Frequency-Domain Methods for Demosaicking of Bayer-Sampled Color Images

Eric Dubois, *Fellow, IEEE*

Abstract—This letter presents a new and simplified derivation of the frequency-domain representation of color images sampled with the Bayer color filter array. Two new demosaicking algorithms based on the frequency-domain representation are described and shown to give excellent results.

Index Terms—Bayer sampling, color filter array, demosaicking, digital cameras, interpolation.

I. INTRODUCTION

MOST digital color cameras use a single CCD sensor with a color filter array (CFA). Typically, estimates of red (R), green (G), and blue (B) tristimulus values are obtained on three interleaved sampling structures whose union forms the CFA sampling structure, as shown in Fig. 1. A postprocessor must interpolate the three color planes to obtain R, G, and B samples at each location of the CFA sampling structure; this process is referred to as *demosaicking*. It is possible to separately interpolate each color plane from other samples of the same color using linear interpolating filters, but this generally yields unsatisfactory results. Thus, numerous methods have been devised to exploit the correlation between the color planes to give improved results. Gunturk *et al.* [1] recently reviewed and compared a large number of such methods. Interestingly, most of the techniques reviewed were characterized as "heuristic methods." The method of projections onto convex sets (POCS) [2] gave the best performance among those compared in [1] in terms of objective performance measures and subjective quality and is usually considered to be the benchmark against which other methods are compared. Virtually all methods suffer from characteristic distortions in critical areas of the images: areas containing high luminance detail produce false colors, while sharp color transitions introduce false luminance patterns.

Recently, Alleysson *et al.* [3] elegantly showed that the CFA image can be represented as the combination of an achromatic or *luma* component at baseband and two *chrominance* components modulated at high spatial frequency. This is analogous to the spatial-frequency-domain component multiplexing used in analog television standards such as NTSC and PAL [4]. These color television systems introduce very similar distortions, commonly referred to as cross-color and cross-luminance. In the
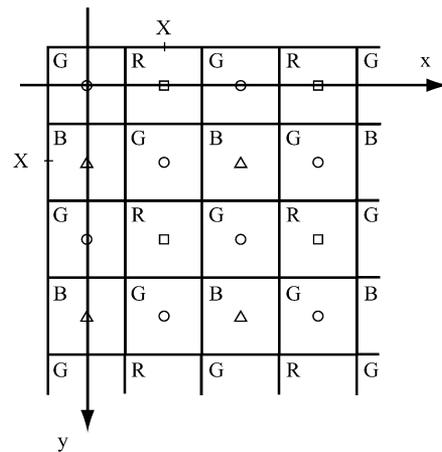
Fig. 1. Upper left portion of the Bayer CFA sampling structure, showing the constituent sampling structures $\Lambda_R(\square)$, $\Lambda_G(\bigcirc)$, and $\Lambda_B(\triangle)$.

case of color television, the luma and chrominance components are available prior to multiplexing and can thus be spatially prefiltered to avoid spectral overlap upon multiplexing, significantly reducing these component crosstalk effects [4]. In the case of the CFA signal, this is not possible, and thus, component crosstalk is unavoidable. CFA demultiplexing methods must try to exploit signal characteristics and correlations as much as possible to obtain the best signal quality. This letter shows that the frequency-domain approach of [3] can yield new insights into the demosaicking problem and indeed lead to new algorithms as good as or better than other techniques. I first present a new and simplified derivation of the frequency-domain representation of the CFA signal and then describe two novel demosaicking algorithms based on this representation.

II. FREQUENCY-DOMAIN REPRESENTATION OF THE CFA SIGNAL

Let $f_{\mathrm{CFA}}[n_1, n_2]$ represent the output of the CFA sensor, sampled on a rectangular sampling lattice Λ. The horizontal and vertical sample spacings are equal, and this sample spacing X is used as the unit of length, called the pixel height (px). Thus, the lattice Λ is simply the integer lattice $\mathbb{Z}^2$. In the Bayer CFA, Λ is partitioned into three disjoint subsets that are each shifted sublattices of Λ, referred to as Λ_R, Λ_G, and Λ_B, illustrated in Fig. 1. The origin is set at the upper left corner of the image on a green sample with a red sample to the right and a blue sample below.

Assume that there exist underlying R, G and B component signals defined on Λ that we are trying to estimate, denoted $f_R[n_1, n_2], f_G[n_1, n_2]$, and $f_B[n_1, n_2]$, respectively. The CFA

signal is obtained by subsampling $f_i[n_1, n_2]$ on Λ_i and adding the three subsampled signals. The subsampling can be represented as multiplication by subsampling functions $m_i[n_1, n_2]$ that take the value 1 on Λ_i and zero otherwise. The specific subsampling functions for the Bayer CFA can be expressed simply as follows:

$$m_G[n_1, n_2] = \tfrac{1}{2}\left(1 + (-1)^{n_1+n_2}\right) \tag{1}$$

$$m_R[n_1, n_2] = \tfrac{1}{4}\left(1 - (-1)^{n_1}\right)\left(1 + (-1)^{n_2}\right) \tag{2}$$

$$m_B[n_1, n_2] = \tfrac{1}{4}\left(1 + (-1)^{n_1}\right)\left(1 - (-1)^{n_2}\right). \tag{3}$$

The CFA signal can then be expressed as

$$
\begin{aligned}
f_{\mathrm{CFA}}[n_1, n_2] &= \sum_{i \in \{R,G,B\}} f_i[n_1, n_2] m_i[n_1, n_2] \\
&= \tfrac{1}{2} f_G[n_1, n_2](1 + (-1)^{n_1+n_2}) \\
&\quad + \tfrac{1}{4} f_R[n_1, n_2](1 - (-1)^{n_1})(1 + (-1)^{n_2}) \\
&\quad + \tfrac{1}{4} f_B[n_1, n_2](1 + (-1)^{n_1})(1 - (-1)^{n_2}).
\end{aligned} \tag{4}
$$

This can be rearranged as

$$
\begin{aligned}
&f_{\mathrm{CFA}}[n_1, n_2] \\
&= \left(\tfrac{1}{4} f_R[n_1, n_2] + \tfrac{1}{2} f_G[n_1, n_2] + \tfrac{1}{4} f_B[n_1, n_2]\right) \\
&\quad + \left(-\tfrac{1}{4} f_R[n_1, n_2] + \tfrac{1}{2} f_G[n_1, n_2] - \tfrac{1}{4} f_B[n_1, n_2]\right) \\
&\quad\quad \times (-1)^{n_1+n_2} \\
&\quad + \left(-\tfrac{1}{4} f_R[n_1, n_2] + \tfrac{1}{4} f_B[n_1, n_2]\right)\left((-1)^{n_1} - (-1)^{n_2}\right) \\
&\triangleq f_L[n_1, n_2] + f_{C1}[n_1, n_2](-1)^{n_1+n_2} \\
&\quad + f_{C2}[n_1, n_2]\left((-1)^{n_1} - (-1)^{n_2}\right).
\end{aligned} \tag{5}
$$

Noting that $-1 = \exp(j\pi)$, this can be rewritten as

$$
\begin{aligned}
&f_{\mathrm{CFA}}[n_1, n_2] \\
&= f_L[n_1, n_2] + f_{C1}[n_1, n_2]\exp(j2\pi(n_1 + n_2)/2) \\
&\quad + f_{C2}[n_1, n_2](\exp(j2\pi n_1/2) - \exp(j2\pi n_2/2)).
\end{aligned} \tag{6}
$$

This can be interpreted as a baseband luma component f_L, a chrominance component f_{C1} modulated at the spatial frequency $(0.5, 0.5)$, and a second chrominance component f_{C2} modulated at the two spatial frequencies $(0.5, 0)$ and $(0, 0.5)$, where spatial frequencies are expressed in cycles per pixel height (c/px). In matrix notation, the relationship between the luma/chrominance components and the RGB components is given by

$$
\begin{bmatrix} f_L \\ f_{C1} \\ f_{C2} \end{bmatrix} = \begin{bmatrix} \tfrac{1}{4} & \tfrac{1}{2} & \tfrac{1}{4} \\ -\tfrac{1}{4} & \tfrac{1}{2} & -\tfrac{1}{4} \\ -\tfrac{1}{4} & 0 & \tfrac{1}{4} \end{bmatrix} \begin{bmatrix} f_R \\ f_G \\ f_B \end{bmatrix} \tag{7}
$$

$$
\begin{bmatrix} f_R \\ f_G \\ f_B \end{bmatrix} = \begin{bmatrix} 1 & -1 & -2 \\ 1 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} f_L \\ f_{C1} \\ f_{C2} \end{bmatrix}. \tag{8}
$$

Taking the Fourier transform of (6), we obtain

$$
\begin{aligned}
F_{\mathrm{CFA}}(u, v) &= F_L(u, v) + F_{C1}(u - 0.5, v - 0.5) \\
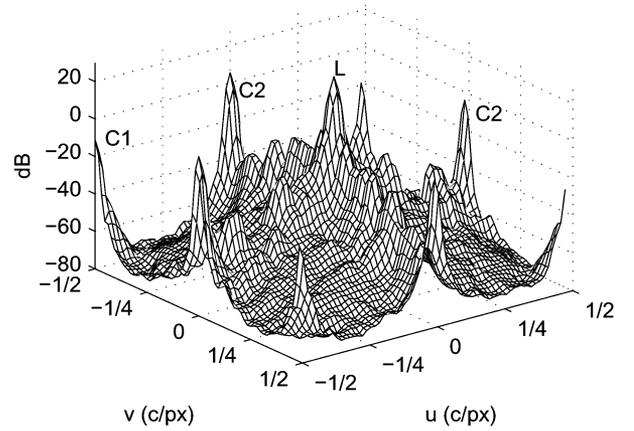&\quad + F_{C2}(u - 0.5, v) - F_{C2}(u, v - 0.5).
\end{aligned} \tag{9}
$$

Fig. 2.   Power spectral density estimate of the lighthouse CFA image.

Note that for any achromatic signal in which $f_R = f_G = f_B$, the two chrominance components are identically zero. Typically, the chrominance components will have lower energy and spatial bandwidth than the baseband luma component. This is illustrated in Fig. 2, which shows an estimate of the power spectral density of the CFA signal for the standard "lighthouse" image. From this figure, we see that there is interference or crosstalk between components and that the crosstalk is strongest between the luma f_L and the modulated chrominance component f_{C2}.

III. FREQUENCY-DOMAIN DEMOSAICKING ALGORITHMS

From the spectrum of Fig. 2, it appears that we could recover the luma component from the CFA signal using a lowpass filter and the two chrominance components using appropriate bandpass filters. The R, G, and B signals could then be recovered from these components using the inverse transformation (8). The algorithm in [3] uses a different, indirect approach, separating the luma from the sum of the modulated chrominance components with a pair of complementary lowpass/highpass filters. The chrominance term is subsampled on each of the sampling structures Λ_i (which is like a demodulation), and then each of these is interpolated back to Λ to give three color difference signals of the form $f_R - f_L, f_G - f_L, f_B - f_L$, which can be added to f_L to recover f_R, f_G, and f_B. Both of these approaches suffer from the significant spectral overlap of the luma and the modulated $C2$ components. However, by observing that there are *two* separate modulated copies of the chrominance component C2, each having different spectral overlap with the luma component, it is possible to significantly reduce the impact of this spectral crosstalk. This observation is the main contribution of this letter, and two methods that exploit it are proposed.

A. Demosaicking Using Complementary Asymmetric Filters for C2

Let the modulated C2 components at frequencies $(0.5, 0.0)$ and $(0.0, 0.5)$ be denoted as $f_{C2\mathrm{ma}}[n_1, n_2] = f_{C2}[n_1, n_2](-1)^{n_1}$ and $f_{C2\mathrm{mb}}[n_1, n_2] = -f_{C2}[n_1, n_2](-1)^{n_2}$, respectively. Note that the high horizontal frequencies in $f_{C2\mathrm{ma}}$ mainly overlap with high horizontal frequencies of f_L, while
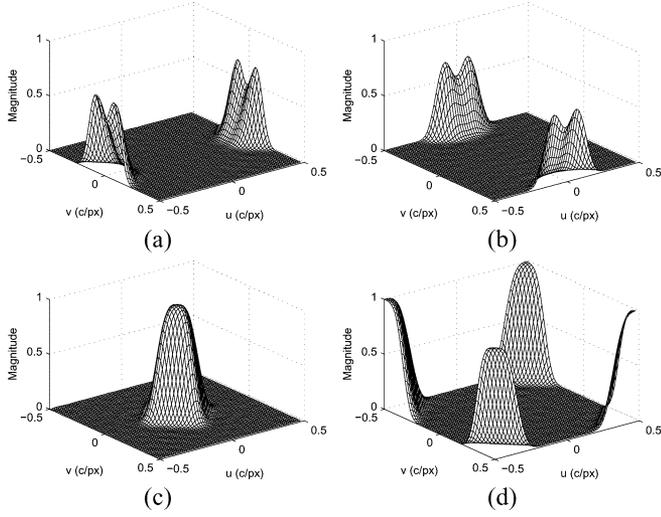
Fig. 3. Frequency response of filters used in the asymmetric filtering method. (a) h_{2a}. (b) h_{2b}. (c) h_{LP}. (d) h_1.

Fig. 4. Ideal frequency response used to design filters with the window method. (a) Contours. (b) Perspective view.

the high vertical frequencies in f_{C2mb} mainly overlap with high vertical frequencies of f_L. The vertical frequencies in f_{C2ma} are mostly crosstalk-free, as are the horizontal frequencies in f_{C2mb}. Thus, we would like to recover the vertical frequency components of f_{C2} mainly from f_{C2ma} and the horizontal components from f_{C2mb}. This can be accomplished using asymmetric filters (with respect to horizontal and vertical frequencies) to extract f_{C2ma} and f_{C2mb} from f_{CFA}, such that together, the entire passband of f_{C2} is recovered. Fig. 3(a) and (b) shows such a pair of filters, where Fig. 3(c) shows the sum of the two filter responses when their passbands are shifted to baseband, denoted h_{LP}. The demosaicking algorithm is then as follows.

1) Filter f_{CFA} with a bandpass filter h_1 centered at frequency $(0.5, 0.5)$ to extract $\hat{f}_{C1m} = f_{CFA} * h_1$, and shift it to baseband to estimate $\hat{f}_{C1}[n_1, n_2] = \hat{f}_{C1m}[n_1, n_2](-1)^{n_1+n_2}$. Fig. 3(d) shows the frequency response of a suitable h_1.
2) Filter f_{CFA} with h_{2a} to get $\hat{f}_{C2ma} = f_{CFA} * h_{2a}$, and demodulate to baseband $\hat{f}_{C2a}[n_1, n_2] = \hat{f}_{C2ma}[n_1, n_2](-1)^{n_1}$. Similarly, $\hat{f}_{C2mb} = f_{CFA} * h_{2b}$ and $\hat{f}_{C2b}[n_1, n_2] = -\hat{f}_{C2mb}[n_1, n_2](-1)^{n_2}$. Their sum forms the estimate $\hat{f}_{C2} = \hat{f}_{C2a} + \hat{f}_{C2b}$.
3) Estimate the luma by $\hat{f}_L[n_1, n_2] = f_{CFA}[n_1, n_2] - \hat{f}_{C1m}[n_1, n_2] - \hat{f}_{C2}[n_1, n_2]((-1)^{n_1} - (-1)^{n_2})$.
4) Estimate the RGB components $\hat{f}_R, \hat{f}_G, \hat{f}_B$ from $\hat{f}_L, \hat{f}_{C1}$, and $\hat{f}_{C2}$ using (8).

B. Adaptive Frequency-Domain Demosaicking Algorithm

The most visible artifacts in demosaicked images tend to be cross-color caused by luma energy near the modulation frequencies $(0.5, 0.0)$ and $(0.0, 0.5)$. Neither the original algorithm of [3] nor the fixed frequency-domain method of the previous section can disambiguate the high-frequency luma in this region from the baseband chrominance, leading to very visible rainbow artifacts such as the one frequently shown on the picket fence in the lighthouse image [1, Fig. 10]. However, in any particular local image region, the spectral overlap most often occurs in
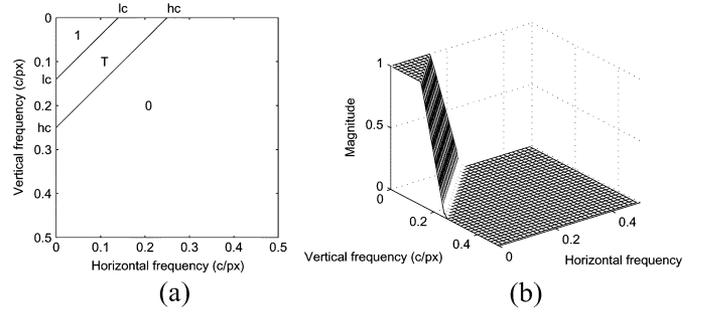
only one of the two frequency directions, and these are the situations where the distortion is most visible. Thus, a possible approach is to separately form two estimates of the C2 component with the full bandwidth, one from each of f_{C2ma} and f_{C2mb}, and to form the final estimate of C2 as a *locally adaptive* weighted sum of these, giving more weight to that component assumed to suffer less from crosstalk. To do this, we can monitor the average local energy of the CFA signal in two frequency bands along the horizontal and vertical frequency axes, respectively, denoted e_X and e_Y, centered at frequencies $(f_m, 0)$ and $(0, f_m)$. Comparing the two average energies, the larger one would be assumed to suffer from greater crosstalk and thus be given less weight in forming the C2 estimate. The resulting algorithm is thus very similar to that of the previous section. However, the filters h_{2a} and h_{2b} now have the same passband shape and can be obtained by modulating a filter like that of Fig. 3(c) to the frequencies $(0.5, 0.0)$ and $(0.0, 0.5)$, respectively. The estimate of C2 is obtained as $\hat{f}_{C2}[n_1, n_2] = w[n_1, n_2]\hat{f}_{C2a}[n_1, n_2] + (1 - w[n_1, n_2])\hat{f}_{C2b}[n_1, n_2]$. In my implementation, I used the weighting coefficient $w = e_Y/(e_X + e_Y)$. The average energies e_X and e_Y were estimated using modulated Gaussian filters with standard deviation of 3.5 pixels, centered at frequencies $(0.375, 0.0)$ and $(0.0, 0.375)$ c/px, respectively, followed by smoothing of the squared output with a 5-by-5 moving average filter. The filter h_2 was obtained using the window method to design a lowpass filter approximating the piecewise planar ideal response shown in Fig. 4, with $lc = 0.14$ and $hc = 0.25$, and multiplying the resulting unit-sample response by $(-1)^{n_1+n_2}$ to shift it to the center frequency $(0.5, 0.5)$. The filters h_{2a} and h_{2b} were designed in the same way but with $lc = 0.11$ and $hc = 0.20$ and multiplying by $(-1)^{n_1}$ and $(-1)^{n_2}$ to shift to the center frequencies $(0.5, 0.0)$ and $(0.0, 0.5)$, respectively. These cutoff values were roughly optimized by trial and error to give the best average performance.

IV. RESULTS

The proposed algorithms were tested on the 24 Kodak images used for the comparison in [1]. They were compared with the frequency-domain algorithm of Alleysson *et al.* [3] and the POCS algorithm of Gunturk *et al.* [2]. The latter algorithm was used as a benchmark, since it gave the lowest mean-square error (MSE) values in [1]. It was used with one level of decomposition, eight iterations, and a threshold of 0.0, as described in [2].

TABLE I
MSE FOR R, G, AND B COMPONENTS DEMOSAICKED USING FOUR METHODS.
(A) POCS [2]. (B) FREQUENCY-DOMAIN METHOD OF [3]. (C) FREQUENCY
DOMAIN WITH ASYMMETRIC FILTERS. (D) ADAPTIVE FREQUENCY DOMAIN

Image		Method			
		A	B	C	D
Lighthouse	R	7.85	15.07	11.72	**7.31**
	G	3.55	5.07	4.66	**3.06**
	B	8.07	14.74	12.43	**7.95**
Sails	R	4.71	7.54	5.49	**4.63**
	G	2.58	3.42	2.63	**2.16**
	B	5.79	8.76	6.44	**5.47**
Statue	R	**5.47**	5.74	5.85	5.57
	G	2.95	2.89	2.90	**2.68**
	B	6.45	7.03	6.64	**6.15**
Window	R	4.92	7.56	4.59	**4.58**
	G	2.60	3.41	2.17	**2.06**
	B	5.78	8.68	**5.39**	5.42
Average	R	11.68	15.65	12.59	**10.79**
	G	5.39	6.06	5.34	**4.60**
	B	11.53	15.48	13.01	**11.15**

Table I shows the MSE for the R, G and B components of four selected images, referred to as "lighthouse," "sails," "statue," and "window" in [3] and as images 16, 7, 14, and 5, respectively, in [2], as well as the average MSE over the 24 images. Different authors have used different version of these images, so these numerical results should not be directly compared with those in other published papers. A border of six pixels around the image was excluded from the MSE computations to reduce the impact of edge effects. The lowest MSE in each row is highlighted in bold. The adaptive frequency-domain algorithm gave the lowest *average* MSE for all components and gave the lowest MSE for all the components of the four test images except for two. The nonadaptive method using asymmetric filters gave a marked improvement over the method of Alleysson *et al.* for most images and on average but was not able to match the performance of the POCS method. Of course, a full evaluation requires the assessment of the subjective quality of the demosaicked images. Preliminary results show that the adaptive frequency-domain method gives excellent subjective performance. For example, the familiar crosstalk effects in the "lighthouse" image are virtually eliminated, showing a very clear improvement over the POCS result. All 24 images demosaicked with the four methods can be found in [5].

V. CONCLUSION

This letter has presented a new and simplified derivation of the frequency-domain representation of Bayer-sampled color images. The frequency-domain formulation has inspired two new demosaicking algorithms with excellent performance; the adaptive algorithm gave lower MSE values than other published techniques. There was little systematic attempt made to optimize the filters used in this letter, which were all designed with the window method and were of high order (21 by 21). The new methods involve a single pass through the image and appear to be considerably faster than the iterative method of [2] based on software implementation (by a factor of more than five). However, since I believe that the system can be further simplified, a detailed complexity analysis is premature. Further work will seek filters of lowest order that optimize subjective quality. A more rigorous approach to select the weighting coefficients in the adaptive frequency-domain method will also be investigated.

REFERENCES

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

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

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

[4] E. Dubois and W. F. Schreiber, "Improvements to NTSC by multidimensional filtering," *SMPTE J.*, vol. 97, pp. 504–511, Jun. 1988.

[5] E. Dubois. (2005) Frequency Domain Methods for Demosaicking of Bayer-Sampled Color Images: Additional Results. [Online]. Available: http://www.site.uottawa.ca/~edubois/demosaicking/