

Skip Prediction and Early Termination for Fast Mode Decision in H.264/AVC

Yury V. Ivanov, C. J. Bleakley

*School of Comp. Science and Informatics, University College Dublin, Belfield, Dublin 4, Ireland
e-mail: {yury.ivanov, chris.bleakley}@ucd.ie*

Abstract

This paper proposes a fast mode decision algorithm for H.264/AVC based on skip prediction and early termination techniques. The skip decision is based on a partially computed Sum of Absolute Differences metric combined with utilization of the Lagrangian Rate-Distortion cost function from the previous frame. A statistical analysis of the spatio-temporal characteristics of the Rate Distortion cost function and SAD metric for skip decisions is provided. Experimental results show that the new algorithm outperforms existing ones in performance providing a 55% reduction in total computational complexity, a slight reduction in bit rate and negligible impact on visual quality.

1. Introduction

The H.264 standard [1] developed by the Joint Video Team (JVT) provides better coding efficiency than MPEG-2 and H.263 at low bit rates [2]. Mode Decision (MD) comprises a significant portion of encoder complexity as the standard allows for seven block sizes, or modes, each having a separate motion vector [3].

The conventional Mode Decision technique can be significantly improved by so-called ‘early termination’. This assumes that some block modes can be eliminated from the mode search by predicting that they will never be chosen in the final decision.

Early termination techniques can be further improved by performing an early skip decision. In the standard MD algorithm the skip mode is checked last. If a positive skip decision is made at the start of the MD process then mode search can be eliminated entirely.

In order not to increase the total complexity of MD the additional skip check must be of low complexity.

Incorrect skip decisions cause poor macroblock (MB) segmentation and lead to bit rate increase or visual quality degradation. Thus, the key is to use an efficient and accurate metric for skip prediction.

The novel algorithm presented in this paper is a combination of an early termination technique, which uses the Rate-Distortion cost function as a metric, plus skip prediction which utilizes partially computed Sum of Absolute Differences (SAD8x8). The work is based on the results of statistical analysis which show a high correlation between the RD cost function of the macroblocks and SAD8x8.

The paper is organized as follows. Section 2 gives a brief overview of MD in H.264. Section 3 reviews related work in the field. Section 4 provides a statistical analysis of the RD cost function and of mode transitions. The skip prediction algorithm is described in Section 5 and experimental results are presented in Section 6. Section 7 concludes the paper.

2. Overview of Mode Decision and Skip

H.264 Mode Decision utilizes the Lagrangian Rate-Distortion cost function [4] when computing mode costs for available MB segmentation options:

$$\min \{J(s, c, MODE | QP, \lambda_{MODE})\} \quad (1)$$
$$J(s, c, MODE | QP, \lambda_{MODE}) = SSD(s, c, MODE | QP) + \lambda_{MODE} \cdot R(s, c, MODE | QP)$$

where QP is the macroblock quantization parameter and λ_{MODE} is the Lagrange multiplier. $MODE$ is the mode chosen from the set of potential modes:

$$MODE \in \{INTRA4 \times 4, INTRA16 \times 16, SKIP, 16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4\} \quad (2)$$

$R(s, c, MODE | QP)$ is the transmitted bit rate associated with $MODE$ and QP and SSD is the Sum of the Squared Differences. The MD algorithm [5] implemented in the reference JM encoder [6]

sequentially calculates J for all segmentation options. After that, the algorithm selects the mode with minimum J cost.

In order to make a skip decision, after computing all the modes, JM performs the following additional checks:

- Best mode selected is INTER16x16
- Reference frame is the previous one
- MV found is equal to predicted MV
- Transform coefficients are all quantized to zero

If all these conditions are met then skip is selected.

3. Related Work

In this section a brief overview of several fast mode decision algorithms for H.264 is provided [7-11],[14] and [15]. The algorithms are variants of the early termination approach.

The MD algorithm described in [7] uses Variable Block Size (VBS) prediction from the surrounding MBs. The method suffers from the disadvantage that modes are predicted only from frame border MBs.

The method in [8] is based on the correlation of motion vectors across the various MB partitions. For the reported average complexity reduction of 39% the bit rate gain is quite significant – 3% on average.

The algorithm in [9] uses a selective Intra decision based on average boundary error metric combined with an early skip decision. Intra prediction based on local edge information with early termination of the Rate Distortion Optimization (RDO) calculation is proposed in [10]. These algorithms report about 20-35% of total complexity reduction on “IPPP” GOP sequences, mainly due to a reduction in Intra coding complexity.

MB mode correlation and classification of MBs according to their motion-energy ratio for MB mode prediction and partial RDO decision is used in [11]. The algorithm operates in one hybrid RDO mode in contrast to the other methods that can operate in two modes (i.e. RDO on/off).

Authors of [14] propose to use All-Zero Coefficient Block (AZCB) metric for early termination. In [15] early prediction is made by estimating a Lagrangian rate-distortion cost function using an adaptive model for the Lagrange multiplier based on local sequence statistics. Method achieves computational saving of 19%–67%, depending on the source sequence.

In these algorithms different metrics for skip and early termination are used. Less computationally

complex metrics give greater complexity reduction but lead to greater errors in mode decisions. Use of J from the previous frame as a metric has the advantage that it is calculated as part of the MD process and does not require an additional processing step. Analysis in the next section shows that partially computed SAD8x8 and J from the previous frame are highly correlated.

4. Statistical Analysis

4.1. Spatio-temporal Correlation of J

In the first series of experiments we investigated if previously calculated values of the Rate-Distortion cost function can be used as a prediction metric for the early termination procedure.

Pearson correlation coefficients were calculated between J values obtained using full search for blocks and their neighbors in the same and the previous frame according to:

$$r_j = \frac{\sum_{i=1}^N (X_i - M_X)(Y_i - M_Y)}{(N-1) \cdot S_X \cdot S_Y} \quad (3)$$

where N is the number of 16x16 blocks in the video sequence, X_i is the J value for the current block and Y_i is the J value for the neighboring block. For each value of j , the positions of the neighboring blocks are fixed relative to the current block (e.g. previous frame, one MB up). M_X , S_X , M_Y , and S_Y are the mean and standard deviation of X_i and Y_i respectively.

In the experiments, video sequences with different content and motion intensity were encoded using the standard JM encoder (see Section 6). Results for a typical video sequence are provided in Figure 1.

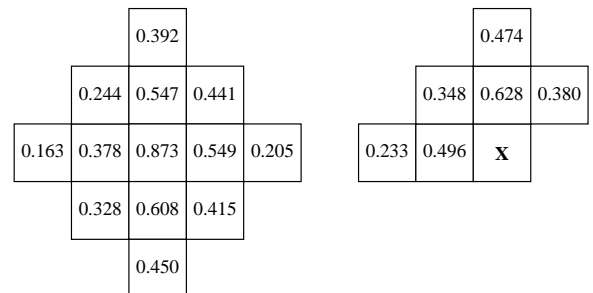


Figure 1. Correlation coefficients r_j for the current frame (right) and the previous frame (left) for Foreman, QCIF.

It can be clearly seen from this example that the RD cost function J is highly correlated between neighboring blocks. Correlation values between J are much higher than the same coefficients calculated between motion vectors. According to [12], correlation values for the motion vectors are only around 0.15–0.4. These results indicate that J is more suitable for VBS prediction than a motion intensity metric.

4.2. Correlation of J and SAD in skip decisions

The goal of the second set of experiments was to determine if SAD and J for skipped macroblocks are also correlated. If they are then SAD can be used to assess the current macroblock for skip prediction instead of J .

SAD for each skipped macroblock of the sequence was calculated and plotted against the corresponding J value as shown in Figure 2. It can be seen that the SAD and J metrics for skipped MBs are highly correlated. The correlation coefficient is 0.985.

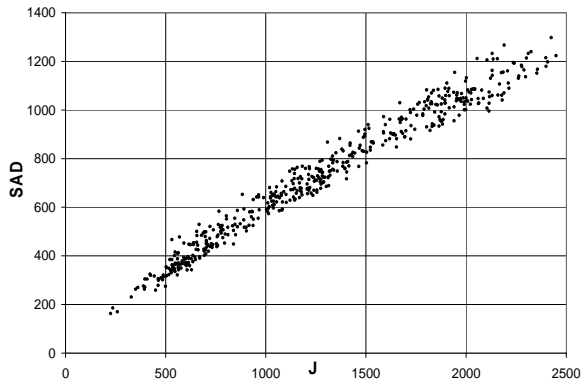


Figure 2. Correlation between SAD and J for skipped MBs. Akiyo, Hall and Carphone QCIF video sequences.

However, when SAD is calculated for every MB prior mode decision, it leads to complexity increase if the skip rate is not high enough. Since SAD and J highly correlated it is advantageous to use a similar metric with lower complexity, such as partially computed SAD8x8.

The correlation between J and partially computed SAD8x8 was investigated. Computation of partial SAD is defined as follows:

$$SAD_{8x8} = \sum_{x=0, y=0}^{7,7} |(s_i[2x, 2y] - c_i[2x, 2y])| \quad (4)$$

where $c[x,y]$ and $s[x,y]$ represent the reconstructed and original MB pixel values, respectively. The results of the analysis are plotted in the Figure 3. The correlation coefficient in this case is only $r=0.626$.

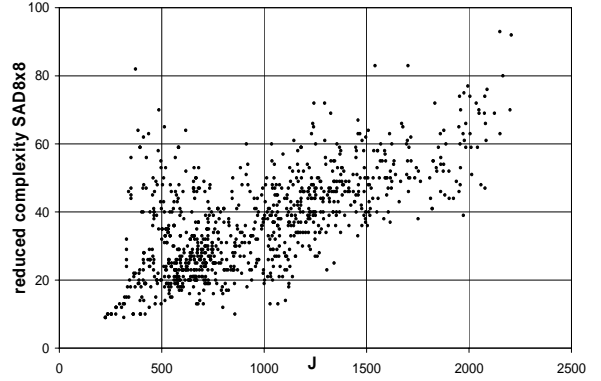


Figure 3. Correlation between partially computed SAD8x8 and J for skipped MBs. Akiyo, Hall and Carphone QCIF video sequences.

It can be concluded that both metrics have a correlation with J . However, using only reduced SAD8x8 will lead to a higher error causing quality degradation. It is proposed to use reduced SAD8x8 for predicting skip decisions, but for the skip case calculate full SAD and determine J_{SKIP} using a linear approximation based on the full SAD.

5. Skip Prediction Algorithm

The proposed algorithm consists of two parts. The first part is SKIP prediction which involves calculating SAD for each MB. Partially computed SAD8x8 is calculated first. This SAD8x8 score is then converted to a J score and a SKIP decision is made based on comparison with the mean J score for the previous frame. If the decision is SKIP then full SAD is calculated. Full SAD is needed in order to provide a more accurate J value for prediction in the next frame. In order to reduce execution time when calculating SAD8x8, the SAD with summation truncation scheme [13] is used. Thus, the calculation of SAD8x8 is terminated if the sum is already greater than mean J .

The second part, which is executed only if the SKIP decision fails, is a mode decision algorithm with early termination. The algorithm utilizes J values from the previous frame in order to omit unnecessarily MD computations.

The full algorithm is summarized in the following pseudo-code:

```

for (all MBs of the frame) {
    J[all modes] = maximum;
// SKIP prediction part
    Calculate SAD8x8 for MB;
    if (JconvertedSAD8x8 <= mean_J) {
        Calculate rest of SAD;
        J[SKIP] = converted SAD;
        best_j = J[SKIP];
        best_mode = SKIP;
    }
    else // MD with early termination
    { end_search = false;
      for (all modes or until end_search){
        if (mode transition is enabled) {
          encode MB with current mode;
          calculate J[mode];
          if (inter & J[mode] <= mean_J)
            end_search = true;
          if (intra & J[mode] <= mean_Ji)
            end_search = true;
        }
        if (end_search) {
            best_J = J[mode];
            best_mode = mode;
        }
      }
      if (not end_search) {
        best_J = minimum J[mode];
        best_mode = mode with best_J;
      }
    }
// finish MB encoding
    transmit MB using best_mode;
    prev_J[MB] = best_J;
    prev_mode[MB] = best_mode;
// finish frame encoding
    mean_J = mean of prev_J for all MBs;
    mean_J_intra = mean of prev_J for
        intra MBs only;
    move to the next frame encoding;
}

```

Mode search is in the order Skip, Inter16x16, Inter16x8, Inter8x16, Inter8x8, Intra4x4, Intra6x16. The choice of predictors was driven by the maximum correlation (see Section 4) and a mean threshold was deemed adequate based on analysis of the results.

6. Experimental Results

In the experiments, the performance of the proposed method was compared with the reference JM9.5 encoder [6]. Several QCIF video sequences of 300 frames were encoded with an “IPPP” GOP structure at 30fps on a 3GHz Pentium 4 PC. Reference encoding used all seven VBS modes, search range size

of 8, CABAC entropy coder and RDO off. In the experiments, *QP* settings ranged from 28 to 32.

The results are provided in Table 1. Bit rate increase $\Delta Bits$, quality degradation $\Delta PSNR$ and execution time change Δt are calculated according to Eqs. (5)–(7), where the minus sign (–) means improvement for the new method. In Eq. (7), total encoding time is used.

The total encoding time for the new method is reduced by almost 55% on average. Bit rate is even lower than for the original JM by 0.38%. Quality drop is acceptable (only 0.11dB). Visual examination of the decoded video sequences did not reveal any anomalies or blocking artifacts.

$$\Delta Bits = \frac{Bits_{method} - Bits_{JM}}{Bits_{JM}} \cdot 100\% \quad (5)$$

$$\Delta PSNR = PSNR_{JM} - PSNR_{method} \quad (6)$$

$$\Delta t = \frac{t_{method} - t_{JM}}{t_{JM}} \cdot 100\% \quad (7)$$

Table1. Performance comparison of the proposed fast MD versus reference JM.

Video sequence	QP	$\Delta Bits$, %	$\Delta PSNR$, dB	Δt , %
Hall, QCIF	28	-7.18	0.14	-57.42
	30	-5.67	0.09	-55.83
	32	-2.88	0.03	-51.44
Claire, QCIF	28	-1.16	0.17	-54.81
	30	3.54	0.07	-54.88
	32	-0.38	0.06	-52.11
Container, QCIF	28	2.89	0.19	-50.68
	30	3.23	0.06	-52.63
	32	2.47	0	-52.27
Akiyo, QCIF	28	0.12	0.12	-57.17
	30	0.22	0.22	-56.07
	32	0.2	0.2	-55.76
Mean		-0.38	0.11	-54.25

The complexity of the fast MD method is roughly equivalent to performing MD using, on average, one VBS mode instead of seven.

7. Conclusions

A skip prediction algorithm for fast mode decision based on early termination techniques was proposed. Using partially computed SAD8x8 and Lagrangian J as prediction metrics, the algorithm can significantly

reduce the computational complexity of H.264 for Intra- and Inter- MB encoding. Visual quality is maintained and the bit rate is even lower than for the conventional algorithm.

Correlation between RD cost function values for neighboring blocks and SAD versus J for skips have also been investigated in a series of experiments.

Future work includes study of other RD-based mode predictors.

Acknowledgements

This research project is sponsored by University College Dublin, Ireland.

8. References

- [1] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H264|ISO/IEC 14496-10 AVC)", document JVT-G050d35.doc, 7th Meeting: Pattaya, Thailand, March, 2003.
- [2] Implementation Studies Group of ISO/IEC, "Main Results of the AVC Complexity Analysis", *ISO/IEC JTC1/SC29/WG11*, Klagenfurt, July 2002.
- [3] 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, 2003.
- [4] H. Everett III, "Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources", *Operations Research*, vol. 11, pp. 399–417, 1963.
- [5] ITU-T Q.15/SG16, "H.26L Test Model Long Term Number5 (TML-5) draft0", document Q15-K-59, Sept. 2000.
- [6] JVT reference software JM 9.5, on the Web: <http://iphome.hhi.de/suehring/tml>.
- [7] G. L. Li, M. J. Chen, H. J. Li and C. T. Hsu, "Efficient Motion Search and Mode Prediction Algorithms for Motion Estimation in H.264/AVC", *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS'05)*, vol. 6, pp. 5481 – 5484, 23-26 May 2005.
- [8] T. Y. Kuo and C.H. Chan, "Fast Macroblock Partition Prediction for H.264/AVC", *Proc. IEEE Int. Conf. Multimedia and Expo (ICME'04)*, vol. 1, pp. 675-678, 27-30 June 2004.
- [9] J. Lee and B. Jeon, "Fast Mode Decision for H.264", *Proc. IEEE Int. Conf. Multimedia and Expo (ICME'04)*, vol. 2, pp. 1131-1134, June 2004.
- [10] F. Pan, X. Lin, S. Rahardja K.P. Lim, Z.G. Li, D. Wu and S. Wu, "Fast Intra Mode Decision Algorithm for H.264/AVC Video Coding", *Proc. IEEE Int. Conf. Image Processing (ICIP'04)*, vol. 2, pp. 781-784, Oct. 2004.
- [11] Z. Hong, W. Cheng-Ke, W. Yang-Li and F. Yong, "Fast Mode Decision for H.264/AVC based on Macroblock Correlation", *Proc. IEEE 19th Int. Conf. Advanced Information Networking and Applications (AINA'05)*, vol. 1, pp. 775-780, 28-30 March 2005.
- [12] F. Kossentini, Y.W. Lee, M.J.T. Smith and R.K. Ward, "Predictive RD Optimized Motion Estimation for Very Low Bit-Rate Video Coding", *IEEE Journal on Selected Areas in Comm.*, vol.15, no. 9, pp. 1752-1763, Dec.1997.
- [13] P. Kuhn, *Algorithms, Complexity Analysis and VLSI Architecture for MPEG-4 Motion Estimation*, Kluwer Academic Publishers, Boston, 1999.
- [14] Y. Kim, Y. Choe, Y. Choi, "Fast Mode Decision Algorithm using AZCB Prediction", *Int. Conf. on Consumer Electronics (ICCE'06) Digest of Technical Papers*, pp.33-34, 2006.
- [15] C. S. Kannangara, I. E.G. Richardson, M. Bystrom, J. R. Solera, Yafan Zhao, A. MacLennan and R. Cooney, "Low-Complexity Skip Prediction for H.264 Through Lagrangian Cost Estimation", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 2, 202-208, 2006.