ATHENA UNIVERSITÄT KLAGENFURT

CODA: Content-aware Frame Dropping Algorithm for High Frame-rate Video Streaming

Vignesh V Menon, Hadi Amirpour, Mohammad Ghanbari, Christian Timmerer
Christian Doppler Laboratory ATHENA, Institute of Information Technology (ITEC), University of Klagenfurt, Austria

Data Compression Conference (DCC)
24 March 2022

Outline

Introduction

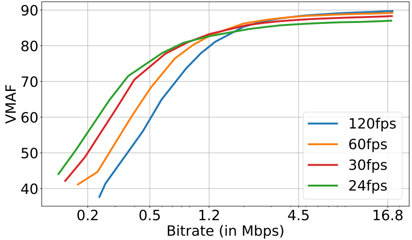
Introduction

High Framerate (HFR) videos

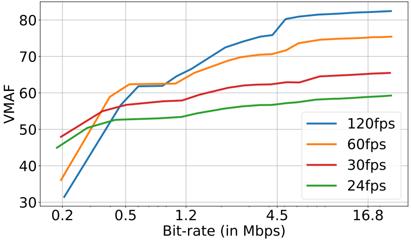
- *High Framerate* (HFR) video streaming enhances the viewing experience and improves visual clarity.[1]
- However, it may lead to an increase of both encoding time complexity and compression artifacts, particularly at lower bitrates.

[1] ITU-R BT.2020-2. "Parameter values for ultra-high definition television systems for production and international programme exchange". In: 2015.

Introduction



(a) HoneyBee



(b) Lips

Figure: Rate-Distortion (RD) curves of UHD encodings of (a) *HoneyBee* and (b) *Lips* sequences for multiple framerates.

Introduction

Variable Framerate (VFR) coding

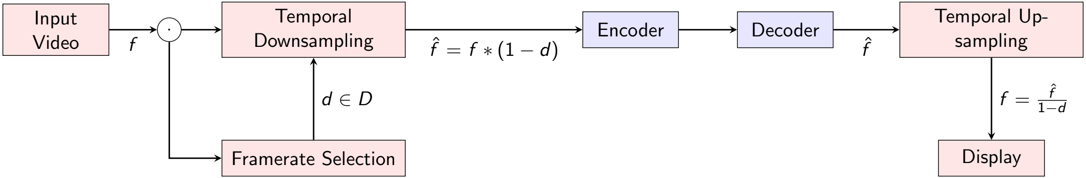


Figure: Block diagram of a variable framerate (VFR) coding scheme[2] in the context of video encoding. f and $\hat{f}$ denote the original framerate of the video and the framerate at which the video is encoded. d represents the frame dropping factor.

[2] G. Herrou et al. "Quality-driven Variable Frame-Rate for Green Video Coding in Broadcast Applications". In: *IEEE Transactions on Circuits and Systems for Video Technology* (2020), pp. 1–1. DOI: 10.1109/TCSVT.2020.3046881.

CODA

CODA

Phase 1: Feature Extraction

Compute texture energy per block

A DCT-based energy function is used to determine the block-wise feature of each frame defined as:

$$H_k = \sum_{i=1}^{w} \sum_{j=1}^{h} e^{\left|\left(\frac{ij}{wh}\right)^2 - 1\right|} \left|DCT(i-1, j-1)\right| \quad (1)$$

where w and h are the width and height of the block, and $DCT(i,j)$ is the (i,j)th DCT component when $i + j > 2$, and 0 otherwise.

The energy values of blocks in a frame is averaged to determine the energy per frame.[3]

[3] Michael King, Zinovi Tauber, and Ze-Nian Li. "A New Energy Function for Segmentation and Compression". In: *2007 IEEE International Conference on Multimedia and Expo*. 2007, pp. 1647–1650. DOI: 10.1109/ICME.2007.4284983.

Proposed Algorithm

Phase 1: Feature Extraction

h_k: SAD of the block level energy values of frame k to that of the previous frame $k-1$.

$$h_k = \frac{SAD(H_k(i) - H_{k-1}(i))}{M} \tag{2}$$

where M denotes the number of CTUs in frame k.

CODA

Phase 2: Framerate prediction

Inputs:

E, h : spatial and temporal complexities

r : video resolution

f_{max} : original framerate

D : set of all frame drop factors $\tilde{d}$

B : set of all target bitrates b (in kbps)

Output: $\hat{f}(b) \ \forall \ b \in B$

Step 1: Determine $\hat{d}(b)$.

$$\hat{d}(b) = d_0 e^{-\frac{\beta_{MA}(r, f_{max}) \cdot h \cdot b}{E}}$$

Step 2: The predicted optimized framerate for the video is given by:

$$\hat{f}(b) = f_{max} \cdot (1 - \hat{d}(b))$$

Evaluation

Evaluation

E and h values are extracted using VCA open-source software.[4]

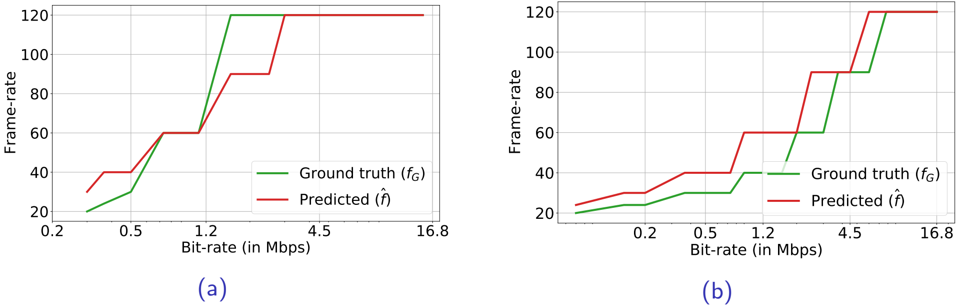
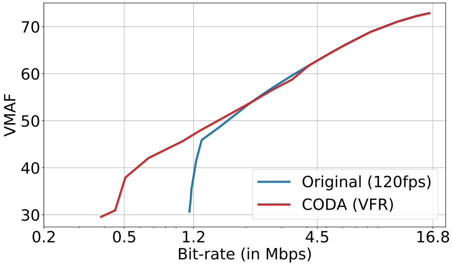


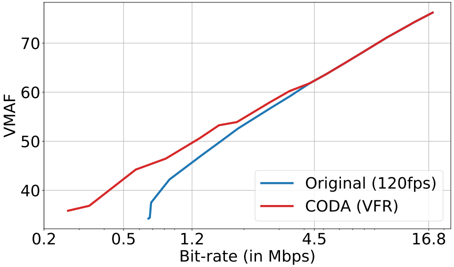
Figure: Optimized framerate prediction results of *Beauty* (a) and *ShakeNDry* (b) sequences. Please note that, depending on the content, the optimized framerate in various bitrates are different.

[4] V. V Menon, C. Feldmann, and H. Amirpour. *VCA*. Version 1.0.0. Available: https://github.com/cd-athena/VCA. Feb. 2022.

Evaluation



(a)



(b)

Figure: Rate-Distortion (RD) results of *Beauty* (a) and *ShakeNDry* (b) sequences for the default encoding and *CODA*-based VFR encoding.

Conclusions and Future Directions

Conclusions

- Presented a content-aware frame dropping algorithm for video streaming applications, especially for HFR videos.
- Predicts the optimized framerate for a set of bitrates defined in a bitrate ladder and resolution for every video, which helps in improving the overall performance of HFR video streaming in terms of encoding time and compression efficiency.
- UHD encoding using the proposed algorithm requires 15.87% fewer bits to maintain the same PSNR and 18.20% fewer bits to maintain the same VMAF as compared to the original framerate encoding. An overall encoding time reduction of 21.82% is also observed.

Q & A

Thank you for your attention!

Vignesh V Menon (vignesh.menon@aau.at)
Hadi Amirpour (hadi.amirpourazarian@aau.at)
Mohammad Ghanbari (ghan@essex.ac.uk)
Christian Timmerer (christian.timmerer@aau.at)