MPEG-4 AVC/H.264 Video Codecs Comparison

Short version of report

Project head: Dmitriy Vatolin
Measurements, analysis: Dmitriy Kulikov,
Alexander Parshin

Codecs:

- Dicas H.264
- Elecard H.264
- Intel IPP H.264
- MainConcept H.264
- x264
- XviD (MPEG-4 ASP codec)

May 2009

CS MSU Graphics&Media Lab

Video Group

http://www.compression.ru/video/codec_comparison/index_en.html

videocodec-testing@graphics.cs.msu.ru

Contents

1 Acknowledgments

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

- **dicas digital image coding GmbH**
- **Elecard Ltd**
- **Intel Corporation**
- **MainConcept GmbH**
- **x264 Development Team**
- **XviD**

The Video Group would also like to thank these companies for their help and technical support during the tests.

2 Overview

2.1 Sequences

Table 1. Summary of video sequences.

Sequence	Number of frames	Frame rate	Resolution and color space
1. Akiyo	300	30	352x288(YV12)
2. Foreman	300	30	352x288(YV12)
3. Paris	1065	30	352x288(YV12)
4. Stefan	300	30	352x288(YV12)
5. Battle	1599	24	704x288(YV12)
6. Wendys	260	30	720x576(YV12)
7. State Enemy	6500	24	720x304(YV12)
8. Indiana Jones	5000	30	704x288(YV12)
9. Mobile Calendar	504	50	1280x720(YV12)
10. Stockholm	604	50	1280x720(YV12)
11. Troy	300	24	1920x1072(YV12)
12. Pedestrian Area	375	24	1920x1080(YV12)

Brief descriptions of the sequences used in our comparison are given in Table 1. More detailed descriptions of these sequences can be found in Appendix 4. Test Set of Video Sequences.

2.2 Codecs

Table 2. Short codec descriptions

Codec	Developer	Version
1. dicas' mpegable H.264 Command-line encoder	dicas digital image coding GmbH	3.9.4
2. Elecard AVC Video Encoder 8-bit edition	Elecard Ltd	build Feb 17 2009
3. Intel IPP H.264 Encoder	Intel Corp.	part of "IPP 6.1 Gold" media samples
4. MainConcept AVC/H.264 Video Encoder Console Application	MainConcept GmbH	MainConcept Codec SDK 8.1
5. x264	x264 Development Team	x264 core:67 r1123M 3d78062
6. Xvid raw mpeg4 bitstream encoder	Xvid	version for 24.08.2007

Brief descriptions of the codecs used in our comparison are given in Table 2. XviD was used as a good quality MPEG-4 ASP reference codec for comparison purposes. Detailed descriptions of all codecs used in our comparison can be found in Appendix 5. Tested Codecs.

3 Objectives and Testing Rules

3.1 H.264 Codec Testing Objectives

The main goal of this report is the presentation of a comparative evaluation of the quality of new H.264 codecs using objective measures of assessment. The comparison was done using settings provided by the developers of each codec.

3.2 Testing Rules

- The entire test set was divided into three primary types of applications. These applications differ by resolution, bitrate and encoding speed requirements:
    - Videoconferences (bitrates of 50-400 kbps)
    - Movies (bitrates of 500-1500 kbps)
    - High-definition television (“HDTV”; bitrates of 1-10 mbps)
- There are special presets and speed limitations for every type of application:
    - Videoconferences (speed requirements for 200 kbps CIF sequences):
        - Minimum 100 fps for "High Speed" preset
        - Minimum 50 fps for "High Quality" preset
    - Movies (speed requirements for 750 kbps 4CIF sequences):
        - Minimum 40 fps for "High Speed" preset
        - Minimum 12 fps for "High Quality" preset
    - HDTV (speed requirements for 3 mbps 1280x720 sequences):
        - Minimum 10 fps for "High Speed" preset
        - Minimum 3 fps for "High Quality" preset
- The developer of each codec provided settings for each type of Application. The individual parameters of each setting to a large extent were chosen by the developers.
- Each codec was tested for speed three times; the median score (the middle value of the three measurements) was then used as the representative time.
- During the testing process, source video sequences were in the YV12 format (.yuv file extension)
- For all measurements the PRO version of the MSU Video Quality Measurement Tool was used (http://www.compression.ru/video/quality_measure/vqmt_pro_en.html#start).
- The following computer configuration was used for the main tests, except for multi-core encoding:

OS Name	Microsoft Windows XP Professional
Processor	4-cores processor: Intel Core Quad Q6600
Total Physical Memory	4093.42 MB
Video Adapter Type	NVIDIA GeForce 8500 GT

During the evaluation the following measures were used: PSNR (Y, U, V components) and SSIM (Y, U, V components). More detailed information about these measures may be found on the Internet at the following URL: http://www.compression.ru/video/quality_measure/info.html

4 Comparison Results

4.1 Video conferences

This is the short version of the report. Only part of relative quality analysis is present below. All the other Video Conferences results (RD curves, bitrate handling, encoding speed, etc) can be found in the full version. You can purchase the full version of the report at the comparison web-page. Y-PSNR results could be found in full version also.

4.1.1 Relative Quality Analysis

Table 3 and Table 4 contain relative bitrate for the same quality for all the encoders.

The MainConcept codec is the leader for all presets according to all objective quality metrics, and it is followed by the x264 codec. The Intel IPP H.264 encoder holds third place. For High Speed preset the results of Elecard encoder are very close to Intel IPP H.264 encoder results. All H.264 encoders show better bitrate ratio comparing to XviD encoder using Y-SSIM as quality metric rather than Y-PSNR.

Note, that each the number in tables below corresponds to some segment of bitrates (see Appendix 6. Figures Explanation for more details). Unfortunately, those segments can be rather different because of different quality of compared encoders. This fact can lead to some inadequate results in case of three and more codecs comparisons. This comparison technique will be improved in the future.

Table 3. Average bitrate ratio for a fixed output quality using videoconference sequences and the High Speed preset (Y-SSIM metric).

	dicas	Elecard	IPP H.264	MainConcept	x264	XviD
dicas	100.00%	95.24%	97.37%	65.06%	70.38%	109.01%
Elecard	105.00%	100.00%	101.21%	66.94%	73.94%	115.43%
IPP H.264	102.70%	98.80%	100.00%	67.56%	72.04%	112.80%
MainConcept	153.71%	149.40%	148.01%	100.00%	104.42%	170.76%
x264	142.09%	135.24%	138.80%	95.77%	100.00%	153.54%
XviD	91.73%	86.63%	88.65%	58.56%	65.13%	100.00%

Table 4. Average bitrate ratio for a fixed output quality using videoconference sequences and the High Quality preset (Y-SSIM metric).

	dicas	IPP H.264	MainConcept	x264	XviD
dicas	100.00%	82.13%	64.80%	69.06%	105.58%
IPP H.264	121.76%	100.00%	83.07%	83.91%	130.94%
MainConcept	154.33%	120.38%	100.00%	102.75%	164.22%
x264	144.80%	119.17%	97.32%	100.00%	151.57%
XviD	94.71%	76.37%	60.89%	65.98%	100.00%

Figure 1 and Figure 2 visualize data in the tables above. Each line in those figures corresponds to one codec. Values in vertical axis are average relative bitrate comparing to the codecs in horizontal axis. The lower bitrate is the better relative results have the codec.

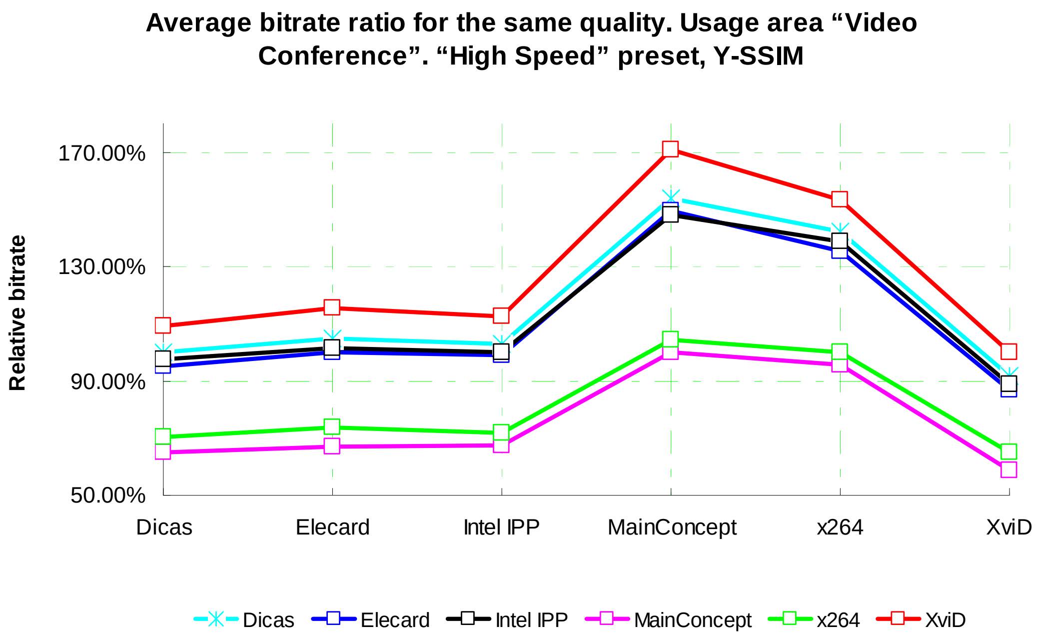


Figure 1. Average bitrate ratio for the same quality. Usage area "Video Conferences". "High Speed" preset, Y-SSIM.

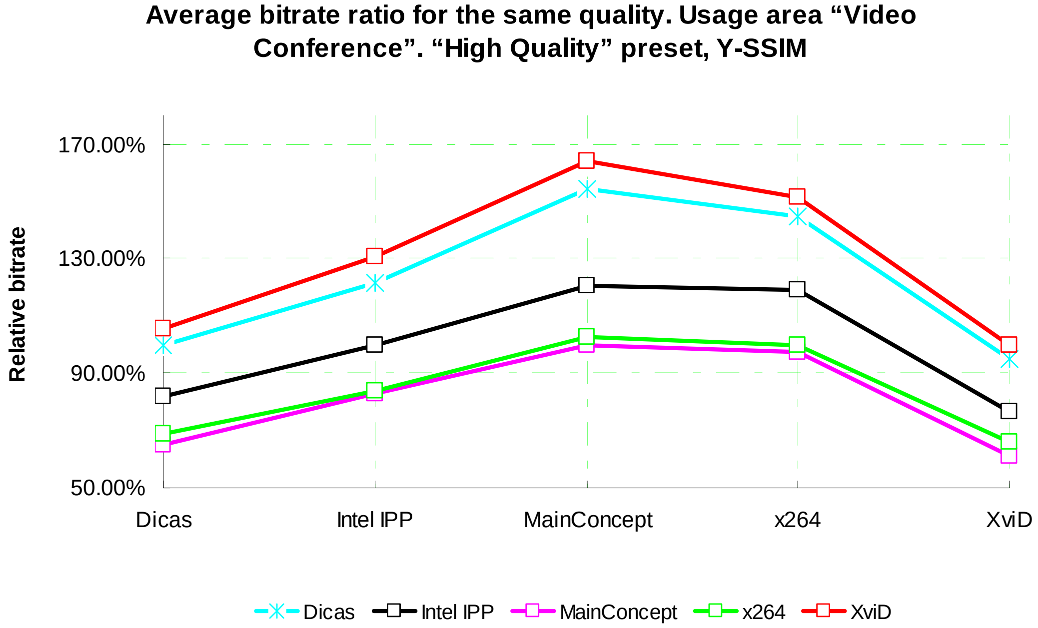


Figure 2. Average bitrate ratio for the same quality. Usage area "Video Conferences". "High Quality" preset, Y-SSIM.

4.2 Movies

This is the short version of the report. Only few examples of figures and part of conclusions are present below. You can purchase the full version of the report at the comparison web-page.

4.2.1 RD Curves

4.2.1.1 High Quality Preset

The High Quality preset results for each sequence are presented in Figure 3 through Figure 8. The first four pictures show the Y-PSNR results and the last four pictures show the Y-SSIM results. Used metric significantly influence on results: leader of PSNR metric is MainConcept, leader of SSIM is x264. Using PSNR as standard of comparison dicas encoder (with HVSAQM enabled) demonstrates the lowest quality among H.264 encoders. Intel IPP H.264 encoder is better than dicas and worth than x264 and MainConcept.

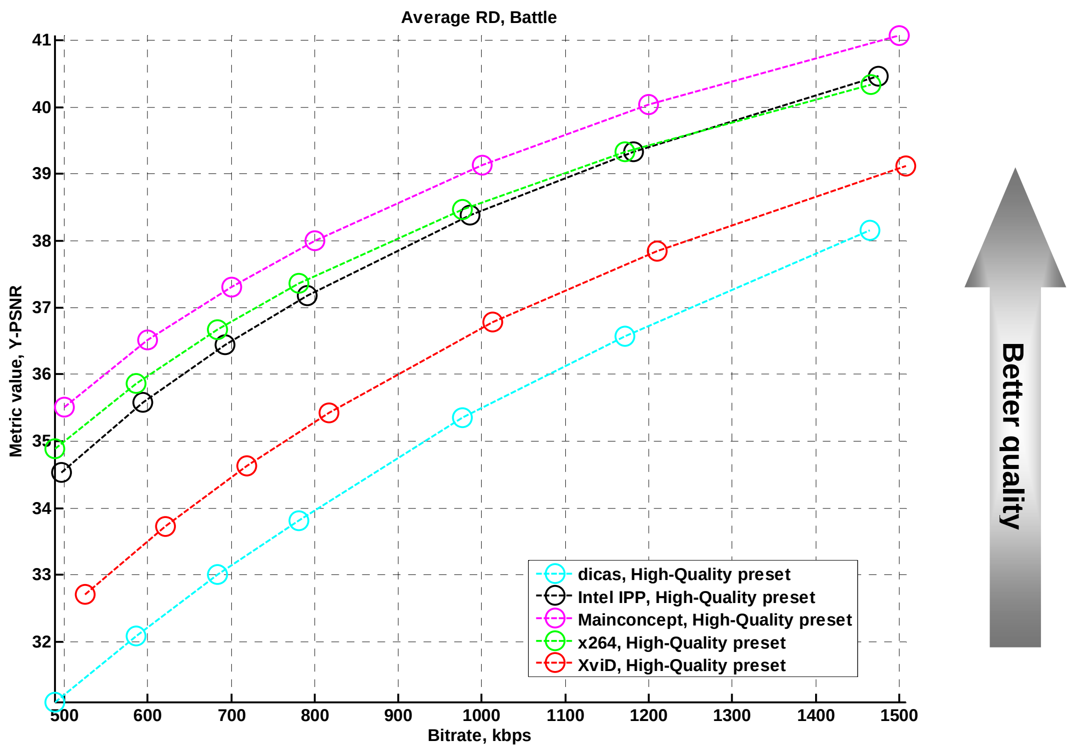


Figure 3. Bitrate/Quality. Usage area “Movies”, “Battle” sequence, “High Quality” preset, Y-PSNR

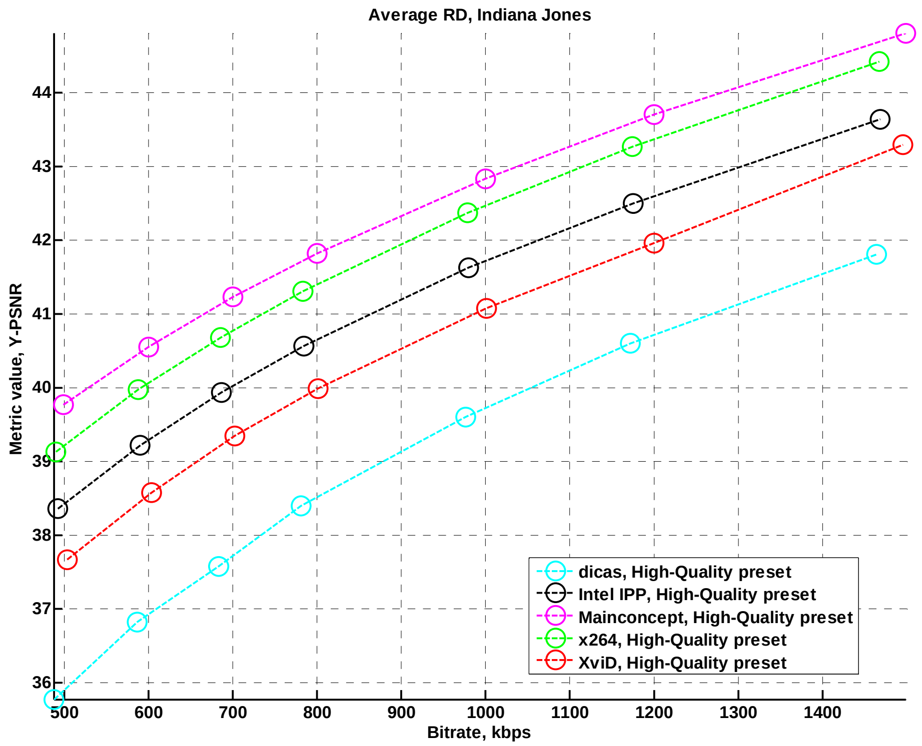


Figure 4. Bitrate/Quality. Usage area “Movies”, “Indiana Jones” sequence, “High Quality” preset, Y-PSNR

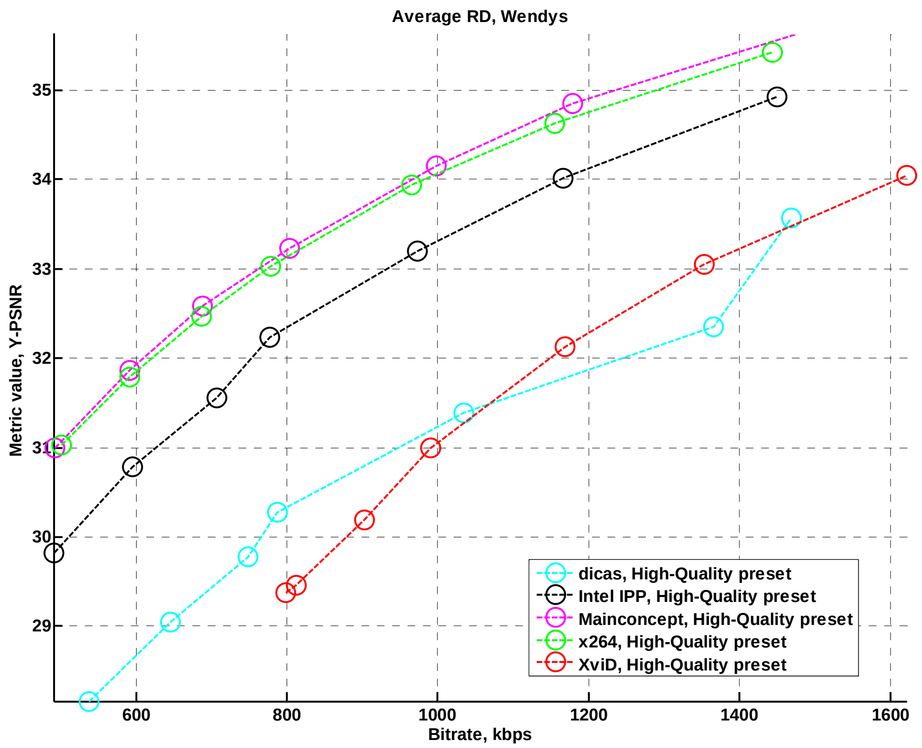


Figure 5. Bitrate/Quality. Usage area “Movies”, “Wendys” sequence, “High Quality” preset, Y-PSNR

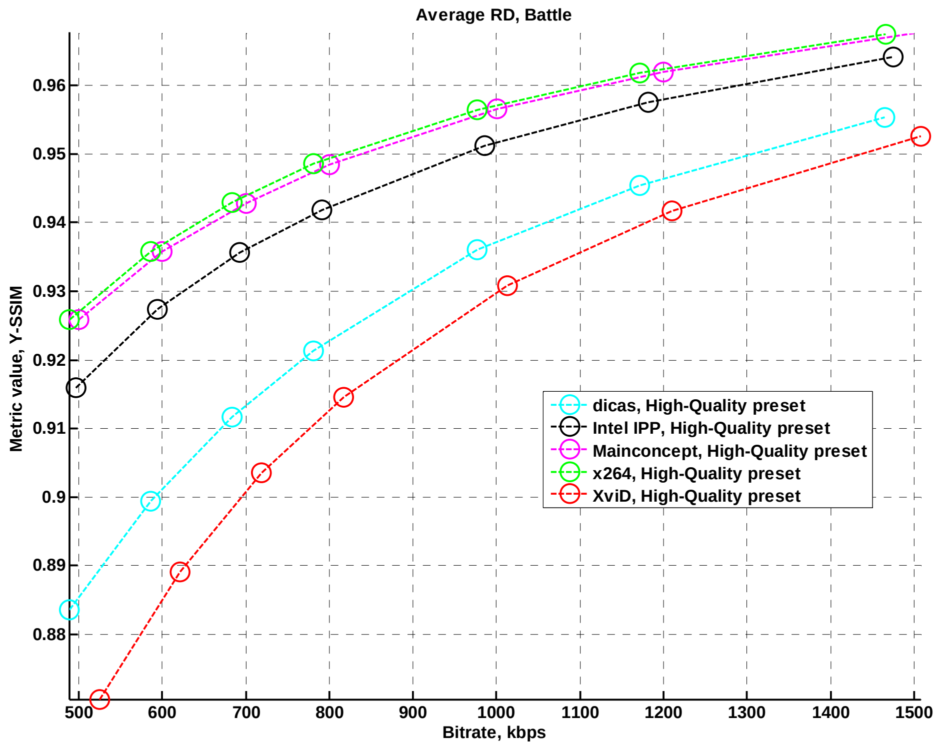


Figure 6. Bitrate/Quality. Usage area "Movies", "Battle" sequence, "High Quality" preset, Y-SSIM

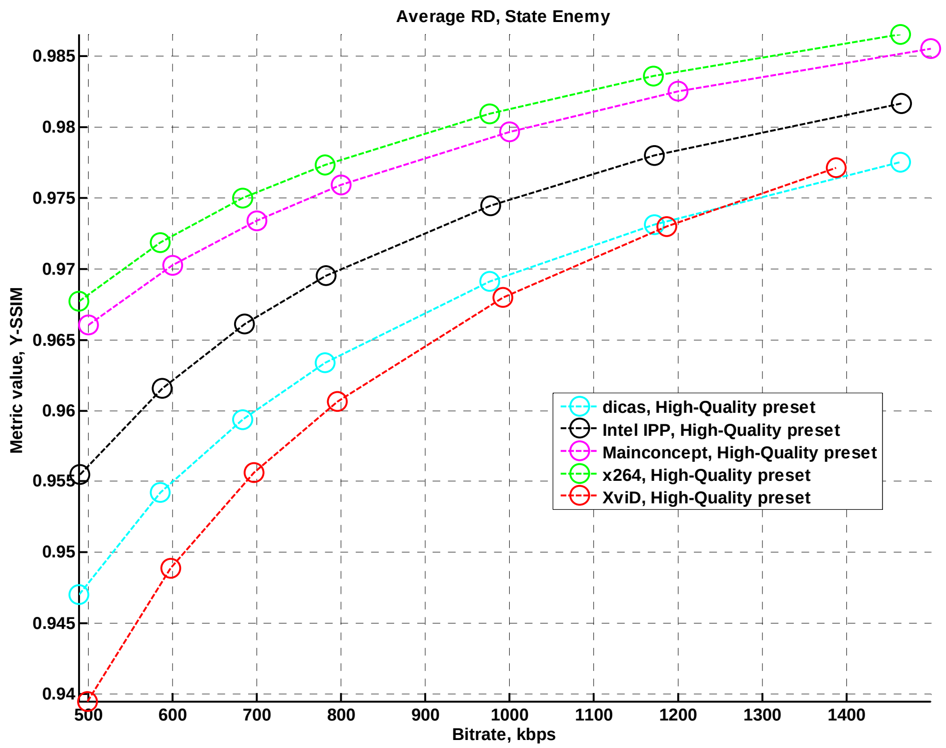


Figure 7. Bitrate/Quality. Usage area "Movies", "State Enemy" sequence, "High Quality" preset, Y-SSIM

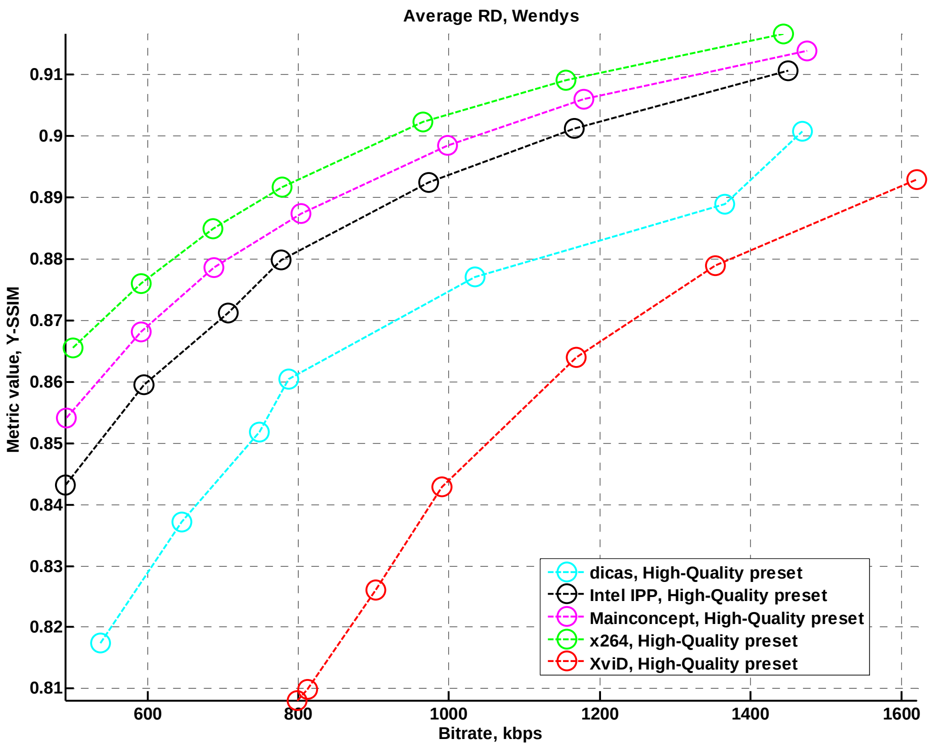


Figure 8. Bitrate/Quality. Usage area "Movies", "Wendys" sequence, "High Quality" preset, Y-SSIM

4.2.1.2 High Speed Preset

The RD curves for the High Speed preset are shown in Figure 9 through Figure 12. The situation is similar to High Quality preset: leader is MainConcept when PSNR is used and x264 when SSIM is used. Third place have Intel IPP H.264 codecs, followed by Elecard and dicas encoders.

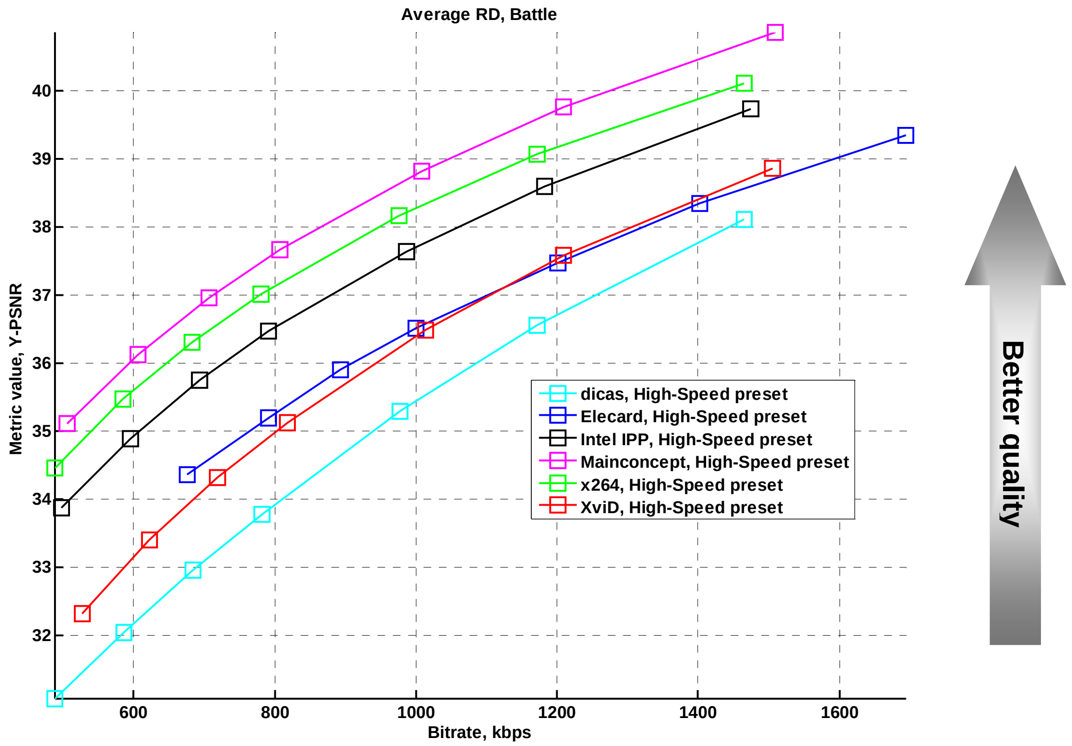


Figure 9. Bitrate/Quality. Usage area "Movies", "Battle" sequence, "High Speed" preset, Y-PSNR

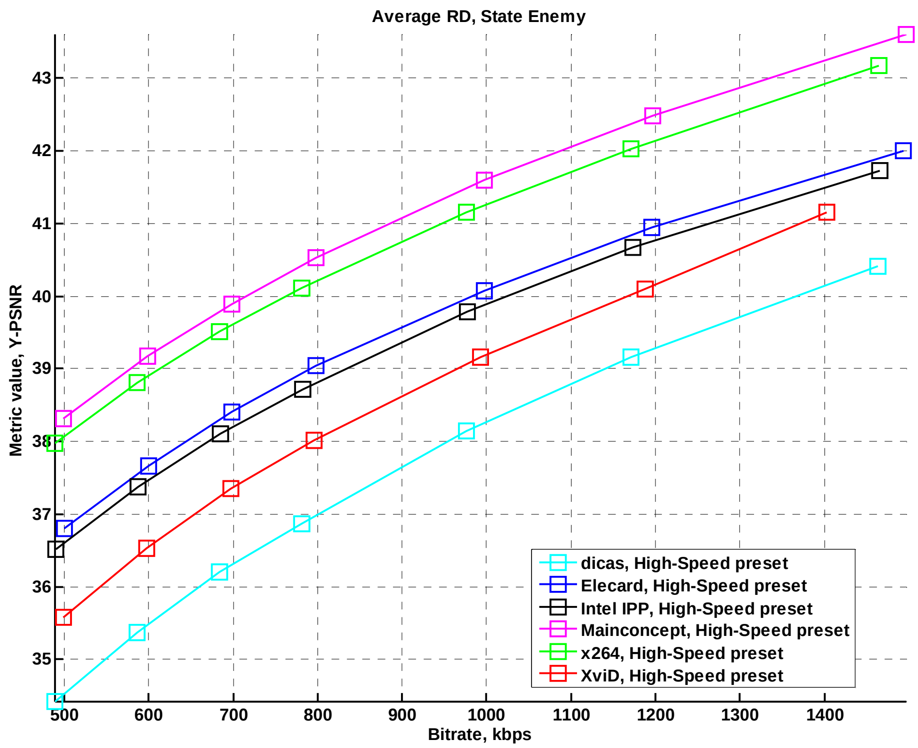


Figure 10. Bitrate/Quality. Usage area "Movies", "State Enemy" sequence, "High Speed" preset, Y-PSNR

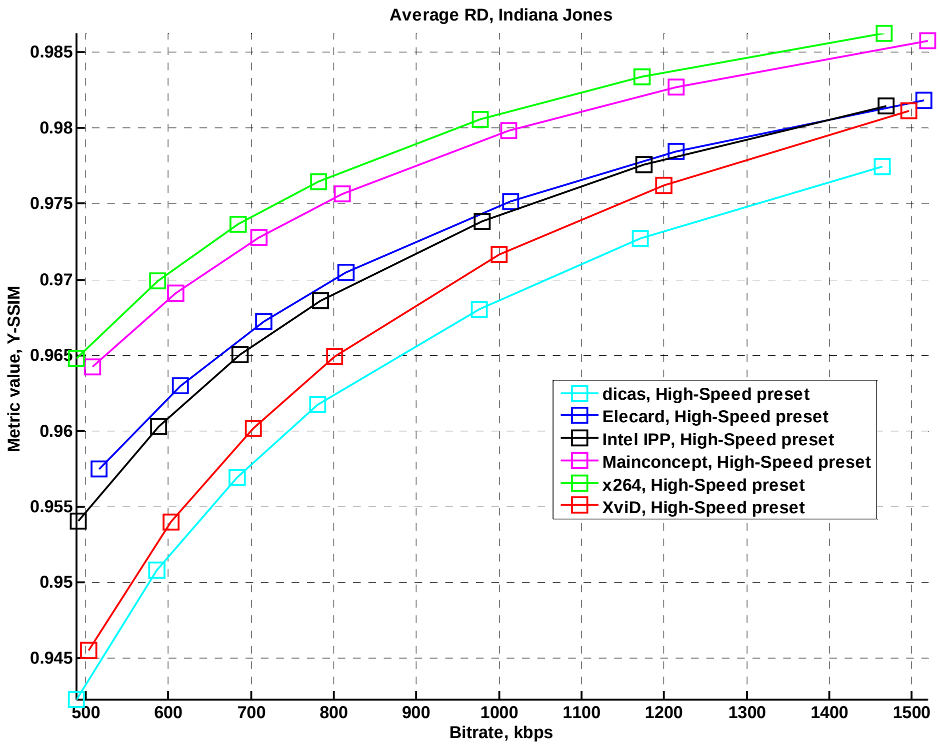


Figure 11. Bitrate/Quality. Usage area "Movies", "Indiana Jones" sequence, "High Speed" preset, Y-SSIM

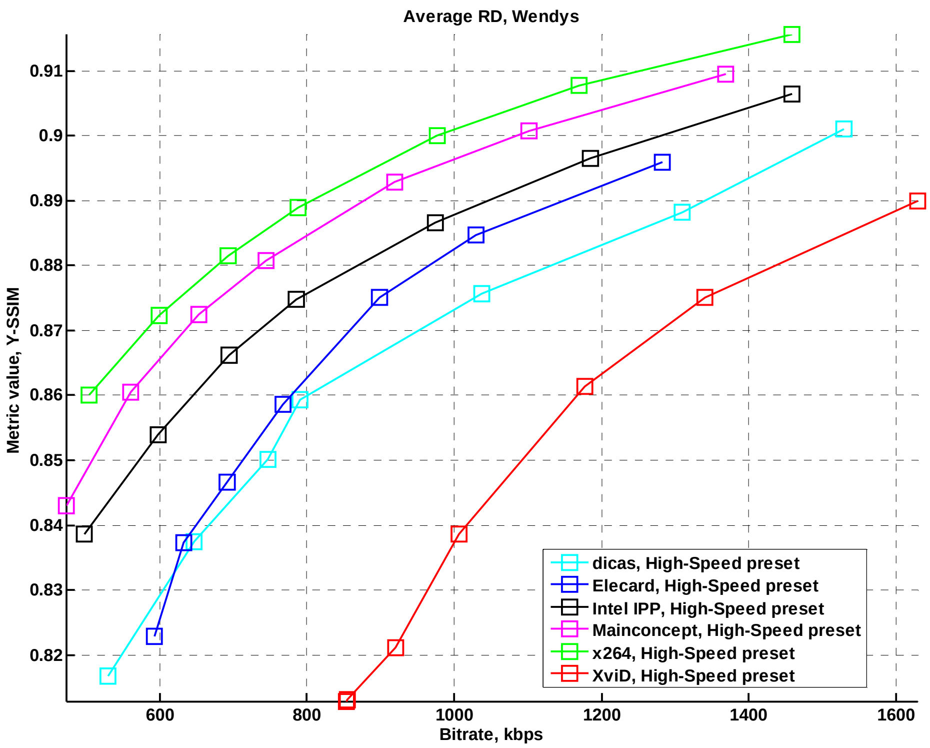


Figure 12. Bitrate/Quality. Usage area "Movies", "Wendys" sequence, "High Speed" preset, Y-SSIM

4.2.2 Encoding Speed

4.2.2.1 High Speed Preset

Absolute speed results are presented in Figure 13 and Figure 14.

The fastest encoder with High Speed preset is Elecard, the slowest one is x264(best fitted our speed requirements). Note unstable results of Elecard encoder at “State Enemy” sequence.

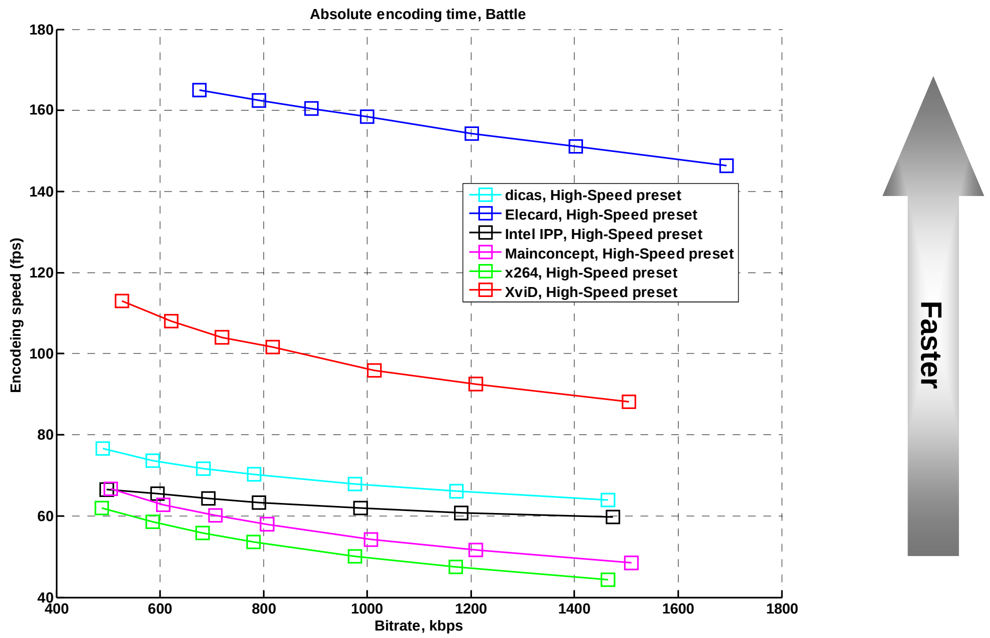


Figure 13. Encoding speed. Usage area “Movies”, “Battle” sequence, “High Speed” preset

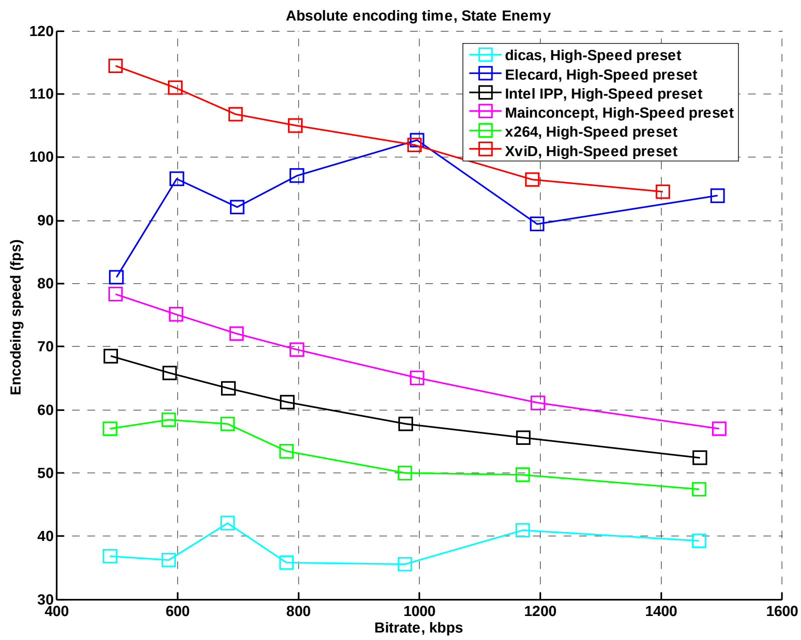


Figure 14. Encoding speed. Usage area “Movies”, “State Enemy” sequence, “High Speed” preset

4.2.3 Speed/Quality Tradeoff

Detailed descriptions of the speed/quality trade-off graphs can be found in Appendix 6. Figures Explanation. Sometimes codec results are not present in the particular graph. The reason for that are extremely poor results of the codec. Its RD curve has no intersection with reference's RD curve.

The speed/quality trade-off graphs simultaneously show relative quality and encoding speed for the encoders tested in this comparison. Again, XviD is the reference codec with both quality and speed normalized to unity for all of the below graphs. The terms "better" and "worse" are used to compare codecs in the same manner as in previous portions of this comparison.

Please note that the averaging method among all sequences suppose that all codecs have the results for each sequence. When it's not the case, then only existing results are taking into account.

4.2.3.1 High Speed Preset

Figure 15 through Figure 20 show results for the High Speed preset. In considering the cumulative results for all sequences, it becomes apparent that the Elecard encoder is better than the XviD codec. The x264 and MainConcept encoders results depend on used metric similar to High Quality results. Another compatible pair is dicas and Intel IPP H.264 encoder, where dicas is totally worse than Intel.

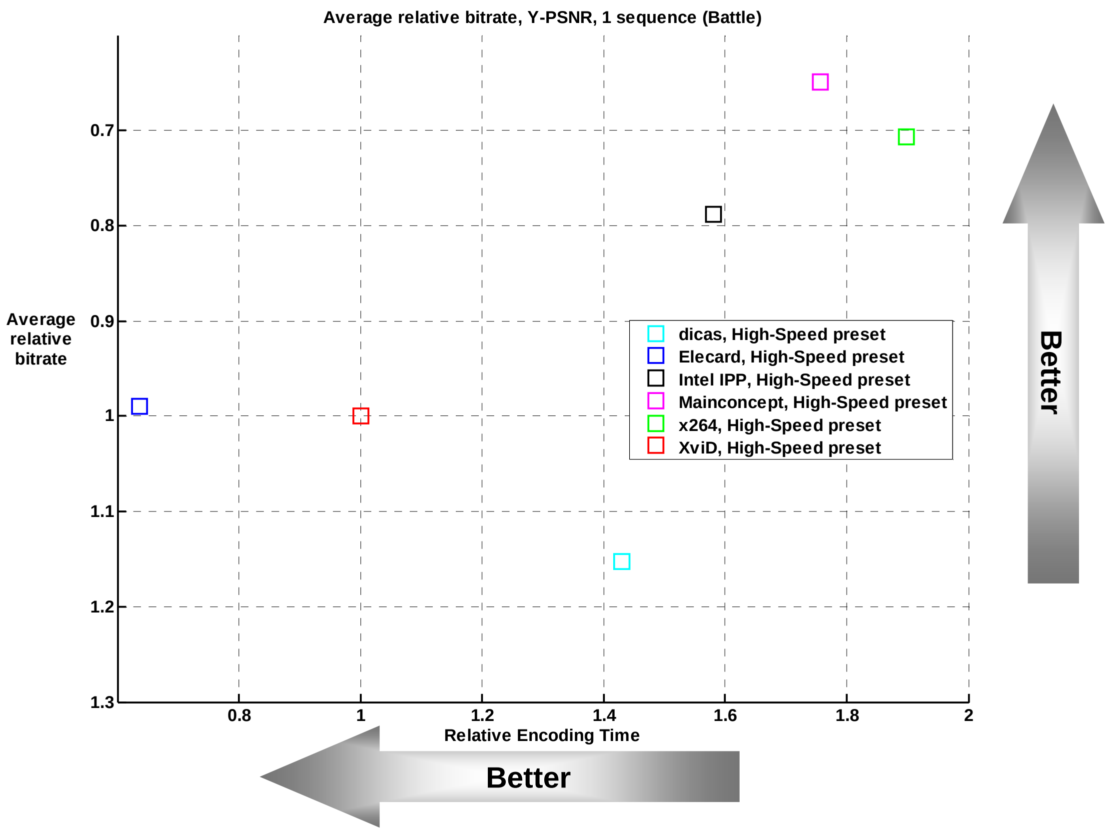


Figure 15. Speed/Quality tradeoff. Usage area "Movies", "Battle" sequence, "High Speed" preset, Y-PSNR

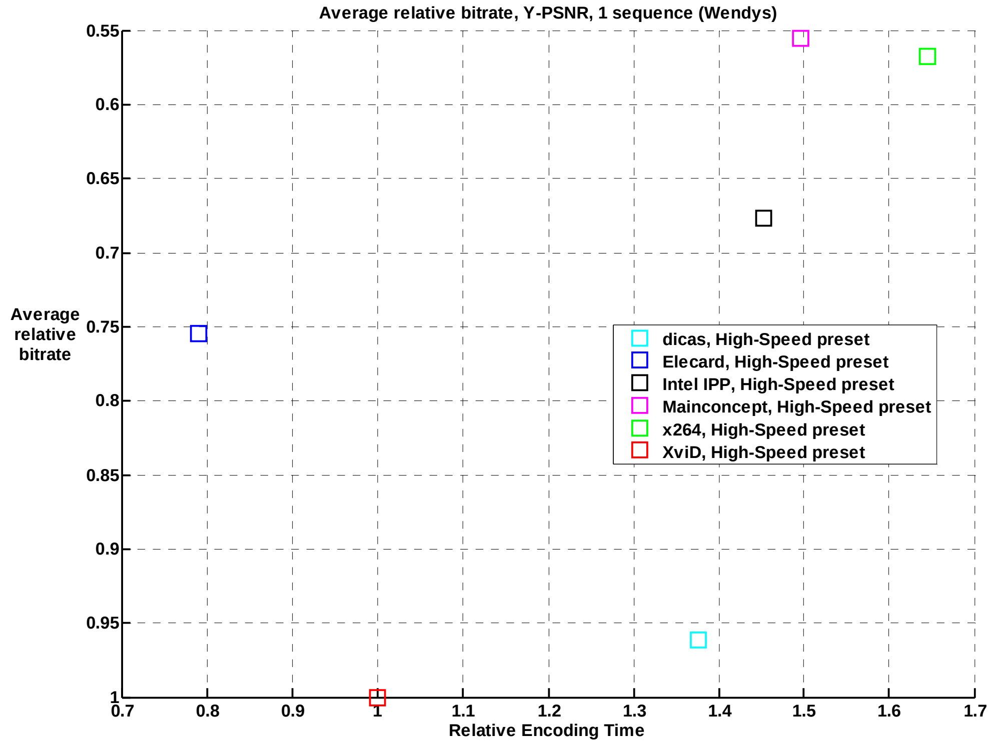


Figure 16. Speed/Quality tradeoff. Usage area "Movies", "Wendys" sequence, "High Speed" preset, Y-PSNR

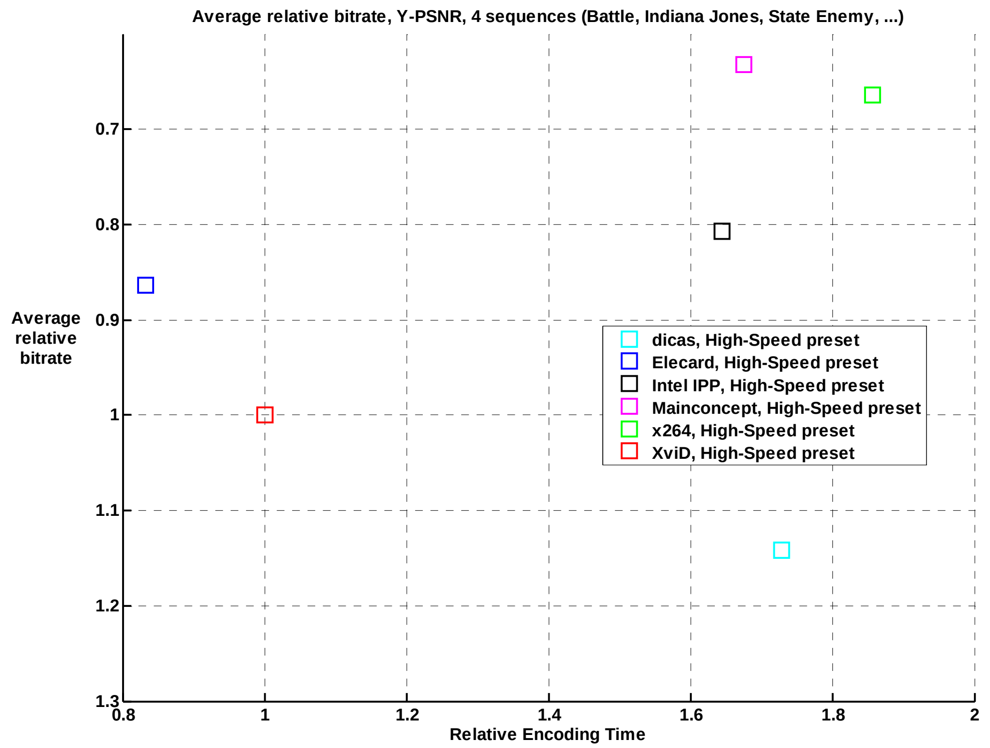


Figure 17. Speed/Quality tradeoff. Usage area "Movies", all the sequences, "High Speed" preset, Y-PSNR

SSIM

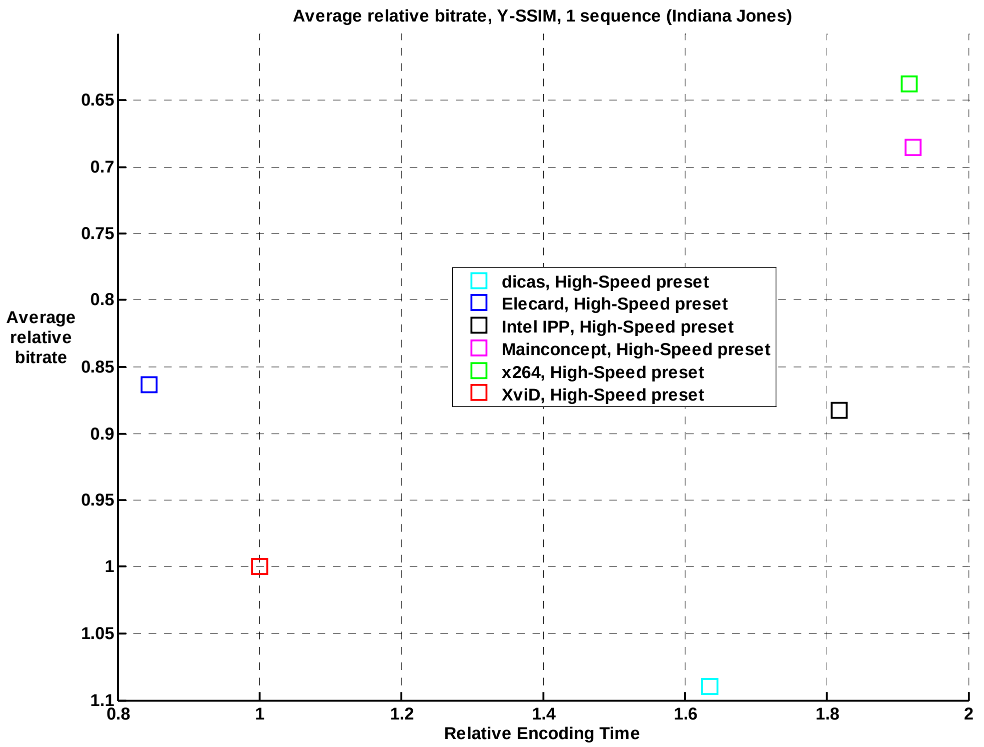


Figure 18. Speed/Quality tradeoff. Usage area "Movies", "Indiana Jones" sequence, "High Speed" preset, Y-SSIM

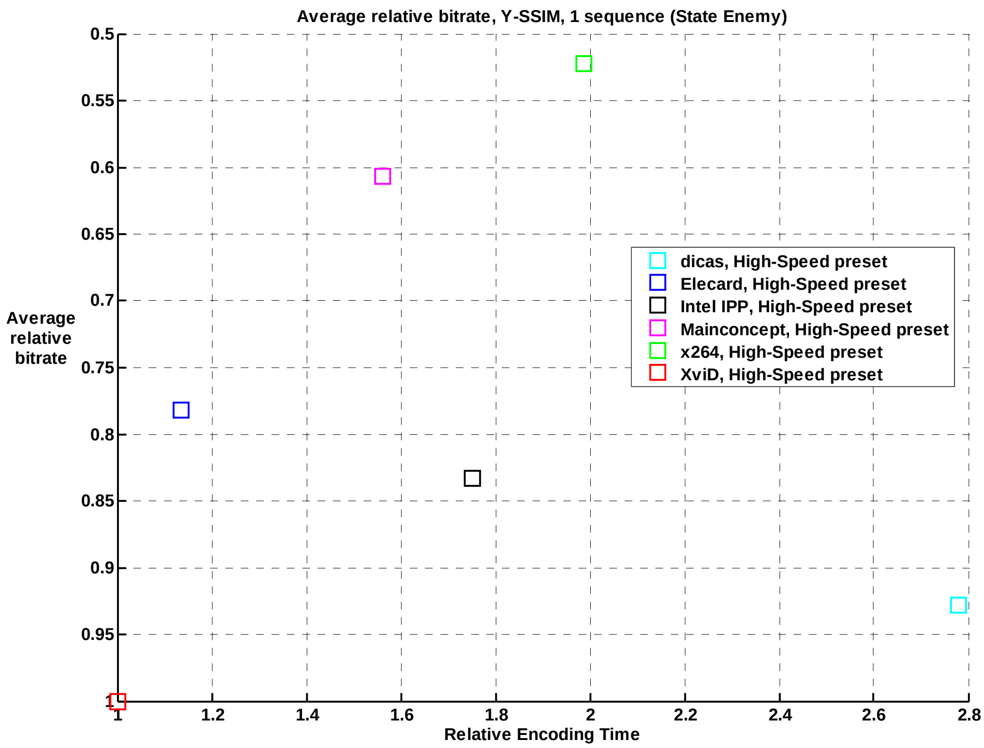


Figure 19. Speed/Quality tradeoff. Usage area "Movies", "State Enemy" sequence, "High Speed" preset, Y-SSIM

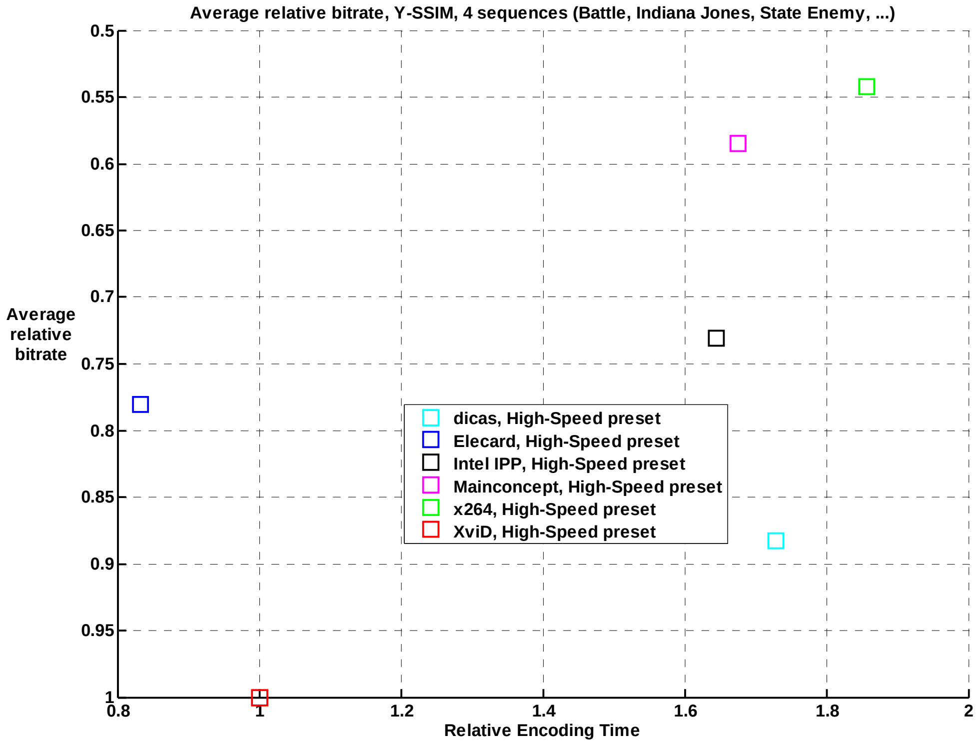


Figure 20. Speed/Quality tradeoff. Usage area "Movies", all sequences, "High Speed" preset, Y-SSIM

4.2.4 Bitrate Handling

With High Speed preset Elecard encoder demonstrates bitrate increasing at "Battle" and "Wendys" sequences.

4.2.4.1 High Speed Preset

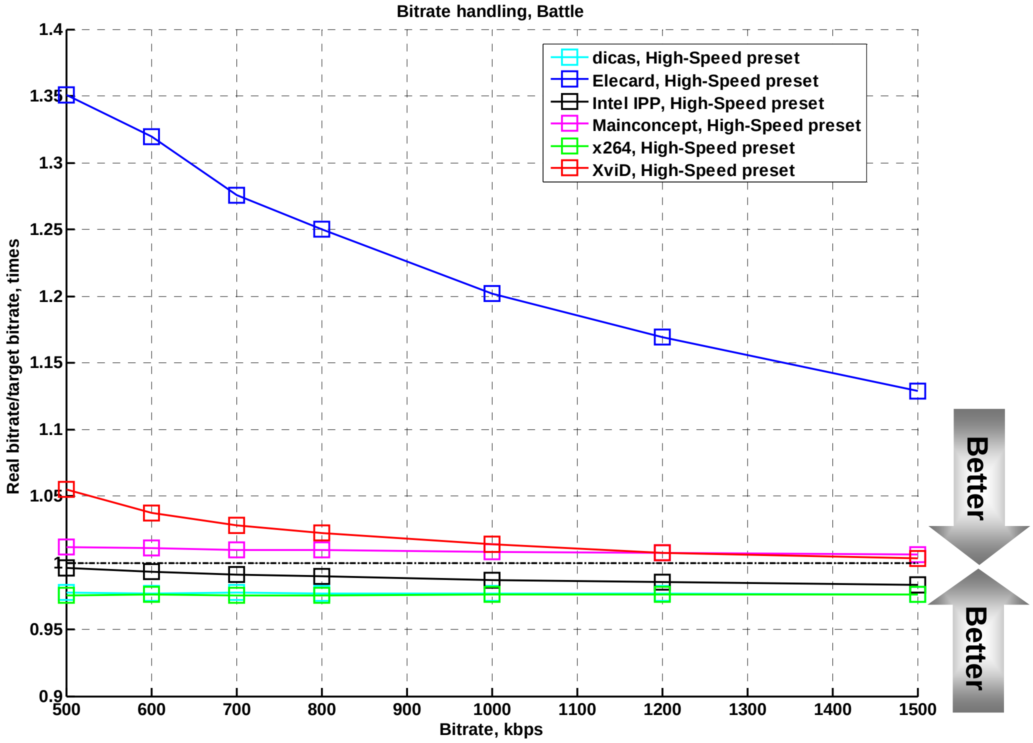


Figure 21. Bitrate Handling. Usage area "Movies", "Battle" sequence, "High Speed" preset

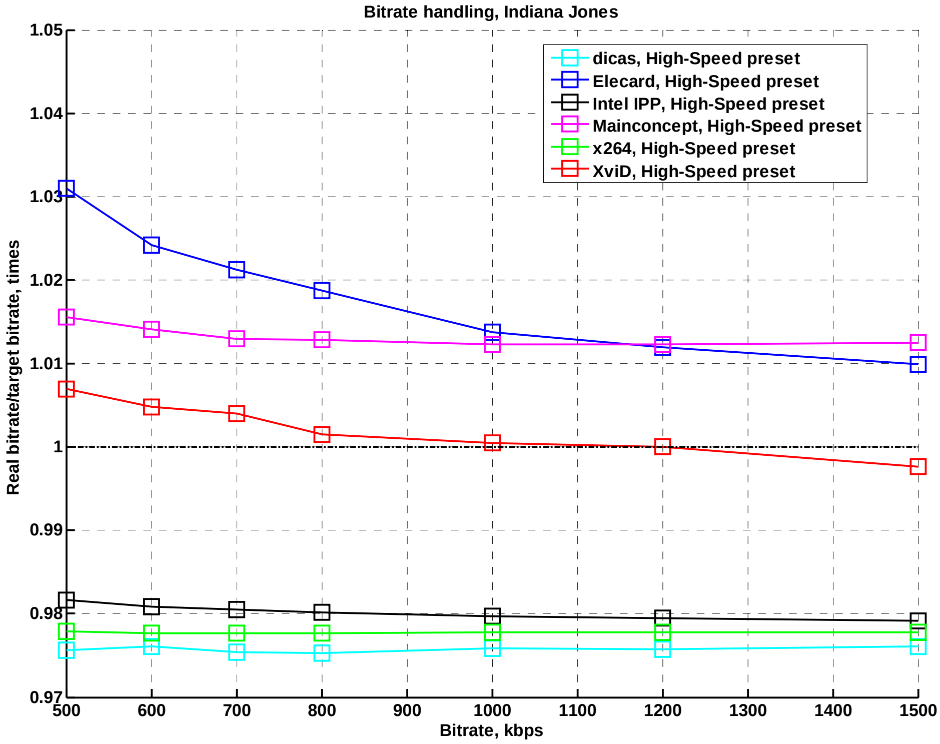


Figure 22. Bitrate Handling. Usage area “Movies”, “Indiana Jones” sequence, “High Speed” preset

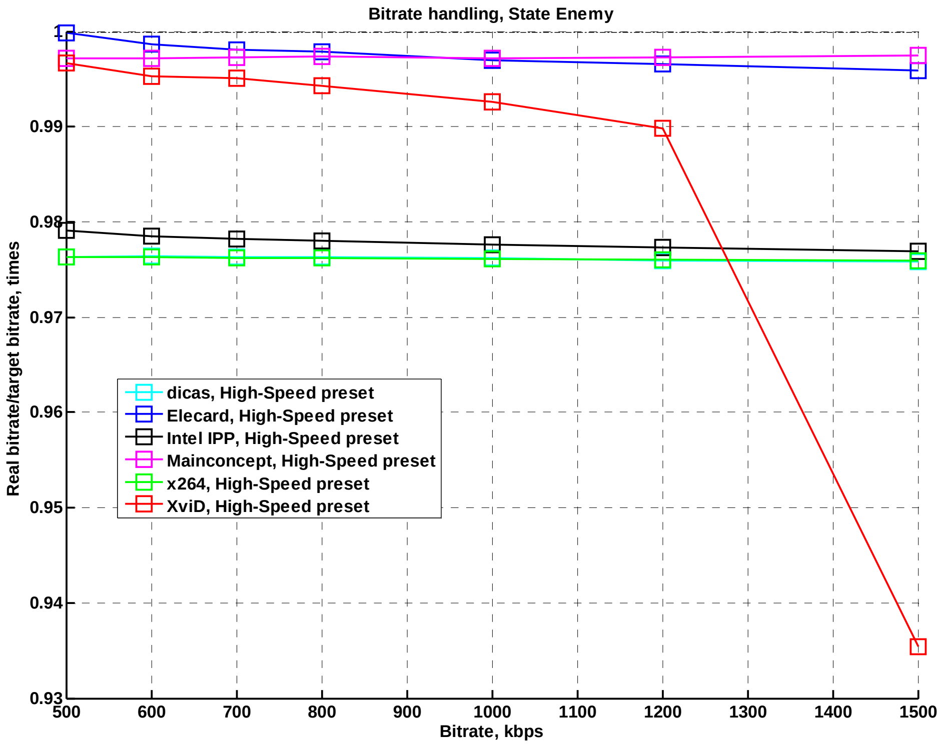


Figure 23. Bitrate Handling. Usage area “Movies”, “State Enemy” sequence, “High Speed” preset

4.2.5 Relative Quality Analysis

Table 5 through Table 8 show relative bitrates for a fixed quality output for all codecs and presets. Note that these tables do not include information about the speed of the encoder.

Note, that each the number in tables below corresponds to some segment of bitrates (see Appendix 6. Figures Explanation for more details). Unfortunately, those segments can be rather different because of different quality of compared encoders. This fact can lead to some inadequate results in case of three and more codecs comparisons. This comparison technique will be improved in the future.

Consider the High Speed preset (Y-PSNR results in Table 5, Y-SSIM results in Table 6). Note that XviD results depend on used quality metric. At average, there are two leaders: x264 and MainConcept encoders.

Table 7 and Table 8 present the High Quality preset results for the Y-PSNR and Y-SSIM quality metrics, respectively. Results are similar to High Speed preset: leaders are x264 and MainConcept depending on used quality metric (difference is 18% of bitrate for the same quality).

Table 5. Average bitrate ratio for the same quality. Usage area "Movie". "High Speed" preset, Y-PSNR.

	Dicas	Elecard	IPP H.264	MainConcept	x264	XviD
Dicas	100%	76%	70%	55%	57%	88%
Elecard	132%	100%	94%	73%	77%	116%
IPP H.264	143%	107%	100%	77%	81%	124%
MainConcept	182%	138%	129%	100%	105%	158%
x264	175%	131%	123%	95%	100%	151%
XviD	114%	86%	81%	63%	66%	100%

Table 6. Average bitrate ratio for the same quality. Usage area "Movie". "High Speed" preset, Y-SSIM.

	Dicas	Elecard	IPP H.264	MainConcept	x264	XviD
Dicas	100%	87%	82%	64%	59%	113%
Elecard	115%	100%	94%	74%	68%	128%
IPP H.264	123%	106%	100%	78%	72%	137%
MainConcept	156%	136%	128%	100%	92%	171%
x264	170%	147%	140%	109%	100%	184%
XviD	88%	78%	73%	59%	54%	100%

Table 7. Average bitrate ratio for the same quality. Usage area "Movie". "High Quality" preset, Y-PSNR.

	Dicas	IPP H.264	MainConcept	x264	XviD
Dicas	100%	64%	51%	55%	85%
IPP H.264	155%	100%	79%	87%	131%
MainConcept	197%	126%	100%	110%	165%
x264	180%	115%	91%	100%	151%
XviD	117%	76%	60%	66%	100%

Table 8. Average bitrate ratio for the same quality. Usage area “Movie”. “High Quality” preset, Y-SSIM.

	Dicas	IPP H.264	MainConcept	x264	XviD
Dicas	100%	74%	62%	57%	110%
IPP H.264	135%	100%	84%	77%	146%
MainConcept	162%	119%	100%	92%	172%
x264	177%	129%	109%	100%	185%
XviD	91%	69%	58%	54%	100%

Figure 24 through Figure 27 visualize data in the tables above. Each line in those figures corresponds to one codec. Values in vertical axis are average relative bitrate comparing to the codecs in horizontal axis. The lower bitrate is the better relative results have the codec.

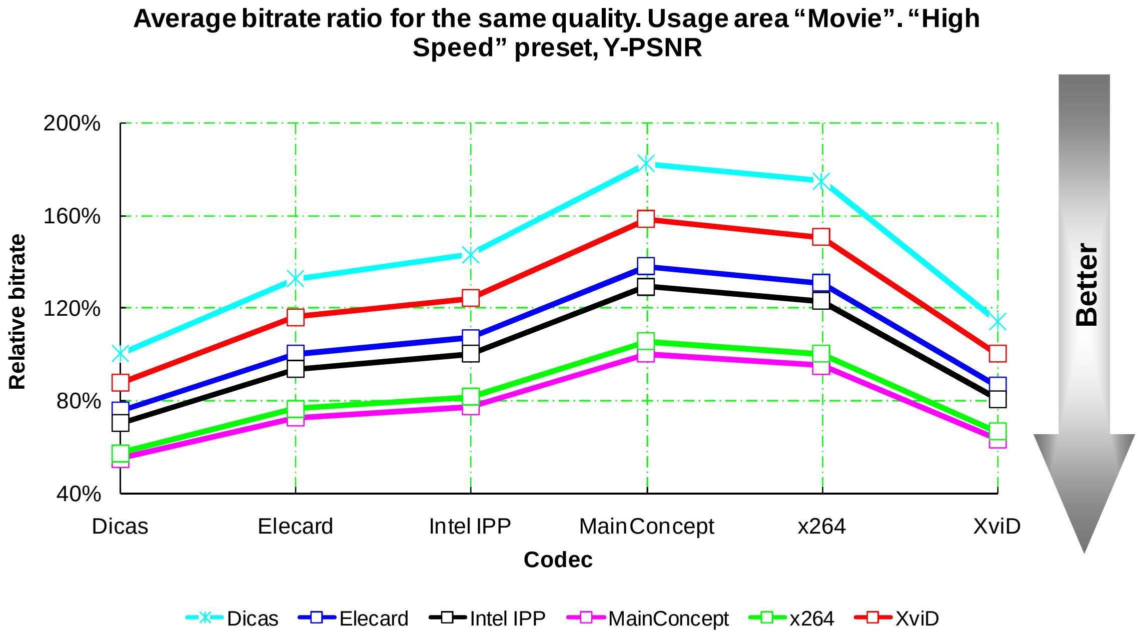


Figure 24. Average bitrate ratio for the same quality. Usage area “Movie”. “High Speed” preset, Y-PSNR.

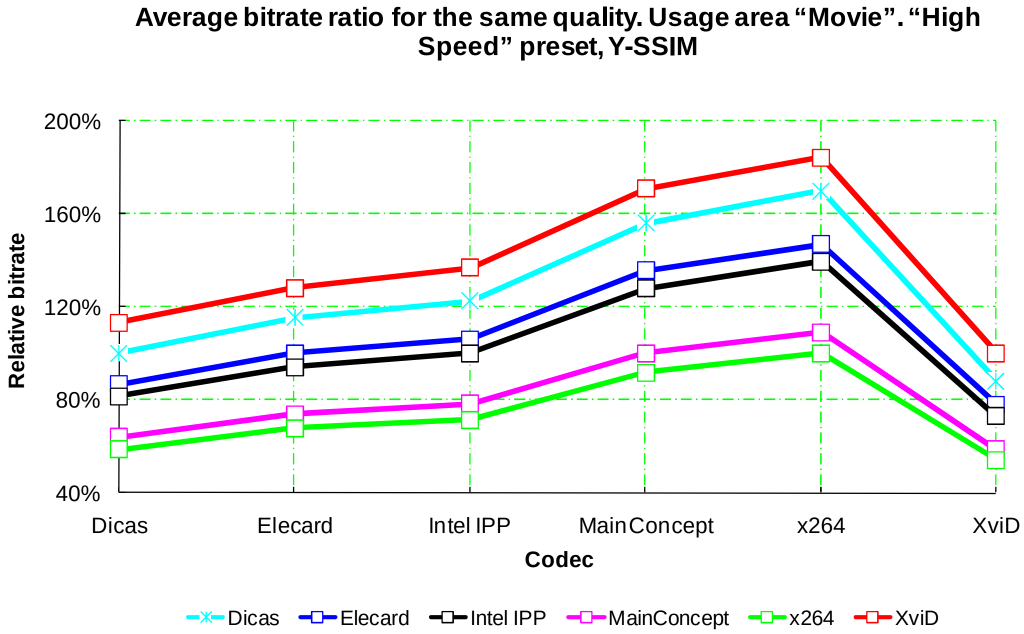


Figure 25. Average bitrate ratio for the same quality. Usage area “Movie”. “High Speed” preset, Y-SSIM.

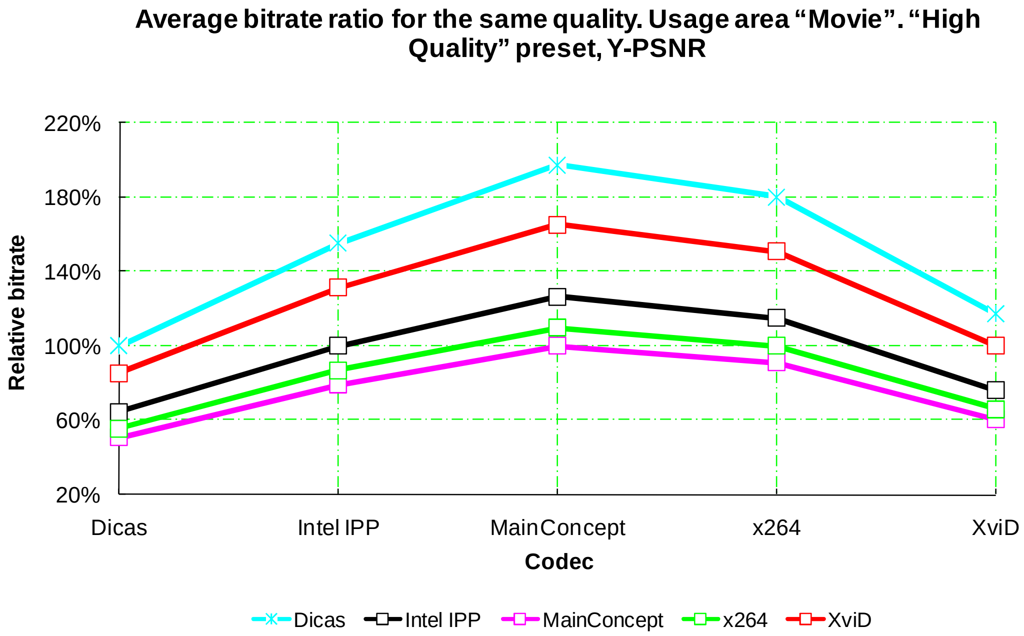


Figure 26. Average bitrate ratio for the same quality. Usage area “Movie”. “High Quality” preset, Y-PSNR.

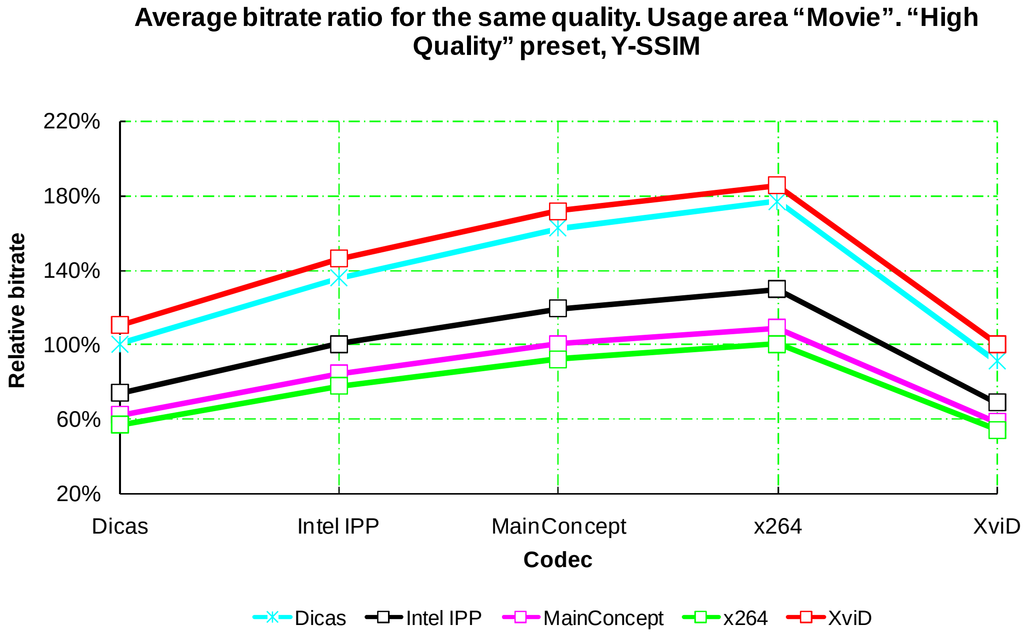


Figure 27. Average bitrate ratio for the same quality. Usage area “Movie”. “High Quality” preset, Y-SSIM.

4.3 HDTV

This is the short version of the report. Only part of relative quality analysis is present below. All the other HDTV results (RD curves, bitrate handling, encoding speed, etc) can be found in the full version. You can purchase the full version of the report at the comparison web-page.

4.3.1 Relative Quality Analysis

Table 9 through Table 12 contain relative bitrate data for a fixed quality output for all the encoders.

Note, that each number in tables below corresponds to some segment of bitrates (see Appendix 6. Figures Explanation for more details). Unfortunately, those segments can be rather different because of different quality of compared encoders. This fact can lead to some inadequate results in case of three and more codecs comparisons. This comparison technique will be improved in the future.

The MainConcept codec is the leader for all presets according to all objective quality metrics, and it is followed by the x264 codec. The Intel IPP H.264 encoder holds third place, except “High Quality” preset, Y-PSNR where it reaches second place and “High Speed” preset, Y-PSNR where Intel IPP H.264 and x264 results are very close to each other. For High Speed preset the results of Elecard encoder are for fourth place. All H.264 encoders show better bitrate ratio comparing to XviD encoder using Y-SSIM as quality metric rather than Y-PSNR. This difference could be easily noted for dicas encoder with HVSAQM enabled – if using Y-PNSR it shows lower results than XviD, but using Y-SSIM it shows better results than XviD.

Table 9. Average bitrate ratio for the same quality. Usage area “HDTV”. “High Speed” preset, Y-PSNR.

	Dicas	Elecard	IPP H.264	MainConcept	x264	XviD
Dicas	100.00%	67.44%	58.84%	45.50%	55.97%	86.07%
Elecard	148.29%	100.00%	86.02%	65.90%	83.93%	128.01%
IPP H.264	169.96%	116.25%	100.00%	77.62%	98.33%	147.20%
MainConcept	219.76%	151.73%	128.83%	100.00%	128.61%	192.91%
x264	178.66%	119.14%	101.70%	77.75%	100.00%	154.24%
XviD	116.19%	78.12%	67.93%	51.84%	64.83%	100.00%

Table 10. Average bitrate ratio for the same quality. Usage area “HDTV”. “High Speed” preset, Y-SSIM.

	Dicas	Elecard	IPP H.264	MainConcept	x264	XviD
Dicas	100.00%	77.98%	67.92%	51.94%	55.74%	103.13%
Elecard	128.24%	100.00%	86.77%	64.90%	70.49%	135.09%
IPP H.264	147.23%	115.24%	100.00%	75.46%	80.95%	153.74%
MainConcept	192.51%	154.07%	132.51%	100.00%	104.68%	205.12%
x264	179.40%	141.86%	123.54%	95.53%	100.00%	184.13%
XviD	96.97%	74.02%	65.04%	48.75%	54.31%	100.00%

Table 11. Average bitrate ratio for the same quality. Usage area "HDTV". "High Quality" preset, Y-PSNR.

	Dicas	IPP H.264	MainConcept	x264	XviD
Dicas	100.00%	53.57%	45.03%	53.20%	86.67%
IPP H.264	186.66%	100.00%	88.52%	104.99%	157.62%
MainConcept	222.06%	112.97%	100.00%	122.21%	188.13%
x264	187.96%	95.25%	81.83%	100.00%	159.64%
XviD	115.39%	63.44%	53.15%	62.64%	100.00%

Table 12. Average bitrate ratio for the same quality. Usage area "HDTV". "High Quality" preset, Y-SSIM.

	Dicas	IPP H.264	MainConcept	x264	XviD
Dicas	100.00%	61.35%	53.03%	52.74%	106.49%
IPP H.264	163.00%	100.00%	88.51%	86.09%	167.71%
MainConcept	188.57%	112.98%	100.00%	96.94%	195.28%
x264	189.60%	116.16%	103.16%	100.00%	193.50%
XviD	93.91%	59.63%	51.21%	51.68%	100.00%

Figure 28 through Figure 31 visualize data in the tables above. Each line in those figures corresponds to one codec. Values in vertical axis are average relative bitrate comparing to the codecs in horizontal axis. The lower bitrate is the better relative results have the codec.

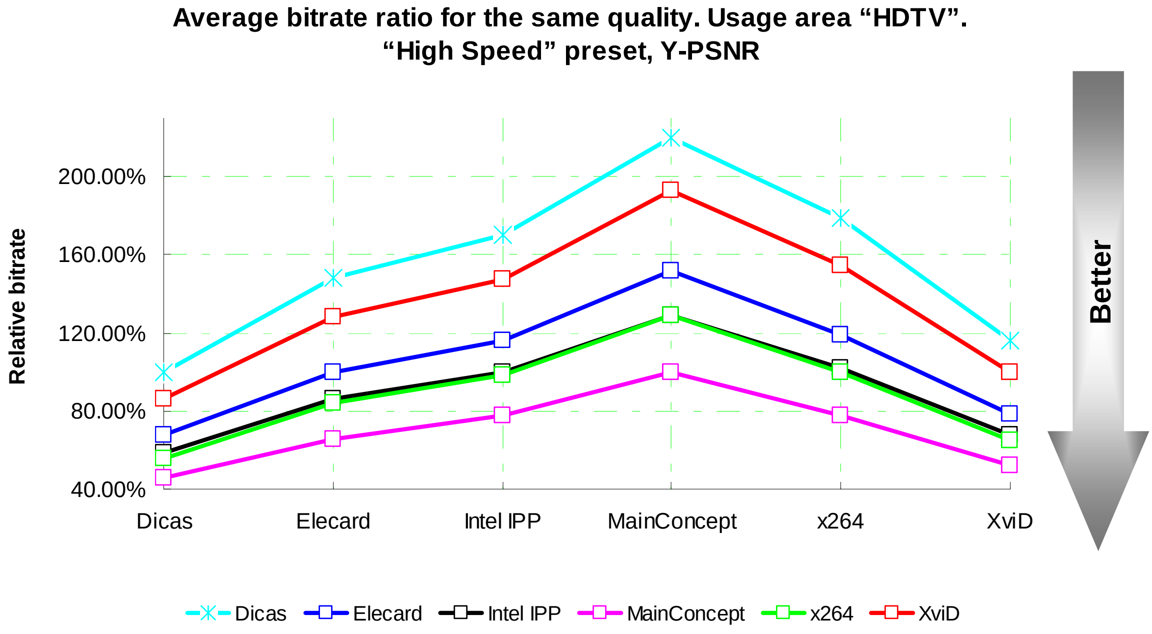


Figure 28. Average bitrate ratio for the same quality. Usage area "HDTV". "High Speed" preset, Y-PSNR.

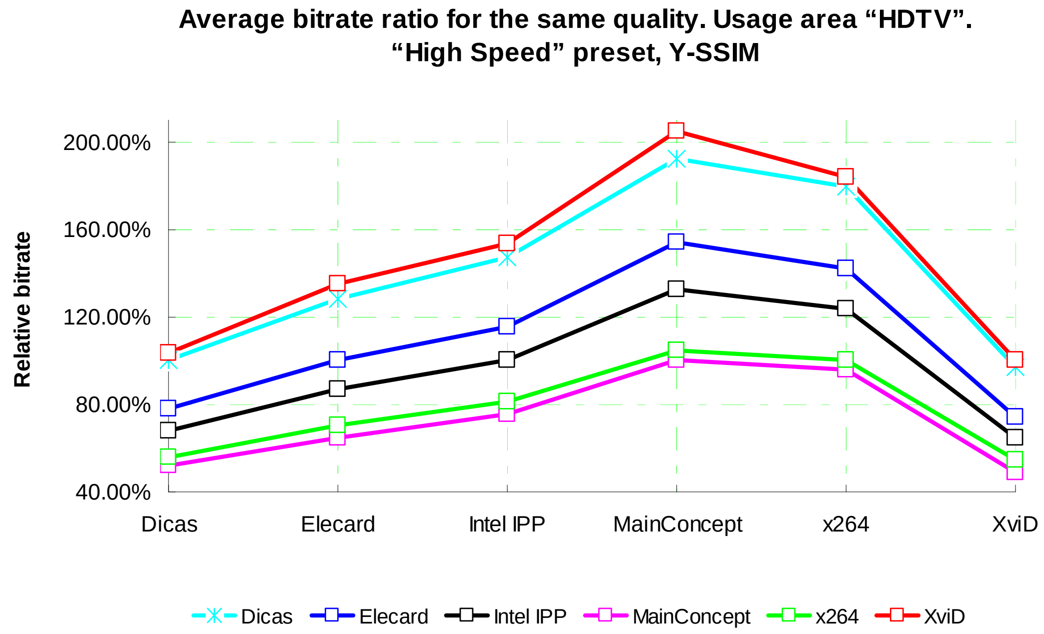


Figure 29. Average bitrate ratio for the same quality. Usage area “HDTV”. “High Speed” preset, Y-SSIM.

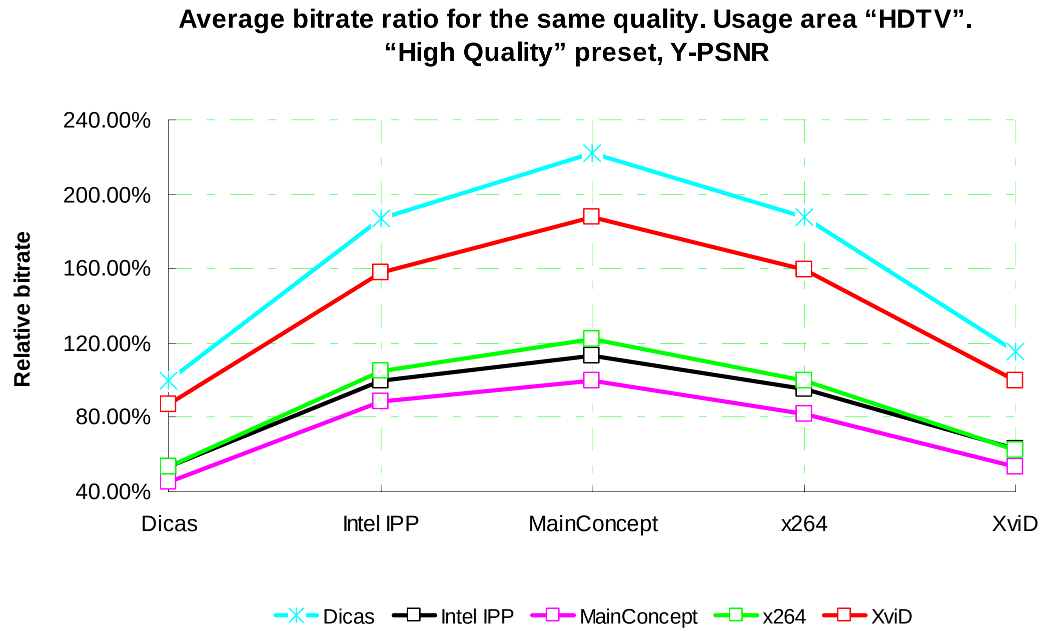


Figure 30. Average bitrate ratio for the same quality. Usage area “HDTV”. “High Quality” preset, Y-PSNR.

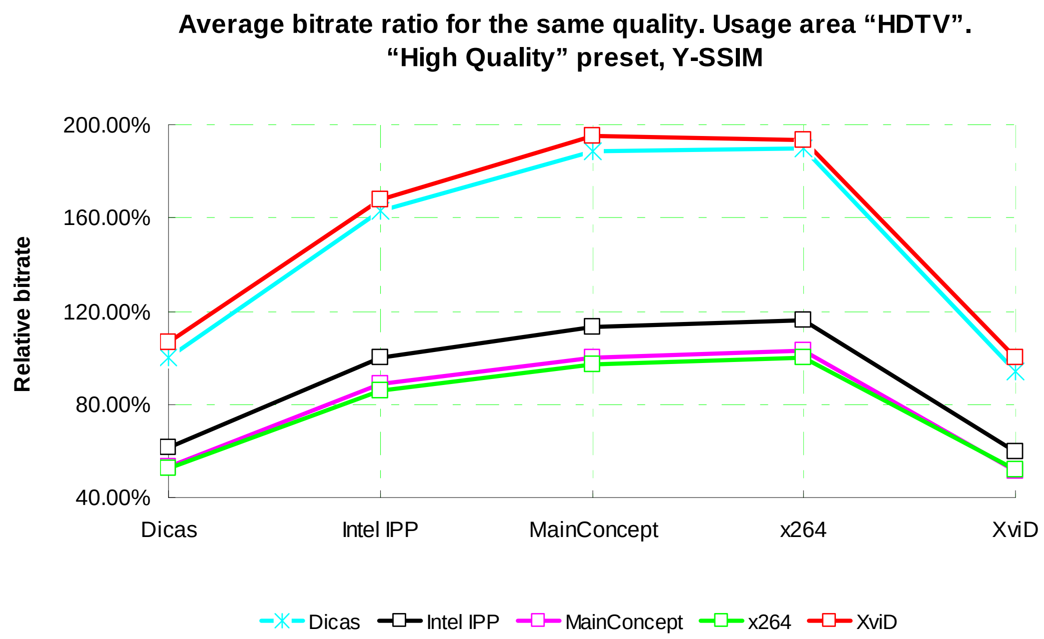


Figure 31. Average bitrate ratio for the same quality. Usage area “HDTV”. “High Quality” preset, Y-SSIM.

4.4 Conclusions

4.4.1 Video conferences

All the analysis is being made with help of Y-SSIM as main quality metric. Leaders in the videoconference area are the x264 and MainConcept codecs, with MainConcept being the slightly better alternative. The worst quality is demonstrated by the XviD encoder as a presenter of MPEG-4 ASP encoder. The dicas codec is the only one which was provided with authentic video conference presets (disabled B-frames, low VBV-buffer) which have a negative impact on picture quality.

4.4.1.1 High Quality preset

MainConcept demonstrates the best quality for all sequences. The top three codecs for this preset are the following:

1. MainConcept
2. x264
3. Intel IPP H.264

The top three codecs also demonstrate acceptable bitrate handling.

4.4.1.2 High Speed preset

MainConcept demonstrates the best quality for all sequences. The top three codecs for this preset are the following:

1. MainConcept
2. x264

3,4. Elecard, Intel IPP H.264

Elecard is being very slightly better than Intel IPP H.264. The first four codecs also demonstrate acceptable bitrate handling.

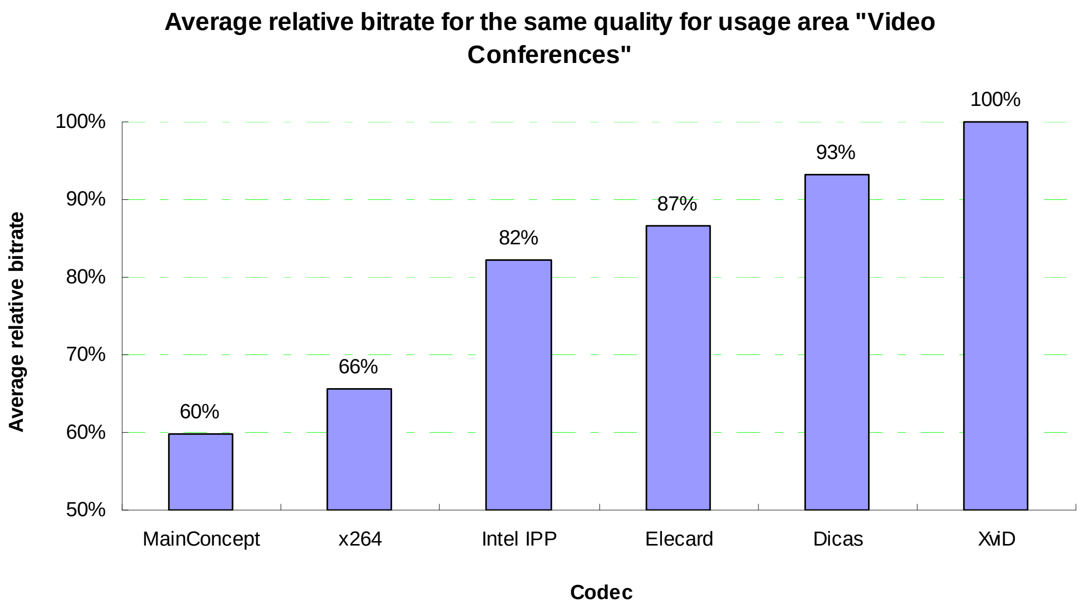


Figure 32. Average bitrate ratio for the same quality. Usage area “Video Conferences”. All presets, Y-SSIM.

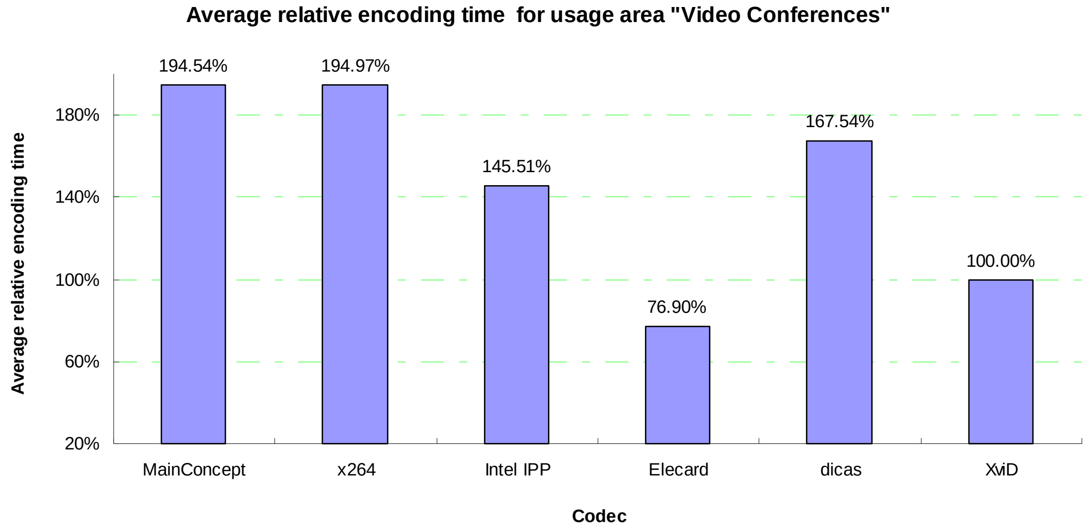


Figure 33. Average relative encoding time. Usage area “Video Conferences”. All presets.

4.4.2 Movies

The leading encoders in this category are MainConcept and x264. The quality of the XviD encoder is again rather low.

4.4.2.1 High Quality Preset

The x264 encoder demonstrates better quality at average, and MainConcept show slightly lower quality. The bitrate handling algorithm of these codecs is acceptable for this category. The Intel IPP H.264 codec once again holds third place and it has faster encoding speed than x264 and MainConcept. The top three codecs for this preset are the following:

1. x264
2. MainConcept
3. Intel IPP H.264

4.4.2.2 High Speed Preset

The results for this preset are similar to those of the High Quality preset. The leaders are the x264 and MainConcept codecs. In third place, once again, is the Intel IPP H.264 encoder, but it has faster encoding speed than x264 and MainConcept. The top three codecs for this preset are the following:

1. x264
2. MainConcept
3. Intel IPP H.264

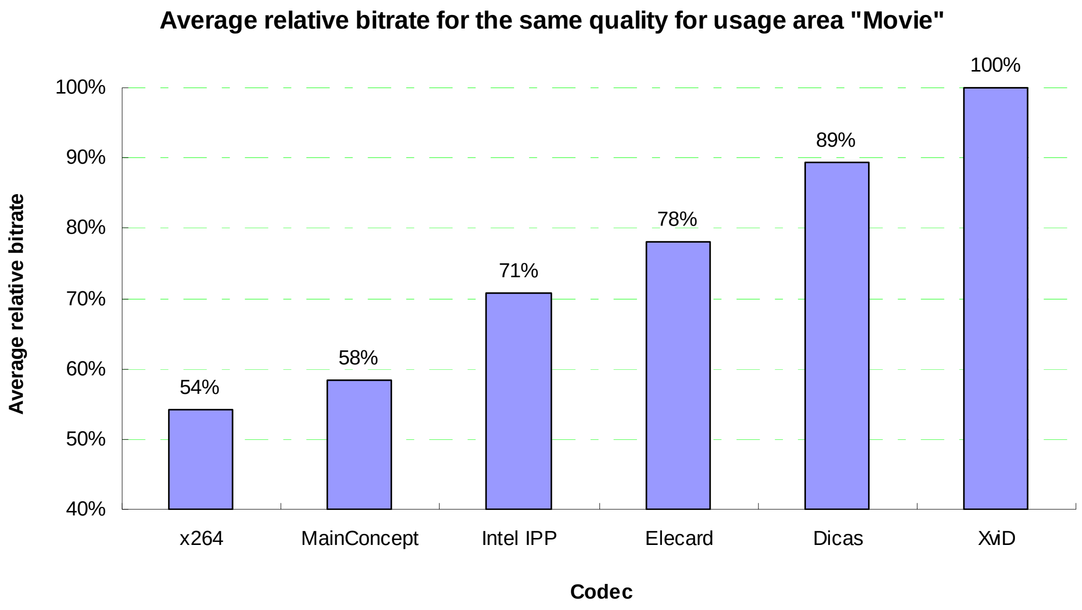


Figure 34. Average bitrate ratio for the same quality. Usage area "Movie". All presets, Y-SSIM.

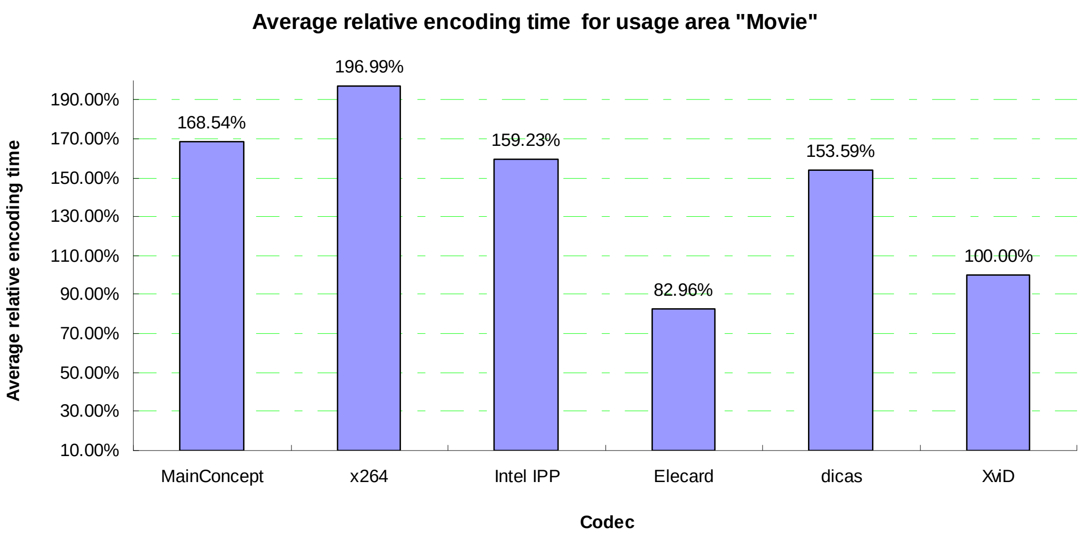


Figure 35. Average relative encoding time. Usage area "Movie". All presets.

4.4.3 HDTV

4.4.3.1 High Quality Preset

The MainConcept and x264 codecs demonstrate the highest quality among all the codecs tested in this comparison. The encoding quality of the MainConcept codec is better than quality the x264 encoder and the speed is faster. The third-place encoder, rated by quality, is the Intel IPP H.264 codec. And, it is faster than MainConcept and x264. The top three codecs for this preset are the following:

1. MainConcept
2. x264
3. Intel IPP H.264

4.4.3.2 High Speed Preset

The situation is almost the same as for High Quality preset. The top three codecs for this preset are the following:

1. MainConcept
2. x264
3. Intel IPP H.264

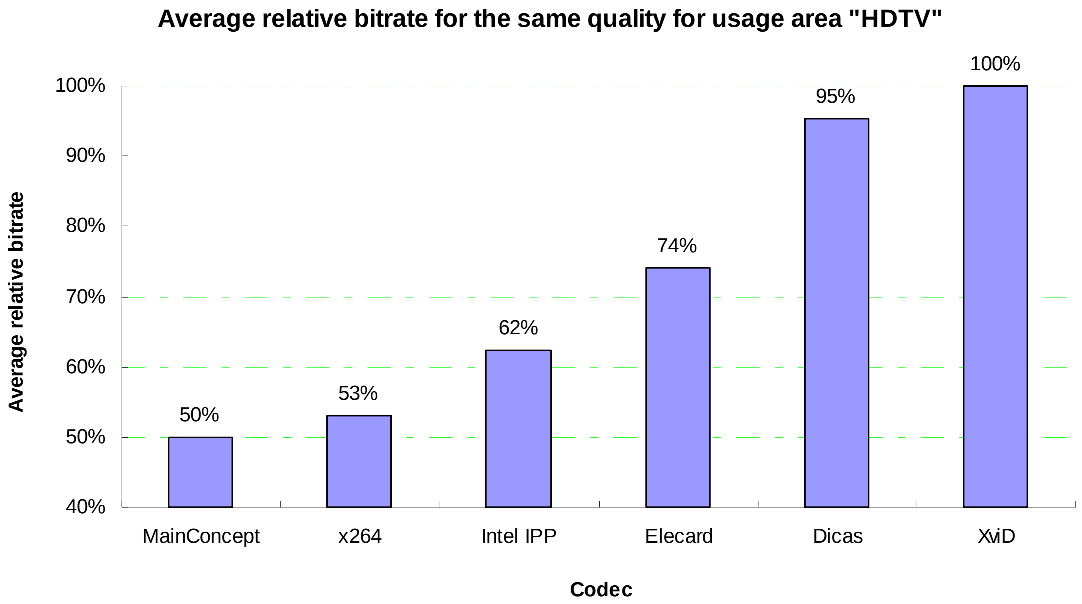


Figure 36. Average bitrate ratio for the same quality. Usage area “HDTV”. All presets, Y-SSIM.

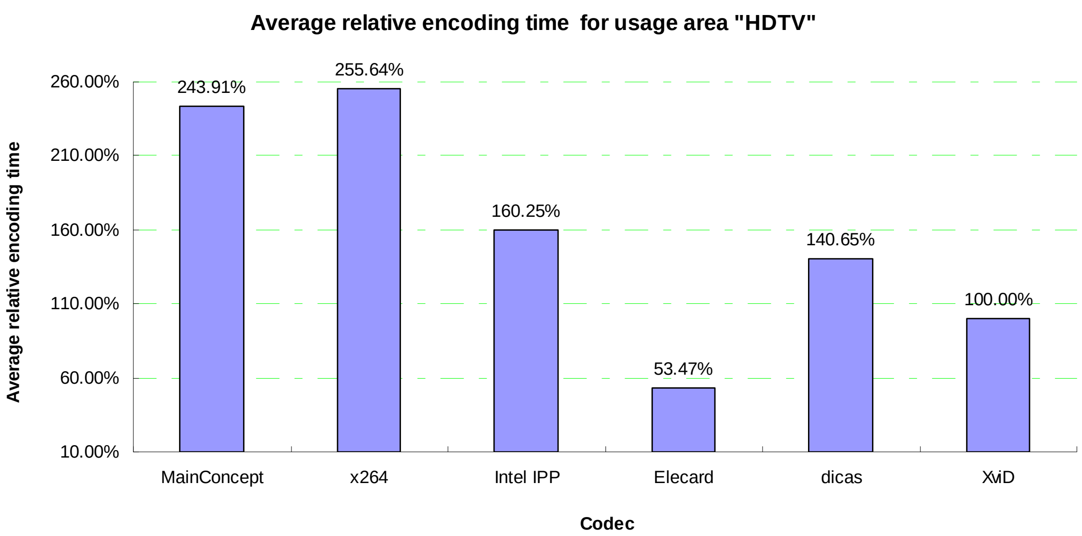


Figure 37. Average relative encoding time. Usage area “HDTV”. All presets.

4.4.4 Overall Conclusions

Overall, the leaders in this comparison are the MainConcept and x264 encoders, with the Intel IPP H.264 encoder taking a strong third place. The XviD (MPEG-4 ASP) codec is the last at average and it demonstrates difficulties with bitrate handling algorithms.

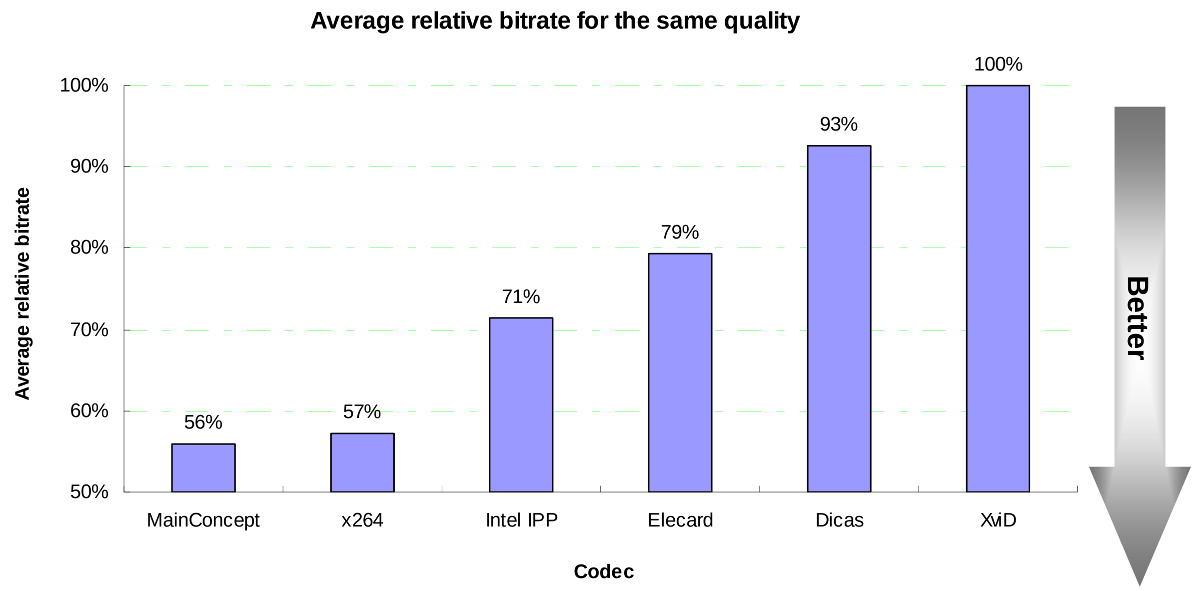


Figure 38. Average bitrate ratio for a fixed quality for all categories and all presets (Y-SSIM).

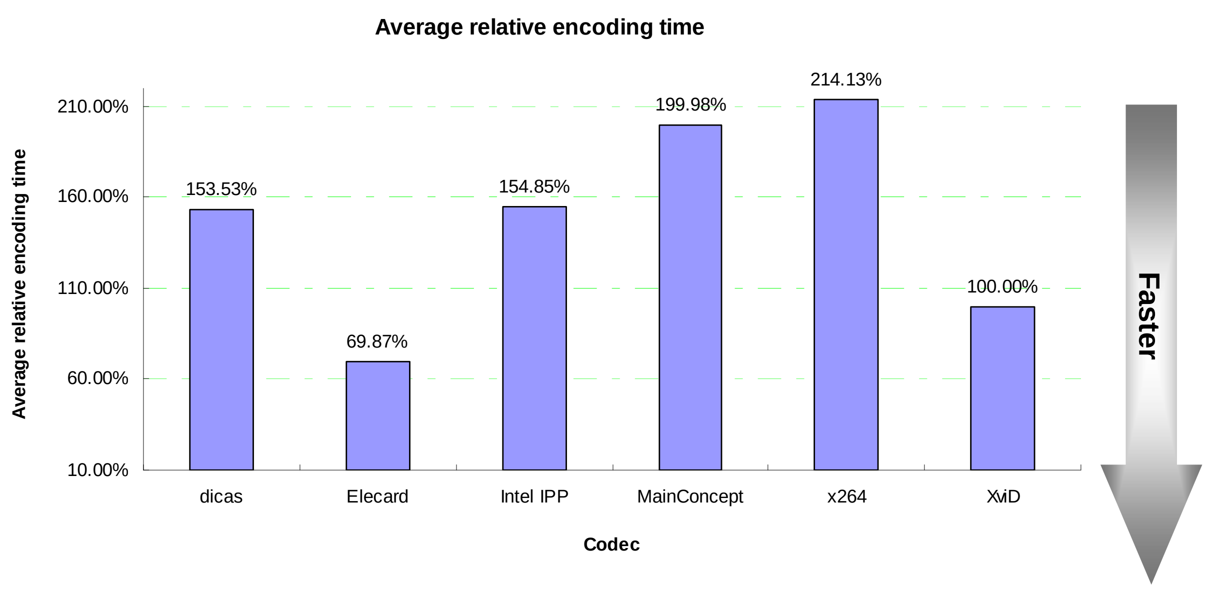


Figure 39. Average relative encoding time for all categories and all presets.

The overall ranking of the codecs tested in this comparison is as follows:

1. MainConcept
2. x264
3. Intel IPP H.264
4. Elecard
5. Dicas
6. XviD

This rank based only on quality results of encoders (see Figure 38). Encoding speed is not considered here.

The difference between the MainConcept and x264 encoders is not overly significant, so these two encoders are both the clear leaders in this comparison. The developers of the Elecard encoder do not provide a High Quality preset, so its ranking is based solely on the results for the High Speed preset. And its encoding speed even faster than for XviD.

4.4.5 Comments from developers

dicas digital image coding GmbH:
"The dicas codec comes off badly mainly because of statistical reasons. It is the only codec which was provided with authentical video conference presets (disabled B-frames and low VBV-buffer) which have a negative impact on picture quality and a HVSAQM algorithm which countervails PSNR as standard of measurement."

Intel Corporation:
"Along with Intel Integrated Performance Primitives (Intel IPP) library, the threaded library of highly optimized functions for multimedia and data processing applications, several media-encoding samples are delivered, including tested H.264 encoder. Originally this encoder was only meant to illustrate how Intel IPP functions can be used for efficient video-encoding, and we are proud that it is capable to compete well with such industry leaders as MainConcept. To our opinion, this year's competition demonstrates that, along with some advanced quality control features like adaptive MB quantization we are currently lacking, major role in defining final quality difference and resulting places distribution plays different rate-control strategies. By design our encoder was aimed to be included into video-processing pipelines and for that reason our rate control is providing near-constant bitrate locally for the whole encoded sequence making it difficult to efficiently redistribute bits between scenes with different complexity. That is why it comes to no surprise that two-pass rate control outperforms it at relatively low bitrates and at the same time that in most cases we managed to provide better bitrate handling than our competitors.

For the next competition we are suggesting to include more information about contenders, their availability for customers, licensing terms and pricing information, together with some overview of the encoders market in general. It would be better to include into benchmark codec products that are representatives of the current encoder market. It would be also very interesting to see how hardware-assisted encoder solutions will perform against software-only ones. Overall, this was an interesting and challenging competition, thank you MSU, and keep up a good work!"

5 Appendix 1. One Pass High Quality Encoding

In the main comparison part for High Quality encoding some encoders use 2-pass encoders and some encoders use 1-pass encoding, and because of it this special part was introduced to compare 1-pass High Quality encoding presets for Movie and HDTV usage areas.

This is the short version of the report. Only part of relative quality analysis is present below. All the other One Pass High Quality Encoding results (RD curves, bitrate handling, encoding speed, etc) can be found in the full version. You can purchase the full version of the report at the comparison web-page.

5.1 Movie

5.1.1 Relative Quality Analysis

Table 13 and Table 14 show relative bitrates for a fixed quality output for all codecs and presets. Note that these tables do not include information about the speed of the encoder.

As one can see the overall result strongly depends on what quality metric was used. At average, there are two leaders: x264 and MainConcept encoders, but if to use Y-PSNR the leader is MainConcept, but for Y-SSIM the leader is x264. The Intel IPP H.264 encoder has third place.

Table 13. Average bitrate ratio for the same quality. Usage area "Movie". "High Quality one pass" preset, Y-PSNR.

	IPP H.264	MainConcept	x264	XviD
IPP H.264	100%	86%	94%	131%
MainConcept	116%	100%	109%	152%
x264	106%	92%	100%	140%
XviD	76%	66%	71%	100%

Table 14. Average bitrate ratio for the same quality. Usage area "Movie". "High Quality one pass" preset, Y-SSIM.

	IPP H.264	MainConcept	x264	XviD
IPP H.264	100%	88%	82%	146%
MainConcept	114%	100%	93%	165%
x264	123%	107%	100%	177%
XviD	69%	61%	57%	100%

Figure 40 and Figure 41 visualize data in the tables above. Each line in those figures corresponds to one codec. Values in vertical axis are average relative bitrate comparing to the codecs in horizontal axis. The lower bitrate is the better relative results have the codec.

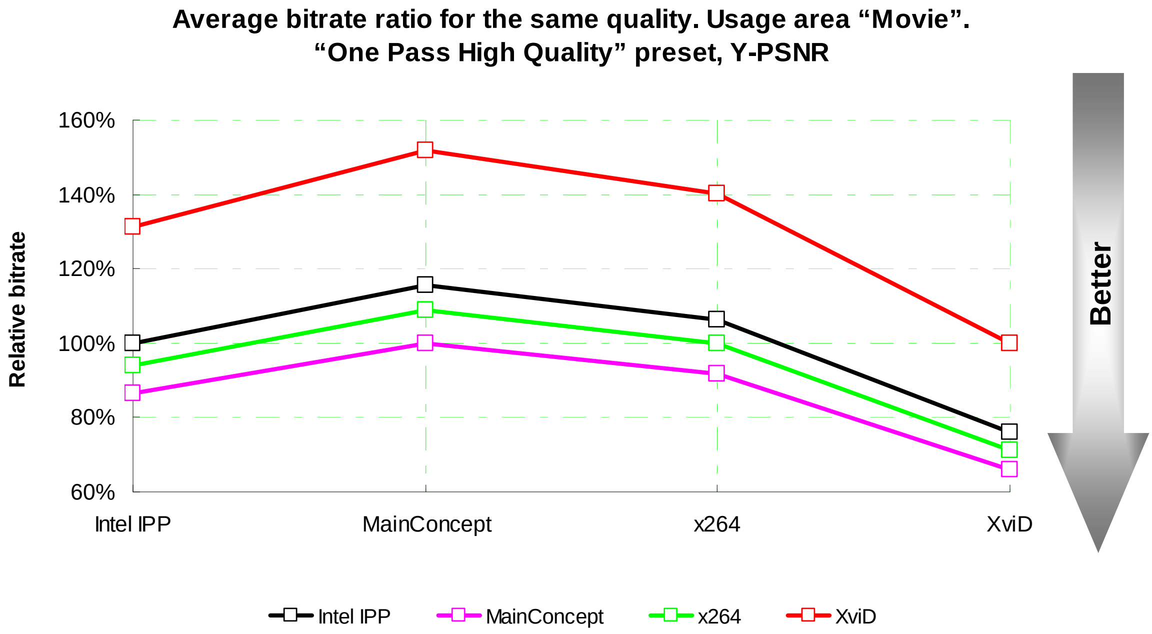


Figure 40. Average bitrate ratio for the same quality. Usage area “Movie”. “One Pass High Quality” preset, Y-PSNR.

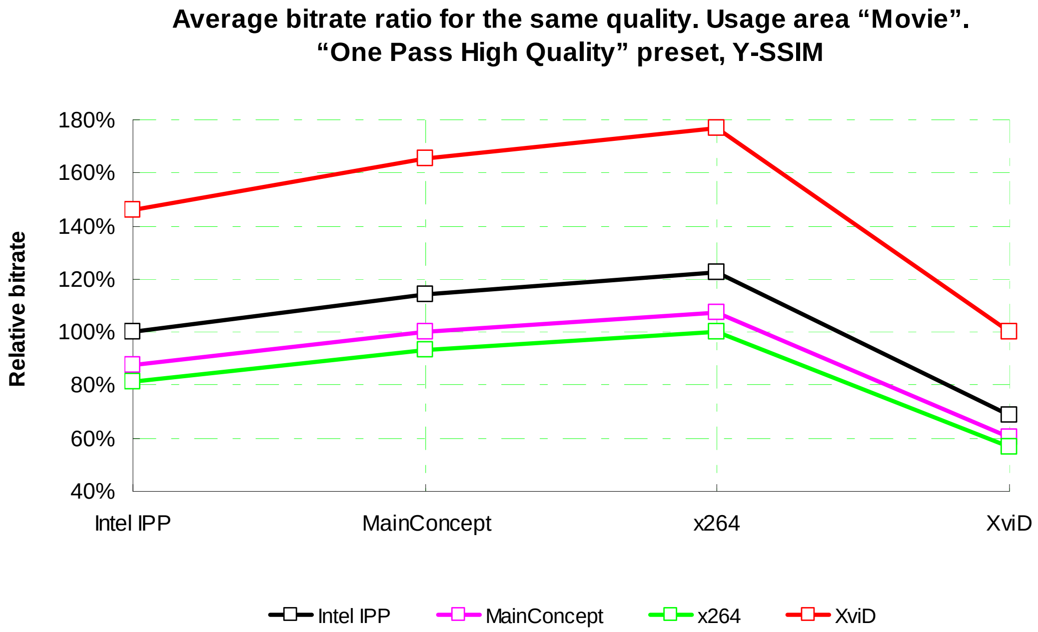


Figure 41. Average bitrate ratio for the same quality. Usage area “Movie”. “One Pass High Quality” preset, Y-SSIM.

5.2 HDTV

5.2.1 Relative Quality Analysis

Table 15 and Table 16 show relative bitrates for a fixed quality output for all codecs and presets. Note that these tables do not include information about the speed of the encoder.

The leader is MainConcept and second place depends on quality chosen metric – for Y-PSNR Intel IPP H.264 is better than x264 and for Y-SSIM x264 is better than Intel IPP H.264 and for Y-SSIM the difference between MainConcept and x264 is very small.

Table 15. Average bitrate ratio for the same quality. Usage area "HDTV". "High Quality one pass" preset, Y-PSNR.

	IPP H.264	MainConcept	x264	XviD
IPP H.264	100%	91%	112%	158%
MainConcept	110%	100%	126%	184%
x264	89%	79%	100%	149%
XviD	63%	54%	67%	100%

Table 16. Average bitrate ratio for the same quality. Usage area "HDTV". "High Quality one pass" preset, Y-SSIM.

	IPP H.264	MainConcept	x264	XviD
IPP H.264	100%	89%	91%	168%
MainConcept	112%	100%	101%	194%
x264	110%	99%	100%	181%
XviD	60%	52%	55%	100%

Figure 42 and Figure 43 visualize data in the tables above. Each line in those figures corresponds to one codec. Values in vertical axis are average relative bitrate comparing to the codecs in horizontal axis. The lower bitrate is the better relative results have the codec.

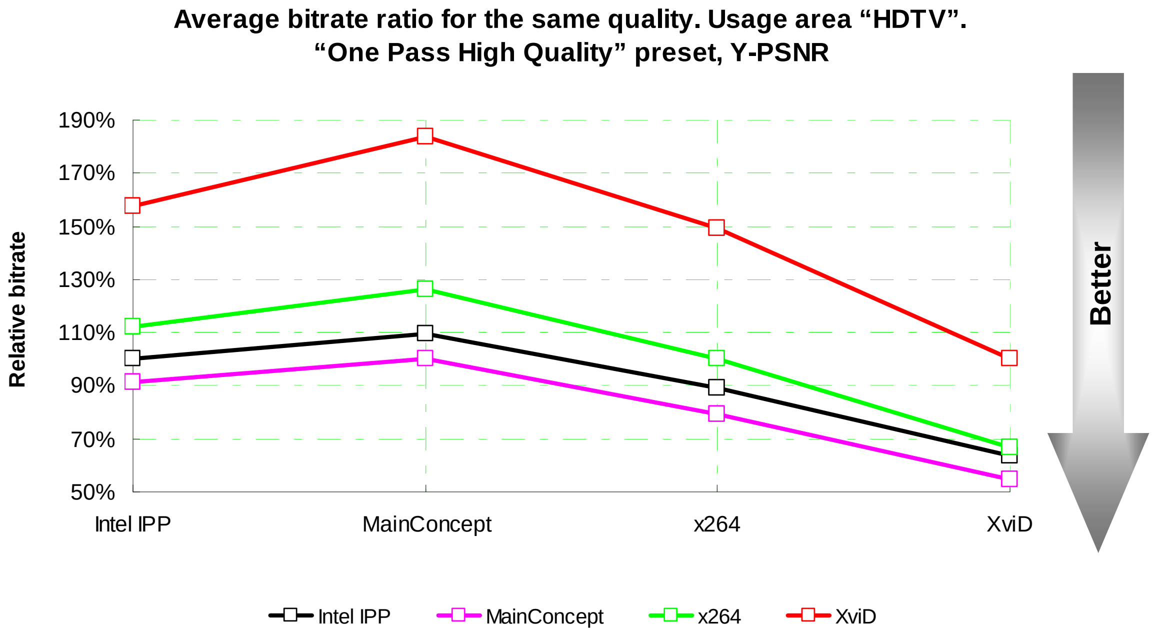


Figure 42. Average bitrate ratio for the same quality. Usage area "HDTV". "One Pass High Quality" preset, Y-PSNR.

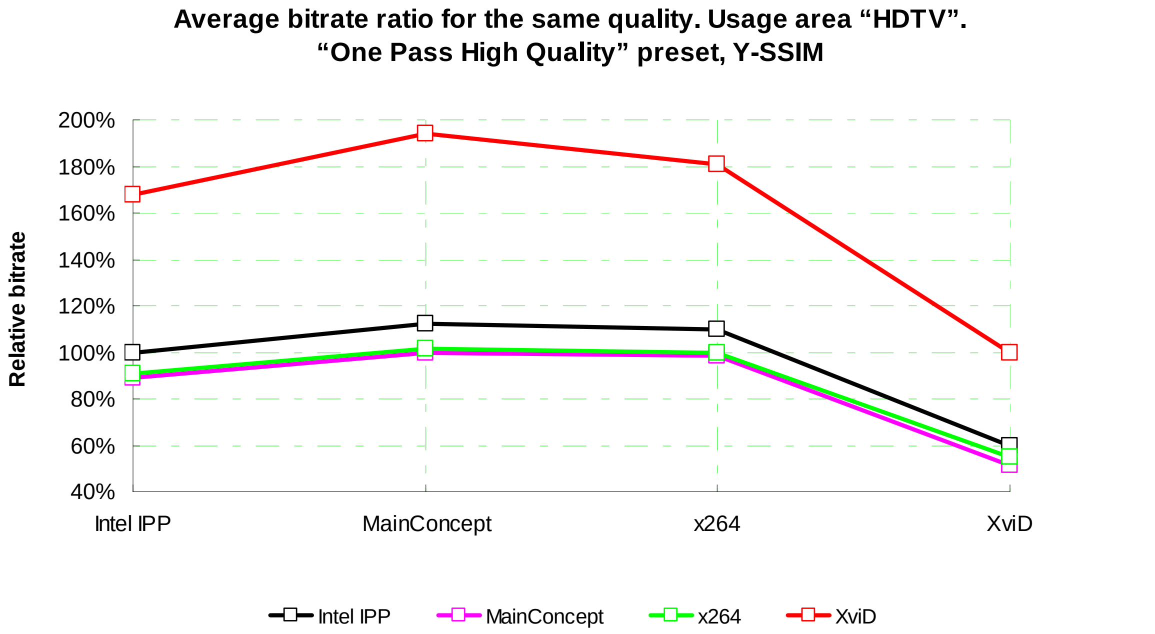


Figure 43. Average bitrate ratio for the same quality. Usage area "HDTV". "One Pass High Quality" preset, Y-SSIM.

6 Appendix 2. Codecs Analysis using Synthetic Sequences

This appendix contains some examples of codec analysis using synthetic sequences. The main purpose of this method is to track some important codec properties using specially created synthetic video sequences. Two types of analysis have been performed:

- Sequences for motion compensation quality estimation
- Sequences for moving object tail area distortion estimation

6.1 Estimation of Motion Compensation Quality

Synthetic sequences are used for detailed analysis of the motion compensation algorithm. The sequence consists of squares that move from frame to frame. Each object can be described with the following parameters:

- Texture
- Size
- Position (X, Y)
- Speed (V_x, V_y)

Textures for moving objects and for backgrounds are created using superposition of sinusoids.

The size of each object is selected randomly using a normal distribution with parameters that depend on the frame resolution.

The initial position of each square is random. Later, for each frame *i+1*, the position is calculated using the following formulas:

$$X_{i+1} = X_i + V_x^i$$
$$Y_{i+1} = Y_i + V_y^i$$

Calculation of the speed of the square has two stages:

1. Addition of a random component to the speed: a uniform random variable in the range [0, *MAX_SPEED*]. The *MAX_SPEED* constant is used to control sequence complexity.
2. Calculation of the correlation component for the speeds of different objects. The correlation component is used for emulating the correlation between the motion of different objects in the scene.

Figure 44 depicts an example frame from a synthetic sequence.

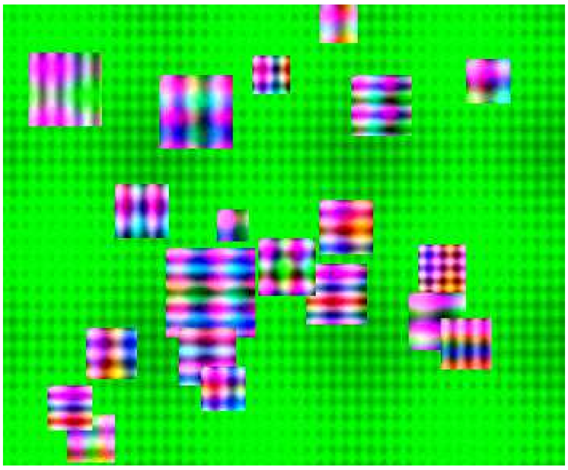


Figure 44. Example frame from a synthetic sequence used for motion analysis.

Two synthetic sequences with different complexities are used in this test. The first step is relative values calculation. The codec under test is launched for each synthetic sequence, and the resulting output quality is compared to that of the reference codec. Therefore, for each sequence and codec, there is one descriptive number (the average bitrate ratio, relative to the reference codec, for a fixed quality).

The final score consist of two parts:

- Encoding quality for the first version of the motion sequence (with simple motion)
- Quality variation for the second version of the motion sequence (with complex motion)

This is the short version of the report. The final results are presented only in full version of report. You can purchase the full version of the report at the comparison web-page.

6.2 Analysis of Distortion in Tail Area

The main purpose of tail area analysis is to test the ability of a codec to properly deal with newly appeared objects and areas.The synthetic sequence for this analysis consists of a static background and a single moving square. The square moves from right to left with a slight random variation in speed. When the object reaches the left border of the frame, it appears again at the right border with a new Y coordinate.

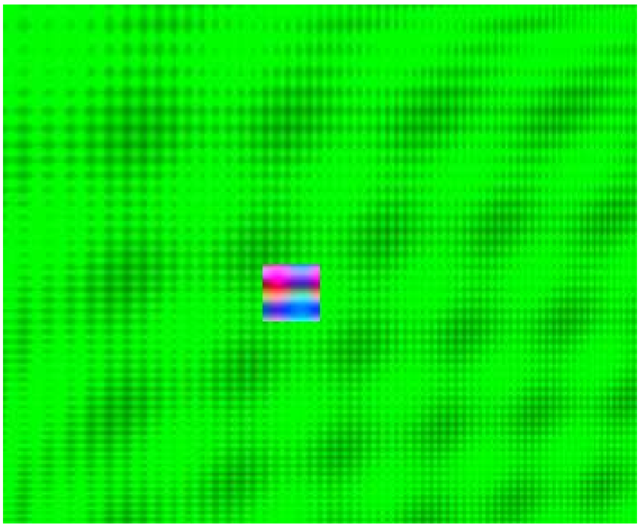


Figure 45. Synthetic sequence for tail area analysis.

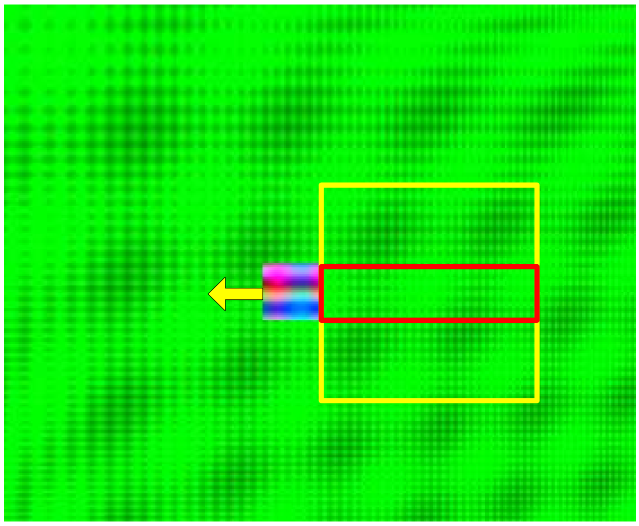


Figure 46. Areas for calculating the quality metric for tail area analysis.

The first step of the final score calculation involves computing the tail area metric for one frame. A rectangle of a specific size, located directly to the right of the moving objects, is considered. This rectangle is divided into two areas: a center area, traversed by the moving object (the red rectangle in figure above) and the remaining area (the yellow rectangle). The ratio of per-frame distortions in those areas is the tail area metric for this frame. The metric for the sequence is the average of the per-frame ratios.

Figure 45 and Figure 46 show an example of a frame from a synthetic sequence as well as the areas used for the calculation of the metrics.

6.3 Other Types of Analysis

The Graphics & Media Lab Video Group performs extensive research in the field of automatic codec analysis. The results of this section are just examples of codec analysis methods developed by the Video Group. Some analyzers currently being developed at the laboratory include the following:

- Blurring analyzer – test codec performance for high-texture sources
- Decimation analyzer – test the quality of the codec's motion estimation (ME) algorithm
- Frames with noise analyzer – test the stability of the codec's frame-level rate control
- Frames with variable noise and noise macro-blocks analyzers – test the quality of the codec's MB-level rate control
- Frame rotation analyzer – test for errors in the codec's ME
- Synthetic motion analyzer – test the quality of the codec's ME
- Tail area analyzer – analysis of codec mode selection
- Edge capture analyzer – analysis of codec MB subdivision selection
- Border quality analyzer – analysis of codec performance for sequences with sharp edges

7 Appendix 3. x264 Over-years Comparison

To analyze the progress of codec development we have decided to compare the quality of one of H.264 codecs over-years n the same video sequence. The x264 encoder was chosen for that task because it presents in almost every MSU VIDEO MPEG-4 AVC/H.264 codecs comparison and it has good results comparing to other encoders. Figure 47 through Figure 50 show the position of x264 codec comparing to other codecs on "Battle" sequence. At all years except 2005 x264 shows the best results. For 2006-2009 years we have shown results using Y-SSIM as quality metrics, for 2005 we have not use this metric as main. Because of these results x264 could be a good reference encoder for analyzing the overall progress of H.264 encoders during last years.

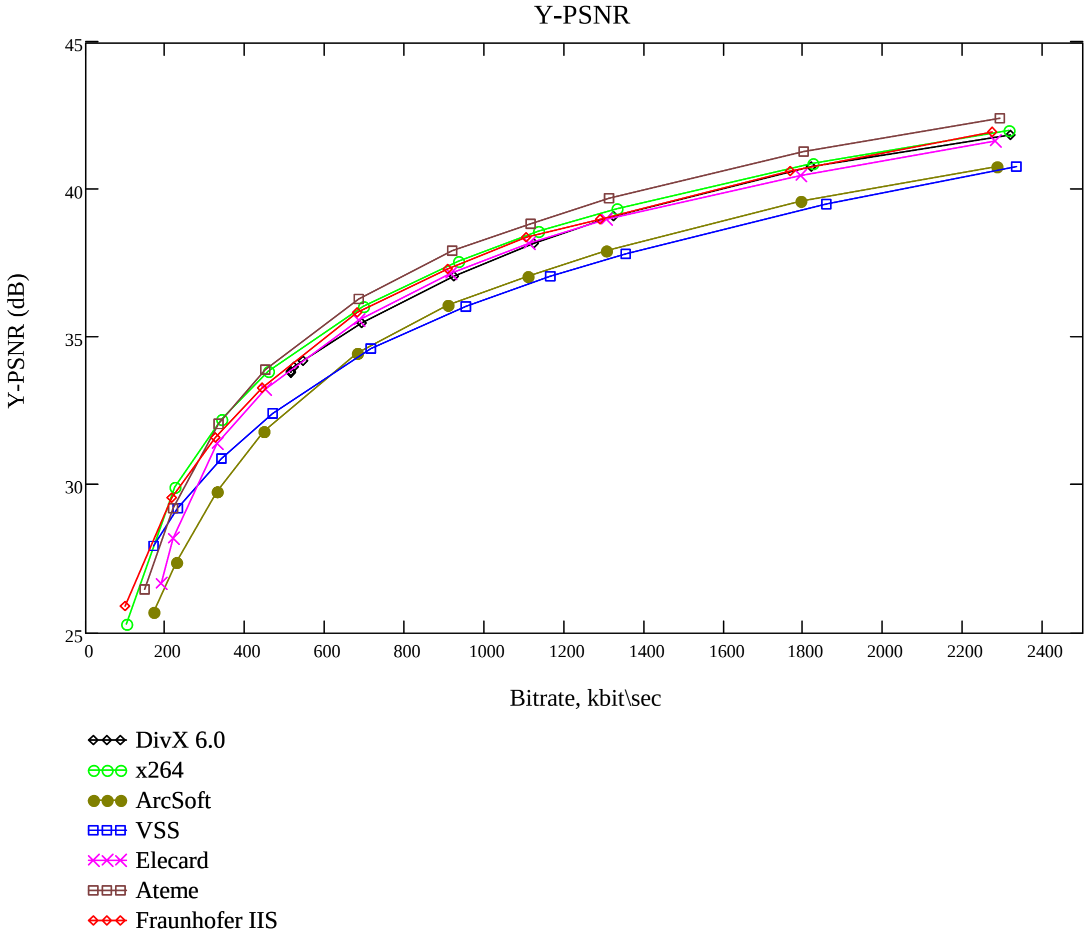


Figure 47. 2005 year. Bitrate/Quality. Usage area "Movies", "Battle" sequence, "High Quality" preset, Y-PSNR

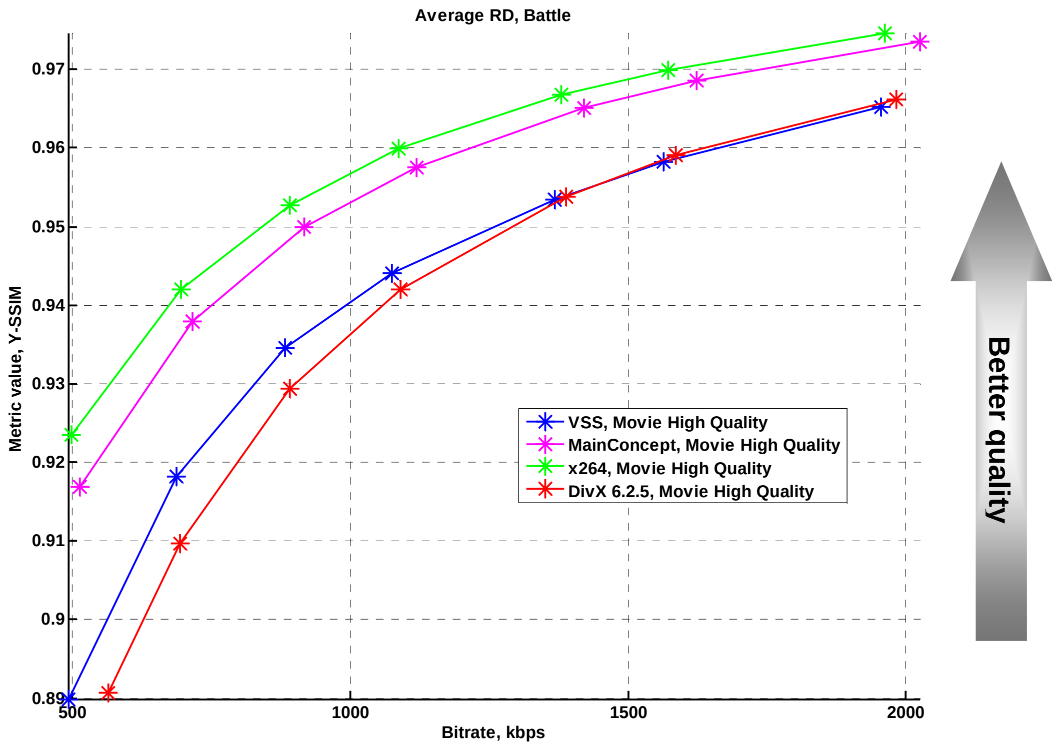


Figure 48. 2006 year. Bitrate/Quality. Usage area "Movies", "Battle" sequence, "High Quality" preset, Y-SSIM

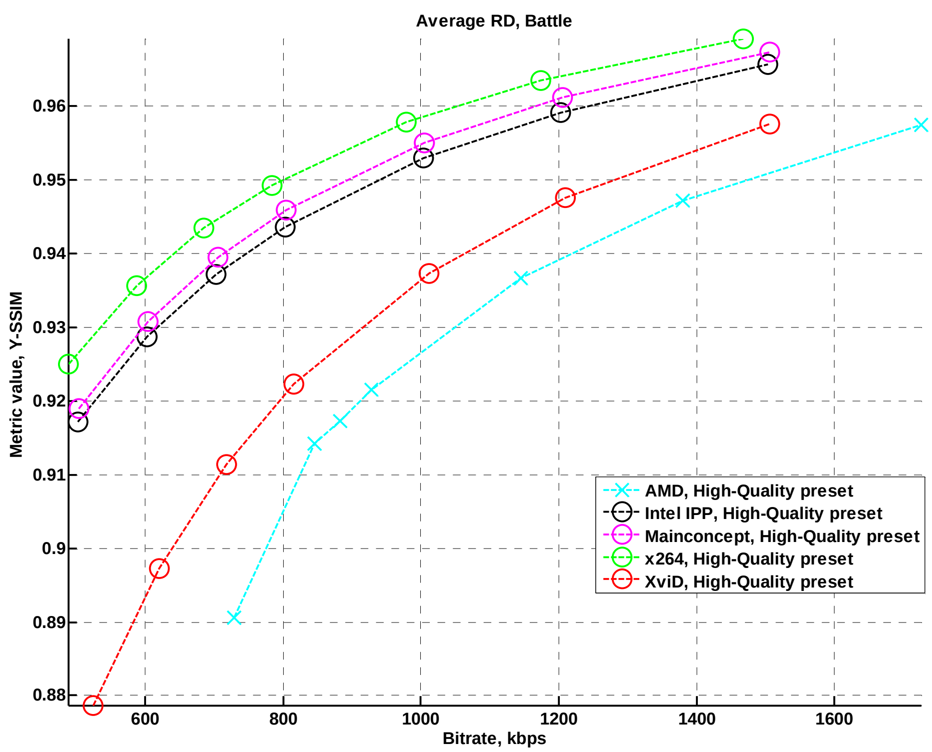


Figure 49. 2007 year. Bitrate/Quality. Usage area "Movies", "Battle" sequence, "High Quality" preset, Y-SSIM

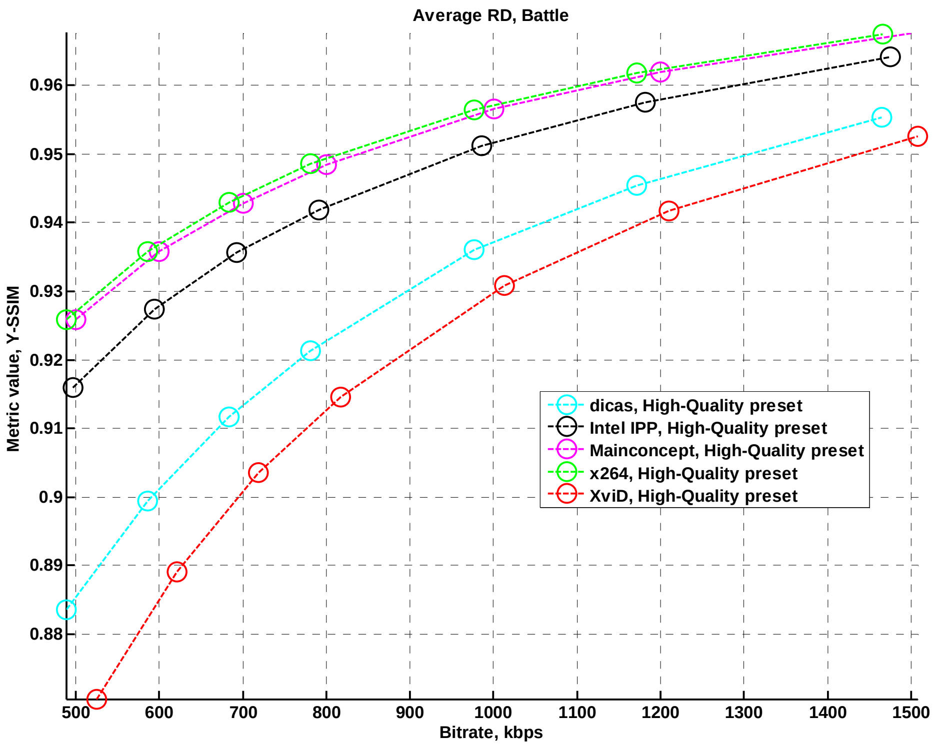


Figure 50. This year. Bitrate/Quality. Usage area “Movies”, “Battle” sequence, “High Quality” preset, Y-SSIM

Figure 51 shows RD curve for “Battle” sequence for x264 encoders of different years. As one can see the best encoder is x264 of current year, and worse is for 2005 year. X264_2006 is better than x264_2007, but it could be explained by different encoding speed, that is shown at Figure 52. This figure show the overall progress very good, x264 encoder became more faster and have better quality over years.

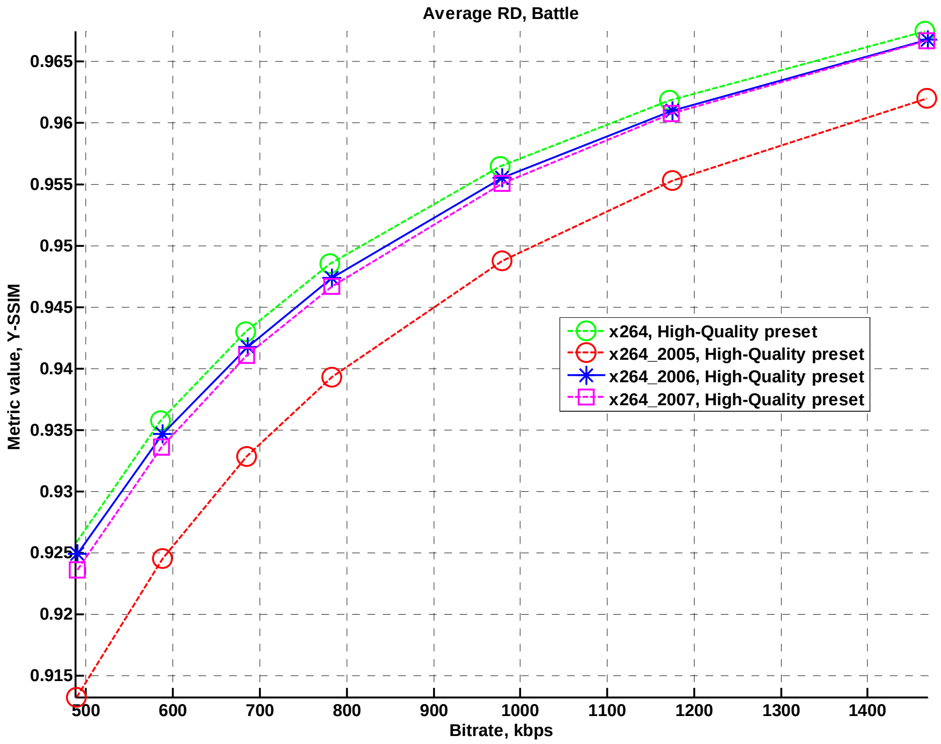


Figure 51. Different versions of x264 encoder. Bitrate/Quality. Usage area "Movies", "Battle" sequence, "High Quality" preset, Y-SSIM

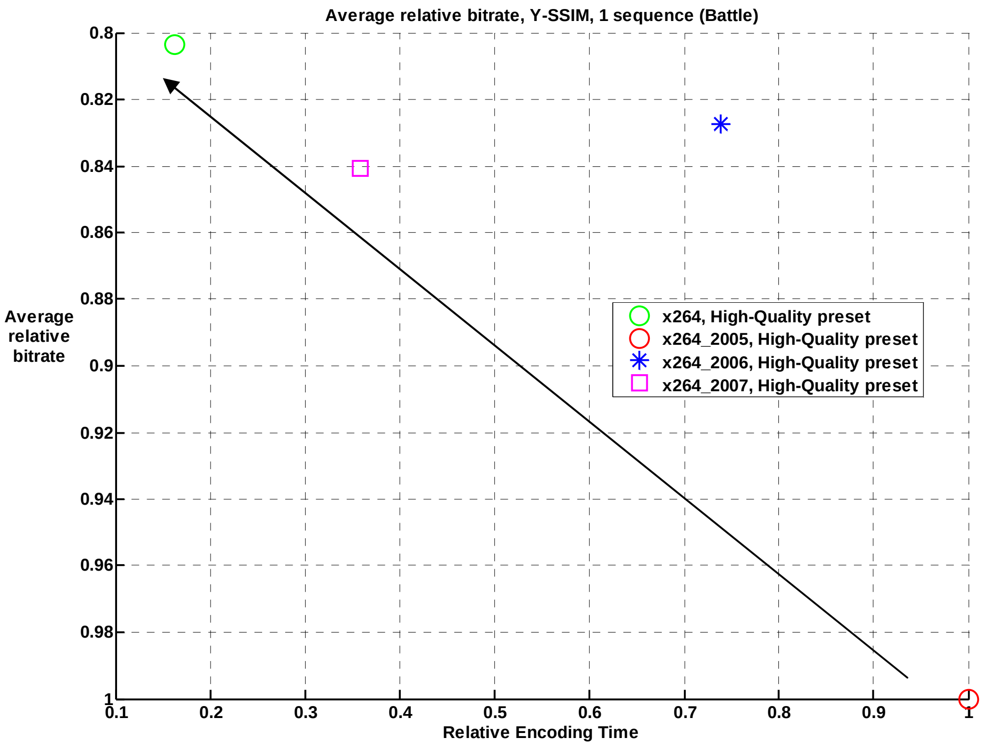


Figure 52. The progress of the x264 encoder over years. Y-SSIM

Bitrate handling mechanism for x264 encoder is quite good for every version, as one can see at Figure 53. Per frame analysis presented at Figure 54 shows that main encoding mechanism did not changed strongly.

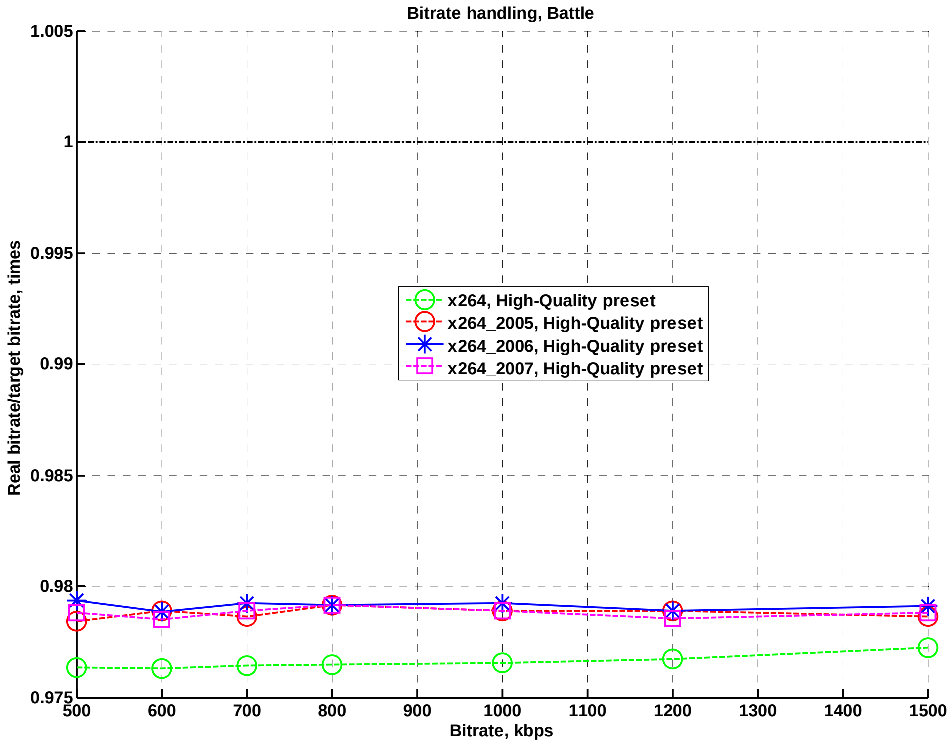


Figure 53. Different versions of x264 encoder. Bitrate handling.

Figure 54. x264 different versions. Per frame analysis, “Battle” sequence, 700kbps

8 Appendix 4. Test Set of Video Sequences

8.1 Videoconference Sequences

8.1.1 "Akiyo"

Sequence title	Foreman
Resolution	352x288
Number of frames	300
Color space	YV12
Frames per second	30
Source	Uncompressed (standard sequence), progressive

Figure 55. Akiyo sequence, frame 1

Figure 56. Akiyo sequence, frame 190

This is well-known sequence. This sequence includes static background and foreground with very low motion – only announcer's face with not very rich mimic. As a result, this sequence can e used to test the behavior of the codec for almost static scene.

8.1.2 "Foreman"

Sequence title	Foreman
Resolution	352x288
Number of frames	300
Color space	YV12
Frames per second	30
Source	Uncompressed (standard sequence), progressive

Figure 57. Foreman sequence, frame 77

Figure 58. Foreman sequence, frame 258

This is one of the most well-known sequences. The sequence includes a face with very rich mimic. There is not a high level of motion, but the motion that is present is disordered and does not have any forward characteristics. The intricate character of the motion creates problems for the motion compensation process. In addition, the camera is shaking, thus making the image unsteady. At the end of the sequence, the camera suddenly turns to the building site, and another scene with almost no motion follows. As a result, this sequence can also be used to test the behavior of the codec for a static scene that follows one with abundant motion.

8.1.3 “Paris”

Sequence title	Paris
Resolution	352x288
Number of frames	1065
Color space	YV12
Frames per second	30
Source	Uncompressed (standard sequence), progressive

Figure 59. Paris sequence, frame 1

Figure 60. Paris sequence, frame 754

This well-known sequence presents two television announcers in front of a static background. In this sequence there are some little objects with high level of motion, such as|: pen, fingers and a ball. Therefore, this sequence can be used to test the behavior of a codec for a mostly static scene with an area of intensive motion.

8.1.4 “Stefan”

Sequence title	Stefan
Resolution	352x288
Number of frames	300
Color space	YV12
Frames per second	30
Source	Uncompressed (standard sequence), progressive

Figure 61. Stefan sequence, frame 1

Figure 62. Stefan sequence, frame 251

This standard sequence presents a sport video with high level of motion. There are two types of motion – the camera motion and sportsmen motion. The first one is not very difficult to estimate, but the second one is very difficult ot estimate and to compensate. Therefore, this sequence can be used to test the behavior of a codec for sport translation or sequences with high level of motion.

8.2 Movie Sequences

8.2.1 “Indiana Jones”

Sequence title	Indiana Jones
Resolution	704x288
Number of frames	5000
Color space	YV12
Frames per second	30
Source	MPEG-2 (DVD), FlaskMPEG deinterlace

Figure 63. Indiana Jones sequence, frame 1

This sequence is a fragment from the “Indiana Jones” movie. In terms of compression, this sequence difficult because of two main reasons: of low contrast parts present and high level of motion in different scenes. And there are scenes with highly different types of motion – from almost static scenes with talking people to scenes with strong motions, like for example scene where stones fall.

8.2.2 "Battle"

Sequence title	Battle
Resolution	704x288
Number of frames	1599
Color space	YV12
Frames per second	24
Source	MPEG-2 (DVD), FlaskMPEG deinterlace

Figure 64. Battle sequence, frame 839

This sequence is a fragment from the beginning of the "Terminator 2" movie. In terms of compression, this sequence is the most difficult among all of the sequences that were used in the analysis. This difficulty is due to three main reasons: continual brightness variation (resulting from explosions and laser flashes as seen in the picture above), very fast motion and frequent scene changes. These characteristics often cause codecs to compress frames as I-frames.

8.2.3 “State Enemy”

Sequence title	State Enemy
Resolution	720x304
Number of frames	6500
Color space	YV12
Frames per second	24
Source	MPEG-2 (DVD), FlaskMPEG deinterlace

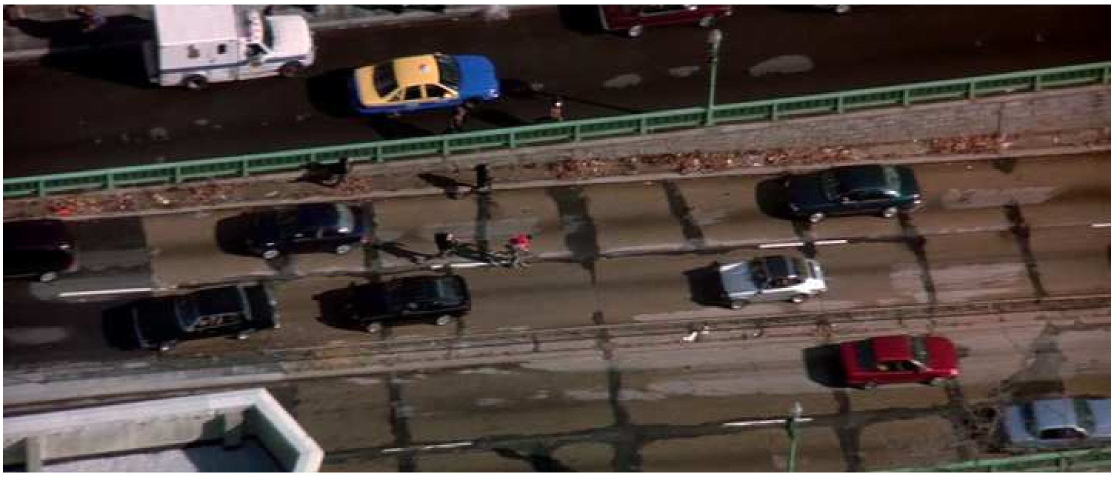


Figure 65. State Enemy sequence, frame 1115

This sequence is a fragment from the “Enemy of the State” movie. This sequence includes outdoor scenes with strong motion at the beginning when bicyclist runs, scenes with low motion and indoor scenes with normal motion. This sequence has scenes with different lightning conditions.

8.2.4 “Wendys”

Sequence title	Wendys
Resolution	720x576
Number of frames	260
Color space	YV12
Frames per second	30
Source	Uncompressed, progressive

Figure 66. Wendys sequence, frame 97

This sequence presents Wendy’s restaurant inside environment. This sequence is interest for strong grain during all the sequence and strong motion sometimes – these factors could lead to difficulties during compression for all tested codecs.

8.3 HDTV Sequences

8.3.1 “Mobile Calendar”

Sequence title	Mobile Calendar
Resolution	1280x720
Number of frames	504
Color space	YV12
Frames per second	50
Source	Uncompressed, progressive

Figure 67. Mobile Calendar sequence, frame 416

Similar to Mobile&Calendar. Close up. Moving calendar with text and a detailed photo of the Vasa ship. Moving train with colorful toys. Background with two types of wallpaper, one brown with details and one yellow with drawn figures. Very detailed and normally demanding. The main compression difficulty that could take place is many little sharp details in the sequence on the calendar and on the background.

8.3.2 "Pedestrian Area"

Sequence title	Pedestrian Area
Resolution	1920x1080
Number of frames	375
Color space	YV12
Frames per second	25
Source	Uncompressed, progressive

Figure 68. Pedestrian Area sequence, frame 129

This is a shot of a pedestrian area. This sequence has low camera position, and people pass by very close to the camera. High depth of field. Static camera. This sequence is interesting for compression because of static camera and areas of different focusing and blurring/sharpening.

8.3.3 “Stockholm”

Sequence title	Stockholm
Resolution	1280x720
Number of frames	604
Color space	YV12
Frames per second	50
Source	Uncompressed, progressive

Figure 69. Pedestrian Area sequence, frame 574

Panning view over the Old Town of Stockholm. Detailed houses, water and moving cars. Panning view over the Old Town of Stockholm. Detailed houses, water and moving cars. This sequence is interesting for compression because of high level of noise and sharp details in the scenes and moving camera and objects such as cars a and water.

8.3.4 "Troy"

Sequence title	Troy
Resolution	1920x1072
Number of frames	300
Color space	YV12
Frames per second	24
Source	MPEG-2 (DVD)

Figure 70. Troy sequence, frame 1

This sequence is a fragment of the "Troy" movie and contains three parts with sharp scene changes. The video includes medium scene motion and slow camera motion. In terms of compression, this sequence is difficult to compress because of the many small details.

9 Appendix 5. Tested Codecs and Presets

9.1 Codecs

9.1.1 dicas' mpegable H.264 Command-line encoder

- Console encoding program
- dicas decoder was used for decoding
- Codec and presets were provided by dicas digital image coding GmbH specifically for this test

Remarks:

No remarks.

```
Usage: H264VidEncoder <yuv src file> <options>

-out <output_file>        Name of the compressed output-file

-size <width> <height>    Original image size

-resize <width> <height> Image size to use for encoding
-region <xoffset> <yoffset> <width> <height>
                          active region in encoded image to show on decoder side
-key   <val>              keyframe period (0=only first frame, 1=all)

-frames <nr>              Number of frames to decode

-rec <filename>           Reconstructed output file

-snr                      Enables SNR calculation

-speed <nr>               Between 0 (fastest) and 4 (slowest)

-nofilter                 Disable deblocking filter

-cavlc                    Use CAVLC encoding (Baseline Profile)

-mbaff                    Enables the MB-Aff coding mode

-bframe <cnt>             Number of max B-Frames

-adaptive_bframes                  Enables adaptive B-Frames. Use -bframe to set
the max number.
-bitrate <KBits per sec> Enable rate control with given bitrate

-vbvrate <KBits per sec> vbv drain per second (default=bitrate).

-vbv <KBits>              set vbv buffer size

-noscenecut               disable scene-cut detection

-fps <frames per second> Assumed framerate for rate control

-lookahead <num>          Num Frames to look ahead (default=100, min=50, max=1024
)
-quant <quantizer>        use constant quantizer (0..51) instead of constant bitr
ate
-c                        Disables all platform specific optimizations( enables C
-code )
-use_nal_del              Insert NAL-Access-Delimiters at the start of each slice
```

Figure 71. icas' mpegable H.264 Command-line encoder

9.1.2 Elecard AVC Video Encoder 8-bit edition

- Console encoding program version 3.9.4
- Reference decoder JM 9.8 was used for decoding
- Codec and presets were provided by Elecard Ltd Company specifically for this test

Remarks:

No remarks.

```
C:\WINDOWS\system32\cmd.exe
Elecard AVC Video Encoder 8-bit edition, ver. 0.0 (build Feb 17 2009)
usage: avcenc config.cfg [parameters list]
```

Figure 72. Elecard AVC Video Encoder 8-bit edition

9.1.3 Intel IPP H.264 enc

- Console encoding program is released as part of “IPP 6.1 Gold” media samples
- Reference decoder JM 9.8 was used for decoding
- Codec and presets were provided by Intel Corp specifically for this test

Remarks:

No remarks.

```
C:\WINDOWS\system32\cmd.exe
T:\H264_2009\framework\codecs\intel>video_enc_con
Error: Destination file name not found
Usage1: video_enc_con [m2|m4|h264|h263|h261|dv_sd|dv_50|dv_hd|vc1|avs] [Options]
 InputParFile OutputEncodedFile
Usage2: video_enc_con [m2|m4|h264|h263|h261|dv_sd|dv_50|dv_hd|vc1|avs] [Options]
 -i InputYUVFile -o OutputEncodedFile
Options:
   [-a file] - file for auxiliary pictures for alpha blending for H.264 encoder
   [-s file] - file for reconstructed frames (H.264 only!!!).
   [-i inputFile] - input YUV video file
   [-w width]     - video width
   [-h height]    - video height
   [-f frameRate] - video frame rate (frames per second)
   [-b bitRate]   - encoded bit rate (bits per second)
   [-n numFrames] - max frames to encode
   [-t num]       - number of threads for encoding
   [-y dir]       - directory of YUV files, names are taken from par files
   [-u file]      - file for dumping perfomance info
   [-p file.csv]  - file for dumping perfomance info into csv format
Note: Options -i,-w,-h,-f,-b,-n override corresponding values inside InputParFil
e.
      If InputParFile not specified, at least options -w,-h should be specified
and -o should be added before OutputEncodedFile.
```

Figure 73. Intel IPP H.264 encoder

9.1.4 MainConcept H.264/AVC encoder

- Console encoding program based on MainConcept CodecSDK 8.1
- Reference decoder JM 9.8 was used for decoding
- Codec and presets were provided by MainConcept AG Company specifically for this test

Remarks:

No remarks.

```
C:\WINDOWS\system32\cmd.exe
MainConcept (Broadcast) H.264/AVC encoder (build 8.0.0 at 2009/02/11 rev. 41900)

Copyright (c) 2009 MainConcept GmbH

THIS SOFTWARE IS FOR EVALUATION PURPOSES ONLY!
```

Figure 74. MainConcept H.264/AVC encoder

9.1.5 x264 encoder

- Console encoding program
- Reference decoder JM 9.8 was used for decoding
- Codec and presets were provided by developers specifically for this test

Remarks:

No remarks.

```
x264 core:67 r1123M 3d78062
Syntax: x264 [options] -o outfile infile [widthxheight]

Infile can be raw YUV 4:2:0 (in which case resolution is required),
  or YUV4MPEG 4:2:0 (*.y4m),
  or AVI or Avisynth if compiled with AVIS support (yes).
Outfile type is selected by filename:
 .264 -> Raw bytestream
 .mkv -> Matroska
 .mp4 -> MP4 if compiled with GPAC support (no)

Options:

  -h, --help                  List the more commonly used options
      --longhelp              List all options

Frame-type options:

  -I, --keyint <integer>      Maximum GOP size [250]
  -b, --bframes <integer>     Number of B-frames between I and P [0]
      --b-pyramid             Keep some B-frames as references
      --no-cabac              Disable CABAC
  -r, --ref <integer>         Number of reference frames [1]
  -f, --deblock <alpha:beta>  Loop filter AlphaC0 and Beta parameters [0:0]
      --interlaced            Enable pure-interlaced mode
```

Figure 75. x264 encoder

9.1.6 Xvid encoder

- Console encoding program
- Codec and presets were provided by developers especially for this test

Remarks:

No remarks.

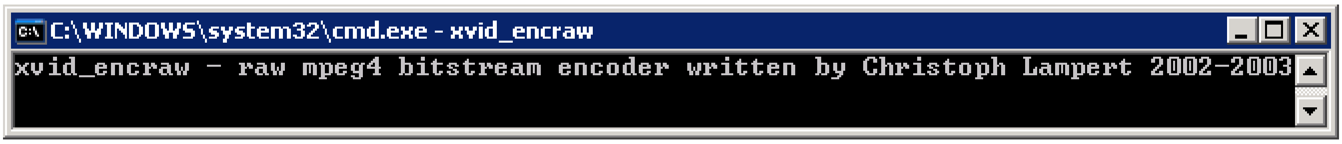


Figure 76. Xvid encoder

9.2 Presets

The table below lists the settings used in this comparison for all of the codecs.

Codec	Preset Name	Preset
Xvid	VideoConference “High Speed”	-max_bframes 2 -quality 6 -vhqmode 1 -bvhq -qpel -turbo -single
	VideoConference “High Quality”	-max_bframes 2 -quality 6 -vhqmode 4 -bvhq -qpel -turbo -single
	Movie “High Speed”	-max_bframes 2 -quality 6 -vhqmode 1 -bvhq -single
	Movie “High Quality”	-max_bframes 2 -quality 6 -vhqmode 4 -bvhq -gmc -single
	HDTV “High Speed”	-max_bframes 2 -quality 6 -vhqmode 1 -bvhq -turbo -single
	HDTV “High Quality”	-max_bframes 2 -quality 6 -vhqmode 4 -bvhq -gmc -single
MainConcept	VideoConference “High Speed”	HadamardTransform = 0 FastMRME = 1 NumRefFrames = 7 SearchRange = 63 FastIntraDecision = 0 FastSBME = 0
	VideoConference “High Quality”	HadamardTransform = 1 FastMRME = 0
	Movie “High Speed”	NumRefFrames = 4 SearchRange = 127 FastIntraDecision = 2 FastSBME = 2
	Movie “High Quality”	*First pass:* NumRefFrames = 4 SearchRange = 127 FastIntraDecision = 1 FastSBME = 1 *Second pass:* SearchRange = 255 HadamardTransform = 1 FastMRME = 0
	Movie “High Quality one Pass”	SearchRange = 255 HadamardTransform = 1 FastMRME = 0
	HDTV “High Speed”	NumRefFrames = 5 SearchRange = 511 HadamardTransform = 1 FastIntraDecision = 2 FastSBME = 2

	HDTV “High Quality”	*First pass:* NumRefFrames = 5 SearchRange = 511 FastIntraDecision = 1 FastMRME = 0 *Second pass:* NumRefFrames = 7 SearchRange = 511 HadamardTransform = 1 FastMRME = 0
	HDTV “High Quality one pass”	SearchRange = 511 HadamardTransform = 1 FastMRME = 0
Elecard	VideoConference “High Speed”	Profile 77 // Main Intra8x8InI 0 Intra8x8InP 0 Intra8x8InB 0
	“Movie “High Speed”	Profile 100 // High Intra8x8InI 1 Intra8x8InP 1 Intra8x8InB 1
	HDTV “High Speed”	Profile 100 // High ModeDecision 0 // 0 - SAD Intra8x8InI 1 Intra8x8InP 1 Intra8x8InB 1
Intel	“High Speed”	2 1 /* number of B frames between I (or P) and the next P, treat B as a reference (only 0 is supported!) */ 6 1 0 /* num_ref_frames (2-16), minimum length of list1 for backward prediction (only 1 is supported!), number of slices. */ 3 /* frame_rate_code */ 2 2 12 12 /* ME method (1-6), subblock split, search_x, search_y */
	“High Quality”	1 1 /* number of B frames between I (or P) and the next P, treat B as a reference (only 0 is supported!) */ 1 1 /* number of B frames between I (or P) and the next P, treat B as a reference (only 0 is supported!) */ 5 /* frame_rate_code */ 2 0 8 8 /* ME method (1-6), subblock split, search_x, search_y */

x264	VideoConference "High Speed"	--no-psnr --no-ssim --bframes 3 --b-pyramid --keyint 500 --threads auto --psy-rd 0 --subme 7 --ref 7 --aq-strength 1.25 --8x8dct --me umh --trellis 2 --b-adapt 0 --scenecut -1 --ratetol 10 --mixed-refs --partitions all
	VideoConference "High Quality"	--no-psnr --no-ssim --bframes 3 --b-pyramid --keyint 500 --threads auto --psy-rd 0 --subme 7 --ref 7 --aq-strength 1.25 --8x8dct --me umh --trellis 2 --b-adapt 0 --scenecut -1 --ratetol 10 --mixed-refs --partitions all
	Movie "High Speed"	*First pass:* --no-psnr --no-ssim --bframes 3 --b-pyramid --keyint 500 --threads auto --psy-rd 0 --pass 1 --subme 1 --direct auto --b-adapt 1 --partitions none Second pass: --no-psnr --no-ssim --bframes 3 --b-pyramid --keyint 500 --threads auto --psy-rd 0 --pass 2 --subme 8 --direct auto --ref 4 --8x8dct --me umh --merange 8 --trellis 1
	Movie "High Quality"	*First pass:* --no-psnr --no-ssim --bframes 3 --b-pyramid --keyint 500 --threads auto --psy-rd 0 --pass 1 --subme 7 --direct auto --b-adapt 2 Second pass: --no-psnr --no-ssim --bframes 3 --b-pyramid --keyint 500 --threads auto --psy-rd 0 --pass 2 --subme 11 --direct auto --ref 7 --8x8dct --me umh --trellis 2 --mixed-refs
	Movie "High Quality one pass"	--no-psnr --no-ssim --bframes 3 --b-pyramid --mixed-refs --keyint 500 --threads auto --psy-rd 0 --subme 11 --ref 8 --8x8dct --me umh --trellis 2 --ratetol 10 --b-adapt 2 --partitions all
	HDTV "High Speed"	*First pass:* --no-psnr --no-ssim --bframes 3 --b-pyramid --keyint 500 --threads auto --psy-rd 0 --pass 2 --subme 11 --direct auto --ref 7 --8x8dct --me umh --trellis 2 --mixed-

		refs Second pass: --no-psnr --no-ssim --bframes 3 --b-pyramid --keyint 500 --threads auto --psy-rd 0 --pass 2 --subme 11 --direct auto --ref 7 --8x8dct --me umh --trellis 2 --mixed-refs
	HDTV "High Quality"	*First pass:* --no-psnr --no-ssim --bframes 3 --b-pyramid --keyint 500 --threads auto --psy-rd 0 --pass 1 --subme 7 --direct auto --b-adapt 2 Second pass: --no-psnr --no-ssim --bframes 3 --b-pyramid --mixed-refs --keyint 500 --threads auto --psy-rd 0 --subme 10 --ref 5 --8x8dct --me esa --trellis 2 --ratetol 10 --b-adapt 2 --partitions all
	HDTV "High Quality one pass"	--no-psnr --no-ssim --bframes 3 --b-pyramid --mixed-refs --keyint 500 --threads auto --psy-rd 0 --subme 10 --ref 5 --8x8dct --me esa --trellis 2 --ratetol 10 --b-adapt 2 --partitions all

10 Appendix 6. Figures Explanation

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

10.1.1.1 RD curves

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

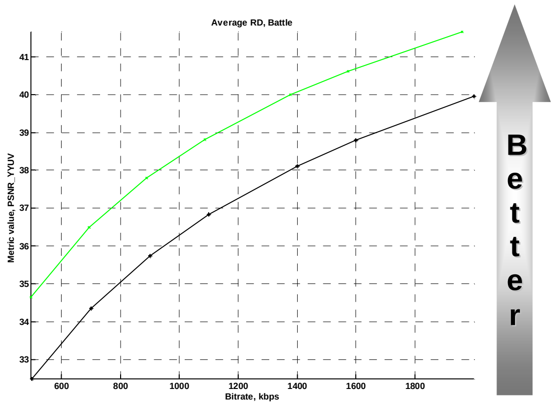
10.1.1.2 Relative Bitrate/Relative Time Charts

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

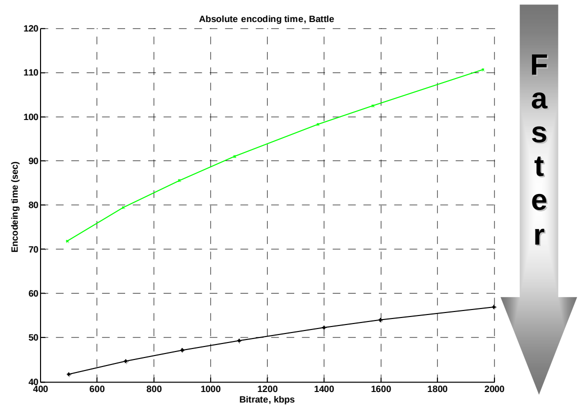
The X-axis shows the relative encoding time for the codec under test. Larger values indicate a slower codec. For example, a value of 2.5 means that the codec under test works 2.5 times slower, on average, than the reference codec.

10.1.1.3 Graph Example

Figure 77 shows a case where these graphs can be useful. In the top left graph, it is apparent that the “Green” codec encodes with significantly 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: it is clearly visible in the bottom graph that one of the codecs is slower, but yields higher visual quality, and that the other codec is faster, but yields lower visual quality.



RD curve. "Green" codec is better!



Encoding time (seconds). "Green codec is slower!

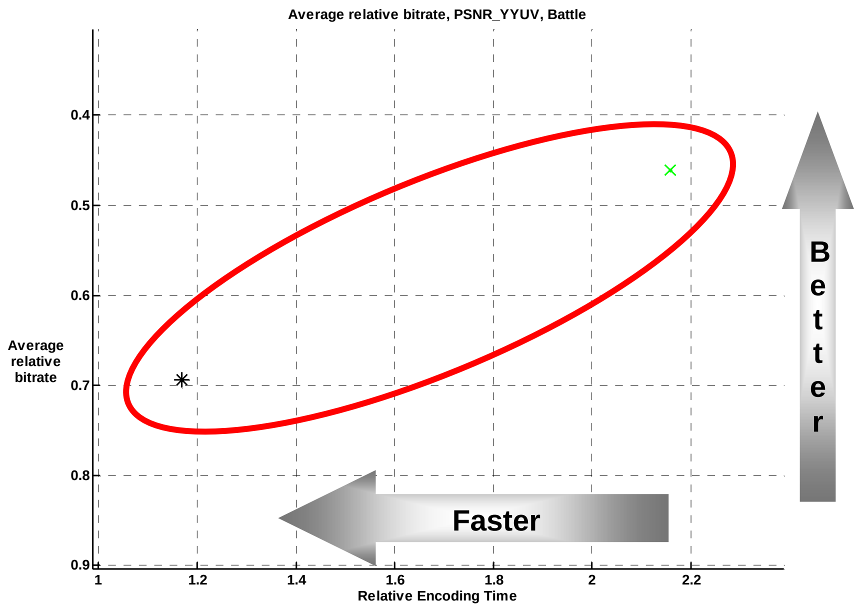


Figure 77. Integral situation with codecs. This plot shows the situation more clearly.

As a result of these advantages, relative bitrate/relative time graphs are used frequently in this report since they assist in the evaluation of the codecs in the test set, especially when number of codecs is large.

A more detailed description of the preparation of these graphs is given below.

10.2 Bitrates Ratio with the Same Quality

The first step in computing the average bitrate ratio for a fixed quality is inversion of the axes of the bitrate/quality graph (see Figure 79). All further computations are performed using the inverted graph.

The second step involves averaging the interval over which the quality axis is chosen. Averaging is performed only over those segments for which there are results for both codecs. This limitation is due to the difficulty of developing extrapolation methods for classic RD curves; nevertheless, for interpolation of RD curves, even linear methods are acceptable.

The final step is calculation of the area under the curves in the chosen interpolation segment and determination of their ratio (see Figure 80). This result is an average bitrate ratio for a fixed quality for the two codecs. If more than two codecs are considered, then one of them is defined as a reference codec and the quality of others is compared to that of the reference.

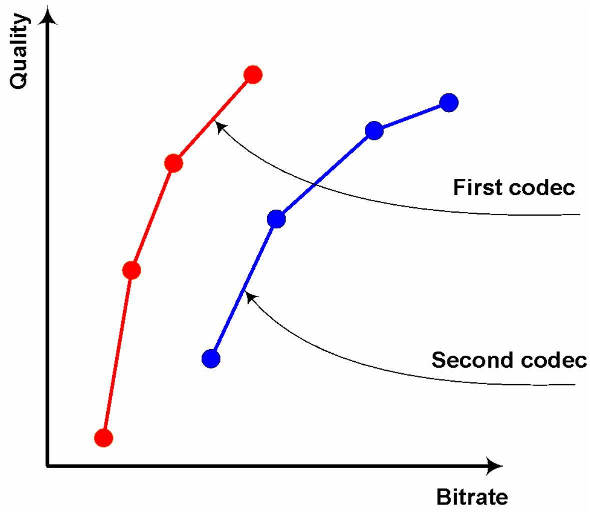


Figure 78. Source Data

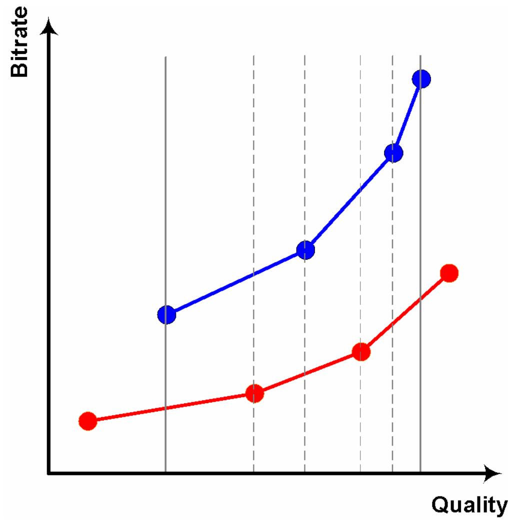


Figure 79. Axes' Inversion and Averaging Interval Choosing

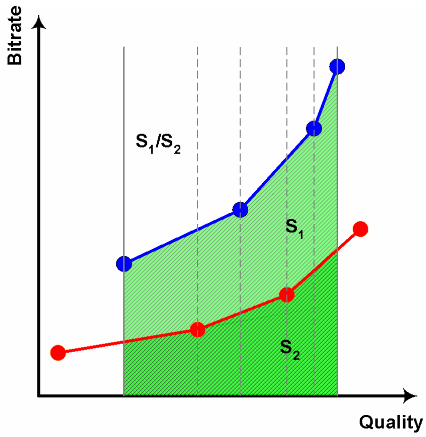


Figure 80. Areas' under Curves Ratio

10.3 Relative Codec Encoding Time Computation

To compute the relative processing time of two codecs for a particular video sequence, the encoding time is calculated for both codecs (the encoding times are summed for all bitrates) and the ratio is taken. For three or more codecs, one codec is chosen as a reference and the ratio of its encoding time to that of the others is calculated.

For multiple sequences, each codec is assigned an arithmetic mean of average relative encoding times for each sequence.

11 Appendix 7. Objective Quality Metrics Description

11.1 PSNR (Peak Signal-to-Noise Ratio)

11.1.1 Brief Description

This metric, which is often used in actual practice, is called the peak signal-to-noise ratio, or PSNR.

$$d(X,Y) = 10 \cdot \log_{10} \frac{255^2 \cdot m \cdot n}{\sum_{i=1, j=1}^{m,n} (x_{ij} - y_{ij})^2},$$

Where *d(X,Y)* – PSNR value between *X* and *Y* frames

x_{ij} – the pixel value for *(i,j)* position for the *X* frame

y_{ij} – the pixel value for *(i,j)* position for the *Y* frame

m,n – frame size *mxn*

Generally, this metric has the same form as the mean square error (MSE), but it is more convenient to use because of the logarithmic scale. It still has the same disadvantages as the MSE metric, however.

In MSU Video Quality Measurement Tool the PSNR can be calculated for all YUV and RGB components and for the L component of LUV color space. The PSNR value is quick and easy to calculate, but it is sometimes inappropriate as relates to human visual perception.

A maximum deviation of 255 is used for the PSNR for the RGB and YUV color components because, in YUV files, there is 1 byte for each color component. The maximum possible difference, therefore, is 255. For the LUV color space, the maximum deviation is 100.

The values of the PSNR in the LUV color space are in the range [0, 100]; the value 100 means that the frames are identical.

11.1.2 Examples

PSNR visualization uses different colors for better visual representation:

- Black – value is very small (99 – 100)
- Blue – value is small (35 – 99)
- Green – value is moderate (20 – 35)
- Yellow –value is high (17 – 20)
- Red –value is very high (0 – 17)

The following is an example of the PSNR metric:

Original Processed PSNR

Figure 81. PSNR example for two frames

The following are further examples demonstrating how various distortions can influence the PSNR value.

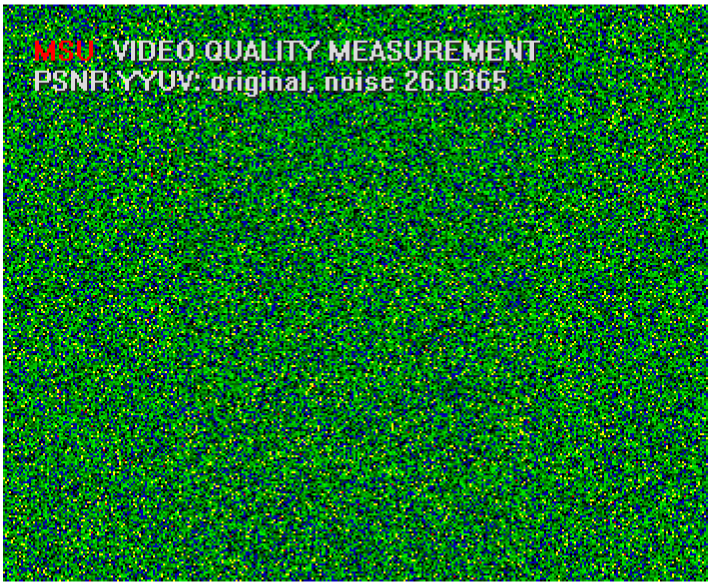
Original image Image with added noise

Blurred image Sharpen image

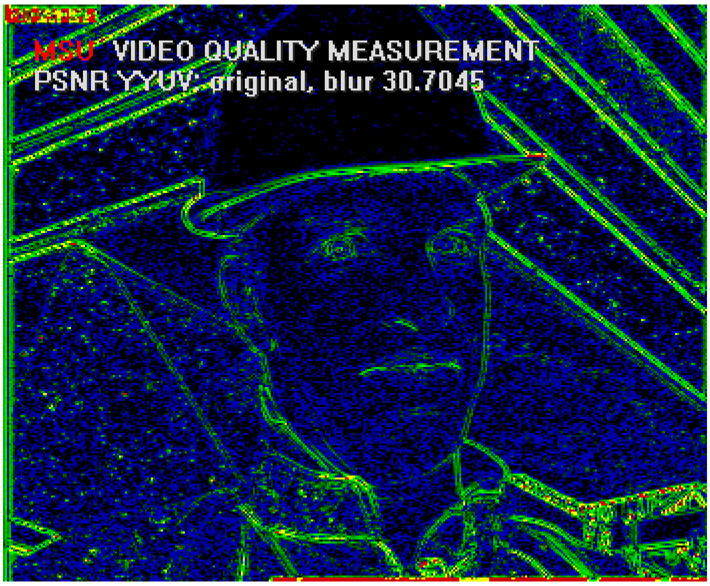
Figure 82. Original and processed images (for PSNR example)

Next are the PSNR values for the Y–plane for these images

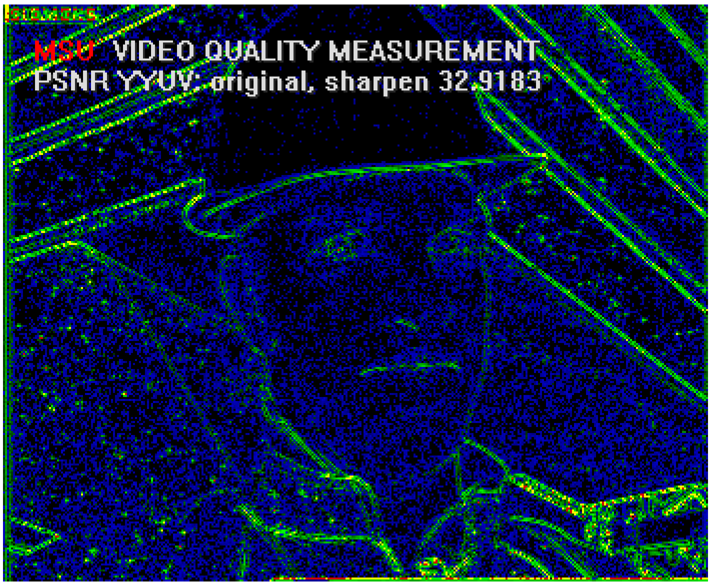
PSNR for image with itself, value = 0



PSNR for image with noisy image, value = 26.0365



PSNR for image with blurred image, value = 30.7045



PSNR for image with sharpen image, value = 32.9183

Figure 83. PSNR values for original and processed images

11.2 SSIM (Structural SIMilarity)

11.2.1 Brief Description

The original paper on the SSIM metric was published by Wang, et al.[1] The paper can be found at the following URL: http://ieeexplore.ieee.org/iel5/83/28667/01284395.pdf

The SSIM author homepage is found at the following URL: http://www.cns.nyu.edu/~lcv/ssim/

The scheme of SSIM calculation can be presented as follows. 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 the following:

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

where

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

$$\sigma_x = \left(\sum_{i=1}^{N} \omega_i (x_i - \mu_x) \right)^{\frac{1}{2}}$$

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

The constants C_1 and C_2 are defined according to the following expressions:

$$C_1=(K_1L)^2$$
$$C_2=(K_2L)^2$$

where L is the dynamic range of the pixel values (255 for 8-bit grayscale images), and K_1, K_2 << 1.

The values $K_1 = 0.01$ and $K_2 = 0.03$ were used for the comparison presented in this report, and the matrix filled with a value “1” in each position to form a filter for the result map.

For the implementation used in this comparison, 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 of the advantages of the SSIM metric is that it better represents human visual perception than does PSNR. SSIM is more complex, however, and takes more time to calculate.

[1] 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.

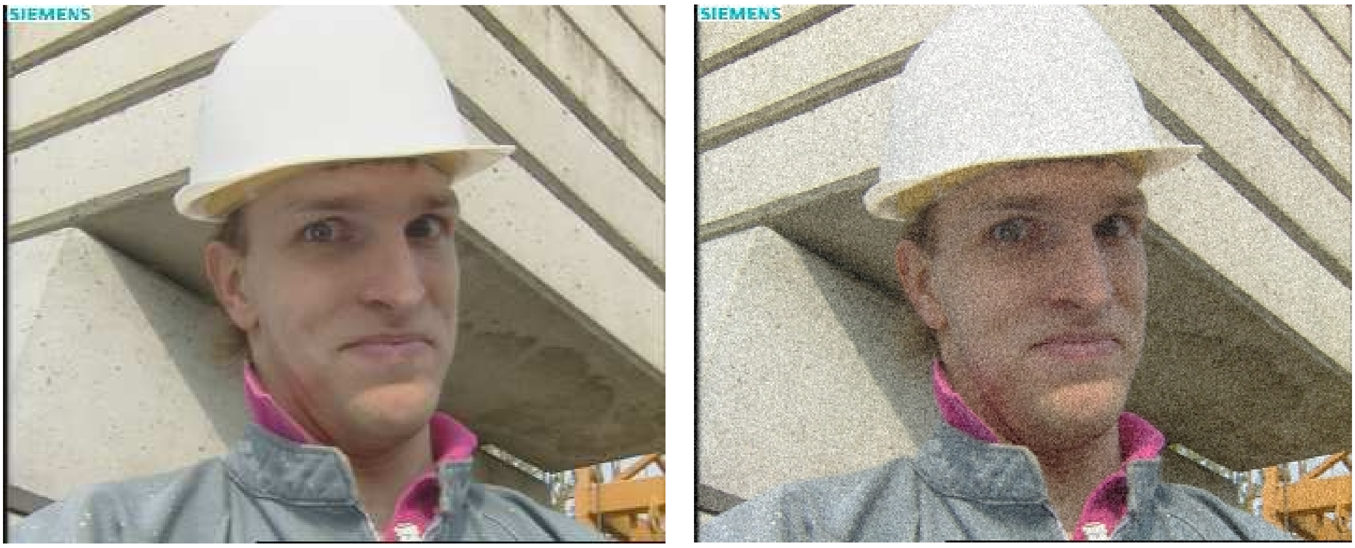
11.2.2 Examples

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

Original Processed SSIM

Figure 84. SSIM example for compressed image

The following are more examples how various types of distortion influence the SSIM value.



Original image Image with added noise

Blurred image Sharpen image

Figure 85. Original and processed images (for SSIM example)

The SSIM values for the Y-plane for these images are given below.

SSIM for image with itself, value = 1

SSIM for image with noisy image, value = 0.552119

SSIM for image with blurred image, value = 0.9225

SSIM for image with sharpen image, value = 0.958917

Figure 86. SSIM values for original and processed images

12 List of Figures

13 About the Graphics & Media Lab Video Group

GRAPHICS & MEDIA LAB
VIDEO GROUP

The Graphics & Media Lab Video Group is part of the Computer Science Department of 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.

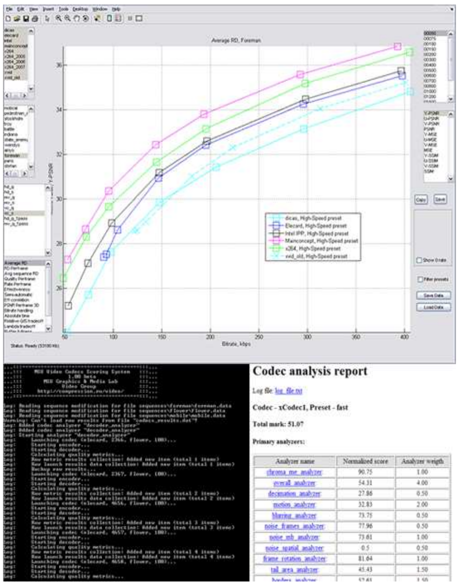
The Video Group has also worked for many years with companies like Intel, Samsung and RealNetworks.

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

This comparison was performed with ViCoS – Video Codec Scoring System

About the Video Codec Scoring System



ViCoS is a fully automatic quality evaluation system for video codecs and video processing algorithms.

It is an advanced system with client-server architecture and relational data base support. It allows robust codec launches with user-friendly interface and functions for video codec or video filter analysis with easy-to-use visualizations of results. With ViCoS you can:

1. **Perform QA with much lesser resources**
   ViCoS usage allows to do Quality Assurance tasks in a highly automatic way. Now video codec features or entire codec quality can be tested very easily without big number of QA specialists.
2. **Perform codec testing without subjective codec testing**
   ViCoS implements many different quality analyzers that can replace expensive subjective quality evaluation for almost every task.
3. **Fast comparison to competitors**
   ViCoS provides functionality for video codecs comparison. Now codec developers can compare their video codec quality to competitors very fast and easily.
4. **Choose optimal default and predefined parameters**
   ViCoS can help to choose optimal (speed/quality trade-off) encoding parameters using preset analysis subsystem.
5. **Compare different versions of a product easily**
   ViCoS helps to perform quick speed and quality comparison of different versions of a codec or video processing software.

And much more.

Main key features of the system:

1) *Client-server architecture*.
2) *Easy modifications* to add a new codec, preset or video sequence.
3) *Robust launches* – if a codec fails the system continues to work, marking the error for this codec
4) *DB usage* – all results can be saved in a data base (almost any relational data base management systems: MySQL, MSSQL, Oracle, etc.)
5) *Result visualization* – all obtained results can be visualized very quickly with user friendly-interface.
6) *Huge Amount of Data Processing* – during ViCoS work huge amount of data is produced, it is processed and categorized very easily and user friendly.
7) *Specific Analysis Types* – ViCoS uses specific types of analysis: well-known and specially developed (Edge capture, Borders quality, Tail area, Blurring, Synthetic motion, and more than 10 other types).

More information could be found at http://yuvsoft.com/technologies/vicos/

E-mail: vicos@yuvsoft.com

YUVsoft Corp. was born out of the research of the Moscow State University video group. The technologies and solutions offered by YUVsoft are based upon more than 10 years of experience in video codec analysis, video processing, image processing and multimedia compression.

MSU Video Quality Measurement Tool

MSU Graphics & Media Lab. Video Group.

GRAPHICS & MEDIA LAB
VIDEO GROUP

Main Features

1. 12 Objective Metric + 5 Plugins

PSNR several versions,
MSAD,
Delta,
MSE,
SSIM Fast,
SSIM Precise,
VQM,

MSU Blurring Metric,
MSU Brightness Flicking Metric,
MSU Brightness Independent PSNR,
MSU Drop Frame Metric,
MSU Noise Estimation Metric,
MSU Scene Change Detector,
MSU Blocking Metric.

2. More Than 30 Supported Formats, Extended Color Depth Support

*.AVI,
*. YUV:
- YUV,
- YV12,
- IYUV,
- UYVY,
- Y,
- YUY2,

*.BMP,

*.AVS:
- *.MOV,
- *.VOB,
- *.WMV,
- *.MP4,
- *.MPG,
- *.MKV,
- *.FLV,

etc.,

Extended Color Depth:
- P010, P014,
- P016, P210,
- P214, P216,
- P410, P414,
- P416,
- P410_RGB,
- P414_RGB,
- P416_RGB.

3. Multi-core Processors Support

MMX, SSE and OpenMP Optimizations

4. Comparative Analysis

Comparison of 3 files at a time

5. ROI Support

Metric calculation for ROI (Region of Interest)

6. GUI & Batch Processing

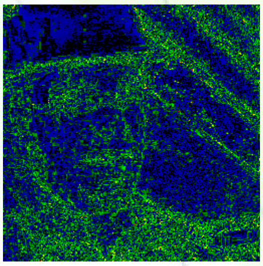
GUI and command line tools

7. Plugins Interface

You can easily develop your own metric

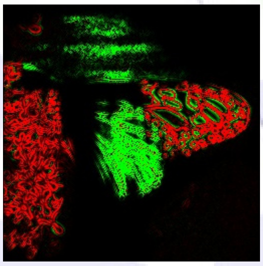
Visualization Examples

Allows easily detect where codec/filter fails



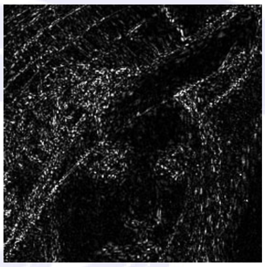
Y-YUV PSNR

Y-YUV Delta

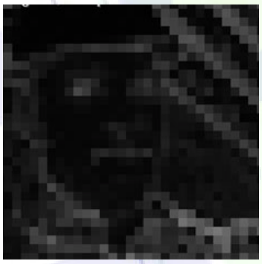


MSU Blurring Metric

MSU Blocking Metric



Y-YUV MSE



VQM

8. Universal Format of Results

Results are saved in *.csv files

9. HDTV Support

10. Open-Source Plugins Available

11. Metric Visualization

Fast problem analysis, see examples above.

http://www.compression.ru/video/quality_measure/index_en.html

Tool was downloaded more than 100 000 times!

Free and Professional versions are available

Big thanks to our contributors:

Apple Inc., intel, skype, nVIDIA, NASA, DOLBY, CISCO, Sun microsystems, NBC UNIVERSAL, DIVX