LCEVC x264 Report: Live Sports & eGames, ABR ladder

Jan Ozer, April 2021

APPENDIX

Table of Contents

Core Report (separate document)

Appendix (this document)

Appendix I: Detailed VMAF and MOS results

- VMAF and MOS BD-Rates
- Individual clips results - eGames
- Individual clips results - Sports

VMAF and MOS BD-Rates

(negative = LCEVC is better)

eGames ABR ladder

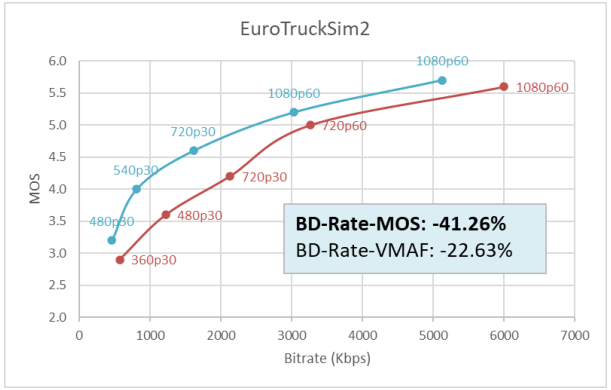
#	Content	BD-Rate-Vmaf	BD-Rate-MOS
1	EuroTruckSim2	-22.63%	-41.26%
2	fallout4	-20.09%	-
3	GTAV	-29.45%	-
4	GTAV2	-48.94%	-37.92%
5	minecraft	-25.15%	-50.40%
6	RUST	-33.38%	-
7	starcraft	-51.07%	-73.00%
8	WITCHER3	-25.66%	-48.08%
Total eGames		**-33.08%**	-
Total Subjectively tested		**-35.08%**	**-54.24%**

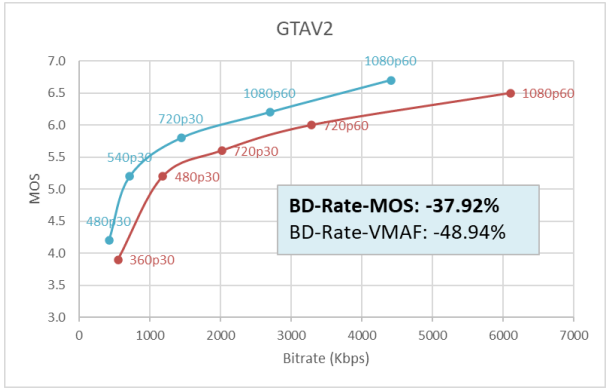
Sports ABR ladder

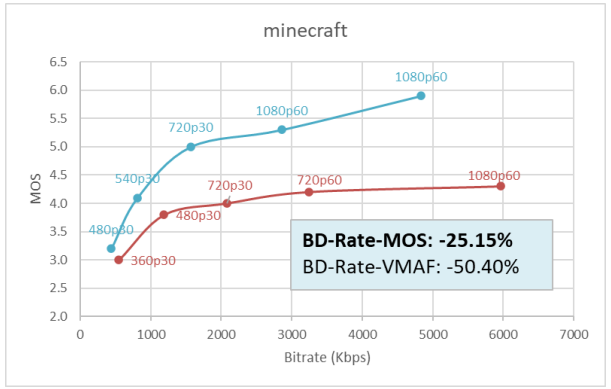
#	Content	BD-Rate-Vmaf	BD-Rate-MOS
1	Rugby	-44.94%	-
2	Soccer-Diego	-22.80%	-
3	F1	-41.53%	-44.39%
4	Horserun	-24.96%	-36.87%
5	Mountainbike	-9.08%	-27.81%
6	Skateboard	-33.07%	-
7	Sports_2_Football	-24.62%	-
8	New_Soccer_Close	-34.84%	-
9	New_Soccer_Mid	-37.54%	-41.76%
10	ElFuente_Box	-31.72%	-53.07%
11	Wimbledon	-47.53%	-49.54%
12	xph_speed_bag	-40.64%	-
Total Sports		**-38.17%**	-
Total Subjectively tested		**-32.34%**	**-44.04%**

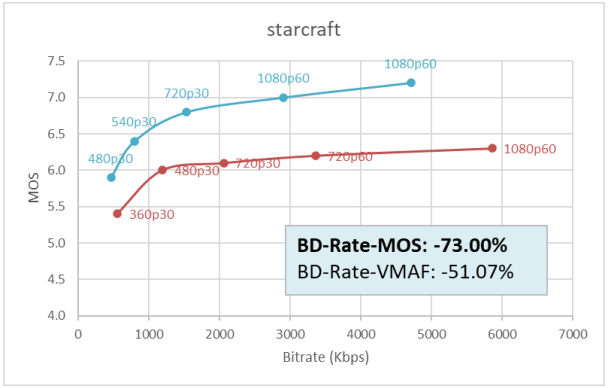
eGames – MOS results

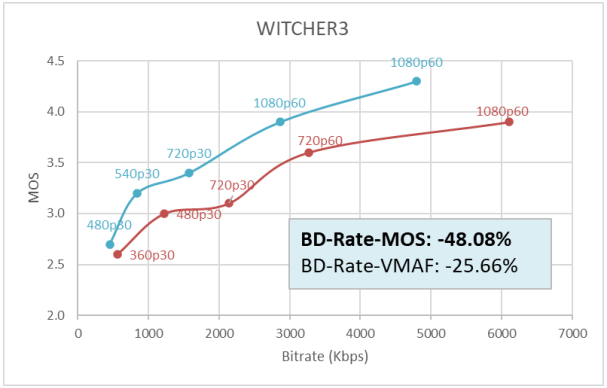
LCEVC x264
x264



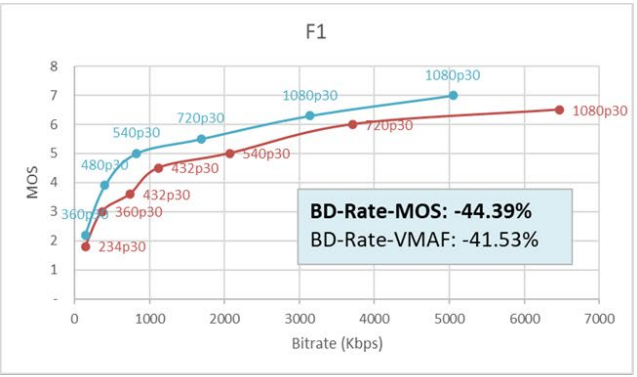


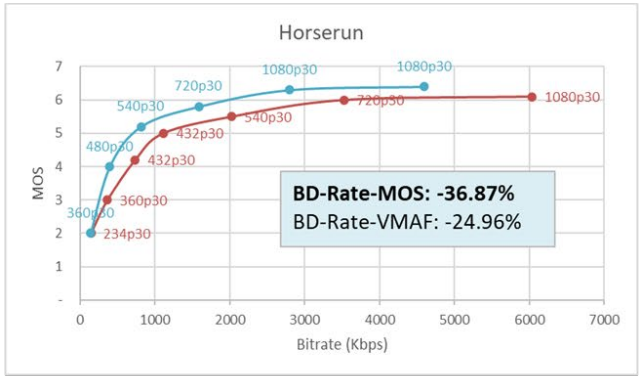


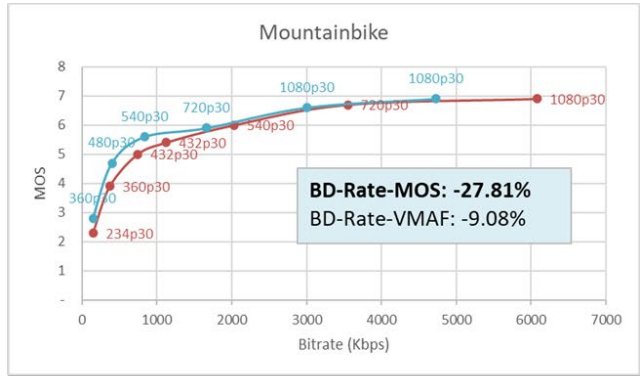


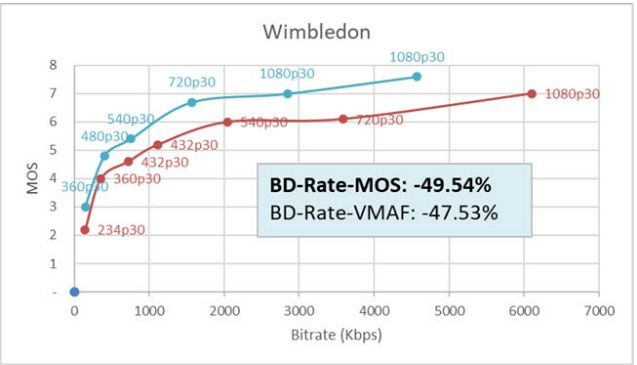


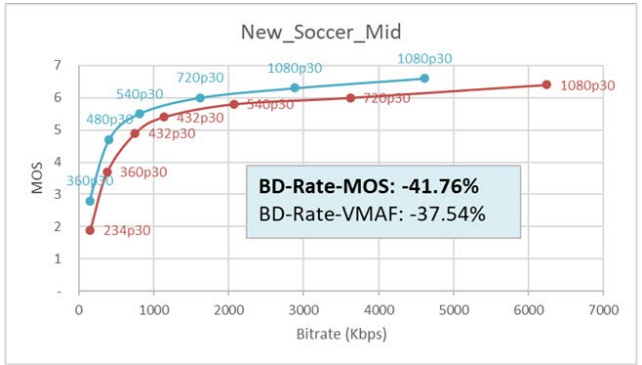
Sports – MOS results

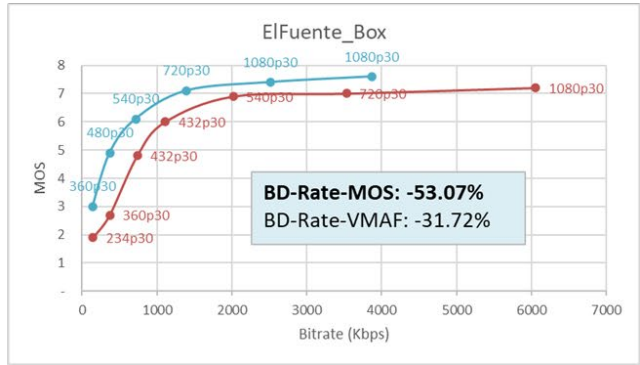












LCEVC x264
x264

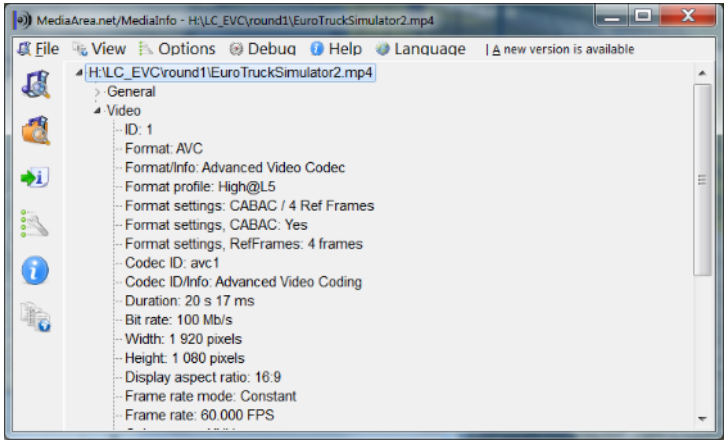
Appendix II: Test clip description

- eGames
- Sports

eGames

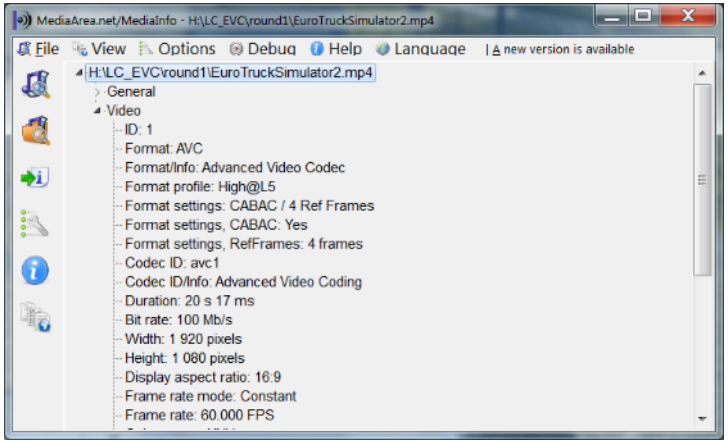
- EuroTruck Simulator 2
- Fallout4
- GTAV
- GTAV2
- Minecraft
- Rust
- Starcraft
- Witcher3

EuroTruck Simulator 2 Description



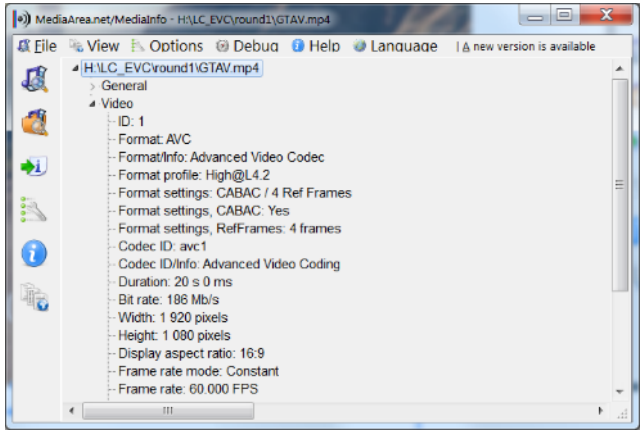
- Highly realistic, finely detailed, and well lit
- Converted from original source so high quality input

Fallout4 Description



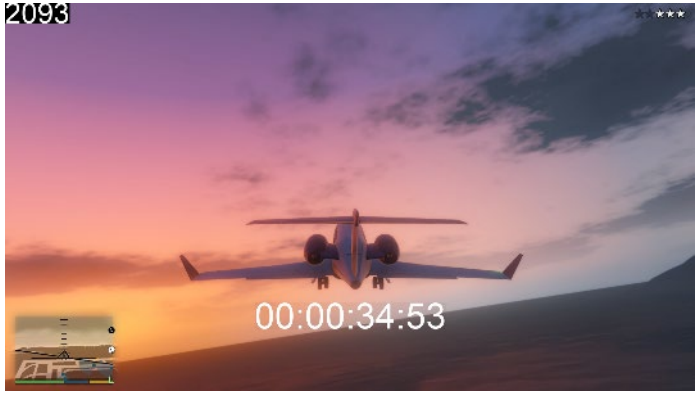
- First person shooter; busy screen but not a lot of fine detail
- Converted from original source so high quality input

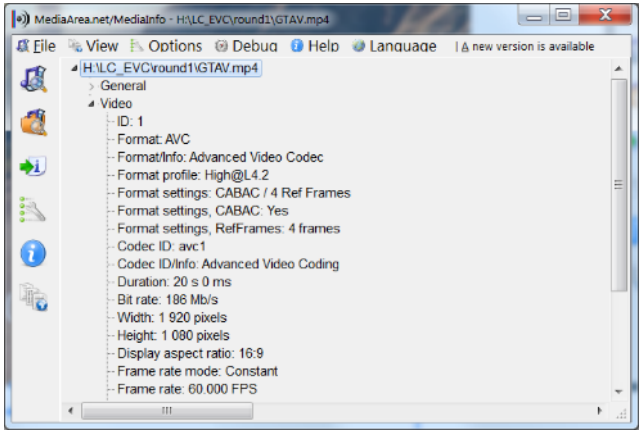
GTAV Description



- Classic car chase game
- Very high quality source

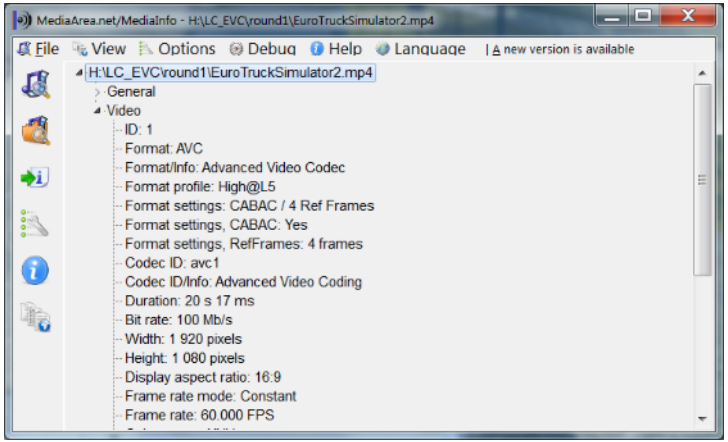
GTAV2 - Description





- Classic car chase game
- Primarily dark and low detail airplane sequence

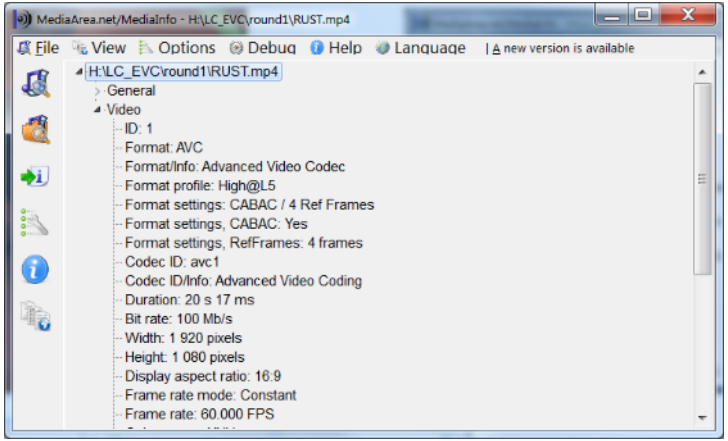
Minecraft Description



- Blocky and low detail but fast moving game

- Converted from original source so high quality input

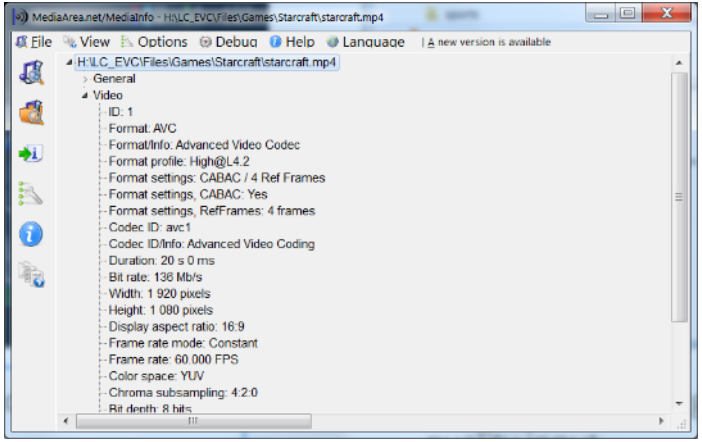
Rust Description



- Dark, not too highly detailed first-person game

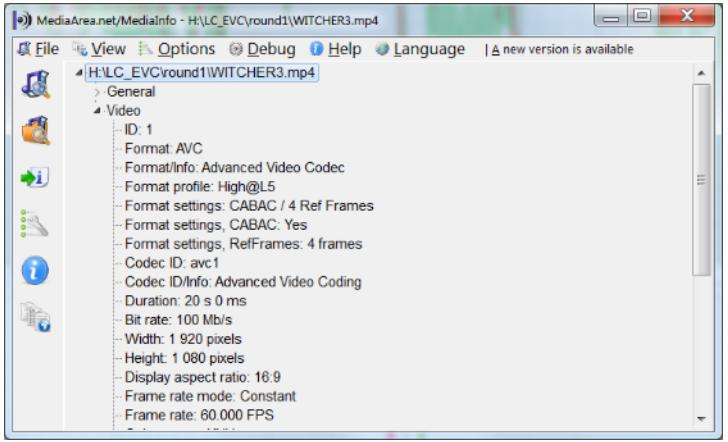
- Taken from original source so very high quality input
- Brightened input before output to improve visibility

StarCraft Description



- Dark, 2D game with lots of tiny moving parts
- Converted from original source so high quality input

Witcher3 Description



- Low detail, high action game
- Captured from YUV source so input quality was very high

Sports

- Rugby
- Soccer-Diego
- Formula 1
- Horserun
- Mountainbike
- Skateboard
- Football
- New_Soccer_Close
- New_Soccer_Mid
- Wimbledon
- Xph_speed_bag

Rugby

- 10 seconds rugby footage
- A mix of fast moving scenes and details (players, grass)

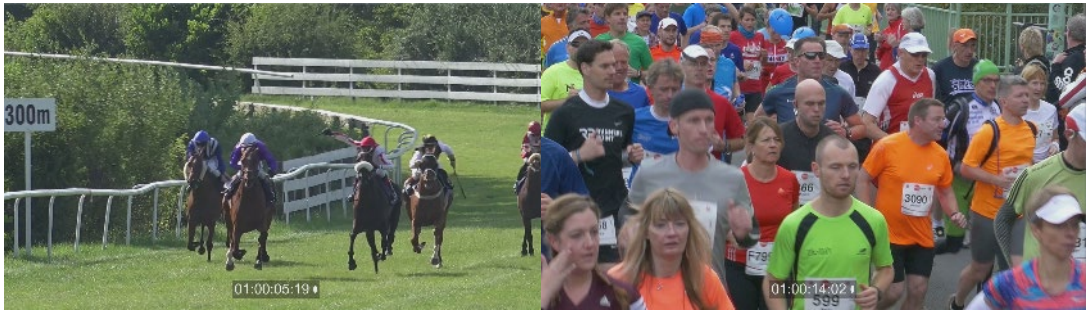
Soccer-Diego Description

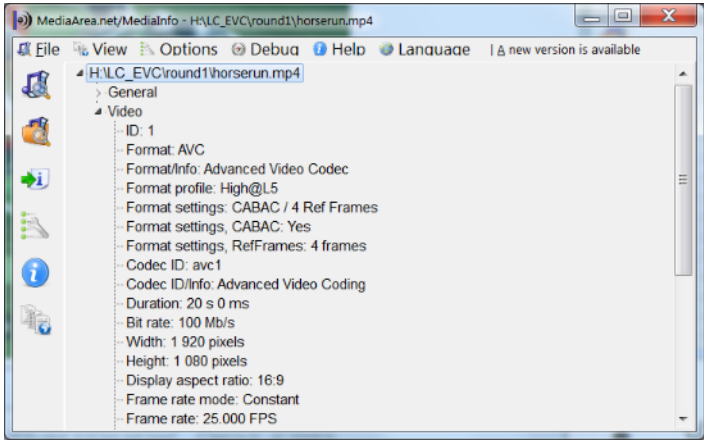
- Soccer clip - high action, small bodies
- Input doesn’t seem as high quality as other soccer clip (not so much detail)

Formula 1 Description

- Fast moving car turning corner
- Fairly high quality input

Horse Run Description

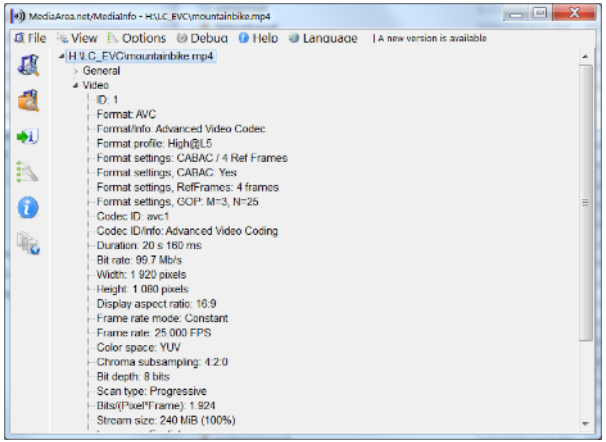




- 10 seconds horserace/ten seconds race close up
- Both from high quality source stock art clips

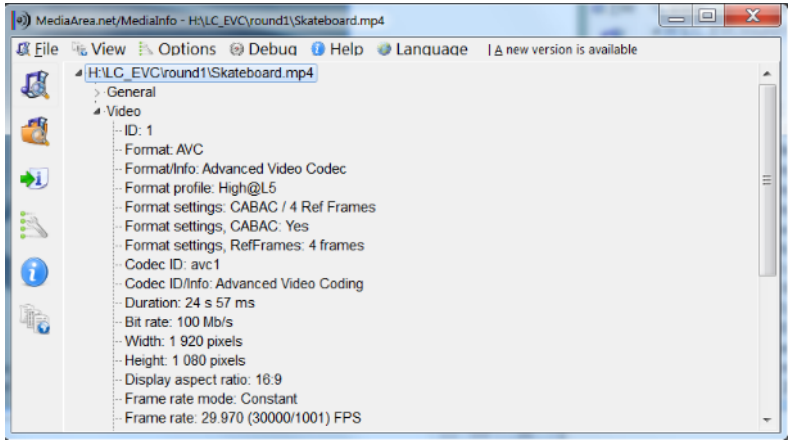
- Challenging footage
  - Horserace - high motion and high detail
  - Run - very high motion, much detail
  - Test clips from purchased stock video footage clips so very high

Mountain Bike Slow Mo Description



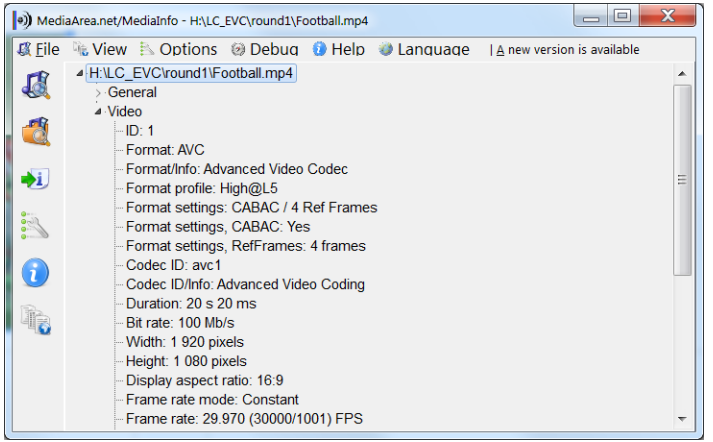
- 20 seconds of slow motion
- Pretty simple footage because it’s slow motion

Skateboard Description



- 30 fps version of Harmonic test clip
- Low detail
- High motion

Football Description



- 30 fps version of Harmonic test clip
- Lots of detail (artificial turf, uniforms)
- Lots of motion
- Lots of issues first time around

- Very relevant for US sporting channels

Formula 1 Description

- Fast moving car turning corner
- Fairly high quality input

New_Soccer_Close

- 9 seconds soccer clip
- A mix of soccer close-up scenes and crowd

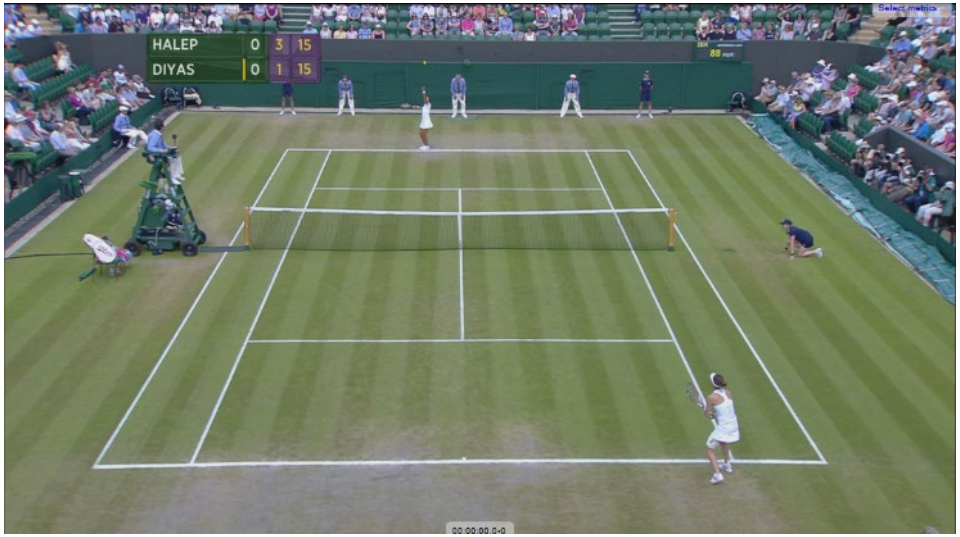
New_Soccer_Mid

- 8 seconds soccer clip
- Soccer action with lot of movement

El Fuente Boxing Description

- Boxing practice with high motion but low detail
- Taken from Netflix El Fuente test set

Wimbledon Description



- Wimbledon tennis match
- Relatively small figures and lots of hard straight lines

Xph_speed_bag

- 18 seconds
- From Xiph test set
- A mix of motion and fine details

Appendix III: GB Tech Report

- Methodology overview
- Profiles of test coordinators
- DSIS methodology
- Lab set-up
- Scope of analysis
- MOS results – summary

MPEG-5 LCEVC subjective testing – ABR ladder

Overview of DSIS MOS results

Vittorio Baroncini, Giacomo Baroncini
Rome, March 2021

Methodology overview

- Production of **DSIS MOS (Mean Opinion Score) according to the ITU-R Recommendation BT 500** (https://www.itu.int/rec/R-REC-BT.500-14-201910-I/en)
- Experiment designed and overseen by the MPEG Test Chair Vittorio Baroncini and performed by the independent laboratory GBtech under the supervision of the Test Administrator Giacomo Baroncini
- DSIS – or **Double Stimulus Impairment Scale** – protocol is known to be an efficient (in term of cost and human resources) and but effective and reliable methodology. The double-stimulus method is cyclic in that the assessor is first presented with an unimpaired reference, then with the same image impaired (in this case, either compressed with x264 or LCEVC). Following this, he is asked to vote on the second, keeping in mind the first. In sessions, which last up to 15 minutes, the assessor is presented with a series of sequences in random order and with random impairments covering all required combinations. At the end of the series of sessions, the mean score for each test condition, grouped by test images, is calculated, together with the 95% Confidence Interval value (that provides the maximum range of oscillation, around the mean score, that can be expected).

Profiles of test coordinators

Vittorio Baroncini (VABTECH ltd)

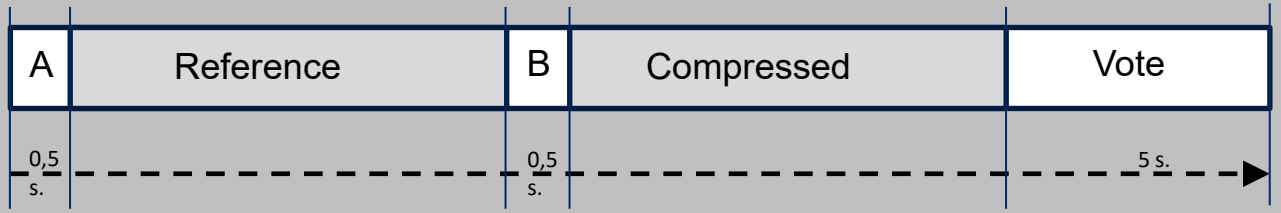
Working in the area of video since late 80s, member of ISO and ITU since mid 90s, expert in subjective and objective video quality assessment; co-founder of the VQEG (1998) and Chair of ITU-R WP6Q (2000-2009) releasing the first standard on objective TV quality metric. Chair (from 2002) of the SC29 WG11 (MPEG) test sub-group, designed and conducted all the Call for Evidence, Competition and Verification tests made in MPEG for Digital Cinema, AVC, HEVC and VVC standards. Designed several new objective and subjective test methods. Run over hundreds subjective test experiments serving many industries all over the world. At the moment member of BSI (British Standard Insitute)

Giacomo Baroncini (GBTech)

Active in the area of subjective video quality assessment since 2006; Contributing to the visual testing activities of SC29 WG11 (MPEG) test sub-group since 2013, acting as Test Administrator in all the Call for Evidence, Competition and Verification tests made in MPEG for HEVC and VVC standards. Run more than a hundreds subjective test experiments serving many industries all over the world.

DSIS methodology

BTC example



- DSIS test method is based on a Basic Test Cells (BTC) for each test point. Each BTC foresees the sequential presentation of:
  - Announcement of Reference video clip (letter A, 0.5 s.)
  - Reference uncompressed video clip (i.e., source)
  - Announcement of Compressed video clip (letter B, 0.5 s.)
  - Compressed video clip (either x264 or LCEVC in this experiment)
  - Voting time (announced by a message on screen – 5 s.)

- A DSIS test sessions includes:
  - One BTC for each test point (coding condition)
  - A stabilization phase; three BTC showing high, mid and low quality
  - One or two check BTC in which Reference is compared to Reference
- A test sessions stands for around 15 minutes
- When more time is required more test sessions are designed
- Each test session has “stabilization phase” and check BTCs

DSIS protocol – Impairment scale

- The DSIS test is based on the 11 grades impairment scale reported on the right
- The scale is a unilateral Likert scale, measuring the difference between the “reference” (source) and the “coded” video clips
- The 11 grades impairment scale adopted is taken from Rec. BT-2095 (EVP)

DSIS impairment scale

Score	Impairment item	
10	Imperceptible	
9	Slightly perceptible	somewhere
8		everywhere
7	Perceptible	somewhere
6		everywhere
5	Clearly perceptible	somewhere
4		everywhere
3	Annoying	somewhere
2		everywhere
1	Severely annoying	somewhere
0		everywhere

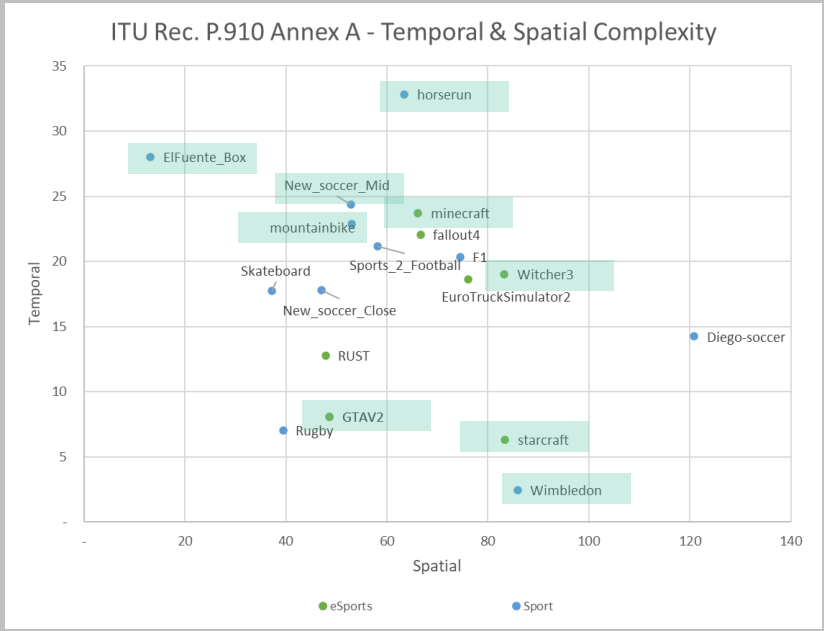
Lab set-up

- Custom SW player designed for formal subjective assessment (MUP)
- 32” Professional computer monitor (ASUS ProArt PA329Q)
- Calibration made by Xrite i1 Display Pro
- Viewing distance 3H
- Viewing angle < 30° (two viewers)
- Low ambient light (< 30 nits) behind the monitor
- Non reflecting dark gray floor, walls and ceiling

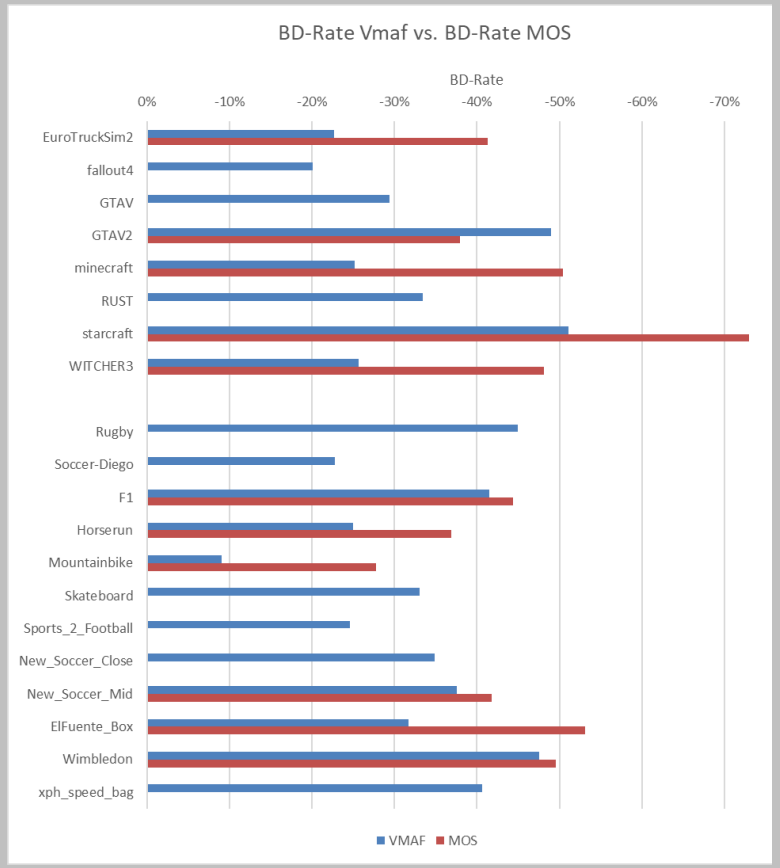
Scope of analysis

Sequences subjectively tested

- A representative set of 11 sequences subjectively tested (on a total of 20 sequences in the experiment) across 2 content types: 6 for sports, 5 for eGames
- Test sequences widely spread across the ITU P.910 temporal and spatial complexity (see chart on the right)
- Test across the full ABR ladders for Sport and eGames, from FullHD 1080p for the top rung, down to the bottom profiles at low resolutions.
- Two tested codecs, CBR real-time encoding: x264 (medium preset), MPEG-5 LCEVC x264 (base layer: x264 medium)
- A total of 128 observations: 78 for Sports (7 rungs for x264, 6 rungs LCEVC, across 6 clips), 50 for Sports (5 rungs for x264, 5 rungs LCEVC, across 5 clips)
- MOS scores for each sequence used to calculate BD-Rate-MOS, computed using an excel implementation of Bjontegaard function and compared to BD-Rate-Vmaf provided by V-Nova



BD-rate MOS results - summary



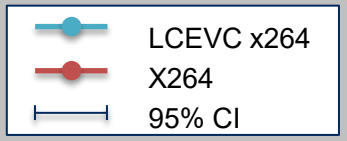
eGames ABR ladder

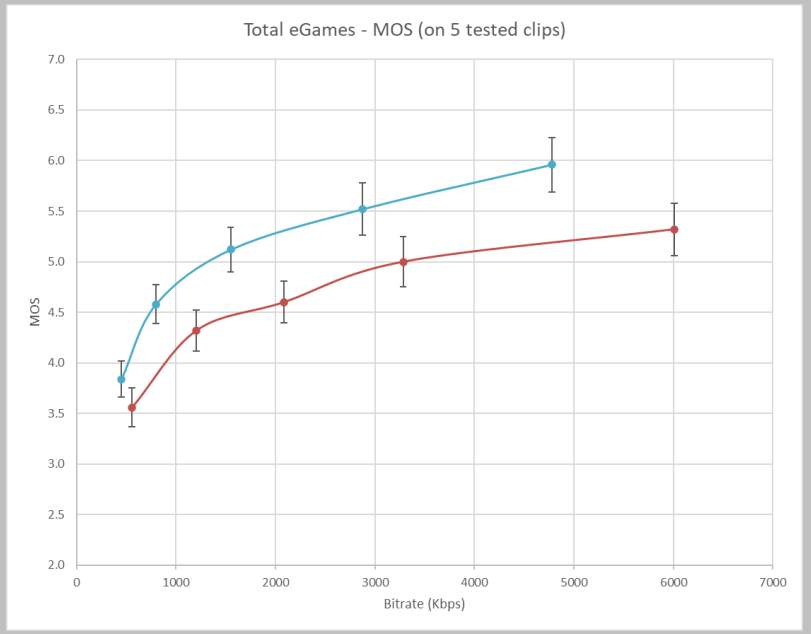
#	Content	BD-Rate-Vmaf	BD-Rate-MOS
1	EuroTruckSim2	-22.63%	-41.26%
2	fallout4	-20.09%	-
3	GTAV	-29.45%	-
4	GTAV2	-48.94%	-37.92%
5	minecraft	-25.15%	-50.40%
6	RUST	-33.38%	-
7	starcraft	-51.07%	-73.00%
8	WITCHER3	-25.66%	-48.08%
Total eGames		**-33.08%**	-
Total Subjectively tested		**-35.08%**	**-54.24%**

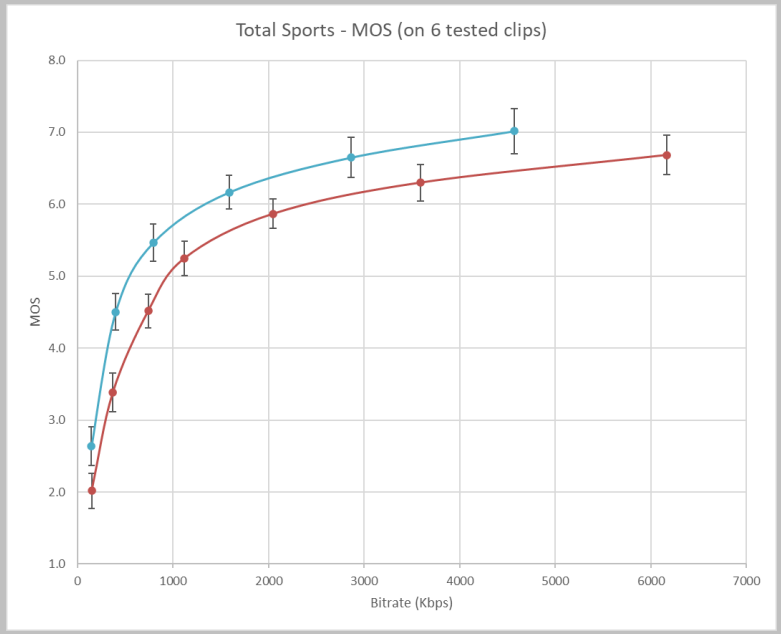
Sports ABR ladder

#	Content	BD-Rate-Vmaf	BD-Rate-MOS
1	Rugby	-44.94%	-
2	Soccer-Diego	-22.80%	-
3	F1	-41.53%	-44.39%
4	Horserun	-24.96%	-36.87%
5	Mountainbike	-9.08%	-27.81%
6	Skateboard	-33.07%	-
7	Sports_2_Football	-24.62%	-
8	New_Soccer_Close	-34.84%	-
9	New_Soccer_Mid	-37.54%	-41.76%
10	ElFuente_Box	-31.72%	-53.07%
11	Wimbledon	-47.53%	-49.54%
12	xph_speed_bag	-40.64%	-
Total Sports		**-38.17%**	-
Total Subjectively tested		**-32.34%**	**-44.04%**

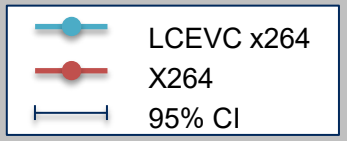
MOS summary results

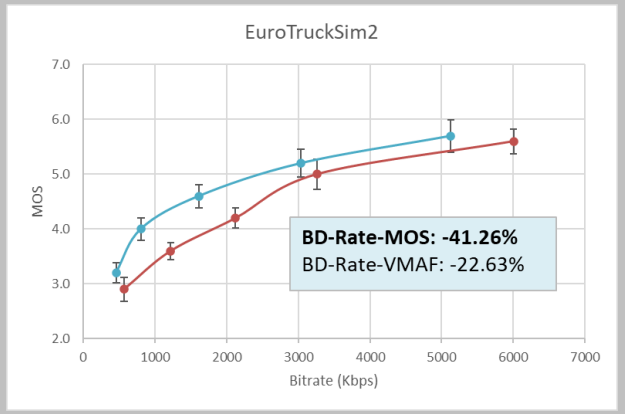


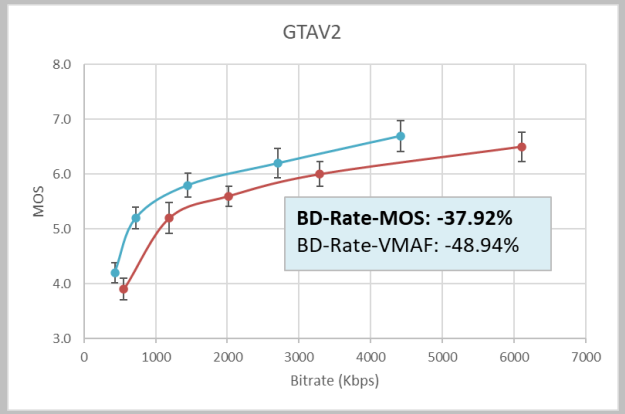


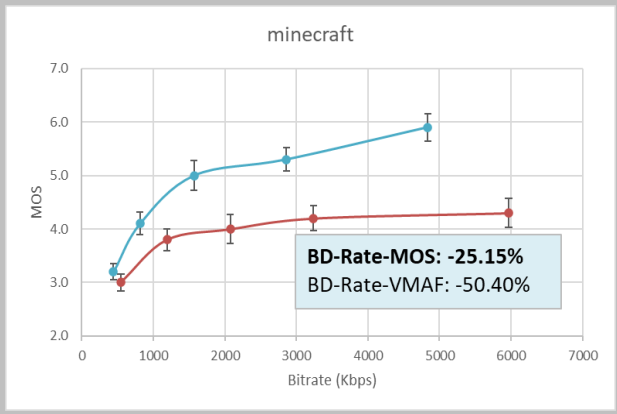


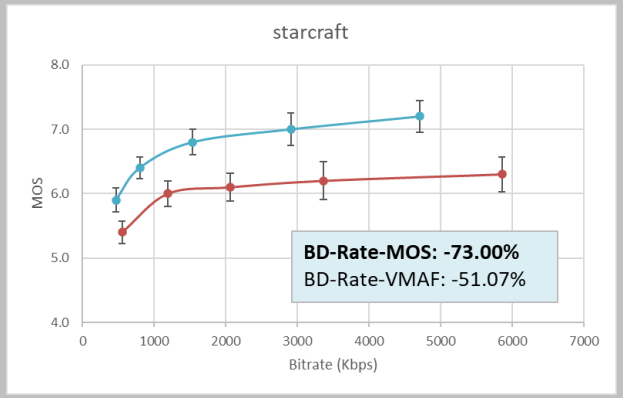
MOS results - eGames

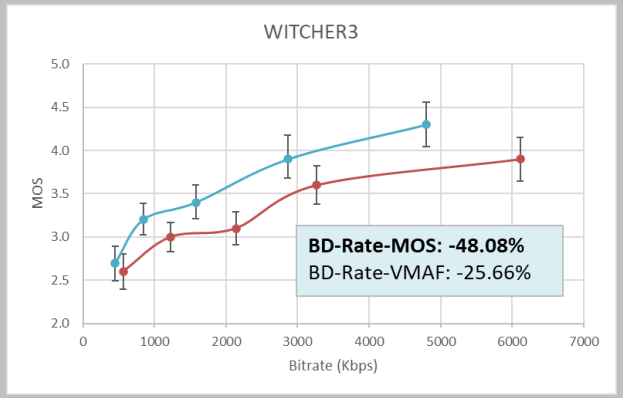












MOS results – Sports