An Enhanced Low Rank Approximation SVD-Based Method for Image Denoising

N. Sreekanth

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).

1. 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. To overcome the problems caused using the fixed transforms, Aharon et al. [5] proposed an adaptive representation technique using K-means and singular value decomposition (called K-SVD), which uses a greedy algorithm to learn an over complete dictionary for image representation and denoising. Under the assumption that each image patch can be represented by the learned dictionary, Elad and Aharon [6] 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. Dabov et al. [7] 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 de-noising results, its computational cost is very high.

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 [8]. This implies that we can achieve a good estimation of the group by taking only a few largest singular values and corresponding singular vectors. 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 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, X can be represented as

$$X = \sum_{i=1}^{N} a_i \phi_i$$  \hspace{1cm} (1)

where $a_i$ (i = 1, ..., N) are the representation coefficients of the image X in terms of the basis functions $\phi_i$ (i = 1, ..., N). $\phi_i$ can either be chosen as a prespecified basis, such as wavelet, curvelet, contourlet, shearlet, 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. [5] 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 X to form a data matrix $S = \{s_1, s_2, ..., s_m\} \in \mathbb{R}_n^{m \times n}$, where $m$ is the number of pixels in each patch, $s_j$ (i = 1, ..., n) are the image patches ordered as columns of S and n is the number of patches. Then the dictionary is learned by solving the following optimization problem:

$$\min_{\Phi, \Lambda} \sum_{i=1}^{n} \|s_i - \Phi a_i\|^2_2 \quad s.t. \|a_i\|_0 \leq \beta$$  \hspace{1cm} (2)

Where $\Phi \in \mathbb{R}^{m \times p}$ is the dictionary of p column atoms, $\Lambda = (a_1, a_2, ..., a_n) \in \mathbb{R}^{p \times n}$ is a matrix of coefficients, $\beta$ indicates the desired sparsity level of the solution, and the notation $\|a_i\|_0$ stands for the count of the nonzero entries in $a_i$. Based on the learned dictionary $\Phi$, S can be represented as

$$S = \Phi \Lambda$$  \hspace{1cm} (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 S. The principal components transform of S is calculated as

$$A = \Phi^T (S - E(S))$$  \hspace{1cm} (4)

with $\Phi$ defined by $\Omega_s = \Phi \Lambda \Phi^T$  \hspace{1cm} (5)

where E(S) is the mean of matrix vectors, $\Omega_s$ is the covariance matrix of S, $\Phi$ is the eigenvector matrix, and $\Lambda = \text{diag}(\lambda_1, ..., \lambda_m)$ is the diagonal eigenvalue matrix with $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_m$  \hspace{1cm} (6)

It can easily be derived that the covariance matrix of the matrix A equals

$$\Omega_s = \Phi^T \Omega_s \Phi = \Lambda$$  \hspace{1cm} (7)

which implies that the entries of 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 S is a decomposition of the form [5].

$$s = U \sum \sigma_i v_i^T = \sum_{i=1}^{n} \sigma_i u_i^T v_i$$  \hspace{1cm} (8)

where $U = (u_1, ..., u_n) \in \mathbb{R}_n^{m \times m}$ and $V = (v_1, ..., v_n) \in \mathbb{R}_n^{n \times n}$ are the matrices with orthonormal columns, $U^T U = V^T V = I$, and where the diagonal matrix $\Sigma = \text{diag}(\sigma_1, ..., \sigma_n)$ has nonnegative diagonal elements appearing in non-increasing order such that $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n \geq 0$  \hspace{1cm} (9)

The diagonal entries $\sigma_i$ of $\Sigma$ are called the singular values of S, while the vectors $u_i$ and $v_i$ are the left and right singular vectors of S, respectively. The product $u_i v_i^T$ in (8) can be considered as an adaptive basis, and $\sigma_i$ as the representation coefficient. In fact, SVD and PCA are intimately related. PCA can be performed by calculating the SVD of the data matrix $(1/\sigma_i^2)^{1/2} s_i$. 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. 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 $Y$ denote a noisy image defined by

$$Y = X + E$$  \hspace{1cm} (10)

where $X$ is the noise-free image, and $E$ represents the additive white Gaussian noise (AWGN) with the standard deviation $\tau$. In this paper, we use a vectorized version of the model

$$y = x + e$$  \hspace{1cm} (11)

Given a noisy observation $y$, our aim is to estimate $x$ as accurately as possible.
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 \( x \) 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 \( y \) is first divided into \( M \) overlapping patches denoted by \( \{ y_{i}^{M} \} \), where \( y_{i} \) is a vectorized format of the \( i \) th image patch. For each patch \( y_{j} \), its similar patch group is formed by searching similar patches from \( \{ y_{i}^{M} \} \). Next, each similar patch group is denoised by the low-rank approximation in SVD domain. Third, the denoised image \( \hat{X}_{0} \) 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 \( \hat{Y} \) 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 \( y_{j} \) with size \( \sqrt{m} \times \sqrt{m} \), the block-matching method finds its similar patches from \( \{ y_{i}^{M} \} \) by a similarity metric. In [9], 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 [7], 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

\[
s(y_{j}, y_{c}) = \left\| y_{j} - y_{c} \right\| \tag{12}
\]

where \( \left\| \cdot \right\| \) denotes the Euclidean distance and \( y_{c} \) is a candidate patch. The smaller \( S(y_{j}, y_{c}) \) is, the more similar \( y_{j} \) and \( y_{c} \) are. The reference patch \( y_{j} \) and its \( L \)-most similar patches denoted by \( \{ y_{c, i} \}^{L}_{i=1} \) are chosen to construct a group matrix using each similar patch as a column of the group matrix, and its corresponding group matrix \( P_{j} \) is formed by

\[
P_{j} = \begin{bmatrix} y_{j}, y_{c,1}, \ldots, y_{c,L} \end{bmatrix} \tag{13}
\]

Due to \( P_{j} \) being made up of the noisy patches, it can be represented as

\[
P_{j} = Q_{j} + N_{j} \tag{14}
\]

where \( Q_{j} \) and \( N_{j} \) 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. 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 \( Q \) and \( P \) instead of \( Q_{j} \) and \( P_{j} \) by a slight abuse of notation. Now our task is to estimate the noise-free group matrix \( Q \) from its noisy version \( P \) as accurately as possible. Ideally, the estimate \( \hat{Q} \) should satisfy

\[
\left\| P - \hat{Q} \right\| = \tau \tag{15}
\]

where \( \left\| \cdot \right\| \) is the Frobenius norm and \( \tau \) is the standard deviation of noise.

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, is to develop an iterative version for the basic denoising method. 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.

\begin{algorithm}
\textbf{Algorithm 1: Proposed Denoising Algorithm}
\textbf{Input:} Noisy image \( Y \)
\textbf{Output:} Denoised image \( \hat{X} \)
\begin{enumerate}
\item \( \tau \) \quad \text{Estimate noise standard deviation by computing the median absolute deviation (MAD) of the finest wavelet coefficients.}
\item \( \{ P_{j} \}^{C}_{j=1} \) \quad \text{Group image patches with the similarity metric defined by Eq. (12) and form group matrices;}
\item \( \text{For each } j \in [1, C] \) \quad \text{do}
\item \( \hat{Q}_{j} \) \quad \text{Calculate the LRA of } P_{j} \text{ via Eq. (21);}
\item \( W_{j} \) \quad \text{Compute the weight for } \hat{Q}_{j} \text{ via Eq. (24);}
\item \text{end for}
\item \( \hat{X}_{0} \) \quad \text{For each patch existing multiple different estimates, aggregate its estimates via Eq. (23);}
\item \( \hat{X} \) \quad \text{Aggregate all } \hat{Q}_{j} \text{ via Eq. (25);}
\item \( \hat{Y} \) \quad \text{Generate a new noisy image via the back projection described by Eq. (29);}
\item \( \hat{X} \) \quad \text{Obtain the final denoised image by performing Step 2 to Step 8 for } \hat{Y} \text{.}
\end{enumerate}
\end{algorithm}
The basic idea of back projection is to generate a new noisy image by adding filtered noise back to the denoised image.

\[ \hat{Y} = \hat{X}_0 + \delta (Y - \hat{X}_0) \]  

(16)

where \( \delta \in (0, 1) \) is a constant projection parameter and \( \hat{X}_0 \) is the denoised result produced by the first stage. Note that when \( \delta \to 0 \), \( \hat{Y} \to \hat{X}_0 \). On the contrary, if \( \delta \to 1 \), \( \hat{Y} \to Y \). 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 \( \hat{X}_0 \) by denoising \( \hat{Y} \) with the proposed three processing steps in Sections III-A–III-C, i.e., patch grouping, SVD-based denoising, and aggregation, respectively. It is necessary to point out that the noise variance of \( \hat{Y} \), denoted by \( \tilde{\sigma}^2 \), needs to be updated in the SVD-based denoising step. We employ the estimator presented in [26] to determine \( \tilde{\sigma}^2 \), which is written as

\[ \tilde{\sigma}^2 = \sqrt{\gamma^2 \|Y - \hat{X}_0\|^2} \]  

(17)

where \( \gamma \) 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 \( \tau \) 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 Structure-similarity (SSIM) index, 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. SSIM 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.

B. Denoising Performance

To quantitatively evaluate the denoising performance of our method, we compare it with two popular existing state-of-the-art image denoising methods: 1) K-SVD [6]; 2) BM3D-SAPCA [7]. Both of these denoising methods utilize the self-similarity of natural images to suppress noise. These denoising algorithms contain some control parameters, which should be tuned according to the noise level of the image. In our experiments, we use the default parameters settings suggested by the respective authors. The source codes of these denoising methods can be downloaded from the respective authors’ websites. In addition, the proposed method was implemented in MATLAB programming language due to its simplicity. In our experiments, we empirically set \( L = 85 \), \( \delta = 0.5 \), and \( \gamma = 0.65 \) for all noise levels. Depending on the amount of noise present in the image, we set the patch size \( 9 \times 9 \) if \( \tau < 20 \), \( 10 \times 10 \) if \( 20 \leq \tau < 40 \), and \( 11 \times 11 \) if \( \tau \geq 40 \). Visual comparisons of denoising results on different standard test images corrupted by AWGN with standard deviation \( \tau \) = 10. Viz. (a) Original image. (b) Grey image (c) AWGN Noisy image. (d) Existing Method. (e) BM3D-SAPCA [7]. (f) Proposed (Patch Grouping-LRA-SVD) method are shown in figure 4.1 to 4.8.
In Table I, we quantify the performances of Existing and Proposed algorithms for the test images with different noise levels in terms of PSNR and SSIM. 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 exploit the nonlocal redundancies in these images by the patch grouping procedure described in Section III-A. The SSIM results of different algorithms are also tabulated in Table I. It can be observed that our method has higher SSIM measures than other methods except BM3D-SAPCA in a word, the quantitative results by our method are competitive with BM3D-SAPCA and clearly superior to K-SVD. In terms of visual quality, our method also is comparable and even superior to the state-of-the-art denoising methods.

Fig. 4.9. Visual comparisons of denoising results on Lena image corrupted by AWGN with standard deviation 10. (a) Original image. (b) Grey image (c) AWGN Noisy image. (d) K-SVD (e) BM3D-SAPCA. (f) Proposed (Patch Grouping-LRA-SVD) method.

V. DISCUSSION

The nonlocal self-similarity of natural images plays an important role in image denoising. 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. 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 K-SVD and BM3D-SAPCA by better preserving the local geometric structure.

2) 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-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-SAPCA, the second denoising stage is directly applied on the original noisy image that is grouped into 3-D data bases on the patch similarities from the denoised image in the first 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 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 popular K-SVD. 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. 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.

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. M. Aharon, M. Elad, and A. Bruckstein, “K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation,” IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
6. 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.
7. J. Sareendran , R. Sridhar „Rajiv arya. “GPRS BASED MONITORING SYSTEM USING IMAGE SENSOR”, International Journal of Pure and Applied Mathematics Volume 119 No. 12 2018, 14387-14392 ISSN: 1314-3395.
8. 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.
9. 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.
10. Guo, Qian, et al. “An efficient SVD-based method for image denoising.” IEEE transactions on Circuits and Systems for Video Technology 26.5 (2016): 868-880.
11. N. Nithyanaandam, K. Venkatesh, M. Rajesh, Transfer The Levels Of The Monitored Carbon, Nitrogen Gases From The Industries, International Journal of Recent Technology and Engineering, Volume-7 Issue-6S3 April 2019.