

Journal of Electronic Imaging

JElectronicImaging.org

Demosaicking based on channel-correlation adaptive dictionary learning

Chenyan Bai
Jia Li
Zhouchen Lin



Chenyan Bai, Jia Li, Zhouchen Lin, "Demosaicking based on channel-correlation adaptive dictionary learning," *J. Electron. Imaging* **27**(4), 043047 (2018),
doi: 10.1117/1.JEI.27.4.043047.

Demosaicking based on channel-correlation adaptive dictionary learning

Chenyan Bai,^{a,*} Jia Li,^b and Zhouchen Lin^{b,c}

^aCapital Normal University, College of Information Engineering, Beijing, China

^bPeking University, Key Laboratory of Machine Perception (MOE), School of Electronic Engineering and Computer Science, Beijing, China

^cShanghai Jiao Tong University, Cooperative Medianet Innovation Center, Shanghai, China

Abstract. Image demosaicking is the problem of reconstructing color images from raw images captured using a digital camera with a color filter array. Sparse representation-based demosaicking method achieves superior performance on the commonly used Kodak dataset; however, it performs inferior on the IMAX dataset. We analyze that the factor of the sparse representation-based demosaicking methods that perform inconsistently is channel-correlation, which we define as the mean value of correlation coefficients between the RGB channels. Accordingly, we propose a channel-correlation adaptive dictionary learning-based demosaicking method. Different from the sparse representation-based demosaicking methods that use a fixed dictionary, our method trains a general dictionary on training image patches with various channel-correlations. Then, we learn a function matrix between the general dictionary and channel-correlations. For a raw image patch with an estimated channel-correlation, we compute a dictionary corresponding to its channel-correlation through the function matrix. Finally, we demosaick it with the corresponding dictionary using the sparse representation model. Experiments confirm that the proposed method performs adaptively well on raw images with various channel-correlations. © 2018 SPIE and IS&T [DOI: [10.1117/1.JEI.27.4.043047](https://doi.org/10.1117/1.JEI.27.4.043047)]

Keywords: demosaicking; adaptive; dictionary learning; channel-correlation.

Paper 180216 received Mar. 12, 2018; accepted for publication Aug. 6, 2018; published online Aug. 20, 2018.

1 Introduction

In color digital imaging, a color image is reconstructed from a raw image captured by a digital camera that has a color filter array (CFA)^{1,2} (Fig. 1). The reconstruction process is called demosaicking. Many demosaicking methods have been developed to improve the quality of the demosaicked images,^{3–7} such as the interpolation-based methods,^{5,8–13} the frequency domain methods,^{14,15} the sparse representation-based methods,^{4,16,17} and the deep learning-based methods.^{18,19}

Among these methods, sparse representation-based demosaicking methods^{16,17} achieve excellent performance. They also can be easily adapted to handle non-Bayer CFAs, e.g., those CFAs with panchromatic pixels or pixels with optimized colors.⁴ However, most of the existing methods use only a fixed dictionary, making the demosaicked results not always satisfactory, especially when the training and test datasets are in different domains. For example, the learned simultaneous sparse coding (LSSC) method¹⁷ obtains high CPSNR on the commonly used Kodak dataset.²⁰ However, it performs inferior on the IMAX dataset,²¹ which is another popular dataset for demosaicking. Therefore, developing demosaicking algorithms that work consistently on raw images with different characteristics is an important issue.

To solve the above problem, adaptive dictionary learning-based demosaicking methods have been proposed. For example, Zhang and Tao²² trained dictionaries on different patch classes determined by energy exclusiveness feature and adaptively selected the most suitable dictionary for a raw image patch. Wu et al.²³ first trained a regressor on

directional difference, then they demosaicked raw images by properly integrating the interpolation algorithm MLRI¹² and the trained regressor.

In this paper, we propose an adaptive dictionary learning-based demosaicking method that can demosaick adaptively to channel-correlation. The exploit of the spatial-spectral correlations among the RGB channels has been well studied.^{5,8–13,23,24} In particular, Duran and Buades²⁴ proposed an algorithm that introduced a clear manner of balancing how much spatial-spectral correlations must be taken advantage of. We also explicitly explore the correlations among the RGB channels. We first define a measure of the correlation, which is called channel-correlation. Then, we analyze how the channel-correlation affects the performance of sparse representation-based demosaicking^{16,17} (see the analysis in Sec. 2). Finally, we propose a channel-correlation adaptive dictionary learning-based demosaicking method (CADLD) that can demosaick adaptively to the channel correlation.

The CADLD performs superiorly on raw images with various channel-correlations (see the flowchart in Fig. 4). We use training image patches with various channel-correlations to train a general dictionary. Then, we learn a function matrix between the general dictionary and channel-correlations. For a raw image patch, we estimate its channel correlation and compute a specific dictionary through the function matrix. Finally, we demosaick it with the specific dictionary.

The contributions of this paper are as follows:

- We define the channel correlation as a measure of the correlations among the RGB channels and show that it is an important factor to affect the performance of

*Address all correspondence to: Chenyan Bai, E-mail: cymbai@cnu.edu.cn

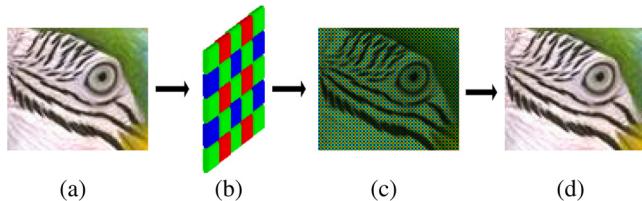


Fig. 1 The process of image demosaicking. (a) A full color image to simulate a real scene, (b) Bayer CFA, (c) Bayer raw image, and (d) Bayer demosaicked image. Images in this paper are best viewed on screen.

the existing sparse representation-based demosaicking methods using fixed dictionaries.

- We propose the CADLD method that can demosaick adaptively to raw images with various channel correlations, which reduces the impact of channel correlation on the sparse representation-based demosaicking.

The rest of the paper is organized as follows: in Sec. 2, we define the channel correlation and analyze its effect on the sparse representation-based demosaicking. Then, we introduce our CADLD method in Sec. 3. In Sec. 4, we conduct experiments to verify our method. Finally, we conclude the paper in Sec. 5.

2 Analysis of Channel-Correlation Effect on Sparse Representation-Based Demosaicking

To study how the correlations among the RGB channels affect the sparse representation-based demosaicking,^{4,16,17} we first define the channel-correlation as a measure of the correlations. Then, we analyze its effect on the sparse representation-based demosaicking.^{4,16,17} In this paper, we use upper- and lower-case bold letters to denote matrices and vectors, respectively. Both images and patches are represented in a column vector in the order of R , G , and B . The main notations used in the paper are shown.

2.1 Channel-Correlation

We define the channel-correlation as the usual correlation coefficient between the RGB channels. Let $\mathbf{x} \in \mathbb{R}^{3m \times 1}$ be a color image patch, with $\mathbf{x}_R \in \mathbb{R}^{m \times 1}$, $\mathbf{x}_G \in \mathbb{R}^{m \times 1}$, and $\mathbf{x}_B \in \mathbb{R}^{m \times 1}$ being its red, green, and blue channels,

respectively, where m is the pixel number in each channel of \mathbf{x} . The channel correlation between \mathbf{x}_R and \mathbf{x}_G is defined as

$$c_{RG} = \frac{\text{Cov}(\mathbf{x}_R, \mathbf{x}_G)}{\sigma_{\mathbf{x}_R} \sigma_{\mathbf{x}_G}}, \quad (1)$$

where $\text{Cov}(\mathbf{x}_R, \mathbf{x}_G)$ is the covariance between \mathbf{x}_R and \mathbf{x}_G , and $\sigma_{\mathbf{x}_R}$ and $\sigma_{\mathbf{x}_G}$ are the standard deviations of \mathbf{x}_R and \mathbf{x}_G , respectively. Similarly, we can get the channel-correlation c_{RB} and c_{GB} . Finally, we define the channel-correlation of patch \mathbf{x} as

$$c = \frac{c_{RG} + c_{RB} + c_{GB}}{3}. \quad (2)$$

With the definition of patch channel-correlation, we can statistic the channel-correlation distribution of a dataset. For example, we compute the channel-correlations of the Kodak²⁰ and IMAX²¹ datasets using Eq. (2). We use the sliding window approach to sample patches. The patch size is $8 \times 8 \times 3$, and the step size is $4 \times 4 \times 1$, i.e., the patches are sampled in every four pixels for both vertical and horizontal directions. We compute the patches' channel-correlations and get their statistics histograms, which are shown in Figs. 2(a) and 2(b). In the same way, we can get that of the PASCAL VOC'07²⁵ dataset, which is shown in Fig. 2(c).

2.2 Channel-Correlation Effect on Sparse Representation-Based Demosaicking

We analyze the channel-correlation effect on the sparse representation-based demosaicking methods.¹⁷ We randomly sample patches from the PASCAL VOC'07²⁵ and DDR training datasets.²³ We test on three patch groups with channel-correlations in 0.25 ± 0.05 , 0.5 ± 0.05 , and 0.75 ± 0.05 , respectively. For each patch group, we obtain 2×10^6 patches with a size of $8 \times 8 \times 3$. Then, we randomly split them into 1.5×10^6 training and 0.5×10^6 test patches. Some training patches of the three groups are shown in the first row of Fig. 3.

We use the SPAMS toolbox^{17,26} to learn three dictionaries on the three different channel-correlation training datasets. All the dictionaries are of the same size 192×256 , which are denoted as $\mathbf{D}_{0.25}$, $\mathbf{D}_{0.5}$, and $\mathbf{D}_{0.75}$. The dictionaries are shown in the second row of Fig. 3. We choose PSRD,⁴ which is a sparse representation-based demosaicking method,

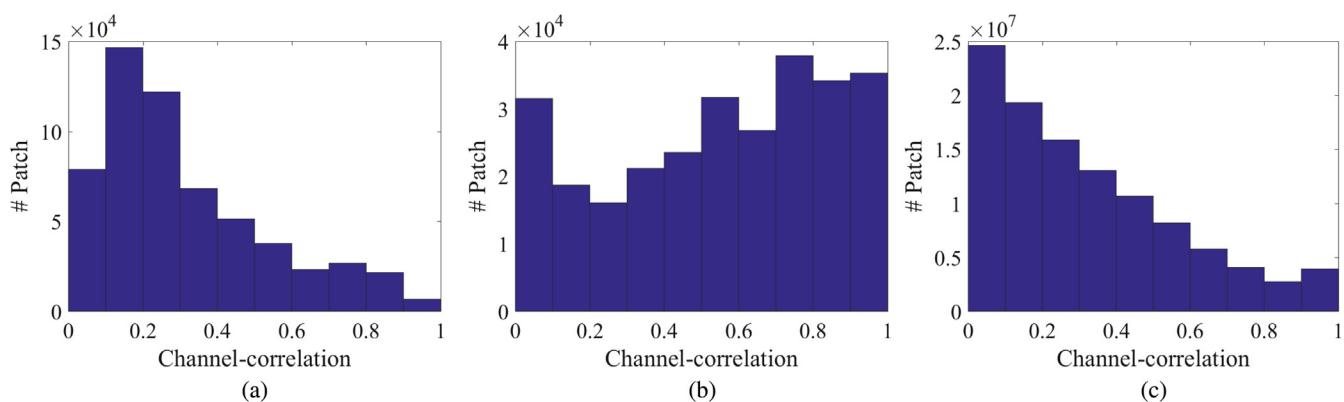


Fig. 2 The patch channel-correlation distributions of the (a) Kodak,²⁰ (b) IMAX,²¹ and (c) PASCAL VOC'07²⁵ datasets.

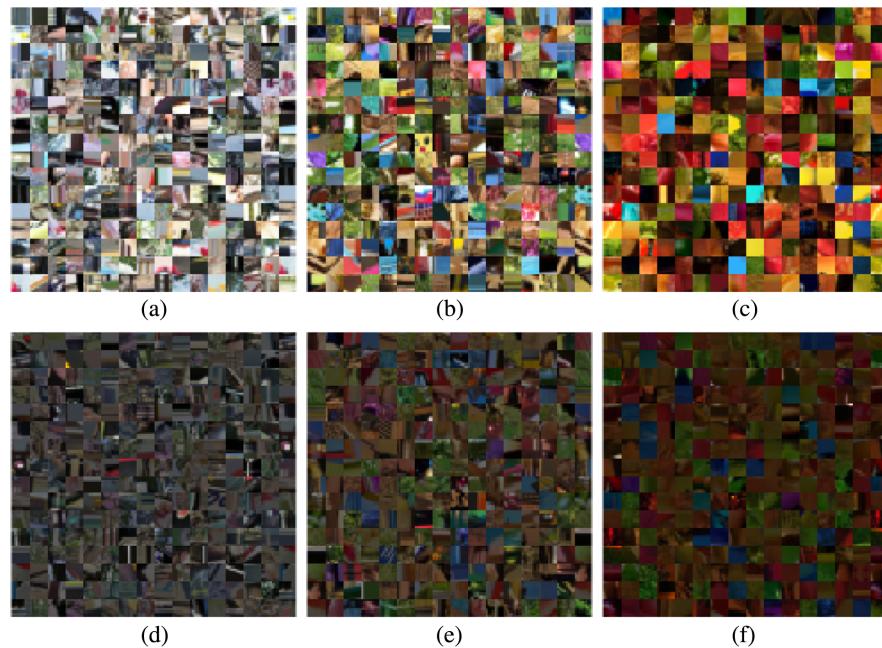


Fig. 3 Illustration of the training image patches and the corresponding dictionaries with different channel-correlations. (a)–(c) Training patches with channel-correlations 0.25, 0.5, and 0.75, respectively. (d)–(f) Dictionaries learned on the corresponding training patches.

to demosaick the three test patch groups with dictionaries $\mathbf{D}_{0.25}$, $\mathbf{D}_{0.5}$, and $\mathbf{D}_{0.75}$, respectively.

For each test patch group, we record which dictionary achieves the highest CPSNR. Then, we count the number of the highest CPSNR for each dictionary, which is denoted by $N_h(\mathbf{D}_j)$, $j \in \{0.25, 0.5, 0.75\}$. For a test patch group with channel-correlation c , we compute the following ratio matrix \mathbf{R}

$$\mathbf{R}(c, j) = \frac{N_h(\mathbf{D}_j)}{N}, \quad (3)$$

where $N_h(\mathbf{D}_j)$ means the number of the highest CPSNR for dictionary \mathbf{D}_j , and N is the total number of the test patches.

The ratio matrix \mathbf{R} is shown in Table 1. We can see that the test patches achieve the most highest CPSNR (numbers in bold) when demosaicked by the dictionaries learned on training patches with an identical channel-correlation. For example, the test patch group with channel-correlation

Table 1 Channel-correlation effect on the sparse representation based demosaicking. We choose test patches with three channel-correlations 0.25, 0.5, and 0.75. Then we demosaick them by three dictionaries $\mathbf{D}_{0.25}$, $\mathbf{D}_{0.5}$, and $\mathbf{D}_{0.75}$ learned on three training patches with channel-correlations 0.25, 0.5, and 0.75, respectively. The table shows the ratio matrix \mathbf{R} computed by Eq. (3). It gets the most highest CPSNR (numbers in bold) when the test and the training patches are within the same channel-correlation.

Channel-correlation	$\mathbf{D}_{0.25}$ (%)	$\mathbf{D}_{0.5}$ (%)	$\mathbf{D}_{0.75}$ (%)
Test _{0.25}	65.50	28.36	6.14
Test _{0.5}	21.87	55.66	22.46
Test _{0.75}	2.42	37.24	60.34

0.25 has 65.50% patches that get the highest CPSNR with the dictionary learned on the patches with channel-correlation 0.25. In addition, we can find that the larger the difference in channel-correlation between the training and test patches, the less percentage that get the highest CPSNR. This motivates us to develop a channel-correlation adaptive demosaicking method.

2.3 Analysis and Discussions

Specifically, we analyze the channel-correlation effect on LSSC,¹⁷ which achieves good performance among the sparse representation-based demosaicking methods. The histograms of the Kodak²⁰ and IMAX datasets²¹ are shown in Figs. 2(a) and 2(b), respectively. We can see that the channel-correlations of the Kodak dataset²⁰ are mainly in the range (0 and 0.4), which is more similar with that of the PASCAL VOC'07 dataset [Fig. 2(c)], whereas that of the IMAX dataset²¹ is far from it. So, we can get an intuitive explanation of why LSSC¹⁷ with dictionary learned on the PASCAL VOC'07 dataset achieves excellent performance on the Kodak dataset, but it performs not so satisfactory on the IMAX dataset. This inspires us to propose the CADLD, which is described in the following section.

3 Channel-Correlation Adaptive Dictionary Learning-Based Demosaicking

Our CADLD is based on the sparse representation-based demosaicking^{16,17} and domain adaptive dictionary learning.²⁷ It can demosaick adaptively to the channel-correlations. The flowchart is shown in Fig. 4. The CADLD has a training phase and a demosaicking phase. In the training phase, we first learn a general dictionary on the training dataset, which includes image patches with various channel-correlations. Then, we reconstruct a mapping function between the channel-correlations and the general dictionary. In the demosaicking phase, for a raw patch, we first estimate its

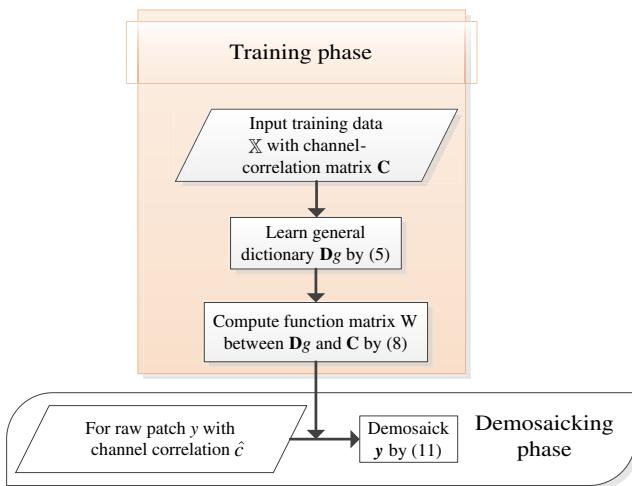


Fig. 4 The flowchart of the proposed CADLD. In the training phase, we learn a general dictionary \mathbf{D}_g on training data with different channel-correlations \mathbf{C} by Eq. (5). Then, we compute the function matrix \mathbf{W} between \mathbf{D}_g and \mathbf{C} by Eq. (8). In the demosaicking phase, we demosaick the raw patch \mathbf{y} according to its channel-correlation by Eq. (11).

channel-correlation. Then, we compute a specific dictionary through the channel-correlation and the mapping function. The dictionary is corresponding to the raw patch's channel-correlation. Finally, we demosaick the raw patch with the specific dictionary using sparse representation-based model. We give the details below.

3.1 Training Phase

As the sparse representation-based demosaicking^{16,17} cannot work well when the training and test datasets are with different channel-correlations, we develop a channel-correlation adaptive demosaicking method based on dictionary learning.

3.1.1 Learning a general dictionary

To deal with the situation when the training and test data are in different domains, Qiu et al.²⁷ proposed a domain adaptive dictionary learning method for the face recognition problem. It first learns a general dictionary on training face data with different viewpoints. Then for each to be recognized test face data, it estimates the viewpoint of the test data and computes a specific dictionary for it. Finally, the test face data are processed with the specific dictionary.

Inspired by Qiu et al.,²⁷ we propose the CADLD method. We denote $\mathbf{x} = (\mathbf{x}_R^T, \mathbf{x}_G^T, \mathbf{x}_B^T)^T \in \mathbb{R}^{3m \times 1}$ as a color image patch, $\mathbf{X}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{in}) \in \mathbb{R}^{3m \times n}$ is a training dataset, which consists of n patches with a specific channel-correlation. Let $\{\mathbf{X}_i\}_{i=1}^s$ be the training datasets with s different channel-correlations, we aim to learn dictionaries on these training datasets, which are with the same sparse coefficient matrix $\mathbf{A} \in \mathbb{R}^{k \times n}$. The model is as follows:

$$\begin{aligned} & \arg \min_{\mathbf{D}_i, \mathbf{A}} \sum_{i=1}^s \|\mathbf{X}_i - \mathbf{D}_i \mathbf{A}\|_F^2 + \lambda \sum_{j=1}^n \|\mathbf{A}(:, j)\|_1 \\ & \text{s.t. } \forall i \|\mathbf{D}_i(:, l)\|_1 = 1, \quad l = 1, 2, \dots, k, \end{aligned} \quad (4)$$

where $\mathbf{D}_i \in \mathbb{R}^{3m \times k}$ ($k > 3m$) is the dictionary learned on the training dataset \mathbf{X}_i , λ is a parameter that balances the sparsity of $\mathbf{A}(:, j)$ and fidelity of the approximation to \mathbf{X}_i .

Let $\mathbf{X} = (\mathbf{X}_1^T, \mathbf{X}_2^T, \dots, \mathbf{X}_s^T)^T \in \mathbb{R}^{3ms \times n}$ be the whole training dataset with s different channel-correlations, through mathematic deduction and omit the constraints, Eq. (4) can be formulated as

$$\arg \min_{\mathbf{D}_g, \mathbf{A}} \|\mathbf{X} - \mathbf{D}_g \mathbf{A}\|_F^2 + \lambda \sum_{j=1}^n \|\mathbf{A}(:, j)\|_1, \quad (5)$$

where $\mathbf{D}_g = (\mathbf{D}_1^T, \mathbf{D}_2^T, \dots, \mathbf{D}_s^T)^T$ is the general dictionary. We use SPAMS toolbox^{17,26} to learn the general dictionary. We project \mathbf{D}_i to satisfy the constraints in Eq. (4) for each iteration when updating the dictionary.

3.1.2 Learning a function between dictionary and channel-correlation

Following Ref. 27, we model the dictionary as a p degree polynomial function of the channel-correlation. Let $\mathbf{d}_i = \text{vec}(\mathbf{D}_i) \in \mathbb{R}^{3mk \times 1}$, we have

$$\mathbf{d}_i = \mathbf{W} \mathbf{c}_i, \quad (6)$$

where $\mathbf{W} \in \mathbb{R}^{3mk \times (p+1)}$ is the coefficient matrix of the polynomial function, and $\mathbf{c}_i = (1, c_i, c_i^2, \dots, c_i^p)^T$. The parameter p is empirically set as 2 in all experiments.

We define \mathbf{D}_g^{VT} as

$$\begin{aligned} \mathbf{D}_g^{VT} &= [\text{vec}(\mathbf{D}_1), \text{vec}(\mathbf{D}_2), \dots, \text{vec}(\mathbf{D}_s)] \\ &= (\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_s) \\ &= (\mathbf{W} \mathbf{c}_1, \mathbf{W} \mathbf{c}_2, \dots, \mathbf{W} \mathbf{c}_s) \\ &= \mathbf{W} \mathbf{C}, \end{aligned} \quad (7)$$

where $\mathbf{C} = (\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_s)$.

We can compute the function parameter matrix \mathbf{W} between the dictionary and channel-correlation by Eq. (7), that is

$$\mathbf{W} = \mathbf{D}_g^{VT} \mathbf{C}^\dagger. \quad (8)$$

where \mathbf{C}^\dagger is the Pseudo-inverse of matrix \mathbf{C} .

3.2 Channel-Correlation Adaptive Demosaicking

In the noiseless case, the model of color imaging with a CFA is as follows:

$$\mathbf{y} = \mathbf{Mx} = (\mathbf{M}_R, \mathbf{M}_G, \mathbf{M}_B)(\mathbf{x}_R^T, \mathbf{x}_G^T, \mathbf{x}_B^T)^T, \quad (9)$$

where $\mathbf{y} \in \mathbb{R}^{m \times 1}$ is a raw image patch, $\mathbf{x} \in \mathbb{R}^{3m \times 1}$ is a color image patch, with \mathbf{x}_R , \mathbf{x}_G , and \mathbf{x}_B being its red, green, and blue channels, respectively, $\mathbf{M} \in \mathbb{R}^{m \times 3m}$ is the mosaicking matrix, and \mathbf{M}_R , \mathbf{M}_G , and \mathbf{M}_B are the diagonal matrices whose diagonal elements are specified by the R , G , and B channels of the CFA, respectively. Formally, let the CFA be $\mathbf{F} \in \mathbb{R}^{\sqrt{m} \times \sqrt{m} \times 3}$, then $\mathbf{M}_R = \text{Diag}\{\text{vec}[\mathbf{F}(:, :, 1)]\}$, $\mathbf{M}_G = \text{Diag}\{\text{vec}[\mathbf{F}(:, :, 2)]\}$, and $\mathbf{M}_B = \text{Diag}\{\text{vec}[\mathbf{F}(:, :, 3)]\}$.

For a raw image patch \mathbf{y} , we use the bilinear interpolation to demosaick \mathbf{y} . Then, we estimate its channel-correlation \hat{c} by the demosaicked color patch. We next compute a specific dictionary corresponding to \mathbf{y}

$$\hat{\mathbf{D}} = \text{mat}(\mathbf{W} \hat{\mathbf{c}}), \quad (10)$$

Table 2 Evaluation of the proposed method on the IMAX dataset.²¹ The individual and average CPSNR values are reported. “Avg.” stands for “average.”

Image ID	SC ¹⁶	SSC ¹⁷	LSC ¹⁷	LSSC ¹⁷	PSRD ⁴	PAMD ²²	DDR ²³	CADLD
01	26.32	26.47	27.58	27.77	26.38	26.31	29.50	29.21
02	33.25	33.51	33.68	34.09	33.22	33.07	35.33	34.87
03	32.43	32.53	32.79	32.89	32.47	32.42	34.41	33.39
04	34.51	34.64	36.18	36.62	34.61	35.25	38.84	37.76
05	30.68	30.97	31.47	31.98	30.76	30.54	34.73	34.04
06	33.10	34.17	35.13	37.47	33.22	33.91	38.94	39.48
07	39.30	39.83	39.04	39.61	39.09	40.16	37.28	39.68
08	37.80	38.28	38.09	38.79	37.80	37.31	39.26	38.51
09	34.97	35.42	35.81	36.46	34.97	34.36	37.54	37.53
10	36.42	36.96	36.98	37.92	36.33	36.24	39.27	38.79
11	37.20	38.01	37.93	39.21	37.07	37.24	40.14	39.87
12	36.58	36.83	37.13	37.68	36.49	36.08	39.62	39.30
13	38.60	39.03	38.99	39.72	38.41	38.38	41.02	40.99
14	36.74	37.28	37.29	38.00	36.67	36.60	39.06	39.10
15	37.23	37.74	37.59	38.43	37.13	36.94	39.31	39.42
16	29.48	29.69	31.12	31.63	29.60	29.53	34.55	34.88
17	29.46	29.83	31.04	31.84	29.53	29.61	33.70	34.32
18	33.56	33.65	34.02	34.60	33.56	33.46	35.87	35.82
Avg.	34.31	34.71	35.10	35.82	34.29	34.30	37.13	37.05

Note: The highest CPSNR of each row is shown in bold.

where \mathbf{W} is the function parameter matrix between the dictionary and channel-correlation, which is computed by Eq. (8), $\hat{\mathbf{c}} = (1, \hat{c}, \hat{c}^2, \dots, \hat{c}^p)^T$ is the channel-correlation vector, and $\text{mat}(\mathbf{W}\hat{\mathbf{c}})$ is a matrix formed by vector $\mathbf{W}\hat{\mathbf{c}}$, whose size is $3m \times k$.

We use the following sparse representation-based demosaicking model to demosaick \mathbf{y}

$$\min_{\alpha} \frac{1}{2} \|\mathbf{y} - \mathbf{M}\hat{\mathbf{D}}\alpha\|_2^2 + \gamma\|\alpha\|_1, \quad (11)$$

where \mathbf{M} is the mosaicking matrix, $\hat{\mathbf{D}}$ is the adaptive dictionary, α is the sparse coefficient, and γ is a parameter that balances the sparsity of α and the fidelity of the approximation to \mathbf{y} .

The optimization Eq. (11) is solved by the alternating direction method,²⁸ which has been widely used in image processing.

4 Experiments

In this section, we verify our CADLD on the commonly used IMAX²¹ and Kodak²⁰ datasets. We first detail the experimental settings. Then, we choose the latest sparse representation-based demosaicking methods and compare the proposed method with them.

4.1 Experimental Settings

In the experiments, we sample training patches and learn a function parameter between the dictionary and channel-correlation. Then for an arbitrary test patch, we estimate its channel-correlation and demosaick it adaptively using CADLD.

We randomly extract patches from two datasets: the PASCAL VOC’07 dataset²⁵ and the DDR training dataset.²³ We sample patches whose channel-correlations are in 0.2 ± 0.05 , 0.4 ± 0.05 , 0.6 ± 0.05 , and 0.8 ± 0.05 , respectively, where four channel-correlation clusters (i.e., 0.2, 0.4, 0.6, and 0.8) are used. Our empirical results show that the demosaicking performance gets slightly better with an increase in the cluster number. However, more clusters need more training images and computation. For each channel-correlation cluster, we sample 2×10^6 patches with a size of $8 \times 8 \times 3$. In all experiments, we set parameters $\lambda = 0.15$ in Eq. (5) and $\gamma = 0.01$ in Eq. (11).

4.2 Compared Methods

We compare our CADLD with the representative sparse representation-based demosaicking methods. They are the SC,¹⁶ SSC,¹⁷ LSC,¹⁷ LSSC,¹⁷ and PSRD.⁴ We also compare

Table 3 Evaluation of the proposed method on the Kodak dataset.²⁰ The individual and average CPSNR values are reported. “Avg.” stands for “Average.”

Image ID	SC ¹⁶	SSC ¹⁷	LSC ¹⁷	LSSC ¹⁷	PSRD ⁴	PAMD ²²	DDR ²³	CADLD
01	40.85	41.33	40.93	41.38	40.51	41.74	40.05	41.43
02	41.75	41.95	42.01	42.19	41.90	41.58	41.48	42.16
03	43.20	43.40	43.94	44.21	43.20	43.10	43.94	44.05
04	42.24	42.47	42.45	42.81	42.20	41.09	41.62	42.56
05	38.79	38.96	39.22	39.46	38.88	38.20	39.26	39.39
06	41.34	41.69	41.41	41.74	41.20	41.57	41.38	41.61
07	43.26	43.56	43.55	43.96	43.30	42.68	43.60	43.78
08	37.47	37.60	37.43	37.58	37.44	37.64	37.47	37.34
09	43.21	43.42	43.73	43.91	43.13	43.16	43.70	43.67
10	42.89	42.99	43.07	43.24	42.82	42.65	43.35	43.19
11	41.25	41.40	41.34	41.55	41.22	40.85	41.28	41.28
12	44.28	44.62	44.47	44.88	44.26	44.44	44.54	44.62
13	36.21	36.36	36.34	36.47	36.00	37.13	36.15	36.83
14	37.78	37.95	38.61	38.83	37.93	36.87	38.16	38.32
15	41.03	41.45	41.23	41.76	41.00	40.38	40.25	41.43
16	44.36	44.84	44.42	44.90	44.35	44.72	44.64	44.89
17	41.81	41.94	41.86	42.02	41.80	41.94	42.28	42.08
18	38.13	38.23	38.30	38.46	38.07	37.90	38.32	38.47
19	41.86	42.10	42.02	42.34	41.65	41.49	41.96	42.08
20	41.92	41.96	42.21	42.19	41.75	41.72	42.32	42.15
21	40.60	40.68	40.64	40.70	40.47	40.94	40.53	40.42
22	38.81	39.03	39.04	39.33	38.83	38.64	39.39	39.16
23	43.47	43.79	43.93	44.27	43.48	43.10	43.96	44.24
24	35.53	35.53	35.77	35.81	35.50	35.58	35.76	36.22
Avg.	40.92	41.14	41.16	41.42	40.87	40.80	41.06	41.31

Note: The highest CPSNR of each row is shown in bold.

Table 4 Total average of CPSNR over the IMAX²¹ and Kodak²⁰ datasets. “Avg.” stands for “Average.”

Dataset	SC ¹⁶	SSC ¹⁷	LSC ¹⁷	LSSC ¹⁷	PSRD ⁴	PAMD ²²	DDR ²³	CADLD
IMAX	34.31	34.71	35.10	35.82	34.29	34.30	37.13	37.05
Kodak	40.92	41.14	41.16	41.42	40.87	40.80	41.06	41.31
Avg.	37.62	37.93	38.13	38.62	37.58	37.55	39.10	39.18

Note: The highest CPSNR of each row is shown in bold.

with the adaptive dictionary learning methods: PAMD²² and DDR.²³ When comparing different methods, we use their associated dictionaries, with default parameters set by their respective authors.

4.3 Results

We evaluate demosaicking performance on the IMAX²¹ and Kodak datasets.²⁰ We compare the experimental results by CPSNR and visual quality. To eliminate boundary effects,

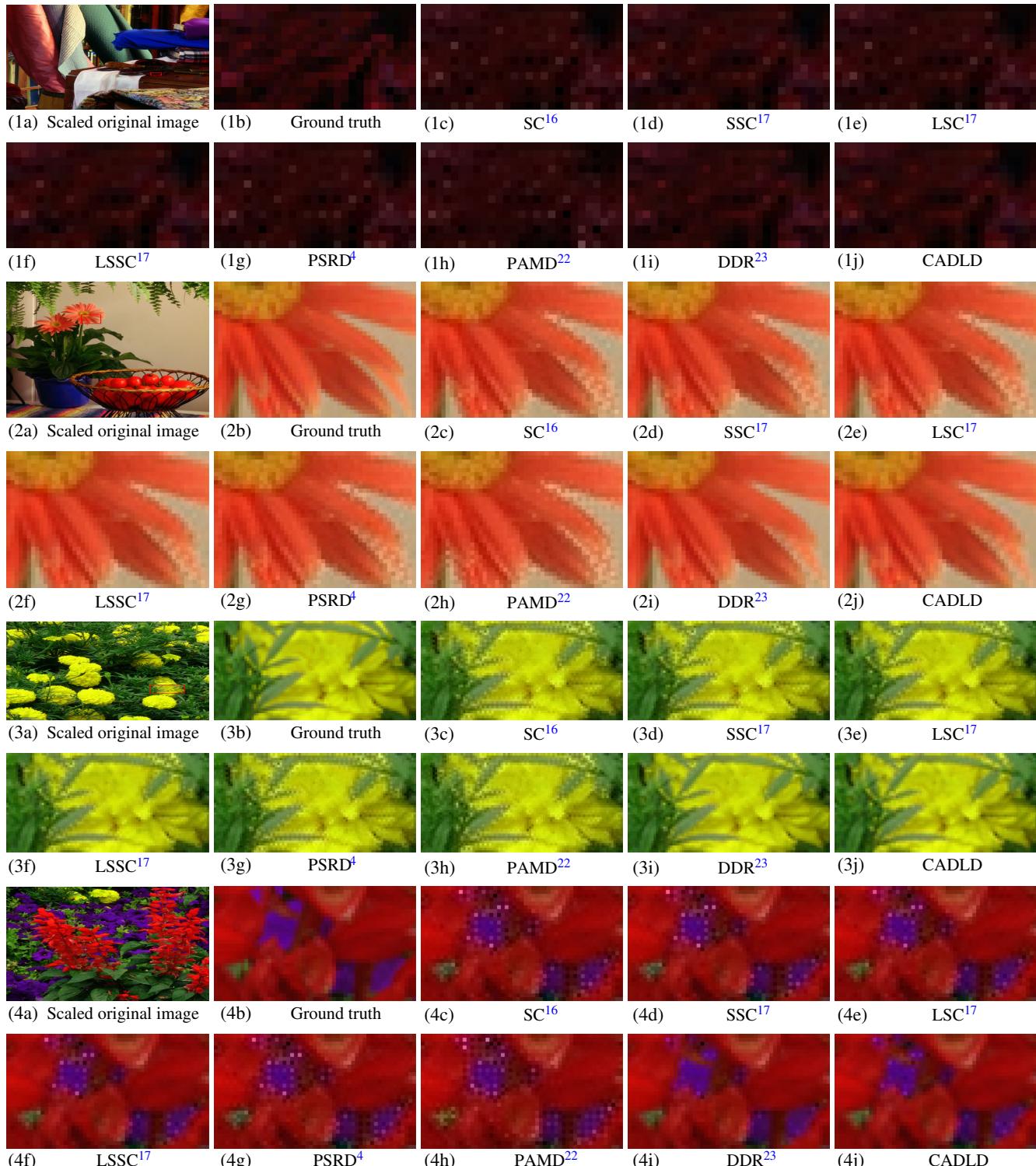


Fig. 5 Blowups of some demosaicked images in the IMAX dataset. From top to bottom, the images are from #2, #9, #16, and #17 images of the IMAX dataset, respectively. In each group, (a) is the scaled original image, in which the red rectangle indicates the selected patch to blow up, (b) is the ground truth, (c)–(i) are the images demosaicked by other dictionary learning methods, and (j) are the images demosaicked by our CADLD. From all the four groups of images, we can clearly see that the images demosaicked by others have severe zipper effect, while those by DDR and our CADLD have better visual quality.

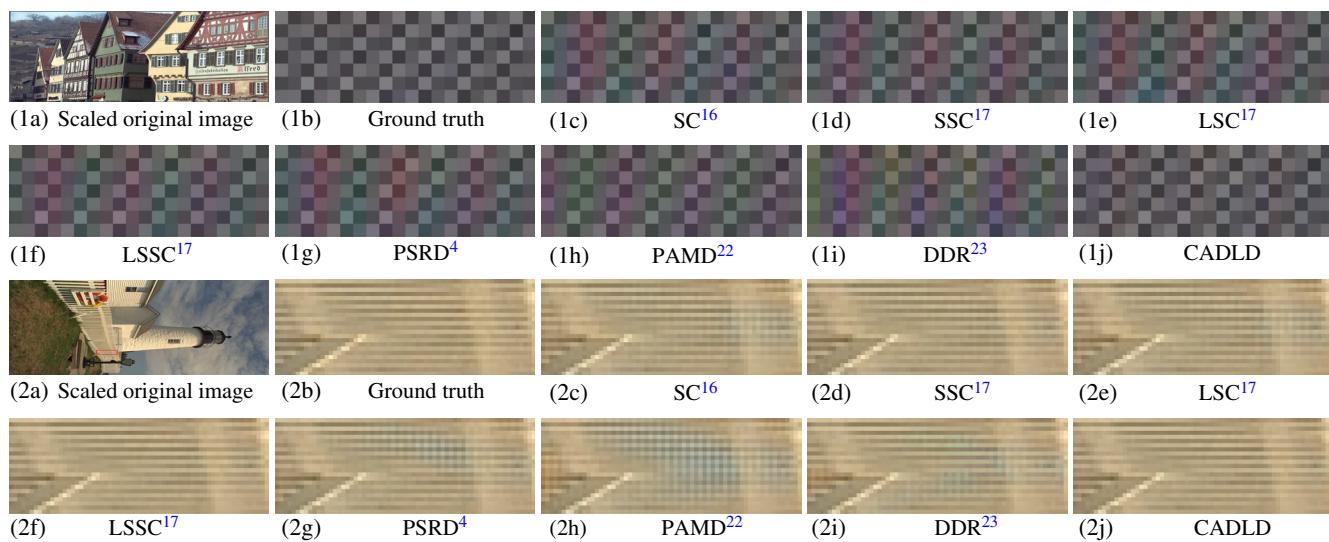


Fig. 6 Blowups of some demosaicked images in the Kodak dataset. From top to bottom, the images are from #8 and #19 images of the Kodak dataset, respectively. In each group, (a) is the scaled original image, in which the red rectangle indicates the selected patch to blow up, (b) is the ground truth, (c)–(i) are the images demosaicked by other dictionary learning methods, and (j) are the images demosaicked by our CADLD. From all the two groups of images, we can clearly see that the images demosaicked by others have severe false color [the first group and (2g)–(2i)] or zipper effect [(2c)–(2f) of the second group], while those by our CADLD have better visual quality.

Table 5 The average running time to demosaick one image from the IMAX dataset.²¹

Methods	SC ¹⁶	SSC ¹⁷	LSC ¹⁷	LSSC ¹⁷	PSRD ⁴	PAMD ²²	DDR ²³	CADLD
Time (s)	59.74	92.31	174.44	273.82	79.51	354.77	12.78	850.63

we exclude an eight-pixel border. The CPSNR values on the IMAX dataset are shown in Table 2. We can see that our CADLD is the second only to DDR²³ on individual image and the whole dataset. The results on the Kodak dataset are shown in Table 3. We can see that our CADLD is the second only to LSSC.¹⁷ This is because that LSSC¹⁷ uses the online dictionary learning method to every image. Overall, our method is superior on both the two datasets, which are significantly different in terms of channel-correction [see Figs. 2(a) and 2(b)]. We also give the total average of CPSNR over the two datasets, which is shown in Table 4. We can see that our CADLD gets the highest average CPSNR value on the two datasets. In addition, our CADLD performs better than PAMD²² on both the two datasets, which also verifies the effectiveness of our channel-correlation defined in Eq. (2).

As the visual quality is typically an important evaluation for demosaicking, we present part of the visual comparison in Figs. 5 and 6. We can see that the visual quality of the CADLD is superior to that of other demosaicking methods, especially in reconstructing along the edges (see Fig. 5) and removing false color at the highly textured regions (see Fig. 6). (Please read detail descriptions on visual difference in the captions.)

We also evaluate the running time of all the compared methods. These methods share the same environment—Intel(R) Core(TM)i7-3632QM CPU @ 2.20 GHz with 16.0 GB RAM. We report the average running time to

demosaick one image from the IMAX dataset,²¹ which is shown in Table 5.

5 Conclusions

In this paper, we propose a channel-correlation adaptive demosaicking method based on dictionary learning. We first define the channel-correlation to measure the correlations among the RGB channels. Then, we analyze the channel-correlation effect on the sparse representation-based demosaicking. We next detail the proposed CADLD method. Experimental results show that our CADLD method works well on raw images with various channel-correlations. Future work will include adapting our method to various CFAs.

Acknowledgments

The work of Chenyan Bai was supported in part by the National Natural Science Foundation of China (NSFC) under Grant No. 61802269, the Key Laboratory of Machine Perception of Peking University (Ministry of Education, Grant Nos. K-2017-06 and K-2018-04), the Excellent Dissertation Foundation of Beijing Jiaotong University (Grant No. 134557522), the Youth Innovative Research Team of Capital Normal University (Grant No. 025185305000/156), and the NVIDIA Corporation. The work of Zhouchen Lin was supported in part by the 973 Program of China under Grant No. 2015CB352502, in part by the National Natural Science Foundation of

China (NSFC) under Grant Nos. 61272341 and 61231002, and in part by the Microsoft Research Asia Collaborative Research Program.

References

1. C. Bai et al., "Automatic design of color filter arrays in the frequency domain," *IEEE Trans. Image Process.* **25**(4), 1793–1807 (2016).
2. J. Li et al., "Automatic design of high-sensitivity color filter arrays with panchromatic pixels," *IEEE Trans. Image Process.* **26**(2), 870–883 (2017).
3. C. Bai et al., "Penrose demosaicking," *IEEE Trans. Image Process.* **24**(5), 1672–1684 (2015).
4. J. Li et al., "Optimized color filter arrays for sparse representation-based demosaicking," *IEEE Trans. Image Process.* **26**(5), 2381–2393 (2017).
5. L. Zhang et al., "Color demosaicking by local directional interpolation and nonlocal adaptive thresholding," *J. Electron. Imaging* **20**(2), 023016 (2011).
6. K.-H. Chung and Y.-H. Chan, "Low-complexity color demosaicing algorithm based on integrated gradients," *J. Electron. Imaging* **19**(2), 021104 (2010).
7. L. Chang and Y.-P. Tan, "Hybrid color filter array demosaicking for effective artifact suppression," *J. Electron. Imaging* **15**(1), 013003 (2006).
8. L. Zhang and X. Wu, "Color demosaicking via directional linear minimum mean square-error estimation," *IEEE Trans. Image Process.* **14**(12), 2167–2178 (2005).
9. D. Paliy et al., "Spatially adaptive color filter array interpolation for noiseless and noisy data," *Int. J. Imaging Syst. Technol.* **17**(3), 105–122 (2007).
10. I. Pekkucuksen and Y. Altunbasak, "Multiscale gradients-based color filter array interpolation," in *17th IEEE Int. Conf. on Image Processing*, pp. 137–140, IEEE (2010).
11. D. Kiku et al., "Minimized-Laplacian residual interpolation for color image demosaicking," *Proc. SPIE* **9023**, 90230L (2014).
13. Y. Monno et al., "Adaptive residual interpolation for color image demosaicking," in *IEEE Int. Conf. on Image Processing*, pp. 3861–3865 (2015).
14. D. Alleysson, S. Susstrunk, and J. Héault, "Linear demosaicing inspired by the human visual system," *IEEE Trans. Image Process.* **14**(4), 439–449 (2005).
15. B. Leung, G. Jeon, and E. Dubois, "Least-squares luma–chroma demultiplexing algorithm for Bayer demosaicking," *IEEE Trans. Image Process.* **20**(7), 1885–1894 (2011).
16. J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. Image Process.* **17**(1), 53–69 (2008).
17. J. Mairal et al., "Non-local sparse models for image restoration," in *Proc. of IEEE Int. Conf. on Computer Vision*, pp. 2272–2279, IEEE (2009).
18. R. Tan et al., "Color image demosaicking via deep residual learning," in *IEEE Int. Conf. on Multimedia and Expo (ICME)*, pp. 793–798, IEEE (2017).
19. M. Gharbi et al., "Deep joint demosaicking and denoising," *ACM Trans. Graphics (TOG)* **35**(6), 1–12 (2016).
20. "Kodak lossless true color image suite," <http://r0k.us/graphics/kodak/> (1 December 2013).
21. "IMAX dataset," http://www4.comp.polyu.edu.hk/~cslzhang/CDM_Dataset.htm (1 April 2010).
22. M. Zhang and L. Tao, "A patch aware multiple dictionary framework for demosaicing," *Lect. Notes Comput. Sci.* **9005**, 236–251 (2014).
23. J. Wu, R. Timofte, and G. L. Van, "Demosaicing based on directional difference regression and efficient regression priors," *IEEE Trans. Image Process.* **25**(8), 3862–3874 (2016).
24. J. Duran and A. Buades, "Self-similarity and spectral correlation adaptive algorithm for color demosaicking," *IEEE Trans. Image Process.* **23**(9), 4031–4040 (2014).
25. M. Everingham et al., "The PASCAL visual object classes challenge 2007 (VOC2007) results," <http://www.pascal-network.org/challenges/VOC/voc2007/workshop/index.html> (9 February 2017).
26. J. Mairal et al., "SPArse modeling software," <http://spams-devel.gforge.inria.fr/> (10 May 2016).
27. Q. Qiu et al., "Domain adaptive dictionary learning," *Lect. Notes Comput. Sci.* **7575**, 631–645 (2012).
28. Z. Lin, R. Liu, and H. Li, "Linearized alternating direction method with parallel splitting and adaptive penalty for separable convex programs in machine learning," *Mach. Learn.* **99**(2), 287–325 (2015).

Chenyen Bai received her PhD in computer science from Beijing Jiaotong University in 2016. She is now an assistant professor at the College of Information Engineering, Capital Normal University. Her research interest is image processing.

Jia Li received his PhD in computer science from Beijing Jiaotong University in 2017. He is now a postdoctor at the Key Laboratory of Machine Perception (Ministry of Education), School of Electronic Engineering and Computer Science, Peking University. His research interests are image processing and machine learning.

Zhouchen Lin received his PhD in applied mathematics from Peking University in 2000. Currently, he is a professor at the Key Laboratory of Machine Perception (Ministry of Education), School of Electronic Engineering and Computer Science, Peking University. His research interests include computer vision, image processing, machine learning, pattern recognition, and numerical optimization. He is an associate editor of *IEEE Transactions on Pattern Analysis and Machine Intelligence* and *International Journal of Computer Vision*.