VISUALLY OPTIMIZED TWO-PASS RATE CONTROL FOR VIDEO CODING USING THE LOW-COMPLEXITY XPSNR MODEL

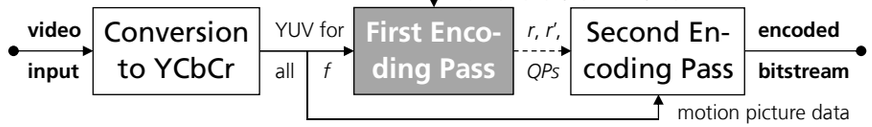
VCIP 2021

Christian R. Helmrich, Ivan Zupancic, Jens Brandenburg, Valeri George, Adam Wieckowski, and Benjamin Bross – Fraunhofer HHI

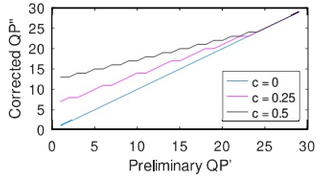
INTRODUCTION

- Rate control (**RC**) methods required when generating compressed video catalogs
- Used by streaming service providers & traditional broadcasters for online content
- Goal: encode such that target bitrate R_T is achieved, on average, across the video
- Videos are often encoded (or transcoded) offline, allowing two-pass RC schemes
- Two-pass RC typically offers better performance than a single-pass RC equivalent
- Fast 1st RC pass for analysis, 2nd RC pass for final encoding based on 1st-pass data

RC functional flowchart



Effect of corrective second *R-QP* step



MOTIVATION

- Few studies of visual quality optimization combined with RC, especially two-pass
- No published two-pass RC design for Versatile Video Coding (**VVC**, GOP[1] size 32)
- Existing RC methods were found to perform sub-optimally in the context of VVC

CONTRIBUTIONS

- Simple and implementation friendly two-step *R-**QP*** (instead of *R*-λ) model for RC
- Use of XPSNR low-complexity psychovisual model for visual QP adaptation, **SCD**[2]

1: group of pictures, 2: scene cut detection, 3: https://github.com/fraunhoferhhi/vvenc, 4: SDR common test conditions, extended to 10 s

TWO-STEP R-QP MODEL replaces traditional hyperbolic $\lambda = \alpha \cdot r^{\beta}$ model

- 1st RC pass: encode with $QP_{\text{base}} = \text{round}\left(40 - \sqrt{\frac{3840 \cdot 2160 \cdot R_T}{W \cdot H \cdot 500000}}\right)$ and derived QP_f, λ_f
- 2nd pass: set $QP'_{\text{base}} = QP_{\text{base}} - \check{c} \cdot \sqrt{QP_{\text{base}}} \cdot \log_2\left(\frac{R_T \cdot F}{fps \cdot \sum_f r_f}\right)$ and QP'_f with 1st-pass r_f
- 2nd model step: correct QP'_{base} and QP'_f as in left figure; set $\lambda''_f = \lambda_f \cdot 2^{(QP''_f - QP_f)/3}$

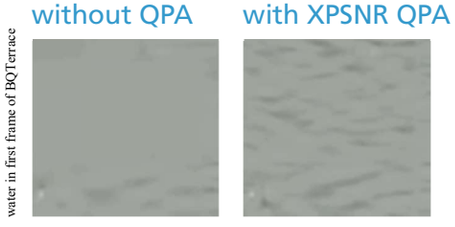
1st model step: $QP''_f = \text{round}\left(QP'_f + \hat{c} \cdot \max\left(0;\ QP_{\text{start}} - QP'_f\right)\right)$

XPSNR MODEL

- CTU-level QP adaptation based on visual sensitivity measure of **XPSNR** algorithm
- Ratio between successive frame visual activities usable for **SCD**[2]. See also tutorial

EXPERIMENTAL RESULTS

- Integrated & tested in open VVC encoder **VVenC**[3], random-access configuration
- Results measured in Bjøntegaard delta-rate (BD-R), JVET's set (CTC[4]) + HHI videos
- **Subjective** quality gain via XPSNR, **objective** BD-R gains; timing relative to VTM



relative to fixed-QP equivalent encoding. Gains of 4 – 12 percentage

Resolution	VTM 12, no QPA		VVenC, no QPA		VVenC, vis. QPA	
Class	PSNR	Runtime	PSNR	Runtime	XPSNR	Runtime
UHD A½	11.2%	98.1%	0.10%	9.60%	0.57%	9.61%
UHD HHI	9.13%	101%	2.30%	9.12%	4.30%	10.1%
HD B	6.31%	105%	0.54%	9.34%	1.36%	9.50%
HD HHI	14.0%	109%	1.72%	9.99%	2.47%	11.1%
SD C	4.66%	102%	0.23%	10.1%	0.51%	10.5%

Outlook: modifications for live streaming operation

Fraunhofer