

A Theory of Plenoptic Multiplexing

Ivo Ihrke*

ivoihrke@cs.ubc.ca

The University of British Columbia

(*) now at Saarland University

Gordon Wetzstein

wetzstel@cs.ubc.ca

The University of British Columbia

Wolfgang Heidrich

heidrich@cs.ubc.ca

The University of British Columbia

Abstract

Multiplexing is a common technique for encoding high-dimensional image data into a single, two-dimensional image. Examples of spatial multiplexing include Bayer patterns to capture color channels, and integral images to encode light fields. In the Fourier domain, optical heterodyning has been used to acquire light fields.

In this paper, we develop a general theory of multiplexing the dimensions of the plenoptic function onto an image sensor. Our theory enables a principled comparison of plenoptic multiplexing schemes, including noise analysis, as well as the development of a generic reconstruction algorithm. The framework also aides in the identification and optimization of novel multiplexed imaging applications.

1. Introduction

Two-dimensional images are the primary means by which humans represent the three-dimensional world surrounding them. The introduction of photography resulted in unprecedented levels of realism in these representations. From the earliest works of color and integral imaging, an often-stated goal in photography has been to increase the amount of visual information that can be acquired.

In fact the “ultimate” camera would capture the full plenoptic function [2]. That is, within the limits of the uncertainty principle it would acquire the position, direction, wavelength, and time of arrival of each individual photon incident on the image sensor. From this information it would then be possible to extract different styles of images using computation. A practical challenge to achieving this goal is that real world image sensors always integrate over a finite area, incident angle, wavelength, and time.

The most common solution to this problem is to trade spatial resolution for an encoding of additional dimensions into the image plane. This *plenoptic multiplexing* can be achieved with something as simple as a color filter array like the Bayer pattern [7], or a more complicated encoding of multiple dimensions using Assorted Pixels [24]. An encoding of light fields can, for instance, be achieved by multiplexing directional light variation into different spatial frequency bands using optical heterodyning [32].

In this paper, we introduce a general theory for describing and analyzing plenoptic multiplexing systems. This

theory allows us to cast a large variety of existing multiplexed imaging approaches into a common framework for reconstruction, performance comparisons, and noise analysis. The framework also allows us to explore the space of multiplexed imaging systems in a principled way, and to derive new multiplexing schemes optimized for given tasks. Our specific contributions are:

- A mathematical framework for multiplexed imaging, which subsumes both spatial and Fourier reconstruction methods proposed in the literature.
- A spatial reconstruction method for Fourier multiplexed data, which can significantly increase the resulting image quality using our framework.
- A discussion of case studies using many existing multiplexing systems for color filter arrays and light fields.
- A comparative noise analysis of competing methods for color imaging and light field photography, which is enabled by our framework.

2. Background and Related Work

Image sensors integrate over a finite range along each dimension of the plenoptic function. In order to capture multiple samples along any of the non-spatial dimensions, three fundamental approaches are available.

Multi-Sensor Capture refers to an approach where multiple image sensors simultaneously capture different samples of the plenoptic function. Examples of this approach include multi-camera systems for capturing light fields and related information [35], as well as 3-chip cameras, which have a separate image sensor for the red, green, and blue channel of a color image. A similar approach is available for HDR imaging [3, 22].

Time Sequential Capture can be used when the hardware requirements of multi-sensor capture would be prohibitive. For example, light fields can also be captured by moving a single camera to different positions [18, 11]. Color images can be acquired sequentially by applying different color filters (e.g. [33]), which is particularly attractive when a large number of color channels is desired, for example in microscopy. HDR images can be acquired by sequentially capturing different exposures [9, 23], or using generalized image mosaicing [28].

A downside of time-sequential capture is of course the difficulty of capturing dynamic environments and videos. If neither multi-sensor, nor time sequential capture are feasible, one can employ the third approach:

Multiplexed Imaging refers to the encoding of multiple images into a single image/sensor plane. In effect, this approach trades spatial resolution for the ability to simultaneously capture multiple slices of the plenoptic function. Multiplexing can either be performed in the spatial domain or in the Fourier domain.

In spatial multiplexing, the pixels in a local neighborhood of the image sensor represent samples from different slices of the plenoptic function. Most often, the image is comprised of repeating tiles, which we call *super-pixels*. Corresponding pixels in each super-pixel come from the same slice of the plenoptic function. The most common example of spatial image multiplexing is the use of color filter arrays (CFAs) such as the Bayer pattern [7], in order to capture a color image in a single photograph. General filter arrays can be used to capture high dynamic range or multi-spectral images [24]. Integral and light field photography can be seen as multiplexing the light field onto a 2D image plane [20, 1, 25]. Here, the super-pixels correspond to different viewpoints, while corresponding points within each super-pixel correspond to the same viewing direction.

Fourier-space multiplexing is closely related to spatial multiplexing, but the spatial layout of corresponding modulation masks is optimized to achieve certain properties in the Fourier domain. In particular, spatial patterns are chosen such that the different slices of the plenoptic function are encoded into different frequency bands. In computer graphics, this *optical heterodyne* approach has so far been used for capturing light fields [32, 31], occluder information [14], and high dynamic range color photographs [34]. Spatially encoded light fields were recently analyzed in Fourier space and it was demonstrated that Fourier reconstruction algorithms apply as well [10]. In color imaging, a similar approach has been used to analyze filter patterns for Bayer-style color mosaics [6].

More exotic light field camera designs include the focused plenoptic camera [21], which captures the light field at a higher spatial, but lower angular resolution at a specific depth plane. Furthermore, light field reconstruction with priors can be formulated in a Bayesian framework [15, 17], which was demonstrated to achieve super-resolution [8].

Various other methods have been proposed that employ combined mask or lens-based optical light modulation and computational reconstruction. These approaches demonstrate how to capture images and scene depth [16] as well as light fields [13] or light fields and temporal light variation [4] using coded apertures, remove motion blur from a single photograph [26] and video sequences [5], or remove veiling glare from photographs [27, 30].

3. Plenoptic Multiplexing

Standard light sensing devices integrate over all dimensions of the plenoptic function [2]. These include spatial \vec{x} , directional $\vec{\theta}$ and temporal t variation as well as the color

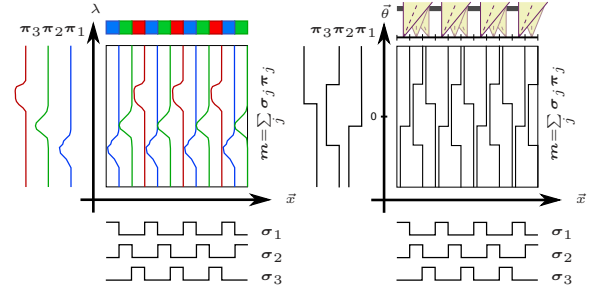


Figure 1. Illustration of the Cartesian product between spatial and plenoptic basis in case of a one-dimensional sensor with a spectral filter basis (left) and with an array of finite apertures (right).

spectrum λ incident on the sensor surface. The most common approach to acquire parts or all of this visual information with a single photograph is to perform an optical pre-modulation of the captured light. For this purpose, we introduce a general *plenoptic modulator* m that is capable of multiplexing different slices of the plenoptic function, such as color channels, onto the sensor image. A reconstruction can then be performed by interpolating each slice to every position; an example for this is color demosaicking. Reconstruction algorithms commonly assume that the acquired signal is spatially band-limited, which is in practice implemented using optical anti-aliasing filters [12].

A continuous, linearized, noise-free sensor image i can be modeled as

$$i(\vec{x}) = \int_{\mathcal{P}} m(\vec{x}, \vec{p}) l_{\lambda}(\vec{x}, \vec{p}) d\vec{p}, \quad (1)$$

where \vec{x} consists of the spatial dimensions on the image sensor, and \vec{p} contains the remaining parameters (direction, wavelength and time) of the plenoptic domain \mathcal{P} . The plenoptic function l_{λ} , returning spectral radiance, is defined *on the sensor plane*. It is usually a non-linearly distorted version of the plenoptic function in world space $l_{\lambda}(\vec{x}, \vec{p}) = l_{\lambda}^{world} \circ g^{-1}(\vec{x}, \vec{p})$, where the distortion g^{-1} is introduced by the optics of the camera's refractive elements. In this paper, we focus on multiplexing and demultiplexing the plenoptic function on the sensor. This approach separates the geometric optics changes in ray direction from the modulation function, which represents terms such as a finite aperture and, for real-world lens systems, any photometric terms. This separation is common practice in the optics literature. Note, however, that we also show in a supplemental document how our multiplexing theory can be directly applied to certain classes of refractive optical elements.

Our theory is based on the separation of the plenoptic modulator into a sum of mutually independent spatial and plenoptic basis functions:

$$m(\vec{x}, \vec{p}) = \sum_j \sigma_j(\vec{x}) \times \pi_j(\vec{p}). \quad (2)$$

A visualization of this separation is shown in Figure 1. In Section 4 we show how many modulators that correspond to state-of-the-art acquisition systems can be decomposed in this way.

Real-world imaging systems can only acquire a discrete subset of the plenoptic slices. Therefore, we introduce a set of $j = 1 \dots N$ plenoptic coefficients $\rho_j(\vec{x}) = \int \pi_j(\vec{p}) l_\lambda(\vec{x}, \vec{p}) d\vec{p}$, which describe a projection of the plenoptic function into the *spatially constant* plenoptic basis. Combining Equations 1 and 2 yields

$$\begin{aligned} i(\vec{x}) &= \sum_{j=1}^N \sigma_j(\vec{x}) \int_{\mathcal{P}} \pi_j(\vec{p}) l_\lambda(\vec{x}, \vec{p}) d\vec{p} \\ &= \sum_{j=1}^N \sigma_j(\vec{x}) \rho_j(\vec{x}). \end{aligned} \quad (3)$$

The spatial basis $\sigma(\vec{x})$ is required to *multiplex* the different plenoptic coefficients into a 2D sensor image. Specifically, $\sigma_j(\vec{x})$ describes the sampling pattern for a single plenoptic basis function $\pi_j(\vec{p})$. In the case of imaging through a color filter array, the spatial basis defines the layout of the individual color filters. We assume the spatial basis functions to be super-pixel periodic so that the visual information is uniformly sampled over the sensor.

In summary, a captured sensor image $i(\vec{x})$ contains measurements of all plenoptic coefficients $\rho_j(\vec{x})$. Their sampling layout is determined by the spatial basis $\sigma(\vec{x})$. Note that the spatial basis functions σ_j can have overlapping support. Therefore, a measured image value may represent linear combinations of all plenoptic coefficients at that location. Each pixel within a super-pixel samples a *different* linear combination of plenoptic coefficients, while corresponding pixels located in different super-pixels sample the *same* linear combination. We introduce the notation $c_k = \sum_j \sigma_j^k \rho_j(\vec{x})$ for each of these different linear combinations of plenoptic coefficients, where σ_j^k is the k^{th} value of the discretized periodic spatial basis function $\sigma_j(\vec{x})$. A sensor image directly samples each of these *measurement channels* at a different location within a super-pixel.

3.1. Spatial Reconstruction

In order to perform a spatial reconstruction, all measurement channels c_k need to be estimated at each spatial location \vec{x} (see Fig. 2, upper center). Initially, only one sample of each measurement channel per super-pixel is recorded by the sensor. An intuitive example is that of color demosaicking, where different color channels of a raw sensor image need to be interpolated to each position of the resulting color image. Under the assumption of an underlying spatially band-limited signal, this is a standard reconstruction (or interpolation) problem. We express this concept as

Theorem 1 (*Plenoptic Spatial Multiplexing, PSM*). *The interpolated measurement channels at a particular location \vec{x} are locally related to the plenoptic function as*

$$\vec{c}(\vec{x}) = \Sigma \vec{\rho}(\vec{x}) = \Sigma \Pi^T \vec{l}_\lambda(\vec{x}). \quad (4)$$

Here, the matrix Σ encodes the spatial basis functions σ_j in its columns and the matrix Π^T encodes the plenoptic bases

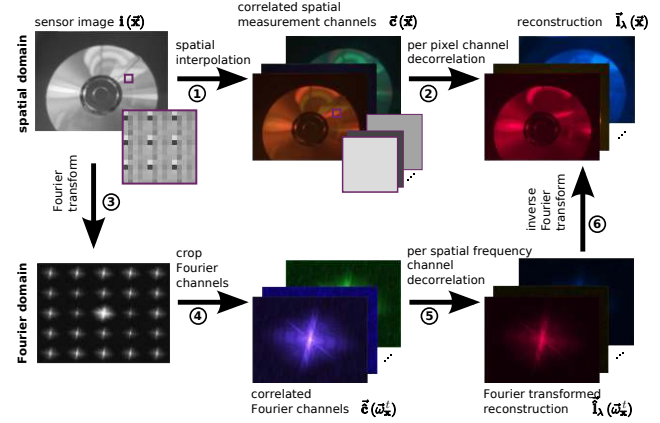


Figure 2. Overview of multiplexed image reconstruction. The plenoptic function can be reconstructed by interpolating all measurement channels and performing a local decorrelation in the spatial domain (upper row). Alternatively, it can be reconstructed in the Fourier domain by cropping and locally decorrelating the Fourier channels created by the plenoptic modulator (lower row).

$\pi_j(\vec{p})$ in its rows. Vector $\vec{c}(\vec{x})$ contains the N measurement channels interpolated to a common position, $\vec{\rho}(\vec{x})$ the corresponding plenoptic coefficients and $\vec{l}_\lambda(\vec{x})$ the plenoptic function. In many multiplexing applications, Equation 4 is a square, well-conditioned system, and can be solved directly. However, in some cases matrices may be non-square or singular, for example because the same plenoptic coefficients are encoded using different spatial basis functions (e.g. the Bayer pattern encodes the green channel at two different locations). In this case, a least squared error solution to Equation 4 is given using the matrices' pseudo-inverses: $\vec{l}_\lambda(\vec{x}) = (\Pi^T \Sigma + \Sigma^+)^{-1} \vec{c}(\vec{x})$. This reconstruction corresponds to the upper row of Figure 2.

The PSM theorem shows that we can reconstruct the plenoptic function $\vec{l}_\lambda(\vec{x})$ from sensor samples $i(\vec{x})$ by performing a *local* decorrelation on the interpolated measurement channels $c_k(\vec{x})$. However, Theorem 1 not only shows that the correlation between the measured channels is spatially local, but also that the correlation is in fact a linear operator, yielding

Corollary 1 *Any linear filter can be applied to the measurement channels $c_k(\vec{x})$ prior to decorrelation while yielding equivalent results to application after the decorrelation.*

Image processing operations such as upsampling, edge detection, blurring, sharpening, etc. can thus be performed on the correlated image without affecting the end result. Although this only applies to linear filters in theory, we show in supplemental material that non-linear filters can achieve high-quality reconstruction results in practice. Non-linear filters are already the preferred choice for color demosaicking; we show that these can also be applied to light field reconstruction.

3.2. Fourier Reconstruction

In recent literature, multiplexing strategies have often been discussed in the Fourier domain [32, 31, 14, 10]. For this reason we provide the dual Fourier view of plenoptic multiplexing and reconstruction in the following.

By applying the convolution theorem, the Fourier transform of an acquired image (Eq. 3) is given as

$$\mathcal{F}_x \{i\} = \mathcal{F}_x \left\{ \sum_{j=1}^N \sigma_j(\vec{x}) \rho_j(\vec{x}) \right\} = \sum_{j=1}^N \hat{\sigma}_j(\vec{\omega}_x) \otimes \hat{\rho}_j(\vec{\omega}_x), \quad (5)$$

where $\hat{\cdot}$ denotes the Fourier transformed version of a quantity, and $\vec{\omega}_x$ are the spatial frequencies.

The Poisson summation formula dictates that the Fourier transform of a periodic function is a weighted set of Dirac peaks. Thus, the Fourier transform of the spatial basis functions is given by

$$\hat{\sigma}_j(\vec{\omega}_x) = \sum_{k=1}^N \hat{\sigma}_j^k \delta(\vec{\omega}_x - k\Delta\vec{\omega}_x), \quad (6)$$

where $\Delta\vec{\omega}_x$ is the frequency offset or distance between successive Dirac peaks¹ and the values $\hat{\sigma}_j^k$ are the Fourier weights for basis function j . These weights correspond to the Fourier transform of a single period of that specific basis function. Combining Equations 5 and 6 as

$$\mathcal{F}_x \{i\} = \sum_{k=1}^N \left(\delta(\vec{\omega}_x - k\Delta\vec{\omega}_x) \otimes \sum_{j=1}^N \hat{\sigma}_j^k \hat{\rho}_j(\vec{\omega}_x) \right) \quad (7)$$

shows that N different linear combinations of Fourier transformed plenoptic coefficients $\hat{\rho}_j$, which we refer to as *Fourier channels* $\hat{c}_k(\vec{\omega}_x) = \sum_j \hat{\sigma}_j^k \hat{\rho}_j(\vec{\omega}_x)$, are created in the frequency domain. Given the same band-limitation requirements on the original signal as in the spatial case, the Fourier channels do not overlap and a single linear combination of Fourier plenoptic coefficients $\hat{\rho}_j$ is measured in each of the channels (see Fig. 2, bottom left). In the following we consider the Fourier channels to be cropped from the multiplexed sensor image (Fig. 2, bottom center). To account for the reduced resolution of a channel compared to the full multiplexed image we introduce a new set of spatial frequencies $\vec{\omega}'_x$. Similar to the spatial case, the correlation of the Fourier channels is *local in the spatial frequencies* $\vec{\omega}'_x$. In analogy to the PSM theorem, Equation 4, we can therefore state the following

Theorem 2 (*Plenoptic Fourier Multiplexing, PFM*). *The Fourier channels at a particular spatial frequency $\vec{\omega}'_x$ are locally related to the Fourier transform of the plenoptic function as*

$$\vec{c}(\vec{\omega}'_x) = \hat{\Sigma} \vec{\rho}(\vec{\omega}'_x) = \hat{\Sigma} \Pi^T \vec{l}_\lambda(\vec{\omega}'_x). \quad (8)$$

¹For a 2D sensor image the bases are periodic in both spatial dimensions, but we will omit the second one in our notation of k for clarity.

The correlation matrix $\hat{\Sigma}_{jk} = \hat{\sigma}_j^k$ is determined by the Fourier weights of the spatial basis which can be written as $\hat{\Sigma} = F\Sigma$ in matrix form. The plenoptic basis functions encoded in Π^T remain unchanged because of their independence of \vec{x} (see Equation 2). A least-squared decorrelation in the Fourier domain can again be performed using the matrices' pseudo-inverses as $\vec{l}_\lambda(\vec{\omega}'_x) = (\Pi^T)^+ \hat{\Sigma}^+ \vec{c}(\vec{\omega}'_x)$, which is illustrated in the lower row of Figure 2.

In summary, the PSM theorem leads to a generic reconstruction algorithm in the spatial domain by interpolating all spatially correlated image channels and decorrelating them locally. The PFM theorem, on the other hand, shows that previously proposed Fourier reconstruction techniques [32, 31, 14, 10] are a special case of this method. The cropping operation in Fourier space is a multiplication with a rect function, which is equivalent to a spatial sinc filter. Therefore, all previously proposed Fourier reconstruction methods use a fixed spatial reconstruction filter, which is a sinc. We will see in the next section that this choice negatively affects the quality of the demultiplexing operation.

Another consequence of our analysis is that both spatial and Fourier multiplexing schemes have nominally the same band-limitation requirements. However, spatial reconstruction can be made more resilient to residual high frequencies than Fourier reconstructions, where the high frequencies get directly transferred into other channels.

4. Case Studies

In the following we analyze two important methods of plenoptic multiplexing and show that they are closely related. Imaging with color filter arrays is well-understood and serves as an intuitive example for our general framework. This is then extended to light field multiplexing, which has recently gained a lot of interest in the vision and graphics community. We propose a novel reconstruction approach for Fourier multiplexed light fields in the spatial domain and show that it yields higher quality results than previous approaches.

4.1. Imaging through Color Filter Arrays

Color filter arrays (CFAs) are spectral filters placed in front of individual sensor elements in a super-pixel like layout as considered in Section 3.1. The most widely used CFA is the Bayer pattern [7]. It consists of a 2×2 pixel-periodic pattern containing two green color filters, as well as a red and a blue one.

Expressed in our multiplexing framework, the plenoptic dimensions reduce to the color spectrum $\vec{p} = (\lambda)$. Disregarding the wavelength dependency of the sensor, the plenoptic basis functions $\pi = \{\pi_r, \pi_g, \pi_b\}$ are the spectral responses of the individual color filters. These are usually chosen so that a transformation into standardized color spaces such as sRGB or XYZ are easily possible. In this application it is not necessary to further process the plenoptic coefficients $\rho_{\{r,g,b\}}(\vec{x})$ in order to extract estimates of the spectral distribution $l_\lambda(\vec{x}, \lambda)$ of the light, because they already represent the desired plenoptic quantities.

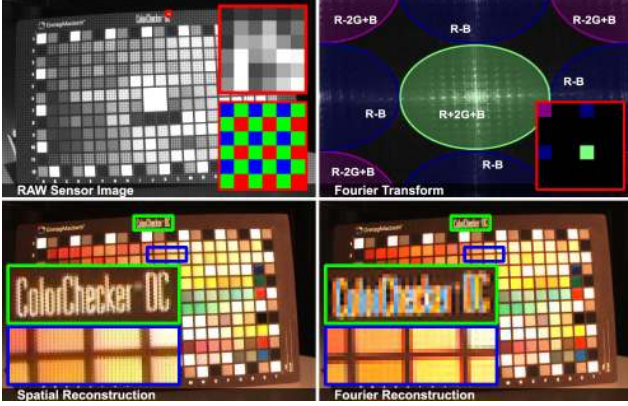


Figure 3. Raw sensor image with magnified part and corresponding CFA (upper left). Fourier transform with channel correlations illustrated for the entire image and the magnified CFA (upper right). Reconstruction of the non-perfectly bandlimited signal in the spatial (lower left) and Fourier (lower right) domain reveal different aliasing artifacts.

The spatial basis functions $\sigma_j(\vec{x})$ are in this case differently shifted Dirac peaks. Their specific position depends on the actual CFA, but the most common layout is shown in Figure 3 (upper left). Although the color channels are uncorrelated in the spatial domain, each of the Fourier channels contains differently weighted contributions from all color channels. The Fourier weights $\hat{\sigma}_j^k$ for the Bayer pattern are illustrated in Figure 3 (upper right), with their sum being color coded in the magnification.

A reconstruction of the original signal can thus be performed in either the spatial or the Fourier domain, which is known for the case of CFA imaging [19]. Many high quality image color priors exist that can be integrated into a spatial reconstruction and yield superior quality. Figure 3 shows the result of spatial reconstruction (lower left) compared to a cropping of the different copies followed by a channel decorrelation in the Fourier domain (lower right). This corresponds to applying least-squares reconstructions using the PSM (Eq. 4) and the PFM (Eq. 8) theorems, respectively. As the original signal is not properly band-limited, the reconstruction shows different aliasing artifacts. A spatial reconstruction produces much better results in this case.

4.2. Light Field Acquisition

We now show that many light field acquisition approaches that aim at capturing directional light variation can be seamlessly integrated into our framework. For simplicity of notation and without loss of generality, we restrict the discussion in this subsection to the 1D case of imaging a 2D light field. We reparameterize the light field by replacing the angular dimension θ by a spatial position $v = \tan(\theta)$ on a plane at unit distance (see Fig. 4, left). With these notational simplifications, our dimensional vectors are $\vec{x} = (x)$ and $\vec{v} = (v)$.

4.2.1 General Non-Refractive Modulators

Attenuation masks that do not include refractive optical elements have recently been popularized for light field acquisition [32, 31, 14]. All of these approaches are associated with a reconstruction in the Fourier domain (see Section 3.2). Here, we show that the employed periodic attenuation masks can be separated into a spatial and a plenoptic basis. This allows the aforementioned acquisition techniques to be expressed in our unifying framework, which enables a spatial reconstruction with a superior quality.

As illustrated in Figure 4 (left), the plenoptic modulator for an attenuation mask at a distance z to a sensor is given by reparameterizing its spatial dimension and including the attenuation pattern as $\mathbf{m}(x, v) = \mathbf{m}(x - zv)$. This models light transport in free space in front of the sensor and serves as a physical means for achieving linearly combined modulators on the sensor plane. By substituting $\mathbf{m}(x, v)$ in Equation 1 with the inverse of its Fourier transform, we can separate it into a spatial and a plenoptic part:

$$\begin{aligned} i(x) &= \int_v \mathbf{l}_\lambda(x, v) \mathbf{m}(x - zv) dv \\ &= \int_v \mathbf{l}_\lambda(x, v) \int_{\omega_x} \hat{\mathbf{m}}(\omega_x) e^{2\pi i(x-zv)\omega_x} d\omega_x dv \quad (9) \\ &= \int_{\omega_x} \hat{\mathbf{m}}(\omega_x) e^{2\pi i x \omega_x} \int_v \mathbf{l}_\lambda(x, v) e^{-2\pi i z v \omega_x} dv d\omega_x. \end{aligned}$$

Discretizing the integrals as $\mathbf{i} = \mathbf{F}^{-1} \hat{\mathbf{M}} \mathbf{F} \mathbf{l}_\lambda$, with $\hat{\mathbf{M}} = \text{diag}(\hat{\mathbf{m}}_j)$, allows us to choose differently weighted inverse Fourier basis functions as the spatial basis $\mathbf{\Sigma} = \mathbf{F}^{-1} \hat{\mathbf{M}}$, and the Fourier basis as the plenoptic basis $\mathbf{\Pi}^T = \mathbf{F}$. This is also illustrated in Figure 4 (center right column).

A spatial reconstruction can therefore be performed by applying a least-squares solution to the PSM theorem (Eq. 4). The spatial reconstruction is performed by upsampling the captured image as outlined in Section 3.1, which allows filters other than a sinc to be applied. The upsampled measurement channels $\vec{c}(\vec{x})$ are then decorrelated as

$$\vec{l}_\lambda(\vec{x}) = (\mathbf{\Pi}^T)^+ \mathbf{\Sigma}^+ \vec{c}(\vec{x}) = \mathbf{F}^{-1} \hat{\mathbf{M}}^{-1} \mathbf{F} \vec{c}(\vec{x}). \quad (10)$$

This is a deconvolution along the directional dimension, which is independently performed for each spatial position. This result shows that deconvolution for sum-of-sinusoids patterns, as employed for instance by Veeraraghavan et al. [32], represents a high-pass filter. Unfortunately, this decreases the signal-to-noise ratio (SNR) of the reconstructed light field significantly as we will show in Section 5.

A Fourier-based reconstruction, as previously proposed, can be performed by directly applying the least-squares solution to the PFM theorem (Eq. 8):

$$\vec{l}_\lambda(\vec{\omega}'_x) = (\mathbf{\Pi}^T)^+ \hat{\mathbf{\Sigma}}^+ \vec{c}'(\vec{\omega}'_x) = \mathbf{F}^{-1} \hat{\mathbf{M}}^{-1} \vec{c}'(\vec{\omega}'_x). \quad (11)$$

Equation 11 shows that the Fourier channels $\vec{c}'(\vec{\omega}'_x)$ need to be re-weighted according to the applied mask and inverse

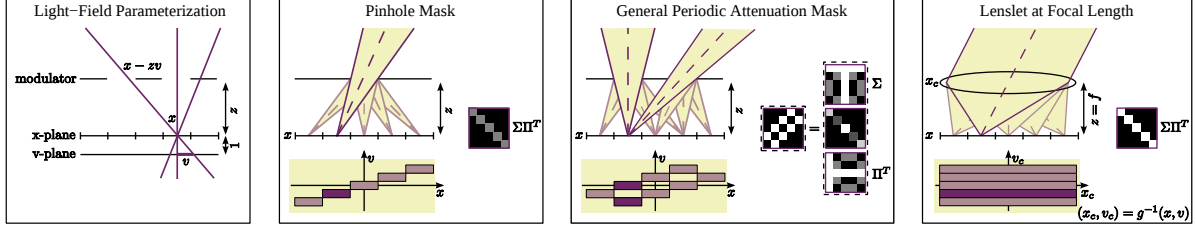


Figure 4. A light field can be parameterized with an x-plane and a v-plane at unit distance (left). Columns 2 to 4 illustrate different plenoptic camera configurations (upper left), corresponding integration surfaces in light field space for the individual sensor pixels (lower left), and the spatial and plenoptic basis functions (right). The convolution of a periodic attenuation mask can be separated into a spatial and plenoptic part using the Fourier basis (column 3, right). Integration surfaces for refractive optical elements already include the mapping g^{-1} from sensor space to camera or world space on the microlens plane.

Fourier transformed in order to invert the plenoptic basis. Another inverse Fourier transform needs to be performed to get the desired plenoptic samples $\vec{l}_\lambda(\vec{x})$ from $\vec{l}_\lambda(\vec{\omega}'_x)$, which is equivalent to the previously proposed inverse 4D Fourier transform of cropped and stacked Fourier tiles in 2D light fields [32, 31, 14].

arrays can further increase the fidelity of the reconstructed imagery as demonstrated in the supplemental material.

4.2.2 Lens-Based and Pinhole Modulators

The combined spatial and plenoptic basis functions for a pinhole attenuation mask and a lens-based plenoptic camera design are shown in Figure 4. The pinhole case and the standard plenoptic camera case, where the lenslets are focused at infinity are trivial, because uncorrelated directional samples of the plenoptic function are directly measured by each sensor element. Therefore, the combined discrete spatial and plenoptic basis is the identity matrix. The two cases differ in light transmission and depth-of-field effects due to large aperture of the microlenses. A similar angular resolution for both cases can be achieved with pinholes that have the size of a sensor pixel.

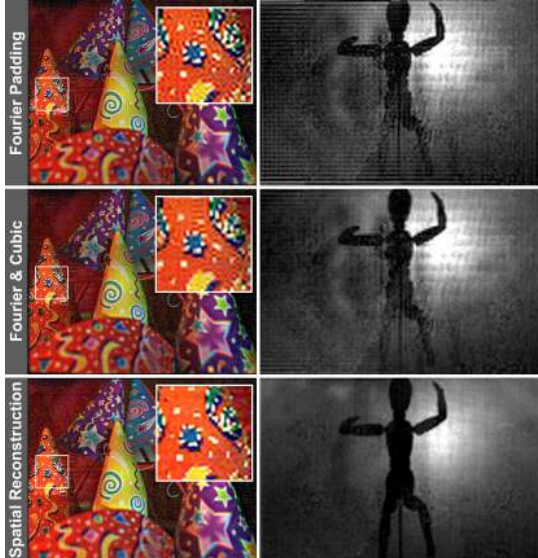


Figure 5. Comparison of reconstruction quality for Cones data set (Veeraraghavan et al. [2007]) and Mannequin data set (Lanman et al. [2008]). All results are three-times upsampled during reconstruction. Top row: upsampling by zero-padding the 4D inverse FFT. Middle row: low resolution 4D inverse FFT followed by bicubic upsampling. Bottom row: bicubic up-sampling of correlated measurement channels followed by local decorrelation.

We have shown in Section 3.2 that the 4D inverse Fourier transform uses a spatial sinc reconstruction filter. A spatial reconstruction based on our theory, on the other hand, can benefit from more sophisticated interpolation schemes. A comparison of reconstructions obtained with the inverse Fourier transform and with our local reconstruction algorithm is shown in Figure 5. Even with a simple spatial interpolation scheme such as cubic interpolation, common ringing artifacts associated with Fourier-based techniques can be avoided. Advanced interpolation methods as commonly applied in demosaicking techniques for color filter

5. Reconstruction Noise Analysis

Our theory, specifically the PSM theorem (Eq. 4), allows us to perform a reconstruction noise analysis and a comparison of alternative plenoptic multiplexing schemes. To the best knowledge of the authors, this is the first time that the reconstruction noise of Fourier multiplexed light fields can be analyzed and the performance of Fourier multiplexing schemes can be directly compared to their spatial alternatives.

We employ a noise model that is commonly applied in computer vision applications [36, 29]. The total noise variance ς^2 of a camera image is modeled as the combination of a signal independent additive term ς_c^2 that includes dark current and amplifier noise as well as a signal dependent photon shot noise term $i(\vec{x})\varsigma_p^2$. The image intensity can be approximated by the mean light transmission τ of the plenoptic modulator, yielding the following noise variance in the captured image:

$$\varsigma^2 = \varsigma_c^2 + \tau\varsigma_p^2. \quad (12)$$

In order to compare alternative plenoptic modulators, we need to propagate the sensor noise ς of a specific setup to the demultiplexed reconstruction. For this purpose, we define a noise amplification term $\alpha = \sqrt{\frac{1}{N}\text{trace}(\mathbf{\Pi}\mathbf{\Sigma}^T\mathbf{\Sigma}\mathbf{\Pi}^T)}$. The signal-to-noise ratio in

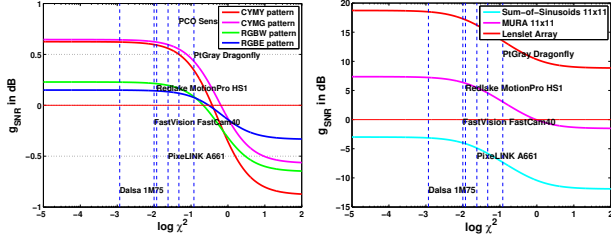


Figure 6. A SNR comparison of various alternative CFA patterns to the Bayer pattern (left). SNR comparison of different light field capture setups (right). The vertical lines indicate χ^2 values for several machine vision cameras tested by Schechner et al. [2007]. All cameras are operating in the gain region of the filters, i.e. $g_{\text{SNR}} > 0$. Note, however, that the lines can be moved left and right along the χ^2 -axis by increasing and decreasing the gain setting of a camera, respectively.

the demultiplexed plenoptic function requires expressions for the signal and for the noise term. Assuming that $\int \mathcal{L}_\lambda(\vec{x}, \vec{p}) d\vec{p} = 1$ and that all plenoptic bases are orthogonal, the signal term in the demultiplexed signal can be approximated by $1/N$. The reconstruction noise term is the standard deviation of the propagated sensor noise $\alpha\zeta$, resulting in an SNR of

$$\text{SNR} = 10 \log_{10} \left(\frac{1}{N \alpha \sqrt{\zeta_c^2 + \tau \zeta_p^2}} \right). \quad (13)$$

We define the SNR in dB. The gain of SNR for this demultiplexed signal compared to a reference signal is then

$$g_{\text{SNR}} = \text{SNR} - \text{SNR}_{\text{ref}} = 10 \log_{10} \left(\frac{\alpha_{\text{ref}} \zeta_{\text{ref}}}{\alpha \zeta} \right). \quad (14)$$

A plot of SNR gain for a standard Bayer CFA filter compared to some of the alternatives is shown in Fig. 6 (left). The figure shows that all alternative filters produce a slightly better SNR than the Bayer pattern when the additive noise term dominates (left part of the plot). However, performance drops below that of the Bayer pattern once photon shot noise becomes dominant. The CYMG design performs best with a gain of ≈ 0.6 dB. We employ the notation introduced in [36], where a parameter $\chi = \zeta_p/\zeta_c$ describes the ratio of signal-dependent and signal-independent noise variance. This makes it more convenient for plotting the performance of a multiplexing scheme for different camera noise parameters up to a global scale as seen in Figure 6. Based on the plot we can see that, from the perspective of SNR gain, alternative CFA patterns are not an effective tool to obtain better noise performance for color cameras with modern sensors, where the noise is dominated by photon shot noise. Note that decreased camera gain settings, i.e. operation in well lit scenes, move the camera curves indicated in Fig. 6 towards the right.

Similarly, we compare the noise performance of various light field acquisition approaches, where the pinhole attenuation mask serves as the reference. Sensor quantization and other non-linearities are disregarded. For this plot we

used attenuation masks that have a resolution of 11×11 for each super-pixel and a similar sensor resolution. The plot in Figure 6 (right) shows that lenslets always perform best in terms of SNR. Among the non-refractive multiplexing methods, MURA-based attenuation masks perform very well for a dominating additive noise term, i.e. at high camera gain settings in low-light conditions. Their SNR gain drops below that of a pinhole for an increasingly dominating photon noise term. Sum-of-sinusoids masks, as employed by Dappled Photography [32], always perform worse than a pinhole. The same result can be inferred from Figure 7, where we simulated the acquisition of a light field with different plenoptic cameras. The exposure times are equivalent and the chosen sensor noise characteristics correspond to those of the PG Dragonfly in Figure 6 (right), i.e. $\log(\chi^2) = -0.93$. As expected, the lenslet-based design has the best noise performance, with the MURA attenuation pattern being very close in this case and the sum-of-sinusoids mask not performing very well. The dataset shows golgi stained neurons captured with a light field microscope (lightfield.stanford.edu).

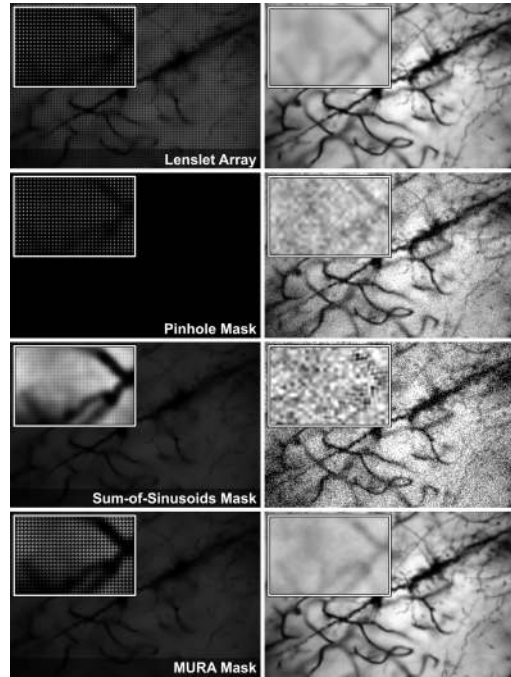


Figure 7. Light field acquisition with different plenoptic cameras. The left column shows simulated sensor images with contrast enhanced close-ups. Individual views of the reconstructed light fields from noisy sensor images are shown in the right column. Noise amplification is in this case marginal for MURA attenuation masks, while sum-of-sinusoids perform worse than pinholes.

6. Discussion and Conclusions

In this paper, we have derived a new theoretical framework for analyzing plenoptic multiplexing techniques. We demonstrate in particular a close relationship between spatially- and Fourier-multiplexed patterns. General spatial patterns such as the Bayer pattern can be reconstructed

with a Fourier-space algorithm, while patterns derived from Fourier space can also be reconstructed by spatial interpolation, followed by a local decorrelation of the individual channels. In particular, we show the Fourier reconstruction to be equivalent to spatial reconstruction with a sinc filter kernel, resulting in well-known ringing artifacts. We demonstrate that even simple spatial filters such as cubic interpolation can improve the reconstructions significantly. Our theory also allows us to apply edge preserving and other non-linear filters, which further improve the reconstructions.

We use our unified framework for a direct comparison of the noise characteristics of different multiplexing methods. With this analysis, it is for the first time possible to compare in a principled way methods such as light field photography with lens arrays, pinholes, dappled photography, and MURA patterns. In the future, it will be possible to derive optimal multiplexing masks for a given task using the same theoretical framework.

Finally, and maybe most interestingly, our theoretical framework has explicitly enumerated a number of variables controlling plenoptic multiplexing approaches. With this analysis we can now begin to develop and analyze novel multiplexing schemes involving other plenoptic dimensions, or possibly combinations of plenoptic dimensions that have not been explored so far.

Acknowledgments This work was supported by Dolby under Wolfgang Heidrich's Dolby Research Chair. Ivo Ihrke was supported by the Alexander von Humboldt Foundation and the DFG Cluster of Excellence 'Multimodal Computing and Interaction', and Gordon Wetzstein by a Walter C. Koerner Fellowship and a UBC Four Year Fellowship. We would like to thank the Stanford Graphics group and the MIT Camera Culture group for making their datasets publicly available. We would also like to thank the anonymous reviewers for their valuable comments.

References

- [1] E. Adelson and J. Wang. Single Lens Stereo with a Plenoptic Camera. *IEEE Trans. PAMI*, 14(2):99–106, 1992. 2
- [2] E. H. Adelson and J. R. Bergen. The Plenoptic Function and the Elements of Early Vision. In *Computational Models of Visual Processing*, pages 3–20. MIT Press, 1991. 1, 2
- [3] M. Aggarwal and N. Ahuja. Split Aperture Imaging for High Dynamic Range. *Int. J. Comp. Vis.*, 58(1):7–17, 2004. 1
- [4] A. Agrawal, A. Veeraraghavan, and R. Raskar. Reinterpretable Imager: Towards Variable Post-Capture Space, Angle and Time Resolution in Photography. In *Proc. Eurographics*, pages 1–10, 2010. 2
- [5] A. Agrawal, Y. Xu, and R. Raskar. Invertible Motion Blur in Video. *ACM Trans. Graph. (Siggraph)*, 28(3):1–8, 2009. 2
- [6] D. Alleyson, S. Süsstrunk, and J. Hérault. Linear Demosaicing inspired by the Human Visual System. *IEEE Trans. Im. Proc.*, 14(4):439–449, 2005. 2
- [7] B. E. Bayer. Color imaging array, July 1976. US Patent 3,971,065. 1, 2, 4
- [8] T. Bishop, S. Zanetti, and P. Favaro. Light-Field Superresolution. In *Proc. IEEE ICCP*, pages 1–9, 2009. 2
- [9] P. E. Debevec and J. Malik. Recovering High Dynamic Range Radiance Maps from Photographs. In *Proc. ACM Siggraph*, pages 369–378, 1997. 1
- [10] T. Georgiev, C. Intwala, S. Babacan, and A. Lumsdaine. Unified Frequency Domain Analysis of Lightfield Cameras. In *Proc. ECCV*, pages 224–237, 2008. 2, 4
- [11] S. Gortler, R. Grzeszczuk, R. Szelinski, and M. Cohen. The Lumigraph. In *Proc. ACM Siggraph*, pages 43–54, 1996. 1
- [12] J. Greivenkamp. Color Dependant Optical Prefilter for the Suppression of Aliasing Artifacts. *Applied Optics*, 29(5):676–684, 1990. 2
- [13] R. Horstmeyer, G. Euliss, R. Athale, and M. Levoy. Flexible Multimodal Camera Using a Light Field Architecture. In *Proc. ICCP*, pages 1–8, 2009. 2
- [14] D. Lanman, R. Raskar, A. Agrawal, and G. Taubin. Shield Fields: Modeling and Capturing 3D Occluders. *ACM Trans. Graph. (Siggraph Asia)*, 27(5):1–10, 2008. 2, 4, 5, 6
- [15] A. Levin and F. Durand. Linear View Synthesis Using a Dimensionality Gap Light Field Prior. In *Proc. IEEE CVPR*, pages 1–8, 2010. 2
- [16] A. Levin, R. Fergus, F. Durand, and W. Freeman. Image and Depth from a Conventional Camera with a Coded Aperture. *ACM Trans. Graph. (Siggraph)*, 26(3):70:1–70:9, 2007. 2
- [17] A. Levin, W. T. Freeman, and F. Durand. Understanding Camera Trade-Offs through a Bayesian Analysis of Light Field Projections. In *Proc. ECCV*, pages 88–101, 2008. 2
- [18] M. Levoy and P. Hanrahan. Light Field Rendering. In *Proc. ACM Siggraph*, pages 31–42, 1996. 1
- [19] X. Li, B. Gunturk, and L. Zhang. Image Demosaicing: a Systematic Survey. In *SPIE Conf. on Visual Comm. and Image Proc.*, 2008. 5
- [20] G. Lippmann. La Photographie Intégrale. *Academie des Sciences*, 146:446–451, 1908. 2
- [21] A. Lumsdaine and T. Georgiev. The Focused Plenoptic Camera. In *Proc. IEEE ICCP*, pages 1–8, 2009. 2
- [22] M. McGuire, W. Matusik, H. Pfister, B. Chen, J. F. Hughes, and S. K. Nayar. Optical Splitting Trees for High-Precision Monocular Imaging. *IEEE Comput. Graph. & Appl.*, 27(2):32–42, 2007. 1
- [23] T. Mitsunaga and S. K. Nayar. Radiometric Self Calibration. In *Proc. IEEE CVPR*, pages 374–380, 1999. 1
- [24] S. Narasimhan and S. Nayar. Enhancing Resolution along Multiple Imaging Dimensions using Assorted Pixels. *IEEE Trans. PAMI*, 27(4):518–530, Apr 2005. 1, 2
- [25] R. Ng. Fourier Slice Photography. *ACM Trans. Graph. (Siggraph)*, 24(3):735–744, 2005. 2
- [26] R. Raskar, A. Agrawal, and J. Tumblin. Coded Exposure Photography: Motion Deblurring using Fluttered Shutter. *ACM Trans. Graph. (Siggraph)*, 25(3):795–804, 2006. 2
- [27] R. Raskar, A. Agrawal, C. A. Wilson, and A. Veeraraghavan. Glare Aware Photography: 4D Ray Sampling for Reducing Glare Effects of Camera Lenses. *ACM Trans. Graph. (Siggraph)*, 27(3):56:1–56:10, 2008. 2
- [28] Y. Schechner and S. Nayar. Generalized Mosaicing: High Dynamic Range in a Wide Field of View. *IJCV*, 53(3):245–267, 2003. 1
- [29] Y. Schechner, S. Nayar, and P. Belhumeur. Multiplexing for Optimal Lighting. *IEEE Trans. PAMI*, 29(8):1339–1354, 2007. 6
- [30] E.-V. Talvala, A. Adams, M. Horowitz, and M. Levoy. Veiling Glare in High Dynamic Range Imaging. *ACM Trans. Graph. (Siggraph)*, 26(3):37:1–37:9, 2007. 2
- [31] A. Veeraraghavan, R. Raskar, A. Agrawal, R. Chellappa, A. Mohan, and J. Tumblin. Non-Refractive Modulators for Encoding and Capturing Scene Appearance and Depth. In *Proc. IEEE CVPR*, 2008. 2, 4, 5, 6
- [32] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. Dappled Photography: Mask Enhanced Cameras for Heterodyned Light Fields and Coded Aperture Refocussing. *ACM Trans. Graph. (Siggraph)*, 26(3):69:1–69:12, 2007. 1, 2, 4, 5, 6, 7
- [33] S. Wang and W. Heidrich. The Design of an Inexpensive Very High Resolution Scan Camera System. *Computer Graphics Forum (Eurographics)*, 23(10):441–450, 2004. 1
- [34] G. Wetzstein, I. Ihrke, and W. Heidrich. Sensor Saturation in Fourier Multiplexed Imaging. In *Proc. IEEE CVPR*, pages 1–8, 2010. 2
- [35] B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy. High Performance Imaging using Large Camera Arrays. *ACM Trans. Graph. (Siggraph)*, 24(3):765–776, 2005. 1
- [36] A. Wuttig. Optimal Transformations for Optical Multiplex Measurements in the Presence of Photon Noise. *Appl. Opt.*, 44(14):2710–2719, 2005. 6, 7