

ADAPTIVE PLENOPTIC SAMPLING

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 present a method of non-uniformly sampling a scene, with a smoothly varying surface, given a finite number of samples. This method approximates such a scene with a set of slanted planes subject to the constraint of finite number of samples. We use the recent spectral analysis of a single slanted plane to determine a piecewise constant spatial sampling rate for the scene. Finally, we show that this sampling rate results in a non-uniform sampling scheme that reconstructs the plenoptic function beyond that of uniform sampling.

Index Terms— Plenoptic sampling, Image-Based Rendering, spectral analysis, adaptive 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 viewpoint and orientation of the camera can be selected by the user. 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.

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. 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].

Chai et al. [3] 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 a uniform sampling rate and optimal reconstruction filter. Using a functional scene model, [4] 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 present a method of non-uniformly sampling a scene with a smoothly varying surface and complex exponential texture. The non-uniform sampling scheme is based on approximating the surface with a set of slanted planes. The slanted plane ap-

proximation is optimised using Lagrange multipliers and results in a non-uniform sample allocation per plane. Consequently, using the spectral analysis of a slanted plane, presented in [5], we determine a piecewise constant spatial sample rate. This leads to a reconstruction of the plenoptic function beyond that of normal uniform sampling for that number of samples.

The outline of this paper is as follows. Section 2 covers the modelling of the scene using the functional framework presented in [4]. In Section 3, we cover the spectral analysis of a slanted plane [5] and define a maximum uniform spatial sampling period. In Section 4, we present a method for non-uniform sampling of a smoothly varying scene surface based upon the previous spectral analysis. Lastly, conclusions 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 [6] or lumigraph [7]. 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) [4].

A further simplification, used in [3, 4], is to fix u and y , corresponding to the situation where the camera positions are constrained to a 1D camera line and only one scan-line is considered in each image, see Figure 1. 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 for a slanted plane using the framework presented in [4]. The scene model consists of three functions that are 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. 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, which is bandlimited to a frequency ω_s . As a result, the scene geometry equations for a finite slanted plane with complex exponential texture are as follows

$$\mathcal{G}_s = \begin{cases} z(x) = s \sin(\phi) + z_{min} \\ x(s) = s \cos(\phi) + x_1 \\ g(s) = e^{j\omega_s s} \end{cases} \quad (1)$$

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

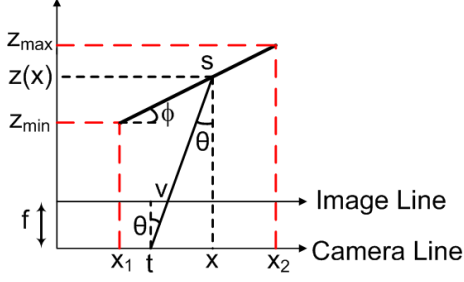


Fig. 1. Scene model of a slanted plane, where s is the curvilinear coordinate, f is the focal length, θ is the viewing angle, ϕ is the angle of slant, z_{min} and z_{max} are the minimum and maximum depth.

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

hence $s \in [0, T]$. A diagram of the scene is shown in Figure 1.

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,

$$t = x - z(x) \tan(\theta) = x - z(x) \frac{v}{f}, \quad (2)$$

where θ is the viewing angle and, assuming finite field-of-view (FFoV), $v \in [-v_m, v_m]$. This relationship is shown in Figure 1. To enforce the no-occlusion condition and make (2) a one to one mapping, [4] places the following constraint on $z(x)$

$$\frac{f}{v_m} > |z'(x)| = |\tan(\phi)|, \quad (3)$$

where $z'(x)$ is the first derivative of z with respect to x . Note that [8] uses a similar framework but permits occlusions.

3. PLENOPTIC SAMPLING OF A SLANTED PLANE

Using the framework defined previously, [5] derived an exact closed-form expression for the plenoptic spectrum of a slanted plane, with bandlimited texture. In this derivation, the plane is constrained to a finite width and the field of view of the cameras is finite, resulting in a spectrum that is band-unlimited in both ω_t and ω_v . The band-unlimited property of the spectrum means that perfect reconstruction of the EPI for a slanted plane, assuming complex exponential texture in this paper, cannot be achieved. However, if we define the essential bandwidth as the region that contains 90% of the signal's energy and assume the plenoptic function is bandlimited to this region, we can reconstruct the EPI up to a certain aliasing error.

In this paper we parameterise the essential bandwidth using a four parameter model, as presented in [5]. The essential bandwidth model is shown in Figure 2 superimposed over the actual plenoptic spectrum. The region covered by the model is defined as

$$\left\{ \omega_t \in [-\Omega_t, \Omega_t], \omega_v \in \left[\frac{z_{opt}}{f} \left(\omega_t - \frac{A}{2} \right), \frac{z_{opt}}{f} \left(\omega_t + \frac{A}{2} \right) \right] \right\},$$

where z_{opt}/f is the slant of the region, A is the width in the ω_t direction, Ω_t is maximum value in ω_t and Ω_v is the maximum value in ω_v . For a slanted plane Ω_t and Ω_v are defined as

$$\Omega_t = \frac{\omega_s f}{f \cos(\phi) - v_m |\sin(\phi)|} + \frac{2\pi}{T}, \quad (4)$$

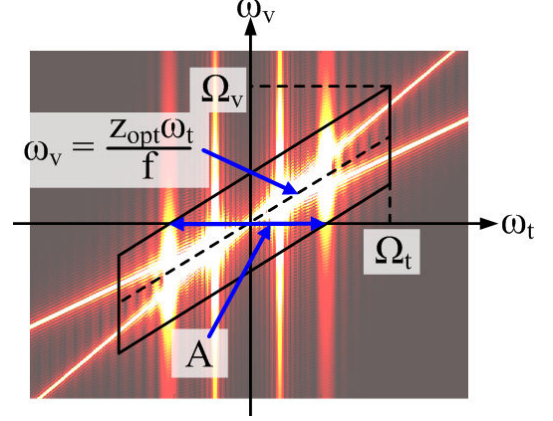


Fig. 2. Parameterisation of the essential bandwidth for a slanted plane superimposed on its plenoptic spectrum. Note that the texture signal is $g(s) = e^{j\omega_s s} + e^{-j\omega_s s}$ in this example.

and

$$\Omega_v = \frac{z_{max} \Omega_t}{f} + \frac{\pi}{v_m}, \quad (5)$$

respectively. Using these fixed points, the other two parameters are defined as

$$z_{opt} = \frac{z_{max} + z_{min}}{2}, \quad (6)$$

and

$$A = \frac{T |\sin(\phi)| \Omega_t}{z_{opt}} + \frac{2\pi f}{z_{opt} v_m}. \quad (7)$$

As a result the maximum spatial sampling period, Δt , for a slanted plane with complex exponential texture is given by

$$\Delta t = \frac{2\pi}{A} = \frac{2\pi z_{opt} v_m}{v_m \Omega_t T |\sin(\phi)| + 2\pi f} \quad (8)$$

and the support of the reconstruction filter, \mathcal{R} , is defined in the (ω_t, ω_v) -domain as follows

$$\mathcal{R} = \left\{ \omega_v \in \left[-\frac{\pi}{\Delta v}, \frac{\pi}{\Delta v} \right], \omega_t \in \left[\frac{\omega_v f}{z_{opt}} - \frac{\pi}{\Delta t}, \frac{\omega_v f}{z_{opt}} + \frac{\pi}{\Delta t} \right] \right\}, \quad (9)$$

where Δv is the finite resolution of the cameras. However, in general, the scene surface is more complex than a slanted plane. The next section will examine sampling a scene with a smoothly varying surface.

4. SAMPLING COMPLEX SCENES

If we approximate the smoothly varying surface, $z(x)$, with a set of slanted planes, we can sample and reconstruct each plane using the theory from the previous section. This approach results in non-uniform sampling of the scene. The approximation of the surface function with L slanted planes is defined as

$$\hat{z}(x) = (x - x_i) \tan(\phi_i) + z(x_i), \quad \text{for } i = 1, \dots, L + 1, \quad (10)$$

where $x \in [x_i, x_{i+1}]$ and the angle of slant for the i th plane is

$$\phi_i = \tan^{-1} \left(\frac{z(x_{i+1}) - z(x_i)}{x_{i+1} - x_i} \right).$$

The plenoptic function of the surface defined in (10) is

$$\hat{p}(t, v) = \begin{cases} \hat{p}_i(t, v), & \text{for } t, v \in \mathcal{S}_i, \quad i = 1, \dots, L + 1 \\ 0, & \text{else} \end{cases} \quad (11)$$

where $\hat{p}_i(t, v)$ is the plenoptic function for the i th plane and \mathcal{S}_i is its region of support in the (t, v) -domain.

However, we are constrained with a finite number of samples, N_T . Therefore, we need to determine the optimum slanted plane approximation of the surface, given N_T samples, and use it to non-uniformly sample the scene. The rest of the section focuses on different aspects of this approach.

4.1. The Sample Allocation Problem

The sample allocation problem can be defined in terms of minimising a distortion function given N_T samples, i.e.

$$\min \left\{ \sum_{i=0}^L D_i(N_i) \right\} \text{ s.t. } N_T = \sum_{i=0}^L N_i, \quad (12)$$

where D_i is the distortion in the plenoptic domain due to the i th plane, and N_i is the number of samples allocated to the i th plane.

4.2. The Distortion Function

The distortion function is a measure of the error in the plenoptic domain. It is the addition of the geometric error, γ , and the aliasing error, α . The geometric error is due to approximating the scene surface, $z(x)$, with a set of slanted planes. The geometric error due to the i th plane is calculated as follows

$$\gamma_i = \underbrace{\int \int_{t, v \in \mathcal{S}_i} |p(t, v) - \hat{p}_i(t, v)|^2 dt dv}_{}, \quad (13)$$

where \mathcal{S}_i is the region of support for $\hat{p}_i(t, v)$ in the (t, v) -domain.

The second part of the distortion is the aliasing error due to undersampling the approximation of the scene surface. In other words, $\alpha_i(N_i)$ is the error caused by undersampling plane i with N_i samples. We assume the worst case scenario for the aliasing error, which is twice the total energy of the plenoptic spectrum that is outside the support of the reconstruction filter in the frequency domain. Hence

$$\alpha_i(N_i) = 2 \underbrace{\int \int_{\omega_t, \omega_v \notin \mathcal{R}_i} |\hat{P}_i(\omega_t, \omega_v)|^2}_{\omega_t, \omega_v \notin \mathcal{R}_i} \omega_t \omega_v, \quad (14)$$

where \mathcal{R}_i is the support of the reconstruction filter in the (ω_t, ω_v) -domain for the i th plane, defined in (9).

However, due to the complexity of the expression for the plenoptic spectrum of a slanted plane, we approximate the aliasing error by assuming the spectrum decays like K/ω outside the support region \mathcal{R}_i . Bearing this in mind, the integrals in (14) simplify to two regions, the first region is defined as

$$\omega_v \notin \left[-\frac{\pi}{\Delta v}, \frac{\pi}{\Delta v} \right],$$

and the second region is

$$\omega_v \in \left[-\frac{\pi}{\Delta v}, \frac{\pi}{\Delta v} \right], \quad \omega_t \notin \left[\frac{\omega_v f}{z_{opt,i}} - \frac{\pi}{\Delta t}, \frac{\omega_v f}{z_{opt,i}} + \frac{\pi}{\Delta t} \right],$$

where $z_{opt,i}$ is defined by the local z_{min} and z_{max} for i th plane. We assume that only the energy in the second region is dependent upon the number of samples for the i th plane, N_i . Consequently the approximation of the aliasing error is

$$\alpha_i(N_i) \approx \frac{16K_i \Delta v}{A_i \pi} + \frac{4A_i K_i \Delta v}{\pi} + \frac{8K_i W_i}{\Delta v (N_i - 1)}, \quad (15)$$

where A_i is defined by (7) for the i th plane, W_i is the width of the region in t that the i th plane is visible to the cameras, and K_i is calculated by setting Δt to satisfy (8) and equating the total energy outside \mathcal{R}_i to 10% of the overall energy. Thus the constant is

$$K_i = \frac{v_m T_i \cos(\phi_i)}{5} \left(\frac{8\Delta v}{A_i \pi} + \frac{2A_i \Delta v}{\pi} + \frac{8\pi}{A_i \Delta v} \right)^{-1}. \quad (16)$$

Consequently, the distortion due to the i th plane is given by $D_i(N_i) = \gamma_i + \alpha_i(N_i)$. Therefore if the approximation of the surface is exact the distortion is only due to undersampling. However if we are in an oversampling situation the distortion is due to the geometric error of the approximation.

4.3. Determining N_i

The sample allocation per plane, N_i , is determined by solving the minimisation in (12) using a Lagrange multiplier, λ . Hence the Lagrangian cost function is

$$\sum_{i=1}^L (D_i(N_i) + \lambda N_i). \quad (17)$$

Solving this minimisation, the sample allocation for the i th plane, assuming $N_i \geq 1 \forall i$, is given by

$$N_i = \sqrt{\frac{8K_i W_i}{\Delta v \lambda}} + 1. \quad (18)$$

The multiplier, λ , is determined by substituting (18) into the finite number of samples constraint in (12) to give

$$N_T = L + \sum_{i=1}^L \sqrt{\frac{8K_i W_i}{\Delta v \lambda}}, \quad (19)$$

and then rearranging to obtain

$$\lambda = \frac{\left[\sum_{i=1}^L \sqrt{8K_i W_i} \right]^2}{\Delta v (N_T - L)^2}. \quad (20)$$

Note that this solution assumes $N_T > L$. Therefore, substituting (20) into (18) gives the sample allocation per plane.

4.4. Merging Planes

We approach the problem of finding the optimum approximation of the surface from the bottom-up. Therefore, starting from an initial, fine-grain, approximation, the optimum solution, given N_T samples, is reached by merging adjacent planes. Adjacent planes are merged if the overall distortion of the approximation, \mathcal{D} , which is given by the Lagrange cost function, is reduced. If the pair of planes in question are merged, then the process is repeated with the next adjacent plane. However, if the planes are not merged, we move to the next plane and repeat the process.

4.5. Non-Uniform Sampling

Given the optimum approximation of the surface with a sample allocation per plane, then (8) gives the maximum spatial sampling period for each plane. However, the sample allocation does not take into account the fact that multiple planes may be visible in a single image. As a result, the sample rate at location t is the maximum of the sample rates for the planes visible in the camera's field of view

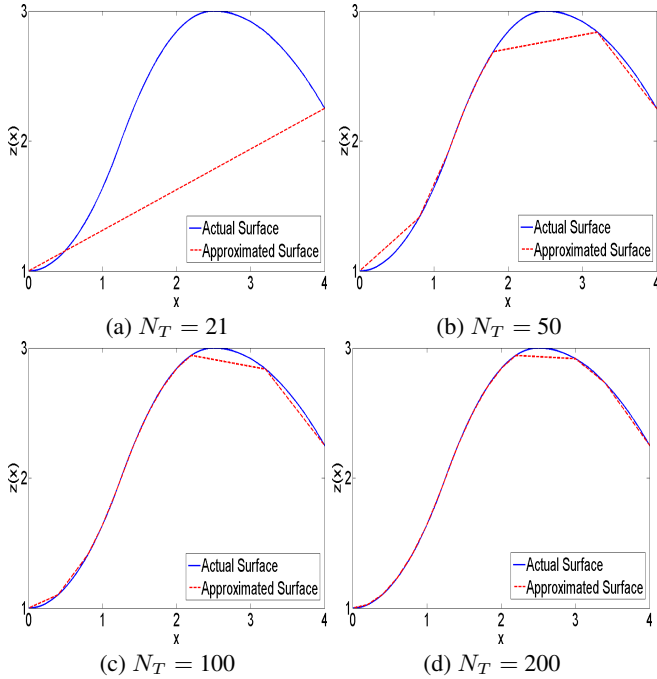


Fig. 3. Examples of approximating a smooth surface, given a fixed number of samples, N_T . The number of planes, L , in each approximation are (a) $L = 1$, (b) $L = 10$, (c) $L = 11$, and (d) $L = 14$.

at that location, e.g. if three planes are visible, with three different sample rates, we take the maximum of these sample rates. In doing this we generate a sample rate profile in t , which can then be used to generate the non-uniform samples.

In terms of reconstruction, the piecewise constant sample rate sections are reconstructed in the Fourier domain using the appropriate filter. A local interpolation is used around the transition region from one rate to another in order to smooth any discontinuities.

4.6. Adaptive Sampling Algorithm

Summarising the above, the algorithm works as follows:

- Step 1.** Initially split the scene surface, $z(x)$, into $L+1$ equal region at the points $x_i, i = 1, \dots, L$. Then generate L slanted planes by joining the points $(x_i, z(x_i))$ to form the approximation.
- Step 2.** Find the solution to the minimisation problem expressed in (12) using (20). This results is an initial overall distortion, \mathcal{D}_{init} , for the surface.
- Step 3.** Starting at plane $i = 1$ calculate the new overall distortion, \mathcal{D}_{new} if it and its neighbour are merged to form a new plane and compare to the previous overall distortion:
 - If Lower.** Merge planes and repeat with the new neighbour.
 - If Higher.** Move to the next plane, $i = i + 1$.
- Step 4.** Using the approximation of the surface, derive the maximum sample rate required for each plane as a function of t . This sample rate then dictates the non-uniform sampling.

The end result of the algorithm is a set of non-uniform samples based upon the slanted plane approximation of the surface. If merging has occurred then the approximation is non-uniform in x .

N ^o of Samples	Non-Uniform PSNR	Uniform PSNR
50	18.92dB	18.00dB
100	22.25dB	21.23dB
150	23.97dB	23.16dB
200	25.23dB	24.52dB

Table 1.

4.7. Simulations

Figure 3 shows an example of approximating a surface, consisting of quadratic pieces, using different values of N_T . For this example the initial, fine-grain, approximation consists of 20 slanted planes, thus the minimum number of samples is 21. When $N_T = 21$ the optimum approximation is a single plane, Figure 3(a). The figure illustrates how the number of planes in the approximation increases with the number of samples.

Using the same quadratic surface, Table 1 shows that the reconstruction of the plenoptic function, using the non-uniform samples, surpasses that of normal uniform sampling for the same number of samples.

5. CONCLUSION

We have presented a method of non-uniformly sampling a scene with a smoothly varying surface function and complex exponential texture. This sampling scheme is based on approximating the surface with a set of slanted planes. We optimise the slanted plane approximation subject to the constraint of N_T samples using Lagrange multipliers. The optimisation allocates the N_T samples between the L planes that form the approximation. Consequently, using the surface approximation and sample allocation, coupled with the spectral analysis for a slanted plane, we derive a piecewise constant sample rate. Finally, this sample rate profile results in a non-uniform sampling scheme that outperforms uniform sampling.

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 elements of early vision," in *Computational Models of Visual Processing*, pp. 3–20. MIT Press, Cambridge, MA, 1991.
- [3] J.X. Chai, S.C. Chan, H.Y. Shum, and X. Tong, "Plenoptic sampling," in *Proc. SIGGRAPH*, 2000, pp. 307–318.
- [4] M.N. Do, D Marchand-Maillet, and M. Vetterli, "On the bandwidth of the plenoptic function," *IEEE Trans. Image Process.*, 2009, Preprint.
- [5] C. Gilliam, P.L. Dragotti, and M. Brookes, "A closed-form expression for the bandwidth of the plenoptic function under finite field of view constraints," in *IEEE International Conference on Image Processing*, 2010, pp. 3965–3968.
- [6] M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. SIGGRAPH*, 1996, pp. 31–40.
- [7] S. Gortler, R. Grzeszczuk, R. Szeliski, and M. Cohen, "The lumigraph," in *Proc. SIGGRAPH*, 1996, pp. 43–54.
- [8] 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.