Predicting the Quality of Demosaiced Images Using the Sparsity of Chroma Gradients

Carlo Noel Ochotorena, Cecille Adrianne Ochotorena, and Yukihiko Yamashita

***Abstract*—The design of most modern cameras utilizes a color filter array that downsamples and interleaves the red, green, and blue pixels of an image into a single mosaiced image. Such a design makes it necessary to interpolate the missing pixels for each color channel using a process known as demosaicing. While it is possible to fill in these pixels, the resulting images are inexact estimates of the true image, with different algorithms offering various levels of success. However, this degree of success cannot be directly quantified in the absence of the true image, making it difficult to design adaptive algorithms for demosaicing. This paper explores a no-reference simple metric for inferring the quality of the estimated image by measuring the sparsity of chroma gradients along four directions (SCG4). The said measure is shown to be significantly correlated with respect to the PSNR in simulations using the Kodak image database.**

***Index Terms*—demosaicing, color filter array, gradient, sparsity.**

I. Introduction

WITH the growing presence of imaging systems in the modern world, it is not surprising that the underlying technology behind such imaging devices have received a proportional amount of attention. In particular, digital imaging sensors have continuously been developed through the years. One of the most apparent aspects of this growth is seen in the resolution of the imaging sensors. As consumer video is pushing for 4K video resolutions and higher, sensor technology has to cope with the market demand. Alongside the increasing resolution [1], [2], [3], there is also a growing commercial interest in high speed video captures [4], [5], [6].

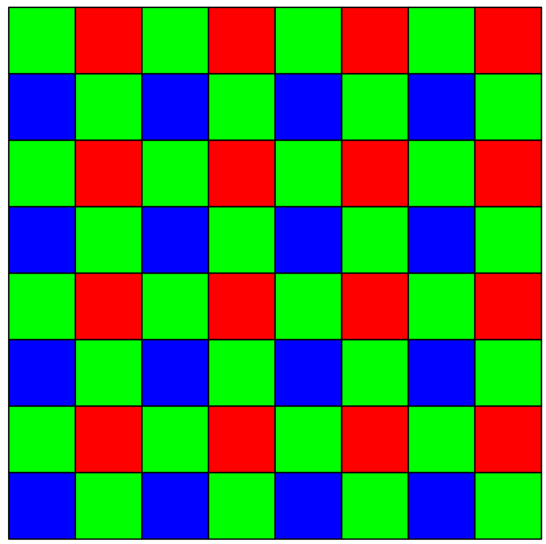

Fig. 1. Bayer color filter array (CFA). Each color site is comprised of a color filter (typically a chemical dye) placed in from of a photodetector.

Driven by the progressively rising requirements, researchers have come up with ways to improve sensor technologies. For instance, the construction of smaller sensors have enabled consumer devices such as mobile phones to capture high-resolution images and video [1], [2], [3]. Sensors with fast readout capabilities can now be utilized to capture high framerate videos (e.g. 1500 fps [6], 10000 fps [5], etc). Improving sensitivities have allowed images and videos to be captured under poor lighting conditions [7], [8]. In a similar manner, developments in sensitivity have also allowed for the lower levels of noise in captured images.

Despite the changes in sensor design throughout the years, one aspect of imaging sensors has remained vastly unchanged since the 1970s—the manner by which color is captured. At the very core of most imaging sensors is an array of photosensitive devices designed to capture incident light and translate the intensity into electronic signals [9]. However, these devices can only describe the intensity of

C. N. Ochotorena is a Ph.D. candidate of International Development Engineering at the Tokyo Institute of Technology, Japan (e-mail: carlo.ochotorena@dlsu.edu.ph)

C. A. Ochotorena is a Ph.D. candidate of Electronics and Communications Engineering at the De La Salle University, Philippines (e-mail: cec.ochotorena@dlsu.edu.ph)

Y. Yamashita is with the International Development Engineering Department, Tokyo Institute of Technology, Japan

light captured over a broad spectrum and not for specific wavelengths, thus making them incapable of quantifying the intensity of different colors.

To address this limitation, wavelength-specific filters are placed in front of the individual sensor cells to allow them to capture color-specific intensities. Collectively, these filters form what is known as a color filter array (CFA). A prominent example of this is the Bayer CFA [9] (shown in Figure 1), which is still used in many imaging sensors today. The downside to such an approach is that each cell in the array can only measure one particular color, effectively downsampling the individual color channels. The resulting captured image appears to be monochromatic (see Figure 2) as only a single flat image is obtained from the sensor. However, this flat image actually represents the interleaving of the intensities of three color channels.

As a result of the interleaving of color channels, the resulting image is often not in a usable form and has to be deinterleaved to form the individual color channels. This leaves missing pixels for each color channel which has to be interpolated to reconstruct the colored image. The interpolation process for images obtained through a CFA is specifically known as demosaicing and is generally a non-trivial reconstruction task as information is readily lost during the downsampling process. For this reason, demosaicing has been a subject of interest to many researchers since the conception of the CFA [11].

As with almost any reconstruction task, the true image, and subsequently the quality of the reconstruction, is unknown to the process. If such quality information were made available, an "oracle" process would be able to make optimal decisions during the reconstruction. Using this premise, this work proposes a no-reference metric that predicts the quality of demosaicing in the absence of the true image. By utilizing such a measure, more effective demosaicing algorithms can be designed. To understand how such a metric can be developed, we first introduce Bayer CFA along with some technical aspects of this array in Section II. An overview of some demosaicing paradigms is provided in Section III. Following this, the proposed metric is discussed in Section IV along with some experiments in Section V.

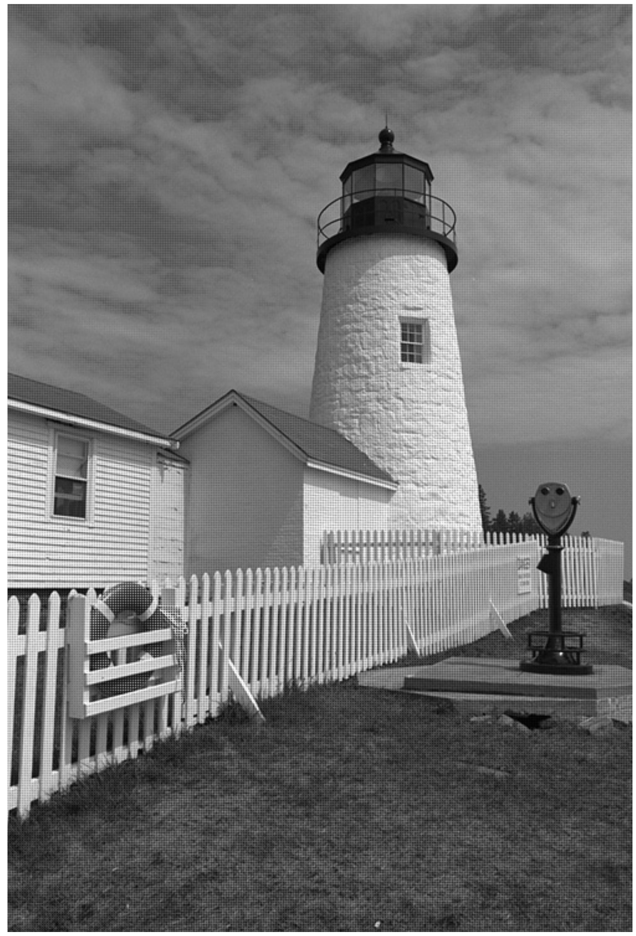

Fig. 2. A mosaiced image of the *Lighthouse* image from the *Kodak* color image database [10] passed through a simulated Bayer CFA.

II. The Bayer CFA

While there have been many color filter array (CFA) designs proposed, the Bayer CFA is still one of the most widely used patterns today [9], [11] This CFA (as illustrated in Figure 1) creates a repeated array of red, green, and blue filters in such a way that every 2×2 pixel area of the entire array contains exactly two green pixels and one each of the red and blue pixels. This construction is consistent with the observation that the human eyes are generally more sensitive to green wavelengths thus making the spatial resolution of the green channel more significant towards our perception of visual quality. To describe the Bayer CFA, we begin by defining the intensities of the red, green, and blue color channels of an image as $f_{\mathrm{R}}(x, y)$, $f_{\mathrm{G}}(x, y)$, and $f_{\mathrm{B}}(x, y)$, respectively for a given pixel coordinate (x, y). Using this notation, any CFA can be described using three masks corresponding to the color channels—$m_{\mathrm{R}}(x, y)$, $m_{\mathrm{G}}(x, y)$, and $m_{\mathrm{B}}(x, y)$. Using these masks, the mosaiced image obtained by the sensor is described as:

$$\begin{aligned} f_{\mathrm{CFA}}(x, y) = & f_{\mathrm{R}}(x, y) m_{\mathrm{R}}(x, y) + \\ & f_{\mathrm{G}}(x, y) m_{\mathrm{G}}(x, y) + \\ & f_{\mathrm{B}}(x, y) m_{\mathrm{B}}(x, y) \end{aligned} \tag{1}$$

With slight abuse of notation, we omit the coordinates for the individual channels in the succeeding discussion for brevity thus using the expressions f_{CFA}, f_{R}, f_{G}, and f_{B} instead.

In the specific case of the Bayer CFA, these masks are position dependent and repeat for every 2×2 region of the grid. While there are variations in the phase convention for the Bayer array, in this work, we followed the convention defined in [12] and define the masks as:

$$m_{\mathrm{R}}(x,y) = \begin{cases} 1, & x \text{ is odd and } y \text{ is even} \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

$$m_{\mathrm{G}}(x,y) = \begin{cases} 1, & x+y \text{ is even} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

$$m_{\mathrm{B}}(x,y) = \begin{cases} 1, & x \text{ is even and} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

A more mathematical way of expressing (2) – (4) is described by [12] in the form:

$$m_{\mathrm{R}}(x,y) = \frac{1}{4}\left[1-(-1)^x\right]\left[1+(-1)^y\right] \tag{5}$$

$$m_{\mathrm{G}}(x,y) = \frac{1}{2}\left[1+(-1)^{x+y}\right] \tag{6}$$

$$m_{\mathrm{B}}(x,y) = \frac{1}{4}\left[1+(-1)^x\right]\left[1-(-1)^y\right] \tag{7}$$

Combining (1) with (5) – (7) leads to the expression:

$$\begin{aligned} f(x,y) = & \frac{1}{4} f_{\mathrm{R}}\left[1-(-1)^x\right]\left[1+(-1)^y\right] + \\ & \frac{1}{2} f_{\mathrm{G}}\left[1+(-1)^{x+y}\right] + \\ & \frac{1}{4} f_{\mathrm{B}}\left[1+(-1)^x\right]\left[1-(-1)^y\right] \end{aligned} \tag{8}$$

which can subsequently be expanded and regrouped to form:

$$\begin{aligned} f(x,y) = & \left[\frac{1}{4}f_{\mathrm{R}} + \frac{1}{2}f_{\mathrm{G}} + \frac{1}{4}f_{\mathrm{B}}\right] + \\ & \left[-\frac{1}{4}f_{\mathrm{R}} + \frac{1}{2}f_{\mathrm{G}} - \frac{1}{4}f_{\mathrm{B}}\right](-1)^{x+y} + \\ & \left[-\frac{1}{4}f_{\mathrm{R}} + \frac{1}{4}f_{\mathrm{B}}\right]\left[(-1)^x - (-1)^y\right] \end{aligned} \tag{9}$$

The first term in the above expression is roughly analogous to a luma channel, f_L, while the remaining two terms can be interpreted as chroma channels, f_{C1} and f_{C2}, multiplied by distinct modulation functions. These individual components can be described as:

$$f_{\mathrm{L}} = \frac{1}{4}f_{\mathrm{R}} + \frac{1}{2}f_{\mathrm{G}} + \frac{1}{4}f_{\mathrm{B}} \tag{10}$$

$$f_{\mathrm{C1}} = -\frac{1}{4}f_{\mathrm{R}} + \frac{1}{2}f_{\mathrm{G}} - \frac{1}{4}f_{\mathrm{B}} \tag{11}$$

$$f_{\mathrm{C2}} = -\frac{1}{4}f_{\mathrm{R}} + \frac{1}{4}f_{\mathrm{B}} \tag{12}$$

This formulation implies that the mosaiced image is essentially a grayscale image (i.e., luma component) corrupted by "noise" from the chroma components. If the chroma components are fully known, the original image can easily be reconstructed from the mosaiced image. For this reason, the chroma components play a pivotal role in our proposed metric.

III. DEMOSAICING OVERVIEW

Equipped with a fair understanding on the design of the Bayer CFA, the next task is to design a system for reconstructing the original image from the mosaiced samples. The most trivial approach to this problem is to simply interpolate the missing pixels of each channel given the known pixels. Since the red and blue channels are effectively downsampled to half of their resolution, image interpolation techniques can be applied to restore the resolution. The green channel, on the other hand, is sampled in a quincunx pattern and has more samples available, thus making interpolation more effective. A clear downside to naïve interpolation is that the high frequency components of the red and blue channels that were discarded during the downsampling process cannot be restored upon doubling [11].

Beyond simply the loss of information, handling each color channel independently leads to more problematic artifacts in the image. Take, for instance, the well-known *Lighthouse* image from the *Kodak* database (see Figure 3). After mosaicing this image and subsequently applying a bicubic interpolant to each color channel, some visible color bands become apparent in the reconstructed image. These bands, often referred to as color moiŕe, result from incorrect and out-of-phase interpolation decisions. An early attempt to

Fig. 3. Reconstruction of a mosaiced image using bicubic interpolation. Naive interpolation can lead to color artifacts in the reconstructed image (right) that are not present in the original image (left).

address such artifacts is to adaptively alter the interpolation direction based on horizontal and vertical gradients of the color channels [11]. This allows the interpolation to adapt to edges in the underlying image to achieve better results. Similar approaches can adapt the weights of neighboring pixels to avoid edge discontinuities [13].

Another demosaicing paradigm is to exploit the relationship between color channels. Since color channels in images are naturally correlated with one another, many demosaicing techniques utilize this relationship to obtain better interpolation results [13], [14], [15], [16], [11], [17], [18]. For instance, many approaches operate under the assumption that hues in an image are relatively slow-changing allowing the differences or ratios between color channels to be utilized for prediction [13], [14], [15], [11]. Other approaches interpolate the green channel independently and use it to guide the red and blue interpolation [16], [17], [18]. A state-of-the-art, Minimized-Laplacian Residual Interpolation (MLRI), tackles the problem using a similar perspective by using guided upsampling process with the interpolated green channel to generate an initial estimate for the red and blue channels [17]. This estimate is then updated using an approximate Laplacian minimization criteria to arrive at a refined solution.

Aside from relying on color differences, some techniques also use adaptive fusion of different estimates of the image. A popular algorithm, known as Adaptive Homogeneity-Directed (AHD) demosaicing [19], performs both horizontal and vertical interpolations of the color channels and merges the two estimates based on a homogeneity metric designed to indicate the presence of color artifacts. Using a very different approach, the frequency-domain approach [12] models the chroma modulation process (discussed in Section II) in the Fourier domain. In this domain, it becomes apparent that color artifacts appear from crosstalk between the luma channel f_{L} and the second chroma channel f_{C2}. The authors of the said work proposed using a non-adaptive least-squares (LS) filter to extract the first chroma component f_{C1} and two non-adaptive LS filters to estimate f_{C2} based on the horizontally and vertically modulated contributions to the image. This results in two images which, similar to AHD, are adaptively merged to form the final image.

Despite the wide variations in approach of these different works, a common feature for many of them is that the optimization criteria is often carried out in the ℓ_2 domain. While this offers a simplistic solution to the problem, many properties of natural images are inherently sparse and not sufficiently served by the ℓ_2 domain. In the succeeding section, we highlight some of the said properties and propose a potential improvement to this criteria by exploiting such sparsity.

IV. Gradient Sparsity

A. Properties of Natural Images

In order to develop our metric, it is first useful to characterize certain properties of natural images. However, as the term "chroma" is used in an ambiguous manner in literature and may refer to different mixtures of the color channels, we explicitly present a series of experiments carried out using the 24 images of the *Kodak* color image database. This allows us to disambiguate the chroma definition and focus on the properties of the specific variant of chroma defined in equations (11) and (12). For these experiments, each of the 24 images are mapped into their respective luma and chroma components, thus generating a set of ground truths which are then characterized.

1) Sparsity of Gradients: One of the central features of our proposed metric is the use of gradients of components. To be specific, we define the gradient in four directions of an arbitrary component f as follows:

$$\nabla_{-x} f \equiv f(x,y) - f(x-1,y) \tag{13}$$
$$\nabla_{+x} f \equiv f(x,y) - f(x+1,y) \tag{14}$$
$$\nabla_{-y} f \equiv f(x,y) - f(x,y-1) \tag{15}$$
$$\nabla_{+y} f \equiv f(x,y) - f(x,y+1) \tag{16}$$

Using the above definitions, we calculate the gradients in four directions for all pixels in the *Kodak* database. To make these measures more meaningful, we obtain the histograms of each gradient direction and each component with a bin size of 1. The resulting plots, as seen in Figure 4, make it apparent that the majority of chroma gradients take on a zero value. The remaining non-zero gradients are still highly likely to be close to zero. On the other hand, the luma component does not exhibit the same behavior. Here, many of the gradients take nonzero values. This test demonstrates how the chroma components of natural images are *approximately* sparse (i.e., contain few large values).

2) Sparsity of Gradients in Four Directions: While the assertion that the gradient of chroma components are sparse, is useful in itself, we delve further into the properties of this sparsity. The previous experiment focused primarily on the individual gradient components without considering how these interact. Looking back at our definition of the four gradients (equations (13)–(16)), we find that these gradients are centered around a given pixel. We make use of this by counting the number of significant gradients around that pixel.

In this second experiment, we consider that every pixel can have between 0 to 4 non-zero gradients associated with it. Previously, we have established that the chroma

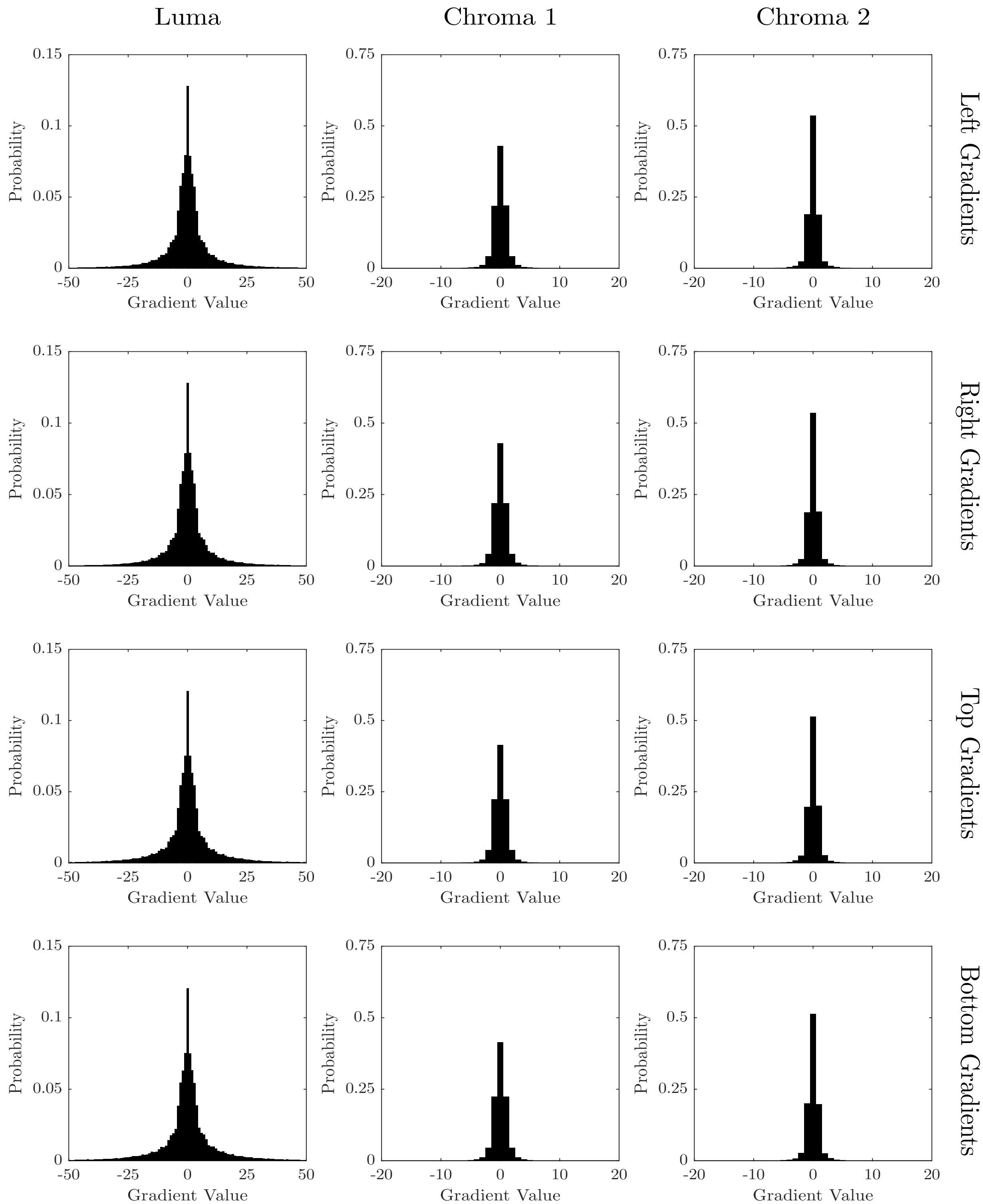

Fig. 4. Histogram of image gradients taken along four directions of various image components. Generally, the probabilities of a zero gradient are significantly higher for that of the chroma components compared to that of the luma component, regardless of the gradient direction.

components are *approximately* sparse due to factors such as noise. As such, a simple criteria such as dividing gradients between zeros and nonzeros is not necessarily reliable. Instead of applying a fixed threshold of 1, we model the probabilistic occurrence around a certain threshold. For each pixel center, the number of gradient magnitudes above or equal to the threshold is recorded. The probabilities of each number (from 0 to 4) are computed from all pixels accumulated from the *Kodak* database. The stacked bar graph in Figure 5 illustrates the trend at various threshold levels.

This test demonstrates how likely the gradients in four directions around a pixel are significant. In either chroma component, it is again apparent that most pixels have insignificant (i.e., below a given threshold) gradients. More importantly, this makes it clear that for a given pixel, there is a high probability that only one of the four gradients are of significant magnitude.

3) Correlating Gradients between Chroma Components: Having established a sparsity trend for each pixel, it is possible to apply the said observation towards developing our metric. However, there is one more property of natural images that is of some interest. An intuitive notion for images is that an edge occurring inside a given region often accompanies a change in color. If this were to hold true, such an edge would present itself in all three components of an image. To verify this, we consider the gradient magnitude at each direction. For each discretized magnitude in the first chroma component, we find the resulting distribution on the second chroma component. This can be visualized using the box plot in Figure 6.

If there exists no correlation between the gradients in two chroma components, we would expect the entire plot to remain close to zero (based on the observations from the previous experiment). The box plots obtained from the sample data, however, clearly demonstrate rising median and quartile levels as the chroma 1 gradient magnitude increases. Such a trend is consistent regardless of the gradient direction being studied. This test establishes that there is, indeed, some form of structural correlation between the two chroma components.

4) Summary: Given the results of the three experiments conducted on natural images, we have arrived at several key observations as summarized below:

1) Gradients of the chroma components are *approximately* sparse;
2) For a given pixel, there is a high probability that, at most, one of the four gradient directions are significant;
3) When a gradient in one of the four directions is significant in one chroma component, there is a non-negligible chance that it is also significant in the other.

B. The Proposed Metric

1) Reconstruction Quality of Luma-Chroma Components: For a predictive metric to be useful, it should not be dependent on any knowledge of the target image. This premise runs contradictory to the properties discussed in the previous section as these are dependent on the complete luma and chroma components. Furthermore, the use of the Bayer array prevents the calculation of any single gradient quantity from the mosaiced image. In order to tackle this limitation, in this work, we reconstruct one of the components and use it as *a priori* information for measuring the quality of the reconstruction.

As there are several components for a given image, a natural question that arises is: which one is suitable for use *a priori*? To answer this we characterize the reconstruction quality of each of the three components over different demosaicing methods. In particular, we use bilinear interpolation, bicubic interpolation, adaptive homogeneity directed (AHD) demosaicing [19], alternating projections (AP) [16], [20], Contour Stencils (CS) [21], [22], Directional Linear Minimum Mean Square-Error (DLMMSE) [23], [24], Least-squares Luma-Chroma Demosaicing (LSLCD) [12], Malvar-He-Cutler (MHC) demosaicing [25], [26], Minimized-Laplacian Residual Interpolation (MLRI) [17], and Successive Approximation (SA) [27]. Running each of these demosaicing algorithms across all the *Kodak* images, we arrive at various PSNR measures for the luma and chroma components. The medians of these measures can be seen in Figure 7.

An apparent observation from this test is that the chroma channels are better preserved in the reconstruction. Furthermore, the first chroma channel is typically reconstructed with higher fidelity. This is consistent with the assertion in [12] that this particular component suffers from minimal crosstalk by being modulated at a high frequency for both the horizontal and vertical axis. This choice of component is also convenient as the first chroma component can easily be approximated using the least-squares filters developed in [12] which can be applied through simple convolution.

2) Sparsity of Chroma Gradients: Having chosen a component to use *a priori*, we proceed to developing our proposed metric, the Sparsity of Chroma Gradients in Four Directions (SCG4). In the previous section, we assert that there exists a correlation between the occurrence of gradients of the two chroma components. This would intuitively lead to the notion that we can model the probability of significance of one component with respect to the significance of the other. However, doing so will enforce a strict measure along each gradient direction. We relax this criteria by, instead,

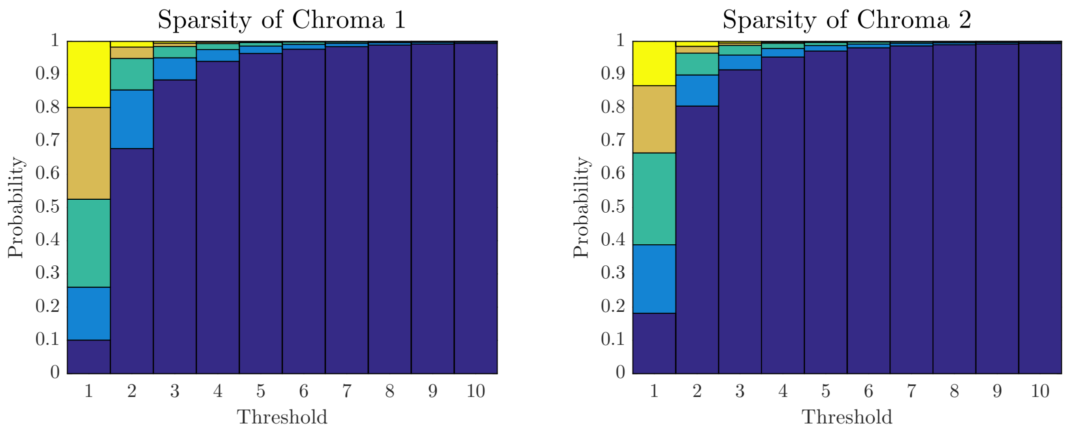

Fig. 5. Probability distribution of gradient magnitudes for each chroma channel above or equal to a given threshold. Since there are four gradient directions, each pixel in a given chroma component can have 0 (shown in blue) to 4 (shown in yellow) significant gradients occuring at different probabilities.

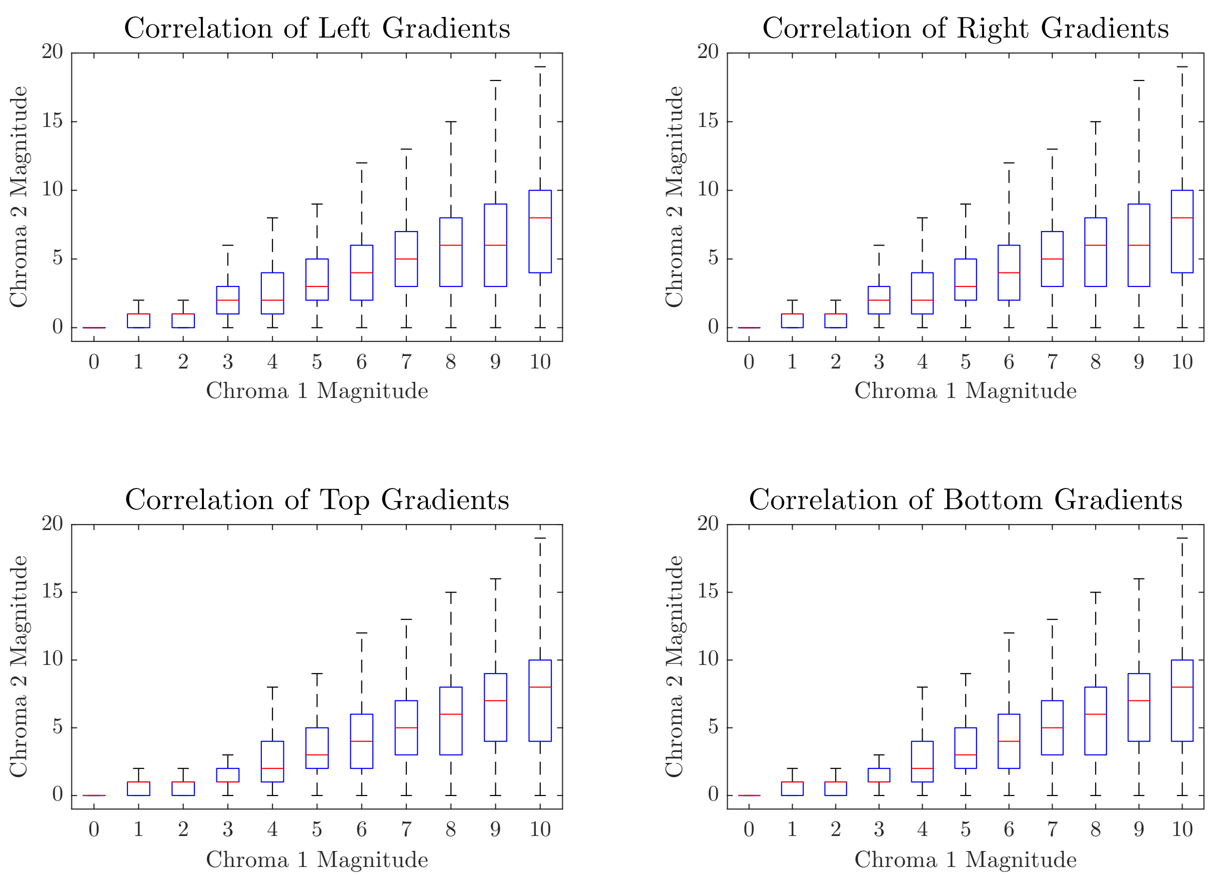

Fig. 6. Probability distribution of the gradients of the second chroma component at different magnitudes of the first chroma component. This box plot demonstrates that, regardless of the gradient direction, there exists a notable correlation between the gradients of the two chroma components for the medians (red line) and quartiles (blue box). It should also be noted that when the first chroma component has a magnitude of 1, the corresponding second chroma component has a very high probability of being zero resulting in a box plot with no visible quartiles and limits.

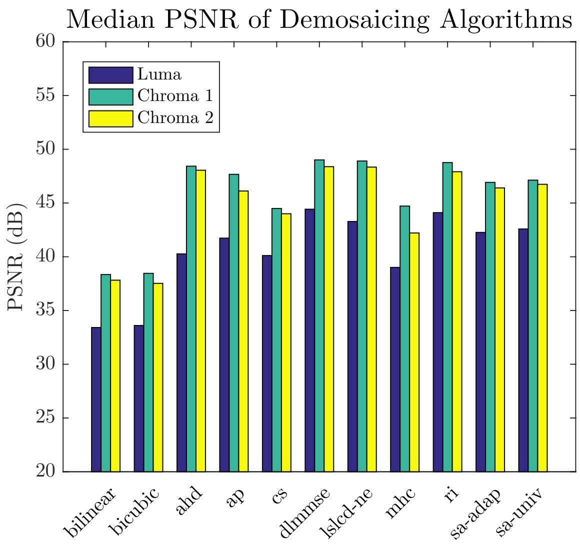

Fig. 7. Median PSNR of the luma and chroma components for the Kodak color image database.

measuring the number of significant gradients in each component defined as:

$$\mathcal{N}_t(x,y) = \mathcal{T}_t\left(|\nabla_{-x} f|\right) + \mathcal{T}_t\left(|\nabla_{+x} f|\right) + \mathcal{T}_t\left(|\nabla_{-y} f|\right) + \mathcal{T}_t\left(|\nabla_{+y} f|\right) \tag{17}$$

where the thresholding function $\mathcal{T}_t(x)$ is defined as:

$$\mathcal{T}_t(x) = \begin{cases} 0, & x < t \\ 1, & x \geq t \end{cases} \tag{18}$$

for an arbitrary threshold t. Using this measure, we can probabilistically model the relationship between two components. Since we use the first chroma as *a priori* information, we reconstruct it directly from the CFA by filtering and count the number of significant gradients for each pixel. We can then obtain the conditional probability of a certain number of significant gradients in the second chroma component exceeding that of the first:

$$p_{\mathrm{C2}}(n) \equiv p\left(\mathcal{N}_{\mathrm{C2}}(x,y) \geq n \mid \mathcal{N}_{\mathrm{C1}}(x,y) = n\right) \tag{19}$$

Such a probability is of interest in this work due to the correlation between the chroma components. If the second chroma component has more significant gradients than the first, we hypothesize that these gradients are more likely to be caused by color artifacts rather than actual image structures. We enforce this idea using empirical testing in the next section.

From a training perspective, these probabilities can readily be obtained from observations made between the filter-reconstructed C1 component, and the ground truth C2 component. Once the probabilities are obtained from the training set, they can simply be stored as a lookup-table for use in our metric. The metric itself is based on observations using a filter reconstructed C1 and the C2 component derived from the demosaiced image. This can be described using the following steps:

1) Given the CFA, use the filter from [12] to estimate the C1 component;
2) Given the demosaiced image, compute the C2 component using equation (12);
3) Calculate the four gradients at each pixel in the C1 and C2 components;
4) Determine the number of significant gradients around each pixel in each component;
5) Find the probability for each observed C2 count given the observed C1 count using equation (19);
6) The SCG4 metric is defined as the mean probability across all pixels in the image.

By using mean probabilities, we ensure that the final metric stays within the range of 0 to 1. A higher mean probability indicates a stronger correlation between the two chroma components and thus a better estimated quality. Conversely, a metric closer to zero indicates a C2 component that is inconsistent with the observations from the C1 component. In the succeeding section, we describe several tests that were used to verify the proposed metric.

V. Experiments

A. Subjective Quality

The first experiment conducted to verify the usefulness of the proposed metric is to evaluate it on a particularly

Fig. 8. A cropped region of the Lighthouse image sorted in descending order according to the SCG4 metric. Listed from left to right are the demosaiced images obtained using LSLCD [12], DLMMSE [23], [24], AP [16], [20], RI [17], AHD [19], SA-Universal and SA-Adaptive [27], CS [21], [22], MHC [25], [26], Bicubic interpolation, and Bilinear interpolation, respectively.

problematic region within the *Lighthouse* image from the *Kodak* database. To accomplish this, the probabilities were first obtained using the images statistics of all the images that are part of the *Kodak* database except for the *Lighthouse* image. This avoids self-training and is intended to strengthen the validity of the experiment.

Given the image statistics learned during the training process, we calculate the metric for the different demosaicing methods discussed in the previous section. In particular, the metric is only applied to a localized region within the Lighthouse image where a portion of the fence contains high-frequency information that is known to result in severe color moiŕe. These sub-images, sorted by the SCG4 metric, can be seen in Figure 8. While there are some deviations in the perceptual order, this figure shows that there is a subjective correlation between the visual quality and the proposed SCG4 metric.

B. Objective Quality with Cross-validation

Expanding on the methodology from the previous test, we obtained objective results from the *Kodak* color image database. Following a cross-validation procedure, each of the 24 images in the database are evaluated using statistics obtained from the other 23 images (excluding the image under test). Each of the demosaiced images are then objectively evaluated with respect to their original images using the peak signal-to-noise ratio (PSNR), a metric

TABLE I
Correlation Coefficients of the 24 Kodak Images at Various Thresholds

Kodak Image Number	Significance Threshold (t)									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	**0.468**	0.940	0.961	0.971	0.970	0.959	0.945	0.930	0.920	0.920
2	**-0.817**	0.866	0.888	0.890	0.885	0.879	0.848	0.768	0.735	0.735
3	**-0.902**	0.915	0.925	0.928	0.923	0.920	0.920	0.920	0.918	0.918
4	**-0.829**	0.918	0.946	0.962	0.969	0.971	0.974	0.977	0.977	0.977
5	**-0.705**	0.931	0.953	0.969	0.977	0.981	0.981	0.979	0.975	0.975
6	**-0.755**	0.948	0.970	0.975	0.972	0.964	0.954	0.945	0.935	0.935
7	**-0.920**	0.918	0.946	0.953	0.958	0.960	0.955	0.950	0.945	0.945
8	**0.002**	0.943	0.957	0.966	0.972	0.972	0.967	0.960	0.952	0.952
9	**-0.900**	0.952	0.965	0.967	0.964	0.960	0.957	0.950	0.940	0.940
10	**-0.883**	0.961	0.974	0.976	0.976	0.975	0.975	0.975	0.974	0.974
11	**-0.866**	0.956	0.973	0.977	0.978	0.976	0.973	0.968	0.961	0.961
12	**-0.923**	0.969	0.970	0.972	0.970	0.961	0.955	0.945	0.931	0.931
13	**0.930**	0.923	0.947	0.959	0.964	0.966	0.966	0.965	0.963	0.963
14	**-0.755**	0.889	0.918	0.922	0.918	0.911	0.903	0.897	0.895	0.895
15	**-0.890**	0.936	0.951	0.957	0.961	0.963	0.964	0.963	0.961	0.961
16	**-0.835**	0.940	0.943	0.933	0.919	0.908	0.899	0.895	0.884	0.884
17	**-0.916**	0.969	0.979	0.982	0.980	0.976	0.971	0.968	0.968	0.968
18	**-0.782**	0.931	0.952	0.964	0.969	0.971	0.972	0.972	0.971	0.971
19	**-0.915**	0.951	0.973	0.980	0.979	0.975	0.966	0.952	0.940	0.940
20	**-0.729**	0.870	0.919	0.930	0.937	0.940	0.941	0.940	0.939	0.939
21	**-0.903**	0.940	0.963	0.975	0.978	0.976	0.970	0.962	0.953	0.953
22	**-0.743**	0.894	0.922	0.935	0.941	0.942	0.939	0.936	0.932	0.932
23	**-0.885**	0.923	0.934	0.938	0.945	0.951	0.956	0.958	0.954	0.954
24	**-0.882**	0.932	0.956	0.963	0.964	0.963	0.963	0.963	0.965	0.965
Median	**-0.851**	0.934	0.952	0.963	0.966	0.963	0.960	0.955	0.949	0.944

Note: Entries in boldface are below the one-tailed critical value of 0.685 commonly used for objective image quality. Alongside the PSNR, the SCG4 metric is also calculated for each of the demosaiced images.

Given these two metrics, the goal is to establish a significant correlation between them. To accomplish this, the Pearson correlation coefficient is used to analyze the degree of correlation between the two variables (PSNR and SCG4 metric). The resulting correlation coefficients are shown in Table 1. As there are 11 methods used to evaluate the correlation, there are 9 degrees of freedom (DoF) corresponding to a critical value of 0.685 for a p-value of 0.01. Apart from the correlation values when the significance threshold is set to 1, the metric is shown to be statistically correlated to the PSNR value. Beyond simply establishing the correlation, we also show that a threshold value of 5 results in the best median correlation for the *Kodak* database.

VI. CONCLUSION

In this work, we developed the Sparsity of Chroma Gradient in Four Directions (SCG4) metric that can be used to predict the quality of a demosaiced image obtained through a Bayer color filter array. We demonstrated certain properties of natural images particularly in relation to the chroma components and utilized these properties as a prior for our proposed metric. Experimental results show a strong correlation between SCG4 and both the subjective and objective quality of the resulting demosaiced image.

An interesting aspect of our proposed metric is its highly localized nature. Each probability estimate utilizes only 10 pixels—the central pixels for C1 and C2, and the four neighboring pixels for each of these components. This not only allows for efficient calculation but also increases the flexibility of the metric. While the smallest feasible area at which the metric can operate has yet to be investigated, it has been shown to operate with relatively small areas such as the cropped region from the Lighthouse image.

The development of such a metric, in itself, is useful in further studies of demosaicing because it can enable adaptivity in many aspects. For instance, one may develop a compound demosaicing technique that switches between various demosaicing algorithms depending on the performance on a local region in an image. More importantly, it can be used as an optimization criteria for a single demosaicing technique. Further work is being done into developing such an adaptive algorithm.

REFERENCES

[1] L. Lindgren, J. Melander, R. Johansson, and B. Moller, “A multiresolution 100-GOPS 4-Gpixels/s programmable smart vision sensor for multisense imaging,” *IEEE Journal of Solid-State Circuits*, vol. 40, no. 6, pp. 1350–1359, June 2005.

[2] T. Yamashita, R. Funatsu, T. Yanagi, K. Mitani, Y. Nojiri, and T. Yoshida, “A camera system using three 33-megapixel CMOS image sensors for UHDTV2,” *SMPTE Motion Imaging Journal*, vol. 120, no. 8, pp. 24–31, Nov 2011.

[3] J. Galt, B. Petljanski, and S. Campbell, “Development of a 70mm, 25-megapixel electronic cinematography camera with integrated flash recorder,” *SMPTE Motion Imaging Journal*, vol. 122, no. 2, pp. 42–48, March 2013.

[4] S. Kleinfelder, S. Lim, X. Liu, and A. E. Gamal, “A 10000 frames/s CMOS digital pixel sensor,” *IEEE Journal of Solid-State Circuits*, vol. 36, no. 12, pp. 2049–2059, Dec 2001.

[5] R. Berger, D. D. Rathman, B. M. Tyrrell, E. J. Kohler, M. K. Rose, R. A. Murphy, T. S. Perry, H. F. Robey, F. A. Weber, D. M. Craig, A. M. Soares, S. P. Vernon, and R. K. Reich, “A 64 64-pixel CMOS test chip for the development of large-format ultra-high-speed snapshot imagers,” *IEEE Journal of Solid-State Circuits*, vol. 43, no. 9, pp. 1940–1950, Sept 2008.

[6] R. Xu, B. Liu, and J. Yuan, “A 1500 fps highly sensitive 256 256 CMOS imaging sensor with in-pixel calibration,” *IEEE Journal of Solid-State Circuits*, vol. 47, no. 6, pp. 1408–1418, June 2012.

[7] M. Furuta, Y. Nishikawa, T. Inoue, and S. Kawahito, “A high-speed, high-sensitivity digital CMOS image sensor with a global shutter and 12-bit column-parallel cyclic A/D converters,” *IEEE Journal of Solid-State Circuits*, vol. 42, no. 4, pp. 766–774, April 2007.

[8] T. Takemoto, H. Yamashita, T. Yazaki, N. Chujo, Y. Lee, and Y. Matsuoka, “A 25-to-28 Gb/s high-sensitivity (-9.7 dBm) 65 nm CMOS optical receiver for board-to-board interconnects,” *IEEE Journal of Solid-State Circuits*, vol. 49, no. 10, pp. 2259–2276, Oct 2014.

[9] B. E. Bayer, “Color imaging array,” July 1976, US Patent 3,971,065.

[10] “Kodak lossless true color image suite,” 2013, Retrieved 2016-11. [Online]. Available: http://r0k.us/graphics/kodak/ [11] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau, “Demosaicking: color filter array interpolation,” *IEEE Signal Processing Magazine*, vol. 22, no. 1, pp. 44–54, Jan 2005.

[12] B. Leung, G. Jeon, and E. Dubois, “Least-squares lumachroma demultiplexing algorithm for bayer demosaicking,” *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 1885–1894, July 2011.

[13] R. Kimmel, “Demosaicing: image reconstruction from color CCD samples,” *IEEE Transactions on Image Processing*, vol. 8, no. 9, pp. 1221–1228, Sep 1999.

[14] J. E. Adams, Jr., “Interactions between color plane interpolation and other image processing functions in electronic photography,” *Proc. SPIE*, vol. 2416, pp. 144–151, 1995. [Online]. Available: http://dx.doi.org/10.1117/12.204825

[15] S.-C. Pei and I.-K. Tam, "Effective color interpolation in CCD color filter array using signal correlation," in *Proceedings 2000 International Conference on Image Processing* (Cat. No.00CH37101), vol. 3, 2000, pp. 488–491 vol.3.

[16] B. K. Gunturk, Y. Altunbasak, and R. M. Mersereau, "Color plane interpolation using alternating projections," *IEEE Transactions on Image Processing*, vol. 11, no. 9, pp. 997–1013, Sep 2002.

[17] D. Kiku, Y. Monno, M. Tanaka, and M. Okutomi, "Beyond color difference: Residual interpolation for color image demosaicking," *IEEE Transactions on Image Processing*, vol. 25, no. 3, pp. 1288–1300, March 2016.

[18] T. Yu, W. Hu, W. Xue, and W. Zhang, "Colour image demosaicking via joint intra and inter channel information," Electronics Letters, vol. 52, no. 8, pp. 605–607, 2016. [19] K. Hirakawa and T. W. Parks, "Adaptive homogeneitydirected demosaicing algorithm," *IEEE Transactions on Image Processing*, vol. 14, no. 3, pp. 360–369, March 2005.

[20] P. Getreuer, "Gunturk–Altunbasak–Mersereau alternating projections image demosaicking," *Image Processing On Line*, 2011.

[21] ——, "Color demosaicing with contour stencils," in 2011 *17th International Conference on Digital Signal Processing (DSP)*, July 2011, pp. 1–6.

[22] ——, "Image demosaicking with contour stencils," *Image Processing On Line*, vol. 2, pp. 22–34, 2012.

[23] L. Zhang and X. Wu, "Color demosaicking via directional linear minimum mean square-error estimation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2167–2178, Dec 2005.

[24] P. Getreuer, "Zhang-Wu directional LMMSE image demosaicking," *Image Processing On Line*, 2011.

[25] H. S. Malvar, L. wei He, and R. Cutler, "High-quality linear interpolation for demosaicing of bayer-patterned color images," in 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 3, May 2004, pp. iii–485–8 vol.3.

[26] P. Getreuer, "Malvar-He-Cutler linear image demosaicking," *Image Processing On Line*, vol. 1, 2011.

[27] X. Li, "Demosaicing by successive approximation," *IEEE Transactions on Image Processing*, vol. 14, no. 3, pp. 370–379, March 2005.