<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Pareto-Optimal Macroblock Classification for Fast Mode Decision in H.264</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Authors(s)</strong></td>
<td>Ivanov, Yuri; Bleakley, Chris J.</td>
</tr>
<tr>
<td><strong>Publication date</strong></td>
<td>2007-11-27</td>
</tr>
<tr>
<td><strong>Conference details</strong></td>
<td>2007 IEEE International Conference on Signal Processing and Communications (ICSPC), Dubai, United Arab Emirates, 24 - 27 November 2007</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>IEEE</td>
</tr>
<tr>
<td><strong>Item record/more information</strong></td>
<td><a href="http://hdl.handle.net/10197/7119">http://hdl.handle.net/10197/7119</a></td>
</tr>
<tr>
<td><strong>Publisher's statement</strong></td>
<td>(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</td>
</tr>
<tr>
<td><strong>Publisher's version (DOI)</strong></td>
<td>10.1109/ICSPC.2007.4728600</td>
</tr>
</tbody>
</table>
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 \( J_{\text{prev}}, SAD_{8}\times8 \) 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:

\[
    W = \Delta R + \mu \Delta D \tag{1}
\]

where \( \Delta R \) is bit rate change as a percentage, \( \Delta D \) is quality loss in dB and \( \mu \) is a constant relating bit rate loss and distortion increase (\( \mu=\frac{13}{2} \)). The choice of \( \mu \) 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.](image)

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.](image)

### 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]:

\[
    FD = \sum_{i=0}^{15} \sum_{j=0}^{15} \begin{cases} 
        1 & |C(x,y) - P(x,y)| \geq T_{\text{diff}} \\
        0 & |C(x,y) - P(x,y)| < T_{\text{diff}} 
    \end{cases} \tag{2}
\]

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 \( T_{\text{diff}} \) 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 \( J_{\text{prev}} \)) can be utilized. In [2], it was demonstrated that \( J_{\text{prev}} \) is highly correlated between successive blocks and that the correlation is much higher than for motion vectors [2]. \( J_{\text{prev}} \) is also a good indicator of how efficiently the MB was coded. Herein, inactive MBs with poor coding efficiency, i.e. \( J_{\text{prev}} \approx T_1 \), 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:

\[
    J_{\text{SKIP, predicted}} = f_1(SAD_{8x8}) = \alpha \cdot SAD_{8x8} + C_1 \tag{3}
\]

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

\[
    J_{\text{SKIP, P}} = f_2(SAD_{full}) = \beta \cdot SAD_{full} + C_2 \tag{4}
\]

\[
    J_{\text{SKIP, B}} = \gamma \cdot f_3(SAD_{full}) \tag{5}
\]

The coefficients \( \alpha, \beta, \gamma \), \( C_1 \) 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 \( SAD_{8x8} \approx T_2 \) and \( J_{\text{SKIP, predicted}} \geq J_{\text{mean}} \) 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.

![MD Class Decision Tree](image)

**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.

![Proposed Fast MD Algorithm](image)

**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>
<thead>
<tr>
<th>Encoding parameter</th>
<th>Config. A</th>
<th>Config. B</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOP structure</td>
<td>IPPP</td>
<td>IBBP</td>
</tr>
<tr>
<td>Ref. frames</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>RDO</td>
<td>off</td>
<td>on</td>
</tr>
<tr>
<td>Search method</td>
<td>Full search</td>
<td>CABAC</td>
</tr>
<tr>
<td>QP</td>
<td>28, 32, 36 and 40</td>
<td></td>
</tr>
</tbody>
</table>

The results of bit rate change $\Delta BR$, PSNR drop $\Delta PSNR$ and total encoding time change $\Delta t$ 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.

\[
\Delta Bits = \frac{\text{Bits}_{\text{method}} - \text{Bits}_{\text{JM}}}{\text{Bits}_{\text{JM}}} \times 100\% \tag{6}
\]

\[
\Delta PSNR = PSNR_{\text{JM}} - PSNR_{\text{method}} \tag{7}
\]

\[
\Delta t = \frac{t_{\text{method}} - t_{\text{JM}}}{t_{\text{JM}}} \times 100\% \tag{8}
\]

<table>
<thead>
<tr>
<th>Video sequence</th>
<th>$\Delta BR$, %</th>
<th>$\Delta PSNR$, dB</th>
<th>$\Delta t$, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table tennis, QCIF</td>
<td>2.52</td>
<td>0.25</td>
<td>–60.1</td>
</tr>
<tr>
<td>News, QCIF</td>
<td>1.54</td>
<td>0.21</td>
<td>–65.7</td>
</tr>
<tr>
<td>Container, QCIF</td>
<td>2.21</td>
<td>0.22</td>
<td>–71.3</td>
</tr>
<tr>
<td>Akiyo, QCIF</td>
<td>–0.26</td>
<td>0.18</td>
<td>–72.6</td>
</tr>
<tr>
<td>Hall, CIF</td>
<td>–2.06</td>
<td>0.13</td>
<td>–60.2</td>
</tr>
<tr>
<td>Paris, CIF</td>
<td>0.72</td>
<td>0.12</td>
<td>–58.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Video sequence</th>
<th>$\Delta BR$, %</th>
<th>$\Delta PSNR$, dB</th>
<th>$\Delta t$, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>News, QCIF</td>
<td>1.33</td>
<td>0.24</td>
<td>–62.3</td>
</tr>
<tr>
<td>Hall, QCIF</td>
<td>–0.99</td>
<td>0.10</td>
<td>–66.0</td>
</tr>
<tr>
<td>Akiyo, QCIF</td>
<td>–2.11</td>
<td>0.09</td>
<td>–70.0</td>
</tr>
<tr>
<td>Mobile, CIF</td>
<td>0.07</td>
<td>0.25</td>
<td>–56.8</td>
</tr>
<tr>
<td>Paris, CIF</td>
<td>2.17</td>
<td>0.23</td>
<td>–66.3</td>
</tr>
</tbody>
</table>
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 ΔPSNR 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.

![Fig. 5. The R-D curve of the sequence Paris, CIF.](image)

Table 4. Accuracy of the class decision metrics, %

<table>
<thead>
<tr>
<th>Video sequence</th>
<th>Class II</th>
<th>Class IV</th>
<th>Class V</th>
</tr>
</thead>
<tbody>
<tr>
<td>News, QCIF</td>
<td>96.4</td>
<td>82.7</td>
<td>96.9</td>
</tr>
<tr>
<td>Hall, QCIF</td>
<td>95.8</td>
<td>88.3</td>
<td>92.4</td>
</tr>
<tr>
<td>Akiyo, QCIF</td>
<td>94.8</td>
<td>97.4</td>
<td>99.1</td>
</tr>
<tr>
<td>Mobile, CIF</td>
<td>96.9</td>
<td>57.0</td>
<td>81.2</td>
</tr>
<tr>
<td>Paris, CIF</td>
<td>95.1</td>
<td>87.7</td>
<td>92.6</td>
</tr>
</tbody>
</table>

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


