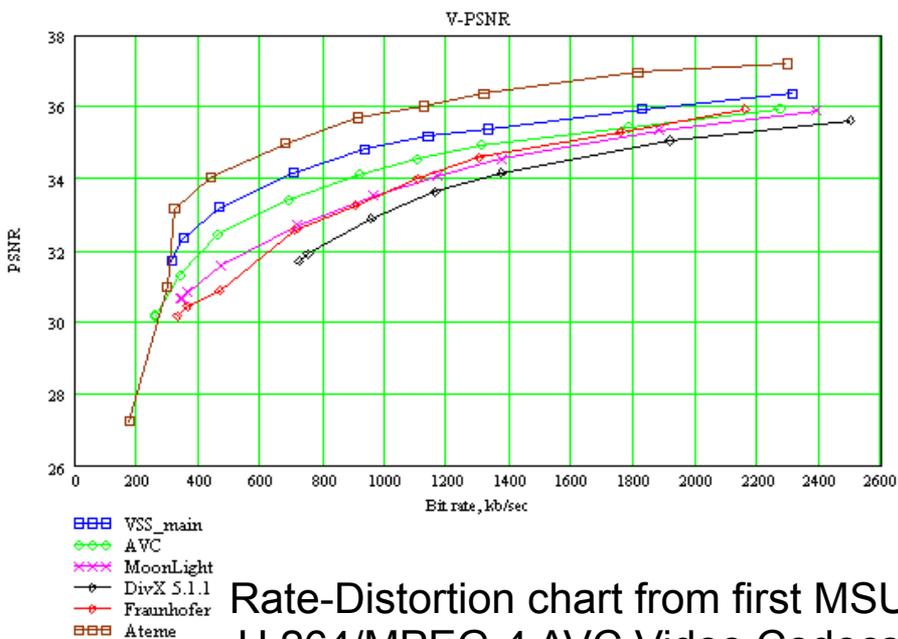


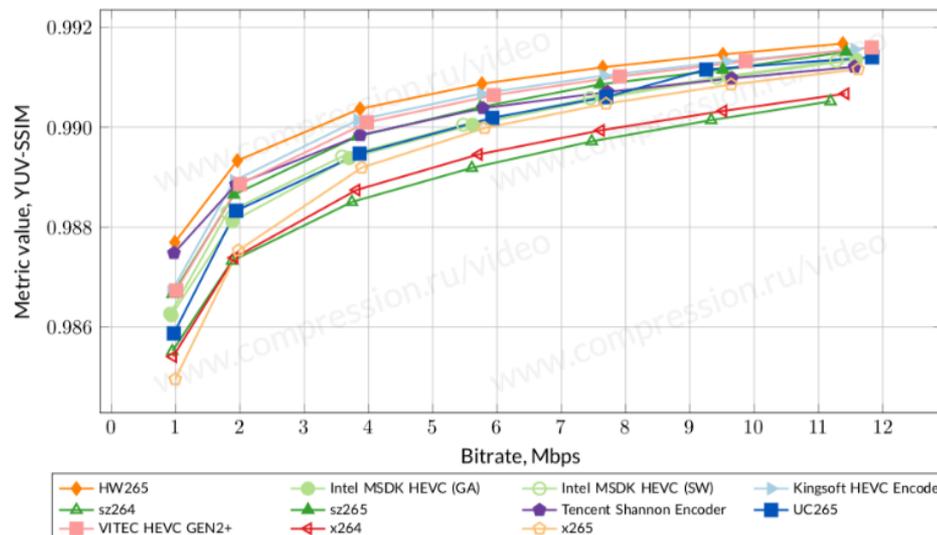
MSU Annual Video Codecs Comparisons Methodology

About MSU Codecs Comparisons

- **15 years** in codecs comparisons (from 2003)
- Released **22 reports** (30000+ charts, 2500+ pages)
- **210+** codecs were analyzed



Rate-Distortion chart from first MSU H.264/MPEG-4 AVC Video Codecs Comparison (2005)



Rate-Distortion chart from 13th HEVC Codecs Comparison (2018)

History of MSU Codecs Comparisons

#	Year	Title	#codecs	#pages	#charts
1	2003	MSU Video Codecs Comparison	33	190	319
2	2005	MPEG-4 SP/ASP Codecs Comparison	7	68	65
3	2004	Lossless Video Codecs Comparison 2004	15	23	9
4	2005	JPEG 2000 Image Codecs Comparison	9	27	216
5	2005	1-st Annual MPEG-4 AVC/H.264 Codecs Comparison	7	67	63
6	2005	MSU Subjective Comparison of Modern Video Codecs	4	34	22
7	2006	MPEG-2 Video Decoders Comparison	8	43	41
8	2006	HD Photo and JPEG 2000 Comparison	10	27	12
9	2005	2-nd Annual MSU MPEG-4 AVC/H.264 Video Codec Comparison	8	88	140
10	2007	3-rd Annual MSU MPEG-4 AVC/H.264 Video Codec Comparison	7	157	185
11	2007	Lossless Video Codecs Comparison 2007	16	130	122
12	2007	4-th Annual MSU MPEG-4 AVC/H.264 Video Codec Comparison	6	162	213
13	2008	Options Analysis of Codec x264	1	41	38
14	2009	5-th MSU MPEG-4 AVC/H.264 Video Codec Comparison	6	204	279
15	2010	6-th MPEG-4 AVC/H.264 Video Codecs Comparison - Short Version	8	407	488
16	2011	7-th MPEG-4 AVC/H.264 Video Codecs Comparison - Standard Version	10	199	1522
17	2012	8-th MPEG-4 AVC/H.264 Video Codecs Comparison - Standard Version	9		2092
18	2013	9-th MPEG-4 AVC/H.264 Video Codecs Comparison	4		1500+
19	2015	HEVC Codec Comparison - 2015	12		5500+
20	2016	HEVC Codec Comparison - 2016	8		4000+
21	2017	HEVC Codec Comparison - 2017	11	552	5000+
22	2018	HEVC Codec Comparison - 2018	14	150+	7118+

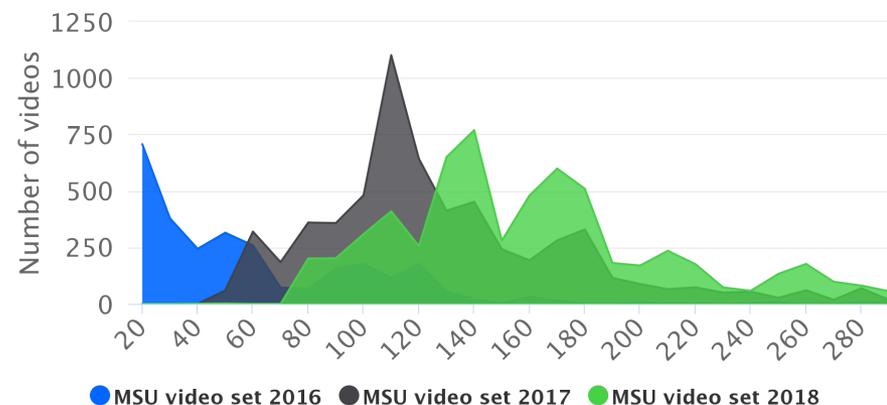
Outline

- Video selection
- Encoding presets selection
- Encoders launches
- Quality measure
- Encoders score calculation
 - Special cases
- Calculation of overall score (BSQ-rate)
- Call-for-codecs 2019
 - Use cases speed limits
 - Conditions for encoders
 - Hardware configuration

Video selection (1)

- Test video set is chosen from MSU video collection (15000+ videos)
 - All videos have high bitrate, they were generated by real users and downloaded from different resources (Vimeo, xiph.org, etc)
 - All videos have an appropriate license (CCBY)
- All videos are in YUV420P (name in ffmpeg), SDR, 8-bit

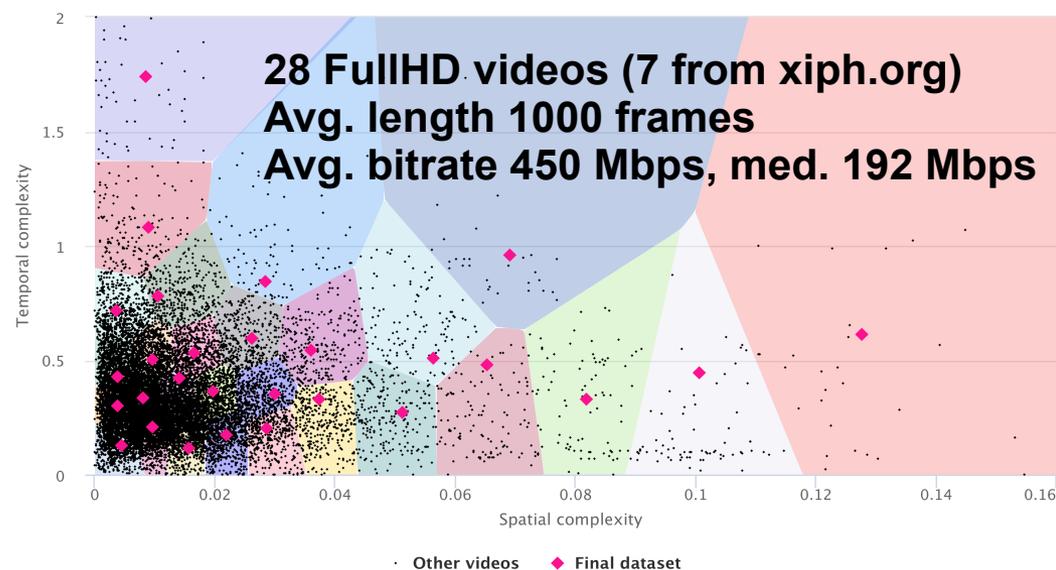
Year	# FullHD videos	# FullHD samples	# 4K videos	# 4K samples	Total # of videos	Total # of samples
2016	3	7	882	2902	885	2909
2017	1996	4738	1544	4561	3540	9299
2018	4342	10330	1946	5503	6288	15833



Bitrate distribution of video sequences from MSU video set by year

Video selection (2)

- Video Collections are divided into clusters by spatiotemporal complexity [1]
- A number of clusters is equal to a number of required test set
- Videos that are close to the clusters centres are chosen as candidates. Final video set also includes videos from previous comparisons and other famous video collections



Video set used in 2018 MSU video codecs comparison (Part I: FullHD)

Encoding presets selection

- All encoding presets are requested and provided by codecs developers
 - If presets were not provided, MSU chooses presets
- All provided presets are checked for compliance with speed limitations
 - Speed limits depend on the encoding use case, they may vary year to year, so they are presented in **call-for-codecs**
- Conditions for encoding presets are also presented in **call-for-codecs**
 - For example, permission for multi-pass encoding, type of rate-control etc.

Encoders launches

- Each configuration (codec, video, use case, bitrate) is launched 3 times for excluding external effects (emissions due to overheating, caching, etc.), and the minimum operating time is taken as a result
- Measurement of encoding time
 - An external time (full working time of encoder binary) measurement is used (including time of reading, writing of encoded video)
 - RAM is used to store FullHD, SSD — for 4K videos
- Hardware configuration may vary from year to year (it is presented in **call-for-codecs**)

Quality measure

- Main metric of the comparison: YUV-SSIM
 - YUV-SSIM is computed as the weighted average of SSIM values for each channel individually
 - $(4*Y\text{-SSIM}+V\text{-SSIM}+U\text{-SSIM})/6$
- Other used metrics: Y-VMAF, Y-SSIM, U-SSIM, V-SSIM, YUV-PSNR, Y-PSNR, U-PSNR, V-PSNR (additional metrics may vary and listed in **call-for-codecs**)
- Metrics calculation: via MSU VQMT http://compression.ru/video/quality_measure/video_measurement_tool.html (version is specified in **call-for-codecs**)

Encoders score calculation

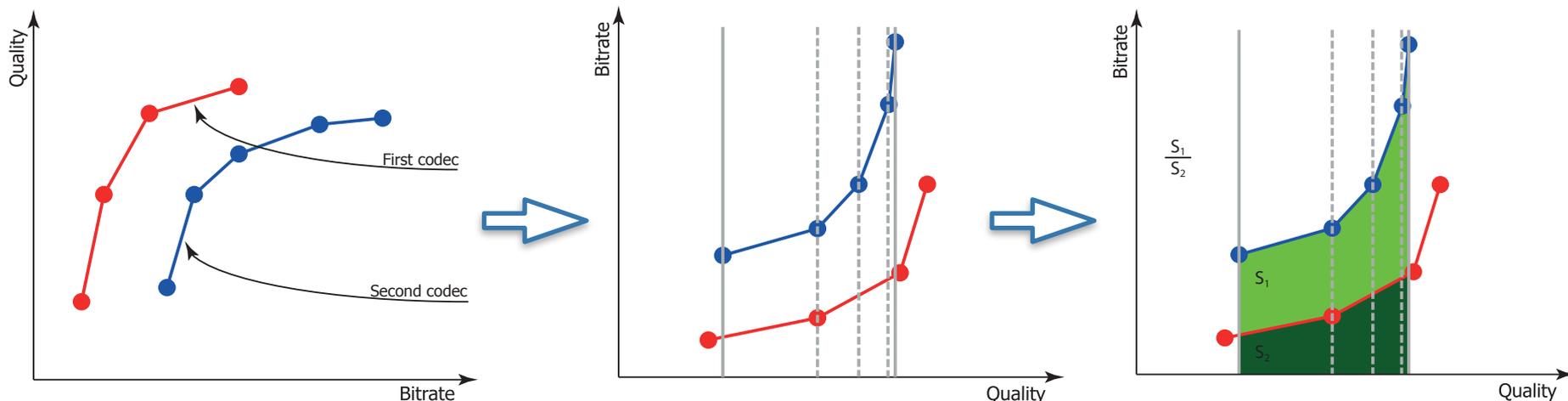
- Averaging of per-frame metrics results for the whole video — simple mean (average) value
- Linear interpolation of RD-curves
- Reference encoder — x264 (newest version, depends on the year)
- Results are averaged across all testing videos: simple mean (average) value, for speed and for quality

Calculation of overall score

Bitrate for the same quality rate (BSQ-rate)

1. Compute the average bitrate ratio for a fixed quality is to invert the axes of the bitrate/quality graph
2. Average the interval over which the quality axis is chosen. The averaging is only over overlapping segments
3. Calculation of the area under the curves in the chosen interpolation segment and determination of their ratio

$$BSQ - rate = \frac{S_1}{S_2}$$

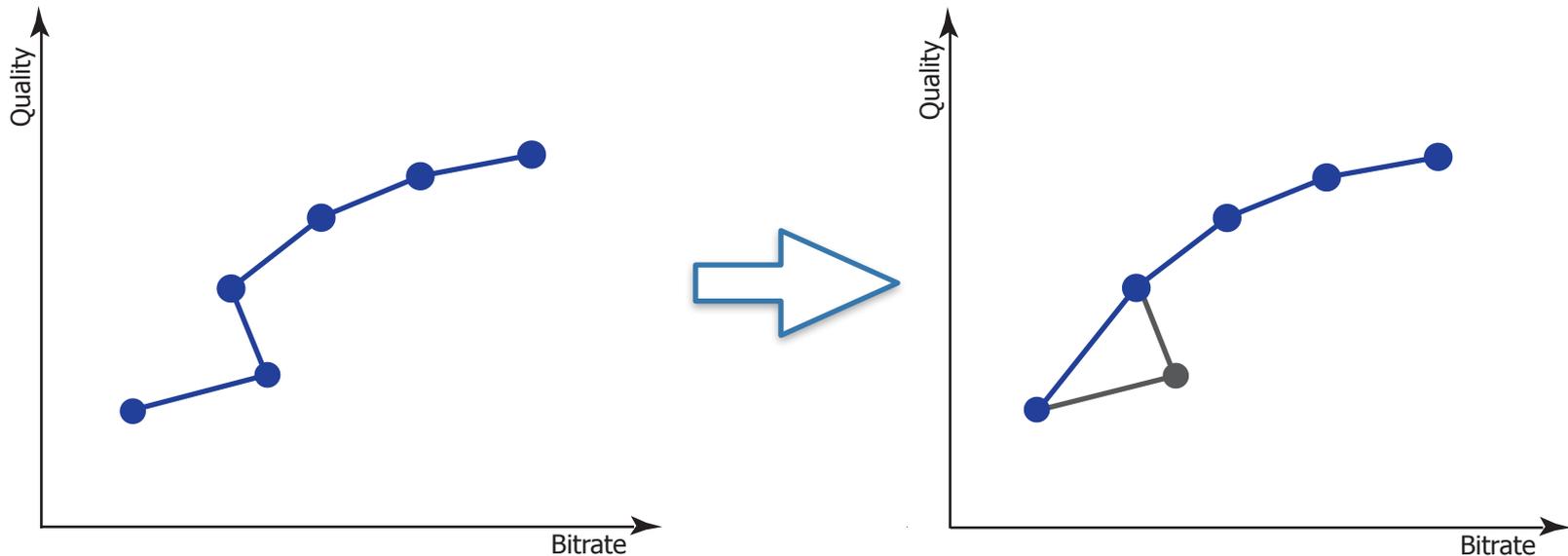


Encoders score calculation

Special cases



Non-monotonic RD curve: outliers are excluded before the interpolation



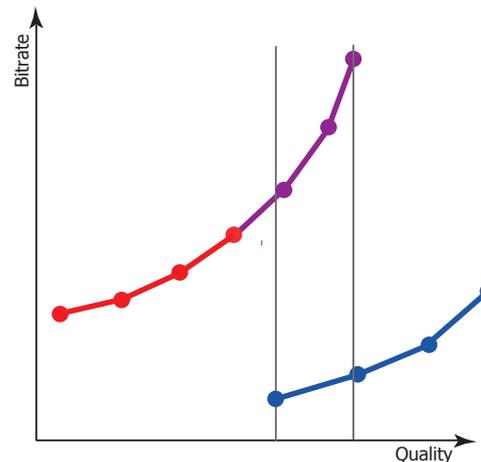
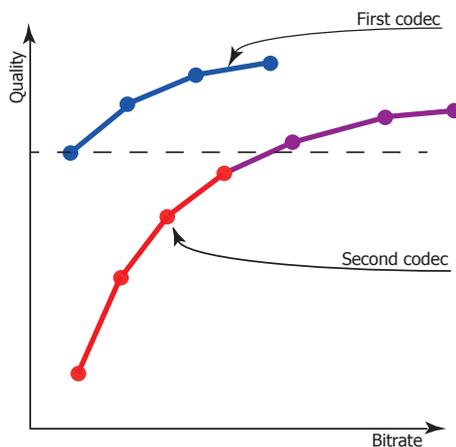
Encoders score calculation

Special cases



Non-overlapping RD curves: additional measurements are made at lower/higher bitrates

- Reference encoder is measured on lower bitrates
- For comparison of two non-reference encoders: lower bitrates for leaders are measured until the intersection



Purple part of the curve - additionally measured bitrates to reach the intersection

Call-for-codecs 2019

Use cases speed limits



	Resolution	FPS
Fast/High Density	FullHD (1080p)	60
Universal/Broadcast VQ	FullHD (1080p)	25
Ripping/Pristine VQ	FullHD (1080p)	1 SSIM-RD curve better than x264-veryslow
Ultra-Ripping	FullHD (1080p)	5
Subjective comparison	FullHD (1080p)	1
4K comparison	4K (UHD)	20

Call-for-codecs 2019

Conditions for encoders



- Participant can use multi-threading
- MSU doesn't limit GOP size and intra-period
- MSU doesn't limit the number of passes in the encoding. Total encoding time should fulfil the requirements
- VBR mode is used during all tests (other modes are discussed individually)
- Encoder should be compatible with reference decoder
- Preferable codec interface — console codec version (with batch processing support — bitrate and file names must be possible to assign from the command line)

Call-for-codecs 2019

Hardware configuration



- **CPU:** Intel Socket 1151 Core i7 8700K (Coffee Lake) (3.7Ghz, 6C12T, TDP 95W)
- **Mainboard:** ASRock Z370M Pro4
- **RAM:** Crucial CT16G4DFD824A 2x16GB (totally 32 GB) DIMM DDR4 2400MHz CL15
- **OS:** Windows 10 x64

Contact

For any questions about comparison methodology
and participation:

videocodec-testing@graphics.cs.msu.ru