Joint Color Decrosstalk and Demosaicking for CFA Cameras

Xiaolin Wu and Xiangjun Zhang

Department of Electrical and Computer Engineering

McMaster University, Canada

(e-mail: xwu@ece.mcmaster.ca)

Abstract

In interest of low cost, low power consumption, and compact size, most digital cameras adopt a design of single sensor array coupled with a color filter array. This design inevitably suffers, due to physical characteristics of the optical and semiconductor components and the imperfection of manufacturing, from the problem of crosstalk between different color channels. Channel crosstalk can desaturate colors and blur image details, but the problem was seemingly overlooked by existing color demosaicking algorithms. To rectify this deficiency we propose a new joint demosaicking and decrosstalk technique that counters the effects of channel crosstalks by adaptive least-squares inverse filtering. The new technique integrates the operations of deconvolution for crosstalk removal and interpolation for color demosaicking, and it introduces a general framework in which any spatially varying crosstalks and varying spatial-spectral correlations can be modeled and factored into the color reproduction. Simulation results show that the proposed technique is highly effective and capable to obtain both high color fidelity and sharp, clean spatial details.

Key words: Color crosstalk, color filter array, digital cameras, color demosaicking, image restoration.

I. INTRODUCTION

The prevailing design of digital color cameras is based on a single sensor array coupled with a color filter array (CFA). We call such cameras CFA cameras as opposed to those that employ multiple sensor arrays CFA cameras enjoy the advantages of low cost, low power consumption, and compact size over multisensor cameras, which are highly desired features for a wide range of products and systems, such as those used in cell phones, laptop computers, security systems, tracking systems, miniature cameras embedded in medical devices, automotive vision systems, and of course consumer still and video cameras. CFA cameras are also found in high end applications. For example, two popular 4K resolution digital movie cameras, Red and DALSA, are of the single-sensor design; the former uses CMOS technology and the latter uses CCD technology.

By interleaving different spectral samples through the use of CFA mosaic, CFA cameras sacrifice spatial resolution for color reproduction. An input image is downsampled in three or four color channels according to the chosen CFA mosaic (see [1] for a list of popular CFA mosaics in use). A color demosaicking process interpolates the downsampled image in all these channels and reconstructs a full color image at the resolution of the sensor array. Almost all existing color demosaicking algorithms work on the premise that image signal is bandlimited in both spatial and spectral domains, particularly relying on high spectral correlations. Intuitively, one would expect that as the advance of semiconduct technology steadily increases the resolution of the sensor array, demosaicking algorithms for CFA cameras will benefit from even higher spatial and spectral correlations and accordingly improve the fidelity of reconstructed images. However, there are two issues, which can, if improperly dealt with, disassociate the image fidelity from native sensor resolution and negate the folklore that "megapixels" directly translate to crisp, vivid, and clean images.

The first issue is physical limits of imaging devices. Ultra-high spatial resolution means that the pixel size (i.e., the cross section of a sensor area) diminishes, and consequently the signal-to-noise ratio (SNR) of the acquired image decreases [2]. For CFA cameras the problem is further aggravated by a phenomenon called color crosstalk [1], [3]. As pixels in CFA are densely packed, optical and electronic interferences between channels (or color crosstalk) are inevitable. Photon leakages between pixels are caused by optical refraction and electron leakages by minority carriers. The closer the neighboring pixels get in geometry, the more severe the color crosstalk becomes. The second issue is that the high spectral correlation assumption, which was made by most demosaicking methods up to now, may not hold everywhere in a natural scene. Highly saturated colors of no or weak spectral correlation do exist and they are semantically

important in visual perception and image understanding.

The effects of color crosstalk in CFA, if left untreated, on different types of color signals can be seen in Fig. 1. The figure presents simulation results that are color images reproduced from Bayer CFA [4] by the PCSD demosaicking method [5], one of the best performing demosaicking methods. As one can expect, on colors of low saturation (e.g., the cloth patterns) color crosstalk blurs image details, not unlike the point spread function (PSF) of gray-level cameras. But highly vivid colors (e.g., the flowers) can be severely desaturated by color crosstalk, which may be an even more objectionable artifact to many than simple blurring.

In the camera industry the removal of color crosstalk, called decrosstalk in this paper, is performed by adopting more sophisticated manufacturing techniques [3], [6], [7], [8], and/or by empirically chosen linear transforms. But these solutions are costly and cannot eliminate the crosstalk problem completely. Furthermore, color crosstalk of CFA data can be caused and aggravated by camera or object motions, for which there exists no hardware solution.

Even when color crosstalk of a CFA camera is sufficiently small that spectral distortions become invisible to naked eyes, research on color decrosstalk for CFA cameras is still warranted. For many applications in medicine, sciences and space, the criterion of color reproduction is beyond merely meet the eyes. The highest spectral precision is sought for it directly affects the validity of many image analysis, classification and cognition tasks.

If treated as an image restoration problem color decrosstalk is more involved than the task of conventional image deconvolution, due to spatially interleaved downsampling of spectral bands by CFA. It appears that previous researchers of color demosaicking overlooked the issue of crosstalk. Recently, Hirakawa proposed a color correction technique that reverses the effect of color crosstalk by matrix inversion [1]. But this technique is a signal-independent linear operation in tandem with demosaicking. As such, it is suboptimal when being coupled with locally adaptive demosaicking algorithms that employ spatially varying interpolators. In this work, we propose a joint demosaicking and decrosstalk technique that factors in color crosstalk when estimating missing spectral samples in CFA. Unlike in [1] color crosstalk is integrated into a process of model-based adaptive inverse filtering by being a regularization term. To suppress the noises in decrosstalk and demosaicking while preserving spatially varying image structures, we perform maximal likelihood estimation, which is facilitated by a piecewise autoregressive (PAR) image model. The new joint demosaicking and decrosstalk technique is general in that any spatially varying channel crosstalk and varying spatial-spectral correlations can be modeled and accounted for in the color reproduction.

(a) Original

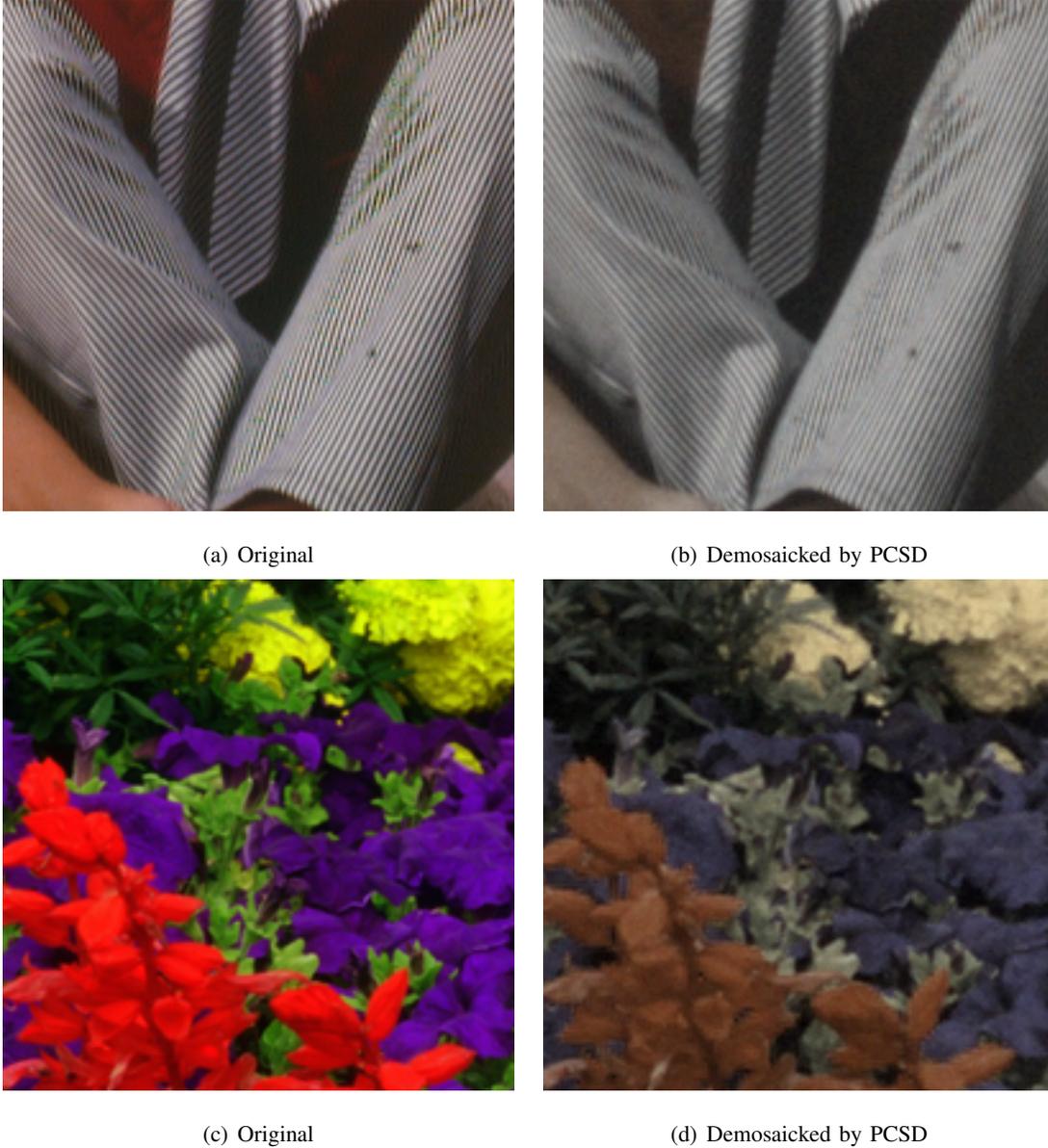(b) Demosaicked by PCSD

(c) Original

(d) Demosaicked by PCSD

Fig. 1. Effects of color crosstalk in CFA on colors of low saturation (top) and colors of high saturation (below). Both blurring of spatial details and distortions of hue and saturation are visible.

The rest of the paper is organized as follows. In section II we formulate the problem of color reproduction from CFA data contaminated by channel crosstalks. In section III we develop our joint model-based decrosstalk-demosaicking technique, and then in section IV we discuss how to adapt the model to varying spectral correlations in the image space to prevent desaturation. Experimental results are reported in section V. Section VI concludes.

II. PROBLEM STATEMENT

For the purposes of decrosstalk and demosaicking let us consider the "true" color image x to be the one produced by an ideal camera of three noise-free sensor arrays of Dirac sampling, namely, $x : \mathbb{Z}^2 \to \mathbb{R}^3$. Denote by $x(m, n) = (x_R(m, n), x_G(m, n), x_B(m, n))$, $(m, n) \in \mathbb{Z}^2$, the RGB tristimulus values at pixel location (m, n). Let $y : \mathbb{Z}^2 \to \mathbb{R}$ be the two-dimensional square lattice of sensor readings after x passes through CFA. The spatio-spectral downsampling process of CFA can be modeled as

$$y(i, j) = \sum_{m,n} \mathbf{c}(m, n)^T \mathbf{x}(m, n) h_{i,j}(i - m, j - n) + e(i, j) \tag{1}$$

where the vector $\boldsymbol{c}(m, n) = (c_R, c_G, c_B)^T$ is the RGB coordinates of the recorded color at pixel location (m, n) (i.e., the color filtered through CFA). The function $\boldsymbol{c} : \mathbb{Z}^2 \to [0, 1]^3$ is periodical, and it takes on values $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ if the CFA consists of red, green and blue samples (e.g., the popular Bayer CFA). For general CFA designs, the components of $\boldsymbol{c}$ can be arbitrary values in normalized interval $[0, 1]$ such that $c_R + c_G + c_B = 1$. The term $e(i, j)$ is an additive noise that is independent of image signal. This additive noise was considered in [1], and it is included in our model as $e(i, j)$ is inherent of CFA sensors.

Of the key role in (1) is the sensor crosstalk-convolution function (CCF) $h : \mathbb{Z}^2 \to \mathbb{R}$. The CCF specifies the combined effect of optical diffraction and minority carrier. Optical diffraction refers to the phenomenon that at a high incidence angle of the light some photons stray away from the center of the pixel in question and get recorded by neighboring sensor(s) [7], [8] As such optical diffraction is location dependent. Minority carrier refers to the fact that electrons can escape from the target after the charge is collected [6]. A precise characterization of spatially varying $h_{i,j}$ is challenging and is being pursued by researchers in microelectronics and sensor technologies. In this paper we assume CCF to be known *a prior* and focus on the image restoration problem of recovering the true color image $\mathbf{x}$ from CFA data $\mathbf{y}$.

Given the CCF h one may contemplate a cascade approach: perform conventional deconvolution of the CFA image followed by color demosaicking. But the CCF h is spatially-variant, thus a global deconvolution method, such as Weiner filtering in frequency domain, cannot work. Spatial deconvoluting of a CFA image is difficult because the CFA image contains rich high frequency contents and one cannot rely on a smoothness prior (e.g., the minimum total variation criterion) as typically done in image deconvolution. The intricate nature of the problem demands a more principled approach that performs decrosstalk and demosaicking jointly.

III. MODEL-BASED JOINT DECROSSTALK-DEMOSAICKING

It follows from (1) that recovering the true color image $\mathbf{x}$ from CFA data y is an ill-posed inverse problem. The linear system of (1) is grossly underdetermined, with the number of unknowns being three times as many as the number of observed samples. To best resolve the uncertainty we seek to regulate the estimated color image $\mathbf{x}$ with as much prior preknowledge about $\mathbf{x}$ as possible. The common preknowledge used in color demosaicking is high spectral correlation among x_r, x_g, and x_b, or the smoothness of the color difference images $d_{G,R}(m,n) = x_G(m,n) - x_R(m,n)$ and $d_{G,B}(m,n) = x_G(m,n) - x_B(m,n)$. While the assumption of low-pass $d_{G,R}$ and $d_{G,B}$ is supported by the fact that common scenes in nature comprise of pastoral colors, highly saturated colors do happen in many types of scenes (e.g., man-made signs/objects and plants/animals of vivid colors). In the presence of highly saturated colors spectral correlation becomes weak or nonexistent. Alternative and more salient regularization terms are therefore needed in solving (1) for $\mathbf{x}$. This is particularly important if we consider the effect of color crosstalk in demosaicking, otherwise severe color desaturation (see Fig. 1) can be caused by erroneously assuming high spectral correlation.

To impose stronger and more robust constraints on the solution of the inverse problem (1), we use a 2D piecewise autoregressive (PAR) model for each of the red, green and blue bands.

$$x_k(i,j) = \sum_{(s,t) \in S_k} a_k(s,t) x_k(i-s, j-t) + \epsilon_k, \quad k \in \{R, G, B\}. \tag{2}$$

For each color band k, the PAR model has parameters $\boldsymbol{a}_k$ and support S_k. The term ϵ_k is a random perturbation independent of spatial location (i,j) and the image signal x_k. There are multitudes of reasons for the use of the PAR model in joint decrosstalk and demosaicking. First, the PAR model is versatile and able to model image waveforms ranging from smooth shades, periodic textures to transients like edges. Indeed, the Gauss-Markov process of the form $X_i = -\sum_{t=1}^{p} a_t X_{i-t} + \eta_i$, where the η_i's are i.i.d. $\sim \mathcal{N}(0, \sigma^2)$, is the maximum entropy rate stochastic process, if a_t's are chosen such that $E\{X_i X_{i+t}\} = a_t$, $1 \le t \le p$, for all i (p. 274, [9]). Second, natural images are piecewise stationary, having slowly varying second-order statistics. This property makes it possible to estimate the PAR model parameters $\boldsymbol{a}_k$ in a moving window in spectral band k. Third, the PAR model is computationally efficient and it allows the maximum likelihood estimate of true color image $\mathbf{x}$ from CFA data y to be computed

by solving the following constrained linear least-square problem

$$\min_{\mathbf{x}} \sum_{(i,j)\in W} \sum_{k\in\{R,G,B\}} \left(x_k(i,j) - \sum_{(s,t)\in S_k} a_k(s,t)x_k(i-s,j-t)\right)^2$$

$$\text{subject to} \quad \sum_{(i,j)\in W} \left(\sum_{m,n} \mathbf{c}(m,n)^T \mathbf{x}(m,n)h_{i,j}(i-m,j-n) - y(i,j)\right)^2 = \sigma \tag{3}$$

where σ is the variance of the additive noise e, and W is a moving local window. The windowed solution is the ML estimate of a block of color pixels $\mathbf{x} \in W$ if the image $\mathbf{x}$ has sufficiently stationary second-order statistics in W, which holds for most natural scenes.

Now we turn to the PAR model parameters $\boldsymbol{a}_k$ in (3), $k \in \{R,G,B\}$. Since each color band k is downsampled by CFA, the model parameters $\boldsymbol{a}_k$ have to be estimated from missing data. This estimation problem is solved by the well-known iterative EM algorithm. Specifically, initial estimates $\boldsymbol{a}_k^{(0)}$ are made and plugged in (3) to solve for $\mathbf{x}$. The resulting estimated full resolution color image $\hat{\mathbf{x}}^{(0)}$ is in turn used to reestimate $\boldsymbol{a}_k$, $k \in \{R,G,B\}$, and obtain improved estimate $\boldsymbol{a}_k^{(1)}$, and so forth. The quality of the reconstructed color image largely depends on the initial condition $\boldsymbol{a}_k^{(0)}$. The initial model parameters $\boldsymbol{a}_k^{(0)}$ can be determined using the results of a good color demosaicking algorithm, ignoring the CCF at the beginning. In this paper we use the PCSD algorithm [5] to generate an initial RGB image $\mathbf{x}^{(0)} = (\tilde{\mathbf{x}}_R, \tilde{\mathbf{x}}_G, \tilde{\mathbf{x}}_B)$, and then estimate $\boldsymbol{a}_k^{(0)}$ using the second-order statistics of $\tilde{\mathbf{x}}_k$, $k \in \{R,G,B\}$ as:

$$\boldsymbol{a}_k^{(0)}(i,j) = \arg\min_{\boldsymbol{a}} \sum_{(i,j)\in W} \left(\tilde{\mathbf{x}}_k(i,j) - \sum_{(s,t)\in S_k} a(s,t)\tilde{\mathbf{x}}_k(i-s,j-t)\right)^2. \tag{4}$$

The above linear least squares problem has the closed form solution

$$\boldsymbol{a}_k^{(0)} = (\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\tilde{\mathbf{x}}_k \tag{5}$$

where the vector $\tilde{\mathbf{x}}_k$ consists of the estimated pixel values $\tilde{x}_k(i,j)$ in band k such that $(i,j) \in W$. The two quadratic terms $(\boldsymbol{D}^T\boldsymbol{D})$ and $\boldsymbol{D}^T\tilde{\mathbf{x}}_k$ are both given by the second-order statistics of $\tilde{\mathbf{x}}_k$. In matrix $\boldsymbol{D}$ each row contains the elements of $\tilde{\mathbf{x}}_k(i-s,j-t)$ for $(s,t) \in S_k$.

IV. Model Adaptation to Varying Spectral Correlations

As discussed in the previous section, the EM algorithm is applied to estimate the PAR model parameters and solve the ML problem (3). In the $(t+1)^{th}$ iteration, the PAR model parameters $\boldsymbol{a}_k^{(t+1)}$ are estimated from the three color bands $\mathbf{x}_k^{(t)}$ resulted from the tth iteration, $k \in \{R,G,B\}$. With respect to a given band k, we can choose one of the two sample sets to estimate the PAR parameter vector $\boldsymbol{a}_k^{(t+1)}$: the pixels of bank k only in local window W, or the pixels of all three bands in W. Via this choice the EM algorithm can fine tune the PAR model according to the spectral correlation in local window W.

In most areas of a natural image the colors have high spectral correlation. Indeed, a principle component analysis of the color distribution of a typical scene reveals a high degree of linear correlation between the different color components [10]. In fact, this is the statistical property that almost color demosaicking algorithms exploit to recover three color bands from CFA data. In our model-based joint decrosstalk-demosaicking technique, if high spectral correlation exists in window W, this knowledge is incorporated by making the PAR models for the R, G, and B bands the same, $\boldsymbol{a} = \boldsymbol{a}_R = \boldsymbol{a}_G = \boldsymbol{a}_B$, and updating the model parameters in the EM process as follows

$$\boldsymbol{a}^{(t+1)}(i,j) = \arg\min_{\boldsymbol{a}} \sum_{(i,j)\in W} \sum_{k\in\{R,G,B\}} \left(\mathbf{x}_k^{(t)}(i,j) - \sum_{(s,t)\in S_k} a(s,t)\mathbf{x}_k^{(t)}(i-s,j-t) \right)^2. \qquad (6)$$

In [5] the authors argued for and demonstrated the importance of keeping interpolation direction consistent in all three color bands, which they called primary consistency, in demosaicking of the so-called pastoral colors (i.e., less saturated). The primary consistency can be imposed with ease in our objective function (3) via the modeling technique described above, but it is quite difficult with other demosaicking techniques.

As we discussed earlier, although not as common, highly saturated colors do occur in daily scenes and they are semantically important. In areas of saturated colors spectral correlation is weak or nonexistent. In such cases many existing color demosaicking algorithms can produce artifacts near edges because they work on the premise of high spectral correlation. Also to compound the problem, channel crosstalk, without being corrected, can severely desaturate colors. Therefore, when there is lack of spectral correlation we have to abandon the estimation scheme (6). Instead, the PAR parameters $\boldsymbol{a}_k^{(t)}$ should be updated using the data of band k image $\mathbf{x}_k^{(t-1)}$ only. In other words, each band uses a different PAR model that is independently estimated from the other bands, namely

$$\boldsymbol{a}_k^{(t+1)}(i,j) = \arg\min_{\boldsymbol{a}} \sum_{(i,j)\in W} \left(\mathbf{x}_k^{(t)}(i,j) - \sum_{(s,t)\in S_k} a(s,t)\mathbf{x}_k^{(t)}(i-s,j-t) \right)^2, \quad k \in \{R,G,B\}. \qquad (7)$$

To choose between the interband and intraband estimated PAR models, we need to discriminate the image signal in local window W according to spectral correlation. This requires a binary classification of the CFA data in the current window W into classes of strong and weak spectral correlation. Solving the classification problem calls for the correlations $\rho_{G,R} = E\{(x_G - \mu_G)(x_R - \mu_R)\}$ between the red and green signals and $\rho_{G,B} = E\{(x_G - \mu_G)(x_B - \mu_B)\}$ between the blue and green signals in window W. However, $\rho_{G,R}$ and $\rho_{G,B}$ can only be estimated using the initially demosaicked image $\mathbf{x}^{(0)}$. The resulting estimates are unreliable in the areas of high color saturation and/or edges because $\mathbf{x}^{(0)}$ is estimated on the premise of high spectral correlation in the first place. We seek for salient features extracted directly

from the raw CFA data to build the classifier, and make two sets of measurements of CFA samples in local window W:

$$V_{G,k} = \left\{ \frac{x_G(i-1,j) + x_G(i+1,j) + x_G(i,j-1) + x_G(i,j+1)}{4} - x_k(i,j) \Big| x_k(i,j) \in W \right\} \tag{8}$$
$$k \in \{R, B\}$$

Then we use the sample variances $\sigma_{G,k}^2$ of the two sets $V_{G,k}$, $k \in \{R, B\}$, as class features. Here $\sigma_{G,k}^2$ is an estimated variance of the color difference signal $d_{G,k} = \mathbf{x}_G - \mathbf{x}_k$, and used as a proxy of $\rho_{G,k}$, $k \in \{R, B\}$.

Another important feature is the local color saturation defined below

$$s = 1 - \frac{3}{\bar{x}_R + \bar{x}_G + \bar{x}_B} \min(\bar{x}_R, \bar{x}_G, \bar{x}_B)$$
$$\bar{x}_k = \frac{1}{|W|} \sum_W x_k, \quad k \in \{R, G, B\} \tag{9}$$

Consider the binary event Y that signifies whether the interband estimation scheme (7) or the intraband scheme (6) offers a more accurate color reproduction of $\mathbf{x} \in W$. Having the three local features s, $\sigma_{G,R}$, and $\sigma_{G,B}$, we can learn the conditional probability $P(Y|s, \sigma_{G,R}, \sigma_{G,B})$ from a training set. Then the above decision problem can be solved by a maximum a posterior probability technique, or a minimum-risk Bayesian-type classifier [11].

In addition to their role in adapting the PAR parameters $\mathbf{a}_k$ to spatially varying spectral correlations, $\sigma_{G,R}^2$ and $\sigma_{G,B}^2$ may be used to judiciously include extra regularization term(s) into the objective function (3). If $\sigma_{G,k}^2$ is small in W, indicating the local smoothness of the color difference signal $d_{G,k} = \mathbf{x}_G - \mathbf{x}_k$, then the reconstruction of $\mathbf{x}$ can benefit from the knowledge that the magnitude of the Laplacian $\|\nabla^2 d_{G,k}\|$ is small. Putting all the above factors together, our decrosstalk-demosaicking algorithm development culminates to

$$\min_{\mathbf{x}} \{ \|(\boldsymbol{A}(\sigma_{G,R}, \sigma_{R,B}, s) - \boldsymbol{I})\mathbf{x}\|_2 + \lambda_R \|\nabla^2 d_{G,R}\|_2 + \lambda_B \|\nabla^2 d_{G,B}\|_2 \}$$
$$\text{subject to} \quad \|(\mathbf{c}^T \mathbf{x}) * \boldsymbol{h} - \mathbf{y}\|_2 = \sigma \tag{10}$$

where $\mathbf{x} \in \mathbb{N}^{3|W|}$ is the vector of all color samples of local window W in a lexicographical order, and $\mathbf{y} \in \mathbb{N}^{|W|}$ is the vector of CFA samples in W. Matrix $\boldsymbol{A}(\sigma_{G,R}, \sigma_{R,B}, s)$ specifies the PAR model that is adapted to spatial and spectral correlations of $\mathbf{x} \in W$ as discussed above. The Lagrangian multiplier λ_k, $k \in \{R, B\}$, is set to monotonically decrease as $\sigma_{G,k}$ increases.

V. SIMULATION RESULTS

Simulations are carried out to evaluate the performance of the proposed joint decrosstalk-demosaicking method. For the generality and validity of the experimental results, we choose a test image set that includes a broad range of color contents from low to high saturation (i.e., high to low spectral correlations). Fig. 2 lists five sample images in our test set. The CCF used in the simulation is a 2D Gaussian kernel with standard deviation of 0.4, and the noise is simulated by a Gaussian random noise with standard deviation of 0.01 (with the pixel values normalized in the range of [0,1]). This CCF is very close to the crosstalk kernel function given by experts on camera sensors [3]. Our joint decrosstalk-demosaicking technique is compared with the color correction method in [1], and with the method of performing decrosstalk (deconvolution) and demosaicking in tandem. The last approach is referred to as the separation method. To make the comparison fair, we assume the CCF to be space invariant so that the frequency domain deconvolution can be applied to decrosstalk, although the proposed joint technique can accommodate any space variant CCF. In the separation approach, a CFA images is first deconvoluted by regularized least squares (RLS) method or the Lucy-Richardson method, with known CCF and noise power. Then the deconvoluted CFA image is demosaicked by the PCSD algorithm to reconstruct the color images. For the color correction method, the CFA image is first demosaicked by the PCSD algorithm. Then the color is corrected with the color correction matrix in [1].

The PSNR results of the reconstructed images in the RGB color space by different methods are listed in Table I. The results of the separation methods are listed in columns "RLS-PCSD" and "Lucy-PCSD", according to the deconvolution method used. The approach with color correction method in [1] is denoted as "PCSD-[1]". Among all the methods, the PCSD method operating on CFA data directly ignoring CCF has the lowest PSNR, which manifests the negative impacts of channel crosstalk. The separation methods and the method PCSD-[1] get higher PSNR than PCSD by trying to reverse the CCF, but they underperform the proposed technique on all test images. The gain made by the proposed technique over the second best method, the Lucy-PCSD method, is up to 5dB. To make the evaluation more meaningful in terms of human perception and in particular to measure the distortion in color saturation, we also list the PSNR results in the HSL (Hue, Saturation, Light) color space in Table II. We call readers' attention to the fact that the new technique achieves significantly higher precision than its competitors in color saturation. For instance, on test image Flower the PSNR value of the proposed technique in saturation is $5 \sim 8$ dB higher than the other methods. The data of Table II demonstrate our success in the prevention of color desaturation due to crosstalks.
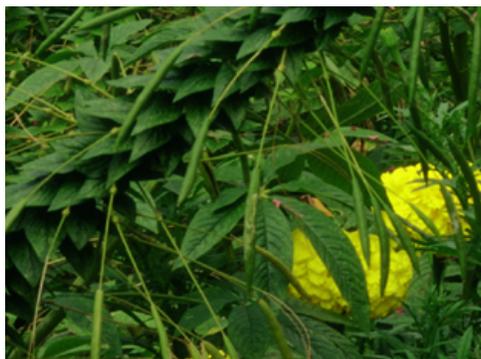
(a) Barb

(b) Flower

(c) Leaves

(d) Bike

(e) Bush

Fig. 2.   Five example images in our test set.

TABLE I

PSNR (dB) RESULTS OF DIFFERENT METHODS IN R, G, B BANDS.

Image	Channel	Methods				
		PCSD	RLS-PCSD	Lucy-PCSD	PCSD-[1]	Proposed
Barb	R	31.06	31.67	32.17	32.16	**33.18**
	G	33.98	34.47	34.41	34.27	**36.35**
	B	31.49	32.14	32.21	32.18	**33.02**
Flower	R	23.01	28.05	29.18	29.16	**31.58**
	G	26.16	30.88	32.50	32.52	**34.32**
	B	23.35	26.52	27.23	27.37	**31.79**
Leaves	R	29.11	30.12	31.68	31.26	**32.03**
	G	30.79	30.83	32.86	32.21	**35.97**
	B	27.35	28.44	29.75	29.49	**30.20**
Bike	R	31.29	32.48	34.08	33.99	**34.43**
	G	33.13	33.25	35.70	35.29	**35.97**
	B	28.45	30.12	33.36	33.26	**33.44**
Bush	R	30.89	32.13	33.14	33.11	**34.22**
	G	29.22	32.53	34.42	34.37	**34.44**
	B	26.47	29.84	31.39	31.52	**35.25**

To assess the visual quality of the different methods, we show the reconstructed images by different methods in Fig. 3. The desaturation effect caused by the crosstalk in CFA is clearly visible in the PCSD results for all the test images, especially highly saturated colors (e.g., the red flowers in Flower image and the green leaves in Bush image in Fig. 3). The Lucy-PCSD method, which performs a deconvolution of CFA prior to PCSD demosaicking, alleviates the problem of color desaturation to an extent but its color reproduction is visibly inferior to PCSD-[1] and the proposed method in some cases. For examples, in Fig. 3 the colors of the Bush and Flower images reproduced by the proposed method appear more vibrant and closer to the original than the Lucy-PCSD method. Another noticeable difference between the proposed method and the Lucy-PCSD method is that the latter suffers from speckle color noises near edges while the former is immune from the problem. This is because in separation approach any signal-dependent deconvolution noises will be magnified by the demosaicking process. Method [1] reproduces tones quite well, but it preserves spatial details not as sharp and clean as the proposed method. This can be seen by comparing the Bush images reconstructed by PCSD-[1] and the proposed method in Fig. 3. In

TABLE II

PSNR (dB) RESULTS OF DIFFERENT METHODS IN HUE, SATURATION AND LUMINANCE COMPONENTS.

Image	Channel	Methods				
		PCSD	RLS-PCSD	Lucy-PCSD	PCSD-[1]	Proposed
Barb	H	11.06	10.91	11.18	11.06	**12.08**
	S	25.57	26.01	26.30	26.22	**26.95**
	L	33.83	34.62	34.98	34.91	**36.16**
Flower	H	21.05	21.32	21.30	21.56	**22.98**
	S	13.82	17.07	17.56	17.61	**22.54**
	L	28.59	30.97	31.39	31.43	**35.25**
Leaves	H	9.51	9.53	9.61	9.51	**9.81**
	S	16.58	15.71	15.02	14.36	**18.61**
	L	34.16	34.55	34.09	33.65	**35.17**
Bike	H	19.10	18.96	19.18	18.76	**19.68**
	S	18.41	20.14	22.14	21.92	**22.85**
	L	34.48	34.67	36.14	36.00	**36.44**
Bush	H	25.10	25.93	26.90	26.36	**27.44**
	S	14.09	16.68	17.12	17.28	**20.83**
	L	34.97	34.74	34.69	34.71	**37.07**

summary, the proposed method appears to achieve the best visual quality by preserving both color fidelity and spatial structures better than others in the comparison group, while largely free of demosaicking color noises.

For cameras of lower quality whose CCF is flatter, or when CFA data are degraded by motions and thus the channel crosstalk is compounded by motion blur, the color de-saturation effect can become more severe. Two such examples are given in Fig. 4 , where the CCF used in simulation is a Gaussian low pass filter with standard deviation 0.6. In this case the color fidelity of the Lucy-PCSD method deteriorates to a level of unacceptable tone misrepresentation, while the proposed technique still manages to recover the vivid colors in correct tone. Reflecting the above observation, the gap between the Lucy-PCSD method and the proposed method in PSNR becomes much larger (the PSNR values of different methods in the RGB space are listed in Table III). Method [1] recovers the color better than Lucy-PCSD method, but is inferior to the proposed method, being prone to color noises as those in the Bike images shown in Fig. 4.

Fig. 3. Comparison of different methods for Gaussian CCF of standard deviation of 0.4 and noise of $\sigma = 0.01$. From top to bottom: PCSD, Lucy-PCSD, PCSD-[1], the proposed method.
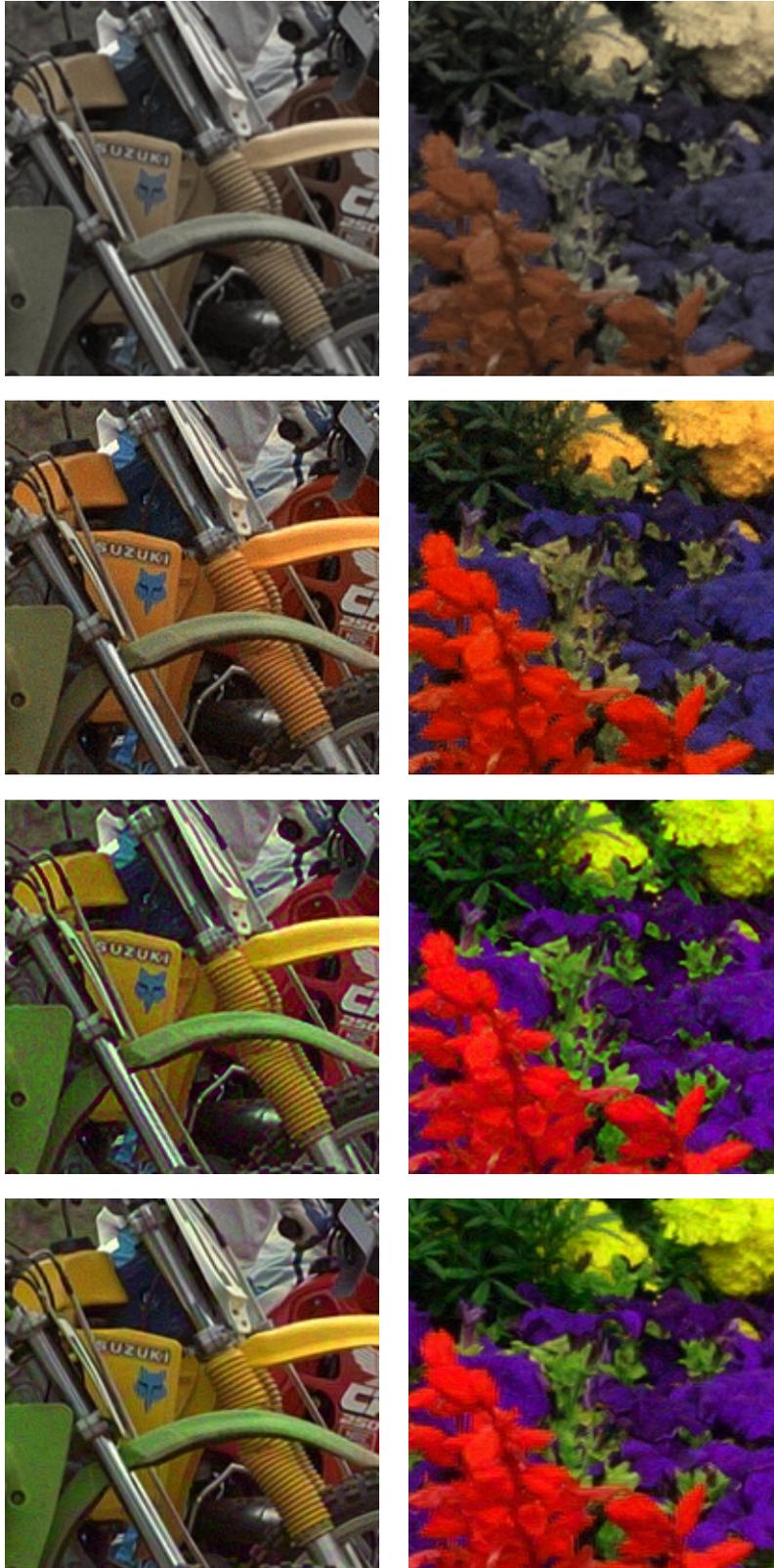
Fig. 4. Comparison of different methods for Bike and Flower images in presence of severe channel crosstalk (Gaussian CCF of standard deviation 0.6 and noise of $\sigma = 0.01$). From top to bottom: PCSD, Lucy-PCSD, PCSD-[1], the proposed method.

TABLE III

PSNR (dB) RESULTS OF DIFFERENT METHODS IN R, G, B BANDS (GAUSSIAN CCF OF STANDARD DEVIATION 0.6).

Image	Channel	Methods			
		PCSD	Lucy-PCSD	PCSD-[1]	Proposed
Bike	R	22.43	26.69	27.63	**31.31**
	G	25.34	26.72	25.93	**32.07**
	B	19.03	26.01	27.75	**30.30**
Flower	R	13.51	20.68	26.95	**27.11**
	G	17.68	20.93	27.86	**28.98**
	B	15.04	20.35	26.52	**27.15**

The proposed new technique is more robust than others when the noise in (1) becomes large. An example is shown in Fig. 5, where the standard deviation of the noise increases to 0.04 (with the pixel values normalized in the range of [0,1]). Comparing our result with those of other methods, one can see the side benefit of denoising of our joint demosaicking and decrosstalk approach. The robustness of the proposed method is improved over [1] by including the noise energy term in the constraint in (10), rather than assuming that the CFA image $\mathbf{y}$ relates the true color image $\mathbf{x}$ only via CCF.

In our simulation the CCF is assumed to be known. If an estimated CCF is used, the error in CCF can adversely affect the performance of the proposed algorithm. Fig. 6 shows what happens when there is 12.5% error in CCF. The error causes the proposed technique a loss of 6.1dB in PSNR comparing with that in Fig. 3(k). Nevertheless, our technique is still significantly better than PCSD. The blind version of our problem of joint decrosstalk and demosaicking when the CCF is unknown, either due to lack of a good CFA camera model or/and unknown camera/object motions, is an interesting and challenging extension of blind image deconvolution. The research on this is under way.

VI. CONCLUSION

A joint demosaicking and decrosstalk technique is proposed to reproduce colors from CFA data contaminated by channel crosstalks. The operations of deconvolution for crosstalk removal and interpolation for color demosaicking are integrated into a unified image restoration process. A piecewise autoregressive multispectral image model is used to accommodate spatially varying channel crosstalks and varying spatial-spectral correlations. Simulation results demonstrate the efficacy of the proposed technique and its advantage over the separation approach of performing deconvolution and demosaicking in cascade.

(a) PCSD(R/G/B: 22.01/24.34/22.21 dB)   (b)   Lucy-PCSD   (R/G/B: 25.37/27.06/24.68 dB)

(c) PCSD-[1](R/G/B: 26.00/27.83/25.43 dB)   (d) The proposed method (R/G/B: 26.20/27.96/26.40 dB)

Fig. 5. Reconstructed image Flower in presence of large noise (standard deviation 0.04).

REFERENCES

[1] K. Hirakawa, "Cross-talk explained," in *Proc. IEEE Image Processing*, Oct. 2005, pp. 677–680.

[2] H. Stark and P. Oskoui, "High-resolution image recovery from image-plane arrays using convex projections," *J. Opt. Soc. Am. A*, vol. 6, no. 11, pp. 1715–1726, Nov. 1989.

[3] H. Tian, Q. Sun, J. Li, and A. Domine, "Crosstalk challenges cmos sensor design," *LaserFocusWorld*, vol. 41, no. 1, Jan. 2005.

[4] B. Bayer and E. K. Company, "Color imaging array," U.S. Patent 3971015, Jul. 20 1976.

[5] X. Wu and N. Zhang, "Primary-consistent soft-decision color demosaicking for digital cameras," *IEEE Tran. Image Processing*, vol. 13, no. 9, pp. 1263–1274, 2004.

[6] G. Agranov, V. Berezin, and R. H. Tsai, "Crosstalk and microlens study in a color CMOS image sensor," *IEEE Tran. Electron Devices*, vol. 50, no. 1, pp. 4–11, Jan. 2003.

[7] H. Rhodes, G. Agranov, C. Hong, U. Boettiger, R. Mauritzson, J. Ladd, I. Karasev, J. McKee, E. Jenkins, W. Quinlin,
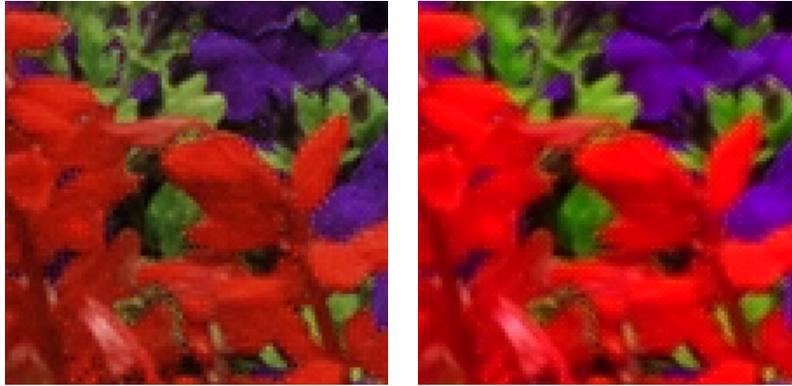
(a) PCSD(R/G/B: 23.01/26.16/23.35 dB) (b) The proposed method (R/G/B: 25.02/27.94/26.42 dB)

Fig. 6.   Reconstructed image Flower when there is a 12.5% error in CCF.

I. Patrick, J. Li, X. Fan, R. Panicacci, S. Simith, C. Mouli, and J. Bruce, "CMOS imager technology shrinks and image performance," *IEEE Microelectronics and Electron Devices*, pp. 7–18, 2004.

[8] I. Shcherback, T. Danov, and O. Yadid-Pecht, "A comprehensive CMOS APS crosstalk study: Photoresponse model, technology, and design trends," *IEEE Trans. Electron Devices*, vol. 51, no. 12, pp. 2033–2041, Dec. 2004.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*.   John Wiley & Sons Canada, Ltd., 1991.

[10] X. Wu, "Color quantization by dynamic programming and principal analysis," *ACM Transactions on Graphics*, vol. 11, no. 4, pp. 348–372, Oct. 1992.

[11] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 2nd ed.   Academic Press, Feb. 2003.