MSU Codec Comparison 2019

Part III: 4K Content, Objective Evaluation

Graphics & Media Lab
Video Group

Video group head	Dr. Dmitriy Vatolin
Project head	Dr. Dmitriy Kulikov
Measurements & analysis	Dr. Mikhail Erofeev,
	Egor Sklyarov,
	Anastasia Antsiferova

Free version

Codecs:

H.265

- Bytedance V265 Encoder
- HW265
- MainConcept HEVC
- SVT-HEVC
- sz265
- x265

Non H.265

- SIF Encoder
- SVT-AV1
- SVT-VP9
- VP9
- WZAurora AV1 Encoder
- x264

CS MSU Graphics & Media Lab, Video Group
March 7, 2020
http://www.compression.ru/video/codec_comparison/index_en.html
videocodec-testing@graphics.cs.msu.ru

1. REPORT VERSIONS

	Free version	Enterprise version
Per-sequence-results	2 of 11 sequences	All 11 sequences (in interactive charts)
Metric: YUV-SSIM	✓	✓
Description of video sequences	✓	✓
Codec info (developer, version number, website link)	✓	✓
Relative quality analysis	✗	✓
Other objective metrics (Y-VMAF, Y-SSIM, U-SSIM, V-SSIM, YUV-PSNR, Y-PSNR, U-PSNR, V-PSNR)	✗	✓
Per-frame metrics results (in HTML report)	✗	All metrics for all sequences (1700+ charts)
Download links for video sequences	✗	✓
Encoders presets description	✗	✓

Contents

2. ACKNOWLEDGMENTS

The Graphics & Media Lab Video Group would like to thank the following companies for providing the codecs and settings used in this report:

- ByteDance Inc.
- Huawei Technologies Co., Ltd.
- MainConcept GmbH
- MulticoreWare, Inc.
- Nanjing Yunyan
- SIF Encoder Team
- Visionular

We are also grateful to these companies for their help and technical support during the tests.

3. OVERVIEW

3.1. Sequences

	Sequence	Number of frames	Frame rate	Resolution
1.	Crowd run	500	50	3840×2160
2.	Ducks take off	500	50	3840×2160
3.	Flat Tour	926	60	3840×2160
4.	Grass	576	24	3840×2160
5.	Hiking	1303	24	3840×2160
6.	Hotel tour	1077	30	3840×2160
7.	Mallorca	1104	30	3840×2160
8.	New Masons	1359	24	3840×2160
9.	OPCW	1015	25	3840×2160
10.	Photo shoot	1053	30	3840×2160
11.	Wedding	1015	24	3840×2160

Table 1: Summary of video sequences

Brief descriptions of the sequences used in our comparison appear in Table 1. Appendix A provides more-detailed descriptions of these sequences.

3.2. Codecs

Codec	Developer	Version
Bytedance V265 Encoder	ByteDance Inc.	v1.2.6
HW265	Huawei Technologies Co., Ltd.	V0.7.3
MainConcept HEVC	MainConcept GmbH	HEVC/H.265 Encoder SDK v11
SIF Encoder	SIF Encoder Team	v1.84.0
SVT-AV1	Open Visual Cloud	0.8.0
SVT-HEVC	Open Visual Cloud	1.4.1
SVT-VP9	Open Visual Cloud	0.1.0
sz265	Nanjing Yunyan	v1.0
VP9	The WebM Project	v1.8.0-424-ge50f4e411
WZAurora AV1 Encoder	Visionular	v1.0
x264	x264 Developer Team	core:157 r2969 d4099dd
x265	MulticoreWare, Inc.	3.2+15-04db2bfee5d6

Table 2: Short codecs' descriptions

Brief descriptions of the codecs used in our comparison appear in Table 2. We used x264 as a good-quality AVC reference codec. Appendix C provides detailed descriptions of all codecs in our comparison.

4. OBJECTIVES AND TESTING RULES

In this report we use objective assessment methods to compare the encoding quality of recent HEVC encoders as well as encoders implementing other standards. This effort employed 11 video sequences at 4K resolution to evaluate codec performance. To choose our test set, we analyzed 384,946 video sequences and selected representative examples (a detailed description of the selection process appears in Appendix B).

Our comparison consists of two use cases. For each use case we offered the codec developers the option to provide encoding parameters for our tests. Nevertheless, the parameters had to satisfy a minimum speed requirements for their respective use case:

- Fast—20fps
- Universal—1fps

Our comparison used a computer with the following configuration: based on an Intel Core i7-8700K (Coffee Lake) processor @ 3.7GHz with 32 GB of RAM running Windows 10. For objective quality measurements we used the YUV-SSIM metric (see Appendix E.1).

5. 4K UNIVERSAL (1FPS)

5.1. RD Curves

Judging from the mean quality scores (computed using the method described in Section D), first place in the quality competition goes to **WZAurora AV1 Encoder**, second place goes to **HW265**, and third place to **MainConcept HEVC**, **Bytedance V265 Encoder**, and **SVT-AV1**.

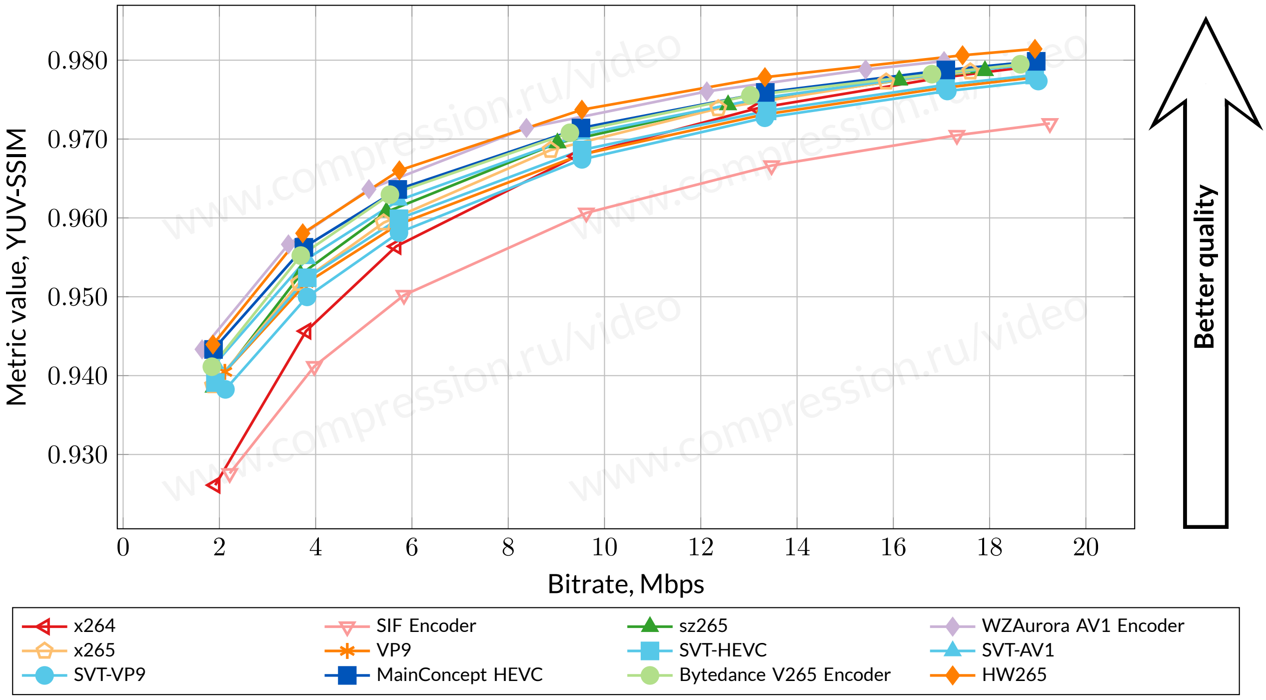


Figure 1: Bitrate/quality—use case "4K Universal (1fps)," *Grass* sequence, YUV-SSIM metric.

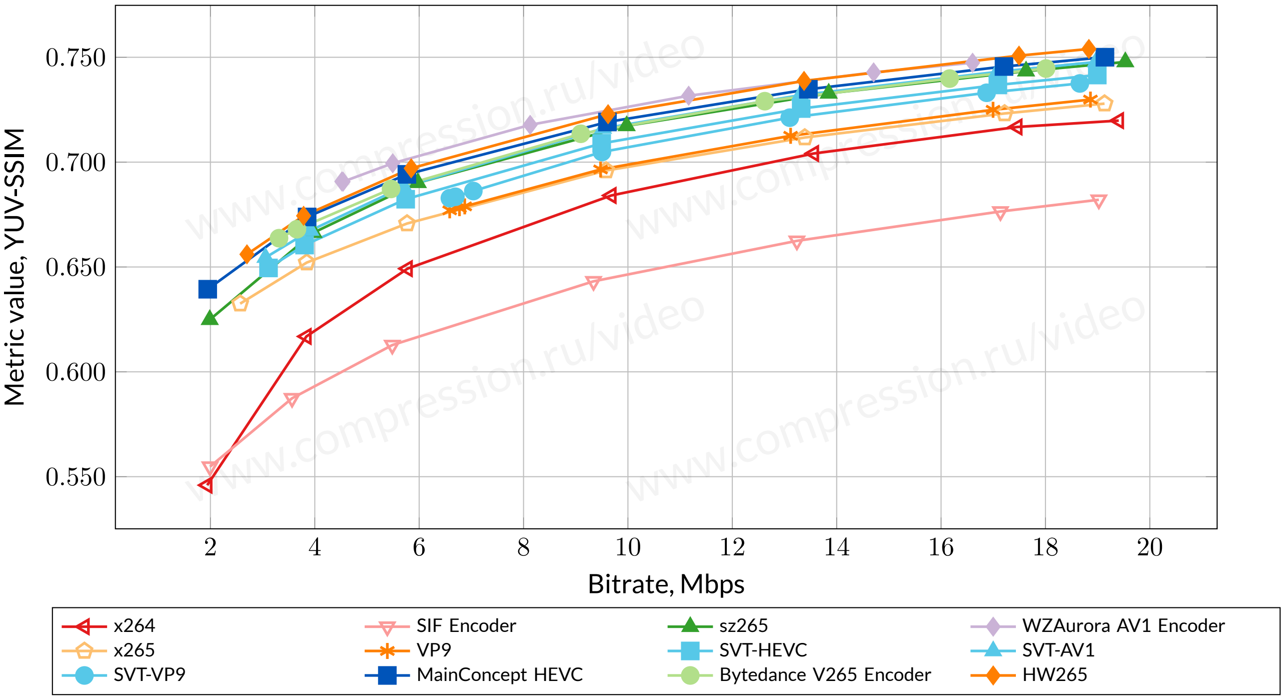
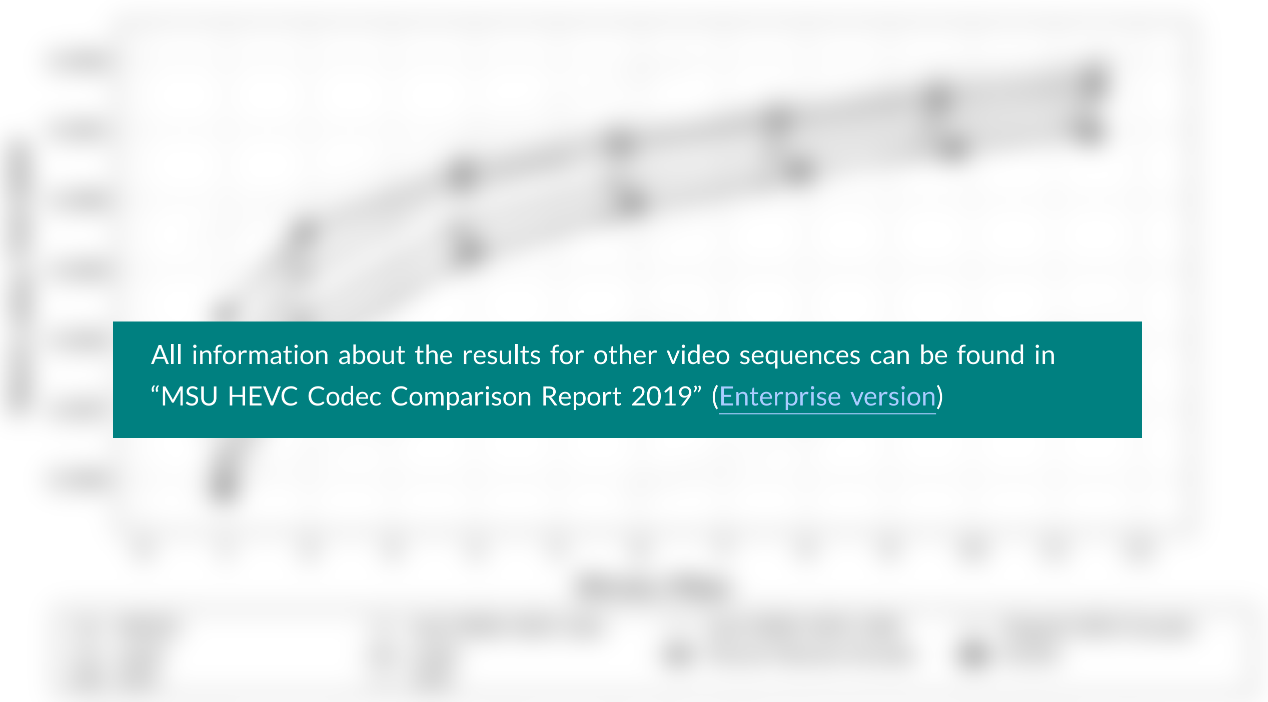


Figure 2: Bitrate/quality—use case “4K Universal (1fps),” *Ducks take off* sequence, YUV-SSIM metric.



All information about the results for other video sequences can be found in “MSU HEVC Codec Comparison Report 2019” (Enterprise version)

5.2. Encoding Speed

Judging from the mean speed scores (computed using the method described in Section D), first place in the speed competition goes to **SIF Encoder**, second place goes to **Bytedance V265 Encoder**, and third place to **SVT-AV1**.

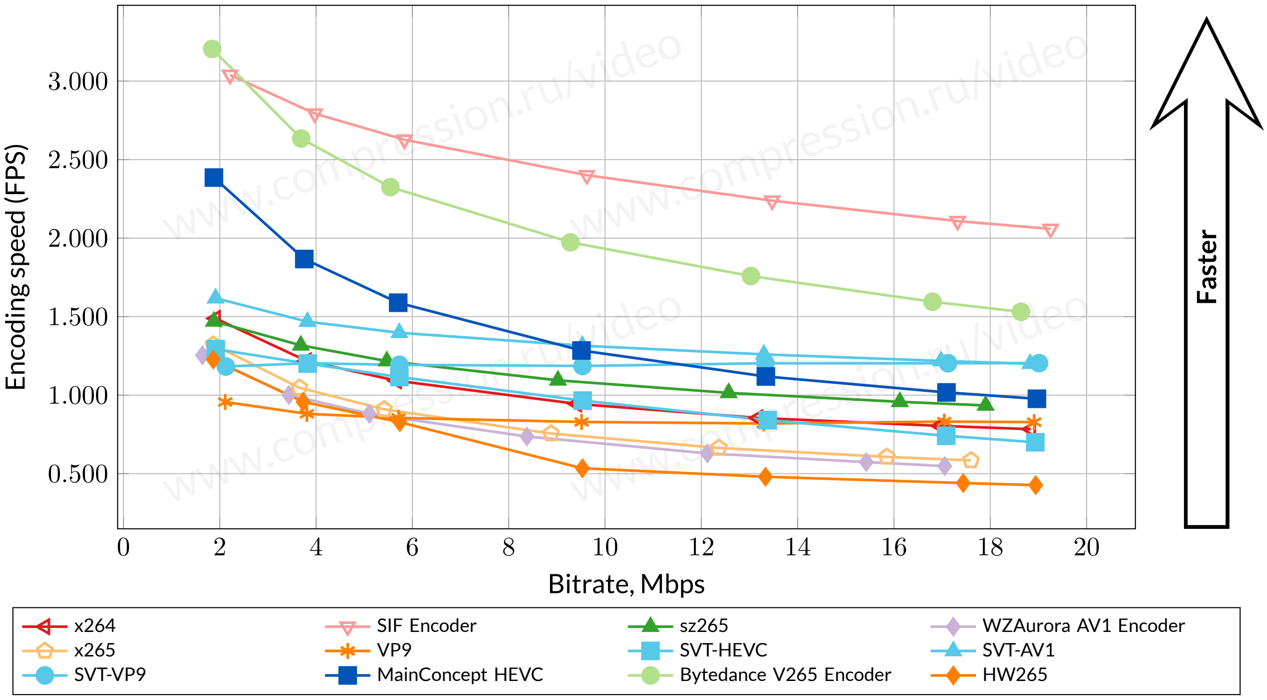


Figure 3: Encoding speed—use case "4K Universal (1fps)," *Grass* sequence.

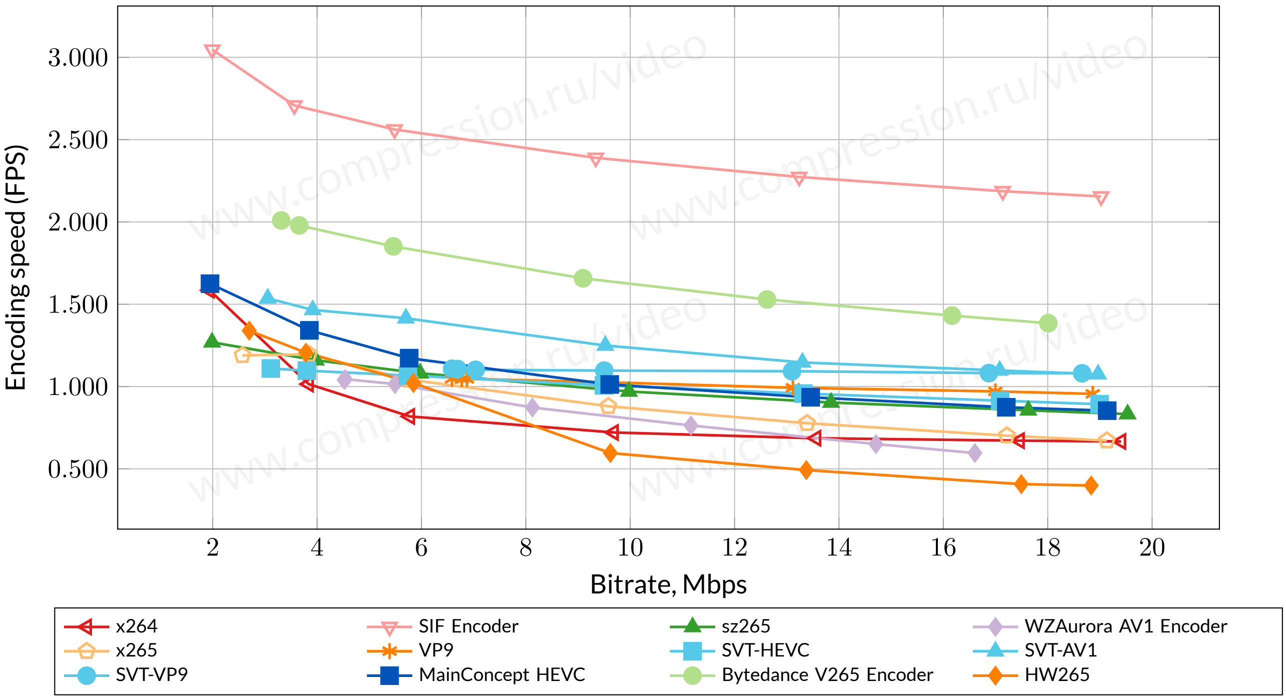
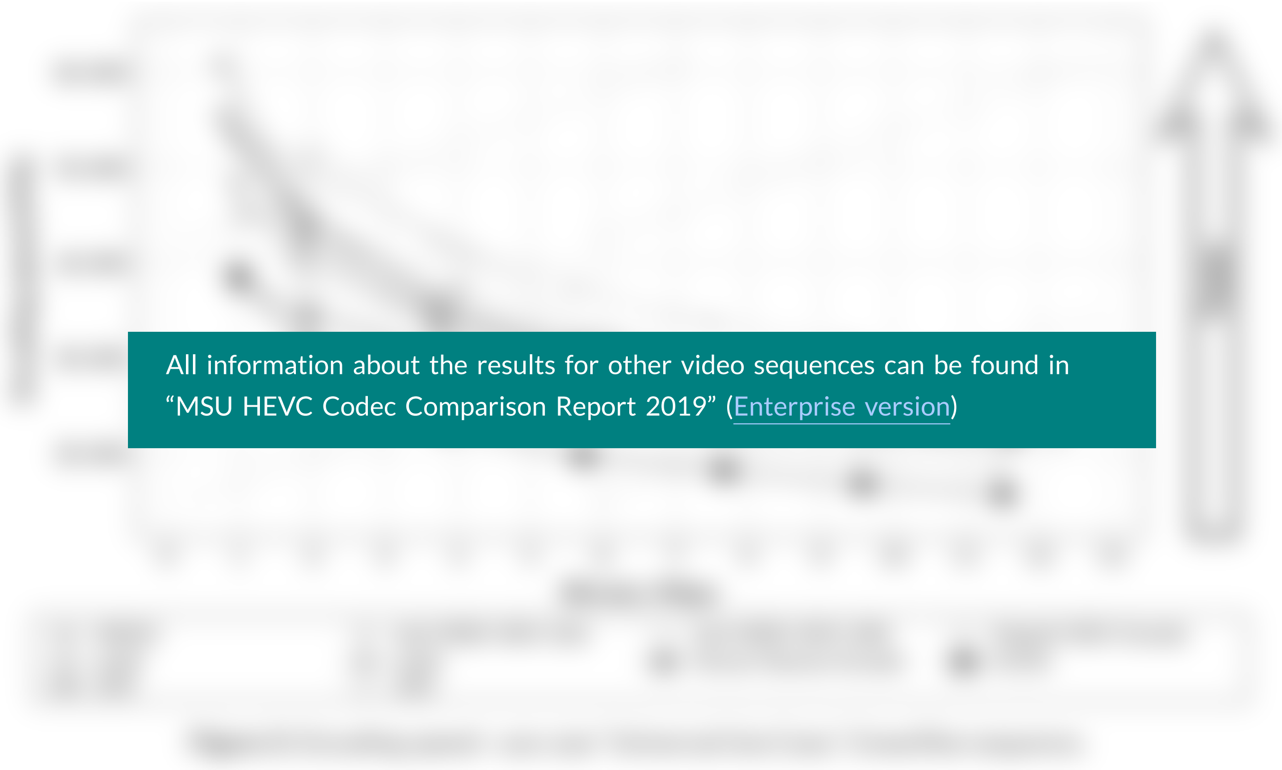


Figure 4: Encoding speed—use case "4K Universal (1fps)," *Ducks take off* sequence.

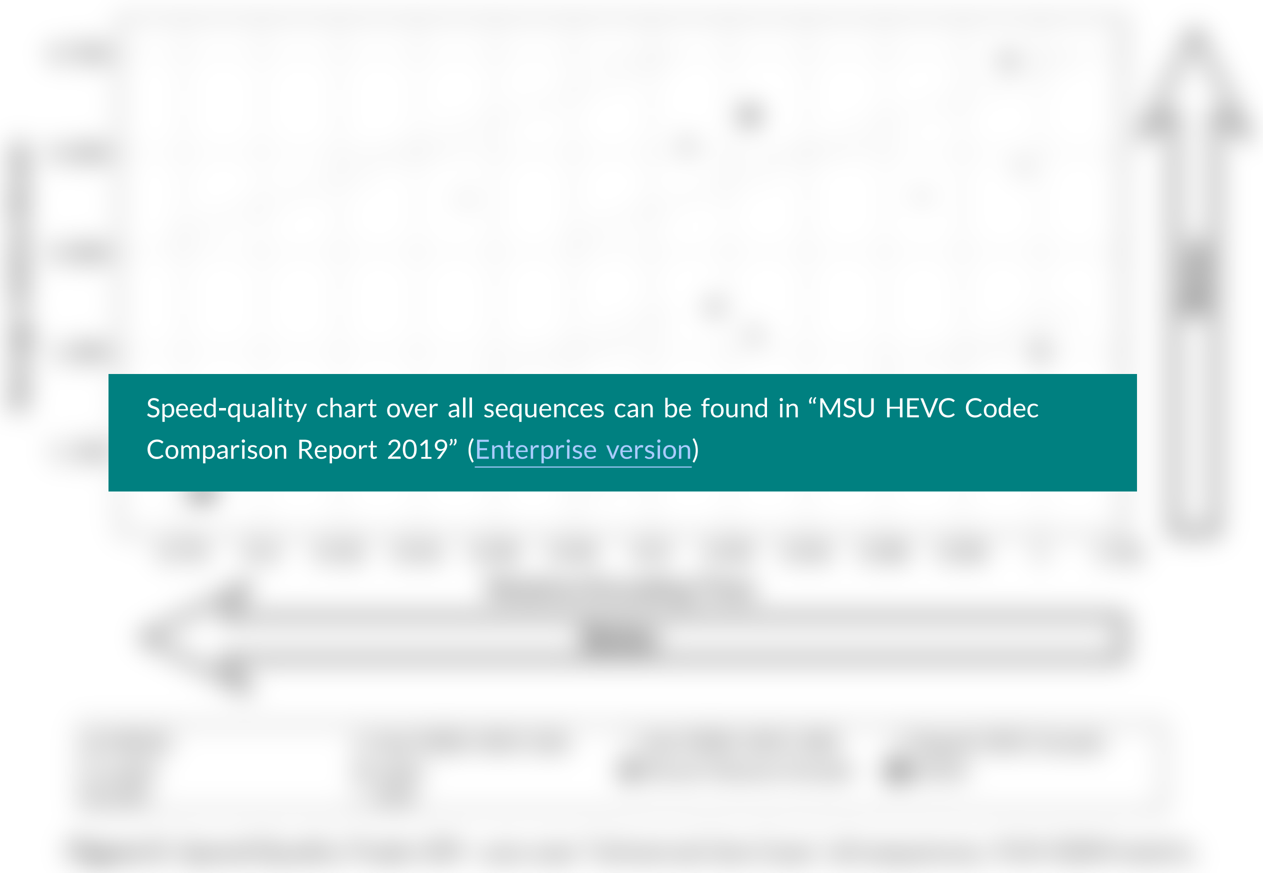


All information about the results for other video sequences can be found in "MSU HEVC Codec Comparison Report 2019" (Enterprise version)

5.3. Speed/Quality Trade-Off

Detailed descriptions of the speed/quality trade-off graphs are in Appendix D. Some graphs omit the results for a particular codec owing to that codec's extremely poor performance (i.e., its RD curve fails to intersect with the reference RD curve).
The speed/quality trade-off graphs show both relative quality and speed scores for the encoders under comparison. Since we chose x264 as the reference codec, we normalized all scores to the x264 scores.

There are five Pareto-optimal encoders: **WZAurora AV1 Encoder**, **HW265**, **MainConcept HEVC**, **Bytedance V265 Encoder**, and **SIF Encoder**.



Speed-quality chart over all sequences can be found in “MSU HEVC Codec Comparison Report 2019” (Enterprise version)

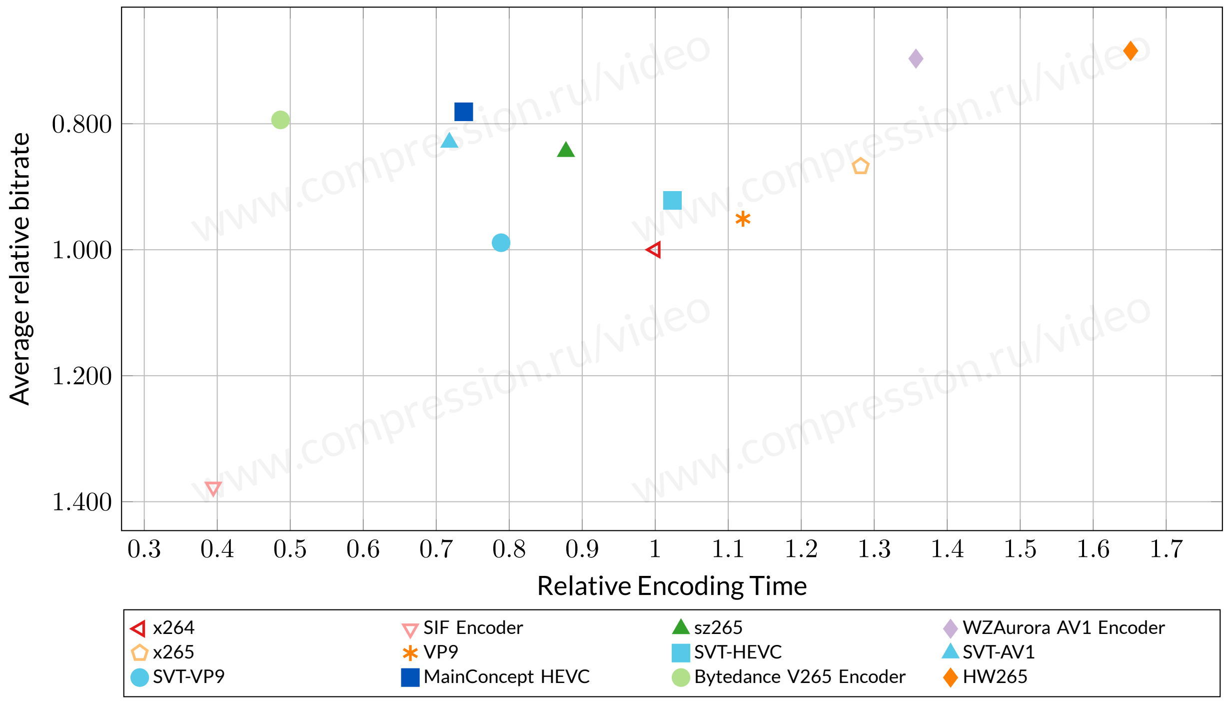


Figure 5: Speed/Quality Trade-Off—use case “4K Universal (1fps),” *Grass* sequence, YUV-SSIM metric.

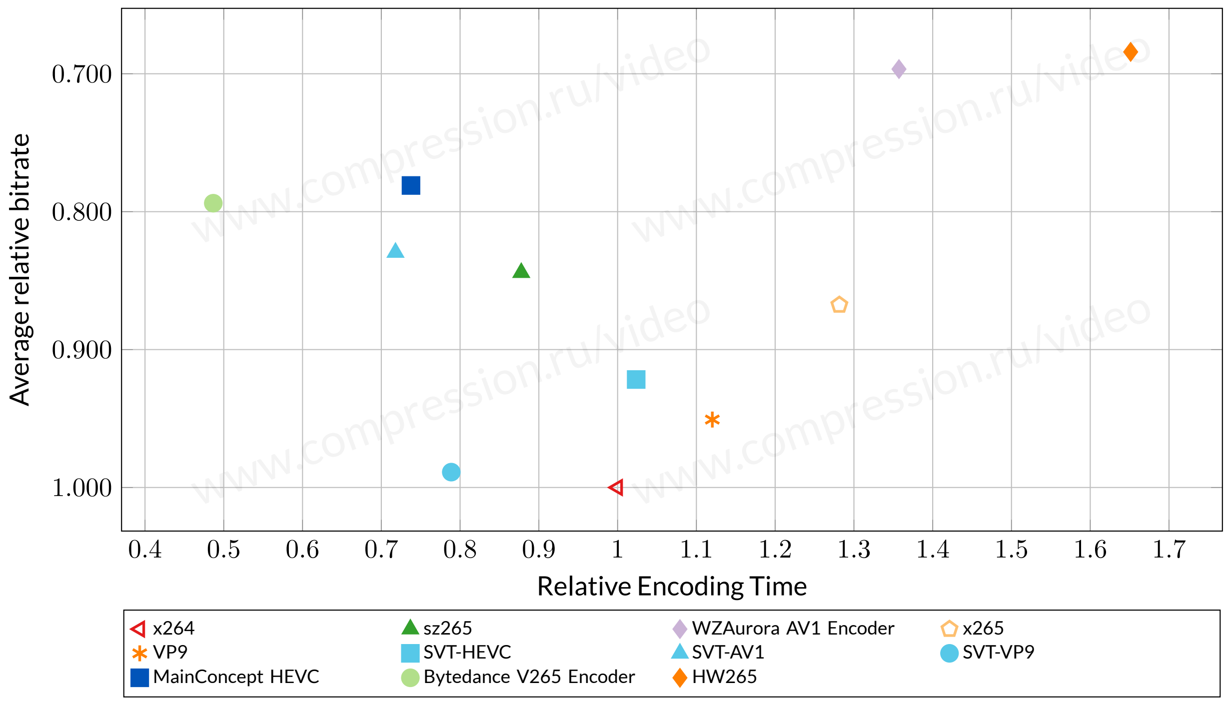


Figure 6: Speed/Quality Trade-Off—use case "4K Universal (1fps)," *Grass* sequence, YUV-SSIM metric, without SIF Encoder.

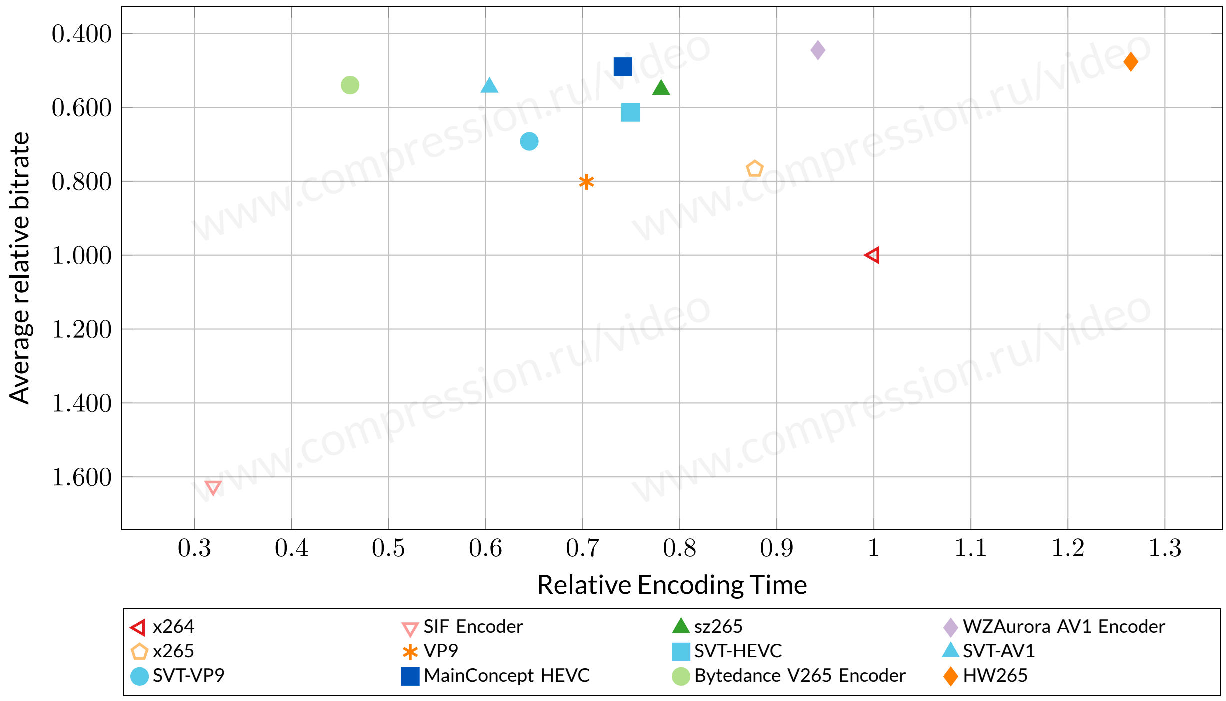


Figure 7: Speed/Quality Trade-Off—use case "4K Universal (1fps)," *Ducks take off* sequence, YUV-SSIM metric.

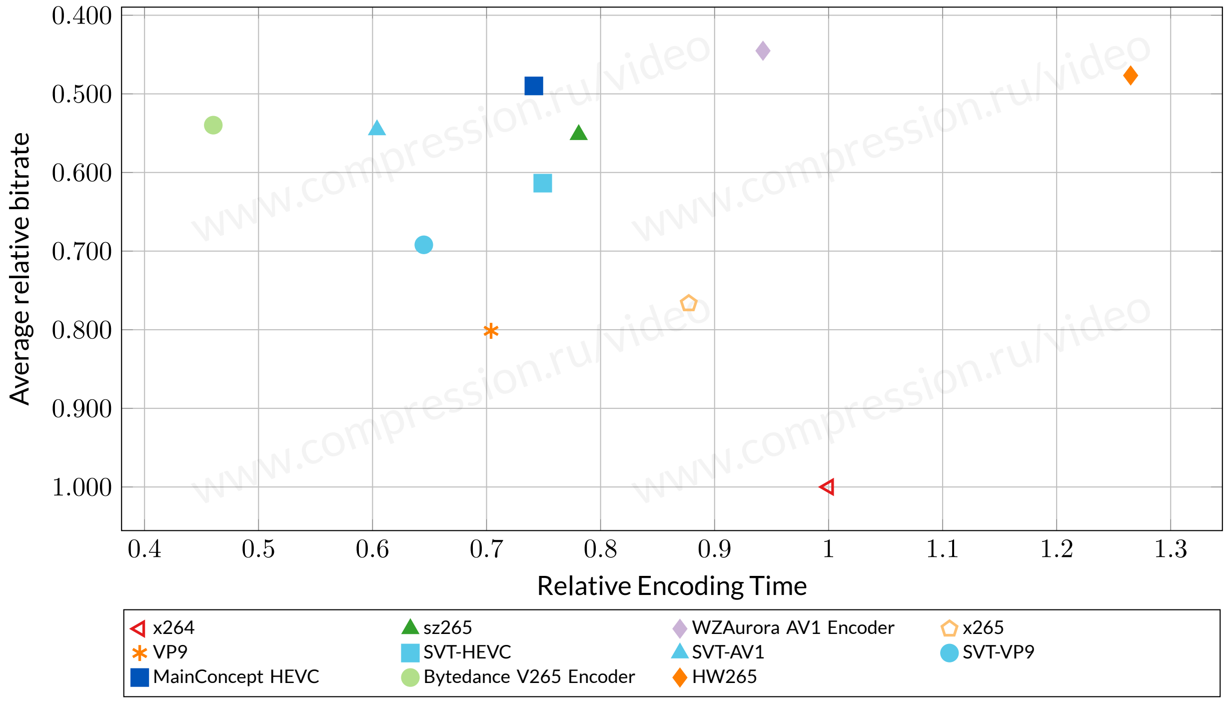
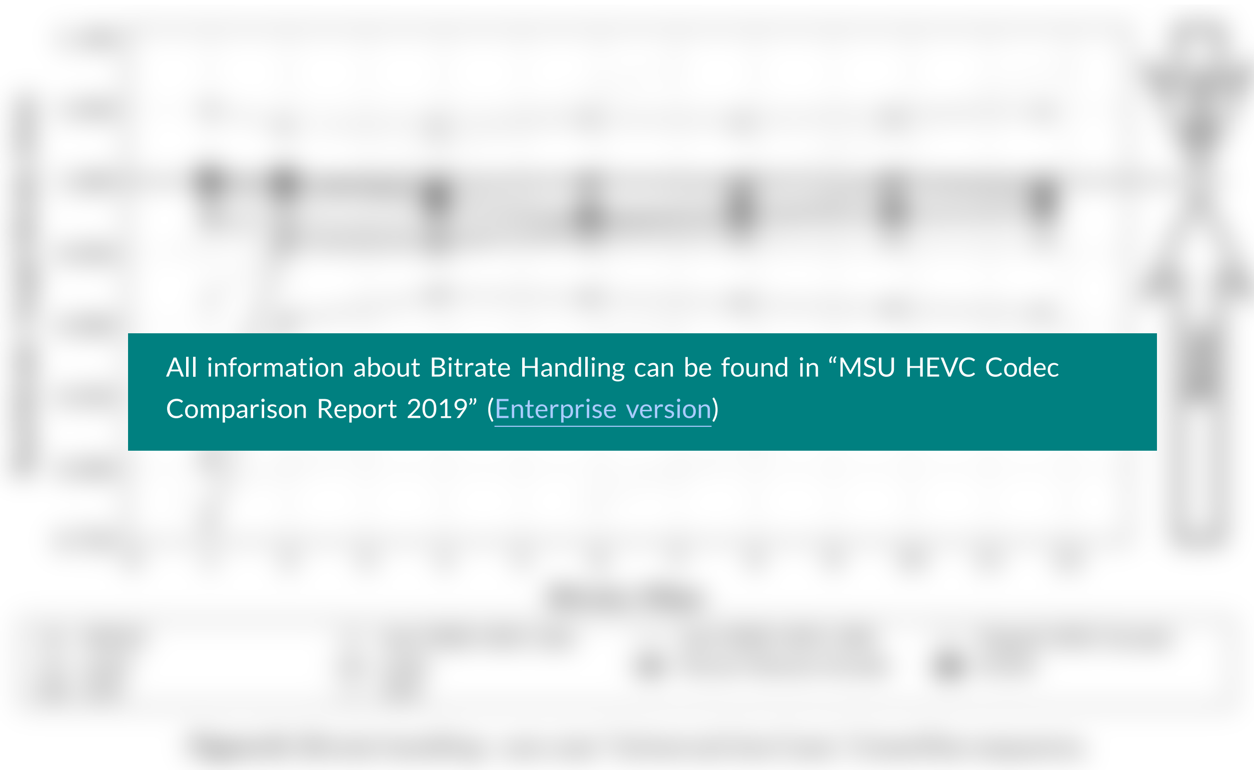


Figure 8: Speed/Quality Trade-Off—use case "4K Universal (1fps)," *Ducks take off* sequence, YUV-SSIM metric, without SIF Encoder.

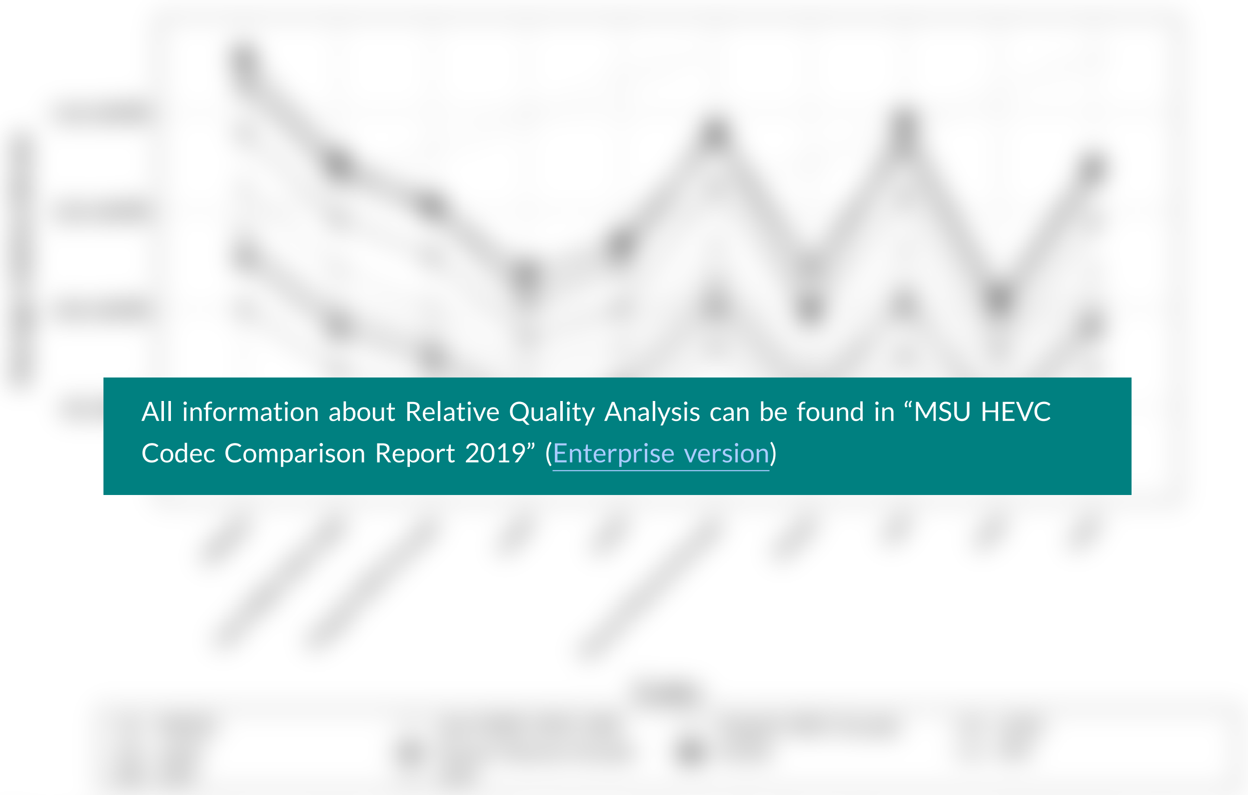
5.4. Bitrate Handling



All information about Bitrate Handling can be found in "MSU HEVC Codec Comparison Report 2019" (Enterprise version)

5.5. Relative Quality Analysis

Note that each number in the tables below corresponds to some range of bitrates (see Appendix D.5). Unfortunately, these ranges can differ significantly because of differences in the quality of compared encoders. This situation can lead to some inadequate results when three or more codecs are compared.



All information about Relative Quality Analysis can be found in "MSU HEVC Codec Comparison Report 2019" (Enterprise version)

6. 4K FAST (20FPS)

6.1. RD Curves

Judging from the mean quality scores (computed using the method described in Section D), first place in the quality competition goes to **HW265**, second place goes to **Bytedance V265 Encoder** and **MainConcept HEVC**, and third place to **sz265**.

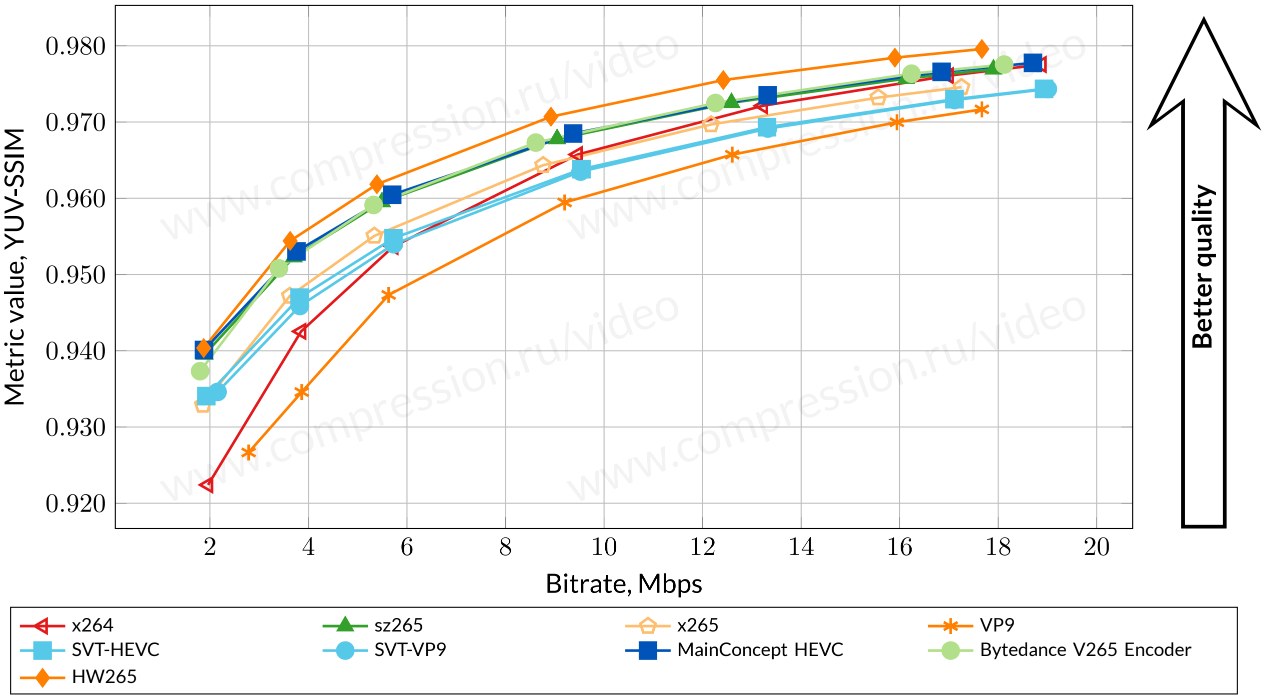


Figure 9: Bitrate/quality—use case "4K Fast (20fps)," *Grass* sequence, YUV-SSIM metric.

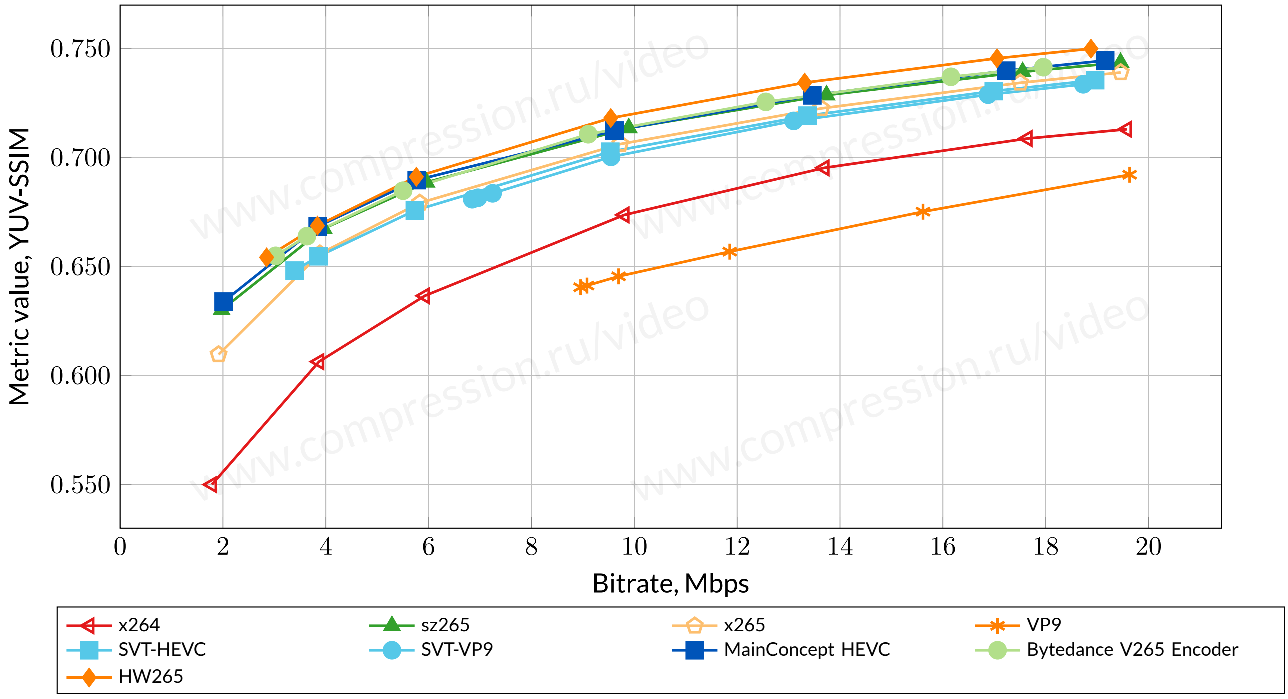
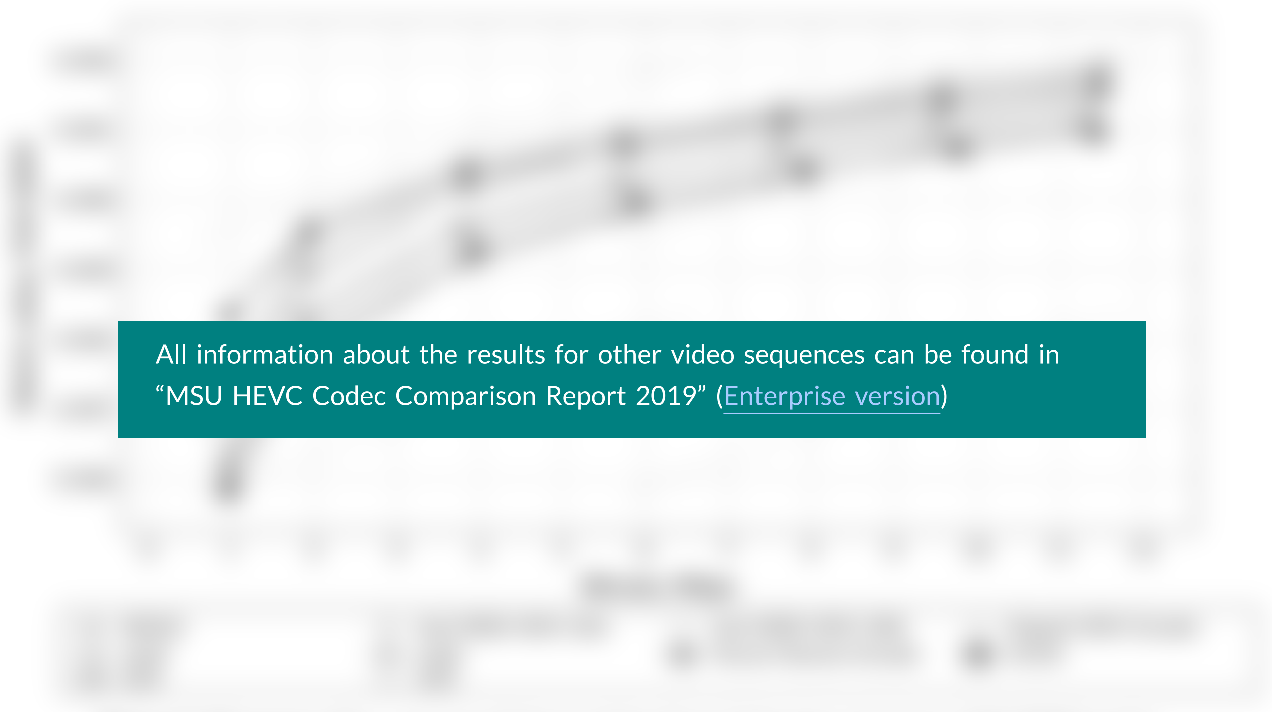


Figure 10: Bitrate/quality—use case "4K Fast (20fps)," *Ducks take off* sequence, YUV-SSIM metric.



All information about the results for other video sequences can be found in "MSU HEVC Codec Comparison Report 2019" (Enterprise version)

6.2. Encoding Speed

Judging from the mean speed scores (computed using the method described in Section D), first place in the speed competition goes to **VP9** and **SVT-HEVC**, second place goes to **SVT-VP9**, and third place to **x265**.

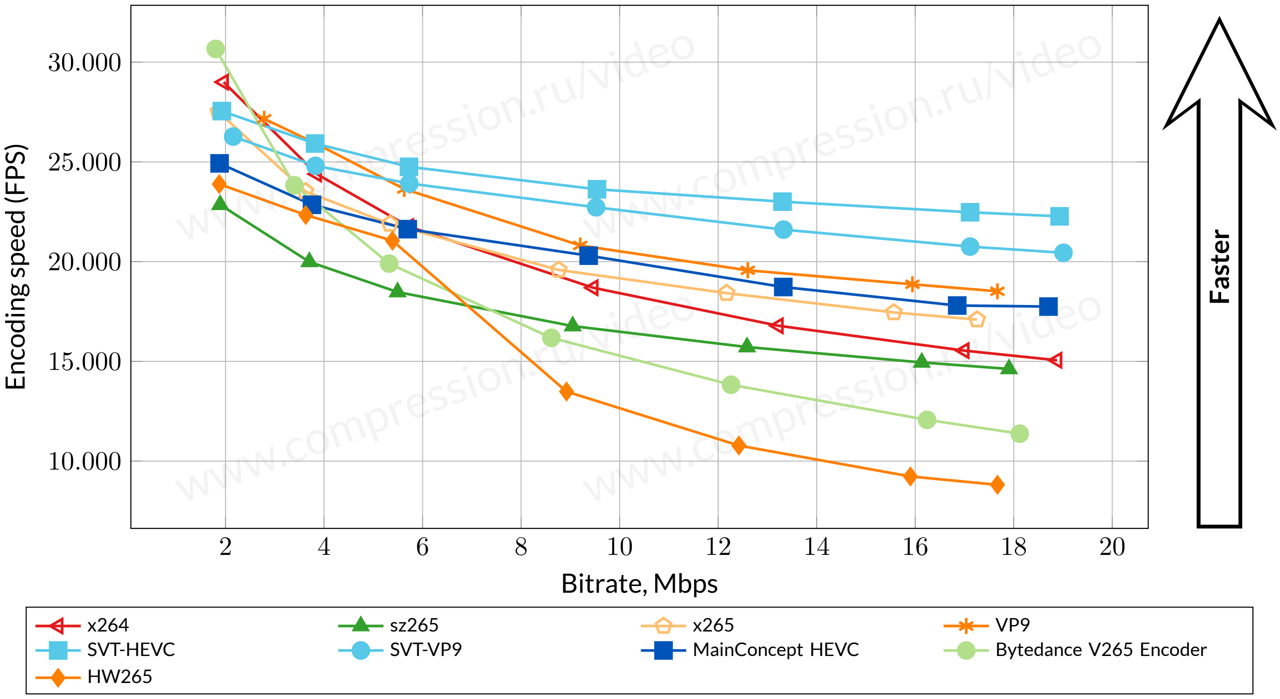


Figure 11: Encoding speed—use case "4K Fast (20fps)," *Grass* sequence.

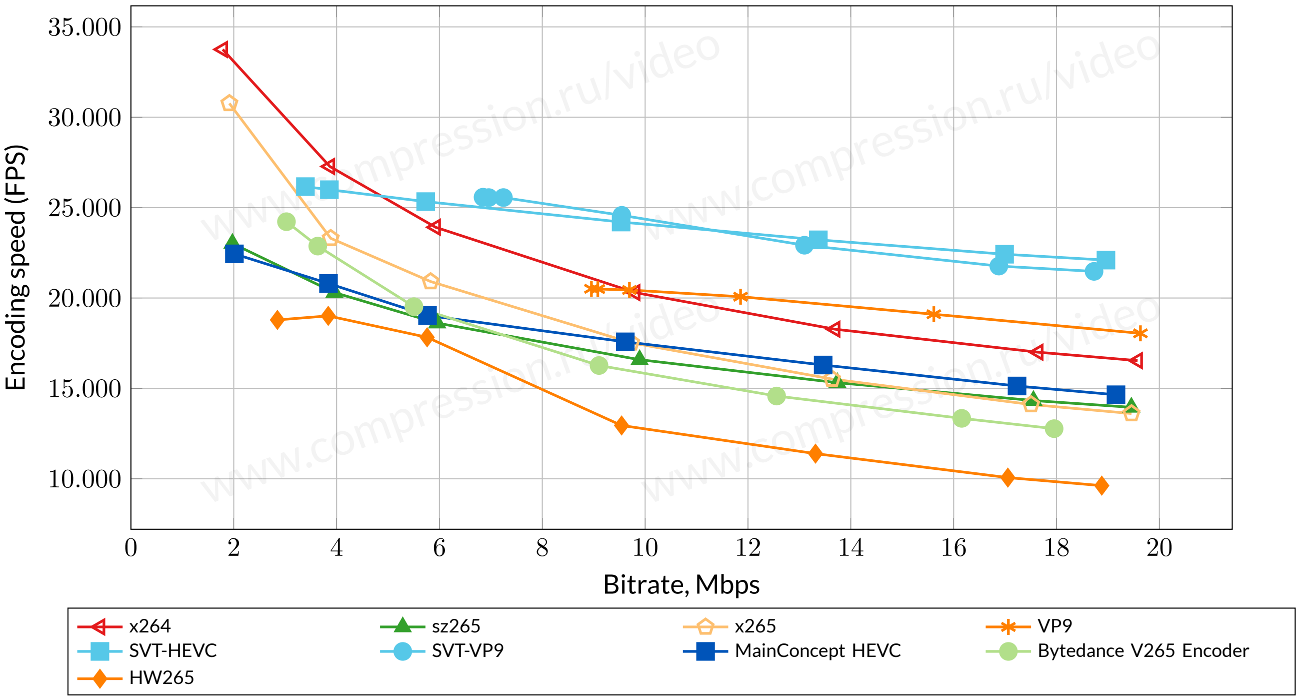
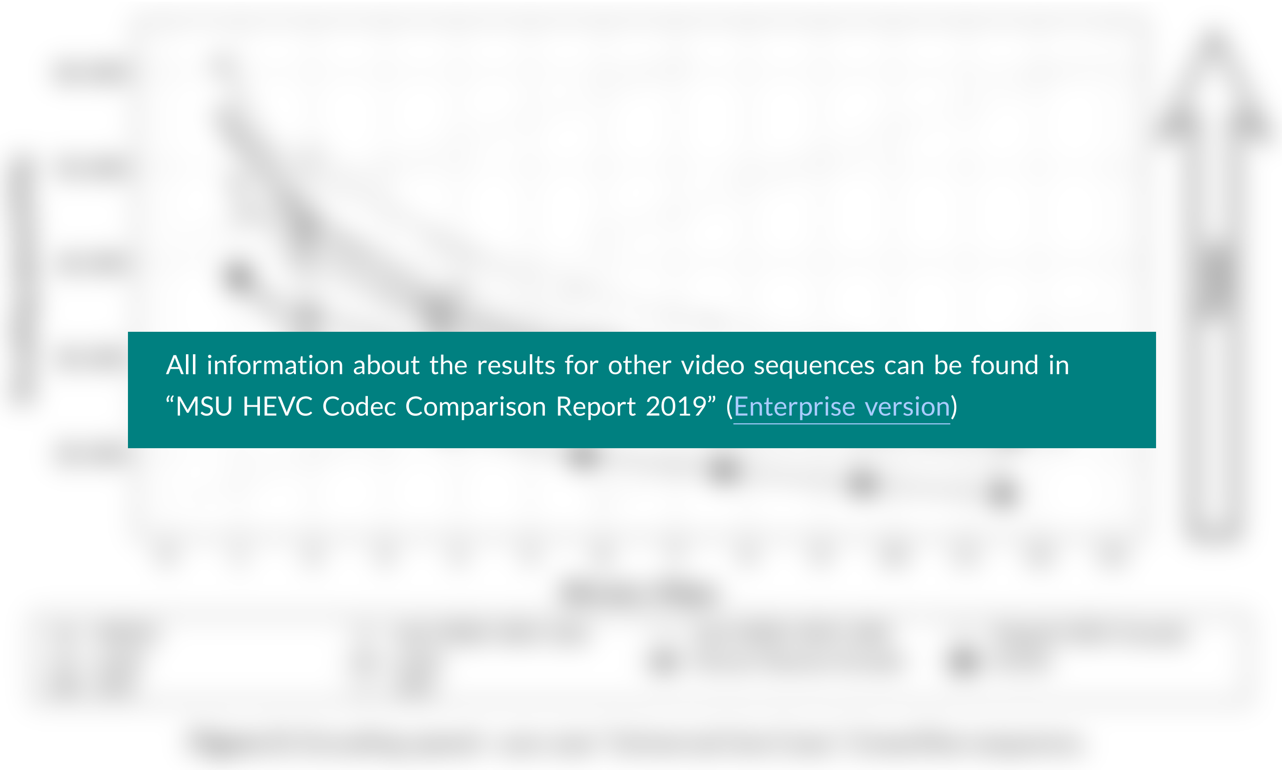


Figure 12: Encoding speed—use case "4K Fast (20fps)," *Ducks take off* sequence.

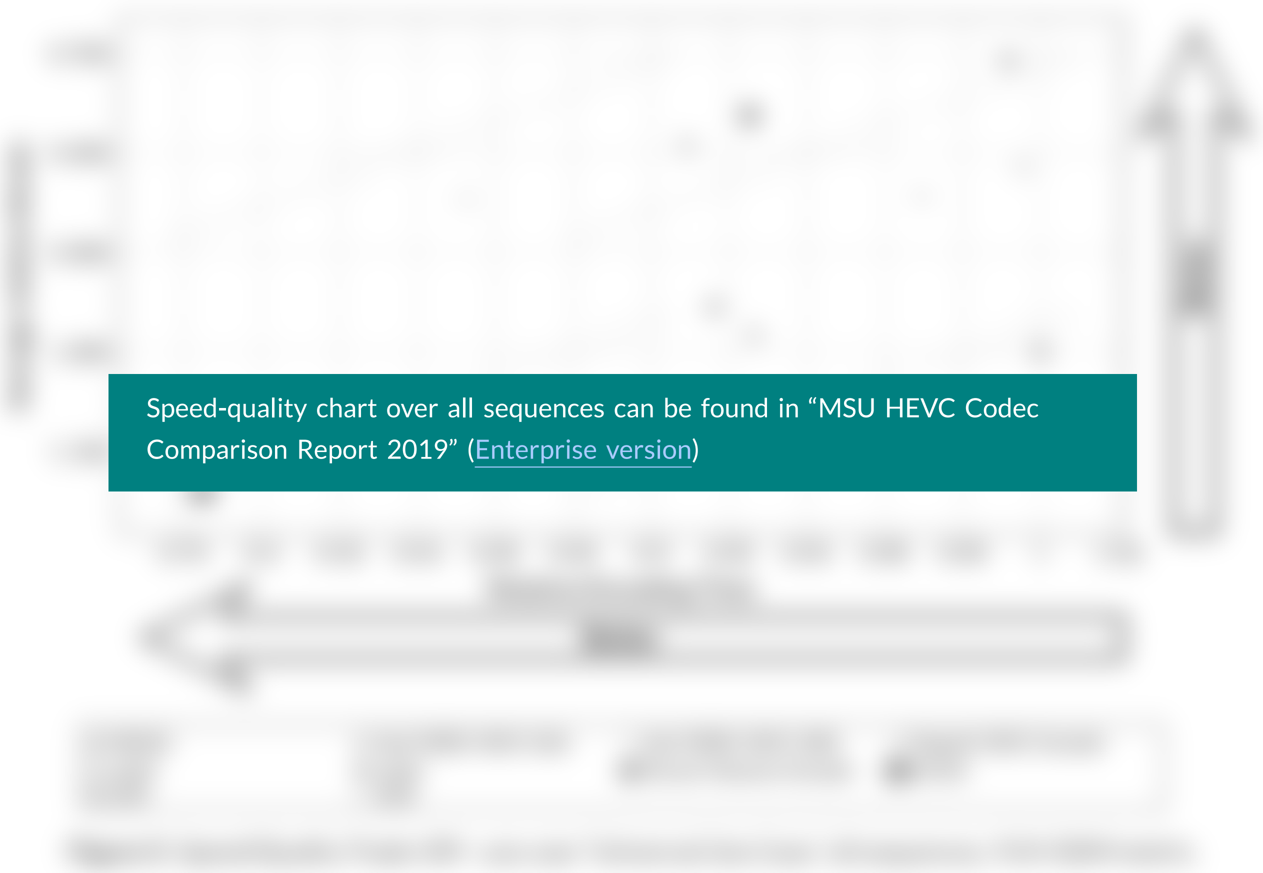


All information about the results for other video sequences can be found in "MSU HEVC Codec Comparison Report 2019" (Enterprise version)

6.3. Speed/Quality Trade-Off

Detailed descriptions of the speed/quality trade-off graphs are in Appendix D. Some graphs omit the results for a particular codec owing to that codec's extremely poor performance (i.e., its RD curve fails to intersect with the reference RD curve).
The speed/quality trade-off graphs show both relative quality and speed scores for the encoders under comparison. Since we chose x264 as the reference codec, we normalized all scores to the x264 scores.

There are six Pareto-optimal encoders: **HW265**, **Bytedance V265 Encoder**, **MainConcept HEVC**, **x265**, **SVT-HEVC**, and **VP9**.



Speed-quality chart over all sequences can be found in “MSU HEVC Codec Comparison Report 2019” (Enterprise version)

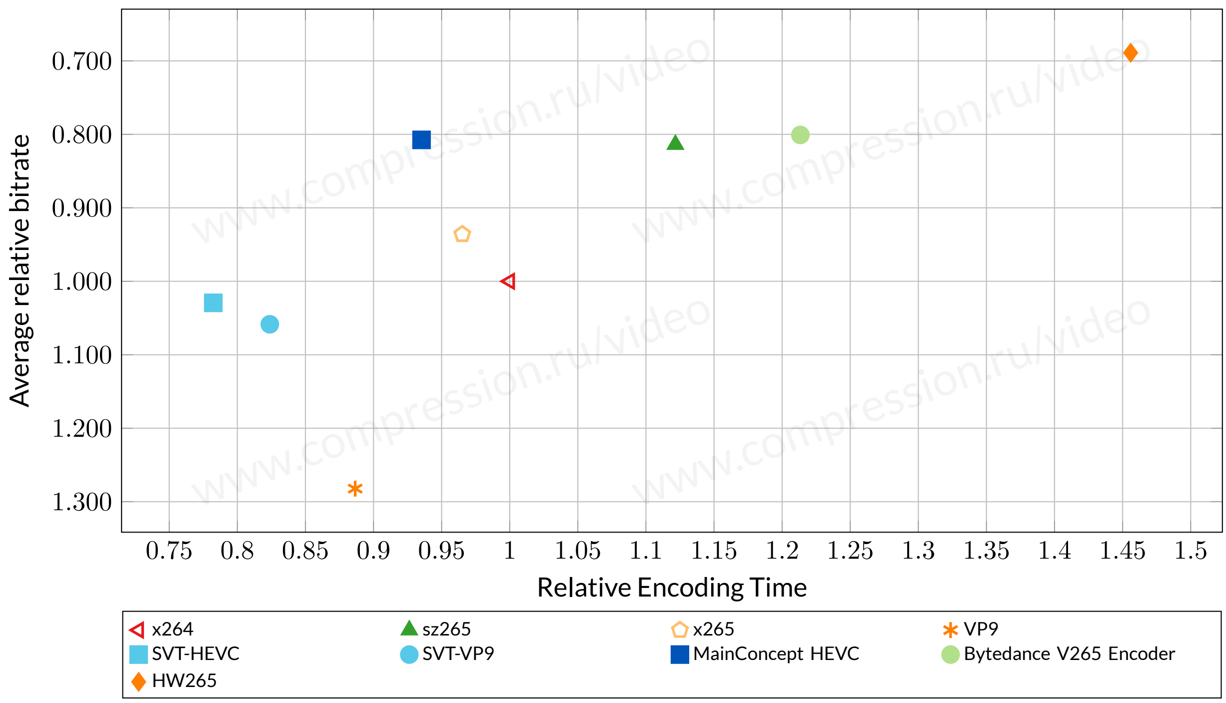


Figure 13: Speed/Quality Trade-Off—use case “4K Fast (20fps),” *Grass* sequence, YUV-SSIM metric.

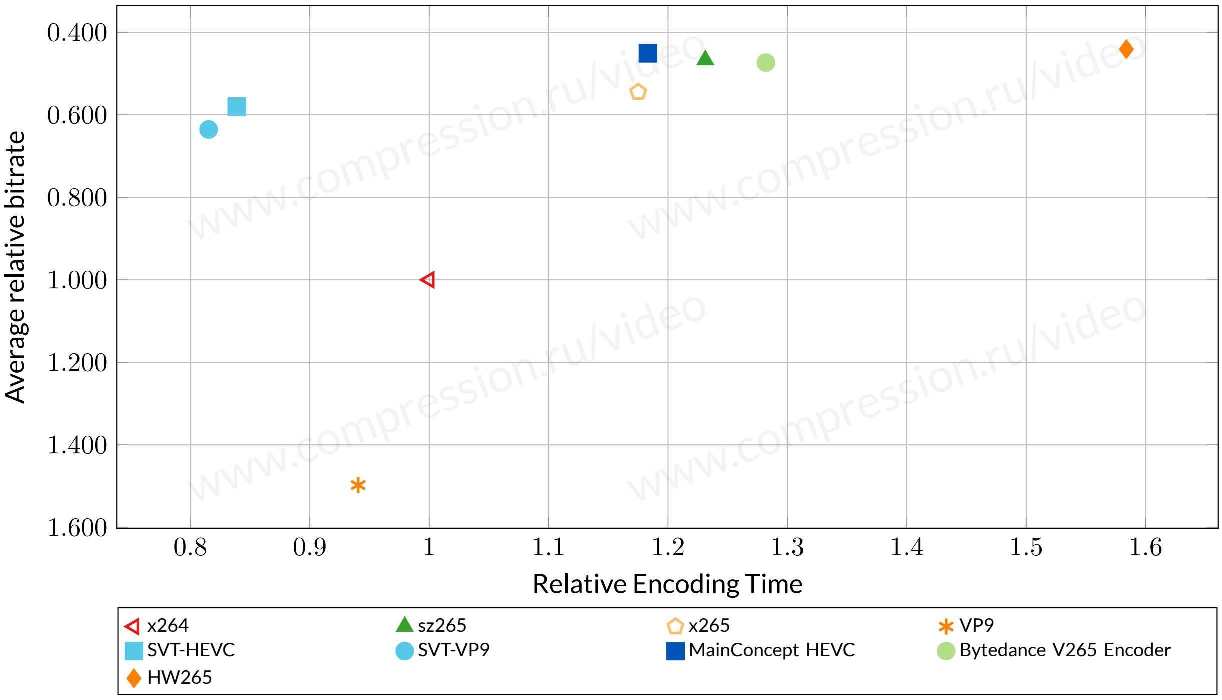
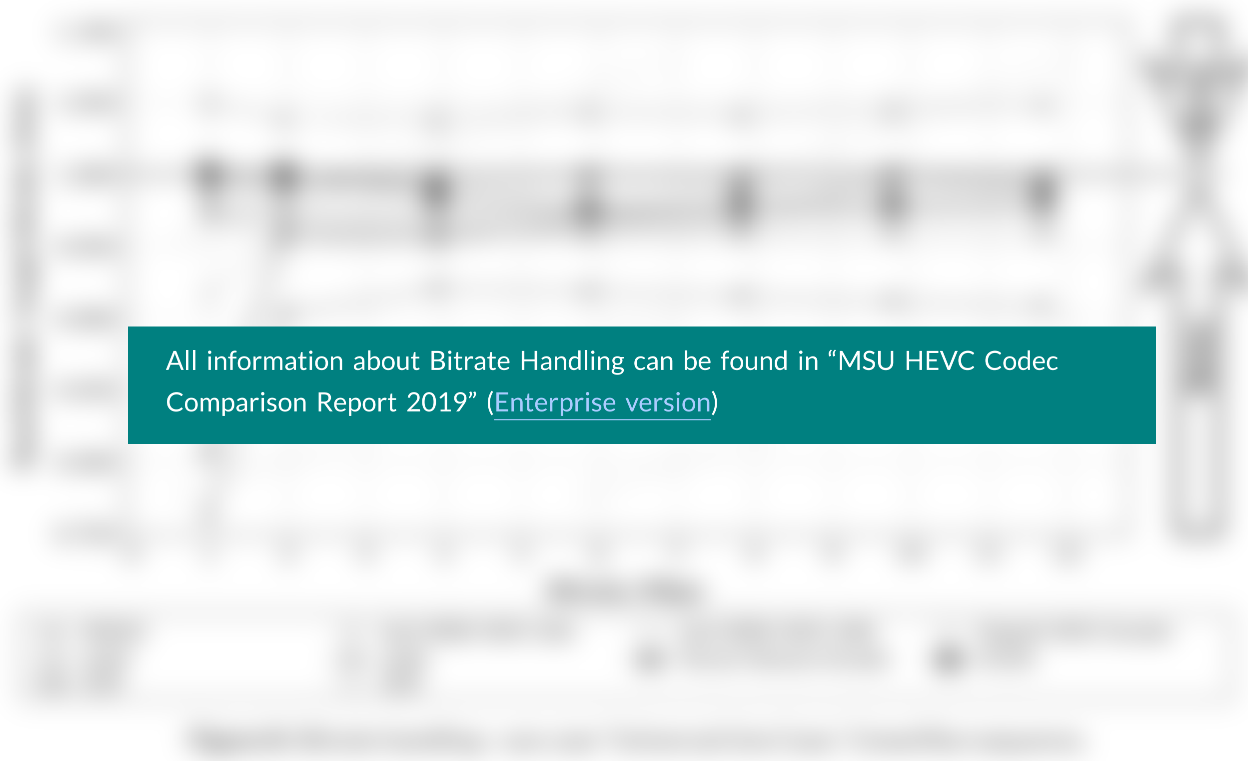


Figure 14: Speed/Quality Trade-Off—use case “4K Fast (20fps),” *Ducks take off* sequence, YUV-SSIM metric.

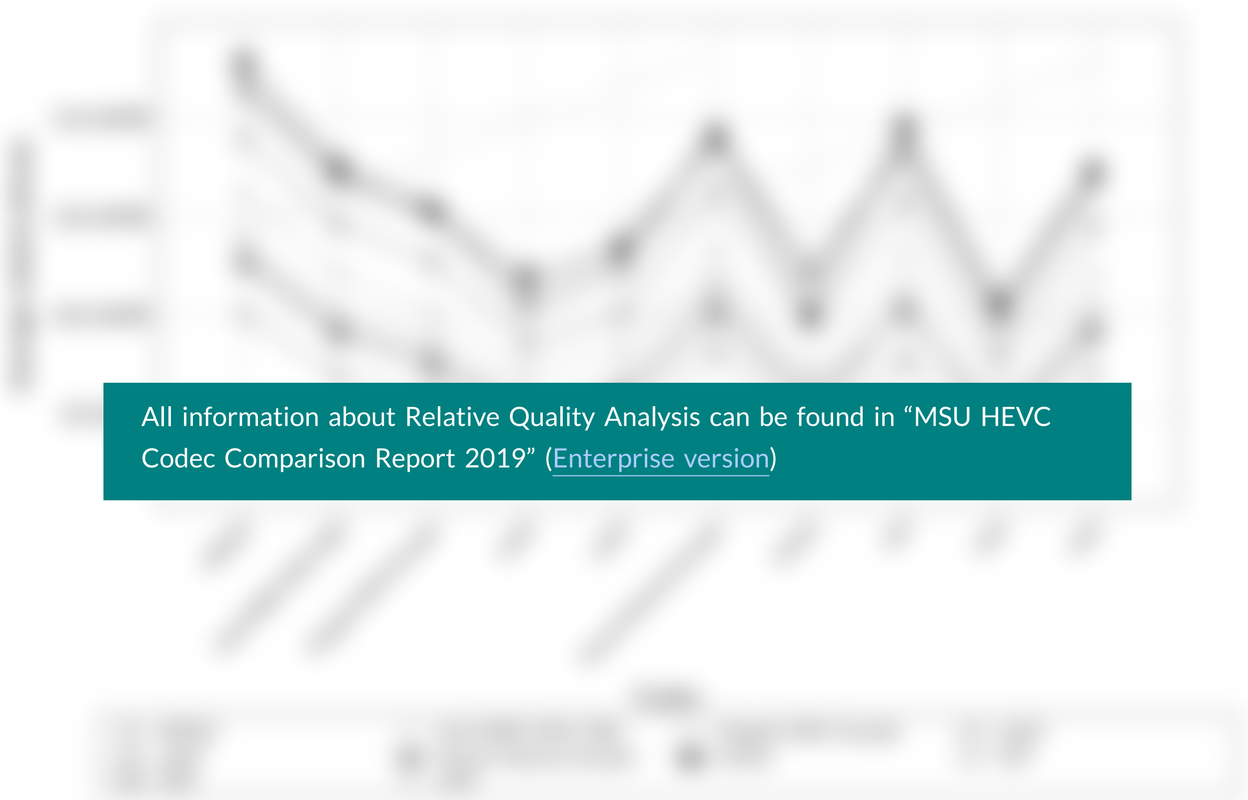
6.4. Bitrate Handling



All information about Bitrate Handling can be found in “MSU HEVC Codec Comparison Report 2019” (Enterprise version)

6.5. Relative Quality Analysis

Note that each number in the tables below corresponds to some range of bitrates (see Appendix D.5). Unfortunately, these ranges can differ significantly because of differences in the quality of compared encoders. This situation can lead to some inadequate results when three or more codecs are compared.



All information about Relative Quality Analysis can be found in "MSU HEVC Codec Comparison Report 2019" (Enterprise version)

7. CONCLUSION

7.1. 4K Universal (1fps)

The plot below shows overall quality scores for the encoders in our comparison (see Section D for a description of the integral-score computation method). First place in the quality competition goes to **WZAurora AV1 Encoder**, second place goes to **HW265**, and third place to **MainConcept HEVC**, **Bytedance V265 Encoder**, and **SVT-AV1**.

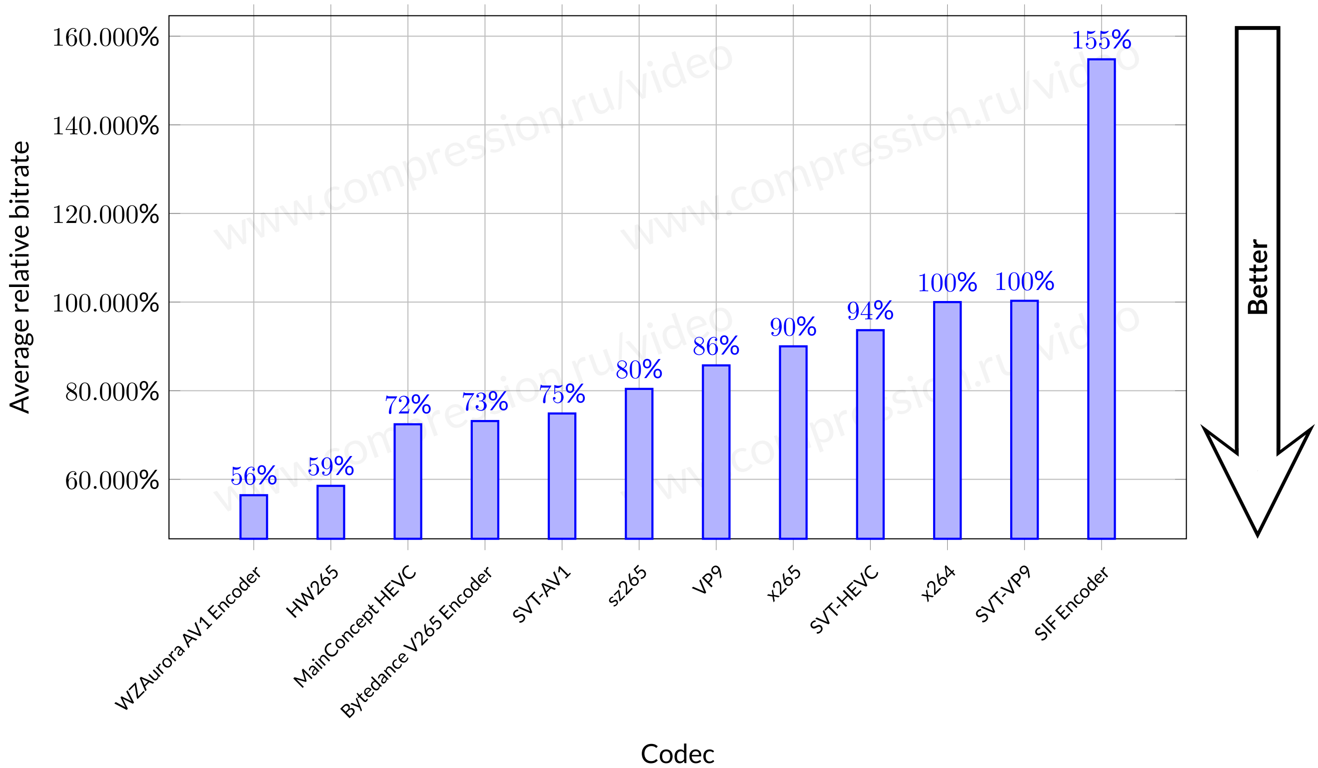


Figure 15: Average bitrate ratio for a fixed quality—use case "4K Universal (1fps)," all sequences, YUV-SSIM metric.

7.2. 4K Fast (20fps)

The plot below shows overall quality scores for the encoders in our comparison (see Section D for a description of the integral-score computation method). First place in the quality competition goes to **HW265**, second place goes to **Bytedance V265 Encoder** and **MainConcept HEVC**, and third place to **sz265**.

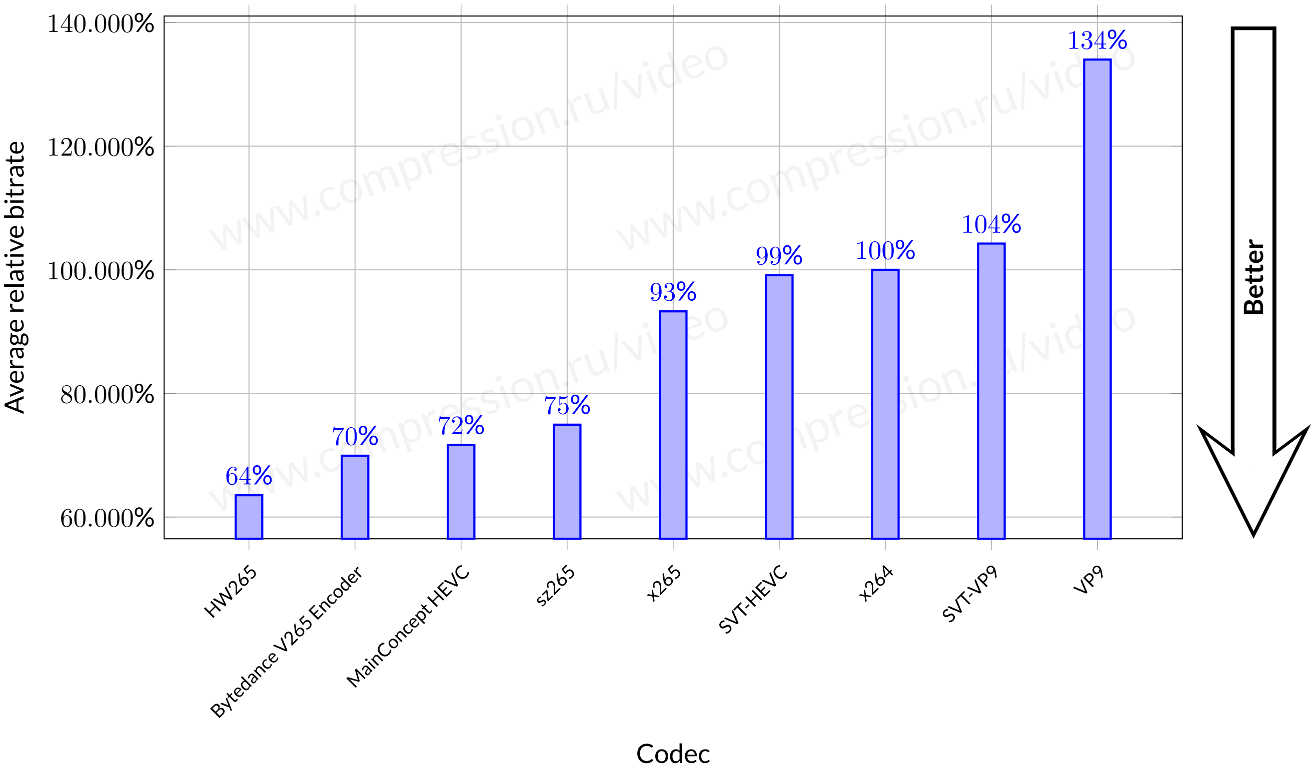


Figure 16: Average bitrate ratio for a fixed quality—use case "4K Fast (20fps)," all sequences, YUV-SSIM metric.

7.3. Overall

The overall score contains only those codecs which participated in all use cases of the comparison. There are several codecs which were compared only in a subset of comparison use cases, and their results are presented in separate charts for each encoding use case.

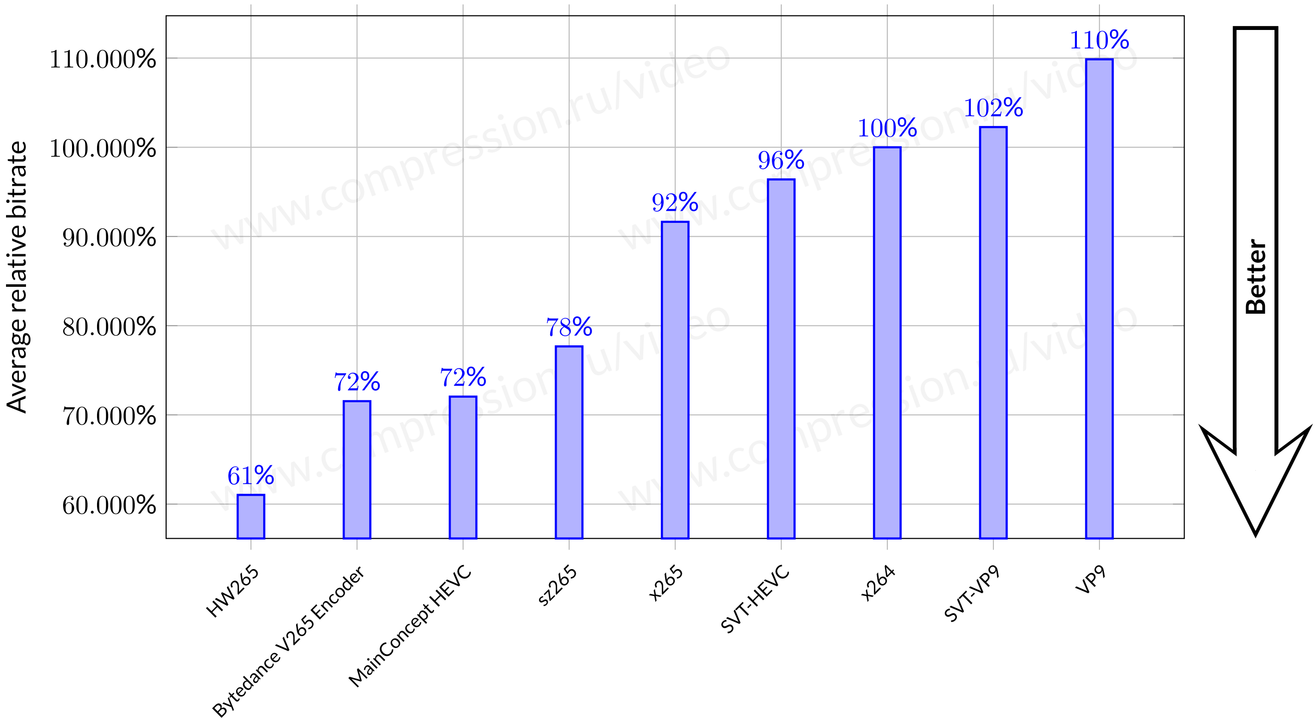


Figure 17: Average bitrate ratio for a fixed quality—all sequences, YUV-SSIM metric.

7.4. Overall over VMAF metric

The overall score contains only those codecs which participated in all use cases of the comparison. There are several codecs which were compared only in a subset of comparison use cases, and their results are presented in separate charts for each encoding use case.

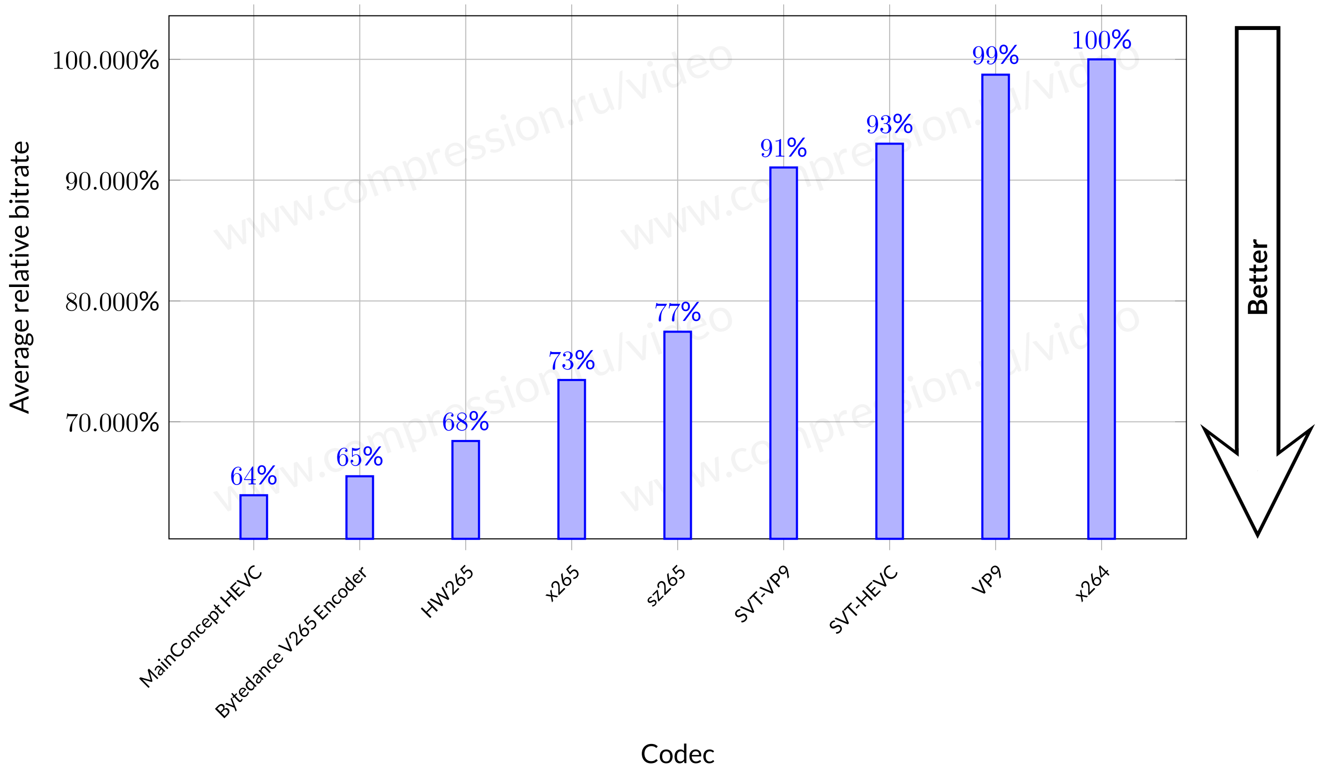
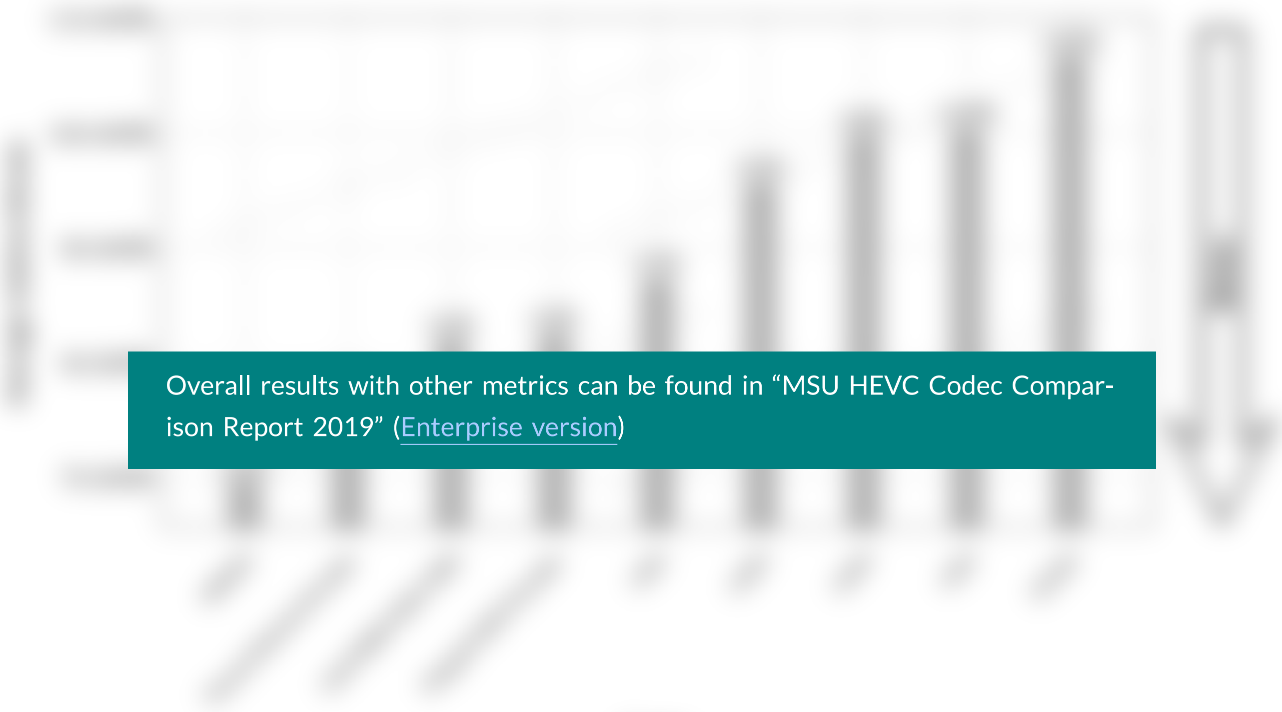


Figure 18: Average bitrate ratio for a fixed quality—all sequences, Y-VMAF (v0.6.2) metric.



Overall results with other metrics can be found in “MSU HEVC Codec Comparison Report 2019” (Enterprise version)

A. SEQUENCES

Direct download links to video sequences used in this comparison can be found in "MSU HEVC Codec Comparison Report 2019" (Enterprise version)

A.1. Crowd run

Sequence title	Crowd run
Resolution	3840×2160
Number of frames	500
Color space	YV12
Frames per second	50
Source resolution	4K
Bitrate	4976.64

A crowd of sportsmen runs while the camera slowly moves left and right.

Figure 19: Crowd run sequence, frame 404

A.2. Ducks take off

Sequence title	Ducks take off
Resolution	3840×2160
Number of frames	500
Color space	YV12
Frames per second	50
Source resolution	4K
Bitrate	4976.64

The flock of ducks takes off the pond.

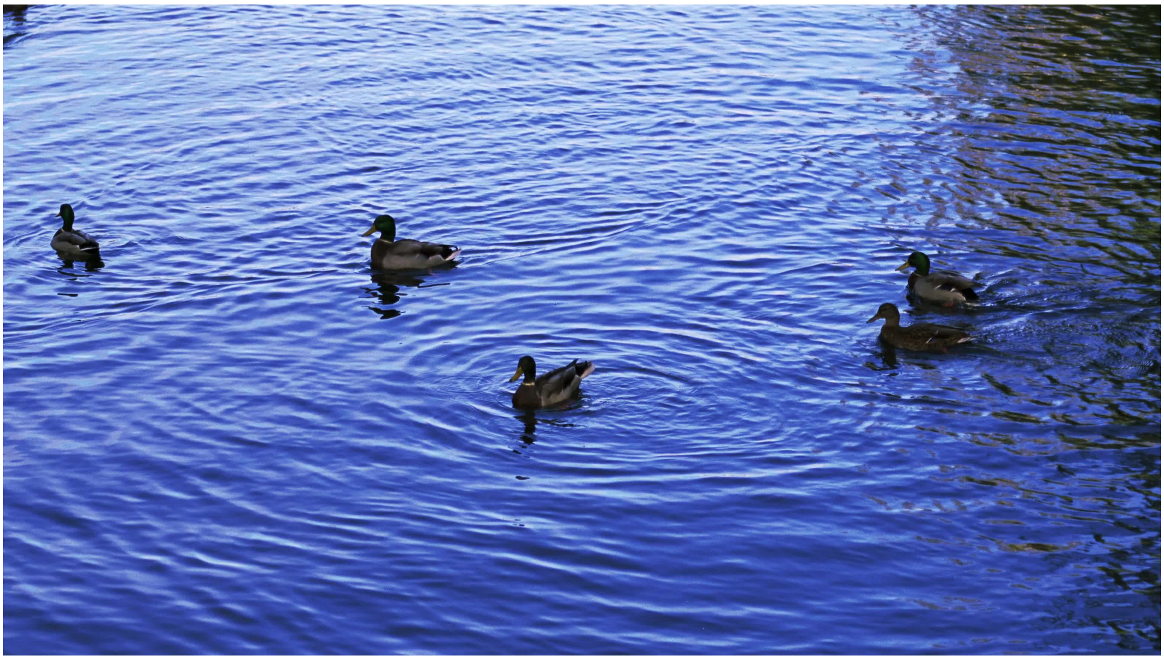


Figure 20: Ducks take off sequence, frame 27

A.3. Flat Tour

Sequence title	Flat Tour
Resolution	3840×2160
Number of frames	926
Color space	YV12
Frames per second	60
Source resolution	4K
Bitrate	124.77

Flat tour smooth scenes.

Figure 21: Flat Tour sequence, frame 243

A.4. Grass

Sequence title	Grass
Resolution	3840×2160
Number of frames	576
Color space	YV12
Frames per second	24
Source resolution	4K
Bitrate	83.31

Grass in the wind at sunset.

Figure 22: Grass sequence, frame 1

A.5. Hiking

Sequence title	Hiking
Resolution	3840×2160
Number of frames	1303
Color space	YV12
Frames per second	24
Source resolution	4K
Bitrate	151.77

Hiking in hills.

Figure 23: Hiking sequence, frame 599

A.6. Hotel tour

Sequence title	Hotel tour
Resolution	3840×2160
Number of frames	1077
Color space	YV12
Frames per second	30
Source resolution	4K
Bitrate	545.32

Hotel tour drone scenes.

Figure 24: Hotel tour sequence, frame 984

A.7. Mallorca

Sequence title	Mallorca
Resolution	3840×2160
Number of frames	1104
Color space	YV12
Frames per second	30
Source resolution	4K
Bitrate	466.58

Seacoasts drone views with fast transitions.

Figure 25: Mallorca sequence, frame 976

A.8. New Masons

Sequence title	New Masons
Resolution	3840×2160
Number of frames	1359
Color space	YV12
Frames per second	24
Source resolution	4K
Bitrate	113.66

Noisy musical clip with short scenes.

Figure 26: New Masons sequence, frame 341

A.9. OPCW

Sequence title	OPCW
Resolution	3840×2160
Number of frames	1015
Color space	YV12
Frames per second	25
Source resolution	4K
Bitrate	252.45

Short scenes with dynamic background.

Figure 27: OPCW sequence, frame 62

A.10. Photo shoot

Sequence title	Photo shoot
Resolution	3840×2160
Number of frames	1053
Color space	YV12
Frames per second	30
Source resolution	4K
Bitrate	873.05

Photo shoot behind scenes.

Figure 28: Photo shoot sequence, frame 787

A.11. Wedding

Sequence title	Wedding
Resolution	3840×2160
Number of frames	1015
Color space	YV12
Frames per second	24
Source resolution	4K
Bitrate	132.55

Wedding party footage.

Figure 29: Wedding sequence, frame 119

B. SEQUENCE SELECTION

In "MSU Video Codecs Comparison 2016" we introduced a new technique for selecting test sequences. This technique create a data set containing representative sequences that encoders face in everyday situations. For this report we use the same method, but we updated the video database from which we sample videos.

We analyzed 384,946 videos at Vimeo, looking for 4K examples with high bitrates (we chose 50 Mbps as our minimum). Doing so enabled us to find and download 145 new 4K videos. Figure 30 shows the bitrate distributions for last year's data set and for the updated data set.

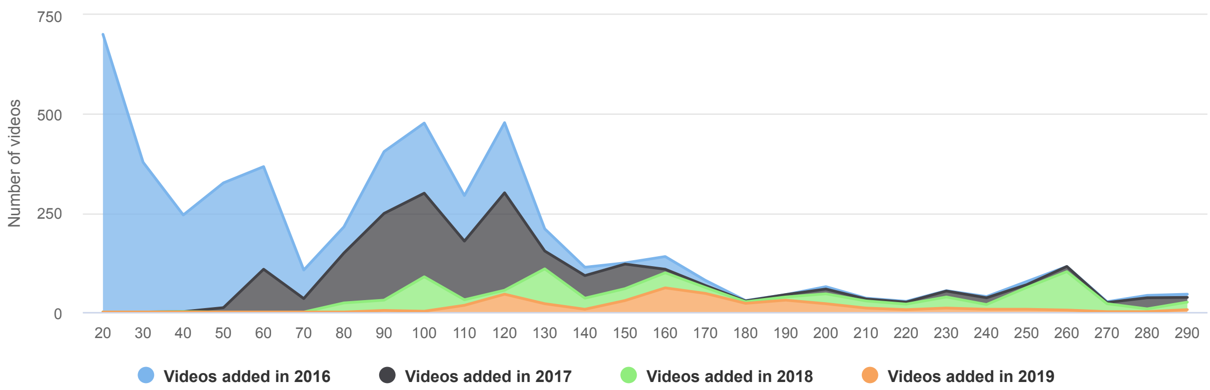


Figure 30: Bitrate distributions for video data set.

For 4K Appendix we use only 4K videos from downloaded collection. All videos were cut at scene change points to samples, with 1000 frames approximate length. Besides 513 samples from newly downloaded videos, we also used 965 samples from "MSU Video Codecs Comparison 2018", 1,659 samples from "MSU Video Codecs Comparison 2017" and 2,906 samples from "MSU Video Codecs Comparison 2016". Thus, our sample database for this year consisted of 6,043 items.

To evaluate spatial and temporal complexity, we encoded all samples using x264 with a constant quantization parameter (QP). We calculated the temporal and spatial complexity for each scene, defining spatial complexity as the average size of the I-frame normalized to the sample's uncompressed frame size. Temporal complexity in our definition is the average size of the P-frame divided by average size of I-frame. [1]

In this year we slightly changed the temporal and spatial complexity calculation process by adding an additional preprocessing step. We use sourse videos from Vimeo, that was uploaded by users, so they all have different chroma subsampling which affects the results of videos evaluated complexity. Therefore to unificate the spatial and temporal complexity results of analysed videos, they all were converted to YUV 4:2:0 chroma subsample. Distribution of obtained samples compared to samples from previous codec comparisons is shown in Figure 31.

[1] C. Chen et. al., "A Subjective Study for the Design of Multi-resolution ABR Video Streams with the VP9 Codec," 2016.

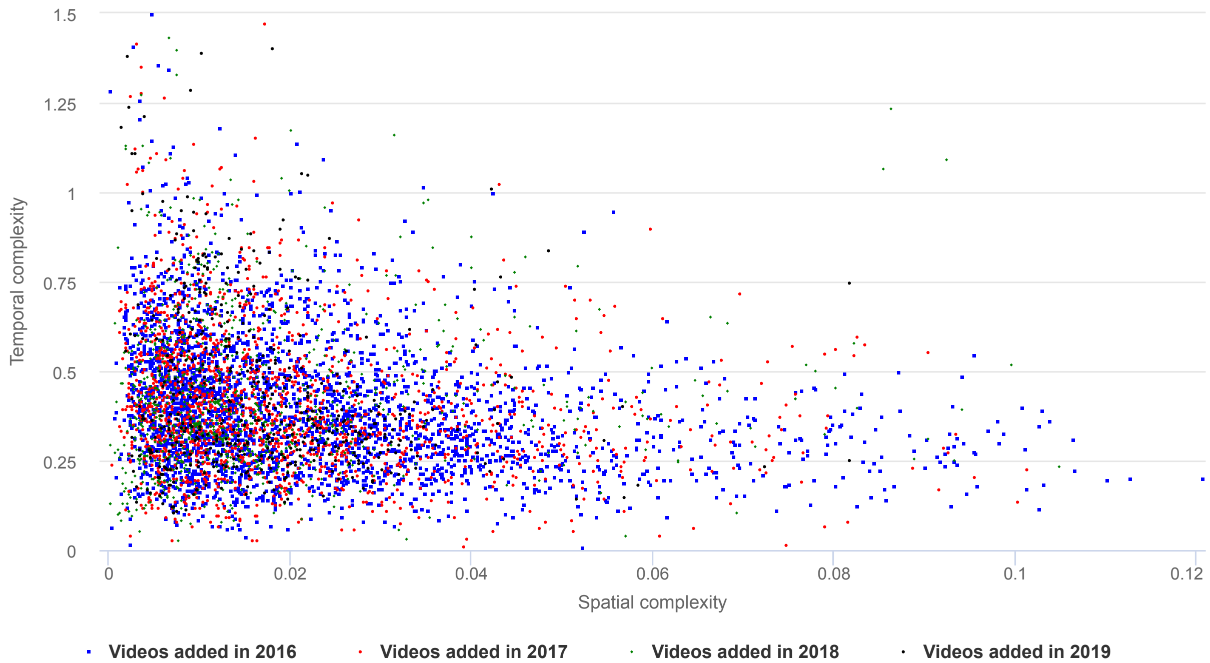


Figure 31: Distribution of obtained samples.

Figure 31 reveals that the new samples have a distribution similar to that of samples from "MSU Video Codecs Comparison 2017". We used the following process to prepare the data set.

We divided the video database into 11 clusters. To avoid completely changing the sequence list, we gave sequences from last year's 4K data set 35 times greater weight than other sequences. For each cluster we selected the video sequence that's closest to the cluster's center and that has a license enabling derivatives and commercial use. Figure 32 shows the cluster boundaries and constituent sequences.

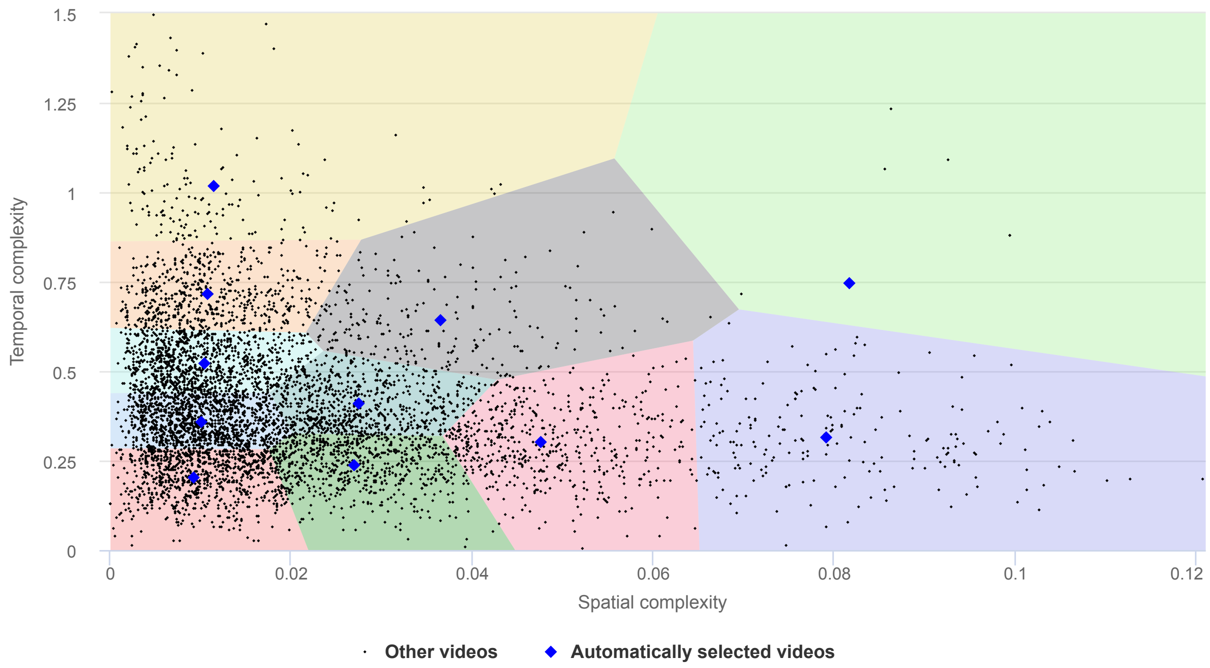


Figure 32: Segmentation of samples.

Figure 33 shows the correspondence of sequences from the previous data set to the newly selected ones. As the figure demonstrates, after adding a preprocessing step for video sequences some clusters don't include videos from old data set.

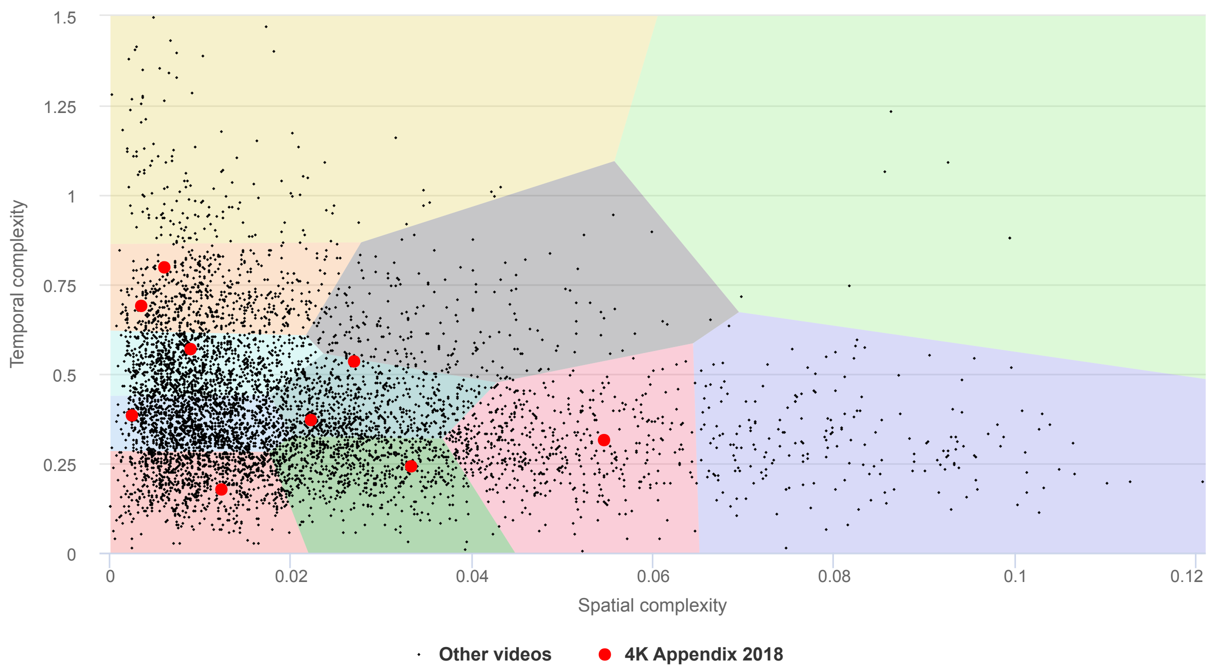


Figure 33: Segmentation of samples relative to old data set.

Some automatically chosen samples contain company names or have other copyright issues, so we removed them from their respective clusters and replaced them with other samples having a suitable license. Figure 34 illustrates these adjustments.

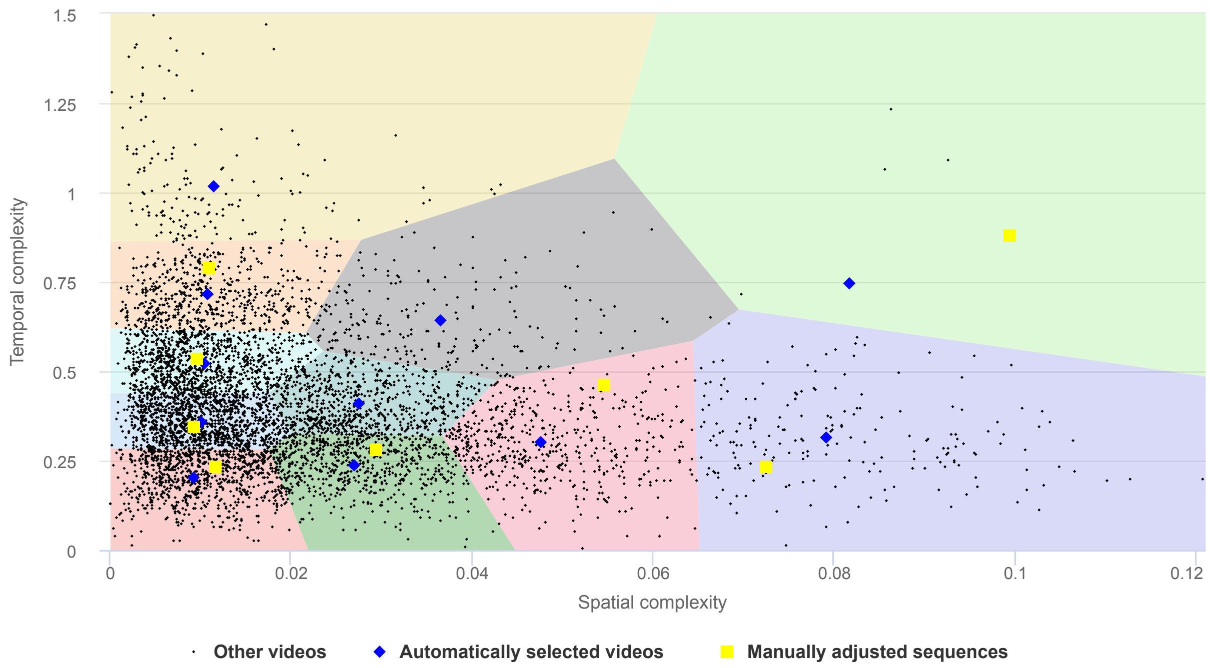


Figure 34: Adjustments to test data set.

Figure 35 shows the final distribution of sequences in the data set.

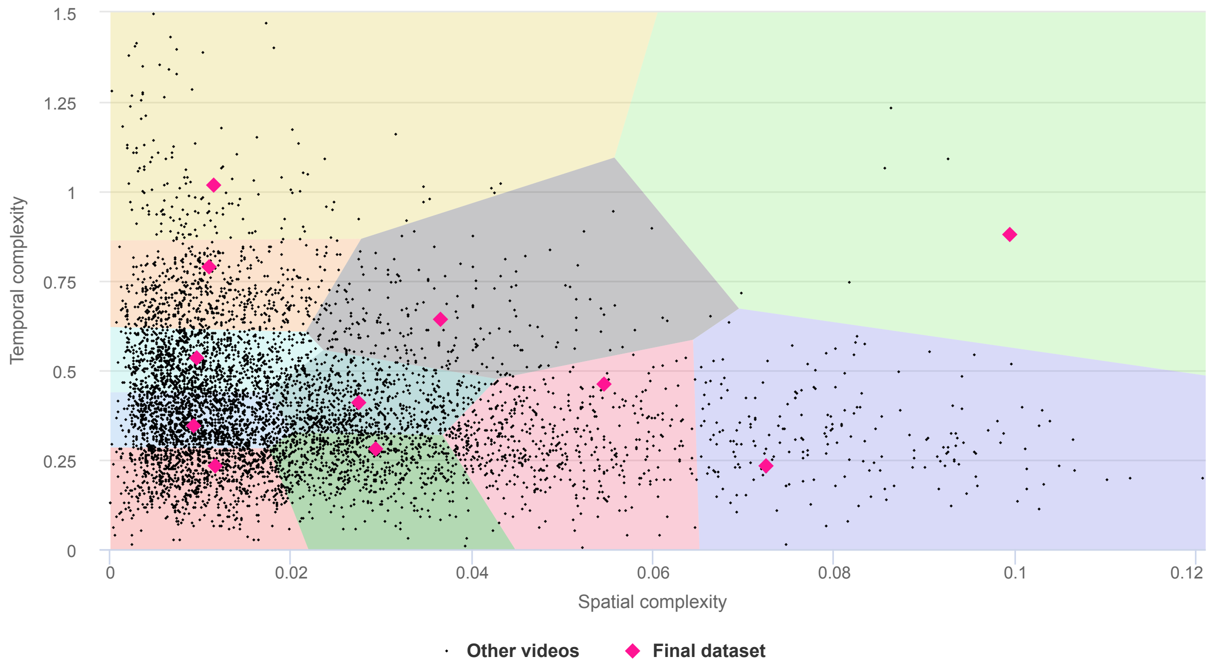


Figure 35: Distribution of sequences in final data set.

The new data set consists of 11 sequences: 9 new ones from Vimeo and 2 from xiph.org. The average bitrate for all sequences in the final set is 1,154.25 Mbps, median – 252.45 Mbps. The complete list of sequences for new data set appears in Appendix A.

C. CODECS

All tested encoders presets can be found in "MSU HEVC Codec Comparison Report 2019" (Enterprise version)

D. FIGURE EXPLANATION

The main charts in this comparison are classic RD curves (quality/bitrate graphs) and relative-bitrate/relative-time charts. Additionally, we also used bitrate-handling charts (the ratio of real to target bitrates) and per-frame quality charts.

D.1. RD Curves

The RD charts show variation in codec quality by bitrate or file size. For this metric, a higher value presumably indicates better quality.

D.2. Relative-Bitrate/Relative-Time Charts

Relative-bitrate/relative-time charts show the average bitrate's dependence on relative encoding time for a fixed-quality output. The y-axis shows the ratio of a codec's bitrate under test to the reference codec's bitrate for a fixed quality. A lower value (that is, a higher the value on the graph) indicates a better-performing codec. For example, a value of 0.7 means the codec can encode the sequence in a file that's 30% smaller what the reference codec produces.

The x-axis shows the relative encoding time. Larger values indicate a slower codec. For example, a value of 2.5 means the codec works 2.5 times slower, on average, than the reference codec.

D.3. Graph Example

Figure 36 shows a situation where these graphs can be useful. In the top-left graph, the "Green" codec clearly produces better quality than the "Black" codec. On the other hand, the top-right graph shows that the "Green" codec is slightly slower. Relative-bitrate/relative-time graphs can be useful in precisely these situations: the bottom graph clearly shows that one codec is slower but yields higher visual quality, whereas the other codec is faster but yields lower visual quality.

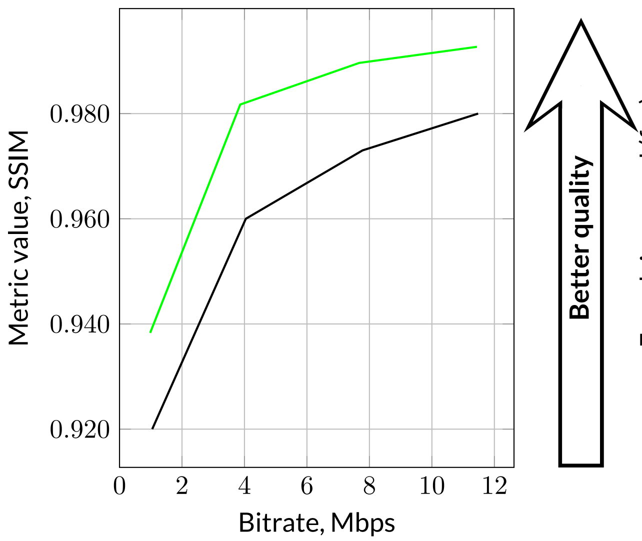
Owing to these advantages, we frequently use relative-bitrate/relative-time graphs in this report because they assist in evaluating the codecs in the test set, especially when the number of codecs is large.

A more detailed description of how we prepared these graphs appears below.

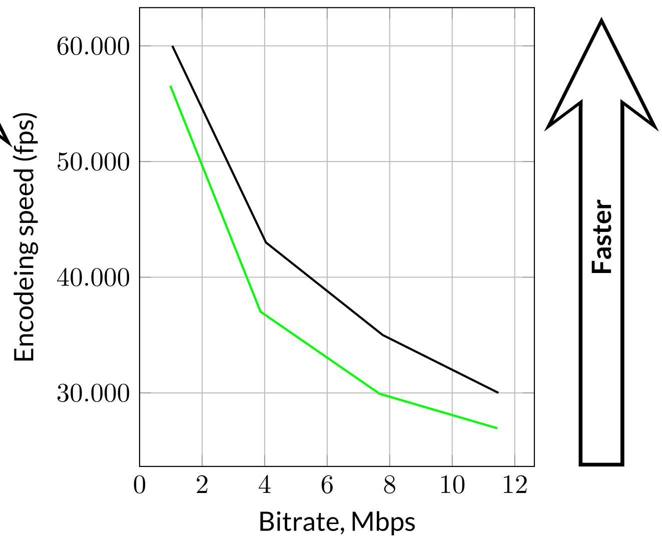
D.4. Bitrate Ratio for the Same Quality

The first step in computing the average bitrate ratio for a fixed quality is to invert the axes of the bitrate/quality graph (see Figure 37b). All further computations use the inverted graph.

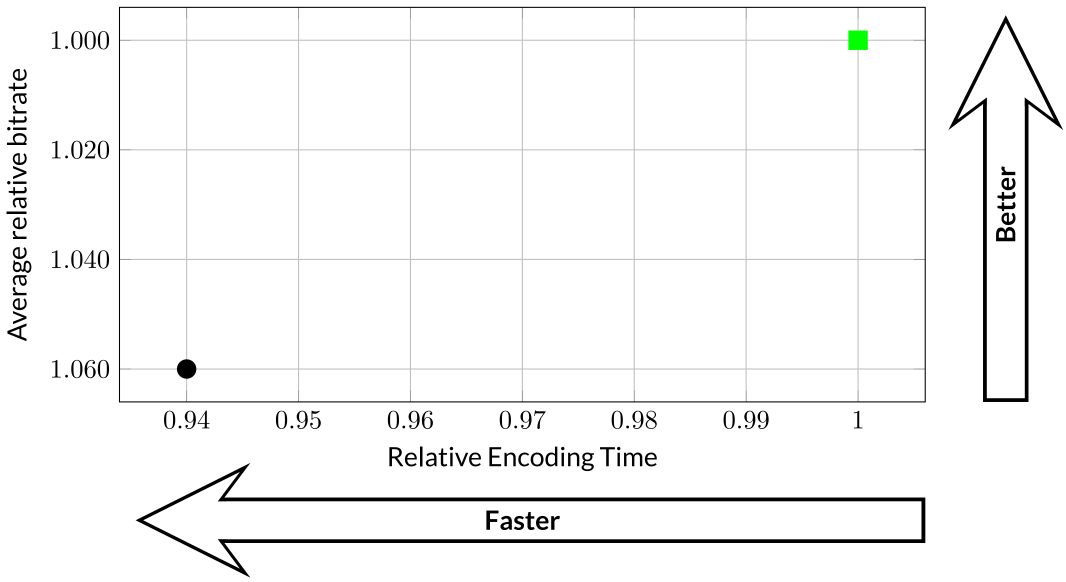
The second step involves averaging the interval over which the quality axis is chosen. The averaging is only over those segments for which both codecs yield results. This limitation is due to the difficulty of developing extrap-



(a) RD curve. “Green” codec is better!

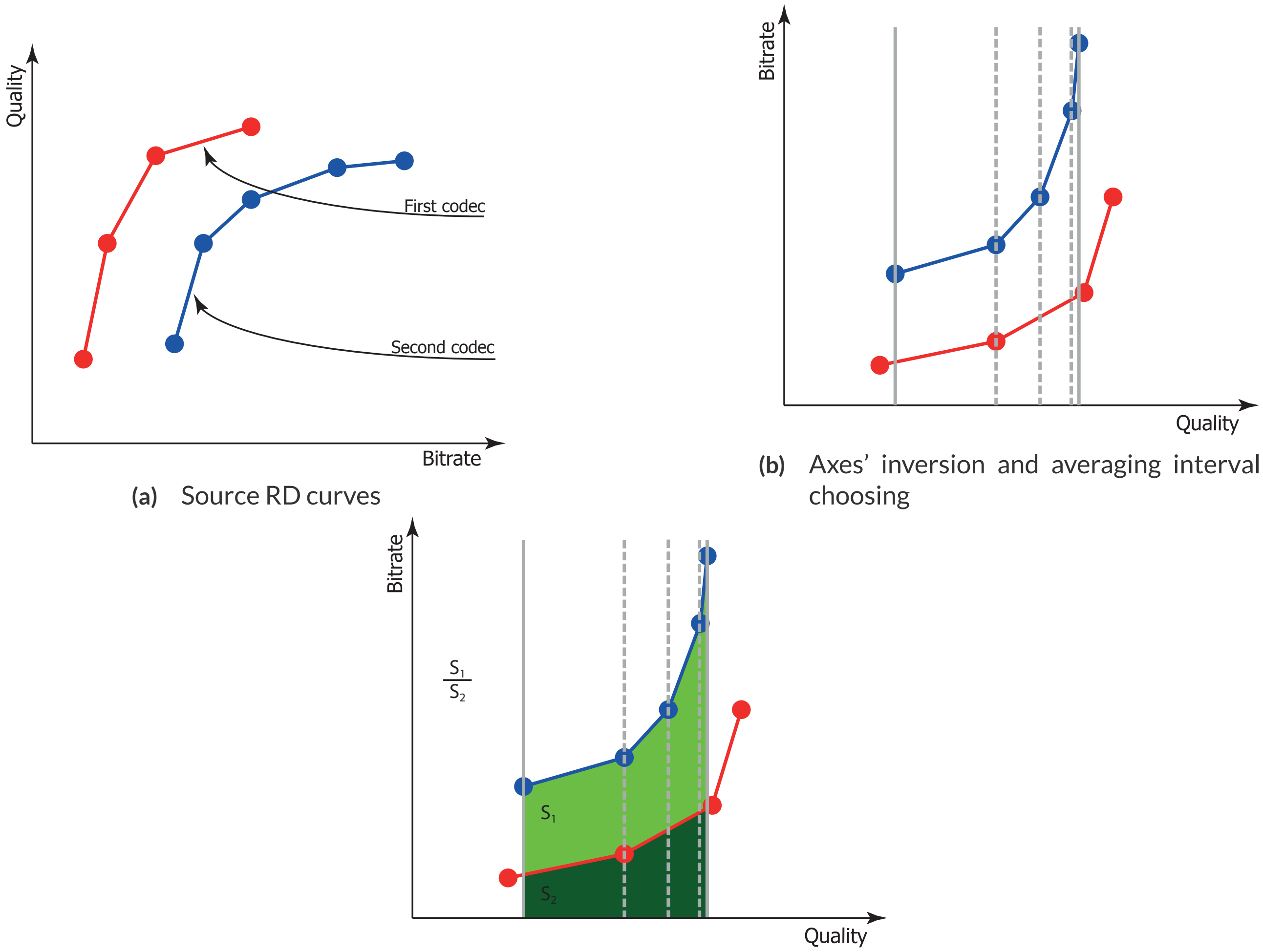


(b) Encoding speed (frames per second). “Green” codec is slower!



(c) Integral situation with codecs. This plot shows the situation more clearly

Figure 36: Speed/Quality trade-off example



(a) Source RD curves

(b) Axes' inversion and averaging interval choosing

(c) Areas under curves ratio

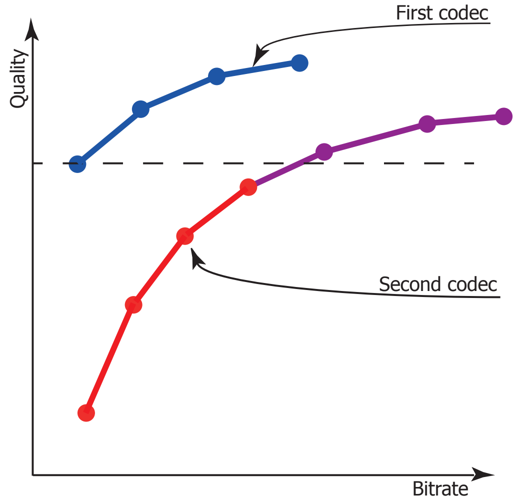
Figure 37: Average bitrate ratio computation

olation methods for classic RD curves; nevertheless, even linear methods are acceptable when interpolating RD curves.

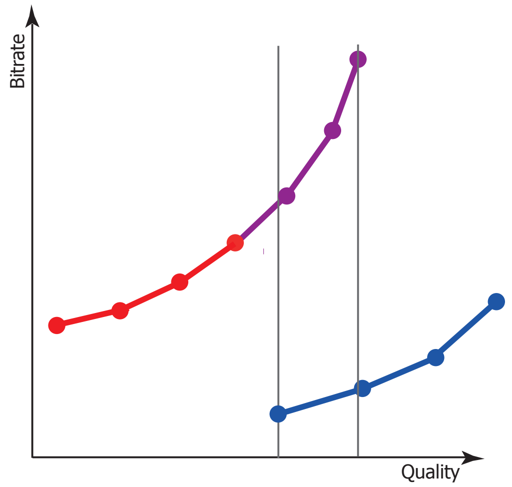
The final step is calculation of the area under the curves in the chosen interpolation segment and determination of their ratio (see Figure 37c). This result is an average bitrate ratio at a fixed quality for the two codecs. When considering more than two codecs, one of is defined as a reference codec, and the quality of the others is compared with that of the reference.

D.4.1. When RD Curves Fail to Cross the Quality Axis

If no segment exists for which two codecs both produce encoding results, we measured the results for additional higher and/or lower bitrates. The schematic example (Figure 38) shows that the results for these extra bitrates (purple) cross with codec two and enable a comparison with codec one.



(a) Source RD curves, purple color indicates results for extra bitrates

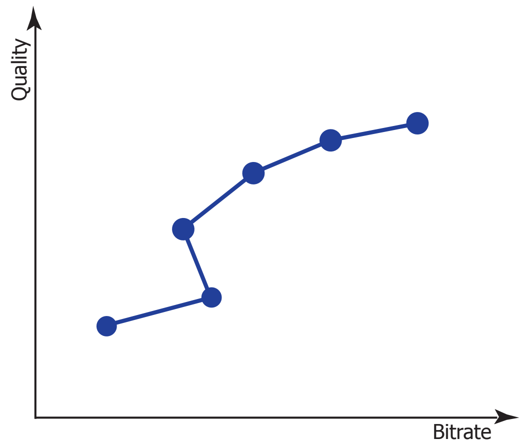


(b) Axes' inversion and averaging interval choosing

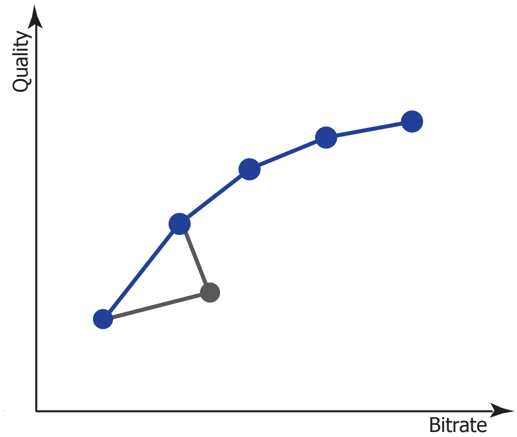
Figure 38: Measuring codec on additional bitrates to make it cross with other codecs over the quality axis.

D.4.2. When RD Curves Are Non-monotonic

Sometimes, especially on complex videos, the encoding results for neighboring bitrates vary greatly owing to the codec's operating characteristics. This situation leads to a non-monotone RD curve, which we process as follows: for each point, use the next point at the target bitrate that has greater or equal quality. This technique yields the reduced monotonic curve, which appears in the example of Figure 39.



(a) Non-monotonic RD-curve.



(b) Points that were used to calculate integral.

Figure 39: Processing non-monotonic RD-curves.

D.5. Relative Quality Analysis

Although most figures in this report provide codec scores relative to a reference encoder (i.e., x264), the "Relative Quality Analysis" sections provide the bitrate ratio at a fixed quality score (see Section D.4) for each codec pair.

This approach may be useful when comparing codec A relative with codec B only.

Below is a simplified example table showing the average bitrate ratio, given a fixed quality, for just two codecs.

	A	B
A	100% ☺	75% ☹
B	134% ☹	100% ☺

Table 3: Example of average bitrate ratio for a fixed quality table

Consider column "B", row "A" of the table, which contains the value 75%. This number should be interpreted in the following way: the average bitrate for Codec B at a fixed quality is 75% less than that for codec A. The icon in the cell depicts the confidence of this estimate. If projections of RD curves on the quality axis (see Figure 37) have large common areas, the cell contains a happy icon. If this overlapping area is small, and thus the bitrate-score calculation is unreliable, the cell contains a sad icon.

Plots of the average bitrate ratio for a fixed quality are visualizations of these tables. Each line in the plot depicts values from one column of the corresponding table.

E. OBJECTIVE-QUALITY METRIC DESCRIPTION

E.1. SSIM (Structural Similarity)

We used the YUV-SSIM objective-quality metric in this report to assess the quality of encoded video sequences. We compute YUV-SSIM as the weighted average of SSIM values for each channel individually (Y-SSIM, U-SSIM and V-SSIM):

$$\text{YUV-SSIM} = \frac{4\,\text{Y-SSIM} + \text{U-SSIM} + \text{V-SSIM}}{6}. \tag{1}$$

Below is a brief description of SSIM computation.

E.1.1. Brief Description

Wang, et al.[2] published the original paper on SSIM. This paper available at http://ieeexplore.ieee.org/iel5/83/28667/01284395.pdf. The SSIM author homepage is http://www.cns.nyu.edu/~lcv/ssim/

The main idea that underlies the structural-similarity (SSIM) index is comparison of the distortion of three image components:

- Luminance
- Contrast
- Structure

The final formula, after combining these comparisons, is

$$\textbf{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x + \mu_y + C_1)(\sigma_x + \sigma_y + C_2)}, \tag{2}$$

where

$$\mu_x = \sum_{i=1}^{N} \omega_i x_i, \tag{3}$$

$$\sigma_x = \sqrt{\sum_{i=1}^{N} \omega_i (x_i - \mu_x)}, \tag{4}$$

$$\sigma_{xy} = \sum_{i=1}^{N} \omega_i (x_i - \mu_x)(y_i - \mu_y). \tag{5}$$

Finally, $C_1 = (K_1 L)^2$ and $C_2 = (K_2 L)^2$, where L is the dynamic range of the pixel values (e.g. 255 for 8-bit greyscale images), and $K1, K2 \ll 1$.

We used $K_1 = 0.01$ and $K_2 = 0.03$ were used for the comparison presented in this report, and we filled the matrix with a value "1" in each position to form a filter for the results map.

[2] Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh and Eero P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," IEEE Transactions on Image Processing, Vol. 13, No. 4, April 2004.

For our implementation, one SSIM value corresponds to two sequences. The value is in the range $[-1, 1]$, with higher values being more desirable (a value of 1 corresponds to identical frames). One advantage of the SSIM metric is that it better represents human visual perception than does PSNR. SSIM is more complex, however, and takes longer to calculate.

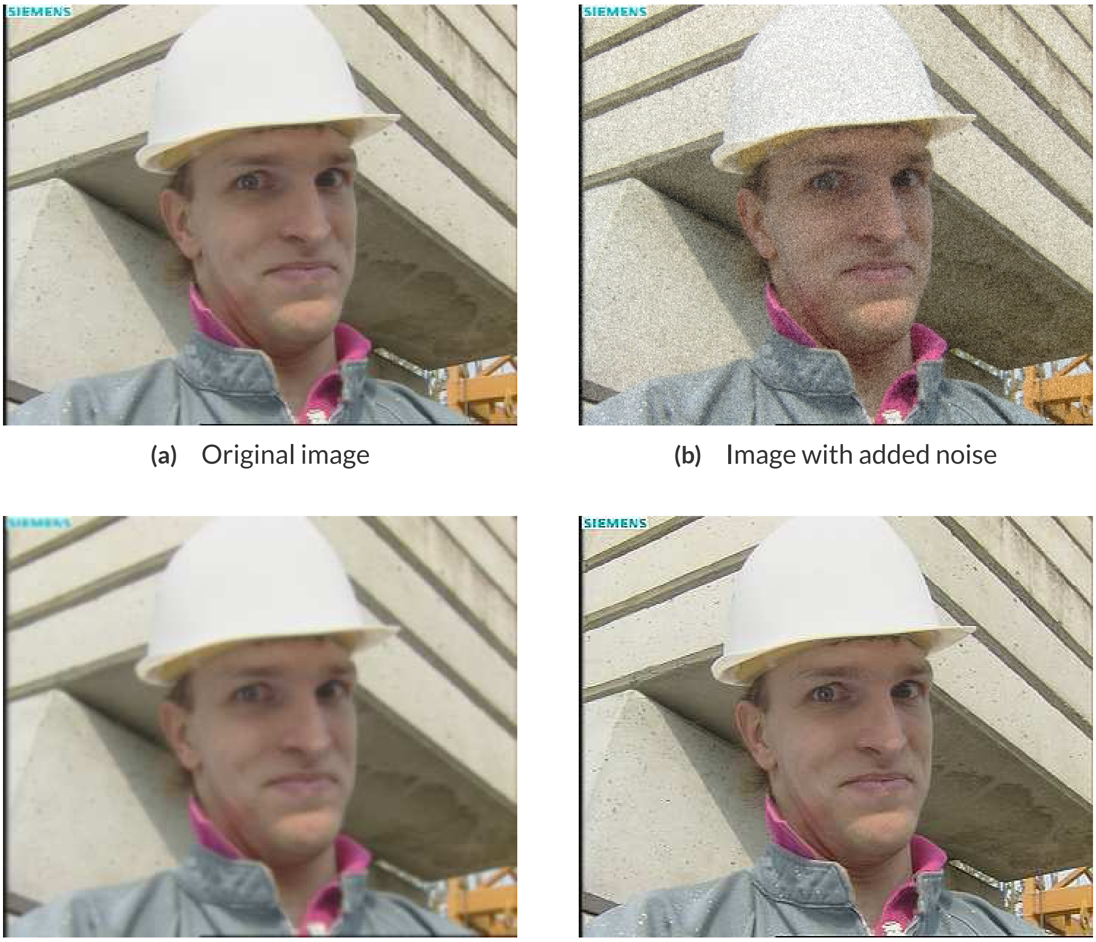
E.1.2. Examples

Figure 40 shows an example SSIM result for an original and processed (compressed with lossy compression) image. The value of 0.9 demonstrates that the two images are very similar.

(a) Original (b) Compressed (c) SSIM

Figure 40: SSIM example for compressed image

Figure 41 depicts various distortions applied to the original image, and Figure 42 shows SSIM values for these distortions.



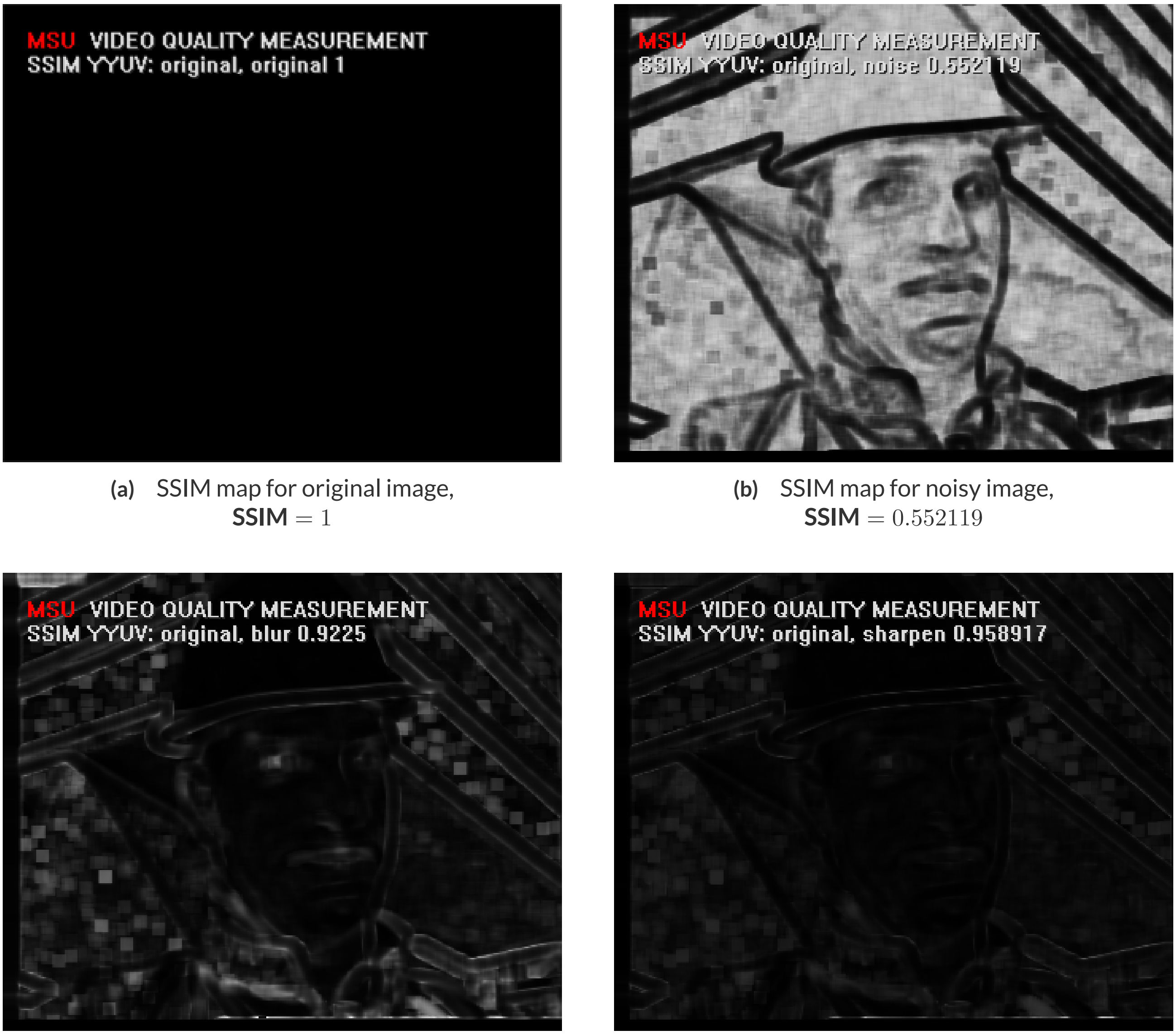
(a) Original image

(b) Image with added noise

(c) Blurred image

(d) Sharpen image

Figure 41: Examples of processed images



(a) SSIM map for original image, SSIM $= 1$

(b) SSIM map for noisy image, SSIM $= 0.552119$

(c) SSIM map for blurred image, SSIM $= 0.9225$

(d) SSIM map for sharpen image, SSIM $= 0.958917$

Figure 42: SSIM values for original and processed images

E.1.3. Measurement method

We used the MSU Video Quality Measurement Tool (VQMT) to calculate objective metrics for the encoded streams. The tool can be downloaded or purchased at `http://compression.ru/video/quality_measure/vqmt_download.html#start`.

Run the command

```
vqmt -in "{original_yuv}" IYUV {width}×{height} -in "decoded_yuv" IYUV
{width}×{height} metrics_list -subsampling -json -json_file "{json_filename}" -threads
3
```

where `input_yuv` is the encoded stream name, `width` and `height` are the size of encoded stream in pixels, `metrics_list` is a list of metrics to measure (e.g., "-metr ssim_precise YYUV -metr ssim_precise UYUV -metr ssim_precise VYUV"), and `json_filename` is the name of the output file containing the metric results.

F. ABOUT THE GRAPHICS & MEDIA LAB VIDEO GROUP

The Graphics & Media Lab Video Group is part of the Computer Science Department of Lomonosov Moscow State University. The Graphics Group began at the end of 1980's, and the Graphics & Media Lab was officially founded in 1998. The main research avenues of the lab include areas of computer graphics, computer vision and media processing (audio, image and video). A number of patents have been acquired based on the lab's research, and other results have been presented in various publications.

The main research avenues of the Graphics & Media Lab Video Group are video processing (pre- and post-, as well as video analysis filters) and video compression (codec testing and tuning, quality metric research and codec development).

The main achievements of the Video Group in the area of video processing include:

- High-quality industrial filters for format conversion, including high-quality deinterlacing, high-quality frame rate conversion, new, fast practical super resolution and other processing tools.
- Methods for modern television sets, such as a large family of up-sampling methods, smart brightness and contrast control, smart sharpening and more.
- Artifact removal methods, including a family of denoising methods, flicking removal, video stabilization with frame edge restoration, and scratch, spot and drop-out removal.
- Application-specific methods such as subtitle removal, construction of panorama images from video, video to high-quality photo conversion, video watermarking, video segmentation and practical fast video deblur.

The main achievements of the Video Group in the area of video compression include:

- Well-known public comparisons of JPEG, JPEG-2000 and MPEG-2 decoders, as well as MPEG-4 and annual H.264 codec testing; codec testing for weak and strong points, along with bug reports and codec tuning recommendations.
- Video quality metric research; the MSU Video Quality Measurement Tool and MSU Perceptual Video Quality Tool are publicly available.
- Internal research and contracts for modern video compression and publication of MSU Lossless Video Codec and MSU Screen Capture Video Codec; these codecs have one of the highest available compression ratios.

The Video Group has also worked for many years with companies like Intel, Samsung and RealNetworks.

In addition, the Video Group is continually seeking collaboration with other companies in the areas of video processing and video compression.

E-mail: video@graphics.cs.msu.ru