Machine Vision and Applications (2021) 32:110
https://doi.org/10.1007/s00138-021-01232-3

ORIGINAL PAPER

Sparse representation with enhanced nonlocal self-similarity for image denoising

Tao Zhou[1] · Chen Li[1] · Xuan Zeng[2] · Yuhang Zhao[1]

Received: 8 October 2020 / Revised: 7 June 2021 / Accepted: 16 July 2021 / Published online: 21 August 2021

Abstract
In the past decade, the sparsity prior of image is investigated and utilized widely as the development of compressed sensing theory. The dictionary learning combined with the convex optimization methods promotes the sparse representation to be one of the state-of-the-art techniques in image processing, such as denoising, super-resolution, deblurring, and inpainting. Empirically, the sparser of image representation, the better of image restoration. In this work, the non-local clustering sparse representation is applied with optimized matching strategies of self-similar patches, which break through the bottleneck of search window (localization) and improve the estimation effect of the sparse coefficient. The experimental results show that the proposed method provides an effective suppression on noise, preserves more details of image and presents more comfortable visual experience.

Keywords Image Denoising · Global Matching · Nonlocal Self-Similarity · Sparse Representation

1 Introduction

Vision is taken as the most advanced source of information for human beings, and images play the most important role in human vision. Noise corruption is inevitable during the sensing process, and it may heavily degrade our visual experience. Removing noise is an essential step in various image processing and vision tasks, such as image segmentation, image coding, and target detection. In particular, many image restoration problems could be addressed by solving a sub-problems of denoising, which further extends the meaning of exploring image denoising techniques.

Many classical denoising methods have emerged in recent years [1–8], such as wavelet, dictionary leaning and multi-scale feature fusion. The methods could be roughly divided into two categories: spatial domain filtering and transform domain filtering. The former directly processes the pixels of the image and the representative one is the non-local means (NLM) [1]. NLM is wonderful algorithm because of its simplicity and effectiveness, which creatively utilized the non-local self-similarity (NSS) of image for denoising task. The core of NLM is to group the similar patches and perform an average operation for noise suppression. Natural images often have NSS prior, as shown in Fig. 1, the contours are not randomly distributed and have a clear correlation/similarity at different positions. The NSS prior is widely used in many image processing methods [2–6], such as the benchmark denoising algorithm-block matching three-dimensional filtering (BM3D) [3].

The other category mainly uses a basis function (atom) to transform the image to another domain, where separates the noise and effective information of image. The representative method is sparse representation. For natural images, the meaningful information usually possesses sparsity in a transform domain (e.g. wavelet domain), where the sparse signal is enhanced and noise remains the same, by using a simple thresholding method, the noise component could be removed and the useful information is retained, and finally the inverse transform is performed to recover the image in spatial domain. There are two typical solutions, one is based on the analytic basis [9] and the other is the dictionary learning. With the flexibility of atom design, the redundant and over-complete dictionary representation could achieve higher sparsity when compared with the analytic

✉ Chen Li
lichen@icrd.com.cn

[1] Shanghai Integrated Circuits R&D Center Co., Ltd, Shanghai 201203, China

[2] State Key Lab of ASIC & System, School of Microelectronics, Fudan University, Shanghai 201203, China

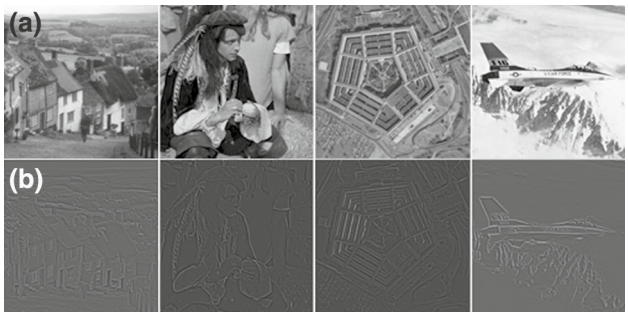


Fig. 1 Examples of NSS features extracted from images. **a** some natural images from the standard test dataset; **b** the contours according to **a** (extraction with second atom by using PCA method)

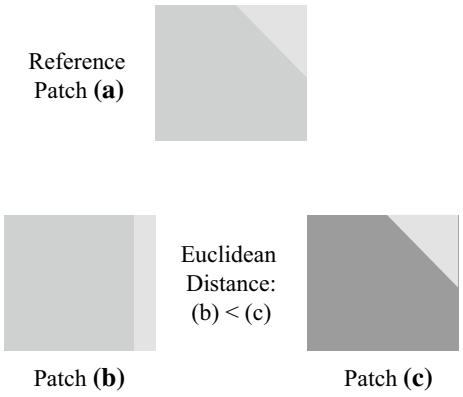


Fig. 2 Illustration of mismatching with common-used Euclidean distance. **a** The reference patch; **b** The patch with the same intensity and different structure; **c** The patch with different intensity and the same structure

basis methods, and indeed obtained much better denoising performance, such as K-means Singular Value Decomposition (K-SVD) [10], Learned Simultaneous Sparse Coding (LSSC) [5], Expected Patch Log Likelihood (EPLL) [11], Nonlocally Centralized Sparse Representation (NCSR) [12]; furthermore, the visual data often has an intrinsic low-rank structure [13], the low-rankness of patch matrix could be viewed as a 2D sparsity prior as compared with the dictionary sparse representation, and based on the low rank matrix completion theory [14], some impressive restoration methods have been proposed [15–24]. The key point of sparse method is to improve the sparsity of representation, and the NSS feature is one of the most important and widely used prior for sparsity improvement, especially in the case of image corrupted by noise. Several very competitive algorithms reported [12, 16, 18], taking full advantage of the NSS prior of natural images, have proven the effectiveness of this prior.

In this work, in order to better employ the NSS prior of image and further improve the denoising performance, the centralized sparse representation (CSR) [25]-based algorithm is proposed with enhanced NSS from 3 aspects: (1) the similar patches from external images are combined with the internal noise patches to refine the dictionary; (2) a global matching strategy is adopted to facilitate the estimation of model parameters; (3) an improved similarity measurement with cosine distance is used to eliminate the influence of the dimensional difference.

Among the sparse representation methods, the benchmarks [12, 25] introduced the sparse coding noise (SCN) in the object function and minimized SCN along with the fidelity term to obtain a superior performance. However, the sub-dictionary is learned from the centralized patches from noisy image itself and neglects the external priors, and the estimation of latent clean image is performed in a local window, which limits the quality of patch grouping with NSS prior. By introducing external reference images and global matching strategies, the clustering is improved to make better dictionary learning and more accurate estimation of latent clean patches, and then, the improvement of denoising performance is expected. In particular, using Euclidean distance (ED) is easy to cause mismatch by dimensional error, as shown in Fig. 2, and we adopt the cosine distance instead of ED to remove the dimensional difference to achieve a better matching effect.

The structure of the paper is constructed as follows: the second part introduces the proposed method, and third part illustrates the operations of NSS enhancement as well as the principle of denoising method; the fourth part demonstrates the denoising results and the comparison with several state-of-the-art algorithms based on two kinds of datasets, one is the standard image processing pictures with added Gaussian white noise (AGWN); the other is the actual low-light image with real noise; the last part summarizes the content of the paper, points out the places need to be improved in the future.

2 The proposed method

For sparse representation, the image $x \in R^N$, and its sparse coefficients $x \approx \Phi\alpha_x$, $\alpha_x \in R^M$, where $M << N$ and most entries of α_x are 0 by dictionary coding (dictionary $\Phi \in R^{MxN}$), the key point of the method is sparse decomposition with dictionary Φ, and generally it could be obtained by solving the l_0-norm minimization problem [26–28], $\alpha_x = \text{argmin}\|\alpha\|_0$, s.t. $\|x-\Phi\alpha_x\|^2 \le \varepsilon$; however, the l_0 norm minimization is a NP-hard problem, and often replaced by its closest convex relaxation-l_1 norm minimization:

$$\alpha_x = \arg\min_{\alpha}\{\|x - \Phi\alpha\|_2^2 + \lambda\|\alpha\|_1\} \quad (1)$$

where the constant λ denotes the regularization parameter, which balance the fidelity term (approximation error) and the sparse prior. And Eq. (1) could be minimized efficiently

by an analytic method such as iterative shrinkage algorithm [29]. In particular, coding with dictionaries learned from natural images could get a better performance than with fixed basis [30].

For image restoration, the degraded image signal y could be generally written as $y = Hx + n$, where H is the degradation operator and for denoising H is specified as identical matrix, x is the latent clean image and n is the additive noise, and generally set as Gaussian distribution $N(0, \sigma_n^2)$. Due to intensity difference in sparse domain between noise and image signal, denoising could be realized by a simple thresholding method, which removes the noise components and retains the structure information [31]. To recover the latent clean image x from its noisy version y with respect to the dictionary Φ, the sparse model could be written in a patched way as below:

$$\alpha_y = \arg\min_{\alpha}\{\|y - \Phi\alpha\|_2^2 + \lambda \sum_i \|\alpha_i\|_1\} \tag{2}$$

However, simply using the sparse coding for the degraded image to recover the latent clean one is a very challenging task, for ill-posed nature of denoising for high-dimensional nature of natural images. On the other hand, it is known that the sparse coefficients of natural images are not randomly distributed, as shown in Fig. 1, and they have a NSS feature obviously. The strong correlations could be used to develop a more effective sparse model by exploring the nonlocal redundancies. Indeed, several classic methods are derived from this point; among them, the most representative one is NSCR, which introduced a sparse coding noise (SCN) to incorporate the NSS prior to help improve the image restoration performance. The objective function of NCSR [12] is:

$$\alpha_y = \arg\min_{\alpha}\{\|y - \Phi\alpha\|_2^2 + \lambda \sum_i \|\alpha_i - \beta_i\|_1\} \tag{3}$$

where parameters Φ and β paly the most important role for performance improvement. Φ is indeed a sub-dictionary which is learned from each cluster grouped in a global area and β is the estimated latent clean patch used to form the SCN (α-β), obtained by the corresponding sub-dictionary coding on an averaged patch with local matching method. The regularization parameter λ could be derived from the interpretation of the model in the Maximum a Posterior way [32].

The NCSR model achieved the state-of-the-art performance at that time. However, new requirements raised later have surpassed its ability, such as the heavy noise condition where the noise components are dominating in the images, it is very difficult to extract effective characteristics if only based on the original one. There is still room for improvement of the estimated β by sub-dictionary coding on a specified global searching instead of a local window. In order to improve the accuracy of sparse coding and cover more noise levels, we propose an enhanced NSS improvement for sparse representation embed in the NCSR basic framework to pursue a better denoising performance.

3 External and global matching strategy

3.1 The effectiveness and feasibility of external images

The prior information could be extracted from different sources. In general, learning a dictionary from the natural image library will be time-consuming, however, in many occasions, a large number of similar images could be obtained and simplify the selection of external priors. With the help of the correlated images, the extracted sparse prior is much more accurate than that learned from the noisy image only. For example, in the surveillance situations, the image taken at night corrupted by a strong noise due to its low light environment, and in the same environment, images taken during the day have less noise while the major components remain the same which could be a worth reference and provide effective guidance for denoising at night.

3.2 Similarity measurement

For image processing, patch operation is widely adopted for consideration of computation efficiency. Therefore, the patch matrix M of image is constructed with each column as a patch vector. To cluster the patches, the K-means method is adopted for similarity calculation, the cosine distance is used to replace the ED since ED ignores the correlation interference between data/patch (as shown in Fig. 2), which may lower the matching accuracy and in turn affects quality of sparse representation.

The cosine distance could be viewed as a normalized ED; it is insensitive to the absolute values of patches and more closed to the humane perception under a neuroscience perspective [33]. Based on the concept of cosine distance, image block similarity measurement is proposed:

$$d_{r,i} = 1 - \cos(y_i, y_r) \tag{4}$$

where r represents the reference image patch, i represents the test patch, $d_{r,i}$ is the 'distance' between the two patches y_r, y_i, function $\cos(\cdot)$ is the cosine similarity of y_r, y_i as two vectors. One patch is classified into a certain class according to the size of $d_{r,i}$.

3.3 Global clustering for dictionary learning

For the noisy image, external priors are added to improve the dictionary learning of corresponding features, aiming to improve the denoising performance, steps include:

(1) Image retrieval: selecting the similar external images from some ready-made datasets or on the web. For the tested dataset set12, the external image could be selected easily by using a search engine, such as TinEye or Google Images. Also, one can use a much more sophisticated method for image searching, such as [34]. For the real-noise dataset, the external similar images are captured with different conditions such as light intensity, exposure time and different view-angles.

(2) Patch matrix formation: decomposing the selected images into patches with specified size, merging the patches sequentially to form a matrix M (each column represents a patch vector);

(3) Feature extraction:

(a) Flat feature extraction, for a specific patch $y_i = M(:, i)$ in M, calculating its variance:

$$\sigma_i = \frac{1}{n}\sum_{p=1}^{b}\sum_{q=1}^{b}(y_i(p,q) - \tilde{y}_i(p,q))^2 \tag{5}$$

where n is the pixel number in a patch, p and q represent the pixel locations, b is the patch size (square shape). $\tilde{y}_i$ is the Gaussian filtered image of patch y_i. Then a threshold method is used to determine whether a patch belongs to the flat region. As for natural images, the flat area generally accounts for a large proportion and the flat structure are simple, a small local window is sufficient for its extraction as well as atom learning.

(b) Detail feature extraction: for the detail patches left in M, the K-means method is used for grouping. The similarity measurement adopts the method described in Sect. 2.2. In consideration of the instability of the global dictionary learning method [10], a sub-dictionary learning followed the Ref.12 is adopted based on PCA method [35]. The number of sub-dictionary atom is consistent with the final optimized clustering centers (assuming K). Given the flat region atom, the total atom number is K + 1, and the final dictionary is represented as:

$$D = \{\varphi_0, \varphi_1, \varphi_3, ..., \varphi_K\} \tag{6}$$

3.4 Global matching for sparse coefficient estimation

The method of searching similar patches in local window is widely used because it is simple and effective. However, for natural images, similar structures may be distributed at different locations across the whole image as well as similar external images. Local windows exclude a large number of high-quality structures outside the window, as shown in Fig. 3. Therefore, we propose a matching method by searching the similar patches in a global area (patch matrix M) instead of a local window.

For the i-th patch y_i of target image, first finding the class S_k it belongs to, for all the image blocks y_j in the class S_k, calculating the distance $d_{i,j}$ by using (4); then choosing the first t minimum distance patches $y_{i,j}$ as the similar patches of y_i. The estimated patch of y_i is calculated as:

$$e_i = \frac{1}{\sum_{j=1}^{t} w_{i,j}}\sum_{j=1}^{t} w_{i,j}y_{i,j} \tag{7}$$

where the averaging weights of $w_{i,j}$ is determined by:

$$w_{i,j} = \exp(-d_{i,j}^2/h^2) \tag{8}$$

Finally, the key parameter of sparse coefficient is calculated by the learned sub-dictionary D (φ_k):

$$\beta_i = \varphi_k^T e_i \tag{9}$$

Additionally, the regularization parameter λ could be derived from applying the maximum posterior probability MAP [28] for NCSR model representation:

$$\lambda_{i,j} = \frac{2\sqrt{2}\sigma_n^2}{\sigma_{i,j}} \tag{10}$$

where σ_n^2 is the noise level of original image, and $\sigma_{i,j}$ is the noise level of jth pixel in ith patch, and the object function is:

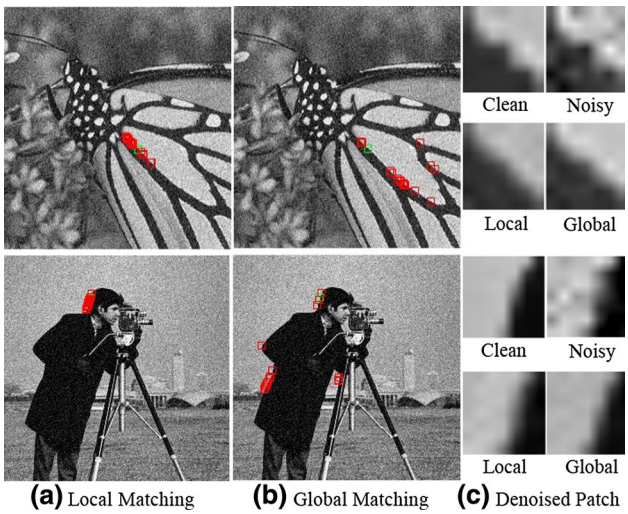


Fig. 3 Comparison of finding similar patches with a local and global matching: **a** The similar patches with local matching; **b** The similar patches with global matching (the reference patch is indicated by green square, and the similar patches indicated by red square); **c** comparison of the averaging effect for local and global matching

$$\alpha_y = \arg\min_{\alpha}\{\|y-\Phi\alpha\|_2^2 + \sum_i\sum_j \lambda_{i,j}\|\alpha_i(j)-\beta_i(j)\|_1\} \quad (11)$$

The global matching strategy offers several distinct advantages: (1) the external prior is used to enhance the NSS feature for sub-dictionary learning and latent clean patch estimation; (2) making the sparse coding of patch consistent with the sub-dictionary learning process in a global nature; (3) the denoising effect has been improved remarkably with different noise levels.

The object function (11) could be solved by an iterative shrinkage method efficiently [29], and the algorithm of this work combined with a global match strategy is illustrated in Algorithm 1.

Algorithm 1: Image denosing with global matching

Input: the noisy image y, residual factor δ, the noise-level σ;

Output: the denoised image $\hat{y}$;

Initialization:

1. set the initial estimated image $x^{(0)} = y$;
2. incorporate the correlated external images $\{z_r\}_{r=1}^{R}$ with a crop and alignment operation;
3. set the residual factor δ and the noise-level σ;

Outer loop (clustering and dictionary learning): Iteration on t = 1, 2, … , T

4. Clustering image patches with global matching with K-means with image $x^{(t-1, S)}$;
5. Update the sub-dictionary $\{\varphi_k\}$ with PCA;
6. Inner loop(latent clean patch estimation and iterative shrinkage): Iteration on s = 1, 2, …, S
   (a) Update the image with residual information to enhance details $x^{(t, s-1/2)} = x^{(t, s-1)} + \delta(y - x^{(t, s-1)})$;
   (b) Estimated the latent clean patch e_i with patch averaging (7) based on $x^{(t, s-1/2)}$;
   (c) Compute the sparse coefficients β_i with corresponding sub-dictionary φ_k according to (9);
   (d) Update the regularization parameter $\lambda_{i,j}$ with (10);
   (e) Compute the denoised sparse coefficients α_i with shrinkage algorithms and $x_i = \varphi_k\alpha_i$;
   (f) Update the image $x^{(t, s)}$ with a least-square solution [36] based on patches x_i;
7. Output $\hat{y} = x^{(T, S)}$;

4 Experimental results and discussions

4.1 Evaluation indicators and parameters setting

In order to verify the effectiveness of NSS enhancement, the proposed method will be tested on the standard test image dataset (set12) and the real-noise photograph dataset, respectively. The results are compared with the classic denoising algorithms K-SVD, EPLL, BM3D, NCSR and WNNM. Peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) are used for evaluation of image denoising performance.

The parameters in K-SVD, EPLL, BM3D, NCSR and WNNM are used as default. Except that there are several parameters to be illustrated for patch matching in NCSR and proposed method. Experimentally, the patch size is set to 6×6, 7×7, 8×8 and 9×9 for $\sigma_n \le 20$, $20 < \sigma_n \le 40$, $40 < \sigma_n \le 70$ and $70 < \sigma_n$, respectively as well as T is set to 2, 3, 3, and 4 according to the noise levels. Additionally, the thresholding value for flat classification is proportional to the total noise level and the default scale coefficient is 1. When used for real photograph denoising, the noise level is evaluated by the method [37].

4.2 Test on standard dataset (set12)

Test is performed on 12 standard test images (as shown in Fig. 4) with additive white Gaussian noise (AWGN) for performance testing. The AWGN is set with a mean value of 0 and the standard deviation σ_n is 10, 30, 50, 70, and 100, respectively.

Table 1 and Table 2 show the PSNR and SSIM results for various sparse representation algorithms, respectively, and the bold values represent the best performance in the comparison. It could be seen that the proposed method has achieved a better performance in almost all noise levels, even in the case of strong noise ($\sigma_n >= 50$). The remarkable improvements under strong noise ($\sigma_n >= 50$) prove the effectiveness of the NSS enhancement, since that the external images with similar details/structures merged have offered extra useful information and improved the quality of image prior, while the other methods that only rely on the single noisy image suffer from heavy degradation as the noise level increases and could not be recovered independently and effectively.

For subjective evaluation, three standard images (Boat, Monarch, Peppers) under strong noise (σ_n=100) were selected for display (shown in Figure 5). The details of the same area of each figure are enlarged (marked by a green square). Fig. 5 indicates that some details in the original image are almost indistinguishable due to severe damage of strong noise. By comparison, it is found that the K-SVD is overall blurred after denoising, image details are lost. As for EPLL and BM3D, some visual artifacts are suppressed, but the water ripple effect appears in the smooth area of the image, mainly caused by the patch processing of the image, while the NCSR and WNNM have an over-smooth phenomenon and some texture information is lost. Local window search limits the quantity and quality for patch matching, resulting in obvious errors in the reconstructed structures.

Fig. 4 The standard test image dataset

As for the proposed method with global matching, more details of the original image are reconstructed, while the flat area is visually pleasant.

4.3 Test on real photograph dataset

To further verify the effectiveness of the proposed algorithm, we applied the algorithm for real-noise image taken at low-light condition in a laboratory environment. The dataset contains 8 scenes, 5 levels of light intensity below 3×10^{-3} lx and 3 exposure times have been chosen for each scene, and 5 photos are taken under every condition. The total number of our self-made image dataset is $8\times5\times3\times5=600$. Several low-light typical scenes are shown in Fig. 6.

The overall image is dark, with low signal-to-noise ratio and dead pixels. Usually, such image needs to be pre-processed before denoising, which mainly includes defect corrections and luminance enhancement. The dataset contains images taken

Table 1 PSNR (dB) results by different denoising methods

σ_n		Barbara	Boat	Camera	Couple	Finger	Hill	House	Lena	Man	Monar	Pepper	Straw	Average
10	K-SVD	34.766	33.963	33.751	33.895	31.456	34.194	35.963	35.662	33.730	33.664	34.237	30.969	33.854
	EPLL	34.310	34.166	34.023	34.115	30.919	34.323	35.747	35.726	33.969	34.275	34.538	30.825	33.911
	BM3D	35.305	34.265	34.188	34.394	31.451	34.499	36.714	36.237	34.045	34.124	34.684	30.917	34.235
	NCSR	35.422	34.284	34.186	34.364	31.634	34.484	36.726	36.218	34.172	34.400	34.676	31.405	34.331
	WNNM	35.874	34.455	34.353	34.515	31.885	34.604	36.850	36.532	34.282	35.004	34.940	31.690	34.582
	Proposed	**36.516**	**35.633**	**35.771**	**35.802**	**33.836**	**35.423**	**37.550**	**37.459**	**35.287**	**36.171**	**35.915**	**33.496**	**35.738**
30	K-SVD	29.030	28.065	28.042	27.646	25.153	28.382	31.187	29.717	28.106	27.874	28.772	24.642	28.051
	EPLL	29.138	28.521	28.358	28.328	24.790	29.014	31.228	30.007	28.562	28.351	29.164	24.747	28.351
	BM3D	29.834	28.598	28.638	28.471	25.655	29.261	32.087	30.454	28.730	28.364	29.280	24.941	28.693
	NCSR	29.757	28.334	28.534	28.243	25.611	29.019	32.029	30.341	28.609	28.305	29.135	25.150	28.589
	WNNM	30.036	28.710	28.810	28.511	25.775	29.304	32.537	30.668	28.783	28.923	29.515	25.475	28.921
	Proposed	**30.104**	**29.023**	**29.936**	**28.854**	**26.654**	**29.348**	**32.779**	**30.910**	**28.983**	**29.630**	**30.162**	**25.986**	**29.364**
50	K-SVD	26.121	25.366	25.712	25.091	22.166	25.870	27.996	26.703	25.592	25.302	26.101	21.536	25.297
	EPLL	26.673	26.088	26.025	25.879	22.133	26.713	28.765	27.415	26.309	25.776	26.626	22.003	25.867
	BM3D	27.260	26.091	26.131	25.973	23.415	27.017	29.694	27.884	26.406	25.819	26.683	22.406	26.231
	NCSR	27.212	25.835	26.105	25.798	23.050	26.766	29.672	27.918	26.296	25.564	26.549	22.477	26.103
	WNNM	27.459	26.321	26.476	26.048	23.396	27.076	30.368	28.148	26.492	26.333	26.965	22.918	26.500
	Proposed	**27.522**	**26.450**	**27.243**	**26.275**	**23.745**	**27.151**	**30.499**	**28.399**	**26.652**	**26.806**	**27.380**	**23.014**	**26.761**
70	K-SVD	24.098	23.689	23.863	23.553	19.547	24.606	25.440	24.714	24.070	23.308	24.053	19.758	23.392
	EPLL	25.038	24.565	24.514	24.452	20.223	25.312	27.038	25.745	24.873	24.071	24.899	20.357	24.257
	BM3D	25.565	24.644	24.620	24.580	22.080	**25.610**	27.912	26.332	25.059	24.242	25.066	21.017	24.727
	NCSR	25.364	24.351	24.520	24.321	21.540	25.287	27.675	26.192	24.805	23.789	24.723	20.884	24.454
	WNNM	**25.713**	**24.871**	24.954	24.622	22.052	25.604	**28.694**	**26.634**	25.056	24.671	25.371	**21.479**	24.977
	Proposed	25.559	24.836	**25.277**	**24.678**	**22.138**	25.589	28.504	26.476	**25.119**	**24.843**	**25.472**	21.436	**24.994**
100	K-SVD	22.390	22.292	21.582	22.335	17.800	23.323	23.612	22.986	22.665	20.688	21.967	18.674	21.693
	EPLL	23.423	23.124	22.856	23.036	18.252	23.948	25.190	24.077	23.473	22.230	23.080	19.018	22.642
	BM3D	23.686	23.298	23.081	23.168	**20.657**	**24.160**	25.872	24.587	23.618	22.518	23.395	19.589	23.136
	NCSR	23.451	22.904	22.773	22.837	20.018	23.828	25.520	24.340	23.292	21.768	22.806	19.321	22.738
	WNNM	**23.878**	23.333	23.435	**23.276**	20.637	24.049	**26.705**	**25.046**	23.588	**23.006**	23.526	19.824	**23.359**
	Proposed	23.764	**23.344**	**23.552**	23.180	20.554	24.113	26.179	24.535	**23.694**	22.759	**23.599**	**20.082**	23.279

Bold number represents the highest value achieved by a certain algorithm at each noise level

Table 2 SSIM results by different denoising methods

σ_n		Barbara	Boat	Camera	Couple	Finger	Hill	House	Lena	Man	Monar	Pepper	Straw	Average
10	K-SVD	0.936	0.923	0.926	0.925	0.969	0.914	0.906	0.946	0.918	0.949	0.923	0.958	0.933
	EPLL	0.931	0.931	0.935	0.932	0.967	0.920	0.903	0.948	0.924	0.956	0.927	0.958	0.936
	BM3D	0.945	0.930	0.932	0.936	0.970	0.920	0.922	0.954	0.923	0.956	0.928	0.959	0.939
	NCSR	0.947	0.930	0.932	0.935	0.970	0.919	0.923	0.954	0.925	0.957	0.927	0.963	0.940
	WNNM	0.951	0.930	0.931	0.936	0.971	0.920	0.920	0.956	0.925	0.960	0.929	0.964	0.941
	Proposed	**0.953**	**0.941**	**0.943**	**0.946**	**0.981**	**0.930**	**0.930**	**0.959**	**0.935**	**0.964**	**0.934**	**0.977**	**0.949**
30	K-SVD	0.828	0.777	0.815	0.759	0.851	0.746	0.830	0.853	0.771	0.867	0.837	0.804	0.812
	EPLL	0.838	0.803	0.832	0.800	0.861	0.787	0.834	0.863	0.797	0.879	0.847	0.823	0.830
	BM3D	0.855	0.807	0.837	0.808	0.879	0.797	0.848	0.877	0.803	0.882	0.850	0.829	0.839
	NCSR	0.856	0.795	0.839	0.800	0.872	0.786	0.848	0.879	0.799	0.884	0.850	0.836	0.837
	WNNM	**0.859**	0.808	0.840	0.809	0.883	0.796	0.852	**0.884**	0.804	0.893	0.856	0.853	0.845
	Proposed	**0.859**	**0.813**	**0.862**	**0.814**	**0.904**	**0.799**	**0.860**	0.881	**0.808**	**0.897**	**0.859**	**0.870**	**0.852**
50	K-SVD	0.742	0.674	0.748	0.642	0.696	0.637	0.763	0.768	0.672	0.796	0.771	0.580	0.707
	EPLL	0.767	0.712	0.762	0.695	0.736	0.693	0.784	0.791	0.715	0.812	0.783	0.649	0.742
	BM3D	0.790	0.716	0.783	0.706	0.793	0.712	0.812	0.817	0.722	0.820	0.794	0.688	0.763
	NCSR	0.794	0.705	0.782	0.698	0.756	0.696	0.817	0.825	0.717	0.822	0.798	0.691	0.758
	WNNM	0.798	0.727	0.784	0.709	0.797	0.714	0.822	0.828	0.728	0.834	0.801	0.734	0.773
	Proposed	**0.802**	**0.731**	**0.806**	**0.724**	**0.819**	**0.718**	**0.829**	**0.836**	**0.733**	**0.849**	**0.812**	**0.745**	**0.784**
70	K-SVD	0.664	0.597	0.680	0.563	0.478	0.580	0.685	0.695	0.603	0.728	0.705	0.396	0.614
	EPLL	0.707	0.642	0.708	0.620	0.585	0.626	0.737	0.732	0.650	0.753	0.731	0.487	0.665
	BM3D	0.735	0.654	0.743	0.637	0.721	0.649	0.775	0.768	0.665	0.767	0.748	0.574	0.703
	NCSR	0.739	0.644	0.746	0.624	0.666	0.631	0.783	0.779	0.658	0.768	0.752	0.557	0.696
	WNNM	**0.742**	0.664	0.740	0.642	0.731	0.650	0.792	**0.783**	0.671	0.783	0.752	0.640	0.716
	Proposed	0.739	**0.666**	**0.768**	**0.647**	**0.745**	**0.651**	**0.799**	0.782	**0.672**	**0.795**	0.763	**0.644**	**0.723**
100	K-SVD	0.586	0.525	0.575	0.497	0.275	0.524	0.607	0.615	0.536	0.619	0.622	0.275	0.521
	EPLL	0.636	0.562	0.635	0.539	0.351	0.554	0.670	0.658	0.578	0.677	0.665	0.330	0.571
	BM3D	0.662	0.591	0.693	0.563	0.626	0.580	0.720	0.702	0.600	0.702	0.688	0.422	0.629
	NCSR	0.668	0.579	0.697	0.552	0.550	0.565	0.734	0.718	0.594	0.698	0.697	0.373	0.619
	WNNM	0.671	0.587	0.688	0.565	0.635	0.577	0.745	**0.722**	0.604	0.720	0.694	0.460	0.639
	Proposed	**0.673**	**0.599**	**0.717**	**0.570**	**0.652**	**0.582**	**0.752**	0.719	**0.612**	**0.723**	**0.708**	**0.529**	**0.653**

Bold number represents the highest value achieved by a certain algorithm at each noise level

under different ambient light levels and under the same shooting parameters, generally, the lower the brightness, the lower the signal-to-noise ratio and stronger noise intensity after digital amplifier. Since latent truth of images is extremely difficult to obtain, we only make a subjective comparison, as shown in Figure 7.

The comparison is similar to that on the standard dataset. Overall, the EPLL and BM3D results still have slight artifacts in the flat region and edge structure; the NCSR and WNNM have a smooth trend. As for the protection of details and textures such as characters and edge lines, the EPLL, BM3D and NCSR are not as good as the proposed algorithm.

5 Conclusions

In this paper, the limitations of reported sparse representation for image denoising are studied. Based on the non-local clustering algorithm framework, we proposed a NSS enhanced denoising algorithm. By combining external reference images and global matching strategies, the

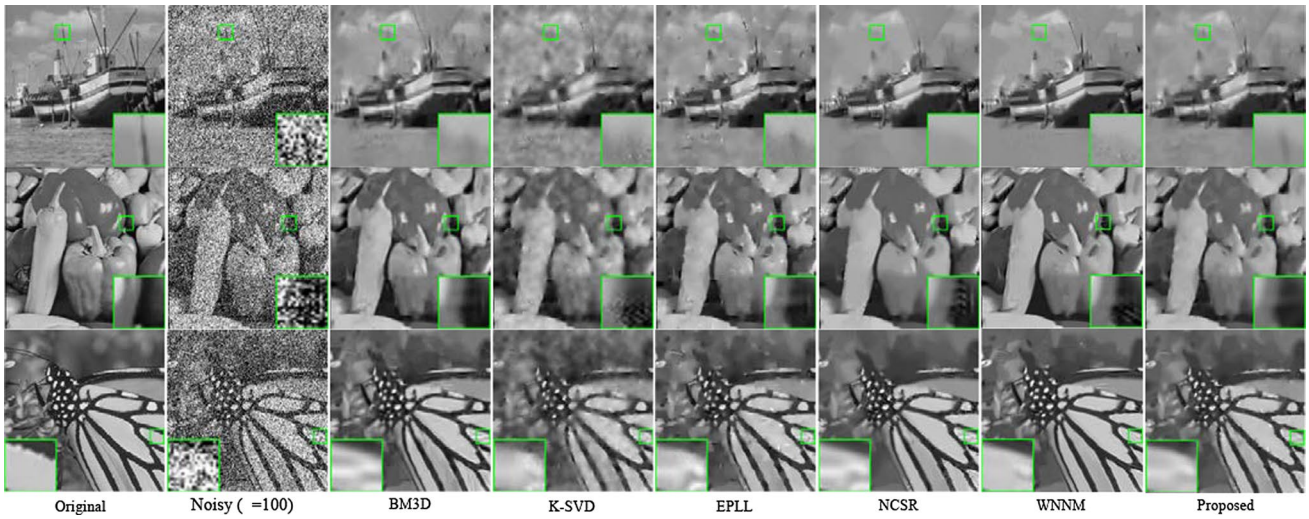


Fig. 5 Comparison of denoising effect for subjective evaluation, the 3 selected images are boat, peppers and Monarch. The classical sparse representation algorithms are selected: K-SVD [10], EPLL [11], BM3D [3], NCSR [12], WNNM [18]. The enlarged details are indicated by green squares

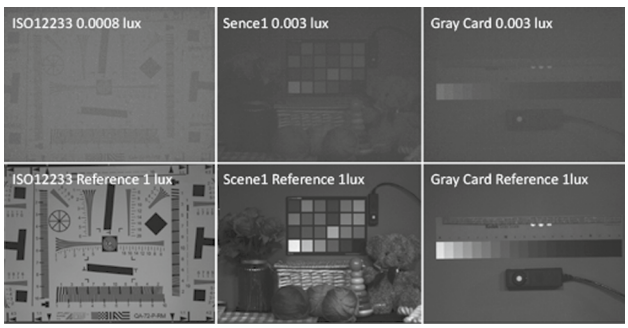


Fig. 6 The real low-light photograph dataset. The pictures are collected with luma target $\leq 3\times10^{-3}$ lx and exposure time = 1/25 s. The top row are the low-light photographs for example, the bottom row are the reference images taken at 1 lx

quantity and quality of similar patches for sparse prior are improved, making the NSS-based method more efficient. By testing on the standard test dataset and the real low-light dataset, the results indicate that the proposed algorithm could restore the details and textures effectively during denoising and result in visually pleasant images. The method proposed surpasses the previous methods alike and achieves competitive performance even compared with remarkable low rank methods.

The method based on sparse model optimization has promoted many advanced image techniques; however, these methods suffer from shortcoming of time-consuming which limits their applications in real world. In particular, with the fast development of machine learning technology, many deep learning-based denoising methods have been invented [38–41], which are classified as the discriminative learning method with the advantages of excellent performance on specific problems and fast processing speed, and disadvantages of poor generalization ability and difficult to collect the training pairs.

On the contrary, the model optimization method, such as sparse representation, is of good generalization and provides relatively satisfactory denoising performance, and how to improve the processing speed of the method is our next focus. Inspired by the deep learning work on image denoising, a natural optimization method could be raised by converting the solution of a sparse model to a network like the deep learning network, and it is expected to take both advantages of retaining the image prior for better generalization and transferring the time-consuming work to the pre-training part. The final testing could be high efficiency with GPU hardware accelerations.

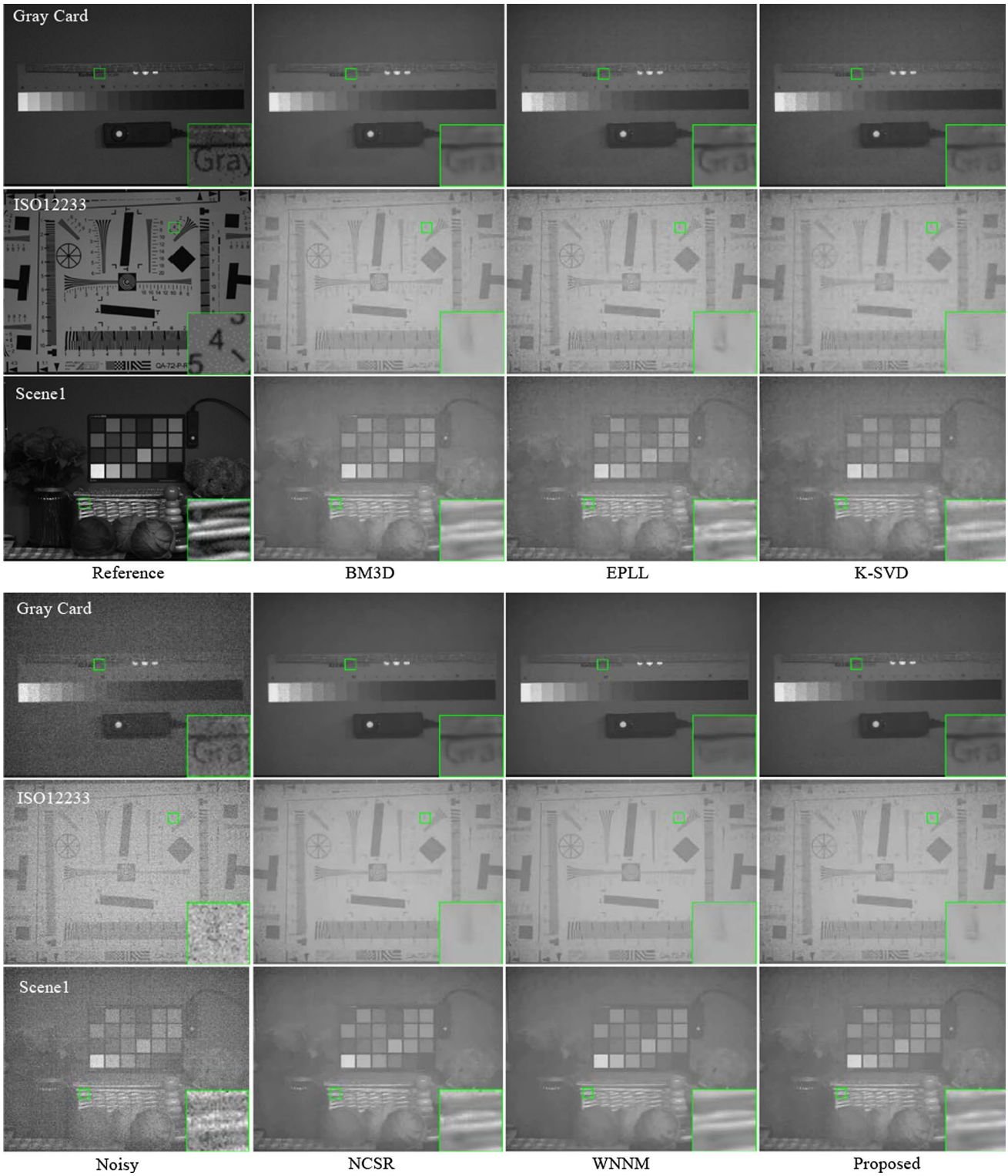


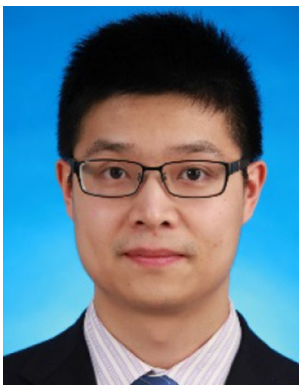
Fig. 7 Comparison of denoising performance for subjective evaluation on real photograph dataset. The input noisy images are after pre-processing, and 3 scenes are selected for demonstration, the enlarged details are indicated by green squares

References

1. Buades, A., Coll, B., Morel, J.M.: A review of image de-noising algorithms with a new one. J. Multiscale Model. Simul. **4**(5), 490–530 (2010)

2. Zhang, L., Dong, W., Zhang, D., Shi, G.: Two-stage image denoising by principal component analysis with local pixel grouping. Pattern Recogn. **43**(4), 1531–1549 (2010)
3. Dabov, K., Foi, A., Katkovnik, V., et al.: Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans. Image Process. **16**(8), 2080–2095 (2007)
4. Xu, J., Zhang, L., Zuo, W., Zhang, D., Feng, X.: Patch group based nonlocal self-similarity prior learning for image denoising, Proc. *IEEE International Conference on Computer Vision*, October (2015) pp. 64–78
5. Mairal, J., Bach, F., Ponce, J., Sapiro, G, Zisserman, A: Non-local sparse models for image restoration. Proc. *IEEE International Conference on Computer Vision*, September 2009, pp. 2272–2279
6. Dong, W., Shi, G., Li, X.: Nonlocal image restoration with bilateral variance estimation: a low-rank approach. IEEE Trans. Image Process. **22**(2), 700–711 (2012)
7. Huang, W., Wang, Q., Li, X.L.: Denoising-based multiscale feature fusion for remote sensing image captioning. IEEE Geosci. Remote Sens. Lett. **18**(3), 436–440 (2021)
8. Yuan, Y., Lin, J.Z., Wang, Q.: Hyperspectral image classification via multitask joint sparse representation and stepwise MRF optimization. IEEE Transactions Cybern. **46**(12), 2966–2977 (2016)
9. Kaur, M., Sharma, K., et al.: Image denoising using wavelet thresholding. Proc. Int. J. Eng. Comput. Sci. **2**(10), 2932–2935 (2012)
10. Elad, M., et al.: Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans. Image Process. **15**(12), 3735–3746 (2006)
11. Zoran, D., Weiss, Y.: From learning models of natural image patches to whole image restoration, Proc. IEEE International Conference on Computer Vision, November (2011) pp. 479–486
12. Dong, W., Zhang, L., Shi, G., Li, X.: Nonlocally centralized sparse representation for image restoration. IEEE Image Processing Transactions **22**(4), 1620–1630 (2012)
13. Wang, S., Zhang, L., Liang, Y.: Nonlocal spectral prior model for low-level vision, Proc. 11th Asian conference on Computer Vision, November (2012)
14. Candes, E.J., Recht, B.: Exact matrix completion via convex optimization. Found. Comput. Math. **9**, 717–772 (2009)
15. Chen, F., Zhang, L., Yu, H. M.: External Patch Prior Guided Internal Clustering for Image Denoising, Proc. IEEE International Conference on Computer Vision, December (2015) pp 603–611
16. Guo, Q., Zhang, C.M., et al.: An efficient SVD-based method for image denoising. IEEE Trans. Circuits Syst. Video Technol. **26**(5), 868–880 (2016)
17. Gu, S. H., Zhang, L. et al.: Weighted nuclear norm minimization with application to image denoising, Proc. IEEE Conference on Computer Vision and Pattern Recognition, June 2014, pp. 1–8
18. Xie, Y., Gu, S.H., et al.: Weighted schatten p-norm minimization for image denoising and background subtraction. IEEE Trans. Image Process. **25**(10), 4842–4857 (2016)
19. Wu, T., Zhang, R., Jiao, Z. H. et al.: Adaptive spectral rotation via joint cluster and pairwise structure. IEEE Transactions on Knowledge and Data Engineering (Early Access), April (2021)
20. Zhang, R., Li, X. L., Zhang, H. Y. et al., "Geodesic Multi-Class SVM with Stiefel Manifold Embedding", IEEE Transactions on Pattern Analysis and Machine Intelligence (Early Access), March (2021)
21. Zhang, R., Li, X.L.: Unsupervised feature selection via data reconstruction and side information. IEEE Trans. Image Process. **29**, 8097–8106 (2020)
22. Zhang, R., Zhang, H.Y., Li, X.L.: Robust multi-task learning with flexible manifold constraint. IEEE Trans. Pattern Anal. Mach. Intell. **43**(6), 2150–2157 (2021)
23. Zhang, R., Tong, H. H.: Robust Principal Component Analysis with Adaptive Neighbors, 33rd Conference on Neural Information Processing Systems, (2019)
24. Gu, S.H., Xie, Q., et al.: Weighted nuclear norm minimization and its applications to low level vision. Int. J. Comput. Vision **121**(7), 183–208 (2017)
25. Dong, W., Zhang, L. et al.: Centralized sparse representation for image restoration, Proc. IEEE International Conference on Computer Vision, November 2011, pp 1259–1266
26. Candes, E. J.: Compressive Sampling, Proceedings of the International Congress of Mathematicians, Madrid, Spain, (2006)
27. Candes, E.J., Romberg, J., Tao, T.: Robust uncertainty principles exact signal reconstruction from highly incomplete frequency information. IEEE Transactions Information Theory **52**(2), 489–509 (2016)
28. Candes, E.J., Romberg, J.: Sparsity and incoherence in compressive sampling. Inverse Prolems **23**, 969–985 (2017)
29. Daubechies, I., Defrise, M., Dye Mol, C.: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Commun. Pure Appl. Math. **57**(11), 1413–1457 (2004)
30. Zhang, L., Zuo, W.M.: Image restoration: from sparse and low-rank priors to deep priors. IEEE Signal Process. Mag. **34**(5), 172–179 (2017)
31. Donoho, D.L.: De-noising by soft-thresholding. IEEE Trans. Inf. Theory **41**(3), 613–627 (1995)
32. Sendur, L., Selesnick, I.W.: Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency. IEEE Trans. Signal Process. **50**(11), 2744–2756 (2002)
33. Olshausen, B.A., Field, D.J.: Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature **381**(13), 607–609 (1996)
34. Gordo, A., Almazan, J. et al.: Deep image retrieval: learning global representations for image search, Proc. European Conference on Computer Vision, September 2016, pp 241–257
35. Shlens, J: A Tutorial on Principal Component Analysis, arXiv: 1404.1100, April (2014)
36. Elad, M., Aharon, M.: Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans. Image Process. **15**(12), 3736–3745 (2006)
37. Zoran, D., Weiss, Y. et al.: Scale invariance and noise in natural images, Proc. IEEE 12th International Conference on Computer Vision, October 2009, pp 2209–2216
38. Chen, Y., Pock, T.: Trainable nonlinear reaction diffusion: a flexible framework for fast and effective image restoration. IEEE Transactions Pattern Anal. Mach. Intell. **39**(6), 1256–1272 (2016)
39. Zhang, K., Zuo, W.M., et al.: Beyond a Gaussian Denoiser: residual learning of deep CNN for image denoising. IEEE Image Process. Transactions **26**(7), 3142–3155 (2017)
40. Zhang, K., Zuo, W.M., Zhang, L.: Ffdnet: toward a fast and flexible solution for cnn based image denoising. IEEE Image Process. Transactions **27**(9), 4608–4622 (2018)
41. Dong, W., Wang, P.Y., et al.: Denoising prior driven deep neural network for image restoration. IEEE Trans. Pattern Anal. Mach. Intell. **41**(10), 2305–2318 (2019)

Dr. Zhou Tao graduated from School of Microelectronics and Solid State Electronics, University of Electronic Science and Technology of China in 2009 and obtained his Ph.D. at Shanghai Institute of Microsystems and Information Technology, Chinese Academy of Sciences in 2014. He joined Shanghai Integrated Circuits R&D Center Co. Ltd. (ICRD) in 2018, and he engaged in CMOS image sensor (CIS) imaging optimization and image analysis. At present, he mainly studies the measurement and characterization of images from CIS, and focuses on various image processing problems in fab manufacturing.



Dr. Li Chen received the B.S. degree in physics from Peking University, China, in 2005, and received the Ph.D. degree at the Institute of Microelectronics, Peking University, China, in 2010. He was a visiting scholar with the Department of Electrical Engineering, Columbia University, from January 2008 to January 2009. In 2010, he joined Shanghai Integrated Circuits R&D Center Co. Ltd. (ICRD), China, where he is currently the Director of the Department of Artificial Intelligence (AI) Technology. His research interests include high-end CMOS image sensor, AI chip architecture and algorithm, leading-edge IC technology, etc. He has authored or coauthored over 30 papers and is the holder of over 100 Chinese and USA patents.

Xuan Zeng (M'97) received the B.S. and Ph.D. degrees in electrical engineering from Fudan University, Shanghai, China, in 1991 and 1997, respectively. She is currently a Full Professor with the Department of Microelectronics, Fudan University. She was a Visiting Professor at the Department of Electrical Engineering, Texas A&M University, College Station, TX, USA, and the Department of Microelectronics, Technische Universiteit Delft, Delft, The Netherlands, in 2002 and 2003, respectively. She was the Director of the State Key Laboratory of ASIC & System from 2008 to 2012. Her current research interests include analog circuit modeling and synthesis, design for manufacturability, high-speed interconnect analysis and optimization, circuit simulation, and ASIC design. She has published more than 200 refereed journal/conference papers and 14 filed Chinese patents.

Zhao Yuhang born in 1972, received Ph.D. degrees of science from Fudan University, Shanghai, China. He is mainly engaged in the research and management of microelectronics and integrated circuit technology, and has more than 60 domestic authorized invention patents, 7 international authorized patents, and completed nearly 20 major national, provincial and ministerial technology research and development projects. He is currently the chief engineer of Shanghai Huahong (Group) Co., Ltd and the Chairman of Shanghai Integrated Circuits R&D Center Co., Ltd (ICRD).