Compressive Demosaicing

Abdolreza Abdolhosseini Moghadam#1, Mohammad Aghagolzadeh#2, Mrityunjay Kumar*3, and Hayder Radha#4
Dept. of Electrical & Computer Eng., Michigan State University, * Eastman Kodak Company, Rochester, NY, USA.
1 abdolhos@msu.edu, 2 aghagol1@msu.edu , 4 radha@msu.edu, 3 mrityunjay.kumar@kodak.com

***Abstract*—A typical consumer digital camera uses a Color Filter Array (CFA) to sense only one color component per image pixel. The original three-color image is reconstructed by interpolating the missing color components. This interpolation process (known as *demosaicing*) corresponds to solving an underdetermined system of linear equations. In this paper, we show that by replacing the traditional CFA with a random panchromatic CFA, recent results in the emerging field of *Compressed Sensing* (CS) can be used to solve the demosaicing problem in a novel way. Specifically, during the image reconstruction process, we exploit the fact that the multi-dimensional color of each pixel has a *compressible* representation in a (possibly overcomplete) color system. While adhering to the "single color per pixel sensing" constraint at the sensing stage, during the reconstruction process we utilize the inter-pixel correlation by exploiting the compressible representation of the overall image in some sparsifying bases. Depending on the CFA, sparsifying bases and the color system, we form an underdetermined system of linear equations and find the sparsest solution for the color image by utilizing a CS solver. We illustrate that, for natural images, the proposed Compressive Demosaicing (CD) framework visually performs better (mostly significantly better) than leading demosaicing approaches. Furthermore, CD could achieve virtually perfect reconstruction on most tested images.**

I. Introduction

Motivated by cost constraints, most low-cost consumer grade digital camera systems are currently designed to (a) sense only one color component per image pixel and (b) interpolate the other missing color components (at each pixel) during reconstruction. The sensing process, which employs a Color Filter Array (CFA), maps each pixel to a single color based on a color pattern. The CFA color pattern and the interpolation process (widely known as *demosaicing*) have a significant impact on the quality of the reconstructed image. The most popular CFA pattern is the Bayer color pattern that employs two green filters, one red, and one blue filter in each 2×2 block within the CFA. Many other CFA patterns have been proposed including ones that are based on secondary colors [3]. There has been a great deal of attention paid to the demosaicing problem, and consequently, a flurry of algorithms has been proposed [2]-[11]. Several recent papers on image demosaicing provide an excellent overview of leading approaches and their classification (e.g., spatial versus frequency domains) [1]. In general, demosaicing algorithms exploit the correlation that exists among adjacent pixels (inter-pixel correlation) and among color planes (inter-channel correlation) [1]-[11].

Meanwhile, the area of *Compressed Sensing* (CS) [12] has attracted a great deal of attention recently. The problem of CS targets the sparsest solution of an underdetermined system of linear equations. Similarly, the problem of demosaicing is basically an attempt to finding a solution to an underdetermined system of linear equations where for each pixel, one linear sample of three color components is sensed. In principle, CFA-based image capture represents a three-to-one compressed sensing. Hence, utilizing the rich results developed in the CS area to solve the demosaicing problem seems plausible. In this paper, we present *Compressive Demosaicing* (CD), a framework to demosaic natural images by employing aspects from the theory of CS. More specifically, instead of finding the missing color components of a pixel, we find an equivalent compressible description of the same image. This equivalent description of the image is essentially the redundant representation of that image with minimal inter-channel and inter-pixel correlations. In words, given the CFA samples, the proposed CD framework finds the transform coefficients of the image (with respect to a sparsifying frame or basis) in a redundant color space, by algorithms developed in the CS area to reconstruct the three-color image. We employ a random panchromatic CFA during the sensing stage of our proposed framework.

It is important to highlight that the proposed compressive demaosaicing framework differs significantly from other recent attempts for combining CS and CFA sensing. In particular, the utility of CS for sensing color images has been proposed in [19]. Our proposed CD framework departs from prior work in many ways both in terms of the problem objectives and also the approach to solve that problem. For instance, [19] requires a compressed sensing camera [18] (where for each pixel, a linear measurement of the whole image is sensed) and hence requires drastic changes in the design of digital cameras which might not be feasible (at least at the present time). On the other hand, in our method, for each pixel, we only sense a linear combination of color components of that (single) pixel, which can be achieved simply by employing a random panchromatic CFA. Hence, we strictly adhere to the "single color per pixel" constraint. Second, [19] utilizes a joint sparsity model to recover a sparse representation of the color image. On the other hand, we utilize a novel combination of Equiangular Tight Frames (ETFs) along with YUV color system to de-correlate the color components of an image.

The remainder of the paper is organized as follows. In section 2, we review the sensing process during image capture and also briefly introduce the problem of CS and how it is related to the proposed compressive demosaicing. In section

3, we formulate the compressive demosaicing problem and describe the redundant sparse/compressible *equivalent form* of the image and how we demosaic that image. Simulation results for natural images are presented in Section 4. Section 5 concludes the paper.

II. Sensing for Compressive Demosaicing

In this section, we review the sensing process during image capture and link the demosaicing problem to the CS problem [12]. Assume that the image of interest consists of three color planes, red ($\mathbf{R}$), green ($\mathbf{G}$) and blue ($\mathbf{B}$). In words, the color of the pixel located at the Cartesian location (i,j) is in the form of $(\mathbf{R}_{i,j}, \mathbf{G}_{i,j}, \mathbf{B}_{i,j})$ in the RGB color system. Using a generic CFA, an $n_1 \times n_2$ image sensed by a "single color per pixel" digital camera can be represented as:

$$\forall (i,j) \in [n_1] \times [n_2] : \mathbf{y}_{i,j} = \alpha_{i,j}\mathbf{R}_{i,j} + \beta_{i,j}\mathbf{G}_{i,j} + \gamma_{i,j}\mathbf{B}_{i,j} \quad (1)$$

where $\forall q \in \mathbb{N},\ [q] := \{1, 2, \ldots, q\}$, $\mathbf{y}_{i,j}$ is the (single) sensed color at pixel (i,j) and $\alpha_{i,j}$, $\beta_{i,j}$ and $\gamma_{i,j}$ are some positive weights associated with the red, green and blue wavelengths at pixel location (i,j), respectively, with the constraint $\forall i,j$: $\alpha_{i,j} + \beta_{i,j} + \gamma_{i,j} = 1$ [3]. Extending this formulation to the whole image yields: $\mathbf{y} = \alpha \odot \mathbf{R} + \beta \odot \mathbf{G} + \gamma \odot \mathbf{B}$ where $\odot$ stands for the Hadamard (point wise) product. Equation (1) suggests that for any pixel (i,j) we have an underdetermined system of one linear equation $\mathbf{y}_{i,j}$ and three unknowns $\mathbf{R}_{i,j}, \mathbf{G}_{i,j}, \mathbf{B}_{i,j}$. Indeed, the objective of all demosaicing algorithms is to find these unknowns. However, from elementary linear algebra, we know that an underdetermined system of linear equations has an infinite set of solutions and hence the problem of demosaicing might not be solved in a linear algebraic way. Therefore, researchers over the past few decades proposed many alternative approaches to tackle this problem including bilinear interpolation, filtering, demodulating and many others [1]. Most of these demosaicing methods assume some hypothesis (or prior model) about the image (to name a few, the colors of adjacent pixels obey a certain relationship or the image edges are horizontally or vertically oriented) and then design their recovery algorithms based on these assumptions. Thus, demosaicing methods work quite well when the corresponding assumptions are satisfied; however the same methods can fail in a significant way once the underlying assumptions are violated. For instance, the presence of a diagonal edge in the image usually causes visible artifacts.

In this paper, we show that recent advances in the emerging field of Compressed Sensing [12] enable us to design a demosaicing algorithm by exploiting the fact that natural images have sparse or compressibe representations in some transforms such as Discrete Cosine Transform (DCT), Contourlets [21], directional wavelets [16]-[17] and so on. Before presenting our algorithm, let us briefly review the Compressed Sensing (CS) problem.

CS targets the *sparsest* solution of an underdetermined system of equations [12], [13]:

$$(P_0): \ \arg\min \|x\|_0 : y_{m\times 1} = P_{m\times n} x_{n\times 1}, \ m < n \quad (2)$$

where $x \in \mathbb{R}^n$ is the target sparse or compressible[1] unknown vector, y is the measurement vector (set of equations), P is the measurement matrix and hence $\|x\|_0$ counts the number of non-zero elements of x. Solving (P_0) directly is NP-hard. It has been shown that under some conditions [12]-[14], the solution x of problem (P_0) is the same as the solution to the following problem:

$$(P_1): \ \arg\min \|x\|_1 : y_{m\times 1} = P_{m\times n} x_{n\times 1}, \ m < n \quad (3)$$

Now (P_1) is a convex optimization problem and can be solved tractably, for instance using Basis Pursuit (BP) [15].

Naturally, the problem of demosaicing is a set of seemingly independent underdetermined system of linear equations (1) and hence utilizing the rich results in the area of CS seem plausible. Here, we show that replacing traditional CFAs by a *random panchromatic CFA* enables us to demosaic the image by CS algorithms. By random panchromatic CFA we imply that for each pixel, the weights α, β and γ in (1) are some positive random numbers that add up to one. The reason that we have utilized random panchromatic CFA will be discussed in more detail in the subsequent sections; meanwhile, an intuitive justification can be highlighted as follows: when we sense only one of the primary colors (red, green or blue) for a particular pixel, then we are discarding the information about the other two colors, which have not been sensed for that particular pixel; and this is an irreversible mapping. On contrary, if we sense a *linear combination* of color components of a pixel, then we are retaining information about all three colors; yet, we now have to face the problem of separating these unknowns (color components) from the equation (one sample).

To employ CS in the problem of demosaicing without introducing dramatic changes in hardware, we have to address some issues and overcome some obstacles, few of which we list here. First, in CS, we would traditionally sense $m > 1$ linear samples from the same signal; however in digital cameras, in the best case when a random panchromatic CFA is employed (or equivalently $\forall i,j : \alpha_{i,j} \neq 0, \beta_{i,j} \neq 0, \gamma_{i,j} \neq 0$) we are sensing only one compressive sample ($m = 1$) from three unknowns (the color components of a specific pixel). In words, for each pixel we have a (seemingly independent) system of only one equation and three unknowns and these unknowns might not be sparse in the RGB color coordinate (and CS does not apply to this type of problem). Second, most of the popular CS decoding algorithms require the underlying signal to have a high dimension ($x \in \mathbb{R}^n$ where $n \gg 1$). Again if we attempt to recover the color components of each pixel individually, even if the pixel has a sparse representation in the RGB domain, then there is no guarantee that the underlying CS decoder would find that solution. Finally, Basis Pursuit (BP), which is arguably the most reliable and best performing CS decoder (in terms of quality of the reconstructed signal)

[1]By k-sparse we mean x is non-zero in k indices ($k = \|x\|_0 = \{\#i : x_i \neq 0\}$). Similarly x is k-compressible if x has k significant non-zero coefficients and the rest of coefficients are very small.

requires $m \approx 5k$ compressive samples to reconstruct a k-sparse/compressible signal with "an acceptable" error. This suggests that the signal which we are trying to find, has to be (approximately) 20% compressible which is not the case for RGB color planes of a natural image. Therefore, we need to recover another (yet equivalent) form of the image such that this equivalent form must be sufficiently sparse.

In our work, and while strictly adhering to the "single color per pixel" CFA at the sensing stage, we address the above challenges by applying our demosaicing algorithm on blocks of the image and employing redundant color spaces in the solver. More specifically, we *jointly* demosaic (blocks of) an image (as opposed to pixel by pixel demosaicing). By applying block based demosaicing, one might exploit the fact that the underlying block has a compressible representation in a basis or frame (for instance DCT, directional wavelets [16]). Furthermore, inter-channel correlations are exploited by utilizing a redundant color space, instead of a traditional three-color space. To summarize, given the CFA samples, we utilize the inter-pixel correlations by looking for the sparsest solution within some transform coefficients of the image; and exploit the inter-channel correlations by looking for these transform coefficients in a redundant color space. Below, we outline the problem formulation of the proposed compressive demosaicing framework.

III. Problem Formulation of Compressive Demosaicing

As before, suppose the target color image is composed of three color planes $\mathbf{R}_{n_1\times n_2}$, $\mathbf{G}_{n_1\times n_2}$ and $\mathbf{B}_{n_1\times n_2}$ (hence the size of the image is $n_1 \times n_2$ pixel). Without loss of generality and to simplify the equations, let us consider the *vectorized* form of the image. By vectorized form we mean that we stack the columns of the image on top of each others, that is:

$$l = (j-1)n_1 + i : \; R_l = \mathbf{R}_{i,j}, \; G_l = \mathbf{G}_{i,j}, \; B_l = \mathbf{B}_{i,j} \quad (4)$$

Throughout this paper, we denote 2D forms of images and also samples by bold-face letters ($\mathbf{A}$) and use the same letter in the normal size italic font (A) to show its vectorized form. Let $N = n_1 n_2$ be the total number of pixels in the image of interest. Then we can re-express (1) in the vectorized form: $\forall l \in [N] = \{1, \ldots, N\} : y_l = \alpha_l R_l + \beta_l G_l + \gamma_l B_l$, or equivalently in the matrix form by:

$$y = \phi [R^T \; G^T \; B^T]^T \quad (5)$$

where for $b \in \{y, R, G, B\} : \; b = [b_1 \; \ldots \; b_N]^T$ and ϕ can be defined by means of matrices $\bar{\alpha}$, $\bar{\beta}$ and $\bar{\gamma}$ as:

$$\phi = [\; \bar{\alpha} \quad \bar{\beta} \quad \bar{\gamma} \;] \quad (6)$$

and $\bar{\alpha}$, $\bar{\beta}$ and $\bar{\gamma}$ are diagonal matrices in the following form.

$$\bar{\alpha}_{i,j} = \begin{cases} \alpha_i & i = j \\ 0 & i \neq j \end{cases}, \; \bar{\beta}_{i,j} = \begin{cases} \beta_i & i = j \\ 0 & i \neq j \end{cases}, \; \bar{\gamma}_{i,j} = \begin{cases} \gamma_i & i = j \\ 0 & i \neq j \end{cases}$$

Hence in the demoasicing problem, we have ϕ (the CFA) and y (sensed image) and we are searching for the R, G and B vectors. It is important to highlight that at this stage, we may not apply a CS decoding algorithm to recover the missing color components due to the following reasons: 1) The vector $[R^T \; G^T \; B^T]^T$ is not necessarily sparse or compressible; and 2) even if $[R^T \; G^T \; B^T]^T$ is sparse and the CFA is random panchromatic (each row of ϕ is non-zero in three column indices), the matrix ϕ is ill-conditioned in terms of what is known as the *Restricted Isometry Constant* (RIC) measure [12]-[14]. This makes ϕ unsuitable for most CS decoders[2]. Hence, we need to change the problem of finding a solution to (5) into an equivalent problem that is suitable for CS solvers. For instance, we need to formulate a variant to (5), for example $y = P\zeta$, which is better-conditioned; meaning, the solution vector ζ is compressible and the RIC measure for P is smaller compared to ϕ. These issues are addressed in the following subsections.

A. Exploiting inter-pixel correlations

In the CS framework, each measurement y_i is usually a linear combination of all (or subsets of large size) of the unknown signal x. This is usually achieved by employing dense measurement matrices P in (2). However, in demosaicing of an $N = n_1 n_2$ pixel digital image, it seems that we have N independent small CS problems in the form of one measurement and three unknown color components for each pixel. As stated before, this type of problem is not generally solvable under CS. Therefore, the first step of our proposed method shall be: re-shaping the demosaicing problem into a format suitable for CS without introducing any dramatic changes in the hardware of digital cameras. To that end: 1) instead of recovering (or interpolating) the missing color components individually for each pixel, we recover these missing colors jointly for blocks of an image and 2) instead of attempting to directly recover RGB planes, we look for the transform coefficients of an alternative RGB planes in a redundant color space. By doing so, we achieve several goals simultaneously which we describe below.

Note that the value of each pixel in any color plane can be considered as a linear combination of transform coefficients of that image in some space. The number of transform coefficients that contribute to the color value is dictated by the nature of the transform. For instance, for Fourier, DCT or any global transform domain, this subset is the set of all transform coefficients (the value of each pixel is a linear combination of all of Fourier/DCT transform coefficients). For local transforms such as wavelets and similar transforms, the size of the set depends on the support size of basis elements. Consequently, with a random panchromatic CFA, each sample would be a linear combination of transform coefficients of RGB color planes. Now this is the format desirable for CS. Moreover, we can exploit the high inter-pixel correlations among adjacent pixels in our proposed method to make the objective signal (representing the same image) sparse and

[2]Broadly speaking, most of CS solvers require that any full rank sub-matrix of the underlying measurement matrix (P) behaves like an orthogonal system.

consequently help the CS solver. Recall that high inter-pixel correlation translates to sparse/compressible representation of these pixels in another transform. For instance, it has been known that the DCT coefficients of texture regions are sparse. Similarly different kinds of directional wavelets [16], [17] represent edges effectively. Again, this motivates searching for transform coefficients of RGB planes (instead of attempting to recover RGB planes directly). Finally, as stated before, most CS decoding algorithms succeed (with some tolerable error) when the underlying signal is in high dimensions. Now instead of finding (R_l, G_l, B_l) for each pixel, if we attempt to find the block transform coefficients of these color planes $(\hat{R}, \hat{G}, \hat{B})$, the length of the solution vector equals (at least) the number of pixels in that block which is advantageous for CS decoders. One might even *virtually* lengthening the solution vector furthermore by utilizing redundant frames in the decoder. In the rest of this subsection we show that by targeting the transform coefficients, we improve the conditions of the virtual measurement matrix P for the recovery process.

As long as a transform is linear, we can express the analysis/synthesis steps in a matrix form. For instance, assume that we want to represent the red plane of the image in a separable transform (A, B): $\mathbf{R} = A\hat{\mathbf{R}}B$. It is straightforward to show that the transform coefficients in the vectorized form are in the form of: $R = \psi_R \hat{R}$ where $\psi_R = B^T \otimes A$ and $\otimes$ is the Kronecker tensor product operator (recall that R is the vectorized form of $\mathbf{R}$). Now assume R, G and B have sparse/compressible representations $(\hat{R}, \hat{G}, \hat{B})$ in the transform domains ψ_R, ψ_G and ψ_B respectively, that is: $R = \psi_R \hat{R}$, $G = \psi_G \hat{G}$ and $B = \psi_B \hat{B}$. Define Ψ as:

$$\Psi = \begin{bmatrix} \psi_R & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \psi_G & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \psi_B \end{bmatrix} \tag{7}$$

then, we have: $[R^T\ G^T\ B^T]^T = \Psi \left[\hat{R}^T\ \hat{G}^T\ \hat{B}^T\right]^T$. Define $\eta = \phi\Psi$. Then (5) would become:

$$y = \eta[\hat{R}^T\ \hat{G}^T\ \hat{B}^T]^T \tag{8}$$

Note that combining (7) and (6) simplifies η by:

$$\eta = \left[\ \bar{\alpha}\psi_R \quad \bar{\beta}\psi_G \quad \bar{\gamma}\psi_B\ \right] \tag{9}$$

Finally depending on the nature of bases used in the matrix Ψ, the RIC of η might be more suitable than the same parameter for the matrix ϕ, and hence $y = \eta[\hat{R}\ \hat{G}\ \hat{B}]^T$ has a better chance of recovery by any CS decoding algorithm (compared to (1)). In the next section, we exploit the inter channel correlations to further sparsify the solution vector we are searching for.

B. Exploiting inter-channel (color) correlation

It is well known that YUV or similar color spaces are more efficient color coordinates than the RGB space for compression applications. For instance, if the color of a pixel in the RGB space is in the form of (a, b, c), then this color can be expressed in the form of (e, f, g) in the YUV color space where (e, f, g) decays faster compared to (a, b, c) when represented in some transform space (e.g. DCT). Similar to the idea of extending bases to frames, one might extend the color coordinate basis vectors to an *over-complete color coordinate* system to represent the colors of an image even more sparsely. In matrix form, one can express the RGB color system using another set of colors $\{c_1, c_2, \ldots, c_q\}$:

$$[r\ g\ b]^T = \theta_{3\times q}[c_1\ c_2\ \ldots\ c_q]^T,\ q \geq 3 \tag{10}$$

where θ is the matrix for converting colors $\{c_1, \ldots, c_q\}$ into the RGB color system. Clearly, by increasing q (the number of colors in the color system), we are increasing the likelihood of expressing the color of a pixel sparsely. Note that $\{c_1, \ldots, c_q\}$ are used for analysis (and not synthesis) at the demosaicing solver only. In words, by targeting a color space $\{c_1, \ldots, c_q\}$ with higher sparsity levels than traditional YUV or other 3D color spaces, neither we utilize these sparisfying over-complete colors for displaying the reconstructed images nor we require sensing a larger number of colors when capturing the image. These colors solely facilitate expressing the color of any pixel in a redundant and sparse format and hence help the CS solver during the decoding process. In our proposed compressive demosaicing framework, we propose the utility of a novel color space that includes Equiangular Tight Frames (ETF) [22] along with YUV in the color transform θ. Before describing the task of θ, let us briefly review the key properties of ETFs.

A real valued (n, k)-equiangular-tight frame (where $n > k$) is a set of n unit norm vectors $\{f_1, \ldots, f_n\}$ in $\mathbb{R}^k$ with the strong property of: $\forall q \neq p \in [n] : |\langle f_q, f_p\rangle| = \chi$, that is the absolute value of inner-product of any two different vectors in the frame is constant. It can be verified [22] that these vectors correspond to finding n lines in $\mathbb{R}^k$ for which the closest pair (in terms of angle) is as far apart as possible. The motivation for using a random panchromatic CFA in conjunction with ETF's in the proposed CD framework should be clear now. Recall that one major drawback of using a Bayer pattern is that for each pixel, the measured color is projected onto only one of the 3D color coordinate bases; hence, we completely lose the information about the other two colors (for that particular pixel). For an ETF, none of the vectors are orthogonal to each other and meanwhile all color vectors have the same angular distances with respect to each other. In words, in the sensing process (which is random panchromatic in our method), we are not discriminating any color over the other ones, and at the solver side an ETF system with sufficient color coordinates can sparsify the color. For sparsifying the RGB components of a pixel, we have used the redundant color coordinate frame composed of a (6,3)-ETF and YUV color space (hence $q = 9$). Let $\lambda = \frac{1+\sqrt{5}}{2}$ and form $(6, 3)$-ETF by:

$$\theta_{ETF} = \frac{1}{\sqrt{1+\lambda^2}} \begin{bmatrix} 0 & 0 & 1 & 1 & \lambda & -\lambda \\ 1 & 1 & \lambda & -\lambda & 0 & 0 \\ \lambda & -\lambda & 0 & 0 & 1 & 1 \end{bmatrix} \tag{11}$$

then θ that we have used for our simulations, is in the form of:

$$\theta = [\theta_{YUV}\ \theta_{ETF}] \tag{12}$$

Now let us denote the identity matrix of size $N \times N$ by I_N and define Θ (which is a $3N \times qN$ matrix) as $\Theta = \theta \otimes I_N$; and as before let R, G and B be, respectively, the vectorized form of the red, green and blue color planes of the image of interest. Then it can be easily verified that:

$$\left[R^T \; G^T \; B^T\right]^T = \Theta \left[\zeta_1 \; \ldots \zeta_{qN}\right]^T \tag{13}$$

where $\{\zeta_1, \ldots, \zeta_{qN}\}$ represents the colors of the same image in the color system $\{c_1, \ldots, c_q\}$. Note that the transform coefficients $\hat{R}$, $\hat{G}$ and $\hat{B}$ in (8) belong to color planes R, G and B respectively. Hence a similar equality holds true for the transform coefficients of the image in different color planes:

$$\left[\hat{R}^T \; \hat{G}^T \; \hat{B}^T\right]^T = \Theta \left[\hat{\zeta}_1 \; \ldots \hat{\zeta}_{qN}\right]^T \tag{14}$$

where $\hat{\zeta} = [\hat{\zeta}_1 \; \ldots \; \hat{\zeta}_{qN}]^T$ might be thought of as sparse (or compressible) and redundant color components of transform coefficients of an image in the color space $\{c_1, \ldots, c_q\}$. Note that having $\hat{\zeta}$, (the vectorized form of) the image can be reconstructed uniquely by $[R^T \; G^T \; B^T]^T = \Psi\Theta\hat{\zeta}$. The objective of compressed demosaicing is finding $\hat{\zeta}$. In the next sub-section, we summarize our proposed compressed demosaicing framework.

C. Integrating inter-pixel and inter-channel correlations

Now we can explicitly express the sensing and the demosaicing stages of our proposed method. As stated before, the only change that we require in the hardware, is substituting a traditional Bayer CFA by a random panchromatic CFA. Hence, the sensed image is in the form of (1). We use y the vectorized form of the captured image (5) along with ϕ in (6) (the matrix describing the CFA) in the image decoder. Integrating (8) and (14) yields:

$$y = \eta\Theta[\hat{\zeta}_1 \; \ldots \; \hat{\zeta}_{qN}]^T \tag{15}$$

Note that the vector $\hat{\zeta} = [\hat{\zeta}_1 \; \ldots \; \hat{\zeta_{qN}}]^T$ expresses the same image with the minimal correlations both among adjacent pixels and among the colors of any pixel. Hence $\hat{\zeta}$ is a compressible signal. Meanwhile it is easy to verify that $P = \phi\Psi\Theta = \eta\Theta$ is a dense matrix and with high probability any subset of its columns is full-rank (because of random entries in ϕ and also the nature of Θ in our method). Now giving y and P, any generic CS decoder (for instance BP) recovers $\hat{\zeta}$ by solving:

$$x = \arg\min \|\bar{\hat{\zeta}}\|_1 : y_{N\times 1} = P_{N\times qN}\bar{\hat{\zeta}}_{qN\times 1} \tag{16}$$

After recovering x (the estimate of $\hat{\zeta}$) we estimate the vectorized forms of RGB color planes of the target image by $\left[R^T \; G^T \; B^T\right]^T = \Psi\Theta x$. Reshaping R, G and B into $\mathbf{R}$, $\mathbf{G}$ and $\mathbf{B}$ (the matrix form) we display the demosaiced image.

IV. Simulation Results

We have tested our proposed method to demosaic natural images and compared our results with some prominent demosaicing methods. The number of demosaicing methods, proposed in the community is so large that we can present only some of the leading approaches to compare them with the proposed CD framework. Among the leading demosaicing approaches we present here are the Homogeneity-directed interpolation [2], Alternating Projections (POCS) [4], Successive Approximations [6] and the method of Hirakawa & Wolfe. The first three methods were applied on a Bayer CFA image while a Spatio-Spectral [3] CFA image and a random panchromatic CFA data were generated for the method of Hirakawa and our proposed CD method. In this section, we show the results for demosaiced images of the well-known "Lighthouse" image (most popular in the demosaicing literature), and we also show results for the "Barbara" image. Both images contain quite challenging high-frequency regions where a demosaicing method can easily fail. Due to space limitations, some cropped regions of these sets of demosaiced images are presented in this section. For example, we focus on the well-know fence area of the "Lighthouse", where most demosaicing papers focus on. In all these simulations, we have chosen YUV and $(6,3)$-ETF for the sparsifying color coordinate in (12) and DCT as the sparsifying transform in (7) and run our algorithm on blocks of size 16×16 pixel. To eliminate the blocking effect, we considered 12 pixel of overlap between adjacent blocks and selected the medians of values calculated for each pixel (for each color). For the solver, we have used the Basis Pursuit algorithm to recover $\hat{\zeta}$ in (16). Finally after the recovery, we have used a median filter to reduce color artifacts [5].

As clearly demonstrated in Fig. 2 and Fig. 1, CD has achieved virtually perfect reconstruction for these parts of the tested images while other methods introduce some artifacts in the same images. These (almost) perfect reconstructions by CD are due to the fact that (the overlapping) sub-blocks of these images can be described in a compressible format in the DCT domain and in the color space $\{(6,3)\text{-ETF, YUV}\}$. More importantly, since the proposed method does not employ any kind of filtering for interpolation process, none of the demosaiced images was blurred.

V. Conclusion

In this paper, we introduced Compressive Demosaicing (CD), a new framework for demosaicing images. As opposed to traditional demosaicing methods, CD exploits inter-channel and inter-pixel correlations to find the compressible transform coefficients of the image of interest in an over-complete color space by utilizing an optimal compressed sensing solver. The only hardware change we require for our method, is to employing a random panchromatic CFA.

Acknowledgment

This work was supported in part by Kodak Research Laboratories and the National Science Foundation.

References

[1] B.K.Gunturk, J.Glotzbach and et. al, "Demosaicking: color filter array interpolation," IEEE SP Magazine, 22(1):44-54, Jan 2005.

[2] K. Hirakawa and T. W. Parks, "Adaptive homogeneity-directed demosaicing algorithm," IEEE Transactions on IP, 14(3):360-369, Mar 2005.

[3] K. Hirakawa and P. J. Wolfe, "Spatio-spectral color filter array design for optimal image recovery," IEEE Trans. IP, 17(10):18761890, Oct. 2008.

Fig. 1. From left to right and top to bottom: A small portion of the "Barbara" image demosaicked by different algorithms: original image, POCS [4], Successive Approximation [6], Homogeneity Directed interpolation [2], Spatio-Spectral CFA [3] and our proposed method.

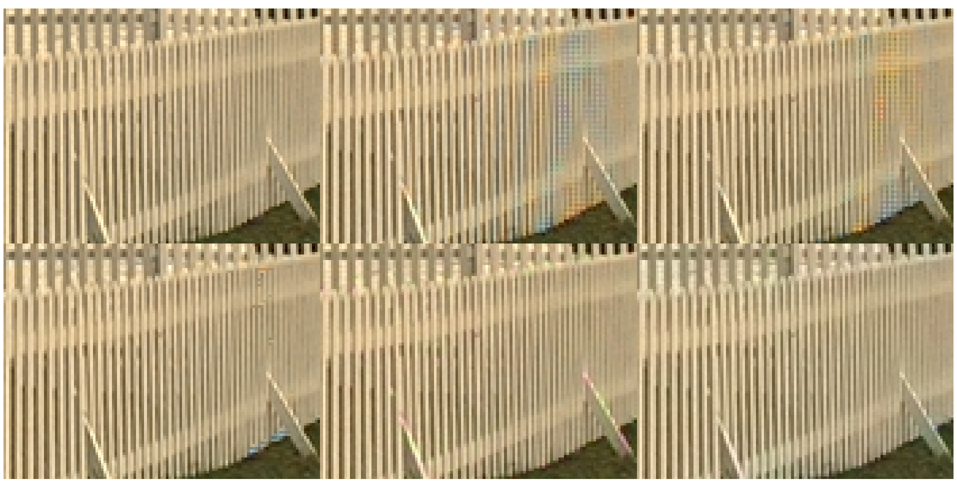

Fig. 2. From left to right and top to bottom: A small portion of the "lighthouse" image demosaicked by different algorithms: original image, POCS [4], Successive Approximation [6], Homogeneity Directed interpolation [2], Spatio-Spectral CFA [3] and our proposed method.

[4] B.K.Gunturk, Y. Altunbasak, R. M. Mersereau, "Color plane interpolation using alternating projections," IEEE Trans. IP, 11(9):997-1013, 2002.

[5] W. T. Freeman, "Median filter for reconstructing missing color samples," *U.S. Patent* 4 724 395, 1988.

[6] X. Li, "Demosaicing by successive approximation," IEEE Transactions on IP, 14(2):370-379, 2005.

[7] W. Lu and Y. Tan. "Color filter array demosaicking: New method and performance measures". IEEE Trans. on IP, 12(10):1194-1210, Oct. 2003.

[8] R. Kimmel, "Demosaicing: image reconstruction from color CCD samples," IEEE Trans. on IP, 8(9):1221-1228, 1999.

[9] X. Wu and N. Zhung, "Primary-consistent soft-decision color demosaicking for digital cameras", IEEE Trans. on IP, 13(9):1263 - 1274, Sep 2004.

[10] D. Paliy, V. Katkovnik and et. al., "Spatially adaptive color filter array interpolation for noiseless and noisy data". International Journal of Imaging Systems and Technology, 17(3): 105 - 122, Oct. 2007.

[11] D. Menon, S. Andriani, G. Calvagno, "Demosaicing with directional filtering and a posteriori decision". IEEE Trans. IP, Jan. 2007.

[12] David Donoho, "Compressed sensing". IEEE Trans. on IT, 52(4):1289-1306, Apr 2006.

[13] E. Cands, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information". IEEE Trans. on IT, 52(2) pp. 489 - 509, Feb. 2006.

[14] David Donoho and Yaakov Tsaig, "Extensions of compressed sensing". Signal Processing, 86(3), pp. 533-548, Mar. 2006.

[15] S. Chen, , D. Donoho, and M. Saunders. "Atomic decomposition by basis pursuit". Tech. Report 479, Dep. of Stat., Stanford Univ. May 1995.

[16] I. W. Selesnick, R. G. Baraniuk and N. G. Kingsbury, "The Dual-Tree Complex Wavelet Transform," IEEE SP Magazine, 22(6):123-151, 2005.

[17] R. Eslami and H. Radha, "New Image Transforms using Hybrid Wavelets and Directional Filter Banks: Analysis and Design," ICIP, Sep. 2005.

[18] M. Duarte, M. Davenport and et. al "Single-pixel imaging via compressive sampling." IEEE SP Magazine, 25(2) pp. 83-91, March 08.

[19] P. Nagesh and B. Li, "Compressive Imaging of Color Images". IEEE ICASSP, Taipei, Taiwan, 09.

[20] http://sparselab.stanford.edu

[21] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," IEEE Trans. on IP, 14(12):2091-2106.

[22] Mtys A. Sustika, Joel A. Troppb, Inderjit S. Dhillona and Robert W. Heath Jr. "On the existence of equiangular tight frames", preprint.