Color Reproduction From Noisy CFA Data of Single Sensor Digital Cameras

Lei Zhang, *Member, IEEE*, Xiaolin Wu, *Senior Member, IEEE*, and David Zhang, *Senior Member, IEEE*

Abstract—Single sensor digital color still/video cameras capture images using a color filter array (CFA) and require color interpolation (demosaicking) to reconstruct full color images. The color reproduction has to combat sensor noises which are channel dependent. If untreated in demosaicking, sensor noises can cause color artifacts that are hard to remove later by a separate denoising process, because the demosaicking process complicates the noise characteristics by blending noises of different color channels. This paper presents a joint demosaicking-denoising approach to overcome this difficulty. The color image is restored from noisy mosaic data in two steps. First, the difference signals of color channels are estimated by linear minimum mean square-error estimation. This process exploits both spectral and spatial correlations to simultaneously suppress sensor noise and interpolation error. With the estimated difference signals, the full resolution green channel is recovered. The second step involves in a wavelet-based denoising process to remove the CFA channel-dependent noises from the reconstructed green channel. The red and blue channels are subsequently recovered. Simulated and real CFA mosaic data are used to evaluate the performance of the proposed joint demosaicking-denoising scheme and compare it with many recently developed sophisticated demosaicking and denoising schemes.

Index Terms—Bayer pattern, color demosaicking, color filter array (CFA), denoising, wavelet.

I. INTRODUCTION

SINGLE sensor digital cameras use a color filter array (CFA), such as the Bayer pattern [1], to capture images. At each pixel, only one of the primary colors (e.g., red, green, and blue) is sampled. The sensor readings are corrupted by noises, which are channel dependent (i.e., the noise statistics vary with different channels). This poses a challenge to color image restoration: the interpolation of missing spectral samples amongst sensor noises. If color interpolation (color demosaicking) and denoising are performed in tandem, as in current practice, then the demosaicking process complicates the task of denoising by blending the sensor noises across color

channels, which may even correlate the noises with signals. As a result, color demosaicking, if done without treating the sensor noises, can introduce color artifacts that are difficult to remove by a separate denoising process. To prevent demosaicking from producing additional artifacts, we propose a new approach that performs color demosaicking and image denoising jointly.

Up to now most demosaicking techniques [3]–[24] assumed noiseless CFA data. The noiseless assumption is clearly invalid in practice. Not only for the low cost and/or resource constrained imaging devices, corruptive noise is inherent to image acquisition, but also for the high-end imaging devices such as high resolution digital cinema cameras, noise corruption can be severe since the pixel is very small and prone to noise. A convenient solution is to employ a separate denoising process after demosaicking. However, this separation approach is problematic. If left untreated at the demosaicking step, sensor noises can be root of new color artifacts that become more difficult to remove by a subsequent denoising process. In general sensor noises exhibit distinct statistics in different channels and, hence, are easier to model and remove without being blended by the demosaicking process. Interpolating spectral noises together with the signal in the demosaicking process produces compounded noises that are difficult to characterize. To aggravate the problem, the demosaicking process can introduce dependency between the compound noises and image signal.

Recently, Hirakawa and Parks [38] developed a joint demosaicking-denoising algorithm. They studied signal-dependent noises of CMOS digital cameras. Treating both demosaicking and denoising as a problem of estimating a sample from its neighbors, the authors tried to adaptively find a filter to accomplish the two tasks simultaneously. They used the total least square (TLS) denoising technique in [37] to determine the filter under some constraints of the CFA pattern. This scheme is shown to perform better than many "demosaicking first + denoising later" methods.

This paper presents a new, computationally efficient joint demosaicking-denoising scheme to restore the full color image from the noisy mosaic data. The new technique is developed for ubiquitous Bayer pattern, but it can be extended to other CFA patterns. To take advantage of spectral correlations, we estimate the red-green and blue-green difference images, called primary difference signals (PDS), rather than directly recover the missing color samples. A linear model is developed to represent and estimate the red-green and blue-green PDS signals. The observed PDS measurements are represented as the sum of true PDS, diffused sensor noise (DSN) and color interpolation error (IE). Based on the second order statistics of these components, an adaptive LMMSE method is used to estimate the

PDS from the noisy measurements. From the estimated PDS, we derive a full resolution green image, on which the original sensor noise is superimposed. The sensor noise is, however, nonstationary, for that it is channel-dependent. This renders traditional denoising algorithms invalid that assume stationary additive noise. A new wavelet-based denoising algorithm is, thus, developed to remove this specific nonstationary noise from the green channel. Anchored on the denoised green image, the red and blue channels are, consequently, recovered.

The rest of the paper is structured as follows. Section II introduces the concepts of primary difference signal, DSN and color IE. Section III discusses the estimation of true PDS from the noisy measurements. Section IV presents the wavelet-based denoising scheme to remove the particular type of nonstationary noise from the green channel. Section V turns to the reconstruction of red and blue channels. Experimental results on simulated and real CFA data are shown in Section VI. Section VII concludes the paper.

II. Modeling Noisy Primary Difference Signal

The spectral correlations between color channels play a central role in color demosaicking, and they are also valuable prior knowledge for removing sensor noises. The high spectral correlations are due to the fact that most natural scenes consist predominantly of pastoral colors. Fully saturated colors are rare. Based on this observation, our approach of joint demosaicking-denoising works with the so-called PDS, defined as $X_{G,R} = G - R$ or $X_{G,B} = G - B$, rather than individual color component signals. Since $X_{G,R}$ and $X_{G,B}$ are much smoother than the original R, G, and B signals, they can be better estimated from noisy measurements. In the definition of PDS, we use the green channel as the anchor, because in the Bayer CFA pattern (by far the most popular mosaic pattern used in digital cameras), the sampling frequency of green channel is twice that of red and blue channels. The sensitivity of the human visual system peaks at the green wavelength, which lies in between those of red and blue in the visible spectrum. The green-red and green-blue correlations are statistically greater than the red-blue correlation [11].

Due to the symmetry of R and B with respect to G, it suffices to study the signal $X_{G,R}$ in the following discussions. All results can be extended to $X_{G,B}$ using the symmetry. To simplify the notations, we will drop the subscripts and refer to $X_{G,R}$ as X in the following development. Given an estimate $\hat{X}$ of X, one can reconstruct the missing green samples as $G = R + \hat{X}$ at the red pixels or $G = B + \hat{X}$ at the blue pixels.

We assume that sensor noises are additive but channel-dependent. For each channel, we assume the noise is Gaussian white with zero mean. The sensor readings are

$$\tilde{R} = R + v_r, \quad \tilde{G} = G + v_g, \quad \tilde{B} = B + v_b \quad (1)$$

where v_r, v_g, and v_b are the sensor noise in the red, green, and blue channels, respectively, and R, G, and B are the true sample values to be recovered. The symbol "$\sim$" means that the sample is noise corrupted. The statistics of v_r, v_g, and v_b may be different, but the noises are mutually uncorrelated between
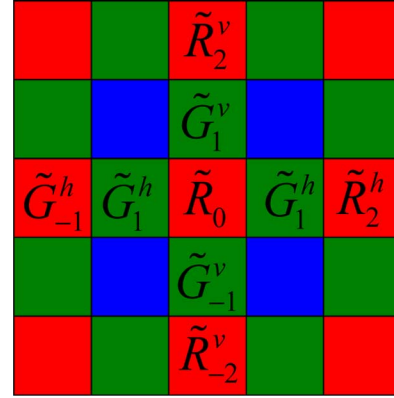

Fig. 1. Column and row of alternating green and red samples intersect at a red sampling position.

the channels. The standard deviations of v_g, v_r, and v_b are denoted by σ_g, σ_r, and σ_b.

In [38], Hirakawa and Parks assumed the sensor noise to be signal-dependent, being proportional to the signal magnitude plus a base level. This product noise model may fit some sensors better, but it makes denoising computationally far more expensive. For this reason, most existing image denoising techniques adopt an additive signal-independent noise model. The channel-dependent noise model adopted by this paper is a tradeoff between the signal-dependent product noise model and the signal-independent additive noise model. We allow the noise statistics to vary in different channels because a given type of sensors behaves differently in different wavelengths. On the other hand, we let sensor noise be independent of signal within each channel to simplify the denoising algorithm. This simplification does not materially degrade the visual quality of denoised images because the signal-to-noise ratio (SNR) is high anyway when the signal amplitude is high.

For concreteness and clarity, the following discussions refer to the mosaic configuration of Fig. 1, in which the PDS X is to be estimated. In Fig. 1, $\tilde{R}_0$ at the center of the window is a noisy red sensor reading. Its interlaced green and red neighbors in horizontal direction are labeled, respectively, as $\tilde{G}_i^h$, $i \in \{\ldots, -3, -1, 1, 3, \ldots\}$, and $\tilde{R}_i^h$, $i \in \{\ldots, -4, -2, 2, 4, \ldots\}$; similarly, the green and red vertical neighbors of $\tilde{R}_0$ are, respectively, $\tilde{G}_i^v$, $i \in \{\ldots, -3, -1, 1, 3, \ldots\}$, and $\tilde{R}_i^v$, $i \in \{\ldots, -4, -2, 2, 4 \ldots\}$. By convention, $\tilde{R}_0 = \tilde{R}_0^h = \tilde{R}_0^v$. We denote by R_i^d and G_i^d the noiseless counterparts of $\tilde{R}_i^d$ and $\tilde{G}_i^d$, $d \in \{h, v\}$, i.e.,

$$\tilde{R}_i^d = R_i^d + v_{r,i}^d, \quad \tilde{G}_i^d = G_i^d + v_{g,i}^d, \quad d \in \{h, v\} \quad (2)$$

where $v_{r,i}^d$ and $v_{g,i}^d$ are sensor noises at the corresponding red and green positions of the CFA pattern.

To obtain some initial measurements of X, we need to interpolate the missing green samples at the red pixels and the missing red samples at the green pixels. It is important to interpolate along the edge direction to prevent interpolation color artifacts. We adopt the well-known directional color interpolation filter proposed by Adams and Hamilton for its simplicity [4]–[6]. Referring to Fig. 1, at the position of each red sample

R_i, we compute both the horizontal and vertical interpolation of the missing green sample

$$\bar{G}_i^d = \frac{1}{2}\left(\tilde{G}_{i-1}^d + \tilde{G}_{i+1}^d\right) + \frac{1}{4}\left(2 \cdot \tilde{R}_i^d - \tilde{R}_{i-2}^d - \tilde{R}_{i+2}^d\right)$$
$$d \in \{h, v\}. \tag{3}$$

In Section III-C, we will see how the two directional estimation values are optimally fused to improve the directional estimates of missing green samples.

At the positions of green samples G_i, the missing red samples are either interpolated horizontally or vertically

$$\bar{R}_i^d = \frac{1}{2}\left(\tilde{R}_{i-1}^d + \tilde{R}_{i+1}^d\right) + \frac{1}{4}\left(2 \cdot \tilde{G}_i^d - \tilde{G}_{i-2}^d - \tilde{G}_{i+2}^d\right)$$
$$d \in \{h, v\} \tag{4}$$

depending on whether the green sample has horizontal red neighbors or vertical red neighbors.

Let $\hat{G}_i^d$ and $\hat{R}_i^d$ be the corresponding noise free interpolation results if noisy sensor readings $\tilde{R}_i^d$ and $\tilde{G}_i^d$ are replaced by the noiseless values R_i^d and G_i^d in (3) and (4).

Using the interpolated samples $\bar{G}_i^d$ and $\bar{R}_i^d$, we can obtain coarse measurements of PDS X at the positions of red and green samples in horizontal and vertical directions, respectively

$$\bar{X}_i^d = \begin{cases} \bar{G}_i^d - \tilde{R}_i^d, & \text{green is interpolated} \\ \tilde{G}_i^d - \bar{R}_i^d, & \text{red is interpolated} \end{cases}$$
$$d \in \{h, v\} \tag{5}$$

$\bar{X}_i^d$ can be rewritten as

$$\bar{X}_i^d = \hat{X}_i^d + v_i^d \tag{6}$$

where the noise free term is

$$\hat{X}_i^d = \begin{cases} \hat{G}_i^d - R_i^d, & \text{green is interpolated} \\ G_i^d - \hat{R}_i^d, & \text{red is interpolated} \end{cases} \tag{7}$$

and the noise term is shown in (8) at the bottom of the page

The term v_i^d is the effect of sensor noise on estimated PDS $\bar{X}_i^d$ via directional interpolation, which we call directional DSN. The term $\hat{X}_i^d$ in (6) is another estimate of PDS X_i by interpolating the noiseless CFA data. We denote the difference between $\hat{X}_i^d$ and true PDS X_i as

$$\varepsilon_i^d = \hat{X}_i^d - X_i. \tag{9}$$

Note that the term ε_i^d is purely the directional IE without any contribution of sensor noises, and it depends only on the true underlying image downsampled by CFA. Combining (6) and

(9), we have a linear model of PDS X

$$\bar{X}_i^d = X_i + \varepsilon_i^d + v_i^d \tag{10}$$

where $\bar{X}_i^d$ is the available noisy measurement of unknown PDS X_i that contains both the DSN term v_i^d and the IE term ε_i^d. The DSN term v_i^d is a linear combination of channel-dependent sensor noises v_g, v_r and v_b.

If left untreated, DSN v_i^d can be a significant contributor to color artifacts in the reconstructed image. However, as will be discussed in Section III, a reduction of DSN can be achieved by exploring its properties different from X_i and ε_i^d.

III. ESTIMATING THE PRIMARY DIFFERENCE SIGNAL

Now with the linear model (10) of PDS X, we proceed to compute two directional estimates of X_i separately using noisy measurements $\bar{X}_i^h$ and $\bar{X}_i^v$ of X_i [refer to (5)]. By optimally fusing these two directional estimates, we generate a full resolution estimate of the green channel, which is then processed in Section IV.

To simplify the notations, we denote by x the unknown PDS X_i, y the associated measurement $\bar{X}_i^h$ or $\bar{X}_i^v$. Similarly, we write ε_i^d as ε and v_i^d as v when the context makes the association clear between these error terms and the estimates in question. Consequently, (10) is rewritten as

$$y = x + \varepsilon + v. \tag{11}$$

A. Power Spectrums of PDS x, DSN v, and IE ε

Since DSN term v is a linear combination of channel-dependent sensor noises so that it is uncorrelated with PDS x and IE ε. It was shown in [11] that the IE ε is zero-mean and nearly uncorrelated with x, and the PDS x is a low-pass signal and ε a bandpass process in natural scenes. This section will further show that DSN v is a low-pass process. Together, these pieces of prior information allow us to compute an optimal estimate of x from noisy measurements y.

Referring to (8), we see that at the red pixel positions where green is interpolated, v is generated by applying filter $F = \begin{bmatrix} -1/4 & 1/2 & -1/2 & 1/2 & -1/4 \end{bmatrix}$ to noises v_r and v_g in the CFA image; at the green pixel positions where red is interpolated, v is obtained by applying $-F = \begin{bmatrix} 1/4 & -1/2 & 1/2 & -1/2 & 1/4 \end{bmatrix}$ to v_g and v_r. Because v_g and v_r are spatially interlaced with each other and have different statistics, the output noise image v will not be stationary. Nevertheless, v can be considered as a "half stationary" process, which means that the statistics of $v(n)$ is the same as that of $v(n + 2l)$. At the red sample positions in

$$v_i^d = \begin{cases} \dfrac{\left(v_{g,i-1}^d + v_{g,i+1}^d\right)}{2} - \dfrac{\left(2 \cdot v_{r,i}^d + v_{r,i-2}^d + v_{r,i+2}^d\right)}{4}, & \text{green is interpolated} \\ \dfrac{\left(2 \cdot v_{g,i}^d + v_{g,i-2}^d + v_{g,i+2}^d\right)}{4} - \dfrac{\left(v_{r,i-1}^d + v_{r,i+1}^d\right)}{2}, & \text{red is interpolated} \end{cases} \tag{8}$$
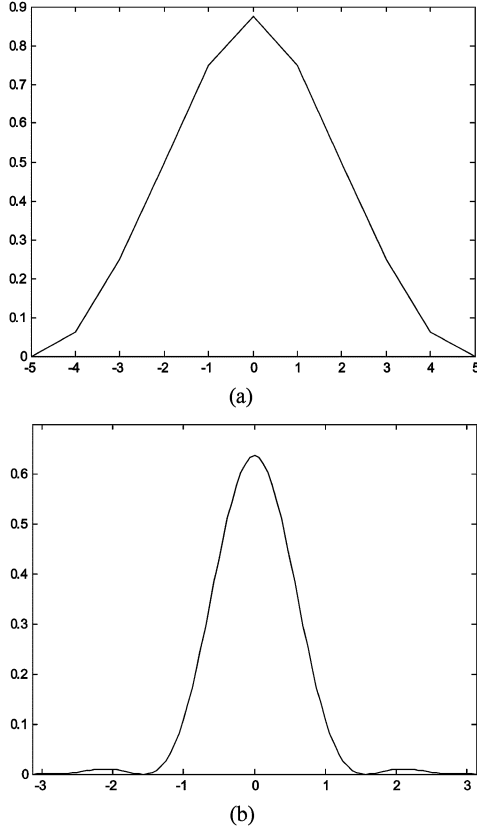
Fig. 2. (a) Autocorrelation function and (b) the power spectrum density of DSN $v(n)$.

Fig. 3. Power spectrum densities of (a) x and (b) ε for a typical nature image.

CFA, it can be derived that the autocorrelation function of v is

$$f_r(l) = E[v(n)v(n+l)] = \begin{cases} \frac{3}{8}\sigma_r^2 + \frac{4}{8}\sigma_g^2, & l = 0 \\ \frac{3}{8}\sigma_r^2 + \frac{3}{8}\sigma_g^2, & l = \pm 1 \\ \frac{1}{4}\sigma_r^2 + \frac{1}{4}\sigma_g^2, & l = \pm 2 \\ \frac{1}{8}\sigma_r^2 + \frac{1}{8}\sigma_g^2, & l = \pm 3 \\ \frac{1}{16}\sigma_r^2, & l = \pm 4 \\ 0, & |l| > 4. \end{cases} \quad (12)$$

Similarly, at the green pixel positions in CFA, we have

$$f_g(l) = E[v(n)v(n+l)] = \begin{cases} \frac{3}{8}\sigma_g^2 + \frac{4}{8}\sigma_r^2, & l = 0 \\ \frac{3}{8}\sigma_g^2 + \frac{3}{8}\sigma_r^2, & l = \pm 1 \\ \frac{1}{4}\sigma_g^2 + \frac{1}{4}\sigma_r^2, & l = \pm 2 \\ \frac{1}{8}\sigma_g^2 + \frac{1}{8}\sigma_r^2, & l = \pm 3 \\ \frac{1}{16}\sigma_g^2, & l = \pm 4 \\ 0, & |l| > 4. \end{cases} \quad (13)$$

It is observed that with the employed filter F, $f_r(l)$ and $f_g(l)$ are different only at $l = 0$ and $l = \pm 4$. If v_g and v_r are of the same intensity level, then DSN v is stationary and $f_r(l) = f_g(l)$. If the difference between σ_r and σ_g is not significant, we see that the difference between $f_r(0)$ and $f_g(0)$ (i.e., $(1/8)(\sigma_g^2 - \sigma_r^2)$) and the difference between $f_r(\pm 4)$ and $f_g(\pm 4)$ (i.e., $(1/16)(\sigma_r^2 - \sigma_g^2)$) are not significant. Therefore, we consider $f_r(l)$ and $f_g(l)$ to be approximately the same. Letting $\sigma_g = \sigma_r = 1$, the autocorrelation function of $f_r(l)$ and the power spectrum density function $f_p(\omega) = (1/2\pi)\sum_{l=-\infty}^{\infty} f_r(l)e^{-il\omega}$ are plotted in Fig. 2(a) and
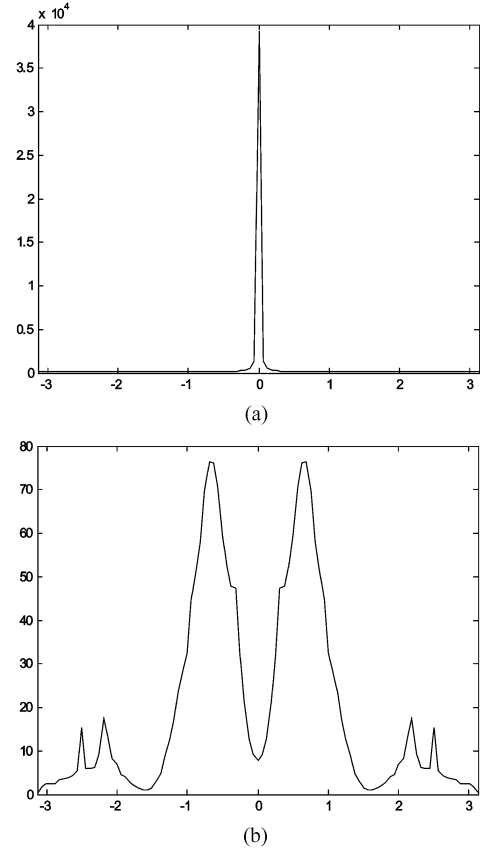
(b), respectively. The power spectral density of $f_r(l)$ clearly shows that v is a low-pass process. For comparison, we also plot in Fig. 3 the power spectrum densities of x and ε for a typical natural image (other images exhibit similar results). Comparing Fig. 2 with Fig. 3, we see that the power of x concentrates in a very low-frequency band, while the low-pass signal v has a much wider bandwidth than x. In contrast, the IE ε is bandpass and is relatively high-frequency dominated.

B. Directional Estimation of the PDS

Based on the linear model (11) and the different characteristics of x, ε and v discussed in the previous section, we estimate the unknown PDS x from its measurement y by using a spatially adaptive linear minimum mean square-error estimation (LMMSE) technique. The LMMSE of x from y is computed as

$$\hat{x} = E[x] + \frac{\text{Cov}(x,y)}{\text{Var}(y)}(y - E[y]). \quad (14)$$

Since DSN v and IE ε are zero-mean and x, v and ε are mutually uncorrelated, it follows from (11) and (14) that

$$\hat{x} = \mu_x + \frac{\sigma_x^2}{\sigma_y^2}(y - \mu_x) \quad (15)$$

where $\mu_x = E[x]$, $\sigma_x^2 = \text{Var}(x)$, $\sigma_y^2 = \text{Var}(y)$. We also have

$$\sigma_y^2 = \sigma_x^2 + \sigma_\varepsilon^2 + \sigma_v^2 \quad (16)$$

where $\sigma_\varepsilon^2 = \text{Var}(\varepsilon)$ and $\sigma_v^2 = \text{Var}(v)$.

From (15), we see that to compute the spatially adaptive LMMSE estimate of $x(n)$, we need three parameters $\mu_x(n)$, $\sigma_x(n)$ and $\sigma_y(n)$. However, these three parameters are not known as a prior. The only available data are the observations $y(n)$. Fortunately, with the linear model (11) and the different characteristics of x, ε and υ, it is possible for us to estimate $\mu_x(n)$, $\sigma_x(n)$ and $\sigma_y(n)$ adaptively from the neighborhood of $y(n)$.

We estimate the statistics of $x(n)$ and $y(n)$ from the neighborhood of $y(n)$ by assuming that process $x(n)$ is ergodic and stationary. Denote by

$$Y_n = [\,y(n - L) \quad \ldots \quad y(n) \quad \ldots \quad y(n + L)\,] \qquad (17)$$

the $(2L + 1) \times 1$ window centered at $y(n)$. Because υ and ε are zero mean, the mean of $y(n)$ is equal to the mean of $x(n)$. $\mu_x(n)$ can be computed by

$$\mu_x(n) = \mu_y(n) = \frac{1}{2L + 1} \sum_{k=1}^{2L+1} Y_n(k). \qquad (18)$$

The variance of $y(n)$ is computed as

$$\sigma_y^2(n) = \frac{1}{2L + 1} \sum_{k=1}^{2L+1} (Y_n(k) - \mu_y(n))^2. \qquad (19)$$

Now only $\sigma_x(n)$ is left to be estimated.

Since that the bandpass IE signal ε has very little energy in the low-frequency band while the low-pass PDS signal x has a very narrow bandwidth (referring to Figs. 2 and 3), passing y through a low-pass filter can remove almost all the energy of ε. Therefore, we smooth y. Denote by $\{h(k)\}$ the response sequence of a low-pass filter[1] and let

$$y_s(n) = (y * h)(n) = \sum_{k=-\infty}^{\infty} y(n - k) \cdot h(k) \qquad (20)$$

where "*" is the convolution operator. Since signal y_s has negligible contribution of IE signal ε, we have approximately

$$y_s \approx x + \upsilon_s \qquad (21)$$

where υ_s is the filtering output of bandpass noise signal υ

$$\upsilon_s(n) = (\upsilon * h)(n) = \sum_{k=-\infty}^{\infty} \upsilon(n - k) \cdot h(k). \qquad (22)$$

Since bandpass signal υ has much energy concentrated in the low-frequency band, its low-pass filtering output υ_s can not be neglected.

We will estimate $\sigma_x^2(n)$ from the smoothed signal y_s. Denote by

$$Y_n^s = [\,y_s(n - L) \quad \ldots \quad y_s(n) \quad \ldots \quad y_s(n + L)\,] \qquad (23)$$

[1]A Gaussian low-pass filter can be employed. The optimal scale of the Gaussian filter depends on the input signal. To reduce the complexity, we fix the filter as $h = [\,4 \quad 9 \quad 15 \quad 23 \quad 26 \quad 23 \quad 15 \quad 9 \quad 4\,]/128$ in the experiments.

the $(2L+1) \times 1$ window centered at $y_s(n)$. Because x and υ_s are uncorrelated, $\sigma_x^2(n)$, the variance of $x(n)$, can then be computed by

$$\sigma_x^2(n) = \frac{1}{2L + 1} \sum_{k=1}^{2L+1} (Y_n^s(k) - \mu_x(n))^2 - \sigma_s^2(n) \qquad (24)$$

where $\sigma_s^2(n)$ is the variance of $\upsilon_s(n)$. Next, we discuss the calculation of $\sigma_s^2(n)$.

As we discussed in Section III-A, process υ (and, hence, υ_s) is a "half stationary" process due to interlaced structure of red and green noises υ_r and υ_g. The variances of $\upsilon_s(n)$ are different at the positions of red and green pixels. At the red pixel positions, $\sigma_s^2(n)$ is computed by

$$\sigma_s^2(n) = R_r(0) \qquad (25)$$

where $R_r(l)$ is the autocorrelation function of $\upsilon_s(n)$ at red pixel positions. Since $\upsilon_s(n)$ is obtained by convolving $\upsilon(n)$ with $\{h(k)\}$ and the autocorrelation of $\upsilon(n)$ at the red pixel positions is $f_r(l)$, $R_r(l)$ is computed by [40]

$$R_r(l) = f_r(l) * h(l) * h(-l). \qquad (26)$$

Similarly, at the green pixel positions, $\sigma_s^2(n)$ is computed as

$$\sigma_s^2(n) = R_g(0), \quad R_g(l) = f_g(l) * h(l) * h(-l). \qquad (27)$$

Now we could be able to adaptively estimate parameters $\mu_x(n)$, $\sigma_x^2(n)$ and $\sigma_y^2(n)$ by using (18), (19), and (24), respectively. Then the spatially adaptive LMMSE of PDS $x(n)$, $\hat{x}(n)$, can be computed using (15). Let $e(n) = \hat{x}(n) - x(n)$ be the estimation error of $x(n)$, the variance of $e(n)$ is

$$\sigma_e^2(n) = E[e^2(n)] = \sigma_x^2(n) - \frac{\sigma_x^2(n)}{\sigma_y^2(n)}. \qquad (28)$$

C. Fusing the Directional Estimates

Using the scheme in Section III-B, we obtain a horizontal estimate and a vertical estimate of PDS $x(n)$ at the positions of red pixels. Let $\hat{x}_h(n)$ and $\hat{x}_v(n)$ be the two estimates, and $e_h(n)$ and $e_v(n)$ be the corresponding estimation errors, namely

$$\begin{cases} \hat{x}_h(n) = x(n) + e_h(n) \\ \hat{x}_v(n) = x(n) + e_v(n). \end{cases} \qquad (29)$$

The variances of estimation errors $e_h(n)$ and $e_v(n)$ are denoted by $\sigma_h^2(n)$ and $\sigma_v^2(n)$.

Estimate $\hat{x}_h(n)$ or $\hat{x}_v(n)$ exploits the CFA spectral-spatial correlation information only in horizontal or vertical direction. They should be fused to offer a more accurate estimate of $x(n)$. Let the fused estimate be the weighted average of $\hat{x}_h(n)$ and $\hat{x}_v(n)$

$$\hat{x}(n) = w_h(n) \cdot \hat{x}_h(n) + w_v(n) \cdot \hat{x}_v(n) \qquad (30)$$

where $w_h(n) + w_v(n) = 1$. Denote by $e(n) = \hat{x}(n) - x(n)$ the estimation error of $\hat{x}(n)$. The optimal weights $w_h(n)$ and $w_v(n)$
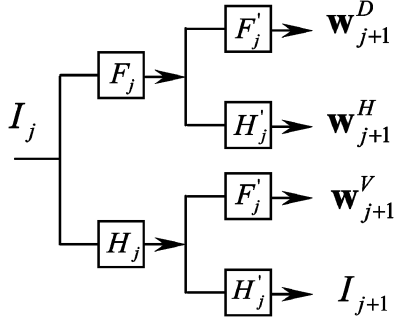
Fig. 4. One-stage decomposition of the 2-D OWE. H' and F' are the transposes of H or F. $\mathbf{w}_j^H$, $\mathbf{w}_j^V$ and $\mathbf{w}_j^D$ represent the wavelet coefficient matrices in horizontal, vertical and, diagonal directions.

should be the pair to minimize the mean square error (MSE) of $\hat{x}(n)$, i.e.,

$$(w_h, w_v) = \min_{w_h + w_v = 1} \arg \left\{ E[e^2(n)] \right\}.$$

Empirically, we found that the correlation between errors $e_h(n)$ and $e_v(n)$ is weak in natural images, especially in the area of rich high-frequency contents that are prone to visible color artifacts. In cases where $e_h(n)$ and $e_v(n)$ are highly correlated (implying that two estimates $\hat{x}_h$ and $\hat{x}_v$ are close to each other), the weighted estimate $\hat{x}$ is insensitive to w_h and w_v anyway. Assuming that $e_h(n)$ and $e_v(n)$ are approximately uncorrelated, i.e., $E[e_h(n) \cdot e_v(n)] \approx 0$, we have

$$\begin{aligned}
\sigma_e^2(n) &= E[e^2(n)] \\
&= w_h^2(n) \cdot \sigma_h^2(n) + w_v^2(n) \cdot \sigma_v^2(n) \\
&\quad + 2 \cdot w_h(n) \cdot w_v(n) \cdot E[e_h(n) \cdot e_v(n)] \\
&\approx w_h^2(n) \cdot \sigma_h^2(n) + w_v^2(n) \cdot \sigma_v^2(n) \\
&= w_h^2(n) \cdot \left(\sigma_h^2(n) + \sigma_v^2(n) \right) + \sigma_v^2(n) \\
&\quad - 2 \cdot w_h(n) \cdot \sigma_v^2(n).
\end{aligned}$$

To minimize $\sigma_e^2(n)$, we let the partial differential of $\sigma_e^2(n)$ with respect to $w_h(n)$ be zero

$$\frac{\partial \sigma_e^2(n)}{\partial w_h(n)} = 2 \cdot w_h(n) \cdot \left(\sigma_h^2(n) + \sigma_v^2(n) \right) - 2 \cdot \sigma_v^2(n) = 0.$$

Finally, we have the optimal weights minimizing the MSE of $\hat{x}(n)$ are

$$w_h(n) = \frac{\sigma_v^2(n)}{\sigma_h^2(n) + \sigma_v^2(n)}$$

$$w_v(n) = \frac{\sigma_h^2(n)}{\sigma_h^2(n) + \sigma_v^2(n)}. \tag{31}$$

IV. RECONSTRUCTING THE GREEN CHANNEL

In Bayer CFA pattern the green channel retains much of image details. The restoration quality of red and blue channels relies on a good estimate of the green channel. In Section III, we have obtained an optimal estimate $\hat{X}$ (i.e., $\hat{X}_{G,R}$ or $\hat{X}_{G,B}$) of PDS X. The effect of DSN υ is much reduced in $\hat{X}$ and we could approximately have $\hat{X} = X$. Adding back the PDS X to red and blue channels leads to the recovery of full resolution

(1)

(2)

(3)

Fig. 5. Test images. (1) and (2) are full color images and we down sample them to Bayer CFA images. (3) is a real CFA image captured by a digital camera.

green channel $G = R + X_{G,R}$ at red pixels or $G = B + X_{G,B}$ at blue pixels. However, this PDS-induced green channel is noisy, because sensor readings $\tilde{R} = R + \upsilon_r$, $\tilde{G} = G + \upsilon_g$ and $\tilde{B} = B + \upsilon_b$ are noise corrupted. In this section, we will develop an approach to reconstruct the green channel from the channel-dependent sensor noises υ_r, υ_g, and υ_b.

For the notation convenience, we use 2-D coordinates (m, n) to label a CFA position in this section. Considering the PDS estimate $\hat{X}$ to be approximately the true PDS X, the noisy missing green samples at the red/blue positions are recovered as

$$\begin{aligned}
\tilde{G}(m, n) &= \tilde{R}(m, n) + X_{G,R}(m, n) \text{ or} \\
\tilde{G}(m, n) &= \tilde{B}(m, n) + X_{G,B}(m, n)
\end{aligned} \tag{32}$$

and, consequently

$$\begin{aligned}
\tilde{G}(m, n) &= G(m, n) + \upsilon_r(m, n) \text{ or} \\
\tilde{G}(m, n) &= G(m, n) + \upsilon_b(m, n)
\end{aligned} \tag{33}$$

where $G = R + X_{G,R}$ or $G = B + X_{G,B}$ is the true green sample value at a red or blue position. In addition, $\tilde{G}(m, n) = G(m, n) + \upsilon_g(m, n)$ at a green pixel position, with $\tilde{G}$ being the sensor reading before any processing.

TABLE I
RUNNING TIME (SECONDS) FOR A 512×768 MOSAIC IMAGE BY DIFFERENT DEMOSAICKING AND DENOISING METHODS

Demosaicking methods	[6]			[7]			[12]			[20]			Joint demosaicking-denoising method [38]	Proposed joint demosaicking-denoising method
Running time	1.2			17			105			4.2				
Denoising methods	[29]	[34]	[36]	[29]	[34]	[36]	[29]	[34]	[36]	[29]	[34]	[36]		
Running time	50	60	308	50	60	308	50	60	308	50	60	308		
Total running time	51	61	309	67	77	325	155	165	413	54	64	312	2406	93

In removing the noises $v_c, c \in \{r, g, b\}$, from $\tilde{G}$ to restore the green image G, we have to respect the fact that the noise is non-stationary. The nonstationarity is due to the mosaic pattern of CFA and channel-dependent sensor noises. Most existing image denoising methods assume that the noise is additive and stationary. In our case, only when the spatially interleaved noises v_r, v_g, and v_b have the same statistics, the underlying process can be considered as stationary. With this in mind we develop a wavelet-based denoising algorithm for denoising $\tilde{G}$.

The wavelet transform (WT) has proved to be very effective in image denoising [27]–[34]. As a time-frequency analysis tool, WT preserves local sharp changes and prevents over-smoothing of pure frequency domain denoising approaches. The over-complete wavelet expansion (OWE) is particularly suited to denoising by utilizing redundant information [28]–[31]. One stage of 2-D OWE decomposition is illustrated in Fig. 4. No downsampling occurs in OWE but the analytic filters vary with scales. Filter H_j is interpolated by putting $(2^{j-1} - 1)$ zeros between each of the coefficients of original filter H_0, so does for F_j. The filter bandwidth decrease is accomplished by the zero-padding of filter coefficients instead of the downsampling of wavelet coefficients. For more detailed information on wavelet and its application to denoising, refer to [25]–[36].

In the remainder of this section, bold vector symbols denote wavelet coefficient matrices. Since OWE is a linear transform, the OWE of noisy green image $\tilde{G} = G + v_c$, where $c \in \{r, g, b\}$ varies with the spatial position of $\tilde{G}$, is

$$\mathbf{w}_j^d = \mathbf{x}_j^d + \boldsymbol{v}_j^d, \quad d \in \{H, V, D\} \tag{34}$$

where matrix $\mathbf{w}_j^d$ represents the OWE of noisy image $\tilde{G}$ at scale j and in horizontal (H), vertical (V) or diagonal (D) direction, and matrices $\mathbf{x}_j^d$ and $\boldsymbol{v}_j^d$ represent the OWEs of noiseless image G and noise v_c respectively.

The LMMSE of $\mathbf{x}_j^d(m, n)$ can be computed from noisy wavelet coefficients in $\mathbf{w}_j^d$ to achieve denoising. Let the variances of $\boldsymbol{v}_j^d(m, n)$ and $\mathbf{x}_j^d(m, n)$ be σ_j^2 and $\sigma_{\mathbf{x}_j}^2$. Since in wavelet domain both $\boldsymbol{v}_j^d(m, n)$ and $\mathbf{x}_j^d(m, n)$ are zero mean and they are mutually uncorrelated, the LMMSE of $\mathbf{x}_j^d(m, n)$ is

$$\hat{\mathbf{x}}_j^d(m, n) = \frac{\sigma_{\mathbf{x}_j}^2}{\sigma_{\mathbf{x}_j}^2 + \sigma_j^2} \mathbf{w}_j^d(m, n). \tag{35}$$

Let us discuss the estimation of noise variance σ_j^2 first. Referring to Fig. 4, $\mathbf{w}_j^D$ is obtained through

$$\mathbf{w}_{j+1}^D = I_0 * L_j^D \tag{36}$$

where I_0 is the input image and 2-D filter L_j^D is

$$L_j^D = H_0 * H_0' * \cdots * H_{j-1} * H_{j-1}' * F_j * F_j'. \tag{37}$$

Similarly, we have

$$\mathbf{w}_{j+1}^H = I_0 * L_j^H, \quad \mathbf{w}_{j+1}^V = I_0 * L_j^V \tag{38}$$

where

$$L_j^H = H_0 * H_0' * \cdots * H_{j-1} * H_{j-1}' * F_j * H_j' \tag{39}$$
$$L_j^V = H_0 * H_0' * \cdots * H_{j-1} * H_{j-1}' * H_j * F_j'. \tag{40}$$

If noises $v_c, c \in \{r, g, b\}$ have the same variance $\sigma_g^2 = \sigma_r^2 = \sigma_b^2 = \sigma^2$, the noise wavelet coefficient $\boldsymbol{v}_j^d$ is stationary and its standard deviation (in horizontal, vertical, or diagonal direction) is

$$\sigma_j = \left\| L_{j-1}^d \right\| \sigma, \quad d \in \{H, V, D\}$$

where $\|\bullet\|$ is the norm operator $\|L\| = \sqrt{\sum_l \sum_k L^2(l, k)}$. However, σ_g, σ_r and σ_b may be different, so $\boldsymbol{v}_j^d(m, n)$ is non-stationary and its variances vary in the positions of red, green and blue samples in CFA.

Denote by $L_{j-1}^d(0, 0)$ the zero-time response of 2-D filter L_{j-1}^d. The coefficient $\boldsymbol{v}_j^d(m, n)$ is obtained by

$$\boldsymbol{v}_j^d(m, n) = \sum_{l=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \boldsymbol{v}_c(m - l, n - k) L_{j-1}^d(l, k). \tag{41}$$

To account for the nonstationarity of noises, we need to examine $\boldsymbol{v}_j^d(m, n)$ in four separate cases corresponding to different sample relations in the Bayer CFA pattern.

Case 1) The position $(m, n) = (\text{odd}, \text{odd})$ corresponds to a *green* sample, i.e., the noise at $\tilde{G}(m, n)$ is v_g. We denote by $\sigma_j^{(1)}$ the standard deviation of $\boldsymbol{v}_j^d(m, n)$ at those positions. At positions $(m + 2z_m, n + 2z_n + 1)$, where z_m and z_n are integers, the noise is v_r; at positions $(m + 2z_m + 1, n + 2z_n)$ the noise is v_b and at other positions the noise is v_g. Define a matrix

$$P_{j-1}^{(1)}(l, k) = \begin{cases} L_{j-1}^d(l, k) \cdot \sigma_r, & l = 2z_m + 1, k = 2z_n \\ L_{j-1}^d(l, k) \cdot \sigma_b, & l = 2z_m, k = 2z_n + 1 \\ L_{j-1}^d(l, k) \cdot \sigma_g, & \text{other positions.} \end{cases} \tag{42}$$

Then at position $(m, n) = (\text{odd}, \text{odd})$, the standard deviation of green noise wavelet coefficients $\boldsymbol{v}_j^d(m, n)$ in this case is $\sigma_j^{(1)} = \left\| P_{j-1}^{(1)} \right\|$.

Case 2) The position $(m, n) = (\text{even}, \text{even})$ also corresponds to a *green* sample but different from **Case 1**, now positions $(m + 2z_m, n + 2z_n + 1)$ are blue samples of noise v_b, positions $(m + 2z_m + 1, n + 2z_n)$ are red samples of noise v_r and other positions are green samples of noise v_g. We denote by $\sigma_j^{(2)}$ the standard deviation of $\boldsymbol{v}_j^d(m, n)$ at
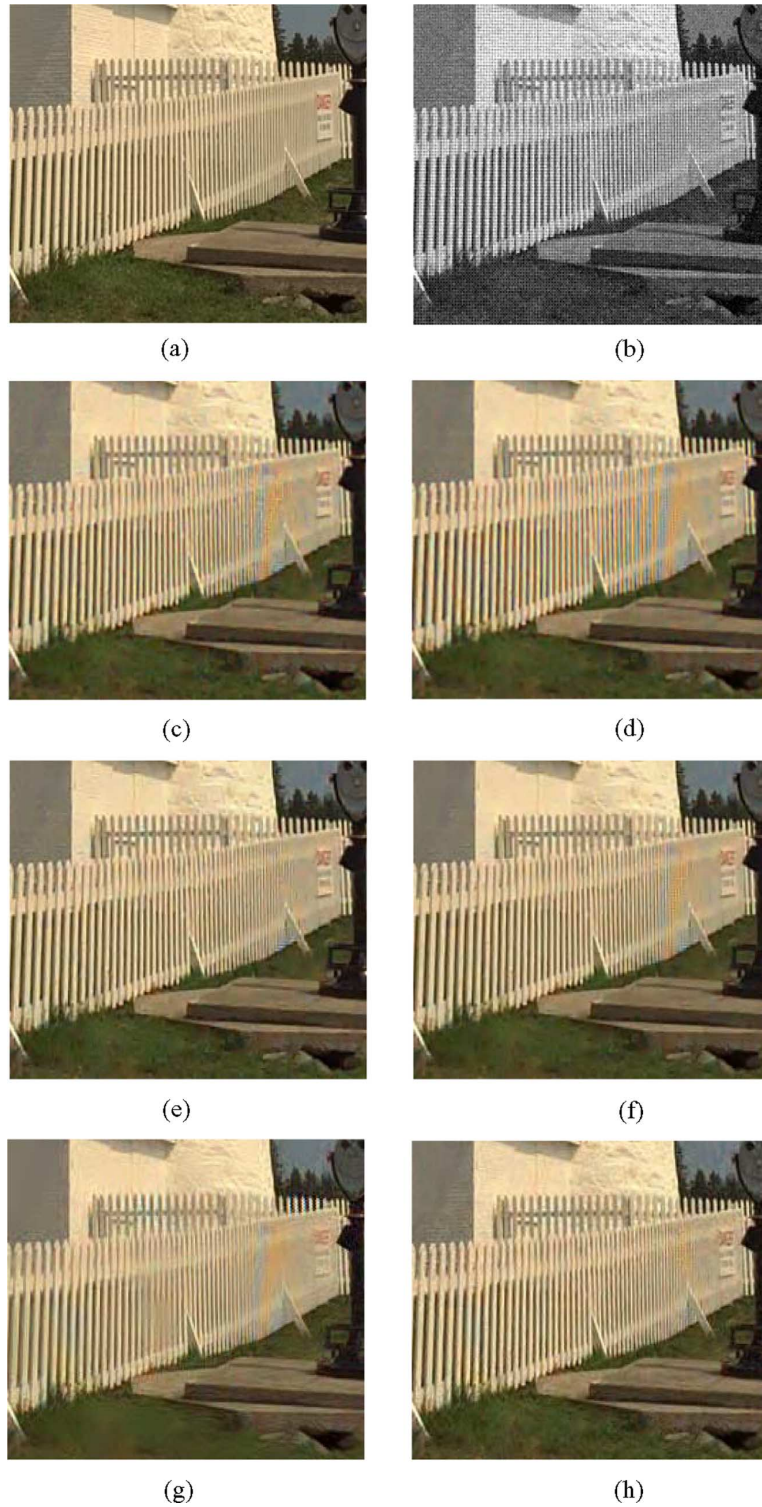
Fig. 6. Full color reconstruction for image "fence." (a) Original image; (b) noisy CFA data ($\sigma_r = 13$, $\sigma_g = 12$, $\sigma_b = 10$); (c)–(f) reconstructed images by demosaicking methods [6], [7], [12], and [20] followed by denoising method [36]; (g) reconstructed image by method [38]; (h) reconstructed image by the proposed method.

those positions. Define a matrix

$$P_{j-1}^{(2)}(l,k) = \begin{cases} L_{j-1}^d(l,k) \cdot \sigma_b, & l = 2z_m + 1, k = 2z_n \\ L_{j-1}^d(l,k) \cdot \sigma_r, & l = 2z_m, k = 2z_n + 1 \\ L_{j-1}^d(l,k) \cdot \sigma_g, & \text{other positions.} \end{cases}$$

(43)

At positions $(m,n) = (\text{even}, \text{even})$, the standard deviation of $\boldsymbol{v}_j^d(m,n)$ is $\sigma_j^{(2)} = \left\| P_{j-1}^{(2)} \right\|$.

Case 3) $(m,n) = (\text{odd}, \text{even})$ corresponds to *red* sample positions. We denote by $\sigma_j^{(3)}$ the standard deviation of $\boldsymbol{v}_j^d(m,n)$ at those positions. Similarly, we can compute
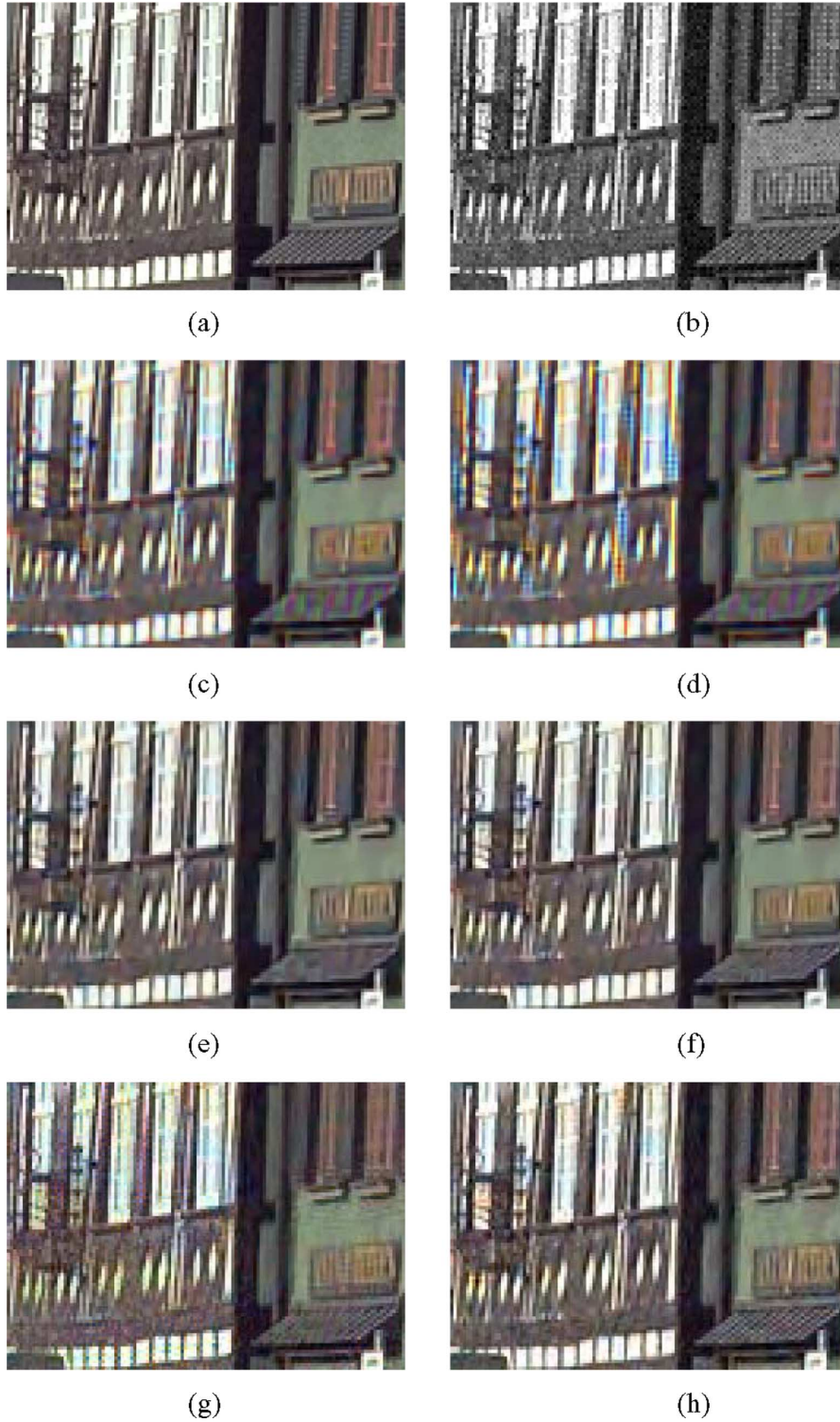
Fig. 7. Full color reconstruction for image "house." (a) Original image; (b) noisy CFA data ($\sigma_r = 13$, $\sigma_g = 12$, $\sigma_b = 10$); (c)–(f) reconstructed images by demosaicking methods [6], [7], [12], and [20] followed by denoising method [36]; (g) reconstructed image by method [38]; (h) reconstructed image by the proposed method.

$\sigma_j^{(3)}$ as $\sigma_j^{(3)} = \left\| P_{j-1}^{(3)} \right\|$, where

$$P_{j-1}^{(3)}(l,k) = \begin{cases} L_{j-1}^d(l,k) \cdot \sigma_r, & l = 2z_m, k = 2z_n \\ L_{j-1}^d(l,k) \cdot \sigma_b, & l = 2z_m + 1, k = 2z_n + 1 \\ L_{j-1}^d(l,k) \cdot \sigma_g, & \text{other positions.} \end{cases}$$

(44)

Case 4) Finally, $(m,n) = (\text{even}, \text{odd})$ corresponds to *blue* sample positions. We denote by $\sigma_j^{(4)}$ the standard devia-

tion of $\boldsymbol{v}_j^d(m,n)$ at those positions and we have $\sigma_j^{(4)} = \left\| P_{j-1}^{(4)} \right\|$, where

$$P_{j-1}^{(4)}(l,k) = \begin{cases} L_{j-1}^d(l,k) \cdot \sigma_b, & l = 2z_m, k = 2z_n \\ L_{j-1}^d(l,k) \cdot \sigma_r, & l = 2z_m + 1, k = 2z_n + 1 \\ L_{j-1}^d(l,k) \cdot \sigma_g, & \text{other positions.} \end{cases}$$

(45)

TABLE II
PSNR (DECIBEL) RESULTS OF THE RECONSTRUCTED FENCE IMAGES

Demosaicking Methods	Denoising Methods	PSNR Result (dB) $\sigma_r=\sigma_g=\sigma_b=12$			$\sigma_r=13,\sigma_g=12,\sigma_b=10$		
		R	G	B	R	G	B
[6]	[29]	28.7	30.3	28.8	28.3	30.3	29.4
[6]	[34]	28.0	30.3	28.8	28.4	30.3	28.6
[6]	[36]	28.1	30.6	28.3	27.9	30.6	28.6
[7]	[29]	27.3	30.7	27.5	26.9	30.8	27.9
[7]	[34]	27.2	30.7	27.5	26.9	30.8	27.8
[7]	[36]	27.0	30.6	27.2	26.8	30.6	27.5
[12]	[29]	29.9	30.7	30.1	29.7	30.7	30.4
[12]	[34]	29.9	30.7	30.1	29.7	30.7	30.3
[12]	[36]	30.2	31.1	30.6	30.0	31.2	30.9
[20]	[29]	30.4	31.5	30.8	30.2	31.6	31.2
[20]	[34]	30.4	31.5	30.7	30.2	31.5	31.1
[20]	[36]	30.4	31.5	31.0	30.2	31.6	31.3
[38]		27.1	28.6	28.3	27.1	28.8	28.5
Proposed Method		30.7	31.5	31.2	30.5	31.6	31.5

TABLE III
PSNR (DECIBEL) RESULTS OF THE RECONSTRUCTED HOUSE IMAGES

Demosaicking Methods	Denoising Methods	PSNR Result (dB) $\sigma_r=\sigma_g=\sigma_b=12$			$\sigma_r=13,\sigma_g=12,\sigma_b=10$		
		R	G	B	R	G	B
[6]	[29]	25.8	27.9	25.7	25.6	27.9	26.0
[6]	[34]	25.9	27.9	25.9	25.8	27.9	26.2
[6]	[36]	25.6	28.0	25.5	25.5	28.0	25.8
[7]	[29]	24.2	28.4	24.0	24.0	28.4	24.2
[7]	[34]	24.2	28.4	24.0	24.1	28.4	24.2
[7]	[36]	24.0	27.9	23.9	23.9	28.0	24.1
[12]	[29]	27.5	28.4	27.6	27.4	28.5	27.8
[12]	[34]	27.4	28.3	27.5	27.4	28.4	27.7
[12]	[36]	27.6	28.5	27.8	27.6	28.7	28.1
[20]	[29]	28.5	29.6	28.7	28.4	29.6	29.0
[20]	[34]	28.4	29.5	28.6	28.3	29.6	28.8
[20]	[36]	28.6	29.6	28.9	28.5	29.7	29.2
[38]			24.7	26.3	24.9	24.7	26.3	25.0
Proposed Method		28.3	29.3	28.5	28.2	29.3	28.8

Regarding to the estimation of $\sigma_{\mathbf{x}_j}^2$, the variance of $\mathbf{x}_j^d(m,n)$, which is required in evaluating (35), we use an $M \times N$ window centered at (m,n). The wavelet coefficients in the window are grouped into four sets corresponding to the four cases. The ith set, denoted by Θ_i, contains those coefficients falling into **Case 1** $\sigma_{\mathbf{x}_j}^2$ can be locally estimated by

$$\sigma_{\mathbf{x}_j}^2 = \frac{1}{M \cdot N} \sum_{i=1}^{4} \sum_{(l,k)\in\Theta_i} \left(\left(\mathbf{w}_j^d(l,k)\right)^2 - \left(\sigma_j^{(i)}\right)^2 \right). \quad (46)$$

Finally, if $\mathbf{x}_j^d(m,n)$ is falling into **Case 1**, $i \in \{1,2,3,4\}$, it is estimated by using (35)

$$\hat{\mathbf{x}}_j^d(m,n) = \frac{\sigma_{\mathbf{x}_j}^2 \mathbf{w}_j^d(m,n)}{\left(\sigma_{\mathbf{x}_j}^2 + \left(\sigma_j^{(i)}\right)^2\right)}$$

where $\sigma_{\mathbf{x}_j}^2$ is computed as in (46) and $\sigma_j^{(i)} = \left\| P_{j-1}^{(i)} \right\|$.

V. RECONSTRUCTING THE RED/BLUE CHANNEL

Once an estimate of green channel is obtained, the restoration of red/blue channel becomes straightforward. The denoising process of green channel in Section IV yields an approximation of the noise v_c as a side product. Let $\mathcal{G}$ be the denoised version of $\tilde{G}$, the additive noise v_c in the mosaic data can be estimated as

$$v_c \approx \tilde{G} - \mathcal{G}. \quad (47)$$

Denote by $\tilde{I}$ the noisy CFA mosaic image. Since the additive noise v_c has been estimated in (47), we can compute the denoised CFA image I as

$$I = \tilde{I} - v_c. \quad (48)$$

Now the problem is transformed to interpolate the red R and blue B channels with the denoised CFA image I and the demosaicked-denoised green image $\mathcal{G}$.

Any existing demosaicking method can be employed. However, to be compatible with the PDS-based technique used in this paper, we recommend the method in [11] to reproduce the missing red and blue samples. This method recovers the chrominance channels through a fast bilinear average operation on the estimated green-red and green-blue PDS signals. It is simple and efficient once we have a robust estimate of the green channel.

The other way to reconstruct the red/blue channel is that we first reproduce the missing red and blue samples using the PDS bilinear interpolation method in [11] and then apply the denoising method in Section IV to the three color channels separately. The color reproduction result by using this scheme is a little better than the previous one but the price is that we run three times the denoising process. In the experiments in Section VI, we choose the second scheme. The first scheme is a better choice when the available computation resource is limited.

VI. EXPERIMENTAL RESULTS

This section presents experimental results and evaluates the performance of the proposed joint demosaicking-denoising
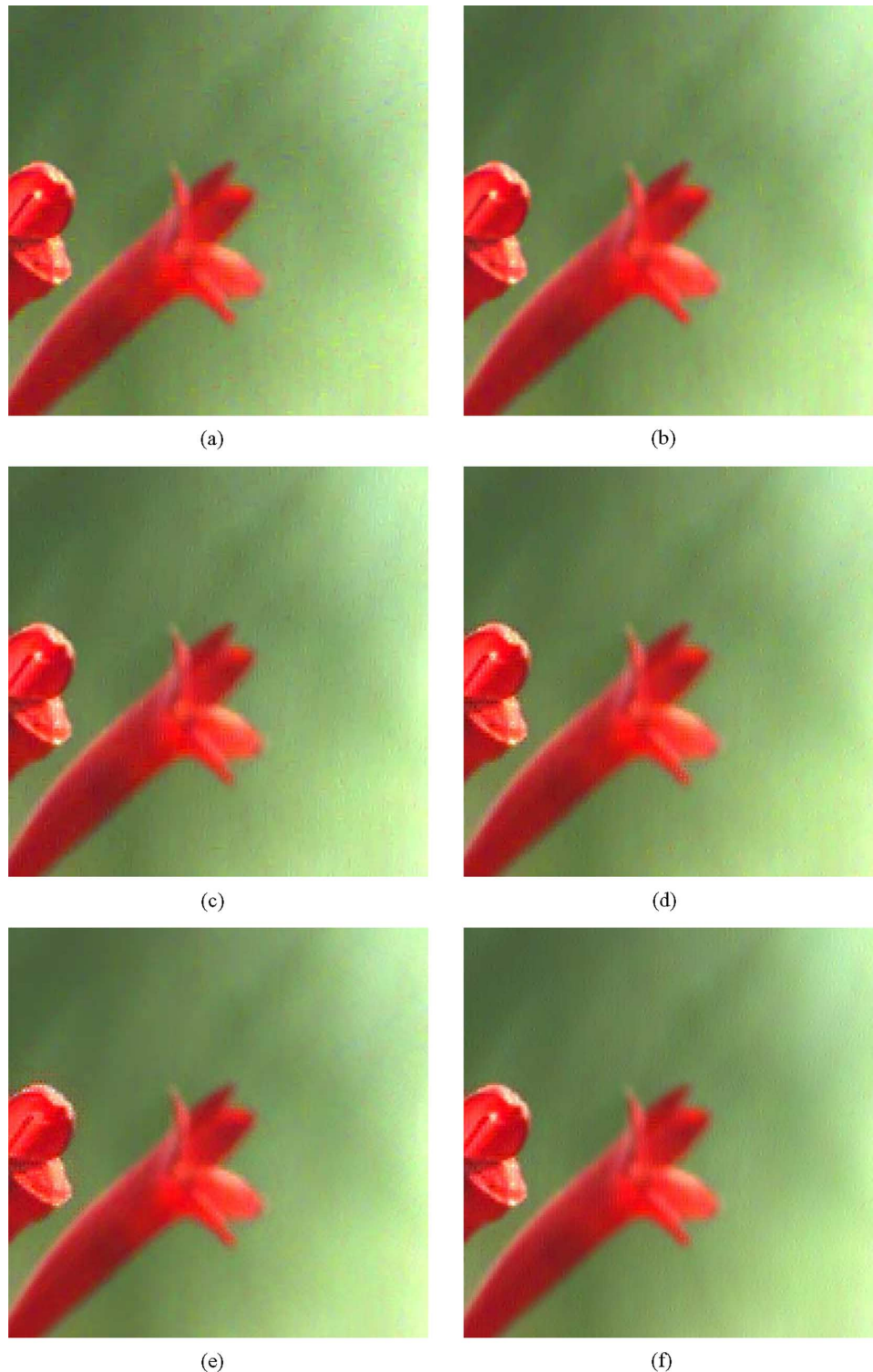
Fig. 8. Full color reconstruction for image "flower." (a)–(d) Reconstructed images by demosaicking methods [6], [7], [12], and [20] followed by denoising method [36]; (e) reconstructed image by method [38]; (f) reconstructed image by the proposed method.

method. The results of three test images (see Fig. 5) are included. The first two images (512 × 768) are from the Kodak set of color images, which are widely used in the literature for evaluating demosaicking algorithms. These full color images were downsampled into CFA mosaic images according to the Bayer pattern. To simulate channel-dependent sensor noises, Gaussian white noises of different variances were added separately to red, blue and green channels of the mosaic images.

For the simulated noisy CFA data for which the original is known, we can compute the PSNR measurement of the restored color images. The third image is a real CFA image, which is captured by the digital cinema camera developed by the DALSA Corporation. This 500 × 750 image was cropped from a very large frame.

Our comparison study is fairly comprehensive. Five demosaicking-denoising schemes are tested for comparison on the

noisy CFA data: the recently developed joint demosaicking-denoising scheme [38]; two well-known gradient based demosaicking methods [6], [7] and two recently developed sophisticated demosaicking methods [12], [20] followed by three competitive wavelet-based denoising techniques [29], [34], [36].[2] The denoising is performed separately on the red, green and blue channels of the demosaicked images. These denoising algorithms assume signal-independent stationary noises of known variances. For a fair evaluation of them on channel-dependent noises, we let these algorithms use an equivalent channel-independent Gaussian white noise of standard deviation $\sigma_a = (1/2)\sqrt{\sigma_r^2 + 2\sigma_g^2 + \sigma_b^2}$ (i.e., the average energy of the noises). Also, we evaluate both stationary and nonstationary noise situations, namely when v_r, v_g and v_b are equal and when they are not equal. In the second case, we make the noise level of each channel proportional to the signal energy of that channel. For example, for the first and second test images we set $\sigma_r = 13$, $\sigma_g = 12$, $\sigma_b = 10$.

To roughly assess the complexity of different demosaicking-denoising solutions, in Table I we list the running times of their Matlab 7.0 implementations to process a 512×768 CFA image. The hardware platform consists of an Intel Pentium 4, 2.26-GHZ CPU and 1.5-GB RAM. It should be stressed that the comparison only serves as a reference because the running time also heavily depends on the optimization of the program codes. It is observed that the complexity of the joint demosaicking-denoising method [38] is significantly higher than the other methods. The complexity of denoising method [36] is much higher than the other two denoising methods, and the demosaicking method [12] requires the greatest running time among all the demosaicking methods.

The reported experimental results can be found at http://www.comp.polyu.edu.hk/~cslzhang/dmdn.htm. For each test image, 14 restored versions (12 different combinations of demosaicking-denoising schemes, the joint demosaicking-denoising scheme [38] and the proposed method) are created. To save space, we only show parts of the reconstructed images here. Among the three denoising schemes, [36] gives the best visual quality. Therefore, we couple this denoising method with the four demosaicking techniques [6], [7], [12], and [20], and compare the outputs with those of the joint demosaicking-denoising scheme [38] and the proposed technique.

Fig. 6(a) is a part of the test image *fence* and Fig. 6(b) is the simulated noisy CFA image. The noise levels are $\sigma_r = 13$, $\sigma_g = 12$, $\sigma_b = 10$. Fig. 6(c)–(f) are the denoised images (by method [36]) of the outputs of the demosaicking methods [6], [7], [12], and [20]. Fig. 6(g) is the result of [38] and Fig. 6(h) is the result of the proposed joint demosaicking-denoising technique. Similarly, Fig. 7(a)–(h) shows the results of image *house* produced by the same set of methods. The demosaicking methods in [6] and [7] produce many color artifacts that resist subsequent denoising process, even the sophisticated ones like [36]. The color artifacts in the high-frequency areas are mainly caused by the IE, while those in smooth areas are culprit of the DSN. The demosaicking

methods [12] and [20] perform better, but they still suffer from color artifacts and blurred details after going through a separate denoising process. This can be seen in the grass portion of the *fence* image and the window panel portion of the *house* image. The joint demosaicking-denoising method [38] also generates many color artifacts. The proposed technique is very effective in removing color artifacts caused by both IE and DSN while preserving fine image structures, apparently achieving the best visual quality. In Tables II and III, the PSNR results of the restored images are listed. The proposed method is also competitive in PSNR measure.

The third test image *flower* consists of real CFA mosaic data. To perform denoising, we first estimate the noise energy of each channel from the CFA data. Specifically, we divide the $N \times M$ CFA image into $4\ N/2 \times M/2$ subimages: two green subimages, one red subimage and one blue subimage. We apply one-stage OWT to each subimage, and estimate the corresponding noise energy as $\sigma = \text{Median}(\mathbf{w}^V)/0.6475$ [27], where $\mathbf{w}^V$ is the diagonal sub-band. For the green channel, the noise level is the average of those of the two green subimages. Fig. 8 presents the results in the same experiment setting as for the other two test images. Similar conclusions can be drawn from Fig. 8 as from Figs. 6 and 7. The proposed method appears to be the best in visual quality.

VII. CONCLUSION

This paper presented a new full color reconstruction method of noisy CFA data through an LMMSE filtering of the green-red and green-blue PDS and a wavelet-based denoising process. It is observed that the PDS is a low-pass process, and it is uncorrelated with the IE, which is a bandpass process, and the DSN, which is a low pass process but with relatively wide bandwidth. Based on these properties we estimated the PDS in both horizontal and vertical directions and then optimally fused them. With the estimated PDS we got a full resolution green image, on which the additive noise is imposed. Considering that the additive noise is channel-dependent, we proposed a specific wavelet-based denoising algorithm to remove the noise from the green channel. The resulted green channel was used to guide the reconstruction of the red and blue samples. The experiments verified that the proposed joint demosaicking-denoising color reconstruction scheme significantly suppresses the noise caused color artifacts while preserving well the image details. It outperforms the state-of-the-art color demosaicking and denoising methods.

---

[2]The source codes of demosaicking methods [12], [20] and denoising methods [34], [36] are obtained from the authors. We thank the authors for sharing their codes.

REFERENCES

[1] B. E. Bayer, "Color imaging array," U.S. patent 3 971 065, 1975.
[2] P. Longère, X. Zhang, P. B. Delahunt, and D. H. Brainard, "Perceptual assessment of demosaicing algorithm performance," *Proc. IEEE*, vol. 90, no. 1, pp. 123–132, Jan. 2002.

[3] D. R. Cok, "Signal processing method and apparatus for producing interpolated chrominance values in a sampled color image signal," U.S. Patent 4 642 678, 1987.

[4] J. E. Adams, "Intersections between color plane interpolation and other image processing functions in electronic photography," *Proc. SPIE*, vol. 2416, pp. 144–151, 1995.

[5] J. E. Adams and J. F. Hamilton Jr, "Adaptive color plane interpolation in single color electronic camera," U.S. Patent 5.506 619, 1996.

[6] J. E. Adams, "Design of practical color filter array interpolation algorithms for digital cameras," *Proc. SPIE*, vol. 3028, pp. 117–125, 1997.

[7] E. Chang, S. Cheung, and D. Y. Pan, "Color filter array recovery using a threshold-based variable number of gradients," *Proc. SPIE*, vol. 3650, pp. 36–43, 1999.

[8] R. Kimmel, "Demosaicing: Image reconstruction from CCD samples," *IEEE Trans. Image Process.*, vol. 8, no. 9, pp. 1221–1228, Sep. 1999.

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

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

[11] L. Zhang and X. Wu, "Color demosaicking via directional linear minimum mean square-error estimation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2167–2178, Dec. 2005.

[12] K. Hirakawa and T. W. Parks, "Adaptive homogeneity-directed demosaicing algorithm," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 360–369, Mar. 2005.

[13] R. Kakarala and Z. Baharav, "Adaptive demosaicing with the principal vectormethod," *IEEE Trans. Consum. Electron.*, vol. 48, no. 11, pp. 932–937, Nov. 2002.

[14] W. Lu and Y. Tan, "Color filter array demosaicking: New method and performance measures," *IEEE Trans. Image Process.*, vol. 12, no. 10, pp. 1194–1210, Oct. 2003.

[15] S. C. Pei and I. K. Tam, "Effective color interpolation in CCD color filter arrays using signal correlation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 6, pp. 503–513, Jun. 2003.

[16] R. Lukac, B. Smolka, K. Martin, K. N. Plataniotis, and A. N. Venetsanopoulos, "Vector filtering for color imaging," *IEEE Signal Process. Mag.*, vol. 22, no. 1, pp. 74–86, Jan. 2005.

[17] R. Lukac and K. N. Plataniotis, "Color filter arrays: Design and performance analysis," *IEEE Trans. Consum. Electron.*, vol. 51, no. 4, pp. 1260–1267, Nov. 2005.

[18] R. Lukac and K. N. Plataniotis, "Data-adaptive filters for demosaicking: A framework," *IEEE Trans. Consum. Electron.*, vol. 51, no. 2, pp. 560–570, May 2005.

[19] R. Lukac, K. Martin, and K. N. Plataniotis, "Demosaicked image postprocessing using local color ratios," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 6, pp. 914–920, Jun. 2004.

[20] X. Li, "Demosaicing by successive approximation," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 370–379, Mar. 2005.

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

[22] S. Farsiu, M. Elad, and P. Milanfar, "Multiframe demosaicing and super-resolution of color images," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 141–159, Jan. 2006.

[23] D. D. Muresan and T. W. Parks, "Demosaicing using optimal recovery," *IEEE Trans. Image Process.*, vol. 14, no. 2, pp. 267–278, Feb. 2005.

[24] B. Tao, I. Tastl, T. Cooper, M. Blasgen, and E. Edwards, "Demosaicing using human visual properties and wavelet interpolation filtering," in *Proc. IS&T/SID 7th Color Imaging Conf.: Color Science, Systems, Applications*, Scottsdale, AZ, Nov. 1999, vol. 7, pp. 252–256.

[25] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. Pure Appl. Math.*, vol. 41, pp. 909–996, 1988.

[26] A. Cohen, I. Daubechies, and J. C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 45, pp. 485–560, 1992.

[27] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.

[28] R. R. Coifman and D. L. Donoho, "Translation-invariant de-noising," in *Wavelet and Statistics*, A. Antoniadis and G. Oppenheim, Eds., 1st ed.   Berlin, Germany: Springer, 1995.

[29] S. G. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1522–1531, Sep. 2000.

[30] Q. Pan, L. Zhang, G. Dai, and H. Zhang, "Two denoising methods by wavelet transform," *IEEE Trans. Signal Process.*, vol. 47, no. 12, pp. 3401–3406, Dec. 1999.

[31] L. Zhang, P. Bao, and X. Wu, "Multiscle LMMSE-based image denoising with optimal wavelet selection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 4, pp. 469–481, Apr. 2005.

[32] M. K. Mihçak, I. Kozintsev, K. Ramchandran, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Process. Lett.*, vol. 6, no. 12, pp. 300–303, Dec. 1999.

[33] A. Pizurica, W. Philips, I. Lamachieu, and M. Acheroy, "A joint inter- and intrascale statistical model for Bayesian wavelet based image denoising," *IEEE Trans. Image Process.*, vol. 11, no. 5, pp. 545–557, May 2002.

[34] A. Pizurica and W. Philips, "Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 654–665, Mar. 2006.

[35] J. Liu and P. Moulin, "Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients," *IEEE Trans. Image Process.*, vol. 10, no. 11, pp. 1647–1658, Nov. 2001.

[36] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.

[37] K. Hirakawa and T. W. Parks, "Image denoising for signal-dependent noise," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing*, Mar. 2005, vol. 2, pp. 29–32.

[38] K. Hirakawa and T. W. Parks, "Joint demosaicking and denoising," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2146–2157, Aug. 2006.

[39] S. Mallat, *A Wavelet Tour of Signal Processing*.   New York: Academic, 1998.

[40] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed.   New York: McGraw-Hill, 1991.

Lei Zhang (M'04) received the B.S. degree in 1995 from the Shenyang Institute of Aeronautical Engineering, Shenyang, China, and the M.S. and Ph.D. degrees in electrical and engineering from Northwestern Polytechnical University, Xi'an, China, in 1998 and 2001, respectively.

From 2001 to 2002, he was a Research Associate with the Department of Computing, The Hong Kong Polytechnic University. From January 2003 to January 2006, he was a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. Since January 2006, he has been an Assistant Professor with the Department of Computing, The Hong Kong Polytechnic University. His research interests include image and video processing, biometrics, pattern recognition, multisensor data fusion, optimal estimation theory, etc.

Xiaolin Wu (SM'96) received the B.Sc. degree from Wuhan University, China, in 1982, and the Ph.D. degree from University of Calgary, Calgary, AB, Canada, in 1988, both in computer science.

He is currently a Professor with the Department of Electrical and Computer Engineering, McMaster University, ON, Canada, and holds the NSERC-Dalsa Research Chair in digital cinema. His research interests include network-aware multimedia communication, joint source-channel coding, signal quantization and compression, and image processing. He has published over 170 research papers and holds two patents in these fields.

Dr. Wu is an Associate Editor of the IEEE TRANSACTIONS ON MULTIMEDIA.

David Zhang (SM'95) graduated in computer science from Peking University, China, and received the M.Sc. degree in computer science and the Ph.D. degree from the Harbin Institute of Technology (HIT), China, in 1982 and 1985, respectively. In 1994, he received the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada.

From 1986 to 1988, he was a Postdoctoral Fellow at Tsinghua University, China, and then an Associate Professor at the Academia Sinica, Beijing, China. Currently, he is a Chair Professor at the Hong Kong Polytechnic University where he is the Founding Director of the Biometrics Technology Centre (UGC/CRC) supported by the Hong Kong SAR Government in 1998. He also serves as a Visiting Chair Professor in Tsinghua University, and Adjunct Professor in Shanghai Jiao Tong University, Harbin Institute of Technology, and the University of Waterloo. He is the author of more than ten books and 190 journal papers.

Dr. Zhang is the Founder and Editor-in-Chief of the *International Journal of Image and Graphics* (IJIG); Editor of the Springer International Series on Biometrics (KISB); Organizer of the the International Conference on Biometrics Authentication (ICBA); Associate Editor of more than ten international journals, including the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART A: SYSTEMS AND HUMANS, the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, and *Pattern Recognition*; and Technical Committee Chair of the IEEE CIS. He is a Croucher Senior Research Fellow, Distinguished Speaker of the IEEE Computer Society, and a Fellow of the International Association of Pattern Recognition (IAPR).