IMAGE CODING USING 2-D ANISOTROPIC DUAL-TREE DISCRETE WAVELET TRANSFORM

Jingyu Yang[*1], *Jizheng Xu*[2], *Feng Wu*[2], *Qionghai Dai*[1] *and Yao Wang*[3]

[1]Department of Automation, Tsinghua University, Beijing 100084, China
[2]Microsoft Research Asia, Beijing 100081, China
[3]Electrical and Computer Engineering Dept, Polytechnic University, Brooklyn, NY 11201

ABSTRACT

We propose an image coding scheme using 2-D anisotropic dual-tree discrete wavelet transform (DDWT). First, we extend 2-D DDWT to anisotropic decomposition, and obtain more directional subbands. Second, an iterative projection-based noise shaping algorithm is employed to further sparsify anisotropic DDWT coefficients. At last, the resulting coefficients are rearranged to preserve zero-tree relationship so that they can be efficiently coded with SPIHT. Experimental results show that our proposed scheme outperforms JPEG2000 and SPIHT at low bit rates despite the redundancy of DDWT.

***Index Terms*—** 2-D DDWT, anisotropic decomposition, redundant wavelet transform

1. INTRODUCTION

2-D Discrete Wavelet Transform (DWT) has received great success in image coding. It captures point singularities efficiently, but fails to capture directional structures which are often anisotropic at different orientations. Many tools have been invented to incorporate directional representation into the multiscale analysis framework [1][2][3].

DDWT proposed by Kingsbury is one of promising tools with the following three main advantages: direction selectivity, limited redundancy, and shift invariance [4][5]. The first two features are especially appealing for image coding. Basis functions with direction selectivity can characterize directional structures efficiently. Limited redundancy would facilitate sparse representation without imposing too much overhead of coding redundant locations.

Due to these advantages, Wang *et al.* first proposed a DDWT-based video coding scheme without motion compensation in which 3-D DDWT achieves better coding performance than 3-D DWT [6]. 3-D DDWT is later extended to anisotropic decomposition [7]. Unlike conventional dyadic decomposition, anisotropic decomposition produces directional basis functions of elongated shape. It thus contributes to some improvements on representation. However, only nonlinear approximation in terms of PSNR vs. numbers of retrained nonzero coefficients is investigated and no coding result has been presented in [7]. Recently, we propose an image coding scheme using DDWT, and the coding performance is comparable to that of JPEG2000 [8].

In this paper, we extend our previous work to 2-D Anisotropic DDWT (ADDWT) for efficient representation of directional features in images. For each level, 10 basis functions, instead of 6 basis function in 2-D DDWT, are obtained with ADDWT. An iterative projection-based noise shaping method [12] is then performed to get sparser coefficients so that only a small portion of them need to be coded. A rearrangement for resulting coefficients is employed to preserve zero-tree relationship for subband coding using SPIHT. According to experimental results, our proposed image coding scheme shows better performance compared with previous work in [8], and outperforms two popular DWT-based image coding schemes, SPIHT and JPEG2000, in terms of both objective metric (PSNR) and visual quality at low bit-rates.

The rest of this paper is organized as follows. Anisotropic decomposition of 2-D DDWT is described in Section 2. Section 3 introduces noise shaping for sparser representation. Experimental results are presented in Section 4. Finally, this paper is concluded in Section 5.

2. 2-D ANISOTROPIC DUAL-TREE DISCRETE WAVELET TRANSFORM (ADDWT)

In this section, we first briefly introduce 2-D DDWT and then describe its extension to anisotropic decomposition.

2.1. 2-D DDWT

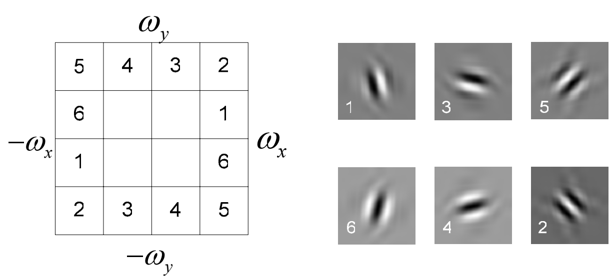
DDWT is a complex transform whose wavelet function is restrained to have single-sided spectrum. Either the real part or the imaginary part can be used as a stand-alone transform since they both guarantee perfect reconstruction. Meanwhile, DDWT is an overcomplete transform with redundancy of

[*] This work was done during the author's internship in Microsoft Research Asia.

$2^m:1$ for m-dimensional signals. Only the real part of DDWT is taken in coding applications to reduce the introduced redundancy [6][7]. For example, the redundancy will be reduced to 2:1 from 4:1 for 2-D case. The real part of DDWT is simply referred to as DDWT hereafter, unless otherwise stated.

The implementation of 2-D DDWT consists of two steps. Firstly, an input image is decomposed up to a desired level by two separable 2-D DWT braches, branch a and branch b, whose filters are specifically designed to meet the Hilbert pair requirements [4]. Then six high-pass subbands are generated: HL_a, LH_a, HH_a, HL_b, LH_b, and HH_b, at each level. Secondly, every two corresponding subbands which have the same pass-bands are linearly combined by either averaging or differencing. As a result, subbands of 2-D DDWT at each level are obtained as $(HL_a+HL_b)/2$, $(HL_a-HL_b)/2$, $(LH_a+LH_b)/2$, $(LH_a-LH_b)/2$, $(HH_a+HH_b)/2$, $(HH_a-HH_b)/2$.

The frequency tiling of 2-D DDWT is illustrated in Fig. 1a, where ω_x and ω_y are the cutoff frequency of 2-D input signal. Fig. 1b shows six wavelets at level 5 whose selective directions are $\pm 75^{\circ}$, $\pm 15^{\circ}$ and $\pm 45^{\circ}$ respectively. These directions are evenly distributed over the 2-D plane. The imaginary part of 2-D DDWT has similar basis functions as the real part. For more details on DDWT, please refer to [4][5].



(a) (b)

Fig. 1. (a) Frequency tiling of 2-D DDWT and (b) Basis functions of 2-D DDWT. A basis function and its ideal spectrum support are associated with the same number.

2.2. 2-D Anisotropic DDWT

The extension from dyadic decomposition to anisotropic decomposition of wavelet packets shows better adaptability to the features of images [9]. In anisotropic decomposition, subbands are allowed to be only decomposed vertically or horizontally rather than along both directions sequentially. In this way, anisotropic wavelet packets based on DWT provides basis functions with different aspect ratios which are thus anisotropic. However, the directions of these basis functions are still only horizontal, vertical, or diagonal. Noting that DDWT subbands are directional, incorporating anisotropic decomposition into DDWT (ADDWT) will generate anisotropic yet directional basis functions. For example, performing vertical decomposition on the 1st subband in Fig. 1 will produce two new subbands. The resulting basis functions and corresponding idealized spectrum supports are illustrated in Fig. 2. It can be observed that the resulting basis functions are indeed anisotropic and directional.

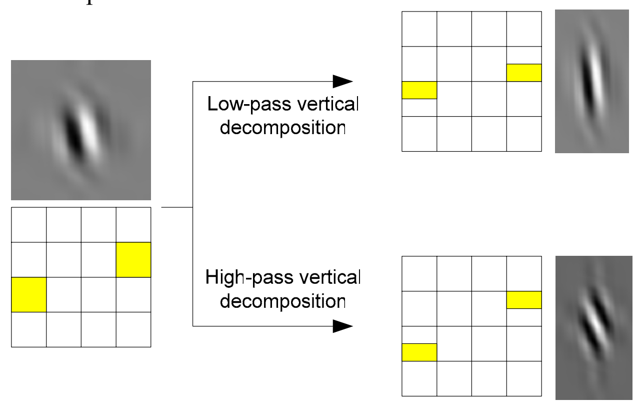
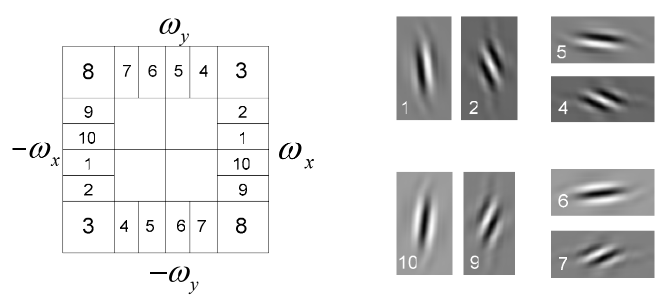


Fig. 2. Illustration of anisotropic decomposition on the 1st pass-band of 2-D DDWT.

Different decomposition strategies would generate different basis functions. Naturally, one would like to select basis functions that best adapt to the regarding image as in DWT-based anisotropic wavelet packets [9]. However, for compression applications, ADDWT coefficients need to be sparsified with noise shaping which will be introduced in Section 3. The computation complexity will increase dramatically if these two stages are jointly optimized. So we use a fixed decomposition pattern to ensure the anisotropy of resulting basis functions so that directional features in images are efficiently captured. The rules for anisotropic decomposition are given as follows.

- Decompose 1st and 6th subbands vertically
- Decompose 3rd and 4th subbands horizontally
- Leave 2nd and 5th subbands without decomposition



(a) (b)

Fig. 3. (a) Frequency tilling of 2-D ADDWT and (b) Basis functions of 2-D ADDWT. A basis function and its ideal support are associated with the same number.

As a result, ADDWT produces 10 subbands whose frequency tiling are shown in Fig. 3a. The corresponding basis functions are illustrated in Fig. 3b. They orient at the directions of $\pm 81^{o}$, $\pm 63^{o}$, $\pm 45^{o}$, $\pm 27^{o}$, and $\pm 9^{o}$ respectively. The 3rd and 8th basis functions are not shown in Fig. 1b since they are the same as the 2nd and 5th basis functions respectively. The reason for the third rule is that further decomposition on 2nd and 5th subbands does not give anisotropic basis functions. The advantages of ADDWT are twofold. On the one hand, locally high frequency components of images are characterized much more precisely by finer division of high-pass subbands. On the other, edges and contours in images are more efficiently represented by anisotropic basis functions oriented in finer directions.

3. SPARSIFY ADDWT COEFFICIENTS WITH NOISE SHAPING

To get a sparse representation for a given signal and a set of redundant dictionaries equals to solve an undetermined linear equation under sparseness constraints. Basis pursuit minimizes the l_1 norm of obtained coefficients via linear programming, and is thus computationally demanding [10]. Matching pursuit, a greedy algorithm, selects the atom which best matches the regarding signal at each iteration [11]. Although computational complexity can be reduced, matching pursuit is often trapped into suboptimum due to its greedy nature. With fast analysis and synthesis transform, noise shaping achieves desirable trade-off between sparseness and computation complexity.

Let $\mathbf{x} \in \mathbb{R}^{MN\times 1}$ be the vector form of an $M \times N$ input image (via vectorization). $\mathbf{A} \in \mathbb{R}^{2MN\times MN}$ represents the analysis matrix of ADDWT. $\mathbb{S} = \left\{ \mathbf{y} \mid \mathbf{y} = \mathbf{Ax}, \mathbf{x} \in \mathbb{R}^{MN\times 1} \right\}$ denotes the range space of $\mathbf{A}$ while its orthogonal complementary space is $\mathbb{S}^{\perp} = \left\{ \mathbf{y} \mid \mathbf{y}^T\mathbf{z} = 0,\ \mathbf{y} \in \mathbb{R}^{2MN\times 1}, \mathbf{z} \in \mathbb{S} \right\}$. Note that $\mathbb{S}^{\perp}$ is also the null space of synthesis matrix $\mathbf{R} \in \mathbb{R}^{MN\times 2MN}$, where $\mathbf{R} \simeq [\mathbf{A}^T\mathbf{A}]^{-1}\mathbf{A}^T$. So a solution $\mathbf{y}^{\mathbb{S}}$ lies in $\mathbb{S}$ plus any component $\mathbf{y}^{\perp}$ in $\mathbb{S}^{\perp}$ will give the same reconstruction as $\mathbf{y}^{\mathbb{S}}$ itself, i.e. $\mathbf{R}(\mathbf{y}^{\mathbb{S}} + \mathbf{y}^{\perp}) = \mathbf{R}\mathbf{y}^{\mathbb{S}}$. Naturally one would like to seek a proper $\mathbf{y}^{\perp}$ in $\mathbb{S}^{\perp}$ which makes $\mathbf{y} = \mathbf{y}^{\mathbb{S}} + \mathbf{y}^{\perp}$ as sparse as possible while perfect reconstruction is naturally guaranteed. Noise shaping introduces components of $\mathbb{S}^{\perp}$ into $y^{\mathbb{S}}$ with an iterative algorithm. More specifically, at each step, coefficients are quantized via thresholding. The reconstructed error in image domain will be projected onto $\mathbb{S}$, and then added back to the quantized coefficients. In this way, quantization error lies in $\mathbb{S}^{\perp}$ will be retained to strengthen large coefficients while weaken small coefficients. The effectiveness of NS has been verified in video coding [6][7]. For more details about noise shaping, please refer to [12].

4. EXPERIMENTS AND RESULTS

We evaluate the coding performance of our proposed method in this section. Three 512×512 grayscale images, *Barbara*, *Baboon*, and *Lena* are tested. Two state-of-the-art DWT-based coding methods, SPIHT with arithmetic coding and JPEG2000, are compared with ADDWT-based image coding scheme. To verify the contribution of anisotropic decomposition, results of DDWT-based image coding are also given. We code DDWT coefficients and ADDWT coefficients with SPIHT followed by arithmetic coding. Test images are decomposed up to 6 levels in each coder. The CDF 9/7 biorthogonal filters are employed for DWT and anisotropic decomposition stage of ADDWT. For DDWT, CDF 9/7 biorthogonal filters are used at the first level decomposition, and Qshift filters in [4] are used for the rest. For noise shaping, the initial threshold is set to 128, and then is decreased to zero by the step size of 1 for the remaining iteration.

To code DDWT coefficients with SPIHT, we concatenate two corresponding subbands of two DDWT trees horizontally. The structure of obtained coefficients looks as if it is generated by decomposing a 512×1024 image with 2-D DWT. For ADDWT, rearrangement is needed before this concatenation since zero-tree relationship is destroyed by anisotropic decomposition. We adopt the same way in [13], ensuring that zero-tree structure is preserved to facilitate efficient subband coding with SPIHT. For every two resulting subbands of anisotropic decomposition, every co-located 2×2 block in these two subbands are concatenated along the decomposition direction. Through this procedure, we virtually get the dyadic decomposition structure for which SPIHT is designed.

Coding performances for three test images are presented in Table I, Table II, and Table III respectively. ADDWT outperforms other three schemes at low bit-rates for most cases. Improvement is significant especially for images with rich directional features such as *Barbara*. Compared with 2-D DDWT, anisotropic decomposition gains about 0.3 dB on average. The reason is that, with anisotropic directional basis functions, 2-D ADDWT can represent directional structures with fewer coefficients. Fig. 4 shows enlarged reconstructed patches of *Barbara* and *Baboon* of JPEG2000 and the proposed scheme to compare subjective visual quality. It can be observed that ADDWT preserves the directional features in these two images much better. Obviously, ADDWT-based method produces much more visually appealing reconstructed images than conventional DWT does.

Table I Performance comparison for *Barbara*.

Bit-rate (bpp)	SPIHT	JPEG2000	DDWT	ADDWT
0.1	24.26	24.64	24.53	**24.81**
0.2	26.66	27.27	27.12	**27.77**
0.3	28.56	29.18	28.89	**29.59**
0.4	30.10	30.82	30.85	**31.33**
0.5	31.40	32.26	32.08	**32.54**

Table II Performance comparison for *Lena*.

Bit-rate (bpp)	SPIHT	JPEG2000	DDWT	ADDWT
0.1	30.22	29.86	30.31	**30.55**
0.2	33.11	32.93	33.31	**33.45**
0.3	34.87	34.79	34.94	**34.99**
0.4	36.15	35.98	36.24	**36.30**
0.5	37.08	**37.12**	**37.12**	37.00

Table III Performance comparison for *Baboon*.

Bit-rate (bpp)	SPIHT	JPEG2000	DDWT	ADDWT
0.1	21.34	21.35	21.28	**21.47**
0.2	22.69	22.63	22.51	**22.71**
0.3	23.76	23.65	23.78	**23.91**
0.4	24.66	24.62	24.60	**24.74**
0.5	**25.64**	25.55	25.39	25.49

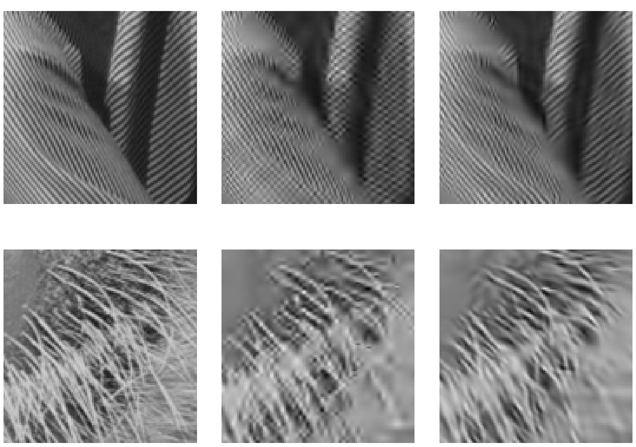


Fig. 4. Enlarged patches of *Barbara* (top row) and *Baboon* (bottom row) at 0.2 bpp. *Left*: original images. *Middle*: results of JPEG2000. *Right*: results of ADDWT.

5. CONCLUSIONS

In this paper, we propose an image coding scheme using 2-D anisotropic dual-tree discrete wavelet transform. To enhance the capability of capturing directional features at different scales, anisotropic decomposition is performed on DDWT to get anisotropic basis functions with more directions. Noise shaping is then employed to further sparsify ADDWT coefficients. With SPIHT the subband method, it is demonstrated that the proposed coding scheme outperforms two popular DWT-based schemes, JPEG2000 and SPIHT, at low bit-rates in terms of both PSNR and subjective visual quality.

6. ACKNOWLEDGEMENTS

This work is partially supported by the Joint Research Fund for Overseas Chinese Young Scholars of NSFC under grant No. 60528004.

7. REFERENCES

[1] M. N. Do and M. Vetterli, “The contourlet transform: an efficient directional multiresolution image representation”, *IEEE Trans. Image Processing*, vol. 14, pp. 2091-2106, Dec. 2005.

[2] E. L. Pennec and S. Mallat, “Sparse geometric image representations with bandelets”, *IEEE Trans. Image Processing*, Vol. 14, pp. 423-438, Apr. 2005.

[3] V. Velisavljevic, B. Beferull-Lozano, M. Vetterli, and P. L. Dragotti, "Directionlets: anisotropic multidirectional representation with separable filtering", *IEEE Trans. Image Processing*, Vol. 15, pp.1916-1933, July 2006.

[4] N. G. Kingsbury, “Complex wavelets for shift invariant analysis and filtering of signals”, *Applied Computational Harmonic Anal.*, vol. 10, pp. 234-253, May 2001.

[5] I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury, “The dual-tree complex wavelet transform”, *IEEE Signal Processing Magazine*, vol. 22, pp. 123-151, Nov. 2005.

[6] B. Wang, Y. Wang, I. W. Selesnick, and A. Vetro, “Video coding using 3-D dual-tree wavelet transforms”, in *Proc. of Int. Conf. on Acoustics, Speech, and Signal Processing*, Philadelphia, USA, pp. 61-64, Mar. 2005.

[7] B. Wang, Y. Wang, and I. W. Selesnick, “An investigation of the efficiency of 3-D dual-tree discrete wavelet transforms and its variants for video representation”, in *Proc. of Picture Coding Symposium*, Beijing, China, Apr. 2006.

[8] J. Yang, W. Xu, Q. Dai, and Y. Wang, "Image compression using 2-D dual-tree discrete wavelet transform (DDWT)", accepted by *IEEE Int. Symp. Circuits and Systems*, New Orleans, May 2007.

[9] D. Xu and M. N. Do, “Anisotropic 2-D wavelet packets and rectangular tiling: Theory and algorithms”, in *Proc. of SPIE Conf. on Wavelets X*, San Diego, USA, pp. 619–630, Aug. 2003.

[10]S. S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic decomposition by basis pursuit”, *SIAM J. Scientific Comp.*, vol. 20, pp. 33-61, 1999.

[11]S. G. Mallat and Z. Zhang, “Matching pursuits with time frequency dictionaries”, *IEEE Trans. Signal Processing*, Vol. 41, pp. 3397-3415, Dec. 1993.

[12]T. H. Reeves and N. G. Kingsbury, “Overcomplete image coding using iterative projection-based noise shaping”, in *Proc. of Int. Conf. Image Processing*, New York, pp. 597-600, Sep. 2002.

[13]H. D. Cho and J. B. Ra, “A rearrangement algorithm of wavelet packet coefficients for zerotree coding,” in *Proc. of Int. Conf. Image Processing*, Kobe, Japan, pp. 556-559, Oct. 1999.