2008

Spatio-Spectral Sampling and Color Filter Array Design

Keigo Hirakawa
University of Dayton, khirakawa1@udayton.edu

Patrick J. Wolfe
Harvard University

5

Spatio-Spectral Sampling and Color Filter Array Design

Keigo Hirakawa and Patrick J. Wolfe

5.1 Introduction

Owing to the growing ubiquity of digital image acquisition and display, several factors must be considered when developing systems to meet future color image processing needs, including improved quality, increased throughput, and greater cost-effectiveness [1], [2], [3]. In consumer still-camera and video applications, color images are typically obtained via a spatial subsampling procedure implemented as a color filter array (CFA), a physical construction whereby only a single component of the color space is measured at each pixel location [4], [5], [6], [7]. Substantial work in both industry as well as academia has been dedicated to postprocessing this acquired raw image data as part of the so-called *image processing pipeline*, including in particular the canonical *demosaicking* task of reconstructing a full color image from the spatially subsampled and incomplete data acquired using a CFA [8], [9], [10], [11], [12], [13]. However, as we detail in this chapter, the inherent shortcomings of contemporary CFA designs mean that subsequent processing steps often yield diminishing returns in terms of image quality. For example, though distortion may be masked to some extent by motion blur and compression, the loss of image quality resulting from all but the most computationally expensive state-of-the-art methods is unambiguously apparent to the practiced eye. Refer to Chapters 1 and 3 for additional information on single-sensor imaging fundamentals.

As the CFA represents one of the first steps in the image acquisition pipeline, it largely determines the maximal resolution and computational efficiencies achievable by subsequent processing schemes. Here we show that the attainable spatial resolution yielded by a particular choice of CFA is quantifiable, and propose new CFA designs to maximize it [14], [15]. In contrast to the majority of the demosaicking literature, we explicitly consider the interplay between CFA design and properties of typical image data, and its implications for spatial reconstruction quality. Formally, we pose the CFA design problem as simultaneously maximizing the allowable spatio-spectral support of luminance and chrominance channels, subject to a partitioning requirement in the Fourier representation of the sensor data. This classical *aliasing-free* condition preserves the integrity of the color image data and thereby guarantees exact reconstruction when demosaicking is implemented as demodulation (demultiplexing in frequency).

Surprisingly, from this perspective we can show the suboptimality of CFA designs based on pure tristimulus values [15]—a standard design approach long taken by industry, particularly as manifested by the popular Bayer pattern [4]. Such designs are less resilient to spatial aliasing as image resolution increases, requiring both stronger assumptions about the image data as well as more computationally demanding nonlinear demosaicking methods to avoid reconstruction artifacts. Here our interest lies in quantifying the trade-offs between performance and complexity for different classes of CFA design; we consider the purely linear reconstruction of typical images as an indication of baseline performance, and interpret the resultant degree of aliasing as providing a measure of the maximally attainable spatio-spectral resolution.

As an alternative to existing CFA patterns, we provide a constructive method to generate feasible CFA designs that exhibit robustness to prior assumptions on color channel bandlimitedness and yield high performance while implying only low complexity for subsequent processing steps in the imaging pipeline. Because our emphasis is on the efficiencies of the overall color image acquisition pipeline, we omit an explicit comparison of demosaicking strategies. However, our analysis yields a general class of linear demosaicking methods that provide state-of-the-art performance and enjoy complexity comparable to simple bilinear interpolation. In addition, our proposed CFA designs are also designed for increased noise robustness: the color filters themselves are panchromatic, alleviating difficulties in low-light conditions, and the linear reconstruction methods we propose can also be expected to enable more tractable noise modelling [15].

The remainder of this chapter is organized as follows. We begin in Section 5.2 by examining the spatio-spectral properties of typical CFA designs in the Fourier domain, and discuss their susceptibility to aliasing. We propose in Section 5.3 a constructive method to specify a physically realizable CFA pattern in terms of its spatio-spectral properties. The resultant CFA designs admit fast, optimal linear reconstruction schemes, which we outline in Section 5.4. In Section 5.5 we give several explicit examples of these new patterns, and provide empirical evaluations on standard color image test sets. We summarize and conclude with a discussion in Section 5.6.

5.2 Spatio-Spectral Analysis of Existing Patterns

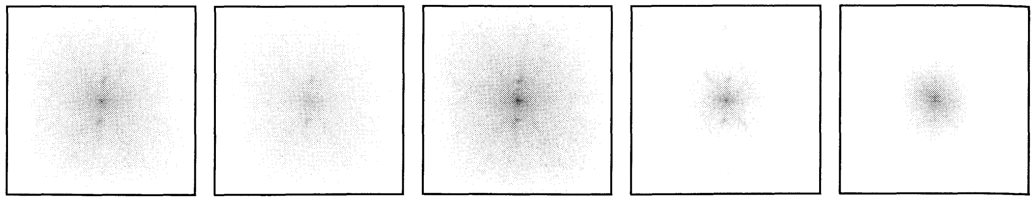
In this section, the spatio-spectral properties of the sampling induced by existing CFA patterns are analyzed. In single-sensor cameras, the pixel sensor at each spatial location is equipped with a color filter, a physical device whose pigments absorb a portion of the electro-magnetic wave in the visible spectrum while passing the rest to the photosensitive element beneath this filter. The measured value at each location is therefore an inner product resulting from a spatio-temporal integration of the incident light over each pixel's physical area and exposure time, taken with respect to the corresponding color filter's spectral response. This is similar to the acquisition process in the retina, where each cone measures the intensity of the light with respect to its spectrally-shifted response [4], [5], [6], [7]. Because the spectral response functions of the cones can be taken to span a three-dimensional space, and cone and sensor measurements are largely proportional to the intensity of the light (i.e., linear), the observed light can be uniquely represented (up to linear transformation) by a color triple. We therefore adopt the standard convention and identify these filters by their color names such as red, green, and blue—though these may not be synonymous with perceived color, which is a function of the environmental illuminant [1]. As the goal of this chapter is the identification and optimization of relevant objective metrics, rather than subjective metrics related to perception, we make no further attempt to elaborate on the issues of color science.

5.2.1 Color Filter Arrays

Here we begin with the Fourier analysis of the spatio-spectral properties of the CFA patterns [10], [13]. This spatially global perspective is a logical starting point for a number of reasons (a spatially local perspective is provided in the next section). First, color filter arrays are physical constructions that are fixed prior to image acquisition, and therefore not adapted to local image properties. Second, color filter arrays typically comprise a repetitive tiling of the image plane formed by the union of alternating color samples.[1] As we describe below, the global spatial periodicity of CFA sampling patterns may be understood in terms of lattices, with a so-called dual or reciprocal lattice determining the resultant spectral periodicity under Fourier transform. Finally, the linear reconstruction methods we consider in the interest of evaluating computation-quality trade-offs preclude adaptation to local statistics of the image under consideration.

To motivate our analysis, let us first consider the interplay between color channels of the acquired image. Let $\boldsymbol{x}(\boldsymbol{n}) = [x_r(\boldsymbol{n}), x_g(\boldsymbol{n}), x_b(\boldsymbol{n})]^T$ denote the RGB tristimulus value of the desired color image at pixel location $\boldsymbol{n} \in \mathbb{Z}^2$. Define $\boldsymbol{c}(\boldsymbol{n}) = [c_r(\boldsymbol{n}), c_g(\boldsymbol{n}), c_b(\boldsymbol{n})]^T$ as the corresponding CFA color combination, so that the measured sensor value $y(\boldsymbol{n})$ at location $\boldsymbol{n}$ can be expressed as the inner product $y(\boldsymbol{n}) = \boldsymbol{c}(\boldsymbol{n})^T \boldsymbol{x}(\boldsymbol{n})$. For the moment, we restrict our attention to $\boldsymbol{c}(\boldsymbol{n}) \in \{[1,0,0]^T, [0,1,0]^T, [0,0,1]^T\}$ as a model for CFA schemes

[1] Pseudo-random CFA patterns have also been considered in the past [7]. Despite their potential theoretical advantages, we omit them from our discussion, as the corresponding reconstruction schemes incur much greater computational expense.



(a) (b) (c) (d) (e)

FIGURE 5.1

Log-magnitude spectra of a typical color image (i.e., image *flower* here) illustrating the lowpass nature of difference channels x_α and x_β relative to x_r, x_g, and x_b. Individual spectra correspond to: (a) red channel, (b) green channel, (c) blue channel, (d) difference x_α, and (e) difference x_β.

that multiplex color samples; note that $c_r + c_g + c_b = 1$. Each pixel sensor thus measures

$$y(\boldsymbol{n}) = \boldsymbol{c}(\boldsymbol{n})^T \boldsymbol{x}(\boldsymbol{n}) = \begin{bmatrix} c_r(\boldsymbol{n}) & c_g(\boldsymbol{n}) & c_b(\boldsymbol{n}) \end{bmatrix} \begin{bmatrix} x_r(\boldsymbol{n}) \\ x_g(\boldsymbol{n}) \\ x_b(\boldsymbol{n}) \end{bmatrix} = \begin{bmatrix} c_r(\boldsymbol{n}) & 1 & c_b(\boldsymbol{n}) \end{bmatrix} \begin{bmatrix} x_\alpha(\boldsymbol{n}) \\ x_g(\boldsymbol{n}) \\ x_\beta(\boldsymbol{n}) \end{bmatrix}, \quad (5.1)$$

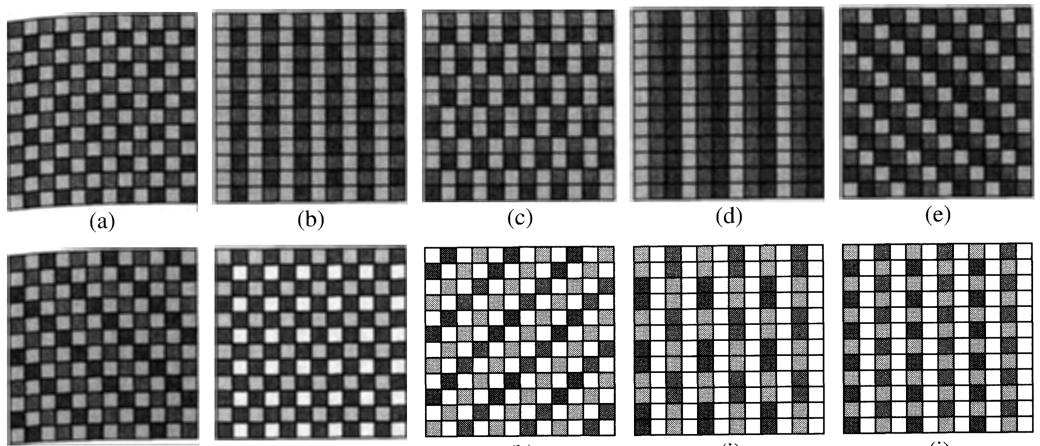
where $x_\alpha = x_r - x_g$ and $x_\beta = x_b - x_g$ are difference channels. As noted in References [10] and [13], this $\{x_\alpha, x_g, x_\beta\}$ representation offers an advantage over the original $\{x_r, x_g, x_b\}$ formulation; the difference channels x_α and x_β serve as a proxy for chrominance components, which enjoy rapid decay in the spatial frequency domain, whereas x_g can be taken to represent the image luminance component, which embodies edge and texture information. In fact, the Pearson correlation coefficient measured between the high-frequency components of the color channels $\{x_r, x_g, x_b\}$ is typically larger than 0.9 [8]—and because of this high degree of redundancy, it is often assumed that x_α and x_β are lowpass relative to $\{x_r, x_g, x_b\}$; see Figure 5.1.

The key observation to be gleaned from Equation 5.1 is that y constitutes a sum of the green channel x_g and the subsampled difference images $c_r \cdot x_\alpha$ and $c_b \cdot x_\beta$. In order to understand the limitations of existing color filter array designs, it is helpful to consider the geometric and algebraic structure of subsampling patterns c_r and c_b through the notion of point lattices [15]. To this end, we say a (nonsingular) sampling matrix $\boldsymbol{M} \in \mathbb{R}^{2\times 2}$ generates a lattice $\boldsymbol{M}\mathbb{Z}^2$. Certain sampling patterns c_r and c_b can in turn be rewritten as two-dimensional pulse trains using lattice notation:

$$c_r(\boldsymbol{n}) = \sum_{\boldsymbol{n}_0 \in \{\boldsymbol{m}_r + \boldsymbol{M}_r \mathbb{Z}^2\}} \delta(\boldsymbol{n} - \boldsymbol{n}_0); \qquad c_b(\boldsymbol{n}) = \sum_{\boldsymbol{n}_0 \in \{\boldsymbol{m}_b + \boldsymbol{M}_b \mathbb{Z}^2\}} \delta(\boldsymbol{n} - \boldsymbol{n}_0), \quad (5.2)$$

where $\boldsymbol{M}_r, \boldsymbol{M}_b$ are 2×2 sampling matrices; $\boldsymbol{m}_r, \boldsymbol{m}_b \in \mathbb{Z}^2$ are termed coset vectors; and $\delta(\boldsymbol{n})$ is the Kronecker delta function.[2] Lattices themselves admit the notion of a Fourier transform as specified by a dual lattice $2\pi \boldsymbol{M}^{-T}\mathbb{Z}^2$; if we define $Y(\boldsymbol{\omega})$ as the Fourier transform (in angular frequency $\boldsymbol{\omega}$) of sensor data $y(\boldsymbol{n})$, it follows from Equations 5.1 and 5.2 that

[2] In fact, Equation 5.2 represents a special case in which sampling patterns c_r and c_b are each themselves lattices. More generally, they are defined in terms of unions of lattice cosets [15]; however, this does not change the fundamentals of our present discussion.



(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

FIGURE 5.2 (See color insert.)

Examples of existing CFAs: (a) Bayer [4], (b) Yamanaka [5], (c) Lukac [7], (d) vertical stripes [7], (e) diagonal stripes [7], (f) modified Bayer [7], (g) cyan-magenta-yellow, (h) Kodak I [16], (i) Kodak II [16], (j) Kodak III [16].

$Y(\boldsymbol{\omega})$ over the region $[-\pi, \pi) \times [-\pi, \pi)$ is given by

$$\begin{aligned} Y(\boldsymbol{\omega}) = X_g(\boldsymbol{\omega}) &+ |\det(\boldsymbol{M}_r)|^{-1} \sum_{\boldsymbol{\lambda}_r \in \{2\pi \boldsymbol{M}^{-T}\mathbb{Z}^2 \cap [-\pi,\pi)^2\}} e^{-j\boldsymbol{m}_r^T \boldsymbol{\omega}} X_\alpha(\boldsymbol{\omega} - \boldsymbol{\lambda}_r) \\ &+ |\det(\boldsymbol{M}_b)|^{-1} \sum_{\boldsymbol{\lambda}_b \in \{2\pi \boldsymbol{M}^{-T}\mathbb{Z}^2 \cap [-\pi,\pi)^2\}} e^{-j\boldsymbol{m}_b^T \boldsymbol{\omega}} X_\beta(\boldsymbol{\omega} - \boldsymbol{\lambda}_b). \end{aligned} \quad (5.3)$$

The key point of Equation 5.3 is that these dual lattices specify the *carrier* frequencies $\{\boldsymbol{\lambda}_r, \boldsymbol{\lambda}_b\}$ about which spectral copies of the difference channels x_α and x_β are replicated in the Fourier domain. The popular Bayer CFA [4], for instance, can be specified as $\boldsymbol{M}_r = \boldsymbol{M}_b = 2\boldsymbol{I}$, $\boldsymbol{m}_r = [0,0]^T$, and $\boldsymbol{m}_b = [1,1]^T$—implying dual lattices equal to $\pi\mathbb{Z}^2$, with nonzero $\{\boldsymbol{\lambda}_r, \boldsymbol{\lambda}_b\}$ given by $[-\pi, 0]^T$, $[0, -\pi]^T$, and $[-\pi, -\pi]^T$.

Examples of several existing CFAs $\boldsymbol{c}(\boldsymbol{n})$ and the corresponding spectra $Y(\boldsymbol{\omega})$ of typical sensor data are illustrated in Figure 5.2 and Figure 5.3, respectively; note that aliasing occurs when, for nonzero $\boldsymbol{\lambda}_r$ or $\boldsymbol{\lambda}_b$, the spectral supports of $X_g(\boldsymbol{\omega})$ and $X_\alpha(\boldsymbol{\omega} - \boldsymbol{\lambda}_r)$ or $X_\beta(\boldsymbol{\omega} - \boldsymbol{\lambda}_b)$ overlap.

Despite its widespread use, the spectral periodization about $[-\pi, 0]^T$ and $[0, -\pi]^T$ induced by the Bayer CFA severely limits allowable spectral bandwidth for X_g. In fact, all CFAs depicted in Figure 5.2 are suboptimal in at least one of two ways: First, as shown in Figure 5.3a to Figure 5.3d and Figure 5.3g to Figure 5.3j, spectral copies of the difference channels appear along the horizontal and/or vertical axes of the Fourier representation, leaving the baseband channel X_g vulnerable to the horizontal and vertical features that frequently dominate natural images [17]. Second, as shown in Figure 5.3d to Figure 5.3f and Figure 5.3h to Figure 5.3j, maximal separation between $X_g(\boldsymbol{\omega})$ and $X_\alpha(\boldsymbol{\omega} - \boldsymbol{\lambda}_r), X_\beta(\boldsymbol{\omega} - \boldsymbol{\lambda}_b)$ is precluded unless all nonzero carrier frequencies $\{\boldsymbol{\lambda}_r, \boldsymbol{\lambda}_b\}$ lie elsewhere along the perimeter of $[-\pi, \pi) \times [-\pi, \pi)$.

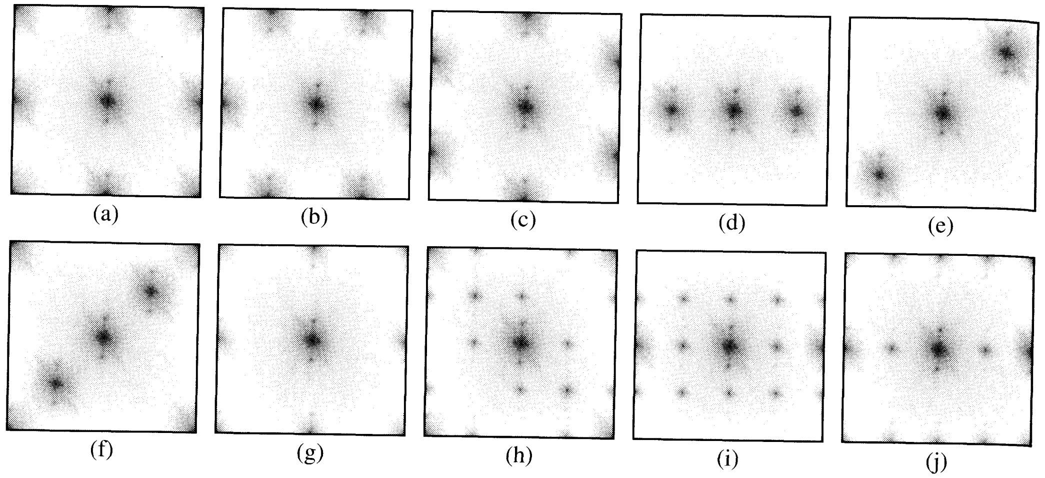


FIGURE 5.3 (See color insert.)

Log-magnitude spectra of a typical color image (i.e., image *flower* here) sampled with CFAs corresponding to Figure 5.2. Color coding is used to distinguish different components, with the $x_g(\boldsymbol{n})$ component shown in green, $x_\alpha(\boldsymbol{n}) = x_r(\boldsymbol{n}) - x_g(\boldsymbol{n})$ in red, and $x_\beta(\boldsymbol{n}) = x_b(\boldsymbol{n}) - x_g(\boldsymbol{n})$ in blue. Individual spectra correspond to: (a) Bayer [4], (b) Yamanaka [5], (c) Lukac [7], (d) vertical stripes [7], (e) diagonal stripes [7], (f) modified Bayer [7], (g) cyan-magenta-yellow, (h) Kodak I [16], (i) Kodak II [16], (j) Kodak III [16].

In fact, these two conditions can be used to formulate a precise statement of CFA suboptimality [15]: any CFA design of the form $\boldsymbol{c}(\boldsymbol{n}) \in \{[1,0,0]^T, [0,1,0]^T, [0,0,1]^T\}$ that places all spectral replicates on the perimeter of $[-\pi,\pi) \times [-\pi,\pi)$, while avoiding $[-\pi,0]^T$ and $[0,-\pi]^T$, can only support *two* distinct colors. While we show in Section 5.3 how panchromatic designs can overcome this restriction, those that have emerged to date (including four-color CFAs) fail to satisfy the above two conditions.

5.2.2 Aliased Sensor Data and Demosaicking

Because the suboptimal CFA designs detailed above are prone to aliasing, linear reconstruction methods no longer suffice as the spectral support of $X_g(\omega)$ increases. Reconstruction is then an ill-posed problem, meaning that stronger assumptions about the signal are needed to recover the full-color image from aliased sensor data. To this end, the most common approach is to invoke the principle that *local* image features are sparse in some canonical representation. One explicit form of this principle is *directionality*—the notion that image features are assumed to be oriented in one direction, and thus that the energy of the corresponding local Fourier coefficients is concentrated accordingly. If X_g is sparse in the direction parallel to an image feature orientation, then aliasing can in turn be avoided; this principle is exploited either explicitly or implicitly by many state-of-the-art demosaicking methods [9], [10], [11], [12], [13]. In a similar manner, under a transformation that is local in both space and frequency, the signal energy may be assumed to be compressed into a few transform coefficients; regularization in the transform domain then helps to recover the full color image [8], [12]. However, demosaicking methods that exploit these assumptions are usually highly nonlinear and computationally demanding. Indeed, effective detection of image feature orientation (especially under the influence of noise) is an active area of research, and the determination of local image statistics requires additional computation. Moreover, subsequent interpolation steps are tightly coupled to estimates of feature directionality; this type of nonlinearity is effectively a data-driven switching mechanism that is expensive to implement in ASIC or DSP hardware. On the other hand, wavelet- and filterbank-based methods often employ iterative reconstruction schemes that may not easily be implemented in portable imaging devices.

The difficulties posed by nonlinear reconstruction methods are especially evident in today's digital video camera architectures. In order to meet the required frame rate with limited computational complexity, for example, it is common to implement demosaicking using methods such as bilinear interpolation that fail to yield satisfactory results. Other processing schemes may introduce pixel flickering artifacts, for instance, interframe oscillation or toggling of pixel colors caused by the susceptibility of edge-detection techniques to noise. Finally, nonlinear demosaicking methods are themselves subject to perturbations due to noise. Although simultaneous image denoising and interpolation methods have emerged in recent years (see, for example, Reference [12]), the difficulties of characterizing noise statistics after nonlinear demosaicking often render stand-alone image denoising methods ineffective. In contrast, the statistics of noise that undergoes only linear processing remain highly tractable, suggesting that a combination of denoising and demosaicking may indeed be possible.

5.3 Spatio-Spectral Color Filter Array Design

By simultaneously considering both the spectral support of luminance and chrominance components, and the spatial sampling requirements of the image acquisition process, we may conceive of a new paradigm for designing CFAs. With robustness to aliasing achieved via ensuring that spectral replicates lie along the perimeter of the Fourier-domain region $[-\pi,\pi) \times [-\pi,\pi)$ while avoiding the values $[-\pi,0]^T$ and $[0,-\pi]^T$ along the horizontal and vertical axes, our CFA design methodology aims to preserve the integrity of color images by way of subsampled sensor data. Images acquired in this manner are easily manipulated, enjoy simple reconstruction schemes, and admit favorable computation-quality trade-offs with the potential to ease subsequent processing in the imaging pipeline [14], [15].

5.3.1 Frequency-Domain Specification of Color Filter Array Designs

Let $0 \le c_r(\boldsymbol{n}), c_g(\boldsymbol{n}), c_b(\boldsymbol{n}) \le 1$ indicate the CFA projection values at a particular spatial location, where $c_r(\boldsymbol{n}), c_g(\boldsymbol{n}), c_b(\boldsymbol{n})$ now assume continuous values and hence represent a *mixture* of prototype channels. With the additional constraint that $c_r + c_g + c_b = \gamma$, it follows in analogy to Equation 5.1 that

$$y(\boldsymbol{n}) = \boldsymbol{c}(\boldsymbol{n})^T \boldsymbol{x}(\boldsymbol{n}) = \begin{bmatrix} c_r(\boldsymbol{n}) & \gamma & c_b(\boldsymbol{n}) \end{bmatrix} \begin{bmatrix} x_\alpha(\boldsymbol{n}) \\ x_g(\boldsymbol{n}) \\ x_\beta(\boldsymbol{n}) \end{bmatrix},$$

and we may determine the modulation frequencies of difference channels $x_\alpha(\boldsymbol{n})$ and $x_\beta(\boldsymbol{n})$ by our choice of $c_r(\boldsymbol{n})$ and $c_b(\boldsymbol{n})$. Recalling Equation 5.3, we seek choices such that Fourier transforms of the frequency-modulated difference images $X_\alpha(\boldsymbol{\omega}-\boldsymbol{\lambda}_r), X_\beta(\boldsymbol{\omega}-\boldsymbol{\lambda}_b)$ are maximally separated from the baseband spectrum $X_g(\boldsymbol{\omega})$.

In the steps outlined below, we first specify candidate carrier frequencies $\{\boldsymbol{\tau}_i\}$ and corresponding weights $s_i, t_i \in \mathbb{C}$ for color filters $c_r(\boldsymbol{n})$ and $c_b(\boldsymbol{n})$. Recalling that for constants ν, κ we have that $\mathscr{F}\{\kappa c_r + \nu\}(\boldsymbol{\omega}) = \kappa \mathscr{F} c_r(\boldsymbol{\omega}) + \nu\delta(\boldsymbol{\omega})$, we see that it is possible to manipulate our candidate color filter values until the realizability condition $0 \le c_r(\boldsymbol{n}), c_g(\boldsymbol{n}), c_b(\boldsymbol{n}) \le 1$ is met. This notion leads to the following algorithm for frequency-domain specification of color filter array designs (with $\bar{\cdot}$ denoting complex conjugation, and Figure 5.4 illustrating the algorithmic steps):

ALGORITHM 5.1 Frequency-domain color filter array design.

1. Specify initial values $\{\boldsymbol{\tau}_i, s_i, t_i\}$. Set modulation frequencies:

$$c_r^{(0)} = \mathscr{F}^{-1} \sum_i s_i \delta(\boldsymbol{\omega}+\boldsymbol{\tau}_i) + \bar{s}_i \delta(\boldsymbol{\omega}-\boldsymbol{\tau}_i)$$
$$c_b^{(0)} = \mathscr{F}^{-1} \sum_i t_i \delta(\boldsymbol{\omega}+\boldsymbol{\tau}_i) + \bar{t}_i \delta(\boldsymbol{\omega}-\boldsymbol{\tau}_i).$$

2. Subtract a constant $\nu_r = \min c_r^{(0)}(\boldsymbol{n}), \nu_b = \min c_r^{(0)}(\boldsymbol{n})$ (non-negativity):

$$c_r^{(1)} = c_r^{(0)} - \nu_r, \qquad c_b^{(1)} = c_b^{(0)} - \nu_b.$$

3. Scale by $\kappa = (\max_{\boldsymbol{n}} c_r^{(1)}(\boldsymbol{n}) + c_b^{(1)}(\boldsymbol{n}))^{-1}$ (convex combination):

$$c_r^{(2)} = \kappa c_r^{(1)}, \qquad c_b^{(2)} = \kappa c_b^{(1)}.$$

4. Find green: $c_g^{(2)} = 1 - c_r^{(2)} - c_b^{(2)}$.
5. Scale by $\gamma = (\max\{c_r^{(2)}(\boldsymbol{n}), c_g^{(2)}(\boldsymbol{n}), c_b^{(2)}(\boldsymbol{n})\})^{-1}$:

$$c_r = \gamma c_r^{(2)}, \qquad c_g = \gamma c_g^{(2)}, \qquad c_b = \gamma c_b^{(2)}.$$

In the first step, candidate carrier frequencies are determined by taking the inverse Fourier transform of $\delta(\boldsymbol{\omega} \pm \boldsymbol{\tau}_i)$. The conjugate symmetry in this step guarantees a real-valued color filter array; in general, however, the resultant design is not physically realizable (points in Figure 5.4a fall outside of the first quadrant, for example). Constants ν_r, ν_b are then subtracted to ensure non-negativity of color filters (Figure 5.4b). A scaling by κ and computation of the green component in the next two steps projects candidate values onto the unit simplex, ensuring convexity and a maximum component value of unity (Figure 5.4c and Figure 5.4d). Finally, multiplication by γ maximizes the quantum efficiency of the color filters (Figure 5.4e). The resultant CFA is physically realizable, with observed spectral data Y given by the sum of baseband components and modulated versions of X_α and X_β:

$$\begin{aligned} Y(\boldsymbol{\omega}) = {} & \gamma X_g(\boldsymbol{\omega}) - \gamma\kappa\nu_r X_\alpha(\boldsymbol{\omega}) - \gamma\kappa\nu_b X_\beta(\boldsymbol{\omega}) \\ & + \gamma\kappa \sum_i \{s_i X_\alpha + t_i X_\beta\}(\boldsymbol{\omega}+\boldsymbol{\tau}_i) + \{\bar{s}_i X_\alpha + \bar{t}_i X_\beta\}(\boldsymbol{\omega}-\boldsymbol{\tau}_i). \end{aligned}$$

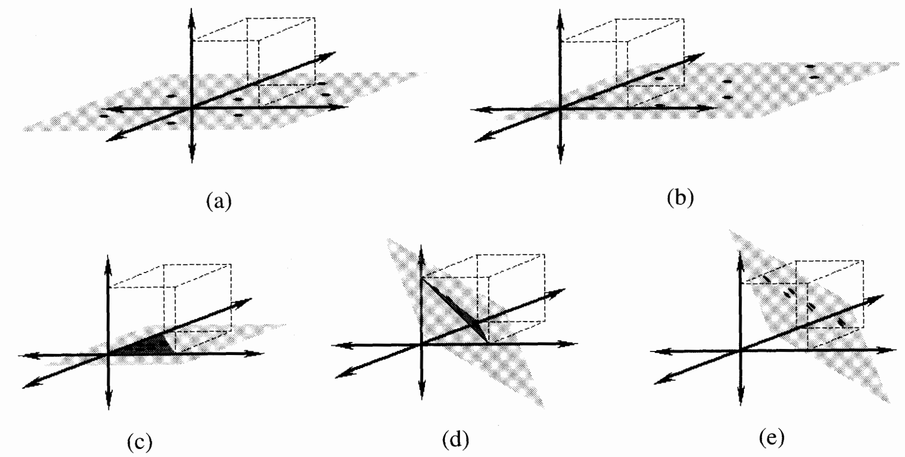


FIGURE 5.4
Color filter array design visualized in Cartesian coordinates (c_r, c_b, c_g), with the dotted cube representing the space of physically realizable color filters ($0 \le c_r(\boldsymbol{n}), c_g(\boldsymbol{n}), c_b(\boldsymbol{n}) \le 1$). Steps 1 to 5 in Algorithm 5.1 are shown as (a) to (e), respectively.

This approach enables the specification of CFA design parameters directly in the Fourier domain, by way of carrier frequencies $\{\boldsymbol{\tau}_i\}$ and weights $\{s_i, t_i\}$. In doing so, we ensure that nonzero carrier frequencies lie along the perimeter of $[-\pi,\pi) \times [-\pi,\pi)$, while avoiding the values $[-\pi,0]^T$ and $[0,-\pi]^T$ as desired.

5.3.2 Analysis and Design Trade-Offs

In this section, some notable features of the above CFA design strategy are considered; readers are referred to Reference [15] for a thorough analysis of design trade-offs. We first note that CFA designs resulting from Algorithm 5.1 are panchromatic, with the resultant filters comprising a mixture of red, green, and blue colors at each spatial location. As color filters are commonly realized by pigment layers of cyan, magenta, and yellow dyes over an array of pixel sensors (i.e., subtractive colors) [18], designs for which $\gamma > 1$ suggest improved quantum efficiency. Furthermore, it becomes easier to control for sensor saturation, as the relative quantum efficiency at each pixel location is approximately uniform ($c_r + c_g + c_b = \gamma$). We also note that the space of feasible initialization parameters $\{\boldsymbol{\tau}_i, s_i, t_i\}$ corresponding to Algorithm 5.1 is underconstrained, offering flexibility in optimizing the CFA design according to other desirable characteristics such as demosaicking complexity, pattern periodicity, resilience to illuminant spectrum, and numerical stability [15].

Our design strategy assumes bandlimitedness of the difference images x_α and x_β, and therefore its robustness hinges on how well this claim holds in various practical situations (e.g., under changes in illuminant). Even as the bandwidths of the modulated difference spectra grow, the increased distance between these channels and the baseband component serves to reduce the risk of aliasing, effectively increasing the spatial resolution of the imaging sensor. Consequently, local interpolation methods are less sensitive to the directionality of image features, and a linear demosaicking method then suffices for many applications.

As described earlier, linearization of the demosaicking step is attractive for several reasons: it can be coded more efficiently in DSP chips, it eliminates the temporal toggling pixel problems in video sequences, it provides a more favorable setup for deblurring, and it yields more tractable noise and distortion characterizations.

5.4 Linear Demosaicking via Demodulation

In this section, we show that the processing pipeline of a typical digital camera can be exploited to greatly reduce the complexity of reconstruction methods [14]. Suppose the conjugate modulation sequences $c_\alpha(\boldsymbol{n}) = c_r^{(0)}(\boldsymbol{n})^{-1}$ and $c_\beta(\boldsymbol{n}) = c_b^{(0)}(\boldsymbol{n})^{-1}$ exist;[3] when these sequences are orthogonal, the modulated signal can be recovered via a multiplication by the conjugate carrier frequency followed by a lowpass filter. Assuming mutual exclusivity of the supports of X_g, X_α, and X_β in the frequency domain, we expect an exact reconstruction according to

$$\hat{\boldsymbol{x}}(\boldsymbol{n}) = \begin{bmatrix} \hat{x}_r(\boldsymbol{n}) \\ \hat{x}_g(\boldsymbol{n}) \\ \hat{x}_b(\boldsymbol{n}) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1/(\gamma\kappa) & 0 & 0 \\ \nu_r/\gamma & 1/\gamma & \nu_b/\gamma \\ 0 & 0 & 1/(\gamma\kappa) \end{bmatrix} \begin{bmatrix} h_\alpha * \{c_\alpha y\} \\ h_g * y \\ h_\beta * \{c_\beta y\} \end{bmatrix}, \tag{5.4}$$

where $*$ denotes the discrete convolution operator, and the passbands of lowpass filters h_α, h_g, h_β match the respective bandwidths of the signals x_α, x_g, x_β.

Given the mutual exclusivity of the signals x_α, x_g, x_β in the Fourier domain, we assume $c_r^{(0)} h_\alpha + h_g + c_b^{(0)} h_\beta = \delta$, where $\delta(\boldsymbol{n})$ is again a Kronecker delta function. Using the linearity and modulation properties of convolution, we obtain:

$$\begin{aligned} h_g * y &= (\delta - c_r^{(0)} h_\alpha - c_b^{(0)} h_\beta) * y \\ &= y - \{c_r^{(0)} h_\alpha\} * y - \{c_b^{(0)} h_\beta\} * y \\ &= y - c_r^{(0)}\{h_\alpha * \{c_\alpha y\}\} - c_b^{(0)}\{h_\beta * \{c_\beta y\}\}. \end{aligned}$$

The demodulation in Equation 5.4 in turn takes the following simplified form:

$$\begin{aligned} \hat{\boldsymbol{x}}(\boldsymbol{n}) &= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1/(\gamma\kappa) & 0 & 0 \\ \nu_r/\gamma & 1/\gamma & \nu_b/\gamma \\ 0 & 0 & 1/(\gamma\kappa) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -c_r^{(0)}(\boldsymbol{n}) & 1 & -c_b^{(0)}(\boldsymbol{n}) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_\alpha * \{c_\alpha y\} \\ y \\ h_\beta * \{c_\beta y\} \end{bmatrix} \\ &= \begin{bmatrix} 1/(\gamma\kappa) + \nu_r/\gamma - c_r^{(0)}(\boldsymbol{n})/\gamma & 1/\gamma & \nu_b/\gamma - c_b^{(0)}(\boldsymbol{n})/\gamma \\ \nu_r/\gamma - c_r^{(0)}(\boldsymbol{n})/\gamma & 1/\gamma & \nu_b/\gamma - c_b^{(0)}(\boldsymbol{n})/\gamma \\ \nu_r/\gamma - c_r^{(0)}(\boldsymbol{n})/\gamma & 1/\gamma & 1/\gamma + \nu_b/\gamma - c_b^{(0)}(\boldsymbol{n})/\gamma \end{bmatrix} \begin{bmatrix} h_\alpha * \{c_\alpha y\} \\ y \\ h_\beta * \{c_\beta y\} \end{bmatrix}. \end{aligned} \tag{5.5}$$

The first term in Equation 5.5 is a 3×3 matrix multiplication (a completely pixelwise operation), whereas the spatial processing component is contained in its second term. In the usual layout of a digital camera architecture, a color conversion module follows immediately, converting the tristimulus output from demosaicking to a standard color space representation through another 3×3 matrix multiplication on a per-pixel basis. The two cascading matrix multiplication steps can therefore be performed together in tandem, with the combined matrix computed offline and preloaded into the camera system.

[3] In this chapter, we do not discuss cases in which there are zeros; however, the results presented here generalize easily to such cases via an appropriate multiplicative constant.

Given sufficient separation of the modulated signals in the frequency domain, crudely designed low-pass filters suffice for the reconstruction task. Suppose we choose to implement Equation 5.5 using a separable two-dimensional odd-length triangle filter — a linear-phase filter with a modest cutoff in the frequency domain. Four cascading boxcar filters can be used to implement a filter of length $2q - 1$ having the following Z transform, with Z_1 and Z_2 corresponding to delay lines in horizontal and vertical directions, respectively:

$$H_\alpha(\boldsymbol{Z}) = H_\beta(\boldsymbol{Z}) = \left(\frac{1 - Z_1^{-q}}{1 - Z_1^{-1}}\right)\left(\frac{1 - Z_1^{-q}}{1 - Z_1^{-1}}\right)\left(\frac{1 - Z_2^{-q}}{1 - Z_2^{-1}}\right)\left(\frac{1 - Z_2^{-q}}{1 - Z_2^{-1}}\right). \tag{5.6}$$

The computational complexity of the above system is eight adders for h_α and h_β each. Moreover, in 4×4 repeating CFAs, the carrier frequencies $c_r^{(0)}$ and $c_b^{(0)}$ are often proportional to sequences of ± 1's (and by extension, c_α and c_β also). In this case, the multiplication by -1 before addition in Equation 5.6 simply replaces adders with subtracters, which is trivial to implement. The overall per-pixel complexity of the demodulation demosaicking in Equation 5.5 is therefore comparable to that of bilinear interpolation (16 add/subtract operations per full pixel), despite its state-of-the-art image quality performance.

5.5 Examples and Analysis

In this section we provide several examples of CFA designs and analyze their performance. These designs, shown in Figure 5.5 and detailed in Table 5.1, were generated in the spirit of Algorithm 5.1 by employing an exhaustive search over a restricted parameter space $\{\tau_i, s_i, t_i\}$ [15]. Though some CFAs in Figure 5.5 have rectangular geometries, we see that nevertheless every pixel sensor has an equal number of neighboring colors, a condition that helps mitigate cross-talk noise due to leakages of photons and electrons. Their

TABLE 5.1
Example CFA patterns specified in terms of parameter values $\{\tau_i, s_i, t_i\}$.

pattern		$i=0$	$i=1$	pattern		$i=0$	$i=1$	pattern		$i=0$	$i=1$
	τ_i	$(\pi, \frac{\pi}{2})$	(π, π)		τ_i	$(\pi, \frac{2\pi}{3})$	$(\frac{2\pi}{3}, \pi)$		τ_i	$(\pi, \frac{\pi}{2})$	(π, π)
A	red s_i	$1+1j$	1	C	red s_i	$1j$	$1j$	E	red s_i	$1+1j$	1
	blue t_i	$1+1j$	-1		blue t_i	$1j$	$-1j$		blue t_i	$1+1j$	-1
	τ_i	$(\pi, \frac{\pi}{2})$	(π, π)		τ_i	$(\pi, \frac{\pi}{3})$	(π, π)		τ_i	$(\pi, \frac{\pi}{2})$	(π, π)
B	red s_i	$1+1j$	0	D	red s_i	$3+4j$	1	F	red s_i	$1+1j$	0
	blue t_i	0	1		blue t_i	$3-4j$	1		blue t_i	0	1

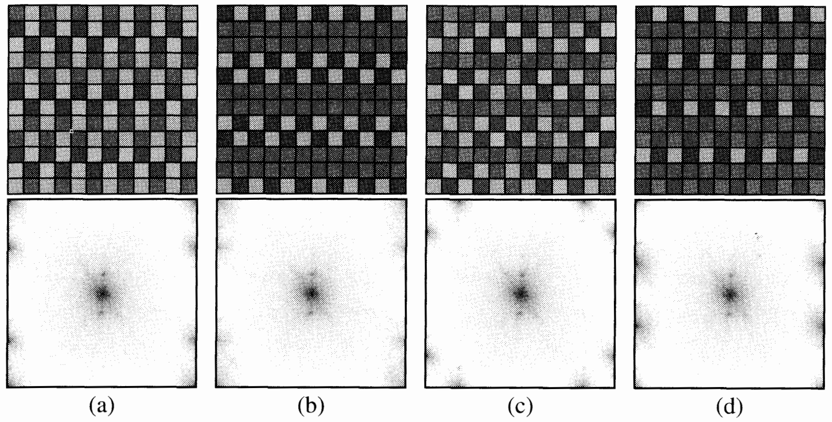


FIGURE 5.5 (See color insert.)

Proposed CFAs (top) and resultant log-magnitude spectra (bottom) of a typical color image (i.e., image *flower* here). Color coding is used as in Figure 5.3 to distinguish components X_α, X_g, and X_β. Subfigures correspond to: (a) *pattern A*, (b) *pattern B*, (c) *pattern C*, and (d) *pattern D*.

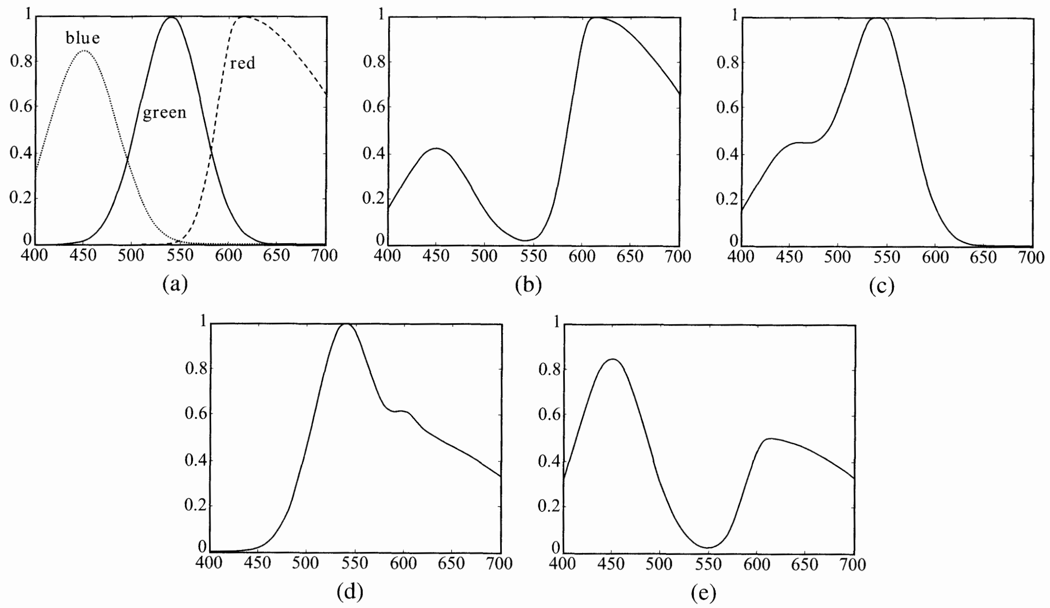


FIGURE 5.6

Spectral sensitivity characteristics (a) of a typical Sony CCD sensor [19], and (b-e) the corresponding *pattern A* color filters derived from these characteristics.

designs are given in Table 5.1 as combinations of prototype red, green, and blue filters; the precise color specifications used in subsequent demosaicking experiments were derived from a popular Sony CCD quantum efficiency function [19] shown in Figure 5.6a. The resultant spectral responses, shown in Figure 5.6b, may be implemented using subtractive color pigments such as cyan, magenta, and yellow.

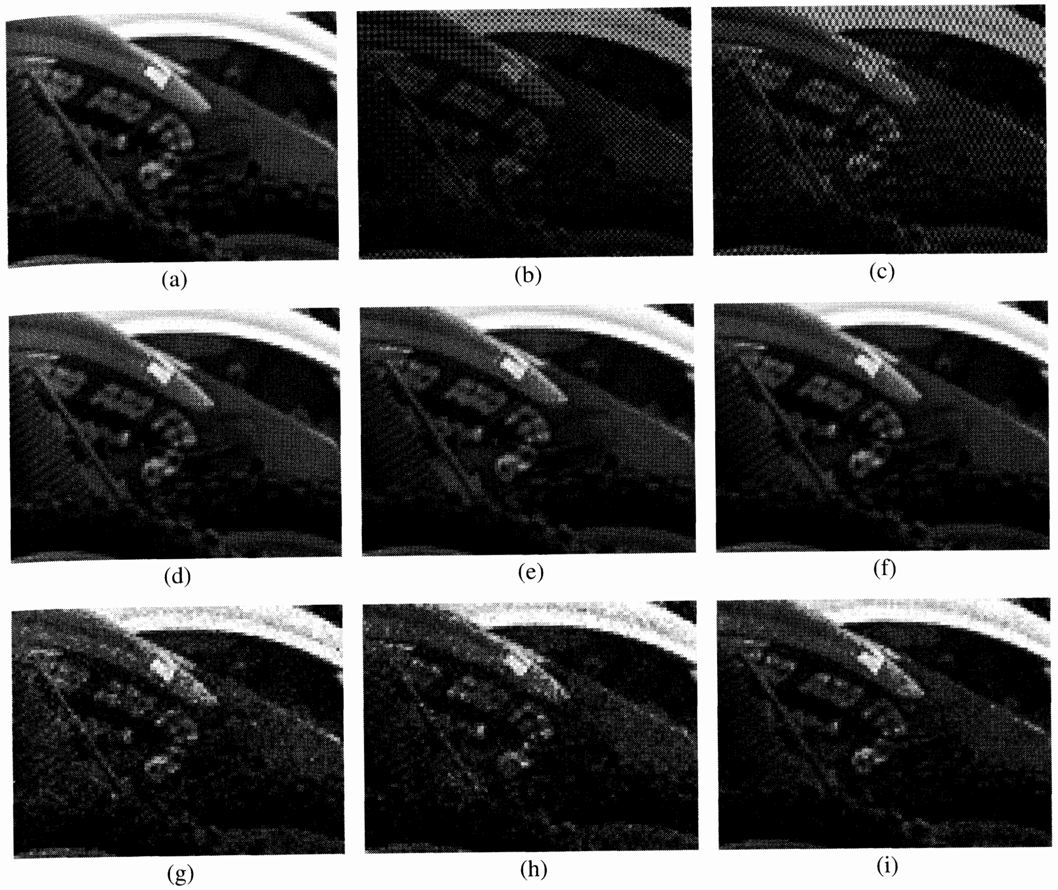


FIGURE 5.7 (See color insert.)

Bike image sensor data (top row), with nonlinear and linear reconstruction methods shown for the case of clean (middle row) and noisy (bottom row) sensor data. Individual images correspond to: (a) original image, (b) Bayer CFA sampling, (c) *pattern A* sampling, (d) nonlinear Bayer reconstruction [8], (e) linear Bayer reconstruction, (f) linear *pattern A* reconstruction, (g) noisy nonlinear Bayer reconstruction, (h) noisy linear Bayer reconstruction, and (i) noisy *pattern A* linear reconstruction.

In comparing Figure 5.3 and Figure 5.5, we see that in the latter case spectral copies of x_α and x_β are placed farther from the Cartesian axes and the origin, thus achieving a better separation of channels in the Fourier domain. The implications of this design improvement may be seen in the demosaicking examples of Figure 5.7; while demosaicking performance is both algorithm- and CFA-dependent, we may consider state-of-the-art methods for demosaicking Bayer CFA data along with the linear reconstruction methodology outlined in Section 5.4, using the well-known *bike* test image shown in Figure 5.7a.

To this end, Figure 5.7b and Figure 5.7c show simulated sensor data $y(\boldsymbol{n}) = \boldsymbol{c}(\boldsymbol{n})^T \boldsymbol{x}(\boldsymbol{n})$ for the *bike* image $\boldsymbol{x}(\boldsymbol{n})$, acquired under $\boldsymbol{c}(\boldsymbol{n})$ representing the Bayer CFA and *pattern A* of Figure 5.5, respectively. Figure 5.7d to Figure 5.7f show demosaicked images corresponding respectively to a reconstruction of a color image from Bayer CFA data using the iterative, nonlinear method of Reference [8], the linear demosaicking algorithm of Section 5.4, and the same linear method applied to the *pattern A* sampled data. This latter reconstruction is competitive with the nonlinear Bayer reconstruction of Figure 5.7d, and exhibits significantly reduced zipper artifacts. On the other hand, compared to the purely linear Bayer demosaicking shown in Figure 5.7e, the linear *pattern A* reconstruction shows a significant gain in fidelity for equal hardware resolution and computational cost. Finally, Figure 5.7g to Figure 5.7i demonstrate its improved resilience to noise, by way of showing the same three reconstructions applied to sensor data corrupted by simulated Poisson noise. Compared to the reconstructions using Bayer CFA data depicted in Figure 5.7g and Figure 5.7h, the *pattern A* linear reconstruction of Figure 5.7i renders contributions from signal-dependent noise far less noticeable.

5.6 Conclusion

By considering the interplay between color filter arrays and typical images, we have posed here the CFA design problem as one of simultaneously maximizing the spectral support of luminance and chrominance channels subject to their mutual exclusivity in the Fourier domain. From this perspective, current design practices were seen to be suboptimal: as image resolution increases, existing CFAs are prone to aliasing, linear reconstruction methods no longer suffice, stronger assumptions must be made about the underlying signal, and additional computational resources are needed to reconstruct the full-color image.

Key to our design paradigm was the notion that the measurement process, an inner product between the color filter array and the image data, induces a modulation in the frequency domain. To this end, we chose to modulate the chrominance spectra away from the baseband luminance channel, and in doing so we proposed a constructive method to design a physically realizable CFA by specifying these modulation frequencies directly. This method generates panchromatic CFA designs that mitigate aliasing and admit favorable computation-quality trade-offs. As we have shown, our corresponding linear demosaicking method yields state-of-the-art performance with an order of complexity comparable to that of bilinear interpolation.

References

[1] K. Parulski and K.E. Spaulding, *Digital Color Imaging Handbook*, ch. Color image processing for digital cameras, G. Sharma (ed.), Boca Raton, FL: CRC Press, 2002, pp. 728–757.

[2] R. Ramanath, W.E. Snyder, Y. Yoo, and M.S. Drew, "Color image processing pipeline," *IEEE Signal Processing Magazine*, vol. 22, no. 1, pp. 34–43, January 2005.

[3] R. Lukac and K.N. Plataniotis, *Color Image Processing: Methods and Applications*, ch. Single-sensor camera image processing, R. Lukac and K.N. Plataniotis (eds.), Boca Raton, FL: CRC Press / Taylor & Francis, October 2006, pp. 363–392.

[4] B.E. Bayer, "Color imaging array," U.S. Patent 3 971 065, July 1976.

[5] S. Yamanaka, "Solid state color camera," U.S. Patent 4 054 906, August 1977.

[6] M. Parmar and S.J. Reeves, "A perceptually based design methodology for color filter arrays," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Montreal, Canada, May 2004, vol. III, pp. 473–476.

[7] R. Lukac and K.N. Plataniotis, "Color filter arrays: Design and performance analysis," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 4, pp. 1260–1267, November 2005.

[8] B.K. Gunturk, J. Glotzbach, Y. Altunbasak, R.W. Schafer, and R.M. Mersereau, "Demosaicking: Color filter array interpolation in single chip digital cameras," *IEEE Signal Processing Magazine*, vol. 22, no. 1, pp. 44–54, January 2005.

[9] K. Hirakawa and T.W. Parks, "Adaptive homogeneity-directed demosaicing algorithm," *IEEE Transactions on Image Processing*, vol. 14, no. 3, pp. 360–369, March 2005.

[10] D. Alleysson, S. Süsstrunk, and J. Hérault, "Linear demosaicing inspired by the human visual system," *IEEE Transactions on Image Processing*, vol. 14, no. 4, pp. 439–449, April 2005.

[11] R. Lukac and K.N. Plataniotis, "Universal demosaicking for imaging pipelines with an RGB color filter array," *Pattern Recognition*, vol. 38, no. 11, pp. 2208–2212, November 2005.

[12] K. Hirakawa and T.W. Parks, "Joint demosaicking and denoising," *IEEE Transactions on Image Processing*, vol. 15, no. 8, pp. 2146–2157, August 2006.

[13] E. Dubois, "Filter design for adaptive frequency-domain Bayer demosaicking," in *Proceedings of the IEEE International Conference on Image Processing*, Atlanta, GA, USA, October 2006, pp. 2705–2708.

[14] K. Hirakawa and P.J. Wolfe, "Second-generation CFA and demosaicking design," in *Proceedings of the IS&T/SPIE 19th Annual Symposium on Electronic Imaging*, San Jose, CA, USA, January 2008.

[15] K. Hirakawa and P.J. Wolfe, "Spatio-spectral color filter array design for enhanced image fidelity," in *Proceedings of the IEEE International Conference on Image Processing*, San Antonio, TX, USA, September 2007, vol. 2, pp. 81–84. Extended version submitted to *IEEE Transactions on Image Processing*, October 2007.

[16] T. Kijima, H. Nakamura, J. Compton, and J. Hamilton, "Image sensor with improved light sensitivity," U.S. Patent Application 2007 0 177 236, August 2007.

[17] D.M. Coppola, H.R. Purves, A.N. McCoy, and D. Purves, "The distribution of oriented contours in the real world," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 7, pp. 4002–4006, March 1998.

[18] R. Ramanath and W.E. Snyder, "Adaptive demosaicking," *Journal of Electronic Imaging*, vol. 12, no. 4, pp. 633–642, October 2003.

[19] Sony Corporation, "Diagonal 6 mm (type 1/3) progressive scan CCD image sensor with square pixel for color cameras," http://products.sel.sony.com/semi/PDF/ICX204AK.pdf, 2004.