



Universiteit
Leiden
The Netherlands

Fourier-domain filtering analysis for color-polarization camera demosaicking

Hagen, N.; Stockmans, T.A.; Otani, Y.; Buranasiri, P.

Citation

Hagen, N., Stockmans, T. A., Otani, Y., & Buranasiri, P. (2024). Fourier-domain filtering analysis for color-polarization camera demosaicking. *Applied Optics*, 63(9), 2314-2323. doi:10.1364/AO.516696

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/4210919>

Note: To cite this publication please use the final published version (if applicable).



Fourier-domain filtering analysis for color-polarization camera demosaicking

NATHAN HAGEN,^{1,*}  THIJS STOCKMANS,²  YUKITOSHI OTANI,¹ 
AND PRATHAN BURANASIRI^{3,4} 

¹Utsunomiya University, Department of Optical Engineering, 7-2-1 Yoto, Utsunomiya, Tochigi 321-8585, Japan

²Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands

³Artificial Intelligence Photonics Advanced Research Laboratory, Department of Physics, School of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand

⁴Electronic and Optoelectronic Device Research Unit, School of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand

*nh@hagenlab.org

Received 28 December 2023; revised 20 February 2024; accepted 24 February 2024; posted 26 February 2024; published 18 March 2024

We review Fourier-domain methods for demosaicking Bayer-filter color cameras and monochrome polarization cameras, and then generalize the approach for the quad-Bayer-filter mosaic and for color-polarization cameras. For each of these four mosaic filter types, we provide theoretical expressions for the sampling functions, the Fourier-domain channels, and the linear combination of reconstructed channels (the demosaicking algorithm) needed to estimate the input (presampled) image. A useful advantage of the Fourier-domain approach is that it provides a direct means of visualizing and quantifying when aliasing is likely or unlikely to be present. For the Bayer and quad-Bayer-filter types, we provide simulated images, while for the polarization camera types we provide experimental images and videos to illustrate the algorithm and analyze crosstalk error. © 2024 Optica Publishing Group

<https://doi.org/10.1364/AO.516696>

1. INTRODUCTION

Mosaic filter arrays are repeated patterns of spectral and polarization filters placed on detector array pixels. In recent years, increasing numbers of these mosaic filters have become commercially available. While the most common pattern available remains the Bayer-filter pattern used for RGB color cameras, micropolarizer filter arrays [1,2], combined color and polarization filters [3], and multispectral filter arrays [4,5] have also become available. Because the mosaic approach uses filters that are not coregistered to one another, spatial variation of intensity in the image can be misinterpreted as spatial variation in color, polarization, or spectrum. Thus, any image containing sharp spatial features will appear to have an unusual coloration, polarization, or spectrum along image feature edges.

However, if we carefully design the demosaicking algorithm by which we estimate the spatially registered color, polarization, or spectral data from the mosaic-sampled image, then we can minimize these errors [6]. The Fourier-domain approach to demosaicking of mosaic-filter-array images was first developed by Dubois in 2005 for Bayer-filter RGB cameras [7]. The basic idea is that we can interpret the color sampling grid as a multiplication of the original image by a sinusoidal function. This multiplication causes the original image to be replicated in the Fourier domain, with the original image data centered at the baseband (zero) frequency, and also two copies centered

at the positive and negative frequencies corresponding to the frequency of the sampling function. The advantage of this approach to analyzing the demosaicking problem is that we can see directly how any overlap in the Fourier-domain data can cause crosstalk between the different channels. For a Bayer-filter pattern, the Fourier-domain channel centered at baseband contains the luminosity information of the image, while the sideband channels contain the chromaticity information. Overlap between these Fourier-domain channels will cause errors in the intensity and in the color, which we see as artificial coloration of edge features in the image.

In general, the highest-performance demosaicking methods used today do not use a Fourier-domain approach, but rather use nonlinear algorithms that take advantage of localized spatial correlations in typical images [8,9]. However, the Fourier-domain approach allows us an easy means to visualize mosaic-filter images and assess how prone a given image is to crosstalk (often described also as aliasing error). One can, for example, use this as a real-time tool to find the optimal defocus for the camera lens, such that we retain as much image resolution as possible while also minimizing crosstalk error. We can also note that researchers have explored Fourier-domain analysis for the purpose of developing new mosaic filter patterns [10–12].

In the discussion below, we start with a review of the Fourier-domain approach to analyzing and demosaicking Bayer RGB filter array images and monochromatic polarization filter array

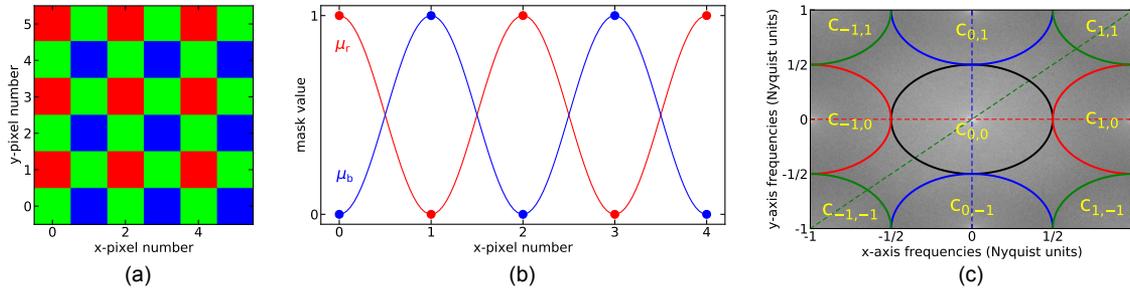


Fig. 1. (a) Bayer-filter mosaic, showing the 6×6 subregion near the image origin. (b) Sinusoidal patterns, with their discrete sampling points, indicating the locations of the red and blue pixels in the filter mosaic. The blue curve is the sampling function for row 0, while the red curve shows the pattern for row 1. (c) Fourier-domain representation of the image after sampling with a Bayer-filter mosaic, with channel labels C_{ij} . The dashed lines indicate horizontal, vertical, and diagonal cross sections through the data. Note that the channel circles' colors do not directly correspond to RGB color, but rather to WUV luminance–chrominance.

images. Using the same basic approach, we show how to adapt this analysis to the case of a Bayer RGB filter pattern in which the individual filters cover not individual detector pixels but 2×2 pixel groups—the “quad-Bayer-filter array” pattern. This is a natural step towards analyzing the case of color-polarization images, which combines single-pixel polarization filters with this quad-Bayer pattern. In each case, we give expressions for the sampling functions, the channel distributions in the Fourier domain, and Fourier-based demosaicking algorithms. Finally, we analyze the crosstalk in an example image, illustrate the horizontal- and vertical-axis bias in the crosstalk signal, and indicate how this affects algorithm and filter design.

2. BAYER-FILTER RGB MOSAIC

The Bayer-filter mosaic consists of a 2×2 set of red, green, and blue color filters replicated across the face of a detector array (Fig. 1). If we write the original image (before filtering) as $I(x, y, \lambda)$, then the red, green, and blue filtered images will be

$$\left. \begin{aligned} r(x, y) &= \int I(x, y, \lambda) \tau_r(\lambda) \mu_r(x, y) d\lambda \\ g(x, y) &= \int I(x, y, \lambda) \tau_g(\lambda) \mu_g(x, y) d\lambda \\ b(x, y) &= \int I(x, y, \lambda) \tau_b(\lambda) \mu_b(x, y) d\lambda \end{aligned} \right\}, \quad (1)$$

where λ is the wavelength, $\tau(\lambda)$ is the color filter efficiency spectrum (filter transmittance and detector quantum efficiency), and $\mu(x, y)$ is the filter sampling (or modulation) function:

$$\left. \begin{aligned} \mu_r &= \frac{1}{4} [1 - \cos(\pi x)] [1 + \cos(\pi y)] \\ \mu_g &= \frac{1}{2} [1 + \cos(\pi x) \cos(\pi y)] \\ \mu_b &= \frac{1}{4} [1 + \cos(\pi x)] [1 - \cos(\pi y)] \end{aligned} \right\}. \quad (2)$$

Two views of μ_r and μ_b are shown in Fig. 1. If these functions are sampled at discrete locations $x = \{0, 1, 2, \dots\}$ and $y = \{0, 1, 2, \dots\}$, then the resulting pattern becomes binary, as shown in Fig. 2.

For Eq. (2), the green filter is placed at the image origin $(x, y) = (0, 0)$, and the image origin is located at the bottom left of the image. It is also common to locate the image origin at the upper left corner, and so showing the explicit origin location is important for clarity. Finally, some Bayer patterns place the red filter at the image origin, and this changes the sampling functions to the alternate representation

$$\left. \begin{aligned} \mu_r(x, y) &= \frac{1}{4} [1 + \cos(\pi x)] [1 + \cos(\pi y)] \\ \mu_g(x, y) &= \frac{1}{2} [1 - \cos(\pi x) \cos(\pi y)] \\ \mu_b(x, y) &= \frac{1}{4} [1 - \cos(\pi x)] [1 - \cos(\pi y)] \end{aligned} \right\}. \quad (3)$$

In all of the simulations and experimentally obtained images used in this article, we will be using expressions corresponding to the green filter origin Eq. (2).

The raw image data obtained from the detector array provides a single digital value at each pixel location. If we take the integrals over λ at each pixel Eq. (1), we can represent the raw image using the r , g , and b color values at each pixel:

$$\begin{aligned} I_{\text{raw}}(x, y) &= r(x, y) \mu_r(x, y) + g(x, y) \mu_g(x, y) \\ &\quad + b(x, y) \mu_b(x, y). \end{aligned} \quad (4)$$

The demosaicking problem is to reconstruct $I(x, y, c)$ for color set $c \in \{r, g, b\}$ in a way that least distorts the underlying data. Figure 3 shows an example color image of a spider, prior to sampling by the filter mosaic. The fine hairs on the spider's body are high-spatial-frequency features in the image that help to make color crosstalk errors visible. Figure 3(b) shows a closeup view of the spider's hairs in the image before mosaic sampling.

Taking the Fourier transform of the mosaic-filtered image gives the distribution shown in Fig. 1(c). Due to the high-spatial-frequency information of these features, cross-sectional views through the Fourier-domain data [Fig. 3(c)] show that the channels are not clearly separated. Clearly, there is crosstalk occurring between the central luminance channel and the chrominance sideband channels, as we can see in the closeup [Fig. 3(d)]. However, if we blur the image prior to sampling (such as by defocusing the camera lens), then we obtain a

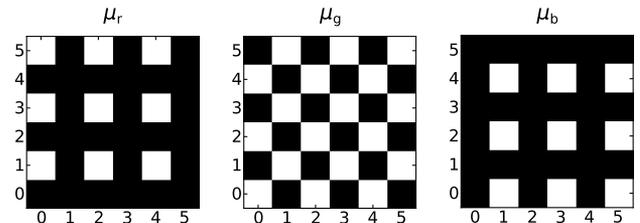


Fig. 2. Sampling functions μ_r , μ_g , and μ_b Eq. (2) for the Bayer-filter mosaic pixels 0 through 5, for the case of the green filter at $(0, 0)$. Black indicates a value of 0, white a value of 1.

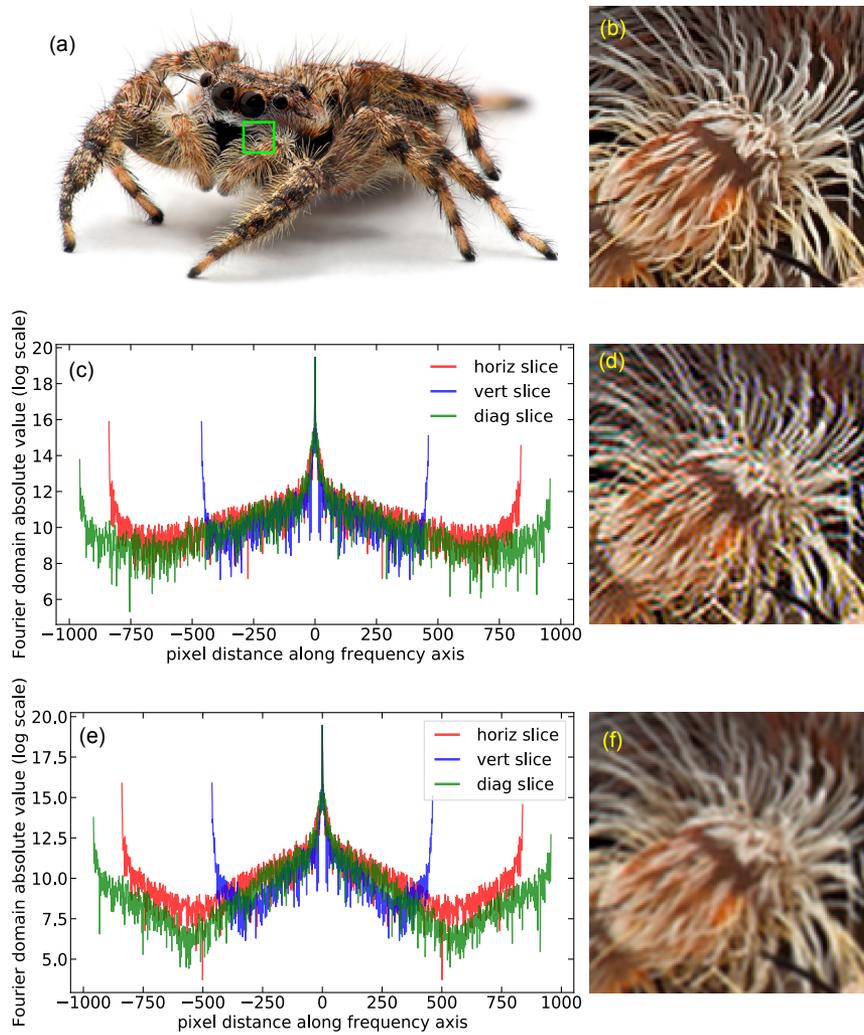


Fig. 3. (a) Example color image and (b) closeup. (c) Cross sections through the image (a), after sampling with a Bayer-filter pattern. The display shows the absolute value of the Fourier-domain data, in logarithmic scale. A second closeup (d) showing the same region after mosaic-sampling and reconstruction. Color artifacts on the thin hairs of the spider appear as a result of color crosstalk. (e) After blurring is applied to the example, the Fourier domain shows separation between channels (deep dips between peaks) and (f) a closeup of the reconstructed blurred image. Because the blurring has separated the channels, the color crosstalk is no longer visible. (Image from Ref. [13].)

Fourier-domain distribution that shows well-separated channels [Fig. 3(e)]. As a result, the color artifacts largely disappear [Fig. 3(f)].

In order to express the image data distribution in the Fourier domain, we take the Fourier transform of the raw sampled image Eq. (4), which gives the lengthy expression

$$\begin{aligned}
 \mathcal{F}\{I_{\text{raw}}(x, y)\} = & R(\xi, \eta) \left[\frac{1}{4} \delta(\xi, \eta) - \frac{1}{8} \delta(\xi - 1, \eta) - \frac{1}{8} \delta(\xi + 1, \eta) + \frac{1}{8} \delta(\xi, \eta - 1) + \frac{1}{8} \delta(\xi, \eta + 1) - \frac{1}{16} \delta(\xi - 1, \eta - 1) \right. \\
 & - \frac{1}{16} \delta(\xi + 1, \eta + 1) - \frac{1}{16} \delta(\xi - 1, \eta + 1) - \frac{1}{16} \delta(\xi + 1, \eta - 1) \left. \right] + G(\xi, \eta) \left[\frac{1}{2} \delta(\xi, \eta) + \frac{1}{8} \delta(\xi - 1, \eta - 1) \right. \\
 & + \frac{1}{8} \delta(\xi + 1, \eta + 1) + \frac{1}{8} \delta(\xi - 1, \eta + 1) + \frac{1}{8} \delta(\xi + 1, \eta - 1) \left. \right] + B(\xi, \eta) \left[\frac{1}{4} \delta(\xi, \eta) + \frac{1}{8} \delta(\xi - 1, \eta) + \frac{1}{8} \delta(\xi + 1, \eta) \right. \\
 & - \frac{1}{8} \delta(\xi, \eta - 1) - \frac{1}{8} \delta(\xi, \eta + 1) - \frac{1}{16} \delta(\xi - 1, \eta - 1) - \frac{1}{16} \delta(\xi + 1, \eta + 1) - \frac{1}{16} \delta(\xi - 1, \eta + 1) - \frac{1}{16} \delta(\xi + 1, \eta - 1) \left. \right].
 \end{aligned} \tag{5}$$

Here (ξ, η) are the Fourier-domain coordinates corresponding to the spatial domain (x, y) , with ξ and η given here in Nyquist

frequency units, so that $+1$ and -1 represent the maximum and minimum frequencies obtained by the discrete Fourier transform of the image. In Eq. (5), we can see that the coefficients multiplying the baseband channel $\delta(\xi, \eta)$ for each color correspond to the average transmittance—since only one in four

pixels transmits red, the R color coefficient has $1/4$, half of all pixels transmit green, and so the green coefficient is $1/2$, etc.

If we take advantage of symmetries in the data, we can define the luminance W and chrominances U and V as

$$\left. \begin{aligned} W &= \frac{1}{4}R + \frac{1}{2}G + \frac{1}{4}B, & R &= W + U + 2V \\ U &= \frac{1}{4}R - \frac{1}{2}G + \frac{1}{4}B, & G &= W - U \\ V &= \frac{1}{4}R - \frac{1}{4}B, & B &= W + U - 2V \end{aligned} \right\}. \quad (6)$$

(Although the symbol Y is often used for the luminance in the literature, we use W instead here in order to avoid confusion with the coordinate y .) Using these instead of the RGB colors directly, we obtain the simpler form

$$\begin{aligned} \mathcal{F}\{I_{\text{raw}}(x, y)\} &= W(\xi, \eta)\delta(\xi, \eta) \\ &\quad - \frac{1}{4}U(\xi, \eta)[\delta(\xi - 1, \eta - 1) + \delta(\xi + 1, \eta + 1) \\ &\quad + \delta(\xi - 1, \eta + 1) + \delta(\xi + 1, \eta - 1)] \\ &\quad + \frac{1}{4}V(\xi, \eta)[\delta(\xi, \eta - 1) + \delta(\xi, \eta + 1) \\ &\quad - \delta(\xi - 1, \eta) - \delta(\xi + 1, \eta)]. \end{aligned} \quad (7)$$

Each of the Dirac delta functions here can be considered as the center location for a Fourier-domain “channel.” If we number the channels $C_{i,j}$ by their frequencies, then the channel representation of the Fourier-domain data can be written as

$$\mathcal{F}\{I_{\text{raw}}(x, y)\} = \sum_{i=(-1,0,1)} \sum_{j=(-1,0,1)} C_{i,j}(\xi, \eta)$$

for the nine channels

$$\left. \begin{aligned} C_{0,0}(\xi, \eta) &= W(\xi, \eta) \\ C_{\pm 1,0}(\xi, \eta) &= -\frac{1}{4}V \\ C_{0,\pm 1}(\xi, \eta) &= \frac{1}{4}V(\xi, \eta) \\ C_{\pm 1,\pm 1}(\xi, \eta) &= -\frac{1}{4}U(\xi, \eta) \end{aligned} \right\}. \quad (8)$$

These are the channels drawn in Fig. 1(c). As implied by the figure, however, the $+1$ and -1 Nyquist frequencies are actually the same thing, so that they should be added together as one single channel. This consolidates the above set of nine channels into just four:

$$\left. \begin{aligned} C_{0,0}(\xi, \eta) &= W(\xi, \eta) \\ C_{1,0}(\xi, \eta) &= -\frac{1}{2}V(\xi, \eta) \\ C_{0,1}(\xi, \eta) &= \frac{1}{2}V(\xi, \eta) \\ C_{1,1}(\xi, \eta) &= -U(\xi, \eta) \end{aligned} \right\}. \quad (9)$$

The four channels of Eq. (9) are illustrated in Fig. 1(c) by using four colors to represent each consolidated channel.

Because of the Hermitian symmetry of the Fourier transform, a channel lying symmetric to its partner channel along a line passing through the origin will have the same content but conjugated, e.g.,

$$C_{+1,+1}(\xi, \eta) = [C_{-1,-1}(\xi, \eta)]^*. \quad (10)$$

Throughout the paper, we will use $C(\xi, \eta)$ for channels in the Fourier domain and $c(x, y)$ for channels that have been shifted to the baseband and inverse Fourier-transformed. It may appear that Hermitian symmetry does not agree with Eq. (8). However, all of the noncentral channels in Eq. (8) straddle the Nyquist boundary, which means that $+1$ and -1 channels are colocated, and thus summed together. Due to the Hermitian symmetry,

this sum means that the imaginary component of the channel will cancel, leaving only the real-valued component.

In order to reconstruct the image color information, we can take the inverse Fourier transform of each channel C_{ij} individually and calculate weighted sums. Because we have taken advantage of the data symmetry, reconstructing the WUV luminance–chrominance images is simple. Writing the inverse transform channels as $c_{ij}(x, y) = \mathcal{F}^{-1}\{C_{ij}(\xi, \eta)\}$, we obtain the reconstruction algorithm for the luminance–chrominance images:

$$\left. \begin{aligned} \hat{w}(x, y) &= c_{00}(x, y) \\ \hat{u}(x, y) &= c_{11}(x, y) \\ \hat{v}(x, y) &= \frac{1}{2}[c_{01}(x, y) - c_{10}(x, y)] \end{aligned} \right\}. \quad (11)$$

Obtaining the RGB colors from the luminance–chrominance images is likewise straightforward: [14]

$$\left. \begin{aligned} \hat{r}(x, y) &= \hat{w} + \hat{u} + 2\hat{v} \\ \hat{g}(x, y) &= \hat{w} - \hat{u} \\ \hat{b}(x, y) &= \hat{w} + \hat{u} - 2\hat{v} \end{aligned} \right\}. \quad (12)$$

Substituting Eq. (11) into Eq. (12) gives the direct RGB color reconstruction for the case of having a green pixel at the origin:

$$\left. \begin{aligned} \hat{r}(x, y) &= c_{00}(x, y) - c_{11}(x, y) - c_{01}(x, y) + c_{10}(x, y) \\ \hat{g}(x, y) &= c_{00}(x, y) + c_{11}(x, y) \\ \hat{b}(x, y) &= c_{00}(x, y) - c_{11}(x, y) + c_{10}(x, y) - c_{01}(x, y) \end{aligned} \right\} \quad (13)$$

and, for the case of a red pixel at the origin,

$$\left. \begin{aligned} \hat{r}(x, y) &= c_{00}(x, y) + c_{11}(x, y) + c_{01}(x, y) - c_{10}(x, y) \\ \hat{g}(x, y) &= c_{00}(x, y) - c_{11}(x, y) \\ \hat{b}(x, y) &= c_{00}(x, y) + c_{11}(x, y) - c_{10}(x, y) + c_{01}(x, y) \end{aligned} \right\}. \quad (14)$$

The operation of extracting a single channel for the inverse Fourier transform involves first shifting the desired channel to the baseband and then multiplying the Fourier-domain data by a mask function that reduces all but the baseband channel to zero. While Fig. 1 may seem to imply that the shape of each channel is an elliptical domain, the rectangular symmetry of the data makes a rectangle more appropriate, allowing the entire Fourier-transform domain to be filled by the four channels.

If the image sampled by the mosaic filter is bandlimited to within the boundaries of a single channel in the Fourier domain, such that the overlap between channels is negligible, then the ideal mask to use is a rectangular mask (see Fig. 4) that does not attenuate any of the spatial frequencies. If, however, there is nonnegligible overlap between channels, then a better approach is to apply a shaped mask that attenuates higher frequencies, which are more prone to crosstalk. Common mask functions

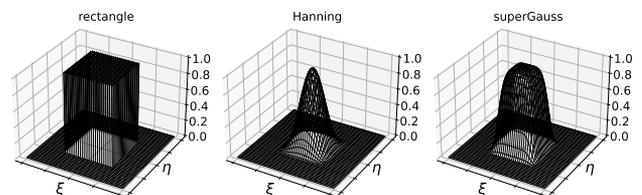


Fig. 4. Three mask functions for weighting the Fourier-domain frequency data.

include the Hanning filter, Blackman filter, or superGauss filter (Fig. 4).

While these mask functions provide a means of reducing crosstalk artifacts in the reconstructed image, they also trade resolution for artifact reduction. Nonlinear algorithms that can make use of the orientation of edge features in the image can actually do a significantly better job at optimizing this tradeoff [15]. While Fig. 1 shows Fourier-domain channels that are uniform in size, it is actually more common in Fourier-domain approaches to allow the luminance channel $W(\xi, \eta)$ to be larger than the chrominance channels U and V , since luminance gradients in natural images tend to be much larger than chrominance gradients.

3. MONOCHROME POLARIZATION FILTER MOSAIC

The polarization camera uses a filter mosaic similar to that of the Bayer filter, but follows the distribution illustrated in Fig. 5(a) [12,16,17]. While one can develop sampling functions for each of the individual micropolarizer orientation angles (0° , 45° , 90° , and 135°), a more efficient approach is once again to take advantage of the symmetries in the sampling functions, and to represent the sampling functions in terms of the Stokes vector elements s_1 and s_2 . Since the s_1 component can be considered as the difference in the quantity of light between the 0° orientation and 90° orientation, the sampling function uses $+1$ for the 0° micropolarizer and -1 for the 90° micropolarizer, as shown in Fig. 5(b). Therefore, the sampling functions can be written as

$$\left. \begin{aligned} \mu_{s0} &= 1 \\ \mu_{s1} &= \frac{1}{2}[\cos(\pi x) - \cos(\pi y)] \\ \mu_{s2} &= \frac{1}{2}[\cos(\pi x) + \cos(\pi y)] \end{aligned} \right\} \quad (15)$$

We can further simplify this set of sampling functions by replacing s_1 and s_2 with the polarization sum and difference $s_+ = s_2 + s_1$, and $s_- = s_2 - s_1$, which are simply [18]

$$\left. \begin{aligned} \mu_{s+} &= \cos(\pi x) \\ \mu_{s-} &= \cos(\pi y) \end{aligned} \right\} \quad (16)$$

The sum and difference sampling functions are therefore Nyquist-frequency modulations in the x and y directions. As a result, we can expect to see a total of three channels in the Fourier domain: channel $C_{0,0}$ centered at the origin, $C_{\pm 1,0}$ straddling the left–right boundary, and $C_{0,\pm 1}$ straddling the top–bottom boundary.

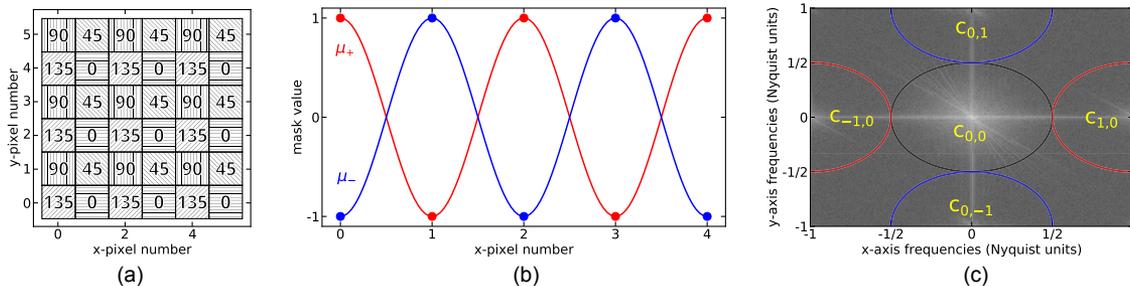


Fig. 5. (a) Polarization filter mosaic, showing the 6×6 subregion near the image origin. (b) Discrete and continuous sinusoidal patterns that are used to represent the sampling locations and modulation mask values for the $s_+ = s_1 + s_2$ and $s_- = s_1 - s_2$ Stokes vector images. (c) Channel centers in the Fourier-domain data.

For the case of the polarization camera, the Fourier-domain reconstruction is therefore straightforward:

$$\left. \begin{aligned} \hat{s}_0(x, y) &= c_{0,0}(x, y) \\ \hat{s}_1(x, y) &= c_{1,0}(x, y) - c_{0,1}(x, y) \\ \hat{s}_2(x, y) &= c_{1,0}(x, y) + c_{0,1}(x, y) \end{aligned} \right\} \quad (17)$$

Figure 6 shows two reconstructions of an example image captured from a monochromatic polarization camera. The “direct” reconstruction involves calculating the s_0 , s_1 , and s_2 values for each 2×2 tile in the mosaic, without interpolation. From the resulting s_1 and s_2 images, we can see that this direct approach causes spurious polarization signatures along edges in the image. In the Fourier-based approach, these spurious signatures also show up if we use the rectangular filter (i.e., no attenuation of high spatial frequencies), but if we apply the “superGauss” filter shaped mask, then the signatures become too small to be visible.

Visualization 1 and Visualization 2 illustrate the spatial frequency distribution among the Fourier-domain channels as we adjust polarization camera lens focus in two scenes.

4. DOUBLE-WIDE BAYER-FILTER MOSAIC

Color-polarization cameras include microfilters for both color and polarization. In order to combine the two, the color filters are given double width, so that an individual filter covers a 2×2 tile of pixels, inside which is a set of four polarization microfilters. Before going straight to the color-polarization mosaic, it is useful to first consider the case of these double-wide color filters and how these operate in the Fourier domain. First we define two auxiliary sampling functions

$$\left. \begin{aligned} \mu_x &= \sqrt{2} \cos\left[\frac{\pi}{4}(2x - 1)\right] \\ \mu_y &= \sqrt{2} \cos\left[\frac{\pi}{4}(2y - 1)\right] \end{aligned} \right\} \quad (18)$$

so that the double-wide filter sampling functions are defined exactly as in Eq. (2), viz,

$$\left. \begin{aligned} \mu_r &= \frac{1}{4}(1 - \mu_x)(1 + \mu_y) \\ \mu_g &= \frac{1}{2}(1 + \mu_x \mu_y) \\ \mu_b &= \frac{1}{4}(1 + \mu_x)(1 - \mu_y) \end{aligned} \right\} \quad (19)$$

or in terms of the WUV luminance–chrominance as

$$\left. \begin{aligned} \mu_w &= 1 \\ \mu_u &= \mu_x \mu_y \\ \mu_v &= \frac{1}{2}(\mu_y - \mu_x) \end{aligned} \right\} \quad (20)$$

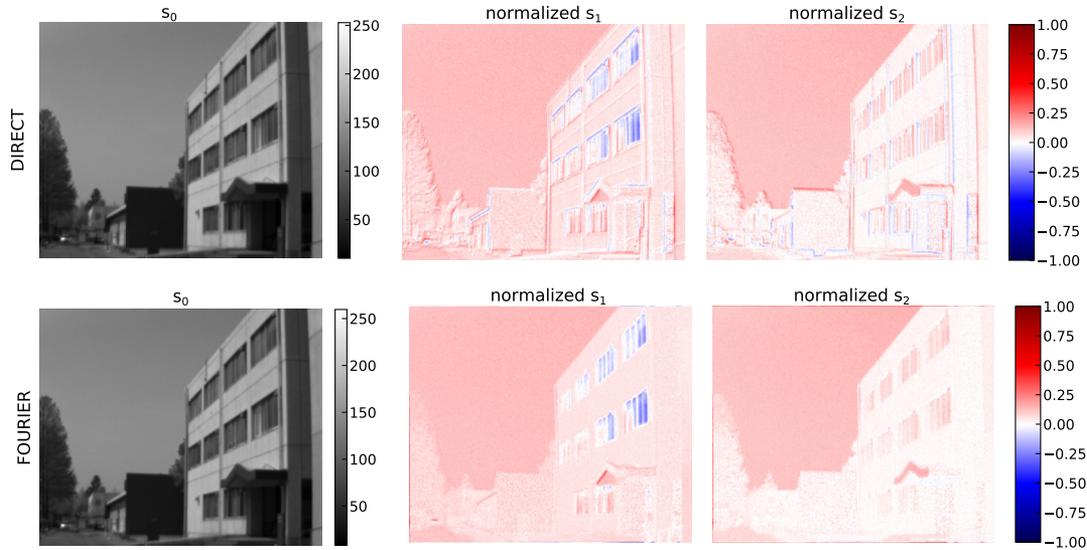


Fig. 6. Example monochromatic polarization camera image reconstruction, showing the results of a “direct” reconstruction, and of a Fourier-based reconstruction that uses a superGauss filter. The direct method shows more polarization artifacts at edge features, whereas the Fourier-domain filter helps to suppress these. (A rectangular Fourier-domain filter obtains a result very similar to that of the direct approach.) [Visualization 1](#) and [Visualization 2](#) show the raw image, reconstructed polarization images, and Fourier-domain distribution while adjusting lens focus.

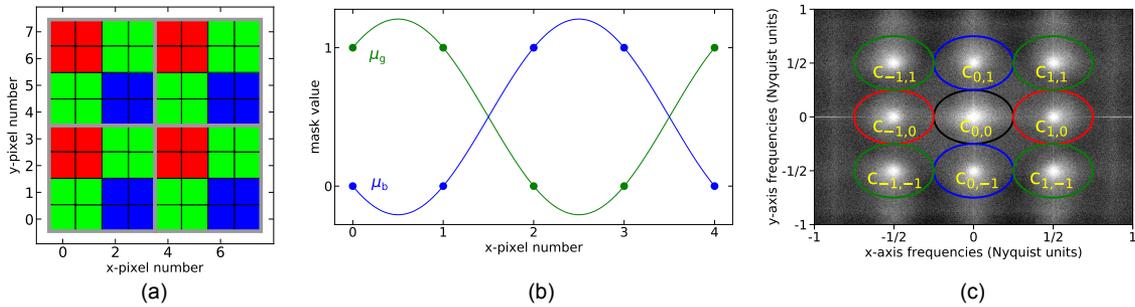


Fig. 7. (a) Double-wide Bayer-filter mosaic, with each superpixel outlined in gray. (b) Discrete and continuous sinusoidal sampling functions of the green and blue pixels in the bottom row of the filter mosaic. (c) Fourier transform of an example mosaic-sampled image. Note that the channel circles’ colors do not directly correspond to RGB colors, but rather to WUV luminance–chrominance.

The resulting sampling functions are shown in Fig. 7, where we can see that the sampling function frequencies are half those of the standard Bayer μ_s , and that there is also a shift — the sampling points no longer lie at a peak or valley of the sinusoidal pattern, but somewhere in between.

Once again, we can simplify the expressions using the WUV luminance–chrominance functions Eq. (6), such that the Fourier-domain representation of the modulation functions shows nine channels in the frequency domain:

$$\begin{aligned} \mathcal{F}\{I_{\text{raw}}(x, y)\} = & W(\xi, \eta)\delta(\xi, \eta) - \frac{1}{4}U(\xi, \eta)[\delta(2\xi - 1, 2\eta + 1) + \delta(2\xi + 1, 2\eta - 1) + i\delta(2\xi - 1, 2\eta - 1) \\ & - i\delta(2\xi + 1, 2\eta + 1)] + \frac{1}{4}V(\xi, \eta)[(1 + i)\delta(\xi, 2\eta - 1) + (1 - i)\delta(\xi, 2\eta + 1) - (1 - i)\delta(2\xi + 1, \eta) \\ & - (1 + i)\delta(2\xi - 1, \eta)]. \end{aligned} \quad (21)$$

This channel distribution is illustrated in Fig. 7(c). As a result of the lower frequency modulations, the channel dimensions are halved, and the channels are no longer straddling the Nyquist boundaries. The smaller channels means that the image must be more tightly bandlimited (image gradients must be smaller) in

order to prevent crosstalk. This is expected, since the tile size has doubled, and so we can expect that bandlimiting the image by defocusing the camera lens must be done more aggressively than in the case of the regular Bayer filter.

As with the Bayer pattern, after masking and inverse Fourier transforming the frequency-domain channels for the quad-Bayer mosaic image, we calculate linear combinations of channels to generate our estimated registered color image:

$$\begin{aligned} \hat{r}(x, y) = & c_{00} + \frac{1}{2}[ic_{-1,-1} + c_{-1,1} + c_{1,-1} + ic_{1,1}] \\ & + \frac{1}{2}[(1 + i)c_{-1,0} + (1 - i)c_{1,0}] \\ & + (1 + i)c_{0,-1} + (1 - i)c_{0,1}, \end{aligned} \quad (22)$$

$$\hat{g}(x, y) = c_{00} - \frac{1}{2}[ic_{-1,-1} + c_{-1,1} + c_{1,-1} + ic_{1,1}], \quad (23)$$

$$\begin{aligned} \hat{b}(x, y) = & c_{00} + \frac{1}{2}[ic_{-1,-1} + c_{-1,1} + c_{1,-1} + ic_{1,1}] \\ & - \frac{1}{2}[(1+i)c_{-1,0} + (1-i)c_{1,0}] \\ & + (1+i)c_{0,-1} + (1-i)c_{0,1}. \end{aligned} \quad (24)$$

Using the Hermitian symmetry of the channels Eq. (10), and by ignoring channels with redundant data, we can simplify Eqs. (22)–(24) by considering only the positive-frequency channels, giving

$$\hat{r}(x, y) = c_{00} + 2ic_{11} + \frac{2}{1+i}(c_{01} - c_{10}), \quad (25)$$

$$\hat{g}(x, y) = c_{00} - 2ic_{11}, \quad (26)$$

$$\hat{b}(x, y) = c_{00} + 2ic_{11} - \frac{2}{1+i}(c_{01} - c_{10}). \quad (27)$$

5. COLOR-POLARIZATION FILTER MOSAIC

The color-polarization camera filter mosaic pattern [Fig. 8(a)] is a simple multiplication of the monochromatic polarization mosaic with the double-wide Bayer mosaic [19,20]. Thus, the full set of sampling functions, for each of the nine RGB color and Stokes vector elements, can be written as

$$\begin{aligned} \mu_{gs0} = \mu_g\mu_s0 &= \frac{1}{2}(1 + \mu_x\mu_y), \\ \mu_{gs+} = \mu_g\mu_{s+} &= \frac{1}{2}(1 + \mu_x\mu_y) \cos(\pi x), \\ \mu_{gs-} = \mu_g\mu_{s-} &= \frac{1}{2}(1 + \mu_x\mu_y) \cos(\pi y), \\ \mu_{rs0} = \mu_r\mu_s0 &= \frac{1}{4}(1 - \mu_x)(1 + \mu_y), \\ \mu_{rs+} = \mu_r\mu_{s+} &= \frac{1}{4}(1 - \mu_x)(1 + \mu_y) \cos(\pi x), \\ \mu_{rs-} = \mu_r\mu_{s-} &= \frac{1}{4}(1 - \mu_x)(1 + \mu_y) \cos(\pi y), \\ \mu_{bs0} = \mu_b\mu_s0 &= \frac{1}{4}(1 + \mu_x)(1 - \mu_y), \\ \mu_{bs+} = \mu_b\mu_{s+} &= \frac{1}{4}(1 + \mu_x)(1 - \mu_y) \cos(\pi x), \\ \mu_{bs-} = \mu_b\mu_{s-} &= \frac{1}{4}(1 + \mu_x)(1 - \mu_y) \cos(\pi y), \end{aligned} \quad (28)$$

where μ_x and μ_y are the auxiliary sampling functions defined in Eq. (18). Figure 8(b) shows two of these sampling functions (only two are shown for clarity), each of which have the repeated sequence 0, 0, +1, -1, ...

As before, using the WUV luminance–chrominance representation for color, and the $s_0s_+s_-$ representation for polarization, takes advantage of the mosaic symmetry to simplify the Fourier-domain expressions. Once we look into the Fourier-domain representation, we quickly see that some of the modulations occur at higher frequencies than the Nyquist limit, which means that these will get wrapped. A modulation of $3/2$

in Nyquist units will therefore wrap around and appear at $-1/2$ instead.

Since there are now 21 Fourier-domain channels—a 5×5 array of channels in the Fourier domain, except the corners—if we number the channels in half-Nyquist frequency units, then the indices go from -2 to $+2$, where ± 2 lies on the Nyquist boundary and ± 3 wraps around to become ∓ 1 . With this wrapping, and the consolidating of channels lying on the Nyquist boundary, a total of 15 channels remain. Because the expressions are quite lengthy, the steps leading up to the final expression are placed in Supplement 1, and only the final result is show here. The algorithm of reconstructing the color and polarization from combinations of channels is

$$\begin{aligned} r_{s0} = & 2c_{0,0} - (1+i)c_{1,0} - c_{1,-1} - (1-i)c_{-1,0} - c_{-1,1} \\ & + ic_{-1,-1} - ic_{1,1} + (1-i)c_{0,-1} + (1+i)c_{0,1}, \\ r_{s1} = & \frac{1}{2}[2c_{2,0} - 2c_{0,2} + 2ic_{1,-1} - 2ic_{-1,1} \\ & - (1-i)c_{1,0} + (1+i)c_{2,1} + (1-i)c_{2,-1} - (1+i)c_{-1,0} \\ & - (1-i)c_{0,1} - (1+i)c_{0,-1} + (1+i)c_{1,2} + (1-i)c_{-1,2}], \\ r_{s2} = & \frac{1}{2}[-2c_{2,0} - 2c_{0,2} + 2c_{1,1} + 2c_{-1,-1} \\ & + (1-i)c_{1,0} - (1+i)c_{2,1} - (1-i)c_{2,-1} + (1+i)c_{-1,0} \\ & - (1-i)c_{0,1} - (1+i)c_{0,-1} + (1+i)c_{1,2} + (1-i)c_{-1,2}], \\ g_{s0} = & 2c_{0,0} + ic_{1,1} + c_{1,-1} + c_{-1,1} - ic_{-1,-1}, \\ g_{s1} = & c_{2,0} - c_{0,2} - ic_{1,-1} + ic_{-1,1}, \\ g_{s2} = & -c_{2,0} - c_{0,2} - c_{1,1} - c_{-1,-1}, \\ b_{s0} = & 2c_{0,0} - ic_{1,1} - c_{1,-1} - c_{-1,1} + ic_{-1,-1} \\ & - (1+i)c_{0,1} - (1-i)c_{0,-1} + (1+i)c_{1,0} + (1-i)c_{-1,0}, \\ b_{s1} = & \frac{1}{2}[2c_{2,0} - 2c_{0,2} + 2ic_{1,-1} - 2ic_{-1,1} \\ & + (1-i)c_{1,0} - (1+i)c_{2,1} - (1-i)c_{2,-1} + (1+i)c_{-1,0} \\ & + (1-i)c_{0,1} + (1+i)c_{0,-1} - (1+i)c_{1,2} - (1-i)c_{-1,2}], \\ b_{s2} = & \frac{1}{2}[-2c_{2,0} - 2c_{0,2} + 2c_{1,1} + 2c_{-1,-1} \\ & - (1-i)c_{1,0} + (1+i)c_{2,1} + (1-i)c_{2,-1} - (1+i)c_{-1,0} \\ & + (1-i)c_{0,1} + (1+i)c_{0,-1} - (1+i)c_{1,2} - (1-i)c_{-1,2}], \end{aligned} \quad (29)$$

where r, g, b, s_j , and $c_{m,n}$ are each images, and thus have (x, y) coordinates. Figure 9 shows an example of using Eq. (29) to reconstruct a color-polarization camera image. As in most polarization images, we can see that the three colors show strong correlations among one another. Visualization 3 shows the raw image and Fourier-domain distribution while adjusting lens focus.

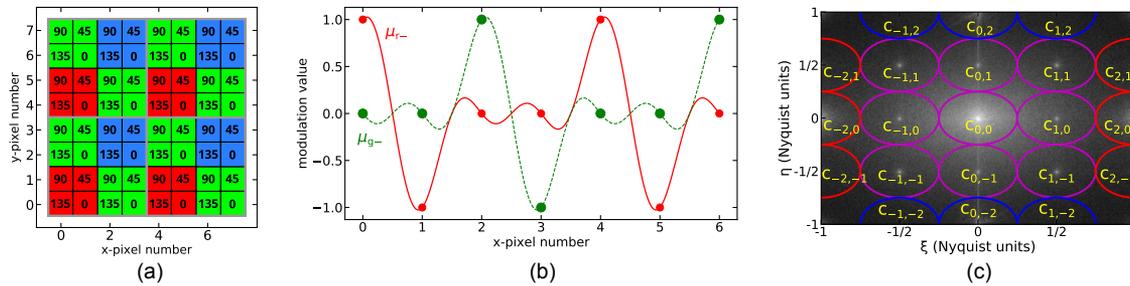


Fig. 8. (a) Color-polarization camera filter mosaic. (b) Sampling functions for the red and green s_{-} filters in the second row of the mosaic (i.e., $y = 1$). (c) Frequency-domain representation of the mosaic-sampled image (simulated using the image from Fig. 3(a), after blurring with a Gaussian filter of width $\sigma = 4$ pixels).

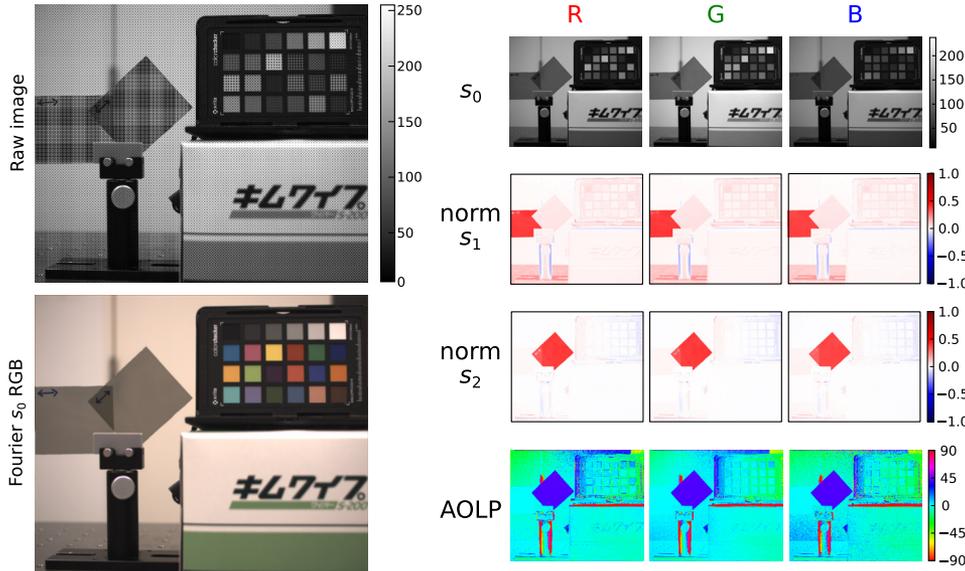


Fig. 9. Example color-polarization camera image reconstruction, showing the Stokes vector images for each of the R, G, and B colors separately. (Visualization 3 shows the raw image and Fourier-domain distribution of a scene captured by a color-polarization camera while adjusting lens focus.)

6. CHANNEL CROSSTALK ANALYSIS PROCEDURE

In order to quantify the Fourier-domain channel crosstalk, we use the scene shown in Fig. 10, captured with an RGB polarization camera, with the lens at best focus (the first row) and severely defocused (the second row). Although the scene may not appear to have strong horizontal and vertical features, the Fourier-domain distributions Figs. 10(b) and 10(e) show that the frequency amplitudes are concentrated near the horizontal and vertical axes, particularly in the defocused case. As a result, we can see that the baseband distribution is bleeding across into channel $C_{0,1}$ and will create a false color-polarization signature there.

In order to see this quantitatively, Figs. 10(c) and 10(f) show the horizontal, vertical, and diagonal cross sections through the Fourier-domain distribution. In the focused image data, we see that the vertical axis has the strongest crosstalk emanating from the baseband signal, the horizontal axis somewhat less, and the diagonal less still. When we defocus the lens, the Fourier-domain amplitudes decrease, but the decrease is greatest for the diagonal, modest for the horizontal, and smallest for the vertical

axis data. As a result, the crosstalk emanating from the baseband is even more strongly concentrated around the axes in the defocused case than the focused case, though the overall crosstalk has gone down. From this, we can conclude that for standard images it is difficult indeed to reduce crosstalk to a negligible amount if the sideband channels are located along the axes. One could of course rotate the camera with respect to the local horizontal axis, since this rotates the Fourier-domain data with respect to the channel distribution. However, user preference for images oriented parallel or perpendicular to the gravity axis means that such images would need a second interpolation step, and to be cropped as well. This reduces resolution and field of view. Clearly, it is preferable to place the sidebands off-axis, if possible [6,12].

In order to estimate the crosstalk quantitatively, we use the data from Figs. 10(c) and 10(f). In each case, we fit a curve (shown as the dashed curves in the figures) to the distribution emanating from the baseband and ignoring the sideband peaks. The amplitude of the dashed curve at the location of the sideband peak (corresponding to the $C_{0,1}$ and $C_{0,-1}$ channel data) gives a measure of the crosstalk amount. Likewise, the ratio

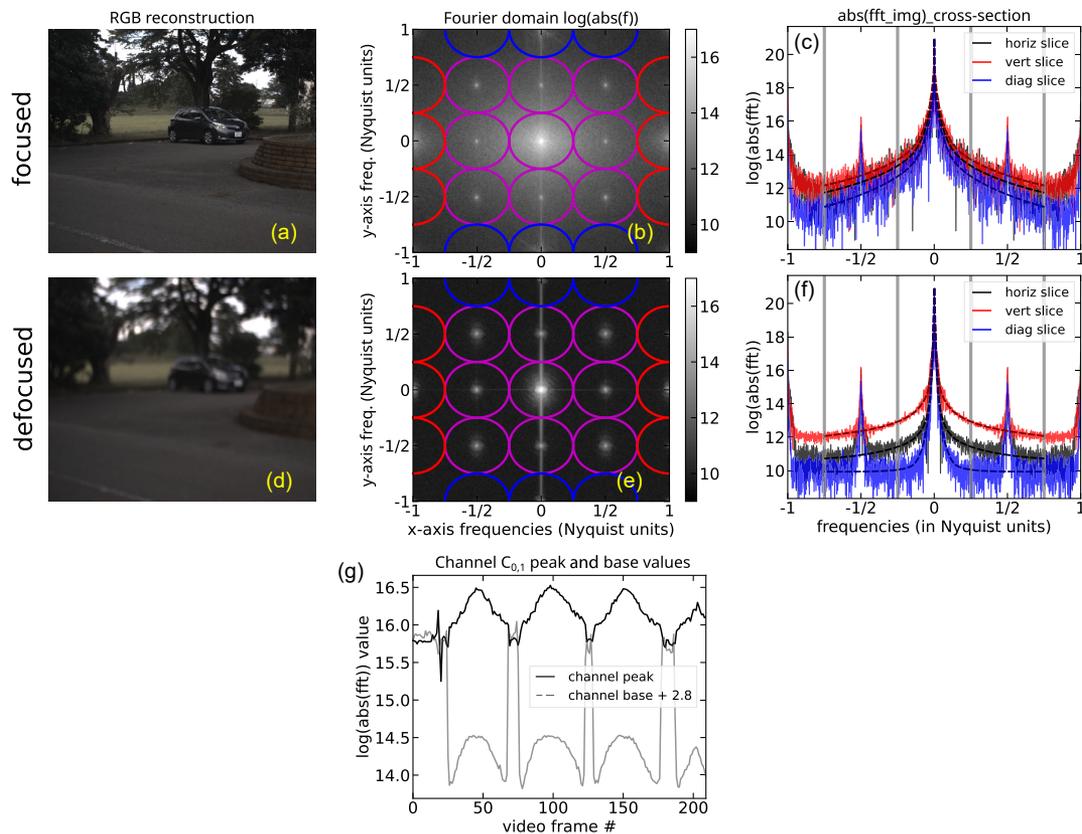


Fig. 10. Example scene captured with an RGB polarization camera, with the lens set (a)–(c) at best focus or (d)–(f) severely defocused. (b), (e) show the corresponding Fourier-domain distributions, while (c), (f) show the horizontal, vertical, and diagonal cross sections through the Fourier-domain data. Subfigure (g) shows the amplitude (in log units) of the Fourier-domain peak for channel $C_{0,1}$ (black curve) together with the estimated amplitude of crosstalk leakage from the baseband (gray curve), for a video in which the lens is repeatedly focused and defocused (see Visualization 4).

between the crosstalk estimate and the channel peak value gives an estimate of the fraction of measured polarization that is due to crosstalk. Here we see that in the focused data the channel peak has a value of 15.8 (the logarithm of the Fourier-domain absolute value), while the crosstalk is estimated as 13.0, so that the ratio is $\exp(13.0 - 15.8) = 0.06$. Thus, about 6% of the overall color-polarization signal in that channel will be false. In the defocused case, the estimated crosstalk is 11.3, and the channel peak is 16.1, so that the ratio is $\exp(11.3 - 16.1) = 0.008$ —the crosstalk fraction has been reduced below 1%.

7. CONCLUSION

In the discussion above, we have collected together the expressions for Fourier-domain demosaicking from the existing literature, given them a consistent notation, and augmented them with expressions for the quad-Bayer-filter mosaic and the color-polarization camera mosaic. For each case, we have shown how the various spatial frequency channels in the Fourier domain can be combined to give the estimated color and polarization images. This linear algorithm approach is particularly useful for visualizing the sampling behavior of each sensor mosaic, providing a direct means of estimating the crosstalk. For any given scene, using our prior knowledge of the channel distribution in the Fourier domain, we can select a focus setting that provides an optimal tradeoff (for the given task at hand) between

image resolution and color-polarization fidelity. Providing an estimate for the optimal focus tradeoff point for crosstalk versus resolution would require defining a specific task, as well as a specific image set on which to evaluate task performance. Thus it is both task-specific and object-specific, and defining a proper value for general-purpose use is a significant but important undertaking.

One place where the Fourier-domain approach can be useful is to augment existing nonlinear demosaicking algorithms to take advantage of the specific nonuniform crosstalk distribution in a given image. Since the primary source of crosstalk is the baseband distribution bleeding into the sidebands, one can estimate the crosstalk, and attempt to compensate for it, prior to reconstructing the sideband data. This would be a Fourier-based nonlinear demosaicking approach, taking advantage of the visualizable information provided by the Fourier-domain distribution.

The algorithms discussed above have also been implemented in an open-access code repository written in Python, together with example use cases that generate the figures used in the manuscript.

Funding. The Development of Lecturer's Research Capabilities Program through the International Network, School of Science, King Mongkut's Institute of Technology Ladkrabang (2023).

Acknowledgment. This research is supported by the Development of Lecturer's Research Capabilities program through the International

Network from the School of Science, King Mongkut's Institute of Technology Ladkrabang, for the 2023 fiscal year.

Disclosures. The authors declare no conflict of interest.

Data availability. The data and algorithms presented in this paper are publicly available in a repository [21]. Other data underlying the results presented in this paper may be obtained from the authors upon request.

Supplemental document. See Supplement 1 for supporting content.

REFERENCES

1. V. Gruev, R. Perkins, and T. York, "CCD polarization imaging sensor with aluminum nanowire optical filters," *Opt. Express* **18**, 19087–19094 (2010).
2. N. Hagen, S. Shibata, and Y. Otani, "Calibration and performance assessment of microgrid polarization cameras," *Opt. Eng.* **58**, 082408 (2019).
3. Y. Maruyama, T. Terada, T. Yamazaki, *et al.*, "3.2-MP back-illuminated polarization image sensor with four-directional air-gap wire grid and 2.5- μm pixels," *IEEE Trans. Electron Devices* **65**, 2544–2551 (2018).
4. K. Shinoda, Y. Ohtera, and M. Hasegawa, "Snapshot multispectral polarization imaging using a photonic crystal filter array," *Opt. Express* **26**, 15948–15961 (2018).
5. T. W. Sawyer, M. Taylor-Williams, R. Tao, *et al.*, "Opti-MSFA: a toolbox for generalized design and optimization of multispectral filter arrays," *Opt. Express* **30**, 7591–7611 (2022).
6. C. Bai, J. Li, Z. Lin, *et al.*, "Automatic design of color filter arrays in the frequency domain," *IEEE Trans. Image Process.* **25**, 1793–1806 (2016).
7. E. Dubois, "Frequency-domain methods for demosaicking of Bayer-sampled color images," *IEEE Signal Process. Lett.* **12**, 847–850 (2005).
8. K. Hirakawa and T. W. Parks, "Adaptive homogeneity-directed demosaicking algorithm," *IEEE Trans. Image Process.* **14**, 360–369 (2005).
9. B. Arad, R. Timofte, R. Yahel, *et al.*, "NTIRE 2022 spectral demosaicking challenge and data set," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2022)*, pp. 882–896.
10. K. Hirakawa and P. J. Wolfe, "Spatio-spectral color filter array design for optimal image recovery," *IEEE Trans. Image Process.* **17**, 1876–1890 (2008).
11. S. P. Jaiswal, L. Fang, V. Jakhetiya, *et al.*, "Adaptive multispectral demosaicking based on frequency-domain analysis of spectral correlation," *IEEE Trans. Image Process.* **26**, 953–968 (2016).
12. A. S. Alenin, I. J. Vaughn, and J. S. Tyo, "Optimal bandwidth micropolarizer arrays," *Opt. Lett.* **42**, 458–461 (2017).
13. <https://species.wikimedia.org/wiki/Chelicerata#/media/File:PlatycryptusUndatusFemale.jpg>.
14. D. Menon and G. Calvagno, "Color image demosaicking: an overview," *Signal Process. Image Commun.* **26**, 518–533 (2011).
15. A. Stojkovic, I. Shopovska, H. Luong, *et al.*, "The effect of the color filter array layout choice on state-of-the-art demosaicking," *Sensors* **19**, 3215–3230 (2019).
16. D. A. LeMaster and K. Hirakawa, "Improved microgrid arrangement for integrated imaging polarimeters," *Opt. Lett.* **39**, 1811–1814 (2014).
17. B. M. Ratliff and G. C. Sargent, "Alternative linear microgrid polarimeters: design, analysis, and demosaicking considerations," *Appl. Opt.* **60**, 5805–5818 (2021).
18. J. S. Tyo, C. F. LaCasse, and B. M. Ratliff, "Total elimination of sampling errors in polarization imagery obtained with integrated microgrid polarimeters," *Opt. Lett.* **34**, 3187–3189 (2009).
19. M. Morimatsu, Y. Monno, M. Tanaka, *et al.*, "Monochrome and color polarization demosaicking using edge-aware residual interpolation," *arXiv*, arXiv:2007.14292 (2020).
20. N. Hagen and Y. Otani, "Fourier domain filtering method for demosaicking color polarization camera images," *Proc. SPIE* **12327**, 123270Z (2023).
21. N. Hagen, "Fourier demosaicking: initial release," Zenodo, 2023, <https://doi.org/10.5281/zenodo.8149453>.