

# A SPECTRAL ANALYSIS FOR LIGHT FIELD RENDERING

S. C. Chan

Department of Electrical and Electronic Engineering,  
The University of Hong Kong.

H. Y. Shum

Microsoft Research, China.

## ABSTRACT

Image based rendering using the plenoptic function is an efficient technique for re-rendering at different viewpoints. In this paper, we study the sampling and reconstruction problem of plenoptic function as a multidimensional sampling problem. The spectral support of plenoptic function is found to be an important quantity in the efficient sampling and reconstruction of such function. A spectral analysis for the light field, a 4D plenoptic function, is performed. Its spectrum, as a function of the depth function of the scene, is then derived. This result enables us to estimate the spectral support of the light field given some prior estimate of the depth function. Results using a piecewise constant depth model show significant improvement in rendering of the light field images. The design of the reconstruction filter is also discussed.

## 1. INTRODUCTION

Images and videos are effective means to convey information of objects, environment or scenes. To provide the users with better experience such as navigation through a virtual environment and to interact with virtual objects, virtual reality techniques are becoming more and more popular. Recently, image based techniques have attracted much attention as an excellent alternative to re-render images at a collection of viewpoints, due to their superior image quality and much lower computational requirement for rendering, compared with 3-D model building. Central to image based rendering is the *plenoptic function* [1], which describes all of the radiance energy perceived by the observer at any point  $(V_x, V_y, V_z)$  in space and time  $t$ . At any point  $(V_x, V_y, V_z)$  in the free space, we can select any of the viewable rays by choosing an azimuth and elevation angle  $(\theta, \phi)$  as well as a band of wavelengths,  $\lambda$ . For a dynamic scene, there is an additional variable time  $t$ . This gives rise to the following 7 dimensions plenoptic function,

$$p = P(\theta, \phi, \lambda, V_x, V_y, V_z, t). \quad (1)$$

In principle, one can reconstruct the view at any point in space and time from the plenoptic function by substituting the appropriate values of  $(V_x, V_y, V_z)$ ,  $(\theta, \phi)$ , and  $t$  into (1).

In practice, only samples of the plenoptic function are available. The reconstruction problem can be stated as follows [5]: *Given a set of discrete samples (complete or incomplete) from the plenoptic function, the goal of image-based rendering is to generate a continuous representation of that function.* Due to difficulties in *capturing* and *storing* the plenoptic function, various simplifications have been advocated [2-7]. The simplest two-dimensional plenoptic function is the *panorama* (cylindrical [2] or spherical [6]). The *lumigraph* of Gortler et al [3] extends the work of Quicktime VR and plenoptic modeling [5] by further developing the idea of capturing the plenoptic function in a region of the environment. By assuming that air is transparent, and using the ray space representation, the 5-

dimension plenoptic function (ignoring wavelength and time) can be reduced to 4 dimensions. Independently, Levoy and Hanrahan proposed a similar 4D plenoptic function called the light field [4]. In light field and lumigraph, images are taken on certain points in a 2-D plane, such as the rectangular grid in the  $(u, v)$ -plane of Fig. 1. At each grid point, an image of the object or scene is taken. We shall call the image plane the  $(s, t)$ -plane. Using this *2D array of images*, it is possible to create different views of the object or scene at different viewing angles through a process called *rendering*. *Lumigraph* differs from light field in that *approximate geometric model* of the object is used to improve the quality of the reconstruction at *lower sampling densities*. Apart from the reconstruction problem described above, a related but more interesting problem is the *plenoptic sampling problem*: "How many samples of the plenoptic function associated with a scene are required for its reconstruction?" This problem and in fact the reconstruction problem can be answered using the sampling theorem of multidimensional signal. The rest of this paper is organized as follows: Section 2 is devoted to the sampling and reconstruction problem of plenoptic function. A spectral analysis for the light field is given in section 3. The spectral support of the light field is then estimated in section 4. The design of the reconstruction filter is briefly discussed in section 5.

## 2. SAMPLING THE PLENOPTIC FUNCTION

Let the Fourier transform (FT) of the plenoptic function  $p_a(t) = p(\theta, \phi, \lambda, V_x, V_y, V_z, t)$  be  $P_a(j\Omega)$ . Suppose  $p_a(t)$  is sampled at the points  $\{\mathbf{Vn} : \mathbf{n} \in N\}$  to obtain the sequence  $p(\mathbf{n}) = p_a(\mathbf{Vn})$ ,  $\mathbf{V}$  is called the sampling matrix and the set of points  $\{\mathbf{Vn} : \mathbf{n} \in N\}$  is called the lattice generated by the matrix  $\mathbf{V}$ , and is denoted by  $LAT(\mathbf{V})$ . The lattices  $LAT(\mathbf{V}^{-1})$  and  $LAT(2\pi\mathbf{V}^{-1})$  are called respectively the reciprocal and scaled reciprocal lattices of  $LAT(\mathbf{V})$ . It can be shown that the discrete-time Fourier transform of  $p(\mathbf{n})$  is related to that of  $P_a(j\Omega)$  by the following

$$P(\omega) = \frac{1}{|\det \mathbf{V}|} \sum_{\mathbf{k} \in N} P_a(j\mathbf{V}^{-T}(\omega - 2\pi\mathbf{k})), \quad (2)$$

where  $N$  is the set of all 7-component integers. Denote  $2\pi\mathbf{V}^{-1}$  by  $\mathbf{U}$ . (2) tells us that  $P(\omega)$  is obtained by adding all the copies of  $P_a(j\Omega)$  shifted to the lattice points on  $LAT(\mathbf{U})$  and perform the transformation  $\Omega = \mathbf{V}^{-T}\omega$ . If any two copies in this sum overlap, aliasing is said to occur. The sampling density  $\rho$  of  $\mathbf{V}$ , which is the number of lattice points per unit volume, is equal to  $1/|\det \mathbf{V}|$ . Suppose  $p_a(t)$  is bandlimited. If the sampling density  $\rho$  is sufficiently high, the terms on the *RHS* of (2) will not overlap. In other words, we can recover  $p_a(t)$  from its samples using a lowpass filter whose passband should include the region of support of

$P_a(j\Omega)$  and stopband should remove all the other aliasing copies in (2). For a given spectral support of  $P_a(j\Omega)$  and sampling matrix  $V$ , the minimum  $\rho$ ,  $\bar{\rho}_v$ , such that no aliasing occurs is called the minimum sampling density of  $p_a(t)$  with  $V$ . Obviously,  $\bar{\rho}_v$  depends on the spectral support of  $P_a(j\Omega)$  as well as the sampling matrix  $V$ . Different sampling geometry will give rise to different minimum sampling density. Therefore, if we know the spectral support of the plenoptic function  $p_a(t)$  associated with a given scene, the minimum sampling density for a given sampling geometry or lattice can be determined. Moreover, an appropriate sampling lattice can be designed for  $p_a(t)$  to reduce its sampling density. It is a very useful result because the information associated with the plenoptic function is usually huge. Oversampling means more intensive data acquisition and of course more storage. It is clear that the reconstruction problem is equivalent to the reconstruction of  $p_a(t)$  given its aliased spectrum  $P(\omega)$ . If  $p_a(t)$  is bandlimited, the sampling theorem assures us that it is possible to recover  $p_a(t)$  without any distortion after proper filtering. The most important difference between sampling a plenoptic function, or simply plenoptic sampling, and that of an arbitrary 7-dimensional function is the differences in their spectral supports. Due to the close relationship between the plenoptic function and the geometry associated with the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between adjacent pixels of an object in a picture, the correlation between successive image frames in a video, or the correlation between successive images in the light field or lumigraph. All these go to the geometry of the 3D world and its variation with time captured by the plenoptic function. From the above discussion, it can be seen that the spectral support of  $p_a(t)$  is of vital important for efficient sampling and reconstruction of plenoptic function. In the following, a spectral analysis for the lightfield is performed and useful expressions for estimating its spectral support will be derived.

### 3. SPECTRAL ANALYSIS OF LIGHT FIELD

As mentioned earlier, light field is a 4D signal parameterized by the variables  $(u, v, s, t)$ . Denote the light field under consideration by  $l(u, v, s, t)$ . Suppose that we know the depth function of the scene and it is equal to  $z(u, v)$ . Further, if we assume that the BRDF model of the scene is Lambertian, then the radiance received at camera position  $(s, t)$  is given by

$$l(u, v, s, t) = l\left(u - \frac{f \cdot s}{z(u, v)}, v - \frac{f \cdot t}{z(u, v)}, 0, 0\right), \quad (3)$$

where  $f$  is the focal length of the camera. Here, for simplicity, we also assume that there is no occlusion in the scene. With the help of the depth function, which is a kind of geometrical information of the scene, two out of the four variables in  $l(u, v, s, t)$  becomes zero, while the other two become functions of  $(u, s, z(u, v))$  and  $(v, t, z(u, v))$ , respectively. The Fourier transform of  $l(u, v, s, t)$  is then given by

$$L(\Omega_u, \Omega_v, \Omega_s, \Omega_t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} l\left(u - \frac{f \cdot s}{z(u, v)}, v - \frac{f \cdot t}{z(u, v)}, 0, 0\right) \cdot e^{-j\Omega^T x} dx \cdot e^{-j(\Omega_s s + \Omega_t t)} ds \cdot dt, \quad (4)$$

where  $x^T = [u, v]$  and  $\Omega^T = [\Omega_u, \Omega_v]$ . Changing the variables

$$\tilde{u} = u - f \cdot s \cdot h(u, v); \quad \tilde{v} = v - f \cdot t \cdot h(u, v), \quad (5)$$

$$\text{we have } L(\Omega_u, \Omega_v, \Omega_s, \Omega_t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} l(\tilde{u}, \tilde{v}, 0, 0) \cdot |J(\tilde{u}, \tilde{v})| \cdot e^{-j\Omega^T x} e^{-j(\Omega_s s + \Omega_t t)} dx \cdot e^{-j(\Omega_s s + \Omega_t t)} ds \cdot dt \quad (6)$$

where  $J(\tilde{u}, \tilde{v}) = \begin{bmatrix} \frac{\partial u}{\partial \tilde{u}} & \frac{\partial v}{\partial \tilde{u}} \\ \frac{\partial u}{\partial \tilde{v}} & \frac{\partial v}{\partial \tilde{v}} \end{bmatrix}$  is the Jacobian matrix,

$p^T = [s, t]$ ,  $h(u, v) = z^{-1}(u, v)$ , and  $x^T = [\tilde{u}, \tilde{v}]$ . It can be shown that the partial derivatives in  $J(\tilde{u}, \tilde{v})$  are given by

$$\frac{\partial u}{\partial \tilde{v}} = f \cdot s \cdot \frac{\partial h(u, v)}{\partial v} \cdot K(u, v), \quad \frac{\partial u}{\partial \tilde{u}} = \left(1 - f \cdot t \cdot \frac{\partial h(u, v)}{\partial v}\right) \cdot K(u, v),$$

$$\frac{\partial v}{\partial \tilde{u}} = f \cdot t \cdot \frac{\partial h(u, v)}{\partial u} \cdot K(u, v), \quad \frac{\partial v}{\partial \tilde{v}} = \left(1 - f \cdot s \cdot \frac{\partial h(u, v)}{\partial v}\right) \cdot K(u, v), \quad (7)$$

where  $[K(u, v)]^{-1} = 1 - f \left( s \cdot \frac{\partial h(u, v)}{\partial u} + t \cdot \frac{\partial h(u, v)}{\partial v} \right)$ . Using (7), we obtain the determinant of the Jacobian matrix as follows

$$|J(\tilde{u}, \tilde{v})| = \frac{1}{1 - f \left( s \cdot \frac{\partial h(u, v)}{\partial u} + t \cdot \frac{\partial h(u, v)}{\partial v} \right)}. \quad (8)$$

Substituting (8) into (6), we have

$$L(\Omega_u, \Omega_v, \Omega_s, \Omega_t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_h(\tilde{u}, \tilde{v}, \Omega_s, \Omega_t) \cdot l(\tilde{u}, \tilde{v}, 0, 0) \cdot e^{-j\Omega^T x} dx \quad (9)$$

where

$$W_h(\tilde{u}, \tilde{v}, \Omega_s, \Omega_t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{e^{-j(\Omega_s s + \Omega_t t) \cdot h(u, v)} \cdot e^{-j(\Omega_u \tilde{u} + \Omega_v \tilde{v})}}{1 - f \left( s \cdot \frac{\partial h(u, v)}{\partial u} + t \cdot \frac{\partial h(u, v)}{\partial v} \right)} ds \cdot dt.$$

From the convolution theorem, (9) can also be written as

$$L(\Omega_u, \Omega_v, \Omega_s, \Omega_t) = F W_h(\Omega_u, \Omega_v, \Omega_s, \Omega_t) * L_{00}(\Omega_u, \Omega_v). \quad (10)$$

where  $F W_h(\Omega_u, \Omega_v, \Omega_s, \Omega_t)$  is the 2D Fourier transform of  $W_h(\tilde{u}, \tilde{v}, \Omega_s, \Omega_t)$  with respect to the variables  $(\tilde{u}, \tilde{v})$ , and  $L_{00}(\Omega_u, \Omega_v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} l(\tilde{u}, \tilde{v}, 0, 0) \cdot e^{-j\Omega^T x} dx$  is the 2D Fourier transform of the light field image at the origin  $l(\tilde{u}, \tilde{v}, 0, 0)$ . For notational convenience, we have used  $(u, v)$  instead of  $(\tilde{u}, \tilde{v})$  inside the integral of  $W_h(\tilde{u}, \tilde{v}, \Omega_s, \Omega_t)$ . In fact, expressing  $(u, v)$  entirely in  $(\tilde{u}, \tilde{v})$  requires us to solve the nonlinear equations (5). Furthermore,  $h(u, v)$ ,  $\partial h(u, v)/\partial u$  and  $\partial h(u, v)/\partial v$  are now functions of  $(\tilde{u}, \tilde{v})$  as well as  $(s, t)$ . (9) is seen to be the Fourier transform of a windowed 2D signal  $l(\tilde{u}, \tilde{v}, 0, 0)$ . The window  $W_h(\tilde{u}, \tilde{v}, \Omega_s, \Omega_t)$ , however, is a 4-D function, which causes the 2D signal to spread in the frequency domain, requiring higher sampling rate. If  $z(u, v)$  is equal to a constant,  $\partial h(u, v)/\partial u$  and  $\partial h(u, v)/\partial v$  will equal to zero, and  $W_h(\tilde{u}, \tilde{v}, \Omega_s, \Omega_t)$  will become an impulse function. Similarly situation occurs when  $z(u, v)$  is sufficiently large. In this case,

$L(\Omega_u, \Omega_v, \Omega_s, \Omega_t)$  will be a 2D slice in the 4D space of frequencies. This allows us to sub-sample the light field in the  $s$  and  $t$  directions, without aliasing [8]. From (8), it is seen that  $|J(\tilde{u}, \tilde{v})|$  is singular when  $f\left(s \frac{\partial h(u, v)}{\partial u} + t \frac{\partial h(u, v)}{\partial v}\right) = 1$ . Although  $\partial h(u, v) / \partial u$  and  $\partial h(u, v) / \partial v$  are in general very small quantities, if  $s$  and  $t$  are allowed to vary from  $-\infty$  to  $\infty$ ,  $|J(\tilde{u}, \tilde{v})|$  will eventually become singular. This suggests that the analysis is a local one and we have to multiply  $|J(\tilde{u}, \tilde{v})|$  by a window  $w(s, t)$  with finite support to prevent it from being singular. Here, the Gaussian window  $g(s, t) = e^{-s^2/2\sigma_s} e^{-t^2/2\sigma_t} / (2\pi\sqrt{\sigma_s\sigma_t})$  is used because it has the smallest time-frequency product. Using the Taylor series expansions of  $h(u, v)$ ,  $\partial h(u, v) / \partial u$  and  $\partial h(u, v) / \partial v$  at  $(\tilde{u}, \tilde{v})$ ,

$$\begin{aligned} h(u, v) &= h(\tilde{u}, \tilde{v}) \left[ 1 + f \left( s \frac{\partial h(u, v)}{\partial u} \right) \Big|_{(\tilde{u}, \tilde{v})} + t \frac{\partial h(u, v)}{\partial v} \Big|_{(\tilde{u}, \tilde{v})} \right] + \dots \\ \frac{\partial h(u, v)}{\partial u} &= \frac{\partial h(u, v)}{\partial u} \Big|_{(\tilde{u}, \tilde{v})} + \frac{\partial^2 h(u, v)}{\partial u \partial u} \Big|_{(\tilde{u}, \tilde{v})} (f \cdot s \cdot h(\tilde{u}, \tilde{v})) \\ &\quad + \frac{\partial^2 h(u, v)}{\partial u \partial v} \Big|_{(\tilde{u}, \tilde{v})} (f \cdot t \cdot h(\tilde{u}, \tilde{v})) + \dots \\ \frac{\partial h(u, v)}{\partial v} &= \frac{\partial h(u, v)}{\partial v} \Big|_{(\tilde{u}, \tilde{v})} + \frac{\partial^2 h(u, v)}{\partial v \partial u} \Big|_{(\tilde{u}, \tilde{v})} (f \cdot s \cdot h(\tilde{u}, \tilde{v})) \\ &\quad + \frac{\partial^2 h(u, v)}{\partial v \partial v} \Big|_{(\tilde{u}, \tilde{v})} (f \cdot t \cdot h(\tilde{u}, \tilde{v})) + \dots \end{aligned}$$

and the geometric series expansion of  $|J(\tilde{u}, \tilde{v})|$ , we have

$$W_h(\tilde{u}, \tilde{v}, \Omega_s, \Omega_t) \approx \frac{A(\Omega_u, \Omega_v, \Omega_s, \Omega_t)}{2\pi} \exp(-\Omega'_s \sigma_s + \Omega'_t \sigma_t / 2),$$

where

$$\begin{aligned} A(\Omega_u, \Omega_v, \Omega_s, \Omega_t) &= \left[ 1 - jf h'_u(\tilde{u}, \tilde{v}) \sigma_s (\Omega_u + \Omega'_s - \sigma_s \Omega_u (\Omega'_s)^2) \right. \\ &\quad \left. - jf h'_v(\tilde{u}, \tilde{v}) \sigma_t (\Omega_v + \Omega'_t - \sigma_t \Omega_v (\Omega'_t)^2) \right]. \quad (11) \end{aligned}$$

The FT of  $W_h(\tilde{u}, \tilde{v}, \Omega_s, \Omega_t)$  wrt to  $(\tilde{u}, \tilde{v})$  is then given by

$$FW_h(\Omega_u, \Omega_v, \Omega_s, \Omega_t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(\Omega_u, \Omega_v, \Omega_s, \Omega_t) e^{-j\left(\frac{\Omega'_s \sigma_s + \Omega'_t \sigma_t}{2}\right)} e^{-j\Omega_u \tilde{u}} e^{-j\Omega_v \tilde{v}} d\tilde{u} d\tilde{v}. \quad (12)$$

(12) is very difficult to evaluate analytically because  $h'_u = \partial h(\tilde{u}, \tilde{v}) / \partial u$ ,  $h'_v = \partial h(\tilde{u}, \tilde{v}) / \partial v$  and  $h(u, v)$  are functions depending on the scene. Further insight into  $FW_h(\Omega_u, \Omega_v, \Omega_s, \Omega_t)$  can, however, be obtained by digitizing it into an infinite sum. Replacing the variables  $\tilde{u}$  and  $\tilde{v}$  by  $n \cdot \Delta u$  and  $m \cdot \Delta v$ , (13) becomes

$$\begin{aligned} FW_h(\Omega_u, \Omega_v, \Omega_s, \Omega_t) &\approx \frac{1}{2\pi} \sum_n \sum_m A_{n,m}(\Omega_u, \Omega_v, \Omega_s, \Omega_t) \\ &\cdot e^{-j\left(\frac{(\Omega'_s(n,m))^2 \sigma_s + (\Omega'_t(n,m))^2 \sigma_t}{2}\right)} e^{-j\Omega_u \Delta u} e^{-j\Omega_v \Delta v} \Delta u \cdot \Delta v. \quad (13) \end{aligned}$$

where  $A_{n,m}(\Omega_u, \Omega_v, \Omega_s, \Omega_t) = 1 - jf h'_u(n \cdot \Delta u, m \cdot \Delta v)$

$$\begin{aligned} &\cdot \sigma_s (\Omega_u + \Omega'_s(n, m) - \sigma_s \Omega_u (\Omega'_s(n, m))^2) \\ &- jf h'_v(n \cdot \Delta u, m \cdot \Delta v) \sigma_t (\Omega_v + \Omega'_t(n, m) - \sigma_t \Omega_v (\Omega'_t(n, m))^2), \end{aligned}$$

$$\Omega'_s(n, m) = \Omega_s + f \Omega_u h(n \cdot \Delta u, m \cdot \Delta u),$$

and  $\Omega'_t(n, m) = \Omega_t + f \Omega_v h(n \cdot \Delta u, m \cdot \Delta u)$ .

It can be seen that  $FW_h(\Omega_u, \Omega_v, \Omega_s, \Omega_t)$  is the sum of an infinite set of amplitude modulated Gaussian functions. The major axis of the projection of the Gaussian function with index  $(n, m)$  along the  $(\Omega_u, \Omega_s)$  plane is given by the line  $\Omega_s + f \Omega_u h(n \cdot \Delta u, m \cdot \Delta u) = 0$ . Thus each portion of the scene will contribute an amplitude modulated Gaussian functions to  $FW_h(\Omega_u, \Omega_v, \Omega_s, \Omega_t)$ . Its orientation along the  $(\Omega_u, \Omega_s)$  and  $(\Omega_v, \Omega_t)$  planes are given by the lines  $\Omega_s + f \Omega_u h(n \cdot \Delta u, m \cdot \Delta u) = 0$  and  $\Omega_t + f \Omega_v h(n \cdot \Delta u, m \cdot \Delta u) = 0$ , respectively. It can be seen that the orientation depends on the depth value at that point,  $h(n \cdot \Delta u, m \cdot \Delta u)$ , while its amplitude depends on  $h(n \cdot \Delta u, m \cdot \Delta u)$  as well as its local derivatives  $h'_u(n \cdot \Delta u, m \cdot \Delta v)$  and  $h'_v(n \cdot \Delta u, m \cdot \Delta v)$ . Given the bounds of  $h'_u(\tilde{u}, \tilde{v})$  and  $h'_v(\tilde{u}, \tilde{v})$  and that of  $h(u, v)$ , (13) allows us to estimate the spectral support of the light field  $l(u, v, s, t)$ . In the following, we show that significant quality improvement in the quality of rendering of the light field can be obtained even with the use of a simple piecewise constant depth model. Because the model is relatively simple, the spectral supports for the rectangular sampling lattices can be determined explicitly [8].

#### 4. PIECEWISE CONSTANT DEPTH MODEL

If the spread of  $A_{n,m}(\Omega_u, \Omega_v, \Omega_s, \Omega_t)$  in (13) is small, the spectrum of the light field can be approximated by a number of 2D slices with different orientations governed by the depth of the scene. Fig.2 shows the projection of the light field in the  $(\Omega_v, \Omega_t)$  plane. Therefore, the spectral support of the light field is approximately bounded by its maximum and minimum depth [8]. For the commonly used rectangular sampling lattice, the minimum sampling density is obtained when the sampled light field is packed as in Fig. 2b. The minimum sampling rate in the  $t$  direction is  $f_t = K_\Omega f \cdot h_d$ , where  $K_\Omega = \min(B_v, 1/\delta_v)$ ,  $h_d = (z_{\min}^{-1} - z_{\max}^{-1})/2$ .  $B_v$  and  $\delta_v$  are respectively the bandwidth and output resolution of the light field in the  $v$  direction. Similar expression applies to the  $s$  direction. It is also possible to determine the minimum sampling densities for the quincunx and hexagonal sampling lattices. Example spectral support and rendering results are shown in Fig. 3 and 4. From Fig. 2, it can be seen that a reconstruction filter is required for downsampling the light field signal from a previously captured light field with oversampling, or upsampling the light field for rendering at a higher resolution. This is discussed in the following section.

#### 5. DESIGN OF THE RECONSTRUCTION FILTER

Suppose that the spectral support of the light field allows us to reduce its sampling rate in the  $s$  and  $t$  dimensions by a factor of say  $d_s$  and  $d_t$ . Alternately, the downsampled lightfield can be upsampled by a similar factor for rendering. Without loss of generality, we only consider the  $t$  dimension. In practice an integer decimation factor  $D$  equal to or just higher than  $d_t$  is chosen. The corresponding decimation matrix  $M$  can be chosen as  $\begin{bmatrix} D & 0 \\ \alpha & 1 \end{bmatrix}$ , where  $\alpha$  is chosen to cover the desired spectrum. The spectral support of the filter

is  $\{-\pi \leq D\omega_0 + \alpha\omega_1 \leq \pi\} \cap \{-\pi \leq \omega_1 \leq \pi\}$ . Fig. 5 shows the spectral support of  $M$  with  $D=4$ , and  $\alpha=0,1,2$ . This filter can be efficiently designed and implemented using the transformation method in [9]. A 1-D lowpass filter  $P(\omega)$  with cutoff frequency  $\pi/D$  is first designed. Form the multi-dimensional filter  $h_p(\mathbf{n}) = p(n_0)p(n_1)$ , where  $p(n)$  is the impulse response of  $P(\omega)$  and  $\mathbf{n} = [n_0 \ n_1]^T$ . The impulse response of the desired filter  $H_1(\omega)$ ,  $h_1(\mathbf{n})$ , is obtained by decimating  $h_p(\mathbf{n})$  with the matrix  $\hat{M} = \begin{bmatrix} 1 & 0 \\ -\alpha & D \end{bmatrix}$ :  $h_1(\mathbf{n}) = D \cdot h_p(\hat{M}\mathbf{n})$ . Similar filter  $h_s(\mathbf{n}')$  ( $\mathbf{n}' = [n_2 \ n_3]^T$ ) can be designed for the  $s$  direction. The resulting filter is the product of  $h_1(\mathbf{n})$  and  $h_s(\mathbf{n}')$ . Further properties of this method can be found in [9].

### CONCLUSION

A spectral analysis for the light field rendering is presented. It enables us to estimate the spectral support of the light field given some prior estimate of the depth function. Results using a piecewise constant depth model show significant improvement in rendering of the light field images. The design of the reconstruction filter is also discussed.

### REFERENCES

- [1] E. H. Adelson and J. Bergen, "The plenoptic function and the elements of early vision." In *Computational Models of Visual Processing*, pages 3-20. MIT Press, Cambridge, MA, 1991.
- [2] S. E. Chen, "QuickTime VR - an image-based approach to virtual environment navigation," *Computer Graphics (SIGGRAPH'95)*, pages 29-38, August 1995.
- [3] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The lumigraph," *Computer Graphics (SIGGRAPH'96)*, pp. 43-54, August 1996.
- [4] M. Levoy and P. Hanrahan, "Light field rendering," in *Computer Graphics Proceedings. Annual Conference Series*, pages 31-42, Proc. SIGGRAPH'96 (New Orleans), August 1996.
- [5] L. McMillan and G. Bishop, "Plenoptic modeling: An image-based rendering system," *Computer Graphics (SIGGRAPH'95)*, pages 39-46, August 1995.
- [6] R. Szeliski and H. Y. Shum, "Creating full view panoramic image mosaics and texture-mapped models," *Computer Graphics (SIGGRAPH'97)*, pages 2521-2528, August 1997.
- [7] H. Y. Shum and L. W. He, "Rendering with Concentric Mosaic" *Computer Graphics (SIGGRAPH'99)*, pp. 299 - 306, August 1999.
- [8] J. X. Chai, X. Tong, S. C. Chan and H. Y. Shum, "Plenoptic sampling," to appear in *SIGGRAPH'2000*.
- [9] P. P. Vaidyanathan, *Multirate systems and filter banks*. Englewood Cliffs, NJ: Prentice Hall, 1992.

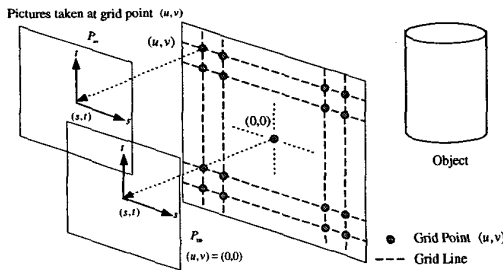


Fig. 1: Light Field construction.

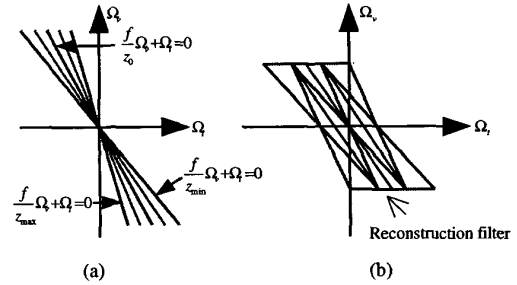


Fig. 2. Projection of light field onto the  $(\Omega_x, \Omega_y)$  plane: (a) spectral predicted by the constant depth model, (b) sampled spectrum and reconstruction using the reconstruction filter ( $z_0$  is the mean depth).

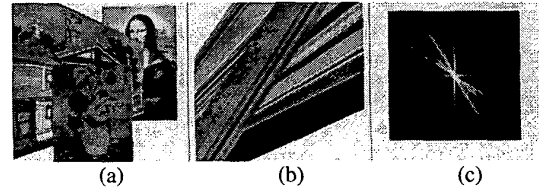


Fig. 3. Example spectral support of light field (a) raw image, (b) EPI, (c) spectral support in the  $(\Omega_x, \Omega_y)$  plane.

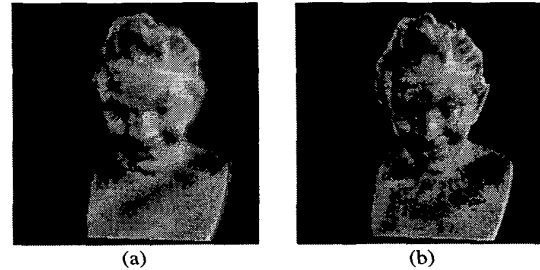


Fig 4: Reconstructed images : (a) at the minimum depth,  $z = 2326$ ; (b) at the optimal plane,  $z = 3323$ ; The focal plane is approximately on the forehead.

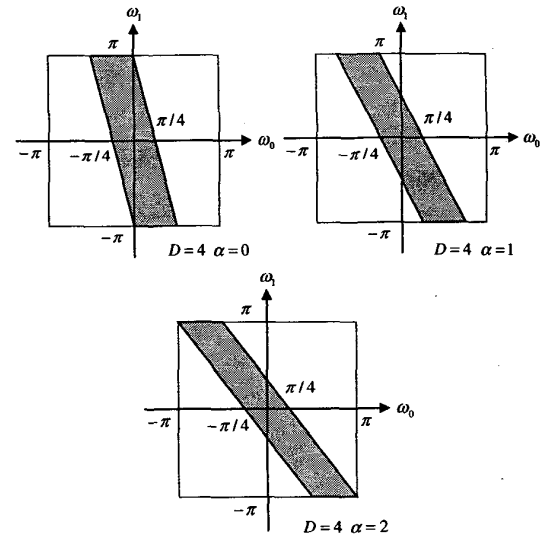


Fig. 5. Spectral support of reconstruction filters with  $D=4$  and different values of  $\alpha$ .