

High Quality and Fast H.264/AVC Video Encoder : Enhanced Intra 4x4 Mode Decision for H.264/AVC coders

Author : Jar-Ferr Yang

Co-authors : Hung-Ming Wang, Chao-Hsuing Tseng

Institute of Computer and Communication Engineering

IEEE Transaction on Circuit and Systems for Video Technology, vol. 15, no. 8, August 2006

The 21st century is in the high-tech information epoch of multimedia communications. Along with progresses of high efficient video and audio compression technologies and high speed and low cost personal computer and VLSI chip, The International Standard Organization (ISO) finalized many multimedia compression standards such that people can store MP3 music in 10 times, JPEG image in 20 times, and MPEG-2 movies in 80 times of efficiency. Currently, the Joint Video Team (JVT) formed by the ISO Motion Picture Expert Group (MPEG) and International Telecommunication Union (ITU) suggested a new video coding standard, H.264/AVC, whose compression efficient is about 3 times of MPEG-2. Due to high quality and high efficiency, H.264/AVC will be effectively and widely adopted by many multimedia transmission and storage systems such that the cost of the popular multimedia services could be further reduced in many applications.

As shown in Fig. 1, H.264/AVC utilizes many computational intensive compression algorithms to increase its compression efficiency. Hence, the H.264/AVC acquires more complicated coding structure with higher computational complexity than the traditional video coding standards. With such complications, the H.264/AVC becomes more difficult to achieve the real-time compression and decompression than the others. For example, for 4x4 intra prediction, the intra mode prediction in the H.264/AVC is conducted in 4x4 block bases. As shown in Fig. 2, the target block is with pixels {b11~b44} while the coded pixels {A, B, ..., L, X} on the top and left of the block are used for predictions. In the H.264/AVC, there are 9 different prediction modes. Fig. 3 shows 8 different direction predictions, while Mode 2 denotes DC Mode, which is non-directional. In order to achieve high compression, H.264/AVC reference software suggests Rate-Distortion (RD) optimization cost function as

$$\phi_1 = D + \lambda_1(Q_p)R \tag{1}$$

where D is the sum of square differences between the original and the reconstructed blocks, R is the data rate, which is used to encode the block, and $\lambda_1(Q_p)$ is Lagrange

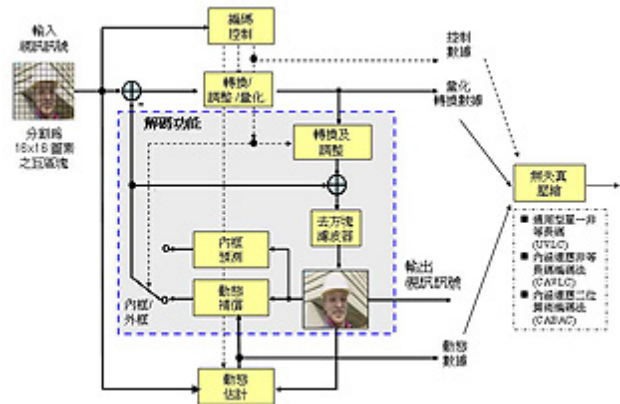


Figure 1. Block diagram of H.264/AVC video encoder

M	A	B	C	D	E	F	G	H
I	b ₁₁	b ₁₂	b ₁₃	b ₁₄				

multiplier, which is dependent on the quantization parameter, Q_p . By using (1), we can achieve the best compression quality; however, it acquires a huge computation to attain the true optimization. For all 9 prediction modes, we need to perform 9 times encoding and decoding processes such that we can compute the true bit rate and true distortion. After 9 time true coding processes, we then substitute distortions and rates into (1), the optimal intra prediction will be the mode, which has the minimum cost function, ϕ_1 .

I	b_{11}	b_{12}	b_{13}	b_{14}
J	b_{21}	b_{22}	b_{23}	b_{24}
K	b_{31}	b_{32}	b_{33}	b_{34}
L	b_{41}	b_{42}	b_{43}	b_{44}

Figure 2. Diagram of 4x4 intra prediction, where target pixels $\{b_{11}\sim b_{44}\}$ of predicting 4x4 block and predictive coded pixels, $\{A\sim M\}$

In order to achieve fast decision of prediction mode, H.264/AVC reference software also suggests two simple cost functions as

$$\phi_2 = SAD + 4P\lambda_2(Q_p) \tag{2}$$

$$\phi_3 = SATD + 4P\lambda_3(Q_p) \tag{3}$$

where SAD denotes the sum of absolute differences, SATD represents the sum of absolute Hadamard transformed differences, P is determined if the detected mode is the most probable mode or not. If it is the most probable mode, $P = 1$; otherwise $P = 0$. In general, the cost function defined in (2) or (3), which does not require any encoding and decoding process, acquires much less computation than that depicted in (1) but with less compression efficiency. Usually, the cost function defined in (3) performs better than that addressed in (2). With the above discussions, we need a huge computation to achieve the better compression of 4x4 block. Hence, how to reduce the computational complexity at the same time to maintain the compression efficient becomes the main objective of this research.

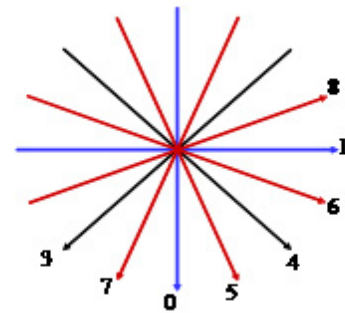
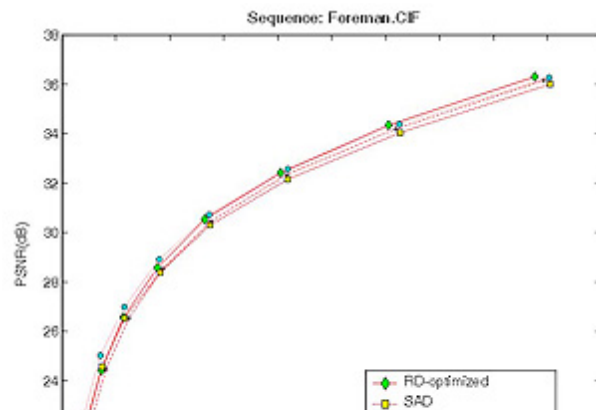


Figure 3. Intra prediction modes and their corresponding prediction directions

In this research, to enhance the coding efficiency, we propose a new cost function as:

$$\phi_4 = SAITD + (4T_c - T_o + 4P) \cdot \lambda_2(Q_p). \tag{4}$$

In ϕ_4 , we not only modify the distortion part but also improve the rate estimation. In (4), the distortion is modified as the sum of absolute integer-transformed differences (SAITD) and the rate is further estimated by T_c , T_o , where T_c and T_o denote the number of non-zero coefficients and the number of absolute-one coefficients after quantization of 4x4 transformed coefficients. T_c and T_o are the important parameter used in the



context adaptive variable length codes (CAVLC). Since we did not perform any real encoding process and the bit rate is predicted by T_c and T_o , the computation of the proposed cost function will not increase too much. However, it could greatly improve the coding performance. After several experimental simulations, Fig. 4 shows the coding performances in PSNR- Bit Rate plots for all cost functions. The results exhibit that the proposed cost function achieves very good coding performance, even for low bit rate video, the proposed method is better than the RD optimization, which acquire a lot of computation.

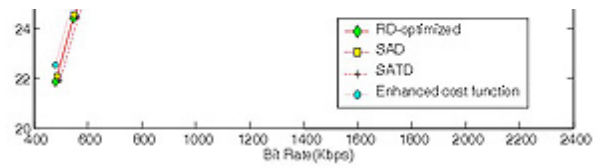


Figure 4. Comparison of coding performance in PSNR-data rate Curves

For further reduction of the computation, we also suggest a fast algorithm to reduce the computation of the SAITD. In the H.264/AVC, we all know that we need to compute 9 cost functions to determine the prediction mode. The total computation for the SAITD, we need to compute the block difference between the original block B, i.e., {b11~b44} and the prediction value block A and the transformation of the block difference. Among all functions, the integer transform demands the most computation. In H.264/AVC, the 4x4 integer transform acquires 64 additions and 16 binary shifts. In intra prediction, for 9 modes, we should spend 576 additions and 144 shifts in integer transforms. So, we further suggestion a fast computation to reduce its computation. In H.264/AVC, the integer transform is a linear transform, which satisfies

$$T(A - B) = T(A) - T(B) \tag{5}$$

where $T()$ denotes the linear transform and A and B are matrices. In other words, the transformation of two subtracted matrices is equal to the subtraction of two transformed matrices. Since the prediction matrix A in most cases are with fixed structure with many symmetrical properties. The computation of transformation for each prediction matrix is enlisted in Table 1. Table 2 shows the computation acquired by the original and the proposed method. From Table 1, we know that the total 9 integer transforms only need 320 additions and 103 binary shifts. Table 2 depicts the total computation for integer transforms. The original method suggested in (3) totally requires 576 additions and 144 shifts, however, the proposed method only needs 384 additions and 118 shifts. If we treated complexity of the addition is the same as that of the shift, we found that the proposed method only needs about 69.8% complexity of the original method. The proposed method not only increases the coding efficiency but also reduces the computational complexity of the existing algorithms.

Table 1 Computational complexity for transformation of predicted values in intra prediction modes

模式 \ 轉換	整數轉換	
模式 0	加法: 8	位移: 6
模式 1	加法: 8	位移: 6
模式 2	加法: 0	位移: 1
模式 3	加法: 34	位移: 13
模式 4	加法: 38	位移: 13
模式 5	加法: 58	位移: 16
模式 6	加法: 58	位移: 16
模式 7	加法: 58	位移: 16
模式 8	加法: 58	位移: 16
總和	加法: 320	位移: 103

In summary, this research proposed a new cost function to achieve the optimal intra prediction for H.264/AVC encoders. We use the SAITD to replace the SATD and further use T_c the number of non-zero coefficient and T_o the number of absolute-one

Table 2. Computation complexity for intra prediction of the original and the proposed methods

原始方法		提出方法	
加法	位移	加法	位移

coefficients to achieve better distortion and rate estimations. Simulations show that the proposed method can improve the coding performance. Furthermore, we also propose a fast realization of the proposed method to reduce the computation. The proposed technology can improve the coding performance and reduce the computation time for H.264 intra prediction.

加法	位移	加法	位移
576	144	384	119

Copyright 2009 National Cheng Kung University