# Color Filter Arrays for Quanta Image Sensors

Omar A. Elgendy , *Member, IEEE*, and Stanley H. Chan , *Senior Member, IEEE*

*Abstract*—**Quanta image sensor (QIS) is a new type of image sensors with single-photon sensitivity. However, the existing literature on QIS has been focusing on monochromatic signals. In this article, we discuss how color imaging can be made possible by designing color filter arrays for QIS and other small pixels. Designing color filter arrays for small pixels is challenging because maximizing the light efficiency while suppressing aliasing and crosstalk are conflicting tasks. We present an optimization-based framework which unifies several mainstream color filter array design methodologies. Our method offers greater generality and flexibility. Compared to existing methods, the new framework can simultaneously handle luminance sensitivity, chrominance sensitivity, cross-talk, anti-aliasing, manufacturability and orthogonality. Extensive experimental comparisons demonstrate the effectiveness of the framework.**

*Index Terms*—**Quanta image sensor (QIS), single-photon detector, demosaicking, denoising, image reconstruction, color filter array.**

## I. Introduction

### A. Quanta Image Sensor

QUANTA Image Sensor (QIS) is a new type of image sensors proposed by E. Fossum in 2005 [1]–[3] as a candidate for the third generation digital image sensor after CCD and CMOS. The sensor comprises a massive array of sub-diffraction limit single-photon detectors, called "jots", with a pixel pitch of $1.1 \ \mu$m. Having read-out noise of 0.21e- and dark current of 0.16e-/sec, QIS can count incoming photons to produce a digital output of bit-depth in the range of 1–5 bits, assuming that the exposure does not saturate the sensor. As reported in [2] and [3], the latest QIS sensor can achieve a resolution of $1024 \times 1024$ jots and and a frame rate up rate up to 1000 fps by using the commercial 45/65 nm 3D-stacked backside illumination CMOS process. These unique characteristics of QIS have led to a new image formation process. Unlike CMOS image sensors that accumulate photons, QIS oversamples the scene by producing binary measurements as depicted in Fig. 1. See also [4]–[6] for detailed discussions.

Despite the rapid advancement in QIS hardware [1]–[3] and algorithms [4]–[7], all reported findings, to-date, are based on monochromatic data. The first color QIS imaging was recently presented by Gnanasambandam *et al.* [8], where they demonstrated how to reconstruct a color image from the sensor using a Bayer color filter array. In this paper, we discuss how to design color filter array for better image acquisition.

### B. Color Filter Arrays for QIS

A color filter array (CFA) is a mask placed on top of the sensor to select (filter) wavelengths. As light passes through the color filter array, the resulting image is a mosaic pattern of the three tri-stimulus RGB colors. Traditionally, CFA is organized as a periodic replica of a 2D kernel called the *color atom*. The de facto color atom is the Bayer pattern [9] because of its simplicity and the readily available demosaicking algorithms, see for instance [10]–[19]. More sophisticated CFAs have been proposed [20]–[30] to improve the Bayer CFA.

When designing a CFA, there are three factors that should be taken into consideration:

- **Aliasing**: Since color filtering is a sampling process, aliasing happens when the sampling rate is less than Nyquist. Aliasing causes false color artifacts at color edges, called the *Moiré* artifacts [21]. Color filters that are susceptible to aliasing, such as the Bayer CFA, require sophisticated demosaicking algorithms to suppress the Moiré artifacts. In contrast, a robust CFA can use less complicated demosaicking algorithms.
- **Sensitivity**: Since CFA is a filter, it blocks part of the incoming light. This reduces the sensor sensitivity and makes the image more susceptible to noise. A good CFA design should maximize the sensitivity by allowing transparent or "panchromatic" color filters that block as few wavelengths as possible.
- **Crosstalk**: Crosstalk can be either optical or electrical [31]. If not treated, crosstalk will make colors look pale or desaturated. Crosstalk desaturation is corrected by pixel-wise multiplication of the RGB color vector using a color correction matrix. However, color correction enhances residual noise in the image [31], [32]. The situation is worsen in QIS because of its small size.

The three factors above are conflicting: Optimizing one generally degrades the others. For conventional CMOS image sensors, crosstalk is not severe, and so most CFA designs in the literature consider aliasing and sensitivity only. The first QIS color filter array design is proposed by Anzagira and Fossum [31]. However, aliasing is not considered because QIS is sufficiently small. We generalize these prior work by considering all three factors so that our framework can be applied to both QIS and CMOS image sensors.
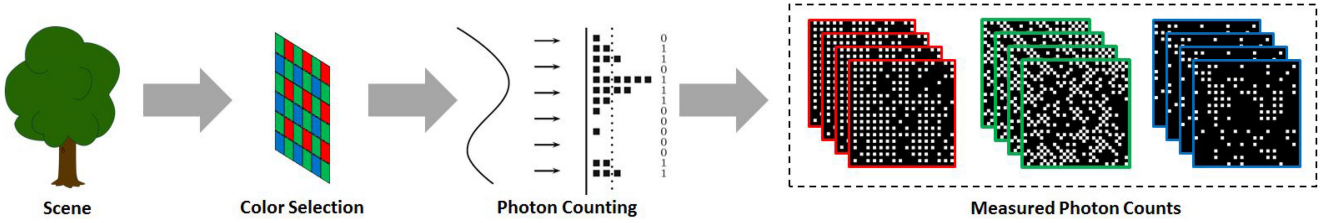
Fig. 1. **QIS Imaging Model [8]**. When the scene image arrives at the sensor, the CFA first selects the wavelength according to the colors. Each color pixel is then sensed using a photon-detector and reports a binary value based on whether the photon counts exceeds certain threshold or not. The measured data contains three subsampled sequences, each representing a measurement in the color channel.

## C. Related Work

The design framework we propose in this paper unifies several mainstream CMOS-based color filter arrays design approaches. To put our paper in the proper context in the literature, we list a few of the known works.

**Spatial CFA Design**: By suppressing the Moiré artifacts and crosstalk while keeping the demosaicking algorithm simple, Lukac and Plataniotis [20] proposed a CFA and compared it with other CFAs using a universal demosaicking method. However, their work did not provide a mathematical framework to analyze the CFA optimality.

**Spatio-Spectral CFA Design**: Hirakawa and Wolfe [21] proposed a method through the spatial and spectral domain analysis. Their CFA reduces aliasing in the frequency domain, and possesses high sensitivity and numerical stability. Condat [33] extended the framework by optimizing luminance and chrominance sensitivity. He defined a new form of orthogonality between chrominance channels in the frequency domain. Hao *et al.* [24] and Wang *et al.* [25] proposed a framework based on symbolic discrete Fourier transform (DFT). Their CFA maximizes the numerical stability of linear demosaicking process under aliasing and physical constraints. Li *et al.* proposed spatio-spectral CFA design methods that are optimized for sensitivty [28], [29].

**Learning-Based CFA Design**: By minimizing the average error on a color dataset, Lu and Vetterli [22] used an iterative algorithm to solve a least squares CFA design problem. Chakrabarti [34] and Henz *et al.* [35] proposed to learn the optimal CFA pattern by using a deep neural network.

Besides these CFA designs, there are a number of other CFA designs such as [23], [26], [27], [30]. On the hardware side, Biay-Cheng *et al.* [36], [37] took into the account that color filter fabrication technology lags the image sensor technology in terms of miniaturization. They proposed a hardware-friendly CFA assuming that the color filter size is $1.5\times$ pixel size.

## D. Scope and Contributions

In this paper, we propose an optimization framework that encompasses aliasing, sensitivity and crosstalk in a unified model. This is the first time a quantitative crosstalk metric is incorporated in an optimization framework for CFA design. Our framework is general as it works with any crosstalk kernel depending on the sensor. Moreover, it works with any color atom size by a mild change in the formulation.

The main contribution of this paper is a general and flexible framework for CFA design. Compared to the existing CFA design framework, the new framework is able to simultaneously (Section III)

- Improve CFA's luminance and chrominance sensitivity,
- Reduce aliasing between luminance and chrominance channels,
- Suppress crosstalk between spectral sub-bands, and
- Enforce orthogonality between chrominance channels to permit simple linear demosaicking.

The design framework is presented in the form of optimization. We have two designs: A convex optimization and a non-convex optimization. In addition to the formulation, we also present an algorithm to solve the non-convex optimization. (Section IV).

For evaluation, we generalize the demosaicking algorithm of Condat [33] in Section V so that it now works with arbitrary CFAs. This pipeline comprises a demosaicking by frequency selection algorithm for removing the CFA masking effect followed by a color correction step for removing the desaturation effect of crosstalk. Experimental evaluation on the DIV2K evaluation dataset in Section VI shows the robustness of our proposed CFAs compared to other CFAs in literature.

## II. BACKGROUND AND NOTATIONS

To facilitate readers to understand the design framework, in this section we introduce a few notations and terminologies. We will start in Section II-A by describing the image formation using a CFA, then we discuss CFA in different domains in Sections II-B and II-C. Afterwards, in Section II-D, we define the optimization variables to simplify the design framework.

## A. Color Image Formation

Consider a color image $\mathbf{im}_{rgb}$ of size $H \times W$. We denote the normalized light intensities in the red, green and blue channels for the ($m$-th, $n$-th) pixel of the color image as

$$\mathbf{im}_{rgb}(m,n) = \begin{bmatrix} \text{im}_r(m,n) \\ \text{im}_g(m,n) \\ \text{im}_b(m,n) \end{bmatrix}, \qquad (1)$$

where $m \in \{0, \ldots, H-1\}, n \in \{0, \ldots, W-1\}$.

**Color Filtering**: To obtain color, we place a color filter on top of each jot to collect light for one of the RGB colors. The CFA
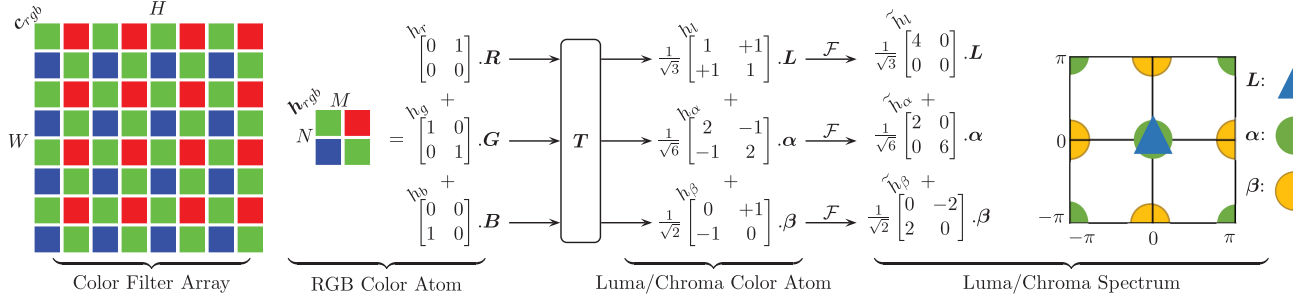
Fig. 2. Our terminology illustrated on the Bayer CFA example. The building unit of a CFA is a color atom. A transformation $\boldsymbol{T}$ is applied to the color atom to transform it from the canonical RGB color space to a luma/chroma color space to simplify the design process. Fourier transform is applied afterwards to obtain the color atom spectrum in the luma/chroma space.

is a periodic pattern of the same resolution of $\mathbf{im}_{rgb}$, defined as

$$\boldsymbol{c}_{rgb}(m,n) = \begin{bmatrix} c_r(m,n) \\ c_g(m,n) \\ c_b(m,n) \end{bmatrix}, \qquad (2)$$

where $c_r(m,n), c_g(m,n), c_b(m,n) \in [0,1]$ are the opacity rates for the red, green and blue pixels, respectively. For example, a red color filter is defined as $\boldsymbol{c}_{rgb}(m,n) = [1,0,0]^T$ as it only passes the red color. The light exposure on the QIS after passing through the CFA is denoted as $\theta(m,n)$, which is a linear combination of the tri-stimulus colors:

$$\theta(m,n) = \eta \boldsymbol{c}_{rgb}(m,n)^T \mathbf{im}_{rgb}(m,n)$$
$$= \eta \sum_{i \in \{r,g,b\}} c_i(m,n) \, \mathrm{im}_i(m,n). \qquad (3)$$

Here, $\eta$ is a positive scalar representing the sensor gain factor.

**Photon Arrival**. The photon arrival is modeled as a Poisson process. Let $\boldsymbol{Y} \in \mathbb{N}^{HW}$ be a vector of non-negative random integers denoting the number of photons arriving at QIS jots according to the light exposure $\boldsymbol{\theta}$. Then, the probability of observing a photon count $Y_j = y_j, j \in \{1, \ldots, HW\}$ is

$$\mathbb{P}(Y_j = y_j) = \frac{\theta_m^{y_j} e^{-\theta_j}}{y_j!}. \qquad (4)$$

In this work, we assume *single-bit QIS* [38] that quantizes the photon counts by QIS jots to a binary values $\boldsymbol{B} \in \{0,1\}^{HW}$ with $B_j = 1$ if $Y_j \geq q$ and $B_j = 0$ if $Y_j < q$, where $q > 0$ is a threshold. The probability of observing $B_j = b_j$ is

$$\mathbb{P}(B_j = b_j) = \Psi_q(\theta_j)^{1-b_j} (1 - \Psi_q(\theta_j))^{b_j}, \qquad (5)$$

where $\Psi_q(\cdot)$ is the incomplete Gamma function [6].

**Temporal Oversampling**. With frame rates that reach 1000 fps, QIS is able to catch the scene movement by taking $T$ temporal samplings for the same scene. This allows us to utilize multiple independent measurements over time to improve the statistics and decrease noise. Hence, for every jot with light exposure $\theta_j$, we have a set of $T$ independent binary measurements $\mathcal{B}_j = \{b_{j,0}, \ldots, b_{j,T-1}\}$.

### B. Color Filter Array Analysis in Different Color Spaces

Since the CFA $\boldsymbol{c}_{rgb}(m,n)$ is a periodic pattern, it is sufficient to use a *color atom* as the optimization variable when designing the CFA. The color atom takes the form

$$\boldsymbol{h}_{rgb}(m,n) = \begin{bmatrix} h_r(m,n) \\ h_g(m,n) \\ h_b(m,n) \end{bmatrix}, \qquad (6)$$

where each of $h_r$, $h_g$ and $h_b$ is an $M \times N$ array. For instance, the GRBG Bayer pattern has the following color atom (when $M = N = 2$):

$$h_r = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \; h_g = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; h_b = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$

because the Bayer pattern has one red pixel and one blue pixel located at two opposite diagonals, and two green pixels located in the remaining two positions. Fig. 2 illustrates the idea.

While the primal RGB color is common for making the CFA, it would be more convenient if the colors are *decorrelated*. To this end, we change the image representation from the canonical RGB basis to an orthornormal basis using a transformation matrix [33], [39]:

$$\boldsymbol{T} = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix}. \qquad (7)$$

This transformation maps an RGB image $\mathbf{im}_{rgb}$ to an image $\mathbf{im}_{l\alpha\beta} = [\mathrm{im}_l, \mathrm{im}_\alpha, \mathrm{im}_\beta]^T$ as follows (we drop the spatial indices $(m,n)$ for simplicity)

$$\mathbf{im}_{l\alpha\beta} = \begin{bmatrix} \mathrm{im}_l \\ \mathrm{im}_\alpha \\ \mathrm{im}_\beta \end{bmatrix} = \boldsymbol{T} \begin{bmatrix} \mathrm{im}_r \\ \mathrm{im}_g \\ \mathrm{im}_b \end{bmatrix} = \begin{bmatrix} (\mathrm{im}_r + \mathrm{im}_g + \mathrm{im}_b)/\sqrt{3} \\ (-\mathrm{im}_r + 2\mathrm{im}_g + \mathrm{im}_b)/\sqrt{6} \\ (\mathrm{im}_r - \mathrm{im}_b)/\sqrt{2} \end{bmatrix},$$

where $\mathrm{im}_l$ is a luminance (luma) component that contains the high frequency components such as edges and textures, whereas $\mathrm{im}_\alpha$ and $\mathrm{im}_\beta$ are chrominance (chroma) components that carry the color information.

Since $\boldsymbol{T}$ is orthonormal (i.e., $\boldsymbol{T}^T \boldsymbol{T} = \boldsymbol{I}$), we can rewrite the sampling process in Equation (3) in the luma/chroma space:

$$\theta(m,n) = \eta \boldsymbol{c}_{rgb}(m,n)^T \, \boldsymbol{T}^T \boldsymbol{T} \, \mathbf{im}_{rgb}(m,n)$$

$$= \eta \boldsymbol{c}_{l\alpha\beta}(m,n)^T \mathbf{im}_{l\alpha\beta}(m,n)$$

$$= \eta \sum_{i \in \{l,\alpha,\beta\}} c_i(m,n) \mathrm{im}_i(m,n), \qquad (8)$$

where $c_l(m,n)$, $c_\alpha(m,n)$ and $c_\beta(m,n)$ are the luma/chroma representation of the CFA, with

$$\boldsymbol{c}_{l\alpha\beta}(m,n) = \boldsymbol{T}\,\boldsymbol{c}_{rgb}(m,n). \qquad (9)$$

The luma/chroma representation of the CFA has a corresponding color atom $h_l(m,n)$, $h_\alpha(m,n)$ and $h_\beta(m,n)$. For instance, the luma/chroma color atom of the GRBG Bayer pattern is

$$\boldsymbol{h}_{l\alpha\beta}(m,n) = \begin{bmatrix} h_l(m,n) \\ h_\alpha(m,n) \\ h_\beta(m,n) \end{bmatrix}, \qquad (10)$$

where the individual components are

$$h_l = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, h_\alpha = \frac{1}{\sqrt{6}}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, h_\beta = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

*Remark 1:* In principle, there are infinite choices for the luma/chroma basis $\boldsymbol{T}$. We choose the one in Equation (7) because it makes the components of natural images statistically uncorrelated in the first order approximation.

### C. Color Filter Array in Fourier Space

When analyzing the aliasing effects of the CFAs, we need to transform the color atom into the Fourier domain. We denote the Fourier transform of a signal by putting a tilde on top of the symbol, e.g., $h \overset{\mathcal{F}}{\to} \widetilde{h}$. The 2D discrete Fourier transform (DFT) of the $i$-th color atom is

$$\widetilde{h}_i(u,v) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} h_i(m,n)e^{-j2\pi\left(\frac{mu}{M}+\frac{nv}{N}\right)} \qquad (11)$$

where $u \in \{0,\dots,M-1\}$, $v \in \{0,\dots,N-1\}$.

For example, the discrete Fourier transform of the luma/chroma color atoms in Equation (10) are

$$\widetilde{h}_l = \frac{1}{\sqrt{3}}\begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}, \widetilde{h}_\alpha = \frac{1}{\sqrt{6}}\begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}, \widetilde{h}_\beta = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}.$$

Here, we observe that the Fourier transform of the color atom has the same size as the original color atom. The luminance channel has only one baseband components at $(0,0)$, whereas the $\alpha$ chrominance channel has one baseband component and a component at $(\pi,\pi)$. The $\beta$ chrominance channel has two components at $(0,\pi)$ and $(\pi,0)$. Fig. 3 illustrates how these frequency locations are identified from a $3 \times 3$ color atom.

While the Fourier transform of the color atom is useful, for demosaicking we also need to analyze the spectrum of the entire CFA. As shown by Hao *et al.* [24], the Fourier transform of the entire CFA can be written in terms of the Fourier transform of the color atoms:

$$\widetilde{c}_i(\boldsymbol{\omega}) = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} \widetilde{h}_i(u,v)\delta\left(\boldsymbol{\omega} - \boldsymbol{\omega}(u,v)\right), \qquad (12)$$
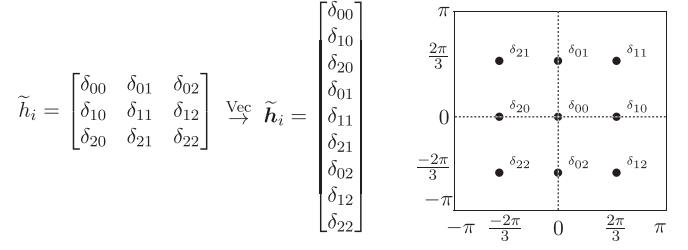


Fig. 3. The Fourier representation of an arbitrary $3 \times 3$ color atom $i$. From left to right: The atom representation, the vector representation and the 2D frequency plane representation. Notice that the frequency plane is divided into 9 regions of size $2\pi/3$, and the spectrum comprises pure sinusoids placed at $\left(\frac{2\pi u}{3}, \frac{2\pi v}{3}\right) \forall u,v \in \{0,1,2\}$.

where $i \in \{l,\alpha,\beta\}$, $\boldsymbol{\omega}$ is the 2D angular frequency, and

$$\boldsymbol{\omega}(u,v) = \left(\frac{2\pi u}{M}, \frac{2\pi v}{N}\right) \begin{array}{l} \forall\,u \in \{0,\dots,M-1\} \\ \forall\,v \in \{0,\dots,N-1\} \end{array}. \qquad (13)$$

is the ($u$-th,$v$-th) 2D angular frequency. It is worth noting that the Fourier transform of the CFA comprises pure sinusoids of amplitudes $\widetilde{h}_i(u,v)$. These sinusoids are placed at $MN$ discrete 2D frequencies $\boldsymbol{\omega}(u,v)$ that divide the 2D frequency plane $[-\pi,\pi] \times [-\pi,\pi]$ into $MN$ equal regions. Therefore, the spectrum of the mosaicked image $\widetilde{\theta}(\boldsymbol{\omega})$ can be written as

$$\widetilde{\theta}(\boldsymbol{\omega}) = \mathcal{F}\left(\sum_{i \in \{l,\alpha,\beta\}} c_i\,\mathrm{im}_i\right) = \sum_{i \in \{l,\alpha,\beta\}} \widetilde{c}_i(\boldsymbol{\omega}) \circledast \widetilde{\mathrm{im}}_i(\boldsymbol{\omega})$$

$$= \sum_{i \in \{l,\alpha,\beta\}}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} \widetilde{h}_i(u,v)\,\widetilde{\mathrm{im}}_i(\boldsymbol{\omega} - \boldsymbol{\omega}(u,v)), \quad (14)$$

where $\circledast$ is the standard 2D convolution operator. Having the spectrum of the mosaicked image $\widetilde{\theta}(\boldsymbol{\omega})$, we can now discuss the optimization variables in our problem.

### D. Design Variables

To formulate the CFA design problem as an optimization problem, we define the following variables. We denote $\boldsymbol{h}_r$, $\boldsymbol{h}_g$ and $\boldsymbol{h}_b$ the vectorized representations of the red, green and blue color atoms, respectively. To ensure physical realizability, we require $\boldsymbol{h}_r, \boldsymbol{h}_g, \boldsymbol{h}_b \in [0,1]^{K\times1}$, where $K \overset{\text{def}}{=} MN$, and we stack all design variables into one long vector

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{h}_r \\ \boldsymbol{h}_g \\ \boldsymbol{h}_b \end{bmatrix} \in \mathbb{R}^{3K\times1}.$$

The design variable $\boldsymbol{x}$ is related to the vectorized RGB and luma/chroma color atoms as

$$\begin{bmatrix} \boldsymbol{h}_r \\ \boldsymbol{h}_g \\ \boldsymbol{h}_b \end{bmatrix} = \begin{bmatrix} \boldsymbol{Z}_r \\ \boldsymbol{Z}_g \\ \boldsymbol{Z}_b \end{bmatrix}\boldsymbol{x} \quad \text{and} \quad \begin{bmatrix} \boldsymbol{h}_l \\ \boldsymbol{h}_\alpha \\ \boldsymbol{h}_\beta \end{bmatrix} = \begin{bmatrix} \boldsymbol{Z}_l \\ \boldsymbol{Z}_\alpha \\ \boldsymbol{Z}_\beta \end{bmatrix}\boldsymbol{x},$$

TABLE I
CFA DESIGN CRITERIA

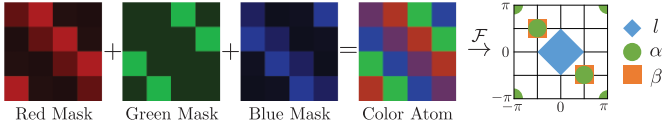| Criterion | Purpose | Regular Pixels | | | | QIS | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | | [21] | [33] | [24] | [36] | [31] | Ours |
| Proposition 1 | To minimize noise power after linear demosaicking | × | ✓ | × | × | × | ✓ |
| Proposition 2 | To simplify denoising of luminance channel | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Proposition 3 | To maximize spatial resolution | ✓ | ✓ | ✓ | × | × | ✓ |
| Proposition 4 | To mitigate cross-talk | × | × | × | ✓ | ✓ | ✓ |
| Definition 4 | Enforce total orthogonality | ✓ | × | ✓ | × | × | ✓ |
| Definition 4 | Enforce quadrature orthogonality | × | ✓ | × | × | × | ✓ |



Fig. 4. A $4 \times 4$ CFA generated by our design framework. Luminance sensitivity $\gamma_l$ and chrominance sensitivity $\gamma_c$ are maximized to improve robustness to noise (Section III-A). No chrominance components ($\alpha$ or $\beta$) are modulated on the vertical and horizontal frequency axes (Section III-B) to mitigate aliasing with the luminance component $l$. The total variation of the red, green and blue masks is upper-bounded by TV$_{\max}$ to mitigate crosstalk (Section III-C).

where the $\boldsymbol{Z}$ matrices are defined by Equation (7) as

$$\boldsymbol{Z}_r = [\boldsymbol{I}, \boldsymbol{0}, \boldsymbol{0}] \quad\quad \boldsymbol{Z}_l = [\boldsymbol{I}, \boldsymbol{I}, \boldsymbol{I}]/\sqrt{3},$$
$$\boldsymbol{Z}_g = [\boldsymbol{0}, \boldsymbol{I}, \boldsymbol{0}] \text{ and } \boldsymbol{Z}_\alpha = [-\boldsymbol{I}, 2\boldsymbol{I}, -\boldsymbol{I}]/\sqrt{6},$$
$$\boldsymbol{Z}_b = [\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{I}] \quad\quad \boldsymbol{Z}_\beta = [\boldsymbol{I}, \boldsymbol{0}, -\boldsymbol{I}]/\sqrt{2}.$$

Given the design variable $\boldsymbol{x}$, we also need to analyze its spectrum. We write the 2D Fourier transform Equation (11) as a matrix-vector product by using the Fourier transform matrix $\boldsymbol{F} \in \mathcal{C}^{K \times K}$. Hence, the vectorized spectra of the luma/chroma color atoms can be written in terms of $\boldsymbol{x}$ as

$$\widetilde{\boldsymbol{h}}_i = \boldsymbol{F}\boldsymbol{h}_i = \boldsymbol{F}\boldsymbol{Z}_i\boldsymbol{x}, \; i \in \{l, \alpha, \beta\}. \tag{15}$$

where $\widetilde{\boldsymbol{h}}_i \in \mathcal{C}^{K \times 1}$, for $i \in \{l, \alpha, \beta\}$. The relation between the matrix and the vector forms of the Fourier transform is:

$$\widetilde{h}_i(u, v) = \text{vec}^{-1}(\widetilde{\boldsymbol{h}}_i) \tag{16}$$

where $\widetilde{h}_i(u, v)$ is the Fourier coefficient.

## III. DESIGN CRITERIA

We now present the design criteria. Our criteria unify the three major approaches in the literature: (i) Sensitivity of luma/chroma channels to noise by Condat [33]; (ii) Aliasing between different color components in the frequency domain by Hirakawa and Wolfe [21]; (iii) Crosstalk between neighboring pixels in the spatial domain by Anzagira and Fossum [31]. Note that the first two criteria were developed for CMOS, whereas the third criterion was developed for QIS. The proposed framework integrates all these criteria into a unified formulation. Table I summarizes the difference between this paper and the previous works. Fig. 4 shows an example CFA, and its frequency representation.

In the following subsections, we present the design criteria and express them in terms of matrix-vectors for the optimization framework in Section IV.

### A. Luminance and Chrominance Sensitivity

*Definition 1:* The luminance sensitivity $\gamma_l$ and the chrominance sensitivity $\gamma_c$ of a CFA with color atom $\{\boldsymbol{h}_l, \boldsymbol{h}_\alpha, \boldsymbol{h}_\beta\}$ of size $M \times N$ are defined as

$$\gamma_l \overset{\text{def}}{=} \frac{1}{K}||\widetilde{\boldsymbol{h}}_l||_2, \text{ and } \gamma_c \overset{\text{def}}{=} \frac{1}{K}\min\left(||\widetilde{\boldsymbol{h}}_\alpha||_2, ||\widetilde{\boldsymbol{h}}_\beta||_2\right). \tag{17}$$

where $K = MN$ is a normalization factor.

Intuitively, the luminance and chrominance sensitivity are measures of the signal power that can be transmitted through the color filter. A more transparent color filter allows more light to pass through, and hence the signal power is higher. This is reflected by the magnitudes $||\widetilde{\boldsymbol{h}}_i||_2$ for $i \in \{l, \alpha, \beta\}$, which according to Parseval's Theorem they are equivalent to $||\boldsymbol{h}_i||_2$.

The following proposition shows how can we compute $\gamma_l$ and $\gamma_c$ in terms of the optimization vector $\boldsymbol{x}$.

*Proposition 1:* For a CFA with color atoms represented by the vector $\boldsymbol{x}$, the luminance and chrominance sensitivity can be calculated as

$$\gamma_l(\boldsymbol{x}) = \frac{1}{K}\mathbf{1}^T\boldsymbol{Z}_l\boldsymbol{x} = \boldsymbol{b}^T\boldsymbol{x}$$

$$\gamma_c(\boldsymbol{x}) = \min\left(\sqrt{\boldsymbol{x}^T\boldsymbol{Q}_\alpha\boldsymbol{x}}, \; \sqrt{\boldsymbol{x}^T\boldsymbol{Q}_\beta\boldsymbol{x}}\right), \tag{18}$$

where $\boldsymbol{b} = \frac{1}{K}\boldsymbol{Z}_l^T\mathbf{1}$, $\boldsymbol{Q}_\alpha = \boldsymbol{Z}_\alpha^T\boldsymbol{Z}_\alpha$ and $\boldsymbol{Q}_\beta = \boldsymbol{Z}_\beta^T\boldsymbol{Z}_\beta$.

*Proof:* See Appendix A. ∎

The luminance sensitivity and the chrominance sensitivity cannot be arbitrarily chosen. One practical consideration is to ensure uniform noise power across the luma channel so that the denoising procedure can be simplified (because the noise will be i.i.d.). Thus, the luminance color atom $h_l$ should be constant, i.e., $h_l(m, n) = c, \forall m, n$, where $c$ is a positive constant. Taking Fourier transform, this means that $\widetilde{h}_l$ comprises only one impulse at baseband $\widetilde{h}_l(0, 0)$, and no impulses at all other frequencies. In vector form, we need

$$\widetilde{\boldsymbol{h}}_l - \text{diag}(\boldsymbol{e}_1)\widetilde{\boldsymbol{h}}_l = \mathbf{0}, \tag{19}$$

where $\boldsymbol{e}_1 = [1, 0, \dots, 0]^T$ is the standard basis. Putting in terms of the optimization variable $\boldsymbol{x}$, we have a constraint.

*Proposition 2. (Uniform Luminance Constraint):* If a CFA has a uniform luminance sensitivity, then $\boldsymbol{x}$ needs to satisfy

$$(\boldsymbol{I} - \text{diag}(\boldsymbol{e}_1))\boldsymbol{F}\boldsymbol{Z}_l\boldsymbol{x} = \mathbf{0}. \tag{20}$$

*Proof:* Using Equation (15), substitute $\widetilde{\boldsymbol{h}}_l = \boldsymbol{F}\boldsymbol{Z}_l\boldsymbol{x}$ into Equation (19). ∎

## B. Anti-Aliasing

In the frequency representation of a mosaicked image, the luminance controls the baseband whereas the chrominance controls the sideband of the spectrum. To minimize spectral interference, i.e., aliasing, CFA design methods modulate the chrominance as far as possible from the baseband. Typically, these methods do not put chrominance on the vertical and horizontal axes to prevent aliasing. However, since QIS is very small and it oversamples the scene, it is possible to relax the aliasing constraint. To this end, we allow chrominance channels to be placed on the vertical and horizontal axis as long as they are placed at frequencies higher than or equal $\pi/2$. This enlarges the feasible set of the solutions.

Mathematically, the anti-aliasing requirement is formulated by forcing the Fourier coefficients of the chrominance color atoms at $(\pm\pi, v)$ and $(u, \pm\pi)$ to zero for all $u$ and $v$ greater than $\pi/2$. In terms of our design variable $\boldsymbol{x}$, we require the following constraint.

*Proposition 3. (Anti-Aliasing Constraint):* The chrominance placed on frequencies less than $\pi/2$ in the vertical and horizontal directions must be set to 0. Hence, $\boldsymbol{x}$ must satisfy

$$\begin{bmatrix} \boldsymbol{W}_\alpha \\ \boldsymbol{W}_\beta \end{bmatrix} \boldsymbol{x} = \boldsymbol{W}\boldsymbol{x} = \boldsymbol{0} \quad (21)$$

where $\boldsymbol{W}_\alpha$ and $\boldsymbol{W}_\beta$ are the matrices formed by choosing the rows in $\boldsymbol{FZ}_\alpha$ and $\boldsymbol{FZ}_\beta$, respectively, that correspond to vertical and horizontal frequency components that are less than $\pi/2$.

To quantify the amount of aliasing for every CFA, we define the following aliasing criterion.

*Definition 2:* For a CFA, aliasing between luminance and chrominance channels is measured by

$$J_l \stackrel{\text{def}}{=} \frac{1}{HW} \int_{[-\pi,\pi)^2} \frac{(S_l(\boldsymbol{\omega})S_\alpha(\boldsymbol{\omega}) + S_l(\boldsymbol{\omega})S_\beta(\boldsymbol{\omega}))}{S_\theta(\boldsymbol{\omega})} d\boldsymbol{\omega}, \quad (22)$$

where $S_l$, $S_\alpha$, $S_\beta$ and $S_\theta$ denote the power spectral density of the luminance channel $\textbf{im}_l$, the two chrominance channels $\textbf{im}_\alpha$ and $\textbf{im}_\beta$, and the mosaicked image $\boldsymbol{\theta}$, respectively.

## C. Crosstalk

Crosstalk is caused by the leakage of electrical and optical charge from a pixel to its adjacent pixels [31], [40]. Crosstalk leads to color de-saturation. To model crosstalk, we follow [31] by defining three scalars $\delta_r$, $\delta_g$, and $\delta_b$ representing the proportion of leaked charges to neighboring pixels. These three scalars then form a *crosstalk kernel*,

$$g_i = \begin{bmatrix} 0 & \delta_i/4 & 0 \\ \delta_i/4 & 1-\delta_i & \delta_i/4 \\ 0 & \delta_i/4 & 0 \end{bmatrix}, \ i \in \{r,g,b\}, \quad (23)$$

which can be considered as a spatial lowpass filter of the mosaicked image. Applying the crosstalk kernel to the CFA is equivalent to a spatially invariant convolution

$$h_i^{\text{ctk}} = g_i \circledast h_i, \ i \in \{r,g,b\},$$



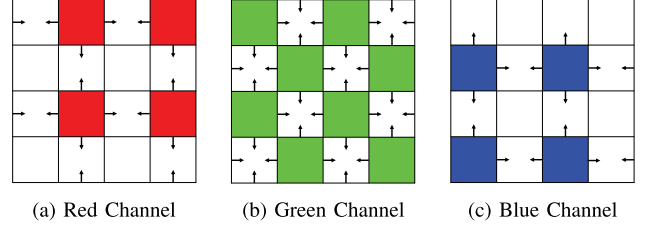(a) Red Channel      (b) Green Channel      (c) Blue Channel

Fig. 5. Crosstalk in Bayer color atom. Each color pixel leak some of its charge to its horizontal and vertical neighbors. Amount of leakage is parametrized by the positive scalars $\delta_r$, $\delta_g$ and $\delta_b$.

where $h_i^{\text{ctk}}$ denotes the effective CFA in the presence of crosstalk.

The effect of crosstalk is more severe when the adjacent colors are different. For example, in Fig. 5, the red and blue pixels are surround by 8 neighbors of different colors and the green pixels are surrounded by 4 neighbors of different colors. This is equivalent to saying that there is a red pixel having a value 1 and is surrounded by pixels having the value 0. Using similar argument, we can see that if the color atoms have more rapid changes of colors, then the resulting CFA is more susceptible to crosstalk.

We propose to quantify the variation of the color atoms (and hence crosstalk) by means of measuring the total variation of the color atom. The total variation is a proxy of the complexity of the color filter array. A color filter array with high total variation means a more complicated pattern and so it is more susceptible to crosstalk. Our total variation is defined as follows.

*Definition 3. (Total Variation):* For a CFA defined by the color atoms $\boldsymbol{h}_r$, $\boldsymbol{h}_g$ and $\boldsymbol{h}_b$, the weighted total variation is defined as

$$\text{TV}(\boldsymbol{x}) \stackrel{\text{def}}{=} \sum_{i \in \{r,g,b\}} \delta_i \|\boldsymbol{Dh}_i\|_1 = \sum_{i \in \{r,g,b\}} \delta_i \|\boldsymbol{DZ}_r\boldsymbol{x}\|_1 \quad (24)$$

where $\boldsymbol{D} \stackrel{\text{def}}{=} [\boldsymbol{D}_x, \boldsymbol{D}_y]^T$ is an operator that computes the vertical and horizontal derivatives with circular boundary conditions, and $\delta_i$ is the leakage factor defined in the crosstalk kernel in Equation (23).

To control the amount of variations in the CFA (so that we can limit the amount of crosstalk), we upper bound the total variation by a scalar $\text{TV}_{\max}$. This leads to the following constraint.

*Proposition 4. (Crosstalk Constraint):* The crosstalk is limited by upper-bounding the total variation metric $\text{TV}(\boldsymbol{x})$:

$$\text{TV}(\boldsymbol{x}) = \sum_{i \in \{r,g,b\}} \delta_i \|\boldsymbol{DZ}_r\boldsymbol{x}\|_1 \leq \text{TV}_{\max}. \quad (25)$$

Fig. 6 shows two CFAs proposed in literature. The first one, proposed in [24] is more robust to aliasing than the second one proposed in [31]. This is because the chrominance channels are modulated at high frequencies which are far from baseband luminance. However, [31] is more robust to crosstalk than [24] because the color atom have less variation in colors. We can also see this in the total variation values: 0.206 compared to 0.131 at $(\delta_r, \delta_g, \delta_b) = (0.23, 0.15, 0.10)$. This trade-off constitutes a gap in literature: Color filter designs can improve robustness of

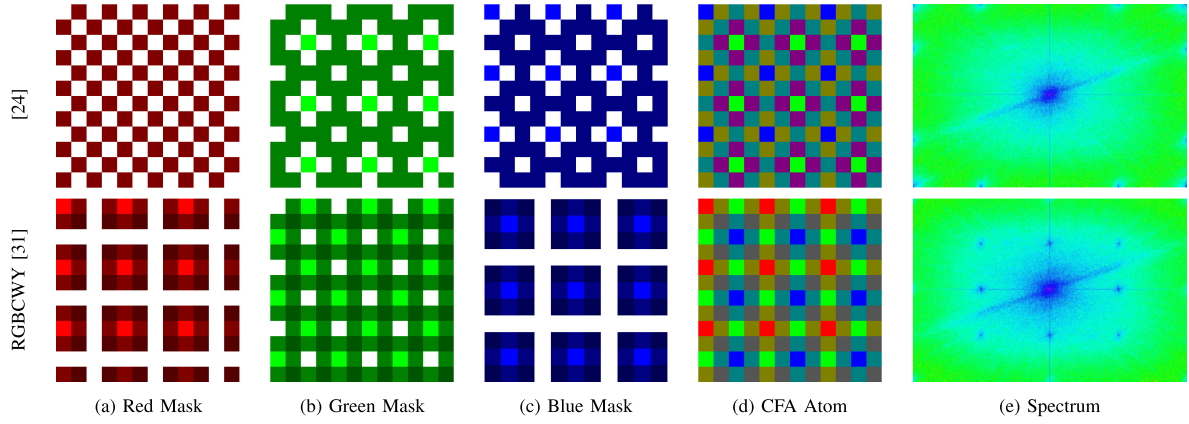|  | (a) Red Mask | (b) Green Mask | (c) Blue Mask | (d) CFA Atom | (e) Spectrum |

Fig. 6. Examples of two CFAs. (Top row) proposed in [24], this array has good aliasing properties, where chrominance channels are placed far away from luminance channel, but it has bad crosstalk properties: $\text{TV}(\boldsymbol{x}) = 0.413$. (Bottom row) Proposed in [31], this array has good crosstalk properties $\text{TV}(\boldsymbol{x}) = 0.263$, but it has bad aliasing properties.

either aliasing or crosstalk, but not for both. Our proposed design framework allows us to optimize them simultaneously.

Since the best value for $TV_{\max}$ changes for different atom sizes, we choose it according to the best value in the state-of-the-art CFA atoms. For example, we choose $TV_{\max} = 0.131$ for $4 \times 4$ atoms since this is the lowest value achieved by Anzagira and Fossum $4 \times 4$ RGBCWY CFA [31]. Other $4 \times 4$ CFAs have higher $TV$ values.

### D. Orthogonality of Chrominance Channels

When designing a CFA, one should take into consideration of the complexity of the demosaicking process. Recall Equation (14) where we show that

$$\widetilde{\theta}(\boldsymbol{\omega}) = \sum_{i \in \{l, \alpha, \beta\}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \widetilde{h}_i(u, v) \, \widetilde{\text{im}}_i(\boldsymbol{\omega} - \boldsymbol{\omega}(u, v)).$$

This is a modulation of the signal by a modulating frequency $\boldsymbol{\omega}(u, v)$. Therefore, to reconstruct the signal, one approach is to demodulate by shifting the channels to the baseband by multiplying pure sinusoids and then applying a lowpass filter [33]. Demodulation can be done efficiently if there is orthogonality between the channels. Following the literature, our optimization takes into account of two forms of orthogonality.

- Total Orthogonality [24] and [21]: The idea is to make one chroma channel zero and the other non-zero at any $(u, v)$. For example, $\widetilde{h}_\alpha(u, v) = 0.9$ and $\widetilde{h}_\beta(u, v) = 0$.
- Quadrature Orthogonality [33]: The idea is to make one chroma channel real and the other imaginary at any $(u, v)$, i.e., the two channels are modulated by a frequency $\boldsymbol{\omega}(u, v)$ but in quadrature phase. Translating the spatial domain, this means that

$$h_\alpha(m, n) = \gamma_c \sqrt{2} \cos\left( \boldsymbol{\omega}(u, v)^T \begin{bmatrix} m \\ n \end{bmatrix} - \phi \right) \quad (26a)$$

$$h_\beta(m, n) = \gamma_c \sqrt{2} \sin\left( \boldsymbol{\omega}(u, v)^T \begin{bmatrix} m \\ n \end{bmatrix} - \phi \right) \quad (26b)$$

where $m \in \{0, \ldots, M-1\}$, $n \in \{0, \ldots, N-1\}$, and $\phi$ is the phase angle. In this way, the two channels can be easily separated during the demosaicking process using the orthogonality of cosine and sine functions.

We formulate the orthogonality criteria as a penalty function that is a surrogate of both approaches.

*Definition 4. (Orthogonality Penalty):* For a CFA having chrominance channels with spectra $\widetilde{h}_\alpha(\boldsymbol{\omega})$ and $\widetilde{h}_\beta(\boldsymbol{\omega})$, the orthogonality penalty is defined as

$$\rho(\boldsymbol{h}_\alpha, \boldsymbol{h}_\beta) \stackrel{\text{def}}{=} \sum_{u=0}^{M} \sum_{v=0}^{N} \left( |\Re\widetilde{h}_\alpha(u, v)| + |\Re\widetilde{h}_\beta(u, v)| \right)$$
$$+ \sum_{u=0}^{M} \sum_{v=0}^{N} \left( |\Im\widetilde{h}_\alpha(u, v)| + |\Im\widetilde{h}_\beta(u, v)| \right) \quad (27)$$

which can be written as a function in $\boldsymbol{x}$ as follows

$$\rho(\boldsymbol{x}) = \left( \|\Re\boldsymbol{F}\boldsymbol{Z}_\alpha\boldsymbol{x}\|_1 + \|\Re\boldsymbol{F}\boldsymbol{Z}_\beta\boldsymbol{x}\|_1 \right)$$
$$+ \left( \|\Im\boldsymbol{F}\boldsymbol{Z}_\alpha\boldsymbol{x}\|_1 + \|\Im\boldsymbol{F}\boldsymbol{Z}_\beta\boldsymbol{x}\|_1 \right) \quad (28)$$

Looking at the first summation in Equation (27), we notice that for every 2D frequency $(u, v)$, the term $|\Re\widetilde{h}_\alpha(u, v)| + |\Re\widetilde{h}_\beta(u, v)|$ is the $\ell_1$-norm of a 2-dimensional vector $[\Re\widetilde{h}_\alpha(u, v), \Re\widetilde{h}_\beta(u, v)]^T$. Therefore, minimizing this $\ell_1$-norm promotes either one of the components to zero (or both). Similar argument applies for the imaginary components in the second summation. As a result, the total variation can be regarded as a proxy to the orthogonality condition.

### E. Condition Number

When designing a color filter array, one should also be aware of the simplicity of the demosaicking algorithm. Since the luminance/chrominance transformation, color filtering and crosstalk are all linear processes, we can represent them by an overall color acquisition matrix $\boldsymbol{A}$. To demosaic the image, in principle we need to invert the $\boldsymbol{A}$ matrix. To avoid the amplification of the estimation error of luminance and chrominance channels, the

condition number of $A$ should be minimized for numerical stability. This metric was discussed in [24], but only the condition number of the luminance/chrominance transformation matrix $T$ was considered. In this paper, we generalize this metric by taking the color filtering and crosstalk into account as well.

To represent the image acquisition in frequency domain as a linear process, we assume the crosstalk kernels for red, green and blue pixels are the same $g_r = g_g = g_b$. Define the following frequency domain variables:

$$\widetilde{\mathbf{im}}_{rgb} = \begin{bmatrix} \widetilde{\mathbf{im}}_r^T \\ \widetilde{\mathbf{im}}_g^T \\ \widetilde{\mathbf{im}}_b^T \end{bmatrix}, \widetilde{H} = [\widetilde{\boldsymbol{h}}_l, \widetilde{\boldsymbol{h}}_\alpha, \widetilde{\boldsymbol{h}}_\beta], \text{ and } \widetilde{G} = \mathrm{diag}(\widetilde{\boldsymbol{g}})$$

(29)

where $\widetilde{\boldsymbol{g}} = (\mathcal{F}\boldsymbol{g})$ is the vectorized version of the $M \times N$ discrete Fourier transform of the crosstalk kernel $g$. Hence, the mosaicked image $\widetilde{\theta}$ can be written as

$$\widetilde{\boldsymbol{\theta}} = \widetilde{G}\widetilde{H}\widetilde{T}\widetilde{\mathbf{im}}_{rgb} = A\widetilde{\mathbf{im}}_{rgb} \tag{30}$$

where we define the color acquisition matrix as $A \stackrel{\text{def}}{=} \widetilde{G}\widetilde{H}\widetilde{T}$. Denote by $\kappa(A)$ the condition number of $A$, i.e.,

$$\kappa(A) = \mathrm{cond}(A) \in [1, \infty] \tag{31}$$

Low values of $\kappa(A)$ imply stable demosaicking process that involves mild amplification of estimation errors in the luminance and chrominance components.

## IV. FORMULATION OF OPTIMAL CFA DESIGN PROBLEM

Using the variables and constraints defined in the previous section, we present two different optimization formulations of the CFA design problem in this section: (i) A non-convex formulation that integrates all the above information into a single optimization, and (ii) convex relaxation that makes the problem more tractable.

### A. Non-Convex CFA Design

The non-convex CFA optimization puts the objectives and constraints defined in the previous section into a single optimization problem. This gives us

$$\underset{\boldsymbol{x}}{\text{maximize}} \quad \gamma_c(\boldsymbol{x}) + \lambda_l \gamma_l(\boldsymbol{x}) - \lambda_\rho \rho(\boldsymbol{x})$$

subject to

$$\begin{array}{ll} \boldsymbol{x} \in [0,1]^{3L} & \text{(Realizability) (a)} \\ (\boldsymbol{I} - \mathrm{diag}(\boldsymbol{e}_1))\boldsymbol{F}\boldsymbol{Z}_l\,\boldsymbol{x} = \boldsymbol{0} & \text{(Proposition 2) (b)} \\ \boldsymbol{W}\boldsymbol{x} = \boldsymbol{0} & \text{(Proposition 3) (d)} \\ \mathrm{TV}(\boldsymbol{x}) \leq \mathrm{TV}_{\max} & \text{(Proposition 4) (e)} \end{array}$$

(32)

where $\lambda_l$ and $\lambda_\rho$ are the regularization parameters controlling the relative weights of the luminance sensitivity and the orthogonality penalty. The penalty function $\rho(\boldsymbol{x})$ is added to the objective with a negative sign so that it is minimized. By lower bounding $\gamma_c(\boldsymbol{x})$ with a constant $\tau$, we can rewrite Equation (32) as

---

**Algorithm 1:** Successive Convex Approximations.

**Require:** Initial guess $\boldsymbol{x}^{(0)}$, $k = 0$.

  **while** $\gamma_c$ not converge **do**

    Replace the quadratic terms $\boldsymbol{x}^T\boldsymbol{Q}_\alpha\boldsymbol{x}$ and $\boldsymbol{x}^T\boldsymbol{Q}_\beta\boldsymbol{x}$ in inequalities (33)(f-g) by their first order Taylor approximations around $\boldsymbol{x}^{(k)}$:

$$\boldsymbol{x}^T\boldsymbol{Q}_\alpha\boldsymbol{x} \approx \boldsymbol{x}^{(k)T}\boldsymbol{Q}_\alpha\boldsymbol{x}^{(k)} + 2(\boldsymbol{x} - \boldsymbol{x}^{(k)})\boldsymbol{Q}_\alpha\boldsymbol{x}^{(k)} \geq \tau^2$$

$$\boldsymbol{x}^T\boldsymbol{Q}_\beta\boldsymbol{x} \approx \boldsymbol{x}^{(k)T}\boldsymbol{Q}_\beta\boldsymbol{x}^{(k)} + 2(\boldsymbol{x} - \boldsymbol{x}^{(k)})\boldsymbol{Q}_\beta\boldsymbol{x}^{(k)} \geq \tau^2$$

    Solve the convex approximation of (33) to get $\gamma_c^{(k)}$

    $k = k + 1$

  **end while**

  **return** $\boldsymbol{x}$

---

$$\underset{\boldsymbol{x},\tau}{\text{maximize}} \quad \tau + \lambda_l \gamma_l(\boldsymbol{x}) - \lambda_\rho \rho(\boldsymbol{x})$$

subject to

$$\begin{array}{ll} \boldsymbol{x} \in [0,1]^{3L} & \text{(Realizability) (a)} \\ (\boldsymbol{I} - \mathrm{diag}(\boldsymbol{e}_1))\boldsymbol{F}\boldsymbol{Z}_l\,\boldsymbol{x} = \boldsymbol{0} & \text{(Proposition 2) (b)} \\ \boldsymbol{W}\boldsymbol{x} = \boldsymbol{0} & \text{(Proposition 3) (d)} \\ \mathrm{TV}(\boldsymbol{x}) \leq \mathrm{TV}_{\max} & \text{(Proposition 4) (e)} \\ \boldsymbol{x}^T\boldsymbol{Q}_\alpha\boldsymbol{x} \geq \tau^2 & \text{(Proposition 1) (f)} \\ \boldsymbol{x}^T\boldsymbol{Q}_\beta\boldsymbol{x} \geq \tau^2 & \text{(Proposition 1) (g)} \end{array}$$

(33)

In this optimization problem, the objective and constraints are convex except for (33)(f) and (33)(g). This is because these inequalities include convex quadratic form in the "$\geq$" side, where convexity comes from the fact that $\boldsymbol{Q}_\alpha$ and $\boldsymbol{Q}_\beta$ are positive semidefinite matrices. Hence, the optimization problem is non-convex.

### B. Solving the Optimization

While problem (33) is non-convex, we can find a local minimum by successive convex approximations [41]. The idea of successive convex approximation is to replace the quadratic terms in the non-convex constrains (33)(f) and (33)(g) by first order approximations around the initial guess $\boldsymbol{x}^{(0)}$. Since the quadratic form is convex, its first order approximation constitutes a lower bound. Hence, we are replacing the non-convex constraints (33)(f) and (33)(g) with convex but tighter constraints that limit the feasible set of $\boldsymbol{x}$. The algorithm repeats until $\tau$ converges to a fixed-point, which is the final chrominance sensitivity.

The overall algorithm is summarized in Algorithm 1. Fig. 7 shows the convergence of Algorithm 1 for designing a $4 \times 4$ color atom. We notice the monotonic increase of $\tau$ until it converges to a fixed point. Since the original problem is non-convex, solution to the problem could be a local minimum depending on how the initialization is done. In practice, we solve the problem for multiple instances with different randomly generated initial guesses which approximately cover the design space (e.g., using the Latin hypercube sampling [42]), and pick the best solution among them.
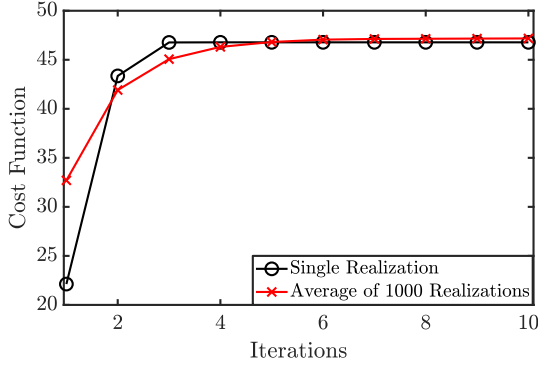
Fig. 7. Convergence of Algorithm 1 for $4 \times 4$ color filter design for single realization and average of 1000 realizations.



(a) Convex Formulation
$\gamma_c(\boldsymbol{x}) = 0.08$

(b) Non-Convex Formulation
$\gamma_c(\boldsymbol{x}) = 0.12$

Fig. 8. $4 \times 4$ color atoms and corresponding spectra obtained using convex and non-convex formulations. Spectra are obtained from mosaicking the "Bikes" image in Kodak color dataset by the color atoms. Both have the same luminance sensitivity $\gamma_l(\boldsymbol{x}) = 0.577$ and same Total variation $TV(\boldsymbol{x}) = 0.131$.

### C. Convex CFA Design

The relaxation from non-convex to convex can be done by explicitly forcing part of the chrominance components to zero. Specifically, we modulate the chrominance channels on the same frequency $\boldsymbol{\omega}(u, v) = (\frac{2\pi u}{M}, \frac{2\pi v}{N})$ using the quadrature orthogonality mentioned in Equation (26). In terms of $\boldsymbol{x}$, these two equations can be written as:

$$\boldsymbol{Z}_\alpha \boldsymbol{x} = \gamma_c \boldsymbol{x}_c, \quad \boldsymbol{Z}_\beta \boldsymbol{x} = \gamma_c \boldsymbol{x}_s \quad (34)$$

where $\boldsymbol{x}_c$ and $\boldsymbol{x}_s$ are constant vectors that represent the vectorized version of the cosine and sine signals on the right hand side of Equations (26a) and (26b), respectively, i.e.,

$$\boldsymbol{x}_c = \text{vec} \left\{ \sqrt{2} \cos \left( \boldsymbol{\omega}(u, v)^T \begin{bmatrix} m \\ n \end{bmatrix} - \phi \right)_{m=0, n=0}^{M-1, N-1} \right\} \quad (35a)$$

$$\boldsymbol{x}_s = \text{vec} \left\{ \sqrt{2} \sin \left( \boldsymbol{\omega}(u, v)^T \begin{bmatrix} m \\ n \end{bmatrix} - \phi \right)_{m=0, n=0}^{M-1, N-1} \right\} \quad (35b)$$

Since we explicitly choose the modulation frequencies of chrominance channels manually, we can drop the aliasing constraint in Proposition 3. However, we still need the uniform luminance constraint in Propositions 2. Since the luminance and chrominance gains are conflicting, the objective of this formulation is to maximize their weighted sum. To this end, the problem is written as:

$$\underset{\boldsymbol{x}, \gamma_c}{\text{maximize}} \quad \gamma_c + \lambda_l \gamma_l(\boldsymbol{x})$$

subject to
$$\begin{aligned}
\boldsymbol{x} &\in [0, 1]^{3L} & \text{(Realizability)} & \quad \text{(a)} \\
(\boldsymbol{I} - \text{diag}(\boldsymbol{e}_1))\boldsymbol{F}\boldsymbol{Z}_l\,\boldsymbol{x} &= \boldsymbol{0} & \text{(Proposition 2)} & \quad \text{(b)} \\
\text{TV}(\boldsymbol{x}) &\leq \text{TV}_{\max} & \text{(Proposition 4)} & \quad \text{(d)} \\
\boldsymbol{Z}_\alpha \boldsymbol{x} - \gamma_c \boldsymbol{x}_c &= 0 & & \quad \text{(e)} \\
\boldsymbol{Z}_\beta \boldsymbol{x} - \gamma_c \boldsymbol{x}_s &= 0 & & \quad \text{(f)} \\
& & & (36)
\end{aligned}$$

In our terminology, the optimization problem of [33] is obtained from Equation (36) by removing the crosstalk constraint (36)(d). Hence, our optimization limits the search space of the optimization in [33] to get CFAs that have acceptable crosstalk performance.
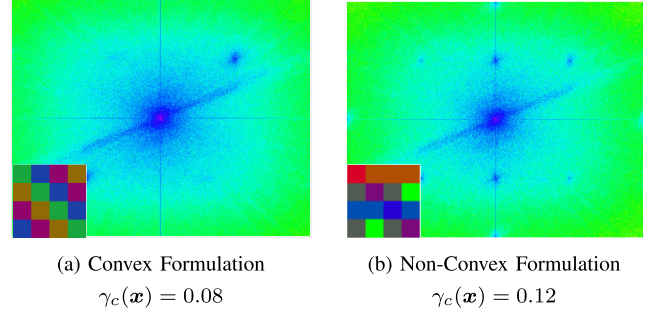
Fig. 8 shows two color atoms obtained using the convex and non-convex formalizations. In the convex formulation, we select the modulation frequency to be $\omega_0 = [\pi, \pi]$ and the phase that maximizes $\gamma_c$ at this frequency is found to be $\phi = \pi/12$. Then, we solve the problem to get $(\gamma_l, \gamma_c, TV) = (0.573, 0.08, 0.263)$. As for the non-convex formulation, we let the optimization to choose modulation frequencies subject to crosstalk and aliasing constraints. Solving the non-convex formulation yields $(\gamma_l, \gamma_c, TV) = (0.573, 0.12, 0.263)$. We notice that the non-convex formulation achieves higher chrominance sensitivity because of its flexibility in choosing the modulation frequencies.

## V. UNIVERSAL DEMOSAICKING

In this section, we present a universal demosaicking algorithm which can be used to all CFAs presented in this paper. Our algorithm comprises two main parts: (i) a demosaicking step to remove the color filtering effect (Section V-B) and (ii) a color correction step to mitigate the crosstalk effect (Section V-C).

### A. Special Consideration for QIS.

Before we talk about the demosaicking algorithm, we should briefly discuss the photon statistics of QIS. In CMOS, the measured voltage can be modeled as a nominal value corrupted by i.i.d. Gaussian noise. For single-bit QIS, previous work showed that the measured photon counts follow a truncated Poisson process [5]. When averaging over a number of temporal frames, the truncated Poisson becomes a Binomial [6]. If the photon count is sufficiently high, this binomial will approximately approach a Gaussian. Applying the law of large numbers on the distribution of $\boldsymbol{B}$ in Equation (5), the average is

$$\frac{1}{T} \sum_{t=0}^{T-1} b_{j,t} \xrightarrow{a.s.} \mathbb{E}[B_j] = 1 - \Psi_q(\theta_j),$$

and the maximum-likelihood estimate of the signal is

$$\theta_j = \Psi_q^{-1} \left( 1 - \frac{1}{T} \sum_{t=1}^{T} b_{j,t} \right)$$

As discussed in [6], we can regard this equation as a tone-mapping of the photon counts. We regard $\theta_j$ as the $j$-th pixel
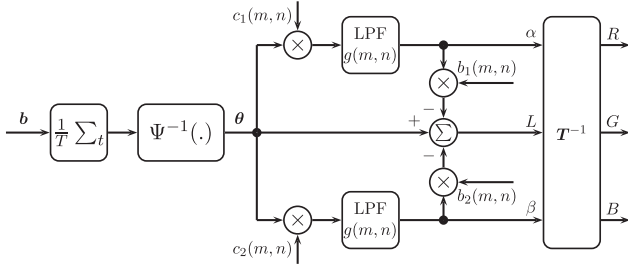
Fig. 9. Illustration of Algorithm 2 of demosaicking by frequency selection for a special case of a CFA that has strictly one replica of the $\alpha$ and $\beta$ chrominance channels. Variable on the figure are defined in Algorithm 2.

of the mosaicked image generated by the CFA. The goal of demosaicking is to reconstruct a color image from $\theta_j$.

### B. Demosaicking by Frequency Selection

Our demosaicking algorithm is based on frequency selection [11]. It generalizes [12] as it works for any CFA as long as it satisfies the orthogonality constraints in Section III-D

The key idea of the algorithm is to shift every chrominance channel to the baseband by multiplying with its carrier, then use a low-pass filter to reconstruct it. For chrominance components that are replicated over distinct carriers, we combine them by simple averaging. After obtaining the $\alpha$ and $\beta$ chrominance channels, they are re-modulated to their original positions and subtracted from the mosaicked image to obtain the luminance channel. This process is summarized in Algorithm 2 for a special case of a CFA that has strictly one replica of the $\alpha$ and $\beta$ chrominance channels. It is also illustrated by Fig. 9. Extension of the algorithm to the general case is straightforward.

To apply Algorithm 2 on CFAs proposed in [24], [31] and [36], they must satisfy the orthogonality constraints in Section III-B. However, this is not satisfied with our choice of the luminance/chrominance basis defined by $T$ in Equation (7). Hence, we use for every CFA the transformation matrix $T$ that makes its luminance and chrominance channel orthogonal. To ensure fairness, we normalize the matrix rows by dividing by their $l_2$ norm so that all luminance and chrominance have the same noise power. The transformation matrices are provided in the supplementary.

### C. Color Correction

The demosaicking algorithm in Algorithm 2 does not take into account the crosstalk effect. Like most of the mainstream image and signal processing (ISP) pipelines, we reduce the cross-talk via a color correction step.

Given the demosaicked color pixel $\widehat{\text{im}}(m, n)$, the color correction multiplies $\widehat{\text{im}}(m, n)$ by a $3 \times 3$ matrix $M$ such that $M\widehat{\text{im}}(m, n)$ is the color-corrected pixel. The matrix $M$ is learned by comparing a measured color pixel to a known color chart value. Mathematically, suppose we have $K$ measured color pixels forming a $3 \times K$ matrix $Q_{\text{False}}$, and a corresponding true color values forming another $3 \times K$ matrix $Q_{\text{GT}}$, we can estimate $M$ by solving

$$M = \arg\min_{M} \epsilon_c(M) \qquad (37)$$

---

**Algorithm 2:** Demosaicking by Frequency Selection.

**Require:** The image $\theta$ which is mosaicked by a CFA of size $M \times N$ as defined in Equation (8), a luminance/chrominance transformation matrix $T$, a low-pass filter $g$, a scalar $K \stackrel{\text{def}}{=} MN$ and a scalar $r = \begin{cases} 2 & \text{if } \boldsymbol{\omega} = (\pi, \pi) \\ 1 & \text{otherwise} \end{cases}$.

**Ensure:** $\alpha$ and $\beta$ chrominance channels are modulated on carriers $\boldsymbol{\omega}(u_1, v_1)$ and $\boldsymbol{\omega}(u_2, v_2)$, respectively.

1) Reconstruct the $\alpha$ chrominance channel

$$\alpha(m, n) = (\theta(m, n)c_1(m, n)) \circledast g(m, n)$$

where

$$c_1(m, n) = \frac{K}{|a_1|} \cos\left( \boldsymbol{\omega}(u_1, v_1)^T \begin{bmatrix} m \\ n \end{bmatrix} + \angle a_1 \right)$$

and $a_1 = \widetilde{h}_\alpha(u_1, v_1)$

2) Reconstruct the $\beta$ chrominance channel

$$\beta(m, n) = (\theta(m, n)c_2(m, n)) \circledast g(m, n)$$

where

$$c_2(m, n) = \frac{K}{|a_2|} \cos\left( \boldsymbol{\omega}(u_2, v_2)^T \begin{bmatrix} m \\ n \end{bmatrix} + \angle a_2 \right)$$

and $a_2 = \widetilde{h}_\beta(u_2, v_2)$

3) Reconstruct the luminance channel

$$L(m, n) = \theta(m, n) - \alpha(m, n)b_1(m, n) - \beta(m, n)b_2(m, n)$$

where

$$b_1(m, n) = \frac{2|a_1|^2}{rK^2} c_1(m, n) \text{ and } b_2(m, n) = \frac{2|a_2|^2}{rK^2} c_2(m, n)$$

**return** $[R, G, B]^T = T^{-1}[L, \alpha, \beta]^T$

---

where $\epsilon_c(M) = \text{Tr}\{(MQ_{\text{False}} - Q_{\text{GT}})^T (MQ_{\text{False}} - Q_{\text{GT}})\}$ is the squared color error. To minimize the noise amplification, it is advised to add regularization when estimating $M$ [43]. Since a standard color chart comprises 24 color patches, we can estimate the noise by computing the norm of covariance matrix of every color patch, and get the average value over the 24 color patches. Hence, the optimization problem is rewritten as

$$M = \arg\min_{M} \epsilon_c(M) + \mu \sum_{i=1}^{24} ||\text{Cov}(MQ_{\text{False}}^{(i)})||_2^2 \qquad (38)$$

where $Q_{\text{False}}^{(i)}$ represents the pixels of the $i$th color patch, and $\mu$ is a positive scalar that controls the noise amplification effect. By varying $\mu$, we can draw a tradeoff curve between color reproduction accuracy and noise amplification.

Fig. 10 shows reconstructed images before and after color correction. In this figure, the crosstalk parameters are $(\delta_r, \delta_r, \delta_r) = (0.23, 0.15, 0.1)$. We can see the effect of color correction in the more saturated red and yellow feathers and in the green leaves in the background.

TABLE II
CFA PARAMETERS AND RECONSTRUCTION QUALITY MEASURED BY YSNR AND SMI METRICS ON MACBETH COLORCHECKER AND AVERAGE CPSNR ON DIV2K EVALUATION DATASET. AN ARROW IS PLACED AFTER EACH METRIC TO SHOW WHETHER HIGHER OR LOWER IS BETTER

| Size | CFA Pattern | Experiment 1 | | | | | Experiment 2 | | | | Experiment 3 | |
| | | CFA Parameters | | | | | YPSNR (dB) ↑ | | SMI (%) ↑ | | CPSNR (dB) ↑ | |
| | | $\gamma_l$ ↑ | $\gamma_c$ ↑ | TV ↓ | $J_l$ ↓ | $\kappa(\boldsymbol{A})$ ↓ | w/o Ctk | w/ Ctk | w/o Ctk | w/ Ctk | w/o Ctk | w/ Ctk |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 × 4 | RGBCY [31] | 0.679 | **0.125** | 0.156 | 6.53 | 2.376 | 23.24 | 22.15 | 93.65 | 92.90 | 30.65 | 30.63 |
| | RGBCWY [31] | 0.597 | 0.102 | **0.131** | 2.67 | 2.143 | 23.57 | 22.45 | 93.71 | 93.10 | 30.75 | 30.74 |
| | Hao et al. [24] | 0.586 | 0.094 | 0.206 | **1.00** | 2.305 | 24.13 | 18.50 | **94.34** | 91.31 | **31.34** | 30.83 |
| | Ours | 0.577 | 0.121 | **0.131** | 2.86 | **1.795** | **24.16** | **23.17** | 94.07 | **93.47** | 31.10 | **31.12** |
| 3 × 3 | Biay-Cheng et al. [36] | 0.612 | **0.167** | 0.175 | 1.65 | 1.892 | 23.92 | 22.04 | 94.44 | 93.88 | 31.07 | 31.14 |
| | Ours | 0.577 | 0.160 | 0.175 | 1.74 | **1.561** | **24.38** | **23.64** | **94.49** | **94.36** | **31.24** | **31.37** |
| 3 × 2 | Condat [33] | 0.866 | 0.250 | 0.317 | **1.16** | 1.915 | **25.33** | 21.68 | **94.71** | **93.87** | **31.30** | 31.17 |
| | Ours | 0.866 | 0.250 | **0.212** | 1.40 | **1.837** | 25.28 | **24.30** | 94.25 | 93.77 | 31.27 | **31.32** |
| 4 × 2 | Hirakawa-Wolfe [21] | 0.866 | 0.125 | 0.275 | **0.98** | 2.767 | **25.38** | 21.22 | **94.57** | 93.58 | **31.42** | 31.08 |
| | Ours | 0.866 | **0.187** | **0.256** | 1.27 | **2.166** | 24.89 | **22.34** | 94.36 | **94.03** | 31.40 | **31.31** |
| 7 × 7 | Bai et al. [28] | 0.577 | 0.091 | 0.204 | **1.87** | 2.242 | 24.26 | 19.68 | **94.56** | 92.94 | **31.41** | 31.09 |
| | Ours | 0.577 | **0.238** | 0.204 | 1.45 | **1.294** | **24.86** | **23.85** | 94.45 | **94.35** | 31.01 | **31.20** |



Fig. 10. Effect of color correction on retaining vivid image colors.

## VI. EXPERIMENTAL EVALUATION

In this section, we present CFAs obtained using our optimization framework in Section VI-A. Afterwards, we evaluate the performance of different CFAs using Algorithm 2. First, we simulate in Section VI-B the reconstruction performance of a Macbeth ColorChecker under D65 illuminant. Second, we show in Section VI-C a quantitative and qualitative comparison of the reconstruction performance of all CFAs on DIV2K evaluation dataset [44]. This dataset comprises 100 color images of resolution $2040 \times 1356$. We downsample them to $1020 \times 678$ for our experiments. In all experiments, QIS parameters are chosen as $(q, \eta, T) = (1, 2, 1000)$, where the value of $T$ is set as a high value to minimize the noise power.

### A. Experiment 1: Proposed Solutions of CFA Design Problem

We focus on the non-convex formulation (33) since it is more flexible than the convex formulation (36). We set the parameters of formulation (33) as $\lambda_l = 0.1$ and $\lambda_\rho = 0.02$. We run multiple instances of Algorithm 1 (2000 instances) using different random initializations for the color atoms $\boldsymbol{x}^{(0)}$. Then, we pick the solution with the highest chrominance sensitivity. To ensure that the initial guess spans the feasible set of $\boldsymbol{x}$, we use uniform Latin hypercube sampling of the domain $[0, 1]^{3L}$.

Fig. 11 shows our proposed CFAs and their accompanied spectra compared to other CFAs in the literature. For every CFA, we compute 1) the luminance and chrominance gains ($\gamma_l$ and $\gamma_c$) in Equation (17) to measure robustness to noise, 2) the total variation metric $TV(x)$ (Proposition 4) to measure robustness to crosstalk, and 3) the aliasing metric $J_l$ in Equation (22)

to measure robustness to aliasing, and the condition number $\kappa(\boldsymbol{A})$ defined in Section III-E. To calculate the aliasing metric for [21], [24], [31] and [36], we use the transformation matrices that make the luminance and chrominance channels orthogonal as mentioned in Section V-B. Leakage factors $(\delta_r, \delta_g, \delta_b)$ are chosen as $(0.23, 0.15, 0.10)$. Results are summarized in Table II.

- $4 \times 4$: Among all $4 \times 4$ CFAs in Table II, Hao *et al.* [24] is the most robust CFA to aliasing, but the least robust to crosstalk; whereas RGBCWY [31] is the most robust to crosstalk, but it is not as robust to aliasing. Our CFA achieves the best of both worlds by having the same total-variation like RGBCWY, and good aliasing metric. Moreover, it has the lowest condition number and it is more sensitive since it comprises 4 panchromatic or "white" pixels with $(r, g, b) = (1/3, 1/3, 1/3)$.
- $3 \times 3$: Compared to Biay-Cheng *et al.* [36], our CFA has less aliasing. The high aliasing metric of [36] is attributed to its design which overlooks frequency domain aliasing.
- $3 \times 2$: Compared to Condat [33], our CFA is more robust to crosstalk and aliasing. It is worth noting that we get exactly the same CFA as Condat's automatically when we remove the crosstalk constraint.
- $4 \times 2$: Compared to Hirakawa-Wolfe [21], our CFA has higher chrominance sensitivity and it is more robust to crosstalk.
- $7 \times 7$: Compared to Bai *et al.* [28], our CFA is more robust to crosstalk and aliasing.

### B. Experiment 2: Macbeth ColorChecker Reconstruction

In this experiment, we simulate the performance of different CFAs in reconstructing the Macbeth ColorChecker image. Pixel response is determined using the incident photon flux of D65 light and the spectral reflectance of Macbeth ColorChecker integrated over the visible light spectrum. QIS parameters and primary color filters are taken from [31]. For every CFA, we generate mosaicked images under two scenarios: 1) crosstalk kernels with leakage factors $(\delta_r, \delta_g, \delta_b) = (0, 0, 0)$, i.e., no crosstalk, and 2) crosstalk kernels with leakage factors $(\delta_r, \delta_g, \delta_b) = (0.45, 0.30, 0.20)$ as suggested by [31]. For fairness of comparison, we use Algorithm 2 for demosaicing all CFAs including RGBCY and RGBCWY CFAs proposed in [31].
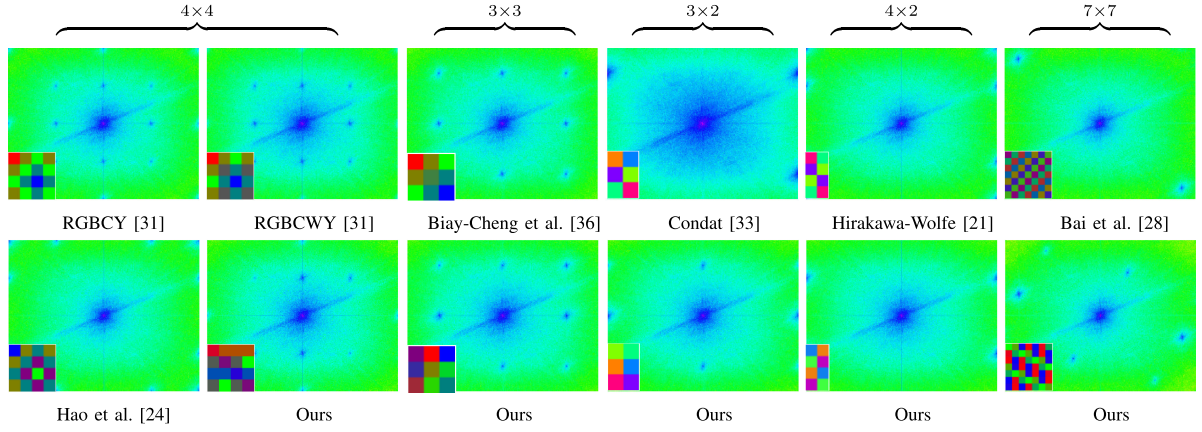
Fig. 11. Our proposed CFAs compared to other CFAs in literature. For every CFA, we show the spectrum of the "bike" Kodak image mosaicked by this CFA. We also show the organization of luminance ($L$) and chrominance channels ($\alpha$ and $\beta$) for CFAs that satisfy orthogonality constraint. Note that the proposed $4 \times 4$ CFA has 3 similar colors in the first row because of the crosstalk constraint. The optimization is still minimized due to the high chrominance sensitivity and high condition number.
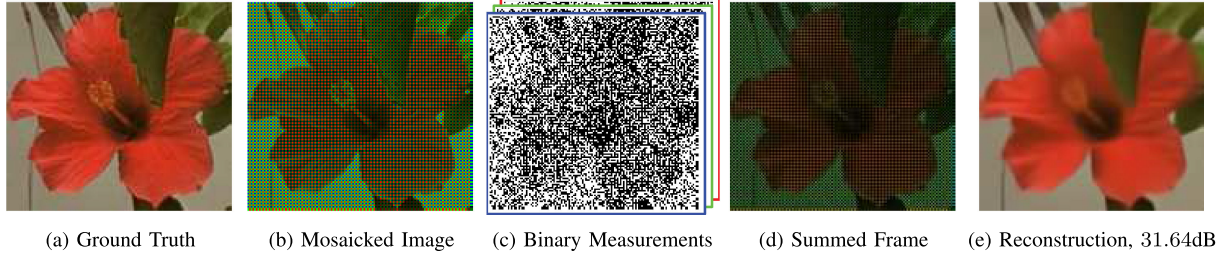


Fig. 12. Simulation Setup: The ground truth image (a) is color filtered by a CFA to produce a mosaicked image (b). QIS generates $T = 1000$ binary frames (c) using the mosaicked image as light exposure. The $T$ binary framed are summed to give an approximately clean image (d). Then, ADMM is applied to obtain the demosaicked image (e). Crosstalk is not added to this example, so there is no need for color correction.
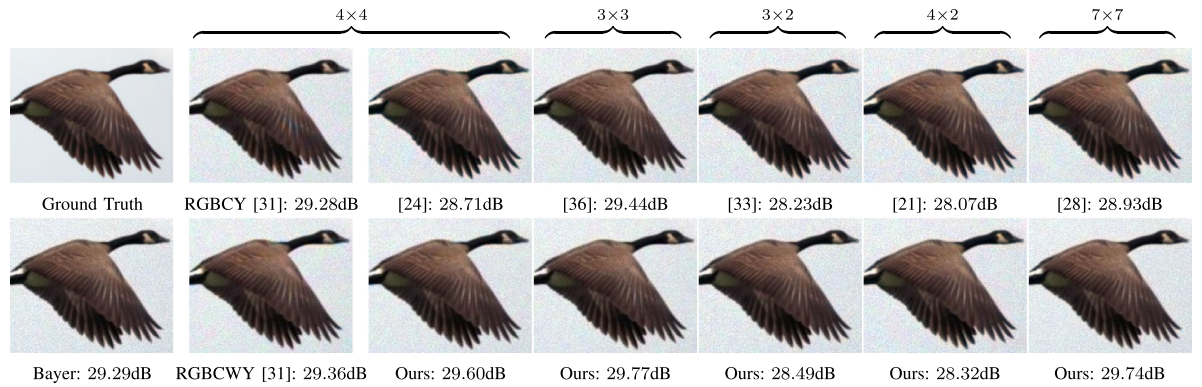


Fig. 13. Reconstructed color images from the QIS measurements on "birds" image. Each subfigure shows the result using a particular CFA design.

Color correction with white balance is performed after color demosaicking for the crosstalk case for removing the crosstalk effect.

We use the following metrics [31] to evaluate the CFAs:

- Sensitivity metamerism index (SMI) which measures the drop in color reproduction accuracy due to crosstalk. It is obtained as a function of the CIEDE2000 color error metric which is obtained by calculating the mean square color difference in the CIELAB color space.
- Luminance SNR (YSNR) which measures the visual noise of luminance channel as defined in ISO 12232 [45].

Table II shows these metrics for different CFAs with and without crosstalk. Our CFAs achieve higher color reproduction accuracy compared to others since they are optimized for crosstalk. This gain in color accuracy happens by trading the noise performance as observed by the drop of YPSNR metric.

### C. Experiment 3: Natural Color Image Reconstruction

In this experiment, we evaluate the performance of different CFAs for natural color image reconstruction. To this end, we use the DIV2K evaluation dataset to generate 100 mosaicked
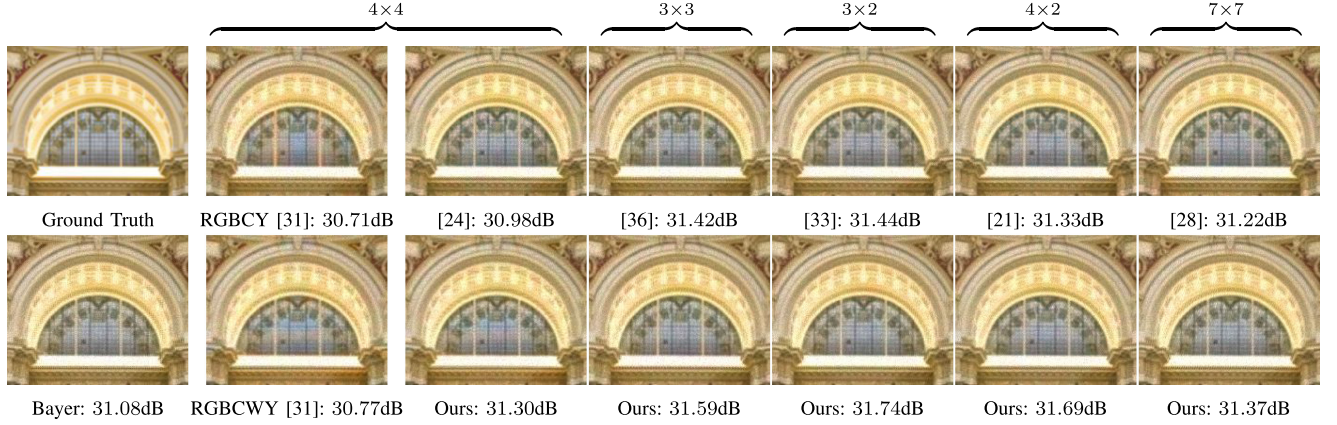
|  | 4×4 | | 3×3 | 3×2 | 4×2 | 7×7 |



| Ground Truth | RGBCY [31]: 30.71dB | [24]: 30.98dB | [36]: 31.42dB | [33]: 31.44dB | [21]: 31.33dB | [28]: 31.22dB |
| Bayer: 31.08dB | RGBCWY [31]: 30.77dB | Ours: 31.30dB | Ours: 31.59dB | Ours: 31.74dB | Ours: 31.69dB | Ours: 31.37dB |

Fig. 14.    Reconstructed color images from the QIS measurements on "building" image. Each subfigure shows the result using a particular CFA design.

images according to QIS model. Two scenarios are assumed: 1) No crosstalk, and 2) Moderate crosstalk with leakage factors $(\delta_r, \delta_g, \delta_b) = (0.23, 0.15, 0.10)$. The low pass filter in Algorithm 2 is chosen as $21 \times 21$ Gaussian with standard deviation $\sigma = 21/3$ and multiplied by a hamming window to mitigate the windowing effect. We simulate the diffraction of light occurring in QIS by blurring the ground truth image with a $5 \times 5$ box kernel

Table II shows the reconstruction quality averaged on 100 images in the DIV2K evaluation dataset. We choose the color PSNR as a quality metric, which is calculated as $\text{CPSNR} = 10 \log_{10}(\frac{1}{\text{MSE}})$ where MSE is the total mean squared error summed over the three color channels. Our CFAs achieve better quality for the crosstalk case. Visually, our CFAs obtain color images with less aliasing artifacts and better details as shown in Fig. 14.

## VII. CONCLUSION

We presented a general and flexible optimization framework to design color filter arrays for QIS. Our framework unifies the CMOS-based color filter array designs and extends to QIS. The color filter arrays designed by our framework are robust to crosstalk between the primary color channels, robust to aliasing between the luminance and chrominance channels, and are robust to noise. We verified the designs both theoretically and numerically through extensive experiments. Our evaluation indicated the effectiveness of the framework in offering trade-offs between different design criteria.

## APPENDIX A

### A. Proof of Proposition 1

Since the luminance channel comprises only one baseband component in the frequency domain, the luminance sensitivity in the amplitude of this component, i.e.,

$$\gamma_l = \frac{1}{K}||\widetilde{\boldsymbol{h}}_l||_2 = \frac{1}{K}\sqrt{\widetilde{h}_l^2(0,0) + 0 + \cdots + 0}$$

$$= \frac{1}{K}\widetilde{h}_l(0,0).$$

Substituting in the DFT equation with $u = v = 0$, we get

$$\gamma_l(\boldsymbol{x}) = \frac{1}{K}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}h_l(m,n)$$

$$= \frac{1}{K}\boldsymbol{1}^T\boldsymbol{h}_l = \frac{1}{K}\boldsymbol{1}^T\boldsymbol{Z}_l\boldsymbol{x} = \boldsymbol{b}^T\boldsymbol{x},$$

where $\boldsymbol{b} \stackrel{\text{def}}{=} \frac{1}{K}\boldsymbol{1}^T\boldsymbol{Z}_l$. As for the chrominance sensitivity $\gamma_c$, by squaring the definition in Equation (17), we get

$$\gamma_c(\boldsymbol{x})^2 = \frac{1}{K^2}\min\left(||\widetilde{\boldsymbol{h}}_\alpha||_2^2, ||\widetilde{\boldsymbol{h}}_\beta||_2^2\right)$$

$$\stackrel{(a)}{=} \min\left(||\boldsymbol{h}_\alpha||_2^2, ||\boldsymbol{h}_\beta||_2^2\right)$$

$$= \min\left(||\boldsymbol{Z}_\alpha x||_2^2, ||\boldsymbol{Z}_\beta x||_2^2\right) = \min\left(\boldsymbol{x}^T\boldsymbol{Q}_\alpha\boldsymbol{x}, \boldsymbol{x}^T\boldsymbol{Q}_\beta\boldsymbol{x}\right), \quad (39)$$

where (a) follows from Parseval theorem, and $\boldsymbol{Q}_\alpha \stackrel{\text{def}}{=} \boldsymbol{Z}_\alpha^T\boldsymbol{Z}_\alpha$ and $\boldsymbol{Q}_\beta \stackrel{\text{def}}{=} \boldsymbol{Z}_\beta^T\boldsymbol{Z}_\beta$ are two positive semidefinite matrices.

## ACKNOWLEDGMENT

## REFERENCES

[1] E. R. Fossum, "The quanta image sensor (QIS): Concepts and challenges," in *Proc. OSA Top. Mtg Comput. Opt. Sens. Imag.*, Jul. 2011, p. JTuE1.
[2] E. R. Fossum, J. Ma, and S. Masoodian, "Quanta image sensor: Concepts and progress," in *Proc. SPIE 9858, Adv. Photon Count. Techn. X*, 2016, vol. 9858, pp. 985804–985804–14.
[3] E. R. Fossum, J. Ma, S. Masoodian, L. Anzagira, and R. Zizza, "The quanta image sensor: Every photon counts," *MDPI Sensors*, vol. 16, no. 8, pp. 1–25, Aug. 2016, Paper 1260.
[4] O. A. Elgendy and S. H. Chan, "Image reconstruction and threshold design for quanta image sensors," in *Proc. IEEE Int. Conf. Image Process. (ICIP'16)*, Sep. 2016, pp. 978–982.
[5] S. H. Chan, O. A. Elgendy, and X. Wang, "Images from bits: Non-iterative image reconstruction for quanta image sensors," *MDPI Sensors*, vol. 16, no. 11, pp. 1–21, Nov. 2016, Paper 1961.
[6] O. A. Elgendy and S. H. Chan, "Optimal threshold design for quanta image sensor," *IEEE Trans. Comput. Imag.*, vol. 4, no. 1, pp. 99–111, Mar. 2018.
[7] F. Yang, Y. M. Lu, L. Sbaiz, and M. Vetterli, "Bits from photons: Oversampled image acquisition using binary poisson statistics," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1421–1436, Apr. 2012.

[8] A. Gnanasambandam, O. A. Elgendy, J. Ma, and S. H. Chan, "Megapixel photon-counting color imaging using quanta image sensor," *Opt. Express*, vol. 27, no. 12, pp. 17298–17310, Jun. 2019.

[9] B. E. Bayer, "Color imaging array," US Patent US3971065A, 1976.

[10] H. S. Malvar and R. Cutler, "High-quality linear interpolation for demosaicing of Bayer-patterned color images," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, May 2004, vol. 3, pp. 485–488.

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

[12] 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.

[13] K. Hirakawa and T. W. Parks, "Adaptive homogeneity-directed demosaicing algorithm," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 360–369, Mar. 2005.

[14] E. Dubois, "Filter design for adaptive frequency-domain bayer demosaicking," in *Proc. 2006 Int. Conf. Image Process.*, Oct. 2006, pp. 2705–2708.

[15] B. Leung, G. Jeon, and E. Dubois, "Least-squares lumachroma demultiplexing algorithm for Bayer demosaicking," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1885–1894, Jul. 2011.

[16] G. Jeon and E. Dubois, "Demosaicking of noisy Bayer-sampled color images with least-squares luma-chroma demultiplexing and noise level estimation," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 146–156, Jan. 2013.

[17] J. T. Korneliussen and K. Hirakawa, "Camera processing with chromatic aberration," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4539–4552, Oct. 2014.

[18] M. Gharbi, G. Chaurasia, S. Paris, and F. Durand, "Deep joint demosaicking and denoising," *ACM Trans. Graph.*, vol. 35, no. 6, pp. 191:1–191:12, Nov. 2016.

[19] H. Tan, X. Zeng, S. Lai, Y. Liu, and M. Zhang, "Joint demosaicing and denoising of noisy Bayer images with ADMM," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2017, pp. 2951–2955.

[20] 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.

[21] 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.

[22] Y. M. Lu and M. Vetterli, "Optimal color filter array design: Quantitative conditions and an efficient search procedure,' presented at the SPIE Conf. Digital Photography V, San Jose, CA, Jan. 20, 2009.

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

[24] 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.

[25] J. Wang, C. Zhang, and P. Hao, "New color filter arrays of high light sensitivity and high demosaicking performance," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2011, pp. 3153–3156.

[26] A. Chakrabarti, W. T. Freeman, and T. Zickler, "Rethinking color cameras," in *Proc. IEEE Int. Conf. Comput. Photography*, May 2014, pp. 1–8.

[27] P. Amba, J. Dias, and D. Alleysson, "Random color filter arrays are better than regular ones," in *Proc. Color Imag. Conf.*, 2016, vol. 2016, no. 1, pp. 294–299.

[28] C. Bai, J. Li, Z. Lin, and J. Yu, "Automatic design of color filter arrays in the frequency domain," *IEEE Trans. Image Process.*, vol. 25, no. 4, pp. 1793–1807, Apr. 2016.

[29] J. Li, C. Bai, Z. Lin, and J. Yu, "Automatic design of high-sensitivity color filter arrays with panchromatic pixels," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 870–883, Feb. 2017.

[30] J. Li, C. Bai, Z. Lin, and J. Yu, "Optimized color filter arrays for sparse representation-based demosaicking," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2381–2393, May 2017.

[31] L. Anzagira and E. R. Fossum, "Color filter array patterns for small-pixel image sensors with substantial cross talk," *J. Opt. Soc. Amer. A*, vol. 32, no. 1, pp. 28–34, Jan. 2015.

[32] C. Chao *et al.*, "Crosstalk metrics and the characterization of 1.1 μm-pixel cis," in *Proc. Int. Image Sensor Workshop*, 2011, pp. 1–4, Paper 7.

[33] 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.

[34] A. Chakrabarti, "Learning sensor multiplexing design through back-propagation," in *Adv. Neural Inf. Process. Syst. 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds., New York: Curran Associates, Inc., 2016, pp. 3081–3089.

[35] B. Henz, E. S. L. Gastal, and M. M. Oliveira, "Deep joint design of color filter arrays and demosaicing," *Comput. Graph. Forum*, vol. 37, no. 2, pp. 389–399, 2018.

[36] H. Biay-Cheng, H. Siddiqui, J. Luo, G. Todor, and A. Kalin, "New color filter patterns and demosaic for sub-micron pixel arrays," in *Proc. Int. Image Sensor Workshop 2015*, Session 4, Paper 1, pp. 1–4.

[37] H. Siddiqui, K. Atanassov, and S. Goma, "Hardware-friendly universal demosaick using non-iterative map reconstruction," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2016, pp. 1794–1798.

[38] E. R. Fossum, "Modeling the performance of single-bit and multi-bit quanta image sensors," *IEEE J. Electron Devices Soc.*, vol. 1, no. 9, pp. 166–174, Sep. 2013.

[39] Y. Hel-Or, "The canonical correlations of color images and their use for demosaicing," Hewlett-Packard Laboratory, Israel, Tech. Rep. HPL-2003-164R1, 2004.

[40] K. Hirakawa, "Cross-talk explained," in *Proc. 15th IEEE Int. Conf. Image Process.*, 2008, pp. 677–680.

[41] Z. Opial, "Weak convergence of the sequence of successive approximations for nonexpansive mappings," *Bull. Amer. Math. Soc.*, vol. 73, no. 4, pp. 591–597, 1967.

[42] M. Stein, "Large sample properties of simulations using latin hypercube sampling," *Technometrics*, vol. 29, no. 2, pp. 143–151, 1987.

[43] T. Yap-Peng and A. Tinku, "Method for color correction with noise consideration," in *Proc. SPIE*, vol. 3963, 1999, pp. 329–337.

[44] E. Agustsson and R. Timofte, "NTIRE 2017 challenge on single image super-resolution: Dataset and study," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit. Workshops*, Jul. 2017, pp. 1122–1131.

[45] *Photography–Digital Still Cameras–Determination of Exposure Index, ISO Speed Ratings, Standard Output Sensitivity, and Recommended Exposure Index*, ISO 12232:2006, 2006.

**Omar A. Elgendy** (S'15–M'19) received the B.Sc. degree in Electronics and Communications Engineering (Distinction with honors) in 2010 and the M.Sc. degree in Engineering Mathematics in 2015, from Faculty of Engineering, Cairo University, and the Ph.D. degree in Electrical and Computer Engineering from Purdue University, West Lafayette, IN, in 2019. He is currently an Image Processing Specialist with GigaJot Tech Inc., Pasadena, CA.
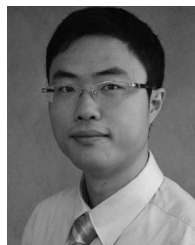
Dr. Elgendy is a recipient of the Best Paper Award of IEEE International Conference on Image Processing 2016. His research interests include statistical image processing with applications to image reconstruction using single-photon imaging sensors, statistical pattern classification, deep learning, and large-scale optimization.



**Stanley H. Chan** (S'06–M'12–SM'17) received the B.Eng. degree (with first class honor) in Electrical Engineering from the University of Hong Kong in 2007, the M.A. degree in Mathematics from the University of California at San Diego in 2009, and the Ph.D. degree in Electrical Engineering from the University of California at San Diego in 2011. From 2012 to 2014, he was a postdoctoral research fellow at Harvard John A. Paulson School of Engineering and Applied Sciences. He is currently an Assistant Professor in the School of Electrical and Computer Engineering and the Department of Statistics at Purdue University, West Lafayette, IN.

Dr. Chan is a recipient of the Best Paper Award of IEEE International Conference on Image Processing 2016, IEEE Signal Processing Cup 2016 Second Prize, Purdue College of Engineering Exceptional Early Career Teaching Award 2019, Purdue College of Engineering Outstanding Graduate Mentor Award 2016, and Eta Kappa Nu (Beta Chapter) Outstanding Teaching Award 2015. He is also a recipient of the Croucher Foundation Fellowship for Postdoctoral Research 2012–2013 and the Croucher Foundation Scholarship for PhD Studies 2008–2010. His research interests include computational imaging and machine learning.

He is an Associate Editor of the IEEE Transactions on Computational Imaging since 2018, an Associate Editor of the OSA Optics Express in 2016–2018, and an Elected Member of the IEEE Signal Processing Society Technical Committee in Computational Imaging since 2015.