

# Adaptive Nonseparable Interpolation for Image Compression With Directional Wavelet Transform

Weisheng Dong, Guangming Shi, *Member, IEEE*, and Jizheng Xu, *Member, IEEE*

**Abstract**—The adaptive directional lifting-based wavelet transform (ADL) locally adapts the filtering directions to the local properties of the image. In this letter, instead of using the conventional interpolation filter for the directional prediction with fractional-pel accuracy, a new two-dimensional nonseparable adaptive interpolation filter is proposed. The adaptive filter is calculated for every fractional-pel direction so as to minimize the energy of the prediction error. The tradeoff between reducing the prediction error and the overhead to code the interpolation filter is discussed. This enables coding gains of up to 0.98 dB, compared to ADL coder, and up to 2.4 dB, compared to the JPEG 2000 for typical test images.

**Index Terms**—Adaptive directional lifting, adaptive 2-D interpolation filter, rate-distortion optimization.

## I. INTRODUCTION

THE discrete wavelet transform (DWT) has become a powerful technique for image compression [1]. Typically, the two-dimensional (2-D) DWT used in these coders is applied as a separable transform by cascading two 1-D transforms, which are along the vertical and the horizontal directions. The main shortcoming of this transform is that it fails to provide an efficient representation for edges and textures, which are not aligned along these two directions. It would result in a large amount of signal energy in high-pass subbands. As a result, the ringing effects around edges can be observed clearly at low bit rates.

Lifting structure is an efficient and popular implementation of wavelet transforms [2]. Using lifting structure, several adaptive wavelet transforms, which adapt the filtering directions to the orientations of edges and textures, have been proposed [3], [4]. Since the constrained direction selection approaches are designed such that they can be reliably repeated at the decoder, these methods do not need to explicitly signal the filtering directions to the decoder. However, this leads to a reduction of direction accuracy. No significant improvement on objective quality measurement over traditional 2-D DWT was obtained.

Recently, new approaches that adaptively select filtering directions using lifting structure have been developed, for example, [5]–[8]. Different from the direction selection methods of the adaptive transforms mentioned above, the directions are accurately determined by minimizing the prediction error

energy and explicitly signaled to the decoder. These approaches have gained significant objective and subjective quality improvements for images with rich textures.

Fractional-pel direction resolutions, such as 1/4-pel in [5] and 1/2-pel in [6], are achieved by the use of interpolation. In [5] and [6], truncated Sinc interpolation filters are applied to interpolate the values of pixels at fractional positions to perform the fractional-pel predictions. However, the truncated Sinc filter is not the optimal choice of interpolation schemes for the images affected by aliasing, suffering from blurred details, and artifacts around edges [9]. In this case, the directional prediction technique becomes less efficient, and thus, the coding efficiency is reduced.

In this letter, a 2-D adaptive filter, similar to the adaptive interpolation filter for video coding [10], [11], is developed and seamlessly integrated into the directional wavelet transform to perform the fractional-pel prediction. Due to its adaptability to local image features, the directional prediction is improved, especially for the regions of edges and textures. Thus, the coding performance of the adaptive directional wavelet transform can be improved.

This letter is organized as follows. Section II presents the general scheme of the ADL transform in [5], while the proposed 2-D nonseparable adaptive filter for directional wavelet transform is described in Section III. In Section IV, the overhead for coding the filter coefficients is considered as a rate-distortion (R-D) optimization problem. Section V shows experimental results, and Section VI concludes this letter.

## II. ADAPTIVE DISCRETE LIFTING-BASED WAVELET TRANSFORM

In this section, we briefly review the general scheme of the ADL presented in [5]. The 2-D wavelet transform is generally separated into two 1-D wavelet transforms. Without loss of generality, we only describe the ADL transform in the vertical direction. It is shown in [2] that each 1-D wavelet transform can be factored into one or multiple lifting stages. A typical lifting stage consists of three steps: split, predict, and update. Let  $x(m, n)_{m, n \in \mathbb{Z}}$  be a 2-D image. First, all samples are split into two subsets, one consisting of all samples at even rows and the other one consisting of all samples at odd rows, i.e.,

$$\begin{cases} x_e(m, n) = x(m, 2n) \\ x_a(m, n) = x(m, 2n + 1). \end{cases} \quad (1)$$

Then, the samples at the odd rows are predicted from the neighboring samples at the even rows. The resulting prediction residuals, or high-pass subband coefficients, are calculated as follows:

$$h(m, n) = x_a(m, n) - P_e(m, n) \quad (2)$$

Manuscript received May 19, 2007; revised October 21, 2007. This work was supported by NSF China (No. 60736043, 60776795, and 60672125), "863 NHTRDP China" (No. 2007AA01Z307), and PCSIRT (No. IRT0645). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Alfred Mertins.

W. Dong and G. Shi are with the School of Electronic Engineering, Xidian University, Xi'an, Shaanxi 710071, China (e-mail: wsdong@mail.xidian.edu.cn; gmshi@xidian.edu.cn).

J. Xu is with Microsoft Research Asia, Haidian, Beijing 100080, China (e-mail: jzxu@microsoft.com).

Digital Object Identifier 10.1109/LSP.2007.914929

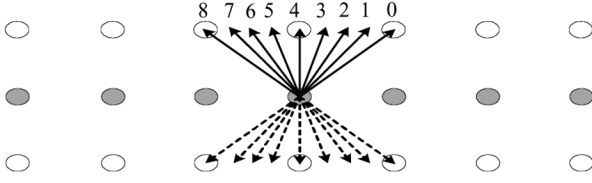


Fig. 1. Directions of the ADL transform.

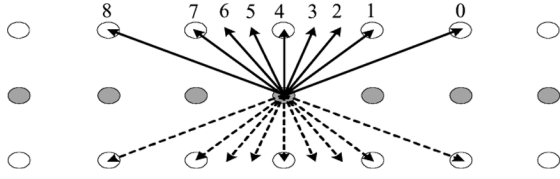


Fig. 2. The 1/3-pel accuracy prediction.

where  $P_e(m, n)$  is the predicting value, which is generated along a selected direction such that the magnitude of the residual  $h(m, n)$  is minimized. The candidate directions with 1/4-pel accuracy are shown in Fig. 1. The prediction along the direction  $d$  is given by

$$P_e(m, n) = \sum_i p_i x_e \left( m + (-1)^{i-1} \frac{d-4}{4}, n + i \right) \quad (3)$$

$$d = 0, 1, \dots, 8$$

where  $p_i$  is given by the wavelet kernel adopted, and  $x_e(m + (-1)^{i-1} \frac{d-4}{4}, n + i)$  is not necessarily sampled at an integer position. The truncated Sinc filter is applied to produce the pixel values at fractional positions. Finally, the samples at even rows are updated along the same direction as that in the prediction step. Consequently, the updating signal of the even rows samples is given as

$$u_h(m, n) = \sum_i u_j h \left( m + (-1)^{i-1} \frac{d-4}{4}, n + i \right) \quad (4)$$

$$d = 0, 1, \dots, 8.$$

By updating the values of odd rows, the low-pass signal can be found as follows:

$$l(m, n) = x_e(m, n) + u_h(m, n). \quad (5)$$

Since ADL can locally adapt the prediction direction to the geometric flow in the image, most of the large coefficients in the high-pass subbands are eliminated. This leads to significant objective and subjective quality improvements for images with directional contents.

### III. DIRECTIONAL WAVELET TRANSFORM WITH NONSEPARABLE 2-D ADAPTIVE INTERPOLATION

In order to improve the efficiency of the fractional-pel prediction, a new nonseparable 2-D adaptive filter scheme is developed. It is shown in [5] that the coding gains of the 1/4-pel accuracy prediction are slightly better than those of 1/2-pel accuracy prediction. Thus, 1/3-pel accuracy prediction can be adopted, reducing the number of fractional-pel directions from 6 to 4, as shown in Fig. 2. For every 1/3-pel direction, an individual set of coefficients is analytically calculated to perform the prediction. Without loss of generality, we only illustrate the prediction along the direction 2. As shown in Fig. 3, the sam-

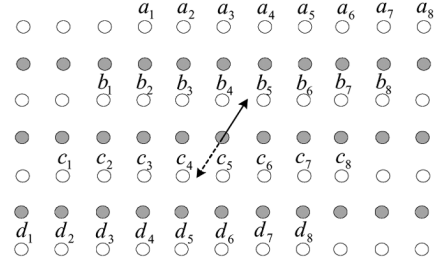


Fig. 3. Prediction along direction 2.

ples  $a_1, \dots, a_8, b_1, \dots, b_8, c_1, \dots, c_8, d_1, \dots, d_8$  are weighted by the coefficients of a 2-D  $4 \times 4$ -taps filter to generate a prediction value for the fractional-pel direction 2. The set of pixels  $a_1, \dots, a_8$  are selected as neighboring integer pixels along the 1/3-pel accuracy directions. For all fractional-pel directions, the filter coefficients are calculated so as to minimize the prediction error energy. However, for integer pixel directions, i.e., directions 0, 1, 4, 7, and 8, there are integer pixels along these directions, and thus, there is no need for interpolation.

Let us assume that  $\alpha_{0,0}, \alpha_{0,1}, \dots, \alpha_{3,7}$ , are coefficients of a 2-D  $4 \times 4$ -taps filter used for a fractional-pel direction. Then the fractional-pel prediction value is computed as follows:

$$P_e(m, n) = \sum_{i=0}^3 \sum_{j=0}^7 P_{i,j} \alpha_{i,j} \quad (6)$$

where  $P_{i,j}$  are pixel values at integer position,  $a_1, \dots, a_8$ , as shown in Fig. 3. The high-pass subband coefficients are

$$h(m, n) = x_a(m, n) - p_e(m, n) \quad (7)$$

where  $x_a(m, n)$  are the pixel values at the odd rows which are illustrated as gray dots in Fig. 3.

The update step followed the prediction step is carried out along the same direction with the same filter coefficients as that in the prediction step. Therefore, the updating signal of the samples at the even rows are given as

$$u_h = \sum_{i=0}^3 \sum_{j=0}^7 h_{i,j} \alpha_{i,j} \quad (8)$$

where  $h_{i,j}$  are the corresponding coefficients in the high-pass subband.

The directional lifting transform with adaptive filters is performed in the following steps.

- 1) Initial prediction directions are estimated using the 8-taps Sinc filter for interpolation purpose for the image to be coded. Among nine candidate directions, one that yields the minimal prediction error is selected, as the direction estimation presented in [5].
- 2) The 2-D filter coefficients  $\alpha_{i,j}$  are calculated for each fractional-pel direction independently by minimization of the prediction error energy, i.e.,

$$D = \sum_{m,n} \left| X_{m,n} - \sum_{i=0}^3 \sum_{j=0}^7 \alpha_{i,j} \times P_{m+2i-3, n+j-(2+i)(-1)^{\lfloor d/4 \rfloor - 7 \lfloor d/4 \rfloor} \right|^2 \quad (9)$$

where  $X_{m,n}$  are the values of pixels at the odd rows,  $P_{m+2i-3,n+j-(2+i)(-1)^{\lfloor d/4 \rfloor} - 7\lfloor d/4 \rfloor}$  are integer pixel values at even rows,  $d$  is the selected fractional-pel direction, and  $\lfloor \cdot \rfloor$  operator is a floor function, which round the fractional value to the nearest integer value. All the blocks, which share the same fractional-pel direction, are used to perform the error minimization. The minimization problem

$$\min_{\alpha_{0,0}, \alpha_{0,1}, \dots, \alpha_{3,7}} D \quad (10)$$

can be solved by the standard least-square method.

- 3) New prediction directions are estimated with the adaptive filters computed in step 2, and then the adaptive directional wavelet transform is performed. This step enables reduction of the signal energy of the high-pass subbands. Thus, the coding efficiency can be improved.

In order to adapt the filter coefficients to the new prediction directions, Steps 2 and 3 are repeated until a predefined quality improvement threshold is achieved. Since the encoder and decoder should use the same filters to ensure a perfect reconstruction, the filter coefficients have to be coded and transmitted as side information.

In this work,  $4 \times 32 = 128$  filter coefficients, corresponding to directions 2, 3, 5, and 6, have to be transmitted. However, the coding gain brought by adaptive 2-D filters can be drastically reduced due to the high overhead. In order to reduce the side information, the number of adaptive 2-D filters is reduced from 4 to 2 by considering the symmetry of statistical properties of an image signal, e.g., direction 2 and 6 can share one filter. The symmetry of the filter-kernels itself is also considered to reduce the number of coefficients of an adaptive 2-D filter from 32 to 16. As a result, only two adaptive 2-D filters have to be calculated and only 32 coefficients have to be transmitted.

Since the adaptive interpolation filters have to be computed for every input image, the computation complexity of the proposed transform is considerable larger than that of ADL transform. However, the inverse proposed transform is slightly more expensive than the inverse ADL transform since the interpolation filter coefficients are transmitted as side information. Thus, the proposed image compression method may be suitable for some internet and wireless applications, which allow offline image compression to achieve best rate distortion performance while ensuring real-time decoding.

#### IV. R-D OPTIMIZED ADAPTIVE INTERPOLATION FILTER

For an image with low-frequency characteristics, the Sinc filter might be an approximately ideal filter for most parts of the image. On the other hand, for the low spatial resolution image, the use of adaptive interpolation filter results in too expensive overhead for coding filter coefficients. In order for this new image codec to benefit from the adaptive filter, a R-D optimization scheme is developed to make a tradeoff between the cost of coding the filter coefficients and the reduction of signal energy of the high-pass subbands. Similar to the work presented in [5], the input image is recursively partitioned into blocks of variable sizes by an optimal quad-tree which minimizes the following cost function:

$$J = D(T) + \lambda R(T) \quad (11)$$

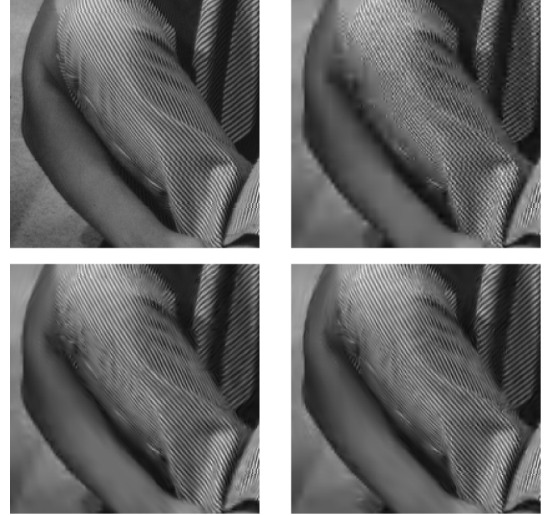


Fig. 4. Visual quality comparison of part of decoded Barbara image at 0.3 bpp. (Top-left) Original. (Top-right) JPEG 2000. (Bottom-left) ADL. (Bottom-right) ADL-Adaptive.

where  $D(T)$  is the distortion of the full quad-tree,  $R(T)$  is the rate of the full quad-tree, and  $\lambda$  is a predefined Lagrangian multiplier which depends on reconstruction quality of interest. Considering the additional overhead for filter coefficients, the cost function is modified to be

$$J = D(T) + \lambda R(T) + \lambda C(T) \quad (12)$$

where  $C(T)$  is the rate of the filter coefficients. To decide whether the image signal is interpolated using adaptive or Sinc filters, three steps are carried out: In the first step, the image is partitioned using Sinc filter for the interpolation purpose and the R-D cost function is calculated using (11). In the second step, the adaptive filter coefficients are calculated for the entire image, as described in the previous section. In the last step, applying the adaptive filter for fractional-pel prediction, the image is partitioned and the R-D cost is calculated again using (12). Then for the coded image, using the Sinc filter and the adaptive filter, the R-D costs are compared. If the adaptive filter outperforms the Sinc filter with respect to the R-D cost, the adaptive filter is chosen. Otherwise, Sinc filter is used for this transform.

#### V. EXPERIMENTAL RESULTS AND OBSERVATION

The adaptive 2-D interpolation for ADL transform (referred to as ADL-Adaptive) is integrated into the reference software VM9.0 of JPEG 2000. We compare the coding performance resulting from the proposed ADL-Adaptive scheme with that from the ADL scheme presented in [5]. The JPEG 2000 which uses conventional separable 2-D DWT implemented in VM9.0 is also compared. The JPEG 2000 testing images, Barbara ( $512 \times 512$ ), Lena ( $512 \times 512$ ), and Air3 ( $1024 \times 1024$ ), are used in this experiment. All experimental results are done with the 5/3 wavelet filter. The filter coefficients of ADL-Adaptive transform are uniformly quantized to 12 bits and a simple fixed length coding scheme is applied.

Fig. 4 presents the parts of the decoded Barbara images by the three approaches, all at 0.3bpp. Compared with the JPEG 2000 coder, both the ADL and ADL-Adaptive coders reconstruct the textures of the trousers very well, and they greatly reduce the

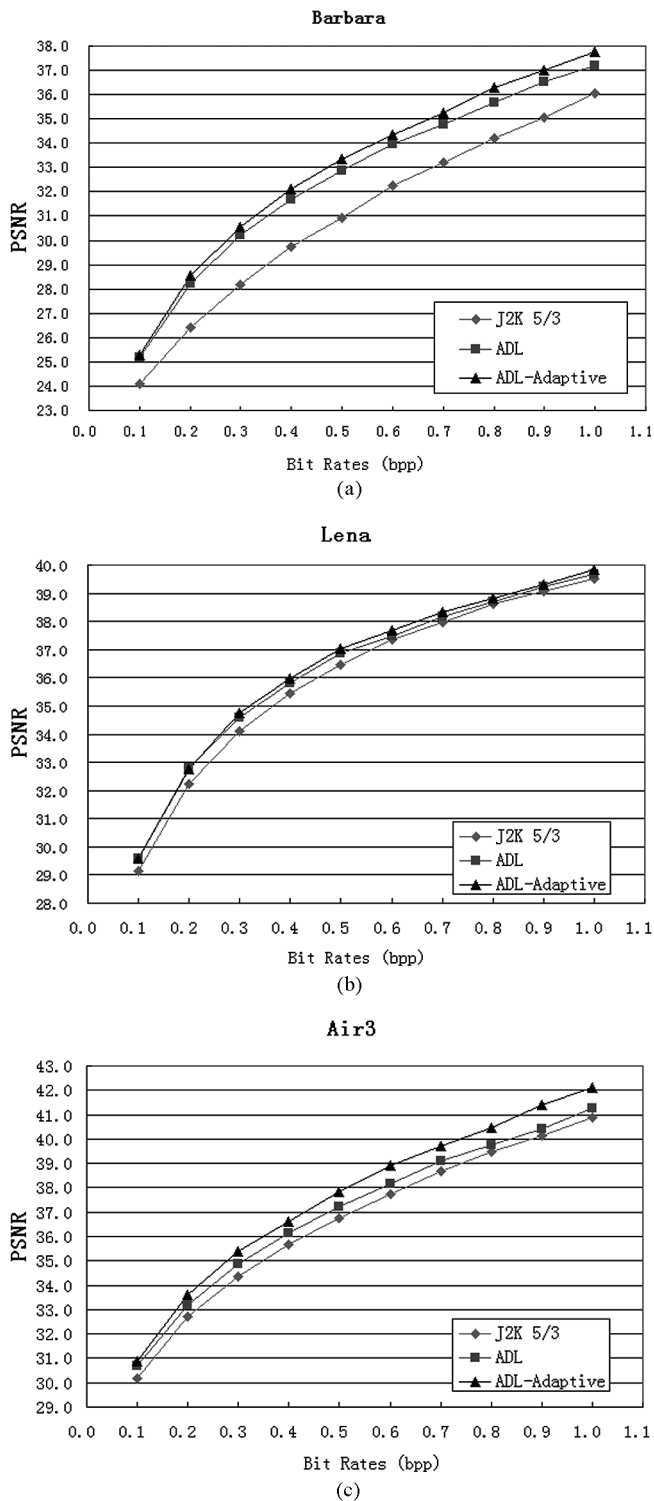


Fig. 5. PSNR versus bpp curves.

ringing effects around edges. It can be observed that the textures on the trousers and edges of the arm produced by ADL-Adaptive coder are sharper and less distorted by ringing effects than those produced by ADL coder. The compression performances for the test images are shown in Fig. 5. For all test images, both

ADL-Adaptive coder and ADL coder outperform JPEG 2000 at all bit-rates. The ADL-Adaptive coder outperforms ADL coder up to 0.6 dB for Barbara, 0.20 dB for Lena, and 0.98 dB for Air3. The coding gain for Lena is much smaller than those for other test images. It can be explained as follows. The adaptive filter is designed to improve the efficiency of directional prediction for high-frequency regions of the image signal. Linear interpolation, such as Sinc function, might be an approximately ideal filter for the images with low-frequency characteristics, leaving little room for improvements. In this case, adaptive filter is not chosen by RDO. Thus, the proposed image codec is especially suitable for the images with high-frequency characteristics.

## VI. CONCLUSIONS

A 2-D nonseparable adaptive filter for the ADL wavelet transform is presented. The adaptive filter is seamlessly integrated into the ADL wavelet transform to perform the fractional-pel prediction. The calculation of the filter coefficients is carried out analytically on the basis of minimizing the prediction error energy of the current image. Thus, the efficiency of the directional prediction is improved, especially for the local image structures, such as sharp edges, and textures. Experimental results demonstrate the effectiveness of the adaptive filter, which can further improve the coding performance of the ADL wavelet transform. A further improvement could be obtained if an efficient entropy coder is employed to code the filter coefficients. This will be done in future work.

## REFERENCES

- [1] D. Taubman and M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice*. Norwell, MA: Kluwer, 2001.
- [2] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *J. Fourier Anal. Appl.*, vol. 4, no. 3, pp. 245–267, 1998.
- [3] D. Taubman, "Adaptive, non-separable lifting transform for image compression," in *Proc. IEEE Int. Conf. Image Processing*, Kobe, Japan, Oct. 1999, vol. 3, pp. 772–776.
- [4] O. N. Gerek and A. E. Cetin, "A 2-D orientation-adaptive prediction filter in lifting structures for image coding," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 106–111, Jan. 2006.
- [5] 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, Feb. 2007.
- [6] C.-L. Chang, A. Maleki, and B. Girod, "Adaptive wavelet transform for image compression via directional quincunx lifting," in *Proc. IEEE Workshop Multimedia Signal Processing*, Shanghai, China, Oct. 2005.
- [7] C.-L. Chang and B. Girod, "Direction-adaptive discrete wavelet transform via directional lifting and bandeletization," in *Proc. IEEE Int. Conf. Image Processing 2006*, Atlanta, GA, Oct. 2006.
- [8] D. Wang, L. Zhang, and A. Vincent, "Curved wavelet transform for image coding," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2413–2421, Aug. 2006.
- [9] M. A. Rodriguez-Florido, J. Ruiz-Alzola, and C. F. Westin, "Artifact reduction in sinc interpolation using adaptive filtering," in *Proc. IEEE Int. Conf. Image Processing*, Thessaloniki, Greece, Oct. 2001, vol. 3, pp. 884–887.
- [10] T. Wedi and H. G. Musmann, "Motion- and aliasing-compensated prediction for hybrid video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 577–586, Jul. 2003.
- [11] Y. Vatis and J. Ostermann, "Locally adaptive non-separable interpolation filter for H.264/AVC," in *Proc. IEEE Int. Conf. Image Processing 2006*, Atlanta, GA, Oct. 2006.