Fast Decision of Block size, Prediction Mode and Intra Block for H.264 Intra Prediction

Gaurav Hansda
1000721849
gaurav.hansda@mavs.uta.edu
Outline

- Introduction to H.264
- Current algorithms for intra prediction
- Proposed algorithms
- Implementation results
- Conclusions
Overview of H.264

• H.264 is an industry standard.
• It defines a format for compressed video data.
• It provides a set of tools that can be used in a variety of ways to compress and communicate visual information.
• Purpose of a standard
  ➢ Define a coded representation (or syntax) that describes visual data in a compressed form and method of decoding the syntax to reconstruct visual information.
  ➢ compliant encoders and decoders can successfully interoperate with each other.
Fig. 1. H.264 profiles [2]
Applications

- Broadcast television
- Streaming video
- Video storage and playback
- Videoconferencing
- Mobile video
- Studio distribution
Coding Process

- The image is divided into macroblocks (16x16 pixels).
- The macroblocks are grouped into slice groups, which are divided into slices.
- Each slice is coded either as an I, P or B slice (There are also types called SI and SP).
- In an I slice, all blocks are coded as I blocks.
- In a P slice, blocks are coded as I or P blocks.
- In a B slice, blocks are coded as I, P or B blocks.
Fig. 2. Typical H.264 encoder [18]
Fig. 3. Typical H.264 decoder [18]
Mode decision of H.264 encoder

Fig. 4. Mode decision hierarchy of an H.264 compliant encoder. [4]
Implication of Hierarchical Structure

1. To ensure the correctness of the decision at upper layer.
2. To ensure early termination is executed accurately and as early as possible.

Most fast mode decision algorithms developed so far, only deal with a single stage of the mode decision hierarchy [5]-[14] and fail to achieve the best possible complexity reduction.
Intra-Prediction

- There are 3 macroblock (MB) modes for intra prediction of luma pixels: intra4x4 (I4MB), intra8x8 (I8MB), and intra16x16 (I16MB).

- Intra4MB and Intra8MB have 9 prediction modes as shown in Fig. 5(a). Intra16MB has only 4 prediction modes as shown in Fig. 5(b).

Fig. 5. Prediction modes for (a) Intra4MB and (b) Intra16MB. [4]
Fig. 6. Prediction flow diagram [18]

Fig. 7. Intra-prediction [18]
Mode Decision

- To achieve a better tradeoff between bit-rate and distortion, H.264 encoder adopts the rate-distortion (R-D) optimization framework and the Lagrangian technique for mode decision [2].
- For intra frames, the best prediction mode of a block is defined as the mode that, among all prediction modes of the block, gives rise to the minimum R-D cost.
- The R-D cost of an MB mode is the sum of the minimum R-D cost of each individual block.
Proposed Algorithm for Block Size Decision

• Block size is highly correlated with texture complexity.
• Variance of a block corresponds to the total energy of the AC coefficients of the block, hence it is good measurement of the texture complexity.
• Thus variance based classification of texture complexity is used [16].
• If variance is above the threshold, Intra4MB and Intra8MB is selected; otherwise, Intra8MB and Intra16MB is chosen. This is simple way to skip the examination of Intra4MB mode.
Fig. 8. Variance-based MB mode decision [4]
Improved Prediction Mode Decision

- Earlier algorithms only consider the edge information of the current block. The correlation between blocks was not considered.
- Hence the Most Probable Mode (MPM) is used.
- The MPM, which takes advantage of the spatial correlation of the prediction modes between the neighboring blocks and the current block for coding, is defined as the prediction mode of the left or the upper neighbor, whichever has the smaller prediction mode number.
Improved Prediction Mode Decision

- Each original image block is evenly divided into four subblocks first. Each subblock is represented by the average pixel magnitude of its pixels.

Fig. 9. Formation of subsampled block for a block of (a) Intra4MB, (b) Intra8MB, and (c) Intra16MB [4]
• Apply the following filters:

<table>
<thead>
<tr>
<th>Vertical</th>
<th>Horizontal</th>
<th>45° diagonal</th>
<th>135° diagonal</th>
<th>No direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_0^v = 1$, $f_1^v = -1$</td>
<td>$f_0^h = 1$, $f_1^h = 1$</td>
<td>$f_0^{45^\circ} = \sqrt{2}$, $f_1^{45^\circ} = 0$</td>
<td>$f_0^{135^\circ} = 0$, $f_1^{135^\circ} = \sqrt{2}$</td>
<td>$f_0^{nd} = 2$, $f_1^{nd} = -2$</td>
</tr>
<tr>
<td>$f_2^v = 1$, $f_3^v = -1$</td>
<td>$f_2^h = -1$, $f_3^h = -1$</td>
<td>$f_2^{45^\circ} = 0$, $f_3^{45^\circ} = \sqrt{2}$</td>
<td>$f_2^{135^\circ} = -\sqrt{2}$, $f_3^{135^\circ} = 0$</td>
<td>$f_2^{nd} = -2$, $f_3^{nd} = 2$</td>
</tr>
</tbody>
</table>

Fig. 10. Five sets of filter coefficients for dominant edge detection. [4]

• Determine the dominant edge.
Fig. 11. Prediction mode decision [4]
Proposed algorithm for Intra Block Decision

- Intra block decision, for inter frames, occupies a considerable percentage of the total computations of inter-frame coding.
- Intra16MB takes much less computation time than the other modes.
- Hence, scaled R-D cost is used [4].
- An MB is less probable to be intra coded if the R-D cost difference between best inter mode and Intra16MB is small. Denoting the scaled R-D cost differences between Intra16MB and the inter MB mode by $d^\wedge J$, and based on the above observation, if $d^\wedge J$ is small both I4MB and I8MB can be skipped.
Joint Model (JM)

- Reference implementation standardized in ISO/IEC JTC1/SC29/WG11
- Decoder implements almost all the features
- Encoder
  - Exercises most of the important coding tools
  - Provides an elaborate list of control parameters
  - Offers a rate distortion optimized implementation
  - Offers several fast computation options
  - Serves as a reference for what is best quality possible using H.264
  - Good description of the reference algorithms exists
- Currently at v17.2
- Can be downloaded from http://iphome.hhi.de/suehring/tml/download/
Sequences used

Hall (CIF and QCIF)

Container (CIF and QCIF)

Mobile (CIF)
The JM reference software version 17.2 is implemented [17].

The conditions of the experiment are as follows.

1. Run on a PC with Intel Core i3 2.27GHz processor and 3.00 GB RAM.
2. Set the QP values to 16, 20, 24, and 28.
3. Number of frames to be coded = 100
4. Enable the R-D optimization.
5. Choose context-adaptive binary arithmetic coding (CABAC) as the entropy coding method.
### Performance

<table>
<thead>
<tr>
<th>QP</th>
<th>PSNR (dB)</th>
<th>Bit rate (kbits/s)</th>
<th>Time (s)</th>
<th>SSIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>38.685</td>
<td>697.64</td>
<td>20.022</td>
<td>0.9766</td>
</tr>
<tr>
<td>24</td>
<td>41.378</td>
<td>961.59</td>
<td>30.875</td>
<td>0.9831</td>
</tr>
<tr>
<td>20</td>
<td>44.118</td>
<td>1359.14</td>
<td>34.094</td>
<td>0.9883</td>
</tr>
<tr>
<td>16</td>
<td>47.085</td>
<td>1930.5</td>
<td>37.767</td>
<td>0.9934</td>
</tr>
</tbody>
</table>

Table I. Performance of Hall (QCIF) in JM reference software version 17.2
# Variance-based block size decision

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Decrease in PSNR (dB)</th>
<th>Increase in bitrate (%)</th>
<th>Decrease in encoding time (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Container (QCIF)</td>
<td>0.099</td>
<td>0.49</td>
<td>23.7</td>
</tr>
<tr>
<td>Hall (QCIF)</td>
<td>0.051</td>
<td>0.27</td>
<td>11.5</td>
</tr>
<tr>
<td>Container (CIF)</td>
<td>0.045</td>
<td>0.26</td>
<td>25.4</td>
</tr>
<tr>
<td>Hall (CIF)</td>
<td>0.054</td>
<td>0.43</td>
<td>24.2</td>
</tr>
<tr>
<td>Mobile (CIF)</td>
<td>0.019</td>
<td>0.07</td>
<td>7.41</td>
</tr>
</tbody>
</table>

Table II. Performance of variance based block size decision in JM reference software version 17.2 compared to using JM v17.2 directly.
Prediction mode decision [8]

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Decrease in PSNR (dB)</th>
<th>Increase in bitrate (%)</th>
<th>Decrease in encoding time (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Container (QCIF)</td>
<td>-0.186</td>
<td>0.95</td>
<td>-31.0</td>
</tr>
<tr>
<td>Hall (QCIF)</td>
<td>-0.181</td>
<td>1.01</td>
<td>-28.6</td>
</tr>
<tr>
<td>Container (CIF)</td>
<td>-0.150</td>
<td>0.83</td>
<td>-29.1</td>
</tr>
<tr>
<td>Hall (CIF)</td>
<td>-0.203</td>
<td>1.60</td>
<td>-32.0</td>
</tr>
<tr>
<td>Mobile (CIF)</td>
<td>-0.174</td>
<td>0.64</td>
<td>-25.4</td>
</tr>
</tbody>
</table>

Table III. Performance of Prediction mode decision [8] in JM reference software version 17.2 compared to using JM v17.2 directly.
### Improved Prediction Mode Decision

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Decrease in PSNR (dB)</th>
<th>Increase in bitrate (%)</th>
<th>Decrease in encoding time (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Container (QCIF)</td>
<td>-0.092</td>
<td>0.47</td>
<td>-31.0</td>
</tr>
<tr>
<td>Hall (QCIF)</td>
<td>-0.118</td>
<td>0.65</td>
<td>-29.1</td>
</tr>
<tr>
<td>Container (CIF)</td>
<td>-0.075</td>
<td>0.41</td>
<td>-29.1</td>
</tr>
<tr>
<td>Hall (CIF)</td>
<td>-0.091</td>
<td>0.70</td>
<td>-32.4</td>
</tr>
<tr>
<td>Mobile (CIF)</td>
<td>-0.136</td>
<td>0.50</td>
<td>-25.9</td>
</tr>
</tbody>
</table>

Table IV. Performance of Improved prediction mode decision in JM reference software version 17.2 compared to using JM v17.2 directly.
Two-stage approach

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Decrease in PSNR (dB)</th>
<th>Increase in bitrate (%)</th>
<th>Decrease in encoding time (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Container (QCIF)</td>
<td>-0.191</td>
<td>0.97</td>
<td>-48.5</td>
</tr>
<tr>
<td>Hall (QCIF)</td>
<td>-0.161</td>
<td>0.90</td>
<td>-36.7</td>
</tr>
<tr>
<td>Container (CIF)</td>
<td>-0.123</td>
<td>0.68</td>
<td>-50.5</td>
</tr>
<tr>
<td>Hall (CIF)</td>
<td>-0.145</td>
<td>1.12</td>
<td>-52.6</td>
</tr>
<tr>
<td>Mobile (CIF)</td>
<td>-0.154</td>
<td>0.57</td>
<td>-35.5</td>
</tr>
</tbody>
</table>

Table V. Performance of two-stage approach incorporating variance based block size decision and improved prediction mode decision in JM reference software version 17.2 compared to using JM v17.2 directly.
Fig. 12. Performance plot for Hall (QCIF) : Bit rate vs PSNR.

Variance based block size decision
Improved Prediction mode or MPM
Two-Stage approach

Fig. 12. Performance plot for Hall (QCIF) : Bit rate vs PSNR.
Conclusions

- The variance-based block size decision achieves on the average 19% (maximum 25.4%) time saving while maintaining almost the same R-D performance as that of the full-search scheme employed in the JM reference software.
- For Mobile sequence that has more texture regions, less time is saved due to the computationally expensive operations for examining I4MB.
- As the texture of an MB becomes homogeneous as the frame size increases, more time is saved for high resolution sequences. Note that the bit-rate increase for all sequences in this test is below 0.5%, which is negligible.
- Up to 52.0% complexity reduction over the JM reference software is achieved, and both the drop in PSNR and the increase in bitrate are very small.
- The proposed algorithms provide significant improvement of computational efficiency over previous methods without noticeable R-D performance loss.
References


    This software is a product of Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG. The latest version of JM Software is 18.0