USING PHASE AND MAGNITUDE INFORMATION OF THE COMPLEX DIRECTIONAL FILTER BANK FOR TEXTURE IMAGE RETRIEVAL

An P.N. Vo, Soontorn Oraintara

University of Texas at Arlington,
Department of Electrical Engineering,
Arlington, TX 76019–0016

Truong T. Nguyen

Paris School of Mines,
Geophysic Research Center,
Fontainebleau, 77305 France

ABSTRACT

This paper discusses how to utilize both magnitude and phase information obtained from the complex directional filter bank (CDFB) for the purpose of texture image retrieval. The relative phase, which is the difference of phases between adjacent CDFB coefficients, has a linear relationship with the angle of dominant orientation within a subband. This information is incorporated to form a new feature vector called CDFB-RP. Texture retrieval performance of the proposed CDFB-RP is compared to those of the conventional transforms including the Gabor wavelet, the contourlet transform, the steerable pyramid and the CDFB. With the same number of features, the CDFB-RP method outperforms all other transforms in texture image retrieval, while keeping lower complexity and computational time.

Index Terms— Directional Filter Bank, Dual-tree, relative phase, texture classification, texture retrieval.

1. INTRODUCTION

The image retrieval problem has recently become more important and necessary because of the rapid growth of multimedia databases and digital libraries. Different search engines use different features to retrieve images. This paper discusses how magnitude and phase information of the complex directional filter bank (CDFB) [1] coefficients can be used to classify texture images.

One of the approaches to texture feature extraction is the filter bank approach in which texture images are decomposed into subbands using a linear transform or filter bank (FB). In [2], texture classification performances of various FB methods are compared, and the conclusion is that no method performs well in all kinds of textures. Several previous works extract texture features based on wavelet packet signatures [3] and tree-structured wavelet transform [4]. Although these methods allow for a multiresolution decomposition, they are limited in directional selectivity and may not be suitable for images with geometric information such as textures.

The 2-D Gabor transform is a directional decomposition that optimally achieves joint resolution in space and in spatial frequency. In [5], the retrieval performance of the Gabor wavelet feature and other multiresolution features such as wavelet transform and tree-structured wavelet transform are compared. The comparisons indicate that the Gabor feature yields the best correct texture retrieval rate. Despite the high performance compared to others, the 2-D Gabor transform produces an over complete representation for images and is very computationally intensive.

Many other multiresolution multi-directional image representations like the steerable pyramid [6], contourlet [7] and complex wavelets [8] have been used to calculate feature vectors. The basic filters of the steerable pyramid are translations and rotations of a single function, and a filter at any orientation can be computed as a linear combination of the basis filters. This property can be used in rotation invariant texture retrieval [9]. The main advantage of the complex wavelet decomposition is that it can produce texture features which are more robust to translation of the image [8].

Recently, the complex directional filter bank is proposed for texture image retrieval [1]. The texture image retrieval performance of the CDFB is close to the 2-D Gabor transform. Although the CDFB yields both magnitude and phase information in the complex coefficients, only the magnitude has been used in the classification process. In this paper, we discuss how phase information can be incorporated explicitely to further improve the classification rate.

The paper is organized as follows. In the next section, the CDFB is briefly reviewed. The linear relationship between the relative phase of the CDFB and the angle of dominant orientation of texture is discussed in Section 3. The procedure to retrieve and classify texture images and the experimental results are presented in Section 4. We discuss the results and conclude the paper in Section 5.

2. THE COMPLEX DIRECTIONAL FILTER BANK

The (energy) shiftable CDFB is a new image decomposition, which is recently introduced in [10]. The CDFB, shown in Fig. 1(a), is an interative multiscale and multidirection FB. Each resolution level consists of a two-channel FB and a pair of directional filter banks (DFBs). The purpose of the CDFB is to provide a shiftable and scalable multiresolution directional decomposition. This transform has some similarities with the complex version of the shiftable pyramid [6], while maintaining a much lower overcomplete ratio.

According to Fig. 1(a), the input image is first smoothed by the lowpass filter $L_0(\boldsymbol{z})$ before passing through the first level of a multiresolution FB. This two-channel FB has two filters, highpass $R_1(\boldsymbol{z})$ and lowpass $L_1(\boldsymbol{z})$. Slices of the frequency responses of these filters at $\omega_2 = 0$ are illustrated in Fig. 1(b) and (c). The high frequency component at the output of filter $R_1(\boldsymbol{z})$ is then decomposed by the dual DFBs, resulting in the highest resolution directional decomposition. The low frequency component, after decimation by $\boldsymbol{D}_2 = 2\boldsymbol{I}$, is fed into the second level decomposition for the second resolution. The block P shows one level of the CDFB, where the 2×2^n decimated outputs of the two DFBs are the real and imaginary parts of 2^n complex-valued subbands. For more detail of the construction of the CDFB, the reader is referred to [10].

The most important property of the CDFB is that all complex directional subbands are shiftable in the sense that there is no significant aliasing in the decimated complex subbands. Therefore, each

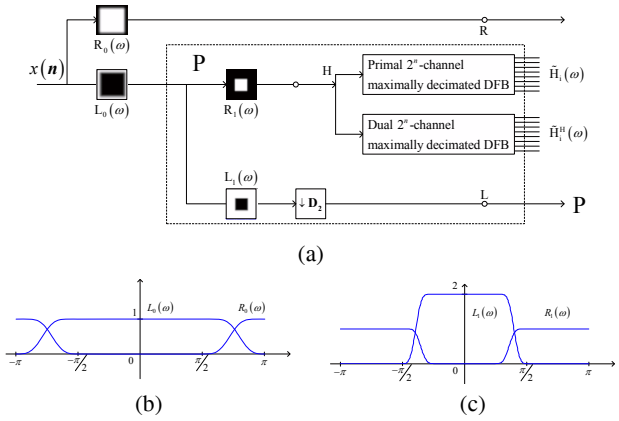


Fig. 1. (a) The FB implemented the pyramidal CDFB image representation. The block P is reiterated to create multi-level decomposition. Slices of the 2-D frequency responses of: (b) $R_0(\boldsymbol{z})$ and $L_0(\boldsymbol{z})$, and (c) $R_1(\boldsymbol{z})$ and $L_1(\boldsymbol{z})$.

complex directional subband provides a shiftable description of image in a specific scale and direction. Note that this description is also very parsimonious because the decimation ratio for the subband is increasing with the number of directions and with the higher scale (lower resolution). An example of the frequency supports of a three-level eight-band CDFB is shown in Fig. 2.

By construction of the CDFB, each pair of corresponding directional filters has the Hilbert transform relation [10]. Therefore, the equivalent directional complex filter for each subband has one-sided frequency support, as illustrated in Fig. 2(a). The real part of the complex filter is symmetric while the imaginary part is anti-symmetric as in Fig. 2(b). The amplitude and phase information of a complex coefficient provides local information on the directional feature of the image at a specific scale and direction. In our previous work, only the amplitude information is used to calculate the feature vector [1]. Thus the objective of the next section is to understand the relation between the phase information and a typical edge so that it can be added to the feature vector.

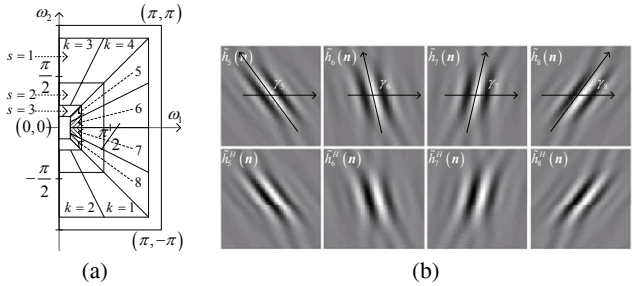


Fig. 2. The essential frequency supports of (a) the complex filters in the three-level eight-band CDFB decomposition, and (b) corresponding spatial impulse responses.

3. RELATIVE PHASES OF THE CDFB

Let an $N \times N$ image be decomposed by the pyramidal CDFB with S scales and K orientations per scale, and let $y_{sk}(i,j)$ be the subband complex coefficient at position (i,j) at scale s and orientation k, where $s = 1, 2, ..., S$ and $k = 1, 2, ..., K$. Fig. 2(a) shows an example of the CDFB decomposition when $S = 3$ and $K = 8$. The relative phase (RP) at (i,j) of one subband is defined as the difference between the phase of coefficient at that location and that of the next complex coefficient. Since the first half of subbands are 'mostly horizontal', therefore, the coefficients are re-ordered in the horizontal direction. The other half of subbands are done in the vertical direction. Specifically, the RP at the spatial location (i,j) of one subband is given as

$$p_{sk}(i,j) = \begin{cases} \angle y_{sk}(i,j) - \angle y_{sk}(i,j+1) \text{ if } 1 \le k \le \frac{K}{2}, \\ \angle y_{sk}(i,j) - \angle y_{sk}(i+1,j) \text{ if } \frac{K}{2} < k \le K, \end{cases} \quad (1)$$

where $\angle$ denotes the phase.

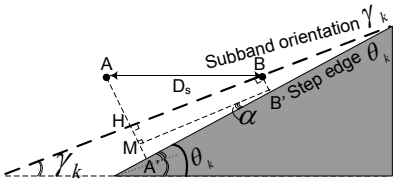


Fig. 3. Relationship between the angle θ_k of an edge and the distances from two horizontally adjacent coefficients located at A and B to the edge in the direction normal to the subband orientation k ($1 \le k \le \frac{K}{4}$) at some arbitrary scale.

We make a simplifying assumption that the angle difference between two consecutive DFB subbands is $\frac{\pi}{K}$. We denote γ_k as the center angle of the CDFB subband k. Thus subband k contains directional information at angles $\theta_k = \gamma_k + \alpha$, where $-\frac{\pi}{2K} < \alpha < \frac{\pi}{2K}$.

Let us consider an edge at angle θ_k in the supported region of subband k with $1 \le k \le \frac{K}{4}$. In this case, the center angle of subband γ_k is an acute angle. Assume that the two horizontally adjacent coefficients A and B are located in the neighborhood of an edge as shown in Fig. 3. AA' and BB' represent the distances from A and B to the edge in the direction normal to subband orientation, respectively. The distance between A and B at scale s is $D_s = 2^s$. We determine the angle of the edge θ_k by determining α in terms of AA' and BB':

$$\begin{aligned} \tan\alpha &= \frac{MA'}{MB'} = \frac{AA' - HM - AH}{MB'}, \\ &= \frac{AA' - BB' - D_s \sin\gamma_k}{D_s \cos\gamma_k} = \frac{AA' - BB'}{D_s \cos\gamma_k} - \tan\gamma_k. \end{aligned}$$

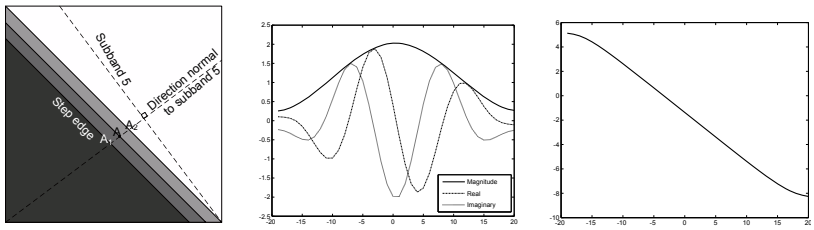
If $K \ge 8$, $|\alpha| \le \frac{\pi}{16}$ and $\alpha \approx \tan\alpha$. Hence, the feature orientation θ_k can be approximated by:

$$\theta_k \approx \gamma_k - \tan\gamma_k + \frac{AA' - BB'}{D_s \cos\gamma_k}. \quad (2)$$

Similarly, the feature orientation θ_k of the other subbands can be approximated by:

$$\theta_k \approx \begin{cases} \gamma_k - \tan\gamma_k - \frac{AA'-BB'}{D_s \cos\gamma_k} & \text{if } \frac{K}{4} < k \le \frac{K}{2}, \\ \gamma_k + \cot\gamma_k + \frac{AA'-BB'}{D_s \sin\gamma_k} & \text{if } \frac{K}{2} < k \le \frac{3K}{4}, \\ \gamma_k + \cot\gamma_k - \frac{AA'-BB'}{D_s \sin\gamma_k} & \text{if } \frac{3K}{4} < k \le K. \end{cases} \quad (3)$$

Fig. 4(b) and (c) show the magnitude and phase of a CDFB coefficient at A (Fig. 4(a)) when the edge (angle $\theta_k = 135^0$) is translated



(a) (b) (c)

Fig. 4. Translation of an edge from A_1 to A_2 for subband $s = 3$ and $k = 5$ when $\theta_k = 135°$: (a) edge translation, and coefficient at A for different positions of the edge: (b) magnitude and (c) phase.

from A_1 to A_2. The x axis represents the translation distance of the edge in horizontal direction. When A lies on the edge, $x = 0$. The distance from A to the edge is approximately $x \sin \theta_k$. We observe that when the edge moves from A_1 to A_2, the corresponding phase at A (Fig. 4(c) varies linearly with respect to the distance to the edge in the direction normal to the subband orientation (γ_k). Hence, the phase at A can be estimated by:

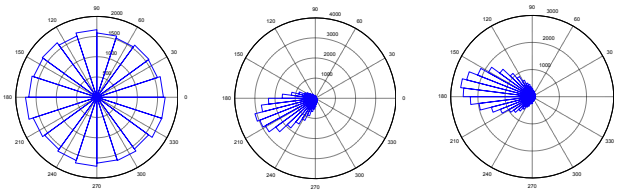
$$\angle y_{sk}(A) = a_{sk} AA' + b_{sk}. \tag{4}$$

The slope a_{sk} and the intercept b_{sk} are constants for each scale s and orientation k. Therefore the term $(AA' - BB')$ in (2) and (3) can be computed from the difference of the phases at A and B:

$$AA' - BB' = \frac{\angle y_{sk}(A) - \angle y_{sk}(B)}{a_{sk}}. \tag{5}$$

From (2), (3) and (5), we can state that the feature orientation θ_k of all CDFB subbands is linearly proportional to the RP ($\angle y_{sk}(A) - \angle y_{sk}(B)$). Because the RP can represent a dominant direction within a directional subband, we use it as a feature in the texture image retrieval application.

Fig. 5 (a) shows the uniform distribution of CDFB phases of texture 'Bark.0000' in the VisTex collection at the finest scale $s = 1$ and orientation $k = 6$. This distribution can not inform us any information of the edges, while the distribution of RPs has a particular direction as in Fig. 5(b). The circular mean m^c of RPs as defined in (6) determines the mean direction of the dominant orientations θ_k in subband k, and the circular variance σ^c in (7) determines the measure of dispersion for these dominant orientations. Note that the RP distributions of different images have the different parameters m_c and σ_c. The RP distribution of texture 'Metal.0002' is in Fig. 5(c) with $m_c = 2.99$ and $\sigma_c = 0.33$. These two parameters are used to form the RP feature vector in the next section.



(a) (b) (c)

Fig. 5. The phase histogram at scale $s = 1$ and orientation $k = 6$ for: (a) the CDFB phases of image 'Bark.0000', (b) the RP phases of image 'Bark.0000', (c) the RP phases of image 'Metal.0002'.

4. CLASSIFICATION METHOD AND EXPERIMENTS

4.1. Experiments

We select two groups of texture images for our experiments. The texture database used in the first experiment contains 116 texture images from the Brodatz album [2], which was used in [5]. The second group of textures consists of 40 images from the VisTex databases used in [9]. Each of these 512×512 images is divided into sixteen 128×128 non-overlapping sub-images, thus creating a database of 1856 texture samples in the first experiment, and 640 samples in the second one. Each original image is treated as a single class and therefore there are 16 samples from each class. To reduce the intensity correlation, all images are normalized to have zero mean and unit variance.

Each image in the database is applied to the following four decompositions: the steerable pyramid, the contourlet transform, the 2-D Gabor transform and the CDFB. The Gabor wavelet is applied with 4 scales and 6 orientations per scale, while each other transform has 3 scales of 8 orientations. Their corresponding feature vectors are computed as in [1]. The RP matrix of each CDFB subband is created as in (1). The mean of the magnitudes of the CDFB coefficients, the circular mean and the circular variance of the RPs which are computed by:

$$m^c(p) = \arctan \frac{\sum_{i,j} \sin p(i,j)}{\sum_{i,j} \cos p(i,j)}, \text{ and} \tag{6}$$

$$\sigma^c(p) = 1 - \frac{\sqrt{\left(\sum_{i,j} \sin p(i,j)\right)^2 + \left(\sum_{i,j} \cos p(i,j)\right)^2}}{L} \tag{7}$$

are used to form the CDFB-RP feature vector, where p is the PR matrix and L is the number of elements. In order to obtain a feature vector which has the same number of features as the Gabor or other feature vectors, the CDFB-RP feature vector is formed by 24 means of the magnitudes of the CDFB coefficients (24 subbands), 16 circular means of RPs (16 finest subbands) and 8 circular variances of RPs (8 finest subbands):

$$f_{yp} = [m_1(y), \ldots, m_{24}(y), m_1^c(p), \ldots, m_{16}^c(p), \sigma_1^c(p), \ldots, \sigma_8^c(p)].$$

The query pattern can be any one of the texture patterns from the image databases. Let f_{xq} and f_{yp} be two CDFB-RP feature vectors. The distance between them is defined by

$$d(f_{xq}, f_{yp}) = \sum_{k=1}^{24} \left| \frac{m_k(x) - m_k(y)}{\alpha(m_k)} \right| + \sum_{k=1}^{16} \left| \frac{m_k^c(q) - m_k^c(p)}{\alpha(m_k^c)} \right| + \sum_{k=1}^{8} \left| \frac{\sigma_k^c(q) - \sigma_k^c(p)}{\alpha(\sigma_k^c)} \right|, \tag{8}$$

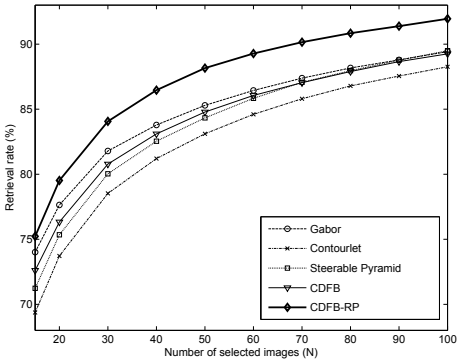


Fig. 6. Average retrieval rate of Brodatz database according to the number of top images considered

Table 1. Comparison of five different feature extraction schemes in texture image retrieval

	Gabor	**Contour**	**Steer**	**CDFB**	**CDFB-RP**
Feature Length	48	48	48	48	48
Redundant Ratio	24	4/3	32/3	8/3	8/3
Feature Time (s)	0.274	0.025	0.048	0.036	0.047
116 Brodatz					
N = 15	74.01	69.37	71.23	72.60	**75.23**
N = 40	83.78	81.21	82.54	83.10	**86.46**
N = 80	88.17	86.79	87.96	87.90	**90.84**
40 VisTex					
N = 15	80.81	75.45	74.65	79.29	**82.26**
N = 40	91.09	88.46	87.94	90.40	**92.07**
N = 65	93.89	92.32	91.47	93.32	**94.69**

where $\alpha(m_k)$, $\alpha(m_k^c)$ and $\alpha(\sigma_k^c)$ are the standard deviations of $m_k(\cdot)$, $m_k^c(\cdot)$ and $\sigma_k^c(\cdot)$ of the entire database. The distances of the Gabor feature vectors and the other vectors are described in [1].

For each query image, N nearest neighbors are selected, and the number of these textures belonging to the same class as the query texture, except for itself, is counted. This number (less than or equal to 15) divided by 15 is defined as the retrieval rate. The performance of the entire class is obtained by averaging this rate over the 16 members which belong to the same class of texture. The average of all classes is the overall performance of the transform.

4.2. Texture Retrieval Results

The second part of Table 1 summarizes overall retrieval rates using different directional transforms for the first experiment. If only the top 15 texture images that are nearest to the search texture are considered, the CDFB-RP gives the best overall retrieval performance of 75.23 %. The CDFB and the Gabor wavelet are at 72.6 % and 74.01 %, while the contourlet and the steerable pyramid are at 69.37 % and 71.23 %. Fig. 6 shows the overall performances for the case of N from 20 to 100. It is clear that the CDFB-RP is consistently better than the Gabor, the contourlet and the steerable pyramid.

In the second experiment with 40 VisTex textures, the retrieval rates for the five directional decompositions are in the third part of Table 1. We observe that with $N = 15$, the overall retrieval performance of the CDFB-RP is highest at 82.26%. The performance of the CDFB (79.29%) is very close to that of the Gabor wavelet (80.81%), while they are significantly higher than those of the contourlet and steerable pyramid (75.45% and 74.65%, respectively).

In order to evaluate the efficiency of overcomplete representation in estimating the feature vector of the texture, we compare the CPU time to calculate the feature vectors (first part of Table 1). The computation is done in MATLAB. As we can see, the Gabor features take the longest time to compute (0.274 s), but it can be reduced if we decimate the lowpass subband of the Gabor filters. The CDFB-RP features (0.047 s) take much less time which is approximately equal to the steerable pyramid features (0.048 s), but more time than the contourlet features (0.025 s) and the CDFB features (0.036 s).

5. CONCLUSION

A new image feature, which we called CDFB-RP, is proposed for feature extraction. The feature is calculated based on the CDFB decomposition, which has several attractive properties such as multiscale, multi-directional and shiftablity. By combining of the magnitude and phase information of the CDFB coefficients, the CDFB-RP feature is used in texture image retrieval. Compared to other directional transforms including the 2-D Gabor wavelet, the contourlet, the steerable pyramid and the CDFB, the CDFB-RP yields best overall performance in classification rate, while keeping the complexity relatively low. Fast decomposition structure and low redundancy make CDFB-RP more efficient in searching and browsing texture images.

6. REFERENCES

[1] An P.N. Vo, Truong T. Nguyen, and Soontorn Oraintara, "Texture image retrieval using complex directional filter banks," in *Proc.ISCAS'06*, 2006, pp. 5495–5498.

[2] Trygve Randen and John Hakon Husøy, "Filtering for texture classification: A comparative study," *IEEE Trans. on PAMI*, vol. 21, no. 4, pp. 291–310, Apr 1999.

[3] Andrew Laine and Jian Fan, "Texture classification by wavelet packet signatures," *IEEE Trans. on PAMI*, vol. 15, no. 11, pp. 1186–1191, Nov 1993.

[4] Tianhorg Chang and C.-C. Jay Kuo, "Texture analysis and classsification with tree-structured wavelet transform," *IEEE Trans. on Image Processing*, pp. 429–441, Oct 1993.

[5] B. S. Manjunath and W.Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. on PAMI*, vol. 18, no. 8, pp. 837–42, Aug 1996.

[6] Eero P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multiscale transform," *IEEE Trans. on Information Theory*, vol. 38, no. 2, pp. 587–607, Mar 1992.

[7] Z. He and M. Bystrom, "Reduced feature texture retrieval using contourlet decomposition of luminance image component," in *Proc. of 2005 ICCCAS*, May 2005.

[8] P. de Rivaz and N. Kingsbury, "Complex wavelet features for fast texture image retrieval," in *Proc. of 1999 ICIP*, 1999, vol. 1, pp. 109 – 113.

[9] M.N. Do and M. Vetterli, "Rotation invariant texture characterization and retrieval using steerable wavelet-domain hmms," *IEEE Trans. on Multimedia*, , no. 4, pp. 517–527, Dec 2002.

[10] Truong T. Nguyen and Soontorn Oraintara, "A shift-invariant multiscale multidirection image decomposition," in *Proc. ICASSP'06*, France, May 2006, pp. 153–156.