**Weighted t-Schatten-p Norm Minimization for Real Color Image Denoising**

MIN LIU (Member, IEEE), XINGGAN ZHANG (Member, IEEE), AND LAN TANG (Member, IEEE)

School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China

Corresponding author: Lan Tang (tanglan@nju.edu.cn)

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant 61976113 and Grant U1936201.

**ABSTRACT** In this paper, to fully exploit the spatial and spectral correlation information, we present a new real color image denoising scheme using tensor Schatten-\( p \) norm (t-Schatten-\( p \) norm) minimization based on t-SVD to recover the underlying low-rank tensor. Similar to matrix Schatten-\( p \) norm, using non-convex t-Schatten-\( p \) (\( 0 < p < 1 \)) norm minimization could obtain better results than the tensor nuclear norm minimization which is a convex relaxation of the nonconvex tensor tubal rank. To avoid over-shrink the tensor tubal rank components, a flexible weighted t-Schatten-\( p \) norm model is proposed with weights assigned to different elements of tensor singular tubes. We adopt the generalized iterated shrinkage algorithm to solve the minimization problem efficiently. Extensive experiments on one synthetic and two realistic datasets demonstrate the effectiveness of our proposed method to remove noise both quantitatively and qualitatively.

**INDEX TERMS** Weighted tensor Schatten-\( p \) norm, low-rank, non-local patch group tensor.

**I. INTRODUCTION**

As a low-level image processing technique, image denoising is crucial to high-level computer vision tasks, for example, segmentation [32], [45], feature extraction, classification and so on. Since noise is ubiquitous in real-world data, a number of image denoising methods have been proposed to date. Most of the existing methods can be roughly categorized into total variation (TV) regularized methods [1], [30], [34], sparse representation methods [7], [8], [10], [11], [36], [37], [39]–[41], nonlocal self-similarity prior based methods [2], [7]–[9], [12], [35], [38]–[41], low-rank regularized methods [9], [12], [21], [35], [38], generative learning methods [23], [24], [29], [47] and discriminative learning methods [3], [6], [13], [42]. Considering color image denoising, one of the representative works is the color BM3D (CBM3D) [7], which first transforms the color image into a less correlated luminance-chrominance space and then applies BM3D in each channel of the transformed space. Xu et al. [36]–[38] perform denoising on the RGB channels simultaneously for better use of the spectral correlation. However, such matrix-based works break the intrinsic structure of the color image and inevitably suffer from loss of structural information and global correlation among color channels. To provide a faithful representation of structural properties for real color images, tensors provide an intuitive way to represent multidimensional data. As a generalization of matrix, tensor can be viewed as a multidimensional array. Real-world color images can be naturally regarded as a third-order tensor with column, row and color modes. There is an increasingly growing interest in recovery of low-rank tensor from the degraded observation, since low-rank tensor can better characterize the local and nonlocal structures of the grouped multidimensional data [27], [33], [46]. Apart from the above model-based methods, discriminative learning based denoising methods are proposed based on the advent of deep learning technique. Although deep learning [13], [42], [43] could powerfully extract features, it is difficult to train an exact network with insufficient training sets.

The low-rank tensor recovery is an effective way to exploit low-dimensional structure in high-dimensional data. However, the tensor rank has several definitions without tight convex relaxations based on different tensor decomposition algebraic frameworks. There are basically three types of tensor factorizations: CANDECOMP/PARAFAC (CP) factorization [4], Tucker factorization [31] and tensor singular value decomposition (t-SVD) [16]. Specifically, given an \( N \)-th order tensor \( \mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N} \), the CP decomposition model is the
sum of rank-one tensors \( A_{cp} = \sum_{r=1}^{R} u_r^{(1)} \circ u_r^{(2)} \circ \cdots \circ u_r^{(N)} \), where \( \circ \) is the outer product, and \( u_r^{(n)} \in \mathbb{R}^{I_n} \) is the \( r \)-th mode-\( n \) factor matrix. The tucker rank is defined as a vector, \( rank(A) := (rank(A^{(1)}), rank(A^{(2)}), \cdots, rank(A^{(N)})) \), where \( A^{(n)} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N} \) is the mode-\( n \) factor matrix. The Tucker model decomposes an \( n \)-th order tensor \( A \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N} \) as \( A = Z \times_1 U^{(1)} \times_2 U^{(2)} \times_3 \cdots \times_N U^{(N)} \), where \( x^n \) denotes the mode-\( n \) matrix product, \( Z \) is the core tensor, and \( U^{(n)} \in \mathbb{R}^{I_n \times R_n} \) is the mode-\( n \) factor matrix. The Tucker rank is defined as a vector, \( rank_k(A) := (rank(A^{(1)}), rank(A^{(2)}), \cdots, rank(A^{(N)})) \), where \( A^{(n)} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N} \) is the mode-\( n \) unfolding (matricization) of \( A \). \( U^{(n)} \) is composed by the orthogonal bases of \( A^{(n)} \). Motivated by that the nuclear norm is the convex relaxation of the matrix rank, some methods [19] adopt sum of matrix nuclear norm (SNN), i.e. \( \sum_n \| A^{(n)} \|_n^\star \) as a convex surrogate of the Tucker rank. Albeit the Tucker rank and its convex relaxation are tractable and widely used, its direct unfolding and folding operation tend to damage intrinsic structure information of the tensor and SNN is not the convex envelope of the Tucker rank according to [28].

Based on the tensor-tensor product (t-product), the t-SVD model decomposes a third-order tensor \( Y \in \mathbb{R}^{I_1 \times I_2 \times I_3} \) as \( Y = U \ast S \ast V^\ast \) similar to the matrix SVD. The tubal rank is defined as the number of nonzero singular tubes of \( S \) [15, 44]. To recover low tubal rank tensor, the tensor nuclear norm (TNN) minimization has achieved competitive performance in [22, 44], since t-SVD based methods are better in exploiting the intrinsic structure and correlation information compared with the matricization operation of the Tucker model. The tensor nuclear norm is defined as the sum of the matrix nuclear norm of each tensor frontal slices in Fourier space. As matrix nuclear norm minimization tends to over-shrink the rank components and treats each rank component equally, the tensor nuclear norm minimization has similar restrictions and the weighted tensor nuclear norm minimization is introduced in [20].

\[
\hat{X} = \arg \min_{X} \| Y - X \|_F^2 + \lambda \| X \|_{W,n}^\star. \tag{1}
\]

Inspired by the fact that using non-convex \( l_p \)-norm minimization (0 < \( p < 1 \)) could get better results than the convex \( l_1 \)-norm minimization [35], we propose the weighted t-Schatten-\( p \) norm minimization model to denoise real color images, which is flexible to deal with different rank components. In [35], WSNM has demonstrated superior denoising performance on grayscale images. However, such matrix-based works have to restructure the three-dimensional data into a matrix, which will lead to an information loss for a real color image. In this paper, we treat a real color image as a third-order tensor with each color channel corresponding to a frontal slice, and introduce low tubal rank minimization to exploit the spatial and spectral information by utilizing the nonlocal self-similarity and DFT (Discrete Fourier Transformation) along the color channel of the image. We first search similar patches for each local patch to form the nonlocal patch group tensor (NPGT) \( Y \). Then we adopt the weighted t-Schatten-\( p \)-norm as the low tubal rank tensor regularization and solve the minimization problem via the generalized iterated shrinkage algorithm. In summary, our contribution is three-fold. Firstly, we propose a weighted t-Schatten-\( p \) norm minimization for real color image denoising. Our method employs the low tubal rank tensor prior to model the spatial nonlocal self-similarity and spectral correlation information of real color images. The weighted t-Schatten-\( p \) norm is introduced as the surrogate functional of low tubal rank constraint to avoid over-shrink the tensor tubal rank components. Secondly, we get the optimal closed-form solution via the generalized iterated shrinkage algorithm. Finally, extensive experiments on both synthetic and realistic noisy image datasets validate that the proposed method could produce competitive performance compared to the state-of-the-art methods.

The rest of this paper is organized as follows. Section II briefly introduces some notations and preliminaries. Section III and IV gives the proposed weighted t-Schatten-\( p \) norm minimization model and the implementation details of the weighted t-Schatten-\( p \) norm minimization for real color image denoising. The experimental results, comparisons with the state-of-the-art algorithms and convergence analysis are reported in Section V. Section VI concludes this paper.

II. NOTATIONS AND PRELIMINARIES

Throughout the paper, we use lowercase letters, bold lowercase letters, bold uppercase letters and calligraphic script letters to denote scalars, vectors, matrices and tensors respectively. For a third-order tensor \( Y \in \mathbb{R}^{I_1 \times I_2 \times I_3} \), its \( i \)-th horizontal slice and frontal slice are denoted as \( Y(i, :, :) \) and \( Y(:, :, i) \) respectively. \( Y(i, j, :) \) and \( Y(:, j, :) \) are the \((i, j)\)-th tube and \((i, j)\)-th slice of tensor \( Y \) respectively. The \( i \)-th frontal slice of \( Y \) is denoted as \( Y^{(i)} \) for brevity. \( \sqrt{\sum_{i,j,k} |y_{ijk}|^2} \) is the Frobenius norm of tensor \( Y \) and is defined as \( \| Y \|_F = \sqrt{\sum_{i,j,k} |y_{ijk}|^2} \). The identity tensor \( I \in \mathbb{R}^{I_1 \times I_1} \) is the tensor whose first frontal slice is the \( I \times I \) identity matrix and other slices are zero matrices. We denote \([a]\) as a nearest integer greater than or equal to \( a \).

We use \( \hat{Y} \) to denote the result of Discrete Fourier Transform (DFT) of \( Y \) along the third dimension, i.e. \( \hat{Y} = \text{fft}(Y, [, 3]) \). And its inverse operator \( \hat{Y} = \text{ifft} \hat{Y}, [, 3] \) also holds. The block circulant matrix of \( Y \) is

\[
b\circ\circ(r,m) = \begin{bmatrix} Y^{(1)} & Y^{(1)} & \cdots & Y^{(2)} \\ Y^{(2)} & Y^{(1)} & \cdots & Y^{(3)} \\ \vdots & \vdots & \ddots & \vdots \\ Y^{(2)} & Y^{(1)} & \cdots & Y^{(1)} \end{bmatrix} \in \mathbb{R}^{I_1 \times I_2 \times I_3}.
\]

Definition 1 (t-Product [16]): Let \( A \in \mathbb{R}^{I_1 \times I_2 \times I_3} \) and \( B \in \mathbb{R}^{I_2 \times I_1 \times I_3} \), the t-product is defined as

\[
C = A \ast B = \text{fold}(b\circ\circ(A).u\text{nfold}(B)) \in \mathbb{R}^{I_1 \times I_2 \times I_3},
\]
where \( \text{unfold}(A) = \begin{bmatrix} A^{(1)}; A^{(2)}; \cdots; A^{(I_3)} \end{bmatrix} \), and 
\( \text{fold}(\text{unfold}(A)) = A \).

\( \tilde{A} \) is denoted as a block diagonal matrix, which is

\[
\tilde{A} = \begin{bmatrix}
A^{(1)} & & \\
& A^{(2)} & \\
& & \cdots \\
& & & A^{(I_3)}
\end{bmatrix} \in \mathbb{C}^{I_1 I_2 \times I_3 I_2}.
\] (4)

According to [22], t-product is equivalent to the familiar matrix multiplication in the Fourier domain, i.e., \( Y = A \ast B \) is equivalent to \( \bar{Y} = \tilde{A} \tilde{B} \) and the property \( \|Y\|_F^2 = \frac{1}{I_1} \|\bar{Y}\|_F^2 \) holds.

**Definition 2 (t-SVD [16], [22]):** For \( Y \in \mathbb{R}^{I_1 \times I_2 \times I_3} \), there exists

\[
\bar{Y} = \mathcal{U} \ast \bar{S} \ast \bar{V}^* \quad \|\bar{Y}\|_F \leq \|Y\|_F,
\] (5)

where \( \mathcal{U} \in \mathbb{R}^{I_1 \times r \times I_3} \), \( \mathcal{V} \in \mathbb{R}^{I_2 \times r \times I_3} \) are orthogonal tensors, \( \bar{V}^* \) is the conjugate transpose of \( \bar{V} \) obtained by conjugate transposing each frontal slice of \( \bar{Y} \) and then reversing the order of transposed frontal slices \( 2 \) through \( I_3 \) as

\[
\text{fold}[\mathcal{V}^{(1)}]; \mathcal{V}^{(1)}; \mathcal{V}^{(I_3-1)}; \cdots; \mathcal{V}^{(I_3)}],
\] and \( \bar{S} \in \mathbb{R}^{r \times r \times I_3} \) is an \( r \)-diagonal tensor with each frontal slice as a diagonal matrix.

The t-SVD computation is detailed in Algorithm 1.

**Algorithm 1 t-SVD [22]**

**Input:** \( Y \in \mathbb{R}^{I_1 \times I_2 \times I_3} \).
**Output:** \( \mathcal{U}, \bar{S}, \mathcal{V} \).

1. Compute \( \bar{Y} = \text{fft}(Y, [1], 3) \).
2. Compute SVD to each frontal slice of \( \bar{Y} \):
   3. for \( k = 1 \) to \( \left\lceil \frac{I_1 + 1}{2} \right\rceil \) do
     4. \( \bar{V}^{(k)} = \text{Ifft}(\bar{S}^{(k)} \bar{V}^{(k)*}) \);
     5. end for
   6. for \( k = \left\lceil \frac{I_1 + 1}{2} \right\rceil + 1 \) to \( I_3 \) do
     7. \( \bar{V}^{(k)} = \text{Ifft}(\bar{S}^{(k-I_1+k+2)}) \);
     8. end for
11. Compute \( \mathcal{U} = \text{Ifft}(\tilde{U}, [1], 3) \), \( \bar{S} = \text{Ifft}(\bar{S}, [1], 3) \) and \( \mathcal{V} = \text{Ifft}(\bar{V}, [1], 3) \).

**III. WEIGHTED T-SCHATTEN-P NORM MINIMIZATION**

The proposed weighted t-Schatten-\( p \) norm of \( \mathcal{X} \) is defined as

\[
\|\mathcal{X}\|_{W,sp} = \left( \frac{1}{I_1} \sum_{k=1}^{I_1} \sum_{i=1}^{\text{min}(I_1, I_2)} \omega_i \|\mathcal{S}_{ik}\|^p \right)^{\frac{1}{p}},
\] (7)

where \( \mathcal{X} = \mathcal{U} \ast \bar{S} \ast \mathcal{V}^* \), and \( \mathcal{W}^{(k)} = \text{diag}(\omega_1; \cdots; \omega_r) \) \( \in \mathbb{R}^{r \times r} \) is a non-negative diagonal matrix. Thus the weighted tensor \( \mathcal{V} \) is a \( r \times r \times I_3 \) tensor, and the the weighted t-Schatten-\( p \) norm with power \( p \) of a tensor \( \mathcal{X} \) is

\[
\|\mathcal{X}\|_{W,sp}^p = \frac{1}{I_1} \sum_{k=1}^{I_1} tr(\mathcal{W}^{(k)} \bar{S}^{(k)p}),
\] (8)

where \( \mathcal{W}^{(k)} \) and \( \bar{S}^{(k)} \) are both diagonal matrices.

Given a tensor \( \mathcal{V} \), the low tubal rank tensor approximation model regularized by the weighted t-Schatten-\( p \) norm is formulated as

\[
\hat{\mathcal{X}} = \arg \min_{\mathcal{X}} \|\mathcal{Y} - \mathcal{X}\|_F^2 + \lambda \|\mathcal{X}\|_{W,sp}^p,
\] (9)

where \( \lambda \) is the compromise parameter to balance the data fidelity term and the regularization term.

By utilizing properties (8) and \( \|\mathcal{Y}\|_F^2 = \frac{1}{I_1} \|\bar{Y}\|_F^2 \), the minimization problem (9) is equivalent to

\[
\hat{\mathcal{X}} = \arg \min_{\mathcal{X}} \frac{1}{I_1} \sum_{k=1}^{I_1} \left( \|ar{Y}^{(k)} - \bar{X}^{(k)}\|_F^2 + \lambda \|\bar{X}^{(k)}\|_{W^{(k)}, sp}^p \right).
\] (10)

The k-th frontal slice of \( \hat{\mathcal{X}} \) can be solved via the generalized iterated shrinkage algorithm [35], [48] illustrated in Algorithm 2 and Algorithm 3.

**Algorithm 2 Generalized Iterated Soft-Thresholding Algorithm (GST) [35]**

**Input:** \( s, \omega, p, J \).
**Output:** \( T_{GST}^p (s; \omega) \).

1. \( \tau_{GST}^p (\omega) = (2\omega(1 - p))^{\frac{1}{2p}} + \omega p(2\omega(1 - p))^{\frac{p-1}{2p}} \);
2. if \( |s| \leq \tau_{GST}^p (\omega) \) then
   4. \( T_{GST}^p (s; \omega) = 0 \);
   else
   5. \( j = 0, s^{(k)} = |s| \);
   6. for \( j = 0, 1, \cdots, J \) do
     7. \( s^{(j+1)} = |s| - \omega p(s^{(j)})^{p-1} \);
     8. \( j = j + 1 \);
   end for
10. \( T_{GST}^p (s; \omega) = \text{sgn}(s) s^{(k)} \).
11. end if
Algorithm 3 t-Schatten-p Norm Minimization

Input: \( Y \in \mathbb{R}^{H \times W \times C} \)

Output: \( \hat{X} \)

1. Compute \( \hat{Y} = \text{fft}(Y, [1, 2]) \).
2. Compute SVD to each frontal slice of \( \hat{Y} \):
3. for \( k = 1 \) to \( \frac{H-1}{2} \) do
4. \( \bar{S}^{(k)} = \hat{U}^{(k)} \hat{S}^{(k)} \bar{V}^{(k)} \) and \( \bar{S}^{(k)} = \text{diag}(s_1, \ldots, s_r) \);
5. for \( i = 1, \ldots, r \) do
6. \( \delta_i = \text{GST}(s_i, \omega_i, p) \)
7. end for
8. \( \Delta^{(k)} = \text{diag}(\delta_1, \ldots, \delta_r) \)
9. \( \hat{X}^{(k)} = \hat{U}^{(k)} \Delta^{(k)} \bar{V}^{(k)} \)
10. end for
11. for \( k = \left( \frac{H+1}{2} \right) \) to \( I_3 \) do
12. \( \hat{X}^{(k)} = \text{conf}((\hat{X}^{(k-l)-(k+2)}) \)
13. end for
14. Compute \( \hat{X} = \text{ifft}(\hat{X}, [1, 2]) \)

IV. APPLYING T-SCHATTEN-P NORM MINIMIZATION TO REAL COLOR IMAGE DENOISING

We treat a real color image as a third-order tensor with each color channel corresponding to a frontal slice, and introduce weighted t-Schatten-p norm minimization to real color image denoising exploiting the spatial and spectral information of the image. For an observed noisy color image \( Y \in \mathbb{R}^{H \times W \times C} \), we separate it into a set of image patches \( \{ P_i \in \mathbb{R}^{p \times p \times C} \} \), where \( H, W \) are spatial height and width, \( C = 3 \) is color channel, \( p \) is the patch size and \( M = (H - p + 1) \times (W - p + 1) \) is total number of patches over the whole image. For each local patch, we search its \( N \) similar patches over the whole image to form a 4D group of \( N \) patches \( G \in \mathbb{R}^{p \times p \times C \times N} \). Then we reshape \( G \) into a third-order tensor \( X \in \mathbb{R}^{p^2 \times N \times C} \) by vectorising its spatial frontal slices, and both the nonlocal similarity and spectral correlation information are well preserved by such representation.

To estimate the underlying clean nonlocal similar patch group tensor \( \hat{X} \) from its noisy observation \( \hat{Y} \), we solve the following optimization problem.

\[
\hat{X} = \arg \min_{\hat{X}} \| \hat{Y} - \hat{X} \|_F^2 + \lambda \| \hat{X} \|_{p,\gamma,S_p}^p.
\]

Following the analysis in Section III, each frontal slice of \( \hat{X} \) could be computed by

\[
\arg \min_{\hat{X}^{(k)}} \frac{1}{\sigma^2} \left( \| \hat{Y}^{(k)} - \hat{X}^{(k)} \|_F^2 + \| \hat{X}^{(k)} \|_{W^{(k)},S_p}^p \right),
\]

where \( \sigma^2 \) is the noise variance, the first term in (12) represents the data conformity term, and the second term indicates the weighted t-Schatten-p norm of low rank regularization. Similar to [35], the diagonal elements of \( W^{(k)} \) are empirically set as

\[
\omega_i = \frac{c \sqrt{N}}{\delta_i^{1/p} + \epsilon},
\]

where \( c \) is a constant, \( N \) is the number of similar patches, \( \delta_i \) is the \( i \)-th singular value of \( \hat{X}^{(k)} \), \( \epsilon \) is a small positive number to avoid dividing by zero. \( \delta_i \) can be initialized as

\[
\delta_i = \sqrt{\max \left( \frac{\| Y_i \|_2^2}{N - \sigma_i^2}, 0 \right)},
\]

The optimization problem (11) could be solved via Algorithm 3. The underlying clean image \( \hat{X} \) could be recovered by aggregating all the estimated \( \hat{X}_i \). To restore clean image iteratively, we utilize the iterative regularization method in [9], [26] to enhance the denoising performance. The central idea of the iterative regularization is to add filtered residual back to the denoised image.

\[
\hat{Y}_l = \hat{X}_{l-1} + \alpha(\hat{Y} - \hat{X}_{l-1}),
\]

where \( \alpha \) is a relaxation parameter. The proposed t-Schatten-p norm minimization for real color image denoising is detailed in Algorithm 4. The estimated standard deviation of noise is monotonically decrease and the underlying clean image is progressively recovered until convergence as the iterations progress.

Algorithm 4 The Proposed Algorithm for Real Color Image Denoising

Input: real color image \( Y \)

Output: denoised image \( \hat{X} \)

1. Initialization: \( \hat{X}^0 = Y \)
2. for \( t = 1 \) to \( T \) do
3. \( \hat{Y}^t = \hat{X}^{t-1} + \alpha(\hat{Y} - \hat{X}^{t-1}) \)
4. Extract local patches \( \{ Y_i^t \}_{1 \leq i \leq L} \) from \( Y^t \)
5. for each patch \( Y_i^t \) do
6. Search nonlocal similar patches to form \( Y_i^t \)
7. Obtain the estimated \( \hat{X}_i^t \) via the weighted t-Schatten-p norm minimization
8. end for
9. Aggregate \( \hat{X}_i^t \) to form the color image \( \hat{X}^t \)
10. end for

V. EXPERIMENTAL RESULTS

In this section, we first evaluate the denoising performance of the proposed method on one synthetic and two public realistic image datasets, and then we compare it with nine representative methods, including color block-matching 3D filtering (CBM3D) [7], multi-channel weighted nuclear norm minimization (MCWNNM) [38], guided image denoising (GID) [36], trilateral weighted sparse coding (TWSC) [37], multi-channel weighted Schatten-p norm minimization (MCWSNM) combining WSNM in [35] and MCWNNM in [38] for color images denoising, color multispectral t-SVD (CMSst-SVD) [18], denoising convolutional neural networks (DnCNN) [42], fast and flexible denoising network (FFDNet) [43] and weighted tensor nuclear norm minimization (WTNNM) [20]. All the denoising results of
TABLE 1. PSNR/SSIM results of competing denoising algorithms on 24 synthetic noisy color images. 

| Method       | PSNR | SSIM |
|--------------|------|------|
| CBM3D        | 58.64| 0.7149|
| MCWNNM       | 58.64| 0.7149|
| GID          | 58.64| 0.7149|
| MCWNNM       | 58.64| 0.7149|
| CTSVM        | 58.64| 0.7149|
| EnCNN        | 58.64| 0.7149|
| PSNet        | 58.64| 0.7149|
| WTNNM        | 58.64| 0.7149|

A. EXPERIMENTAL RESULTS ON SYNTHETIC DATASET

The Kodak dataset (http://r0k.us/graphics/kodak) which includes 24 lossless color images is utilized in our simulated experiments. The simulated noisy images are generated by adding AWGN with \( \sigma = 40 \), \( \sigma = 20 \), \( \sigma = 20 \) to each color channel independently.

We empirically set patch size to 8 \( \times \) 8 and the number of nonlocal similar patches as \( N = 90 \). The involved parameters \( \alpha \), \( c \), \( p \) are set as 0.1, \( 2\sqrt{2} \), and 0.95 respectively.

The current approach is compared with the state-of-the-art methods in terms of PSNR and SSIM indices. The experimental results are listed in Table 1. It can be observed that our method achieves better performance in terms of PSNR and SSIM.

B. EXPERIMENTAL RESULTS ON REAL-WORLD DATASETS

We adopt two publicly available realistic image datasets to evaluate our proposed method. The CC15 dataset includes 15 noisy and ground truth pair images (with size 512 \( \times \) 512 \( \times \) 3) of 11 static scenes, provided in [25] as shown in Figure 2. The CC60 dataset comprises 60 pairs of images of size 500 \( \times \) 500 \( \times \) 3, cropped in [36]. The images in CC15 dataset and CC60 dataset are from different shots. The standard deviation of noise has been given as a parameter for synthetic noisy images dataset. However, when it comes to real color image datasets, the noise level of each color channel is assumed to be Gaussian and could be estimated by the noise estimation method in [5].

The PSNR and SSIM results of CC15 and CC60 datasets are tabulated in Table 3 and 4, and the best results are highlighted in bold. In Table 3, the PSNR and SSIM results of CC15 are categorized into five groups based on different camera settings, and each group comprises three different scenes numbered as 1, 2, and 3. The average PSNR and SSIM results of competing methods on the holistic datasets (CC15 and CC60) are also listed in Table 3 and 4. It can be seen that the proposed method achieves better average PSNR and SSIM results than other competing methods in both datasets.

The visual performance of the comparative denoising methods are shown in Figure 3 and Figure 4. Figure 3 shows the denoised images of the real noisy image with rich fur texture captured by Nikon D800 with ISO = 6400 from CC15 dataset by different methods. It can be observed that our method can achieve better performance in terms of PSNR and SSIM.
our proposed method is effective in removing noise as well as preserving the fur texture, while other competing methods remain some noise in their corresponding denoised images. In comparison, our proposed method could remove the noise without generating much noise. The denoised performance of the realistic image from CC60 dataset in Figure 4 demonstrates that CBM3D, MCWNNM, GID, TWSC, CMS-SVD and FFDNet are likely to generate some color artifacts, while WTNNM and our proposed method are effective in removing noise. It is obvious that our proposed method could yield visually pleasant results.

C. CONVERGENCE ANALYSIS

Since the proposed t-Schatten-\( p \) norm minimization model is non-convex, it is difficult to prove the model’s global convergence theoretically. Here we provide quantitative results to demonstrate the convergence of the proposed method. Figure 5 shows the curves of the PSNR indexes versus iteration numbers about two images from the Kodak Dataset. It can be observed that the PSNR values gradually increase and become flat afterwards as the iteration numbers goes on, illustrating the convergence of our proposed method empirically.

D. COMPARISON OF COMPLEXITY ANALYSIS

We assume that the size of the extracted patch is \( p \times p \), the number of the similar patches within a group is \( N \), and the number of the patch group is \( M \). The main computation cost of the proposed method concentrates on computing FFT,
TABLE 3. PSNR/SSIM results of competing denoising algorithms on CC15 dataset [25].

| Method    | PSNR | SSIM |
|-----------|------|------|
| CBM3D     | 15.85| 0.83 |
| M2CSNMM   | 16.04| 0.85 |
| GID       | 15.68| 0.81 |
| MCWSNMM   | 16.10| 0.86 |
| TWSC      | 15.75| 0.82 |
| CMS-SVD   | 16.27| 0.87 |
| DnCNN     | 15.98| 0.84 |
| PFDDNet   | 16.09| 0.85 |
| W7NNM     | 16.11| 0.87 |
| Ours      | 16.21| 0.88 |

TABLE 4. Average PSNR/SSIM results of competing denoising algorithms on CC60 dataset [25], [36].

| Method    | PSNR | SSIM |
|-----------|------|------|
| CBM3D     | 15.85| 0.83 |
| M2CSNMM   | 16.04| 0.85 |
| GID       | 15.68| 0.81 |
| MCWSNMM   | 16.10| 0.86 |
| TWSC      | 15.75| 0.82 |
| CMS-SVD   | 16.27| 0.87 |
| DnCNN     | 15.98| 0.84 |
| PFDDNet   | 16.09| 0.85 |
| W7NNM     | 16.11| 0.87 |
| Ours      | 16.21| 0.88 |

FIGURE 3. Denoised images of the real noisy image captured by Nikon D800 from CC15 dataset by different methods.

\[
\left\lceil \frac{1}{2} \right\rceil \text{ SVDs of } p^2 \times N \text{ matrices. The computation cost of generalized iterated soft-thresholding (GST) is } O(JN), \text{ where } J \text{ is the iteration number in the GST. Thus the computation complexity is } O(p^2N\log(C) + p^4NC + JNMT). \text{ Suppose there are } T \text{ iterations, the overall complexity of our proposed method is } O((p^2\log(C) + p^4 + JNMT)). \text{ According}
\]
to [18], the time complexity of CBM3D and CMS-t-SVD are $O((p^2\log(p) + p^2\log(N))NM)$ and $O((p^3 + p^4)NM)$. The computation cost of MCWNNM in [38] and TWSC in [37] are $O((p^4N + N^3)MT)$ and $O((\max(p^6, p^2N))MT)$. Thus, our proposed method is less efficient compared to CBM3D and CMS-t-SVD. The running time results of different methods are listed in Table 4 and the fastest result is highlighted in bold. Although DnCNN and FFDNet are much faster than other methods benefitting from GPU computation, they have to take about 1 day to train specific models on GPU. It should be noted that CBM3D, CMS-t-SVD are implemented with compiled C++ mex-function and parallel computation, therefore their computational time is much less than other competing methods like MCWNNM, TWSC, WTNNM and the proposed method which are implemented in Matlab. Although the proposed method is more computational complicated, it delivers better quantitative denoising results with finer details and less artifacts.

VI. CONCLUSION

In this paper, we treat a real color image as a third-order tensor with each color channel corresponding to a frontal slice, and introduce low tubal rank minimization to exploit the spatial and spectral information of the image. We propose a weighted $t$-Schatten-$p$ norm minimization for real color
image denoising. The proposed method is flexible to provide different treatment to different rank components, therefore the denoised image is closer to the latent low rank tensor. Extensive experiments on one synthetic and two realistic image datasets demonstrate that our proposed method could produce competitive performance compared to other state-of-the-art methods for real color image denoising.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments, and appreciate all the authors for providing their code.

REFERENCES

[1] A. Beck and M. Teboulle, “Fast gradient-based algorithms for constrained total variation image denoising and deblocking problems,” IEEE Trans. Image Process., vol. 18, no. 11, pp. 2419–2434, Nov. 2009.

[2] A. Buades, B. Coll, and J.-M. Morel, “A non-local algorithm for image denoising,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR), vol. 2, Jun. 2005, pp. 60–65.

[3] H. C. Burger, J. C. Schuler, and S. Harmeling, “Image denoising: Can plain neural networks compete with BMD?” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 2392–2399.

[4] J. D. Carroll and J.-J. Chang, “Analysis of individual differences in multidimensional scaling via an n-way generalization of ‘Eckart-Young’ decomposition,” Psychometrika, vol. 35, no. 3, pp. 283–319, Sep. 1970.

[5] G. Chen, F. Zhu, and P. A. Heng, “An efficient statistical method for image noise level estimation,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 477–485.

[6] Y. Chen, W. Yu, and T. Pock, “On learning optimized reaction diffusion processes for effective image restoration,” in Proc. ACM Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2015, pp. 5261–5269.

[7] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Color image denoising via sparse 3D collaborative filtering with grouping constraint in luminance-chrominance space,” in Proc. IEEE Int. Conf. Image Process., vol. 1, Sep. 2007, pp. I-313–I-316.

[8] S. Gu, Z. Yan, K. Zhang, W. Zuo, and L. Zhang, “Toward convolutional blind denoising of real photographs,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 1712–1722.

[9] C. J. Hollar and L.-H. Lim, “Most tensor problems are NP-hard,” J. ACM, vol. 60, no. 6, pp. 1–39, Nov. 2013.

[10] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, “Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging,” SIAM J. Matrix Anal. Appl., vol. 34, no. 1, pp. 148–172, Jan. 2013.

[11] 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.

[12] H. Kong, X. Xie, and Z. Lin, “TF-schatten-p norm for low-rank tensor recovery,” IEEE J. Sel. Topics Signal Process., vol. 12, no. 6, pp. 1405–1419, Dec. 2018.

[13] Z. Kong and X. Yang, “Color image and multispectral image denoising using block diagonal representation,” IEEE Trans. Image Process., vol. 28, no. 9, pp. 4247–4259, Sep. 2019.

[14] Z. 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.

[15] Z. Liu, X. Zhang, and T. Zhang, “Robust tensor completion and low-rank representation,” IEEE Trans. Image Process., vol. 24, no. 12, pp. 5469–5478, Dec. 2015.

[16] E. Luo, S. H. Chan, and T. Q. Nguyen, “Adaptive image denoising by mixture adaptation,” IEEE Trans. Image Process., vol. 25, no. 10, pp. 4489–4503, Oct. 2016.

[17] S. Nam, Y. Hwang, Y. Matsuhashi, and S. J. Kim, “A holistic approach to cross-channel image noise modeling and its application to image denoising,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 1683–1691.

[18] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, “An iterative regularization method for total variation-based image restoration,” Multiscale Model. Simul., vol. 4, no. 2, pp. 460–489, Jan. 2005.

[19] Y. Sun, Z. Lin, H. Jin, J. Yang, and Z. Wang, “Image-specific prior adaptation for denoising,” IEEE Trans. Image Process., vol. 31, no. 4, pp. 849–862, Apr. 2013.

[20] B. Romera-Paredes and M. Pontil, “A new convex relaxation for tensor completion,” in Proc. Adv. Neural Inf. Process. Syst., 2013, pp. 2967–2975.

[21] S. Roth and M. J. Black, “Fields of experts,” Int. J. Comput. Vis., vol. 82, no. 2, pp. 205, 2009.

[22] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” Phys. D. Nonlinear Phenomena, vol. 60, nos. 1–4, pp. 259–268, Nov. 1992.

[23] L. R. Tucker, “Some mathematical notes on three-mode factor analysis,” Psychometrika, vol. 31, no. 3, pp. 279–311, Sep. 1966.

[24] S. Wang, Y. Wang, Y. Wu, T. Lin, S. Li, and B. Chen, “Quantification of full left ventricular metrics via deep regression learning with contour guidance,” IEEE Access, vol. 7, pp. 47918–47928, Oct. 2019.

[25] Y. Wu, L. Fang, and S. Li, “Weighted tensor Rank-1 decomposition for nonlocal image denoising,” IEEE Trans. Image Process., vol. 28, no. 6, pp. 2719–2730, Jun. 2019.

[26] Z. Wu, Q. Wang, Z. Wu, and Y. Shen, “Total variation-regularized weighted nuclear norm minimization for hyperspectral image mixed denoising,” J. Electron. Imag., vol. 25, no. 1, Feb. 2016, Art. no. 013037.

[27] Y. Xie, S. Gu, Y. Liu, W. Zuo, W. Zhang, and L. Zhang, “Weighted schatten p-norm minimization for image denoising and background subtraction,” IEEE Trans. Image Process., vol. 25, no. 10, pp. 4842–4857, Oct. 2016.

[28] J. Xu, L. Zhang, and D. Zhang, “External prior guided internal prior learning for real-world noisy image denoising,” IEEE Trans. Image Process., vol. 27, no. 6, pp. 2996–3010, Jun. 2018.

[29] J. Xu, L. Zhang, and D. Zhang, “A trilateral weighted sparse coding scheme for real-world image denoising,” in Proc. Eur. Conf. Comput. Vis. (ECCV), 2018, pp. 20–36.

[30] J. Xu, L. Zhang, D. Zhang, and X. Feng, “Multi-channel weighted nuclear norm minimization for real color image denoising,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 1096–1104.

[31] J. Xu, L. Zhang, W. Zuo, D. Zhang, and X. Feng, “Patch group based nonlocal self-similarity prior learning for image denoising,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 244–252.

[32] Z. Zhu, X. Zhang, Q. Wang, Y. Bai, Y. Chen, L. Tang, and X. Liu, “Group sparsity residual constraint for image denoising with external nonlocal self-similarity prior,” Neurocomputing, vol. 275, pp. 229–236, Jan. 2018.

[33] K. Zhang, W. Zuo, and L. Zhang, “Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising,” IEEE Trans. Image Process., vol. 26, no. 7, pp. 3142–3155, Jul. 2017.

[34] K. Zhang, W. Zuo, and L. Zhang, “FFDNet: Toward a fast and flexible solution for CNN-based image denoising,” IEEE Trans. Image Process., vol. 27, no. 9, pp. 4608–4622, Sep. 2018.
[44] 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., Jun. 2014, pp. 3842–3849.

[45] Z. Zhang, C. Duan, T. Lin, S. Zhou, Y. Wang, and X. Gao, “GVFOM: A novel external force for active contour based image segmentation,” Inf. Sci., vol. 506, pp. 1–18, Jan. 2020.

[46] 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.

[47] D. Zoran and Y. Weiss, “From learning models of natural image patches to whole image restoration,” in Proc. Int. Conf. Comput. Vis., Nov. 2011, pp. 479–486.

[48] W. Zuo, D. Meng, L. Zhang, X. Feng, and D. Zhang, “A generalized iterated shrinkage algorithm for non-convex sparse coding,” in Proc. IEEE Int. Conf. Comput. Vis., Dec. 2013, pp. 217–224.

MIN LIU (Member, IEEE) received the B.S. degree from the Nanjing University of Information Science and Technology, Nanjing, China, in 2008, and the M.S. degree from Nanjing University, in 2011, where she is currently pursuing the Ph.D. degree with the Department of Electronic Science and Engineering. Her current research interests include image denoising, sparse signal representation, tensor analysis, and machine learning.

XINGGAN ZHANG received the B.S. degree in electrical engineering and the S.M. and Ph.D. degrees from the Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China, in 1982, 1988, and 2001, respectively. He joined the Department of Electronic Engineering, NUAA, in 1992, where he was an Associate Professor. He joined the Department of Electronic Science and Engineering, Nanjing University, in 1999, where he is currently a Professor. His research interests include target recognition and image processing.

LAN TANG (Member, IEEE) received the B.S. and M.S. degrees in communication engineering from Jilin University, Jilin, China, in 2002 and 2005, respectively, and the Ph.D. degree in communication and information science from Southeast University, Nanjing, China, in 2009. Since 2009, she has been with the School of Electronic Science and Engineering, Nanjing University, China, where she is currently an Associate Professor. Her current research interests include cooperative communication, energy efficient and energy harvesting wireless communication, wireless information and power transfer, and optimization theory.