Proactive Frame-Skipping Decision Scheme For Variable Frame Rate Video Coding

F. Pan, X. Lin, S. Rahardja, K. P. Lim, Z. G. Li, D. J. Wu, S. Wu

Institute for Infocomm Research
21 Heng Mui Keng Terrace, Singapore 119613

Abstract

Many rate control algorithms focus on the adjustment of quantisation values to retain a certain buffer level, and arbitrary frame-skipping is often needed to keep the buffer from overflow at very low bit-rates. In this paper, a content adaptive rate control is proposed to optimise the balance between spatial and temporal quality via active frame-skipping. The occurrence of frame-skipping is jointly dependent on the temporal and spatial quality of the video, and on the fullness of the buffer. This helps to achieve a consistent spatial and temporal quality and to enhance the overall perceptual quality. Experimental results show that the new scheme is simple but very effective, with large average PSNR gains and consistently improved visual quality, and the improvement in perceptual quality is much significant than that in average PSNR’s.

Keywords: Video compression, rate control, variable frame rate, frame skipping

1 Introduction

Digital video coding plays an important role in the development of multimedia applications. The block-based DCT coding approach is widely used in video compression standards such as MPEG and H.26x families. The main objective of a video compression system is to represent a video sequence with as few bits as possible while preserving the level of image detail and quality for the given application. Due to their wide applications, the problem of rate control has been studied extensively. In conventional rate control approaches [1-5], the frame rate is determined a priori based on experience and is independent of the quantisation performed during the coding process. This type of approach works fine with high bit-rate coding. However, at very low bit-rate video coding, it could result in extremely low spatial quality with unnecessary high temporal resolution, or vice versa, and perceptual quality is greatly undermined. Recently, a number of rate control schemes that jointly adjust the frame rate and quantisation value have been proposed [6-8]. A source model is used in [6] to predict R-D characteristics and the frame rate is adjusted to ensure a minimum picture quality of the coded frames. However, in this scheme, the number of frames to be skipped, fsk, is decided only based on the quantization value of a frame without considering the temporal quality, and a number of iterations are needed before fsk is decided, which is not practical in real-time video coding. In [7], the frame rate is reduced when motion is fast and increased when the motion is slow. Ironically, this is against the perception of human vision system (HVS) where high frame rate is needed if the motion is high. Another drawback of this scheme is its slow response to changes in the video and transmission bit-rate, as its frame rate can only be changed from one basic unit to another (a basic unit has 12 frames). Reed et al [8] presented a constrained bit-rate control scheme where the frame rate and quantisation step-size are jointly decided within the constrained regions. This scheme takes advantage of significant delay and is designed for off-line encoding where the video is stored in a server prior to transmission. Also, the constraint region is determined a priori based on experience, which is itself not a trivial task.

As frame rate is a statistical parameter of encoded video, it is difficult to adjust the frame rate instantaneously. In reality, frame rate is closely related to the number of skipped frames during coding. In most of the existing rate control algorithms, frame-skipping is activated only when the buffer overflows, and thus their occurrences are in an uncontrolled manner. On the other hand, it is commonly believed that HVS is less sensitive to temporal details and more sensitive to spatial details for low motion videos and vice versa. Therefore, for low motion video, more frames can be skipped without much affecting the temporal perception; and the saved bits can be used to enhance the spatial quality; Similarly, for high motion sequence, fewer or no frames should be dropped to achieve a better temporal quality by compromising the spatial quality.

Based on the above observation, a novel variable frame rate-encoding algorithm by using controlled frame-skipping is presented in this paper. The frame rate is controlled via active frame skipping to optimise the spatial and temporal quality. The occurrence of frame skipping is decided by the frame-skipping factor, which is dependent on the spatial and temporal quality of the video, and the fullness of the buffer. This frame-skipping factor is updated together with the quantization value and buffer level after encoding each frame by taking into consideration of HVS behaviour. Therefore a consistent spatial and temporal quality could be achieved and the overall perceptual quality is enhanced.

The rest of the paper is organized as follows. Section 2 describes the active frame-skipping scheme based on the motion information, the spatial complexity and buffer fullness. Section 3 presents the simulation results, and Section 4 is the conclusions.
\section{Active Frame Skipping Scheme}

The content adaptive frame-skipping algorithm is implemented based on framework of MPEG-4 Q2 rate control scheme. The block diagram of the improved rate control algorithm is shown in Figure 1. One of the main differences between the new algorithm and MPEG-4 Q2 is that, in the estimating of the target bits for a frame, the possible frame skipping due to video contents are taking into consideration. The detailed functions of the extra blocks are described in the following subsections.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Content adaptive frame skipping algorithm}
\end{figure}

In this section, we will design a variable frame rate control algorithm by using controlled frame-skipping technique. Let \(f_s\) be the source frame rate of the video, \(f_E\) be the encoding frame rate. A frame-skipping factor \(F_{FS}\) is defined to be the ratio of encoded frame rate versus source frame rate.

\[ F_{FS} = f_s / f_E \]

where \(f_s\) is the initial value of the frame-skipping factor which is decided by the encoding frame frequency. \(F_{FS}\) is set to be

\[ f_s = f_s / f_E \]

\[ (2) \]

Note that \(f_1(MV_{SLD}), f_2(Q_{SLD})\) and \(f_3(B_{DIFF})\) are the three items of the frame-skipping factor that are dependent on the amount of motion and the spatial quality of the video, and the buffer level respectively. The details of these items will be discussed in the following subsections.

\subsection{Frame-skipping factor by motion information}

There are many ways to estimate the amount of motion contained in a video sequence, such as the difference of histogram (DOH), the block histogram difference (BH), the histogram of difference image (HOD), block variance difference (BV) and motion compensation error (MCE) \[9\]. All of these measures need extra calculation.

In this paper, a simple yet very convenient measure for motion information, the sliding window average motion vector, \(MV_{SLD}\) is used. In order to quantify the amount of motion, all the motion vectors of the previously coded P/B frames are obtained. This consists of \(X\) and \(Y\) components for each of the macroblocks in the video frame. Then, all the motion vectors are summed up and averaged by the total number of motion vectors. Note that for those macroblocks that are intra-coded, a maximum motion vector that equals to the size of search region is used. After obtaining the averaged motion vector, sliding-window based average is further used to enhance the robustness of the resulting measure such that it is not too sensitive to noises. The size of sliding-window determines how fast the system responses to a motion change. A large window size means slow responses. In this experiment, a size of 2 is used to cater for faster responses.

Therefore, the first item in Equation (1) is decided by,

\[ f_1(MV_{SLD}) = \alpha \times MV_{SLD} + \beta \]

where \(\alpha\) and \(\beta\) are two parameters that are determined empirically.

\subsection{Frame-skipping factor by spatial quality}

For difficult video sequences with high motion and complicated spatial detail, a good balance between spatial and temporal quality has to be achieved in order to produce perceptually good quality video. However, as mentioned previously, if the video contains high motion, the frame rate should be increased so that good temporal quality is retained. This means that, at very low bit-rate, the quantization value could be very large and spatial quality is greatly compromised, resulting in severely deteriorated spatial quality. To overcome the problem, frame-skipping factor need to be adjusted based on the spatial quality of encoded frames.

There exist many perceptual metrics to quantify the quality of the encoded frame \[10\]. However, the quantization value \(Q_P\) is also a very good indication of how good the spatial quality a frame is coded, and \(Q_P\) is conveniently available. Similar to the previous subsection, we use \(Q_{SLD}\), the sliding window average of previous \(Q_P\)’s to improve the robustness of the scheme.

Based on the above observation and experimental results, \(f_2(Q_{SLD})\) is chosen to be a piecewise linear function as follows,

\[ f_2(Q_{SLD}) = \begin{cases} a_1 \times Q_{SLD} + b_1 & \text{if } Q_{SLD} \leq 6 \\ a_2 \times Q_{SLD} + b_2 & \text{if } Q_{SLD} \geq 20 \\ a_3 \times Q_{SLD} + b_3 & \text{elsewhere} \end{cases} \]

where \(a_1, a_2, a_3, b_1, b_2, b_3\) are the parameters that are determined empirically.
Similar to conventional rate control algorithm, forced frame-skipping will be used in case of buffer overflow; and forced frame-coding will be needed in case of buffer underflow. Additional adjustment of frame-skipping factor based on the difference between actual buffer level and target buffer level is also used. This is implemented using the following functions,

\[ f_3(B_{\text{DIFF}}) = \frac{B_{\text{DIFF}}}{B_{\text{BUF}}} = \frac{B_{\text{ACT}} - B_{\text{TGT}}}{B_{\text{BUF}}} \]  

(5)

where \( B_{\text{ACT}} \) is the actual buffer fullness, \( B_{\text{TGT}} \) is the target buffer level, \( B_{\text{DIFF}} \) is the difference, and \( B_{\text{BUF}} \) is the buffer size. This equation shows that, if the current buffer level is above the target buffer level, \( f_3 \) will be adjusted upwards, and if the current buffer level is lower than the target buffer level, \( f_3 \) will be decreased.

Note that similar as in [5], the target buffer level in this scheme is not constant, but is set to be adaptive to the position of this frame in the GOP, i.e., the target buffer level is reduced gracefully from the beginning of a GOP to the end of GOP. This helps to reduce the number of forced frame-skipping caused by the encoding of an I-frame whose bit-rate is much higher than that of the average P/B-frames.

### 3 SIMULATION RESULTS

Numerous experiments have been conducted to evaluate the performance of the new adaptive frame-rate control algorithm at various bit-rates. Standard video sequences such as ‘Akiyo’, ‘Carphone’, ‘Children2’, ‘Coastguard’, ‘Foreman’, ‘Trevor’, ‘Miss America’, ‘Weather’, etc. are tested, and MoMuSys-FDIS-V1.0-990812 codec was used as the test-bed. Note that in our algorithm, the parameter in Equation (3) and Equation (4) are decided empirically as, \( \alpha = -3, \beta = 3.3, a_1 = 0.6, b_1 = -4, a_2 = 0.6, b_2 = -13, a_3 = 0.067, b_3 = -0.8 \) respectively.

Table 1 shows the performance comparisons between the new scheme and MPEG-4 Q2 rate control algorithm at 32 kbps. Note that the video format in this experiment is QCIF, and the GOP size is set to 50. From the Table we can see that, for low motion video sequence such as ‘Akiyo’, we have an average gain in PSNR up to 1.06 dB. This is due to the fact that we code it at much lower frame, and the saved bits are used to enhance the spatial quality without much loss of temporal resolution. This is proved by the subject test on the reconstructed video. For high motion video such as ‘Carphone’, the frame rate is also reduced slightly, which is reflected in the small gain in PSNR of 0.23 dB. However, as our scheme has well controlled frame-skipping such that the occurrence of frame-skipping is uniformly distributed across the whole sequence, the improvement in perceptual quality is much significant than that in average PSNR’s. Similar observations can be found from the other two sequences listed in Table 1.

**Table 1. Performance Comparisons at 32 Kbps**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Scheme</th>
<th>PSNR (dB)</th>
<th>Frame-rate (f/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akiyo</td>
<td>Q2</td>
<td>34.97</td>
<td>22.40</td>
</tr>
<tr>
<td>New</td>
<td></td>
<td>36.03</td>
<td>11.70</td>
</tr>
<tr>
<td>Carphone</td>
<td>Q2</td>
<td>28.84</td>
<td>20.26</td>
</tr>
<tr>
<td>New</td>
<td></td>
<td>29.07</td>
<td>15.63</td>
</tr>
<tr>
<td>Foreman</td>
<td>Q2</td>
<td>26.36</td>
<td>14.25</td>
</tr>
<tr>
<td>New</td>
<td></td>
<td>26.71</td>
<td>12.4</td>
</tr>
<tr>
<td>Weather</td>
<td>Q2</td>
<td>24.56</td>
<td>11.3</td>
</tr>
<tr>
<td>New</td>
<td></td>
<td>24.73</td>
<td>8.9</td>
</tr>
</tbody>
</table>

**Figure 3. PSNR comparison of sequence ‘Akiyo’**

**Figure 4. PSNR comparison of sequence ‘Carphone’**
Figure 5 and Figure 4 show the PSNR comparison of the sequences ‘Akiyo’ and ‘Carphone’ by using the two schemes. Note that in the PSNR calculation, frames that are skipped will be replaced by the nearest previous encoded frames as suggested by Corbera [3]. These two figures show that the new scheme has consistently better PSNR’s. The regular sawtooth structure in the PSNR shows the well-controlled skipped frame across the sequence. It should point out that, because of the uniform distribution of the skipped frames in our new scheme, the PSNR could be increased substantially by using simple frame-interpolation schemes to replace the skipped frame. Figure 5 and Figure 6 show the reconstructed frames of the video sequences ‘Akiyo’ and ‘Carphone’ by using both the MPEG-4 Q2 rate control algorithm and the scheme presented in this paper. Obviously the new scheme outperforms Q2 scheme significantly.

4 CONCLUSIONS

This paper has presented a new content adaptive rate control algorithm. This algorithm uses active frame-skipping such that the occurrence of frame skipping is planned and well controlled based on the contents of the video sequence, the buffer information as well as the HVS’s perceptual behaviours. Experimental results show that this algorithm can achieve a good balance between spatial and temporal quality for very low bit-rate video coding with significant average PSNR gains and consistently improved visual quality, and the improvement in perceptual quality is much significant than that in average PSNR’s.

5 REFERENCES