CODING BLOCK-LEVEL PERCEPTUAL VIDEO CODING FOR 4:4:4 DATA IN HEVC

Lee Prangnell, Miguel Hernández-Cabronero and Victor Sanchez
Department of Computer Science, University of Warwick, Coventry, England, UK

Uniform Reconstruction Quantization (URQ) in HEVC

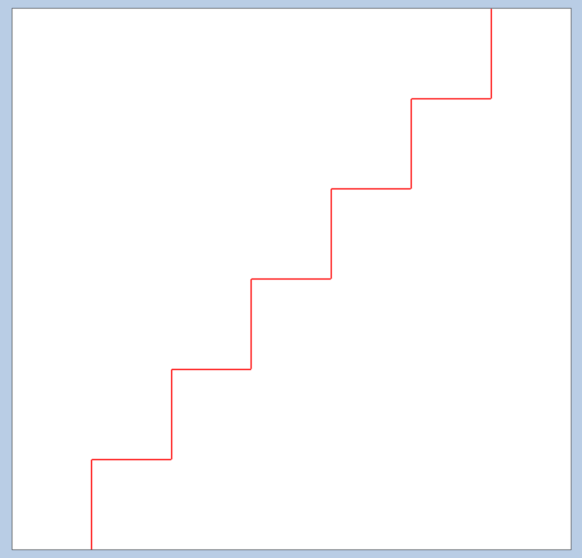

Figure 1: Uniform Quantization Step Size (QStep) of URQ in HEVC. The Quantization Parameter (QP) has a binary logarithmic relationship with the QStep.

- URQ uniformly quantizes transform coefficients based on a QStep/QP (see Figure 1) without taking into account the perceptual characteristics of luma and chroma sample data in a Coding Unit (CU) [1].
- Bits are, therefore, wasted on perceptually insignificant luma and chroma pixel regions. URQ is not a perceptually optimized quantization technique, which constitutes a significant drawback.

AdaptiveQP in HEVC

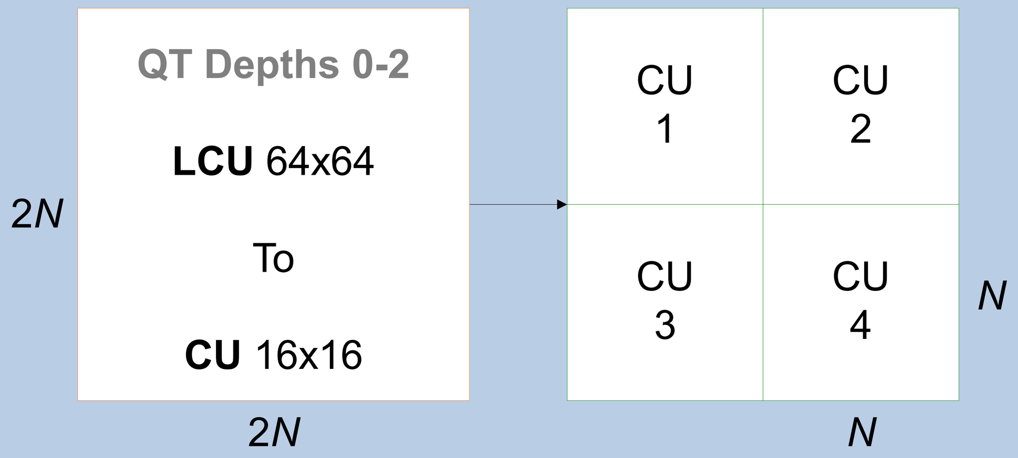

Figure 2: In AdaptiveQP, the $2N \times 2N$ CUs at QuadTree (QT) depth levels 0-2 are partitioned into four $N \times N$ CUs, where N=32 (level 0), N=16 (level 1) or N=8 (level 2). Each CU is then partitioned into four sub-blocks [1].

$$PQ_Y = Q + \left[6 \times \log_2(L)\right] \quad (1)$$

$$L = \frac{f \cdot l + t_Y}{l + f \cdot t_Y} \quad (2)$$

$$f = 2 \quad (3)$$

$$l = 1 + \min\left(\sigma^2_{Y,d}\right), \quad d = 1,\dots,4 \quad (4)$$

PQ_Y — CU-level perceptual QP.
Q — Frame-level QP.
L — Normalized spatial activity in a luma CB.
f — Scaling factor (default QP adaptation range in HM).
l — Non-normalized spatial activity in a luma CB.
t_Y — Mean spatial activity for all $2N \times 2N$ luma CBs.
$\sigma^2_{Y,d}$ — Variance of pixels in sub-block d of a luma CB.

- AdaptiveQP is a luma-based perceptual quantization technique in JCT-VC HEVC HM [2].
- Compared with URQ, it can decrease bitrates without incurring a loss of perceptual quality [3].
- AdaptiveQP increases or decreases the QP of an entire CU based on the variance of pixels in sub-block d of a Y Coding Block (CB) only, which constitutes a shortcoming. See equations (1)-(4) and Figure 2.

Novel Full Color Perceptual Quantization Method (FCPQ)

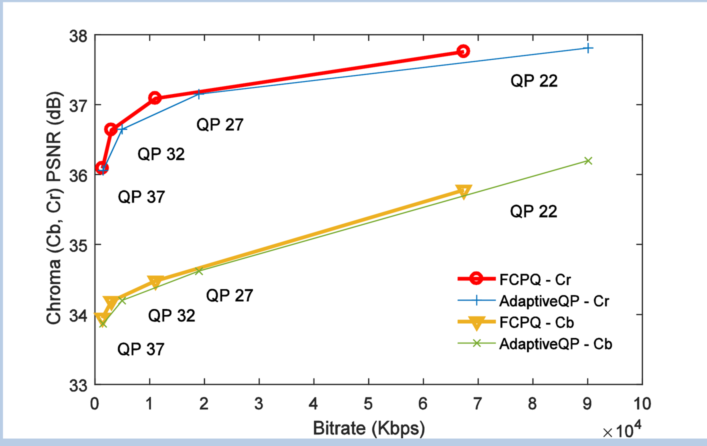

Figure 3: Overall bitrate reductions and PSNR (dB) values attained by FCPQ and AdaptiveQP on the OldTownCross YCbCr 10-bit 4:4:4 HD sequence (RA).

(a) (b)

Figure 4: The YCbCr 4:4:4 10-bit 1080p HD sequence OldTownCross. Subfigure (a): FCPQ. Subfigure (b): AdaptiveQP (both coded with initial $Q = 37$ and RA).

$$PQ_{Cb} = Q + \left[6 \times \log_2(B)\right] \quad (5)$$

$$PQ_{Cr} = Q + \left[6 \times \log_2(R)\right] \quad (6)$$

$$B = \frac{f \cdot b + t_{Cb}}{b + f \cdot t_{Cb}} \quad (7)$$

$$R = \frac{f \cdot r + t_{Cr}}{r + f \cdot t_{Cr}} \quad (8)$$

$$b = 1 + \min\left(\sigma^2_{Cb,k}\right), \quad k = 1,\dots,4 \quad (9)$$

$$t_{Cb} = \frac{1}{C_{Cb}} \sum_{n=1}^{C_{Cb}} b_n \quad (10)$$

$$r = 1 + \min\left(\sigma^2_{Cr,z}\right), \quad z = 1,\dots,4 \quad (11)$$

$$t_{Cr} = \frac{1}{C_{Cr}} \sum_{n=1}^{C_{Cr}} r_n \quad (12)$$

PQ_{Cb} — CB-level perceptual chroma Cb QP.
PQ_{Cr} — CB-level perceptual chroma Cr QP.
B — Normalized spatial activity in a chroma Cb CB.
R — Normalized spatial activity in a chroma Cr CB.
b — Non-normalized spatial activity in a chroma Cb CB.
r — Non-normalized spatial activity in a chroma Cr CB.
t_{Cb} — Mean spatial activity for all $2N \times 2N$ chroma Cb CBs.
t_{Cr} — Mean spatial activity for all $2N \times 2N$ chroma Cr CBs.

- FCPQ is CB-level perceptual quantization technique in HEVC for YCbCr 4:4:4 and RGB video data.
- FCPQ computes the variances of raw data in all three CBs in a CU. See equations (5)-(12).
- The CB-level chroma QPs are signaled in the PPS by exploiting the CU-level QP offset technique [4].
- Bitrates are significantly decreased (See Figure 3) without affecting perceptual quality (see Figure 4).

Table 1: Average bitrate and PSNR (dB) percentage differences over QPs = {22,27,32,37} attained by FCPQ compared with AdaptiveQP on six YCbCr 4:4:4 and six RGB sequences. The All Intra test results are on the left and the Random Access test results are on the right.

FCPQ versus AdaptiveQP – Bitrate (%) and PSNR dB (%) – All Intra								
	YCbCr 4:4:4				RGB			
Sequence	Bitrate	Y	Cb	Cr	Bitrate	G	B	R
BirdsInCage	−25.9	0.0	−0.5	−0.1	−14.6	0.0	−0.9	0.0
DuckAndLegs	−14.6	−0.3	−1.1	−0.6	−9.5	−0.7	−1.1	−0.4
Kimono	−25.5	−0.5	−1.2	−0.5	−23.8	−0.7	−1.8	−1.0
OldTownCross	−28.7	−0.5	−2.1	−1.2	−25.2	−0.9	−2.6	−1.6
ParkScene	−18.4	−0.5	−1.2	−0.6	−16.6	−0.8	−1.7	−1.0
Traffic	−18.8	−0.9	−1.3	−0.9	−16.8	−1.1	−2.2	−1.5

FCPQ versus AdaptiveQP – Bitrate (%) and PSNR dB (%) – Random Access								
	YCbCr 4:4:4				RGB			
Sequence	Bitrate	Y	Cb	Cr	Bitrate	G	B	R
BirdsInCage	−12.6	0.0	0.0	0.0	−8.3	0.0	0.0	0.1
DuckAndLegs	−12.4	0.2	0.1	0.0	−9.8	0.0	0.0	0.1
Kimono	−13.8	0.1	−0.1	0.0	−11.3	−0.2	0.0	−0.2
OldTownCross	−28.6	0.3	−0.3	−0.1	−21.4	−0.1	0.2	0.0
ParkScene	−12.9	−0.1	0.0	−0.1	−11.3	−0.2	0.0	−0.2
Traffic	−11.6	−0.6	−0.2	−0.3	−9.3	−0.6	−0.4	−0.7

- **Best Overall Bitrate Reductions:** 28.7% OldTownCross YCbCr (AI) — See Table 1.
- **Discussion:** FCPQ attains superior bitrate reduction results when applied to high variance video data.
- **Conclusion:** Full color CB-level QP adjustment is superior to luma-only CU-level QP selection.
- **Future Work:** JND-Based CB-level perceptual quantization of 4:4:4 high bit-depth video data.

[1] G. Sullivan, J-R. Ohm, W. Han and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649-1668, 2012.
[2] K. McCann, C. Rosewarne, B. Bross, M. Naccari, K. Sharman and G. J. Sullivan (Editors), "HEVC Test Model 16 (HM 16) Encoder Description," *JCT-VC R1002, 18th Meeting of JCT-VC*, Sapporo, JP, 2014, pp. 1-59.
[3] K. Sato, M. Budagavi, M. Coban, H. Aoki and X. Li, "CE4: Summary report of Core Experiment on quantization," *JCT-VC F024, 6th Meeting of JCT-VC*, Torino, IT, 2011, pp. 1-20.
[4] D. Flynn, N. Nguyen, D. He, A. Tourapis, G. Cote and D.Singer, "RExt: CU-adaptive chroma QP offsets," *JCT-VC O0044, 15th Meeting of JCT-VC*, Geneva, CH, 2013, pp. 1-4.