<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>A spectral analysis for light field rendering</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Chan, SC; Shum, HY</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>IEEE International Conference on Image Processing, 2000, v. 2, p. 25-28</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2000</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/46176">http://hdl.handle.net/10722/46176</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.; ©2000 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.</td>
</tr>
</tbody>
</table>
A SPECTRAL ANALYSIS FOR LIGHT FIELD RENDERING

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

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,

\[
p = P(\theta, \phi, \lambda, V_x, V_y, V_z, t).
\]

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

\[
P(\omega) = \frac{1}{|detV|} \sum_{n \in N} P_\mu(j\omega^{-1} n - 2\pi \mathbf{a}),
\]

where \(N\) is the set of all 7-component integers. Denote \(2\pi V^{-1}\) by \(U\). (2) tells us that \(P(\omega)\) is obtained by adding all the copies of \(P_\mu(\Omega)\) shifted to the lattice points on \(LAT(U)\) and performing the transformation \(\Omega = V^{-1} \omega\). If any two copies in this sum overlap, aliasing is said to occur. The sampling density \(\rho\) of \(V\), which is the number of lattice points per unit volume, is equal to \(1/|detV|\). Suppose \(p_\mu((\theta,\phi,\lambda, V_x, V_y, V_z, t))\) 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 \(p_\mu((\theta,\phi,\lambda, V_x, V_y, V_z, t))\) from its samples using a lowpass filter whose passband should include the region of support of...
\[ P_s(\mathbf{f}) \] and stopband should remove all the other aliasing copies in (2). For a given spectral support of \( P_s(\mathbf{f}) \) and sampling matrix \( V \), the minimum \( \rho \cdot \tilde{\rho} \), such that no aliasing occurs is called the minimum sampling density of \( P_s(\mathbf{f}) \) with \( V \). Obviously, \( \tilde{\rho} \) depends on the spectral support of \( P_s(\mathbf{f}) \) 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 \( p_s(t) \) 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 \( p_s(t) \) 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 \( p_s(t) \) given its aliased spectrum \( P(\omega) \). If \( p_s(t) \) is bandlimited, the sampling theorem assures us that it is possible to recover \( p_s(t) \) 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 successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image frames in a video, or the correlation between adjacent pixels of an object in a picture, differences in their spectral supports. Due to the close relationship between the plenoptic function and the scene, there is significant amount of redundancy in the plenoptic function. This can be the correlation between successive image
It can be seen that \( \mathcal{F}w(\Omega_0,\Omega_1,\Omega_2,\Omega_3) \) 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 \((\Omega_0,\Omega_1)\) plane is given by the line \( \Omega_0 + \beta \Omega_1 h(n \cdot \Delta u,m \cdot \Delta v) = 0 \). Thus each portion of the scene will contribute an amplitude modulated Gaussian function to \( \mathcal{F}w(\Omega_0,\Omega_1,\Omega_2,\Omega_3) \). Its orientation along the \((\Omega_0,\Omega_1)\) and \((\Omega_2,\Omega_3)\) planes are given by the lines \( \Omega_0 + \beta \Omega_1 h(n \cdot \Delta u,m \cdot \Delta v) = 0 \) and \( \Omega_2 + \beta \Omega_3 h(n \cdot \Delta u,m \cdot \Delta v) = 0 \), respectively. It can be seen that the orientation depends on the depth value at that point, \( h(n \cdot \Delta u,m \cdot \Delta v) \), while its amplitude depends on \( h(n \cdot \Delta u,m \cdot \Delta v) \) as well as its local derivatives \( h'(n \cdot \Delta u,m \cdot \Delta v) \) and \( h''(n \cdot \Delta u,m \cdot \Delta v) \). Given the bounds of \( h'(n \cdot \Delta u,m \cdot \Delta v) \) and \( h''(n \cdot \Delta u,m \cdot \Delta v) \) 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 \( A_\alpha(\Omega_0,\Omega_1,\Omega_2,\Omega_3) \) 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 \((\Omega_0,\Omega_1)\) 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 \( s \) direction is \( f_s = K_{\alpha} f - h_s \), where \( K_{\alpha} \) is \( \min(\beta,1/d) \), \( h_s = \frac{z_m - z_{\alpha}}{2} \). \( B_s \) and \( d_s \) 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 \( s \) dimension. In practice an integer decimation factor can be achieved by 

\[
\left[ \begin{array}{c} D \\ 0 \end{array} \right],
\]

where \( \sigma \) is chosen to cover the desired spectrum. The spectral support of the filter matrix \( M \) can be chosen as...
is \(-\pi \leq D\omega_0 + \omega_0 \leq \pi\) \(\cap \{-\pi \leq \omega_1 \leq \pi\}\). 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 \(P(o)\) with cutoff frequency \(\pi/D\) is first designed. Form the multi-dimensional filter \(h_p(n) = p(n_0)p(n_1)\), where \(p(n)\) is the impulse response of \(P(o)\) and \(n=[n_0, n_1]^T\). The impulse response of the desired filter \(H,(o)\), \(h_p(n)\), is obtained by decimating \(h_p(n)\) with the matrix

\[
M = \begin{bmatrix} 1 & 0 \\ -\alpha & D \end{bmatrix}, \quad h_p(n) = D \cdot h_p(Mn)\]

Similar filter \(h,(o')\)

\((n'=[n_1, n_2]^T)\) can be designed for the \(s\) direction. The resulting filter is the product of \(h,(n)\) and \(h,(o')\). 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