Contourlet Image Coding Based on Adjusted SPIHT

Haohao Song, Songyu Yu, Li Song, and Hongkai Xiong

Institute of Image Communication and Information Processing,
Shanghai Jiao Tong University, SHANGHAI 200030, P.R. China
{songhaohao, syyu, xionghongkai}@sjtu.edu.cn
songli@qantsoft.com

Abstract. Contourlet is a new image representation method, which can efficiently represent contours and textures in images. In this paper, we analyze the distribution of significant contourlet coefficients in different subbands and propose a contourlet image coding algorithm by constructing a virtual low frequency subband and adjusting coding method of SPIHT (Set Partitioning in Hierarchical Trees) algorithm according to the structure of contourlet coefficients. The proposed coding algorithm can provide an embedded bit stream, which is very desirable in heterogeneous networks. Our experiments demonstrate that the proposed coding algorithm can achieve better or competitive compression performance when compared with traditional wavelet transform with SPIHT and wavelet-based contourlet transform with SPIHT, which both are embedded image coding algorithms based on two non-redundant transforms. At the same time, benefiting from genuine contourlet adopted in the proposed coding algorithm, more contours and textures in the coded images are preserved to ensure superior subjective quality.

1 Introduction

Efficient representation of visual information lies at the heart of image compression. For 1-D piecewise smooth signals, wavelets have been served as a right representation tool. In addition, fast transform and convenient tree structures (e.g. zerotree and spatial-orientation tree) provide the key factors for the success of wavelets in image compression. EZW [1], SPIHT [2], and EBCOT used in JPEG2000 [3] are famous embedded wavelet image coding algorithms.

However, two-dimensional signals are smooth away from discontinuities across smooth curves. Such signals resemble natural images where discontinuities are generated by edges – referred to points in the image with sharp contrast in the intensity, whereas edges are often gathered along smooth contours that are created by typically smooth boundaries of physical objects. The commonly used separable wavelets in 2-D obtained by a tensor-product of 1-D wavelets are only good at capturing the discontinuities at edge points, but do not see the smoothness along contours. Thus, more powerful schemes are needed in higher dimensions [4] [5].

In recent years, many efficient representations are proposed for two-dimensional signals such as images. Ridgelet [6] by Candès and Donoho achieves high efficient representation of linear singularities in images. Ridgelet transform is done by viewing ridgelet analysis as a form of wavelet analysis in the Radon domain. Many attempts have been made to use ridgelet for image compression [7] [8].

Y.-S. Ho and H.J. Kim (Eds.): PCM 2005, Part I, LNCS 3767, pp. 629 – 640, 2005.

However, in images, edges are typically curved rather than straight. Ridgelet cannot represent curves efficiently. Curvelet [9] by Candès and Donoho can be suited for objects which are smooth away from discontinuities across smooth curves. Later, the authors proposed the second-generation curvelet [10].

Additionally, bandelet [11] by Pennec and Mallat, wedgelet [12] by Donoho and beamlet [13] by Donoho and Huo all can represent contours of objects in images efficiently.

Contourlet [4] [5] is presented by Do and Vetterli as a new image representation method. Although Contourlet transform (CT) is a overcomplete transform with a redundancy factor 4/3 that would increase the rate for a given distortion, it can efficiently represent image containing contours and textures [14]. Recently, some approaches have been attempted to use CT for image compression.

A low bit-rate image coding using CT was proposed in [15]. It uses a scalar quantizer with a zero bin that is twice as large as the other bins to code contourlet coefficients. Recently, a coding technique based on a mixed CT and wavelet transform (WT) was presented [16]. The transform is optimized through an iterative projection process in the transform domain in order to minimize the quantization error in the image domain. Both algorithms cannot produce embedded bit stream.

In [17], a wavelet-based contourlet transform (WBCT) is proposed, which is a non-redundant transform. Authors used WBCT in conjunction with SPIHT to propose an embedded image coding algorithm. Experimental results show that the coding algorithm is competitive to WT with SPIHT, especially for a category of images that have a significant amount of textures and oscillatory patterns and therefore are not "wavelet-friendly" images. But due to the frequency scrambling resulting from the downsampling of the high frequency (HF) subbands in wavelet, WBCT non-linear approximation results might result in larger distortion than contourlet results in some image regions such as contours and textures. It means that CT might be more efficient in representing contours and textures than WBCT. At the same time, the computing complexity of the coding algorithm in [17] will increase fast due to the repositioning algorithm when moving forward along the scales.

Because wavelet incurs frequency scrambling and genuine CT is more efficient in representing contours and textures, we propose a new contourlet image coding algorithm based on adjusted SPIHT. It combines genuine CT with SPIHT efficiently by (1) adopting adaptive subband coding order scheme, (2) constructing a virtual low frequency (LF) subband and (3) adjusting coding structure of SPIHT properly. The proposed coding algorithm can produce embedded bit stream. Experimental results show that compared with traditional WT with SPIHT and WBCT with SPIHT, the proposed coding algorithm has better or comparative objective quality of reconstructed images with lower computational complexity. Meanwhile, more contours and textures are preserved in the proposed coding algorithm and the visual effect is better than those mentioned above.

The rest of the paper is organized as follows. Section 2 will present contourlet transform for image. In Section 3, SPIHT algorithm is introduced briefly. Contourlet image coding algorithm based on adjusted SPIHT is addressed in Section 4. Section 5 contains experimental results. Section 6 concludes the paper and discusses possible future work.

2 Contourlet Transform

Contourlet was proposed by Do and Vetterli at 2001. It is a double filter bank structure for obtaining sparse expansions for typical images with smooth contours. In this double filter bank, the Laplacian pyramid (LP) is first used to capture the point discontinuities, then followed by a directional filter bank (DFB) to link point discontinuities into linear structures. The overall result is an image expansion using basic elements like contour segments, and thus is named contourlet [5].

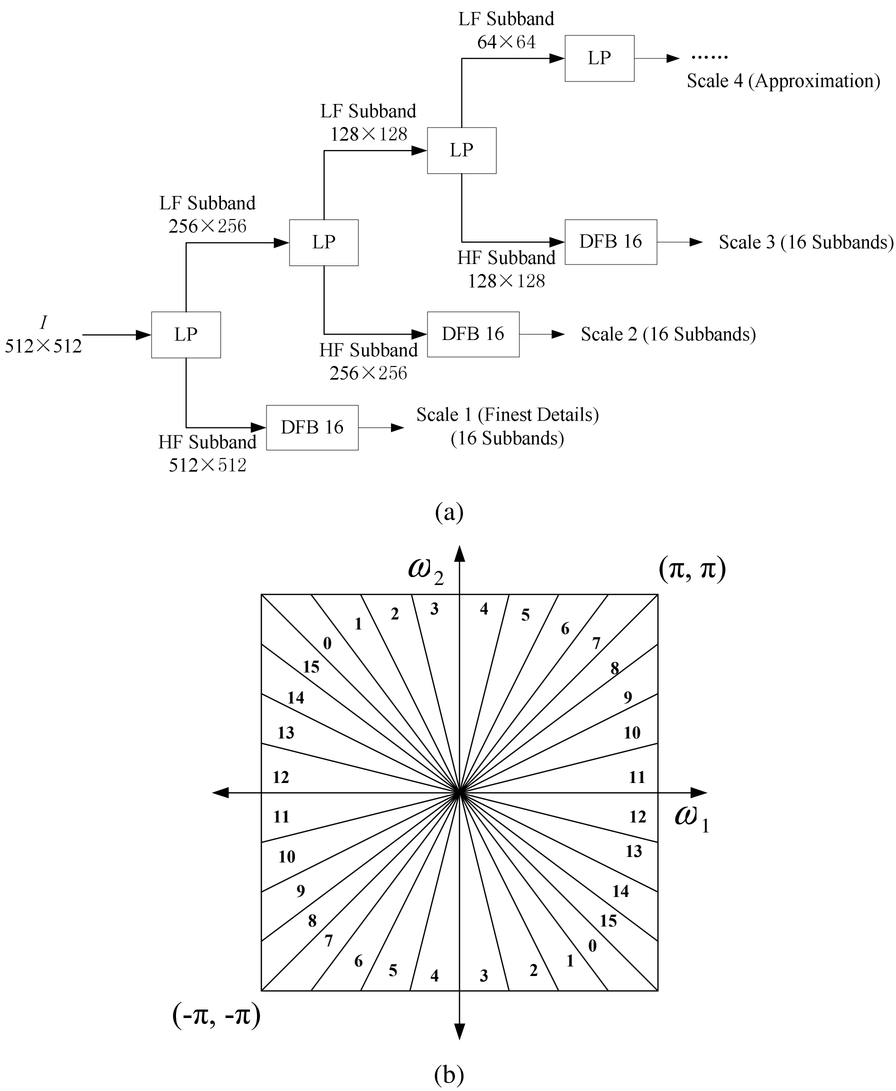


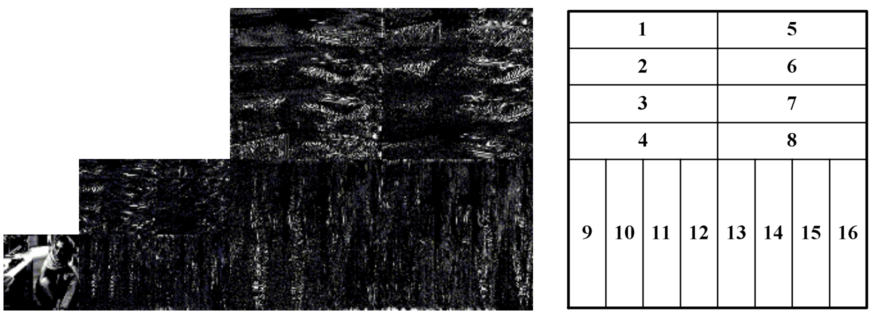
Fig. 1. CT for a 512 ×512 image. (a) Transform flowchart. (b) Frequency partitioning with 16 real wedge-shaped frequency bands.

Fig. 1 shows CT for a 512 ×512 image. HF subband images from the LP are fed into a DFB so that directional information can be captured. The scheme can be iterated on LF subband image. CT decomposes images into directional subbands at multiple scales.

CT has several distinguishing properties as follows: (1) seamless translation to the discrete world; (2) 2-D frequency partition on centric squares; (3) fast filter bank algorithms and convenient tree structures; (4) compactly supported contourlet frames; (5) flexible refinements for the spatial resolution and the angular resolution.

In contrast to the critically sampled wavelet scheme, the LP has the distinguishing feature that each pyramid level generates only one bandpass image (even for multidimensional cases), and this image does not have "scrambled" frequencies. This frequency scrambling happens in the wavelet filter bank when a highpass channel, after downsampling, is folded back into the low frequency band, and thus its spectrum is reflected. In the LP, this effect is avoided by downsampling the lowpass channel only [5].

Fig. 2 shows the example of CT on "Barbara" image. We noticed that only contourlets which match both location and direction of image contours produce significant coefficients.



(a) (b)

Fig. 2. Example of CT on "Barbara" image. (a) For clear visualization, image is only decomposed into two pyramidal levels, which are then decomposed into 16 directional subbands respectively. Small coefficients are shown in black while large coefficients are shown in white. (b) Corresponding sequence numbers of 16 subbands.

3 SPIHT

Because different types of terminals and nodes in heterogeneous networks have different processing, memory and display capabilities, resolution, temporal, and SNR scalability of the coded video bit stream are needed essentially. It is achievable by realizing progressive transmission and embeddedness of information in decreasing order of its information content [1], which is an important characteristic of set partitioning and significance evaluation on hierarchical structures of transformed images [2]. This recognition has spawned more algorithms in image processing which include EZW, SPIHT, and EBCOT. It is noticeable that these three algorithms are all constructed on the basis of WT.

SPIHT was proposed by Said and Pearlman at 1996. It utilizes three basic concepts: (1) searching for sets in spatial-orientation trees in a WT; (2) partitioning the WT coefficients in these trees into sets defined by the level of the highest significant bit in a bit-plane representation of their magnitudes; and (3) coding and transmitting bits associated with the highest remaining bit planes first.

The following sets of coordinates are used in SPIHT:

$O(i, j)$: set of coordinates of all offspring of node (i, j).
$D(i, j)$: set of coordinates of all descendants of node (i, j).
$L(i, j) = D(i, j) - O(i, j)$.

In the practical implementation, SPIHT defines three ordered lists to store the significance information: list of insignificant sets (LIS), list of insignificant pixels (LIP), and list of significant pixels (LSP). In all lists each entry is identified by a coordinate (i, j), which in LIP and LSP represents individual pixels, and in LIS represents either set $D(i, j)$ or $L(i, j)$. To differentiate between them, the algorithm defined that a LIS entry is of type A if it represents $D(i, j)$, and of type B if it represents $L(i, j)$.

SPIHT is extremely fast in execution. It has better compression performance than EZW and comparable compression performance with EBCOT but with lower computing complexity. It is the reason why SPIHT is adopted in this paper.

4 Contourlet Image Coding Based on Adjusted SPIHT

Advanced modern image coding algorithm should satisfy three main properties as follow: (1) producing embedded bit stream, the reception of code bits can be truncated at any point and the image can still be decompressed and reconstructed to adapt different types of terminals and .5(t)01 Tm(0 9.96 42 Tc32048 O.6(m)28.2(n).7(s)8t ar on the analysis of distribution of significant contourlet coefficients in different subbands, we propose an adaptive subband coding order scheme that can satisfy the demand of embedded coding.

Table 1 shows distribution of significant contourlet coefficients of 4 standard test images (8 bit/pixel (bpp), 512×512) at different thresholds. 4-scale LP and 16 directional decomposition at each scale is performed on each HF image.

Table 1. Distribution of significant contourlet coefficients of 4 standard test images at different thresholds. There, threshold 1=128, threshold 2=64, threshold 3=32. 16 subbands are sorted by the number of significant coefficients above corresponding threshold in subbands in descending order. 01-16 are the sequence numbers of 16 subbands as illustrated in Fig. 2 (b).

Image	Threshold 1	Threshold 2	Threshold 3
Barbara	06/04/05/02/13/01/03/ 12/09/10/07/16/11/14/ 15/08	06/04/05/02//03/01/13/ 07/16/09/10/12/11/15/ 14/08	06/04/05/03/02/07/01/ 16/13/09/10/12/15/11/ 14/08
Goldhill	13/04/05/12/11/14/16/ 03/06/15/02/10/09/07/ 01/08	13/04/12/05/11/14/16/ 03/10/15/06/02/09/07/ 01/08	13/04/12/05/11/14/16/ 15/03/10/06/02/01/07/ 09/08
Mandrill	13/12/14/16/04/11/03/ 06/15/09/05/02/10/01/ 07/08	13/12/14/11/16/15/04/ 10/03/06/09/02/05/01/ 07/08	13/12/14/16/11/15/04/ 10/03/06/09/02/05/01/ 07/08
Fingerprint	13/12/14/16/11/15/10/ 09/04/07/01/05/02/03/ 06/08	13/12/14/11/15/16/04/ 10/09/01/07/05/02/03/ 06/08	13/12/14/11/15/16/04/ 10/09/05/03/06/02/07/ 01/08

By analyzing Table 1, we found that although distribution of significant contourlet coefficients varies with images, the orderliness of distribution at different thresholds is kept for every image. In general, if a subband has more significant coefficients at higher threshold, it will have more significant coefficients too at lower threshold in all probability. Based on the analysis above, we propose an adaptive subband coding order scheme that can scan and code significant coefficients as much as possible when bit stream is truncated at any point in order to satisfy the demand of embedded coding.

Our proposed coding algorithm is initialized orderly according to subband significance. Subband significance is defined as the number of significant coefficients in the subband. The more a subband contains significant coefficients, the more significant the subband is. This ensures that significant coefficients can be scanned and coded as much as possible. By testing many images, we found that the order of subband significance basically does not change when threshold descends at 32 (i.e. when bit-plane descends at 5). Therefore, the order of subband significance is initialized as its order at threshold = 32.

In order to decode correctly, the initialization order of subbands need to be coded. As there are only 16 subbands in each scale in the proposed coding algorithm, if the order of first 15 subbands has been decided, the sixteenth subband must have been decided too. This means that only the sequence numbers of first 15 subbands need be coded in-order. Relative to being able to coding more significant coefficients and all output bits, the bits used to code the sequence numbers are negligible.

4.2 Construction of Virtual LF Subband

In the proposed coding algorithm, image is decomposed by LP firstly. Subsequently, each scale HF image is decomposed by DFB. The number of directional decomposition at each scale HF image is same (N=16) as illustrated in Fig. 2 (a). In contourlet HF image, the size of ith scale is 4 times as that of $(i+1)$th, and they are all decomposed into 16 subbands. It is easy and clear to find RFC among HF images i.e. tree structure like zerotree in EZW and spatial-orientation tree in SPIHT.

However, there is not apparent direction in approximation image i.e. LF image, especially when the size of LF image is very small in CT. LF image is not performed directional decomposition further. So we can not construct tree structure between LF image and HF images directly.

It is well known that the more zero coefficients an algorithm can cluster together, the higher compression ratio it can achieve. As for tree structure, the higher tree is, the easily above condition is achieved. Therefore, we want to construct RFC between LF image and HF image, which can make tree structure higher in contourlet coefficients.

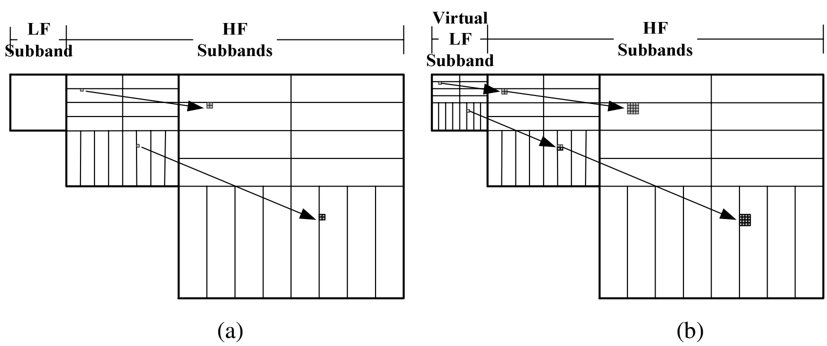


Fig. 3. The comparison of tree structures. (a) Tree structure in contourlet coefficients without virtual LF subband. (b) Tree structure in contourlet coefficients after constructing virtual LF subband.

In our proposed coding algorithm, a virtual LF subband is constructed. It has the same size and contains the same pixels as the original LF image, but is partitioned imaginarily. Unlike HF images, virtual LF subband is partitioned spatially into 16 subbands but not performed directional decomposition. It makes the construction of RFC between LF image and HF images easy. Fig. 3 shows the comparison of tree structures between in contourlet coefficients without virtual LF subband and in contourlet coefficients after constructing virtual LF subband. It is obvious that when image is decomposed by LP into the same scales, tree structure in contourlet coefficients after constructing virtual LF subband is higher than that in contourlet coefficients without virtual LF subband (i.e. the height of tree structure in contourlet coefficients is added by 1). Therefore, the construction of virtual LF subband can achieve better compression performance.

4.3 Adjusted SPIHT for Contourlet

Because SPIHT was proposed based on WT, its definition of some concepts, initialization content and coding structure is not suitable for CT. In order to make the most of SPIHT to code contourlet coefficients, it must be adjusted properly.

Unlike wavelet decomposition, where each scale image except the highest scale only contains 3 subbands LH, HL and HH, each scale image in contourlet decomposition contains more subbands. In our proposed coding algorithm, 16 subbands are contained in each HF image. In SPIHT, spatial-orientation trees are constructed as shown in Fig. 4 (a). In order to adopt SPIHT to code contourlet coefficients, we organize scales and subbands into a new format.

In the proposed coding algorithm, each scale image including virtual LF image is regarded as a whole firstly and is set in the same direction. Because virtual LF image is decomposed into 16 directions too, tree structure can be constructed easily from virtual LF image to the highest frequency image as shown in Fig. 4 (b).

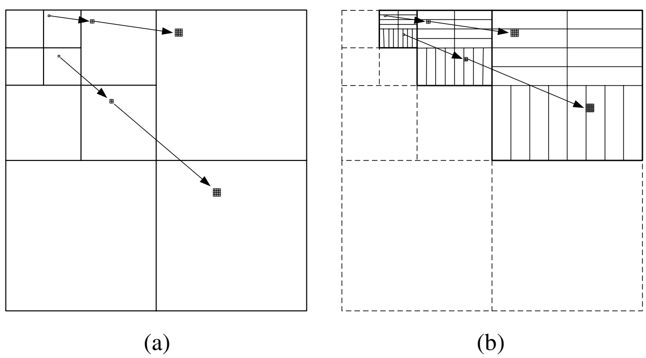


Fig. 4. Tree structures in two coding algorithms. (a) Tree structure in SPIHT. (b) Tree structure in the proposed coding algorithm. There, virtual LF image has been decomposed into 16 subbands and is constructed RFC with HF images.

In initialization, all coefficients belonging to the highest scale are added to LSP and only those with descendants are also attached to the LIS as type A entries in SPIHT. However, in our proposed coding algorithm, genuine LF coefficients are added to LSP and each subband of virtual LF image is added orderly to LIS according to subband significance as type A entries. By proper adjusting in initialization, the proposed coding algorithm can easily make use of SPIHT to code contourlet coefficients.

Furthermore, because the highest bit-plane of coefficients in LF image is higher than that of coefficients in HF images with a great probability, we output two quantization-steps (QP) *n1* and *n2*. QP *n1* is the value of the highest bit-plane of coefficients in LF image, and QP *n2* is the value of the highest bit-plane of coefficients in all HF images. In the proposed coding algorithm, LIS list will be checked only when *n1* decreases to the value of *n2*. Thus, many bits can be saved for signing spatial-orientation zerotrees in higher bit-planes and compression performance is improved further. In practical coding, only *n1* and the difference between *n1* and *n2* are output.

In order to illuminate our adjustment in the process of initialization, the proposed initialization for contourlet coefficients in SPIHT is presented in its entirety below.

Initialization:

1) output QP *n1* and the difference between QP *n1* and QP *n2*.
2) add coordinates $(i, j) \in$ genuine LF subband to LIP.
3) add each subband of virtual LF image constructed in 4.2 orderly to LIS according to the subband code order (subband significance) determined in 4.1 as type A entries.

By determining subband code order, constructing virtual LF subband and adjusting the initialization of SPIHT, our proposed algorithm can scan and code contourlet coefficients efficiently.

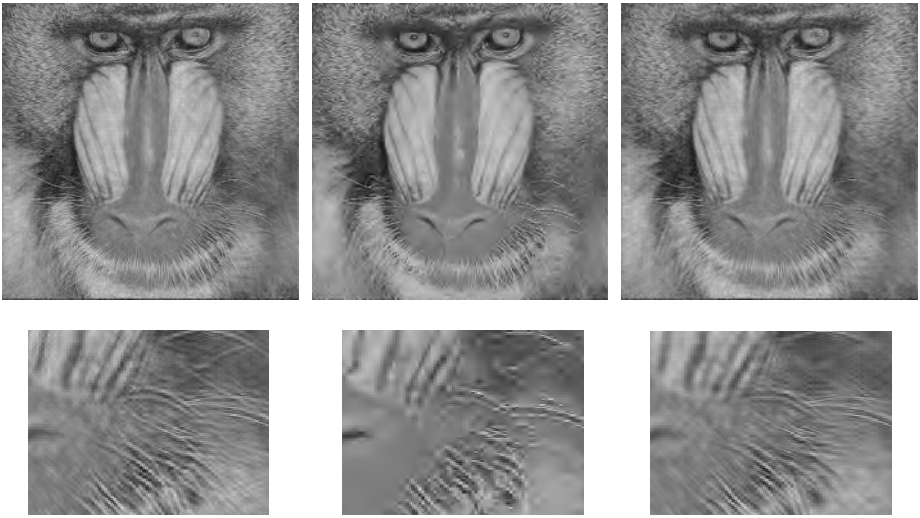
5 Experimental Results

In order to validate the high efficiency of the proposed image coding algorithm, extensive experiments have been carried out on 4 standard test images (8 bit/pixel (bpp), 512×512) at bit-rates of 0.15 bpp and 0.25 bpp. We presented 4-scale decomposition and used the 9/7 biorthogonal filters [18] for WT and the filters from Phoong *et al.* [19] for LP and the directional decomposition for each scale with N=16 separately. The output bit stream isn't further compressed with an arithmetic coder.

Table 2. Comparison of coding results on three coding algorithms (average PSNR in dB)

Image	Bit-Rate (bpp)	The proposed coding algorithm (dB)	Traditional WT with SPIHT(dB)	WBCT with SPIHT(dB)
Barbara	0.15	23.42	23.59	21.41
	0.25	26.15	26.46	24.73
Goldhill	0.15	26.74	25.77	25.22
	0.25	27.70	27.30	27.04
Mandrill	0.15	20.59	19.16	19.05
	0.25	21.71	20.39	19.88
Fingerprint	0.15	23.59	22.35	21.81
	0.25	26.02	26.44	25.06

Table 2 shows that average PSNR results with the proposed coding algorithm are 1.4 dB higher than original WT with SIPHT at best, and are 0.7-2.0 dB higher than WBCT with SPIHT in [17] respectively, which are both embedded image coding algorithm. Fig. 5 and Fig. 6 present the visual quality comparison of the proposed coding algorithm with original WT with SIPHT and WBCT with SPIHT, respectively. It is clear that the better subjective quality can be acquired by the proposed coding algorithm as genuine contourlet is adopted and more contours and textures are preserved. At the same time, the proposed coding algorithm has lower computing complexity than the coding algorithm in [17] as repositioning algorithm is discarded.



(a) (b) (c)

Fig. 5. The coding results of "Mandrill" image at 0.25 bpp. (a) The proposed coding algorithm. (b) WT with SPIHT. (c) WBCT with SPIHT.

(a) (b) (c)

Fig. 6. The coding results of "Barbara" image at 0.25 bpp. (a) The proposed coding algorithm. (b) WT with SPIHT. (c) WBCT with SPIHT.

6 Conclusion

This paper analyzed distribution of significant contourlet coefficients in different subbands and proposed a contourlet image coding algorithm based on adjusted SPIHT by adopting adaptive subband coding order scheme, constructing a virtual LF subband and adjusting coding structure of SPIHT properly. Experimental results show that the proposed coding algorithm has better or competitive compression performance than traditional WT with SPIHT and WBCT with SPIHT. At the same time, because genuine contourlet is adopted, the visual effect of the proposed coding algorithm is superior to them by preserving more contours and textures in the coded images.

In future work, we will investigate the correlation of contourlet coefficients in different directional subbands and different scales. CT for motion compensated image, expanding CT to 3-D and corresponding application for 3-D object image coding will be studied.

Acknowledgement

The authors acknowledge the financial support of RFDP No.20040248047 and Shanghai NSF No.04ZR14082.

References

1. Shapiro, J.M.: Embedded Image Coding Using Zerotrees of Wavelet Coefficient. IEEE Trans. on Signal Processing **41** (1993) 3445-3462
2. Said, A., Pearlman, W.A.: A New, Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees. IEEE Trans. Circuits Syst. Video Technol **6** (1996) 243-250
3. Taubman, D.: High Performance Scalable Image Compression with EBCOT. IEEE Trans. on Image Processing **9** (2000) 1158-1170
4. Do, M.N., Vetterli, M.: Contourlets: A Directional Multiresolution Image Representation. In: The IEEE International Conference on Image Processing (ICIP'2002), Rochester. Vol. 1. (2002) 357-360
5. Do, M.N., Vetterli, M.: The Contourlet Transform: An Efficient Directional Multiresolution Image Representation. Submitted to IEEE Trans. on Image Processing
6. Candès, E.J., Donoho, D.L.: Ridgelets: a Key to Higher-Dimensional Intermittency? Phil. Trans. R. Soc. Lond. A. (1999) 2495-2509
7. Do, M.N., Vetterli, M.: Orthonormal Finite Ridgelet Transform for Image Compression. In: The IEEE International Conference on Image Processing (ICIP'2000), Vancouver, Canada. Vol. 2 (2000) 367-370
8. Granai, L.: Radon and Ridgelet Transform Applied to Motion Compensated Images. EPFL, No 02.10 (2002)
9. Candès, E.J., Donoho, D.L.: Curvelets – a surprisingly effective nonadaptive representation for objects with edges. In: Cohen, A., Rabut, C., Schumaker, L.L. (eds.): Curve and Surface Fitting. Vanderbilt University Press, Saint-Malo (1999)
10. Candès, E.J., Donoho, D.L.: New Tight Frames of Curvelets and Optimal Representations of Objects with Smooth Singularities. Department of Statistics, Stanford University, Tech. Rep. (2002)

11. Pennec, E.L., Mallat, S.: Sparse geometric image representation with bandelets. IEEE Trans. on Image Processing **14** (2005) 423-438
12. Donoho, D.L.: Wedgelets: nearly-minimax estimation of edges. Ann. Statist. **27** (1999) 859–897
13. Donoho, D.L., Huo X.: Beamlets and multiscale image analysis. Multiscale and Multiresolution Methods. Springer Lecture Notes in Computational Science and Engineering. Vol. 20 (2001) 149-196
14. Do, M.N.: Directional multiresolution image representations. Ph.D.thesis, EPFL, Lausanne, Switzerland (2001)
15. Eslami, R., Radha, H.: On low bit-rate coding using the contourlet transform. In: Proc. of Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, USA. (2003) 1524-1528
16. Chappelier, V., Guillemot, C., Marinkovic, S.: Image Coding With Iterated Contourlet and Wavelet Transforms. In: The IEEE International Conference on Image Processing (ICIP '2004). Vol. 5 (2004) 3157 – 3160
17. Eslami, R., Radha, H.: Wavelet-based contourlet coding using an SPIHT-like algorithm. In: Proc. of Conference on Information Sciences and Systems, Princeton. (2004) 784-788
18. Antonini, M., Barlaud, M., Mathieu, P., Daubechies, I.: Image coding using wavelet transformation. IEEE Trans. Image Processing **1** (1992) 205-220
19. Phoong, S.M., Kim, C.W., Vaidyanathan, P.P., Ansari, R.: A new class of two-channel biorthogonal filter banks and wavelet bases. IEEE trans. on Signal Processing **43** (1995) 649-665