

# Optimum Spectral Sensitivity Functions for Single Sensor Color Imaging

Zahra Sadeghipoor<sup>a</sup>, Yue M. Lu<sup>b</sup>, and Sabine Süsstrunk<sup>a</sup>

<sup>a</sup>School of Computer and Communication Sciences

École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

<sup>b</sup>Harvard School of Engineering and Applied Sciences, Cambridge, MA 02138, USA

## ABSTRACT

A cost-effective and convenient approach for color imaging is to use a single sensor and mount a color filter array (CFA) in front of it, such that at each spatial position the scene information in only one color channel is captured. To estimate the missing colors at each pixel, a demosaicing algorithm is applied to the CFA samples. Besides the filter arrangement and the demosaicing method, the spectral sensitivity functions of the CFA filters considerably affect the quality of the demosaiced image. In this paper, we propose an algorithm to compute the optimum spectral sensitivities of filters in the single sensor imager. The proposed algorithm solves a constrained optimization problem to find optimum spectral sensitivities and the corresponding linear demosaicing method. An important constraint for this problem is the smoothness of spectral sensitivities, which is imposed by modeling these functions as a linear combination of several smooth kernels. Simulation results verify the effectiveness of the proposed algorithm in finding optimal spectral sensitivity functions, which outperform measured camera sensitivity functions.

**Keywords:** Single sensor imaging, spectral sensitivity functions, color filter array, demosaicing.

## 1. INTRODUCTION

One technique to capture three-channel RGB images is to use three sensors. In such a system, a beam-splitter is necessary to divide the incoming light into three paths, each captured at full resolution. Despite the accuracy of this method, using three sensors and a beam-splitter significantly increases both the production cost and the size of the imaging system. Thus, in consumer color cameras, a single sensor usually measures the visual scene information. To capture color information, a color filter array (CFA) is mounted in front of the sensor. Each filter in the CFA transmits a portion of visible radiation and blocks the rest. Thus in each pixel of the sensor output image, the value of only one color channel is known and others are missing. Because of this spatial color sampling, the output of the sensor is referred to as the mosaiced image. To estimate missing color information in each pixel, a post processing algorithm, called “demosaicing”, is then applied to the mosaiced image. Although the quality of the output is not as high as the three sensor imager output, its low cost and smaller size have motivated many researchers to study design issues of a single sensor camera.

The most popular CFA in color imaging is the Bayer pattern,<sup>1</sup> shown in Fig. 1. Many demosaicing algorithms assume that the camera uses the Bayer CFA, in which the sampling rate of the green channel is twice that of the other two colors. Therefore, the general framework of several methods<sup>2-5</sup> is to first reconstruct the green channel, since it is sampled with higher frequency and thus easier to accurately reconstruct as compared to the blue and red information. The less aliased high frequency components of green are used to increase the estimation quality of high frequencies in other channels.

Besides demosaicing, designing the optimum arrangement of filters in the CFA and sampling frequencies of different channels have also been addressed. In reference 6, Alleysson *et al.* studied the frequency representation of the mosaiced image. They proposed a frequency selective demosaicing technique. They showed that if in the Bayer CFA, the sampling frequency of the green channel is decreased from 1/2 to 1/4, linear demosaicing

---

Further author information: (Send correspondence to Zahra Sadeghipoor)

E-mail: zahra.sadeghipoor@epfl.ch, Telephone: +41 21 693 13 28

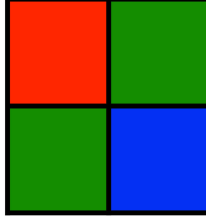


Figure 1. The Bayer CFA.

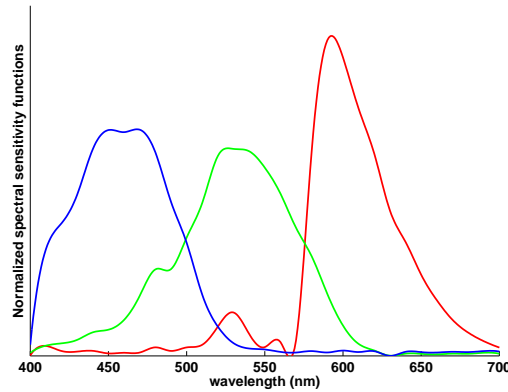


Figure 2. The spectral sensitivity functions measured for a consumer camera.

would achieve higher reconstruction accuracy. Hirakawa and Wolfe<sup>7</sup> also followed the idea of examining the mosaicing and demosaicing processes in the frequency domain and proposed a frequency based approach to find the optimum arrangement of filters in the CFA. They assumed that each pixel of the CFA is a convex combination of three primary colors (red, green, and blue) and found the coefficients of this combination such that the overlap between the modulated chrominance and baseline spatial information in the frequency domain is minimized.

The demosaicing algorithm and the arrangement of filters in the CFA are not the only factors that affect the quality of acquisition in a single sensor camera. As Alleysson *et al.* showed in reference 8, the system spectral response, which is determined by the sensor sensitivity and transmittance of CFA filters, also affects the spatial reconstruction and the color reproduction quality of demosaiced images. Figure 2 illustrates the overall spectral responses of a sensor and three color filters in a consumer camera. In this paper, we refer to the system response simply as the filter spectral sensitivity functions. In reference 8 it has been explained that increasing the correlation between the responses of different filters in the CFA results in more accurate spatial reconstruction, while it degrades the color reproduction acuity. In the extreme case, when spectral sensitivities fully overlap, the camera would be a black and white imager with no color reproduction. Alleysson *et al.* modeled the spectral sensitivities as Gaussian functions and empirically found the parameters of these functions (mean and bandwidth) such that the designed spectral responses provide the best solution to the trade-off between color reproduction and spatial reconstruction accuracy.

Some other studies have also addressed designing the spectral responses. However, most of them do not consider the single sensor imaging scenario.<sup>9-11</sup> They investigate designing spectral sensitivities only in terms of high color reproduction quality. In fact, they assume that three sensors are used to acquire the color image, therefore demosaicing artifacts are not taken into account in their algorithms. Parmar and Reeves<sup>12</sup> presented an algorithm that designs spectral sensitivities of filters in the CFA, when a single sensor is used to capture images. They consider a specific arrangement of filters in the CFA (for instance the Bayer pattern) and find the appropriate spectral sensitivity functions for this pattern by maximizing the demosaicing quality.

In this paper, we introduce an optimization procedure that jointly designs spectral sensitivities of CFA

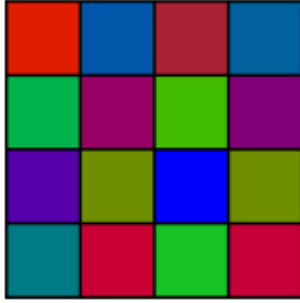


Figure 3. The CFA designed by the optimum CFA algorithm.<sup>13</sup>

filters and linear demosaicing, which result in the highest spatial reconstruction and color reproduction quality. The proposed algorithm extends the work of Lu and Vetterli,<sup>13</sup> where they presented a method to find the optimum arrangement of filters in the CFA and the demosaicing. We refer to their method as the “optimum CFA” algorithm. In their paper, Lu and Vetterli assumed that the value of each pixel in the color filter array is a weighted average of three primary colors (red, green, and blue) and a linear demosaicing, which can be implemented as a matrix multiplication, reconstructs the full resolution color image. As proposed in their paper, the optimum weights for each CFA pixel and also the optimum demosaicing matrix are computed by minimizing the error of reconstructing full resolution color images. Figure 3 shows a CFA designed by the optimum CFA algorithm.

In this paper, we change the block-diagram used in<sup>13</sup> for designing CFA and demosaicing and consider a scheme that better models the single-sensor imaging pipeline. By using the new block-diagram, instead of modeling each filter as a weighted average of primary colors, we are able to find the optimum spectral sensitivities. The optimum values for spectral responses and the demosaicing matrix are computed by minimizing the estimation error. The goal of our algorithm is to design physically realizable filters, thus a smoothness constraint is imposed to the proposed optimization problem. To this end, we assume that each spectral sensitivity is modeled as a linear combination of some predefined smooth functions. The spectral sensitivities and the demosaicing algorithm designed by our method consistently show better sampling and reconstruction performance as compared against some consumer camera spectral responses.

In the next section, the imaging pipeline used to model the single sensor camera is explained. In this section, we also present the optimization problem that finds the spectral sensitivity functions and the demosaicing matrix. The smoothness condition of filters and how to impose it to the design process are explained in Section 3. Furthermore, this section discusses the optimization procedure. The experimental setup and simulation results, which verify the effectiveness of the proposed algorithm, are presented in Section 4. Finally Section 5 concludes the article.

## 2. DESIGN SPECTRAL SENSITIVITY FUNCTIONS

In this section, we present a framework for finding optimum spectral responses and demosaicing. To this end, we consider the block-diagram shown in Fig. 4 as a simple single-sensor imaging pipeline, when the input is the scene radiance and the output is a three channel color image.

According to a simple image formation model, the pixel intensity in the output of the sensor can be computed as the following:

$$C(i, j) = \int E(i, j, \lambda)R(i, j, \lambda)S(\lambda)d\lambda, \quad (1)$$

where  $E(i, j, \lambda)$  and  $R(i, j, \lambda)$  are, respectively, the light power and the object reflectance at spatial position  $(i, j)$  and wavelength  $\lambda$ .  $S(\lambda)$  is the sensor sensitivity at wavelength  $\lambda$ .

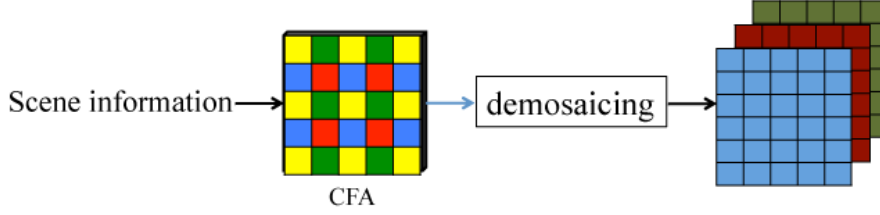


Figure 4. The single sensor imaging pipeline.

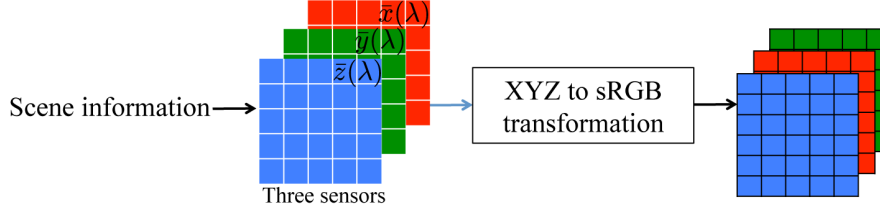


Figure 5. The three sensor imaging pipeline.

In single-sensor imaging, as shown in Fig. 4, only one filter is placed at each spatial position and the sensor output is a one channel image. Thus in this case, (1) can be re-written as the following:

$$\mathbf{Y}(i, j) = \int E(i, j, \lambda) R(i, j, \lambda) \phi_k(\lambda) d\lambda, \quad (2)$$

where  $\mathbf{Y}$  is a one channel mosaiced image and  $\phi_k$  is the sensitivity of the  $k^{th}$  filter in the CFA. In the above equation, the index of the filter ( $k$ ) depends on the spatial position ( $i, j$ ). For instance in Fig. 4, a yellow filter is placed in the first pixel, thus for  $i = j = 1$ ,  $\phi_k(\lambda)$  is the spectral response of this filter.

After forming the mosaiced image ( $\mathbf{Y}$ ), the demosaicing procedure is applied to reconstruct the full resolution image:

$$\hat{\mathbf{X}} = D(\mathbf{Y}) \quad (3)$$

where  $D(\cdot)$  is the demosaicing operator and  $\hat{\mathbf{X}}$  is the reconstructed color image.

Equations (2) and (3) show that in the single sensor imaging pipeline, the spectral sensitivity functions ( $\phi_k$ 's) and the demosaicing operator ( $D$ ) are the main factors that control the quality of the reconstructed image ( $\hat{\mathbf{X}}$ ). In this paper, we propose to design spectral responses of the CFA and demosaicing such that the reconstruction accuracy is maximized or, equivalently, the reconstruction error is minimized.

To minimize the reconstruction error, the difference between the reconstructed image ( $\hat{\mathbf{X}}$ ) and an ideal color image, which represents the same scene, needs to be computed. Note that artifacts in the reconstructed image are due to the spatial sampling behavior of the CFA. These artifacts can be avoided if three sensors (no spatial sampling in different color channels) are used to capture the color image. Thus, the ideal color image is the output of the three sensor imaging system. According to the image formation model (1), to compute the pixel intensities in the ideal image, spectral responses of three sensors are needed to be known. Here we assume that the main goal of imaging is to display images on a monitor, so the final image should be represented in the sRGB color encoding, which is the appropriate space for most displays.<sup>14</sup> The pipeline for forming the ideal color image is represented in Fig. 5. We model the sensitivities of the three sensors as the CIE-XYZ color matching functions.<sup>15</sup> As a result, the captured image is in XYZ space. Then a conventional XYZ to sRGB transformation<sup>14</sup> can be used to compute the sRGB representation of the image.

To be able to compare two images, the reconstructed image needs to be represented in the same color space as the ideal image. However, after demosaicing, the reconstructed image is still in the space given by the CFA's spectral sensitivities. Therefore, the demosaiced image is transformed to the sRGB color space as shown in

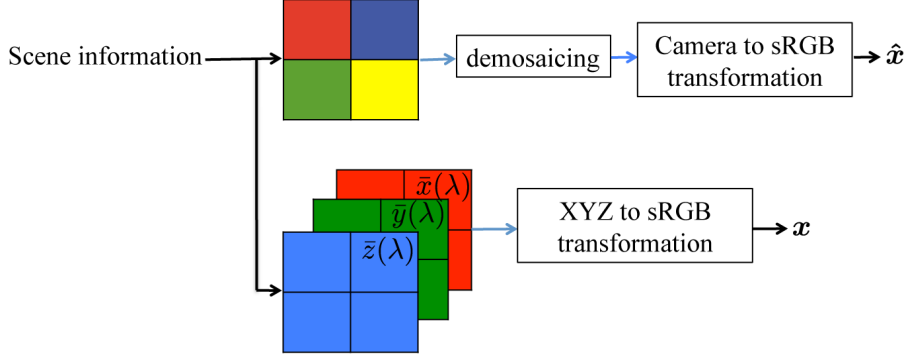


Figure 6. The block-diagram considered in the proposed optimization problem. The first branch shows the single sensor and the second illustrates the three sensor color imaging system.

Fig. 6. This figure is quite similar to the block-diagram used in,<sup>8</sup> where Alleysson *et al.* empirically found the suitable parameters for spectral sensitivities. Using the notations of Fig. 6, the reconstruction error energy can be written as:

$$e = E\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 \quad (4)$$

where  $E$  denotes the mathematical expectation.  $\mathbf{x}$  and  $\hat{\mathbf{x}}$  are the vector representation of the ideal and the demosaiced images, respectively.

To formulate the optimization problem, we approximate the image formation equation, shown in (2), by the following discrete model:

$$\mathbf{Y}(i, j) = \sum_{p=1}^M E(i, j, \lambda_p) R(i, j, \lambda_p) \phi_k(\lambda_p). \quad (5)$$

In the above equation, the scene information and the spectral sensitivity function are spectrally sampled at  $M$  different spectral bands.  $\lambda_p$  represents the  $p^{\text{th}}$  wavelength band along the electromagnetic spectrum. Thus, scene information can be approximated by a multispectral image with  $M$  channels.

Assume a block of size  $\sqrt{m} \times \sqrt{m}$  in the multispectral image. The spectral information of each pixel in this block can be represented in a column vector of size  $M$ . By stacking the vectors of all pixels in this block, the following vector can be obtained:

$$\mathbf{z} = [\mathbf{z}_1^T, \mathbf{z}_2^T, \dots, \mathbf{z}_m^T]^T, \quad (6)$$

where  $\mathbf{z}_i$  includes spectral information of the  $i^{\text{th}}$  pixel.

Similar to the multispectral information of the scene, a column vector of size  $M$ , called  $\phi_i$ , contains the values of  $i^{\text{th}}$  filter spectral sensitivity at  $M$  spectral bands. We assume that  $N$  filters form the CFA (in conventional color imaging,  $N = 3$ ), and the following matrix contains the spectral responses of all CFA filters:

$$\Phi_0^{N \times M} = [\phi_1^{M \times 1}, \phi_2^{M \times 1}, \dots, \phi_N^{M \times 1}]^T \quad (7)$$

where  $i^{\text{th}}$  row of  $\Phi_0$  is the spectral response of the  $i^{\text{th}}$  filter and superscripts denote dimensions.

According to the discrete image formation model in (5), the intensity of pixel  $i^{\text{th}}$  in the mosaiced image can be computed by the following matrix multiplication:

$$\mathbf{y}_i = \boldsymbol{\alpha}^{1 \times N} \Phi_0^{N \times M} \mathbf{z}_i^{M \times 1} \quad (8)$$

where  $\mathbf{y}_i$  is the  $i^{\text{th}}$  pixel in the mosaiced image and  $\mathbf{z}_i$  is the corresponding pixel in the multispectral image. In the above equation  $\boldsymbol{\alpha}$  is a simple selection vector, which selects the  $j^{\text{th}}$  filter in  $\Phi_0$ . The  $j^{\text{th}}$  filter is the filter located at the position of  $\mathbf{z}_i$ . For instance, in Fig. 6, if  $\mathbf{z}_i$  is the pixel located at the position of the red filter

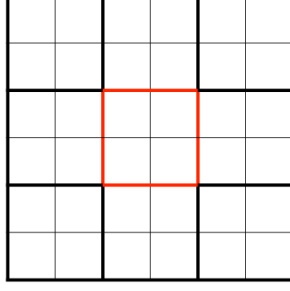


Figure 7. A block of size  $6 \times 6$  and its central block (with red borders) of size  $2 \times 2$  ( $\sqrt{m} = 6$  and  $\sqrt{n} = 2$ ).

(the first pixel in the CFA), then  $\alpha$  selects the spectral sensitivity of the red filter. In fact,  $\alpha$  models the spatial sampling of CFA.

To compute the block  $\mathbf{y}$  in the mosaiced image, which corresponds to a block of size  $\sqrt{m} \times \sqrt{m}$  in the multispectral image ( $\mathbf{z}$ ), the following matrix is formed:

$$\Phi^{mN \times mM} = \mathbf{I}^{m \times m} \otimes \Phi_0^{N \times M} \quad (9)$$

where  $\otimes$  denotes the Kronecker product and  $\mathbf{I}$  is the identity matrix. As a result, we can write:

$$\mathbf{y}^{m \times 1} = \mathbf{A}^{m \times mN} \Phi^{mN \times mM} \mathbf{z}^{mM \times 1} \quad (10)$$

Here,  $\mathbf{A}$  is a selection matrix containing only ones and zeros. This matrix at each spatial position extracts the information corresponding to the CFA filter in that pixel (similar to vector  $\alpha$  in (8)).

To reconstruct the full resolution  $N$  channel block in the demosaicing step, as proposed in reference 13, the corresponding block and also the intensities of neighboring pixels in the mosaiced image are used. According to the lowpass behavior of natural images, i.e. high correlation between pixel intensities in a local neighborhood, incorporating the extra neighboring information in demosaicing increases the quality of reconstruction. Here we assume that a block of size  $\sqrt{m} \times \sqrt{m}$  in the mosaiced image is used to demosaic its central block of size  $\sqrt{n} \times \sqrt{n}$  ( $n < m$ ). Figure 7 shows these blocks for the case of  $\sqrt{m} = 6$  and  $\sqrt{n} = 2$ . There is a trade-off in choosing the optimum value for the  $m/n$  ratio. If this ratio is very large, the intensities of the demosaiced image are affected by values in a large neighborhood in the mosaiced image. Consequently, it is possible that sharp edges in a small region are blurred. It also increases the computational complexity of demosaicing. On the other hand, if  $m/n$  is too small, the amount of information used in demosaicing might not be enough to have a high quality reconstruction.

We apply a linear demosaicing scheme to reconstruct the full resolution image. Demosaicing is one of the most computationally demanding steps in single-sensor imaging and it needs to be done in real time. Therefore, a linear operation, which can be implemented as an efficient matrix multiplication, can be an appropriate choice for demosaicing. By applying a linear demosaicing, a block of size  $\sqrt{n} \times \sqrt{n}$  in the  $N$  channel demosaiced image ( $N$  is the number of filters in the CFA) can be obtained as:

$$\hat{\mathbf{w}}^{nN \times 1} = \mathbf{D}^{nN \times m} \mathbf{y}^{m \times 1} \quad (11)$$

where  $\mathbf{y}$  and  $\hat{\mathbf{w}}$  are vector representations of a block in the sampled and reconstructed images, respectively, and  $\mathbf{D}$  is the demosaicing matrix.

The last block in the first branch of Fig. 6 transforms the reconstructed image from the  $N$ -dimensional camera space to sRGB. Here, to avoid high complexity in the optimization process, we only compute the linear sRGB values of the image, so each pixel intensity can be transformed to sRGB using the following linear operation:

$$\hat{\mathbf{x}}_i^{3 \times 1} = \mathbf{T}_0^{3 \times N} \hat{\mathbf{w}}_i^{N \times 1} \quad (12)$$

where  $\hat{\mathbf{w}}_i$  and  $\hat{\mathbf{x}}_i$  contain the pixel information in the  $N$ -dimensional camera and three-dimensional sRGB spaces, respectively. Matrix  $\mathbf{T}_0$  is obtained by concatenating a  $3 \times N$  matrix, which maps the camera to the standard XYZ space, and the  $3 \times 3$  XYZ to linear sRGB transformation matrix. To change the space for a block of size  $n$ , matrix  $\mathbf{T} = \mathbf{I}^{n \times n} \otimes \mathbf{T}_0$  can be used as follows:

$$\hat{\mathbf{x}}^{3n \times 1} = \mathbf{T}^{3n \times Nn} \hat{\mathbf{w}}^{Nn \times 1} \quad (13)$$

Combining (10), (11), and (12), the output of the first branch in Fig. 6 can be computed by the following matrix multiplications:

$$\hat{\mathbf{x}} = \mathbf{TDA}\Phi\mathbf{z}, \quad (14)$$

where  $\mathbf{z}$  is the vector representation for a multispectral block of size  $\sqrt{m} \times \sqrt{m}$ ,  $\Phi$  is the matrix of filter spectral sensitivities,  $\mathbf{A}$  is the CFA selection matrix,  $\mathbf{D}$  is the demosaicing matrix, and  $\mathbf{T}$  is the camera to sRGB transformation matrix.

Using the image formation model formulated in (5), a block of size  $\sqrt{m} \times \sqrt{m}$  in the multispectral image can be used to compute the pixel intensities in its central block (with size  $\sqrt{n} \times \sqrt{n}$ ) in the output of the three-sensor imager (second branch of Fig. 6) as the following:

$$\mathbf{x} = \mathbf{MF}_{XYZ}\mathbf{S}\mathbf{z} \quad (15)$$

where  $\mathbf{S}$  is a selection matrix that extracts the central block of  $\mathbf{z}$ .  $\mathbf{F}_{XYZ}$  is a matrix containing the CIE-XYZ color matching functions, which maps the multispectral data to XYZ space, and  $\mathbf{M}$  is the XYZ to linear sRGB transformation matrix.

By computing the reconstructed image ( $\hat{\mathbf{x}}$ ), and the ideal color image ( $\mathbf{x}$ ), the reconstruction error can be written as:

$$e = \text{E} \|\mathbf{MF}_{XYZ}\mathbf{S}\mathbf{z} - \mathbf{TDA}\Phi\mathbf{z}\|_2^2 \quad (16)$$

It can be easily proved, as it has been shown in,<sup>13</sup> that the reconstruction error can be written in the following way:

$$e = \|\mathbf{MF}_{XYZ}\mathbf{S}\mathbf{P} - \mathbf{TDA}\Phi\mathbf{P}\|_F^2. \quad (17)$$

In the above equation,  $\mathbf{P}$  is the square root of the multispectral images' correlation matrix ( $\mathbf{C}$ ), i.e.  $\mathbf{C} = \mathbf{P}\mathbf{P}$ .

Then, sensitivity functions and the demosaicing matrix are designed by solving the following optimization problem:

$$\{\Phi_0^*, \mathbf{D}^*\} = \underset{\Phi_0, \mathbf{D}}{\text{argmin}} \|\mathbf{MF}_{XYZ}\mathbf{S}\mathbf{P} - \mathbf{TDA}\Phi\mathbf{P}\|_F^2 \quad (18)$$

where  $\Phi = \mathbf{I}^{m \times m} \otimes \Phi_0$  and  $\Phi_0^*, \mathbf{D}^*$  are optimum values for these matrices.

### 3. PHYSICALLY REALIZABLE SPECTRAL SENSITIVITIES

Since no constraint is imposed on the proposed optimization problem (18), the resultant spectral sensitivities might contain sudden changes and sharp transitions. However, practical filters are smooth, and a filter with sharp transitions is not physically realizable. Because of this practical concern in designing the CFA filters, it is essential to impose the smoothness of spectral responses on problem (18).

To ensure that the optimum sensitivities found by solving (18) are smooth, we assume that the response of each filter can be modeled as a linear combination of several positive smooth functions:

$$\phi_i(\lambda) = \sum_{j=1}^K \beta_{ij} \mathbf{b}_j(\lambda), \quad \beta_{ij} \geq 0 \quad (19)$$

where  $\phi_i$  is the  $i^{\text{th}}$  filter response.  $\mathbf{b}_j$ 's are  $K$  smooth functions and  $\beta_{ij}$ 's are coefficients of the linear combination. By modeling the sensitivities as shown in the above equation, to find matrix  $\Phi$ , optimum coefficients of linear combination ( $\beta_{ij}$ ) are computed in the optimization procedure. The function  $\mathbf{b}$  in (19) can be any smooth

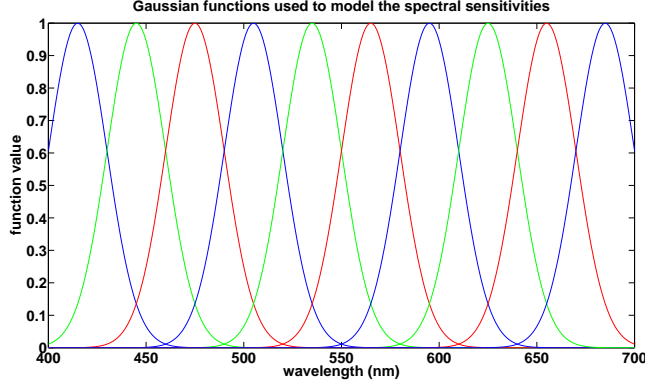


Figure 8. The smooth Gaussian functions used to model the spectral sensitivities of CFA filters.

positive kernel. For instance, it is possible to use Gaussian functions, shown in Fig. 8, to model the spectral sensitivities.

Increasing the number of smooth functions ( $K$ ) in (19) enlarges the search space and increases degrees of freedom in the optimization problem. Thus, theoretically more accurate results can be obtained by solving the problem. On the other hand, the complexity of the proposed algorithm increases as  $K$  increases. Including one more smooth function in (19) means that  $N$  (the number of CFA filters) more coefficients need to be computed.

Let us assume that  $\mathbf{B}$  is the matrix of smooth functions where each row includes values for one smooth function, i.e.,:

$$\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_K]^T. \quad (20)$$

Here,  $\mathbf{b}_j$  is a column vector, which represents the  $j^{\text{th}}$  smooth function. The matrix  $\beta$  is formed by all linear combination coefficients ( $\beta_{ij}$ ). Using these notations and (19), the matrix of spectral responses can be written as:

$$\Phi_0^{N \times M} = \beta^{N \times K} \mathbf{B}^{K \times M} \quad (21)$$

Substituting (21) in (18) results in the following optimization problem:

$$\{\beta^*, \mathbf{D}^*\} = \underset{\beta, \mathbf{D}}{\operatorname{argmin}} \|\mathbf{M}\mathbf{F}_{\text{XYZ}}\mathbf{S}\mathbf{P} - \mathbf{T}\mathbf{D}\mathbf{A}\Phi\mathbf{P}\|_F^2 \quad \text{s.t. } \Phi = \mathbf{I}^{n \times n} \otimes \Phi_0, \Phi_0 = \beta\mathbf{B}. \quad (22)$$

To solve this optimization problem, we use the following iterative minimization approach, which also was used by Lu and Vetterli in<sup>13</sup> for their specific optimization problem:

1. Initialization

- $k = 0$ .
- Assign random value to matrix  $\beta$  (matrix of linear combination coefficients) and call it  $\beta^{(0)}$ .
- Assign an arbitrary large value to  $e^{(0)}$ .

2. Compute  $\mathbf{D}$  as

$$\mathbf{D}^{(k)} = \mathbf{T}^\dagger \mathbf{M}\mathbf{F}_{\text{XYZ}}\mathbf{S}\mathbf{P}(\mathbf{A}\Phi^{(k)}\mathbf{P})^\dagger, \quad (23)$$

which is the solution of problem (22) when  $\beta$  is fixed.  $(\cdot)^\dagger$  in the above equation denotes the Moore-Penrose pseudo-inverse of a matrix and  $\Phi^{(k)} = \mathbf{I} \otimes (\beta^{(k)}\mathbf{B})$ .

3. Solve the constrained quadratic optimization problem (22) assuming that the value of  $\mathbf{D}$  is  $\mathbf{D}^{(k)}$  and update  $\beta$ . Call the solution  $\beta^{(k+1)}$ .



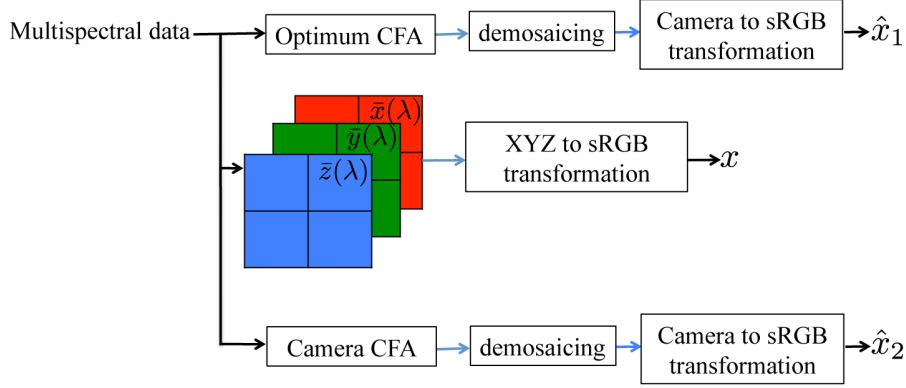


Figure 9. The block-diagram used to investigate the performance of the optimum sensitivity functions. First branch: single-sensor imager with optimum sensitivities, second branch: three sensor imager, and last branch: single sensor imager with camera sensitivity functions.

4.  $k = k + 1$ .
5. Compute  $e^{(k)}$  using (17).
6. If  $|e^{(k)} - e^{(k-1)}| > \delta$ , go back to step 2, otherwise stop.

In this procedure  $\delta$  is a constant, which defines the stop criterion of optimization. Clearly, both computational complexity and accuracy of the solution are inversely proportional to  $\delta$ .

Note that to solve (22), we need to compute the correlation matrix of the multispectral images ( $\mathbf{C} = \mathbf{E} \{ \mathbf{z} \mathbf{z}^T \}$ ) and its square root ( $\mathbf{P}$ ). In practice, the correlation matrix is approximated by the following ensemble summation:

$$\mathbf{C} = \sum \mathbf{z} \mathbf{z}^T. \quad (24)$$

To compute  $\mathbf{C}$  and solve the optimization problem we use a set of multispectral images as training.

#### 4. EXPERIMENTAL RESULTS

In this section, we examine the performance of spectral sensitivities and demosaicing designed by our algorithm in single sensor imaging. The performance of the optimum CFA is compared against some spectral sensitivities measured for consumer cameras.

To this end, we consider the block-diagram shown in Fig. 9. In this block-diagram both first and third branches show parts of the single sensor imaging pipeline, and the second branch is the three sensor color camera. The main difference between the first and third pipelines is that the *optimum* sensitivities are used in the first branch, while in the third one the camera CFA samples the multispectral image. We can evaluate the performance of optimum sensitivities and the demosaicing matrix by computing the difference between  $\hat{\mathbf{x}}_1$  (output of the first branch) and  $\mathbf{x}$  (three sensor camera output) as follows:

$$\text{CPSNR}_{\text{opt}} = 10 \log \frac{255^2}{\frac{1}{3Q} \|\hat{\mathbf{x}}_1 - \mathbf{x}\|_2^2}, \quad (25)$$

where  $Q$  is the number of pixels in each image. Similarly for the real camera we have:

$$\text{CPSNR}_{\text{cam}} = 10 \log \frac{255^2}{\frac{1}{3Q} \|\hat{\mathbf{x}}_2 - \mathbf{x}\|_2^2}, \quad (26)$$



Figure 10. Color representations of some images in the multispectral dataset presented in<sup>16</sup> and used in our experiments.

In all experiments, we used a multispectral dataset captured by Chakrabarti and Zickler.<sup>16</sup> This dataset contains 50 multispectral images with 31 channels ( $\lambda = 400 : 10 : 700$ ). Figure 10 shows color representations of some images in this dataset.

The dataset is randomly split into training and test sets, each set containing 25 images. The training set is used to find optimum sensitivities and the demosaicing matrix as explained in Section 3. For the case of real single sensor camera (third branch of Fig. 9), the correlation matrix of the training set is used to compute the optimum linear demosaicing, which corresponds to camera sensitivity functions. To compute this matrix, equation (23) is evaluated when  $\Phi$  is replaced with camera spectral responses. After optimization, CFA spectral sensitivities and demosaicing matrices are applied to multispectral data. To remove the possible bias in comparison, due to specific training and test sets, both optimum and camera sensitivities are tested with 20 different sets of training and test data. In each experiment, both sets are chosen randomly.

We compare the performance of optimum sensitivities against three cameras, called *cam1*, *cam2*, and *cam3*. The spectral sensitivity functions of these cameras are shown in Fig. 11.

In the first set of experiments, we assume that the number of filters in the CFA is three ( $N = 3$ ). The pattern shown in Fig. 12-a is used to arrange filters. The Gaussian functions shown in Fig. 8 are the smooth kernels that model the filter spectral responses. The optimum spectral sensitivities, designed by our algorithm for this case, are shown in Fig. 13.

Table 1 summarizes the average CPSNR values, which show the performance of optimum and camera sensitivity functions. In this table, the similarity between reconstructed *linear* sRGB and the ideal sRGB images are computed. It can be seen that our sensitivity functions perform better than those of the three commercial cameras.

We also studied the quality of reconstructed *nonlinear* sRGB images. Results of this study are shown in Table 2. Although the proposed optimization problem only minimizes the reconstruction error of linear sRGB images, the optimum spectral sensitivities still outperform camera sensitivity functions in reconstructing nonlinear images.

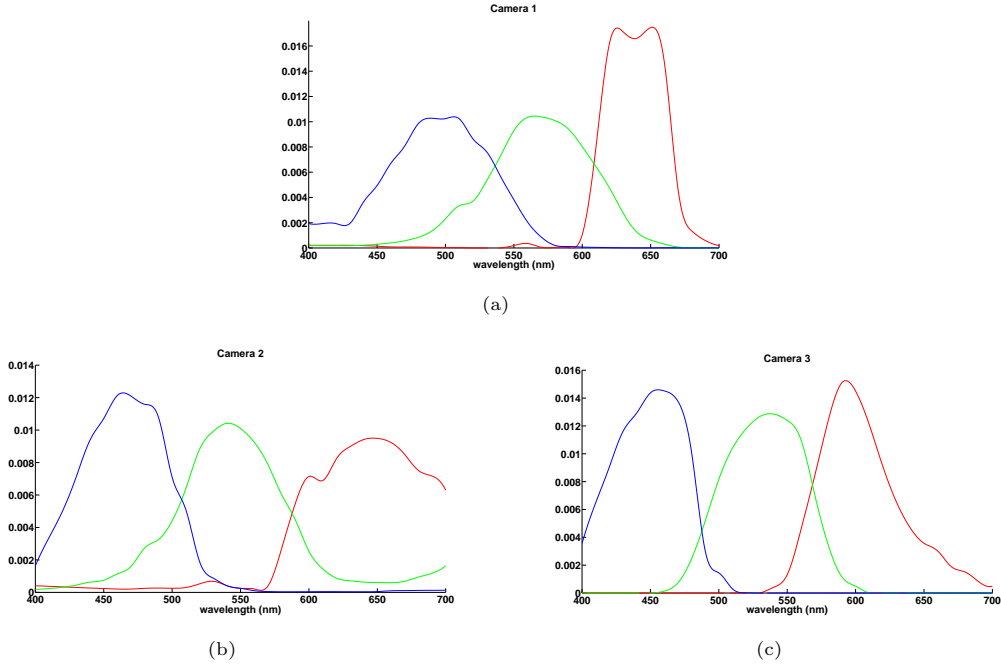


Figure 11. The camera spectral sensitivity functions used in our experiment. (a) *cam1*, (b) *cam2*, and (c) *cam3*.

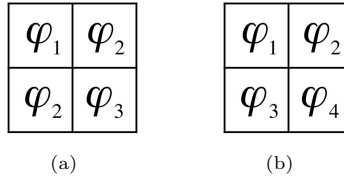


Figure 12. The filter arrangements used in optimizing (a) three and (b) four spectral sensitivities for CFA filters.

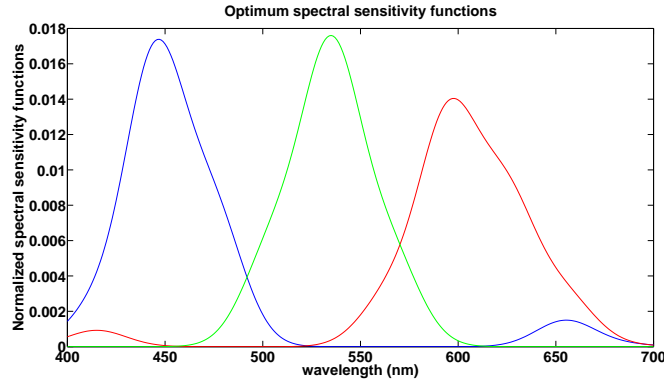


Figure 13. Three optimum sensitivity functions designed by the proposed algorithm.

	Optimum	<i>cam1</i>	<i>cam2</i>	<i>cam3</i>
average	35.01	34.58	34.10	34.28
std.	0.52	0.52	0.44	0.48

Table 1. The CPSNR demosaicing results of optimum sensitivity functions and camera sensitivities for linear sRGB images. Each reported CPSNR is the average over 20 runs of the algorithm with different train and test. The last row shows standard deviations over different runs.

	Optimum	<i>cam1</i>	<i>cam2</i>	<i>cam3</i>
average	32.84	31.78	32.30	32.48
std.	0.32	0.34	0.32	0.37

Table 2. The CPSNR demosaicing results of optimum sensitivity functions and camera sensitivities for nonlinear sRGB images. Each reported CPSNR is the average over 20 runs of the algorithm with different train and test. The last row shows standard deviations over 20 different runs.

To visually compare the results, an image directly rendered to sRGB (the second branch of Fig. 9) and outputs of camera and optimum sensitivities are shown in Fig. 14. For this image, sensitivities designed with our algorithm are more successful in color reproduction (see the color of the purple car in the bottom right of the image). Additionally, less false colors (demosaicing artifacts) are visible in the result of the proposed method.

We also used the proposed optimization problem to design four different filters for the CFA ( $N = 4$ ). In this case, the arrangement of Fig. 12-b is considered. The resulting sensitivity functions are illustrated in Fig. 15. As can be seen in this figure, the optimum sensitivities of two filters, located in the red part of the spectrum, are significantly similar. This observation suggests that for color imaging, when the goal is to represent images in the sRGB space, using three filters in the CFA is optimum. Another conclusion drawn from this result is that the sampling rate of information in the red part of the spectrum (long wavelengths) is twice the rate of the blue and green channels. However, in the Bayer pattern, the green channel has the highest sampling rate. Alleysson *et al.*<sup>6</sup> showed that decreasing the green sampling frequency from 1/2 to 1/4 results in higher reconstruction quality when their frequency selection algorithm is used for demosaicing. In this paper, we came to the same conclusion by optimizing the spectral sensitivity functions.

## 5. CONCLUSION

We proposed an iterative algorithm to find the optimum shape of spectral sensitivities for a CFA and the optimum demosaicing matrix. The proposed algorithm solves a constrained optimization problem to compute the optimum spectral sensitivity functions. The main constraint imposed is that the optimum sensitivities need to be smooth. Experimental results proved the effectiveness of our algorithm. Using the optimum spectral sensitivities results in higher reconstruction quality in color acquisition as compared to using three different sets of measured camera sensitivity functions.

## Acknowledgments

This work is supported by the Swiss National Science Foundation under the grant number 200021-124796/1 and the Xerox Foundation.

## REFERENCES

1. Bayer, B. E., “Color imaging array,” (1976). U.S. Patent 3971065.
2. Hirakawa, K. and Parks, T. W., “Adaptive homogeneity-directed demosaicing algorithm,” *IEEE Transactions on Image Processing* **14**(3) (2005).
3. Gunturk, B. K., Altunbasak, Y., and Mersereau, R. M., “Color plane interpolation using alternating projections,” *IEEE Transactions on Image Processing* **11**(9) (2002).
4. Chang, L. and Tan, Y., “Effective use of spatial and spectral correlations for color filter array demosaicking,” *IEEE Transactions on Consumer Electronics* **50**(1) (2004).
5. Lu, W. and Tan, Y., “Color filter array demosaicking: New method and performance measures,” *IEEE Transactions on Image Processing* **12**(10) (2003).
6. Alleysson, D., Süsstrunk, S., and Héroult, J., “Linear demosaicing inspired by the human visual system,” *IEEE Transactions on Image Processing* **14**(4), 439–449 (2005).
7. Hirakawa, K. and Wolfe, P., “Spatio-spectral color filter array design for optimal image recovery,” *IEEE Transactions on Image Processing* **17**(10) (2008).



(a)



(b)



(c)



(d)



(e)

Figure 14. Visual comparison of optimum spectral sensitivities and different cameras results. (a) The image directly rendered to the sRGB space, the results of (b) optimum sensitivity functions, (c) *cam1*, (d) *cam2*, and (e) *cam3* sensitivities.

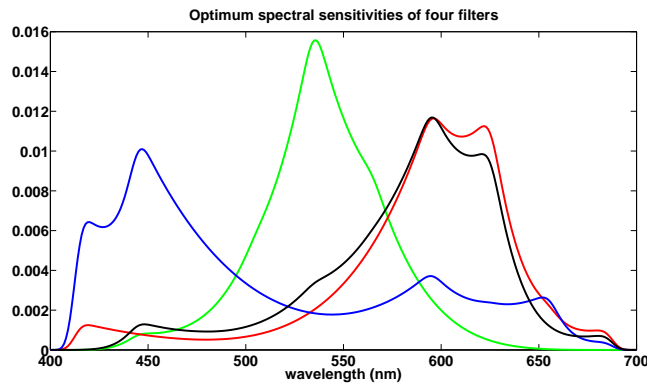


Figure 15. Four optimum sensitivity functions designed by the proposed algorithm.

8. Alleysson, D., Süsstrunk, S., and Marguier, J., "Influence of spectral sensitivity functions on color demosaicing," in [*Proceedings of IS&T/SID 11th Color Imaging Conference*], **11**, 351–357 (2003).
9. Vora, P. and Trussell, H., "Mathematical methods for the analysis of color scanning filters," *IEEE Transactions on Image Processing* **6**(2) (1997).
10. Kuniba, H. and Berns, R. S., "Spectral sensitivity optimization of color image sensors considering photon shot noise," *Journal of Electronic Imaging* **18**(2) (2009).
11. Yasuma, F., Mitsunaga, T., Iso, D., and Nayar, S. K., "Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum," *IEEE Transactions on Image Processing* **19**(9) (2010).
12. Parmar, M. and Reeves, S. J., "Selection of optimal spectral sensitivity functions for color filter arrays," *IEEE Transactions on Image Processing* **19**(12) (2010).
13. Lu, Y. M. and Vetterli, M., "Optimal color filter array design: Quantitative conditions and an efficient search procedure," in [*Proceedings of IS&T/SPIE EI: Digital Photography V*], (2009).
14. *IEC 61966 2-1:1999. Multimedia systems and equipment - color measurement and management- Part 2-1: color management-default RGB color space - sRGB* (1999).
15. Wyszecki, G. and Stiles, W., [*Color Science: Concepts and Methods, Quantitative Data and Formulas*], Wiley (1982).
16. Chakrabarti, A. and Zickler, T., "Statistics of real-world hyperspectral images," in [*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*], (2011).