Politecnico
di Torino
1859

Master degree program
in Cinema and Media Engineering

AY 2022-2023
March-April graduation session 2023

Master Thesis Presentation

The Dynamic Optimizer Framework

Video encoding, assessment and comparison

Author
Chemin Davide

Supervisor
Masala Enrico

Background

- From its introduction, in the mid-eighties, **digital video** has undergone a remarkable development up to the present day. Significant advances has been made in the research and development of coding technologies and systems.
- Video compression is essential for saving of **storage** space, the compliance with **bandwidth limits** and the reduction of **delivery costs**. For these reasons OTT services providers are always looking towards more and more efficient compression techniques.
- **Coding standards** evolve in this direction. New released codecs ensure better compression rates for the same perceptual quality compared to past versions, as with HEVC and the new VVC.
- We analysed a new approach to bitrate reduction - or quality gain - originally conceived by Netflix and based on the **optimization** of the encoding parameters without any improvement on the existing encoders.
  - ▸ *Katsavounidis, I. (2018), "Dynamic optimizer – a perceptual video encoding optimization framework", Netflix tech blog.*

Video coding and compression

Digital video compression

- Lossless
  - Sampling
  - Run-lenght coding
  - Entropy coding
  - Prediction

- Lossy
  - Quantization
  - DCT (JPEG standards)
  - Motion Compensation (MPEG standards)

Video coding standards

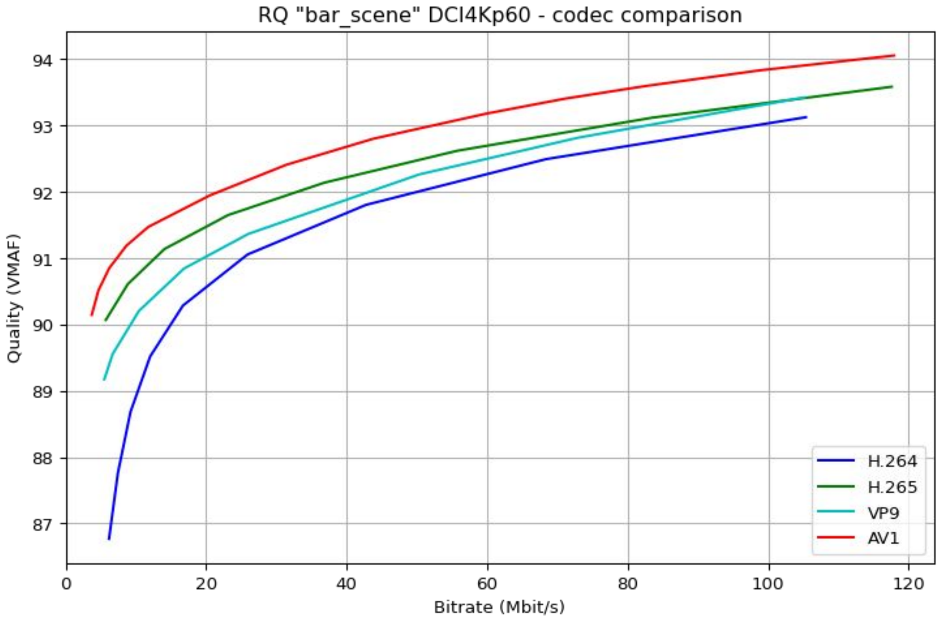
Representative codecs comparison

	x264	**x265**	**VP9**	**AV1**
x264	0	22.22	32.49	73.58
x265	-17.73	0	7.28	39.97
VP9	-23.82	-6.72	0	31.54
AV1	-41.96	-28.54	-23.96	0

Average bitrate savings in percent for the selected codecs.
doi: 10.1109/MCNA50957.2020.9264275

Encoders assessment

The Rate-Distortion curve



01. Rate and **bitrate control**

The **Constant Rate Factor** (CRF) allows to attempt to achieve a certain output quality having the bitrate fluctuating as necessary without particular constraints.

02. Distortion and **quality metrics**

Subjective, like MOS, and objective, like MSE, PSNR, SSIM and **VMAF**.

Dynamic optimization

Netflix Research | Home | Research Areas | Business Areas | Publications | Search Research Site

Video codec comparison using the dynamic optimizer

Streaming | Encoding & Quality | Publication

August 15, 2018

Authors

Ioannis Katsavounidis, Senior Research Scientist

Liwei Guo , Senior Software Engineer

Abstract

We present a new methodology that allows for more objective comparison of video codecs, using the recently

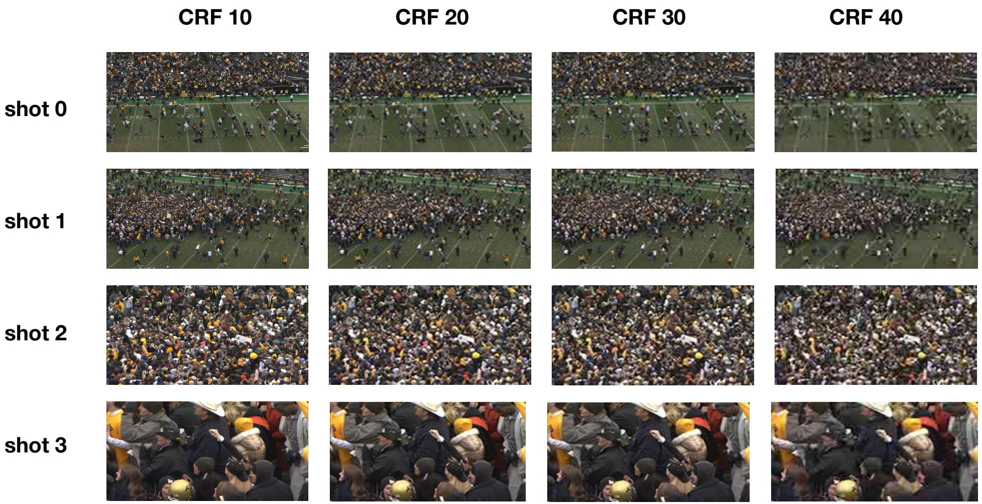
Dynamic optimization

Introduction (1)

1. The video sequence is split into **shots**, also called coding units;

2. Each shot is encoded using a number of quantization parameters or **CRF values**, based on the encoder capabilities. These are called elemental encodes.
   - *For example, the H.264 ffmpeg encoder allows to encode at 52 CRF points within the the range [0-51].*

3. Each elemental encode is assessed using the **VMAF quality metric**;



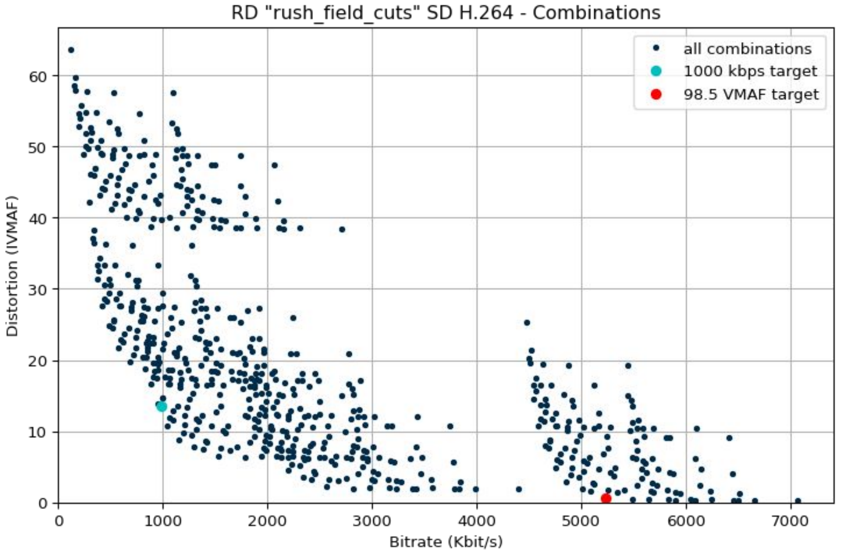
Coding units and elemental encodes.

Dynamic optimization

Introduction (2)

4. For each shot, we can identify a set of **RQ points**: the bitrate (bit/s) and quality values (VMAF) pairs. They produce the RQ curve of a shot, or the **RD curve**, if we convert quality to distortion (IVMAF = 100 - VMAF).

5. This cloud of points is used to find the **optimal combination** of elemental encodes able to:
   a. Minimize the distortion for a given target rate or,
   b. Minimize the bitrate for a given target quality.



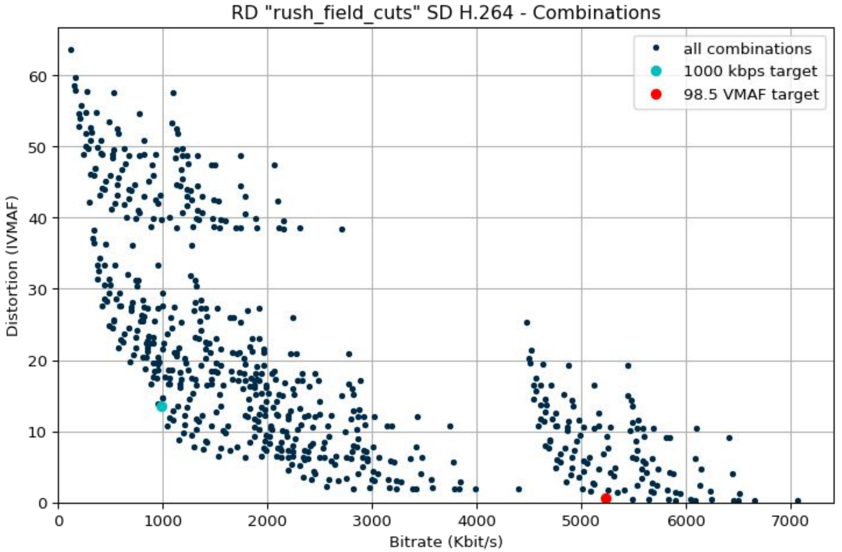
Rate-Distortion possible combinations with the optimal solutions for the specified rate and quality targets.

Implemented methods

	01. **Brute force approach**	**02. Lagrangian optimization** with **exhaustive coding**	**03. Lagrangian optimization** with **curve fitting**
computational complexity	Exponential	Low	Low
coding costs	High	High	Low
solution approximation	No	Low	High

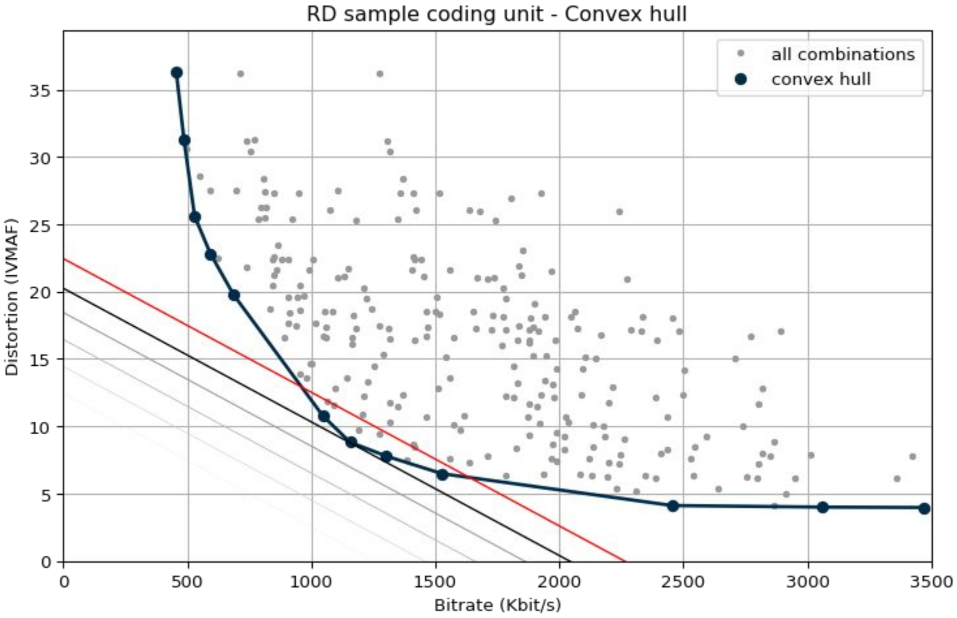
01. Brute force approach - BF

1. Encode each shot at all CRFs values;
2. Get the RD values for each elemental encode;
3. Create all **possible combinations** of elemental encodes for the entire sequence;
   - *For example, a 5-shot scene encoded at 8 points produces 8^5 combinations*
4. Find the optimal solution: **iterate all options** to find the one that:
   a. Maximize the quality (or minimize distortions) without exceeding the bitrate target or,
   b. Minimize the rate without falling below the quality target.



Rate-Distortion possible combinations with the optimal solutions for the specified rate and quality targets.

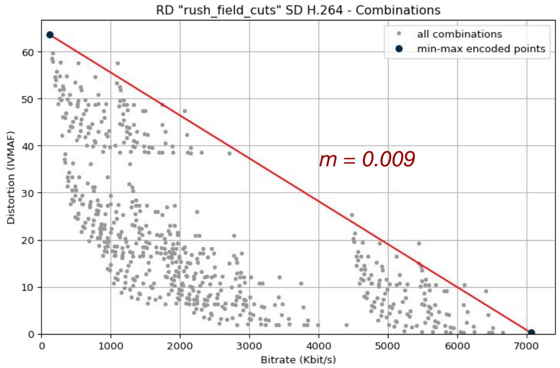
02. Lagrangian optimization - LG



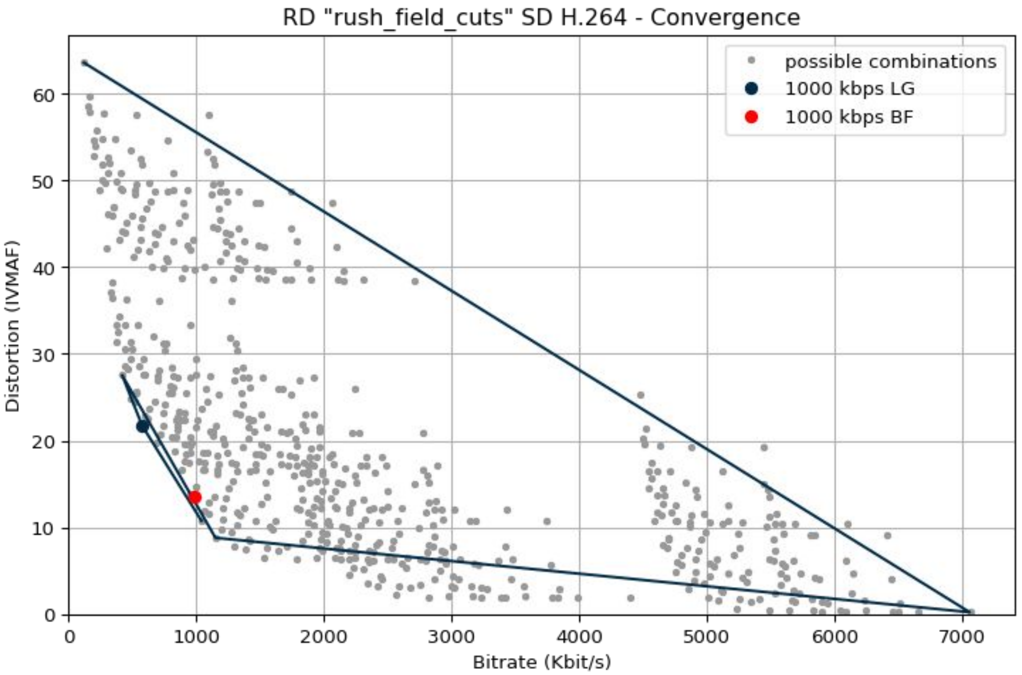
- According to Lagrange theory, it is possible to optimize the encoding parameters - in this case the choice of the best CRF value - for **each shot independently** of each other.
- The **convex hull** represents the achievable compression performance for the given video sequence.

02. Lagrangian optimization - LG

1. Encode each shot at all CRFs values;
2. Get the RD values for each elemental encode;
3. Get the RD values of the entire sequence at the minimum and maximum CRF value;
   - *For example, all shots encoded at 10 and 45 in the CRF range [10,45]*
4. Assuming to place these points in the Cartesian plane, **compute the slope** of the line that intersects them, called total slope;
5. For each shot, find the elemental encode with the slope of the **tangent** to the convex hull closest to the total slope;
6. Check whether the new combination is **exceeding the target** or not and update the minimum or maximum RD sequence values and its slope accordingly;
7. Iterate points 5 and 6 until **convergence.**
   - *For example, when the new combination is the same as the last one*



02. Lagrangian optimization - LG

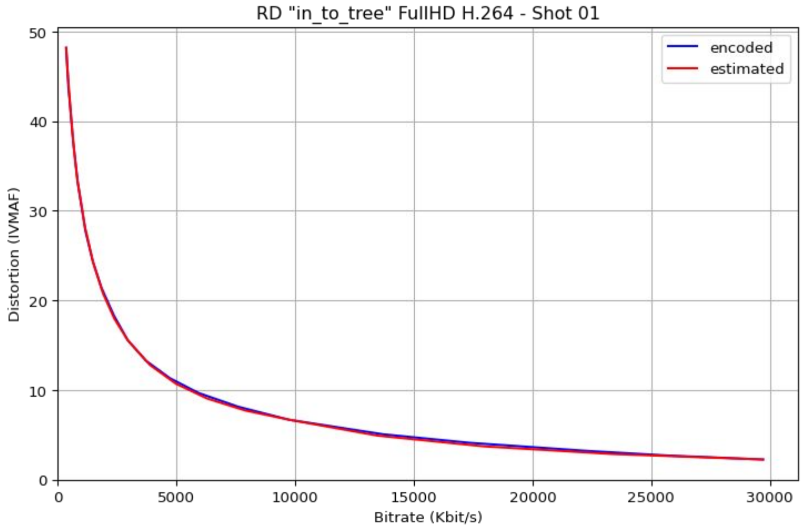


Convergence in the lagrangian approach. The total slope changes because it's endpoints move in reaching the optimal solution.

Unlike the brute force method, the lagrangian only considers solutions that lie on the convex hull.

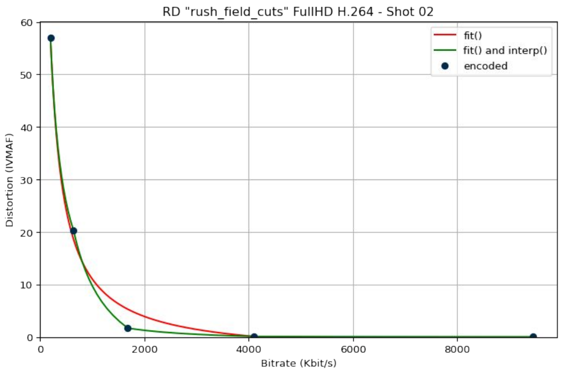
03. Lagrange with curve fitting - CF

1. Encode each shot at a limited number of CRFs values;
   - *For example, just 5 or 6 in the CRF range [10,45]*
2. Get the RD values for each elemental encode;
3. **Curve fitting**: estimate an interpolating curve for these points;
4. Estimate the position of the missing points along the curve according to the curve distribution function. For a better fitting to real values, move these points **compressing or expanding** them having the few computed as a reference;
5. Repeat steps from 3 to 7 of the Lagrangian optimization to reach the convergence.



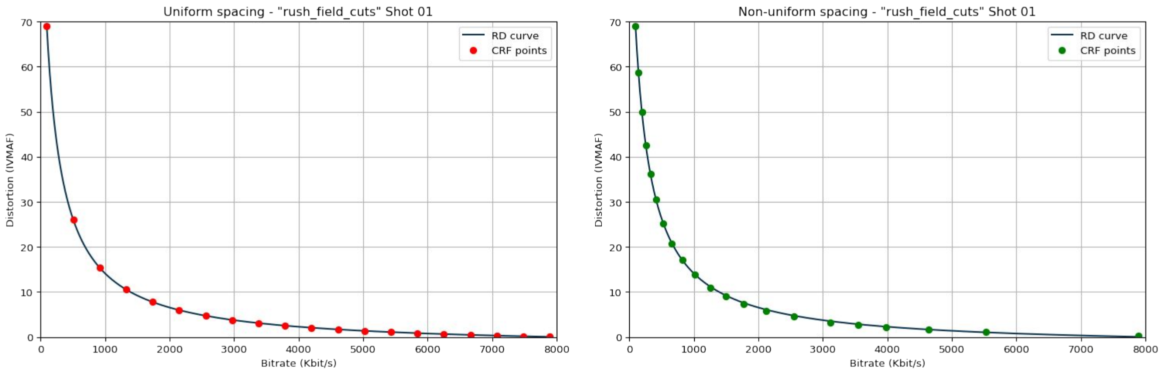
03. Lagrange with curve fitting - CF

Interpolation



Interpolation with the actual values.

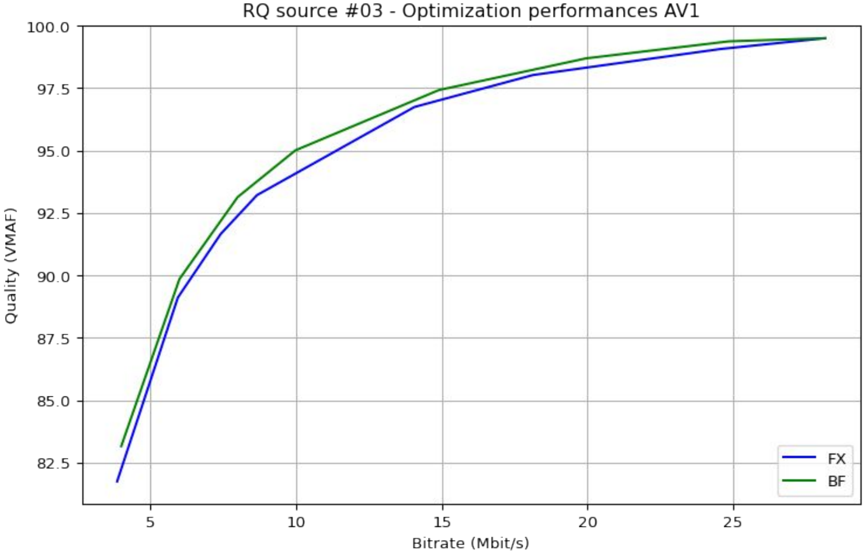
Spacing

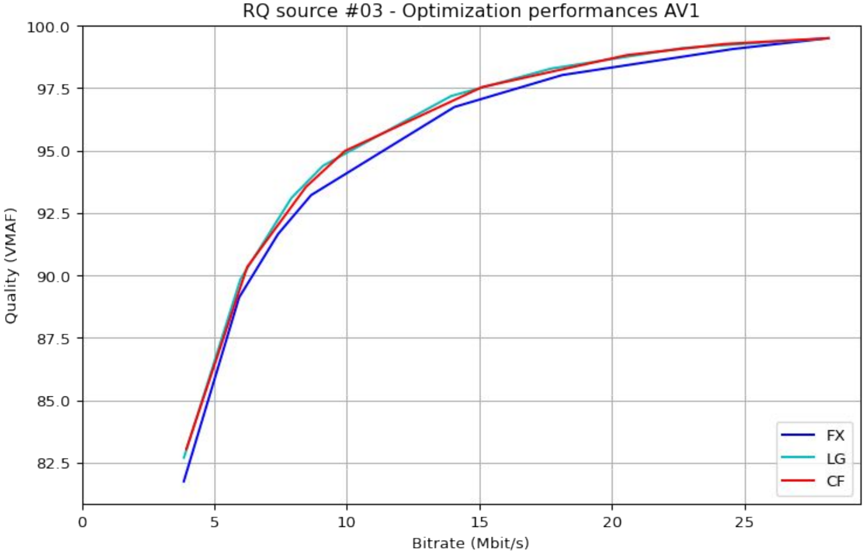


CRF points spaced according to the distribution function of the estimated RD curve.

Assessment and results - 01

Methods comparison

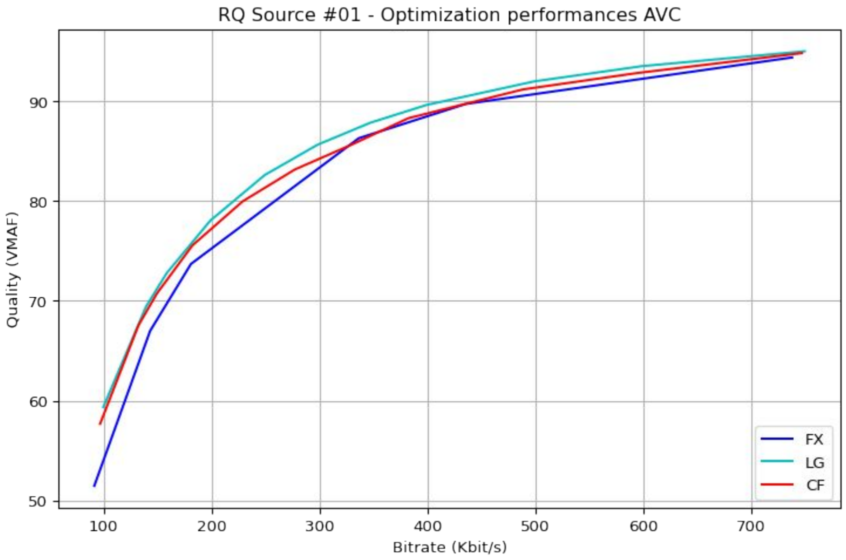




FullHD sequence of 4 detected shots, encoded in AV1 at all CRF points in the range [15,40]. 5 points instead for CF.

Assessment and results - 02

Methods comparison



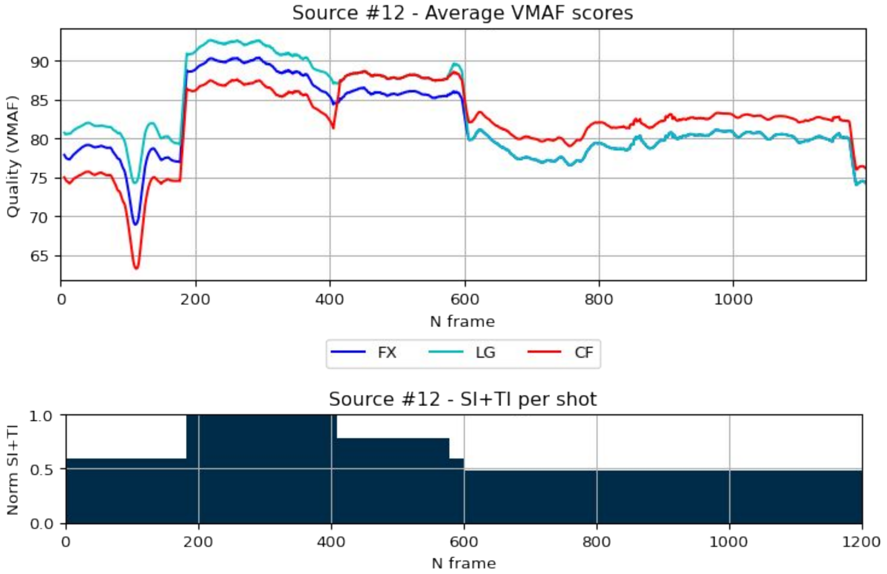
SD source, 126 detected shots.

4K source, 46 detected shots.

Assessment and results - 03

Tested conditions

1. Optimization methods performances
2. Number of elemental encodes per shot
3. Coding standards
4. Distortion metric
5. Content complexity
6. Synthetic and natural content



4K sequence of 5 detected shots. At the top, averaged per-frame VMAF scores over time. At the bottom, Spatial and Temporal Information scores, a measure of the content complexity.

Key findings

Methods comparison

Politecnico di Torino

01. Brute force approach

- Not suitable for long sequences with many shots and many elemental encodes or CRF points.
- Mainly used for benchmarking. It returns the best optimal solution, used to assess the performances of the other two solutions.

02. Lagrangian optimization with **exhaustive coding**

- The lagrangian solution is a satisfactory approximation when the points of the convex hull are dense enough.
- It does not guarantee optimality but it is very close to optimal performance.
- Suitable when time and computational costs are not relevant.

03. Lagrangian optimization with **curve fitting**

- The estimation can differs from actual values, therefore results might not be as accurate as the other techniques.
- It best performs when the number of elemental encodes is low, hence allowing a great reduction in the number and the costs for encoding.

Master degree program
in Cinema and Media Engineering

AY 2022-2023
March-April graduation session 2023

Open-source release of the **software** implementation, made of python scripts for the optimization procedures and FFmpeg commands for the encoding and assessment of the video files.

- *Dynopt - Github repository*

Author
Chemin Davide

Supervisor
Masala Enrico

Master degree program
in Cinema and Media Engineering

AY 2022-2023
March-April graduation session 2023

Master Thesis Presentation

The Dynamic Optimizer Framework

Video encoding, assessment and comparison

Author
Chemin Davide

Supervisor
Masala Enrico