Automatic Design of Color Filter Arrays in the Frequency Domain

Chenyan Bai, Jia Li, Zhouchen Lin, *Senior Member, IEEE*, and Jian Yu

***Abstract*— In digital color imaging, the raw image is typically obtained through a single sensor covered by a color filter array (CFA), which allows only one color component to be measured at each pixel. The procedure to reconstruct a full color image from the raw image is known as demosaicking. Since the CFA may cause irreversible visual artifacts, the CFA and the demosaicking algorithm are crucial to the quality of demosaicked images. Fortunately, the design of CFAs in the frequency domain provides a theoretical approach to handling this issue. However, almost all the existing design methods in the frequency domain involve considerable human effort. In this paper, we present a new method to automatically design CFAs in the frequency domain. Our method is based on the frequency structure representation of mosaicked images. We utilize a multi-objective optimization approach to propose frequency structure candidates, in which the overlap among the frequency components of images mosaicked with the CFA is minimized. Then, we optimize parameters for each candidate, which is formulated as a constrained optimization problem. We use the alternating direction method to solve it. Our parameter optimization method is applicable to arbitrary frequency structures, including those with conjugate replicas of chrominance components. Experiments on benchmark images confirm the advantage of the proposed method.**

***Index Terms*— Color filter array (CFA), demosaicking, multi-objective optimization, alternating direction method (ADM).**

I. Introduction

COLOR images contain at least three color components at each pixel, such as red (R), green (G), and blue (B), or cyan (C), magenta (M), and yellow (Y). To produce a color image, a digital camera would need one sensor for each color component to record its values. However, multiple sensors are expensive and have difficulty in precise registration. So most digital cameras use a single sensor covered by a color filter array (CFA). A CFA is a hardware which has the same size as the sensor and allows only one color component to be sensed at each pixel. The process to recover a full color image from the image obtained from a single sensor with a CFA is called *demosaicking* [1]–[3]. Both the CFA and the demosaicking algorithm affect the quality of the reconstructed full color image. As a demosaicking algorithm takes the images mosaicked with a CFA as input, one can design the CFA to make the subsequent demosaicking process much more robust to visual artifacts. Note that whenever we refer to CFA design, the CFA is periodic and defined on the square lattice, in which the minimum periodic array is called a CFA pattern. For other types of CFAs, e.g., random CFA [4] or irregular CFA [5], we refer the readers to [5].

Most of the existing CFAs are designed empirically in the spatial domain under different considerations [6], [7]. The Bayer CFA [8] is the most popular CFA in the consumer market (Fig. 1(1a)) and hence the majority of demosaicking algorithms are proposed for it [1]–[3]. The Bayer CFA was designed based on the human visual system's (HVS) greater sensitivity to green light. However, spectral characteristic analysis [9] has shown that aliasing artifacts are inherent to the Bayer CFA. We can see from Fig. 1(2a) that there are chrominance components of the image mosaicked with Bayer CFA located on the horizontal and the vertical axes, where the luminance component has a high spectral density. To overcome the limitation of the Bayer CFA, many other CFAs have been proposed [3], [6], [7], [10]–[12]. Recently, Fujifilm X-Trans CFA [13] was presented to mimic the irregular and randomly arranged particles in silver halide film. It is claimed to be more resistant to Moiré effects than the Bayer CFA.

Some more systematic CFA design methods have also been proposed. Parmar and Reeves [17] developed a CFA design method using only RGB color components. They proposed an error criterion that incorporates the HVS effect in evaluating the perceived quality of the reconstructed image. They also presented a sequential backward selection (SBS) scheme to optimize the criterion. Their designed CFA was shown to perform better than the Bayer CFA in terms of their error criterion. A similar method was proposed by Lu and Vetterli [18]. The color components of their CFA are weighted combinations of R, G, and B. They optimized the CFA by minimizing the reconstruction error of the linear minimum mean square error demosaicking [19]. This approach was extended in [20] to design a CFA that can simultaneously capture visible and near-infrared image pairs.

Manuscript received October 14, 2015; revised January 19, 2016; accepted February 8, 2016. Date of publication February 18, 2016; date of current version March 8, 2016. The work of J. Yu was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61370129 and in part by the Ph.D. Programs Foundation of Ministry of Education of China under Grant 20120009110006. The work of Z. Lin was supported in part by the 973 Program of China under Grant 2015CB352502, in part by the NSFC under Grant 61231002 and Grant 61272341, and in part by the Microsoft Research Asia Collaborative Research Program. The associate editor coordinating the review of this manuscript and approving it for publication was Mr. Pierre-Marc Jodoin.

C. Bai, J. Li, and J. Yu are with the Beijing Key Laboratory of Traffic Data Analysis and Mining, School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, P.R. China (e-mail: baichenyan@gmail.com; jiali.gm@gmail.com; jianyu@bjtu.edu.cn).

Z. Lin is with the Key Laboratory of Machine Perception (Ministry of Education), School of Electronic Engineering and Computer Science, Peking University, Beijing 100871, P.R. China, and also with the Cooperative Medianet Innovation Center, Shanghai Jiao Tong University, Shanghai 200240, P.R. China (e-mail: zlin@pku.edu.cn).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org., provided by the author. The material contains the proof of Theorem 1. The total size of the file is 191 MB. Contact baichenyan@gmail.com for further questions about this work.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2016.2531287

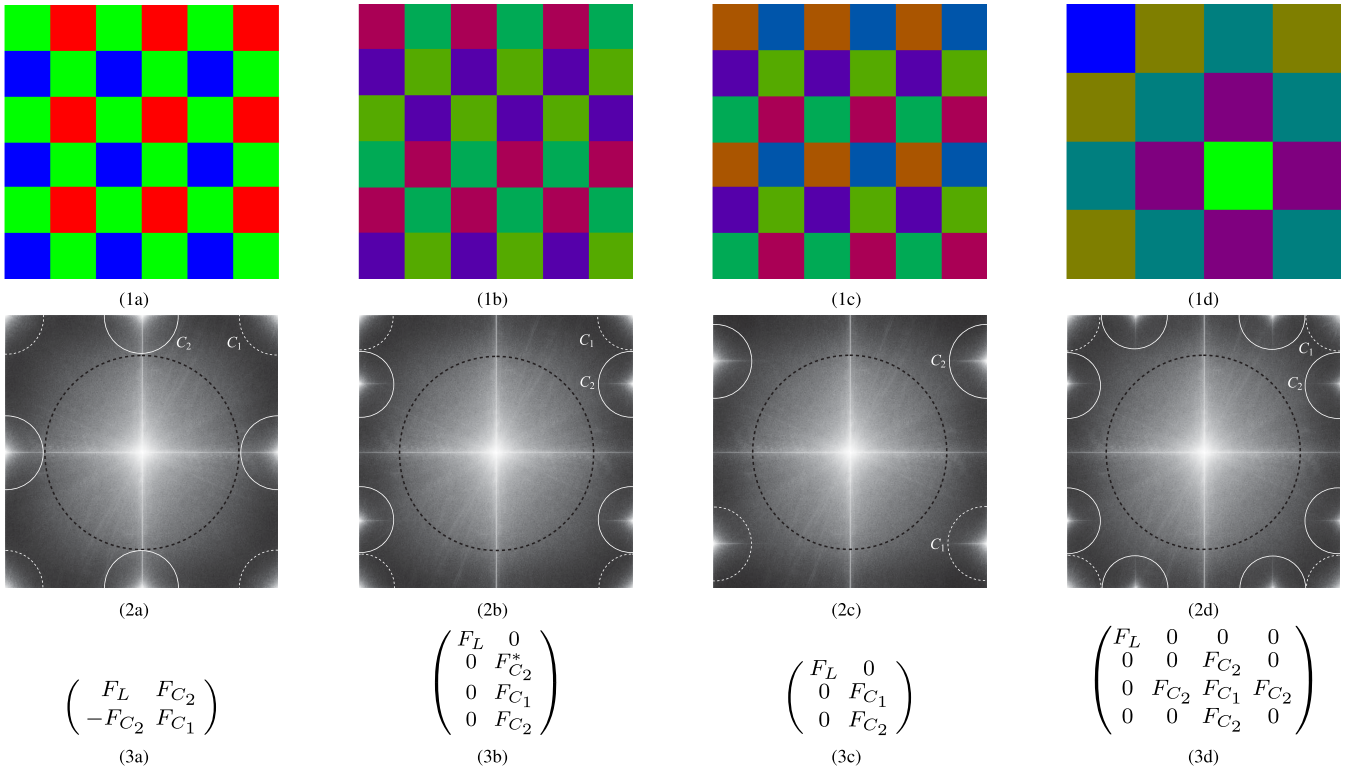


Fig. 1. Four existing CFA patterns and their corresponding spectra and frequency structures. The first row are the (a) Bayer CFA [8], (b) Hirakawa CFA [14], (c) Condat CFA [15], and (d) Hao CFA [7], respectively. The sum across color channels of all the CFAs are normalized to be all-one matrices. The second row are the average spectra of all 24 images in the Kodak dataset [16] mosaicked with the corresponding CFAs in the first row. From the second row, we can clearly see that the luminance spectrum is concentrated in the center and on the horizontal and the vertical axes. The chrominance spectra are located on the borders in which the dashed and solid circles refer to the replicas of the two chrominance components C_1 and C_2, respectively. The third row are the corresponding frequency structures [7] of CFAs in the first row, which record the spectral components and their positions in the second row. Note that by convention the basebands are placed at the center in the second row, but all the Discrete Fourier Transform (DFT) spectra are periodic in both horizontal and vertical directions. However, the frequency origin $(0, 0)$ of frequency structures in the third row is at the top-left of the matrix. **Images in this paper are best viewed on screen!**

In [21], Parmar and Reeves presented a design method for three-color CFAs (e.g., RGB, or CMY). They first demonstrated that the spectral sensitivity functions of the color filter affect both the color reproduction and the quality of demosaicked images. Then they selected spectral sensitivity functions via minimizing the reconstruction error in the CIELAB space [22] over several illuminants. The obtained CFAs were shown to perform better than several existing RGB and CMY CFAs in terms of both perceptual evaluation and objective image quality measure S-CIELAB [22]. Sadeghipoor et al. [23] incorporated the smoothness prior of spectral sensitivities into the optimization to select spectral sensitivity functions.

Since the seminal work by Alleysson et al. [9], the frequency representation of mosaicked images has provided new insights into demosaicking algorithm [9], [24] and CFA design [14]. The CFA design in the frequency domain [7], [14], [15] provides a theoretical approach to producing full color images with fewer visual artifacts. Alleysson et al. found that the images mosaicked with the Bayer CFA consist of a luminance component (luma for short) at the baseband and multiple modulated replicas[1] of two independent chrominance components[2] (chroma for short) at the high frequency bands. We show the spectrum of Bayer CFA in Fig. 1(2a), where the luma is denoted by the black dashed circle and the modulated chromas are denoted by the white dashed and solid circles. Hirakawa and Wolfe [14] extended the spectral characteristic of Bayer CFA to arbitrary rectangular and periodic CFAs. Then the CFA design was converted to a parameter search problem, where the minimum distance between luma and modulated chromas is maximized. Hirakawa and Wolfe conducted exhaustive search in the parameter space with several carefully designed constraints. The optimality of each parameter was empirically evaluated by its demosaicking performance on an image set. So their approach is time-consuming and depends on the demosaicking algorithm and the image set. One of the obtained CFA is shown in Fig. 1(1b). We can see that the chromas of images mosaicked with their CFA are modulated far away from the center and the horizontal and the vertical axes (Fig. 1(2b)). Condat [15] followed this approach and developed a constructive method to manually determine the parameters step by step. He argued that a CFA should be robust to noise as well as aliasing, especially for photography in

[1]In general, the replicas of a chrominance component C can be its conjugate C^* or aC, where a is a real or complex scalar.

[2]Note that representing all color requires one luminance component and at least two independent chrominance components. In order to avoid redundancy, it is natural to use only two chrominance components.

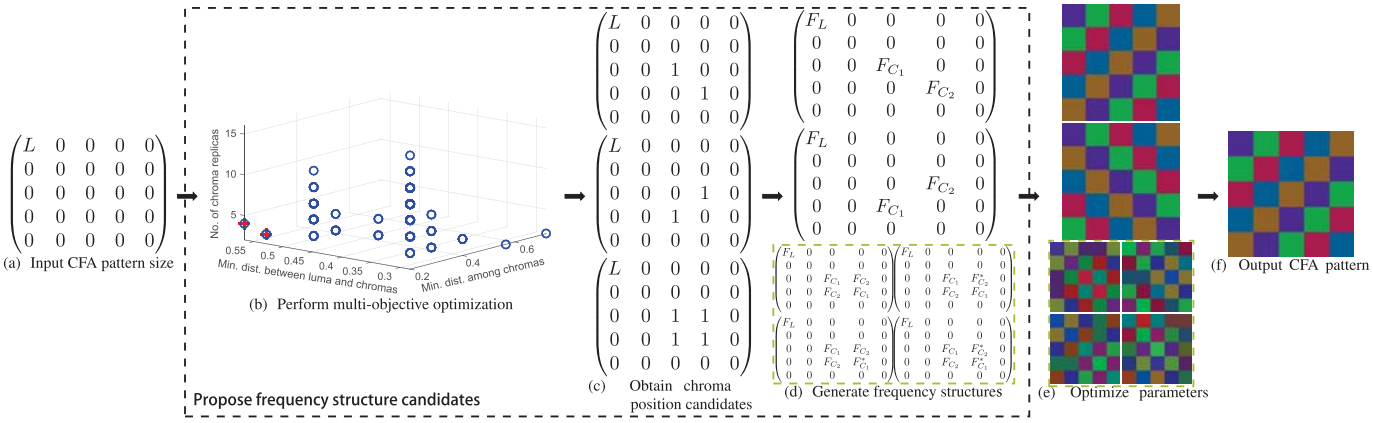


Fig. 2. Overview of the proposed automatic CFA design method. From left to right: (a) is the input CFA pattern size, in which 'L' marks the position of luma and '0' marks the available chroma positions. Our method first performs multi-objective optimization (b) to obtain chroma position candidates (c), where '1' marks the selected chroma positions. Then it generates frequency structures (d) according to the chroma position candidates. It next optimizes parameters for every frequency structure to produce CFAs (e) and finally outputs the obtained CFA (f).

low-light conditions. So he chose the parameters by simultaneously maximizing the minimum distance between luma and chromas and the sensitivity of the CFA. He proposed a 2×3 CFA pattern using six color components (Fig. 1(1c)). Both approaches need a considerable human effort in parameter optimization. Also, an appropriate luminance/chrominance basis should be chosen carefully.

An alternative approach for CFA design in the frequency domain was proposed by Hao et al. [7]. Their method is based on the frequency structure, which records all the luma and chromas of mosaicked images at their corresponding frequencies in a matrix (see the third row of Fig. 1). They first manually specified a frequency structure with some guidelines. Then for the given frequency structure, the parameter optimization was formulated as a constrained optimization problem. With certain assumptions, they used a geometric method to solve it. The proposed CFA is shown in Fig. 1(1d). Their geometric method is attractive for it is intuitive and visual. However, the user has to choose the vertices of optimal triangle on the boundary of the feasible region. Moreover, it cannot work when the frequency structure contains conjugate replicas of a chroma (e.g., the one shown in Fig. 1(3b)). Additionally, the geometric method is derived from the Frobenius norm of the inverse of color transformation matrix, which is an approximation of its spectral norm. Details on this issue are provided in subsection II-B.

Inspired by Hao et al. [7], in this paper we propose an automatic method for CFA design in the frequency domain, which requires no human interaction. As shown in Fig. 2, our method consists of two main steps. For a given CFA pattern size, it first proposes frequency structure candidates (the dashed box in Fig. 2), where the overlap between frequency components is minimized. Then it optimizes parameters for each candidate by maximizing the numeric stability of color transformation (Fig. 2(e)). We summarize the comparison of automaticity of our method with that of other methods in Table I. More details on the comparison are presented in subsection II-B.

TABLE I

COMPARISON OF THE AUTOMATICITY OF DIFFERENT METHODS IN CFA DESIGN PROCESS. NOTE THAT THE PARAMETER OPTIMIZATION METHOD DEVELOPED BY HAO [7] CANNOT WORK WHEN THE FREQUENCY STRUCTURE CONTAINS CONJUGATE CHROMA REPLICAS, WHILE THE PROPOSED ONE CAN HANDLE THIS CASE

Method	Specification of luma/chroma basis	Minimization of the overlap among frequency components	Parameter optimization
Hirakawa [14]	nonautomatic	**automatic**	nonautomatic
Condat [15]	nonautomatic	**automatic**	nonautomatic
Hao [7]	**automatic**	nonautomatic	nonautomatic
Our	**automatic**	**automatic**	**automatic**

The contributions of this paper are:

- Based on the frequency structure [7], we propose a new approach for designing CFAs in the frequency domain, which is fully automatic.
- We use multi-objective optimization to propose frequency structure candidates, which discards a vast majority of unpromising frequency structures automatically.
- For a given frequency structure, we formulate parameter optimization as a constrained optimization problem, which directly works on the spectral norm. We use the alternating direction method (ADM) to solve it. Our formulation and solution process are both derived from the spectral norm, and they are applicable to arbitrary frequency structures, including those with conjugate chroma replicas.

The remaining part of this paper is organized as follows. In Section II, we introduce the existing CFA design methods in the frequency domain. Then we introduce our CFA design method in Section III and the solution process in the Appendix. In Section IV, we conduct experiments to show the effectiveness of our design method. Finally, we conclude the paper in Section V.

II. RELATED WORK

In this section, we first introduce the frequency representation of mosaicked images. Then we review the

existing design methods of CFAs in the frequency domain.

A. Frequency Representation of Mosaicked Images

Alleysson et al. [9] showed that an image mosaicked with the Bayer CFA can be interpreted in the frequency domain as the sum of a luma and multiple subsampled replicas of two independent chromas. The subsampling of chromas is implemented by multiplying with modulation functions. More importantly, they demonstrated that the luma and the modulated (or subsampled) chromas have different locations in the frequency domain, i.e., luma is at the baseband, while the modulated chromas are at the high frequency bands (Fig. 1(2a)). This leads to designing a frequency selection based demosaicking algorithm to recover the luma and the modulated chromas [9], [24]. It then demodulates the modulated chromas. If there exist multiple replicas of a chroma, one can combine all the estimations adaptively to obtain a more accurate one [24]. The RGB components are finally estimated using the transformation from luma/chroma basis to RGB basis. An example of basis transformation is shown in (2). Frequency selection based demosaicking is linear which provides a good compromise between the quality of demosaicked images and computational complexity.

Although the above analysis is induced by the Bayer CFA, the characterization of mosaicked images that luma and modulated chromas have different frequency locations can extend to arbitrary periodic CFAs defined on the square lattice [14] (see (3) in subsection II-B). Hao et al. [7] proposed using a matrix called frequency structure to record all the information about the frequency representation of images mosaicked with a CFA. The frequency structure contains the luma and all the replicas of chromas, and with the positions of which one can obtain all the modulated ones (see Fig. 1). Moreover, Hao et al. [7] proved that the frequency structure can be easily computed using a symbolic Discrete Fourier Transform (DFT). The symbolic DFT is a standard DFT which treats symbols as parameters [25]. For example, for a sequence of N symbols $s_0, s_1, \cdots, s_{N-1}$, its 1D symbolic DFT is defined as a sequence of linear polynomials $S_0, S_1, \cdots, S_{N-1}$, where $S_k = \frac{1}{N}\sum_{u=0}^{N-1} s_u \exp(-2\pi iku/N)$, $k \in \{0, 1, \cdots, N-1\}$, and i is the imaginary unit. This definition of 1D symbolic DFT can be generalized to the 2D case. We take the Hirakawa CFA [14] (see Fig. 1(1b)) as an example. Since the Hirakawa CFA in the RGB basis is

$$CFA_{\mathrm{H}} = \frac{1}{3}\left(\begin{pmatrix}2&0\\1&1\\1&1\\0&2\end{pmatrix}R + \begin{pmatrix}0&2\\0&2\\2&0\\2&0\end{pmatrix}G + \begin{pmatrix}1&1\\2&0\\0&2\\1&1\end{pmatrix}B\right),$$

we can compute its frequency structure as follows:

$$F_{\mathrm{H}} = \mathrm{DFT}\,(CFA_{\mathrm{H}}) = \begin{pmatrix}F_L & 0\\ 0 & F^*_{C_2}\\ 0 & F_{C_1}\\ 0 & F_{C_2}\end{pmatrix}, \tag{1}$$

where R, G, and B are symbols, * means symbolic conjugate, i.e., the coefficients of R, G, and B in $F^*_{C_2}$ are all conjugate to those in F_{C_2}, and F_L, F_{C_1}, and F_{C_2} refer to the luma and the two chromas used in frequency structure, respectively, which are given as:

$$\begin{pmatrix}F_L\\F_{C_1}\\F_{C_2}\end{pmatrix} = \frac{1}{12}\begin{pmatrix}4 & 4 & 4\\ 2 & 0 & -2\\ 1+i & -2-2i & 1+i\end{pmatrix}\begin{pmatrix}R\\G\\B\end{pmatrix}. \tag{2}$$

This implies $F_{C_2} = \frac{1+i}{12}R + \frac{-2-2i}{12}G + \frac{1+i}{12}B$, and hence we have $F^*_{C_2} = \frac{1-i}{12}R + \frac{-2+2i}{12}G + \frac{1-i}{12}B$. Note that the rows and columns of frequency structure are indexed by $(0, 1, \cdots, n_r - 1)$ and $(0, 1, \cdots, n_c - 1)$, which represent the frequency points of $2\pi\left(0, \frac{1}{n_r}, \cdots, \frac{n_r-1}{n_r}\right)$ and $2\pi\left(0, \frac{1}{n_c}, \cdots, \frac{n_c-1}{n_c}\right)$, respectively, where the size of the CFA pattern is $n_r \times n_c$.

B. CFA Design Methods in the Frequency Domain

The frequency representation of mosaicked images also allows us to understand the visual artifacts in demosaicked images as the aliasing between luma and modulated chromas [9]. Namely, if luma and modulated chromas overlap in the frequency domain, some frequency components contain the sum of luma and chromas instead of each of them separately. Then the demosaicking algorithm can produce visual artifacts when it recovers luma and chromas independently. Thus we consider that these artifacts are inherent to the CFAs. This motivates the design of CFAs by reducing the spectra overlap between luma and modulated chromas. According to the motivation, many CFA design methods in the frequency domain have been presented [7], [14], [15], [26].

Inspired by the spectral characteristic analysis of Bayer CFA [9], Hirakawa and Wolfe [14] proposed the first CFA design method in the frequency domain. Instead of directly using the RGB basis, they empirically chose G, R−G, and B−G as the basis to decorrelate the image channels. Let $\mathbf{c}(\mathbf{n}) = (c_R(\mathbf{n}), c_G(\mathbf{n}), c_B(\mathbf{n}))^T$ be the color pixel of the CFA at $\mathbf{n}$, where $\mathbf{n} \in \mathbb{Z}^2$ and $\mathbb{Z}$ denotes the set of integers. So it is physically realizable, i.e., it is real, non-negative and lies in [0, 1]. They further required that it satisfies $c_R(\mathbf{n}) + c_G(\mathbf{n}) + c_B(\mathbf{n}) = \gamma$. Let $\mathbf{x}(\mathbf{n}) = (x_R(\mathbf{n}), x_G(\mathbf{n}), x_B(\mathbf{n}))^T$ denote the color pixel of the full color image at $\mathbf{n}$, $x_{C_1} = x_R - x_G$, and $x_{C_2} = x_B - x_G$. Then the noise-free mosaicked image y would be:

$$\begin{aligned}
y(\mathbf{n}) &= \mathbf{c}(\mathbf{n})^T\mathbf{x}(\mathbf{n}) = \mathbf{c}(\mathbf{n})^T\mathbf{I}\mathbf{x}(\mathbf{n})\\
&= \mathbf{c}(\mathbf{n})^T\begin{pmatrix}1&1&0\\0&1&0\\0&1&1\end{pmatrix}\begin{pmatrix}1&-1&0\\0&1&0\\0&-1&1\end{pmatrix}\mathbf{x}(\mathbf{n})\\
&= \begin{pmatrix}c_R(\mathbf{n})\\ \gamma\\ c_B(\mathbf{n}))\end{pmatrix}^T\begin{pmatrix}x_{C_1}(\mathbf{n})\\ x_G(\mathbf{n})\\ x_{C_2}(\mathbf{n})\end{pmatrix}\\
&= \begin{pmatrix}c_R(\mathbf{n})\\ \gamma\\ c_B(\mathbf{n}))\end{pmatrix}^T\begin{pmatrix}1 & 0 & 0\\ -\frac{\mu_1}{\gamma} & 1 & -\frac{\mu_2}{\gamma}\\ 0 & 0 & 1\end{pmatrix}\\
&\quad\times\begin{pmatrix}1 & 0 & 0\\ \frac{\mu_1}{\gamma} & 1 & \frac{\mu_2}{\gamma}\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x_{C_1}(\mathbf{n})\\ x_G(\mathbf{n})\\ x_{C_2}(\mathbf{n})\end{pmatrix}\\
&= (c_R(\mathbf{n}) - \mu_1, \gamma, c_B(\mathbf{n}) - \mu_2)\begin{pmatrix}x_{C_1}(\mathbf{n})\\ x_L(\mathbf{n})\\ x_{C_2}(\mathbf{n})\end{pmatrix},
\end{aligned} \tag{3}$$

where $x_L(\mathbf{n}) = x_G(\mathbf{n}) + (\mu_1/\gamma)x_{C_1}(\mathbf{n}) + (\mu_2/\gamma)x_{C_2}(\mathbf{n})$ represent the luma, x_{C_1} and x_{C_2} represent the two chromas, and $(\cdot)^T$ denotes matrix transpose. So all the parameters are γ, μ_1, μ_2, and the Fourier coefficients of the Fourier transforms of c_R and c_B. They next conducted parameter search, so that the resultant CFA is physically realizable and the chromas are modulated far away from the luma. Minimizing the overlap between luma and chromas is achieved by enforcing a constraint during parameter search that chromas should be located at the spectrum border. They also empirically imposed that the red-green-blue ratio in luma should be $1:1:1$ or $1:2:1$. The spectrum of images mosaicked with their proposed CFA is shown in Fig. 1(2b). We can see that the modulated chromas are far away from the center and the horizontal and the vertical axes, where the luma has a high spectrum density. We can also see from Fig. 1(2a) that the modulated chromas of Bayer CFA overlap with the luma on the horizontal and the vertical axes.

Condat [15] followed the approach of Hirakawa and Wolfe [14]. However, he argued that for modern cameras the robustness of a CFA to noise is more important than to aliasing, especially in low-light conditions. So he proposed a new CFA that is robust to both aliasing and noise (Fig. 1(1c)). In comparison with the work of Hirakawa and Wolfe, he used an orthonormal basis: $L = (R+G+B)/\sqrt{3}$, $C_1 = (-R+2G-B)/\sqrt{6}$, and $C_2 = (R-B)/\sqrt{2}$, which is claimed to maximally decorrelate the image channels. So his model was simplified as:

$$\begin{aligned} y(\mathbf{n}) = \mathbf{c}(\mathbf{n})^T \tfrac{1}{6} \begin{pmatrix} 2 & -1 & 3 \\ 2 & 2 & 0 \\ 2 & -1 & -3 \end{pmatrix} \operatorname{diag}\begin{pmatrix} \sqrt{3} \\ \sqrt{6} \\ \sqrt{2} \end{pmatrix} \\ \operatorname{diag}\begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{6} \\ 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ -1 & 2 & -1 \\ 1 & 0 & -1 \end{pmatrix} \mathbf{x}(\mathbf{n}) \\ = (\gamma/\sqrt{3}, c_1(\mathbf{n}), c_2(\mathbf{n})) \left(x_L(\mathbf{n}), x_{C_1}(\mathbf{n}), x_{C_2}(\mathbf{n})\right)^T, \quad (4) \end{aligned}$$

where $c_1(\mathbf{n}) = (-c_R(\mathbf{n}) + 2c_G(\mathbf{n}) - c_B(\mathbf{n}))/\sqrt{6}$, $c_2(\mathbf{n}) = (c_R(\mathbf{n}) - c_B(\mathbf{n}))/\sqrt{2}$, and $\operatorname{diag}(\cdot)$ converts a vector to a diagonal matrix whose j-th diagonal element is the j-th element of the vector. Then he used a constructive approach to manually determine all the parameters step by step. Different from the other design methods, he selected parameters to simultaneously maximize the minimum distance between luma and chromas and the sensitivity of the CFA, which can reduce the noise level in demosaicked images. In order to maximally reduce the overlap between luma and chromas, he imposed that the two chromas are conjugate and each of them has only one replica. The designed 2×3 CFA is shown in Fig. 1(1c). It has six distinct color components. The spectrum of his CFA is shown in Fig. 1(2c).

Based on the frequency structure, Hao et al. [7] designed CFAs from a new perspective. The design of CFAs leaves many parameters to be chosen. Since the luma and the two chromas constitute a basis, there exists an invertible conversion between it and the RGB basis. Formally, we have the following relationship:

$$(F_L, F_{C_1}, F_{C_2})^T = \mathbf{M}\ (R, G, B)^T, \quad (5)$$

where F_L, F_{C_1}, and F_{C_2} denote the luma and the two chromas, respectively, R, G, and B refer to the red, green, and blue color components, respectively, $\mathbf{M} \in \mathbb{C}^{3\times 3}$ is invertible and is called the color transformation matrix, and $\mathbb{C}$ denotes the set of complex numbers. In frequency selection based demosaicking, the RGB full color image is recovered from the estimated F_L, F_{C_1}, and F_{C_2} via solving (5). However, the estimations of F_L, F_{C_1}, and F_{C_2} contain errors. Accordingly, one should control the error in demosaicked images that results from the estimation errors. Formally, we denote $\mathbf{y} = (\Delta F_L, \Delta F_{C_1}, \Delta F_{C_2})^T$ as the estimation errors and $\mathbf{x} = (\Delta R, \Delta G, \Delta B)^T$ as the error that results from $\mathbf{y}$. Then according to (5), we have $\mathbf{y} = \mathbf{M}\mathbf{x}$. Consequently, the amplification factor of estimation errors is:

$$\frac{\|\mathbf{x}\|_2}{\|\mathbf{y}\|_2} = \frac{\|\mathbf{M}^{-1}\mathbf{y}\|_2}{\|\mathbf{y}\|_2} \le \max_{\mathbf{y}\neq\mathbf{0}} \frac{\|\mathbf{M}^{-1}\mathbf{y}\|_2}{\|\mathbf{y}\|_2} = \|\mathbf{M}^{-1}\|_2, \quad (6)$$

where $\mathbf{M}^{-1}$ is the inverse of $\mathbf{M}$, $\|\mathbf{M}^{-1}\|_2$ is the spectral norm of $\mathbf{M}^{-1}$ which is its largest singular value, and $\|\mathbf{x}\|_2$ is the l_2 norm of vector $\mathbf{x}$. This implies that decreasing $\|\mathbf{M}^{-1}\|_2$ can greatly enhance the numerical stability of color transformation. With the help of frequency structure, they formulated parameter optimization as a constrained optimization problem to maximize the numerical stability of the color transformation. Meanwhile, the problem of minimizing the aliasing between luma and chromas is converted into a frequency structure selection problem. For a selected frequency structure, Hao et al. formulated the parameter optimization problem as follows:

$$\begin{aligned} &\min_{\mathbf{M}} \ \|\mathbf{M}^{-1}\|_F \\ &\text{s.t. } c_j \in [0,1], \textstyle\sum_j c_j = \mathbf{1},\ j \in \{R, G, B\}, \end{aligned} \quad (7)$$

where $\mathbf{1}$ denotes the all-one matrix, c_R, c_G, and c_B denote the three channels of the CFA, and $\|\mathbf{M}^{-1}\|_F$ is the Frobenius norm of $\mathbf{M}^{-1}$ to approximate $\|\mathbf{M}^{-1}\|_2$. They further imposed that $\mathbf{M}$ should be real, which implies that the frequency structures cannot contain conjugate replicas of the chromas. Then they proposed a geometric method to solve (7). Although they provided several guidelines for manual frequency structure choice, the computation for all the candidates still requires immense resources for a reasonably sized CFA pattern. Moreover, the proposed geometric method needs the user to specify the optimal triangle, which contains the origin as its inner point and minimizes $\|\mathbf{M}^{-1}\|_F$.

III. Proposed Automatic CFA Design Method

In this section, we first outline the proposed method. Then we describe each step in detail.

A. Method Overview

As shown in Fig. 2, our CFA design process consists of two main steps. For a given CFA pattern size, we first perform multi-objective optimization to propose frequency structure candidates (the dashed box in Fig. 2), where the minimum distances between their respective frequency points are maximized (Fig. 2(b)-2(c)) and the replicate relations among the chromas are specified (Fig. 2(d)). Then we optimize

parameters for each frequency structure candidate to obtain the corresponding CFA (Fig. 2(e)). A CFA pattern size does not always result in a unique frequency structure as well as CFA (see Fig. 2(d) and 2(e)). So we finally output the CFA that has the best demosaicking performance on a training image set (Fig. 2(f)). Below we explain the two steps in detail.

For a given CFA pattern size, maximizing the minimum distance between frequency components in the frequency domain is equivalent to finding a frequency structure that the minimum distance between its frequency points is maximized [7]. Exhaustive tests of all the feasible frequency structures are computationally intractable even for a moderate CFA pattern size. For example, for a 5×5 CFA pattern, if we exclude the horizontal and the vertical axes of luma, the number of different chroma position allocations is 255, and the number of frequency structures further grows significantly. One approach to compromise computation and optimality of the produced CFA is to use a fast but approximate method to discard a majority of unpromising ones. Since missed frequency structures cannot be recovered in the subsequent stage, it is also important to contain all the possibly optimal ones. To account for these requirements, we formulate the generation of frequency structure candidates as a multi-objective optimization problem. We will describe it in subsection III-B.

Similar to Hao et al. [7], for a given frequency structure candidate, we formulate the parameter optimization as a constrained optimization problem. However, we enhance the computational stability of color transformation by directly minimizing $\|\mathbf{M}^{-1}\|_2$, rather than its approximation $\|\mathbf{M}^{-1}\|_F$. Also, we enforce that the CFA is physically realizable and the sum across its color channels is an all-one matrix. Besides that, we make no assumptions on frequency structure and CFA. We will introduce our model in subsection III-C and the solution process in the Appendix.

B. Propose Frequency Structure Candidates

For a given CFA pattern size, we argue that the minimum distance between luma and chromas as well as between chromas of the frequency structure should be as large as possible. They are our first two objectives. If the given size of a CFA pattern is larger than 2×2, all chromas should not locate on the horizontal and the vertical axes of luma. Moreover, with redundant chroma replicas, we can estimate each chroma more accurately by fusing all its estimations adaptively [7], [24]. So the number of chroma replicas should also be as large as possible [7], which is our third objective. The three objectives are in conflict and hence we cannot find a single solution that is optimal for all of them. We propose a multi-objective optimization approach to find an appropriately balanced solution.

1) Multi-Objective Optimization: Multi-objective optimization [27] refers to the simultaneous minimization or maximization of more than one objective functions. More formally, it studies the problem as follows:

$$\max_{\mathbf{x}} \ \{f_1(\mathbf{x}), f_2(\mathbf{x}), \cdots, f_m(\mathbf{x})\}, \quad \text{s.t. } \mathbf{x} \in \Omega, \tag{8}$$

where we have $m \geq 2$ objective functions f_j and want to maximize all the functions simultaneously, $\mathbf{x}$ is the decision variable, and Ω is the feasible region which can be formed by various constraints. Note that we assume that all the objective functions are to be maximized for simplicity. If an objective function f_j is to be minimized, it is equivalent to maximizing the function $-f_j$.

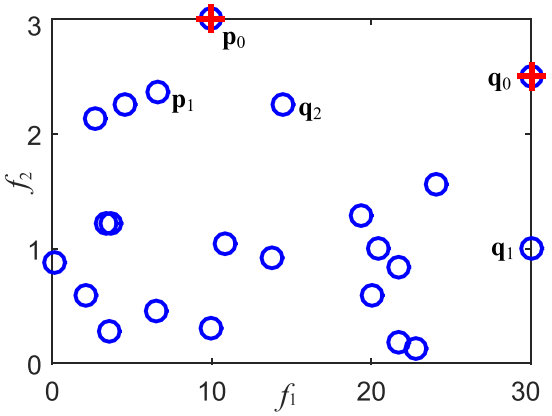


Fig. 3. Pareto optimal solutions of a multi-objective optimization problem with two objectives. The blue circles are the feasible points, while the red crosses are the Pareto optimal solutions.

The objective functions can be incommensurable, i.e., in different units. For example, in Fig. 3, $f_1 \in [0, 30]$ and $f_2 \in [0, 3]$ have different value ranges. Also, there is only partial ordering in the objective space, e.g., we cannot compare $(f_1(\mathbf{x}_1), f_2(\mathbf{x}_1))^T = (3, 2.5)^T$ with $(f_1(\mathbf{x}_2), f_2(\mathbf{x}_2))^T = (2, 3)^T$. Furthermore, in general, there may be partial conflicts among the objective functions, i.e., maximizing one function can decrease the values of the others. Because of the possible incommensurability and conflict among the objective functions, it is not possible to composite a global objective function as a weighted sum of all the objective functions, or find a single solution that is optimal w.r.t. every objective function. The solutions of a multi-objective optimization problem are called *Pareto optimal* solutions. We state a more formal definition in the following:

Definition 1: A decision variable $\mathbf{x}_1$ is said to be *dominated* by $\mathbf{x}_2$ if $f_j(\mathbf{x}_1) \leq f_j(\mathbf{x}_2)$ for all $j = 1, 2, \ldots, m$ and $f_k(\mathbf{x}_1) < f_k(\mathbf{x}_2)$ for at least one index k.

For example, in Fig. 3, $\mathbf{p}_1$ is dominated by $\mathbf{p}_0$, and $\mathbf{q}_1$ and $\mathbf{q}_2$ are both dominated by $\mathbf{q}_0$. Since $f_2(\mathbf{p}_0) > f_2(\mathbf{q}_0)$ and $f_1(\mathbf{p}_0) < f_1(\mathbf{q}_0)$, $\mathbf{p}_0$ and $\mathbf{q}_0$ are not dominated by each other.

Definition 2: A decision variable $\mathbf{x}^* \in \Omega$ is *Pareto optimal* if $\mathbf{x}^*$ cannot be dominated by any variable $\mathbf{x} \in \Omega$.

In Fig. 3, $\mathbf{p}_0$ and $\mathbf{q}_0$ cannot be dominated by any other feasible points. So they are both Pareto optimal to the problem. All the Pareto optimal solutions constitute the *Pareto optimal set* of the problem, e.g., $\{\mathbf{p}_0, \mathbf{q}_0\}$ is the Pareto optimal set to the multi-objective optimization problem illustrated in Fig. 3.

2) Obtain Chroma Position Candidates: Note that the only luma is fixed at the top-left of frequency (0,0) in the frequency structure (see Fig. 2(a)). So we only need to choose the replicas of the two chromas and their positions in the matrix to finally determine a frequency structure. As noted before, the rows and columns of frequency structure are indexed by $(0, 1, \cdots, n_r - 1)$ and $(0, 1, \cdots, n_c - 1)$, which

represent the frequency points of $2\pi\left(0, \frac{1}{n_r}, \cdots, \frac{n_r-1}{n_r}\right)$ and $2\pi\left(0, \frac{1}{n_c}, \cdots, \frac{n_c-1}{n_c}\right)$, respectively, where $n_r \times n_c$ is the CFA pattern size. In the following discussion, we omit 2π from all frequency points for simplicity. Since the designed CFA is real, once the position of a chroma frequency point $(n_x/n_c, n_y/n_r)$ in the frequency structure is chosen, the position $((1-n_x/n_c) \bmod 1, (1-n_y/n_r) \bmod 1)$ must also be chosen [7], where mod is the modulo operation, $n_x \in \{0, 1, \cdots, n_c - 1\}$, and $n_y \in \{0, 1, \cdots, n_r - 1\}$. If the two positions are different, we call them a conjugate position pair, otherwise we say that the position is self-conjugate, e.g., $(\frac{1}{2}, \frac{1}{2})$, or $(\frac{1}{2}, 0)$. If the matrix has m_p conjugate position pairs and m_s self-conjugate positions, there are $2^{m_p+m_s} - m_s - 1$ feasible chroma position allocations. Also, if the CFA pattern size is larger than 2×2, we first discard those allocations that contain chroma positions on the horizontal and the vertical axes of luma. Then we perform multi-objective optimization on the rest of allocations.

More formally, the multi-objective optimization problem is:

$$\max_{\mathbf{x}} \ \{f_1(\mathbf{x}), f_2(\mathbf{x}), f_3(\mathbf{x})\}$$
$$\text{s.t. } \mathbf{x} \in \text{the set of feasible chroma position allocations,} \quad (9)$$

where f_1 denotes "the minimum distance between luma and chroma positions", f_2 denotes "the minimum distance between chroma positions", and f_3 denotes "the number of chroma replicas". Since frequency structure is periodic in both horizontal and vertical directions (please read the caption of Fig. 1), we compute the distance between two positions in it as follows. Suppose the two positions are (x_1, y_1) and (x_2, y_2). Then the distances along the horizontal and the vertical directions are $d_x = \min(|x_1 - x_2|, 1 - |x_1 - x_2|)$ and $d_y = \min(|y_1 - y_2|, 1 - |y_1 - y_2|)$, respectively, where $|x|$ is the absolute value of the scalar x. So the Euclidean distance between the two positions is $\sqrt{d_x^2 + d_y^2}$. We take the frequency structure F_{H} in (1) as an example. The distance between F_L and F_{C_1} is $\left(\min(1/2, 1-1/2)^2 + \min(2/4, 1-2/4)^2\right)^{1/2} = \sqrt{2}/2$. The distance between F_L and $F_{C_2}^*$ is $\left(\min(1/2, 1-1/2)^2 + \min(1/4, 1-1/4)^2\right)^{1/2} = \sqrt{5}/4$. The distance between F_L and F_{C_2} is $\left(\min(1/2, 1-1/2)^2 + \min(3/4, 1-3/4)^2\right)^{1/2} = \sqrt{5}/4$. So $f_1(F_{\mathrm{H}})$ is $\min(\sqrt{2}/2, \sqrt{5}/4) = \sqrt{5}/4$. Similarly, we can compute the $f_2(F_{\mathrm{H}})$. Thus solving problem (9) is equivalent to finding the Pareto optimal set from a given point set (see Fig. 3). We use the non-dominated sorting scheme to solve it [27]. The objective value of f_1 for the Bayer CFA is 0.5. Since f_1 is more important than f_2 and f_3, we reject the chroma position candidates whose objective values of f_1 below 0.5.

3) Generate Frequency Structure: We generate all the frequency structures according to the chroma position candidates. For each candidate, we divide its selected positions into two non-overlapping groups. The two position groups are for the replicas of F_{C_1} and F_{C_2}, respectively. It is important to note that F_{C_1} and F_{C_2} are symmetric, i.e., swapping them does not result in a new frequency structure. Then without loss of generality, we only assume equal or conjugate replicas of a chroma, i.e., the replicas of a chroma C are all in $\{C, C^*\}$. It may produce multiple frequency structures (see Fig. 2(d)).

C. Optimize Parameters

Following [7], we parameterize the complex color transformation matrix $\mathbf{M}$ as $\mathbf{M}_1 + i\mathbf{M}_2$, where $\mathbf{M}_1$ and $\mathbf{M}_2$ are the real and imaginary parts of $\mathbf{M}$, respectively, and they are both real. Then F_L, F_{C_1}, and F_{C_2} can be linearly parameterized by $\mathbf{M}$. We apply the inverse symbolic DFT to the parameterized frequency structure and obtain the *vectorized* CFA pattern denoted by $\mathbf{CM}_1 + \mathbf{DM}_2$, where $\mathbf{C}$ and $\mathbf{D}$ are the complex coefficient matrices for $\mathbf{M}_1$ and $\mathbf{M}_2$, respectively. Let $\mathbf{c}_j$ be the j-th channel of the RGB CFA pattern with a size of $n_r \times n_c$, where $j \in \{R, G, B\}$. The vectorized CFA pattern is $(\mathrm{vec}(\mathbf{c}_R), \mathrm{vec}(\mathbf{c}_G), \mathrm{vec}(\mathbf{c}_B))$ with a size of $n_r n_c \times 3$, where $\mathrm{vec}(\cdot)$ is the operator to convert a matrix into a vector.

We take the frequency structure of Hirakawa CFA [14] as an example. We first write the color transformation in (5) in more detail:

$$\begin{pmatrix} F_L \\ F_{C_1} \\ F_{C_2} \end{pmatrix} = \begin{pmatrix} M_{11}^{(1)} + iM_{11}^{(2)} & M_{12}^{(1)} + iM_{12}^{(2)} & M_{13}^{(1)} + iM_{13}^{(2)} \\ M_{21}^{(1)} + iM_{21}^{(2)} & M_{22}^{(1)} + iM_{22}^{(2)} & M_{23}^{(1)} + iM_{23}^{(2)} \\ M_{31}^{(1)} + iM_{31}^{(2)} & M_{32}^{(1)} + iM_{32}^{(2)} & M_{33}^{(1)} + iM_{33}^{(2)} \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}, \quad (10)$$

where the superscripts (1) and (2) indicate that the elements are from $\mathbf{M}_1$ and $\mathbf{M}_2$, respectively. The conjugate of F_{C_2} is given as:

$$F_{C_2}^* = \left(M_{31}^{(1)} - iM_{31}^{(2)}, \quad M_{32}^{(1)} - iM_{32}^{(2)}, \quad M_{33}^{(1)} - iM_{33}^{(2)}\right)\mathbf{P}, \quad (11)$$

where $\mathbf{P} = (R, G, B)^T$. Then we substitute (10) and (11) into the frequency structure of Hirakawa CFA shown in Fig. 1(3b). We next apply the inverse symbolic DFT to the frequency structure and we have the expression shown at the bottom of this page. So the vectorized CFA pattern in the RGB basis can

$$\begin{pmatrix} ((1,1,2)\mathbf{M}_1 + (i,i,0)\mathbf{M}_2)\mathbf{P}, & ((1,-1,-2)\mathbf{M}_1 + (i,-i,0)\mathbf{M}_2)\mathbf{P} \\ ((1,-1,0)\mathbf{M}_1 + (i,-i,2)\mathbf{M}_2)\mathbf{P}, & ((1,1,0)\mathbf{M}_1 + (i,i,-2)\mathbf{M}_2)\mathbf{P} \\ ((1,1,-2)\mathbf{M}_1 + (i,i,0)\mathbf{M}_2)\mathbf{P}, & ((1,-1,2)\mathbf{M}_1 + (i,-i,0)\mathbf{M}_2)\mathbf{P} \\ ((1,-1,0)\mathbf{M}_1 + (i,-i,-2)\mathbf{M}_2)\mathbf{P}, & ((1,1,0)\mathbf{M}_1 + (i,i,2)\mathbf{M}_2)\mathbf{P} \end{pmatrix}$$

be denoted by $\mathbf{CM}_1 + \mathbf{DM}_2$ with a size of 8×3, where

$$\mathbf{C} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \\ 1 & -1 & -2 \\ 1 & 1 & 0 \\ 1 & -1 & 2 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} i & i & 0 \\ i & -i & 2 \\ i & i & 0 \\ i & -i & -2 \\ i & -i & 0 \\ i & i & -2 \\ i & -i & 0 \\ i & i & 2 \end{pmatrix}.$$

The produced CFA pattern in the RGB basis should be physically realizable, i.e., $\mathbf{CM}_1+\mathbf{DM}_2$ is real and lies in [0, 1]. Also, the sum across color channels of CFA pattern should be an all-one matrix, i.e., the vectorized CFA pattern satisfies $(\mathbf{CM}_1 + \mathbf{DM}_2)(1, 1, 1)^T = \mathbf{1}$. Accordingly, we propose the following parameter optimization model:

$$\begin{aligned} &\min_{\mathbf{M}} \|\mathbf{M}^{-1}\|_2 \\ &\text{s.t. } \mathbf{C}\Re(\mathbf{M})+\mathbf{D}\Im(\mathbf{M}) \geq \mathbf{0}, (\mathbf{C}\Re(\mathbf{M})+\mathbf{D}\Im(\mathbf{M}))\mathbf{a}=\mathbf{e}, \end{aligned} \tag{12}$$

where $\mathbf{M}^{-1}$ is the inverse of $\mathbf{M}$, $\mathbf{a} = (1, 1, 1)^T$, $\mathbf{e} = \mathbf{1}_{n_r n_c \times 1}$, $\geq$ stands for componentwise greater than or equal to, $\mathbf{0}$ denotes the zero matrix, $\mathbf{1}$ denotes the matrix whose elements are all 1, and $\Re(\cdot)$ and $\Im(\cdot)$ are the linear operators to extract the real and the imaginary parts of a complex vector or matrix, e.g., $\Re(\mathbf{M}) = \Re(\mathbf{M}_1 + i\mathbf{M}_2) = \mathbf{M}_1$ and $\Im(\mathbf{M}) = \mathbf{M}_2$.

As noted in [7], the constraint $(\mathbf{C}\Re(\mathbf{M}) + \mathbf{D}\Im(\mathbf{M}))\mathbf{a} = \mathbf{e}$ in (12) is equivalent to a simpler one: $\mathbf{Ma} = \mathbf{b}$, where $\mathbf{b} = (1, 0, 0)^T$. So we reformulate (12) into an equivalent one:

$$\min_{\mathbf{M}} \|\mathbf{M}^{-1}\|_2, \text{ s.t. } \mathbf{C}\Re(\mathbf{M})+\mathbf{D}\Im(\mathbf{M}) \geq \mathbf{0}, \mathbf{Ma}=\mathbf{b}. \tag{13}$$

In order to improve the readability, we move the details of solving for (13) to the Appendix.

IV. Experiments

In this section, we carry out experiments to validate the effectiveness of our proposed CFA design method. We first compare our parameter optimization method with those of other CFA design methods. Then we produce a new CFA using our design method and compare it with other CFAs. All the experiments are performed on the commonly used Kodak PhotoCD dataset [16].

A. Experimental Settings

1) Compared Design Methods and CFAs: We test our CFA design method by comparing with three existing CFA design methods in the frequency domain. They are proposed by Hirakawa and Wolfe [14], Condat [15], and Hao et al. [7], respectively. Since the corresponding three CFAs have been shown to outperform the existing CFAs, we only involve the three CFAs and the most popular Bayer CFA for comparison.

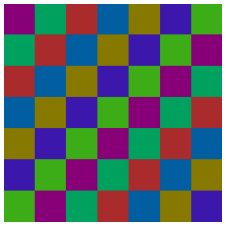
2) Frequency Selection Based Demosaicking: We briefly describe the main steps of frequency selection based demosaicking mentioned in subsection II-A. The different frequency locations of luma and modulated chromas allow us to estimate them directly from the mosaicked image by selecting corresponding frequencies [9], [15], [24], [28]. So these methods can be viewed as demultiplexing. They first use appropriate bandpass filters to estimate the modulated chromas. Then they demodulate the estimations. If there exist multiple estimations of a chroma, one can fuse them properly to obtain a more robust one. The luma is estimated by subtracting the remodulated chromas from the mosaicked image. The RGB full color image is finally recovered using the inverse transformation of (5).

In order to be fair, when we compare the parameter optimization of all design methods, we adapt the least-square luma-chroma demultiplexing algorithm (LSLCD) [28] to all the compared CFAs, in which the filters are learned to minimize the mean-squared demosaicking error over a training set. LSLCD was first proposed for the Bayer CFA, the source code of which is publicly available.[3] It was extended to arbitrary regular and periodic CFAs in [29]. Note that if a chroma has more than one replica, an adaptive method can give a more accurate estimate of the chroma. For example, Dubois [24] observed that in the Bayer CFA the two modulated replicas of C_2 suffer differently from the overlap between luma and chromas (see Fig. 1(2a)), i.e., one replica mainly overlaps with high horizontal frequencies of luma, while the other one mainly overlaps with the high vertical frequencies. Thus demosaicking algorithm can exploit this characteristic to reduce the impact of spectral overlap. However, we focus only on evaluating the effectiveness of parameter optimization methods instead of the CFAs. So if a chroma has multiple replicas, we simply average all its estimations. As in [29], frequency selection based demosaicking learns the filters to minimize the mean-squared demosaicking error over a training set. So one needs to specify the filter size and the training image set. Following [15], we use a 13×13 filter size and learn the filters on all 24 images in the Kodak dataset. Accordingly, we exclude a 13-pixel border (the filter width) to eliminate the boundary effects.

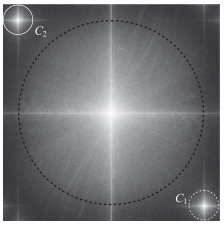
When comparing different CFAs, we use their respective associated demosaicking algorithms, with default parameters set by their respective authors. These parameters have been optimized on the Kodak dataset. For the Bayer CFA, we use the LSLCD specified for it, where an adaptive weighting strategy is employed [28]. For the Hirakawa CFA and the Condat CFA, we use the demosaicking algorithm developed by Condat [15], the source code of which is publicly available.[4] For the Hao CFA, we use the code provided by the authors. They also apply the adaptive weighting scheme [24] to combine multiple estimations of a chroma. We adapt the LSLCD to our designed CFAs. If a chroma has multiple replicas, we simply average all its estimations. Also, we exclude a 13-pixel border when computing CPSNR values. We want to note that *the demosaicking algorithms also highly affect the image quality, especially the combination strategies for CFAs that have more than one replica of a chroma [7], [24]. However, an extensive comparison of demosaicking algorithms is out of the scope of this work.*

[3] http://www.site.uottawa.ca/∼edubois/lslcd/.

[4] http://www.gipsa-lab.grenoble-inp.fr/∼laurent.condat/publications.html.



(1)



(2)

$$\begin{pmatrix} F_L & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & F_{C_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & F_{C_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(3)

$$\mathbf{R} = \begin{pmatrix} 0.52988 & 0.00000 & 0.66075 & 0.00000 & 0.52988 & 0.23581 & 0.23581 \\ 0.00000 & 0.66075 & 0.00000 & 0.52988 & 0.23581 & 0.23581 & 0.52988 \\ 0.66075 & 0.00000 & 0.52988 & 0.23581 & 0.23581 & 0.52988 & 0.00000 \\ 0.00000 & 0.52988 & 0.23581 & 0.23581 & 0.52988 & 0.00000 & 0.66075 \\ 0.52988 & 0.23581 & 0.23581 & 0.52988 & 0.00000 & 0.66075 & 0.00000 \\ 0.23581 & 0.23581 & 0.52988 & 0.00000 & 0.66075 & 0.00000 & 0.52988 \\ 0.23581 & 0.52988 & 0.00000 & 0.66075 & 0.00000 & 0.52988 & 0.23581 \end{pmatrix}$$

$$\mathbf{G} = \begin{pmatrix} 0.00000 & 0.63045 & 0.16963 & 0.36955 & 0.47012 & 0.08898 & 0.67521 \\ 0.63045 & 0.16963 & 0.36955 & 0.47012 & 0.08898 & 0.67521 & 0.00000 \\ 0.16963 & 0.36955 & 0.47012 & 0.08898 & 0.67521 & 0.00000 & 0.63045 \\ 0.36955 & 0.47012 & 0.08898 & 0.67521 & 0.00000 & 0.63045 & 0.16963 \\ 0.47012 & 0.08898 & 0.67521 & 0.00000 & 0.63045 & 0.16963 & 0.36955 \\ 0.08898 & 0.67521 & 0.00000 & 0.63045 & 0.16963 & 0.36955 & 0.47012 \\ 0.67521 & 0.00000 & 0.63045 & 0.16963 & 0.36955 & 0.47012 & 0.08898 \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} 0.47012 & 0.36955 & 0.16962 & 0.63045 & 0.00000 & 0.67521 & 0.08898 \\ 0.36955 & 0.16962 & 0.63045 & 0.00000 & 0.67521 & 0.08898 & 0.47012 \\ 0.16962 & 0.63045 & 0.00000 & 0.67521 & 0.08898 & 0.47012 & 0.36955 \\ 0.63045 & 0.00000 & 0.67521 & 0.08898 & 0.47012 & 0.36955 & 0.16962 \\ 0.00000 & 0.67521 & 0.08898 & 0.47012 & 0.36955 & 0.16962 & 0.63045 \\ 0.67521 & 0.08898 & 0.47012 & 0.36955 & 0.16962 & 0.63045 & 0.00000 \\ 0.08898 & 0.47012 & 0.36955 & 0.16962 & 0.63045 & 0.00000 & 0.67521 \end{pmatrix}$$

(4)

Fig. 4. The proposed Diag7 CFA pattern. (1) is the proposed 7×7 diagonal stripe CFA pattern. (2)-(4) are its spectrum, frequency structure, and color values, respectively. The notations in (2) and (3) are the same as those in Fig. 1.

TABLE II

EVALUATION OF THE NUMERICAL STABILITY OF COLOR TRANSFORMATIONS. THREE EXISTING CFA DESIGN METHODS ARE COMPARED WITH THE PROPOSED ONE IN PARAMETER OPTIMIZATION. "ORIG." STANDS FOR "ORIGINAL". NOTE THAT OUR METHOD PRODUCES THE IDENTICAL **M** AS THAT OF THE CONDAT CFA [15] AND THE HAO CFA [7], RESPECTIVELY

Numerical Stability	Hirakawa [14]		Condat [15]		Hao [7]	
	Orig.	Ours	Orig.	Ours	Orig.	Ours
$\|\mathbf{M}^{-1}\|_2$	4.24	**3.73**	3.00	3.00	5.66	5.66

TABLE III

EVALUATION OF THE PROPOSED PARAMETER OPTIMIZATION METHOD ON THE KODAK PHOTOCD DATASET. "AVG." STANDS FOR "AVERAGE". THE INDIVIDUAL AND AVERAGE CPSNR VALUES ARE REPORTED. NOTE THAT OUR NEWLY DESIGNED CFAS ARE IDENTICAL TO THE CONDAT CFA [15] AND THE HAO CFA [7], RESPECTIVELY. SO THEIR CPSNR VALUES ARE THE SAME

Image ID	Hirakawa [14]		Condat [15]		Hao [7]	
	Orig.	Ours	Orig.	Ours	Orig.	Ours
01	39.78	**39.89**	39.69	39.69	40.78	40.78
02	40.69	**40.83**	40.41	40.41	41.66	41.66
03	41.17	**41.23**	42.21	42.21	40.00	40.00
04	40.04	**40.17**	40.66	40.66	41.87	41.87
05	36.87	**36.96**	37.15	37.15	36.34	36.34
06	41.29	**41.51**	40.83	40.83	41.15	41.15
07	41.18	**41.24**	42.09	42.09	40.20	40.20
08	37.67	**37.79**	37.45	37.45	37.84	37.84
09	41.94	**42.06**	42.26	42.26	40.85	40.85
10	42.42	**42.53**	42.82	42.82	42.18	42.17
11	40.48	**40.68**	40.31	40.31	40.96	40.96
12	42.86	**42.96**	43.62	43.62	43.14	43.14
13	35.58	**35.79**	35.19	35.19	35.72	35.72
14	35.94	**36.14**	36.07	36.07	35.07	35.07
15	39.38	**39.56**	39.75	39.75	40.28	40.28
16	44.79	**45.01**	44.66	44.66	43.99	43.98
17	41.69	**41.86**	41.24	41.24	41.70	41.70
18	37.64	**37.78**	37.56	37.56	37.28	37.28
19	41.59	**41.75**	41.29	41.29	41.31	41.31
20	41.59	**41.83**	41.24	41.24	40.55	40.55
21	40.60	**40.81**	40.22	40.22	40.22	40.22
22	38.47	**38.58**	38.74	38.74	38.11	38.11
23	41.92	**42.03**	42.05	42.05	40.75	40.75
24	36.57	**36.71**	36.29	36.29	36.39	36.39
Avg.	40.09	**40.24**	40.16	40.16	39.93	39.93

B. Comparison of Parameter Optimization

For a given CFA, we can determine its coefficient matrices **C** and **D** as well as color transformation matrix **M**. However, for a given frequency structure, we can only compute its **C** and **D**. We have to use a parameter optimization method to obtain an **M**. Then with the **C**, **D**, and **M**, we can determine the CFA. This gives us a chance to compare with other CFA design methods in parameter optimization.

We first compute the frequency structure of a CFA (e.g., the Hirakawa CFA [14]) produced by an existing CFA design method in the frequency domain, and with which we can calculate the corresponding **C** and **D**. Then we obtain a new **M** with our parameter optimization method and also a newly designed CFA. We next compare the newly designed CFA with the original one.

We first compare the numerical stability of color transformations. The values of $\|\mathbf{M}^{-1}\|_2$ are reported in Table II. A smaller value may indicate more numerical stability. From left to right, the three group comparisons are corresponding to the Hirakawa CFA [14], the Condat CFA [15], and the Hao CFA [7], respectively. In each group, the better value is

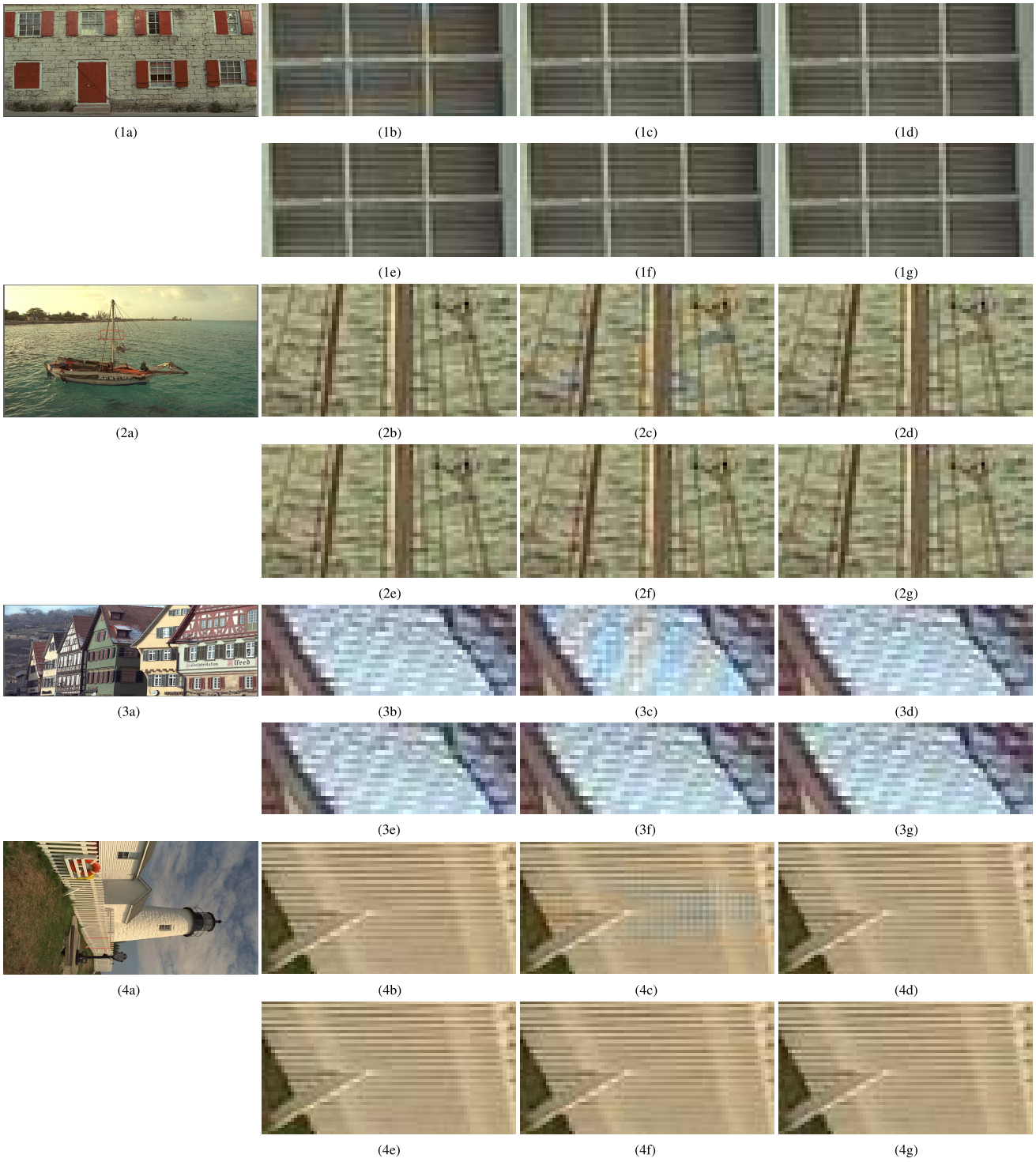


Fig. 5. Blowups of some demosaicked images in the Kodak dataset. From top to bottom, the images are from #1, #6, #8, and #19 images of the Kodak dataset, respectively. In each group, (a) is the scaled original image, in which the red rectangle indicates the selected patch to blow up; (b) is the ground truth; (c) is the image demosaicked from the image mosaicked with the Bayer CFA; (d)-(g) are the images demosaicked from the images mosaicked with the optimized CFAs. From all the four groups of images, we can clearly see that the images demosaicked from the images mosaicked with the Bayer CFA have severe false color artifacts, while those by the optimized CFAs have better subjective quality. (1a) Scaled original image. (1b) Ground truth. (1c) Bayer CFA [8]. (1d) Hirakawa CFA [14]. (1e) Condat CFA [15]. (1f) Hao CFA [7]. (1g) Diag7. (2a) Scaled original image. (2b) Ground truth. (2c) Bayer CFA [8]. (2d) Hirakawa CFA [14]. (2e) Condat CFA [15]. (2f) Hao CFA [7]. (2g) Diag7. (3a) Scaled original image. (3b) Ground truth. (3c) Bayer CFA [8]. (3d) Hirakawa CFA [14]. (3e) Condat CFA [15]. (3f) Hao CFA [7]. (3g) Diag7. (4a) Scaled original image. (4b) Ground truth. (4c) Bayer CFA [8]. (4d) Hirakawa CFA [14]. (4e) Condat CFA [15]. (4f) Hao CFA [7]. (4g) Diag7.

in boldface. Note that the $\|\mathbf{M}^{-1}\|_2$ cannot be directly compared across different CFAs since it increases with the size of CFA pattern and the number of chroma replicas. We can see that our new color transformation system generated from the frequency structure of the Hirakawa CFA [14] is more stable than that of the original one (see the first group of Table II).

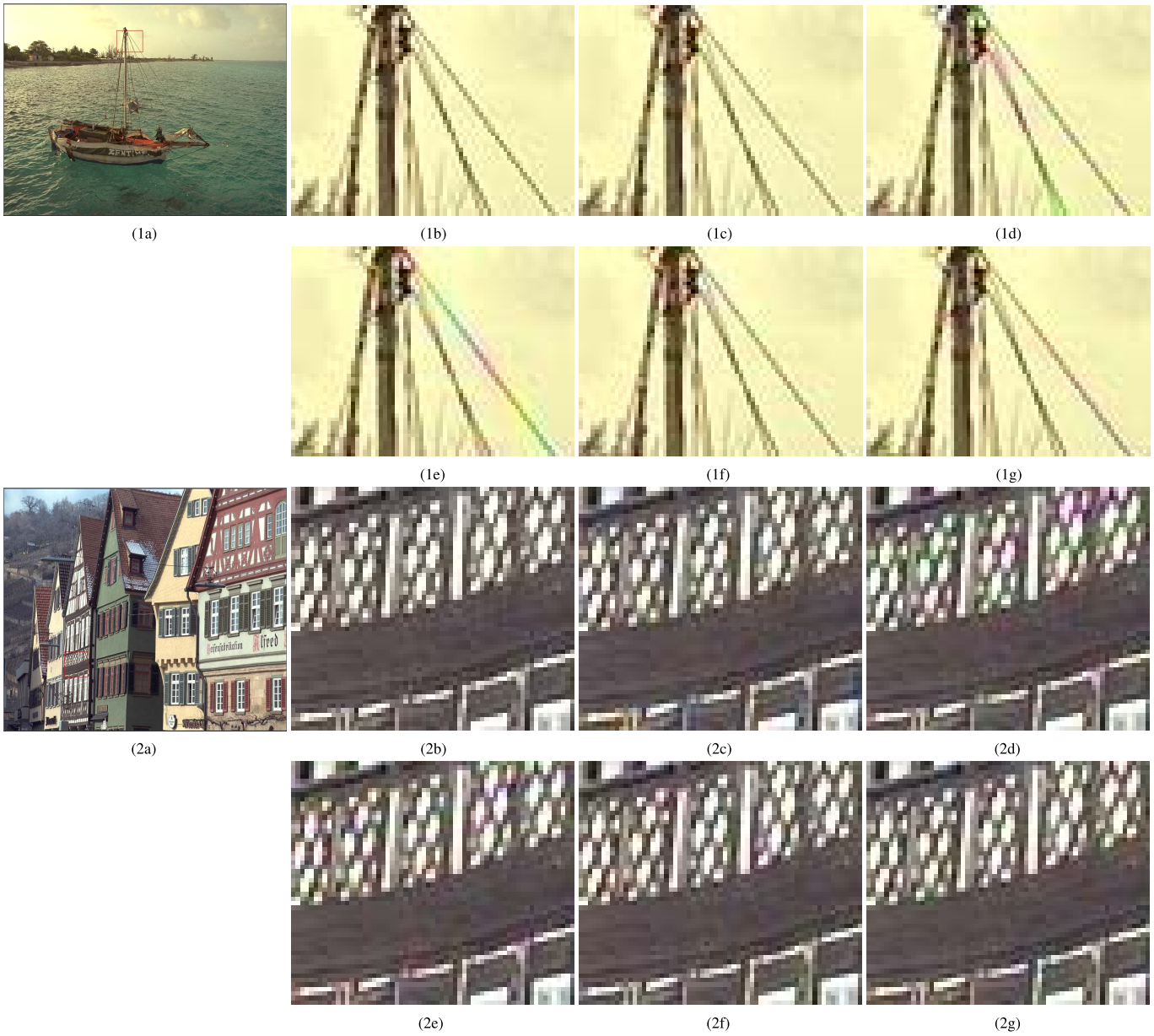


Fig. 6. Two more blowups of demosaicked images in the Kodak dataset. From top to bottom, the images are from #6 and #8 images of the Kodak dataset, respectively. From the two groups of images, we can see that our Diag7 CFA results in similar or better visual quality as that of other CFAs. (1a) Scaled original image. (1b) Ground truth. (1c) Bayer CFA [8]. (1d) Hirakawa CFA [14]. (1e) Condat CFA [15]. (1f) Hao CFA [7]. (1g) Diag7. (2a) Scaled original image. (2b) Ground truth. (2c) Bayer CFA [8]. (2d) Hirakawa CFA [14]. (2e) Condat CFA [15]. (2f) Hao CFA [7]. (2g) Diag7.

Our method results in the identical $\mathbf{M}$ as that of the Condat CFA [15] and the Hao CFA [7], respectively. So their values of $\|\mathbf{M}^{-1}\|_2$ are the same (see the last two groups of Table II). It implies that the parameters of the Condat CFA and the Hao CFA are already optimal for their respective frequency structures. However, we want to note that the frequency structures of the Condat CFA and the Hao CFA do not contain conjugate chroma replicas (see Fig. 1(3c) and (3d)), whereas the one of the Hirakawa CFA has a conjugate chroma replica (see Fig. 1(3b)).

Then we compare the newly designed CFAs with their respective original ones on the Kodak dataset using the same demosaicking algorithm. The individual and average CPSNR values are given in Table III. The better values in each group are in boldface. Similar to the above analysis of $\mathbf{M}$, our newly designed CFA according to the frequency structure of the Hirakawa CFA performs better than the original one on both individual image and the whole dataset (see the first group of Table III). Also, our newly designed CFAs are identical to the Condat CFA and the Hao CFA, respectively. So they have the same CPSNR values (see the last two groups of Table III).

C. Comparison of CFA Design

We first generate a new CFA using our design method and then compare it with the others on the Kodak dataset.

TABLE IV

EVALUATION OF THE PROPOSED CFA ON THE KODAK PHOTOCD DATASET. FOUR EXISTING CFAs ARE COMPARED WITH THE PROPOSED ONE. THE INDIVIDUAL AND AVERAGE CPSNR VALUES ARE REPORTED

Image ID	Bayer [8]	Hirakawa [14]	Condat [15]	Hao [7]	Diag7
01	38.65	40.03	39.62	40.31	**40.78**
02	40.79	40.88	40.33	**41.61**	40.99
03	**42.36**	41.75	42.08	41.37	42.19
04	41.07	40.55	40.54	**41.82**	40.55
05	**38.35**	37.30	36.94	37.54	37.84
06	40.42	41.27	40.74	41.15	**41.37**
07	**42.87**	41.86	42.00	41.62	41.74
08	35.72	37.80	37.37	**38.07**	37.90
09	**42.80**	42.27	42.21	42.10	42.57
10	42.70	42.69	42.70	42.85	**43.10**
11	40.31	40.60	40.25	**40.92**	40.38
12	43.43	43.39	43.52	43.85	**44.11**
13	35.13	**35.57**	35.21	35.24	35.52
14	**36.92**	36.27	35.94	35.95	35.68
15	40.05	39.78	39.61	**40.54**	39.72
16	44.08	**44.76**	44.63	44.18	44.59
17	**41.80**	41.68	41.17	41.76	41.47
18	**37.88**	37.82	37.53	37.48	37.60
19	40.86	41.58	41.19	41.49	**41.64**
20	41.27	41.58	41.13	41.27	**41.63**
21	39.29	**40.62**	40.19	40.26	40.15
22	38.65	38.80	38.67	38.64	**38.89**
23	**43.12**	42.41	41.94	42.32	42.18
24	35.61	**36.62**	36.23	36.40	36.52
Avg.	40.17	40.33	40.07	40.36	**40.38**

The whole process for a given CFA pattern size is demonstrated in Fig. 2. It is worth noting that, if we specify the CFA pattern size to 4×2, 3×2, and 4×4, respectively, the produced frequency structure candidates by our proposal strategy include those of the Hirakawa CFA [14], the Condat CFA [15], and the Hao CFA [7]. Besides the three ones, we check all pattern sizes that are equal to or smaller than 9×9.

The CFA pattern that performs the best in our experiments is shown in Fig. 4. It is a 7×7 diagonal stripe CFA pattern using optimized color components (Fig. 4(1)), which we call the Diag7 CFA pattern. From Fig. 4(2), we can see that all its chroma replicas are located far away from the center and also the horizontal and the vertical axes. The Diag7 CFA pattern has two chromas F_{C_1} and F_{C_2}, each of which having only one replica (Fig. 4(3)). Accordingly, the two chromas are conjugate, i.e., $F_{C_1} = F^*_{C_2}$ and $F_{C_2} = F^*_{C_1}$. All the color values of the Diag7 CFA pattern are shown in Fig. 4(4), and the corresponding color transformation matrix $\mathbf{M}$ is, as shown at the bottom of this page.

The CFAs with redundant chroma replicas do not achieve the best performance. One possible reason is that we simply use the average strategy to fuse all estimations of a chroma. We take the 5×5 pattern size shown in Fig. 2 as an example. Our method finally produces two CFA patterns with two chromas and four CFA patterns with four chromas (see Fig. 2(e)). The average CPSNR values of the two CFA patterns with two chromas on the Kodak dataset are 40.00dB and 39.88dB, respectively, while the best CFA pattern with four chromas results in 39.15dB. On the other hand, the average CPSNR value of the Hao CFA [7] with five chromas (see Fig. 1(3d)) is improved by 0.43dB (from 39.93dB to 40.36dB) when a tailored combination strategy is adopted (see Table III and IV). However, developing a combination strategy that is beneficial to all CFAs with redundant chroma replicas is non-trivial and out of the scope of this work. We will consider the problem in the future.

We next investigate the demosaicking performance of our new CFA. The individual and average CPSNR values of all tested CFAs on the Kodak dataset are shown in Table IV. The best values are in boldface. We can see that our Diag7 CFA yields similar individual CPSNR values as other CFAs and slightly outperforms the others on the whole dataset.

We also present part of the visual comparison in Fig. 5 and 6. From Fig. 5, we can see that the visual quality of optimized CFAs is better than that of the Bayer CFA, especially in removing aliasing artifacts. This testifies that reducing the spectral overlap between luma and chromas is important for producing demosaicked images with high visual quality. Fig. 6 further shows the perceptually superiority of the Diag7 CFA over other three optimized CFAs.

V. CONCLUSIONS

In this paper, we present an automatic CFA design method in the frequency domain based on the frequency structure [7]. To accomplish this, we develop a multi-objective optimization approach to automatically rule out a majority of unpromising frequency structures. Then for each frequency structure candidate, we present a new parameter optimization method that is appropriate for arbitrary frequency structures, including those with conjugate chrominance replicas. Our work provides an automatic approach to designing CFAs that are advantageous during the subsequent demosaicking process in producing fewer visual artifacts. Extensive experiments on standard test images demonstrate the superiority of our design method.

APPENDIX A
SOLVING THE PARAMETER OPTIMIZATION MODEL

The optimization problem (13) can be solved by the alternating direction method (ADM) [30]. When applying ADM to (13), we first let $\mathbf{C} = \mathbf{C}_1 + i\mathbf{C}_2$ and $\mathbf{D} = \mathbf{D}_1 + i\mathbf{D}_2$, then the constraint $\mathbf{C}\Re(\mathbf{M}) + \mathbf{D}\Im(\mathbf{M}) \geq \mathbf{0}$ in (13) can be written in more detail as: $\mathbf{C}_1\Re(\mathbf{M}) + \mathbf{D}_1\Im(\mathbf{M}) \geq \mathbf{0}$ and $\mathbf{C}_2\Re(\mathbf{M}) + \mathbf{D}_2\Im(\mathbf{M}) = \mathbf{0}$. Since $\mathbf{C}_1, \mathbf{D}_1, \mathbf{C}_2$, and $\mathbf{D}_2$ are all real, the two new constraints are linear w.r.t. $\mathbf{M}$,

$$\mathbf{M} = \begin{pmatrix} 0.31316 & 0.34342 & 0.34342 \\ 0.10836 + 0.13588i & -0.17171 + 0.02579i & 0.06335 - 0.16167i \\ 0.10836 - 0.13588i & -0.17171 - 0.02579i & 0.06335 + 0.16167i \end{pmatrix}$$

which is the requirement of ADM. So we rewrite (13) as follows:

$$\begin{aligned}&\min_{\mathbf{M}} \|\mathbf{M}^{-1}\|_2\\ &\text{s.t. } \mathbf{C}_1\Re(\mathbf{M})+\mathbf{D}_1\Im(\mathbf{M})\geq \mathbf{0},\\ &\qquad \mathbf{C}_2\Re(\mathbf{M})+\mathbf{D}_2\Im(\mathbf{M})=\mathbf{0},\quad \mathbf{Ma}=\mathbf{b}.\end{aligned}\tag{14}$$

We next introduce three auxiliary variables $\mathbf{N}_1$, $\mathbf{N}_2$, and $\mathbf{S}$ to reformulate our model as follows:

$$\begin{aligned}&\min_{\mathbf{M},\mathbf{N}_1,\mathbf{N}_2,\mathbf{S}} \|\mathbf{M}^{-1}\|_2+\mathcal{I}_{\mathbb{R}_+}(\mathbf{S})\\ &\text{s.t. } \mathbf{M}=\mathbf{N}_1+i\mathbf{N}_2,\quad (\mathbf{N}_1+i\mathbf{N}_2)\mathbf{a}=\mathbf{b},\\ &\qquad \mathbf{C}_1\mathbf{N}_1+\mathbf{D}_1\mathbf{N}_2=\mathbf{S},\mathbf{C}_2\mathbf{N}_1+\mathbf{D}_2\mathbf{N}_2=\mathbf{0},\end{aligned}\tag{15}$$

where $\mathcal{I}_{\mathbb{R}_+}(\cdot)$ is an indicator function, i.e., $\mathcal{I}_{\mathbb{R}_+}(\mathbf{y})=\begin{cases}0, & \mathbf{y}\in\mathbb{R}_+;\\ +\infty, & \text{otherwise.}\end{cases}$, and $\mathbb{R}_+$ denotes the set of real non-negative matrices. It should be noted that $\mathbf{N}_1$ and $\mathbf{N}_2$ are not necessarily required to be real. This is because the frequency structure candidates satisfy the conjugate symmetry. So if all the constraints are met, it can guarantee that $\mathbf{N}_1$ and $\mathbf{N}_2$ are real.

The augmented Lagrangian function of problem (15) is:

$$\begin{aligned}\mathcal{L}(\mathbf{M},\mathbf{N}_1,\mathbf{N}_2,\mathbf{S},\mathbf{X},\mathbf{x},\mathbf{Y},\mathbf{Z})&=\|\mathbf{M}^{-1}\|_2+\mathcal{I}_{\mathbb{R}_+}(\mathbf{S})\\ &+\langle\mathbf{X},\mathbf{M}-(\mathbf{N}_1+i\mathbf{N}_2)\rangle+\langle\mathbf{x},(\mathbf{N}_1+i\mathbf{N}_2)\mathbf{a}-\mathbf{b}\rangle\\ &+\langle\mathbf{Y},\mathbf{C}_1\mathbf{N}_1+\mathbf{D}_1\mathbf{N}_2-\mathbf{S}\rangle+\langle\mathbf{Z},\mathbf{C}_2\mathbf{N}_1+\mathbf{D}_2\mathbf{N}_2\rangle\\ &+\frac{\beta}{2}\Big(\|\mathbf{M}-(\mathbf{N}_1+i\mathbf{N}_2)\|_F^2+\|(\mathbf{N}_1+i\mathbf{N}_2)\mathbf{a}-\mathbf{b}\|_2^2\\ &+\|\mathbf{C}_1\mathbf{N}_1+\mathbf{D}_1\mathbf{N}_2-\mathbf{S}\|_F^2+\|\mathbf{C}_2\mathbf{N}_1+\mathbf{D}_2\mathbf{N}_2\|_F^2\Big),\end{aligned}\tag{16}$$

where $\mathbf{X}$, $\mathbf{x}$, $\mathbf{Y}$, and $\mathbf{Z}$ are the Lagrange multipliers, $\langle\cdot,\cdot\rangle$ is the inner product, and $\beta>0$ is the penalty parameter which is updated during iterations.

Then by ADM, we can solve problem (15) via the following iterations:

$$\begin{aligned}\mathbf{M}^{k+1}&=\arg\min_{\mathbf{M}}\mathcal{L}(\mathbf{M},\mathbf{N}_1^k,\mathbf{N}_2^k,\mathbf{S}^k,\mathbf{X}^k,\mathbf{x}^k,\mathbf{Y}^k,\mathbf{Z}^k)\\ &=\arg\min_{\mathbf{M}}\|\mathbf{M}^{-1}\|_2+\frac{\beta}{2}\|\mathbf{M}-(\mathbf{N}_1^k+i\mathbf{N}_2^k)+\mathbf{X}^k/\beta\|_F^2\\ &=\arg\min_{\mathbf{M}}\frac{1}{\beta}\|\mathbf{M}^{-1}\|_2+\frac{1}{2}\|\mathbf{M}-\mathbf{W}^k\|_F^2,\end{aligned}\tag{17}$$

$$\begin{aligned}&\{\mathbf{N}_j^{k+1}\}_{j=1}^2\\ &=\arg\min_{\mathbf{N}_1,\mathbf{N}_2}\mathcal{L}(\mathbf{M}^{k+1},\mathbf{N}_1,\mathbf{N}_2,\mathbf{S}^k,\mathbf{X}^k,\mathbf{x}^k,\mathbf{Y}^k,\mathbf{Z}^k)\\ &=\arg\min_{\mathbf{N}_1,\mathbf{N}_2}\frac{1}{2}\|\mathbf{M}^{k+1}-(\mathbf{N}_1+i\mathbf{N}_2)+\mathbf{X}^k/\beta\|_F^2\\ &\quad+\frac{1}{2}\|(\mathbf{N}_1+i\mathbf{N}_2)\mathbf{a}-\mathbf{b}+\mathbf{x}^k/\beta\|_2^2\\ &\quad+\frac{1}{2}\|\mathbf{C}_1\mathbf{N}_1+\mathbf{D}_1\mathbf{N}_2-\mathbf{S}^k+\mathbf{Y}^k/\beta\|_F^2\\ &\quad+\frac{1}{2}\|\mathbf{C}_2\mathbf{N}_1+\mathbf{D}_2\mathbf{N}_2+\mathbf{Z}^k/\beta\|_F^2,\end{aligned}\tag{18}$$

$$\begin{aligned}\mathbf{S}^{k+1}&=\arg\min_{\mathbf{S}}\mathcal{L}(\mathbf{M}^{k+1},\mathbf{N}_1^{k+1},\mathbf{N}_2^{k+1},\mathbf{S},\mathbf{X}^k,\mathbf{x}^k,\mathbf{Y}^k,\mathbf{Z}^k)\\ &=\arg\min_{\mathbf{S}}\mathcal{I}_{\mathbb{R}_+}(\mathbf{S})+\frac{\beta}{2}\|\mathbf{C}_1\mathbf{N}_1^{k+1}+\mathbf{D}_1\mathbf{N}_2^{k+1}-\mathbf{S}+\mathbf{Y}^k/\beta\|_F^2\\ &=\max(\mathbf{0},\Re(\mathbf{C}_1\mathbf{N}_1^{k+1}+\mathbf{D}_1\mathbf{N}_2^{k+1}+\mathbf{Y}^k/\beta)),\end{aligned}\tag{19}$$

$$\mathbf{X}^{k+1}=\mathbf{X}^k+\beta(\mathbf{M}^{k+1}-(\mathbf{N}_1^{k+1}+i\mathbf{N}_2^{k+1})),\tag{20}$$

$$\mathbf{x}^{k+1}=\mathbf{x}^k+\beta((\mathbf{N}_1^{k+1}+i\mathbf{N}_2^{k+1})\mathbf{a}-\mathbf{b}),\tag{21}$$

$$\mathbf{Y}^{k+1}=\mathbf{Y}^k+\beta(\mathbf{C}_1\mathbf{N}_1^{k+1}+\mathbf{D}_1\mathbf{N}_2^{k+1}-\mathbf{S}^{k+1}),\tag{22}$$

$$\mathbf{Z}^{k+1}=\mathbf{Z}^k+\beta(\mathbf{C}_2\mathbf{N}_1^{k+1}+\mathbf{D}_2\mathbf{N}_2^{k+1}),\tag{23}$$

where $\mathbf{W}^k=(\mathbf{N}_1^k+i\mathbf{N}_2^k)-\mathbf{X}^k/\beta$ in (17).

The solution to problem (17) is given by Theorem 1. We provide the proof of Theorem 1 in the Supplementary Material.

Theorem 1: The solution to problem (17) is:

$$\mathbf{M}^{k+1}=\mathbf{U}^k\Sigma^{k+1}(\mathbf{V}^k)^H,\tag{24}$$

where $\mathbf{U}^k\mathbf{\Lambda}^k(\mathbf{V}^k)^H$ is the SVD of $\mathbf{W}^k$, $\mathbf{U}^k$ and $\mathbf{V}^k$ are unitary matrices, $\mathbf{\Lambda}^k=\text{diag}(\boldsymbol{\lambda}^k)$, in which $\text{diag}(\mathbf{y})$ converts the vector $\mathbf{y}$ into a diagonal matrix whose j-th diagonal element is $\mathbf{y}_j$, $\boldsymbol{\lambda}^k=(\lambda_1^k,\lambda_2^k,\lambda_3^k)^T$ is the real vector of singular values of $\mathbf{W}^k$ and satisfies $\lambda_1^k\geq\lambda_2^k\geq\lambda_3^k>0$, and $\Sigma^{k+1}=\text{diag}(\boldsymbol{\sigma}^{k+1})$, in which $\boldsymbol{\sigma}^{k+1}=(\sigma_1^{k+1},\sigma_2^{k+1},\sigma_3^{k+1})^T$ is the solution to the following problem:

$$\min_{\sigma_1\geq\sigma_2\geq\sigma_3>0}\ \frac{1}{\beta\sigma_3}+\frac{1}{2}\sum\nolimits_{j=1}^3(\sigma_j-\lambda_j^k)^2.\tag{25}$$

Problem (18) has a closed-form solution and we show the solution process as follows. Let $\mathbf{E}=\begin{pmatrix}\mathbf{N}_1\\ \mathbf{N}_2\end{pmatrix}$ in (18), then the objective function of (18) is:

$$\begin{aligned}\mathcal{F}(\mathbf{E})&=\frac{1}{2}\left\|\mathbf{M}^{k+1}-(\mathbf{I},i\mathbf{I})\mathbf{E}+\frac{\mathbf{X}^k}{\beta}\right\|_F^2\\ &\quad+\frac{1}{2}\left\|(\mathbf{I},i\mathbf{I})\mathbf{Ea}-\mathbf{b}+\frac{\mathbf{x}^k}{\beta}\right\|_2^2\\ &\quad+\frac{1}{2}\|(\mathbf{C}_1,\mathbf{D}_1)\mathbf{E}-\mathbf{S}^k+\mathbf{Y}^k/\beta\|_F^2\\ &\quad+\frac{1}{2}\|(\mathbf{C}_2,\mathbf{D}_2)\mathbf{E}+\mathbf{Z}^k/\beta\|_F^2.\end{aligned}\tag{26}$$

It is equivalent to the following function:

$$\begin{aligned}&f(\text{vec}(\mathbf{E}))\\ &=\frac{1}{2}\|\mathbf{A}_1\text{vec}(\mathbf{E})-\mathbf{b}_1^k\|_2^2+\frac{1}{2}\|\mathbf{A}_2\text{vec}(\mathbf{E})-\mathbf{b}_2^k\|_2^2\\ &\quad+\frac{1}{2}\|\mathbf{A}_3\text{vec}(\mathbf{E})-\mathbf{b}_3^k\|_2^2+\frac{1}{2}\|\mathbf{A}_4\text{vec}(\mathbf{E})-\mathbf{b}_4^k\|_2^2,\end{aligned}\tag{27}$$

where $\mathbf{A}_1=\mathbf{I}\otimes(\mathbf{I},i\mathbf{I})$, $\mathbf{A}_2=\mathbf{a}^T\otimes(\mathbf{I},i\mathbf{I})$, $\mathbf{A}_3=\mathbf{I}\otimes(\mathbf{C}_1,\mathbf{D}_1)$, $\mathbf{A}_4=\mathbf{I}\otimes(\mathbf{C}_2,\mathbf{D}_2)$, $\mathbf{b}_1^k=\mathbf{A}_1^H\text{vec}(\mathbf{M}^{k+1}+\mathbf{X}^k/\beta)$, $\mathbf{b}_2^k=\mathbf{A}_2^H(\mathbf{b}-\mathbf{x}^k/\beta)$, $\mathbf{b}_3^k=\mathbf{A}_3^H\text{vec}(\mathbf{S}^k-\mathbf{Y}^k/\beta)$, $\mathbf{b}_4^k=\mathbf{A}_4^H\text{vec}(-\mathbf{Z}^k/\beta)$, $\otimes$ denotes Kronecker product, $\mathbf{I}\in\mathbb{R}^{3\times3}$ is the identity matrix, $\mathbf{y}^T$ and $\mathbf{y}^H$ denote the transpose and conjugate transpose of $\mathbf{y}$, respectively, and $\text{vec}(\mathbf{M})$ is the

Algorithm 1 ADM Algorithm for Problem (14)

Input: $\mathbf{C}$, $\mathbf{D}$, $\mathbf{a}$, $\mathbf{b}$, $\beta = 1$, $\beta_{\max} = 10^4$, $\rho_0 = 1.1$, $\varepsilon_1 = 10^{-5}$, and $\varepsilon_2 = 10^{-10}$.

for each $\mathbf{N} \in \Omega$ defined in (33) **do**
  Initialization: $\mathbf{N}_1 = \mathbf{N}$, $\mathbf{N}_2 = \mathbf{N}$, $\mathbf{S} = \frac{1}{3} \times \mathbf{1}$, $\mathbf{X} = \mathbf{0}$, $\mathbf{x} = \mathbf{0}$, $\mathbf{Y} = \mathbf{0}$, $\mathbf{Z} = \mathbf{0}$, $k = 0$.
  while the stop conditions (31) and (32) are not met **do**
    fix the others and update $\mathbf{M}$ by (24).
    fix the others and update $\mathbf{N}_1$ and $\mathbf{N}_2$ by (28).
    fix the others and update $\mathbf{S}$ by (19).
    update the multipliers $\mathbf{X}$, $\mathbf{x}$, $\mathbf{Y}$, and $\mathbf{Z}$ by (20)-(23).
    update β by (29) and (30).
    $k \leftarrow k + 1$.
  end while
  if $\|\hat{\mathbf{M}}^{-1}\|_2 > \|\mathbf{M}^{-1}\|_2$ **then**
    $\hat{\mathbf{M}} \leftarrow \mathbf{M}$.
  end if
end for

Output: $\hat{\mathbf{M}}$.

vectorization of $\mathbf{M}$. Set $\frac{\partial f}{\partial \text{vec}(\mathbf{E})} = 0$, we have the solution to (27) as follows:

$$\text{vec}(\mathbf{E}^{k+1}) = \text{vec}\begin{pmatrix}\mathbf{N}_1^{k+1} \\ \mathbf{N}_2^{k+1}\end{pmatrix} = \mathbf{G}^{-1}(\mathbf{b}_1^k + \mathbf{b}_2^k + \mathbf{b}_3^k + \mathbf{b}_4^k), \quad (28)$$

where $\mathbf{G} = \mathbf{A}_1^H\mathbf{A}_1 + \mathbf{A}_2^H\mathbf{A}_2 + \mathbf{A}_3^H\mathbf{A}_3 + \mathbf{A}_4^H\mathbf{A}_4$.

To accelerate the convergence of the algorithm, we use the following adaptive updating strategy for the penalty parameter β [30]:

$$\beta^{k+1} = \min(\beta_{\max}, \beta^k \rho), \quad (29)$$

where $\beta_{\max}$ is an upper bound of $\{\beta^k\}$. The value of ρ is defined as

$$\rho = \begin{cases} \rho_0, & \text{if } \beta^k \alpha^k < \varepsilon_1, \\ 1, & \text{otherwise}, \end{cases} \quad (30)$$

where $\rho_0 \geq 1$ is a constant and $\alpha^k = \max\{\|\mathbf{M}^{k+1} - \mathbf{M}^k\|_\infty, \|\mathbf{N}_1^{k+1} - \mathbf{N}_1^k\|_\infty, \|\mathbf{N}_2^{k+1} - \mathbf{N}_2^k\|_\infty, \|\mathbf{S}^{k+1} - \mathbf{S}^k\|_\infty\}$. The stopping criteria are:

$$\beta^k \alpha^k < \varepsilon_1 \quad (31)$$

and

$$\max\{\|\mathbf{M}^{k+1} - \mathbf{N}_1^{k+1} - i\mathbf{N}_2^{k+1}\|_\infty, \|(\mathbf{N}_1^{k+1} + i\mathbf{N}_2^{k+1})\mathbf{a} - \mathbf{b}\|_\infty, \|\mathbf{C}_1\mathbf{N}_1^{k+1} + \mathbf{D}_1\mathbf{N}_2^{k+1} - \mathbf{S}^{k+1}\|_\infty, \|\mathbf{C}_2\mathbf{N}_1^{k+1} + \mathbf{D}_2\mathbf{N}_2^{k+1}\|_\infty\} < \varepsilon_2. \quad (32)$$

Problem (14) is nonconvex w.r.t. $\mathbf{M}$ in its entire domain. However, it is convex in each of eight convex cones, which are identified by the following interior point set:

$$\Omega = \{\text{diag}((1,1,1)^T), \text{diag}((1,1,-1)^T), \text{diag}((1,-1,1)^T), \text{diag}((1,-1,-1)^T), \text{diag}((-1,1,1)^T), \text{diag}((-1,1,-1)^T), \text{diag}((-1,-1,1)^T), \text{diag}((-1,-1,-1)^T)\}. \quad (33)$$

The optimization in each cone can simply be achieved by initializing $\mathbf{M}$ with an interior point of the cone.

We summarize the whole solution process of problem (14) in Algorithm 1.

ACKNOWLEDGMENT

The authors would like to thank Dr. David Wipf from Microsoft Research Asia for proofreading this paper. They are also grateful to Eric Dubois, Laurent Condat, and Pengwei Hao for sharing their source codes with us.

REFERENCES

[1] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau, "Demosaicking: Color filter array interpolation," *IEEE Signal Process. Mag.*, vol. 22, no. 1, pp. 44–54, Jan. 2005.

[2] X. Li, B. Gunturk, and L. Zhang, "Image demosaicing: A systematic survey," *Proc. SPIE*, vol. 6822, p. 68221J, Jan. 2008.

[3] D. Menon and G. Calvagno, "Color image demosaicking: An overview," *Signal Process., Image Commun.*, vol. 26, nos. 8–9, pp. 518–533, Oct. 2011.

[4] L. Condat, "Color filter array design using random patterns with blue noise chromatic spectra," *Image Vis. Comput.*, vol. 28, no. 8, pp. 1196–1202, 2010.

[5] C. Bai, J. Li, Z. Lin, J. Yu, and Y.-W. Chen, "Penrose demosaicking," *IEEE Trans. Image Process.*, vol. 24, no. 5, pp. 1672–1684, May 2015.

[6] R. Lukac and K. N. Plataniotis, "Color filter arrays: Design and performance analysis," *IEEE Trans. Consum. Electron.*, vol. 51, no. 4, pp. 1260–1267, Nov. 2005.

[7] P. Hao, Y. Li, Z. Lin, and E. Dubois, "A geometric method for optimal design of color filter arrays," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 709–722, Mar. 2011.

[8] B. E. Bayer, "Color imaging array," U.S. Patent 3 971 065, Jul. 20, 1976.

[9] D. Alleysson, S. Süsstrunk, and J. Hérault, "Linear demosaicing inspired by the human visual system," *IEEE Trans. Image Process.*, vol. 14, no. 4, pp. 439–449, Apr. 2005.

[10] J. Adams, K. Parulski, and K. Spaulding, "Color processing in digital cameras," *IEEE Micro*, vol. 18, no. 6, pp. 20–30, Nov./Dec. 1998.

[11] *Realization of Natural Color Reproduction in Digital Still Cameras, Closer to the Natural Sight Perception of the Human Eye*. [Online]. Available: http://www.sony.net/SonyInfo/News/Press/200307/03-029E/, accessed Dec. 1, 2013.

[12] J. F. Hamilton, Jr., and J. T. Compton, "Processing color and panchromatic pixels," U.S. Patent 8 274 715, Sep. 25, 2012.

[13] *X-Trans CMOS*. [Online]. Available: http://www.fujifilmusa.com/products/digital_cameras/x/fujifilm_x_pro1/features/, accessed Nov. 10, 2014.

[14] K. Hirakawa and P. J. Wolfe, "Spatio-spectral color filter array design for optimal image recovery," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1876–1890, Oct. 2008.

[15] L. Condat, "A new color filter array with optimal properties for noiseless and noisy color image acquisition," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2200–2210, Aug. 2011.

[16] *Kodak Lossless True Color Image Suite*. [Online]. Available: http://r0k.us/graphics/kodak/, accessed Dec. 1, 2013.

[17] M. Parmar and S. J. Reeves, "A perceptually based design methodology for color filter arrays," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 3. May 2004, pp. III-473–III-476.

[18] Y. M. Lu and M. Vetterli, "Optimal color filter array design: Quantitative conditions and an efficient search procedure," *Proc. SPIE*, vol. 7250, pp. 725009-1–725009-8, Jan. 2009.

[19] D. Zhang and X. Wu, "Color demosaicking via directional linear minimum mean square-error estimation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2167–2178, Dec. 2005.

[20] Y. M. Lu, C. Fredembach, M. Vetterli, and S. Süsstrunk, "Designing color filter arrays for the joint capture of visible and near-infrared images," in *Proc. 16th IEEE Int. Conf. Image Process.*, Nov. 2009, pp. 3797–3800.

[21] M. Parmar and S. J. Reeves, "Selection of optimal spectral sensitivity functions for color filter arrays," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3190–3203, Dec. 2010.

[22] X. Zhang and B. A. Wandell, "A spatial extension of CIELAB for digital color-image reproduction," *J. Soc. Inf. Display*, vol. 5, no. 1, pp. 61–63, 1997.

[23] Z. Sadeghipoor, Y. M. Lu, and S. Süsstrunk, "Optimum spectral sensitivity functions for single sensor color imaging," *Proc. SPIE*, vol. 8299, pp. 829904-1–829904-14, Jan. 2012.
[24] E. Dubois, "Frequency-domain methods for demosaicking of Bayer-sampled color images," *IEEE Signal Process. Lett.*, vol. 12, no. 12, pp. 847–850, Dec. 2005.
[25] Y. Li, P. Hao, and Z. Lin, "Color filter arrays: Representation and analysis," Dept. Comput. Sci., Queen Mary Univ. London, London, U.K., Tech. Rep. RR-08-04, 2008.
[26] J. Wang, C. Zhang, and P. Hao, "New color filter arrays of high light sensitivity and high demosaicking performance," in *Proc. 18th IEEE Int. Conf. Image Process.*, Sep. 2011, pp. 3153–3156.
[27] K. Miettinen, *Nonlinear Multiobjective Optimization*. New York, NY, USA: Springer, 1998.
[28] B. Leung, G. Jeon, and E. Dubois, "Least-squares luma–chroma demultiplexing algorithm for Bayer demosaicking," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1885–1894, Jul. 2011.
[29] E. Dubois, *Color-Filter-Array Sampling of Color Images: Frequency Domain Analysis and Associated Demosaicking Algorithms*. Boca Raton, FL, USA: CRC Press, 2009.
[30] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 2. 2011, pp. 612–620.

Chenyan Bai received the B.E. and M.E. degrees in computer science from the School of Mathematics and Computer, Hebei University, China, in 2008 and 2011, respectively. She is currently pursuing the Ph.D. degree with the Beijing Key Laboratory of Traffic Data Analysis and Mining, School of Computer and Information Technology, Beijing Jiaotong University. Her research interest is image processing.

Jia Li received the bachelor's degree in mathematics and the master's degree in computer science from Zhengzhou University, Zhengzhou, China, in 2007 and 2012, respectively. He is currently pursuing the Ph.D. degree with the Beijing Key Laboratory of Traffic Data Analysis and Mining, School of Computer and Information Technology, Beijing Jiaotong University. His research interest is image processing.

Zhouchen Lin (M'00–SM'08) received the Ph.D. degree in applied mathematics from Peking University, in 2000. He is currently a Professor with the Key Laboratory of Machine Perception, School of Electronic Engineering and Computer Science, Peking University. He is also a Chair Professor with Northeast Normal University. He was a Guest Professor with Shanghai Jiao Tong University, Beijing Jiaotong University, and Southeast University. He was also a Guest Researcher with the Institute of Computing Technology, Chinese Academic of Sciences. His research interests include computer vision, image processing, machine learning, pattern recognition, and numerical optimization. He is an Associate Editor of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE and the *International Journal of Computer Vision*.

Jian Yu received the B.S. and M.S. degrees in mathematics and the Ph.D. degree in applied mathematics from Peking University, Beijing, China, in 1991, 1994, and 2000, respectively. He is currently a Professor with Beijing Jiaotong University and the Director of the Beijing Key Laboratory of Traffic Data Analysis and Mining. His research interests include machine learning, image processing, and pattern recognition.