



Image Denoising method based on NSCT bivariate model and Variational Bayes threshold estimation

Wang Deyan^{1,2} · Xiao Yin² · Gao Ya²

Received: 5 June 2018 / Revised: 3 August 2018 / Accepted: 12 September 2018 /

Published online: 25 September 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

In order to reduce the Gaussian noise introduced during image generation, this paper presents an image denoising algorithm based on variational Bayes (V-Bayes) estimation and nonsampled contourlet transform (NSCT) bivariate model. First, the proposed algorithm uses the NSCT's advantages of translation-invariant and multidirection-selectivity, exploits the intra-scale and inter-scale correlations of NSCT coefficients. Then, the corresponding nonlinear bivariate threshold function of the model is deduced by using V-Bayes estimation theory. Finally, the noise-reduced coefficients are inverse-transformed by NSCT to obtain denoised image. The simulation results show that the denoised image has obvious improvement in subjective visual effects and performance indicators, and effectively preserves the details and texture information in the original image.

Keywords Image denoising · Nonsampled Contourlet transform · Bivariate model · Variation Bayes · Threshold estimation

1 Introduction

During the acquisition and transmission process, images are often contaminated by various noises, such as Gaussian white noise in optical images [1, 6, 12]. The presence of noise reduces the resolution of the original image and seriously affects the subsequent classification and recognition of the target. Therefore, image denoising has become an important method for image preprocessing. Image denoising is to improve image quality and highlight the features of the image itself [17]. As Donoho et al. proposed the concept of nonlinear wavelet threshold denoising [21], the denoising method based on wavelet transform has been widely used. In recent years, the research focus for wavelet denoising is the study of statistical model of image

✉ Wang Deyan
wangdy@wxit.edu.cn

¹ Management & Science University, Shah Alam, Malaysia

² School of Internet of Things Technology, Wuxi Institute of Technology, Jiangsu, Wuxi, China

wavelet coefficients, which is to accurately model the non-Gaussian image wavelet coefficients with certain correlations [15]. The basic idea of this type of algorithm is to use the statistical model as a priori probability model of the wavelet coefficients, and then use this prior information to estimate the wavelet coefficients of the original image in the Bayes framework. For example, an image denoising method combined with wavelet domain and SUREShrink threshold estimation (WT + SUREShrink) was proposed in [14]. This method is based on SURE (Stein's Unbiased Risk Estimation) criterion, which is an unbiased estimate of the mean square error criterion, and the SURE threshold approaches the ideal threshold. An image denoising method combined with the wavelet domain and BayesShrink threshold estimation (WT + BayesShrink) was proposed in [7], which is based on the assumption that no-noise image wavelet coefficients obey the generalized Gaussian distribution.

However, with the increasing manifestation of the limitations of wavelet transforms, such as the non-sparseness of high-dimensional time coefficients, lack of multi-directional selectivity, multi-scale geometric analysis methods have emerged. Contourlet transform [4] is a powerful and versatile two-dimensional signal transformation tool. Compared with wavelet transform, it has better multi-resolution, multi-directional performance, and can accurately capture the intrinsic geometry of the image. For example, an image denoising method combined with Contourlet domain and BayesShrink threshold estimation (CT + BayesShrink) was proposed in [10]. However, because Contourlet lacks translation invariance, on this basis, Cunha et al. proposed the nonsubsampling contourlet transform (NSCT) [2], which uses non-subsampling Laplacian pyramid (LP) and directional filter bank (DFB) to construct the decomposition structure. This process can avoid the downsampling process, so as to make the transformation have translation invariance, and further improve the performance of the Contourlet transform in the field of image denoising [3, 22].

In addition, there is a problem in the NSCT coefficient variance estimation algorithm based on the traditional Bayes principle, that is, how to select an appropriate prior probability coefficient model of the original image and the noise coefficient. To solve this problem, the Variational Bayes (V-Bayes) method can be used. The basic idea of V-Bayes is to approximate the true posterior probability distribution with a more easily approximated distribution. It realizes the estimation by minimizing the Kullback-Leibler (KL) divergence between the approximate distribution and the real posterior probability distribution [5, 9].

For this reason, this paper combines NSCT bivariate model and V-Bayes estimation, and proposes an image denoising algorithm. The simulation analysis was performed on the standard test images Lena, Barbara, and Peppers and compared with the existing classical methods. The results show that the proposed algorithm can effectively remove the noise in the image and obtain the highest PSNR value and noise suppression capability.

2 Nonsubsampling Contourlet transforms and bivariate model

2.1 Contourlet transform

Contourlet transform is implemented by Laplacian Pyramid (LP) decomposition and Directional Filter Bank (DFB). The LP decomposition decomposes the original image into low-frequency sub-band and high-frequency sub-band, where the low-frequency sub-band is generated from the original image by two-dimensional low-pass analysis filtering and downsampling. The low-frequency sub-band is subjected to up-sampling and two-dimensional low-

pass synthesis filtering to form the same low-frequency components as the original image size. The original image is subtracted from the low-frequency components to form another high-frequency sub-band. The high-frequency sub-band is further decomposed into 2^{l_j} sub-bands by a directional filter bank (for different scales j and l_j can take different values). Repeating the above process for low-frequency sub-band can achieve multi-scale multi-direction decomposition [13, 18]. Figure 1 is a schematic diagram of the decomposition of the high-frequency sub-band by DFB, and Fig.2 shows the sub-band of the Lena image after 2-layer Contourlet transform.

The Contourlet transform process is expressed as [20]: (1) Using LP to perform the first-order multiscale decomposition of the original image to obtain the first-order low-pass sub-band and the first-order band-pass sub-band; (2) Using DFB to combine singular points that have the same direction and are not continuous into a new coefficient, and the first-order band-pass sub-band is decomposed in multiple directions to obtain a directional sub-band; (3) Repeat steps 1 and 2 for the directional sub-bands to obtain the low-pass sub-bands at different scales; (4) Contourlet reconstruction transformation is performed on LP and DFB to obtain a transformed image matrix.

2.2 Nonsampled Contourlet transform (NSCT)

The contourlet transform can sparsely represent the image in an optimal way, and can efficiently capture curved and oriented geometrical structures in images. Because down-sampling is performed in both LP and DFB, the redundancy of the Contourlet coefficients of the image is greatly reduced (redundancy is only 4/3) [23]. As a result, this transformation lacks translation invariance. Therefore, if the Contourlet transform is used directly for image denoising, there will be a noticeable ringing effect.

In order to solve the above problems, this paper uses the NSCT. The NSCT is a non-orthogonal transform that discards the down-sampling operations in the Contourlet transform, but combines the nonsampled pyramid (NSP) and nonsampled directional filterbanks filter banks (NSDFB). After the transformation, the size of the sub-bands in each direction on each scale is the same as that of the original image, and its redundancy reaches $1 + \sum_{j=1}^J 2^{l_j}$ (J

Fig. 1 Decomposition of high-frequency sub-band by DFB

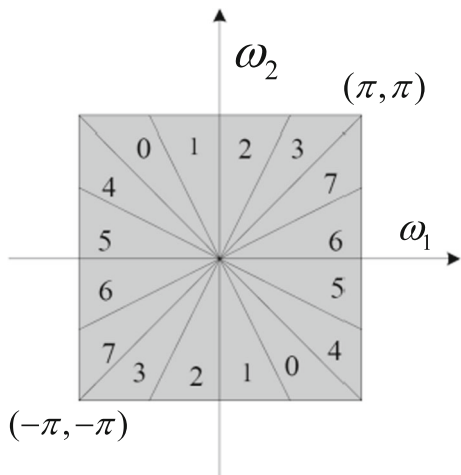
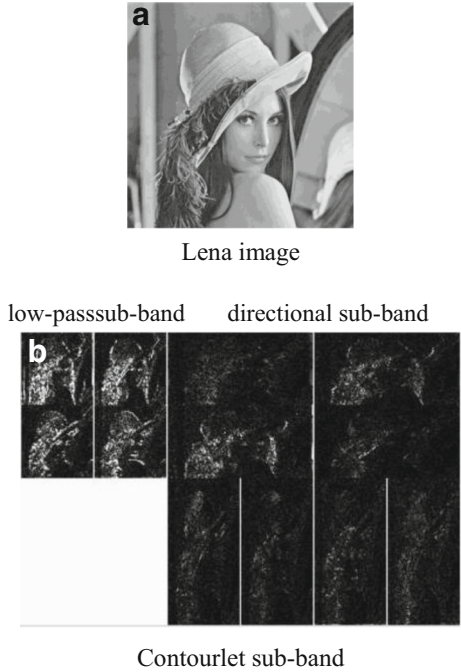


Fig. 2 The sub-band of the Lena image after 2-layer Contourlet transform



denotes the decomposition layer number of NSP). Therefore, the improvement of the redundancy of the coefficients makes the transform have a translation invariance, which is beneficial to the effect of image denoising [16]. Fig.3 is a schematic diagram of a 2-layer nonsubsampling Contourlet transform.

2.3 Bivariate model

A large number of studies have shown that the image wavelet transform coefficients are not independent, but there is a certain degree of correlation, including intra-scale correlation and inter-scale correlation. In addition, the distribution of the wavelet coefficients of each sub-band

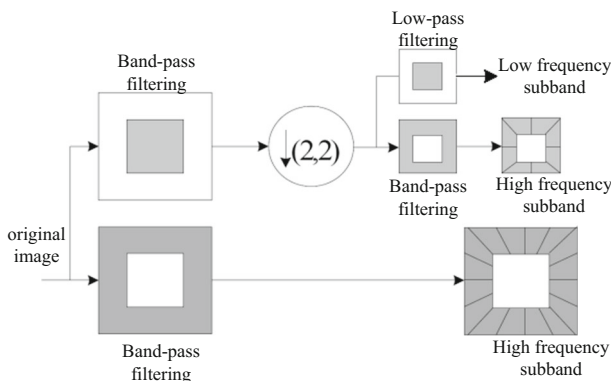


Fig. 3 NSCT decomposition schematic

is not Gaussian, but forms a curve with high peaks and long tails. Based on the above two properties of image wavelet coefficients, various statistical models have been proposed. Among them, the bivariate model fully exploited the inter-scale correlation of image wavelet coefficients [11], used the non-Gaussian density function to model the distribution of “parent” and “child” coefficients, and used Bayes statistical theory to obtain the analytical expression of the maximum posterior estimate. Therefore, it has achieved good results in the application of image denoising.

Because the coefficients in the NSCT transform domain have similar distribution characteristics with the wavelet coefficients, that is, the decomposition coefficients between the adjacent scales have strong correlation. For this reason, the processing idea of the bivariate model in the wavelet domain can be used to extend it into the NSCT transform domain [8].

Considering the relationship between the current coefficient and its parent coefficient, the image observation signal can be expressed as

$$y = \omega + n \tag{1}$$

where $\omega = [\omega_1, \omega_2]$, ω_1 is the NSCT coefficient of the original image scale, and ω_2 is the NSCT coefficient with the same spatial position in the previous scale as the parent coefficient of ω_1 . $y = [y_1, y_2]$ and $n = [n_1, n_2]$ represent the vectors composed of the NSCT decomposition coefficient of observation image and the noise, respectively.

In the NSCT domain, the objective of image denoising is to obtain ω 's estimate $\hat{\omega}$ from y . The maximum posterior probability (MAP) can be used to estimate $\hat{\omega}(y) = \arg \max_{\omega} p_{\omega|y}(\omega|y)$, and then use Bayes' rule to obtain

$$\hat{\omega}(y) = \arg \max_{\omega} [\log(p_n(y-\omega)) + \log(p_{\omega}(\omega))] \tag{2}$$

From the above equation, we must first know the probability distribution of the NSCT decomposition coefficient of noise. Here assume that the noise density function p_n obeys a Gaussian distribution with a mean of 0 and a variance of σ_n^2 :

$$p_n(n) = \frac{1}{2\pi\sigma_n^2} \exp\left(-\frac{n_1^2 + n_2^2}{2\sigma_n^2}\right) \tag{3}$$

According to the distribution of NSCT decomposition coefficients, the model proposed in [19] was modified to fit the distribution of ω by the probability distribution represented by the following formula:

$$p_{\omega}(\omega) = \frac{3}{2\pi\sigma^2} \exp\left(-\frac{\sqrt{3}}{\sigma} \sqrt{\omega_1^2 + \omega_2^2}\right) \tag{4}$$

where σ represents the model edge variance of each NSCT coefficient. Since $p_{\omega}(\omega)$ follows the properties of the convex function and can be derived, so that $f(\omega) = \lg(p_{\omega}(\omega))$, then (2) is equivalent to solving the system of equations:

$$\begin{cases} \frac{y_1 - \hat{\omega}_1}{\sigma_n^2} + f_1(\hat{\omega}) = 0 \\ \frac{y_2 - \hat{\omega}_2}{\sigma_n^2} + f_2(\hat{\omega}) = 0 \end{cases} \tag{5}$$

where f_i is the partial derivative of $f(\omega)$ to ω_i . Then, the applicable bivariate threshold function in the NSCT transform domain can be solved:

$$\hat{\omega}_1 = \frac{\left(\sqrt{y_1^2 + y_2^2} - \frac{\sqrt{3}\sigma_n}{\sigma}\right)_+}{\sqrt{y_1^2 + y_2^2}} y_1 \tag{6}$$

where the symbol $(f)_+$ represents $(f)_+ = 0$ when $f < 0$, otherwise $(f)_+ = f$. Equation (6) is the MAP estimate of the current coefficient ω_1 in the bivariate model and is called the bivariate atrophy function. It can be seen that if we want to obtain $\hat{\omega}_1$, we must estimate the NSCT coefficient variance σ_n^2 of the noise first. In this paper, the variation Bayes estimation method is used to estimate the variance.

3 Variational Bayes threshold estimation

3.1 Bayes estimation

For the estimation of the variance σ_n^2 of the NSCT coefficients, a Bayes estimation method can be used. Also consider the relationship between the original image and noise NSCT decomposition coefficients

$$y = \omega + n \tag{7}$$

In the process of image blurring, the edge area is affected more than flat areas and often contains more information. Therefore, we will perform denoising in the image gradient domain. Let $\nabla\omega$ and ∇y denote the gradient of the NSCT coefficients of the original image and the noisy image, respectively, so the eq. (7) can be converted as

$$\nabla y = \nabla\omega + n \tag{8}$$

Given the gradient ∇y of the noisy image, according to the Bayes principle, the posterior probability of $\nabla\omega$ can be expressed as

$$p(\nabla\omega | \nabla y) = \frac{p(\nabla y | \nabla\omega)p(\nabla\omega)}{p(\nabla y)} \tag{9}$$

According to the (9), one of the key problems of the NSCT coefficient variance estimation algorithm based on the Bayes principle lies in how to select an appropriate noise function and prior probability coefficient model of the original image coefficient and noise coefficient.

3.2 Variational Bayes estimation (V-Bayes)

In order to solve the above problem of traditional Bayes estimation, the variational Bayes method provides an effective solution to the problem. The basic idea is to

approximate the true posterior probability distribution $p(\nabla\omega|\nabla y)$ with an easy-to-handle approximation distribution $q(\nabla\omega)$. The KL divergence can measure the distance between two distributions, so estimation can be achieved by minimizing the KL divergence between the approximate distribution and the real posterior probability distribution.

$$\begin{aligned}
 &KL\left\{q(\nabla\omega, \sigma^2)\|p(\nabla\omega|\nabla y)\right\} \\
 &= \int q(\nabla\omega, \sigma^2) \ln \frac{q(\nabla\omega, \sigma^2)}{p(\nabla\omega|\nabla y)} d\nabla\omega + \ln p(\nabla y) \geq 0
 \end{aligned}
 \tag{10}$$

The above equation holds if and only if $q(\nabla\omega, \sigma^2) = p(\nabla\omega|\nabla y)$.

Note that $p(\nabla y)$ is always a constant throughout the estimation process, so a cost function C_{KL} can be defined to get the optimal value of the approximate distribution:

$$\begin{aligned}
 C_{KL} &= KL\left\{q(\nabla\omega, \sigma^2)\|p(\nabla\omega|\nabla y)\right\} - \ln p(\nabla y) \\
 &= \int q(\nabla\omega) \ln \frac{q(\nabla\omega)}{p(\nabla\omega)} d\nabla\omega \\
 &\quad + \int q(-\sigma^2) \ln \frac{q(-\sigma^2)}{p(-\sigma^2)} d(-\sigma^2)
 \end{aligned}
 \tag{11}$$

The variational Bayes estimation minimizes the cost function through an iterative approach (that is, variation Bayes expectation maximization theorem), thereby estimating the noise function.

4 Image denoising steps

From the previous discussion, it can be seen that the NSCT has the advantages of anisotropy, multi-directional selectivity and translation invariance, while the bivariate model can fully exploit the inter-scale correlation of image wavelet coefficients. At the same time, we have reason to believe that the transform coefficients and the wavelet coefficients of the image Contourlet and the NSCT have very similar properties in terms of correlation and distribution characteristics. Therefore, this paper combines the two and proposes an image denoising algorithm based on NSCT and bivariate model. The detailed steps of the algorithm are as follows:

- Step 1: NSCT is performed on the noisy image, in which the NSP decomposition layer number and the number of NSDFB decomposed direction sub-bands on each scale can be set.
- Step 2: Estimate the NSCT coefficient variance $\sigma_n^2(k)$ of the noise.
 - (2.1) Perform NSCT on noisy images and estimate noise standard deviation σ_n using V-Bayes estimation method.
 - (2.2) Generate a Gaussian white noise image with the same size as the original image with an average value of 0 and a variance of σ_n^2 .
 - (2.3) Perform NSCT on this noise image, and square the transformed coefficients to obtain the NSCT coefficient variance $\sigma_n^2(k)$ of noise.

- (2.4) Go to step (2.2), repeat the above method several times (the number of repetitions in this experiment is 10), take the average, and get the final $\sigma_n^2(k)$.
- Step 3: For each high frequency direction sub-band, use local adaptive method to estimate the model edge variance $\sigma^2(k)$ for each coefficient, ie, $\sigma^2(k) = \max(\sigma_{v_1}^2(k) - \sigma_n^2(k), 0)$, where $\sigma_{v_1}^2(k) = \frac{1}{M} \cdot \sum_{v_1(k) \in N(k)} y_1^2(k)$. Here, $N(k)$ represents a square window centered on the current NSCT coefficient $y_1(k)$, and M is the number of coefficients in the window.
- Step 4: Substitute $\sigma_n^2(k)$ and $\sigma^2(k)$ into (6) to obtain the bivariate threshold ω_1 .
- Step 5: Reconstruct the denoised Contourlet coefficient matrix and perform an inverse NSCT to obtain a denoised image.

5 Simulation and analysis

5.1 Simulation settings

On the PC with Intel i5–7500 CPU and 8G memory, various denoising algorithms are executed through Matlab R2014 software. Three 512×512 standard test images Lena, Barbara, and Peppers were selected for the simulation experiment. Gaussian white noise with zero mean and different levels of standard deviation was added. The noise standard deviations σ_n were 10, 20, 30, 40, 50. For system parameters, appropriate parameters were set according to the analysis of the result of previous experiments. In this experiment, we set the number of NSCT layers to 5, and the number of directional sub-bands was 4, 8, 8, 16, and 16 in order from coarse to fine scale. The window coefficient M is 13×13 .

In order to evaluate the effectiveness of the NSCT domain+V-BayesShrink threshold estimation denoising algorithm (NSCT+V-BayesShrink) proposed in this paper, we compared three typical algorithms in the current denoising field, these are wavelet domain + SUREShrink threshold estimation method (WT + SUREShrink) [14], the wavelet domain + BayesShrink threshold estimation method (WT + BayesShrink) [7], the Contourlet domain + BayesShrink threshold estimation method (CT + BayesShrink) [10].

5.2 Performance parameters

For the real images I , a denoised image S is obtained after denoising. In this paper, we select the following indicators for judging the denoising effect of an image:

1) Peak signal-to-noise ratio (PSNR), which represents the difference between a denoised image and a real image.

$$PSNR = 20 \log \left(\frac{255}{\sqrt{MSE}} \right) \quad (12)$$

where, MSE is the mean square error value of the real image I and the denoised image S , expressed as

$$MSE = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (I(i, j) - S(i, j))^2 \quad (13)$$

2) The noise suppression parameter (ρ), the greater the value, the stronger the suppression of noise.

$$\rho = \frac{\Gamma(I-\bar{I}, S-\bar{S})}{\sqrt{\Gamma(I-\bar{I}, I-\bar{I})} \sqrt{\Gamma(S-\bar{S}, S-\bar{S})}} \quad (14)$$

where, the symbol “ $\bar{\cdot}$ ” means taking an average of the image. The definition of operator Γ is $\Gamma(A, B) = \sum A(i, j)B(i, j)$.

In addition, visually judging the degree of noise pollution of the image, the important edge and detail information retention in the image is also an important criterion for the performance of denoising algorithm.

5.3 Performance comparison

Figure 4 shows the standard Lena image and Gaussian noise image with standard deviation of $\sigma_n = 30$. Figure 5 shows the denoised Lena image processed by various denoising methods. Figure 6 shows the standard Barbara image and Gaussian noise image with standard deviation of $\sigma_n = 40$. Figure 7 shows the denoised Barbara image processed by various denoising methods. From a visual point of view, the NSCT+V-BayesShrink algorithm well preserved the edge details of the image, effectively removed the noise, and the denoised image content became clearer. The parts such as brims, hair, eyes, and eyebrows that are rich in edge details are clearer than the corresponding parts in other denoised images.

Table 1 shows the PSNR of the noise images obtained by adding the noise with the standard deviations σ_n of 10, 20, 30, 40, and 50 to the Lena, Barbara, and Peppers images, respectively. It can be seen that after adding noise, the PSNR of the image is significantly reduced, and the larger the noise standard deviation, the lower the PSNR.

For each image, 10 tests were repeated at each noise level, and calculate the average PSNR. The experimental results are shown in Figs. 8, 9 and 10. Table 2 shows the average PSNR of the various algorithms over the three images. In addition, the average ρ values of various algorithms on three images under different noise levels are also calculated, as shown in Table 3.

From Fig. 8 to Fig. 10, it can be seen that as the noise level increases, the PSNR of the denoised image of various methods will also decrease. In addition, for the three images, the denoising performance on the Barbara image is slightly worse than that of the Lena and Peppers images. This is because the Barbara image has more directional texture information, which makes it more difficult to denoise.

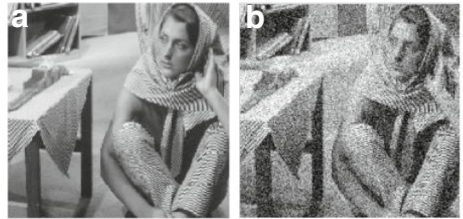
Fig. 4 Lena image, (a) original image, (b) noise image ($\sigma_n = 30$)



Fig. 5 Denoised Lena image, (a) WT + SUREShrink, (b) WT + BayesShrink, (c)CT + BayesShrink, (d)NSCT+V-BayesShrink



Fig. 6 Barbara image, (a) original image, (b) noise image ($\sigma_n = 40$)



The above experimental results show that the NSCT+V-BayesShrink algorithm has good performance both in terms of objective indicators and subjective effects. From Tables 2 and 3, it can be seen that in all simulation experiments, the PSNR and the value of NSCT+V-BayesShrink algorithm are the highest. With different noise levels, the PSNR is about

Fig. 7 Denoised Barbara image, (a) WT + SUREShrink, (b) WT + BayesShrink, (c)CT + BayesShrink, (d)NSCT+V-BayesShrink

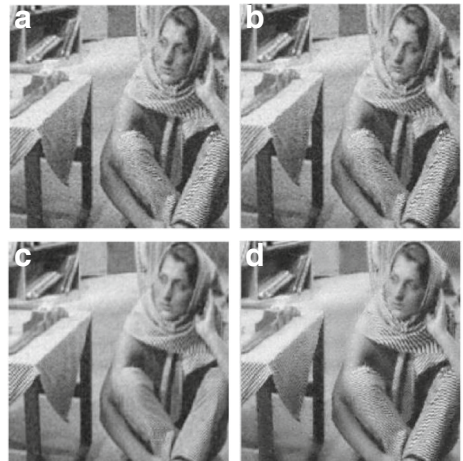


Table 1 PSNR (dB) of the noise-added image on three images

Test image	Noise standard deviation				
	10	20	30	40	50
Lena	26.17	23.43	18.75	16.17	13.75
Barbara	26.20	23.46	18.71	16.10	13.80
Peppers	26.18	23.42	18.79	16.14	13.74

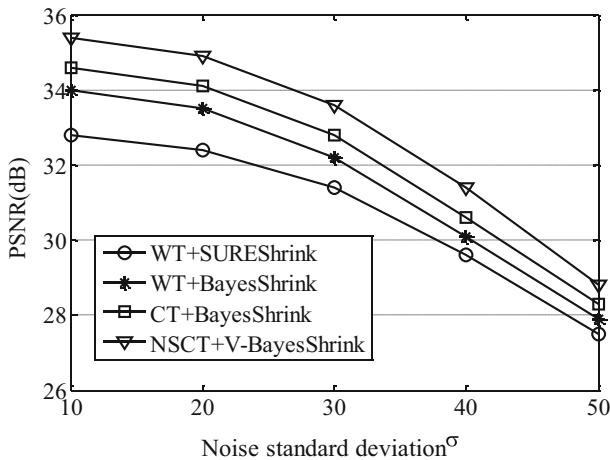


Fig. 8 Comparison of denoising performance on Lena images

0.7 dB higher than CT + BayesShrink algorithm, and the average value is about 0.01 higher than CT + BayesShrink algorithm. The reasons for this are as follows:

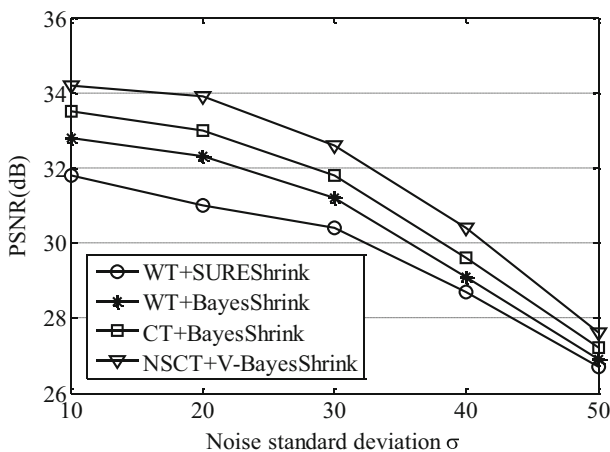


Fig. 9 Comparison of denoising performance on Barbara images

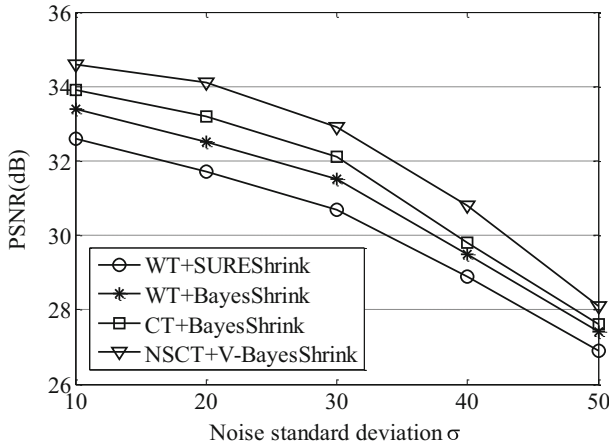


Fig. 10 Comparison of denoising performance on Peppers images

- 1) Use of NSCT. NSCT has many advantages such as anisotropy, translation invariance, multi-direction selectivity and so on, which make the image denoising very advantageous. As a result, the NSCT+V-BayesShrink algorithm achieves better results than the WT-based algorithm, effectively maintains the directional information in the original image, and largely avoids the appearance of the ringing effect. Although the WT-BayesShrink algorithm is also based on the bivariate model, the WT transform only has approximate translation invariance and directional selectivity of 6 directions on each scale, so this algorithm is not as good as NSCT-based algorithm.
- 2) Selection of bivariate models. Bivariate model fully exploits the inter-scale correlation of image coefficients, and has the atrophic function expression associated with the “father” coefficient. In addition, NSCT+V-BayesShrink algorithm uses a local adaptive method when estimating the model edge variance of the current coefficient, that is, taking into account the influence of the surrounding coefficients on the current coefficient, which is equivalent to the use of intra-scale correlation of the coefficients. In other words, the algorithm considers both the intra-scale and inter-scale correlations of the coefficients. Compared with the CT-BayesShrink based on the generalized Gaussian model which only considers intra-scale correlations, the results of NSCT+V-BayesShrink are more reasonable than CT-BayesShrink.
- 3) Use of V-Bayes. V-Bayes estimation is used to estimate the variance of the noise coefficient in NSCT+V-BayesShrink. Compared with the median estimation method proposed by Donoho in the wavelet domain, V-Bayes estimation method is more suitable

Table 2 Average PSNR (dB) of denoised images obtained by various algorithms

Denoising method	Noise standard deviation					Average
	10	20	30	40	50	
WT + SUREShrink	32.42	31.72	30.84	29.06	27.03	30.21
WT + BayesShrink	33.41	32.76	31.62	29.55	27.41	30.95
CT + BayesShrink	34.05	33.44	32.24	30.02	27.72	31.49
NSCT+V-BayesShrink	34.73	34.31	33.05	30.87	28.16	32.22

Table 3 Average ρ of denoised images obtained by various algorithms

Denoising method	Noise standard deviation					Average
	10	20	30	40	50	
WT + SUREShrink	0.9444	0.9432	0.9447	0.9428	0.9436	0.9437
WT + BayesShrink	0.9565	0.9571	0.9554	0.9563	0.9578	0.9666
CT + BayesShrink	0.9663	0.9658	0.9683	0.9666	0.9685	0.9679
NSCT+V-BayesShrink	0.9764	0.9773	0.9765	0.9767	0.9774	0.9769

for the Contourlet domain, and the estimated noise coefficient variance is also more accurate.

6 Conclusion

This paper combines NSCT bivariate model with V-Bayes estimation and proposes an image denoising algorithm. Using the advantages of translation invariance and multi-direction selectivity of the NSCT bivariate model, the V-Bayes valuation theory is used to derive the threshold function. Finally, the denoised coefficients are inverse-transformed by NSCT to obtain denoised images. The simulation analysis on the standard test image shows that the denoised image obtained by the algorithm has a significant improvement in the subjective visual effects, PSNR and noise suppression performance ρ .

Acknowledgements My paper is supported by the Project supported by the National Natural Science Foundation of China (61502204).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Chang SG, Yu B, Vetterli M (2000) Adaptive wavelet thresholding for image denoising and compression[J]. *IEEE Trans Image Process* 9(9):1532–1546
2. Da CA, Zhou J, Do MN (2006) The nonsubsampling contourlet transform: theory, design, and applications.[J]. *IEEE Trans Image Process* 15(10):3089–3101
3. Devanna H, Kumar GAES (2011) Performance analysis of modified nonsubsampling Contourlet transform for image Denoising[J]. *Res J Appl Sci Eng Technol* 3(7):589–595
4. Eslami R, Radha H (2003) The contourlet transform for image denoising using cycle spinning[C]//i. *Signals Syst Comput 2004 Conf Rec Thirty-Seventh Asilomar Conf IEEE* : 1982–1986
5. Fraysse A, Rodet T (2014) A measure-theoretic Variational Bayes algorithm for large dimensional problems[J]. *Siam J Imag Sci* 7(4):212–217
6. Gu S, Zhang L, Zuo W et al. (2014) Weighted nuclear norm minimization with application to image Denoising[C]//i. *Comput Vision Pattern Recogn IEEE* : 2862–2869
7. Hashemi SM, Beheshti S (2011) Adaptive image denoising by rigorous BayesShrink thresholding[C]//i. *Stat Signal Process Works IEEE* : 713–716
8. Jia J, Chen L (2009) Using anisotropic bivariate threshold function for image denoising in NSCT domain[J]. *J Electron Inform Technol* 31(3):349–356
9. Li X (2008) Variational Bayes image processing on stochastic factor graphs[C]//i. *IEEE Int Conf Image Process IEEE* : 1748–1751

10. Li DM, Zhang LJ, Yang JH et al (2016) Research on wavelet-based contourlet transform algorithm for adaptive optics image denoising[J]. *Optik - Int J Light Electron Optics* 127(12):5029–5034
11. Lin XH, Cheng JL, University X et al (2008) Image segmentation based on bivariate models in NSCT domain[J]. *J Image Graph* 13(10):1841–1844
12. Luisier F, Blu T, Unser M (2011) Image Denoising in mixed Poisson–Gaussian noise[J]. *IEEE Tran Image Process Publ IEEE Signal Process Soc* 20(3):696–708
13. Min D, Zhang J, Ma Y (2015) Image denoising via bivariate shrinkage function based on a new structure of dual contourlet transform[J]. *Signal Process* 109:25–37
14. Mohideen SK, Perumal SA, Krishnan N et al (2010) A novel approach to image Denoising by combining Neighshrink and Sureshrink in wavelet domain[J]. *Digit Image Process* 21(3):121–126
15. Naimi H, Adamou-Mitiche ABH, Mitiche L (2015) Medical image denoising using dual tree complex thresholding wavelet transform and wiener filter[J]. *J King Saud Univ Comput Inform Sci* 27(1):40–45
16. Ou Y, Hong B (2013) Image De-noising algorithm using adaptive Bayes threshold by subband based on nonsubsampling Contourlet transform[J]. *Third Int Conf Intell Syst Design Eng Appl IEEE* 2013:832–835
17. Rasti B, Sveinsson JR, Ulfarsson MO et al (2014) Hyperspectral image Denoising using first order spectral roughness penalty in wavelet domain[J]. *IEEE J Select Top Appl Earth Observ Remote Sens* 7(6):2458–2467
18. Sadreazami H, Ahmad MO, Swamy MNS (2014) Contourlet domain image denoising using normal inverse gaussian distribution[C]// *Elect Comput Eng IEEE* : 1–4
19. Sendur L, Selesnick IW (2002) Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency[J]. *IEEE Trans Signal Process* 50(11):2744–2756
20. Shahdoosti HR, Hazavei SM (2017) Image denoising in dual contourlet domain using hidden Markov tree models[J]. *Digit Sign Process* 67(1):17–29
21. Starck JL, Candes EJ, Donoho DL (2002) The curvelet transform for image denoising[J]. *IEEE Trans Image Process Publ IEEE Sign Process Soc* 11(6):670–684
22. Wang XY, Yang HY, Zhang Y et al (2013) Image denoising using SVM classification in nonsubsampling contourlet transform domain[J]. *Inform Sci Int J* 246(14):155–176
23. Zhang S, Moloney C (2008) The nonredundant contourlet transform (NRCT): a multiresolution and multidirection image representation[C]// *Elect Comput Eng IEEE* : 1323–1326



Wang Deyan (1977), female, Lanzhou, Gansu, Master, associate professor, the main research direction is particle swarm algorithm, machine learning algorithm, etc.



Xiao Ying (1979), male, Wuxi, Jiangsu, Master, associate professor, the main research direction for the network technology, cloud computing, etc.



Gao Ya (1987), female, Yancheng, Jiangsu, Ph.D., major research direction for high-performance networks and exchanges, etc.