Robust Tensor Approximation with Laplacian Scale Mixture Modeling for Multiframe Image and Video Denoising

Weisheng Dong, *Member, IEEE*, Tao Huang, Guangming Shi, *Senior member, IEEE*, Yi Ma and Xin Li

***Abstract*—Sparse and low-rank models have been widely studied in the literature of signal processing and computer vision. However, as the dimensionality of data set increases (e.g.,multispectral images, dynamic MRI images and video sequences), the optimality of vector and matrix-based data representations and modeling tools becomes questionable. Inspired by recent advances in sparse and low-rank tensor analysis, we propose a novel robust tensor approximation (RTA) framework with Laplacian Scale Mixture (LSM) modeling for 3D data and beyond. Our technical contributions are summarized as follows: 1) conceptually similar to robust PCA, we consider its tensor extension here - i.e., low-rank tensor approximation in the presence of outliers modeled by sparse noise; 2) built upon previous work on tensor sparsity, we propose to model tensor coefficients with a LSM prior and formulate a Maximum A Posterior (MAP) estimation problem for noisy observations. Both unknown sparse coefficients and hidden LSM parameters can be efficiently estimated by the method of alternating optimization; 3) we have derived closed-form solutions for both subproblems and developed computationally efficient denoising techniques for multiframe images and video. Experimental results on three data sets have shown that the proposed algorithm can better preserve the sharpness of important image structures and outperform several existing state-of-the-art image/video denoising methods (e.g., BM4D/VBM4D and tensor dictionary learning).**

I. Introduction

As sensing and computing technologies advance, more and more high-dimensional data are acquired these days to support a wide range of scientific and engineering applications. For example, in the field of geoscience remote sensing, early multispectral cameras have limited spectral resolution and can acquire a small number of (usually 3-15) spectral bands; while recently developed hyperspectral imaging technology can support contiguous coverage (often hundreds of spectral bands) with unprecedent spectrum resolution. Taking another example, fast advances in MRI technology have evolved from parallel collection of k-space data using SENSE [1] to the acquisition of MRI sequence using k-t SENSE [2]; the addition of temporal dimension creates both opportunities and challenges for so-called functional or dynamic MRI [3]. Other examples of higher-dimensional data include multi-exposure imagery for high dynamic range (HDR) imaging [4], volumetric data set for visualization [5] and multiview video (to support interactive 3D TV) [6].

As the dimensionality reaches beyond 3D, conventional data representations such as vector-based and matrix-based become insufficient. Artificially reshaping multi-dimensional data into 1D vectors or 2D matrices inevitably destroys the intrinsic dependency structures among multiple factors, which is not desirable for further analysis or processing. A better and more principled approach of preserving multi-factor structure is to us higher-order tensors represented by multi-dimensional arrays. However, unlike vectors and matrices, the collection of mathematical and computational tools developed for tensor analysis have been scarce in the literature. A handful exceptions in the existing literature include parallel factor analysis (PARAFAC) [7], multilinear algebra [8], tensor decompositions [9], factor analysis [10], low-rank tensor approximation [11], high-order singular value decomposition [12] and Kronecker-Basis-Representation based tensor sparsity[13]).

In particular, there is an urgent need for developing tensor analysis tools that are *robust* to outliers and noise contaminations for the following reasons. First, multi-dimensional data acquired from real world are often imperfect - either corrupted by noise (e.g., sensor noise during the MRI acquisition) or contain missing samples (e.g., depth maps acquired by Kinect sensors). It is desirable to suppress the interference from the noisy or incomplete observation by exploiting the high-order intrinsic structure underlying data ensembles. Second, even in the situation of clean and complete data, the task of data analysis often requires the separation or segmentation of multi-dimensional data - e.g., for video data, background subtraction has been extensively studied in the literature [14]. When background is the target of estimation, foreground data are treated as outliers and therefore calls for robust processing. Third, there has been a flurry of works on robust statistical tools developed for low-dimensional data such as robust principal component analysis (RPCA) [15], [16], [17], robust subspace learning [18] and robust subspace clustering [19], [20]. Inspired by these recent advances, it is intellectually appealing to pursue their higher-order counterpart (i.e., robust tensor analysis and computationally efficient implementations).

In this paper, we make three-fold (insight, tool and application) contributions as follows. First, we suggest that

A preliminary version of this work was published by the Proceedings of International Conference on Computer Vision (ICCV)'2015.

This work was supported in part by the Natural Science Foundation of China under Grant 61622210, Grant 61471281, Grant 61632019, Grant 61836008, Grant 61621005, and Grant 61390512.

W. Dong, T. Huang and G. Shi are with the School of Artificial Intelligence, Xidian University, Xi'an, 710071, China (e-mail: wsdong@mail.xidian.edu.cn)

Yi Ma is with the Dept. of EECS, University of California at Berkeley, California 94720, USA

Xin Li is with Lane Dep. of CSEE, West Virginia University, Morgantown WV 26506-6109, USA

self-similarity is a fundamental property of multidimensional data acquired from the physical world. To exploit the self-repeating patterns in those data, we propose to generalize existing idea of grouping similar 2D patches into higher-dimensional space. Higher-order tensor becomes a natural and unified framework for organizing those similar low-order tensors. Such new insight allows us to model complex data with multiple uncertainty factors (e.g., multi-spectral, multi-exposure or multi-view) in a principled manner.

Second, we propose to develop robust statistical modeling tools for tensor coefficients via Laplacian scalar mixture (LSM) [21]. Our work can be interpreted as an extension of the previous tensor sparsity using Kronecker-Basis-Representation [13] targeting at a more faithful characterization of uncertainty in tensor coefficients by parameterized models. Various data restoration tasks can be formulated as a Maximum A Posterior (MAP) estimation problem with the LSM prior for tensor coefficients. Both unknown sparse coefficients and hidden LSM parameters can be efficiently estimated by the method of alternating optimization [22], [23].

Third, we have implemented a novel robust tensor approximation framework with Laplacian Scale Mixture (LSM) modeling for multi-frame image and video denoising. It is rigorously shown that both problems of estimating unknown sparse coefficients and hidden LSM parameters admit closed-form solutions leading to computationally efficient implementations. In particular, we have shown how to robustly estimate the variance parameter from noisy observations, which generalizes our previous work [11] focusing on nonlocal self-similarity only. Experimental results have shown that the proposed denoising method significantly outperforms current state-of-the-art multiframe image and video denoising methods including tensor dictionary learning [24] and BM4D [25].

The rest of the paper is organized as follows. In Sec. II, we briefly review the literature of sparse coding and low rank matrix completion by structuring them into vector-based, matrix-based and tensor-based methods. In Sec. III, we propose to exploit nonlocal similarity of multidimensional data by low rank tensor approximation and model the tensor coefficients by Laplacian scalar mixture (LSM). In Sec. IV, we study the application of robust tensor approximation with Laplacian scalar mixture (RTA-LSM) into multiframe/multispectral image denoising and video denosing. In Sec. V, we report our experimental results and comparison against other competing denoising techniques in the literature. In Sec. VI, we make some concluding remarks about this work and future research.

II. Related Works

A. Previous Works on Data Modeling and Representations

In this subsection, we briefly review three classes of data modeling methods in 1D, 2D, 3D and beyond respectively. More complete reviews can be found in a recent book [26] as well as survey papers [27], [28], [29].

Vector-based methods are often built upon the theory of sparse coding (a.k.a. sparse representations). The basic idea behind sparse coding is that the target signal - when represented by a 1D vector in R^N - can be well approximated by the linear combination of a small number of atoms (a.k.a. basis functions) from a dictionary $\mathbf{D} \in R^{M\times N}$. Early works have used fixed dictionaries (e.g., DCT and wavelets) whose elements possess good localization properties in space/time and frequency; later adaptive dictionaries were obtained by so-called dictionary learning approaches which rooted in the classical principal component analysis (PCA) [30], [31], [32]. No matter a fixed or adaptive dictionary is used, the sparsity of targeted signal is often enforced by a soft/hard thresholding operation with respect to sparse coefficients.

An important new insight was brought to the theory of sparse coding by exploiting so-called *nonlocal* self-similarity in real-world signals such as natural images [33]. Thanks to the self-repeating patterns in those data, it is possible to find many patches similar to an exemplar one; when those similar patches are grouped together, a new dimension (reflecting the nonlocal connection) is induced to the data representation [34], [35], [36]. If each 2D image patch is reshaped into a 1D vector, grouping similar patches will form a 2D matrix with a desirable low-rank property; if similar 2D patches are directly packed together, we have a 3D data array with the familiar sparse property (after 3D sparsifying transforms). Either way, nonlocal self-similarity of real-world data boils down to structural constraints that can be conveniently exploited by linear algebraic tools.

Matrix-based methods are built upon the theory of low-rank matrix approximation [37], [38]. As mentioned above, when similar patches are grouped to form a data matrix $\mathbf{Y}$, the rank of $\mathbf{Y}$ is usually low. Therefore, common tasks such as noise removal and missing data completion can be achieved by recovering a low-rank matrix from its noisy/incomplete version; more specifically, singular value decomposition (SVD) has been a standard tool for the class of matrix completion problems [12]. By designing appropriate shrinkage/thresholding operators for singular values, sate-of-the-art image denoising performance can be achieved [37], [38].

Similar to the former case, it is desirable to exploit nonlocal self-similarity in 3D data such as volumetric data and video. To accommodate higher-order data representation, high-order SVD becomes a natural extension for handling the grouped multi-dimensional data arrays. However, only ad-hoc thresholding or Wiener filtering techniques have been considered in the previous work of HOSVD denoising [12]. It is desirable to establish a more systematic framework for modeling and exploiting the self-repeating patterns in 3D data and beyond, which serves as the motivation for developing tensor-based methods next.

Tensor-based methods have been scarcely studied in the literature (e.g., parallel factor analysis (PARAFAC) [7], multilinear algebra [8], tensor decompositions [9], factor analysis [10]). The connection between tensor approximation and sparse representation has only been recently established using Kronecker-Basis-Representation [13]. Accordingly, the applications of tensor-based methods have been limited to volumetric data denoising so far. In [39], [40], [11], a low-rank tensor approximation method was developed for multispectral image denoising. Most recently, Peng [24] proposed an ef-

fective multispectral image denoising method using nonlocal tensor dictionary learning.

To exploit nonlocal spatio-temporal redundancy, 3D image patches are clustered into groups via k-means clustering. Then each set of grouped similar 3D patches are linearly approximated by low-rank tensor approximation. More specifically, the criterion of minimum description length (MDL) [41] can be used to determine the rank for each tensor model. However, existing tensor approximation tools did not take noise interference into consideration and therefore are not *robust* at all. To fill in this gap, we propose to model the tensor coefficients by Laplacian Scale Mixture (LSM) modeling techniques and develop robust tensor approximation tools in this paper.

B. Previous Works on Multi-frame Image and Video Denoising

The field of single-frame image denoising has advanced rapidly in the past decades. Early attacks on denoising are based on sparse representations of images (e.g., [42], [43], [44]); later this line of research has evolved into image denoising methods via dictionary learning [30], [31], [32]. Since 2005, the class of nonlocal image denoising has received increasingly more attention - dozens of competing methods have been proposed inspired by pioneering works such as nonlocal-mean [33] and BM3D [34]. Recent advances include the combination with dictionary learning (e.g., [35]), the extension of sparse representations (e.g., [36]) and the connection with low-rank matrix approximation (e.g., [37], [38]).

By contrast, the field of multi-frame image denoising has been under-researched as of today despite the abundance of volumetric data in the real world (e.g., volume rendering, remote sensing and dynamic MRI). In the literature, there exist several multi-frame image denoising methods built upon straightforward extension of their single-frame counterpart. For instance, sparse coding based denoising was extended into low-rank tensor approximation (LATA) method in [39]; the method of nonlocal-mean filtering was made spatially adaptive for the application of MRI in ANLM3D [45]; the tool of was introduced into hyperspectral image denoising in [40]. Later, state-of-the-art image denoising method BM3D was extended into BM4D [25] for volumetric data restoration and tensor dictionary learning (TensorDL) method [24] was developed for multispectral image denoising. Most recently, KBR-based tensor sparsity has found successful application into multispectral image denoising in [13].

The literature of video denoising is not as rich as its image counterpart. Low-rank based approach toward video denoising was first proposed in [46]; video denoising via reliable motion estimation [47] and spatiotemporal GSM [48] was developed the same year. The nonlocal mean (NLM) method was extended with adaptive regularization for video denoising in [49]; the BM4D [25] approach was extended into video denoising in VBM4D [50]. Later, patch-based video denoising with optical flow estimation was developed in [51] which has shown even more improved performance over VBM4D [50]; global search has been found to boost the performance of video denoising in [52]. Most recently, low rank tensor approximation has been applied into video denoising [53], [54].

III. Robust Tensor Approximation with Laplacian Scale Mixture Modeling

In this paper, we present a novel robust tensor approximation (RTA) framework with Laplacian Scale Mixture (LSM) modeling for high-dimensional data and demonstrate its potential into multiframe image/video denoising. We first introduce a sparse tensor approximation (STA) method for exploiting nonlocal self-similarity; then connect it with Laplacian Scale Mixture (LSM) modeling to derive computationally efficient and robust optimization algorithm; finally we consider the applications into multispectral image denoising, multiframe MRI denoising and color video denoising. Throughout this paper, we use $\mathcal{S}$, $\tilde{\mathcal{S}}$, $\hat{\mathcal{S}}$ to denote clean/original, noisy and reconstructed tensors respectively.

A. Sparse Tensor Approximation for Exploiting Nonlocal Self-Similarity

Sparse tensor approximation consists of two steps: *patch grouping* and *sparse approximation*. The objective of patch grouping is to find similar patches from the given data to form a new representation suitable for sparse coding (a.k.a. simultaneous sparse coding [55] or group sparsity [56]). For 2D photographic images, it has been generally agreed that the source of self-similarity arises from the abundance of regular edges and textures in the physical world. Does such observation extend to higher-dimensional data? We argue that the answer is affirmative at least for the high-dimensional data of interest to us including multispectral/multiview/multiexposure images and video. A unifying view toward those data is that the new dimension is induced by some well-defined physical quantity (e.g., spectral, viewpoint, exposure time and motion). Sampling along this new dimension has to observe the same fundamental physical laws (e.g., Pauli's exclusion principle [57] and limited speed of physical motion) as sampling in space and time; therefore similar 2D image patches tend to remain similar within the same sampling interval along this new dimension.

To group 3D patches from a given (possibly noisy) 3D image of size $H \times W \times L$, we propose the following strategy (note that the issue of noise to patch grouping will be addressed later). For each exemplar patch $\mathcal{P}_i$ (sized by $\sqrt{n} \times \sqrt{n} \times L$ and located at the spatial position i(the collection of spatial positions forms a sublattice of integral lattice in 3D), we search for similar patches within a local window (e.g., $40 \times 40 \times 40$). That is

$$G_i = \{i_j | \|\mathcal{P}_i - \mathcal{P}_{i_j}\|_{L_2} < T\}, \tag{1}$$

where T is a predefined threshold and G_i stores the positions of similar patches. Alternatively, one can form G_i by k-nearest neighbor (k-NN) method. After patch grouping, we can stack similar 3D patches into a 4^{th}-order tensor in which the 4th dimension signifies the nonlocal similarity in 3D data; alternatively, we can also obtain a 3^{rd}-order tensor by reshaping each matrix slice of 3D patches into 1D vectors. Such dimensionality-reduction strategy is conceptually similar to the existing low-rank approximation for 2D image denoising where 2D patches are reshaped into 1D vectors and stacked into a 2D data matrix instead of a 3D data array (note

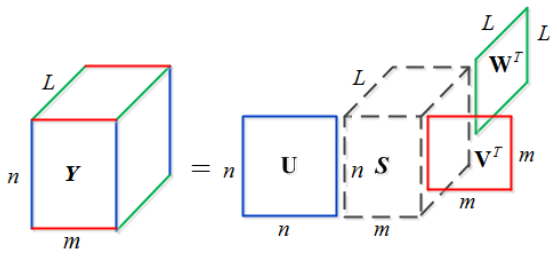


Fig. 1. The HOSVD expansion of a low-rank Tensor $\mathcal{Y}$.

that similar strategy has been adopted in [13] too). One potential benefit for such dimensionality-reduction strategy is to facilitate the task of dictionary learning (e.g., 2D SVD transform in [37] has been shown to work better than 3D sparsifying transform in BM3D [58]).

Sparse approximation of tensors is a natural extension of existing sparse coding problems studied for vectors and matrices [13]. For vectors, the sparsity can be formalized by the number of nonzero coefficients (l_0-norm) or the summation of absolute coefficients (l_1-norm); for matrices, the nuclear norm or the summation of singular values has found a convenient and suitable choice for characterizing the sparsity. Most recently, sparsity measures for tensors were studied in [13] and a Kronecker-Basis-Representation (KBR) based tensor sparsity (a combination of l_0-norm and rank norm) was proposed as the objective of regularization. In view of computational burden with the task of tensor sparsity enforcement, an effective alternating direction method of multipliers (ADMM) algorithm [59] was designed to solve the KBR regularization minimization problem in [13]. By contrast, we propose to study the "convex" version here (i.e., skipping the nonconvex l_0-norm) and directly consider the enforcement of tensor sparsity by a high-order SVD (HOSVD) approach as follows.

Given a noisy tensor $\mathcal{Y}_i$, we first construct its HOSVD as [60], [61]

$$\begin{aligned}\mathcal{Y}_i &= \sum_{r=1}^{n}\sum_{c=1}^{m}\sum_{l=1}^{L}\tilde{\mathcal{S}}_i(r,c,l)\boldsymbol{u}_{i,r}\times\boldsymbol{v}_{i,c}\times\boldsymbol{w}_{i,l}\\ &= \tilde{\mathcal{S}}_i\times_1\mathbf{U}_i\times_2\mathbf{V}_i\times_3\mathbf{W}_i,\end{aligned} \tag{2}$$

where $\mathbf{U}_i = [\boldsymbol{u}_{i,1},\cdots,\boldsymbol{u}_{i,n}] \in \mathbb{R}^{n\times n}$, $\mathbf{V}_i = [\boldsymbol{v}_{i,1},\cdots,\boldsymbol{v}_{i,m}] \in \mathbb{R}^{m\times m}$ and $\mathbf{W}_i = [\boldsymbol{w}_{i,1},\cdots,\boldsymbol{w}_{i,L}] \in \mathbb{R}^{L\times L}$ denote orthogonal matrices, $\tilde{\mathcal{S}}_i \in \mathbb{R}^{n\times m\times L}$ is the transformed 3D coefficient array (so-called "core tensor"), $\tilde{\mathcal{S}}_i(r,c,l)$ are the components of $\tilde{\mathcal{S}}_i$, $\times$ denotes the tensor product (i.e., $\boldsymbol{x}\times\boldsymbol{y} = \boldsymbol{x}\boldsymbol{y}^\top$), and $\times_j$ denotes the $j-th$ model tensor product. Three orthogonal matrices $\mathbf{U}_i$, $\mathbf{V}_i$ and $\mathbf{W}_i$ can be computed from the SVD of mode-j ($j = 1, 2, 3$) flattening of $\mathcal{Y}_i$ respectively. An example of HOSVD decomposition is illustrated in Fig.III-A

Note that since the tensor is formed by grouping similar patches, $\mathcal{Y}_i$ can be approximated by a low-rank representation [11] - i.e.,

$$\begin{aligned}\hat{\mathcal{X}}_i &= \sum_{r=1}^{r_1}\sum_{c=1}^{r_2}\sum_{l=1}^{r_3}\hat{\mathcal{S}}_i(r,c,l)\boldsymbol{u}_{i,r}\times\boldsymbol{v}_{i,c}\times\boldsymbol{w}_{i,l}\\ &= \hat{\mathcal{S}}_i\times_1\hat{\mathbf{U}}_i\times_2\hat{\mathbf{V}}_i\times_3\hat{\mathbf{W}}_i,\end{aligned} \tag{3}$$

where $\hat{\mathbf{U}}_i = [\boldsymbol{u}_{i,1},\cdots,\boldsymbol{u}_{i,r_1}] \in \mathbb{R}^{n\times r_1}$, $\hat{\mathbf{V}}_i = [\boldsymbol{v}_{i,1},\cdots,\boldsymbol{v}_{i,r_2}] \in \mathbb{R}^{m\times r_2}$ and $\hat{\mathbf{W}}_i = [\boldsymbol{w}_{i,1},\cdots,\boldsymbol{w}_{i,r_3}] \in \mathbb{R}^{L\times r_3}$ are the three matrices associated with $\mathbf{U}_i$, $\mathbf{V}_i$ and $\mathbf{W}_i$ respectively, $r_1 \le n$, $r_2 \le m$ and $r_3 \le L$, and $\hat{\mathcal{S}}_i \in \mathbb{R}^{r_1\times r_2\times r_3}$ denotes the smaller core tensor. The triple (r_1, r_2, r_3) is often called the *multirank* of $\mathcal{Y}_i$. To estimate the multirank of a tensor, minimum description length (MDL) method [41] can be used for different mode-k flattening of the tensor. With the estimated rank parameters (r_1, r_2, r_3), a low-rank tensor approximation can be easily obtained by setting the last $n-r_1$, $m-r_2$ and $L-r_3$ slices along different modes in $\tilde{\mathcal{S}}_i$ to be zero matrices.

For matrices, it has been suggested that nuclear norm is a more desirable choice of regularization than rank from a computational perspective [62]. Similarly, we can also obtain the low-rank approximation of tensors by inducing the sparsity on the array of tensor coefficients rather than explicitly estimating multirank parameters - namely

$$\begin{aligned}\hat{\mathcal{S}}_i &= \underset{\mathcal{S}_i}{\operatorname{argmin}}\,\psi(\mathcal{S}_i),\\ &\quad s.\,t.,\ ||\mathcal{Y}_i - \mathcal{S}_i\times_1\mathbf{U}_i\times_2\mathbf{V}_i\times_3\mathbf{W}_i||_F^2 \le \sigma_w^2,\end{aligned} \tag{4}$$

where $\psi(\cdot)$ is a sparse regularization function inducing the sparsity to the components of $\mathcal{S}_i$ (e.g., KBR-based tensor sparsity [13]); and ($\mathbf{U}_i$, $\mathbf{V}_i$ $\mathbf{W}_i$) are orthogonal matrices obtained by the HOSVD of $\mathcal{Y}_i$. Thanks to the orthogonality of those matrices, Eq. (4) can be rewritten into

$$\hat{\mathcal{S}}_i = \underset{\mathcal{S}_i}{\operatorname{argmin}}\,\psi(\mathcal{S}_i),\ s.\,t.,\ ||\tilde{\mathcal{S}}_i - \mathcal{S}_i||_F^2 \le \sigma_w^2, \tag{5}$$

where $\tilde{\mathcal{S}}_i = \mathcal{Y}_i\times_1\mathbf{U}_i^\top\times_2\mathbf{V}_i^\top\times_3\mathbf{W}_i^\top$. It follows that the problem of sparse tensor approximation can be reformulated in the Lagrangian form - i.e.,

$$\hat{\mathcal{S}}_i = \underset{\mathcal{S}_i}{\operatorname{argmin}}\,||\tilde{\mathcal{S}}_i - \mathcal{S}_i||_F^2 + \lambda\psi(\mathcal{S}_i). \tag{6}$$

Similar to the theory of low-rank approximation, popular choices of $\psi(\cdot)$ include the pseudo-norm ℓ_0 and the ℓ_1 norm, which exactly lead to the hard and soft thresholding of tensor coefficient array $\tilde{\mathcal{S}}_i$ respectively. Generally speaking, the selection of threshold λ is non-trial; for this reason, a heuristic two-stage method similar to BM3D [58] has been adopted in [12]. In this work, we propose to develop a more *robust* and *principled* approach toward sparse tensor approximation modeling of tensor coefficients by Laplacian scale mixture (LSM).

B. Laplacian scale mixture modeling for robust tensor approximation

In the literature of sparse coding and low rank approximation, the choice of threshold parameter (equivalent to Lagrangian multiplier) is often solved by modeling sparse coefficients or singular values by a parametric probabilistic distribution. Along this line of reasoning, we propose to formulate a Maximum a Posterior (MAP) estimation problem for inferring unknown tensor coefficients $\mathcal{S}_i$ from corrupted or incomplete $\tilde{\mathcal{S}}_i$. To simplify the notation, we will use $\tilde{\boldsymbol{s}} \in \mathbb{R}^{n\cdot m\cdot L}$ and $\boldsymbol{s} \in \mathbb{R}^{n\cdot m\cdot L}$ to denote the one-dimensional representations of $\tilde{\mathcal{S}}_i$ and $\mathcal{S}_i$ respectively (i.e., the subscript

index i denoting spatial position is dropped). Let $\boldsymbol{s}$ denote the noiseless version of $\tilde{\boldsymbol{s}}$, then we have $\tilde{\boldsymbol{s}} = \boldsymbol{s} + \boldsymbol{n}$, where $\boldsymbol{n} \in \mathbb{R}^{n \cdot m \cdot L}$ denotes the term of additive Gaussian noise. It follows that the MAP estimation of $\boldsymbol{s}$ from $\tilde{\boldsymbol{s}}$ can be written as

$$\boldsymbol{s} = \underset{\boldsymbol{s}}{\operatorname{argmin}}\{-\log P(\tilde{\boldsymbol{s}}|\boldsymbol{s}) - \log P(\boldsymbol{s})\}, \tag{7}$$

where $\log P(\tilde{\boldsymbol{s}}|\boldsymbol{s})$ is given by

$$P(\tilde{\boldsymbol{s}}|\boldsymbol{s}) \propto \exp(-\frac{1}{2\sigma_w^2}||\tilde{\boldsymbol{s}} - \boldsymbol{s}||_2^2), \tag{8}$$

and *a prior* distribution of $\boldsymbol{s}$ is given by

$$P(\boldsymbol{s}) \propto \prod_j \exp(-\frac{\psi(s_j)}{\theta_j}). \tag{9}$$

It is easy to verify that MAP estimation in Eq. (7) boils down to the following weighted ℓ_1 norm minimization problem when $P(\boldsymbol{s})$ is chosen to observe independent and identically distributed (iid) Laplacian - i.e.,

$$\boldsymbol{s} = \underset{\boldsymbol{s}}{\operatorname{argmin}}\, ||\tilde{\boldsymbol{s}} - \boldsymbol{s}||_2^2 + 2\sigma_w^2 \sum_j \frac{1}{\theta_j}|s_j|, \tag{10}$$

where θ_j is the standard derivation of s_j. We note that it has been shown that the *weighted* ℓ_1 norm is more effective than the original ℓ_1 norm in the literature of sparse approximation [63]. With the above setup, the key issue now becomes how to robustly estimate the variance parameters θ_j from noisy observations $\tilde{\boldsymbol{s}}$.

Our proposed robust estimation of variance parameter θ_j from noisy observation is based on the assumption with a Laplacian Scale Mixture (LSM) prior for $\boldsymbol{s}$. Similar to the Gaussian scale mixture (GSM) prior [43], one can decompose $\boldsymbol{s}$ into point-wise product of a Laplacian vector $\boldsymbol{\alpha}$ and a positive hidden scalar multiplier $\boldsymbol{\theta}$ with probability $P(\theta_j)$- i.e., $s_j = \theta_j \alpha_j$. Conditioned on θ_j, s_j observes a Laplacian distribution with standard deviation of θ_j. Assuming that θ_j and α_j are independent, we can write the LSM prior of $\boldsymbol{s}$ as

$$P(\boldsymbol{s}) = \prod_i P(s_j),\ P(s_j) = \int_0^\infty P(s_j|\theta_j)P(\theta_j)d\theta_j. \tag{11}$$

It should be noted that since there is no close-form expression of $P(\boldsymbol{s})$ for many choices of $P(\theta_j)$, it is usually infeasible to analytically solve the above MAP estimation of $\boldsymbol{s}$ with the LSM prior.

However, such difficulty can be overcome by using a joint prior model $P(\boldsymbol{s}, \boldsymbol{\theta})$ as we have in our recent work [64]. More specifically, we can substitute $P(\boldsymbol{s}, \boldsymbol{\theta})$ into the MAP estimation of Eq. (7) and obtain

$$(\boldsymbol{s}, \boldsymbol{\theta}) = \underset{\boldsymbol{s},\boldsymbol{\theta}}{\operatorname{argmin}}\{-\log P(\tilde{\boldsymbol{s}}|\boldsymbol{s}) - \log P(\boldsymbol{s}|\boldsymbol{\theta}) - \log P(\boldsymbol{\theta})\}. \tag{12}$$

where we have adopted a factorial distribution for multipliers and the noninformative Jeffrey's prior - i.e., $P(\theta_j) = \frac{1}{\theta_j}$. With Jeffrey's prior, Eq. (12) can be rewritten into

$$(\boldsymbol{s}, \boldsymbol{\theta}) = \underset{\boldsymbol{s},\boldsymbol{\theta}}{\operatorname{argmin}}\, ||\tilde{\boldsymbol{s}} - \boldsymbol{s}||_2^2 + 2\sqrt{2}\sigma_w^2 \sum_j \frac{|s_j|}{\theta_j} + 2\sigma_w^2 \sum_j \log\theta_j. \tag{13}$$

Note that similar to GSM, we have $\boldsymbol{s} = \boldsymbol{\Lambda}\boldsymbol{\alpha}$ in LSM, where $\boldsymbol{\Lambda} = \operatorname{diag}(\theta_j) \in \mathbb{R}^{n \cdot m \cdot L \times n \cdot m \cdot L}$. It follows that Eq.13 can be rewritten as

$$\begin{aligned}(\boldsymbol{\alpha}, \boldsymbol{\theta}) = &\underset{\boldsymbol{s},\boldsymbol{\theta}}{\operatorname{argmin}}\, ||\tilde{\boldsymbol{s}} - \boldsymbol{\Lambda}\boldsymbol{\alpha}||_2^2 + 2\sqrt{2}\sigma_w^2 \sum_j |\alpha_j| \\ &+ 4\sigma_w^2 \sum_j \log(\theta_j + \epsilon),\end{aligned} \tag{14}$$

where ϵ is a small positive constant ensuring numerical stability. It can be observed from Eq. (14) that robust approximation problem of $\boldsymbol{s}$ has been translated to the problem of jointly estimating $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$, which can be solved by the method of alternative optimization.

C. Alternative Optimization

In the previous work of sparse tensor approximation [13], relaxation (using log-sum form) is necessary to make the computation more tractable. By contrast, LSM-based approach in our work conveniently support computationally tractable solution to Eq. (14) using the method of alternating optimization, which forms the inner loop of the proposed denoising algorithm. More surprisingly, both sub-problems in our formulation admit closed-form solutions implying computationally efficient implementations as we will elaborate next.

For the first subproblem, given an initial estimate of $\boldsymbol{\alpha}$, one can obtain an estimate of $\boldsymbol{\theta}$ by optimizing

$$\boldsymbol{\theta} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\, ||\tilde{\boldsymbol{s}} - \mathbf{A}\boldsymbol{\theta}||_2^2 + 4\sigma_w^2 \sum_j \log(\theta_j + \epsilon), \tag{15}$$

where $\mathbf{A} = \operatorname{diag}(\boldsymbol{\alpha})$. Along the diagonal of $\mathbf{A}$, we can rewrite Eq. (15) into

$$\boldsymbol{\theta} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \sum_j \{a_j\theta_j^2 + b_j\theta_j + c\log(\theta_j + \epsilon)\}, \tag{16}$$

where $a_j = \alpha_j^2$, $b_j = 2\alpha_j\tilde{s}_j$ and $c = 4\sigma_w^2$. It follows that Eq. (16) can be simplified into a sequence of scalar minimization problems

$$\theta_j = \underset{\theta_j}{\operatorname{argmin}}\, a_j\theta_j^2 + b_j\theta_j + c\log(\theta_j + \epsilon), \tag{17}$$

The simplified scalar minimization problem can be solved by taking $\frac{df(\theta_j)}{d\theta_j} = 0$, where $f(\theta)$ is the right hand side of Eq. (17). By taking $\frac{df(\theta_j)}{d\theta_j} = 0$, one can obtain two stationary points - i.e.,

$$\theta_{j,1} = -\frac{b_j}{4a_j} + \sqrt{\frac{b_j^2}{16} - \frac{c}{2a_j}}, \theta_{j,2} = -\frac{b_j}{4a_j} - \sqrt{\frac{b_j^2}{16} - \frac{c}{2a_j}} \tag{18}$$

Case I) $b_j^2/(16a_j^2) - c/(2a_j) \geq 0$. In this case, the global minimization of Eq. (17) is given by comparing $f(0)$, $f(\theta_{j,1})$ and $f(\theta_{j,2})$;

Case II) $b_j^2/(16a_j^2) - c/(2a_j) < 0$. Note that there is no stationary point in the range of $[0, \infty)$. Since ϵ is a small positive constant, $g(0) = b_j + c/\epsilon$ is always positive. It follows that $f(0)$ is the global minimum for this case. Therefore the solution to Eq. (17) can be written as

$$\theta_j = \begin{cases} 0, & \text{if } b_j^2/(16a_j^2) - c/(2a_j) < 0, \\ t_j, & \text{otherwise} \end{cases} \tag{19}$$

where $t_j = \text{argmin}_{\theta_j}\{f(0), f(\theta_{j,1}), f(\theta_{j,2})\}$.

For the second subproblem, for a fixed $\boldsymbol{\theta}$, one can update the estimate of $\boldsymbol{\alpha}$ by solving

$$\boldsymbol{\alpha} = \underset{\boldsymbol{\alpha}}{\text{argmin}} \, ||\tilde{\boldsymbol{s}} - \boldsymbol{\Lambda}\boldsymbol{\alpha}||_2^2 + 2\sqrt{2}\sigma_w^2 \sum_j |\alpha_j|, \tag{20}$$

which also admits the following closed-form solution

$$\alpha_j = \mathcal{S}_{\tau_j}(\frac{\tilde{s}_j}{\theta_j}), \tag{21}$$

where $\mathcal{S}_{\tau_j}(\cdot)$ corresponds to the standard soft-thresholding function with a threshold of $\tau_j = \frac{\sqrt{2}\sigma_w^2}{\theta_j^2}$.

Combining the two sub-problems of Eqs. (15) and (20), we obtain a solution to Eq. (14) by the method of alternating optimization. Then the targeted tensor coefficients $\boldsymbol{s}$ can be given by $\hat{\boldsymbol{s}} = \hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{\alpha}}$, where $\hat{\boldsymbol{\Lambda}}$ and $\hat{\boldsymbol{\alpha}}$ denote the estimates of $\boldsymbol{\Lambda}$ and $\boldsymbol{\alpha}$ respectively. With recovered tensor coefficients, the reconstructed tensor can be finally obtained by

$$\hat{\mathcal{X}} = \hat{\mathcal{S}} \times_1 \mathbf{U} \times_2 \mathbf{V} \times_3 \mathbf{W}, \tag{22}$$

where $\hat{\mathcal{S}}$ is the coefficient array corresponding to $\hat{\boldsymbol{s}}$.

IV. Multiframe/Multispectral Image and Video Denoising with Robust Tensor Approximation

In this section, we apply the proposed robust tensor approximation method to multiframe image and video denoising. Suppose noisy observation data are denoted by $\mathcal{Y} = \mathcal{X} + \mathcal{N}$, where $\mathcal{X} \in \mathbb{R}^{H\times W\times L}$ and $\mathcal{N} \in \mathbb{R}^{H\times W\times L}$ correspond to unknown image/video and additive noise respectively. Let $\mathcal{Y}_i = \tilde{\mathcal{R}}_i\mathcal{Y}$ denote the 3^{rd} tensor formed by similar 3D patches, where $\tilde{\mathcal{R}}_i$ denotes an operator grouping similar patches into a 3^{rd}-order tensor[1]. It follows that the problem of denoising multiframe/multispecal images and video can be formulated as

$$\begin{aligned}(\mathcal{X}, \{\mathcal{S}_i\}) = &\underset{\mathcal{X},\{\mathbf{D}_i\},\{\mathcal{S}_i\}}{\text{argmin}} \; ||\mathcal{Y} - \mathcal{X}||_F^2 \\ &+ \eta \sum_i ||\tilde{\mathcal{R}}_i\mathcal{X} - \mathcal{S}_i \times_1 \mathbf{U}_i \times_2 \mathbf{V}_i \times_3 \mathbf{W}_i||_F^2 \\ &+ 2\sqrt{2}\sigma_w^2 \sum_i ||\boldsymbol{\Lambda}_i\boldsymbol{s}_i||_1 + 2\sigma_w^2 \sum_i \log\boldsymbol{\theta}_i,\end{aligned} \tag{23}$$

where $\mathbf{U}_i, \mathbf{V}_i, \mathbf{W}_i$ denotes the three orthogonal matrices calculated by HOSVD. The above global minimization problem can be decomposed into the following two sub-problems and solved by the method of alternating optimization again.

A. Solving for whole image/video

Let $\hat{\mathcal{X}}_i = \mathcal{X}_i \times_1 \mathbf{U}_i \times_2 \mathbf{V}_i \times_3 \mathbf{W}_i$ denote the reconstructed low-rank tensor with an initial estimate of $\mathcal{S}_i$. Then, for a fixed $\{\mathcal{S}_i\}$, the whole image or video $\mathcal{X}$ can be recovered by solving the following ℓ_2-minimization problem

$$\mathcal{X} = \underset{\mathcal{X}}{\text{argmin}} \, ||\mathcal{Y} - \mathcal{X}||_F^2 + \eta \sum_{i=1}^{N} ||\tilde{\mathcal{R}}_i\mathcal{X} - \hat{\mathcal{X}}_i||_F^2, \tag{24}$$

[1]Instead of forming a 4^{th} order tensor for the set of similar 3D patches, we found that combining them into a 3^{rd}-order tensor leads to more convenient implementation and slightly better denoising performance.

which is equivalent to the following equation (after reshaping tensors into long vectors)

$$\boldsymbol{x} = \underset{\boldsymbol{x}}{\text{argmin}} \, ||\boldsymbol{y} - \boldsymbol{x}||_2^2 + \eta \sum_{i=1}^{N} ||\tilde{\mathbf{R}}_i\boldsymbol{x} - \hat{\boldsymbol{x}}_i||_2^2, \tag{25}$$

where $\boldsymbol{y} \in \mathbb{R}^{H\cdot W\cdot L}, \boldsymbol{x} \in \mathbb{R}^{H\cdot W\cdot L}, \hat{\boldsymbol{x}} \in \mathbb{R}^{\sqrt{n}\cdot\sqrt{n}\cdot L}$ correspond to the vector representations of tensors $\mathcal{Y}, \mathcal{X}, \hat{\mathcal{X}}_i$ respectively; and $\tilde{\mathbf{R}}_i \doteq [\tilde{\mathbf{R}}_{i_0}, \tilde{\mathbf{R}}_{i_1}, \cdots, \tilde{\mathbf{R}}_{i_{m-1}}]$ the operator extracting similar patches. Eq.(25) can be solved in a closed-form by

$$\boldsymbol{x} = (\mathbf{I} + \eta \sum_{i=1}^{N} \tilde{\mathbf{R}}_i^\top \tilde{\mathbf{R}}_i)^{-1}(\boldsymbol{y} + \eta \sum_{i=1}^{N} \tilde{\mathbf{R}}_i^\top \hat{\boldsymbol{x}}_i), \tag{26}$$

where the matrix inversion can be easily calculated since it is diagonal. Similar to K-SVD [30], Eq. (26) can be computed by averaging over each set of reconstructed 3D patches $\hat{\mathcal{X}}_i$.

B. Solving for $\{\boldsymbol{s}_i\}$ and $\{\boldsymbol{\theta}_i\}$

For a fixed $\mathcal{X}$, Eq. (23) reduces to a sequence of low-rank tensor approximation problems - i.e., for each exemplar 3D image patch,

$$(\boldsymbol{s}_i, \boldsymbol{\theta}_i) = \underset{\boldsymbol{s}_i,\boldsymbol{\theta}_i}{\text{argmin}} \, ||\tilde{\boldsymbol{s}}_i - \boldsymbol{s}_i||_2^2 + 2\sqrt{2}\frac{\sigma_w^2}{\eta}||\boldsymbol{\Lambda}_i\boldsymbol{s}_i||_1 + 2\frac{\sigma_w^2}{\eta}\log\boldsymbol{\theta}_i, \tag{27}$$

where we have used $\tilde{\mathcal{S}}_i = \mathcal{X}_i \times_1 \mathbf{U}_i^\top \times_2 \mathbf{V}_i^\top \times_3 \mathbf{W}_i^\top$. This is exactly the problem studied in the previous section.

Putting things together, the proposed multiframe image/video denoising algorithm based on Robust Tensor Approximation with Laplacian Scale Mixture (RTA-LSM) is summarized in **Algorithm 1**. The *robustness* of **Algorithm 1** can be summarized as the following three aspects. First, patch grouping is updated at the outer loop - i.e., improved grouping results will lead to improved tensor approximation and vice versa. We have found that the inner iteration often converges in just few iterations ($J = 2$ in our implementation). Second, low-rank tensor approximation at the inner loop is based on a robust estimation of variance parameters using LSM models. Third, we have adopted the strategy of iterative regularization where noise fed back to the denoised image is controlled by a small positive parameter δ.

V. Experimental results

We have implemented the proposed algorithm under MATLAB. Multispectral images, MR image sequences and color videos are used to verify the performance of the proposed algorithm. To verify the performance of the propose LSM prior, we also implemented a reweighted ℓ_1 nonlocal low-rank tensor approximation (NLTA-reweighted ℓ_1) algorithm by iteratively computing $\boldsymbol{\theta}_i$ in Eq. (10) as $1/(|\hat{s}_j^{(k)}| + \epsilon)$ [63], where $\hat{s}_j^{(k)}$ is the estimate of the k-th iteration and ϵ is a small constant. The basic parameters are set as follows: block size $5 \times 5 \times L$ (L denotes the number of frames), the number of similar patches $m = 100$, the regularization parameter $\delta = 0.12$, and iteration numbers $K_{max} = 7$ and $J = 2$.

Algorithm 1 RTA-LSM based image/video denoising

• **Initialization**:

(a) Set the initial estimate $\hat{\mathcal{X}} = \mathcal{Y}$ and the parameter η;

(b) Obtain the set of tensors $\{\mathcal{X}_i\}$ from $\hat{\mathcal{X}}$ via k-NN search for each exemplar patch.

• **Outer loop**: for $k = 1, 2, \ldots, K_{max}$ **do**

(a) Tensor dataset $\mathcal{X}_i$ construction: grouping a set of similar 3D patches into a 3^{rd} tensor for each exemplar patch;

(b) **Inner loop** (Low-rank tensor approximation by solving Eq. (27)): for $j = 1, 2, \ldots, J$ **do**

(I) Compute $\boldsymbol{\theta}_i$ for fixed $\boldsymbol{\alpha}_i$ via Eq.(19);

(II) Compute $\boldsymbol{\alpha}_i$ for fixed $\boldsymbol{\theta}_i$ via Eq.(21);

(III) Output $\boldsymbol{s}_i = \text{diag}(\boldsymbol{\theta}_i)\boldsymbol{\alpha}_i$ if $j = J$.

End for

(c) Reconstruct $\{\mathcal{X}_i\}$ from $\{\mathcal{S}_i\}$ via Eq.(22).

(d) Reconstruct the whole image $\hat{\mathcal{X}}^{(k+1)}$ from $\{\mathcal{X}_i\}$ by solving Eq.(26).

(e) If $k < K_{max}$ set $\hat{\mathcal{X}}^{(k+1)} = \hat{\mathcal{X}}^{(k+1)} + \delta(\mathcal{Y} - \hat{\mathcal{X}}^{(k+1)})$

End for

• Output $\hat{\mathcal{X}}^{(k+1)}$

A. Multispectral image denoising

The whole CAVE database consisting of 32 hyperspectral images is used as the test set. The images of size $512\times512\times31$ are captured with the wavelengths in the range of $400-700$ nm at a interval of 10 nm. Two sets of experiments are conducted. In the first experiment, additive Gaussian noise with different standard derivations is added to the hyperspectral images; in the second experiment, mixed noise of additive Gaussian and Poisson is added as done in [24]. In the setting of mixed noise, standard derivations of Gaussian noise vary from 10 to 100, and the variance of Poisson noise is fixed at $\boldsymbol{y}/2^k$, wherein $k = 4$. We have compared the proposed method against several recently developed multispectral image denoising methods - including tensor dictionary learning (TensorDL) method [24],band-wise KSVD method [30], BM4D method [25], PARAFAC method [40], low-rank tensor approximation (LATA) method [39], KSVD method[30], ANLM3D method [45], band-wise BM3D method [34], the two-stage HOSVD denoising method [12] and KBRreg method [13][2]. Similar to [24], for the case of mixed noise we have applied the variance-stabilizing transformation (VST) [65] and its inverse to noisy spectral images before and after applying a test denoising method.

Average PSNR results at different noise levels are reported in Table I. From Table I, it can be seen that the proposed method consistently outperforms all other competing methods. The average PSNR improvements over KBRreg [13], HOSVD [12], and NLTA-reweighted ℓ_1 methods are larger than $3.22dB$, $1.25dB$ and $0.97dB$, respectively. In Fig. 2 and Fig. 3, we have compared the portions of denoised images at the 590nm band of *chart_and_stuffed_toy* with Gaussian noise of $\sigma_w = 30$ and the 680nm band of *egyptian_statue* with Gaussian noise of $\sigma_w = 50$. In Fig. 4 and Fig. 5,we show the visual comparisons at different noise levels of mixed noise of additive Gaussian and Poisson. It can be seen that the other test methods all produce visually annoying artifacts; by contrast, the denoised image by our proposed method contain much fewer artifacts. Surprisingly, we can also see that the reconstructed images by the proposed method are even sharper than the original images. This is due to the effective collaborative filtering across both the spectral and the spatial dimensions by the proposed method.

B. 3D MRI denoising

We have also applied the proposed method to 3D MRIs denoising. The $T1$-weighted 3D MRIs are obtained from the Brainweb database[3]. The 3D MRIs is of size $217 \times 181 \times 10$ with $1 \times 1 \times 1mm^3$ resolution. Additive Gaussian noise with different noise levels σ_w is added to simulate noisy 3D MRIs [4] The proposed method is compared with some recently developed 3D MRIs denoising methods, including the ANLM3D method [45], the band-wise BM3D method [34], and the BM4D method [25]. For the reason of completeness, LATA [39], TensorDL [24], HOSVD [12] and NLTA-Reweighted ℓ_1 developed for spectral image denoising are also included here.

Table II shows the PSNR result comparison among competing methods. From Table II we observe that the BM4D method [25] performs much better than the TensorDL method [24]. The reason is that the correlations between the slices are not strong and smaller 3D patches (i.e., $4 \times 4 \times 4$) used in BM4D can better exploit the local correlations. Even though the full slices 3D patches (i.e., $5 \times 5 \times 10$) are used, the proposed method still outperforms the BM4D [25] for all noise levels. The PSNR gain over BM4D method can be up to 1.04 dB. To facilitate visual comparison, we have shown portions of the reconstructed MRI by different methods in Fig. 6. It can be seen that the recovered image by our method is subjectively superior to those by other methods.

C. Color video denoising

We have also applied the proposed method to color video denoising. A total of 19 videos from the Arizona State University (ASU) dataset [5] are used as the test sequences. The selected test dataset involves different types of motion, including translation, rotation, scaling, etc. Additive Gaussian noise with different noise levels σ_w is added to simulate noisy input video. We have compared the proposed method against several recently developed color video denoising methods including frame-wise BM3D denoising (fBM3D) method [34], color BM3D(CMB3D) [66], VIDOSAT method [67] , VBM3D method [68], VBM4D method [25] and joint adaptive sparsity and low-rankness (SALT) method [69]. We have used their publicly released codes in our experimental comparison; for each video, the average PSNR results over 28 frames are calculated. Table III shows the comparison of average PSNR results among competing methods. From Table III, it can be seen that the proposed RTA-LSM consistently outperforms all other competing methods. Fig. 7, Fig. 8 and Fig. 9 include

[2]We thank the authors of [24], [30], [25], [40], [39], [45], [34], [12], [13] for providing their source codes in their websites.

[3]http://brainweb.bic.mni.mcgill.ca/brainweb/

[4]Our method can also be used for Rician noise by using the VST [65].

[5]https://media.xiph.org/video/derf/

TABLE I
AVERAGE PSNR RESULTS OF THE COMPETING METHODS FOR DIFFERENT NOISE LEVELS ON THE SET OF TEST HYPERSPECTRAL IMAGES.

method	Gaussian noise				
	$\sigma_w = 10$	$\sigma_w = 20$	$\sigma_w = 30$	$\sigma_w = 50$	$\sigma_w = 100$
BwKSVD [30]	33.69	30.46	28.76	26.78	24.00
BwBM3D [34]	40.20	36.49	34.35	31.84	28.42
KSVD [30]	35.76	32.61	30.81	28.64	26.64
LRTA [39]	39.37	36.10	34.11	31.56	28.04
PARAFAC [40]	30.87	30.73	30.55	29.33	26.10
TensorDL [24]	42.83	38.94	36.75	33.97	29.92
BM4D [25]	43.23	39.47	37.21	34.32	30.43
ANLM3D [45]	39.83	36.74	35.40	33.10	30.08
HOSVD [12]	46.04	42.54	40.29	37.27	32.89
KBRreg [13]	43.56	40.32	38.57	35.99	30.73
NLTA-Reweighted ℓ_1	45.52	42.41	40.48	37.78	34.27
Proposed RTA-LSM	**46.78**	**43.49**	**41.47**	**38.79**	**34.76**
method	Mixture of Poisson and Gaussian noise with fixed $k = 4$				
	$\sigma_w = 10$	$\sigma_w = 20$	$\sigma_w = 30$	$\sigma_w = 50$	$\sigma_w = 100$
BwKSVD [30]	29.13	27.56	26.09	23.76	20.11
BwBM3D [34]	33.94	31.84	31.65	30.03	27.13
KSVD [30]	30.86	29.8	28.54	25.92	22.27
LRTA [39]	33.60	33.00	31.93	30.44	27.76
PARAFAC [40]	31.30	28.64	27.31	23.86	20.36
TensorDL [24]	37.05	35.54	34.72	32.87	30.04
BM4D [25]	37.39	35.99	34.82	33.03	30.11
ANLM3D [45]	34.68	33.91	33.43	32.27	30.05
HOSVD [12]	40.17	38.95	37.79	35.98	32.68
KBRreg [13]	38.75	36.78	35.29	33.719	30.43
NLTA-Reweighted ℓ_1	40.17	39.37	38.46	36.79	33.82
Proposed RTA-LSM	**41.13**	**39.96**	**38.90**	**37.35**	**34.37**

the visual comparisons at different levels of additive Gaussian noise. Video frames denoised by ours are preferred to others in terms of better preserving image details.

D. Running time comparison

The proposed RTA-LSM algorithm has been implemented under Matlab. The running time comparison between the proposed algorithm and other competing methods is reported in Table IV. The test sequence is *Claire* in the ASU dataset and we have selected 28 frames from that test sequence in our experiment. For color video denoising, the proposed RTA-LSM algorithm is not as efficient as VBM3D and VBM4D (note that they have been optimized by C-coded MEX compiler) but is generally 3-5 times faster than the other competing method SALT. In our current implementation, the computational bottleneck of the proposed method lies in block-matching and low-rank tensor approximation both of which can be implemented more efficiently using C instead of Matlab. Moreover, if tensor decomposition for each set of exemplar patches can be performed in parallel (e.g., implemented on GPU), the running time of the proposed algorithm can be further greatly reduced.

VI. CONCLUSIONS

In this paper we have proposed a robust sparse/low-rank tensor approximation approach for multiframe/multispectral image and video denoising. To fully exploit the spatio-temporal dependency, we group similar 3D image patches into 3^{rd} tensors, which lend themselves to be approximated by low-rank tensors. For the purpose of making low-rank tensor approximation more robust, we propose a new regularization term for sparse coefficients using Laplacian scale mixture model. LSM modeling translates low-rank tensor approximation into an optimization problem with sparse coefficients and hidden scalar variables. We have adopted the method of alternating optimization and shown that both subproblems can be solved in closed-form. Experimental results on both hyperspectral/multiframe images and color video show that the proposed method performs significantly better than existing methods in the open literature.

REFERENCES

[1] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, P. Boesiger *et al.*, "Sense: sensitivity encoding for fast mri," *Magnetic Resonance in Medicine*, vol. 42, no. 5, pp. 952–962, 1999.

[2] J. Tsao, P. Boesiger, and K. P. Pruessmann, "k-t blast and k-t sense: Dynamic mri with high frame rate exploiting spatiotemporal correlations," *Magnetic Resonance in Medicine*, vol. 50, no. 5, pp. 1031–1042, 2003.

[3] J. Polzehl and V. G. Spokoiny, "Functional and dynamic magnetic resonance imaging using vector adaptive weights smoothing," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 50, no. 4, pp. 485–501, 2001.

[4] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann, 2010.

[5] K. Momma and F. Izumi, "Vesta 3 for three-dimensional visualization of crystal, volumetric and morphology data," *Journal of applied crystallography*, vol. 44, no. 6, pp. 1272–1276, 2011.

[6] E. Kurutepe, M. R. Civanlar, and A. M. Tekalp, "Client-driven selective streaming of multiview video for interactive 3dtv," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 11, pp. 1558–1565, 2007.

[7] R. A. Harshman, "Foundations of the parafac procedure: Models and conditions for an" explanatory" multimodal factor analysis," 1970.

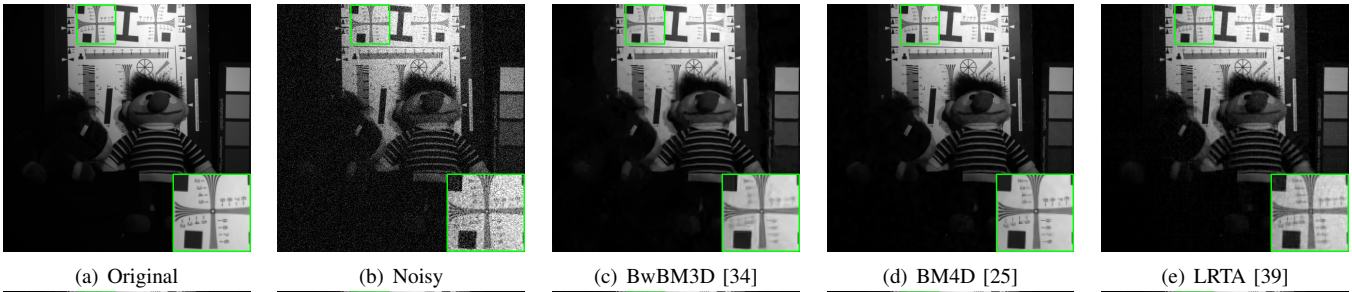
[8] R. Merris, *Multilinear algebra*. CRC Press, 1997, vol. 8.

[9] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.

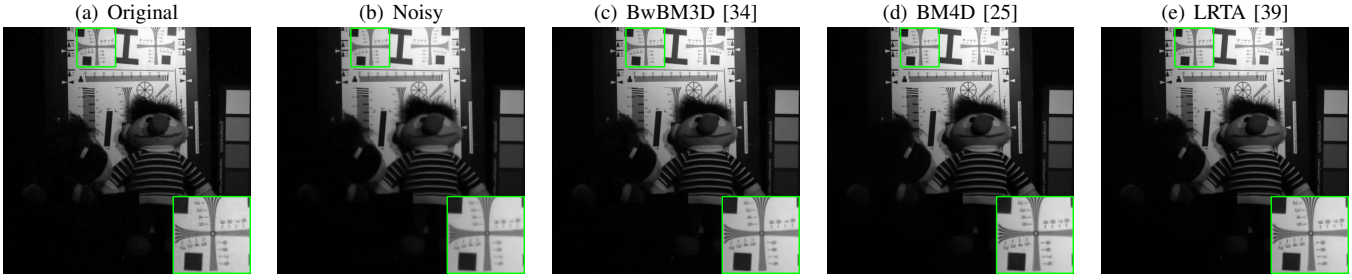
[10] R. P. McDonald, *Factor analysis and related methods*. Psychology Press, 2014.

[11] W. Dong, G. Li, G. Shi, X. Li, and Y. Ma, "Low-rank tensor approximation with laplacian scale mixture modeling for multiframe image denoising," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 442–449.

[12] A. Rajwade, A. Rangarajan, and A. Banerjee, "Image denoising using the higher order singular value decomposition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 35, no. 4, pp. 849C–862, 2013.

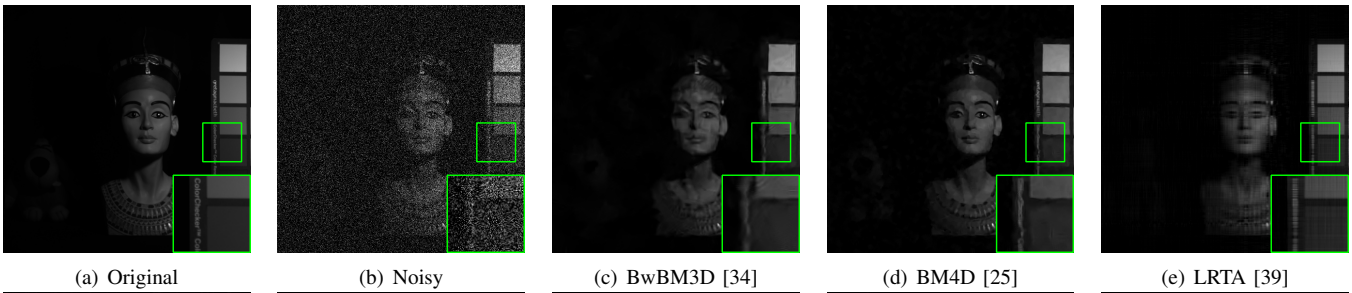


(a) Original (b) Noisy (c) BwBM3D [34] (d) BM4D [25] (e) LRTA [39]

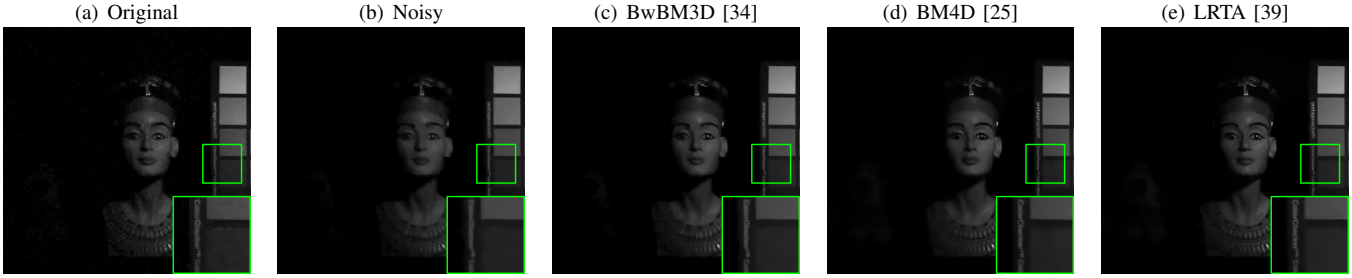


(a) TDL [24] (b) HOSVD [12] (c) KBRreg [13] (d) NLTA-Reweighted ℓ_1 (e) **RTA-LSM**

Fig. 2. (a)Original images at 590nm band of *chart_and_stuffed_toy* in CAVE dataset [70]; (b) The images corrupted by Gussian noise of $\sigma_w = 30$ (PSNR=20.48dB); (c) BwBM3D [34] (PSNR=33.24dB); (d) BM4D [25] (PSNR=36.51dB); (e) LRTA [39] (PSNR=33.42dB); (f) TDL [24] (PSNR=37.30dB); (g) HOSVD [12] (PSNR=34.27dB); (h) KBRreg [13] (PSNR=39.11dB); (i) NLTA-Reweighted ℓ_1 (PSNR=39.70dB); (j)**Proposed RTA-LSM** (PSNR=**40.21**dB).



(a) Original (b) Noisy (c) BwBM3D [34] (d) BM4D [25] (e) LRTA [39]



(a) TDL [24] (b) HOSVD [12] (c) KBRreg [13] (d) NLTA-Reweighted ℓ_1 (e) **RTA-LSM**

Fig. 3. (a)Original images at 680nm band of *egyptian_statue* in CAVE dataset [70]; (b) The images corrupted by Gussian noise of $\sigma_w = 50$ (PSNR=16.77dB); (c) BwBM3D [34] (PSNR=34.44dB); (d) BM4D [25] (PSNR=37.90dB); (e) LRTA [39] (PSNR=33.65dB); (f) TDL [24] (PSNR=36.49dB); (g) HOSVD [12] (PSNR=35.51dB); (h) KBRreg [13] (PSNR=37.46dB); (i) NLTA-Reweighted ℓ_1 (PSNR=38.46dB); (j)**Proposed RTA-LSM** (PSNR=**39.37**dB).

[13] Q. Xie, Q. Zhao, D. Meng, and Z. Xu, "Kronecker-basis-representation based tensor sparsity and its applications to tensor recovery," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.

[14] M. Piccardi, "Background subtraction techniques: a review," in *IEEE International Conference on Systems, Man and Cybernetics*, vol. 4, 2004, pp. 3099–3104.

[15] E. J. Candes, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM (JACM)*, vol. 58, no. 3, p. 11, 2011.

[16] T. Bouwmans and E. H. Zahzah, "Robust pca via principal component pursuit: A review for a comparative evaluation in video surveillance," *Computer Vision and Image Understanding*, vol. 122, pp. 22–34, 2014.

[17] C. Guyon, T. Bouwmans, and E.-h. Zahzah, "Robust principal component analysis for background subtraction: Systematic evaluation and comparative analysis," in *Principal component analysis*. InTech, 2012.

[18] F. De La Torre and M. J. Black, "A framework for robust subspace learning," *International Journal of Computer Vision*, vol. 54, no. 1-3, pp. 117–142, 2003.

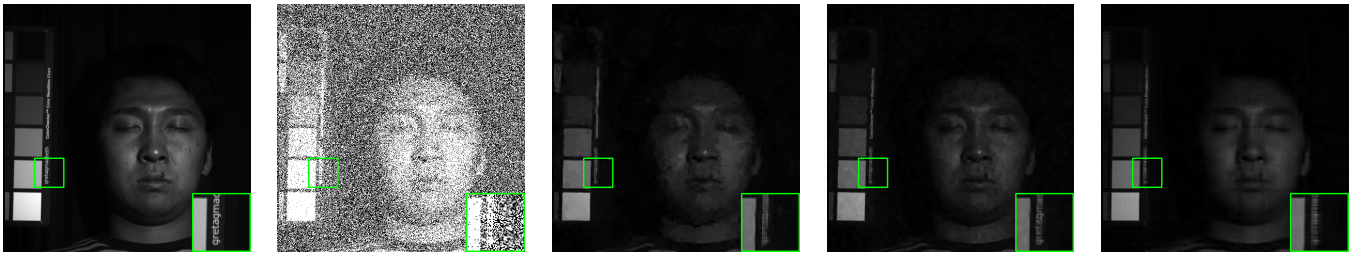
[19] M. Soltanolkotabi, E. Elhamifar, E. J. Candes *et al.*, "Robust subspace clustering," *The Annals of Statistics*, vol. 42, no. 2, pp. 669–699, 2014.

[20] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma, "Rasl: Robust alignment by sparse and low-rank decomposition for linearly correlated images," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 11, pp. 2233–2246, 2012.

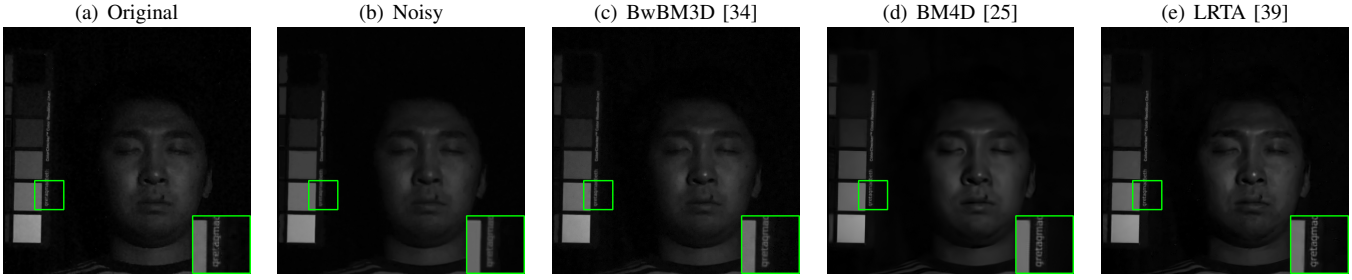
[21] P. Garrigues and B. A. Olshausen, "Group sparse coding with a laplacian scale mixture prior," in *Advances in neural information processing systems*, 2010, pp. 676–684.

[22] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[23] Y. Chen, Y. Wang, M. Li, and G. He, "Augmented lagrangian alternating direction method for low-rank minimization via non-convex approxima-

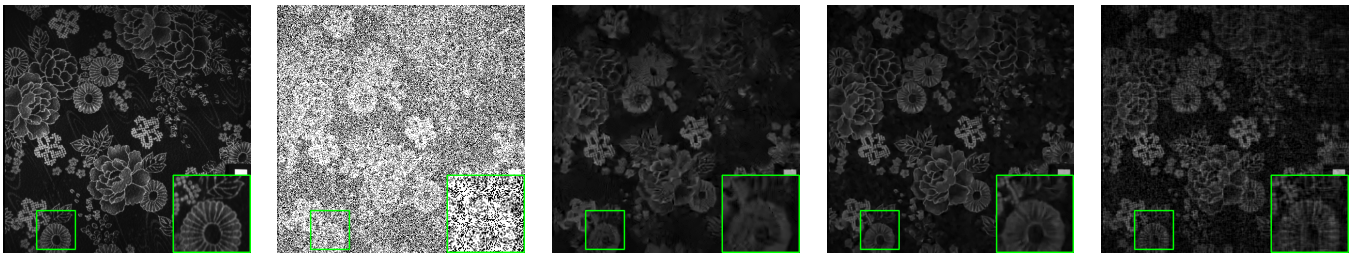


(a) Original (b) Noisy (c) BwBM3D [34] (d) BM4D [25] (e) LRTA [39]

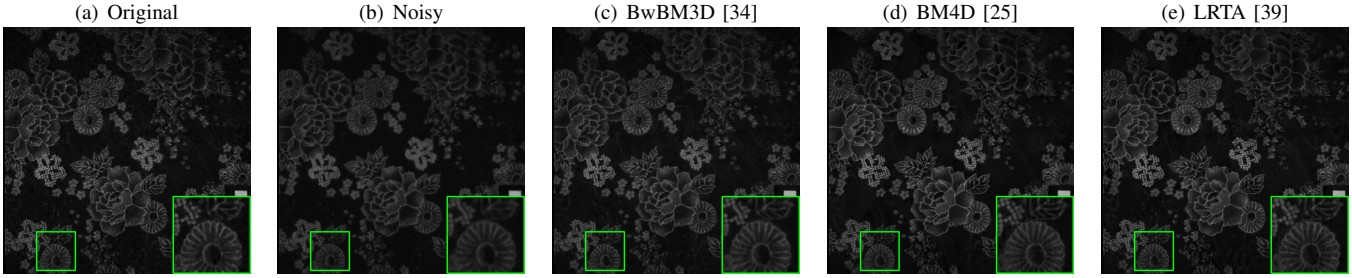


(a) TDL [24] (b) HOSVD [12] (c) KBRreg [13] (d) NLTA-Reweighted ℓ_1 (e) **RTA-LSM**

Fig. 4. (a)Original images at 440nm band of *face_ms* in CAVE dataset [70]; (b) The images corrupted by mixture of Poission and Gussian noise with fixed $k = 4$ and $\sigma_w = 30$ (PSNR=4.04dB); (c) BwBM3D [34] (PSNR=23.03dB); (d) BM4D [25] (PSNR=22.91dB); (e) LRTA [39] (PSNR=22.82dB); (f) TDL [24] (PSNR=22.90dB); (g) HOSVD [12] (PSNR=22.42dB); (h) KBRreg [13] (PSNR=23.02dB); (i) NLTA-Reweighted ℓ_1 (PSNR=23.01dB); (j)**Proposed RTA-LSM** (PSNR=**23.12**dB).



(a) Original (b) Noisy (c) BwBM3D [34] (d) BM4D [25] (e) LRTA [39]



(a) TDL [24] (b) HOSVD [12] (c) KBRreg [13] (d) NLTA-Reweighted ℓ_1 (e) **RTA-LSM**

Fig. 5. (a)Original images at 440nm band of *cloth_ms* in CAVE dataset [70]; (b) The images corrupted by mixture of Poission and Gussian noise with fixed $k = 4$ and $\sigma_w = 50$ (PSNR=3.79dB); (c) BwBM3D [34] (PSNR=20.20dB); (d) BM4D [25] (PSNR=21.44dB); (e) LRTA [39] (PSNR=20.63dB); (f) TDL [24] (PSNR=21.61dB); (g) HOSVD [12] (PSNR=19.70dB); (h) KBRreg [13] (PSNR=21.50dB); (i) NLTA-Reweighted ℓ_1 (PSNR=21.86dB); (j)**Proposed RTA-LSM** (PSNR=**22.11**dB).

tion," *Signal, Image and Video Processing*, vol. 11, no. 7, pp. 1271–1278, 2017.

[24] Y. Peng, D. Meng, Z. Xu, C. Gao, Y. Yang, and B. Zhang, "Decomposable nonlocal tensor dictionary learning for multispectral image denoising," in *Proc. of the IEEE CVPR*, 2014, pp. 4321–4328.

[25] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE Trans. on Image Process.*, vol. 22, no. 1, pp. 119–133, Jan. 2013.

[26] I. Markovsky, *Low rank approximation: algorithms, implementation, applications*. Springer Science & Business Media, 2011.

[27] L. Grasedyck, D. Kressner, and C. Tobler, "A literature survey of low-rank tensor approximation techniques," *GAMM-Mitteilungen*, vol. 36, no. 1, pp. 53–78, 2013.

[28] D. Goldfarb and Z. Qin, "Robust low-rank tensor recovery: Models and algorithms," *SIAM Journal on Matrix Analysis and Applications*, vol. 35, no. 1, pp. 225–253, 2014.

[29] X. Zhou, C. Yang, H. Zhao, and W. Yu, "Low-rank modeling and its applications in image analysis," *ACM Computing Surveys (CSUR)*, vol. 47, no. 2, p. 36, 2015.

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

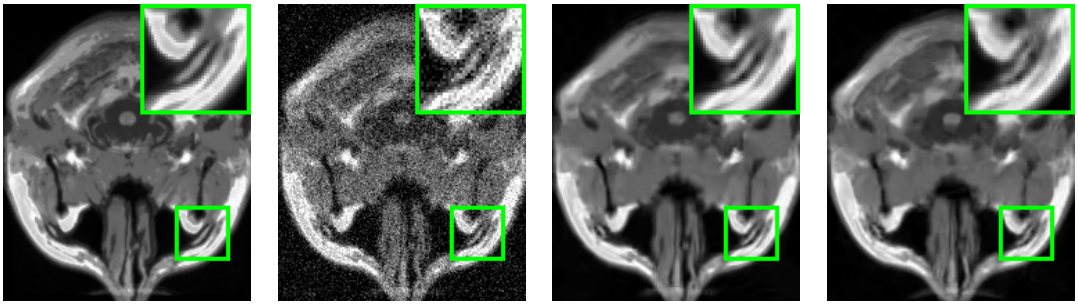
[31] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. on Image Processing*, vol. 17, no. 1, pp. 53–69, Jan. 2008.

[32] L. Zhang, W. Dong, D. Zhang, and G. Shi, "Two-stage image denoising by principal component analysis with local pixel grouping," *Pattern Recognition*, vol. 43, pp. 1531–1549, Apr. 2010.

[33] A. Buades, B. Coll, and J. M. Morel, "A non-local algorithm for image

TABLE II
THE PSNR RESULTS OF COMPETING METHODS FOR ADDITIVE GAUSSIAN NOISE ON THE 3D MRIS.

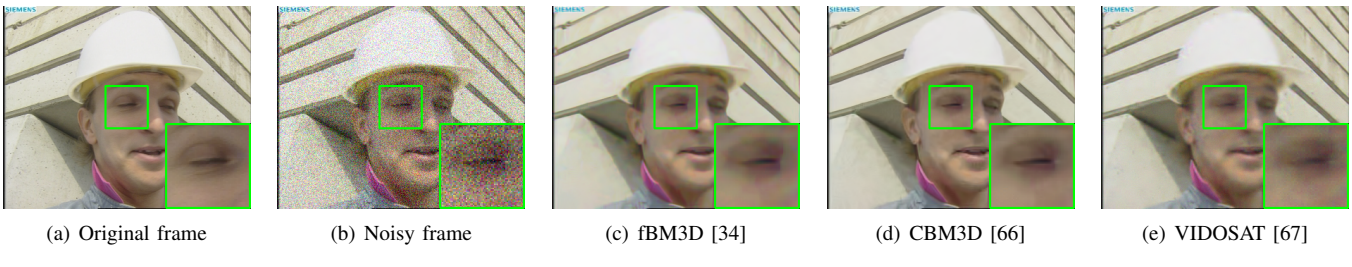
method	3D MRIs				
	$\sigma_w = 10$	$\sigma_w = 20$	$\sigma_w = 30$	$\sigma_w = 50$	$\sigma_w = 100$
ANLM3D [45]	32.85	29.11	27.30	25.49	23.59
LRTA [39]	32.79	28.72	27.08	24.50	21.59
BwBM3D [34]	35.42	31.81	29.78	27.33	23.94
BM4D [25]	36.37	32.85	30.89	28.49	25.39
TensorDL [24]	33.63	30.03	28.48	26.18	23.66
HOSVD [12]	36.46	33.17	31.21	28.75	25.45
NLTA-Reweighted ℓ_1	35.18	31.79	30.28	27.88	25.45
Proposed RTA-LSM	**37.06**	**33.62**	**31.83**	**29.53**	**26.41**



(a) Ori image (b) Noisy image (c) BwBM3D [34] (d) BM4D [25]

(e) TensorDL [24] (f) HOSVD [12] (g) NLTA-Reweighted ℓ_1 (h) **RTA-LSM**

Fig. 6. (a) The original MRI (the 3^{rd} slice); (b) The noisy MRI ($\sigma_w = 30$, PSNR=18.58dB); (c) BwBM3D [34] (PSNR=29.78dB); (d) BM4D [25] (PSNR=30.89dB); (e) TensorDL [24] (PSNR=28.48dB); (f) HOSVD [12] (PSNR=31.21dB); (g) NLTA-Reweighted ℓ_1 (PSNR=30.28dB); (h) **Proposed RTA-LSM** (PSNR= **31.83**dB).



(a) Original frame (b) Noisy frame (c) fBM3D [34] (d) CBM3D [66] (e) VIDOSAT [67]

(f) VBM3D [68] (g) VBM4D [25] (h) SALT [69] (i) NLTA-Reweighted ℓ_1 (j) **RTA-LSM**

Fig. 7. one frame extracted from Sequence Foreman($\sigma_w = 30$):(a) the original frame ; (b) the noisy frame (PSNR=18.91dB); (c) fBM3D [34] (PSNR=31.53dB); (d) CBM3D [66] (PSNR=32.98dB); (e) VIDOSAT [67] (PSNR=31.24dB); (f) VBM3D [68] (PSNR=32.27dB); (g) VBM4D [25] (PSNR=32.64dB);(h) SALT [69] (PSNR=33.14dB); (i) NLTA-Reweighted ℓ_1 (PSNR=34.30dB); (j) **Proposed RTA-LSM** (PSNR=**34.52**dB).

TABLE III
THE PSNR RESULTS OF COMPETING METHODS FOR ADDITIVE GAUSSIAN NOISE ON THE ASU DATASET.

method	ASU dataset (19 videos)			
	$\sigma_w = 10$	$\sigma_w = 30$	$\sigma_w = 50$	$\sigma_w = 100$
fBM3D [34]	34.92	29.13	26.09	21.07
CBM3D [66]	36.57	30.59	27.19	21.30
VIDOSAT [67]	37.67	31.26	28.09	21.53
VBM3D [68]	37.61	31.64	27.87	20.80
VBM4D [25]	38.19	31.91	28.86	24.64
SALT [69]	**38.37**	32.00	28.26	21.70
NLTA-Reweighted ℓ_1	38.16	32.35	29.64	25.48
Proposed RTA-LSM	38.11	**32.61**	**30.08**	**26.84**

(a) Original frame (b) Noisy frame (c) fBM3D [34] (d) CBM3D [66] (e) VIDOSAT [67]

(f) VBM3D [68] (g) VBM4D [25] (h) SALT [69] (i) NLTA-Reweighted ℓ_1 (j) **RTA-LSM**

Fig. 8. one frame extracted from Sequence Bus($\sigma_w = 50$):(a) The original frame ; (b) The noisy frame (PSNR=15.13dB); (c) fBM3D [34] (PSNR=23.56dB); (d) CBM3D [66] (PSNR=25.32dB); (e) VIDOSAT [67] (PSNR=23.59dB); (f) VBM3D [68] (PSNR=23.80dB); (g) VBM4D [25] (PSNR=25.93dB);(h) SALT [69] (PSNR=24.89dB); (i) NLTA-Reweighted ℓ_1 (PSNR=27.02dB); (j) **Proposed RTA-LSM** (PSNR= **27.32**dB).

(a) Original frame (b) Noisy frame (c) fBM3D [34] (d) CBM3D [66] (e) VIDOSAT [67]

(f) VBM3D [68] (g) VBM4D [25] (h) SALT [69] (i) NLTA-Reweighted ℓ_1 (j) **RTA-LSM**

Fig. 9. one frame extracted from Sequence News($\sigma_w = 100$):(a) the original frame ; (b) the noisy frame (PSNR=10.53dB); (c) fBM3D [34] (PSNR=20.33dB); (d) CBM3D [66] (PSNR=20.57dB); (e) VIDOSAT [67] (PSNR=20.91dB); (f) VBM3D [68] (PSNR=20.05dB); (g) VBM4D [25] (PSNR=24.92dB);(h) the SALT [69] (PSNR=20.93dB); (i) NLTA-Reweighted ℓ_1 (PSNR=24.59dB); (j) **Proposed RTA-LSM** (PSNR= **27.32**dB).

denoising," in *IEEE Int. Conf. on CVPR*, vol. 2, 2005, pp. 60–65.

[34] 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.

[35] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *Proc. of the IEEE ICCV*, Tokyo, Japan, 2009.

[36] W. Dong, X. Li, L. Zhang, and G. Shi, "Sparsity-based image denoising via dictionary learning and structural clustering," in *Proc. of the IEEE CVPR*, 2011, pp. 457–464.

[37] W. Dong, G. Shi, and X. Li, "Nonlocal image restoration with bilateral variance estimation: a low-rank approach," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 700–711, Feb. 2013.

[38] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm

TABLE IV
RUNNING TIME(SEC) ON 28 FRAMES OF THE SEQUENCE CLAIRE(EXPERIMENTS PERFORMED ON INTEL CORE i7-7700 CPU)

method	28 frames of the Sequence Claire($144 \times 176 \times 3$)			
	$\sigma_w = 10$	$\sigma_w = 30$	$\sigma_w = 50$	$\sigma_w = 100$
fBM3D [34]	20.6	20.7	25.5	26.0
CBM3D [66]	11.0	10.4	17.9	17.3
VIDOSAT [67]	159.1	81.6	317.5	315.1
VBM3D [68]	4.6	4.5	4.4	3.8
VBM4D [25]	38.5	39.1	39.9	40.3
SALT [69]	2235.0	2048.2	2362.0	2252.1
NLTA-Reweighted ℓ_1	479.7	483.4	720.1	726.9
Proposed RTA-LSM	488.2	492.5	741.1	734.1

minimization with application to image denoising," in *Proc. of the IEEE CVPR*, 2014.

[39] N. Renard, S. Bourennane, and J. Blanc-Talon, "Denoising and dimensionality reduction using multilinear tools for hyperspectral images," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 2, pp. 138–142, 2008.

[40] X. Liu, S. Bourennane, and C. Fossati, "Denoising of hyperspectral images using the parafac model and statistical performance analysis," *IEEE Trans. Geoscience and Remote Sensing*, vol. 50, no. 10, pp. 3717C–3724, Oct. 2012.

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

[42] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2744–2756, Nov. 2002.

[43] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Trans. on Signal Process.*, vol. 12, no. 11, pp. 1338C–1351, Nov. 2003.

[44] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. on Image Process.*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009.

[45] J. V. Manjon, P. Coupe, L. Marti-Bonmati, L. Collins, and M. Robles, "Adaptive non-local means denoising of mr images with spatially varying noise levels," *Journal of Magnetic Resonance Imaging*, vol. 31, no. 1, pp. 192–203, 2010.

[46] H. Ji, C. Liu, Z. Shen, and Y. Xu, "Robust video denoising using low rank matrix completion," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 1791–1798.

[47] C. Liu and W. T. Freeman, "A high-quality video denoising algorithm based on reliable motion estimation," in *European Conference on Computer Vision*. Springer, 2010, pp. 706–719.

[48] G. Varghese and Z. Wang, "Video denoising based on a spatiotemporal gaussian scale mixture model," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 7, pp. 1032–1040, 2010.

[49] C. Sutour, C.-A. Deledalle, and J.-F. Aujol, "Adaptive regularization of the nl-means: Application to image and video denoising," *IEEE Transactions on Image Processing*, vol. 23, no. 8, pp. 3506–3521, 2014.

[50] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, "Video denoising, deblocking, and enhancement through separable 4-d nonlocal spatiotemporal transforms," *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 3952–3966, 2012.

[51] A. Buades, J. L. Lisani, and M. Miladinovic, "Patch-based video denoising with optical flow estimation," *IEEE Transactions on Image Processing*, vol. 25, no. 6, pp. 2573–2586, 2016.

[52] T. Ehret, P. Arias, and J.-M. Morel, "Global patch search boosts video denoising," in *International Conference on Computer Vision Theory and Applications*, vol. 5, 2017, pp. 124–134.

[53] B. Wen, Y. Li, L. Pfister, and Y. Bresler, "Joint adaptive sparsity and low-rankness on the fly: an online tensor reconstruction scheme for video denoising," in *IEEE International Conference on Computer Vision (ICCV)*, 2017.

[54] L. Gui, G. Cui, Q. Zhao, D. Wang, A. Cichocki, and J. Cao, "Video denoising using low rank tensor decomposition," in *Ninth International Conference on Machine Vision (ICMV 2016)*, vol. 10341, 2017, p. 103410V.

[55] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *12th International Conference on Computer Vision (ICCV)*. IEEE, 2009, pp. 2272–2279.

[56] J. Huang and T. Zhang, "The benefit of group sparsity," *The Annals of Statistics*, vol. 38, no. 4, pp. 1978–2004, 2010.

[57] R. L. Liboff, *Introductory quantum mechanics*. Pearson Education India, 2003.

[58] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[59] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," in *Advances in neural information processing systems*, 2011, pp. 612–620.

[60] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Applicat.*, vol. 21, no. 4, pp. 1253–1278, 2000.

[61] G. Bergqvist and E. G. Larsson, "The higher-order singular value decomposition: Theory and an application," *IEEE Signal Processing Magazine*, pp. 151–154, May 2010.

[62] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM review*, vol. 52, no. 3, pp. 471–501, 2010.

[63] E. Candes, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted l_1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 877–905, 2008.

[64] W. Dong, G. Shi, Y. Ma, and X. Li, "Image restoration via simultaneous sparse coding: Where structured sparsity meets gaussian scale mixture," *International Journal of Computer Vision*, vol. 114, no. 2-3, pp. 217–232, 2015.

[65] M. Makitalo and A. Foi, "Optimal inversion of the generalized anscombe transformation for poisson-gaussian noise," *IEEE Trans. on Image Process.*, vol. 22, no. 1, pp. 91–103, 2013.

[66] 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 *IEEE International Conference on Image Processing,*, vol. 1, Sep. 2007.

[67] B. Wen, S. Ravishankar, and Y. Bresler, "Video denoising by online 3d sparsifying transform learning," in *Proc. of the IEEE ICIP*. IEEE, 2015, pp. 118–122.

[68] D. Kostadin, F. Alessandro, and E. KAREN, "Video denoising by sparse 3d transform-domain collaborative filtering," in *European Signal Processing Conference*, vol. 149. Tampere, Finland, 2007.

[69] B. Wen, Y. Li, L. Pfister, and Y. Bresler, "Joint adaptive sparsity and low-rankness on the fly: an online tensor reconstruction scheme for video denoising," in *IEEE Proc. of ICCV*, 2017.

[70] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," *IEEE Trans. on Image Process.*, vol. 19, no. 9, pp. 2241–2253, 2010.