Joint Collaborative Team on Video Coding (JCT-VC)
of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 Document: JCTVC-O0044
15th Meeting: Geneva, CH, 23 October – 1 November 2013

Title: RExt: CU-adaptive chroma QP offsets

Status: Input Document to JCT-VC

Purpose: Proposal

Author(s): David Flynn | *dflynn@blackberry.com*
Nguyen Nguyen | *nnguyen@blackberry.com*
Dake He | *dhe@blackberry.com*
Alexandros Tourapis | *atourapis@apple.com*
Guy Cote | *cote@apple.com*
David Singer | *singer@apple.com*

Source: BlackBerry Ltd. and Apple Inc.

Abstract

In N0292 [1] a method is presented that attempts to address a case where 4:2:0 content is upconverted and encoded as 4:4:4 using a combination of per CU chroma QP offset adjustment and a minimum TU size restriction. During the discussion and presentation of that work, it was suggested that the part of the proposal relating to the chroma QP offset adjustment could be more generally applied than in that specific case.

This contribution provides more general use cases and a method that provides an updated chroma QP offset adjustment mechanism without the minimum TU size restriction.

1 Use cases

Currently the HEVC specification provides a high level technique that permits the relative adjustment of the derived chroma QP (Quantization Parameter) versus a block's luma QP. The lowest granularity at which the relative offset may be controlled is per slice.

Increasing the granularity to the CTU or CU level allows an encoder additional control over rate control and quality control, where implementation is able to adjust the bias between chroma and luma, which can be useful in handling mixed fidelity content (such as a combination of upconverted 4:2:0 and native 4:4:4 content) or in allowing region dependent quantisation.

1.1 Amortizing cost

If an encoder wishes to modify the Chroma QP offset on a per-slice basis, it must either perform some form of per-frame pre-analysis (with a consequential buffering delay) or apply some metric determined during the encoding of a previous picture or slice. A finer grained control mechanism permits an encoder to amortize the cost of any such decisions over the entire picture.

1.2 Rate control

Such adjustments can be useful in rate control where, particularly in 4:4:4 systems, the chroma bitrate can be a substantial fraction, if not the majority of, the total bitrate (see Table 1). At lower QP values, it can be useful to limit the ratio of chroma rate to luma rate to avoid excessive chroma rate due to noise.

True constant bitrate encoders that have a fixed bit allocation per block can also benefit from the ability to trade-off bits between luma and chroma in a content dependent manner.

1.3 Content dependent quantisation control

While experiments performed by the JCT-VC have used constant QP encoding in order to limit variables when analysing the codec's performance, typical encoders will employ delta QP for either rate control or to permit region of interest coding or to enhance the subjective quality.

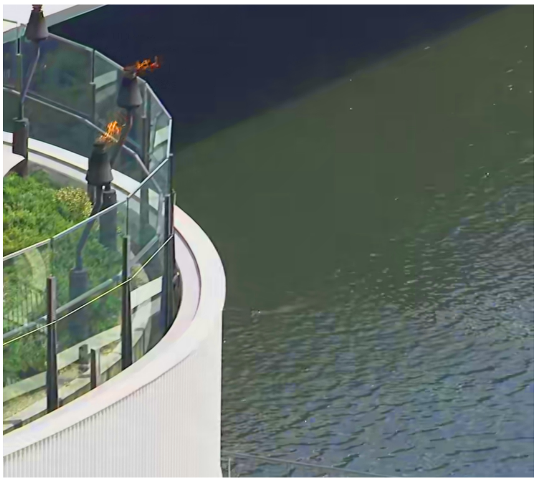
Table 1 – Ratio of Chroma to Luma bitrate for various sequences

	Intra @ QP n						Low Delay @ QP n				
Sequence	12	17	22	27	32	37	17	22	27	32	37
BirdsInCage_1920x1080_60_10bit_444.yuv	2.3	2.6	1.9	1.2	0.4	0.3	4.8	11.7	7.4	0.8	0.4
CrowdRun_1920x1080_50_10bit_444.yuv	1.4	1.2	0.8	0.5	0.3	0.2	1.3	0.7	0.3	0.2	0.1
EBUGraphics_1920x1080_50_10bit_422.yuv	0.7	0.6	0.5	0.3	0.3	0.2	0.6	0.5	0.4	0.3	0.3
EBUHorse_1920x1080_50_10bit_422.yuv	0.5	0.4	0.3	0.1	0.1	0.1	0.5	0.4	0.1	0.1	0
EBUKidsSoccer_1920x1080_50_10bit_422.yuv	0.5	0.3	0.1	0	0	0	0.1	0	0	0	0
EBULupoCandlelight_1920x1080_50_10bit_444.yuv	0.5	0.4	0.2	0.2	0.2	0.2	0.6	0.1	0.1	0.1	0.1
EBURainFruits_1920x1080_50_10bit_444.yuv	0.5	0.4	0.3	0.2	0.1	0.1	0.2	0.1	0.1	0	0
Kimono1_1920x1080_24_10bit_422.yuv	1	1.1	0.7	0.3	0.2	0.1	–	0.5	0.2	0.1	0
Kimono1_1920x1080_24_10bit_444.yuv	2.1	2.4	2	0.6	0.4	0.2	4.1	2.1	0.5	0.2	0.1
Seeking_1920x1080_50_10bit_422.yuv	0.7	0.6	0.4	0.3	0.2	0.1	0.4	0.2	0.2	0.1	0.1
Traffic_2560x1600_30_10bit_422_crop.yuv	0.6	0.5	0.4	0.3	0.2	0.1	0.7	0.5	0.3	0.1	0
Traffic_2560x1600_30_10bit_444_crop.yuv	1.1	0.9	0.7	0.5	0.3	0.2	1.7	1.5	1	0.5	0.2
VenueVu_1920x1080_30_10bit_444.yuv	0.9	0.9	0.8	0.7	0.6	0.5	0.6	0.5	0.5	0.4	0.3

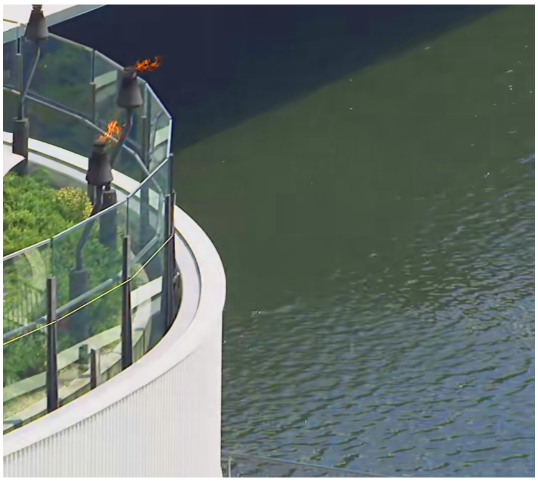
Subjective quality in smooth areas can be significantly improved and is typically done by reducing the QP (luma and chroma) for a particular region based on an activity measure. It isn't necessary though that both luma and chroma QPs are both adjusted. In some situations, luma need only be adjusted, while in others, the chroma QP may be adjusted. Figure 1 shows typical chroma distortion that can be seen in the water of the BQTerrace sequence when encoded at QP 37, where the water can be seen to be comprised of two tones that vary in an unnatural way. By applying a chroma QP offset of -6 relative to that of the luma, the distortion can be reduced. It should be noted that currently, in the scenario that the luma QP is adjusted, chroma is also implicitly impacted, which may, in some situations, be undesirable. In those scenarios, adjusting both parameters synchronously may be preferred.

Figure 1 – Chroma distortion effects (BQTerrace, QP=37)

(a) Anchor



(b) Chroma QP offset = −6



NB, the saturation of these images has been increased to better show the origin of a temporal effect as a single picture.

2 Method

The core of the method allows each CU (up to a given depth) to signal that a chroma QP offset adjustment is to be applied to that CU. The same QP offset may also be applied to subsequent sibling or child CUs within the CTU tree if the maximum depth is reached. This is similar in principle to the operation of delta QP signalling in the current HEVC specification.

To provide additional flexibility in terms of number of offsets, each CU that invokes the mode may signal an index into an offset table. This offset table, which contains a desired set of Chroma QP offset values, is signalled in the PPS so as to limit the rate cost of the provided chroma QP offsets, while allowing implementation discretion as to the exact offsets used. The case where the table contains only a single offset only requires signalling of the mode flag.

The chroma QP adjustment values, at all levels, are limited in range such that the total deviation from the luma QP is limited to the current range of ±12.

Each chroma QP adjustment value is applied equally to both chroma components. I.e., the offset table does not contain individual offsets for each chroma component. The feature is globally enabled through use of a PPS level flag, and locally through a slice header flag. The slice header flag could enable an encoder to disable this feature within a slice if it is determined that signalling these additional indices could be detrimental for the encoding of the current slice, or unnecessary (e.g. for disposable B coded pictures at very high QP values).

3 Encoder implementations

Two independent encoder implementations have been written to investigate chroma QP offsets. This section contains a description of the encoder's non-normative behaviour relating to the proposed mode.

Rather than performing a full RDO test for the mode, simple estimators have been developed. Both perform a pre-analysis on the input picture that is amortized across the encoding of each picture on a CU/CTU basis.

Undoubtedly better estimators or mode decision processes can be developed. It would be expected that using an RDO test would be possible to further exploit the relative trade-offs of the mode.

3.1 Encoder A

This encoder supports multiple QP adjustment values and uses an estimator to select between them. During the mode decision process, the following steps are performed. For the current CTU, compute the variance of the pixel values that comprise the source data. If the variance is within a predefined threshold, the mode is enabled. The magnitude of the chroma QP adjustment used is determined by linearly scaling the variance thresholds to the QP adjustment range. The default variance thresholds are [500, 3000].

3.2 Encoder B

This encoder estimator only supports a single chroma QP adjustment value (ignoring implied zero adjustment). During the mode decision process, the following steps are performed:

1. For the current CU, estimate the transform domain activity of the source data to determine the mode decision. In the case where the current CU is below the maximum depth at which the mode can be signalled, the estimation is performed once for the area covered by the minimum CU size and applies to all smaller CUs in the sub-tree.
2. For a CU with the mode enabled, adjust the chroma lambda to the appropriate value for the effective QP offset. The lambda is calculated at the start of the slice in the usual manner.

The estimator uses the HM transform to analyse the co-located chroma blocks in the source video data to determine the amount of high-frequency activity. The mode decision is based upon a simple threshold that counts the number of high-frequency coefficients with a magnitude above a particular value:

$$est_c = \text{loFiCountThreshold} < \left(\sum |x_{i,j}| > \text{loFiQuant}\right), \text{where } \begin{array}{l} x = T(blk_c) \\ i,j \geq \frac{blkSize}{2} \end{array}$$

The estimation is performed for both chroma components, and if either exceeds the threshold, the mode is asserted. The thresholds for loFiQuant and loFiCountThreshold are 8 and 8 respectively.

The current encoder implementation has a limitation with regard to the setting of the correct lambda for SAO decisions, where the lambda is only set once per CTU. When the maximum depth is set to zero, i.e., the mode operates on a CTU basis, the design behaves as intended.

4 Results

Table 2 – BD-Rate performance of encoder A under All-Intra configuration

(a) Y'CbCr 4:2:2

	Main-tier			High-tier			Super-High-tier		
	Y	U	V	Y	U	V	Y	U	V
Traffic_2560x1600_30_10bit_422_crop.yuv	−4.4%	15.2%	17.3%	−5.2%	13.6%	15.6%	−5.7%	11.9%	14.1%
Kimono1_1920x1080_24_10bit_422.yuv	−5.7%	7.6%	11.8%	−31.7%	30.0%	27.1%	−26.8%	33.3%	22.4%
EBUHorse_1920x1080_50_10bit_422.yuv	−4.8%	29.9%	36.0%	−6.9%	25.7%	32.5%	−7.9%	20.7%	23.8%
EBUGraphics_1920x1080_50_10bit_422.yuv	−10.6%	31.6%	36.6%	−12.4%	30.3%	33.9%	−12.9%	24.4%	24.5%
EBUWaterRocksClose_1920x1080_50_10bit_422.yuv	−5.3%	18.1%	24.1%	−7.2%	19.1%	22.9%	−7.4%	17.3%	16.1%
EBUKidsSoccer_1920x1080_50_10bit_422.yuv	−1.2%	16.3%	15.4%	−1.8%	9.0%	8.3%	−2.6%	5.9%	5.7%
Seeking_1920x1080_50_10bit_422.yuv	−7.3%	31.9%	34.5%	−9.2%	29.0%	30.4%	−9.6%	21.6%	21.7%
Overall	−5.6%	21.5%	25.1%	−10.6%	22.4%	24.4%	−10.4%	19.3%	18.3%

(b) Y'CbCr 4:4:4

	All Intra Main-tier			All Intra High-tier			All Intra Super-High-tier		
	Y	U	V	Y	U	V	Y	U	V
Traffic_2560x1600_30_10bit_444_crop.yuv	−6.4%	14.0%	16.5%	−7.6%	13.2%	15.2%	−8.0%	11.3%	13.6%
Kimono1_1920x1080_24_10bit_444.yuv	−6.3%	6.3%	13.1%	−8.8%	7.8%	9.4%	−8.8%	9.0%	7.5%
EBULupoCandlelight_1920x1080_50_10bit_444.yuv	−1.7%	14.0%	16.3%	−2.1%	10.5%	15.8%	−3.2%	10.5%	15.7%
EBURainFruits_1920x1080_50_10bit_444.yuv	−2.3%	13.3%	13.2%	−2.8%	11.5%	11.7%	−3.6%	10.3%	11.7%
VenueVu_1920x1080_30_10bit_444.yuv	−9.9%	21.8%	21.8%	−10.8%	20.7%	20.6%	−11.3%	22.3%	21.6%
BirdsInCage_1920x1080_60_10bit_444.yuv	−15.4%	28.7%	24.3%	−18.7%	24.3%	25.5%	−13.3%	16.1%	19.7%
CrowdRun_1920x1080_50_10bit_444.yuv	−10.6%	28.7%	31.6%	−14.2%	24.8%	25.4%	−15.5%	19.5%	19.0%
Overall	−7.5%	18.1%	19.5%	−9.3%	16.1%	17.6%	−9.1%	14.2%	15.5%

(c) RGB 4:4:4

	All Intra Main-tier			All Intra High-tier			All Intra Super-High-tier		
	G	B	R	G	B	R	G	B	R
Traffic_2560x1600_30_12bit_444_crop.rgb	−23.4%	21.7%	25.6%	−24.1%	20.8%	23.4%	−23.6%	18.8%	19.8%
Kimono1_1920x1080_24_10bit_444.rgb	−29.9%	24.1%	38.0%	−31.7%	30.0%	27.1%	−26.8%	33.3%	22.4%
EBULupoCandlelight_1920x1080_50_10bit_444.rgb	−20.5%	24.1%	30.9%	−22.6%	22.0%	35.3%	−22.2%	26.0%	31.6%
EBURainFruits_1920x1080_50_10bit_444.rgb	−20.1%	22.3%	25.1%	−20.5%	20.6%	23.8%	−21.0%	18.8%	24.0%
VenueVu_1920x1080_30_8bit_444.rgb	−19.9%	18.7%	20.0%	−19.8%	19.2%	20.2%	−20.2%	22.4%	23.8%
DucksAndLegs_1920x1080_30_10bit_444.rgb	−26.9%	29.1%	30.0%	−22.3%	20.3%	17.9%	−16.6%	8.9%	7.4%
OldTownCross_1920x1080_50_10bit_444.rgb	−33.1%	33.7%	34.6%	−28.9%	33.2%	32.9%	−23.4%	23.3%	22.8%
ParkScene_1920x1080_24_10bit_444.rgb	−26.9%	23.9%	34.8%	−26.7%	24.4%	28.1%	−23.3%	22.8%	20.3%
Overall	−24.8%	24.7%	29.9%	−24.3%	23.8%	26.1%	−22.0%	21.8%	21.5%

5 Specification Text

Please see enclosed specification text document.

BlackBerry Limited 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).

Apple Inc 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).

References

[1] D. Flynn, N. Nguyen, and D. He, "Rext: Fidelity adaptive coding mode," JCTVC-N0292, JCT-VC, Jul. 2013.