Bayesian Framework with Non-local and Low-rank Constraint for Image Reconstruction

Zhonghe Tang, Shengzhe Wang, Jianliang Huo, Hang Guo, Haibo Zhao and Yuan Mei

Department of Guided and Information Engineering, Southwest Institute of Technology and Physics, 630811, Sichuan, China

E-mail: sinawang21@126.com, huoliaang@163.com, 408306714@qq.com, haibozhao125@sina.com, meiyuan1984@foxmail.com

Abstract. Built upon the similar methodology of 'grouping and collaboratively filtering', the proposed algorithm recovers image patches from the array of similar noisy patches based on the assumption that their noise-free versions or approximation lie in a low dimensional subspace and has a low rank. Based on the analysis of the effect of noise and perturbation on the singular value, a weighted nuclear norm is defined to replace the conventional nuclear norm. Corresponding low-rank decomposition model and singular value shrinkage operator are derived. Taking into account the difference between the distribution of the signal and the noise, the weight depends not only on the standard deviation of noise, but also on the rank of the noise-free matrix and the singular value itself. Experimental results in image reconstruction tasks show that at relatively low computational cost the performance of proposed method is very close to state-of-the-art reconstruction methods BM3D and LSSC even outperforms them in restoring and preserving structure.

1. Introduction

The search for efficient image denoising methods still is a valid challenge. The local smoothing methods and the frequency domain filter aim at a noise reduction and at a reconstruction of the main geometrical configurations but not at the preservation of the fine local structure, details and texture. Early local denoising method relied on various smoothness assumptions, such as anisotropic filtering, total variation or image decomposition on fixed basis. In some degree, non local methods all explicitly exploit the self-similarities of natural images and represent each clean image patch as a linear combination of a few elements from a basis set called a dictionary. In essence, they are all data driven representation methods. However, the difference among these denoising algorithms is the source of atoms in the dictionary and the constraints to the representation coefficients. In NLM [1], the atoms are from the noisy image and the coefficients depend on the similarly between two noisy image patches. Based on a nonlocal regularity assumption, the NLM [1] estimates an unknown pixel by a weighted average of local and non-local pixels throughout the entire image. To BM3D [2], it is based on classical fixed orthogonal dictionaries and requires the corresponding representation coefficient to be sparse. It restores noisy patches by finding more similar patches in the noisy image using patch matching, stacking fragments together into a 3D signal group, and denoising the group by using collaborative filtering. Inspired by BM3D[2] , K-SVD[3] scheme, employs a joint sparse decomposition based on a learned, possibly over-complete, dictionaries adapted to specific image

patches and assumes if two image patches are similar[4], the coefficients should be similar in structure or the distribution of nonzero entries.

Compared with conventional NLM [1] and BM3D [2], LSSC [4] takes both the structure and the sparsity into consideration. However, there still exist some shortcomings. To improve universality of dictionary, its size should be huge, but the learning of a huge dictionary and the determining of a high-dimensional but sparse coefficient matrix are very computationally expensive. In essence, the structured similarity between the columns of coefficient matrix is equivalent to the constraint that the image matrix or coefficient matrix [2, 4, 5, 6, 7] has a low rank. Data matrix stacked by similar image patches is born with low rank characteristic [4, 7], but this property is rarely exploited directly in spatial domain.

In this paper, an image denoising method is proposed which is built upon the methodology of 'grouping and collaboratively filtering'. It combines two now classical ideas into a single framework: The low rank joint decomposition and the non-local means idea. The fundamental idea is to convert the problem of removing noise to a low rank matrix approximation problem and a weighted nuclear norm is defined to replace the classical nuclear norm. we focused primarily on low rank constraint in the spatial domain instead of the transform domain. For each image patch centered at the current pixel a similar patch series is found in a search window. The numerical rank of corresponding noise-free matrix is predicted based on the estimation of noise variances and singular values of noisy data matrix, then some significant singular values are retained according the numerical rank. Then, for each retained singular value, an adaptive thresholds or weight are calculated. Then shrinkage is carried out on the retained singular values. Lastly, the recovered image patch can be easily reconstructed based on the shrunk singular values.

2. Related work

Essentially there are three types of image denoising method: local, non-local and the mixture of them. Our approach is a non-local approach. Thus, we will only discuss the most related non-local techniques.

In the non-local image denoising, we always are required to search n-1 similar image patches $s_{i,j}$, $j=1,\ldots,n-1$ for a given reference patch s_i from a search window $\mathbf{W}_i$ sharing the same centre with given reference patch, then group them into an image patch set $\mathbf{S}_i=\{s_i,s_{i,1},\ldots,s_{i,n-1}\}$. Measuring the similarity within patches set $\mathbf{S}_i$ is one of fundamental problems in image denoising. Another key issue is to how to explore the similarity and gather all the helpful information about the noise-free version of s_i distributing among the n image patches to restore the noisy patch s_i. The key prior is these image patches S_i are similar and correlated[8].

In the conventional NLM [1], the weight $w_{i,j}$ is computed based on the similarity between the current reference patch s_i and candidate patch $s_{i,j}$. Then, the noise-free version of the estimated s_i is computed as

$$\hat{s}_i=\sum_{j=1}^{n}\hat{w}_{i,j}s_{i,j},\hat{w}_{i,j}=w_{i,j}/\sum_{j=1}^{n}w_{i,j} \tag{1}$$

To find enough image patches whose noise-free versions are almost same as that of given s_i, the size of search window $\mathbf{W}_i$ should be very large enough, but it will cause the search very computationally expensive. At the same time, only the similarity between s_i and $s_{i,j}$ is considered and exploited instead of any two patches in the n patches set.

In the BM3D [2] method, A shrinkage in 3D transform domain such as wavelet shrinkage and Wiener filter are applied on the 3D coefficient array of patch set $\mathbf{S}_i$ to suppress noise. The result is

obtained by synthesizing a clean image from the de-noised patches. However, essentially fixed orthogonal transform is a kind of local processing, and only local interactions of neighboring pixels and patches are considered or the similarity among a local range is exploited. The definition of local range depend on the grouping order how the 2D patches are grouped into the 3D array. In a word, it is not completely non local de-noising method.

To solve the mentioned issues, some sparse models for de-noising are proposed. In the K-SVD [3], an over-complete dictionary $\mathbf{D}$ adapted to specific patches will be learned. It is assumed the each image patches s_i can be sparsely represented by the over-complete dictionary $\mathbf{D}$. To determine the sparse coefficient, we need solve the following optimization problem

$$\min_{\alpha} \left|\alpha\right|_{l^p},\ 0 \le p \le 1 \text{ and } |\, s_i - \mathbf{D}\alpha \,| < \delta \tag{2}$$

To exploit the structure self-similarity of natural image, in the LSSC, some similar image patches $s_{i,j},\ \ i = 1,\ldots,n-1$ are employed to restrict the sparse coefficient of s_i and its structure. An important assumption in LSSC [4] is that the sparse coefficients of some similar image patches $s_{i,j}$, should possess similar structure and decomposition, especially the distribution or locations of nonzero entries. To determine the sparse coefficient matrix $\mathbf{A}$, we need solve the problem

$$\min_{\mathbf{A}} \left|\mathbf{A}\right|_{p,q}, |\, \mathbf{S}_i - \mathbf{D}\mathbf{A} \,| < \delta \tag{3}$$

In fact, the similarities of the structures of the image patches can be represented from many aspects, not only limited in the distribution of nonzero entries. At same time, all the denoising methods based on overcomplete dictionary suffer from a drawback: the learning of overcomplete dictionary and the determination of structured coefficient vector or matrix are very computationally expensive [9,10].

3. Image denoising model based on non-local and low-rank

3.1. Image denoising and low-rank decomposition

Concretely, firstly let us define a similar image patch set $\mathbf{S}_i$ for the given patch s_i, namely, $\mathbf{S}_i = \{s_{i,j} \,|\, \mathrm{sim}(s_i, s_{i,j}) > \varepsilon, s_{i,j} \subset \mathbf{W}_i\}$, where $\mathrm{sim}(s_i, s_{i,j})$ is used to measure the similarity between reference patch s_i and candidate block $s_{i,j}$. Two more similar patches lead a higher $\mathrm{sim}(s_i, s_{i,j})$ value. And the higher the similarity threshold ε is, the less the image patches with higher similarity are grouped into the set S_i. The effectiveness of de-noising degraded image largely depends on how to measure the similarity between two patches and how to utilize the similarity in the de-noising methods. In this paper, in order to simplify the problem the simple Euclidean distance is used to measure the similarity. The focus is to exploit structure self-similarity of natural image by low-rank approximation or decomposition to restore the noisy patches.

In this paper, the image patch $s_{i,j}$ is stacked as the column vector, and then image patch array $\mathbf{S}_i$ can be considered as a 2D data matrix $\mathbf{S}_i \in R^{m \times n}$, where n is the number of patches in the patch set $\mathbf{S}_i$ and m is the number of the pixels in each image patch. In general case, n is bigger than m. According the analysis above, the patch group $\mathbf{S}_i$ can be decomposed into three parts: commonness information $\mathbf{A}_i \in R^{m \times n}$, the difference $\mathbf{E}_i \in R^{m \times n}$, and the noise $\mathbf{N}_i \in R^{m \times n}$, i.e., $\mathbf{S}_i = \mathbf{A}_i + \mathbf{E}_i + \mathbf{N}_i, \mathrm{rank}(\mathbf{A}_i) \le \min(m, n),\ . \left\|\mathbf{E}_i\right\|_{l^0} < m \times n$ .In fact the commonness and difference are relative. Generally, the bigger the rank of commonness $\mathrm{rank}(\mathbf{A}_i)$ is, the less the energy $\left\|\mathbf{E}_i\right\|_{l^2}$ of difference $\mathbf{E}_i$ is and the sparser the difference is. That means by adjusting the rank of the commonness, we can assign partial difference to the commonness $\mathbf{A}_i$ and other unimportant difference to the

uncertainty information $\mathbf{N_i}$. At same time, to image de-noising problem both $\mathbf{A}_i$ and $\mathbf{E}_i$ are signals we need recover[9,11,12]. It is not necessary to recover the commonness and difference separately. Therefore, the low-rank sparse decomposition can be simplified as a low-rank decomposition or approximation problem mathematically

$$\mathbf{S}_i = \mathbf{A}_i + \mathbf{N}_i, \text{rank}(\mathbf{A}_i) < \min(m,n) \tag{4}$$

Therefore, our focus is to solve the following optimization problem

$$\mathbf{A}_i^* = \min_{\mathbf{A}} \text{rank}(\text{A}), \text{s.t.} \|\text{S}_i\text{-A}\|_{l^2} < \varsigma \tag{5}$$

The parameter ς depends on the variance of noise. Applying the relaxation to the problem (5) yields a new optimization problem as follows

$$\mathbf{A}_i^* = \arg\min_{\mathbf{A}} \{\frac{1}{2}\|\mathbf{A}-\mathbf{S}_i\|_F^2 + \tau \times \text{rank}(\mathbf{A})\} \tag{6}$$

Unfortunately, (6) is a highly no convex optimization problem, and corresponding minimization problem is extremely difficult (it always is NP-hard and hard to approximate). At present, no efficient numerical solution is known[13].

To solve the problem (6), Candes ect. [8] replaced the $\text{rank}(\mathbf{A})$ with the nuclear norm $\|\mathbf{A}\|_* = \sum_i \sigma_i$ where σ_i is the singular value of data matrix $\mathbf{A}$. Simple replacement yields a tractable optimization problem:

$$\mathbf{A}_i^* = \arg\min_{\mathbf{A}} \{\frac{1}{2}\|\mathbf{A}-\mathbf{S}_i\|_F^2 + \tau\|\mathbf{A}\|_*\} \tag{7}$$

Fortunately, this optimization problem (7) is convex and can be solved efficiently. If we denote the singular value decomposition of data matrix S_i as $S_i = \mathbf{U}_i\Sigma_i\mathbf{V}_i^T$, where $\Sigma_i = diag(\lambda_{i,1},\ldots,\lambda_{i,r})$, $r = \min(m,n)$ and $\lambda_{i,j}$ is singular value of data matrix S_i (including zero). Then, the optimal solution of (8) is computed as [14, 15]

$$T_\tau(\mathbf{S}_i) = \arg\min_{A} \{\frac{1}{2}\|\mathbf{A}-\mathbf{S}_i\|_F^2 + \tau\|\mathbf{A}\|_*\} \tag{8}$$

Here, $T_\tau(\mathbf{S}_i)$ is the singular value shrinkage operation and can be defined as

$$T_\tau(\mathbf{S}_i) := \mathbf{U}_i T_\tau(\Sigma_i)\mathbf{V}_i^T, T_\tau(\Sigma_i) = diag\{(\lambda_{i,j}-\tau)_+\} \tag{9}$$

Where $(t)_+$ denote the positive part, defined as $\max(0,t)$. The remarkable difference between $\text{rank}(\mathbf{A})$ and $\|\mathbf{A}\|_*$ is that the former only depends on the number of nonzero singular values and not the specific value, but $\|\mathbf{A}\|_*$ is equal to their sum and depends directly on their values. This is just like the difference between l_0-norm and l_1-norm.

Essentially, the singular value shrinkage operation simply applies a soft-thresholding rule to the singular values of $\mathbf{S}_i$. Just like the common soft thresholding denoising in wavelet domain [16]. The determination of proper parameter or threshold τ is a key issue. It is necessary to utilize an adaptive threshold determined by the similarity of patch set, the variance of noise and the distribution of signal information and noise[17].

3.2. Effect of perturbation on the rank and the truncated SVD

Suppose that $\mathbf{S}_i$ is a matrix containing many column signals that are contaminated with a certain level of noise. The SVD resolves the data in noisy $\mathbf{S}_i$ into r mutually orthogonal components [18] as following

$$\mathbf{S}_i = \mathbf{U}_i \Sigma_i \mathbf{V}_i^T = \sum_{j=0}^{r} \lambda_{i,j} u_{i,j} v_{i,j}^T = \sum_{j=0}^{r} \lambda_{i,j} z_{i,j} \tag{10}$$

Where $z_{i,j} = u_{i,j} v_{i,j}^T$. $u_{i,j}$ and $v_{i,j}$ are column vectors of orthogonal matrix $\mathbf{U}_i$ and $\mathbf{V}_i$. In other words, the SVD decomposition (10) can be regarded as Fourier expansion or wavelet decomposition, and consequently, $\lambda_{i,j}$ can be interpreted as the proportion of signal $\mathbf{S}_i$ lying in the "direction" of $z_{i,j}$ whose rank is low and equal to one. Therefore, the SVD can be a useful tool in applications involving the need to sort through noisy data and lift out relevant information.

The noise-free version $\mathbf{A}_i$ always is assumed to be low-rank. But in most case, the matrix $\mathbf{A}_i$ and its rank is unknown. Whether the rank of noise-free matrix can be estimated accurately based on the noisy matrix is a key problem to low- rank decomposition. The effect of noise on singular values is important for the estimation the rank of $\mathbf{A}_i$. The following theory will show the perturbation of singular values due to a certain level of noise.

Theorem.1 Perturbations and Rank [18]. For $\mathbf{A} \in \Re^{m \times n}$, where m the size of image is patch and n is the number of image patches. Denote the noisy version of $\mathbf{A}$ as $\mathbf{S} = \mathbf{A} + \mathbf{N} \in R^{m \times n}$. where matrix $\mathbf{N}$ contains perturbation or noise. If we let $\{\lambda_1, \lambda_2, \cdots, \lambda_r\}$ and $\{\beta_1, \beta_2, \cdots, \beta_r\}$ denote the singular value (nonzero as well as any zero ones) of matrix $\mathbf{S}$ and $\mathbf{A}$ as, $r = \min\{n, m\}$, then

$$\left|\lambda_t - \beta_t\right| \le \left\|\mathbf{N}\right\|_2^2, \text{ for each } t = 1, \ldots, r \tag{11}$$

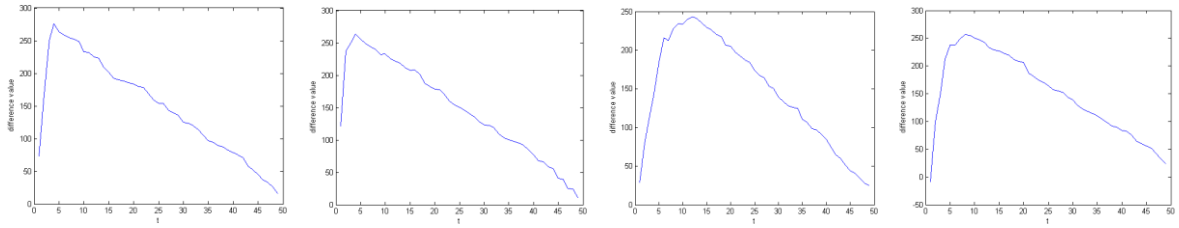
The proof can be found in the literature [18].

This theorem explains why singular values are the primary mechanism for numerically determining the rank of a matrix. If the rank of noise-free signal matrix $\mathbf{A}$ is equal to l, then the singular value $\beta_p, p = l+1, \ldots, r$ should be exactly zero. Therefore, Theory above can guarantee that $r - l$ singular value of the computed $\lambda_p, p = l+1, \ldots, r$ cannot be larger than $\|\mathbf{N}\|_2^2$. Therefore, if the singular values λ_p obeys

$$\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_q > \|\mathbf{N}\|_2^2 = \xi \ge \lambda_{q+1} \ge \ldots \ge \lambda_r \tag{12}$$

Then, it's reasonable to consider q to be the numerical rank of nosiy free signal $\mathbf{A}_i$. ξ is a threshold depending on the standard deviation of noIn the general case, the perturbation $\left\|\mathbf{N}\right\|_2^2$ is unknown. Fortunately, if a perturbation is Gaussian white noise, relatively accurate estimation of $\left\|\mathbf{N}\right\|_2^2$ can be obtained. Based on the estimation, we can predict the difference between the singular values for $\mathbf{A}$ and $\mathbf{S}$ is just related to $\left\|\mathbf{N}\right\|_2^2$.

To verify the theory above, zero mean white Gaussian noise with standard deviation σ =30 is added to the 512×512 Barbara image. From the noisy image and the noise-free image, four 7×7 image patch pairs a_i and s_i are selected randomly. For each noisy patch s_i , the top most similar 60 image patches are choose from a 25×25 search window. Then, the r singular values λ_t, β_t, t =1,…,r (r=49) of $\mathbf{S}_i$ and $\mathbf{A}_i$ are computed. Fig.1 show the difference $e_t = \lambda_t - \beta_t$ between λ_t and β_t.



(a) (b) (c) (d)

Figure 1. the difference of the singular values between the noisy data matrix S_i and the noise-free data matrix A_i. (a) the first example (b) the second example (c)the third example (d) the fourth example

It can be easily observed that there exist many similarities[19]. All the singular values β_t can be divided into two parts. The first parts is $\beta_t, t=1,\ldots,t_s$, and corresponding z_t represent the commonness and the similarity among the $\mathbf{S}_i$ and noise has little effect on the singular values when β_t is bigger or t is small. As the singular values decrease, the difference increases. The second part $\beta_t, t=t_{s+1},\ldots,r$ is small and always close to zero. And corresponding z_t always represents the randomness and uncertainty information, such as noise and the weak difference of contents. As t increase, the difference decrease. When t is close to rank of noise-free data matrix, the difference e_t reaches its peak point. In general case, the most information of signal is concentrated on the first part, and the second part is not important for the signal recovery.

3.3. Threshold determination of the nonlocal and low-rank denoising

For the singular value shrinkage operation, the key issue is to select a reasonable threshold. Obviously, applying a fixed threshold $\lambda_{i,k}$ (k is the truncated location and is equal to $\mathrm{rank}(\mathbf{A}_i)$) to all singular values λ_t is not reasonable. Although, the truncation can remove most noise, much residue still appear in the retained singular values. At same time, the noise is always distributed uniformly across the $z_{i,j}$, but signal is not. Therefore, it is necessary to design an adaptive threshold for each retained singular values $\lambda_t, i=1,\ldots,k$ to remove the noise contained in $\hat{\mathbf{S}}_i$ and preserve the information of signal.

According the analysis and experiment above, a fact can be found that the bigger the singular value $\lambda_t, i=1,\ldots k$ is, the less the difference is between the noisy matrix and the noise-free one. Namely, the bigger singular value is, the smaller the threshold should be. In a word, to adaptively adjust the parameter τ, the parameter τ is assumed to be proportional to noise variance and in inverse proportion to singular value. Therefore, we adopt the following strategy for determining the parameter $\tau_{i,j}$

$$\tau_{i,j}=\begin{cases} c\dfrac{\sigma^2}{\lambda_{i,j}} & if \quad j<k_i \\ \lambda_{i,k} & else \end{cases} \tag{13}$$

where k_i is the estimation of the rank of $\mathbf{A}_i$ and σ^2 is noise variance, $\lambda_{i,j}$ is the $j-\mathrm{th}$ singular value of $\mathbf{S}_i$, and c is predefined constants. For this, the noise standard deviation σ is estimated using a simple gradient-based estimator as [20,21],

$$\sigma=1.4826\times\mathrm{median}\left(\left|\nabla Y-\mathrm{median}(\nabla Y)\right|\right) \tag{14}$$

where ∇Y is the vectorized form of the gradient of the input nosiy image Y, which is calculated as

$$\nabla Y = 0.4082 \times \mathrm{vec}\left(Y * \begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix} \right) \tag{15}$$

Here, $\mathrm{vec}(\bullet)$ and $*$ denote the vectorization operation and the convolution operation, respectively.

3.4. The main frame of proposed denoising scheme

To conclude this section, we will summarize the main step of the proposed image denoising method. The concrete process is shown as follows .

Grouping the similar patches. For the given reference patch s_i, n top most similar patches $s_{i,j}$ are selected based the similarity between s_i and $s_{i,j}$. Then, these vectorized patches are grouped into data matrix $\mathbf{S}_i = [\mathrm{vec}(s_{i,1}), \ldots, \mathrm{vec}(s_{i,n})]$.

Estimating the rank of underlying matrix $\mathbf{A}_i$. Estimate the noise standard deviation σ and parameter $\xi = \kappa n \sigma$, and then compute the singular values of matrix $\mathbf{S}_i$. Lastly, estimate the numerical rank k_i based on theory 1.

Applying the adaptive shrinkage to each singular values. Based the rank estimated rank k_i , truncate the singular values and compute the adaptive parameter $\tau_{i,t}$. Then, apply the shrinkage to the retained singular values $\lambda_{i,t}$.

In our implementation, the image patches are selected with overlapping regions. Thus, each pixel is covered by several de-noised patches. In this paper, the value of each pixel in images is determined by taking the average of all the estimations for same location. In fact, we observe a fact the lower the estimated rank is, the better the quality of recovery is. Intuitively, an aggregation can be performed by a weighted averaging at those pixel positions where there are overlapping block-wise estimates just like BM3D. This is one direction of our future research.

4. Experimental results

In this section, the performance of the proposed denoising scheme is tested. The main innovation in the frame is the combination of non local idea and low rank.

Throughout this study, the patch size in our implementation is set as 7×7 for $\sigma \le 30$ and 9×9 for $\sigma > 30$ .The size of search window is fixed as 25×25. The parameter ξ is set as $\kappa n \sigma$ where n is size of patch and κ is a constant and chosen according to an empirical rule. The parameter ε [4] is set as $32\sigma / n$ for images scaled between 0 and 255 to all test images, which has shown to be appropriate in all of our non local denoising experiments. Iterations times are set as 5. In the experiment, δ and c is fixed as 0.035 and 115.

Table 1. The PSNR(dB) values of the denoised images with respect to different noise levels of Gaussian noise.

Image	σ	NLM	K-SVD	BM3D	LSSC	K-LLD	Proposed
Barbara	5	37.06	38.11	38.33	38.48	36.61	**38.51**
	10	33.16	34.42	34.97	**34.97**	33.03	34.93
	20	30.26	30.81	31.74	**31.57**	29.05	31.59
	25	29.08	29.58	30.72	**30.47**	27.27	30.51
	50	25.66	25.58	27.23	27.06	23.20	27.18
Fingerprint	5	35.15	36.66	36.59	36.70	35.85	**36.71**
	10	30.99	32.42	32.43	32.57	31.55	**32.62**
	20	27.25	28.45	**28.87**	28.78	28.14	**28.87**
	25	26.14	27.25	27.70	27.62	26.99	**27.78**
	50	22.96	23.25	24.53	24.25	22.34	**24.55**
Lena	5	37.94	38.62	**38.71**	38.69	37.61	38.61

	10	34.36	35.51	35.94	35.83	35.12	35.58
	20	31.61	32.38	33.07	32.90	32.37	32.77
	25	30.36	31.36	32.08	31.87	31.34	31.63
	50	27.32	27.75	29.05	28.87	25.15	28.81
Man	5	37.01	37.50	37.80	**37.89**	36.24	37.86
	10	33.11	33.55	33.94	**34.06**	33.09	33.97
	20	29.72	30.05	30.54	**30.64**	30.12	30.60
	25	28.38	29.08	29.62	**29.63**	29.22	29.59
	50	25.33	26.08	**26.81**	26.69	24.38	26.71
Couple	5	36.75	37.30	**37.49**	37.45	35.89	37.39
	10	32.90	33.48	**34.02**	33.98	33.08	33.91
	20	29.01	30.02	**30.76**	30.69	29.92	30.77
	25	27.99	28.88	**29.72**	29.61	29.02	29.71
	50	24.58	25.29	**26.46**	26.30	23.20	26.37

For performance evaluation, several common test images are considered (including Barbara, Lena, Figureprint, Man and Couple) which are corrupted by additive white Gaussian noise with standard deviation $\sigma \in \{5,10,20,25,50\}$. To illustrate the proposed denoising method more clearly and fully demonstrate the performance gains from the combination of non-local and low-rank, the proposed method is compared with five state-of-the-art methods which include NLM[1], BM3D[2], K-SVD[3], LSSC[4], and K-LLD[6]. Except LSSC algorithm, all the results are based on the source codes or executables released by the original authors. The default parameters are employed in the comparison algorithms. We try to realize LSSC. Unfortunately, our results are inconsistent with the original literature. To make a fair comparison, all the evaluation indexes of LSSC are from the literature [4].

4.1. Comparison of quantitative evaluation

The PSNR results on five test images are reported in Table 1 for different noise variance σ^2. The best performance, the second best and the third best is highlighted by red, green, and blue color in each cell, respectively. From table 1, we can conclude that the proposed non-local and low-rank denoising method significantly outperforms NLM, K-SVD and K-LLD. In term of PSNR, BM3D produce state-of-the-art result, however the results of the three schemes are nearly the same. Even on a certain test image containing similar structure pattern (edge and texture), such as Figureprint and Barbara, proposed scheme outperforms the BM3D and LSSC. That demonstrates low-rank is very suitable for capturing and restoring the patches with similar structure. For Barbara, when variance of noise is low, performance of proposed method outperforms other schemes. But as noise level increase, our method is not better than BM3D, but still better than LSSC. The reason is that when noise is weak, it is easy to capture the common pattern contained in data matrix. As the noise increase, the extraction of patterns becomes more difficult. On contrary, for the figureprint image, their local patterns are single and play a main role. No matter how strong the noise is, it is easy to capture the corresponding local pattern. Therefore, for Figureprint our scheme always outperforms the BM3D and LSSC.

It is not surprising that our method shows outstanding performances for the test image containing some textures or structure patterns. For Barbara image, the proposed method outperforms the K-SVD and NLM, and it is superior to K-SVD by 0.4~1.6dB and to LSSC by 0.2 dB. For fingerprint image, proposed method outperforms the BM3D, LSSC, K-SVD and NLM, and on average it is superior to BM3D by 0.1~0.2dB, to LSSC by 0.1~0. 3dB and to K-SVD by 0.1~1.3dB.

Table 2. The SSIM results for different standard images and several denoising methods.

Image	σ	NLM	K-SVD	BM3D	K-LLD	Proposed
Barbara	5	0.944	0.964	0.965	0.951	0.966
	10	0.928	0.934	0.942	0.920	0.941
	20	0.867	0.881	0.923	0.851	0.919
	25	0.825	0.850	0.887	0.804	0.892
	50	0.671	0.713	0.794	0.572	0.805

Fingerprint	5	0.965	0.987	0.987	0.985	0.988
	10	0.948	0.968	0.968	0.964	0.969
	20	0.889	0.922	0.929	0.924	0.933
	25	0.859	0.898	0.911	0.903	0.913
	50	0.712	0.775	0.830	0.778	0.838
Lena	5	0.914	0.946	0.944	0.928	0.951
	10	0.901	0.910	0.916	0.902	0.916
	20	0.837	0.863	0.877	0.864	0.872
	25	0.790	0.843	0.861	0.840	0.863
	50	0.674	0.761	0.799	0.552	0.801
Man	5	0.895	0.951	0.954	0.937	0.954
	10	0.873	0.900	0.907	0.887	0.908
	20	0.784	0.815	0.833	0.817	0.833
	25	0.724	0.779	0.805	0.789	0.804
	50	0.597	0.663	0.705	0.561	0.707
Couple	5	0.895	0.950	0.951	0.923	0.952
	10	0.877	0.897	0.909	0.890	0.909
	20	0.787	0.815	0.848	0.819	0.848
	25	0.745	0.779	0.820	0.793	0.818
	50	0.566	0.632	0.706	0.562	0.695

Compared with PSNR, SSIM is a popular evaluation for measuring the content and structure similarity between two images, which have been proved to be inconsistent with human eye perception. To understand how the combination of low-rank decomposition and non-local idea restores the structure, we also compare the SSIM of the proposed algorithm and other leading denoising techniques in the literature at different noise levels for 5 test images. The SSIM results on five test images are shown in table 2. The best performance and the second best are highlighted by red and green font in each cell, respectively. We conclude the proposed denoising method has achieved competitive SSIM performance to BM3D, K-SVD, KLLD and NLM for the most test images. It is not surprising that BM3D achieved some competitive PSNR, but in term of SSIM the proposed method is best. The success owe to the fact the proposed method aims to restore and preserve the local structure not minimize the numerical error.

Note that the parameters have not been optimized for speed and for quality in this experiment. They are set just based on experience. Optimized parameters may further improve the performances. This is another direction of future research. Because all the source codes are developed by different program language, it is not fair to measure the speed only depending on running time. Intuitively, all the algorithms (K-SVD, BM3D and LSSC) need search some similar image patches, the time spent on the search should be nearly same. But for time spent for the computation of the coefficient matrix, LSSC are computationally expensive. At same time, K-SVD and LSSC must learn a overcomplete dictionary based on a large amount of samples. It is natural conclusion that the proposed method is faster than LSSC and has the nearly same speed as BM3D algorithm.

4.2. Visual comparisons

Even though BM3D and LSSC perform a little better in terms of PSNR, they suffer from classical artifacts, as shown by the example of Fig.2 and Fig.3. Fig.2 shows the denoising results of Barbara image $(\sigma = 30)$ obtained only by BM3D, K-SVD, NLM and proposed method due to space limit. The visual comparisons of the detail and local region, especially the regions containing certain structure and texture, show that proposed method provides better visual quality. Some edges and textures are blurred even lost in BM3D, K-SVD and NLM, but they are better preserved in the proposed method.

In the first row of Fig.3, a certain texture structure is contained in the original image patch (as shown in Fig.3 (a1)). Compared with the results of the K-SVD, NLM and BM3D $(\sigma = 30)$, our result is more clear and complete. On the contrary, in other result there exist some obvious blur and loss. In the second row of Fig.3, in the original image block (as shown in Fig.3 (a2)) contains a straight edge.

Compared with the results of the K-SVD, NLM and BM3D, the edge in our result is clear and straight. However, the edge of compared algorithms is seriously deformed and bent. In the smooth region with a very weak structure (as shown in Fig.3 (a3)), noticeable reconstruction artefacts appear in the results obtained with BM3D, because some noise is considered as detail or structure. Unfortunately, in some degree our algorithm only aims to restore the main structure and loses some weak structure (shown in last row of Fig.3). The main reason is the truncation of SVD due to the inaccurate estimation of the rank of the corresponding noise-free data matrix. The loss of weak structure originates from the inaccurate truncation of singular values. Increasing the size of patch can avoid the weak structure loss. But it will increase the computational complexity of patch matching.

(a) (b) (c) (d)

Figure 2. Visual comparisons of Barbara (a) BM3D (b) K-SVD (c) NLM (d) Proposed method.

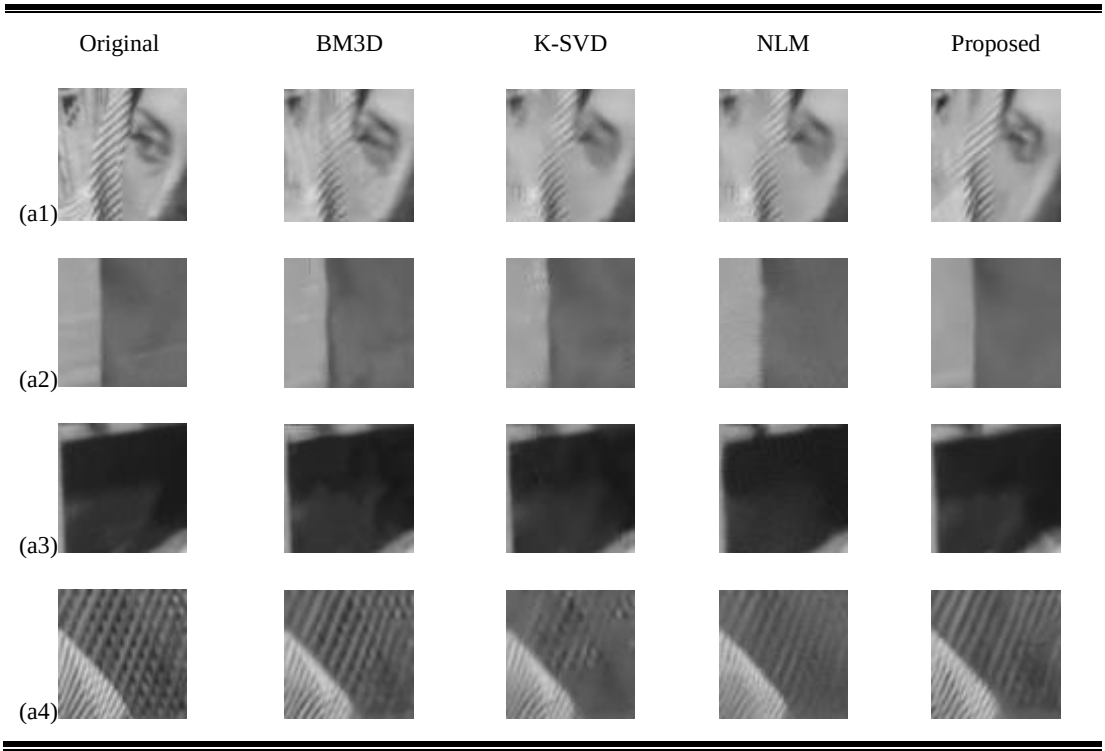


Figure 3. The local detail and structure comparison

5. Conclusion

In this paper, a simple but effective framework for image denoising is presented which works by exploiting the merits of self-similarity and similarity matrices low rank approximation. The weighted

nuclear norm is defined and used to replace the classical nuclear norm to improve adaptivity of the threshold used in the singular value shrinkage. The corresponding problems are effectively tackled. We need to take several practical steps: finding similar patches for each reference patch, computing the singular values of noisy data matrix and estimating the rank of corresponding noise-free data matrix, executing low rank approximation for each data matrix based on the weight matrix, reconstructing low rank matrices and obtaining denoised image. Experimental results demonstrate that at relatively low computational cost the proposed method is very close to state-of-the-art denoising technique BM3D and LSSC even outperforms them in some images containing many structure and texture in terms of both PSNR and SSIM.

Acknowledgements
This work was supported by the Southwest Institute of Technology and Physics. The authors thank Chatterjee and Dong, for providing their results.

References

[1]. Buades, B. Coll, and J. M. Morel, A review of image denoising methods, with a new one, Multiscale Model Simulation, 2(4), pp. 490–530, 2005.
[2]. K. Dabov, A. Foi, V. Katkovnik, and K. O. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, IEEE Transactions on Image Processing, 8(16), pp. 2080–2095, 2007.
[3]. M. Elad and M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries, IEEE Transactions on Image Processing, 12(15), pp. 3736–3745, 2006.
[4]. J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, Non-local sparse models for image restoration, IEEE 12th International Conference on Computer Vision, pp. 2272–2279, 2009.
[5]. W. Dong, X. Li, L. Zhang, and G. Shi, Sparsity-based image denoising via dictionary learning and structural clustering, in 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 457–464, 2011.
[6]. P. Chatterjee and P. Milanfar, clustering-based denoising with locally learned dictionaries, IEEE Transactions on Image Processing, 7 (18), pp. 1438–1451, 2009.
[7]. L. Zhang, W. Dong, D. Zhang, and G. Shi, Two-stage image denoising by principal component analysis with local pixel grouping, Pattern Recognition, 4(43), pp. 1531–1549, 2010.
[8]. E. J. Candes and B. Recht, Exact matrix completion via convex optimization, Foundations of Computational Mathematics, 9(6), pp. 717–772, 2009.
[9]. K. Konstantinides and K. Yao, Statistical analysis of effective singular values in matrix rank determination, IEEE Transactions on Speech Signal Processing, 36(5), pp. 757–763, 1988.
[10]. Hui Jiy, Chaoqiang Liuz, Zuowei Sheny and Yuhong Xu, Robust video denoising using low rank matrix completion, 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.1791–1798, 2010.
[11]. Yigang Peng, Ganesh. A., Wright J., Wenli Xu and Yi Ma, RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images, 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.763-770, 2010.
[12]. E. J. Candès, X. Li, Y. Ma, and J. Wright. Robust Principal Component Analysis? Journal of ACM 58(1), pp.1-37, 2009.
[13]. Tianyi Zhou and Dacheng Tao, GoDec: Randomized Lowrank & Sparse Matrix Decomposition in Noisy Case, In the 28th International Conference on Machine Learning (ICML), pp. 33–40, 2011.
[14]. E. J. Candès and Y. Plan. Matrix completion with noise. Proceedings of the IEEE, 98(6), pp. 925-936, 2009.
[15]. J .F. Cai, E. J. Candès, and Z. W. Shen, A singular value thresholding algorithm for matrix completion, SIAM Journal on Optimization. 20(4), pp. 1956–1982, 2010.

[16]. D. L. Donoho and I. M. Johnstone, Ideal spatial adaptation via wavelet shrinkage, Biometrika, 3(81), pp. 425–455, 1994.
[17]. S. G. Chang, B. Yu, and M. Vetterli, Adaptive wavelet thresholding for image denoising and compression, IEEE Transactions on Image Processing, 9(9), pp. 1532–1546, 2000.
[18]. Carl D.Meyer, Matrix Analysis and Applied Linear Algebra, Society for industrial and Applied Mathematics (SIAM), 2000.
[19]. D. D. Muresan and T.W. Parks, Adaptive principal components and image denoising, In Proceedings of the 2003 International Conference on Image Processing, pp. 1101–1104, 2003.
[20]. P. Chatterjee and P. Milanfar, patch-based near-optimal image denoising patch-based near-optimal image denoising, IEEE Transactions on Image Processing, 21(4), pp. 1635–1649, 2012.
[21]. Jianfeng Cai, Stanley Osher and Zuowei Shen, Split Bregman methods and frame based image restoration, Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal, 8(2), pp. 337-369, 2009