



Title	Pareto-Optimal Macroblock Classification for Fast Mode Decision in H.264
Authors(s)	Ivanov, Yuri, Bleakley, Chris J.
Publication date	2007-11-27
Publication information	Ivanov, Yuri, and Chris J. Bleakley. "Pareto-Optimal Macroblock Classification for Fast Mode Decision in H.264." IEEE, 2007.
Conference details	2007 IEEE International Conference on Signal Processing and Communications (ICSPC), Dubai, United Arab Emirates, 24 - 27 November 2007
Publisher	IEEE
Item record/more information	http://hdl.handle.net/10197/7119
Publisher's statement	(c) 2007 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/ republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works
Publisher's version (DOI)	10.1109/ICSPC.2007.4728600

Downloaded 2024-08-10 04:48:33

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)



© Some rights reserved. For more information

PARETO-OPTIMAL MACROBLOCK CLASSIFICATION FOR FAST MODE DECISION IN H.264

Yuri V. Ivanov, C. J. Bleakley

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

ABSTRACT

This paper presents a novel fast mode decision algorithm for H.264/AVC based on a Pareto-optimal macroblock classification strategy. Previously published H.264 low complexity schemes mostly concentrated on improving class decision metrics, but did not justify the choice of MD classes. Herein, we use Pareto analysis to derive the optimal set of MD classes and to define efficient class decision metrics. For each MD class only rate-distortion optimal complexity settings are used. Experimental results show that the proposed algorithm outperforms previously published algorithms, providing a 57–73% reduction in total computational complexity with some reduction in bit rate and acceptable visual quality.

Index Terms— H.264, Mode Decision, Pareto.

1. INTRODUCTION

The H.264 standard [1] provides better coding efficiency than MPEG-2 and H.263 at low bit rates. Mode Decision (MD) comprises a significant portion of encoder complexity. The standard has a tree-structured partitioning scheme and allows for various macroblock (MB) partition sizes, or modes, each having a separate motion vector.

The conventional MD technique can be significantly improved by so-called ‘early termination’ (ET) and ‘forward SKIP prediction’. These techniques assume that some block modes can be eliminated from the mode search without loss and that correct SKIP decisions may be made at the start of the MD process. The key to the success of these techniques is utilization of fast and efficient decision metrics.

Further MD complexity reduction involves more sophisticated approaches whereby all MBs in the frame are classified according to certain features of the video signal. Different MD search parameters are used for each class. Researchers have mostly concentrated on improving the class decision metrics.

This paper, proposes a novel fast MD algorithm where MB classification is based on the Pareto curve. Thus for each MD class only rate-distortion optimal

complexity settings are used. It is demonstrated that utilization of J_{prev} , SAD_{8x8} and Frame Difference (FD) as a class decision metrics results in an efficient fast MD algorithm.

The paper is organized as follows. Section 2 reviews related work in the field. Section 3 describes the proposed MB classification and class decision metrics. The fast MD algorithm is detailed in Section 4 and experimental results are presented in Section 5. Finally, Section 6 concludes the paper.

2. RELATED WORK

Several approaches have been taken by researchers to speed up the MD process. Techniques such as ‘early termination’ and ‘forward SKIP prediction’ (e.g. [2][3][4][5]) are effective because of their simplicity, particularly for video sequences where the SKIP rate is high.

Dividing MD searches into different classes is more effective. In [6], the authors propose mode group (MGs) classification that utilizes overlapped MGs based on a measure of the residual error with predefined empirical thresholds. The method provides 52% complexity reduction, but only for P frames with Rate-Distortion Optimization (RDO) on. A feature-based approach with a risk minimizing mode decision was in [7]. It is less effective – only reducing encoding time by 20-30%. In [8] simple MB classification is proposed based on the determination of active and inactive regions of the frame. Results showing a 40% complexity reduction indicate that more sophisticated classification is required.

3. ANALYSIS

An MD algorithm based on MB classification should meet the following criteria in order to be effective:

- 1) MD classes should be optimal;
- 2) Class decision metrics should be optimal.

Current approaches, described in Section 2, focus on finding a statistically correct metric for the MD class decision. However, none (except perhaps [6]) discusses in detail or justifies their choice of MD classes. MB modes are grouped empirically or based on correlation with a decision metric. In our opinion, MB classification

is the most important issue in fast MD algorithms and should be performed based on the optimal rate, distortion and complexity settings of the encoder.

3.1. MD classes based on the Pareto curve

In [9], the computational complexity of the H.264 encoder was analyzed with different combinations of encoding tools and parameters. The coding efficiency of these versions of the coder was assessed using a single metric:

$$W = \Delta R + \mu \Delta D \quad (1)$$

where ΔR is bit rate change as a percentage, ΔD is quality loss in dB and μ is a constant relating bit rate loss and distortion increase ($\mu=13$). The choice of μ is discussed in [9].

The results were then projected onto a Pareto curve (Fig. 1). It was clearly demonstrated that not all combinations of encoding tools and parameters are optimal in the rate-distortion versus complexity sense. Optimal complexity scaling is achieved only for coder configurations on the convex hull, i.e. the Pareto curve.

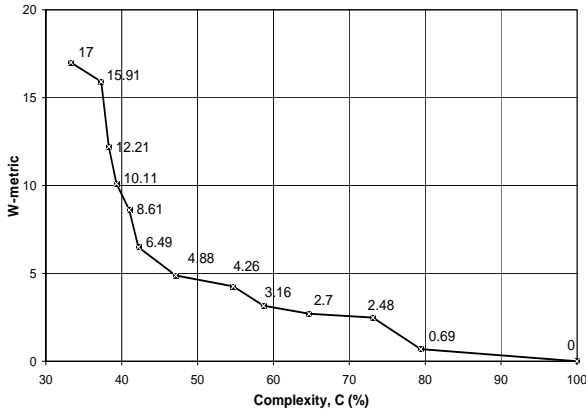


Figure 1. Pareto-optimal H.264 complexity scaling curve.

Herein, all MD classes must be on the convex hull of the Pareto curve. For the effective scaling of computational complexity a subset of the points on the convex hull must be selected. MD classes must be evenly distributed and span the entire curve. Thus, we selected five MD classes as shown in Fig.2.

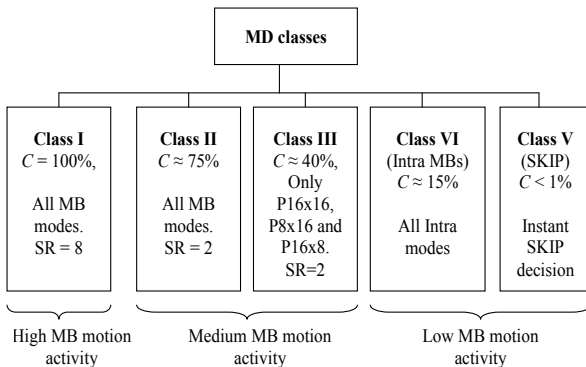


Figure 2. Proposed MB classification.

3.2. Class decision metrics

Class decision metrics are required in order to allocate MBs to MD classes. Metrics must provide an optimal trade-off between the following requirements:

- 1) High accuracy;
- 2) Low computational complexity.

Among the various decision metrics we selected ones that are based on the visual properties of MBs, since they provide the highest correlation with MB distortion [2][8].

To determine the MBs with high motion activity, the Frame Difference (FD) metric is used [8]:

$$FD = \sum_{x=0}^{15} \sum_{y=0}^{15} \begin{cases} 1 & |C(x,y) - P(x,y)| \geq T_diff \\ 0 & |C(x,y) - P(x,y)| \leq T_diff \end{cases} \quad (2)$$

where $C(x,y)$ is a pixel at position (x,y) for the MB in the current frame, $P(x,y)$ is for the previous frame and T_diff is an activity threshold.

According to the value of FD two types of MB can be defined: active and inactive. Class I is allocated to any active MB. Note, that in the case of multi-reference frames, FD metric is only computed for the first previous frame.

For inactive macroblocks the Lagrangian J from the same MB in the previous frame (or J_{prev}) can be utilized. In [2], it was demonstrated that J_{prev} is highly correlated between successive blocks and that the correlation is much higher than for motion vectors [2]. J_{prev} is also a good indicator of how efficiently the MB was coded. Herein, inactive MBs with poor coding efficiency, i.e. $J_{prev} > T_1$, are allocated to class II.

Partially computed SAD_{8x8} is used to separate MD classes further. The high correlation between SAD_{8x8} and J ($r = 0.626$, [2]) allows SKIP and Intra prediction based on calculation of SAD_{8x8} and conversion to a predicted J value for the current MB:

$$J_{SKIP_predicted} = f_1(SAD_{8x8}) = \alpha \cdot SAD_{8x8} + C_1 \quad (3)$$

If a MB is expected to have good coding efficiency, i.e. $J_{SKIP_predicted} < J_{mean}$, and is an inactive MB, then it is allocated to class V (SKIP). J_{mean} is calculated based on J_{prev} . For P- and B- frames separate equations are used to calculate $J_{SKIP_predicted}$:

$$J_{SKIP_P} = f_2(SAD_{full}) = \beta \cdot SAD_{full} + C_2 \quad (4)$$

$$J_{SKIP_B} = \chi \cdot f_3(SAD_{full}) \quad (5)$$

The coefficients α , β , γ , C_1 and C_2 are derived from correlation experiments using a linear approximation. Note, that Eqs. (3)–(5) work only for $J \leq T_1$.

Inactive MB with $SAD_{8x8} > T_2$ and $J_{SKIP_predicted} \geq J_{mean}$ indicates a static MB of low quality, which is allocated to class IV.

The actual MD class decision tree is shown in the Figure 3.

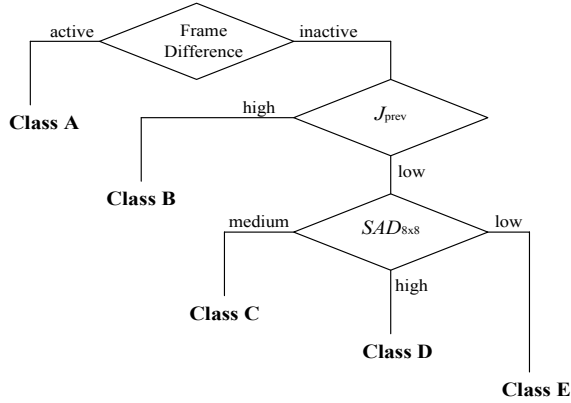


Figure 3. MD class decision tree.

4. ALGORITHM

The proposed MD class selection algorithm with Pareto-optimal distribution of MD classes is shown in Fig.4.

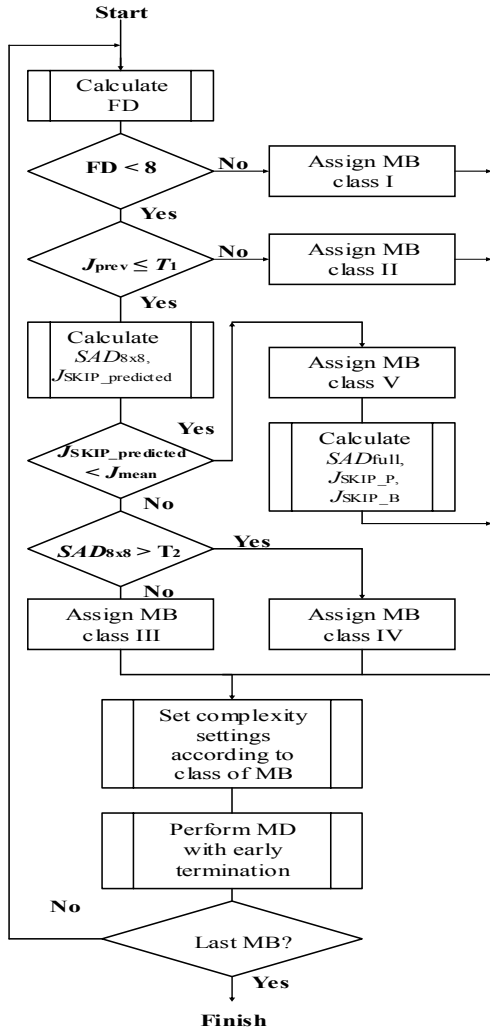


Figure 4. Proposed fast MD algorithm.

The fast mode decision algorithm is based on our earlier work [2]. In order to omit unnecessarily MD computations it utilizes early termination based on the J value from the previous frame. If J is less than J_{mean} for the previous frame then the MD process is terminated early.

5. EXPERIMENTAL RESULTS

The performance of the proposed algorithm was compared with that of the reference JM9.5 encoder [10]. QCIF and CIF video sequences (300 frames each) with different degrees of motion and spatial complexity were encoded at 30fps using the two test configurations shown in Table 1.

Table 1. Test configurations used in experiments

Encoding parameter	Config. A	Config. B
GOP structure	IPPP	IBBP
Ref. frames	1	5
RDO	off	on
Search method	Full search	
Entropy coder	CABAC	
QP	28, 32, 36 and 40	

The results of bit rate change ΔBR , PSNR drop $\Delta PSNR$ and total encoding time change Δt averaged across all QP calculated as in Eqs. (6)–(8) are provided in Tables 2 and 3. The minus sign (–) indicates an improvement for the new method.

$$\Delta Bits = \frac{Bits_{method} - Bits_{JM}}{Bits_{JM}} \cdot 100\% \quad (6)$$

$$\Delta PSNR = PSNR_{JM} - PSNR_{method} \quad (7)$$

$$\Delta t = \frac{t_{method} - t_{JM}}{t_{JM}} \cdot 100\% \quad (8)$$

Table 2. Experimental results for configuration A

Video sequence	ΔBR , %	$\Delta PSNR$, dB	Δt , %
Table tennis, QCIF	2.52	0.25	-60.1
News, QCIF	1.54	0.21	-65.7
Container, QCIF	2.21	0.22	-71.3
Akiyo, QCIF	-0.26	0.18	-72.6
Hall, CIF	-2.06	0.13	-60.2
Paris, CIF	0.72	0.12	-58.4

Table 3. Experimental results for configuration B

Video sequence	ΔBR , %	$\Delta PSNR$, dB	Δt , %
News, QCIF	1.33	0.24	-62.3
Hall, QCIF	-0.99	0.10	-66.0
Akiyo, QCIF	-2.11	0.09	-70.0
Mobile, CIF	0.07	0.25	-56.8
Paris, CIF	2.17	0.23	-66.3

Total encoding time is reduced by roughly 60–73% depending on the video sequence, except for Mobile at 56%. On low motion sequences (i.e. Akiyo, Hall) the bit rate is reduced since the algorithm produces a slightly higher SKIP rate than the original coder. Quality degradation of 0.1–0.25 dB is acceptable. The highly textured sequence Mobile and the high motion sequence Tennis have greatest $\Delta PSNR$ and least Δt . Visual examination of all decoded video sequences did not reveal any anomalies or blocking artifacts.

Fig. 5 presents the R-D curve for a typical sequence, Paris. The R-D curve of the fast MD is almost the same as that produced by the reference coder.

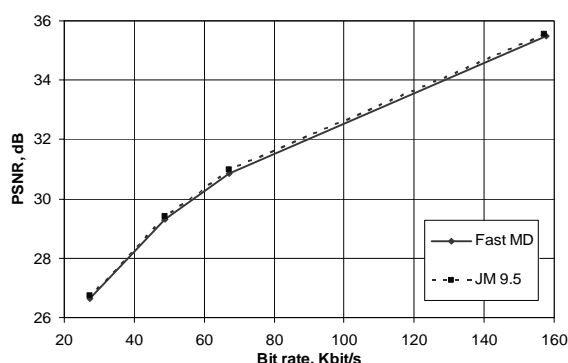


Figure 5. The R-D curve of the sequence Paris, CIF.

The accuracy of the class decision metrics was assessed for configuration B. Class decisions (except I and II) were compared with the final mode selected for the same MB by the reference encoder. The results in Table 4 indicate that the class decision metrics are accurate. The lower accuracy for the Intra decision (class IV), especially for Mobile, indicates that the $SAD_{8 \times 8} > T_2$ condition should be improved, possibly by adding an adaptive threshold scheme for T_2 .

Table 4. Accuracy of the class decision metrics, %

Video sequence	Class II	Class IV	Class V
News, QCIF	96.4	82.7	96.9
Hall, QCIF	95.8	88.3	92.4
Akiyo, QCIF	94.8	97.4	99.1
Mobile, CIF	96.9	57.0	81.2
Paris, CIF	95.1	87.7	92.6

6. CONCLUSIONS

A fast mode decision algorithm for H.264/AVC based on a macroblock classification strategy was proposed. A Pareto Analysis was used to derive optimal MD classes. An efficient combination of class decision metrics was utilized which provides results which are well correlated with full search Mode Decision results.

Experimental results indicate that the proposed algorithm provides a 57–73% reduction in total computational complexity with little impact on bit rate and visual quality.

ACKNOWLEDGEMENTS

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

7. 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] Y.V. Ivanov and C.J. Bleakley, "Skip Prediction and Early Termination for Fast Mode Decision in H.264/AVC", *Proc. Int. Conf. on Digital Communications (ICDT)*, August 2006.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] B. Feng, G.Zhu and W.Liu, "Fast Adaptive Inter Mode Decision Method for P Slices in H.264", *Proc. IEEE 3rd Int. Conf. on Consumer Communications and Networking (CCNC'06)*, vol. 2, pp. 745-748, 2006.
- [7] C. Kim and C.C. Jay Kuo, "A Feature-based Approach to Fast H.264 Intra/Inter Mode Decision", *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS'05)*, vol. 1, pp.308-311, 23-26 May, 2005.
- [8] H.Wang and Z.Zhu, "Fast Mode Decision and Reduction of Reference Frames for H.264 Encoder", *Proc. Int. Conf. On Control and Automation (ICCA'05)*, vol. 2, pp. 1040-1043, 26-29 June 2005.
- [9] Y.V. Ivanov and C.J. Bleakley, "Survey and Pareto Analysis Method for Coding Efficiency Assessment of Low Complexity H.264 Algorithms", *Proc. 10th Irish Machine Vision and Image Processing Conf. (IMVIP)*, pp. 172-179, 2006.
- [10] JVT reference software JM 9.5, on the Web: <http://iphome.hhi.de/suehring/tml>.