A New Image Coding Scheme with Hierarchical Representation and Adaptive Interpolation

Xinfeng Zhang[1,2,3]
[2]Key Lab of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China
[3]Graduate University of Chinese Academy of Sciences Beijing 100049, China

Ruiqin Xiong[1], Siwei Ma[1], Wen Gao[1]
[1]Institute of Digital Media, Peking University, Beijing 100871, China
{xfzhang, rqxiong, swma, wgao}@jdl.ac.cn

***Abstract*—In this paper, we propose a new image coding scheme which combines the advantage of hierarchical (i.e. multi-resolution) representation, adaptive interpolation and rate-distortion optimization capability of block-based coding. We divide the coding process into two layers by partitioning the image pixels into four groups with chessboard structure. In the first layer, pixels in one of these partitions are encoded and reconstructed using the state-of-the-art coding scheme. In the second layer, directional adaptive interpolation is employed to predict and code the pixels in the remaining three partitions. Joint mode decision is performed for the co-located macroblocks of these three partitions. *Skip* mode is introduced to save coding bits for blocks where the interpolation process works very well. Experimental results demonstrate that the proposed coding scheme achieves very promising performance.**

***Index Terms*— video/image coding, adaptive interpolation, intra prediction, multi-resolution**

I. INTRODUCTION

The efficiency of visual signal compression has been continuously improved in the past thirty years. The latest standard H.264/AVC is not only the state-of-the-art scheme for video coding, but also one of the best schemes for image coding [1]. In H.264/AVC, an image is coded and reconstructed in a block-by-block manner. To exploit the spatial correlation among pixels, directional intra prediction from neighboring blocks is employed. The block-by-block processing mechanism not only facilitates the reconstruction feedback loop in intra prediction, but also enables the block-level rate-distortion optimized (RDO) coding mode decision. However, one limitation of H.264/AVC is that it does not fully exploit the spatial correlation among pixels, because the directional intra prediction in H.264/AVC is not very accurate, especially for the pixels far from the boundary of previously reconstructed blocks.

On the other hand, it has been widely recognized that hierarchical/multi-resolution analysis is highly efficient in coding data with strong spatial correlation. One evidence is the state-of-the-art still image coding standard JPEG2000, which achieves much better performance than JPEG. Using multi-level discrete wavelet transform (DWT), JPEG2000 represents image by several subbands. The wavelet analysis and synthesis filters actually provide efficient way for intra prediction. The limitation of JPEG2000, however, is that DWT is a kind of global transform, which makes signal-adaptive coding mode adaptation difficult.

To exploit the anisotropic spatial correlation among pixels, directional discrete wavelet transform (DDWT) has been studied in recent researches [2]-[3]. Although DDWT provides better performance over JPEG2000, it is hard to do rate-distortion optimization (RDO) since the reconstruction process is absent in the wavelet-based coding framework. This limits the further improvement of wavelet based coding.

Motivated by the individual advantages of H.264/AVC and JPEG2000, some novel image/video coding methods are proposed in many literatures (e.g. [4], [5]). In this paper, we propose a new image coding scheme, which combines hierarchical representation, directional prediction/interpolation and block-based reconstruction and optimization. In this scheme, we divided the coding procedure into two layers by partitioning the image pixels into four groups with chessboard structure. In the first layer, the pixels in one of these four partitions are encoded and reconstructed with existing image/video coding methods, e.g. H.264/AVC. In the second layer, pixels in the other three partitions are jointly encoded using block-based coding or line-based coding chosen by RDO criterion [6] [7]. For these two coding modes, different adaptive bidirectional interpolation methods are employed to predict the pixels. Due to the efficiency of adaptive bidirectional interpolation, the prediction errors are usually very small. To present the well-interpolated blocks more compactly, *skip* mode is introduced. In addition, the optimal quantization for different data partitions is also discussed.

The rest of the paper is organized as follows. Section II presents an overview of the proposed coding scheme. Section III describes the adaptive bidirectional interpolation for block-based and line-based coding modes. Joint mode decision and bit allocation are investigated in Section IV. Experimental results are shown in Section V and Section VI concludes this paper.

II. OVERVIEW OF HIERARCHAL IMAGE CODING FRAMEWORK

The framework of proposed image coding scheme is illustrated in Figure 1. Pixels of input image are partitioned into four groups with chessboard structure. The partitions are referred as partition *I*, partition P_i (i=0, 1, 2) as illustrated in Figure 2 and each block represents a pixel. In the first layer, the pixels of partition *I* are coded using existing coding methods (e.g. KTA) and reconstructed. In the second layer, we use an adaptive bidirectional interpolation method to predict pixel values to be coded. And we carry out RDO based mode decision on the three macro-blocks in the co-located position of P_i (i=0…2) jointly at a time and the reconstruction order for them is P_2, P_1, P_0. The adaptive bidirectional interpolation in latter partitions (P_0, P_1) uses reconstructed pixels in partitions *I* and P_2.

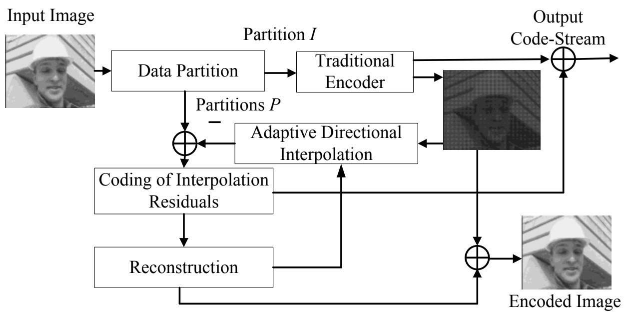


Figure 1. The proposed image coding framework with hierarchal representation and adaptive interpolation

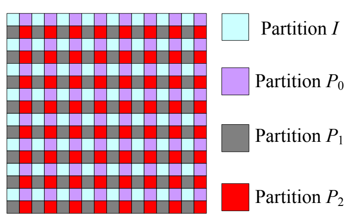


Figure 2. The data partition structure for a 16x16 image

III. ADAPTIVE BIDIRECTIONAL INTERPOLATION

In order to improve intra prediction accuracy in second layer of our scheme, we design two types of adaptive bidirectional interpolation methods as intra prediction methods, auto-regressive model (AR) based interpolation method and a set of fixed directional filters (FDF). In [6] and [7], line-based coding method employs similar intra prediction method and improves coding performance. However, it just employs unidirectional interpolation, because pixels used in prediction are limited on left and upper neighborhoods. In view of the success of line-based image coding method, we design our interpolation method for block and line based coding unit, respectively. In next sub-section, we take partition P_2 as an example to introduce the adaptive bidirectional interpolation method.

A. *Adaptive Interpolation for Block-based Coding*

In Figure 3, it illustrates the adaptive bidirectional interpolation structure for AR based and FDF methods, respectively. In block-based coding method, we only use the AR based interpolation method. The prediction value of current pixel in P_2 is generated by adaptive bidirectional interpolation with its 4-neighborhood pixels illustrated in Figure 3 (1). The interpolation coefficients are acquired by training with reconstructed pixels in neighborhoods. We denote the prediction pixels in training as vector x and the pixels used in prediction as vector y_i, i=0,…, L-1. L is the number of pixels used in AR interpolation. Then the interpolation filter coefficients $\boldsymbol{w}=\{w_i \mid i=0,\ldots,L\text{-}1\}$ are solved by the following optimization problem [4],

$$\boldsymbol{w} = \operatorname{argmin}\left\| \boldsymbol{x} - \sum_{k=0}^{L-1} w_k \cdot \boldsymbol{y_k} \right\|_2^2 \quad (1)$$

In the AR coefficients training process, we select training samples in a symmetry area around the current pixel. First, we select training samples in the same scale as Figure 3 (1). If pixels in this scale are not reconstructed yet, we select training samples in 2 times scale. This is reasonable because there is geometric duality in one image between different resolutions [8]. Therefore, our method not only uses bidirectional pixels in prediction, but also exploits more local samples in training than that in [7], which makes solution for equation (1) more stable.

The adaptive bidirectional interpolation method in P_1 and P_0 is similar with that in P_2. They use the around pixels illustrated in Figure 3 (3) and (4). The distance between prediction pixel and pixels used in prediction is closer than that in P_2. And this makes the prediction for P_1 and P_0 more accuracy than that in P_2.

B. *Adaptive Interpolation for Line-based Coding*

In order to improve the prediction efficiency in P_2, we take advantage the line-based coding type in this partition. In line-based coding type, the prediction and reconstruction in a block are performed line-by-line. Therefore, the coded pixels in prior line are used to predict pixels in the current line. In line-based coding, there are two coding orders, row-by-row and column-by-column. Coding order is selected according to the RDO criteria [6]. In our coding scheme, coding order is predicted from the corresponding position in partition *I*. And we define two kinds of adaptive interpolation methods to prediction pixels to be coded, AR based and FDF. Taking row-by-row order for example, the AR based adaptive interpolation structure is illustrated in Figure 3 (2). The interpolation coefficients are also acquired by solving the optimization problem in (1). The difference from block based coding is that the pixels in prior coded row are used and it makes the interpolation fit the vertical edge more accuracy. In order to depress the unstable feature of AR training, we also define a set of fixed directional interpolation filters (FDF), which are illustrated in Figure 3 (5). The 9 pixels in neighborhood are exploited in prediction. Both unidirectional and bidirectional interpolation filters are designed.

The selection of AR filter and FDF are according to prediction errors in surrounding reconstructed pixels, which generating minimal errors is selected. Therefore, there is no bit to send to index which filter is used. For partition P_1 and P_0, we only use block-based coding method, as a result of the enough accuracy of prediction.

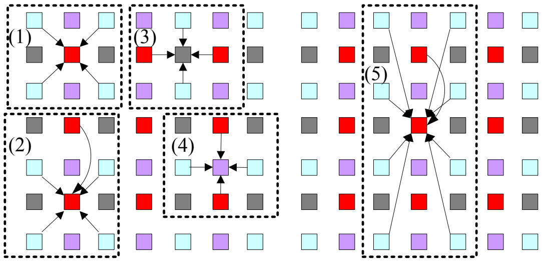


Figure.3. Adaptive bidirectional interpolation structure for pixels in different partitions

IV. THE OPTIMAZATION IN SECOND LAYER CODING

After prediction, the residuals in partition P_i are very few. It has two reasons. First, the pixels used for predicting partition P_i are closer than that in traditional intra prediction. Second, the adaptive bidirectional interpolation predicts unknown pixels very efficiently. Therefore, according to the characteristic of residuals in partition P_i, we propose some new designs to improve coding performance for them.

A. Mode Design in Second Layer Coding

In the second layer, we inherit the intra modes in H.246/AVC, block-based I4MB, I8MB and I16MB and block/line adaptive modes in [7]. Because the co-located three macro-blocks in second layer partitions come from the same image region with 1 pixel shift, they usually have similar structures and residuals characteristic. In order to fully exploit correlations between the three macro-blocks and improve encoding efficiency, we introduce *skip* mode into our image/intra-frame coding scheme, denoted as I_*skip*, and a joint mode, denoted as I_*joint*. We only write one bit for I_*skip* mode to represent that there are no residuals in the three macro-blocks. The reconstructed pixels are the values of the adaptive bidirectional interpolation in block-based I4MB type. If they are I_*joint* mode, we should write another bit to index that and only write one MB type next to it. Finally, if they have different MB types, we write three intra modes for them, respectively. In order to get the best performance, we use the RDO criteria for mode decision. We select mode with minimum RD cost for encoding the three macroblocks.

B. Bit Allocation

In the hierarchal image coding framework, pixel distortions in different partitions have different influence on the whole image. In order to improve the compression efficiency, we optimized the bit allocation for different partitions by changing quantization parameter (QP) for partitions. We denote D as the distortion (measured in MSE) of the whole image. Let the D_I and D_{P0} as the distortion in partition I and P_0. Although adaptive AR coefficients are not the same in most cases, the average of the AR coefficients can be regards equally in different directions. Therefore, we assume a line error propagation filter illustrated in matrices A and B in (2).

In (2), the matrix A represents that the one error in partition I and its propagation in around pixels in whole image. The matrix B illustrates the error propagation in partition P_2.

$$A=\begin{bmatrix} 0 & 1/16 & 0 & 1/16 & 0 \\ 1/16 & 1/4 & 3/8 & 1/4 & 1/16 \\ 0 & 3/8 & 1 & 3/8 & 0 \\ 1/16 & 1/4 & 3/8 & 1/4 & 1/16 \\ 0 & 1/16 & 0 & 1/16 & 0 \end{bmatrix} \quad B=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 0 & 0 \\ 0 & 1/4 & 1 & 1/4 & 0 \\ 0 & 0 & 1/4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{2}$$

Equation (3) and (4) illustrates the relationship D and QP, R and QP, respectively [7].

$$D = a \times 10^{b \bullet Q_p} \tag{3}$$

$$R = c \ln \frac{\sigma^2}{D} = c\left(\ln \sigma^2 - \ln a - \ln\left(10^{b \bullet Q_p}\right)\right) \tag{4}$$

where Q_p is the representative of QP and σ^2 is the signal variance. The relationship between D and D_I is illustrated in the following,

$$D = \left\| vec(A) \right\| \times D_I = wD_I = wa \times 10^{b \bullet Q_p} \tag{5}$$

With the rate-distortion optimization problem,

$$\min J = D + \lambda R \tag{6}$$

We can get relationship between w and Q_p as follows,

$$Q(w) = \frac{1}{b} \log_{10}\left(\frac{-\lambda cb}{wa}\right) \tag{7}$$

Therefore, the QP difference, ΔQ, between partition I and partition P0 can be formulated as follows,

$$\Delta Q = Q(w_{I0}) - Q(w_{P0}) = \frac{1}{b} \log_{10}\left(\frac{w_{P0}}{w_{I0}}\right) \approx -\frac{0.27}{b} \tag{8}$$

Based on experiments and the constant b is usually smaller than 0.1, we set partition I, P_2, P_0 and P_1 with QP-3, QP-2, QP, QP, respectively.

V. EXPERIMENTAL RESULTS

To evaluate the performance of our proposed hierarchal image coding (HIC) method, we implement our proposed method in KTA2.7. The partition I is encoded with the image coding method in [7]. At present, we do not have implemented special de-block filter and MDDT adaptive with our method yet. Therefore, in the experiment of this paper, we disable the de-block filter and MDDT for all the coding method. We compare our method with method in [7] (LIC), KTA2.7 [10] and JPEG2000 [11]. The intra coding is performed on luminance component only. The QP set is {35, 39, 43, 47} and the frame rate is 30 fps.

The experiment results are show in Table 1. We employ Bjontegarrd delta PSNR as described in [12] to calculate the average coding gains in terms of PSNR. The notation ΔB and ΔP are the bit-rate increase and PSNR gain compared with LIC, KTA2.7 and JPEG2000 in Table 1. Figure 4 illustrates some distortion-rate curves for different images. From these results, our proposed HIC method has a significant improvement over KTA2.7, LIC and JPEG2000. And our method has a much better performance for higher resolution 1920x1088. Overall, average PSNR gains of 1.11 dB, 0.97 dB and 0.91dB can be obtained by our proposed HIC coding method at low and medium bitrates comparing with KTA, LIC and JPEG2000, respectively.

TABLE I. COMPARISON OF EXPERIMENT RESULTS

Sequences		**HIC VS. KTA**		**HIC VS. LIC**		**HIC VS.JPEG2000**	
		ΔB(%)	*ΔP(dB)*	*ΔB(%)*	*ΔP(dB)*	*ΔB(%)*	*ΔP(dB)*
1280 x 704	Crew	-11.97	0.38	-8.34	0.24	-34.86	1.31
	Cydists	-23.60	1.14	-19.54	0.91	-34.08	1.77
	Raven	-21.13	0.81	-22.80	0.89	-19.93	0.69
	Vidyo1	-15.83	0.80	-10.78	0.50	-28.34	1.50
1920 x 1088	Riverbed	-22.33	0.99	-22.71	1.01	-3.94	0.15
	Kimono1	-28.17	1.50	-27.97	1.49	-6.35	0.26
	Station2	-29.00	1.26	-23.19	0.95	-29.00	1.08
	Rushhour	-34.45	2.37	-30.66	2.01	-10.56	0.60
	Tennis	-22.57	0.86	-19.43	0.73	-29.04	1.13
	Tractor	-20.59	1.01	-19.20	0.94	-12.16	0.56
Avg		-22.96	1.11	-20.46	0.97	-20.83	0.91

VI. CONCLUSIONS

In this paper we propose a new hierarchal image coding scheme. Adaptive bidirectional interpolation methods are used in intra prediction. Adaptive block/line coding unit is employed. We jointly code the macro-blocks in second layer partitions and introduce I_*skip* mode into our coding scheme. At the same time, approximation optimization bit allocation strategy is derived. Experimental results show that our proposed HIC scheme improves the image/intra coding performance significantly. Our method is implemented in two layers at present, but it is easily to expand to more layers. That will be worth to study in next work.

ACKNOWLEDGEMENT

This work was support in part by National Science Foundation (No. 61073083, No. 60803068), Beijing Natural Science Foundation (No. 4112026), Specialized Research Fund for the Doctoral Program of Higher Education (SRFDP) and National Basic Research Program of China (973 program) under Grant 2009CB320904.

REFERENCES

[1] F. D. Simone, M. Quaret, F. Dufaux, A. G. Tescher and T. Ebrahimi, "A comparative study of JPEG 2000, AVC/H.264, and HD Photo," SPIE Optics and Photonics, Applications of Digital Image Processing XXX, vol. 6696 (2007).

[2] 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.

[3] 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, May 2007.

[4] Xiaolin Wu, Xiangjun Zhang and Xiaohan Wang, "Low Bit-Rate Image Compression via Adaptive Down-sampling and Constrained Least Squares Upconversion," *IEEE Trans. Image Process*., vol. 18, no. 3, pp. 552–561, Mar. 2009.

[5] Weisi Lin, Li Dong, "Adaptive Downsampling to Improve Image Compression at Low Bit Rates," *IEEE Trans. Image Process*., vol. 15, no. 9, pp. 2513–2521, Sep. 2006.

[6] X. Peng, J. Xu and F. Wu, "Line-based image coding using adaptive prediction filters," in Proceedings of IEEE International Symposium on Circuits and Systems, ISCAS, May 2010.

[7] X. Peng, J. Xu and F. Wu, "Improved Line-based Image Coding by Exploiting Long-Distance Correlations," in Proceedings of SPIE Conference on Visual Communications and Image Processing, VCIP, Vol. 7744 77441D, Jul. 2010.

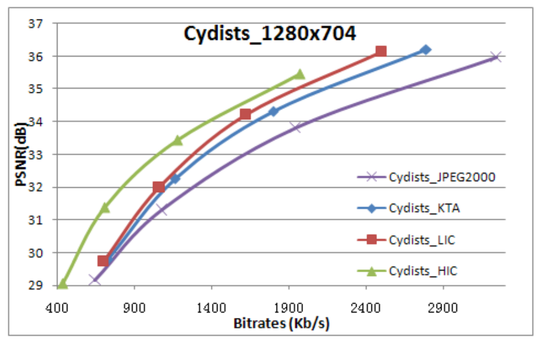
[8] X. Li and M. T. Orchard "New edge-directed interpolation," *IEEE Trans. Image Process*., vol. 10, pp. 1521, Oct. 2001.

[9] K. Takagi, Y. Takishima and Y. Nakajima, "A Study on Rate Distortion Optimization Scheme for JVT Coder," in proceedings of Visual Communications and Image Processing, VCIP, pages 914-923, 2003.

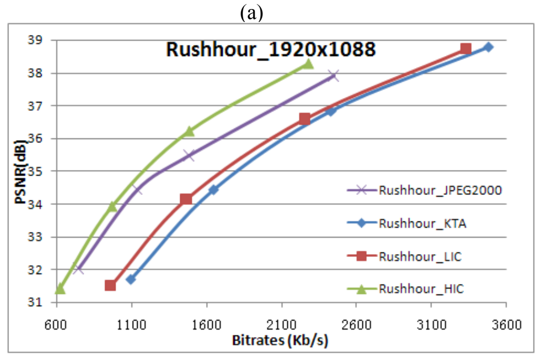
[10] http://iphome.hhi.de/suehring/tml/

[11] http://www.ece.uvic.ca/~mdadams/jasper/

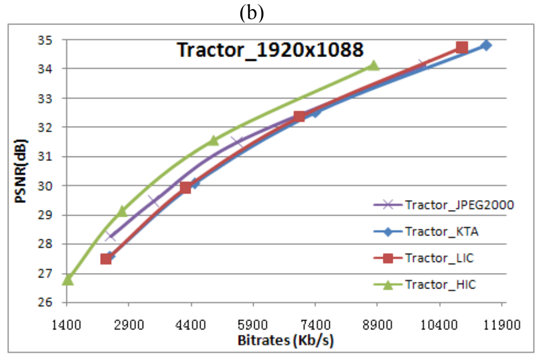
[12] G. Bjontegaard, "Calculation of Average PSNR Differences between RD-curves," ITU-T Q.6/SG16 VCEG, VCEG-M33, Apr. 2001.



(a)



(b)



(c)

Figure 4. Rate-PSNR curves of HIC versus those of KTA, LIC and JPEG2000 on different images