Int J Comput Vis manuscript No.
(will be inserted by the editor)

On Plenoptic Multiplexing and Reconstruction

Gordon Wetzstein · Ivo Ihrke · Wolfgang Heidrich

Received: date / Accepted: date

Abstract Photography has been striving to capture an ever increasing amount of visual information in a single image. Digital sensors, however, are limited to recording a small subset of the desired information at each pixel. A common approach to overcoming the limitations of sensing hardware is the optical multiplexing of high-dimensional data into a photograph. While this is a well-studied topic for imaging with color filter arrays, we develop a mathematical framework that generalizes multiplexed imaging to all dimensions of the plenoptic function. This framework unifies a wide variety of existing approaches to analyze and reconstruct multiplexed data in either the spatial or the frequency domain. We demonstrate many practical applications of our framework including high-quality light field reconstruction, the first comparative noise analysis of light field attenuation masks, and an analysis of aliasing in multiplexing applications.

Keywords Computational Photography · Optical Multiplexing · Plenoptic Function · Light Fields

1 Introduction

Despite the tremendous advances in camera technology throughout the last decades, the basic principle of operation of modern cameras is still the same as that of Joseph Nicéphore Niépce's camera, which he used to capture the first permanent photograph in 1826. Digital sensors have replaced light sensitive resins and on-board image processing using integrated computing hardware is now common practice, even for consumer-grade digital cameras. However, the acquired visual information has always been what a single human eye can perceive: a two-dimensional trichromatic image. Fueled by advances of digital camera technology and computational processing, image acquisition has begun to transcend limitations of film-based analog photography.

Computational photography has emerged as an interdisciplinary field that is dedicated to the exploration of sophisticated approaches to capturing, analyzing, and processing visual information. Most of the proposed techniques aim at acquiring the dimensions of the plenoptic function (Adelson and Bergen 1991) with combined optical modulation and computational processing (Wetzstein et al 2011). The plenoptic function provides a ray-based model of light encompassing most properties of interest for image acquisition, including the color spectrum as well as spatial, temporal, and directional light variation.

A most desirable plenoptic camera would capture all plenoptic dimensions in a single image using *plenoptic multiplexing*. This can be achieved with something as simple as a color filter array or, more generally, consider additional plenoptic quantities (Narasimhan and Nayar 2005). In either case, a full-resolution image is computed from an interleaved sensor image by interpolating the captured data. Alternatively, an encoding of the spatio-angular plenoptic dimensions, commonly referred to as light fields (Levoy and Hanrahan 1996), can be achieved by multiplexing directional light variation into spatial frequency bands using optical heterodyning (Veeraraghavan et al 2007, 2008; Lanman et al 2008).

In this paper, we introduce a mathematical framework for describing and analyzing plenoptic multiplexing systems. This allows us to cast a large variety of

Gordon Wetzstein
MIT Media Lab
E-mail: gordonw@media.mit.edu

Ivo Ihrke
Universität des Saarlandes
E-mail: ihrke@mmci.uni-saarland.de

Wolfgang Heidrich
The University of British Columbia
E-mail: heidrich@cs.ubc.ca

existing multiplexed imaging approaches into a common framework for analysis, reconstruction, and performance evaluation.

1.1 Contributions

We analyze approaches to acquiring the dimensions of the plenoptic function and present a framework that unifies previously proposed reconstruction methods. Specific contributions are as follows:

- Capture and reconstruction of the plenoptic dimensions are very similar in nature. While intuitive, we demonstrate that sophisticated reconstruction methods developed for one dimension can be similarly applied to other dimensions.
- We introduce a mathematical framework for image formation in plenoptic multiplexing applications. This model generalizes both spatial and Fourier multiplexing methods that have been proposed independently in the literature.
- We present, for the first time, spatial reconstructions of Fourier multiplexed light fields and other plenoptic manifolds and show that the resulting image quality can be significantly increased.
- We establish a metric for the quantitative evaluation of attenuation masks used in light field acquisition. We compare a variety of attenuation patterns and analyze their performance with respect to sensor noise amplification.

1.2 Overview of Benefits and Limitations

The framework introduced in Section 3 shares limitations of other image processing methods, such as color demosaicing: the captured sensor images are assumed to be composed of repeating *super-pixels*. Each of these super-pixels contains different samples of the plenoptic function, but the sampling layout within the super-pixels is spatially invariant. While standard color filter arrays (CFAs, e.g. Bayer (1976)) only perform an interpolation of the spatially interleaved samples, our framework targets more sophisticated multiplexing schemes that require additional data processing after the interpolation. A general assumption for all such approaches is that the sampled signal is band-limited. In practice, this is achieved using optical anti-aliasing filters (e.g., Greivenkamp (1990)).

As illustrated in Figure 1, we demonstrate that our framework allows for reconstructions, that is interpolation and subsequent processing, of multiplexed data in both the spatial and Fourier domain. Although this may seem straightforward for some applications, such as color demosaicing, a variety of mask-based light field acquisition approaches have recently been proposed with corresponding analyses and reconstructions being exclusively performed in the Fourier domain (Veeraraghavan et al 2007, 2008; Lanman et al 2008; Georgiev et al 2008; Agrawal et al 2010b). Our framework is the first to generalize optical multiplexing to all plenoptic dimensions and to demonstrate a unified reconstruction approach in either domain.

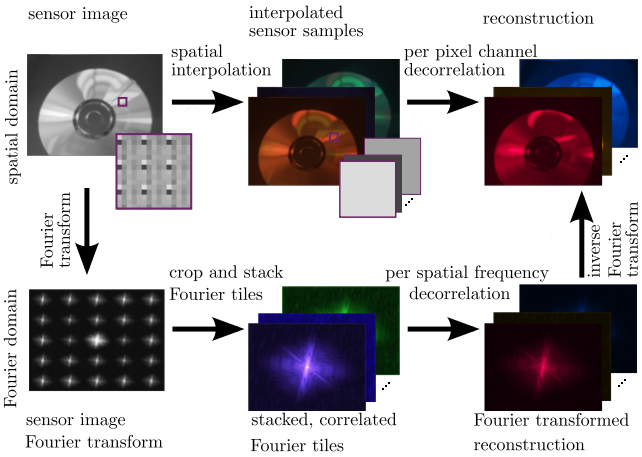


Fig. 1 Overview of multiplexed image reconstruction. The plenoptic function can be reconstructed by interpolating the sensor samples and performing a local decorrelation in the spatial domain (upper row). Alternatively, it can be reconstructed in the Fourier domain by cropping and locally decorrelating Fourier tiles that are created by the periodic structure of the employed optical filters (lower row).

Finally, the proposed formulation allows, for the first time, a quantitative evaluation of attenuation masks for light field acquisition. We compare different designs and demonstrate that the optimal choice, in terms of signal-to-noise ratio, is dependent on camera noise characteristics. We do not propose new optical light modulation techniques to capture any of the plenoptic dimension, but analyze and unify a variety of existing methods; we outline important criteria for the design of optimal light field attenuation masks.

2 Background and Related Work

Standard digital image sensors integrate over all plenoptic dimensions. As a result, most visual information is irreversibly lost during image capture. Three fundamental approaches to overcome these limitations are available: multi-device capture, time-sequential imaging, and single-sensor plenoptic multiplexing.

Multi-device capture refers to approaches that employ multiple image sensors or cameras to simultaneously sample different plenoptic dimensions. Examples include multi-camera systems for capturing light fields (Wilburn et al 2005) or high-speed events (Agrawal et al 2010a) as well as three-chip cameras, which have a sepa-

rate image sensor for the red, green, and blue channel of a color image. Similar approaches have been proposed for high dynamic range (HDR) imaging (Aggarwal and Ahuja 2004) and for general plenoptic image acquisition (McGuire et al 2007).

Time-sequential imaging is commonly used when the hardware requirements or the cost of multi-sensor approaches are prohibitive. For example, light fields can also be captured by moving a single camera to different positions (Levoy and Hanrahan 1996; Gortler et al 1996). Color images can be acquired sequentially by applying different color filters (e.g., Wang and Heidrich (2004)), which is particularly attractive when a large number of color channels is desired. HDR images are usually computed from different exposures (Debevec and Malik 1997; Mitsunaga and Nayar 1999) or using generalized image mosaicing (Schechner and Nayar 2003). An obvious disadvantage of time-sequential capture is the difficulty of capturing dynamic environments and videos.

A common solution for encoding visual information, however, is **plenoptic multiplexing**. Here, multiple differently filtered images are encoded on a single sensor. In effect, this approach trades spatial resolution for the ability to simultaneously capture multiple slices of the plenoptic function. In the literature, plenoptic multiplexing has been analyzed either in the spatial domain or in the Fourier domain.

Spatial multiplexing includes imaging with CFAs to capture color information (e.g., Compton (2007)), parallax barriers (Ives 1903) or lenslet arrays (Lippmann 1908; Adelson and Wang 1992; Ng 2005) for light field acquisition, per-pixel light modulation for high-speed photography (Bub et al 2010), and *Assorted Pixels* (Narasimhan and Nayar 2005; Yasuma et al 2010) as a framework generalizing all these techniques. In each case, a sensor image is comprised of super-pixels that each contain different plenoptic samples at some location on the sensor; interpolating corresponding, interleaved sub-images allows high-resolution imagery to be directly reconstructed.

Optical heterodyning or *Fourier multiplexing* techniques encode different slices of the plenoptic function in different frequency bands. This approach has so far been used for capturing light fields (Veeraraghavan et al 2007, 2008), occluder information (Lanman et al 2008), high-speed events (Agrawal et al 2010b), and high dynamic range photographs (Wetzstein et al 2010). All of these approaches have analyzed and reconstructed captured data exclusively in the Fourier domain. The main difference to spatial multiplexing is that each sensor pixel captures a mixture of all the plenoptic dimensions, which has been shown to improve the light transmission of employed optical filters (Veeraraghavan et al 2007; Lanman et al 2008), but comes at the cost of increased computational complexity. Although spatially encoded data can be analyzed in Fourier space (Alleyson et al 2005; Georgiev et al 2008), we are the first to demonstrate how Fourier multiplexed data can be reconstructed by performing a spatial interpolation followed by a local per-pixel decorrelation of the interpolated data.

Recently, approaches to optically transferring one plenoptic dimension to another, which may then be more convenient to capture, have been proposed. Bando et al (2008) mount a color filter array in the aperture of a standard camera. This encodes the directions of a light field in different colors; a full color light field is subsequently reconstructed using natural image priors. Horstmeyer et al (2009) also insert optical filters in camera apertures, but directly capture a light field with pinhole attenuation masks, where each direction corresponds to a differently filtered version of the photographed scene from slightly different perspectives. Georgiev and Lumsdaine (2010) propose a similar approach using lenslet arrays instead of the pinhole masks. Flexible post-capture resolution tradeoffs have been discussed for spatio-temporal volumes (Gupta et al 2010) and for the optical encoding of temporal light variation in the directions of a light field (Agrawal et al 2010b). The light field in the latter approach is acquired with a sum-of-sinusoids mask, which requires a Fourier-based reconstruction. We demonstrate how the acquisition of mixed plenoptic dimensions, captured with such approaches, can be improved with a spatial reconstruction enabled by our framework.

Compressive sensing (CS) has been introduced as a non-linear reconstruction that aims at beating the Nyquist limit. A single pixel camera, for instance, is proposed by Takhar et al (2006), but requires a large number of photographs; Veeraraghavan et al (2011) and Reddy et al (2011) capture and reconstruct high-speed videos using CS paradigms. Due to the lack of a periodic sampling pattern, these specific CS approaches are not directly supported by our framework.

Other, more exotic camera designs include the focused plenoptic camera (Lumsdaine and Georgiev 2009), light field reconstruction with a Bayesian framework (Levin et al 2008; Levin and Durand 2010), and super-resolution techniques (Bishop et al 2009; Ben-Ezra et al 2005).

3 Plenoptic Multiplexing

A popular approach to capturing high-dimensional visual information with a single photograph is multiplexing. For this purpose, a modulator optically separates this information so that a sensor records an image mosaic containing the desired data. Computational processing is then applied to reconstruct the final full-resolution image.

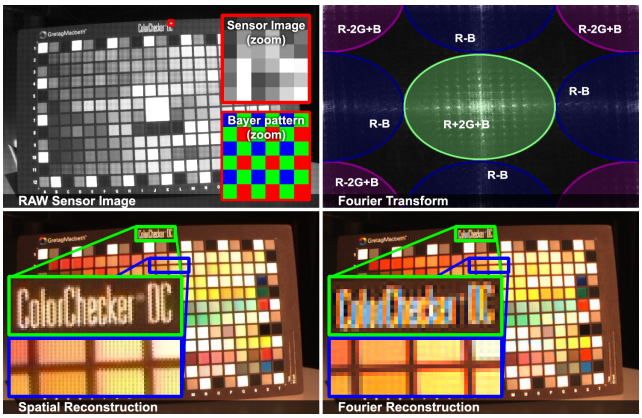


Fig. 2 Upper left: RAW sensor image with close-up and corresponding CFA. Upper right: Fourier transform with channel correlations illustrated for the entire image and the magnified CFA. Reconstructions of the non-perfectly band-limited signal in the spatial (lower left) and Fourier (lower right) domain reveal different aliasing artifacts.

Consider the example shown in Figure 2. A CFA, in this case a Bayer pattern, optically filters the light before it reaches a sensor so that the captured RAW photograph consists of repetitive super-pixels, each encoding four color samples (Fig. 2, upper left). A standard reconstruction or demosaicing interpolates all color channels to every pixel (Fig. 2, lower left). Alternatively, the RAW image can be analyzed in the Fourier domain (Alleyson et al (2005), Fig. 2, upper right), where four different tiles are created that each contain contributions from all color channels. These tiles can be cropped and decorrelated before being transformed back into the spatial domain (Fig. 2, lower right). Although more sophisticated Fourier reconstructions (Li et al 2008) may mitigate visible artifacts, a spatial reconstruction produces much better results in this case, because it is usually more resilient to aliasing artifacts (see Sec. 5).

Imaging with a Bayer pattern is a well-known problem and only serves as an intuitive, motivating example. An additional processing step after interpolating the sensor samples is, in this particular case, not necessary. Here, a Fourier reconstruction is practically not very useful; understanding the process, however, is essential for later parts of the paper, where we consider light field multiplexing approaches that have previously been analyzed exclusively in the Fourier domain.

In the following, we introduce an image formation model for plenoptic multiplexing (Secs. 3.1, 3.2) and demonstrate how this can be used to derive a generic spatial reconstruction algorithm (Sec. 3.3) as well as a corresponding Fourier interpretation (Sec. 3.4). The notation introduced in this section, along with physical interpretations, is summarized in Table 1. All formulations are continuous unless stated otherwise.

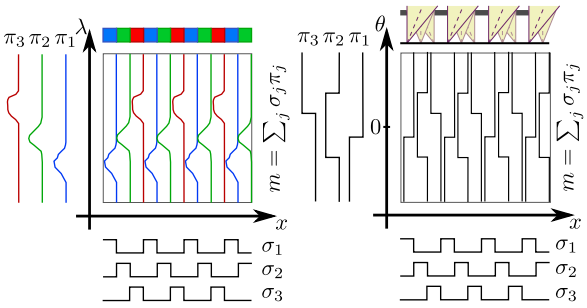


Fig. 3 1D illustration of the plenoptic modulator being separated into a spatial and a plenoptic basis. Left: imaging with CFAs; right: light field capture with an array of pinholes.

3.1 Plenoptic Image Formation

We consider the acquisition of the plenoptic function on the sensor plane, behind the main lens of a camera. A sensor image $\boldsymbol{i}(\vec{x})$ is formed by integrating the plenoptic function $\boldsymbol{l}_\lambda(\vec{x},\vec{p})$ over the plenoptic domain $\mathscr{P} : \Omega_+ \times \mathbb{R} \times \mathbb{R}$ directly above the sensor

$$\boldsymbol{i}(\vec{x}) = \int_{\mathscr{P}} \boldsymbol{m}(\vec{x},\vec{p})\, \boldsymbol{l}_\lambda(\vec{x},\vec{p})\, d\vec{p}. \tag{1}$$

In this formulation, the plenoptic dimensions $\vec{p}$, including directional variation $\vec{\theta}$ over the half-sphere Ω_+, the temporal domain t as well as the color spectrum λ, are separated from the spatial location on the sensor $\vec{x}$. A *plenoptic modulator* $\boldsymbol{m}(\vec{x},\vec{p})$, which is capable of selectively attenuating each dimension at every location, models a generic optical filter. In the case of color imaging this modulator is the CFA, but we show in the following sections that this general formulation includes a wide variety of optical elements. Equation 1 not only accounts for multiplexing different slices of one plenoptic dimension, such as different color channels, onto a sensor, but also for the combined acquisition of multiple dimensions.

3.2 Basis Decomposition

The plenoptic modulator introduced in Equation 1 gives rise to a linear operator describing the projection of the plenoptic function onto the sensor. Transformed into some domain, such linear operators are always separable. The singular value decomposition (SVD) and Fourier analysis are only two examples; in either case a linear operator is decomposed into a basis, and a corresponding inverse, in which the operator is separable. Basis decompositions of this kind provide a powerful tool for theoretical analysis and practical signal processing. We propose a novel basis decomposition for plenoptic imaging; for this purpose, the plenoptic mod-

Symbol	Definition	Physical interpretation	
		Imaging with CFAs	Light field capture
$\boldsymbol{l}_\lambda(\vec{x},\vec{p})$	The plenoptic function, Eqs. 1, 3–5, Sec. 4	Color photograph	Light field
$\hat{\boldsymbol{l}}_\lambda(\vec{\omega}_x,\vec{p})$	Spatial Fourier transform of plenoptic function, Eq. 9		
$\boldsymbol{i}(\vec{x})$	Monochromatic sensor image, Eqs. 1, 4, 6, 8, Sec. 4	RAW sensor mosaic with one sample per pixel	
$\vec{\boldsymbol{i}}(\vec{x})$	Vector of interpolated sensor samples, Eq. 5	Sensor samples interpolated to all positions	
$\hat{\vec{\boldsymbol{i}}}(\vec{\omega}_x)$	Cropped & stacked Fourier tiles of sensor image, Eq. 9	Correlated, high-dimensional FT of $\boldsymbol{l}_\lambda(\vec{x},\vec{p})$	
$\boldsymbol{\sigma}_j(\vec{x})$, $\hat{\boldsymbol{\sigma}}_j(\vec{\omega}_x)$	Spatial basis functions & FT, $j=1\dots N$, Eqs. 2, 4, 6–8	CFA Layout	Layout / spatial frequencies
$\boldsymbol{\pi}_j(\vec{p})$	Plenoptic basis functions, Eqs. 2–4	Spectral transm. of CFA	All angular frequencies
$\boldsymbol{\rho}_j(\vec{x})$, $\hat{\boldsymbol{\rho}}_j(\vec{\omega}_x)$	Plenoptic coefficients & FT, Eqs. 3–6, 8, 9	Color channels	Sampled angular frequencies
$\boldsymbol{\Sigma}$, $\hat{\boldsymbol{\Sigma}}$	Spatial correlation matrix & FT, Eqs. 5, 9	Constant weights, defined by $\boldsymbol{\rho}_j(\vec{x})$, in matrix form	
$\Pi\{\cdot\}$	Projection operator onto the plenoptic basis, Eqs. 5, 9	Projection onto color channels or directions	
$\mathcal{F}\{\cdot\}$	Projection operator onto the Fourier basis, Eqs. 6, 8		
$\boldsymbol{i},\vec{\boldsymbol{i}},\hat{\vec{\boldsymbol{i}}},\vec{\boldsymbol{l}}_\lambda,\boldsymbol{\Pi},\boldsymbol{F}$	Discrete versions of above quantities, Secs. 4, 6		

Table 1 A summary of the notation used in this paper with references to the most important equations.

ulator is separated into a sum of mutually independent spatial and plenoptic basis functions:

$$\boldsymbol{m}(\vec{x},\vec{p}) = \sum_{j=1}^{N} \boldsymbol{\sigma}_j(\vec{x}) \times \boldsymbol{\pi}_j(\vec{p})\,. \tag{2}$$

In this formulation, the plenoptic basis $\boldsymbol{\pi} = \{\boldsymbol{\pi}_j \,|\, j = 1\dots N\}$ is a linear operator that acts on the plenoptic domain, as opposed to the spatial domain, of the plenoptic function. The spatial basis functions $\boldsymbol{\sigma}_j(\vec{x})$, $j = 1\dots N$ describe the layout and mixing of the plenoptic samples on the sensor and only act on the spatial domain of the transformed plenoptic function.

As an example, Figure 3 illustrates possible plenoptic basis functions $\boldsymbol{\pi}_j(\vec{p})$ in color imaging (left) that model the spectral transmissions of the employed color filters, whereas the spatial basis functions $\boldsymbol{\sigma}_j(\vec{x})$ describe the layout of the color samples on the sensor. Figure 3 (right) illustrates these bases for light field cameras with a pinhole array mounted at a slight distance to the sensor.

The proposed basis decomposition, however, is not unique. It is usually related to, but not fully defined by, the optical properties of a specific plenoptic modulator. In fact, any well-known basis can serve as a spatial or plenoptic basis so long as it only acts on either the spatial or the plenoptic domain. In Section 4, for instance, we show that light field acquisition is a convolution allowing us to employ Fourier decomposition as its Eigen-decomposition and similarly as the spatial and plenoptic basis. For the application of general plenoptic image reconstruction and analysis, however, we demonstrate that the proposed decomposition into a spatial and a plenoptic basis facilitates a framework that generalizes multiplexing to all dimensions of the plenoptic function and unifies many existing reconstruction approaches. Furthermore, it proves crucial for analyzing aliasing and noise of a plenoptic imaging system.

In the remainder of this paper, we assume the spatial basis functions to be periodic, thereby implementing the super-pixel concept outlined in Section 1.2. All digital imaging systems are designed to acquire a discrete set of $j = 1\dots N$ plenoptic samples, such as colors or directions, at every pixel. These samples represent projections of the plenoptic function onto the set of plenoptic basis functions

$$\boldsymbol{\rho}_j(\vec{x}) = \int_{\mathscr{P}} \boldsymbol{\pi}_j(\vec{p})\, \boldsymbol{l}_\lambda(\vec{x},\vec{p}) d\vec{p}. \tag{3}$$

We term these projections $\boldsymbol{\rho}_j(\vec{x})$ *plenoptic coefficients* (see Table 1). Their number N often corresponds to the number M of sensor pixels in each super-pixel, but may be lower as in many CFAs where $N = 3$, $M = 4$. Combining Equations 1–3 as

$$\begin{aligned}\boldsymbol{i}(\vec{x}) &= \sum_{j=1}^{N} \boldsymbol{\sigma}_j(\vec{x}) \int_{\mathscr{P}} \boldsymbol{\pi}_j(\vec{p})\, \boldsymbol{l}_\lambda(\vec{x},\vec{p}) d\vec{p} \\ &= \sum_{j=1}^{N} \boldsymbol{\sigma}_j(\vec{x})\, \boldsymbol{\rho}_j(\vec{x})\,,\end{aligned} \tag{4}$$

allows us to model a sensor image at each position $\vec{x}$ as a linear combination of all plenoptic coefficients $\boldsymbol{\rho}_j(\vec{x})$. Spatial multiplexing approaches are designed to directly sample one plenoptic coefficient per sensor pixel. However, Equation 4 similarly models the acquisition of differently weighted linear combinations of all plenoptic coefficients at each pixel.

3.3 Spatial Reconstruction

The goal of a spatial reconstruction is the recovery of all plenoptic quantities at every pixel of the full-resolution sensor image; initially, only one sample is recorded at each pixel. Under the assumptions of an underlying band-limited signal and a super-pixel-periodic spatial basis, this is a standard interpolation or demosaicing of the M interleaved sensor sub-images in $\boldsymbol{i}(\vec{x})$, resulting in the vector-valued image $\vec{\boldsymbol{i}}(\vec{x})$. In order to compute the desired plenoptic samples from $\vec{\boldsymbol{i}}(\vec{x})$, the spatial and

plenoptic bases subsequently need to be inverted in a per-pixel manner.

Before interpolation, every sensor pixel in $\boldsymbol{i}(\vec{x})$ is associated with a single value of all N spatial basis functions $\boldsymbol{\sigma}_j(\vec{x})$ at $\vec{x}$. Interpolating the $k = 1 \dots M$ sub-pixels within the super-pixels to all locations also interpolates the spatial basis. Therefore, each position in $\vec{\boldsymbol{i}}(\vec{x})$ is associated with an array of constants $\boldsymbol{\Sigma} \in \mathbb{R}^{M \times N}$, as defined by the interpolated spatial basis functions. With this notation, we can define

Theorem 1 *(Plenoptic Spatial Multiplexing, PSM). The interpolated sensor samples $\vec{\boldsymbol{i}}(\vec{x})$, representing weighted combinations of all plenoptic coefficients at a particular location $\vec{x}$, are* **locally** *related to the plenoptic function as*

$$\vec{\boldsymbol{i}}(\vec{x}) = \boldsymbol{\Sigma}\vec{\boldsymbol{\rho}}(\vec{x}) = \boldsymbol{\Sigma}\Pi\left\{\boldsymbol{l}_\lambda(\vec{x},\vec{p})\right\}. \tag{5}$$

The operator $\Pi\{\cdot\}$ in Equation 5 projects the plenoptic function onto the plenoptic basis and $\vec{\boldsymbol{\rho}}(\vec{x})$ consists of the N corresponding plenoptic coefficients at each position. Note that $\Pi\{\cdot\}$ is an operator acting on a function, producing a vector $\vec{\boldsymbol{\rho}}(\vec{x})$ with N elements, i.e. $\Pi\{\boldsymbol{l}_\lambda(\vec{x},\vec{p})\} = (<\boldsymbol{\pi}_1, \boldsymbol{l}_\lambda(\vec{x},\vec{p})> \dots <\boldsymbol{\pi}_N, \boldsymbol{l}_\lambda(\vec{x},\vec{p})>)^T$, with $<\cdot,\cdot>$ being the inner product. All quantities are spatially continuous; $\Pi\{\cdot\}$, however, projects the continuous plenoptic function onto a discrete set of plenoptic basis functions. An inversion of this operator can therefore only result in an approximation of the underlying continuous plenoptic quantities. This reconstruction is illustrated in the upper row of Figure 1; the proof for Theorem 1 is included in Appendix A.

The PSM theorem shows that we can reconstruct the plenoptic function $\boldsymbol{l}_\lambda(\vec{x},\vec{p})$ from sensor samples $\boldsymbol{i}(\vec{x})$ by performing a *local* decorrelation on the interpolated measurement samples $\vec{\boldsymbol{i}}(\vec{x})$ followed by an inversion of operator $\Pi\{\cdot\}$; the latter inversion is optional. However, Theorem 1 not only shows that the correlation between the measured samples 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 measured sensor samples $\boldsymbol{i}(\vec{x})$* 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 Section 4 that non-linear filters can achieve high-quality reconstruction results in practice. Non-linear filters are already the preferred choice for color demosaicing; we show that these can also be applied to light field reconstruction.

3.4 Fourier Reconstruction

In recent literature, multiplexing strategies have often been analyzed exclusively in the Fourier domain (Veeraraghavan et al 2007, 2008; Lanman et al 2008; Georgiev et al 2008; Agrawal et al 2010b). 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. 4) is given as

$$\mathcal{F}_x\{\boldsymbol{i}(\vec{x})\} = \mathcal{F}_x\left\{\sum_{j=1}^{N}\boldsymbol{\sigma}_j(\vec{x})\boldsymbol{\rho}_j(\vec{x})\right\} = \sum_{j=1}^{N}\hat{\boldsymbol{\sigma}}_j(\vec{\omega}_x)\otimes\hat{\boldsymbol{\rho}}_j(\vec{\omega}_x), \tag{6}$$

where $\hat{}$ 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 super-pixel-periodic spatial basis functions is given by

$$\hat{\boldsymbol{\sigma}}_j(\vec{\omega}_x) = \sum_{k=1}^{M}\hat{\boldsymbol{\sigma}}_j^{\,k}\,\delta(\vec{\omega}_x - k\Delta\vec{\omega}_x), \tag{7}$$

where $\Delta\vec{\omega}_x$ is the frequency offset or distance between successive Dirac peaks[1] and the values $\hat{\boldsymbol{\sigma}}_j^{\,k}$ are complex weighting factors for spatial basis function j. These weights correspond to the Fourier transform of a single period of that specific basis function. Combining Equations 6 and 7 as

$$\mathcal{F}_x\{\boldsymbol{i}(\vec{x})\} = \sum_{k=1}^{M}\left(\delta(\vec{\omega}_x - k\Delta\vec{\omega}_x)\otimes\sum_{j=1}^{N}\hat{\boldsymbol{\sigma}}_j^{\,k}\hat{\boldsymbol{\rho}}_j(\vec{\omega}_x)\right) \tag{8}$$

shows that, under the band-limit assumption, M different tiles containing linear combinations of Fourier transformed plenoptic coefficients $\hat{\boldsymbol{\rho}}_j(\vec{\omega}_x)$ are created in the frequency domain. As illustrated in Figure 2 (upper right), these tiles can be cropped from the Fourier transformed sensor image and arranged in a stack $\vec{\hat{\boldsymbol{i}}}(\vec{\omega}_x)$ (see Fig. 1, bottom center). The correlation of these stacked tiles is *local in the spatial frequencies $\vec{\omega}_x$*. In analogy to the PSM theorem, Equation 5, we can therefore state the following

Theorem 2 *(Plenoptic Fourier Multiplexing, PFM). Cropping and stacking the individual Fourier tiles of a multiplexed sensor image allows the plenoptic function*

[1] For 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.

to be expressed as a correlation in the Fourier domain that is **local** *for each spatial frequency:*

$$\vec{\hat{\boldsymbol{i}}}(\vec{\omega}_x) = \hat{\boldsymbol{\Sigma}}\, \vec{\hat{\boldsymbol{\rho}}}\,(\vec{\omega}_x) = \hat{\boldsymbol{\Sigma}} \mathit{\Pi} \left\{ \hat{\boldsymbol{l}}_\lambda(\vec{\omega}_x, \vec{p}) \right\}. \tag{9}$$

The correlation matrix $\hat{\boldsymbol{\Sigma}}_{jk} = \hat{\boldsymbol{\sigma}}_j^k$ is determined by the Fourier weights of the spatial basis. In case the spatial basis is the identity $\boldsymbol{\Sigma} = \boldsymbol{I}$, its Fourier weights are the Fourier basis itself $\hat{\boldsymbol{\Sigma}} = \boldsymbol{F}$. The projection onto the plenoptic basis $\mathit{\Pi}\{\cdot\}$ remains unchanged because of its independence of $\vec{x}$ (see Eq. 2), which allows us to change the order of operation, i.e. $\mathcal{F}_x\{\mathit{\Pi}\{\boldsymbol{l}_\lambda(\vec{x},\vec{p})\}\} = \mathit{\Pi}\{\mathcal{F}_x\{\boldsymbol{l}_\lambda(\vec{x},\vec{p})\}\} = \mathit{\Pi}\left\{\hat{\boldsymbol{l}}_\lambda(\vec{\omega}_x,\vec{p})\right\}$. This theorem is illustrated in the lower row of Figure 1; the proof is included in Appendix B.

3.5 Discussion

Previously proposed Fourier multiplexing approaches have analyzed the image formation and reconstruction exclusively in the frequency domain. The mathematical framework presented in this section, however, formulates plenoptic multiplexing and demultiplexing in very general terms. Not only does our framework model the acquisition and reconstruction of arbitrary combinations of plenoptic dimensions, but Theorems 1 and 2 also allow us to analyze and process multiplexed data in either the spatial or the Fourier domain. The notation introduced in this section makes it easy to understand the close connections between the two different interpretations, which is important because each has its own advantages. A spatial reconstruction, that is an interpolation of the sensor samples followed by a per-pixel decorrelation, is generally the preferred method for processing captured data (Sec. 4) and analyzing reconstruction noise (Sec. 6). A Fourier perspective of the same problem, on the other hand, provides a powerful tool for analyzing many important properties such as aliasing (Sec. 5).

Our analysis demonstrates that both spatial and Fourier multiplexing schemes are closely related. The cropping operation in Fourier space is a multiplication with a rect function, which is equivalent to a spatial sinc filter (see App. B). Therefore, all previously proposed Fourier reconstruction methods use a fixed spatial reconstruction filter: the sinc. We demonstrate in the next section that this choice negatively affects the quality of demultiplexed data. More sophisticated apodization approaches (i.e. using a soft roll-off rather than hard cropping) can potentially improve the quality of Fourier-based reconstructions, and are in fact equivalent to using non-sinc linear filters in a spatial reconstruction. Note, however, that non-linear reconstruction filters, including those commonly used for demosaicing, cannot easily be interpreted as Fourier-domain operations.

A consequence of our analysis is that multiplexing schemes have, independent of the reconstruction domain, nominally the same band-limitation requirements. Due to the large choice of linear and non-linear filters, however, a spatial reconstruction can be made more resilient to residual high frequencies, thereby mitigating aliasing artifacts (see Secs. 4, 5).

4 Application to Light Field Reconstruction

In the following, we demonstrate how the general framework introduced in the last section applies to the reconstruction of light fields. We show, for the first time, how Fourier multiplexed light fields captured with non-refractive attenuation masks can be reconstructed with a superior quality in the spatial domain (Sec. 4.1). Although the general plenoptic modulator introduced in Section 3.1 only models selective *attenuation* for each plenoptic dimension, we show in Section 4.2 how similar concepts apply to a variety of acquisition systems with *refractive* optical elements. In Section 4.3, we demonstrate how our framework allows the tempo-directional plenoptic manifolds proposed by Agrawal et al (2010b) to be reconstructed with a higher quality than the originally proposed Fourier processing.

Throughout this section, we employ a two-plane parameterization for light fields. As illustrated for a 1D case in Figure 4, this includes a position $\vec{x}$ on the sensor and the relative distance on a plane at unit distance $\vec{v} = \tan(\vec{\theta})$, which replaces the actual angle $\vec{\theta}$.

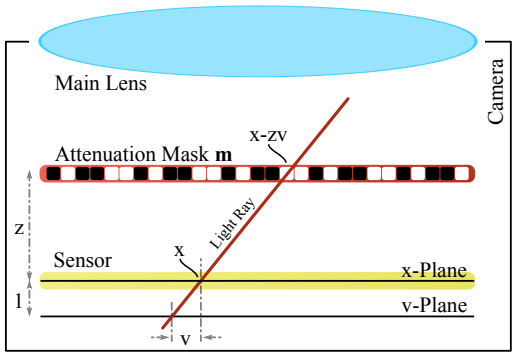


Fig. 4 A light field can be parameterized by a spatial position x on the sensor plane and a relative distance v on a plane at unit distance.

4.1 General Non-Refractive Modulators

Attenuation masks that do not include refractive optical elements have recently been popularized for light

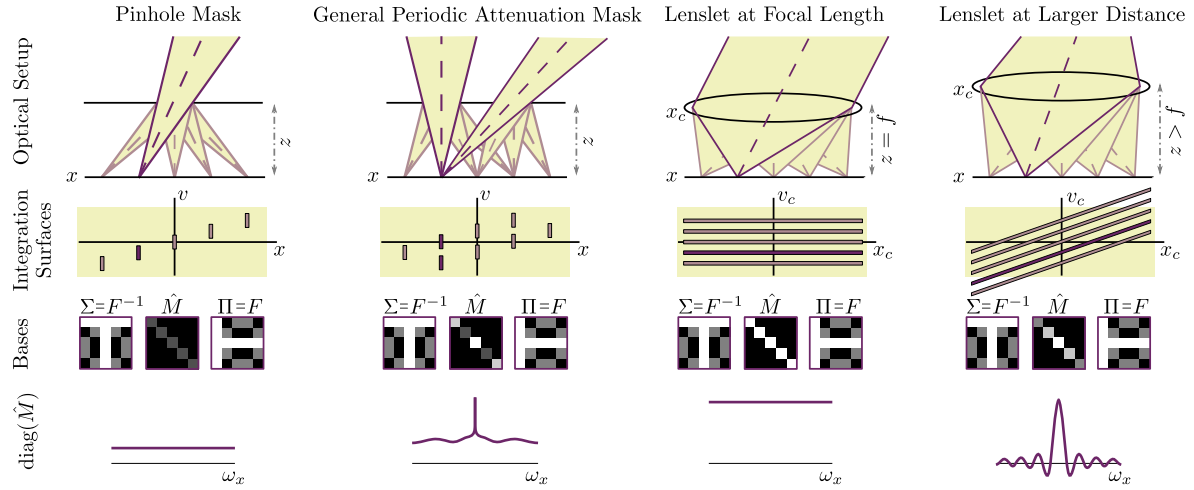


Fig. 5 An illustration of the optical setups, integration surfaces, spatial and plenoptic bases as well as the weighting factors $\hat{\boldsymbol{m}}$ for a variety of super-pixel based light field cameras. The convolution of a light field and a periodic attenuation mask or refractive optical element, resulting in the captured sensor image, can be separated into a spatial and a plenoptic part using the Fourier basis (row 3). This allows us to perform a light field reconstruction in either the spatial or Fourier domain by applying the PSM or PFM theorem, respectively. Please note that the camera and sensor coordinates for non-refractive elements are identical; the integration surfaces for refractive optical elements already include the mapping from sensor space to world space on the microlens plane.

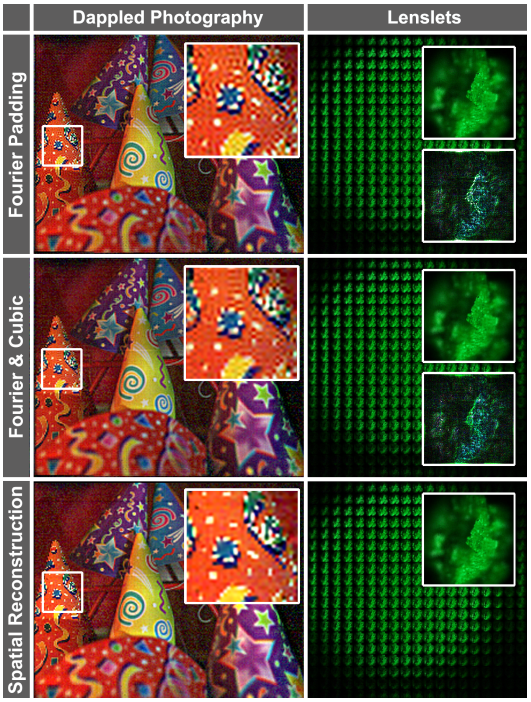


Fig. 6 Comparison of reconstruction quality for Cones data set (Veeraraghavan et al 2007) captured with a non-refractive sum-of-sinusoids mask and lenslet-based Fluorescent Crayon Wax data set (Levoy et al 2006). All results are three-times upsampled during reconstruction. Top row: upsampling by zero-padding the 4D inverse DFT. Middle row: low resolution 4D inverse DFT followed by bicubic upsampling. Bottom row: bicubic up-sampling followed by local decorrelation. For the right column, we show one of the light field views and a contrast enhanced difference image to the spatial reconstruction in the magnifications. Ringing artifacts are clearly visible.

field acquisition (Veeraraghavan et al 2007, 2008; Lanman et al 2008; Georgiev et al 2008). All of these approaches have been analyzed and reconstructed exclusively in the Fourier domain. Here, we show how the employed periodic attenuation masks can be separated into a spatial and a plenoptic basis. This separation allows the aforementioned techniques to be expressed in the framework introduced in Section 3.

As illustrated in Figure 4, the plenoptic modulator (Eq. 1) for attenuation masks at a distance z to a sensor is $\boldsymbol{m}(\vec{x},\vec{v}) = \boldsymbol{m}(\vec{x}-z\vec{v})$. This formulation models the light transport in free space from sensor to mask as well as the attenuation caused by the latter. A separation of this modulator into a purely spatial and a plenoptic, in this case directional, part can be achieved by substituting the modulator $\boldsymbol{m}(\vec{x}-z\vec{v})$ with its inverse-transformed Fourier transform

$$\begin{aligned}\boldsymbol{i}(\vec{x}) &= \int_{\vec{v}} \boldsymbol{l}_\lambda(\vec{x},\vec{v})\boldsymbol{m}(\vec{x}-z\vec{v})\,d\vec{v} \\ &= \int_{\vec{v}} \boldsymbol{l}_\lambda(\vec{x},\vec{v}) \int_{\vec{\omega}_x} \hat{\boldsymbol{m}}(\vec{\omega}_x) \mathrm{e}^{2\pi i(\vec{x}-z\vec{v})\cdot\vec{\omega}_x}\,d\vec{\omega}_x\,d\vec{v} \\ &= \int_{\vec{\omega}_x} \hat{\boldsymbol{m}}(\vec{\omega}_x)\mathrm{e}^{2\pi i\vec{x}\cdot\vec{\omega}_x} \int_{\vec{v}} \boldsymbol{l}_\lambda(\vec{x},\vec{v})\mathrm{e}^{-2\pi i z\vec{v}\cdot\vec{\omega}_x}\,d\vec{v}\,d\vec{\omega}_x.\end{aligned} \tag{10}$$

Equation 10 shows that the modulator projects the light field onto its angular frequencies, which are then multiplexed, with mask-dependent weights $\hat{\boldsymbol{m}}(\vec{\omega}_x)$, into the spatial frequencies of a sensor image. This decomposition into the Fourier and the inverse Fourier basis satisfies the necessary requirements for the Fourier transform to be interpreted as the plenoptic basis, only

acting on the directional components of the incident light field, and the inverse Fourier transform being the spatial basis. An intuitive interpretation for why the Fourier basis and its inverse serve as the plenoptic and spatial bases in this case is that the image formation in Equation 10 is a convolution; a Fourier decomposition is the Eigen-decomposition of any circular convolution process.

For a practical processing of digital images, the integrals can be discretized, both spatially and plenoptically, as $\vec{\boldsymbol{i}} = \boldsymbol{F}^{-1}\hat{\boldsymbol{M}}\boldsymbol{F}\vec{\boldsymbol{l}}_\lambda$, with $\hat{\boldsymbol{M}} = diag(\hat{\boldsymbol{m}})$. In this case, $\vec{\boldsymbol{l}}_\lambda$ is a stack of N images containing the discretized plenoptic function. As illustrated in Figure 5 (center left), this formulation allows us to choose a weighted inverse discrete Fourier transform (DFT) as the spatial basis $\boldsymbol{\Sigma} = \boldsymbol{F}^{-1}\hat{\boldsymbol{M}}$, and the DFT as the discretized plenoptic projection $\boldsymbol{\Pi} = \boldsymbol{F}$. A spatial reconstruction can therefore be performed by solving Equation 5. Again, a discrete spatial reconstruction is performed by interpolating the measured samples in the sensor image $\boldsymbol{i}$ to all positions and then decorrelating the resulting, vector-valued, discrete image $\vec{\boldsymbol{i}}$ in a per-pixel manner as

$$\vec{\boldsymbol{l}}_\lambda = \boldsymbol{\Pi}^{-1}\boldsymbol{\Sigma}^{-1}\vec{\boldsymbol{i}} = \boldsymbol{F}^{-1}\hat{\boldsymbol{M}}^{-1}\boldsymbol{F}\vec{\boldsymbol{i}}. \tag{11}$$

Equation 11 is a *per-image-pixel deconvolution* with a kernel that is defined by the attenuation pattern of the mask. As shown in Section 6, a deconvolution with sum-of-sinusoids patterns (Veeraraghavan et al 2007; Agrawal et al 2010b), for instance, represents a high-pass filter. Unfortunately, this type of filter decreases the signal-to-noise ratio (SNR) of the reconstructed light field significantly by amplifying sensor noise.

Alternatively, the previously proposed discrete Fourier reconstruction can be performed by directly solving Equation 9 as:

$$\hat{\vec{\boldsymbol{l}}}_\lambda = \boldsymbol{\Pi}^{-1}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\vec{\boldsymbol{i}}} = \boldsymbol{F}^{-1}\hat{\boldsymbol{M}}^{-1}\hat{\vec{\boldsymbol{i}}}. \tag{12}$$

Equation 12 shows that the stacked discrete Fourier tiles $\hat{\vec{\boldsymbol{i}}}$ need to be re-weighted on a *per-spatial-frequency* basis; the weighting factors $\hat{\boldsymbol{m}}$ depend on the applied mask. An inverse DFT is required to invert the plenoptic basis $\boldsymbol{\Pi}$. An additional inverse DFT is applied to compute the desired plenoptic samples $\vec{\boldsymbol{l}}_\lambda$ from $\hat{\vec{\boldsymbol{l}}}_\lambda$, which is equivalent to an inverse 4D Fourier transform of cropped and stacked light field Fourier tiles (Veeraraghavan et al 2007, 2008; Lanman et al 2008; Georgiev et al 2008).

Figure 6 (left) shows comparisons of light field reconstructions in the Fourier domain, as previously proposed, and in the spatial domain. Noise and ringing artifacts are significantly reduced in the spatial reconstruction, which is enabled by our framework. Even with a simple linear spatial interpolation scheme such as cubic interpolation, common ringing artifacts associated with Fourier-based techniques can be avoided. A more detailed discussion on aliasing artifacts can be found in Section 5.

Although Corollary 1 is theoretically only valid for linear filters, Figure 7 demonstrates that a practical reconstruction works well with non-linear filters. The presented example shows two different viewpoints of the reconstructed Mannequin dataset (Lanman et al 2008) and, more importantly, a spatial reconstruction with a non-linear joint bilateral filter. This filter is just one of many possible choices for sophisticated spatial reconstruction filters facilitated by Theorem 1. The joint bilateral filter can, in this case, reconstruct a slightly sharper image than a bicubic filter.

4.2 Refractive Modulators

The general plenoptic modulator introduced in Section 3.1 does not directly model ray deflections caused by refractive optical elements. However, the image formation can be modeled as a convolution of the plenoptic function and a point spread function (PSF) $\boldsymbol{k}(\vec{x}, \vec{x}_c, \vec{v}_c)$ (Levin et al 2009):

$$\boldsymbol{i}(\vec{x}) = \int_{\vec{x}_c}\int_{\vec{v}_c} \boldsymbol{k}(\vec{x}, \vec{x}_c, \vec{v}_c)\, \boldsymbol{l}_\lambda(\vec{x}_c, \vec{v}_c)\, d\vec{x}_c d\vec{v}_c, \tag{13}$$

where $\vec{x}$ is the spatial coordinate on the sensor surface and $\vec{x}_c$ is the spatial coordinate defined on the plane of the refractive elements (see Fig. 5, right). Here, $\boldsymbol{l}_\lambda(\vec{x}_c, \vec{v}_c)$ is the light field on the plane of the refractive elements; the directional coordinate $\vec{v}_c$ describes the directional light variation before modulation by the PSF inside the camera behind the main lens. PSFs for different optical elements are well known. Under paraxial approximations and disregarding the element's aperture and wavelength of light, the PSF for most refractive elements is of the form

$$\boldsymbol{k}(\vec{x}, \vec{x}_c, \vec{v}_c) = \delta\left(\vec{x} - z\vec{v}_c - \phi(\vec{x}_c)\right). \tag{14}$$

The specific PSF for a lens at focal distance to the sensor is given by $\phi_{z=f}(\vec{x}_c) = 0$. The term becomes non-zero when the lens is moved away from the focal distance, as proposed by Lumsdaine and Georgiev (2009): $\phi_{z\neq f}(\vec{x}_c) = s\vec{x}_c$, with $s = 1 - z/f$ being the slope of the integration surface. Optical setups and integration surfaces for all of these cases are illustrated in Figure 5.

In the following, we demonstrate how a separation of the PSF into a spatial and a plenoptic basis can be performed. This is shown for a single element of an array

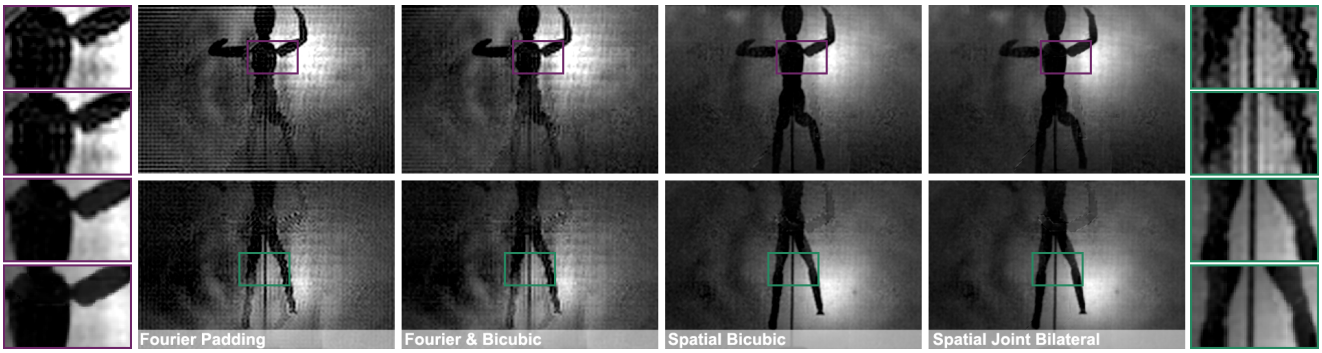


Fig. 7 Reconstruction results for Mannequin data set (Lanman et al 2008). The results are three-times upsampled and show two different views of the reconstruction. The fifth column shows a reconstruction that was computed with a non-linear joint bilateral filter. Close-ups are shown for the top row on the left and for the bottom row on the right.

of optical elements such as lenslets. The plenoptic function is assumed to be spatially band-limited, which in this case requires it to be constant over a *single optical element*, i.e. $\boldsymbol{l}_\lambda(\vec{x}_c, \vec{v}_c) = \boldsymbol{l}_\lambda(\vec{v}_c)$. Combining Equations 13 and 14 and substituting the PSF with its inverse-transformed Fourier transform yields

$$\begin{aligned}\boldsymbol{i}(\vec{x}) &= \int_{\vec{x}_c}\int_{\vec{v}_c} \boldsymbol{l}_\lambda(\vec{v}_c)\,\delta(\vec{x} - z\vec{v}_c - \phi(\vec{x}_c))\,d\vec{x}_c d\vec{v}_c \qquad (15)\\ &= \int_{\vec{x}_c}\int_{\vec{v}_c} \boldsymbol{l}_\lambda(\vec{v}_c) \int_{\vec{\omega}_x} e^{2\pi i(\vec{x} - z\vec{v}_c - \phi(\vec{x}_c))\cdot\vec{\omega}_x} d\vec{\omega}_x d\vec{x}_c d\vec{v}_c\\ &= \int_{\vec{\omega}_x} \hat{\boldsymbol{m}}(\vec{\omega}_x) e^{2\pi i\vec{x}\cdot\vec{\omega}_x} \int_{\vec{v}_c} \boldsymbol{l}_\lambda(\vec{v}_c)\, e^{-2\pi i z\vec{v}_c\cdot\vec{\omega}_x} d\vec{v}_c\, d\vec{\omega}_x,\end{aligned}$$

where the extra term $\hat{\boldsymbol{m}}(\vec{\omega}_x) = \int_{\vec{x}_c} e^{-2\pi i\phi(\vec{x}_c)\cdot\vec{\omega}_x} d\vec{x}_c$ varies for different refractive elements and represents their optical transfer function (OTF, see Figure 5). Lenses at focal distance to the sensor are not affected by the extra term, i.e. $\hat{\boldsymbol{m}}(\vec{\omega}_x) = 1,\ \forall\vec{\omega}_x$.

Figure 6 (right) shows a comparison of spatial and Fourier reconstructions of a light field that was captured inside a microscope using a lenslet array at focal distance to the sensor (Levoy et al 2006). A decorrelation of the interpolated sensor image $\vec{\boldsymbol{i}}(\vec{x})$ is in this case redundant, as every sensor pixel measures uncorrelated directional samples of the light field. Although artifacts are visible in the magnifications of Figure 6 (right), they are more subtle than those in mask-based light field reconstructions (Fig. 6, left). This can be attributed to pixels under each microlens integrating spatial light variation over the entire lenslet area, which provides a proper spatial band-limit and therefore minimizes aliasing (Levoy and Hanrahan 1996) (see Sec. 5).

Employing lenslets at a distance to the sensor that is different to their focal lengths, i.e. $z \neq f$, was explored by Lumsdaine and Georgiev (2009). It was shown that such a setup in combination with a custom resorting algorithm is capable of reconstructing light fields with a higher spatial but reduced angular resolution. In this particular setup, the corresponding OTF is a sinc (Fig. 5, right). Applying our theory would aim at reconstructing the full spatial and directional resolution of the light field, which is ill-conditioned due to the zero-crossings of the OTF, unless additional statistical priors are incorporated.

4.3 Plenoptic Dimension Transfer

Agrawal et al (2010b) propose to equip a camera aperture with a pinhole mask that can be moved throughout the exposure time of a single photograph. This motion encodes temporal light variation in the directions of a light field. The light field itself is acquired by mounting an additional sum-of-sinusoids attenuation mask at a small distance to the camera sensor. Such a setup allows the captured photograph to be reinterpreted as either a high spatial resolution image, a light field, or a video for different parts of the scene in post-processing. Mathematically, the image formation can be formulated as

$$\boldsymbol{i}(\vec{x}) = \int_{\vec{v}}\int_{t} \boldsymbol{l}_\lambda(\vec{x},\vec{v},t)\,\boldsymbol{m}_t(\vec{x},\vec{v},t)\,\boldsymbol{m}_v(\vec{x},\vec{v},t)\,dtd\vec{v}, \quad (16)$$

where $\boldsymbol{m}_v(\vec{x},\vec{v},t) = \boldsymbol{m}(\vec{x} - z\vec{v})$ is the non-refractive attenuation mask, as introduced in Section 4.1, and $\boldsymbol{m}_t(\vec{x},\vec{v},t) = \delta(\vec{v} - \psi_{\vec{v}}(t))$ is the moving pinhole aperture. The pinhole motion is described by $\psi_{\vec{v}}(t)$. In our framework, this can be expressed as

$$\begin{aligned}\boldsymbol{i}(\vec{x}) &= \int_{\vec{v}}\int_{t} \boldsymbol{l}_\lambda(\vec{x},\vec{v},t)\,\delta(\vec{v} - \psi_{\vec{v}}(t))\,\boldsymbol{m}(\vec{x} - z\vec{v})\,dtd\vec{v}\\ &= \int_{t} \boldsymbol{l}_\lambda(\vec{x},\psi_{\vec{v}}(t),t)\,\boldsymbol{m}(\vec{x} - z\psi_{\vec{v}}(t))\,dt\\ &= \int_{\vec{\omega}_x} \hat{\boldsymbol{m}}(\vec{\omega}_x) e^{2\pi i\vec{x}\cdot\vec{\omega}_x}\\ &\qquad \int_{t} \boldsymbol{l}_\lambda(\vec{x},\psi_{\vec{v}}(t),t)\, e^{-2\pi i z\psi_{\vec{v}}(t)\cdot\vec{\omega}_x} dtd\vec{\omega}_x. \qquad (17)\end{aligned}$$



Fig. 8 Two different views of a light field reconstructed with an upsampling factor of three. Each view encodes a different temporal slice within the exposure time of a single photograph, as proposed by Agrawal et al (2010b). The spatial reconstruction (bottom row) increases the reconstruction quality by reducing ringing artifacts that are visible in previously proposed Fourier reconstructions (top and center row).



Fig. 9 Analyzing aliasing artifacts. Multiplexing a light field with 5×5 views onto a sensor is simulated without (left column) and with (right column) a synthetic optical anti-aliasing filter applied. The Fourier tiles of the sensor image slightly overlap without any filtering (upper left). A proper anti-aliasing filter mitigates this overlap and thereby artifacts in the reconstructions (right). A spatial processing of multiplexed data (lower left) is usually more robust to signal aliasing than a corresponding Fourier approach (center left).

The pinhole motion $\psi_{\vec{v}}(t)$ introduces a manifold in the tempo-directional domain of the plenoptic function over which the sensor integrates. A reconstruction can only recover this manifold; temporal and directional light variation are coupled. In our framework, the plenoptic basis for this example is the Fourier transform of the plenoptic manifold, whereas the spatial basis is, just as in the case of attenuation masks and refractive optical elements, the inverse Fourier transform.

Figure 8 shows spatial and Fourier-based reconstructions of one of the datasets used by Agrawal et al (2010b). Columns 1 and 2 show two different views of the reconstructed light field. Each of these views additionally encodes a different time in the animation: the Rubik's cube is moved away from the camera. As seen in the close-ups, a spatial reconstruction, enabled by Theorem 1, can reduce ringing artifacts as compared to previously employed Fourier reconstructions.

5 Analyzing Aliasing

One of the main arguments throughout the last section is that a spatial reconstruction of multiplexed data can improve the image quality. We demonstrate in this section, that the difference in quality is mainly due to aliasing, that is violations of the band-limit assumption. While spatial processing with sophisticated reconstruction filters offers the benefit of improved image quality, a Fourier perspective allows aliasing to be analyzed in a convenient manner.

Consider the experiment in Figure 9. A light field (cs.ubc.ca/~wetzste1/SyntheticLightFields/) is multiplexed onto a single sensor with a MURA attenuation mask. The mask consists of a repetitive pattern of 5×5 pixels, as introduced by Gottesman and Fenimore (1989). Figure 9 (upper left) shows the Fourier transform of a simulated sensor image without any anti-aliasing applied. The grid of 5×5 different Fourier tiles is clearly visible. Without optically filtering the captured signal, the copies slightly overlap, thereby causing pre-aliasing in the reconstructions (Fig. 9, rows 2, 3). As expected from the results presented in Section 4, aliasing artifacts in the Fourier reconstruction (Fig. 9, left center) are much stronger than in the corresponding spatial reconstruction (Fig. 9, left bottom). With a proper anti-aliasing filter applied to the signal before capture, however, the quality differences become more subtle (Fig. 9, right column). While a spatial reconstruction is more resilient to pre-aliasing, artifacts persist that can be mitigated for both types

of reconstruction by optically anti-aliasing the signal before capture.

For this experiment we applied a first-order Butterworth low-pass filter to each input image, before multiplexing them on the sensor. This filter significantly reduces pre-aliasing in the recorded signal by eliminating high spatial frequencies from the light field. In the Fourier domain, reduced aliasing is visible by less overlap of the frequency tiles (Fig. 9, upper right).

Optical anti-aliasing filters are common practice for imaging with CFAs (e.g., Greivenkamp (1990)). While attenuation masks have been optimized for the 4D frequency distribution of natural scenes (Veeraraghavan et al 2008), anti-aliasing mechanisms have yet to be implemented for mask-based light field cameras.

6 Analyzing Light Field Reconstruction Noise

The last sections have discussed a theoretical framework for modeling plenoptic multiplexing with practical applications in high-quality reconstructions. In many multiplexing tasks, however, there is a wide choice of optical filters or modulators that can be employed to acquire the same visual information. Several different CFA designs are, for instance, available for color imaging. While RGB-based color filters are usually preferred in bright lighting conditions, CMY filters provide better noise characteristics in low-light conditions (Sajadi et al 2011). In this section, we analyze and compare a variety of optical multiplexing filters for two different applications: imaging with color filter arrays and light field acquisition. While color imaging serves as an intuitive example to validate the noise model introduced in this section, we present the first comparative analysis of non-refractive light field multiplexing masks. Most of the employed attenuation patterns (Veeraraghavan et al 2007, 2008; Lanman et al 2008) have been exclusively analyzed in the Fourier domain, which makes it difficult to evaluate the noise characteristics of the actual mask pattern. The spatial analysis enabled by Theorem 1 allows us to compare the performance of alternative attenuation masks with respect to the signal-to-noise ratio (SNR) in reconstructed imagery.

We employ a noise model that is commonly applied in computer vision (Wuttig 2005; Schechner et al 2007). The total noise variance ς^2 of a camera image is modeled as the combination of a signal-independent additive term ς_c^2, which includes dark current and amplifier noise, as well as a signal-dependent photon shot noise term $\boldsymbol{i}(\vec{x})\varsigma_p^2$. Following standard practice (Schechner et al 2007), we approximate the image intensity term in the photon noise by the mean light transmission $\boldsymbol{\tau}$ of the plenoptic modulator, yielding the following noise variance in the captured image:

$$\varsigma^2 = \varsigma_c^2 + \boldsymbol{\tau}\varsigma_p^2. \tag{18}$$

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 introduce a noise amplification term $\boldsymbol{\alpha}$ that is based on our discretized image formation (Sec. 4, see Schechner et al (2007) for more details):

$$\boldsymbol{\alpha} = \sqrt{\frac{1}{N} trace\left(\left(\boldsymbol{\Pi}\boldsymbol{\Sigma}\right)^T \boldsymbol{\Sigma}\boldsymbol{\Pi}\right)^{-1}}. \tag{19}$$

The signal-to-noise ratio in the demultiplexed plenoptic function requires expressions for the signal and for the noise term. Assuming a normalized plenoptic function and orthogonal plenoptic basis functions, the signal term in the demultiplexed signal can be approximated by $1/N$, where N is the number of sampled plenoptic coefficients (e.g, color channels or light field views). The reconstruction noise term is the standard deviation of the propagated sensor noise $\boldsymbol{\alpha}\varsigma$, resulting in an SNR of

$$\mathrm{SNR} = 10\log_{10}\left(\frac{1}{N\boldsymbol{\alpha}\sqrt{\varsigma_c^2 + \boldsymbol{\tau}\varsigma_p^2}}\right), \tag{20}$$

where the SNR is defined in dB. The gain of SNR for this demultiplexed signal compared to some demultiplexed reference signal with a noise term of ς_{ref} is then

$$g_{\mathrm{SNR}} = \mathrm{SNR} - \mathrm{SNR}_{ref} = 10\log_{10}\left(\frac{\boldsymbol{\alpha}_{ref}\varsigma_{ref}}{\boldsymbol{\alpha}\varsigma}\right). \tag{21}$$

A positive gain indicates an improved SNR, relative to the reference signal, while a negative gain indicates noise amplification. A plot of SNR gain for a standard Bayer CFA filter, which serves as the reference signal, compared to the more transmissive cyan-yellow-magenta-yellow (CYMY), cyan-yellow-magenta-green (CYMG), and red-green-blue-white (RGBW) CFAs is shown in Figure 10 (left). The plots demonstrate 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. We employ the notation introduced by Wuttig (2005), where a parameter $\chi = \varsigma_p/\varsigma_c$ describes the ratio of signal-dependent and signal-independent noise terms. This makes it more convenient to plot the performance of a multiplexing

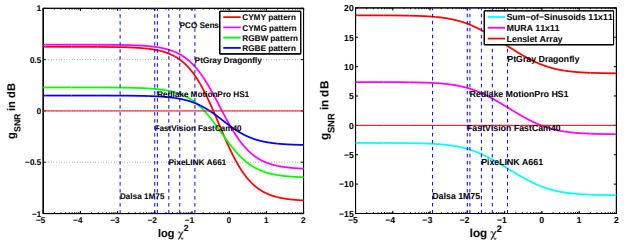


Fig. 10 SNR comparison of various alternative CFA patterns to the Bayer pattern (left). SNR comparison of different light field attenuation masks (right). The vertical lines indicate χ^2 values for several machine vision cameras tested by Schechner et al (2007). All cameras are operating in the gain region of the filters, i.e. $g_{\mathrm{SNR}} > 0$. 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.

scheme with different camera noise parameters up to a global scale, as seen in Figure 10. Our CFA noise analysis helps to determine the exact gain regions of a particular setup, which is especially important for dynamically switchable implementations (Sajadi et al 2011).

Similarly, we compare the noise performance of various light field acquisition approaches, where the pinhole attenuation mask serves as the reference. The size of the simulated pinhole matches that of a sensor pixel. Sensor quantization and other non-linearities are disregarded in this experiment. The plot in Figure 10 (right) shows that lenslets at focal distance to the sensor always perform best in terms of SNR. Among the non-refractive multiplexing methods, MURA-based attenuation masks (Lanman et al 2008) perform very well for a dominating additive noise term, i.e. at high camera gain settings in low-light conditions. However, their SNR gain drops below that of a pinhole for an increasingly dominating photon noise term. Sum-of-sinusoids masks (Veeraraghavan et al 2007; Agrawal et al 2010b) always perform worse than a pinhole.

When considering only additive, signal-independent noise in the captured sensor images, which is most often the case in low-light environments, the noise of the reconstructed plenoptic slices can be quantified by the covariance matrix C:

$$C = \varsigma^2 \left((\boldsymbol{\Sigma\Pi})^T (\boldsymbol{\Sigma\Pi}) \right)^{-1}, \tag{22}$$

where ς^2 is the variance of an additive, zero-mean Gaussian noise distribution in the sensor image $\boldsymbol{i}(\vec{x})$. Figure 11 shows the magnitudes of C for several light field acquisition schemes assuming $\varsigma^2 = 1$. Values larger than 1 amplify noise in the camera image and off-diagonal entries accumulate noise from different regions of the captured images. The covariance matrix of the sum-of-sinusoids (SoS) mask has many large-valued off-diagonal entries, which indicates noise amplification in the reconstruction. The matrix for MURA masks does have off-diagonal entries, but with much smaller magnitudes than SoS masks. Similar interpretations can be inferred from the plots of the singular values of the multiplexing matrix $\boldsymbol{\Sigma\Pi}$ in Figure 11 (right).

Based on the covariance analysis and predicted SNR gain (Fig. 10, right), we expect SoS masks to amplify sensor noise more than both pinhole and MURA attenuation masks for a dominating additive noise term. In order to validate this prediction, we simulate the acquisition of a light field with a variety of different methods and camera noise parameters (Fig. 12). In this experiment, we use attenuation masks with a resolution of 11×11 for each super-pixel and a similar sensor resolution. Each lenslet in the first column of Figure 12 covers the same area as a corresponding mask super-pixel. As expected, for low-light conditions (Fig. 12, row 2) lenslet arrays and MURA masks have a better noise performance than pinhole masks, whereas SoS masks perform worse (Fig. 12, column 3). The value $\log(\chi^2) = -0.93$ corresponds to a PointGrey Dragonfly camera (Schechner et al 2007) where the additive noise term dominates. The lower two rows in Figure 12 show how the noise increases for an increasingly dominating photon noise term up to a point where a pinhole mask performs better than even the MURA mask. The same effect was described by Wenger et al (2005) for Hadamard codes in an illumination multiplexing application. Please note that for our analysis the exposure times of the simulated sensors were equal for each method, resulting in visible intensity differences between the sensor images (Fig. 12, top row).

7 Discussion and Conclusions

In this paper, we have introduced a framework that unifies a variety of plenoptic multiplexing approaches. Previously, these techniques have been analyzed with respect to a specific plenoptic dimension. In most cases, reconstructions have been performed exclusively in either the spatial or the Fourier domain. We have demonstrated the importance of such a unified view: certain properties, such as aliasing, can be theoretically analyzed more conveniently in the frequency domain. Other characteristics, such as noise amplification of the employed optical modulators, are easier to be evaluated in the spatial domain. We show, for the first time, how the quality of practical reconstruction mechanisms for some of the discussed techniques can be increased with spatial processing, rather than previously proposed Fourier-based algorithms. The latter, however, may require fewer computational resources.

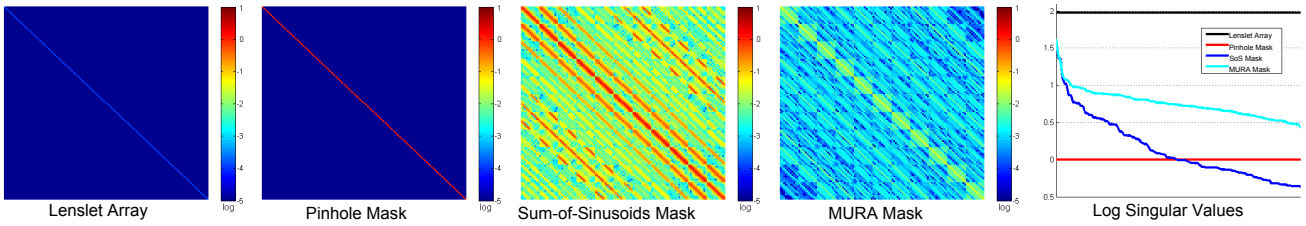


Fig. 11 Covariance matrices and eigenvalues for different light field multiplexing schemes. Large values and especially off-diagonal entries in the covariance matrices indicate amplification of additive noise in the sensor images. The sum-of-sinusoids mask is thus expected to perform worse than a pinhole mask for dominating dark current noise in the sensor image, which can also be inferred from the plots showing the multiplexing matrix's singular values (right).

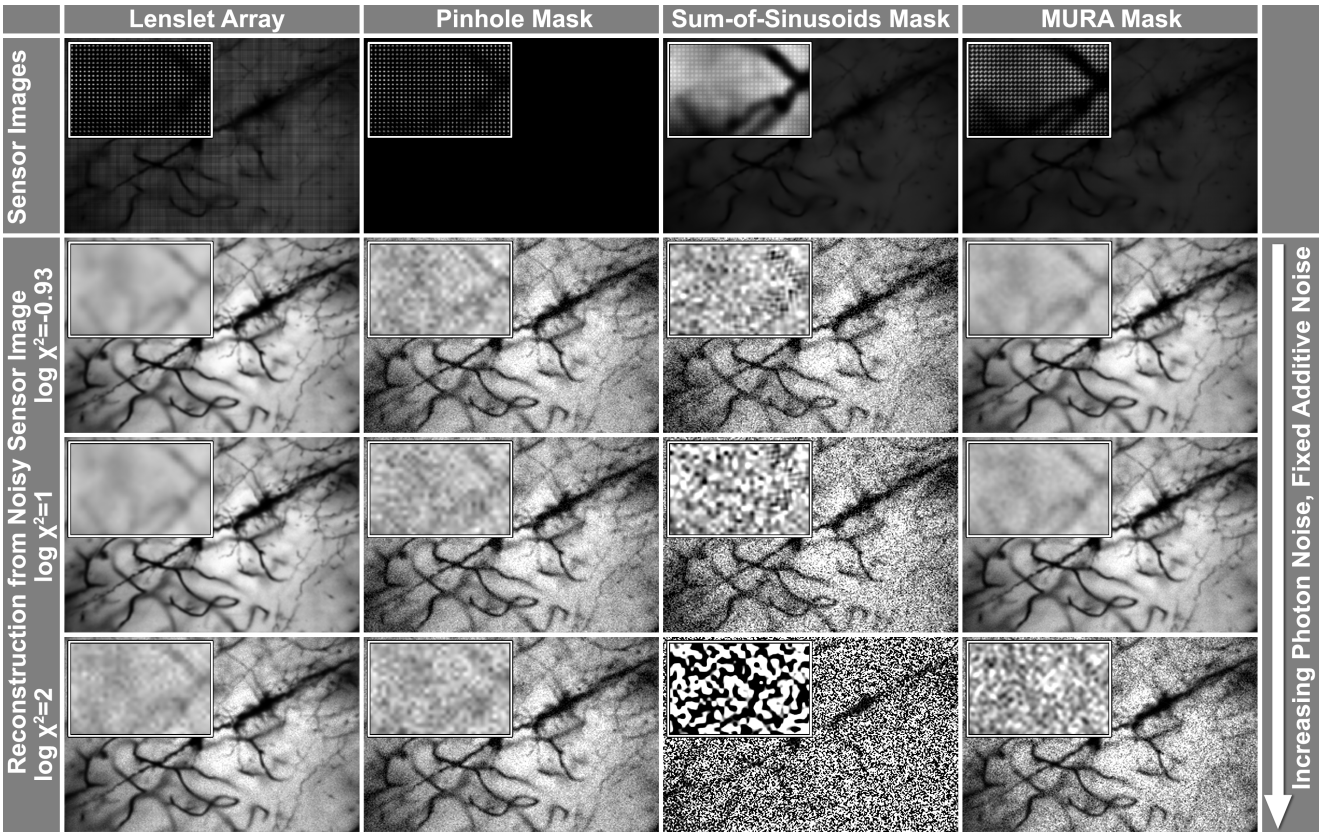


Fig. 12 Comparison of noise amplification for different light field acquisition schemes on the golgi stained neuron dataset (lightfield.stanford.edu). Row 1 shows simulated sensor images with contrast enhanced close-ups. The other rows show a single view of the reconstructed light field from a noisy sensor image. The ratio χ^2 of signal-dependent photon noise and signal-independent dark current noise varies for the different reconstructions. Row 2 simulates a reconstruction with a dominating additive noise term, while rows 3 and 4 show the effect of an increasingly dominating photon noise term in the sensor images.

7.1 Benefits and Limitations

The proposed framework generalizes multiplexing systems where the underlying signal is sampled in a super-pixel-periodic fashion. While this is the most common approach for color image and light field acquisition, several methods that sample the plenoptic function in a completely random manner have been proposed (Veeraraghavan et al 2011; Reddy et al 2011). Due to the lack of a regular sampling structure, these specific approaches are not supported by our framework. However, we envision multiplexing approaches that combine random plenoptic projections with super-pixel-periodic spatial sampling patterns to be an exciting avenue of future research. Our image formation unifies a wide range of previously proposed multiplexing schemes and paves the way for novel multiplexing techniques. We generalize the analysis and reconstruction to either the spatial or the Fourier domain. Practically, this allows for higher quality reconstructions of Fourier multiplexed data and the formulation of optimality criteria of employed optical modulators.

We do not propose new optical multiplexing methods, but evaluate and unify a variety of existing ap-

proaches. The theory presented in this paper allows knowledge of color demosaicing, which has been built up for decades within the computer vision community, to be transferred to the reconstruction of light fields and other dimensions of the plenoptic function.

7.2 Future Work

As mentioned above, we would like to explore strategies that sample the plenoptic function in a random but super-pixel-periodic fashion. A combination of compressive sensing paradigms and traditional, periodic sampling approaches could prove essential in the quest for plenoptic resolution improvements beyond the Nyquist limit. Application-specific reconstruction filters, exploiting natural image statistics, could further push the boundaries of conventional image acquisition. The exploitation of natural image statistics is common practice for imaging with color filter arrays and subsequent demosaicing. However, there is significant potential to develop similar techniques for demosaicing other multiplexed plenoptic information, for instance light fields (Levin and Durand 2010).

7.3 Conclusion

The unifying theory presented in this paper is a crucial step toward the "ultimate" camera capturing all visual information with a single shot. Only within the last few years has the research community started to investigate approaches to acquire the plenoptic function with joint optical modulation and computational processing. Our work ties many of these new techniques to more traditional ways of sampling visual information. The proposed framework is essential for the evaluation and optimization of plenoptic multiplexing schemes of the future.

Appendix A: Proof of PSM Theorem

Throughout this paper, we assume that the spatial basis functions $\boldsymbol{\sigma}_j(\vec{x})$ are super-pixel-periodic[2], i.e. $\boldsymbol{\sigma}_j(\vec{x}) = \boldsymbol{\sigma}_j(\vec{x} + t\Delta\vec{x}_s), \forall t \in \mathbb{Z}$. The offset between successive super-pixels is denoted as $\Delta\vec{x}_s$. This notation allows us to define a sampling operator as

$$\text{Ш}_k(\vec{x}) = \sum_{t\in\mathbb{Z}} \delta\left(\vec{x} + t\Delta\vec{x}_s + \Delta\vec{x}_k\right), \tag{23}$$

where $\Delta\vec{x}_k$ is the offset of individual samples within each super-pixel. This sampling operator basically extracts a sub-image or channel $k = 1 \dots M$ from an interleaved sensor image, where only corresponding sub-pixels within the super-pixels are included in each channel. Sampling such a channel $\widetilde{\boldsymbol{i}}_k(\vec{x})$ from a sensor image $\boldsymbol{i}(\vec{x})$, in combination with Equation 4, results in the following expression

$$\begin{aligned}\widetilde{\boldsymbol{i}}_k(\vec{x}) &= \text{Ш}_k(\vec{x})\boldsymbol{i}(\vec{x}) = \text{Ш}_k(\vec{x})\left(\sum_{j=1}^{N}\boldsymbol{\sigma}_j(\vec{x})\boldsymbol{\rho}_j\left(\vec{x}\right)\right)\\ &= \sum_{j=1}^{N}\boldsymbol{\sigma}_j^k\text{Ш}_k(\vec{x})\boldsymbol{\rho}_j\left(\vec{x}\right).\end{aligned} \tag{24}$$

As discussed in Section 3.3, the spatial basis functions $\boldsymbol{\sigma}_j(\vec{x})$ become spatially-invariant constants $\boldsymbol{\sigma}_j^k$ in the sampled channels because of spatial periodicity of the basis. Each channel $\tilde{\boldsymbol{i}}_k(\vec{x})$ is thus associated with N constants $\boldsymbol{\sigma}_j^k$ defined by the spatial basis. Reconstructing a channel $\boldsymbol{i}_k(\vec{x})$ from its sampled representation $\tilde{\boldsymbol{i}}_k(\vec{x})$ is performed by convolving with a reconstruction filter kernel $\boldsymbol{f}(\vec{x})$:

$$\boldsymbol{i}_k(\vec{x}) = \widetilde{\boldsymbol{i}}_k(\vec{x}) \otimes \boldsymbol{f}(\vec{x}). \tag{25}$$

According to the sampling theorem, the original signal must be spatially band-limited for this reconstruction to be a faithful representation.

Due to the spatial basis being a set of constants at every position in the interpolated channels $\boldsymbol{i}_k(\vec{x})$, and the plenoptic basis being spatially-invariant, a convolution with the filter kernel can be formulated as

$$\begin{aligned}\boldsymbol{i}_k(\vec{x}) &= \left(\sum_{j=1}^{N}\boldsymbol{\sigma}_j^k\text{Ш}_k(\vec{x})\boldsymbol{\rho}_j\left(\vec{x}\right)\right) \otimes \boldsymbol{f}(\vec{x})\\ &= \sum_{j=1}^{N}\boldsymbol{\sigma}_j^k\int_{\mathscr{P}}\boldsymbol{\pi}_j(\vec{p})\left(\text{Ш}_k(\vec{x})\boldsymbol{l}_\lambda(\vec{x},\vec{p}) \otimes \boldsymbol{f}(\vec{x})\right)d\vec{p}\end{aligned} \tag{26}$$

Equation 26 shows that all channels $\boldsymbol{i}_k(\vec{x})$ are locally related to the sampled and reconstructed plenoptic function via a linear combination. It also shows that applying a linear filter to the measured channels, i.e. before decorrelation, is equivalent to applying the same filter to the plenoptic function itself, i.e. after decorrelation. □

[2] Following Sec. 3.4, we omit periodicity in the second spatial dimension in our notation of t for clarity.

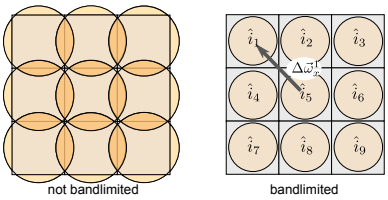


Fig. 13 Information overlaps in the Fourier domain if the signal is not suitably band-limited (left). With the appropriate band-limitation, the Fourier representation decomposes into distinct correlated Fourier tiles $\hat{\boldsymbol{i}}_k$. The arrow in the right figure indicates the position $\Delta\vec{\omega}_x^k$ of the Dirac peaks, i.e. the center frequencies, of a Fourier tile.

Appendix B: Proof of PFM Theorem

The proof of Theorem 2 follows Equations 6-8. To provide additional detail, we start with Equation 8

$$\mathcal{F}_x\left\{\boldsymbol{i}(\vec{x})\right\} = \sum_{k=1}^{M}\left(\delta\left(\vec{\omega}_x - \Delta\vec{\omega}_x^k\right) \otimes \sum_{j=1}^{N} \hat{\boldsymbol{\sigma}}_j^k \hat{\boldsymbol{\rho}}_j\left(\vec{\omega}_x\right)\right), \quad (27)$$

where $k\Delta\vec{\omega}_x$ has been replaced by $\Delta\vec{\omega}_x^k$, a vector to the center frequency of a Fourier tile (Fig. 13, right), to allow for generalized sampling patterns. If the signal is properly band-limited the image's Fourier transform separates into disjoint sets, each encoding one Fourier tile $\hat{\boldsymbol{i}}_k(\vec{\omega}_x')$ (Fig. 13, right). To separate the notation for spatial frequencies in the sensor image and for Fourier tiles we use the substitution $\vec{\omega}_x' = \vec{\omega}_x - \Delta\vec{\omega}_x^k$.

The Fourier tiles can be cropped from the Fourier transformed sensor image by applying a rect filter[3]:

$$\begin{aligned}\hat{\boldsymbol{i}}_k(\vec{\omega}_x') &= \mathrm{rect}^k(\vec{\omega}_x')\delta\left(\vec{\omega}_x'\right) \otimes \sum_{j=1}^{N} \hat{\boldsymbol{\sigma}}_j^k \hat{\boldsymbol{\rho}}_j\left(\vec{\omega}_x'\right) \quad (28)\\ &= \mathrm{rect}^k(\vec{\omega}_x') \sum_{j=1}^{N} \hat{\boldsymbol{\sigma}}_j^k \int_{\mathscr{P}} \boldsymbol{\pi}(\vec{p})\hat{\boldsymbol{l}}_\lambda(\vec{\omega}_x', \vec{p})d\vec{p}.\end{aligned}$$

The convolution with a Dirac train in Equation 8 reduces to a convolution with a single Dirac peak $\delta\left(\vec{\omega}_x'\right)$ because of band-limitation. This is a unit operation and thus removed from the equations. □

In addition, by inverse Fourier transforming and sampling the Fourier tile $\hat{\boldsymbol{i}}_k(\vec{\omega}_x')$ we see that the reconstruction filter is indeed a sinc:

$$\boldsymbol{i}_k(\vec{x}) = \sum_{j=1}^{N} \boldsymbol{\sigma}_j^k \int_{\mathscr{P}} \boldsymbol{\pi}_j(\vec{p})\left(\mathrm{Ш}_k(\vec{x})\boldsymbol{l}_\lambda(\vec{x}, \vec{p}) \otimes \mathrm{sinc}(\vec{x})\right)d\vec{p} \quad (29)$$

[3] For other sampling patterns, corresponding, e.g. hexagonal, masking functions would be used. In addition, apodization functions can be used to reduce ringing artifacts in the reconstruction at the expense of decreasing the effective resolution, see e.g. (Veeraraghavan et al 2007). This does not affect the proof.

References

Adelson E, Wang J (1992) Single Lens Stereo with a Plenoptic Camera. IEEE Trans PAMI 14(2):99–106

Adelson EH, Bergen JR (1991) The Plenoptic Function and the Elements of Early Vision. In: Computational Models of Visual Processing, MIT Press, pp 3–20

Aggarwal M, Ahuja N (2004) Split Aperture Imaging for High Dynamic Range . Int J Comp Vis 58(1):7–17

Agrawal A, Gupta M, Veeraraghavan A, Narasimhan S (2010a) Optimal Coded Sampling for Temporal Super-Resolution. In: Proc. IEEE CVPR, pp 374–380

Agrawal A, Veeraraghavan A, Raskar R (2010b) Reinterpretable Imager: Towards Variable Post-Capture Space, Angle and Time Resolution in Photography. In: Proc. Eurographics, pp 1–10

Alleyson D, Süsstrunk S, Hérault J (2005) Linear Demosaicing inspired by the Human Visual System. IEEE Trans Im Proc 14(4):439–449

Bando Y, Chen BY, Nishita T (2008) Extracting depth and matte using a color-filtered aperture. ACM Trans Graph (Siggraph Asia) 27(5):134

Bayer BE (1976) Color imaging array. US Patent 3,971,065

Ben-Ezra M, Zomet A, Nayar S (2005) Video Superresolution using Controlled Subpixel Detector Shifts. IEEE Trans PAMI 27(6):977–987

Bishop T, Zanetti S, Favaro P (2009) Light-Field Superresolution. In: Proc. ICCP, pp 1–9

Bub G, Tecza M, Helmes M, Lee P, Kohl P (2010) Temporal Pixel Multiplexing for Simultaneous High-Speed, High-Resolution Imaging. Nature Methods 7:209–211

Compton J (2007) Color filter array 2.0. `http://johncompton.pluggedin.kodak.com`

Debevec PE, Malik J (1997) Recovering High Dynamic Range Radiance Maps from Photographs. In: Proc. ACM Siggraph, pp 369–378

Georgiev T, Lumsdaine A (2010) Rich Imamge Capture with Plenoptic Cameras. In: Proc. ICCP, pp 1–8

Georgiev T, Intwala C, Babacan S, Lumsdaine A (2008) Unified Frequency Domain Analysis of Lightfield Cameras. In: Proc. ECCV, pp 224–237

Gortler S, Grzeszczuk R, Szelinski R, Cohen M (1996) The Lumigraph. In: Proc. ACM Siggraph, pp 43–54

Gottesman SR, Fenimore EE (1989) New Family of Binary Arrays for Coded Aperture Imaging. Applied Optics 28(20):4344–4352

Greivenkamp J (1990) Color Dependant Optical Prefilter for the Suppression of Aliasing Artifacts. Applied Optics 29(5):676–684

Gupta M, Agrawal A, Veeraraghavan A, Narasimhan SG (2010) Flexible Voxels for Motion-Aware Videography. In: Proc. ECCV, pp 100–114

Horstmeyer R, Euliss G, Athale R, Levoy M (2009) Flexible Multimodal Camera Using a Light Field Architecture. In: Proc. ICCP, pp 1–8

Ives FE (1903) Parallax stereogram and process of making same. U.S. Patent 725,567

Lanman D, Raskar R, Agrawal A, Taubin G (2008) Shield Fields: Modeling and Capturing 3D Occluders. ACM Trans Graph (Siggraph Asia) 27(5):131

Levin A, Durand F (2010) Linear View Synthesis Using a Dimensionality Gap Light Field Prior. In: Proc. IEEE CVPR, pp 1–8

Levin A, Freeman WT, Durand F (2008) Understanding Camera Trade-Offs through a Bayesian Analysis of Light Field Projections. In: Proc. ECCV, pp 88–101

Levin A, Hasinoff SW, Green P, Durand F, Freeman WT (2009) 4D Frequency Analysis of Computational Cameras for Depth of Field Extension. ACM Trans Graph (Siggraph) 28(3):97

Levoy M, Hanrahan P (1996) Light Field Rendering. In: Proc. ACM Siggraph, pp 31–42

Levoy M, Ng R, Adams A, Footer M, Horowitz M (2006) Light Field Microscopy. ACM Trans Graph (Siggraph) 25(3):924–934

Li X, Gunturk B, Zhang L (2008) Image Demosaicing: a Systematic Survey. In: SPIE Conf. on Visual Comm. and Image Proc., pp 68,221J–68,221J–15

Lippmann G (1908) La Photographie Intégrale. Academie des Sciences 146:446–451

Lumsdaine A, Georgiev T (2009) The Focused Plenoptic Camera. In: Proc. ICCP, pp 1–8

McGuire M, Matusik W, Pfister H, Chen B, Hughes JF, Nayar SK (2007) Optical Splitting Trees for High-Precision Monocular Imaging. IEEE Comput Graph & Appl 27(2):32–42

Mitsunaga T, Nayar SK (1999) Radiometric Self Calibration. In: Proc. IEEE CVPR, pp 374–380

Narasimhan S, Nayar S (2005) Enhancing Resolution along Multiple Imaging Dimensions using Assorted Pixels. IEEE Trans PAMI 27(4):518–530

Ng R (2005) Fourier Slice Photography. ACM Trans Graph (Siggraph) 24(3):735–744

Reddy D, Veeraraghavan A, Chellappa R (2011) P2C2: Programmable Pixel Compressive Camera for High Speed Imaging. In: Proc. IEEE CVPR, pp 1–8

Sajadi B, Majumder A, Hiwada K, Maki A, Raskar R (2011) Switchable Primaries Using Shiftable Layers of Color Filter Arrays. ACM Trans Graph (Siggraph) 30(3):1–8

Schechner Y, Nayar S (2003) Generalized Mosaicing: High Dynamic Range in a Wide Field of View. IJCV 53(3):245–267

Schechner Y, Nayar S, Belhumeur P (2007) Multiplexing for Optimal Lighting. IEEE Trans PAMI 29(8):1339–1354

Takhar D, Laska JN, Wkin M, Duarte MF, Baron D, Sarvotham S, Kelly KF, Baraniuk RG (2006) A New Compressive Imaging Camera Architecture using Optical-Domain Compression. IS&T/SPIE Computational Imaging IV 6065

Veeraraghavan A, Raskar R, Agrawal A, Mohan A, Tumblin J (2007) Dappled Photography: Mask Enhanced Cameras for Heterodyned Light Fields and Coded Aperture Refocussing. ACM Trans Graph (Siggraph) 26(3):69

Veeraraghavan A, Raskar R, Agrawal A, Chellappa R, Mohan A, Tumblin J (2008) Non-Refractive Modulators for Encoding and Capturing Scene Appearance and Depth. In: Proc. IEEE CVPR, pp 1–8

Veeraraghavan A, Reddy D, Raskar R (2011) Coded Strobing Photography: Compressive Sensing of High Speed Periodic Videos. IEEE Trans PAMI 33(4):671–686

Wang S, Heidrich W (2004) The Design of an Inexpensive Very High Resolution Scan Camera System. Computer Graphics Forum (Eurographics) 23(10):441–450

Wenger A, Gardner A, Tchou C, Unger J, Hawkins T, Debevec P (2005) Performance Relighting and Reflectance Transformation with Time-Multiplexed Illumination. ACM Trans Graph (Siggraph) 24(3):756–764

Wetzstein G, Ihrke I, Heidrich W (2010) Sensor Saturation in Fourier Multiplexed Imaging. In: Proc. IEEE CVPR, pp 1–8

Wetzstein G, Ihrke I, Lanman D, Heidrich W (2011) State of the Art in Computational Plenoptic Imaging. In: Proc. Eurographics (STAR), pp 1–24

Wilburn B, Joshi N, Vaish V, Talvala EV, Antunez E, Barth A, Adams A, Horowitz M, Levoy M (2005) High Performance Imaging using Large Camera Arrays. ACM Trans Graph (Siggraph) 24(3):765–776

Wuttig A (2005) Optimal Transformations for Optical Multiplex Measurements in the Presence of Photon Noise. Appl Opt 44(14):2710–2719

Yasuma F, Mitsunaga T, Iso D, Nayar SK (2010) Generalized Assorted Pixel Camera: Post-Capture Control of Resolution, Dynamic Range and Spectrum. IEEE Trans Im Proc 99