IMAGE DEMOSAICING BY USING ITERATIVE RESIDUAL INTERPOLATION

Wei Ye and Kai-Kuang Ma

School of Electrical and Electronic Engineering
Nanyang Technological University, Singapore 639798
ye0003ei@e.ntu.edu.sg, ekkma@ntu.edu.sg

ABSTRACT

A new demosaicing approach has been introduced recently, which is based on conducting interpolation on the generated *residual* fields rather than on the color-component difference fields as commonly practiced in most demosaicing methods. In view of its attractive performance delivered by such *residual interpolation* (RI) strategy, a new RI-based demosaicing method is proposed in this paper that has shown much improved performance. The key success of our approach lies in that the RI process is iteratively deployed to all the three channels for generating a more accurately reconstructed G channel, from which the R channel and the B channel can be better reconstructed as well. Extensive simulations conducted on two commonly-used test datasets have clearly demonstrated that our algorithm is superior to the existing state-of-the-art demosaicing methods, both on objective performance evaluation and on subjective perceptual quality.

Index Terms— Image demosaicing, regression filter, iterative residual interpolation.

1. INTRODUCTION

A color image captured by a single image sensor is denoted as a *mosaiced* image. This is commonly achieved by using a *color filter array* (CFA), on which the pixels are arranged in the Bayer pattern [1], such that only one of the three primary color components (i.e., R, G, or B) is recorded at each pixel location. *Demosaicing* is the reverse process of mosaicing, which restores an approximate full color image based on the acquired mosaiced image. One strategy commonly practiced in the majority of demosaicing methods [2–7] is to interpolate those missing pixels on the G channel first, for this channel has the highest sampling density among all three channels in the Bayer pattern; to be exact, the number of pixels with known values are extactly twice that of the R channel and of the B channel, respectively. The missing pixels on the R and B channels are then restored by interpolating the color-component difference fields R-G and B-G, respectively. This is based on the observation that the color-component differences are usually locally smooth.

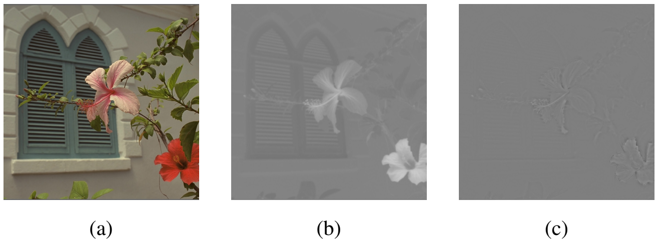


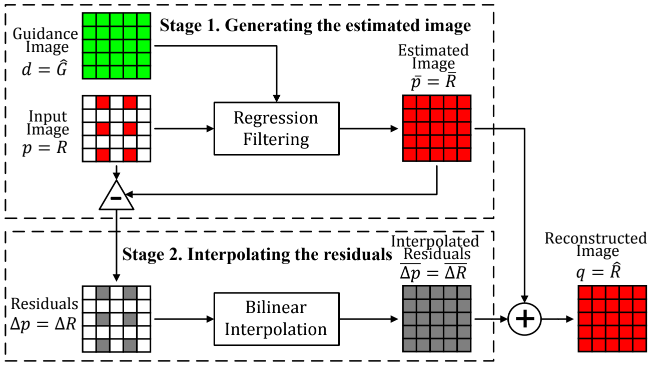
Fig. 1. Comparison of color-component differences and residuals: (a) a full color image, “Kodak 3”; (b) the color-component difference field, R-G; (c) the residual field, R-$\bar{R}$.

Recently, a new demosaicing approach is proposed in [8], which is based on the so-called *residual interpolation* (RI). Instead of conducting interpolation on the color-component difference fields, the demosaicing algorithm as described in [8] performs interpolation over the *residual* fields. Note that a *residual* incurred at each pixel is the difference yielded between a known color value (i.e., ground truth) and its estimated value. If the estimated values are sufficiently close to the ground truth, the resulted residual field will be much smoother than the color-component difference field, especially when there is a sharp *color* transition (e.g., Fig. 1). This indicates that interpolation conducted on the residual fields has a potential to yield a better reconstructed image.

However, it has been observed that the RI-based demosaicing method as described in [8] did not fully explore the RI-based strategy, since the reconstruction of the G channel is mainly based on another approach, called the *gradient-based threshold free* (GBTF) [5]. In this paper, a new demosaicing algorithm is developed by exploiting the RI-based strategy to the G channel, using a new reconstruction scheme that iteratively employs the RI process to all the three channels. Simulation results obtained from extensive experiments conducted on the Kodak and the IMAX datasets have clearly shown that the proposed iterative RI-based demosaicing algorithm outperforms all existing state-of-the-art demosaicing methods.

The rest of this paper is organized as follows. A succinct review of the RI process employed by [8] is given in Section 2. Section 3 describes the proposed demosaicing algorithm in detail. Section 4 conducts performance evaluation of several comparable methods. Section 5 concludes the paper.

978-1-4799-5751-4/14/$31.00 ©2014 IEEE


Stage 1. Generating the estimated image
Guidance Image $d = \hat{G}$
Input Image $p = R$
Regression Filtering
Estimated Image $\bar{p} = \bar{R}$
$-$
Stage 2. Interpolating the residuals
Interpolated Residuals $\overline{\Delta p} = \overline{\Delta R}$
Residuals $\Delta p = \Delta R$
Bilinear Interpolation
$+$
Reconstructed Image $q = \hat{R}$

△ denotes that the subtraction operation is only applied to pixels with known R values

Fig. 2. The RI-based reconstruction scheme for the R channel and for the B channel (not shown).

2. RESIDUAL INTERPOLATION

Residual interpolation (RI) is introduced in [8] for conducting image demosaicing to restore those missing pixel values on an incomplete image p under the guidance of a complete image d. Fig. 2 shows its RI-based reconstruction scheme for restoring the R channel (likewise, for the B channel).

For each color channel under reconstruction, the RI process consists of two stages: (1) generate the estimated channel image and compute the corresponding residual field, and (2) interpolate the residual field for compensating it back to the estimated channel image. In Stage (1), the estimated image $\bar{p}$ is obtained by using the *regression filter* (also known as the *guided filter*) [9] as

$$\bar{p}(i,j) = \frac{\sum_{(u,v)\in w_{i,j}} a(u,v)}{MN} \cdot d(i,j) + \frac{\sum_{(u,v)\in w_{i,j}} b(u,v)}{MN}, \tag{1}$$

where $w_{i,j}$ is the $M \times N$ local window centered at pixel (i,j). The coefficients $a(u,v)$ and $b(u,v)$ are first computed at each pixel location (u,v) by performing linear regression between p and d over the $M \times N$ local window $w_{u,v}$ centered at (u,v). For ease of presentation, the above-described regression filtering process is denoted as: $\bar{p} = \Re(p|d)$. Then, the residuals can be computed as $\Delta p = p - \bar{p}$, involving only those known values. In Stage (2), the bilinear interpolation is exploited to Δp to obtain the estimated $\overline{\Delta p}$ for being compensated back to the estimated image $\bar{p}$ to generate the reconstructed image q.

3. PROPOSED DEMOSAICING ALGORITHM WITH ITERATIVE RESIDUAL INTERPOLATION

A new RI-based demosaicing algorithm is proposed, in which the RI process is deployed through a unified and iterative way to all the three channels to reconstruct a much more accurate G channel, from which the R and B channels can be more accurately restored as well. The proposed iterative RI-based reconstruction scheme will be conducted along the horizontal direction and the vertical direction individually (the order is immaterial). Consequently, two G channel images will be generated and are subject to be linearly combined together to produce the final reconstructed G channel. The horizontal processing flow is depicted in Fig. 3. For the vertical direction, the entire operations will be performed in the same way, except along the vertical direction.

3.1. Iterative Residual Interpolation

Given a Bayer mosaiced image (Fig. 3 (a)), a simple *finite impulse response* (FIR) filter, with three coefficient taps $[1/2, 1, 1/2]$, is first applied to each channel for conducting linear interpolation to produce three interpolated images, $\tilde{R}^{(1)}$, $\tilde{G}^{(1)}$, and $\tilde{B}^{(1)}$ (Fig. 3 (b)). The RI process is then employed to refine each interpolated channel image (as the input image, p) under the guidance of another interpolated channel image (as the guidance image, d). In our approach, the RI process is used as a *refining* process, rather than a reconstruction process as employed in [8]; the RI is exploited not only to the initially interpolated pixel values but also to the ones in the follow-up iterations.

Note that all the three channels *should* help each other during the refining process, since the known values of the three channels are non-overlapped on the Bayer mosaiced image. Therefore, each image in Fig. 3(b) can serve both roles — i.e., as the input image p and as the guidance image d. By using the regression filter, three estimated images, $\bar{R}^{(1)}$, $\bar{G}^{(1)}$, and $\bar{B}^{(1)}$, are generated (Fig. 3 (c)). That is,

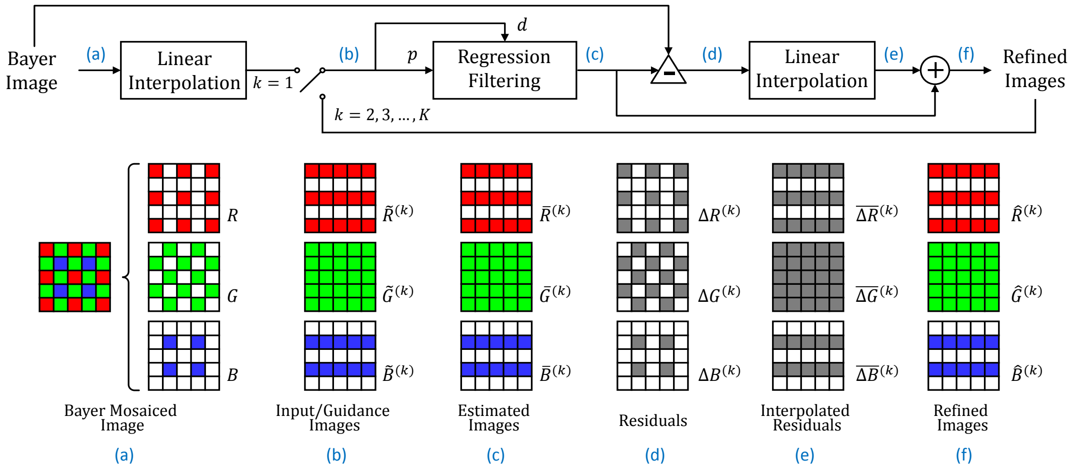
$$\bar{R}^{(1)} = \Re\left(\tilde{R}^{(1)}|\tilde{G}^{(1)}_{\tilde{R}}\right) \text{ and } \bar{B}^{(1)} = \Re\left(\tilde{B}^{(1)}|\tilde{G}^{(1)}_{\tilde{B}}\right), \tag{2}$$

where $\tilde{G}^{(1)}_{\tilde{R}} \cup \tilde{G}^{(1)}_{\tilde{B}} = \tilde{G}^{(1)}$, and the subscript $\tilde{R}$ in $\tilde{G}^{(1)}_{\tilde{R}}$ denotes that only those rows in $\tilde{G}^{(1)}$ that correspond to the rows in $\tilde{R}^{(1)}$ that have values will be involved in the regression filtering process. (The same interpretation goes to $\tilde{G}^{(1)}_{\tilde{B}}$, likewise). As discussed previously, the initially interpolated images $\tilde{G}^{(1)}_{\tilde{R}}$ and $\tilde{G}^{(1)}_{\tilde{B}}$ can be also refined under the guidance of $\tilde{R}^{(1)}$ and of $\tilde{B}^{(1)}$, separately. Therefore,

$$\bar{G}^{(1)}_{\tilde{R}} = \Re\left(\tilde{G}^{(1)}_{\tilde{R}}|\tilde{R}^{(1)}\right) \text{ and } \bar{G}^{(1)}_{\tilde{B}} = \Re\left(\tilde{G}^{(1)}_{\tilde{B}}|\tilde{B}^{(1)}\right). \tag{3}$$

The estimated images as shown in Fig. 3(c) are then subtracted from their corresponding Bayer mosaiced channel images in Fig. 3(a), respectively, to generate the residual fields $\Delta R^{(1)}$, $\Delta G^{(1)}$, and $\Delta B^{(1)}$ as shown in Fig. 3(d). The 3-tap FIR filter used previously is exploited here again to conduct linear interpolation for each residual field along the horizontal direction to generate the interpolated residual fields $\overline{\Delta R}^{(1)}$, $\overline{\Delta G}^{(1)}$, and $\overline{\Delta B}^{(1)}$ as shown in Fig. 3(e). These fields are then added back to the estimated channel images to give the refined images $\hat{R}^{(1)}$, $\hat{G}^{(1)}$, and $\hat{B}^{(1)}$as illustrated in Fig. 3(f). The above-described refining process will be repeated for further refining the channel images through iterations (i.e., $k = 2, 3, \ldots, K$).

978-1-4799-5751-4/14/$31.00 ©2014 IEEE



△ denotes that the subtraction operation is only applied to pixels with known values on the mosaiced channel image

Fig. 3. The proposed iterative RI process in the *horizontal* direction only; likewise, for vertical direction (not shown). Figures (a)-(f) denote the locations of the pixels that are involved in the processing points corresponding to the labels (a)-(f) as indicated in the block diagram, respectively.

3.2. Two Design Issues of the Iterative Process

Two important issues in the iterative RI process need to be further investigated: 1) the window size of the regression filter, and 2) the stopping criterion for the iterative RI process.

It has been observed that if the window size of the regression filter, $M \times N$, is fixed throughout the iterative RI process, the performance turns out to be quite poor. Instead, the window size is gradually increased in our approach, starting from $M^{(1)} = 3$ and $N^{(1)} = 5$ (for $k = 1$). In each subsequent iteration, the window size is set according to $M^{(k)} = M^{(k-1)} + 2$, $N^{(k)} = N^{(k-1)} + 4$. (For the vertical-direction processing, the regression window size becomes $N^{(k)} \times M^{(k)}$.)

As to the stopping criterion, the *magnitude* (denoted as γ) and the *smoothness* (denoted as δ) of the residual fields are considered, since the former is an effective indicator regarding how well the estimated values fit the ground truth, while the *smoothness* is also considered because it reveals how difficult it is to interpolate the residuals accurately. These two quantities are measured for the residual fields generated in each iteration; that is, for iteration k,

$$\gamma^{(k)}(i,j) = \left[\Delta p^{(k)}(i,j)\right]^2, \tag{4}$$

$$\delta^{(k)}(i,j) = \left|\Delta p^{(k)}(i,j+1) - \Delta p^{(k)}(i,j-1)\right|, \tag{5}$$

where $\Delta p^{(k)}$ is $\Delta R^{(k)}$, $\Delta G^{(k)}$, or $\Delta B^{(k)}$. These two quantities, $\gamma^{(k)}(i,j)$ and $\delta^{(k)}(i,j)$, will be further multiplied, and such product computed at each pixel (i,j) will be averaged over the entire image. The computed average value at iteration k is then compared with the one obtained in the previous iteration $(k-1)$. If the current average value is smaller than the previous one, the iterative process will be continued for the next iteration. Otherwise, the iterative process will be stopped (i.e., $k = K$), and the reconstructed G channel image along the horizontal direction will be set as $\hat{G}_H = \hat{G}^{(K-1)}$.

To obtain the reconstructed G channel image along the *vertical* direction, the entire above-described process will be applied to the initially-acquired mosaiced image, except along the vertical direction in each stage of Fig. 3.

3.3. Final Reconstructions of Color Channels

To linearly combine $\hat{G}_H$ and $\hat{G}_V$, the weightage at each pixel location (i,j) needs to be computed and denoted as $w_H(i,j)$ for $\hat{G}_H$ and as $w_V(i,j)$ for $\hat{G}_V$. It is expected that the smaller the magnitude of the residuals and the smoother the residual field, the more reliable the restored value. Hence, the weight can be computed by simply taking the inverse of the product of the magnitude and the smoothness. That is,

$$w_H(i,j) = \left[\gamma_H(i,j) \cdot \delta_H(i,j)\right]^{-1}, \tag{6}$$

where $\gamma_H = f * \gamma^{(K-1)}$ and $\delta_H = f * \delta^{(K-1)}$. Operator $*$ stands for the convolution operation. The Gaussian filter f is deployed here to take the neighboring pixels into consideration for yielding more reliable measurements on $\gamma_H(i,j)$ and $\delta_H(i,j)$. The 5×5 Gaussian kernel with the standard variation 1 is empirically set for f. To compute $w_V(i,j)$, the same operations are applied to the vertical direction. Finally, $\hat{G}_H$ and $\hat{G}_V$ can be linearly combined similarly as in [5–7]

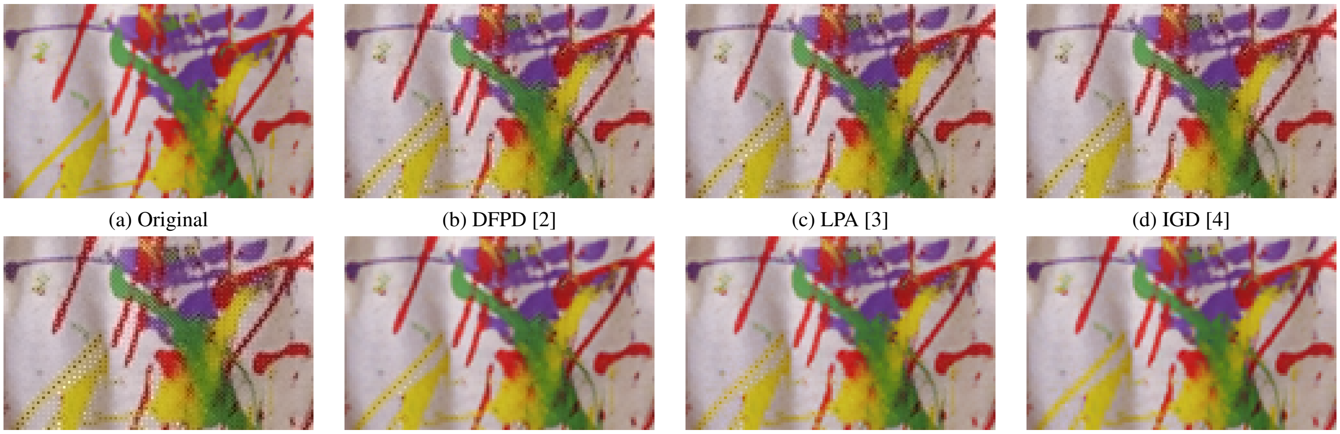
$$\hat{G}(i,j) = \frac{\omega_H(i,j)}{\omega_C(i,j)} \cdot \hat{G}_H(i,j) + \frac{\omega_V(i,j)}{\omega_C(i,j)} \cdot \hat{G}_V(i,j), \tag{7}$$

where $\omega_C(i,j) = \omega_H(i,j) + \omega_V(i,j)$.

Based on the reconstructed G channel, the R channel and the B channel can be individually restored as performed in [8] (i.e., via Fig. 2), except using a different window size for the

Table 1. Average PSNR and CPSNR results (in dB) for the Kodak and the IMAX datasets

Method	Kodak				IMAX				All			
	PSNR			CPSNR	PSNR			CPSNR	PSNR			CPSNR
	R	G	B		R	G	B		R	G	B	
DFPD [2]	40.15	42.45	39.73	40.61	33.78	37.18	32.97	34.25	36.33	39.29	35.68	36.80
LPA [3]	41.59	44.37	40.94	42.05	34.34	37.86	33.29	34.70	37.24	40.46	36.35	37.64
IGD [4]	**41.71**	**44.85**	**41.10**	**42.25**	34.33	37.38	33.45	34.69	37.28	40.37	36.51	37.72
GBTF [5]	41.63	44.75	40.95	42.14	33.54	36.57	32.72	33.92	36.78	39.84	36.01	37.21
LDI-NAT [6]	38.30	40.49	37.94	38.77	36.28	39.76	34.39	36.20	37.09	40.05	35.81	37.23
RI [8]	39.64	42.17	38.87	39.99	36.07	39.99	35.35	36.48	37.50	40.86	36.76	37.88
Proposed	40.26	43.47	39.71	40.84	**36.62**	**40.28**	**35.78**	**36.98**	**38.07**	**41.56**	**37.35**	**38.52**



(a) Original (b) DFPD [2] (c) LPA [3] (d) IGD [4]

(e) GBTF [5] (f) LDI-NAT [6] (g) RI [8] (h) Proposed

Fig. 4. Visual comparison of the demoaicing results for a region in the "IMAX 5" image

regression filtering — 11 × 11 as suggested in [8], but 7 × 7 used in our work due to a better PSNR performance is yielded.

4. EXPERIMENTAL RESULTS

The proposed method is evaluated on the Kodak (12 images) and the IMAX (18 images) datasets that have been commonly used for evaluating demosaicing performance (e.g., [10]). Each color image is first downsampled according to the Bayer pattern [1], followed by conducting demosaicing process using various methods under comparison. The differences between the original and the reconstructed images are measured in terms of the PSNR measured for each channel and the CPSNR computed by combining all three channels together. Our proposed method is compared with six other demosaicing methods: directional filtering and posteriori decision (DFPD) [2], local polynomial approximation (LPA) [3], integrated gradient (IGD) [4], gradient based threshold free demosaicing (GBTF) [5], local directional interpolation and nonlocal adaptive thresholding (LDI-NAT) [6], and the original RI-based demosaicing method (RI) [8].

The PSNRs and CPSNRs are reported in Table 1. In particular, the proposed method outperforms the original RI-based method [8] in both datasets. It has been pointed out in [8] that LPA, IGD, and GBTF work quite well for the Kodak dataset. Despite that our proposed method is slightly inferior to these three methods in terms of objective measurement, however, the subjective performance in terms of perceptual quality is superior. Furthermore, if both datasets are combined and jointly evaluated, our method outperforms all other methods in PSNR, as shown in Table 1.

Besides the objective quality evaluation, our proposed method also shows clear advantages on the subjective quality evaluation. A visual comparison of the demosaicing results for a region with busy colors of the "IMAX 5" image is shown in Fig. 4. Several other test images also clearly demonstrate such superior performance on visual quality assessment.

5. CONCLUSION

In this paper, a novel color image demosaicing method that further exploits the potential of the RI-based demosaicing strategy is proposed by introducing an iterative residual interpolation scheme for reconstructing a much improved G channel, from which the restoration of the R and B channels will be greatly benefited as well. Extensive simulation results on commonly-used test datasets have clearly shown that our proposed method is able to produce superior demomsaiced images in terms of both objective evaluation in PSNR and subjective assessment on perceptual quality.

6. REFERENCES

[1] B. E. Bayer, “Color imaging array,” U.S. Patent 3971065, 1976.

[2] D. Menon, S. Andriani, and G. Calvagno, “Demosaicing with directional filtering and a posteriori decision,” *IEEE Trans. on Image Processing*, vol. 16, no. 1, pp. 132–141, 2007.

[3] D. Paliy, V. Katkovnik, R. Bilcu, S. Alenius, and K. Egiazarian, “Spatially adaptive color filter array interpolation for noiseless and noisy data,” *International Journal of Imaging Systems and Technology*, vol. 17, no. 3, pp. 105–122, 2007.

[4] K. H. Chung and Y. H. Chan, “Low-complexity color demosaicing algorithm based on integrated gradients,” *Journal of Electronic Imaging*, vol. 19, no. 2, pp. 021104, 2010.

[5] I. Pekkucuksen and Y. Altunbasak, “Gradient based threshold free color filter array interpolation,” in *Proc. of IEEE Int. Conf. on Image Processing*, Hong Kong, 2010, pp. 137–140.

[6] L. Zhang, X. Wu, A. Buades, and X. Li, “Color demosaicking by local directional interpolation and nonlocal adaptive thresholding,” *Journal of Electronic Imaging*, vol. 20, no. 2, pp. 023016, 2011.

[7] I. Pekkucuksen and Y. Altunbasak, “Multiscale gradients based color filter array interpolation,” *IEEE Trans. on Image Processing*, vol. 22, no. 1, pp. 157–165, 2013.

[8] D. Kiku, Y. Monno, M. Tanaka, and M. Okutomi, “Residual interpolation for color image demosaicking,” in *Proc. of IEEE Int. Conf. on Image Processing*, Melbourne, Austrilia, 2013, pp. 2304–2308.

[9] K. He, J. Sun, and X. Tang, “Guided image filtering,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397–1409, 2013.

[10] X. Li, B. Gunturk, and L. Zhang, “Image demosaicing: A systematic survey,” in *Proc. of SPIE*, 2008, pp. 68221J–68221J–15.

978-1-4799-5751-4/14/$31.00 ©2014 IEEE