Fourier Slice Photography

Ren Ng
Stanford University

Abstract

This paper contributes to the theory of photograph formation from light fields. The main result is a representation of (lens-based) photographic imaging in the Fourier domain: the 2D Fourier transform of a photograph is a 2D slice in the 4D Fourier spectrum of the light field. Photographs focused at different depths correspond to slices at different rotations in the 4D space. The paper demonstrates the utility of this theorem in two different ways. First, the theorem is used to analyze the performance of digital refocusing, where one computes photographs focused at different depths from a single light field. The analysis shows in closed form that the sharpness of refocused photographs increases linearly with directional resolution. Second, the theorem yields a Fourier-domain algorithm for digital refocusing, where we extract the appropriate 2D slice of the light field's Fourier transform, and perform an inverse 2D Fourier transform. This method is faster than previous approaches.

Keywords: Digital photography, Fourier transform, projection-slice theorem, refocusing, synthetic aperture photography, plenoptic sampling.

1 Introduction

A light field is a representation of the light flowing allowing all rays in free-space. We can synthesize pictures by computationally tracing these rays to where they would have terminated in a desired imaging system. In synthetic photography, for example, we conceptually trace the rays through an imaginary camera and sum them on the synthetic film plane. Classical light field rendering assumes a pin-hole camera model [Levoy and Hanrahan 1996; Gortler et al. 1996], but we have seen increasing interest in modeling a realistic camera with a lens that creates finite depth of field [Isaksen et al. 2000; Vaish et al. 2004; Levoy et al. 2004]. *Digital refocusing* is the process by which we control the film plane of the synthetic camera to produce photographs focused at different depths in the scene (see bottom of Figure 8).

Synthetic photography of traditional photographic subjects, including portraits, high-speed action and macro close-ups, is made possible with a hand-held *plenoptic camera* [Anonymous 2005; Adelson and Wang 1992]. It is practical to turn a conventional camera into a plenoptic camera by simply inserting a microlens array in front of the photosensor. The pixels under each microlens measure the amount of light striking that microlens along each incident ray. In this way, the sensor samples the in-camera light field in a single photographic exposure.

This paper presents a new mathematical theory about photographic imaging by deriving its Fourier-domain representation. The theory is derived from the geometrical optics of image formation, and makes use of the well-known Fourier Slice Theorem [Bracewell 1956]. The end result is the *Fourier Slice Photography Theorem* (Section 3.2), which states that the 2D Fourier transform of a photograph is a 2D slice in the 4D Fourier spectrum of the light field. Photographs focused at different depths correspond to slices at different rotations in the 4D space. This Fourier representation is mathematically simpler than the more common, spatial-domain representation, which is based on integration rather than slicing.

The closest related work, which inspired this paper, is the plenoptic sampling analysis of Chai et al. [2000]. They study the sampling rates for classical light field rendering in terms of the Fourier spectrum of the light field. A key difference between our work is that Chai et al. focus on the classical problem of rendering pin-hole images from light fields, whereas this paper analyzes the formation of photographs through lenses.

Sections 4 and 5 apply the Fourier Slice Photograph Theorem in two very different ways. Section 4 uses it to theoretically analyze the performance of digital refocusing with a band-limited plenoptic camera. The theorem enables a closed-form analysis showing that the sharpness of refocused photographs increases linearly with the number of samples under each microlens. This result supports recent experimental work on the plenoptic camera [Anonymous 2005], and provides theoretical tools for future camera designs.

Section 5 applies the theorem in a very different manner to derive a fast *Fourier Slice Digital Refocusing* algorithm. This algorithm computes photographs by extracting the appropriate 2D slice of the light field's Fourier transform and performing an inverse Fourier transform. The asymptotic complexity of this algorithm is $O(n^2 \log n)$, compared to the $O(n^4)$ approach of existing algorithms, which are essentially different approximations of numerical integration in the 4D spatial domain.

2 Background

Consider the light flowing along all rays inside the camera. Let $\overline{L}_F$ be a two-plane parameterization of this light field, where $\overline{L}_F(u, v, x, y)$ is the radiance along the ray traveling from position (u, v) on the lens plane to position (x, y) on the sensor plane (see Figure 1). F is the distance between the lens and the sensor. Let us consider how photographs are formed from the light field in conventional cameras and plenoptic cameras [Anonymous 2005].

Conventional Camera The image that forms inside a conventional camera is proportional to the irradiance [Stroebel et al. 1986], which is equal to a weighted sum of the radiance coming through the lens:

$$E_F(x, y) = \frac{1}{F^2} \iint \overline{L}_F(u, v, x, y) \cos^4 \phi \, du \, dv, \qquad (1)$$

where F is the separation between the lens and the film, $E_F(x, y)$ is the irradiance on the film at (x, y), and ϕ is the angle between ray (u, v, x, y) and the film plane normal. The integration is a physical process that takes place on the sensor surface, such as the accumulation of electrons in a pixel of a CCD that is exposed to light.

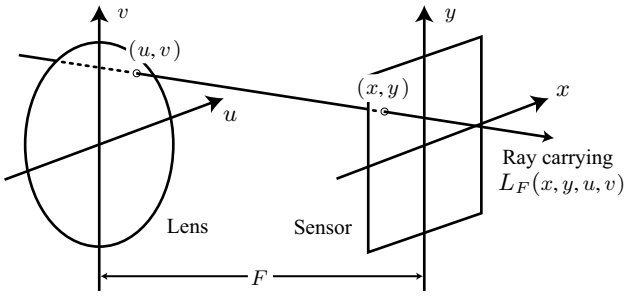

Figure 1: We parameterize the 4D light field, L_F, inside the camera by two planes. The uv plane is the principal plane of the lens, and the xy plane is the sensor plane. $L_F(u, v, x, y)$ is the radiance along the given ray.

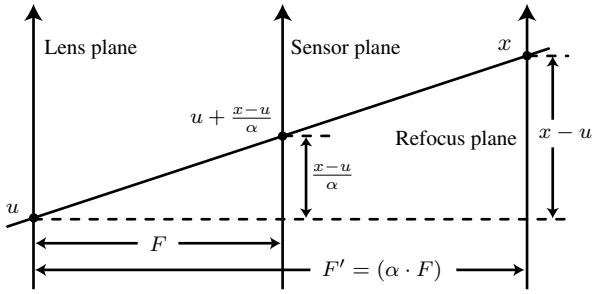

Figure 2: Reparameterizing the light field by moving the sensor plane from F to $F' = (\alpha \cdot F)$. The diagram shows the simplified 2D case involving only x and u. By similar triangles, the illustrated ray that intersects the lens at u, and the F' plane at x, also intersects the F plane at $u + (x - u)/\alpha$.

The derivations below assume that the uv and xy planes are infinite in extent, and that $\overline{L}$ is simply zero beyond the physical bounds of the lens and sensor. To shorten the equations, they also absorb the $\cos^4\phi$ into the light field itself, be defining $L(u, v, x, y) = \overline{L}(u, v, x, y)\cos^4\phi$. This contraction is possible because ϕ depends only on the angle that the ray makes with the light field planes.

As a final note about Eq. 1, it is worth mentioning that the light field inside the camera is related to the light field in the world via the focal length of the lens and the thin lens equation. To keep the equations as simple as possible, however, the derivations deal exclusively with light fields inside the camera.

Plenoptic Camera In the case of a plenoptic camera that measures light fields, image formation involves two steps: measurement and processing. For measurement, this kind of camera uses a flat sensor that provides a directional sampling of the radiance passing through each point on the sensor. Hence, if the sensor is at a depth F from the lens, it samples L_F directly. During processing, this light field can be used to compute the conventional photograph at *any* depth $E_{F'}$, where F' need not be the same as F. This is done by reparameterizing L_F to produce $L_{F'}$ and then applying Eq. 1.

A simple geometric construction (see Figure 2) shows that if we let $\alpha = F'/F$,

$$\begin{aligned} L_{F'}(u, v, x, y) &= L_{(\alpha\cdot F)}(u, v, x, y) \\ &= L_F(u,\ v,\ u + (x - u)/\alpha,\ v + (y - v)/\alpha) \\ &= L_F\left(u,\ v,\ u(1 - 1/\alpha) + x/\alpha,\ v(1 - 1/\alpha) + y/\alpha\right). \end{aligned} \quad (2)$$

In other words, $L_{F'}$ is a 4D shear of L_F. Combining Eqs. 1 and 2 leads to the central definition of this section, which codifies the fundamental relationship between photographs and light fields:

Photography Operator Let $\mathcal{P}_\alpha$ be the operator that transforms a light field at a sensor depth F into the photograph formed on film at depth $(\alpha \cdot F)$. If $\mathcal{P}_\alpha[L_F]$ represents the application of $\mathcal{P}_\alpha$ to light field L_F, then

$$\mathcal{P}_\alpha[L_F](x, y) = E_{(\alpha\cdot F)}(x, y) = \quad (3)$$
$$\frac{1}{\alpha^2 F^2}\iint L_F\left(u,\ v,\ u(1-1/\alpha) + x/\alpha,\ v(1-1/\alpha) + y/\alpha\right)\, du\, dv.$$

This definition is the basis for digital refocusing, in that it explains how to compute photographs at different depths from a single measurement of the light field inside the camera. Note that the photography operator can be thought of as shearing the 4D space, and then projecting down to 2D.

3 Photographic Imaging in Fourier-Space

The key to analyzing Eq. 3 in the Fourier domain is the Fourier slice theorem (also known as the Fourier projection-slice theorem and the central section theorem), which was discovered by Bracewell [1956] in the context of radio astronomy. This theorem is the theoretical foundation of many medical imaging techniques [Macovski 1983] including CT (computed tomography) and MR (magnetic resonance). The classical version of the Fourier slice theorem [Deans 1983] states that a 1D slice of a 2D function's Fourier spectrum is the Fourier transform of an orthographic integral projection of the 2D function. The projection and slicing geometry is illustrated in Figure 3.

Conceptually, the theorem works because the value at the origin of frequency space gives the DC value (integrated value) of the signal, and rotations do not fundamentally change this fact. From this perspective, it makes sense that the theorem generalizes to higher dimensions. It also makes sense that the theorem works for shearing operations as well as rotations, because shearing a space is equivalent to rotating and dilating the space.

These observations mean that we can expect that the photography operator, which we have observed is a shear followed by projection, should be proportional to a dilated 2D slice of the light field's 4D Fourier transform. With this intuition in mind, Section 3.1 and 3.2 are simply the mathematical derivations in specifying this slice precisely, culminating in Eqs. 9 and 10.

3.1 Generalization of the Fourier Slice Theorem

Let us first digress to study a generalization of the Fourier Slice Theorem to higher dimensions, so that we can apply it in our 4D space. Similar generalizations certainly exist [Mersereau and Oppenheim 1974; Deans 1983], but these tend to deal with projection from N dimensions down to $(N - 1)$ dimensions, which is not directly suitable for our application.

An important idea in the generalization presented here is to observe that all projections and slices of a function, in any dimension, can be expressed as an appropriate rotation of the function followed by a canonical projection or slicing. This approach is embodied in the following operator definitions.

Integral Projection Let $\mathcal{I}_M^N$ be the canonical projection operator that reduces an N-dimensional function down to M-dimensions by integrating out the last $N - M$ dimensions: $\mathcal{I}_M^N[f](x_1, \ldots, x_M) = \int f(x_1, \ldots, x_N)\, dx_{M+1} \ldots dx_N$.

Slicing Let $\mathcal{S}_M^N$ be the canonical slicing operator that reduces an N-dimensional function down to an M dimensional one by zero-ing out the last $N - M$ dimensions: $\mathcal{S}_M^N[f](x_1, \ldots, x_M) = f(x_1, \ldots, x_M, 0, \ldots, 0)$.

Rotation Let $\mathcal{R}$ denote an operator for an arbitrary rotation of an N-dimensional function. It is convenient to also allow $\mathcal{R}$ to act on N-dimensional column vectors as an $N \times N$ rotation matrix, so that $\mathcal{R}[f](\mathbf{x}) = f(\mathcal{R}^{-1}\mathbf{x})$, where $\mathbf{x}$ is an N-dimensional column vector, and $\mathcal{R}^{-1}$ is the inverse of $\mathcal{R}$.

Fourier Transform Let $\mathcal{F}^N$ denote the N-dimensional Fourier transform operator, and let $\mathcal{F}^{-N}$ be its inverse.

With these four definitions, we can state a generalization of the Fourier slice theorem as follows:

THEOREM (GENERALIZED FOURIER SLICE). *Let f be an N-dimensional function. If we rotate f, integral-project it down to M of its dimensions, and Fourier transform the resulting function, the result is equivalent to Fourier transforming f, transforming the result by the same rotation and slicing it down to M dimensions. Compactly in terms of operators, the theorem says:*

$$\boxed{\mathcal{F}^M \circ \mathcal{I}_M^N \circ \mathcal{R} \equiv \mathcal{S}_M^N \circ \mathcal{R} \circ \mathcal{F}^N.} \tag{4}$$

A proof of the theorem is presented in Appendix A.

Figure 4 summarizes the relationships implied by the theorem between the N-dimensional signal, M-dimensional projected signal, and their Fourier spectra. A trivial point to note about the theorem is that it reduces to the classical version for $N = 2$ and $M = 1$ (compare Figures 3 and 4).

3.2 Fourier Slice Photography

We are now ready to express the Photography Operator (Eq. 3) in the Fourier domain, using the Generalized Fourier Slice Theorem (Eq. 4).

3.2.1 Photographic Imaging as Integral Projection

Remember that the photography operator (Eq. 3) corresponds to a shear and integral projection of the light field. A rotation-projection may be factorized into a shear-projection-dilation, a fact sometimes used in volume rendering systems [Lacroute and Levoy 1994]. Conceptually, this means that the photography operator can be factorized into a rotation, projection and dilation:

$$\mathcal{P}_\alpha \equiv \mathcal{W}_\alpha^2 \circ \mathcal{I}_2^4 \circ \mathcal{R}_\alpha^4, \tag{5}$$

which relies on the following two additional operators:

Photography Warp $\mathcal{W}_\alpha^2$ is a 2D scaling and dilation such that $\mathcal{W}_\alpha^2[f](x, y) = (\cos\theta/(\alpha \cdot F))^2 f(x \cos\theta/\alpha, y \cos\theta/\alpha)$, where $\theta = -\arctan(1 - 1/\alpha)$.

Photography Rotation $\mathcal{R}_\alpha^4$ is a 4D rotation operator defined by $\mathcal{R}_\alpha^4[f](x, y, u, v) = f(u\cos\theta - x\ \sin\theta, v\cos\theta - y\sin\theta,\ x\cos\theta + u\sin\theta, y\cos\theta + v\sin\theta)$, where $\theta = -\arctan(1 - 1/\alpha)$.

Given these definitions, verifying that Eq. 5 is consistent with Eq. 3 is a matter of straightforward but rather tedious algebra, which is omitted here.

3.2.2 Photographic Imaging as Fourier Slice

The point of these manipulations is that since $\mathcal{R}_\alpha^4$ is a 4D rotation operator, we can apply the Fourier Slice Theorem to turn the projection into a slice. Substituting $(\mathcal{F}^{-2} \circ \mathcal{S}_2^4 \circ \mathcal{R}_\alpha^4 \circ \mathcal{F}^4)$ for $(\mathcal{I}_2^4 \circ \mathcal{R}_\alpha^4)$, we arrive at the following result:

$$\mathcal{P}_\alpha \equiv \mathcal{W}_\alpha^2 \circ \mathcal{F}^{-2} \circ \mathcal{S}_2^4 \circ \mathcal{R}_\alpha^4 \circ \mathcal{F}^4, \tag{6}$$

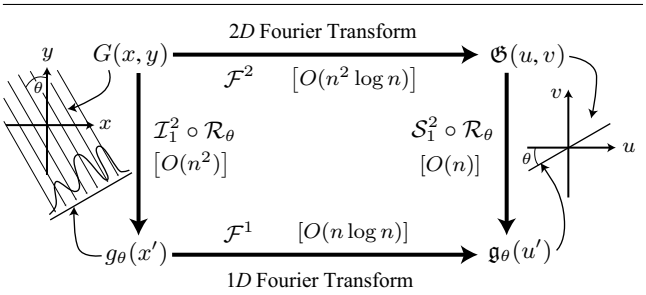

Figure 3: Classical Fourier Slice Theorem, using the operator notation developed in Section 3.1. Computational complexities for each transform are given in square brackets, assuming n samples in each dimension.

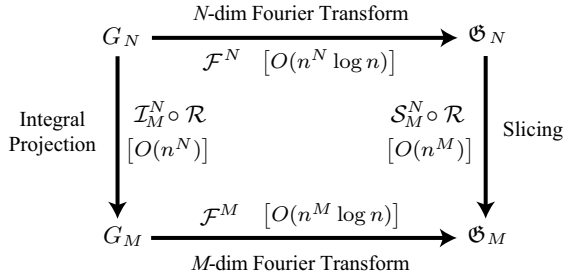

Figure 4: Generalized Fourier Slice Theorem (Eq. 4). Transform relationships between an N-dimensional function G_N, an M-dimensional integral projection of it, G_M, and their respective Fourier spectra, $\mathfrak{G}_N$ and $\mathfrak{G}_M$. n is the number of samples in each dimension.

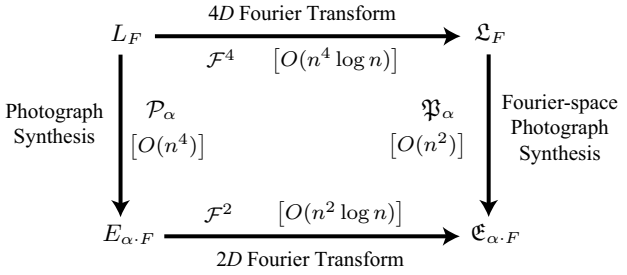

Figure 5: Fourier Slice Photography Theorem (Eq. 10). Transform relationships between the 4D light field L_F, a lens-formed 2D photograph $E_{\alpha \cdot F}$, and their respective Fourier spectra, $\mathfrak{L}_F$ and $\mathfrak{E}_{\alpha \cdot F}$. n is the number of samples in each dimension.

namely that a lens-formed photograph is obtained from the 4D Fourier spectrum of the light field by: extracting an appropriate 2D slice ($\mathcal{S}_2^4 \circ \mathcal{R}_\alpha^4$), applying an inverse 2D transform ($\mathcal{F}^{-2}$), and scaling and dilating the resulting image ($\mathcal{W}_\alpha^2$).

3.2.3 Notational Improvements

It is worth making two notational changes to bring us to an elegant form for the final theorem. First, we can fold the warp into the frequency domain to arrive at the following identity:

$$\mathcal{P}_\alpha \equiv \mathcal{F}^{-2} \circ \mathfrak{W}_\alpha^2 \circ \mathcal{S}_2^4 \circ \mathcal{R}_\alpha^4 \circ \mathcal{F}^4, \tag{7}$$

where $\mathfrak{W}_\alpha^2$ is the dual dilation scaling operator to $\mathcal{W}_\alpha^2$, such that $\mathfrak{W}_\alpha^2[G](k_x, k_y) = (1/F^2)\, G(k_x(\alpha/\cos\theta), k_y(\alpha/\cos\theta))$ where

$\theta = -\arctan(1 - 1/\alpha)$. Eq. 7 follows from Eq. 6 with the identity that $\mathcal{F}^2 \circ \mathcal{W}_\alpha^2 \equiv \mathfrak{W}_\alpha^2 \circ \mathcal{F}^2$. This is a consequence of the Fourier Similarity Theorem [Bracewell 1986].

Second, let us define one last operator that combines all the operators that act in the Fourier domain:

Fourier Photography Operator

$$\mathfrak{P}_\alpha \equiv (\mathfrak{W}_\alpha^2 \circ \mathcal{S}_2^4 \circ \mathcal{R}_\alpha^4). \tag{8}$$

It is easy to verify that $\mathfrak{P}_\alpha$ has the following explicit definition, directly from the definitions of $\mathfrak{W}_\alpha^2$, $\mathcal{S}_2^4$ and $\mathcal{R}_\alpha^4$. This explicit form is required for calculations:

$$\mathfrak{P}_\alpha[G](k_x, k_y) \tag{9}$$
$$= (1/F^2)\, G((1-\alpha)\cdot k_x,\ (1-\alpha)\cdot k_y,\ \alpha \cdot k_x,\ \alpha \cdot k_y).$$

This equation shows that, of course, $\mathfrak{P}_\alpha$ is still an efficient slicing operation, although not an orthographic slice.

3.2.4 Final Theorem

Applying Eq. 8 to Eq. 7 brings us, finally, to our goal.

THEOREM (FOURIER SLICE PHOTOGRAPHY).

$$\boxed{\mathcal{P}_\alpha \equiv \mathcal{F}^{-2} \circ \mathfrak{P}_\alpha \circ \mathcal{F}^4} \tag{10}$$

A photograph is the inverse 2D Fourier transform of a dilated 2D slice in the 4D Fourier transform of the light field.

Figure 5 illustrates the relationships implied by this theorem.

From an intellectual standpoint, the value of the theorem lies in the fact that $\mathfrak{P}_\alpha$, a slicing operator, is conceptually simpler than $\mathcal{P}_\alpha$, an integral operator. This point is made especially clear by reviewing the explicit definitions of $\mathfrak{P}_\alpha$ (Eq. 9) and $\mathcal{P}_\alpha$ (Eq. 3) . By providing a frequency-based interpretation, the theorem contributes insight by providing two equivalent but very different perspectives on photographic imaging. In this regard, the Fourier Slice Photography Theorem is not unlike the Fourier Convolution Theorem, which provides equivalent but very different perspectives of convolution in the two domains.

From a practical standpoint, the theorem provides a faster computational pathway for certain kinds of light field processing. The computational complexities for each transform are illustrated in Figure 5, but the main point is that slicing via $\mathfrak{P}_\alpha$ $(O(n^2))$ is asymptotically faster than integration via $\mathcal{P}_\alpha$ $(O(n^4))$. This fact is the basis for the algorithm in Section 5.

4 Theoretical Limits of Digital Refocusing

The overarching goal of this section is to demonstrate the theoretical utility of the Fourier Slice Photography Theorem. Section 4.1 presents a general signal-processing theorem, showing exactly what happens to photographs when a light field is distorted by a convolution filter. Section 4.2 applies this theorem to analyze the performance of a band-limited light field camera. In these derivations, we will often use the Fourier Slice Photography Theorem to move the analysis into the frequency domain, where it becomes simpler.

4.1 Photographic Effect of Filtering the Light Field

A light field produces exact photographs focused at various depths via Eq. 3. If we distort the light field by filtering it, and then form photographs from the distorted light field, how are these photographs related to the original, exact photographs? The following theorem provides the answer to this question.

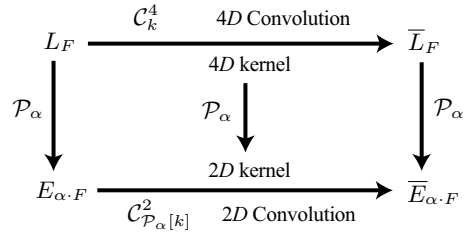

Figure 6: Filtered Light Field Photograph Theorem (Eq. 11). L_F is the input 4D light field, and $\overline{L}_F$ is a 4D filtering of it with 4D kernel k. $E_{\alpha\cdot F}$ and $\overline{E}_{\alpha\cdot F}$ are the best photographs formed from the two light fields, where the photographs are focused with focal plane depth $(\alpha \cdot F)$. The theorem shows that $\overline{E}_{\alpha\cdot F}$ is a 2D filtering of $E_{\alpha\cdot F}$, where the 2D kernel is the photograph of the 4D kernel, k.

THEOREM (FILTERED LIGHT FIELD PHOTOGRAPHY). *A 4D convolution of a light field results in a 2D convolution of each photograph. The 2D filter kernel is, somewhat surprisingly, simply the photograph of the 4D filter kernel focused at the same depth. Compactly in terms of operators,*

$$\boxed{\mathcal{P}_\alpha \circ \mathcal{C}_k^4 \equiv \mathcal{C}^2_{\mathcal{P}_\alpha[k]} \circ \mathcal{P}_\alpha,} \tag{11}$$

where we have expressed convolution with the following operator:

Convolution $\mathcal{C}_k^N$ is an N-dimensional convolution operator with filter kernel k (an N-dimensional function), such that $\mathcal{C}_k^N[F](\mathbf{x}) = \int F(\mathbf{x} - \mathbf{u})\ k(\mathbf{u})\ d\mathbf{u}$ where $\mathbf{x}$ and $\mathbf{u}$ are N-dimensional vector coordinates and F is an N-dimensional function.

Figure 6 illustrates the theorem diagramatically. It is worth noting that in spite of its simplicity and plausibility, the theorem is not obvious, and proving it in the spatial domain is quite difficult. Appendix B presents a proof of the theorem in the frequency-domain. At a high level, the approach is to apply the Fourier Slice Photography Theorem and the Convolution Theorem to move the analysis into the frequency domain. In that domain, photograph formation turns into a simpler slicing operator, and convolution turns into a simpler multiplication operation.

This theorem is useful because it is simple and general. Note that it is not at all specific to Fourier slice methods, but applies generally to photographic imaging as a physical process. The next section contains a concrete example of how to use the theorem, but it should be emphasized that the theorem is much more broadly applicable. It will be worth exploiting it in general analysis of light field acquisition, where the system impulse response is the filter kernel, $k(u, v, x, y)$, and light field processing, where the resampling strategy defines $k(u, v, x, y)$.

4.2 Band-Limited Plenoptic Camera

This section analyzes digital refocusing from a plenoptic camera [Anonymous 2005], to answer the following questions. What is the quality of the photographs refocused from the acquired light fields? How are these photographs related to the exact photographs, such as those that might be taken by a conventional camera that were optically focused at the same depth?

The central assumption here, from which we will derive significant analytical leverage, is that the plenoptic camera captures band-limited light fields. While perfect band-limiting is physically impossible, it is a plausible approximation in this case because the

camera system blurs the incoming signal through imperfections in its optical elements, through area integration over the physical extent of microlenses and photosensor pixels, and ultimately through diffraction.

The band-limited assumption means that the acquired light field, $\hat{L}_{F_L}$, is a simply the *exact* light field, L_{F_L}, convolved by a perfect low-pass filter, a 4D sinc:

$$\hat{L}_{F_L} = \mathcal{C}^4_{\text{lowpass}}\left[L_{F_L}\right], \text{ where} \tag{12}$$

$$\text{lowpass}(k_u, k_v, k_x, k_y) = 1/(\Delta x \Delta u)^2 \cdot \text{sinc}(k_u/\Delta u,\ k_v/\Delta u,\ k_x/\Delta x,\ k_y/\Delta x). \tag{13}$$

In this equation, Δx and Δu are the linear spatial and directional sampling rates of the integrated light field camera, respectively. The $1/(\Delta x \Delta u)^2$ is an energy-normalizing constant to account for dilation of the sinc. Also note that, for compactness, we use multi-dimensional notation so that $\text{sinc}(u, v, x, y) = \text{sinc}(x)\,\text{sinc}(y)\,\text{sinc}(u)\,\text{sinc}(v)$.

4.2.1 Analytic Form for Refocused Photographs

Our goal is an analytic solution for the digitally refocused photograph, $\hat{E}_F$, computed from the band-limited light field, L_{F_L}. This is where we apply the Filtered Light Field Photography Theorem. Letting , $\alpha = F/F_L$,

$$\begin{aligned}\hat{E}_F &= \mathcal{P}_\alpha\left[\hat{L}_{F_L}\right] = \mathcal{P}_\alpha\left[\mathcal{C}^4_{\text{lowpass}}\left[L_{F_L}\right]\right] \\ &= \mathcal{C}^2_{\mathcal{P}_\alpha[\text{lowpass}]}\left[\mathcal{P}_\alpha\left[L_{F_L}\right]\right] = \mathcal{C}^2_{\mathcal{P}_\alpha[\text{lowpass}]}\left[E_F\right],\end{aligned} \tag{14}$$

where E_F is the *exact* photograph at depth F. This derivation shows that the digitally refocused photograph is a 2D-filtered version of the exact photograph. The 2D kernel is simply a photograph of the 4D sinc function interpreted as a light field, $\mathcal{P}_\alpha\left[\text{lowpass}\right]$.

Somewhat surprisingly, it turns out that photographs of a 4D sinc light field are simply 2D sinc functions:

$$\begin{aligned}&\mathcal{P}_\alpha\left[\text{lowpass}\right] \\ &\quad= \mathcal{P}_\alpha\left[1/(\Delta x \Delta u)^2 \cdot \text{sinc}(k_u/\Delta u, k_v/\Delta u, k_x/\Delta x, k_y/\Delta x)\right] \\ &\quad= 1/D_x^2 \cdot \text{sinc}(k_x/D_x, k_y/D_x),\end{aligned} \tag{15}$$

where the Nyquist rate of the 2D sinc depends on the amount of refocusing, α:

$$D_x = \max(\alpha\Delta x, |1-\alpha|\Delta u). \tag{16}$$

This fact is difficult to derive in the spatial domain, but applying the Fourier Slice Photography Theorem moves the analysis into the frequency domain, where it is easy (see Appendix C).

The end result here is that, since the 2D kernel is a sinc, the band-limited camera produces digitally refocused photographs that are just band-limited versions of the exact photographs. The performance of digital refocusing is defined by the variation of the 2D kernel band-width (Eq. 16) with the extent of refocusing.

4.2.2 Interpretation of Refocusing Performance

Notation Recall that the spatial and directional sampling rates of the camera are Δx and Δu. Let us further define the width of the camera sensor as W_x, and the width of the lens aperture as W_u. With these definitions, the spatial resolution of the sensors is $N_x = W_x/\Delta x$ and the directional resolution of the light field camera is $N_u = W_u/\Delta u$.

Exact Refocusing Since $\alpha = (F/F_L)$ and $\Delta u = W_u/N_u$, it is easy to verify that

$$\begin{aligned}&|\alpha\Delta x| \geq |(1-\alpha)\Delta u| \\ \Leftrightarrow\ &|F - F_L| \leq \Delta x(N_u F/W_u).\end{aligned} \tag{17}$$

The claim here is that this is the range of focal depths, F_L, where we can achieve "exact" refocusing, i.e. compute a sharp rendering of the photograph focused at that depth. What we are interested in is the Nyquist-limited resolution of the photograph, which is the number of band-limited samples within the field of view.

Precisely, by applying Eq. 17 to Eq. 16, we see that the bandwidth of the computed photograph is $(\alpha\Delta x)$. Next, the field of view is not simply the size of the light field sensor, W_x, but rather (αW_x). This dilation is due to the fact that digital refocusing scales the image captured on the sensor by a factor of α in projecting it onto the refocus focal plane (see Eq. 3). If $\alpha > 1$, for example, the light field camera image is zoomed in slightly compared to the conventional camera. Figure 7 illustrates this effect.

Thus, the Nyquist resolution of the computed photograph is

$$(\alpha W_x)/(\alpha\Delta x) = W_x/\Delta x. \tag{18}$$

This is simply the spatial resolution of the camera, the maximum possible resolution for the output photograph. This justifies the assertion that the refocusing is "exact" for the range of depths defined by Eq. 17. Note that this range of exact refocusing *increases linearly with the directional resolution, N_u.*

Inexact Refocusing If we exceed the exact refocusing range, i.e.

$$|F - F_L| > \Delta x(N_u F/W_u). \tag{19}$$

then the band-limit of the computed photograph, $\hat{E}_F$, is $|1-\alpha|\Delta u < \alpha\Delta x$ (see Eq. 16), and the resulting resolution is not maximal, but rather $(\alpha W_x)/(|1-\alpha|\Delta u)$, which is less than $W_x/\Delta x$. In other words, the resulting photograph is blurred, with reduced Nyquist-limited resolution.

Re-writing this resolution in a slightly different form provides a more intuitive interpretation of the amount of blur. Since $\alpha = F/F_L$ and $\Delta u = W_u/N_u$, the resolution is

$$\frac{\alpha W_x}{|1-\alpha|\Delta u} = \frac{W_x}{W_u/(N_u \cdot F) \cdot |F - F_L|}. \tag{20}$$

Since $((N_u F)/W_u)$ is the f-number of a lens N_u times smaller than the actual lens used on the camera, we can now interpret $W_u/(N_u \cdot F) \cdot |F - F_L|$ as the size of the conventional circle of confusion cast through this smaller lens when the film plane is misfocused by a distance of $|F - F_L|$.

In other words, when refocusing beyond the exact range, we can only make the desired focal plane appear as sharp as it appears in a conventional photograph focused at the original depth, with a lens N_u times smaller. Note that the sharpness *increases linearly with the directional resolution, N_u.*

4.3 Summary

It is worth summarizing the point of the analysis in Section 4. On a meta-level, this section has demonstrated the theoretical utility of the Fourier Slice Photography Theorem, applying it several times in deriving Eqs. 11, 14 and 15.

At another level, this section has derived two end results that are of some importance. The first is the Filtered Light Field Photography Theorem, which is a simple but general signal-processing tool for analyzing light field imaging systems. The second is the fact that making a simple band-limited assumption about plenoptic

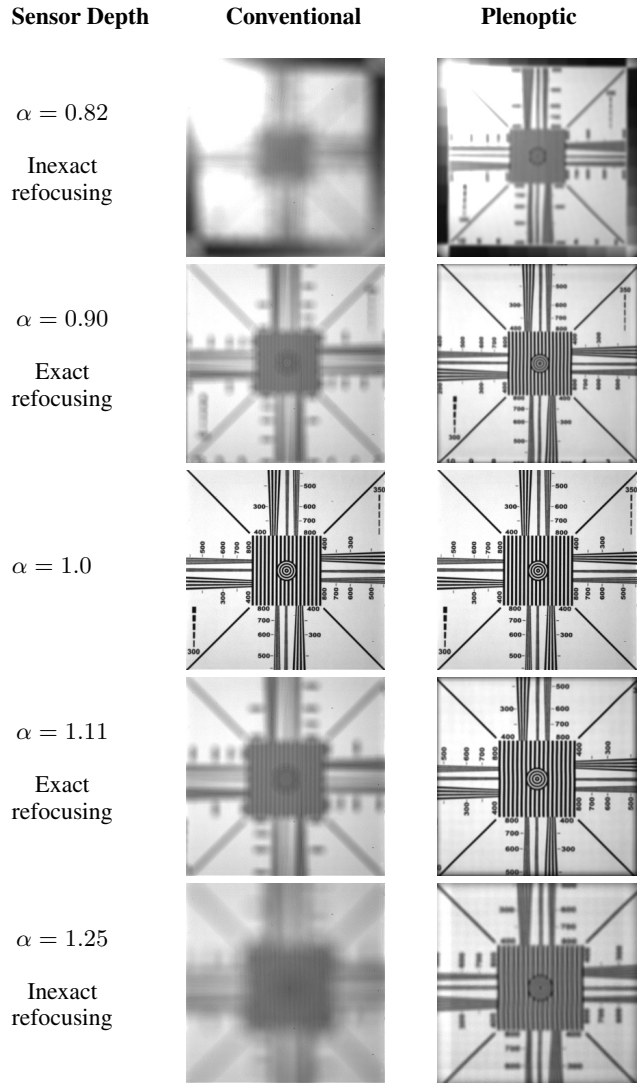

Figure 7: Photographs produced by conventional and plenoptic cameras [Anonymous 2005], using digital refocusing in the latter case. The sensor depth is given as a fraction of the focal depth that brings the target into focus. Note that minor digital refocusing provides the plenoptic camera with a wider effective depth of focus than the conventional system. Also note how the field of view changes slightly with the sensor depth, a change due to divergence of the light rays from the lens aperture.

cameras yields an analytic proof that limits on digital refocusing improve linearly with directional resolution. This phenomenon has been noted experimentally, and yields exciting applications in digital photography [Anonymous 2005]. The analysis in this section provides the first theoretical backing for the phenomenon.

5 Fourier Slice Digital Refocusing

This section applies the Fourier Slice Photography Theorem in a very different way, to derive an asymptotically fast algorithm for digital refocusing. The presumed usage scenario is as follows: an in-camera light field is available (perhaps having been captured by a plenoptic camera [Anonymous 2005]). The user wishes to digitally refocus in an interactive manner, i.e. select a desired focal plane and view a synthetic photograph focused on that plane (see bottom two rows of Figure 8).

In previous approaches to this problem [Isaksen et al. 2000; Levoy et al. 2004; Anonymous 2005], spatial integration via Eq. 3 results in an $O(n^4)$ algorithm, where n is the number of samples in each of the four dimensions. The algorithm described in this section provides a faster $O(n^2 \log n)$ algorithm, with the penalty of a single $O(n^4 \log n)$ pre-processing step.

5.1 Algorithm

The algorithm follows trivially from the Fourier Slice Photography Theorem:

Preprocess Prepare the given light field, L_F, by pre-computing its 4D Fourier transform, $\mathcal{F}^4[L]$, via the Fast Fourier Transform. This step takes $O(n^4 \log n)$ time.

Refocusing For each choice of desired world focus plane, W,

- Compute the conjugate virtual film plane depth, F', via the thin lens equation: $1/F' + 1/W = 1/f$, where f is the focal length of the lens.
- Extract the dilated Fourier slice (via Eq. 9) of the preprocessed Fourier transform, to obtain $(\mathfrak{P}_\alpha \circ \mathcal{F}^4)[L]$, where $\alpha = F'/F$. This step takes $O(n^2)$ time.
- Compute the inverse 2D Fourier transform of the slice, to obtain $(\mathcal{F}^{-2} \circ \mathfrak{P}_\alpha \circ \mathcal{F}^4)[L]$. By the theorem, this final result is $\mathcal{P}_\alpha[L_F] = E_{F'}$ the photo focused on world plane W. This step takes $O(n^2 \log n)$ time.

Figure 8 illustrates the steps of the algorithm.

5.2 Implementation and Results

The complexity in implementing this simple algorithm has to do with ameliorating the artifacts that result from discretization, resampling and Fourier transformation. These artifacts are conceptually the same as the artifacts tackled in Fourier volume rendering [Levoy 1992; Malzbender 1993], and Fourier-based medical reconstruction techniques [Jackson et al. 1991] such as those used in CT and MR. The interested reader should consult these citations and their bibliographies for further details.

5.2.1 Sources of Artifacts

In general signal-processing terms, when we sample a signal it is replicated periodically in the dual domain. When we reconstruct this sampled signal with convolution, it is multiplied in the dual domain by the Fourier transform of the convolution filter. The goal is to perfectly isolate the original, central replica, eliminating all other replicas. This means that the ideal filter is band-limited: it is of unit value for frequencies within the support of the light field, and zero for all other frequencies. Thus, the ideal filter is the sinc function, which has infinite extent.

In practice we must use an imperfect, finite-extent filter, which will exhibit two important defects (see Figure 9). First, the filter will not be of unit value within the band-limit, gradually decaying to smaller fractional values as the frequency increases. Second, the filter will not be truly band-limited, containing energy at frequencies outside the desired stop-band.

The first defect leads to so-called *rolloff artifacts* [Jackson et al. 1991], the most obvious manifestation of which is a darkening of the borders of computed photographs (see Figure 10). Decay in the filter's frequency spectrum with increasing frequency means that the spatial light field values, which are modulated by this spectrum, also "roll off" to fractional values towards the edges.

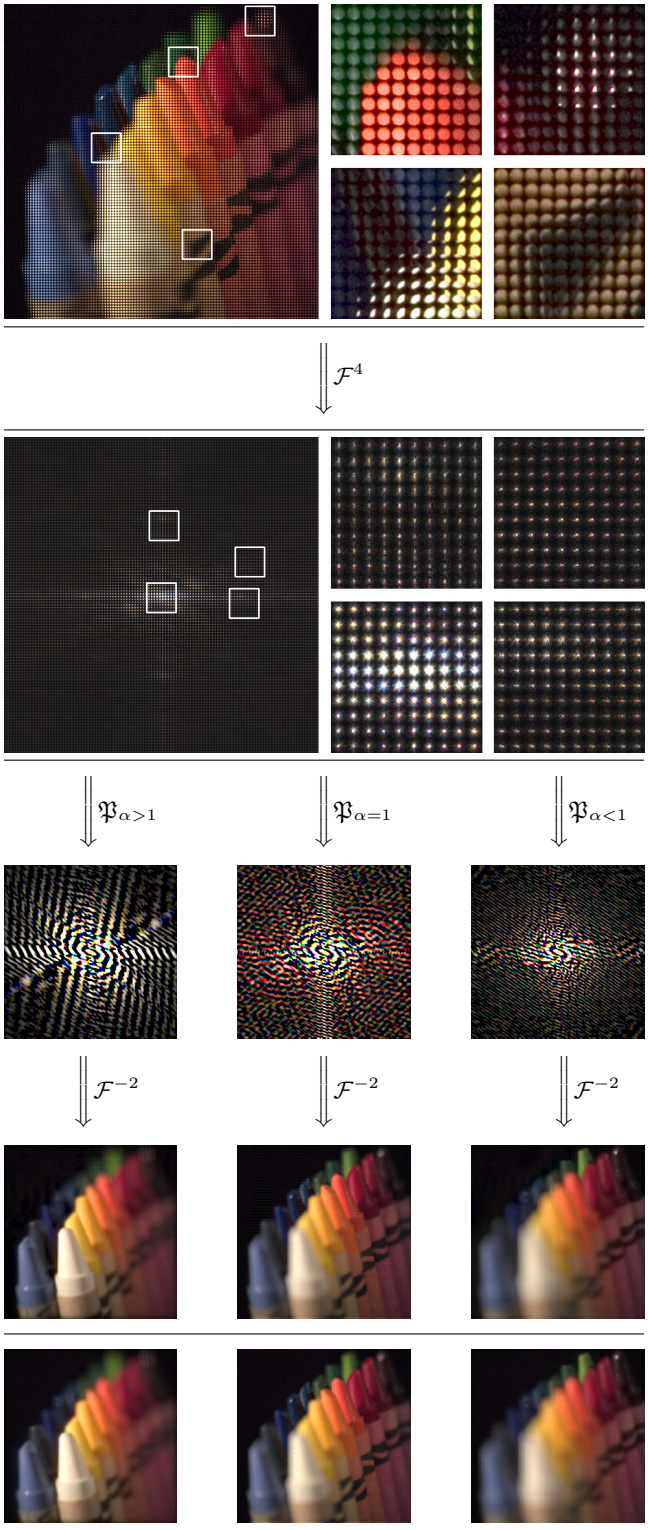

Figure 8: Fourier Slice Digital Refocusing algorithm. *Row 1:* 4D light field captured from a plenoptic camera, with four close-ups. *Row 2:* 4D Fourier transform of the light field, with four close-ups. *Row 3:* Three 2D slices of the 4D Fourier transform, extracted by $\mathfrak{P}_\alpha$ (Eq. 9) with different values of α. *Row 4:* Inverse 2D transforms of row 3. These images are conventional photographs focused on the closest, middle and furthest crayons. *Row 5:* Same images as row 4 except computed with brute-force integration in the spatial domain via Eq. 3, for comparison.

The second defect, energy at frequencies above the band-limit, leads to *aliasing artifacts* in computed photographs (see Figure 10). The non-zero energy beyond the band-limit means that the periodic replicas are not fully eliminated, leading to two kinds of aliasing. First, the replicas that appear parallel to the slicing plane appear as 2D replicas of the image encroaching on the borders of the final photograph. Second, the replicas positioned perpendicular to this plane are projected and summed onto the image plane, creating ghosting and loss of contrast.

5.2.2 Correcting Rolloff Error

Rolloff error is a well understood effect in medical imaging and Fourier volume rendering. The standard solution is to multiply the affected signal by the reciprocal of the filter's inverse Fourier spectrum, to nullify the effect introduced during resampling. In our case, directly analogously to Fourier volume rendering [Malzbender 1993], the solution is to spatially pre-multiply the input light field by the reciprocal of the filter's 4D inverse Fourier transform (see Figure 11). This is performed prior to taking its 4D Fourier transform in the pre-processing step of the algorithm.

Unfortunately, this pre-multiplication tends to accentuate the energy of the light field near its borders, maximizing the energy that folds back into the desired field of view as aliasing.

5.2.3 Suppressing Aliasing Artifacts

The three main methods of suppressing aliasing artifacts are oversampling, superior filtering and zero-padding. Oversampling within the extracted 2D Fourier slice $(\mathfrak{P}_F \circ \mathcal{F}^4)[L]$ increases the replication period in the spatial domain. This means that less energy in the tails of the in-plane replicas will fall within the borders of the final photograph.

Exactly what happens computationally will be familiar to those experienced in discrete Fourier transforms. Specifically, increasing the sampling rate in one domain leads to an increase in the field of view in the other domain. Hence, by oversampling we produce an image that shows us more of the world than desired (see Figure 12), not a magnified view of the desired portion. Aliasing energy from neighboring replicas falls into these outer regions, which we crop away to isolate the central image of interest.

Oversampling is appealing because of its simplicity, but oversampling alone cannot produce good quality images. The problem is that it cannot eliminate the replicas that appear perpendicular to the slicing plane, which are projected down onto the final image as described in the previous section.

This brings us to the second major technique of combating aliasing: superior filtering. As already stated, the ideal filter is a sinc function with a band-limit matching the spatial bounds of the light field. Our goal is to use a finite-extent filter that approximates this perfect spectrum as closely as possible. The best methods for producing such filters use iterative techniques to jointly optimize the band-limit and narrow spatial support, as described in Jackson et al. [1991] in the medical imaging community, and Malzbender [1993] in the Fourier volume rendering community.

Jackson et al. show that a much simpler, and near-optimal, approximation is the Kaiser-Bessel function. They also provide optimal Kaiser-Bessel parameter values for minimizing aliasing. Figure 13 illustrates the striking reduction in aliasing provided by such optimized Kaiser-Bessel filters compared to inferior quadrilinear interpolation. Surprisingly, a Kaiser-Bessel window of just width 2.5 suffices for excellent results.

The third and final method to combat aliasing is to pad the light field with a small border of zero values before pre-multiplication and taking its Fourier transform [Levoy 1992; Malzbender 1993]. This pushes energy slightly further from the borders, and minimizes

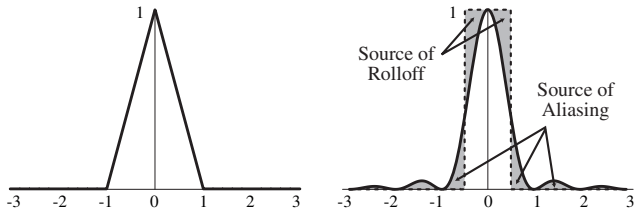

Figure 9: Source of artifacts. *Left*: Triangle reconstruction filter. *Right*: Frequency spectrum of this filter (solid line) compared to ideal spectrum (dotted line). Shaded regions show deviations that lead to artifacts.

Figure 10: Two main classes of artifacts. *Left*: Gold-standard image produced by spatial integration via Eq. 3. *Middle*: Rolloff artifacts using Kaiser-Bessel filter. *Right*: Aliasing artifacts using quadrilinear filter.

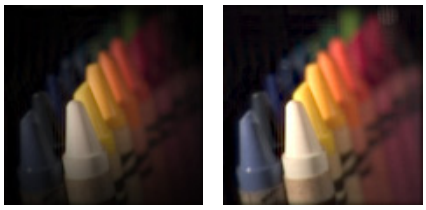

Figure 11: Rolloff correction by pre-multiplying the input light field by the reciprocal of the resmpling filter's inverse Fourier transform. *Left*: Kaiser-Bessel filtering without pre-multiplication. *Right*: With pre-multiplication.

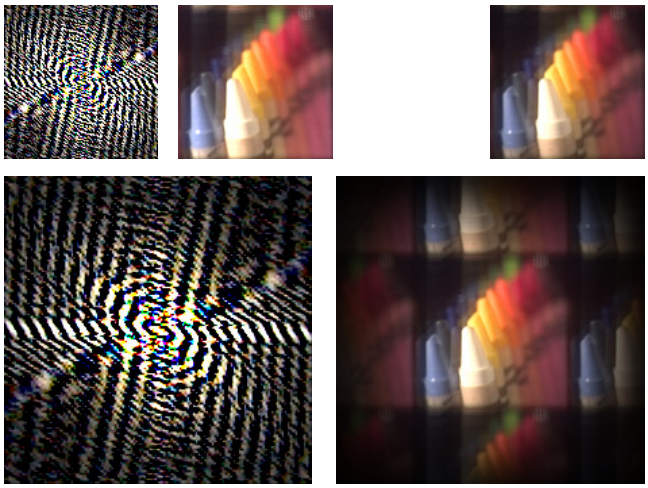

Figure 12: Aliasing reduction by oversampling. Quadrilinear filter used to emphasize aliasing effects for didactic purposes. *Top left pair*: 2D Fourier slice and inverse-transformed photograph, with unit sampling. *Bottom*: Same as top left pair, except with $2\times$ oversampling in the frequency domain. *Top right:* Cropped version of bottom right. Note that aliasing is reduced compared to version with unit sampling.

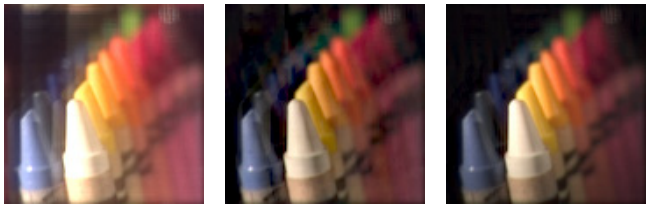

Figure 13: Aliasing reduction by superior filtering. Rolloff correction is applied. *Left*: Quadrilinear filter (width 2). *Middle*: Kaiser-Bessel filter, width 1.5. *Right*: Kaiser-Bessel filter, width 2.5.

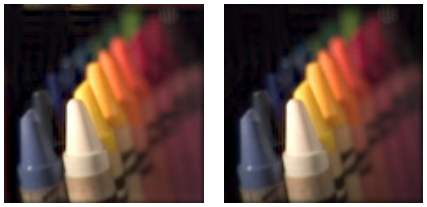

Figure 14: Aliasing reduction by padding with border of zero values. *Left*: Kaiser-Bessel filtering with no zero border. Note slight aliasing in top left. *Right*: With 5% zero border.

the amplification of aliasing energy by the pre-multiplication described in 5.2.2 (see Figure 14).

5.2.4 Implementation Summary

We directly discretize the algorithm presented in 5.1, applying the following four techniques (from 5.2.2 and 5.2.3) to suppress artifacts. In the pre-processing phase,

1. We pad the light field with a small border (5% of the width in that dimension) of zero values.
2. We pre-multiply the light field by the reciprocal of the Fourier transform of the resampling filter.

In the refocusing step where we extract the 2D Fourier slice,

3. We use a linearly-separable Kaiser-Bessel resampling filter. width 2.5 produces excellent results. For fast previewing, an extremely narrow filter of width 1.5 produces results that are superior to (and faster than) quadrilinear interpolation.
4. We oversample the 2D Fourier slice by a factor of 2. After Fourier inversion, we crop the resulting photograph to isolate the central quadrant.

The bottom two rows of Figure 8 compare this implementation of the Fourier Slice algorithm with spatial integration.

5.2.5 Performance Summary

We sampled the performance of our algorithm compared to spatial-domain methods. Tests were performed on a 3.0 Ghz Pentium IV processor. The FFTW-3 library [Frigo and Johnson 1998] was used to compute Fourier transforms efficiently.

For a light field with 256×256 st resolution and 16×16 uv resolution, spatial-domain integration achieved 1.68 fps (frames per second) using nearest-neighbor quadrature and 0.13 fps using quadrilinear interpolation. In contrast, the Fourier Slice method presented in this paper achieved 2.84 fps for previewing (Kaiser-Bessel filter, width 1.5 and no oversampling) and 0.54 fps for higher quality (width 2.5 and 2x oversampling). At these resolutions, performance

between spatial-domain and Fourier-domain methods are comparable.

The Fourier Slice method outperforms the spatial methods as the directional uv resolution increases, because the number of light field samples that must be summed increases for the spatial integration methods, but the cost of slicing stays constant per pixel. For a light field with 128×128 st resolution and 32×32 uv resolution, spatial-domain integration achieved 1.63 fps using nearest-neighbor, and 0.10 fps using quadrilinear interpolation. The Fourier Slice method achieved 15.58 fps for previewing and 2.73 fps for higher quality. At these resolutions, the Fourier slice methods are an order of magnitude faster.

6 Conclusions and Future Work

The main contribution of this paper is the Fourier Slice Photography Theorem. By describing how photographic imaging occurs in the Fourier domain, this simple but fundamental result provides a versatile tool for algorithm development and analysis of light field imaging systems. This paper tries to present concrete examples of this utility in analyzing a band-limited plenoptic camera, and in developing the Fourier Slice Digital Refocusing algorithm.

The band-limited analysis provides theoretical backing for recent experimental results demonstrating exciting performance of digital refocusing from plenoptic camera data [Anonymous 2005]. The derivations also yield a general-purpose tool (Eq. 11) for analyzing plenoptic camera systems.

The Fourier Slice Digital Refocusing algorithm is asymptotically faster than previous approaches. It is also efficient in a practical sense, thanks to the optimized Kaiser-Bessel resampling strategy borrowed directly from the tomography literature. Continuing to exploit this connection with tomography will surely yield further benefits in light field processing.

A clear line of future work would extend the algorithm to increase its focusing flexibility. Presented here in its simplest form, the algorithm is limited to full-aperture refocusing. Support for different apertures could be provided by appropriate convolution in the Fourier domain that results in a spatial multiplication masking out the undesired portion of the aperture. This is related to work in Fourier volume shading [Levoy 1992].

A very different class of future work might emerge from looking at the footprint of photographs in the 4D Fourier transform of the light field. It is a direct consequence of the Fourier Slice Photography Theorem (consider Eq. 10 for all α) that the footprint of all full-aperture photographs lies on the following 3D manifold in the 4D Fourier space:

$$\big\{(\alpha \cdot k_x,\ \alpha \cdot k_y, (1-\alpha) \cdot k_x,\ (1-\alpha) \cdot k_y) \\ \text{where } \alpha \in [0, \infty), \text{ and } k_x, k_y \in \mathbb{R}\big\} \quad (21)$$

Three possible lines of research are as follows.

First, attempting to extend this footprint to cover the entire 4D space leads to the intriguing idea of *Light Field Tomography*, where we try to reconstruct the light field inside the camera by taking sets of conventional photographs similarly to the way one reconstructs 3D volumes from x-ray projections in CT scanning. Preliminary theoretical work suggests that light field tomography might be possible if we use a slit aperture and vary the position of the slit in acquiring the sets of photographs.

Second, if our only goal is to compute refocused photographs, then we only require the information on the 3D manifold described above. Thus, it might be possible to optimize light field camera designs to provide greater fidelity on this manifold, at the expense of the vast remainder of the space that does not contribute to refocused photographs. One could also compress light fields for refocusing by storing only the data on the 3D manifold.

Finally, thinking about the footprint suggests a different style of depth analysis. The idea is that a photograph focused at a particular depth will contain sharp details (hence high frequencies) only if an object exists at that depth. This observation suggests a simple *Fourier Range Finding* algorithm: search for regions of high spectral energy on the 3D manifold at a large distance from the origin. The rotation angle of these high energy regions gives the depth (via the Fourier Slice Photography Theorem) of focal planes that intersect features in the visual world.

Acknowledgments

Thanks to Marc Levoy, Pat Hanrahan, Ravi Ramamoorthi, Mark Horowitz and Kayvon Fatahalian for discussions, help with references and reading a draft. Thanks also to Dwight Nishimura and Brad Osgood. This work was supported by NSF grant 0085864-2 (*Interacting with the Visual World*), and by a Microsoft Research Fellowship.

References

ADELSON, T., AND WANG, J. Y. A. 1992. Single lens stereo with a plenoptic camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence 14*, 2 (Feb), 99–106.

ANONYMOUS. 2005. Light field photography with a hand-held plenoptic camera. *Submitted to SIGGRAPH 05*.

BRACEWELL, R. N. 1956. Strip integration in radio astronomy. *Aust. J. Phys. 9*, 198–217.

BRACEWELL, R. N. 1986. *The Fourier Transform and Its Applications, 2nd Edition Revised.* WCB / McGraw-Hill.

CHAI, J., TONG, X., AND SHUM, H. 2000. Plenoptic sampling. In *SIGGRAPH 00*, 307–318.

DEANS, S. R. 1983. *The Radon Transform and Some of Its Applications*. Wiley-Interscience.

FRIGO, M., AND JOHNSON, S. G. 1998. Fftw: An adaptive software architecture for the fft. In *ICASSP conference proceedings*, vol. 3, 1381–1384.

GORTLER, S. J., GRZESZCZUK, R., SZELISKI, R., AND COHEN, M. F. 1996. The lumigraph. In *SIGGRAPH 96*, 43–54.

ISAKSEN, A., MCMILLAN, L., AND GORTLER, S. J. 2000. Dynamically reparameterized light fields. In *SIGGRAPH 2000*, 297–306.

JACKSON, J. I., MEYER, C. H., NISHIMURA, D. G., AND MACOVSKI, A. 1991. Selection of convolution function for fourier inversion using gridding. *IEEE Transactions on Medical Imaging 10*, 3, 473–478.

LACROUTE, P., AND LEVOY, M. 1994. Fast volume rendering using a shear-warp factorization of the viewing transformation. 451–458.

LEVOY, M., AND HANRAHAN, P. 1996. Light field rendering. In *SIGGRAPH 96*, 31–42.

LEVOY, M., CHEN, B., VAISH, V., HOROWITZ, M., MCDOWALL, I., AND BOLAS, M. 2004. Synthetic aperture confocal imaging. In *To appear in SIGGRAPH 2004*.

LEVOY, M. 1992. Volume rendering using the fourier projection-slice theorem. In *Proceedings of Graphics Interface '92*, 61–69.

MACOVSKI, A. 1983. *Medical Imaging Systems*. Prentice Hall.

MALZBENDER, T. 1993. Fourier volume rendering. *ACM Transactions on Graphics 12*, 3, 233–250.

MERSEREAU, R. M., AND OPPENHEIM, A. V. 1974. Digital reconstruction of multidimensional signals from their projections. *Proc IEEE 62*, 1319–1338.

STROEBEL, L., COMPTON, J., CURRENT, I., AND ZAKIA, R. 1986. *Photographic Materials and Processes*. Focal Press.

VAISH, V., WILBURN, B., JOSHI, N., AND LEVOY, M. 2004. Using plane + parallax for calibrating dense camera arrays. In *Proceedings of CVPR*, 2–9.

Appendices: Proofs and Derivations

A Generalized Fourier Slice Theorem

Theorem (Generalized Fourier Slice).

$$\mathcal{F}^M \circ \mathcal{I}_M^N \circ \mathcal{R} \equiv \mathcal{S}_M^N \circ \mathcal{R} \circ \mathcal{F}^N.$$

Proof. The following proof is inspired by one common approach to proving the classical 2D version of the theorem. The first step is to note that

$$\mathcal{F}^M \circ \mathcal{I}_M^N = \mathcal{S}_M^N \circ \mathcal{F}^N, \tag{22}$$

because substitution of the basic definitions shows that for an arbitrary function, f, both $\left(\mathcal{F}^M \circ \mathcal{I}_M^N\right)[f](u_1, \ldots, u_M)$ and $\left(\mathcal{S}_M^N \circ \mathcal{F}^N\right)[f](u_1, \ldots, u_M)$ are equal to

$$\int f(x_1, \ldots, x_N) \exp\left(-2\pi i \left(x_1 u_1 + \cdots + x_M u_M\right)\right) dx_1 \ldots dx_N.$$

The next step is to observe that if rotation operators commute with Fourier transforms, that is $\mathcal{R} \circ \mathcal{F}^N \equiv \mathcal{F}^N \circ \mathcal{R}$, then the proof of the theorem would be complete because for every function f we would have

$$\begin{aligned} \left(\mathcal{F}^M \circ \mathcal{I}_M^N \circ \mathcal{R}\right)[f] &= \left(\mathcal{F}^M \circ \mathcal{I}_M^N\right)[\mathcal{R}[f]] \qquad (23)\\ &= \left(\mathcal{S}_M^N \circ \mathcal{F}^N\right)[\mathcal{R}[f]] \quad \text{(by Eq. 22)}\\ &= \left(\mathcal{S}_M^N \circ \mathcal{R} \circ \mathcal{F}^N\right)[f]. \quad \text{(commute)} \end{aligned}$$

Thus, the final step is to show that $\mathcal{R} \circ \mathcal{F}^N \equiv \mathcal{F}^N \circ \mathcal{R}$. Directly substituting the operator definitions establishes these two equations:

$$\left(\mathcal{R} \circ \mathcal{F}^N\right)[f](\mathbf{u}) = \int f(\mathbf{x}) \exp\left(-2\pi i \left(\mathbf{x}^T \left(\mathcal{R}^{-1}\mathbf{u}\right)\right)\right) d\mathbf{x}, \tag{24}$$

$$\left(\mathcal{F}^N \circ \mathcal{R}\right)[f](\mathbf{u}) = \int f\left(\mathcal{R}^{-1}\mathbf{x}'\right) \exp\left(-2\pi i \left(\mathbf{x}' \cdot \mathbf{u}\right)\right) d\mathbf{x}'. \tag{25}$$

where $\mathbf{x}$ and $\mathbf{u}$ are N-dimensional column vectors, and the integral is taken over all of N-dimensional space.

Let us now apply the change of variables $\mathbf{x} = \mathcal{R}^{-1}\mathbf{x}'$ to Eq. 25, noting that $\mathbf{x} = \mathcal{R}\mathbf{x}'$, and $d\mathbf{x} = d\mathbf{x}'$ because the Jacobian of a rotation matrix is unity. Making these substitutions,

$$\begin{aligned} \left(\mathcal{F}^N \circ \mathcal{R}\right)[f](\mathbf{u}) &= \int f(\mathbf{x}) \exp\left(-2\pi i \left((\mathcal{R}\mathbf{x})^T \mathbf{u}\right)\right) d\mathbf{x}\\ &= \int f(\mathbf{x}) \exp\left(-2\pi i \left(\left(\mathbf{x}^T \mathcal{R}^T\right)\mathbf{u}\right)\right) d\mathbf{x}\\ &= \int f(\mathbf{x}) \exp\left(-2\pi i \left(\mathbf{x}^T \left(\mathcal{R}^{-1}\mathbf{u}\right)\right)\right) d\mathbf{x}, \qquad (26) \end{aligned}$$

where the last two lines rely on linear algebra rules for transposing matrix products and the fact that the transpose of a rotation matrix is its inverse. Equations 24 and 26 show that $\mathcal{R} \circ \mathcal{F}^N \equiv \mathcal{F}^N \circ \mathcal{R}$, completing the proof. □

B Filtered Light Field Photography Thm.

Theorem (Filtered Light Field Photography).

$$\mathcal{P}_F \circ \mathcal{C}_k^4 \equiv \mathcal{C}_{\mathcal{P}_F[k]}^2 \circ \mathcal{P}_F,$$

To prove the theorem, let us first establish a lemma involving the closely-related modulation operator:

Modulation $\mathcal{M}_\beta^N$ is an N-dimensional modulation operator, such that $\mathcal{M}_\beta^N[F](\mathbf{x}) = F(\mathbf{x}) \cdot \beta(\mathbf{x})$ where $\mathbf{x}$ is an N-dimensional vector coordinate.

Lemma. *Multiplying an input 4D function by another one, k, and transforming the result by $\mathfrak{P}_F$, the Fourier photography operator, is equivalent to transforming both functions by $\mathfrak{P}_F$ and then multiplying the resulting 2D functions. In operators,*

$$\mathfrak{P}_F \circ \mathcal{M}_k^4 \equiv \mathcal{M}_{\mathfrak{P}_F[k]}^2 \circ \mathfrak{P}_F \tag{27}$$

Algebraic verification of the lemma is direct given the basic definitions, and is omitted here. On an intuitive level, however, the lemma makes sense because $\mathfrak{P}_\alpha$ is a slicing operator: multiplying two functions and then slicing them is the same as slicing each of them and multiplying the resulting functions.

Proof of theorem. The first step is to translate the classical Fourier Convolution Theorem (see, for example, Bracewell [1986]) into useful operator identities. The Convolution Theorem states that a multiplication in the spatial domain is equivalent to convolution in the Fourier domain, and vice versa. As a result,

$$\mathcal{F}^N \circ \mathcal{C}_k^N \equiv \mathcal{M}_{\mathcal{F}^N[k]}^N \circ \mathcal{F}^N \tag{28}$$

$$\text{and} \quad \mathcal{F}^N \circ \mathcal{M}_k^N \equiv \mathcal{C}_{\mathcal{F}^N[k]}^N \circ \mathcal{F}^N. \tag{29}$$

Note that these equations also hold for negative N, since the Convolution Theorem also applies to the inverse Fourier transform.

With these facts and the lemma in hand, the proof of the theorem proceeds swiftly:

$$\begin{aligned} \mathcal{P}_F \circ \mathcal{C}_k^4 &\equiv \mathcal{F}^{-2} \circ \mathfrak{P}_F \circ \mathcal{F}^4 \circ \mathcal{C}_k^4\\ &\equiv \mathcal{F}^{-2} \circ \mathfrak{P}_F \circ \mathcal{M}_{\mathcal{F}^4[k]}^4 \circ \mathcal{F}^4\\ &\equiv \mathcal{F}^{-2} \circ \mathcal{M}_{(\mathfrak{P}_F \circ \mathcal{F}^4)[k]}^2 \circ \mathfrak{P}_F \circ \mathcal{F}^4\\ &\equiv \mathcal{C}_{(\mathcal{F}^{-2} \circ \mathfrak{P}_F \circ \mathcal{F}^4)[k]}^2 \circ \mathcal{F}^{-2} \circ \mathfrak{P}_F \circ \mathcal{F}^4\\ &\equiv \mathcal{C}_{\mathcal{P}_F[k]}^2 \circ \mathcal{P}_F, \end{aligned}$$

where we apply the Fourier Slice Photography Theorem (Eq. 10) to derive the first and last lines, the Convolution Theorem (Eqs. 28 and 29) for the second and fourth lines, and the lemma (Eq. 27) for the third line. □

C Photograph of a 4D Sinc Light Field

This appendix derives Eq. 15, which states that a photo from a 4D sinc light field is a 2D sinc function. The first step is to apply the Fourer Slice Photography Theorem to move the derivation into the Fourier domain.

$$\begin{aligned} &\mathcal{P}_\alpha\left[1/(\Delta x \Delta u)^2 \cdot \mathrm{sinc}(x/\Delta x, y/\Delta x, u/\Delta u, v/\Delta u)\right]\\ =&1/(\Delta x \Delta u)^2 \cdot \left(\mathcal{F}^{-2} \circ \mathfrak{P}_\alpha \circ \mathcal{F}^4\right)[\mathrm{sinc}(x/\Delta x, y/\Delta x, u/\Delta u, v/\Delta u)]\\ =&\left(\mathcal{F}^{-2} \circ \mathfrak{P}_\alpha\right)[\sqcap(k_x \Delta x, k_y \Delta x, k_u \Delta u, k_v \Delta u)]. \qquad (30) \end{aligned}$$

Now we apply the definition for the Fourier photography operator $\mathfrak{P}_\alpha$ (Eq. 8), to arrive at

$$\begin{aligned} &\mathcal{P}_\alpha\left[1/(\Delta x \Delta u)^2 \cdot \mathrm{sinc}(x/\Delta x, y/\Delta x, u/\Delta u, v/\Delta u)\right] \qquad (31)\\ &= \mathcal{F}^{-2}\left[\sqcap(\alpha k_x \Delta x,\ \alpha k_y \Delta x,\ (1-\alpha) k_x \Delta u,\ (1-\alpha) k_y \Delta u)\right]. \end{aligned}$$

Note that the 4D rect function now depends only on k_x and k_y, not k_u or k_v. Since the product of two dilated rect functions is equal to the smaller rect function,

$$\begin{aligned} &\mathcal{P}_\alpha\left[1/(\Delta x \Delta u)^2 \cdot \mathrm{sinc}(x/\Delta x, y/\Delta x, u/\Delta u, v/\Delta u)\right]\\ &\quad = \mathcal{F}^{-2}\left[\sqcap(k_x D_x, k_y D_x)\right] \qquad (32)\\ &\text{where } D_x = \max(\alpha \Delta x, |1-\alpha| \Delta u). \qquad (33) \end{aligned}$$

Applying the inverse 2D Fourier transform completes the proof:

$$\begin{aligned} &\mathcal{P}_\alpha\left[1/(\Delta x \Delta u)^2 \cdot \mathrm{sinc}(x/\Delta x, y/\Delta x, u/\Delta u, v/\Delta u)\right]\\ &\quad = 1/D_x^2 \cdot \mathrm{sinc}(k_x/D_x, k_y/D_x). \qquad (34) \end{aligned}$$