EXTENDED CONJUGATE POLAR FOURIER TRANSFORM IN CONVOLUTION NETWORK

Can Xu *Wenrui Dai* *Hongkai Xiong**

Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

ABSTRACT

This paper proposes an extended conjugate polar Fourier transform (ECPFT), to design iterated radial filter bank (RFB) and directional filter bank (DFB) convenient for accurate multiscale and multidirectional decomposition in discretization over a convolution network. With conjugated symmetric form, ECPFT would convert complex directional wavelets in original spatial domain to real ones in the inverse Fourier domain of ECPFT. Furthermore, it can contribute to changing the design of nonseparable RFB and DFB to decomposition in scales and directions with 1-D filter bank in each dimension separably in the ECPFT domain. The generated properties of conjugate symmetry and periodicity of 2π in both angular and radial dimensions, also guarantee convolution and sampling operations in both dimensions within the inverse Fourier domain of ECPFT. The fast algorithm for discretization is developed to reduce complexity.

***Index Terms*—** Polar Fourier transform, directional filter bank, wavelet, curvelet, convolution network

1. INTRODUCTION

The paces of signal processing has moved ahead since the ability to efficiently approximate signals containing pointwise singularities achieves success from Fourier analysis to wavelet analysis [1]. When applied to two or higher dimensional signals, specially with singularity of line or surface, the separable wavelets have a bias towards horizontal and vertical directions and leads to block effect. In fact, the 2-D Fourier transform can be realized in terms of polar coordinates instead of the usual Cartesian coordinates, which is extremely useful when analyzing directional information [2, 3].

Multiscale geometric analysis (MGA) to construct an optimal representation of high-dimensional function has ever become a significant subject, e.g. ridgelet [4], curvelet [5]. The second generation of curvelet [6] transform is initially constructed in frequency domain on the polar coordinate which is separable in angles and radius. To be concrete, Candes and Donoho only construct a curvelet in polar Fourier domain and then convert it back to the Cartesian coordinates in spatial domain for convolution with signal. This allows for good sensitivity to directions and simple construction in the continuous domain. However, it leads to problems in discrete images because of the mismatch of rectangular sampling grid of images and the polar coordinate. It indicates a lack of complete interpretation of the polar version of standard convolution. This fact motivates the development of a new directional multiresolution transform directly in the discrete domain - the contourlet transform [7]. The directional filter bank (DFB) is applied to the high frequency components to proceed directional decomposition. It is worth mentioning that the design of 2-D filters with different shapes and corresponding sampling matrix is not an easy work. A modified pseudo-polar Fourier transform was proposed in [8, 9], to help design directional filter banks. Unlike the aforementioned effort, it enables convolution in the angular dimension, while disregards the desirable convolution and sampling in the radial dimension. Apart from MGA, several efforts focus on efficient representation of geometrical regularity, e.g. bandelet [10], wedgelet [11], and shearlet [12]. Their limitation lies in that they lack finer decomposition of directional high-frequency subbands. Impressively, a scattering transform has been proposed [13] which is a deep convolutional network with cascades of wavelet filters and average operators. The decomposition of one layer contains a low frequency part and several high frequency parts with different scales and directions. Furthermore, these filters are iterated layer by layer on high frequency parts to implement finer frequency partition. This differs from aforementioned several transforms in which directional subbands will not be applied to multiscale decomposition later.

To achieve the accurate multiscale and multidirectional decomposition in discretization over a convolution network, this paper proposes an extended conjugate polar Fourier transform (ECPFT), to make iterated radial filter bank (RFB) and directional filter bank (DFB) convenient. The ECPFT which is supposed to be conjugated symmetric, would convert complex directional wavelets in original spatial domain to real ones in the inverse Fourier domain of ECPFT. Also, it can contribute to change the design of nonseparable RFB and DFB, which conduct decomposition in scales and directions, to that of 1-D filter bank in each dimension separably in the ECPFT domain. The generated properties of conjugate

*The work has been partially supported by NSFC under Grant 61501294, No. 61529101, Grant 61425011, the China Postdoctoral Science Foundation under Grant 2015M581617, and the Program of Shanghai Academic Research Leader under Grant 17XD1401900.

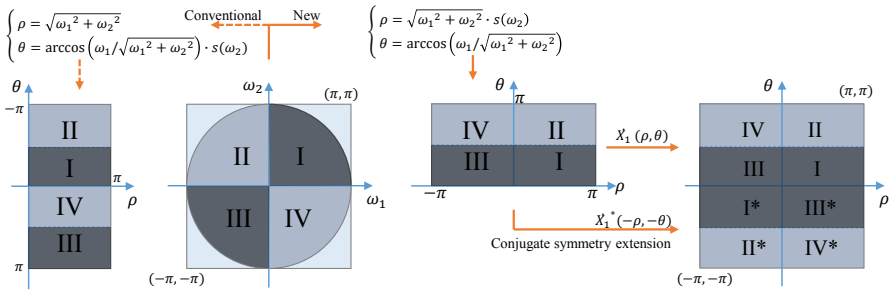

Fig. 1. The four sub-figures from left to right are: conventional polar Fourier transform (PFT) domain, Fourier transform (FT) domain, conjugated polar Fourier transform (CPFT) domain and extended conjugate polar Fourier transform (ECPFT) domain.

symmetry and periodicity of 2π in both angular and radial dimensions, also guarantee convolution and sampling operations in both dimensions in the inverse Fourier domain of ECPFT. To the best of our knowledge, this has never been achieved in the literature.

2. EXTENDED CONJUGATE POLAR FOURIER TRANSFORM IN FILTER BANKS

The complicated process of constructing multiscale and multidirectional filter banks can be simplified via the combination of ECPFT and 1-D filter banks. The ECPFT is composed of the following three steps. i) Fourier transform ii) conjugated polar conversion iii) conjugate symmetric extension and periodization. In particular, the underlying second and third steps differ from conventional PFT or CPFT, make it distinguished.

2.1. Conjugate Polar Conversion

We take coordinate conversion to two-dimensional frequency $(\omega_1, \omega_2) \in [-\pi, \pi] \times [-\pi, \pi]$. The conventional polar conversion and proposed conjugate polar conversion formulas are given for comparison in (1) and (2). Applying the two conversions to Fourier transform would lead to polar Fourier transform (PFT) and conjugate polar Fourier transform (CPFT).

$$\begin{cases} \omega_1 = \rho\cos\theta \\ \omega_2 = \rho\sin\theta \end{cases} \quad \begin{cases} \rho = \sqrt{\omega_1^2+\omega_2^2} \\ \theta = \arccos\left(\omega_1/\sqrt{\omega_1^2+\omega_2^2}\right)\cdot s(\omega_2), \end{cases} \tag{1}$$

$$\begin{cases} \omega_1 = \rho\cos\theta \\ \omega_2 = \rho\sin\theta \end{cases} \quad \begin{cases} \rho = \sqrt{\omega_1^2+\omega_2^2}\cdot s(\omega_2) \\ \theta = \arccos\left(\omega_1/\sqrt{\omega_1^2+\omega_2^2}\right), \end{cases} \tag{2}$$

where

$$s(\omega) = \begin{cases} 1, & \text{if } \omega \geq 0 \\ -1, & \text{others.} \end{cases}$$

The Fourier transform of a signal $x(u_1, u_2)$ is

$$X(\omega_1,\omega_2) = \sum_{u_1,u_2\in\mathbb{Z}^2} x[u_1,u_2]e^{-j\omega_1 u_1}e^{-j\omega_2 u_2}. \tag{3}$$

Apply (2) to (ω_1, ω_2) in (3) derives the CPFT:

$$\begin{aligned} \dot{X}(\rho,\theta) &= X(\rho\cos\theta, \rho\sin\theta) \\ &= \sum_{u_1,u_2\in\mathbb{Z}^2} x[u_1,u_2]e^{-j\rho\cos\theta u_1}e^{-j\rho\sin\theta u_2} \end{aligned} \tag{4}$$

The main difference between PFT and CPFT is the definition domain. In conventional case, $(\rho, \theta) \in [0, \pi] \times [-\pi, \pi]$. If the input signal $f(u_1, u_2)$ is real in spatial domain, then it is conjugate symmetric in frequency domain and conjugate periodic by π in polar Fourier domain:

$$X_p(\rho,\theta) = X_p^*(\rho,\theta+\pi), \tag{5}$$

where X_p denotes the PFT of signal x.

The definition domain of CPFT is:

$$\{(\rho,\theta)|(\rho,\theta)\in[-\pi,\pi]\times[0,\pi]\}. \tag{6}$$

The conjugate symmetry corresponds to the conjugate symmetry about axis θ:

$$\dot{X}(\rho,\theta) = \dot{X}^*(-\rho,\theta). \tag{7}$$

This property is desirable if we take one-dimensional filtering/convolution operations along the ρ-dimension (radial dimension) with a real impulse response function. Moreover, this operation is equivalent to 2-D convolution in the original spatial domain. This can be better understood by an example in Fig. 1. The fan-shaped pair of frequency subbands of directions between 0 and $\pi/2$ are mapped to two rectangular subbands separated by a distance π in the PFT domain, which satisfies (5). If mapped to the CPFT domain, it corresponds to one rectangular subband which are conjugate symmetric about ρ axis, as represented in (7).

Since the definition domain of CPFT in (6) corresponds to only a part of the whole frequency domain, $X(\omega_1, \omega_2)$ is thereby divided into two parts:

$$\begin{aligned} X(\omega_1,\omega_2) &= X_1(\omega_1,\omega_2) + X_2(\omega_1,\omega_2) \\ X_1(\omega_1,\omega_2) &= \begin{cases} X(\omega_1,\omega_2), & \text{if } \sqrt{\omega_1^2+\omega_2^2} < \pi \\ 0, & \text{else} \end{cases} \end{aligned} \tag{8}$$

Considering definition domain, $\dot{X}$ equals to $\dot{X}_1$.

2.2. Conjugate extension and periodization

Convolution demands periodicity in frequency domain, thus we extend the definition domain of CPFT from $[-\pi, \pi]\times[0, \pi]$ to $[-\pi,\ \pi] \times [-\pi,\ \pi]$ by conjugate symmetry to make it have a period of 2π along each dimension, maintaining the period in accordance with that of FT of 1-D digital signals. The extended CPFT (ECPFT) of x_1 is denoted by:

$$\ddot{X}_1(\rho,\theta) = \begin{cases} \dot{X}_1(\rho,\theta), & \text{if } 0 \le \theta \le \pi \\ \dot{X}_1^*(-\rho,-\theta), & \text{if } -\pi \le \theta < 0. \end{cases} \tag{9}$$

Note to pay attention that values of $\ddot{X}_1(\rho,\theta)$ for $\theta < 0$ do *not* satisfy (4) if replacing $\dot{X}(\rho,\theta)$ with $\ddot{X}_1(\rho,\theta)$. Taking inverse Fourier transform of $\ddot{X}_1(\rho,\theta)$ would convert it from the ECPFT domain to the inverse ECPFT domain, which differs with the original spatial domain. The signal in the inverse ECPFT domain, denoted by $\ddot{x}_1[v_1, v_2]$, can be calculated by:

$$\ddot{x}_1[v_1,v_2] = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \ddot{X}_1(\rho,\theta)e^{j\rho v_1}e^{j\theta v_2}d\rho d\theta. \tag{10}$$

Since $\ddot{X}_1(\rho,\theta)$ is conjugate symmetric if the original is real, then $\ddot{x}_1(v_1, v_2)$ is real. It guarantees the computational efficiency by avoiding the trouble of calculating with complex numbers. $\ddot{x}_1(v_1, v_2)$ can also be obtained by sampling $\ddot{X}_1(\rho,\theta)$ and taking DFT. We uniformly sample $\ddot{X}_1(\rho,\theta)$ by the same interval in each dimension. The sampled value can obtained by plugging $\rho = \frac{2\pi}{N_1}n_1, \theta = \frac{2\pi}{N_2}n_2$ into (10) and (9).

2.3. Fast Approximation

To reduce computation complexity, we develop a fast approximate algorithm (FAA) to calculate $\ddot{x}_1[v_1, v_2]$ with minor errors, along with an inverse fast algorithm (IFAA) to calculate $x[u_1, u_2]$. Fig. 2 (a) and (b) provide two analysis and synthesis diagrams of ECPFT: direct discrete conversion and perfect reconstruction.

Suppose the input two-dimensional signal is in size of $N_1 \times N_2$ (N_1, N_2 is the integer power of 2). Taking FFT we get the same size $X[m_1, m_2]$ which are sampled values of $X(\omega_1, \omega_2)$. The circular area with a radius of π corresponds to the upper area of $\ddot{X}_1(\rho,\theta)$, i.e. the whole area of $\dot{X}_1(\rho,\theta)$. In order to preserve energy, we take $N_1 \times N_2$ samples of $\dot{X}_1(\rho,\theta)$ uniformly. We denote the sampled value $\dot{X}_1(\frac{2\pi}{N_1}n_1, \frac{\pi}{N_2}n_2)$ by $\dot{X}_1[n_1, n_2]$. Instead of directly calculating $\dot{X}_1(\frac{2\pi}{N_1}n_1, \frac{\pi}{N_2}n_2)$ through (4), we could: 1) map its coordinate $(\frac{2\pi}{N_1}n_1, \frac{\pi}{N_2}n_2)$ into the Cartesian coordinate, 2) find the nearest point $(\frac{2\pi}{N_1}m_1, \frac{2\pi}{N_2}m_2)$, and then 3) take $X[m_1, m_2]$ as its value.

2.4. Designing Filter Banks

Fig. 3 illustrates two types of filter banks as the foundation of the underlying convolution network. The support configurations for decomposition filter pairs are shown in Fig. 4,

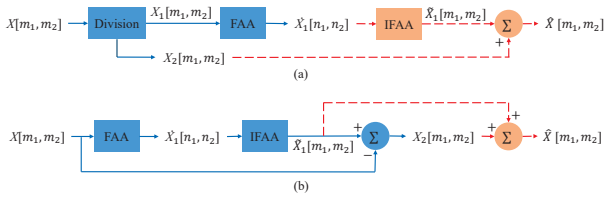

Fig. 2. (a) The direct discrete implementation of coordinate conversion. (b) The scheme guarateeing perfect reconstrction.

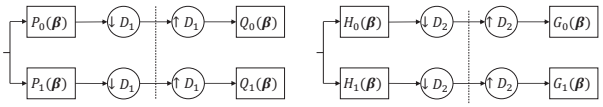

Fig. 3. RFB (left) and DFB (right). The Left side of each filter bank is the analysis side for decomposition. The right side of each filter bank is the synthesis side for reconstruction.

with their sampling matrix given in (11). RFB corresponds to multiscale decomposition and DFB the multidirection decomposition. Though described in 2-D form, it is in practice implemented by 1-D convolution and sampling along each dimension, respectively.

$$D_1 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \tag{11}$$

For a better understanding, we take an example of 2-level 2-channel RFB and DFB in Fig. 5. The frequency supports of the four scale and four directional subbands indexed from 0 to 3 are shown in red and blue boxes, respectively. We only take one filtering operation once a time to divide the low-frequency and high-frequency part. The subbands are obtained by filtering $\ddot{X}_1(\rho,\theta)$ with a 2-level 1-D FB along the ρ dimension and the θ dimension. Finer directional decompositon can obtained by iterating the same filter pair and downsamping factor layer by layer. None undesirable subbands are obtained in comparison to 2-D wavelet decomposition which demands twice filtering along each dimension for each layer.

3. EXPERIMENTS

For validation, we make ECPFT in images, followed by the inverse Fourier transform, so as to obtain the representation in the inverse ECPFT domain. Take convolution and sampling with 1-D filters along each dimension, respectively, to accomplish image decomposition in scales and angles. For reconstruction, we perform upsampling and filtering, followed by an inverse process of ECPFT. Twelve gray scale images of 512×512 or 256×256 in size are used for testing, and 3% most significant coefficients are used for approximation.

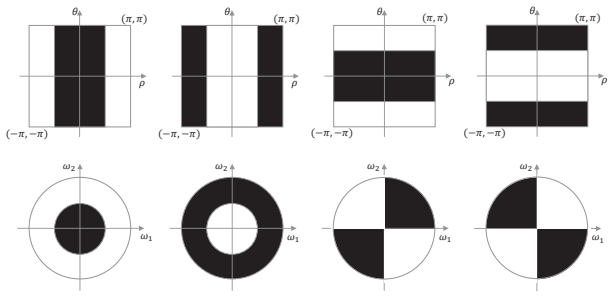

Fig. 4. Support configurations for the filters in the RFB and DFB. The top line is the frequency view of new polar coordinate and the bottom line corresponds to the frequency domain of the original coordiante. The first and second columns are the support configurations of radial filter pair $P_0[\boldsymbol{\beta}], P_1[\boldsymbol{\beta}]$; The third and fourth columns are the support configurations of directional filter pair $H_0[\boldsymbol{\beta}], H_1[\boldsymbol{\beta}]$.

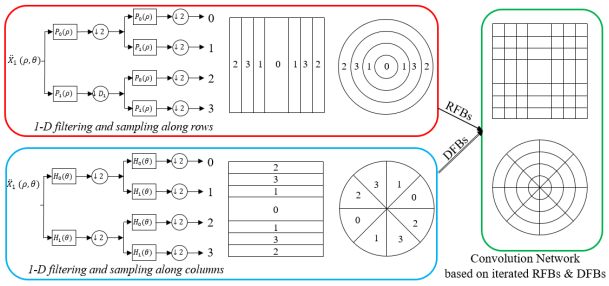

Fig. 5. An example of a full convolution network with RFBs and DFBs of depth $(l_1, l_2) = (2, 2)$.

3.1. Fast approximation

Since the input signal is divided into two parts x_1 and x_2, x_1 is first recovered from decomposition coefficients and then combined with x_2 to obtain the reconstructed signal. In continuous case, the order of division and ECPFT of x_1 makes no difference. In discrete case, different orders have different effects, as shown in Fig. 6 with different visual effects from Fig. 2 (a) and (b). Table 1 depicts the the energy occupation of X_2 are less than 1.2% for all four sizes, which shows the values decrease with the increase of image size.

3.2. Nonlinear Approximation

Considering that curvelet transform uses PFT to construct curvelet functions, we compare their performance of image representation by nonlinear approximation (NLA), using wavelet and contourlet transforms as reference. For the 12 test images, we select 3% most significant coefficients in each transform domain, and then compare the reconstructed images from these sets of coefficients. Table.2 shows that the proposed structure achieves the best performance among three methods for all testing images. Especially in *barbara* and *fingerprint*, there is a significant gain of 0.48 dB and 0.71 dB in peak signal-to-noise ratio (PSNR), respectively.

Fig. 6. Reconstrction. Leftmost: the original image; Middle left: reconstructed by removing $X_2[m_1, m_2]$; Middle right and rightmost are reconstructed images with two coordinate conversion schemes in Fig. 2 (a) and (b), respectively.

Table 1. Energy ratio and Running time of calculating $\dot{X}_1[n_1, n_2]$ from both direct formula and fast approximation

size	128	256	512	1024
energy ratio $\|\|x_2\|\|^2/\|\|x\|\|^2$	0.66%	0.66%	0.21%	0.09%
Running time of (4)(10)	0.77	8.70	92.08	1672.96
Running time of FAA	0.02	0.08	0.31	1.29

Table 2. PSNR of reconstructed images with different transforms

PSNR(dB)	wavelet	contourlet	curvelet	proposed
barbara	17.68	22.28	21.97	**22.86**
cameraman	9.36	20.87	20.48	**20.93**
peppers	17.97	25.33	24.38	**26.70**
fingerprint	17.83	17.04	17.22	**17.93**
lena	17.87	25.39	24.95	**25.41**
baboon	15.57	21.97	21.24	**22.06**
bridge	16.00	21.65	20.62	**22.01**
lax	19.16	22.05	21.50	**22.21**
woman	19.43	25.04	24.17	**25.95**
crowd	19.68	23.08	22.22	**23.92**
milkdrop	22.26	27.41	27.37	**27.75**
plane	18.92	22.78	23.46	**23.72**

4. CONCLUSION

In order to achieve accurate multiscale and multidirectional decomposition in discretization over a convolution network, this paper proposed an extended conjugate polar Fourier transform which is conjugate symmetric and periodic. These properties are guaranteed by conjugate polar conversion, extension and periodization operations. It would help change the design of nonseparable RFB and DFB to decomposition in scales and directions with 1-D filter bank in each dimension separably in the ECPFT domain. Also, it guarantees convolution and sampling operations in both dimensions within the inverse Fourier domain of ECPFT.

5. REFERENCES

[1] S. Mallat, "A wavelet Tour of Signal Processing: the sparse way." New York: Academic, 2008.

[2] A. Averbuch, R. Coifman, D. Donoho, M. Elad, and M. Israeli, "Accurate and fast discrete polar fourier transform," in Proc. Thirty-Seventh Asilomar Conference on Signals, Systems and Computers, 2003, pp. 1933-1937.

[3] Q. Wang, O. Ronneberger, and H. Burkhardt, "Rotational invariance based on fourier analysis in polar and spherical coordinates," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 31, no. 9, pp. 1715-1722, 2009.

[4] M. N. Do and M. Vetterli, "The finite ridgelet transform for image representation," *IEEE Trans. on Image Processing*, vol. 12, no. 1, pp. 16-28, 2003.

[5] J.-L. Starck, E. J. Cand'es, and D. L. Donoho, "The curvelet transform for image denoising," *Image Processing, IEEE Trans. on*, vol. 11, no. 6, pp. 670-684, 2002.

[6] E. Candes and D. Donoho "New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities," *Communications on Pure and Applied Mathematics*, vol. 57, no. 2, pp. 219-266, Feb. 2004.

[7] M. Do, and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Trans. Image Process*, vol. 14, no. 12, pp. 2091-2106, Dec. 2005.

[8] G. Shi, L. Liang, and X. Xie, "Design of directional filter banks with arbitrary number of subbands," *IEEE Trans. on Signal Processing*, vol. 57, no. 12, pp. 4936-4941, 2009.

[9] L. Liang, G. Shi, and X. Xie, "Nonuniform directional filter banks with arbitrary frequency partitioning," *IEEE Trans. on Image Processing*, vol. 20, no. 1, pp. 283-288, 2011.

[10] E. Pennec and S. Mallat "Sparse geometric image representation with bandelets," *IEEE Trans. Image Process*, vol. 14, no. 4, pp. 423-438, Apr. 2005.

[11] J. K. Romberg, M. Wakin, and R. Baraniuk, "Multiscale wedgelet image analysis: fast decompositions and modeling, in *Proc. IEEE Intl. Conference on Image Processing*, 2002, pp. 585-588.

[12] W.-Q. Lim, "The discrete shearlet transform: A new directional transform and compactly supported shearlet frames," *Image Processing, IEEE Trans. on*, vol. 19, no. 5, pp. 1166-1180, 2010.

[13] J. Bruna and S. Mallat "Invariant scattering convolution networks," *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 35, no. 8, pp. 1872-1886, 2013.