Tensor Nuclear Norm-Based Low-Rank Approximation With Total Variation Regularization

Yongyong Chen, Shuqin Wang, and Yicong Zhou, *Senior Member, IEEE*

***Abstract*—Some existing low-rank approximation approaches either need to predefine the rank values (such as the matrix/tensor factorization-based methods) or fail to consider local information of data (e.g., spatial or spectral smooth structure). To overcome these drawbacks, this paper proposes a new model called the tensor nuclear norm-based low-rank approximation with total variation regularization (TLR-TV) for color and multispectral image denoising. TLR-TV uses the tensor nuclear norm to encode the global low-rank prior of tensor data and the total variation regularization to preserve the spatial-spectral continuity in a unified framework. Including the hyper total variation (HTV) and spatial-spectral total variation (SSTV), we propose two TLR-TV-based algorithms, namely TLR-HTV and TLR-SSTV. Using the alternating direction method of multiplier, we further propose two simple algorithms to solve TLR-HTV and TLR-SSTV. Extensive experiments on simulated and real-world noisy images demonstrate that the proposed TLR-HTV and TLR-SSTV outperform the state-of-the-art methods in color and multispectral image denoising in terms of quantitative and qualitative evaluations.**

***Index Terms*—Low-rank tensor approximation, tensor nuclear norm, hyper total variation, spatial-spectral total variation, image denoising.**

I. Introduction

EXPLOITING low-dimensional structure from high-dimensional data, low-rank matrix approximation (LRMA) and low-rank tensor approximation (LRTA) [1]–[3] have received increasing attention in the fields of computer vision and pattern recognition. Because images are inevitably corrupted by noise during the image acquisition and transmission procedures, this poses great challenges to further image processing, such as image inpainting [4], image classification [5], and object detection [6]. Therefore, image denoising is an important and necessary step to recover a clean image data from contaminated observations.

The core of image denoising based on LRMA and LRTA [1], [3], [7]–[11] is how to exactly build a proper low-rank regularizer to measure the global structure of the underlying image data due to the fact that the high-dimensional data inherently possess a low-rank structure [1]–[3]. As a traditional method, principal component analysis (PCA) uses the l_2-norm fidelity to measure the error. However, because l_2-norm is optimal to suppress additive Gaussian noise, PCA fails to measure outliers. To address this issue, Candès *et al.* [7] and Wright *et al.* [1] proposed a robust PCA (RPCA) model in which the error (including outliers, gross corruptions) is measured by l_1-norm and the low-rank property is explored by the nuclear norm. Then, the noisy observation matrix Y is decomposed into a low-rank term L and a sparse component S. Many variants of LRMA [9]–[14] have been developed. For more details about LRMA, one refers to [15], [16]. Most of the recent approaches for LRMA can be roughly classified into two categories: 1) matrix rank minimization, which is often achieved by minimizing its convex relaxation; and 2) matrix factorization, where the factor matrices naturally upper-bound the rank of the approximation. For example, shang *et al.* [17] proposed a scalable robust bilinear factorization method to recover low-rank and sparse matrices. The low-rank matrix $L \in \mathbf{R}^{n_1 \times n_2}$ is factorized as the product of two small factor matrices $U \in \mathbf{R}^{n_1 \times r}$ and $V \in \mathbf{R}^{r \times n_2}$, i.e., $L = UV$. Thus, the low-rank property can be guaranteed by $rank(L) = rank(UV) \leq \min\{rank(U), rank(V)\}$ and $r \ll \min\{n_1, n_2\}$. The basic assumption of matrix-factorization-based methods [17]–[19] is that the low-rank prior r is known. However, this assumption is impractical since the true rank value of matrix data is unknown in real applications. On the other hand, to handle the tensor data, the above methods should unfold the tensor data to produce matrix Y among the spectral mode, indicating that LRMA-based denoising approaches are exploited only the spectral correlation of tensor data. This is insufficient and may cause the loss of useful information along other two dimensions [3]. The last limitation of LRMA-based approaches is that they can handle only the matrix data [2], [3].

Recent advances [2], [3], [20], [21] have investigated that tensor-based methods have great potential in image processing. Unfortunately, the tensor rank is difficult to define and implement in real applications. In the literature, two common tensor decomposition forms are CANDECOMP/PARAFAC (CP) [22] and Tucker decomposition [23]. These result in different definitions of tensor rank. For example, the CP rank is defined as the

Manuscript received April 14, 2018; revised August 12, 2018; accepted September 7, 2018. Date of publication October 1, 2018; date of current version December 17, 2018. This work was supported in part by the Macau Science and Technology Development Fund under Grant FDCT/189/2017/A3 and in part by the Research Committee at University of Macau under Grants MYRG2016-00123-FST and MYRG2018-00136-FST. The guest editor coordinating the review of this manuscript and approving it for publication was Prof. Thierry Bouwmans. *(Corresponding author: Yicong Zhou.)*

Y. Chen and Y. Zhou are with the Department of Computer and Information Science, University of Macau, Macau 999078, China (e-mail: yongyongchen.cn@hotmail.com; yicongzhou@umac.mo).

S. Wang is with the College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266590, China (e-mail: ShuqinWang.cn@hotmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTSP.2018.2873148

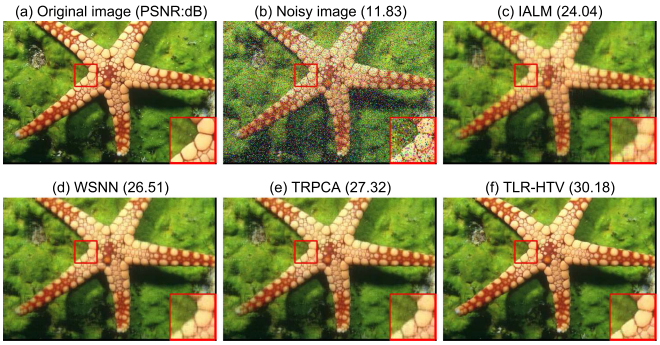

Fig. 1. Comparison of color image denoising. (a) Original image, (b) noisy image corrupted by salt&pepper noise, (c)–(f) reconstructed by IALM [29], WSNN [21], TRPCA [3], and our TLR-HTV, respectively. We can see that the images reconstructed by low-rank matrix and tensor-based methods lose some spatial structure of the color image, and our proposed TLR-HTV can recover the fine structure of the color image.

smallest number of rank-1 factors since it decomposes a tensor as a number of rank-1 factors. As a result, it may suffer from high computation cost [24]. The best low-rank approximation is unknown because it is difficult to compute the CP rank [3]. Tucker decomposition [23] decomposes a tensor data into one small core tensor and a set of matrices [3], [25]. However, the ranks along all modes must be manually predefined before running the algorithm [26]. This may be intractable in practice. Following the similar idea of matrix factorization, tensor factorization [27], [28] is also adopted to encode the low-rank property. Apart from the challenging issue of the unknown rank value, the local minima of matrix factorization and tensor factorization may lead to performance degradation.

These LRMA-and LRTA-based approaches still face two main challenges: 1) some existing low-rank approximation approaches may need to manually pre-define the rank values (especially matrix factorization and tensor factorization-based methods [17]–[19], [27]). 2) LRMA-and LRTA-based approaches explore only the high correlation along the spectral dimension, without consideration of the spatial piecewise smooth structure of image data. Their performances are not satisfactory for many real applications. As shown in Fig. 1, the images reconstructed by LRMA-and LRTA-based methods still remain noise and suffer from the loss of local details (see the zoomed window). Therefore, ignoring local structure of image data leads to the fact that current LRMA-and LRTA-based approaches are still lack of robustness to recover details of images. Many efforts using total variation (TV) regularization [26], [30]–[32] have been developed to capture the local characteristics in color and multispectral images which are of crucial importance in magnetic resonance image super-resolution [33], hyperspectral image denoising [26], [31]. These phenomena motivate us to combine the global property (low-rank prior) with the local structure information of the image data in the tensor level.

To address these issues, we propose a new model called tensor nuclear norm-based low-rank approximation with total variation regularization (TLR-TV) in a tensor framework. It treats the noisy color and multispectral images as the third-order tensor data, taking advantage of the inherently spatial-spectral relation. TLR-TV can simultaneously explore the high correlation of image data and the spatial-spectral local consistency for effective image recovery. Instead of using the traditional nuclear norm [1], [29], matrix factorization or tensor factorization regularizers [17], [19], [27], [28], the proposed method regularizes the inherently low-rank property by the tensor nuclear norm because it has superior power of capturing the "spatial-shifting" correlation [8] and avoids pre-estimating the rank prior. The tensor nuclear norm of the tensor data $\mathcal{X}$ is defined as the sum of singular values of all frontal slices of $\hat{\mathcal{X}}$, where $\hat{\mathcal{X}}$ is the tensor obtained by the fast Fourier transformation along the tube fibers of $\mathcal{X}$. And the tensor nuclear norm has been proved to be the tightest convex envelope of the tensor multi-rank [25]. The tensor nuclear norm is based on a novel tensor decomposition framework [8], named tensor-SVD. It has a similar structure of the matrix singular value decomposition. The resulting algebraic operation is close to that of matrix algebraic operation (please see Definition 3.2 in Section III). Our main contributions are as follows.

- We propose a new TLR-TV model as a unified framework. TLR-TV integrates the global low-rank property and local piecewise smooth characteristic for image denoising. The proposed TLR-TV uses the tensor nuclear norm to model the low-rank prior and the total varaition regularization to preserve the spatial consistency for effective image recovery.
- Integrating the hyper total variation (HTV) and spatial-spectral total variation (SSTV) to our TLR-TV model, we further propose two algorithms for color and multispectral image denoising, namely TLR-HTV and TLR-SSTV. To solve the optimization problems of TLR-HTV and TLR-SSTV, we propose two easy-to-implement algorithms via the alternating direction method of multipliers (ADMM).
- Extensive experiments on simulated and real noisy images validate the effectiveness of our proposed methods qualitatively and quantitatively. The results demonstrate that the proposed TLR-HTV is better to exploit the local smooth structure for color image and TLR-SSTV is the better one for multispectral image.

The outline of this paper is structured as follows. Section II reviews the related works. Section III presents several notations and definitions. Section IV introduces our method for image denoising, and designs two efficient algorithms based on ADMM. Section V evaluates the performance of our proposed methods and Section VI concludes this paper.

II. Related Work

In this section, we roughly divide the related works of image denoising into three categories: low-rank matrix approximation (LRMA), low-rank tensor approximation (LRTA), and low-rank plus total variation-based methods.

A. LRMA-Based Methods

LRMA aims to decompose the unfolding matrix data into a superposition of a low-rank component representing the clean image and a sparse term denoting the sparse noise, such as

impulse noise, stripes, dead lines [9], [11]. Lin *et al.* [29] proposed a fast and efficient solver for robust principal component analysis. Based on this framework, Zhang *et al.* [9] proposed an image restoration model named low-rank matrix recovery (LRMR) by the bilateral random projection-based GoDec [34]. However, LRMR needs to estimate the matrix rank value, which is hard to determine. Thus the result of LRMR may be suboptimal. Essentially, LRMA-based methods [9]–[11], [35] consider the low-rank structure in only the mode-3 unfolding matrix of a given tensor data. In other words, they effectively exploit only the spectral redundancy among mode-3. Thus, these aforementioned approaches may face several deficiencies. Firstly, these approaches are able to handle only the matrix data, which severely limits their scalability. Secondly, the process of transforming the multi-dimensional data into a matrix data unavoidably causes the structure information loss without taking the spatial correlation within a channel (of color image) or band (of multispectral image) into account.

B. LRTA-Based Methods

To overcome the first shortcoming of LRMA-based methods, many works [2], [3], [21], [25], [30] extend the low-rank property into the tensor level and possess the strong performance guarantee over the matrix methods. For example, considering both the spatial and spectral information, Xie *et al.* [36] proposed to denoise multispectral images based on Tucker and CP low-rank decomposition for a general tensor. From both theoretical and computational aspects, Goldfarb *et al.* [2] presented a convex optimization framework HoRPCA and a nonconvex model for LRTA. Then, the model HoRPCA was applied for image denoising in [37]. Considering that the contributions of different modes for the Tucker rank may be not equal, Huang *et al.* [21] proposed to interpret the weighted strategy for all unfolding matrices into the HoRPCA framework.

C. Low-Rank Plus Total Variation-Based Methods

These above-mentioned LRMA and LRTA approaches share several limitations: they may either need to manually predefine the rank values (especially matrix/tensor factorization-based methods) or ignore the local structure information of the image data. Several efforts [18], [27] have been made to estimate the rank values of clean data. However, these methods may overestimate or underestimate the rank values of the recovered data since the true rank values of matrix and tensor data are unknown in real scenarios. Therefore, works [26], [32], [33] proposed to exploit the total varation for preserving edge information and the local piecewise smoothness. In [26], He *et al.* integrated the hyper total variation norm and the nuclear norm into a rank-constrained LRMA model. Shi *et al.* [33] focused on magnetic resonance image super-resolution from their low-resolution counterparts using the LRMA and TV regularization.

III. Preliminaries

In this section, we first give the definition of t-product (Definition 3.1) which provides a matrix-like multiplication of third-order tensors. The tensor-SVD (Definition 3.2) is then efficiently computed by matrix singular value decomposition in the Fourier domain. Finally, the tensor nuclear norm (Definition 3.3) is proposed and transformed into the nuclear norm of block-diagonal matrix in the Fourier domain (see (4)). This results in the superior power of capturing the intrinsically "spatial-shifting" correlation [8]. They will be used in the formulation and the optimization of our proposed method. Table I lists some notations used in this paper.

TABLE I
Explanation of Notation in This Paper

$\mathcal{X}$	a tensor	$\|\mathbf{X}\|_*$	the matrix nuclear norm
$\mathbf{X}$	a matrix	$\langle \mathcal{X}, \mathcal{Y} \rangle$	$\mathbf{vec}(\mathcal{X})^T \mathbf{vec}(\mathcal{Y})$
$\mathbf{x}$	a vector	$\|\mathcal{X}\|_F$	$\sqrt{\langle \mathcal{X}, \mathcal{X} \rangle}$
$\mathcal{X}(i,j,t)$	the (i,j,t)-th entry of $\mathcal{X}$	$\|\mathcal{X}\|_1$	$\sum_{i,j,t} \|\mathcal{X}_{i,j,t}\|$
$\mathcal{X}(i,j,:)$	the tube fiber of $\mathcal{X}$	$\|\mathcal{X}\|_\infty$	$\max_{i,j,k} \|\mathcal{X}(i,j,k)\|$
$\mathcal{X}^{(t)}$	the t-th frontal slice of $\mathcal{X}$	$\hat{\mathcal{X}}$	$\mathbf{fft}(\mathcal{X}, [], 3)$
$\mathcal{X}^T$	The transpose of $\mathcal{X}$	$\|\mathcal{X}\|_\circledast$	the tensor nuclear norm

Definition 3.1: **(t-product)** Let $\mathcal{X} \in \mathbf{R}^{n_1 \times n_2 \times n_3}$ and $\mathcal{Y} \in \mathbf{R}^{n_2 \times n_4 \times n_3}$. Then the t-product $\mathcal{X} * \mathcal{Y}$ is an $n_1 \times n_4 \times n_3$ tensor,

$$\mathcal{X} * \mathcal{Y} = \mathbf{fold}(\mathbf{circ}(\mathcal{X}) * \mathbf{TenMat}(\mathcal{Y})), \tag{1}$$

where $\mathbf{circ}(\mathcal{X})$ is defined as an $n_1 n_3 \times n_2 n_3$ matrix

$$\mathbf{circ}(\mathcal{X}) = \begin{bmatrix} \mathcal{X}^{(1)} & \mathcal{X}^{(n_3)} & \cdots & \mathcal{X}^{(2)} \\ \mathcal{X}^{(2)} & \mathcal{X}^{(1)} & \cdots & \mathcal{X}^{(3)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{X}^{(n_3)} & \mathcal{X}^{(n_3-1)} & \cdots & \mathcal{X}^{(1)} \end{bmatrix}; \tag{2}$$

The **TenMat** and **fold** operations are defined as

$$\mathbf{TenMat}(\mathcal{X}) = \begin{bmatrix} \mathcal{X}^{(1)} \\ \mathcal{X}^{(2)} \\ \vdots \\ \mathcal{X}^{(n_3)} \end{bmatrix}, \quad \mathbf{fold}(\mathbf{TenMat}(\mathcal{X})) = \mathcal{X}. \tag{3}$$

Based on the above t-product (Definition 3.1), the tensor-SVD is defined as follows.

Definition 3.2: **(Tensor-SVD)** For $\mathcal{X} \in \mathbf{R}^{n_1 \times n_2 \times n_3}$, its tensor-SVD is given by

$$\mathcal{X} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T,$$

where $\mathcal{U} \in \mathbf{R}^{n_1 \times n_1 \times n_3}$ and $\mathcal{V} \in \mathbf{R}^{n_2 \times n_2 \times n_3}$ are orthogonal tensors, $\mathcal{S} \in \mathbf{R}^{n_1 \times n_2 \times n_3}$ is an f-diagonal tensor. Each frontal slice of $\mathcal{S}$ is a diagonal matrix. $*$ is the t-product.

Definition 3.3: **(Tensor nuclear norm)** The tensor nuclear norm (**TNN**) of a tensor $\mathcal{X} \in \mathbf{R}^{n_1 \times n_2 \times n_3}$, denoted as $\|\mathcal{X}\|_\circledast$, is defined as the sum of singular values of all frontal slices of $\hat{\mathcal{X}}$ and further transformed into the nuclear norm of block-diagonal matrix in the Fourier domain, i.e.,

$$\|\mathcal{X}\|_\circledast = \frac{1}{n_3} \sum_{i=1}^{n_3} \|\hat{\mathcal{X}}^{(i)}\|_* = \|\mathbf{bdiag}(\hat{\mathcal{X}})\|_*, \tag{4}$$

where $\mathbf{bdiag}(\hat{\mathcal{X}}) \in \mathbf{R}^{n_1 n_3 \times n_2 n_3}$ is a block-diagonal matrix of $\mathcal{X}^T$ in the Fourier domain,

$$\mathbf{bdiag}(\hat{\mathcal{X}}) = \begin{bmatrix} \hat{\mathcal{X}}^{(1)} & & & \\ & \hat{\mathcal{X}}^{(2)} & & \\ & & \ddots & \\ & & & \hat{\mathcal{X}}^{(n_3)} \end{bmatrix}. \tag{5}$$

IV. Tensor Nuclear Norm-Based Low-Rank Approximation With Total Variation Regularization

In this section, we first develop a Tensor nuclear norm-based Low Rank approximation with Total Variation regularization (TLR-TV) model that integrates the tensor nuclear norm, l_1-norm, and TV regularization in a unified framework. In TLR-TV, two extensions of the traditional TV regularization, i.e., HTV and SSTV are introduced to handle the tensor data. Sections IV-B and IV-C propose two algorithms by ADMM, in which our model is split into several easier sub-problems. All variables are updated by minimizing the objective function in a Gaussian-Seidel manner and thus the per-iteration cost is quite low.

A. TLR-TV

Many LRMA-and LRTA-based methods [1]–[3], [11] have achieved promising performance on image denoising. However, it may be insufficient that they consider only the low-rank prior and ignore the local structure information of image data. To overcome this issue, when recovering the tensor data $\mathcal{L}$ using LRTA, we also impose the TV constraint on $\mathcal{L}$ to preserve the local information. Thus, we develop a new model called the tensor nuclear norm-based low-rank approximation with total variation regularization, i.e.,

$$\begin{aligned} &\min_{\mathcal{L},\mathcal{S},\mathcal{G}} \ \alpha\|\mathcal{L}\|_{\circledast} + \mu\|\mathcal{L}\|_{TV} + \lambda\|\mathcal{S}\|_1 + \gamma\|\mathcal{G}\|_F^2 \\ &\text{s.t.} \quad \mathcal{L} + \mathcal{S} + \mathcal{G} = \mathcal{Y}, \end{aligned} \tag{6}$$

where $\alpha, \mu, \lambda, \gamma$ are trade-off parameters to balance all four terms. $\mathcal{Y}$ is the observed tensor data, such as the noisy color and multispectral images. $\mathcal{L}$ is the low-rank tensor term, which is the clean image data. $\mathcal{S}$ is the sparse error term denoting outliers and non-Gaussian noise, such as salt&pepper noise and $\mathcal{G}$ represents the Gaussian noise. $\|\mathcal{L}\|_{\circledast}$ denotes the tensor nuclear norm in Definition 3.3.

Compared with several state-of-the-art approaches, such as RPCA [1], HoRPCA [2], TRPCA [3], one advantage of our TLR-TV model is that TLR-TV not only employs the high correlation among the spectral mode (the tensor nuclear norm), but also encodes the spatial-spectral continuity (the TV norm) in a tensor framework. Besides, several existing matrix factorization or Tucker decomposition-based approaches [26] inevitably require to predefine the rank values of three unfolding matrices. They may overestimate or underestimate the rank of the recovered data since the true rank values of matrix and tensor data are unknown in real scenarios. Our TLR-TV uses the tensor nuclear norm to encode the low-rank prior by automatically determining the rank information and the superior power of capturing the "spatial-shifting" correlation [8]. This is the main advantage and difference of tensor nuclear norm compared to CP-decomposition-and Tucker-decomposition-based rank. Due to the aforementioned aspects, the proposed TLR-TV can well reconstruct clean images from their noisy ones.

In TRL-TV, the tensor nuclear norm $\|\mathcal{L}\|_{\circledast}$ is exploited to encode the global low-rank prior of tensor data. Then the remaining issue of TRL-TV is how to select a proper TV regularization to handle the tensor data. The TV regularization first introduced in [38] was used to remove Gaussian noise and preserve edges of images. This is due to the fact that TV could measure the difference between a pixel and its neighbors. There are two popular TV norms, l_1-based anisotropic TV and l_2-based isotropic TV. Here, we focus on the anisotropic TV norm for image denoising. Anisotropic TV norm is defined as

$$\|L\|_{TV} = \sum_{i,j} \|\mathcal{D}_h L\|_1 + \|\mathcal{D}_v L\|_1, L \in \mathbb{R}^{n_1 \times n_2} \tag{7}$$

where $\mathcal{D}_h$ and $\mathcal{D}_v$ are horizontal and vertical 2-D finite-difference operators.

Here, we introduce two novel TV norms, namely HTV and SSTV norms, to preserve the local piecewise continuity of tensor data. They can be considered as extensions of the traditional anisotropic TV norm (7). For a tensor data $\mathcal{L} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, HTV is defined as follows

$$\mathcal{L}_{HTV} = \sum_{t=1}^{n3} \|\mathcal{L}^{(t)}\|_{TV}, \tag{8}$$

and SSTV is calculated by

$$\mathcal{L}_{SSTV} = \sum_{i,j,t} \|\mathcal{D}_h \mathcal{L}\|_1 + \|\mathcal{D}_v \mathcal{L}\|_1 + \|\mathcal{D}_s \mathcal{L}\|_1. \tag{9}$$

where $\mathcal{D}_s$ is the 2-D finite-difference operator along the spectral or channel direction.

The HTV norm is in a band-by-band way, which means that the traditional TV norm is computed in spatial domain separately, while the SSTV norm considers the consistency among the both spatial and spectral dimensions.

Note that the first two terms of TLR-TV in (6), namely $\|\mathcal{L}\|_{\circledast}$ and $\|\mathcal{L}\|_{TV}$, are coupled, because they share the same variable $\mathcal{L}$. An intuitive interpretation is that we impose both global low-rankness and local piecewise priors into the tensor data to be recovered. Then, we employ the variable splitting technique [39] to solve the TLR-TV model (6) in the next subsection. To distinct our methods with two different TV norms, we use TLR-HTV and TLR-SSTV to stand for TLR-TV using HTV and SSTV, respectively.

B. Iterative Optimization Procedure for TLR-HTV

Introducing the HTV norm and an auxiliary variable $\mathcal{Z}$ to the TLR-TV model (6), our TLR-HTV model can be

reformulated as

$$\min_{\mathcal{L},\mathcal{Z},\mathcal{S},\mathcal{G}} \alpha\|\mathcal{L}\|_{\circledast} + \mu\|\mathcal{Z}\|_{HTV} + \lambda\|\mathcal{S}\|_1 + \gamma\|\mathcal{G}\|_F^2$$
$$\text{s.t.} \quad \mathcal{L}+\mathcal{S}+\mathcal{G}=\mathcal{Y},\ \mathcal{L}=\mathcal{Z}. \tag{10}$$

Due to the fact that the objective function of model (10) *w.r.t.* the four variables $\mathcal{L},\mathcal{Z},\mathcal{S},\mathcal{G}$ is separable, we solve the model (10) using the ADMM framework by minimizing the augmented Lagrangian function of model (10) which is

$$\begin{aligned} L_\rho(\mathcal{L},\mathcal{Z},\mathcal{S},\mathcal{G};\mathcal{W},\mathcal{E}) &= \alpha\|\mathcal{L}\|_{\circledast} \\ &+ \mu\|\mathcal{Z}\|_{HTV} + \lambda\|\mathcal{S}\|_1 + \gamma\|\mathcal{G}\|_F^2 \\ &+ \frac{\rho}{2}\left(\left\|\mathcal{L}+\mathcal{S}+\mathcal{G}-\mathcal{Y}+\frac{\mathcal{W}}{\rho}\right\|_F^2 + \left\|\mathcal{L}-\mathcal{Z}+\frac{\mathcal{E}}{\rho}\right\|_F^2\right), \end{aligned} \tag{11}$$

where $\mathcal{W}$ is the Lagrange multiplier (or dual variable) associated with constraint $\mathcal{L}+\mathcal{S}+\mathcal{G}=\mathcal{Y}$, and $\mathcal{E}$ is the Lagrange multiplier corresponding to constraint $\mathcal{L}=\mathcal{Z}$. $\rho>0$ is a positive penalty parameter.

Then, we try to find a saddle point of $L_\rho(\mathcal{L},\mathcal{S},\mathcal{Z},\mathcal{G};\mathcal{W},\mathcal{E})$, which is also the solution of the original problem (6). With simple manipulations, we have the following iterative scheme using the general framework of ADMM:

1) Update $\mathcal{L}$: Restoration: Fixing other variables except for $\mathcal{L}$ in (11), we can obtain the following sub-problem:

$$\begin{aligned} \mathcal{L}_{k+1} &= \arg\min_{\mathcal{L}} L_{\rho_k}(\mathcal{L},\mathcal{Z}_k,\mathcal{S}_k,\mathcal{G}_k;\mathcal{W}_k,\mathcal{E}_k); \\ &= \arg\min_{\mathcal{L}} \frac{\alpha}{\rho_k}\|\mathcal{L}\|_{\circledast} + \frac{1}{2}\|\mathcal{L}-\mathcal{D}_k\|_F^2, \end{aligned} \tag{12}$$

where $\mathcal{D}_k = \frac{\rho_k(\mathcal{Y}+\mathcal{Z}_k-\mathcal{S}_k-\mathcal{G}_k)-(\mathcal{W}_k+\mathcal{E}_k)}{2\rho_k}$.

Using equation (4), the above (12) can be transformed into the Fourier domain. It is equivalent to the following formula:

$$\hat{\mathcal{L}}_{k+1} = \arg\min_{\hat{\mathcal{L}}} \frac{\alpha}{\rho_k}\|\mathbf{bdiag}(\hat{\mathcal{L}})\|_* + \frac{1}{2}\|\hat{\mathcal{L}}-\hat{\mathcal{D}}_k\|_F^2; \tag{13}$$

Obviously, (13) can be separated into n_3 independent minimization problems where the t-th problem is

$$\begin{aligned} \hat{\mathcal{L}}_{k+1}^{(t)} &= \arg\min_{\hat{\mathcal{L}}^{(t)}} \frac{\alpha}{\rho_k}\left\|\hat{\mathcal{L}}^{(t)}\right\|_* + \frac{1}{2}\left\|\hat{\mathcal{L}}^{(t)}-\hat{\mathcal{D}}_k^{(t)}\right\|_F^2 \\ &= \Gamma_{\alpha/\rho_k}\left(\hat{\mathcal{D}}_k^{(t)}\right) \end{aligned} \tag{14}$$

where $t=1,2,\ldots,n_3$, and $\Gamma_\rho(\mathbf{X})$ is the matrix singular value thresholding operator [40]: $\Gamma_\rho(\mathbf{X}) := \mathbf{U}\text{diag}(\bar{\sigma})\mathbf{V}^T, \mathbf{X} = \mathbf{U}\text{diag}(\sigma)\mathbf{V}^T$ is the SVD of matrix $\mathbf{X}$ and $\bar{\sigma}=\max\{\sigma-\rho,0\}$. Due to the independence of the sub-problems, it can be efficiently computed in parallel.

2) Update $\mathcal{Z}$: Spatial and Spectral Smoothness Preservation: Fixing other variables except for $\mathcal{Z}$ in (11), we can obtain the following sub-problem:

$$\begin{aligned} \mathcal{Z}_{k+1} &= \arg\min_{\mathcal{Z}} L_{\rho_k}(\mathcal{L}_{k+1},\mathcal{Z},\mathcal{S}_k,\mathcal{G}_k;\mathcal{W}_k,\mathcal{E}_k); \\ &= \arg\min_{\mathcal{Z}} \frac{\mu}{\rho_k}\|\mathcal{Z}\|_{HTV} + \frac{1}{2}\|\mathcal{Z}-\mathcal{T}_k\|_F^2, \end{aligned} \tag{15}$$

where $\mathcal{T}_k = \mathcal{L}_{k+1} + \mathcal{E}_k/\rho_k$.

To solve (15), we can solve each frontal slice $\mathcal{Z}_{k+1}^{(t)}$ of $\mathcal{Z}_{k+1}$, $1\le t\le n_3$ separately,

$$\mathcal{Z}_{k+1}^{(t)} = \arg\min_{\mathcal{Z}^{(t)}} \frac{\mu}{\rho_k}\|\mathcal{Z}^{(t)}\|_{TV} + \frac{1}{2}\|\mathcal{Z}^{(t)}-\mathcal{T}_k^{(t)}\|_F^2. \tag{16}$$

Many efforts have been made to develop efficient and scalable algorithms for the TV problem. Here, we adopt the fast gradient-based algorithm [41] to solve the sub-problem (16).

3) Update $\mathcal{S}$: Impulse Noise Removal: Fixing other variables except for $\mathcal{S}$ in (11), we can obtain the following sub-problem:

$$\begin{aligned} \mathcal{S}_{k+1} &= \arg\min_{\mathcal{S}} L_{\rho_k}(\mathcal{L}_{k+1},\mathcal{Z}_{k+1},\mathcal{S},\mathcal{G}_k;\mathcal{W}_k,\mathcal{E}_k); \\ &= \arg\min_{\mathcal{S}} \frac{\lambda}{\rho_k}\|\mathcal{S}\|_1 + \frac{1}{2}\|\mathcal{S}-\mathcal{M}_k\|_F^2, \end{aligned} \tag{17}$$

where $\mathcal{M}_k = \mathcal{Y}-\mathcal{L}_{k+1}-\mathcal{G}_k-\mathcal{W}_k/\rho_k$. Then the closed-form solution of (17) can be obtained by resorting to the element-wise shrinkage operator [42], that is,

$$\mathcal{S}_{k+1} = \mathcal{S}_{\lambda/\rho_k}(\mathcal{M}_k). \tag{18}$$

And $\mathcal{S}_\rho(\mathbf{x}) := \text{sign}(\mathbf{x}) * \max\{|\mathbf{x}|-\rho,0\}$.

4) Update $\mathcal{G}$: Gaussian Noise Removal: Fixing other variables except for $\mathcal{G}$ in (11), we can obtain the following sub-problem:

$$\begin{aligned} \mathcal{G}_{k+1} &= \arg\min_{\mathcal{G}} L_{\rho_k}(\mathcal{L}_{k+1},\mathcal{Z}_{k+1},\mathcal{S}_{k+1},\mathcal{G};\mathcal{W}_k,\mathcal{E}_k); \\ &= \arg\min_{\mathcal{G}} \gamma\|\mathcal{G}\|_F^2 + \frac{\rho_k}{2}\|\mathcal{G}-\mathcal{N}_k\|_F^2, \end{aligned} \tag{19}$$

where $\mathcal{N}_k = \mathcal{Y}-\mathcal{L}_{k+1}-\mathcal{S}_{k+1}-\mathcal{W}_k/\rho_k$. This is a standard least squares regression problem with closed-form solution:

$$\mathcal{G}_{k+1} = (2\gamma+\rho_k)^{-1}(\rho_k\mathcal{N}_k). \tag{20}$$

5) Update $\mathcal{W},\mathcal{E},\rho$: Lagrangian Multipliers and Penalty Parameter: These variables are updated by the following equations:

$$\mathcal{W}_{k+1} = \mathcal{W}_k + \rho_k(\mathcal{L}_{k+1}+\mathcal{S}_{k+1}+\mathcal{G}_{k+1}-\mathcal{Y}); \tag{21}$$
$$\mathcal{E}_{k+1} = \mathcal{E}_k + \rho_k(\mathcal{L}_{k+1}-\mathcal{Z}_{k+1}); \tag{22}$$
$$\rho_{k+1} = \min\{\beta\rho_k,\rho_{\max}\}. \tag{23}$$

It is worth noting that we select $\beta>1$ to further facilitate the convergence speed [43].

For the stopping criterion, we select the following formulas:

$$chg\mathcal{Y} = \|\mathcal{L}_{k+1}+\mathcal{S}_{k+1}+\mathcal{G}_{k+1}-\mathcal{Y}\|_\infty; \tag{24}$$
$$chg\mathcal{M} = \|\mathcal{L}_{k+1}-\mathcal{Z}_{k+1}\|_\infty; \tag{25}$$
$$chg\mathcal{L} = \|\mathcal{L}_{k+1}-\mathcal{L}_k\|_\infty; \tag{26}$$
$$chg\mathcal{Z} = \|\mathcal{Z}_{k+1}-\mathcal{Z}_k\|_\infty; \tag{27}$$
$$chg\mathcal{S} = \|\mathcal{S}_{k+1}-\mathcal{S}_k\|_\infty; \tag{28}$$
$$chg\mathcal{G} = \|\mathcal{G}_{k+1}-\mathcal{G}_k\|_\infty. \tag{29}$$

where $\|\mathcal{L}\|_\infty = \max_{i,j,t}\mathcal{L}_{i,j,t}$. At the k-th iteration, we will terminate our TLR-HTV algorithm by the following final

criterion:

$$Chg = \max\{chg\mathcal{Y}, chg\mathcal{M}, chg\mathcal{L}, chg\mathcal{Z}, chg\mathcal{S}, chg\mathcal{G}\} \leq \epsilon, \tag{30}$$

where $\epsilon > 0$ is a given tolerance. In our experiments, we set $\epsilon = 10^{-5}$.

C. Iterative Optimization Procedure for TLR-SSTV

Introducing the SSTV norm and two auxiliary variables $\mathcal{Z}, \mathcal{X}$ to the TLR-TV model (6), our TLR-SSTV model can be written as

$$\begin{aligned} &\min_{\mathcal{L},\mathcal{Z},\mathcal{S},\mathcal{X},\mathcal{G}} \alpha\|\mathcal{L}\|_{\circledast} + \mu\|\mathcal{X}\|_1 + \lambda\|\mathcal{S}\|_1 + \gamma\|\mathcal{G}\|_F^2 \\ &\text{s.t.} \quad \mathcal{L} + \mathcal{S} + \mathcal{G} = \mathcal{Y},\ \mathcal{L} = \mathcal{Z},\ \mathcal{D}(\mathcal{Z}) = \mathcal{X}, \end{aligned} \tag{31}$$

where $\mathcal{D} = [\mathcal{D}_h; \mathcal{D}_v; \mathcal{D}_s]$ is the concatenated difference operation. Then, the augmented Lagrangian function of problem (31) is

$$\begin{aligned} L_\rho(\mathcal{L}, \mathcal{Z}, \mathcal{S}, \mathcal{X}, \mathcal{G}; \mathcal{W}, \mathcal{E}, \Pi) &= \alpha\|\mathcal{L}\|_{\circledast} + \mu\|\mathcal{X}\|_1 \\ &+ \lambda\|\mathcal{S}\|_1 + \gamma\|\mathcal{G}\|_F^2 + \frac{\rho}{2}\left(\left\|\mathcal{L} + \mathcal{S} + \mathcal{G} - \mathcal{Y} + \frac{\mathcal{W}}{\rho}\right\|_F^2\right. \\ &\left. + \left\|\mathcal{L} - \mathcal{Z} + \frac{\mathcal{E}}{\rho}\right\|_F^2 + \left\|\mathcal{D}(\mathcal{Z}) - \mathcal{X} + \frac{\Pi}{\rho}\right\|_F^2\right), \end{aligned} \tag{32}$$

where Π is the Lagrange multiplier associated with constraint $\mathcal{D}(\mathcal{Z}) = \mathcal{X}$. Then, the iterative procedures for TLR-SSTV are summarized as follows. Due to the page limitation, we give only the sub-problems of TLR-SSTV different from these of TLR-HTV.

1) Update $\mathcal{Z}$: Fixing other variables except for $\mathcal{Z}$ in (32), we can obtain the following sub-problem:

$$\begin{aligned} \mathcal{Z}_{k+1} &= \arg\min_{\mathcal{Z}} \frac{\rho_k}{2} \\ &\times \left(\left\|\mathcal{L}_{k+1} - \mathcal{Z} + \frac{\mathcal{E}_k}{\rho_k}\right\|_F^2 + \left\|\mathcal{D}(\mathcal{Z}) - \mathcal{X}_k + \frac{\Pi_k}{\rho_k}\right\|_F^2\right). \end{aligned} \tag{33}$$

It can be solved by the following linear system:

$$\rho_k(I + \mathcal{D}^*\mathcal{D})\mathcal{Z} = \rho_k(\mathcal{L}_{k+1} + \mathcal{D}^*(\mathcal{X}_k)) + \mathcal{E}_k - \mathcal{D}^*(\Pi_k), \tag{34}$$

where $\mathcal{D}^*$ denotes the adjoint operator of $\mathcal{D}$. Finally, by the 3D Fourier transform (fftn) and its inverse transform (ifftn), we can obtain the closed-form solution

$$\begin{aligned} \mathcal{Z}_{k+1} &= \text{ifftn} \\ &\times \left(\frac{\text{fftn}(\rho_k(\mathcal{L}_{k+1} + \mathcal{D}^*(\mathcal{X}_k)) + \mathcal{E}_k - \mathcal{D}^*(\Pi_k))}{\rho_k(\mathbf{1} + |\text{fftn}(\mathcal{D}_h)|^2 + |\text{fftn}(\mathcal{D}_v)|^2 + |\text{fftn}(\mathcal{D}_s)|^2)}\right). \end{aligned} \tag{35}$$

2) Update $\mathcal{X}$: Fixing other variables except $\mathcal{X}$ in (32), we can obtain the following sub-problem:

$$\mathcal{X}_{k+1} = \arg\min_{\mathcal{X}} \frac{\mu}{\rho_k}\|\mathcal{X}\|_1 + \frac{1}{2}\left\|\mathcal{X} - \left(\mathcal{D}(\mathcal{Z}_{k+1}) + \frac{\Pi_k}{\rho_k}\right)\right\|_F^2, \tag{36}$$

Algorithm 1: TLR-HTV.

Input: $\mathcal{Y}$, ϵ, α, $\lambda = 1/\sqrt{\max(n_1, n_2)n_3}$, μ, γ, $\beta = 1.1$.
Initialize: $\mathcal{L}_0$, $\mathcal{Z}_0$, $\mathcal{S}_0$, $\mathcal{G}_0$, $\mathcal{W}_0$, $\mathcal{E}_0$ initialized to $\mathbf{0}$, $\rho_0 = 10^{-4}$, $k = 0$.

```
while not converged do
    Compute D̂_k = fft(D_k, [], 3);
    Update L̂_{k+1} by
    for i = 1, 2, ..., n_3 do
        Compute L̂^{(i)}_{k+1} using (14);
    end for
    Compute L_{k+1} = ifft(L̂_{k+1}, [], 3);
    Update Z_{k+1}, S_{k+1}, G_{k+1}, W_{k+1}, E_{k+1}, and ρ_{k+1}
     by (15), (18), (20), (21), (22), (23), respectively;
    Check the convergence conditions
        Chg ≤ ε.
end while
```

Output: $\mathcal{L}_k$.

Algorithm 2: TLR-SSTV.

Input: $\mathcal{Y}$, ϵ, α, λ, μ, γ, $\beta = 1.1$.
Initialize: $\mathcal{L}_0$, $\mathcal{Z}_0$, $\mathcal{S}_0$, $\mathcal{X}_0$, $\mathcal{G}_0$, $\mathcal{W}_0$, $\mathcal{E}$, Π_0 initialized to $\mathbf{0}$, $\rho_0 = 10^{-4}$, $k = 0$.

```
while not converged do
    Compute D̂_k = fft(D_k, [], 3);
    Update L̂_{k+1} by
    for i = 1, 2, ..., n_3 do
        Compute L̂^{(i)}_{k+1} using (14);
    end for
    Compute L_{k+1} = ifft(L̂_{k+1}, [], 3);
    Update Z_{k+1}, S_{k+1}, X_{k+1}, G_{k+1}, W_{k+1}, E_{k+1},
     Π_{k+1}, and ρ_{k+1} by (35), (18), (37), (20), (21), (22),
     (38), (23), respectively;
    Check the convergence conditions
end while
```

Output: $\mathcal{L}_k$.

and its close-form solution is

$$\mathcal{X}_{k+1} = \mathcal{S}_{\mu/\rho_k}(\mathcal{D}(\mathcal{Z}_{k+1}) + \Pi_k/\rho_k). \tag{37}$$

3) Update Π: Π_{k+1} is updated by

$$\Pi_{k+1} = \Pi_k + \rho_k(\mathcal{D}(\mathcal{Z}_{k+1}) - \mathcal{X}_{k+1}). \tag{38}$$

Note that the updating rules of variables $\mathcal{S}_{k+1}$, $\mathcal{G}_{k+1}$, $\mathcal{W}_{k+1}$, $\mathcal{E}_{k+1}$, and ρ_{k+1} are the same with those of TLR-HTV.

Algorithms 1 and 2 show the pseudo codes of our specifically designed algorithms: TLR-HTV and TLR-SSTV, respectively.

D. Connection With Several Existing Works

In this section, we emphasize the connection and difference between our TLR-TV and several existing works [2]–[4], [21], [30], [32], [44] in the literature. The works in [2], [21] and [3] model the globally high correlation among different channels/bands using the tensor unfolding and the tensor nuclear

norm, respectively. However, their common shortcoming is the lack of preserving local information. Our methods incorporate the hyper total variation (HTV) and spatial-spectral total variation (SSTV) regularizations to characterize the local piecewise smoothness; The method in [32] has to manually predefine the rank values along three modes before running the algorithm. There are no off-the-shelf methods to set the rank values in a unified way. Our proposed methods can automatically determine the rank values. Besides, the corresponding sub-problem with respect to the low-rank term of LRMA and LRTA focus only on the real space while our sub-problem is in the real and Fourier spaces (see the sub-problem (13)). [30] proposed a novel tensor completion model using the weighted sum of the nuclear norm (SNN) [21] and total variation simultaneously. The weighted SNN (WSNN) is defined as the weighted sum of nuclear norm of all unfolding matrices. It solves the corresponding problem by the primal-dual splitting technique. Similarly, [44] uses the tensor unfolding among all modes and Tucker decomposition to enable the low-rank prior. However, our methods use the tensor nuclear norm which is different from Tucker decomposition [44] and (weighted) SNN [4], [21], [30]. The comparison between the tensor nuclear norm and the CP-and Tucker-decomposition-based rank is listed as follows:

- The CP rank is defined as the smallest number of rank-1 factors. However, CP decomposition faces several challenges: (1) the local minima and high computation burden due to the nonconvex optimization process; (2) heavy ill-posedness because the best rank prior is generally unknown in practice;
- The Tucker rank (or multi-rank) is a vector of ranks of all unfolding matrices. As a convex surrogate of Tucker rank, the sum of nuclear norm is defined as the sum of the nuclear norms of all unfolding matrices. However, the unfolding process, transforming the tensor data into matrices, may lose the inherent spatial-spectral correlation. Because different modes of tensor data may have various contributions to Tucker rank, the weighted SNN [21] was proposed. One challenging issue is how to choose a proper weight for each mode.
- The tensor nuclear norm of the tensor data $\mathcal{X}$ is defined as the sum of the singular values of all frontal slices of $\hat{\mathcal{X}}$, obtained by the fast Fourier transformation along the tube fibers of $\mathcal{X}$. Then, the tensor nuclear norm can be transformed into the nuclear norm of block-diagonal matrix in the Fourier domain. This results in the superior power of capturing the "spatial-shifting" correlation.

These differences motivated us to use the tensor nuclear norm to encode the low-rank prior.

V. Experimental Results

In this section, we conduct experiments on several challenging datasets to evaluate the performance of our algorithms. All experiments are run in MATLAB R2016a on a 64-bit personal computer with a E5-2609 1.90 GHz CPU and 16 GB memory. We apply our TLR-HTV and TLR-SSTV for removing noise from color and multispectral images, and analyze our models with respect to several important parameters.

TABLE II
Comparison of State-of-the-Art Methods and Their Properties

Methods	Categories	Low rank?	TV ?
SSTV [31]	Matrix	No	SSTV
LRTV [26]	Matrix	Mode-3 unfolding	HTV
IALM [29]	Matrix	Mode-3 unfolding	No
LRMR [9]	Matrix	Mode-3 unfolding	No
WSNN [21]	Tensor	Mode-1,2,3 unfolding	No
TRPCA [3]	Tensor	Tensor-SVD	No
TLR-HTV	Tensor	Tensor-SVD	HTV
TLR-SSTV	Tensor	Tensor-SVD	SSTV

A. Experimental Setting

Baseline approaches: Our TLR-HTV and TLR-SSTV are compared with several state-of-the-art approaches including: low-rank matrix approximation method: IALM [29]; low-rank tensor approximation methods: tensor robust principal component analysis (TRPCA) [3], weighted sum of nuclear norm (WSNN) [21]; TV-based methods: spatial-spectral total variation (SSTV) [31]. low-rank matrix approximation and TV-based method: LRTV [26]. We briefly describe these competing methods in Table II.

Datasets

We use four datasets, including two color image datasets: Berkeley segmentation dataset, Kodak PhotoCD dataset, a multispectral image dataset: Columbia Multispectral dataset, and one real noisy color image dataset. The statistics of these datasets are briefly summarized below:

- Berkeley Segmentation Dataset (BSD):[1] There are 300 clean color images of size $481 \times 321 \times 3$ in the whole dataset. In our experiments, we randomly select 60 images for the impulse noise removal task.
- Kodak PhotoCD Dataset (Kodak):[2] The whole dataset consists of 24 clean color images of size $512 \times 768 \times 3$. We select all images in our experiment.
- Columbia Multispectral Database (CAVE):[3] The CAVE dataset contains 32 clean multispectral images. Each of them has 31 bands of size 512×512. In our second experiment, we select two multispectral images: *Toy* and *Flowers* for the multispectral image denoising task.
- Real noisy color image:[4] There are 60 real noisy color images of different sizes in this dataset. Here, we choose two images: *Dog* and *Frog* of sizes 400×400 and 375×500 with red, blue, and green channels, respectively.

Quantitative assessment:

Except visual quality, we employ four quantitative quality indexes, including the peak signal-to-noise ratio (PSNR), structure similarity (SSIM), feature similarity (FSIM), erreur relative globale adimensionnelle de synthese (ERGAS) [45] to measure the reconstruction accuracies. In general, an image with higher PSNR, SSIM, and FSIM values, or lower ERGAS value is closer to the ground truth. For a given image I of size $m \times n$, PSNR

[1] https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/
[2] http://r0k.us/graphics/kodak/
[3] http://www1.cs.columbia.edu/CAVE/databases/multispectral/
[4] http://demo.ipol.im/demo/125/

TABLE III
QUANTITATIVE EVALUATION OF DIFFERENT DENOISING ALGORITHMS ON 4 BSD AND 4 KODAK COLOR IMAGES WITH 10%/20% IMPULSE NOISE

Methods	Indexes								
Noisy	PSNR	14.39/11.42	14.80/11.83	14.80/12.37	14.74/11.74	15.25/12.26	16.65/12.64	15.08/12.09	15.18/12.19
	SSIM	0.28/0.147	0.264/0.137	0.282/0.149	0.268/0.145	0.129/0.057	0.157/0.078	0.128/0.057	0.247/0.133
	FSIM	0.736/0.589	0.716/0.568	0.74/0.606	0.706/0.545	0.623/0.503	0.664/0.554	0.624/0.500	0.744/0.646
	ERGAS	554.6/780.76	509.02/716.73	385/543.62	464.92/656.79	454.22/640.05	369.49/522.42	452.35/638.19	433.13/610.80
SSTV [31]	PSNR	23.2/17.74	23.69/18.75	23.69/18.76	23.41/18.49	23.96/18.65	24.08/18.54	24.16/18.78	23.66/18.10
	SSIM	0.76/0.463	0.742/0.457	0.766/0.463	0.733/0.452	0.607/0.266	0.615/0.271	0.611/0.269	0.715/0.394
	FSIM	0.936/0.834	0.908/0.811	0.933/0.83	0.906/0.793	0.884/0.737	0.896/0.757	0.892/0.746	0.928/0.822
	ERGAS	200.68/376.07	180.75/319.28	135.85/260.64	174.85/302.25	166.16/305.96	139.93/264.72	158.78/294.60	163.20/309.25
IALM [29]	PSNR	24.51/23.31	25.35/24.04	25.35/24.41	24.93/24.03	28.51/27.76	27.55/26.76	28.04/27.41	22.70/22.17
	SSIM	0.833/0.757	0.822/0.754	0.783/0.718	0.803/0.744	0.862/0.837	0.875/0.845	0.910/0.887	0.783/0.732
	FSIM	0.886/0.858	0.881/0.853	0.861/0.836	0.875/0.851	0.919/0.908	0.925/0.910	0.959/0.950	0.903/0.886
	ERGAS	171.93/197.19	152.8/177.06	120.51/135.94	142.85/158.41	98.17/107.17	93.94/102.79	101.84/109.46	182.41/193.73
WSNN [21]	PSNR	27.24/25.8	28.02/26.51	28.02/26.59	26.9/25.85	31.97/30.42	31.81/29.98	31.65/30.22	26.30/24.68
	SSIM	0.920/0.868	0.904/0.849	0.893/0.837	0.883/0.832	0.937/0.847	0.949/0.864	0.957/0.887	0.904/0.735
	FSIM	0.938/0.913	0.929/0.903	0.919/0.892	0.918/0.894	0.967/0.954	0.975/0.960	0.984/0.974	0.964/0.936
	ERGAS	125.38/147.76	110.87/131.35	89.98/105.77	113.92/128.55	66.07/78.88	57.49/70.97	67.16/79.16	120.54/145.07
TRPCA [3]	PSNR	29.43/27.5	29.18/27.32	29.18/28.68	28.27/26.84	30.09/29.00	31.07/29.76	30.78/29.60	25.75/24.75
	SSIM	0.950/0.907	0.925/0.868	0.945/**0.900**	0.913/0.863	0.927/0.900	0.948/**0.922**	0.948/0.925	0.915/**0.873**
	FSIM	0.962/0.939	0.945/0.918	0.958/0.932	0.939/0.913	0.950/0.938	0.968/0.956	0.975/0.966	0.959/0.945
	ERGAS	97.22/121.38	95.79/118.35	66.18/83.18	97.23/114.74	82.54/93.72	62.69/72.84	74.20/85.02	128.30/143.80
TLR-HTV	PSNR	**32.79/29.92**	**32.79/30.18**	32.79/29.40	**31.76/29.37**	**34.36/33.05**	**34.38/32.72**	**36.09/34.38**	27.28/26.31
	SSIM	**0.955/0.917**	**0.938/0.896**	0.912/0.864	0.927/**0.886**	0.928/**0.914**	0.933/0.919	**0.959/0.947**	0.895/0.864
	FSIM	0.970/**0.951**	**0.967/0.947**	0.960/0.936	**0.958/0.939**	**0.980/0.974**	**0.985/0.978**	**0.992/0.988**	0.967/**0.958**
	ERGAS	**66.35/92.28**	**63.57/85.67**	55.63/76.59	**65.15/85.82**	**50.03/58.18**	**42.73/51.78**	**40.31/49.07**	108.15/120.75
TLR-SSTV	PSNR	32.03/28.92	31.57/28.01	**32.99/29.73**	30.67/27.93	32.23/30.68	33.66/31.72	33.49/31.75	**28.09/26.45**
	SSIM	**0.955**/0.876	0.937/0.810	**0.953**/0.876	**0.930**/0.813	**0.938**/0.874	**0.952**/0.902	0.950/0.899	**0.939**/0.849
	FSIM	**0.973**/0.942	0.962/0.918	**0.971/0.941**	0.957/0.913	0.968/0.956	0.981/0.970	0.983/0.974	**0.975/0.958**
	ERGAS	72.08/103.11	72.82/109.64	**50.64/73.66**	73.89/101.35	64.35/76.78	46.53/58.16	54.33/66.38	**98.11/118.35**

is defined as

$$PSNR = 20\log\Big(\frac{S_{peak}}{\sqrt{MSE}}\Big),$$

where S_{peak} is the peak pixel value of image I. $\hat{I}$ denotes the restored image of I. The mean square error (MSE) between I and $\hat{I}$ is equal to $\sum_{x=1}^{m}\sum_{y=1}^{n}[\hat{I}_i(x,y) - I_i(x,y)]^2/(mn)$. SSIM is defied as

$$SSIM = \frac{(2\mu_I\mu_{\hat{I}} + C_1)(2\sigma_{I,\hat{I}} + C_2)}{(\mu_I^2 + \mu_{\hat{I}}^2 + C_1)(\sigma_I^2 + \sigma_{\hat{I}}^2 + C_2)},$$

where μ_I and $\mu_{\hat{I}}$ denote the average values of image I and $\hat{I}$, respectively. Their corresponding variances are denoted as σ_I and $\sigma_{\hat{I}}$. $\sigma_{I,\hat{I}}$ represents the covariance of I and $\hat{I}$. $C_1 = (0.01 * L)^2$ and $C_2 = (0.03 * L)^2$. L is the dynamic range of the pixel values. FSIM is defined as

$$FSIM = \frac{\sum_{z\in\Omega} S_L(z) * PC_m(z)}{\sum_{z\in\Omega} PC_m(z)},$$

where Ω demotes the whole image spatial domain. The phase congruency for position z of image I is denoted as $PC_x(I)$, then $PC_m(z) = \max\{PC_I(z), PC_{\hat{I}}(z)\}$. $S_L(z)$ is the gradient magnitude for position z. ERGAS is defined as

$$ERGAS = 100\alpha\sqrt{\frac{1}{P}\sum_{i=1}^{P}\Big(\frac{MSE_i^2}{\mu_{\hat{I}i}^2}\Big)},$$

where α is the zoom factor. P is the number of spectral bands. For example, if I is a color image, $P = 3$.

B. Color Image Denoising

In this subsection, we aim to evaluate the performance of our methods for the impulse noise removal and investigate which of HTV and SSTV is more suitable for color image denoising. We use a third-order tensor $\mathcal{Y} \in R^{n_1\times n_2\times 3}$ to represent a color image of spatial size $n_1 \times n_2$ with red, blue and green channels. Many efforts [1], [10]–[12] have demonstrated that transforming each band of a multispectral image or each frame of a surveillance video into a vector and stacking each vectorized image as a column of matrix, the generated matrix is low-rank. However, this scheme is unsuitable for color image denoising. Therefore, we apply the IALM method to reconstruct each channel of a color image independently due to the fact that their top singular values dominate the main information [2], [3]. The regularization parameters λ of IALM and TRPCA are set to $1/\sqrt{n_{(1)}}$ and $1/\sqrt{n_{(1)}n3}$, respectively, where $n_{(1)} = \max\{n1, n2\}$. Here, we compared our methods with the WSNN algorithm, and the weights of three unfolding matrices for low-rank tensor approximation are set to $[15, 15, 1.5]$ as suggested in [3]. The default parameters of other competing methods are adopted. We randomly select 60 images from BSD dataset and all images from Kodak dataset for this experiment. 8 examples of selected images are shown in the first row of Table III. And the noisy color images are generated by adding with different ratios τ of the salt&pepper noise. Here, the ratios τ are set to 0.1 and 0.2, respectively.

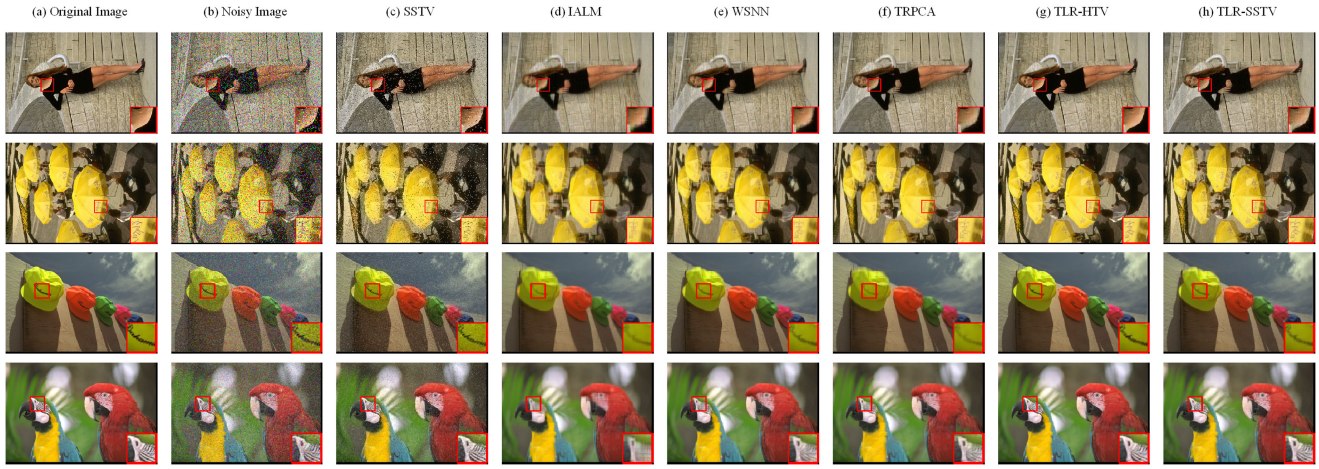

Fig. 2. Denoised images on 2 BSD images and 2 Kodak images by all competing methods (ratio $\tau = 0.2$). Columns from left to right: (a) Original image, (b) noisy image, the denoised image reconstructed by (c) SSTV [31], (d) IALM [29], (e) WSNN [21], (f) TRPCA [3], (g) TLR-HTV, and (h) TLR-SSTV. The figure is viewed better in zoomed PDF.

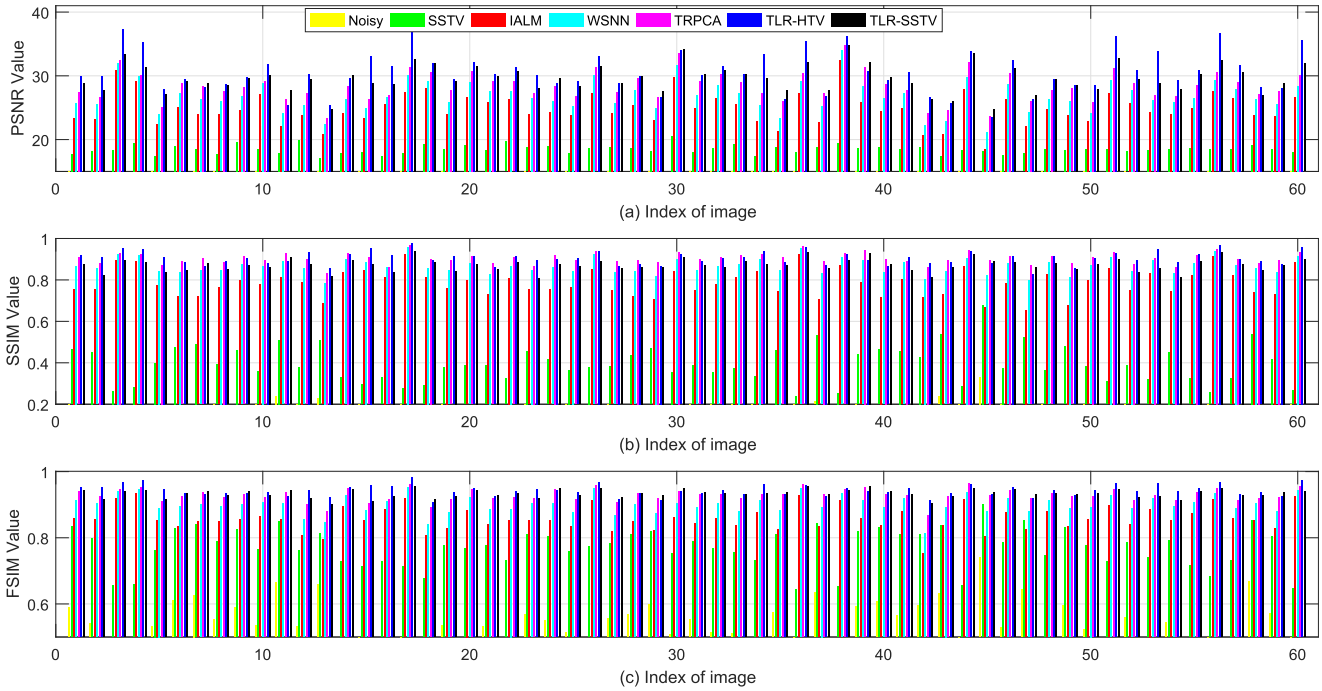

Fig. 3. Comparison of PSNR, SSIM, and FSIM results of different methods for color image denoising on 60 BSD images. (a) PSNR values, (b) SSIM values, and (c) FSIM values. The figure is viewed better in zoomed PDF.

We compute all quantitative quality indexes on 4 BSD and 4 Kodak color images as reported in Table III and all quantitative quality indexes on all selected color images are shown in Figs. 3 and 4. The best results are highlighted in hold and the second-best results of each quality index are underlined for clear comparison. And several typical recovered results are shown in Fig. 2.

From Table III and Figs. 2, 3, and 4, several observations can be drawn:

- For the color image denoising task, low-rank tensor approximation-based approaches perform better than low-rank matrix approximation-based approaches. For example, all quality indexes of WSNN and TRPCA are better than those of LRMA. The main reason is that more channel information of a color image is exploited in low-rank tensor approximation-based approaches. While IALM generates artifacts, because IALM treats three channels of a color image independently.
- Compared with the second-best denoising method TRPCA that also exploits the tensor nuclear norm to portray the low-rank structure of the tensor data, our proposed TLR-HTV and TLR-SSTV achieve average 2.18/1.68 dB and 1.24/1.55 dB improvement on all BSD/Kodak images, respectively, further verifying their advantages and superiority.

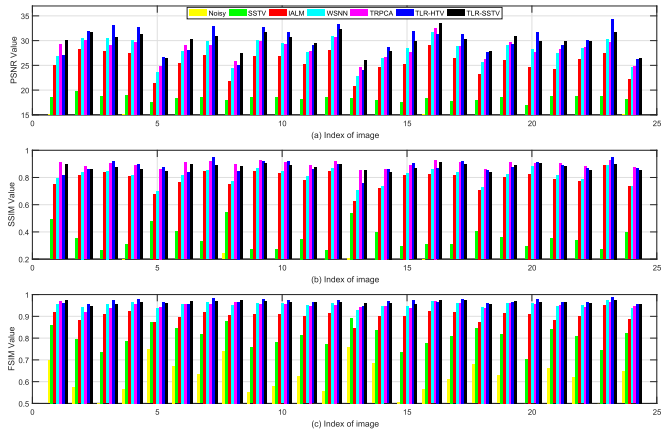

Fig. 4. Comparison of PSNR, SSIM, and FSIM results of different methods for color image denoising on 24 Kodak images. (a) PSNR values, (b) SSIM values, and (c) FSIM values. The figure is viewed better in zoomed PDF.

TABLE IV
AVERAGE PSNR (DB) RESULTS OF ALL DENOISING METHODS ON TWO TESTING COLOR IMAGE DATASETS

Methods	BSD Dataset		Kodak Dataset	
	$\tau = 0.1$	$\tau = 0.2$	$\tau = 0.1$	$\tau = 0.2$
Noisy	15.16	12.15	15.14	12.13
SSTV [31]	24.00	18.49	23.79	18.32
IALM [29]	25.87	24.93	26.21	25.43
WSNN [21]	28.05	26.86	30.34	28.05
TRPCA [3]	30.06	28.41	29.66	28.29
TLR-HTV	**33.02**	**30.60**	31.35	**29.97**
TLR-SSTV	32.36	29.65	**32.05**	29.84

- In general, our proposed methods outperform other competing methods. This is mainly due to the fact that our models exploit the global spectral high relationships between entries, and the local spatial or/and spectral smoothness. One can obtain that considering only the low-rank prior of a color image may cause distortion of local details. Specifically, IALM, WSNN, and TRPCA methods suffer from the hue bias issue (see the color of woman's Siamese skirt in the first row of Fig. 2).
- In almost cases, TLR-HTV achieves higher PSNR, SSIM and FSIM values than TLR-SSTV.

Table IV reports the average PSNR results of all denoising methods on two testing color image datasets. From this table, we can see that our TLR-HTV and TLR-SSTV achieve the highest PSNR values in all cases. TLR-HTV and TLR-SSTV outperform the best competing methods at least 1.01 dB and 1.24 dB, respectively. To further investigate the denoising performance of our methods, we change the ratio τ from 0.1 to 0.9 and compute the PSNR and SSIM values as shown in Fig. 5. It can be observed that our proposed TLR-HTV achieves higher PSNR and SSIM values in all cases over other competing methods, especially when τ is in the interval $[0.3, 0.7]$. Our TLR-SSTV outperforms other methods when τ is relatively small. Therefore, we can conclude that TLR-HTV is more robust than TLR-SSTV to the changes of ratios of impulse noise on color images.

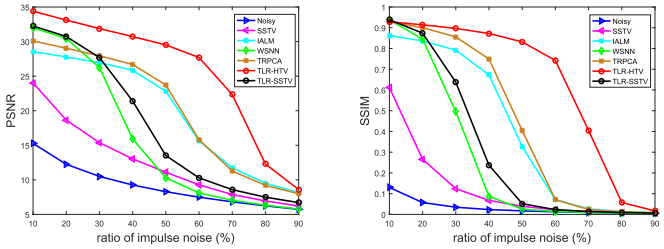

Fig. 5. Left: PSNR values with different ratios of impulse noise. Right: SSIM values with different ratios of impulse noise.

C. Multispectral Image Denoising

A multispectral image essentially can be viewed as a three-order tensor whose first two dimensions are spatial information and third dimension is band [2], [9]. In this subsection, two multispectral images, *Boy* and *Flowers* are selected to investigate the performance of our proposed methods. Before the denoising process, the gray values of each band were scaled to $[0, 1]$. To validate the robustness of our proposed methods for multispectral image denoising, two typical kinds of noises were added into the original clean multispectral images as described as follows:

- **Case1: Impulse noise** In the first denoising case, we added the salt&pepper noise with the ratio $\tau = 0.15$.
- **Case2: Mixed noise** Except for the salt&pepper noise in the first denoising case, the zero-mean Gaussian noise with variance 0.075 was added into all bands of two multispectral images.

Note that the IALM method is replaced by LRMR which is specially designed multispectral image denoising approach. And thus LRMR can be regarded as the high version of IALM for multispectral image denoising task. Another multispectral image mixed noise removal method: LRTV is a rank-constraint RPCA model with HTV regularization. As suggested in [26], we set the rank prior of mode-3 unfolding matrix $r = 10$ and the regularization parameter $\lambda = 20/\sqrt{n_{(1)}}$ for LRTV. We found the the weights $[15, 15, 1.5]$ of three unfolding matrices for low-rank tensor approximation are unsatisfactory for multispectral image denoising. Thus, we omit the WSNN method in this experiment.

All quantitative quality indexes on these two multispectral images in all cases are provided in Table V with the best results highlighted in bold and the second-best results underlined. Note that the mean values of each quality index of all bands are computed and denoted as MPSNR, MSSIM, MFSIM, and MERGAS, respectively. For the mixed noise removal case, our proposed TLR-SSTV achieves the highest MPSNR, MSSIM, and MFSIM values over all competing methods. LRTV yields the similar performance on both two multispectral images with our TLR-SSTV. However, LRTV performs bad for impulse noise removal. For impulse noise removal case, TRPCA obtains comparable or even slightly better results in terms of MPSNR and MSSIM on *Flowers* data over TLR-SSTV, but TRPCA fails to remove the mixed noise. In general, our proposed TLR-SSTV shows remarkable performance over TLR-HTV. The main reason is that the difference between consecutive bands of a multispectral image is relative small. So minimizing the SSTV regularization is more reasonable for multispectral image denoising.

TABLE V
QUANTITATIVE EVALUATION OF THE DIFFERENT DENOISING ALGORITHMS ON *Boy* AND *Flowers* FOR ALL CASES

Images	Indexes	Methods (case1/case2)						
		Noisy	SSTV [31]	LRMR [9]	LRTV [26]	TRPCA [3]	TLR-HTV	TLR-SSTV
Boy	MPSNR	12.30/12.09	41.40/32.73	22.85/29.41	36.03/32.51	42.25/28.30	31.60/29.60	**42.58/32.86**
	MSSIM	0.085/0.071	0.962/0.720	0.534/0.649	0.977/0.911	**0.989**/0.506	0.945/0.886	0.987/**0.915**
	MFSIM	0.481/0.464	0.990/0.959	0.841/0.937	0.990/0.958	**0.992**/0.902	0.961/0.918	**0.992/0.969**
	MERGAS	1002.50/1027.07	40.62/**97.28**	298.21/140.58	67.80/98.74	40.70/160.32	113.69/138.53	**38.94**/99.73
Flowers	MPSNR	12.16/11.97	44.07/32.92	25.91/28.64	43.75/36.15	**46.30**/28.97	38.53/33.78	45.77/**36.45**
	MSSIM	0.050/0.036	0.946/0.609	0.664/0.541	0.986/0.855	**0.988**/0.434	0.961/0.868	0.984/**0.872**
	MFSIM	0.431/0.415	**0.994**/0.952	0.872/0.913	0.993/0.943	0.990/0.887	0.970/0.906	0.990/**0.962**
	MERGAS	1512.93/1546.44	38.57/138.17	310.47/224.29	39.07/**93.47**	**38.19**/215.50	72.11/123.51	38.55/95.36

TABLE VI
QUANTITATIVE EVALUATION OF THE DIFFERENT DENOISING ALGORITHMS ON FOUR MULTISPECTRAL IMAGE DATASETS FOR MIXED NOISE REMOVAL

Images	Indexes	Methods (case2)						
		Noisy	SSTV [31]	LRMR [9]	LRTV [26]	TRPCA [3]	TLR-HTV	TLR-SSTV
CD_ms	MPSNR	11.97	32.04	29.82	33.13	27.57	35.49	**36.99**
	MSSIM	0.029	0.747	0.660	**0.946**	0.441	**0.946**	0.920
	MFSIM	0.359	0.924	0.899	0.943	0.848	0.929	**0.956**
Fake1	MPSNR	11.52	33.01	28.87	35.82	28.64	33.62	**35.93**
	MSSIM	0.033	0.687	0.594	**0.919**	0.460	0.914	0.891
	MFSIM	0.387	0.945	0.914	**0.958**	0.878	0.934	**0.958**
Fake2	MPSNR	11.87	33.34	29.18	36.25	28.99	34.01	**36.33**
	MSSIM	0.034	0.694	0.601	**0.921**	0.474	0.916	0.894
	MFSIM	0.392	0.947	0.917	**0.960**	0.882	0.936	0.958
Fake3	MPSNR	11.89	33.18	29.36	34.20	28.60	32.79	**35.13**
	MSSIM	0.038	0.714	0.632	**0.919**	0.479	0.914	0.906
	MFSIM	0.444	0.960	0.925	0.949	0.900	0.922	**0.965**

Fig. 7 shows the PSNR and SSIM values of all bands of *Flowers* with mixed noise. It is easy to see that our proposed TLR-SSTV achieves higher PSNR and SSIM values in most bands of *Flowers*, meaning that TLR-SSTV outperforms other methods for mixed noise removal. We also conducted experiments for mixed noise removal on four additional multispectral image datasets, including CD_ms, fake_and_beers (Fake1), fake_and_food (Fake2), fake_and_lemon (Fake3) as reported in Table VI. Our TLR-SSTV shows consistently superior performance over other competing methods with respect to MPSNR and MFSIM. Similar observations can be drawn from Fig. 6 which shows the visual comparison of all methods.

D. Real Noisy Color Image Denoising

In this subsection, we aim to test the Gaussian noise removal on two real noisy color images: Dog and Frog. These two real noisy images are mainly contaminated by Gaussian noise. Since there is no ground truth of the real noisy images, we can not compute the four above quantitative indexes and give only the visual quality comparison of different denoised real color images. Our TLR-HTV method is compared with a basic method: BM3D [46], one blind image denoising method: NC [47], and SSTV [31]. Note that the denoised results of NC are downloaded from the NC's webpage. Other methods are not compared here since they aim at removing the impulse noise. Figs. 8 and 9 show the denoised results on these two real color images by all competing approaches. One can observe that TLR-HTV can reconstruct more details while other methods either partially remove noise or oversmooth the results. This further demonstrates the superiority of the proposed TLR-HTV method.

E. Model Analysis

In this subsection, we aim to investigate how to tune parameters in our TLR-HTV. The parameter analysis of TLR-SSTV can be carried out in the similar way. Note that the experiments used for parameter setting are for the color image denoising. We select the Kodak color image: *Cap* as the test data as shown in the first row of Table III. There are four important parameters including α, μ, λ, and γ, in TLR-HTV (10). For simility, we set $\alpha + \mu + \lambda + \gamma = 1$ and $\lambda = 1/\sqrt{n_{(1)}n3}$, following TRPCA [3]. If α, μ, λ are fixed, γ is equal to $1 - \alpha - \mu - \lambda$. Thus, we need to tune only parameters α and μ.

1) Sensitivity Analysis of Parameters α and μ: In our model (10), α and μ are the parameters to balance the contributions of the low-rank term and TV regularization. These two terms dominate the performance of TLR-HTV. Therefore, there is a need to investigate how TLR-HTV is sensitive to α and μ. α and μ are selected from the set of $[0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]$ and $[0.001, 0.005, 0.01, 0.03, 0.05, 0.07, 0.1, 0.2, 0.3, 0.4]$, respectively. In Fig. 10, we show the PSNR and SSIM values of TRL-HTV with different combinations of α and μ. From Fig. 10, we

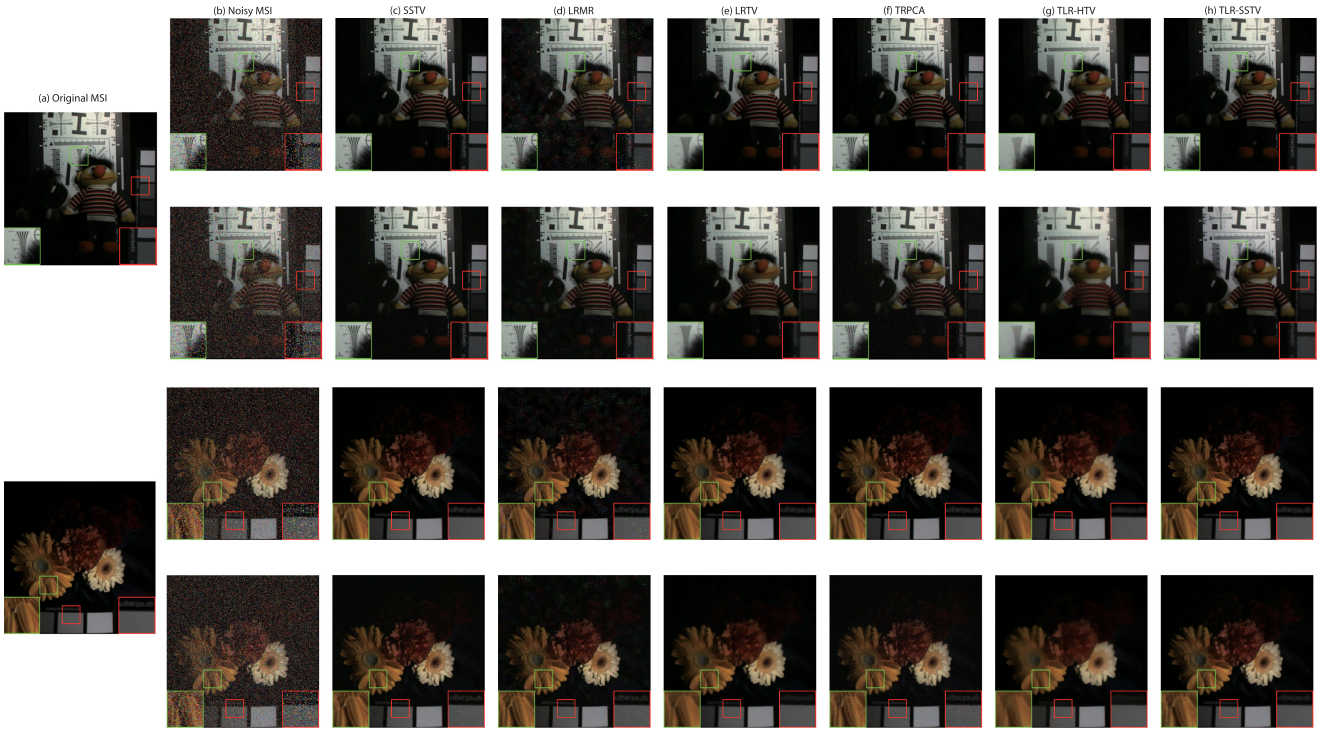

Fig. 6. Denoising results of images *Boy* (the first two rows) and *Flowers* (the last two rows) on CAVE dataset. The first and third rows are for Case 1 impulse noise removal while the second and last rows are for Case 2 mixed noise removal. (a) Original false-color image (R:25, G:20, B:16). (b) Noisy image. The denoised image reconstructed by (c) SSTV [31], (d) LRMR [9], (e) LRTV [26], (f) TRPCA [3], (g) TLR-HTV, and (h) TLR-SSTV. The figure is viewed better in zoomed PDF.

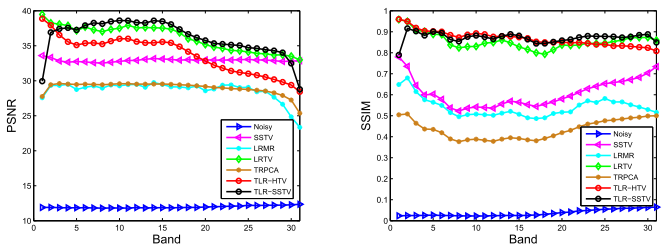

Fig. 7. Left: PSNR values of all bands of *Flowers* corrupted by mixed noise; Right: SSIM values of all bands of *Flowers* corrupted by mixed noise.

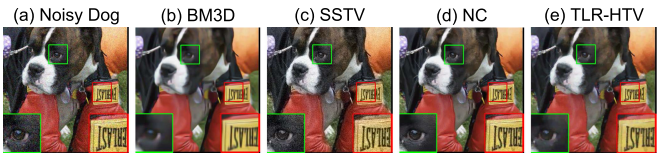

Fig. 8. Denoised images of the real noisy image "Dog." (a) Real noisy image: "Dog." The denoised image reconstructed by (b) BM3D [46], (c) SSTV [31], (d) NC [47], and (e) TLR-HTV. The figure is viewed better in zoomed PDF.

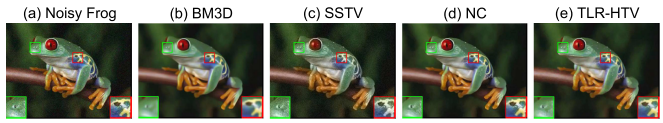

Fig. 9. Denoised images of the real noisy image "Frog." (a) Real noisy image: "Frog." The denoised image reconstructed by (b) BM3D [46], (c) SSTV [31], (d) NC [47], and (e) TLR-HTV. The figure is viewed better in zoomed PDF.

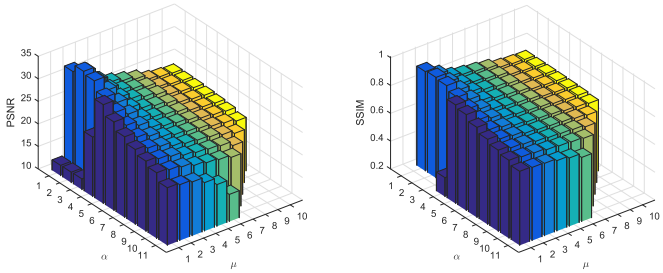

Fig. 10. PSNR and SSIM values of the TLR-HTV with different combinations of α and μ. Left: PSNR versus (α and μ); Right: SSIM versus (α and μ).

can see that with larger α and μ, TLR-HTV may cause the performance degradation. When $\alpha = 0.05, 0.1$ and $\mu = 0.005$, TLR-HTV achieves the highest PSNR and SSIM values. In our all experiments, α, μ, and λ of TLR-HTV are set to $0.1, 0.005, 1/\sqrt{n_{(1)}n3}$, and then γ is equal to $1 - \alpha - \mu - \lambda$.

2) Empirical Convergence of TLR-HTV: Fig. 11 shows the empirical analysis for the convergence of TLR-HTV. The first three subgraphs of Fig. 11 plot the PSNR, SSIM, and FSIM values versus the iteration number of the TLR-HTV. We can observe that after 60 iterations, these three values become almost a constant. The Change defined in (30) is exploited to illustrate the convergence rate as reported in the last subgraph of Fig. 11. After 20 iterations, the stopping criterion experiences a monotonic and rapid decline and stays stable after the iteration reaches 90.

3) Running Time of TLR-HTV: The detailed comparison of average running time of all denoising methods is reported in Table VII. As stated in (14), (13) can be separated into n_3

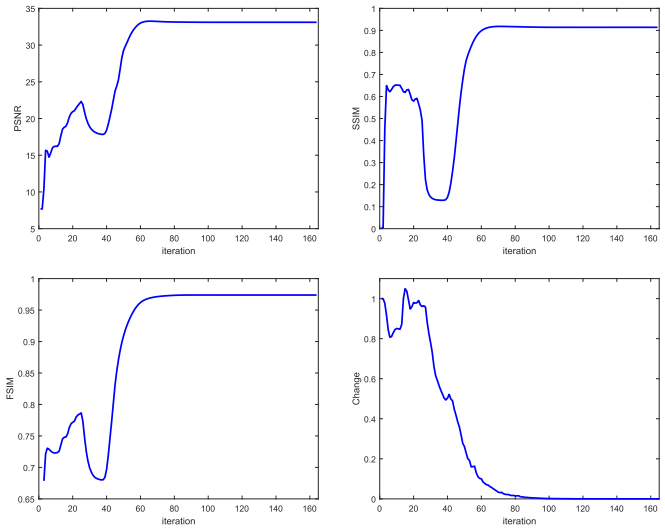

Fig. 11. PSNR, SSIM, FSIM, and change values versus the iteration number of the TLR-HTV. Upper Left: PSNR versus iteration. Upper right: SSIM versus iteration. Lower left: FSIM versus iteration. Lower right: Change (stopping criterion) versus iteration.

TABLE VII
COMPARISON OF AVERAGE RUNNING TIME (IN SECONDS) FOR THE COLOR IMAGE DENOISING ON ALL KODAK IMAGES

SSTV [31]	IALM [29]	WSNN [21]	TRPCA [3]	TLR-HTV
98.38	104.15	166.97	163.60	907.52

independent minimization problems, which means that (13) can be efficiently computed in parallel due to the independence of all sub-problems. However, we still update the n_3 independent minimization problems in a serial manner for fair comparison. From Table VII, one can see that SSTV is the fastest denoising method among all competing methods and IALM is the second-fastest method. WSNN and TRPCA have almost the same computation costs. Compared with IALM, WSNN consumes more time because WSNN has to compute the nuclear norms of two additional unfolding matrices. Unfortunately, our proposed TLR-HTV has more running time than all competing methods. The main computation burden of TLR-HTV is performing the tensor-SVD and the total variation constraint. It is natural because we impose the global low-rank prior and local piecewise continuity into the tensor data to be recovered. However, our TLR-HTV has achieved significant improvement. Therefore, it is worthy to achieve significantly better denoising performance by sacrificing computation efficiency.

VI. CONCLUSION

In this paper we proposed a new method for color and multispectral image denoising using the globally and locally intrinsic characteristics of the underlying clean image. More specifically, the low-rank structure of underlying data is depicted by the tensor nuclear norm under the tensor-SVD framework. To further improve the denoising performance, two TV regularizations for tensor data were exploited to preserve the local details of images. Extensive experiments on both simulated and real color images as well as multispectral images have demonstrated superior performance of the proposed methods over several state-of-the-art LRMA and LRTA-based denoising methods.

Several directions may be considered in our future work. Firstly, we consider to develop fast and scalable algorithms for our proposed model. One potential scheme is to utilize the tensor factorization technique [27] instead of tensor-SVD to capture the low-rank property. Secondly, our proposed method can be extended to tensor completion with applications to color image inpainting and video recovery.

REFERENCES

[1] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma, "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization," in *Proc. Neural Inf. Process. Syst.*, Dec. 2009, pp. 2080–2088.

[2] D. Goldfarb and Z. Qin, "Robust low-rank tensor recovery: Models and algorithms," *SIAM J. Matrix Anal. Appl.*, vol. 35, no. 1, pp. 225–253, 2014.

[3] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 5249–5257.

[4] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013, doi: 10.1109/TPAMI.2012.39.

[5] Y. Zhang, Z. Jiang, and L. S. Davis, "Learning structured low-rank representations for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 676–683.

[6] X. Shen and Y. Wu, "A unified approach to salient object detection via low rank matrix recovery," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 853–860.

[7] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, May 2011, Art. no. 11.

[8] M. E. Kilmer and C. D. Martin, "Factorization strategies for third-order tensors," *Linear Algebra Appl.*, vol. 435, no. 3, pp. 641–658, Aug. 2011.

[9] H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image restoration using low-rank matrix recovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4729–4743, Aug. 2014.

[10] Y. Xie, Y. Qu, D. Tao, W. Wu, Q. Yuan, and W. Zhang, "Hyperspectral image restoration via iteratively regularized weighted Schatten p-norm minimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4642–4659, Aug. 2016.

[11] Y. Chen, Y. Guo, Y. Wang, D. Wang, C. Peng, and G. He, "Denoising of hyperspectral images using nonconvex low rank matrix approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5366–5380, Sep. 2017.

[12] Z. Kang, C. Peng, and Q. Cheng, "Robust PCA via nonconvex rank approximation," in *Proc. IEEE Int. Conf. Data Mining*, 2015, pp. 211–220.

[13] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, "Weighted nuclear norm minimization and its applications to low level vision," *Int. J. Comput. Vis.*, vol. 121, no. 2, pp. 183–208, 2017.

[14] Y. Chen, Y. Wang, M. Li, and G. He, "Augmented Lagrangian alternating direction method for low-rank minimization via non-convex approximation," *Signal, Image Video Process.*, vol. 11, no. 7, pp. 1271–1278, 2017.

[15] X. Zhou, C. Yang, H. Zhao, and W. Yu, "Low-rank modeling and its applications in image analysis," *ACM Comput. Surv.*, vol. 47, no. 2, 2015, Art. no. 36.

[16] T. Bouwmans, A. Sobral, S. Javed, S. K. Jung, and E.-H. Zahzah, "Decomposition into low-rank plus additive matrices for background/foreground separation: A review for a comparative evaluation with a large-scale dataset," *Comput. Sci. Rev.*, vol. 23, pp. 1–71, 2017.

[17] F. Shang, Y. Liu, H. Tong, J. Cheng, and H. Cheng, "Robust bilinear factorization with missing and grossly corrupted observations," *Inf. Sci.*, vol. 307, pp. 53–72, 2015.

[18] Z. Wen, W. Yin, and Y. Zhang, "Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm," *Math. Program. Comput.*, vol. 4, no. 4, pp. 333–361, 2012.

[19] S. Wang, Y. Wang, Y. Chen, P. Pan, Z. Sun, and G. He, "Robust PCA using matrix factorization for background/foreground separation," *IEEE Access*, vol. 6, pp. 18 945–18 953, 2018.

[20] O. Semerci, N. Hao, M. E. Kilmer, and E. L. Miller, "Tensor-based formulation and nuclear norm regularization for multienergy computed tomography," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1678–1693, Apr. 2014.
[21] B. Huang, C. Mu, D. Goldfarb, and J. Wright, "Provable low-rank tensor recovery," *Optim. Online*, vol. 4252, 2014. [Online]. Available: http://www.optimization-online.org/DB_FILE/2014/02/4252.pdf
[22] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Sep. 2009.
[23] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.
[24] W. Hu, D. Tao, W. Zhang, Y. Xie, and Y. Yang, "The twist tensor nuclear norm for video completion," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 12, pp. 2961–2973, Dec. 2017.
[25] Z. Zhang, G. Ely, S. Aeron, N. Hao, and M. Kilmer, "Novel methods for multilinear data completion and de-noising based on tensor-SVD," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 3842–3849.
[26] W. He, H. Zhang, L. Zhang, and H. Shen, "Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 178–188, Jan. 2016.
[27] P. Zhou, C. Lu, Z. Lin, and C. Zhang, "Tensor factorization for low-rank tensor completion," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1152–1163, Mar. 2018.
[28] Y. Xu, R. Hao, W. Yin, and Z. Su, "Parallel matrix factorization for low-rank tensor completion," *Inverse Problems Imag.*, vol. 9, no. 2, pp. 18 945–18 953, 2015.
[29] Z. Lin, M. Chen, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," 2010, *arXiv:1009.5055*.
[30] T. Yokota and H. Hontani, "Simultaneous visual data completion and denoising based on tensor rank and total variation minimization and its primal-dual splitting algorithm," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 3732–3740.
[31] H. K. Aggarwal and A. Majumdar, "Hyperspectral image denoising using spatio-spectral total variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 442–446, Mar. 2016.
[32] Y. Wang, J. Peng, Q. Zhao, D. Meng, Y. Leung, and X.-L. Zhao, "Hyperspectral image restoration via total variation regularized low-rank tensor decomposition," 2017, *arXiv:1707.02477*.
[33] F. Shi, J. Cheng, L. Wang, P.-T. Yap, and D. Shen, "LRTV: MR image super-resolution with low-rank and total variation regularizations," *IEEE Trans. Med. Imag.*, vol. 34, no. 12, pp. 2459–2466, Dec. 2015.
[34] T. Zhou and D. Tao, "Godec: Randomized low-rank & sparse matrix decomposition in noisy case," in *Proc. Int. Conf. Mach. Learn.*, 2011, pp. 33–40.
[35] W. He, H. Zhang, L. Zhang, and H. Shen, "Hyperspectral image denoising via noise-adjusted iterative low-rank matrix approximation," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 8, no. 6, pp. 3050–3061, Jun. 2015.
[36] Q. Xie *et al.*, "Multispectral images denoising by intrinsic tensor sparsity regularization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1692–1700.
[37] C. Li, Y. Ma, J. Huang, X. Mei, and J. Ma, "Hyperspectral image denoising using the robust low-rank tensor recovery," *J. Opt. Soc. Amer. A*, vol. 32, no. 9, pp. 1604–1612, Sep. 2015.
[38] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1–4, pp. 259–268, 1992.
[39] R. Courant, "Variational methods for the solution of problems of equilibrium and vibrations," *Bull. Amer. Math. Soc.*, vol. 49, no. 1, pp. 1–23, 1943.
[40] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, Jan. 2010.
[41] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009.
[42] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *J. Roy. Statist. Soc. B*, vol. 58, pp. 267–288, 1996.
[43] S. Boyd *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
[44] X. Li, Y. Ye, and X. Xu, "Low-rank tensor completion with total variation for visual data inpainting," in *Proc. AAAI Conf. Artif. Intell.*, 2017, pp. 2210–2216.
[45] L. Wald, *Data Fusion. Definitions and Architectures: Fusion of Images of Different Spatial Resolutions*. Paris, France: Presses des MINES, 2002.
[46] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
[47] M. Lebrun, M. Colom, and J.-M. Morel, "Multiscale image blind denoising," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3149–3161, Oct. 2015.

Yongyong Chen received the B.S. and M.S. degrees from the College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao, China, and visited the National Key Lab for Novel Software Technology, Nanjing University, Nanjing, China, as an exchange student in 2017. He is currently working toward the Ph.D. degree with the Department of Computer and Information Science, University of Macau, Macau, China. His research interests include (non-convex) low-rank and sparse matrix/tensor decomposition models, with applications to image processing, data mining, and computer vision.

Shuqin Wang is currently working toward the M.S. degree with the College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao, China. Her current research interests include sparse representation, low-rank matrix/tensor approximation, and multispectral/hyperspectral image recovery.

Yicong Zhou (M'07–SM'14) received the B.S. degree from Hunan University, Changsha, China, and the M.S. and Ph.D. degrees from Tufts University, Medford, MA, USA, all in electrical engineering.

He is currently an Associate Professor and the Director with the Vision and Image Processing Laboratory, Department of Computer and Information Science, University of Macau, Macau, China. His research interests include chaotic systems, multimedia security, computer vision, and machine learning.

Dr. Zhou is a senior member of the International Society for Optical Engineering. He is currently an Associate Editor for *Neurocomputing*, *Journal of Visual Communication and Image Representation*, and *Signal Processing: Image Communication*. He is the Co-Chair of Technical Committee on Cognitive Computing in the IEEE Systems, Man, and Cybernetics Society. He was the recipient of the Third Price of Macau Natural Science Award in 2014.