

Research Article

A Novel Directionlet-Based Image Denoising Method Using MMSE Estimator and Laplacian Mixture Distribution

Yixiang Lu, Qingwei Gao, Dong Sun, Dexiang Zhang, Yi Xia, and Hui Wang

School of Electrical Engineering and Automation, Anhui University, Hefei 230601, China

Correspondence should be addressed to Qingwei Gao; qingweigao@ahu.edu.cn

Received 1 October 2014; Revised 29 December 2014; Accepted 23 February 2015

Academic Editor: Igor Djurović

Copyright © 2015 Yixiang Lu 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.

A novel method based on directionlet transform is proposed for image denoising under Bayesian framework. In order to achieve noise removal, the directionlet coefficients of the uncorrupted image are modeled independently and identically by a two-state Laplacian mixture model with zero mean. The expectation-maximization algorithm is used to estimate the parameters that characterize the assumed prior model. Within the framework of Bayesian theory, the directionlet coefficients of noise-free image are estimated by a nonlinear shrinkage function based on weighted average of the minimum mean square error estimator. We demonstrate through simulations with images contaminated by additive white Gaussian noise that the proposed method is very competitive when compared with other methods in terms of both peak signal-to-noise ratio and visual quality.

1. Introduction

The recent advancement in multiscale geometric analysis has promoted an enormous amount of research in the field of image processing, including image denoising. The distortion of images by additive white Gaussian noise is common during the images acquisition, processing, compression, and transmission. The goal of image restoration is to reconstruct a plausible estimate of the original image from the distorted image or observation and keep the image features as much as possible at the same time.

As a representative of multiscale geometric analysis tools, wavelet transform (WT) [1] played an important role in the application of image denoising and has achieved great success over the last decade. There are plenty of works on image denoising using wavelet thresholding since Donoho [2, 3] has done the pioneering work for signal denoising using WT. The method described in [3] has near-optimal properties in the minimax sense and performs well in simulation studies of one-dimensional curve estimation. The image denoising methods using threshold mainly consist of three steps: (1) representing the image sparsely using WT; (2) determining an appropriate threshold and comparing the coefficients with it, preserving or modifying the few high-magnitude transform coefficients that convey mostly the true-signal energy

and discarding the rest which are mainly due to noise; (3) reconstructing the denoised image using inverse WT based on the thresholded coefficients. Due to their effectiveness and simplicity that operate on one wavelet coefficient at a time, various wavelet-based thresholding methods were proposed for image denoising [4–6]. Although the denoising performance of the wavelet thresholding is better than that of spatial filters (such as the filters proposed in [7–9]), the problem of how to develop an optimal threshold used for distinguishing between the coefficients that are mainly due to signal and those mainly due to noise has not been solved well. Thus, for the issue of image denoising, the wavelet-based thresholding algorithms with threshold are not an optimal strategy.

To improve the denoised results, the prior probabilistic knowledge of the noise-free image and noise image in wavelet domain is exploited in the process of image denoising. In the statistical wavelet-based denoising approaches, instead of using an arbitrary threshold to shrink the noisy wavelet coefficients, the shrinkage function is designed by minimizing a Bayesian risk, typically under the minimum mean square error (MMSE) criterion [10], minimum mean absolute error (MMAE) criterion [11], or maximum a posteriori (MAP) criterion [12]. In general, modeling the statistics of natural images is a challenging task, partly because of the high dimension of the signal. However, the power of statistical

image models can be substantially improved by representing the spatial image using WT, since the WT nearly decorrelates the dependencies between image pixels. By exploiting the nature of the histogram of the wavelet coefficients, different noise reduction algorithms try to incorporate the prior model that best matches the true image coefficients. In [13], Mallat proposed that the wavelet coefficients show non-Gaussian behavior and sharp peak centering around zero with heavy-tailed distribution and used the generalized Gaussian distribution (GGD) with the shape parameter s to fit the coefficients. Since then, the GGD has been commonly used to model the image wavelet coefficients and accordingly derived different image denoising algorithms based on some criteria [14, 15]. For example, Moulin and Liu [15] first assumed that the wavelet coefficients follow a simple independent and identically distributed GGD prior, then estimated the shape parameter using Mallat's algorithm proposed in [13], and finally developed a shrinkage estimator using MAP criterion based on the assumed prior. In addition to the GGD prior, several other PDFs have also been proposed to model the wavelet coefficients of the image, such as Gaussian scale mixture (GSM) distribution [16, 17], α -stable distribution [18], and Bessel K -form distribution [19]. Among these proposed models used for the purpose of denoising, the GSM model accounts for both marginal and pairwise joint distributions of the wavelet coefficients, and thus the corresponding denoising methods outperform the methods using other prior models. Although the α -stable model provides a more accurate fitness to the histogram of the WT, this probability density function (PDF) has no closed-form expression even though its characteristic function has one, and its parameter estimation is poor, specially in the presence of noise [20]. This disadvantage hampers its wide application for image denoising in practice. It has been shown that the application of prior knowledge in the noise reduction algorithms results in a performance better than that of the wavelet-based denoising methods using thresholds. However, due to the inherent drawbacks of the wavelet bases which lack anisotropy and have poor directional selectivity, wavelets fail to capture the geometric regularity along the singularities of surfaces (e.g., the edges and contours), thus seriously affecting the performances of the wavelet-based methods.

To exploit the anisotropic regularity of a surface along edges, several image representations with elongated basis functions have been proposed to capture the geometric regularity of a given image, including curvelets [21], bandelets [22], and contourlets [23]. They were also introduced into the field of image denoising because of their ability of efficiently representing the anisotropic regularity [24, 25]. For example, in [25], Eslami and Radha modeled the joint PDF of parents and children of the coefficients in semi-translation-invariant contourlet transform and implemented image denoising by constructing a bivariate shrinkage function under MAP criterion. However, since the directional filter banks used in these representation tools are inseparable, the construction of directional filters is very complex, which is the major disadvantage of these transforms. Moreover, the process of image decomposition using directional filters is more time consuming than decomposition using wavelets.

Apart from these image representation tools mentioned above, another related image representation scheme called sparse representation [26, 27] was proposed and extended to the image denoising community. Sparse representation accounts for most or all information of a signal with a linear combination of a small number of elementary signals called atoms which are chosen from a so-called overcomplete dictionary. The results reported in [27] demonstrated that the denoising performance based on learned dictionaries surpasses that of wavelet-based methods. However, the algorithm imposes a very high computational burden, since it utilizes highly overcomplete dictionaries obtained via a preliminary training procedure.

In this paper, a new denoising method based on directionlet transform is proposed. Because of anisotropy and multidirectionality of directionlets, they allow us to capture more image features which can enhance the denoising performance. To recover the noise-free image using the Bayesian technique, we model the directionlet coefficients to be a two-state Laplacian mixture PDF with zero mean and use the expectation-maximization (EM) algorithm which is a specialization to the mixture density estimation problem to estimate the parameters involved in the assumed PDF. The noise-free coefficients are estimated by using average MMSE estimator.

The rest of this paper is organized as follows. The directionlet transform is briefly reviewed in Section 2. The statistical model of the directionlet coefficients of the noise-free image is studied in Section 3 as well as the discussion on the goodness of fit. Moreover, the design of the MMSE estimator is also presented in this section. The performance of the proposed algorithm is evaluated by using experimental results and compared with other existing denoising methods in Section 4. Finally, some conclusions are drawn in Section 5.

2. Directionlet Transform

Discontinuity curves (i.e., the edges and contours) presented in the images are highly anisotropic and they are characterized by a geometrical coherence. These features are not properly captured by the standard two-dimension (2D) WT that uses isotropic basis functions. Because of this, when an image is decomposed by the standard 2D WT, many wavelet bases intersect the discontinuity curves and result in many large magnitude coefficients. Obviously, in such cases, WT is not an optimal sparse representation for images. To efficiently capture these elongated structures characterized by geometric regularity along different directions (not only the horizontal and vertical), the directionlet transform [28] based on lattice is proposed recently. This transform is constructed by using two concepts: directionality and anisotropy, and they are realized through two steps, respectively. First, use an integer lattice which is represented by a nonunique generator matrix

$$M_{\Lambda} = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix}, \quad a_1, a_2, b_1, b_2 \in Z \quad (1)$$

to partition the discrete space, where the two vectors \mathbf{d}_1 (with slope $r_1 = b_1/a_1$) and \mathbf{d}_2 (with slope $r_2 = b_2/a_2$) determine

two directions which are called transform direction and alignment direction, respectively. By sampling the image with generator matrix M_Λ , one can get $|\det(M_\Lambda)|$ cosets. Second, achieve the anisotropy by using an unequal number of 1D wavelet transform steps along the two directions; that is, the 1D transform is applied more along one direction than the other direction. It should be noticed that if $|\det(M_\Lambda)| > 1$, then the iterated 1D transform should be applied in each coset separately. For notational simplicity, the directionlet transform is denoted as $\text{SAWT}(M_\Lambda, n_1, n_2)$ in the rest of this paper, where n_1 and n_2 are the numbers of 1D WT along transform direction and alignment direction in each iteration, respectively. The basis functions of SAWT are called directionlets, which include elongated functions that are very suitable for image representation. The frequency decomposition of $\text{SAWT}(M_\Lambda, 2, 1)$ along 0° for images is shown in Figure 1.

From the construction principle of directionlets, it can be seen that the $\text{SAWT}(M_\Lambda, n_1, n_2)$ retains the separable filtering and subsampling, the simplicity of computations and filter design from the standard 2D WT, unlike the case for some other directional transform constructions (e.g., curvelets, contourlets, or bandelets). The corresponding asymmetric basis functions have directional vanishing moments along any two directions with rational slopes, which allow for a sparser representation of elongated and oriented features. As a consequence of the improved sparsity, directionlets provide an efficient tool for nonlinear approximation of images, significantly outperforming the standard 2D WT.

3. PDF Modeling and MMSE Estimator

We consider the problem of image restoration as an MMSE estimate of the noisy image. Let us denote the observed data by $f = r + \epsilon$, where r is the clean data and ϵ is the additive Gaussian noise. Applying the $\text{SAWT}(M_\Lambda, n_1, n_2)$ to the observed data, we have

$$y = g + n, \quad (2)$$

where y , g , and n signify the directionlet coefficients of f , r , and ϵ , respectively. Since we would like to use MMSE estimator to recover the noise-free image, the PDF of variable g is needed.

3.1. Statistical Model of Directionlet Coefficients. In this subsection, we investigate the statistical distribution of directionlet coefficients of the image and show some results on modeling natural images. By casually observing natural images, one can find that they are highly inhomogeneous; that is, they typically contain many regions that are smooth, interspersed with “features” such as contours, or surface marking. After applying the standard 2D WT on the images, the marginal distributions of the detail coefficients which reflect the inhomogeneous characteristics show a large peak at zero and tails that fall significantly slower than a Gaussian of the same variance [13] (i.e., non-Gaussianity). Based on the observation, various PDFs [15, 17–19] have been proposed to model this non-Gaussianity in the past decades. Due to the directionlets including elongated basis functions that are

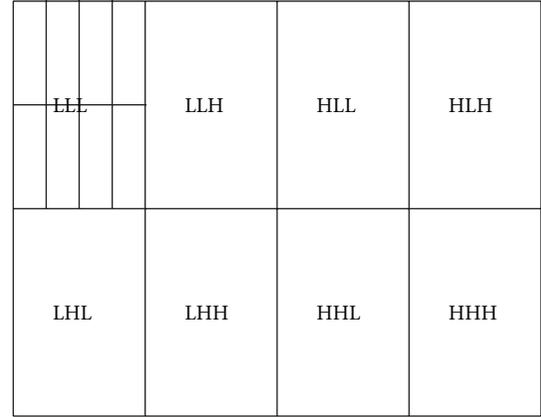


FIGURE 1: Frequency decomposition of $\text{SAWT}(M_\Lambda, 2, 1)$ along 0° , where two decomposition levels are used. H and L stand for high pass and low pass, respectively.

nearly parallel to the edges, the coefficients obtained by using $\text{SAWT}(M_\Lambda, n_1, n_2)$ contain more zeros than those obtained by using WT. Correspondingly, the marginal distributions of the detail directionlet subbands at every scale will exhibit stronger non-Gaussianity with heavier tails. Thus, in this paper, a mixture density of two Laplacian distributions with zero mean and heavy tails is proposed to model the directionlet coefficients of the images. This prior mixture PDF can be expressed as

$$p_g(g) = \sum_{i=0,1} a_i \cdot p_i(g) = \sum_{i=0,1} a_i \cdot \frac{1}{\sqrt{2}\sigma_i} \exp\left(-\frac{\sqrt{2}|g|}{\sigma_i}\right), \quad (3)$$

where $a_i \geq 0$ and $\sum_{i=0}^1 a_i = 1$ and $p_i(g)$ represents the i th Laplacian PDF denoted as $\text{Laplace}(g; \sigma_i)$ and where σ_i is the standard deviation of the i th Laplacian PDF. To determine whether the assumed prior PDF can provide an appropriate fit for the directionlet coefficients, we visually examine the degree of match between the histogram of the coefficients and the adopted Laplacian mixture PDF whose parameters are estimated from the coefficients. Figure 2(a) displays a natural image called “Couple” (<http://www.cipr.rpi.edu/resource/stills/misc1.html>). Figure 2(b) displays the histogram of the HHH_1 subband of the image shown in Figure 2(a) and a Laplacian mixture PDF fitted to this subband. Also the corresponding Gaussian density function with the variance estimated from the coefficients is displayed in Figure 2(b). It is clear that the Laplacian mixture PDF tends to fit the directionlet coefficients far better than the Gaussian PDF.

Naturally, another real question is whether the Laplacian mixture model describes the directionlet coefficients more accurately than other PDFs proposed in the literature. Here, we compare the assumed PDF fit with Gaussian PDF and Gaussian mixture PDF proposed in [29] and [16], respectively. The comparison is conducted by using a distance [30] between the modeled CDF and the empirical CDF of a given

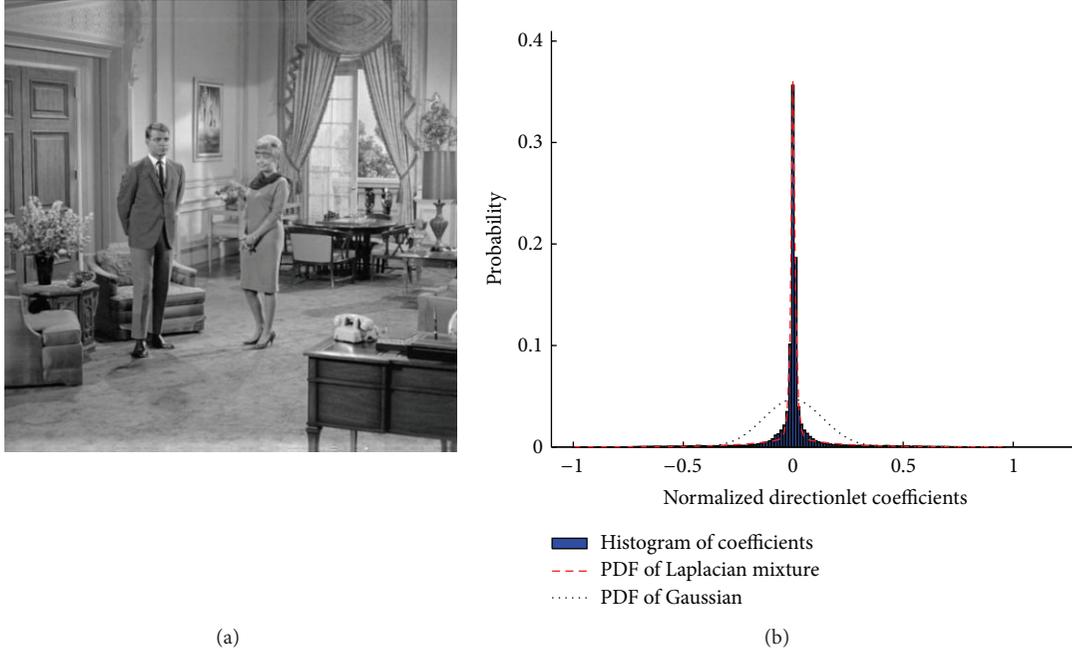


FIGURE 2: A natural image and its corresponding fitting models in directionlet domain. (a) Couple image. (b) Histogram of the normalized directionlet coefficients for image and the Gaussian and Gaussian mixture PDFs used to model the coefficients.

directionlet subband of the test image to determine which prior PDF approaches the empirical PDF more precisely. The distance was originally defined as

$$D = \max_{-\infty < x < +\infty} |P(x) - G(x)|, \quad (4)$$

where $P(x)$ and $G(x)$ denote the modeled CDF and the empirical CDF, respectively. To get the data samples, we use the specified SAWT($M_\Lambda, 2, 1$) with a Daubechies' Symlet 4 wavelet to decompose the test image along two directions 0° and 45° , respectively. Then, in each detail subband, the histogram of the coefficients is computed to generate the empirical CDF, and the modeled CDF is computed by estimating the corresponding parameters from the directionlet coefficients. It should be noted that when the transform direction 45° is adopted in SAWT($M_\Lambda, 2, 1$), two cosets will be produced in image decomposition and there will be too many detail subbands. But here, we only compute the distance on the subbands belonging to one of the two cosets due to the limited space. The resultant distances computed from Gaussian, Gaussian mixture, and the proposed PDFs for image "Couple" are reported in Table 1. It can be observed that the Laplacian mixture density provides smaller D values than the other compared models in most cases. In a few cases, even though the Gaussian mixture distribution produces a slightly smaller value of D than the proposed model, the latter, in fact, is already statistically very close to the histogram of the directionlet coefficients. This provides a strong evidence that the Laplacian mixture PDF fits to the directionlet subbands of the image more accurately than others. A similar conclusion applies to the other detail subbands at different composition levels.

3.2. MMSE Estimator. Having known the prior model of the directionlet coefficients of the noise-free image, now we derive the shrinkage function based on (2). Since the directionlet transform is orthogonal and linear, the variable n is still following a Gaussian distribution with the same variance as ϵ . In the Bayesian framework, the denoising task is performed by statistical estimation. For a given noisy observation y , there are infinitely many possible signals g that may have generated the observation. These candidates are not all equally likely, however, and follow the posterior probability distribution $p_{g|y}(g | y)$. One can obtain the most possible candidate $\hat{g}(y)$ under a given criterion. A common criterion used for getting $\hat{g}(y)$ is the MMSE which minimizes the expected squared error $\Delta g = \int p_{g|y}(g | y) \|\hat{g}(y) - g\|^2 dg$. The minimum error (MMSE estimator) is achieved for

$$\hat{g}(y) = E\{g | y\} = \int g p_{g|y}(g | y) dg, \quad (5)$$

where $E\{\cdot\}$ represents the mathematical expectation, the a posterior mean (i.e., the optimal value for g).

Using Bayes' rule, the MMSE estimator of g given noisy observation y can be easily derived as

$$\hat{g}(y) = \frac{\int g p_{y|g}(y | g) p_g(g) dg}{p_y(y)}, \quad (6)$$

where $p_y(y)$ is the convolution of $p_g(g)$ with $p_n(n) = (1/\sqrt{2\pi\sigma_n^2}) \exp(-n^2/(2\sigma_n^2))$. In (6), substituting various PDFs for $p_g(g)$ combined with (2) results in different estimators. For a simple prior PDF, one can directly derive the corresponding MMSE estimator. For example, if g obeys

TABLE 1: The cumulative distribution distance D values of Gaussian, Gaussian mixture, and Laplacian mixture PDFs, respectively, and the corresponding parameters estimated from two level subbands along two directions of Couple image.

Orientation (coset)	Model	Image subbands						
		HHH	HHL	LHH	HLH	LHL	HLL	LLH
		Level I						
0°	Gauss	0.2370	0.2365	0.2556	0.2730	0.2365	0.2253	0.2650
	GM	0.0214	0.0474	0.0267	0.0372	0.0327	0.0288	0.0342
	LM	0.0218	0.0281	0.0283	0.0211	0.0130	0.0313	0.0209
		Level II						
45° (coset 2)	Gauss	0.2215	0.2130	0.2575	0.2454	0.2130	0.2266	0.2651
	GM	0.0208	0.0232	0.0351	0.0389	0.0232	0.0312	0.0387
	LM	0.0155	0.0188	0.0157	0.0219	0.0187	0.0126	0.0217

a Laplacian distribution, then y follows a Laplacian-Gaussian distribution, which can be expressed as

$$\begin{aligned}
 p_y(y) &= \text{Laplace}(y; \sigma) * p_n(y) \\
 &= \frac{\exp(-y^2 / (2\sigma_n^2))}{2\sqrt{2}\sigma} \\
 &\quad \cdot \left[\text{erfcx}\left(\frac{\sigma_n}{\sigma} - \frac{y}{\sqrt{2}\sigma_n}\right) + \text{erfcx}\left(\frac{\sigma_n}{\sigma} + \frac{y}{\sqrt{2}\sigma_n}\right) \right], \quad (7)
 \end{aligned}$$

where $*$ denotes the convolution operator and $\text{erfcx}(x) = 2/\sqrt{\pi} \int_0^\infty e^{-t^2 - 2tx} dt$ is the complementary error function. For notational simplicity, throughout this paper we use $\text{LapGauss}(x; \sigma, \sigma_n)$ to denote the Laplacian-Gaussian distribution expressed in (7) and also use $\text{erfcx}(x_+, \sigma, \sigma_n)$ and $\text{erfcx}(x_-, \sigma, \sigma_n)$ to denote $\text{erfcx}(\sigma_n/\sigma + x/(\sqrt{2}\sigma_n))$ and $\text{erfcx}(\sigma_n/\sigma - x/(\sqrt{2}\sigma_n))$, respectively. Finally, the corresponding MMSE estimator can be derived as

$$\begin{aligned}
 \hat{g}(y) &= \left(\left(y - \frac{\sqrt{2}\sigma_n^2}{\sigma} \right) \text{erfcx}(y_-, \sigma, \sigma_n) \right. \\
 &\quad \left. + \left(y + \frac{\sqrt{2}\sigma_n^2}{\sigma} \right) \text{erfcx}(y_+, \sigma, \sigma_n) \right) \\
 &\quad \cdot (\text{erfcx}(y_-, \sigma, \sigma_n) + \text{erfcx}(y_+, \sigma, \sigma_n))^{-1}. \quad (8)
 \end{aligned}$$

However, it is usually difficult to derive the expression of the MMSE estimator for mixture PDFs. To get a closed-form estimator for Laplacian mixture prior, we introduce a hidden variable into (3).

Once the hidden variable is introduced, (3) can be rewritten as

$$p_g(g) = \sum_{i=0,1} p_s(s=i) \cdot p_{g|s}(g | s=i), \quad (9)$$

where $p(g | s=i) \sim \text{Laplace}(g; \sigma_i)$ and s is a binary hidden variable which can take on the value 0 (insignificant coefficient) or 1 (significant coefficient). Obviously, $p_s(s=0)$

and $p_s(s=1)$ are equal to a_0 and a_1 , respectively. The posterior probability $p_{g|y}(g | y)$ in (5) becomes

$$\begin{aligned}
 p_{g|y}(g | y) &= \frac{\sum_s p_{g,y,s}(g, y, s)}{p_y(y)} \\
 &= \sum_s p_{g|y,s}(g | y, s) p_{s|y}(s | y). \quad (10)
 \end{aligned}$$

Inserting (10) into the expression of the MMSE estimator (6) and exchanging the order of integration and summation yielded

$$\hat{g}(y) = \sum_s p_{s|y}(s | y) \int g p_{g|y,s}(g | y, s) dg. \quad (11)$$

From (11), one can see that the integral term corresponds to a Laplacian MMSE estimator $\hat{g}_{s=i}(y)$ defined in (8) under a given hidden state s . In fact, this is a weighted average of two MMSE estimates, where the weighting is determined by $p_{s|y}(s | y)$. Using Bayes' rule, the weightings $p_{s|y}(s | y)$ can be computed as

$$p_{s|y}(s | y) = \frac{p_{y|s}(y | s) p_s(s)}{\sum_s p_{y|s}(y | s) p_s(s)}, \quad (12)$$

where $p_{y|s}(y | s)$ follows a Laplacian-Gaussian distribution defined in (7) under hidden state s . Substituting (7), (8), and (12) into (11) yields the full Bayes MMSE estimator for the Laplacian mixture prior

$$\hat{g}(y) = \frac{\sum_{i=0,1} a_i \cdot \text{LapGauss}(y; \sigma_i, \sigma_n) \cdot \hat{g}_{s=i}(y)}{\sum_{i=0,1} a_i \cdot \text{LapGauss}(y; \sigma_i, \sigma_n)}, \quad (13)$$

where $a_i \geq 0$ and $\sum_{i=0}^1 a_i = 1$, y is the directionlet coefficient of noisy image, and σ_i and σ_n are the standard deviations of the i th Laplacian PDF and the Gaussian noise, respectively.

In order to implement image denoising using (13), we must compute the parameters involved in the estimator, including the standard deviation σ_n and other three independent parameters σ_1 , σ_2 , and a_0 (or a_1) that characterize the Laplacian mixture prior. In this paper, we assume that the standard deviation σ_n is known. If this is not the case, we

use a robust median estimator proposed in [31] to estimate σ_n from the finest decomposition scale of the noisy wavelet coefficients; that is, $\sigma_n = \text{Median}(|Y_{i,j}|)/0.6745$, where Median represents the median absolute deviation operator and $Y_{i,j}$ represents the noisy wavelet coefficients in the finest decomposition level. The other three unknown parameters σ_1 , σ_2 , and a_0 can be estimated with the EM algorithm described in [32]. Here, we only give the corresponding estimation formulas which are as follows:

$$\begin{aligned} a_i^{t+1} &= \frac{1}{N} \sum_{k=1}^N r_{i,k}, \quad k = 0, 1, \dots, N, \\ \sigma_i^{t+1} &= \sqrt{\max\left(\frac{\left(\sum_{k=1}^N r_{i,k} y_k^2\right)}{\sum_{k=1}^N r_{i,k}} - \sigma_n^2, 0\right)}, \end{aligned} \quad (14)$$

where $r_{i,k} = a_i^t p(y_k | \sigma_i^t, \lambda = i) / \sum_{i=0,1} a_i^t p(y_k | \sigma_i^t, \lambda = i)$ is the responsibility factor, t denotes the t th iteration of the algorithm, and N is the number of samples. It can be seen that once the initial values of σ_1 , σ_2 , a_0 , and a_1 are given, all the parameters can be iteratively estimated. For a more detailed discussion on EM algorithm, please refer to [33].

4. Experimental Results

In this section, we present some simulation results obtained by processing several test images using the proposed MMSE estimator in directionlet domain. To assess the performance of our method, we compare our algorithm with other commonly used methods including SureShrink [3], BayesShrink [31], BLS-GSM [17], and K-SVD [27] which is not a wavelet-based method. The directionlet used in our method is constructed using Haar wavelet, and the image is decomposed into three levels. To obtain better results, the specified SAWT($M_\Lambda, 2, 1$) is implemented along 0° , 45° , 90° , and -45° , respectively, and the final denoised image is obtained by averaging the four results. For all other algorithms used for comparison, if not stated otherwise, the used wavelets and the free parameters are set as suggested in the reference papers. Since both the decimated wavelet transform and directionlet transform lack shift invariance, all the wavelet-based and directionlet-based algorithms are conducted in undecimated transform domain to minimize the side effects like pseudo-Gibbs phenomena.

4.1. Results Using Additive Gaussian Noise. Five test images, “Boat” and “Peppers” from <http://sipi.usc.edu/database/database.php?volume=misc> and “Couple,” “Cameraman,” and “Lena” from <http://www.cipr.rpi.edu/resource/stills/misc1.html>, are used for gauging the performance of the proposed denoising algorithm. The five images are all of size 512×512 . In order to obtain the noisy images, we add white Gaussian noise on the original test images. A series of noise levels σ_n ranging from 5 to 100 are selected in our experiments. To quantify the achieved performance improvement, a measure called peak signal-to-noise ratio (PSNR) is computed, which is defined as

$$\text{PSNR}_{\text{dB}} = 10 \log_{10} \frac{255^2}{\text{MSE}}, \quad (15)$$

where MSE is the mean square error between the original and the denoised images.

The PSNR values obtained from all the used methods are listed in Table 2, with the best one highlighted in bold font. Even though, in general, the wavelet-based thresholding methods (SureShrink and BayesShrink) work well with an improvement of several decibels with respect to the noisy images, more sophisticated denoising methods prove definitely superior in all cases. From the table, it can be seen that (1) the PSNR values provided by the proposed method are larger than those of other methods. This indicates that the directionlet transform achieves a more sparse representation for images than other transforms such as wavelet transform and steerable pyramid transform; (2) the K-SVD achieves a comparable performance with our approach and is even better in some cases at low noise levels, but the performance degrades faster than the proposed method at high noise levels. This is because the increase in noise intensity results in a decrease in the number of effective sparse atoms, which further results in the substantial degradation of the denoising performance; (3) the behavior of BLS-GSM is similar to our algorithm but with a high degradation rate of denoising performance when the noise level is high.

For visual comparison, the noisy image ($\sigma_n = 20$) of “Couple” and the denoised images obtained by various methods are illustrated in Figures 3(a)–3(f), respectively. By observing these figures, we find that the visual evaluation of the denoised results is consistent with the quantitative evaluation in terms of PSNR. Specifically, the results with SureShrink and BayesShrink exhibit plenty of visual artifacts and Gibbs phenomena near discontinuities; thus they achieve bad visual effects. Comparing with other algorithms, the proposed method provides fewer artifacts as well as better preservation of edges and other details. Even though the K-SVD and BLS-GSM can obtain satisfactory visual results, the former blurs the edges and other details which can be seen clearly from the floor and the desk, and the latter also introduces some artifacts while removing noise.

In addition to the PSNR and visual comparison, we also conduct a significance test to further justify the improvement. To examine whether there is a significant difference between the performances of the proposed method and the other methods, the Mann-Whitney test was applied on the obtained PSNRs. In this test, additional 25 images from the USC-SIPI Image Database (<http://sipi.usc.edu/database/database.php?volume=misc>) as well as the above 5 test images were used to compute the PSNR values. The differences between the proposed method and each of the other four methods were considered statistically significant if P values are less than 0.05. Statistical analyses were performed using SPSS 17.0. The test results are listed in Table 3.

From the table, it can be seen that all the P values are less than 0.05. Thus, the difference between the proposed method and other methods is significant; that is, the improvement is significant.

4.2. Results Using Speckle Noise. To make the proposed method more convincing, we also conducted the experiments



(a)



(b)



(c)



(d)



(e)



(f)

FIGURE 3: Comparison of the denoised images obtained from different denoising algorithms on test image Couple with $\sigma_n = 20$. (a) Noisy image. (b) Denoised image using SureShrink. (c) Denoised image using BayesShrink. (d) Denoised image using BLS-GSM. (e) Denoised image using K-SVD. (f) Denoised image using the proposed method.

TABLE 2: PSNR comparison (Gaussian noise) of the proposed method with other methods (in dB).

σ_n	5	10	20	30	50	75	100
Boat							
Noisy	34.1172	28.1398	22.1890	18.7310	14.6045	11.7486	10.1119
SureShrink	32.8096	30.6670	27.8854	26.0042	23.1738	20.6964	19.3363
BayesShrink	35.7005	32.5481	29.1581	27.2197	24.9616	22.9441	21.4948
BLS-GSM	36.7329	33.4831	30.2994	28.4360	26.1782	24.5648	23.5208
K-SVD	37.2160	33.5918	30.3368	28.4117	25.8962	24.0504	22.7441
Proposed	37.6733	34.0517	30.7789	29.3584	26.6035	24.8389	23.6182
Peppers							
Noisy	34.1349	28.1728	22.2225	18.7757	14.7374	11.8613	10.2132
SureShrink	33.4827	31.7831	29.1256	26.9193	24.2321	21.1825	19.6603
BayesShrink	36.1450	33.3432	30.1000	28.1650	25.9694	23.2449	21.7568
BLS-GSM	37.1139	33.9436	30.7529	28.6675	26.0842	24.2591	22.5839
K-SVD	37.7215	34.2628	30.8189	29.0493	26.1453	23.6957	21.8496
Proposed	37.1641	34.2156	31.5728	29.8722	27.4658	25.5356	24.1813
Couple							
Noisy	34.1468	28.1501	22.1580	18.7253	14.5570	11.7173	10.1078
SureShrink	32.5729	28.9800	26.6306	24.9866	22.3759	20.2493	19.3832
BayesShrink	36.2527	32.2714	28.6094	26.8294	24.6925	22.8233	21.5051
BLS-GSM	36.9564	32.4654	29.3840	27.2101	24.7357	23.0515	21.5715
K-SVD	37.2685	33.4258	29.6725	27.5475	25.1199	23.4921	22.4298
Proposed	37.0631	33.0186	29.9705	27.7835	25.5059	23.9003	22.8189
Cameraman							
Noisy	31.4893	27.0812	21.9267	18.8280	14.9846	12.2984	10.6453
SureShrink	33.6523	31.0300	27.9941	25.8978	22.8099	19.8985	18.7697
BayesShrink	38.2318	34.1862	30.4611	28.1938	25.0584	22.5210	20.4949
BLS-GSM	40.3564	36.5456	32.9820	29.8171	27.6347	25.0651	22.7713
K-SVD	40.9217	37.1257	33.2066	30.9474	28.4564	25.9567	24.0799
Proposed	40.8315	37.2493	33.4672	31.2645	28.8026	26.7494	25.2453
Lena							
Noisy	30.0189	26.2837	21.3969	18.3500	14.5940	11.9378	10.3692
SureShrink	31.6837	29.7113	27.5191	25.9838	23.6352	21.2366	20.1037
BayesShrink	36.8307	33.4175	30.2196	28.5905	26.2983	24.0695	22.3581
BLS-GSM	38.4729	35.6441	32.5985	29.7101	28.1149	26.5862	25.6407
K-SVD	38.5806	35.4874	32.3522	30.4133	27.7653	25.7731	24.4709
Proposed	38.7493	35.8493	33.1672	31.2645	28.2026	26.4494	25.0453

TABLE 3: The statistical comparison (Gaussian noise $\sigma_n = 20$) between the proposed method and the other methods.

Methods	Proposed	K-SVD	BLS-GSM	BayesShrink	SureShrink
PSNR (mean \pm SD ¹)	31.90 \pm 1.55	31.23 \pm 2.12	31.10 \pm 1.53	29.71 \pm 1.03	27.62 \pm 0.96
<i>P</i> values (proposed versus # ²)	—	0.038	<0.01	<0.01	<0.01

¹SD: standard deviation of the PSNR series; ²#: one of the other four methods (i.e., SureShrink, BayesShrink, BLS-GSM, and K-SVD).

on the speckle noise which is a multiplicative noise with a χ distribution

$$p_A(\mu) = \frac{2L^L e^{2\mu L} e^{-Le^{2\mu}}}{\Gamma(L)}, \quad (16)$$

where $L \geq 1$ and $\Gamma(\cdot)$ denotes the gamma function. It should be noted that the smaller the value of L , the higher the level

of the speckle noise. To adapt the above methods to speckle noise, we must modify some steps involved in each algorithm. The specific modifications are as follows: (1) for the wavelet-based methods including the proposed one, we first take logarithmic operation on the noisy images to convert the multiplicative noise into additive one. Then the denoising algorithm as used in the case of additive noise experiment is performed. Finally, the denoised result is obtained by

TABLE 4: PSNR comparison (speckle noise) of the proposed method with the other methods (in dB).

L	1	6	12	20	32
Couple					
Noisy	15.0949	22.7065	25.7106	27.9160	29.9535
SureShrink	20.0332	25.8413	27.0039	27.6973	28.2485
BayesShrink	19.3898	25.5683	27.0603	28.9848	30.8208
BLS-GSM	21.9560	26.4502	29.4929	30.5549	32.4309
K-SVD	21.3902	26.2782	29.6724	31.1942	33.2195
Proposed	22.4516	27.0145	30.2409	31.6080	33.9895
Cameraman					
Noisy	14.8480	22.4357	25.4539	27.6675	29.7107
SureShrink	20.2833	25.4857	27.3508	29.1526	30.8918
BayesShrink	19.6823	25.1516	27.9765	30.0169	31.3640
BLS-GSM	21.6407	27.0976	28.7225	31.49293	33.1347
K-SVD	21.2120	26.6703	28.6297	32.3785	34.6483
Proposed	22.2975	27.5493	29.7168	32.6638	34.2609
Lena					
Noisy	14.8853	22.4857	25.4966	27.7053	29.7469
SureShrink	20.4569	27.5477	28.6355	29.2568	29.7915
BayesShrink	21.8229	29.8639	31.8440	33.2049	34.4592
BLS-GSM	22.8901	30.5525	32.6039	33.5162	35.9007
K-SVD	22.2918	30.1317	32.1205	33.0830	35.3714
Proposed	23.1152	30.9638	32.9499	33.2915	35.7910

taking an inverse logarithmic operation on the result of the denoising algorithm; (2) since the original K-SVD denoising algorithm is designed for homogeneous white Gaussian noise and the speckle noise cannot be assumed to be white and Gaussian, we use the local noise variance instead of a constant global noise level to train the dictionary. Then, compute the denoised output image using the trained overlapping patches. For more details about the modification on K-SVD, please see [34].

Three images “Couple,” “Cameraman,” and “Lena” are used as the test images in this experiment. For each image, five noise levels, that is, $L = 1, 6, 12, 20, 32$, are considered in this experiment. The obtained PSNR values are listed in Table 4. From the table, it can be seen that the performances of the algorithms are consistent with the case of additive noise except for the K-SVD. This is because when L is small, the speckle noise is strongly non-Gaussian, which degrades the performance of the K-SVD, and the speckle gradually approaches Gaussian distribution with the increase of L . It can also be found that the proposed method obtains larger PSNR values in most cases, which indicates that the proposed method outperforms the other methods. For visual comparison, the speckled Lena image ($L = 12$) and the corresponding denoised results are illustrated in Figure 4. It can be observed that the results are consistent with the case of additive Gaussian noise.

5. Conclusion

The main issues regarding image denoising were addressed in this paper. The directionlet coefficients of the image were modeled as Laplacian mixture distribution with zero mean and at the same time the goodness of fit was analyzed by using a distance between the CDF of the assumed prior and the CDF of the directionlet coefficients. To recover the noise-free image, the MMSE estimator was constructed based on the assumed models. The experimental results confirm that the proposed MMSE approach allows efficient removal of both white Gaussian noise and speckle noise, outperforming the denoising methods published in the literature in terms of PSNR. Furthermore, our method is seen to generate fewer artifacts as well as better preservation of edges and other details. Although the computational cost of the proposed method is higher than the wavelet-based methods due to the directionlet transform producing more subbands, it is still less than the steerable pyramid transform due to the separable filters of directionlets and far less than the K-SVD that needs training highly overcomplete dictionaries.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.



FIGURE 4: Comparison of the denoised images obtained from different denoising algorithms on test image Lena with $L = 12$. (a) Noisy image. (b) Denoised image using SureShrink. (c) Denoised image using BayesShrink. (d) Denoised image using BLS-GSM. (e) Denoised image using K-SVD. (f) Denoised image using the proposed method.

Acknowledgments

This work is supported by National Natural Science Foundation of China (NSFC) for Youth (61402004, 61402003) and NSFC (60872163, 61370110, 61272025) and is also supported by Research Funding of Anhui University for Doctors (J10113190021).

References

- [1] I. Dagher and C. Taleb, "Image denoising using fourth order wiener filter with wavelet quadtree decomposition," *Journal of Electrical and Computer Engineering*, vol. 2014, Article ID 108393, 9 pages, 2014.
- [2] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.
- [3] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Association*, vol. 90, no. 432, pp. 1200–1224, 1995.
- [4] T. R. Downie and B. W. Silverman, "The discrete multiple wavelet transform and thresholding methods," *IEEE Transactions on Signal Processing*, vol. 46, no. 9, pp. 2558–2561, 1998.
- [5] S. G. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1522–1531, 2000.
- [6] P. Bao and L. Zhang, "Noise reduction for magnetic resonance images via adaptive multiscale products thresholding," *IEEE Transactions on Medical Imaging*, vol. 22, no. 9, pp. 1089–1099, 2003.
- [7] R. A. Peters II, "New algorithm for image noise reduction using mathematical morphology," *IEEE Transactions on Image Processing*, vol. 4, no. 5, pp. 554–568, 1995.
- [8] A. Restrepo and A. C. Bovik, "Adaptive trimmed mean filters for image restoration," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 8, pp. 1326–1337, 1988.
- [9] O. A. Ojo and T. G. Kwaaitaal-Spassova, "An algorithm for integrated noise reduction and sharpness enhancement," *IEEE Transactions on Consumer Electronics*, vol. 46, no. 3, pp. 474–480, 2000.
- [10] M. Protter, I. Yavneh, and M. Elad, "Closed-form MMSE estimation for signal denoising under sparse representation modeling over a unitary dictionary," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3471–3484, 2010.
- [11] M. I. H. Bhuiyan, M. O. Ahmad, and M. N. S. Swamy, "Spatially adaptive wavelet-based method using the cauchy prior for denoising the SAR images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 4, pp. 500–507, 2007.
- [12] S. M. M. Rahman, M. O. Ahmad, and M. N. S. Swamy, "Video denoising based on inter-frame statistical modeling of wavelet coefficients," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 2, pp. 187–198, 2007.
- [13] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.
- [14] Z. Zeng and I. Cumming, "Bayesian speckle noise reduction using the discrete wavelet transform," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS '98)*, pp. 7–9, July 1998.
- [15] P. Moulin and J. Liu, "Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors," *IEEE Transactions on Information Theory*, vol. 45, no. 3, pp. 909–919, 1999.
- [16] J. K. Romberg, H. Cho, and R. G. Baraniuk, "Bayesian tree-structured image modeling using wavelet-domain hidden Markov models," *IEEE Transactions on Image Processing*, vol. 10, no. 7, pp. 1056–1068, 2001.
- [17] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1338–1351, 2003.
- [18] A. Achim, A. Bezerianos, and P. Tsakalides, "Novel Bayesian multiscale method for speckle removal in medical ultrasound images," *IEEE Transactions on Medical Imaging*, vol. 20, no. 8, pp. 772–783, 2001.
- [19] J. M. Fadili and L. Boubchir, "Analytical form for a Bayesian wavelet estimator of images using the BESsel K form densities," *IEEE Transactions on Image Processing*, vol. 14, no. 2, pp. 231–240, 2005.
- [20] L. Boubchir and J. M. Fadili, "A closed-form nonparametric Bayesian estimator in the wavelet domain of images using an approximate α -stable prior," *Pattern Recognition Letters*, vol. 27, no. 12, pp. 1370–1382, 2006.
- [21] E. J. Candès and D. L. Donoho, "New tight frames of curvelets and optimal representations of objects with piecewise c^2 singularities," *Communications on Pure and Applied Mathematics*, vol. 57, no. 2, pp. 219–266, 2004.
- [22] 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.
- [23] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091–2106, 2005.
- [24] X. Yang, W. Li, and L. Jiao, "Image denoising based on second generation bandelets and multi-level thresholding," in *Proceedings of the 6th World Congress on Intelligent Control and Automation (WCICA '06)*, vol. 2, pp. 9509–9513, June 2006.
- [25] R. Eslami and H. Radha, "Translation-invariant contourlet transform and its application to image denoising," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3362–3374, 2006.
- [26] X. Zhu and F. Kui, "Image recovery algorithm based on learned dictionary," *Mathematical Problems in Engineering*, vol. 2014, Article ID 964835, 6 pages, 2014.
- [27] 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.
- [28] 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.
- [29] M. K. Mihçak, I. Kozintsev, K. Ramchandran, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Letters*, vol. 6, no. 12, pp. 300–303, 1999.
- [30] H. Xie, L. E. Pierce, and F. T. Ulaby, "Statistical properties of logarithmically transformed speckle," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 3, pp. 721–727, 2002.
- [31] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532–1546, 2000.

- [32] H. Xie, L. E. Pierce, and F. T. Ulaby, "SAR speckle reduction using wavelet denoising and Markov random field modeling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 10, pp. 2196–2212, 2002.
- [33] R. A. Redner and H. F. Walker, "Mixture densities, maximum likelihood and the EM algorithm," *SIAM Review*, vol. 26, no. 2, pp. 195–239, 1984.
- [34] D. Bartuschat, A. Borsdorf, H. Kostler, R. Rubinstein, and M. Sturmer, *A Parallel K-SVD Implementation for CT Image Denoising*, Department of Computer Science, 2009.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

