Attenuation-based Light Field Displays

Bachelorarbeit

der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von

Adrian Wälchli

2015

Leiter der Arbeit:
Prof. Dr. Matthias Zwicker
Institut für Informatik und angewandte Mathematik

Abstract

Abstract goes here

Acknowledgements

Special thanks goes to Prof. Dr. Matthias Zwicker for his supervision and support of this thesis and for giving me the opportunity to work together with the Computer Graphics Group at the University of Berne. A big thank you to secondary advisor Siavash Bigdeli for his time and effort to help me throughout the project. Lastly, I would like to thank my family, friends and colleagues for the great time I had with them during the time of my bachelor studies.

Contents

1 Introduction **1**
1.1 Related Work 1

2 Capturing a Light Field **2**
2.1 The Plenoptic Function and the Light Field 2
2.2 Light Field Acquisition 3
2.3 Visualization 5
2.4 The Plenoptic Camera 6

3 Light Field Tomography **7**
3.1 A Model for Light Attenuation 7
3.2 Discrete Attenuation Layers 8
3.3 Ray Casting 9
3.4 Iterative Reconstruction 11

4 Spectral Analysis **12**
4.1 Definitions 12
4.2 Spectral Support of Light Fields 13
4.3 Spectral Support of Multiplicative Displays 13
4.4 The Fourier Slice Theorem 13

A ap1 **15**
A.1 apsec1 15

List of Tables **16**

List of Figures **16**

Bibliography **17**

Chapter 1

Introduction

1.1 Related Work

Chapter 2

Capturing a Light Field

2.1 The Plenoptic Function and the Light Field

The plenoptic function, as introduced by Adelson and Bergen [1991], is a 7D function that describes the intensity of light for every frequency, along every light ray in space, at any time. It is defined as

$$\begin{aligned} P\colon \mathbb{R}^3 \times [0, 2\pi) \times [0, \pi] \times \mathbb{R}^2 &\to \mathbb{R}^+ \\ (x, y, z, \theta, \phi, t, \lambda) &\mapsto P\left(x, y, z, \theta, \phi, t, \lambda\right), \end{aligned}$$

where the parameters (x, y, z) are the coordinates of a point in 3D space and the angles (θ, ϕ) describe the direction of an incoming light ray at time t. The light's intensity is given for every wavelength λ and thus, the plenoptic function not only captures the visible frequency spectrum but all electromagnetic waves. A commonly used measure for light is the radiance, which is obtained from P by integrating over all wavelengths: $R\left(x, y, z, \theta, \phi, t\right) = \int_{\mathbb{R}} P\left(x, y, z, \theta, \phi, t, \lambda\right) \mathrm{d}\lambda$.

In practice, it is impossible to acquire all the data needed to model the 7D plenoptic function and hence it is reasonable to consider only a subset of the parameters. Dropping the time parameter t in $R\left(x, y, z, \theta, \phi, t\right)$ yields a 5D function for the radiance in a static scene. As described by Levoy and Hanrahan [1996], this five dimensional representation can further be reduced to four dimensions in the following way. The radiance along a line is constant in free space and so, the 5D plenoptic function holds redundant information for the points on this line. Ignoring this redundancy leads to the equivalent 4D parameterization of the ray space. Levoy and Hanrahan [1996] propose a parameterization by two parallel planes, as seen in figure 2.1, where the coordinates of the lines (rays) are given by the intersections with the two planes. The **4D light field** $L(u, v, s, t)$ is therefore defined as the radiance along the line intersecting the two planes at coordinates (u, v) and (s, t). This two-plane parameterization of the light field is the most common one seen in literature, but there are many ways to choose a parameterization. For instance, one can use a plane and two angles to define each ray passing this plane, which would result in a light field $L(u, v, \theta, \phi)$ where $\theta, \phi \in (0, \pi)$.

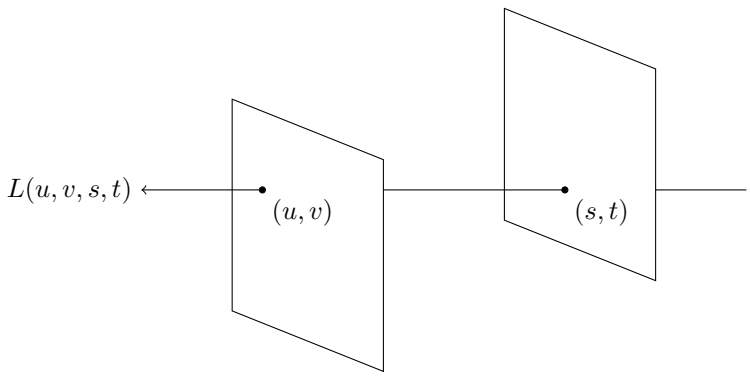


Figure 2.1: Parametrization of the light field with two planes.

2.2 Light Field Acquisition

For practical applications, the light field must be discretized and so, an appropriate sampling method needs to be chosen. This means that only a slice of the actual light field can be captured and the two planes are clipped to form rectangles. In this work, the term *light field* is used for both the infinite, continuous light field as well as the discrete collection of data samples.

Oblique Projection

Oblique projection, as shown in figure 2.2(a), is a special case of orthographic projection: The parallel rays do not need to be perpendicular to the image plane of the camera. The advantage is that there is a one-to-one correspondence between camera position and ray angle, since all rays in one camera are parallel. This means that the angular resolution is simply the number of cameras, and the spatial resolution is the number of pixels in the image plane. The angular extent from $\theta_{\min}$ to $\theta_{\max}$ is called the **field of view** (FOV) of the light field and should not be confused with the field of view of a conventional camera. For a uniform angular sampling with resolution $N_\theta \times N_\phi$, the angles θ_i and ϕ_j are

$$\theta_i = \theta_{\min} + (i-1)\frac{\mathrm{FOV}_\theta}{N_\theta - 1}, \qquad \phi_j = \phi_{\min} + (j-1)\frac{\mathrm{FOV}_\phi}{N_\phi - 1}, \tag{2.1}$$

where $i = 1, 2, \ldots, N_\theta$ and $j = 1, 2, \ldots, N_\phi$.

Given a light field $L(u, v, s, t)$ and the distance d between the two planes, a re-parameterization $L'(\theta, \phi, s, t)$ can be obtained according to figure 2.2(b) by the transformation

$$\theta = \arctan\left(\frac{u-s}{d}\right), \qquad \phi = \arctan\left(\frac{v-t}{d}\right). \tag{2.2}$$

Note that uniform sampling in angular dimension does not yield a uniform grid in the (u, v)-plane. Despite the simplicity of this projection type, it is not feasible to build cameras of this type and so, oblique projection is left to be used exclusively by computers for rendering synthetic scenes.

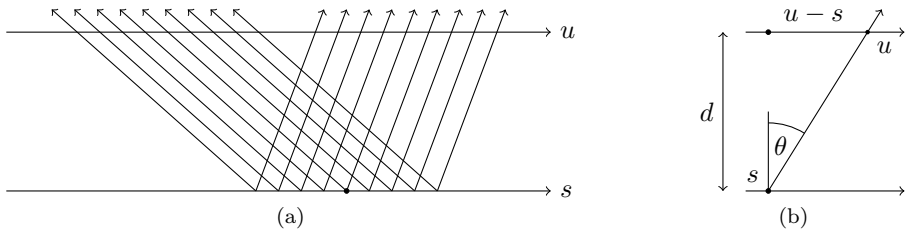


Figure 2.2: (a) Light field acquisition using oblique projection. (b) Re-parameterization of the two-plane representation to angular coordinates.

Perspective Projection

Another way to capture the light field is with a grid of optical systems, e.g. cameras. Typically, the (u, v)-plane is sampled on a grid $G_{uv} = \{(u_i, v_j) \mid i = 1, \ldots, n, j = 1, \ldots, m\}$ on the (u, v)-plane with a resolution $n \times m$. The extent in horizontal (vertical) direction is called the horizontal (vertical) **baseline**. Although it is strictly speaking not correct, the resolution of the (u, v)-plane is often referred to as the angular resolution. The angles of the rays in a light field captured by perspective projections are determined by the focal length, the sensor size and the sensor resolution of the camera. For a camera light field, typically it is expected that

- All cameras are placed at grid positions in G_{uv} on the same plane, called the (u, v)-plane,
- The optical axes of the cameras are orthogonal to the (u, v)-plane,
- All cameras have the same intrinsic parameters (e.g. focal length).

In this case, the focal planes of all cameras coincide with a common focal plane, the (s, t)-plane. Figure 2.3(a) shows this scenario for three cameras in two dimensions. Given images $I_{uv}(x, y)$ with respect to a coordinate system centered at the camera position (u, v), the coordinates on the (s, t)-plane are $s = u + x$, and $t = v + y$. Thus, the light field in continuous coordinates is obtained by

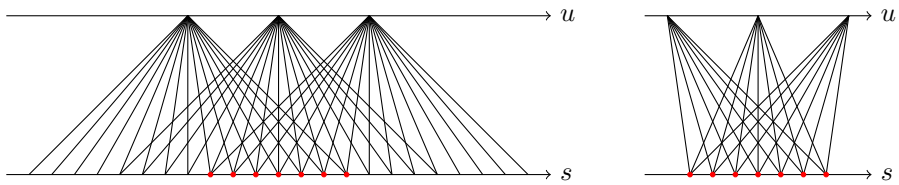
$$L(u, v, s, t) = L(u, v, u + x, v + y) = I_{uv}(x, y). \tag{2.3}$$

In the discrete case, each camera captures sample points on the (s, t)-plane, but not everyone of these sample points on the (s, t)-plane is captured by every camera. So, as demonstrated in figure 2.3(b), the camera images need to be rectified such that all discrete coordinates (u, v, s, t) correspond to valid rays. This rectification process is equivalent to a re-parameterization L' of the continuous light field L, given by the formula

$$L'(u, v, s', t') = L\left(u, v, \gamma(s' - u) + u, \gamma(t' - v) + v\right), \tag{2.4}$$

where $\gamma = \frac{d}{d'}$ and d' is the distance between the (u, v)-plane and the new (s', t')-plane. As derived by Isaksen et al. [2000], this re-parameterization is equivalent to a 4D shear.

A different way to understand this coordinate change is to imagine the (u, v)- and (s, t)-plane being the aperture and sensor planes respectively, resulting in



(a) (b)

Figure 2.3: Perspective projections of a scene. (a) Projections with three pinhole cameras. (b) Discarding unused rays corresponds to cropping the camera images.

one big camera in which a light field is formed. Changing the distance between the two planes is now equivalent to changing the focal length of this one camera. The effect on the light field inside is similar to refocusing, except that in a conventional camera, the image on the sensor is formed by a weighted integral over u and v such that the angular information vanishes. Objects at focal distance from the camera would appear sharp and objects away from the focal point would become blurred.

From stereo vision, it is known that the displacement of the projections in the image planes of two cameras is only dependent on the focal length f, the baseline Δu and the distance z, and the relation is given by $\Delta x = f\Delta u/z$. This knowledge can directly be applied to the two-plane parameterization. For the continuous light field, it amounts to

$$\mathrm{d}s = \frac{z - Z_{st}}{z - Z_{uv}}\,\mathrm{d}u \qquad \text{and} \qquad \mathrm{d}t = \frac{z - Z_{st}}{z - Z_{uv}}\,\mathrm{d}v, \tag{2.5}$$

with Z_{uv} and Z_{st} denoting the placement of the (u, v)- and (s, t)-planes in Z-direction. Usually, the coordinate system is chosen such that $Z_{uv} = 0$. In the discrete case, the displacement Δs or Δt is also called the **disparity** and is often measured in pixel units.

2.3 Visualization

The epipolar-plane image (EPI) allows for a very intuitive visualization of depth from a 4D light field. It was first defined by Bolles et al. [1987] as follows. Consider a point P in 3D space and a pair of cameras with the optical axis pointing in the same direcion. The plane passing through P and the two centers of projection is called the **epipolar plane**. The epipolar plane projects to a line on each of the camera image planes, named the **epipolar line**. This line represents a constraint for the projection of P in each of the images and it is used to solve the correspondance problem in computer vision. The notion of epipolar lines can be directly applied to a multiple camera setup. In figure 2.4, a synthetic scene is rendered in 500 different positions along a horizontal baseline. Since the camera movement is in horizontal direction only, the epipolar lines correspond to a fixed pixel row in each image. The EPIs shown in figures 2.4(b) and 2.4(c) are created by collecting the chosen pixel row (scanline) in every image and stacking it up.

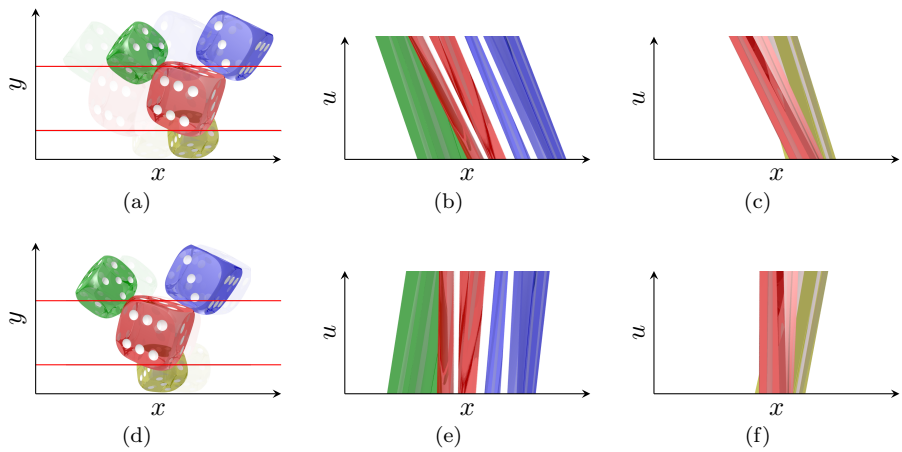


Figure 2.4: (a) Raw 3D light field rendered from 500 positions along a horizontal baseline. Two scanlines are extracted from every image. (b) The feature paths of the blue and green dice have a steeper slope than those of the red die. (c) Feature paths of the yellow die have an even steeper slope, indicating greater depth. (d) The light field is rectified according to figure 2.3(b) such that the disparities of the red die are approximately zero. (e) - (f) EPIs from the same scanlines. The slopes of the feature paths stay the same relative to each other.

As described in the previous section, the depth component of P occurs as a displacement of the projections in consecutive images. Under the assumption that the (u, v)-plane is sampled uniformly, the disparity D with respect to P stays constant from one image to the next. Thus, following the projection of the point P in every image corresponds to a line in the EPI with a slope proportional to $1/D$. Bolles et al. [1987] refer to this line as the **feature path**. This means that points farther away from the camera will appear as a feature path in the EPI with steeper slope than points close to the camera. Note that the depth range in the light field can immediately be determined by identifying the maximum and minimum slope in the EPI. Also, for a perfectly Lambertian scene, each line in the EPI has a uniform color.

2.4 The Plenoptic Camera

Chapter 3

Light Field Tomography

3.1 A Model for Light Attenuation

The light field display is modeled by a volumetric attenuator $\mu(x, y, z)$ that attenuates the light traveling through its material. According to the Beer-Lambert law, the intensity of a light ray $\mathcal{R} \subset \mathbb{R}^3$ passing through the material decreases exponentially over distance:

$$I = I_0 e^{-\int_{\mathcal{R}} \mu(r)\,\mathrm{d}r}. \tag{3.1}$$

The incident intensity I_0 is the intensity of the ray before it enters the attenuator. Equation 3.1 can be rewritten into

$$\bar{I} := \log\left(\frac{I}{I_0}\right) = -\int_{\mathcal{R}} \mu(r)\,\mathrm{d}r. \tag{3.2}$$

Now, let the attenuator $\mu(x, y, z)$ be a cubic slab of height d in Z-direction and let $L(u, v, s, t)$ be the two-plane parameterization of the light field such that the (s, t)-plane coincides with the (x, y)-plane of the attenuator and the (u, v)-plane is at distance d. The set of points describing the ray defined by the coordinates (u, v, s, t) is

$$\mathcal{R} = \left\{ \lambda a + b \;\middle|\; a = \begin{pmatrix} u - s \\ v - t \\ d \end{pmatrix}, b = \begin{pmatrix} s \\ t \\ 0 \end{pmatrix}, \lambda \in \mathbb{R} \right\}. \tag{3.3}$$

A point $p = (x, y, z)^T$ is part of the ray $\mathcal{R}$ if and only if

$$\exists \lambda \in \mathbb{R} : p = \lambda a + b \qquad \Longleftrightarrow \qquad a \times (p - b) = 0, \tag{3.4}$$

where $\times$ denotes the cross product. Now, I can be replaced with the light field L and the right hand side of equation 3.2 can be written as an integral over $\mathbb{R}^3$:

$$\bar{L}(u, v, s, t) = -\int_{\mathbb{R}^3} \mu(p)\delta(a \times (p - b))\,\mathrm{d}p. \tag{3.5}$$

Here, δ denotes the Dirac delta function on $\mathbb{R}^3$ and μ is zero outside the boundaries of the slab. This means that the integrand is only non-zero for points on the ray with coordinates (u, v, s, t).

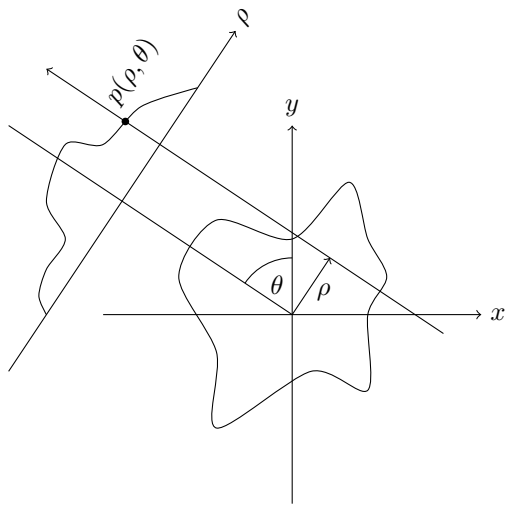


Figure 3.1: The 2D Radon transform of the ray (ρ, θ) passing a material with density $f(x, y)$.

Combining equation 3.1 and 3.5 gives the light field emitted by the attenuator. The goal is to produce such an attenuation display that emits a given target light field.

In computed tomography, the **Radon transform** of a real valued and compactly supported, continuous function $f(x, y)$ on $\mathbb{R}^2$ is defined as

$$p(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\delta(x \cos\theta - y \sin\theta - \rho) \, \mathrm{d}x \, \mathrm{d}y, \tag{3.6}$$

where $(\rho, \theta) \in \mathbb{R} \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ defines a ray as shown in figure 3.1. Because the Radon transform is essentially a line integral, it can be generalized to three or more dimensions. Adapting the notation from the two-plane parameterization, the Radon transform of the attenuation map μ along ray $\mathcal{R}$ becomes

$$p(u, v, s, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu(x, y, z)\delta\left(a \times \left((x, y, z)^T - b\right)\right) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z, \tag{3.7}$$

which is equivalent to equation 3.5. This shows that

$$\bar{L}(u, v, s, t) = -p(u, v, s, t), \tag{3.8}$$

or with the words of Wetzstein et al. [2011]: "The logarithm of the emitted light field is equivalent to the negative Radon transform of the attenuation map."

3.2 Discrete Attenuation Layers

The previous section introduced a continuously varying attenuation map to model the display. Wetzstein et al. [2011] propose to represent the attenuator with a set of N two-dimensional layers, also called masks.

Let $L_{ijkl} = L(u(i), v(j), s(k), t(l))$ be the matrix of samples from the light field and for simplicity, let $m \coloneqq m(i, j, k, l)$ be a linear index of the 4D indices. Equation 3.1 suggests a per-ray constraint in the form

$$L_m = L_0 \prod_{n=1}^{N} t^{(n)}(h(m, n)), \tag{3.9}$$

where $h(m, n)$ is the (discrete) 2D coordinate of the intersection of the m-th ray with the n-th layer, and $t^{(n)}(\xi)$ is the **transmittance** of layer n at that coordinate. Having a constraint for each ray, the goal is to solve for the transmittance t. However, the system of equations in 3.9 is non-linear and cannot directly be solved. One can obtain a linear system of equations by taking the logarithm in 3.9:

$$\bar{L}_m = \sum_{n=1}^{N} \log\left(t^{(n)}(h(m, n))\right) = -\sum_{n=1}^{N} a^{(n)}(h(m, n)) = -P_m \alpha. \tag{3.10}$$

Here, $a^{(n)} \coloneqq -\log t^{(n)}$ denotes the **absorbance** of layer n. This relation between transmittance and absorbance also directly follows from the Beer-Lambert law. Here, $P_m = \left(P_m^{(1)}, \ldots, P_m^{(N)}\right)$ is a binary row vector, encoding the intersection of the ray with the pixels on each layer. The unknown absorbance is represented by the column vector $\alpha = \left(\alpha^{(1)}, \ldots, \alpha^{(N)}\right)^T$. Each $\alpha^{(i)}$ is just a flattened representation of the absorbance matrix $a^{(i)}$. Note that equation 3.10 is the equivalent of the continuous version in 3.8, since P_m encodes the Radon transform. Finally, the above equations indexed by m can be combined into one large linear system $P\alpha = -\bar{L}$.

In most cases, P is not a square matrix and the system can become overdetermined, which means that it has no solution in general. However, it is still possible to find values for α such that the error $\left\|P\alpha + \bar{L}\right\|$ is small. Thus, the objective becomes

$$\begin{aligned} \underset{\alpha}{\text{argmin}} \quad & \left\|P\alpha + \bar{L}\right\|^2 \\ \text{subject to} \quad & 0 \leq \alpha \leq \infty. \end{aligned} \tag{3.11}$$

Finally, when optimal values α are found, the transmittance used to fabricate the layers is obtained by calculating $e^{-\alpha}$. Also note that the matrix P is very sparse because it is assumed that a ray passes through each layer at exactly one pixel no more than once and inter-reflections between the layers are not supported by the model. Thus, P can be efficiently stored using an appropriate data structure.

3.3 Ray Casting

To obtain the linear system P, the intersections between the rays and the attenuation layers have to be calculated. This calculation depends on the parameterization of the light field. For continuous light fields, it is always possible to apply a re-parameterization to get the desired representation (e.g. the two-plane parameterization) and then compute the intersection in a standard way. For discrete light fields however, this would require a suitable interpolation in

ray-space, which gives poor results when the distribution of samples in the target space becomes too sparse.

What follows is a description of two methods to compute the indices for the non-zero entries in P. For simplicity, only a two-dimensional attenuator of size w is assumed, consisting of N layers at various depths $Z_{\min} = z_1 < z_2 < \cdots < z_N = Z_{\max}$. It is also assumed that the image plane s of the virtual cameras is bisecting the attenuator in the middle, at depth $Z_s = \frac{Z_{\max} - Z_{\min}}{2}$.

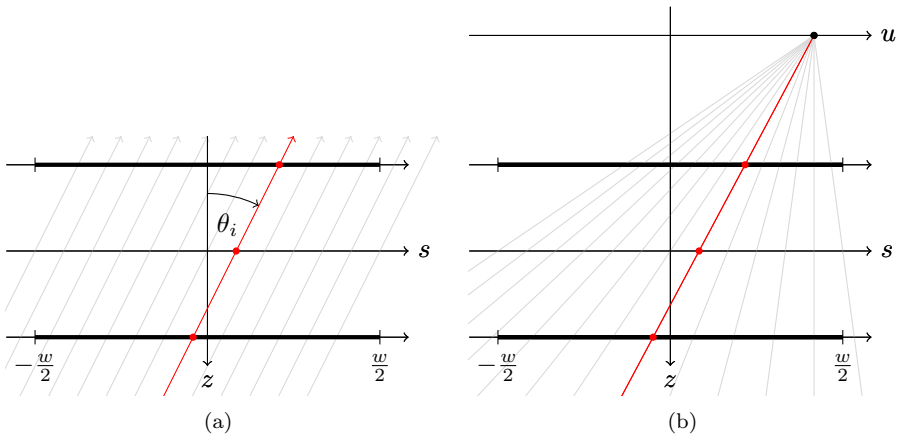


Figure 3.2: Computation of the ray-layer intersections from oblique (a) and perspective (b) projections. Two attenuation layers are drawn (top and bottom) with the virtual image plane in the center. Light rays intersect the layers at positions to be calculated.

Oblique Projection

The setup for the oblique projection type is illustrated in figure 3.2(a). Let θ_i denote the angle of the i-th oblique view from the light field. Following the notation in previous sections, the linear index $m = m(i, k)$ identifies the ray $(\theta_i, s(k))$. The intersection of ray m with the n-th layer is simply

$$h(m, n) = s(k) + \Delta z \tan(\theta_i), \tag{3.12}$$

where $\Delta z = Z_s - z_n$ is the displacement of layer n from the image plane. The next step is to compute the pixel index at the point $h(m, n)$. The shift in pixel units can directly be derived from the shift in equation 3.12 given the pixel size Δs, yielding

$$\Delta k = \left[\frac{\Delta z \tan(\theta_i)}{\Delta s} \right], \tag{3.13}$$

and the new index is $k' = k + \Delta k$. The brackets in the above equation denote the rounding operation. Finally, this information is stored in the propagation matrix with the assignment $P^{(n)}_{mk'} = 1$.

Perspective Projection

For the perspective projection, it is assumed that the (u, v)-plane (or u-plane here) is at depth Z_u, with a distance $Z_s - Z_u$ from the image plane. Again, let $m = m(i, k)$ be the index that identifies the ray $(u(i), s(k))$. Finding the coordinates of the ray (u, s) on a layer is similar to the re-parameterization in equation 2.4. Setting $\gamma = \frac{Z_s - Z_u}{z_n - Z_u}$ for layer n, the formula for the intersection is

$$h(m, n) = \gamma(s(k) - u(i)) + u(i). \tag{3.14}$$

As before, the new pixel index $k' = k + \Delta k$ is computed from the shift in the number of pixels

$$\Delta k = \left[\frac{h(m, n) - s(k)}{\Delta s}\right], \tag{3.15}$$

and the assignment $P^{(n)}_{mk'} = 1$ is made.

3.4 Iterative Reconstruction

The optimization problem in equation 3.11 is essentially a fitting problem. Theoretically, it can be solved in a least squares sense using the normal equation $P^T P\alpha = P^T \bar{L}$ and by inverting the matrix $P^T P$. For high resolution light fields, the matrix P becomes extremely large and it is unfeasible to compute the inverse of $P^T P$.

In general, the approach to solve these kind of problems is to use iterative methods. The choice of the method depends on the type of problem and the structure of the design matrix. In computed tomography, a variety of iterative solvers have been developed to solve the exact same problem. Among the different methods is the Simultaneous Algebraic Reconstruction Technique (SART) first proposed by Andersen and Kak [1984]. The update rule of SART for iteration $k = 0, 1, 2, \ldots$ is

$$\alpha^{(k+1)} = \alpha^{(k)} + \lambda C P^T R \left(\bar{L} - P\alpha^{(k)}\right), \tag{3.16}$$

where λ is a relaxation factor. R and C denote the diagonal matrices with entries $R_{ii} = \frac{1}{r_i}$ and $C_{ii} = \frac{1}{c_i}$, where r_i and c_i are the sum of the elements in the i-th row and column of P respectively. The parts in 3.16 involving P and P^T are also referred to as the **forward-** and **back-projection** respectively.

The convergence of SART has been studied by Jiang and Wang [2001]. They have proven that it converges to a weighted least squares solution.

Chapter 4

Spectral Analysis

This chapter is intended to give an overview of the spectral properties and limitations specific to multiplicative light field displays. Spectral analysis is a crucial method for the quality assessment and it is the origin of a comprehensive understanding of 3D displays. A light field emitted by the display can be interpreted as a signal that is composed of sine waves with different amplitude, phase and frequency. Section 4.1 introduces the Fourier transform, an operation that decomposes such a signal into the frequencies that produce it. The spectral support, i.e. the range of frequencies the display is able to produce, is analyzed in section 4.3.

4.1 Definitions

The **Fourier transform** $\hat{f}$ of an integrable function $f\colon \mathbb{R}^n \to \mathbb{C}$ is defined as

$$\hat{f}(\xi) = \mathcal{F}(f)(\xi) \coloneqq \int_{\mathbb{R}^n} f(x) e^{-2\pi \mathrm{i} x \cdot \xi} \,\mathrm{d}x \tag{4.1}$$

for any $\xi \in \mathbb{R}^n$. According to the Fourier integral theorem, if both f and $\hat{f}$ are absolutely integrable and f is continuous, then the inverse transform

$$f(x) = \mathcal{F}^{-1}(\hat{f})(x) \coloneqq \int_{\mathbb{R}^n} \hat{f}(\xi) e^{2\pi \mathrm{i} x \cdot \xi} \,\mathrm{d}\xi \tag{4.2}$$

is well-defined. The domain of f is called the **spatial domain** and the domain of $\hat{f}$ is referred to as the **frequency domain**. An important property of the Fourier transform is that a convolution in the spatial domain becomes a multiplication in the frequency domain, or in other words,

$$\widehat{(f * g)}(\xi) = \hat{f}(\xi) \cdot \hat{g}(\xi) \tag{4.3}$$

for integrable functions $f, g\colon \mathbb{R}^n \to \mathbb{C}$. On the other hand, a multiplication in the spatial domain becomes a convolution in the frequency domain after applying the Fourier transform, that is

$$\widehat{(f \cdot g)}(\xi) = (\hat{f} * \hat{g})(\xi). \tag{4.4}$$

4.2 Spectral Support of Light Fields

Consider a scene with a bounded depth range between $Z_{\min}$ and $Z_{\max}$. The two objects at the boundaries are shown in figure 4.1(a), with the virtual image s plane between them. The consequent 2D light field $L(u, s)$ (or EPI) is shown in figure 4.1(b). From equation 2.5 it follows that objects appear in the EPI with a slope $\frac{\mathrm{d}u}{\mathrm{d}s} = \frac{z-Z_u}{z-Z_s}$. Substituting z with $Z_{\min}$ and $Z_{\max}$ gives the slopes for the red and blue objects at the boundary, defining the range of slopes in the EPI for objects between the two.

Applying the Fourier transform to the continuous light field reveals that the frequency response is non-zero on lines $\frac{\mathrm{d}s}{\mathrm{d}u}\xi_s + \xi_u = 0$. Again, for the scene with bounded depth range, this yields two lines representing the limits of the spectral support as shown in figure 4.1(c). Objects between the red and blue ones will also have a frequency response within the fan spanned by the two lines. Therefore, the region of support for a continuous light field with bounded depth range can be defined in the following way.

$$\mathcal{S}(\xi_u, \xi_s) \coloneqq \begin{cases} 1, & \text{if } Z_{\min} \leq \dfrac{Z_u\xi_u + Z_s\xi_s}{\xi_u + \xi_s} \leq Z_{\max} \\ 0, & \text{otherwise} \end{cases} \tag{4.5}$$

A similar expression follows for the 4D light field, defining a 4D hyperfan for the region of support as derived by Dansereau et al. [2015]. Note that occlusions as well as specular reflections are not incorporated in the above expression. These effects introduce additional discontinuities in the EPI that result in a high frequency response possibly outside the fan defined in equation 4.5.

4.3 Spectral Support of Multiplicative Displays

Now, consider a single attenuation layer. The light field emitted by this layer has constant depth and thus, the lines in the EPI all have the same slope.

4.4 The Fourier Slice Theorem

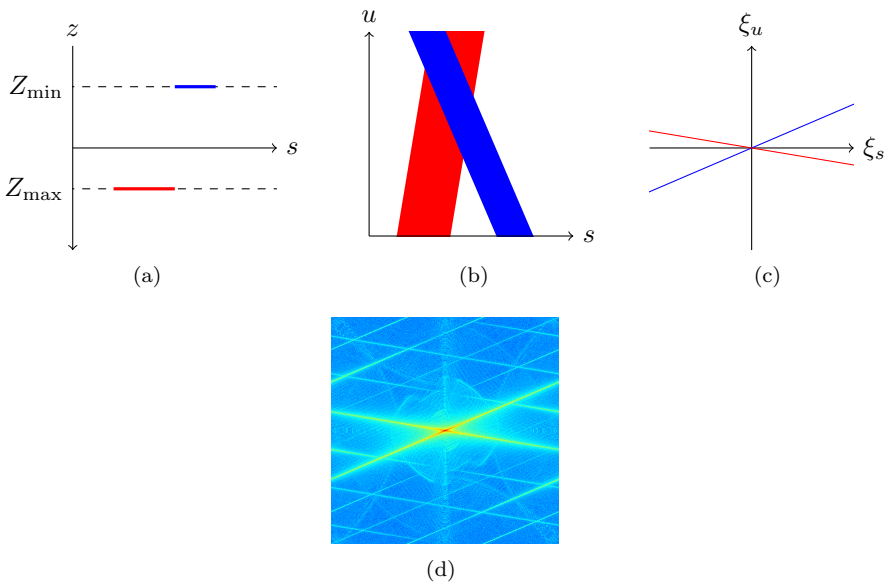


Figure 4.1: (a) Two objects (red and blue) placed at the bounds of the depth range. (b) The EPI representing the 2D light field of the scene. (c) Fourier transform of the EPI. The red and blue line mark the bounds for the spectral support. (d) Discrete Fourier transform of the EPI. Absolute values are represented with colors on a logarithmic scale.

Appendix A

ap1

A.1 apsec1

List of Tables

List of Figures

2.1 Parametrization of the light field with two planes 3
2.2 Parameterization for light fields from oblique projections 4
2.3 Parameterization for light fields from perspective projections . . 5
2.4 Visualization of the light field with epipolar plane images 6

3.1 The Radon transform 8
3.2 Computation of the ray-layer intersections from oblique and perspective projections 10

4.1 Spectral analysis for light fields with bounded depth range 14

Bibliography

E. H. Adelson and J. Bergen. The plenoptic function and the elements of early vision. *Computational Models of Visual Processing*, pages 3–20, 1991.

A. H. Andersen and A. C. Kak. Simultaneous algebraic reconstruction technique (SART): A superior implementation of the ART algorithm. *Ultrasonic Imaging*, 6(1):81–94, 1984.

R. C. Bolles, H. H. Baker, and D. H. Marimont. Epipolar-plane image analysis: An approach to determining structure from motion. *International Journal of Computer Vision*, pages 7–55, 1987.

J.-X. Chai, X. Tong, S.-C. Chan, and H.-Y. Shum. Plenoptic sampling. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '00, pages 307–318. ACM, 2000.

D. G. Dansereau, O. Pizarro, and S. B. Williams. Decoding, calibration and rectification for lenselet-based plenoptic cameras. In *Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1027–1034. IEEE Computer Society, 2013.

D. G. Dansereau, O. Pizarro, and S. B. Williams. Linear volumetric focus for light field cameras. *ACM Trans. Graph.*, 34(2):15:1–15:20, 2015.

F. Durand, N. Holzschuch, C. Soler, E. Chan, and F. X. Sillion. A frequency analysis of light transport. *ACM Trans. Graph.*, 24(3):1115–1126, 2005.

S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen. The lumigraph. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '96, pages 43–54, New York, NY, USA, 1996. ACM.

F.-C. Huang, G. Wetzstein, B. A. Barsky, and R. Raskar. Eyeglasses-free display: Towards correcting visual aberrations with computational light field displays. *ACM Trans. Graph.*, 33(4):59:1–59:12, 2014.

A. Isaksen, L. McMillan, and S. J. Gortler. Dynamically reparameterized light fields. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '00, pages 297–306. ACM, 2000.

M. Jiang and G. Wang. Convergence of the simultaneous algebraic reconstruction technique (SART). In *Conference Record of the Thirty-Fifth Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 360–364, 2001.

C. Kim, A. Hornung, S. Heinzle, W. Matusik, and M. Gross. Multi-perspective stereoscopy from light fields. *ACM Trans. Graph.*, 30(6):190:1–190:10, 2011.

M. Levoy and P. Hanrahan. Light field rendering. In *Proceedings of the International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '96)*, pages 31–42, 1996.

C.-K. Liang and R. Ramamoorthi. A light transport framework for lenslet light field cameras. *ACM Trans. Graph.*, 34(2):16:1–16:19, 2015.

Y. Lu, W. Wang, S. Chen, Y. Xie, J. Qin, W.-M. Pang, and P.-A. Heng. Accelerating algebraic reconstruction using CUDA-enabled GPU. In *Sixth International Conference on Computer Graphics, Imaging and Visualization*, pages 480–485. IEEE, 2009.

R. Narain, R. A. Albert, A. Bulbul, G. J. Ward, M. S. Banks, and J. F. O'Brien. Optimal presentation of imagery with focus cues on multi-plane displays. *ACM Trans. Graph.*, 34(4):59:1–59:12, 2015.

R. Ng. Fourier slice photography. *ACM Trans. Graph.*, 24(3):735–744, 2005.

R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan. Light field photography with a hand-held plenoptic camera. *Computer Science Technical Report CSTR*, 2(11), 2005.

V. Vaish, B. Wilburn, N. Joshi, and M. Levoy. Using plane + parallax for calibrating dense camera arrays. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1. IEEE, 2004.

G. Wetzstein, D. Lanman, W. Heidrich, and R. Raskar. Layered 3D: Tomographic image synthesis for attenuation-based light field and high dynamic range displays. *ACM Trans. Graph.*, 30(4):95:1–95:12, 2011.

G. Wetzstein, D. Lanman, M. Hirsch, and R. Raskar. Tensor displays: Compressive light field synthesis using multilayer displays with directional backlighting. *ACM Trans. Graph.*, 31(4):80:1–80:11, 2012.

M. Yan. Convergence analysis of SART by bregman iteration and dual gradient descent. pages 1–15, 2010.

M. Zwicker, W. Matusik, F. Durand, H. Pfister, and C. Forlines. Antialiasing for automultiscopic 3D displays. In *ACM SIGGRAPH 2006 Sketches*, SIGGRAPH '06, New York, NY, USA, 2006. ACM.

Erklärung

gemäss Art. 28 Abs. 2 RSL 05

Name/Vorname: .....

Matrikelnummer: .....

Studiengang: ……

Bachelor ☐ Master ☐ Dissertation ☐

Titel der Arbeit: .....

.....

.....

LeiterIn der Arbeit: .....

.....

Ich erkläre hiermit, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe o des Gesetztes vom 5. September 1996 über die Universität zum Entzug des auf Grund dieser Arbeit verliehenen Titels berechtigt ist.

.....

Ort/Datum

.....

Unterschrift