This ICCV paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the version available on IEEE Xplore.

External Patch Prior Guided Internal Clustering for Image Denoising

Fei Chen[1], Lei Zhang[2], and Huimin Yu[3]

[1]College of Mathematics and Computer Science, Fuzhou University, Fuzhou, China
[2]Dept. of Computing, The Hong Kong Polytechnic University, Hong Kong, China
[3]College of Information Science & Electronic Engineering, Zhejiang University, China
[3]State Key Laboratory of CAD & CG, China
chenfei314@fzu.edu.cn,cslzhang@comp.polyu.edu.hk,yhm2005@zju.edu.cn

Abstract

Natural image modeling plays a key role in many vision problems such as image denoising. Image priors are widely used to regularize the denoising process, which is an ill-posed inverse problem. One category of denoising methods exploit the priors (e.g., TV, sparsity) learned from external clean images to reconstruct the given noisy image, while another category of methods exploit the internal prior (e.g., self-similarity) to reconstruct the latent image. Though the internal prior based methods have achieved impressive denoising results, the improvement of visual quality will become very difficult with the increase of noise level. In this paper, we propose to exploit image external patch prior and internal self-similarity prior jointly, and develop an external patch prior guided internal clustering algorithm for image denoising. It is known that natural image patches form multiple subspaces. By utilizing Gaussian mixture models (GMMs) learning, image similar patches can be clustered and the subspaces can be learned. The learned GMMs from clean images are then used to guide the clustering of noisy-patches of the input noisy images, followed by a low-rank approximation process to estimate the latent subspace for image recovery. Numerical experiments show that the proposed method outperforms many state-of-the-art denoising algorithms such as BM3D and WNNM.

1. Introduction

Image denoising aims to restore the latent clean image $\mathbf{x}$ from its noise-corrupted version $\mathbf{y} = \mathbf{x} + \mathbf{v}$, where $\mathbf{v}$ is commonly assumed to be additive white Gaussian noise of standard deviation σ. As a classical and fundamental problem in low level vision, image denoising has been widely studied [28, 29, 6, 25, 19, 20, 8]. From a probability-based perspective, denoising could be treated as a Bayesian posterior estimation problem: $p(\mathbf{x}|\mathbf{y}) = p(\mathbf{y}|\mathbf{x})p(\mathbf{x})/p(\mathbf{y})$, where $p(\mathbf{y})$ is a constant once $\mathbf{y}$ is given. The conditional probability $p(\mathbf{y}|\mathbf{x})$ is called the data model, which can be written as

$$p(\mathbf{y}|\mathbf{x}) \propto \exp\left(-\frac{\|\mathbf{y}-\mathbf{x}\|_2^2}{\sigma^2}\right). \tag{1}$$

The distribution of probability $p(\mathbf{x})$ is called the prior model since it specifies a priori bias among the targeted patterns and is independent of data observation. The classic total variation (TV) [28] method actually assumes that natural image gradients exhibit heavy-tailed distributions, which can be fitted by Laplacian or hyper-Laplacian models [17, 34]. Many statistical models of wavelet coefficients have also been proposed, such as generalized Gaussian [12] and Gaussian scale mixture [25] models, etc.

The seminal work of nonlocal means (NLM) brings a new era of denoising by finding nonlocal similar patches within the image [2]. Such an internal self-similarity prior is widely exploited in patch-based denoising methods and has achieved a great success [7, 22, 15, 30]. A typical example is the so-called BM3D algorithm [10], which uses collaborative filtering in transformed domain by grouping nonlocal similar patches. The BM3D algorithm is very effective and it has been a benchmark in image denoising. Based on the principle of sparse and redundant representations, another effective category of denoising methods is the dictionary-based models. Largely owe to the success of the seminal work of KSVD [1], learning dictionaries from natural image patches has become a hot topic in image processing and computer vision [27, 21, 7].

Since image denoising is a typical ill-posed problem, the use of external natural patch priors to regularize the denoising process has proved to be very successful. Chatterjee and Milanfar [9] learned denoising bounds based on clustered natural patch database. Zoran and Weiss [33] learned

clean natural image patches using Gaussian mixture models (GMMs), and reconstructed the latent image by maximizing the expected patch log likelihood (EPLL). The MLP [3] method uses a multi-layer perceptron network to learn image denoising procedure from training examples, which consist of pairs of noisy and noise-free patches. Zontak and Irani [31] proposed an internal parametric prior to evaluate the nonlocal patch recurrence. They also found that patch recurrence holds across scales [32]. Though external patch priors based image denoising methods have shown competitive results, they do not make good use of image internal self-similarity. Mosseri et al. [24] combined the power of internal and external priors to improve image denoising performance. However, there is not global objective functional in this method, but a combining strategy to select the internal and external information. Obviously, how to integrate external patch priors and internal self-similarity into one framework to improve image denoising performance is an interesting problem to investigate.

Recently, low-rank approximation methods have exhibited exciting performance on denoising [13, 26, 14, 18, 5, 16]. It is accepted that the latent structure underlying image similar patches forms a low-dimensional subspace. Given a noisy observation set of similar patches $Y = [\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_m] \in \mathbb{R}^{n\times m}$ and $Y = X + V$, where X and V are the corresponding patch matrices of original image and noise, respectively, the independence of noises at the different pixels implies that

$$p(Y|X) \propto \exp\left(-\frac{\|Y - X\|_F^2}{\sigma^2}\right). \tag{2}$$

Under the assumption that the image patches in Y have similar structures, the latent clean data matrix X has a low-rank property, i.e., $p(X) \propto \exp(-\|X\|_*)$. In terms of logarithmic likelihood $E = -\ln p$, we can minimize the posterior energy:

$$E(X) = \frac{1}{\sigma^2}\|Y - X\|_F^2 + \tau\|X\|_*, \tag{3}$$

where τ is a positive constant. It is shown in Cai et al. [4] that the optimal solution to this problem is $\hat{X} = US_\tau(\Sigma)V^T$ where $Y = U\Sigma V^T$ is the SVD of Y and $S_\tau(\Sigma)$ is the soft-thresholding function on Σ with parameter τ. In [18], the nuclear norm minimization has been extended to weighted nuclear norm minimization (WNNM) and has achieved excellent denoising results. However, such methods face two challenges:

1. How to measure the similarity of patches for low-rank denoising method? For example, Fig. 1 shows a given reference patch (a) and two candidate patches (b) and (c). One commonly used similarity measure is the Euclidean distance (such as in BM3D [10] and WNNM [18]). It means that only patches whose l_2 distance to the reference patch is smaller than a fixed threshold are considered as similar patches to the reference one. However, there is no guarantee that Euclidean distance is a good choice for patch similarity measure, since the image patch space is not a Euclidean space. We might conclude that patch (b) is closer to (a) than (c) by Euclidean distance. However, by visual comparison, (a) and (c) have the same edge structure, and their collaborative filtering could preserve patch structure better. Therefore, subspace structure and some structure measure should be taken into account in patch grouping.

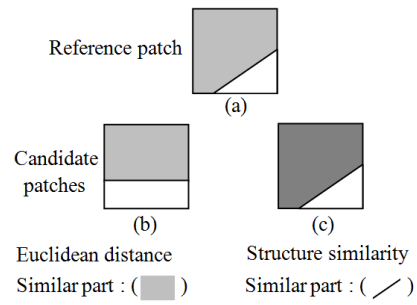


Figure 1. Similarity measure between the reference patch and the candidate patches.

2. How to find similar patches on a global scope? Block matching [10] is a simple and effective method for finding image patches similar to a given reference one. This method has shown convincing results, but its efficiency is low. In practice, block matching is often performed within a relatively large local window instead of globally. This stems from the assumption that many similar patches can be found in the local region. However, some patches with salient structures such as round edges or corners do not have repetitive patterns within the neighborhood. Moreover, enforcing local similarity in non-repetitive patterns will in particular cause ringing artifacts (please refer to Fig. 2). In such case, globally searching for the most similar patches in the whole image becomes very important.

In this work, we propose a new framework of patch clustering based low-rank regularization to address the above two issues. In order to obtain the latent patch subspaces with different structures, we perform patch clustering guided by a learned Gaussian mixture prior. By utilizing external patch prior to guide internal clustering, an energy function with low-rank regularization is constructed for image denoising. By minimizing the proposed objective functional, denoising and patch clustering procedures are carried out simultaneously. The proposed model effectively exploits external natural patch priors and internal self-similarity priors to improve the denoising performance.

Clean image | Noisy image ($\sigma = 35$) | BM3D[10]: PSNR=25.44dB | WNNM[18]: PSNR= 26.12dB | Ours: PSNR= 26.12dB

Figure 2. Local block matching causes ringing artifacts in human face.

2. External prior guided internal clustering for patch based image denoising

Since image patch space is not a ball like Euclidean space, using the Mahalanobis distance characterized by the patch covariance matrix could be a better choice for patch similarity measure. In order to illustrate it, we uniformly extract 299,000 image patches (size: 8×8) from 5 images (*Lena, House, Cameraman, Monarch, Peppers*). These patches are divided into 3 groups according to their variance v_p: smooth areas ($v_p \leq 0.002$), structural region ($0.002 < v_p \leq 0.02$), and textural area ($v_p > 0.02$). The correct matching rates (%) of Mahalanobis distance (MD) and Euclidean distance (ED) under different noise levels (σ) are listed in Table 1. It can be seen that MD indeed has much better performance than ED in matching noisy patches. By observing the close relationship between covariance matrix and SVD, we find that a collection of patches with similar structure could benefit much the low-rank matrix recovery for image restoration. In this section, we introduce a data clustering method to group similar image patches by a set of Gaussian mixture models (GMM). We will then show how GMM priors can be incorporated in low-rank method for image denoising.

Table 1. Mahalanobis distance (MD) vs. Euclidean distance (ED).

σ_{noise}	Smooth		Structural		Textural	
	MD	ED	MD	ED	MD	ED
$\sigma = 35$	6.2	**8.1**	**36.4**	23.0	**65.6**	57.3
$\sigma = 55$	4.2	**7.2**	**24.5**	13.4	**53.1**	36.5
$\sigma = 75$	3.4	**6.9**	**17.0**	10.0	**43.0**	24.3

2.1. Problem formulation

Before giving the problem formulation, we first analyze the relationship between covariance matrix and SVD. To simplify the notation, we assume that without loss of generality the patch matrix $Y \in \mathbb{R}^{n \times m}$ has zero mean. Let $C = \frac{1}{m} Y Y^T$ be the $n \times n$ covariance matrix. We can use SVD to decompose Y, and rewrite the covariance matrix as $C = \frac{1}{m} U \Sigma V^T V \Sigma U^T = \frac{1}{m} U \Sigma^2 U^T$. Therefore, the eigenvectors of the covariance matrix C are the same as those of matrix U. This implies that a good low-rank approximation of the data matrix Y can be computed using only a small subset of the eigenvectors and eigenvalues. Moreover, the higher the level of correlation or structure similarity among patches in Y, the more energy are concentrated in the first a few eigenvalues. As described in SVD-based low-rank matrix approximation, the small eigenvalues are truncated and therefore grouping the patches with structure similarity to maximize the retained eigenvalues is beneficial for low-rank matrix recovery.

In order to use the full covariance matrix for patch clustering, an intuitive method is to use GMM. As discussed in Section I, we can assume that the latent structures of image patches form K low-dimensional subspaces. Thus, the probability of a given image patch $\mathbf{x}_i$ could be defined as a weighted sum of K Gaussians:

$$p(\mathbf{x}_i|\Theta) = \sum_{k=1}^{K} \omega_k p_k(\mathbf{x}_i|\theta_k), \tag{4}$$

where $\Theta = (\omega_1, \cdots, \omega_K, \theta_1, \cdots, \theta_K)$ is the set of parameters and $\sum_{k=1}^{K} \omega_k = 1$. Note that each θ_k describes a Gaussian density function p_k parameterized by the mean μ_k and covariance matrix Σ_k, and $p_k(\mathbf{x}_i|\theta_k) = c \cdot \exp\big(-\frac{1}{2}(\mathbf{x}_i - \mu_k)^T \Sigma_k^{-1} (\mathbf{x}_i - \mu_k)\big)$. Here, the negative exponent is called the Mahalanobis distance between $\mathbf{x}_i$ and the center μ_k, and c is the normalization constant. Θ can be determined by maximum likelihood, typically using the Expectation-Maximization algorithm [11].

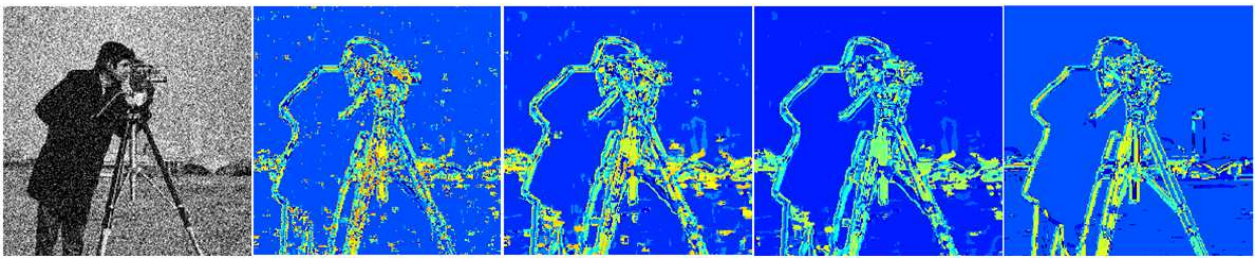
Given an image $\mathbf{x}$, we can extract all overlapped patches into a set, denote by $R\mathbf{x} \doteq (R_1\mathbf{x}, \cdots, R_m\mathbf{x})$, where R_i is a matrix which extracts the i-th patch from $\mathbf{x}$. To simplify the likelihood expression, we introduce class label $C = (c_1, c_2, \cdots, c_m)$ and $c_i \in \{1, \cdots, K\}$ denotes which Gaussian $R_i\mathbf{x}$ is from. Note that $p(R_i\mathbf{x}, c_i|\Theta)$ is the likelihood of the i-th patch belonging to class c_i under the prior Θ. Motivated by EPLL[33], we assume independence among patches and propose the following patch cluster log likelihood term:

$$\begin{aligned} &\log p(R\mathbf{x}, C|\Theta) = \log \prod_{i=1}^{m} p(R_i\mathbf{x}, c_i|\Theta) = \\ &\sum_{i=1}^{m} \log p(R_i\mathbf{x}|c_i) p(c_i) = \sum_{i=1}^{m} \log(\omega_{c_i} p_{c_i}(R_i\mathbf{x}|\theta_{c_i})). \end{aligned} \tag{5}$$

Obviously, our modeling is different from EPLL[33]

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

Figure 3. Local block-matching vs. global searching. (a) A noisy patch. (b) The most similar patches to it by using block matching in a local window. (c) The most similar patches to it by using our global searching. (d) The reconstructed patches by the two sets of similar patches using the same low-rank minimization model in (9).



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

Figure 4. (a) is a noisy image (σ=50). (b)$\sim$(d) show the patch clustering results in iterations 1, 4 and 7. Different colors represent the different clusters. (e) shows the clustering result on the clean image for reference.

which tries to maximize the expected patch log likelihood: $EPLL_p(\mathbf{x}) = \log p(R\mathbf{x}|\Theta)$. If we know the value of C, it gives rise to a partition of all overlapped patches $R\mathbf{x}$ into K classes. Let $\bar{R}_k\mathbf{x} \doteq [R_{k_1}, R_{k_2}, \cdots, R_{k_{d(k)}}]$ denote the matrix formed by the set of vectorized patches from the k-th class, where $d(k)$ is the number of the corresponding patches. It is reasonable to assume that these vectorized patches in each class form a low-rank matrix and we can model the data matrix as $\bar{R}_k\mathbf{x} = Z_k + V_k$, where Z_k and V_k denote the low-rank matrix and the Gaussian noise matrix respectively. Then the rank minimization problem can be approximately solved by minimizing the following functional:

$$E(Z_k) = \frac{1}{\sigma^2}\|\bar{R}_k\mathbf{x} - Z_k\|_F^2 + \tau\|Z_k\|_*. \tag{6}$$

Given a corrupted image $\mathbf{y}$, we want to restore the latent clean image $\mathbf{x}$. With the data fidelity term (1) over the whole image, we propose to solve the following global objective functional for image denoising by combining (5) and the low-rank regularization term (6):

$$(\hat{\mathbf{x}}, \hat{C}, \{\hat{Z}_k\}) = \arg\min_{\mathbf{x}, C, \{Z_k\}} \frac{\lambda}{\sigma^2}\|\mathbf{y} - \mathbf{x}\|_2^2 - \sum_{i=1}^{m} \log p(R_i\mathbf{x}, c_i|\Theta) + \sum_{k=1}^{K}\left(\frac{1}{\sigma^2}\|\bar{R}_k\mathbf{x} - Z_k\|_F^2 + \tau\|Z_k\|_*\right), \tag{7}$$

where λ is a positive constant. Our energy functional effectively incorporates image internal self-similarity and external patch priors into the whole image denoising model. There are two advantages of such clustering based low-ranks regularization.

Firstly, it can globally search the similar patches by patch clustering within the whole image rather than a limited search window. Fig. 3(a) shows a noisy image and indicates a noisy patch; Fig. 3(b) shows the most similar patches to it by using block matching in a local window (i.e., the patch searching method used in WNNM); Fig. 3(c) shows the most similar patches to it by using our global searching. One can see that global searching can find more and better similar patches to the noisy one. In Fig. 3(d), we use the same low-rank minimization model (9) to reconstruct the original patch by using the two sets of patches, respectively. Clearly, global search leads to much better reconstruction.

Secondly, similarity is measured not by Euclidean distance but by Mahalanobis distance based on a learned patch space. Therefore, the proposed model could be very effective to preserve the image edge or corner structure. Fig. 4 visualizes how external GMM prior helps computing internal self-similarity.Fig. 4(a) is a noisy image. Figs. 4(b) $\sim$ 4(d) visualize the patch clustering results in iterations 1, 4 and 7. The similar clusters are represented in similar colors. In Fig. 4(e), we show the clustering result on the clean image as the ground-truth. It can be clearly seen that the

our method effectively exploits the external GMM prior to guide the similar patch grouping across the whole images. The clustering accuracy is also increased with the iteration.

2.2. Optimization and algorithm

We first learn a finite GMM over the pixels of natural image patches. Let $\Theta^{'}$ be the learned parameter of GMM from a training set of clean image patches. Under the white noise assumption with zero mean and standard deviation σ, the Θ can be estimated from $\Theta^{'}$ by a simple eigenvalue-based computation $\Sigma_k = \Sigma_k^{'} + \sigma^2 I, k = 1, \cdots, K$.

When the parameter Θ is known, the proposed model (7) has three kinds of unknowns: the unknown image $\mathbf{x}$, class label C and low-rank matrixes $\{Z_k\}$. We use an alternating minimization procedure to solve the three unknowns. We start with some initial guess $\mathbf{x}^0$. In the ℓ-th iteration, $\ell \in \mathbb{N}$, the solutions to the alternating minimization scheme are detailed as follows.

1. Fix $\mathbf{x}^\ell$, solve for the class label C^ℓ.

   For each patch $R_i\mathbf{x}^\ell$, we use the PDF of the learned GMM to calculate the likelihood within each class,

$$p(k|R_i\mathbf{x}^\ell) = \frac{\omega_k p_k(R_i\mathbf{x}^\ell|\theta_k)}{\sum_{j=1}^K \omega_j p_j(R_i\mathbf{x}^\ell|\theta_j)}, k = 1, \cdots, K, \quad (8)$$

   and find the class of which the Gaussian component generates the maximum likelihood.

2. Fix C^ℓ, solve for the low-rank matrixes $Z_k^\ell, k = 1, \cdots, K$.

$$\hat{Z}_k^\ell = \arg\min_{Z_k} \frac{1}{\sigma^2}\|\bar{R}_k\mathbf{x} - Z_k\|_F^2 + \tau\|Z_k\|_*. \quad (9)$$

   Eq. (9) can be solved by the weighted singular value thresholding algorithm [18]. Let $\sigma_j(Z_k)$ be the j-th singular value of Z_k and $U\Sigma V^T$ be the SVD of $\bar{R}_k\mathbf{x}$. It is shown that the optimal solution to this problem is given by [18]

$$\hat{Z}_k^\ell = U(\Sigma - \tau\,\mathrm{diag}(\alpha))_+ V^T, \quad (10)$$

   where $\alpha_j = 1/(\sigma_j(Z_k) + \epsilon)$ and $(x)_+ = \max\{x, 0\}$.

3. Fix $\{Z_k^\ell\}$, solve for $\mathbf{x}^{\ell+1}$.

   In practice, the patch clustering term $\sum_{i=1}^m \log p(R_i\mathbf{x}, c_i|\Theta)$ can be neglected in this step since the reconstruction of the whole image depends mainly on the low-rank matrixes $\{Z_k^\ell\}$. After solving for each Z_k^ℓ, we can reconstruct the whole image by solving the following minimization problem

$$\hat{\mathbf{x}}^{\ell+1} = \arg\min_{\mathbf{x}} \sum_{k=1}^K \|\bar{R}_k\mathbf{x} - Z_k^\ell\|_F^2 + \lambda\|\mathbf{y} - \mathbf{x}\|_2^2. \quad (11)$$

   This is a quadratic optimization problem, and it can be solved in closed form:

$$\hat{\mathbf{x}}^{\ell+1} = \left(\lambda I + \sum_k \bar{R}_k^T\bar{R}_k\right)^{-1}\left(\lambda\mathbf{y} + \sum_k \bar{R}_k^T Z_k^\ell\right). \quad (12)$$

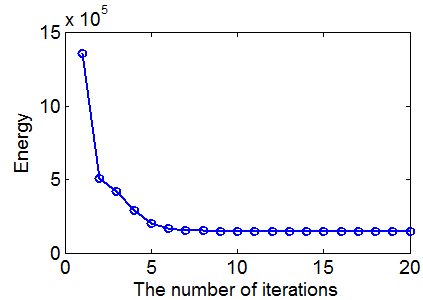


Figure 5. The convergence curve of the proposed algorithm on image *Cameraman*.

After obtaining an improved estimate of the unknown image $\mathbf{x}^{\ell+1}$, the noise standard deviation σ can be updated by utilizing the feedback of filtered noise. The updated σ is then used to improve the estimate $\mathbf{x}^{\ell+2}$. Such a process is iterated until convergence. The complete optimization procedure is summarized as Alg. 1.

Alg. 1 GMM guided patch clustering with low-rank regularization for image denoising

---

1 **Input**: noisy image $\mathbf{y}$, noise standard deviation σ, learned GMM model parameter $\Theta^{'}$ and K.

2 **Output**: denoised image $\mathbf{x}$.

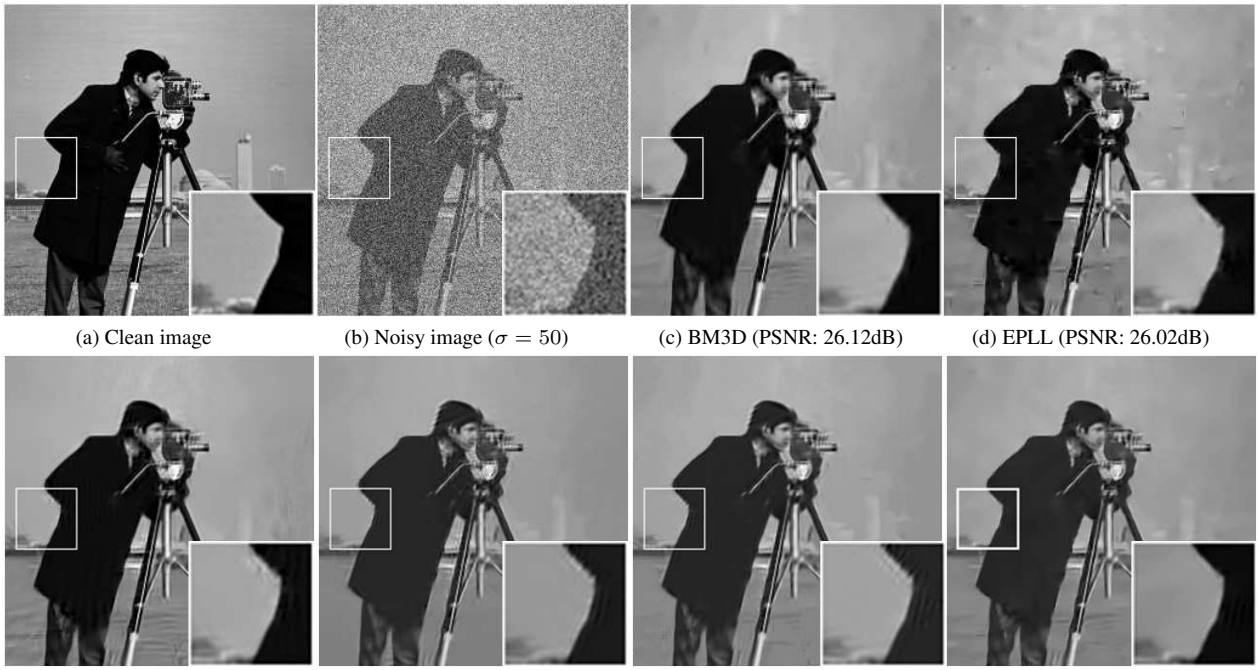
3 **Initialization**:

1) Choose a reasonable scaling factor γ for controlling the re-estimation of noise variance;
2) Initialize $\mathbf{x}^0 = \mathbf{y}, \sigma^0 = \sigma$.

4 **Optimization**:

1) Outer loop: for $\ell = 1 : L$ do
2) Inner loop: for $k = 1 : K$ do
   (i) Update $\Sigma_k = \Sigma_k^{'} + (\sigma^{\ell-1})^2 I$;
   (ii) Calculate conditional probabilities $p(k|R\mathbf{x}^{\ell-1})$ via Eq. (8).
   End for
3) Compute C^ℓ by finding the class which has the highest conditional probability for each patch.
4) Inner loop: for $k = 1 : K$,
   (i) Create $\bar{R}_k\mathbf{x}^{\ell-1}$ by stacking the patches from the k-th class;
   (ii) Compute Z_k^ℓ via Eq. (10).
   End for
5) Compute $\mathbf{x}^\ell$ via Eq. (12);
6) Update σ^ℓ, such that $(\sigma^\ell)^2 = \gamma(\sigma^2 - \|\mathbf{y} - \mathbf{x}^\ell\|_2^2)$.

End for



(a) Clean image (b) Noisy image ($\sigma = 50$) (c) BM3D (PSNR: 26.12dB) (d) EPLL (PSNR: 26.02dB)

(e) LSSC (PSNR: 26.35dB) (f) SAIST (PSNR: 26.15dB) (g) WNNM (PSNR: 26.42dB) (h) Ours (PSNR: 26.55dB)

Figure 6. Denoising results on image *C. Man* by different methods ($\sigma = 50$) .

2.3. Applying a fine classification

As illustrated in Table. 1, patch cluster could not work well on smooth areas, and as a result, a great number of smooth patches (more than 10^4) could be clustered into one class. In order to improve the accuracy of patch cluster, we employ a simple K-means for fine classification by using the mean intensity and location of each patch. If the patch number of some class is too small (less than 10), we merge these patches to the other classes most likely.

2.4. Complexity and convergence

The proposed model in Eq. (7) is nonconvex, and thus the proposed algorithm cannot be guaranteed to converge to a global optimum. However, it is empirically found that our algorithm converges rapidly. Fig. 5 shows an example convergence curve of the proposed algorithm on image *Cameraman* with $\sigma = 50$. One can see that the proposed algorithm converges within 15 iterations.

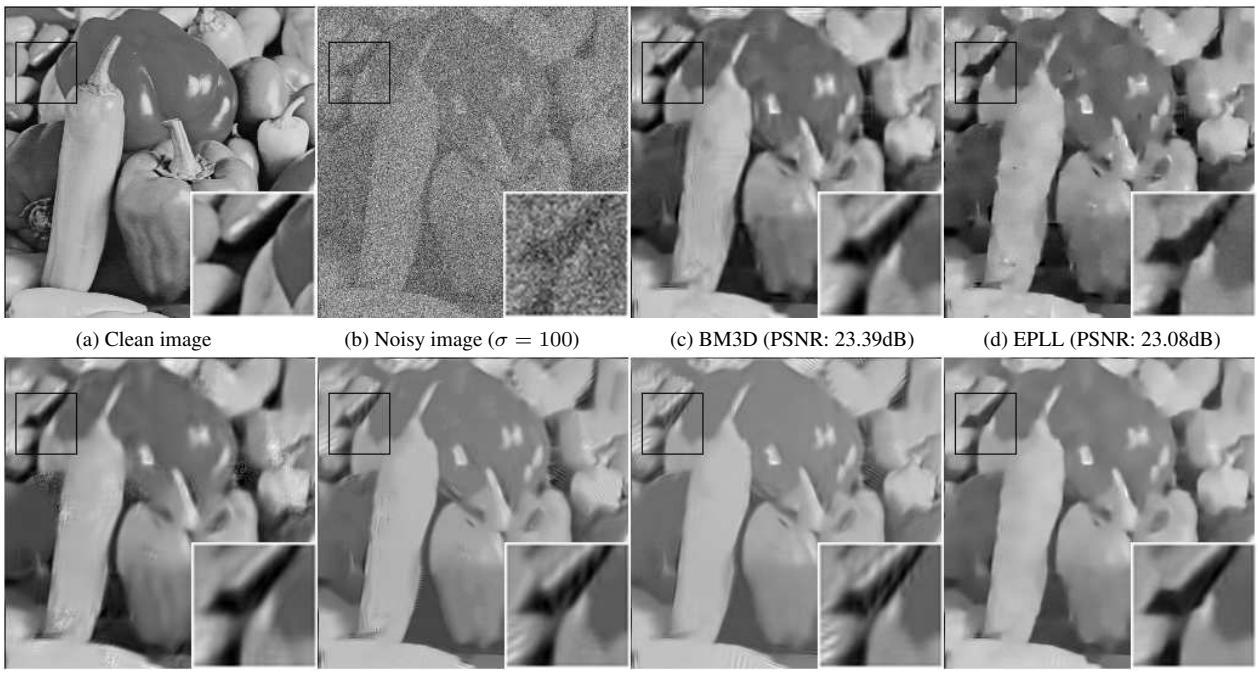
Suppose that we have m patches in k clusters and each patch has n dimensions. The main computational cost in a single iteration of the proposed algorithm consists of three parts. The first part is to evaluate posterior probabilities (8). We need to compute the determinant of the covariance matrix which needs $O(n^3)$ operations. Thus, the first part of patch clustering costs $O(mkn^3)$. The second part is low-rank matrix recovery (10). The SVD for all k clusters costs approximately $O(mkn^2)$. The third part is to average all of the grouped similar patches for each exemplar patch, which costs $O(mn)$. Suppose that our algorithm converges in t iterations, the total computational complexity is $O(tmkn^3)$ without considering fine classification. In our MATLAB implementation, it takes approximately 3 minutes to denoise a 256×256 image with standard deviations $\sigma = 20$ on a laptop with Intel Xeon E3 CPU (3.40 GHz).

3. Experimental results

In this section, we validate the performance of the proposed algorithm and compare it with recently proposed state-of-art denoising methods, including BM3D [10], EPLL[33], LSSC [22], SAIST [13] and WNNM [18]. For all the competing methods the source codes are obtained from their original authors. We used the default parameters in their software packages. As in[33], we learn the GMM models $\Theta^{'}$ with 250 mixture components from a set of 2×10^6 patches, uniformly sampled from the 200 training images from the Berkeley segmentation dataset (BSD) [23]. By experiment experience, we set patch size to 7×7, 8×8, 9×9 and 10×10 for $\sigma \leq 20$, $20 < \sigma \leq 40$, $40 < \sigma \leq 60$ and $\sigma > 60$, respectively. The involved parameters λ and γ in the proposed algorithm are fixed to 0.18 and 0.67. The source code of the proposed method can be downloaded at `http://www4.comp.polyu.edu.hk/~cslzhang/code/PCLR.zip`.

3.1. Results on 12 test images

We first evaluate the proposed algorithm and its competing algorithms on 12 popularly used test images. Gaussian



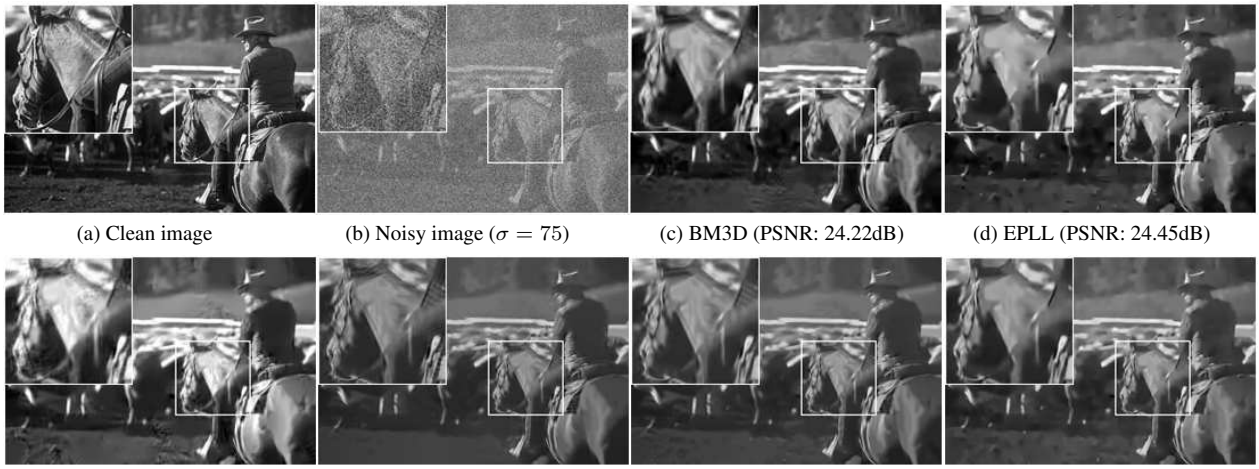
(a) Clean image (b) Noisy image ($\sigma = 100$) (c) BM3D (PSNR: 23.39dB) (d) EPLL (PSNR: 23.08dB)

(e) LSSC (PSNR: 23.20dB) (f) SAIST (PSNR: 23.32dB) (g) WNNM (PSNR: 23.46dB) (h) Ours (PSNR: 23.69dB)

Figure 7. Denoising results on image *Peppers* by different methods ($\sigma = 100$) .

Table 2. Denoising PSNR (dB) results by different denoising methods.

Images	$\sigma = 10$						$\sigma = 30$					
	BM3D	EPLL	LSSC	SAIST	WNNM	Ours	BM3D	EPLL	LSSC	SAIST	WNNM	Ours
C.man	34.18	34.02	34.24	34.30	34.44	**34.46**	28.64	28.36	28.63	28.36	28.80	**28.82**
Monarch	34.12	34.27	34.44	34.76	**35.03**	34.96	28.36	28.35	28.20	28.65	**28.92**	28.83
House	36.71	35.75	**36.95**	36.66	36.94	36.83	32.09	31.23	32.41	32.30	**32.52**	32.17
Parrot	33.57	33.58	33.62	33.66	**33.81**	33.80	28.12	28.07	27.99	28.12	28.33	**28.34**
Peppers	34.68	34.54	34.80	34.82	34.95	**34.96**	29.28	29.16	29.25	29.24	29.49	**29.56**
Montage	37.35	36.49	37.26	37.46	**37.84**	37.48	31.38	30.17	31.10	31.06	**31.65**	31.35
Lena	35.93	35.58	35.83	35.90	**36.03**	35.98	31.26	30.79	31.18	31.27	**31.43**	31.36
Barbara	34.98	33.61	34.98	35.24	**35.51**	35.14	29.81	27.57	29.60	30.14	**30.31**	29.70
Boat	33.92	33.66	34.01	33.91	**34.09**	34.07	29.12	28.89	29.06	28.98	**29.24**	29.21
Couple	34.04	33.85	34.01	33.96	**34.14**	34.12	28.87	28.62	28.77	28.72	**28.98**	28.89
Hill	33.62	33.48	33.66	33.65	**33.79**	33.72	29.16	28.90	29.09	29.06	**29.25**	29.15
Man	33.98	33.97	34.10	34.12	**34.23**	34.21	28.86	28.83	28.87	28.81	**29.00**	**29.00**
Images	$\sigma = 50$						$\sigma = 100$					
	BM3D	EPLL	LSSC	SAIST	WNNM	Ours	BM3D	EPLL	LSSC	SAIST	WNNM	Ours
C.man	26.12	26.02	26.35	26.15	26.42	**26.55**	23.07	22.86	23.15	23.09	23.36	**23.48**
Monarch	25.82	25.78	25.88	26.10	**26.32**	26.25	22.52	22.23	22.24	22.61	**22.95**	22.93
House	29.69	28.76	29.99	30.17	**30.32**	29.78	25.87	25.19	25.71	26.53	**26.68**	25.96
Parrot	25.90	25.84	25.82	25.95	26.09	**26.14**	22.96	22.71	22.79	23.04	23.19	**23.28**
Peppers	26.68	26.63	26.79	26.73	26.91	**27.02**	23.39	23.08	23.20	23.32	23.46	**23.69**
Montage	27.90	27.17	28.10	28.00	**28.27**	28.20	23.89	23.42	23.77	23.98	24.16	**24.24**
Lena	29.05	28.42	28.95	29.01	**29.24**	29.12	25.95	25.30	25.96	25.93	**26.20**	26.16
Barbara	27.23	24.82	27.03	27.51	**27.79**	27.11	23.62	22.14	23.54	24.07	**24.37**	23.66
Boat	26.78	26.65	26.77	26.63	26.97	**26.99**	23.97	23.71	23.87	23.80	24.10	**24.16**
Couple	26.46	26.24	26.35	26.30	**26.65**	26.56	23.51	23.32	23.27	23.21	23.55	**23.65**
Hill	27.19	26.96	27.14	27.04	**27.34**	27.24	24.58	24.43	24.47	24.29	24.75	**24.83**
Man	26.81	26.72	26.72	26.68	**26.94**	**26.94**	24.22	24.07	23.98	24.01	24.36	**24.43**



(a) Clean image (b) Noisy image ($\sigma = 75$) (c) BM3D (PSNR: 24.22dB) (d) EPLL (PSNR: 24.45dB)

(e) LSSC (PSNR: 24.34dB) (f) SAIST (PSNR: 24.38dB) (g) WNNM (PSNR: 24.56dB) (h) Ours (PSNR: 24.69dB)

Figure 8. Denoising results on image *horse* from the BSD dataset by different methods ($\sigma = 75$) .

white noise with standard deviations $\sigma = 10, 30, 50, 100$ are added to those test images. It can be seen that our method has almost the same PSNR results as WNNM on all noise levels, and higher PSNR than other methods. The proposed method effectively exploits the external GMM prior to guide the similar patch grouping, and it can search for the patches in the whole image.This makes the proposed method very robust to reduce strong noise in textural area. In particular, significant improvements can be observed for images *Peppers* and *C. Man* since these images contain strong textures and weak local similarity.

In addition, the proposed method demonstrates higher superiority to other methods in terms of visual quality. The visual comparisons of competing denoising methods at noise levels 50 and 100 are shown in Fig. 6 and Fig. 7, respectively. One can see that the visual quality improvements achieved by the proposed method are more convincing. It can preserve more textures and fine details than the competing methods. Although WNNM[18] has demonstrated a good trade-off between noise removal and edge preservation, it still tends to over-smooth image details and cause ringing artifacts. By more effectively exploit the image external priors and internal priors, the proposed method shows very strong capability to reconstruct the image latent structure from strong noise corruption.

3.2. Results on the BSD dataset

We also comprehensively evaluate the proposed method on 100 test images from the BSD dataset [23]. Table 3 shows qualitative comparisons of competing denoising methods on four noise levels ($\sigma = 15, 35, 55, 75$). One can see that the proposed method can achieve very competitive denoising performance. The visual comparisons of the denoising methods are shown in Fig. 8. More examples of visual comparison can be found in the supplementary file.

Table 3. Denoising results (PSNR) by different denoising methods on 100 test images with additive Gaussian noise.

σ	BM3D	EPLL	LSSC	SAIST	WNNM	Ours
15	30.87	30.99	31.05	31.00	31.15	**31.18**
35	26.91	26.98	26.98	26.96	**27.13**	**27.13**
55	25.12	25.13	25.11	25.19	**25.34**	**25.34**
75	24.07	23.97	23.93	24.06	**24.25**	24.21

4. Conclusion

In this paper, a new denoising approach based on image internal self-similarity prior and external patch priors was presented. Our approach differs from the low-rank based models (e.g., models in [13] and [18]) in two ways. Firstly, we conducted low-rank regularization of similar patches based on global patch clustering but not local block matching. Therefore, we can globally search the similar patches within the whole image. Secondly, we learned a GMM prior model from image patches to guide the patch clustering and the subsequent low-rank subspace learning. Such a clustering based low-rank approximation makes the latent patch reconstruction very robust to noise. Experimental results showed that the proposed algorithm can achieve very competitive denoising performance. In particular, it can preserve much better the image texture structures than the other state-of-the-art denoising algorithms under severe noise environment .

Acknowledgements

This work is supported by NSFC grant (61401098, 61471321), the HK RGC GRF grant (PolyU 5313/13E), and the Open Project Program of the State Key Lab of CAD&CG (Grant No. A1415), Zhejiang University.

References

[1] M. Aharon, M. Elad, and A. Bruckstein. K-svd: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Process.*, 54(11):4311–4322, 2006.

[2] A. Buades, B. Coll, and J. M. Morel. A non-local algorithm for image denoising. *Proc. CVPR*, 2005.

[3] H. Burger, C. Schuler, and S. Harmeling. Image denoising: Can plain neural networks compete with bm3d? *Proc. CVPR*, 2012.

[4] J. Cai, E. Candes, and Z. Shen. A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.*, 20(4):1956–1982, 2010.

[5] E. J. Candes and Y. Plan. Matrix completion with noise. *Proceedings of the IEEE*, 2010.

[6] S. G. Chang, B. Yu, and M. Vetterli. Adaptive wavelet thresholding for image denoising and compression. *IEEE Trans. Image Process.*, 9(9):1532 –1546, 2007.

[7] P. Chatterjee and P. Milanfar. Clustering-based denoising with locally learned dictionaries. *IEEE Trans. Image Process.*, 18(7):1438 – 1451, 2009.

[8] P. Chatterjee and P. Milanfar. Is denoising dead? *IEEE Trans. Image Process.*, 19(4):895–911, 2010.

[9] P. Chatterjee and P. Milanfar. Learning denoising bounds for noisy images. *ICIP*, 2010.

[10] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Trans. Image Process.*, 16(8):2080– 2095, 2007.

[11] A. Dempster, N. Laird, and D. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B*, 1977.

[12] M. Do and M. Vetterli. Wavelet-based texture retrieval using generalized gaussian density and kullbackleibler distance. *IEEE Trans. Image Process.*, 11(2):146–158, 2002.

[13] W. Dong, G. Shi, and X. Li. Nonlocal image restoration with bilateral variance estimation: a low-rank approach. *IEEE Trans. Image Process.*, 22(2):700–711, 2013.

[14] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang. Compressive sensing via nonlocal low-rank regularization. *IEEE Trans. Image Process.*, 23(8):3618–3632, 2014.

[15] W. Dong, L. Zhang, G. Shi, and X. Li. Nonlocally centralized sparse representation for image restoration. *IEEE Trans. Image Process*, 22(4):1620 –1630, 2013.

[16] D. L. Donoho, M.Gavish, and A. Montanari. The phase transition of matrix recovery from gaussian measurements matches the minimax mse of matrix denoising. *PNAS*, 2013.

[17] R. Fergus, B. Singh, A. Hertzmann, S. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *Proc. ACM SIGGRAPH*, 2006.

[18] S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted nuclear norm minimization with application to image denoising. *Proc. CVPR*, 2014.

[19] A. Levin and B. Nadler. Natural image denoising: Optimality and inherent bounds. *CVPR*, 2011.

[20] A. Levin, B. Nadler, F. Durand, and W. T. Freeman. Patch complexity, finite pixel correlations and optimal denoising. *ECCV*, 2012.

[21] J. Mairal, F. Bach, and J. Ponce. Task-driven dictionary learning. *IEEE Trans. Patt. Anal. Mach. Intell.*, 32(4):791–804, 2012.

[22] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. *Proc. ICCV*, 2009.

[23] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. *Proc.ICCV*, 2001.

[24] I. Mosseri, M. Zontak, and M. Irani. Combining the power of internal and external denoising. *Proc.ICCP*, 2013.

[25] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli. Image denoising using scale mixtures of gaussians in the wavelet domain. *IEEE Trans. Image Process.*, 12(11):1338 – 1351, 2003.

[26] A. Rajwade, A. Rangarajan, and A. Banerjee. Image denoising using the higher order singular value decomposition. *IEEE Trans. Patt. Anal. Mach. Intell.*, 35(4):849–862, 2013.

[27] R. Rubinstein, A. Bruckstein, and M. Elad. Dictionaries for sparse representation modeling. *Proc. Special Issue on Applications of Compressive Sensing and Sparse Representation*, 98(6):1045– 1057, 2010.

[28] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1):259–268, 1992.

[29] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. *Proc. ICCV*, 1998.

[30] G. Yu, G. Sapiro, and S. Mallat. Solving inverse problems with piecewise linear estimators: from gaussian mixture models to structured sparsity. *IEEE Trans. Image Process.*, 21(5):2481–2499, 2012.

[31] M. Zontak and M. Irani. Internal statistics of a single natural image. *Proc. CVPR*, 2011.

[32] M. Zontak, I. Mosseri, and M. Irani. Separating signal from noise using patch recurrence across scales. *Proc. CVPR*, 2013.

[33] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. *ICCV*, 2011.

[34] W. Zuo, L. Zhang, C. Song, and D. Zhang. Texture enhanced image denoising via gradient histogram preservation. *Proc. CVPR*, 2013.