Wavelet-based Contourlet Packet Image Coding

R. Eslami and H. Radha
Department of Electrical and Computer Engineering
Michigan State University
2120 Engineering Building
East Lansing, Michigan 48824
e-mail: {eslamira, radha} @ egr.msu.edu

Abstract –

The contourlet transform is a new directional transform, which is capable of capturing contours and fine details in images. We recently introduced the *wavelet-based contourlet transform* (WBCT) that is a non-redundant version of the contourlet transform, and appropriately used this transform for image coding. In this paper, we introduce the concept of *wavelet-based contourlet packets* (WBCP), which is similar to the notion of wavelet packets (WP). Using WBCP, we have the flexibility of choosing the most proper basis based on a criterion. In this work, we utilize WBCP for image coding to extend our previous work that was based on WBCT for image coding. Our simulation results show that the proposed WBCT packets provide both visual and PSNR improvements over WBCT. Moreover, for texture images the results outperform those of WP, visually, while achieve comparable PSNR values.

I. INTRODUCTION

Wavelets have shown their ability in representing natural images that contain smooth areas separated with edges. However, wavelets cannot efficiently take advantage of the fact that the edges usually found in natural images are smooth curves. Incorporating some tree algorithms such as SPIHT [19] into the wavelet image coding schemes, one can offset this deficiency to some extent. However, many images do not lie in the category of images that are piece-wise smooth and hence, wavelet-friendly. Texture images and natural images that possess mainly oscillatory patterns and fine details are examples of signals that wavelets are incapable of representing efficiently, and in particular for image denoising [8] and low bit-rate image compression [11] applications. Recently, several new methods and image transforms have been proposed for image denoising and coding. These methods have shown to be more powerful in representing the mentioned images than wavelets. The steerable pyramid [20], curvelet transform [3], contourlet transform [7], [6], [8] and wavelet-based contourlet transform [11] are a few examples in literature.

The contourlet transform is a directional transform that has been introduced recently. In [9], it is shown that in spite of the redundancy of the contourlet transform, by using this transform in an image coding system, one can obtain better visual results for texture and contour parts of images in comparison to a wavelet coder. This observation motivated us to investigate for a new transform based on the contourlet transform that provides two new features in addition to the former important properties such as directionality and multiresolution. These two new important features that facilitate the transform to be used for image coding are (a) non-redundancy and (b) a multiresolution structure that is similar to the 2-D wavelet decomposition. The latter feature is imperative for employing efficient scanning schemes similar to the ones used by leading wavelets-based coders. Consequently, we introduced the wavelet-based contourlet transform (WBCT) [11] that is a non-redundant version of the contourlet transform.

In the contourlet transform (Fig. 1), there are two stages: a Laplacian pyramid followed by a directional filter banks (DFB). That is, the DFB is applied to the output of the Laplacian pyramid where the number of directions is decreased as one moves through the coarser scales of the Laplacian pyramid. Laplacian pyramids provide a multiresolution system while directional filter banks grant directionality to the contourlet transform. Due to the redundancy of Laplacian pyramids, the overall transform of contourlet is redundant, and so, is not a proper choice for image coding. In the WBCT, we replaced the Laplacian pyramid stage of the contourlet transform with the wavelet transform, which is by construction a non-redundant multiresolution system, and hence, this results in the overall non-redundant WBCT transform. Based on WBCT, we developed a new coding approach: *CSPIHT* (Contourlet Set Partitioning in Hierarchical Trees), which combines the benefits of the non-redundant WBCT transform and the efficiency of a modified version of the SPIHT algorithm [10], [11]. Our previous work on CSPIHT showed the potential of this scheme for coding a variety of images such as fingerprints and textures.

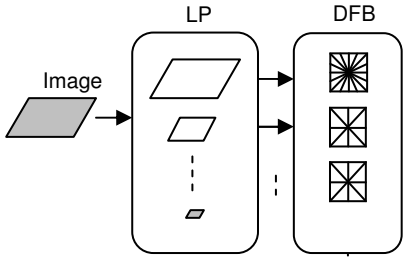

Fig. 1. A flow graph of the contourlet transform. The image is first decomposed into subbands by the Laplacian pyramid and then each detail image is analyzed by the directional filter banks.

In this work, we extend our previous effort to *Wavelet-based Contourlet Packets* (WBCP) and use this family of orthogonal basis for coding textures and fingerprint images.

The paper is organized as follows. The next section briefly explains the WBCT and also WBCP. Section III describes the coding scheme we used to utilize WBCP including a best basis selection algorithm, quantization and scanning. In Section IV, we provide some of the experimental results we obtained and finally, our conclusions are given in Section V.

II. WBCT AND WAVELET-BASED CONTOURLET PACKETS

A. Wavelet-based Contourlet Transform

Similar to the contourlet transform, the WBCT consists of two filter bank stages. The first stage provides subband decomposition, which in the case of the WBCT is a wavelet transform, in contrast to the Laplacian pyramid used in contourlets [7]. The second stage of the WBCT is a directional filter bank, which provides angular decomposition (Fig. 2).

At each level (j) in the wavelet transform, we obtain the traditional three highpass bands corresponding to the LH, HL, and HH bands. We apply DFB with the same number of directions to each band in a given level (j). Starting from the desired maximum number of directions $N_d = 2^{L_d}$ on the finest level of the wavelet transform J, we decrease the number of directions at every other dyadic scale when we proceed through the coarser levels ($j < J$). This way we could achieve the anisotropy scaling law; that is $width \approx length^2$ [3]. Fig. 3 illustrates a schematic plot of the WBCT using 3 wavelet levels and $L_d = 3$ directional levels.

Fig. 4 shows an example of the WBCT coefficients of the *Peppers* image. Here we used 3 wavelet levels and 8 directions at the finest level. You can see that most of the coefficients in the HL subbands are in the vertical directional subbands (the lower half of the subbands) while those in the LH subbands are in the horizontal directional subbands (the upper half of the subbands).

B. Wavelet-based Contourlet Packets

One of the major advantages of the WBCT is that we can construct *Wavelet-based Contourlet Packets* (WBCP) in much the same way as we have *Wavelet Packets* (WP). That is, keeping in mind the anisotropy scaling law (the number of directions is doubled at every other wavelet levels when we refine the scales), we allow quad-tree decomposition of both

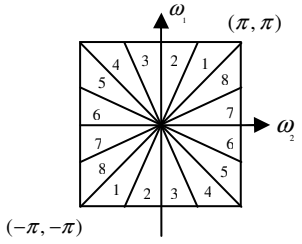

Fig. 2. Directional filter bank frequency partitioning.

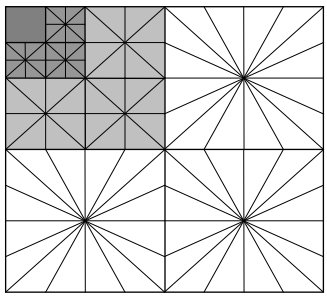

Fig. 3. A schematic plot of the WBCT using 3 dyadic wavelet levels and 8-4-4 directions (8 directions at the finest level ($N_d = 8$)). The directional decomposition is overlaid the wavelet subbands.

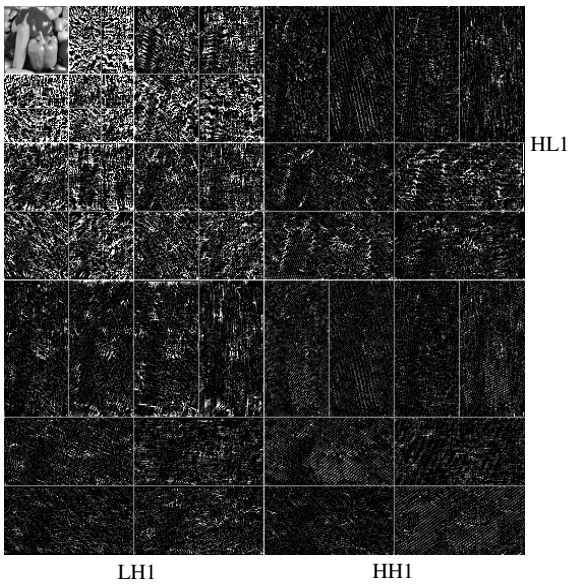

Fig. 4. The WBCT coefficients of the *Peppers* image. Here, 3 wavelet scales and 8-4-4 directions are used. For better visualization, the transform coefficients are clipped between 0 and 7.

lowpass and highpass channels in wavelets and then apply the DFB to each subband. Fig. 5 schematically illustrates an example of the wavelet-based contourlet packets. Below we present a multiresolution analysis for WBCP.

For an input image of size $N \times N$, where $N^{-1} = 2^L$, we associate the approximation space $\mathbf{V}_L^2 = \mathbf{V}_L \otimes \mathbf{V}_L$ to the original image or the root of wavelet packet tree [12]. That is, we use separable subspaces, and construct image wavelet packets using separable products of one-dimensional wavelet packets. For a two-dimensional decomposition of wavelet packets, we label each node of the quad-tree with a scale j, and two values $0 \le p < 2^{j-L}$ and $0 \le q < 2^{j-L}$ (where $j - L > 0$) corresponding to the frequencies of one-dimensional subspaces

$$\mathbf{W}_j^{p,q} = \mathbf{W}_j^p \otimes \mathbf{W}_j^q, \qquad (1)$$

which satisfy $\mathbf{W}_j^{p,q} \subset \mathbf{V}_L^2 = \mathbf{W}_j^{0,0}$. For the one-dimensional spaces we have

$$\mathbf{W}_j^p = \mathbf{W}_{j+1}^{2p} \oplus \mathbf{W}_{j+1}^{2p+1} \text{ and } \mathbf{W}_j^q = \mathbf{W}_{j+1}^{2q} \oplus \mathbf{W}_{j+1}^{2q+1}. \qquad (2)$$

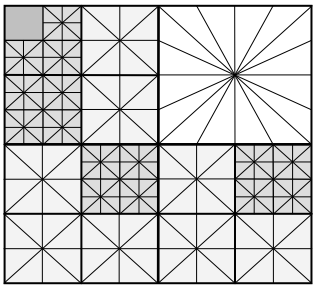

Fig. 5. Schematic diagram of an example of the wavelet-based contourlet packets.

By inserting (2) in (1) we obtain

$$\mathbf{W}_j^{p,q} = \mathbf{W}_{j+1}^{2p,2q} \oplus \mathbf{W}_{j+1}^{2p+1,2q} \oplus \mathbf{W}_{j+1}^{2p,2q+1} \oplus \mathbf{W}_{j+1}^{2p+1,2q+1}, \quad (3)$$

which corresponds to the four children subspaces of the quad-tree node. Now if $\{j_i, p_i, q_i\}_{1\le i\le I}$ are the indices of the terminal nodes (or leaves) of the wavelet packet tree, we have the following reconstruction relation:

$$\mathbf{W}_L^{0,0} = \oplus_{i=1}^{I} \mathbf{W}_{j_i}^{p_i,q_i}. \quad (4)$$

Let $\{\psi_j^p(x_1 - 2^j n_1)\}_{n_1\in\mathbb{Z}}$, $\{\psi_j^q(x_2 - 2^j n_2)\}_{n_2\in\mathbb{Z}}$, and $\{\psi_j^{p,q}(x - 2^j n)\}_{n\in\mathbb{Z}^2}$, where $x = (x_1, x_2)$, be the elements of corresponding spaces $\mathbf{W}_j^p$, $\mathbf{W}_j^q$, and $\mathbf{W}_j^{p,q}$. The reconstruction relation (4) implies that the corresponding wavelet packet bases

$$\{\psi_{j_i}^{p_i,q_i}(x - 2^{j_i} n)\}_{(n_1,n_2)\in\mathbb{Z}^2, 1\le i\le I}, \quad (5)$$

construct an orthonormal basis of $\mathbf{V}_L^2 = \mathbf{W}_j^{0,0}$.

Now we provide formulation for the DFB stage in WBCP and relate it to wavelet packets. For an l-level DFB, we have 2^{l_d} directional subbands with $G_k^{(l_d)}$, $(0 \le k < 2^{l_d})$ equivalent synthesis filters and the overall downsampling matrices of $S_k^{(l_d)}$, $(0 \le k < 2^{l_d})$ defined as [7]:

$$S_k^{(l_d)} = \begin{cases} \begin{bmatrix} 2^{l_d-1} & 0 \\ 0 & 2 \end{bmatrix}, & \text{if } 0 \le k < 2^{l_d-1} \\ \begin{bmatrix} 2 & 0 \\ 0 & 2^{l_d-1} \end{bmatrix}, & \text{if } 2^{l_d-1} \le k < 2^{l_d} \end{cases}. \quad (6)$$

Then, $\{g_k^{(l_d)}[n - S_k^{(l_d)} m]\}$, $(0 \le k < 2^{l_d}, m \in \mathbb{Z}^2)$ is a directional basis for $l^2(\mathbb{Z}^2)$; where $g_k^{(l_d)}$ is the impulse response of the synthesis filter $G_k^{(l_d)}$, n is the spatial index, and m is a shift in position. Now, if we apply l_d directional levels to the space $\mathbf{W}_{j_i}^{p_i,q_i}$ at a terminal node, where $p_i + q_i \ne 0$ (we exclude the approximation image), we obtain 2^{l_d} directional subbands for $\mathbf{W}_{j_i}^{p_i,q_i}$ where

$$\mathbf{W}_{j_i}^{p_i,q_i} = \oplus_{k=0}^{2^{l_d}-1} \mathbf{W}_{j_i,k}^{p_i,q_i,(l_d)}. \quad (7)$$

The directional level l_d as mentioned before, is dependent on the scale j of wavelet packets, that is, for the terminal nodes at depth $j - L = 1$ we apply a DFB with the maximum number of desired directions $N_d = 2^{L_d}$ to the corresponding wavelet packets at those nodes and then as we increase the depth, we decrease L_d at every other scale (note that we do not apply DFB to the leaf node $\mathbf{W}_{j_0}^{0,0}$ during this procedure). Let us define the elements of wavelet-best contourlet packets as

$$\eta_{j_i,k,n}^{p_i,q_i,(l_d)} = \sum_{m\in\mathbb{Z}^2} g_k^{(l_d)}[m - S_k^{(l_d)} n] \psi_{j_i,m}^{p_i,q_i}. \quad (8)$$

Using a similar proof to the one provided in [7], we can show that the family $\{\eta_{j_i,k,n}^{p_i,q_i,(l_d)}\}_{n\in\mathbb{Z}^2, 1\le i\le I, 0\le k<l_d}$ form an orthonormal basis for $\mathbf{V}_L^2 = \mathbf{W}_j^{0,0}$.

III. Wavelet-based Contourlet Packet Coding

Similar to wavelet packets (WP), wavelet-based contourlet packets (WBCP) provide a large library of bases. The number of WBCP (or WP) bases in a full quad-tree decomposition of depth J is at least $2^{4^{J-1}}$ [12]. Therefore, using a proper algorithm for choosing a "best" basis from the wavelet packets library, one can adaptively provide a best representation for an image and improve the non-linear approximation of wavelets. Using the same philosophy, we design a WBCP image coding scheme to improve our previous results in which we used WBCT in conjunction with the CSPIHT algorithm for image coding. Below we describe the proposed coding system.

A. Best WBCP Basis Selection

A best wavelet packet or wavelet-based contourlet packet is a basis that is best adapted to a particular signal. Using a brute force search in the library of $2^{4^{J-1}}$ bases would require more than $N^2 2^{4^{J-1}}$ operation (for an image of size $N \times N$), which is computationally prohibitive. In [5], a bottom-up pruning scheme of the fully decomposed wavelet packets tree is presented that is based on dynamic programming. The authors suggested the Shannon entropy measure as the cost function; however, the cost function can be any additive function of the form $M(x) = \sum_{k\in\mathbb{Z}^2} \mu(|x(k)|)$; $\mu(0) = 0$. The Shannon entropy measure is different from the entropy of a random variable that is not an additive function, and the authors in [15] observed that it is incapable of choosing a meaningful basis from the library of WP. Another method is to optimizing the rate-distortion function using Lagrange multipliers [18]. However, this approach is computationally exhaustive. Other approaches including matching pursuit [14] and basis pursuit [4] have also been proposed. If we consider a redundant dictionary of bases $D = \{g_p\}_{0\le p<P}$ with $P > N$ vectors, they provide a solution for choosing N linearly independent vectors out of the dictionary D. The solution, however, is a suboptimal solution. Although these approaches supply better adapted bases that can represent signals not well adapted with the WP tree structure, they are computationally prohibitive for large signals [12].

For a best basis selection, we used the dynamic programming approach proposed in [5]. Furthermore, we used

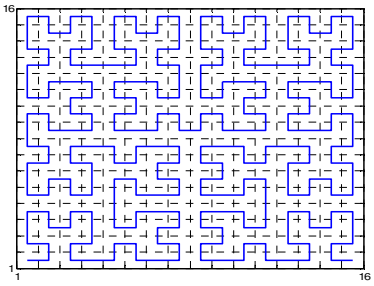

Fig. 6. A Hilbert filling curve of order 4.

a direct transform coding approach as proposed in [13] that is based on non-linear approximation. In this approach we threshold the transform coefficients and then quantize them. However, we have to code also the positions of these *significant* coefficients, in which we use a binary bitstream as the *significance map*. For this coding approach, a proper cost function would be a one that estimates the actual entropy of the generated bitstreams, which are the quantized coefficients and significance map. Therefore, we used the cost function proposed in [15], which is the addition of two costs:

- $\text{cost1}(x) = \sum_{i,(Q(x_i)\neq 0)} \max(0, \log_2(|Q(x_i)|))$,
- $\text{cost2}(x) = N^2(-p\log_2(p) - (1-p)\log_2(1-p))$,

where *cost1* is a first order approximation of the entropy and *cost2* is formulated assuming a Bernoulli random variable for the significance map. Hence, considering a global threshold, we use the above cost function to find the best basis of WBCP for a given image.

B. Quantization and Scanning

Similar to the contourlet transform [17], we found that the WBCT coefficients are non-Gaussian, and hence a uniform quantization is not optimal. For a Laplacian distribution, it is shown [21] that if one takes the zero-bin twice as large as the other bins, then a near-optimal quantization can be achieved. Assuming a Laplacian distribution for the WBCT coefficients, we use the mentioned quantizer for our approach.

Using a proper method for scanning the transform coefficients, one can take advantage of the dependencies among them to improve the resulting rate. One of the scanning method is employing a Hilbert filling curve. If one uses this mapping curve for scanning the wavelet coefficients, an improvement in the output rate is gained when compared with the case of row-wise scanning [1]. Fig. 6 illustrates a Hilbert filling curve of order 4. Another approach is to scan the coefficients taking into account the orientation of each subband. The scanning of the wavelet scales is performed in order of increasing frequency. In WBCP each wavelet subband (except the approximation image) is decomposed into several directional subbands; therefore, at each wavelet subband, we scan the directional subbands individually. We employed the following three scanning approaches: row-wise (*or* column-wise) scanning, scanning based on the orientation of directional subbands in which we scan horizontal (vertical) directional subbands, horizontally (vertically), and scanning using Hilbert filling curves. We observed that Hilbert filling curves yield the best results.

After scanning, we run-length-code the significance map and entropy-code the resulting bitstreams. Of course, we are required to code the shape of the selected best WBCP tree, as well. To do so, we assign 1 to the terminal nodes of the best tree and zero to the other existing nodes, then scan these values. A run-length coding could also be used for the resulting bitstream. The length of this bitstream is maximally 4^J for a packet tree of depth J. For small J, the overhead of coding the shape of tree is not significant.

IV. Experimental Results

We tested the proposed WBCP coder scheme as well as the WBCT/CSPIHT coder on a variety of images with fine details such as textures and fingerprints, each having a size of 512x512. For the sake of comparison, we also provide the results of the wavelets-based SPIHT coder and WP coder. The WBCT and WBCP use 5 wavelet levels and 16 directions at the finest scale. We used non-separable fan filters of support sizes (23, 23) and (45, 45) as described in [16]. We also used depth-five WP and utilized biorthogonal Daubechies 9-7 wavelets in all the schemes. Fig. 7 depicts the rate-distortion curves obtained for the texture images *D16* and *D80* acquired from the Brodatz textures collection [2] and *Fingerprint* image. The rates of the coders are computed using the entropy of the output bitstreams. However, one can approach the entropy using the arithmetic entropy coder. As Fig. 7 demonstrates, we could clearly improve our previous PSNR performance resulting from the WBCT/CSPIHT coder. When compared with the WP coder, for the texture images mentioned above, the proposed WBCP coder provides competitive PSNR results. To compare the above approaches visually, we present some of the coded images at very low bit-rate. Fig. 8 shows the coded results of the image *D80* at $r = 0.07$ bpp. As seen, more details are preserved in the coded image of the WBCP coder when compared with the other methods. Fig. 9 depicts another example where the coded results of the *Fingerprint* image at $r = 0.05$ bpp are shown. Again, the best result is obtained through the proposed WBCP coding approach. Although it is clear that the WBCT/CSPIHT coder preserved more ridges when compared with the popular wavelets SPIHT coder, there are still more ridges retained in the coded image using the proposed WBCP coder. Because of the smooth regions around the *Fingerprint* image, the PSNR values are degraded in the CSPIHT and WBCP coding results for this image. As we reported in [11], due to the downsampling of highpass channel of the wavelet transform in the WBCT, this scheme introduces pseudo-Gibbs phenomena artifacts in the smooth regions. However, the WBCT and WBCP schemes provide significant improvement in the visual results of texture images and other images that contain mostly contours and oscillatory patterns.

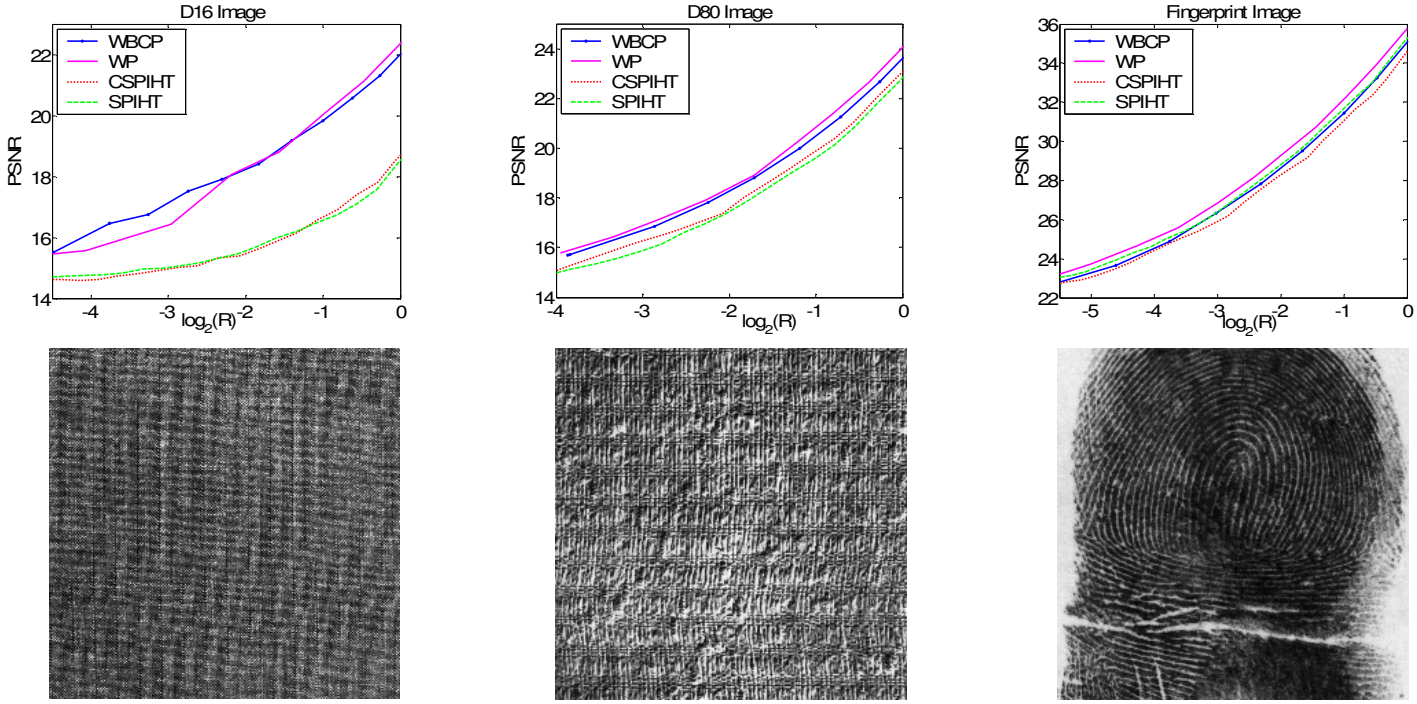

Fig. 7. The PSNR-rate curves for the images *D16*, *D80 [2]*, and *Fingerprint* using the coding schemes of WBCP coder, WP coder, WBCT/CSPIHT coder, and SPIHT coder. The original images are shown at the bottom.

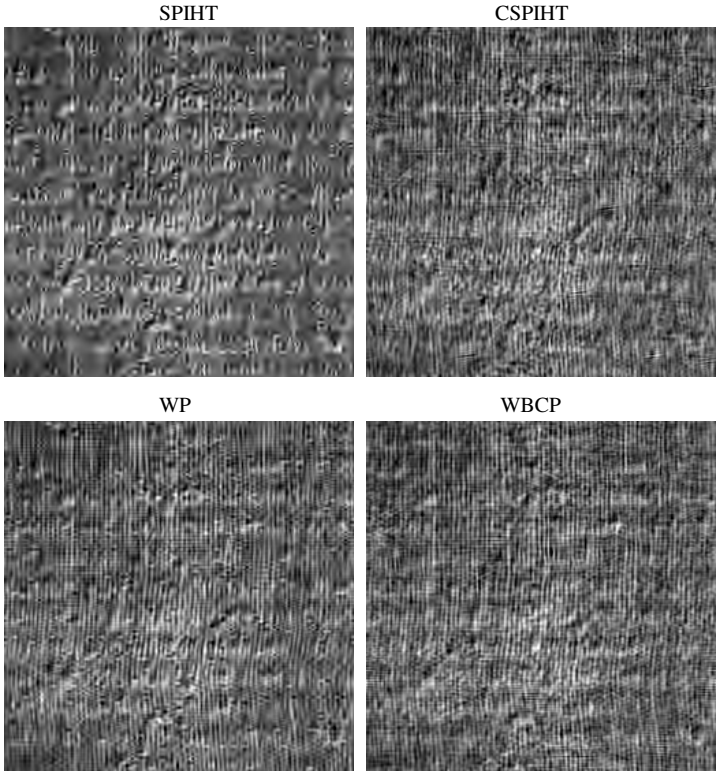

Fig. 8. The *D80* image coded at rate $r = 0.07$ bpp.

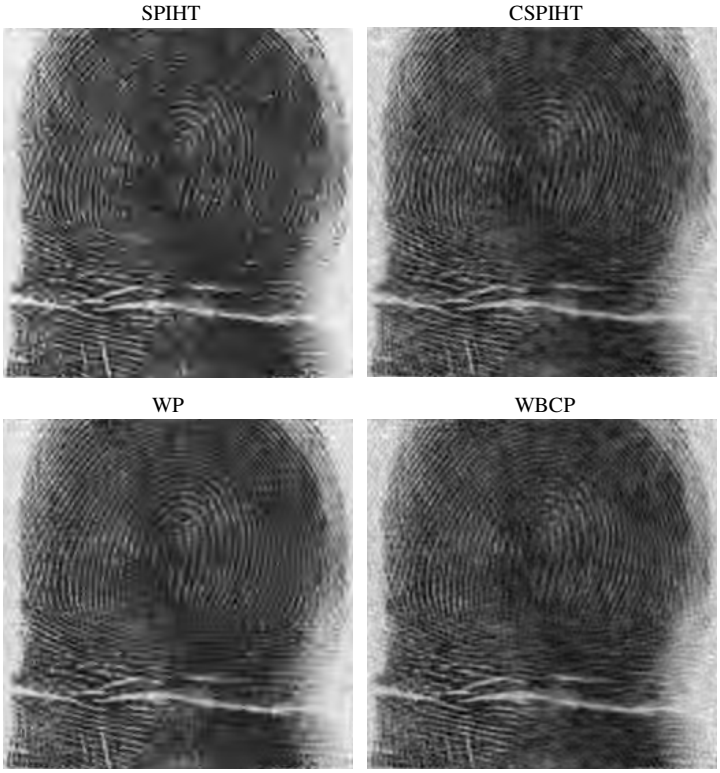

Fig. 9. The *Fingerprint* image coded at rate $r = 0.05$ bpp.

V. Conclusions

We extended our recent work (WBCT/CSPIHT coding) and proposed a coding scheme based on WBCP. With WBCP, we could gain an overall (significant) improvement in PSNR values. Our experiments also indicated an enhancement in the visual results when compared with the WBCT/CSPIHT coder as well as the WP coder.

REFERENCES

[1] W. Brghorn, T. Boskamp, M Lang, and H.-O. Peitgen, "Fast variable run-length coding for embedded progressive wavelet-based image compression," *IEEE Trans. Image Processing*, vol. 10, pp. 1781-1790, Dec. 2001.

[2] Brodatz Textures Collection, *available on* http://www.ux.his.no/~tranden/brodatz.html

[3] E. J. Candes and D. L. Donoho, "Curvelets – a surprisingly effective nonadaptive representation for objects with edges," *in Curve and Surface Fitting*, Saint- Malo, 1999, Vanderbuilt Univ. Press.

[4] S. Chen and D. Donoho, "Atomic decomposition by basis pursuit," *in SPIE Intl. Conf. on Wavelets,* San Diego, Jul. 1995.

[5] R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Information Theory*, vol. 38, pp. 713-718, 1992.

[6] M. N. Do, *Directional multiresolution image representations.* Ph.D. thesis, EPFL, Lausanne, Switzerland, Dec. 2001.

[7] M. N. Do and M. Vetterli, "Contourlets," *in Beyond Wavelets*, Academic Press, New York, 2003.

[8] R. Eslami and H. Radha, "Image denoising using translation-invariant contourlet transform," *to appear in proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Mar. 2005.

[9] R. Eslami and H. Radha, "On low bit-rate coding using the contourlet transform," *in proc. of Asilomar Conference on Signals, Systems, and Computers,* pp. 1524-1528, Pacific Grove, USA, 2003.

[10] R. Eslami and H. Radha, "Wavelet-based contourlet coding using an SPIHT-like algorithm," *in proc. of Conference on Information Sciences and Systems*, pp. 784-788, Princeton, March 2004.

[11] R. Eslami and H. Radha, "Wavelet-based contourlet transform and its application to image coding," *to appear in proc. of IEEE International Conference on Image Processing*, Singapore, Oct. 2004.

[12] S. Mallat, *A wavelet tour of signal processing.* Academic Press, 2nd Ed., 1998.

[13] S. Mallat and F. Falzon, "Analysis of low bit-rate image transform coding," *IEEE Trans. Signal Processing,* vol. 46, no. 4, pp. 1027-1042, Apr. 1998.

[14] S. Mallat and Z. Zhang, "Matching pursuit with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, pp. 3397-3415, Dec. 1993.

[15] F. G. Mayer, A. Z. Averbuch, and J.-O. Stromberg, "Fast adaptive wavelet packet image compression," *IEEE Trans. Image Processing,* vol. 9, no. 5, pp. 792-800, May 2000.

[16] S. M. Phoong, C. W. Kim, P. P. Vaidyanathan, and R. Ansari, "A new class of two-channel biorthogonal filter banks and wavelet bases," *IEEE Trans. Signal Processing,* vol. 43, no. 3, pp. 649-665, Mar. 1995.

[17] D. D.-Y. Po and M. N. Do, "Directional multiscale modeling of images using the contourlet transform," submitted to *IEEE Trans. Image Processing*, 2003.

[18] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense," *IEEE Trans. Image Processing*, vol. 2, pp. 160-175, Apr. 1993.

[19] A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits and Systems for Video Technology,* vol. 6, pp. 243-250, June 1996.

[20] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation" *in proc. Intl. Conf. on Image Processing,* Washington DC, 1995.

[21] G. J. Sullivan, "Efficient scalar quantization of exponential and Laplacian random variables," *IEEE Trans. Information Theory*, vol. 42, pp. 1365-1374, Sep. 1996.