P. Hanrahan, Light field rendering, in: Proceedings of SIGGRAPH, New Orleans, 1996, pp. 31–42.
- [14] J. Gortler, Steven, R. Grzeszczuk, R. Szeliski, F. Cohen, Michael, The lumigraph,

- in: Proceedings of SIGGRAPH, New Orleans, 1996, pp. 43–54.
- [15] R.C. Bolles, H.H. Baker, D.H. Marimont, Epipolar-plane image analysis: an approach to determining structure from motion, *IJCV* 1 (1) (1987) 7–55.
- [16] D. Wood, D. Azuma, K. Aldinger, B. Curless, T. Duchamp, D. Salesin, Surface light fields for 3d photography, in: Proceedings of SIGGRAPH, New Orleans, 2000, pp. 287–296.
- [17] A. Isaksen, L. McMillan, S.J. Gortler, Dynamically reparameterized light fields, in: Proceedings of SIGGRAPH, New Orleans, 2000, pp. 297–306.
- [18] J. Pearson, M. Brookes, P. Dragotti, Plenoptic layer-based modeling for image based rendering, *IEEE Trans. Image Process.* 22 (9) (2013) 3405–3419.
- [19] C. Zhang, T. Chen, Light Field Sampling, vol. 2, 2006, <http://dx.doi.org/10.2200/S00035ED1V01Y200601VM006>.
- [20] R. Ng, M. Levoy, M. Brädf, G. Duval, M. Horowitz, P. Hanrahan, Light field photography with a hand-held plenoptic camera, in: Stanford Technical Report, 2005.



Peng Zhou received the B.S. degree in Communication Engineering from Taiyuan University of Technology in 2004, the M.S. degree in Communication and Information System from Yangtze University, Jingzhou, China in 2009. He is a Ph.D. candidate in Huazhong University of Science and Technology. His research interests are plenoptic sampling and spectral analysis, currently focusing on sampling theory in 3D.



Li Yu received the B.S. degree in Electronics and Information Engineering, the M.S. degree in Communication and Information System and the Ph.D. degree in Electronics and Information Engineering, all from Huazhong University of Science and Technology (HUST), Wuhan, China, in 1995, 1997, and 1999, respectively. In 2000, she joined the Electronics and Information Engineering Department, HUST, where she has been a professor since 2005. She is a co-sponsor of China AVS standard special working group and working as the key member of China AVS standard special working group. Her team has applied more than 10 related patents and submitted 79 proposals to AVS standard organization. Her current research interests include multimedia communication and processing, computer network, wireless communication.



Gang Zhong received a B.S. degree in communication engineering from Huazhong University of Science and Technology, Wuhan, China, in 2008. His research interests are video coding and image processing, currently focusing on depth image super resolution.