A Novel Macroblock Level Rate Control Method for Stereo Video Coding

Gaofeng Zhu, 1 Mei Yu, 1 Gangyi Jiang, 1 Zongju Peng, 1 Feng Shao, 1 Fen Chen, 1 and Yo-Sung Ho 2

1 Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211, China
2 Department of Information and Communication, Gwangju Institute of Science and Technology, Gwangju 500-712, Republic of Korea

Correspondence should be addressed to Gangyi Jiang; jianggangyi@126.com

Received 30 August 2013; Accepted 2 January 2014; Published 4 March 2014

Copyright © 2014 Gaofeng Zhu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To compress stereo video effectively, this paper proposes a novel macroblock (MB) level rate control method based on binocular perception. A binocular just-notification difference (BJND) model based on the parallax matching is first used to describe binocular perception. Then, the proposed rate control method is performed in stereo video coding with four levels, namely, view level, group-of-pictures (GOP) level, frame level, and MB level. In the view level, different proportions of bitrates are allocated for the left and right views of stereo video according to the prestatistical rate allocation proportion. In the GOP level, the total number of bitrates allocated to each GOP is computed and the initial quantization parameter of each GOP is set. In the frame level, the target bits allocated to each frame are computed. In the MB level, visual perception factor, which is measured by the BJND value of MB, is used to adjust the MB level bit allocation, so that the rate control results in line with the human visual characteristics. Experimental results show that the proposed method can control the bitrate more accurately and get better subjective quality of stereo video, compared with other methods.

1. Introduction

Three-dimensional (3D) videos, including stereo video multiview video and have currently been coming to the home through various approaches, such as Blu-Ray disc, cable and satellite transmission, terrestrial broadcast, and streaming and download through Internet [1]. Moreover, 3D video technologies have been gradually matured to be moved into mobile platforms, for example, 3D mobile phone and mobile 3D television [2]. However, there are many crucial technologies that need further research for stereo video coding. Rate control is a key issue in stereo/multiview video coding and transmission. Lu et al. [3] proposed a rate control scheme for multiview video coding (MVC), on the basis of the new quantitative measure for stereo video quality. And the macroblock (MB) quantization parameters (QPs) are modified according to the QP of the neighboring MB on the purpose of eliminating block effect in stereo video. Shao et al. proposed a rate control method for asymmetric stereo video coding [4]. They use a fixed threshold to quantize binocular psychovisual redundancy and establish the relationship between distortion and quantization for asymmetric stereo video coding. Zheng et al. [5] presented a rate control algorithm for MVC by analyzing the rate allocation proportion among different types of views, and the frame complexity was used to regulate the target bit for each frame. Liu et al. [6] presented a rate control technique for multiview video plus depth (MVD) based 3D video coding, and an image-stitching method was utilized to simultaneously encode video and depth. While these rate control algorithms in stereo/multiview video coding made useful explorations, however, these algorithms were only designed to frame layer bit allocation. In order to more finely rate control, it is necessary to develop the rate control techniques on the MB layer.

Moreover, the above algorithms did not take into account human visual characteristics well. Due to physiological and
psychological mechanisms of human visual system (HVS), human eyes do not detect all of image or video distortion [7]. Thus, the concept of just-noticeable difference (JND) has played an important role in understanding HVS [8, 9]. JND implies a visibility threshold in which human eyes can perceive changes in image values and depends on the luminance and contrast of local image region. Liu et al. [10] utilized a JND model to separate edge and textured regions of the image. Recently, several studies had concentrated on constructing visibility threshold for 3D image/video. Based on modeling and incorporating the binocular combination of injected noises, luminance masking, and contrast masking, binocular JND (BJND) was reported as the first model to measure the perceptible distortion of stereo images [11]. It was empirically demonstrated that human cannot realize distortion in stereo image if the distortion in one view image is less than the BJND value.

In this paper, we propose a novel MB level rate control method for stereo video coding based on visual perception. The proposed rate control method is performed in stereo video coding with four levels, namely, view level, group-of-pictures (GOP) level, frame level, and MB level. The BJND model based on the parallax matching is first exploited. Then, visual perception factor, designed based on BJND, is used to adjust the MB level bit allocation. So the rate control results in line with human visual characteristics. The rest of this paper is organized as follows. Section 2 describes the BJND model and Section 3 describes the proposed method. Section 4 shows experimental results, and finally, a conclusion is given in Section 5.

2. The BJND Model

To account for the mutual effect on the minimum distortion in one view that evokes perceptual differences in stereo image, the BJND model was proposed by Zhao et al. and measured with a set of psychophysical stimuli [11]. It has been demonstrated that the BJND is binocular combination of injected noises, luminance masking, and contrast masking in binocular viewing.

In [11], Zhao et al. did not take into account the disparity between left and right views of stereo image in their experiments, and they obtained BJND values under assumption about that the disparities between left and right views were zero. This assumption is not reasonable since in most cases the disparity of stereo image is not zero. In practice, binocular disparity is attributed to the primary visual cortical areas, and it is one of the most important stereo cues. Stereo matching itself has been a difficult task in computer vision, and a comprehensive review of stereo matching is given in [12]. Here, a state-of-the-art stereo matching technique in [13] is adopted.

To improve Zhao’s model, a modified BJND model is presented with the consideration of binocular disparity. Given the left and right views of stereo image with the disparity image corresponding to left view, by incorporating the models of the binocular combination of injected noises, luminance, and contrast masking effects, a BJND at the right view is defined as follows:

\[
BJND_r(i, j, d) = BJND_r(bg_l(i + d, j), eh_l(i + d, j), n_l(i + d, j))
\]

\[
= T_C(bg_l(i + d, j), eh_l(i + d, j)) \times \left( 1 - \left( \frac{n_l(i + d, j)}{T_C(bg_l(i + d, j), eh_l(i + d, j))} \right)^{1/\lambda} \right),
\]

where \(i\) and \(j\) are the pixel coordinates, \(d\) is the disparity at the position of \((i, j)\) in right view, and \(T_C\) is a contrast masking threshold. Note that BJND, is dependent on the background luminance level \(bg_l\), the edge height \(eh_l\), and the noise amplitude \(n_l\) of the corresponding pixel position in the left view. Here, the original sequence is used, so there is no noise in the left view; that is, \(n_l = 0\).

\[T_C\] is defined by

\[
T_C(bg_l, eh_l) = A_{lim}(bg_l) + K(bg_l) \cdot eh_l,
\]

where \(A_{lim}()\) is a luminance masking function and \(K()\) is a fitting function.

\(A_{lim}(bg_l)\) and \(bg_l\) are calculated in

\[
A_{lim}(bg_l) = \begin{cases} 
0.0027 \cdot (bg_l^2 - 96 \cdot bg_l) + 8, & \text{if } 0 \leq bg_l < 48, \\
0.0001 \cdot (bg_l^2 - 32 \cdot bg_l) + 1.7, & \text{if } 48 \leq bg_l \leq 255,
\end{cases}
\]

\[
bg_l(x, y) = \frac{1}{32} \sum_{i=-5}^{5} \sum_{j=0}^{5} f(x - 3 + i, y - 3 + j) \cdot B(i, j).
\]

The fitting function \(K(bg_l)\) of \(bg_l\) is defined as

\[
K(bg_l) = -10^{-6} \cdot (0.7 \cdot bg_l^2 + 32 \cdot bg_l) + 0.07,
\]

\[bg_l \in [0, 255].\]

The edge height \(eh_l\) is given by

\[
eh_l(i, j) = \sqrt{E_H^2(i, j) + E_V^2(i, j)},
\]
3. The Proposed BJND-Based Rate Control Method

Here, we present a new rate control method for the stereo video based on the modified BJND. The framework of the proposed method is illustrated in Figure 1, with rate control in four levels, including view level, group-of-pictures (GOP) level, frame level, and MB level.

3.1. View Level Rate Control. Hierarchical B Pictures (HBP) coding structure in stereo video is shown in Figure 2. There are two views to be encoded. It is seen that there is no inter-view predication in left view, and right view is encoded with unidirectional interview prediction from the reconstructed left view.

In [14], we obtained the proportion between left view and right view according to preencoding the first GOP. The target bit rate for each view is computed by

$$T_{\text{view}}(k) = T_{\text{total}} \times w(k),$$

where $T_{\text{total}}$ denotes the total bit rate of left and right views. $w(k)$ is bit allocation proportion for each view. $k$ is the coding order index of view, and the value of $k$ is 0 or 1, which represents the left view and right view, respectively; $T_{\text{view}}(k)$ is target bit rate for the $k$th view.

3.2. GOP Level Rate Control. Within the left view or the right view, the total number of bits allocated to each GOP is computed and the initial QP of each GOP is set in GOP level.

where $E_k(i, j)$ is $5 \times 5$ Sobel operator and the formula is as follows:

$$E_k(i, j) = \frac{1}{24} \sum_{h=1}^{5} \sum_{v=1}^{5} p(i - 3 + h, j - 3 + v) \ast G_k(h, v),$$

$$k = H, V,$$

$$G_H = \begin{bmatrix}
-1 & -2 & 0 & 2 & 1 \\
-2 & -3 & 0 & 3 & 2 \\
-3 & -5 & 0 & 5 & 3 \\
-2 & -3 & 0 & 3 & 2 \\
-1 & -2 & 0 & 2 & 1 
\end{bmatrix},$$

$$G_V = \begin{bmatrix}
1 & 2 & 3 & 2 & 1 \\
2 & 3 & 5 & 3 & 2 \\
0 & 0 & 0 & 0 & 0 \\
-2 & -3 & -5 & -3 & -2 \\
-1 & -2 & -3 & -2 & -1 
\end{bmatrix}.$$
level rate control. The total number of bits allocated for the ith GOP is computed by
\[ T_r (n_{i,0}) = \frac{T_{\text{view}}(k)}{F_r} \times N_{\text{gop}} - \left( \frac{B_i}{8} - B_C (n_{i-1, N_{\text{gop}}}) \right), \] (9)
where \( F_r \) is frame rate, \( N_{\text{gop}} \) denotes total number of frames in the current GOP, and \( B_C (n_{i-1, N_{\text{gop}}}) \) is actual buffer occupancy after encoding the \((i-1)\)th GOP. The buffer occupancy should be kept at \( B_i/8 \) after encoding each GOP. In the case of the constant bit rate (CBR), \( T_r \) is updated frame by frame as follows:
\[ T_r (n_{i,j}) = T_r (n_{i,j-1}) - A (n_{i,j-1}), \] (10)
where \( A(n_{i,j-1}) \) is the actual encoding bits for the \((j-1)\)th frame.

The initial QP of the first GOP is a predefined QP, denoted as \( Q_{0p} \). As shown in Figure 2, 10 and BI frames in the first GOP of the left view and P0 and BI frames in the first GOP of the right view are encoded by \( Q_{0p} \). The first frame QP in other GOP is computed by
\[ QP_{st} = \frac{\sum B_{\text{BQP}}}{N_b} - 1 - \frac{8T_r (n_{i-1, N_{\text{gop}}})}{T_r (n_{i,0})} - \frac{N_{\text{gop}}}{15}, \] (11)
where \( N_b \) is total number of bits in encoding B frame in a GOP and \( \sum B_{\text{BQP}} \) is the sum of QPs for all B frames in the previous GOP.

3.3. Frame Level Rate Control. Within each GOP, the target bits for each B frame are calculated according to the target buffer fullness level and the number of remaining bits in frame level rate control.

The actual buffer occupancy, \( T_{\text{buf}} (n_{i,j}) \), is computed as
\[ T_{\text{buf}} (n_{i,j}) = \frac{T_{\text{view}}(k)}{F_r} + \frac{\gamma (T_{\text{b}} (n_{i,j-1}) - T_{\text{b}} (n_{i,j}) - B_C (n_{i,j}))}{N_b - 1}, \] (12)
where \( \gamma \) is a constant and its typical value is 0.75 [15]. \( T_{\text{b}} (n_{i,j-1}) \) is the target buffer level of the \( (j-1) \)th frame in the \( i \)th GOP, \( T_{\text{b}} (n_{i,j}) \) is the initial value of target buffer level, \( B_C (n_{i,j}) \) is the actual buffer occupancy after coding the \( j \)th B frame in the \( i \)th GOP.

Meanwhile, the number of remaining bits should also be considered when the target bit is computed.
\[ T_{\text{rem}} (n_{i,j}) = \frac{T_r (n_{i,j})}{N_b (j-1)}, \] (13)
where \( N_b (j-1) \) is the number of remaining B frames in the \( i \)th GOP.

The target bits, \( T (n_{i,j}) \), are a weighted combination of \( T_{\text{buf}} (n_{i,j}) \) and \( T_{\text{rem}} (n_{i,j}) \) and represented by
\[ T (n_{i,j}) = \beta \times T_{\text{rem}} (n_{i,j}) + (1 - \beta) \times T_{\text{buf}} (n_{i,j}), \] (14)
where \( \beta \) is a constant and its typical value is 0.5 [15].

3.4. MB Level Rate Control. Let \( T_{\text{mb}} (n_{i,j}) \) and \( N \) denote the number of remaining bits for all non-coded macroblock in the current frame and the number of MBs, respectively. Let \( T_{\text{mb}} (i, j, k) \) and \( \sum p=1 \text{MAD}^2 (i, j, k) \) denote target bit rate and MAD value in the \( k \)th MB, respectively. Consider
\[ T_{\text{mb}} (i, j, k) = T_{\text{rb}} (n_{i,j}) \times \frac{\text{MAD}^2 (i, j, k)}{\sum p=1 \text{MAD}^2 (i, j, p)} . \] (15)

With the current MB's BJND value and the average BJND value of removing the left, right, upper, and lower boundary MBs, the tolerable distortion degree, \( \mu(i, j, k) \), of the current MB in the whole frame is measured as follows:
\[ \mu (i, j, k) = \frac{\text{BJND} (i, j, k)}{\sum_{u=2}^{X-1} \sum_{v=2}^{Y-1} \text{BJND} (i, j, u, v)} . \] (16)

where \( X \) and \( Y \) denote the number of MBs in a row and a column for each frame, respectively. The fluctuation of \( \mu(i, j, k) \) is large, thus, it is normalized and then added with 0.5 so as to get the perception factor weighting \( \omega(i, j, k) \), which is done by
\[ \omega (i, j, k) = \frac{\mu (i, j, k) - \mu_{\text{min}}}{\mu_{\text{max}} - \mu_{\text{min}}} + 0.5, \]
\[ \mu_{\text{min}} = \min \left\{ \sum_{u=2}^{X-1} \sum_{v=2}^{Y-1} \text{BJND} (i, j, u, v) \right\} , \] (17)
\[ \mu_{\text{max}} = \max \left\{ \sum_{u=2}^{X-1} \sum_{v=2}^{Y-1} \text{BJND} (i, j, u, v) \right\} , \]
where \( \min \) indicates the minimum function and \( \max \) indicates the maximum function.

The perception factor \( \omega(i, j, k) \) is used to adjust the target bit allocation for the MB level. The larger BJND value represents more sensitive to distortion for human eye; that is, the value of \( \omega(i, j, k) \) is larger, so the MB should be allocated more bits. Conversely, the MB with smaller BJND value should be allocated smaller bits. Consider
\[ T_{\text{mb}} (i, j, k) = T_{\text{rb}} (n_{i,j}) \times \omega (i, j, k) \times \frac{\sum \text{MAD}^2 (i, j, k)}{\sum p=1 \text{MAD}^2 (i, j, p)} . \] (18)

Based on the allocated target bits, the quantization step size can be computed by the quadratic R-Q model:
\[ T_{\text{mb}} (i, j, k) = \left( \frac{X_1}{Q_{\text{step}}} + \frac{X_2}{Q_{\text{step}}} \right) \times \text{MAD} (i, j, k) . \] (19)
The left view (S9T0)

(b) The right view (S10T0)

(c) The BJND map of the right view in pixel domain

(d) The BJND mask on the basis of (c) by MB processing

**Figure 3:** The BJND of the BookArrival sequence.

where $Q_{\text{step}}$ is the quantization step size of the current MB. $X_1$ and $X_2$ are the model parameters, which need to be updated using a linear regressive technique.

Furthermore, QP of the current MB can be computed by

$$ QP = 6 \times \log_2(Q_{\text{step}}) + 4. $$

(20)

### 4. Experimental Results and Discussions

In order to evaluate the performance of the proposed MB level rate control method for stereo video coding based on visual perception, six representative stereo video sequences with $1024 \times 768$ spatial resolution, including BookArrival, AltMoabit, DoorFlowers, LeavingLaptop, Newspaper, and Kendo, are used in the experiments. BookArrival, AltMoabit, DoorFlowers, and LeavingLaptop are provided by Fraunhofer HHI [16]. Newspaper is provided by Gwangju Institute of Science and Technology [17]. Kendo sequence is captured by Nagoya University with the moved camera array [18].

In the experiments, we use the revised MVC software JMVC7.0 to implement the rate control methods. The test conditions of the six sequences are shown in Table 1. Two middle views (view 9 and view 10) for BookArrival, DoorFlowers, and LeavingLaptop, two middle views (view 5 and view 6) for AltMoabit and Newspaper, the two middle views (view 4 and view 5) for Kendo are used to simulate the left and right views of stereo video in Figure 2.

<table>
<thead>
<tr>
<th>Table 1: Test conditions.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame rate</td>
</tr>
<tr>
<td>GOP length</td>
</tr>
<tr>
<td>Search range</td>
</tr>
<tr>
<td>Reference frames</td>
</tr>
</tbody>
</table>

In order to avoid the influence of BJND results caused by occlusion and exposure in the left and right views of stereo video, the left, right, upper, and lower MBs in each frame of the right view sequence are not processed. And the BJND results of BookArrival and AltMoabit are given in this paper, as shown in Figures 3 and 4, respectively. Figures 3(a) and 3(b) show the original left and the right view images. Figure 3(c) is the BJND map of the right view in pixel domain, and Figure 3(d) is the BJND mask on the basis of Figure 3(c) by MB processing. In the same way, Figure 4 shows the similar results.

#### 4.1. Rate Control Accuracy in Stereo Video Coding

Under the test conditions shown in Table 1, three methods, Zheng’s in [5], SMBRC, and the proposed method, are compared. SMBRC denotes the stereo video MB level rate control method; that is, the MB level monoscopic video JVT-G012 algorithm is extended to stereo video coding and implemented in JMVC.
The detailed average control accuracies of three methods are shown in Table 2. In Table 2, the target bitrate and the actually controlled bitrate are the average value of two views. Rate control error ratio, $E_{RC}$, is used to measure the accuracy of the bitrate estimation and defined by

$$E_{RC} = \frac{|R_{actual} - R_{target}|}{R_{target}} \times 100\%,$$

where $R_{target}$ and $R_{actual}$ denote the target bitrate and the actual coding bitrate, respectively.

Zheng's method is a frame-level rate control method, while the proposed method and SMBRC are MB level rate control scheme, so they have better control accuracy. This can be evidently seen from Table 2. Table 2 also indicates that the absolute inaccuracy of the proposed stereo video rate control method is within 0.192%. It is obvious that the proposed method can more precisely control the bitrate in stereo video coding.

4.2. Rate Distortion Performance Comparison. Figure 5 shows the rate-distortion (R-D) performance comparison results of the three methods. In Figure 5, the average bitrate is the average value of two views, and the peak signal-to-noise ratio (PSNR) value is also the average value of two views. It is clear that, compared with Zheng's and SMBRC methods, the proposed method can achieve the better R-D performance under different target bitrates since more accurate binocular perception information is used in the proposed method.

4.3. Subjective Quality Comparison. Different sequences possess different motion and scene characteristics; therefore the video frames which can best reflect the differences in the subjective quality are selected in accordance with the characteristics of each sequence. For the outdoor scene AltMoabit sequence, we select the 52th frame in view 6 as an example, shown in Figure 6. The bus and the words on the bus in the image are the areas of human visual attention in Figure 6, and more bits are allocated for these areas under the visual perception factor based on BJND. Therefore the subjective quality of the proposed method is better than the other methods.

In order to eliminate the limitation of only one video frame, we select the 43rd frame to 47th frame for the indoor scene LeavingLaptop sequence, shown in Figure 7. The details of the human body and head and the book are allocated more bitrates, so the subjective quality of the proposed method is better than the other methods. Moreover, for quantization parameter selection, if we select a small value $QP$, the subjective quality of the three methods is too good to reflect the difference. Hence, $QP$ is set to 37.
Figure 5: The R-D performance comparisons of three methods.
### Table 2: The detailed control accuracies of the three schemes.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Target bitrate (kbps)</th>
<th>Actual bitrate (Kbps)</th>
<th>${E_{RC}}$ (%)</th>
<th>Zheng</th>
<th>SMBRC</th>
<th>Proposed</th>
<th>Zheng</th>
<th>SMBRC</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>BookArrival</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2454.588</td>
<td>2461.882</td>
<td>2457.641</td>
<td>2456.328</td>
<td>0.297</td>
<td>0.124</td>
<td>0.071</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1087.928</td>
<td>1091.650</td>
<td>1089.203</td>
<td>1088.316</td>
<td>0.342</td>
<td>0.117</td>
<td>0.036</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>569.213</td>
<td>571.351</td>
<td>571.052</td>
<td>569.973</td>
<td>0.376</td>
<td>0.323</td>
<td>0.133</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>328.962</td>
<td>329.536</td>
<td>329.185</td>
<td>329.137</td>
<td>0.174</td>
<td>0.068</td>
<td>0.053</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AltMoabit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2898.506</td>
<td>2895.074</td>
<td>2902.553</td>
<td>2901.796</td>
<td>0.118</td>
<td>0.140</td>
<td>0.113</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1482.783</td>
<td>1481.250</td>
<td>1485.130</td>
<td>1482.949</td>
<td>0.103</td>
<td>0.158</td>
<td>0.011</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>766.879</td>
<td>771.416</td>
<td>768.551</td>
<td>767.346</td>
<td>0.592</td>
<td>0.218</td>
<td>0.061</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>422.260</td>
<td>425.840</td>
<td>424.130</td>
<td>422.078</td>
<td>0.848</td>
<td>0.443</td>
<td>0.043</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Newspaper</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2624.359</td>
<td>2629.377</td>
<td>2630.822</td>
<td>2626.997</td>
<td>0.191</td>
<td>0.246</td>
<td>0.101</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1363.146</td>
<td>1365.172</td>
<td>1366.715</td>
<td>1365.219</td>
<td>0.149</td>
<td>0.262</td>
<td>0.152</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>700.768</td>
<td>703.104</td>
<td>702.132</td>
<td>702.384</td>
<td>0.431</td>
<td>0.333</td>
<td>0.195</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>387.959</td>
<td>389.030</td>
<td>388.711</td>
<td>388.285</td>
<td>0.276</td>
<td>0.194</td>
<td>0.084</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kendo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1664.009</td>
<td>1669.532</td>
<td>1673.897</td>
<td>1666.769</td>
<td>0.332</td>
<td>0.594</td>
<td>0.166</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>889.142</td>
<td>894.546</td>
<td>895.593</td>
<td>887.494</td>
<td>0.608</td>
<td>0.726</td>
<td>0.185</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>518.328</td>
<td>521.145</td>
<td>515.668</td>
<td>520.191</td>
<td>0.543</td>
<td>0.513</td>
<td>0.359</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>323.106</td>
<td>325.386</td>
<td>320.659</td>
<td>325.120</td>
<td>0.706</td>
<td>0.757</td>
<td>0.623</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DoorFlowers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2081.433</td>
<td>2088.817</td>
<td>2085.565</td>
<td>2082.963</td>
<td>0.355</td>
<td>0.199</td>
<td>0.073</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>878.540</td>
<td>882.652</td>
<td>879.054</td>
<td>877.835</td>
<td>0.468</td>
<td>0.059</td>
<td>0.080</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>448.642</td>
<td>448.616</td>
<td>448.214</td>
<td>444.370</td>
<td>0.006</td>
<td>0.095</td>
<td>0.952</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>259.811</td>
<td>259.945</td>
<td>259.116</td>
<td>258.028</td>
<td>0.051</td>
<td>0.268</td>
<td>0.686</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LeavingLaptop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2343.246</td>
<td>2347.945</td>
<td>2347.679</td>
<td>2344.263</td>
<td>0.201</td>
<td>0.189</td>
<td>0.043</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1029.631</td>
<td>1023.200</td>
<td>1031.212</td>
<td>1027.668</td>
<td>0.625</td>
<td>0.154</td>
<td>0.191</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>534.344</td>
<td>531.283</td>
<td>533.795</td>
<td>533.878</td>
<td>0.573</td>
<td>0.103</td>
<td>0.087</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>307.944</td>
<td>304.321</td>
<td>307.543</td>
<td>308.274</td>
<td>1.176</td>
<td>0.130</td>
<td>0.107</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.398</td>
<td>0.267</td>
<td>0.192</td>
</tr>
</tbody>
</table>

Figure 6: The subjective image comparison among three methods of AltMoabit sequence (QP = 37, the 52th frame in the view 6).
Figure 7: Continued.
5. Conclusion

A novel macroblock (MB) level rate control method for stereo video is proposed in this paper. The proposed method is performed with four levels, namely, view level, group-of-picture (GOP) level, frame level, and MB level. In the view level, different proportions of bitrates are allocated for left view and right view of stereo video according to the prestatistical rate allocation proportion. In the GOP level, the total number of bitrates allocated to each GOP is computed and the initial quantization parameter of each GOP is set. In the frame level, the target bits allocated to each frame are computed. In the MB level, the binocular just-noticeable difference (BJND) model based on the parallax matching is first given. Then, visual perception factor, which is measured by the ratio between the BJND value of the current MB and the average BJND value of removing the left, right, upper, and lower boundary MBs, is used to adjust the MB level bit allocation. Experimental results show that the proposed method can control the bitrate more accurately and the average of rate control error is only about 0.19%. Compared with other schemes, the proposed method gets better rate-distortion performance and subjective quality of stereo images.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the Natural Science Foundation of China under Grant nos. U1301257, 61271270, 61311140262, 61171163, and 61271021. It was also sponsored by K. C. Wong Magna Fund in Ningbo University.

References


Submit your manuscripts at
http://www.hindawi.com