JOINT DENOISING AND DEMOSAICKING OF NOISY CFA IMAGES BASED ON INTER-COLOR CORRELATION

Xingyu ZHANG[*], *Ming-Ting SUN*[†], *Lu FANG*[‡] *and Oscar C. AU*[*]

[*]The Hong Kong University of Science and Technology, {eexyzhang, eeau}@ust.hk
†University of Washingtong, sun@ee.washingtong.edu
‡University of Science and Technology of China, fanglu@ustc.edu.cn

ABSTRACT

Most digital cameras use a single sensor coupled with a Color Filter Array (CFA) to capture images, and apply demosaicking to interpolate the full color images. In reality, the CFA image is noisy, which causes problems in the demosaicking process. This paper proposes a Joint Denoising and Demosaicking based on inter-Color correlation (JDDC) scheme. We propose a new framework that linearly combines an extracted luminance image and a low-passed RGB images to get a full color image. Given the noise in the extracted luminance image and the low-passed RGB images are non-stationary and partially correlated, we modify the classical Non-Local Means (NLM) filter to denoise the extracted luminance image and the low-passed RGB images before the combination. Experimental results verify the effectiveness of the proposed scheme both objectively and subjectively.

Index Terms— CFA, Bayer, Denoising, Demosaicking.

1. INTRODUCTION

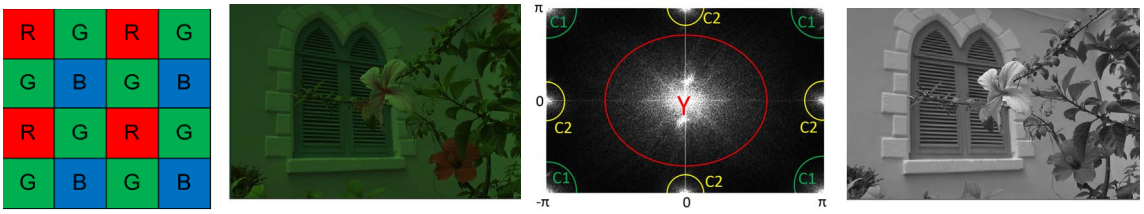
Digital color images usually have three color values (i.e. red, green and blue) at each pixel location. To reduce the cost of an imaging system, a Color Filter Array (CFA) is overlaid on a single sensor to capture only one color at each pixel location. There are several patterns of CFA. The most commonly used pattern is the Bayer pattern CFA [1] as shown in Fig. 1(a). To recover the full color image, the other two missing color values must be interpolated at each pixel location. Such an interpolation process is called demosaicking.

Many demosaicking algorithms [1] have been proposed in the literature. However, most of them assume noiseless CFA data, which is rarely true in reality. For almost all kinds of digital imaging devices, contaminative noise is inherent due to the process of converting photons to electrons and the electronic amplifier noise *etc*. In order to obtain a full color image, the most straightforward solution is to perform demosaicking first and then denoising, since there are many effective denoising algorithms existing for noisy gray-scale or full color images. However, the presence of noise poses a great challenge to the demosaicking task, and the compound noise produced in demosaicking is difficult to remove by the subsequent denoising process. The second approach is to perform denoising first and then demosaicking. Many adavanced monochromatic and color image denoising methods have been proposed in the literature [2][3]. However, they are not directly applicable to CFA images due to the underlying mosaic structure of CFA. One simple solution is to partition the CFA image into one red, one blue, and two green sub-images, and then denoise them separately. However this solution does not well exploit the spatial and inter-channel correlation properties in natural images. [4] designed a denoising algorithm for the CFA images based on the principle component analysis (PCA) of local image statitics. By adaptively computing the co-variance matrix of each variable block, the PCA could transform the noisy signal into another space to compact the signal energy and suppress the noise effectively.

Another alternative solution is to perform denosing and demosaicking jointly. In light of the observation that image interpolation and denoising are both estimation problem, [5] develops a novel technique to combine demosaicking and denoising procedures systematically into a single operation through the total least squares (TLS) filter using available neighboring noisy pixels. Later, [6] proposed a joint scheme outperformed [5]. [6] estimates the color difference signal using the Minimum Mean Square Error (MMSE) technique which exploits both spectral and spatial correlations to simultaneously suppress the sensor noise and interpolation error. Then a wavelet-based denoising process is applied to remove the channel-dependent noise.

In this paper, we propose a novel Joint Denoising and Demosaicking based on inter-Color correlation (JDDC) scheme. With the assumption that the high frequency components of R, G, and B images resemble each other in a natural image [8], JDDC linearly combines the extracted luminance and low-passed RGB images to get a full color image. Given the noise in the extracted luminance and low-passed RGB images are non-stationary and partially correlated, we modify the Non-Local Means (NLM) filter to denoise the extracted luminance and low-passed RGB images before the combination. Simulation results show that the proposed scheme achieves better

978-1-4799-2893-4/14/$31.00 ©2014 IEEE



(a) (b) (c) (d)

Fig. 1: (a), Bayer CFA; (b), a CFA image; (c), frequency representation of (a); (d), an extracted luminance image of (b).

performance compared to previous reported schemes.

The rest of this paper is organized as follows. In Section 2, the frequency analysis of CFA images [7] is introduced. Section 3 introduces the detail of our proposed scheme. Finally, experimental results are shown in Section 4, with a conclusion drawn in Section 5.

2. FREQUENCY ANALYSIS OF CFA IMAGES

Denote the clean pixel value of an RGB image as $C_q(i,j)$, where $q \in S : \{r, g, b\}$ and location $(0, 0)$ indicates the location of the top left pixel inside a image. Its CFA image can be viewed as subsampling the R, G, B color values at each pixel location with the sampling function $m_q(i,j)$,

$$\begin{aligned} m_r(i,j) &= [1+(-1)^i][1+(-1)^j]/4 \\ m_g(i,j) &= [1-(-1)^{(i+j)}]/2 \\ m_b(i,j) &= [1-(-1)^i][1-(-1)^j]/4. \end{aligned} \tag{1}$$

Therefore, the representation of a CFA image is

$$\begin{aligned} I(i,j) &= \sum_{q\in S} C_q(i,j)m_q(i,j) \\ &= \frac{1}{4}[C_r(i,j)+2C_g(i,j)+C_b(i,j)] \\ &+ \frac{1}{4}[C_r(i,j)-2C_g(i,j)+C_b(i,j)](-1)^{(i+j)} \\ &+ \frac{1}{4}[C_r(i,j)-C_b(i,j)][(-1)^i+(-1)^j]. \end{aligned} \tag{2}$$

We define $Y(i,j) = \frac{1}{4}[C_r(i,j)+2C_g(i,j)+C_b(i,j)]$, $C_1(i,j) = \frac{1}{4}[C_r(i,j)-2C_g(i,j)+C_b(i,j)]$, and $C_2(i,j) = \frac{1}{4}[C_r(i,j)-C_b(i,j)]$. Let $\check{}$ represents the discrete time Fourier transform. Then the Fourier transform of I is

$$\begin{aligned} \check{I}(u,v) &= \check{Y}(u,v)+\check{C}_1(u-\pi,v-\pi) \\ &+\check{C}_2(u-\pi,v)+\check{C}_2(u,v-\pi). \end{aligned} \tag{3}$$

A typical frequency representation of a CFA image Fig. 1(b) is shown in Fig. 1(c). It can be seen that the luminance component Y and the chrominance components C_1, C_2 cover different regions in the frequency domain. Therefore, it is possible to filter out a luminance image from the CFA image. Although an adaptive filter may extract the luminance image with higher quality, later in the experiment, we use a fixed 5x5 filter as proposed in [7] for simplicity .

3. PROPOSED SCHEME

As in Fig. 1(d), the extracted luminance image contains details of the full color image, i.e. the high frequency information of the image. For a full color image of natural scene, the high frequency components of different color channels are observed to resemble each other [8]. Motivated by this observation, our proposed JDDC scheme adds the high frequency information from the luminance image onto the low-passed R, G, and B images to get the full color image. Denoising the extracted luminance image and the low-passed RGB images are needed before the combination.

3.1. Estimating color values via inter-color correlation

Each color channel, i.e. R, G and B, can be decomposed into a low frequency component and a high frequency component. With the assumption that the high frequency components of R, G, and B resemble each other, we have

$$C_q - C_q^{lp} = C_p - C_p^{lp}, \quad q,p \in S \tag{4}$$

where C_q^{lp} represents the low frequency component of the color channel q. Taking Eq. 4 into the definition of $Y(i,j)$, i.e. $Y(i,j)=\frac{1}{4}[C_r(i,j)+2C_g(i,j)+C_b(i,j)]$, $C_q(i,j)$ can thus be estimated as

$$\hat{C}_q(i,j) = Y(i,j) - \frac{1}{4}\boldsymbol{\alpha}_q^T \mathbf{C}^{lp}(i,j), \tag{5}$$

where the vector $\mathbf{C}^{lp} = [C_r^{lp} \;\; C_g^{lp} \;\; C_b^{lp}]^T$, $\boldsymbol{\alpha}_r = [-3 \;\; 2 \;\; 1]^T$, $\boldsymbol{\alpha}_g = [1 \;\; -2 \;\; 1]^T$ and $\boldsymbol{\alpha}_b = [1 \;\; 2 \;\; -3]^T$. The low-passed RGB image C_p^{lp} can be obtained via a low pass filter a,

$$\begin{aligned} C_q^{lp}(i,j) &= ((C_q \cdot m_q) * a)(i,j) \\ &= \sum_{(k,l)\in B_{i,j}^a} C_q(k,l)m_q(k,l)a(k-i,l-j), \end{aligned} \tag{6}$$

The operation $*$ means the convolution operation. Set $B_{i,j}^a$ contains the pixel indices that kernel a will cover when calculating the convolution for location (i,j).

However, for the noisy CFA image, the luminance image $Y(i,j)$ and the noise free low-passed color images C_q^{lp} are

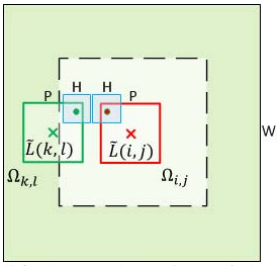


Fig. 2: Modified search region of the NLM filter.

not directly available. Therefore we perform denoising on an extracted luminance image and the noisy low-passed RGB images before applying the principle developed in Eq. 5 to estimate the full color image.

3.2. Denoising with modified NLM filter

In this work, we use the NLM filter for the denoising purpose. The NLM filter is firstly proposed in [2] for monochromatic image denoising and has demonstrated considerable noise reduction capability. In our problem, the noise on the extracted luminance image and the low-passed images are found to be non-stationary and partially correlated. In our JDDC scheme, we modify the NLM filter by adjusting its search region and the definition of the weight function, so that the weight can be easily calculated in the modified NLM filtering.

Let the extracted luminance image of image I be L

$$L(i,j) = (I * h)(i,j) = \sum_{(t,s)\in B_{i,j}} I(t,s)h(t-i, s-j). \quad (7)$$

The filter h is used to extract the luminance image from the CFA image. Set $B_{i,j}$ contains the pixel indices that filter h covers in the convolution for location (i,j). According to [7], filter h is designed with a constraint that the ratio of the used R, G and B pixels are exactly 1:2:1 in the generated $L(i,j)$ at different locations, which is the same as $Y(i,j)$. Thus, if we replace $Y(i,j)$ by $L(i,j)$, Eq. 5 approximately holds true.

We assume the noisy color values of a full color image is $\tilde{C}_q(i,j) = C_q(i,j) + n_q(i,j)$, where n_q is the stationary signal-independent noise with zero mean and variance σ_q^2. σ_q^2 can be different for different channels. Then the noisy CFA image is $\tilde{I}(i,j) = \sum_{q\in S} \tilde{C}_q(i,j)m_q(i,j) = I(i,j) + N(i,j)$, where $N(i,j) = \sum_{q\in S} n_q(i,j)m_q(i,j)$. The CFA image noise $N(i,j)$ is obviously uncorrelated. $N(i,j)$ is non-stationary since its variance $\sigma_N^2(i,j) = \sum_{q\in S} m_q(i,j)\sigma_q^2(i,j)$ is not a constant, but location dependent. When $m_q(i,j) = 1$, $\sigma_N^2(i,j) = \sigma_q^2(i,j)$.

Let the noisy extracted luminance image of $\tilde{I}$ be $\tilde{L}(i,j)$

$$\tilde{L}(i,j) = (\tilde{I} * h)(i,j) = Y(i,j) + n_L(i,j). \quad (8)$$

Noise $n_L(i,j) = \sum_{(t,s)\in B_{i,j}} N(t,s)h(t-i,s-j)$ is for the noisy extracted luminance image with zero mean and variance σ_L^2. $n_L(i,j)$ is also partially correlated since the overlapping

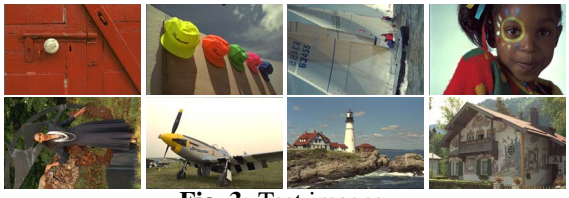


Fig. 3: Test images.

property of set $B(i,j)$ at different locations.

$$\sigma_L^2(i,j) = E[n_L(i,j)^2] = E[(\sum_{(t,s)\in B_{i,j}} N(t,s)h(t-i,s-j))^2]$$
$$= \sum_{(t,s)\in B_{i,j}} h^2(t-i,s-j)E[N^2(t,s)] \quad (9)$$

When h is selected to be the fixed 5×5 filter designed by [7],

$$\sigma_L^2(i,j) = \begin{cases} \frac{1}{4096}(44\sigma_r^2 + 2528\sigma_g^2 + 44\sigma_b^2), & \text{if } m_g(i,j) = 1 \\ \frac{1}{4096}(2464\sigma_r^2 + 88\sigma_g^2 + 64\sigma_b^2), & \text{if } m_r(i,j) = 1 \\ \frac{1}{4096}(64\sigma_r^2 + 88\sigma_g^2 + 2464\sigma_b^2), & \text{if } m_b(i,j) = 1 \end{cases}$$

By using the NLM filter, the estimated clean $\hat{L}(i,j)$ of noisy $\tilde{L}(i,j)$ is achieved by the weighted average of multiple noisy $\tilde{L}(k,l)$ in a predefined search region:

$$\hat{L}(i,j) = \frac{\sum_{(k,l)\in B_{i,j}^s} w(i,j,k,l)\tilde{L}(k,l)}{\sum_{(k,l)\in B_{i,j}^s} w(i,j,k,l)}, \quad (10)$$

with $w(i,j,k,l)$ being the weight for $\tilde{L}(k,l)$ when computing $\hat{L}(i,j)$, and $B_{i,j}^s$ being the set of location indices of reference pixels. The weight $w(i,j,k,l)$ depends on the resemblance between a local noise-free patch centered at (i,j) and a local noise-free patch centered at (k,l). Higher resemblance leads to a larger weight. However, in practice, only noisy pixels are available. Then a natural approach is to use the sum of squared difference between noisy patches as an approximation to measure the resemblance between corresponding noise-free patches. When the noise is correlated, [9] proposed to replace the Euclidean distance between the current patch and the reference patch by the Mahalanobis distance, which takes the noise co-variance matrix into account for whitening the noise. However, in the whitening process, the inverse of the co-variance matrix needs to be calculated. For example, when the patch size is 5×5, the size of the co-variance matrix is 25×25. In our problem, $n_L(i,j)$ is partially correlated and also non-stationary, which means high computation complexity for inverse calculation of the co-variance matrix.

In JDDC, we propose to modify the search region in the NLM filtering process, so that the co-variance matrix between the current patch and the reference patch is diagonal. The inverse co-variance matrix is then also diagonal and can be easily calculated. The property of such a modified search region is that the noise n_L in the reference patch and the current patch are uncorrelated. Let the size of the patch, the size of filter h, and the outer boundary length of the search window be $P\times P$, $H\times H$, and $W\times W$ respectively. For the current pixel

location (i,j), the modify search region $B^s_{i,j}$ includes the pixels sandwiched between a $(2P+2H+1)\times(2P+2H+1)$ region and a $W \times W$ region, both of which are centered at location (i,j). For example, as shown in Fig. 2, two cross points are the current pixel $\tilde{L}(i,j)$ and a reference pixel $\tilde{L}(k,l)$ with $l=j$ and $k<i$. The local patches of $\tilde{L}(i,j)$ and $\tilde{L}(k,l)$ are denoted as $\Omega_{i,j}$ and $\Omega_{k,l}$. For the noise $n_L(i-\frac{H}{2}, j-\frac{H}{2})$ in $\Omega_{i,j}$, it is contributed by the noise $N(t,s)$ with $(t,s)\in B_{i-\frac{H}{2},j-\frac{H}{2}}$, i.e. noise N in an $H\times H$ region that kernel h covers when calculating $\tilde{L}(i-\frac{H}{2}, j-\frac{H}{2})$. Similarly, noise $n_L(k+\frac{H}{2}, l-\frac{H}{2})$ in $\Omega_{k,l}$ is contributed by the noise $N(t,s)$ with $(t,s) \in B_{k+\frac{H}{2},l-\frac{H}{2}}$. Therefore, as long as the horizontal distance between $\tilde{L}(i,j)$ and $\tilde{L}(k,l)$ are larger than $P+H$, i.e $(i-k) > (P+H)$, the noise in $\Omega_{i,j}$ and the noise in $\Omega_{k,l}$ will be uncorrelated. The similarity measurement can therefore be defined as

$$d(i,j,k,l) = \sum_{b_1,b_2 \in B_1} \frac{[\tilde{L}(k+b_1, l+b_2) - \tilde{L}(i+b_1, j+b_2)]^2}{\sigma_L(i+b_1, j+b_2)\sigma_L(k+b_1, l+b_2)}. \quad (11)$$

Set $B_1 = \{0, 1, ..., \frac{P}{2}\}$. Note that, we will consider the current pixel as a reference pixel as well, with $d(i,j,i,j) = 0$. Correspondingly, the weight $w(i,j,k,l)$ is defined as

$$w(i,j,k,l) = exp\{-\frac{d(i,j,k,l)}{Ng^2}\} \quad (12)$$

where N is the total number of pixels in a patch, g acts as the filtering parameter.

For the denoised estimates of the low-passed RGB images, the same process of denoising $\tilde{L}$ can be used. Taking $\hat{L}$ and the denoised low-passed average color images into Eq. 5 instead of Y and $\mathbf{C}^{lp}$, we can get a denoised full color image.

4. EXPERIMENTAL RESULTS

The test images we used to evaluate the proposed scheme are listed left to right and from the first row to the second row in the order of Image 1 to Image 8 as shown in Fig. 3. The noisy CFA images of the test images are obtained by mosaicking those full color images in the Bayer pattern, and adding with the simulated white Gaussian noise. The peak signal-to-noise ratio (PSNR) of the estimated full color image with respect to the original noise-free image is used as the objective measure. We compared our proposed JDDC scheme with several schemes, including demosaicking [10]+denoising [11] scheme, an advanced joint denoising and demosaicking scheme [6], denoising using the NLM filter on the four sub images of the CFA image+demosaicking [10], denoising [4]+demosaicking [10], and an extended variant of our approach in which the recovered RGB image is mosaicked again and then demosaicked by [10] for a fair comparison with [4]+[10] scheme. Table 1 shows that our proposed JDDC and its variant provides better PSNR performance than other schemes on average. Some visual comparisons of different schemes are demonstrated in Fig. 4. Results from our proposed JDDC scheme and its variant also provide sharper edges and cleaner smooth regions.

Table 1: PSNRs of the reconstructed RGB images.

Image		σ_r,σ_g,σ_b=15,20,25					
		[10]+[11]	[6]	Proposed JDDC	SubNL +[10]	[4]+[10]	Proposed JDDC +[10]
1	R	30.98	33.66	33.38	32.27	33.90	33.37
	G	30.79	34.52	35.85	32.29	35.21	35.88
	B	29.62	32.61	35.02	31.21	34.47	35.08
2	R	31.00	35.12	35.43	32.60	35.52	35.45
	G	30.79	35.11	35.80	32.37	35.51	35.86
	B	29.88	32.96	34.86	31.43	34.85	34.90
3	R	30.73	34.66	34.88	32.19	34.84	34.92
	G	30.56	34.44	34.75	31.99	34.63	34.81
	B	30.22	33.14	34.51	31.52	34.51	34.58
4	R	31.81	34.11	34.39	32.66	34.51	34.44
	G	31.55	34.66	35.83	32.85	35.42	35.89
	B	30.79	33.03	34.99	31.78	34.79	35.04
5	R	30.86	32.57	32.81	31.63	32.77	32.77
	G	30.53	32.65	33.07	31.39	32.81	33.09
	B	29.74	31.75	32.99	30.87	32.68	32.96
6	R	34.00	35.69	36.11	34.71	36.10	36.15
	G	33.48	35.81	36.31	34.21	36.14	36.37
	B	32.21	33.19	34.69	32.51	34.86	34.85
7	R	30.62	33.09	33.31	31.74	33.28	33.31
	G	30.42	33.19	33.61	31.58	33.41	33.63
	B	29.90	31.96	32.88	30.90	32.83	32.87
8	R	30.91	32.94	33.26	31.71	33.11	33.27
	G	30.52	32.72	33.17	31.38	32.88	33.20
	B	29.84	31.60	32.22	30.67	32.42	32.23
Average		30.91	33.55	34.34	32.02	34.23	34.37

(a) Original (b) [6] (c) Proposed JDDC

(d) SubNL+[10] (e) [4]+[10] (f) Proposed JDDC+[10]

Fig. 4: Visual comparison of different schemes on Image 2.

5. CONCLUSION

A simple JDDC scheme is proposed. The full color values of each pixel are estimated by a linear combination of an estimated luminance image and the estimated low-passed RGB images. Both subjective and objective experimental results demonstrate that the proposed scheme is effective in the joint denoising and demosaicking task.

6. REFERENCES

[1] D. Menon and G. Calvagno, "Color image demosaicking: an overview", *Signal Processing: Image Communication*, vol. 26, no. 8, Oct. 2011, pp. 518-533.

[2] A. Buades, B. Coll and J. M. Morel, "A review of image denoising algorithms, with a new one", *Multiscale Modeling and Simulation*, vol. 4, no. 2, pp. 490-530, 2005.

[3] J. Dai *et al.*, "Multichannel Nonlocal Means Fusion for Color Image Denoising", *IEEE Trans on Circuits and Systems for Video Technolog*, vol. 23, no. 11, pp. 1873-1886, Nov. 2013.

[4] L. Zhang *et al.*,"PCA-based spatially adaptive denoising of CFA images for single-sensor digital cameras", *IEEE Trans on Image Processing*, vol. 18, no. 4, pp. 797-812, Apr. 2009.

[5] K. Hirakawa and T. W. Parks, "Joint demosaicing and denoising", *IEEE Trans. on Image Processing*, vol. 15, no. 8, pp. 2146-2157, Aug. 2006.

[6] L. Zhang, X. Wu and D. Zhang, "Color Reproduction from Noisy CFA Data of Signal Sensor Digital Cameras", *IEEE Trans. on Image Processing*, vol. 16, no. 9, pp 2184-2197, Sept. 2007.

[7] D. Alleysson, S. Süsstrunk and J. Hérault, "Linear demosaicing inspired by the human visual system", *IEEE Trans. on Image Processing*, vol. 14, no. 4, pp. 439-449, April 2005.

[8] L. Fang *et al.*, "Joint Demosaicing and Subpixel-Based Down-Sampling for Bayer Images: A Fast Frequency-Domain Analysis Approach", *IEEE Trans. on Multimedia*, vol. 14, no. 4, pp. 1359-1369, Aug. 2013.

[9] B. Goossens *et al.*, "An improved non-local denoising algorithm", in *2008 International Workshop on Local and Non-Local Approximation in Image Processing (LNLA 2008)*, pp. 143-156, 2008.

[10] K. Hirakawa and T.W. Parks,"Adaptive homogeneity-directed demosaicing algoritm", *IEEE Trans on Image Processing*, vol. 14, no. 3, pp. 360-369, March 2005.

[11] K. Dabov *et al.*,"Image denoising by sparse 3-D transform-domain collaborative filtering", *IEEE Trans on Image Processing*, vol. 16, no. 8, pp. 2080-2095, Aug. 2007.