A Novel Weighted Cross Prediction for H.264 Intra Coding

Liping Wang, Lai-Man Po, Yusuf Md.Salah Uddin, Ka-Man Wong, and Shenyuan Li

lipiwang@cityu.edu.hk, eelmpo@cityu.edu.hk, ymdsalahu2@student.cityu.edu.hk,
kmwong@ee.cityu.edu.hk, shenyuli@cityu.edu.hk
Department of Electronic Engineering, City University of Hong Kong, Hong Kong SAR, China

Abstract—**In this paper, a novel weighted cross prediction (WCP) mode is proposed to replace DC mode in Intra_4x4 prediction of H.264/AVC. In the proposed scheme, the upper right part of one 4x4 block mainly employs vertical prediction while the lower left part mainly uses horizontal prediction, predicting both in vertical and horizontal directions in one block. This scheme uses simple prediction equations with fixed weighting coefficients. Experimental results show that WCP has improvement compared to H.264 and it is very competitive while comparing to other Intra_4x4 prediction algorithms.**

Index Terms—**video coding, prediction mode, intra prediction, weighted cross prediction (WCP)**

I. INTRODUCTION

Joint Video Team (JVT) of ITU-T Video Coding Experts Group (VCEG) and ISO Motion Picture Experts Group (MPEG) has released the latest standard for video coding, which is known as H.264 or MPEG-4 Part 10 Advanced Video Coding (AVC) [1-2]. The directional spatial prediction for intra coding is one of the significant contributions of H.264/AVC. In contrast with previous standards, such as H.263 [3] and MPEG-4 [4], intra prediction in H.264 is applied in spatial domain instead of transform domain, which utilizes the correlation between adjacent blocks to remove spatial redundant information in an image.

Many efforts have been done to simplify the intra coding in H.264. Fast intra mode decision algorithms were proposed to reduce the number of modes that needed calculation according to some criterion [5-6]. Besides, fewer efforts on the intra prediction modes innovations were raised, including altering the reference pixels [7], exchanging the order of intra prediction [8], and so on. Most of the algorithms can reduce the time complexity remarkably, but few can improve the coding efficiency.

Recently, distance-based weighted prediction (DWP) method and its simplified version integral DWP (iDWP) [9] were proposed to improve prediction efficiency, which are used to replace DC mode in Intra_4x4 prediction. The weighted prediction is done by coefficients which are inversely proportional to the distances between current pixel and its reference pixels. Although these methods give some improvements, they are computationally complex. This inspires us to propose a novel intra prediction mode called weighted cross prediction (WCP), which combines the vertical and horizontal predictions to replace the DC mode in Intra_4x4 prediction. The fixed weighting coefficients in our approach have a lower computation complexity with a comparable performance to DWP and iDWP.

The paper is organized as follows. We briefly introduce the intra prediction in H.264 in Section 2. In section 3, we describe our proposed method in details and analysis about the simulation results will be given in section 4. Finally, conclusions will be drawn in section 5.

II. INTRA PREDICTION IN H.264/AVC

The intra predictions in H.264/AVC are performed in a block-based manner, by referring to the neighboring samples of previously-decoded blocks which are left to and/or above the block to be predicted. Two primary types of luminance intra coding are supported: The Intra_4x4 prediction is suitable for the parts with significant details, while the Intra_16x16 is applied to the smoother areas.

Nine prediction modes are provided for every 4x4 block, named as vertical, horizontal, DC, diagonal-down-left, diagonal-down-right, vertical-right, horizontal-down, vertical-left and horizontal-up prediction. On the other hand, vertical, horizontal and DC modes, which are specified similar to the modes in Intra_4x4, with plane prdiction compose all the four prediction modes for 16x16 luma block. The encoder chooses the best prediction mode from the above 13 modes to minimize the Lagrangian cost function, which takes both distortion and bit rate into consideration [1].

Figure 1 illustrates 16 samples of 4x4 block (labeled as a-p) which are predicted by previously-decoded samples in the neighboring blocks (labeled as A-M), when using the Intra_4x4 prediction. For vertical or horizontal modes, the pixel values are extrapolated by upper samples [A, B, C, D] or left samples [I, J, K, L], while other directional modes utilize the linear weighted average of reference samples [A, B,…, M]. For the DC mode, if all samples A, B, C, D, I, J, K, L are available, all the predicted pixels are the same and formed by means of upper and left samples as formula (1).
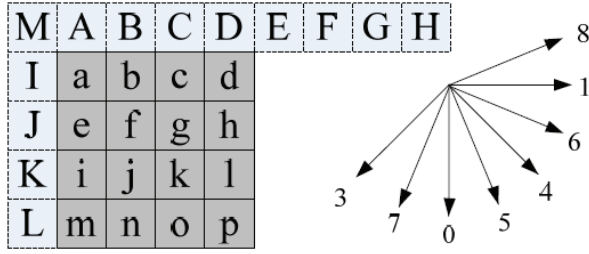
Figure 1. Intra_4x4 prediction modes

$$a = b = ... = p = (A + B + C + D + I + J + K + L + 4) >> 3. \quad (1)$$

In Intra_4x4 prediction, each block is predicted using either the DC mode or one of the eight directional modes. It is well known that these eight directional modes are used to predict the regions with unified orientations. Therefore, it is better if the remainder DC mode can be used to predict some areas where the textures have no unified orientation. However to predict such areas using DC mode is not very accurate because it uses one value to predict all pixels in the block, which can not adapt any kind of variations between the pixels. Since the Intra_4x4 DC mode cannot provide accurate prediction for some areas without unified directions, we replace DC mode with WCP mode, which is much simpler than DWP and iDWP [9].

III.   PROPOSED INTRA PREDICTION ALGORITHM

It has been shown from the statistics that the vertical and horizontal predictions are more frequently used than other modes. This is implying higher correlations between the reference samples and the pixels to be predicted in these two directions [10]. Therefore, it is possible to enhance the intra prediction accuracy by employing more such directional predictions. This inspires us to design WCP to replace DC mode in Intra_4x4 by combining vertical and horizontal predictions.

Furthermore, because only the upper and left reference samples are available, the block can be divided into three parts: the diagonal, upper right and lower left part. As the diagonal pixels have the equal distance between the upper and left reference samples so that they can be predicted by both the corresponding upper and left reference samples. However, the upper right pixels are closer to the upper reference samples while the lower left pixels are closer to the left ones. Therefore, we mainly use the upper reference samples as the major component for the upper right part prediction while using the left reference samples for the lower left part. In other words, we use the vertical prediction as the major component for the upper right part, while employing the horizontal prediction as the major component for the lower left part.

Since correlation between pixels exists in both vertical and horizontal directions; we cannot only use the vertical or horizontal prediction directly in a 4x4 block. Also, the
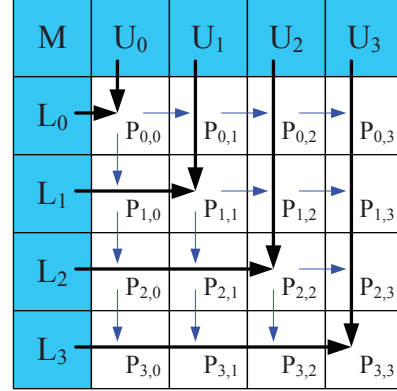

Figure 2.   Illustration of intra 4x4 prediction

correlations between the adjacent pixels are higher than the correlations between the reference samples and the pixels. As a result, when we predict the upper right part, we use the left predicted pixel to adjust the vertical prediction, while using the upper predicted pixel as an adjuster to the horizontal prediction for the lower left part, which is similar to [11].

As illustrated in Figure 2, $P_{i,j}$ ($0 <= i,j <= 3$) denotes the pixels to be predicted in the ith row and jth column of current 4x4 block, U_j and L_i denote the reference samples reconstructed from upper and left blocks respectively.

In order to make a weighted cross prediction, we assign the same coefficient 1/2 to both upper and left reference samples to predict the diagonal pixels because of their equal distances to the reference samples. On the other hand, when we predict upper right and lower down parts, a larger coefficient 3/4 is assigned to the major reference samples while a smaller coefficient 1/4 to the adjustment ones. The mathematical expression is as formula (2):

$$P_{i,j} = \begin{cases} \dfrac{1}{4}(3U_j + P_{i,j-1}) & i < j; \\ \dfrac{1}{2}(L_i + U_j) & i = j; \\ \dfrac{1}{4}(3L_i + P_{i-1,j}) & i > j. \end{cases} \quad (2)$$

Furthermore, the denominators and the multiplications can be implemented by low level operations like shift operation, such that $P_{i,j}$ can be calculated as formula (3):

$$P_{i,j} = \begin{cases} (3U_j + P_{i,j-1} + 2) >> 2 = [(U_j << 1) + U_j + P_{i,j-1} + 2] >> 2 & i < j; \\ (L_i + U_j + 1) >> 1 & i = j; (3) \\ (3L_i + P_{i-1,j} + 2) >> 2 = [(L_i << 1) + L_i + P_{i-1,j} + 2] >> 2 & i > j. \end{cases}$$

Since some of the predicted values are based on the other predicted values, we have to compute the predicted values in the following order. The diagonal pixels should be computed first, which are only based on U_j and L_i. Then, the upper right and lower left parts can be computed. The mathematical expressions of WCP are listed as formula (4)-(19). It can be seen that the longer the distance between the reference sample and the pixel is, the smaller the coefficient becomes.

$$P_{0,0} = \frac{1}{2}U_0 + \frac{1}{2}L_0; \tag{4}$$

$$P_{1,1} = \frac{1}{2}U_1 + \frac{1}{2}L_1; \tag{5}$$

$$P_{2,2} = \frac{1}{2}U_2 + \frac{1}{2}L_2; \tag{6}$$

$$P_{3,3} = \frac{1}{2}U_3 + \frac{1}{2}L_3; \tag{7}$$

$$P_{0,1} = \frac{3}{4}U_1 + \frac{1}{4}P_{0,0} = \frac{3}{4}U_1 + (\frac{1}{8}U_0 + \frac{1}{8}L_0); \tag{8}$$

$$P_{0,2} = \frac{3}{4}U_2 + \frac{1}{4}P_{0,1} = \frac{3}{4}U_2 + (\frac{3}{16}U_1 + \frac{1}{32}U_0 + \frac{1}{32}L_0); \tag{9}$$

$$P_{0,3} = \frac{3}{4}U_3 + \frac{1}{4}P_{0,2} = \frac{3}{4}U_3 + (\frac{3}{16}U_2 + \frac{3}{64}U_1 + \frac{1}{128}U_0 + \frac{1}{128}L_0); \tag{10}$$

$$P_{1,2} = \frac{3}{4}U_2 + \frac{1}{4}P_{1,1} = \frac{3}{4}U_2 + (\frac{1}{8}U_1 + \frac{1}{8}L_1); \tag{11}$$

$$P_{1,3} = \frac{3}{4}U_3 + \frac{1}{4}P_{1,2} = \frac{3}{4}U_3 + (\frac{3}{16}U_2 + \frac{1}{32}U_1 + \frac{1}{32}L_1); \tag{12}$$

$$P_{2,3} = \frac{3}{4}U_3 + \frac{1}{4}P_{2,2} = \frac{3}{4}U_3 + (\frac{1}{8}U_2 + \frac{1}{8}L_2); \tag{13}$$

$$P_{1,0} = \frac{3}{4}L_1 + \frac{1}{4}P_{0,0} = \frac{3}{4}L_1 + (\frac{1}{8}U_0 + \frac{1}{8}L_0); \tag{14}$$

$$P_{2,0} = \frac{3}{4}L_2 + \frac{1}{4}P_{1,0} = \frac{3}{4}L_2 + (\frac{3}{16}L_1 + \frac{1}{32}U_0 + \frac{1}{32}L_0); \tag{15}$$

$$P_{2,1} = \frac{3}{4}L_2 + \frac{1}{4}P_{1,1} = \frac{3}{4}L_2 + (\frac{1}{8}U_1 + \frac{1}{8}L_1); \tag{16}$$

$$P_{3,0} = \frac{3}{4}L_3 + \frac{1}{4}P_{2,0} = \frac{3}{4}L_3 + (\frac{3}{16}L_2 + \frac{3}{64}L_1 + \frac{1}{128}U_0 + \frac{1}{128}L_0); \tag{17}$$

$$P_{3,1} = \frac{3}{4}L_3 + \frac{1}{4}P_{2,1} = \frac{3}{4}L_3 + (\frac{3}{16}L_2 + \frac{1}{32}U_1 + \frac{1}{32}L_1); \tag{18}$$

$$P_{3,2} = \frac{3}{4}L_3 + \frac{1}{4}P_{2,2} = \frac{3}{4}L_3 + (\frac{1}{8}U_2 + \frac{1}{8}L_2). \tag{19}$$

The comparison of computational complexity between WCP, iDWP and DWP has been tabulated in Table I. To predict one pixel in iDWP requires 1 multiplication, 2 additions, 1 subtraction, and 2 shifts, which has much lower complexity than DWP [9]. As in the proposed technique the number of operation to predict one pixel is 2.75 additions and 1.75 shifts in average, which is less complex than iDWP.

TABLE I.   COMPARISON OF COMPUTATIONAL COMPLEXITY BETWEEN WCP, iDWP AND DWP

operation	WCP	iDWP	DWP
division	0	0	1
multiplication	0	1	2
shift	1.75	2	6
addition	2.75	2	0
subtraction	0	1	0

IV.   EXPERIMENTAL RESULTS AND DISCUSSIONS

The proposed algorithm is simulated on the JM software (version 10.2 [12]). The performance of WCP is compared with intra prediction scheme of H.264 in terms of the average

TABLE II.   SIMULATION RESULTS OF CIF SEQUENCES

Sequence (CIF)	WCP		iDWP	
	Δ PSNR (dB)	Δ bitrate (%)	Δ PSNR (dB)	Δ bitrate (%)
Bus	0.09	-0.96	0.13	-1.35
Stefan	0.08	-0.78	0.10	-0.95
Mobile	0.06	-0.48	0.05	-0.44
Foreman	0.02	-0.23	0.02	-0.30
average	0.06	-0.61	0.08	-0.76

TABLE III.   SIMULATION RESULTS OF QCIF SEQUENCES

Sequence (QCIF)	WCP		iDWP	
	Δ PSNR (dB)	Δ bitrate (%)	Δ PSNR (dB)	Δ bitrate (%)
Salesman	0.07	-0.86	0.11	-1.28
News	0.06	-0.55	0.09	-0.89
Claire	0.04	-0.53	0.05	-0.67
average	0.06	-0.65	0.08	-0.95

TABLE IV.   SIMULATION RESULTS OF BUS SEQUENCE

bus	H.264		WCP		iDWP	
QP	PSNR (dB)	Bitrate (kbits/s)	PSNR (dB)	Bitrate (kbits/s)	PSNR (dB)	Bitrate (kbits/s)
20	42.63	7790.02	42.65	7739.39	42.66	7718.59
24	39.13	5706.39	39.15	5665.60	39.16	5647.07
28	35.80	4064.10	35.82	4031.32	35.82	4016.01
32	32.56	2724.45	32.58	2701.40	32.58	2690.79

TABLE V.   SIMULATION RESULTS OF SALESMAN SEQUENCE

salesman	H.264		WCP		iDWP	
QP	PSNR (dB)	Bitrate (kbits/s)	PSNR (dB)	Bitrate (kbits/s)	PSNR (dB)	Bitrate (kbits/s)
20	43.07	1753.39	43.07	1741.58	43.07	1736.65
24	39.58	1242.61	39.59	1231.60	39.61	1228.34
28	36.32	846.24	36.34	840.45	36.33	835.98
32	33.22	542.81	33.20	538.82	33.22	538.23

PSNR (dB) and the bitrate (kbits/s) for several test sequences including both CIF and QCIF format. In the experiment, all the sequences have 100 frames (30Hz), compressed with all Intra prediction by applying four quantization parameters: QP=20, 24, 28, and 32. The test conditions are indicated as follows.

(a) Intra 16x16 and Intra 4x4 are used;

(b) RD Optimization is used;

(c) The entropy coding method is CABAC;

(d) The 8x8 transform is disabled;

(e) The adaptive rounding is disabled.

Since our proposed method is for low complexity intra prediction, we only compare the simulation results of WCP and iDWP here. The simulation results for WCP and iDWP are listed in TABLE II and TABLE III. The average PSNR
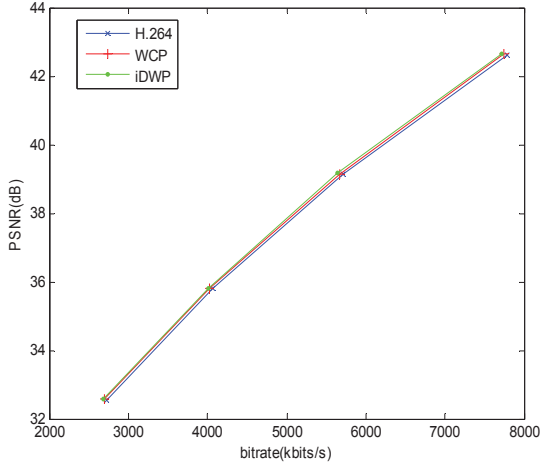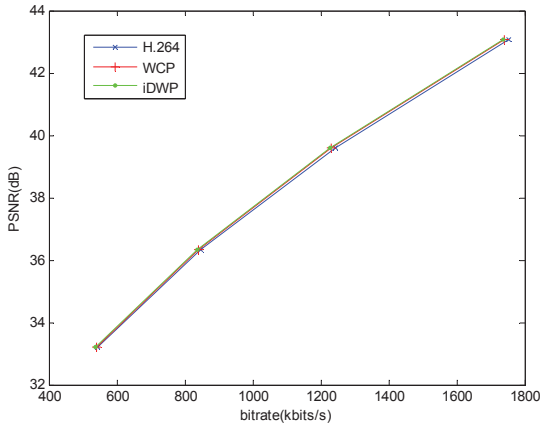
Figure 3. RD curve for "Bus" sequence

Figure 4. RD curve for "Salesman" sequence

and bitrate differences are based on the method described in [13]. The simulation results of Bus and Salesman sequences are listed in TABLE IV and TABLE V and their rate-distortion (RD) curves are shown in Figure 3 and Figure. 4. All the performance obtained by WCP is over H.264 with an average PSNR improvement of 0.06 dB which reaches its maximum of 0.09 dB in CIF sequences. In QCIF sequences, the average PSNR increase is 0.06 dB. It can be seen that the result of WCP is very competitive to iDWP. In mobile sequence, the WCP performance is even better than iDWP. There are improvements of WCP compared to H.264 for all tested sequences.

V. CONCLUSIONS

In this paper, WCP is proposed to replace DC mode in Intra_4x4 to solve the problem that DC mode used in the regions with no unified direction. To improve the prediction accuracy, we utilize the vertical prediction as the major approach on the upper right part, while the horizontal prediction as the major approach on the lower left part, which includes both vertical and horizontal predictions in one block. The experimental results show the improvements compared to the original H.264 with DC mode and WCP gives comparable RD performance to iDWP with lower computational complexity**.**

REFERENCES

[1] ITU-T Recommendation H.264 and ISO/IEC 14496-10, "Advanced video coding for generic audiovisual services," May 2003 (and subsequent amendment and corrigenda).

[2] T. Wiegand, G.J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," IEEE Trans. Circuits Syst. Video Technol., vol.13, no.7, pp.560-576, July 2003.

[3] ITU-T, "Video coding for low bit rate communication," ITU-T, Recommendation H.263 version 1, 1995.

[4] ISO/IEC 14496-2 (MPEG-4), "Coding of moving pictures and audio,", ISO/IEC, 1999.

[5] Bojun Meng, and Oscar C. Au, "Fast intra-prediction mode selection for 4x4 blocks in H.264," International Conference on Acoustics, Speech, and Signal Processing, 2003, vol. 3, 6-10 pp. III-389-392.

[6] Kim C S, Hsuan-Huei Shih, and C.C. Jay Kuo, "Feature-based intra-prediction mode decision for H.264," IEEE Proceedings of International Conference on Image Processing, Singapore, vol.2, October 2004, pp.769-772.

[7] Min-hua Zhou, "Intra prediction with simplified prediction modes," ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT 3rd Meeting, JVT-D026, Fairfax, Virginia, USA, 6-10 March, 2002.

[8] Woo-Shik Kim, Dea-Sung Cho, and Kyung-Wook Kim, "Intra prediction enhancements," ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT 5th Meeting, JVT-E051, Geneva, CH, 9-17 October, 2002.

[9] Shengsheng Yu, Yi Gao, Jiazhong Chen, and Jingli Zhou, "Distance-based weighted prediction for H.264 Intra Coding," IEEE IET International Conference on Audio, Language and Image Processing 2008, July 2008, pp.1477-1480.

[10] Nan Zhang, Baocai Yin, Dehui Kong, and Wenying Yue, "Spatial prediction based intra-coding," 2004 IEEE International Conference on Multimedia and Expo (ICME), Taipei, Taiwan, June 2004, pp.97-100.

[11] Baocai Yin, Lei Sun, Dehui Kong, and Pengfei Ji, "Adjacent piexels-based intra prediction algorithm," Journal of Beijing University of Technology, vol.32, no.11, pp.1037-1042, Nov. 2006.

[12] JM10.2, http://iphome.hhi.de/suehring/tml/.

[13] Gisle Bjontegaard, "Calculation of average PSNR differences between RD-curves," ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, VCEG 13th Meeting, VCEG-M33, Austin, Texas, USA, April 2001.