COLOR DEMOSAICKING VIA NONLOCAL TENSOR REPRESENTATION

Lili Huang[⋆†] *Xuan Wu*[†] *Wenze Shao*[*] *Hongyi Liu*[†] *Zhihui Wei*[†] *Liang Xiao*[†]

[⋆]Guangxi University of Science and Technology, Liuzhou, China
[†]Nanjing University of Science and Technology, Nanjing, China
[*]Nanjing University of Posts and Telecommunications, Nanjing, China

ABSTRACT

A single sensor camera can capture scenes by means of color filter array. Each pixel samples only one of the three primary colors. Color demosaicking (CDM) is a process of reconstruction a full color image from this sensor data. In this paper, we propose a novel CDM scheme based on learned simultaneous sparse coding over nonlocal tensor representation. First, similar 2D patches are grouped to form a three-order tensor, that is, 3D array. Then, three sub-dictionaries, which characterize the coherent structures that appear in each dimension of the grouped tensor, are learned jointly by using Tucker decomposition. The consequent coefficient tensor is imposed by the grouped-block-sparsity constraint, which forces the similar patches to share the same atoms of the dictionaries in their sparse decomposition. Experimental results demonstrate the effectiveness both in the average CPSNR and visual quality.

Index Terms— Color demosaicking, dictionary learning, tensor decomposition, nonlocal method, sparse representation

1. INTRODUCTION

The common approach in single sensor digital cameras is to use a color filter array (CFA) to capture one of the three primary colors at each location. In such digital cameras, an important part of image processing pipeline is color demosaicking (CDM), which estimates the other two missing color components to reconstruct a full color image.

Most recently, sparse representation (SR) and nonlocal self-similarity (NSS) models have shown powerful capability in solving numerous problems in image processing and pattern recognition. They are also popular for CDM methods, such as the self-similarity driven demosaicking (SDD) [1], the learned sparse coding (LSC) demosaicking [2], the learned simultaneous sparse coding (LSSC) demosaicking [3], etc.

Traditional SR and NSS based methods work on vector spaces by representing patches as vectors. However, a patch is intrinsically a matrix, or the second order tensor with two modes including spatial height and width. Vectorization results in damaging nature structure and correlation in the original tensorial data. Therefore, it is desired to carry out sparse coding directly on the tensor objects rather than its vectorized versions. This not only preserves higher-order structure of the image, but can improve learnability of dictionary, especially in cases that the number of training samples is small [4, 5, 6].

The proposed method, as all above methods, which reformulates the demosaicking as the denoising problem, includes two stages. In the first stage, the full color image is initially estimated to recover regions that fit with the NSS assumption. We consider this estimation as noisy signal, and regard interpolation errors as *demosaicking noises*. In the second stage, we use denoising technique based on LSSC over nonlocal tensor representation (LSSC-NTR) to remove the errors, and obtain the improved estimate as result of denoising procedure. Specifically, similar 2D patches are first grouped to form a three-order tensor. Then, three spatially adaptive subdictionaries are learned jointly by using Tucker decomposition, i.e., a form of multilinear principle component analysis (MPCA),which seeks the bases in each dimension that characterize the most representative information regarding the underlying structure of the tensor and thus removes the noise and trivial information [7]. The consequent coefficient tensor is imposed by the grouped-block-sparsity constraint, which forces the similar patches to share the same atoms of the dictionaries in their sparse decomposition. This can implicitly average out the errors not only across the spatial height and width, but also in the third dimension so as to keep the similarity between intensity values at corresponding pixels for the different patches. Finally, each group of similar patches is reconstructed collaboratively by these learned subdictionaries along three ways. Averaging and summing these smaller overlapping patches, we obtain the desired CDM result.

2. NOTATIONS AND PRELIMINARIES

A tensor is a multiway array or multidimensional matrix. The order of a tensor is referred to as the number of its dimensions, also called ways or modes. In this paper, scalars (zero-order tensors) are denoted by lowercase letters; vectors (one-order

The research is supported in part by the National Natural Science Foundation (NSF) of China (61302178, 91538108, 11431015, 61571230, 61402239, 61301217, 61301215), NSF of Guangxi Province (2014GXNSFAA118360), Jiangsu Province (BK2012800, BK20130868, BK20130883).

tensors) by boldface lowercase letters; matrices (two-order tensors) by boldface capital letters; higher-order tensors (order three or higher) by calligraphic capital letters. In general, we denote a N-order tensor by $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ and its elements by $x_{i_1 i_2 \ldots i_N}$, where index $i_n \in \{1, 2, \ldots, I_n\}$ for $1 \leqslant n \leqslant N$. By varying index i_n while fixing the others, we obtain a series of vector-valued subtensor called mode-n fibers. Then mode-n unfolded matrix of $\mathcal{X}$, denoted by $\mathbf{X}_{(n)} \in \mathbb{R}^{I_n \times (I_1 \ldots I_{n-1} I_{n+1} \ldots I_N)}$, is formed by arranging all the mode-n fibers as columns of the matrix. The column rank of $\mathbf{X}_{(n)}$ is known as n-rank of $\mathcal{X}$ denoted by r_n. The N-tuple $(r_1, r_2, \ldots, r_N)$ is called the multilinear rank of $\mathcal{X}$. The Frobenius norm of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is defined as: $\|\mathcal{X}\|_F = \left(\sum_{i_1 i_2 \ldots i_N} x^2_{i_1 i_2 \ldots i_N} \right)^{1/2}$.

The mode-n product of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ by a matrix $\mathbf{U} \in \mathbb{R}^{J \times I_n}$, being the product of all mode-n fibers by $\mathbf{U}$, is a tensor denoted as $\mathcal{Y} = \mathcal{X} \times_n \mathbf{U} \in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_N}$, with elements

$$y_{i_1 \ldots i_{n-1} j i_{n+1} \ldots i_N} = \sum\nolimits_{i_n=1}^{I_n} x_{i_1 \ldots i_n \ldots i_N} u_{j i_n}.$$

The idea can also be expressed according to unfolded tensors:

$$\mathcal{Y} = \mathcal{X} \times_n \mathbf{U} \Leftrightarrow \mathbf{Y}_{(n)} = \mathbf{U} \mathbf{X}_{(n)}. \tag{1}$$

The Tucker decomposition is a form of high-order PCA [8]. It treats $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ as a multilinear transformation of a core tensor by a matrix along each mode, given by

$$\mathcal{X} = \mathcal{G} \times_1 \mathbf{U}^1 \times_2 \mathbf{U}^2 \cdots \times_N \mathbf{U}^N. \tag{2}$$

Here, factor matrix $\mathbf{U}^n \in \mathbb{R}^{I_n \times R_n}$ can be seen as the principal component along nth mode. Elements of the core tensor $\mathcal{G} \in \mathbb{R}^{R_1 \times R_2 \times \cdots \times R_N}$ show the level of interaction among the different components. Via the Kronecker products, the matricized version of the Tucker decomposition (2) is

$$\mathbf{X}_{(n)} = \mathbf{U}^n \mathbf{G}_{(n)} \left(\mathbf{U}^N \otimes \cdots \otimes \mathbf{U}^{n+1} \otimes \mathbf{U}^{n-1} \otimes \cdots \otimes \mathbf{U}^1 \right)^T.$$

It is easy to see that the mode-n fibers of $\mathcal{X}$ belongs to the space spanned by the columns of the matrix $\mathbf{U}^n$.

Finally, for the convenience of presenting the proposed algorithm, as shown in Section 3.2, we introduce the following notation:

$$\mathcal{X} \times_{-n} \{\mathbf{U}\} \triangleq \mathcal{X} \times_1 \mathbf{U}^1 \cdots \times_{n-1} \mathbf{U}^{n-1} \times_{n+1} \mathbf{U}^{n+1} \cdots \times_N \mathbf{U}^N.$$

3. PROPOSED COLOR DEMOSAICKING METHOD

The proposed CDM method has the following steps. 1) Initial interpolation of the green (**G**) channel; 2) LSSC-NTR based enhancement of G; 3) initial interpolation of red (**R**) and blue (**B**) with the enhanced G; 4) LSSC-NTR enhancement of R and B. For step 1 and step 3, we use the Hamilton-Adams interpolation [9] due to its simplicity and efficiency. For step 2 and 4, the used LSSC-NTR scheme is the same. Therefore, in the following we only describe the LSSC-NTR of initially interpolated G channel.

3.1. Proposed model

As previously stated, the interpolated **G** image can be modeled as $\mathbf{Y} = \mathbf{X} + \mathbf{N}$, where $\mathbf{Y}, \mathbf{X}, \mathbf{N} \in \mathbb{R}^{D_1 \times D_2}$ are respectively the initially recovered **G** channel, the unknown true **G** channel, and the CDM error. It is well known that recovering $\mathbf{X}$ from $\mathbf{Y}$ is an ill-posed inverse problem, and the regularization technique is often utilized to find a stable solution. Thus, we resort to the following optimization problem

$$\min_{\mathbf{X}} \lambda \|\mathbf{Y} - \mathbf{X}\|_F^2 + \mathrm{Reg}(\mathbf{X}), \tag{3}$$

where λ is a positive parameter, the first term is the image fidelity term, and the second term is the regularization term that exploits prior knowledge about $\mathbf{X}$. Next we fix our attention on the modeling for regularizer.

3.1.1. From LSC to LSSC

In the traditional SR model, an image patch of size $\sqrt{n} \times \sqrt{n}$ is reshaped into a column vector $\mathbf{x} \in \mathbb{R}^n$, and assumed to be compactly expressed, or efficiently approximated, as a linear combination of a few columns (referred to as atoms) from the given redundant dictionary $\mathbf{D} \in \mathbb{R}^{n \times m}$ with $m > n$. This is achieved by defining the representation as the solution of

$$\min_\alpha \|\alpha\|_0, \text{ s.t. } \mathbf{x} \approx \mathbf{D}\alpha, \tag{4}$$

where $\|\alpha\|_0$ counts the nonzero elements in α, and we expect a sparse outcome, $\|\alpha\|_0 = r \ll n$. Solving (4) is commonly referred to as atomic decomposition or sparse coding [10].

It is essential to choose the dictionary for employment of the above SR model. Compared with an off-the-shelf dictionary, learning the dictionary adapted to a set of patches $\{\mathbf{x}_i \in \mathbb{R}^n\}_{i=1}^I$ encourages higher sparsity and leads to more effective modeling. This can be formulated as the problem

$$\min_{\mathbf{D}, \mathbf{A}} \sum\nolimits_{i=1}^I \|\alpha_i\|_0, \text{ s.t. } \forall i \, \mathbf{x}_i \approx \mathbf{D}\alpha_i, \tag{5}$$

where $\mathbf{A} \in \mathbb{R}^{m \times I}$ contains I representation coefficients $\{\alpha_i \in \mathbb{R}^m\}_{i=1}^I$.

Solving (5) can be called learned sparse coding (LSC). This model assumes that the representation of each patch has its own nonzero pattern, which discounts relationships that may exist between them. This problem can be addressed by grouping similar patches and using the joint-sparsity (also known as grouped-sparsity) model, as illustrated in Fig. 1. It amounts to solve the following problem

$$\min_{\mathbf{D}, \{\mathbf{A}_i\}_{i=1}^I} \sum\nolimits_{i=1}^I \|\mathbf{A}_i\|_{0,\infty}, \text{ s.t. } \forall i \, \mathbf{X}_i \approx \mathbf{D}\mathbf{A}_i, \tag{6}$$

where $\mathbf{X}_i = \left[\mathbf{x}_j^i\right]_{j \in S_i} \in \mathbb{R}^{n \times |S_i|}$, and $\mathbf{A}_i = \left[\alpha_j^i\right]_{j \in S_i} \in \mathbb{R}^{m \times |S_i|}$. Here, $S_i = \left\{ j \,\middle|\, 1 \leqslant j \leqslant I, \, \left\|\mathbf{x}_j^i - \mathbf{x}_i\right\|_2^2 \leqslant \xi \right\}$ is a index set of similar patches to $\mathbf{x}_i$, and ξ is a preset threshold. The grouped-sparsity regularizer measured by pseudo

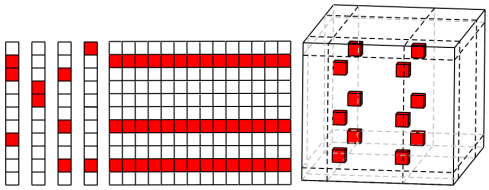


Fig. 1: Sparsity vs. grouped-sparsity vs. grouped-block-sparsity: Red squares represents nonzeros values in vectors (left), matrix (middle), or three-order tensor (right).

norm $\|\mathbf{A}_i\|_{0,\infty} \triangleq \sum_r \psi\left(\|\mathbf{A}_i(r,:)\|_\infty\right)$ counts the number of nonzero rows, where $\mathbf{A}_i(r,:)$ is the r-th row of $\mathbf{A}_i$ and ψ is the indicator function defined as: $\psi(a) = 1$ for $|a| > 0$, otherwise $\psi(a) = 0$. Solving (6) can be called learned simultaneous sparse coding (LSSC), which ensures that similar patches admit similar decompositions.

3.1.2. LSSC on nonlocal tensor representation

In this subsection, we generalize the grouped-sparsity prior of vectors to tensors so as to reduce the initial interpolation errors and enhance the color reproduction of **G** channel.

Let $\left\{\mathbf{X}_i \in \mathbb{R}^{d_1 \times d_2}\right\}_{i=1}^{I}$ denote a set of overlapped patches extracted in every location in the image $\mathbf{X}$, where $I = (D_1 - d_1 + 1)(D_2 - d_2 + 1)$ is the number of the patches. In order to accurately represent the diverse local structures of the image, we group these patches into K clusters via K-means++ method [11]. Similar patches in each cluster are stacked along third mode to create a three-order tensor with size $d_1 \times d_2 \times d_3$ denoted by $\mathbf{R}_k(\mathbf{X}) \triangleq \mathcal{X}_k$ $(k = 1, 2, \ldots, K)$, where d_3 is the number of similar patches in the k-th cluster, and $\mathbf{R}_k$ denotes an operator grouping similar 2D patches into the three-order tensor. From Eq. (1), the problem (6) can be rewritten as

$$\min_{\mathbf{D},\{\mathbf{A}_i\}_{i=1}^{I}} \sum_{i=1}^{I} \|\mathbf{A}_i\|_{0,\infty}, \text{ s.t. } \forall i\ \mathbf{X}_i \approx \mathbf{A}_i \times_1 \mathbf{D}. \quad (7)$$

Then we can extend this problem into tensor case:

$$\begin{aligned} &\min_{\{\mathbf{D}^j\}_{j=1}^{3},\{\mathcal{A}_k\}_{k=1}^{K}} \sum_{k=1}^{K} \|\mathcal{A}_k\|_{GBS}, \\ &\text{s.t. } \forall k\ \mathcal{X}_k \approx \mathcal{A}_k \times_1 \mathbf{D}^1 \times_2 \mathbf{D}^2 \times_3 \mathbf{D}^3, \end{aligned} \quad (8)$$

where $\mathcal{A}_k \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ is the coefficient tensor of the tensorial data $\mathcal{X}_k$, and $\{\mathbf{D}^j \in \mathbb{R}^{d_j \times m_j}\}_{j=1}^{3}$ are three redundant dictionaries, whose columns can be selected to encode the local and nonlocal information of $\mathcal{X}_k$ in three different modes: two spatial modes and one clustering mode, respectively, which is not the same as traditional learned dictionary that encodes all the information in one low dimensional subspace. The pseudo norm $\|\mathcal{A}_k\|_{GBS}$, called grouped-block-sparsity regularizer, generalizes the concept of grouped-sparsity of vectors to tensor, which can be easily understood as illustrated in Fig. 1 and has the following definition [12, 13]:

Definition (grouped-block-sparsity). A tensor $\mathcal{X}_k \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ is $\left(r_k^1, r_k^2, r_k^3\right)$-grouped-block sparse with respect to the dictionaries $\{\mathbf{D}^j\}_{j=1}^{3}$, if it admits a sparse representation based only on few selected columns of each dictionary $\left(r_k^j << m_j,\ j = 1, 2, 3\right)$, i.e., $\|\mathcal{A}_k\|_{GBS} \leqslant \left(r_k^1, r_k^2, r_k^3\right)$ if and only if each smallest subset of index, denoted by $\mathcal{I}^j$ $(j = 1, 2, 3)$, which contains r_k^j $(j = 1, 2, 3)$ elements, satisfies $a_{i_1 i_2 i_3} = 0$ for all $(i_1, i_2, i_3) \notin \mathcal{I}^1 \times \mathcal{I}^2 \times \mathcal{I}^3$.

Therefore, the optimization task (8) can be changed to be

$$\begin{aligned} &\min_{\{\mathbf{D}^j\}_{j=1}^{3},\{\mathcal{A}_k\}_{k=1}^{K}} \sum_{k=1}^{K} \left\|\mathcal{X}_k - \mathcal{A}_k \times_1 \mathbf{D}^1 \times_2 \mathbf{D}^2 \times_3 \mathbf{D}^3\right\|_F^2, \\ &\text{s.t. } \forall k\ \|\mathcal{A}_k\|_{GBS} \leqslant \left(r_k^1, r_k^2, r_k^3\right). \end{aligned}$$

Integrating the above prior knowledge into the general model (3) brings about the following denoising scheme:

$$\begin{aligned} &\min_{\mathbf{X},\{\mathbf{D}^j\}_{j=1}^{3},\{\mathcal{A}_k\}_{k=1}^{K}} \left\{\lambda \|\mathbf{Y} - \mathbf{X}\|_F^2 \right. \\ &\qquad \left. + \sum_{k=1}^{K} \left\|\mathcal{X}_k - \mathcal{A}_k \times_1 \mathbf{D}^1 \times_2 \mathbf{D}^2 \times_3 \mathbf{D}^3\right\|_F^2 \right\}, \\ &\text{s.t. } \forall k\ \|\mathcal{A}_k\|_{GBS} \leqslant \left(r_k^1, r_k^2, r_k^3\right). \end{aligned} \quad (9)$$

Thanks to the sparsity structure of the coefficient tensor $\mathcal{A}_k$, we can further simplify above the problem. To this end, let $\mathcal{Z}_k \triangleq sub(\mathcal{A}_k) \in \mathbb{R}^{r_k^1 \times r_k^2 \times r_k^3}$, whose index of element belongs to the index set $\mathcal{I}^1 \times \mathcal{I}^2 \times \mathcal{I}^3$, denote the intrinsic subtensor of $\mathcal{A}_k$. Thus, it can select the columns from the three dictionaries to reconstruct $\mathcal{X}_k$. In other words, these columns constitute the subdictionaries $\{\mathbf{D}_k^j\}_{j=1}^{3}$ that only relate to the cluster $\mathcal{X}_k$. With this in mind, the problem (9) can be equivalently reformulated as:

$$\begin{aligned} &\min_{\mathbf{X},\left\{\mathbf{D}_k^1,\mathbf{D}_k^2,\mathbf{D}_k^3,\mathcal{Z}_k\right\}_{k=1}^{K}} \left\{\lambda \|\mathbf{Y} - \mathbf{X}\|_F^2 \right. \\ &\qquad \left. + \sum_{k=1}^{K} \left\|\mathcal{X}_k - \mathcal{Z}_k \times_1 \mathbf{D}_k^1 \times_2 \mathbf{D}_k^2 \times_3 \mathbf{D}_k^3\right\|_F^2 \right\}, \end{aligned} \quad (10)$$

where $\{\mathbf{D}_k^j \in \mathbb{R}^{d_j \times r_k^j}\}_{j=1}^{3}$ are three compact subdictionaries with $d_j \geqslant r_k^j$, since there are high correlations between atoms of them. This reformulation transforms the original constrained optimization (9) into an unconstrained optimization that easy to solve.

3.2. Proposed algorithm

Since the proposed problem (10) is a non-convex joint optimization, its local optimal solution can be obtained by alternating optimization technique.

Firstly, when $\mathbf{X}$ is fixed, the minimization in (10) is reduced to a series of smaller problems:

$$\min_{\left\{\mathbf{D}_k^j\right\}_{j=1}^{3},\mathcal{Z}_k} \left\|\mathcal{X}_k - \mathcal{Z}_k \times_1 \mathbf{D}_k^1 \times_2 \mathbf{D}_k^2 \times_3 \mathbf{D}_k^3\right\|_F^2, \quad (11)$$

with $k = 1, 2, \ldots, K$. The optimization problem (11) can be readily solved by the Tucker decomposition technique which

is form of MPCA. Particularly, the higher-order orthogonal iteration algorithm is utilized to learn the subdictionaries, which alternatively updates the following formulas [14]:

$$\mathbf{D}_k^j = \mathrm{SVD}\left(\left(\mathcal{X}_k \times_{-j} \left\{\mathbf{D}_k^T\right\}\right)_{(j)}, r_k^j\right),\ j = 1, 2, 3,$$
$$\mathcal{Z}_k = \mathcal{X}_k \times_1 \mathbf{D}_k^{1\,T} \times_2 \mathbf{D}_k^{2\,T} \times_3 \mathbf{D}_k^{3\,T},$$

where r_k^j can be determined by minimizing the Akaike information criterion [15], and the notation $\mathrm{SVD}(\mathbf{W}, r)$ indicates the r leading left singular vectors of matrix $\mathbf{W}$. Then, we obtain the $\mathrm{rank}(r_k^1, r_k^2, r_k^3)$ approximation of the tensor $\mathcal{X}_k$, denoted as

$$\hat{\mathcal{X}}_k = \mathcal{Z}_k \times_1 \mathbf{D}_k^1 \times_2 \mathbf{D}_k^2 \times_3 \mathbf{D}_k^3.$$

Secondly, when $\left\{\mathbf{D}_k^1, \mathbf{D}_k^2, \mathbf{D}_k^3, \mathcal{Z}_k\right\}_{k=1}^K$ are fixed, the problem (10) is quadratic in $\mathbf{X}$, noted that $\mathbf{R}_k(\mathbf{X}) \triangleq \mathcal{X}_k$, and its closed-form solution is

$$\mathbf{X} = \left(\lambda \mathbf{I} + \sum\nolimits_{k=1}^K \mathbf{R}_k^T \mathbf{R}_k\right)^{-1} \left(\lambda \mathbf{Y} + \sum\nolimits_{k=1}^K \mathbf{R}_k^T \hat{\mathcal{X}}_k\right).$$

4. EXPERIMENTAL RESULTS

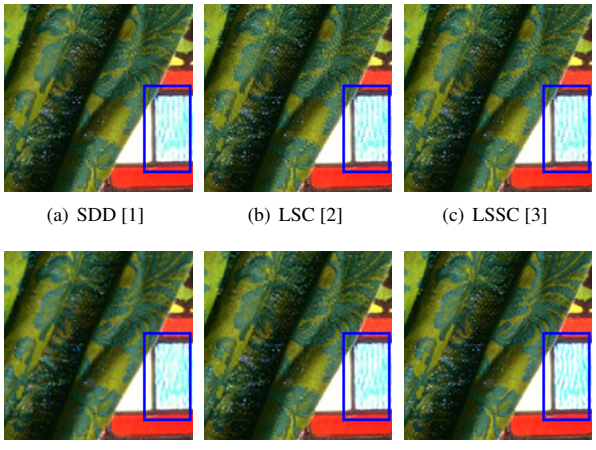
In this section, simulations are performed to evaluate the proposed method. We conduct the simulated experiments on the IMAX dataset[1]. The IMAX dataset consists of 18 images and the image size is 500×500, which are cropped from original 2310×1814 high-resolution images. For the proposed algorithm, the main parameters are empirically tuned for the optimal performance and set as follows: patch size 5×5, the number of clusters $K = 90$, the regularization parameter $\lambda = 0.2$.

Table 1: CPSNR (dB) for different methods on the IMAX dataset, where the bold font represents the best performance.

Image	SDD [1]	LSC [2]	LSSC [3]	PSDD [16]	Proposed
1	27.37	27.47	28.21	28.44	**28.80**
2	34.09	33.64	34.31	34.12	**34.82**
3	32.66	32.67	**32.98**	30.12	32.66
4	35.96	35.98	**36.97**	33.97	36.80
5	31.89	31.19	32.17	33.48	**33.64**
6	35.46	35.16	37.86	**38.27**	37.60
7	36.43	38.93	**39.37**	33.17	35.32
8	37.37	37.93	**38.58**	35.94	37.70
9	36.21	36.02	**37.34**	36.63	37.29
10	37.63	36.84	38.04	37.80	**38.60**
11	38.54	37.98	**39.49**	38.74	39.27
12	37.50	37.03	37.95	37.32	**38.40**
13	40.05	38.92	40.09	40.08	**40.78**
14	38.00	37.19	38.22	38.00	**38.50**
15	38.26	37.62	38.72	38.44	**38.88**
16	31.10	31.17	32.37	33.16	**33.39**
17	30.53	30.90	32.39	**33.60**	33.15
18	34.23	34.20	35.03	34.15	**35.34**
Average	35.18	35.05	36.12	35.30	**36.16**

[1]The IMAX set is downloaded from http://www4.comp.polyu.edu.hk/˜cslzhang/CDM_Dataset.htm.

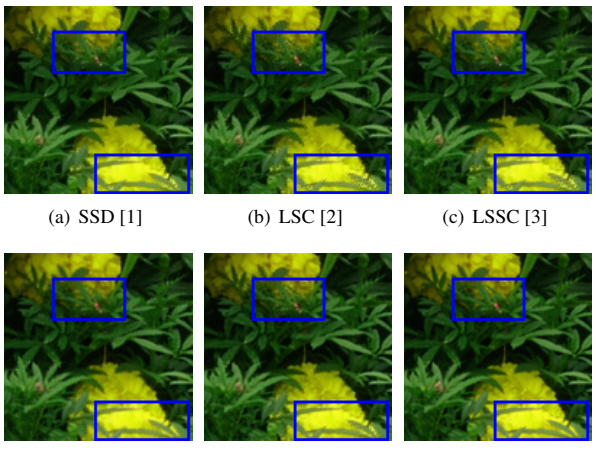
The color peak signal-to-noise ratio (CPSNR) is adopted to assess the quality of the resulting images. Table 1 reports the CPSNR results of five CDM methods including the proposed. Best results are in bold. From Table 1, the average CPSNR of the proposed method is higher than that of SDD, LSC, PSDD and close to that of LSSC. Fig. 2 and Fig. 3 show the visual comparison, which include the cropped and zoomed results of the five methods including the proposed on *images 1* and *16* in the IMAX dataset. It can be clearly seen that the proposed method is more free of speckle color noises and reproduce sharper and cleaner edges without much zipper effects.



(a) SDD [1] (b) LSC [2] (c) LSSC [3]

(d) PSDD [16] (e) Proposed (f) Ground truth

Fig. 2: Cropped and zoomed CDM results on *Image 1* in the IMAX dataset by different methods. Zoom into pdf file for a better view.



(a) SDD [1] (b) LSC [2] (c) LSSC [3]

(d) PSDD [16] (e) Proposed (f) Ground truth

Fig. 3: Cropped and zoomed CDM results on *Image 16* in the IMAX dataset by different methods.

5. REFERENCES

[1] A. Buades, B. Coll, J.M. Morel, et al., "Self-similarity driven color demosaicking," *IEEE Trans. on Image Processing*, vol. 18, no. 6, pp. 1192–1202, 2009.

[2] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. on Image Processing*, vol. 17, no. 1, pp. 53–69, 2008.

[3] J. Mairal, F. Bach, J. Ponce, et al., "Non-local sparse models for image restoration," in *Proc. Int. Conf. Computer Vision*. IEEE, 2009, vol. 30, pp. 2272–2279.

[4] X. He, D. Cai, and P. Niyogi, "Tensor subspace analysis.," *Advances in Neural Information Processing Systems*, pp. 499–506, 2005.

[5] M. Yin, J. Gao, and S Cai, "Image super-resolution via 2d tensor regression learning," *Computer Vision and Image Understanding*, vol. 132, pp. 12–23, 2015.

[6] G. Zhong and M. Cheriet, "Tensor representation learning based image patch analysis for text identification and recognition," *Pattern Recognition*, vol. 48, no. 4, pp. 1211–1224, 2015.

[7] H. Lu, K. Plataniotis, and A. Venetsanopoulos, "Mpca: Multilinear principal component analysis of tensor objects," *IEEE Transactions on Neural Networks*, vol. 19, no. 1, pp. 18–39, 2008.

[8] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *Siam Review*, vol. 66, no. 4, pp. 294–310, 2005.

[9] J.F. Hamilton Jr. and J.E. Adams, "Adaptive color plane interpolation in single sensor color electronic camera," *U.S. Patent 5 629 734*, 1997.

[10] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. on Image Processing*, vol. 15, no. 12, pp. 3736–3745, 2006.

[11] D. Arthur and S. Vassilvitskii, "k-means++: the advantages of careful seeding," in *Proc. Eighteenth Acm-Siam Symposium on Discrete Algorithms*. ACM-SIAM, 2015, pp. 1027–1035.

[12] C. F. Caiafa and A. Cichocki, "Computing sparse representations of multidimensional signals using kronecker bases.," *Neural Computation*, vol. 25, no. 1, pp. 186–220, 2013.

[13] Y. Peng, D. Meng, Z. Xu, et al., "Decomposable nonlocal tensor dictionary learning for multispectral image denoising," in *Proc. Computer Vision and Pattern Recognition*. IEEE, 2014, pp. 2949–2956.

[14] L.D. Lathauwer, B.D. Moor, and J. Vandewalle, "On the best rank-1 and rank-$(r_1, r2, \ldots, r_N)$ approximation of higher-order tensors.," *SIAM J. matrix Anal. appl.*, vol. 21, no. 4, pp. 1324–1342, 2000.

[15] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. on Acoust., Speech, Signal Processing*, vol. 33, no. 2, pp. 387–392, 1985.

[16] J.T. Korneliussen and K. Hirakawa, "Camera processing with chromatic aberration," *IEEE Trans. on Image Processing*, vol. 23, no. 10, pp. 4539–52, 2014.