Note of Visually Optimized Two-Pass Rate Control for Video Coding Using the Low-Complexity XPSNR Model

Ushio Sorakado*

***Abstract*—This document is the note of the paper named: "Visually Optimized Two-Pass Rate Control for Video Coding Using the Low-Complexity XPSNR Model", which has been accepted at the IEEE Int. Conf. on Visual Communications and Image Processing (VCIP), Munich, in December 2021. We access the paper via the following link: https://ieeexplore.ieee.org/document/9675364/. The paper propose a kind of two-pass rate control (RC) scheme and implement it on VVenC-1.0.0 [1]. Experimental results prove that the proposed method achieves accurate rate control with negligible decrease in rate-distortion (R-D) performance.**

***Index Terms*—Two-pass rate control (RC), Visual quality assessment (VQA), quality of experience (QoE).**

I. Introduction

The work of [2] aims to study the combination of two-pass rate control (RC) and direct perceptual encoding optimization controlled by a low-complexity psycho-visual model.

The contribution of [2] is summarized as follows.

Theoretical support of the method:
Due to camera sensor noise in high-resolution video recordings, the relationship between bit rate and quantization parameter (QP) is significantly different in high and low bit rate scenarios. Specifically, in low bit rate scenarios, the logarithm of the bit rate change appears to be almost linearly related to QP and λ. However, the increase in bit rate rises significantly as the QP becomes smaller, in high bit rate scenarios. Thus, the R-Q or R-λ model should combine two part, where the first part describes its behaviour in low bit rate scenarios, while the second part is designed to deal with high bit rate scenarios. In this paper, the first pass rate control determined preliminary QP $QP^{'}$ for the second pass rate control based on the first part of R-Q model, and the sceond part of R-Q model is used to adjust $QP^{'}$ to the corrected QP $QP^{''}$ in the second pass rate control.

Two-Pass rate control method:
In the first pass rate control, fix QP QP_{base} is used for encoding. QP_{base} is required for e.g. loop filter initialization and QP delta-coding. The calculation of QP_{base} is defined as:

$$QP_{base} = round\left(40 - \sqrt{\frac{3840 \cdot 2160}{W \cdot H} \cdot \frac{R_{target}}{500000}}\right). \tag{1}$$

After the first pass encoding, the bit consumption of each frame so-called r_f is recorded. Then, the QP of each frame is caculated from:

$$QP_f^{'} = QP_f - c_{low} \cdot \sqrt{max(1; QP_f)} \cdot log_2\left(\frac{r_f^{'}}{r_f}\right), \tag{2}$$

in the second pass, where c_{low} is set as 0.82, $QP_f^{'}$ and QP_f are the preliminary second-pass QP and integer first-pass QP of f-th frame, while $r_f^{'}$ and r_f are target second-pass and resulting first-pass bit counts of f-th frame. Specifically, the second-pass per-frame target bit counts $\widehat{r}_f$ are initialized as

$$\widehat{r}_f = round\left(r_f \cdot \frac{R_{target} \cdot F}{fps \cdot \sum r_f}\right), \tag{3}$$

Then, before final encoding of each frame, $\widehat{r}_f$ is refined to

$$r_f^{'} = max\left(1; \widehat{r}_f + \left(\sum \widehat{r}_c - r_c^{''}\right) \cdot d \cdot \frac{r_f}{g_f}\right), \tag{4}$$

to better match the target bit rate as the encoding progresses,where C is the set of all frames already encoded in the second pass and $r_c^{''}$ is the final bit consumption of each frame c in C. Constant d equals 1 for all f in the last encoded GOP, else 0.5. g_f is the sum of all bits r in the GOP to which f belongs. In addition, the preliminary base QP $QP_{base}^{'}$ of second-pass rate control is initialized as

$$QP_{base}^{'} = QP_{base} - c_{low} \cdot \sqrt{QP_{base}} \cdot log_2\left(\frac{R_{target} \cdot F}{fps \cdot \sum r_f}\right) \tag{5}$$

The work of [2] further corrects QP from $QP_f^{'}$ to $QP_f^{''}$ based on:

$$QP_f^{''} = QP_f^{'} + c_{high} \cdot max\left(0; QP_{start} - QP_f^{'}\right), \tag{6}$$

where QP_{start} is the correction threshold and $0 \leq c_{high} \leq 1$ serves as a parameter to control the strength of the correction. Specifically, $c_{high} = \frac{1}{8} \cdot max\left(0; round\left(log_2 H\right) - 7\right)$. Similarly, $QP_{base}^{''}$ is caculated from $QP_{base}^{'}$ based on (6).

The Combination of Two-Pass rate control and perceptual rate-distortion optimization:
As the proposed two-pass rate control algorithm allocates QP at frame level, the low-complexity XPSNR model [3] is adopted to allocates QP offset at coding-tree-unit (CTU) level to optimize perceptual rate-distorion (R-D) performance. Specifically, in both pass rate control, the QP offset ΔQP_k of the CTU with index k will be calculated and added with corresponding frame-level QP. The formulation of ΔQP is given by

$$\Delta QP_k = -round\left(3 \cdot log_2\left(\omega_k\right)\right), \tag{7}$$

where ω_k is the perceptual weight of k-th CTU, and the calculation of it is omitted here for brevity. Then, combined with corresponding frame-level QP QP_f and $lambda\ lambda_f$, the QP and λ of k-th CTU is calculated as

$$QP_k = QP_f - \Delta QP_k, \qquad \lambda_k = \frac{\lambda_f}{\omega_k}, \tag{8}$$

REFERENCES

[1] Adam Wieckowski, Jens Brandenburg, Tobias Hinz, Christian Bartnik, Valeri George, Gabriel Hege, Christian Helmrich, Anastasia Henkel, Christian Lehmann, Christian Stoffers, Ivan Zupancic, Benjamin Bross, and Detlev Marpe, "Vvenc: An open and optimized vvc encoder implementation," in *Proc. IEEE International Conference on Multimedia Expo Workshops (ICMEW)*, 2021, pp. 1–2.

[2] Christian R Helmrich, Ivan Zupancic, Jens Brandenburg, Valeri George, Adam Wieckowski, and Benjamin Bross, "Visually optimized two-pass rate control for video coding using the low-complexity xpsnr model," in *2021 International Conference on Visual Communications and Image Processing (VCIP)*. IEEE, 2021, pp. 1–5.

[3] Christian R Helmrich, Sebastian Bosse, Heiko Schwarz, Detlev Marpe, and Thomas Wiegand, "A study of the extended perceptually weighted peak signal-to-noise ratio (xpsnr) for video compression with different resolutions and bit depths," *ITU Journal: ICT Discoveries*, vol. 3, no. 1, pp. 65–72, 2020.