Computational Time Reduction Using Low Complexity Skip Prediction for H.264/Avc Standard

ZORAN MILICEVIC,
Signals and IT Department,
Raska 2, Belgrade
REPUBLIC OF SERBIA

ZORAN BOJKOVIC
Faculty of Transport and Traffic Engineering,
Univeristy of Belgrade
Vojvode Stepe 305, Belgrade
REPUBLIC OF SERBIA

Abstract: - In an attempt to reduce the computational time savings using H.264/AVC encoder, this paper seeks to provide that 23-35% can be achieved. The constraints are the following: Fidelity Range Amendment (FRExt) profile with correspondence tools, the skip prediction only for B pictures and rate-distortion optimization mode selection disabled. On the other hand, it is shown that there is not significant loss of rate-distortion performance.

Key-Words: - Fidelity Range Extensions Amendment, Low complexity skip prediction, Rate-distortion optimization, Lagrange multiplier

1 Introduction

The ITU-T Video Coding Experts Group (VCEG) and ISO│IEC 14496-10 AVC Moving Picture Experts Group (MPEG) formed the Joint Video Team (JVT) in 2001 to develop the new video coding standard H.264/Advanced Video Coding (AVC) [12, 10, 4, 11].

H.264 video coding standard has the same basic functional elements as previous standards (MPEG-1, MPEG-2, MPEG-4 part 2, H.261, H.263) [16, 15, 2], i.e., transform for reduction of spatial correlation, quantization for bit rate control, motion compensated prediction for reduction of temporal correlation, entropy encoding for reduction of statistical correlation. However, in order to fulfill better coding performance, the important changes in H.264 occur in the details of each functional element by including intra-picture prediction, a new 4x4 integer transform, multiple reference pictures, variable block sizes and a quarter pel precision for motion compensation, a deblocking filter [7], and improved entropy coding. Improved coding efficiency comes at the expense of added complexity to the coder/decoder [8]. H.264 utilizes some methods to reduce the implementation complexity [19]. Multiplier-free integer transform is introduced. Multiplication operation for the exact transform is combined with the multiplication of quantization [6]. Compared with MPEG-4, H.263 and MPEG-2, H.264/AVC can achieve 39%, 49% and 64% of bit-rate reduction, respectively [13]. The improvement in coding performance comes mainly from the prediction part. Intra prediction significantly improves the coding performance of H.264/AVC intra frame coder. On the other side, inter prediction is enhanced by motion estimation with quarter-pixel accuracy [5], variable block sizes, multiple reference frames and improved spatial/temporal direct mode. Unlike the previous standards, prediction must be always performed before texture coding not only for inter macroblocks but also for intra macroblocks [9].

In H.264/AVC standard, an advanced method called Context Adaptive Binary Arithmetic Coding (CABAC) [3], is included for entropy coding of syntax element value. The use of CABAC has been adopted into the standard as a way of gaining additional performance relative to Context – Based Adaptive Variable Length Coding (CAVLC) coding [1], at the cost of additional complexity. Encoding with CABAC consist of three stages-binarization, context modelling and adaptive binary arithmetic coding [17].

In 2004 JVT added new extensions in the observed standard. These extensions, known as the Fidelity

Range Extensions (FRExt) [18], provide a number of enhanced capabilities relative to the base specification. Specifically, these include: supporting an adaptive block-size for the residual spatial frequency transform, supporting encoder-specified perceptual-based quantization scaling matrices, and supporting efficient lossless representation of specific regions in video content. The FRExt project produced a suite of four new profiles collectively called the High profiles.

This paper is organized as follows. Section two explains low complexity skip prediction, the skip prediction model and skip prediction algorithm. In the next section we discussed about two different test models (original and our modification) to explain experimental results and analyze some of important performances of coder for video coding standard H.264/AVC. Section four concludes the paper.

2 Low complexity skip prediction

The modification software version with algorithm applied reduces computational processing through early identification of macroblocks to be skipped. Prior to coding each macroblock, the encoder estimates the rate-distortion cost of coding or skipping the macroblock. Based on these estimates, the macroblock is either coded as normal (i.e. the encoder processes the macroblock and selects an appropriate coding mode) or skipped (i.e. no further processing is carried out). Skip prediction model aims to reduce computation whilst maintaining or improving rate-distortion performance.

In order to generate the problem we have implemented one know skip prediction algorithm [14] for complex calculations. Comparing to the previous results in the open literature, our research includes:

a) Fidelity Range Extension Profile with correspondence tools,
b) The skip algorithm only for B pictures,
c) Rate distortion optimization mode selection disabled.

2.1 Skip prediction model and algorithm

Let M_i be the coding mode (one of K possible modes) chosen by the encoder for macroblock X_i and let $M_i=K$ denote the "skip" mode. Define the rate-distortion costs of coding or skipping a macroblock as [14]:

$$J(X_i, M_i) = D(X_i, M_i) + \lambda R(X_i, M_i) \quad (1)$$

and

$$J(X_i, K) = D(X_i, K) \quad (2)$$

respectively, where λ is a weighting parameter (Lagrange multiplier). Note that the rate associated with a skipped macroblock is effectively zero. Macroblock X_i should be skipped (not coded) if:

$$D(X_i, M_i) + \lambda R(X_i, M_i) \geq D(X_i, K) \quad (3)$$

The skip prediction algorithm proceeds as follows [14]:

- For every macroblock, calculate $D^n(X_i, K)$ and read previously stored values $D^{(n-1)}\left(X_i^{(n-1)}, M_i^{(n-1)}\right)$ and $R^{(n-1)}\left(X_i^{(n-1)}, M_i^{(n-1)}\right)$
- Calculate F_i:
$$F_i = D^{(n-1)}\left(X_i^{(n-1)}, M_i^{(n-1)}\right) \cdot R^{(n-1)}\left(X_i^{(n-1)}, M_i^{(n-1)}\right) \quad (4)$$
- Calculate $\hat{\lambda}$ using equation:
$$\hat{\lambda} = \left(7.374 \cdot 10^{-8} k_i + 5.239 \cdot 10^{-5}\right) \exp\left[\left(-3.688 \cdot 10^{-5} k_i + 0.3203\right) \cdot QP\right] \quad (5)$$
- Choose "skip" mode if the following expression is true:
$$D^{(n-1)}\left(X_i^{(n-1)}, M_i^{(n-1)}\right) + 0.5\hat{\lambda} R^{(n-1)}\left(X_i^{(n-1)}, M_i^{(n-1)}\right) \geq D^{(n)}(X_i, K) \quad (6)$$
- If "skip" mode is chosen, no further processing is carried out and the macroblock is marked as "skipped". If "code" mode is chosen, process the macroblock as normal.

3 Results and discussion

Experimental work is carried out using test software JVT version JM 10.2 implemented by Fidelity Range Extensions (FRExt) amendment [20] and modification version JM 10.2M.

Our idea is to perform tests and compare different test versions in order to show which improvements are obtained by JM 10.2M.

In all of our experiments five test sequences, i. e., Foreman (pour in detail), Bus and Tempete (more reach) and Stefan and Mobile (the most reach) are used with CIF format (352x288) and 4:2:0 High profile.

For comparison and to analyze the output results of our all experiments well-known the key factors which must be compared are: signal - to – noise ratio (SNR) for luma (Y) picture component, the bit rate (kbps) and the computational time (ms).

We measured SNR only for Y because human visual system is better sensitive on luma then on chroma components of pictures.

We took 25 pictures in one test sequence as enough condition taking into account that human visual system (HVS) is very sensitive to this rate.

Starting from the fact that video coding standard H.264/AVC holds two methods of entropy coding, i. e., Context Adaptive Binary Arithmetic Coding (CABAC) and Context Adaptive Variable Length Coding (CAVLC), we have used CABAC because this method gives better results than CAVLC in the sense of video context coding [11].

Also, we have applied Hadamard transformation because it improves the encoder performance comparing to other transformations [11].

Finally, if we compare the original and the decoded pictures (Forman, Bus, Stefan, Tempete, Mobile). It can be seen that the picture distortion (SNR for Y component) has minimal values.

We taking into account three relevant parameters: the computational time, the SNR and the bit rate in order to compared FRExt and modified encoder.

A number of CIF test sequences were encoded using the "FRExt" encoder. The same sequences were encoded using a modified version of the FRExt encoder, incorporating skip prediction algorithm. Table 1 shows experimental results when both test software version are used.

The bit rate has got the value 21,5 % as it is shown in Table 1.

- Stefan test sequence is shown in Table 1 reached the value up to 30% in computational time and 0,48% in SNR. The bit rate is increased up to 19,3%.
- Tempete test sequence has computational time 29,5% and SNR 0,29%. The bit rate has the value over 13% as it is shown in Table 1.
- Mobile test sequence shows the best results (35% in reduction of computational time) and only 0,295 % in SNR. The bit rate has the smallest increase of 18 % as it is shown in Table 1.

From the Table I it should be noted that our gain in computational time for Bus, Stefan, Mobile and Tempete sequences is higher than the bit rate obtained of the encoder output. The only exception is for the sequence Forman.

On the other hand, if we focus only on B pictures in all sequences values of bit rate are increased in depend on test sequences, too. However, SNR (Y) has a small differences between the original and the modification version. Also, computational time is reduced.

Table 1. Comparison simulation results

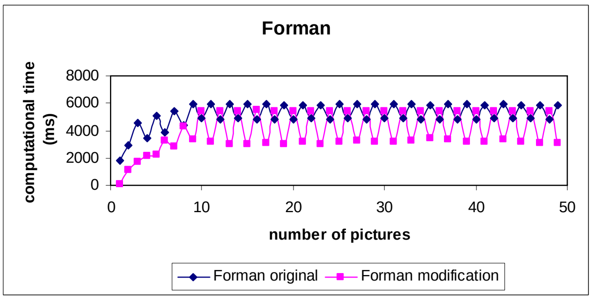
Comparison software version JM 10.2 and JM 10.2M									
Sequence	JM 10.2	JM 10.2M		JM 10.2	JM 10.2M		JM 10.2	JM 10.2M	
	time (s)	time (s)	Total (%)	SNR (Y)	SNR (Y)	Total (%)	Bit rate (kbps)	Bit rate (kbps)	Total (%)
Forman	**252,667**	**192,779**	**-23,702**	**36,91**	**36,49**	**+1,137**	**297,79**	**389,55**	**+30,813**
Bus	**278,546**	**197,421**	**-29,124**	**34,55**	**34,28**	**+0,781**	**998,41**	**1213,60**	**+21,553**
Stefan	**275,094**	**190,890**	**-30,609**	**35,25**	**35,08**	**+0,482**	**856,85**	**1022,63**	**+19,347**
Mobile	**293.576**	**190.544**	**-35,095**	**33,82**	**33,72**	**+0,295**	**1126,86**	**1331,82**	**+18,188**
Templete	**275,411**	**193,938**	**-29,582**	**34,36**	**34,26**	**+0,291**	**835,28**	**944,89**	**+13,122**

It can be seen that there exists a reduction in computational time (ms) in the range from 23 to 35 % depending on the test sequence. Also, there is slightly decrease in SNR up to 1%. However, there is an increment in the bit rate from 13% to 30%.

Comparing different test versions for every sequence the following facts arise:

- There is reduction in computation time about 23% for Forman test sequence, until signal-to-noise ration decreases for 1%. Bit rate is increased up to 30% and it is the worst result in comparison with other sequences as it is shown in Table 1.
- Computational time in Bus test sequence has got the value of 29%, while SNR is a little better.

Figure 1 shows computational time reduction (curve with square dots marks original, while curve with cycle dots marks modification computational time in relations with number of pictures) for all test sequences.



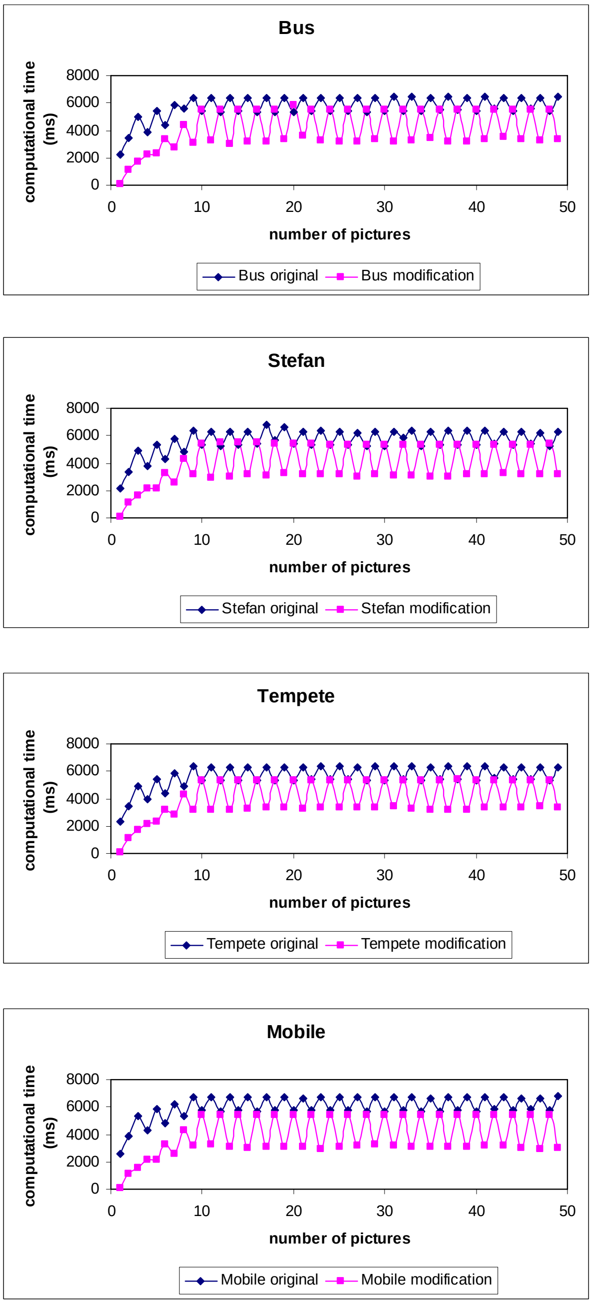


Figure 1. Computational time reduction for all test sequences.

4 Conclusion

This paper compares the two test models (original and modification) for skip prediction in H.264 encoder. The rate-distortion costs of coding or skipping a macroblock are estimated prior to processing. A decision whether to code the macroblock or stop further processing is made based on a Lagrangian cost function. Test results indicate that coding time is reduced by 23-35 % through early identification of macroblocks that are likely to be skipped during the coding process. There is not significant loss of rate-distortion performance. Coding time is substantially reduced because a significant number of macroblocks are not processed by the encoder. This reduction is particularly high for high-activity sequences. It is lower for low-activity sequences such as “Foreman”. The computational saving depends on the activity of the video sequence. It is particularly significant for sequences with medium and high activity.

Acknowledgement

It is a pleasure to acknowledgement the help received from our colleague Tomislav Sajdl associated with company Raking, Belgrade. He is help was in the form of computer programs.

References:

[1]G. Bjontegaard and K. Lillevold, Context-adaptive VLC coding of coefficients, JVT document C028, Fairfax, May 2002.

[2]T. Stockhammer, M. Hannuksela, and S. Wenger, H.26L/JVT coding network abstraction layer and IP based transport, *IEEE ICIP 2002*, Rochester, New York, Vol. 2, pp. 485-488, Sept. 2002.

[3]D. Marple, H. Schwrts, and T. Wiegand, Context-based adaptive binary arithmetic coding in the H.264/AVC video compression standard, *IEEE Transactions on Circuits and System for Video Technology*, vol.13, no.7, pp.620-636, July 2003.

[4] T. Weigand, G. Sullivan, G. Bjontegaard and A. Luthra, Overview of the H.264/AVC Video Coding Standard, *IEEE Transactions on Circuits and System for Video Technology*, vol.13, no.7, pp.560-576, 2003.

[5]T. Wedi, H. G. Musmann "Motion- and Aliasing-Compensated Prediction for Hybrid Video Coding", *IEEE Transactions on Circuits and System for Video Technology*, vol.13, no.7, pp. 577-586 , 2003.

[6] H. Mavar, A. Hallapuro, M. Karczewicz, and L. Kerofsky: "Low-Complexity Transform and Quantization in H.264/AVC", *IEEE Transactions on Circuits and System for Video Technology*, vol.13, no.7, pp. 598-603, 2003.

[7] P. List, A. Joch, J. Lainema, G. Bjontegaard, M. Karczewicz, "Adaptive Deblocking Filter", *IEEE Transactions on Circuits and System for Video Technology*, vol.13, no.7, pp. 614-619, 2003.

[8]J. Ribas-Corbera, P. A. Chou, and S.

Regunathan: "Generalized Hypothetical Reference Decoder for H.264/AVC", *IEEE Transactions on Circuits and System for Video Technology*, vol.13, no.7, pp. 674-687, 2003.

[9] M. Flierl and B. Girod, "Generalized B Pictures and the draft JVT/H.264 Video Compression Standard", *IEEE Transactions on Circuits and System for Video Technology*, vol.13, no.7, pp. 587-597 , 2003.

[10]M Flierl and B. Girod, Video coding with supermodes motion-compensated signals. Applications to H.264 and Beyond, *vol. 70 of Series in Engineering and Computer Science,* Kluwer Accademic Publishers, Boston, Mass, USA, 2004.

[11] A. Puri, H. Chen, A. Luthra, Video Coding using the H.264/MPEG-4 AVC compression standard, *Signal Processing: Image Communication 19*, pp.793-846, 2004.

[12] D. Milovanovic, Z. Bojkovic, Advanced video coding standard H.264/AVC Performance vs. complexity, *WSEAS Transactions on Communications,* vol. 4, pp. 1-6, January 2005.

[13]Y. W. Huang, B. Y. Hsieh, T. Ch. Chen and L. G. Clan, Analysis fast algorithm, and VLSI architecture design for H.264/AVC intra frame coding, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 3, pp. 378-401, March 2005.

[14]C. S. Kannangara, I. E. G. Richardson, M. Bystrom, J. R. Solera, Y. Zhao, A. MacLennan, and R. Cooney, Low-complexity skip prediction for H.264 through lagrangian cost estimation, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 2, pp. 202-208, February 2006.

[15]K.R.Rao, J.J.Hwang, *Techniques and Standards for images, video and audio coding*, Prentice Hall PTR, 1996.

[16]K.R.Rao, Z.S.Bojkovic, D.A.Milovanovic, *Multimedia communication systems: techniques, standards and networks*, Prentice Hall PTR, 2002.

[17] Iain E. G. Richardson, *H.264 and MPEG-4 video compression video coding for next-generation multimedia*, John Viley & Sons Ltd., The Atrium, Southern Gate, Chichester, West Sussex PO 19 8SQ, England, 2003.

[18]G. Sullivan, P. Topiwala and A. Luthra, Overview of the H.264/AVC advanced video coding standard: Owerview and introduction to the fidelity range extensions, *SPIE Conference on Applications of Digital Image Processing*, pn. 5558-53, August, 2004.

[19]Z. M. Milicevic, Z. S. Bojkovic, RD optimization and skip prediction for H.264/AVC standard, *Conference EUROCON'05*, Belgrade, November 2005.

[20]http://bs.hhi.de/~suehring/tml/download/JM10.2