Rate distortion optimal adaptive quantization and coefficient thresholding for MPEG coding
Farin, D.S.; Kaesemann, M.; de With, P.H.N.; Effelsberg, W.W.J.

Published in:
Proceedings of the 23rd international symposium on information theory in Benelux, Louvain-la Neuve, Belgium, May 29-31, 2002

Published: 01/01/2002

Document Version
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Table 3: Compression rates in bits per source byte for with new memory reduction scheme, comparing CTW with CTM when the model bits are coded with a short range K-T estimator. Note that the first column of this Table is a replication of the third numerical column of Table 2. The left most two columns use the same limit on $a + b$ as in Table 1.2. The third and fourth column prune less severely.

<table>
<thead>
<tr>
<th>Pruning Rule</th>
<th>$a + b \leq \min(2, t/20000)$</th>
<th>$a + b \leq t/100000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>CTM</td>
<td>CTW</td>
</tr>
<tr>
<td>bib</td>
<td>1.98</td>
<td>1.89</td>
</tr>
<tr>
<td>book1</td>
<td>2.28</td>
<td>2.27</td>
</tr>
<tr>
<td>book2</td>
<td>2.07</td>
<td>2.05</td>
</tr>
<tr>
<td>geo</td>
<td>4.60</td>
<td>4.59</td>
</tr>
<tr>
<td>news</td>
<td>2.62</td>
<td>2.57</td>
</tr>
<tr>
<td>obj1</td>
<td>4.11</td>
<td>3.81</td>
</tr>
<tr>
<td>obj2</td>
<td>2.91</td>
<td>2.79</td>
</tr>
<tr>
<td>paper1</td>
<td>2.46</td>
<td>2.34</td>
</tr>
<tr>
<td>paper2</td>
<td>2.35</td>
<td>2.28</td>
</tr>
<tr>
<td>paper3</td>
<td>2.02</td>
<td>2.52</td>
</tr>
<tr>
<td>paper4</td>
<td>3.00</td>
<td>2.83</td>
</tr>
<tr>
<td>paper5</td>
<td>3.16</td>
<td>2.95</td>
</tr>
<tr>
<td>paper6</td>
<td>2.56</td>
<td>2.41</td>
</tr>
<tr>
<td>pic</td>
<td>0.80</td>
<td>0.80</td>
</tr>
<tr>
<td>proce</td>
<td>2.54</td>
<td>2.39</td>
</tr>
<tr>
<td>progl</td>
<td>1.83</td>
<td>1.73</td>
</tr>
<tr>
<td>progt</td>
<td>1.84</td>
<td>1.72</td>
</tr>
<tr>
<td>trans</td>
<td>1.65</td>
<td>1.53</td>
</tr>
<tr>
<td>AVG</td>
<td>2.52</td>
<td>2.41</td>
</tr>
</tbody>
</table>
2 Adaptive Quantization

Consider the problem of coding an image at rate with minimum distortion (MSE) . Each image consists of a fixed number of coding units (e.g., macroblocks), which can each be coded with different quantizer settings . Let be the distortion of macroblock when quantized with , and let be the number of bits required for coding the macroblock. The optimization problem can now be formulated as

\[
\min_{q_i} \sum_i D_i(q_i) \quad \text{such that} \quad \sum_i R_i(q_i) \leq R_{\text{max}}.
\]  

(1)

In [5], Shoham showed that by using the Lagrange-multiplier framework, the constrained optimization problem can be written as the equivalent problem

\[
\min_{q_i} \sum_i D_i(q_i) + \lambda R_i(q_i)
\]  

(2)

for a fixed . The paper [5] also provides the proof that each solution of the transformed (unconstrained) problem is also a solution of the original problem with the rate-constraint if the rate-distortion function is convex. As is dependent on , a suitable value of has to be determined to solve the original problem with . The suitable value can be determined using a binary search.

The advantage of the second problem formulation in eq. (2) is that the sum and the minimum operator can be exchanged to

\[
\sum_i \min_{q_i} D_i(q_i) + \lambda R_i(q_i).
\]  

(3)

This formulation obviously reveals that the global optimization can now be carried out independently for each macroblock, making an efficient implementation feasible.

Unfortunately, according to the MPEG standard, changing the quantization scale requires additional bits in the macroblock header to code the new settings. The overhead comprises 2 bits for coding the macroblock mode and 5 bits for the quantization scale, compared to only 1 bit for the macroblock mode when the quantizer is the same as in the last macroblock. Especially at low bit rates, this overhead cannot be ignored. Hence, we introduce the quantizer change overhead as an extra contribution to the rate:

\[
R^{q_i}(q_i, q_{i-1}) = \begin{cases} 
1 & \text{for } q_i = q_{i-1}, \\
7 & \text{for } q_i \neq q_{i-1}, \\
1 & \text{if } q_i \text{ is the first MB in a slice.}
\end{cases}
\]  

(4)

\footnote{In this paper, we use the term rate to denote the number of bits per frame.}

\footnote{In practice, exact equivalence of rate cannot be guaranteed, and a suitable tolerance has to be accepted.}

\footnote{Note that additional header fields exist. As they have constant size, they can be ignored in the minimization problem. However, they have to be considered when calculating the total rate.}

After adding the overhead to the functional in (2), our optimization problem reads as

\[
\min_{q_i} \sum_i D_i(q_i) + \lambda R_i(q_i) + \lambda R^{q_i}(q_i, q_{i-1}).
\]  

(5)

This can no longer be solved independently for each macroblock, but can be determined using a dynamic programming approach. Consider the graph in Figure 1, in which each column of nodes represents a macroblock and each row defines a quantizer scale. Each path through the graph corresponds to a possible coding of the frame. Traversing the nodes induces associated costs and graph edges from row to row have costs . Hence, the total path cost is equivalent to functional (5), and the minimum cost path defines the solution of the above minimization problem.

Fig. 1. Equivalent graph search problem to the Lagrangian minimization problem. Note that there are 31 different quantizer scales in MPEG instead of only four.

3 Thresholding

Ramchandran [4] introduced coefficient thresholding as a post-processing step after quantization to further reduce the bit rate while still retaining as much image quality as possible. The idea is to drop coefficients (set to zero) when the additional distortion is small compared to the number of bits saved. Considering thresholding as a separate post-processing step increases the difficulty that it is not clear how to choose quantization parameters. If a target rate of is requested, obviously the rate after quantization has to be greater, so that thresholding can be used to further decreasing the rate. However, the exact rate is unknown. We solved this problem by incorporating coefficient thresholding together with adaptive quantization into a single Lagrangian framework.

The following algorithm exploits a useful property of the DCT which leads to an efficient implementation of thresholding. As the DCT does not change the norm of a vector, the MSE of a block can either be computed in the
spatial domain or, equivalently, in the frequency domain. This property enables
to calculate efficiently the additional distortion that is introduced by modifying
(or even omitting) a single coefficient.

To simplify notation, we concentrate on a single DCT block, consisting of
coefficients \( c_i \). We denote quantization by \( q_i = Q(c_i) \) and dequantization by
\( c_i = Q^{-1}(q_i) \). Let \( C = \{(p_i, c_i)\} \) be the ordered set (ascending \( p_i \))
of non-zero quantized coefficients (\( p_i \neq 0 \)), where \( p_i \) is the position of the coefficient (in zig-zag order). Hence, by using a table of the Huffman code-lengths \( r(\text{run, value}) \),
the bits needed to code coefficient \( i \) are \( r(p_i - p_{i-1} + 1, c_i) \). Omitting the coefficient
would induce additional distortion \( c_i \). Note that coefficient \( i = 0 \) is always the
DC coefficient which cannot be omitted. Let \( S \subseteq C \) be the subset of coefficients
in the block that we decide to code. Hence, we intend to minimize the Lagrangian
cost associated to a selection \( S \):

\[
\min_{S \subseteq C} \left\{ \sum_{(p_i,c_i) \in C-S} c_i^2 + \sum_{1 \leq k \leq |S|} (c_i - \tilde{c}_i)^2 + \lambda r(p_i - p_{i-1} + 1, c_i) \right\}.
\]

Similar to adaptive quantization in the previous section, this minimization
problem can be solved by computing an equivalent graph search. The corre-
spending graph is depicted in Figure 2. Every non-zero coefficient is represented
by a graph node. A special node \( EOB \) is added as a last node so that skipping
the last coefficient is possible. Each non-skipping edge is attributed with weight
\( (c_i - \tilde{c}_i)^2 + \lambda r(p_i - p_{i-1} + 1, c_i) \), consisting of the quantization error and the
length of the Huffman code. Each skipping edge is attributed with weight \( \sum \tilde{c}_i \),
where the sum includes all skipped coefficients.

![Fig. 2. Equivalent graph search problem to coefficient thresholding.](image)

To visualize the principle of coefficient thresholding, we coded a frame using
fixed quantization scales. Afterwards, we applied coefficient thresholding to fur-
ther reduce the bit rate. The result for a frame of the \( \text{Claire} \) sequence is shown
in Figure 5. It can be seen that for small reductions of rate in the thresholding
step, the slope of the rate-distortion curve is less than that using quantization only.
However, for larger rate reductions, the slope of the thresholding curves
becomes much steeper, corresponding to a faster decrease of image quality.

To optimally join adaptive quantization and thresholding, we merged the
adaptive quantization graph and the thresholding graph into a single combined

![Fig. 3. Combination of adaptive quantization graph and thresholding graph. Each
box shown with a thresholding graph actually contains six concatenated thresholding
graphs (for the six DCT blocks contained in each macroblock).](image)

graph (Fig. 3). In this way, we get the "convex hull" over the curves of Fig. 5,
being the optimal combination of adaptive quantization and thresholding at
every bit rate.

4 Coefficient Amplitude Reduction
In this section, we introduce \textit{coefficient amplitude reduction} (CAR) as a gen-
eralization of coefficient thresholding. The idea is that it can be advantageous
to decrease the value of a coefficient when the number of bits saved outweighs
the additional distortion. Especially when the true coefficient value is near the
lower decision boundary of the quantization interval, reducing the coefficient
amplitude does not introduce much additional distortion (Fig. 4a). On the other
hand, when a slight decrease prevents the run-value pair to be coded with costly
escape sequences, the bit rate gain may be significant. As the MPEG Huffman
table is monotone, increasing coefficients will never lead to shorter codes.

CAR can be implemented by extending the thresholding graph as shown in
Fig. 4b. For each coefficient with value \( q_i \), further \( c_i \) nodes are created, represen-
ting the new (reduced) value of the coefficient (range 1,...,\( q_i \)). Node costs
are assigned accordingly. All new edges and dummy nodes have zero cost.

CAR can be implemented independently of thresholding by omitting the skip-
ning edges in Fig. 4b. A further advantageous property is that the combination
of CAR and thresholding can be implemented by successive application of CAR
first and thresholding afterwards on the modified coefficients. Clearly, the CAR
step selects the optimal coefficient value, equivalent to the shortest path between
every second node. As every sub-path of a shortest path is also a shortest path,
either this path or a skipping edge would be chosen by the thresholding step. By
replacing the thresholding sub-graphs in Fig. 3 with the CAR graphs, we obtain
the theoretically optimal encoding of the frame.
5 Results

We have implemented the above algorithms into the SAMPEG encoder framework [1]/[2]. A single frame was selected from a test sequence and coded with different combinations of quantization and coefficient modification. For quantization, we used three variants: constant quantization (all macroblocks are coded with the same quantization scale), adaptive quantization (as explained above), and adaptive quantization without considering quantizer-scale change overhead (NOO). Furthermore, we used the TMS reference implementation [6] for comparison. For coefficient modification, we used: thresholding alone, CAR alone, both combined, and both disabled.

Table 1 shows the absolute PSNR reached for a fixed rate and the increase in PSNR compared to using constant quantizer scales. Adaptive quantization increases the PSNR by about 0.13 dB. Applying thresholding leads to another 0.2-0.3 dB increase. The PSNR increase obtained from CAR is only marginal and can be neglected. NOO cannot increase PSNR much above the constant quantization heuristic. At low bit rates, it even performs worse because of frequent quantizer changes. Comparable results are obtained for other input images.

According to our results, using frame-constant quantization scales is a good heuristic for PSNR optimal quantization. The heuristic achieves \(-1.2 - 1.8\) dB compared to TMS and is only \(0.3 - 0.5\) dB below the theoretical maximum.

In a second experiment, we examined which coefficients in a DCT block were thresholded most. Approximately 70-80% of the thresholded coefficients were at the end of the DCT block. Skipping these coefficients moves the EOB code to an earlier position, which results in a particularly large reduction of bits. Accordingly, applying thresholding to only the last coefficients of a block results in an about 80% of the PSNR increase. This fact may enable computationally inexpensive heuristics for fast thresholding.

Table 1. Overall results: PSNR in dB (absolute and increase compared to constant quantization). The bits per frame were chosen as if the sequence was coded at 1.2 Mbps (at CIF resolution).

6 Conclusions

We have presented an optimal quantization algorithm for MPEG coded I-frames. It achieves to generate images with the best possible image quality at a given bit rate. Even though it may be computationally too complex for practical encoding applications, it is suitable to serve as a reference to compare other, heuristic, algorithms with. The algorithm can easily be extended to support P- and B-frames by including the additional macroblock mode decisions in a similar way.

References


7 Even though thresholding can be combined with constant quantization, the performance highly depends on the selected rate (see Fig. 5).
Flexible Access to JPEG2000 Codestreams

Marcela Iregui, Jérôme Meessen, Philippe Chevalier, and Benoît Macq

Université Catholique de Louvain, Communications and Remote sensing laboratory, Bâtiment Stevin, Place du Levant, 2 B-1348 Louvain-la-Neuve, Belgium

{iregui, meessen, macq}@ele.ucl.ac.be
http://www.tale.ucl.ac.be

Abstract. This paper presents an efficient way for delivering JPEG2000 (2K) data in a client-server architecture to seamlessly browse images from a remote server. The main advantage of the emerging 2K coding algorithm is its flexibility of the generated codestream, which permits seamless navigation through very large images and sets of images, an useful issue in medical remote diagnosis or remote sensing applications. Herein, we present a data flow strategy based on optimal data parsing along with an interactive communication protocol that allows an efficient data transfer. The relevant portions of the codestream to be sent are selected taking into account several parameters such as: delays, memory, bandwidth and client displaying and decoding capabilities in state-full sessions (i.e. keeping track of the already sent data to avoid redundancies). We propose an optimal use of the available resources according to the user preferences by using optimization techniques.

1 Introduction

In professional applications like remote sensing and medical imaging, there exists a large need to browse images from a remote server in a seamless way. However, this process can be very slow due to the large size of data and bandwidth limitations, so that only a portion of the image can be displayed at a particular time. Moreover in dynamic browsing interfaces the user can change the request at any moment. In consequence, the response time must be minimized by sending the most important portions of the image at each time.

Lately, new coding algorithms as JPEG2000 allow progressive decoding of portions of the compressed images, which is very useful to meet the above requirements. Thus, we propose a server-client architecture where the client can access and browse 2K images from a remote server. If the server deals with

This work was funded by the IST project PRIAM (IST-28646)