Sample selection and reconstruction for array based multispectral imaging

Except where reference is made to the work of others, the work described in this dissertation is my own or was done in collaboration with my advisory committee. This dissertation does not include proprietary or classified information.

Manu Parmar

Certificate of Approval:

Thomas S. Denney, Jr Professor Electrical and Computer Engineering	Stanley J. Reeves, Chair Professor Electrical and Computer Engineering
Jitendra K. Tugnait Professor Electrical and Computer Engineering	John Y. Hung Professor Electrical and Computer Engineering

George T. Flowers
Interim Dean
Graduate School

Sample selection and reconstruction for array based multispectral imaging

Manu Parmar

A Dissertation

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the

Degree of

Doctor of Philosophy

Auburn, Alabama
May 10, 2007

Sample selection and reconstruction for array based multispectral imaging

Manu Parmar

Signature of Author

Date of Graduation

Vita

Manu Parmar was born in Mumbai (formerly, Bombay), India, in 1978. He received the B.E. degree in electrical engineering from the Government College of Engineering, Pune, in 2000. He traveled to the United States in 2000 to join the M.S. program in electrical engineering at Auburn University and received the M.S. degree in 2002. He joined the Ph.D. program in electrical engineering at Auburn University in 2002. His research interests are in in the areas of digital imaging, color restoration, optimal image acquisition, and multispectral imaging.

Dissertation Abstract

Sample selection and reconstruction for array based multispectral imaging

Manu Parmar

Doctor of Philosophy, May 10, 2007
(M.S., Auburn University, Auburn, 2002)
(B.E., Government College of Engineering, Pune University, 2000)

121 Typed Pages

Directed by Stanley J. Reeves

In this work we address the problem of acquisition of multispectral images in a sampled form and the subsequent processing of the acquired signal. The problem is relevant in the context of color imaging in digital cameras, and increasingly, in the field of hyperspectral imaging as applied to remote-sensing and target recognition. The scope of this work encompasses a broad swath across image processing problems and includes: image acquisition, in the problem of optimally selecting sampling rates and patterns of multiple channels; image reconstruction, in the reconstruction of the sparsely sampled data; image restoration, in obtaining an estimate of the true scene from noisy data; and finally, image enhancement and representation, in the problem of presenting the reconstructed image in a color-space that allows for transformations that achieve best perceived quality.

Acquisition of multispectral images in the simplest form entails either the use of multiple sensor arrays to sample separate spectral bands in a scene, or the use of a single sensor array with a mechanism that switches overlaying band-pass filters. Due to the nature of the acquisition process, both these methods suffer from shortcomings in terms of weight, cost,

time of acquisition, etc. An alternative scheme widely in use only uses one sensor array to sample multiple bands. An array of filters, referred to as a *mosaic*, is overlaid on the sensor array such that only one color is sampled at a given pixel location. The full color image is obtained during a subsequent reconstruction step commonly referred to as *demosaicking*. This scheme offers advantages in terms of cost, weight, mechanical robustness and the elimination of the related post-processing step since registration in this case is exact.

Three main issues need to be addressed in such a scheme, viz., the shape and arrangement of the sampling pattern, selection of the sensitivities of the spectral filters, and the design of the reconstruction algorithm. Each of the above problems is contingent on multiple factors. Sensor sampling patterns are constrained by the limitations of electronic devices and manufacturing processes, spectral sensitivities are affected by the material properties of the colors painted on the array to form filters, and the reconstruction methods are limited by computational resources.

In this research, we address the above problems from a signal processing perspective and attempt to develop parametric algorithms that can accommodate external limitations and constraints. We have developed methodologies for the selection of optimal sampling patterns that will allow for ordered, repeated array blocks. In addition we have developed an algorithm for demosaicking of CFA data based on Bayesian techniques. We have also proposed a formulation for the selection of optimal spectral sensitivities for individual color filters.

Acknowledgments

The culmination of this work, and the entire enterprise of a completed Ph.D., is due to the continued support of many individuals. First, I would like to thank my advisor Dr. Stanley J. Reeves for the steady support, constant encouragement, and the super advise throughout the course of this work. I am deeply indebted to him for his great patience with me and my work, his glee at discovering new problems and their solutions, and the amazing example he sets with his approach to work, research, signal processing, and life in general.

I am indebted to Dr. John Y. Hung, my M.S. thesis advisor for his support and his exceptional qualities as a guide and mentor. I am thankful to Dr. Thomas S. Denney for serving on my graduate committee, reviewing my work, and his wonderful ideas and enthusiasm in myriad areas. I thank him for supporting me and my work on many occasions along the line. I also thank Dr. Jitendra K. Tugnait for serving on my graduate committee, Dr. Victor Nelson, the ECE graduate program coordinator, Ms. Jo Ann Loden, and the wonderful people administering the department for all their help and their efforts in making the Ph.D. experience at Auburn a pleasant one. I am grateful for the continued financial support accorded me by the ECE department.

Last but not the least, I would like to thank my parents Cdr. Ram Singh and Mrs. Varinder Parmar and my sister Ekta Parmar for their love and inspiration. I am grateful for the love, support, and the wonderful company of my fiancée, Dr. Jhilmil Jain and the great bunch of friends I had the good fortune to meet in my time at Auburn.

Style manual or journal used Journal of Approximation Theory (together with the style known as "aums"). Bibliograpy follows van Leunen's *A Handbook for Scholars.*

Computer software used The document preparation package TEX (specifically LATEX) together with the departmental style-file `aums.sty`.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Statement of the problem

In digital image acquisition, the optical sensor is either a charge coupled device (CCD) or complementary metal oxide semiconductor (CMOS) device that is inherently monochromatic [1]. At a particular pixel location on a sensor-array, the photosensitive device integrates the incident energy over its entire spectrum to generate a charge that is indicative of intensity. The sensor array is thus capable of acquiring only a grayscale representation of the imaged scene. In color or multispectral imaging where different bands along the signal spectrum carry distinct information about the scene, the incident energy needs to be sampled along the wavelength range of interest. In these applications, color filter overlays (typically color pigment dyes) are used to cover the optical sensor-array such that the array only captures energy in a particular range of wavelengths. In consumer applications such as digital cameras, where the object is to produce a color image that may be displayed either on a display device (a cathode ray tube (CRT) or liquid crystal display (LCD)) or printed on paper, at least three color channels or bands must be sampled along the range of visible wavelengths.

Typically, digital color cameras sample three (with wavelengths centered around the red, green, and blue regions of the visible spectrum), or four (cyan, magenta, yellow, and white) bands while document scanners with special applications sometimes sample up to six bands. One way to achieve multi-band acquisition is to use multiple sensor-arrays overlaid

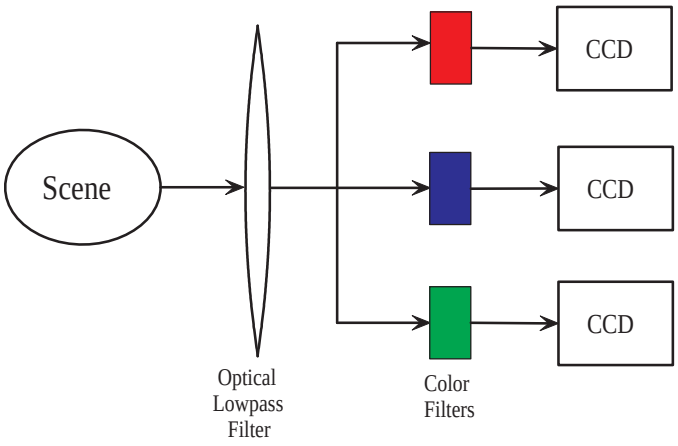


Figure 1.1: Image acquisition with multiple sensor-arrays

with color filters such that energy in a distinct band is incident on a particular sensor-array. In this case the number of sensor-arrays equals the number of bands to be sampled. Figure 1.1 illustrates such a scheme where three distinct channels (red, green, and blue) are sampled.

The optical sensor and its accompanying circuitry form a significant portion of the total cost of a camera (up to 25% [2]), and multi-sensor arrays are limited only to the most expensive digital cameras meant for professional use. Also, the beam-splitting arrangement, which typically is a dichroic prism, adds weight to the imager. Finally, since the color bands are acquired at different planes, an additional step of image registration is added to the imaging pipeline.

An alternative arrangement uses sequential color sampling. A full color image is produced by taking multiple exposures while switching the color filter cascaded with the sensor-array. The color filter in this case may be transmissive, dichroic, or a tunable liquid crystal filter. The main disadvantage in this case is that the system is extremely sensitive to motion. Only a few cameras targeted for studio use apply this technique.

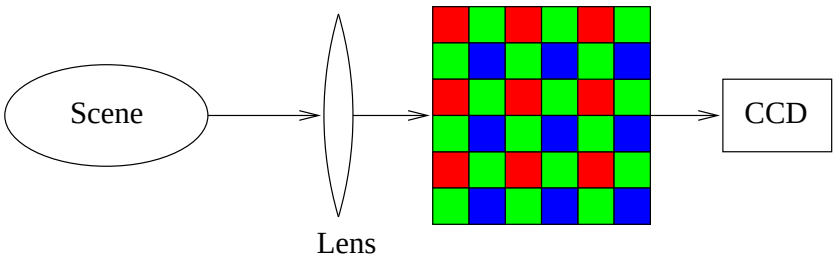


Figure 1.2: Image acquisition with a single sensor-array

Lately, manufacturers of consumer-level cameras (including Digital single lens reflex (SLR) cameras) and video cameras have predominantly used another alternative scheme that eliminates the limitations in the above schemes at the cost of added digital image processing. In this scheme only one sensor-array (Fig. 1.1) is used to acquire the full-color image. An array of filters, referred to as a mosaic, is overlaid on the sensor-array such that only one color is sampled at a given pixel location. The full color image is obtained during a subsequent reconstruction step commonly referred to as demosaicking. This scheme offers multiple advantages in terms of cost, weight, mechanical robustness, and the elimination of the image registration step since registration in this case is exact.

Such a mosaic-based sampling scheme for multispectral imaging presents a slew of new challenges and has attracted much research interest. The main issues that need to be addressed are:

- selection of the shape, arrangement, and sampling rates of mosaic filters to ensure optimal reconstruction
- selection of spectral sensitivities of the mosaic filters to ensure optimal performance (color reproduction in case of color cameras)

- the design of the reconstruction algorithm.

Each of the above problems is affected by multiple factors. The choice of a sampling scheme for the mosaic or color filter array (CFA) depends not only on the suitability of a particular pattern from the point of view of image reconstruction quality, but also on material properties of the color filter pigments and the semiconductor photosensitive elements. For example, it is desirable from an image quality perspective that the sampling pattern be random. This ensures that there are no reconstruction artifacts due to fixed patterns in the imaged scene. On the other hand, from a strict semiconductor devices perspective, it is desirable to have fixed repeated sampling patterns to prevent color inconsistencies due to cross-contamination among adjacent colors on the array. Demosaicking algorithms present trade-offs in terms of reconstruction quality and computational time. The selection of spectral sensitivities for the color filters is dependent on particular applications and viewing conditions for the final image.

1.2 Scope of the thesis

The research problems listed in Section 1.1 have been addressed to a large extent as independent problems in the literature. Recently, demosaicking algorithms have been a subject of extensive research and various new approaches have been used to reconstruct full-color images from sub-sampled data: projections on convex sets [3], wavelet domain processing [4], decision-theory [5], neural networks [6] etc. Traditional image reconstruction techniques have also been used to address the problem of demosaicking [7, 8, 9]. The problem of selection of spectral sensitivities has been addressed only from the point of view of color reproduction accuracy when areas of uniform colors are sampled [10, 11, 12, 13].

The problem of selection of sampling patterns has seen surprisingly little interest in the open literature while actual sampling schemes and algorithms used by camera manufacturers remain closely guarded proprietary information. Sampling schemes that have been patented or published in the literature are predominantly based on heuristics and on convenience of sensor-array read-out [14, 15, 16, 17, 18].

The unique problem of simultaneous spectral and spatial sampling presented by mosaic-based sampling schemes does not appear to be addressed in the open literature. In this work, we will propose methods to solve the above problems using unified approaches based on signal processing principles. In addition, the methods proposed are parametric and are flexible to the addition of constraints due to external factors.

Chapter 2 provides an overview of the fundamentals of human color vision and color image processing. The subject of colorimetry, the measurement of color, is introduced. The chapter also describes perceptually uniform color spaces that are commonly used to form measures for color reproduction accuracy. Also, generalized image formation models for multispectral image acquisition are detailed.

In Chapter 3 we present an algorithm for the recovery of color images from sparsely sampled, noisy data. The proposed algorithm is based on the Bayesian framework, which allows for the effective use of prior information in finding estimates for full-color *true* images. We present results for a number of test images and demonstrate the efficacy of the proposed algorithm.

In Chapter 4 we propose a method for the selection of optimal spectral sensitivities for the color filters used in the CFA mosaic. The proposed method is based on a unique joint spatial-spectral treatment that accounts for the simultaneous sampling in the spectral

and spatial domains, which is a characteristic of CFA-based imaging. Optimal color filter transmittance functions for a number of common CFA arrangements are derived and shown to perform better than standard RGB and CMY color filters in terms of both spatial reconstruction quality and color fidelity.

In Chapter 5 we propose two methods for the selection of sampling arrangements for CFAs. Both methods are based on optimization of criteria formed using standard image processing techniques and incorporate the effects of human color vision in their mathematical modeling.

In Chapter 6 we discuss the results obtained in previous chapters and summarize the problems yet to be solved.

Chapter 2

Background

One of the primary features desirable in a color imaging system is an ability to faithfully reproduce colors in a scene. The imaging system must also preserve the original colors during the transfer and further processing of the acquired signal among different devices (e.g., camera to printer to scanner). To this end, it is critical that the imaging system account for the mechanisms of color vision in the human visual system (HVS) and the limitations of various devices in the imaging system regarding the processing of color signals.

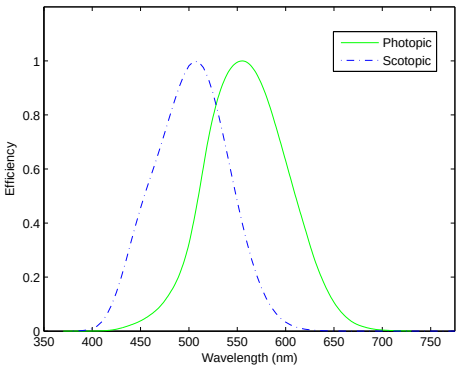
2.1 Color fundamentals and human color vision

The foundations of color theory and the spectral nature of visible light originate with the work of Isaac Newton. His experiments with prisms led to the understanding that the visible part of electromagnetic radiation (the wavelength region between λ_{min}=360 nm and λ_{max}=830 nm) can be decomposed into monochromatic components. It is important to understand that although it is common to refer to radiation or objects possessing certain colors, they only possesses the ability to trigger a sensation that is perceived as a particular color by the HVS. The appearance of a color is also dependent on viewing conditions, foreground and background color, spatial characteristics of the scene, and ambient light. In addition, color appearance is very subjective and differs widely among observers.

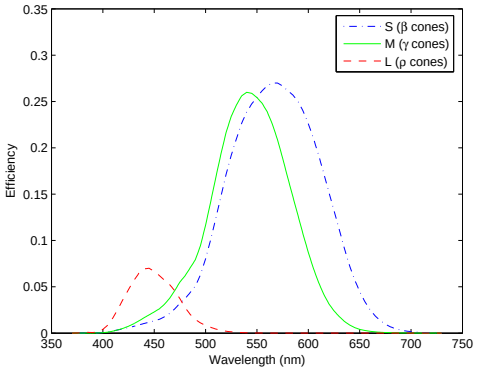
A consistent method for the specification and measurement of color (colorimetry) is not possible without an understanding of the HVS properties. Sharma and Trussel [19] summarize a history of the development of the understanding of color vision:

The wider acceptance of the wave theory of light paved the way for a better understanding of both light and color [20], [21]. Both Palmer [22] and Young [20] hypothesized that the human eye has three receptors, and the difference in their responses contributes to the sensation of color. However, Grassmann [23] and Maxwell [24] were the first to clearly state that color can be mathematically specified in terms of three independent variables. Grassmann also stated experimental laws of color matching that now bear his name [[25], p. 118]. Maxwell [26], [27] demonstrated that any additive color mixture could be matched by proper amounts of three primary stimuli, a fact now referred to as trichromatic generalization or trichromacy. Around the same time, Helmholtz [28] explained the distinction between additive and subtractive color mixing and explained trichromacy in terms of spectral sensitivity curves of the three "color sensing fibers" in the eye.

It has been determined that the the human retina has two kinds of receptors, viz., rods and cones. The primary function of the rods is to provide monochromatic vision under low illumination levels (scotopic vision). A photosensitive pigment called rhodopsin that is sensitive primarily in the blue-green region of the spectrum is responsible for sensing radiation in the rods. Under normal illumination, the rods are saturated and the cones contribute to vision (photopic luminosity). There are three types of cones, each sensitive in a portion of the visible spectrum and thus named L (long wavelengths), M (medium wavelengths), and S (small wavelengths) types of cones. The spectral sensitivities of the cones have been determined through microspectrophotometric measurements [29], [30]. Figure 2.1(a) shows the luminous response of rods and the aggregated response of the three cones and represents



(a) Photopic and Scotopic luminosity functions for the HVS.



(b) Cone sensitivities corrected for peak optical transmittance of the ocular media and the internal QE of the photoisomerization

Figure 2.1: Sensitivities of human rods and cones.

luminosity under scotopic and photopic conditions respectively. Figure 2.1(b) shows the sensitivities of the three cones as determined by Stockman et al. [30] and is a representation of the color sensitivity of the HVS.

2.1.1 Trichromacy

The responses of the three cones to radiation emitted or reflected by a scene can be modeled by a linear system under fixed ambient conditions. For an incident radiation with a spectral distribution given by $f(\lambda)$, where λ represents wavelength, the responses of the three cones are given by the 3×1 vector

$$c_i = \int_{\lambda_{min}}^{\lambda_{max}} s_i(\lambda) f(\lambda)\, d\lambda, \qquad i = 1, 2, 3, \tag{2.1}$$

where, $s_i(\lambda)$ is the sensitivity of the i^{th} type of cone and the visible range of the electromagnetic spectrum is between $\lambda_{min} = 360$ nm and $\lambda_{max} = 830$ nm. The cone responses are a projection of the incident spectrum onto the three dimensional space spanned by the cone sensitivity functions of. This space is called the human visual subspace (HVSS). Although the actual colors perceived by the HVS are due to further non-linear processing by the human nervous system, under similar viewing conditions and ocular adaptation, a color may be approximately specified by the responses obtained at the three types of cones.

Equation (2.1) may be written in the discrete form as

$$c = S^T f \tag{2.2}$$

where c is a 3×1 vector such that each element of c specifies the response obtained at one type of cone, f is a $n \times 1$ vector that contains samples of the incident spectrum along the wavelength range, and S is a $n \times 3$ matrix. The columns of S are the sampled cone sensitivity functions. Typically, the visible range of wavelengths is sampled every 10 nm such that $n = 31$. A higher sampling rate is used in applications involving fluorescent lamps that have sharp spectral peaks [19].

Consider the vectors p_i, $i = 1, 2, 3$, such that $S^T p_i$ are linearly independent. The vectors p_i are said to constitute a set of color primaries. They are colorimetrically independent in that no one color can be formed as a linear combination of the other two and the matrix $S^T P$, where $P = [p_1 \quad p_2 \quad p_3]$, is non-singular. For any spectrum f, we define the vector $a(f) = (S^T P)^{-1} S^T f$ such that $S^T f = S^T P a(f)$. This implies that for any spectrum f, there exists a linear combination of the primaries that elicits the same response at the cones and thus matches the spectrum in color. This result, referred to as the principle of

trichromacy, is used in color matching experiments where the color of a particular spectrum is matched to the color obtained by a linear combination of a set of primaries.

Consider the set of unit-intensity orthonormal spectra given by $\{e_i\}_{i=1}^{n}$, where e_i is an $n \times 1$ vector having a 1 in the i^{th} position and zeros elsewhere. This set forms an orthonormal basis for all visible spectra. Let a_i be the vector that denotes the weights applied to a set of primaries to colorimetrically match the spectrum of e_i ($S^T = S^T P a_i$). For $A = [a_1, a_2, \cdots, a_n]^T$, we can form the color matching matrix A such that

$$S^T I_N = S^T P A^T. \tag{2.3}$$

The columns of A are referred to as the color matching functions (CMFs) associated with the primaries that are the columns of P. Any spectrum f may be represented as a weighted sum of $\{e_i\}_{i=1}^{N}$ as

$$f = \sum_{i=1}^{n} f_i e_i, \tag{2.4}$$

where f_i are the elements of f. From (2.3), it follows that the spectrum of f is colorimetrically matched by weighting the primaries with the elements of

$$\sum_{i=1}^{n} f_i e_i = A^T f. \tag{2.5}$$

$A^T f$ is a 3×1 vector that represents the relative intensities of the primaries P that match the color of f and is referred to as a tristimulus vector.

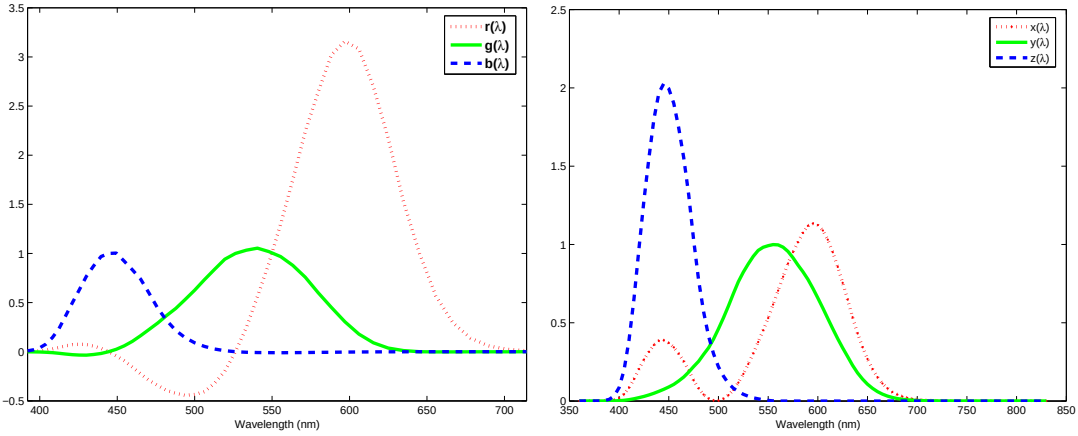
2.2 Colorimetry

To offer a consistent means of measurement and comparison, tristimulus values obtained from different experiments need to be defined with respect to a standard set of color matching functions (CMFs). The International Commission on Illumination, CIE, has defined a set of such CMFs that are used as standards in the industry. The CIE 1931 recommendations define a standard colorimetric observer by providing two equivalent sets of CMFs.

The CIE RGB CMFs ($\bar{r}(\lambda)$, $\bar{g}(\lambda)$, and $\bar{b}(\lambda)$) are associated with monochromatic primaries at wavelengths of 700.0, 546.1, and 435.8 nm respectively. The radiant intensities are adjusted so that the tristimulus values for the constant spectral power distribution (SPD) spectrum are equal. The CIE XYZ CMFs ($\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$) are obtained by a linear transformation of the CIE RGB CMFs, with the additional constraints that the XYZ CMFs have no negative values, the choice of $y(\lambda)$ is coincident with the luminous efficiency function (the relative sensitivity of the human eye at each wavelength [31]), and the tristimulus values are equal for the equi-energy spectrum. The CIE XYZ tristimulus values are most commonly used in color research and applications. The Y tristimulus value is referred to as the luminance and closely represents the perceived brightness or intensity of a radiant spectrum. The X and Z tristimulus values contain information about color or chrominance.

2.3 Perceptually uniform color spaces

A unit for color difference that is commonly used in color research is the *just noticeable difference* (JND). It has been established through psychovisual experiments that the JND is highly variable across the CIE XYZ space and the space is perceptually non-uniform [32].



(a) CIE $\bar{r}(\lambda)$, $\bar{r}(\lambda)$, $\bar{r}(\lambda)$ color matching functions

(b) CIE $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ color matching functions

Figure 2.2: CIE XYZ and CIE RGB color matching functions

Equal distances in the XYZ space do not correspond to equal differences in perceived color and thus Euclidian distance between two points in the XYZ space can not be used as a reliable objective measure of the perceived difference between two colors.

A perceptually uniform color space is highly desirable in defining tolerances in color reproduction systems and in objectively measuring the performance of various image processing algorithms. There has been much research directed at defining suitable perceptually uniform color spaces [31], [33]. The CIE has proposed two uniform color spaces for practical applications, viz., the CIE 1976 $L^*U^*V^*$ (CIELUV) space and the CIE 1976 $L^*a^*b^*$ (CIELAB) space. The CIELAB space is most commonly used in the imaging and printing industry as the preferred device independent color space.

2.3.1 The CIELAB space

The L^*, a^*, and b^* components of the CIELAB space are defined in terms of the X, Y, and Z components of the CIE XYZ space by the nonlinear transformation

$$\begin{aligned} L^* &= 116\, f\left(\frac{Y}{Y_n}\right) - 16, \\ a^* &= 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right], \\ b^* &= 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right], \end{aligned} \tag{2.6}$$

where X_n, Y_n, and Z_n are the D_{65} white point values in the XYZ color space and

$$f(x) = \begin{cases} 7.787x + \frac{16}{116}, & \text{if } 0 \le x \le 0.008856 \\ x^{\frac{1}{3}}, & \text{if } 0.008856 < x \le 1. \end{cases} \tag{2.7}$$

The D_{65} white point values are the XYZ tristimulus values obtained for the D_{65} standard illumination specified by CIE to correspond to daylight at the temperature of 6500K. The distance between two color stimuli in the CIELAB space is denoted by ΔE^*_{ab} and a ΔE^*_{ab} of 2.3 corresponds approximately to a JND. In the CIELAB space L^* indicates lightness, chroma is indicated by the radial distance $\sqrt{(a^*)^2 + (b^*)^2}$ in the a^*b^* plane, and hue is indicated by $\left(\arctan \frac{a^*}{b^*}\right)$.

2.4 The sCIELAB space

The CIELAB and CIELUV color spaces suffer from a critical limitation in terms of their application to determining color reproduction errors. Both color spaces are not truly

uniform and deviations from uniformity occur in different regions of these color spaces. They are thus not suitable for applications that require color manipulations in these non-uniform regions. Also, the transformation from device color spaces to the CIELAB space (which is most commonly used to measure color reproduction errors) is highly non-uniform and does not lend itself to easy manipulation in conjunction with commonly used computation models and techniques in image processing. In addition, the color difference measure ΔE^*_{ab} is suitable only for measuring differences in large regions of uniform color. The CIELAB space does not account for the spatial characteristics of human vision, and a point-wise comparison between two images using CIELAB measures is not relevant. The HVS has a low-pass nature and attenuates high-frequency components of incident spectra. The frequency response of the eye has been studied in much detail using experiments with luminance and chrominance spatial patterns [34] and it has been determined that the bandwidth of the luminance channel is significantly larger than the bandwidth of the chrominance channels. This difference is exploited in the design of color processing algorithms and in numerous other applications like the transmission of color signals. A complete model that incorporates HVS spatial characteristics has not yet been developed although numerous models that independently explain different aspects of the psychophysics of human color vision have been proposed [35], [36].

The sCIELAB model [37] is a spatial extension of the CIELAB model and provides a measure that increases the accuracy in determining color errors. It's use has been demonstrated by the authors [38] on the visibility of textures in printed halftone patterns. The sCIELAB space is derived by a spatial filtering operation that simulates blurring by the HVS on color images. The color image is first transformed into three opponent color planes

O_1, O_2, and O_3, corresponding to black-white, red-green, and yellow-blue components respectively. The three planes are defined as a linear transformation on the XYZ tristimulus values as

$$\begin{bmatrix} O_1 \\ O_2 \\ O_3 \end{bmatrix} = \begin{bmatrix} 0.279 & 0.72 & -0.107 \\ -0.449 & 0.29 & -0.077 \\ 0.086 & 0.59 & -0.501 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \tag{2.8}$$

Each opponent-color channel is filtered by two-dimensional separable kernels of the form

$$f(x, y) = k_i \sum_i w_i E_i, \tag{2.9}$$

where $E_i = \exp\big(-(x^2 + y^2)/\sigma_i^2\big)$, and w_i are weights on the three opponent planes. E_i for each opponent color plane is determined by considering the HVS response to the particular color plane. The resulting three channels are transformed back to CIELAB space via the CIEXYZ inverse transform. The ΔE^*_{ab} measure on this new image is referred to as ΔE_s . The sCIELAB difference measure reflects both spatial and color sensitivity, and equals ΔE^*_{ab} over uniform regions of the image where the low-pass nature of the HVS does not cause appreciable degradation.

2.5 Image formation

The spectrum incident at the optical sensor of a recording device is a result of multiple transformations on the spectrum emanating from a scene. The scene may contain only an

illumination source, or as is more common, the scene may be a combination of light sources and reflecting surfaces. Figure 2.5 shows the illumination spectra of common light sources.

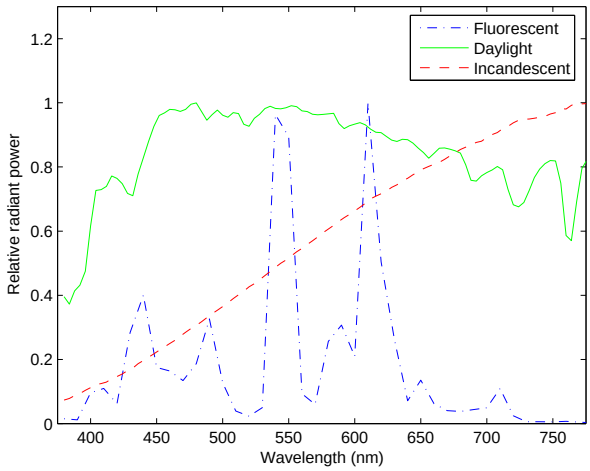


Figure 2.3: Relative spectral power distributions of common light sources

The generalized response obtained at a single location on a sensor-array can be described by

$$\begin{aligned} t_i &= \int_{-\infty}^{\infty} f_i(\lambda) d(\lambda) r(\lambda) l(\lambda) d\lambda + \eta_i \qquad i = 1, 2, .., k \\ &= \int_{-\infty}^{\infty} m_i(\lambda) r(\lambda) l(\lambda) d\lambda + \eta_i, \qquad i = 1, 2, .., k \end{aligned} \tag{2.10}$$

where k distinct color channels are sampled at each location, and

$$
\begin{aligned}
f_i(\lambda) &= \text{spectral transmittance of color filter,} \\
g(\lambda) &= \text{sensitivity of sensor,} \\
r(\lambda) &= \text{spectral reflectance of imaged surface,} \\
l(\lambda) &= \text{spectral power density of illuminant,} \\
\eta_i &= \text{measurement noise, and} \\
m(\lambda) &= f_i(\lambda)g(\lambda).
\end{aligned}
$$

The model may be expressed in the discrete form as

$$t = M^T L r + \eta, \tag{2.11}$$

where

t = $k \times 1$ vector of measurements,

r = $p \times 1$ vector of scene reflectance samples, p is the number of times the spectrum is sampled (usually, $p = 31$),

L = $p \times p$ diagonal matrix with samples of radiant spectrum of illuminant along its diagonal, and

M = $p \times k$ matrix that describes the combined filter-sensor response.

For a complete 2-dimensional image of dimension $M \times N$ and k color channels, the image formation model incorporates the blur due to the taking lens and thermal effects in

the optical sensor and is given by

$$y = Ax + n \tag{2.12}$$

Where x $(kMN \times 1)$ represents the channel intensities of the original scene. The vector x is of the form $x = [x_1^T, x_2^T, \cdots, x_k^T]^T$, and is a color-stacked form of the original column-ordered scene. The vector y is the similarly arranged data, and n represents system noise. The matrix A represents the blur and is of the form

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,k} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,k} \\ \vdots & \vdots & \vdots & \vdots \\ A_{k,1} & A_{k,2} & \cdots & Ak,k \end{bmatrix}, \tag{2.13}$$

where $A_{i,j}$ are $MN \times MN$ matrices that represent the blurring effect on the j^{th} color channel due to the i^{th} color channel.

We will use the models in Eqs. (4.2), and (2.12) in the next few chapters to design methods for CFA sample selection, demosaicking, and the selection of color filter spectral sensitivities.

Chapter 3

Demosaicking of Color Filter Array Data

3.1 Introduction

The multiple stages in the imaging pipeline of typical digital color imaging device are illustrated in Fig. 3.1. Although the sequence of operations may differ in various devices, the basic operations do remain the same. In the initial stage, the image is acquired via the taking optics after due exposure and focus control. Pre-processing attempts to correct for optical distortions and eliminate artifacts due to the electronics of the imager. The next few stages involve digital image processing and various manufacturers implement these operations in different configurations. Full-color images are formed from sparsely sampled CFA data in the demosaicking step. Post-processing involves the operations of sharpening, deblurring, color correction, etc., and in almost all implementations, these operations follow the demosaicking operation. The demosaicking operation that occurs earlier is inherently a nonlinear blurring operation, and operations like deblurring carried out at a late stage in the imaging pipeline can not take advantage of linear models of image formation. In this chapter, a generalized framework for the recovery of color images is proposed that addresses this issue. The proposed framework may be used to simultaneously demosaic, deblur, and denoise color images acquired as sparsely sampled CFA data.

The problem of reconstruction of full-color images from sparsely sampled CFA data is illustrated in Fig. 3.2. The CFA image may be separated into distinct sub-sampled color channels. At each pixel location, color channels that have not been sampled are

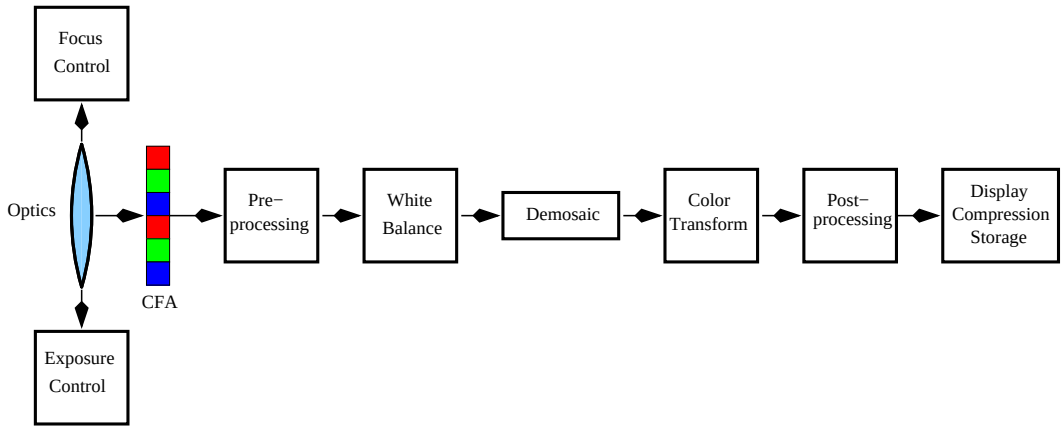


Figure 3.1: Image processing pipeline in a digital camera

reconstructed using information from adjoining pixels that are sampled. Both inter- and intra-channel intensity information may be used in the process.

If the color planes are considered as three separate images, the demosaicking problem may be thought of as analogous to the image interpolation problem. This naive approach to demosaicking using conventional interpolation strategies leads to artifacts (the *zipper* effect) unique to the demosaicking process. Figure 3.3 illustrates the process of CFA acquisition and demosaicking. The CFA acquires the intensity of incident energy at each spatial location. The resultant CFA image is separated into three color channels (red, green, and blue) and missing data is found by bilinear interpolation of color channels. The resulting full-color image shows severe color artifacts.

The demosaicking problem is effectively addressed by considering it as a problem of recovery of special multichannel signals. These special multichannel signals are color images that have a well-understood image formation model (Section **??**), and there is a clear set of desirable features required in the final reconstructed image, viz., its closeness in color appearance to the original scene. The properties of the human visual system and its response

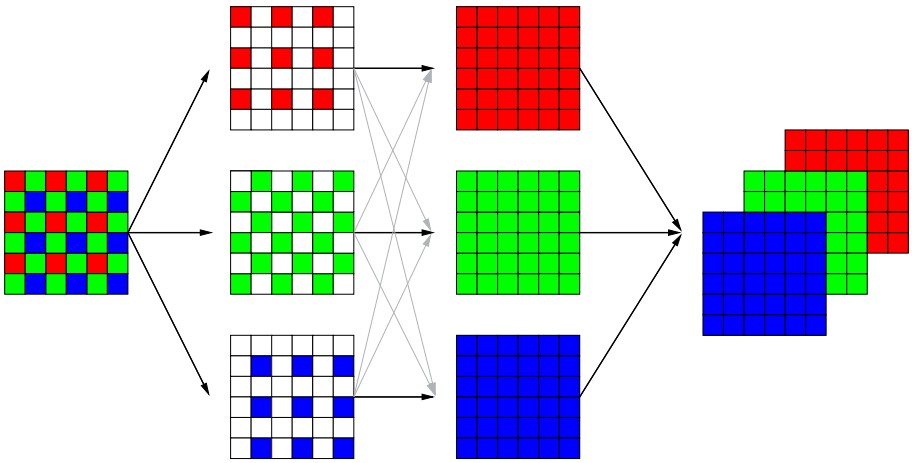
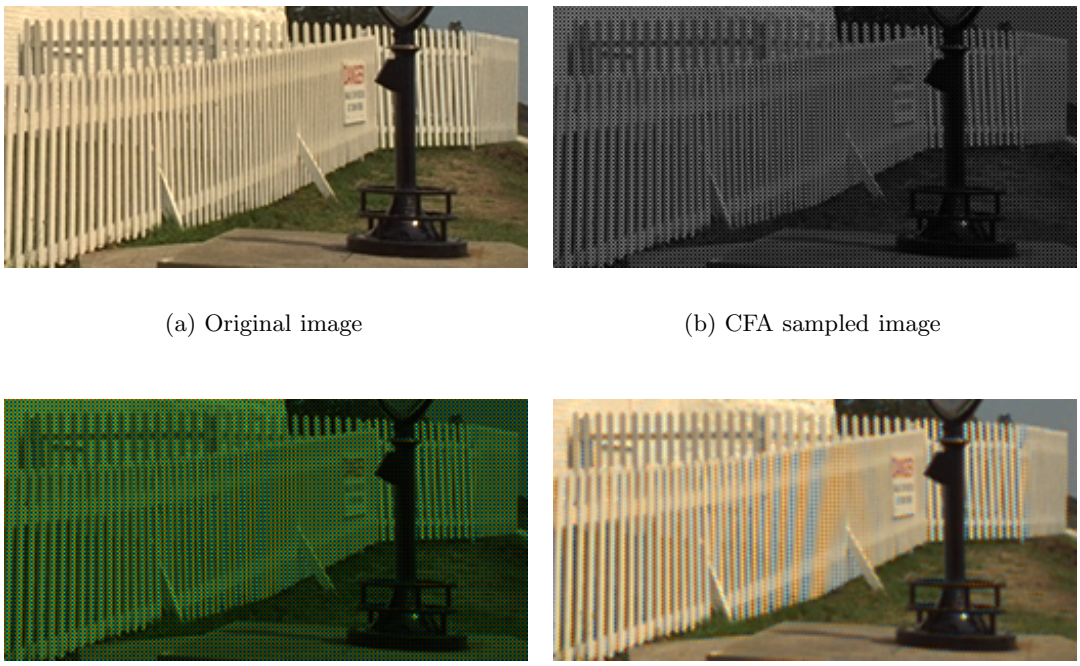


Figure 3.2: CFA sampling and demosaicking

to color signals should thus be an important component of demosaicking techniques. A number of specialized algorithms for demosaicking have been presented in the literature [39, 40, 41, 42, 43, 44, 3, 45, 46]. A majority of these algorithms can be classified into two broad groups – algorithms that are based on heuristics (these algorithms may account for cross-channel correlation and the presence of edges) and algorithms that are based on the solution of a mathematical problem. The most notable feature (as seen in the survey by Gunturk et al. [47]) of the most successful demosaicking algorithms is that they effectively use inter-channel correlation in addition to the spatial correlations in images.

A common thread in all these algorithms, especially conspicuous in heuristic algorithms, is that they are tailored for the Bayer array [14]. In light of the multitude of CFA patterns present both in the open literature and in proprietary frameworks used by manufacturers of DSCs [14, 18, 48, 49, 50, 51, 52, 17], it would be fruitful to design an algorithm that works for a general CFA pattern. The framework for color image restoration introduced in this chapter is not constrained by CFA arrangements and can be used for the joint deblurring and demosaicking of data obtained from arbitrary CFAs.



(a) Original image

(b) CFA sampled image

(c) CFA image with colors

(d) Bilinear reconstructed

Figure 3.3: (a) An image with information about three colors (red, green, and blue) at each spatial location. (b) Representation of the image as it would be acquired with a CFA-based imager. (c) CFA data shown with sampled colors at each location. (d) Result of bilinear reconstruction of CFA data.

3.2 Bayesian restoration

The problem of restoration of images from sparse and noisy data is generally ill-posed, and easy inversion is rarely possible. Effective solutions require accurate knowledge of the degradation model and the characteristics of the imaging system. In deterministic techniques, prior information about the *true* image is typically introduced to ensure a solution to the inversion problem. For instance, the assumption of local smoothness in images holds in the general case and is typically used for regularization of the least squares

inversion problem. In non-deterministic techniques, prior information is in the form of a distribution or a probability density function for the image that is to be estimated.

Bayesian methods have been used extensively in imaging applications like restoration, segmentation, computed tomography, etc. The Bayesian approach is well suited to the problem of image restoration since it allows for flexible and effective means to incorporate prior information into the solution. Within the Bayesian paradigm, a common choice for the estimator is the maximum a posteriori (MAP) solution, which provides an estimate for the true image x from the noisy data y as the maximum of the posterior probability density function $\mathrm{p}(x|y)$, which is in effect the most likely image given the occurrence of the data y. The choice of $\mathrm{p}(x)$, the prior density function, greatly affects the quality of the solution.

Markov random field (MRF) image models effectively describe the smoothness and local nature of features in the general natural scene and have been extensively used to define image priors. MRFs are derived from potential functions that may be thought of as representing a quantity akin to energy. The desirability of configurations of local intensities depends on the value of the resulting potential function. Common potential functions of note are functions of the type $\psi(x_i - x_j)$, where x_i and x_j are intensity values at the i^{th} and j^{th} pixel respectively. A common prior model is the Gauss-Markov random field (GMRF) where ψ is a quadratic. A major issue with the GMRF model is its behavior across edges. Cost functions based on GMRF models will penalize large intensity differences between pixels and characteristically oversmooth across edges. Typically, local features and edges are accommodated by augmenting GMRF models with line processes [53] that lend a degree of adaptation to the estimation procedure. The line process l acts to inhibit smoothing across edges and encourages smoothing across pixels that do not lie across edges. The

MAP estimate in this case is found as

$$\hat{x}, \hat{l} = \arg\max_{x,l} \mathrm{p}(x, l|y). \tag{3.1}$$

Bayesian restoration of multichannel images has been addressed in the literature. Molina et al. [54] use multiple line processes to define a potential function that has cross-channel line-process terms. The cost function derived from the resulting GMRF prior reduces the penalty on intensity differences in a particular channel if line processes in other channels indicate the presence of an edge. This approach is useful since edge features in scenes typically appear at boundaries of objects and are typically reflected across all color channels. The Bayesian approach has also been used to address the specific problem of color reconstruction from single-sensor data. Parmar et al. [55] use a GMRF prior that includes only one line process that reflects the probability of the presence of an edge in all three color channels of a color image.

Here, we propose a novel prior model for color images that incorporates cross-channel edge information. The proposed model is the result of a GMRF model augmented with line processes that attenuate the penalty on large intensity differences to prevent smoothing across edges. In addition, the prior has cross channel terms that describe spatial smoothing in the color-difference channels. Section 3.3 details the reasoning behind this approach. The image model and restoration algorithm are developed in Section 3.4. In Section 3.5 the efficacy of the proposed algorithm is demonstrated on the problem of reconstruction of color images acquired by single-sensor digital cameras.

3.3 Color image model

Digital color images are obtained either by direct acquisition with digital cameras or by scanning prints or slides obtained by film-based cameras. In both cases, the digital imaging device typically acquires three color bands in the red, green, and blue (RGB) regions of the spectrum. The forward model in the discrete form for the signal acquired at a pixel location is given by

$$c_i = \mathcal{F}\left(\sum_{k=1}^{N} r_k s_{i_k} + \eta\right), \qquad i = \mathrm{R}, \mathrm{G}, \mathrm{B} \tag{3.2}$$

where c_i is the intensity of the i^{th} color, r_k and s_{i_k} are samples of scene irradiance and sensitivity of the i^{th} sensor at wavelength k nm respectively, η is noise, and $\mathcal{F}(.)$ is a non-linear function that describes the characteristics of the imaging system.

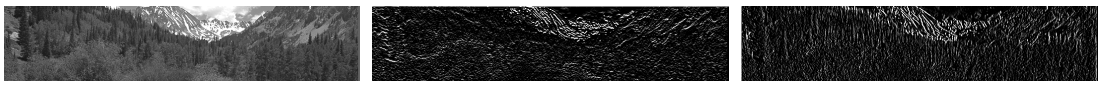
Channel intensities in (3.2) are functions of the inner products of the scene spectral content described by the irradiance and the respective sensor sensitivities. The irradiance incident on a sensor due to the general natural scene is a consequence of the reflectance of the scene and the radiance of the illuminant. Typically, the illuminant (sunlight, camera flash, fluorescent light, etc.) has broad support across the spectrum. Also, sensor sensitivities have considerable overlapping support, especially among adjoining bands (red and green, green and blue). As a result, the channel intensities detected at a spatial location are well correlated.

Color images can be decomposed into the luminance channel, which describes brightness, and two chrominance channels, which convey information about color [56]. It is well

Figure 3.4: Sample images from Eastman Kodak's PhotoCD PCD0992.

known that most significant structures in color images manifest predominantly in the luminance channel. Color image enhancement algorithms take advantage of this feature; for instance, unsharp masking as applied to color images is performed only on the luminance channel. Luminance is commonly defined as a linear combination of the color channels, although the weights for each channel differ among various treatments. The appearance of major features in the luminance channel suggests that feature edges in typical color images

are very well correlated. This phenomenon was demonstrated by Gunturk et al. [3] by decomposing each color channel of a set of images into 4 bands by filtering with directional (0, $\pi/4$, $\pi/2$, and $3\pi/4$) high-pass filters. It was shown that corresponding high frequency components across the color channels are highly correlated. An equivalent assumption is that the color-difference channels (R-G and G-B) are band limited. The smoothness of the R-G and R-B color-difference channels in a natural scene is illustrated in the images in the left column in Figs. 3.5(b) and 3.5(c) respectively.



(a) From left to right: The red channel shown in grayscale, ∇_HR; and ∇_VR

(b) From left to right: The R-G channel difference image; ∇_H(R-G); and ∇_V(R-G)

(c) From left to right: The R-B channel difference image; ∇_H(R-B); and ∇_V(R-B)

Figure 3.5: A representation of horizontal and vertical gradients obtained as the first differences in the respective directions.

It follows from the high correlation among high-frequency components of the color channels that the gradients of the color channels will be highly correlated. This is illustrated in Fig. 3.5. Figure 3.5(a) shows the red channel of an image from Kodak's PhotoCD PCD0992 [57] (Fig. 3.4) and the horizontal and vertical gradients found as the respective first differences of the red channel image (denoted hereafter with the symbols ∇_H and ∇_V

Table 3.1: Correlation coefficients found for the database of images in Eastman Kodak's PhotoCD PCD0992.

ρ	i = R j = G	i = G j = B	i = B j = R
i, j	0.8534	0.9230	0.7560
$\nabla_H i, \nabla_H j$	0.9777	0.9765	0.9561
$\nabla_V i, \nabla_V j$	0.9751	0.9776	0.9565
$\nabla_{DL} i, \nabla_{DL} j$	0.9760	0.9778	0.9546
$\nabla_{DR} i, \nabla_{DR} j$	0.9751	0.9774	0.9532

respectively). Figs. 3.5(b) and 3.5(c) show the color-difference (∇_HR - ∇_HG, and ∇_VR - ∇_VG) images and the corresponding images for the R-B channel respectively. The color-difference images are characteristically smooth and the difference images of channel gradients illustrate their small magnitude. Table 3.1 shows the correlation coefficients among color channels and color channel gradients for an image obtained by stitching together all images in Eastman Kodak's PhotoCD PCD0992 [57]. Correlation coefficients for channel gradients in each direction are very high in each case, and highest for adjacent bands ($\nabla_\theta R \nabla_\theta G$ and $\nabla_\theta B \nabla_\theta R$).

3.3.1 Degradation Model

The degradation model is presented in the matrix-vector form as:

$$y = Ax + w, \tag{3.3}$$

where $y \in \mathbb{R}^{3MN \times 1}$ is formed by stacking the three column-ordered color channels. The degradation to the true image $x \in \mathbb{R}^{3MN \times 1}$ is described by A, which is a block matrix that

has expressions for intra-channel blur as its diagonal blocks and cross-channel blur as the off-diagonal blocks. The additive noise w is zero-mean white Gaussian noise with variance σ_w^2. The color channels are ordered column-wise and stacked such that $x = [x^{\mathrm{R}^T}, x^{\mathrm{G}^T}, x^{\mathrm{B}^T}]^T$ (y is similarly arranged).

3.3.2 Prior model

From the discussion in Section 3.3, it follows that a suitable prior model for color images will be formed from potential functions of the type

$$\psi\left(\left(x_i^k - x_j^k\right) - \left(x_i^{k'} - x_j^{k'}\right)\right), \tag{3.4}$$

where i, j are pixel locations, k, k' are R,G,B, and $k \neq k'$. The cost functions derived from such priors will penalize the difference between the first-difference of pixel intensities between color channels.

For an $M \times N$ color image, the true image x is defined as a realization of a random process defined on a 3-D rectangular lattice S with $3MN$ points. In addition we introduce two sets of line processes l_θ and c_θ, $\theta = H, V, DL, DR$, for the horizontal, vertical, and left and right diagonal directions respectively. The line processes l_θ model intra-channel intensity transitions and c_θ model the transitions in the color difference channels. Fig. 3.6 shows all line processes associated with the RGB intensities at a spatial location in the image.

The prior joint density for x, l_θ, and c_θ is defined as a Gibbs density to ensure that the resulting field is a MRF (Hammersley-Clifford theorem) [58]. A Gibbs field has a density

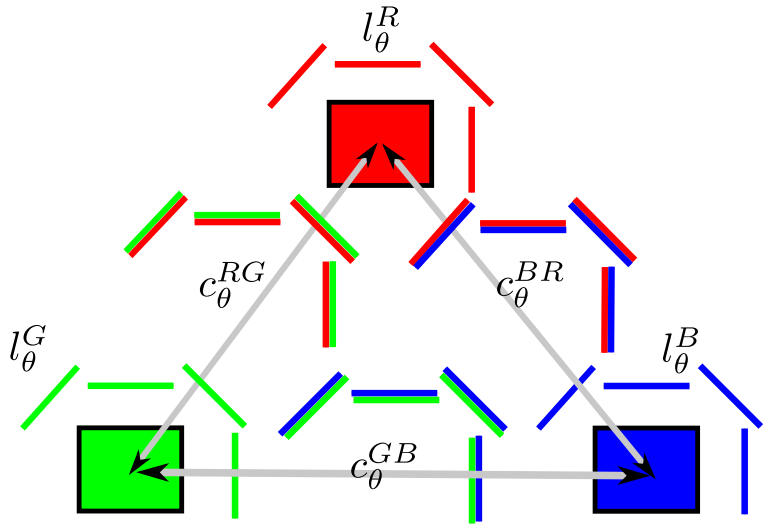


Figure 3.6: Representation of a point in the 3-D lattice with associated line processes. Red, green and blue pixels are shown surrounded by the respective line processes that denote intra-channel edges (l_θ^k). Line processes for the cross-channel terms ($c_\theta^{kk'}$) are appropriately labeled.

function of the form

$$p(x) = \frac{\exp\left(-\sum_{i\in\mathcal{C}} \tilde{v}_i(x)\right)}{Z_x}, \tag{3.5}$$

where Z_x in the denominator normalizes the density function, $\tilde{v}_i(x)$ is the potential function defined over the set of cliques $\mathcal{C}$. A subset $\mathcal{C}$ of S is a clique if every pair of distinct sites in $\mathcal{C}$ are neighbors. Figure 3.7 shows the set of cliques used to define the proposed prior model

$$p(x, l_\theta, c_\theta) = \frac{1}{Z_x} \exp\Big(-\frac{1}{2\sigma_x^2} \sum_i \sum_{k,k'} \sum_\theta \zeta(x_i^k)^2 \quad + \big((x_i^k - x_{i:+\theta}^k) - (x_i^{k'} - x_{i:+\theta}^{k'})\big)^2 \tilde{c}_{\theta_i}^{kk'} + (x_i^k - x_{i:+\theta}^k)^2 \tilde{l}_{\theta_i}^k \Big) \tag{3.6}$$

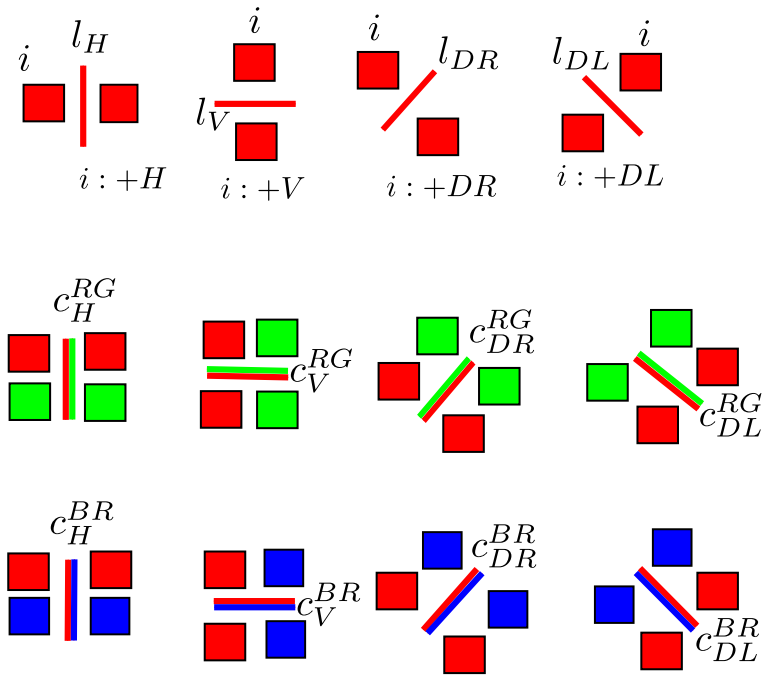


Figure 3.7: The set of cliques associated with a red pixel at location i. Locations of $i : +\theta$, $\theta = H, V, DL, DR$ are labeled.

where $1 \leq i \leq MN$; $\theta = H, V, DL, DR$; $\tilde{l}_{\theta_i} = 1 - l_{\theta_i}$, $\tilde{c}_{\theta_i} = 1 - c_{\theta_i}$; $k, k' = R, G, B$, $k \neq k'$ (x^k are random processes with intensities of the three color channels); and the term ζx_{ij}^2 keeps the density in (3.6) from being improper. ζ is set to a number that is small enough for this term to have little effect on the solution. We used $\zeta = 10^{-3}$ for all reconstructions in this paper. The index $i : +\theta$ refers to the pixel location adjacent to i in the direction of θ (illustrated in Fig. 3.7), and $i : -\theta$ will refer to the spatial location $i : +\theta + \pi$.

3.4 Algorithm Derivation

The maximum a posteriori (MAP) estimates of x and the line processes given the prior described in (3.6) are obtained by maximizing

$$p(x, l_\theta, c_\theta|y) = \frac{p(y \mid x)\, p(x, l_\theta, c_\theta)}{p(y)}. \tag{3.7}$$

As $p(y)$ is constant with respect to x, l_θ, and c_θ, the optimal values of x, l_θ, and c_θ are the solution to the following optimization problem

$$\hat{x}, \hat{l}_\theta, \hat{c}_\theta = \arg\max_{x, l_\theta, c_\theta} p(y \mid x)\, p(x, l_\theta, c_\theta), \tag{3.8}$$

where $p(y \mid x)$ is the likelihood function in the presence of additive, uncorrelated Gaussian noise and is given by

$$p(y \mid x) = \frac{1}{(2\pi)^{3MN} \sigma_w^{6MN}} e^{\left(-\frac{1}{2\sigma_w^2}[y-Ax]^T[y-Ax]\right)}. \tag{3.9}$$

Optimizing the cost function in (4.18) simultaneously over x, l_θ and c_θ is non-convex and computationally prohibitive. Instead, we iteratively update the estimate of x and then update the estimate of the edge variables. Results derived in [59] based on the iterated conditional modes (ICM) algorithm [60] and the iterated conditional average (ICA) technique are used to update x, l_θ and c_θ.

3.4.1 The ICM iterations for pixel update

The ICM is a greedy iterative algorithm that sequentially updates values of pixels by maximizing their conditional posterior probability. Specifically, it finds the value of x_i^k that maximizes the conditional probability of x_i^k given all the remaining pixels x_j^k, $x_j^{k'}$, $x_i^{k'}$, $i \neq j$, and $k \neq k'$, and the associated edge variables. The required conditional probability is given by

$$\begin{aligned} p(x_i^k \mid l_\theta^k, c_\theta^{kk'}, x_{\setminus ik}) &= p(y \mid x_i^k, x_{\setminus ik})\, p(x_i^k \mid x_{\setminus ik}, l_\theta^k, c_\theta^{kk'}) \\ &= p(y \mid x)\, p(x_i^k \mid x_{\text{nb}}, l_\theta^k, c_\theta^{kk'})\,, \end{aligned} \tag{3.10}$$

where $x_{\setminus ik}$ is the set of all pixels in the image except the i^{th} pixel of color k, and x_{nb} are neighboring pixels of x_i^k.

It can be shown [59] that the conditional posterior density of a pixel contingent on its neighboring elements x_{nb} is

$$p\left(x_i^k \mid y, x_{\text{nb}}, l_{\theta_{\text{nb}}}, c_{\theta_{\text{nb}}}\right) = \frac{1}{\sqrt{2\pi}\,\sigma_i} e^{\left(-(x_i^k-\mu_i)^2/2\,\sigma_i^2\right)}, \tag{3.11}$$

where

$$\mu_i = \frac{a_i^T\,(y - Ax_{-i}) + (\sigma_w^2/\sigma_x^2)\rho_i^k}{a_i^T a_i + (\sigma_w^2/\sigma_x^2)\gamma_i^k}\,, \tag{3.12}$$

and a_i is the column of A corresponding to pixel x_i, x_{-i} is x with a zero in the i^{th} pixel, and

$$\rho_i^k = \sum_{\theta,k,k'} \left(x_{i:+\theta}^k \tilde{l}_{\theta_{i:+}}^k + (x_{i:+\theta}^k + x_i^{k'} - x_{i:+\theta}^{k'})\, \tilde{c}_{\theta_{i:+}}^{kk'} + x_{i:-\theta}^k \tilde{l}_{\theta_i}^k + (x_{i:-\theta}^k + x_i^{k'} - x_{i:-\theta}^{k'})\, \tilde{c}_{\theta_i}^{kk'} \right)$$

$$\gamma_i^k = \zeta + \sum_{\theta,k,k'} \left(\tilde{l}_{\theta_{i:+}}^k + \tilde{c}_{\theta_{i+}}^{kk'} + \tilde{l}_{\theta_i}^k + \tilde{c}_{\theta_i}^{kk'} \right) \tag{3.13}$$

The conditional mean is also the conditional mode and maximizes the probability in Equation (3.11). The ICM algorithm consists of iteratively replacing pixel x_i^k with its conditional mean μ_i^k, i.e.

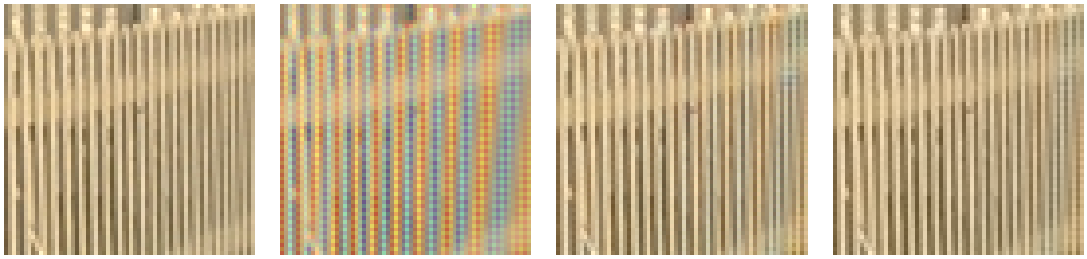
$$x_i^{k^{p+1}} = \frac{a_i^T \left(y - A x_{-i}^p \right) + (\sigma_w^2/\sigma_x^2)\rho_i^p}{c + (\sigma_w^2/\sigma_x^2)\gamma_i^p}\,, \tag{3.14}$$

where $[.]^p$ denotes the value of a variable after the p^{th} iteration. The ICA algorithm [59] is used to update the line processes by iteratively updating a single line variable with its mean conditioned on x and all the other edge variables. The restoration algorithm iterates alternately between the pixel updates and line variable updates until convergence. Updated values are used in subsequent iterations. Convergence of this iteration for fixed l_θ and c_θ is assured [60] because

$$p(x \mid y) = p(x_i^k \mid y, x_{\setminus ik})\, p(x_{\setminus ik} \mid y)$$

does not decrease at each iteration.

(a) Original image (b) Bilinear (c) POCS (d) Proposed



(e) Cropped image (f) Bilinear (g) POCS (h) Proposed

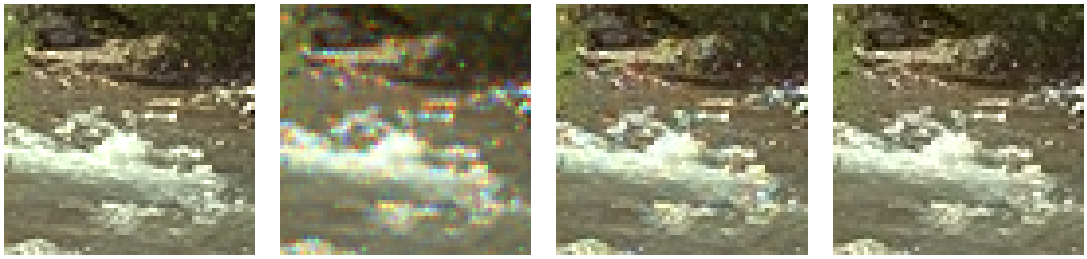
Figure 3.8: Reconstruction results for image 19 in Kodak PhotoCD PCD0992

3.4.2 Edge Variable Update

The ICA algorithm [59] is used to update the inter and intra-channel line processes. The ICA algorithm iteratively updates a single edge variable with its mean conditioned on x and all the other edge variables. It can be shown that the density of an intra-channel edge variable conditioned on x and all the other edge variables has the truncated exponential density

$$p(l^k_{\theta_i} \mid x^k_i, l_{\theta_{\setminus ik}}, c_\theta) = C_1 \exp\{\alpha_{l^k_{\theta_i}} l^k_{\theta_i}\} \qquad 0 \le l^k_{\theta_i} \le 1\,, \tag{3.15}$$

(a) Original image (b) Bilinear (c) POCS (d) Proposed



(e) Cropped image (f) Bilinear (g) POCS (h) Proposed

Figure 3.9: Reconstruction results for image 13 in Kodak PhotoCD PCD0992

where C_1 is a constant, $l_{\theta_{\setminus ik}}$ represents all the edge variables except $l^k_{\theta_i}$ variable, and

$$\alpha_{l^k_{\theta_i}} = \frac{1}{2\sigma_x^2} \sum_k (x^k_i - x^k_{i:-\theta})^2 \,. \tag{3.16}$$

Similarly, the density of an inter-channel edge variable conditioned on x and all the other edge variables has the truncated exponential density

$$p(c^{kk'}_{\theta_i} \mid x^k_i, c_{\theta_{\setminus ikk'}}, l_\theta) = C_1 \exp\{\alpha_{c^{kk'}_{\theta_i}} c^{kk'}_{\theta_i}\} \qquad 0 \le c^{kk'}_{\theta_i} \le 1 \,, \tag{3.17}$$

(a) Original image (b) Bilinear (c) POCS (d) Proposed

(e) Cropped image (f) Bilinear (g) POCS (h) Proposed

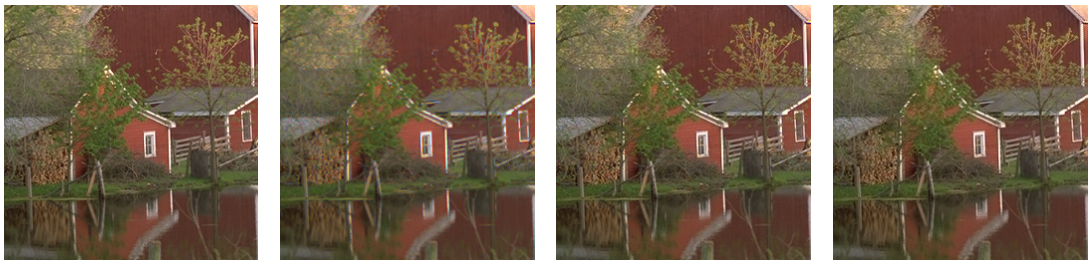
Figure 3.10: Reconstruction results for image 11 in Kodak PhotoCD PCD0992

where C_1 is a constant, $l_{\theta_{\setminus ik}}$ represents all the edge variables except $l^k_{\theta_i}$ variable, and

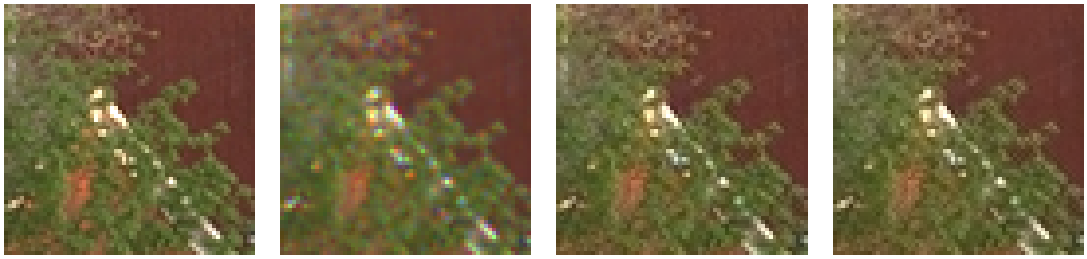
$$\alpha_{c^{kk'}_{\theta_i}} = \frac{1}{2\sigma_x^2} \sum_{kk'} \left((x^k_i - x^k_{i:+\theta}) - (x^{k'}_i - x^{k'}_{i:+\theta}) \right)^2 . \tag{3.18}$$

The conditional distributions in (3.15) and (3.17) have a mode of one, so an ICM update of l^k_θ and $c^{kk'}_\theta$ would result in turning all the edges variables on. For this reason, we use the ICA algorithm, which uses the iteration

$$l^{k^{p+1}}_{\theta_i} = \bar{l}^{k^p}_{\theta_i} , \tag{3.19}$$



(a) Original image (b) Bilinear (c) POCS (d) Proposed



(e) Cropped image (f) Bilinear (g) POCS (h) Proposed

Figure 3.11: Reconstruction results for image 22 in Kodak PhotoCD PCD0992

where $\bar{l}_{\theta_i}^{k^p}$ is the mean of the truncated exponential random variable in (3.15) and is given by

$$\bar{l}_{\theta_i}^{k} = \frac{\alpha_{l_{\theta_i}^k} \exp\left(\alpha_{l_{\theta_i}^k}\right) - \exp\left(\alpha_{l_{\theta_i}^k}\right) + 1}{\alpha_{l_{\theta_i}^k}\left[\exp\left(\alpha_{l_{\theta_i}^k}\right) - 1\right]}. \tag{3.20}$$

If the $\alpha_{l_{\theta_i}^k}$ term in (3.20) gets large, $\exp\left(\alpha_{l_{\theta_i}^k}\right)$ can exceed the floating-point limits on a typical computer, and this can cause numerical instability problems. To remedy this situation, we multiply the top and bottom of (3.20) by $\exp\left(-\alpha_{l_{\theta_i}^k}\right)$, which yields the more stable expression:

$$\bar{l}_{\theta_i}^{k} = \frac{\alpha_{l_{\theta_i}^k} - 1 + \exp\left(-\alpha_{l_{\theta_i}^k}\right)}{\alpha_{l_{\theta_i}^k}\left[1 - \exp(-\alpha_{l_{\theta_i}^k})\right]}. \tag{3.21}$$

Similar expressions hold for $c_{\theta_i}^{kk'}$.

(a) Original image (b) Bilinear (c) POCS (d) Proposed

(e) Cropped image (f) Bilinear (g) POCS (h) Proposed

Figure 3.12: Reconstruction results for image 21 in Kodak PhotoCD PCD0992

3.4.3 Demosaicking Algorithm

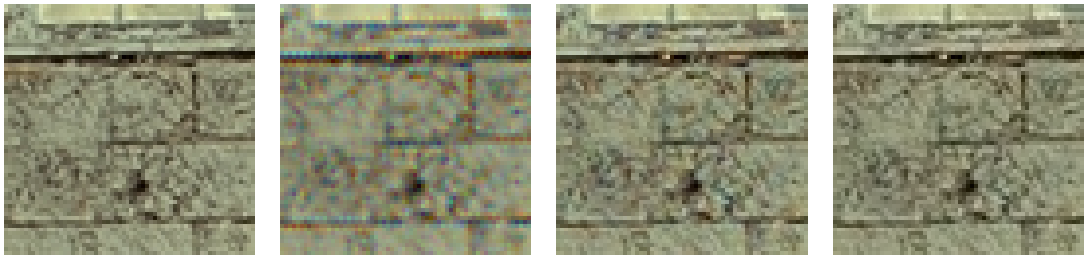
The demosaicking algorithm consists of two steps that are performed alternately until convergence. First, pixel values (x) are updated using Equation (3.14). In the second step, the edge variables $l^k_{\theta_i}$, $c^{kk'}_{\theta_i}$ are updated with their conditional means using Equation (3.19). The reconstruction algorithm is summarized as follows:

1. Initialize x, l_θ, and c_θ using the result of bilinear interpolation on CFA data.

2. Starting from the red pixel at $i = 1$, update the value of color intensity using (3.14) using previously updated values for subsequent pixels.

3. Given x^{p+1}, l^p_θ, and c^p_θ, obtain l^{p+1}_θ, and c^{p+1}_θ using the ICA iteration in (3.19).

(a) Original image (b) Bilinear (c) POCS (d) Proposed



(e) Cropped image (f) Bilinear (g) POCS (h) Proposed

Figure 3.13: Reconstruction results for image 1 in Kodak PhotoCD PCD0992

4. Repeat until convergence.

3.5 Experiments

We demonstrate the proposed algorithm on the problem of reconstruction of full-color images acquired with single-sensor digital cameras. The system model in this is case is

$$y = SHx + w \tag{3.22}$$

where S is a sampling matrix that and H represents blur. The proposed algorithm was used to reconstruct test images from the Kodak PhotoCD PCD0992 collection (Fig. 3.4) that have been sampled by the Bayer CFA. The blur matrix H for our experiments was set

at the identity to facilitate comparison with demosaicking methods that do not include a deblurring operation. The values for σ_x, and σ_w were set at 10^{-2} and 10^{-1} respectively. Experimental results are shown in Figs. 3.8-3.13. Table 4.2 gives the RMSE values for the color channels of the reconstructed images.

Table 3.2: Channel RMS errors for images listed in the first column

	Bilinear			POCS [3]			Proposed		
Image	Red	Green	Blue	Red	Green	Blue	Red	Green	Blue
Fig. 3.8	17.5996	9.3686	16.8524	4.6703	2.5134	4.3881	4.1491	2.5991	3.9239
Fig. 3.9	17.8050	12.2653	18.4221	6.3757	3.9813	6.8792	4.4562	3.7071	4.7347
Fig. 3.10	14.7799	9.3768	14.7081	5.4179	2.7658	4.5654	5.1074	2.7888	3.9457
Fig. 3.11	11.1448	7.1655	11.6881	4.3177	2.7996	4.4672	4.1826	3.0986	3.4819
Fig. 3.12	15.1467	9.1876	14.7369	5.2312	2.7563	4.8997	4.8315	2.8907	3.8166
Fig. 3.13	14.7886	9.3236	14.7624	4.6903	2.5793	4.4692	3.6143	2.2998	2.9378

The results obtained with the proposed algorithm are compared with results obtained with the algorithm in [3] based on projections on to convex sets. The proposed algorithm provides reconstructed results that have lower RMSE when compared with the original image for which intensities of all three colors are available at each location. Subjectively, it is seen in Figs. 3.8-3.13 that the proposed algorithm leads to reconstructed images that display fewer color artifacts. The improved results can be seen most clearly in textured regions in images where the proposed prior that accounts for cross-channel correlations allows for better reconstruction. For instance, in the cropped images in Fig. 3.9, false color is clearly seen in the POCS reconstructed image in the white water, while the results of the proposed algorithm have very few false color artifacts.

3.6 Conclusions

In this chapter we have proposed a novel cross-channel GMRF prior model for color images. The proposed prior model takes advantage of the high correlation between high-frequency edge content across the color channels of an image. We also propose a Bayesian algorithm that uses the new prior model to obtain MAP estimates for the restoration of color images. The proposed algorithm can be used to jointly demosaic and deblur images acquired with single-sensor digital cameras. The efficacy of the algorithm is demonstrated in experiments in which we demosaic sparsely sampled color images to arrive at a good estimate of the full-color image.

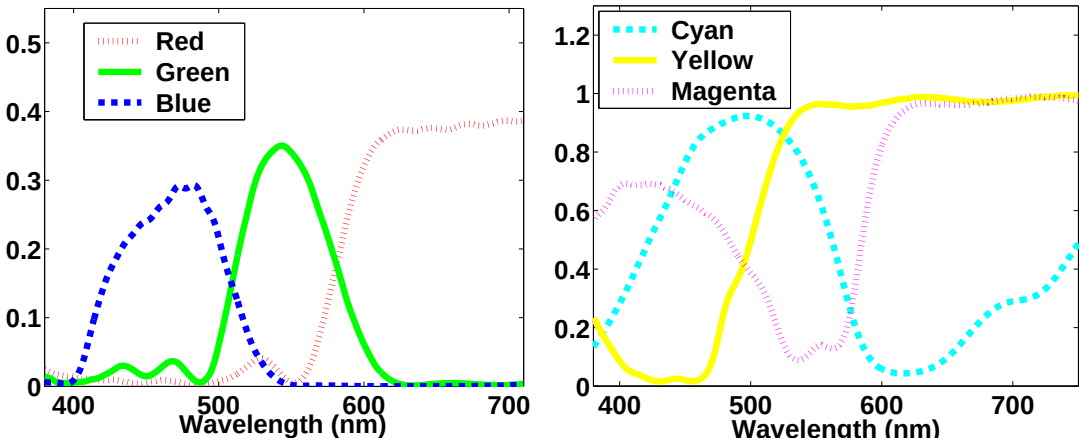
Chapter 4

Selection of sensor spectral sensitivities

4.1 Background

Figures 4.1(a) and 4.1(b) show the responses of typical sensor-arrays that use RGB and CMY color filters respectively. The responses shown are for sensor-arrays that are cascaded with IR blocking filters and the camera lens. CMY sensors provide greater dynamic range and SNR as compared to RGB sensors since the RGB primary pigments are created using multiple layers of the subtractive primaries C, M, and Y. Each camera will have a different spectral response due to tolerances in sensors and color filters. The response is compensated for in the digital image processing steps to produce a standard color response. Typically, the response of cameras is measured and this information is used in later color correction steps [10].

The colorimetric accuracy of any acquisition device in terms of its ability to exactly reproduce the colors in a scene, as well as the ability of the system to reproduce luminance information (exceedingly important even in color images [61]), depends fundamentally on the spectral sensitivity functions of the color filters. The design of spectral sensitivity functions is thus an important design parameter. Much work has been directed at the problem of selection of spectral sensitivity functions for improving the reproduction of color when multiple colors are acquired at each spatial location [62, 63, 13, 64, 65, 11, 12, 66, 67]. The main issues in this problem are the accurate reproduction of color under varying illuminants and in the presence of noise [13, 65]. Wolski et al. [66] propose a method for the selection of optimal sensor sensitivities based on their ability to reproduce color under



(a) RGB

(b) CMY

Figure 4.1: Spectral sensitivity functions. Ordinates represent transmittance, abscissae are wavelength in nm. (a),(b) RGB and CMY transmittances respectively from ImagEval's vCamera toolbox.

varying illuminants using a regularization approach. Vora and Trussell [11], [12] perform a mathematical analysis of sensor responses in terms of their color reproduction abilities and propose a method for the design of spectral sensitivity functions.

The selection of spectral sensitivity selection is all the more critical for CFA-based acquisition systems, but the problem has received very little attention in the research community. In addition to the obvious effect on color reproduction, the spatial-chromatic sampling nature of CFA-based schemes enforces a dependence of spatial reconstruction quality on the spectral sensitivity functions [68]. Sensitivity functions that project incident spectra to a tristimulus space that is highly correlated will yield improved spatial reconstruction results, since information about unsampled colors at specific locations can be well deduced from the values of the sampled colors due to high inter-channel correlation. However, a high inter-channel correlation has a detrimental effect on the discriminability of colors.

Alleysson et al. [69] demonstrate this trade-off between color discrimination and spatial reconstruction quality. Alleysson's analysis follows from a unique model of the CFA image considered as a single-band signal. This CFA image has a localized luminance-chrominance response in the frequency domain. They arrive at optimal values of sensitivity bandwidths and maximum values for Gaussian sensitivity functions based on PSNR values obtained after a frequency-domain reconstruction of Bayer-sampled images.

A common theme in most research on the reconstruction of full-color images from CFA data is the choice of a forward model that simulates CFA acquisition by subsampling a three-color image. This is a major deficiency, as this approach bypasses the very significant effect that color filter spectral sensitivity functions have on the appearance of reproduced color. It is critical to accurate judgment of reconstruction quality vis-à-vis the acquired scene that the forward model incorporate the transformation to the final tristimulus space of image representation.

We are not aware of any work in the literature that addresses the problem of sampling the color spectrum and sampling in the spatial domain in a unified framework. This dual-sampling is a characteristic of mosaic-based sampling schemes, and in this chapter we propose a method for the design of optimal spectral sensitivities that addresses this deficiency. We demonstrate the great impact that the spectral transmittance functions have on the quality of reconstruction and propose a method for the design of optimal spectral sensitivities for the color filters used in CFAs. The sensitivities are optimal in the sense that they minimize the color reproduction error between the tristimulus values obtained for a scene and the tristimulus values obtained for the reconstructed image that represents the same scene in the linearized CIELAB perceptually uniform color space. The optimality

criterion is reduced to a form that is expressed in terms of inter- and intra-pixel wavelength correlation matrices and does not depend on particular images.

The chapter is organized as follows: In Section 4.2 we present a forward model for image formation in the case of CFA acquisition. In Section 4.3 we develop the optimality criterion based on a multi-channel Wiener reconstruction from CFA data, and in Section 4.4 we present a model for spatial-spectral correlation. Sections 4.5 and 4.6 present experimental results and conclusions.

4.2 Image formation Model

Figure 4.2 depicts the image formation process for an image acquired at a spatial location in a sensor array. Radiation from a light source is incident on a point in the scene with reflectance $x(\lambda)$. The reflected beam then travels through a color filter to an optical detector (CCD or CMOS device) that has a sensitivity $d(\lambda)$. The signal obtained at the detector is given by

$$c = \int_{\lambda_{min}}^{\lambda_{max}} f(\lambda)d(\lambda)x(\lambda)l(\lambda)d\lambda + \eta, \tag{4.1}$$

where $l(\lambda)$ is the spectral power density of the illuminant, $f(\lambda)$ is the spectral transmittance of the color filter, and η is the measurement noise. The detector is sensitive in the wavelength range $(\lambda_{min}, \lambda_{max})$.

As is customary in the field, the visible spectrum is assumed to be adequately sampled when sampled every 10 nm in the range 400–700 nm to give a total of 31 samples. The

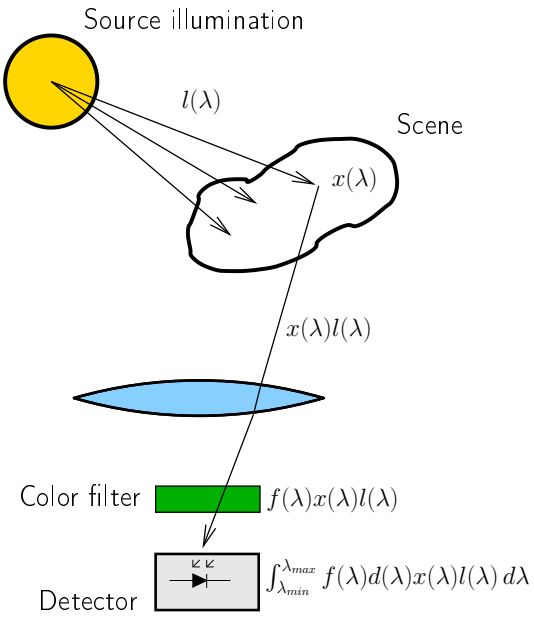


Figure 4.2: Representation of the image formation process in color image acquisition with color filters

image formation model may be expressed in the discrete form as

$$c = M^T Lx + \eta, \tag{4.2}$$

where $x \in \mathbb{R}^{31\times 1}$ contains scene reflectance samples, $L \in \mathbb{R}^{31\times 31}$ is a diagonal matrix with samples of the radiant spectrum of the illuminant along its diagonal, and $M \in \mathbb{R}^{31\times 1}$ describes the combined filter-sensor response.

We extend (4.2) to describe the image formation model for an $m \times n$ sensor array that samples p color channels at each pixel location. Let $f_i \in \mathbb{R}^{31\times 1}$, $i = 1, 2, \cdots, p$, describe the sampled spectral sensitivity functions of the p colors and let $F = [f_1^T, f_2^T, \cdots, f_p^T]^T$. The

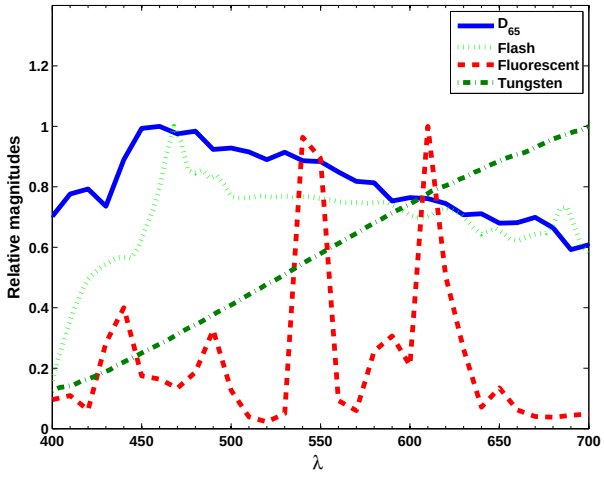


Figure 4.3: Sampled spectra of common illuminants in the range 400-700 nm

image acquired at the sensor is described by the vector

$$y = \bar{F}\bar{L}x + \eta. \tag{4.3}$$

The scene is described by the vector of reflectance values $x = [x_1^T, x_2^T, \cdots, x_{mn}^T]^T$, where $x_i \in \mathbb{R}^{31\times 1}, 1 \leq i \leq mn$ are the sampled reflectance spectra acquired at the mn distinct pixel locations in the sensor array. The p-color image, $y \in \mathbb{R}^{pmn\times 1}$, is of the form $y = [y_1^T, y_2^T, \cdots, y_p^T]^S$, where $y_i \in \mathbb{R}^{mn\times 1}$ are the column-ordered color channels.

The spectrum incident at the i^{th} location is obtained as $L_i x_i$, where $L_i \in \mathbb{R}^{31\times 31}$ is diagonal with sampled values of the illuminant spectrum as its diagonal elements. Figure 4.3 shows the sampled values of common illuminant spectra. The matrix of illuminant spectra $\bar{L}$ in Eq. (4.3) is formed as $\bar{L} = \text{diag}(L_1, L_2, \cdots, L_{mn})$.

The matrix $\bar{F}$ is of the form

$$\bar{F} = \begin{bmatrix} I_{mn} \otimes f_1^T \\ I_{mn} \otimes f_2^T \\ \vdots \\ I_{mn} \otimes f_p^T \end{bmatrix}, \tag{4.4}$$

where $\otimes$ denotes the Kronecker matrix product and I_{mn} is the $mn \times mn$ identity matrix. In CFA-based image acquisition, only one color is sampled at a particular location. The image formation model in this case is

$$g = Sy. \tag{4.5}$$

S is the subsampling matrix that reduces the pmn samples of y to the vector g of size $mn \times 1$ such that we are left with only one color sample at each pixel location. S is block-diagonal and of the form

$$S = \begin{bmatrix} E_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & E_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & E_p \end{bmatrix}, \tag{4.6}$$

where E_i, $i = 1, 2, \cdots, p$, are row-deficient identity matrices with rows corresponding to missing color samples removed. The dimensions of E_i depend on the CFA sampling arrangement. For instance, in the case of the Bayer-type pattern that samples three distinct colors, $p = 3$, and $E_1, E_3 \in \mathbb{R}^{mn/4 \times mn}$ and $E_2 \in \mathbb{R}^{mn/2 \times mn}$.

The CIEXYZ color space is the most commonly used device-independent color space in colorimetry [56]. Let

$$z = \bar{A}\bar{L}x \tag{4.7}$$

be the column ordered representation of the scene in the CIEXYZ color space when viewed under the illuminant L. The matrix $\bar{A}$ is formed from the color matching functions $\bar{x}$, $\bar{y}$, and $\bar{z}$ that describe the CIEXYZ space such that the product $\bar{A}\bar{L}x$ is the projection of the scene radiances to the CIEXYZ space.

We consider the noise-free case and assume wide-sense stationary signals. The multi-channel Wiener filter estimate $\hat{z}$ of z with respect to g is given by

$$\hat{z} = R_{zg}R_{gg}^{-1}g, \tag{4.8}$$

where $R_{zg} = \mathrm{E}\left\{zg^T\right\}$, $R_{gg} = \mathrm{E}\left\{gg^T\right\}$. Substituting explicit expressions for R_{zg} and R_{gg} gives

$$\hat{z} = \left(\bar{A}\bar{L}R_{xx}\bar{L}^T\bar{F}^TS^T\right)\left(S\bar{F}\bar{L}R_{xx}\bar{L}^T\bar{F}^TS^T\right)^{-1}g, \tag{4.9}$$

where $R_{xx} = \mathrm{E}\left\{xx^T\right\}$.

4.3 Error Criterion

The XYZ tristimulus errors at each location of the array are arranged in a column-ordered form to form the error vector

$$e = z - \hat{z} = (\bar{A} - P)\bar{L}x, \tag{4.10}$$

where

$$P = \left(\bar{A}\bar{L}R_{xx}\bar{L}^T\bar{F}^TS^T\right)\left(S\bar{F}\bar{L}R_{xx}\bar{L}^T\bar{F}^TS^T\right)^{-1}S\bar{F}. \tag{4.11}$$

The XYZ space is not perceptually uniform, i.e., errors in XYZ space do not reflect perceived differences in color. A criterion based on (4.10) will not quantify perceived color reproduction errors and is not suitable for the purpose of defining an accurate objective criterion. A more suitable error criterion would be based on tristimulus differences in a perceptually uniform color space like the CIELAB space. The transformation from XYZ to CIELAB is nonlinear and is given by

$$\begin{aligned} L^* &= 116\left(\frac{Y}{Y_w}\right)^{1/3} - 16, \\ a^* &= 500\left[\left(\frac{X}{X_w}\right)^{1/3} - \left(\frac{Y}{Y_w}\right)^{1/3}\right], \\ b^* &= 200\left[\left(\frac{Y}{Y_w}\right)^{1/3} - \left(\frac{Z}{Z_w}\right)^{1/3}\right], \end{aligned} \tag{4.12}$$

where X_w, Y_w, and Z_w are the XYZ tristimulus values of a reference white stimulus. Let Δu_{XYZ} and $\Delta u_{L^*a^*b^*}$ represent a small distance in the XYZ space and the corresponding

distance in the CIE $L^*a^*b^*$ space respectively. For small differences in XYZ space, the corresponding difference in the nonlinear CIE $L^*a^*b^*$ space may be approximated about a point (X_0, Y_0, Z_0) in XYZ space by the linear transformation

$$\Delta u_{_{L^*a^*b^*}} = J \Delta u_{_{XYZ}}, \tag{4.13}$$

where

$$J = \frac{1}{3} \begin{bmatrix} 0 & 116 & 0 \\ 500 & -500 & 0 \\ 0 & 200 & -200 \end{bmatrix} \cdot \begin{bmatrix} X_w^{-\frac{1}{3}} X_0^{-\frac{2}{3}} & 0 & 0 \\ 0 & Y_w^{-\frac{1}{3}} Y_0^{-\frac{2}{3}} & 0 \\ 0 & 0 & Z_w^{-\frac{1}{3}} Z_0^{-\frac{2}{3}} \end{bmatrix} \tag{4.14}$$

is the gradient of the $XYZ \rightarrow$ CIELAB transform. We linearize the transform from XYZ to CIELAB space about the XYZ coordinates of the illuminant under which the image is acquired.

An error criterion is formed as the expectation of the 2-norm of the aggregated CIELAB errors at all spatial locations of the reconstructed p-color image as

$$\Delta E = \mathrm{E}\left\{ \|\bar{J}e\|_2^2 \right\} = \mathrm{E}\left\{ \|\bar{J}(\bar{A} - P)\bar{L}x\|_2^2 \right\}, \tag{4.15}$$

where $\bar{J} = J \otimes I_{mn}$ achieves the transformation in (4.14) for all mn tristimulus values. Since the 2-norm of a vector is identical to its Frobenius norm, we can use the identity $\mathrm{tr}(ABC) = \mathrm{tr}(BCA)$ to reduce the error criterion to the form

$$\Delta E = \mathrm{tr}\left(\bar{L} R_{xx} \bar{L}^T (\bar{A} - P)^T \bar{J}^T \bar{J} (\bar{A} - P) \right) \tag{4.16}$$

Note that, due to the expression of P in the form in (4.11) the error criterion in (4.16) is dependent only on the general statistics of scene reflectances as described by the block matrix R_{xx} and not on particular scenes.

In order to be physically realizable, the filter sensitivities should be reasonably smooth. To this end, we define a cost on the roughness of the individual filter sensitivities as

$$\varepsilon = k_s \sum_k \|Lf_i\|_2^2, \tag{4.17}$$

where k_s is a scaling factor that can be used to specify the relative importance of smoothness with respect to data fidelity. L is the Laplacian matrix that yields a roughness estimate. The criterion in (4.16) is amended to incorporate the weight on roughness in (4.17). To account for image acquisition under different illuminants, the criterion is optimized over a sum of the standard D_{65}, D_{75}, and uniform illuminants (radiance at each wavelength is one). The revised criterion is given by

$$\Phi = \Delta E_{D_{65}} + \Delta E_{D_{75}} + \Delta E_{\text{uniform}} + \varepsilon. \tag{4.18}$$

Optimal values of filter sensitivities that minimize color reproduction errors in the perceptually uniform linearized CIELAB space are obtained by minimizing (4.18) with respect to the filter sensitivities in F.

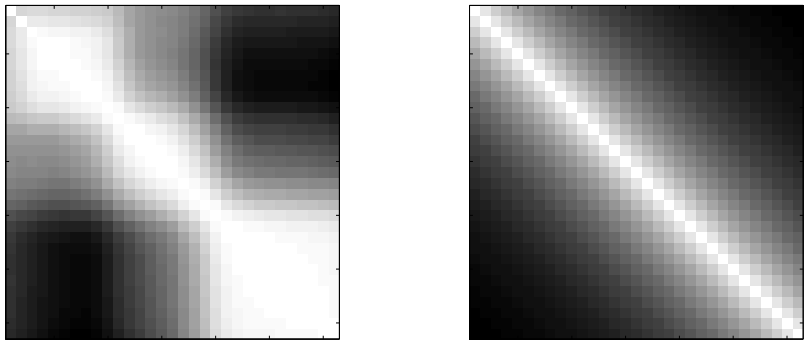
4.4 Correlation matrix model

For a fixed set of filter sensitivities F, the optimization criterion is a function of image statistics represented by R_{xx}. The block matrix R_{xx} has the form

$$R_{xx} = \mathrm{E}\left\{\begin{bmatrix} x_1x_1^T & x_1x_2^T & \cdots & x_1x_{mn}^T \\ x_2x_1^T & x_2x_2^T & \cdots & x_2x_{mn}^T \\ \vdots & \vdots & \ddots & \vdots \\ x_{mn}x_1^T & x_{mn}x_2^T & \cdots & x_{mn}x_{mn}^T \end{bmatrix}\right\}$$
$$= \begin{bmatrix} R^{(1,1)} & R^{(1,2)} & \cdots & R^{(1,mn)} \\ R^{(2,1)} & R^{(2,2)} & \cdots & R^{(2,mn)} \\ \vdots & \vdots & \ddots & \vdots \\ R^{(mn,1)} & R^{(mn,2)} & \cdots & R^{(mn,mn)} \end{bmatrix}, \tag{4.19}$$

and has spectral autocorrelation matrices for each pixel location as its diagonal blocks and inter-pixel spectral crosscorrelation matrices as its off-diagonal blocks.

A correlation matrix of the form in (4.19) contains elements due to the spectral correlation at a single spatial location and spatial correlations across the image. We will rely on a numerical optimization of (4.18) using Matlab's `fmincon` routine to arrive at an optimal value for F. The routine carries out a gradient-based search for the minimum and requires multiple computations of Φ and $\partial\Phi/\partial f_i$. This requires multiple products of large-dimension matrices of the type $\left(S\bar{F}\bar{L}R_{xx}\bar{L}^T\bar{F}^TS^T\right)$ seen in (4.11). The challenge in modeling a generalized R_{xx} applicable for all acquired scenes lies in forming blocks of R_{xx} that give it a regular structure that lends itself to optimized computation while not defying the statistical



(a) Super-image (b) Proposed model

Figure 4.4: The spectral correlation matrix $R^{(1,1)}$ for (a) the super-image obtained by accumulating spectral data from all 22 sample images together and (b) for the proposed model

properties of any particular scene. In this section we propose such a model for R_{xx} and suggest means to exploit the regular structures of the matrices involved to reduce computation times.

It is not unreasonable to assume that for most images, natural or otherwise, spatial correlations cease to be significant at a distance. This assumption follows from the local smoothness seen in most acquired images. In this work, for computational tractability, we assume that spatial correlation is insignificant at a distance greater than 16 pixel-widths. We further assume that the spectral correlation at a particular location is separable from the spatial correlation. This allows a representation of R_{xx} that contains blocks that are all functions of $R^{(1,1)}$.

Here, we make the following assumptions in forming R_{xx}:

1. The spectral correlation at a particular spatial location is separable from the spatial correlation in samples across the acquired image in a particular wavelength band.

Figure 4.5: Sample multispectral images from Hordley et al. [70] rendered in sRGB space for the D_{65} illuminant.

2. The correlation coefficients that form the elements of $R^{(k,k)}$ (4.19) are an exponential function of wavelength separation.

3. The correlation matrices $R^{(k,l)}$, that incorporate spatial correlations in the image bands are an exponential function of spatial distance.

4. The acquired image is spatially periodic. This yields an R_{xx} that is block-circulant with block-circulant Toeplitz blocks.

Let $r_{(k,l)}^{(i,j)}$, $1 \leq i,j \leq 31$, be an element of $R^{(k,l)}$, $1 \leq k,l \leq mn$ in the correlation matrix R_{xx}. Consider the diagonal blocks of R_{xx} $(k = l)$. As $\mathrm{E}\left\{x_k^i x_k^j\right\} = \mathrm{E}\left\{x_k^j x_k^i\right\}$, $r_{(k,k)}^{(i,j)} = r_{k,k}^{(j,i)}$

and due to the assumption of stationarity of the reflectance spectrum, $R^{(k,k)}$ are symmetric, Toeplitz matrices.

We define the elements of R_{xx} as follows.

$$r_{(1,1)}^{(i,j)} = a_1 + b_1 e^{-\alpha_1/\beta_1} \tag{4.20}$$

$$R^{(k,l)} = \left(a_2 + b_2 e^{-\alpha_2/\beta_2}\right) R^{(1,1)}, \tag{4.21}$$

where the parameters a_1, a_2, b_1, and b_2 are set at 0.25, 0.25, 0.75, and 0.75 respectively. Also, $\beta_1 = \beta_2 = 10$, and $\alpha_1 = |i-j|$, $\beta_1 = 10$, $\alpha_2 = \sqrt{(|l-k| \mod m)^2 + \text{floor}\,(|l-k|/n)^2}$ represent distances in one and two dimensions respectively.

Due to the nature of column-ordering of the scene reflectances and the periodicity of the image, R_{xx} may be expressed as

$$R_{xx} = D \otimes R^{(1,1)}, \tag{4.22}$$

where $D \in R^{mn \times mn}$ is block-circulant with circulant blocks.

Hordley et al. [70] have provided a database of multi-spectral images (Fig. 4.5) captured using a Spectracube© camera. These spectral images were acquired under controlled conditions under the D_{75} illuminant. A correction for the illuminant was applied to arrive at reflectance values sampled every 10 nm in the range 400–700 nm. Figure 4.4 shows representations of $R^{(1,1)}$ for the proposed model and for the super-image formed by accumulating spectral data from all 22 images shown in Fig. 4.5.

Computing Φ

The computation of P requires an inversion of the $mn \times mn$ matrix $\left(S\bar{F}\bar{L}R_{xx}\bar{L}^T\bar{F}^TS^T\right)$ and a computation of the similar matrix $\bar{A}\bar{L}R_{xx}\bar{L}^T\bar{F}^T$. Note that $\bar{A}$ and $\bar{F}$ are formed from the x,y,z color matching functions and f_i respectively as given in (4.4). The matrix product $\bar{F}W\bar{F}^T$, where $W = \bar{L}R_{xx}\bar{L}^T$, is given by

$$\bar{F}W\bar{F}^T = \begin{bmatrix} F_1WF_1^T & F_1WF_2^T & \cdots & F_1WF_p^T \\ F_2WF_1^T & F_2WF_2^T & \cdots & F_2WF_p^T \\ \vdots & \vdots & \ddots & \vdots \\ F_pWF_1^T & F_pWF_2^T & \cdots & F_pWF_p^T \end{bmatrix}, \tag{4.23}$$

where $F_k = I_{mn} \otimes f_k^T, 1 \leq k \leq p$. Now,

$$\bar{F}W = \begin{bmatrix} F_1W \\ F_2W \\ \vdots \\ F_pW \end{bmatrix}, \tag{4.24}$$

where F_kW is obtained as in (4.25) and the blocks of $\bar{F}W\bar{F}^T$ are formed as in (4.26), where $1 \leq j, k \leq p$.

$$F_k W = \begin{bmatrix} f_k^T LR^{(1,1)} L^T & f_k^T LR^{(1,2)} L^T & \cdots & f_k^T LR^{(1,mn)} L^T \\ f_k^T LR^{(2,1)} L^T & f_k^T LR^{(2,2)} L^T & \cdots & f_k^T LR^{(2,mn)} L^T \\ \vdots & \vdots & \ddots & \vdots \\ f_k^T LR^{(mn,1)} L^T & f_k^T LR^{(mn,2)} L^T & \cdots & f_k^T LR^{(mn,mn)} L^T \end{bmatrix} \tag{4.25}$$

$$F_j W F_k^T = \begin{bmatrix} f_j^T LR^{(1,1)} L^T f_k^T & f_j^T LR^{(1,2)} L^T f_k^T & \cdots & f_j^T LR^{(1,mn)} L^T f_k^T \\ f_j^T LR^{(2,1)} L^T f_k^T & f_j^T LR^{(2,2)} L^T f_k^T & \cdots & f_j^T LR^{(2,mn)} L^T f_k^T \\ \vdots & \vdots & \ddots & \vdots \\ f_j^T LR^{(mn,1)} L^T f_k^T & f_j^T LR^{(mn,2)} L^T f_k^T & \cdots & f_j^T LR^{(mn,mn)} L^T f_k^T \end{bmatrix} \tag{4.26}$$

Since $LR^{(1,1)}L^T$ is symmetric, we let $f_j^T LR^{(1,1)} L^T f_k^T = f_k^T LR^{(1,1)} L^T f_j^T = R_{(|j-k|)}$. $\bar{F} W \bar{F}^T$ can be written as

$$\bar{F} W \bar{F}^T = \bar{R}_F \otimes D, \tag{4.27}$$

where

$$\bar{R}_F = \begin{bmatrix} R_{(0)} & R_{(1)} & \cdots & R_{(p-1)} \\ R_{(1)} & R_{(0)} & \cdots & R_{(p-2)} \\ \vdots & \vdots & \ddots & \vdots \\ R_{(p-1)} & R_{(1)} & \cdots & R_{(0)} \end{bmatrix} \tag{4.28}$$

Consider now the product $S\bar{F}W\bar{F}^TS^T$. Pre and post-multiplication of a matrix by S and S^T is equivalent to eliminating rows and columns according to the structures of the E_i that constitute S. For instance, in the Bayer case, the following rows and columns will be eliminated due to E_1: $2:2:m, m+1:2m, 2m+2:2:3m, 3m+1:4m, \quad \cdots \quad mn$. $S\bar{F}W\bar{F}^TS^T$ can be written as

$$S\bar{F}W\bar{F}^TS^T = \left(\bar{R}_F \otimes D\right)_{\downarrow}, \tag{4.29}$$

where $\{.\}_{\downarrow}$ is decimation by inspection of S. We can use (4.27) and (4.29) to reduce the computation of P in (4.11) from a successive product of large matrices to the very much simpler Kronecker products to get

$$P = \left(\bar{R}_A \otimes D\right)\left(\left(\left(\bar{R}_F \otimes D\right)_{\downarrow}\right)^{-1}\right)_{\uparrow}, \tag{4.30}$$

where $\{.\}_{\uparrow}$ is the addition (by inspection) of rows and columns of zeros to a $mn \times mn$ matrix to get a matrix of size $3mn \times mn$ and $\bar{R}_A$ is formed from the XYZ color matching functions similarly to $\bar{R}_F$ in (4.28).

4.5 Experiments and Discussion

The varied design parameters for CFA patterns have led to a wide variety of distinct CFA arrangements. A number of CFA arrangements have appeared in the literature and been used commercially [48, 52, 18]. Figure 4.6 shows the periodic CFA arrangements considered in our experiments. All but the Gindele CFA shown in Fig. 4.6 are 3-color (typically,

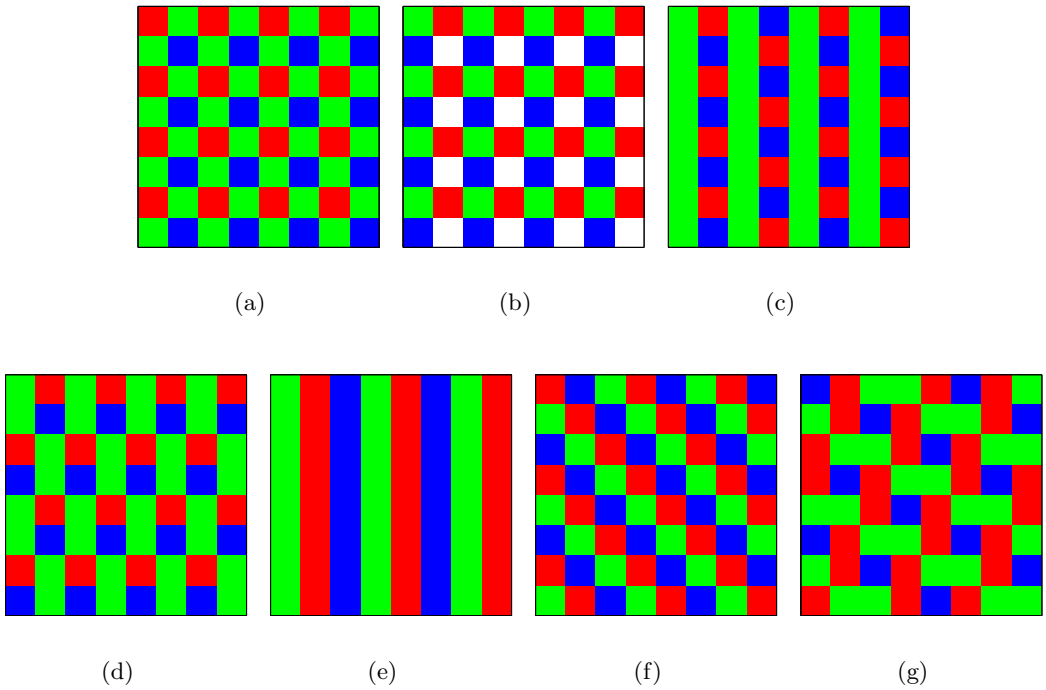


Figure 4.6: Common periodic CFAs. (a) Bayer [14], (b) Gindele [18], (c) Yamanaka [48], (d) Lukac [49], (e) striped, (f) diagonal striped [49], (g) CFA based on the Holladay halftone pattern [50].

RGB) CFAs. The Gindele CFA has a photo-site that samples luminance (represented by white in Fig. 4.6 (b)) in addition to RGB. In our experiments we set the transmittance function of the luminance photosites to a perfectly transmissive filter (one at all wavelengths) and optimize for only three colors.

The optimization framework detailed in Section 4.3 is used to obtain optimal color filter transmittance functions for the CFA patterns shown in Fig. 4.6. In each case, the sampling matrix in (4.6) is constructed according to the respective CFA arrangements in Fig. 4.6. The `fmincon` routine in MATLAB® is used to carry out a constrained minimization of the criterion in (4.18) to arrive at optimal estimates for f_i. The constraints on f_i are that the

elements of $f_i \in [0,1]$. Since a global minimum is not assured, we generate a number of distinct initial conditions and choose the least value of the criterion to arrive at the filter sensitivities f_i. Initial values of f_i are assumed to be Gaussian curves in the wavelength domain and are given by

$$f_i(\lambda) = e^{-\frac{(\lambda-\mu_i)^2}{\sigma_i^2}}. \tag{4.31}$$

An initial condition is defined by the vectors $\mu = [\mu_1, \mu_2, \cdots, \mu_p]^T$ and $\sigma = [\sigma_1, \sigma_2, \cdots, \sigma_p]^T$ that contain the means and standard deviations respectively of the p Gaussian curves that represent the filter sensitivities. The elements of μ and σ are uniform random variables defined over the visible spectrum with the following constraints: $400 \leq \mu_i \leq 700$ and $\rho \leq \sigma_i \leq 10\rho$, where $\rho = 30$. Optimal filter sensitivities obtained for the CFA patterns in Fig. 4.6 are shown in Fig. 4.7.

4.5.1 Evaluation

The simulation pipeline shown in Fig. 4.8 was used to evaluate the performance of the optimized color filter transmittance functions obtained in Section 4.5 with respect to typical RGB and CMY filter transmittances obtained from ImagEval's vCamera toolbox [71] (Fig. 4.1). Figure 4.8 illustrates the methods used for objective and subjective comparisons.

The s-CIELab ΔE error metric [37] is a measure of perceptual difference between color images that takes into account the lowpass nature of the human visual system and its specific response to luminance and chrominance components of images. The s-CIELab ΔE is thus a more accurate measure of perceptual errors between color images than MSE values in sRGB space. The s-CIELab error metric was used as an objective measure to quantify

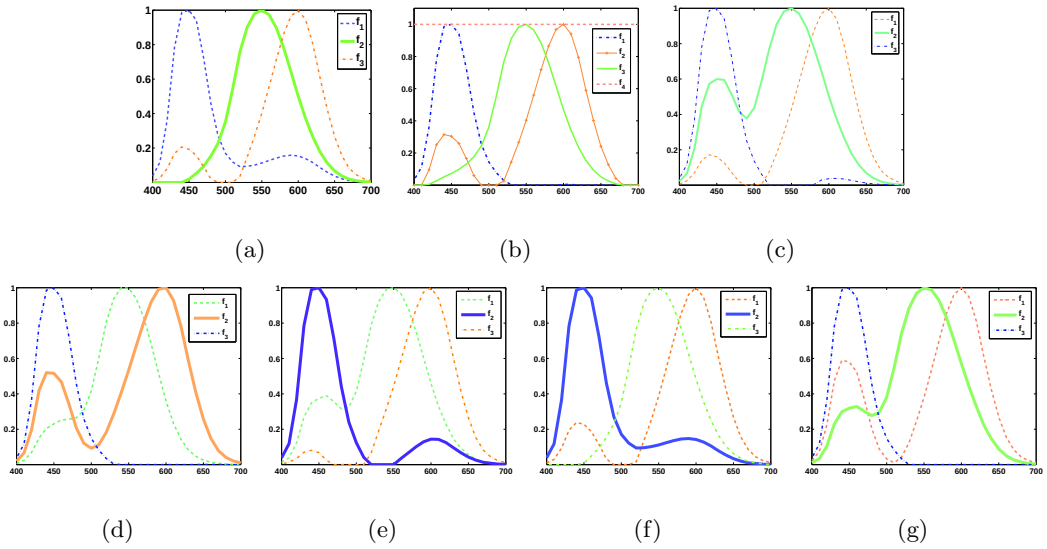


Figure 4.7: (a)-(g) Optimal spectral sensitivity functions obtained for the CFA patterns shown in Figs. 1(a)-1(g) respectively. Ordinates represent normalized transmittances. The colors of transmittance curves are sRGB values for the respective spectra. Bolder lines correspond to the optimal sensitivities obtained at the location of the green filter in the respective CFA patterns.

the error between the original image as obtained from the CIEXYZ tristimulus values of the multispectral images and the CIEXYZ values of corresponding demosaicked images.

Table 4.1 gives average s-CIELab ΔE values for all images from the multispectral image database for each of the CFA patterns in Fig. 4.6 as acquired with the RGB, CMY, and the optimized sensitivities. These results demonstrate that the optimized color filters perform significantly better than the standard RGB and CMY filters for all CFA arrangements. Table 4.2 gives the average s-CIELab ΔE values for the images in Fig. 4.5 when acquired using the Bayer CFA with the RGB, CMY, and optimized transmittances respectively under the standard D_{65} and D_{75} illuminants. Table 4.2 shows that the optimal transmittances perform significantly better than the typical RGB and CMY color filters across the range of images represented in the multispectral database.

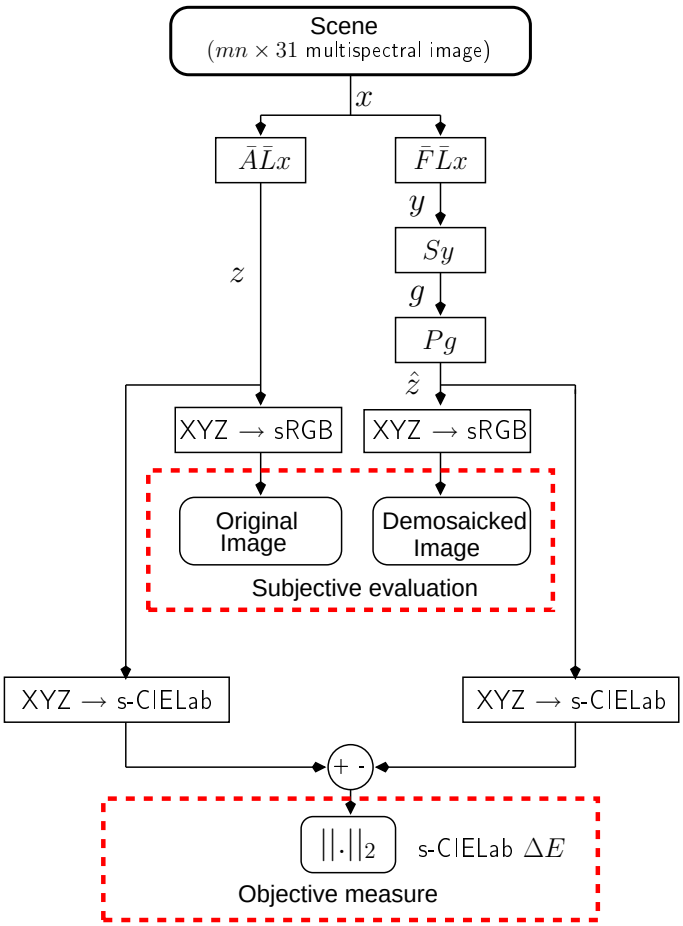


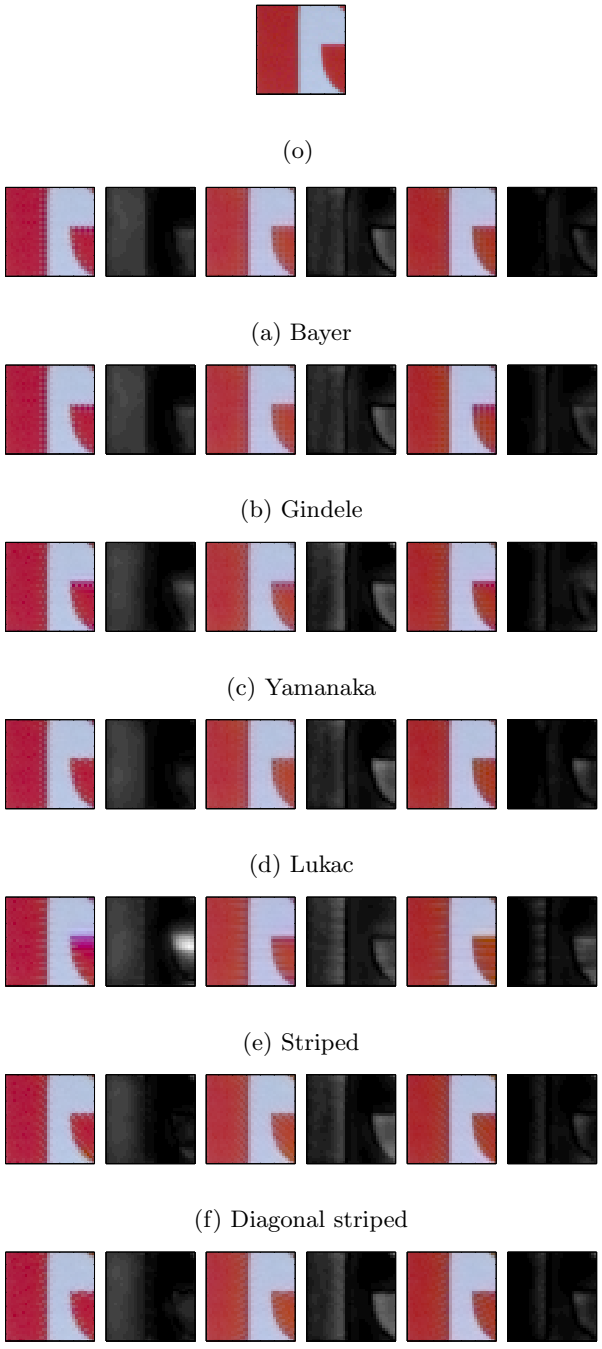
Figure 4.8: The simulation pipeline. All variables are as described in preceding sections.

Subjective evaluation involves visual inspection of the original multispectral images from the database in [70] rendered in the standard sRGB color space [72] against the corresponding demosaicked images obtained after subsampling with the CFA patterns being considered. Note that since the objective here is to compare the performance of color filter sensitivity functions, we used the linear MMSE estimator in (4.9) for demosaicking even though many sophisticated demosaicking algorithms are known.

Table 4.1: Average s-CIELab ΔE values. All images in [70] are sampled according to the patterns in Fig. 4.6 and the demosaicked results are compared with the original multispectral scene as shown in Fig. 4.8. The smallest error values are shown in bold font.

Pattern	ΔE_S RGB	ΔE_S CMY	ΔE_S optimal
(a) Bayer	6.0911	5.9310	**3.2724**
(b) Gindele	6.5438	5.8246	**3.9508**
(c) Yamanaka	6.5051	6.6076	**3.9725**
(d) Lukac	6.2251	6.3891	**3.3761**
(e) Striped	7.3646	8.4028	**5.1840**
(f) Diagonal	6.3924	6.3206	**3.2203**
(g) Utah-dot	6.6039	6.7034	**3.6788**

Figures 4.9 and 4.10 show results of a subjective comparison for two images cropped from images 3 and 4 in Fig. 4.5 respectively. The top row in each figure shows the original image. Subsequent rows show results of demosaicking of images sampled with the CFA patterns in Fig. 4.6. On each row, for a particular CFA pattern, demosaicking results of images sampled with RGB, CMY, and optimized filter sensitivities appear from left to right. In addition, the s-CIELab ΔE error images corresponding to each reconstructed image is shown to its right. Note that reconstruction errors appear both as chrominance errors (hue-shifts most apparent in smooth areas of the image) and luminance errors (spatial artifacts most apparent at edges). The images obtained by sampling with optimized sensitivity functions show fewer luminance and chrominance artifacts. The reduced s-CIELAB ΔE is apparent from the error images in both figures. The improvement is most striking for the striped CFA pattern (Figs. 4.9(e) and 4.10(e)),



(o)

(a) Bayer

(b) Gindele

(c) Yamanaka

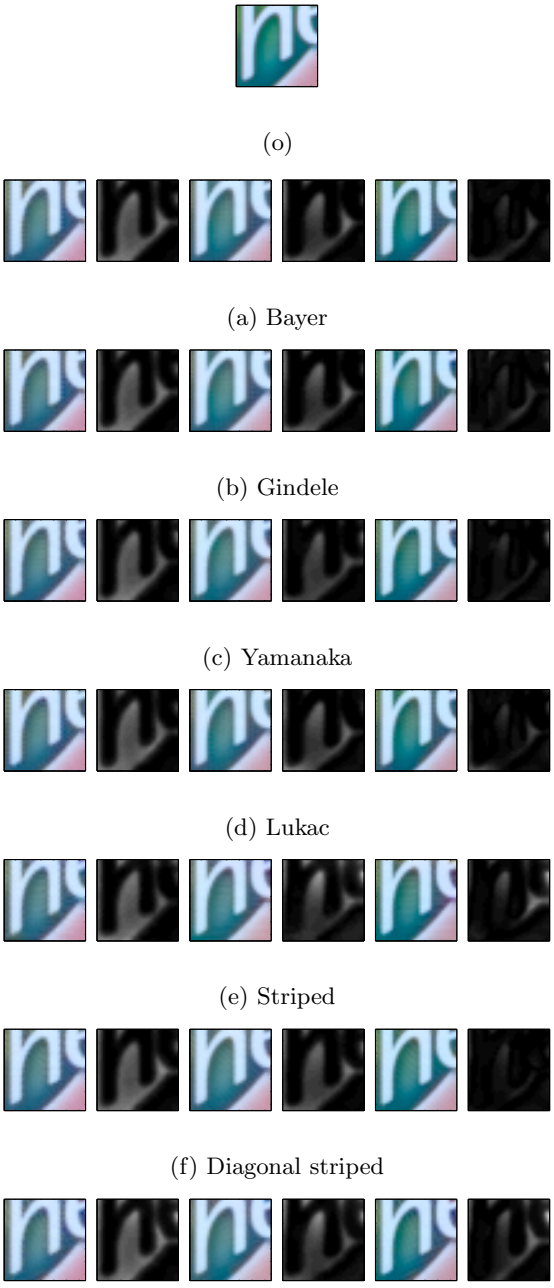
(d) Lukac

(e) Striped

(f) Diagonal striped

(g) Holladay halftone based CFA

Figure 4.9: sRGB representations (for the D_{65} illuminant) of an image cropped from image 3 from the database of multispectral images [70]. (o) Original image. (a)-(g) From left to right — Images reconstructed from the CFA sampled images obtained from the RGB, CMY, and optimized color filters respectively. s-CIELab ΔE error images appear to the right of each reconstructed image.



(o)

(a) Bayer

(b) Gindele

(c) Yamanaka

(d) Lukac

(e) Striped

(f) Diagonal striped

(g) Holladay halftone based CFA

Figure 4.10: sRGB representations (for the D_{65} illuminant) of an image cropped from image 4 from the database of multispectral images [70]. (o) Original image. (a)-(g) From left to right — Images reconstructed from the CFA sampled images obtained from the RGB, CMY, and optimized color filters respectively. s-CIELab ΔE error images appear to the right of each reconstructed image.

Table 4.2: Average s-CIELab ΔE values for images obtained from images in [70] (Fig. 4.5) for the Bayer CFA. The smallest error values are shown in bold font.

Image	D_{65}			D_{75}		
	ΔE_S RGB	ΔE_S CYM	ΔE_S 3-color	ΔE_S RGB	ΔE_S CYM	ΔE_S 3-color
1	4.0427	3.7018	2.6672	3.6251	3.9869	2.5705
2	**1.6187**	2.5591	1.7399	**1.6010**	2.6234	1.7249
3	5.2446	5.1769	**2.1377**	5.2770	4.8766	**2.0960**
4	5.3974	4.0390	**2.3506**	5.4185	3.9630	**2.3482**
5	7.4964	6.1835	**3.0427**	7.5834	5.8516	**3.0135**
6	17.5901	15.5977	**4.8360**	16.8534	14.9227	**4.3995**
7	5.2221	3.7120	**3.1922**	5.2917	3.8785	**3.2260**
8	4.3750	5.3385	**2.8245**	4.2959	5.7677	**2.8491**
9	1.4525	1.8873	**1.2786**	1.4795	1.8842	**1.2937**
10	6.4883	12.2503	**5.8876**	6.4916	12.1880	**5.6871**
11	1.9786	2.1576	**1.7307**	2.0047	2.1994	**1.7530**
12	**2.3159**	3.8119	3.0739	**2.4500**	3.9217	3.1051
13	5.9047	5.8125	**2.8353**	6.0141	5.7663	**2.8907**
14	5.3764	4.8572	**2.6532**	5.4475	4.8765	**2.6403**
15	**5.3252**	8.4890	6.8539	**5.5803**	8.6885	6.9568
16	**1.8925**	2.7593	2.2233	**1.9273**	2.9487	2.1937
17	6.8992	6.1207	**4.1638**	6.8540	5.6755	**3.9818**
18	**2.3528**	2.7197	3.7209	**2.2742**	2.9810	3.6197
19	18.7389	12.4805	**2.8845**	18.6495	12.4153	**2.8835**
20	7.7903	3.4098	**2.7930**	7.9236	3.2726	**2.6811**
21	7.8062	5.7991	**1.3551**	7.8522	6.3095	**1.3747**
22	8.6948	11.6179	**7.7472**	8.6027	11.5274	**7.7411**
Average	6.0911	5.9310	**3.2724**	6.3638	6.2054	**3.4178**

where the RGB and CMY transmittance functions lead to very visible artifacts at edges, while the optimized transmittance functions greatly reduce these artifacts.

4.6 Conclusions

In this chapter we have demonstrated the significant effect that the color filter spectral sensitivity functions have on the quality of reconstruction of CFA images. A unified spatial-chromatic sampling framework is proposed for the optimization of the color filter sensitivity functions. The proposed method optimizes an error criterion that is dependent not on particular images but on the general statistics of multispectral images. Optimal color filter sensitivity functions for several periodic CFAs are obtained, and optimized color filter transmittances are shown to perform significantly better than CFAs with standard RGB and CMY color filters in terms of the s-CIELab ΔE metric and subjective comparisons.

Chapter 5

Sample selection in color filter arrays

5.1 Introduction

To be able to display a color image in print or on a display device, we need full information about at least three color primaries at each pixel location. As described in Chapter 2, although multiple image sensors that capture distinct spectral channels would provide best results, only a few professional digital cameras employ multiple sensors. To avoid high costs, optical complexity, and problems with image registration that occur due to acquisition at different planes, most color cameras rely on a single sensor-array to acquire a color image. A *mosaic* of color filters is overlaid on the imaging sensor to achieve sparse sampling. At each location the color filter reflects all wavelengths of light except the wavelengths in the range that describe the desired color. The missing color samples at each location are reconstructed in a post-processing step commonly referred to as *demosaicking*. In this chapter, we propose two methodologies for the selection of sample locations in a color filter array (CFA). Both methods allow for the effects of the HVS and use sequential selections methods in conjunction with image restoration techniques to arrive at optimal sampling patterns. The two methods differ in the choice of restoration techniques and color spaces used for restoration. The first method uses regularization for restoration in the RGB space while the second method uses Wiener filtering in a perceptually uniform color space for restoration.

5.2 Background

A variety of sampling patterns have been proposed by researchers in the field. One of the earliest and most popular array patterns was developed by Bayer [14] in 1975. Most reconstruction algorithms proposed in the literature are tailored for the Bayer array. The Bayer array (Fig. 5.2) differentiates between luminance sensitive and chrominance sensitive elements and is configured such that each element type appears in a repeated pattern. Since the HVS is more sensitive to degradation in luminance and detects intensity variations more readily than color variations, the Bayer pattern uses an excess of luminance elements. The HVS luminance response (Fig. 2.1(a)) correponds very closely with the spectrum of the color green and thus green is sampled at every other pixel location along a row and is staggered by one pixel in adjacent rows. The blue and red channels are used in alternate rows to complete the mosaic.

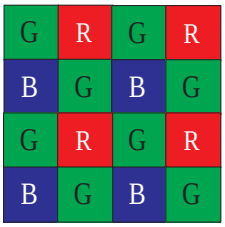


Figure 5.1: The Bayer Array

Various improvements to the Bayer array have been suggested in recent years. Yamagami et al. [15] describe an array pattern that addresses the problem of saturation in the green channel of the Bayer array. The Yamagami array has alternating green and luminance samples in a row adjacent to alternating red, luminance, and blue samples. The improvement in color reproduction is achieved at the cost of spatial resolution. Hamilton et al. [16] propose a CFA that uses the subtractive primaries cyan, magenta, and yellow in addition to

green filters to address the issue of photon acquisition in low-light conditions. Zhu et al. [17] address the issue of aliasing artifacts that arise due to high-frequency periodic patterns in an image. They propose a CFA pattern based on random arrays derived from a blue noise pattern and also present a demosaicking scheme for their array. Gindele and Gallagher [18] address saturation issues and propose an array that contains luminance samples in addition to the color primaries red, green, and blue.

A CFA that uses the subtractive primaries cyan, magenta, and yellow in its mosaic offers advantages in terms of the signal to noise ratio and dynamic range. But, after demosaicking, the color values must be converted to red, blue, and green for display. This conversion is achieved via a matrix multiplication that amplifies noise and also leads to color distortion [73]. In the following sections we only consider CFAs that consist of red, green, and blue photosites. In Chapter 4 we propose a method for the selection of optimal sensitivity functions for individual color channels.

5.3 Sample selection based on regularization

The appearance of the reconstructed image depends fundamentally on the characteristics of the HVS. The perceived image is the result of processing of color and spatial variances in the image by the HVS. This motivates the use of an HVS model to evaluate the performance of color image reconstruction. The sample selection method described in this section uses a simple model of the human visual system to characterize the perceptual error in an image reconstructed from a sub-sampled CFA. A sequential algorithm is used to select samples that minimize an error criterion that incorporates the effect of the HVS.

5.3.1 Human color vision model

The image processing flow for an image captured with a digital camera and viewed by an observer has multiple steps (Fig. 5.2). To get an accurate description of the perceived image, the PSFs of the demosaicking process and the HVS must be known precisely. In this treatment, we use a rudimentary model for the PSFs of the three color channels based on a functional model of the low-contrast photopic modulation transfer function (MTF) of the HVS described by Sullivan et al. [74]. We assume that the MTF of the entire work-flow of the digital camera retains the dominating characteristics of the HVS in that:

1. it is more sensitive to spatial frequencies in the vertical and horizontal directions, and
2. the response of chrominance channels falls faster than the response of the luminance channel.

We also also assume that the green channel corresponds closely to the luminance response.

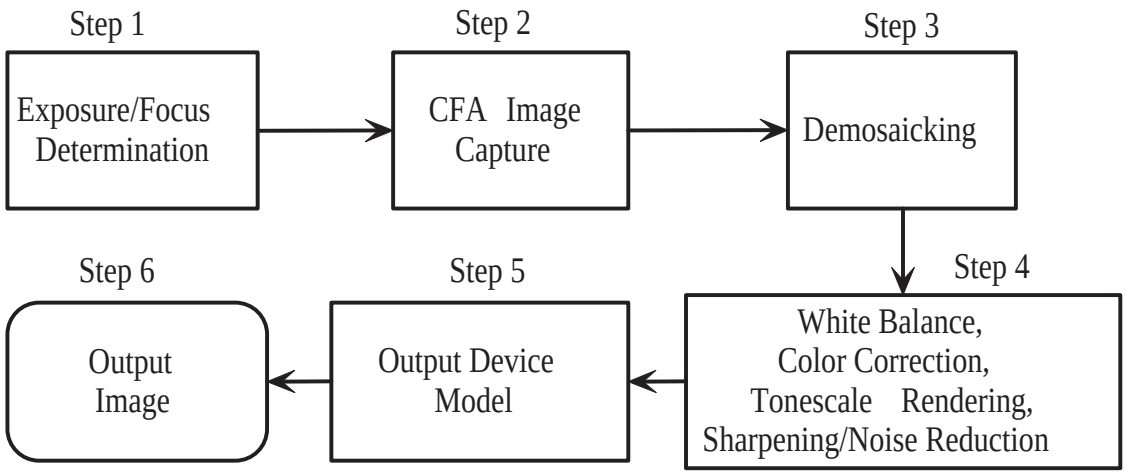


Figure 5.2: A typical image processing pipeline in a color digital camera

The MTF of the green channel is obtained from the MTF as described by Sullivan as

$$V_{G_{ij}} = \begin{cases} a(b + c\bar{f}_{ij}) \exp\left(-\left(c\bar{f}_{ij}\right)^d\right), & \text{if } \bar{f}_{ij} > f_{max} \\ 1.0, & \text{otherwise,} \end{cases} \tag{5.1}$$

where the constants a, b, c, and d are calculated from empirical data to be 2.2, 0.192, 0.114 and 1.1 respectively; $\bar{f}_{ij}$ is the radial spatial frequency in cycles/degree as subtended by the image on the human eye scaled for the viewing distance, and f_{max} is the frequency corresponding to the peak of V_{ij}. Since we need the MTF in terms of discrete linear frequencies along the vertical and horizontal directions (f_i, f_j), we must express (f_i, f_j) in terms of the radial frequency $\bar{f}_{ij}$. The discrete frequencies along the horizontal and vertical directions depend on the pixel pitch Δ of the output device (print or display device) and the total number of frequencies M. A location (i, j) in the frequency domain corresponds to the following f_i and f_j in cycles/mm:

$$\begin{aligned} f_i &= \frac{i-1}{\Delta M}, \\ f_j &= \frac{j-1}{\Delta M}. \end{aligned} \tag{5.2}$$

The linear frequencies are scaled for the viewing distance s and converted to radial frequency as

$$f_{ij} = \frac{\pi}{180 \arcsin\left(\frac{1}{\sqrt{1+s^2}}\right)} \sqrt{f_i^2 + f_j^2}. \tag{5.3}$$

The MTF is not uniform along all directions. The HVS is most sensitive to spatial variation along the horizontal and vertical directions. To account for this variation, the MTF is normalized by an angle dependent function $s(\theta_{ij})$ such that

$$\bar{f}_{ij} = \frac{f_{ij}}{s(\theta_{ij})}, \tag{5.4}$$

where

$$s(\theta_{ij}) = \frac{1-w}{2}\cos(4\theta_{ij}) + \frac{1+w}{2}, \tag{5.5}$$

with w being a symmetry parameter and

$$\theta_{ij} = \arctan\left(\frac{f_j}{f_i}\right). \tag{5.6}$$

The response obtained for the green channel for $w = 0.7$, and a viewing distance of 45 cm and a pixel pitch of 0.27 mm is shown in Fig. 5.3.

The response of the HVS to chrominance, or the contrast sensitivity to spatial variations in the chrominance channels, falls off faster than the response to the luminance channel. A simple chrominance response model corresponding to a decaying exponential is chosen as a basis for the HVS response to the blue and red channels. The red and blue channel response is modelled as

$$V_{B,R}(f_{ij}) = e^{(-0.15 f_{ij})}, \tag{5.7}$$

The response obtained for the red and blue channels is shown in Fig. 5.4.

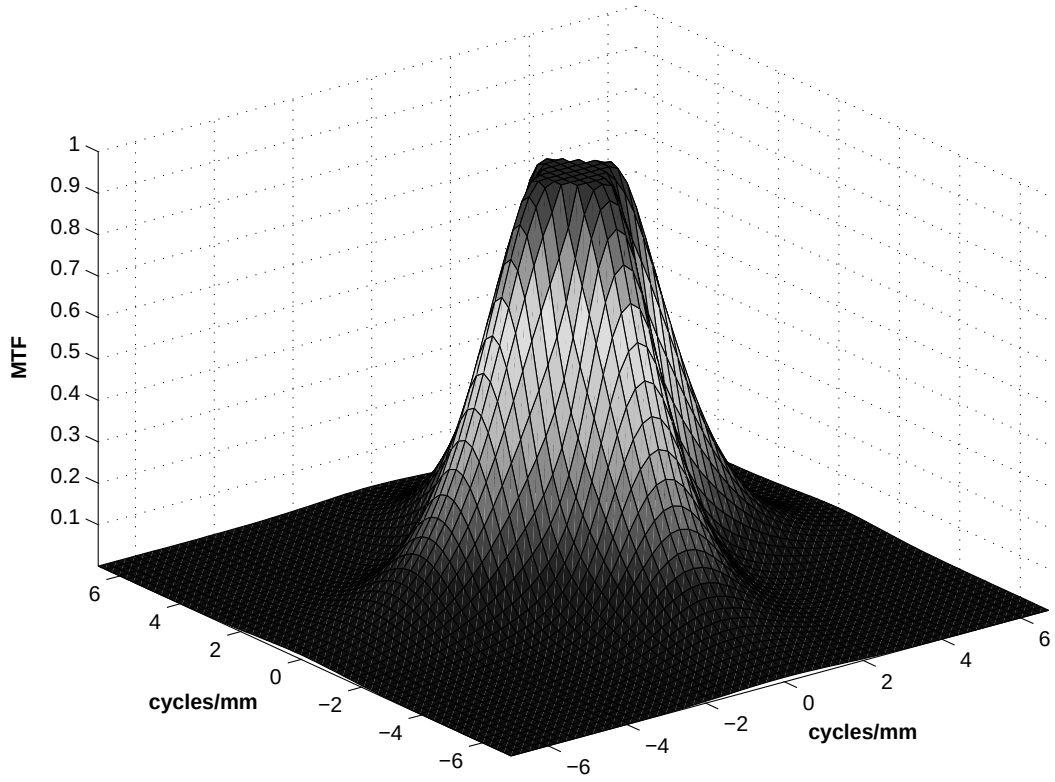


Figure 5.3: HVS green channel MTF

The HVS point spread functions h_i for i = Red, green, blue are obtained as

$$\begin{aligned} h_G &= \mathcal{F}^{-1}\left\{V_G(i,j)\right\}, \\ h_{R,B} &= \mathcal{F}^{-1}\left\{V_{R,B}(i,j)\right\}. \end{aligned} \tag{5.8}$$

The matrices H_i are constructed from h_i such that multiplication of a column-ordered image by H_i yields the 2-D convolution of the image by the point spread function h_i.

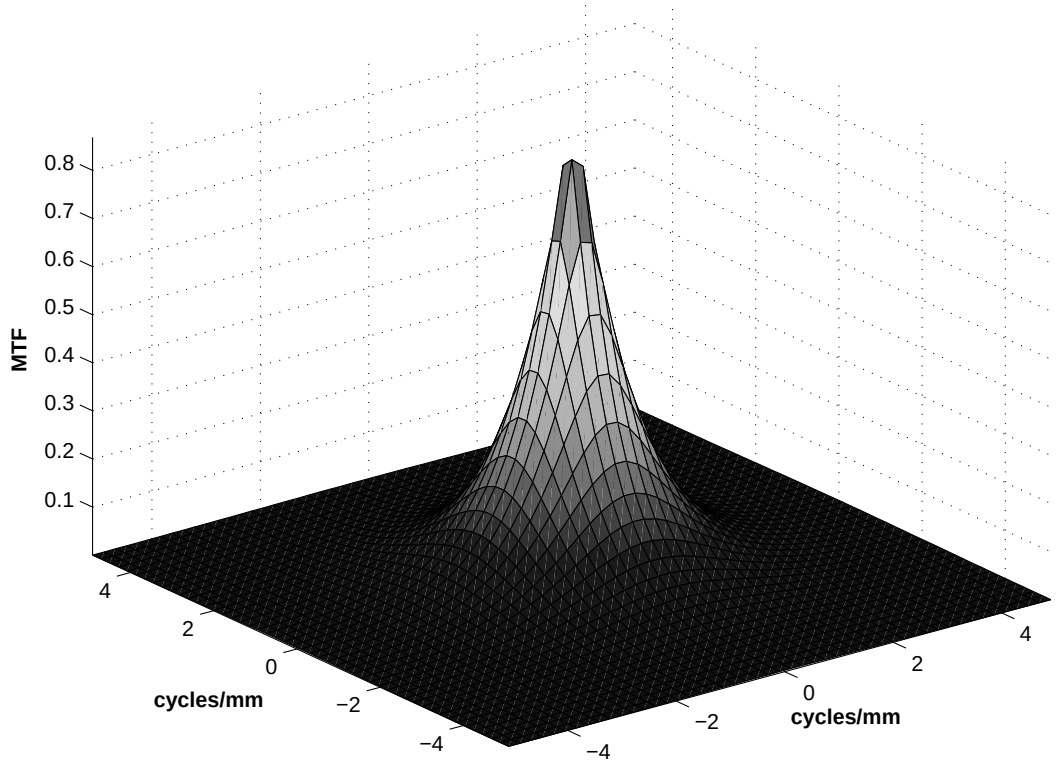


Figure 5.4: HVS red and blue channel MTFs

5.3.2 Mathematical model

We model the sub-sampled image as a linear transformation that maps the full-color image to an image that contains only one color value at a particular pixel location. The sub-sampled image is represented as

$$y_i = A_i x_i + u_i, \qquad i = \text{red}, \text{green}, \text{blue}, \tag{5.9}$$

where x_i, $(mn \times 1)$ and y_i, $(mn \times 1)$ are the red, green and blue channels of the original and the sub-sampled $m \times n$ images arranged in a column-ordered form, and u_i, $(mn \times 1)$,

are the similarly arranged noise terms. The matrices A_i are the *sampling matrices*. For the fully-sampled case, A_i are identical to the $mn \times mn$ identity matrix. For the sub-sampled case, the matrices A_i contain only the rows corresponding to a sampled pixel location. We assume that the image and noise are uncorrelated.

We form a regularization functional for each channel that contains an energy bound on the residual $A_i x - y_i$ and a penalty on the roughness as:

$$\Phi_i = \|A_i x_i - y_i\|_2^2 + \mu_i L_i {x_i}^2. \tag{5.10}$$

The estimate of x_i found on minimizing the constrained least squares problem in (5.10) is

$$\hat{x}_i = (A_i^H A_i + \mu_i L_i^H L_i)^{-1} A_i^H y_i, \tag{5.11}$$

where A^H is the Hermitian transpose of A. To obtain the best estimate for the perceived image, we minimize the discrepancy in the reconstructed image when viewed through the HVS. Let the matrices H_i, i = Red, green, blue, represent the filtering effect corresponding to the point spread functions (PSFs) of the red, green and blue channels of the HVS respectively. We form a discrepancy function for one channel (dropping the subscript) as

$$d = E\{\|Hx - H\hat{x}\|_2^2\}, \tag{5.12}$$

where E{.} represents Expectation, and $\|.\|_F$ denotes the matrix 2-norm.

$$\begin{aligned} d &= \mathrm{E}\left\{\|Hx - H(A^HA + \mu L^HL)^{-1}A^HAx\|_2^2\right\} + \mathrm{E}\left\{\|H(A^HA + \mu L^HL)^{-1}A^Hn\|_2^2\right\} \\ &= \mathrm{E}\left\{\|H(A^HA + \mu L^HL)^{-1}\mu L^HLx\|_2^2\right\} + \mathrm{E}\left\{\|H(A^HA + \mu L^HL)^{-1}A^Hn\|_2^2\right\}. \end{aligned} \tag{5.13}$$

Let $P = (A^HA + \mu L^HL)$, such that

$$d = \mathrm{E}\left\{\|HP^{-1}\mu L^HLx\|_2^2\right\} + \mathrm{E}\left\{\|HP^{-1}A^Hn\|_2^2\right\}. \tag{5.14}$$

Now,

$$\begin{aligned} E\{\|HP^{-1}A^Hn\|_2^2\} &= \mathrm{E}\left\{\mathrm{tr}\left(n^HAP^{-H}H^HHP^{-1}A^Hn\right)\right\} \\ &= \mathrm{tr}\left(AP^{-H}H^HHP^{-1}A^HR_n\right), \end{aligned} \tag{5.15}$$

where R_n is the correlation matrix for n and is described by the relation $R_n = \mathrm{E}\left\{nn^H\right\}$. We assume that the noise is independent, identically distributed such that $R_n = \mu I$. Also, P is symmetric and $P^H = P$. Thus, Eq. (5.15) reduces to

$$\mathrm{E}\left\{\|HP^{-1}A^Hn\|_2^2\right\} = \mu\,\mathrm{tr}\left(AP^{-1}H^HHP^{-1}A^H\right). \tag{5.16}$$

Also,

$$\begin{aligned} \mathrm{E}\left\{\|HP^{-1}\mu L^HLx\|_2^2\right\} &= \mathrm{E}\left\{\mathrm{tr}\left(x^H\mu L^HLP^{-1}H^HHP^{-1}\mu L^HLx\right)\right\} \\ &= \mu^2\,\mathrm{tr}\left(L^HLP^{-1}H^HHP^{-1}L^HLR_x\right), \end{aligned} \tag{5.17}$$

where R_x is the correlation matrix for x and is described by the relation $R_x = \mathrm{E}\left\{xx^H\right\}$. From Eqs. (5.16) and (5.17), we have

$$d = \mu \,\mathrm{tr}\left(P^{-1}H^H H P^{-1}\left(A^H A + \mu L^H L R_x L^H L\right)\right). \tag{5.18}$$

For $L = R_x^{-\frac{1}{2}}$, $L^H L = R_x^{-1}$, and Eq. (5.18) reduces to

$$d = \mu \,\mathrm{tr}\left(P^{-1}H^H H\right). \tag{5.19}$$

We define an error function as a weighted sum of the channel discrepancy functions as

$$e = \sum_i \kappa_i d_i = \kappa_i \sum \mu_i \,\mathrm{tr}\left((A_i^H A_i + \mu_i R_{x_i}^{-1})^{-1} H_i^H H_i\right), \tag{5.20}$$

where κ_i are scaling factors that reflect the perceptual importance of the fidelity in a particular channel.

5.3.3 Sampling Strategy

The goal is to sample only one color channel at each sample location. Thus, we have to select mn samples from a set of $3mn$ samples. The error criterion defined in (5.20) may be used to optimize the selection procedure. The criterion does not depend on the scene being imaged and may be used for sub-sampling a general scene if the statistical properties (R_x and R_n) of the fully sampled image are defined accurately.

Each row in the matrices A_i in (5.20) corresponds to a sample in the respective channel. The error criterion defined in (5.20) may be used to obtain the row that when eliminated would cause the least error in the reconstructed signal when viewed through the HVS.

An exhaustive optimization would require the computation of the error criterion for all combinations of eliminated rows, and would require $\frac{(3mn)!}{(2mn)!\,(mn)!}$ computations of the error criterion. For a reasonably sized array, this computation would require immense resources.

The authors in [75] use a greedy algorithm for sequential backward selection (SBS) of samples for signal reconstruction. The sequential backward selection algorithm can not be guaranteed to provide optimal results, but the authors in [76] have shown that the algorithm consistently provides good results with a relatively tight upper bound on the error criterion. We devise an SBS scheme for optimizing the criterion as follows. We start with a fully sampled image with all mn samples in each channel. The error criterion is computed after eliminating one row from one of the matrices A_i, and the row that gives the least value for the criterion is eliminated. In the next step, The matrix A_i from which the row is eliminated is of dimension $(m-1) \times n$. The error criterion is computed again after eliminating one row from A_i, and rows of A_i are successively eliminated with the constraint that the three channels are sampled in a mutually exclusive manner.

Computation of the error criterion requires the computation of the inverse of the matrix P for each eliminated row. For an $m \times n$ array, P is of dimension $mn \times mn$, and the inversion requires considerable computation even for small arrays. The error criterion may be simplified using the Sherman-Morrison matrix inversion formula such that we need find only an update term after each elimination. Also, the matrices H_i are circulant block-circulant and the matrix products involving H_i may be computed using DFTs. In spite of these simplifications, the computation of the criterion is cumbersome since in the form of (5.35), it requires the storage of at least the three $mn \times mn$ initial matrices P_i^{-1}.

5.3.4 Experiments

The power spectral density of a random process is given by the Wiener-Khinchine relation, $S_x(j\omega) = \mathcal{F}\{R_x\}$. We obtained an R_x representative of a general scene imaged by a digital camera from the mean, S_{avg}, of the power spectra of a large number of images reflecting various image types as $R_x = \mathcal{F}^{-1}\{S_{avg}\}$. The images used to obtain S_{avg} span a wide range of categories including natural scenes, landscapes, portraits, and a few color test images obtained from the USC-SIPI [77] image database.

The sample selection procedure detailed in Section 5.3.3 was applied for fully-sampled RGB arrays of different sizes. The error criterion values obtained for a Bayer array ($e_{_{Bayer}}$) and an array obtained by the SBS scheme ($e_{_{SBS}}$) detailed in Sec. 5.3.3 are shown in Table 5.1. The weights on the individual channel errors are $\kappa_{_{Red}} = 1$, $\kappa_{_{green}} = 1.6$, and $\kappa_{_{blue}} = 1$. The values of κ_i reflect the relative importance of the green channel on image quality and precise values may be obtained through psychovisual experiments. An 8×8 array obtained using SBS is shown in Fig. 5.5.

Table 5.1: Comparison of error criterion values with a Bayer array

Array size	$e_{_{Bayer}}$	$e_{_{SBS}}$
8×8	28.8083	27.5952
12×12	46.0583	44.3362
16×16	74.9760	72.3530
32×32	218.4921	211.1279

5.4 Sample selection based on Wiener filtering

In the following sections we describe a design method for an *RGB* type CFA based on the Wiener filtering of the sub-sampled CFA image. Since color differences in the *RGB*

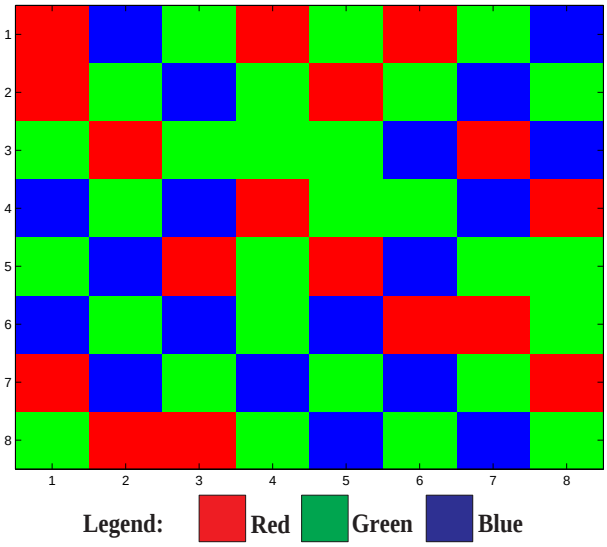


Figure 5.5: An 8×8 array

space do not correspond to perceptual differences, in this work, we consider a model of the HVS based on a uniform color space to quantify perceptual effects.

5.4.1 The $Y_y C_x C_z$ color space

Various models have been proposed in the literature that use perceptually uniform color spaces like the CIE $L^*a^*b^*$to describe the modulation transfer functions (MTFs) of the HVS. In this work, we use a model first described by Flohr et al. [78] to define the MTFs of the HVS luminance and chrominance channels. This model served as a basis for the HVS model used in Section 5.3.1. The Flohr model is channel-independent and is based on a color space that is a linearization of the CIE $L^*a^*b^*$color space. The transformation from CIE $L^*a^*b^*$to RGB is nonlinear and Flohr et al. propose a linearization about the D_{65}

white-point to form a color space characterized by the channels Y_y, C_x, and C_z as

$$\begin{aligned} Y_y &= 116\frac{Y}{Y_n} - 16, \\ C_x &= 500\left[\frac{X}{X_n} - \frac{Y}{Y_n}\right], \\ C_z &= 200\left[\frac{Y}{Y_n} - \frac{Z}{Z_n}\right]. \end{aligned} \tag{5.21}$$

The Y_y component in this color space corresponds to luminance and C_x and C_z are similar to $R-G$ and $B-Y$ opponent color chrominance components respectively.

In this work, we derive an MSE criterion in the $Y_yC_xC_y$ space to obtain an RGB array, and we will need to transform the error to the RGB space. From Eq. (5.21), the transformation from $Y_yC_xC_y$ to XYZ may be obtained as

$$\begin{aligned} X &= \frac{C_xX_n}{500} + \frac{1}{116}(Y_y + 16), \\ Y &= \frac{Y_n}{116}(Y_y + 16), \\ Z &= \frac{Z_n}{116}(Y_y + 16) - \frac{Z_n}{200}C_z. \end{aligned} \tag{5.22}$$

The transformation from XYZ to RGB about the D_{65} white point is performed as

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.240479 & -1.537150 & -0.498535 \\ -0.969256 & 1.875992 & 0.041556 \\ 0.055648 & -0.204043 & 1.057311 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \tag{5.23}$$

The transformation from $Y_yC_xC_z$ space to RGB space is achieved via the cascaded transformation $Y_yC_xC_z \rightarrow XYZ \rightarrow RGB$ as

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.240479 & -1.537150 & -0.498535 \\ -0.969256 & 1.875992 & 0.041556 \\ 0.055648 & -0.204043 & 1.057311 \end{bmatrix} \left(\begin{bmatrix} \frac{1}{16} & \frac{X_n}{500} & 0 \\ \frac{Y_n}{116} & \frac{16Y_n}{116} & 0 \\ \frac{Z_n}{116} & 0 & -\frac{Z_n}{200} \end{bmatrix} \begin{bmatrix} Y_y \\ C_x \\ C_z \end{bmatrix} + \begin{bmatrix} \frac{16}{116} \\ \frac{16Y_n}{116} \\ \frac{16Z_n}{116} \end{bmatrix} \right),$$

where the values X_n, Y_n, and Z_n for the D_{65} white point are 0.3127, 0.3290, and 0.3583 respectively such that

$$\begin{aligned} \begin{bmatrix} R \\ G \\ B \end{bmatrix} &= \begin{bmatrix} 0.0220356 & -0.067728 & 0.000893 \\ 0.0138047 & 0.085737 & -0.000074 \\ 0.0031668 & -0.009224 & -0.001894 \end{bmatrix} \begin{bmatrix} Y_y \\ C_x \\ C_z \end{bmatrix} + \begin{bmatrix} 0.352569 \\ 0.220875 \\ 0.050669 \end{bmatrix} \\ &= T_1 \begin{bmatrix} Y_y \\ C_x \\ C_z \end{bmatrix} + t \end{aligned} \tag{5.24}$$

5.4.2 The HVS MTFs

Flohr et al. propose a model that is a combination of the models detailed by Näsänen [79] and Sullivan et al. [74]. The Luminance MTF is modelled by an exponential that is similar to the MTF of the green channel in (5.1) as

$$V_{Y_y}(\bar{f}_{ij}) = K(L)e^{-\alpha(L)\bar{f}_{ij}}, \tag{5.25}$$

where $\bar{f}_{ij}$ is the radial spatial frequency in cycles/degree as subtended by the image on the human eye, and is a weighted magnitude of the linear frequency vector $[f_i \ f_j]^T$. L is the average luminance for the display, $K(L) = aL^b$,

$$\alpha(L) = \frac{a}{c \ln(L) + d}, \tag{5.26}$$

and $a = 131.6$, $b = 0.3188$, $c = 0.525$, $d = 3.91$.

An approximation to experimental results obtained by Mullen [80] is used to obtain the chrominance MTFs as

$$V_{C_x,C_z}(f_{ij}) = Ae^{(-\alpha f_{ij})}, \tag{5.27}$$

where $\alpha = 0.419$ and $A = 400$ as determined by Kolpatzik and Bouman [81]. As evident from Eqs. (5.26)-(5.27) the HVS model has a lowpass nature for both the luminance and the chrominance channels. The MTF of the chrominance channels decays at a greater rate and the luminance channel MTF has lesser sensitivity at odd multiples of $\pi/4$.

The HVS point spread functions (PSFs) h_i for $i = Y_y, C_x, C_z$ are obtained by taking the two-dimensional inverse Fourier transforms of $V_{Y_y}(\bar{f}_{ij})$ and $V_{C_x,C_z}(f_{ij})$ as follows:

$$\begin{aligned} h_{Y_y} &= \mathcal{F}^{-1}\left\{V_{Y_y}\right\}, \\ h_{C_x,C_z} &= \mathcal{F}^{-1}\left\{V_{C_x,C_z}\right\}. \end{aligned} \tag{5.28}$$

5.4.3 Sampling Strategy

Consider the image processing pipeline for a typical digital color camera depicted in Fig. 5.2. We propose a variation in the pipeline for the purpose of determining an error criterion (Fig. 5.6). During image acquisition, all three color channels are acquired at each sample location and full information about Y_y, C_x, and C_z channels is available. Intensity values obtained from RGB sensors may be transformed into the $Y_yC_xC_z$ space to obtain the required values. The image is then sub-sampled so that we are left with only one channel at a particular location and a demosaicking process is used to reconstruct the image. We propose a reconstruction method based on the Wiener filter for this stage of the pipeline. The HVS model detailed in Section 5.4.2 is used to characterize the perceptual error between the original and the reconstructed image. Since we need to determine sample locations for an RGB array, a color space transformation is applied to the output image obtained after convolution with the HVS PSF to convert the values to RGB space.

An error criterion is defined as the MSE between the reconstructed and the original image when passed through the HVS and after a color transformation into RGB space. We start with the fully sampled image with all three color channels available at each pixel location. The error criterion is then evaluated after eliminating all samples one at a time. The sample value that leads to the least increase in the error criterion is eliminated and the procedure is repeated with the remaining samples until only one channel is left at each pixel location. The resulting sampling arrangement assures the least perceptual degradation in the original fully-sampled image due to sparse sampling. The procedure neglects the effect of color space transforms and quantization associated with the enhancement processes in

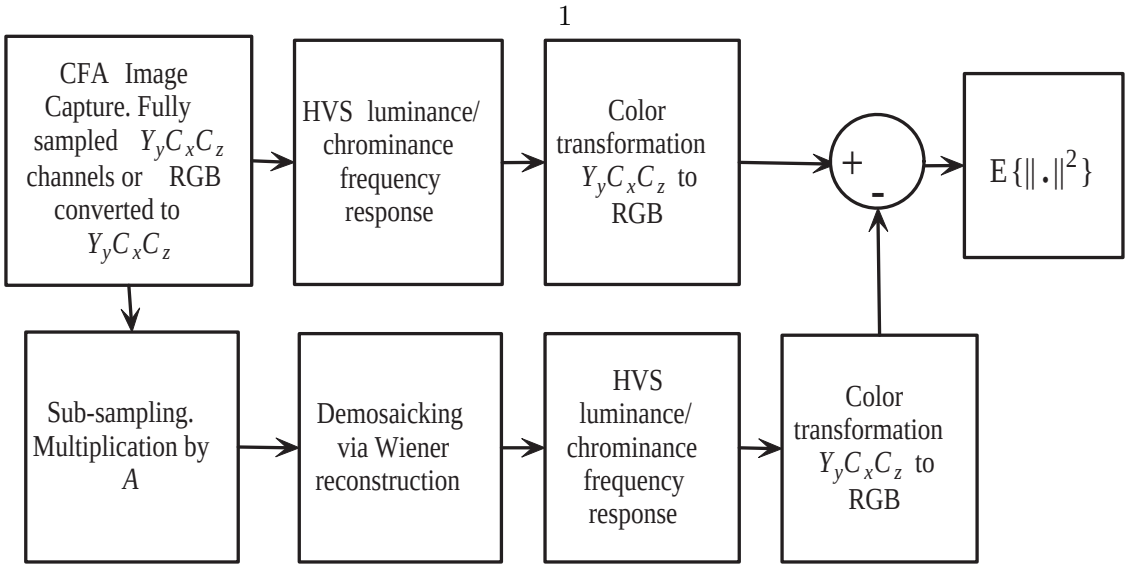


Figure 5.6: Block diagram for calculating the error criterion

Step 4 (Fig. 5.2), and the display device model in Step 5. In effect, we assume that color channel values obtained during acquisition are translated with reasonable fidelity to Step 4. Fig. 5.6 depicts the calculation of the error criterion in the form of a block diagram.

5.4.4 Mathematical Model

We assume that the effect of noise in the sub-sampling process may be neglected due to its much lower magnitude when compared to pixel intensities. For an original image I containing $m \times n$ pixels, the sub-sampled image is modelled as

$$y = Ax, \tag{5.29}$$

where $x \in \mathbb{C}^{(3mn \times 1)}$ is the fully sampled image and consists of the luminance and opponent chrominance channels (viz. the Y_y, C_x, and C_z values) in column-ordered form and takes

the form $x = [x_{Y_y}^T \quad x_{C_x}^T \quad x_{C_z}^T]^T$. Thus, the k^{th}, $2k^{th}$, and $3k^{th}$ elements of x ($k < mn$) represent the three channel values for the same pixel location. The vector $y \in \mathbb{C}^{(mn \times 1)}$ is the similarly arranged sub-sampled image, and contains only one channel at a particular pixel location. The matrix $A \in \mathbb{C}^{(mn \times 3mn)}$ is a sampling matrix that represents a linear transformation that maps the fully-sampled image to an image that is sub-sampled such that only one color channel is sampled at a particular location.

The Wiener filter solution for the estimate $\hat{x}$ of x in Eq. (5.29) is found as

$$\hat{x} = R_{xy} R_y^{-1} y, \tag{5.30}$$

where $R_{xy} = \mathrm{E}\left\{xy^T\right\}$ and $R_y = \mathrm{E}\left\{yy^T\right\}$; $\mathrm{E}\left\{.\right\}$ represents expectation. Substituting explicit expressions for R_{xy} and R_y gives

$$\begin{aligned}
\hat{x} &= \mathrm{E}\left\{xy^T\right\} \left(\mathrm{E}\left\{yy^T\right\}\right)^{-1} y \\
&= \mathrm{E}\left\{x(Ax)^T\right\} \left(\mathrm{E}\left\{Ax(Ax)^T\right\}\right)^{-1} Ax \\
&= \mathrm{E}\left\{xx^T A^T\right\} \left(\mathrm{E}\left\{Axx^T A^T\right\}\right)^{-1} Ax \\
&= R_x A^T \left(A R_x A^T\right)^{-1} Ax,
\end{aligned} \tag{5.31}$$

$$\text{where } R_x = \mathrm{E}\left\{xx^T\right\} = \mathrm{E}\left\{\begin{bmatrix} x_{Y_y}^T \\ x_{C_x}^T \\ x_{C_y}^T \end{bmatrix}\begin{bmatrix} x_{Y_y} & x_{C_x} & x_{C_z} \end{bmatrix}\right\} = \mathrm{E}\left\{\begin{bmatrix} x_{Y_y}^T x_{Y_y} & x_{Y_y}^T x_{C_x} & x_{Y_y}^T x_{C_z} \\ x_{C_x}^T x_{Y_y} & x_{C_x}^T x_{C_x} & x_{C_x}^T x_{C_z} \\ x_{C_z}^T x_{Y_y} & x_{C_z}^T x_{C_x} & x_{C_z}^T x_{C_z} \end{bmatrix}\right\}$$

$$= \begin{bmatrix} R_{x_{Y_y}} & R_{x_{Y_y}x_{C_x}} & R_{x_{Y_y}x_{C_z}} \\ R_{x_{Y_y}x_{C_x}} & R_{x_{C_x}} & R_{x_{C_x}x_{C_z}} \\ R_{x_{Y_y}x_{C_z}} & R_{x_{C_x}x_{C_z}} & R_{x_{C_z}} \end{bmatrix} \tag{5.32}$$

The elements on the diagonals of R_x are the autocorrelation matrices for the three channels and the off-diagonal elements are the channel crosscorrelation matrices. An error functional is formed as the mean square error of the original image and the reconstructed image when viewed through the HVS and converted to RGB space as

$$e = \mathrm{E}\left\{\|THx - TH\hat{x}\|_2^2\right\}, \tag{5.33}$$

where $\|.\|_2$ denotes the Frobenius matrix norm. The matrix H is constructed such that multiplication of a column-ordered image by H yields the 2-D convolution of the image by the PSFs h_i obtained in Eq. (5.28). The three channels of the HVS model are assumed to be independent such that H is block diagonal and of the form

$$H = \begin{bmatrix} H_{Y_y} & 0 & 0 \\ 0 & H_{C_x} & 0 \\ 0 & 0 & H_{C_z} \end{bmatrix}, \tag{5.34}$$

where the matrices H_i represent convolution of the individual channels by their respective PSFs and have a circulant block circulant structure. The matrix T is obtained from Eq. (5.24) such that multiplication of a column ordered image by T achieves the color transformation from $Y_yC_xC_z$ space to RGB space. T may be represented as a Kronecker matrix product of the form $T = T_1 \otimes I_{mn}$, where I_{mn} is the $mn \times mn$ identity matrix.

The error criterion is thus

$$\begin{aligned} e &= \mathrm{E}\left\{ \left\| THx - THR_xA^T\left(AR_xA^H\right)^{-1}Ax \right\|_2^2 \right\} x \\ &= \mathrm{E}\left\{ \left\| TH\left(I - R_xA^T\left(AR_xA^T\right)^{-1}A\right)x \right\|_2^2 \right\} \\ &= \mathrm{E}\left\{ \mathrm{tr}\left(x^T\left(I - R_xA^T\left(AR_xA^T\right)^{-1}A\right)^T H^TT^TTH\left(I - R_xA^T\left(AR_xA^T\right)^{-1}A\right)x \right) \right\}, \end{aligned}$$

where tr(.) represents the trace of a matrix. Let $P = \left(I - R_xA^T\left(AR_xA^T\right)^{-1}A\right)$, such that

$$e = \mathrm{E}\left\{ \mathrm{tr}\left(x^HP^HH^HT^HTHPx \right) \right\} = \mathrm{tr}\left(P^HH^HT^HTHPR_x \right). \tag{5.35}$$

Note that the criterion described by Eq. (5.35) does not depend on a particular scene being imaged. We only need to know the statistical properties of the scene as described by the elements of R_x to evaluate the criterion.

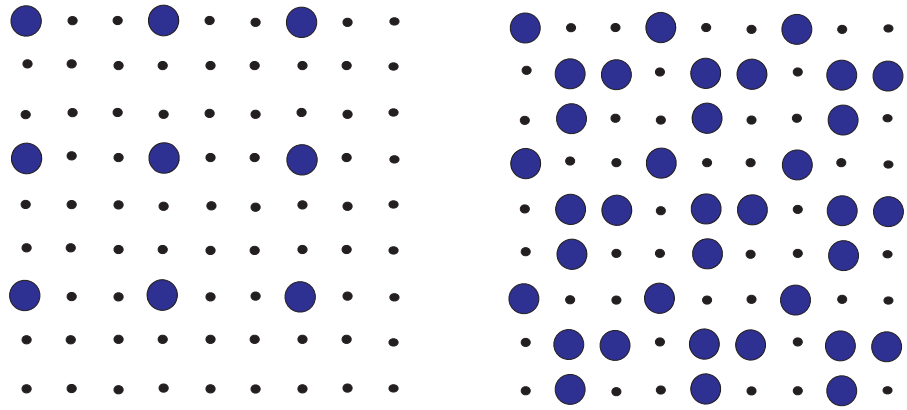
5.4.5 Sampling Procedure

Two different sampling procedures are detailed in this section. In the first case, we start with a fully-sampled image x with information about all three color channels. The goal is to eliminate samples such that we are left with only one color channel at each pixel location. As described in Section 5.4.3, we begin by eliminating the samples one at a time.

The error criterion is evaluated after each elimination and the sample that leads to the least increase in the error criterion is eliminated. Initially, the matrix A is of size $3mn \times 3mn$ and each row of A corresponds to a sample of the original image. Eliminating a sample from the original image is equivalent to eliminating a row from A. The error criterion defined in Eq. (5.35) may be used to obtain the row that when eliminated would cause the least error. Since the optimization requires immense computational resources, we once again use the SBS technique (Section 5.3.3) to elimintae samples one at a time.

In the second case, we once again start with the fully-sampled image but instead of eliminating a single sample, we eliminate a sub-array of samples from the original image. Figure 5.7(a) represents one channel of the image. The light dots represent pixel locations and the heavy dots represent a sub-array of samples. At each iteration, a shifted version of this sub-array is eliminated. This leads to a periodic replication of a non-periodic sampling pattern (Fig. 5.7(b)). The arrangement depicted in Figs. 5.7(a) and 5.7(b) leads to a 4×4 block periodic pattern. Such a block sampling pattern offers advantages in terms of computational simplicity and ease in the design of demosaicking algorithms.

In both cases, computation of the error criterion may be simplified using the Sherman-Morrison matrix inversion formula [82]. Instead of computing the inverse terms at each iteration we can find only an update term after each elimination. Also, the block circulant structure of H may be exploited for performing matrix multiplication via DFTs. In spite of these simplifications, the algorithm places a great demand on computational and storage resources.



(a) Sampling sub-array (b) Periodic pattern example

Figure 5.7: Rod and cone sensitivities

5.4.6 Experiments

We considered a 12×12 array. A variety of images that span a wide range of categories including natural scenes, landscapes, portraits, and a few color test images were obtained from the USC-SIPI [77] image database. The RGB channel values were converted to the $Y_y C_x C_z$ color space. Mean power spectra S_{m_i} for the individual channels and the mean crossspectra $S_{m_{ij}}$ were found from the power spectra of the available images. Using the Wiener-Khinchine relation for the power spectral density of a random process $S_x(j\omega) = \mathcal{F}\{R_x\}$, we obtained the elements of an R_x representative of a general scene imaged by a digital camera from the mean spectra S_{m_i} and $S_{m_{ij}}$ as $R_i = \mathcal{F}^{-1}\{S_{m_i}\}$, and $R_{ij} = \mathcal{F}^{-1}\{S_{m_{ij}}\}$.

The sample selection procedures detailed in Section 5.4.5 were applied for a fully-sampled 12×12 RGB array. Figure 5.8 shows the array obtained using the first method where the samples are eliminated one at a time. Figures 5.9(a), 5.9(b), and 5.9(c) show the array patterns obtained using the second method with 6×6, 4×4, and 3×3 blocks respectively. Figure 5.9(d) shows the array obtained with a 2×2 repeating block. This array is identical to the Bayer array. The error criterion values obtained for these cases are shown in Table 5.2.

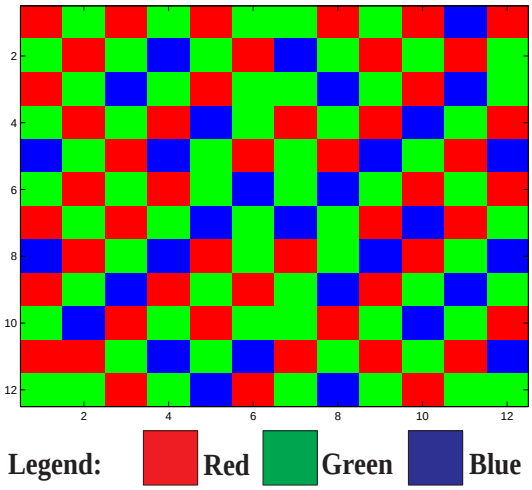
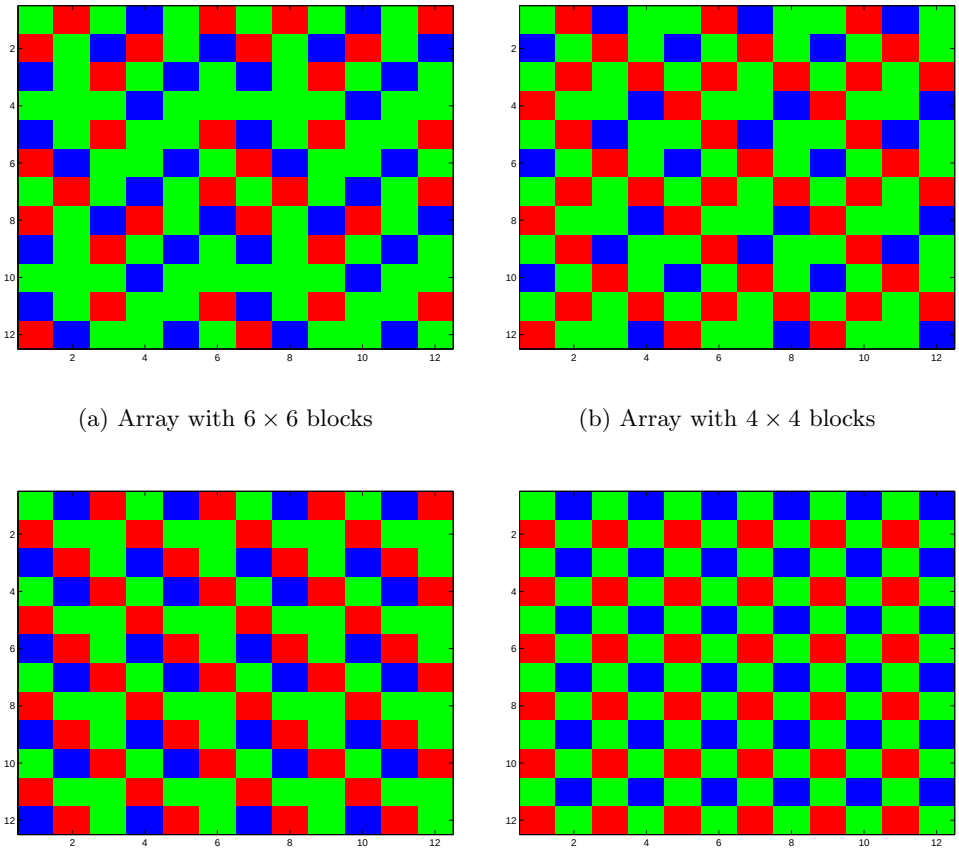


Figure 5.8: Array obtained by eliminating samples one at a time

Table 5.2: Comparison of error criterion values for a 12x12 array

Block size	d
12×12	610.0892
6×6	656.1477
4×4	673.0023
3×3	684.8360
2×2	692.3486



(a) Array with 6×6 blocks

(b) Array with 4×4 blocks

(c) Array with 3×3 blocks

(d) Array with 2×2 blocks

Figure 5.9: Block based array patterns

5.5 Conclusions and discussion

In Sections 5.3 and 5.4 we proposed two design methodologies for selection of color samples in CFAs. Both methods minimize error criteria obtained after reconstructing sub-sampled images. The first method uses regularization for restoration and defines an error criterion in the RGB space while the second method uses Wiener filtering for restoration and

defines an error criterion in the perceptually uniform $Y_yC_xC_z$ space. The SBS algorithm is used to sequentially eliminate samples until we arrive at an optimal sampling arrangement.

The results of experiments are listed in Tables 5.1 and 5.2. Both algorithms give error criterion values that are smaller than that obtained for the Bayer array. For the second algorithm, the error is least when samples are eliminated one at a time rather than in blocks. The error increases progressively as the block size is reduced and is maximum for the 2×2 case (which is identical to the Bayer array). The error criterion has a value smaller than the error criterion value for the Bayer array for all other cases.

The second algorithm is more interesting since:

1. It defines the error criterion in a perceptually uniform space where the magnitude of the error corresponds to the error perceived by a human observer.

2. It provides an ability to select block-based sampling patterns. This is useful for a number of reasons, primarily, since it results in symmetric array patterns, it is simpler to design adaptive demosaicking algorithms for the resulting arrays. Also, block-based patterns lend themselves to simplifiction in computation as the criterion in this case may be reduced to a structured form (circulant or Toeplitz). Finally, it is important that a particular color sample be surrounded by an identical set of color samples everywhere in the array. This is due to the phenomenon of spectral bleeding that occurs in closely spaced photosensitive elements in the sensor-array. A particular element in the array that is covered by a color filter will also generate some current due to the spill-over from neigboring elements. This contaminates the expected spectral response of the element in question. A consistent arrangement allows the image processor to account for the spectral bleeding.

5.6 Future work

The algorithms proposed in this work are extremely computationally intesive. We have shown that the resulting sampling arrangements perform better than the most commonly used array pattern (the Bayer array), but to validate the efficacy of the resulting sampling patterns, we need to design larger arrays. At this time, due to memory contraints, we can only design array patterns for images of size upto 12×12. The second method has a block structure and we are exploring ways to simplify computations to enable the design of larger arrays.

Conventionally, images are stored and displayed such that individual pixels are rectangular in shape. In this work we have considered rectangular sensor elements in CFAs. It has been shown that hexagonal arrangements have many advantages [83], [84], [85]. In particular, a hexagonal sampling grid allows two-dimensional sub-sampling at sub-nyquist frequencies. Also, in hexagonal arrays, the distance between a particular element and it's immediate neighbors is the same and this property can be used effectively in demosaicking algorithms. The selection of sampling patterns for hexagonally sampled arrays is an interesting problem to be considered in the future.

Chapter 6

Summary

6.1 Summary of results

The acquisition of multispectral images in the mosaicked form presents many advantages in terms of cost, simplicity of design, and the elimination of the registration step required in multi-sensor cameras. At the same time, mosaicked imaging presents many new challenges. The mosaicked image must be reconstructed to form full-color images, and a suitable algorithm must be designed for the purpose. The sampling arrangement and the sampling rate for the color samples must be chosen, and spectral sensitivity functions must be chosen for the colors used in the mosaic. In this work we have developed methods that address each of the above issues.

In Chapter 3 we proposed a general framework for the recovery of color images from sparse data [55]. An algorithm based on the Bayesian paradigm that may be used for simultaneous deblurring, denoising, and demosaicking of CFA data [86] was developed. The proposed algorithm relies on a hierarchical Bayesian formulation for the image model that accounts for the high correlation among color channels of a typical image. The ICM algorithm was then used to locally arrive at optimal pixel values given their neighboring elements. The proposed algorithm does not assume any particular CFA sampling arrangement and can be used for demosaicking of arbitrary CFA arrangements.

A novel joint spatial-chromatic sampling framework for the optimization of CFA based imaging parameters was proposed in Chapter 4 [68]. We addressed the problem of optimization of spectral sensitivity functions for the color filters in the sensor-array. An objective

criterion was introduced incorporates the effects of both spatial and spectral sampling in one unified framework. is introduced. Experimental results indicate that the optimized transmittance functions found by minimizing the objective criterion greatly outperform standard RGB and CMY color filters. Optimized color filter transmittances lead not only to reduced chromatic errors, but they also lead to fewer spatial artifacts in the reconstructed images [87]. Optimized transmittances were found for various common CFA arrangements and shown to outperform standard color filters in each case [88].

Two design methods for the selection of CFA sampling patterns were proposed in Chapter 5 [51, 52]. Both methods incorporate the effects of the human visual system in determining reconstruction quality of CFA sampled images. The quality of reconstructed images is used to derive objective criteria which may be minimized with respect to CFA sampling arrangements to derive optimal arrangements. The second method provides an ability to select block-based sampling patterns which leads to ease in the design of demosaicking algorithms and color filters with consistent effective transmittances across sensor-arrays.

6.2 Future work

There are several unresolved issues in the problem of multispectral imaging using focal-plane arrays. In light of the methods proposed in this work, future work is called for in the following areas:

1. In Chapters 4 and 5, objective criteria are derived to describe the distance between original images and images reconstructed from sub-sampled CFA data. The efficacy of the criterion hinges on the ability of the multi-dimensional autocorrelation matrix R_{xx} to describe faithfully the properties of a natural scene. In this work we based our

correlation model on the key assumption that both spatial and spectral correlations decay with distance in space and wavelength respectively. Spatial correlation does indeed fall with distance in the general scene, but the nature of the relation between elements of the autocorrelation function along the wavelength dimension is not easily modeled. Research in this area will help refine the results obtained in this work.

2. Recently, researchers have started to explore the problem of CFA-based imaging for multiple number of color bands (>4) [89, 90, 91, 92]. There is great potential of realizing the benefits of multispectral imaging with CFAs because of the steady increase in sensor-array sizes. In Chapter 5 we have demonstrated that full-color images of reasonable quality may be reconstructed from CFAs with sparse spatial sampling of a particular color. For instance, blue is sampled at a much lower rate than green in the 5×5 optimal block-based array, without a great loss in the quality of reconstructed images. This suggests that the sparse sampling of particular colors due to an increase in the number of color bands is a reasonable trade-off and should be investigated in the context of the joint spatial-chromatic sampling framework developed in this work.

3. The effect of noise has not been considered in the development of the spatial-chromatic reconstruction method proposed in this work. Effective noise models will greatly increase the usefulness of methods proposed here.

Bibliography

[1] S. Abdallah, B. Saleh, and A. Aboulsoud, "A general overview of solid state imaging sensors types," in *Photonics and Its Application at Egyptian Engineering Faculties and Institutes, Third workshop on*, 2002, pp. 1–10.

[2] J. Adams, K. Parulski, and K. Spaulding, "Color processing in digital cameras," *IEEE Micro*, vol. 18, no. 4, pp. 20–30, Nov/Dec 1998.

[3] B. Gunturk, Y. Altunbasak, and R. Mersereau, "Color plane interpolation using alternating projections," *Image Processing, IEEE Transactions on*, vol. 11, no. 9, pp. 997–1013, Sept. 2002.

[4] J. Driesen and P. Scheunders, "Wavelet-based color filter array demosaicking," in *International Conference on Image Processing, Proceedings of*, 2004, pp. V: 3311–3314.

[5] X. Wu and N. Zhang, "Primary-consistent soft-decision color demosaicking for digital cameras (patent pending)," *Image Processing, IEEE Transactions on,*, vol. 13, no. 9, pp. 1263–1274, Sept. 2004.

[6] J. Go, K. Sohn, and C. Lee, "Interpolation using neural networks for digital still cameras," *Consumer Electronics, IEEE Transactions on*, vol. 46, no. 3, pp. 610–616, Aug 2000.

[7] H. Trussell, "A MMSE estimate for demosaicking," *Proceedings of the International Conference on Image Processing*, vol. 3, pp. 358–361, 2001.

[8] H. Trussell and R. Hartwig, "Mathematics for demosaicking," *IEEE Trans. Image Processing*, vol. 11, no. 4, pp. 485–492, April 2002.

[9] D. Taubman, "Generalized wiener reconstruction of images from colour sensor data using a scale invariant prior," *International Conference on Image Processing, Proceedings of*, vol. 3, pp. 801–804, 2000.

[10] P. Vora, J. Farrell, J. Tietz, and D. Brainard, "Image capture: simulation of sensor responses from hyperspectral images," *IEEE Trans. Image Processing*, vol. 10, no. 2, pp. 307–316, Feb 2001.

[11] P. Vora and H. Trussell, "Mathematical methods for the analysis of color scanning filters," *IEEE Trans. Image Processing*, vol. 6, no. 2, pp. 321–327, Feb 1997.

[12] ——, "Mathematical methods for the design of color scanning filters," *IEEE Trans. Image Processing*, vol. 6, no. 2, pp. 312–320, Feb 1997.

[13] M. J. Vrhel and H. J. Trussell, "Optimal color filters in the presence of noise," *IEEE Trans. Image Processing*, vol. 4, no. 6, pp. 814–823, June 1995.

[14] B. Bayer, "Color imaging array," U.S. Patent 3971065, July 1976.

[15] T. Yamagami, T. Sasaki, and A. Suga, "Image signal processing apparatus having a color filter with offset luminance filter elements," U.S. Patent 5323233, June 1994.

[16] J. Hamilton, J. Adams, and D. Orlicki, "Particular pattern of pixels for a color filter array which is used to derive luminanance and chrominance values," U.S. Patent 6330029 B1, Dec. 2001.

[17] W. Zhu, K. Parker, and M. A. Kriss, "Color filter arrays based on mutually exclusive blue noise patterns," *Journal of Visual Communication and Image Representation*, vol. 10, pp. 245–247, 1999.

[18] E. Gindele and A. Gallagher, "Sparsely sampled image sensing device with color and luminance photosites," U.S. Patent 6476865 B1, Nov. 2002.

[19] G. Sharma and H. J. Trussell, "Digital color imaging," *IEEE Transactions on Image Processing*, vol. 6, no. 7, pp. 901–932, 1997.

[20] T.Young, "An account of some cases of the production of colors, not hitherto described," *Philos. Trans. R. Soc. London*, p. 387, 1802.

[21] ——, "On the theory of light and colors," *Philos. Trans. R. Soc. London*, vol. 92, pp. 20–71, 1802.

[22] G. Palmer, *Theory of Light*. Paris, France: Hardouin and Gattey, 1786.

[23] H. Grassmann, "Zur theorie der farbenmischung," *Poggendorf', Annalen Phys. Chemie,*, vol. 89, pp. 69–84, 1853.

[24] J. Maxwell, "Theory of the perception of colors," *Trans. R. Scottish Soc. Arts*, vol. 4, pp. 394–400, 1856.

[25] G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd Edition*. New York: Wiley, 1982.

[26] J. Maxwell, "The diagram of colors," *Trans. R. Soc. Edinburgh*, vol. 21, pp. 275–298, 1857.

[27] ——, "Theory of compound colors and the relations to the colors of the spectrum," *Proc. R. Soc. Lond.*, vol. 10, pp. 404–409, 1860.

[28] H. L. F. von Helmholtz, "Theory of compound colors and the relations to the colors of the spectrum," *Phys. Opt.*, 1866.

[29] H. J. A. Dartnall, J. K. Bowmaker, and J. D. Mollon, *Microspectrophotometry of human photoreceptors*. Academic, 1983.

[30] A. Stockman, D. I. A. MacLeod, and N. E. Johnson, "Spectral sensitivities of the human cones," *Journal of the Optical Society of America*, vol. 10, pp. 2491–2521, 1993.

[31] R. S. Gentile, *Device independent color in PostScript*, ser. Proc. SPIE: Human Vision, Visual Processing, and Digital Display IV. SPIE, 1993, vol. 1913, pp. 419–432.

[32] G. Wyszecki and G. H. Fielder, "Color difference matches," *Journal of the Optical Society of America*, vol. 62, pp. 1501–1513, 1971.

[33] A. D. North and M. D. Fairchild, "Measuring color-matching functions. Part I," *Color Res. Appl.*, vol. 18, no. 3, pp. 155–162, June 1993.

[34] Y. Nayatani, K. Takahama, H. Sobagaki, and K. Hashimoto, "Color appearance model and chromatic-adaptation transform," *Color Res. Appl.*, vol. 15, pp. 210–221, Feb 1990.

[35] S. L. Guth, "Model for color vision and light adaptation," *J. Opt. Soc. Amer. A*, vol. 8, pp. 976–993, Jun 1991.

[36] Y. Nayatani, K. Hashimoto, H. Sobagaki, and K. Takahama, "Comparison of color-appearance models," *Color Res. Appl*, vol. 15, pp. 272–284, Oct 1990.

[37] X. Zhang and B. Wandell, "A spatial extension of CIELab for digital color image reproduction," in *Proc. Soc. Inform. Display 96 Digest*, 1996, pp. 731–734.

[38] X. Zhang, D. Silverstein, J. Farrell, and B. Wandell, "Color image quality metric S-CIELAB and its application on halftone texture visibility," in *COMPCON97 Digest of Papers, IEEE*, 1997, pp. 44–48.

[39] D. R. Cok, "Reconstruction of CCD images using template matching," *Proceedings of IS&T's Annual Conference/ICPS*, pp. 380–385, 1994.

[40] R. Kimmel, "Demosaicing: image reconstruction from color CCD samples," *Image Processing, IEEE Transactions on*, vol. 8, no. 9, pp. 1221–1228, Sept. 1999.

[41] J. Glotzbach, R. Schafer, and K. Illgner, "A method of color filter array interpolation with alias cancellation properties," *International Conference on Image Processing, Proceedings of*, vol. 1, pp. 141–144, 2001.

[42] D. Su and P. Willis, "Demosaicing of color images using pixel level data-dependent triangulation," in *Theory and Practice of Computer Graphics, 2003. Proceedings*, 3-5 June 2003, pp. 16–23.

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

[44] E. Dubois, "Frequency-domain methods for demosaicking of Bayer-sampled color images," *Signal Processing Letters*, vol. 12, no. 12, pp. 847–850, December 2005.

[45] L. Chang and Y.-P. Tan, "Effective use of spatial and spectral correlations for color filter array demosaicking," *Consumer Electronics, IEEE Transactions on*, vol. 50, no. 1, pp. 355–365, Feb 2004.

[46] K. Hirakawa and T. Parks, "Joint demosaicing and denoising," *Image Processing, IEEE Transactions on*, vol. 15, no. 8, pp. 2146–2157, Aug. 2006.

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

[48] S. Yamanaka, "Solid state camera," U.S. Patent 4054906, 1977.

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

[50] C. M. Hains, "Personal communication," Xerox Corp., July 2006.

[51] M. Parmar and S. J. Reeves, "Color filter array design based on a human visual model," *Proceedings of IS&T/SPIEs International Conference on Electronic Imaging, Computational Imaging II*, vol. 5299, no. 5299, pp. 73–82., 2004.

[52] ——, " A perceptually based design methodology for color filter arrays ," *Acoustics, Speech, and Signal Processing, 2004. Proceedings, IEEE International Conference on*, vol. 53, pp. 473–476., 2004.

[53] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–741, 1984.

[54] R. Molina, J. Mateos, A. Katsaggelos, and M. Vega, "Bayesian multichannel image restoration using compound gauss-markov random fields," *Image Processing, IEEE Transactions on*, vol. 12, no. 12, pp. 1642–1654, Dec. 2003.

[55] M. Parmar, S. J. Reeves, and T. S. Denney, Jr., "Bayesian edge-preserving color image reconstruction from color filter array data," in *Computational Imaging III*, C. A. Bouman and E. L. Miller, Eds., vol. 5674, no. 1. SPIE, 2005, pp. 259–268.

[56] G. Sharma, Ed., *Digital Color Imaging Handbook*. Boca Raton, FL, USA: CRC Press, Inc., 2002, ch. one: Color fundamentals for digital imaging.

[57] Eastman Kodak Company, "PhotoCD PCD0992," (http://r0k.us/graphics/kodak/).

[58] D. Griffeath, *Introduction to Markov Random Fields*. Springer, 1976, chapter 12 of *Denumerable Markov Chains* by Kemeny, Knapp, and Snell (2nd edition).

[59] T. S. Denney, Jr. and S. J. Reeves, "Bayesian image reconstruction from Fourier-domain samples using prior edge information," *Journal of Electronic Imaging*, vol. 14, no. 4, p. 043009, 2005.

[60] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society, series B*, vol. 48, pp. 259–302, 1986.

[61] R. Hunt, "Why is black-and-white so important in color," in *Color Imaging Conference*, 1995, pp. 54–57.

[62] M. J. Vrhel and H. J. Trussell, "Color filter selection for color correction in the presence of noise," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1993, pp. 313–316.

[63] ——, "Filter consideration in color correction," *IEEE Trans. Image Processing*, vol. 3, no. 2, pp. 147–161, 1994.

[64] G. Sharma, H. Trussell, and M. Vrhel, "Optimal nonnegative color scanning filters," *IEEE Transactions on Image Processing*, vol. 7, no. 1, pp. 129–133, Jan 1998.

[65] N. Shimano, "Optimization of spectral sensitivities with gaussian distribution functions for a color image acquisition device in the presence of noise," *Optical Engineering*, vol. 45, no. 1, p. 013201, 2006.

[66] M. Wolski, C. Bouman, J. Allebach, and E. Walowit, "Optimization of sensor response functions for colorimetry of reflective and emissive objects," *IEEE Trans. Image Processing*, vol. 5, no. 3, pp. 507–517, Mar 1996.

[67] D.-Y. Ng and J. P. Allebach, "A subspace matching color filter design methodology for a multispectral imaging system," *IEEE Transactions on Image Processing*, vol. 15, no. 9, pp. 2631–2643, Sep. 2006.

[68] M. Parmar and S. J. Reeves, "Selection of optimal spectral sensitivity functions for color filter arrays," in *International Conference on Image Processing, Proceedings of*, Oct. 2006, pp. 1005–1008.

[69] D. Alleysson, S. Süsstrunk, and J. Marguier, "Influence of spectral sensitivity functions on color demosaicing," *Proceedings of the Eleventh Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications*, pp. 351–357, November 2003.

[70] S. D. Hordley, G. D. Finlayson, and P. Morovic, "A multi-spectral image database and an application to image rendering across illumination," in *Proceedings of Third*

International Conference on Image and Graphics, Hong-Kong China,, December 2004, pp. 349–355.

[71] J. E. Farrell, F. Xiao, P. B. Catrysse, and B. A. Wandell, "A simulation tool for evaluating digital camera image quality," in *Proceedings of IS&T/SPIEs Electronic Imaging 2004: Image Quality and System Performance*, vol. 5294, Jan. 2004.

[72] I. 61966-2-2, *Multimedia systems and Equipment — Colour measurement and management – Part 2-1: Colour management- Default RGB colour space – sRGB.* International Electrotechnical Comission, 1999. [Online]. Available: http://www.srgb.com

[73] M. A. Kriss, "Color filter arrays for digital electronic still cameras," *Proceedings of IS&T's 49th Annual Conference*, pp. 272–278, 1996.

[74] J. Sullivan, L. Ray, and R. Miller, "Design of minimum visual modulation halftone patterns," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 21, no. 1, pp. 33–38, 1991.

[75] S. J. Reeves and L. P. Heck, "Selection of observations in signal reconstruction," *IEEE Transactions on Signal Processing*, vol. 43, no. 3, pp. 788–791, March 1995.

[76] S. J. Reeves and Z. Zhao, "Sequential algorithms for observation selection," *IEEE Transactions on Signal Processing*, vol. 47, no. 1, pp. 123–132, January 1999.

[77] USC-SIPI, "Color image database," available at: http://sipi.usc.edu/services/database/Database.html.

[78] T. J. Flohr, B. W. Kolpatzik, R. Balasubramanian, D. A. Carrara, C. A. Bouman, and J. P. Allebach, "Model Based Color Image Quantization," *Human Vision, Visual Processing, and Digital Display IV (1993)*, vol. SPIE 1913, pp. 270–281, 1993.

[79] R. Näsänen, "Visibility of halftone dot textures," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 14, no. 6, pp. 920–924, 1984.

[80] K. T. Mullen, "The contrast sensitivity of human color vision to red-green and blue-yellow chromatic gratings," *J. Physiol.*, vol. 359, pp. 381–400, 1985.

[81] B. Kolpatzik and C. Bouman, "Optimized error diffusion for image display," *Journal of Electronic Imaging*, vol. 3, no. 3, pp. 277–292, July 1992.

[82] G. H. Golub and C. F. Van Loan, *Matrix Computations.* Baltimore, MD: The Johns Hopkins University Press, 1996.

[83] R. M. Mersereau, "The processing of hexagonally sampled two-dimensional signals," *IEEE Proceedings*, vol. 67, pp. 930–949, June 1979.

[84] M. Golay, "Hexagonal parallel pattern transformations," *IEEE Trans Comp*, vol. 18, pp. 733–740, 1969.

[85] I. Her, "Geometric transformations on the hexagonal grid," *IEEE Trans. Image Processing*, vol. 4, no. 9, pp. 1213–1222, September 1995.

[86] M. Parmar, S. J. Reeves, and T. S. Denney, Jr., "Bayesian restoration of color images using a non-homogenous cross-channel prior,," in *submitted to The International Conference on Image Processing*, 2007.

[87] M. Parmar and S. J. Reeves, "Optimization of spectral sensitivity functions for color filter arrays," *in review*.

[88] ——, "Optimization of color filter sensitivity functions for color filter array based image acquisition," in *Proceedings of the Fourteenth Color Imaging Conference*, Nov. 2006, pp. 96–101.

[89] R. Ramanath, W. Snyder, H.Du, H. Qi, and X. Wang, "Band selection using independent component analysis for hyperspectral image processing," *Proceedings AIPR workshop*, pp. 93–98, 2003.

[90] R. Ramanath, W. E. Snyder, and H. Qi, "Mosaic multispectral focal plane array cameras," in *Infrared Technology and Applications XXX. Edited by Andresen, Bjorn F.; Fulop, Gabor F. Proceedings of the SPIE, Volume 5406, pp. 701-712 (2004).*, B. F. Andresen and G. F. Fulop, Eds., Aug. 2004, pp. 701–712.

[91] L. Miao and H. Qi, "The design and evaluation of a generic method for generating mosaicked multispectral filter arrays," *IEEE Trans. Image Processing*, vol. 15, no. 9, pp. 2780–2791, August 2006.

[92] L. Miao, H. Qi, R. Ramanath, and W. Snyder, "Binary tree-based generic demosaicking algorithm for multispectral filter arrays," *IEEE Trans. Image Processing*, vol. 15, no. 11, pp. 3550–3558, November 2006.