

# Adaptive Multispectral Demosaicking Based on Frequency-Domain Analysis of Spectral Correlation

Sunil Prasad Jaiswal, *Student Member, IEEE*, Lu Fang, *Member, IEEE*, Vinit Jakhetya, *Member, IEEE*, Jiahao Pang, *Member, IEEE*, Klaus Mueller, *Senior Member, IEEE*, and Oscar C. Au, *Fellow, IEEE*

**Abstract**—Color filter array (CFA) interpolation, or 3-band demosaicking, is a process of interpolating the missing color samples in each band to reconstruct a full color image. In this paper, we are concerned with the challenging problem of multispectral demosaicking, where each band is significantly undersampled due to the increment in the number of bands. Specifically, we demonstrate a frequency-domain analysis of the subsampled color-difference signal and observe that the conventional assumption of highly correlated spectral bands for estimating undersampled components is not precise. Instead, such a spectral correlation assumption is image dependent and rests on the aliasing interferences among the various color-difference spectra. To address this problem, we propose an adaptive spectral-correlation-based demosaicking (ASCD) algorithm that uses a novel anti-aliasing filter to suppress these interferences, and we then integrate it with an intra-prediction scheme to generate a more accurate prediction for the reconstructed image. Our ASCD is computationally very simple, and exploits the spectral correlation property much more effectively than the existing algorithms. Experimental results conducted on two datasets for multispectral demosaicking and one dataset for CFA demosaicking demonstrate that the proposed ASCD outperforms the state-of-the-art algorithms.

**Index Terms**—Color difference, demosaicking, frequency analysis, multispectral color filter array, spectral correlation.

## I. INTRODUCTION

### A. Background

STANDARD color images are usually represented by three spectral components, namely, red (R), green (G), and blue (B). However, multispectral imaging systems capture more than three spectral bands, which means they retain more information of a scene than standard color imaging systems, with each band representing a narrow wavelength of visible light. Due to the increase in the number of bands, multispectral imaging can reveal much more information, and thus it has

This work is supported in part by Natural Science Foundation of China (NSFC) under contract No. 61303151 and 61331015 and supported in part by the GRF 16211615. This research was also partially supported by NSF (USA) grant IIS 1527200 as well as the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the “IT Consilience Creative Program (ITCCP)” (NIPA-2013-H0203-13-1001), directed by NIPA.

S. P. Jaiswal, L. Fang, V. Jakhetya, O. C. Au are/were with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology (e-mail: spjaiswal@connect.ust.hk; eefang@ust.hk; vjakhetya@connect.ust.hk; eeau90@gmail.com). Dr. Lu Fang is the corresponding author of the paper.

J. Pang is with SenseTime Group Limited, Hong Kong. (e-mail: pangjiahao@sensetime.com).

K. Mueller is with the Department of Computer Science, Stony Brook University-State University of NY (SUNY), USA and SUNY Korea (e-mail: mueller@cs.stonybrook.edu).

become very useful in the fields of satellite imaging [1], computer vision [2], and medical imaging [3].

In recent years, a variety of multispectral imaging systems have been proposed, with different mechanisms to capture images. These systems can be classified into three categories [4]: (i) Single-camera-multi-shot systems [5-6], which capture images at a high spectral resolution but require multiple shots to obtain the images, and thus need special equipment, such as a high-speed lighting system [7], for real-time imaging; (ii) multi-camera-one-shot systems [8], which capture images using several cameras with different color filters in one shot, resulting in a system that is quite complicated as it requires perfect alignment of multiple cameras; and (iii) single-camera-one-shot systems [9-14], such as consumer RGB cameras (for three bands), which overcome the problems of the first two categories of systems in terms of size, cost, and real-time imaging. To reduce the production cost, such RGB cameras use a single image sensor, instead of three sensors, with a color filter array (CFA). The CFA allows only one color to be measured at each pixel location, and thus two missing color elements at each pixel must be estimated from the adjacent pixels. This process is called CFA interpolation or demosaicking. The most famous CFA is the Bayer CFA [15], which is shown in Fig.1(a). Single-camera-one-shot systems, such as

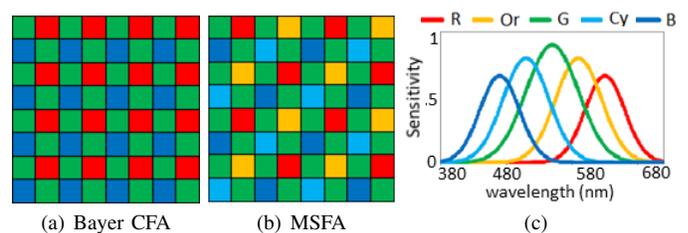


Fig. 1. (a) Bayer CFA (US Patent 3971065), (b) 5-band MSFA [4,12], (c) Schematic spectral sensitivities of different bands [4,12].

RGB cameras, can be extended to multispectral imaging by replacing the CFA with a multispectral filter array (MSFA), where more than three spectral bands are subsampled [9-14]. Fig.1(b) shows a particular 5-band MSFA sensor, and its schematic spectral sensitivities are shown in Fig. 1(c). Several other MSFA patterns are shown in Fig. 2 and it can be seen that due to the increment in the number of bands in the MSFA pattern, each band is significantly undersampled, which makes the multispectral demosaicking a challenging problem.

So far practical realizations of MSFA cameras have been achieved by [10-13]. Geelen *et al.* [10] designed a camera that

captures near infra-red bands, whereas Martinez *et al.*'s [11] and Monno *et al.*'s [12] cameras operate in the visible range. Thomas *et al.* [13] designed a prototype of a camera which captures both visible and near infra-red bands on the same sensor chip. Another architecture for multispectral imaging using CFA is presented in [14], but it requires four sensors to capture the spatial and spectral information in a scene.

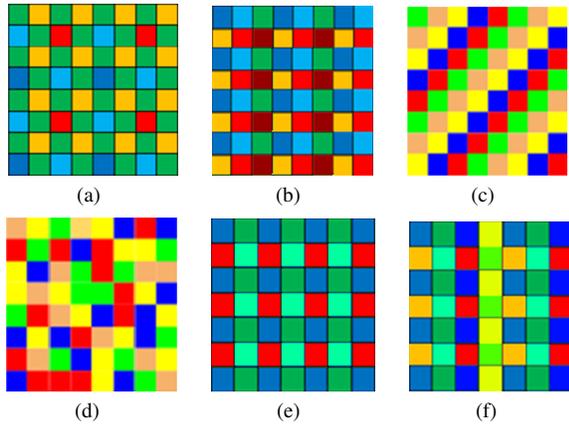


Fig. 2. Several MSFA patterns [9]: (a) BTES [16], (b) Brauers [17], (c) Uniform [18], (d) Random [18], (e) 4-band pattern, (f) 8-band pattern.

## B. Related Works

In this subsection, we provide a brief history of MSFAs and discuss the related works on MSFA demosaicking.

1) *Brief history of MSFA patterns*: A comprehensive review of MSFA sensors was given by Lapray *et al.* in [9]. Miao *et al.* [16] introduced the first systematic way to generate MSFA patterns for multispectral imaging. They proposed a binary tree-based generic scheme, which generates MSFA patterns (Fig. 2(a)) by recursively separating the checkerboard pattern based on a binary tree decomposition defined by the number of spectral bands and the probability of the appearance of each band. Brauers *et al.* [17] introduced a six-band MSFA pattern (Fig. 2(b)), which consists of color filter blocks of dimension  $3 \times 2$ , while Aggarwal *et al.* proposed two MSFAs [18], known as a random MSFA and a uniform MSFA, which can be generalized to any number of bands. Recently, Monno *et al.* [12] proposed a 5-band MSFA pattern based on the density requirement and the derivative requirement. In this MSFA pattern, which is shown in Fig. 1(b), the density of the G-band is higher among all the bands, and it also satisfies the derivative requirement, where each spectral band must be arrayed so that derivatives can be calculated at every pixel location from the raw MSFA data. Fig. 2(e) and Fig. 2(f) shows a particular 4-band and 8-band MSFA pattern respectively.

2) *Related works on MSFA demosaicking*: Several demosaicking algorithms with the popular CFA pattern [15] have been proposed in the literature. But very few algorithms address the challenging problem of multispectral demosaicking. One way to reconstruct the undersampled components is to extend the CFA demosaicking methods to MSFA patterns. Yasuma *et al.* [19] used the classical linear interpolation, Wang *et al.* [20] extended the classical median filtering to

MSFA demosaicking, and Brauers *et al.* [17] extended one of the widely used techniques in CFA demosaicking known as color-difference interpolation (CDI) [15,31-37]. CDI-based techniques assume that the color channels have similar edges or texture structures, and the method in [17] extends this idea to MSFA demosaicking to reconstruct multispectral images by taking into account inter-color correlation. These extension-based MSFA demosaicking algorithms are computationally very simple and easily extended to all the MSFA patterns, but their performance is insufficient and produces artifacts.

Miao *et al.* [21] proposed a generic demosaicking method, known as the binary tree edge sensing (BTES) method, based on the MSFA design in [16]. The BTES method uses the same binary tree that generates the MSFA pattern [16] and interpolates the missing pixels for each band by utilizing the edge correlation information. The algorithm iteratively performs edge-sensing interpolation to generate a full multispectral image. Although the BTES method is a generic method and reconstructs all of the MSFA patterns, its performance suffers in cases of severely undersampled patterns which have *sparsely sampled bands*, such as the pattern in Fig. 1(b).

Monno *et al.* [4,12,22] proposed several algorithms to address the interpolation of the MSFA pattern shown in Fig. 1(b). In [4] and [22] they first generated a guided image from the subsampled (G) band and then used it as a reference image for the interpolation of other subsampled components. In [4], they used an adaptive-kernel-based method for the reconstruction of each component, and then improved it in [22] by a guided filter. Both of these algorithms ([4],[22]) follow the assumption of color-difference smoothness, i.e., the color channels have similar edges and texture structures. Thus these algorithms are image dependent and the reconstructed results have visible artifacts around the edges. To improve the performance of [22], the authors developed multiple guided images in [12] and used them to drive the interpolation of different bands.

Recently, Wang *et al.* [23] proposed a linear-interpolation-based generic MSFA demosaicking scheme inspired by the mathematical analysis of CFA demosaicking in [24]. Basically, they combine the linear minimum mean square error (LMMSE) technique and the residual interpolation method [25] to get an accurate demosaicking. The LMMSE between the original and the reconstructed images is achieved by Wiener estimation and the artifacts are further reduced in the second step by the residual interpolation method [25]. Mihoubi *et al.* [26] proposed an MSFA demosaicking method in which they assume that both the spatial and spectral correlations are correlated and then combine them to estimate the missing pixels. Wang *et al.* [27] extended the DWT-based CFA methods [32] to MSFA demosaicking by estimating the low-frequency and high-frequency components of unknown bands in the wavelet domain. The high-frequency component is estimated using inter-channel correlation, whereas the low-frequency bands are interpolated individually and linearly.

In summary, the existing works on MSFA demosaicking have the following limitations:

- 1) Classical CFA algorithms, like median filtering, linear demosaicking, etc., can be easily extended to any MSFA

pattern and are computationally very simple. However, the results are not impressive as they produce artifacts.

- 2) The performance of the methods in [21] and [23] depend on the pattern and the post-processing step (second step), respectively. Other demosaicking schemes are either dependent on spectral correlation [4,22,26,27] or dependent on spatial correlations [26], and their results are lacking in terms of quality.

### C. Our Contribution

The color-difference interpolation (CDI)-based approaches [15,31-37] assume that the high-frequency components, such as edges and textures of different channels, are very similar or correlated, which makes the color-difference signal a low-pass signal. These algorithms further assume that applying a low-pass filter to a subsampled color-difference signal extracts the color-difference signal, and they take advantage of this property to estimate the missing samples. These assumptions have been widely used in the history of CFA demosaicking because of their low computational consumption. We have studied these assumptions and observe that they are not valid for all images, which reduces the overall quality of the reconstructed image. A similar observation is discussed in [38-40], but it has not been analyzed properly and is limited to CFA demosaicking. The main contributions of our paper are as follows:

- 1) To overcome the problems in existing MSFA and CFA demosaicking, we efficiently demonstrate a frequency-domain analysis of the subsampled color-difference signal and analyze the conventional assumptions of the CDI-based scheme to find the conditions under which they are valid.
- 2) Based on the frequency-domain analysis, we propose an adaptive spectral-correlation-based demosaicking (ASCD) algorithm that modifies the assumption of the CDI technique, and we then integrate it with an intra-prediction scheme to reconstruct the undersampled components. The integration is done on an LMMSE basis to generate a more accurate prediction for demosaicking.
- 3) Our approach can reduce the level of artifacts since it exploits the relationship of inter-color correlation among the different bands more effectively than existing algorithms. Moreover, our algorithm is computationally very simple and can be easily extended to any CFA or MSFA patterns.

Extensive simulation results demonstrate that the proposed algorithm achieves the best performance in most cases, both in the case of CFA demosaicking and MSFA demosaicking, as compared to the state-of-the-art demosaicking methods. A preliminary version of this work has been presented in [41] but is limited to CFA demosaicking.

The remainder of the paper is organized as follows. In Section II, we formally introduce the assumptions of the CDI-based algorithm and then analyze them in the frequency-domain. Section III describes the proposed ASCD scheme, which exploits the spectral correlations effectively. Simulation results are shown in Section IV, and concluding remarks are given in Section V.

## II. FREQUENCY-DOMAIN ANALYSIS OF THE COLOR DIFFERENCE INTERPOLATION (CDI)-BASED TECHNIQUE

### A. Overview of the CDI-based Technique and its Assumptions

The CDI-based technique has been used in 3-band demosaicking methods to exploit the spectral correlation of images [15,31-37]. These methods make a key assumption that the high-frequency components between different spectral bands are not only correlated but are very similar to each other. Based on this assumption, they have shown that color-difference signals (subtracting one band from another) contain mainly low-frequency components, i.e., color-difference signals are low-pass signals. We observe, however, that this assumption is image dependent and may not work for all images.

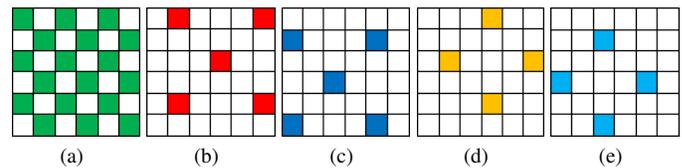


Fig. 3. MSFA pattern [12] : (a)  $G_{MSFA}$  pattern, (b)  $R_{MSFA}$  pattern, (c)  $B_{MSFA}$  pattern, (d)  $Or_{MSFA}$  pattern, (e)  $Cy_{MSFA}$  pattern.

To prove the same, we make an analysis of this assumption in the frequency-domain for MSFA demosaicking. We choose a 5-band MSFA pattern for the analysis, which is shown in Fig. 3. Our analysis can be easily extended to other MSFA and CFA patterns. Fig. 3. shows that, except for the  $G$  band, all the other spectral bands are significantly undersampled, which makes multispectral demosaicking a challenging task. We consider only the reconstruction of the  $G$  band and  $R$  band in the analysis, as reconstruction of other spectral bands may be treated in a similar manner.

A brief overview of the existing CDI-based algorithm is presented as follows. First we interpolate the green component since the  $G$  pixels are densely sampled and thus can preserve high frequencies of the  $G$  channel. We denote the reconstructed green component as  $\hat{G}$  in our discussion. For the reconstruction of the red component, the CDI algorithm estimates the color-difference  $K_R = R_s - \hat{G}_s$  at the available  $R$  pixels, where  $R_s$  and  $\hat{G}_s$  denote the  $R$  and  $\hat{G}$  components sampled at the available  $R$  positions, as shown in Fig. 3(b). The formal definition of  $K_R$  is presented in (5) in the next subsection. Each color component, such as  $R$  (or  $G$ ), can be decomposed into a low-frequency term,  $R_l$  (or  $G_l$ ), and a high-frequency term,  $R_h$  (or  $G_h$ ), and the reason behind the color-difference phenomenon can be interpreted in the two assumptions made by CDI.

- 1) *Correlation Assumption (Assumption-1)*: The first assumption of the CDI technique is about the similarity of inter-color high-frequency components [15, 31-37], i.e.,  $R_h \approx G_h$ , based on which we have

$$R - G = (R_l + R_h) - (G_l + G_h) \approx (R_l - G_l). \quad (1)$$

The difference signal ( $R - G$ ) does not contain high-frequency components, and it is therefore called a low-frequency signal ( $R_l - G_l$ ).

2) *Low-Pass Assumption (Assumption-2)*: Based on this analysis, CDI further assumes that the low-frequency signal ( $R_l - \hat{G}_l$ ) can be estimated from the corresponding subsampled versions, as given below:

$$\zeta\{K_R\} = \zeta\{R_s - \hat{G}_s\} \approx R_l - \hat{G}_l, \quad (2)$$

where  $\zeta\{\cdot\}$  denotes low-pass filtering, i.e., the low-pass filtering of  $R_s$  and  $\hat{G}_s$  gives the low frequencies of  $\mathbf{R}$  and  $\hat{G}$  respectively.

Using both the assumptions in (1) and (2), we can also write  $\zeta\{K_R\} \approx (R_l - \hat{G}_l) \approx (R - \hat{G})$ . Therefore, if an interpolated  $\hat{G}$  is fully available by some method, then the red component interpolated by the CDI method can be recovered using the assumptions in (1) and (2),

$$\hat{R}_{CDI} = \zeta\{K_R\} + \hat{G} \approx (R_l - \hat{G}_l) + \hat{G} \approx R. \quad (3)$$

Here  $\hat{R}_{CDI}$  refers to the red component reconstructed by the CDI method. We first make an analysis of these assumptions in the frequency-domain, and based on this analysis, we propose a demosaicking scheme that works for all sets of images.

### B. Frequency-Domain Analysis of the Subsampled Color Difference Signal

Frequency-domain analysis [42-48] emerged as an outstanding tool to address the issues of CFA demosaicking, and has also been extended to other problems, such as sub-pixel downsampling [49-50]. This tool has been mainly used to study the characteristics of CFA patterns either to recover the full demosaicked image or to design an optimal CFA pattern, as shown by Hirakawa *et. al.* [43]. However, in this paper, we perform frequency-domain analysis of the subsampled color-difference signal ( $K_R$ ), and will particularly analyze the validity of both of the assumptions of the CDI-based technique.

Let  $\mathbf{H}$  represent an original high-resolution input image of dimension  $\mathbf{M} \times \mathbf{N}$ , where  $\mathbf{H} \in \{\mathbf{R}, \mathbf{G}, \mathbf{B}, \mathbf{Or}, \mathbf{Cy}\}$ . We use  $H_{MSFA}$  to represent the incomplete image obtained by applying the corresponding MSFA pattern to the original high-resolution  $\mathbf{H}$  image. For instance,  $R_{MSFA}$  is obtained by applying the corresponding MSFA pattern (Fig. 3(b)) to the original  $\mathbf{R}$  image. A  $4 \times 4$  block of  $R_{MSFA}$  will always have two red components, as shown in Fig. 3(b), and the missing color components in  $R_{MSFA}$  are represented as zero. We can write  $R_{MSFA}$  in terms of  $\mathbf{R}$ , and it is given as

$$R_{MSFA}(i, j) = \begin{cases} R(i, j), & \text{for } (i = 4k - 3, j = 4l - 2) \\ & \text{and } (i = 4k - 1, j = 4l) \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

where  $R(i, j)$  is the red component at the  $(i, j)^{th}$  co-ordinate of the image and  $k = 1, \dots, M/4$ ,  $l = 1, \dots, N/4$  and  $i = 1, \dots, M$ ,  $j = 1, \dots, N$ . For the analysis of the subsampled

color-difference signal, the color-difference  $K_R = R_{MSFA} - \hat{G}$  is estimated at the available  $\mathbf{R}$  pixels and is given by

$$K_R(i, j) = \begin{cases} R_{MSFA}(i, j) - \hat{G}(i, j), & \text{for } (i = 4k - 3, j = 4l - 2) \\ & \text{and } (i = 4k - 1, j = 4l) \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

We can further write (5) as

$$K_R(i, j) = (R(i, j) - \hat{G}(i, j))\Delta_{MSFA}^R(i, j), \quad (6)$$

where  $\Delta_{MSFA}^R(i, j)$  is the modulation signal for the color-difference signal and is given by

$$\Delta_{MSFA}^R(i, j) = \frac{1}{16}(1 + \sin \frac{\pi i}{2})(1 - (-1)^i)(1 - \cos \frac{\pi j}{2})(1 + (-1)^j) + \frac{1}{16}(1 - \sin \frac{\pi i}{2})(1 - (-1)^i)(1 + \cos \frac{\pi j}{2})(1 + (-1)^j). \quad (7)$$

Taking the discrete-time Fourier transform (DTFT) of (7), we get

$$\tilde{\Delta}_{MSFA}^R(u, v) = \frac{1}{8}\mathbf{\Delta}(\mathbf{u})^T \begin{bmatrix} \frac{-a}{4} & 0 & \frac{a}{2} & 0 & \frac{-a}{4} \\ 0 & \frac{3a}{2} & 0 & \frac{3a}{2} & 0 \\ \frac{-a}{2} & 0 & 1 & 0 & \frac{a}{2} \\ 0 & \frac{3a}{2} & 0 & \frac{3a}{2} & 0 \\ \frac{-a}{4} & 0 & \frac{a}{2} & 0 & \frac{-a}{4} \end{bmatrix} \mathbf{\Delta}(\mathbf{v}), \quad (8)$$

where the DTFT is denoted as  $[\cdot]$ ,  $\mathbf{\Delta}(\mathbf{u}) = [\delta(u + \frac{1}{2}), \delta(u + \frac{1}{4}), \delta(u), \delta(u - \frac{1}{4}), \delta(u - \frac{1}{2})]^T$ ,  $\mathbf{\Delta}(\mathbf{v}) = [\delta(v + \frac{1}{2}), \delta(v + \frac{1}{4}), \delta(v), \delta(v - \frac{1}{4}), \delta(v - \frac{1}{2})]^T$ ,  $a = e^{j\pi/2}/8$ , and  $[\cdot]^T$  denotes transposition. Taking the Fourier transform of  $K_R$  in (6) and using the result of (8), we get

$$\tilde{K}_R(u, v) = \frac{1}{8} \cdot \mathbf{1}_3^T \cdot \begin{bmatrix} \frac{3a}{2}\tilde{K}(u + \frac{1}{4}, v + \frac{1}{4}) & 0 & \frac{3a}{2}\tilde{K}(u + \frac{1}{4}, v - \frac{1}{4}) \\ 0 & \tilde{K}(u, v) & 0 \\ \frac{3a}{2}\tilde{K}(u - \frac{1}{4}, v + \frac{1}{4}) & 0 & \frac{3a}{2}\tilde{K}(u - \frac{1}{4}, v - \frac{1}{4}) \end{bmatrix} \mathbf{1}_3, \quad (9)$$

where  $\tilde{K}(u, v) = \tilde{R}(u, v) - \tilde{\hat{G}}(u, v)$  and  $\mathbf{1}_3 = [\mathbf{1} \ \mathbf{1} \ \mathbf{1}]^T$ . We only show the  $3 \times 3$  spectrum in (9) because the rest of the spectrum appears periodically in the  $(u, v)$  space.

We evaluate the magnitude of  $\tilde{K}_R(u, v)$  and show all 13 replicated spectra of  $\tilde{K}$  in Fig. 4. Let  $A = 1/4$  be the horizontal and vertical shift of the spectra such that the main spectrum is at location  $(0, 0)$ , the nearest diagonal spectra are at  $(+A, -A)$ ,  $(-A, +A)$  and the nearest anti-diagonal spectra are at  $(+A, +A)$ ,  $(-A, -A)$ . We can see that the main spectrum at the  $(0, 0)$  location is free from aliasing in the horizontal and vertical directions, but it may be affected by the diagonal and anti-diagonal spectra. If we apply a low-pass

filter ( $\zeta\{\cdot\}$ ) on  $\tilde{K}_R$  in (9), it may not exactly give the low frequencies of  $\tilde{K}$  and thus violates the *Low-Pass Assumption* (*Assumption-2*) of the CDI algorithm. This assumption remains valid only when there is no interference between the diagonal and the center spectra, and only in such a special case does applying a low-pass filter (LPF) on  $\tilde{K}_R$  give exactly the low frequency of  $\tilde{K}$ .

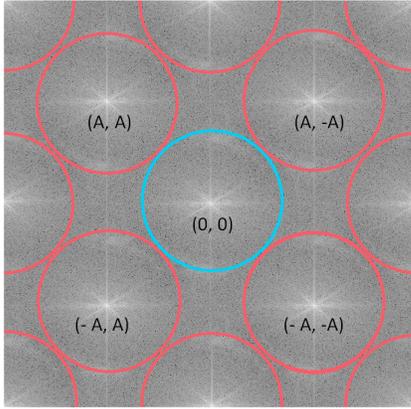


Fig. 4. Frequency spectra of sub-sampled color-difference signal ( $\tilde{K}_R(u, v)$ ).

To get a better insight into the interferences, we plot some possible cases of interference caused by the diagonal and the center spectra in Fig. 5. The blue curve represents the center spectrum ( $\tilde{K}(u, v)$ ) at the  $(0, 0)$  location, and the red curve represents the diagonal spectra at  $(+A, -A)$  in  $\tilde{K}_R$  in (9). Fig. 5(a) depicts the case when the *Low-Pass Assumption* is satisfied, i.e., the center spectrum has no interference with the diagonal spectra, and thus applying an LPF on  $\tilde{K}_R$  will extract  $\tilde{K}$ . Fig. 5(b) suggests the case when the center spectrum has interference with the diagonal spectra, and the red stripes under the red curve within the LPF ( $f \in (0, f_c)$ ) indicate the aliasing interference. In such a case, applying the LPF on  $\tilde{K}_R$  will give a low-pass version of  $\tilde{K}(u, v)$  with aliasing interference and thus violates the *Low-Pass Assumption*.

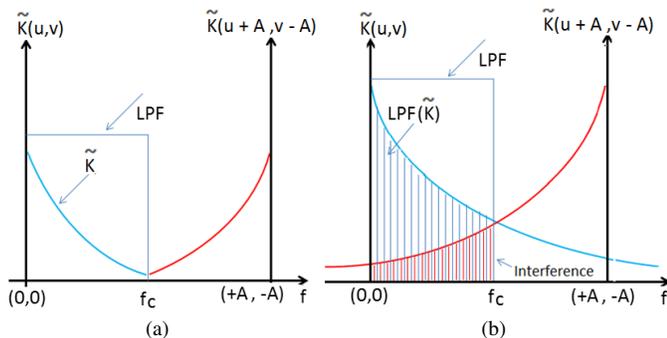


Fig. 5. Spectrum interference analysis: (a) no interference, (b) interference.

In summary, there will be no interference among the spectra only when the spectrum  $\tilde{K}$  is compact, which implies that the color-difference signal ( $K$ ) has to be a low-frequency signal, i.e., the *Correlation Assumption* (*Assumption-1*) has to be satisfied. The *Correlation Assumption* is possible only when the high-frequency components between different bands

TABLE I  
AVERAGE CORRELATION COEFFICIENT ( $\rho$ ) BETWEEN HIGH-FREQUENCY COMPONENTS OF DIFFERENT BANDS

Dataset	$\rho_{R_h G_h}$	$\rho_{B_h G_h}$	$\rho_{O_r G_h}$	$\rho_{C_h G_h}$
Monno (5-band)	0.57	0.56	0.67	0.72
Cave (5-band)	0.82	0.81	0.88	0.75
McMaster (3-band)	0.57	0.55	–	–
Kodak (3-band)	0.79	0.77	–	–

are highly correlated. We observe that such an assumption is image dependent and remains valid only for a limited set of images. For different image datasets, we have calculated the correlation coefficient ( $\rho$ ) between the high-frequency components of  $G_h$  with different bands, and the results are shown in Table I.

For this particular experiment, we apply the Canny edge detection method on the different bands to extract the edges/textures and consider them as a high-frequency component of the bands. From Table I, we can see that the average correlation coefficient is not around 1; some coefficients are higher (e.g., Kodak), while some are lower (e.g., McMaster). A similar observation was drawn for different datasets by Wang *et al.* in [27], where they analyzed the correlation coefficient between the high-frequency sub-bands in the wavelet domain.

Based on this observation and analysis, we propose to modify the *Low-Pass Assumption* in order to estimate the distortion of the CDI method in terms of spectral correlation. We can then show the effect of both of the assumptions on the estimated distortion and then propose an adaptive spectral-correlation-based demosaicking scheme that works for all types of images.

### C. Generalized Low-Pass Assumption of CDI Method

We propose to generalize the *Low-Pass Assumption* of the CDI method in terms of spectral correlation, i.e., applying the LPF on the  $K_R$  image gives the low-pass version of  $K$  with some interference artifacts, as shown in Fig. 5(b):

$$\zeta\{K_R\} = \zeta\{R_s - \hat{G}_s\} = (R_l - \hat{G}_l) + (R - \hat{G})_{AI}, \quad (10)$$

where  $(R - \hat{G})_{AI}$  refers to the diagonal/anti-diagonal spectra that cause interferences with the center spectrum. Assuming the LPF operation  $\zeta$  is linear, we can further write (10) as

$$\zeta\{R_s - \hat{G}_s\} = \zeta\{R_s\} - \zeta\{\hat{G}_s\} = (R_l + R_{AI}) - (\hat{G}_l + \hat{G}_{AI}), \quad (11)$$

where  $R_{AI}$  and  $\hat{G}_{AI}$  are the aliasing interference components occurring when the LPF ( $\zeta$ ) is applied on the individual subsampled components  $R_s$  and  $\hat{G}_s$  respectively. Comparing (10) and (11), we can write  $(R - \hat{G})_{AI} = R_{AI} - \hat{G}_{AI}$ . If there is no interference on the center spectrum in Fig. 5(b), then  $(R - \hat{G})_{AI} = 0$ , which implies either both  $R_{AI}$  and  $\hat{G}_{AI}$  are zero or  $(R_{AI} - \hat{G}_{AI}) = 0$ .

### D. Distortion Analysis of CDI Method in Terms of Spectral Correlation

In this subsection, we estimate the distortion of the CDI method in terms of spectral correlation to study the effect of

both assumptions on it. The interpolated  $\hat{R}_{CDI}$ , by substituting (11) into (3), can be written as

$$\hat{R}_{CDI} = \zeta\{R_s - \hat{G}_s\} + \hat{G} = (R_l + R_{AI}) + (\hat{G}_h - \hat{G}_{AI}), \quad (12)$$

where  $\hat{G}_h$  (or  $\hat{G} - \hat{G}_l$ ) is the high-pass filtered image of the  $\hat{G}$  component. The distortion ( $D_{CDI}$ ) between the original  $\mathbf{R}$  component and interpolated  $\hat{R}_{CDI}$  component can be estimated as

$$\begin{aligned} D_{CDI} &= R - \hat{R}_{CDI} = R - (R_l + R_{AI}) - (\hat{G}_h - \hat{G}_{AI}) \\ &\implies D_{CDI} = (R_h - R_{AI}) - (\hat{G}_h - \hat{G}_{AI}), \end{aligned} \quad (13)$$

where  $R_h = R - R_l$  is the high-pass filtered image of the  $\mathbf{R}$  component. In the case of no interference artifacts, i.e.,  $(R_{AI} - \hat{G}_{AI}) = 0$ , the above equation will reduce to  $D_{CDI} = R_h - \hat{G}_h$ . To simplify the process and without loss of generality, we denote  $R_{h-} = R_h - R_{AI}$ ,  $\hat{G}_{h-} = \hat{G}_h - \hat{G}_{AI}$ , substitute them back into (13), and take the square of both sides to get

$$D_{CDI}^2 = (R_{h-} - \hat{G}_{h-})^2 = R_{h-}^2 + \hat{G}_{h-}^2 - 2 \times R_{h-} \times \hat{G}_{h-}. \quad (14)$$

Now  $\hat{G}_{h-}$  and  $R_{h-}$  contain high-frequency components and  $D_{CDI}$  is the residual image. As a well-known fact, high-frequency components and a residual image can be modeled as a zero mean Laplace distribution [49-51], i.e.,  $E[D_{CDI}] = 0$ ,  $E[R_{h-}] = 0$ , and  $E[\hat{G}_{h-}] = 0$ . In such a case, the variance can be given by  $\sigma_{R_{h-}}^2 = E[R_{h-}^2]$ ,  $\sigma_{\hat{G}_{h-}}^2 = E[\hat{G}_{h-}^2]$  and  $\sigma_{D_{CDI}}^2 = E[D_{CDI}^2]$  respectively. Then the average distortion is calculated by taking the expected value on both sides of (14), i.e.,

$$E[D_{CDI}^2] = E[R_{h-}^2] + E[\hat{G}_{h-}^2] - 2 \times E[R_{h-} \times \hat{G}_{h-}]. \quad (15)$$

The correlation coefficient estimated between  $\hat{G}_{h-}$  and  $R_{h-}$  is given by  $\rho_{R_{h-}\hat{G}_{h-}} = E[R_{h-} \times \hat{G}_{h-}] / (\sigma_{R_{h-}} \times \sigma_{\hat{G}_{h-}})$ , and the above equation becomes

$$\sigma_{D_{CDI}}^2 = \sigma_{R_{h-}}^2 + \sigma_{\hat{G}_{h-}}^2 - 2 \times \sigma_{\hat{G}_{h-}} \times \sigma_{R_{h-}} \times \rho_{R_{h-}\hat{G}_{h-}}. \quad (16)$$

The above equation is the variance of the distortion obtained by the CDI method for the  $\mathbf{R}$  component. In (16),  $\rho_{R_{h-}\hat{G}_{h-}} \in (-1, +1)$ , and it tells the similarity between the high-frequency components of two channels. A higher value of  $\rho_{R_{h-}\hat{G}_{h-}}$  means the edge structures among the components are quite similar, and a smaller  $\rho_{R_{h-}\hat{G}_{h-}}$  means the edges are not similar. If the correlation coefficient among the bands increases, then  $\sigma_{D_{CDI}}^2$  decreases. In the case of no aliasing interference, i.e.,  $(R_{AI} - \hat{G}_{AI}) = 0$ , then  $D_{CDI} = (R_h - \hat{G}_h)$ , and (16) can be reduced to

$$\sigma_{D_{CDI}}^2 = \sigma_{R_h}^2 + \sigma_{\hat{G}_h}^2 - 2 \times \sigma_{\hat{G}_h} \times \sigma_{R_h} \times \rho_{R_h\hat{G}_h}. \quad (17)$$

In Table II, we show the distortion of the CDI method ( $\sigma_{D_{CDI}}^2$ ) in both the case of aliasing and non-aliasing. We can observe that poor spectral correlation leads to higher distortions in both cases. Thus we propose the following to reduce the effects of the two assumptions of the CDI-based scheme:

- 1) To reduce the effect of poor correlation among the bands, i.e., when the *Correlation Assumption* fails, we

TABLE II  
EFFECT OF CORRELATION COEFFICIENT ON CDI METHOD

	Aliasing	No Aliasing
$\rho_{R_{h-}\hat{G}_{h-}} = -1$	$(\sigma_{R_{h-}} + \sigma_{\hat{G}_{h-}})^2$	$(\sigma_{R_h} + \sigma_{\hat{G}_h})^2$
$\rho_{R_{h-}\hat{G}_{h-}} = 0$	$\sigma_{R_{h-}}^2 + \sigma_{\hat{G}_{h-}}^2$	$\sigma_{R_h}^2 + \sigma_{\hat{G}_h}^2$
$\rho_{R_{h-}\hat{G}_{h-}} = 1$	$(\sigma_{R_{h-}} - \sigma_{\hat{G}_{h-}})^2$	$(\sigma_{R_h} - \sigma_{\hat{G}_h})^2$

propose to interpolate the undersampled components independently, without using the other bands (intra-prediction).

- 2) To reduce the effect of aliasing interference among the spectra, i.e., when the *Low-Pass Assumption* fails, we propose to use a differently shaped LPF that can reduce the effect of aliasing and preserve more frequencies than the traditional rectangular LPF.

We then integrate the intra-prediction scheme with the modified CDI on an optimal basis to generate a more accurate prediction for reconstruction of undersampled components.

### III. PROPOSED ADAPTIVE SPECTRAL-CORRELATION-BASED DEMOSAICKING (ASCD) ALGORITHM

In this section, we present our proposed ASCD algorithm and choose the 5-band MSFA pattern, which is shown in Fig. 3, for the explanation. Under a similar rationale, our proposed method can be generalized to other MSFA patterns. Since the density of the G pixels is much higher than that of other bands, we first address the reconstruction of the green ( $\mathbf{G}$ ) component and then use it to drive the interpolation of the other undersampled bands.

#### A. Interpolation of G Component

We use  $G_{MSFA}$  to represent the incomplete image obtained by applying the corresponding MSFA pattern (Fig. 3(a)) to the original high-resolution  $\mathbf{G}$  image, where a  $2 \times 2$  block of  $G_{MSFA}$  will always have two green components, and the missing color components in  $G_{MSFA}$  are represented as zero. We can write  $G_{MSFA}$  in terms of  $\mathbf{G}$ , and it is given as

$$G_{MSFA}(i, j) = G(i, j) \Delta_{MSFA}^G(i, j), \quad (18)$$

where  $\Delta_{MSFA}^G(i, j)$  is the modulation function, and it is given as

$$\begin{aligned} \Delta_{MSFA}^G(i, j) &= \frac{1}{4}(1 - (-1)^i)(1 - (-1)^j) \\ &\quad + \frac{1}{4}(1 + (-1)^i)(1 + (-1)^j). \end{aligned} \quad (19)$$

Using a similar derivation to (8), the Fourier transform of  $G_{MSFA}$  in (18) can be written as

$$\tilde{G}_{MSFA}(u, v) = \frac{1}{2}.$$

$$\mathbf{1}_3^T \cdot \begin{bmatrix} \tilde{G}(u + \frac{1}{2}, v + \frac{1}{2}) & 0 & \tilde{G}(u + \frac{1}{2}, v - \frac{1}{2}) \\ 0 & \tilde{G}(u, v) & 0 \\ \tilde{G}(u - \frac{1}{2}, v + \frac{1}{2}) & 0 & \tilde{G}(u - \frac{1}{2}, v - \frac{1}{2}) \end{bmatrix} \cdot \mathbf{1}_3. \quad (20)$$

We evaluate the magnitude of  $\tilde{G}_{MSFA}(u, v)$ , and it is shown in Fig. 6(a). We can see that there are five replicated spectra of  $\tilde{G}$ . Let  $A = \frac{1}{2}$  be the horizontal and vertical shift of the replicated spectra in (20) such that the center spectrum is at location (0,0). The horizontal and vertical shift in  $\tilde{G}_{MSFA}$  ( $A = \frac{1}{2}$ ) is two times more than in  $\tilde{K}_R(u, v)$  ( $A = \frac{1}{4}$ ) in (9), and thus we can say that the frequency localizations of  $\tilde{G}$  are highly separated from each other. We can see that a rectangular filter can extract the center spectrum, but we observe that a circular filter can reduce the aliasing effect much better than the rectangular filter and thus leads to better results.

In Fig. 6(a) and Fig. 6(b), we show both a rectangular and circular LPF respectively. In both cases, there is no aliasing in the horizontal and vertical frequency, but there can be aliasing among the diagonal spectra. It can be theoretically shown that a circular filter can extract a higher passband area of the center spectrum than the rectangular filter. The other filter that can be used to retain more frequency of the center spectrum is a band-pass filter, as discussed in [48]. In the band-pass filter, we need to manually select  $f$  based on the image content. For simplicity, we choose a circular LPF in the simulation. In practical implementation, one can choose a butter-worth filter of order  $n$  for the implementation of different LPFs. If the value of  $n$  in the butter-worth filter is kept 1, the LPF shows a smooth transition at the cut-off frequency and limits the artifacts. When the value of  $n$  increases the transition gets sharper. One can also consider the Gaussian filter with different  $\sigma$  values for the implementation of different LPFs.

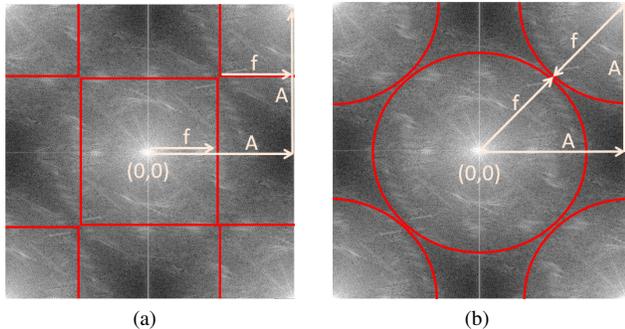


Fig. 6. Filter study: (a) rectangular filter, (b) circular filter.

The larger the cut-off frequency, the higher the passband area of the center spectrum that will be retained, and thus a better reconstruction of  $\tilde{G}(u, v)$  can be achieved. Since the magnitudes of the spectra in (20) are equal, we choose the cut-off frequency for the circular filter as half of the distance between the center spectrum and the diagonal spectrum, i.e., the cut-off frequency is given as  $\sqrt{A^2 + A^2}/2 = A/\sqrt{2}$ .

Choosing a circular filter or band-pass filter may not totally eliminate the aliasing effect on the center spectrum, but it can reduce the effect of aliasing much better than the traditional rectangular LPF and thus leads to better results. We use the circular filter in (9) as well to extract the color-difference spectrum ( $\tilde{K}$ ) to reduce the effects of aliasing interferences, and thus address the issue of the *Low-Pass Assumption* of the CDI method. In a similar way, we compute the cut-off frequency in (9) to extract  $\tilde{K}$  as it is done in (20).

## B. Interpolation of Other Components

Based on the frequency-domain analysis of the assumptions of the CDI method in the previous section, we observe that the method is highly reliant on a strong correlation coefficient, and in the case of poor correlation, it fails to work. Therefore, we propose an adaptive spectral-correlation-based demosaicking, which incorporates an intra-prediction scheme (to be analyzed in Section III-B (1)) for the poorly-correlated case and the CDI method for the highly-correlated case, so as to make a better reconstruction.

1) *Analysis of intra-prediction based interpolation method:* The simplest way to do multispectral demosaicking is to apply a LPF to each undersampled band independently, without using the other bands. We denote the interpolated red component by an independent LPF as  $\hat{R}_{LPF}$ , and it is given as

$$\hat{R}_{LPF} = \zeta\{R_s\} = R_l + R_{AI}. \quad (21)$$

We use a circular LPF for interpolation in (21) to reduce the effect of aliasing and keep the same cut-off frequency as used in (10), and (11) for the CDI method. Thus  $R_{AI}$  in (21) (aliasing interference component) also remains the same as it is in (11). The distortion between the original  $\mathbf{R}$  component and the interpolated  $\hat{R}_{LPF}$  component can be estimated as

$$D_{LPF} = R - \hat{R}_{LPF} = R - R_l - R_{AI} = R_h - R_{AI} = R_{h-}, \quad (22)$$

where  $R_h = R - R_l$  is the high-pass filter output of the  $\mathbf{R}$  component. Assuming  $D_{LPF}$  follows the Laplace distribution [49-51] with zero mean, i.e.,  $E[D_{LPF}] = 0$ , then the variance of the final distortion ( $\sigma_{D_{LPF}}^2$ ) is equal to the variance of  $R_{h-}$  ( $\sigma_{R_{h-}}^2$ ), i.e.,

$$\sigma_{D_{LPF}}^2 = E[(R - \hat{R}_{LPF})^2] = E[R_{h-}^2] = \sigma_{R_{h-}}^2. \quad (23)$$

Here  $\sigma_{D_{LPF}}^2$  is the variance of the distortion obtained when an LPF is applied to an undersampled component independently, without using the other bands.

2) *Analysis of the proposed demosaicking scheme:* The proposed ASCD algorithm makes judicious use of both the CDI method and the independent LPF method by taking into account the spectral correlation. We first compare the performance of both methods by analyzing the spectral correlations, and based on this analysis, we then propose to optimally combine both schemes to generate a more accurate prediction for reconstruction.

By comparing  $\sigma_{D_{CDI}}^2$  in (16) and  $\sigma_{D_{LPF}}^2$  in (23), we can write the distortion of the CDI method in terms of the distortion of the independent LPF method, and it is given as

$$\sigma_{D_{CDI}}^2 = \sigma_{D_{LPF}}^2 + \sigma_{\hat{G}_{h-}} \times (\sigma_{\hat{G}_{h-}} - 2 \times \sigma_{R_{h-}} \times \rho_{R_{h-}\hat{G}_{h-}}). \quad (24)$$

We know that  $\sigma_{\hat{G}_{h-}} > 0$ , and if  $\sigma_{D_{CDI}}^2 < \sigma_{D_{LPF}}^2$ , it implies that the CDI method has lower distortion than the independent LPF method. In such a case, from (24), we can write

$$\begin{aligned} \sigma_{\hat{G}_{h-}} - 2 \times \sigma_{R_{h-}} \times \rho_{R_{h-}\hat{G}_{h-}} &< 0, \\ \implies \rho_{R_{h-}\hat{G}_{h-}} &> 0.5 \times \sigma_{\hat{G}_{h-}} / \sigma_{R_{h-}}. \end{aligned} \quad (25)$$

If the above condition is satisfied, then we can say that the CDI method can achieve a better prediction than the independent

LPF method by exploiting the correlation property effectively; otherwise we should use the independent LPF method for reconstruction to increase the prediction accuracy of the demosaicking. By rearranging the terms in (25), we have generalized the condition for both cases, and it is shown in (26):

$$\rho_{R_{h-}\hat{G}_{h-}} \times \sigma_{R_{h-}} / \sigma_{\hat{G}_{h-}} = \begin{cases} > 0.5 & \text{if } \sigma_{D_{CDI}}^2 \leq \sigma_{D_{LPF}}^2 \\ < 0.5 & \text{if } \sigma_{D_{CDI}}^2 > \sigma_{D_{LPF}}^2. \end{cases} \quad (26)$$

The condition in (26) plays a crucial role in determining which method should be used for reconstruction. In view of this observation, we propose a linear combination of the LPF method and CDI method on a block-by-block basis to generate a more accurate prediction for demosaicking, and optimal weights are estimated in a LMMSE sense for each block. We denote the proposed interpolated red component by  $\hat{R}_{ASCD}$ , and it is expressed as

$$\hat{r}_{ASCD} = w_1 \hat{r}_{LPF} + w_2 \hat{r}_{CDI}. \quad (27)$$

Here  $\hat{r}_{ASCD} \in \hat{R}_{ASCD}$  is the interpolated block of size  $m \times n$ , and  $\hat{r}_{LPF} \in \hat{R}_{LPF}$  and  $\hat{r}_{CDI} \in \hat{R}_{CDI}$  are reconstructed blocks of the same size obtained by the LPF method and CDI method respectively.  $w_1$  and  $w_2$  are the weighted coefficients for the combinations. Thus the distortion ( $d_{ASCD}$ ) between the original block ( $r \in \mathbf{R}$ ) and the interpolated block ( $\hat{r}_{ASCD}$ ) can be written as

$$\begin{aligned} d_{ASCD} &= r - \hat{r}_{ASCD} = r - (w_1 \hat{r}_{LPF} + w_2 \hat{r}_{CDI}) \\ &\implies d_{ASCD} = w_1 d_{LPF} + w_2 d_{CDI}. \end{aligned} \quad (28)$$

Here  $d_{LPF} \in D_{LPF}$  and  $d_{CDI} \in D_{CDI}$  are the distorted blocks obtained by the LPF and CDI methods respectively. To obtain the optimal weights ( $w_1, w_2$ ) in the LMMSE sense, the problem can be formulated as follows:

$$\begin{aligned} \min_{w_1, w_2} & E[d_{ASCD}^2] \\ \text{s.t.} & \sum_{i=1,2} w_i = 1. \end{aligned} \quad (29)$$

To minimize the residual energy ( $E[d_{ASCD}^2]$ ) of each block, we differentiate it with respect to  $w_1$  and  $w_2$ , and the optimal weights can be expressed as

$$\begin{cases} w_1 = E[d_{CDI}(d_{CDI} - d_{LPF})] / E[(d_{CDI} - d_{LPF})^2], \\ w_2 = E[d_{LPF}(d_{LPF} - d_{CDI})] / E[(d_{CDI} - d_{LPF})^2]. \end{cases} \quad (30)$$

By substituting (30) into (27), we can efficiently interpolate the missing samples of the block for the red component. Unfortunately, to estimate  $d_{LPF}$  and  $d_{CDI}$  in (30), we need the original block ( $r$ ), which is not available in practice. In our experiments, we propose to use the resultant image of an existing method as an initialization to our algorithm. We choose the method in [12] as it is computationally very simple and has already been implemented in the hardware set up in [12]. However, other existing algorithms could also be chosen for the initialization, and this is discussed (Table V) in the next section.

We find that the optimal weights ( $w_1, w_2$ ) satisfy the condition in (26), and this validates the efficiency of the proposed ASCD method. A brief description is as follows: We know  $d_{LPF} = R_{h-}$  from (22), and substituting it into (14) results in  $d_{CDI} = d_{LPF} - \hat{G}_{h-}$ , which implies  $d_{LPF} - d_{CDI} = \hat{G}_{h-}$ . Substituting both of them into (30) and using the definition of the correlation coefficient ( $\rho_{R_{h-}\hat{G}_{h-}} = E[R_{h-} \times \hat{G}_{h-}] / (\sigma_{R_{h-}} \times \sigma_{\hat{G}_{h-}})$ ) and variance ( $\sigma_{\hat{G}_{h-}}^2 = E[\hat{G}_{h-}^2]$ ), we can write  $w_2$  in (30) as

$$w_2 = E[R_{h-}\hat{G}_{h-}] / E[\hat{G}_{h-}^2] = \rho_{R_{h-}\hat{G}_{h-}} \times \sigma_{R_{h-}} / \sigma_{\hat{G}_{h-}}. \quad (31)$$

Now suppose the correlation coefficient among  $R_{h-}$  and  $G_{h-}$  is very high. This implies that the *Correlation Assumption* of the CDI method is satisfied, and thus we can say that the CDI method can do a better reconstruction than an intra-prediction method, i.e., we can write  $\sigma_{D_{CDI}}^2 < \sigma_{D_{LPF}}^2$ . In this case the CDI method should be given more weight than the independent LPF method in (27) for effective utilization of the spectral correlation property. So by comparing (26) and (31) in the case of  $\sigma_{D_{CDI}}^2 < \sigma_{D_{LPF}}^2$ , we get  $w_2 > 0.5$  and  $w_1 = 1 - w_2 < 0.5$ , which validates that the CDI method is given more weight in (27). A similar analysis can be drawn when the correlation among the bands is poor, i.e., when the *Correlation Assumption* fails. From the above analysis, we can say that the proposed algorithm is adaptive to spectral correlation and assigns the optimal weights judiciously to both of the methods.

### C. Distortion Comparison of LPF, CDI, and Proposed ASCD

The proposed ASCD algorithm has the potential to work well, even in the case of poor spectral correlation. To prove this, we express the distortion of the proposed ASCD in terms of the spectral correlation. Substituting  $d_{LPF} = R_{h-}$ ,  $d_{CDI} = R_{h-} - \hat{G}_{h-}$ , and  $w_1 + w_2 = 1$  into (28), we can write

$$d_{ASCD} = R_{h-} - w_2 \hat{G}_{h-}, d_{ASCD}^2 = (R_{h-} - w_2 \hat{G}_{h-})^2. \quad (32)$$

Taking the expected value on both sides, we get

$$E[d_{ASCD}^2] = E(R_{h-}^2) + w_2^2 E(\hat{G}_{h-}^2) - 2w_2 E(R_{h-}\hat{G}_{h-}). \quad (33)$$

Using the result of (31), i.e.,  $E(R_{h-}\hat{G}_{h-}) = w_2 \times E[\hat{G}_{h-}^2]$ , and incorporating it into the above equation, we get

$$\begin{aligned} E[d_{ASCD}^2] &= E(R_{h-}^2) + w_2^2 E(\hat{G}_{h-}^2) - 2w_2^2 \times E[\hat{G}_{h-}^2] \\ &\implies E[d_{ASCD}^2] = E(R_{h-}^2) - w_2^2 E(\hat{G}_{h-}^2). \end{aligned} \quad (34)$$

Since we know  $E[R_{h-}^2] = \sigma_{R_{h-}}^2$  and  $E[\hat{G}_{h-}^2] = \sigma_{\hat{G}_{h-}}^2$ , and substituting  $w_2$  from (31) into the above equation, we get

$$\begin{aligned} \sigma_{D_{ASCD}}^2 &= \sigma_{R_{h-}}^2 - (\rho_{R_{h-}\hat{G}_{h-}} \times \sigma_{R_{h-}} / \sigma_{\hat{G}_{h-}})^2 \sigma_{\hat{G}_{h-}}^2 \\ &\implies \sigma_{D_{ASCD}}^2 = \sigma_{R_{h-}}^2 (1 - (\rho_{R_{h-}\hat{G}_{h-}})^2). \end{aligned} \quad (35)$$

We plot a curve of  $\sigma_{D_{LPF}}^2$  given in (23),  $\sigma_{D_{CDI}}^2$  shown in (16), and  $\sigma_{D_{ASCD}}^2$  as a function of the correlation coefficient, and it is shown in Fig. 7. In the curve, the distortion of the independent LPF method remains constant as it is not dependent on other bands during reconstruction, while the distortion of

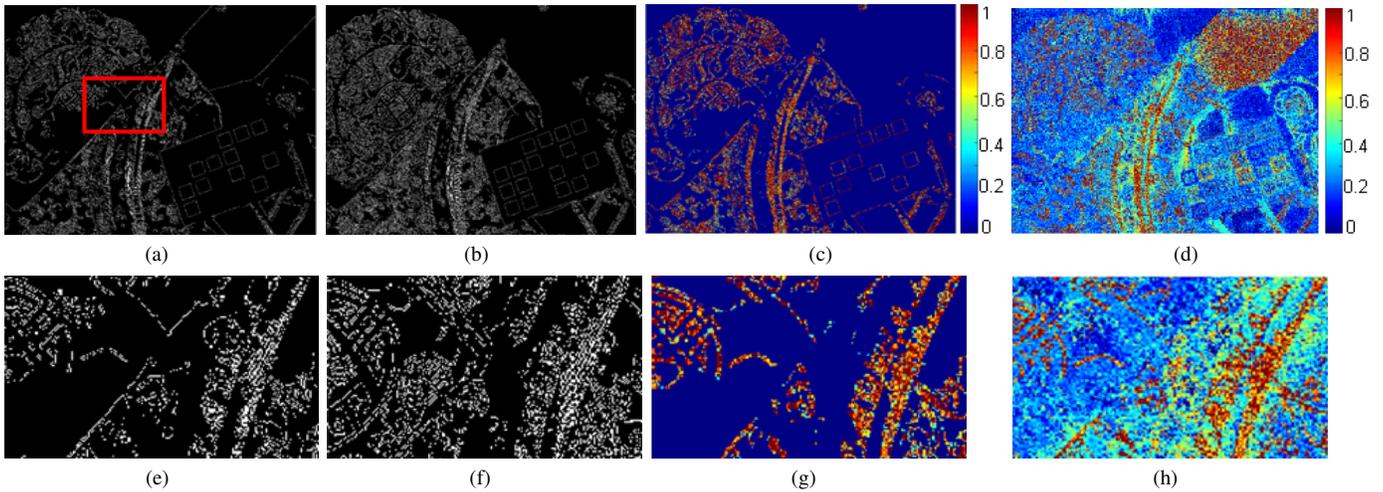


Fig. 8. Visual analysis of weight ( $w_2$ ) map for the CHINADDRESS image in the Monno Dataset: (a) high-frequency component of  $\mathbf{R}$  band, (b) high-frequency component of  $\mathbf{G}$  band, (c) correlation coefficient map of (a) and (b), (d) optimal weight ( $w_2$ ) map. Bottom row is a zoomed-in version of the red box in (a).

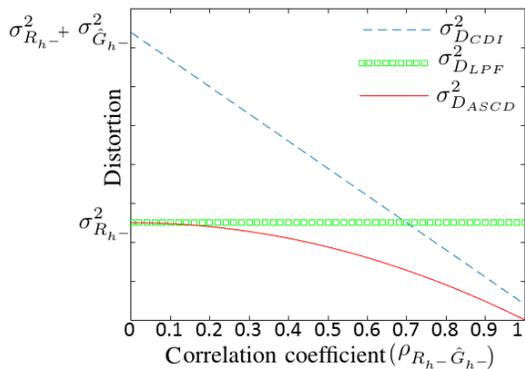


Fig. 7. Distortion analysis of methods in terms of correlation coefficient.

the CDI method is much higher in the case of poor correlation and then decreases with an increase of the correlation among the bands. We can see that only after a certain point does the CDI method work better than the independent LPF method. This implies that the CDI method can work only when the high frequencies of different spectral bands are highly correlated, whereas the distortion of the proposed method is a downward parabola, which is maximum at zero correlation and works well at all correlation coefficients.

#### D. Algorithm Steps

In summary, the main steps of the algorithm are as follows:

- 1) The green component ( $\hat{G}$ ) is interpolated by using a circular filter to get a precise reconstruction.
- 2) To interpolate the missing samples of a block in the red component, its corresponding  $\hat{r}_{CDI}$  and  $\hat{r}_{LPF}$  are computed. To compute  $\hat{r}_{CDI}$ , we use the circular filter, as shown in Fig. 4, to suppress the aliasing interference artifacts in (9). The same filter is used with the same cut-off frequency for computing  $\hat{r}_{LPF}$  in (21) by the independent LPF method (intra-prediction).
- 3) Then the optimal weighted parameters ( $w_1, w_2$ ) are calculated using (30), and  $\hat{r}_{ASCD}$  is estimated from (27).

- 4) Step 2 and step 3 are repeated to reconstruct the whole image.

In a similar manner, other undersampled components can likewise be reconstructed efficiently.

#### IV. SIMULATION RESULTS AND DISCUSSION

We implement the proposed algorithm, test its performance on different datasets, and compare the results with different state-of-the-art algorithms. For the evaluation of the 5-band demosaicking methods, we use two standard multispectral image datasets, the Cave dataset [19] and Monno dataset [12]. The Cave dataset consists of 31-band multispectral images of 32 scenes, and the image size is  $512 \times 512$ . These 31-band images are acquired at every 10 nm from 400 nm to 700 nm, and the ground-truth 5-band images are simulated from these 31-band images. The other dataset is the Monno dataset [12], which consists of 5-band multispectral images of 12 scenes, and the image size is  $1368 \times 1800$ . For CFA demosaicking, we test the performance of the proposed algorithm on the IMAX dataset [39] since it is commonly used for evaluating performance, and it contains 18 test images. All three datasets consist of images having highly textured regions as well as smooth regions, and thus the efficiency of the demosaicking algorithms can be tested thoroughly.

##### A. 5-band Demosaicking

For the performance comparison, we first down-sample a 5-band ground truth image according to the MSFA pattern given in Fig. 1(b) and then conduct the demosaicking process by various methods. We also show the performance of different methods on the MSFA pattern shown in Fig. 2(a). The number of green and cyan samples are identical in both the MSFA patterns, whereas the number of samples is changed for other colors. We follow a similar strategy for the reconstruction of undersampled components for the MSFA pattern in Fig. 2(a). We use CIE D65 (daylight) [52] illumination for the evaluation. Our proposed algorithm is compared with state-of-the-art demosaicking methods: practical one-shot multispectral

TABLE III  
AVERAGE PSNR (IN DB) RESULTS OF THE MONNO AND CAVE DATASETS (HIGHEST PSNR IS HIGHLIGHTED IN BOLDFACE).

MSFA Pattern	Monno Dataset							Cave Dataset					
	Algorithm	R	G	B	Or	C	Mean	R	G	B	Or	C	Mean
MSFA [12]	BTES [21]	45.73	47.90	43.64	44.72	41.39	44.67	42.60	46.54	40.46	39.41	37.84	41.37
	LI [23]	46.91	48.45	44.96	45.82	42.90	45.80	43.79	47.05	41.05	40.65	39.12	42.33
	IID [26]	52.40	48.48	47.10	49.61	45.95	48.72	44.10	46.31	43.34	43.12	42.52	43.87
	GF [22]	52.70	49.02	47.23	50.83	46.95	49.34	44.61	47.65	43.31	42.13	41.25	43.79
	PROPOSED	<b>54.64</b>	<b>51.87</b>	47.88	<b>51.72</b>	<b>47.80</b>	<b>50.78</b>	<b>45.81</b>	47.85	<b>44.94</b>	<b>45.20</b>	44.60	<b>45.68</b>
MSFA [16]	BTES [21]	43.35	47.89	40.04	47.92	41.35	44.10	39.26	46.54	37.30	42.11	37.84	40.61
	LI [23]	45.52	48.45	41.19	49.21	42.91	45.45	40.22	47.04	38.19	43.32	39.11	41.57
	IID [26]	50.30	48.48	44.56	51.32	46.96	48.32	41.95	46.32	39.42	44.18	42.50	42.87
	GF [22]	50.18	49.05	38.83	53.43	46.90	47.67	41.67	47.65	38.83	44.43	41.27	42.77
	PROPOSED	<b>52.95</b>	<b>51.87</b>	46.45	<b>54.16</b>	<b>47.84</b>	<b>50.65</b>	<b>44.05</b>	47.85	<b>42.79</b>	<b>45.86</b>	44.60	<b>45.03</b>

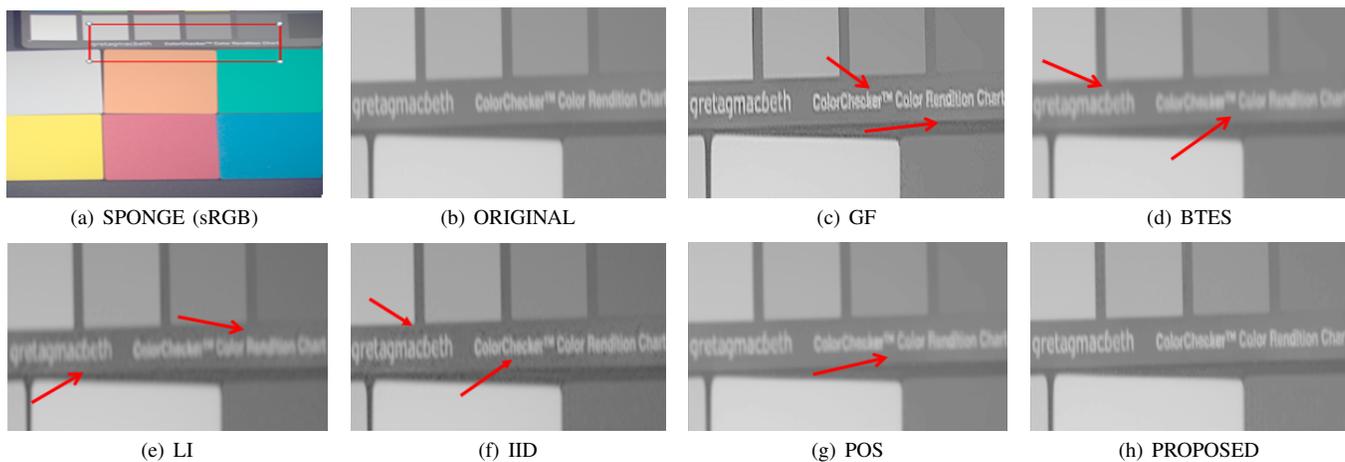


Fig. 9. Visual comparison of Or band of SPONGE image in Cave dataset. (Gamma correction is applied for the display.)

demaicking (POS) [12], BTES [21], guided-filter (GF) [22], linear interpolation (LI) [23], and iterative intensity difference (IID) [26]. We show both the peak signal to noise ratio (PSNR) and structural similarity (SSIM) [53] for quantitatively evaluating the objective performance of the demosaicking process. For fair comparison, we also investigate the subjective quality of the reconstructed image obtained from these methods.

The proposed ASCD algorithm adaptively assigns higher weights to the CDI scheme in the case of higher spectral correlation, whereas in the case of lower spectral correlation it assigns higher weights to the intra-prediction scheme in (27) to achieve an accurate prediction. In Fig. 8, we present a visual analysis of the CHINADDRESS image (Fig. 10) to support the rationale of our algorithm. Fig. 8(a) and 8(b) are the high-frequency components of the **R** and **G** band respectively, where we can observe that most of the edges/textures are not similar to each other and have different orientations (see (e) and (f) for better visual analysis). We also compute the spectral correlation or correlation coefficient of these two components, and this is shown in Fig. 8(c), where the red dots denote the correlated edges and the blue dots denote the uncorrelated edges. Hence we can see that the spectral bands are not highly correlated, and the *Correlation Assumption* of the CDI method fails in such cases. Thus any algorithm that is based on such an assumption reduces the quality of the reconstructed image. In

Fig. 8(d), we show the proposed optimal weight ( $w_2$  in (30)) map, which assigns smaller weights to the CDI scheme in the case of uncorrelated edges, whereas in the case of correlated edges, it assigns higher weights. Thus we can say that the proposed ASCD is adaptive to the spectral correlation.

We can see, however, some inconsistencies between Fig. 8(c) and (d) at the top-right region, where high weights are assigned for low correlation. The reason behind this is as follows: In Fig. 8(b), we can see that there is no high-frequency (HF) component at the top-right region of the **G** band, which implies that the standard deviation of the HF component of the **G** band ( $\sigma_{\hat{G}_{h-}}$ ) in that region is a small number ( $\approx 0$ ). If we substitute a small value of  $\sigma_{\hat{G}_{h-}}$  into (31), we get a higher  $w_2$  in that region. This is why there are inconsistencies in that region; however, they will not affect the performance of the proposed algorithm. We can make this conclusion because, intuitively, in the top-right region, for the reconstruction of the **R** band, there is no use of **G** band since the HF component of the **G** band is zero. Thus, we only need an intra-prediction scheme for the reconstruction in that region. This intuition holds true in the case of proposed ASCD method and it can be explained as follows: since  $w_2 \approx 1$  and  $w_1 \approx 0$  in that region, substituting these weights into (27) results  $\hat{r}_{ASCD} \approx \hat{r}_{CDI} \approx \hat{r}_{LPF} + \hat{G}_{h-} \approx \hat{r}_{LPF}$ . Thus, we can see that our algorithm satisfies the intuition that an

TABLE IV  
AVERAGE SSIM RESULTS ON THE MONNO AND CAVE DATASET (HIGHEST SSIM IS HIGHLIGHTED IN BOLDFACE).

MSFA	Monno Dataset							Cave Dataset					
	Algo.	R	G	B	Or	C	Mean	R	G	B	Or	C	Mean
MSFA [12]	BTES [21]	0.9598	0.9934	0.9574	0.9781	0.9557	0.9689	0.9724	0.9801	0.9710	0.9610	0.9524	0.9674
	LI [23]	0.9610	0.9937	0.9611	0.9801	0.9641	0.9720	0.9780	0.9889	0.9791	0.9671	0.9612	0.9749
	IID [26]	0.9891	0.9937	0.9797	0.9859	0.9780	0.9843	0.9795	0.9874	0.9802	0.9701	0.9807	0.9796
	GF [22]	0.9899	0.9943	0.9804	0.9939	0.9805	0.9878	0.9805	0.9910	0.9801	0.9770	0.9790	0.9815
	POS [12]	0.9889	0.9939	<b>0.9834</b>	0.9942	0.9801	0.9881	0.9831	<b>0.9922</b>	0.9822	0.9840	<b>0.9825</b>	0.9848
	PROPOSED	<b>0.9975</b>	<b>0.9965</b>	0.9771	<b>0.9949</b>	<b>0.9826</b>	<b>0.9897</b>	<b>0.9841</b>	0.9917	<b>0.9856</b>	<b>0.9865</b>	0.9821	<b>0.9860</b>
MSFA [16]	BTES [21]	0.9598	0.9934	0.9418	0.9875	0.9557	0.9676	0.9664	0.9801	0.9512	0.9770	0.9524	0.9654
	LI [23]	0.9610	0.9936	0.9589	0.9895	0.9641	0.9734	0.9687	0.9889	0.9561	0.9791	0.9612	0.9708
	IID [26]	0.9891	0.9936	0.9745	0.9941	0.9780	0.9859	0.9703	0.9874	0.9610	0.9803	0.9807	0.9759
	GF [22]	0.9899	0.9943	0.9325	0.9972	0.9805	0.9789	0.9728	0.9910	0.9589	0.9831	0.9790	0.9770
	POS [12]	0.9889	0.9939	<b>0.9812</b>	0.9964	0.9801	0.9881	0.9765	<b>0.9922</b>	0.9742	0.9858	<b>0.9825</b>	0.9822
	PROPOSED	<b>0.9975</b>	<b>0.9965</b>	0.9801	<b>0.9981</b>	<b>0.9826</b>	<b>0.9910</b>	<b>0.9782</b>	0.9917	<b>0.9791</b>	<b>0.9875</b>	0.9821	<b>0.9837</b>

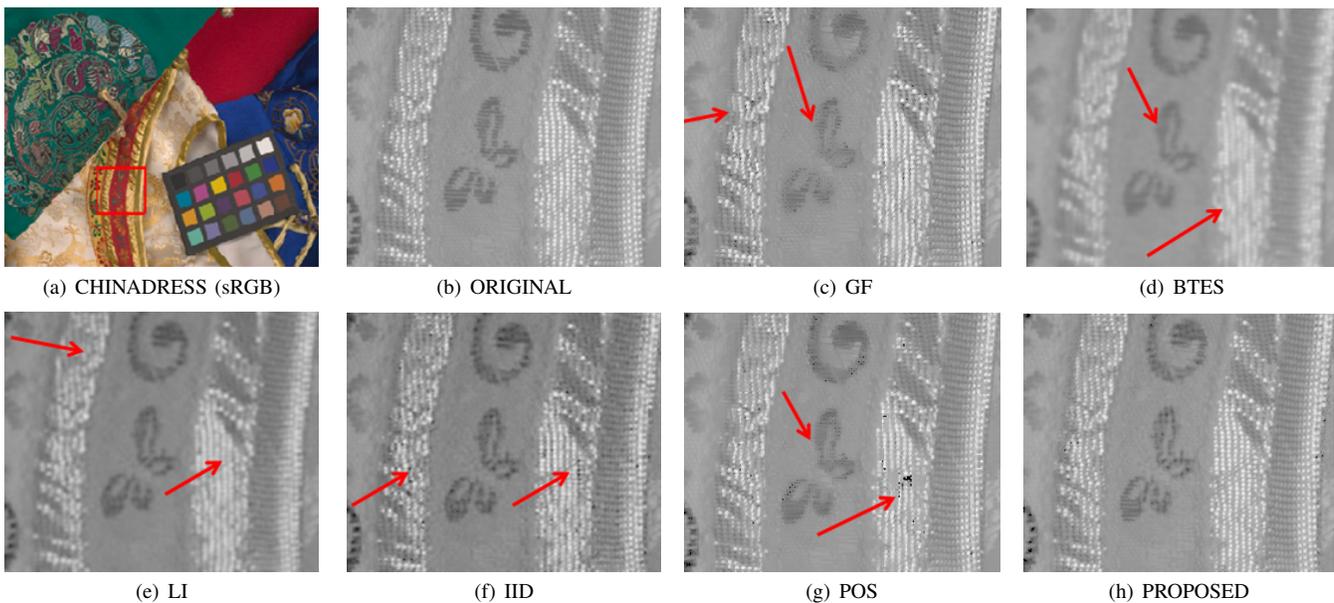


Fig. 10. Visual comparison of R band of the CHINADDRESS image in the Monno dataset. (Gamma correction is applied for the display.)

intra-prediction scheme is needed for the reconstruction in this region.

Table III gives the PSNR performance of the different methods on both the Cave and Monno datasets. In the case of the MSFA pattern (Fig. 1(b)) [12], we can observe that our algorithm achieves the highest average PSNR in almost all the bands for both datasets. Similarly, for the MSFA pattern (Fig. 2(a)) [16], we can see that our algorithm consistently performs well in all the band images. The Cave dataset consists of images having more smooth objects with higher spectral correlation (Table I), and thus the PSNR performances of the other algorithms are closer to that of the proposed algorithm. Images in the Monno dataset consist of textured objects with lower spectral correlation, and thus the proposed algorithm significantly outperforms the other methods, as it is adaptive to spectral correlations.

Table IV gives the SSIM performance comparison of the different algorithms for both datasets. The SSIM model is more consistent with human visual perception, and when the SSIM value is close to 1, the reconstructed image is similar

to the original image. We can clearly see that the proposed algorithm achieves a higher SSIM in almost all the bands for both datasets. From Table III and IV, we can say that the objective performance of the proposed algorithm, both in terms of PSNR and SSIM, outperforms the existing methods.

For the subjective quality evaluation we show the reconstructed images of the different methods and compare them with the proposed algorithm. Fig. 9 shows a visual comparison of the demosaicked Or band of the SPONGE image produced by different methods. The reconstructed image by the GF [22] method looks sharper than the original image and has artifacts, as indicated by the red arrows, whereas the image reconstructed by the BTES [21] method has lost all its detail. Similarly, the reconstructed images of LI [23] and POS [12] show artifacts as well as unclear text, whereas the proposed algorithm preserves the edges as well as produces clearer text. We also observe that although IID [26] uses both spatial and spectral correlations for reconstruction, its performance is lacking when compared to the proposed algorithm. Our algorithm is designed in such a way that it is adaptive to

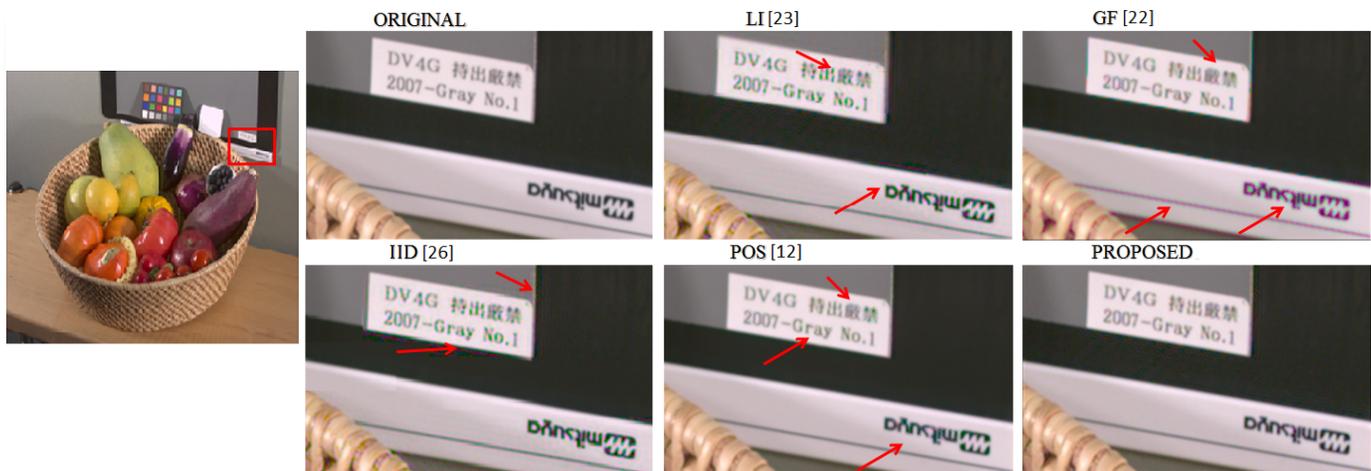


Fig. 11. Visual comparison of sRGB of the FRUIT image in the Monno dataset. (Gamma correction is applied for the display.)

TABLE V  
AVERAGE PSNR PERFORMANCE OF PROPOSED ALGORITHM USING DIFFERENT TECHNIQUES AS INITIALIZATION ON THE MONNO DATASET.

	R	B	Cy	Or	Mean
GIF [57]	53.65	46.90	49.99	46.65	49.29
JBU [56]	53.92	47.11	50.99	47.01	49.75
AKU [4]	54.55	47.11	51.25	47.47	50.09
GF [22]	54.95	47.32	51.77	47.92	50.49
POS [12]	54.63	47.88	51.72	47.80	50.50
IDEAL	<b>56.07</b>	<b>48.63</b>	<b>53.52</b>	<b>50.22</b>	<b>52.11</b>

spectral correlations and assigns the weights judiciously to both the intra- and inter-prediction methods.

In Fig. 10, we show a visual comparison of the demosaicked R band of the CHINADRESS image that consists of textured regions. From the visual comparison, we can see that the existing methods either lose the fine details or add extra artifacts in the reconstructed results, whereas our algorithm suppresses these artifacts and produces a better reconstruction as compared to the existing algorithms.

Fig. 11 shows a visual comparison of the FRUIT image in the Monno dataset. We convert both the original and reconstructed 5-band images to the standard RGB (sRGB) domain for illustration purposes, and the transformation is done using spatio-spectral Wiener estimation [4,12]. From the visual comparison, we can see that the existing methods either blur the edges or produce significant color artifacts on the text, whereas the proposed algorithm reduces the artifacts significantly and also preserves the edges at the same time. From both the objective and subjective quality assessments, we can say that the proposed algorithm has the ability to reduce the artifacts as compared to the existing algorithms.

To estimate the optimal weights in (30), the proposed algorithm uses an existing method for initialization. This is akin to most of the multispectral demosaicking algorithms, which also need an initial interpolation for initialization and then incorporate the result into their reconstruction frameworks. All the RI-based algorithms [28,29,30,54,55] generate an initial interpolation using a guided filter [57] and then use it in their

proposed framework to drive the interpolation of each band. AKU [4] uses the joint bilateral upsampling (JBU) method [56] to generate a guided image, whereas GF [22] and POS [12] use a guided image filtering (GIF) [57] method to develop a reference image.

To make a fair comparison, we investigate the performance of the proposed algorithm using these techniques as the initialization for our framework, and the results are shown in Table V. We can see that most of them yield a similar performance, but we use POS [12] for our framework as it is computationally very simple and is feasible in a real hardware setup. We also show the performance of the proposed algorithm in the ideal case using a ground-truth image for initialization, and as expected, the proposed algorithm in the ideal case has *superior* performance than in the practical cases.

### B. 3-band Demosaicking

Our proposed algorithm can easily be extended to CFA patterns for 3-band demosaicking. The density of the G pixels in the CFA pattern (Fig. 1(a)) and the MSFA pattern is the same, whereas the density of the other bands in the CFA is higher than that in the MSFA pattern and thus eases the interpolation process. The reconstruction of the G band is done in exactly the same manner as is done in the case of 5-band demosaicking. For the R and B band, the modulation signal in (8) is changed, and thus the color-difference signal  $K_R(u, v)$  in (9) will be changed as well. We follow a similar strategy in estimating the cut-off frequency for the circular LPF in (9) and in (21) in the case of CFA demosaicking.

We evaluate the performance of the proposed algorithm on the IMAX dataset, which consists of images dominated by highly textured regions with low spectral correlations. We compare the performance of the proposed algorithm with state-of-the-art demosaicking methods: the integrated gradient (IGD) [36], gradient-based threshold free demosaicking (GBTf) [37], local directional interpolation and nonlocal adaptive thresholding (NAT) [39], four-direction interpolation (FDI) [54], minimized-Laplacian residual interpolation (MLRI) [55] and iterative residual interpolation (IRI) [30].

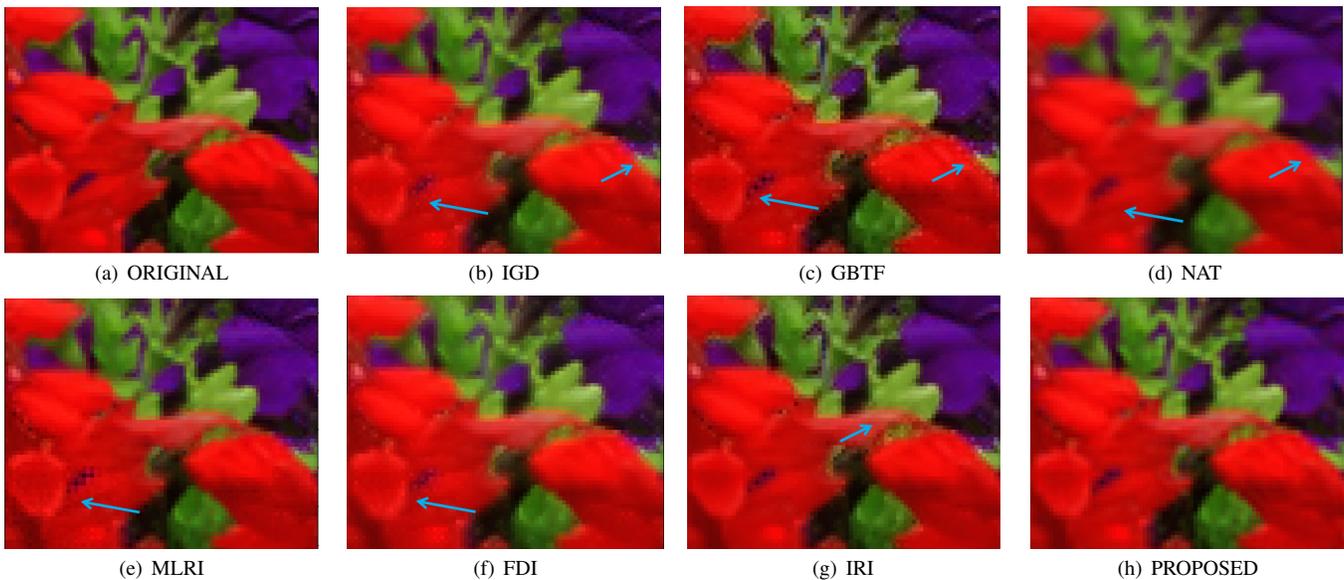


Fig. 12. Visual comparison of FLOWER image in IMAX dataset.

TABLE VI  
AVERAGE PSNR AND SSIM RESULTS ON THE IMAX DATASET.

Algo.	AVG. PSNR			AVG. SSIM		
	R	G	B	R	G	B
IGD [36]	34.83	37.38	33.96	0.9423	0.9622	0.9082
GBT [37]	33.48	36.59	32.71	0.9317	0.9558	0.8927
SSD [38]	35.02	38.27	33.80	0.9471	0.9779	0.9114
NAT [39]	36.31	39.82	34.50	0.9604	0.9787	0.9266
MLRI [55]	36.35	39.90	35.36	0.9601	0.9785	0.9384
IRI [30]	36.80	<b>40.28</b>	35.42	0.9633	<b>0.9818</b>	0.9411
FDI [54]	36.31	39.99	35.02	0.9601	0.9790	0.9391
PROPOSED	<b>37.18</b>	39.80	<b>35.71</b>	<b>0.9822</b>	0.9782	<b>0.9599</b>

In Table VI, we present the performances of the aforementioned methods and present the PSNR and SSIM performances on the IMAX dataset. For G band reconstruction, our algorithm is lacking both in terms of PSNR and SSIM as we do not use other bands as prior information. In this paper, we focus more on the other subsampled bands by exploiting the property of spectral correlation in a more adaptive manner. This is why our algorithm achieves better results in other bands, both in terms of PSNR and SSIM, as shown in Table VI.

In Fig. 12, we show a visual comparison of a 3-band demosaicked image, FLOWER, with different methods and compare the same with the proposed algorithm. We can observe that the existing methods produce artifacts (shown by blue arrows) near the red petals and are not able to preserve the edges, whereas the proposed algorithm reduces the artifacts and thus achieves a superior performance to the existing algorithms.

In summary, our algorithm outperforms the existing residual interpolation (RI)-based approaches [30,55] as well as algorithms based on the color-difference assumption [15,31-37] both in terms of objective and subjective quality. Algorithms based on color-difference assumptions are sensitive to spectral correlation. In the case of lower spectral correlation, the *Correlation Assumption* of such methods fails, and thus aliasing interferences increase in (9), which results in the

failure of the *Low-Pass Assumption*. Differently, we introduce a circular LPF and an intra-prediction scheme to address these issues. Our algorithm focuses more on other bands as they are significantly undersampled, and thus reconstruction becomes a challenging task. To improve the performance of the proposed algorithm, especially for the G band, we can use some post-processing techniques, as used in the CFA demosaicking method [30], to achieve a superior performance.

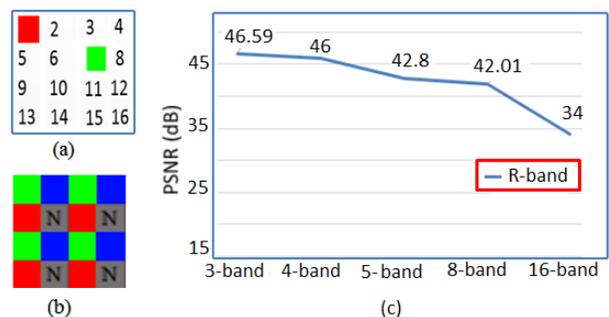


Fig. 13. (a) 16-band MSFA, (b) RGB\_NIR [9,12], (c) Performance analysis of proposed algorithm on different MSFA patterns (R component of CHINADRESS image)

### C. Performance of Proposed Algorithm on Different Patterns

In the previous subsections, we show the performance of the proposed algorithm on 5-band MSFA and 3-band MSFA patterns, where the sampling density of the G band is higher than that of other bands. However, our proposed algorithm is not limited to such patterns and is designed in such a way that it can be extended to any *arbitrary pattern*. In this subsection, we discuss the performance of the proposed algorithm with other MSFA patterns and also study its behavior with different sampling densities of the G band.

In Fig. 13 (c), we show the performance of the proposed ASCD algorithm on 3-band MSFA (Fig. 1(a)), 5-band MSFA

TABLE VII  
AVERAGE PSNR PERFORMANCE OF PROPOSED ALGORITHM WITH DIFFERENT BLOCK SIZES.

Size	R band			B band		
	Cave	Monno	IMAX	Cave	Monno	IMAX
4 × 4	45.78	54.55	37.08	44.91	47.81	35.64
6 × 6	<b>45.81</b>	<b>54.64</b>	<b>37.18</b>	<b>44.94</b>	<b>47.88</b>	<b>35.71</b>
8 × 8	45.76	54.55	37.08	44.90	47.82	35.63
10 × 10	45.70	54.42	37.01	44.85	47.71	35.55
12 × 12	45.63	54.38	36.90	44.80	47.63	35.62

(Fig. 1(b)), 4-band MSFA (Fig. 2(e)), 8-band MSFA (Fig. 2(f)) and 16-band MSFA (Fig. 13 (a)). A general observation that can be drawn from Fig. 13(c) is that the performance of the proposed algorithm decreases from the 3-band to the 16-band MSFA which is very obvious. This is because if we array more bands in the MSFA pattern, the sampling densities of each spectral band become lower and it becomes a challenging task for the demosaicking. More observations are as follows:

- 1) Although the sampling density of the R band is the same for the 3-band (Fig. 1(a)) and 4-band MSFA (Fig. 2(e)), the reconstruction performance is lower in the case of the 4-band MSFA as the density of the G band is decreased. A similar analysis can be made for the 5-band and 8-band patterns.
- 2) We have shown the reconstruction for a 16-band MSFA pattern, where each color appears only once in a 4 × 4 pattern (Fig. 13(a)), and thus it can be considered as a case in which all the bands are *severely undersampled*. In this case, we can see that the performance decreases significantly as compared to the other MSFA patterns. However, we believe that the existing algorithms may suffer too, because most of them either use inter-color correlation or spatial correlation for the reconstruction.

With the performance analysis given above, we can say that when all the bands are severely undersampled, the proposed method will show a drastic change in performance. The performance also depends on how other bands are arranged on the image sensor.

Our proposed algorithm can also be extended to MSFA patterns that consist of spectral bands beyond the visible range. For reference, we have shown a pattern in Fig. 13 (b), which consists of RGB and near-infrared (NIR) bands. In this particular pattern, all the bands are equally sampled, and thus the demosaicking problem in such a case is equivalent to the one we show for the 16-band pattern in Fig. 13(a). Both in the case of the 16-band pattern and the RGB\_NIR pattern, we follow the same strategy for reconstruction as is done in the case of the 5-band MSFA.

#### D. Discussion of Block Size and Computational Complexity

Our algorithm works on a block-by-block basis (non-overlapping) to generate a more accurate prediction for reconstruction. We test the performance of the proposed algorithm on all the datasets using different block sizes, and the results are shown in Table VII. We can observe that a block size of 6 × 6 has a reasonable performance as compared to the other block sizes for all the datasets. In the case of higher spectral

correlation (Cave dataset), even when we increase or decrease the block size, the spectral correlation among the bands does not vary much, and thus the validity of both the assumptions remains unchanged with different block sizes. This is why the performance of the proposed algorithm doesn't vary much for the Cave dataset with different block sizes in Table VII. However, in the case of lower spectral correlation (IMAX and Monno), with a change in the block size, spectral correlation may vary, and thus the performance also varies with different block sizes.

Our algorithm is computationally very simple and requires only basic operations, in contrast to other existing methods. For each block, we perform a 2-D low pass filtering on (9) and (21) to estimate  $\hat{r}_{CDI}$  and  $\hat{r}_{LPF}$  respectively and then estimate the optimal parameters ( $w_1, w_2$ ) from (30). On a desktop computer with an Intel Core i7 CPU, our MATLAB implementation of the proposed algorithm takes about 0.7 sec to reconstruct an image of size 512 × 512 and 1.7 sec for an image of size 1368 × 1800 using a block of size 6 × 6.

## V. CONCLUSION

In this paper, we have addressed the challenging problem of multispectral demosaicking by demonstrating a frequency-domain analysis of the subsampled color-difference signal. With the help of this analysis, we studied the important assumption of spectral correlation and explained why it is image dependent, which leads to aliasing interferences among the color-difference spectra. To address the issues, we proposed an adaptive spectral-correlation-based demosaicking (ASCD) that uses a novel anti-aliasing filter to reduce the interferences, and we incorporated an intra-prediction scheme to generate a more accurate reconstruction of the undersampled components. The proposed scheme can be easily extended to any CFA and MSFA patterns by exploiting the spectral correlation property more effectively than the existing algorithms. Our approach is computationally very simple and outperforms the existing methods both in terms of objective and subjective quality.

## ACKNOWLEDGMENT

We would like to thank Dr. Yusuke Monno, Tokyo Institute of Technology, Japan for providing codes and datasets for performance comparison.

## REFERENCES

- [1] C. Yang, J. H. Everitt, Q. Du, B. Luo, and J. Chanussot, "Using high-resolution airborne and satellite imagery to assess crop growth and yield variability for precision agriculture," in *Proceedings of the IEEE*, vol.101, no.3, pp.582-592.
- [2] A. Zia, J. Liang, J. Zhou, and Y. Gao, "3D reconstruction from hyperspectral images," in *Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on*, vol., no., pp.318-325, 5-9 Jan. 2015.
- [3] G. Lu and B. Fei, "Medical hyperspectral imaging: A review," in *Journal of Biomedical Optics*, vol.19(1), pp. 10901, Jan. 2014.
- [4] Y. Monno, M. Tanaka, and M. Okutomi, "Multispectral demosaicking using adaptive kernel upsampling," in *Image Processing (ICIP), IEEE International Conference on* vol., no., pp.3157-3160, 11-14 Sept. 2011.
- [5] H. Fukuda, T. Uchiyama, H. Haneishi, M. Yamaguchi, and N. Ohyama, "Development of 16-bands multispectral image archiving system," in *Proc. of SPIE*, vol. 5667, pp.136-145, 2005.
- [6] C. Cui, H. Yoo and M. Ben-Ezra, "Multi-spectral imaging by optimized wide band illumination," in *International Journal of Computer Vision*, vol. 86, no. 2-3, pp.140-151, 2010.

- [7] S. Han, I. Sato, T. Okabe, and Y. Sato, "Fast spectral reflectance recovery using DLP projector," in *Proc. of Asian Conf. on Computer Vision (ACCV)*, pp. 318-330, 2010.
- [8] K. Ohsawa, T. Ajito, H. Fukuda, Y. Komiya, H. Haneishi, M. Yamaguchi, and N. Ohya, "Six-band HDTV camera system for spectrum-based color reproduction," in *Journal of Imaging Science and Technology*, vol. 48, no. 2, pp. 85-92 (2004).
- [9] P. J. Lapray, X. Wang, J. B. Thomas, and P. Gouton, "Multispectral filter arrays: Recent advances and practical implementation," in *Sensors*, 14(11):21626-21659, 2014.
- [10] B. Geelen, N. Tack, A. Lambrechts, "A compact snapshot multispectral imager with a monolithically integrated per-pixel filter mosaic," in *Proc. SPIE*, 2014.
- [11] M. A. Martinez, E. M. Valero, J. H. Andrs, J. Romero, G. Langfelder, "Combining transverse field detectors and color filter arrays to improve multispectral imaging systems," in *Application Optical*, vol. 53, pp. C14-C24, 2014.
- [12] Y. Monno, S. Kikuchi, M. Tanaka, and M. Okutomi, "A practical one-shot multispectral imaging system using a single image sensor," in *Image Processing, IEEE Transactions on*, vol.24, no.10, pp.3048-3059, Oct. 2015.
- [13] J.-B. Thomas, P.-J. Lapray, P. Gouton, C. Clerc, "Spectral Characterization of a Prototype SFA Camera for Joint Visible and NIR Acquisition," in *Sensors*, 16, 993, 2016.
- [14] Y. Murakami, M. Yamaguchi, N. Ohya, "Hybrid-resolution multispectral imaging using color filter array," in *Optical express*, vol. 20, pp. 7173-7183, 2012.
- [15] X. Li, B. Gunturk, and L. Zhang, "Image demosaicing: A systematic survey," in *proc. SPIE*, vol.6822, p. 68221J, 2008.
- [16] L. Miao and H. Qi, "The design and evaluation of a generic method for generating mosaicked multispectral filter arrays," in *Image Processing, IEEE Transactions on*, vol.15, no.9, pp.2780-2791, Sept. 2006.
- [17] J. Brauers and T. Aach, "A color filter array based multispectral camera," in *Workshop Farbbildverarbeitung*, 2006.
- [18] H. K. Aggarwal and A. Majumdar, "Compressive sensing multi-spectral demosaicing from single sensor architecture," in *Signal and Information Processing (ChinaSIP), 2014 IEEE China Summit & International Conference on*, vol., no., pp. 334-338, 2014.
- [19] F. Yasuma, T. Mitsunaga, D. Iso, and S.K. Nayar, "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," in *Image Processing, IEEE Transactions on*, vol.19, no.9, pp.2241-2253, Sept. 2010.
- [20] X. Wang, J. B. Thomas, J. Y. Hardeberg, and P. Gouton, "Median filtering in multispectral filter array demosaicing," in *Proc. of SPIE*, vol. 8660, pp. 86600E, 2013.
- [21] L. Miao, H. Qi, R. Ramanath, and W.E. Snyder, "Binary tree-based generic demosaicing algorithm for multispectral filter arrays," in *Image Processing, IEEE Transactions on*, vol.15, no.11, pp.3550-3558, Nov. 2006.
- [22] Y. Monno, M. Tanaka, and M. Okutomi, "Multispectral demosaicing using guided filter," in *Proc. of SPIE*, vol.8299, pp. 001-007, 2012.
- [23] C. Wang, X. Wang, and J. Hardeberg, "A linear interpolation algorithm for spectral filter array demosaicing," in *In Abderrahim Elmoataz, Olivier Lezoray, Fathallah Nouboud, and Driss Mammass, editors, Image and Signal Processing, Springer International Publishing*, vol. 8509, pp 151-160, 2014.
- [24] B. C. de Lavarne, D. Alleysson, and J. Herault, "Practical implementation of Immse demosaicing using luminance and chrominance spaces," in *Computer Vision and Image Understanding*, vol.107(1), no., pp.3-13, 2007.
- [25] D. Kiku, Y. Monno, M. Tanaka, and M. Okutomi, "Residual interpolation for color image demosaicing," in *Image Processing (ICIP), IEEE International Conference on*, vol., no., pp.2304-2308, 15-18 Sept. 2013.
- [26] S. Mihoubi, O. Losson, B. Mathon and L. Macaire, "Multispectral demosaicing using intensity-based spectral correlation," in *Image Processing Theory, Tools and Applications (IPTA), 2015 International Conference on*, vol., pp. 461-466, 2015.
- [27] X. Wang, J. B. Thomas, J. Y. Hardeberg, P. Gouton, "Discrete wavelet transform based multispectral filter array demosaicing," in *In Proceedings of the Colour and Visual Computing Symposium (CVCS)*, vol., pp. 1-6., 2013.
- [28] S. P. Jaiswal, L. Fang, V. Jakhethiya, M. Kuse and O. C. Au, "Optimized high-frequency based interpolation for multispectral demosaicing," in *Proc. of Image Processing (ICIP), IEEE International Conference on*, 2016.
- [29] Y. Monno, D. Kiku, S. Kikuchi, M. Tanaka, and M. Okutomi, "Multispectral demosaicing with novel guide image generation and residual interpolation," in *Image Processing (ICIP), IEEE International Conference on*, vol., no., pp.645-649, 27-30 Oct. 2014.
- [30] W. Ye and K. K. Ma, "Color image demosaicing using iterative residual interpolation," in *Image Processing, IEEE Transactions on*, vol.24, no.12, pp.5879-5891, Dec. 2015.
- [31] S. C. Pei and I. K. Tam, "Effective color interpolation in CCD color filter arrays using signal correlation," in *Circuits and Systems for Video Technology, IEEE Transactions on*, vol.13, no.6, pp.503-513, June 2003.
- [32] B.K. Gunturk, J. Glotzbach, Y. Altunbasak, R.W. Schafer, R.M. Mersereau, "Demosaicking: color filter array interpolation," in *Signal Processing Magazine, IEEE*, vol.22, no.1, pp.44-54, Jan. 2005.
- [33] D. Zhang and X. Wu, "Color demosaicing via directional linear minimum mean square-error estimation," in *Image Processing, IEEE Transactions on*, vol.14, no.12, pp.2167-2178, Dec. 2005.
- [34] Xin Li, "Demosaicing by successive approximation," in *Image Processing, IEEE Transactions on*, vol.14, no.3, pp.370-379, March 2005.
- [35] C.-Y. Su, "Highly effective iterative demosaicing using weighted-edge and color-difference interpolations," in *IEEE Trans. on Consumer Electronics*, vol. 52, no. 2, pp. 639-645, 2006.
- [36] K.-H. Chung and Y.-H. Chan, "Low-complexity color demosaicing algorithm based on integrated gradients," in *Journal of Electronic Imaging*, vol. 19, no. 2, p. 021104, 2010.
- [37] I. Pekkucuksen and Y. Altunbasak, "Gradient based threshold free color filter array interpolation," in *Image Processing (ICIP), IEEE International Conference on*, vol., no., pp.137-140, 26-29 Sept. 2010.
- [38] A. Buades, B. Coll, J. M. Morel, and C. Sbert, "Self-similarity driven color demosaicing," in *Image Processing, IEEE Transactions on*, vol.18, no.6, pp.1192-1202, June 2009.
- [39] L. Zhang, X. Wu, A. Buades, and X. Li, "Color demosaicing by local directional interpolation and nonlocal adaptive thresholding," in *Journal of Electronic imaging*, vol. 20, No. 2, 2011.
- [40] J. Duran and A. Buades, "Self-similarity and spectral correlation adaptive algorithm for color demosaicing," in *Image Processing, IEEE Transactions on*, vol.23, no.9, pp.4031-4040, Sept. 2014.
- [41] S.P. Jaiswal, O.C. Au, V. Jakhethiya, Y. Yuan, and H. Yang, "Exploitation of inter-color correlation for color image demosaicing," in *Image Processing (ICIP), 2014 IEEE International Conference on*, vol., no., pp.1812-1816, 27-30 Oct. 2014.
- [42] D. Alleysson, S. Susstrunk and J. Herault, "Linear demosaicing inspired by the human visual system," in *Image Processing, IEEE Transactions on*, vol. 14, no. 4, pp. 439-449, April 2005.
- [43] K. Hirakawa and P. J. Wolfe, "Spatio-Spectral Color Filter Array Design for Optimal Image Recovery," in *Image Processing, IEEE Transactions on*, vol. 17, no. 10, pp. 1876-1890, Oct. 2008.
- [44] K. Hirakawa and T.W. Parks, "Adaptive homogeneity-directed demosaicing algorithm," in *Image Processing, IEEE Transactions on*, vol.14, no.3, pp.360-369, March 2005.
- [45] G. Jeon and E. Dubois, "Frequency-domain methods for demosaicing of Bayer-sampled color images," in *Signal Processing Letters, IEEE*, vol.12, no.12, pp.847-850, Dec. 2005.
- [46] B. Leung, G. Jeon, and E. Dubois, "Least-Squares Luma-Chroma Demultiplexing Algorithm for Bayer Demosaicing," in *Image Processing, IEEE Transactions on*, vol.20, no.7, pp.1885-1894, July 2011.
- [47] G. Jeon and E. Dubois, "Demosaicing of Noisy Bayer-Sampled Color Images With Least-Squares Luma-Chroma Demultiplexing and Noise Level Estimation," in *Image Processing, IEEE Transactions on*, vol.22, no.1, pp.146-156, Jan. 2013.
- [48] N.X. Lian, L. Chang, Y. P. Tan, and V. Zagorodnov, "Adaptive Filtering for Color Filter Array Demosaicing," in *Image Processing, IEEE Transactions on*, vol.16, no.10, pp.2515-2525, Oct. 2007.
- [49] L. Fang, O.C. Au, K. Tang, and A.K. Katsaggelos, "Antialiasing filter design for subpixel downsampling via frequency-domain analysis," in *Image Processing, IEEE Transactions on*, vol.21, no.3, pp.1391-1405, March 2012.
- [50] L. Fang, O.C. Au, Y. Chen, A.K. Katsaggelos, H. Wang, and X. Wen, "Joint Demosaicing and Subpixel-Based Down-Sampling for Bayer Images: A Fast Frequency-Domain Analysis Approach," in *Multimedia, IEEE Transactions on*, vol.14, no.4, pp.1359-1369, Aug. 2012.
- [51] J. Zhou and O.C. Au, "Determining the capacity parameters in PEE-based reversible image watermarking," in *Signal Processing Letters, IEEE*, vol.19, no.5, pp.287-290, May 2012.
- [52] CIE, <http://www.cie.co.at/index.php/LEFTMENU/DOWNLOADS>.
- [53] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," in *Image Processing, IEEE Transactions on*, vol.13, no.4, pp.600-612, April 2004.

- [54] Y. Kim and J. Jeong, "Four direction residual interpolation for demosaicking," in *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. PP, no. 99, pp. 1-1, 2015.
- [55] D. Kiku, Y. Monno, M. Tanaka, and M. Okutomi, "Minimized-laplacian residual interpolation for color image demosaicking," in *IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics*, p. 90230L-90230L, 2014.
- [56] J. Kopf, M. F. Cohen, D. Lischinski, and M. Uyttendaele, "Joint bilateral upsampling," in *ACM Transactions on Graphics*, vol. 26(3), no. 96, 2007.
- [57] K. He, J. Sun, and X. Tang, "Guided image filtering," in *Proc. of the 11th European Conference on Computer Vision (ECCV)*, pp. 1-14, 2010.



**Sunil P. Jaiswal** received the B. Tech degree from the LNM Institute of Information Technology (LNMIIT), Jaipur, India, in 2012. He is currently pursuing the Ph.D. degree in electronic and computer engineering in the Hong Kong University of Science and Technology (HKUST), Hong Kong since 2012. He also used to be a visiting scholar in the visual analytics and imaging lab in Stony Brook University, NY, USA and SUNY Korea in 2015/2016. He co-authored a best paper award in student category in I2MTC'2011. His research interest includes multi-

spectral imaging, 3-D medical imaging and Image quality assessment.



**Lu Fang** received her Ph.D in Electronic and Computer Engineering from HKUST in 2011, and B.E. from USTC in 2007, respectively. Dr. Fang's research interests include image/video processing, vision for intelligent robot, 3D reconstruction and tracking. Dr. Fang serves as TC member in Multimedia Signal Processing Technical Committee (MMSP-TC) in IEEE Signal Processing Society.



**Vinit Jakhetiya** received the B.Tech. degree in computer and communication engineering from LNM Institute of Information Technology, India, in 2011 and Ph.D degree in electronics and computer engineering at Hong Kong University of Science and Technology, Hong Kong, in 2016. From January 2015 to December 2015, he was a Visiting Student with the School of Computer Science and Engineering, Nanyang Technological University, Singapore, and later he joined as a Project Officer at the same university. His research interests include image/video

processing, image quality assessment, and visual perceptual modeling. Dr. Jakhetiya was the recipient of the best undergraduate paper award from the 28th IEEE International Instrumentation and Measurement Technology Conference (2011).



**Jiahao Pang** received the B.Eng. degree from South China University of Technology, Guangzhou, China, in 2010, and the M.Sc. and Ph.D. degrees from The Hong Kong University of Science and Technology, Hong Kong, in 2011 and 2016, respectively. He is currently a Researcher with SenseTime Group Limited, Hong Kong. His research interests include image/video processing, graph signal processing and deep learning.



**Klaus Mueller** holds a PhD in computer science from The Ohio State University. He is currently a professor of computer science at Stony Brook University, NY, USA and an adjunct senior scientist at Brookhaven National Lab.

His research interests are visual analytics, HCI, medical imaging, and high performance computing. He has published extensively in all of these areas and his work was cited more than 7,000 times. He won the NSF EARLY CAREER award in 2001, the SUNY Chancellor Award for Excellence in Scholarship and Creative Activity in 2011, and the IEEE Meritorious Service Certificate in 2016. He is a senior member of the IEEE and an Associate Editor-in-Chief of IEEE Transactions on Visualization and Computer Graphics. For more info see <http://www.cs.sunysb.edu/mueller/>



**Oscar C. Au** (S'87-M'90-SM'01-F'12) received the B.A.Sc. degree from the University of Toronto, in 1986, and the M.A. and Ph.D. degrees from Princeton University, in 1988 and 1991, respectively. After being a postdoctoral position at Princeton University for one year, he joined The Hong Kong University of Science and Technology (HKUST), as an Assistant Professor, in 1992, and promoted to Associate Professor and then Full Professor. He left HKUST in 2014 and migrated to USA.

His main research contributions are on video/image coding and processing, watermarking/light weight encryption, and speech/audio processing. Research topics include fast motion estimation for H.261/3/4/5, MPEG-1/2/4, and AVS, optimal and fast sub-optimal rate control, mode decision, transcoding, denoising, deinterlacing, post-processing, multiview coding, view interpolation, depth estimation, 3DTV, scalable video coding, distributed video coding, subpixel rendering, JPEG/JPEG2000, HDR imaging, compressive sensing, halftone image data hiding, GPU-processing, and software-hardware co-design. He has authored 70 technical journal papers, more than 370 conference papers, three book chapters, and more than 70 contributions to international standards.

His fast motion estimation algorithms were accepted into the ISO/IEC 14496-7 MPEG-4 international video coding standard and the China AVS-M standard. His lightweight encryption and error resilience algorithms are accepted into the China AVS standard. He was the Chair of Screen Content Coding AdHoc Group in JCTVC for HEVC. He has more than 25 granted U.S. patents and is applying for 70+ more on his signal processing techniques.

Dr. Au is a fellow of IEEE and HKIE. He is/was an Associate/Senior Editor of several IEEE journals (TCSVT, TIP, TCAS1, SPL, JSTSP, SigView) and non-IEEE journals (JVCIR, JSPS, TSIP, JMM, JFI, and SWJ). He is a Guest Editor of some special issues in JSTSP and TCSVT. He was a BoG member and Vice President Technical Activity of APSIPA. He was the Chair of three technical committees (TCs): the IEEE CAS MSA TC, the IEEE SPS MMSP TC, and APSIPA IVM TC. He was a member of five other TCs: the IEEE CAS VSPC TC, the DSP TC, the IEEE SPS IVMSPTC, the IFS TC, and the IEEE ComSoc MMC TC. He served on two steering committees: the IEEE TMM, and the IEEE ICME. He also served on organizing committee of many conferences, including the ISCAS 1997, ICASSP 2003, and ISO/IEC 71st MPEG in 2005, and ICIP 2010. He was the General Chair of several conferences: PCM 2007, ICME 2010, and PV 2010. He received five best paper awards: SIPS 2007, PCM 2007, MMSP 2012, ICIP 2013, and MMSP 2013. He was the IEEE Distinguished Lecturer (DL) in 2009 and 2010, APSIPA DL in 2013 and 2014, and has been a keynote speaker multiple times.