A NONLINEAR STEERABLE COMPLEX WAVELET DECOMPOSITION OF IMAGES

Zikai Sun and Thierry Blu

The Chinese University of Hong Kong
zksun@link.cuhk.edu.hk, thierry.blu@m4x.org

1. ABSTRACT

Signal and image representations that are steerable are essential to capture efficiently directional features. However, those that are successful at achieving directional selectivity usually use too many subbands, resulting in low computational efficiency. In this paper, we propose a two-dimensional nonlinear transform that uses only two subbands to achieve rotation invariance property, and enjoys a mirror reconstruction making it similar to a "tight frame". The two-subband structure is merged into a unique, concise, complex-valued subband that approximates a Wirtinger gradient which is naturally steerable. Complete steerability, though, is achieved by utilizing the Fourier-Argand representation, which provides a steerable filter able to estimate the amplitude and direction of image features, even in the presence of very high noise. We demonstrate the efficiency of the representation by comparing how it performs in wavelet-based denoising algorithms.

Index Terms— Complex wavelet, nonlinear transform, rotation-invariance, steerable, Fourier-Argand representation.

2. INTRODUCTION

The discrete wavelet transform (DWT)[13, 5], as a popular multi-resolution signal representation, has been intensively applied in image processing, such as segmentation, texture classification[21], edges detection[1]and denoising[15]. However, separable wavelet transform performs less efficiently on 2D image signals compared to its remarkable success on 1D signals. One reason is that DWT regards 2D images as an extension of 1D signal and neglects the underlying intrinsic geometric information.

To mitigate the limitation of wavelets in higher dimension data, multiscale decompositions that enjoy some form of steerability have been proposed. For example, ridgelet[4], curvelet[19], contourlet[7], directionlet[22] and shearlet [10, 12], etc. However, edges with different orientations are isolated in different subbands, which inevitably uses many filters to achieve multiple orientations. Besides, the directional selectivity is limited, and the rotation sensitivity problem still exists: slight rotation in the input signal will completely change the coefficient pattern. Simoncelli's Steer-

able Pyramid[18] claims approximate rotational symmetry, but uses many subbands and has high redundancy. DT-CWT[16] use the complex filter to render it shift-invariant. Its 2D version has six oriented subbands, which is makes it possible to distinguish orientations by steps of 15 degrees.

Many methods also take nonlinear operation into consideration. Lifting scheme[20, 6] enables a locally adaptive wavelet transform. Based on this, [9] proposed an edge adaptive lifting structure. [8] use an interpolation-based direction-adaptive DWT on directional lifting. There are also some adaptive schemes proposed in recent years. [17] suitably adjust the transform based on spatial segmentation. ADR [11] use a curve detector to adaptively deals with line and curve information. These adaptive methods take the strength of different transform. However, The idea behind it is a scheme that uses spatial segmentation and selects proper transform for different regions of the image rather than an adaptive transform itself.

This paper contributes a novel nonlinear transform for 2D signals. Contradict from previous ideas that expanding many subbands to enhance the limited orientational selectivity, we use nonlinearity to achieve direction selectivity. This indicates that the transform is no longer a tight frame. However, similar property remains and it still guarantees perfect reconstruction. The proposed decomposition method has a concise two subbands structure: one subband for lowpass and another for anisotropic features, such as edges in images. Besides, this method is rotation invariance. This property provides stability concerning rotation, which is a much more robust and valuable property than the orientational selectivity.

3. METHOD

3.1. Overview

Throughout paper we denote by $z = (z_1, z_2) \in \mathbb{C}^2$ the two-dimensional variable in the z-plane. Define $H(z)$ and $G(z)$ as the z-transform of the two dimensional lowpass and highpass filters, h and g, respectively. We write our filter banks as,

$$\begin{cases} H(z) = L(z) \\ G(z) = G_1(z) + iG_2(z) \end{cases} \quad (1)$$

Where $L(z)$ is a traditional 2D wavelet lowpass filter. $G(z)$ is a complex highpass filter, whose real and imaginary parts

typically represent two orthogonal gradients. One of the parts, such as $G_1(z)$, directly represents the edge information of the image. This allows us to express meaningful edge information of the natural image. Based on the description above, the highpass filter should have the following form,

$$G_1(z) = \mathfrak{Re}\{D(z)e^{-i\theta(x)}\}$$
$$G_2(z) = \mathfrak{Re}\{D(z)e^{-i(\theta(x)-\frac{\pi}{2})}\} \tag{2}$$
$$D(z) = D_1(z) + iD_2(z)$$

Where $\theta(x)$ is a (non-rigorous) attempt to indicate that the phase of the output of the filter $D(z)$ is further modified locally (dependence on x). In the complex plane, together with the "phase change" term $e^{-i\theta}$, $G(z)$ can be regarded as a local "projection" of $D(z)$ on two orthogonal directions: $\theta(x)$ and $\theta(x) - \frac{\pi}{2}$. $\theta(x)$ various according to the position of the pixel $x \in \mathbb{R}^2$ and represent a direction along the edge of the nature image. $D(z)$ is designed to be a Wirtinger gradient filter.

By designing the highpass complex filter $D(z)$, we can make the $G(z)$ channel represent edge information. And by assigning the value of $\theta(x)$ locally related to the content of the natural image, we can create a rotation invariant subband. As for a standard DWT decomposition, the multiresolution 2D filters are given by

$$H^l(z) = H(z^{2^{l-1}})H^{l-1}(z) = \prod_{k=0}^{l-1} H(z^{2^k})$$
$$G^l(z) = G(z^{2^{l-1}})H^{l-1}(z) = G(z^{2^{l-1}})\prod_{k=0}^{l-2} H(z^{2^k}) \tag{3}$$
$$l = 1, ..., N$$

Where l is the scale index. The signal is iteratively decomposed on the lowpass analysis channel, which gives a redundant signal representation.

3.2. Design of $D(z)$ and PR condition

To perfectly reconstruct a 2D signal with a synthesis filterbank that mirrors the analysis (tight-frame, so as to guarantee the identity between errors in the transformed domain, and in the image domain), the analysis filter bank needs to be designed under the constraint that $H(z)$ and $G(z)$ are power complementary. This perfect reconstruction condition is written as follows,

$$H(z)H(z^{-1}) + \sum_{i=1}^{2} G_i(z)G_i(z^{-1}) = \mathrm{C} \tag{4}$$

Where C is a constant. Rewrite Eq. 4 in matrix form, we can define $P_\theta(z), M(z) \in \mathbb{R}^{2 \times 2}$ as below,

$$\underbrace{\begin{bmatrix} G_1(z^{-1}) & G_2(z^{-1}) \\ -G_2(z) & G_1(z) \end{bmatrix}}_{P_\theta(z^{-1})^T} \underbrace{\begin{bmatrix} G_1(z) & -G_2(z^{-1}) \\ G_2(z) & G_1(z^{-1}) \end{bmatrix}}_{P_\theta(z)}$$
$$= \underbrace{\begin{bmatrix} C - H(z^{-1})H(z) & 0 \\ 0 & C - H(z^{-1})H(z) \end{bmatrix}}_{M(z)} \tag{5}$$

Decompose $P_\theta(z^{-1})$, we then have $P_\theta(z^{-1}) = R_\theta P(z^{-1})$, where

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, P(z) = \begin{bmatrix} D_1(z) & -D_2(z^{-1}) \\ D_2(z) & D_1(z^{-1}) \end{bmatrix} \tag{6}$$

Since $R_\theta(z^{-1})^T R_\theta(z) = I$, combine this with Eq. 5, we have

$$M(z) = P_\theta(z^{-1})^T P_\theta(z)$$
$$= P(z^{-1})^T (R_\theta(z^{-1})^T R_\theta(z))P(z) \tag{7}$$
$$= P(z^{-1})^T P(z)$$

This indicate that filters only needs to meet following condition to satisfy Eq. 4,

$$H(z)H(z^{-1}) + \sum_{i=1}^{2} D_i(z)D_i(z^{-1}) = \mathrm{C} \tag{8}$$

Observe that θ cancels out in this condition, so the value can be set locally arbitrarily from pixel to pixel. Hence $D(z)$ can be designed independently of θ, in agreement with this condition.

To design the $D(z)$, we start our analysis from the 2D separable wavelet. Let $A(z_i)$ and $B(z_i)$ be the lowpass and highpass wavelet filter in one dimension separately. Define $U(x_i) = A(z_i)A(z_i^{-1})$ and $V(x_i) = B(z_i)B(z_i^{-1})$, where $x_i = z_i + z_i^{-1}$. Then from the PR condition of $A(z_i), B(z_i)$ in 2D, we have

$$C = \sum_{M,N \in \{U,V\}} M(x_1)N(x_2)$$
$$= U(x_1)U(x_2) + V(x_1)(U(x_2) + \frac{1}{2}V(x_2)) \tag{9}$$
$$+ V(x_2)(U(x_1) + \frac{1}{2}V(x_1))$$

Compare Eq.9 with Eq. 8, we can design lowpass filter as $H(z) = A(z_1)A(z_2)$, and design $D(z)$ by Factorize following equation,

$$D_1(z)D_1(z^{-1}) = V(x_1)(U(x_2) + \frac{1}{2}V(x_2))$$
$$D_2(z)D_2(z^{-1}) = V(x_2)(U(x_1) + \frac{1}{2}V(x_1)) \tag{10}$$

Theoretically, $A(z_i)$ and $B(z_i)$ can be any type of wavelet, such as Haar, Daubechies, Symlet, etc. Here we use the Haar wavelet as an example. Writing $D(z)$ in following form,

$$D_1(z_1, z_2) = B(z_1)Q(z_2)$$
$$D_2(z_1, z_2) = D_1(z_2, z_1) \tag{11}$$

Where $B(z_i) = 1 - z_i^{-1}$ is a highpass filter that acts as a gradient. And $Q(z_i) = a + bz_i^{-1}$ is a lowpass filter with two variables to solve. D_1 and D_2 are gradients that perpendicular to each other, together formulate the Wirtinger gradient ($D = \frac{\partial f}{\partial z_1} + i\frac{\partial f}{\partial z_2}$) on the complex plane. After simplifications, we find

$$(2 + x_1)(2 + x_2) + (2 - x_1)((a^2 + b^2) + 2abx_2) + (2 - x_2)((a^2 + b^2) + 2abx_1) = C \tag{12}$$

for arbitrary values of x_1 and x_2. The values of a, b, c that ensure this identity are given by $a = 1 + \frac{1}{\sqrt{2}}$, $b = 1 - \frac{1}{\sqrt{2}}$, $C = 16$. This result means that taking any position of the image as the origin of z-transform, we can assign an arbitrary value of θ, and use $D(z)$ as before to formulate $G(z)$. Then, the analysis and synthesis meet the conditions of perfect reconstruction.

3.3. Design of $\theta(x)$ and orientation selection

For natural images with noise, the structure signal is usually continuous, or at least continuous along one direction for the edge part. Therefore, for 2D highpass subbands, locally decomposing images along the edge direction can maximize the distinction between signal and noise.

We parametrize the edge model in natural images as the derivative of an elongated Gaussian function noted as h with two shaping parameters a and b, where $a \ll b$.

$$h(x,y) = \frac{-2x}{\pi a^2 b} \exp\left(-\frac{x^2}{a^2} - \frac{y^2}{b^2}\right) \tag{13}$$

We then utilize the Fourier-Argand representation [23] to represent this function in a steerable way. Eq.13 can then be written as the sum of a series of basis functions H_k,

$$h(x,y) = \sum_{k \in \mathbb{Z}} h_k(|x+iy|)\left(\frac{x+iy}{|x+iy|}\right)^k = \sum_{k \in Z} H_k(x+iy) \tag{14}$$

where h_k is the Fourier series decomposition of h,

$$h_k(r) = \frac{1}{2\pi} \int_0^{2\pi} h(r\cos\theta, r\sin\theta) e^{-ik\theta} d\theta \tag{15}$$

Let the z-transform of the input image be $F(z)$, the output after the filter is then $F^l(z) = F(z)\prod_{k=0}^{l-1} H(z^{2^k})$ We then define the max Fourier-Argand orientation in l level decomposition of the input image f as

$$\theta(x) = \arg\max_{\alpha \in [0, 2\pi)} \sum_{n=1}^{N} (\mathcal{Z}^{-1}\{F^l(z)\} * H_k) e^{-ina} \tag{16}$$

The angle θ is then a robust direction to indicate the local edge orientation.

4. FRAMEWORK OF MC SURE-LET DENOISING

A noisy image is usually modeled as $y = x + b$, where x is the clean image and $b \sim \mathcal{N}(0, \sigma)$ is a Gaussian noise image with statistically independent pixels of variance σ^2. Several levels of wavelet decomposition on the noisy image provide subbands that can then be processed. We write the whole process for one subband as

$$F_k(\mathbf{y}) = g_k^{-1}(\Theta_k(g_k(\mathbf{y}))) \tag{17}$$

where Θ_k is a thresholding function, e.g., $\Theta_k(w(x)) = w(x) \cdot (1 - \exp(-(\frac{w(x)}{3\sigma_k})^8))$. where σ_k is the estimated noise variance of the subband. After filtering, we use the inverse
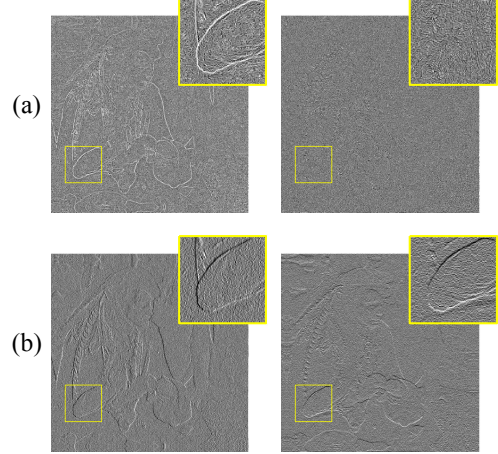

Fig. 1: Visualization of the subbands of the proposed method, with(a) and without(b) the Fourier-Argand orientation selection.

filter to get the response of the subband. The final denoising result and its matrix form can be written as

$$\mathbf{F}(\mathbf{y}) = \sum_k a_k F_k(\mathbf{y}) = \mathbf{F}\mathbf{a} \tag{18}$$

Where a_k is the coefficient of the linear combination of subbands. The expected error between the final result and the clean image can be minimized by optimizing the coefficients in $\mathbf{a}$. The optimization target follows [3] is,

$$\text{SURE} = \frac{1}{N}||\mathbf{y} - \mathbf{F}(\mathbf{y})||^2 - \sigma^2 + \frac{2\sigma^2}{N}\text{div}_\mathbf{y}\{\mathbf{F}(\mathbf{y})\} \tag{19}$$

Using the Monte-Carlo method[14] to estimate the divergence term, we know that $\text{div}_\mathbf{y}\{\mathbf{F}(\mathbf{y})\} \approx \frac{1}{\epsilon N}\tilde{n}^T(\mathbf{F}(\mathbf{y} + \epsilon\tilde{n}) - \mathbf{F}(\mathbf{y}))$. Where N is the number of pixels of the image. Since $\mathbf{F}(\mathbf{y}) = \mathbf{F}\mathbf{a}$, we have,

$$\text{SURE} \approx \frac{1}{N}||\mathbf{y}^T\mathbf{y} - 2\mathbf{a}^T\mathbf{F}^T\mathbf{y} + \mathbf{a}^T\mathbf{F}^T\mathbf{F}\mathbf{a}||^2 - \sigma^2$$
$$+ \frac{2\sigma^2}{\epsilon N^2}\tilde{n}^T(\mathbf{F}(\mathbf{y} + \epsilon\tilde{n}) - \mathbf{F}(\mathbf{y})) \tag{20}$$

This quadratic function can then minimized by calculating $\nabla_\mathbf{a}\text{SURE} = 0$, leading to

$$\mathbf{a} = (\mathbf{F}^T\mathbf{F})^{-1}(\mathbf{F}^T\mathbf{y} - \frac{\sigma^2}{\epsilon N}\tilde{n}^T(\mathbf{F}(\mathbf{y} + \epsilon\tilde{n}) - \mathbf{F}(\mathbf{y}))) \tag{21}$$

Note that this form does not require the wavelet transformation in $\mathbf{F}(\mathbf{y})$ to satisfy the linearity, so with the Monte-Carlo technique, we can easily apply FA-CWT in the SURE-LET framework.

5. ANALYSIS AND EXPERIMENTS

We visually plot the wavelet responses of the first level, shown in Fig.1. The upper two figures (a) show the channel's real and imaginary parts using the proposed transform. The lower two figures (b) are the real and imaginary parts after using the Wirtinger gradient filter $D(z)$.
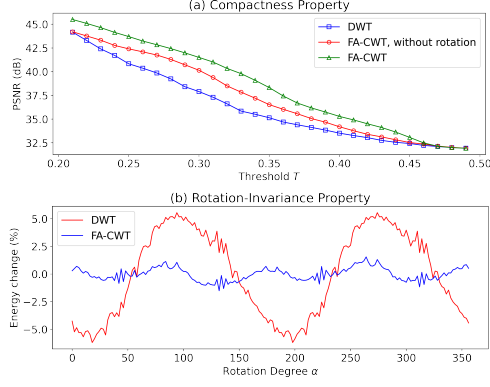
Fig. 2: (a) Proposed method can retain more signals when the same threshold is used. (b) Proposed method stays stable with different rotation degrees, whereas conventional wavelet oscillates.

5.1. Compactness property

From Fig.1, we can intuitively observe that signals can be gathered on the real part of the subband more compactly with fewer filters, while noises still evenly distributed in each channel. Therefore, our filter can concentrate signals enery in fewer pixels. We verify This property by using different thresholds T for each subband of the clean image and use PSNR to measure in what degree subbands decays. The result is shown in Fig.2 (a).

5.2. Rotation invariance property

Rotation invariance is a property that is often overlooked. A similar property is shift-invariance, in which people usually use cycle-spinning or dual-tree method to achieve this property[2]. In contrast, the similar property but on rotation is rarely investigated. Since we decompose highpass subbands along edge direction, rotate image will only result in a rotation of subbands, while illuminate stays still. (See Fig. 1(a)).

We test this property as follow: Get the subbands of the rotated image $w_\alpha(x)$ We then calculate the changes in signal energy as $E_\alpha = \frac{\langle w_\alpha(x), w_\alpha(x) \rangle}{\frac{1}{N} \sum_\theta \langle w_\theta(x), w_\theta(x) \rangle} - 1$. Results in Fig.2(b) show that the proposed method stays stable with various angles, while wavelet results oscillates.

5.3. Comparison with other transforms

We compare our method, the Fourier-Argand Wavelet Transform (FA-CWT), with other wavelet transforms, i.e., Undecimated Discrete Wavelet Transform (UDWT), Dual-tree complex wavelet transform (DTCWT)[16] and Non-subsampled shearlet transform (NSST)[12]. Each transform uses the same pipeline described as following: (1). Decompose noisy image by the transform. (2). Threshold each subband by $n\sigma$, where n is selected to optimize the performance. (3). Reconstruct the image from the denoised subbands. The results are

Table 1: Comparison between different transforms. Our method achieves competitive results (in PSNR/dB) with the use of fewer subbands. $\mathbb{R}$ and $\mathbb{C}$ means the subbands are real or complex type.

data	σ / PSNR	Methods (Number of subbands)			
		UDWT ($12\,\mathbb{R}$)	DTCWT ($24\,\mathbb{C}$)	NSST ($40\,\mathbb{R}$)	FA-CWT ($4\,\mathbb{C}$)
boat	25 / 20.32	28.25	27.22	27.59	**28.31**
	50 / 14.62	23.9	24.02	24.87	**25.33**
	75 / 11.75	21.41	22.03	22.76	**23.34**
	100 / 10.13	19.83	20.67	21.11	**21.74**
	150 / 8.45	17.89	18.83	18.9	**19.55**
man	25 / 20.28	28.37	27.25	27.83	**28.62**
	50 / 14.68	24.21	24.47	25.29	**25.91**
	75 / 11.84	21.67	22.43	23.14	**23.86**
	100 / 10.19	19.97	20.91	21.33	**22.06**
	150 / 8.47	17.83	18.81	18.83	**19.53**

Table 2: Results using FA-CWT in SURE-LET (PSNR/dB)

Methods	#bands	σ				
		25	50	75	100	150
SURE-LET	$12\,\mathbb{R}$	20.26	14.63	11.80	10.16	8.46
	$18\,\mathbb{R}$	29.23	24.64	21.96	20.20	18.05
SURE-LET with FA-CWT	$3\,\mathbb{C}$	30.25	25.84	23.04	21.07	18.60
	$4\,\mathbb{C}$	29.83	26.48	23.93	21.91	19.27

shown in Tab.1. We constrained each method to four levels to make methods comparable. The proposed method represents the high-frequency information only by 4 complex subbands, while others need 12/24/40 real or complex subbands. Meanwhile, our method brings significant improvements, especially in high noise levels.

5.4. Application of FA-CWT for SURE-LET denoising

We use FA-CWT as the transform in the SURE-LET denoising framework; i.e., we use Stein's Unbiased Risk Estimate (SURE) as the criterion to estimate the optimal coefficients of the denoising process. Since FA-CWT is a non-linear transform, we have to rely on the Monte-Carlo method to estimate the SURE. We compare results on *lena* with and without FA-CWT. The results in Fig.2 show that replacing the wavelet transform with FA-CWT improves the denoising performance significantly, while requiring much fewer subbands.

6. CONCLUSION

This paper proposed a steerable wavelet transform that can adaptively decompose images based on edge features. The method achieves rotation-invariance properties in a more concise way. We apply it to denoising to show its efficiency. In the future, we will explore it in other image processing tasks, such as image super-resolution, deconvolution, edge enhancement.

References

[1] Mariví Tello Alonso, Carlos López-Martínez, Jordi J Mallorquí, and Philippe Salembier. Edge enhancement algorithm based on the wavelet transform for automatic edge detection in sar images. *IEEE transactions on geoscience and remote sensing*, 49(1):222–235, 2010.

[2] Adriaan Barri, Ann Dooms, and Peter Schelkens. The near shift-invariance of the dual-tree complex wavelet transform revisited. *Journal of mathematical analysis and applications*, 389(2):1303–1314, 2012.

[3] Thierry Blu and Florian Luisier. The sure-let approach to image denoising. *IEEE Transactions on Image Processing*, 16(11):2778–2786, 2007.

[4] Emmanuel J Candès and David L Donoho. Ridgelets: A key to higher-dimensional intermittency? *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 357(1760):2495–2509, 1999.

[5] Ingrid Daubechies. *The wavelet transform, time-frequency localization and signal analysis*. Princeton University Press, 2009.

[6] Ingrid Daubechies and Wim Sweldens. Factoring wavelet transforms into lifting steps. *Journal of Fourier analysis and applications*, 4(3):247–269, 1998.

[7] Minh N Do and Martin Vetterli. The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on image processing*, 14(12):2091–2106, 2005.

[8] Zhijun Fang, Naixue Xiong, Laurence T Yang, Xingming Sun, and Yan Yang. Interpolation-based direction-adaptive lifting dwt and modified spiht for image compression in multimedia communications. *IEEE systems journal*, 5(4):584–593, 2011.

[9] Omer N Gerek and A Enis Çetin. A 2-d orientation-adaptive prediction filter in lifting structures for image coding. *IEEE Transactions on Image Processing*, 15(1):106–111, 2005.

[10] Kanghui Guo, Gitta Kutyniok, and Demetrio Labate. Sparse multidimensional representations using anisotropic dilation and shear operators, 2006.

[11] Qiangui Huang, Boya Hao, and Sheng Chang. Adaptive digital ridgelet transform and its application in image denoising. *Digital Signal Processing*, 52:45–54, 2016.

[12] Gitta Kutyniok and Demetrio Labate. *Shearlets: Multiscale analysis for multivariate data*. Springer Science & Business Media, 2012.

[13] Stéphane Mallat. *A wavelet tour of signal processing*. Elsevier, 1999.

[14] Sathish Ramani, Thierry Blu, and Michael Unser. Monte-carlo sure: A black-box optimization of regularization parameters for general denoising algorithms. *IEEE Transactions on image processing*, 17(9):1540–1554, 2008.

[15] Sachin D Ruikar and Dharmpal D Doye. Wavelet based image denoising technique. *IJACSA) International Journal of Advanced Computer Science and Applications*, 2(3), 2011.

[16] Ivan W Selesnick, Richard G Baraniuk, and Nick C Kingsbury. The dual-tree complex wavelet transform. *IEEE signal processing magazine*, 22(6):123–151, 2005.

[17] R Sethunadh and Tessamma Thomas. Spatially adaptive image denoising using inter-scale dependence in directionlet domain. *IET Image Processing*, 9(2):142–152, 2015.

[18] Eero P Simoncelli and William T Freeman. The steerable pyramid: A flexible architecture for multi-scale derivative computation. In *Proceedings., International Conference on Image Processing*, volume 3, pages 444–447. IEEE, 1995.

[19] Jean-Luc Starck, Emmanuel J Candès, and David L Donoho. The curvelet transform for image denoising. *IEEE Transactions on image processing*, 11(6):670–684, 2002.

[20] Wim Sweldens. The lifting scheme: A construction of second generation wavelets. *SIAM journal on mathematical analysis*, 29(2):511–546, 1998.

[21] Michael Unser. Texture classification and segmentation using wavelet frames. *IEEE Transactions on image processing*, 4(11):1549–1560, 1995.

[22] Vladan Velisavljevic, Baltasar Beferull-Lozano, Martin Vetterli, and Pier Luigi Dragotti. Directionlets: anisotropic multidirectional representation with separable filtering. *IEEE Transactions on Image Processing*, 15(7):1916–1933, 2006.

[23] Tianle Zhao and Thierry Blu. The fourier-argand representation: An optimal basis of steerable patterns. *IEEE Transactions on Image Processing*, 29:6357–6371, 2020.