A NEW COMBINATION OF 1D AND 2D FILTER BANKS FOR EFFECTIVE MULTIRESOLUTION IMAGE REPRESENTATION

Yuichi Tanaka[†‡*], *Masaaki Ikehara*[†], *and Truong Q. Nguyen*[‡]

[†]Department of Electronics and Electrical Engineering
Keio University, Yokohama, Kanagawa 223-8522 Japan
[‡]Department of Electrical and Computer Engineering
University of California, San Diego, La Jolla, CA 92093-0407
E-mail: tanaka.ych@gmail.com, ikehara@tkhm.elec.keio.ac.jp, nguyent@ece.ucsd.edu

ABSTRACT

In this paper, an effective multiresolution image representation using the combination of 2D quincunx filter bank (FB) and directional wavelet transform (WT) is presented. The proposed method yields simple implementation and low calculation costs compared to the other 1D and 2D FB combinations or adaptive directional WTs. Furthermore, it is a nonredundant transform and realizes quad-tree like multiresolution representation. In applications on nonlinear approximation and image coding, the proposed filter bank shows visual quality improvements and has higher PSNR.

***Index Terms*—** wavelet transforms, directional filter banks, image coding, nonlinear approximation.

1. INTRODUCTION

The traditional scheme to realize multiresolution image representation (MIR) is to apply 1D filters separately to horizontal and vertical directions, commonly referred to as "separable" transform. In contrast, "nonseparable" transforms consist of 2D filters and 2D downsampling matrices which cannot be factorized into 1D filter/downsampling pairs. The traditional wavelet transform (WT) is categorized as a separable transform, which is used in various applications. However, it has poor diagonal orientation selectivity since frequencies which represent different orientation are gathered into one subband in each resolution. For example, in image coding, these diagonal high-frequency coefficients are often truncated (not transmitted to the decoder side) and thus, the reconstructed image has blurred regions for diagonal orientations.

To reduce the artifact, some combinations of 1D separable and 2D nonseparable filter banks (FBs) have been proposed. The contourlet transform [1] is one of the most well-known transforms in this category, which obtains MIR as Laplacian pyramid [2]. The method shows good results in nonlinear approximation (NLA), however, it is a redundant transform due to the Laplacian pyramid. Redundancy is not desirable in many kinds of practical signal processing applications and thus its critically-sampled improvement, CRISP-contourlet [3], was recently developed. However, it requires several sets of 2D FBs which have specific (unusual) passband supports.

In [4], a simple tree-structure of 2D FBs, called uniform-quincunx DFB (uqDFB), was presented where "quincunx" means the downsampling matrix $\mathbf{Q}$ for 2D FBs is defined as $\mathbf{Q} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. The uqDFB consists of only diamond and fan support shape FBs. Furthermore, it satisfies the permissible condition which is required to have good frequency characteristics. The uqDFB can also merge its two lowest frequency resolutions to represent the low-frequency region effectively. It is called nonuniform quincunx DFB (nuqDFB) and the low-frequency resolution can be transformed by the separable WT. The nuqDFB shows better results in NLA than the contourlet due to its nonredundancy. However, it still remains a problem that the transformed image cannot generate the traditional quad-tree MIR and hence, the quad-tree coding cannot be used. To overcome this problem on the 2D FB-based MIR, directional FB (DFB) [5] is applied to the highpass subbands transformed by the separable WT in [6]. This hybrid wavelet and directional transform (HWD) is suitable for the conventional image coding method since every subband transformed by DFBs corresponds to that of the separable WT. The HWD yields better visual quality in reconstructed images than the separable WT when the set partitioning in hierarchical trees (SPIHT) [7] is employed as an image encoder. However, many transforms by 2D FBs are required when we need good orientation selectivity. Therefore, the computation cost for the HWD is much higher than that of the separable WT.

Another method is not to discard high-frequency regions for image coding at low bit rates. It needs only 1D WT filters and is based on transform along the "curves" in images [8, 9, 10]. In other words, 1D filters are rotated (or skewed) to fit lines in images (such as edges between objects and background). We denote this type of WTs as "directional WT" hereafter. First, the directional WT transforms an original image with the vertical downsampling matrix $\mathbf{M}_v = \text{diag}(1, 2)$, and then transforms with the horizontal one $\mathbf{M}_h = \text{diag}(2, 1)$ (or vice versa). After both downsamplings of the directional WT, one has four subbands corresponding to the traditional WT's LL, LH, HL, and HH. Obviously the LL subband can be transformed recursively, thus finally the quad-tree MIR is obtained. The directional WTs are suitable for the conventional quad-tree coders.

By transforming along the curves in images, the directional WTs preserve both high-frequency and low-frequency information even in low bit rates where the separable WTs often yield blurred artifacts. However, they have some drawbacks compared to the traditional separable WT. The main drawback is the computation cost to determine the curves in images. At each pixel, transforms are needed for several directions to decide the direction which maximizes the energy in LL subband (i.e. minimizes the energies in LH, HL and

*This work was supported by JSPS (Japan Society for the Promotion of Science) and Global COE Program "High-Level Global Cooperation for Leading-Edge Platform on Access Spaces (C12)".

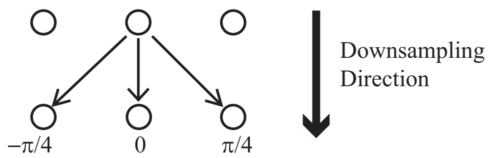

Fig. 1. Transform directions.

HH subbands). Furthermore, directional WTs support sub-pel and quarter-pel accuracy for each direction which contributes to better representation of curves along with a significantly higher computation cost than that of the separable WTs. Furthermore, they need to transmit the side information of the curves to decoder. Although this barely affects the entire bit budget to be encoded, it usually requires a careful manipulation to store or transmit since it should be lossless data.

To overcome the problems as previously mentioned, we propose a new combination of 1D and 2D FBs based on 2D quincunx FBs and the directional WT (QDW). It does not require any adaptive processing as in the conventional directional WTs. The proposed combination is simple, but preserves both low- and high-frequency information even after the quantization/truncation of many transformed coefficients.

2. QDW FRAMEWORK

In this section, we present the QDW framework. It is based on a simple method to avoid the direction search part on directional WTs. Furthermore, the QDW can be used iteratively and the resulting multiresolution image is similar to that using the traditional WT. The transform directions are defined as shown in Fig. 1.

2.1. 2D Part in QDW

In the QDW, the original image is firstly decomposed by two 2D quincunx FBs (see “Quincunx FB Part” in Fig. 3). One of them consists of a diamond filter and its complementary filter, and the other is a fan filter pair. The fan FB is fundamentally realized to modulate horizontally (or vertically) the diamond FB by $e^{j\pi}$, thus we now consider required properties for a diamond FB. The most important requirement in the QDW or other image processing using FBs is regularity of filters. The QDW transforms images by 2D FBs first, and then it performs the directional 1D part. Hence avoiding DC leakage in highpass filters is a challenging issue. In this paper, we employ the diamond FB proposed in [11] which is based on the direct optimization of its filter coefficients. Both lowpass and highpass filter sizes are 11×11 and there are two zeros at aliasing frequencies. There are many methods to design 2D diamond or fan FBs, however, the FB used here has a short highpass filter length. The high-frequency regions usually exist locally. The short highpass filter does not spread them into neighborhood regions. The fan FB is determined as a modulated version of the diamond FB.

Fig. 2 shows the frequency plane partition using quincunx FBs and the subband 2 of the 512 × 512 grayscale image *Barbara* after 1-level decomposition[1] by 2D quincunx FBs. The subband has the size of 256 × 256. The pixel values are emphasized to clearly show the directional characteristics. One notes that it captures the

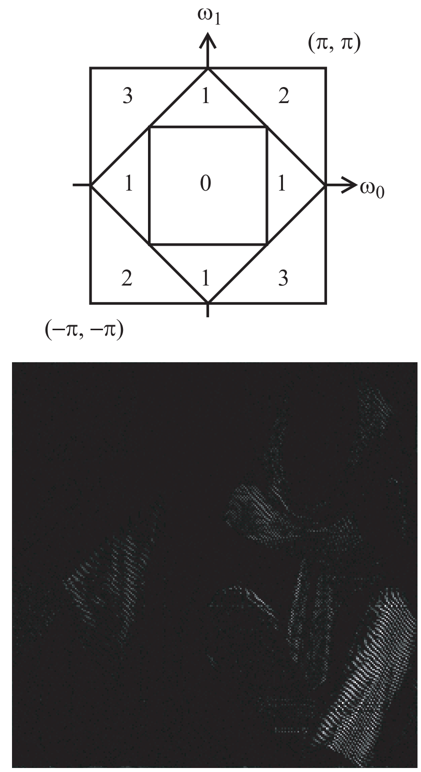

Fig. 2. Pertaining to the 2D QDW framework. (Top) Frequency plane partition where numbers represent subband indices. (Bottom) Subband 2 after 1-level decomposition which represents the direction along $\pi/4$.

high-frequency coefficients which are distributed on specific direction along $\pi/4$. Consequently, the 2D part in the QDW framework provides a new MIR with the downsampling matrix $\mathrm{diag}(2, 2)$: 1) lowpass, 2) highpass with the horizontal-vertical curves, 3) highpass with the curves along $\pi/4$, and 4) highpass with the curves along $-\pi/4$.

2.2. Directional 1D Part in QDW

The previous subsection shows that the 2D quincunx FBs decompose directional high-frequency energy in the original image well. The entire QDW framework is shown in Fig. 3. In this figure, arrowheads in “Separable Directional WT Part” represent the transform directions. In this paper, the 9/7 WT is used for subbands 1 to 3. It is realized by the lifting factorization of their filter coefficients which consists of two prediction and two update steps, and one scaling [12]. For subbands 2 and 3, the 9/7 directional lifting WT is applied in the direction $\pi/4$ and $-\pi/4$ first, and then in the direction $-\pi/4$ and $\pi/4$, respectively. The directional lifting WT for the direction $\pi/4$ with the vertical downsampling is shown in Fig. 4. In this figure, “p” and “u” denote coefficients for prediction and update steps, respectively. Additionally, white and gray circles indicate the even and the odd rows, respectively. Please refer to [10] for the implementation of the directional WTs as 2D FBs.

In this part, the curve determining process required in the adaptive directional WTs is no longer needed for the QDW since its 2D part divides the curves along the diagonal directions. After the di-

[1] The downsampling matrix of n-level decomposition is $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}^n$.

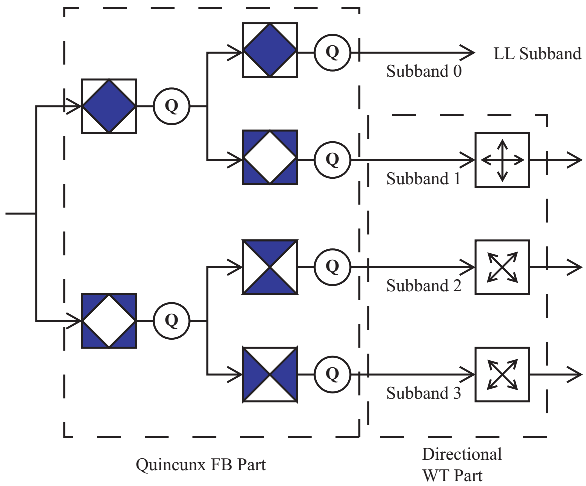

Fig. 3. QDW decomposition.

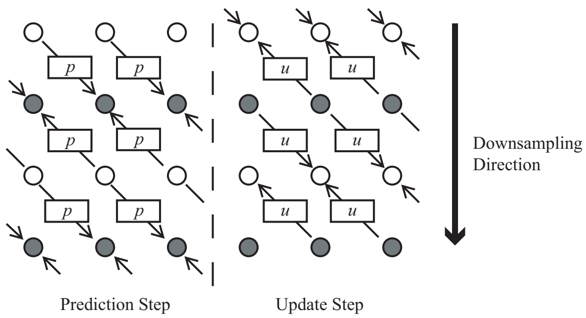

Fig. 4. Directional lifting WT along the direction $\pi/4$.

rectional part, there are 12 subbands corresponding to the subband 0. Obviously, the downsampling matrix for the subband 0 is $\mathrm{diag}(2, 2)$, hence its size is one fourth of the original image. The QDW can be applied to the subband 0 recursively if needed.

3. APPLICATION

In this section, we discuss applications using the QDW in NLA and image coding. For fair comparison, two benchmark results are presented. One is the separable 9/7 WT, and the other is the contourlet. 5-level decomposition is used for all of the transforms including the QDW. For the QDW, the decomposition level for 2D quincunx FBs can be changed. Hence, we show results from two designs: 1) 1-level QDW and 4-level WT (abbreviation is QDW(1, 4)) and 2) 2-level QDW and 3-level WT (abbreviation is QDW(2, 3)). In each level for QDW, an input subband is decomposed by quincunx FBs followed by the directional WTs (see Fig. 3). In this paper, *Barbara* image is used for all applications since it has rich textures.

3.1. Nonlinear Approximation

NLA is a good measure to estimate the potential of transform methods for MIR. All 5-level decomposed *Barbara* images are set to have the same number of the largest coefficients and the remaining coefficients are set to zeros. The comparison of NLA is presented in Fig. 5. Obviously both of the QDWs accomplish uniformly higher PSNR

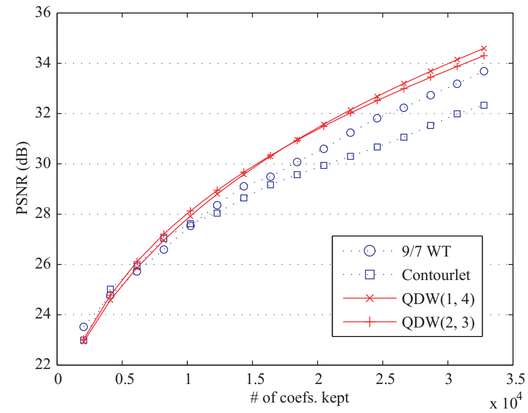

Fig. 5. NLA comparison.

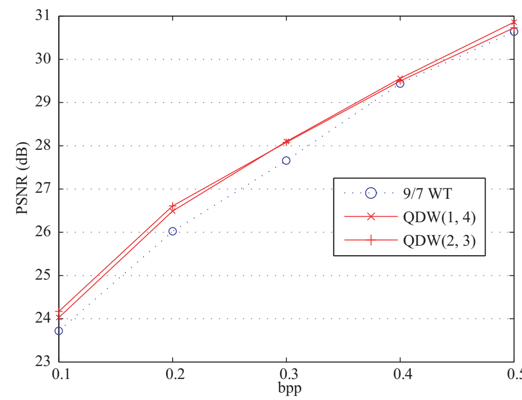

Fig. 7. SPIHT image coding results.

than the conventional methods. The contourlet has worse results when the number of kept coefficients increase since it is a redundant transform. The QDW(2, 3) is slightly worse than the QDW(1, 4) in the large number of kept coefficients. The reconstructed images with 5% of kept coefficients are shown in Fig. 6. The QDWs capture high-frequency regions (such as trousers) as well as the contourlets. Furthermore, smooth regions are reproduced as well as the 9/7 WT.

3.2. Image Coding

Image coding is one of the key applications using MIR. Based on the NLA results, the QDW is expected to have good image coding performance. In this subsection, we compare the proposed method with the 9/7 WT on the set partitioning hierarchical trees (SPIHT) progressive image transmission algorithm [7]. The results in various bitrates are depicted in Fig. 7. In low bitrate, the QDWs show better results than the 9/7 WT. To compare the perceptual visual quality, the reconstructed images for 0.25 bpp are shown in Fig. 8. Clearly the textures on the tablecloth are different. The QDW(2, 3) keeps the best textures, whereas the QDW(1, 4) has better texture presentation comparing to that using the 9/7 WT. There is a trade-off issue to select the QDW(1, 4) or (2, 3) between PSNR and textures, however, both show good visual quality.

Fig. 6. Reconstructed image used NLA. From left to right: 9/7 WT, Contourlet, QDW(1, 4) and QDW(2, 3).

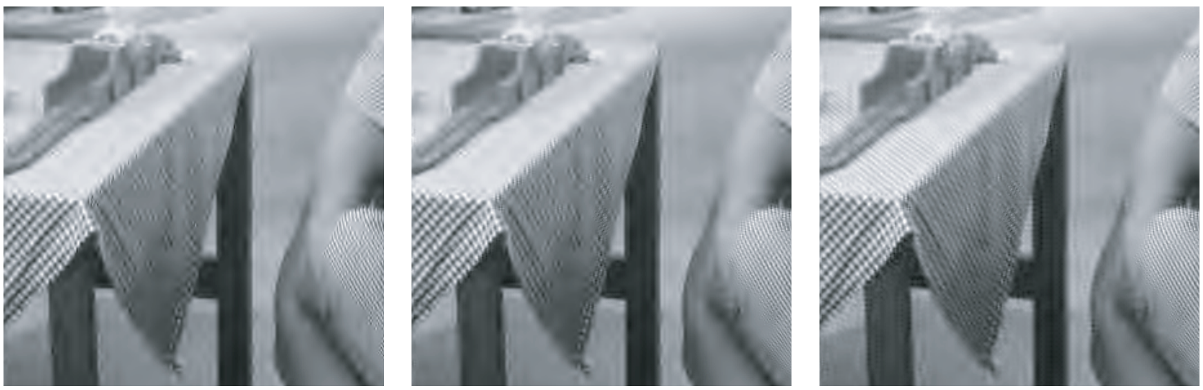

Fig. 8. Reconstructed image used SPIHT. From left to right: 9/7 WT, QDW(1, 4) and QDW(2, 3).

4. CONCLUSIONS

In this paper, we propose a new 1D and 2D FB combination named the QDW for better MIR. It yields better results than the contourlet or the 9/7 WT, and it can be implemented easily than the traditional 1D-2D framework due to small number of filterings by 2D FBs. Furthermore, the calculation of the curves in images is no longer needed in contrast to the adaptive directional WTs. The QDW is expected to have good performance in the other fields that require an efficient MIR (such as denoising). We plan to investigate these in future work.

5. REFERENCES

[1] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, 2005.

[2] P. Burt and E. Adelson, "The laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532–540, 1983.

[3] Y. Lu and M. N. Do, "CRISP-contourlet: a critically sampled directional multiresolution image representation," in *Proc. SPIE Conf. on Wavelet Applications in Signal and Image Processing*, 2003, pp. 655–665.

[4] T. T. Nguyen and S. Oraintara, "A class of multiresolution directional filter bank," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 949–961, 2007.

[5] R. H. Bamberger and M. J. T. Smith, "A filter bank for the directional decomposition of images: theory and design," *IEEE Trans. Signal Process.*, vol. 40, no. 4, pp. 882–893, 1992.

[6] R. Eslami and H. Radha, "A new family of nonredundant transforms using hybrid wavelets and directional filter banks," *IEEE Trans. Image Process.*, vol. 16, no. 4, pp. 1152–1167, 2007.

[7] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 243–250, 1996.

[8] D. Wang, L. Zhang, A. Vincent, and F. Speranza, "Curved wavelet transform for image coding," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2413–2421, 2006.

[9] W. Ding, F. Wu, X. Wu, S. Li, and H. Li, "Adaptive directional lifting-based wavelet transform for image coding," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 416–427, 2007.

[10] C.-L. Chang and B. Girod, "Direction-adaptive discrete wavelet transform for image compression," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1289–1302, 2007.

[11] T. T. Nguyen and S. Oraintara, "Multidimensional filter banks design by direct optimization," in *Proc. ISCAS'05*, 2005, pp. 1090–1093.

[12] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *Journal of Fourier Analysis and Applications*, vol. 4, no. 3, pp. 247–269, 1998.