Impact of Redundancy and Gaussian Filtering on Contourlet-Based Texture Retrieval

Nadia Baaziz and Momar Diop
Département d'Informatique et d'Ingénierie
Université du Québec en Outaouais
101 rue Saint Jean Bosco, Gatineau (Québec), J8X 3X7 Canada
e-mail: nadia.baaziz@uqo.ca; diom31@uqo.ca

***Abstract*— Multiscale image representations, such as contourlets and wavelets are crucial to significant feature extraction for texture retrieval. In this paper, the aim is to highlight the positive impact of added redundancy and Gaussian filtering in multiscale image decompositions for texture retrieval. Using energy-based retrieval framework on Vistex database, conducted experiments on five multiscale transforms (contourlets, wavelets and their redundant counterparts), show the competitive enhancement provided by the redundant contourlet decomposition in terms of discriminant features and improvement of retrieval rates.**

Keywords*— *texture retrieval; contourlet; feature extraction.

I. INTRODUCTION

Retrieving image data from large scale databases lead to great challenging issues in the research field of content-based image retrieval (CBIR). A special attention is given to texture retrieval because of the omnipresence of this visual feature in most real-world images [1]. Textures are prominent in natural images (as in grasslands, brick walls, fabrics, biological tissues, etc.) and many important image properties are revealed through texture analysis such as granularity, smoothness, coarseness, periodicity, geometric structure, orientation and so on [2]. Therefore, texture retrieval is relevant to CBIR since texture characteristics are powerful in discriminating between images. Successful texture retrieval relies strongly on relevant texture feature extraction and effective similarity measurement steps yielding to high image retrieval accuracy while preserving low level of computational complexity.

There are renowned methodologies for texture feature extraction operating in the spatial domain (e.g., gray level co-occurrence matrices), in the frequency domain (e.g., Fourier spectrum measurements), or in the spatial-frequency domain (e.g., energy of wavelet coefficients) [1].

Spatial-frequency transforms, also known as multiscale representations, decompose the image spectrum into a set of localized frequency-partitions exhibiting image details at multiple scales and directions. Linear filter banks and down-sampling operators are the main tools to perform such decompositions yielding various image representations with specific properties such as multiple scales, frequency selectivity, directionality, redundancy or compactness, perfect reconstruction and shift invariance. The redundancy property is directly related to the amount of oversampling at the decomposition stage. Non-subsampled decompositions are known to be completely redundant and shift invariant. Examples of multiscale transforms include discrete wavelets, Laplacian pyramids and contourlets.

Recent studies have reported the achievement of remarkable outcomes due to the development of a variety of new texture feature extraction techniques operating on these multiscale representations [3][4][5]. This probably was motivated by two main facts: a) the human visual system adequacy to the spatial-frequency representation of image signals, and b) the inherent nature of texture patterns in terms of presence of edges, relation between primitive texture elements and variation in scales and orientations [1][2].

The major contribution of this paper is to study and reveal the benefits of incorporating redundancy and Gaussian filtering in multiscale image transforms for texture feature extraction and retrieval. The remainder of this paper is organized as follows. Section II recalls the main properties of discrete wavelet, discrete contourlet and their redundant variants that are subject to exploration in this work. A special focus is made on redundancy properties. Section III details the incorporated feature extraction methods and the texture retrieval framework. Section IV discusses experimental results and main achievements while Section V concludes the paper.

II. MULTISCALE TEXTURE REPRESENTATION

A. Discrete wavelets and stationary wavelet transforms

The discrete wavelet transform (DWT) is efficiently implemented by means of iterative linear filtering and critical down-sampling of the original image yielding three high-frequency sub-bands at each scale level in addition to one low-frequency sub-band usually known as image approximation. The DWT provides a highly compact image representation, that is, the transform is orthogonal, and incorporated down-sampling rates result into a total number of wavelet coefficients equal to the image size. Since its development, the DWT gave rise to many renowned methods and techniques in various fields of image processing and particularly in image compression. However, its use for texture analysis has revealed some limitations in

 ISBN: 978-1-61208-487-9

capturing relevant information. In fact, major drawbacks were reported in many studies; lack of shift invariance, poor frequency selectivity and poor directionality (only horizontal, vertical and diagonal orientations).

The stationary wavelet transform (SWT) has been introduced as an improved extension of the DWT. This non-subsampled variant is implemented through the so-called "algorithme à trous" in French (word trous translates into holes in English). The SWT achieves shift invariance property at the cost of substantial redundancy of wavelet coefficients. Indeed, the L-level SWT representation of a $K\times M$ image results into one approximation sub-band and $3L$ detail sub-bands. Thus, total number of wavelet coefficients is equal to $(3L+1)KM$.

B. Standard contourlet transform(SCT)

The discrete contourlet transform as introduced in [6] is designated here as the standard contourlet transform (SCT). Multi-Directionality, non-separable 2-D filtering and small amount of redundancy are among the new ingredients in this geometric transform. The decomposition is performed with high computational efficiency by combining two distinct stages. First, a multiscale decomposition uses a Laplacian pyramid (LP) scheme to transform the image into one coarse version plus a set of Laplacian sub-images. Second, a directional stage (DFB) applies iteratively 2-D filtering (DFB) and critical subsampling to further partition each LP sub-band into different and flexible numbers of frequency wedge-shaped sub-bands, thus capturing geometric structures and directional information in natural images.

When compared to the DWT, the SCT with its extra feature of directionality is almost critically sampled with a small redundancy factor up to 4/3 due to the Laplacian pyramid. SCT leads to efficient representation of smooth object boundaries with a small number of local coefficients in the right directional sub-bands.

C. Nonsubsampled contourlet transform (NSCT)

NSCT is a non-subsampled variant of the SCT [7]. All down-sampling operations are di carded from decomposition stages thus eliminating aliasing problems and allowing for full shift-invariance. However, major drawback lies in the rapid increase of computational cost as large number of directional sub-bands are generated. For example, the NSCT of a $K\times M$ image with 3 scale levels and 4 directions per level results into one approximation sub-band and 3×4 directional sub-bands. Thus, total number of wavelet coefficients is equal to $13KM$ since each sub-band size is the same as original image size.

D. Redundant contourlet transform (RCT)

The redundant contourlet transform (RCT) has been introduced in [8][9]. The RCT variant aims at reducing aliasing problems in SCT by discarding sub-sampling operations and allowing for more redundancy in the multiscale decomposition scheme (see Fig. 1).

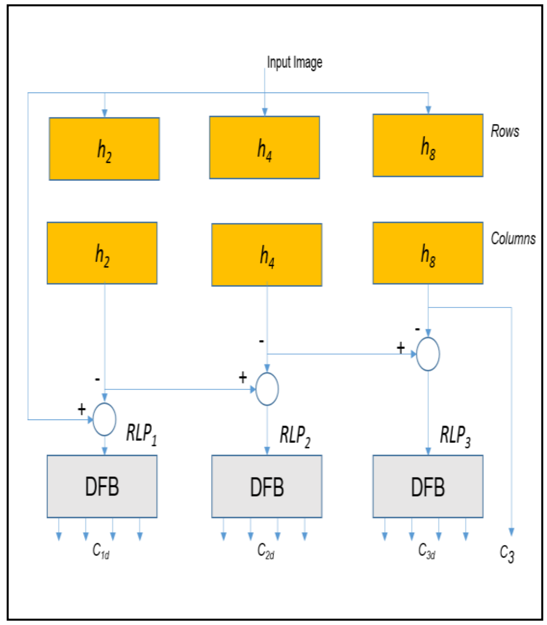

Figure 1. RCT block diagram (3 scale levels) and its frequency partition. Down-sampling is discarded from the Laplacian stage.

As for the standard contourlet transform, the RCT decomposition scheme is divided into two parts: a multiscale decomposition and a directional filter bank (DFB) using two-dimensional filtering and critical down-sampling. While DFB stage is kept unchanged, the multiscale stage is replaced by a redundant Laplacian pyramid (RLP) construct using a set of linear phase low-pass filters with pseudo-Gaussian properties. Filter impulse responses $h_b(n)$ are given in (1) where increasing values of the factor b decreases the filter passband:

$$h_b(n) = e^{-2\left(\frac{n}{b}\right)^2} - e^{-2}\left\{ e^{-2\left(\frac{n-b}{b}\right)^2} + e^{-2\left(\frac{n+b}{b}\right)^2} \right\}. \tag{1}$$

L filters (with $b=2^l$, $l=1\ldots L$) may be used to build a pyramid having $L+1$ equal-size sub-images: L detail sub-image and one coarse image approximation C_L. Then, a DFB with $D=4$ orientations and $1{:}4$ critical down-sampling is applied on each of the RLP sub-bands to obtain $4L$ equal-size directional sub-bands $\{C_{ld};\ l = 1\ldots L;\ d = 1\ldots D\}$. Thus, the redundancy factor for the RCT is $L+1$ since each RLP sub-band size is the same as original image size.

III. TEXTURE FEATURE EXTRACTION AND RETRIEVAL

Multiscale energy-based approach for texture feature extraction consists in calculating energy (L^1 norm, L^2 norm or some combination of both) and characterizing its distribution through multiscale sub-band images. The energy-based approach assumes that different texture

 ISBN: 978-1-61208-487-9

patterns have different energy distribution in the space-frequency domain. This approach is very appealing due to its low computational complexity involving mainly the calculation of first and second order moments of sub-bands coefficients [3]. Given a multiscale decomposition yielding L scale levels and D directional sub-bands C_{ld} at each level, two feature vectors E_1 and E_2 are formed as follow:

$$E(l,d) = \frac{1}{KM} \sum_{i=1}^{K} \sum_{j=1}^{M} \left| C_{ld}(i,j) \right|, \tag{2}$$

$$F(l,d) = \sqrt{\frac{1}{KM} \sum_{i=1}^{K} \sum_{j=1}^{M} \left[C_{ld}(i,j) \right]^2}\ , \tag{3}$$

$$E_1(l,d) = \left(E(l,d), \sqrt{E(l,d)^2 + F(l,d)^2} \right), \tag{4}$$

$$E_2(l,d) = \left(E(l,d), F(l,d) \right), \tag{5}$$

$$E_1 = \left\{ E_1(l,d); \quad l = 1 \ldots L;\ d = 1 \ldots D \right\}, \tag{6}$$

$$E_2 = \left\{ E_2(l,d); \quad l = 1 \ldots L;\ d = 1 \ldots D \right\}. \tag{7}$$

With $K \times M$ being the size of a given sub-band C_{ld}. Similarity measurement is computed as the Euclidean distance between two compared feature vectors.

Given a database with P images, a visual index is constructed by computing, for each image, its feature vector E_1 (or E_2) as described in (2)-(7). Retrieving similar images to the user query Q is done through the calculation of M Euclidean distances between the query feature vector and the visual index features. The N smallest distances in a ranked order are then selected as *Top-N* matches and corresponding images are retrieved.

IV. EXPERIMENTAL RESULTS

To evaluate the impact of added redundancy and pseudo-Gaussian filtering on the performance of texture retrieval, we conducted experiments using VisTex database [10]. We selected 40 grayscale images from various texture categories (as shown in Fig. 2). Each image of size 512×512 is subdivided into 16 overlapping sub-images of size 256×256, and thus, a database with 640 sub-images is constructed. To avoid any trivial discrimination based on local mean and variance, we normalized grayscale values to zero mean and unit variance. Texture retrieval performance is measured in terms of the retrieval rate (%), which is calculated as the percentage of relevant images found among the *Top-N* retrieved images (with $N=15$).

All retrieval results presented in this paper are obtained by averaging the retrieval rates corresponding to 640 queries. We compared five distinct multiscale transforms combined to two distinct texture retrieval frameworks using energy E_1 and energy E_2 features, respectively. The implemented multiscale decompositions are as follow:

1. DWT: Discrete Wavelet Transform using 4-tap *Daubechies* filters;
2. SWT: Stationary Wavelet Transform using 4-tap *Daubechies* filters;
3. SCT: Standard Contourlet Transform, with *pkva* filters, yielding 4 directional sub-bands at each scale level;
4. NSCT: Non Subsampled Contourlet Transform, with *pkva* filters, yielding 4 directional sub-bands at each scale level;
5. RCT: Redundant Contourlet Transform, with pseudo-Gaussian and *pkva* filters, yielding 4 directional sub-bands at each scale level.

Average retrieval rates (%) obtained by each of the compared methods are given in Table I. In most cases, feature extraction using energy E_2 gives slightly better results than energy E_1. Retrieval is performed with different combinations of scale levels. Each additional scale level increases the performance of the retrieval rate. Top results correspond to 3 scale levels as shown in Table I. The worst retrieval rates are obtained from wavelet-based retrieval (DWT and SWT). This is probably due to the poor directional selectivity of these transforms. We can clearly observe that retrieval rates improve as transform redundancy increases, i.e., SWT leads to better rates than DWT, and NSCT rates are better than SCT ones. However, despite the fact that RCT has a partial redundancy; it yields the best texture retrieval performance (about 3 points higher) thanks to pseudo-Gaussian filters in the Laplacian pyramid stage (see Table II).

The benefit from pseudo-Gaussian filtering is also illustrated in Table III. Indeed, texture retrieval is performed on Laplacian sub-bands, namely, NSCTLap and RCTLap that correspond to the multiscale pyramid stage of NSCT and RCT respectively. Two observations may be drawn from this experiment: 1) retrieval rates are substantially lower than in Table I due to the lack of directional selectivity in the Laplacian decomposition, and 2) performance of RCTLap is better than NSCTLap thanks to pseudo-Gaussian filters, which allow for good texture discrimination.

We believe that, despite the simplicity of the energy-based model for feature extraction, all these successful results in texture retrieval rates have manifested the potential of added redundancy and Gaussian filtering, through RCT, in extracting discriminant texture features.

 ISBN: 978-1-61208-487-9

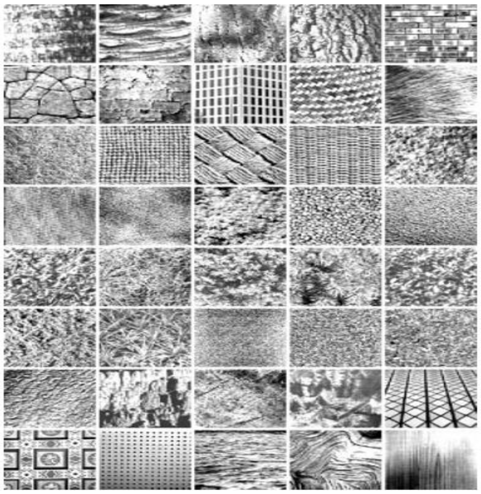

Figure 2. Texture images from the VisTex collection that are used in the experiments; from left to right and top to bottom: Bark0, Bark6, Bark8, Bark9, Brick1, Brick4, Brick5, Buildings9, Fabric0, Fabric4, Fabric7, Fabric9, Fabric11, Fabric14, Fabric15, Fabric17, Fabric18, Flowers5,Food0, Food5, Food8, Grass1, Leaves8, Leaves10, Leaves11, Leaves12, Leaves16, Metal0, Metal2, Misc2, Sand0, Stone1, Stone4, Terrain10, Tile1, Tile4, Tile7, Water5, Wood1, and Wood2.

V. CONCLUSION

The quality of texture retrieval is subject to effective image representation and relevant feature extraction that discriminate among different textures. In this paper, we successfully demonstrated the positive impact of image transform redundancy and Gaussian filtering on the efficiency of retrieval rates. Moreover, the conducted experiments using multiscale energy-based feature vectors have shown the superiority of the redundant contourlet transform (RCT) for texture discrimination and retrieval, in comparison to other multiscale transforms, namely DWT, SWT, SCT and NSCT. Subsequent ongoing research directions focus on rotation invariance of such multiscale retrieval schemes.

TABLE I. AVERAGE RETRIEVAL RATES (%) IN THE TOP-15 IMAGES VS. THE NUMBER OF DECOMPOSITION LEVELS. FIVE MULTISCALE TRANSFORMS ARE COMPARED.

Transform type	**2 levels**		**3 levels**	
	Feature E_1	***Feature E_2***	***Feature E_1***	***Feature E_2***
DWT	57.68	57.84	57.43	57.95
SWT	58.13	58.35	62.96	63.54
SCT	60.78	60.56	62.96	62.89
NSCT	61.67	61.95	61.48	62.44
RCT	**63.39**	**63.54**	**64.94**	**65.16**

TABLE II. RCT AVERAGE RETRIEVAL RATES (%) COMPARED TO OTHER TRANSFORMS. RATE DIFFERENCES ARE SHOWN.

Transform type	**2 levels**		**3 levels**	
	Feature E_1	***Feature E_2***	***Feature E_1***	***Feature E_2***
RCT -DWT	+5.71	+5.70	+7.51	+7.21
RCT-SWT	+5.26	+5.19	+1.98	+1.62
RCT-SCT	+2.61	+2.98	+1.98	+2.27
RCT-NSCT	+1.72	+1.59	+3.46	+2.72
RCT	**63.39**	**63.54**	**64.94**	**65.16**

TABLE III. AVERAGE RETRIEVAL RATES (%) IN THE TOP-15 IMAGES VS. THE NUMBER OF DECOMPOSITION LEVELS. TWO REDUNDANT LAPLACIAN TRANSFORMS ARE COMPARED.

Transform type	**2 levels**		**3 levels**	
	Feature E_1	***Feature E_2***	***Feature E_1***	***Feature E_2***
NSCTLap	41.29	42.25	44.13	45.48
RCTLap	44.05	44.11	44.73	45.32
Difference	+2.76	+1.86	+0.60	-0.14

REFERENCES

[1] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: ideas, influences, and trends of the new age," ACM Computing Surveys 40/2:1-60, 2008.

[2] M. Mirmehdi, X. Xie, and J. Suri, Handbook of Texture Analysis. Imperial College Press, London, 2008.

[3] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," IEEE Trans. on Image Processing 11/2:146-158, 2002.

[4] D. Po and M. N. Do, "Directional multiscale modeling of images using the contourlet transform," IEEE Trans. on Image Processing, 15/6:1610-1620, 2006.

[5] M. S. Allili, N. Baaziz, and M. Mejri, "Texture modeling using contourlets and finite mixtures of generalized Gaussian distributions and applications," IEEE Trans. on Multimedia, 16(3), pp. 772-784, 2014.

[6] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," IEEE Trans. on Image Processing, 14/12:2091-2106, 2005.

[7] A. L. Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: theory, design, and applications," IEEE Trans. on Image Processing, 15/10:3089-3101, 2006.

[8] N. Baaziz, "Adaptive watermarking schemes based on a redundant contourlet transform," Proc. IEEE Int. Conf. on Image Processing (ICIP), pp. I-221-4, 2005.

[9] N. Baaziz, O. Abahmane, and R. Missaoui, "Texture feature extraction in the spatial-frequency domain for content-based image retrieval," CoRR Information Retrieval and Multimedia, arXiv:1012.5208, 2010.

[10] Vision Texture Database. [Online]. Available from: http://vismod.media.mit.edu/ pub/VisTex/ 2016.05.16

 ISBN: 978-1-61208-487-9