



**HAL**  
open science

## Temporal Adaptive Quantization using accurate estimations of Inter and Skip probabilities

Maxime Bichon, Julien Le Tanou, Michael Ropert, Wassim Hamidouche, Luce Morin, Lu Zhang

► **To cite this version:**

Maxime Bichon, Julien Le Tanou, Michael Ropert, Wassim Hamidouche, Luce Morin, et al.. Temporal Adaptive Quantization using accurate estimations of Inter and Skip probabilities. Picture Coding Symposium (PCS), Jun 2018, San Francisco, United States. 10.1109/PCS.2018.8456275 . hal-01793025

**HAL Id: hal-01793025**

**<https://hal.science/hal-01793025>**

Submitted on 16 May 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Temporal Adaptive Quantization using accurate estimations of Inter and Skip probabilities

Maxime Bichon<sup>\*,†</sup>, Julien Le Tanou<sup>\*</sup>, Michael Ropert<sup>\*</sup>

<sup>\*</sup>Ericsson R&D, TV & Media

35136 Saint-Jacques-de-la-Lande, France

firstname.lastname@ericsson.com

Wassim Hamidouche<sup>†</sup>, Luce Morin<sup>†</sup> and Lu Zhang<sup>†</sup>

<sup>†</sup>IETR Lab UMR CNRS 6164, INSA Rennes

35708 Rennes, France

firstname.lastname@insa-rennes.fr

**Abstract**—Hybrid video coding systems use spatial and temporal predictions in order to remove redundancies within the video source signal. These predictions create coding-scheme-related dependencies, often neglected for sake of simplicity. The R-D Spatio-Temporal Adaptive Quantization (RDSTQ) solution uses such dependencies to achieve better coding efficiency. It models the temporal distortion propagation by estimating the probability of a Coding Unit (CU) to be Inter coded. Based on this probability, each CU is given a weight depending on its relative importance compared to other CUs. However, the initial approach roughly estimates the Inter probability and does not take into account the Skip mode characteristics in the propagation. It induces important Target Bitrate Deviation (TBD) compared to the reference target rate. This paper provides undeniable improvements of the original RDSTQ model in using a more accurate estimation of the Inter probability. Then a new analytical solution for local quantizers is obtained by introducing the Skip probability of a CU into the temporal distortion propagation model. The proposed solution brings  $-2.05\%$  BD-BR gain in average over the RDSTQ at low rate, which corresponds to  $-13.54\%$  BD-BR gain in average against no local quantization. Moreover, the TBD is reduced from  $38\%$  to  $14\%$ .

## I. INTRODUCTION

Rate Distortion Optimization (RDO) concept [1] aims to minimize the Distortion  $D$  subject to a rate constraint  $R \leq R_T$ .  $\lambda$  is the Lagrange multiplier which controls the trade-off between  $D$  and  $R$  [2]. RDO process minimizes the Rate-Distortion (R-D) cost function  $J = D + \lambda R$ .

Video encoders based on recent Motion Picture Expert Group (MPEG) video compression standards, such as High Efficiency Video Coding (HEVC) [3], rely on closed-loop predictive coding scheme processed on block-level. The prediction process refers to reconstructed samples that have been affected by the lossy quantization process. Obviously, the distortion made on the reference samples impacts the coding efficiency of samples, or Coding Unit (CU), referring to. These distortion propagation effects are stated in this article as Inter-CU Dependencies (*ICUD*). In order to select the best encoding parameters, RDO is commonly performed on each CU without considering *ICUD*. It results in a suboptimal source compression from an R-D standpoint, as further explained by Ortega and Ramchandran in [4].

Several studies have tried to model the coding dependencies in order to improve the global coding efficiency [5]-[13]. These solutions differ from each other in terms of the considered

dependencies (frame-level or CU-level) and/or the coding parameters to be estimated (prediction mode, quantization parameter, target bitrate, etc).

Recent papers consider dependencies from the Rate-Control (RC) perspective [5], [6], [7]. In [5], Xie et al. construct a temporal propagation chain by assuming consistent motion over frames, which allows the estimation of a propagation factor for each Coding Tree Unit (CTU). The propagation factor is used to derive the target bitrate for each CTU, supposed optimal for the whole sequence. Li et al. [6] estimate the distortion propagation frame by frame, using a pre-analysis, in order to provide a consistent video quality over an entire Group of Pictures (GOP). Fiengo et al. [7] express the distortion as a convex function of all frames bitrate. Primal-Dual Proximal Algorithm is further used to solve the convex optimization problem and achieve near optimal RC. However, this solution is too complex to be applied for real-time encoding.

The *ICUD* have been considered for Intra coding in [8], [9]. In [8], Qingbo et al. experimentally estimate a linear distortion propagation model used for  $\lambda$  computation, coupled with an off-line learning and a multiple Lagrange Multiplier framework. *ICUD* consideration for Intra prediction mode optimization has been evaluated in [9].

Adaptive quantization has been optimized by modeling dependencies either at the pixel level [10], transform coefficient level [11] or CU level [12], [13]. In [10], Valenzise and Ortega estimate a temporal dependency tree, further used to design an adaptive quantization based on the tree depth. Winken et al. [11] measure the dependencies between coefficients levels after Discrete Cosine Transform (DCT)/Discrete Sine Transform (DST), leading to an optimization problem solved by an iterative approach. Ropert et al. [12] propose the Rate Distortion Spatio-Temporal Quantization (*RDSTQ*), as a generalization of Macroblock-Tree framework designed for x264 open-source encoder [13]. The main interest of this approach is to model the temporal distortion propagation between CUs from an R-D standpoint.

Despite the high R-D efficiency of *RDSTQ* against no adaptive quantization ( $-19,4\%$  Structural Similarity (SSIM)-based bits savings in average), the distortion propagation model shows room for improvements. First, the Inter probability estimation is conservative and does not fit well the ground truth, leading to sub-optimal R-D behavior. Second, the *RD-*

*STQ* optimization aims to keep the average bitrate of the GOP unchanged, or equivalently its average quantizer, compared to a reference without adaptive quantization. However, coding performance of the Skip mode is not considered in the *RDSTQ*, which leads to significant Target Bitrate Deviation (*TBD*) as detailed in section II-C.

In this paper we propose two main improvements to the *RDSTQ* solution: an accurate estimation of the Inter prediction probability and the consideration of the Skip mode probability into the distortion propagation model. These two enhancements enable a Bjøntegaard Delta Bit Rate (BD-BR) saving of  $-2.05\%$  in average against *RDSTQ* solution using the x.265 HEVC video encoder, while the *TBD* is reduced from 38% to 14%.

The rest of the paper is organized as follows. An overview of the *RDSTQ* solution and its weaknesses is first presented in Section II. The proposed improvements to overcome identified limitations are described in Section III. Experimental results, demonstrating the benefits of the improved model, are exhibited in Section IV. Finally, conclusions are given in Section V.

## II. *RDSTQ* MODEL OVERVIEW

We define  $i$  as the CU index in the frame,  $t$  the frame index,  $N$  the number of CUs in a frame, and  $T$  the GOP length. *RDSTQ* model aims to find the set of local quantizers  $\{q_{i_t}\}_{i=0, t=0}^{i=N-1, t=T-1}$  noted  $\{q\}$  for all CUs within the GOP, in order to minimize the total distortion  $D_{Tot}$

$$\begin{aligned} \{q^*\} &= \arg \min_{\{q\}} D_{Tot} \\ &= \arg \min_{\{q\}} \sum_{t=0}^{T-1} \sum_{i=0}^{N-1} \Psi_{i_t} D_{i_t}(\{q\}) \\ \text{s.t.} & \sum_{t=0}^{T-1} \sum_{i=0}^{N-1} R_{i_t}(q_{i_t}) = R_{Tot}. \end{aligned} \quad (1)$$

The *RDSTQ* is designed to optimize various quality metrics such as SSIM by using a psycho-visual weighting factor  $\Psi$  based on the local pixel variance. For the particular case of  $\Psi = 1$  for all CUs, the model minimizes the Mean Square Error (MSE) and is named Rate Distortion Temporal Quantization (*RDTQ*). After formalization of a temporal distortion propagation model, the total distortion expression is exhibited, and the optimization problem (1) is analytically solved under few assumptions. We point out that no psycho-visual function was considered in this paper, i.e.  $\Psi = 1$ , hence only *RDTQ* is mentioned hereafter.

### A. Temporal Distortion Propagation Model

The *RDTQ* defines the distortion  $D_{i_t}$  of a given CU  $i_t$  as a sum of its local distortion  $d_{i_t}$  and the distortion  $D_{j_{t_{ref}}}$  propagating from its reference CUs  $j_{t_{ref}}$ . The propagation formula is given as follows:

$$D_{i_t} = d_{i_t}(q_{i_t}) + p_{i_t} \underbrace{\sum_{j_{t_{ref}} \in Ref(i_t)} r_{j_{t_{ref}}, i_t} D_{j_{t_{ref}}}}_{\eta_{i_t}}. \quad (2)$$

with  $Ref(i_t)$  is the set of references used for motion compensation,  $p_{i_t}$  is the probability of a CU to be Inter coded

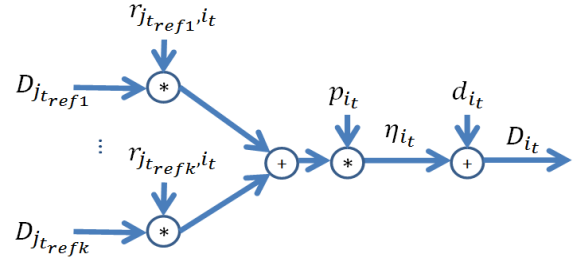


Fig. 1. Temporal Distortion Propagation Scheme

and  $r_{j_{t_{ref}}, i_t}$  the pixel surface ratio involved in the motion compensation to go from pixels of  $j_{t_{ref}}$  to  $i_t$ .  $d_{i_t}(q_{i_t})$  is the intrinsic distortion, i.e. the distortion that only depends of the local quantizer. Fig. 1 illustrates the temporal distortion propagation model. The main simplification made for solving problem (1) is to assume that  $\eta_{i_t}$  and  $p_{i_t}$  are independent of  $q_{i_t}$ . Considering  $p_{i_t}$  and  $\eta_{i_t}$  are independent of  $q_{i_t}$  is the basic element allowing to achieve analytical solution.

### B. Local Quantization and Analytical Solution

Generalizing the propagation over the entire GOP based on (1) and (2) leads to the total distortion formula (3).

$$D_{Tot} = \sum_{t=0}^{T-1} \sum_{i=0}^{N-1} \Psi_{i_t} \left( p_{i_t} \sum_{i_{t-1} \in Ref(i_t)} r_{i_{t-1}, i_t} \left( \dots \left( p_{i_1} \sum_{i_0 \in Ref(i_1)} r_{i_0, i_1} d_{i_0} + d_{i_1} \right) + \dots \right) + d_{i_t} \right). \quad (3)$$

The total distortion derivative is equal to the local distortion derivative multiplied by the accumulation factor  $U_{k_\tau}$  for each CU  $k_\tau$  as

$$\frac{\partial D_{Tot}}{\partial Q_{k_\tau}} = \frac{\partial d_{k_\tau}}{\partial Q_{k_\tau}} U_{k_\tau}. \quad (4)$$

Accumulation factor  $U_{k_\tau}$  is recursively defined by

$$U_{k_{\tau-1}} = \sum_{i_\tau} p_{i_\tau} r_{k_{\tau-1}, i_\tau} U_{i_\tau} + \Psi_{k_{\tau-1}}, U_{k_{T-1}} = \Psi_{k_{T-1}}. \quad (5)$$

Finally, in considering the independence of rates, the R-D Shannon bound and the high bitrate assumption  $d = Q^2/12$ , the optimal local delta quantization parameter results in

$$\Delta q_{k_\tau} = -3 \left( \log_2(U_{k_\tau}) - \frac{\sum_{t=0}^{T-1} \sum_{i=0}^{N-1} \log_2(U_{i_t})}{T \times N} \right). \quad (6)$$

### C. Model Limitations

One defines  $\omega_{k_\tau}^{Intra} > 0$  and  $\omega_{k_\tau}^{Inter} > 0$  the Sum of Absolute Transform Differences (SATD) prediction costs of Intra and Inter modes, respectively. The SATD costs are estimated in the look-ahead analysis. Probability of Inter prediction is defined as a function of  $\omega_{k_\tau}^{ratio} = \omega_{k_\tau}^{Intra} / \omega_{k_\tau}^{Inter}$ . The Inter probability estimator used in [12] is given by

$$p_{k_\tau} = 1 - \min\left(1; \frac{1}{\omega_{k_\tau}^{ratio}}\right). \quad (7)$$

This formula assumes that if SATD costs are equivalent, i.e.  $\omega_{k_\tau}^{ratio} = 1$ , Inter probability should be null and there is no propagation, i.e.  $p_{k_\tau} = 0$ . However, close Intra/Inter prediction costs should intuitively lead to equiprobable Intra and Inter modes. Moreover, neither theoretical nor experimental proof of the correctness of (7) has been given. Section III-A gives more insights and proposes a solution which improves the R-D efficiency.

The optimization problem (1) is solved in [12] assuming that the bitrate constraint is fulfilled or equivalently that the average quantizer over the GOP should remain unchanged against a configuration without adaptive quantization. However, the Skip mode is ignored in the initial solution which leads to an important *TBD*. In terms of quantization, Skip mode is equivalent to an infinite quantization step, i.e.  $D = \sigma^2$  and  $R = 0$ . The *TBD* of *RDTQ* and the Skip mode consideration for adaptive quantization are deeply analyzed in Section III-B. A new analytical solution to (1) with respect to Skip is exhibited.

### III. ENHANCED *RDTQ* MODEL

Accurate Inter prediction probability estimation and Skip mode consideration for adaptive quantization are both described in the following subsections.

#### A. Inter Probability

Based on Statistical inference, we estimate the Inter probability  $p$  as the Likelihood function  $L(\omega^{ratio}|mode) \propto P(mode = Inter|\omega^{ratio})$  and plot it against the function (7) on Fig. 2.  $\omega_{k_\tau}^{ratio}$  is the prior information known beforehand while the event for a CU  $k_\tau$  to be Inter coded is the evidence.

Two other functions have been experimented. The first one, defined by (8), conserves the same shape as (7) but assumes Inter probability to be equal to 0.5 in the case of  $\omega_{k_\tau}^{Intra} = \omega_{k_\tau}^{Inter}$ .

$$p_{k_\tau} = \frac{\omega_{k_\tau}^{ratio} - 1 + \max(\omega_{k_\tau}^{ratio}; 1)}{2 * \max(\omega_{k_\tau}^{ratio}; 1)} \quad (8)$$

The second one, defined by (9), is a gamma distribution fitting the ground truth extracted from a RDO. Functions (8) and (9) are both plotted on Fig. 2.

$$p_{k_\tau} = \frac{1}{\Gamma(\alpha)} \gamma\left(\alpha, \frac{\omega_{k_\tau}^{ratio}}{\beta}\right) \quad (9)$$

The performances of each function are discussed in Section IV.

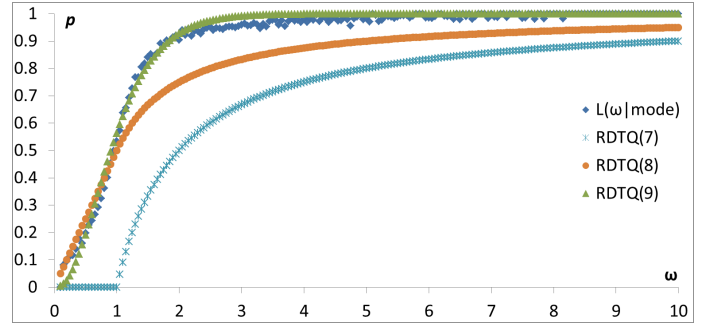


Fig. 2. Inter Probabilities  $p$  according to  $\omega$  estimate by (7), (8) and (9).

#### B. Skip consideration

Let  $c_{k_\tau}$  the probability of the CU  $k_\tau$  to be coded and  $1 - c_{k_\tau}$  the probability of the CU to be skipped. We define  $D_{k_\tau}^C$  and  $R_{k_\tau}^C$  the distortion and rate of a coded CU  $k_\tau$  while  $D_{k_\tau}^S$  and  $R_{k_\tau}^S$  the distortion and rate of a skipped CU, such

$$D_{k_\tau}^C = d_{k_\tau} + p_{k_\tau} \eta_{k_\tau}, \quad D_{k_\tau}^S = \sigma_{k_\tau}^2 + p_{k_\tau} \eta_{k_\tau}. \quad (10)$$

$D_{k_\tau}^C$  is the distortion given in (2).  $D_{k_\tau}^S$  is also composed of the propagated distortion from references  $\eta_{k_\tau}$ , but does not include the intrinsic distortion, since Skip mode do not apply any quantization. However, it is related to innovation  $\sigma_{k_\tau}^2$ , the unpredictable part of the signal, defined as the variance of the residue. One can rewrite (2) in introducing the Skip probability as follow:

$$D_{k_\tau} = c_{k_\tau} D_{k_\tau}^C + (1 - c_{k_\tau}) D_{k_\tau}^S \\ = c_{k_\tau} d_{k_\tau} + [1 - c_{k_\tau}] \sigma_{k_\tau}^2 + p_{k_\tau} \eta_{k_\tau}. \quad (11)$$

We sum up all distortions and derivate the total distortion according to the local quantizer  $q_{k_\tau}$ . After some manipulations, we obtain (12).

$$\frac{\partial D_{Tot}}{\partial Q_{k_\tau}} = \left( \frac{\partial d_{k_\tau}}{\partial Q_{k_\tau}} c_{k_\tau} + \frac{\partial c_{k_\tau}}{\partial Q_{k_\tau}} d_{k_\tau} - \frac{\partial c_{k_\tau}}{\partial Q_{k_\tau}} \sigma_{k_\tau}^2 \right) U_{k_\tau} \quad (12)$$

$Q_{k_\tau} = 2^{\frac{q_{k_\tau} - 4}{6}}$  is the local quantizer step according to the quantizer parameter  $q_{k_\tau}$ .  $c_{k_\tau}$  is chosen and derived as

$$c_{k_\tau} = \frac{12\sigma_{k_\tau}^2}{12\sigma_{k_\tau}^2 + Q_{k_\tau}^2} \quad (13)$$

$$\frac{\partial c_{k_\tau}}{\partial Q_{k_\tau}} = \frac{-24\sigma_{k_\tau}^2 Q_{k_\tau}}{144\sigma_{k_\tau}^4 + Q_{k_\tau}^4 + 24\sigma_{k_\tau}^2 Q_{k_\tau}^2}. \quad (14)$$

According to (14) we assume  $\frac{\partial c_{k_\tau}}{\partial Q_{k_\tau}} \approx 0$  for sufficiently high values of  $\sigma_{k_\tau}^2$  and  $Q_{k_\tau}^2$ . Finally, (12) is simplified as (15)

$$\frac{\partial D_{Tot}}{\partial Q_{k_\tau}} = \frac{\partial d_{k_\tau}}{\partial Q_{k_\tau}} c_{k_\tau} U_{k_\tau}. \quad (15)$$

The rate  $R_{k_\tau}$  of a CU  $k_\tau$  can also be rewritten as function of  $c_{k_\tau}$ . However, rate of skipped CUs is theoretically equal to zero. Thus the R-D cost to minimize  $J_{Tot}$  is defined by

$$J_{Tot} = D_{Tot} + \lambda \left( \sum_{t=0}^{T-1} \sum_{i=0}^{N-1} c_{i_t} R_{i_t} - R_{Tot} \right). \quad (16)$$

The necessary condition to find the minimum of  $J_{Tot}$  is determined by the condition of all derivatives equal to zero  $\forall k \in \{1, \dots, N\}, \forall \tau \in \{0, \dots, T-1\}$

$$\frac{\partial J_{Tot}}{\partial Q_{k_\tau}} = \frac{\partial D_{Tot}}{\partial Q_{k_\tau}} + \lambda \frac{\partial}{\partial Q_{k_\tau}} \sum_{t=0}^{T-1} \sum_{i=0}^{N-1} c_{i_t} R_{i_t} = 0. \quad (17)$$

We assume independence of rates and uses the same approximation as for (14). The total rate derivate results in

$$\frac{\partial}{\partial Q_{k_\tau}} \sum_{t=0}^{T-1} \sum_{i=0}^{N-1} c_{i_t} R_{i_t} = \frac{\partial R_{k_\tau}}{\partial Q_{k_\tau}} c_{k_\tau}, \quad (18)$$

and from (15) we obtain

$$\frac{\partial J_{Tot}}{\partial Q_{k_\tau}} = \frac{\partial d_{k_\tau}}{\partial Q_{k_\tau}} c_{k_\tau} U_{k_\tau} + \lambda \frac{\partial R_{k_\tau}}{\partial Q_{k_\tau}} c_{k_\tau} = 0. \quad (19)$$

Consequently,  $\lambda$  is independent of  $c_{k_\tau}$  and

$$\lambda^* = 2\ln(2)U_{k_\tau}D_{k_\tau}. \quad (20)$$

Summing the weighted log values on both side over all CUs of the GOP and setting  $\lambda' = \lambda/(2\ln(2))$ , we have

$$\log_2(\lambda') \underbrace{\sum_{t=0}^{T-1} \sum_{i=0}^{N-1} c_{i_t}}_{=N_{Tot}} = \sum_{t=0}^{T-1} \sum_{i=0}^{N-1} c_{i_t} \log_2(U_{i_t}D_{i_t}) \quad (21)$$

$$\frac{\sum_{t=0}^{T-1} \sum_{i=0}^{N-1} c_{i_t} \log_2(U_{i_t}D_{i_t})}{N_{Tot}} = \log_2(U_{k_\tau}D_{k_\tau}). \quad (22)$$

At the other side, let's compute  $\frac{2R_{Tot}}{N_{Tot}}$  based on the R-D Shannon bound and combine with the previous equality

$$\frac{2R_{Tot}}{N_{Tot}} = \frac{2}{N_{Tot}} \sum_{t=0}^{T-1} \sum_{i=0}^{N-1} c_{i_t} R_{i_t} \quad (23)$$

$$\frac{2R_{Tot}}{N_{Tot}} = \frac{\sum_{t=0}^{T-1} \sum_{i=0}^{N-1} c_{i_t} (\log_2(c\sigma_{i_t}^2) - \log_2(d_{i_t}))}{N_{Tot}}. \quad (24)$$

By mixing (22) with (24), we exhibit

$$\begin{aligned} \frac{2R_{Tot}}{N_{Tot}} &= -\log_2(U_{k_\tau}) + 2R_{k_\tau} - \log_2(c\sigma_{k_\tau}^2) \\ &+ \frac{\sum_{t=0}^{T-1} \sum_{i=0}^{N-1} c_{i_t} \log_2(c\sigma_{i_t}^2 U_{i_t})}{N_{Tot}}. \end{aligned} \quad (25)$$

By using the high bitrate rate approximation  $d_{k_\tau} = \frac{Q_{k_\tau}^2}{12}$  with (25), the optimal quantization parameter  $q_{k_\tau}$  is

$$q_{k_\tau} = 3 \left[ - \left( \log_2(U_{k_\tau}) - \frac{\sum_{t=0}^{T-1} \sum_{i=0}^{N-1} c_{i_t} \log_2(U_{i_t} c \sigma_{i_t}^2)}{N_{Tot}} \right) - \frac{2R_{Tot}}{N_{Tot}} + \log_2(12) \right] + 4. \quad (26)$$

Finally, if we compute  $\frac{2R_{Tot}}{N_{Tot}}$  assuming the whole sequence is encoded with a unique quantizer and combine it with (26), the optimal delta quantizer  $\Delta q_{k_\tau}$  is

$$\Delta q_{k_\tau} = -3 \left( \log_2(U_{k_\tau}) - \frac{\sum_{t=0}^{T-1} \sum_{i=0}^{N-1} c_{i_t} \log_2(U_{i_t})}{\sum_{t=0}^{T-1} \sum_{i=0}^{N-1} c_{i_t}} \right). \quad (27)$$

#### IV. EXPERIMENTAL RESULTS

We use the x265 software HEVC encoder [14] in order to have similar test conditions than [12]. We consider the Common Test Conditions (CTC) defined by the Joint Collaborative Team on Video Coding (JCT-VC) [15]. The videos are encoded in Random Access (RA) coding configuration for five Quantization Parameters (QP) values  $\in \{22, 27, 32, 37, 42\}$ . We add the QP value of 42 to highlight the Skip mode influence since it is statistically more used at low bitrate.

Coding performance is measured using BD-BR metric [16]. Negative BD-BR value reflects the percentage of bit savings achieved at equivalent YUV distortion, measured with Peak Signal to Noise Ratio (PSNR), between the anchor and the proposed solution. The anchor is the x265 encoder without adaptive quantization algorithm. Three Inter probability models, defined in (7), (8) and (9) are compared.

The BD-BR results and the corresponding bitrate deviations for the five considered QP values are presented in Table I. Moreover, since Skip consideration should be statistically more efficient at low bitrate, similar results presented in Table II consider only the four highest QP values  $\{27, 32, 37, 42\}$ .

TABLE I  
CODING EFFICIENCY OF INTER PROBABILITIES AND SKIP CONSIDERATION OVER NO LOCAL QUANTIZATION IN x265.

Models	$RDTQ$ (7)	$RDTQ$ (8)	$RDTQ$ (9)
	Without Skip probability: $c_{k_\tau} = 1 \forall k, \tau$		
<b>Average</b>	<b>-10.38%</b>	<b>-12.18%</b>	<b>-11.96%</b>
<b>Best</b>	<b>-19.00%</b>	<b>-20.77%</b>	<b>-22.37%</b>
<b>Worst</b>	<b>0.24%</b>	<b>-2.10%</b>	<b>0.91%</b>
With Skip probability			
<b>Average</b>	<b>-10.07%</b>	<b>-11.87%</b>	<b>-11.84%</b>
<b>Best</b>	<b>-19.01%</b>	<b>-20.43%</b>	<b>-22.04%</b>
<b>Worst</b>	<b>-1.60%</b>	<b>-4.57%</b>	<b>-3.39%</b>
Without Skip probability: $c_{k_\tau} = 1 \forall k, \tau$			
<b>Average</b>	<b>37.35%</b>	<b>53.88%</b>	<b>75.01%</b>
With Skip probability			
<b>Average</b>	<b>10.3%</b>	<b>12.67%</b>	<b>17.53%</b>

TABLE II  
CODING EFFICIENCY OF INTER PROBABILITIES AND SKIP CONSIDERATION  
OVER NO LOCAL QUANTIZATION IN x265 FOR LOW RATE.

	Models	$RDTQ$ (7)	$RDTQ$ (8)	$RDTQ$ (9)	
	Without Skip probability: $c_{k,\tau} = 1 \forall k, \tau$				
BD-BR PSNR	Class A (8bits)	-8.55%	-10.70%	-11.39%	
	Class B	-9.42%	-12.05%	-12.14%	
	Class C	-15.43%	-16.76%	-17.23%	
	Class D	-11.99%	-12.99%	-13.45%	
	Class E	-10.99%	-15.08%	-16.28%	
	<b>Average</b>	<b>-11.49%</b>	<b>-13.66%</b>	<b>-14.17%</b>	
	<b>Best</b>	<b>-22.20%</b>	<b>-22.78%</b>	<b>-23.15%</b>	
	<b>Worst</b>	<b>-5.91%</b>	<b>-6.72%</b>	<b>-5.66%</b>	
	With Skip probability				
	Class A (8bits)	-8.36%	-10.32%	-10.91%	
Class B	-8.99%	-11.67%	-12.22%		
Class C	-14.95%	-16.15%	-16.85%		
Class D	-11.36%	-12.21%	-12.70%		
Class E	-9.94%	-13.69%	-14.21%		
<b>Average</b>	<b>-10.93%</b>	<b>-12.97%</b>	<b>-13.54%</b>		
<b>Best</b>	<b>-21.85%</b>	<b>-22.63%</b>	<b>-23.80%</b>		
<b>Worst</b>	<b>-4.66%</b>	<b>-6.31%</b>	<b>-5.57%</b>		
TBD	Without Skip probability: $c_{k,\tau} = 1 \forall k, \tau$				
	Class A (8bits)	29.10%	40.59%	52.65%	
	Class B	39.40%	60.16%	86.24%	
	Class C	23.85%	37.04%	48.87%	
	Class D	34.79%	51.08%	67.70%	
	Class E	65.03%	85.06%	126.78%	
	<b>Average</b>	<b>38.05%</b>	<b>54.98%</b>	<b>76.84%</b>	
	With Skip probability				
	Class A (8bits)	6.20%	7.64%	9.97%	
	Class B	9.30%	13.49%	20.07%	
Class C	3.75%	4.25%	5.52%		
Class D	5.35%	6.98%	9.39%		
Class E	20.11%	17.93%	23.90%		
<b>Average</b>	<b>8.65%</b>	<b>10.08%</b>	<b>13.98%</b>		

We can observe from Table I higher bitrate savings for  $RDTQ$  (8) and  $RDTQ$  (9) over  $RDTQ$  (7) whether the Skip consideration is enabled or not.  $RDTQ$  (8) saves in average  $-1.80\%$  on top of  $RDTQ$  (7), while  $RDTQ$  (9) has slightly lower performance than  $RDTQ$  (8) with  $-1.58\%$  bitrate reduction. The three Inter probability models suffer from a slight bitrate increase around  $0.31\%$  and  $0.12\%$  when Skip is considered in the optimization.

Regarding the  $TBD$ ,  $RDTQ$  (8) and  $RDTQ$  (9) logically induce higher bitrate deviation, since the two functions tends to propagate more weight back to referenced CUs than  $RDTQ$  (7) (Fig. 2). The introduction of Skip probability considerably reduces the  $TBD$  whatever the Inter probability estimator used, going from  $75.01\%$  to  $17.53\%$  for  $RDTQ$  (9) for instance. We point out that the new Inter probability models together with Skip consideration bring systematic bitrate savings over original model while significantly reducing the  $TBD$ .

The same observations stand for the lowest bitrates, where the relative bitrate savings and the  $TBD$  reduction are even more significant. Focusing on the results provided by  $RDTQ$  (9) with the Skip consideration, BD-BR are improved of  $-2.05\%$  in average, while the  $TBD$  is reduced from  $38.05\%$  to  $13.98\%$ .

## V. CONCLUSION

In this paper, we demonstrate the benefits of considering Skip probability for adaptive quantization, along with the use of accurate Inter probability estimator. First, we provide an Inter probability estimation that better correlates to the ground truth. It brings  $-1.5\%$  to  $-2\%$  BD-BR gain, at the cost of higher target bitrate deviation compared to the reference. In addition, the Skip mode probability consideration allows to better fulfill the bitrate constraint. New analytical solution is given for the optimal quantizers with a  $TBD$  reduced to an acceptable level of  $13.98\%$  while bitrate savings holds; especially for the low bitrates. However, the proposed model still suffers from one minor issue: we assume that the Inter probability is independent from the local quantizer. We should better consider the observation that the more the local quantization step decreases, the more the Intra probability of the CU increases; which cuts off the temporal propagation. This issue will be considered in our future work.

## REFERENCES

- [1] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Processing Magazine*, vol. 15, no. 6, 1998.
- [2] T. Wiegand and B. Girod, "Lagrange multiplier selection in hybrid video coder control," in *IEEE International Conference on Image Processing (ICIP)*, Thessaloniki, October 2001, pp. 542–545.
- [3] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, vol. 22, no. 12, pp. 1649–1668, 2012.
- [4] A. Ortega and K. Ramchandran, "Rate-distortion methods for image and video compression," *IEEE Signal Processing Magazine*, vol. 15, no. 6, pp. 23–50, 1998.
- [5] J. Xie, L. Song, R. Xie, Z. Luo, and X. Wang, "Temporal dependent bit allocation scheme for rate control in HEVC," in *IEEE Workshop on Signal Processing Systems (SiPS)*, 2015.
- [6] Y. Li, H. Jia, X. Xie, and T. Huang, "Rate control for consistent video quality with inter-dependent distortion model for HEVC," in *Visual Communications and Image Processing (VCIP)*, 2016.
- [7] A. Fiengo, G. Chierchia, M. Cagnazzo, and B. Pesquet-Popescu, "Rate allocation in predictive video coding using a convex optimization framework," *IEEE Transactions on Image Processing*, vol. 26, no. 1, pp. 479–489, 2017.
- [8] W. U. Qingbo, J. Xiong, L. U. O. Bing, C. Huang, and X. U. Linfeng, "A novel joint rate distortion optimization scheme for intra prediction coding in H.264/AVC," *IEICE Transactions on Information and Systems*, vol. 97, no. 4, pp. 989–992, 2014.
- [9] M. Bichon, J. Le Tanou, M. Ropert, W. Hamidouche, L. Morin, and L. Zhang, "Inter-block dependencies consideration for intra coding in H.264/AVC and HEVC standards," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017.
- [10] G. Valenzise and A. Ortega, "Improved video coding efficiency exploiting tree-based pixelwise coding dependencies," in *Visual Information Processing and Communication*, 2010.
- [11] M. Winken, A. Roth, H. Schwarz, and T. Wiegand, "Multi-frame optimized quantization for high efficiency video coding," in *Picture Coding Symposium (PCS)*, 2015.
- [12] M. Ropert, J. Le Tanou, M. Bichon, and M. Blestel, "R-D spatio-temporal adaptive quantization based on temporal distortion backpropagation in HEVC," in *IEEE International Workshop on Multimedia Signal Processing (MMSP)*, 2017.
- [13] J. Garrett-Glaser, "A novel macroblock-tree algorithm for high performance optimization of dependent video coding in H.264/AVC," 2011.
- [14] x265, "[Online]. Available: <https://bitbucket.org/multicoreware/x265>."
- [15] F. Bossen, "Common test conditions and software reference configurations," *Tech. Rep. JCTVC-L1100*, Jan. 2013.
- [16] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," in *ITU-T VCEG-M33, Texas, USA, Austin, Texas, Apr. 2001*.