Article

Optimized Multi-Spectral Filter Arrays for Spectral Reconstruction

Renjie Wu [1], Yuqi Li [1,*], Xijiong Xie [1] and Zhijie Lin [2]

[1] College of Information Science and Engineering, Ningbo University, Ningbo 315000, China
[2] School of Information and Electronic Engineering, Zhejiang University of Science and Technology, Hangzhou 310000, China
* Correspondence: liyuqi1@nbu.edu.cn

check for updates

Abstract: Multispectral filter array (MSFA)-based imaging is a compact, practical technique for snapshot spectral image capturing and reconstruction. The imaging and reconstruction quality is highly influenced by the spectral sensitivities and spatial arrangement of channels on MSFAs, and the used reconstruction method. In order to design a MSFA with high imaging capacity, we propose a sparse representation based approach to optimize spectral sensitivities and spatial arrangement of MSFAs. The proposed approach first overall models the various errors associated with spectral reconstruction, and then uses a global heuristic searching method to optimize MSFAs via minimizing the estimated error of MSFAs. Our MSFA optimization method can select filters from off-the-shelf candidate filter sets while assigning the selected filters to the designed MSFA. Experimental results on three datasets show that the proposed method is more efficient, flexible, and can design MSFAs with lower spectral construction errors when compared with existing state-of-the-art methods. The MSFAs designed by our method show better performance than others even using different spectral reconstruction methods.

Keywords: filter array design; channel selection; spectral reconstruction

1. Introduction

Multi-spectral images are widely used in many applications such as material classification, object recognition, and color constancy. In recent decades, many snapshot spectral imaging techniques [1] have been proposed. Among these techniques, multi-spectral filter array (MSFA)—based imaging is a compact, convenient, and robust way to acquire and reconstruct spectral images. The method using a camera with mosaic of the multiple filters on the sensor has become commercially available, e.g., SILIOS technologies developed a manufacturing technique called COLOR SHADES® to produce MSFAs [2].

However, most of the existing MSFAs are usually designed in an ad-hoc manner, e.g., the spectral sensitivity function of channels are designed as a series of band-pass functions with a close full width at half maximum (FWHM) and the same intervals. In fact, to design an optimal MSFA, the influence of spectral sensitivity functions of channels, channel spatial distribution, and reconstruction method are mutually correlated [3]. Previous works [4–7] optimize either (a) the spectral sensitivity functions of channels or (b) the arrangement of channels on periodic mosaic pattern. These methods treat the two issues separately thus cannot guarantee the optimal MSFA design. Besides, the MSFAs design methods do not take advantage of multispectral image statistics thus cannot assure an accurate spectral recovery. In addition, previous works optimize spectral sensitivity functions of channels based on theoretical transmittance function (e.g., radial basis function [4], Fourier basis function [8]). However, due to the limitation of manufacturing, MSFAs confirming to such theoretically perfect functions or

arbitrary manually set functions are usually not available in practice [9]. Therefore, compared with theoretical optimization, selecting a few filters from a candidate commercial filter set is more practical for MSFA design.

It is worth noting that the MSFA design also depends on the applied spectral reconstruction method. Joint design of a MSFA and a spectral reconstruction algorithm is potentially effective. Here spectral reconstruction is an operation that recovering spectrum at the position of each pixel from incomplete samples output from a MSFA sensor. The most widely applied solution first obtains the full resolution image of multiple channels through multispectral demosaicing [10], and then recovers spectral information of each pixel using the obtained multiple responses [11–13]. Recently, some works have demonstrated the great capability of sparse coding either for color demosaicing [14,15] or for spectral recovery [16]. However, the two-step spectral reconstruction may introduce propagation error. Thus reconstruction quality can be further improved using an overall sparse representation method [17,18].

Based on these considerations, we propose a sparse representation overall model to estimating reconstruction errors with prior knowledge. We firstly introduced a sparse coding method to directly recover multispectral images from raw outputs of MSFA sensors; then we estimate and qualify the reconstruction errors of the spectral reconstruction method to evaluate the capability of MSFA sensors; finally, the optimal MSFA can be achieved by (1) selecting a subset from a large candidate set of commercial filters; (2) assigning the selected filters onto MSFA mosaic pattern that minimizes the estimated reconstruction errors using a heuristic global searching method.

2. Related Works

In recent decades, many efforts have been made to study training based spectral image reconstruction and filter optimization for multispectral cameras. As shown in Figure 1, these works mainly focus on the following issues:
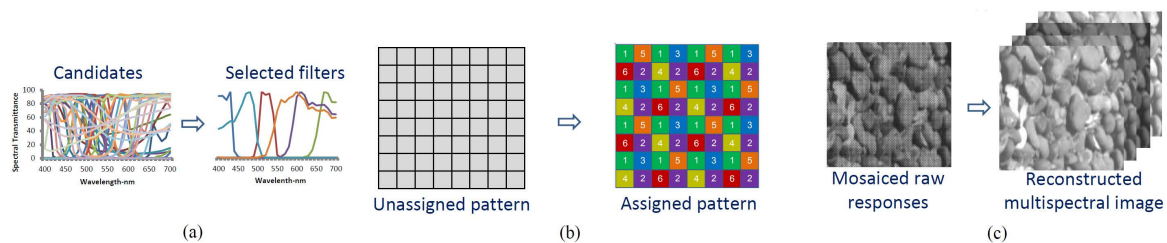
Figure 1. Three mutually dependent issues in MSFA design. (**a**) Selecting the used filters from commercial candidate filter set; (**b**) Assigning the selected channels on Multispectral filter array (MSFA) pattern; (**c**) Spectral reconstruction from mosaiced raw response images to reconstructed multispectral images.

2.1. Spectral Reconstruction

Spectral images can be recovered from multiple-channel raw response via learning a mapping between the spectra and the corresponding responses for a given multispectral camera. The mapping is learned from a training set consisting of the sample pairs of "spectra-responses" images. Some works pose spectral reconstruction as a traditional high dimensional interpolation problem and solve it using pseudo-inverse method [19], locally linear regression [20] or nonlinear radial-basis-function approximation [21,22].

Recently, other ways of reconstructing spectral images from raw response images (usually RGB images) have been explored, including sparse reconstruction and deep learning. Sparse reconstruction methods exploit the sparsity of spectra, recover spectra for every pixels using an learnt overcomplete dictionary [23–25]. Based on sparse reconstruction, some literature has proved that the reconstruction accuracy would benefit from combining the prior of local manifold structure [16,26]. To further improve reconstruction quality, the convolutional neural network based learning methods which utilize the spatial information of images have begun to appear [27,28]. Various network structures have been proposed for a contest to show the potential of deep learning in spectral construction [29]. However,

none of these spectral reconstruction methods assume the raw response images are mosaicked, thus cannot be applied to our problem. A more recent literature [30] utilized random printed masks as filter arrays and reconstructed multispectral images using an end-to-end network.

2.2. Spectral Filters Selection

Filter selection is to select a few used filters from a candidate filter set while ensuring reconstruction accuracy. Imai et al. compared the behaviours of narrow-band and wide-band filters on spectral reflectance estimation [31]. Chi et al. minimized the condition number of spectral sensitivity of filters to make the spectral construction more robust to noises [7]. Other methods take both the prior of spectral information and imaging noises into account, e.g., Ansari et al. [32] used a Wiener filter estimation method to estimate reconstruction error of Munsell spectra; Wang et al. [33] presented a model that estimates the spectral reconstruction error of multi-spectral imaging system using the prior of spectral correlation and discussed the tradeoff between choosing narrow filters and wide filters; Shen et al. [34] also estimated the reconstruction error and applied binary differential evolution algorithm to find the optimal combination of filters. Instead of estimating reconstruction error, other methods used brute-force strategy that recovering all spectral images in training set using possible filters combination and calculating the real total reconstruction error [4,35]. However, compared with error estimation, actual calculation is much more time-consuming, especially when the number of possible filters combination is too large. In fact, filter selection is an NP-hard problem thus inefficient evaluation of solutions would make the selection unacceptable and impractical. Taking advantage of a pre-trained deep spectral reconstruction network, Fu et al. [36] used a Lasso-based channel pruning algorithm to select filters, but their method can only select filters for specific filter arrays.

2.3. MSFA Pattern Design

Previous works [37,38] have shown that a good channel spatial arrangement requires satisfying three properties: (i) channels should be spatially uniform to ensure robustness against image sensor defects; (ii) channel arrangement should be regular to ensure image reconstruction efficiency; (iii) channels should be neighbor-consistent (each channel has the same neighbor channel) to avoid the inconsistent optical crosstalk between adjacent pixels. The most popular design is the binary tree-based MFSA proposed by Miao et al. [38], the design strategy ensures all the above properties and is widely applied in real imaging systems [6,39]. However, the channel spatial arrangement of MSFAs are set manually thus cannot guarantee optimal design. Besides, the MSFA pattern must be designed with the filter selection and the used spectral reconstruction algorithm to achieve optimal imaging capacity.

2.4. Overall Design

The overall design takes all of the above factors into consideration. In the area of color imaging, Henz et al. [40] proposed an overall color filter array optimization method using deep convolutional neural networks. Similarly, Nie et al. [41] jointly optimized filters and recovered spectral images by optimizing the weights in CNN. Li et al. [42] adopted sparse representation to model the pipeline of color imaging and demosaicking, and then optimized filters arrays via minimizing mutual coherence. However, all of these continuous optimizations are not suitable for the discrete combination problem involved filter selection and channel arrangement. In contrast, other overall design methods [3,43] applied heuristic global search algorithm to find a near-optimal MSFA. Motivated by these overall design methods, we model the total errors in the pipeline of imaging and reconstruction by considering the spectral sensitivities of channels, the spatial arrangement of channels, and the statistical prior of spectral images as a whole.

3. Spectral Reconstruction Using Sparse Coding

Assume that a scene is captured by a camera with a c-channel MSFA sensor, a spectral image patch of the captured scene is denoted as S. Let the resolution of S be $w \times h \times l$, where h and w is the

vertical and horizontal spatial resolution respectively, and l is the spectral resolution. The model of spectral imaging with the MSFA is as follows:

$$X(u,v) = \sum_{i=1}^{c} m(u,v,i) \sum_{j=1}^{l} f(i,\lambda_j) S(u,v,\lambda_j) \tag{1}$$

where $X(u,v)$ denotes the response at spatial coordinate (u,v) on the response image patch, $S(u,v,\lambda)$ denotes the spectral power distribution (SPD) function of the spectrum at spatial coordinate (u,v), $f(i,\lambda)$ denotes the spectral sensitivity function of the $i(i=0,1,...,c)$ th channel. The mask function $m(u,v,i)$ with value 1 indicates the position (u,v) is assigned to the ith channel, while those with value 0 indicates the position is assigned to other channels.

The above equation can be written into matrix form:

$$\mathbf{x} = \mathbf{\Phi}\, \mathbf{s} \tag{2}$$

where $\mathbf{s}$ denotes the $whl \times 1$ vector vectorized from the spectral image patch S, $\mathbf{x}$ denotes the $wh \times 1$ vector vectorized from the response image patch X, $\mathbf{\Phi}$ denotes the $wh \times whl$ matrix simulating the imaging process from spectral images to response images. The matrix $\mathbf{\Phi}$ combines the spectral sensitivity function $f()$ of selected channels and the mask function $m()$ of the channels. Once the matrix Φ is decided, the corresponding MSFA is uniquely designed. Therefore, our goal is to construct the optimal matrix $\mathbf{\Phi}$ to ensure an accurate spectral reconstruction.

The pipeline of the spectral reconstruction method is shown in Figure 2. Taking full advantages of natural spectral image statistics, a natural spectral image patch can be sparsely represented by an overcomplete dictionary:

$$\mathbf{s} = \sum_{j=1}^{q} \mathbf{d_j} \alpha_j \tag{3}$$

where $\mathbf{D} = (\mathbf{d_1}, \mathbf{d_2}, ..., \mathbf{d_q}) \in \mathbb{R}^{whl \times q}$ is the dictionary containing q atoms, and $(\alpha_1, \alpha_2, ..., \alpha_q)^T \in \mathbb{R}^q$ is sparse coefficient vector that most of the coefficients in α are zeros.
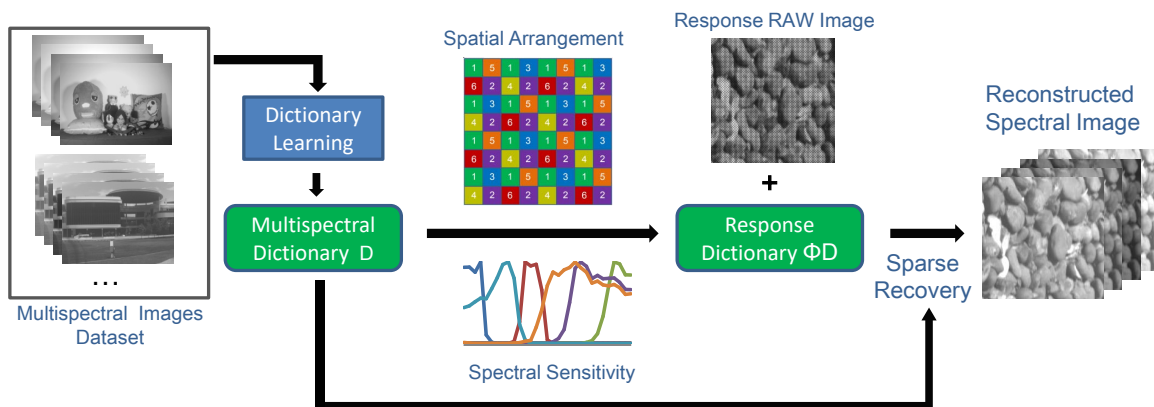

Figure 2. The flowchart of our spectral reconstruction.

The overcomplete dictionary is learned from a large training dataset. The dictionary learning can be formulated as:

$$\min_{\{\mathbf{D}, \mathcal{A}\}} ||\mathbf{T} - \mathbf{D}\,\mathcal{A}||_2^2; \quad s.t.\, \forall i = 1, 2, ..., o,\, ||\mathcal{A}_i||_0 \leq k \tag{4}$$

where $\mathbf{T} \in \mathbb{R}^{whl \times o}$ is the training dataset consisting of o spectral image patches, and $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_o) \in \mathbb{R}^{q \times o}$ is a coefficient matrix composed of o k-sparse vectors in its column. The optimization can be solved efficiently using k-SVD algorithm [44].

According to the theorem of compressive sensing, a spectral image patch can be recovered from a raw response image patch **y** by using sparse representation of the learned dictionary **D** and the sensing matrix **Φ**. Let us assume that the spectral image patches are *k*-sparse, it means the patches can be represented by no more than *k* atoms. Use Λ to denote the recovered support set consisting of the indexes of the non-zero coefficients. Note that the columns of **ΦD** are normalized to unit vectors.

$$\min_{\hat{\alpha}} ||\mathbf{y} - \mathbf{\Phi D_\Lambda}\hat{\alpha}||_2^2, \quad s.t. \; size(\Lambda) \leq k \tag{5}$$

Adopting batch-OMP algorithm [45] can select the support set Λ and optimize the coefficients $\hat{\alpha}$. Then the spectral image patch can be recovered as:

$$\hat{\mathbf{s}} = \mathbf{D_\Lambda}\,\hat{\alpha} \tag{6}$$

In the optimization, due to the limitation of computation capability, the resolution of used raw image patches are usually small (e.g., 32×32). To handle large raw response images, the raw response images must be segmented into small patches with overlapping for optimization. Then, the small reconstructed spectral image patches can be fused to obtain the desired large reconstructed spectral image.

4. Estimation of Reconstruction Error

Most of the existing works [46,47] optimize the sensing matrix via minimizing the mutual coherence. Such mutual coherence minimization algorithms use a simple and efficient objective function to reduce correlations between the atoms in the learned dictionary. However, such algorithm may not guarantee the reconstruction error minimization in our problem, as we observed in our simulations. Unlike these works, we estimate the mean-square error (MSE) of reconstruction and treat the MSE as the objective function to be optimized. Based on the proposed sparse reconstruction method, analyzing the behavior of various noise and estimating the spectral reconstruction error are as follows.

Suppose a camera with a MSFA sensor acquires a noisy response vector $\mathbf{y} \in \mathbb{R}^{whl}$, given by:

$$\mathbf{y} = \mathbf{\Phi s} + \mathbf{e} = \mathbf{\Phi D_\Lambda}\alpha + \mathbf{e} \tag{7}$$

where **Φs** is the noise-free responses, $\mathbf{e} \in \mathbb{R}^{whl}$ are zero-mean random variables with variances σ_e. To simplify the expressions, here we assume the imaging noises are signal-independent and spatial-independent. Sparse recovery for a particular response **y** can be treated as a two-steps process: first decide the support set for the response; and then optimize a least square problem using the selected support set. Therefore, with a given support set Λ, the coefficients $\hat{\alpha}$ can be optimized as:

$$\hat{\alpha} = (\mathbf{\Phi D_\Lambda})^+ \mathbf{y} \tag{8}$$

According to Equation (5), the average spectral reconstruction MSE can be defined as the integral of the estimated error in least square optimization [3], with respect to support set. The MSE is modeled as:

$$
\begin{aligned}
MSE &\triangleq \int_\Lambda p(\Lambda)\, E\{||\mathbf{D_\Lambda}\hat{\alpha} - \mathbf{s}||_2^2\} \\
&= \int_\Lambda p(\Lambda)\, E\{||\mathbf{D_\Lambda}(\hat{\alpha} - \alpha)||_2^2\} \\
&= \int_\Lambda p(\Lambda)\, E\{||\mathbf{D_\Lambda}(\mathbf{\Phi D_\Lambda})^+\mathbf{e} + \mathbf{D_\Lambda}(\mathbf{I} - (\mathbf{\Phi D_\Lambda})^+(\mathbf{\Phi D_\Lambda}))\hat{\alpha}||_2^2\} \\
&= \int_\Lambda p(\Lambda)\, Tr\{\sigma_e^2 \mathbf{D_\Lambda}(\mathbf{\Phi D_\Lambda})^+((\mathbf{\Phi D_\Lambda})^+)^T \mathbf{D_\Lambda}^T \\
&\quad + \mathbf{D_\Lambda}(\mathbf{I} - (\mathbf{\Phi D_\Lambda})^+(\mathbf{\Phi D_\Lambda}))Corr_{\hat{\alpha}_\Lambda}(\mathbf{I} - (\mathbf{\Phi D_\Lambda})^+(\mathbf{\Phi D_\Lambda}))^T \mathbf{D_\Lambda}^T\}
\end{aligned}
\tag{9}
$$

where $p(\Lambda)$ is the possibility of support set Λ, $Tr\{\}$ denotes the trace of a square matrix defined as the sum of the elements on the main diagonal, $()^+$ denotes pseudo-inversion, **I** denotes a $whl \times whl$ identity matrix, and $Corr_{\hat{\alpha}_\Lambda}$ denotes the correlation matrix of the coefficients $\hat{\alpha}$ in the support set Λ.

However, it is impractical to calculate the sum of the estimated error for all possible support sets. In order to approximately calculate the integral result, we adopt the importance sampling strategy that only takes account of the errors of the 10 most frequent support sets. The correlation matrix and the support sets are obtained as priors from dictionary learning. We will verify the accuracy of the estimated errors in Section 6.

The calculation for the estimated reconstruction error can be accomplished in milliseconds by GPU-accelerated parallel computing, while actual computing methods may take seconds [35] or hours [48] to calculate the reconstruction error of only a single 1000×1000 image. Although their methods do not use GPU-acceleration, our method is still much more acceptable and practical than previous works.

5. Heuristic Search for MSFA Design

Without loss of generality, to design an MSFA, we need to select n distinct filters from a candidate filter set $\mathcal{F}$ consisting of m filters ($n \ll m$), and assigning these selected filters on a $t \times t$ square pattern. Here the $t \times t$ pattern is the minimum periodic array on the MSFA. Since the number of distinct filters usually decides the thickness of the MSFA, considering the manufacturing complexity and the practicality of the sensor, the number n must be small ($n \le t^2$). We assume that the square pattern is much smaller than an image patch ($t \ll w$ and $t \ll h$).

For any MSFA with known spectral sensitivity functions of filters and the arrangement of filters on periodic mosaic pattern, its spectral reconstruction capability can be evaluated using Equation (9). However, selecting filters from candidates and assigning the filters on sensor are compute-intensive and time-consuming. Based on such a situation, to make the problem solvable, we use a Genetic Algorithm to heuristically search for a near-optimal solution in global space. The basic definitions and manipulations of the applied Genetic Algorithm are as follows:

Individual Encoding—Each individual MSFA is represented by a set $\tau = \{idx_1, idx_2, ..., idx_{t^2}\}$ consisting of t^2 indexes of selected filters in the candidate set. Each set τ contains n distinct numbers. We denote the sensing matrix of the MSFA τ as Φ_τ.

Fitness Function—The fitness function $fitness()$ is defined using the reciprocal of the estimated reconstruction error shown in Equation (9). We denote the fitness function of the individual τ as:

$$fitness(\tau) = \frac{1}{MSE_{\Phi_\tau}} \tag{10}$$

where MSE_{Φ_τ} is the estimated MSE error with the sensing matrix Φ_τ of the individual τ. The lower the estimated reconstruction error level, the higher the fitness of the designed MSFA.

Select Operation—Selecting the p_s% individuals from current generation to the next generation using Roulette Wheel selection strategy. The greater the fitness function value, the higher the probability of being selected to survive.

Mutation Operation—In order to avoid trapping into the local optimal solution, as shown in Figure 3, we define three mutation operations of an individual, each mutation operation has its own execution possibility p_m%:

(a) randomly select an unselected index value to replace a selected index value in the individual, but the unselected filter must be one of the neighbors of the replaced filter, where the distance between two filters is defined using the cosine distance between the spectral sensitivities of the filters;

(b) randomly exchange two index values in the individual to slightly adjust the distribution of filters in the individual;

(c) randomly replace the value at a position with the value at another position in the individual to slightly adjust the proportion of filters in the individual.
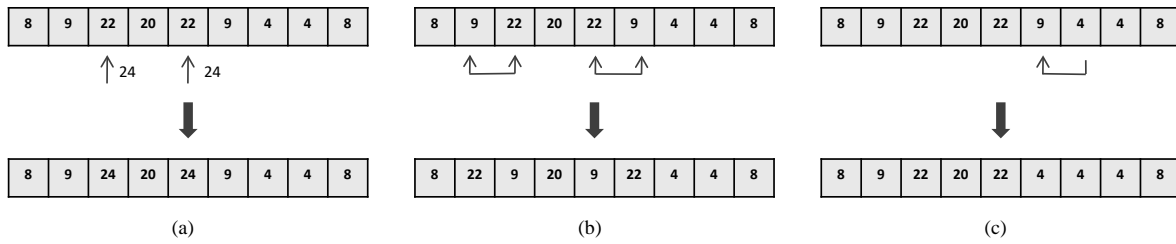
Figure 3. Three mutation operations. (**a**) Replace a selected filter with an unselected filter; (**b**) Exchange two selected filters; (**c**) Replace a channel with another selected channel.

Crossover Operation—To enhance the diversity of populations, as shown in Figure 4, the crossover operation of two individuals is defined as exchanging the different index values stored in two individuals τ_1 and τ_2. The exchanging pairs are selected randomly with a possibility p_c%. This operation actually swaps the filter combinations and filter distributions of the two individuals.
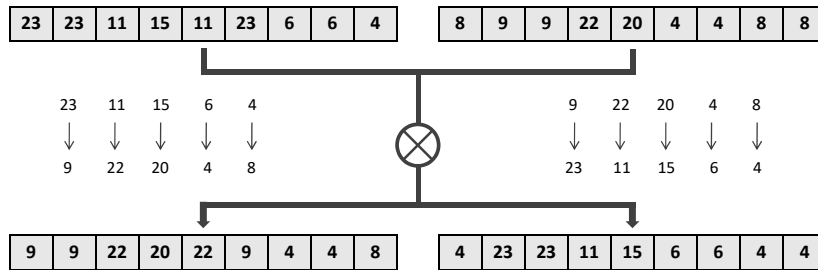
Figure 4. The crossover operation of two individuals.

6. Results and Discussion

To verify our estimated error model, we randomly generated 50 different MSFAs by selecting filters from Roscolux filter set (It is available online along with their specifications as in http://www.rosco.com/filters.) and simulated to calculate the spectral reconstruction mean square error of each MSFA on CAVE [6] multispectral image dataset. Figure 5a shows a strong linear relationship between the objective function and the estimated spectral reconstruction error, confirming the high accuracy of our objective function as a predictor for spectral reconstruction error.
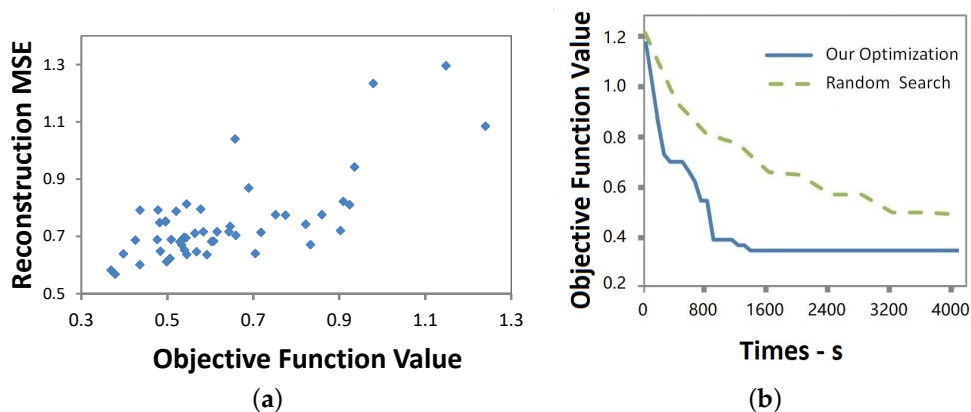
Figure 5. Estimated and actual errors Convergence rate (**a**) Relationship between objective function values and the actual reconstruction mean square error; (**b**) Comparison of convergence rates of the proposed search method and random search.

We evaluated the performance of the proposed heuristics search by comparing it with a random search which randomly generates MSFAs and calculating their estimated MSE for spectral

reconstruction. The convergence rates of both processes are presented in Figure 5b. Since the plot of our optimization strategy achieves a low value and almost flattens out after 1000 s, thus only the optimized values of both strategies in the first 4000 s are shown. The higher convergence rate proved that the proposed search method can effectively accelerate the search.

In order to prove the efficiency and effectiveness of our MSFA design method, we evaluated our optimized MSFA by comparing the spectral construction error of our MSFA with the errors of three MSFAs designed by three state-of-art methods (Yanagi's [43], Fu's [36], and Arad's [35]) in simulations on the three multispectral image datasets(CAVE [6], Harvard's [49], and ICVL's [24]). It is worth noting that for fair comparisons, we did not compare our designed MSFAs with SILIOS and other commercial MSFAs since their candidate filter set is unknown. To the best of our knowledge, Yanagi's method is the only overall MSFA design method based on training data, and other two methods select filters for Bayer patterns without pattern design. In addition, the spectral sensitivity functions and the number of filters used by Yangi's and other two MSFAs are quite different (e.g., nine narrow-band-pass filters used by Yanagi's, three wide-band filters in other twos). Therefore, we designed MSFAs using corresponding candidate filter sets and compared our optimized MSFAs with their solutions respectively.

The training set we used is 100,000 spectral image patches selected from each dataset using volume maximization strategy [13]. Applying K-svd algorithm, we learned a multispectral patch dictionary consisting of $q = 3000$ atoms for sparse reconstruction, where each atom is a $whl \times 1$ vector. The spectral resolution l is set to 31 in the range of visible wavelengths from 400 to 700 nm, and we set $w = h = 8$ for 2×2 MSFA pattern in the comparisons with Fu's and Arad's MSFAs while setting $w = h = 9$ for 3×3 MSFA pattern in the comparison with Yangi's MSFA. In the genetic algorithm, we used the population size 5000 and the algorithm usually converged after 30th generations. The possibility of selection, crossover, and mutation operations are empirically set to $p_s\% = 0.7$, $p_c\% = 0.5$, and $p_m\% = 0.15$. Our MSFA is optimized under moderate noise level with the standard deviation $\sigma_e = 0.03$.

All spectral images were reconstructed using batch-omp algorithm [45] with sparsity = 10. We reconstructed each image under the noiseless case and noisy case (in presence of added noise $SNR \approx 30$ db). We ran the experiments on a 3.6 GHz Intel Core i7 workstation with an NVIDIA TITAN Xp GPU and 64 G main memory.

First, we compared our method with Yanagi's method. We used both of methods to choose nine filters from a candidate set consisting of 31 uniformly distributed narrow band-pass functions from 400 nm to 700 nm, with 10 nm width; and assigned the nine filters onto 3×3 filter patterns. The two designed MSFAs are shown in Figure 6. Taking only five minutes, our method can design an MSFA which is close to the designed MSFA of Yanagi's method, while their method took 20 h. We also showed the reconstruction error and the PSNR of the reconstructed images of the two designed MSFAs in Table 1. Our designed MSFA shows superior results to Yanagi's results irrespective of the dataset used. It is worth noting that the patterns of the two MSFAs match, thus we simulate to calculate reconstruction errors using other MSFAs with the same spectral sensitivities but a different spatial arrangement. We found no better spatial arrangement for the used channels of both methods. It is evident that both of the methods can design the optimal pattern.

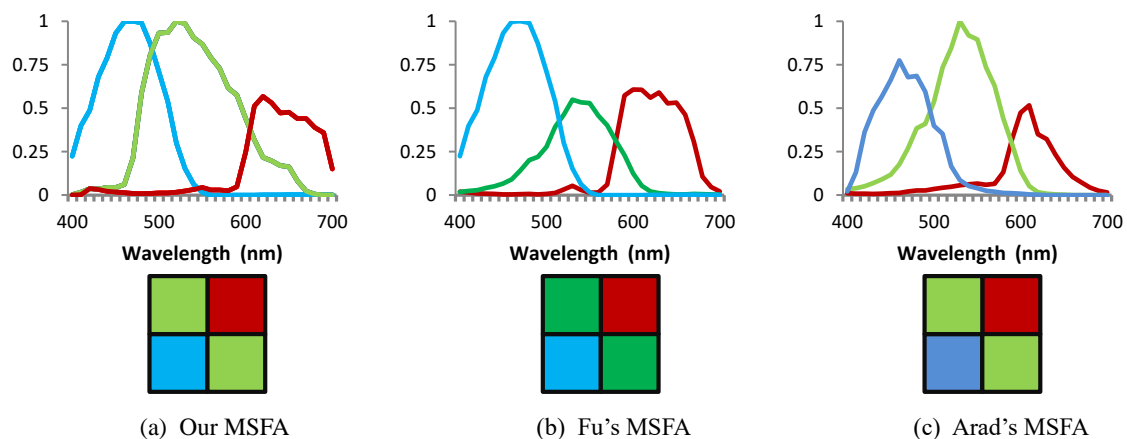(a) Our MSFA　　　　　　　　　　　　　　　　　　　　(b) Yanagi's MSFA
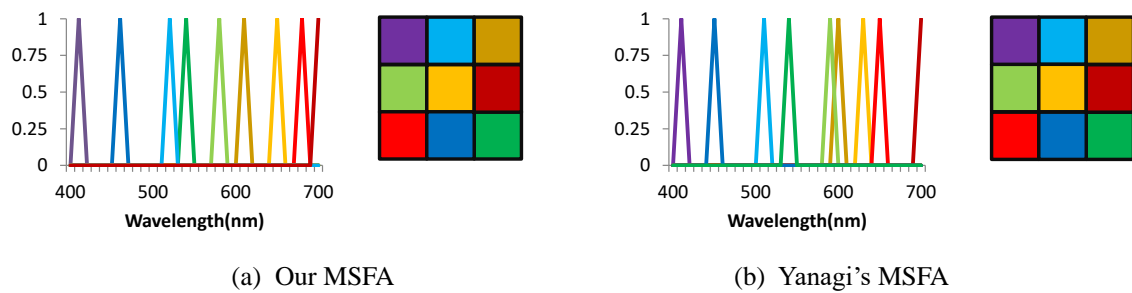
Figure 6. The spectral sensitivity functions and channel spatial arrangement of two MSFAs designed by Yanagi's method and our method.

Table 1. Quantitative evaluation of the reconstructions using Yanagi's and our MSFA (see Figure 6) on synthetic datasets. Note that the bold font indicates the minimal error.

		Dataset					
	MSFA	**CAVE**		**Harvard**		**ICVL**	
		RMSE	**PSNR**	**RMSE**	**PSNR**	**RMSE**	**PSNR**
Noisy	Ours	**0.172**	**21.38**	**0.165**	**22.97**	**0.215**	**19.80**
(SNR = ∞)	Yanagi's	0.203	20.67	0.190	21.18	0.268	18.42
Noisy	Ours	**0.198**	**20.19**	**0.184**	**21.26**	**0.244**	**19.15**
(SNR ≈ 30 db)	Yanagi's	0.231	19.39	0.217	19.90	0.293	17.43

Then we compared our method with Fu's and Arad's method. We selected three channels from a candidate set consisting of RGB channels of 28 off-the-shelf digital cameras [50], and arranged the three channels on 2 × 2 MSFAs. Spectral sensitivities and spatial arrangement of the two designed MSFAs are shown in Figure 7. Note that although the three patterns match, our pattern is obtained by our optimization method while their patterns are manually set, thus our searching space is much larger than theirs. Besides, our method took only 10 min to achieve the results while Arad's method consumed 10 h. To quantify the reconstruction accuracy of multispectral images, we calculated the MSE and PSNR of the reconstructed spectral images of the three designed MSFAs and presented the results in Table 2. It can be observed that our designed MSFA showed superior results to their results, especially in the noisy case.

(a) Our MSFA　　　　　　　　　　(b) Fu's MSFA　　　　　　　　　　(c) Arad's MSFA

Figure 7. The spectral sensitivity functions and channel spatial arrangement of three MSFAs designed by Arad's, Fu's and our method.

Table 2. Quantitative evaluation of the sparse reconstructions using Fu's, Arad's and our MSFA (see Figure 7) on synthetic datasets. Note that the bold font indicates the minimal error.

		Dataset					
	MSFA	**CAVE**		**Harvard**		**ICVL**	
		RMSE	**PSNR**	**RMSE**	**PSNR**	**RMSE**	**PSNR**
Noisy (SNR = ∞)	Ours	**0.258**	**20.21**	**0.212**	**21.97**	**0.071**	**28.59**
	Fu's	0.322	19.89	0.219	21.63	0.086	28.57
	Arad	0.386	19.75	0.232	21.14	0.094	28.51
Noisy (SNR ≈ 30db)	Ours	**0.281**	**17.86**	**0.246**	**17.97**	**0.095**	**24.26**
	Fu's	0.328	16.36	0.257	16.77	0.109	22.33
	Arad	0.440	15.32	0.288	16.03	0.156	20.98

For visualization, we also show three example images from CAVE's, Harvard's dataset and ICVL's dataset and the normalized absolute reconstruction RMSE of each pixel in Figure 8a. A spectral image was reconstructed from all of its overlap patches by filling in the patches from left to right, top to bottom and fusing adjacent patches. It is worth noting that the absolute reconstruction errors of our MSFA are much more uniform in the spatial domain. The comparison of three examples of reconstructed spectra by different MSFAs is presented in Figure 8b. Although our results are not perfect, they are the closest to the ground truth (please zoom in for details), The imperfect results are due to the lack of freedom of filter selection since we chose filters from the filters of 28 very similar commercial cameras.
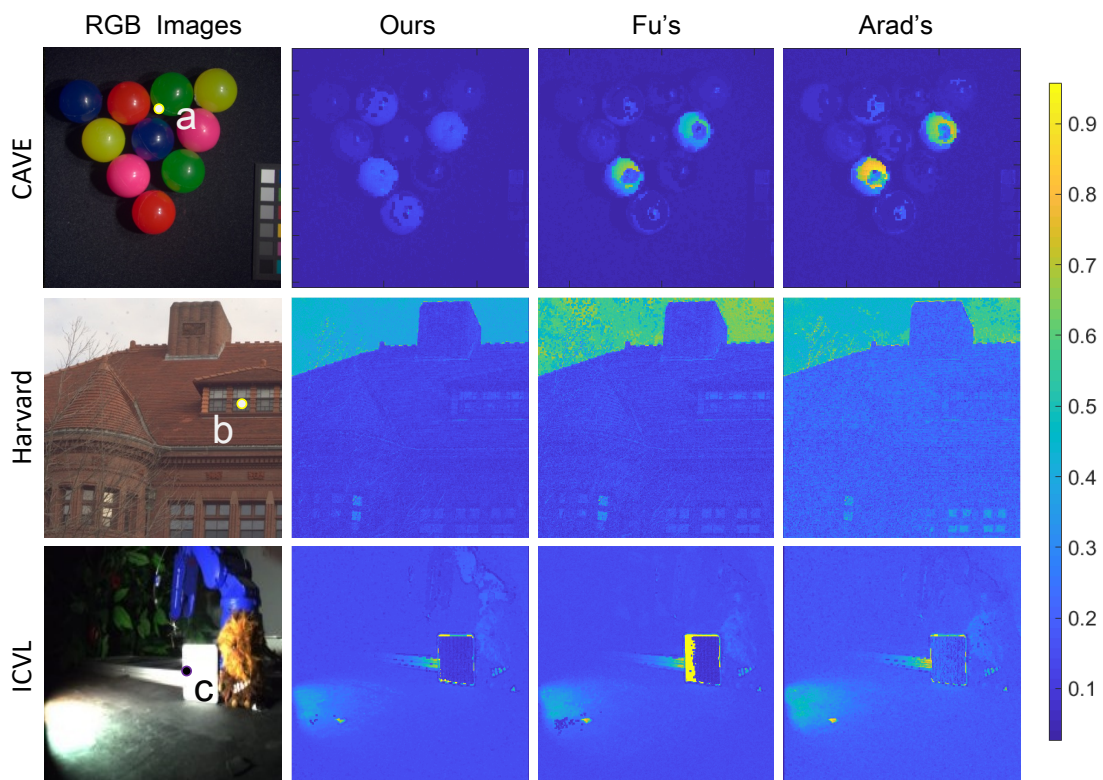
(**a**) Comparison of the absolute spectral reconstruction error of three images using Fu's, Arad's, and our method.

Figure 8. *Cont.*
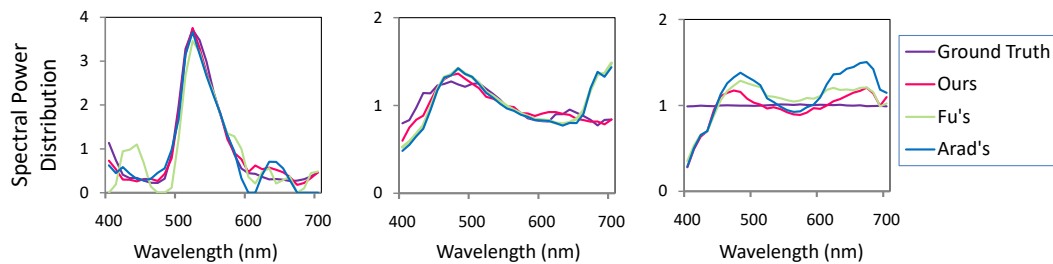
(**b**) Comparison of the reconstructed spectra at position a (left), b (middle), c (right) using Fu's, Arad's, and our method.

Figure 8. The comparison of reconstruction accuracy of the three MSFAs (see Figure 7).

We also evaluated the MSFAs by using deep-learning-based spectral reconstruction method, such as the HSCNN method [27], which won the NTIRE spectral reconstruction contest [29] recently. Our designed MSFA still outperformother MSFAs although the reconstruction method has been changed. The results are shown in Table 3.

Table 3. Quantitative evaluation of the deep-learning-based reconstructions using Fu's, Arad's and our MSFA (see Figure 7) on synthetic datasets.

		Dataset					
	MSFA	**CAVE**		**Harvard**		**ICVL**	
		RMSE	**PSNR**	**RMSE**	**PSNR**	**RMSE**	**PSNR**
Noisy (SNR = ∞)	Ours	**0.148**	**24.46**	**0.091**	**29.96**	**0.054**	**31.12**
	Nie's	0.166	23.76	0.114	27.07	0.068	29.87
	Arad	0.161	23.54	0.119	26.49	0.069	29.89
Noisy (SNR = 30 db)	Ours	**0.163**	**23.08**	**0.097**	**27.23**	**0.059**	**30.02**
	Nie's	0.172	22.59	0.118	25.72	0.072	27.93
	Arad	0.186	22.47	0.122	25.06	0.072	28.28

The two experiments above show that our MSFA design method is more efficient and effective than the state-of-the-art methods, irrespective of the candidate filters and dataset used, thus our method can be applied to design MSFAs using much larger candidate filter sets.

7. Conclusions

In this paper, we introduced an MSFA design method based on sparse representation. The proposed method takes full advantage of spectral image statistics and optimizes filter selection and spatial arrangement. We presented an analysis of spectral reconstruction error to describe the imaging and reconstruction capability of different MSFA. Through globally searching the MSFA which has the minimum reconstruction error, the near-optimal MSFA can be achieved. Experiments demonstrate that our method is robust to different image datasets and filter sets; our optimized MSFA outperforms that of other methods and can provide valuable references or guidance for MSFA designers.

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

References

1. Hagen, N.; Kudenov, M.W. Review of Snapshot Spectral Imaging Technologies. *Opt. Eng.* **2013**, *52*, 090901. [CrossRef]

2. Silios Technologies: Color Shades. Available online: https://www.silios.com/cms-cameras-1 (accessed on 31 May 2019).

3. Li, Y.; Majumder, A.; Zhang, H.; Gopi, M. Optimized Multi-Spectral Filter Array Based Imaging of Natural Scenes. *Sensors* **2018**, *18*, 1172. [CrossRef] [PubMed]

4. 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 2012 19th IEEE International Conference on Image Processing, Orlando, FL, USA, 30 September–3 October 2012; pp. 2137–2140.

5. 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. [CrossRef] [PubMed]

6. Yasuma, F.; Mitsunaga, T.; Iso, D.; Nayar, S.K. Generalized Assorted Pixel Camera: Postcapture Control of Resolution, Dynamic Range, and Spectrum. *Trans. Image Process.* **2010**, *19*, 2241–2253. [CrossRef] [PubMed]

7. Chi, C.; Yoo, H.; Ben-ezra, M. Multi-spectral imaging by optimized wide band Illumination. *Int. J. Comput. Vis.* **2010**, *56*, 140. [CrossRef]

8. Jia, J.; Barnard, K.; Hirakawa, K. Fourier Spectral Filter Array For Optimal Multispectral Imaging. *Trans. Image Process.* **2016**, *25*, 1530–1543. [CrossRef] [PubMed]

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

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

11. Liu, Z.; Liu, Q.; ai Gao, G.; Li, C. Optimized spectral reconstruction based on adaptive training set selection. *Opt. Express* **2017**, *25*, 12435–12445. [CrossRef]

12. Liang, J.; Wan, X. Optimized method for spectral reflectance reconstruction from camera responses. *Opt. Express* **2017**, *25*, 28273–28287. [CrossRef]

13. Li, Y.; Wang, C.; Zhao, J.; Yuan, Q. Efficient spectral reconstruction using a trichromatic camera via sample optimization. *Vis. Comput.* **2018**, *34*, 1773–1783. [CrossRef]

14. Sadeghipoor, Z.; Lu, Y.M.; Süsstrunk, S. A novel compressive sensing approach to simultaneously acquire color and near-infrared images on a single sensor. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 1646–1650.

15. Zhang, J.; Luo, H.; Liang, R.; Ahmed, A.; Zhang, X.; Hui, B.; Chang, Z. Sparse representation-based demosaicing method for microgrid polarimeter imagery. *Opt. Lett.* **2018**, *43*, 3265–3268. [CrossRef] [PubMed]

16. Li, Y.; Wang, C.; Zhao, J. Locally Linear Embedded Sparse Coding for Spectral Reconstruction From RGB Images. *IEEE Signal Process. Lett.* **2018**, *25*, 363–367. [CrossRef]

17. Aggarwal, H.K.; Majumdar, A. Compressive sensing multi-spectral demosaicing from single sensor architecture. In Proceedings of the 2014 IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP), Xi'an, China, 9–13 July 2014; pp. 334–338.

18. Sadeghipoor, Z.; Thomas, J.B.; Süsstrunk, S. Demultiplexing visible and near-infrared information in single-sensor multispectral imaging. In Proceedings of the 24th Color and Imaging Conference, San Diego, CA, USA, 7–11 November 2016; pp. 76–81.

19. Imai, F.H.; Berns, R.S. Spectral estimation using trichromatic digital cameras. In Proceedings of the International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives, Chiba, Japan, 21–22 October 1999; pp. 1–8.

20. Zhang, W.F.; Tang, G.; Dai, D.Q.; Nehorai, A. Estimation of reflectance from camera responses by the regularized local linear model. *Opt. Lett.* **2011**, *36*, 3933–3935. [CrossRef] [PubMed]

21. Nguyen, R.M.; Prasad, D.K.; Brown, M.S. Training-based spectral reconstruction from a single RGB image. In Proceedings of the European Conference on Computer Vision (ECCV), Zurich, Switzerland, 6–12 September 2014; pp. 186–201.

22. Heikkinen, V.; Cámara, C.; Hirvonen, T.; Penttinen, N. Spectral imaging using consumer-level devices and kernel-based regression. *JOSA A* **2016**, *33*, 1095–1110. [CrossRef] [PubMed]

23. Robles-Kelly, A. Single image spectral reconstruction for multimedia applications. In Proceedings of the 23rd ACM international conference on Multimedia, Brisbane, Australia, 26–30 October 2015; pp. 251–260.

24. Arad, B.; Ben-Shahar, O. Sparse recovery of hyperspectral signal from natural rgb images. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; pp. 19–34.

25. Fu, Y.; Zheng, Y.; Zhang, L.; Huang, H. Spectral Reflectance Recovery From a Single RGB Image. *IEEE Trans. Comput. Imaging* **2018**, *4*, 382–394. [CrossRef]

26. Jia, Y.; Zheng, Y.; Gu, L.; Subpa-Asa, A.; Lam, A.; Sato, Y.; Sato, I. From RGB to spectrum for natural scenes via manifold-based mapping. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 4715–4723.

27. Xiong, Z.; Shi, Z.; Li, H.; Wang, L.; Liu, D.; Wu, F. Hscnn: Cnn-based hyperspectral image recovery from spectrally undersampled projections. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 518–525.

28. Aeschbacher, J.; Wu, J.; Timofte, R. In defense of shallow learned spectral reconstruction from rgb images. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 471–479.

29. Arad, B.; Ben-Shahar, O.; Timofte, R.; Van Gool, L.; Zhang, L.; Yang, M.H. Ntire 2018 challenge on spectral reconstruction from rgb images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake, UT, USA, 18–22 June 2018; pp. 1042–104209.

30. Zhao, Y.; Guo, H.; Ma, Z.; Cao, X.; Yue, T.; Hu, X. Hyperspectral Imaging With Random Printed Mask. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long beach, CA, USA, 16–20 June 2019; pp. 10149–10157.

31. Imai, F.H.; Rosen, M.R.; Berns, R.S. Comparison of spectrally narrow-band capture versus wide-band with a priori sample analysis for spectral reflectance estimation. In *Color and Imaging Conference, Proceedings of the 8th Color and Imaging Conference, Scottsdale, AZ, USA, 7–10 November 2000*; Society for Imaging Science and Technology: Springfield, VA, USA; pp. 234–241.

32. Ansari, K.; Thomas, J.B.; Gouton, P. Spectral band Selection Using a Genetic Algorithm Based Wiener Filter Estimation Method for Reconstruction of Munsell Spectral Data. *Electron. Imaging* **2017**, *2017*, 190–193. [CrossRef]

33. Wang, X.; Thomas, J.B.; Hardeberg, J.Y.; Gouton, P. Multispectral imaging: narrow or wide band filters? *J. Int. Colour Assoc.* **2014**, *12*, 44–51.

34. 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. [CrossRef]

35. Arad, B.; Ben-Shahar, O. Filter selection for hyperspectral estimation. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 21–26.

36. Fu, Y.; Zhang, T.; Zheng, Y.; Zhang, D.; Huang, H. Joint Camera Spectral Sensitivity Selection and Hyperspectral Image Recovery. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 788–804.

37. Lukac, R.; Plataniotis, K.N. Color filter arrays: Design and performance analysis. *IEEE Trans. Consumer Electron.* **2005**, *51*, 1260–1267. [CrossRef]

38. Miao, L.; Qi, H.; Ramanath, R.; Snyder, W.E. Binary Tree-based Generic Demosaicking Algorithm for Multispectral Filter Arrays. *IEEE Trans. Signal Process* **2006**, *15*, 3550–3558.

39. 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. [CrossRef] [PubMed]

40. Henz, B.; Gastal, E.S.; Oliveira, M.M. Deep Joint Design of Color Filter Arrays and Demosaicing. *Comput. Graphics Forum* **2018**, *37*, 389–399. [CrossRef]

41. Nie, S.; Gu, L.; Zheng, Y.; Lam, A.; Ono, N.; Sato, I. Deeply Learned Filter Response Functions for Hyperspectral Reconstruction. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake, UT, USA, 18–22 June 2018; pp. 4767–4776.

42. Li, J.; Bai, C.; Lin, Z.; Yu, J. Optimized color filter arrays for sparse representation-based demosaicking. *IEEE Trans. Image Process.* **2017**, *26*, 2381–2393. [CrossRef] [PubMed]

43. Yanagi, Y.; Shinoda, K.; Hasegawa, M.; Kato, S.; Ishikawa, M.; Komagata, H.; Kobayashi, N. Optimal transparent wavelength and arrangement for multispectral filter array. *Electron. Imaging* **2016**, *2016*, 1–5. [CrossRef]

44. Aharon, M.; Elad, M.; Bruckstein, A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Process.* **2006**, *54*, 4311–4322. [CrossRef]

45. Rubinstein, R.; Zibulevsky, M.; Elad, M. Efficient Implementation of the K-SVD Algorithm Using Batch Orthogonal Matching Pursuit. Available online: http://www.cs.technion.ac.il/users/wwwb/cgi-bin/tr-get.cgi/2008/CS/CS-2008-08.pdf (accessed on 27 June 2019).

46. Obermeier, R.; Martinez-Lorenzo, J.A. Sensing Matrix Design via Mutual Coherence Minimization for Electromagnetic Compressive Imaging Applications. *IEEE Trans. Comput. Imaging* **2017**, *3*, 217–229. [CrossRef]

47. Lu, C.; Li, H.; Lin, Z. Optimized projections for compressed sensing via direct mutual coherence minimization. *Signal Process.* **2018**, *151*, 45–55. [CrossRef]

48. Kawakami, R.; Matsushita, Y.; Wright, J.; Ben-Ezra, M.; Tai, Y.W.; Ikeuchi, K. High-resolution hyperspectral imaging via matrix factorization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, CO, USA, 20–25 June 2011; pp. 2329–2336.

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

50. Jiang, J.; Liu, D.; Gu, J.; Süsstrunk, S. What is the space of spectral sensitivity functions for digital color cameras? In Proceedings of the 2013 IEEE Workshop on Applications of Computer Vision (WACV), Tampa, FL, USA, 15–17 January 2013; pp. 168–179.