

## A Spatially Adaptive Denoising Algorithm Based on Curvelet Transform

Peng Feng, Feng Yang

Key Laboratory of Opto-electronics Technology &  
System Ministry of Education  
Chongqing University  
Chongqing, China, 400044

Biao Wei, Deling Mi

Key Laboratory of Opto-electronics Technology &  
System Ministry of Education  
Chongqing University  
Chongqing, China, 400044

**Abstract**—A new approach for image denoising based on the Curvelet transform is presented in this paper. The existing theory for Curvelet and Ridgelet suggests that these new approaches can outperform wavelet method in certain image processing including image denoising, edge detection and image enhancement. However in original digital Curvelet transform it uses a simple hard-thresholding rule for filtering the noisy Curvelet coefficient which of course causes some problems such as killing too many signal Curvelet coefficients that might contain useful image information. Here we introduce BayesShrink denoising scheme into Curvelet domain that is an adaptive, data-driven thresholding approach for image denoising, namely CurShrink. The threshold is derived in a Bayesian framework, and the prior used on the Curvelet coefficients is the generalized Gaussian distribution (GGD) widely used in image processing applications. The approach is valid because Curvelet transform produce correlated Curvelet coefficients, along the edge or curve of the image; the large Curvelet coefficients maybe have large Curvelet coefficients as it neighbors. Experimental results show that the proposed method is better than hard-thresholding denoising scheme in wavelet and curvelet domain.

**Keywords**- Curvelet Transform, image denoising, Multi-scale Geometric Analysis(MGA), CurShrink

### I. INTRODUCTION

Image data is generally vulnerable to be contaminated by noise during the acquisition and transportation process. The denoising of a natural image corrupted by noise is not only a classical but also a very difficult problem in image processing[1,2]. Researchers have proposed all kinds of noise removal methods according to the statistical and spectrum distribution property of noise which can be classified into two categories: spatial domain denoising approaches and frequency domain denoising approaches[3]. Typically, noise is characterized by high spatial frequencies in an image, and Fourier-based methods usually try to suppress high-frequency components, which also tend to reduce edge sharpness. Due to this Ingrid Daubechies and Stephen Mallat introduced a new transform namely: Wavelet transform[4]. And Donoho and his coworkers at Stanford pioneered a wavelet denoising scheme by thresholding the wavelet coefficients arising from the standard discrete wavelet transform[5]. Now this work has been widely used in image processing and engineering application and so on. However, because of the inherent limit plaguing wavelet denoising of images which arises from the well-known and

frequently depicted fact that the 2-D wavelet transform of image exhibits large wavelet coefficients even at fine scale, all along the important edges in the image, and the limitation of wavelet used in detecting highly anisotropic elements, such as alignments in an image or sheets in a cube, a new promising analyzing approach were developed: Multi-scale Geometric Analysis which include Ridgelet transform[6], MonoScale Ridgelet transform[7] and Curvelet transform[8], and these are very different from wavelet-like systems. Curvelets and Ridgelets take the form of basis elements which exhibit very high directional sensitivity and are highly anisotropic. The Curvelet transform are based on Multi-scale Ridgelets combined with a spatial bandpass filtering operation to isolate different scales. Its definition is as follows[8]:

The Curvelet tight frame for  $L_2(\mathbb{R}^2)$  is a collection of analyzing elements  $\gamma_\mu = \gamma_\mu(x_1, x_2)$  indexed by tuples  $\mu \in M'$  to be described below. It has been defined in [8] and has the following key properties:

Transform Definition:

$$\alpha_\mu \equiv \langle f, \gamma_\mu \rangle, \quad \mu \in M' \quad (1)$$

Parseval Relation:

$$\|f\|_2^2 = \sum_{\mu \in M'} |\alpha_\mu|^2 \quad (2)$$

L2 Reconstruction Formula:

$$f = \sum_{\mu \in M'} \langle f, \gamma_\mu \rangle \gamma_\mu \quad (3)$$

So according to the theory for a certain continuous-space model, discrete Ridgelet transforms and discrete Curvelet transforms provide near-ideal sparsity of representation of both smooth objects and of objects with edges. And In certain statistical theory, this implies that simple thresholding of noisy Curvelet coefficients in these expansions is a near-optimal method of denoising in the presence of white Gaussian noise.

In wavelet domain, Chang et. al.[9] proposed an adaptive thresholding scheme by taking the immediate neighbour coefficients into account. Here in this paper, we extend Chang's idea from wavelet domain to Curvelet domain. For image, we will consider a neighborhood window which is around the Curvelet coefficients to be thresholded, and introduce one way to threshold the Curvelet coefficient which is named: CurShrink. Experiments show that by using CurShrink thresholding rule to adjust the Curvelet coefficients we get the higher PSNR for the tested image and we also find that in image denoising case this approach

outperforms the traditional hard-thresholding algorithm in 2-D wavelet domain and Curvelet domain which is used by E L Starck and David Donoho.

The contents of this paper are as follows. In Section 2, we first introduce the definition of Curvelet transform and some of its advantages and shortcomings. Next we will explain how to use the CurShrink method to remove the noise in image data. Then experimental results will be shown in Section 4. And finally we'll give the conclusion and future work to be done in Section 5.

## II. DIGITAL CURVELET TRANSFORM

### A. Continuous Ridgelet Transform

The 2-D continuous Ridgelet transform in [2] can be defined as follows. We pick a smooth univariate function  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  with sufficient decay and satisfying the admissibility condition

$$\int |\hat{\psi}(\xi)|^2 / |\xi|^2 d\xi < \infty \quad (4)$$

For each  $a > 0$ , each  $b \in \mathbb{R}$  and each  $\theta \in [0, 2\pi)$ , the 2-D continuous Ridgelet transform in [2] can be defined as follows:

$$\psi_{a,b,\theta}(x) = a^{-1/2} \psi[(x_1 \cos \theta + x_2 \sin \theta - b) / a] \quad (5)$$

This function is constant along lines  $x_1 \cos \theta + x_2 \sin \theta - b = \text{const}$ . Transverse to these ridges it is a wavelet. Given an integrable bivariate function  $f(x)$ , we define its Ridgelet coefficients by:

$$R_f(a, b, \theta) = \int \psi_{a,b,\theta}(x) f(x) dx \quad (6)$$

And we have the exact reconstruction formula:

$$f(x) = \int_0^{2\pi} \int_{-\infty}^{+\infty} \int_0^{\infty} R_f(a, b, \theta) \psi_{a,b,\theta}(x) \frac{da}{a^3} \frac{db}{4\pi} \quad (7)$$

### B. Radon Transform

The basic tool for calculating Ridgelet coefficient is to view Ridgelet analysis as a form of wavelet analysis in the Radon domain. The Radon transform of an object  $f$  is the collection of line integrals indexed by  $(\rho, \theta) \in [0, 2\pi) \times \mathbb{R}$

$$Rf(\rho, \theta) = \int f(x, y) \delta(\rho - x \cos \theta - y \sin \theta) dx dy \quad (8)$$

where  $\delta$  is the Dirac function. The Ridgelet coefficients  $Rf(a, b, \theta)$  of an object  $f$  are given by analysis of the Radon transform via

$$R_f(a, b, \theta) = \int Rf(\rho, \theta) a^{-1/2} \psi[(\rho - b) / a] d\rho \quad (9)$$

From above formula we can find that the Ridgelet transform is precisely the application of a one-dimensional wavelet transform to the slices of the Radon transform where the angular parameter  $\theta$  is constant and the distance parameter  $\rho$  is varying.

The implementation of digital Ridgelet transform follows a widely used approach in the literature of medical imaging

and is based on fast Fourier transforms. The key component is to obtain approximate digital samples from the Fourier transform on a polar grid, i.e., along lines going through the origin in the frequency plane. Fig. 1 represents the flowgraph of the Ridgelet transform. For more information about Ridgelet transform, please see the reference[4]. The output of Ridgelet transform of a digital array of size  $n \times n$  is an  $2n \times 2n$  array and hence introduces a redundancy factor equal to 4.

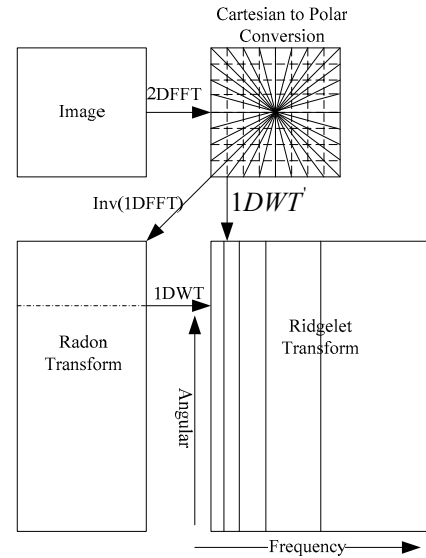


Figure 1. Ridgelet transform.

### C. Curvelet Transform

The idea of Curvelet is to represent a curve as a super-position of functions of various lengths and widths obeying the scaling law  $\text{width} \approx \text{length}^2$ . This can be done by first decomposing the image into subbands, i.e., separating the object into a series of disconnected scales. Each scale is then analyzed by means of a local Ridgelet transform.

Curvelets are based on Multi-scale Ridgelets combined with a spatial bandpass filtering operation to isolate different scales. This spatial bandpass filter nearly kills all Multi-scale Ridgelets which are not in the frequency range of the filter. In other words, a Curvelet is a Multi-scale Ridgelet which lives in a prescribed frequency band. The bandpass is set so that the Curvelet length and width at fine scales are related by a scaling law  $\text{width} \approx \text{length}^2$  and so the anisotropy increases with decreasing scale like a power law.

There is a very special relationship between the depth of the Multi-scale pyramid and the index of the dyadic subbands; the side length of the localizing windows is doubled at every other dyadic subband, hence maintaining the fundamental property of the Curvelet transform which says that elements of length about  $2^{-j/2}$  serve for the analysis and synthesis of the  $j$ th subband. While Multi-scale Ridgelets have arbitrary dyadic length and arbitrary dyadic widths, Curvelets have a scaling obeying  $\text{width} \approx \text{length}^2$ . Loosely speaking, the Curvelet dictionary is a subset

of the Multi-scale Ridgelet dictionary, but which allows reconstruction. In our opinion the “à trous” subband filtering algorithm is especially well-adapted to the needs of the digital Curvelet transform. The algorithm decomposes an  $n \times n$  image  $I$  as a superposition of the form

$$I(x, y) = c_J(x, y) + \sum_{j=1}^J w_j(x, y) \quad (10)$$

Where  $C_J$  is a coarse or smooth version of original image  $I$  and  $w_j$  means “the details of  $I$  at scale  $2^{-j}$ ”. See [8] for more information. Thus, this algorithm outputs  $J+1$  subband arrays of size  $n \times n$ .

As a side comment, we note that the coarse description of the image  $C_J$  is not processed. We used the default value pixels  $b_{min} = 16$  in our implementation. Fig. 2 gives one overview of the organization of the algorithm.

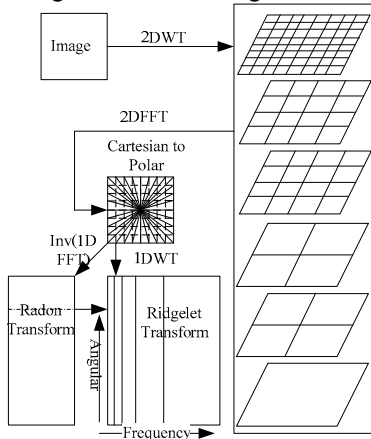


Figure 2. Curvelet transform

This implementation of the Curvelet transform is redundant. The redundancy factor is equal to  $16J+1$  whenever  $J$  scales are employed. Finally, the method enjoys exact reconstruction and stability, because each step of the transform is both invertible and stable.

### III. IMAGE DENOSING WITH CURSHRINK

Because of noise in the course measure, data from optical triangular scanners is preprocessed, which includes smoothing, removal noise and deleting exceptional points, before reconstructing surface is cable of identification clouds and process, the minimum distance is found automatically and tandem compound to acquire point group. Cut point group along a certain direction to obtain cross-section these are continuous and in a sequence [4].

The denoising algorithm is grounded on the empirical observation that the wavelet coefficients in a subband of a natural image can be summarized adequately by a generalized Gaussian distribution (GGD) with zero mean, we extend this assumption into curvelet domain which means that the curvelet coefficients in a directional subband can be described by a zero-mean GGD model too. Here we use adaptive threshold estimation via neighboring Bayes

thresholding estimation method, namely CurShrink. The CurShrink algorithm can be described as follows:

Perform forward 2-D Curvelet decomposition on the noisy image.

Apply the proposed shrinkage scheme to threshold the Curvelet coefficients using

The noise variance of each directional subband needs to be estimated first. In some situations, it may be possible to measure based on information other than the corrupted image. If such is not the case, it is estimated from the subband by the robust median estimator

$$\sigma_{k,l} = \text{Median}(|H_{k,l}|) / 0.6745 \quad (11)$$

Where  $H_{k,l}$  is directional subband at the  $l$ th level and the  $k$ th direction and  $\sigma_{k,l}$  is the noise variance of this subband.

Calculate the coefficients variance of local window of noisy image.

$$\sigma_{W_{i,j}}^2 = \sum_{(m,n) \in W_{i,j}} [c_{m,n} - \overline{c_{i,j}}]^2 \quad (12)$$

Where  $\sigma_{W_{i,j}}^2$  is variance estimation of local window  $W_{i,j}$  at position  $(i,j)$ .  $\overline{c_{i,j}}$  is the mean value of coefficients of window  $W_{i,j}$  (and we will explain how to choose the window in the next paragraph)  $N$  is the width of window and empirically it equals to 7,  $m \in [i-(N-1)/2, i+(N-1)/2]$ ,  $n \in [i-(N-1)/2, i+(N-1)/2]$ ; Note that if the center point of the window is on the boundary of subband, i.e.,  $c_{i,j}$  is on the first/last column/row, we don't process this coefficient.

estimate the coefficient variance of image without noise.

$$\sigma_{W_{i,j},k,l} = \begin{cases} \sqrt{\sigma_{W_{i,j}}^2 - \sigma_{k,l}^2}, & \sigma_{W_{i,j}}^2 > \sigma_{k,l}^2 \\ 0, & \sigma_{W_{i,j}}^2 < \sigma_{k,l}^2 \end{cases} \quad (13)$$

Where  $\sigma_{W_{i,j},k,l}$  means the estimated coefficients variance of local window  $W_{i,j}$  at the  $l$ th level and the  $k$ th direction for the noise-free image.

Calculate the thresholding value for each directional subband:

$$T_{W_{i,j},k,l} = \frac{\sigma_{k,l}}{\sigma_{W_{i,j},k,l}} \quad (14)$$

Perform inverse 2-D Curvelet transform on the thresholded curvelet coefficients.

Due to the nonorthogonality of digital Curvelet transform and quite redundant[4], as a consequence the noisy coefficients are correlated. That's why we should use local information to design our denoising scheme just as step 2(2). And of course, Curvelet coefficients are correlated in a small neighborhood. Therefore, here we incorporate neighbor

Curvelet coefficients within the window into thresholding process.

Suppose  $c_{i,j}$  is the set of Curvelet coefficients of contaminated image. For every Curvelet coefficient  $c_{i,j}$  of our interest, we need to “cover” a neighborhood window  $W_{i,j}$  above it. We choose the window as a square which means it includes  $3 \times 3$  or  $5 \times 5$  or  $7 \times 7$  pixels and the pixel to be thresholded is at the centre of the window square. Fig 3 illustrates a  $3 \times 3$  neighborhood window centred at the Curvelet coefficient to be thresholded. Note that the size of window has influence on the performance of denoising with our method. The larger the window, the relatively smaller the threshold is. If the size of window is too large, it is not only meaningless for keeping the edge of image, but also a lot of noise will be treated as signal due to the too small thresholding value. Another reason is that the implementation of the Curvelet transform is too redundant. The redundancy factor is equal to  $16J+1$  whenever  $J$  scales are employed. The larger the window, the more complex the computation is. So an intermediate window size of  $3 \times 3$  and  $5 \times 5$  is recommended.

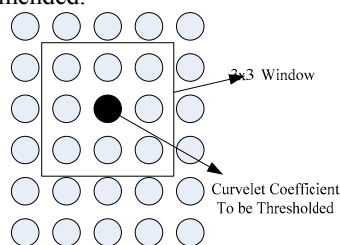


Figure 3. An illustration of the neighborhood window and the curvelet coefficients

#### IV. EXPERIMENTAL RESULTS

We perform our experiments on the well-known images Lena, Babara and Peppers. For comparison we implement wavelet, Curvelet denoising method with hard-thresholding and our proposed method. We choose db9/7 biorthogonal wavelet with 8 vanishing moments for wavelet decomposition and the decomposition level is 4. We threshold the all detailed scales except the four coarsest scales using the threshold for wavelet denoising. For Curvelet transform the decomposition level is 4 and four subbands of the finest scale are used. For different Gaussian white noise level, i.e., the different variance of white noise, Table 1 list the experimental results of 3 images for PSNR(Peak Signal to Noise Ratio). From Fig 4 to Fig 6 we illustrate the 3 images with corrupted by white noise with  $\sigma = 40$  and the images processed by the wavelet transform, the original Curvelet transform and the optimized Curvelet transform, respectively.

These experimental results show obviously that the Curvelet is superior to wavelet in terms of PSNR and the proposed CurShrink method outperforms hard-thresholding approach no matter it's implemented in wavelet or Curvelet domain. In recovery of linear features and curves the wavelet is worst, many edges has been blurred. But Curvelet-based approaches are good at recovering edge (such as the edge of Lena's hat).

#### V. CONCLUSION

In this paper, we study image denoising by incorporating neighboring Curvelet coefficients. Experimental results show that the CurShrink threshold rules give better results than hard-thresholding method no matter it is used by wavelet transform or Curvelet transform under all experiments. We conclude that CurShrink scheme can be used for practical image denoising application. The future work should be concentrated on how to utilize the correlation of the noise Curvelet coefficients and reduce the computation redundancy of Curvelet transform.

#### REFERENCES

- [1] Stephen Mallet, A wavelet tour of signal processing[M], Academic Press, 2008
- [2] Rafael Gonzalez, et. al., Digital Image Processing Using MATLAB[M], USA: Prentice Hall, 2003
- [3] XIE Jie-cheng, ZHANG Da-li, XU Wen-li, Overview on Wavelet Image Denoising [J], Journal of Image and Graphics, 2002, 7A(3):209-217
- [4] Stephen Mallet, WEN Liang-hwang. Singularity detection and processing with wavelets[J]. IEEE Trans Inform Theory, 1992, 38(11):617-643.
- [5] Stephen Mallet, ZHONG S. Characterization of signals from multiscale edges[J]. IEEE Trans PAMI, 1992, 14(7): 710-732.
- [6] Emmanuel J. Candès. Ridgelets: Theory and Applications[D]. USA: Department of Statistics, Stanford University, 1998.
- [7] Emmanuel J. Candès. Monoscale Ridgelets for the Representation of Images with Edges[R]. USA: Department of Statistics, Stanford University, 1999.
- [8] David L. Donoho, Mark R. Duncan. Digital Curvelet transform: strategy, implementation and experiments[A]. Proc. Aerosense 2000, Wavelet Applications VII[C]. SPIE, 2000. 4056. 12 - 29.
- [9] Grace Chang, Bin Yu, Martin Vetterli. Adaptive wavelet thresholding for image denoising and compression[J]. IEEE Trans. Image Processing, 2000, 9(9):1532-1546.
- [10] JIAO Li-cheng, TAN Shan, Development and Prospect of Image Multiscale Geometric Analysis[J], Acta Electronica Sinica, 2003, 31(12A):1975-1981
- [11] FENG Peng, PAN Yingjun, WEI Biao et. al., Enhancing Retinal Image by the Contourlet Transform[J], Pattern Recognition Letters, 2007, 28(4):516-522

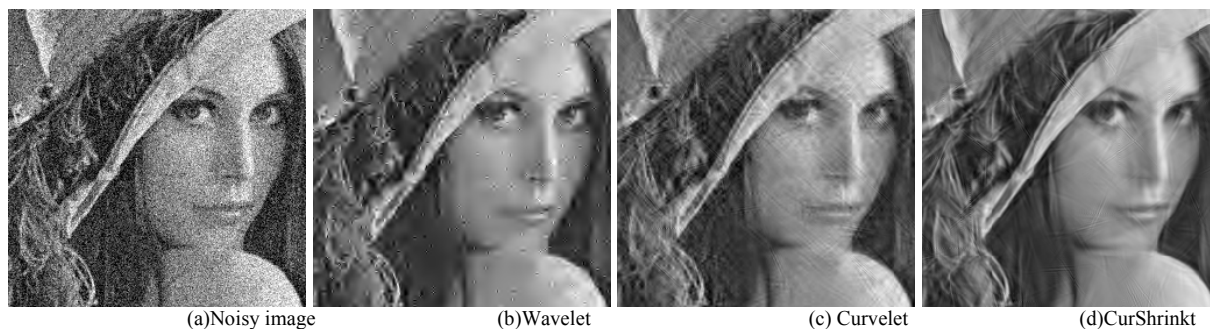


Figure 4. The “Zoom in” part of denoised Lena image with different algorithm( $\sigma=30$ )

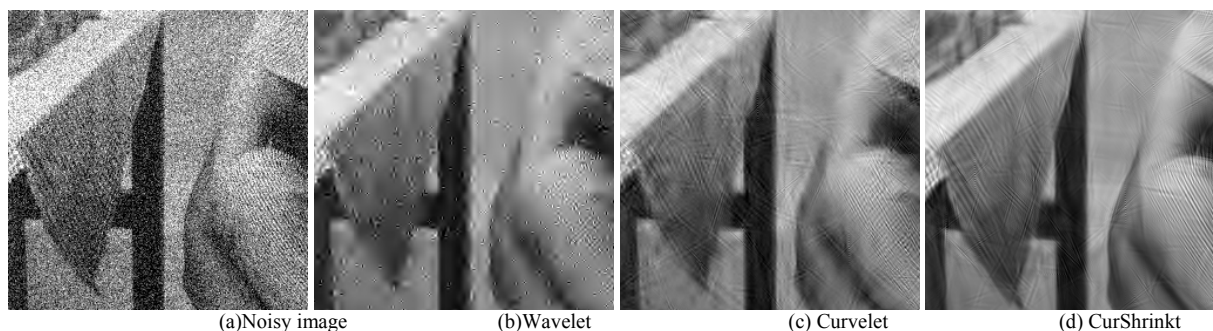


Figure 5. The “Zoom in” part of denoised Barbara image with different algorithm( $\sigma=40$ )

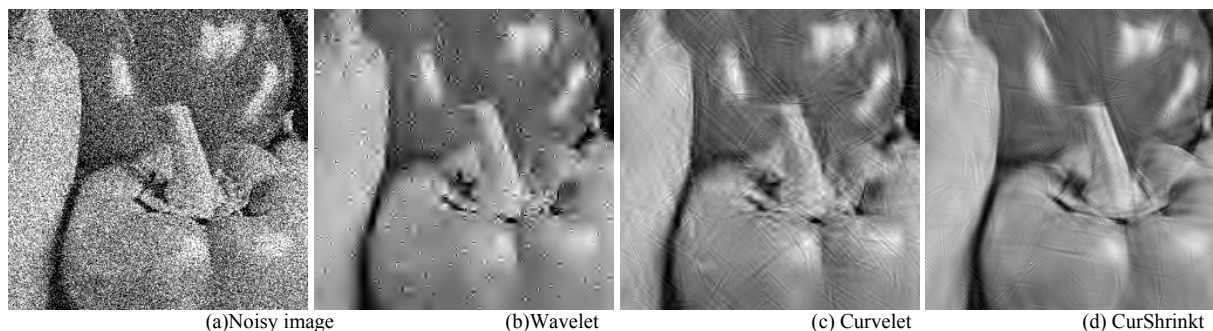


Figure 6. The “Zoom in” part of denoised Peppers image with different algorithm( $\sigma=50$ )

TABLE I. THE COMPARISON OF DENOISED IMAGES OF DIFFERENT ALGORITHMS IN TERMS OF PSNR

$\sigma$	Lena			Barbara			Peppers		
	30	40	50	30	40	50	30	40	50
Noisy	18.63	16.11	14.17	18.63	16.13	14.17	18.62	16.14	14.18
9/7 Wavelet	26.33	24.69	23.41	23.34	22.06	21.14	26.22	24.67	23.40
Curvelet	26.77	25.55	24.32	24.15	23.07	22.20	26.46	25.11	24.22
CurShrinkt	29.45	28.01	26.97	26.50	25.20	24.21	28.97	27.60	26.69