Joint Image Denoising and Demosaicking by Low Rank Approximation and Color Difference Model

Xia Zhai¹, Weiwei Guo², Yongqin Zhang³, Jinsheng Xiao⁴, and Xiaoguang Hu⁵

¹Arts Department, Henan University of Economics and Law, China
²Department of Information and Communication Engineering, University of Electronic Science and Technology, China
³School of Information Science and Technology, Northwest University, China
⁴School of Electronic Information, Wuhan University, China
⁵School of Criminal Science and Technology, People’s Public Security University of China, China

Abstract: Digital cameras generally use a single image sensor which surface is covered by a color filter array. The Color Filter Array (CFA) limits each sensor pixel by sampling one of the three primary color values (red, green or blue), whereas the other two missing color values would be acquired by the post-processing procedure called demosaicking. From the noisy CFA data, the full color images are reconstructed through an imaging pipeline of demosaicking and denoising. However, image denoising in the RGB space has expensive computation cost. In this paper, to increase the efficiency and the color fidelity, we propose a novel joint denoising and demosaicking strategy to reconstruct the noiseless full color image from the input noisy CFA data. The low-rank approximation technique is first used to remove the noise from CFA data. Then, image demosaicking using both color difference space and signal correlation are applied to the denoised CFA data to obtain the noise-less full color image. The experimental results show that the proposed algorithm not only improves the quality of full color image but also outperforms the existing state-of-the-art methods both subjectively and objectively.

Keywords: Image denoising, image demosaicking, CFA, low rank approximation, color difference model.

Received May 8, 2013; accepted June 1, 2014

1. Introduction

The digital cameras are important imaging equipment for image processing and computer vision. The full color digital images are usually composed of three different spectral components, i.e., red, green and blue. It is desirable of having three Complementary Metal Oxide Semiconductor (CMOS)/Charge Coupled Device (CCD) image sensors to capture full color digital images. However, in order to decrease the production cost, most digital color cameras only use a single sensor with only one Color Filter Array (CFA) instead of three. Therefore, the color interpolation or demosaicking is applied to the CFA data to reconstruct the full color images. Generally speaking, image denoising is carried out after image demosaicking in many applications. In the real world, the noise of digital cameras mainly comes from electrical components. The pixel sensor of the CMOS devices typically consists of a photodiode and three transistors, which are the major sources of noise. Consequently, image denoising is an essential preprocessing step to reconstruct the full color image. In recent decades, many image denoising methods have appeared and received considerable attention, such as wavelet-based methods [8], the bilateral filtering [9], the non-local means [2], the low rank technique [15] and the others [1, 10, 14]. However, these methods did not consider that image demosaicking has the amplification effects of noise in CFA data, which caused the blurred edges and smoothed textures of the output images.

Image demosaicking plays a very important role to generate full color images. So far, great progress has been made in theoretical research of image demosaicking and many methods have been introduced into the studies. They can be mainly divided into two categories. One is directly to interpolate each color channel separately, which produces good results when applied to gray-scale images. Unfortunately, when applied to color images, they often lead to visible color artifacts in the reconstructed images. The other is to interpolate through the signal correlation among the color channels [4, 6, 11], which has better performance in color fidelity. Recently, many demosaicking methods have been proposed. Gunturk et al. [4] exploited the inter-channel correlation in an alternating projection scheme, which outperforms the nearest-neighbor replication, bilinear interpolation and cubic spline interpolation. Yuk et al. [11] developed another demosaicking method that selects direction adaptively for interpolation using the direction similarity in different color spaces. However, these demosaicking methods often lead to color artifacts in reconstructed images. In fact, the acquired CFA data is generally noisy, which causes the interpolation error.
and color distortion. However, most demosaicking methods are used to reconstruct the full color images on the assumption that the CFA data is noise-free. In order to make the assumption more realistic, some researchers have put forward novel methods that generate the full color image from the noisy CFA signal [3, 5, 7, 12, 13, 16]. Hirakawa and Parks [5] proposed a joint of demosaicking and denoising algorithm. They treated both demosaicking and denoising as estimation problems, and proposed an optimal image filter based on the Total Least Square (TLS) technique. Zhang et al. [13] exploited both spectral and spatial correlations to suppress sensor noise and interpolation error simultaneously by using the Linear Minimum Mean Square-Error (LMMSE) technique. Menon and Calvagno [7] adopted the space-varying filters and the evaluation of noise autocorrelation to reconstruct the full resolution color image. Zhang et al. [12] developed a spatially adaptive method based on Principle Component Analysis (PCA), which worked on the noisy CFA data by supporting window to analyze the local image statistics. Condat and Mosaddegh [3] presented a joint demosaicking and denoising framework by total variation minimization to construct the full color image.

In this paper, we present a united denoising and demosaicking framework based on both the low rank approximation and color difference model. The proposed algorithm employs a low rank approximation strategy for the local patches to achieve resolution enhancement and noise removal. The low rank approximation adopts the refined parallel analysis with Monte Carlo simulation to select signal subspace dimensionality of the noisy patches. In order to obtain the full color image, the adaptive edge-preserving and edge-directed interpolation method is proposed for color reproduction from the denoised CFA data. Finally, according to the image observation model, the simulation results of the synthetic images with variant noise levels demonstrate that the proposed algorithm outperforms the state-of-the-art methods both quantitatively and qualitatively.

The remainder of this paper is organized as follows. In section 2, we briefly review the image observation model, and then describe the proposed algorithm in detail. Section 3 gives the experimental evaluations and comparisons. At last, section 4 draws the conclusions.

2. Proposed Algorithm

To produce high quality images acquired by digital cameras, many methods have been proposed to reconstruct the full color images. There are mainly three schemes to reconstruct the full color image from the noisy CFA data. One is demosaicking first and followed denoising, the second is joint demosaicking and denoising together, and the third is denoising first and demosaicking later. For the first scheme, it is still a challenging problem in image processing. The noise in CFA data will affect the demosaicked results and caused color artifacts, which is difficult to be removed in the following denoising processing. The first scheme also has the drawbacks of high computational cost. Consequently, the second and third schemes are generally used to derive the full color images. In this paper, we will only study the third scheme to build the full color images.

2.1. Image Observation Model

The mathematical model of the observed noisy images [5] is given as follows:

\[
y(i,j) = X(i,j) + k + k X(i,j) \quad (1)
\]

Where \( x(i, j) \) and \( y(i, j) \) are separately the ideal and measured sensor values at pixel location \((i, j)\), \( \delta \sim N(0,1) \) is the noise, \( k_0 \) and \( k_1 \) are sensor dependent parameters. Specifically, the noisy color images can be described in another formula [13]:

\[
\hat{r} = r + v_r, \quad \hat{g} = g + v_g, \quad \hat{b} = b + v_b \quad (2)
\]

Where \( v_r, v_g \) and \( v_b \) are the noise in the red, green and blue channel, \( r, g \) and \( b \) are the ideal values, \( \hat{r}, \hat{g} \) and \( \hat{b} \) are the noisy signals. In fact, the acquired images from the CCD/CMOS sensors often contain the additive and/or multiplicative noises. However, only noisy CFA image data is available in practice. In order to reconstruct the noiseless full color image, we propose the joint denoising and demosaicking algorithm that employs the low rank approximation and color difference model for the noisy CFA data. The proposed blind algorithm mainly consists of two steps: The CFA denoising and image demosaicking.

2.2. CFA Denoising Step

Different from the existing patch-based denoising methods, the proposed algorithm utilizes the projections of local patch stacks onto signal subspace. For each patch of the CFA data, the signal dimensionality was determined by the refined parallel analysis with Monte Carlo simulation. After the projection of each patch onto its signal subspace, the denoised CFA data is acquired by patch alignment and the averaging method.

First of all, the denoising step is implemented to remove noise of the noisy CFA data. Assume that \( x(i, j) \) is the true pixel value at the spatial location \((i, j)\) of the observed image \( X \in \mathbb{R}^{M \times N} \) and \( Y \) is a patch-based representation of \( X \). Those \( \sqrt{P} \times \sqrt{P} \) pixels around the center pixel \( x(i, j) \) in the image \( X \) constitute a reshaped column vector from \( Y \). That is to say, \( Y \) is a matrix of
size $M \times N \times P$, which consists of the pixels in the image $X$.

Suppose that, the image patch centered at $x(i, j)$ is in a $\sqrt{L} \times \sqrt{L}$ sliding search window. The image matrix $X$ will be divided into $M_D \times N_D$ image blocks of the same size $\sqrt{L} \times \sqrt{L}$. To ensure that the search window covers all the related image blocks, we define $\sqrt{L} \sqrt{L} + \sqrt{P} - 1$. For each image block, its similar patches are found by block matching method. Then these similar patches are reorganized as the matrix $Z_{m,n}$ with the size of $L \times P$, where $m=1, ..., M_D$ and $n=1, ..., N_D$.

According to Singular Value Decomposition (SVD) of the similar patch group, the most energy of the image signal concentrates on the first few principal components, while the energy of noise will evenly spread over the whole components. Hence, the optimal dimensionality of PCA can be used for noise removal. By removing the mean value from each column vectors, the zero-centered image matrix is given as:

$$\overline{Z}_{m,n} = Z_{m,n}(l,p) - \frac{1}{L} \sum_{l=1}^{L} Z_{m,n}(l,p)$$

Where $l=1, ..., L$, $p=1, ..., P$, $m=1, ..., M_D$, $n=1, ..., N_D$. To reduce the calculation time on the condition of $L>P$, we replace $\overline{Z}_{m,n}$ with the covariance matrix $\overline{Z}_{m,n}^T \overline{Z}_{m,n}$. Thus, we have:

$$\overline{Z}_{m,n}^T \overline{Z}_{m,n} U = U^2 V^T$$

Where $U = [u_1, u_2, ..., u_p]$ is the $P \times P$ unitary matrix consisted of eigenvectors derived from $\overline{Z}_{m,n}^T \overline{Z}_{m,n}$. Define $\Sigma = diag \{ \lambda_1, \lambda_2, ..., \lambda_p \}$ as the $P \times P$ diagonal matrix, $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p \geq 0$, $r = \text{rank}(\overline{Z}_{m,n}^T \overline{Z}_{m,n})$, where $\lambda_1, ..., \lambda_r$ are the singular values of $\overline{Z}_{m,n}^T \overline{Z}_{m,n}$. By this projection, $\overline{Z}_{m,n}^T \overline{Z}_{m,n}$ can be decorrelated, and $\hat{Z}_{m,n}$ is projected onto the basis $U$, the reformed uncorrelated matrix $\hat{Z}_{m,n}$ is given as:

$$\hat{Z}_{m,n} = \overline{Z}_{m,n} \times U$$

According to principle component analysis, we find that the images can be reconstructed only by several important components. The output images would not lose much information if the small singular values are ignored and the image patch size is large enough. Therefore, we can only use the first $K$ eigenvectors to reconstruct the noiseless CFA data, which not only removes the noise but also preserves the most image information. In our algorithm, we adopt the refined parallel analysis with Monte Carlo simulation to select the signal components.

Let the singular values of the image patch $\overline{Z}_{m,n}$ be denoted by $\lambda_p$ sorted in the descending order, where $p=1, ..., P$. Assume $\alpha_p$ as the singular values of the artificial data. So, the parallel analysis can estimate the signal dimensionality of noisy image data as follows:

$$K = \max \{ p \neq 1 \alpha_p, P \}$$

Where $\alpha_p$ is a threshold and the singular value $\lambda_p$ less than $\alpha_p$ is discarded to remove noise.

In order to estimate the signal subspace dimensionality from noisy image data, we generate the artificial data with the same size of the image patches. Let $C \in \mathbb{R}^{L \times P}$ denote the artificial matrix. Similarly, we use $C^T C$ instead of $C$ for the projection of the artificial data. The SVD of this artificial data matrix is given as:

$$C^T C = V S V^T$$

Where $V$ is the unitary matrix of eigenvectors derived from the artificial data matrix, and $S = \text{diag} \{ \beta_1, \beta_2, ..., \beta_p \}$ is a diagonal matrix with its singular value in the descending order. Let $\beta_p$ for $p=1, ..., P$ denote the singular values of the artificial data matrix. Then we compare $\lambda_p$ with $\beta_p$ to determine the signal subspace dimensionality:

$$K = \max \{ p \neq 1 \beta_p, P \}$$

Through the refined parallel analysis, the signal subspace dimensionality is determined so that we can reconstruct the noiseless image. For each image block, the straightforward way to reconstruct a noiseless image block is to project the noisy column vectors directly onto the subspace of the first $K$ eigenvectors in a low rank representation. Hence, each projected weight matrix of signal subspace is:

$$W_{m,n}(k,p) = \hat{Z}_{m,n}(l,k) / \hat{Z}_{m,n}(l,p)$$

Where $k=1, 2, ..., K$ and $p=1, 2, ..., P$. Then, the noiseless image patch stack $R_{m,n}$ is acquired by the weighted subspace as follows:

$$R_{m,n} = \hat{Z}_{m,n}(l,k) W_{m,n}(k,p) + \frac{1}{L} \hat{Z}_{m,n}(l,p)$$

Finally, we can reconstruct the noiseless image $F$ by patch alignment technique and overlap averaging scheme for the whole image. Each pixel value at the position $(i, j)$ would be described by this formula:

$$F(i, j) = \frac{1}{n_i} \sum_{m,n,p} R_{m,n}(l,p)$$

Where $n_i$ is the number of pixels used in the image patch stacks for the whole image.

Our proposed denoising algorithm is used to reduce the noise of the CFA data in the blind condition. In order to reconstruct a full color image, the proposed interpolation step called demosaicking is performed on the denoised CFA data.
2.3. Image Demosaicking Step

Most consumer cameras capture images by a Bayer CFA pattern that only one color value is sampled at the location of each pixel. The Bayer pattern is shown in Figure 1, which provides an array of red (r), green (g) and blue (b) color components. Since, only a single color component is available at each spatial location, the demosaicking is used to interpolate the other two missing color components.

Yuk et al. [11] proposed a simple demosaicking method which interpolates in color difference space. In order to avoid interpolation across the edge and reduce color artifacts, it interpolates color images using the direction similarity. However, this method ignores that the adjacent pixels have different correlation along the different directions. To reduce color artifacts, another demosaicking method [16] exploits the adaptive spatial correlation to interpolate the pixels in the color difference space for the full color image reconstruction. In our algorithm, we employ the color difference model for image demosaicking due to its low contrast.

Assume that $R$, $G$ and $B$ are the desired image color values obtained by a digital camera, and $R'$, $G'$ and $B'$ are interpolated values. Their normalized values are between 0 and 1. We separately compute the differences between color channels, i.e., $K_R$ (green minus red) and $K_B$ (green minus blue), and then interpolated them. Thus, the full color image is reconstructed by the interpolation method.

Firstly, we describe how to compute $K_R/K_B$ at the red/blue pixels. According to the structural similarity, the interpolation direction should be determined before the interpolation operation. Let $L$, $R$, $U$ and $D$ is the index of $R$, $G$ and $B$. $H$ and $V$ denote the image’s horizontal and vertical interpolation direction, respectively.

As is shown in Figure 2, we interpolate the pixel at the location $R$, the similarity between horizontal and vertical directions can be calculated as:

$$
\begin{align*}
V &= K_x - K_{x_{2 \downarrow_2}} + K_x - K_{x_{2 \uparrow_2}} \\
H &= K_x - K_{x_{2 \downarrow_2}} + K_x - K_{x_{2 \uparrow_2}}
\end{align*}
$$

(12)

After computing the similarity between horizontal and vertical directions, we select the dominant direction for the interpolation. That is, if their values are comparatively different, the direction of interpolation is selected according to smoother direction. Otherwise, we should select a bigger window to determine the direction of interpolation. In our algorithm, if $|H-V| < c |H+V|$, then we will choose the bigger window and calculate $H$ and $V$ again as follows:

$$
\begin{align*}
V &= |K_x - K_{x_{2 \downarrow_2}}| + K_{x_{2 \downarrow_2}} - K_{x_{2 \uparrow_2}} + |K_x - K_{x_{2 \uparrow_2}}| \\
H &= |K_x - K_{x_{2 \downarrow_2}}| + K_{x_{2 \downarrow_2}} - K_{x_{2 \uparrow_2}} + |K_x - K_{x_{2 \uparrow_2}}|
\end{align*}
$$

(13)

If $H>V$, $K_R=0.5(G_{U+} + G_{D0})-0.25(R_{U\downarrow} + 2R + R_{D\uparrow})$, otherwise, $K_R=0.5(G_{L+} + G_{B})-0.25(R_{L\downarrow} + 2R + R_{B\uparrow})$.

The estimation of $K_B$ at $B$ pixels is performed in the similar way. Next we compute $K_R/K_B$ at the blue/red pixels. Here, we take the estimation of $K_R$ for example. According to Figure 3, we compute $K_R$ at $R$ pixels with diagonal neighboring locations in the following form:

$$
K_x = 0.25(K_{x_1} + K_{x_2} + K_{x_3} + K_{x_4})
$$

(14)

Next, we compute $K_R/K_B$ at the green pixels. Referring to Figure 4, since the interpolated $K_R/K_B$ values at the blue/red pixels are available by the previous procedures, we calculate $K_R/K_B$ at green pixels as the weighted sum of vertical and horizontal direction similarity. The process is approximately similar with the computation of $K_R/K_B$ at red/blue pixels.

$$
\begin{align*}
V &= 0.5|K_{y_{2 \downarrow_2}} - K_{y_{2 \uparrow_2}}| + 0.5|K_{y_{1 \downarrow_2}} - K_{y_{1 \uparrow_2}}| \\
H &= 0.5|K_{y_{1 \downarrow_2}} - K_{y_{1 \uparrow_2}}| + 0.5|K_{y_{1 \downarrow_2}} - K_{y_{1 \uparrow_2}}|
\end{align*}
$$

(15)
If \(|H-V|<c \) \(|H+V|\), we choose a larger window and compute \(H\) and \(V\) again. Then, we have:

\[
K_x = \begin{cases} 
0.5(K_{x_{d1}} + K_{x_{r1}}), & \text{if } H > V \\
0.5(K_{x_{e1}} + K_{x_{r1}}), & \text{otherwise}
\end{cases}
\]  
(16)

The same process is performed to compute \(K_y\) values at green pixels.

After all \(K_R/K_B\) are obtained at each pixel, we calculate the missing pixel values to generate the full color image. Finally, at \(R\) channel, the pixel value is:

\[
R' = R, \quad G' = R + K_x, \quad B' = G' - K_y
\]  
(17)

At \(G\) channel, the pixel value is:

\[
G' = G, \quad R' = G - K_x, \quad B' = G - K_y
\]  
(18)

At \(B\) channel, the pixel value is:

\[
B' = B, \quad G' = G + K_y, \quad R' = G' - K_x
\]  
(19)

### 3. Simulation Results

To evaluate the performance of the proposed algorithm comprehensively, the experiments were carried out. A set of test images is selected from the well-known Kodak dataset. In our experiments, we compare our proposed algorithm with the PCA-based spatially adaptive denoising [12] and the joint demosaicing and denoising [5]. The performances of these different methods were evaluated on the test images by Peak Signal-To-Noise Ratio (PSNR) and Structural Similarity (SSIM).

In the experiment, the test color images were mosaicked with the Bayer CFA and corrupted by the speckle noise model with different deviations 0.0100, 0.0225 and 0.0400, respectively. The parameters of our proposed algorithm is set with the variable search window of 25×25 pixels, the image block of 17×17 pixels and the neighborhood patch of 7×7 pixels for noise removal. Moreover, the color demosaicking with adaptive window size is used to reconstruct the full color images from the CFA data. According to the similarity between horizontal and vertical directions, the demosaicking step adaptively selects one of three different windows of size 9×9, 13×13 and 17×17.

Figure 5 shows the simulated noisy color images generated from ‘kodim21’ corrupted by the speckle noise with different deviations 0.0100, 0.0225 and 0.0400, respectively. The visual comparison of these different methods for the test image ‘kodim21’ is given in Figure 6. For a set of test Kodak color images, the PSNR(dB)/SSIM results are presented in Table 1. As can be seen from the results, the proposed algorithm has less color artifacts and sharper edges. The experimental results demonstrate that our proposed algorithm can suppress the noise and preserve the image details efficiently, and outperforms the state-of-the-art methods both objectively and subjectivel.
4. Conclusions

In this paper, we present a joint denoising and demosaicing algorithm to reconstruct the full color image from noisy CFA data in the blind condition. The proposed algorithm is based on the low rank approximation and color difference model. The experimental results show that our proposed algorithm can achieve higher resolution, less noise and sharper edges. For noisy CFA data, our proposed method not only can remove the noise and the color artifacts effectively, but also can recover rich textures and details. It is also superior to other state-of-the-art methods both quantitatively and qualitatively.

References


**Xia Zhai** received the Bachelor’s degree in Art and Design in 2011 from Guangzhou Academy of Fine Arts, P. R. China, and the Master’s degree in 2013 from Manchester Metropolitan University, United Kingdom. She is currently a teacher at the Arts Department, Henan University of Economics and Law, P.R. China. Her research interests include digital arts and jewelry art.

**Weiwei Guo** received the Bachelor’s degree in Communication Engineering in 2011 from North China University of Water Resources and Electric Power, P. R. China, and the Master’s degree in Information and Communication Engineering in 2014 from University of Electronic Science and Technology of China, P. R. China, respectively. Her research interests include the signal processing and image restoration.

**Yongqin Zhang** received the B.S.E. degree in Electronics Science and Technology in 2005 from Zhengzhou University, P.R. China, the Ph.D. degree in Communication and Information Systems in 2010 from Wuhan University, P.R. China, respectively. He served as an assistant researcher at Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences and later as a postdoctoral researcher at the Institute of Computer Science and Technology, Peking University, P.R. China. He is currently a lector at the College of Computer and Information Engineering, Henan University of Economics and Law, P.R. China. His research interests include magnetic resonance imaging, sparse representation and image restoration.

**Jinsheng Xiao** received the Ph.D. degree in school of mathematics and statistics in 2001 from Wuhan University, Wuhan, P. R. China. He is currently an associate professor at School of Electronic Information, Wuhan University, Wuhan, P.R. China. His current research interests include computer vision, and image/video processing.

**Xiaoguang Hu** received the Ph.D. degree in Photogrammetry and remote sensing in 2012 from Wuhan University, P.R. China. He is currently a postdoctoral researcher at the College of Engineering, Peking University, P.R. China. His research interests are in biologically-inspired computational vision with technological applications to target detection, and robotics.