WEIGHTED ADAPTIVE LIFTING-BASED WAVELET TRANSFORM

Yu Liu and King Ngi Ngan

Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong SAR
Email: {yuliu,knngan}@ee.cuhk.edu.hk

ABSTRACT

In this paper, we propose a new weighted adaptive lifting (WAL)-based wavelet transform that is designed to solve the problems existing in the previous adaptive directional lifting (ADL) approach. The proposed approach uses the weighted function to make sure that the prediction and update stages are consistent, the directional interpolation to improve the orientation property of interpolated image, and adaptive interpolation filter to adjust to statistical property of each image. Experimental results show that the proposed WAL-based wavelet transform for image coding outperforms the conventional lifting-based wavelet transform up to 3.02 dB in PSNR and significant improvement in subjective quality is also observed. Compared with the ADL approach, up to 1.18 dB improvement in PSNR is reported.

***Index Terms*—** Image coding, wavelet, weighted function, directional interpolation, adaptive interpolation filter

1. INTRODUCTION

Spatial transform has played an important role in most image and video coding methods. Wavelet transform has many advantages, such as multi-resolution representation, good energy compaction and decorrelation, it was adopted by JPEG-2000 standard.The wavelet-based JPEG2000 not only presents superior coding performance over the DCT-based JPEG, but also provides scalabilities in rate, quality and resolution.

The wavelet transform can be efficiently implemented by the lifting scheme [1], where the FIR wavelet filter can be factored into lifting stages. The key idea of the lifting scheme contains three stages: the first stage splits the samples $x[n]$ into even subset $x_e[n]$ and odd subset $x_o[n]$; the second stage uses the even subset $x_e[n]$ to predict the odd subset $x_o[n]$ and calculates the high-pass subband H as the prediction residual; the third stage uses the high-pass subband H to update the even subset $x_e[n]$ to ensure the preservation of moments in the low-pass subband L. The three stages are briefly referred as *split*, *predict* and *update* stage, respectively. The conventional lifting scheme uses the elements in neighbor horizontal or vertical direction, as shown in Fig.1(a). Although this kind of 2D lifting structure is very efficient in representing the horizontal and vertical edges, it doesn't work well when the edges are neither horizontal nor vertical. As a matter of fact, natural images often contain richly directional attributes, which can be commonly approximated as linear edges on a local level. These edges may be neither vertical nor horizontal. If such a fact is not taken into account, it would result in large magnitude in these high frequency coefficients. This problem has been pointed out by many researchers [2, 3].

In order to remove further the spatial redundancy in these directional attributes, directionally spatial prediction is incorporated into the conventional lifting scheme, which results in the adaptive directional lifting (ADL)-based wavelet transform [2]. In the ADL scheme, the odd subset $x_o[n]$ are predicted from the neighboring even subset $x_e[n]$ with an optimal direction in the predict stage. And in the update stage, the even subset $x_e[n]$ are updated from the high-pass coefficients with the inverse direction of the predict stage. Fig.1(b) shows the predict and update processes with an optimal direction in the ADL scheme, where the integer pixels are marked by black circles and the sub-pixels by white circles. For the sub-pixel interpolation, the popular *Sinc* interpolation is adopted, which is always performed in either the horizontal or vertical direction.

However, there are several problems existing in the ADL scheme. The first problem is the mismatch between the predict and update stages. When the optimal direction in the predict stage is located in sub-pixel precision, the high-pass coefficients cannot update exactly the pixels they are predicted, as shown in Fig.1(b). The second one is that, since the sub-pixel interpolation is always performed in either horizontal or vertical direction, it only favors horizontal or vertical direction and also may blur the orientation property existing in original images. The third is that the *Sinc* interpolation filter with constant coefficients is adopted for all images. However, different cameras, which have different low-pass filters, produce different aliasing components in the image signal. These varying aliasing components can not be considered by invariant filters [4].

In this paper, we propose a new weighted adaptive lifting (WAL)-based wavelet transform for image coding. The proposed WAL scheme provides a series of solutions to those problems that existed in the previous ADL scheme. In the following Section 2, various components of the proposed WAL scheme are described. Experimental results are reported in Section 3. Conclusion is drawn in Section 4.

This work was supported in part by a grant from the Chinese University of Hong Kong Direct Grant for Research Scheme under Project 2050383.

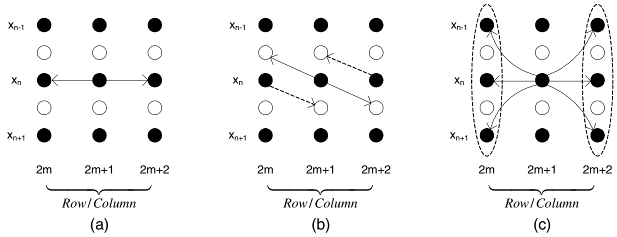


Fig. 1. Different lifting schemes, (a) conventional lifting, (b) adaptive directional lifting (ADL), (c) weighted adaptive lifting (WAL)

2. WEIGHTED ADAPTIVE LIFTING SCHEME

2.1. Weighted function

In the lifting stage, when the odd subset are predicted, all the even subset are available. An even better prediction can be expected if more available neighboring pixels are used as sources, as shown in Fig.1(c). Instead of a single pixel in the even subset as in the ADL scheme, a function of a set of integer pixel values is used as the input to the lifting stage. Function f is called as weighted function, which can take any integer pixels in the even subset as variables. The weighted function for directional prediction can be generally formulated as

$$f = \sum_k w_k x_e[n+k] \quad (1)$$

where w_k is the weighted parameter. When the displacement in directional prediction is located in integer pixel precision, the weighted function is defined as

$$f{=}\sum_k w_k x_e[n+k]{=}x_e[n+a], (w_{k=a}{=}1, w_{k\neq a}{=}0) \quad (2)$$

where a is the integer position. When the displacement in directional prediction is located in sub-pixel precision, the weighted function is defined as

$$f = \sum_k w_k x_e[n+k] = \sum_k \alpha_k x_e[n+k], (w_k = \alpha_k) \quad (3)$$

where α_k is the coefficient factor of the interpolation filter. In other word, the weighted function is the sub-pixel value calculated from neighboring integer pixels by a certain interpolation filter.

As depicted above, when the displacement in directional prediction is located in sub-pixel precision, the prediction and update stages may have mismatch. Thus, a weighted lifting scheme is proposed to solve this problem. The basic idea is that in the update stage the obtained high-pass coefficients are likewise distributed to those pixels that are used to calculate the high-pass coefficient in the predict stage. A similar principle is also used in the motion compensated temporal filtering (MCTF) for temporal transform [5], but is developed independently.

Since the high-pass coefficient $d[n]$ is obtained as follows:

$$d[n] = x_o[n] + \sum_i \sum_k p_i w_k x_e[n+k] \quad (4)$$

where w_k is the weighted parameter. Following the idea, the low-pass coefficient $c[n]$ is calculated as follows:

$$c[n] = x_e[n] + \sum_j \sum_l u_j w_l d[n+l] \quad (5)$$

It means that the high-pass coefficient will be added exactly to the pixels they are predicted. Now, the prediction and update stages are consistent.

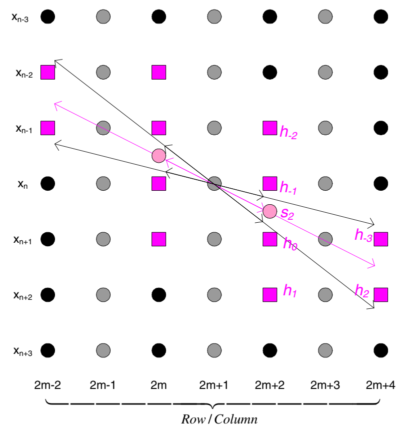


Fig. 2. Directional interpolation

Table 1. Applied filters of directional interpolation process

Filter name	Position	Filter coefficients
Bilinear filter	1/4	$(3, 1)/4$
	1/2	$(2, 2)/4$
	3/4	$(1, 3)/4$
Telenor 4-tap filter	1/4	$(-1, 13, 5, -1)/16$
	1/2	$(-2, 10, 10, -2)/16$
	3/4	$(-1, 5, 13, -1)/16$
2-tap filter	N/A	$(-1, 5)/4$

2.2. Directional interpolation

Since the interpolation used in the ADL scheme is always performed in either the horizontal or vertical direction, it may blur the orientation property existing in original images. To solve this problem, we propose a directional interpolation to improve the orientation property of interpolated image. As shown in Fig.2, for different sub-pixel position, we use different integer pixels to interpolate the sub-pixel. The interpolation is related to the predicted direction. For example, in order to interpolate the half pixel s_2, not only the integer pixels $\{h_{-2}, h_{-1}, h_0, h_1\}$in horizontal/vertical direction, but also the integer pixels $\{h_{-3}, h_2\}$ along the predicted direction are used. The proposed directional interpolation makes use of the directional information which is obtained from the weighted lifting scheme. Therefore, there is no any additional cost for the directional interpolation.

Directional interpolation filter is constructed by three kinds of filters: Bilinear filter, Telenor 4-tap filter [6] and 2-tap filter. Fig.3 shows the block diagram of the directional interpolation process, and Table 1 lists the coefficients of applied filters of the directional interpolation process. In the directional interpolation process, the pixels $\{h_{-3}, h_2\}$ are used as the inputs of Bilinear filtering process and the pixels $\{h_{-2}, h_{-1}, h_0, h_1\}$ are used as the inputs of the Telenor 4-tap filtering process, then the 2-tap filtering process takes the outputs from the Bilinear filtering and Telenor 4-tap filtering processes as its inputs. The final output of 2-tap filtering process is the directional interpolated sub-pixel. In fact, the filtering process of the block diagram in Fig.3 is equivalent to that of the directional interpolation filter listed in Table 2.

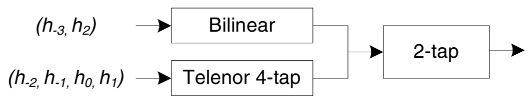


Fig. 3. Block diagram of directional interpolation process

2.3. Adaptive interpolation filter

In the ADL scheme, the popular *Sinc* interpolation filter with constant coefficients is adopted. However, the *Sinc* interpolation filter is designed to interpolate the horizontal or vertical direction of the image, regardless of the directional prediction. The filter coefficients are invariant, and the same interpolation filter is used for all images. However, different cameras, i.e. different image acquisition processes with different low-pass filters, produce different aliasing components in the image signal. These varying aliasing components can not be considered by invariant filters [4].

In order to reduce the directional prediction error and improve the coding efficiency, for each image the optimal interpolation filter, instead of invariant interpolation filter, is designed to minimize the energy of the high subband:

$$D^2 = \sum(|d[n]|)^2 = \sum(x_o[n] + \sum_i \sum_k p_i w_k x_e[n+k])^2 \quad (6)$$

The minimization problem

$$w_k = \arg\min E\left[(x_o[n] + \sum_i \sum_k p_i w_k x_e[n+k])^2\right] \quad (7)$$

can be solved by the Wiener-Hopf equation.

To decide whether the calculated optimal filter or the default filter in Table 2 is adopted, the rate constrained filter selection is performed by minimizing a Lagrangian cost function, which will be discussed in the next subsection in detail. Finally, the selected directional interpolation filter will be incorporated into the weighted lifting scheme, which leads to the proposed weighted adaptive lifting scheme.

2.4. Side information

Since the directional prediction is incorporated into the wavelet transform for image coding, the direction should be produced as side information. Therefore, one of important issues is to reduce the overhead bits from the direction information. Ideally, one optimal direction would be assigned to each pixel. But, in order to decrease the number of bits required for coding the direction information, the image is divided into 16x16 macroblocks. Like H.264/AVC, a tree-structured macroblock partition scheme is employed. From the view point of bit allocation strategies, the various modes relate to various bit-rate partitions.

To find the best direction vector, the rate constrained direction estimation for a given subblock S_i is performed by minimizing the following Lagrangian cost function

$$J_{dir}(S_i, w) = D_{S_i}(dv, w) + \lambda R(dv - dvp) \quad (8)$$

where D_{S_i} denotes the sum of absolute value of high-pass subband coefficients of current subblock, dv and dvp denote the estimated direction vector and its corresponding direction vector predictor, and $R(dv - dvp)$ denotes the number of bits to transmit dv. For the direction vector coding, only the differences between true direction vectors and their predictors

Table 2. Directional interpolation filter

Position	Filter coefficients
1/4	$(-4, -5, 65, 25, -5, -12)/64$
1/2	$(-8, -10, 50, 50, -10, -8)/64$
3/4	$(-12, -5, 25, 65, -5, -4)/64$

from the neighboring blocks are coded. The predictors can be generated using median prediction or directional segmentation prediction.

To find the best mode, the rate constrained mode decision for a given macroblock MB_j is performed by minimizing the following Lagrangian cost function

$$J_{mode}(MB_j, w) = \sum_i J_{dir}(S_i, w) + \lambda R(mode\_type) \quad (9)$$

where $R(mode\_type)$ denotes the number of bits to transmit the mode type. For the mode type coding, there are six modes including DEFAULT, DIRECT, 16x16, 16x8, 8x16, and 8x8. If the 8x8 mode is chosen, each of the four 8x8 sub-macroblocks may be further divided into 5 sub-modes, namely SM DIRECT, SM 8x8, SM 8x4, SM 4x8, and SM 4x4. The mode DEFAULT denotes the default horizontal/ vertical direction. The mode type will be entropy coded by UVLC (universal variable length coding).

To decide whether the calculated optimal filter or the default filter in Table 2 is adopted, the rate constrained filter selection is performed by minimizing the following Lagrangian cost function

$$J_{filter} = \sum_j J_{mode}(MB_j, w) + \lambda R(w - wp) \quad (10)$$

where w and wp denote the corresponding interpolation filter and its filter predictor. The predictor can be the filter in the preceding transform or the default filter in Table 2. For the coding of adaptive interpolation filter coefficients, a differential coding scheme is applied, where the differences to the filter-coefficients in the preceding transform or default filter-coefficients are coded. Since the coefficients are transmitted once per image, the additional bit rate is required for the coefficients is negligible.

In Eq.(8)-(10), λ is the Lagrangian multiplier which is used to control the rate-distortion contribution in these cost functions. Generally, at low bitrate, the slope of the rate-distortion curve is sharper than that at high bitrate, thus λ should be larger at low bitrate to properly estimate the rate cost percentage.

3. EXPERIMENTAL RESULTS

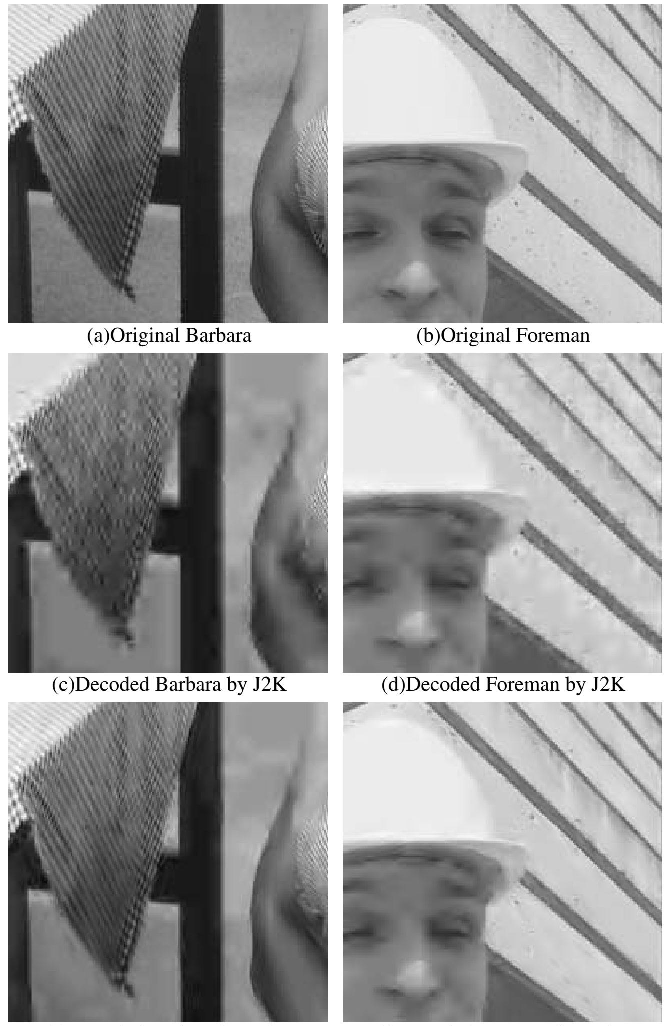
In the experiments, we compare three lifting schemes: conventional lifting scheme [1], adaptive directional lifting (ADL) scheme [2], and the proposed weighted adaptive lifting (WAL) scheme. In order to evaluate the WAL performance objectively, we simply replace the conventional lifting-based wavelet transform module of JPEG2000 with the proposed WAL transform and use the same bit-plane coding and EBCOT technique as in JPEG2000. The ADL approach also adopts the same substitution method to evaluate the performance.

We report the experimental results of four testing images: Barbara (512x512), Lena (512x512), Baboon (512x512), and

Table 3. Coding performance comparison(in dB) of different lifting schemes

Images	bpp	J2K	ADL	WAL
Barbara (512x512)	0.125	24.59	25.95	26.30
	0.25	27.38	29.22	30.40
	0.5	30.95	32.95	33.84
	1.0	36.04	37.24	37.99
Lena (512x512)	0.125	30.11	30.68	30.77
	0.25	33.22	33.88	33.90
	0.5	36.45	36.94	37.17
	1.0	39.51	39.73	39.81
Baboon (512x512)	0.125	21.40	21.39	21.57
	0.25	22.87	23.04	23.25
	0.5	25.17	25.34	25.81
	1.0	28.62	28.88	29.03
Foreman (352x288)	0.125	28.71	29.94	30.74
	0.25	32.32	33.53	33.78
	0.5	35.89	37.25	37.44
	1.0	40.60	41.30	41.43

the first frame of Foreman (352x288) video sequence. All images are decomposed by three-level wavelet transform with the same 5/3-tap wavelet filter in the three schemes. To compare the coding performance of the three schemes, we present the PSNR results of all test images at 0.125, 0.25, 0.5, and 1.0 bpp in Table 3. The coding gain can be up to 3.02dB for Barbara compared with J2K, and 1.18dB compared with ADL. For relatively smooth images like Lena, the WAL still shows advantages over the J2K and ADL. Fig.4 presents the decoded Barbara and Foreman images by J2K and WAL, both at the rate 0.25bpp. From the Barbara image decoded by J2K, there are severe pattern aliasing and blur effects on the scarf area which significantly damage the texture information of the original image. Also, edge ringing artifacts are clearly visible in both the J2K decoded Barbara and Foreman images. In contrast, the WAL method preserves texture information better, and greatly reduces the ringing artifacts around the edges.



(a)Original Barbara (b)Original Foreman

(c)Decoded Barbara by J2K (d)Decoded Foreman by J2K

(e)Decoded Barbara by WAL (f)Decoded Foreman by WAL

Fig. 4. Part of the decoded Barbara and Foreman images at rate 0.25bpp.

4. CONCLUSION

In this paper, we propose a new weighted adaptive lifting (WAL)-based wavelet transform for image coding. The main contribution of the proposed approach lies in three parts. First of all, the weighted function is used in the lifting stage to make sure that the prediction and update stages are consistent. Secondly, the directional interpolation is employed to improve the orientation property of interpolated image. Finally, the coefficients of interpolation filter are optimized to adapt to statistical property of each image. Experimental results show that the proposed WAL-based wavelet transform for image coding outperforms the conventional lifting-based wavelet transform up to 3.02 dB in PSNR and significant improvement in subjective quality is also observed. Compared with the ADL approach, up to 1.18 dB improvement in PSNR is reported.

5. ACKNOWLEDGMENT

The authors would like to thank Dr. Feng Wu, Microsoft Research Asia, for many insightful discussions on the topic.

6. REFERENCES

[1] W. Sweldens, ”The lifting scheme: A custon-design construction of biorthogonal wavelets,” Appl. Comput. Harmon. Anal, vol. 3, no. 2, pp. 186-200, 1996

[2] W. Ding, F. Wu, S. Li, ”Lifting-based wavelet transform with directionally spatial prediction”, Picture Coding Symposium 2004, USA, Dec. 2004

[3] O.N. Gerek, A.E. Cetin, ”A 2D orientation-adaptive prediction filter in lifting structures for image coding,” IEEE Trans. Image Process., vol.15, no. 1, Jan 2006

[4] T. Wedi, ”Adaptive interpolation filters and high-resolution displacements for video coding”, IEEE Trans. Circuits Syst. Video Technol., Vol.16, No.4, pp.484-491, April 2006

[5] R. Xiong, F. Wu, J.Xu, S. Li and Y.-Q. Zhang, ”Barbell lifting wavelet transform for highly scalable video coding,” Picture Coding Symposium 2004, USA, Dec 2004

[6] T. Wedi, ”Interpolation Filter for motion compensated prediction with 1/4 and 1/8-pel accuracy”, ITU-T SG16, doc.Q15-J-14, Osaka, Japan, May. 2000.