Joint Video Experts Team (JVET)
of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11
13th Meeting: Marrakech, MA, 9–18 Jan. 2019

Document: JVET-M0091-v1

Title:	**Clean-up and finalization of perceptually optimized QP adaptation method in VTM**		
Status:	Input document to JVET (AHG 10)		
Purpose:	Proposal		
Author(s) or Contact(s):	Christian Helmrich	Email:	christian.helmrich@hhi.fraunhofer.de
Source:	Fraunhofer HHI, Einsteinufer 37, 10587 Berlin, Germany		

Abstract

This contribution proposes a clean-up and a completion of the perceptually optimized QP adaptation (QPA) algorithm already integrated into the VTM codec software. Specifically, the following points are addressed:

1. for HD and smaller input sequences, the previously employed reduction of the CTU size is removed
2. for HD and smaller input, a depth-1 QPA (4 QPs per CTU) is used instead of the CTU size reduction
3. the QPA parameter a_{pic} is now defined as a function of the picture size instead of by a case-statement
4. an extension of the QPA algorithm for better visual handling of CTUs with glaring colors is provided
5. some remaining obsolete QPA related is code removed and some DC offset calculations are unified.

The proposed changes, whose integration into the next VTM software version is suggested, result in slightly reduced bitstream sizes for HD or smaller sequences and the *Campfire* UHD sequence and reportedly provide between 1.7 and 2.5% additional luma BD-rate gain on the random-access coded sequences of the SDR common test conditions (CTC). However, the proposal does not affect the CTC since QPA is off by default.

1 Introduction to Perceptual QP Adaptation

The perceptually optimized QP adaptation (QPA) approach introduced in JVET-H0047 [1] and investigated further in JVET-K0206 [2] and L0181 [3] is a non-normative encoder configuration which can be activated in the VTM reference software [4] via the options --*PerceptQPA* =1, --*SliceChromaQPOffsetPeriodicity* =1. When enabled, it improves the subjective coding quality of some reconstructed videos considerably by way of a *CTU-wise* adaptation of the quantization parameter QP_k, which is delta-coded conventionally [2], [5].

Part of the original QPA algorithm of H0047 was the psychovisually motivated reduction of the CTU size from 128×128 to 64×64 luma samples for high-definition (HD), 2K (2048×1080 luma samples) and smaller sequences, thereby preventing visibly coarse QPA at the cost of a slightly reduced overall coding efficiency. This contribution reports on the implementational details of a *sub-CTU-wise* QPA for said HD and smaller sequences, also called *depth-one* QPA since it utilizes a delta-QP coding depth of one; see also JVET-L0362 [5]. This sub-CTU QPA provides a further slight visual quality improvement over the K0206/L0181 variant currently integrated into VTM 3.x and avoids the aforementioned efficiency reduction due to a change of the CTU size. The five modifications made for this QPA revision are discussed in more detail in Section 2.

2 Proposed Modifications to QPA Algorithm

2.1 Removal of CTU Size Reduction

In function *parseCfg*( ) of the VTM 3 encoder, the CTU size is halved from 128×128 to 64×64 luma samples if QPA is active, $\text{QP}_{\text{base}} \leq 37$, and the picture size is more than 512×320 and less than or equal to 2048×1280.

Instead of reducing the CTU size, it is proposed to, when the abovementioned conditions are met, set the maximum delta-QP coding depth to one (*diff_cu_qp_delta_depth* =1 according to the HEVC-based syntax).

2.2 Introduction of Depth-One QPA

In case of activated QPA and *diff_cu_qp_delta_depth* =1 according to Subsection 2.1, a *sub-CTU-wise* QPA algorithm, calculating a sub-CTU QP_q for each of the four 64×64-sample quadrants q of every CTU, is used for bit-allocation. This sub-CTU QPA follows the method previously described in K0206 [2] to derive the adapted QP_k for each block k except that, now, $k = q$ is a subblock and, accordingly, the high-pass values $h[x, y]$ and the visual activities a_k are calculated separately for each subblock of size $|B_q|$ (instead of $|B_k|$).

Each quantization group (see the delta-QP coding syntax adopted from L0362 [5]) is assigned the respective subblock QP_q for the CTU quadrant q in which that quantization group is located. For this operation to work acceptably in terms of visual quality, split sizes of 128×64 or 64×128 luma samples (i. e., binary-tree splits at the CTU level) are prevented inside the encoder-side search loop when the depth-one QPA is used. Note that CTU-level quad-tree splits or no splits (one CTU-sized CU) are still allowed for best efficiency. Moreover, the Lagrange parameter λ_k is still adapted and applied once per block, i. e., on a CTU-by-CTU basis.

2.3 Simplification of a_{pic} Definition

In K0206, the picture size dependent visual activity normalization factor is given by a three-case definition:

$$a_{\text{pic}} = 2^{BD} \cdot \begin{cases} 16 & \text{for UHD sequences} > 2048 \times 1280 \text{ pels,} \\ 32 & \text{for HD sequences} > 1024 \times 640 \text{ pels,} \\ 64 & \text{otherwise.} \end{cases}$$

A simpler and less coarse specification of a_{pic} can be reached by the following straightforward curve-fitting:

$$a_{\text{pic}} = 2^{BD+4} \cdot \sqrt{\frac{3840 \cdot 2160}{W \cdot H}},$$

where W and H are the picture's luma width and height, respectively, and BD is the coder internal luma bit-depth. Of all (U)HD material used by the JVET, this change only affects the *Cosmos1* sequence (1920×856).

2.4 Extension of Chromatic QPA

The human visual system (HVS) is known to be relatively insensitive to subtle intensity differences – such as those introduced by quantization in lossy coding – in spatial areas dominated by glaring colors like bright yellow, green, or cyan [6]. To exploit this HVS inaccuracy in perceptual bit-allocation applications, a simple extension of the QPA algorithm, increasing the (sub)block-wise QP_k in case of glaring colors, was devised:

1. obtain the DC offset $\tilde{p}_k = \text{mean}(s_p[x, y]), [x, y] \in B_k^p$ of the block B_k^p in each plane $p \in \{\text{Y}, \text{Cb}, \text{Cr}\}$
2. derive the luma-chroma difference $d_k = \tilde{\text{Y}}_k - \tilde{\text{C}}_{\min}$ with $\tilde{\text{C}}_{\min} = \min(\tilde{c}_k), c \in \{\text{Cb}, \text{Cr}\}$ for every k
3. if $d_k > 2^{BD-1}$, then $QP'_k = QP_k + \lfloor 6 \cdot \log_2(d_k \cdot 2^{1-BD}) \rfloor$, otherwise do nothing (i. e., $QP'_k = QP_k$).

In other words, when, in a (sub)block k, the mean luma sample value $\tilde{\text{Y}}_k$ is high (bright) while the minimum $\tilde{\text{C}}_{\min}$ of the mean chroma sample values is very low (far below the grayscale level), the (sub)block input is dominated by a glaring color, and its luma-channel QP_k is increased proportionally by an offset of up to 6.

Of all natural video sequences currently in use by the JVET, the glaring-colors QPA only affects the coding of the *Campfire* recording (see Fig. 1), where it leads to a mean bit-stream size reduction by 5% in all-intra and 7.6% in random-access coding without causing any visible degradation in the subjective coding quality.

2.5 Clean-up of Obsolete Code Parts

The QPA source code in VTM 3 includes some obsolete parts, particularly such which are encapsulated in the *#else* paths of the macros `GLOBAL_AVERAGING` and `SHARP_LUMA_DELTA_QP`. Moreover, picture-wise and block-wise DC offset calculations (mean component sample value) are implemented independently at various places in the encoder. Therefore, it is proposed to remove the obsolete code and to merge the DC offset calculations into a unified, possibly SIMD optimized helper function to be executed instead of the old code.

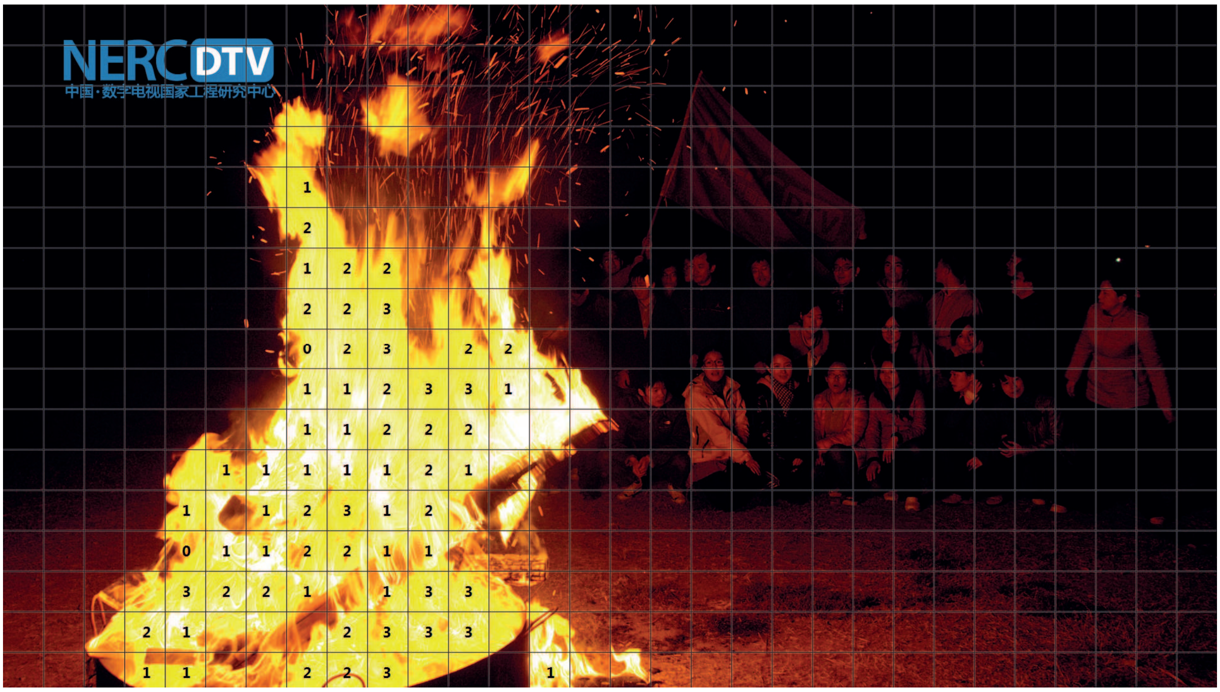

***Figure 1**. Additional CTU-wise QP offsets $\mathrm{QP}_k' - \mathrm{QP}_k$ due to glaring-colors QPA on first frame of Campfire sequence (SDR, $|B_k|$ = 128×128). In CTUs without a number, $d_k \leq 2^{BD-1}$, so the QP offset is 0.*

3 Experimental Results

Bjøntegaard delta (BD) rate gains on the SDR-category common test conditions (CTC) classes of sequences [7], [8] were measured to verify the presence of significant coding efficiency increases due to the improved QPA changes without notable efficiency decreases on any class. For this experiment, a VTM version 3.0.1 with all QP-coding related merged adoptions [3], [5] and fixes [4, request !115] as of Dec. 27, 2018 is used.

Tables 1 and 2 list the (W)PSNR BD-rate results for the modified sub-CTU-wise QPA, with added glaring-color QPA, relative to the reference condition of VTM 3.0.1, providing only a CTU-wise QPA along with said CTU size reduction for the HD-and-smaller random-access (RA) coded sequences. It can be stated that

- on the *Campfire* video in class A1, some coding gain is moved from the Cb to the other components,
- the BD rates for class A2 remain unchanged since all QPA modifications proposed here are inactive,
- the BD rates (PSNR and, partially, WPSNR) for classes *B–F* notably increase in RA due to the depth-one QPA instead of depth-zero QPA with CTU size reduction, particularly in the chromatic channels.

More detailed statistics, including those for low-delay B and P configurations, as well as an alternative MS-SSIM analysis, are provided in three dedicated Excel files included in the Zip file containing this document.

4 Summary and Conclusion

This contribution suggested some finalizing non-normative modifications to the perceptually optimized QP adaptation (QPA) method already integrated into the VTM reference software with the objective of a further increase in both objective and subjective coding efficiency, as well as a clean-up of the QPA source code.

The described algorithmic changes reportedly yield between 1.7 and 2.5% additional luma PSNR BD-rate gain on the random-access coded classes *A1* and *B–F* of sequences of JVET's SDR common test conditions. This gain primarily results from the removal of an encoder-side decision to reduce the CTU size on smaller sequences, combined with the introduction of a sub-CTU *depth-one* QPA instead of the CTU size reduction, when QPA is on. The sub-CTU QPA uses the existing *diff_cu_qp_delta_depth* related QP coding syntax [9].

In light of the outcome of this study, it is requested to adopt the proposed changes to the QPA algorithm in the next VTM software version. Since the modifications are non-normative, the VVC syntax is not affected.

Table 1. *PSNR BD-rate [7] results for VTM 3.0.1 with vs. without proposed changes to QPA, SDR CTC [8].*

All Intra	**Gain Y** (%)	**Gain Cb** (%)	**Gain Cr** (%)	**Time Enc.** (%)	**Time Dec.** (%)
Class A1	–0.98	**4.68**	–0.78	100	97
Class A2	0.00	0.00	0.00	100	98
Class B	0.84	0.77	0.75	102	97
Class C	1.08	0.10	0.06	103	100
Class E	1.31	–1.06	–1.39	111	101
Overall	0.53	0.84	–0.14	103	99
Class D	0.30	**–3.60**	**–3.58**	111	109
Class F	1.44	0.28	0.32	103	100
Random Access	**Gain Y** (%)	**Gain Cb** (%)	**Gain Cr** (%)	**Time Enc.** (%)	**Time Dec.** (%)
Class A1	–1.71	**2.26**	–1.66	99	98
Class A2	0.00	0.00	0.00	101	100
Class B	**–2.48**	**–4.24**	**–4.18**	103	101
Class C	**–2.11**	**–2.55**	**–2.44**	103	101
Class E					
Overall	–1.73	–1.64	**–2.38**	102	100
Class D	–0.33	**–5.24**	**–4.83**	119	109
Class F	**–2.01**	**–2.54**	**–3.20**	106	98

Table 2. *WPSNR BD-rate [7] results for VTM 3.0.1 with vs. without proposed QPA changes, SDR CTC [8].*

All Intra	**Gain Y** (%)	**Gain Cb** (%)	**Gain Cr** (%)	**Time Enc.** (%)	**Time Dec.** (%)
Class A1	–0.52	**6.17**	–0.54	100	100
Class A2	0.00	0.00	0.00	100	96
Class B	–0.03	–0.03	–0.06	102	95
Class C	0.12	–0.76	–0.67	103	98
Class E	0.40	**–2.20**	**–2.56**	111	107
Overall	0.00	0.48	–0.68	103	99
Class D	0.30	**–3.66**	**–3.64**	111	110
Class F	0.11	–0.80	–0.87	103	102
Random Access	**Gain Y** (%)	**Gain Cb** (%)	**Gain Cr** (%)	**Time Enc.** (%)	**Time Dec.** (%)
Class A1	–1.04	**3.18**	–1.29	99	98
Class A2	0.00	0.00	0.00	101	99
Class B	–0.93	**–2.82**	**–2.73**	103	101
Class C	0.88	0.39	0.79	103	103
Class E					
Overall	–0.28	–0.20	–0.95	102	101
Class D	–0.56	**–5.55**	**–4.88**	119	110
Class F	1.19	0.43	0.20	106	87

5 References

[1] S. Bosse, C. Helmrich, H. Schwarz, D. Marpe, T. Wiegand, “Perceptually Optimized QP Adaptation and Associated Distortion Measure,” JVET-H0047, Macao, CN, Oct. 2017.
[2] C. Helmrich, H. Schwarz, D. Marpe, T. Wiegand, “Improved Perceptually Optimized QP Adaptation and Associated Distortion Measure,” JVET-K0206, Ljubljana, SI, July 2018.
[3] C. Helmrich, B. Bross, J. Erfurt, “Corrected Operation of ALF Encoding with Perceptually Optimized QP Adaptation,” JVET-L0181, Macao, CN, Oct. 2018.
[4] JVET/Fraunhofer, “VVCSoftware_VTM,” https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM.
[5] Y. Zhao, H. Yang, J. Chen, “Quantization Parameter Signaling,” JVET-L0362, Macao, CN, Oct. 2018.
[6] A. Valberg, *Light Vision Color*, Wiley, Mar. 2005, and https://arxiv.org/abs/1703.04421.
[7] G. Bjøntegaard, “Calculation of average PSNR differences between RD-curves,” VCEG-M33, 2001.
[8] F. Bossen, J. Boyce, X. Li, V. Seregin, K. Sühring, “JVET common test conditions and software reference configurations for SDR video,” JVET-L1010, Macao, CN, Oct. 2018.
[9] B. Bross, J. Chen, S. Liu, “Versatile Video Coding (Draft 3),” JVET-L1001, ver. 7, Macao, Dec. 2018.

6 Patent Rights Declaration(s)

Fraunhofer HHI may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).