DISTANCE-BASED WEIGHTED PREDICTION FOR ADAPTIVE INTRA MODE BIT SKIP IN H.264/AVC

Lai-Man Po, Liping Wang, Kwok-Wai Cheung, Ka-Man Wong, Ka-Ho Ng, Shenyuan Li, and Chi-Wang Ting

Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong SAR, China
Email: eelmpo@cityu.edu.hk

ABSTRACT

Adaptive Intra Mode Bit Skip (AIMBS) technique using boundary pixels smoothness has been shown to achieve coding efficiency improvement for H.264/AVC's Intra_4x4 coding in relatively large QPs. However, the DC mode in the Multiple-Prediction of the AIMBS becomes much less effective. To tackle this problem and further improve the coding efficiency, distance-based weighted prediction (DWP) is proposed to replace DC mode in Multiple-Prediction for predicting blocks without directional preferences. The proposed method is named as AIMBS-DWP that can enhance the robustness of AIMBS in much larger range of QPs and achieving higher rate-distortion performance. Experimental results show that an average bitrate reduction of 3.79% with lower computational requirement can be obtained by AIMBS-DWP as compared with H.264/AVC. The improvement is especially obvious in high visual quality configurations with small QPs.

Index Terms - H.264/AVC, video coding, intra prediction, mode bit skip, weighted prediction.

1. INTRODUCTION

Joint Video Team (JVT) of ITU-T VCEG and ISO 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]. By adopting several new coding techniques, H.264/AVC achieves 50% bitrate reduction comparing with the previous standards. One of the major contributions to this high coding efficiency improvement is the directional spatial prediction for intra coding. Different from the previous standards (such as H.263 [3] and MPEG-4 [4]), intra prediction in H.264/AVC is mainly applied in the spatial domain instead of the transform domain. It uses the correlation between adjacent blocks to remove the spatial pixels redundancy. In H.264/AVC, Intra_4×4 prediction provides nine directional prediction modes for each 4×4 block. Jointly usage of these nine prediction modes achieves much higher coding efficiency than just using the DC prediction in transform domain of previous video coding standards. For each 4×4 block, however, 1 or 4 bits are required to indicate the intra mode when Baseline profile is used [5]. Therefore, $16 \sim 64$ bits are required to represent the Intra_4×4 modes in each of the macroblock (MB). To reduce this burden, an adaptive intra mode bit skip (AIMBS)

method is proposed by Kim [6], in which Intra_4x4 prediction is split into Single-Prediction or Multiple-Prediction modes depending on the smoothness of neighbor blocks boundary pixels. The mode bit is skipped in Single-Prediction, which is used in low-detailed picture areas. Meanwhile, the computational complexity is also significantly reduced due to avoiding the mode selection process from these 9 prediction modes. However, the DC mode in the Multiple-Prediction of the AIMBS becomes much less effective and the order of the 9 modes is required to rearrange with DC mode assigned as the last mode.

Recently, a distance-based weighted prediction (DWP) method is proposed in [7] for replacing the DC mode in Intra_4x4 prediction. The DWP is suitable for predicting in the detailed areas without directional preference. It is because the DWP is designed based on the theory that the correlation between current pixel and its reference samples is inversely proportional to their distances. In this paper, to tackle the ineffectiveness of the DC mode in the Multiple-Predictions of AIMBS, DWP is proposed to replace DC prediction and the proposed method is called AIMBS-DWP. In this way, we can maintain rate-distortion improvement and complexity reduction in the smooth areas by AIMBS and enhancing the prediction accuracy in the complex areas by DWP. It is also proven that the AIMBS-DWP is much more robust with performance improvement over larger range of QPs especially in low QPs configurations.

This paper is organized as follows. In section II, a brief review of the AIMBS is given for providing the detail of this technique. After that the enhanced method of AIMBS-DWP is introduced in section III. Simulation results to confirm the improvement as compared with the original H.264/AVC's intra prediction are given in section IV. Finally, conclusions are drawn in Section V.

2. ADAPTIVE INTRA MODE BIT SKIP (AIMBS)

The intra prediction in H.264/AVC is performed in a block-based manner. The current block is predicted by referring to the reconstructed pixels of neighbor blocks, which are above or to the left of it. Three types of luminance intra prediction are provided: Intra_4×4, Intra_8x8 (supported only in Fidelity Extension profiles [8], hence no consideration is given in this paper) and Intra_16x16. The Intra_4x4 prediction has 9 spatial prediction modes, which are suitable to be employed in the complex areas. While Intra_16x16
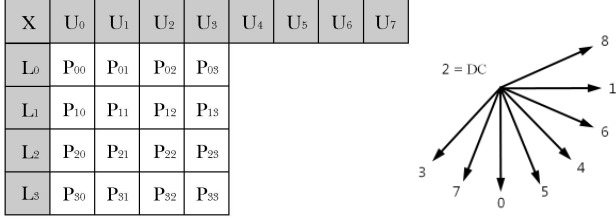
Fig. 1: Intra_4x4 prediction.

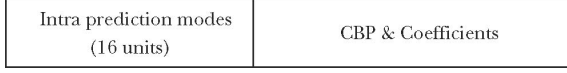| Intra prediction modes (16 units)	CBP & Coefficients

Fig. 2: Intra_4x4 prediction bitstream in H.264/AVC.

prediction has 4 modes, which is mainly used in smooth areas. The mode selection is based on the minimization of the Lagrangian cost function [9].

For Intra_4x4 prediction, the 8 directional prediction modes and DC mode are shown in Fig.1, where the 16 pixels of the 4×4 block (labeled as P_{ij}) are predicted by the upper reconstructed pixels (labeled as U_j) and left reconstructed pixels (labeled as L_i) and M. For DC mode, the predicted pixel's values are the same and equal to the average of the upper 4 pixels (U_0 to U_3) and left 4 pixels (L_0 to L_3). If the upper reconstructed pixels and the left reconstructed pixels are available, the DC predicted value is defined as

$$P_{ij} = \frac{1}{8}\left(\sum_{k=0}^{3} U_k + \sum_{k=0}^{3} L_k\right) \tag{1}$$

After the mode decision, the selected modes and residual coefficients are encoded for sending to the decoder side. The bitstream format for Intra_4x4 coding of H.264/AVC is shown in Fig. 2, in which the bits for Intra prediction modes (16 units) are in preceding of CBP & coefficients. For each intra_4x4 block, 1 or 4 bits are needed to indicate the intra mode. Then, overall 16 ~ 64 bits are required for a MB, and that is a large bitrate consumption in intra frame coding [5].

To reduce this burden of intra_4x4 mode bits, the AIMBS was proposed in [6]. The principle of AIMBS is that if the neighbor reconstructed pixels have the same or similar values, all the 9 prediction modes result in the same or similar prediction. In this case, it is unnecessary to perform all the 9 prediction modes and go through the complicated rate-distortion optimization (RDO) process to select the best one. The simplest DC mode can be used as the only prediction. Thus, the bits for indicating prediction mode can be saved. In practice, the AIMBS only uses the variance value of 8 reconstructed pixels on left and above (M and U_4 to U_7 are not considered) to classify smoothness. In this way, each 4x4 block is classified into Single-Prediction or Multiple-Prediction. In Single-Prediction only DC mode is used where intra mode bits can be skipped while 9 prediction modes are used in Multiple-Prediction that is same as Intra_4x4 of H.264/AVC. The decision of using Single-Prediction or Multiple-Prediction is based on the

relationship between the variance (σ_p) of the left and above reconstructed pixel values and a predefined threshold value Th based on quantization parameter (QP). The σ_p and Th are determined by the following two equations:

$$\sigma_p = \left\lfloor \left[\sum_{j=0}^{3}(\mu - U_j)^2 + \sum_{i=0}^{3}(\mu - L_i)^2 + 4 \right] / 8 \right\rfloor, \tag{2}$$

$$Th = \left\lfloor \frac{Qstep^2 + 8}{16} \right\rfloor, \tag{3}$$

where μ in eq.(2) is the mean value of the neighbor pixels ($U_0 \sim U_3$ and $L_0 \sim L_3$) as shown in Fig. 1.

In AIMBS, the prediction mode bits are required only for the blocks using Multiple-Prediction. Thus, the bits for intra prediction modes contain variable units. In this case, the decoder cannot distinguish the intra prediction mode bits and CBP & Coefficients bits. In order to solve this problem, the format changing of the bitstream is adopted. In addition, the DC mode become much less effective in the Multiple-Prediction as it is seldom used. To release this problem, the DC mode is assigned to be the last mode (mode number 8) in AIMBS.

3. DISTANCE-BASED WEIGHTED PREDICTION FOR AIMBS

The major problem of the DC mode in Multiple-Prediction of AIMBS is mainly due to most of the smooth 4x4 blocks are classified into Single-Prediction. In addition, the DC mode is not very effective in predicting complex block without directional preference in Multiple-Prediction. In order to tackle this problem and further improve the coding efficiency, distance-based weighted prediction (DWP) [7] is proposed to replace DC mode in Multiple-Prediction. The proposed solution can take advantages of both AIMBS and DWP and improve the robustness. Meanwhile it can compensate the shortages of them. Thus it performs well in both the smooth and complex areas.

3.1 Distance-based Weighted Prediction (DWP)
Let us first introduce the DWP that mainly focuses on improving the prediction accuracy of DC mode in complex areas of a picture. In Intra_4x4 prediction, 8 directional modes are used to predict the regions with unified directions and it is hoped that the DC mode can be used to predict the areas where the textures have no unified directions. However, DC mode cannot adapt any kind of variations. DWP is proposed to replace DC mode to predict some non-unified varying areas, which commonly occurs in ordinary videos. This mode is designed based on the empirical model that the correlation between two pixels is inversely proportional to the distance between them. Consequently, the prediction value of P_{ij} can be calculated by eq.(4) as shown below:

$$P_{ij} = (L_i \times (i+1) + U_j \times (j+1) + \Delta)/(i+j+2), \tag{4}$$

where Δ is the adjusting coefficient.

3.2 AIMBS-DWP

In AIMBS method, Multiple-Prediction is used to predict the detailed areas but the DC mode cannot adapt any kind of variations and it is rarely used in Multiple-Prediction. To overcome this shortage, the DWP mode can be used to replace the DC mode in Multiple-Prediction, which is more suitable for predicting the detailed areas. Thus, the accuracy of the Multiple-Prediction can be significantly improved. As the computational requirement of the DWP mode is higher, the conventional DC mode is remained in the Single-Prediction part of the proposed AIMBS-DWP method. In this design, the DWP mode is used only when it is necessary. This proposed combination method of AIMBS-DWP can reduce bitrate and encoding time in the low detailed areas by AIMBS and improve the prediction accuracy in detailed areas by DWP.

The flowchart of the proposed AIMBS-DWP method is shown in Fig. 3. In which Intra_4x4 prediction is performed in the following procedure:

Step 1: Determination of Single-Prediction or Multiple-Prediction
  If $\sigma_p < Th$, go to Step 2 (Single-Prediction)
  If $\sigma_p \geq Th$, go to Step 3 (Multiple-Prediction)
Step 2: Single-Prediction:
  Only perform DC mode without RDO mode selection;
  The bits for indicating modes can be skipped;
  Go to Step 4.
Step 3: Multiple-Prediction:
  Use eight directional modes and DWP mode with RDO mode selection;
  Go to Step 4.
Step 4: Calculation of the residual coefficients, transformation and quantization.
Step 5: Entropy coding:
  Skip the mode bits for the blocks using Single-Prediction;
  Encode the coefficients & CBP first and then the Intra_4x4 mode bits.

4. EXPERIMENTAL RESULTS

To verify the efficiency of the proposed AIMBS-DWP method, the simulation is implemented on H.264/AVC reference software JM 11.0 [10]. For the purpose of comparison, we used the same sequences as in [6]. Three CIF sequences (Mobile, Paris, and Foreman), four 720P sequences (City, Night, Crew and Shuttlestart) and one 1080P sequence (Rolling Tomato) were used in the experiments. All the sequences are compressed with intra prediction only and the QPs of 22, 27, 32, and 37 are used. The test conditions are indicated as follows.
  (a) Intra 16x16 and Intra 4x4 are used;
  (b) RDO Optimization is used;
  (c) The entropy coding method is CAVLC;
  (d) The 8x8 transform is disabled;
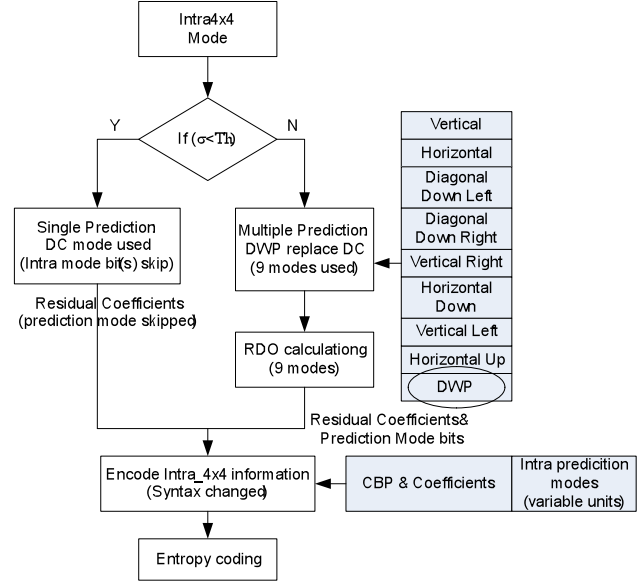 (e) The adaptive rounding is disabled.

Fig. 3: Flowchart of the proposed method AIMBS-DWP.

The performance of DWP, AIMBS and AIMBS-DWP are compared with anchor intra prediction scheme in H.264/AVC. The average PSNR gains and bitrate reductions are listed in the Tables I and II, which are calculated based on the BD-PSNR [11]. Table I shows the average bitrate reduction of the AIMBS-DWP method is 3.79%, which is a little greater than the sum of bitrate reduction of DWP (0.67%) and AIMBS (3.08%). In addition, the average PSNR gain is 0.19 dB for all sequences. The encoding time of Foreman sequence is shown in Table III. It is seen that the AIMBS-DWP maintains the time saving rate of the AIMBS method.

The AIMBS-DWP was also tested at small QPs range: QP = 20, 24, 28, and 32. From Table II, we can find that the proposed method can achieve 2.84% bitrate reduction with 0.15 dB PSNR increase. The performance of original AIMBS in small QPs is not as good as in large QPs. In this situation, the proposed method can improve the performance significantly. Moreover, we can see that AIMBS-DWP has a greater improvement in the sequences with complicated content, such as Paris, City and Night sequences. Small QPs and complicated content imply high quality coding requirement. The proposed method performs better to meet high quality requirement, which indicating that it has higher prediction accuracy.

5. CONCLUSIONS

In this paper, a novel AIMBS-DWP method is proposed to improve the coding efficiency of Intra_4x4 prediction of H.264/AVC. Experimental results show that the performance of AIMBS-DWP is better than sum of the improvements of AIMBS and DWP. It can achieve 3.79% and 2.84% average bitrate reductions as compared with H.264/AVC in medium and small QPs respectively. It can also improve the robustness of AIMBS in much larger range of QPs especially in high visual quality applications.

Table I. Simulation results with QP={22, 27, 32, 37}.

Sequence	Resolution	Number of frames	DWP		AIMBS		AIMBS-DWP	
			ΔBitrate(%)	ΔPSNR(dB)	ΔBitrate(%)	ΔPSNR(dB)	ΔBitrate(%)	ΔPSNR(dB)
Mobile	CIF	300	-0.47	0.05	-0.49	0.05	-1.01	0.11
Paris		150	-0.83	0.08	-1.01	0.09	-1.87	0.17
Foreman		300	-0.47	0.03	-2.95	0.18	-3.53	0.22
Average coding efficiency for CIF			**-0.59**	**0.05**	**-1.48**	**0.11**	**-2.14**	**0.17**
City	720P	150	-1.38	0.10	-1.68	0.11	-3.08	0.21
Night		150	-1.43	0.10	-1.91	0.13	-3.28	0.23
Crew		150	-0.27	0.01	-4.41	0.18	-4.65	0.19
Shuttlestart		150	-0.69	0.03	-5.58	0.22	-6.33	0.25
Rolling Tomato	1080P	60	0.13	-0.00	-6.62	0.11	-6.57	0.10
Average coding efficiency for HD			**-0.73**	**0.05**	**-4.04**	**0.15**	**-4.78**	**0.20**
Average			**-0.67**	**0.05**	**-3.08**	**0.13**	**-3.79**	**0.19**

Table II. Simulation results with QP={20, 24, 28, 32}.

Sequence	Resolution	Number of frames	DWP		AIMBS		AIMBS-DWP	
			ΔBitrate(%)	ΔPSNR(dB)	ΔBitrate(%)	ΔPSNR(dB)	ΔBitrate(%)	ΔPSNR(dB)
Mobile	CIF	300	-0.49	0.06	-0.25	0.03	-0.75	0.10
Paris		150	-0.82	0.08	-0.55	0.06	-1.35	0.14
Foreman		300	-0.66	0.05	-1.66	0.11	-2.29	0.16
Average coding efficiency for CIF			**-0.66**	**0.06**	**-0.82**	**0.07**	**-1.46**	**0.13**
City	720P	150	-1.64	0.13	-0.74	0.06	-2.30	0.18
Night		150	-1.55	0.12	-1.09	0.08	-2.54	0.19
Crew		150	-0.34	0.01	-2.99	0.12	-3.47	0.14
Shuttlestart		150	-0.86	0.04	-3.98	0.16	-4.90	0.20
Rolling Tomato	1080P	60	0.18	-0.00	-5.16	0.06	-5.13	0.06
Average coding efficiency for HD			**-0.84**	**0.06**	**-2.79**	**0.10**	**-3.67**	**0.15**
Average			**-0.77**	**0.06**	**-2.05**	**0.08**	**-2.84**	**0.15**

Table III. Encoding time of Foreman sequence.

Foreman(CIF)	H.264	DWP	AIMBS	AIMBS-DWP
QP	Time(s)	Time(s)	Time(s)	Time(s)
22	429.818	429.634	379.096	380.202
27	371.297	370.816	279.042	279.244
32	328.338	327.638	173.022	173.164
37	299.763	300.155	108.425	108.352

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] Iain E. G. Richardson, *H.264 and MPEG-4 Video Compression: Video Coding for Next Generation Multimedia*, Wiley, England, 2003.

[6] D.Y. Kim, K.H. Han, and Y.L. Lee, "Adaptive intra mode bit skip in intra coding," IEEE Asia Pacific Conference on Circuits and Systems 2008, pp. 446-449, Nov. 2008.

[7] S.S Yu, Y Gao, J.Z. Chen, and J.L. Zhou, "Distance-based weighted prediction for H.264 intra coding," IEEE IET International Conference on Audio, Language and Image Processing 2008, pp.1477-1480, July 2008.

[8] G.J. Sullivan, P. Topiwala, and A. Luthra, "The H.264/AVC advanced video coding standard: Overview and introduction to the fidelity range extensions," presented at the *SPIE Conference on Applications of Digital Image Processing XXVII*, vol. 5558, pp. 53-74, Special Session on Advances in the New Emerging Standard: H.264/AVC, Aug. 2004.

[9] G.J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," IEEE Signal Process. Mag., vol. 15, no. 6, pp. 74-90, Nov. 1998.

[10] H.264 Joint Video Team (JVT) Reference Software version 11.0, http://iphome.hhi.de/suehring/tml/.

[11] G. 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.