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

Xin Lu, Bixing Zhou
Harbin Institute of Technology, China

Xuesong Jin
Harbin University of Commerce, China

Graham Martin
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 $R_{ctu}^i$ is determined by:

$$R_{ctu}^i = \frac{R_{Rem}}{\sum_{j \in YetToBeCoded} \omega_{ctu}^j} \times \omega_{ctu}^i$$

where $R_{Rem}$ is the number of remaining bits in the bit budget; $\omega_{ctu}^i$ is weight for bit allocation.

### 2. QP value determination

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

$$QP_{ctu} = 4.2005 \times ln\lambda_{ctu} + 13.7122$$

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

$$\lambda_{ctu} = \alpha_{ctu} \left(\frac{C_{ctu}}{R_{ctu}}\right)^{\beta_{ctu}}$$

where $\alpha_{ctu}$ and $\beta_{ctu}$ denote the model coefficients, $C_{ctu}$ 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 μ and v.

$$QP_{ctu} = \mu \times lnR_{ctu} + v$$

where $\mu = -4.2005\beta_{ctu}$ and

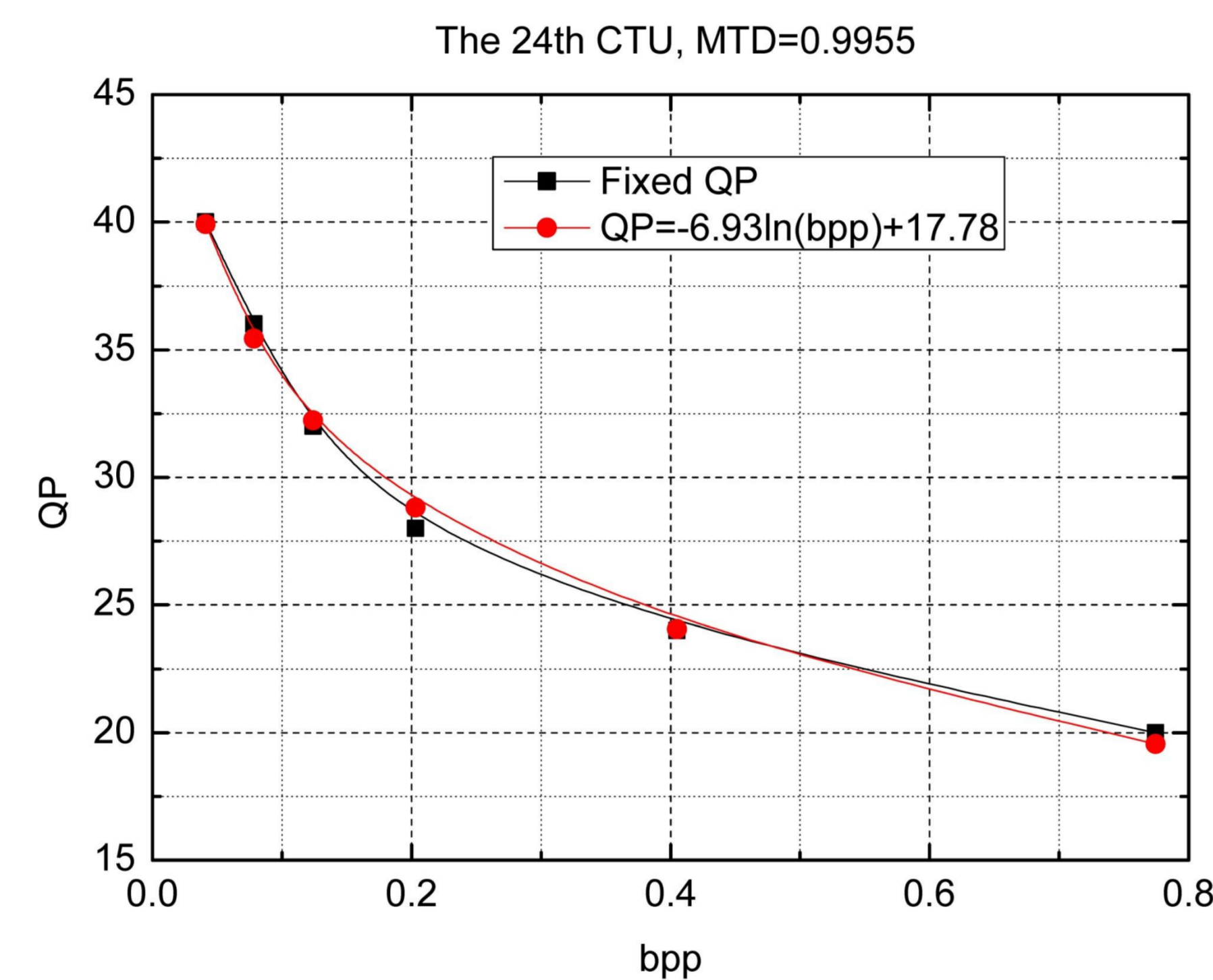$$v = 4.2005\beta_{ctu} \times lnC_{ctu} + 4.2005ln\alpha_{ctu} + 13.7122$$



*Fig.1: Relationship between QP and R*

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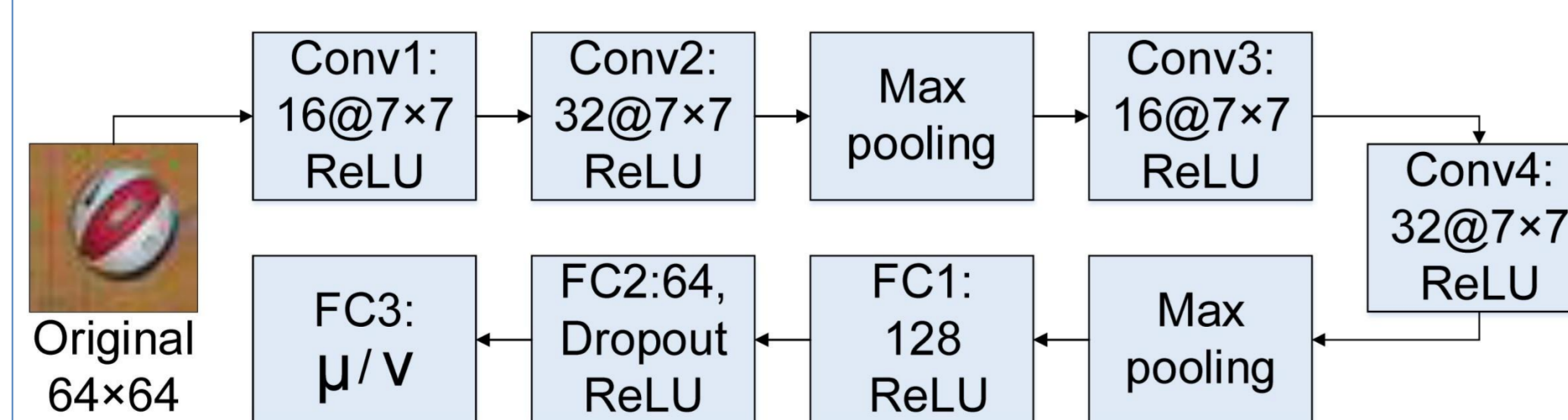


*Fig 2: The CNN structure for the model parameter prediction*

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

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

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

## 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.
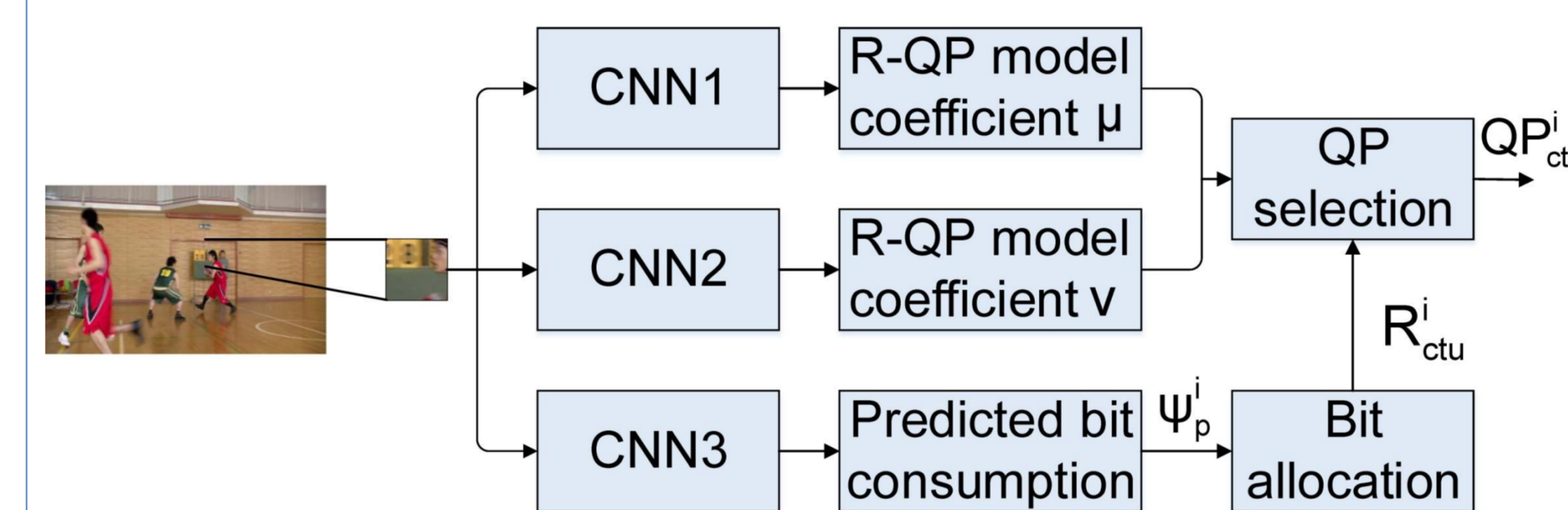


*Fig. 3: The proposed CNN-based rate control scheme*

Considering the picture content in the CTUs, the number of target bits for each CTU is further adjusted as

$$R_{ctu}^i = R_f \times \left(\frac{\omega_{ctu}^i}{\omega_f} + \frac{\psi_p^i}{\Psi_p}\right) \times \frac{1}{2}$$

where $R_{ctu}^i$ is the number of target bits for the *i*th CTU; $R_f$ denotes the overall bit budget for the to-be-coded frame; $\omega_{ctu}^i/\omega_f$ represents the proportion of the bit budget for the *i*th CTU in the overall bit budget for the entire frame. $\psi_p^i$ is the weight of bit allocation for the *i*th CTU, and it is also the predicted bit consumption for the CTU.

$$\Psi_p = \sum_{i=1}^{N} \psi_p^i$$

where *N* is the number of CTUs contained in the to-be-coded frame, $\Psi_p$ denotes the overall predicted bit consumption for the entire frame.

## 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.

| Class | Sequence | Proposed R-QP scheme BDBR(%) | Proposed R-QP scheme BDPSNR(dB) | Proposed RC scheme BDBR(%) | Proposed RC scheme BDPSNR(dB) |
|---|---|---|---|---|---|
| A | Traffic | -0.96 | 0.05 | -0.97 | 0.05 |
| B | ParkScene | -0.93 | 0.04 | -1.51 | 0.07 |
| B | Kimono1 | -0.53 | 0.02 | -0.68 | 0.02 |
| B | BQTerrace | -0.85 | 0.05 | -1.49 | 0.10 |
| B | Cactus | -2.11 | 0.08 | -3.01 | 0.12 |
| C | BQMall | -0.72 | 0.05 | -1.02 | 0.07 |
| C | RaceHorses | -0.88 | 0.06 | -0.21 | 0.01 |
| D | BasketballPass | -0.30 | 0.02 | -0.26 | 0.02 |
| D | BQSquare | -0.64 | 0.06 | -1.11 | 0.09 |
| E | SlideShow | -1.91 | 0.27 | -2.80 | 0.42 |
| E | Johnny | -0.58 | 0.03 | -0.92 | 0.04 |
| E | Kristen | -1.18 | 0.07 | -1.60 | 0.11 |
| E | FourPeople | -0.91 | 0.06 | -0.92 | 0.06 |
| E | BasketballDrive | -1.60 | 0.08 | -2.17 | 0.11 |
| Average | - | -1.01 | 0.07 | -1.33 | 0.09 |

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.

| Class | Average BDBR Li's algorithm | Average BDBR Proposed R-QP scheme | Average BDBR Proposed RC scheme |
|---|---|---|---|
| A | -1.0% | -1.0% | -1.0% |
| B | -1.0% | -1.2% | -1.7% |
| C | -0.2% | -0.8% | -0.6% |
| D | -0.4% | -0.5% | -0.6% |
| E | -1.0% | -1.2% | -1.7% |
| Average | -0.7% | -0.9% | -1.1% |

Compared with Li's algorithm, our schemes achieve better coding efficiency, namely the average BDBR is reduced by 0.2% and 0.4% on average.

## 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.