Rate-Distortion Optimized Intra Update for Error Resilience in MPEG-4 Video Coding

Woo-Shik Kim, Rae-Hong Park
Department of Electronic Engineering, Sogang University, CPO Box 1142, Seoul 100-611, Korea

ABSTRACT: Motion compensation is a powerful method to compress a video sequence. However, once an error occurs, the error propagates through the frames. Recently, the intra update method was proposed to stop the error propagation at the cost of the compression efficiency. This paper proposes an intra update method based on rate-distortion optimization in error prone environments. The rate and the distortion are estimated using the Lagrangian optimization method to simultaneously select the coding mode and the quantization step size. The proposed method is applied to moving picture experts group (MPEG)-4 codec, and experimental results show that it is robust to the packet loss than the conventional ones.

Key words: intra update; error resilience; video coding; MPEG-4

I. INTRODUCTION

The moving picture experts group (MPEG)-4 standard enlarges upon not only video compression but also manipulation of multimedia data such as audio, still images, and computer graphics, by adopting advanced technologies including object-based coding, scalable coding, and error resilience tools (ISO/IEC, 2000b; Ebrahimi and Horne, 2000). The error resilience tools are efficient especially in the packet-based communication environment, where packet losses bring severe distortions to the decoded images of compressed video (Wang and Zhu, 1998; Wang et al., 2000; Wenger et al., 1998).

Though motion compensation methods are effective for video compression by reducing the redundancies between contiguous frames (Rao and Hwang, 1996), one of their main drawbacks is the error propagation caused by using the erroneous area in the previous frame in motion compensation.

The intra update algorithms have been proposed to stop the error propagation by forcing not to use the motion compensation. They reduce the error propagation at the cost of the bit rate increase. The rate control should be considered to adjust the amount of the intra-updated area.

In this article, an intra update algorithm is proposed, in which the rate-distortion optimized selection of the coding mode and the quantization parameter is used for each macroblock (MB) according to the packet loss rate.

The rest of the article is organized as follows. In Section II, the conventional intra update and rate-distortion optimization algorithms are reviewed including the mode selection scheme in MPEG-4. In Section III, the proposed algorithm is presented and experimental results are discussed in Section IV. Finally, conclusions are given in Section V.

II. CONVENTIONAL ALGORITHMS

A. Intra Update Algorithms. In MPEG-4 visual standard annex E (ISO/IEC, 1998; Budagavi et al., 2000), adaptive intra refresh is specified in which the fixed amount of motion area is coded in intra mode, because the motion area gives a large distortion by error concealment. To choose the motion area, the sum of the absolute differences in each MB is measured during motion estimation.

This method is similar to the method presented in Haskell and Messerschmitt (1992), where the motion area is intra updated and the cycle of intra update of the whole area is smaller than the reciprocal of the error rate. In this algorithm, the mode decision is done heuristically by motion estimation.

An intra update algorithm was proposed to encode the block in which a large error occurs in Liao and Villasenor (2000), and a method to determine the amount of the intra-updated area was presented in Stuhlmueller et al. (2000). A length-based intra refreshment method was proposed, where the MB with a high bit rate in the previous frame is selected as the MB to be coded in intra mode, because such an MB also requires a high bit rate (Lin et al., 2000).

In these algorithms, rate-distortion optimization was not performed in selecting the coding mode.

In Zhang et al. (2000), Côté and Kossentini (1999), Wenger and Côté (1999), and Côté et al. (2000), intra update algorithms were proposed, where the rate distortion optimization method was applied to determine the coding mode. In these algorithms, the distortion was estimated considering the packet loss rate and the concealment effect; however, the selection scheme of a quantization parameter was not described in detail. Also implementation of these algorithms to H.263 codec was presented (ITU-T, 1998).

B. Rate-Distortion Optimization. The limit of the storage space and bandwidth in communication channels requires rate control (Ortega and Ramchandran, 1998; Sullivan and Wiegand, 1998). Selection of a coding mode and a quantization parameter plays a main role in rate control. If the quantization parameter increases, the bit rate decreases while the distortion increases. Intra mode coding not performing motion compensation requires high bit rates, whereas inter mode coding using motion compensation reduces the...
bit rate. Distortions from the previous frames generated not only by the motion estimation errors but also by the packet losses are propagated.

To improve the quality of decoded images under the bit rate constraint, several rate-distortion optimization algorithms have been proposed. A rate control algorithm was proposed, where the quantization parameter was calculated using Lagrangian optimization (Ribas-Corbera and Lei, 1999). In this algorithm, the bit rate and the distortion were estimated using statistical information of the discrete cosine transform (DCT) coefficients. The rate-distortion optimized mode selection methods were proposed in Lee and Dickinson (1997), Wiegand et al. (1996), and Schuster and Katsaggelos (1996), where the mode and the quantization parameter were determined using Lagrangian optimization and dynamic programming (Hiller and Lieberman, 1995; Smith, 1991). The Lagrangian multiplier was changed according to the coding mode (Wiegand and Andrews, 1998).

**C. Mode Selection in MPEG-4.** In the MPEG-4 verification model (VM) (ISO/IEC, 2000a), both intra and inter modes are used to code each MB in the inter-coded frame. Additionally, the intra+Q and inter+Q modes are supported to adjust the quantization parameter in the difference quantizer. The difference quantizer is denoted by an integer, of which the range is from −2 to 2.

To select the coding mode of each MB, the sum of the absolute differences and the deviation, calculated during the motion estimation process, are used. Let \( p_{i,j} \) and \( q_{i,j} \) denote the gray levels at \((i, j)\) in the current and previous frames, respectively. Let the sum of the absolute difference SAD be defined by

\[
SAD(x, y) = \sum_{i=1}^{16} \sum_{j=1}^{16} |p_{i,j} - q_{i+x,j+y}|, \tag{1}
\]

where \((x, y)\) denotes the motion vector. The deviation \( A \) is calculated as

\[
A = \sum_{i=1}^{16} \sum_{j=1}^{16} |p_{i,j} - \mu_{MB}|, \tag{2}
\]

where the sample mean \( \mu_{MB} \) of gray level of an MB is given by

\[
\mu_{MB} = \frac{1}{(16 \times 16)} \sum_{i=1}^{16} \sum_{j=1}^{16} p_{i,j}, \tag{3}
\]

and \( 16 \times 16 \) MBs are assumed. If \( A \) is smaller than \( SAD(x, y) \), the MB is encoded in intra mode.

**III. PROPOSED ALGORITHM**

The intra update can stop the error propagation, at the cost of the bit rate increase. In conventional algorithms, the coding mode is selected considering the distortion caused by the packet loss. To improve the efficiency of the intra update, the coding mode should be chosen by considering the distortion and by adjusting the quantization parameter for each MB. The difference quantization in MPEG-4 is adopted to adjust the quantization parameter for each MB. To determine the coding mode and the difference quantizer of the current MB, the rate and the distortion of the MB are estimated based on Lagrangian optimization.

**A. Estimation of the Rate and the Distortion.** The rate and the distortion are estimated separately in intra and inter modes. If these values are calculated from the encoded and decoded images for each mode and each difference quantizer, the computational complexity is too high. Thus these values should be estimated with a low computational load by using the values computed beforehand, for example, using the sum of absolute differences calculated during motion estimation.

The rate and the distortion of the MB in intra mode are estimated as follows. The rate \( R_{n,\text{intra}} \) of the \( n \)th MB in intra mode is calculated as

\[
R_{n,\text{intra}} = \frac{\sum_{i=1}^{16} \sum_{j=1}^{16} |p_{i,j} - \mu_{n}|}{QP}, \tag{3}
\]

where \( QP \) denotes the quantization parameter of the MB, \( p_{i,j} \) represents the gray level at pixel \((i, j)\) in the MB, and \( \mu_{n} \) denotes the average gray level of the \( n \)th MB. Note that the numerator in (3) is calculated beforehand during the motion estimation procedure.

In intra mode, the distortion caused by error propagation is not considered. The distortion \( D_{n,\text{intra}} \) of the \( n \)th MB in intra mode is calculated as

\[
D_{n,\text{intra}} = \frac{QP^2}{12}, \tag{4}
\]

which reflects the distortion caused by the DCT and the quantization (Ribas-Corbera and Lei, 1999).

Similarly, the rate and the distortion of the MB in inter mode are estimated as follows.

The rate \( R_{n,\text{inter}} \) of the \( n \)th MB in inter mode is calculated as

\[
R_{n,\text{inter}} = \frac{\sum_{i=1}^{16} \sum_{j=1}^{16} |p_{i,j} - q_{i+x,j+y}|}{QP}, \tag{5}
\]

where \( q_{i+x,j+y} \) signifies the gray level at pixel \((i + x, j + y)\), with \((x, y)\) representing the motion vector, and \( C \) denoting the number of bits required to code the motion vector.

The distortion \( D_{n,\text{inter}} \) of the \( n \)th MB in inter mode is calculated as

\[
D_{n,\text{inter}} = (1 - P_{\text{loss}}) \frac{QP^2}{12} + P_{\text{loss}} \sum_{i=1}^{16} \sum_{j=1}^{16} |p_{i,j} - q_{i+x,j+y}|, \tag{6}
\]

where \( P_{\text{loss}} \) represents the packet loss rate. In (6), the distortion consists of two terms as shown in the right-hand side. The first term corresponds to the case in which the packet error does not occur, where the distortion is estimated considering the distortion caused by the DCT and the quantization. If the packet error occurs, error concealment is performed by copying the region in the previous frame specified by the motion vector. The distortion caused by error concealment is estimated as expressed in the second term that is the sum of absolute differences precomputed during the motion estimation step.

Even though the rate and the distortion can be estimated differently, the proposed method shows a simple estimation procedure with a small computational load by using the values already computed in the motion estimation step. The error concealment method is simply assumed as described above, noting that more advanced error concealment techniques can give a smaller distortion term.

Figure 1. Comparison of the decoded images (Foreman sequence, 1% packet loss rate). (a) Original images; (b) Images decoded by MPEG-4 VM; (c) Images decoded by the proposed algorithm.
B. Rate-Distortion Optimized Intra Update.

The rate and the distortion are calculated for each mode (intra or inter) and each difference quantizer value ($\frac{|H_1|}{H_1^2}$ and $\frac{|H_2|}{H_2^2}$). Also the case not using difference quantization is to be considered. Thus the rate and the distortion are estimated for each mode according to the value of the difference quantizer value. Then, the Lagrangian optimization function $J$ defined as

$$J = \sum_n \left( D_{n,\text{mode},DQ} + \lambda_{\text{mode}} R_{n,\text{mode},DQ} \right)$$  \hspace{1cm} (7)

is calculated for all modes and all values of difference quantization, where $\text{mode}$ is either intra or inter, and $DQ$ is one of $\{2, -1, 1, 0\}$. If $DQ$ is equal to zero, difference quantization is not used; otherwise the MB mode selects intra+$Q$ or inter+$Q$ as described in Section IIc. The constant $\lambda_{\text{mode}}$ denotes the Lagrangian multiplier defined as Sullivan and Wiegand (1998):

$$\lambda_{\text{mode}} = 0.85 \times QP^2. \hspace{1cm} (8)$$

The increase in the bit rate can cause buffer overflow, which degrades the quality of the decoded images by frame skipping. In the proposed algorithm, if the buffer is as full as a predetermined threshold value, for example, 80% of the buffer size, the coding mode is set to inter mode.

The computational complexity of the proposed algorithm is quite low, because the rate and the distortion terms used are estimated beforehand during the motion estimation step, such as the sum of the absolute differences between the current MB and the region in the previous frame referred by the motion vector.

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

The proposed algorithm is implemented on the MPEG-4 reference software and its performance is compared with that of the MPEG-4.

![Figure 2](image)

**Figure 2.** Comparison of the decoded images (Dancer sequence, 1% packet loss rate). (a) MPEG-4 VM; (b) proposed algorithm.

![Figure 3](image)

**Figure 3.** Comparison of the average PSNR as a function of the frame index (Foreman sequence). (a) 50 Kbps; (b) 100 Kbps.
VM codec that uses a mode selection method described in Section IIC. In the MPEG-4 VM, the bit rate and the distortion caused by motion compensation are considered to select the coding mode for each MB, whereas not only the bit rate but also the distortion caused by the packet loss and the error concealment is considered in the proposed method. The MPEG-4 VM rate control algorithm is applied to both methods in selecting a quantization parameter of the frame. Among the error resilience techniques supported in MPEG-4, resynchronization and data partitioning are used, where the resynchronization packet size is set to 800 bits.

To simulate the packet loss, the H.323-based packet network system is assumed, where the bit error rate is set to zero and the packet loss rate varies from 0 to 30%. The packet size is assumed as 100 bits. The random number generator is used with the same seed values to make the packet loss concurrently in both bitstreams generated by the codec implementing the proposed algorithm and the MPEG-4 VM codec. Fifty different seed values are used to obtain the average performance of both codecs. Note that only the luminance component is used to calculate the peak signal to noise ratio (PSNR).

The Foreman and Dancer test sequences are used. In the Foreman sequence with a quarter common intermediate format (QCIF) of $176 \times 144$ pixel size and 4:2:0 color format, prominent motions between frames generated by the unstable movement of the camera give the global motions of the scene, and the man shows an expressive facial look. One-third of the original frame rate of the sequence is used to simulate the abrupt motions. The Dancer sequence has common intermediate format (CIF) of $352 \times 288$ pixel size and 4:2:0 color format and shows moderate motions compared with the former one; however, the background and the dancers in the scene steadily change from frame to frame. For this sequence, the same frame rate as that of the original sequence is used.

![Figure 4. Comparison of the average PSNR as a function of the frame index (Dancer sequence). (a) 400 Kbps; (b) 1.2 Mbps.](image)

![Figure 5. Comparison of the average PSNR as a function of the bit rate (1% packet loss rate). (a) Foreman sequence; (b) Dancer sequence.](image)
Figures 1 and 2 show the decoded frames of the Foreman and Dancer sequences, respectively, where the bitstream undergoes 1% packet loss. Figure 1(a) shows the original images, Figure 1(b) illustrates the images decoded by the MPEG-4 VM codec, and Figure 1(c) displays those by the proposed algorithm. In Figure 1(b), the change of the facial expressions becomes obscure because of the error propagation, whereas in Figure 1(c) it is clear.

Figure 2(a) shows the image decoded by the MPEG-4 VM codec, where small square patterns appear because of the error concealment and the error propagation. On the other hand, the proposed algorithm gives a finer appearance as shown in Figure 2(b). Note that the rectangular region in Figure 2(b) represents the intra-updated area in which error propagation is effectively removed.

The performance of the proposed algorithm is evaluated at various bit rates from 30 to 500 Kbps for the Foreman sequence, whereas from 120 Kbps to 2 Mbps for the Dancer sequence. In these experiments, the packet loss rate is set to 1%, of which the size is 100 bits. Figures 3 and 4 show a frame-by-frame average PSNR of the Foreman and Dancer sequences, respectively. Figure 3(a,b) shows the result at the bit rate of 50 and 100 Kbps, respectively. Figure 4(a,b) shows the result at the bit rate of 400 Kbps and 1.2 Mbps, respectively. In Figures 3 and 4, at first the average PSNR of the proposed algorithm is smaller than that of the conventional one, because of using a smaller quantization parameter in difference quantization to code more MBs in intra mode. However, as the frame number increases, the average PSNR of the proposed algorithm becomes larger than that of the conventional one, because of the effective reduction of the error propagation caused by the packet loss, by using the optimized intra update method.

Figure 5(a,b) illustrates the average PSNR as a function of the bit rate for Foreman and Dancer sequences, respectively, with the packet loss rate of 1%. The proposed algorithm shows the enhanced performance by 1–2 dB in terms of the average PSNR compared with the conventional one. The average PSNR of the proposed algorithm at several different bit rates for two test sequences is compared in Table I. Note that the degradation of the average PSNR as the bit rate increases is mainly due to the increment of the amount of the lost packets, which yields severe distortions by the increased error propagation.

In Figure 6, the performance of the proposed algorithm is compared with the conventional one for various packet loss rates from 0.1% to 5%, where the Foreman sequence is used at the bit rate of 100 Kbps. Note that the log scale is used to represent the packet loss rate. When the packet loss rate is very small, the performance of the proposed algorithm is not significant, but as the packet loss rate increases, the improvement in terms of the average PSNR is achieved. Table II lists the average PSNR of the proposed and conventional algorithms.

In Section IV, to show the effectiveness of the proposed algorithm experimental results are shown with various bit rates and packet loss rates. The error concealment is performed during decoding to restore the lost area due to the packet loss. The main objective of the article is to present the rate-distortion optimized intra update with the quantization parameter adjustment. The proposed algorithm can achieve better performance if the rate and the distortion are estimated more precisely. Note that in the proposed algorithm, a simple estimation method is used to reduce a computational load.

Table I. Performance comparison in terms of the average PSNR (dB) at various bit rates.

<table>
<thead>
<tr>
<th>Bit rate (Kbps)</th>
<th>Proposed algorithm</th>
<th>MPEG-4 VM</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreman</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>26.26</td>
<td>25.57</td>
<td>0.69</td>
</tr>
<tr>
<td>50</td>
<td>26.92</td>
<td>25.63</td>
<td>1.29</td>
</tr>
<tr>
<td>100</td>
<td>26.14</td>
<td>25.19</td>
<td>0.95</td>
</tr>
<tr>
<td>300</td>
<td>24.36</td>
<td>22.90</td>
<td>1.46</td>
</tr>
<tr>
<td>120</td>
<td>30.51</td>
<td>29.41</td>
<td>1.10</td>
</tr>
<tr>
<td>Dancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>30.60</td>
<td>29.53</td>
<td>1.07</td>
</tr>
<tr>
<td>400</td>
<td>30.73</td>
<td>29.38</td>
<td>1.35</td>
</tr>
<tr>
<td>1200</td>
<td>27.31</td>
<td>26.37</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Table II. Performance comparison in terms of the average PSNR (dB) at various packet loss rates (Foreman sequence, 100 Kbps).

<table>
<thead>
<tr>
<th>Packet loss rate (%)</th>
<th>Proposed algorithm</th>
<th>MPEG-4 VM</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>33.05</td>
<td>32.94</td>
<td>0.11</td>
</tr>
<tr>
<td>0.5</td>
<td>28.98</td>
<td>27.99</td>
<td>0.99</td>
</tr>
<tr>
<td>1</td>
<td>26.14</td>
<td>25.19</td>
<td>0.95</td>
</tr>
<tr>
<td>5</td>
<td>19.17</td>
<td>18.33</td>
<td>0.84</td>
</tr>
</tbody>
</table>

Figure 6. Comparison of the average PSNR as a function of the packet loss rate (Foreman sequence, 100 Kbps).

V. CONCLUSIONS

In this article, the rate-distortion optimized intra update method in MPEG-4 video coding is proposed to stop the error propagation caused by the packet loss in the packet-based communication channels. The coding mode for each MB is determined considering the bit rate and the distortion due to the error concealment and the error propagation, at the same time the quantization parameter of the MB is selected to meet the bit rate constraint. The Lagrangian optimization method is applied to select jointly the coding mode and the quantization parameter for each MB, in which difference quantization is used. The computational complexity is considerably small, because the bit rate and the distortion estimations are performed using the predetermined values computed during the motion estimation step. The experimental results with various bit rates and packet loss rates show the effectiveness of the proposed algorithm compared with the conventional method. Further research will focus on extension of the proposed algorithm by combining the adaptive buffer control strategy.
REFERENCES


