A Rate Control Scheme for HEVC Intra Coding Using Convolution Neural Network (CNN)

Xin Lu, Bixing Zhou, Xuesong Jin, Graham Martin
Harbin Institute of Technology, China, Harbin University of Commerce, China, University of Warwick, UK

Abstract

A Convolutional Neural Network (CNN)-based rate control scheme for HEVC intra-coding is described. 1. An improved QP determination model (R-QP model) is suggested and a CNN is used to predict the model coefficients; 2. A CNN-based framework is employed to predict the bit consumption for each CTU.

Objectives

1. Simplify the QP calculation process; 2. Achieve an accurate QP prediction; 3. Realise a reasonable bit budget allocation.

Research Background

A rate control scheme usually consists of two stages: bit allocation and QP value determination. 1. Bit allocation

At the CTU level, the number of target bits for each CTU is determined by:

\[ R_{\text{ctu}}^i = \frac{R_{\text{rem}}}{\sum_{j \in \text{CTU}} \alpha_j} \times \alpha_j \]

where \( R_{\text{rem}} \) is the number of remaining bits in the bit budget; \( \alpha_j \) is weight for bit allocation.

2. QP value determination

The R-Q model is employed to predict the QP value for the to-be-coded CTU.

\[ QP_{\text{ctu}} = 4.2005 \times \ln \lambda_{\text{ctu}} + 13.7122 \]

where \( \lambda_{\text{ctu}} \) is the Lagrange multiplier, and for intra modes it can be expressed by:

\[ \lambda_{\text{ctu}} = \alpha_{\text{ctu}} \left( C_{\text{ctu}} / R_{\text{ctu}} \right) \beta_{\text{ctu}} \]

where \( \alpha_{\text{ctu}} \) and \( \beta_{\text{ctu}} \) denote the model coefficients, \( C_{\text{ctu}} \) indicates the complexity of the to-be-coded CTU, which can be roughly measured by the Sum of Absolutely Transformed Difference (SATD) of the CTU.

CNN-Oriented R-QP Model

1. Proposed R-QP model

As the determination of QP value for the CTU is largely dependent on its texture complexity, we propose an improved QP determination model which only involves two model coefficients, namely \( \mu \) and \( v \).

\[ QP_{\text{ctu}} = \mu \times \ln \lambda_{\text{ctu}} + v \]

where \( \mu = -4.2005 \beta_{\text{ctu}} \) and \( v = 4.2005 \beta_{\text{ctu}} \times \ln \alpha_{\text{ctu}} + 4.2005 \ln \alpha_{\text{ctu}} + 13.7122 \)

The 24th CTU, MTD=0.9955

![Fig. 1: Relationship between QP and R](image)

2. CNN-based prediction of the R-QP model coefficients

We use the images from the UCID dataset and the RAISE dataset to train the CNN. The loss function \( L \) for the CNN is defined as:

\[ L = \frac{1}{N} \sum_{n=1}^{N} \left[ GT(n) - P(n) \right]^2 \]

where \( N \) denotes the number of CTUs involved in the training stage; \( GT \) is the label for each CTU; \( P \) represents the predicted value.

![Fig. 2: The CNN structure for the model parameter prediction](image)

Improved Bit Allocation Scheme

Given a fixed QP value, a larger bit consumption is generally required for a CTU containing complex texture, and vice versa.

A CNN-based framework is used to predict the bit consumption for each CTU according to the entire picture content that it contains. The structure of the CNN is shown in Fig. 2. The proposed R-QP model and the improved bit allocation strategy are combined to form a CNN-based rate control scheme for HEVC.

![Fig. 3: The proposed CNN-based rate control scheme](image)

Experimental Evaluation

In the following tables, the ‘Proposed R-QP scheme’ employs the proposed CNN-based R-QP model, and the ‘Proposed RC scheme’ combines the proposed R-QP model and the improved bit allocation strategy.

<table>
<thead>
<tr>
<th>Class</th>
<th>Sequence</th>
<th>Proposed R-QP scheme</th>
<th>Proposed RC scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Traffic</td>
<td>-0.95</td>
<td>-0.97</td>
</tr>
<tr>
<td>B</td>
<td>ParkScene</td>
<td>-0.93</td>
<td>-1.51</td>
</tr>
<tr>
<td>C</td>
<td>Kimono1</td>
<td>-0.53</td>
<td>-0.68</td>
</tr>
<tr>
<td>D</td>
<td>BQTerrace</td>
<td>-0.85</td>
<td>-1.49</td>
</tr>
<tr>
<td>E</td>
<td>Cactus</td>
<td>-2.11</td>
<td>-3.01</td>
</tr>
<tr>
<td>C</td>
<td>QBMell</td>
<td>-0.72</td>
<td>-1.02</td>
</tr>
<tr>
<td>D</td>
<td>RacingHorse</td>
<td>-0.88</td>
<td>-0.21</td>
</tr>
<tr>
<td>E</td>
<td>BasketballPass</td>
<td>-0.30</td>
<td>-0.26</td>
</tr>
<tr>
<td>F</td>
<td>BQSqure</td>
<td>-0.04</td>
<td>-1.11</td>
</tr>
<tr>
<td>G</td>
<td>SlideShow</td>
<td>-1.91</td>
<td>-2.80</td>
</tr>
<tr>
<td>H</td>
<td>Johnny</td>
<td>-0.58</td>
<td>-0.92</td>
</tr>
<tr>
<td>I</td>
<td>Kristen</td>
<td>-1.18</td>
<td>-1.60</td>
</tr>
<tr>
<td>J</td>
<td>FourPeople</td>
<td>-0.91</td>
<td>-0.92</td>
</tr>
<tr>
<td>K</td>
<td>BasketballDrive</td>
<td>-1.60</td>
<td>-2.17</td>
</tr>
</tbody>
</table>

The results show that 1) given the same bit rate, the proposed RC scheme increases the average PSNR by up to 0.42dB, and 2) given the same video quality, the proposed RC scheme produces an average saving in bit rate of up to 3.01%, compared to the default rate control scheme in HM 16.9.

Conclusion

The proposed R-QP model is effective in describing the rate-distortion relationship in HEVC, and the CNN-based bit allocation scheme is appropriate for ensuring that the best use of each bit is made.

Data Compression Conference (DCC)/ 26th March, 2020/ Snowbird, Utah, USA