Fast and efficient intraprediction method for H.264/AVC

Hui Yuan, Yilin Chang, Zhaoyang Lu, and Ming Li
Xidian University, State Key Laboratory of ISN, Xi'an, 710071 China
E-mail: hyuan@mail.xidian.edu.cn

Abstract. A new intraprediction method is proposed to decrease coding complexity as well as improve coding efficiency of H.264/AVC. Whether the predicted values of different prediction modes are similar is determined according to the proposed rules. When the predicted values are similar, the average of such values is used to predict the current block without invoking the rate-distortion optimization process. Experimental results show that by integrating the proposed method in the H.264/AVC codec, 39.25% coding time can be saved; meanwhile, the average bit-rate reduction is 2.62% under the same peak signal-to-noise ratio of reconstructed videos. © *2010 Society of Photo-Optical Instrumentation Engineers.*
[DOI: 10.1117/1.3377968]

1 Introduction

H.264/AVC[1] can give higher coding efficiency than any other video coding standards because it provides many complex modes, both in intra and interframe coding. Rate distortion optimization (RDO)[2] is used to choose the best mode when coding a block. In this work, intraframe coding of H.264/AVC is investigated. Motivated by the observation that different prediction methods can give similar predicted values, a new intraprediction method is proposed to decrease coding complexity, as well as enhance coding efficiency of H.264/AVC.

2 Proposed Intraprediction Mode

In H.264/AVC, a picture is divided into many macroblocks (MB). Each MB can be partitioned into blocks with different sizes, such as 4×4, 8×8, or 16×16. For each block, different intraprediction methods can all be modeled as a mathematical expression, as shown in Eq. (1),

$$t_y = \sum_x (\omega_{\gamma,x} \times s_x) \Big/ \sum_x \omega_{\gamma,x}, \qquad (1)$$

where t and s represent the predicted and prediction pixels, t_y and s_x represent the y'th predicted pixel and the x'th prediction pixel, respectively, $\omega_{\gamma,x}$ is the corresponding weighted coefficient of the x'th prediction pixel under the γ'th prediction mode, and ω_γ represents a set of coefficients of the γ'th prediction mode.

In Eq. (1), when s_x are the same, t_y are the same, and even if s_x are different, t_y may still be similar, because the coefficients ω_γ are also different. Therefore, different prediction modes may give similar predicted values. That means the residual signals derived by different prediction modes are semblable, and the distortions are also semblable. Accordingly, the RDO process can be saved, and there is no need to use so many bits to label different prediction modes.

To determine whether different predicted blocks are semblable, features of different predicted blocks must be extracted, and for easy implementation, means and variances are used. First, the predicted blocks of all the prediction modes are obtained. The mean and variance of each predicted block are then derived by Eqs. (2) and (3),

$$\text{avg}_i = \frac{1}{W \times H} \sum_{m=1}^{W} \sum_{n=1}^{H} \text{pred}_i(m,n), \qquad (2)$$

$$\text{var}_i = \frac{1}{W \times H} \sum_{m=1}^{W} \sum_{n=1}^{H} [\text{pred}_i(m,n) - \text{avg}_i]^2, \qquad (3)$$

where W and H are the width and height of the block, respectively, $\text{pred}_i(m,n)$ is the predicted pixel of the i'th prediction mode at position (m,n) of the predicted block, and avg_i and var_i are the mean and variance of all the predicted pixels in the predicted block under the i'th prediction mode. Furthermore, Eqs. (4) and (5) are then used to compute the variance of avg_i and var_i, respectively,

$$\text{VAR}_{\text{avg}} = \frac{1}{M} \sum_{i=0}^{M-1} \left(\text{avg}_i - \frac{1}{M} \sum_{j=0}^{M-1} \text{avg}_j \right)^2, \qquad (4)$$

$$\text{VAR}_{\text{var}} = \frac{1}{M} \sum_{i=0}^{M-1} \left(\text{var}_i - \frac{1}{M} \sum_{j=0}^{M-1} \text{var}_j \right)^2, \qquad (5)$$

where M is the number of prediction modes of the block, and VAR_{avg} and VAR_{var} represent the variance of avg_i and var_i, respectively. Different prediction modes can give similar predicted values when VAR_{avg} and VAR_{var} are less than some threshold TH. The block, which has similar predicted values under different prediction modes, is named the decoder side prediction method derivable (DSPMD) block for the decoder, and can derive the prediction method without labeling information. For a DSPMD block, any of the predicted values can be chosen as the final value. However,
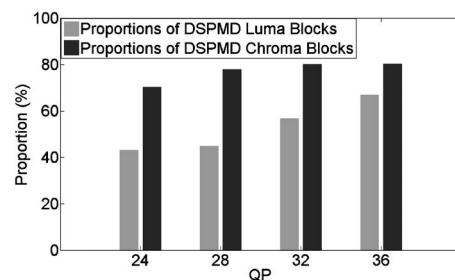

Fig. 1 Proportions of DSPMD blocks in intraframes for the Foreman sequence.
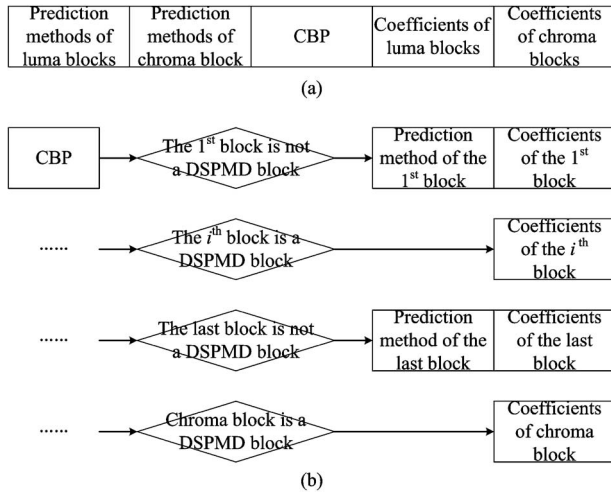
Fig. 2 Bit stream comparison between (a) H.264/AVC and (b) the proposed method.

one cannot guarantee which prediction mode is best. Therefore, to get a better predicted value, the average of different predicted values is used as the final predicted value of the block, as shown in Eq. (6),

$$\text{PRED}_{\text{final}}(m,n) = \frac{1}{M} \sum_{i=0}^{M-1} \text{pred}_i(m,n), \qquad (6)$$

where $\text{PRED}_{\text{final}}(m,n)$ is the average of different predicted values at position (m,n) of the current block.

For a certain picture, the number of DSPMD blocks increases with the increment of *TH*, and $\text{PRED}_{\text{final}}$ of the DSPMD blocks will be distant from their optimal predicted values, which is determined by the RDO process when *TH* increases. But when the quantization step (Q_{step}) is large, *TH* should be increased. That is because even if the predicted values are a little farther from each other, the quantized signals (derived from discrete cosine transform on residual signals and the succeeding quantization) are still similar. To adapt to Q_{step}, *TH* is set to be $Q_{\text{step}}/2$ empirically.

Figure 1 shows the proportions of DSPMD blocks of the Foreman sequence. The proportion of luminance and chrominance DSPMD blocks can reach as high as 60 and 80%. Therefore, the RDO process and prediction mode information can be greatly saved by detecting DSPMD blocks.

In a bit stream of H.264/AVC, mode information of all the blocks of an MB are followed by a coded block pattern (CBP) and coefficient bits, as shown in Fig. 2(a), which means that the mode information of all the blocks must be decoded first without considering whether a block is a DSPMD block. To ensure that DSPMD blocks can be decoded, a bit stream of an MB is changed to Fig. 2(b), where CBP of the MB is decoded first, and then whether a block is a DSPMD block is determined. For a DSPMD block, $\text{PRED}_{\text{final}}$, shown in Eq. (6), is used as the prediction result, and the following bits are decided as coefficients.

3 Experimental Results

In the experiments, H.264/AVC reference software Joint Model (JM) 15.1 was used to encode ten sequences with various representative contents. Each sequence is encoded at four different quantization parameters (QP) i.e., 24, 28, 32, and 36. The corresponding Q_{step} are 10, 16, 26, and 40. The coding complexity and efficiency are evaluated by the percentage of average encoding time savings, and percentage of average bit rate savings under the same PSNR of reconstructed videos (BD bit rate),[3] respectively. Statistic results of all the sequences are shown in Table 1. From Table 1, by integrating the proposed method into JM15.1, the 39.25% coding time can be saved and the BD bit rate is −2.62% on average.

Figure 3 compares bits of each component between the proposed method and JM15.1. It can be concluded that the gains of coding efficiency are embodied in bit reductions of both mode information and coefficients, while keeping the same coding quality. The bit reductions of the coefficients benefit from the combination of all the predicted values, as shown in Eq. (6). To show the results clearly, RD curves for the Bigship sequence are presented in Fig. 4. It can be concluded that the proposed method outperforms JM15.1.

Moreover, the proposed method was compared with the estimation based method in Ref. 4 and the pixel-based direction detection (PDD) method in Ref. 5. Compared with the method in Ref. 4, the BD bit rate of the proposed method is −1.05%; meanwhile 73.47% coding time is

Table 1 Coding efficiency of different test sequences.

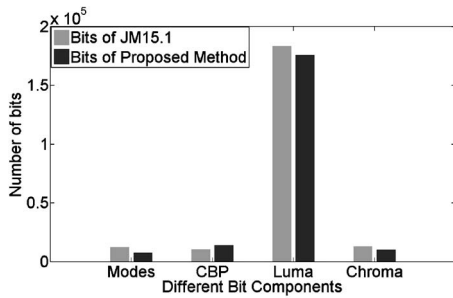| Sequences	BD-Bit Rate (%)	ΔTime (%)	Sequences	BD-Bit Rate (%)	ΔTime (%)
Carphone (cif)	−2.86	−36.82	Bigship (720p)	−5.35	−59.62
Claire (cif)	−1.87	−27.41	City (720p)	−1.93	−30.59
Container (cif)	−1.02	−20.76	Crew (720p)	−2.68	−44.56
Foreman (cif)	−3.72	−45.01	Night (720p)	−1.85	−30.35
PeopleOnStreet (1080p)	−2.45	−41.54	Traffic (1080p)	−2.51	−55.83
Average of all the sequences				**−2.62**	**−39.25**

Fig. 3 Different bit components in the stream of the first frame of Bigship.

saved. Compared with the PDD method, 23.11% average coding time is increased. However, the BD bit rate of the proposed method can achieve −5.33%. Although the coding time of the proposed method is larger than that of the PDD method, coding efficiency of the proposed method is better than that of the PDD method.
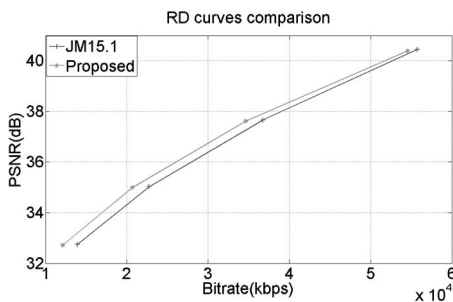

Fig. 4 RD curves for Bigship sequence.

4 Conclusions

A fast and efficient intraprediction method is proposed. Whether a block is a DSPMD block is determined, and then all the predicted values of a DSPMD block are averaged to get the final predicted value. The RDO process can be neglected for those DSPMD blocks, and the identifiers of prediction methods can also be omitted as well. Experimental results show that the coding complexity is decreased, and that coding efficiency is improved as well.

References

1. Telecommunication Standardization Sector of ITU, "Series H: audio visual and multimedia systems Infrastructure of audiovisual services– coding of moving video," Recommendation ITU-T H.264 (Mar. 2009).
2. T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, "Rate-constrained coder control and comparison of video coding standards," *IEEE Trans. Circuits Syst. Video Technol.* **13**(7), 688–703 (Jul. 2003).
3. G. Bjøntegaard, "Improvements of the BD-PSNR model," ITU-T SG16 Q.6 Document, VCEG-AI11, Berlin (Jul. 2008).
4. C. S. Park and S. J. Ko, "Estimation-based intra prediction algorithm for H.264/AVC," *Opt. Eng.* **48**(3), 030506–1–3 (2009).
5. A. C. Tsai, J. F. Wang, J. F. Yang, and W. G. Lin, "Effective subblock-based and pixel-based fast direction detections for H.264 intra prediction," *IEEE Trans. Circuits Syst. Video Technol.* **18**(7), 975–982 (Jul. 2008).