HARDWARE-FRIENDLY UNIVERSAL DEMOSAICK USING NON-ITERATIVE MAP RECONSTRUCTION

Hasib Siddiqui, Kalin Atanassov, Sergio Goma

Qualcomm Technologies, Inc., San Diego, CA

ABSTRACT

Non-Bayer color filter array (CFA) sensors have recently drawn attention due to their superior compression of spectral energy, ability to deliver improved signal-to-noise ratio, or ability to provide high dynamic range (HDR) imaging. Demosaicking methods that perform color interpolation of Bayer CFA data have been widely investigated. However, a bottleneck to the adaption of emerging non-Bayer CFA sensors is the unavailability of efficient color-interpolation algorithms that can demosaick the new patterns. Designing a new demosaick algorithm for every proposed CFA pattern is a challenge. In this paper, we propose a hardware-friendly *universal demosaick* algorithm based on *maximum a-posteriori* (MAP) estimation that can be configured to demosaick raw images captured using a variety of CFA sensors.

The forward process of mosaicking is modeled as a linear operation. We then use quadratic data-fitting and image prior terms in a MAP framework and pre-compute the inverse matrix for recovering the full RGB image from CFA observations for a given pattern. The pre-computed inverse is later used in real-time application to demosaick the given CFA pattern. The inverse matrix is observed to have a Toeplitz-like structure, allowing for hardware-efficient implementation of the algorithm.

We use a set of 24 Kodak color images to evaluate the quality of our demosaick algorithm on three different CFA patterns. The PSNR values of the reconstructed full-channel RGB images from CFA samples are reported in the paper.

Index Terms— Color filter array, Bayer mosaic, universal demosaick, statistical MAP estimation.

1. INTRODUCTION

Bayer pattern is no longer the single dominant pattern in sensor industry. Multiple CFA patterns have recently gained popularity because of their superior spectral-compression performance [1], improved signal-to-noise ratio [2], or ability to provide HDR imaging. Traditionally, raw data captured using a non-Bayer CFA sensor is either (1) resampled to a Bayer pattern and then color-interpolated using Bayer demosaick [3, 4] or (2) processed using a dedicated demosaick algorithm designed for the specific color sampling pattern. The former solution results in sub-optimal interpolation, while the latter solution is hardware inefficient, necessitating implementation of a separate demosaick block in the image signal processor (ISP) to support color-interpolation of a specific CFA raw data. Universal demosaick proposed herein is based on statistical maximum a-posteriori (MAP) framework [5], which can be configured to demosaick raw data acquired from any 3-channel CFA sensor fitted at the front end of the camera.

Demosaick algorithms that rely on iterative MAP reconstruction to estimate the unknown RGB color values at each pixel location given the observed CFA data have already been proposed in the literature [6]. By modeling the forward CFA mosaicking as a linear operation and invoking image smoothness constraints, the demosaicked image given any color subsampling can be obtained by solving a set of algebraic equations. However, traditional algebraic methods are hardware-unfriendly and are not suitable for real-time applications.

Our proposed approach involves pre-computing and storing the inverse operator required for MAP reconstruction of the 3-channel RGB image. The reconstruction operator is learned offline for a given CFA pattern and then approximated using a set of compact 2D spatial filters. In order to ensure data-independent reconstruction, we use quadratic models for both *data-fitness* and *prior model* terms. Since, quadratic image priors lead to blurring of object boundaries, we seek two independent reconstructions: one enforcing horizontal smoothness and the other vertical smoothness in the recovered image. The images obtained from the two directional reconstructions are then combined using homogeneity-directed classifier as proposed by Hirakawa [3].

We present our demosaick results on the three different CFA patterns shown in Figure 1: (a) Traditional 2×2 Bayer. (b) Modified 3×3 Bayer [7] with color filters 1.5 times the pixel size. There are 6 spectral components in this 3×3 pattern, namely R, G, B, $Y = (R + G)/2$; $C = (B + G)/2$; $W = (R + 2G + B)/4$. This format can potentially be used when pixel sizes are in the sub-micron range; fabrication of color filters with equally small sizes is not practical with the current technology. (c) 4×4 Lukac [1]. The CFA images are simulated by spectral sub-sampling of the set of 24 standard Kodak images according to the patterns shown in Figure 1. The full-channel RGB images are then reconstructed using the proposed algorithm and the reconstruction PSNRs are calculated.

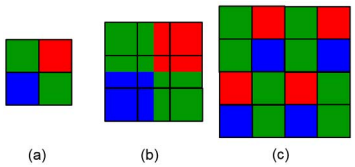


Fig. 1. CFA patterns considered for analysis: (a) 2×2 Bayer; (b) Modified 3×3 Bayer [7] ; (c) 4×4 Lukac [1].

2. UNIVERSAL DEMOSAICK

2.1. Non-Iterative MAP Estimation

To introduce our proposed method, we shall assume that $x \in \mathbb{R}^{3N}$ denotes the full 3-channel RGB image to be reconstructed in lexicographical order; $y \in \mathbb{R}^N$ denotes samples captured using a CFA sensor; $A \in \mathbb{R}^{N \times 3N}$ denotes the spectral-compression operator that performs the CFA color sub-sampling [8]; and $n \in \mathbb{R}^N$ denotes the additive, zero-mean, i.i.d. Gaussian noise with covariance matrix $\Lambda = E\left[n^T n\right] = \sigma^2 I$. Then, the forward model is given by

$$y = Ax + n. \tag{1}$$

Let $T_l, T_{c_1}, T_{c_2} \in \mathbb{R}^{N \times 3N}$ be linear transformations that transform x into its de-correlated luminance/chrominance components: $l \triangleq T_l x, c_1 \triangleq T_{c_1} x, c_2 \triangleq T_{c_2} x \in \mathbb{R}^N$. Further, we use homogeneous Gaussian Markov random field (GMRF) [5] prior models for the opponent color components l, c_1, and c_2, assuming the three GMRFs share a common non-causal prediction matrix $G_{i,j} = g_{i-j}$, but have different non-causal prediction variances. Specifically, we assume the non-causal prediction variances for the luminance and chrominance components are σ_l^2 and σ_c^2 respectively. The MAP estimate of x given the observations y is then given by

$$\begin{aligned} \hat{x} &= \arg\min_{x \in \mathbb{R}^{3N}} \left\{ \frac{1}{2\sigma^2}\|y - Ax\|^2 + \frac{1}{2\sigma_l^2} x^T {T_l}^T B T_l x \right. \\ &+ \left. \frac{1}{2\sigma_c^2} \sum_{k=1,2} x^T T_{c_k}^T B T_{c_k} x \right\}, \end{aligned} \tag{2}$$

where $B \in \mathbb{R}^{3N \times 3N}$ is a symmetric invertible matrix. The entries of B are given by $B_{i,j} = \delta_{i-j} - g_{i-j}$, where δ_k is the Kronecker delta function. The ratios σ/σ_l and σ/σ_c control the trade-off between fitting to the data and the smoothness of luminance/chrominance components of the solution.

The MAP reconstruction of (2) can be computed in closed form as follows:

$$\hat{x} = \left(A^T A + \frac{\sigma^2}{\sigma_l^2} {T_l}^T B T_l + \frac{\sigma^2}{\sigma_c^2} \sum_{k=1,2} T_{c_k}^T B T_{c_k} \right)^{-1} A^T y. \tag{3}$$

Therefore, if we pre-compute the inverse matrix for a given CFA pattern:

$$H \triangleq \left(A^T A + \frac{\sigma^2}{\sigma_l^2} {T_l}^T B T_l + \frac{\sigma^2}{\sigma_c^2} \sum_{k=1,2} T_{c_k}^T B T_{c_k} \right)^{-1} A^T, \tag{4}$$

then we may reconstruct the full 3-channel RGB image x from the CFA observations y by simply computing the matrix-vector product:

$$\hat{x} = Hy. \tag{5}$$

It is important to note that H is independent of the sensor data y and needs to be computed only once for a given CFA pattern. While the non-iterative MAP (NMAP) computation in (5) seems appealing, it is rarely used in practice. In general, H is enormous and non-sparse for most problems and, therefore, iterative solutions of (2) are more efficient.

As we discuss in a later section, for the demosaick problem the inverse matrix H is observed to be sparse and the rows of H repeat with a periodicity of $3MN$, where $M \times N$ is support of a period of the CFA pattern. This special structure of H allows us to efficiently implement computation for NMAP estimation (5) for real-time applications.

0	0	0	0	0
0	0	0	0	0
-1/6	2/3	0	2/3	-1/6
0	0	0	0	0
0	0	0	0	0

Fig. 2. Non-causal prediction matrix $g_k^{(H)}$ used in the Gaussian Markov Random Field (GMRF) directional image prior.

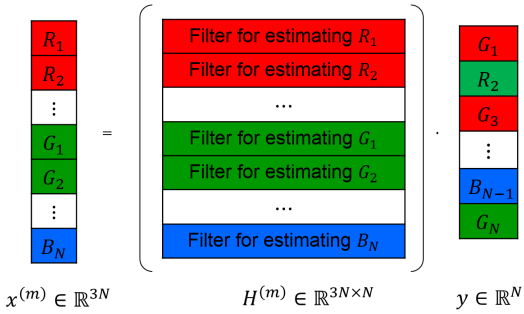


Fig. 3. Demosaicking using non-iterative MAP: $\hat{x}^{(m)}$ is the full 3-channel reconstructed image, $H^{(m)}$ is the reconstruction matrix; $\hat{y}$ is the CFA data.

2.2. Integrating Edge Information

A homogeneous GMRF prior model, used in the afore-stated modeling, assumes the local density is independent of its absolute position within the image. While the homogeneity assumption dramatically reduces the number of non-causal pre-

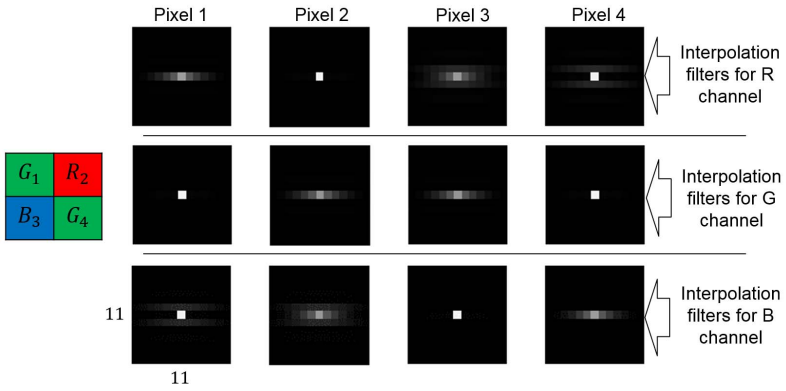


Fig. 4. Horizontal demosaicking filters derived using proposed non-iterative MAP for each of the 4 phases p in the 2×2 Bayer CFA sensor. The filters correspond to unique rows in the inverse matrix $H^{(m)}$ for $m = H$.

diction parameters, it leads to smoothing and aliasing artifacts in the demosaicked image.

Drawing inspiration from birectional (vertical/horizontal) interpolators used frequently in Bayer-CFA demosaikcing [3, 4], we propose to derive bidirectional MAP-based interpolation operators for each of the CFA sensors supported by our universal demosaick algorithm. We obtain a directional demosaick estimate, $\hat{x}^{(m)}$, using a homogeneous GMRF image prior model, described by matrix $B^{(m)}$, that enforces smoothness only along a specified direction:

$$\hat{x}^{(m)} = \left(A^T A + \frac{\sigma^2}{\sigma_l^2} {T_l}^T B^{(m)} T_l + \frac{\sigma^2}{\sigma_c^2} \sum_{k=1,2} T_{c_k}^T B^{(m)} T_{c_k} \right)^{-1} A^T y. \quad (6)$$

Defining the inverse matrix $H^{(m)}$ as in (4), but with the superscript m included, (6) can be written as

$$\hat{x}^{(m)} = H^{(m)} y, \quad (7)$$

where $m \in \{V, H\}$ and $B_{i,j}^{(m)} = \delta_{i-j} - g_{i-j}^{(m)}$. The non-causal prediction parameters $g_k^{(H)}$ used in our paper are given in Figure 2. The values for $g_k^{(V)}$ are given by the transpose of the matrix shown in Figure 2.

A vertical demosaick operator $H^{(V)}$ performs better in areas of the image dominated by vertical edges, while a horizontal demosaick operator $H^{(H)}$ performs betters in areas dominated by horizontal edges. A convex average of the two estimates at each pixel location is used to get the final demosaicked image:

$$\hat{x}_{i,j} = \beta_{i,j} \hat{x}_{i,j}^{(H)} + (1 - \beta_{i,j}) \hat{x}_{i,j}^{(V)}, \quad (8)$$

where $\beta_{i,j} \in [0, 1]$ represents the classifier decision at location (i, j) in the image. For horizontal edges, $\beta_{i,j} = 1$, while for vertical edges, $\beta_{i,j} = 0$. In this manuscript, we use the homogeneity-directed classifier proposed by Hirakawa in [3] to estimate $\beta_{i,j}$.

2.3. Estimation of Inverse Matrix

The columns of the inverse matrix $H^{(m)} = \begin{bmatrix} h_1 & h_2 \dots h_N \end{bmatrix}^T \in \mathbb{R}^{3N \times N}$ represent impulse response of the linear system: $\hat{x}^{(m)} = H^{(m)} y$. Therefore, it is possible to learn the inverse matrix column-by-column with its j-th column computed as:

$$h_j^{(m)} = \arg \min_{x \in \mathbb{R}^{3N}} \left\{ \frac{1}{2\sigma^2} \|e_j - Ax\|^2 + \frac{1}{2\sigma_l^2} x^T {T_l}^T B^{(m)} T_l x + \frac{1}{2\sigma_c^2} \sum_{k=1,2} x^T T_{c_k}^T B^{(m)} T_{c_k} x \right\}, \quad (9)$$

where e_j represents the j-th unit vector.

The reconstruction matrix $H^{(m)}$ needs to be learned apriori only once for a given CFA constellation. The same reconstruction matrix is then used to demosaick all raw data captured using that sensor.

2.4. Filter Extraction From Inverse Matrix

Figure 3 illustrates non-iterative MAP demosaicking based on computing the matrix-vector product: $\hat{x}^{(m)} = H^{(m)} y$. We make following observations about the reconstruction matrix:

- Each row of $H^{(m)}$ represents a directional color interpolation filter.
- For each pixel location and interpolation direction, m, there are 3 demosaicking filters: one for R, one for G, and one for B.
- There are only as many unique rows in $H^{(m)}$ per color channel as the support of the unit CFA cell. Thus, if the support of one CFA period is $M \times N$ pixels, maximum number of unique color interpolation filters per channel per interpolation direction is MN. We denote this set of filters by: $h_{R,p}^{(m)}, h_{G,p}^{(m)}, h_{B,p}^{(m)}$, where $p \in \{1, \dots, MN\}$ and $m \in \{H, V\}$.

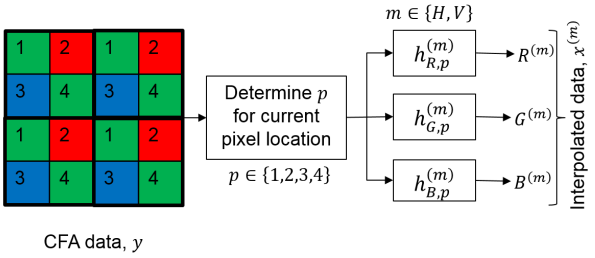


Fig. 5. Real-time implementation of NMAP demosaick.

- Rows of $H^{(m)}$ are sparse and a 2-D interpolation filter represented by a given row has a compact support around the center pixel.

The set of unique rows in $H^{(m)}$ for the standard 2×2 Bayer CFA sensor learned offline using (9), arranged as 2D filters, is shown in Figure 4. We notice that the filter coefficients decay rapidly as the distance from the center increases and are practically zero outside the 11×11 window. For the Bayer example, $M = N = 2$; thus, there are four demosaick filters per channel per interpolation direction.

2.5. Configurable Hardware Architecture

Figure 5 illustrates how the proposed NMAP reconstruction can be computed efficiently in real-time using filtering operations. Using the methodology discussed above, the demosaick filter set $h_{R,p}^{(m)}, h_{G,p}^{(m)}, h_{B,p}^{(m)}$ is pre-computed for each *phase* $p \in \{1, \ldots, MN\}$ in an $M \times N$ CFA pattern the ISP is required to demosaick. The demosaick filter values in the ISP unit are configured according to the CFA sensor attached at the front-end of the camera. The phase p of the current pixel being processed in the observation image y is determined using $modulo - N$ and $modulo - M$ operations, and the corresponding filter set $h_{R,p}^{(m)}, h_{G,p}^{(m)}, h_{B,p}^{(m)}$ is used to estimate the 3-channel color vector $x^{(m)}$.

3. EXPERIMENTAL RESULTS

We used the set of standard 24 Kodak color images to evaluate the performance of the proposed universal demosaick algorithm. The CFA images were generated by color-subsampling of the full 3-channel RGB images according to the patterns shown in Fig. 1. The CFA images were then processed using the NMAP universal demosaick in Figure 5 to obtain the directional estimates $\hat{x}^{(m)}$. We used only horizontal and vertical directions, i.e. $m \in \{H, V\}$, to produce the results. However, the algorithm can easily be extended to incorporate any number of interpolation directions. The directional estimates $\hat{x}^{(H)}$ and $\hat{x}^{(V)}$ were then combined using (8) to obtain the final estimate $\hat{x}$. We used the color homogeneity metric proposed in [3] to estimate the mixing coefficients $\beta_{i,j}$.

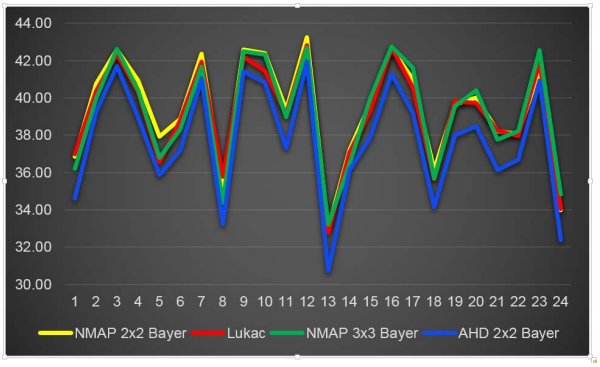


Fig. 6. PSNR values of full 3-channel RGB images reconstructed from three different CFA sampling patterns. Two different demosaick algorithms were used for reconstructions: (a) Proposed non-iterative MAP (NMAP) algorithm was used to demosaick 2×2 Bayer, 3×3 Bayer, and 4×4 Lukac patterns; (b) Adaptive homogeneity-directed (AHD) demosaick was used to demosaick only 2×2 Bayer data.

The demosaick filters for each of the three CFA patterns were learned apriori using the procedure described in Sections 2.3 and 2.4. The filter support selected was 11×11; non-zero values outside this support were discarded.

In Figure 6, we evaluate the performance of our universal demosaick algorithm across our test set of 24 Kodak images using the PSNR measure of image quality. For the Bayer sensor, we also use adaptive homogeneity-directed demosaick (AHD) [3] to perform color interpolation. The simulation results show that our universal demosaick outperforms AHD [3] on the Bayer data, while additionally providing the capability for demosaicking non-Bayer CFA images.

4. CONCLUSIONS

We proposed a universal demosaick algorithm based on statistical MAP estimation that can be configured to demosaick raw images captured using a variety of CFA sensors. Our proposed approach involves pre-computing and storing the inverse operator required for MAP reconstruction of the 3-channel RGB image. The reconstruction operator is learned offline for a given CFA pattern and then approximated using a set of compact 2D spatial filters. The resulting non-iterative MAP estimation, implemented using a filter bank, is computationally efficient and hardware friendly. We use a set of 24 Kodak color images to evaluate the quality of our demosaick algorithm on three different CFA patterns. The high PSNR values of the reconstructed RGB images show the effectiveness of the proposed method.

5. REFERENCES

[1] Keigo Hirakawa and Patrick J Wolfe, "Spatio-spectral color filter array design for optimal image recovery," *Image Processing, IEEE Transactions on*, vol. 17, no. 10, pp. 1876–1890, 2008.

[2] John T Compton and John F Hamilton Jr, "Image sensor with improved light sensitivity," Mar. 20 2012, US Patent 8,139,130.

[3] K. Hirakawa and T.W. Parks, "Adaptive homogeneity-directed demosaicing algorithm," in *Proc. IEEE Int. Conf. Image Processing*, 2003, pp. 669–672.

[4] Hasib Siddiqui and Hau Hwang, "Training-based demosaicing," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1034–1037.

[5] Charles Bouman and Ken Sauer, "A generalized gaussian image model for edge-preserving map estimation," *Image Processing, IEEE Transactions on*, vol. 2, no. 3, pp. 296–310, 1993.

[6] Jayanta Mukherjee, R Parthasarathi, and S Goyal, "Markov random field processing for color demosaicing," *Pattern Recognition Letters*, vol. 22, no. 3, pp. 339–351, 2001.

[7] Biay-Cheng Hseih, Hasib Siddiqui, Jiafu Luo, Todor Georgiev, Kalin Atanassov, Sergio Goma, HY Cheng, JJ Sze, RJ Lin, KY Chou, et al., "New color filter patterns and demosaic for sub-micron pixel arrays," .

[8] Yue M Lu and Martin Vetterli, "Optimal color filter array design: Quantitative conditions and an efficient search procedure," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2009, pp. 725009–725009.