



Faculty Scholarship

2011

Color demosaicking by local directional interpolation and nonlocal adaptive thresholding

Xiaolin Wu

Follow this and additional works at: https://researchrepository.wvu.edu/faculty_publications

Digital Commons Citation

Wu, Xiaolin, "Color demosaicking by local directional interpolation and nonlocal adaptive thresholding" (2011). *Faculty Scholarship*. 88.

https://researchrepository.wvu.edu/faculty_publications/88

This Article is brought to you for free and open access by The Research Repository @ WVU. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of The Research Repository @ WVU. For more information, please contact ian.harmon@mail.wvu.edu.

Journal of Electronic Imaging

SPIEDigitalLibrary.org/jei

Color demosaicking by local directional interpolation and nonlocal adaptive thresholding

Lei Zhang
Xiaolin Wu
Antoni Buades
Xin Li



Color demosaicking by local directional interpolation and nonlocal adaptive thresholding

Lei Zhang

The Hong Kong Polytechnic University
Department of Computing
Hung Hom, Kowloon, Hong Kong, China
E-mail: cslzhang@comp.polyu.edu.hk

Xiaolin Wu

McMaster University
Department of Electrical and Computer Engineering
Hamilton, Ontario, L8G 4K1, Canada

Antoni Buades

Université Paris Descartes
45, rue des Saints Pères, 7ème étage
75270 Paris Cedex 06, France

Xin Li

West Virginia University
Lane Department of Computer Science and Electrical Engineering
PO Box 6109
Morgantown, West Virginia 26506

Abstract. Single sensor digital color cameras capture only one of the three primary colors at each pixel and a process called color demosaicking (CDM) is used to reconstruct the full color images. Most CDM algorithms assume the existence of high local spectral redundancy in estimating the missing color samples. However, for images with sharp color transitions and high color saturation, such an assumption may be invalid and visually unpleasant CDM errors will occur. In this paper, we exploit the image nonlocal redundancy to improve the local color reproduction result. First, multiple local directional estimates of a missing color sample are computed and fused according to local gradients. Then, nonlocal pixels similar to the estimated pixel are searched to enhance the local estimate. An adaptive thresholding method rather than the commonly used nonlocal means filtering is proposed to improve the local estimate. This allows the final reconstruction to be performed at the structural level as opposed to the pixel level. Experimental results demonstrate that the proposed local directional interpolation and nonlocal adaptive thresholding method outperforms many state-of-the-art CDM methods in reconstructing the edges and reducing color interpolation artifacts, leading to higher visual quality of reproduced color images. © 2011 SPIE and IS&T. [DOI: 10.1117/1.3600632]

1 Introduction

Single sensor (CCD/CMOS) digital color cameras capture images with a color filter array (CFA), such as the Bayer

Paper 10182R received Oct. 26, 2010; revised manuscript received May 3, 2011; accepted for publication May 25, 2011; published online Jun. 27, 2011.

1017-9909/2011/20(2)/023016/16/\$25.00 © 2011 SPIE and IS&T

pattern CFA.¹ At each pixel, only one of the three primary colors (red, green, and blue) is sampled; the missing color samples are estimated by a process called color demosaicking (CDM) to reconstruct full color images. The color reproduction quality depends on the image contents and the employed CDM algorithms.² Various CDM algorithms^{3–18} have been proposed in the past decades. The classical second order Laplacian correction^{3,4} (SOLC) algorithm is one of the benchmark CDM schemes due to its simplicity and efficiency. The recently developed methods include the successive approximation-based CDM by Li,⁹ the adaptive homogeneity CDM by Hiraoka et al.,¹⁰ the directional linear minimum mean square-error estimation (DLMMSE)-based CDM method by Zhang et al.,¹² the directional filtering and *a posteriori* decision CDM by Menon et al.,¹³ the sparse representation-based method by Mairal et al.,¹⁴ and the nonlocal means-based self-similarity driven (SSD) method by Buades et al.,¹⁵ etc. A recent review of CDM methods can be found in Ref. 19.

Most of the existing CDM methods assume high local spectral correlations. This assumption may well be valid for images such as those in the Kodak dataset.²⁰ The Kodak dataset was not originally released for CDM but it has been widely used as a benchmark dataset in evaluating CDM algorithms. Inadvertently, the Kodak dataset misled the research of CDM to some extent. It was pointed out in Refs. 15, 16, and 19 that images in the Kodak dataset have much higher spectral correlation, lower color saturation, and smaller chrominance.

matic gradients than images in other datasets, e.g., the McMaster dataset used in this paper (refer to Sec. 3.1 for more information). Compared with the digital color images captured by current digital cameras, the images in Kodak dataset are smoother and less saturated, and hence they are less representative for the applications such as CDM. On the McMaster dataset, the simple SOLC method outperforms many lately developed more complex methods. The reason appears to be that these methods were developed aiming to reproduce the problematic Kodak images, without considering a wider range of test images.

In natural images, the spectral correlation is often weak around object boundaries. Consequently, many CDM algorithms derived under the assumption of high spectral correlation may fail in areas of edges. One way to improve color reproduction near edges is to exploit the nonlocal spatial and spectral redundancies. In natural images, there can be many similar structures/patterns throughout the scene. The most similar pixels to the given one can be far from it. Thus, we can relax the constraint of local neighborhood to nonlocal neighborhood when enhancing the given pixel. The nonlocal means (NLM) filters have been widely used in image processing, such as denoising and deblurring.^{21–26} The mathematical framework of NLM denoising was well established by Buades et al.,²¹ where the given pixel x is estimated as the weighted average of all pixels whose Gaussian neighborhoods look like the neighborhood of x . The recently developed SSD algorithm by Buades et al.¹⁵ is an NLM-based CDM scheme. In Sec. 3 we will see that SSD has similar peak signal-to-noise ratio (PSNR) results to the classical SOLC algorithm on the McMaster dataset but it achieves much better perceptual quality.

In this paper, we propose to couple local directional interpolation (LDI) with nonlocal enhancement for a more effective CDM. The employed CDM strategy is very simple: initial local CDM by LDI, followed by a nonlocal enhancement process. In the initial CDM, only the local spatial-spectral correlation within a compact local window is exploited to avoid CDM errors caused by high color variations around color edges of high saturation. Since directional information is crucial for edge preservation, we use directional filters to interpolate the missing color samples. The obtained directional estimates are then fused according to the local directional gradients. The results of LDI can be augmented by exploiting nonlocal redundancy to reduce initial CDM errors. The similar pixels to the estimated pixel are chosen by patch matching (in practice, a relatively large local window is used), and the matched pixels are used to enhance the initial CDM result.

A straightforward way to utilize nonlocal redundancy is NLM filtering, as in NLM denoising^{21–25} and the SSD method.¹⁵ With NLM, an initially demosaicked pixel is re-estimated as the weighted average of the similar pixels to it. Although NLM can remove much of the CDM noise (i.e., initial CDM errors), it blurs sharp edges and fails to remove bad color artifacts accompanying high saturation object boundaries. To overcome these drawbacks, we propose a novel adaptive thresholding method to make better use of nonlocal redundancy than the NLM filtering. Different from NLM filtering, which directly applies weighted average to the pixel to be enhanced, we model the local patch centered on the pixel as a signal vector and compute the statistics of this vector for

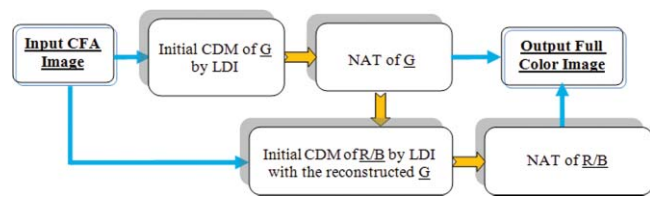


Fig. 1 Flowchart of the proposed CDM method.

processing. By using the nonlocal redundancy, we adaptively compute the optimal transformation domain in which the given patch is decorrelated, and then apply soft thresholding in the transformation domain for filtering. The experimental results in Sec. 3 clearly demonstrate that the proposed LDI and nonlocal adaptive thresholding (NAT)-based method outperforms most of the existing CDM methods, including the recently developed NLM-based SSD algorithm. Compared with NLM filtering, NAT works on the structural level instead of the pixel level. Therefore, it preserves sharp edges much better and removes more color artifacts than NLM.

The rest of the paper is organized as follows. Section 2 describes in detail the proposed local directional interpolation and nonlocal adaptive thresholding (LDI-NAT) scheme for CDM. Section 3 presents the experimental results and Sec. 4 concludes the paper.

2 Proposed Color Demosaicking Algorithm

2.1 Strategy and Flowchart

Figure 1 illustrates the flowchart of the proposed CDM algorithm. First, an initial interpolation is applied to the green (G) channel by LDI and fusion. Second, the NAT is applied to enhance the interpolated G channel. In the third step, the red (R) and blue (B) channels are initially interpolated by the help of the reconstructed G channel. Finally, NAT is applied to the R and B channels so that the whole CDM is completed.

One key issue in the initial CDM is the use of local and directional information. In high saturation areas of natural images, the change of colors is abrupt. Therefore, if we use too many local neighbors to estimate the missing color samples, unexpected errors can be introduced and they can be difficult to remove in the stage of nonlocal enhancement. On the other hand, the preservation of edges is crucial to the visual quality of reconstructed color images. Since edges usually have one or more dominant directions, the interpolation should be along, instead of across, the edge main directions.

With the above considerations, we propose an LDI scheme for initial CDM (the detailed description of LDI is in Sec. 2.2). We will use an example to explain why the strategy of LDI is adopted for initial CDM. Figure 2(a) shows a small patch where there are sharp color transitions (from red to white) in it. Figure 2(b) shows the green and red color difference image (i.e., G-R) of Fig. 2(a). In Fig. 2(c), we plot the color difference signals (with the origin being the center of the patch) along four directions: horizontal (dh), vertical (dv), 45 deg diagonal (d45), and 135 deg diagonal (d135). Some observations can be made from this example.

First, the assumption of smooth color difference used in many CDM methods is invalid. Particularly, from Fig. 2(c) we see that the color differences outside the two-pixel-wide neighborhood are very different from the center one. Therefore, using a big local window (e.g., bigger than 5×5) to

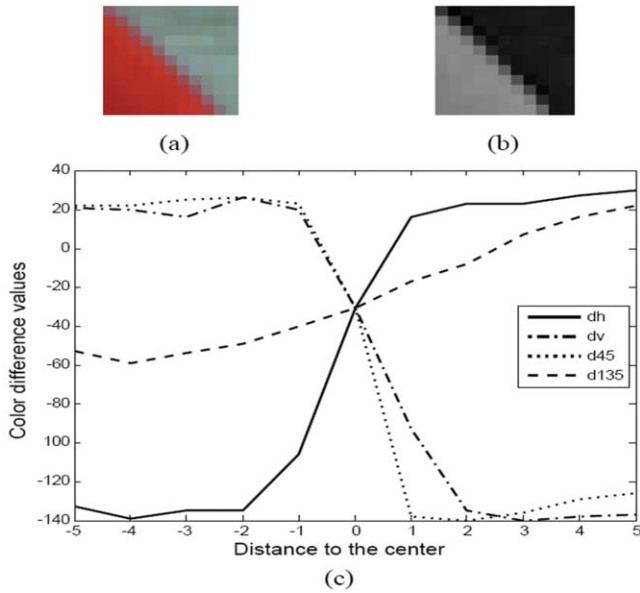


Fig. 2 (a) Cropped and zoomed full color patch; (b) green and red color difference image of (a); and (c) color difference signals along horizontal (dh), vertical (dv), 45 deg diagonal (d45), and 135 deg diagonal (d135) directions at the center of color difference image (b).

estimate the missing color samples can result in unexpected errors. In other words, a compact local window should be used in the initial CDM of high saturation areas. Second, the color edge direction information is very useful for color interpolation. From Fig. 2(c), we see that the color difference along the 135 deg diagonal direction is much smoother than other directions, and hence it should contribute more to the color estimation. Due to the color down-sampling in the mosaic CFA pattern, the color difference signal G-R along diagonal directions cannot be directly calculated. In practice, they are estimated as the weighted average of color differences in other directions.

2.2 Local Directional Interpolation of Green Channel

In various CFA patterns, such as the Bayer pattern,¹ the sampling frequency of G is higher than that of R and B channels. Therefore, the G channel preserves much more image structural information than the other two color channels. Usually, a better reconstruction of G will lead to a better reconstruction of R and B. As shown in Fig. 1, we will initially interpolate the G channel by using local redundancy, and then enhance it by using nonlocal redundancy.

$$\begin{cases} \nabla_n = |G_2 - G_4| + |R_0 - R_{10}| + \frac{1}{2} |G_1 - G_{14}| + \frac{1}{2} |G_3 - G_{15}| + \varepsilon \\ \nabla_s = |G_2 - G_4| + |R_0 - R_{12}| + \frac{1}{2} |G_1 - G_{19}| + \frac{1}{2} |G_3 - G_{18}| + \varepsilon \\ \nabla_w = |G_1 - G_3| + |R_0 - R_9| + \frac{1}{2} |G_2 - G_{13}| + \frac{1}{2} |G_4 - G_{20}| + \varepsilon \\ \nabla_e = |G_1 - G_3| + |R_0 - R_{11}| + \frac{1}{2} |G_2 - G_{16}| + \frac{1}{2} |G_4 - G_{17}| + \varepsilon \end{cases}, \quad (2)$$

where ε is a small positive number to avoid the gradient being zero.

In general, a bigger gradient along a direction means more variations in that direction and, hence, it is more difficult to

The well-known SOLC algorithm^{3,4} is actually a directional interpolation method. In SOLC, at each R or B position two filtering outputs of G are computed along horizontal and vertical directions, respectively, and then one of them is selected based on the gradients in the two directions. However, SOLC has two problems. First, it considers only two directions in the interpolation. This limits its capability in preserving edge structures along other directions. Second, SOLC simply selects one of the two directions for interpolation, but this will lose much useful information in the local area, resulting in many interpolation errors. In this section, we propose to fuse the directional information for more robust color interpolation.

Since there can be sharp color transitions in highly saturated regions, we use a compact local window for the initial interpolation. Considering a CFA block (refer to Fig. 3) and we will focus on the red pixel R_0 , where the green color is to be estimated. (The missing green colors on blue pixels can be similarly interpolated.) Intuitively, if we could know the color difference between G and R at position R_0 , denoted by $d_{gr} = G_0 - R_0$, the missing green sample can then be recovered as $G_0 = R_0 + d_{gr}$. Therefore, how to estimate the color difference d_{gr} is a key in the interpolation of G.

We compute the color difference along four directions: north (n), south (s), west (w), and east (e). Referring to Fig. 3, the four directional estimates, d_{gr}^n , d_{gr}^s , d_{gr}^w , and d_{gr}^e , are calculated as follows:

$$\begin{cases} d_{gr}^n = G_2 - (R_0 + R_{10}) / 2 \\ d_{gr}^s = G_4 - (R_0 + R_{12}) / 2 \\ d_{gr}^w = G_1 - (R_0 + R_9) / 2 \\ d_{gr}^e = G_3 - (R_0 + R_{11}) / 2 \end{cases}. \quad (1)$$

The interpolation error of the four directional estimates relates to the edge direction and color transition at R_0 . In order to evaluate which estimate is better, we calculate the gradients at R_0 along the four directions. There are many forms to define the directional gradients at R_0 . We have the following considerations. First, the gradient should be calculated using the pixels from the same channel; second, to make the calculation of gradients more stable, we could involve neighboring columns/rows of the central column/row in calculation; third, the central column/row should have higher contribution to the gradient than the neighboring columns/rows. Based on the above three considerations, we use the following formula to calculate the gradients along north, south, west, and east directions:

accurately estimate the color difference, vice versa. Therefore, we can use the gradients as indices to weight the four estimates into a final estimate. An optimal weighting scheme needs to know the joint distribution of the gradient and the

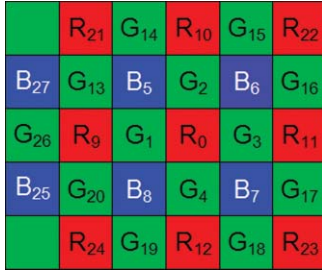


Fig. 3 CFA block.

color difference. However, such information is unknown in advance or difficult to estimate online. In this paper, we simply let the weight assigned to a directional estimate be inversely proportional to the gradient along that direction:

$$w_n = \frac{1}{\nabla_n}, \quad w_s = \frac{1}{\nabla_s}, \quad w_w = \frac{1}{\nabla_w}, \quad w_e = \frac{1}{\nabla_e}. \quad (3)$$

We then normalize the four weights to make the sum of them be 1. There is

$$\bar{w}_n = \frac{w_n}{C}, \quad \bar{w}_s = \frac{w_s}{C}, \quad \bar{w}_w = \frac{w_w}{C}, \quad \bar{w}_e = \frac{w_e}{C}, \quad (4)$$

where $C = w_n + w_s + w_w + w_e$. The four directional estimates are then fused into one estimation

$$\hat{d}_{gr} = \bar{w}_n d_{gr}^n + \bar{w}_s d_{gr}^s + \bar{w}_w d_{gr}^w + \bar{w}_e d_{gr}^e. \quad (5)$$

Finally, the missing green component at R_0 can be estimated as

$$\hat{G}_0 = R_0 + \hat{d}_{gr}. \quad (6)$$

By applying the above procedures to all of the R and B positions, we can reconstruct the G channel.

2.3 Nonlocal Enhancement of G Channel

By using the method described in Sec. 2.2, an initial estimate of each missing green sample can be obtained. Since only the local redundancy in a compact local window is exploited, the interpolation may not be accurate, especially around object boundaries where sharp color or intensity changes will occur. Fortunately, in natural images there are many similar patterns or structures, while a similar structure to the given one may appear far from it. Such nonlocal redundancy can be exploited to enhance the CDM results. The nonlocal means (NLM) technique has been extensively studied and effectively used in image/video denoising and restoration,²¹⁻²⁶ and recently it has also been successfully used in CDM.¹⁵ In this section, we use the nonlocal redundancy to reduce the initial interpolation errors and enhance the color reproduction quality of G channel.

2.3.1 Nonlocal enhancement by nonlocal means filtering

One straightforward way for the nonlocal enhancement of G channel is to apply NLM filtering to the interpolated green sample \hat{G}_0 , as in many NLM-based denoising works.²¹⁻²⁵ To this end, we search for similar pixels (can be either original

green samples or interpolated green samples) to the given \hat{G}_0 in the recovered G image. The searching can be performed in the whole image; however, this is computationally prohibitive and is not necessary. In practice, we search for similar pixels to \hat{G}_0 in a large enough window (e.g., a 31×31 window), denoted by Ω , centered on it. The patch-based method can be used to determine the similarity between \hat{G}_0 and other pixels in Ω . Denoted by \mathbf{P}_0 the $s \times s$ patch is centered on \hat{G}_0 , and by \mathbf{P}_i the $s \times s$ patch is centered on a green pixel G_i in Ω . The l_1 -norm distance between \mathbf{P}_0 and \mathbf{P}_i is computed as

$$d_i = \|\mathbf{P}_0 - \mathbf{P}_i\|_1 = \frac{1}{s^2} \sum_{k=1}^s \sum_{l=1}^s |\mathbf{P}_0(k, l) - \mathbf{P}_i(k, l)|. \quad (7)$$

In general, the smaller the distance d_i is, the more similar G_i is to \hat{G}_0 . Based on d_i , we select the N most similar pixels to \hat{G}_0 (including \hat{G}_0 itself) for the nonlocal enhancement of \hat{G}_0 .

For the convenience of expression, we denote by z_0 the given pixel \hat{G}_0 , by z_n , $n = 1, \dots, N - 1$, the searched similar pixels to \hat{G}_0 , and by d_n the associated distance of z_n . The nonlocal enhancement output of \hat{G}_0 by NLM filtering, denoted by \hat{x}_0 , is computed as the weighted average of z_n

$$\hat{x}_0 = \sum_{n=0}^{N-1} w_n z_n, \quad (8)$$

where the weights w_n are set as

$$w_n = \exp(-d_n/\sigma)/C, \quad (9)$$

with $C = \sum_{n=0}^{N-1} \exp(-d_n/\sigma)$ being the normalization factor to make the sum of w_n be 1. In Eq. (9), parameter σ controls the decay rate of weight w_n w.r.t. distance d_n . In the literature of image denoising, σ is usually preset according to the standard deviation of the noise in the image. In the SSD algorithm for CDM,¹⁵ a coarse-to-fine strategy was used. The nonlocal average process is iterated three times, and the parameter σ is set smaller and smaller in the three iterations.

2.3.2 Nonlocal enhancement by nonlocal adaptive thresholding

The NLM filtering-based nonlocal enhancement of \hat{G}_0 is actually the weighted average of samples z_0, z_1, \dots, z_{N-1} . Although it can suppress many interpolation errors generated in the initial CDM and lead to much better color reproduction than many existing CDM algorithms (refer to Sec. 3.2), it may also smooth the edges and some bad color artifacts around object boundaries can still survive. Nonetheless, in NLM the local neighboring pixels to \hat{G}_0 in the patch \mathbf{P}_0 , which altogether form the local pattern (i.e., structure) on \hat{G}_0 , are only used to determine the weights w_n for averaging. Actually, \mathbf{P}_0 and the similar patches \mathbf{P}_i to it also specify the variations of the local pattern on \hat{G}_0 . This information is not efficiently exploited in NLM weighting. To more effectively exploit the nonlocal redundancy, we propose a nonlocal adaptive thresholding (NAT) scheme in this section.

By viewing the initial CDM error as additive noise, the initial CDM output can be modeled as $\mathbf{y} = \mathbf{x} + \mathbf{v}$, where \mathbf{x} is the true signal to be restored, \mathbf{v} is additive noise, and \mathbf{y} is the initial CDM result. To robustly estimate the original signal \mathbf{x} from the degraded observation \mathbf{y} , a regularized solu-

tion is often desired such that $\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} J(\mathbf{x})$ s.t. $\|\mathbf{y} - \mathbf{x}\|_2 \leq \tau$, where $J(\mathbf{x})$ is the regularization term and τ is a small number. For example, in the total variational (TV) based image restoration,²⁷⁻³⁰ $J(\mathbf{x})$ is the l_1 -norm of the gradients of \mathbf{x} . Recently, the sparsity prior of \mathbf{x} has been successfully used for image restoration.^{14,31-36} By assuming that the signal \mathbf{x} can be sparsely coded (i.e., represented) by a dictionary of atoms Ψ , i.e., $\mathbf{x} \approx \Psi\alpha$ and most of the coefficients in α are small, the sparsity-based estimation of \mathbf{x} from \mathbf{y} can be obtained via l_1 -norm minimization

$$\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_1 \text{ s.t. } \|\mathbf{y} - \Psi\alpha\|_2 \leq \tau. \quad (10)$$

The above l_1 -norm minimization problem can be solved by standard convex optimization techniques³⁷ or by the iterative shrinkage methods.³⁶ The sparse representation modeling has led to many interesting results in image processing, such as compressive sensing³⁸⁻⁴⁰ and denoising.^{33,35}

In our problem of the nonlocal enhancement of \hat{G}_0 , we denote by $\mathbf{y}_0 = [y_0, y_1, \dots, y_{M-1}]^T$, where $M = s^2$. The column vector \mathbf{y}_0 contains the samples in the $s \times s$ patch \mathbf{P}_0 centered on \hat{G}_0 , and it can be viewed as the observation of the unknown true signal $\mathbf{x}_0 = [x_0, x_1, \dots, x_{M-1}]^T$. Then we have $\mathbf{y}_0 = \mathbf{x}_0 + \mathbf{v}_0$, where \mathbf{v}_0 represents the initial CDM error. Once a good estimation of \mathbf{x}_0 , denoted by $\hat{\mathbf{x}}_0$, can be made from \mathbf{y}_0 , the nonlocal enhancement of \hat{G}_0 , denoted by $\hat{\mathbf{x}}_0$, can be readily extracted from $\hat{\mathbf{x}}_0$. Since the elements in patch \mathbf{P}_0 are highly correlated, it can be assumed that signal \mathbf{x}_0 is sparse in some domain Ψ , i.e., $\mathbf{x}_0 \approx \Psi\alpha_0$ and α_0 is a sparse coefficient vector. The enhancement of \hat{G}_0 can be modeled as

$$\hat{\alpha}_0 = \arg \min_{\alpha_0} \|\alpha_0\|_1 \text{ s.t. } \|\mathbf{y}_0 - \Psi\alpha_0\|_2 \leq \tau. \quad (11)$$

Once $\hat{\alpha}_0$ is optimized, the estimated signal can be obtained as $\hat{\mathbf{x}}_0 = \Psi \cdot \hat{\alpha}_0$.

Now the question is how to determine the sparse domain Ψ in Eq. (11) to solve $\hat{\alpha}_0$. Although the wavelet bases or the Fourier bases are often used, these analytically designed bases cannot effectively characterize the many different local patterns across the image. The dictionary learning^{33,34} methods have been recently proposed to learn an over-complete dictionary of bases from a training dataset to span the sparse domain. Nonetheless, for a given signal \mathbf{y}_0 , many atoms in the learned over-complete dictionary will be irrelevant, while the l_1 -norm minimization needs much computational cost. With these considerations, in this paper we propose the NAT scheme to solve Eq. (11) with nonlocal redundancy.

Recall that after the nonlocal similar pixels searching to \hat{G}_0 , we obtain $N-1$ similar patches to \mathbf{P}_0 . The vector \mathbf{y}_0 is formed by stretching \mathbf{P}_0 , and similarly we can form another $N-1$ column vectors by stretching \mathbf{P}_i , $i = 1, 2, \dots, N-1$. Denote by $\mathbf{y}_0 = [y_{0,0}, y_{0,1}, \dots, y_{0,M-1}]^T$ the vector formed by \mathbf{P}_0 , and by $\mathbf{y}_i = [y_{i,0}, y_{i,1}, \dots, y_{i,M-1}]^T$ the vectors formed by other patches. Then, an $M \times N$ data matrix \mathbf{Y} can be established by $\mathbf{Y} = [\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{N-1}]$. Each row of \mathbf{Y} is then centralized by subtracting its mean value. For the convenience of expression, we still use the symbol \mathbf{Y} in the following development.

Since $\mathbf{y}_i = \mathbf{x}_i + \mathbf{v}_i$, where \mathbf{x}_i is the unknown true signal and \mathbf{v}_i is the initial CDM error, we have $\mathbf{Y} = \mathbf{X} + \mathbf{V}$, where $\mathbf{X} = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{N-1}]$ and $\mathbf{V} = [\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{N-1}]$. A good

domain Ψ for \mathbf{X} should be a domain where the vectors \mathbf{x}_i could be sparsely coded; that is, $\mathbf{X} \approx \Psi\Lambda$ and Λ is a sparse matrix. Since only the observation of \mathbf{X} , i.e., \mathbf{Y} , is available, we set the objective function to determine Ψ as

$$\arg \min_{\Lambda, \Psi} \|\Lambda\|_1 \text{ s.t. } \|\mathbf{Y} - \Psi\Lambda\|_F \leq \tau, \quad (12)$$

where $\|\cdot\|_F$ is the Frobenius norm.

Equation (12) is a joint optimization problem of Λ and Ψ , which can be solved by alternatively optimizing Λ and Ψ . Considering that the average power of CDM error \mathbf{V} is not seriously high (but the resulting color artifacts can be visually very unpleasing), here we propose an efficient solution to Eq. (12). By using singular value decomposition (SVD), we can factorize \mathbf{Y} as $\mathbf{Y} = \Phi\Gamma$, where Φ is an orthonormal matrix spanned by the eigenvectors of the covariance matrix of \mathbf{Y} (i.e., $\mathbf{Y}\mathbf{Y}^T$), and $\Gamma = \Phi^T\mathbf{Y}$ is the projection of \mathbf{Y} over Φ^T . We let the desired dictionary $\Psi = \Phi$. If we also let $\Lambda = \Gamma$, then the constraint $\|\mathbf{Y} - \Psi\Lambda\|_F = \|\mathbf{Y} - \Phi\Gamma\|_F = 0 \leq \tau$ is perfectly satisfied but $\|\Lambda\|_1$ will have a certain amount so that $\arg \min_{\Lambda} \|\Lambda\|_1$ is not optimized.

Thus, Γ needs to be further processed for a better solution to Λ .

With $\mathbf{Y} = \mathbf{X} + \mathbf{V}$, we have $\Gamma = \Phi^T\mathbf{Y} = \Phi^T\mathbf{X} + \Phi^T\mathbf{V} = \Gamma_X + \Gamma_V$, where $\Gamma_X = \Phi^T\mathbf{X}$ and $\Gamma_V = \Phi^T\mathbf{V}$. Φ^T will decorrelate true signal \mathbf{X} and many coefficients in Γ_X will be small, while there are a few significant coefficients in Γ_X and they are mainly the projection coefficients of \mathbf{X} on the eigenvectors associated with the most significant eigenvalues of $\mathbf{Y}\mathbf{Y}^T$. The CDM errors \mathbf{V} are very like random noise, and thus, the energy of Γ_V will be evenly spread over the domain spanned by Φ^T . Therefore, we could apply a soft threshold t to Γ to remove Γ_V from Γ so that the desired Λ can be obtained as follows:

$$\Lambda(i, j) = \begin{cases} \text{sign}(\Gamma(i, j)) \cdot (|\Gamma(i, j)| - t) & \text{if } \Gamma(i, j) > t \\ 0 & \text{if } \Gamma(i, j) \leq t \end{cases}. \quad (13)$$

Actually soft-thresholding is widely used to solve the l_1 -norm minimization problems.^{30,36,41} With Eq. (13), the term $\|\Lambda\|_1$ is much reduced while the term $\|\mathbf{Y} - \Psi\Lambda\|_F$ can still be controlled within a small range τ , and finally a good solution to Eq. (12) is obtained.

The selection of threshold t depends on the CDM error level in \mathbf{V} . In practice, we can estimate t as follows. The CDM accuracy of a local area is closely related to its local smoothness. If the local area is smooth, usually the CDM error will be low, and vice versa. Therefore, we empirically estimate the CDM error level based on the local intensity variation. We calculate the average gradient magnitude of all patches in \mathbf{Y} and denoted it by g_Y , and then we let $t = c \cdot g_Y$, where c is a constant. By experience, we set $c = 0.03$ in the experiments.

Once the solution to Λ in Eq. (12) is obtained, the desired solution $\hat{\alpha}_0$ in Eq. (11) is obtained by extracting it from Λ . The nonlocal enhancement result of \hat{G}_0 is $\hat{\mathbf{x}}_0 = \Psi \cdot \hat{\alpha}_0$. From the above description, we can see that NLN applies nonlocal enhancement to \hat{G}_0 by weighted averaging, while NAT applies nonlocal enhancement to the local patch centered on \hat{G}_0 . In other words, NAT lifts NLN from

the pixel level to the structure level. Consequently, NAT can reconstruct much better the image edges than NLM, as we will see in the experimental results in Sec. 3.

2.4 Initial Interpolation of R and B Channels

With the nonlocally enhanced G channel, we first compute the initial estimates of R and B channels by exploiting the local spatial-spectral correlation, and then enhance them by nonlocal redundancy. Since the interpolations of R and B channels are symmetric, in the following we only discuss the reconstruction of B.

We interpolate the missing B samples by using a two-step strategy. First, we interpolate the B samples at the R positions and then, with these interpolated B samples, all the other B samples at the G positions can be interpolated. Referring to Fig. 3, suppose we are to interpolate the missing sample B_0 at R_0 . Note that all the G samples have been recovered and are now available, and we can estimate the color differences between B and G along the four diagonal directions at R_0 as

$$\begin{cases} d_{bg}^{nw} = B_5 - G_5 \\ d_{bg}^{ne} = B_6 - G_6 \\ d_{bg}^{se} = B_7 - G_7 \\ d_{bg}^{sw} = B_8 - G_8 \end{cases}, \quad (14)$$

where the superscripts “nw,” “ne,” “se,” and “sw” represent the north-western, north-east, south-east, and south-western directions, respectively.

The four directional estimates are weighted for a more robust estimate. To determine the weights, the gradients along the four directions are calculated as follows:

$$\begin{cases} \nabla_{nw} = |B_5 - B_7| + |R_{21} - R_0| + |G_5 - G_0| + \varepsilon \\ \nabla_{ne} = |B_6 - B_8| + |R_{22} - R_0| + |G_6 - G_0| + \varepsilon \\ \nabla_{se} = |B_5 - B_7| + |R_{23} - R_0| + |G_7 - G_0| + \varepsilon \\ \nabla_{sw} = |B_6 - B_8| + |R_{20} - R_0| + |G_8 - G_0| + \varepsilon \end{cases}, \quad (15)$$

where ε is a small positive number. As in Eqs. (3) and (4), the four weights are set as

$$\begin{aligned} \bar{w}_{nw} &= \frac{1}{C \cdot \nabla_{nw}}, & \bar{w}_{ne} &= \frac{1}{C \cdot \nabla_{ne}}, \\ \bar{w}_{se} &= \frac{1}{C \cdot \nabla_{se}}, & \bar{w}_{sw} &= \frac{1}{C \cdot \nabla_{sw}}, \end{aligned} \quad (16)$$

where $C = 1/\nabla_{nw} + 1/\nabla_{ne} + 1/\nabla_{se} + 1/\nabla_{sw}$. Then, the final blue and green color difference at position R_0 is estimated by $\hat{d}_{bg} = \bar{w}_{nw}d_{bg}^{nw} + \bar{w}_{ne}d_{bg}^{ne} + \bar{w}_{se}d_{bg}^{se} + \bar{w}_{sw}d_{bg}^{sw}$, and the missing blue component at R_0 is estimated as $\hat{B}_0 = G_0 + \hat{d}_{bg}$.

Once the B samples at the R positions are interpolated as described above, we can consequently interpolate the B samples at all the other G positions. Take the position G_1 in Fig. 3 as an example. Note that the blue samples at R_9 and R_0 have been interpolated, and we denote them as \hat{B}_9 and \hat{B}_0 . The directional estimates of the blue and green color difference at G_1 are computed as

$$\begin{cases} d_{bg}^n = B_5 - G_5 \\ d_{bg}^s = B_8 - G_8 \\ d_{bg}^w = \hat{B}_9 - G_9 \\ d_{bg}^e = \hat{B}_0 - G_0 \end{cases}. \quad (17)$$

The gradients at position G_1 along the four directions are calculated as

$$\begin{cases} \nabla_n = |G_{14} - G_1| + |B_5 - B_8| + \frac{1}{2}|R_{21} - R_9| + \frac{1}{2}|R_{10} - R_0| + \varepsilon \\ \nabla_s = |G_{19} - G_1| + |B_5 - B_8| + \frac{1}{2}|R_{20} - R_9| + \frac{1}{2}|R_{12} - R_0| + \varepsilon \\ \nabla_w = |G_1 - G_{26}| + |R_0 - R_9| + \frac{1}{2}|B_{27} - B_5| + \frac{1}{2}|B_{25} - B_8| + \varepsilon \\ \nabla_e = |G_1 - G_3| + |R_0 - R_9| + \frac{1}{2}|B_5 - B_6| + \frac{1}{2}|B_8 - B_7| + \varepsilon \end{cases}. \quad (18)$$

The associated four weights for the four directions are computed in the same way as in Eqs. (3) and (4), and the fused color difference is obtained as $\hat{d}_{bg} = \bar{w}_n d_{bg}^n + \bar{w}_s d_{bg}^s + \bar{w}_w d_{bg}^w + \bar{w}_e d_{bg}^e$. Finally, the missing B sample at position G_1 is interpolated by $\hat{B}_1 = G_1 + \hat{d}_{bg}$.

2.5 Nonlocal Enhancement of R and B Channels

Once the R and B channels are interpolated with the help of nonlocally enhanced G channel, they can then be enhanced by exploiting nonlocal redundancies in R and B channels, respectively. The process is the same as that for the G channel. For each interpolated red (blue) sample \hat{R}_0 (\hat{B}_0), we search for similar pixels to it in a large window centered on it. The N most similar pixels to \hat{R}_0 (\hat{B}_0), including itself, are used to enhance it via NLM or NAT.

3 Experimental Results

3.1 McMaster Dataset

The Kodak image dataset²⁰ is widely used as a standard dataset in CDM and many other color image processing fields. The Kodak dataset contains 24 full color images,

Table 1 Statistics of the Kodak and the McMaster datasets.

Datasets		Kodak	McMaster
Mean spectral correlation	G and R	0.8712	0.7445
	G and B	0.9050	0.7114
Mean saturation		15.6	45.81
Mean chromatic gradient		1.78	4.54

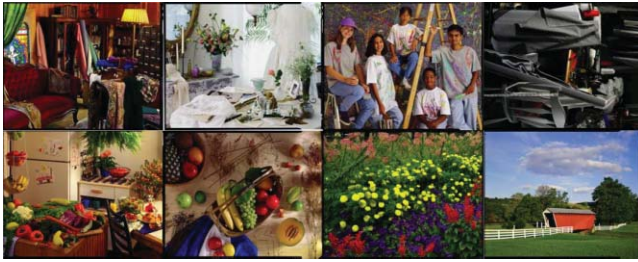


Fig. 4 Scenes of the eight test images in McMaster dataset.

whose spatial size is 768×512 . It is said that these images were originally captured by film and then digitized by scanner. However, in recent years it has been noticed that the statistics of Kodak images are very different from other natural images,^{15,16,19} e.g., the images in the McMaster dataset to be introduced. The images in Kodak dataset look smooth and less saturated, which makes them less representative for the digital color images captured by the current digital cameras and, hence, less representative for applications such as CDM. Specifically, the Kodak images have very high spectral correlation, are smooth in chromatic gradient, and have low saturation (refer to Table 1). It is doubted that these images were post-processed, and they are not suitable for evaluating CDM algorithms.



Fig. 5 Cropped McMaster sub-images (500×500) used in the experiments. From top to bottom and left to right, these sub-images are labeled as 1 to 18.

In this study, we use a new color image dataset, namely the McMaster dataset, for the evaluation of CDM algorithms. This dataset was established in a project of developing new CDM methods by McMaster University, Canada, in collaboration with some industry partners. It has eight

Table 2 Summary of the SSD, LDI-NLM, and LDI-NAT algorithms.

Methods	SSD	LDI-NLM	LDI-NAT
Procedures	<ol style="list-style-type: none"> 1. The full color image is initially interpolated by bilinear interpolator. Denote it by \mathbf{u}_0. 2. For $\sigma = \{16, 4, 1\}$ <ol style="list-style-type: none"> 2a. Apply nonlocal means filtering to \mathbf{u}_0 with scale parameter σ. 2b. Convert \mathbf{u}_0 into YUV color space and apply chromatic regularization to U and V channels. Transform the regularized image back to RGB color space, and denote it by \mathbf{u}. 2c. Let $\mathbf{u}_0 = \mathbf{u}$. <p>End</p>	<ol style="list-style-type: none"> 1. The G channel is initially recovered by LDI. 2. The G channel is nonlocally enhanced by NLM filtering. 3. The R or B channel is initially recovered by LDI with the nonlocally recovered G. 4. The R or B channel is nonlocally enhanced by NLM filtering. 	<ol style="list-style-type: none"> 1. The G channel is initially recovered by LDI. 2. The G channel is nonlocally enhanced by NAT with soft-thresholding. 3. The R or B channel is initially recovered by LDI with the nonlocally recovered G. 4. The R or B channel is nonlocally enhanced by NAT with soft-thresholding.
The way to use nonlocal information	<ol style="list-style-type: none"> 1. A coarse-to-fine strategy is used to iteratively exploit the nonlocal redundancy. In each iteration, the similar pixels to the given one are weighted as the updated estimation. 2. In each iteration, the RGB color space is transformed into the YUV space for chromatic regularization. 	<ol style="list-style-type: none"> 1. The R, G, and B channels are separately enhanced. 2. The similar pixels to the given one are weighted, while the weights are determined based on the distances between similar patches. 	<ol style="list-style-type: none"> 1. The R, G, and B channels are separately enhanced. 2. The pixels in the given patch are taken as a vector signal, which is soft-thresholded in an adaptively computed sparse domain based on the statistics of similar patches.

Table 3 PSNR (dB) results by different CDM methods on the McMaster dataset.

Methods		SOLC (Ref. 3)	AHD (Ref. 10)	SA (Ref. 9)	DLMMSE (Ref. 12)	SSD (Ref. 15)	LDI-NLM	LDI-NAT
1	R	28.26	26.02	23.53	26.94	27.28	28.81	29.29
	G	31.22	29.82	25.17	30.63	30.68	32.31	32.67
	B	26.34	24.04	22.05	24.82	25.12	26.47	26.71
2	R	33.68	32.47	31.63	33.30	33.61	34.66	35.02
	G	37.62	37.20	34.00	37.66	37.81	39.01	39.08
	B	32.11	31.26	30.74	31.86	32.01	32.79	32.92
3	R	30.64	31.10	31.47	32.60	32.81	33.41	33.05
	G	33.73	33.49	32.75	35.28	35.05	35.50	35.51
	B	28.60	29.67	29.80	30.70	30.93	30.99	30.31
4	R	32.80	33.76	34.59	34.70	36.36	37.41	36.25
	G	37.16	35.66	34.05	36.99	38.98	39.01	40.33
	B	30.89	31.48	32.19	32.07	33.49	34.02	33.30
5	R	33.61	29.52	28.60	30.38	31.10	34.50	35.05
	G	36.28	34.73	30.97	35.11	35.43	37.67	38.15
	B	30.47	28.78	28.08	29.41	29.48	31.02	31.16
6	R	37.14	33.92	32.23	34.98	36.09	38.59	39.40
	G	40.30	37.72	32.50	38.61	38.85	41.70	43.42
	B	34.00	29.96	29.14	31.15	31.72	34.21	34.97
7	R	33.85	35.64	37.03	38.30	36.61	36.28	36.09
	G	36.34	37.36	40.39	40.70	37.62	37.66	37.41
	B	32.45	35.07	36.22	37.29	36.38	34.59	34.49
8	R	34.87	34.15	35.31	35.45	35.31	36.89	36.31
	G	39.09	39.45	38.49	41.43	40.34	40.44	40.29
	B	35.04	35.79	35.82	36.99	36.76	36.84	36.67
9	R	34.36	31.54	30.71	32.39	33.72	35.54	35.49
	G	39.62	37.99	33.83	38.73	39.52	41.56	41.73
	B	35.34	34.00	32.54	34.66	35.38	36.54	36.30
10	R	36.86	33.99	34.03	34.70	36.33	37.64	38.26
	G	40.86	39.17	36.15	40.00	40.23	42.19	42.64
	B	36.08	34.88	34.78	35.55	36.13	36.51	36.83
11	R	38.12	36.13	36.16	36.91	38.16	39.25	39.82
	G	40.78	39.34	37.11	40.44	40.19	41.66	42.57
	B	37.19	34.73	34.33	35.75	36.81	37.50	37.66
12	R	37.13	33.60	34.49	34.74	35.37	37.62	38.36
	G	40.17	40.09	37.66	39.59	39.70	41.45	41.49
	B	35.70	36.24	36.24	36.47	37.11	37.51	37.59

Table 3 Cont.

Methods		SOLC (Ref. 3)	AHD (Ref. 10)	SA (Ref. 9)	DLMMSE (Ref. 12)	SSD (Ref. 15)	LDI-NLM	LDI-NAT
13	R	39.80	37.91	38.11	38.66	40.01	42.23	41.77
	G	43.46	42.16	39.90	42.57	43.82	45.55	44.89
	B	37.65	36.20	36.51	36.75	37.19	37.88	38.13
14	R	37.85	37.33	36.82	37.74	38.66	39.28	39.39
	G	41.37	40.65	38.79	41.13	41.93	42.62	42.84
	B	35.64	34.30	34.45	34.78	35.00	35.82	36.12
15	R	36.44	34.88	34.87	35.32	36.23	37.34	36.95
	G	41.20	40.27	38.13	40.71	40.75	42.39	42.68
	B	38.17	36.84	36.52	37.30	37.90	38.49	38.99
16	R	32.75	30.95	28.75	31.95	32.21	34.18	34.97
	G	34.09	32.36	28.60	33.22	32.99	35.00	35.59
	B	31.63	26.85	24.87	28.06	28.30	31.12	31.53
17	R	31.24	27.12	25.35	28.32	29.24	31.60	32.14
	G	35.17	32.13	26.68	33.31	33.62	37.31	37.62
	B	30.69	26.65	25.06	27.77	28.38	30.78	30.91
18	R	32.69	32.30	31.61	33.32	33.24	34.63	34.58
	G	36.20	35.69	33.84	37.02	35.91	37.30	37.27
	B	33.43	31.90	31.11	32.93	33.44	34.87	34.30
Average	R	34.71	33.05	32.68	34.06	34.71	36.10	36.23
	G	38.11	37.10	34.63	38.10	38.08	39.46	39.79
	B	33.41	32.30	31.87	33.15	33.47	34.33	34.38

high resolution (size: 2310×1814) color images that were originally captured by Kodak film and then digitized. The scenes of the eight images are shown in Fig. 4. Since these images are big in size, we crop 18 sub-images (size: 500×500) from them to evaluate the CDM methods. Figure 5 shows the cropped 18 sub-images. In Table 1 we compare the mean spectral correlation, mean chromatic gradient, and mean saturation of the images in the two datasets. [The mean saturation is computed as follows. For each pixel with color components $\{r, g, b\}$, its saturation is computed as $s = \sqrt{[(r - y)^2 + (g - y)^2 + (b - y)^2]}/3$, where $y = (r + g + b)/3$. The mean saturation is obtained by averaging the saturation of all pixels. The mean chromatic gradient is computed as follows. We first convert the image into the YUV space. The chromatic gradient of each pixel is set as the modulus of the gradient in the U and V channels. Then, the mean chromatic gradient is obtained by averaging over the whole image.] We see that the spectral correlation of the Kodak images is obviously higher than that of the McMaster dataset. The McMaster images are more saturated and there are many sharp structures with

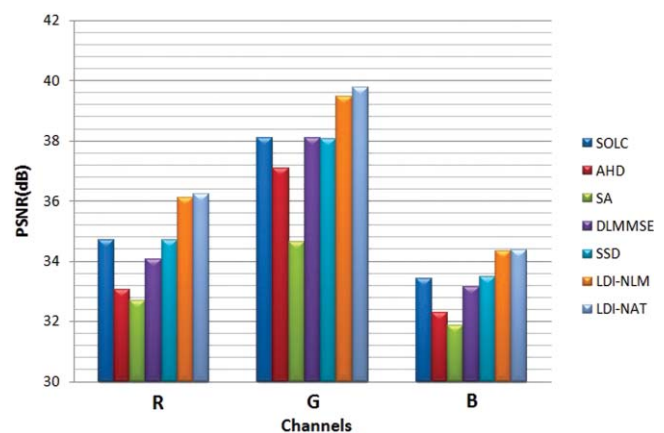


Fig. 6 Graphical presentation of the average PSNR by different methods on the McMaster dataset.

Table 4 Zipper Effect Ratio (ZER) by different CDM methods on the McMaster dataset.

Methods	SOLC (Ref. 3)	AHD (Ref. 10)	SA (Ref. 9)	DLMMSE (Ref. 12)	SSD (Ref. 15)	LDI-NLM	LDI-NAT
1	0.2059	0.1678	0.4348	0.2021	0.0996	0.0748	0.1082
2	0.0939	0.1225	0.1673	0.1249	0.0753	0.0486	0.0682
3	0.1659	0.2336	0.4357	0.2179	0.1044	0.0815	0.1044
4	0.3475	0.3952	0.7680	0.5287	0.0915	0.0930	0.1468
5	0.0996	0.1130	0.1831	0.1144	0.0629	0.0477	0.0591
6	0.0987	0.1206	0.2121	0.1306	0.0563	0.0477	0.0477
7	0.1983	0.1945	0.1459	0.1368	0.1163	0.1282	0.1397
8	0.1249	0.2336	0.1988	0.1344	0.0567	0.0510	0.0686
9	0.1387	0.2160	0.3585	0.2174	0.0915	0.0524	0.0896
10	0.1020	0.1464	0.2212	0.1526	0.0739	0.0434	0.0577
11	0.1425	0.2264	0.3280	0.2260	0.1130	0.0667	0.0801
12	0.1011	0.1835	0.2102	0.1587	0.0481	0.0243	0.0572
13	0.1249	0.2040	0.2098	0.1821	0.0577	0.0200	0.0830
14	0.1135	0.1721	0.2102	0.1649	0.0653	0.0381	0.0682
15	0.1564	0.1955	0.2975	0.2417	0.0982	0.0610	0.1120
16	0.1549	0.2150	0.3852	0.2350	0.1554	0.0920	0.1096
17	0.1788	0.2245	0.3933	0.2584	0.1468	0.1125	0.1254
18	0.1669	0.1444	0.4329	0.1759	0.0768	0.0567	0.0791
Average	0.1508	0.1949	0.3107	0.2001	0.0883	0.0633	0.0891

abrupt color transitions in them. Many CDM methods use the Kodak dataset as the target images in algorithm development and testing, and they assume that the color differences change smoothly. Though this assumption holds well for the Kodak dataset, we can see from Table 1 that it may not hold for the images in the McMaster dataset. The cropped 18

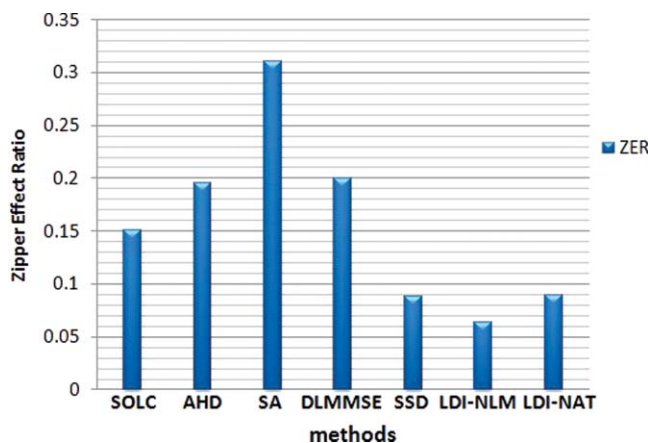


Fig. 7 Graphical presentation of the average ZER by different methods on the McMaster dataset.

sub-images and the source code of the proposed LDI-NLM and LDI-NAT algorithms can be downloaded at http://www.comp.polyu.edu.hk/~cslzhang/CDM_Dataset.htm.

3.2 Color Demosaicking Results

We evaluate the performance of various CDM schemes on the McMaster dataset. The proposed LDI- and NLM-based CDM method is denoted by LDI-NLM and the LDI- and NAT-based CDM method is denoted by LDI-NAT. The following representative CDM methods are used for comparison: the second order Laplacian correction (SOLC) method,^{3,4} the adaptive homogeneity-directed (AHD) method,¹⁰ the successive approximation (SA) method,⁹ the directional linear minimum mean square-error estimation (DLMMSE) method,¹² and the self-similarity driven (SSD) method.¹⁵ Considering that the SSD, LDI-NLM, and LDI-NAT algorithms all exploit the nonlocal redundancy, in Table 2 we summarize the procedures of the three algorithms for a better understanding of the common points and differences between them. These nonlocal methods involve a step of similar patch searching, which is one of the main sources of computational cost. Therefore, SSD, LDI-NLM, and LDI-NAT have higher complexity than the local methods SOLC, AHD, SA, and DLMMSE.

Suppose that the same nonlocal similar patch searching algorithm is used for SSD, LDI-NLM, and LDI-NAT, then

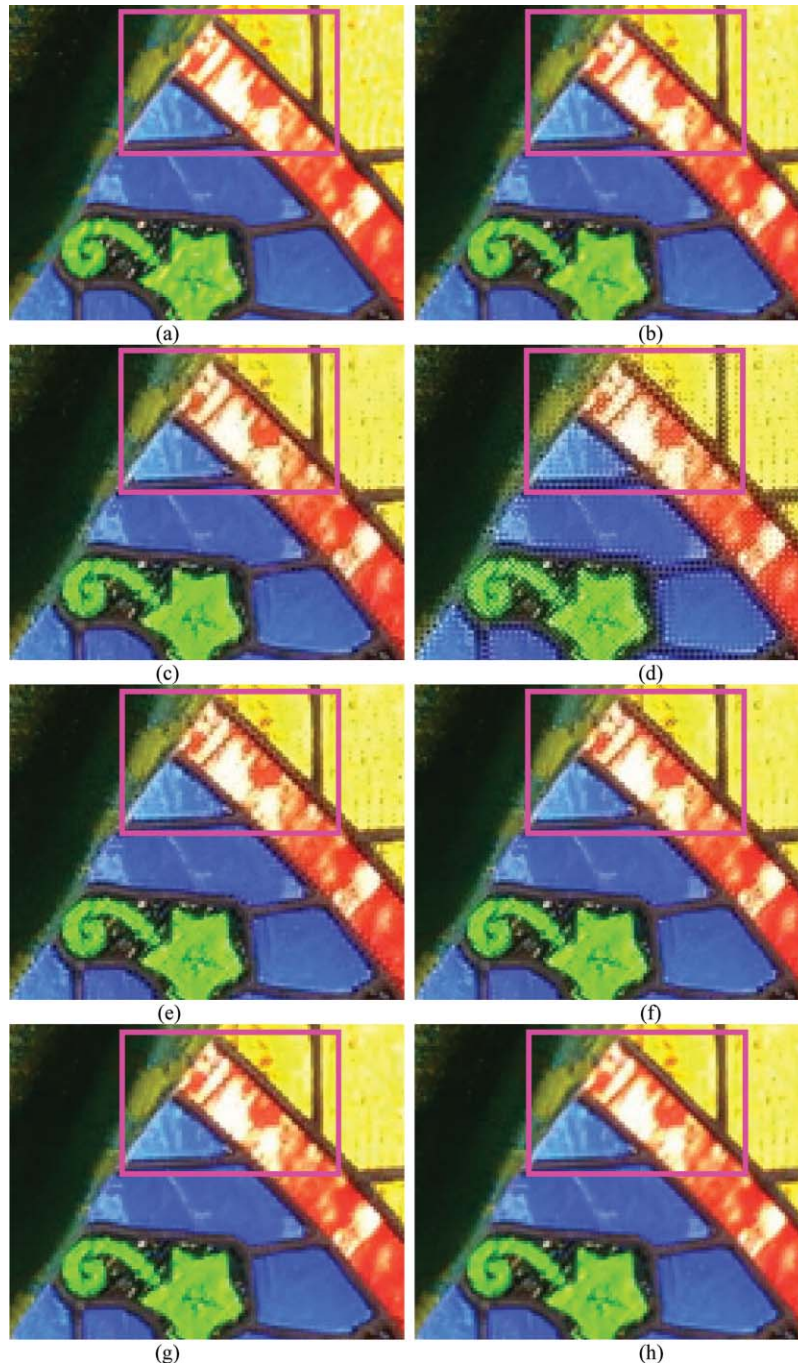


Fig. 8 (a) Original image 1 and demosaicked images by (b) SOLC (Ref. 3); (c) AHD (Ref. 10); (d) SA (Ref. 9); (e) DLMMSE (Ref. 12); (f) SSD (Ref. 15); (g) the proposed LDI-NLM and (h) LDI-NAT.

LDI-NLM has similar complexity to SSD because both of them use weighted average to exploit the nonlocal redundancy. However, LDI-NAT has higher complexity than SSD because it uses PCA to exploit the nonlocal redundancy. The data matrix Y formed by nonlocal similar patches is of size $M \times N$, and the covariance matrix of it is of size $M \times M$. In the PCA transformation, the SVD of the covariance matrix is required and the complexity is $O(M^3)$, which is much higher than that of the weighted average. Hence, the proposed LDI-NAT has the highest complexity among the

competing methods, while LDI-NLM and SSD have similar complexity.

In our implementation of LDI-NLM, 25 similar patches to the given patch (patch size: 5×5) are searched in a 31×31 local window. (Please note that based on our experiments, using more similar patches in NLM filtering will not improve the final CDM performance.) The parameter σ [refer to Eq. (9)] in the NLM filtering is set as 2.5. In our implementation of LDI-NAT, 100 similar patches to the given patch (patch size: 5×5) are searched in a 31×31 window. The threshold used

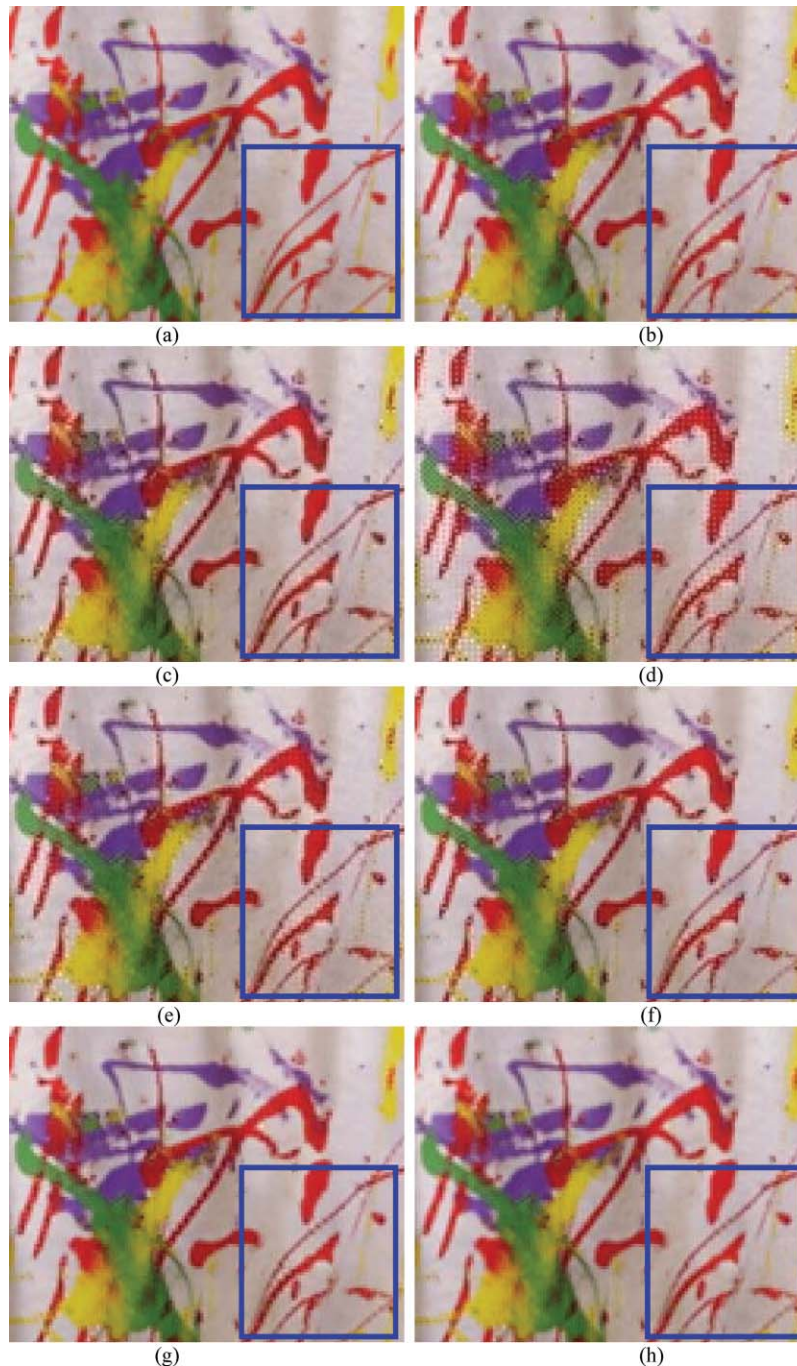


Fig. 9 (a) Original image 5 and demosaicked images by (b) SOLC (Ref. 3); (c) AHD (Ref. 10); (d) SA (Ref. 9); (e) DLMMSE (Ref. 12); (f) SSD (Ref. 15); (g) the proposed LDI-NLM and (h) LDI-NAT.

in Eq. (13) is set as $t = 0.03 \times g_Y$, where g_Y is the average gradient magnitude of the similar patches.

In the experiments, we down-sampled the original color images into CFA images according to the Bayer pattern and then reconstructed the full color images from the CFA mosaic data by using the seven methods. Table 3 lists the PSNR results. We see that the proposed LDI-NLM and LDI-NAT algorithms achieve much higher PSNR measures than other competing algorithms in almost every channel of all of the test images. Although the classical SOLC is simple, it achieves almost the same PSNR results as the recently de-

veloped SSD scheme, while SOLC and SSD outperform the other three methods in the competition. The DLMMSE has similar PSNR results to SOLC and SSD, and the AHD and SA algorithms have the lowest PSNR measures. Figure 6 graphically presents the average PSNR results by various methods on the McMaster dataset.

It is well known that PSNR is not a good indicator of CDM quality because the CDM errors mainly occur around the (color) edges, which account for only a small portion of the image pixels. In Ref. 15, the zipper effect ratio (ZER) was used to evaluate the color edge preservation

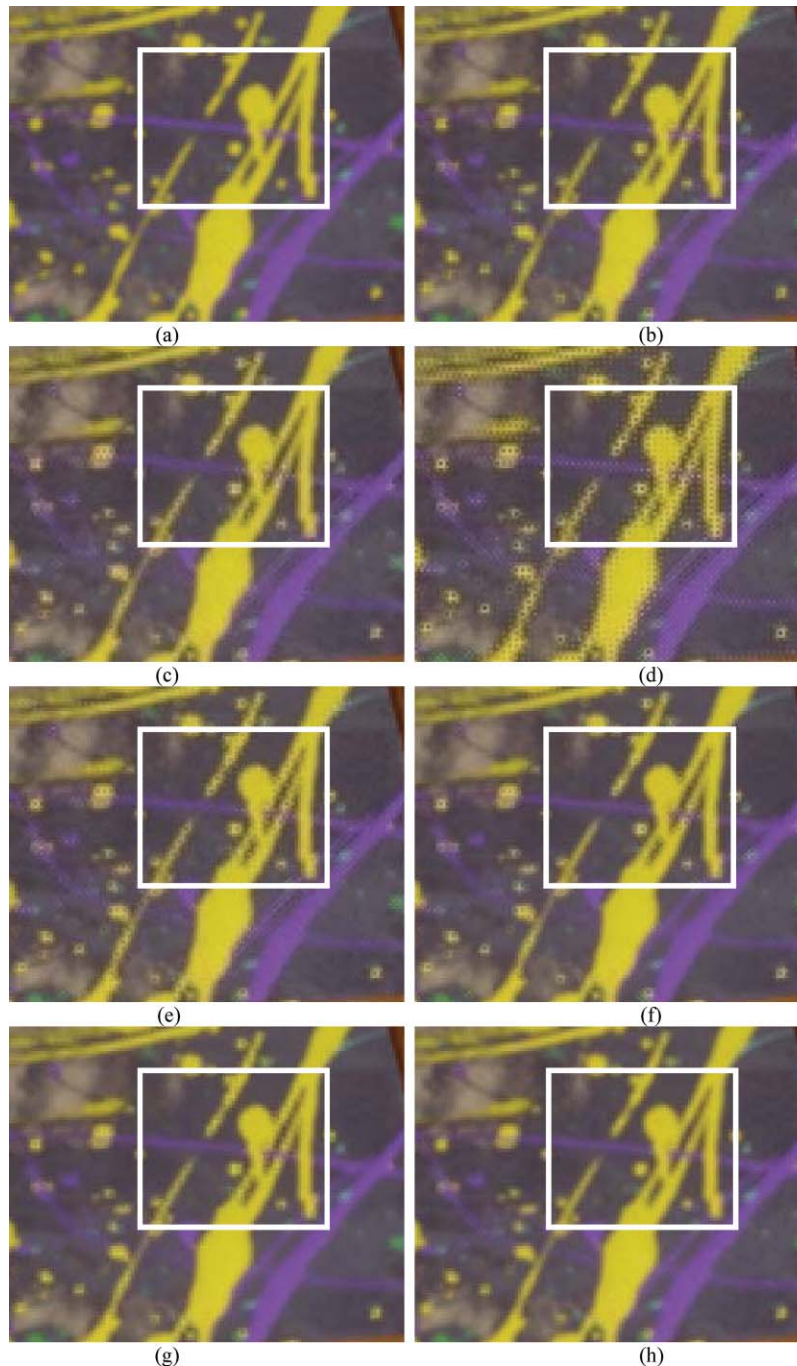


Fig. 10 (a) Original image 6 and demosaicked images by (b) SOLC (Ref. 3); (c) AHD (Ref. 10); (d) SA (Ref. 9); (e) DLMMSE (Ref. 12); (f) SSD (Ref. 15); (g) the proposed LDI-NLM and (h) LDI-NAT.

performance of CDM. Although this metric cannot perfectly reflect the CDM quality, it works better than PSNR in evaluating the CDM performance. Table 4 shows the ZER measures of the seven competing methods. Figure 7 graphically presents the average ZER results by various methods on the McMaster dataset. We see that LDI-NLM, SSD, and LDI-NAT achieve much lower ZER values than other methods. Although SOLC and DLMMSE have similar PSNR results to SSD, their ZER measures are much worse than SSD. This also validates that PSNR is not a good metric to measure image edge preservation. Note that LDI-NLM has lower ZER

values than LDI-NAT. However, LDI-NAT actually has much better edge preservation than LDI-NLM. This is because LDI-NLM results in smooth color edges, while the ZER metric favors smooth images. Nonetheless, how to define a good CDM quality metric is a very difficult problem, and this is beyond the discussion of this paper.

Figures 8–11 show the cropped and zoomed CDM results of the seven methods on images 1, 5, 6, and 16. It can be clearly seen that the proposed two algorithms, especially LDI-NAT, yield much better CDM outputs than the other five methods. The methods SA, AHD, and DLMMSE

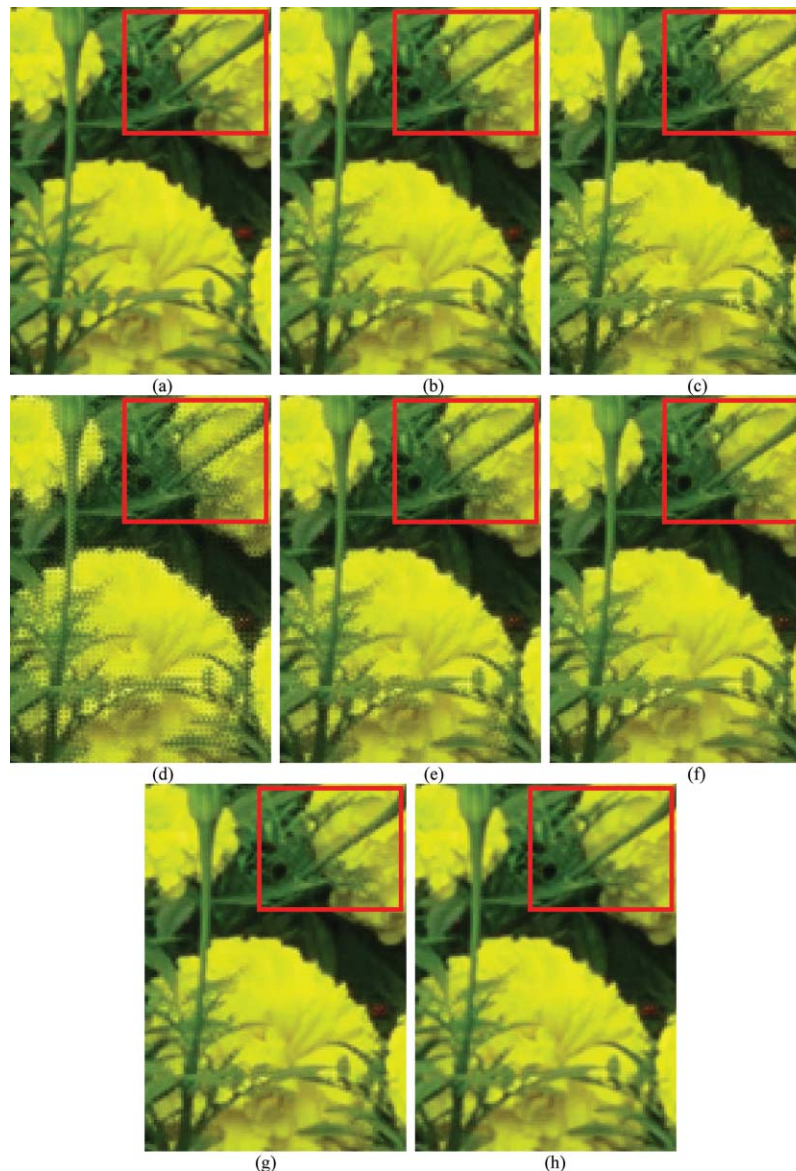


Fig. 11 (a) Original image 16 and demosaicked images by (b) SOLC (Ref. 3); (c) AHD (Ref. 10); (d) SA (Ref. 9); (e) DLMMSE (Ref. 12); (f) SSD (Ref. 15); (g) the proposed LDI-NLM and (h) LDI-NAT.

produce many zipper effects and false colors because they assume smooth color differences but this assumption does not hold well on the McMaster dataset. The SOLC uses a compact (five-tap) filter to interpolate the missing colors, which makes it free of many interpolation errors caused by abrupt color changes. However, SOLC neither fully exploits the local redundancy nor uses any nonlocal redundancy for CDM. The SSD exploits the nonlocal redundancy to iteratively recover the color information but it is not effective in using the image local directional information. As a result, both SOLC and SSD still produce many visible color artifacts, which can be clearly observed in Figs. 8–11. The proposed LDI-NLM and LDI-NAT effectively exploit the image local redundancy and edge direction information in the initial interpolation, and exploit the nonlocal similarity to enhance the CDM output. They significantly reduce the CDM errors and artifacts, recovering more faithfully the missing color samples than SOLC and SSD. Their

higher PSNR and lower ZER measures in Tables 3 and 4 also validate their powerful capability in color reproduction.

At last, we will compare the performance of LDI-NLM and LDI-NAT. As summarized in Table 2, LDI-NLM exploits the nonlocal redundancy by NLM filtering. NLM filtering is powerful in smoothing the initial CDM noise; however, it may also smooth the edges. In addition, for strong zippers and color artifacts caused by sharp color transition in highly saturated areas, NLM filtering is not effective to remove them. Different from NLM filtering, LDI-NAT processes the patch centered on the given pixel as a whole to better preserve the local pattern. The nonlocal redundancy is used to compute the sparse domain, and adaptive soft-thresholding is used to remove the initial CDM errors. Compared with NLM filtering, the NAT exploits the structural statistics in the similar patches, and hence, it can more effectively preserve the edges and reduce the zipper effects and false colors. From Table 3, we see that for images with more smooth areas

(e.g., images 8, 9, 13, and 18), the overall PSNR measures of LDI-NLM can be higher than LDI-NAT. Nonetheless, in the smooth areas, both the methods can have good visual quality. In areas with high color variations, LDI-NAT leads to much better CDM outputs. This can be more clearly seen in the images with more color edges, e.g., images 1, 5, 6, and 16.

4 Conclusion

This paper presented a CDM scheme by effectively exploiting both the local spectral correlation and the nonlocal similarity in the CFA image. Many previous CDM methods assume the high local spectral redundancy in the color interpolation. Such an assumption, however, fails for images with sharp color transitions and high color saturation. Fortunately, the nonlocal redundancy can be used to compensate for the lack of local redundancy in CDM. We first computed four directional local estimates of a missing color component and fused them into one estimate based on the local directional gradients. After the LDI, the nonlocal similar pixels to the given one were searched to enhance the initial CDM results. Apart from the commonly used NLM filtering technique, a NAT scheme was proposed to better preserve the local structure while reducing the initial CDM errors. The proposed LDI-NAT algorithm was tested on the McMaster dataset in comparison with state-of-the-art CDM methods. The experimental results showed that LDI-NAT visually leads to much better demosaicked images, significantly reducing the displeasing zipper effects and false colors that often appear in highly saturated areas.

Acknowledgments

This work is supported by the Hong Kong RGC General Research Fund (PolyU 5375/09E).

References

- B. E. Bayer and Eastman Kodak Company, "Color imaging array," U.S. Patent No. 3,971,065 (1975).
- P. Longère, X. Zhang, P. B. Delahunt, and D. H. Brainard, "Perceptual assessment of demosaicing algorithm performance," *Proc. IEEE* **90**(1), 123–132 (2002).
- J. E. Adams, "Intersections between color plane interpolation and other image processing functions in electronic photography," *Proc. SPIE* **2416**, 144–151 (1995).
- J. E. Adams and J. F. Hamilton Jr., "Adaptive color plane interpolation in single color electronic camera," U.S. Patent No. 5,506,619 (1996).
- R. Kimmel, "Demosaicing: image reconstruction from CCD samples," *IEEE Trans. Image Process.* **8**(9), 1221–1228 (1999).
- B. K. Gunturk, Y. Altunbasak, and R. M. Mersereau, "Color plane interpolation using alternating projections," *IEEE Trans. Image Processing* **11**(9), 997–1013 (2002).
- R. Lukac, K. Martin, and K. N. Plataniotis, "Demosaicked image post-processing using local color ratios," *IEEE Trans. Circuits Syst. Video Technol.* **14**(6), 914–920 (2004).
- D. D. Muresan and T. W. Parks, "Demosaicing using optimal recovery," *IEEE Trans. Image Process.* **14**(2), 267–278 (2005).
- X. Li, "Demosaicing by successive approximation," *IEEE Trans. Image Process.* **14**(3), 370–379 (2005).
- K. Hirakawa and T. W. Parks, "Adaptive homogeneity-directed demosaicing algorithm," *IEEE Trans. Image Process.* **14**(3), 360–369 (2005).
- D. Alleysson, S. Susstrunk, and J. Herault, "Linear demosaicing inspired by the human visual system," *IEEE Trans. Image Process.* **14**(4), 439–449 (2005).
- L. Zhang and X. Wu, "Color demosaicing via directional linear minimum mean square-error estimation," *IEEE Trans. Image Process.* **14**(12), 2167–2178 (2005).
- D. Menon, S. Andriani, and G. Calvagno, "Demosaicing with directional filtering and a posteriori decision," *IEEE Trans. Image Processing*, **16**(1), 132–141 (2007).
- J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. Image Process.* **17**(1), 53–69 (2009).
- A. Buades, B. Coll, J.-M. Morel, and C. Sbert, "Self-similarity driven color demosaicking," *IEEE Trans. Image Process.* **18**(6), 1192–1202 (2009).
- F. Zhang, X. Wu, X. Yang, W. Zhang, and L. Zhang "Robust color demosaicking with adaptation to varying spectral correlations," *IEEE Trans. Image Process.* **18**(12), 2706–2717 (2009).
- X. Wu and L. Zhang, "Color video demosaicking via motion estimation and data fusion," *IEEE Trans. Circuits Syst. Video Technol.* **16**, 231–240 (2006).
- X. Wu and L. Zhang, "Improvement of color video demosaicking in temporal domain," *IEEE Trans. Image Process.* **15**, 3138–3151 (2006).
- X. Li, B. Gunturk, and L. Zhang, "Image demosaicking: a systematic survey," Visual Communications and Image Processing 2008, *Proc. SPIE*, Vol. 6822, pp. 68221J-1–68221J-15, San Jose, California (2008).
- Kodak color image dataset, <http://r0k.us/graphics/kodak/>.
- A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.* **4**(2), 490–530 (2005).
- S. Kindermann, S. Osher, and P. W. Jones, "Deblurring and denoising of images by nonlocal functionals," *Multiscale Model. Simul.* **4**(4), 1091–1115 (2005).
- K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform domain collaborative filtering," *IEEE Trans. Image Process.* **16**(8), 2080–2094 (2007).
- T. Brox, O. Kleinschmidt, and D. Cremers, "Efficient nonlocal means for denoising of textural patterns," *IEEE Trans. Image Process.* **17**(7), 1083–1092 (2008).
- A. Buades, B. Coll, and J. M. Morel, "Nonlocal image and movie denoising," *Int. J. Comput. Vision* **76**(2), 123–139 (2008).
- X. Li and Y. Zheng, "Patch-based video processing: a variational Bayesian approach," *IEEE Trans. Circuits Syst. Video Technol.* **19**(1), 27–40 (2009).
- L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D* **60**, 259–268 (1992).
- P. Blomgren and T. Chan, "Color TV: total variation methods for restoration of vector-valued images," *IEEE Trans. Image Process.* **7**(3), 304–309 (1998).
- T. Chan, S. Esedoglu, F. Park, and A. Yip, "Recent developments in total variation image restoration," in *Mathematical Models of Computer Vision*, N. Paragios, Y. Chen, and O. Faugeras, Eds., Springer-Verlag, New York (2005).
- A. Chambolle, R. A. DeVore, N.-Y. Lee, and B. J. Lucier, "Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage," *IEEE Trans. Image Process.* **7**(3), 319–335 (1998).
- S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.* **43**(1), 129–159 (2001).
- D. Donoho and M. Elad, "Optimal sparse representation in general (nonorthogonal) dictionaries via L_1 minimization," *Proc. Natl. Acad. Sci. U.S.A.* **100**, 2197–2202 (2003).
- M. Aharon, M. Elad, and A. M. Bruckstein, "The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.* **54**(11), 4311–4322 (2006).
- R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proc. IEEE* **98**(6), 1045–1057 (2010).
- M. Protter and M. Elad, "Image sequence denoising via sparse and redundant representations," *IEEE Trans. Image Process.* **18**(1), 27–36 (2009).
- J. M. B. Dias and M. A. T. Figueiredo, "A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Trans. Image Process.* **16**(12), 2992–3004 (2007).
- S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press (2004).
- D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory* **52**(4), 1289–1306 (2006).
- E. Candès and T. Tao, "Near optimal signal recovery from random projections: universal encoding strategies?," *IEEE Trans. Inf. Theory* **52**(12), 5406–5425 (2006).
- E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory* **52**(2), 489–509 (2006).
- D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory* **41**(3), 613–627 (1995).



Lei Zhang received his BS degree in 1995 from Shenyang Institute of Aeronautical Engineering, Shenyang, China, his MS and PhD degrees in automatic control theory and engineering from Northwestern Polytechnical University, Xi'an, China, respectively, in 1998 and 2001. From 2001 to 2002, he was a research associate in the Department of Computing, The Hong Kong Polytechnic University. From January 2003 to January 2006 he worked as a postdoctoral fellow in the Department of Electrical and Computer Engineering, McMaster University, Canada. In 2006, he joined the Department of Computing, The Hong Kong Polytechnic University, as an assistant professor. Since September 2010, he has been an associate professor in the same department. His research interests include image and video processing, biometrics, computer vision, pattern recognition, multi-sensor data fusion and optimal estimation theory, etc. Zhang is an associate editor of *IEEE Trans. on SMC-C*, the leading guest editor of *EURASIP Journal on Image and Video Processing*, Special Issue on Emerging Methods for Color Image and Video Quality Enhancement, and the guest editor of *Sensors*, Special Issue on Hand-based Biometrics Sensors and System. Zhang was awarded the *Merit Award*, *Faculty Award in Research and Scholarly Activities 2010*, the *Best Paper Award of SPIE VCIP2010*. More information can be found in his homepage <http://www4.comp.polyu.edu.hk/~cslzhang/>.



Antoni Buades received his degree in mathematics from the University of Balearic Islands in 2001 and his MS degree in mathematical vision from the ENS Cachan, Paris. He obtained the PhD degree in mathematics from the Universitat Illes Balears in 2006. He is a researcher at CNRS Paris Descartes and Universitat Illes Balears focusing on mathematical analysis of digital images, image restoration, and computer vision.



Xin Li received his BS degree with highest honors in electronic engineering and information science from University of Science and Technology of China, Hefei, in 1996, and his PhD degree in electrical engineering from Princeton University, Princeton, New Jersey, in 2000. He was a member of technical staff with Sharp Laboratories of America, Camas, Washington from August 2000 to December 2002. Since January 2003, he has been a faculty member in Lane Department of Computer Science and Electrical Engineering. His research interests include image/video coding and processing. Li received a Best Student Paper Award at the Conference of Visual Communications and Image Processing as the junior author in 2001, a runner-up prize of Best Student Paper Award at IEEE Asilomar conf. on Signals, Systems and Computers as the senior author in 2006, and a Best Paper Award at the Conference of Visual Communications and Image Processing as the single author in 2010. Li is currently serving as an associate editor for *IEEE Transactions on Image Processing* and is a member of IEEE Technical Committee on Image, Video and Multidimensional Signal Processing (IVMSP).



Xiaolin Wu received his BSc degree from Wuhan University, China, in 1982, and his PhD degree from the University of Calgary, Calgary, Ontario, Canada, in 1988, both in computer science. He started his academic career in 1988 and has since been on the faculty at the University of Western Ontario, New York Polytechnic University, Brooklyn, and currently McMaster University, Hamilton, Ontario, where he is a professor in the Department of Electrical and Computer Engineering, and where he holds the NSERC-DALSA Industrial Research Chair in Digital Cinema. His research interests include image processing, multimedia compression, joint source-channel coding, multiple description coding, and network-aware visual communication. He has published over 180 research papers and holds two patents in these fields. He currently serves as an associate editor for the *IEEE Transactions on Image Processing*.