

An Efficient SVD-Based Method for Image Denoising

Qiang Guo, Caiming Zhang, Yunfeng Zhang, and Hui Liu

Abstract—Nonlocal self-similarity of images has attracted considerable interest in the field of image processing and has led to several state-of-the-art image denoising algorithms, such as block matching and 3-D, principal component analysis with local pixel grouping, patch-based locally optimal wiener, and spatially adaptive iterative singular-value thresholding. In this paper, we propose a computationally simple denoising algorithm using the nonlocal self-similarity and the low-rank approximation (LRA). The proposed method consists of three basic steps. First, our method classifies similar image patches by the block-matching technique to form the similar patch groups, which results in the similar patch groups to be low rank. Next, each group of similar patches is factorized by singular value decomposition (SVD) and estimated by taking only a few largest singular values and corresponding singular vectors. Finally, an initial denoised image is generated by aggregating all processed patches. For low-rank matrices, SVD can provide the optimal energy compaction in the least square sense. The proposed method exploits the optimal energy compaction property of SVD to lead an LRA of similar patch groups. Unlike other SVD-based methods, the LRA in SVD domain avoids learning the local basis for representing image patches, which usually is computationally expensive. The experimental results demonstrate that the proposed method can effectively reduce noise and be competitive with the current state-of-the-art denoising algorithms in terms of both quantitative metrics and subjective visual quality.

Index Terms—Back projection, image denoising, low-rank approximation (LRA), patch grouping, self-similarity, singular value decomposition (SVD).

Manuscript received September 9, 2014; revised December 26, 2014 and February 8, 2015; accepted March 18, 2015. Date of publication March 25, 2015; date of current version May 3, 2016. This work was supported in part by the National Natural Science Foundation of China under Grant 61202150, Grant 61272245, Grant 61373078, Grant 61332015, and Grant 61472220; in part by the National Natural Science Foundation of China Joint Fund with Guangdong under Grant U1201258; in part by the China Post-Doctoral Science Foundation under Grant 2013M531600; and in part by the Program for Scientific Research Innovation Team in Colleges and Universities of Shandong Province. This paper was recommended by Associate Editor A. Kaup.

Q. Guo is with the School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan 250014, China, and with Shandong Provincial Key Laboratory of Digital Media Technology, Jinan 250014, China, and also with the School of Computer Science and Technology, Shandong University, Jinan 250100, China (e-mail: guoqiang@sdufe.edu.cn).

C. Zhang is with the School of Computer Science and Technology, Shandong University, Jinan 250100, China, and also with Shandong Provincial Key Laboratory of Digital Media Technology, Jinan 250014, China (e-mail: czhang@sdu.edu.cn).

Y. Zhang and H. Liu are with the School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan 250014, China, and also with Shandong Provincial Key Laboratory of Digital Media Technology, Jinan 250014, China.

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

Digital Object Identifier 10.1109/TCSVT.2015.2416631

I. INTRODUCTION

DURING acquisition and transmission, images are inevitably contaminated by noise. As an essential and important step to improve the accuracy of the possible subsequent processing, image denoising is highly desirable for numerous applications, such as visual enhancement, feature extraction, and object recognition [1], [2].

The purpose of denoising is to reconstruct the original image from its noisy observation as accurately as possible, while preserving important detail features such as edges and textures in the denoised image. To achieve this goal, over the past several decades, image denoising has been extensively studied in the signal processing community, and numerous denoising techniques have been proposed in the literature. In general, denoising algorithms can be roughly classified into three categories: 1) spatial domain methods; 2) transform domain methods; and 3) hybrid methods [3], [4]. The first class utilizes the spatial correlation of pixels to smooth the noisy image, the second one exploits the sparsity of representation coefficients of the signal to distinguish the signal and noise, and the third one takes advantage of spatial correlation and sparse representation to suppress noise.

Spatial domain methods, also called spatial filters, estimate each pixel of the image by performing a weighted average of its local/nonlocal neighbors, in which the weights can be determined by their similarities and higher weights are given to similar pixels. Therefore, spatial filters can be further divided into local filters and nonlocal filters. Smith and Brady [5] proposed a structure preserving local filter called SUSAN, which uses the intensity distance as a quantitative measure of the similarity between pixels. Tomasi and Manduchi [6] proposed bilateral filtering by generalizing the SUSAN filter, in which both the intensity and spatial distances are used to measure the similarity between pixels. Although these local filters are effective for preserving edges, they cannot perform very well when the noise level is high. The reason is that the severe noise destroys the correlations of pixels within local regions [7]. To overcome this disadvantage of local filters, Buades *et al.* [8] proposed the nonlocal mean (NLM) filter, which estimates each pixel by a nonlocal averaging of all the pixels in the image. The amount of weighting for a pixel is based on the Euclidean distance between the patch centered around the pixel being denoised and the one centered around a given neighboring pixel. In essence, NLM uses the structural redundancy, namely, self-similarity that is inherent in natural images, to estimate each pixel. NLM can be considered as an extension of the bilateral filter by the means of replacing

pointwise photometric distances with patch distances. Several variants of NLM have been proposed to improve the adaptivity of the nonlocal filter [9], [10]. Talebi *et al.* [3] proposed a spatially adaptive iterative filtering (SAIF) to improve the performance of NLM. Recently, there has been a growing interest in exploiting the self-similarity of images to suppress noise. Chatterjee and Milanfar [11], [12] proposed a patch-based locally optimal wiener (PLOW) filter, which also exploits the structural redundancy for image denoising and achieves the near optimal performance in the minimum mean-squared error (MMSE) sense. Zhang *et al.* [13] proposed a two-direction nonlocal (TDNL) variational model for image denoising using the horizontal and vertical similarities in the matrix formed by similar image patches. SAIF, PLOW, and TDNL are currently considered to be state of the art in spatial domain denoising methods.

Transform domain methods assume that the image can be sparsely represented by some representation basis, such as wavelet basis and its directional extensions. Due to the sparsity of representation coefficients, noise is uniformly spread throughout the coefficients in the transform domain, while most of image information is concentrated on the few largest ones. Therefore, noise can be effectively distinguished by different coefficient shrinkage strategies, including *BayesShrink* [14], *ProbShrink* [15], *BiShrink* [16], *MultiShrink* [17], and *SUREShrink* [18], [19]. Despite its remarkable success in dealing with point and line singularities, the fixed wavelet transform fails to provide an adaptive sparse representation for the image containing complex singularities. To overcome the problems caused using the fixed transforms, Aharon *et al.* [20] proposed an adaptive representation technique using K-means and singular value decomposition (called K-SVD), which uses a greedy algorithm to learn an overcomplete dictionary for image representation and denoising. Under the assumption that each image patch can be represented by the learned dictionary, Elad and Aharon [21] proposed a K-SVD based denoising algorithm, in which each image patch can be expressed as a linear combination of few atoms of the dictionary. Although the dictionary-based methods are more robust to noise, they are computationally expensive.

Spatial-based filters and transform-based filters have achieved great success in image denoising. Their overall performance, however, does not generally surpass the hybrid methods. Due to its impressive performance, the most well-known hybrid method for image denoising is the block-matching and 3-D (BM3D) filtering reported in [22], which groups similar patches into 3-D arrays and deals with these arrays by sparse collaborative filtering. To the best of the authors' knowledge, it is the first one that utilizes both nonlocal self-similarity and sparsity for image denoising. However, the fixed 3-D transform is not able to deliver a sparse representation for image patches containing edges, singularities, or textures. Thus, BM3D may introduce visual artifacts. Dabov *et al.* [23] proposed an improved BM3D filter (called BM3D-SAPCA) that exploits adaptive-shape patches and principal component analysis (PCA). Although BM3D-SAPCA achieves state-of-the-art denoising results,

its computational cost is very high (Table III). Zhang *et al.* [24] proposed an adaptive image denoising scheme using PCA with local pixel grouping (LPG-PCA). This method uses block matching to group the pixels with similar local structures, transforms each group of pixels using locally learned PCA basis, and shrinks PCA transformation coefficients using the linear MMSE estimation technique. Both LPG-PCA and BM3D-SAPCA use the PCA basis to represent image patches. A key difference between them is that LPG-PCA applies PCA on 2-D groups of fixed-size image patches, while BM3D-SAPCA applies PCA on 3-D groups of adaptive-shape image patches. He *et al.* [25] presented an adaptive hybrid method called ASVD, which uses SVD to learn the local basis for representing image patches. Another SVD-based denoising method is called spatially adaptive iterative singular-value thresholding (SAIST) [26]. This method uses SVD as a sparse representation of image patches and reduces noise in images by iteratively shrinking the singular values with *BayesShrink*. BM3D-SAPCA and SAIST are considered to be the current state of the art in image denoising.

In this paper, we propose a simple and efficient denoising method by combining patch grouping with SVD. The proposed method first groups image patches by a classification algorithm to achieve many groups of similar patches. Then each group of similar patches is estimated by the low-rank approximation (LRA) in SVD domain. The denoised image is finally obtained by aggregating all processed patches. The SVD is a very suitable tool for estimating each group because it provides the optimal energy compaction in the least square sense [27]. This implies that we can achieve a good estimation of the group by taking only a few largest singular values and corresponding singular vectors. While ASVD uses SVD to learn a set of local basis for representing image patches and SAIST uses SVD as a sparse representation of image patches, the proposed method exploits the optimal energy compaction property of SVD to lead an LRA of image patches. Experiments indicate that the proposed method achieves highly competitive performance in visual quality, and it also has a lower computational cost than most of existing state-of-the-art denoising algorithms.

The rest of this paper is organized as follows. In Section II, we briefly review image representation tools for the sake of completeness. We present the proposed algorithm in detail in Section III, which fuses the nonlocal self-similarity and the LRA using patch clustering and SVD. In Section IV, we report the experimental results of our method to validate its efficacy and compare it with the state-of-the-art methods. In Section V, we discuss the differences between our method and other state-of-the-art methods. Finally, we conclude this paper with some possible future work in Section VI.

II. LINEAR IMAGE REPRESENTATION

Let \mathbf{X} be a grayscale image. The basic principle of linear image representation is that the signal of interest can be decomposed into a weighted sum of a given

representation basis. Thus, \mathbf{X} can be represented as

$$\mathbf{X} = \sum_{i=1}^N a_i \phi_i \quad (1)$$

where $a_i (i = 1, \dots, N)$ are the representation coefficients of the image \mathbf{X} in terms of the basis functions $\phi_i (i = 1, \dots, N)$. ϕ_i can either be chosen as a prespecified basis, such as wavelet [28], curvelet [29], contourlet [30], shearlet [31], and other directional basis, or designed by adapting its content to fit a given set of images. In general, an adaptive basis has better performance than the prespecified one.

Aharon *et al.* [20] proposed a learning method to achieve a set of adaptive basis (also called dictionary). This method extracts all the $\sqrt{m} \times \sqrt{m}$ patches from the image \mathbf{X} to form a data matrix $\mathbf{S} = (\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n) \in \mathcal{R}^{m \times n}$, where m is the number of pixels in each patch, $\mathbf{s}_i (i = 1, \dots, n)$ are the image patches ordered as columns of \mathbf{S} and n is the number of patches. Then the dictionary is learned by solving the following optimization problem:

$$\min_{\Phi, \mathbf{A}} \sum_{i=1}^n \|\mathbf{s}_i - \Phi \mathbf{a}_i\|_2^2 \quad \text{s.t.} \quad \|\mathbf{a}_i\|_0 \leq \beta \quad (2)$$

where $\Phi \in \mathcal{R}^{m \times p}$ is the dictionary of p column atoms, $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) \in \mathcal{R}^{p \times n}$ is a matrix of coefficients, β indicates the desired sparsity level of the solution, and the notation $\|\mathbf{a}_i\|_0$ stands for the count of the nonzero entries in \mathbf{a}_i . Based on the learned dictionary Φ , \mathbf{S} can be represented as

$$\mathbf{S} = \Phi \mathbf{A}. \quad (3)$$

Another method for image representation with adaptive basis selection is PCA [32], which determines the basis from the covariance statistics of the data matrix \mathbf{S} . The principal components transform of \mathbf{S} is calculated as [33]

$$\mathbf{A} = \Phi^T (\mathbf{S} - E(\mathbf{S})) \quad (4)$$

with Φ defined by

$$\Omega_{\mathbf{S}} = \Phi \Lambda \Phi^T \quad (5)$$

where $E(\mathbf{S})$ is the matrix of mean vectors, $\Omega_{\mathbf{S}}$ is the covariance matrix of \mathbf{S} , Φ is the eigenvector matrix, and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_m)$ is the diagonal eigenvalue matrix with

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m. \quad (6)$$

It can easily be derived that the covariance matrix $\Omega_{\mathbf{A}}$ of the matrix \mathbf{A} equals

$$\Omega_{\mathbf{A}} = \Phi^T \Omega_{\mathbf{S}} \Phi = \Lambda \quad (7)$$

which implies that the entries of \mathbf{A} are uncorrelated. This property of PCA can be used to distinguish between the signal and noise. It is because the energy of noise is generally spread over the whole transform coefficients, while the energy of a signal is concentrated on a small amount of coefficients.

One major shortcoming of the adaptive dictionary and PCA is that they impose a very high computational burden.

An alternative method for adaptive basis selection is SVD. The SVD of the data matrix \mathbf{S} is a decomposition of the form [34]

$$\mathbf{S} = \mathbf{U} \Sigma \mathbf{V}^T = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^T \quad (8)$$

where $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_n) \in \mathcal{R}^{m \times n}$ and $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_n) \in \mathcal{R}^{n \times n}$ are the matrices with orthonormal columns, $\mathbf{U}^T \mathbf{U} = \mathbf{V}^T \mathbf{V} = \mathbf{I}$, and where the diagonal matrix $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_n)$ has nonnegative diagonal elements appearing in nonincreasing order such that

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0. \quad (9)$$

The diagonal entries σ_i of Σ are called the singular values of \mathbf{S} , while the vectors \mathbf{u}_i and \mathbf{v}_i are the left and right singular vectors of \mathbf{S} , respectively. The product $\mathbf{u}_i \mathbf{v}_i^T$ in (8) can be considered as an adaptive basis, and σ_i as the representation coefficient.

In fact, SVD and PCA are intimately related. PCA can be performed by calculating the SVD of the data matrix $(1/\sqrt{n})\mathbf{S}^T$ (refer to [35] for more details). In addition, if a matrix is low rank, we can easily estimate it from its noisy version by the LRA in SVD domain. Thus, we propose a new denoising method using SVD instead of PCA in the following section, which has a low computational complexity.

III. PROPOSED METHOD

Based on the analysis of SVD in Section II, we propose an efficient method to estimate the noise-free image by combining patch grouping with the LRA of SVD, which leads to an improvement of denoising performance. The main motivation to use SVD in our method is that it provides the optimal energy compaction in the least square sense, which implies that the signal and noise can be better distinguished in SVD domain. Fig. 1 shows a block diagram of the proposed approach. Concretely, the patch grouping step identifies similar image patches by the Euclidean-distance-based similarity metric. Once the similar patches are identified, they can be estimated by the LRA in the SVD-based denoising step. In the aggregation step, all processed patches are aggregated to form the denoised image. The back projection step uses the residual image to further improve the denoised result.

For ease of presentation, let \mathbf{Y} denote a noisy image defined by

$$\mathbf{Y} = \mathbf{X} + \mathbf{E} \quad (10)$$

where \mathbf{X} is the noise-free image, and \mathbf{E} represents the additive white Gaussian noise (AWGN) with the standard deviation τ that, in practice, can usually be estimated by various methods such as median absolute deviation (MAD) [36], SVD-based estimation algorithm [37], and block-based ones [38], [39]. In this paper, we use a vectorized version of the model (10)

$$\mathbf{y} = \mathbf{x} + \mathbf{e}. \quad (11)$$

Given a noisy observation \mathbf{y} , our aim is to estimate \mathbf{x} as accurately as possible.

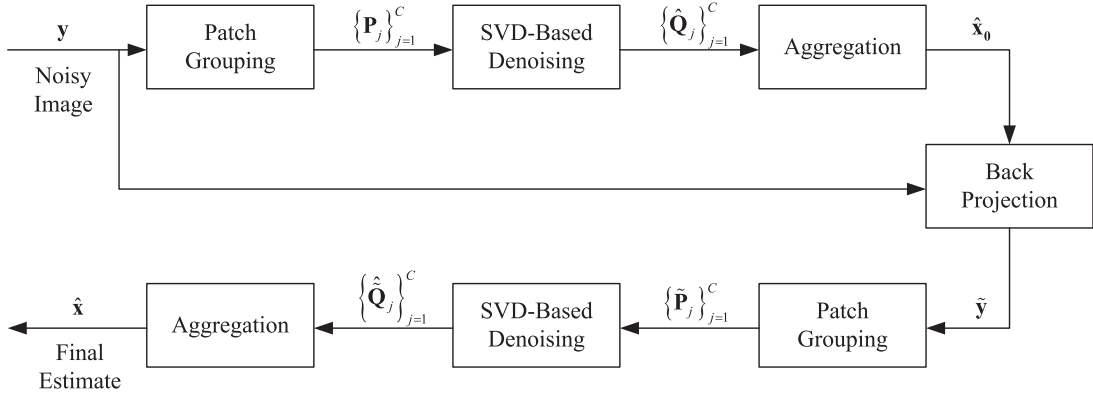


Fig. 1. Block diagram of the proposed denoising algorithm.

As similarly done in BM3D and LPG-PCA, the proposed method also has two stages: 1) the first stage produces an initial estimation of the image \mathbf{x} and 2) the second stage further improves the result of the first stage. Different from them, our method adopts the LRA to estimate image patches and uses the back projection to avoid loss of detailed information of the image. Each stage contains three steps: 1) patch grouping; 2) SVD-based denoising; and 3) aggregation. In the first stage, the noisy image \mathbf{y} is first divided into M overlapping patches denoted by $\{\mathbf{y}_i\}_{i=1}^M$, where \mathbf{y}_i is a vectorized format of the i th image patch. For each patch \mathbf{y}_j , its similar patch group is formed by searching similar patches from $\{\mathbf{y}_i\}_{i=1}^M$. Next, each similar patch group is denoised by the low-rank approximation in SVD domain. Third, the denoised image $\hat{\mathbf{x}}_0$ is achieved by aggregating all denoised patches. In the second stage, the final denoised image is obtained by applying the processing steps described above on the image $\tilde{\mathbf{y}}$ produced by the back projection process. In the rest of this section, the procedures of our proposed method will be described in detail.

A. Patch Grouping

Grouping similar patches, as a classification problem, is an important and fundamental issue in image and video processing with a wide range of applications. While there exist many classification algorithms available in [40], e.g., block matching, K -means clustering, nearest neighbor clustering, and others, we exploit the block-matching method for image patch grouping due to its simplicity.

For each given reference patch \mathbf{y}_j with size $\sqrt{m} \times \sqrt{m}$, the block-matching method finds its similar patches from $\{\mathbf{y}_i\}_{i=1}^M$ by a similarity metric. In [22], the Euclidean distance from the transform coefficients is used to identify the similar square patches. A shape-adaptive version of this similarity metric is presented in [23], whereas it leads to a high computational cost. The simplest measure of similarity between two patches is the Euclidean distance directly in the spatial domain. Thus, we employ the spatial Euclidean distance as our similarity metric, which is defined by

$$S(\mathbf{y}_j, \mathbf{y}_c) = \|\mathbf{y}_j - \mathbf{y}_c\|_2^2 \quad (12)$$

where $\|\cdot\|_2$ denotes the Euclidean distance and \mathbf{y}_c is a candidate patch. The smaller $S(\mathbf{y}_j, \mathbf{y}_c)$ is, the more similar

\mathbf{y}_j and \mathbf{y}_c are. The reference patch \mathbf{y}_j and its L -most similar patches denoted by $\{\mathbf{y}_{c,i}\}_{i=1}^L$ are chosen to construct a group matrix using each similar patch as a column of the group matrix, and its corresponding group matrix \mathbf{P}_j is formed by

$$\mathbf{P}_j = [\mathbf{y}_j, \mathbf{y}_{c,1}, \dots, \mathbf{y}_{c,L}]. \quad (13)$$

Due to \mathbf{P}_j being made up of the noisy patches, it can be represented as

$$\mathbf{P}_j = \mathbf{Q}_j + \mathbf{N}_j \quad (14)$$

where \mathbf{Q}_j and \mathbf{N}_j denote the noise-free group matrix and the noise matrix, respectively.

In general, the number L of similar patches in the group matrix cannot be too small. Too small L leads to too few patches within each group matrix, which makes the SVD-based denoising less robust. On the contrary, too large one leads to dissimilar patches being grouped together, which results in an incorrect estimation of \mathbf{P}_j . Similarly, the patch size $\sqrt{m} \times \sqrt{m}$ also has an impact on the performance of our method. We will discuss the influence of L and the patch size in Section IV-C.

B. SVD-Based Denoising

For simplicity of description, we will use \mathbf{Q} and \mathbf{P} instead of \mathbf{Q}_j and \mathbf{P}_j by a slight abuse of notation. Now our task is to estimate the noise-free group matrix \mathbf{Q} from its noisy version \mathbf{P} as accurately as possible. Ideally, the estimate $\hat{\mathbf{Q}}$ should satisfy

$$\|\mathbf{P} - \hat{\mathbf{Q}}\|_F^2 = \tau^2 \quad (15)$$

where $\|\cdot\|_F$ is the Frobenius norm¹ and τ is the standard deviation of noise.

The similarity between patches within the noise-free image \mathbf{x} leads to a high correlation between them, which means that \mathbf{Q} is a low-rank matrix. Fig. 2 shows the low-rank property of \mathbf{Q} by displaying the singular values of group matrices of *Lena* image with different noise levels, where each point is the average i th singular value over all group matrices. The estimate of \mathbf{Q} can be obtained by the

¹For an $m \times n$ matrix A with elements a_{ij} ($i = 1, \dots, m, j = 1, \dots, n$), the Frobenius norm is defined as the square root of the sum of the absolute squares of its elements, i.e., $\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$.

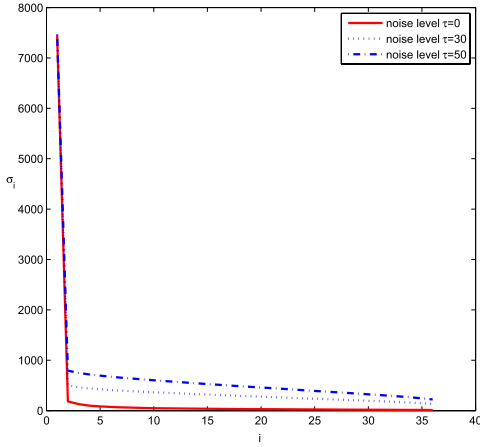


Fig. 2. Singular values of group matrices of *Lena* image with different noise levels.

LRA in the least square sense. Therefore, we can estimate \mathbf{Q} from \mathbf{P} by solving the following optimization problem:

$$\hat{\mathbf{Q}} = \arg \min_{\mathbf{Z}} \|\mathbf{P} - \mathbf{Z}\|_F^2 \quad \text{s.t.} \quad \text{rank}(\mathbf{Z}) = k \quad (16)$$

where $\text{rank}(\cdot)$ denotes the rank of matrix \mathbf{Z} .

In SVD domain, \mathbf{P} can be represented as

$$\mathbf{P} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T. \quad (17)$$

Let

$$\mathbf{P}_k = \mathbf{U}\mathbf{\Sigma}_k\mathbf{V}^T \quad (18)$$

where $\mathbf{\Sigma}_k$ is obtained from the matrix $\mathbf{\Sigma}$ by setting the diagonal elements to zeros but the first k singular values

$$\mathbf{\Sigma}_k = \text{diag}(\sigma_1, \dots, \sigma_k, 0, \dots, 0). \quad (19)$$

\mathbf{P}_k is the solution of (16), which is a classical result given by the Eckart–Young–Mirsky theorem [41], [42].

Theorem 1 (Eckart–Young–Mirsky): For any real matrix \mathbf{P} , if the matrix \mathbf{Q} is of rank k , then

$$\|\mathbf{P} - \mathbf{Q}\|_F^2 \geq \sum_{i=k+1}^n \sigma_i^2 \quad (20)$$

where $\sigma_i (i = 1, \dots, n)$ are the singular values of \mathbf{P} , and equality is attained when $\mathbf{Q} = \mathbf{P}_k$ is defined by (18).

This theorem shows that \mathbf{P}_k is the optimal solution for (16) in the Frobenius norms sense. Thus, we have

$$\hat{\mathbf{Q}} = \mathbf{P}_k. \quad (21)$$

The key issue for this method is to determine the value of k . By comparing (15) with (20), we can find that \mathbf{P}_k is the ideal estimate of \mathbf{P} when $\sum_{i=k+1}^n \sigma_i^2$ is equal to τ^2 . Therefore, k can be determined by the following criterion:

$$\sum_{i=k}^n \sigma_i^2 > \tau^2 \geq \sum_{i=k+1}^n \sigma_i^2. \quad (22)$$

C. Aggregation

Till now, we have estimated each group matrix by applying the LRA defined by (21). Then the denoised patches can be obtained by rearranging column vectors of each denoised group matrix. As a result of taking the L nearest neighbors of each patch to construct a group matrix, a single patch might belong to several groups, and multiple estimates of this patch can be obtained. Thus, we aggregate different estimates of this patch to obtain its denoised version by the following averaging process:

$$\hat{\mathbf{x}}_i = \frac{1}{n} \sum_{j=1}^n \hat{\mathbf{x}}_{i,j} \quad (23)$$

where $\hat{\mathbf{x}}_i$ is the denoised version of a patch \mathbf{y}_i , and $\hat{\mathbf{x}}_{i,j} (j = 1, \dots, n)$ denote n different estimates of \mathbf{y}_i .

The next step is to synthesize the denoised image from the denoised patches. Since the patches are sampled with overlapping regions for avoiding block artifacts at the boundaries of patches, multiple estimates are obtained for each pixel. Thus, these estimates of each pixel in the image need to be aggregated to reconstruct the final denoised image. The common method of combining such multiple estimates is to perform a weighted averaging of them. Meanwhile, the weighted averaging procedure can suppress noise further. The simplest form of aggregation is the uniformly weighted averaging that assigns the same weight to all estimates. However, the uniform weights will lead to an oversmoothed result. In general, the adaptive weights derived from various biased and unbiased estimators, such as variance-based weights, SURE-based weights, and exponential weights [10], can lead to better results. Different from these adaptive weights, in this paper, we exploit the weights depending on the rank k of each group matrix due to its simplicity. For the j th group matrix $\hat{\mathbf{Q}}_j$, our weight is defined by

$$w_j = \begin{cases} 1 - \frac{k}{L+1}, & k < L+1 \\ \frac{1}{L+1}, & k = L+1. \end{cases} \quad (24)$$

If $k < L+1$, it means that patches in the group matrix are linearly correlated. The higher the degree of correlation of patches is, the smaller the rank k of the group matrix is. The estimate of patches yielded from the LRA is better. Thus, this estimate needs to be assigned a high weight. If $k = L+1$, there exists no correlation among patches. The simplest uniform weight is used. Based on the weights defined in (24), the denoised estimate for the i th pixel of the image can be expressed as

$$\hat{x}_i = \frac{1}{W} \sum_{j \in \Gamma(x_i)} w_j \hat{x}_{i,j} \quad (25)$$

where W is a normalizing factor defined by

$$W = \sum_j w_j. \quad (26)$$

$\Gamma(x_i)$ denotes the index set of all similar group matrices containing the pixel x_i , which is described as

$$\Gamma(x_i) = \{j | x_i \in \mathbf{Q}_j, j = 1, \dots, C\} \quad (27)$$

and $\hat{x}_{i,j}$ denotes the denoised estimate of the i th pixel in the j th similar group matrix $\hat{\mathbf{Q}}_j$. Once all pixels are estimated by (25), the final denoised image can be obtained by reshaping the estimates of all pixels.

D. Back Projection

Although most of noise can be removed using the denoising procedures described before, there still exists a small amount of noise residual in the denoised image. The noise residual stems from the fact that noise in the original noisy image affects the accuracy of the patch grouping, which leads to an inaccurate group. The grouping errors in turn affect the SVD-based denoising. In addition, there exists another reason for noise residual. Ideally, based on the discussion in Section III-B, the optimal estimate $\hat{\mathbf{Q}}$ satisfies

$$\begin{aligned} \|\mathbf{P} - \hat{\mathbf{Q}}\|_F^2 = \|\mathbf{P} - \mathbf{Q}\|_F^2 &\implies \|\mathbf{P} - \hat{\mathbf{Q}}\|_F^2 = \|\mathbf{N}\|_F^2 \\ &\implies \sum_{i=k+1}^n \sigma_i^2 = \tau^2. \end{aligned} \quad (28)$$

Unfortunately, the left side of (28) is not usually equal to the right side. In most cases, it is that $\tau^2 > \sum_{i=k+1}^n \sigma_i^2$. Therefore, we need to further improve the denoising performance of our method.

The commonly used way to further improve the performance of a denoising method, as used by the clustering-based denoising method using locally learned dictionaries (named K-LLD) [43] and SAIST, is to develop an iterative version for the basic denoising method. While the iterative strategy for image denoising has been widely used in the literature, it has a very high computational cost, which limits the scope of applications. An alternative way exploited by BM3D and LPG-PCA is the two-stage approach, in which the basic estimate of the noisy image yielded by the denoising method is used as a reference image to perform improved grouping and parameter estimation.

In this paper, unlike the iteration-based or the reference-based strategies, we make use of the two-stage strategy with a back projection step to further suppress the noise residual. Back projection is an efficient method that uses the residual image to improve the denoised result [44], [45]. In fact, the use of the residuals in improving estimates can date at least back to [46], in which this idea is termed *twicing*. This concept is also known by several names, such as Bregman iterations, l_2 -boosting, and biased diffusion. Milanfar [47] provides a good overview of these methods. The basic idea of back projection is to generate a new noisy image by adding filtered noise back to the denoised image

$$\tilde{\mathbf{y}} = \hat{\mathbf{x}}_0 + \delta(\mathbf{y} - \hat{\mathbf{x}}_0) \quad (29)$$

where $\delta \in (0, 1)$ is a constant projection parameter and $\hat{\mathbf{x}}_0$ is the denoised result produced by the first stage. Note that when $\delta \rightarrow 0$, $\tilde{\mathbf{y}} \rightarrow \hat{\mathbf{x}}_0$. On the contrary, if $\delta \rightarrow 1$, $\tilde{\mathbf{y}} \rightarrow \mathbf{y}$. For simplicity, in our experiments, we set $\delta = 0.5$, which is a tradeoff between 1 and 0.

Now we can achieve an improved result of $\hat{\mathbf{x}}_0$ by denoising $\tilde{\mathbf{y}}$ with the proposed three processing steps in Sections III-A–III-C, i.e., *patch grouping*, *SVD-based*

Algorithm 1 Proposed Denoising Algorithm

Input: Noisy image \mathbf{y}

Output: Denoised image $\hat{\mathbf{x}}$

- 1: $\tau \leftarrow$ Estimate noise standard deviation by computing the median absolute deviation (MAD) of the finest wavelet coefficients as described in [36];
 - 2: $\{\mathbf{P}_j\}_{j=1}^C \leftarrow$ Group image patches with the similarity metric defined by Eq. (12) and form group matrices;
 - 3: **for** each $j \in [1, C]$ **do**
 - 4: $\hat{\mathbf{Q}}_j \leftarrow$ Calculate the LRA of \mathbf{P}_j via Eq. (21);
 - 5: $w_j \leftarrow$ Compute the weight for $\hat{\mathbf{Q}}_j$ via Eq. (24);
 - 6: **end for**
 - 7: $\hat{\mathbf{x}}_i \leftarrow$ For each patch existing multiple different estimates, aggregate its estimates via Eq. (23).
 - 8: $\hat{\mathbf{x}}_0 \leftarrow$ Aggregate all $\hat{\mathbf{Q}}_j$ via Eq. (25);
 - 9: $\tilde{\tau} \leftarrow$ Update the noise standard deviation τ via Eq. (30);
 - 10: $\tilde{\mathbf{y}} \leftarrow$ Generate a new noisy image via the back projection described by Eq. (29);
 - 11: $\hat{\mathbf{x}} \leftarrow$ Obtain the final denoised image by performing *Step 2* to *Step 8* for $\tilde{\mathbf{y}}$.
-

denoising, and *aggregation*, respectively. It is necessary to point out that the noise variance of $\tilde{\mathbf{y}}$, denoted by $\tilde{\tau}^2$, needs to be updated in the SVD-based denoising step. We employ the estimator presented in [26] to determine $\tilde{\tau}^2$, which is written as

$$\tilde{\tau} = \gamma \sqrt{\tau^2 - \|\mathbf{y} - \hat{\mathbf{x}}_0\|_F^2} \quad (30)$$

where γ is a scaling factor.

To summarize, the complete procedure of our proposed method is algorithmically described in Algorithm 1.

IV. EXPERIMENTAL RESULTS

To demonstrate the efficacy of the proposed denoising algorithm, in this section, we give our experimental results concerning simulations that have been conducted on ten natural grayscale images with size 512×512 . These images have been commonly used to validate many state-of-the-art denoising methods. The noisy images are generated by adding zero mean white Gaussian noise with different levels to the test images. The noise level τ is from 10 to 50, and the intensity value for each pixel of the images ranges from 0 to 255.

A. Evaluation Criteria

Two objective criteria, namely, peak signal-to-noise ratio (PSNR) and feature-similarity (FSIM) index [48], are adopted to provide quantitative quality evaluations of the denoising results. PSNR is the mostly widely used quality measure in the literature, even though it is often inconsistent with human eye perception. FSIM measures the similarity between two images by combining the phase congruency feature and the gradient magnitude feature, which is based on the fact that human visual system understands an image mainly according to its low-level features. The aforementioned criteria can comprehensively reflect the performance of the denoising methods.



Fig. 3. Visual comparisons of denoising results on *Lena* image corrupted by AWGN with standard deviation 30. (a) Original image. (b) Noisy image. (c) PLOW [11]. (d) K-SVD [21]. (e) LPG-PCA [24]. (f) SAIST [26]. (g) BM3D-SAPCA [23]. (h) Proposed method.

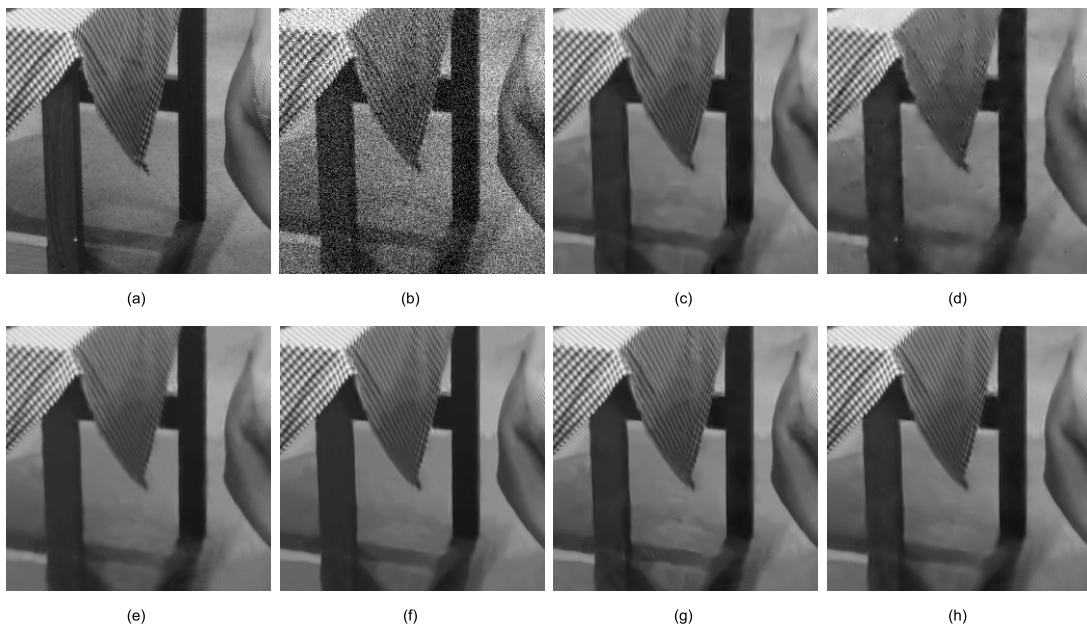


Fig. 4. Visual comparisons of denoising results on *Barbara* image corrupted by AWGN with standard deviation 30. (a) Original image. (b) Noisy image. (c) PLOW [11]. (d) K-SVD [21]. (e) LPG-PCA [24]. (f) SAIST [26]. (g) BM3D-SAPCA [23]. (h) Proposed method.

of noise present in the image, we set the patch size 9×9 if $\tau < 20$, 10×10 if $20 \leq \tau < 40$, and 11×11 if $\tau \geq 40$. The influence of different parameters will be evaluated in Section IV-C.

In Table I, we quantify the performances of six competing algorithms for the test images with different noise levels in terms of PSNR and FSIM. From Table I we can observe that BM3D-SAPCA, which is considered to be the state of the art in image denoising, achieves the highest PSNR values on average, and slightly outperforms our method. However, our method performs better on images with high repeating patterns such as *Elaine*, *Zelda*, and *Barbara*. It is because our method sufficiently exploits the nonlocal redundancies in these images

by the patch grouping procedure described in Section III-A. For example, for *Elaine* image, on average, our method is superior to BM3D-SAPCA by 0.14 dB, to SAIST by 0.28 dB, to LPG-PCA by 0.58 dB, to PLOW by 0.47 dB, and to K-SVD by 0.43 dB, respectively. The FSIM results of different algorithms are also tabulated in Table I. It can be observed that our method has higher FSIM measures than other methods except BM3D-SAPCA. In a word, the quantitative results by our method are competitive with BM3D-SAPCA and SAIST, and clearly superior to LPG-PCA, PLOW, and K-SVD.

In terms of visual quality, our method also is comparable and even superior to the state-of-the-art denoising methods. Fig. 3 shows the denoising results of *Lena* image with a

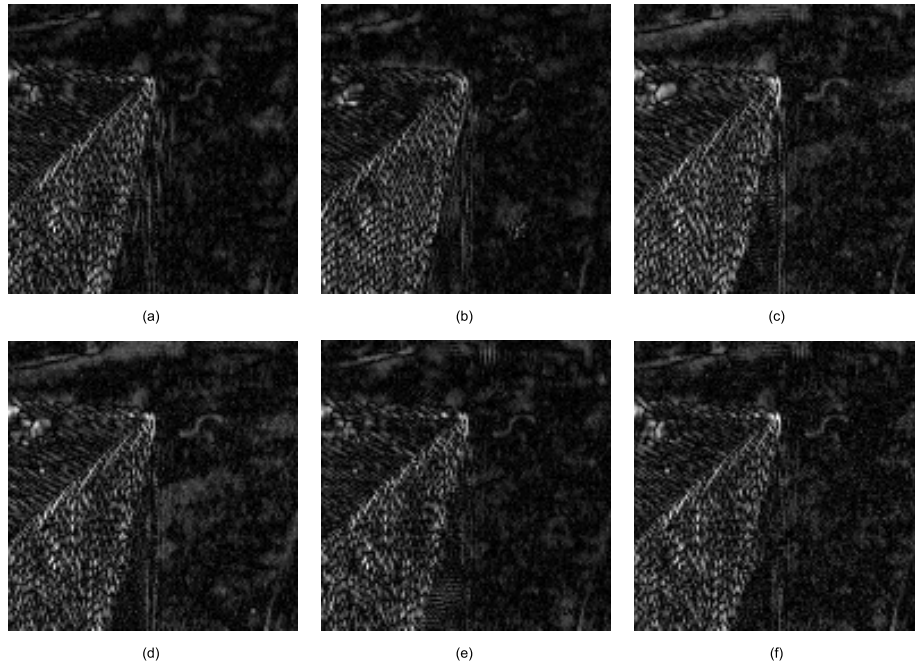


Fig. 5. Zoomed absolute difference between the original *Barbara* image and the denoised version. (a) PLOW [11] (MAE = 4.73). (b) K-SVD [21] (MAE = 5.31). (c) LPG-PCA [24] (MAE = 5.23). (d) SAIST [26] (MAE = 4.75). (e) BM3D-SAPCA [23] (MAE = 4.44). (f) Proposed method (MAE = 4.35).

noise level of $\tau = 30$. As can be seen from it, the results by the proposed method are visually close to SAIST and LPG-PCA, and better than BM3D-SAPCA, PLOW, and K-SVD, especially in the edge and texture regions. The visual comparisons are further illustrated in Fig. 4, which shows the zoomed-in denoising results of the noisy image *Barbara* by different methods. It can be observed that the denoised images by SAIST, LPG-PCA, and the proposed method are very similar in real visual perception, in which some edges and textures are better preserved, and fewer artifacts are introduced. We note that although BM3D-SAPCA has higher PSNR and FSIM measures than our method, denoised results by BM3D-SAPCA contain more noticeable artifacts around edges and in smooth regions than our results. The main reason is that BM3D-SAPCA exploits the orthogonal transform to represent similar image patches and reduces noise by thresholding representation coefficients, in which significantly visible artifacts are produced. (This phenomenon also was discussed in [49].)

As shown in Fig. 5, we also calculated the absolute difference images between the original *Barbara* image and the denoised versions of six denoising algorithms. The mean absolute error (MAE) values of PLOW, K-SVD, LPG-PCA, SAIST, BM3D-SAPCA, and our method are 4.73, 5.31, 5.23, 4.75, 4.44, and 4.35, respectively. The MAE value produced by our method is lower than those by other denoising algorithms. To further demonstrate our performance, we apply the proposed method to some real noisy images.³ Fig. 6 displays the denoised images yielded by PLOW and

our method. Our method can reduce the noise effectively, while preserving the finer features. In short, our denoising results are both quantitatively and visually comparable with the state of the art.

C. Influence of Parameters

In our method, there are four tuning parameters: 1) patch size; 2) number of similar patches in each group matrix L ; 3) projection factor δ ; and 4) scaling factor γ . The patch size plays an important role in the proposed denoising algorithm. On the one hand, a too large patch size can capture the varying local geometry and also lead to a high computational cost. On the other hand, a too small patch size can reduce the denoising performance of the proposed method. To study the influence of the patch size, we set $L = 85$, $\delta = 0.5$, $\gamma = 0.65$, and perform our method on *Barbara* image with different patch sizes and noise levels. The PSNR results are listed in Table II. It can be observed that choosing the patch size from 9×9 to 11×11 in our experiments is a reasonable tradeoff between accuracy and speed.

The projection parameter δ controls the amount of residual image added to the output of the first stage. To analyze the effect of this parameter, we run our algorithm with different values, $\delta = 0.1, 0.2, \dots, 0.9$. Fig. 7(a) shows the denoising performance of the proposed algorithm applied to the noisy image *Barbara* ($\tau = 30$) as a function of the parameter δ . As can be seen, the highest PSNR is reached when $\delta = 0.5$. Similar curves for γ and L are shown in Fig. 7(b) and (c), respectively. The best denoising result is obtained with $\gamma = 0.65$. In addition, we find that our algorithm is insensitive to L in the range [70, 100]. Thus, we choose $L = 85$ as a tradeoff between accuracy and speed.

³Available at <http://users.soe.ucsc.edu/~7epriyam/PLOW/>. In our experiments, we use their grayscale versions.



Fig. 6. Visual comparisons of denoising results on real noisy images with unknown noise characteristics.

TABLE II
PSNR (dB) RESULTS OF THE PROPOSED DENOISING METHOD ON *Barbara* IMAGE WITH DIFFERENT PATCH SIZES AND NOISE LEVELS

	Patch size						
	7 × 7	8 × 8	9 × 9	10 × 10	11 × 11	12 × 12	13 × 13
$\tau = 10$	35.40	35.42	35.43	35.42	35.38	35.35	35.30
$\tau = 20$	32.07	32.18	32.22	32.23	32.22	32.19	32.16
$\tau = 30$	29.86	30.07	30.16	30.21	30.21	30.24	30.24
$\tau = 40$	28.18	28.44	28.58	28.67	28.74	28.78	28.81
$\tau = 50$	26.82	27.14	27.31	27.42	27.50	27.56	27.60

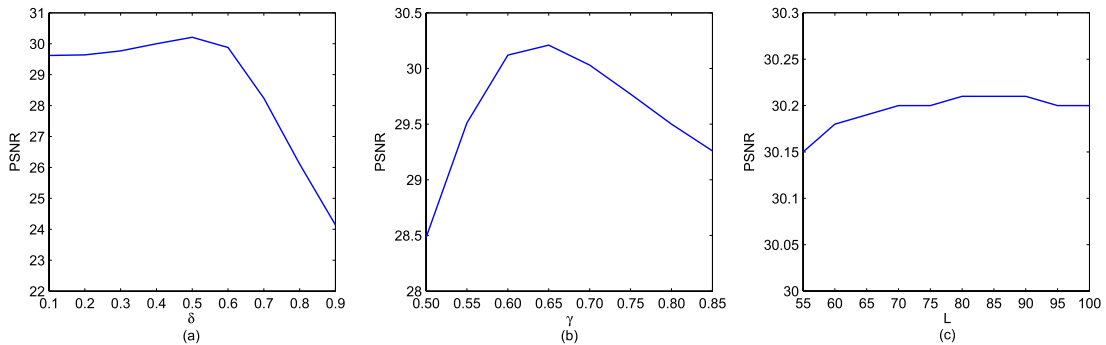


Fig. 7. PSNR results on *Barbara* image ($\tau = 30$) as a function of (a) varying δ with a patch size of 10×10 , $L = 85$, and $\gamma = 0.65$, (b) varying γ with a patch size of 10×10 , $\delta = 0.5$, and $L = 85$, and (c) varying L with a patch size of 10×10 , $\delta = 0.5$, and $\gamma = 0.65$.

D. Analysis of Iterative Denoising

Since the noise reduces the accuracy of the patch grouping, the grouping errors make the residual image of the first stage contain a lot of visual edge or texture details. In the second denoising stage, the new noisy image produced by back projection contains more structural details than the denoised image of the first stage and has a lower noise level than the original noisy image, which improves the accuracy of the patch grouping and reduces the error of the LRA in SVD domain. Therefore, a better result can be reached at the end of the

second denoising round. The previously mentioned experimental results demonstrate that the two-stage denoising scheme based on back projection is very effective in suppressing noise.

Intuitively, the denoised results might be further improved if the proposed denoising scheme is iterated more than twice. Fig. 8 shows the evolution of the PSNR during the iterations. Here, we note that further iterations cannot significantly improve the denoising performance. In theory, the truncated SVD shrinkage used in our method guarantees an optimal approximation of the noisy group matrix in the least square

TABLE III
COMPARISON OF THE COMPUTATIONAL TIME AND THE IMPLEMENTATION LANGUAGE OF DIFFERENT DENOISING METHODS

	PLOW	K-SVD	LPG-PCA	SAIST	BM3D-SAPCA	Proposed
Language	MATLAB	MATLAB	MATLAB	MATLAB	MATLAB	MATLAB
Average running time (s)	1875	393	997	197	925	77

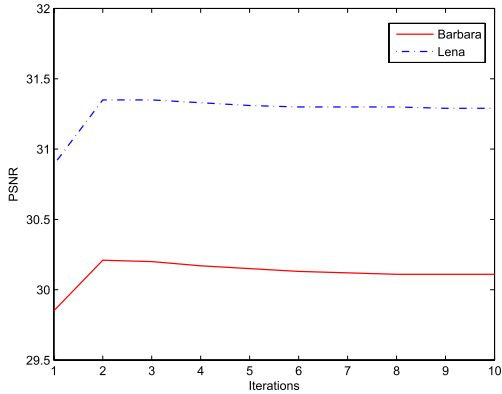


Fig. 8. Evolution of the PSNR with iterations for different images ($\tau = 30$).

sense, which makes most of the noise be suppressed in the first two iterations.

E. Computational Cost

To evaluate the computational cost of six denoising methods, we compare the running time on the ten test images with different noise levels. We have run all the source codes by default throughout all experiments performed on a platform of an Intel Core i7-870 CPU 2.93 GHz with 4-GB memory. Denoising a 512×512 grayscale image, SAIST, K-SVD, BM3D-SAPCA, LPG-PCA, and PLOW take, on average, roughly 197, 393, 925, 997, and 1875 s, respectively. The computational cost of the proposed method is quite low in comparison with the denoising algorithms above. For the test images, our MATLAB implementation requires only 77 s on average. There are two main computational components of our algorithm, one is grouping similar patches to form group matrices and the other is the calculation of SVD for each group matrix. In general, the patch grouping step takes approximately 23% of the execution time, whereas 75% of the time is spent in estimating group matrices by the LRA in SVD domain. The execution time of the various algorithms are presented in Table III. It can be seen that the proposed method provides the fastest running speed among the six denoising algorithms.

V. DISCUSSION

The nonlocal self-similarity of natural images plays an important role in image denoising. The most well-known denoising method based on the nonlocal self-similarity is BM3D, which often produces state-of-the-art denoising results. BM3D-SAPCA is an improved version of BM3D by exploiting PCA and shape-adaptive image patches, which achieves remarkable performance. Our method utilizes the nonlocal self-similarity to construct low-rank group matrices that can be easily estimated by the LRA.

The main differences between these methods are threefold.

- 1) The basis functions of image representations are different. BM3D uses the fixed 3-D basis functions (joint wavelets and cosine bases) that are less adapted to the edges and textures. To improve the denoising performance, BM3D-SAPCA applies PCA to the shape-adaptive patch group, whereas it leads to a high computational cost. Our method uses an adaptive basis derived by SVD, which outperforms BM3D and BM3D-SAPCA by better preserving the local geometric structure.
- 2) BM3D uses the Euclidean distance from the transform coefficients to identify the similar square patches, which can improve the robustness of the block matching. In BM3D-SAPCA, the usual square patches are replaced by the shape-adaptive patches for block matching, whereas it leads to a complex aggregation process with a high computational cost. Different from BM3D and BM3D-SAPCA, the proposed method calculates the similarity metric based on the Euclidean distance directly in the spatial domain due to its simplicity.
- 3) In BM3D and BM3D-SAPCA, the second denoising stage is directly applied on the original noisy image that is grouped into 3-D data arrays based on the patch similarities from the denoised image in the first stage. However, the denoising for the first stage might contain grouping errors due to the effect of noise, which yields an incorrect basic estimate of the noisy image. Therefore, using this basic estimate as the pilot signal is not very ideal because it would decay the accuracy of the second denoising stage. Unlike BM3D and BM3D-SAPCA, the second denoising stage in our method is applied on a new noisy image obtained by adding a part of residual image to the basic estimate, i.e., back projection. The new noisy image contains more structural details than the output of the first stage, which improves the accuracy of the patch grouping and the LRA in SVD domain.

Besides, the existing methods based on the adaptive representation, such as K-SVD and ASVD, often need to learn a set of adaptive basis using the given training images. Unfortunately, the learning process is computationally expensive. The proposed method is intrinsically simpler than K-SVD and ASVD. And it does not need to be trained for each image separately or for a given training data set, which avoids a high computational cost for the training process.

VI. CONCLUSION

In this paper, we have presented a simple and efficient method for image denoising, which takes advantage of the nonlocal redundancy and the LRA to

attenuate noise. The nonlocal redundancy is implicitly used by the block-matching technique to construct low-rank group matrices. After factorizing by SVD, each group matrix is efficiently approximated by preserving only a few largest singular values and corresponding singular vectors. This is due to the optimal energy compaction property of SVD. In fact, the small singular values have little effect on the approximation of the group matrix when it has a low-rank structure. The experimental results demonstrate the advantages of the proposed method in comparison with current state-of-the-art denoising methods.

The computational complexity of the proposed algorithm is lower than that of most of the existing state-of-the-art denoising algorithms, but higher than BM3D. The fixed transform used by BM3D is less complex than SVD, whereas it is less adapted to edges and textures. The main computational cost of our algorithm is the calculation of SVD for each patch group matrix. As each group matrix could potentially be processed independently in parallel, our method is suitable for parallel processing. Therefore, in practice, we can use a parallel implementation to speed it up, which will make it feasible for real-time or near real-time image denoising. In addition, while developed for grayscale images, our method can be extended to shape-adaptive color image and video denoising by taking into account the shape-adaptive patches and the temporal redundancy across color components and frames. This further work will be studied in the future.

ACKNOWLEDGMENT

The authors would like to thank Chatterjee and Milanfar [11], Elad and Aharon [21], Dabov *et al.* [23], Zhang *et al.* [24], and Dong *et al.* [26] for sharing their source codes and Prof. J. Wang with Shandong University and the anonymous reviewers for their helpful comments and suggestions to improve the quality of this paper.

REFERENCES

- [1] S. M. M. Rahman, M. O. Ahmad, and M. N. S. Swamy, "Video denoising based on inter-frame statistical modeling of wavelet coefficients," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 2, pp. 187–198, Feb. 2007.
- [2] F. Luisier, T. Blu, and M. Unser, "SURE-LET for orthonormal wavelet-domain video denoising," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 6, pp. 913–919, Jun. 2010.
- [3] H. Talebi, X. Zhu, and P. Milanfar, "How to SAIF-ly boost denoising performance," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1470–1485, Apr. 2013.
- [4] G. Varghese and Z. Wang, "Video denoising based on a spatiotemporal Gaussian scale mixture model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 7, pp. 1032–1040, Jul. 2010.
- [5] S. M. Smith and J. M. Brady, "SUSAN—A new approach to low level image processing," *Int. J. Comput. Vis.*, vol. 23, no. 1, pp. 45–78, 1997.
- [6] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. IEEE Int. Conf. Comput. Vis.*, Mumbai, India, Jan. 1998, pp. 839–846.
- [7] L. Shao, R. Yan, X. Li, and Y. Liu, "From heuristic optimization to dictionary learning: A review and comprehensive comparison of image denoising algorithms," *IEEE Trans. Cybern.*, vol. 44, no. 7, pp. 1001–1013, Jul. 2014.
- [8] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.
- [9] C. Kervrann and J. Boulanger, "Local adaptivity to variable smoothness for exemplar-based image regularization and representation," *Int. J. Comput. Vis.*, vol. 79, no. 1, pp. 45–69, Aug. 2008.
- [10] C. A. Deledalle, V. Duval, and J. Salmon, "Non-local methods with shape-adaptive patches (NLM-SAP)," *J. Math. Imag. Vis.*, vol. 43, no. 2, pp. 103–120, Jun. 2012.
- [11] P. Chatterjee and P. Milanfar, "Patch-based near-optimal image denoising," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1635–1649, Apr. 2012.
- [12] P. Chatterjee and P. Milanfar, "Is denoising dead?" *IEEE Trans. Image Process.*, vol. 19, no. 4, pp. 895–911, Apr. 2010.
- [13] X. Zhang, X. Feng, and W. Wang, "Two-direction nonlocal model for image denoising," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 408–412, Jan. 2013.
- [14] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1532–1546, Sep. 2000.
- [15] A. Pižurica and W. Philips, "Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 654–665, Mar. 2006.
- [16] 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.
- [17] Q. Guo and S. Yu, "Image denoising using a multivariate shrinkage function in the curvelet domain," *IEICE Electron. Exp.*, vol. 7, no. 3, pp. 126–131, Feb. 2010.
- [18] F. Luisier, T. Blu, and M. Unser, "A new SURE approach to image denoising: Interscale orthonormal wavelet thresholding," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 593–606, Mar. 2007.
- [19] Q. Guo and C. Zhang, "A noise reduction approach based on Stein's unbiased risk estimate," *ScienceAsia*, vol. 38, no. 2, pp. 207–211, Jun. 2012.
- [20] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for denoising overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
- [21] 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.
- [22] 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.
- [23] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "BM3D image denoising with shape-adaptive principal component analysis," in *Proc. Workshop Signal Process. Adapt. Sparse Struct. Represent. (SPARS)*, Saint-Malo, France, Apr. 2009, HAL Id: inria-00369582.
- [24] L. Zhang, W. Dong, D. Zhang, and G. Shi, "Two-stage image denoising by principal component analysis with local pixel grouping," *Pattern Recognit.*, vol. 43, no. 4, pp. 1531–1549, Apr. 2010.
- [25] Y. He, T. Gan, W. Chen, and H. Wang, "Adaptive denoising by singular value decomposition," *IEEE Signal Process. Lett.*, vol. 18, no. 4, pp. 215–218, Apr. 2011.
- [26] 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.
- [27] J.-F. Yang and C.-L. Lu, "Combined techniques of singular value decomposition and vector quantization for image coding," *IEEE Trans. Image Process.*, vol. 4, no. 8, pp. 1141–1146, Aug. 1995.
- [28] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd ed. New York, NY, USA: Academic, 2008.
- [29] E. Candès, L. Demanet, D. Donoho, and X. Ying, "Fast discrete curvelet transforms," *Multiscale Model. Simul.*, vol. 5, no. 3, pp. 861–899, Sep. 2006.
- [30] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.
- [31] G. Easley, D. Labate, and W.-Q. Lim, "Sparse directional image representations using the discrete shearlet transform," *Appl. Comput. Harmon. Anal.*, vol. 25, no. 1, pp. 25–46, Jul. 2008.
- [32] F. W. Mauldin, Jr., D. Lin, and J. A. Hossack, "The singular value filter: A general filter design strategy for PCA-based signal separation in medical ultrasound imaging," *IEEE Trans. Med. Imag.*, vol. 30, no. 11, pp. 1951–1964, Nov. 2011.
- [33] J. J. Gerbrands, "On the relationships between SVD, KLT and PCA," *Pattern Recognit.*, vol. 14, nos. 1–6, pp. 375–381, 1981.
- [34] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Philadelphia, PA, USA: SIAM, 1998.

- [35] J. Shlens. (Apr. 2014). "A tutorial on principal component analysis." [Online]. Available: <http://arxiv.org/abs/1404.1100>
- [36] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, Aug. 1994.
- [37] W. Liu and W. Lin, "Additive white Gaussian noise level estimation in SVD domain for images," *IEEE Trans. Image Process.*, vol. 22, no. 3, pp. 872–883, Mar. 2013.
- [38] D.-H. Shin, R.-H. Park, S. Yang, and J.-H. Jung, "Block-based noise estimation using adaptive Gaussian filtering," *IEEE Trans. Consum. Electron.*, vol. 51, no. 1, pp. 218–226, Feb. 2005.
- [39] A. Amer and E. Dubois, "Fast and reliable structure-oriented video noise estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 1, pp. 113–118, Jan. 2005.
- [40] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, Sep. 1999.
- [41] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, Sep. 1936.
- [42] G. W. Stewart, *Matrix Algorithms: Basic Decompositions*, vol. 1. Philadelphia, PA, USA: SIAM, 1998.
- [43] P. Chatterjee and P. Milanfar, "Clustering-based denoising with locally learned dictionaries," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1438–1451, Jul. 2009.
- [44] W. Dong, L. Zhang, G. Shi, and X. Wu, "Nonlocal back-projection for adaptive image enlargement," in *Proc. 16th IEEE Int. Conf. Image Process.*, Cairo, Egypt, Nov. 2009, pp. 349–352.
- [45] S. Dai, M. Han, Y. Wu, and Y. Gong, "Bilateral back-projection for single image super resolution," in *Proc. IEEE Int. Conf. Multimedia Expo*, Beijing, China, Jul. 2007, pp. 1039–1042.
- [46] J. W. Tukey, *Exploratory Data Analysis*. Reading, MA, USA: Addison-Wesley, 1977.
- [47] P. Milanfar, "A tour of modern image filtering: New insights and methods, both practical and theoretical," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 106–128, Jan. 2013.
- [48] L. Zhang, D. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
- [49] S. G. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1522–1531, Sep. 2000.



Qiang Guo received the B.S. degree in information and computing science from Shandong University of Technology, Zibo, China, in 2002, and the M.S. and Ph.D. degrees in computer science from Shanghai University, Shanghai, China, in 2005 and 2010, respectively.

He is an Associate Professor with Shandong Provincial Key Laboratory of Digital Media Technology, School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan, China. His research interests include

image restoration, sparse representation, and object detection.



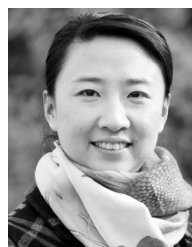
Caiming Zhang received the B.S. and M.S. degrees in computer science from Shandong University, Jinan, China, in 1982 and 1984, respectively, and the Ph.D. degree in computer science from Tokyo Institute of Technology, Tokyo, Japan, in 1994.

He was a Post-Doctoral Fellow with University of Kentucky, Lexington, KY, USA, from 1998 to 1999. He is currently a Professor with the School of Computer Science and Technology, Shandong University, and a Distinguished Professor with the School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan. His research interests include computer aided geometric design, computer graphics, information visualization, and medical image processing.



Yunfeng Zhang received the B.S. degree in computational mathematics and application software from Shandong University of Technology, Jinan, China, in 2000, and the M.S. degree in applied mathematics and the Ph.D. degree in computational geometry from Shandong University, Jinan, in 2003 and 2007, respectively.

He is a Professor with Shandong Provincial Key Laboratory of Digital Media Technology, School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan. His research interests include computer aided geometric design, image processing, computational geometry, and function approximation.



Hui Liu received the B.S., M.S., and Ph.D. degrees in computer science from Shandong University, Jinan, China, in 2001, 2004, and 2008, respectively.

She is a Professor with the Shandong Provincial Key Laboratory of Digital Media Technology, School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan. Her research interests include computer aided geometric design, medical image processing, and shape modeling.