A CLOSED-FORM EXPRESSION FOR THE BANDWIDTH OF THE PLENOPTIC FUNCTION UNDER FINITE FIELD OF VIEW CONSTRAINTS

Christopher Gilliam, Pier Luigi Dragotti, and Mike Brookes

Electrical and Electronic Engineering Department, Imperial College London, UK

ABSTRACT

The plenoptic function enables Image-based rendering (IBR) to be viewed in terms of sampling and reconstruction. Thus the spatial sampling rate can be determined through spectral analysis of the plenoptic function. In this paper we examine the bandwidth of the plenoptic function when both the field of view and the scene width are finite. This analysis is carried out on two planar Lambertian scenes, a fronto-parallel plane and a slanted plane, and in both cases the texture is bandlimited. We derive an exact closed-form expression for the plenoptic spectrum of a slanted plane with sinusoidal texture. We show that in both cases the finite constraints lead to band-unlimited spectra. By determining the essential bandwidth, we derive a sampling curve that gives an adequate camera spacing for a given distance between the scene and the camera line.

Index Terms— Plenoptic function, Image-Based Rendering, spectral analysis, bandwidth, sampling.

1. INTRODUCTION

Visual media is currently undergoing the most extensive change in decades, the move from 2D to 3D scene representation. One such example is free viewpoint TV (FTV), where the user selects the viewpoint and orientation of the camera. This type of 3D scene representation can be achieved using Image-Based Rendering (IBR), where an arbitrary new viewpoint of a scene is rendered from a multi-view image set. An important goal in IBR is to minimise the size of the multi-view image set whilst still achieving good quality rendering.

The principle behind IBR is that each image is considered as capturing a set of light rays travelling from a scene to the camera [1]. The light rays in question are described using a 7D function, known as the *plenoptic function* [2]. It specifies the intensity of the light ray passing through the camera centre at a 3D spatial location for a given viewing direction, wavelength and time [3]. IBR can therefore be seen as the problem of sampling and interpolating the plenoptic function. That is, a finite set of images, with finite resolution, samples the continuous plenoptic function and the rendering of a new viewpoint is the reconstruction from the samples [1]. As a result the minimum spatial sampling rate can be determined by spectral analysis of the plenoptic function, which in turn allows the size of the image set to be minimised.

Chai et al. [4] performed the first spectral analysis for the plenoptic function. Assuming a Lambertian surface with no occlusion, they showed that the spectrum is approximately bounded by lines relating to the maximum and minimum depths of the scene and that finite camera resolution bandlimits the spectrum. This led to an optimal uniform sampling rate and reconstruction filter. Zhang and Chen [5] extended this spectral analysis to more general cases, in particular non-Lambertian and occluded scenes. Using a functional scene model, [6] re-examined the spectral support of the continuous plenoptic function proving that, for a Lambertian scene, the plenoptic spectrum is exactly bound by minimum and maximum depths of the scene. However they went on to show the plenoptic spectrum is band-unlimited unless the scene surface is flat.

In this paper, we examine the bandwidth of the plenoptic spectrum for a Lambertian scene, approximated using planar facets, under two constraints: the first is that the cameras have a finite field of view (FFoV) and the second is finite scene width (FSW). [6] assumes a FFoV in their analysis, however it is only used to enforce a no-occlusion constraint. Therefore our work differs by applying the FFoV to the Fourier analysis of the plenoptic function. Specifically, we examine the plenoptic spectrum for two finite scenes, a slanted plane and a fronto-parallel plane (FPP). We show that applying FFoV and FSW to both types of scene results in band-unlimited spectra. In fact, even under the Lambertian assumption, neither plenoptic spectrum is bounded by lines relating to the depth (or depths) present in the scene. Lastly, we use the idea of the essential bandwidth [7] to derive equations for an adequate camera spacing.

The outlines of this paper is as follows. Section 2 covers the modelling of the scene using the functional framework presented in [6]. In Section 3, we derive the exact close form expression for the plenoptic spectrum of a slanted plane with sinusoidal texture. We then define an adequate reconstruction filter for this plenoptic spectrum. In Section 4, we define the essential bandwidth for a FPP with bandlimited texture. From this, we derive a sampling curve for the camera spacing as the distance from the scene to the cameras varies, and show some reconstruction simulations. Lastly, conclusions and future work are given in Section 5.

2. PLENOPTIC FUNCTION AND SCENE MODEL

The plenoptic function can be parameterised in various ways [1]. For this paper, we will consider a common parameterisation, the light field [8] or lumigraph [9]. This is a two plane parameterisation of the plenoptic function, in which the scene is bounded within a box and each light ray is described using its intersection with two parallel planes, the image plane, (v, y), and camera plane, (t, u). The distance between the two planes is the focal length, f. Therefore $p(t, u, v, y)$ is the intensity of the light ray (t, u, v, y) at camera location (t, u) and pixel location (v, y) [6].

A further simplification, used in [4, 6], is to fix u and y, corresponding to the situation where the camera positions are constrained to a 1D line and only one scan-line is considered in each image. In this case the light field is reduced to two dimensions, $p(t, v)$. Plotting this plenoptic representation in the (t, v)-space leads to the Epipolar Plane Image (EPI). Using this parameterisation, the plenoptic spectrum is defined as $P(\omega_t, \omega_v) = \mathcal{F}_{t,v} \{p(t, v)\}$, where $\mathcal{F}$ is the Fourier transform operator.

We will now define the scene model using the framework presented in [6]. The scene model consists of three functions that are
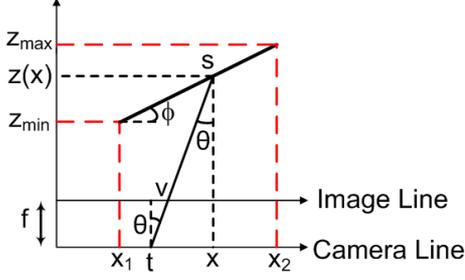
Fig. 1. Scene model of a slanted plane showing the intersection of a light ray (t, v) with the object surface at $(x, z(x))$, where s is the curvilinear coordinate, f is the focal length, z_{min} and z_{max} are the minimum and maximum depth, respectively. Note that θ is the viewing angle and ϕ is the slant of the plane.

collectively known as the scene geometry equations, $\mathcal{G}_s$. The first function, $z(x)$, models the scene surface, hence z is the depth of the surface at a coordinate point x. Note that x is the projection of the surface point onto the camera line, see Figure 1. The next function, $x(s)$, describes the relationship between x and the surface curvilinear coordinate, s. Lastly, $g(s)$, models the texture signal pasted to the scene surface. The texture signal is bandlimited to a frequency ω_s. Moreover the scene is assumed to have a finite width, T, therefore $s \in [0, T]$. Having defined the scene geometry, a light ray arriving at (t, v) can be related to its point of origin on the scene surface using the $(x, z(x))$ coordinates as follows,

$$t = x - z(x)\tan(\theta) = x - z(x)\frac{v}{f}, \tag{1}$$

where θ is the viewing angle. Applying FFoV restricts v such that $v \in [-v_m, v_m]$. Figure 1 shows an example of a scene model and the relationship between a light ray emanating from a point on the surface and its intersection with the t and v lines. The importance of this framework is that it allows the plenoptic function, $p(t, v)$, to be related to the surface light field [5], $l(x, \theta)$, which is the intensity of a light ray emitted from a point x at a viewing angle θ [6]. To enforce the no-occlusion condition and make (1) a one to one mapping, [6] places the following constraint on $z(x)$

$$|z'(x)| < \frac{f}{v_m}, \tag{2}$$

where $z'(x)$ is the first derivative of z with respect to x. Note that this framework is similar to that used in [5], however they allow occlusions to occur. The next section uses the above framework to derive the plenoptic spectrum for a slanted plane with sinusoidal texture.

3. SPECTRAL ANALYSIS OF SLANTED PLANE

A spectral analysis of a slanted plane with sinusoidal texture was presented in [5]. However no constraint was placed on the width of the plane, resulting in a solution that is limited to a certain region in ω_t but independent of ω_v. In order to bound their solution in ω_v, they assumed the approximate result given in [4]. In the following discussion we will derive the plenoptic spectrum for a finite slanted plane, therefore avoiding the need to use the approximation solution from [4] to bound the spectrum.

The scene geometry equations for a slanted plane with sinusoidal texture are as follows

$$\mathcal{G}_s = \begin{cases} x = s\cos(\phi) + x_1 \\ z(x) = s\sin(\phi) + z_{min} \\ g(s) = \sin(\omega_s s) \end{cases} \tag{3}$$

where $x \in [x_1, x_2]$, $z \in [z_{min}, z_{max}]$ and ϕ is the angle between the plane and the line $z = z_{min}$. The finite width of the plane is

$$T = \frac{x_2 - x_1}{\cos(\phi)} = \frac{z_{max} - z_{min}}{\sin(\phi)},$$

hence $s \in [0, T]$. A diagram of the scene geometry is shown in Figure 1. Note that $z'(x) = \tan(\phi)$, therefore the no-occlusion constraint is $|\tan(\phi)| < (f/v_m)$.

3.1. Derivation of Plenoptic Spectrum

Starting with an arbitrary Lambertian scene, the equation for the plenoptic spectrum under the FFoV and FSW constraints becomes

$$P(\omega_t, \omega_v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(t, v)\, e^{-j(\omega_t t + \omega_v v)}\, dt dv,$$

$$\overset{(i)}{=} f\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (1 - z'(x)q)\, l(x)\, e^{-j(\omega_t(x - z(x)q) + \omega_v fq)} dx dq,$$

$$\overset{(ii)}{=} f\int_{x_1}^{x_2} l(x) e^{-j\omega_t x}\int_{-\frac{v_m}{f}}^{\frac{v_m}{f}} (1 - z'(x)q) e^{-j(\omega_v f - z(x)\omega_t)q} dq dx,$$

$$\overset{(iii)}{=} \int_{x_1}^{x_2} l(x)\, \mathrm{sinc}(\omega_I)\, e^{-j\omega_t x}\, dx$$

$$- j\int_{x_1}^{x_2} z'(x)\, l(x)\, \frac{d}{d\omega_I}\{\mathrm{sinc}(\omega_I)\}\, e^{-j\omega_t x}\, dx, \tag{4}$$

where $\quad q = v/f, \quad$ and, $\quad \omega_I = \omega_v v_m - \dfrac{z(x)v_m}{f}\omega_t.$

Step (i) follows from applying (1) and assuming a Lambertian scene so that $p(t, v) = l(x)$, and then using the substitution $q = v/f$. Step (ii) follows from applying the FFoV and FSW constraints, and finally step (iii) from solving the integral in q.

Applying the geometry equations in (3) to equation (4), leads to the following integral

$$P_s = e^{-j\omega_t x_1}\left[\int_0^T g(s)\cos(\phi)\mathrm{sinc}(\hat{\omega}_I)\, e^{-j\omega_t \cos(\phi)s}ds\right.$$

$$\left. -j\frac{v_m\sin(\phi)}{f}\int_0^T g(s)\frac{d}{d\hat{\omega}_I}\{\mathrm{sinc}(\hat{\omega}_I)\}\, e^{-j\omega_t \cos(\phi)s}ds\right], \tag{5}$$

where $\quad \hat{\omega}_I = \omega_v v_m - (s\sin(\phi) + z_{min})\dfrac{v_m}{f}\omega_t.$

Using Euler's formula, the texture signal, $g(s)$, can be split into two complex exponentials, $e^{j\omega_s s}$ and $e^{-j\omega_s s}$. Taking the positive complex exponential first, and defining the following variables

$$a = \omega_v v_m - \omega_t\frac{z_{max}v_m}{f}, \qquad b = \omega_v v_m - \omega_t\frac{z_{min}v_m}{f},$$

$$c_1 = \frac{-f\omega_t\cos(\phi) + \omega_s f}{\sin(\phi)\omega_t v_m}, \quad c_2 = \frac{-f\omega_t\cos(\phi) - \omega_s f}{\sin(\phi)\omega_t v_m}, \tag{6}$$

the solution to (5) is

$$|P_s| = \left|\frac{1}{\omega_t}\left[\mathrm{sinc}(a)\, e^{-jT(\omega_t\cos(\phi) - \omega_s)} - \mathrm{sinc}(b)\right]\right.$$

$$+ \frac{\omega_s f}{2v_m\sin(\phi)\omega_t^2}\left[\zeta(jb(c_1 - 1)) - \zeta(ja(c_1 - 1))\right.$$

$$\left.\left. - \zeta(jb(c_1 + 1)) + \zeta(ja(c_1 + 1))\right]e^{jbc_1}\right|, \tag{7}$$

where $\zeta(jw)$, for $w \in \mathbb{R}$, is the following integral

$$\zeta(jw) = \int_0^{jw}\frac{1 - e^{-h}}{h}dh.$$

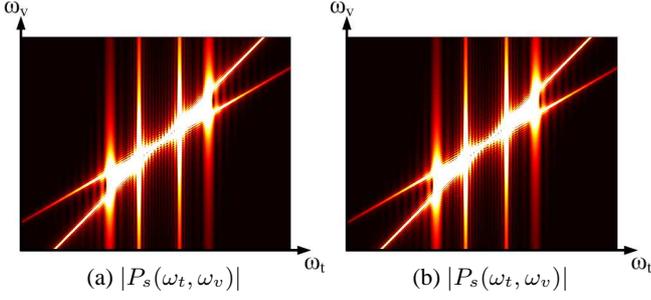(a) $|P_s(\omega_t, \omega_v)|$ (b) $|P_s(\omega_t, \omega_v)|$

Fig. 2. Comparison of the plenoptic spectrum for a slanted plane with sine wave texture, when generated numerically from (5), (a), and when generated using the solution (8), (b).

(a)   (b)

Fig. 3. (a) The essential bandwidth and its parameters. (b) Diagram of the sampling curve for a FPP with bandlimited texture.

The solution to the above integral is

$$
\begin{array}{ll}
\mathrm{E}_1(jw) + \ln(w) + j\frac{\pi}{2} + \gamma & , \text{ if } w > 0 \\
\mathrm{E}_1(-jw) - 2j\mathrm{Si}(-w) + \ln(-w) + j\frac{\pi}{2} + \gamma & , \text{ if } w < 0 \\
0 & , \text{ if } w = 0
\end{array}
$$

where $\mathrm{E}_1(jw)$ is the exponential integral, $\mathrm{Si}(w)$ is the sine integral and γ is Euler's constant; for definitions see [10]. For the negative exponential component of the texture, $g(s)$, ω_s is replaced by $-\omega_s$ and c_1 by c_2.

Therefore the closed form expression for the plenoptic spectrum of a slanted plane with $g(s) = \sin(\omega_s s)$ is

$$
\begin{aligned}
|P_s| = \Bigg| & \frac{j\,\mathrm{sinc}(a)}{\omega_t}\sin(\omega_s T)\,e^{-j\omega_t\cos(\phi)T} \\
& + \frac{\omega_s f}{4 v_m \sin(\phi)\omega_t^2}\{[\zeta(jb(c_1-1)) - \zeta(ja(c_1-1)) \\
& - \zeta(jb(c_1+1)) + \zeta(ja(c_1+1))]\,e^{jbc_1} + [\zeta(jb(c_2-1)) \\
& - \zeta(ja(c_2-1)) - \zeta(jb(c_2+1)) + \zeta(ja(c_2+1))]\,e^{jbc_2}\}\Bigg|.
\end{aligned}
\tag{8}
$$

If $\omega_t = 0$, then

$$
\begin{aligned}
|P_s| = \Bigg| & \sin\left(\frac{\omega_s T}{2}\right)\mathrm{sinc}\left(\frac{\omega_s T}{2}\right)|\cos(\phi)\mathrm{sinc}(\omega_v v_m) \\
& - j\frac{\sin(\phi)v_m}{f}\left(\frac{\cos(\omega_v v_m)}{\omega_v v_m} - \frac{\sin(\omega_v v_m)}{(\omega_v v_m)^2}\right)\Bigg|.
\end{aligned}
$$

Figure 2 compares the plenoptic spectrum calculated numerically using (5) to that calculated using the the close form expression (8). The PSNR between the spectra is 66dB.

3.2. Bandwidth and Filtering

The plenoptic spectrum in Figure 2(b) is band-unlimited in both ω_t and ω_v. Therefore perfect reconstruction cannot be achieved as the finite resolution of the cameras, Δv, leads to lowpass frequency cut-offs in ω_v [4]. However, if the reconstruction filter passband includes most of the energy of the plenoptic function, then the scene can be adequately reconstructed. The bandwidth of the region containing most of the energy is defined as the essential bandwidth [7] of the signal. To clarify, adequate reconstruction leads to a certain reconstruction quality, taking more samples will result in a better quality but at the cost of having more samples. Thus it is a trade-off between reconstruction quality and sample density. For the rest of the paper we will assume that an essential bandwidth of 90% results in an adequate rendering quality.
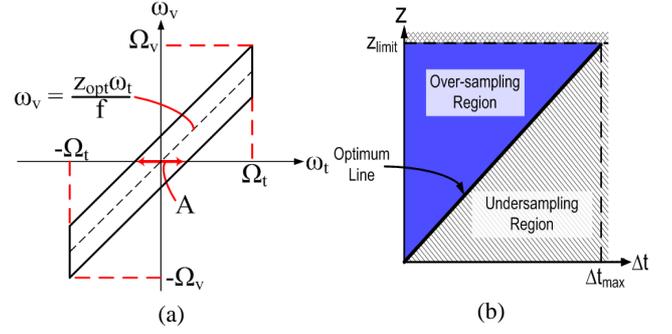
The shape of the essential bandwidth is shown in Figure 3(a). It is parameterised by z_{opt}/f the slant of the region, A, the width in the ω_t direction, Ω_t the maximum value in ω_t and Ω_v the maximum value in ω_v. Assuming the optimum reconstruction filter from [4] leads to

$$
z_{opt} = \frac{2}{(z_{min}^{-1} + z_{max}^{-1})}, \quad \text{and} \quad A \equiv \frac{2\pi}{\Delta t}.
\tag{9}
$$

For the essential bandwidth of a slanted plane, we shall make two assumptions: The first is that Ω_v and Ω_t are related as follows

$$
\Omega_v = \Omega_t \frac{z_{max}}{f}, \quad \text{which leads to} \quad A = \Omega_t\left(\frac{T\sin(\phi)}{z_{min}}\right).
$$

Thus the essential bandwidth depends only on Ω_t. The second assumption is that the spectrum outside the essential bandwidth is characterised by six lines visible in Figure 2; $a=0$, $b=0$, $c_1+1=0$, $c_1-1=0$, $c_2+1=0$ and $c_2-1=0$. If we approximate the decay of these lines and equate their total to 10% of the overall energy, it can be shown that Ω_t is the unique real solution to

$$
\frac{2\alpha}{3X^3} + \left(8v_m^2\sin^2(\omega_s T) + \frac{2(z_{max}+z_{min})\beta}{\sin(\phi)z_{max}f}\right)\frac{1}{X} + \gamma = 0,
\tag{10}
$$

where $\quad X = \Omega_t - (f\omega_s)/(f\cos(\phi) - v_m\sin(\phi))$,

$$
\alpha = \left(\frac{f\omega_s}{\sin(\phi)}\right)^2\left(2\,\mathrm{sinc}(\omega_s T) + \ln\left(\frac{f\cos(\phi)-\sin(\phi)v_m}{f\cos(\phi)+\sin(\phi)v_m}\right)\right)^2,
$$

$$
\beta = T(4\,\mathrm{sinc}(\omega_s T)+1)^2(f^2\cos^2(\phi)+v_m^2\sin^2(\phi)),
$$

$$
\gamma = -\frac{2\pi^2}{5}v_m T\cos(\phi)(1 - \mathrm{sinc}(2\omega_s T)).
$$

Consequently the optimum camera spacing is

$$
\Delta t = \frac{2\pi\, z_{min}}{\Omega_t T\sin(\phi)}.
\tag{11}
$$

4. MAXIMISING THE SPATIAL SAMPLING PERIOD

We shall now consider a special case in which $\phi = 0$ and the scene is a finite fronto-parallel plane (FPP) at a depth z_c. The texture signal is defined as

$$
g(s) \xrightarrow{\mathcal{F}} G(\omega) = 0, \text{ for } \omega \notin [-\omega_s, \omega_s].
\tag{12}
$$

Now applying $z(x) = z_c$ and $x = s$ to (4), results in an integral that can be expressed in terms of the convolution of two Fourier transforms. Therefore the plenoptic spectrum becomes

$$
|P_{FPP}| = \mathrm{sinc}(\omega_I)\left|G(\omega_t) * \mathrm{sinc}\left(\frac{\omega_t T}{2}\right)e^{-j\omega_t\frac{T}{2}}\right|,
\tag{13}
$$

where $*$ is the convolution operator. Again the plenoptic spectrum is band-unlimited in both ω_t and ω_v, requiring the calculation of the essential bandwidth.

The essential bandwidth can be calculated by remembering that the convolution of two signals leads to an addition in bandwidth, and that the essential bandwidth of a sinc function is approximately equal to the width of its main lobe [7]. The width of the main lobe of the sinc functions in (13) are $4\pi/T$ and $2\pi/v_m$. Therefore the essential bandwidth is a slanted elongated oval centred at the origin, the gradient of the slant being equal to z_c/f. Approximating this with the essential bandwidth shown in Figure 3(a) leads to the following parameters

$$A = \frac{2\pi f}{v_m z_c}, \ \ \Omega_t = \omega_s + \frac{2\pi}{T}, \ \text{ and } \Omega_v = \left[\omega_s + \frac{2\pi}{T}\right]\frac{z_c}{f} + \frac{\pi}{v_m}.$$

Thus, using the relationship in (9), the camera spacing is given by

$$\Delta t = \frac{v_m z_c}{f}. \tag{14}$$

The important point here is that by increasing the distance between the plane and the camera line, the camera spacing can be increased. However, due to finite resolution, there is a maximum z_c beyond which the cameras cannot distinguish the texture. This limit is when

$$z_{limit} = \frac{\pi f T}{(\omega_s T + 2\pi)\Delta v}\left[\frac{\pi}{\Delta v} - \frac{\pi}{v_m}\right] \tag{15}$$

Therefore (14) and (15) can be used to create a sampling curve for the scene in terms of z_c against Δt. This sampling curve is illustrated in Figure 3(b); the optimum line in this case is given by (14). For a given value of z_c, the scene is oversampled or undersampled according to whether Δt is chosen to the left or right of this line.

Now consider a scene that consists of two identical FPPs at different depths, but sufficiently far apart so that no occlusion occurs. Under a uniform sampling framework, assuming the geometry is known as a prior allowing separate reconstruction, the sampling period will be dictated by the closest plane. This type of sampling results in one plane being adequately sampled whilst the other is oversampled. However, as the scene is piecewise planar and the starting point of each piece is known, then a non-uniform sampling framework is applicable. Hence the sample rate varies depending on the plane being sampled. This means that both planes are adequately sampled. As a result the overall number of samples can be reduced, or, assuming a fixed number of samples, the reconstruction quality can be increased.

Figure 4 compares two PSNR curves for a reconstructed synthetic EPI as the number of samples increases. The difference between the two is that one uses uniform sampling and the other non-uniform sampling. The EPI consists of two identical FPPs, both with bandlimited texture, at differing depths. The important point is that for a fixed number of samples, non-uniform sampling results in a better quality reconstruction. Note that this simulation assumes the geometry as a prior, thus each plane is reconstructed separately for both uniform and non-uniform sampling.

5. CONCLUSIONS

We have studied the bandwidth of the plenoptic function for two types of Lambertian scenes under finite field of view (FFoV) and finite scene width (FSW) constraints. For a slanted plane with sinusoidal texture signal, we have derive an exact close form expression for the plenoptic spectrum. However the spectrum is band-unlimited
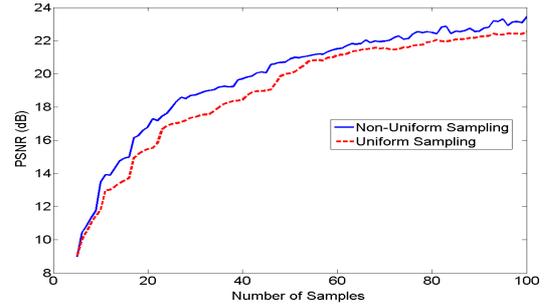

Fig. 4. Comparison of the PSNR of a reconstructed synthetic EPI when sampled uniformly and non-uniformly. The number of samples varies from 5 to 100. The EPI consists of two FPPs at different depths.

due to the constraints. Likewise, applying the constraints to a fronto-parallel plane with bandlimited texture results in a band-unlimited plenoptic spectrum. However using the idea of the essential bandwidth, adequate reconstruction of the plenoptic function is possible. In the case of the fronto-parallel plane, we use the essential bandwidth to derive a sampling curve giving the optimal camera spacing for a certain distance to the scene.

6. REFERENCES

[1] C. Zhang and T. Chen, "A survey on image-based rendering - representation, sampling and compression," *EURASIP Signal Proc.: Image Commun.*, vol. 19, pp. 1–28, 2004.

[2] E.H. Adelson and J.R. Bergen, "The plenoptic function and the elments of early vision," in *Computational Models of Visual Processing*, pp. 3–20. MIT Press, Cambridge, MA, 1991.

[3] H.Y. Shum, S.C. Chan, and S.B. Kang, *Image-Based Rendering*, Springer, 2007.

[4] J.X. Chai, S.C. Chan, H.Y. Shum, and X. Tong, "Plenoptic sampling," in *Proc. SIGGRAPH*, 2000, pp. 307–318.

[5] C. Zhang and T. Chen, "Spectral analysis for sampling image-based rendering data," *IEEE Trans. on Circ. and Syst. for Video Tech.*, vol. 13, no. 11, pp. 1038–1050, 2003.

[6] M.N. Do, D Marchand-Maillet, and M. Vetterli, "On the bandwidth of the plenoptic function," *IEEE Transactions on Image Processing*, 2009, Preprint.

[7] B.P. Lathi, *Modern Digital and Analog Communication Systems*, Oxford University Press, 1998.

[8] M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. SIGGRAPH*, 1996, pp. 31–40.

[9] S. Gortler, R. Grzeszczuk, R. Szeliski, and M. Cohen, "The lumigraph," in *Proc. SIGGRAPH*, 1996, pp. 43–54.

[10] M. Abramowitz and I.A. Stegun, Eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, 1964.