Robust adaptive directional lifting wavelet transform for image denoising

X.T. Wang  G.M. Shi  Y. Niu  L. Zhang

Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, Xidian University, Xi'an 710071, People's Republic of China
E-mail: xtwang@mail.xidian.edu.cn

Abstract: Recent researches have shown that the adaptive directional lifting (ADL) can represent edges and textures in images effectively. This makes it possible to separate noise from image signal distinctly in image denoising. However, a key issue named orientation estimation for ADL becomes inefficient and error prone in the noised circumstance. The authors propose a robust adaptive directional lifting-based (RADL) wavelet transform for image denoising by constructing ADL in an anti-noise way. In our method, a simple model of pixel pattern classification is incorporated into orientation estimation module to strengthen the robustness of this algorithm. Moreover, instead of determining the transform strategy based on sub-blocks, RADL is performed on pixel-level to pursue better denoising results. Experimental results show that the proposed technique demonstrates both PSNR and visual quality improvement on images with rich textures.

1 Introduction

Over the past decades, many kinds of denoising methods have been extensively discussed both in spatial and frequency domain. In essence, the key factor that lies in any successful denoising method is to find the different characteristics between signal and noise [1–9]. For instance, 2D discrete wavelet transform (DWT), which has been fully exploited and widely used in image denoising, shows different amplitude characteristics between signal and noise in wavelet domain [1]. The principle of wavelet denoising is to identify and zero out the coefficients smaller than a certain amplitude. Despite its prevalence, DWT suffers from an inherent limit: 2D DWT is not the optimal way to represent multidimensional signals such as edges and textures. Therefore it is of fundamental importance to overcome this limitation if one wants to achieve better denoising results.

In recent years, many researches have been studied for capturing the anisotropic geometrical structures in images efficiently via fully exploiting the directional correlation in either spatial or frequency domain, such as Multiscale Geometric Analysis (MGA) [7–12], directional filter banks (DFB) [13, 14], directional wavelet [15] and complex wavelets [16]. Many of these theories have been successfully applied in image denoising because of the good approximation behaviour. However, most of the approaches mentioned above are expensive systems with various redundancy ratios.

In contrast, wavelet-based tools and ideas are still very attractive for image processing because of their simplicity and efficiency. The applications of DWT have been extensively studied, offering us plenty of processing algorithms and realising structures, especially the development of lifting structure by Daubechies and Sweldens [17] and Sweldens [18]. This technique greatly decreased the computational complexity of the conventional DWT by factorising filter bank convolution into several lifting steps. However, lifting scheme is just a simplified approach to perform DWT, which does nothing to increase the direction flexibility. Therefore the question turns to: ‘Can we perform the lifting scheme with finer directionality, while still retaining its structure and desirable features?’ Wenpeng and Peng [19] and Chang and Girod [20], have given an affirmative answer by developing the adaptive directional lifting-based (ADL) wavelet transform. ADL integrates orientation transform into the framework of traditional lifting scheme and incorporates local spatial direction prediction into each lifting stage. Instead of applying horizontal and vertical lifting steps alternately, ADL performs the lifting scheme in local windows in the direction of the highest pixel correlation. So far, ADL transform has achieved tempting success in image compression [19–25], whereas little attention has been given to the possible use for image denoising.

An important feature of an efficient transform is the good approximation behaviour to signals. It has been proved that by the optimal approximation, sparse representation can distinguish the clear signal from the noise-corrupted observation [3–5]. Fortunately, ADL can provide a good approximation behaviour by effectively decorrelating the dependencies found over image discontinuities and compacting high-frequency components induced by image signals into the low-pass sub-band. Therefore the dominated singularities left in high-frequency sub-bands are caused by
Although image denoising can benefit from the good approximation behaviour of ADL, the direct application of ADL to image denoising may result in some problems. In ADL, the local dominant orientation estimation plays a very important role. In ADL, the local optimal orientation was determined by minimising the prediction error. This method is noise-sensitive and its performance declines in noised circumstances. Moreover, in the homogeneous regions where no directional information exists at all, orientation estimation is unnecessary. In these regions, ADL may falsely estimate directional information owing to the noise. Performing transform along fake directions may lead to slight scratches that severely affect the denoising performance. To overcome the drawbacks mentioned above, in this paper, we propose a robust adaptive directional lifting (RADL)-based wavelet transform by constructing the ADL in an anti-noise way. Preliminary results of this work were first reported in [26], in which, images are segmented into homogeneous regions and regions containing edges and textures (hereafter referred to as texture regions) before the transform to avoid artefacts in smooth regions. This paper provides a comprehensive description of the RADL and proposes a robust orientation estimation algorithm. First, the image pixels are classified into two sets according to their local activities, depicting smooth regions and texture regions. This procedure can reduce artefacts and unnecessary costs for orientation estimation in smooth regions. Second, in the texture regions where the orientation estimation is performed, the pixel classification model is integrated with the inter-scale correlation of wavelet coefficients to strengthen the robustness of the energy-based estimation algorithm. Finally, RADL is performed on pixel level. Each pixel has its own directional information and transform strategy. However, in the ADL, pixels contained in the same sub-block share a fixed angle.

2 Review of adaptive directional lifting-based wavelet transform

The fundamental differences between the conventional lifting and ADL are the different predict and update operators. Instead of performing the lifting steps always in horizontal or vertical directions, ADL analyses local spatial correlations in all directions, and then chooses a direction for lifting steps in which the prediction error is minimal [19]. Without loss of generality, let \( x[m, n] \) be a 2D signal. First, \( x[m, n] \) is split into two disjoint subsets, \( x_o[m, n] \) and \( x_e[m, n] \), where \( x_o[m, n] \) is the even indexed row subset and \( x_e[m, n] \) is the odd indexed row subset. The predicting and updating procedures are shown as follows

\[
\begin{align*}
    d[m, n] &= x[m, n] - P(x[m, n]) \\
    c[m, n] &= x_e[m, n] + U(d[m, n])
\end{align*}
\]

(1)

where \( d[m, n] \) is the prediction residual and \( c[m, n] \) is the coarse approximation of \( x[m, n] \). In the conventional lifting scheme, the predictor \( P(\cdot) \) is a linear combination of the neighbouring even subset \( x_o[m, n] \), and the updater \( U(\cdot) \) is a linear combination of the neighbouring residual \( d[m, n] \)

\[
\begin{align*}
    P(x_o[m, n], n) &= \sum p_i x_i[m + i, n] \\
    U(d[m, n]) &= \sum u_j d[m + j, n]
\end{align*}
\]

(2)

where \( p_i \) is the high-pass filter coefficient and \( u_j \) is the low-pass filter coefficient. Apparently, 2D conventional lifting scheme applies horizontal and vertical lifting steps alternately. However, nature images usually contain great amount of directional attributes that can be considered as linear edges in a local window. In order to further remove the spatial redundancy of these directional attributes, in ADL, directional spatial prediction is incorporated into the conventional lifting scheme to provide an efficient representation.

In the prediction step, the samples of the odd subset are predicted from the neighbouring even subsets with an optimal direction. This means sub-pixel accuracy is required. Taking vertical transform as an example, as shown in Fig. 1, the integer pixels are represented by the markers ‘•’, the half pixels are marked by ‘×’ and the quarter pixels by ‘+’. Assuming the strongest correlation exists at the angle \( \theta \), the prediction should use the quarter pixels of the even indexed samples identified by the arrows in Fig. 1. In principle, the prediction angle \( \theta \), which illustrates local direction information, should be a continuous variable; however, in practice, nine different directions of correlation are predefined as \( \text{dir} = 0, \pm 1, \pm 2, \pm 3, \pm 4 \), as shown in Fig. 1. Therefore the predictor \( P(\cdot) \) and updater \( U(\cdot) \) can be represented as

\[
\begin{align*}
    P(x_o^*, \text{dir}) &= \sum p_i x_i^*[m + i, n + \text{sign}(i - 1) \ast \text{dir}] \\
    U(d^*, \text{dir}) &= \sum u_j d^*[m + i, n + \text{sign}(i - 1) \ast \text{dir}]
\end{align*}
\]

(3)

where \( x_o^*[m, n] \) is the interpolated version of \( x_o[m, n] \), \( d^*[m, n] \) is the interpolated version of prediction residual \( d[m, n] \) and

\[
\text{sign}(x) = \begin{cases} 
1 & x \geq 0 \\
-1 & x < 0
\end{cases}
\]

Note that the conventional lifting scheme can be viewed as a special case of the ADL, when \( \text{dir} = 0 \).

Apparently, the main difference between the conventional 2D lifting and ADL is that ADL has an additional interpolating step and an orientation estimation step before the lifting stage. Therefore many features of the conventional lifting scheme, such as simple structure, in-place operation, suitable for hardware implementation and easily reverse transform, are inherited.

3 Credibility of ADL in image denoising

Since this proposed transform is constructed based on ADL, which is first proposed and applied in image compression,
in this section, we start from discussing the feasibility of applying ADL in image denoising.

The simple yet powerful wavelet-based denoising method namely soft-thresholding modifies all the coefficients in high-pass sub-bands and zeros out the coefficients smaller than the threshold. This indicates that in the ideal case, the signal information can be totally packed into the low pass-band, whereas leave only noise in the high-pass sub-bands. In this case the noise could be perfectly eliminated without damaging any signal information. However, in fact, wavelet distributes too much signal energy into the high-pass sub-bands. Therefore the damage to signal information is inevitable in the shrinkage denoising. The basic principle of wavelet shrinkage denoising indicates that better denoising performance could be achieved when less signal energy is left in the high-pass sub-bands.

Fortunately, ADL can provide a much better energy compaction property than conventional lifting via the finer directionality. In Fig. 2, we visualise the effect of 2D ADL wavelet transform. Fig. 2a is a test image (pixel values in the allowable range of 0 to 255) composed of textures in different directions. Figs. 2b and c are the HH bands of one level conventional 9/7 lifting and ADL (9/7 filter bank) decomposition, respectively. Obviously, the traditional lifting left huge signal high-frequency information in the HH band while ADL successfully removed the statistical redundancy of all the directional features. In Fig. 2b, the energy summation of all the coefficients is $8.11 \times 10^{4}$, the normalised energy is 4950, whereas the energy summation of Fig. 2c is $1.93 \times 10^{8}$, and the normalised energy is 117.8. Since less signal information is left in the high-pass sub-bands of ADL, one can expect better denoising results. Fig. 2d shows the HH band of one level ADL(9/7 filter bank) decomposition of the noised version of Fig. 2a. It is obvious that the dominate singularities left in the HH band are caused by noise; thus, after the thresholding, noise can be eliminated while the signal information is preserved quite well. The contradiction between noise-eliminating and edge-preserving can be solved in some sense.

Even though image denoising can benefit from the compaction property of ADL, the direct application of ADL to image denoising may lead to some problems:

- Most nature images contain quantitative smooth regions that have no directional information. There is no use performing orientation estimation in these regions; Moreover, ADL may falsely estimate directional information because of the noise in these regions. Performing lifting steps along the false direction would cause slight scratches in smooth regions, which severely affect the denoising performance.
- In the orientation estimation step of ADL, all high-frequency coefficients are used to determine the optimal direction. In noised circumstances, the accuracy of this orientation estimation method could be seriously affected by the noised pixels in smooth regions.
- In ADL, all the pixels in one sub-block should share the same local direction. However, the pixels in the same sub-block can be further classified into two subsets: pixels in texture regions and pixels in smooth regions. Thus it is not appropriate to perform directional lifting to all the pixels in the same dir. Each pixel should has its own dir, and the dir of the pixels belonging to smooth regions should be set to zero.

4 Robust adaptive directional lifting

This section is concerned with the robust adaptive directional lifting (RADL) and its application in image denoising.

To solve the problems stated above, we extend ADL to a new robust adaptive directional lifting (RADL)-based wavelet transform, which involves three additional parts.

Image pixel classification based on noise variance estimation, which classifies all the pixels into two categories: pixels belonging to texture regions and pixels belonging to smooth regions.

Robust orientation estimation based on pixel classification and inter-scale correlation, which provides a more accurate local orientation estimation than the previous method adopted in ADL.

Optimal transform strategy performs the transform on pixel level instead of block level to avoid artefacts in the smooth regions.

4.1 Image pixel classification

Let $\sigma_{im}^{2}$ be the variance of a local window of a clear image $X$. As we know that in a local window [27]

$$
\sigma_{im}^{2} = \begin{cases} 
0 & \text{smooth region} \\
\frac{1}{N} \sum_{i,j \in B} [X(i,j) - \mu]^2 & \text{texture region} 
\end{cases}
$$

Fig. 2  Testing image and its one-level decomposition HH band

a Test image  
b Traditional lifting scheme  
c ADL  
d ADL decomposition for the noised image
Assuming that the noisy image $Y$ is corrupted by additive zero-mean white Gaussian noise whose variance is $\sigma_n^2$. The variance of a local window $\sigma_{n,m}^2$ can be presented as follows

$$\sigma_{n,m}^2 \begin{cases} \approx \sigma_n^2 & \text{smooth region} \\ = \sigma_n^2 + \sigma_s^2 & \text{texture region} \end{cases}$$

(5)

Therefore, in a noisy image, we can set a threshold $T$, for each pixel $Y(i,j)$

$$\begin{cases} \sigma^2_{Y(i,j)} / \sigma_n^2 \leq T, & Y(i,j) \in \text{smooth region} \\ \sigma^2_{Y(i,j)} / \sigma_n^2 > T, & Y(i,j) \in \text{texture region} \end{cases}$$

(6)

where $\sigma^2_{Y(i,j)}$ is the variance of the local window centred by the pixel $Y(i,j)$. $\sigma_n^2$ is the noise variance of the whole image, which is estimated by the widely used estimation method based on robust median estimator [28].

A label flag can be set to each pixel

$$\text{flag}_{Y(i,j)} = \begin{cases} 0, & \sigma^2_{Y(i,j)}/\sigma_n^2 \leq T \\ 1, & \sigma^2_{Y(i,j)}/\sigma_n^2 > T \end{cases}$$

(7)

flag expresses the local activity of each pixel in the image. $\text{flag} = 0$ means the pixel belongs to the smooth region, and $\text{flag} = 1$ means the pixel belongs to the texture region. The threshold $T$ can be obtained from the local variance image field [29]. This threshold is validated through large number of images. Fig. 3a shows a test image ‘Barbara’ corrupted by Gaussian noise with a standard deviation equal to 20. Fig. 3b is the pixel classification result of the noisy image.

In our method, we take two classification steps to complete the pixel classification. First, the image is segmented into sub-blocks. All sub-blocks of the image are classified into: ‘Important Region’ – $B_{IR}$ and ‘Not Important Region’ – $B_{NIR}$. $B_{IR}$ belongs to the texture region where interpolation and orientation estimation are performed. $B_{NIR}$ belongs to the smooth region where no directional information exists at all. In $B_{NIR}$, line singularities are not included. Point singularities can be effectively caught by rectilinear wavelet, therefore the directional information $\text{dir}$ can be set to zero directly. This method is free from the cost of orientation estimation and the artefacts caused by lifting using virtual pixel values in smooth areas. Second, the pixel classification procedure is performed to each pixel in $B_{IR}$ to obtain the classification result as shown in Fig. 3b.

The complexity of this procedure is analysed as follows: Considering a sub-block containing $N$ pixels, orientation estimation needs $72 \times N$ addition and multiplication operations, whereas the sub-block classification needs $4 \times N$ operations and the pixel classification in a sub-block needs $34 \times N$ operations.

Note that in a transform containing more than one level, the orientation estimation needs to be performed in every level. However, the pixel classification only needs to be performed one time before the transform. Assuming an image is segmented into $A$ sub-blocks, in which, $B$ sub-blocks are $B_{IR}$. Each sub-block contains $N$ pixels. Performing orientation estimation directly on every sub-block costs $94.5 \times N \times A$ operations, and our proposed method costs $(128.5 \times N \times B) + (4 \times N \times A)$ operations. Apparently, the complexity of our method is determined by $B/A$. When $B/A < 0.704$, the complexity of our method is lower than the previous method.

4.2 Robust orientation estimation

As mentioned above, the accuracy of the orientation estimation is a key to the good performance of ADL. In order to illustrate the effects of the inaccurate orientation information, we have done some special test experiments. Fig. 4a is a test image whose textures are all along 45°. In our experiment, nine directions (45°–135°) are compelled into the ADL transform respectively to illustrate the effects caused by different orientation information. Table 1 shows the sub-band energy of one-level ADL decomposition for the clear test image (Fig. 4a) carried out in different directions. Apparently, the energy left in high-pass bands increases rapidly when the compelled direction moves far away from the correct direction. Fig. 4b is the noised version of Fig. 4a, contaminated by Gaussian noise with a standard deviation $\sigma$ equal to 25. The PSNR of the noised

---

**Fig. 3** Noised Barbara and the pixel classification result

- a Noised image ($\sigma = 20$, PSNR = 22.0256 dB)
- b Pixel classification result
image (Fig. 4b) is 17.9604 dB. Figs. 4c–k show the denoising results of ADL using different directions. The PSNR comparison is shown in the last column of Table 1. The experimental results indicate that the accuracy of the orientation estimation is a key to good denoising performance in our method.

In ADL, the orientation estimation step analyses the local spatial correlation in all directions. It calculates the energy summation of the prediction error in each direction, and chooses the direction in which the prediction error energy is minimal

$$\text{dir} = \arg \min_{\text{dir}} \{ \| f_{\text{dir}} \| \}$$

where $f$'s are the frequency coefficients of the high-pass sub-band. The energy of $f$ is calculated as follows

$$\| f_{\text{dir}} \| = \sum_{i=1}^{M} \sum_{j=1}^{N} x_{i,j}^2$$

where $M$ and $N$ are the size of the high-pass sub-band.

This energy-based orientation estimation technique takes all the coefficients in the high-pass sub-band into account to

---

Table 1 Sub-band energy and denoising PSNR comparison

<table>
<thead>
<tr>
<th>Compelled direction</th>
<th>LL</th>
<th>LH</th>
<th>HL</th>
<th>HH</th>
<th>PSNR, dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>45° (correct direction)</td>
<td>$1.2601 \times 10^7$</td>
<td>$5.49 \times 10^3$</td>
<td>15.50</td>
<td>7.900</td>
<td>21.45</td>
</tr>
<tr>
<td>56.25°</td>
<td>$1.2604 \times 10^7$</td>
<td>$5.82 \times 10^3$</td>
<td>52.37</td>
<td>22.06</td>
<td>21.30</td>
</tr>
<tr>
<td>67.5°</td>
<td>$1.2593 \times 10^7$</td>
<td>$5.46 \times 10^3$</td>
<td>$1.00 \times 10^3$</td>
<td>410.66</td>
<td>20.57</td>
</tr>
<tr>
<td>78.75°</td>
<td>$1.2591 \times 10^7$</td>
<td>$5.46 \times 10^3$</td>
<td>$1.92 \times 10^4$</td>
<td>$2.53 \times 10^3$</td>
<td>20.03</td>
</tr>
<tr>
<td>90°</td>
<td>$1.2590 \times 10^7$</td>
<td>$3.87 \times 10^3$</td>
<td>$1.33 \times 10^5$</td>
<td>$1.12 \times 10^4$</td>
<td>18.90</td>
</tr>
<tr>
<td>101.25°</td>
<td>$1.2576 \times 10^7$</td>
<td>$3.22 \times 10^3$</td>
<td>$4.94 \times 10^5$</td>
<td>$3.13 \times 10^4$</td>
<td>17.31</td>
</tr>
<tr>
<td>112.5°</td>
<td>$1.2540 \times 10^7$</td>
<td>$2.18 \times 10^3$</td>
<td>$1.17 \times 10^6$</td>
<td>$6.40 \times 10^4$</td>
<td>16.40</td>
</tr>
<tr>
<td>123.75°</td>
<td>$1.2498 \times 10^7$</td>
<td>$1.40 \times 10^3$</td>
<td>$1.96 \times 10^6$</td>
<td>$1.00 \times 10^4$</td>
<td>15.95</td>
</tr>
<tr>
<td>135°</td>
<td>$1.2468 \times 10^7$</td>
<td>$2.01 \times 10^3$</td>
<td>$2.51 \times 10^6$</td>
<td>$1.25 \times 10^4$</td>
<td>15.79</td>
</tr>
</tbody>
</table>

---

Fig. 5 Direction analysis of a texture image

- a Grey-scale image with two edges
- b Prediction residual $H$ of the grey-scale image
- c Pixel classification result of the grey-scale image
- d Down-sampling of pixel classification result

---

Fig. 6 Orientation estimation results

- a Estimation by ADL for clear image
- b Estimation by ADL for noised image
- c Estimation by ROE for noised image
calculate and analyse the energy as shown in formula (9). This approach can achieve rather accurate orientation information in clear images. However, in noised circumstance, if the spatial correlation of the image is not strong enough, the estimation results are usually incorrect. For example in blocks with only a few edges, the accuracy of the orientation estimation will be seriously affected by the disturbance of the coefficients induced by the pixels belonging to the noisy smooth regions beside the edges. This indicates that the disturbance will be greatly decreased, if the coefficients used for the energy calculation only come from the large magnitude coefficients induced by pixels belonging to the texture regions.

In order to increase the robustness of the energy-based orientation estimation algorithm in ADL, we propose a robust orientation estimation (ROE) algorithm by taking the pixel classification result and the wavelet coefficients inter-scale correlation into account.

In this paper, the image discontinuities such as edges and textures are localised by pixel classification performed to the full-resolution image. However, the lifting scheme is a pyramid transform, which means the size of each sub-band changes during the transform. Locations of the large magnitude coefficients that represent the image discontinuities change in different scales. Fortunately, it has been proved that there are strong correlations between the coefficients in different scales [30]. In critically sampled orthogonal wavelet decomposition, the position of large wavelet coefficients out of parents at lower resolutions can be predicted with good

Fig. 7 Denoising results comparison

a Original image
b Noised image (σ = 20, PSNR = 22.88 dB)
c Denoising by ADL transform strategy (PSNR = 24.92 dB)
d Denoising by the proposed transform strategy (PSNR = 25.05 dB)

Fig. 8 Original test images
From top left to bottom right: Barbara, Bike, Zebra, Monarch, Remote1 and Remote2
accuracy [31]. It means that if the positions of image discontinuities in the high resolution are known, one can localise the large wavelet coefficients in low resolutions.

Fig. 5 illustrates the method of the proposed ROE. Fig. 5a is a grey-scale image with two edges. The pixel classification result of Fig. 5a is shown in Fig. 5c. Fig. 5b is the prediction residual $H$ of Fig. 5a. As mentioned above, taking into account the dependencies among the coefficients, one can determine the location of the large coefficients in $H$ by down-sampling the pixel classification result (Fig. 5c) along
the row, which is shown in Fig. 5d. Finally, the numeric matrix, which represents the prediction residual (Fig. 5b), and down-sampled classification result (Fig. 5d) are merged by point multiplication to determine $\Omega$ – the coefficients used for calculating the sub-band energy. The energy of the prediction residual is calculated as follows

$$\|f_{\text{dir}}\| = \sum_{i,j \in \Omega} V_{x}^{2}$$

(10)

In other words, only the coefficients induced by pixels belonging to texture regions are taken into account in the determination of the orientation estimation. In this way, the disturbance of the noise in smooth regions can be avoided.

Figs. 6a and b show the orientation estimation results of the energy-based approach used in ADL for clear and noised version of Barbara, respectively. Obviously, when ADL is performed in noised circumstances, the estimated directions are almost all incorrect where the spatial correlation is not strong, as shown in Fig. 6b. Fig. 6c shows the orientation estimation result of the proposed ROE. Apparently, the accuracy of local dominant orientation estimation in Fig. 6c has been greatly improved compared with Fig. 6b.

It can be seen from this section that the accuracy of the orientation estimation is a key to good denoising results in our method. Inaccurate directional information will seriously affect the denoising results. However, the accuracy of the orientation estimation can be easily diminished by noise in smooth regions. Therefore we propose ROE, which improves the accuracy of the orientation estimation by taking into account only pixels in edge and texture regions. This method further improves the denoising performance.

### 4.3 Direction modification and pixel-level transform strategy

In ADL, once the local dominant direction is determined, all the pixels in one sub-block share the same direction in which the transform is performed. However, in $B_{IR}$, the pixels can be further classified into two subsets: pixels belonging to texture regions and pixels belonging to smooth regions beside the texture regions as shown in Fig. 5. It is inappropriate to perform directional transform to the pixels out of texture regions in $B_{IR}$.

In our method, RADL is only performed to pixels belonging to edge and texture regions. If RADL is performed to pixels in smooth regions, artefacts will be caused in denoising results. As mentioned before, in RADL, intensity values of fractional pixel locations are needed during the transform. Hence, the interpolation of sub-pixels becomes an issue. As we know, a key to good interpolation is the smoothness of signal [32]. However, the smoothness of the image signal does not hold across edges. Therefore interpolation cannot give accurate value for the sub-pixels belonging to smooth regions beside the edges. Using these inaccurate sub-pixels as the inputs of the predictor and updater in formula (3) would cause artefacts in smooth regions as shown in Fig. 7c.

Fortunately, comparing with the previous ADL used in image compression, there is no restriction of side-information in image denoising. More flexible transform strategy can be selected. This makes it possible to implement RADL at pixel level. In

![Fig. 10 Denoising results of Monarch](image-url)

From top left to bottom right: original image, noised image ($\sigma = 20$, PSNR = 22.13 dB, SSIM = 0.52), CT (PSNR = 25.41 dB, SSIM = 0.743), CV (PSNR = 27.83 dB, SSIM = 0.726), RADL-ST (PSNR = 27.91 dB, SSIM = 0.742), RADL-BiS (PSNR = 28.31 dB, SSIM = 0.766)

© The Institution of Engineering and Technology 2011
this paper, we use the pixel pattern classification result again to modify the directional information of each pixel:

\[ \text{dir}'_{i,j} = \text{dir}_{i,j} \times \text{flag}_{i,j} \]  

(11)

where \( \text{dir}_{i,j} \) is the directional information of each pixel determined in the orientation estimation, \( \text{flag}_{i,j} \) is the classification label of each pixel determined in formula (7).

Fig. 7d shows the denoising result of the proposed transform. It can be seen that performing RADL only to pixels belonging to edge and texture regions can effectively improve the denoising result by avoiding artefacts in smooth regions.

5 Experimental results

In order to illustrate the potential of the RADL in image denoising, we tested a set of standard 8-bit grey-scale images with rich textures namely Barbara, Bike, Zebra and Monarch as shown in Fig. 8. We also chose two remote sensing images (Remote1 and Remote2) to highlight the denoising performance of the RADL relative to other denoising techniques. All these images are contaminated by simulated zero-mean additive white Gaussian noise at five different intensity levels: \( \sigma \in [10, 20, 30, 40, 50] \).

This section presents three examples to show the efficiency of this new transform by comparing it with other denoising methods. In the first example, RADL is compared with traditional DWT. Different threshold shrinkages such as soft-thresholding (ST), Bayes-Shrink (Bas), Bivariate-Shrink (Bis) are used to deal with the coefficients of DWT and RADL, respectively. The second example shows the denoising performance of the RADL relative to other transforms such as contourlet (CT) and curvelet (CV), which also focus on catching the geometric structures of image signals. For further illustration, the third example shows and analyses a detail part of the Barbara image to facilitate the readers observe the qualitative differences between the different methods. In all comparisons, we use critically sampled structures for all the transforms over four decomposition levels. DWT and RADL are both constructed on 9/7 filter banks. These reliable comparisons were only possible thanks to the kindness of the various authors who have provided their respective MATLAB codes on their personal web sites.

5.1 Example 1

In this example, a critically sampled orthogonal DWT, with 9/7 filter banks, is used for comparison. Three different threshold shrinkages are adopted to deal with the coefficients. For each image, six denoising methods, including wavelet with Soft-Thresholding [1] (DWT-ST), wavelet with Bayes-Shrink [33] (DWT-BaS), wavelet with Bivariate-Shrink [34] (DWT-BiS), RADL with Soft-Thresholding (RADL-ST), RADL with Bayes-Shrink (RADL-BaS) and RADL with Bivariate-Shrink (RADL-BiS), have been performed to all noise levels. For all the denoising methods, we assumed that the noise level \( \sigma \) is unknown and is estimated by the robust median estimator.

![Fig. 11 Denoising result of Bike](image-url)

From top left to bottom right: original image, noised image (\( \sigma = 20 \), PSNR = 23.32 dB, SSIM = 0.48), CT (PSNR = 25.08 dB, SSIM = 0.646), CV (PSNR = 28.26 dB, SSIM = 0.839), RADL-ST (PSNR = 27.20 dB, SSIM = 0.843), RADL-BiS (PSNR = 28.84 dB, SSIM = 0.860)
The resulting PSNR values of the noised images ($\sigma = 20$) and the average improvements (AI) compared with the noised images over all noise levels are detailed in Table 2, with the best one highlighted in bold font.

As mentioned above, RADL can capture the image structure efficiently, thus it can eliminate noise as well as preserve edge information. In order to prove this point, we adopt a latest objective perceptual quality index structural similarity (SSIM) [35] as another quality evaluation method. This image assessment assumes that the human visual perception is highly adapted for extracting structural information from a scene, and thus compares local patterns of pixel intensities that have been normalised for luminance and contrast. We use this structure-based method to evaluate the subjective quality improvement of all the denoising techniques mentioned above. The comparison of SSIM for each denoising method is also shown in Table 2.
Table 3 High-frequency power analysis

<table>
<thead>
<tr>
<th></th>
<th>(p_{\text{low}} (\times 10^{11}))</th>
<th>(p_{\text{high}} (\times 10^{15}))</th>
<th>(h_{\text{rate}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>original image</td>
<td>2.4189</td>
<td>2.3547</td>
<td>0.0887</td>
</tr>
<tr>
<td>noised image</td>
<td>2.4172</td>
<td>2.8077</td>
<td>0.1041</td>
</tr>
<tr>
<td>CT</td>
<td>2.4154</td>
<td>1.9104</td>
<td>0.0758</td>
</tr>
<tr>
<td>CV</td>
<td>2.3971</td>
<td>2.1428</td>
<td>0.0821</td>
</tr>
<tr>
<td>DWT-ST</td>
<td>2.4037</td>
<td>1.3888</td>
<td>0.0546</td>
</tr>
<tr>
<td>DWT-BaS</td>
<td>2.4054</td>
<td>1.4793</td>
<td>0.0579</td>
</tr>
<tr>
<td>DWT-BiS</td>
<td>2.4048</td>
<td>1.8462</td>
<td>0.0713</td>
</tr>
<tr>
<td>RADL-ST</td>
<td>2.4008</td>
<td>1.9418</td>
<td>0.0748</td>
</tr>
<tr>
<td>RADL-BaS</td>
<td>2.4029</td>
<td>1.2394</td>
<td>0.0818</td>
</tr>
<tr>
<td>RADL-BiS</td>
<td>2.4052</td>
<td>2.2037</td>
<td>0.0639</td>
</tr>
</tbody>
</table>

Table 3 shows the PSNR of the detail part and \(h_{\text{rate}}\) of denoising image of each methods. Apparently, RADL can preserve more high-frequency information than other techniques.

5.2 Example 2
This example shows the denoising results of some typical images to highlight the performance of the RADL relative to other transforms. From Figs. 10 to 12, we visualise the denoising performance of different transforms by zooming into some test images (Monarch, Bike and remote1). CT, CV, RADL-ST, RADL-BiS are used for comparison. It can be seen from the experimental results, RADL outperforms CT and CV both subjectively and objectively. Moreover, image denoising methods based on RADL exhibit less visual artefacts than CT and CV especially in remote sensing images.

5.3 Example 3
For further illustration, this example analyses the high-frequency energy of a detail part of Barbara to facilitate the readers observe the qualitative differences between the different methods as shown in Fig. 13. As we know, textures and edges induce most of the 2D signal high-frequency energy; therefore, it is necessary to preserve high-frequency energy in texture regions as much as possible. Here we define a parameter \(h_{\text{rate}}\) which represents the ratio of the high-frequency energy and total energy of an image

\[
\text{energy}(\text{image}) = \frac{p_{\text{high}}}{p_{\text{high}} + p_{\text{low}}}
\]

where \(p_{\text{high}}\) is the high-frequency energy, \(p_{\text{low}}\) is the low-frequency energy. Table 3 shows the PSNR of the detail part and \(h_{\text{rate}}\) of denoising image of each methods. Apparently, RADL can preserve more high-frequency information than other techniques.

6 Conclusion
In this paper, RADL transform is proposed for image denoising. Using recent research from ADL, directional spatial prediction was incorporated into the conventional lifting scheme to remove the spatial redundancy of the directional attributes. In our method, pixel classification and inter-scale correlation are taken into account to strengthen the robustness of the orientation estimation algorithm. Moreover, the transform is performed at pixel level by only carrying out directional transform on pixels belonging to texture regions. Some simulation tests are made to verify that the texture features of the original signal can be reserved after denoising via this method.

7 Acknowledgment
The authors would like to thank NSF of China (Nos. 60776795, 60736043, 60902031), Specialized Research Fund for the Doctoral Program of Higher Education of China (No. 200807010004, 20070701023), Fundamental Research Funds for the Central Universities of China (No. JY10009092028) for their support. They would also like to thank Dr. F. Wu, Microsoft Research Asia, for many insightful discussions on the topic, and Dr. Minh N. Do for providing the code to his contourlet denoising algorithm and the affirmance to the denoising results.

8 References