

## Research Article

# Image Denoising Algorithm Combined with SGK Dictionary Learning and Principal Component Analysis Noise Estimation

Wenjing Zhao , Yue Chi, Yatong Zhou , and Cheng Zhang

Tianjin Key Laboratory of Electronic Materials and Devices, Hebei University of Technology, Tianjin 300401, China

Correspondence should be addressed to Yatong Zhou; [zhouyatong\\_zw@126.com](mailto:zhouyatong_zw@126.com)

Received 21 August 2017; Revised 31 January 2018; Accepted 13 February 2018; Published 14 March 2018

Academic Editor: Daniel Zaldivar

Copyright © 2018 Wenjing Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

SGK (sequential generalization of  $K$ -means) dictionary learning denoising algorithm has the characteristics of fast denoising speed and excellent denoising performance. However, the noise standard deviation must be known in advance when using SGK algorithm to process the image. This paper presents a denoising algorithm combined with SGK dictionary learning and the principal component analysis (PCA) noise estimation. At first, the noise standard deviation of the image is estimated by using the PCA noise estimation algorithm. And then it is used for SGK dictionary learning algorithm. Experimental results show the following: (1) The SGK algorithm has the best denoising performance compared with the other three dictionary learning algorithms. (2) The SGK algorithm combined with PCA is superior to the SGK algorithm combined with other noise estimation algorithms. (3) Compared with the original SGK algorithm, the proposed algorithm has higher PSNR and better denoising performance.

## 1. Introduction

In the image acquisition and transmission, noise is inevitably carried, which will reduce image quality, so image denoising has a very important significance. Image denoising algorithms can be divided into space domain denoising and frequency domain denoising. The former includes the mean filtering, median filtering, and Wiener filtering. The latter includes Fourier transform [1], Laplace transform [2], and wavelet transform [3]. A series of postwavelet multiscale tools have been developed based on the wavelet theory to filter noise effectively such as curvelet [4], directionlet [5], bandelet [6], and shearlet [7].

In recent years, there are some novel denoising algorithms such as nonlocal mean [8] denoising, Gaussian mixture model denoising [9], and dictionary learning denoising [10] based on sparse representation [11]. An image denoising method based on wavelet and SVD transforms improves denoising performance [12]. Moreover,  $K$ -singular value decomposition ( $K$ -SVD) [13] based on overcomplete sparse representation has recently been the subject of intense research activity within the denoising community [14, 15]. However,  $K$ -SVD increases the iteration number when dealing with large data. So Sujit proposed the SGK [13] dictionary

learning algorithm in 2013, which not only overcomes the drawbacks of ordinary dictionary learning that breaks the sparse coefficient structure but also can be applied to a variety of sparse representations, with the low complexity and fast calculation ability [10].

At present, many image denoising algorithms need to foreknow the noise standard deviation [16], but it is usually unknown in practice. So the noise estimation has been developed in the image denoising community. The classic image filtering in [17] estimates the noise standard deviation by the convolution of image and filter. The DCT of the image patch [18] concentrates the image structure in the low frequency coefficient region, so that the noise estimation can be performed by the high frequency coefficient. It is also common to estimate noise level by the grayscale value of the image [19]. Patch-based local variance [20] generally estimates noise level by robust statistical algorithms. The Bayesian contraction algorithm [21] is used to denoise the image and analyze the autocorrelation of residuals in the range of noise standard deviation to find the true value. The distribution of the sideband filter response [22] can be divided into two parts according to the difference of the image and noise, which is calculated by the expected

maximization [23]. The kurtosis of the edge sideband filter response distribution [24] is constant for the noisy image, and a kurtosis model can be established and the noise standard deviation can be evaluated by finding the best parameters of the model. However, the above algorithms mostly assume that the image is uniform. For images with abundant textures, Pyatykh et al. [25] proposed PCA noise estimation based on the data patch, where the noise standard deviation can be estimated as the minimum eigenvalue of the image patch covariance matrix.

Based on the above considerations, a denoising algorithm combined with SGK dictionary learning and PCA noise estimation is proposed. Firstly, the image with additive Gaussian white noise is segmented, and the noise level is estimated by calculating the minimum eigenvalue of the image patch covariance matrix. Then the estimated noise standard deviation is entered into SGK dictionary learning algorithm to denoise the image. During the denoising process, each image patch is sparse and the sparse representation coefficient is calculated by pursuit algorithm. The dictionary atom is updated with the sparse representation coefficient; therefore a more accurate approximation of the image patch is obtained. The experimental results show that the proposed algorithm is superior to other algorithms in noise level estimation and has better denoising performance.

## 2. SGK Dictionary Learning Denoising Algorithm

2.1. *Image Denoising Problem and SGK Dictionary Learning.* SGK dictionary learning algorithm is a generalization of

the  $K$ -means clustering. It mainly consists of two stages: sparse coding stage and dictionary update stage when using SGK dictionary learning algorithm to perform denoising [13], and the flow chart is shown in Figure 1. SGK algorithm firstly processes image through the original DCT dictionary and then updates dictionary with the sparse representation coefficient. Each local patch extracted in the image is sparse-coded by new training dictionary to achieve the denoising performance.

2.2. *Sparse Coding Stage.* For an image  $\mathbf{A}$  of size  $\sqrt{T} \times \sqrt{T}$  added to additive white Gaussian noise  $\mathbf{W} \in R^{\sqrt{T} \times \sqrt{T}}$ , it constitutes a noisy image  $\mathbf{B}$ :

$$\mathbf{B} = \mathbf{A} + \mathbf{W}. \quad (1)$$

Assume that the dictionary  $\mathbf{D} \in R^{t \times l}$  consists of image atoms  $\mathbf{d}_l \in R^t$ , where  $l = 1, 2, \dots, L$ .  $\mathbf{Q}_{ij}$  represents a  $t \times T$  matrix that extracts patches of size  $\sqrt{t} \times \sqrt{t}$  in image  $\mathbf{A}$ , which is  $\forall_{ij} \{\mathbf{Q}_{ij}\mathbf{A} \in R^t\}$ . For each local patch, the sparse representation  $\mathbf{a} = \mathbf{Q}_{ij}\mathbf{A}$  can be represented by a dictionary  $\mathbf{D}$ :

$$\hat{\beta} = \arg \min_{\beta} \left\{ \eta \|\beta\|_0 + \|\mathbf{D}\beta - \mathbf{a}\|_2^2 \right\}. \quad (2)$$

For any patch in the image,

$$\hat{\beta}_{ij} = \arg \min_{\beta_{ij}} \left\{ \eta_{ij} \|\beta_{ij}\|_0 + \|\mathbf{D}\beta_{ij} - \mathbf{Q}_{ij}\mathbf{A}\|_2^2 \right\} \quad \forall_{ij}. \quad (3)$$

Therefore the global image representation is shown as

$$\{\hat{\mathbf{A}}, \hat{\beta}_{ij}\} = \arg \min_{\mathbf{A}, \beta_{ij}} \left\{ \rho \|\mathbf{B} - \mathbf{A}\|_2 + \sum_{ij} \eta_{ij} \|\beta_{ij}\|_0 + \sum_{ij} \|\mathbf{D}\beta_{ij} - \mathbf{Q}_{ij}\mathbf{A}\|_2 \right\}. \quad (4)$$

For the solution of (4),  $\beta_{ij}$  can be obtained by (3), and then  $\mathbf{B}$  is represented as sparse approximation of  $\mathbf{A}$  by choosing appropriate  $\eta_{ij}$ , so it can be obtained as

$$\begin{aligned} \hat{\beta}_{ij} &= \arg \min_{\beta_{ij}} \|\beta_{ij}\|_0 \\ \text{s.t.} \quad &\|\mathbf{Q}_{ij}\mathbf{B} - \mathbf{D}\beta_{ij}\|_2^2 \leq (C\sigma)^2 \\ &\forall_{ij}. \end{aligned} \quad (5)$$

2.3. *Dictionary Update Stage.* In the dictionary update stage, updating each image's atoms sequentially can minimize sparse representation error, which is denoted as

$$\mathbf{g}_{ij} = \mathbf{Q}_{ij}\mathbf{A} - \mathbf{D}\beta_{ij} = \mathbf{Q}_{ij}\mathbf{A} - \sum_r \mathbf{d}_m \beta_{ij}(r). \quad (6)$$

In (6),  $\beta_{ij}(r)$  is the  $r$ th component of  $\beta_{ij}$ . And the error matrix  $\mathbf{G}$  is composed of all these elements  $\{\mathbf{g}_{ij}\}$ . So the error of the image patch  $\mathbf{d}_l$  can be expressed as

$$\mathbf{g}_{ij}^l = \mathbf{Q}_{ij}\mathbf{A} - \sum_{r \neq l} \mathbf{d}_m \alpha_{ij}(r) = \mathbf{g}_{ij} + \mathbf{d}_l \alpha_{ij}(l). \quad (7)$$

All these  $\{\mathbf{g}_{ij}^l\}$  form the error matrix  $\mathbf{G}^l$  and also form vector  $\beta_l$  containing corresponding  $\{\beta_{ij}(l)\}$ , so it has

$$\begin{aligned} \mathbf{G}^l &= \mathbf{E} + \mathbf{d}_l \beta_l \implies \\ \|\mathbf{G}\|_F^2 &= \|\mathbf{G}^l - \mathbf{d}_l \beta_l\|_F^2. \end{aligned} \quad (8)$$

And  $\|\cdot\|_F$  is Frobenius norm in (8). According to the sequential generalization of  $K$ -means [12], the solution of (8) is

$$\mathbf{d}_l^{(t+1)} = \arg \min_{\mathbf{d}_l} \|\mathbf{G}\|_F^2 = \arg \min_{\mathbf{d}_l} \|\mathbf{G}^l - \mathbf{d}_l \beta_l\|_F^2. \quad (9)$$

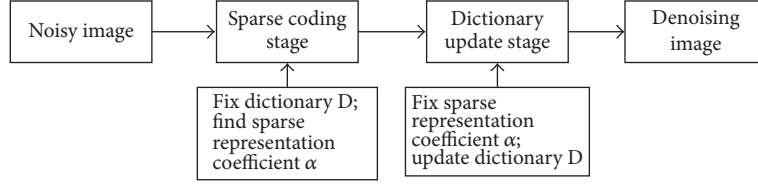


FIGURE 1: SGK algorithm denoising flow chart.

The closed-form solution of (9) is

$$\hat{\mathbf{d}}_l = \mathbf{G}^l \boldsymbol{\beta}_l^T (\boldsymbol{\beta}_l \boldsymbol{\beta}_l^T)^{-1}. \quad (10)$$

It replaces all atoms with  $\mathbf{d} = \hat{\mathbf{d}}_l$  and SGK training dictionary  $\hat{\mathbf{D}}$  is used to obtain the final sparse representation of the component  $\hat{\beta}_{ij}$  for each extracted local patch.  $\hat{\mathbf{A}}$  is obtained as

$$\hat{\mathbf{A}} = \arg \min_{\mathbf{A}} \left\{ \rho \|\mathbf{B} - \mathbf{A}\|_2 + \sum_{ij} \|\mathbf{D} \hat{\beta}_{ij} - \mathbf{Q}_{ij} \mathbf{A}\|_2 \right\}. \quad (11)$$

The final solution of the sparse representation error minimization problem is

$$\hat{\mathbf{A}} = \left( \lambda \mathbf{I}_N + \sum_{ij} \mathbf{Q}_{ij}^T \mathbf{Q}_{ij} \right)^{-1} \left( \lambda \mathbf{B} + \sum_{ij} \mathbf{Q}_{ij}^T \mathbf{D} \hat{\beta}_{ij} \right). \quad (12)$$

### 3. Noise Estimation Theory

**3.1. Noise Estimation Theory Based on PCA.** Suppose that  $\mathbf{A}$  is a clean image with size of  $\sqrt{T} \times \sqrt{T}$ , and  $\mathbf{B}$  represents an image with additive white Gaussian noise  $\mathbf{W}$ . The noise variance is unknown, so it needs to be estimated. For  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{W}$ , each image contains  $H = (\sqrt{T} - \sqrt{M} + 1)(\sqrt{T} - \sqrt{M} + 1)$  patches with size  $M = \sqrt{M} \times \sqrt{M}$ . Since  $\mathbf{W}$  is the independent additive white Gaussian noise, there is  $\mathbf{W} \sim N_M(0, \sigma^2 \mathbf{I})$  and  $\text{cov}(\mathbf{A}, \mathbf{W}) = 0$ .

Suppose that  $\mathbf{J}_A$  and  $\mathbf{J}_B$  are, respectively, the sample covariance matrices of  $\mathbf{A}$  and  $\mathbf{B}$ . Meanwhile,  $\tilde{\mu}_{A,1} \geq \tilde{\mu}_{A,2} \geq \dots \geq \tilde{\mu}_{A,M}$  are the eigenvalues of  $\mathbf{J}_A$ , and the corresponding eigenvectors are  $\tilde{\mathbf{W}}_{A,1}, \dots, \tilde{\mathbf{W}}_{A,M}$ . Similarly,  $\tilde{\mu}_{B,1} \geq \tilde{\mu}_{B,2} \geq \dots \geq \tilde{\mu}_{B,M}$  are the eigenvalues of  $\mathbf{J}_B$ , and corresponding eigenvectors are  $\tilde{\mathbf{W}}_{B,1}, \dots, \tilde{\mathbf{W}}_{B,M}$ .  $\tilde{\mathbf{W}}_{B,1}^T \mathbf{B}, \dots, \tilde{\mathbf{W}}_{B,M}^T \mathbf{B}$  represent the sample principal component of  $\mathbf{B}$  [26], and  $S^2$  represents the sample variance, so it is shown as follows:

$$S^2 (\tilde{\mathbf{W}}_{B,k}^T \mathbf{B}) = \tilde{\mu}_{B,k}, \quad k = 1, 2, \dots, M, \quad (13)$$

where  $S^2$  represents the sample variance.

In order to apply PCA to noise variance estimation, it defines positive integer  $c$ . The clean image  $\mathbf{A}$  satisfies  $\mathbf{A}_i \in \mathbf{W}_{M-c} \subset R^M$ , and its dimension  $M - c$  is less than the number of coordinates  $M$ . So there is

$$\mathbf{E}(|\tilde{\mu}_{B,h} - \sigma^2|) = O\left(\frac{\sigma^2}{\sqrt{H}}\right), \quad H \rightarrow \infty. \quad (14)$$

And it is held for all  $h = M - c + 1, \dots, M$ . When considering the overall principal component,  $\text{cov}(\mathbf{A}, \mathbf{W}) = 0$  represents  $\sum_B = \sum_A + \sum_W$ , where  $\sum_B$ ,  $\sum_A$ , and  $\sum_W$ , respectively, represent the overall covariance matrices of  $\mathbf{B}$ ,  $\mathbf{A}$ , and  $\mathbf{W}$ . Meanwhile, the minimum eigenvalues of  $\sum_W = \sigma^2 \mathbf{I}$  and  $\sum_A$  are zero, so the minimum eigenvalue of  $\sum_B$  is  $\sigma^2$ . With the sample size  $N$  tending to infinity, it meets

$$\lim_{N \rightarrow \infty} \mathbf{E}(|\tilde{\mu}_{B,M} - \sigma^2|) = 0, \quad (15)$$

which represents the fact that  $\tilde{\mu}_{B,M}$  converges to  $\sigma^2$ , so the noise variance can be estimated as  $\tilde{\mu}_{B,M}$  and it is a consistent estimation of the noise level.

If the above assumptions hold, the expected values of  $\tilde{\mu}_{B,M-c+1} - \tilde{\mu}_{B,M}$  can be calculated from the trigonometric inequality and (15):

$$\mathbf{E}(\tilde{\mu}_{B,M-c+1} - \tilde{\mu}_{B,M}) = O\left(\frac{\sigma^2}{\sqrt{H}}\right). \quad (16)$$

The condition of (16) is

$$\tilde{\mu}_{B,M-c+1} - \tilde{\mu}_{B,M} < \frac{T\sigma^2}{\sqrt{H}}, \quad (17)$$

where  $T$  is a fixed value and it satisfies  $T > 0$ .

For the estimation of noise variance  $\sigma_{\text{est}}^2$ , it can be verified by (17). If (17) holds,  $\sigma_{\text{est}}^2$  is the final estimation. But if (17) cannot hold, it is necessary to extract a subset of the image patches with a small standard deviation. Performing noise estimation again to satisfy (17) is satisfied until the final noise estimation is obtained.

### 3.2. Estimate the Noise Level Based on PCA: An Example.

Figure 2 is a PCA noise estimation example of the house image. Figure 2(a) is the house image, and Figure 2(b) shows the noise estimation results under the different noise standard deviation. It can be seen that the PCA noise estimation value is very close to the true value.

## 4. Experimental Results' Analysis

The standard Kodak Photo CD benchmark was used to evaluate the performance of denoising algorithm. The size of some images is  $256 \times 256$ , and the size of other images is  $512 \times 512$ . The patches sizes of all images are  $8 \times 8$ .

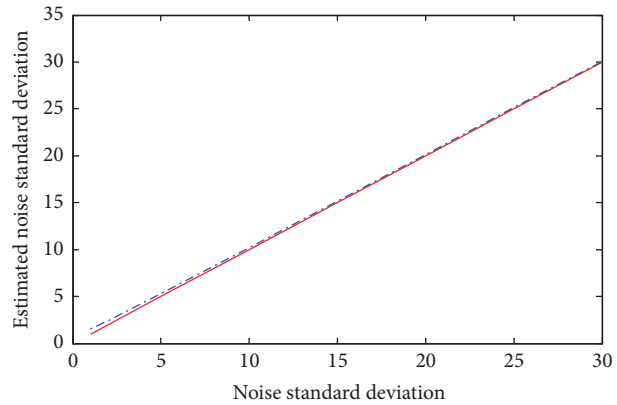
**4.1. Comparison of Four Dictionary Learning Denoising Algorithms.** We conduct experiments for image denoising by using SGK, DCT, Global, K-SVD dictionary learning

TABLE 1: Image denoising results of five dictionary learning algorithms.

Algorithms	Time/s	PSNR/dB	MSE
SGK	17.378	32.077	77.470
DCT	83.884	31.079	113.873
Global	100.479	31.734	89.435
K-SVD	284.181	32.171	72.383
BM3D	12.256	30.717	119.478



(a) House image



(b) Noise level estimation results

FIGURE 2: Noise estimation based on principal component analysis.

algorithms, and BM3D algorithms. The similarity of the first four algorithms is to build a dictionary and then use the dictionary to denoise. DCT algorithm denoises an image by sparsely representing each block with the overcomplete DCT dictionary, thus averaging the represented parts [11]. Global algorithm denoises an image by training a dictionary on patches from the noisy image, sparsely representing each block with this dictionary and averaging the represented parts [11]. K-SVD algorithm uses DCT dictionary to initialize and then uses singular value decomposition for dictionary updating [11]. BM3D algorithm is an image signal denoising method based on transform domain enhancement sparse representation [26].

Throughout this experiment, we use SGK, DCT, Global, K-SVD, and BM3D algorithms to denoise the Barbara image with  $\sigma = 25$  as an example. As shown in Figures 3(b)–3(f), the denoising results of five different algorithms are basically the same, and the image's details are basically well preserved. In the following, we do quantitative comparisons between five algorithms, and the experimental data is shown in Table 1. The PSNR of the SGK is similar to that of K-SVD, and it is superior to Global, DCT, and BM3D algorithms. Encouragingly, we see that SGK runs much faster than K-SVD. As to the value of MSE, SGK is smaller than Global, DCT, and BM3D algorithms. With all the above-mentioned results, the denoising supremacy of SGK over the rest algorithms is demonstrated.

*4.2. Comparison of SGK Combined with Different Noise Estimation Algorithms.* In order to analyze the sensitivity of SGK algorithm to the noise level, the noise standard deviation is set as  $\sigma = 5, 10, 15, 20, 25$ , respectively. SGK algorithm is used to denoise the Peppers image with different noise standard deviation.

Figure 4 shows the denoising experiment's results using the SGK algorithm with different offsets. The five different colors curves, respectively, represent the case where the PSNR varies with the noise offsets if the noise standard deviation is given. It can be found in Figure 4 that the PSNR of denoised image is basically invariant when the noise standard deviation has negative offset of 0~−5% and the forward offset of 0~+5%. When the offset of noise standard deviation continues to increase, it shows a significant downward trend in −5%~−25% of the negative offset and 5% to 25% of forward offset, which indicates that the image's PSNR is significantly reduced. Therefore the PSNR would be changed when the noise standard deviation has negative or forward offset, which shows that the SGK algorithm is sensitive to the offsets of the noise standard deviation. So it is necessary to estimate noise level before image denoising. If the estimated noise level is close to the true noise, the denoising results will be more accurate.

In order to analyze the performances of SGK dictionary denoising combined with different noise estimation



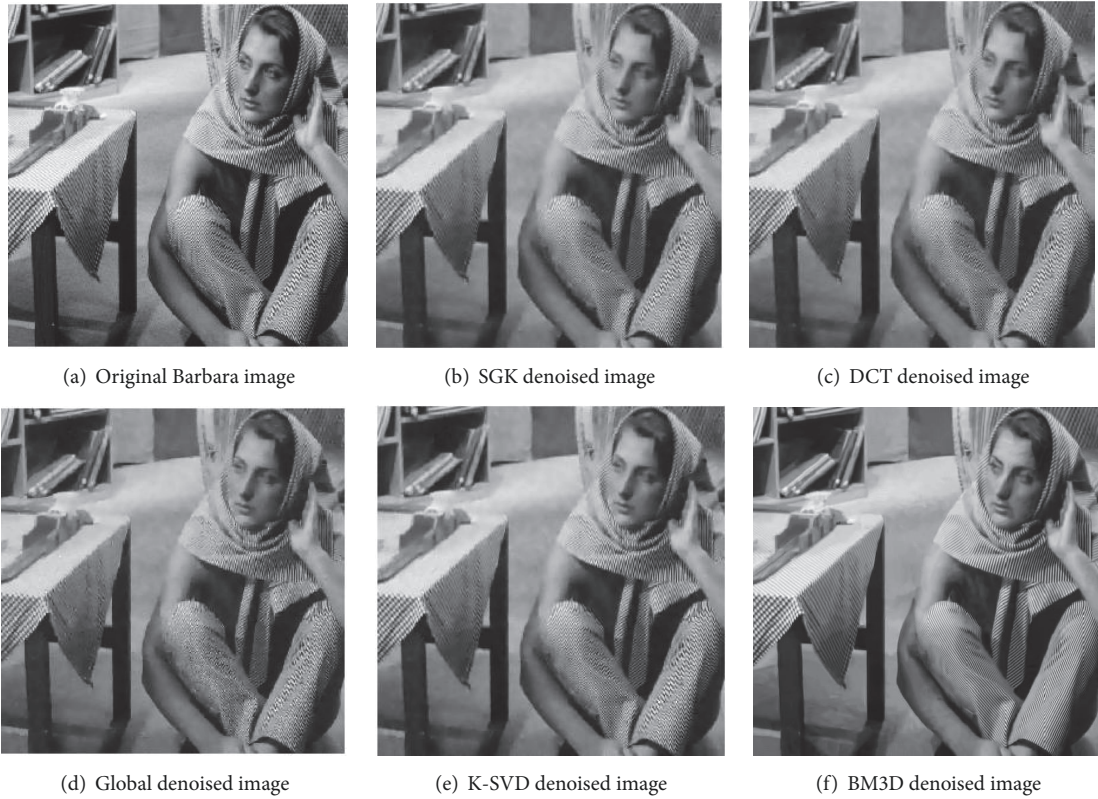


FIGURE 3: Using five kinds of dictionary learning algorithm for image denoising.

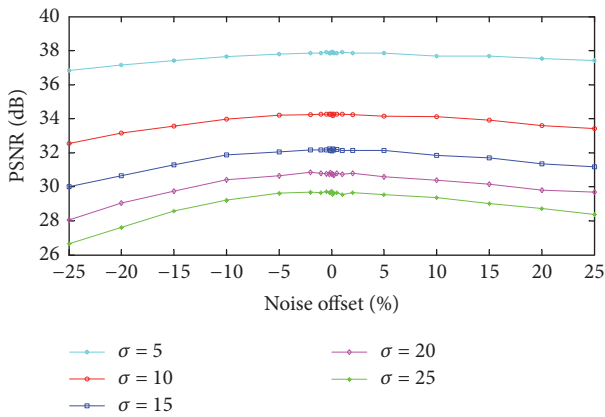


FIGURE 4: SGK denoising performance with different noise offset.

algorithms, this paper also introduces four noise estimation algorithms: Kurtosis [25], local standard deviation distribution mode (Mode) [27], local standard deviation (Med) [27], and local standard deviation minimum (Min) [27].

Kurtosis assumes that the corresponding distribution of the kurtosis edge bandpass filter should be a constant for the noise-free images. However, the kurtosis at the entire scale of a noisy image may vary. Under this assumption, the noise standard deviation can be estimated by the kurtosis model.

Mode can divide image and the noise standard deviation is estimated according to the distribution pattern of the

image's local standard deviation. As the variance of noise is constant throughout the picture, it will affect every local variance value equally. As a result, the maximum of the bell-shaped distribution will reflect the local variance of the degraded image within homogeneous areas. This value is the mode of the distribution, and it is very close to the mode, so we can use the mode to estimate it.

Med estimates noise standard deviation based on the median of the image's local standard deviation. If practical difficulties to properly estimate the mode arise, it may be useful to use the median operator instead, due to its greater simplicity and the fact that both parameters, albeit different, are not far apart in practice. So an alternative estimation procedure can be done.

Min estimates the noise standard deviation based on the minimum value of the local standard deviation of the images. Within a uniform area, the variance of the degraded image equals the variance of noise. According to the previous statement, one straightforward way to estimate standard deviation is to calculate the variance within homogeneous regions, where the variance of the original image is close to zero.

The above noise estimation algorithms are all combined with the SGK dictionary learning algorithm in this paper. Now we combine the five noise estimation algorithms mentioned above with SGK, respectively. The combined algorithms are referred to as PCA + SGK, Kurtosis + SGK, Mode + SGK, Med + SGK, and Min + SGK. Figure 5 shows the



FIGURE 5: Comparison of five kinds of algorithms.

denoising results of the Lena image, where the noise standard deviation is set to  $\sigma = 15$ . Analysis of Figure 5 leads to the following observation: the denoising performance of Mode + SGK in Figure 5(d) is slightly the worst, while the remaining four algorithms are basically the same. They all retain the details of the image in the denoising process.

In order to analyze the denoising performance of five algorithms quantitatively, additive Gaussian white noise images with  $\sigma = 2, 5, 10, 20, 30, 40$  are, respectively, denoised and compared in Table 2. The best results are shown in boldface. Experiments performed on noisy Lena image indicate that the proposed algorithm outperforms, in terms of estimation accuracy  $|\sigma_{\text{est}} - \sigma|$ , estimation time, PSNR, and MSE, the four existing algorithms.

Figure 6 illustrates the performance of algorithms more intuitively varying with standard deviation. Figure 6(a) shows variation of noise estimation absolute error with the noise standard deviation. It can be seen that PCA + SGK has the least value, which is superior to the other four algorithms. So the estimation of PCA + SGK is the most accurate. Figure 6(b) shows the variation of the noise estimation time with the noise standard deviation. For the noise estimation time, Mode + SGK, Med + SGK, and Min + SGK are the least accurate, followed by PCA + SGK and Kurtosis + SGK. Figure 6(c) shows the variation of PSNR with the noise standard deviation. With the increase of standard deviation, PCA + SGK and Kurtosis + SGK keep the PSNR higher, followed by Med + SGK and Mode + SGK, and Min + SGK

has the lowest PSNR. Figure 6(d) shows the variation of MSE with the noise standard deviation. In this experiment, PCA + SGK owns the lowest value of MSE, which means that the denoising performance is better than any of the other four algorithms.

**4.3. Denoising Experiment of Noisy Image with Unknown Standard Deviation.** The above experiments presumably assumed that the standard deviation of the noise contained in the image is known. In order to demonstrate the advantage of the proposed PCA + SGK algorithm, it is used to denoise the noisy image with unknown standard deviation and compare it with the original SGK algorithm. Twelve classic original images are shown in Figure 7. These images are mixed into additive white Gaussian noise with unknown standard deviation. After that, we do PCA + SGK denoising and SGK denoising, respectively. When SGK is used for denoising, the standard deviation can only be guessed based on the noisy image or given a random value because the noise standard deviation is unknown. When PCA + SGK is used for denoising, the noise standard deviation is first estimated by PCA and thus is entered into SGK for denoising. Denoising results of the Cameraman image using these two algorithms are shown in Figure 8. Figure 8(a) shows the noise image with  $\sigma = 10$ . Figure 8(b) shows the denoised image by SGK and Figure 8(c) shows the denoised image by PCA + SGK. The experiment testifies for the good performance of our approach. It can be seen that the denoising performance of

TABLE 2: Denoising indicators on SGK combined with five noise estimation algorithms.

$\sigma$	Algorithms	$\sigma_{est}$	$\sigma_{est} - \sigma$	$ \sigma_{est} - \sigma $	Time/s	PSNR/dB	MSE
$\sigma = 2$	PCA + SGK	<b>2.733</b>	<b>+0.733</b>	<b>0.733</b>	1.799	<b>43.385</b>	<b>2.983</b>
	Kurtosis + SGK	2.841	+0.841	0.841	4.699	43.378	2.987
	Mode + SGK	2.768	+0.768	0.768	0.070	43.222	3.010
	Med + SGK	4.063	+2.063	2.063	0.051	41.811	4.286
	Min + SGK	0.496	-1.504	1.504	<b>0.037</b>	42.119	3.993
$\sigma = 5$	PCA + SGK	<b>5.510</b>	<b>+0.510</b>	<b>0.510</b>	1.829	<b>38.383</b>	<b>9.436</b>
	Kurtosis + SGK	5.166	+0.166	0.166	5.865	38.302	9.614
	Mode + SGK	4.480	-0.520	0.520	0.091	38.139	9.981
	Med + SGK	6.194	+1.194	1.194	0.047	38.251	9.727
	Min + SGK	0.674	-4.326	4.326	<b>0.035</b>	34.175	24.865
$\sigma = 10$	PCA + SGK	<b>10.342</b>	<b>+0.342</b>	<b>0.342</b>	1.604	<b>35.177</b>	<b>19.201</b>
	Kurtosis + SGK	9.576	+0.424	0.424	4.848	34.795	21.559
	Mode + SGK	8.809	+1.191	1.191	0.090	34.074	25.447
	Med + SGK	10.463	+0.463	0.463	0.051	35.149	19.869
	Min + SGK	1.640	-18.360	18.360	<b>0.037</b>	28.244	97.417
$\sigma = 15$	PCA + SGK	<b>15.153</b>	<b>+0.147</b>	<b>0.147</b>	1.995	<b>34.269</b>	<b>30.632</b>
	Kurtosis + SGK	14.196	+0.804	0.804	5.208	33.899	30.760
	Mode + SGK	12.494	+2.506	2.506	0.084	31.381	47.311
	Med + SGK	14.838	-0.162	0.162	0.082	33.221	30.971
	Min + SGK	1.807	-13.193	13.193	<b>0.029</b>	28.274	96.766
$\sigma = 20$	PCA + SGK	<b>20.071</b>	<b>+0.071</b>	<b>0.071</b>	1.618	<b>31.911</b>	<b>41.876</b>
	Kurtosis + SGK	18.943	-1.057	1.057	4.818	31.574	45.256
	Mode + SGK	17.038	-2.963	2.963	0.094	30.186	62.300
	Med + SGK	19.319	-0.681	0.681	0.055	31.777	43.193
	Min + SGK	1.877	-18.123	18.123	<b>0.028</b>	22.152	396.153
$\sigma = 30$	PCA+SGK	<b>29.672</b>	<b>-0.328</b>	<b>0.328</b>	1.393	<b>29.917</b>	<b>66.286</b>
	Kurtosis + SGK	28.480	-1.520	1.520	5.126	29.724	69.292
	Mode + SGK	23.356	-6.644	6.644	0.084	29.044	161.709
	Med + SGK	28.199	-1.801	1.801	0.054	29.614	71.074
	Min + SGK	3.292	-26.708	26.708	<b>0.029</b>	18.675	882.180
$\sigma = 40$	PCA + SGK	<b>39.546</b>	<b>-0.454</b>	<b>0.454</b>	1.865	<b>28.371</b>	<b>94.630</b>
	Kurtosis + SGK	38.238	-1.762	1.762	5.134	28.305	96.062
	Mode + SGK	33.463	-6.537	6.537	0.087	25.813	170.526
	Med + SGK	37.068	-2.932	2.932	0.053	27.910	105.212
	Min + SGK	5.385	-34.615	34.615	<b>0.034</b>	16.222	1152.049

PCA + SGK is better than SGK, which retains more details of the original image.

The denoising results of 12 images are shown in Table 3. It is seen that the PSNR of SGK is less than that of PCA + SGK; that is, PCA + SGK has better denoised performance than PCA. Because the standard deviation of noisy image is not given when using SGK, the noise level can only be guessed and entered into SGK for denoising. While using PCA + SGK to deal with noisy image, PCA is first used to estimate the standard deviation, and then the estimated value is entered into SGK for denoising, so the denoising performance is better. Quantitative comparisons with traditional SGK illustrate the benefits of PCA + SGK.

## 5. Conclusions

In this paper, the algorithm of PCA noise estimation combined with SGK dictionary learning was proposed to denoise image. The noisy image is first divided into patches, and

the noise standard deviation is estimated by calculating the minimum eigenvalue of the image patch covariance matrix. After that, the estimated noise standard deviation is entered into SGK dictionary learning algorithm. The sparse representation of each training sample is obtained by sparse coding, and the dictionary atoms are updated by dictionary updating to denoise the image. This algorithm effectively solves the problem that the SGK algorithm requires a prior noise standard deviation for image denoising. This paper has the following three conclusions.

Firstly, the SGK dictionary learning algorithm is compared with K-SVD, DCT, Global, and BM3D algorithms. The PSNR of SGK algorithm and those of the other four algorithms do not have much difference, and the MSE of SGK algorithm is only higher than K-SVD algorithm. SGK algorithm owns great advantage in denoising time, which is much faster than K-SVD, DCT, and Global algorithms. Therefore, the SGK algorithm has the best denoising performance.



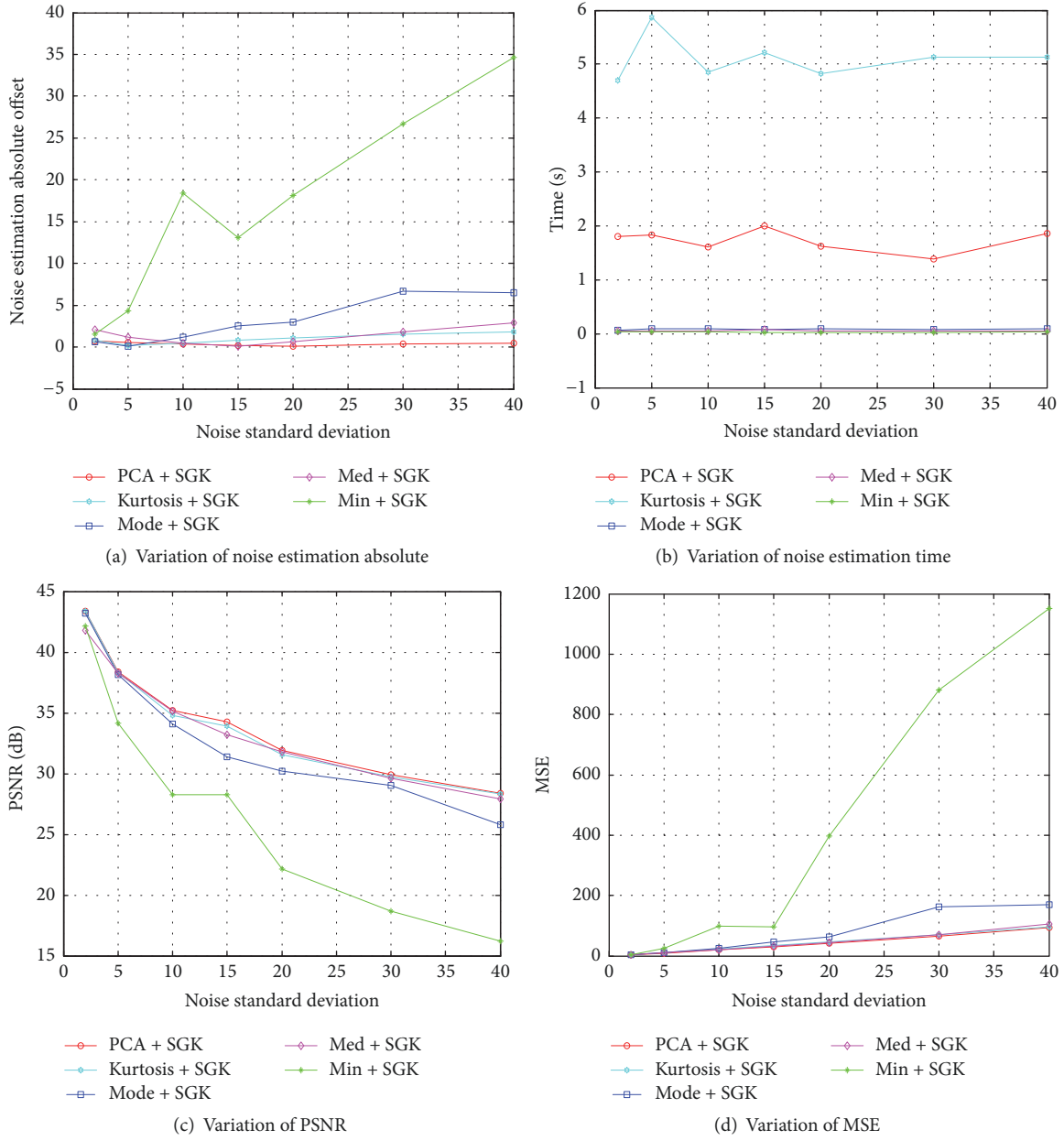


FIGURE 6: Variation of denoising performance with noise standard deviation on five algorithms.



FIGURE 7: Twelve classic images.



TABLE 3: Denoising results of two algorithms for 12 images.

Image	Gessed noise value $\sigma'$	PSNR of SGK	PCA estimated value $\sigma_{est}$	PSNR of PCA + SGK
Barbara	12.15	31.5826	3.25	40.8875
Boat	20.02	31.0877	4.96	36.9938
Bridge	29.78	24.7522	8.67	32.1713
Cameraman	35.11	27.5935	10.22	33.7171
Couple	30.45	26.1943	12.30	32.2721
Fingerprint	20.10	28.0063	15.06	29.7811
Flintstones	4.98	23.4492	19.10	29.0998
Hill	5.23	22.5119	20.21	29.9451
House	10.00	22.8280	22.58	32.4558
Lena	15.15	23.7830	24.34	30.8314
Man	15.50	21.8376	28.04	28.3323
Peppers	18.24	22.4973	30.38	28.7696

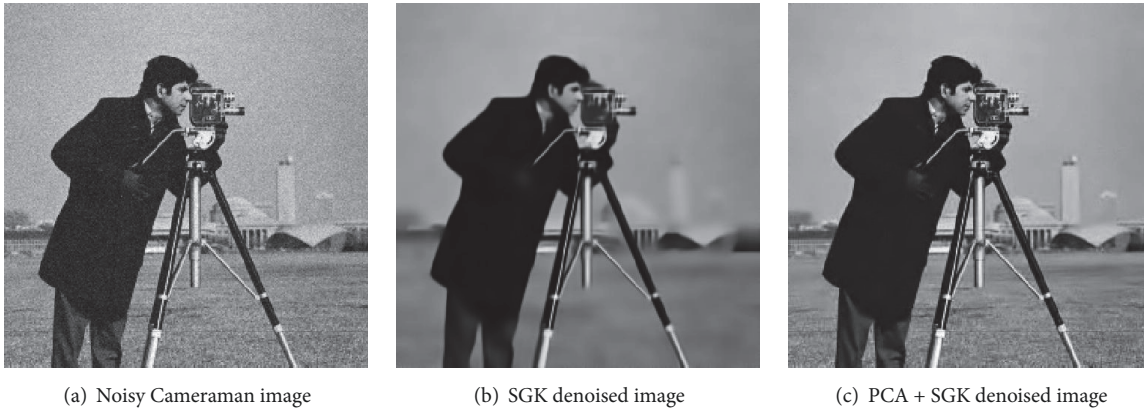


FIGURE 8: Comparison of two denoising algorithms on the Cameraman image.

Secondly, the PCA algorithm is compared with the other four noise estimation algorithms: Kurtosis, Mode, Mad, and Min. The five algorithms are, respectively, combined with the SGK algorithm to denoise the additive Gaussian white noise images with different standard deviation. The absolute deviation of the noise estimated by PCA + SGK is the smallest, and it is better than the other four algorithms; that is, the noise standard deviation estimation of this algorithm is the most accurate. For the noise estimation time, Min + SGK, Med + SGK, and Mode + SGK all have a faster estimation and then it is PCA + SGK algorithm proposed in this paper and Kurtosis + SGK is the slowest. On the other hand, PCA + SGK and Kurtosis + SGK keep the high PSNR, followed by Med + SGK and Mode + SGK. The lowest PSNR is that of Min + SGK. At the same time, the MSE value of PCA + SGK is the lowest. So the denoising performance of proposed algorithm is better than the other four algorithms. It is found that the proposed algorithm is more accurate to estimate the noise standard deviation with faster denoising speed and good denoising performance.

Thirdly, PCA + SGK and SGK are, respectively, used to denoise the image with different standard deviation. Experiments show that PSNR of PCA + SGK is much higher than that of SGK. When using SGK for denoising, the noise

standard is unclear, so the denoising performance is not good, while PCA + SGK firstly uses the PCA to estimate the noise standard deviation, which is close to the true value of the noise level, so the denoising performance is more ideal and the image's details are better preserved. While performance improvement is different for different images, the results nonetheless indicate the potential of proposed algorithm over original SGK algorithms.

### Conflicts of Interest

The authors declare that there are no conflicts of interest.

### Acknowledgments

This work was supported by the Natural Science Foundation of Hebei Province (no. E2016202341) and Humanity and Social Science Foundation of Ministry of Education of China (15YJA630108).

### References

- [1] A. M. Scarfone, "Deformed Fourier transform," *Physica A: Statistical Mechanics and its Applications*, vol. 480, pp. 63–78, 2017.

- [2] N. Yalcin, E. Celik, and A. Gokdogan, "Multiplicative Laplace transform and its applications," *Optik - International Journal for Light and Electron Optics*, vol. 127, no. 20, pp. 9984–9995, 2016.
- [3] Y. Hel-Or and D. Shaked, "A discriminative approach for wavelet denoising," *IEEE Transactions on Image Processing*, vol. 17, no. 4, pp. 443–457, 2008.
- [4] J.-L. Starck, E. J. Candes, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Transactions on Image Processing*, vol. 11, no. 6, pp. 670–684, 2002.
- [5] V. Velisavljević, B. Beferull-Lozano, M. Vetterli, and P. L. Dragotti, "Directionlets: anisotropic multidirectional representation with separable filtering," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 1916–1933, 2006.
- [6] E. Le Pennec and S. Mallat, "Sparse geometric image representations with bandelets," *IEEE Transactions on Image Processing*, vol. 14, no. 4, pp. 423–438, 2005.
- [7] B. G. Bodmann, G. Kutyniok, and X. Zhuang, "Gabor shearlets," *Applied and Computational Harmonic Analysis*, vol. 38, no. 1, pp. 87–114, 2015.
- [8] A. Ben Said, R. Hadjidj, K. Eddine Melkemi, and S. Fofou, "Multispectral image denoising with optimized vector non-local mean filter," *Digital Signal Processing*, vol. 58, pp. 115–126, 2016.
- [9] X. Cong-Hua, C. Jin-Yi, and X. Wen-Bin, "Medical image denoising by generalised Gaussian mixture modelling with edge information," *IET Image Processing*, vol. 8, no. 8, pp. 464–476, 2014.
- [10] S. K. Sahoo and A. Makur, "Dictionary training for sparse representation as generalization of K-means clustering," *IEEE Signal Processing Letters*, vol. 20, no. 6, pp. 587–590, 2013.
- [11] S. Liu, L. Li, Y. Peng, G. Qiu, and T. Lei, "Improved sparse representation method for image classification," *IET Computer Vision*, vol. 11, no. 4, pp. 319–330, 2017.
- [12] M. Wang, Z. Li, X. Duan, and W. Li, "An image denoising method with enhancement of the directional features based on wavelet and SVD transforms," *Mathematical Problems in Engineering*, vol. 2015, Article ID 469350, 9 pages, 2015.
- [13] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3736–3745, 2006.
- [14] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [15] B. Dumitrescu and P. Irofti, "Regularized K-SVD," *IEEE Signal Processing Letters*, vol. 24, no. 3, pp. 309–313, 2017.
- [16] P. Jiang and J.-Z. Zhang, "Fast and reliable noise level estimation based on local statistic," *Pattern Recognition Letters*, vol. 78, pp. 8–13, 2016.
- [17] D.-H. Shin, R.-H. Park, S. Yang, and J.-H. Jung, "Block-based noise estimation using adaptive gaussian filtering," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 1, pp. 218–226, 2005.
- [18] N. N. Ponomarenko, V. V. Lukin, M. S. Zriakhov, A. Kaarna, and J. Astola, "An automatic approach to lossy compression of AVIRIS images," in *Proceedings of the 2007 IEEE International Geoscience and Remote Sensing Symposium, (IGARSS '07)*, pp. 472–475, Spain, June 2007.
- [19] X. Liu, M. Tanaka, and M. Okutomi, "Single-image noise level estimation for blind denoising," *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 5226–5237, 2013.
- [20] B. R. Corner, R. M. Narayanan, and S. E. Reichenbach, "Noise estimation in remote sensing imagery using data masking," *International Journal of Remote Sensing*, vol. 24, no. 4, pp. 689–702, 2003.
- [21] P. Wyatt and H. Nakai, "Developing nonstationary noise estimation for application in edge and corner detection," *IEEE Transactions on Image Processing*, vol. 16, no. 7, pp. 1840–1853, 2007.
- [22] A. Barducci, D. Guzzi, P. Marcoionni, and I. Pippi, "Assessing noise amplitude in remotely sensed images using bit-plane and scatterplot approaches," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 8, pp. 2665–2675, 2007.
- [23] W. Yao, "A note on EM algorithm for mixture models," *Statistics & Probability Letters*, vol. 83, no. 2, pp. 519–526, 2013.
- [24] D. Zoran and Y. Weiss, "Scale invariance and noise in natural images," in *Proceedings of the 12th International Conference on Computer Vision (ICCV '09)*, pp. 2209–2216, October 2009.
- [25] S. Pyatykh, J. Hesser, and L. Zheng, "Image noise level estimation by principal component analysis," *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 687–699, 2013.
- [26] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3D filtering," in *Proceedings of the Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning*, vol. 6064, pp. 354–365, USA, January 2006.
- [27] S. Aja-Fernández, G. Vegas-Sánchez-Ferrero, M. Martín-Fernández, and C. Alberola-López, "Automatic noise estimation in images using local statistics. Additive and multiplicative cases," *Image and Vision Computing*, vol. 27, no. 6, pp. 756–770, 2009.

