Title

Optimized Multi-Spectral Filter Array Based Imaging of Natural Scenes

Permalink

https://escholarship.org/uc/item/64m401dd

Journal

Sensors, 18(4)

ISSN

1424-8220

Authors

Li, Yuqi
Majumder, Aditi
Zhang, Hao
et al.

Publication Date

2018

DOI

10.3390/s18041172

Copyright Information

Peer reviewed

Optimized Multi-Spectral Filter Array Based Imaging of Natural Scenes

Yuqi Li [1,2,*], **Aditi Majumder** [3], **Hao Zhang** [3] **and M. Gopi** [3]

1   College of Information Science and Engineering, Ningbo University, Ningbo 315000, China
2   College of Computer Science and Technology, Zhejiang University, Hangzhou 310000, China
3   Department of Computer Science, University of California, Irvine, CA 92697-3425, USA;
    majumder@ics.uci.edu (A.M.); zhang.hao.alan@gmail.com (H.Z.); gopi@ics.uci.edu (M.G.)
*   Correspondence: liyuqi1@nbu.edu.cn

Abstract: Multi-spectral imaging using a camera with more than three channels is an efficient method to acquire and reconstruct spectral data and is used extensively in tasks like object recognition, relighted rendering, and color constancy. Recently developed methods are used to only guide content-dependent filter selection where the set of spectral reflectances to be recovered are known a priori. We present the first content-independent spectral imaging pipeline that allows optimal selection of multiple channels. We also present algorithms for optimal placement of the channels in the color filter array yielding an efficient demosaicing order resulting in accurate spectral recovery of natural reflectance functions. These reflectance functions have the property that their power spectrum statistically exhibits a power-law behavior. Using this property, we propose power-law based error descriptors that are minimized to optimize the imaging pipeline. We extensively verify our models and optimizations using large sets of commercially available wide-band filters to demonstrate the greater accuracy and efficiency of our multi-spectral imaging pipeline over existing methods.

Keywords: multi-spectral imaging; content independent channel selection; multispectral filter array; demosaicing

1. Introduction

Multi-channel spectral imaging using a single-camera is an efficient way to acquire spectral data. Compared with multi-camera systems that use multiple cameras with modified camera sensitivity, it is more compact and practical. Typically, the multiple different channels of the spectrum are captured (a) over multiple shots using a different color filter in front of the camera sensor for each shot; or (b) in a single shot using a mosaic of the multiple filters on the sensor forming a multiple spectral filter array (MSFA).

Traditionally, narrow band filters are used to provide a large number of channels, thereby providing higher spectral resolution assuring an accurate spectral recovery. However, this severely compromises the acquisition rate in multi-shot capture and the spatial resolution in single shot capture. Since spatial resolution is very important for most scenarios, commercially available multi-spectral cameras are multi-shot. Narrow band filters also significantly compromise light efficiency and therefore almost all multi-spectral cameras use longer exposures, thereby limiting its capture to only static scenes. Recently, use of wide-band filters has been explored to alleviate these problems by reducing the number of channels used [1]. However, the channels are usually chosen in an ad hoc manner, thereby not assuring an accurate spectral recovery. Recently, a work showed that, with channel selection, the spectral reconstruction accuracy could be improved by over 33% [2].

Previous works have shown that both man-made and natural objects or phenomena (like illumination) have smooth spectral power distribution. Statistically, their power spectrum follows a strong power law [3,4]. We exploit this statistical property to carefully select a small number of optimal wide band channels (usually 5–6) for multi-spectral imaging that can ensure an accurate spectral reconstruction when used in an MSFA for a single shot capture. The main contributions of this work are as follows.

- First, we design and develop a novel power-law prior based channel optimization method that models the various errors associated with spectral reconstruction—namely error due to estimation (reconstruction error), noise (imaging error) and demosaicing (demosaicing error). These errors depend only on the camera parameters (e.g., spectral sensitivities of channels, the MSFA pattern, the demosaicing order, and variance of the sensor noise) and not on the content. To the best of our knowledge, this is the first model for defining all the different errors in a content-independent multi-spectral imaging pipeline.
- Second, we construct an objective function that quantifies the total error using a combination of the three above-mentioned errors. Next, we use a discrete particle swarm optimization method to optimize the imaging pipeline by (1) selecting a few channels from a large set of candidate channels; (2) constructing a conducive mosaic pattern with the chosen channels on the MSFA; and (3) selecting a channel ordering during demosaicing that minimizes the objective function and hence the total error in spectral reconstruction.

2. Related Works

Multispectral Imaging and Reconstruction: a large number of the existing imaging techniques today are either bulky, or expensive and require professional calibration [5,6]. Our focus in this paper are the alternate, more compact spectral imaging techniques that are popularly used for commodity applications. There are two existing spectral imaging techniques that fit this bill—(a) compressive spectral imaging and (b) multi-spectral filter array (MSFA) based spectral imaging.

Compressive computational imaging systems randomly encode captured spectra [7], and then reconstruct multi-spectral images using compressive sensing reconstruction techniques [8–10]. Taking full advantages of the sparseness of spectral image data, these reconstruction techniques can recover multi-spectral images using fewer observations than conventionally needed. Furthermore, due to the high light throughput ($\approx$0.5 [9]) of these systems, the quality of reconstructed images is robust to imaging noise. However, the sparse recovery process is time-consuming. For example, it can take minutes or even hours to reconstruct a single $512 \times 512 \times 31$ multi-spectral frame. Therefore, it is not suitable for real-time applications, such as rendering relighted scenes, or object detection.

Contrary to compressive computational imaging, MSFA offers a very time efficient option for spectral reconstruction. Recently, it has been shown that, when using wide-band filters, MSFA based methods can be more accurate in the common case where the "photon to read noise ratios" (PRR) is not very low [11]. Therefore, various multi-channel spectral imaging system with MSFA have been designed recently. For example, Yasuma et al. [12] proposed a sub-micron pixel image sensor design with seven colors and two exposures; Sajadi et al. [13] combined a red-green-blue Bayer color filter array (CFA) with a cyan-magenta-yellow CFA in a layered CFA design to capture images with high color fidelity; Monno et al. [14] presented a high-performance multi-spectral demosaicing algorithm and utilized a single sensor with 5-channel MSFA to capture multi-spectral images. However, none of these single-shot methods focus on optimizing image pipeline parameters (like the channels selected, spatial arrangement of channels on the MSFA, demosaicing order) in a unified manner and hence cannot provide the desired spatial and spectral precision with accuracy of reconstruction.

Channel Selection: references [15,16] optimize spectral sensitivity of channels with a regular three-channel RGB Bayer color filter array (CFA) , while references [17,18] optimize channels to minimize the error in spectral recovery, but only for spectral data that have strong priors like known distribution of the captured spectral functions (e.g., daylight [19]). Such domain-specific priors yield *content-dependent*

methods. More importantly, theoretically conducive spectral sensitivity of the channels are assumed (e.g., radial basis function [20,21], Fourier basis function [22]). This yields poor results in practice due to a large deviation of the sensitivities of commercially available filters from such well-behaved functions [23] (Figure 1).
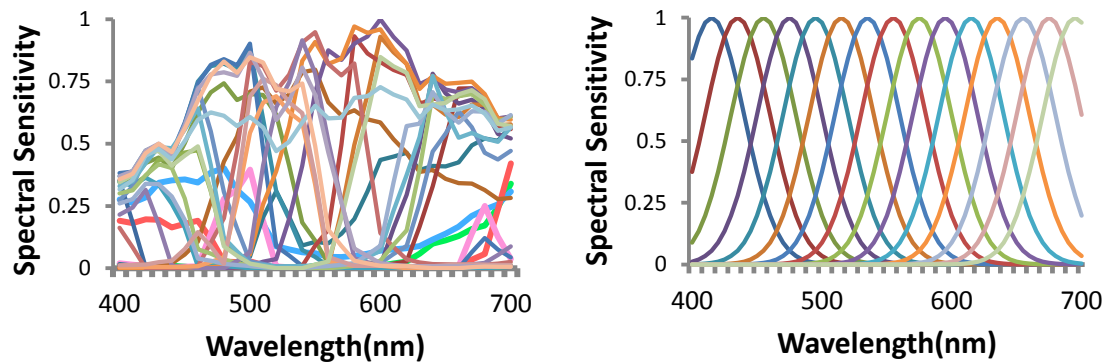

Figure 1. Left: commercial filter set; right: radial basis function.

In contrast, we select filters from a set of commercially available wide-band filters to create a practically feasible spectral imaging pipeline. A handful of earlier works take this route. Yasuma et al. provided a cost function to describe spectral recovery error that is optimized to select filters for the generalized assorted pixel (GAP) camera [12]. However, the design and spatial arrangement of the GAP camera lacks freedom, while impacting the quality of spectral recovery significantly. Chi et al. [1] heuristically minimize the condition number of spectral sensitivity of channels to make the spectral recovery specifically robust to noise. Other work [24] optimizing spectral reflectance targets can also be used to select channels. All these methods cannot guarantee spectral error minimization and therefore optimal spectral reconstruction. Recently, Arad et al. [2] used different three-channel combinations to reconstruct multispectral images and choose the best combination. However, it can not be extended to multi-channel selection directly since it is too time-consuming and may depend on specific spectral reconstruction method.

Demosaicing: demosaicing methods have been devised independent of the other elements of a spectral imaging pipeline. Interestingly, it provides a diverse view of a conducive filter array and demosaicing methods. While some prior works [25] emphasize that the spectral sensitivities of neighboring pixels should be highly correlated to minimize inter-channel demosaicing error, other works (e.g., [26]) demonstrate the benefits of completely independent demosaicing fostered by low correlation between narrow band channels. Many works [21,27] treat the channel with highest sampling rate and largest throughput as the dominant channel and complete missing values of other channels that are considered dependent on the guide channel. In addition, a recent work uses a channel-dependent demosaicing strategy for a non-channel-dominant narrow-band MSFA [28].

Comparison with Proposed Work: the aforementioned literature survey reveals that mutually correlated problems of channel selection, filter array design, demosaicing methods and imaging noise have always been addressed in exclusion and therefore failed to yield solutions that will be optimal in terms of accuracy and efficiency for the final output of the entire spectral imaging pipeline. Therefore, we attempt to address all these correlated issues in a unified manner and optimize them together.

3. Modeling Error in Spectral Recovery

Let the spectral power distribution (SPD) function of the spectrum at spatial coordinate (x, y) be $s(x, y, \lambda)$. Consider a multi-spectral camera with N channels ($N > 3$). Let channel i ($1 \leq i \leq N$) have the spectral sensitivity function $m_i(\lambda)$ that can be obtained from specification sheets or measured

using a broadband light source using some low-cost methods [29]. Let the response of channel i of the camera to $s(x, y, \lambda)$ be x_i.

In a practical system, x_i is comprised of ground truth, $\tilde{x}_i$, and errors, ϵ_i. If the channel i at (x, y) is interpolated during demosaicing, then $\epsilon_i = \zeta_i + \delta_i$ where ζ_i denotes error due to imaging noise and δ_i denotes demosaicing error. For pixels where channel i is captured directly, $\delta_i = 0$. Therefore,

$$
\begin{aligned}
x_i = \tilde{x}_i + \epsilon_i &= \int_{\lambda_1}^{\lambda_2} s(\lambda) m_i(\lambda) d\lambda + \epsilon_i \\
&= \int_{\lambda_1}^{\lambda_2} s(\lambda) m_i(\lambda) d\lambda + \zeta_i + \delta_i.
\end{aligned}
\tag{1}
$$

We can write the above equation using matrices considering all channels as

$$
X = Ms + \epsilon = Ms + \zeta + \delta,
\tag{2}
$$

where $[\lambda_1, \lambda_2]$ is the range of wavelength of visible light (e.g., 380–780 nm), M is a known matrix of dimension $N \times K$ where each row of M denotes one of the channels at a spectral resolution of K, s is the $K \times 1$ SPD function to be recovered and X denotes the $N \times 1$ known response vector and ϵ, ζ and δ are all $N \times 1$ error vectors.

Let us assume that the estimated spectrum $\hat{s}$ is given by multiplying X with reconstruction matrix W. Using a pseudo-inverse matrix $W^+ = M^T (MM^T)^{-1}$ or a Wiener pseudo inversion $W^+ = Corr_s M^T (M Corr_s M^T)^{-1}$ [30] with an approximated autocorrelation matrix $Corr_s$ to represent the reconstruction matrix, the recovered spectrum $\hat{s}$, is given by

$$
\hat{s} = WX = W(Ms + \epsilon) = WMs + W\epsilon.
\tag{3}
$$

We use the mean squared error (MSE) between the original and the recovered spectrum to describe the average error in the estimation as

$$
MSE = \text{Tr} \left\{ \mathrm{E}[(s - \hat{s})(s - \hat{s})^T] \right\}.
\tag{4}
$$

Now, replacing Equation (3) in Equation (4), we get

$$
\begin{aligned}
MSE &= \text{Tr} \left\{ E[(s - WMs - W\epsilon)(s - WMs - W\epsilon)^T] \right\} \\
&= \text{Tr} \left\{ E[(Hs - W\epsilon)(Hs - W\epsilon)^T] \right\},
\end{aligned}
\tag{5}
$$

where $\mathrm{E}[.]$ is the expected values of a variable, $H = (I - WM)$, and $\text{Tr} \{.\}$ denotes the trace of a square matrix defined as the sum of the elements on the main diagonal. Please note that pseudo-inverse reconstruction method only provides a first approximation. Therefore, the MSE is not very accurate and only gives an upper-bound of total errors. Since the errors ϵ are independent from the spectrum s, $\mathrm{E}[s\epsilon^T] = 0$ and $\mathrm{E}[\epsilon s^T] = 0$. Therefore, MSE can be rewritten as

$$
MSE = \text{Tr} \left\{ H Corr_s H^T \right\} + \text{Tr} \left\{ W Corr_\epsilon W^T \right\},
\tag{6}
$$

where $Corr_s$, $Corr_\epsilon$ are the autocorrelation matrices $\mathrm{E}[ss^T]$, $\mathrm{E}[\epsilon\epsilon^T]$, respectively. Interestingly, in the above equation, the first term describes the error due to spectral recovery while the latter accounts for the error due to imaging noise and demosaicing. Next, we analyze the statistical properties of the natural multi-spectral images in order to model and approximate the correlation matrix $Corr_s$ and $Corr_\epsilon$. Finally, we also describe how MSFA patterns and demosaicing interpolation affect demosaicing error and noise, respectively.

3.1. Spectral Characteristics of Natural Images

In order to study this, we use three hyperspectral datasets—CAVE dataset, Harvard dataset, and one that we ourselves capture using a Surface Optics SOC-730 (Surface Optics Corporation, San Diego, CA, USA) camera with a 1024 × 1024 spatial resolution and a 2 nm (400–1000 nm) spectral resolution—this dataset contains 60 images of a sample, which is shown in Figure 2. Prior work has shown that multi-spectral images can be decomposed into Cartesian products of spectral and spatial components [31]. Therefore, we discuss them separately.
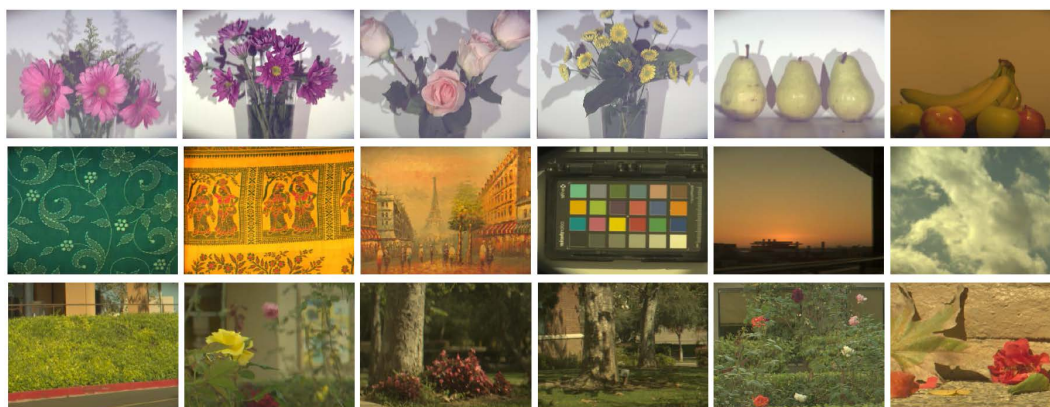

Figure 2. Several sample images of our natural multispectral images dataset. Each image was captured using an SOC-730 hyperspectral camera with a spatial resolution of 600 × 800 and 31 spectral measurements (400–700 nm) at each pixel.

Spectral components—prior works observe the following: (a) illumination and reflectance spectra of natural or man-made objects and phenomena are relatively smooth [32] and therefore frequency-limit functions can be used to approximate them [33]; (b) in addition to being smooth, these illumination and reflectance spectra also show remarkably similar energy distribution in the frequency domain (Figure 3). In fact, the surfaces formed by these distribution functions have a shape close to an exponential function with an exponent of (−2). Therefore, they follow Steven's power law of human perception. Steven's Law is the fundamental empirical law of human perception that most human responses to input environmental stimuli is related by a power function [34].

We illustrate the aforementioned properties in Figure 3. To eliminate the effect of varying illumination and brightness of the images, we first compute the difference spectrum $\Delta s = s - \mu$, where μ is the mean value in spectral direction of multispectral images. For each Δs, we compute

$$E[|\Delta s_{\mathcal{F}}(k)|^2] \propto 1/k^{2-\beta}, \tag{7}$$

where $\Delta s_{\mathcal{F}}(k)$ is the Fourier transform of difference spectrum $\Delta s(\lambda)$, k is the frequency of spectrum, $2 - \beta$ is the frequency exponent, and β clusters around 0 for natural spectra.

Spatial components—it is well known that the RGB real-world images, including both natural landscapes and man-made environments, also follow the Power Law [35,36] and tend to be scale invariant. This property has already been widely used in many computer graphics applications [37], e.g., exposing forgeries, imaging denoising. Interestingly, we also observe this property across different bands of real-world multispectral images (Figure 3). As in spectral components, we can show for spatial components as well that the difference band images, $\Delta I = I - \mu$, where μ is the mean of the spectral band across space, exhibits the same property as:

$$E[|\Delta I_{\mathcal{F}}(f, \theta)|^2] \propto A(\theta)/f^{2-\alpha(\theta)}, \tag{8}$$

where $\Delta I_{\mathcal{F}}$ is the 2D Fourier transform of difference band-images ΔI, (f, θ) is the polar coordinates of 2D frequency of band-images, $A(\theta)$ and $2 - \alpha(\theta)$ are the scaling function and the frequency exponent function respectively for each orientation θ. Here, $\alpha(\theta)$ clusters around 0 for natural images. Furthermore, for natural objects, the scaling function $A(\theta)$ clusters around 1; while, for man-made objects, the value of $A(\theta)$ for oblique orientation, θ_o, is much smaller than its value for horizontal and vertical orientations, θ_h and θ_v, respectively. This is due to the fact that man-made objects usually contain horizontal and vertical edges [35].
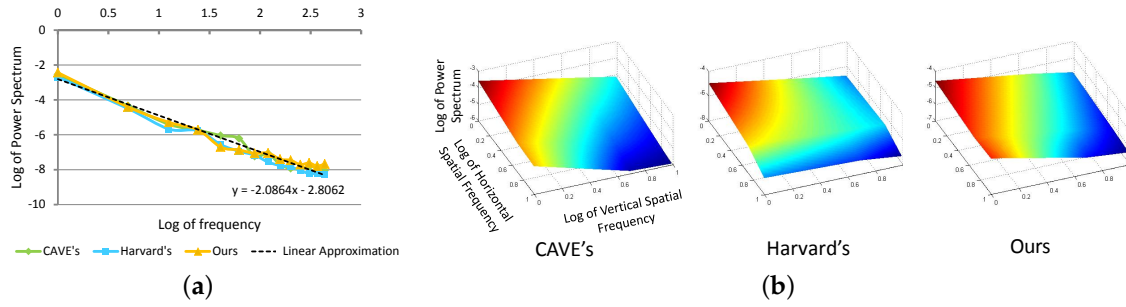

Figure 3. Spectral reflectance in four different data sets (Harvard's [31], CAVE [38], and our dataset) have similar behavior after Fourier transform. (**a**) Log of power spectrum of spectral component in four different data sets; (**b**) log of power spectrum of spatial component in four different data sets.

3.2. Modeling Recovery Error

The statistical fact that spatial and spectral components of multi-spectral images follow a power law is not dependent on any particular data set. Therefore, this provides a prior that we use to approximate recovery error $Tr\left\{ HCorr_s H^T \right\}$ in Equation (6). In order to achieve this, we approximate $Corr_s$ by exploiting the content independent power law and the *Wiener–Khinchin theorem*. Here, every spectral wavelength sample point s_i is treated the same. Mathematically, we assume $s_i = \Delta s_i + \mu$ ($1 \leq i \leq K$) have the same distribution with the mean value μ and the standard deviation σ; therefore, $E[s_i^2] = E[s_j^2] = \mu^2 + \sigma^2$ ($1 \leq i, j \leq K$).

We approximate $Corr_s(i, j)$ in matrix $Corr_s$ using the autocorrelation value $Corr_s(i, j) = \mu^2 + \eta_1 C_s(\Delta \lambda)$, where $\Delta \lambda$ is the distance between spectral band i and j, and η_1 is a scale factor. According to the Wiener–Khinchin theorem, the autocorrelation of a signal could be represented by using the inverse Fourier transform of power spectral function. Thus, a general form of autocorrelation value $C_s(\Delta \lambda) = \int \Delta s(\lambda) \Delta s(\lambda \pm \Delta \lambda) \, d\lambda$ can be described as:

$$C_s = \mathcal{F}^{-1}\{E[|\Delta s_{\mathcal{F}}(k)|^2]\}. \tag{9}$$

Finally, our approximation of $Corr_s$ can be expressed as

$$Corr_s = \mu^2 + \eta_1 \begin{bmatrix} C_s(0) & C_s(1) & \cdots & C_s(n-1) \\ C_s(1) & \ddots & & C_s(n-2) \\ \vdots & & \ddots & \vdots \\ C_s(n-1) & C_s(n-2) & \cdots & C_s(0) \end{bmatrix}. \tag{10}$$

Therefore, the final recovery error is estimated by $Tr\{HCorr_s H^T\}$. Compared with the methods constructing the autocorrelation matrix derived from some datasets, we use this error model because it only depends on the average intensity μ and the scale factor η_1. The effectiveness of the model will be verified in Section 5.1.

3.3. Modeling Demosaicing Error and Imaging Noise

In addition to spectral recovery error, the demosaicing error and imaging noise considerably affect the quality of the reconstruction. However, unlike the spectral recovery error, these errors are heavily dependent on (1) the design of MSFA (multi-spectral filter array), which involves the distribution and placement of the different channels in the filter array; and (2) the demosaicing technique that decides the order of the demosaicing process based on the dependency between channels or lack thereof.

Design of MSFA: prior works [39,40] have shown that a good MSFA pattern should have the following properties: (i) it should be spatially uniform to ensure robustness in the face of image sensor imperfections; (ii) it should be periodic to ensure efficiency of image reconstruction; and (iii) it should be neighbor consistent (each channel has the same neighbor channels) to ensure immunity to optical/electrical cross talk between neighboring pixels. Reference [40] presents a binary tree based MSFA design method that assures all the above properties and is elegant in design and implementation. Therefore, we adapt this design and customize it to suit our needs.

In the binary tree based MSFA design [40], the generation of the MSFA is an iterative process of binary splitting as illustrated in Figure 4a. A parent channel evenly divides into two children channels in each splitting, increasing the number of channels by one while scaling down the sampling rate of the children created by half. Reference [40] shows that, since the MSFA patterns satisfy the three properties above, the sibling channels are exchangeable. We can use binary trees to represent such MSFA patterns (Figure 4b) where a leaf node at level l has a sampling rate of 2^{-l}. Figure 4c,d shows that the patterns generated may not be unique. However, the binary tree for a particular sampling rate combination is unique.
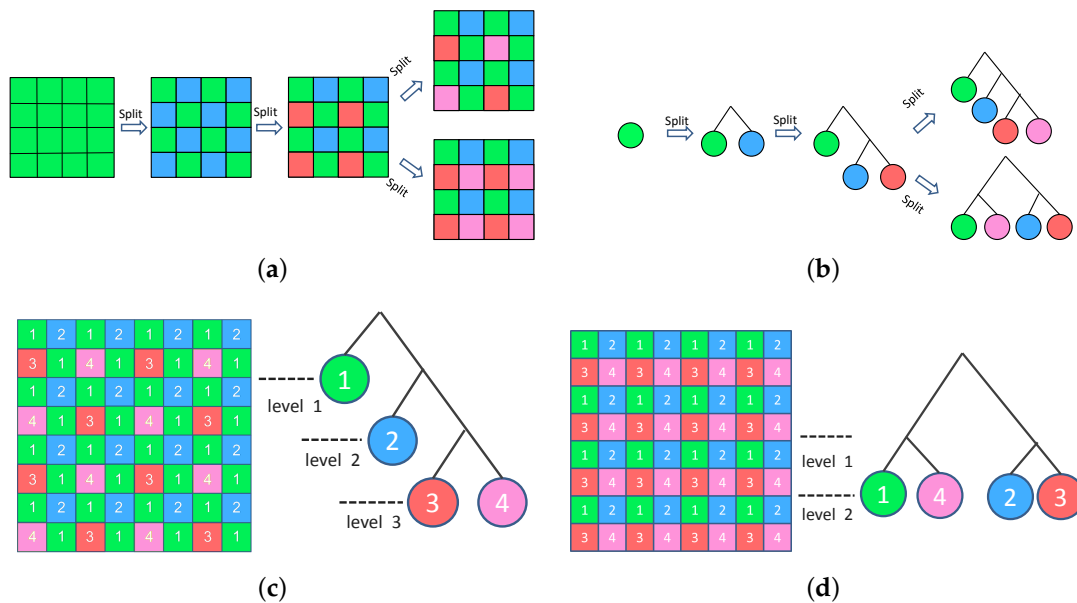

Figure 4. (**a**) Generation of the MSFAs. (**b**) Generation of the binary trees. (**c,d**) Left: two patterns of four channels MSFA; right: the binary trees of the four channels pattern have four leaves, each leaf represents a spectral channel. The four leaf nodes correspond to the four spectral channels with (**c**) sampling rate $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\}$ and (**d**) sampling rate $\{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\}$.

Demosaicing strategy: channel-independent demosaicing completes an undemosaiced channel at a subsampled resolution without dependence of other channels. On the contrary, channel-dependent demosaicing completes an undemosaiced channel at a subsampled resolution using the content from one of more different channels. The channel dependent methods are categorized as fully dependent or partially dependent based on if they use all the channels or a subset thereof for demosaicing, respectively.

Our observation reveals that neither "one-to-many channel-dependent demosaicing" [21] nor "all channel- independent demosaicing" [26] strategy are optimal (see Figure 5). Initially, we start with full resolution mosaic of the most important channel (e.g., green), we use dashed arrows from channel *a* to channel *b* to denote the demosaicing of the subsampled resolution channel *b* with the guidance of the full resolution channel *a*: self loops indicate channel-independent demosaicing of channel *b* without guidance while other arrows indicate channel-dependent demosaicing (Figure 5). It is important to note that channel selection, channel spatial arrangement, demosaicing strategy and imaging noise levels are mutually correlated. Therefore, we devise an adaptive demosaicing based on channels and imaging noise levels. In the following sections, we will model and quantify channel-independent demosaicing and channel-dependent demosaicing errors and use it for the design of the MSFA and the demosaicing method.
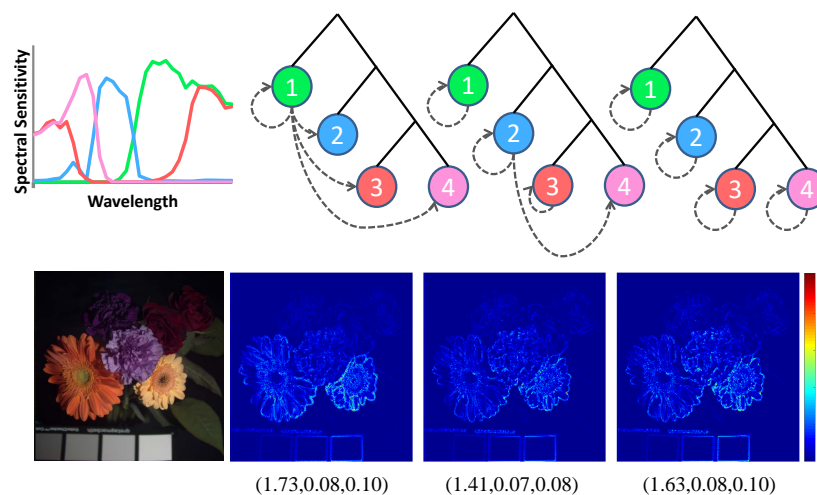

(1.73,0.08,0.10)   (1.41,0.07,0.08)   (1.63,0.08,0.10)

Figure 5. Reconstruction spectral errors of a multispectral image, and the reconstruction error statistics (mean, maximum, standard deviation) of 18 images caused by different demosaicing strategy, the multispectral images are from CAVE dataset. Dashed arrows from channel *a* to channel *b* to denote the demosaicing of the subsampled resolution channel *b* with the guidance of the full resolution channel *a* (from left to right: "one-to-many channel-dependent demosaicing"; partial channel-dependent demosaicing; "all channel-independent demosaicing").

3.4. Channel-Independent Demosaicing Error

First, we consider channel-independent demosaicing for a sensor with resolution R. Let the set of pixels' measuring channel i (at level l in the binary tree) directly on the sensor be I_i. Let us consider a pixel p and $x_i(p)$ denote the response of channel i ($1 \leq i \leq m$) at the pixel p. If pixel $p \in I_i$, the response $x_i(p)$ is the addition of noise-free ground truth response $\tilde{x}_i(p)$ and imaging noise $\zeta_i(p)$. If $p \notin I_i$, i.e., $x_i(p)$ is interpolated from other directly measured pixels during demosaicing, then $x_i(p)$ is the sum of $\tilde{x}_i(p)$, $\zeta_i(p)$ and demosaicing error $\delta_i(p)$.

Therefore, the response x_i at pixel p can be written as:

$$x_i(p) = B_i(p)\left(\tilde{x}_i(p) + \zeta_i(p)\right) + \bar{B}_i(p) \sum_{p' \in P} w_{l_i}(p, p')\left(\tilde{x}_i(p') + \zeta_i(p')\right), \qquad (11)$$

where $B_i(p)$ is a binary flag that 1 for $p \in I_i$ and 0, otherwise, and $w_{l_i}(p, p')$ are corresponding interpolation weights, determined by both the level l_i and the Gaussian filter used. Combining Equations (2) and (11), we find the demosaicing error and imaging noise of demosaiced pixel p as:

$$
\begin{cases}
\zeta_i(p) = B_i(p)\zeta_i(p) + \bar{B}_i(p)\sum_{p'\in P} w_{l_i}(p,p')\zeta_i(p'), \\[2mm]
\delta_i(p) = \bar{B}_i(p)\left(\tilde{x}_i(p) - \sum_{p'\in P} w_{l_i}(p,p')\tilde{x}_i(p')\right).
\end{cases}
\tag{12}
$$

The noise of directly observed pixels come from measurement errors and quantization errors resulting from analog to digital conversion and therefore spatially independent [41]. We use $\bar{\zeta}_i^2$ to denote the average directly observed noise variance in I_i. Noise variance of channel i, $E[\zeta_i^2]$ is a constant that can be estimated using previous method by Shimano [42] as:

$$
\begin{aligned}
E[\zeta_i^2] &= \frac{1}{R}\sum_p\left(B_i(p)\bar{\zeta}_i^2 + \bar{B}_i(p)\sum_{p'\in P} w_{l_i}^2(p,p')\bar{\zeta}_i^2\right) \\
&= \left(2^{-l_i} + (1-2^{-l_i})\varphi_{l_i}\right)\bar{\zeta}_i^2,
\end{aligned}
\tag{13}
$$

where φ_{l_i} denotes $\sum_p\sum_{p'\in P} w_{l_i}^2(p,p')$ for abbreviation. Demosaicing error of channel i at pixel p is given by the square of difference between ground truth response $x_i(p)$ and the interpolated value obtained from the ground truth responses of the neighboring pixels. Let us now assume that, for each pixel p, the square of ground truth response has the same distribution. Thus, for two pixels, p and q, $p \neq q$, we have $E[\tilde{x}_i^2(p)] = E[\tilde{x}_i^2(q)]$. Therefore, demosaicing error variance $E[\delta_i^2]$ of channel i is given by the average of the variance $mean(\delta_i(p)^2)$ of all the pixels in the image and is derived from Equation (14) as:

$$
\begin{aligned}
E[\delta_i^2] &= \frac{1}{R}\sum_p \bar{B}_i(p)\left(\tilde{x}_i(p) - \sum_{p'\in P} w_l(p,p')\tilde{x}_i(p')\right)^2 \\
&= \alpha_l E[\tilde{x}_i^2] + \sum_{p,q}^{p\neq q}\beta_l(p,q)E[\tilde{x}_i(p)\tilde{x}_i(q)],
\end{aligned}
\tag{14}
$$

where α_l and $\beta_l(p,q)$ are coefficients of $E[\tilde{x}_i^2(p)]$ and $E[\tilde{x}_i(p)\tilde{x}_i(q)]$, respectively, in the expansion of Equation (14). We model $E[\tilde{x}_i(p)\tilde{x}_i(q)]$ as $M_i E[\tilde{I}_i(p)\tilde{I}_i(q)^T]M_i^T$, where M_i is the known spectral sensitivity of i th channel. We assume a spatially independent autocorrelation matrix $E[\tilde{I}_i(p)\tilde{I}_i(p)^T]$ that can be simplified to $E[\tilde{I}_i\tilde{I}_i^T]$. Similar to spectral variable s_i, we assume the distribution of $\tilde{I}_i$ has the mean value μ and the the standard deviation σ, where $\tilde{I}_i = \mu + \Delta\tilde{I}_i$. We use $C_I(\Delta u, \Delta v) = \iint \Delta\tilde{I}(u,v)\,\Delta\tilde{I}(u\pm\Delta u, v\pm\Delta v)\,du\,dv$ to represent the ratio between correlation matrix $E[\tilde{I}_i(p)\tilde{I}_i(q)^T]$ and $E[\tilde{I}_i\tilde{I}_i^T]$:

$$
E[\tilde{I}_i(p)\tilde{I}_i(q)^T] = \frac{\mu^2 + \eta_2 C_I(\Delta u, \Delta v)}{\mu^2 + \eta_2 C_I(0,0)}E[\tilde{I}_i\tilde{I}_i^T],
\tag{15}
$$

where $(\Delta u, \Delta v)$ is given by the absolute value of vector difference between locations of pixel p and q, and η_2 is a scale factor. Similar to the approximation of spectral autocorrelation matrix, $C_I(\Delta u, \Delta v)$ can also be approximated using Wiener–Khinchin theorem as:

$$
C_I = \mathcal{F}^{-1}\{E[|\Delta I_{\mathcal{F}}(f,\theta)|^2]\}.
\tag{16}
$$

Since demosaicing error is independent from imaging noise, the variance $E[\epsilon_j^2]$ in correlation matrix $Corr_e$ is the sum of variance of demosaicing error and imaging noise: $E[\epsilon_i^2] = E[\delta_i^2] + E[\zeta_i^2]$.

3.5. Channel-Dependent Demosaicing Error

In channel dependent demosaicing, we denote the guidance channel and the demosaiced channel by g and r from an upper and lower level l_g and l_r respectively. Prior works show that color-difference invariance in local regions is a reasonable assumption for chrominance (red/blue) channel demosaicing [43] yielding low computational complexity [44]. This color-difference invariance results in a linear demosaicing operation that is convenient for representation and computation. Therefore, we construct a demosaicing error model using the same color-difference invariance and express the response $x_r(p)$ of demosaiced channel r at position p as:

$$
\begin{aligned}
x_r(p) =& \tilde{x}_g(p) + \epsilon_g(p) + B_r(p)\left(\tilde{x}_r(p) + \zeta_r(p) - \tilde{x}_g(p) - \epsilon_g(p)\right) \\
&+ \bar{B}_r(p)\sum_{p'\in P} w_{l_r}(p,p')\left(\tilde{x}_r(p') + \zeta_r(p') - \tilde{x}_g(p') - \epsilon_g(p')\right).
\end{aligned}
\tag{17}
$$

From the above equation, demosaicing error variance can be transformed into a form like Equation (14) as:

$$
E[\delta_r^2] = \alpha_{l_r}E[(\tilde{x}_{r\backslash g})^2] + \sum_{p,q}^{p\neq q}\beta_{l_r}(p,q)E[(\tilde{x}_{r\backslash g}(p))(\tilde{x}_{r\backslash g}(q))],
\tag{18}
$$

where $\tilde{x}_{r\backslash g}$ denotes $\tilde{x}_r - \tilde{x}_g$. Here, the demosaicing error just contains the interpolation error caused in this demosaicing. The imaging noise from demosaiced channel and accumulative errors from guide channel are induced into the noise variance as:

$$
\begin{aligned}
E[\zeta_r^2] &= \frac{1}{R}\sum_p\left[B_r(p)\bar{\zeta}_r^2 + \bar{B}_r(p)\sum_{p'\in P}w_{l_r}^2(p,p')\bar{\zeta}_r^2 + \bar{B}_r(p)\left(1 + \sum_{p'\in P}w_{l_r}^2(p,p')\right)\frac{E[\epsilon_g^2] - 2^{-l_g}\bar{\zeta}_g^2}{1 - 2^{-l_g}}\right] \\
&= \left(2^{-l_r} + (1-2^{-l_r})\varphi_{l_r}\right)\bar{\zeta}_r^2 + \frac{(1-2^{-l_r})(1+\varphi_{l_r})(E[\epsilon_g^2] - 2^{-l_g}\bar{\zeta}_g^2)}{1 - 2^{-l_g}},
\end{aligned}
\tag{19}
$$

where $E[\epsilon_g^2]$ denote the error variance of guide channel g. Note that, in Equation (19), the noise variance of channel r consists of two components: (1) the imaging noise variance ζ_r^2 has the same form of channel-independent demosacing (as in Equation (13)); (2) the error variance $E[\epsilon_g^2]$ of guidance channel with the coefficient decided by the level of the channel r and the filter used in the demosaicing algorithm. Similar to the case of channel-independent demosaicing, the error variance of demosaiced channel r can be calculated by $E[\epsilon_r^2] = E[\delta_r^2] + E[\zeta_r^2]$.

The relationship between channels in demosaicing can be represented using a *Demosaicing Forest* as illustrated in Figure 6. In this forest, the channel demosaiced without guidance of other channels is the root of a tree in the forest, and a pair of parent and child nodes in a tree represents a pair of guidance and demosaiced channels in channel-dependent demosaicing. Hence, for each channel i, there is a unique *demosaicing chain* containing all reachable nodes from the root to channel i, denoted by $DC(i)$. Intuitively, if two channels i and j come from different trees, i.e., $DC(i) \cap DC(j) = \varnothing$, then the two channels are independent; otherwise, the two channels would have a *lowest common chain element* k, where the level of k, $l_k = max(level(DC(i) \cap DC(j)))$. The errors of these two channels are then partially correlated due to sharing of the same component from the error of channel k. Thus, for the off-diagonal elements $E[\epsilon_i\epsilon_j]$ (the correlation between the errors of channel i and j) in correlation $Corr_e$, we have:

$$
E[\epsilon_i\epsilon_j] = \begin{cases} 0, & \text{if } DC(i) \cap DC(j) = \varnothing, \\ \dfrac{(1-2^{-l_i})^{\frac{3}{2}}(1-2^{-l_j})^{\frac{3}{2}}(E[\epsilon_k^2] - 2^{-l_k}\bar{n}_k^2)}{\prod_p(1+\varphi_{l_p})^{-\frac{1}{2}}\prod_q(1+\varphi_{l_q})^{-\frac{1}{2}}(1-2^{-l_k})}, & \text{otherwise,} \end{cases}
\tag{20}
$$

$$
s.t. \quad i \neq j, p \in DC(i) - DC(k), q \in DC(j) - DC(k).
$$

Therefore, the objective function *MSE* is built with a few parameters: the Gaussian filter we used in demosaicing, scale factors $\frac{\eta_1}{\mu^2}$ and $\frac{\eta_2}{\mu^2}$, and imaging noise $\bar{\zeta}_i^2$ of each channel.
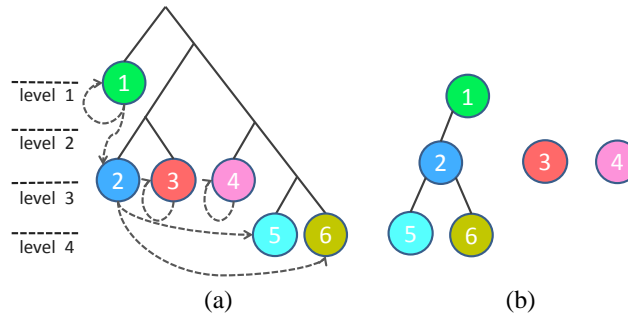

| | (a) | | | (b) | |

Figure 6. (**a**) The binary tree of a MSFA pattern and the demosaicing order; (**b**) the corresponding demosaicing forest.

4. Imaging Optimization Method

In this section, to optimize reconstruction accuracy, we will seek possible combinations of channels, MSFA patterns, and demosaicing patterns that minimize the objective function (Equation (6)) with specific parameters. We propose an optimization method that selects *n* channels from *m* candidate channels ($n \ll m$) and determines the MSFA patterns and demosaicing orders to facilitate accurate recovery of spectral reflectance of most natural and man-made objects.

Considering manufacturing complexity, we assume the number of channels, *n*, on MSFA to be relatively small ($n < 7$). Using the binary-tree-based constraint, we know that the number of possible patterns $\mathcal{P}$ with *n* channels are also small. For example, for $n < 7$, num $(\mathcal{P}) = 2, 3, 5$ when $n = 4, 5, 6$. Therefore, we can enumerate every possible pattern for a specific number of channels, and calculate the optimal channel selection, the mosaic pattern, and demosaicing order for each pattern.

However, searching for an ordered subset of *n* channels from *m* candidates that would minimize the objective function is NP hard. Therefore, an exhaustive search is time-consuming. For example, with $n = 6$ and $m = 100$, we will need to evaluate $P(100, 6) \approx 10^{12}$ permutations. Therefore, we apply a discrete particle swarm optimization (DPSO) method, detailed in Algorithm 1 to search for the optimal imaging parameters.

Particle swarm optimization [45] is a global search method that operates on a population of particles where each particle represents a candidate solution. The method moves these particles around in the search-space and updates the position and velocity of each particle iteratively.

Algorithm 1 The Proposed DPSO Method

1:　**Input:** Set of candidate channels: $\{M_j(\lambda)|j = 1, \ldots, m\}$; Number of output channels: *n*; Pattern $\mathcal{P}$.
2:　**Output:** Selected ordered set of *n* channels; Decided demosaicing order $\mathcal{O}$.
3:　**Initialization:**
4:　Generate population X_i and velocity V_i　　　$(1 \leq i \leq P)$
5:　$[X^{gb}, \mathcal{O}^{gb}] = findbest(X_i^{pb}, \mathcal{P})$
6:　**for** $j = 1 \ldots G$ **do**
7:　　**for** $i = 1 \ldots P$ **do**
8:　　　$X_i^{pb} = f(M_{X_i}, \mathcal{P}) < f(M_{X_i^{pb}}, \mathcal{P}) ? X_i : X_i^{pb}$
9:　　　$V_i = w(V_{i-1} + c_1 r * (X_i^{pb} \ominus X_i) + c_2 r * (X^{gb} \ominus X_i))$
10:　　$\bar{V}_i = reducemat(V_i, p)$
11:　　$X_i = X_i \oplus \bar{V}_i$
12:　　**end for**
13:　$[X^{gb}, \mathcal{O}^{gb}] = findbest(X_i^{pb}, \mathcal{P})$　　$(1 \leq i \leq P)$
14:　**end for**
15:　**return** *n* ordered channels X^{gb}, demosaicing order $\mathcal{O}^{gb}$;

Let the number of particles be P and the number of iteration be G. Note that ith particle X_i denotes ith candidate solution where each solution consists of n ordered channel selected from m candidate channels, and therefore each particle X_i is a $m \times 1$ integer vector. In vector X_i, the element with positive value k indicates the channel is the k th selected channel in pattern $\mathcal{P}$, while those with 0 value indicate the channels that are not selected. The function $findbest()$ finds the global optimal tuple of ordered channels X^{gb} and the optimal demosaicing order $\mathcal{O}^{gb}$ for the specific pattern $\mathcal{P}$. The behavior of the algorithm is adjusted using four parameters: c_1, c_2, w, and p. For each iteration j, the algorithm maintains the global best solution X^{gb} and the local best solution X_i^{pb} of each particle. The objective function $f(M_{X_i}, \mathcal{P})$ calculates the *MSE* of the imaging with selected channels X_i and specific pattern $\mathcal{P}$ for every possible demosaicing orders and then return the minimum *MSE*.

Each iteration of particle i uses standard procedures in a basic particle swarm optimization but with modified operations. We first calculate the velocity from X_i to X_i^{pb} and the velocity from X_i to X^{gb} using the mutation operation $\ominus$. The velocity from vector A to B is defined as a differential matrix of swapping pairs from A to B. Then, the algorithm updates the velocity V_i using weighted sum of previous velocity V_{i-1}, $X_i^{pb} \ominus X_i$, and $X^{gb} \ominus X_i$; here, the weighted values indicate the possibility to swapping pairs. Then, the algorithm uses the function $reducemat()$ to retain the principal component of velocity V_i by clipping the velocity with a threshold probability p; finally, it updates the solution by adding clipped velocity $\bar{V}_i$ to the current solution X_i. The adding operation $X_i \oplus \bar{V}_i$ is defined as swapping the status(selected with index or unselected) of the xth channel and the yth channel in set X_i, where (x, y) is a swapping pair in matrix $\bar{V}_i$ [46].

5. Evaluation and Comparison

In this section, we evaluate our error and noise models in simulation using popular multi-spectral image datasets. Next, we compare our optimization method with GAP camera design and Chi's filter selection method. Finally, we explore the optimal number of channels and how imaging noise level affects MSFA design and demosaicing strategy. The candidate set of channels used in simulation is constructed with spectral transmission data of 30 off-the-shelf coating color filters from http://www.rosco.com/filters (see Figure 7).
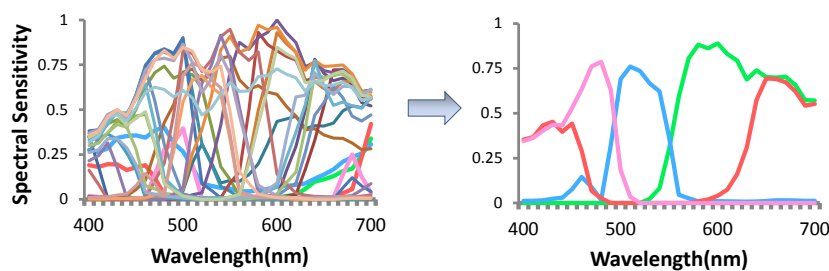

Figure 7. Four channels selected from a candidate set.

5.1. Error Models

We evaluate the estimated spectral recovery from our method and compare it with Chi's [1] and Shen's method [18]. We randomly choose 100 different channel combinations from the candidate set, each channel combination consisting of six different channels. Next, we calculate average estimated spectral recovery error only (ignoring the error due to noise and demosaicing) for each channel combination using different spectral dataset in [47], which contains natural spectral reflectances in abundance.

Figure 8 shows the relationship between objective function values for different channel combinations and the corresponding spectral recovery error on two different datasets for Chi's, Shen's and our method. Note Chi's method is content-independent, but Shen's method uses a prior based on a large generic spectral dataset [32]. Unlike other methods, our method shows a strong

linear relationship between evaluated objective function and the estimated spectral recovery error, confirming the higher accuracy of our objective function as a predictor for spectral recovery error.
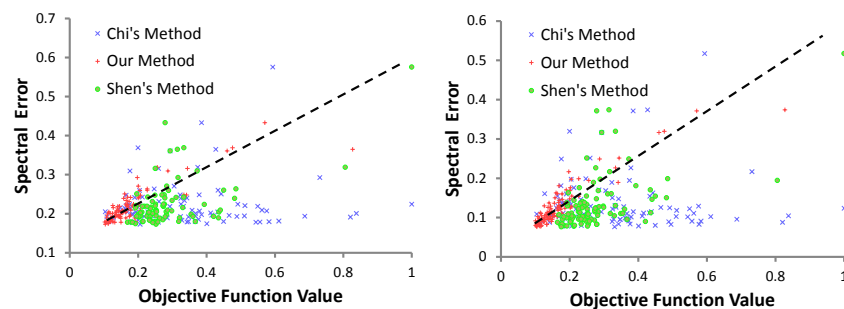

Figure 8. Relationship between spectral estimated error and objective function value of previous methods and our method on "natural" (**left**) and "paper" (**right**) dataset in [47]. Note the scatters of our method distribute along the black dashed line while the scatters of the previous method distribute irregularly.

To evaluate our overall models, we consider three different random combination of channels, MSFA patterns, and demosaicing orders as shown in Figure 9. We adopt the bilinear interpolation method and the binary tree-based generic(BTG) demosaicing method [26] as the channel-independent demosaicing, and adopt state-of-the-art guided filter (GF) method [48] and residual interpolation (RI) method [49] as the channel-dependent demosaicing in reconstruction. In error estimation, we used scale factor $\eta_1 = 0.025$ and $\eta_2 = 0.15$, and random Gaussian noise with $SNR \approx 50$ db to the response of camera. The values of the scale factors are directly obtained from the statistics of the multispectral images in the used datasets.
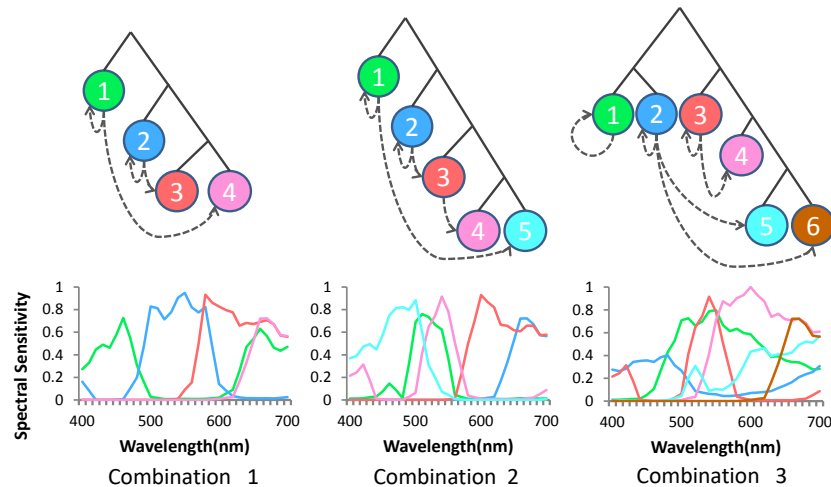

Figure 9. Three examples of combinations of different channels, MSFA patterns and demosaicing orders. Note that the three examples have different numbers of channels (4, 5, 6). The combinations are set casually to verify our overall model.

The comparison between estimated errors and actual reconstruction errors using different demosaicing methods is shown in Figure 10. The actual errors are acquired by calculating the difference between ground truth and reconstructed results. It is worth noting that impact of different combination is much more than that of different demosaicing methods. Therefore, the combination is the primary factor while the methods are secondary. Our estimated errors are close to the average actual errors and illustrate the suitability of our model as a good descriptor for overall reconstruction error.
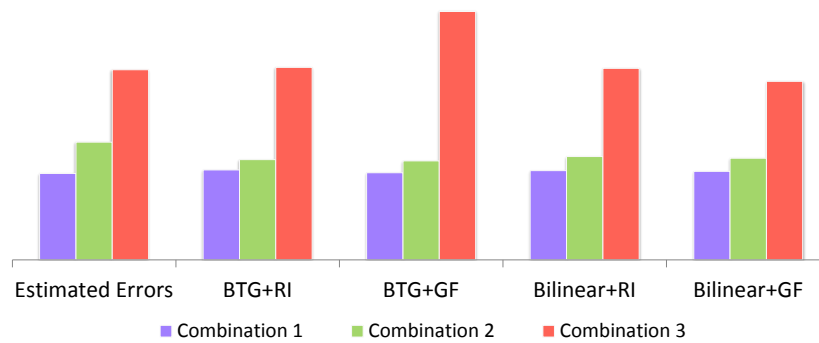
Figure 10. Comparison between estimated errors and actual reconstruction errors using different demosaicing methods (BTG + RI, BTG + GF, Bilinear + RI, Bilinear + GF) with the channels and patterns shown in Figure 9. The estimated errors are scaled for comparison.

5.2. Comparison with Previous Methods

To verify the effectiveness of our optimization method, we select channels from candidate channel set using our method, decide channel arrangement on MSFA pattern and demosaicing order using three methods—GAP camera [12], Chi and Monno's method [1,14], and our method—as shown in Figure 11, and compare spectral reconstruction accuracy of our method with the other two methods on three different datasets—'CAVE' spectral dataset [38], Harvard's spectral dataset [31], and our dataset. Chi and Monno's method is a combination of Chi's channel selection method and Monno's MSFA design and demosaicing order.
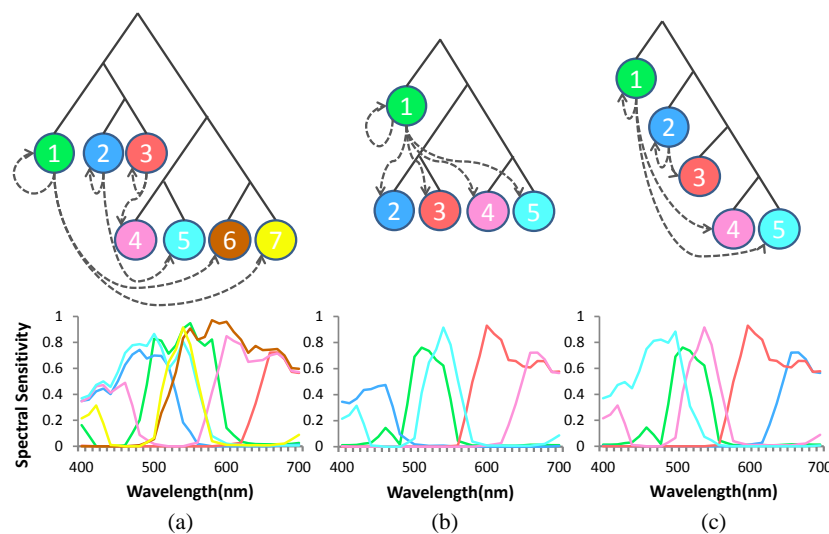

Figure 11. Selected channels, mosaic binary tree, and demosacking order of (**a**) GAP camera; (**b**) Chi and Monno's method; and (**c**) our method.

We used the demosaicing method in [26] and added random Gaussian noise to responses of camera with $SNR \approx 50$ db. To quantify the reconstruction accuracy of multispectral images, especially on edges, we use relative error $|\Delta s(\lambda)|$ defined as:

$$|\Delta s(\lambda)| = \sqrt{\frac{\int (s(\lambda) - \hat{s}(\lambda))^2 d\lambda}{\int s^2(\lambda) d\lambda}}. \tag{21}$$

Table 1 shows reconstruction error of the three optimized channel combinations in Figure 11 along with the error statistics (mean, maximum, standard deviation). The optimized channels of our method show superior results to GAP camera's and Chi and Monno's results irrespective of the dataset used.

Table 1. Spectral reconstruction error of three optimized channels (see Figure 11) of GAP, Chi and Monno's method, and our method.

Methods	CAVE's Dataset			Harvard's Dataset			Our Dataset		
	Max	Mean	Std	Max	Mean	Std	Max	Mean	Std
GAP	0.4518	0.3231	0.0880	0.2849	0.0794	0.0854	0.2602	0.1964	0.0421
Chi and Monno's	0.4381	0.2852	0.0867	0.2498	0.0744	0.0753	0.2231	0.1880	0.0428
Ours	**0.4115**	**0.2775**	**0.0814**	**0.2196**	**0.0629**	**0.0679**	**0.1999**	**0.1586**	**0.0342**

Figure 12 shows comparison of the three methods on two multi-spectral images. Since our method takes into account the demosaicing error, our estimated spectral error near sharp edges in images is significantly smaller than others.
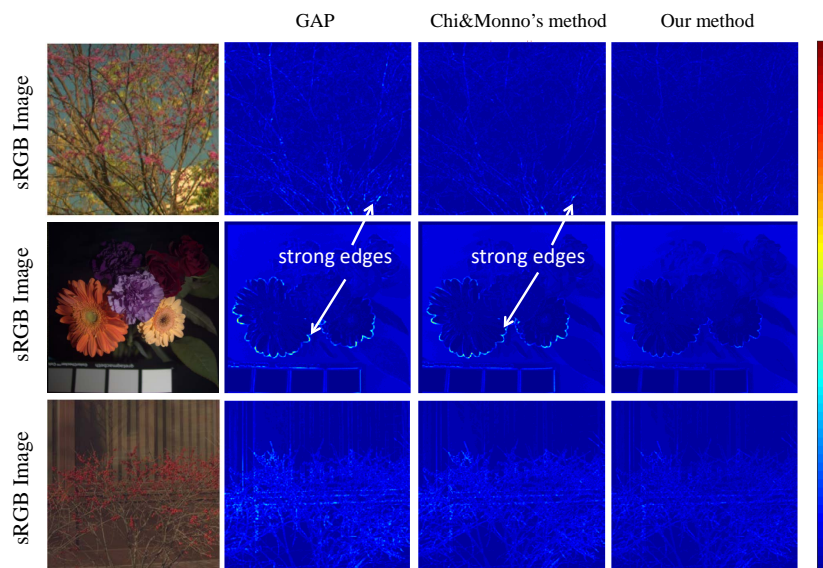

Figure 12. Reconstruction spectral error of multi-spectral images in our database (the 1st row), CAVE spectral database (the 2nd row), and Harvard's database (the 3th row) using a GAP camera, Chi and Monno's method and our mosaic camera (see Figure 11). Please zoom in and see sharp edges in gray images.

6. Discussion

Optimal Number of Channels. In order to explore the best number of mosaiced channels on MSFAs, in Figure 13, we plot the estimated error using our error model and optimization methods against the different number of channels. Note that the recovery error decreases with an increase in number of channels, while the errors caused by demosaicing and imaging noise increase with the increasing number of channels. This is expected since more channels would result in more accurate recovery while sacrificing spatial resolution for each channel. Therefore, the sweet spot is where the sum of these two curves has a minima. In Figure 13, we see that, using our candidate channels, this is around 5–6 channels.
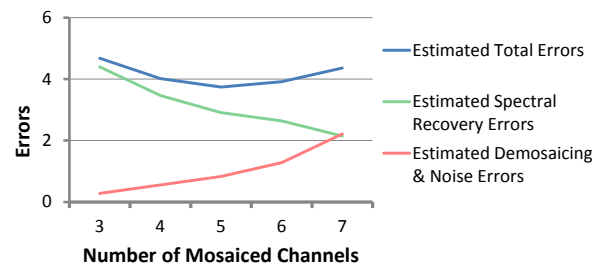
Figure 13. Our estimated error (spectral recovery error, demosaicing error and imaging noise) on "CAVE" spectral dataset [38] with a different number of mosaiced channels. The errors reveal the optimal number of channels to be around 5–6. Here, we use the MSFA and demosaicing method in [26].

Behavior of Imaging Noises. Adopting our model and optimization method, we can also explore how imaging noise affects optimal channel selection, design of MSFA, and demosaicing strategy via simulation. Figure 14 shows the optimized results under different imaging noise levels. It can be seen that, under higher noise levels, the optimized channel combination would have higher light throughput to preserve SNR of cameras (the observation is similar to previous works [13,20]). We also found that, in the presence of higher noise, the MSFA pattern tends to be more uniform, that is, the binary tree is more balanced. Furthermore, the demosaicing method tends to be more channel-independent. This is expected since, under high noise levels, the magnitude of noise is much larger than the demosaicing error, while channel-dependent demosaicing propagates noise between channels, although it can reduce demosaicing error. These conclusions provide guidelines to effectively reduce the search space in our optimization method.
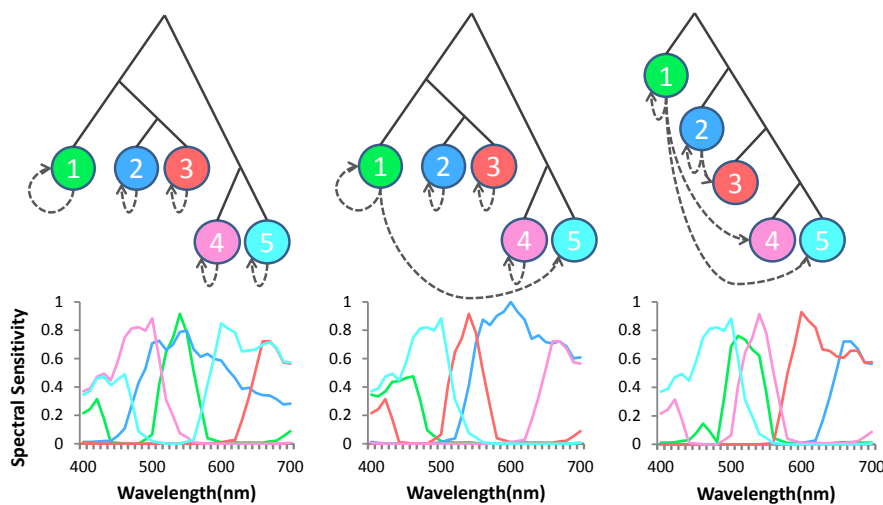

Figure 14. The optimized results under different noise levels (from left to right: SNR = 4, 16, 64 db).

7. Conclusions

In summary, we propose new error models for multi-spectral imaging and utilize the models to select optimal channel combination, the pattern of MSFA, and demosaicing order for multi-spectral imaging. We verified the effectiveness of our method and compared it with previous methods.

Our method can also be applied to other similar problems. For example, projection display is a dual imaging system, and, therefore, selecting few efficient primaries for a spectrally accurate display is a dual problem of our channel selection. In other areas, such as wireless sensor networking, where sensor or observation selection [50] is critical, our optimization might provide an effective solution in acceptable time.

In the future, we plan to apply our method to such varied domains, our immediate interest being in multi-spectral display. Moreover, we plan to explore the relationship between spectral sensitivity of tunable filter camera channels and accuracy of multi-spectral image registration.

Author Contributions: Aditi Majumder and M. Gopi conceived and designed the experiments; Yuqi Li performed the experiments; Yuqi Li and Hao Zhang analyzed the data; Yuqi Li wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chi, C.; Yoo, H.; Ben-Ezra, M. Multi-Spectral Imaging by Optimized Wide Band Illumination. *Int. J. Comput. Vis.* **2010**, *86*, 140–151.
2. Arad, B.; Ben-Shahar, O. Filter Selection for Hyperspectral Estimation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 3153–3161.
3. Dannemiller, J.L. Spectral Reflectance of Natural Objects: How Many Basis Functions Are Necessary? *J. Opt. Soc. Am. A* **1992**, *9*, 507–515.
4. Chiao, C.-C.; Cronin, T.W.; Osorio, D. Color Signals in Natural Scenes: Characteristics of Reflectance Spectra and Effects of Natural Illuminants. *J. Opt. Soc. Am. A* **2000**, *17*, 218–224.
5. Hagen, N.; Kudenov, M.W. Review of Snapshot Spectral Imaging Technologies. *Opt. Eng.* **2013**, *52*, 090901.
6. Gao, L.; Bedard, N.; Hagen, N.; Kester, R.T.; Tkaczyk, T.S. Depth-resolved image mapping spectrometer (IMS) with structured illumination. *Opt. Express* **2011**, *19*, 17439–17452.
7. Mohan, A.; Raskar, R.; Tumblin, J. Agile Spectrum Imaging: Programmable Wavelength Modulation for Cameras and Projectors. *Comput. Graph. Forum* **2008**, *27*, 709–717.
8. Wagadarikar, A.A.; Pitsianis, N.P.; Sun, X.; Brady, D.J. Spectral Image Estimation for Coded Aperture Snapshot Spectral Imagers. In *Image Reconstruction from Incomplete Data V*; SPIE: Bellingham, WA, USA, 2008; p. 707602.
9. Lin, X.; Liu, Y.; Wu, J.; Dai, Q. Spatial-spectral Encoded Compressive Hyperspectral Imaging. *ACM Trans. Graph.* **2014**, *33*, 233.
10. Jeon, D.S.; Choi, I.; Kim, M.H. Multisampling Compressive Video Spectroscopy. *Comput. Graph. Forum* **2016**, *35*, 467–477.
11. Mitra, K.; Cossairt, O.; Veeraraghavan, A. Can We Beat Hadamard Multiplexing? Data Driven Design and Analysis for Computational Imaging Systems. In Proceedings of the 2014 IEEE International Conference on Computational Photography (ICCP), Santa Clara, CA, USA, 2–4 May 2014.
12. Yasuma, F.; Mitsunaga, T.; Iso, D.; Nayar, S.K. Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum. *IEEE Trans. Image Process.* **2010**, *19*, 2241–2253.
13. Sajadi, B.; Majumder, A.; Hiwada, K.; Maki, A.; Raskar, R. Switchable primaries using shiftable layers of color filter arrays. *ACM Trans. Graph.* **2011**, *30*, 65.
14. Monno, Y.; Kikuchi, S.; Tanaka, M.; Okutomi, M. A Practical One-Shot Multispectral Imaging System Using a Single Image Sensor. *IEEE Trans. Image Process.* **2015**, *24*, 3048–3059.
15. Parmar, M.; Reeves, S.J. Selection of Optimal Spectral Sensitivity Functions for Color Filter Arrays. *IEEE Trans. Image Process.* **2010**, *19*, 3190–3203.
16. Sadeghipoor, Z.; Lu, Y.M.; Süsstrunk, S. Optimum spectral sensitivity functions for single sensor color imaging. In Proceedings of the IS&T/SPIE Electronic Imaging, Burlingame, CA, USA, 22–26 January 2012; p. 829904.
17. Park, J.I.; Lee, M.H.; Grossberg, M.D.; Nayar, S.K. Multispectral Imaging using Multiplexed Illumination. In Proceedings of the 2007 IEEE 11th International Conference on Computer Vision, Rio de Janeiro, Brazil, 14–21 October 2007; pp. 1–8.
18. Shen, H.L.; Yao, J.F.; Li, C.; Du, X.; Shao, S.J.; Xin, J.H. Channel selection for multispectral color imaging using binary differential evolution. *Appl. Opt.* **2014**, *53*, 634–642.

19. López-Álvarez, M.A.; Hernández-Andrés, J.; Valero, E.M.; Romero, J. Selecting algorithms, sensors, and linear bases for optimum spectral recovery of skylight. *J. Opt. Soc. Am. A* **2007**, *24*, 942–956.

20. Shimano, N. Optimization of Spectral Sensitivities with Gaussian Distribution Functions for a Color Image Acquisition Device in the Presence of Noise. *Opt. Eng.* **2006**, *45*, 013201.

21. Monno, Y.; Kitao, T.; Tanaka, M.; Okutomi, M. Optimal Spectral Sensitivity Functions for a Single-Camera One-Shot Multispectral Imaging System. In Proceedings of the 19th IEEE International Conference on Image Processing, Orlando, FL, USA, 30 September–3 October 2012; pp. 2137–2140.

22. Jia, J.; Barnard, K.J.; Hirakawa, K. Fourier spectral filter array for optimal multispectral imaging. *IEEE Trans. Comput. Imaging* **2016**, *25*, 1530–1543.

23. Lapray, P.J.; Wang, X.; Thomas, J.B.; Gouton, P. Multispectral Filter Arrays: Recent Advances and Practical Implementation. *Sensors* **2014**, *14*, 21626–21659.

24. Li, Y.; Wang, C.; Zhao, J.; Yuan, Q. Efficient spectral reconstruction using a trichromatic camera via sample optimization. *Vis. Comput.* **2018**, 1–11, doi:10.1007/s00371-017-1469-3.

25. Shinoda, K.; Hamasaki, T.; Hasegawa, M.; Kato, S.; Ortega, A. Quality metric for filter arrangement in a multispectral filter array. In Proceedings of the Picture Coding Symposium (PCS), San Jose, CA, USA, 8–11 December 2013; pp. 149–152.

26. Miao, L.; Qi, H.; Ramanath, R.; Snyder, W.E. Binary tree-based generic demosaicking algorithm for multispectral filter arrays. *IEEE Trans. Image Process.* **2006**, *15*, 3550–3558.

27. Jaiswal, S.P.; Fang, L.; Jakhetiya, V.; Pang, J.; Mueller, K.; Au, O.C. Adaptive multispectral demosaicking based on frequency-domain analysis of spectral correlation. *IEEE Trans. Image Process.* **2017**, *26*, 953–968.

28. Mihoubi, S.; Losson, O.; Mathon, B.; Macaire, L. Multispectral demosaicing using pseudo-panchromatic image. *IEEE Trans. Comput. Imaging* **2017**, *3*, 982–995.

29. Alvarez-Cortes, S.; Kunkel, T.; Masia, B. Practical Low-Cost Recovery of Spectral Power Distributions. *Comput. Graph. Forum* **2016**, *35*, 166–178.

30. Shimano, N.; Terai, K.; Hironaga, M. Recovery of spectral reflectances of objects being imaged by multispectral cameras. *J. Opt. Soc. Am. A* **2007**, *24*, 3211–3219.

31. Chakrabarti, A.; Zickler, T. Statistics of Real-World Hyperspectral Images. In Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Colorado Springs, CO, USA, 20–25 June 2011; pp. 193–200.

32. Parkkinen, J.P.; Hallikainen, J.; Jaaskelainen, T. Characteristic spectra of Munsell colors. *J. Opt. Soc. Am. A* **1989**, *6*, 318–322.

33. Stiles, W.S.; Wyszecki, G.; Ohta, N. Counting metameric object-color stimuli using frequency-limited spectral reflectance functions. *J. Opt. Soc. Am. A* **1977**, *67*, 779–786.

34. Goldstein, E.B.; Brockmole, J. *Sensation and Perception*; Cengage Learning: Boston, MA, USA, 2016.

35. Torralba, A.; Oliva, A. Statistics of Natural Image Categories. *Netw. Comput. Neural Syst.* **2003**, *14*, 391–412.

36. Hyvarinen, A.; Hurri, J.; Hoyer, P.O. *Natural Image Statistics*; Springer: Berlin, Germany, 2009.

37. Pouli, F.; Cunningham, D.W.; Reinhard, E. Image Statistics and their Applications in Computer Graphics. 2010. Aavilable online: https://s3.amazonaws.com/academia.edu.documents/44894836/Image_Statistics_and_their_Applications_20160419-16347-1hlzcca.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1523335119&Signature=52fMxbpkPhRptNNF0p6pfPLE58A%3D&response-content-disposition=inline%3B%20filename%3DImage_statistics_and_their_applications.pdf (accessed on 10 April 2018).

38. Spectral Database of CAVE Laboratory, Columbia University. Available online: http://www.cs.columbia.edu/CAVE/databases/multispectral/ (accessed on 1 February 2016).

39. Lukac, R.; Plataniotis, K.N. Color Filter Arrays: Design and Performance Analysis. *IEEE Trans. Consum. Electron.* **2005**, *51*, 1260–1267.

40. Miao, L.; Qi, H. The Design and Evaluation of a Generic Method for Generating Mosaicked Multispectral Filter Arrays. *IEEE Trans. Image Process.* **2006**, *15*, 2780–2791.

41. Takamatsu, J.; Matsushita, Y.; Ogasawara, T.; Ikeuchi, K. Estimating Demosaicing Algorithms using Image Noise Variance. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 279–286.

42. Shimano, N. Recovery of Spectral Reflectances of Objects Being Imaged Without Prior Knowledge. *IEEE Trans. Signal Process.* **2006**, *15*, 1848–1856.

43. Losson, O.; Macaire, L.; Yang, Y. Comparison of Color Demosaicing Methods. *Adv. Imag. Electron. Phys.* **2010**, *162*, 173–265.

44. Li, X.; Gunturk, B.; Zhang, L. Image demosaicing: A systematic survey. In *Visual Communications and Image Processing 2008*; SPIE: Bellingham, WA, USA, 2008; Volume 6822, p. 68221J.

45. Kennedy, J. Particle Swarm Optimization. In Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia, 27 November–1 December 1995; pp. 1942–1948.

46. Kennedy, J.; Eberhart, R.C. A discrete binary version of the particle swarm algorithm. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, Orlando, FL, USA, USA, 12–15 October 1997; Volume 5, pp. 4104–4108.

47. Spectral Database of Spectral Color Research Group, University of Eastern Finland. Available online: https://www.uef.fi/spectral/spectral-database (accessed on 1 February 2016).

48. Monno, Y.; Tanaka, M.; Okutomi, M. Multispectral demosaicking using adaptive kernel upsampling. In Proceedings of the 18th IEEE International Conference on Image Processing, Brussels, Belgium, 11–14 September 2011; pp. 3157–3160.

49. Monno, Y.; Kiku, D.; Kikuchi, S.; Tanaka, M.; Okutomi, M. Multispectral demosaicking with novel guide image generation and residual interpolation. In Proceedings of the IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 645–649.

50. Joshi, S.; Boyd, S. Sensor Selection via Convex Optimization. *IEEE Trans. Signal Process* **2009**, *57*, 451–462.