A VVC/H.266 Real-time Software Encoder for UHD Live Video Applications

Spin Digital | Whitepaper | August 2nd 2022 | Version 1.0

spin digital

Spin Digital Video Technologies GmbH
Helmholtzstraße 2-9. 10587 Berlin, Germany
www.spin-digital.com | info@spin-digital.com

TABLE OF CONTENTS

1. SUMMARY

Versatile Video Coding (VVC) is the newest video coding standard designed to significantly reduce the bitrate over its predecessor, High Efficiency Video Coding (HEVC), as well as to facilitate next-generation video applications.

Spin Digital has succeeded in developing a VVC real-time software encoder for Ultra HD (4K and 8K) live streaming and broadcasting. Its first version achieves 18% bitrate savings at equal quality compared to Spin Digital's HEVC real-time encoder, an industry-leading encoder for its high performance and compression efficiency.

The new VVC encoder can process 4Kp60 and 8Kp30 10-bit HDR videos in real-time using a single server with a state-of-the-art dual-socket CPU architecture. In the near future, through the use of advanced encoding algorithms and next-generation CPU architectures, new releases of the encoder will enable real-time 8Kp60 encoding and higher compression efficiency.

Using an objective quality criterion based on VMAF and informal subjective tests, the recommended range of bitrates for the target live applications have been determined. 4Kp60 live encoding at broadcast-grade quality can be achieved with 13 to 14 Mbps compared to 16 to 17 Mbps needed by Spin Digital's HEVC real-time encoder. For 8Kp60 live applications the recommended bitrate with the new VVC encoder is 40 Mbps instead of the 50 Mbps required by the optimized HEVC encoder.

Spin Digital's VVC encoder has been integrated into a live streaming framework that includes all the required components for live applications such as: input capture via SDI and TS over IP (TSoIP), pre-processing, pre-analysis, audio and video encoding, advanced rate control, and HTTP streaming and TSoIP broadcasting.

2. INTRODUCTION: TRENDS IN LIVE MEDIA APPLICATIONS

Video traffic accounts for about 80% of the total share of Internet data traffic and nearly 20% of Internet video traffic is already live (Cisco 2018). And it will continue to grow. OTT live and on-demand services are becoming increasingly popular, forcing traditional TV broadcasters to also use the open Internet to boost viewership (Careless 2021) (SES 2022). In addition to professional created content, User-Generated Content (UGC) is being live-streamed on platforms such as YouTube and Facebook, generating significantly higher engagement rates (Bern 2019). Other live applications including e-gaming and video surveillance are also experiencing huge Internet traffic growth (Cisco 2018). The Covid-19 pandemic has also changed the way in which people are interacting in private and professional life: there are more and more video conferencing sessions as well as more live events online. And these trends do not seem to change in the post-pandemic scenario (Wyman 2021).

The next generation of live applications is primarily focused on delivering higher quality video as well as immersive experiences that engage the end user more with the content. The 4K Ultra-High Definition (UHD) High Dynamic Range (HDR) format is becoming mainstream, 8K UHD streaming and broadcasting is already a reality, and 8K 360-/180-degree video is an accepted format for high-end live Virtual Reality (VR) streaming applications.

For these emerging live services, High Efficiency Video Coding (HEVC) represents the state-of-the-art codec for compressing video with lower data rates without sacrificing quality (Spin Digital 2020, Ultra HD Forum 2021). However, HEVC is reaching its maximum compression efficiency capacity. As a result, next-generation codecs are needed to make these enhanced applications more accessible.

Versatile Video Coding (VVC), or H.266, is the latest video coding standard that not only provides significantly higher compression efficiency than HEVC, but is also designed to enable efficient coding of different types of video, such as UHDTV (4K and 8K), VR/360-degree, and screen content. VVC also includes tools for more efficient encoding for adaptive bitrate streaming and scalable encoding. VVC coding solutions, especially for live streaming and broadcasting, will be essential in the near future to facilitate the development of next-generation multimedia applications.

3. THE VVC/H.266 STANDARD AND ITS APPLICATIONS

The VVC/H.266 standard was completed in July 2020 by the Joint Video Experts Team (JVET) of ISO and ITU (Bross et al. 2020). VVC represents the state-of-the-art in video coding providing 50% bitrate reduction with respect to its predecessor, HEVC/H.265, for the same visual quality. This level of compression gain is only possible at the expense of significant computational increase.

In the currently fragmented codec landscape, VVC competes with other coding standards including AV1 by the Alliance for Open Media (AOM), Low Complexity Enhancement Video Coding (LCEVC) and Essential Video Coding (EVC) by MPEG. Independent evaluations have shown that VVC can achieve higher compression efficiency than these other standards: e.g. 19.5% to 20.5% better than AV1 and 27.3% to 30.5% better than EVC for UHD video (Grois et al. 2021).

VVC not only produces higher compression than other standards, but is also designed to achieve more efficient coding of different types of content (footage, screen content, and 360-degree video) and new-generation video formats including UHD 4K and 8K with HDR and wide color gamut.

3.1. VVC TARGET APPLICATIONS

The VVC standard has been designed to make possible next-generation video applications. With VVC high-end 8K live and on-demand services can become more popular, since high-quality 8K content can be transmitted over bandwidth-constrained networks such as terrestrial broadcasting and streaming over the Internet. 4K HDR premium content can also be delivered to more users and with lower streaming costs. Even Full HD can become the new standard resolution with high quality and very low bitrates.

In addition, VVC includes new features to make Adaptive Bitrate (ABR) streaming services more efficient in compression, as well as to enhance the immersive experience of VR events in 360/180-degree formats. VVC also includes scalability, a coding technique to compress video in multiple layers, each layer representing a different resolution or quality of the same video. Thus, video services can be deployed with scalability enabled in order to support networks, receivers and display devices with different capabilities. Unlike the scalable extensions of other codecs, this feature has been simplified and made more practical for easier deployment.

3.2. VVC FOR LIVE

Currently, HEVC is the state-of-the-art video codec for UHD live applications. Existing codec implementations already achieve the performance required to transmit 4K and 8K video live over the Internet as well as broadcast channels. Using latest-generation HEVC real-time encoders, the recommended bitrates to produce broadcast-grade quality 4Kp60 video range from 15 to 18 Mbps (Schwarz 2022) and for 8Kp60 from 48 to 80 Mbps (Schwarz 2 2022). However, these encoding solutions have reached a point of diminishing returns for HEVC under real-time conditions, meaning that adding more computation will result in marginal compression gains.

In order to reduce the bandwidth beyond the capabilities of HEVC and, in this way, make UHD live services more accessible to the public, a new-generation codec is needed as it includes more efficient coding tools and features. Figure 1 shows a comparison of state-of-the-art performance optimized encoding implementations of AVC, HEVC, AV1, and VVC for 4K-UHD video in terms of compression efficiency and encoding complexity. Compression efficiency is measured as bitrate increase at equal quality referred to a baseline encoder, and encoding complexity is calculated as CPU utilization time relative to that of a baseline encoder. These implementations generate different trade-offs between compression efficiency and complexity, adding more computation resulting in higher compression efficiency (lower bitrate for the same quality). With a bitrate increase of 0.0% and CPU time of 1.0x, the baseline encoder is based on HEVC with a configuration that represents the best operational point for real-time applications.

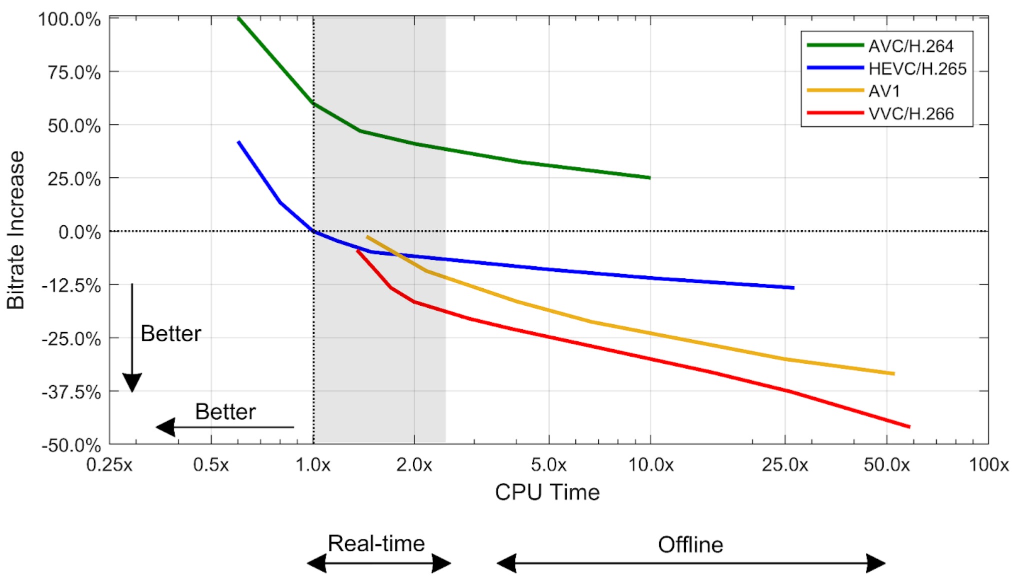


Figure 1: Video codecs (AVC, HEVC, AV1, and VVC) compression efficiency capabilities and encoding complexity. When a video codec reaches a point of diminishing returns the next generation codec offers the path for further compression gains.

The figure illustrates how different generations of optimized encoders evolve from low complexity (i.e. AVC) to very high complexity (i.e. VVC) and how they can translate this complexity into compression efficiency. The baseline HEVC encoder is able to process 4Kp60 video in real time, and with the same computational complexity has a much higher compression efficiency than AVC/H.264. But beyond a relative complexity of 1.5x HEVC and AVC are no longer able to significantly increase compression efficiency. In contrast, AV1 and especially VVC implementations are able to significantly overcome the compression limits of HEVC at very high complexity. Finally, from a complexity higher than 25.0x, it is observed that VVC increases its compression efficiency advantage compared to AV1.

3.3. VVC ENCODERS: VOD VS LIVE

As seen in Figure 1, current performance optimized encoder implementations cover a wide range of encoding complexity, but it will be the requirements of the application that determine in which region the encoder should operate. Thus, two types of encoders can be distinguished: *offline encoders* and *real-time encoders*. On the one hand, this first category is addressed to VoD applications or post-production workflows where the main requirement is to reduce the bitrate as much as possible and, at the same time, achieve high video quality, even if this results in long encoding times. Figure 1 shows that an optimized VVC encoder is able to achieve up to 50% bitrate reduction compared to the baseline HEVC live encoder, but at the cost of 50 times the computational complexity.

On the other hand, the second category of encoders is used in streaming and broadcast applications that require real-time video encoding. The main challenge for these encoders is, therefore, to achieve the highest possible compression efficiency under real-time conditions. According to the results shown in Figure 1, the relative complexity region in which an optimized encoder can potentially compress UHD video (e.g. 4K at 60 fps) in real-time with current CPU-based computing architectures is between 1.0x to 2.5x the complexity of a highly optimized HEVC encoder. These estimations indicate that an optimized VVC live encoder can achieve around 20.0% bitrate reduction at the cost of 2.5x more computation compared to the HEVC baseline. AV1, at the same complexity level (2.5x compared to the HEVC baseline), has a compression efficiency between HEVC and VVC, with a maximum of 12.0% bitrate savings.

4. SPIN DIGITAL'S VVC REAL-TIME ENCODER

Spin Digital has developed the first version of a high-quality and performance-optimized VVC software encoder for UHD HDR live broadcast and streaming. The encoder has been extensively optimized for latest-generation CPU architectures in order to achieve the performance and compression levels required for 4K and 8K video, with optimizations including advanced coding-tool decision algorithms, SIMD processing, memory optimizations, and a multi-level parallelization approach that is able to scale to systems with large numbers of CPU cores. With all these optimizations together, the new encoder enables real-time encoding of 4K at 60 fps and 8K at 30 fps video in 10-bit HDR using a dual-socket server with a total of 76 CPU cores.

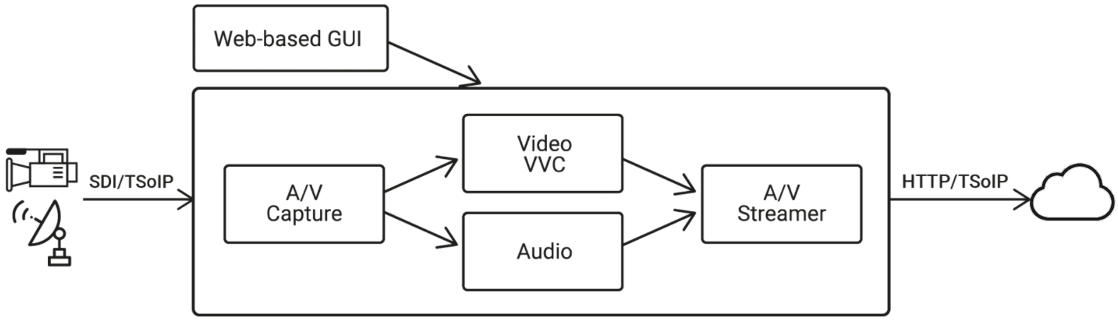


Figure 2: Key components of Spin Digital's VVC live encoder diagram

In addition to having developed an optimized core encoding module, the VVC encoder has been integrated into a complete live streaming framework that includes (see Figure 2): input capture with both SDI and IP interfaces, pre-processing (scaling, color conversion, tone mapping), pre-analysis, advanced rate control, core VVC encoding, audio encoding, and streaming for HTTP (HLS, DASH) or TSoIP delivery (UDP, RTP, SRT, RIST Zixi).

5. ENCODER ASSESSMENT FOR 4K-UHD STREAMING AND BROADCASTING

The new Spin Digital VVC real-time encoder was assessed in terms of compression efficiency, encoding complexity, and multithreaded encoding speed. The results were also compared to five open-source optimized software encoder implementations of different coding standards including H.264/AVC, H.265/HEVC, AV1, and VVC/H.266.

The encoders were configured assuming a **4K-UHD live streaming and broadcasting scenario**. This use case mainly requires that the rate control is enabled and that long Group-of-Picture (GOP) structures are used for maximum compression efficiency with frequent random access points (e.g. 1 to 3 seconds intra period).

A wide set of test video sequences have been selected that are representative of 4K-UHDTV content as well as long enough to stabilize the rate control.

5.1. VIDEO SEQUENCES

A total of 11 1-minute 4K video sequences representing the target use case were selected in these experiments. The test set includes camera footage, animation and videos with visual effects in order to analyze the quality impact of the encoders on heterogeneous textures and motion.

The video sequences were acquired from different providers including: Netflix (Xiph 2015, Netflix 2022), Poznan Supercomputing and Networking Center (PSNC) (Immersify 2018), Fraunhofer HHI and InterDigital (Fraunhofer HHI, InterDigital 2022), and Unigine (Unigine 2017). The master files were preprocessed using Spin Digital's high-precision filters to generate a 4K-UHD broadcast distribution format, which is specified in the table below. The preprocessing operations applied to each sequence are described in Appendix.

Table 1: Technical specifications of the 4K-UHD distribution format

Parameters	4K-UHD distribution format
Resolution	3840x2160 pixels
Frame rate	24, 59.94, 60 fps
Chroma sampling	4:2:0
Bit-depth	10 bits
Transfer function and color gamut	SDR: SDR, BT.709 HDR: PQ, BT.2020

Table 2 presents detailed information of the test sequences: producer, type of content, HDR and color gamut formats, and Spatial Information (SI) and Temporal Information (TI) (ITU-R 2019). As can be observed, seven clips are in High Dynamic Range (HDR) PQ with a BT.2020 color gamut, whereas the remaining four are in SDR BT.709. The obtained SI and TI values cover a wide spectrum of complexities, from low (Meridian, RollerCoaster) to high (SolLevante, TunnelFlag) SI and TI, demonstrating the heterogeneity in complexity of the selected sequences (see Figure 3).

Table 2: Technical information of the test sequences: producer, type of content, format, SI, TI

	Producer	Type	Format	SI	TI
BasketballGame	Netflix	Footage	4Kp59.94 HDR	112.6 (med)	152.2 (med)
BerlinSeqs	Fraunh. HHI & InterDigital	Footage	4Kp60 HDR	172.2 (med)	63.6 (low)
DrivingPOV	Netflix	Footage	4Kp59.94 HDR	150.0 (med)	147.4 (med)
FollowCar	PSNC	Footage	4Kp59.94 SDR	201.9 (high)	111.8 (med)
MC2	PSNC	Footage	4Kp59.94 SDR	280.5 (high)	82.5 (low)
Meridian	Netflix	Footage & CGI	4Kp59.94 HDR	89.8 (low)	23.6 (low)
RollerCoaster	Netflix	Footage	4Kp59.94 HDR	92.3 (low)	84.9 (low)
SolLevante	Netflix	Animation	4Kp24 HDR	237.5 (high)	269.4 (high)
Superposition	Unigine	CGI	4Kp60 SDR	226.5 (high)	82.9 (low)
ToddlerFountain	Netflix	Footage	4Kp59.94 HDR	170.2 (med)	79.3 (low)
TunnelFlag	Netflix	Footage	4Kp59.94 SDR	211.6 (high)	206.8 (high)

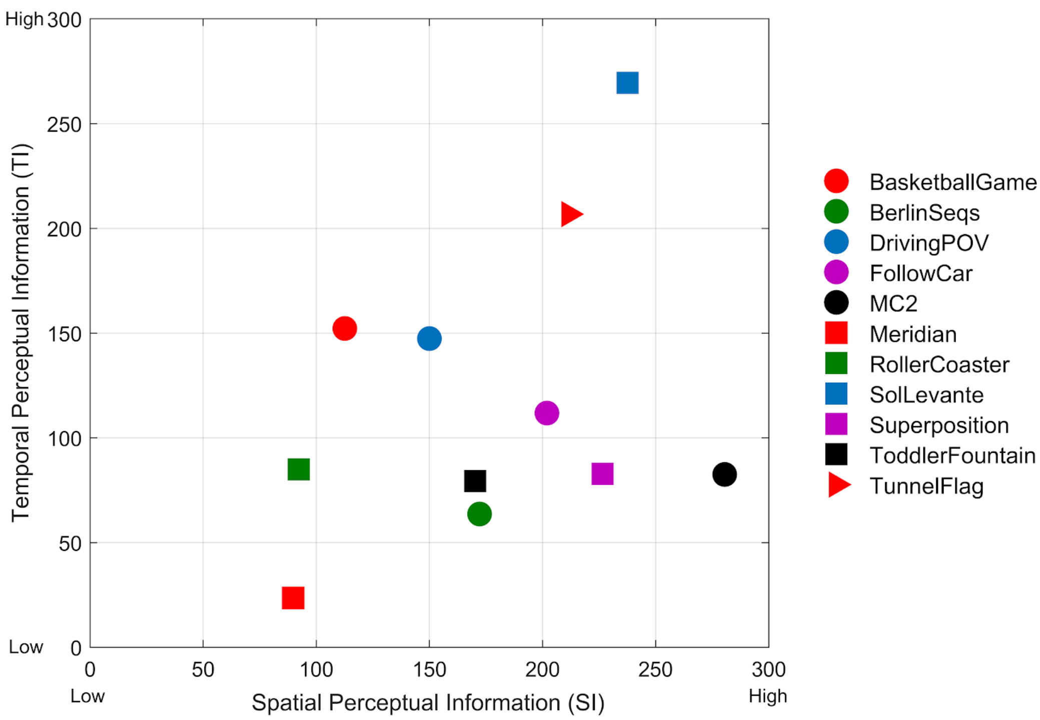


Figure 3: Spatial and Temporal Perceptual Information of the test sequences

5.2. VIDEO ENCODERS

Table 3 shows, for each encoder, the standard on which it is based, the company or institution that develops it, and the version and release date used in the experiments.

Table 3: Coding standard, company, version, and release date for each encoder under assessment

	x264	x265	SVT-HEVC	Spin Digital HEVC	SVT-AV1	VVenC	Spin Digital VVC
Standard	AVC	HEVC	HEVC	HEVC	AV1	VVC	VVC
Company	VideoLAN	Multicore Ware	Intel	Spin Digital	Intel & Netflix	Fraunhofer HHI	Spin Digital
Version	r3075	3.5	1.5.1	1.2	1.0.0	1.3.1	0.1
Release date	Sep. 2021	Mar. 2021	June 2021	June 2022	Apr. 2022	Dec. 2021	June 2022

The table specifies one AVC/H.264 encoder (x264), three implementations of the HEVC standard (x265, SVT-HEVC, Spin Digital HEVC), one AV1 encoder (SVT-AV1), and two VVC encoders (VVenC and Spin Digital VVC). Spin Digital's HEVC encoder is part of a commercially available product called Spin Enc Live (Spin Digital 2022). For the sake of clarity, from now on this encoder will be called Spin Digital HEVC.

5.3. ENCODING SETTINGS

The video encoders were configured as similarly as possible assuming the 4K-UHD broadcast scenario. This use case typically requires random-access encoding mode (long GOP), open GOP, 1-second intra period, and Constant Bit Rate (CBR) rate control with a 1-second buffer size for the Hypothetical Reference Decoder (HRD) model. Some exceptions to this typical configuration are: the Variable Bit Rate (VBR) algorithm of SVT-AV1 was selected, since the CBR method is not supported in random-access mode; and the rate control mechanism of VVenC does not support the HRD constraint.

In addition, the encoders were tuned to maximize the Peak Signal-to-Noise Ratio (PSNR). Other parameters, such as GOP size, GOP structure, and lookahead window were kept at their default values.

The target bitrates were selected based on the recommendations given by the Ultra HD Forum for 4K HEVC final distribution (Ultra HD Forum 2021). In particular, for compression efficiency (BD-rate) analysis a range between 8 and 44 Mbps in steps of 4 Mbps was used. This range is wide enough to accurately compare different generations of encoders.

For each encoder several presets were used in order to analyze different tradeoffs between quality and speed. Spin Digital's HEVC and VVC encoders were evaluated only with the recommended preset for the intended use case.

For the sake of a fair comparison, the number of encoding threads was set to 8 in all encoders. An analysis of maximum encoding speed when using a high-performance multi-core CPU is presented later in this report.

5.4. COMPARISON METRICS

The video encoders were compared from the compression efficiency, encoding complexity, and performance points of view using BD-rate, CPU time and the encoding speed in frames per second, respectively.

Compression efficiency: BD-rate

The Bjontegaard Delta (BD)-rate method (Bjøntegaard 2001, Bjøntegaard 2008) was used to compute compression efficiency. Its goal is to compute the average bitrate increase produced by a test encoder referred to a baseline encoder at the same quality

Spin Digital's HEVC real-time encoder - Spin Digital HEVC - was selected as the baseline encoder. Four quality metrics used in the experiments were: PSNR, Perceptually Weighted PSNR (XPSNR) (Helmrich et al. 2020, Helmrich 2021), Multi-Scale Structural Similarity (MS-SSIM) (Wang, Simoncelli, and Bovik 2003), and Video Multi-method Assessment Function (VMAF) (Netflix 2021). The PSNR, XPSNR, and MS-SSIM metrics were calculated using the luma and chroma components.

Encoding complexity: CPU time

Encoding complexity was measured in terms of average CPU utilization time (including both user-level and system-level CPU utilization) over the target bitrates during the encoding process, relative to a reference encoder. Note that the CPU time is the accumulated time across all cores of the CPU, and therefore it can also be seen as *single-threaded encoding time*. Spin Digital HEVC has been used as a reference.

The platform used to run the encodings and quality metrics for BD-rate and CPU time computation is a server with the following components:

- CPU: 4x Intel Xeon Platinum 8176 CPU @ 2.10GHz (4x 28 cores)
- DRAM: 24x 16 GB DDR4 2666 MHz
- OS: Ubuntu 20.04

Encoding jobs were executed in parallel using the parallel job scheduling framework called *parallel* (Tange 2018).

Maximum performance: multithreaded encoding speed

The maximum performance of the encoders was measured in terms of multithreaded encoding speed. This measurement allowed us to determine the maximum number of frames per second produced by the encoders running on a latest-generation multi-core platform.

In order to achieve maximum performance it is not only important that the encoder is highly optimized for single-threading, but also that the multithreading parallelism is well designed and implemented in order to use the platform resources efficiently. To measure this ability we have also recorded the CPU utilization during the encoding process.

To measure encoding speed, the value provided by the encoding applications was taken instead of using the elapsed time facilitated by the linux *time* command, as the former measures the performance devoted to encoding only, excluding the time consumed for memory allocation and other initialization steps at the beginning of the encoding.

The server specifications used to measure the speed performance of the encoders are listed below:

- CPU: Intel Xeon Platinum 8368@ 2.4 GHz (2x 38 cores)
- DRAM: 16x 16 GB DDR4 3200 MHz
- OS: Red Hat 8.5

5.5. COMPRESSION EFFICIENCY AND ENCODING COMPLEXITY

Table 4 and Figures 4 to 7 show the complete results produced by the encoders under study in terms of BD-rate based on PSNR, XPSNR, MS-SSIM, and VMAF, and CPU time, in which the reference is Spin Digital HEVC.

Table 4: Results in terms of BD-rate based on different quality metrics (PSNR, XPSNR, MS-SSIM, VMAF) and CPU time relative to Spin Digital HEVC for the encoders and presets under evaluation. The bitrate range for BD-rate calculation is from 8 Mbps to 44 Mbps

Encoder - preset	PSNR BD-rate [%]	XPSNR BD-rate [%]	MS-SSIM BD-rate [%]	VMAF BD-rate [%]	CPU Time [times]
x264 r3075 - slower	32.44	53.24	47.13	16.35	4.08
x264 r3075 - med	46.00	69.80	61.59	22.91	1.37
x264 r3075- vfast	100.29	156.16	112.24	80.48	0.73
x265 v3.5 - slower	-13.32	-10.66	-6.75	-23.65	25.55
x265 v3.5 - medium	4.08	7.21	10.67	4.71	2.95
x265 v3.5 - ultrafast	52.08	57.46	53.38	43.82	0.86
SVT-HEVC v1.5.1 - 3	15.05	16.05	25.91	34.64	5.83
SVT-HEVC v1.5.1 - 5	32.77	33.64	54.95	54.71	1.94
SVT-HEVC v1.5.1 - 7	43.88	44.21	60.92	71.12	1.08
SVT-HEVC v1.5.1 - 9	66.30	66.78	77.34	102.76	0.76
SVT-HEVC v1.5.1 - 11	100.69	103.19	101.52	134.83	0.60
SVT-AV1 v1.0.0 - 4	-31.50	-33.51	-30.58	-32.46	24.70
SVT-AV1 v1.0.0 - 6	-23.78	-25.47	-23.74	-28.92	6.65
SVT-AV1 v1.0.0 - 7	-16.00	-17.26	-16.01	-20.52	3.99
SVT-AV1 v1.0.0 - 8	-9.36	-10.61	-11.32	-4.68	2.16
SVT-AV1 v1.0.0 - 9	-1.24	-3.72	-4.12	-0.08	1.43
SVT-AV1 v1.0.0 - 10	16.28	16.94	9.36	14.72	0.87
SVT-AV1 v1.0.0 - 12	44.58	46.73	32.26	30.36	0.57
VVenC v1.3.1 - fast	-37.73	-41.54	-32.87	-36.85	26.00
VVenC v1.3.1 - faster	-31.14	-34.22	-24.48	-27.78	15.12
Spin Digital HEVC v1.2	0.00	0.00	0.00	0.00	1.00
Spin Digital VVC v0.1	-16.57	-17.01	-17.37	-18.11	1.98

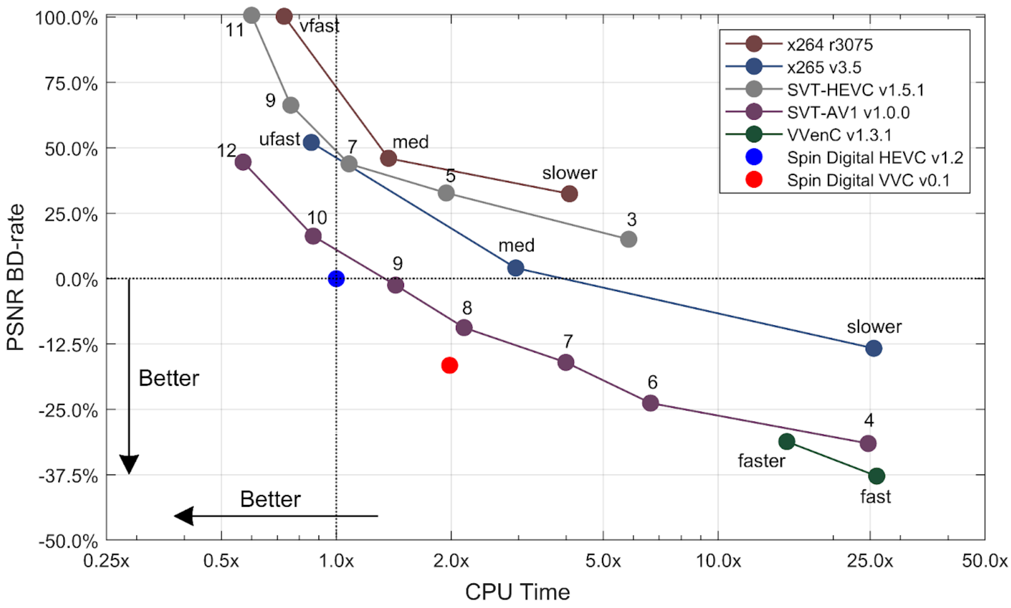


Figure 4: PSNR BD-Rate and CPU utilization time relative to Spin Digital HEVC for the encoders and presets under evaluation

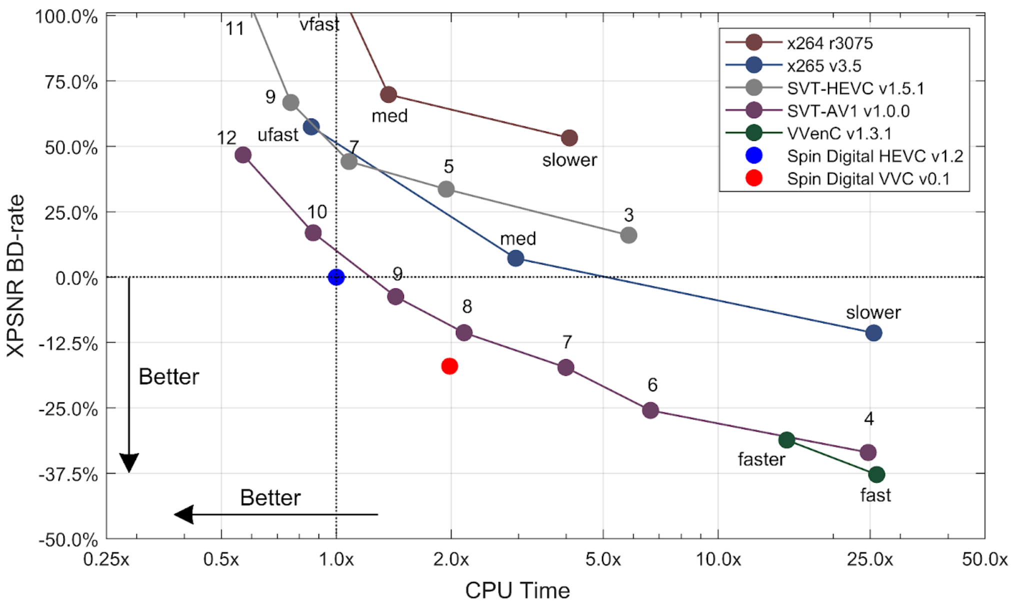


Figure 5: XPSNR BD-Rate and CPU utilization time relative to Spin Digital HEVC for the encoders and presets under evaluation

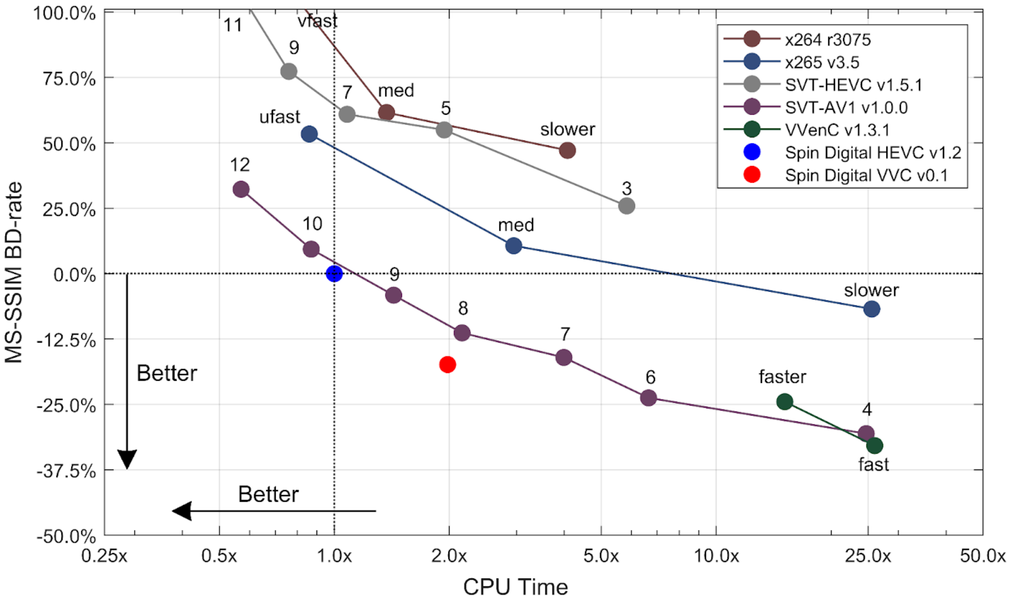


Figure 6: MS-SSIM BD-Rate and CPU utilization time relative to Spin Digital HEVC for the encoders and presets under evaluation

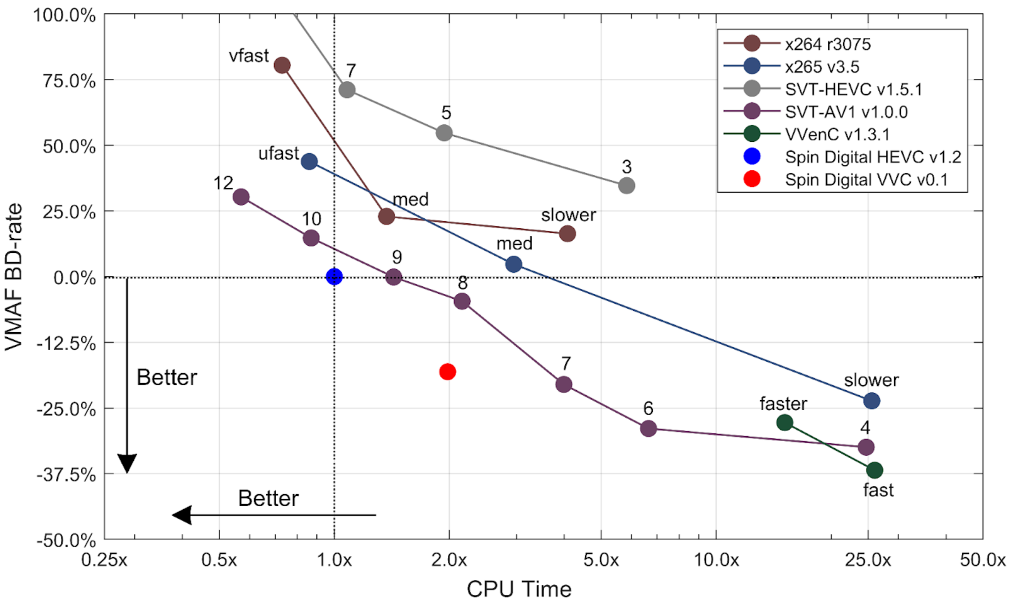


Figure 7: VMAF BD-Rate and CPU utilization time relative to Spin Digital HEVC for the encoders and presets under evaluation

When compared to an optimized HEVC encoder, Spin Digital's VVC encoder produces bitrate savings of 18% at twice the computational complexity and, at the same complexity, outperforms SVT-AV1 (preset 8).

As can be seen in the table and figures, at a CPU utilization time of 2.0x, Spin Digital VVC produces the highest compression efficiency of all encoders, with bitrate savings over Spin Digital HEVC ranging from 16.57% to 18.11%, depending on the quality metric used. With similar complexity, Spin Digital VVC achieves higher BD-rate gains than SVT-AV1 - 8, for example: -16.57% vs -9.36% (PSNR) or -18.11% vs -4.68% (VMAF).

Moreover, it is observed that the bitrate savings achieved by Spin Digital VVC are comparable to those of SVT-AV1 - 7 and x265 - slower, but those encoding modes require 2.0 times and 12.5 times more computation, respectively.

At a comparable compression efficiency, Spin Digital VVC is half as complex as SVT-AV1 (preset 7) and significantly less complex than x265 (preset slower).

Table 5 reports the results in terms of BD-rate and CPU time for a narrower range of bitrates between 6 Mbps and 25 Mbps, which can be considered more realistic for 4K VVC or AV1 video. For the sake of simplicity, the results correspond to Spin Digital's HEVC and VVC encoders, SVT-AV1 - 8, and VVenC - faster.

Table 5: Results in terms of BD-rate based on different quality metrics (PSNR, XPSNR, MS-SSIM, VMAF) and CPU time relative to Spin Digital HEVC for selected encoders and presets under evaluation. The bitrate range for BD-rate calculation is 6 Mbps to 25 Mbps

Encoder - preset	PSNR BD-rate [%]	XPSNR BD-rate [%]	MS-SSIM BD-rate [%]	VMAF BD-rate [%]	CPU Time [times]
SVT-AV1 v1.0.0 - 8	-11.07	-13.13	-13.62	-7.99	2.22
VVenC v1.3.1 - faster	-31.67	-34.24	-27.50	-29.38	14.65
Spin Digital HEVC v1.2	0.00	0.00	0.00	0.00	1.00
Spin Digital VVC v0.1	-16.33	-16.70	-17.89	-16.58	1.87

As can be seen, the results for Spin Digital VVC when using a range between 6 Mbps and 25 Mbps are similar to the results obtained using a larger range of bitrates (8 Mbps to 44 Mbps), which was needed for comparing multiple codec generations. Although slightly higher BD-rate gains are observed for SVT-AV1 - 8, and also for VVenC - faster, in this narrower range, Spin Digital VVC still achieves the highest compression gain at a comparable CPU time.

5.6. COMPRESSION EFFICIENCY PER SEQUENCE

Table 6 provides more detailed information about the BD-rate savings obtained by the new VVC encoder when compared to the HEVC baseline. According to these results, Spin Digital's VVC encoder achieves an average bitrate reduction of 16.57% for the same PSNR, ranging from 11.96% to 26.97%, for twice the complexity of Spin Digital HEVC. If the VMAF metric is used, the VVC encoder achieves an 18% bitrate reduction, going from 8.88% to 33.65%. Other metrics exhibit similar compression gains.

Table 6: BD-rate results for each video sequence achieved by Spin Digital VVC referred to Spin Digital HEVC

Video Sequence	PSNR BD-rate [%]	XPSNR BD-rate [%]	MS-SSIM BD-rate [%]	VMAF BD-rate [%]
BasketballGame	-15.89	-16.40	-14.58	-15.42
BerlinSeqs	-11.96	-12.06	-12.28	-15.67
DrivingPOV	-17.83	-18.58	-17.91	-18.70
FollowCar	-26.97	-26.83	-26.70	-24.08
MC2	-12.22	-13.29	-15.84	-11.97
Meridian	-12.66	-13.62	-14.73	-18.75
RollerCoaster	-16.96	-17.56	-15.78	-16.85
SolLevante	-13.80	-14.75	-14.62	-13.90
Superposition	-18.62	-18.91	-19.51	-33.65
ToddlerFountain	-12.48	-12.42	-13.74	-8.88
TunnelFlag	-21.55	-21.41	-24.10	-18.47
Average	-16.57	-17.01	-17.37	-18.11

5.7. SELECTED QUALITY-BITRATE PLOTS

Unlike the BD-rate metric, which gives a single number representing the compression efficiency of an encoder in average over a range of bitrates, the quality-bitrate (rate-distortion) curves provide a better insight of the compression efficiency of the encoders for different bitrates.

The following figures show the quality-bitrate curves based on PSNR, XPSNR, MS-SSIM, and VMAF for the *DrivingPOV* sequence. In order to reduce the number of encoders and presets displayed, only those that produce comparable complexities around 2.0x (from 1.0x to 3.0x) were included: x264 - medium, x265 - medium, SVT-HEVC - 5, SVT-AV1 - 8, Spin Digital HEVC, and Spin Digital VVC. Although VVenC has no preset falling in the specified range of complexity, VVenC - faster has also been included as a reference of VVC's potential in terms of compression efficiency.

As can be observed, VVenC - faster achieves in general the highest quality at equal bitrate –but at the cost of 15 times higher complexity– followed by Spin Digital VVC and then by SVT-AV1 - 8.

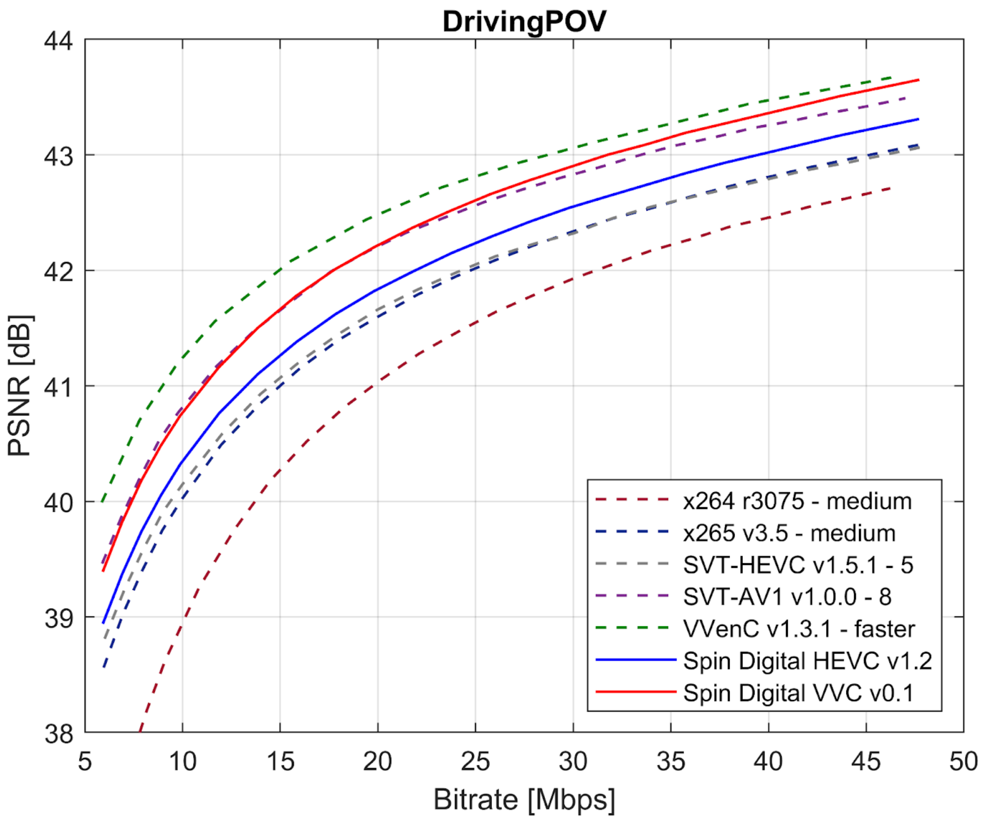


Figure 8: PSNR-bitrate curves for DrivingPOV

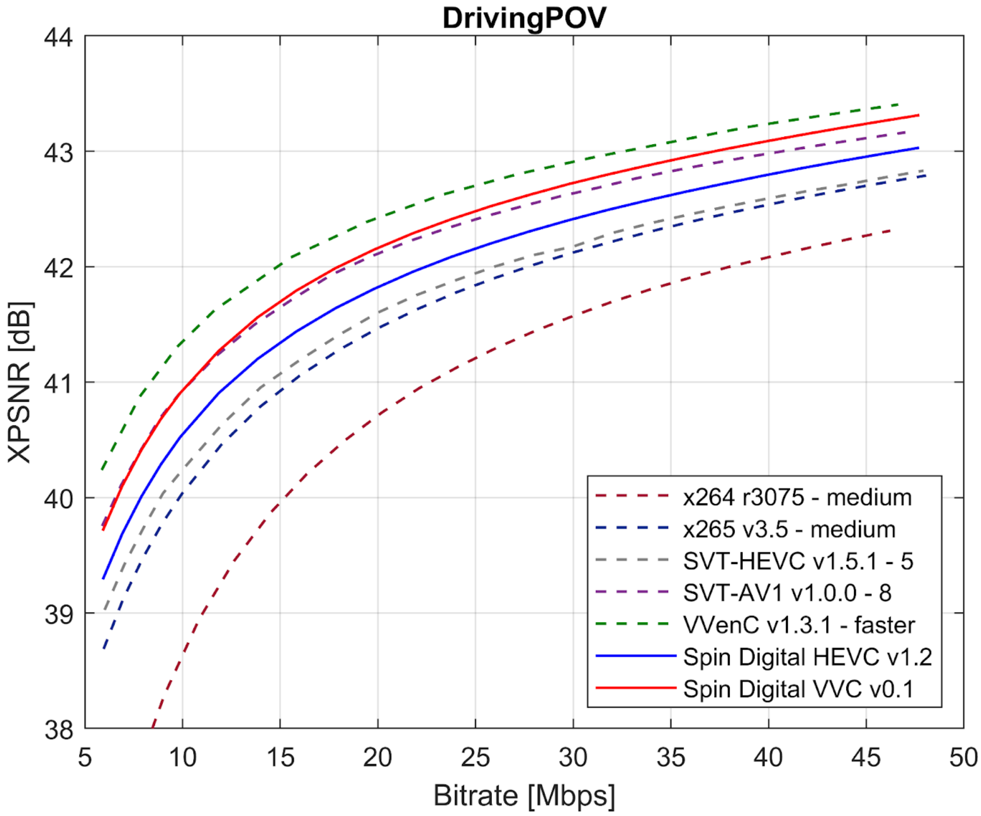


Figure 9: XPSNR-bitrate curves for DrivingPOV

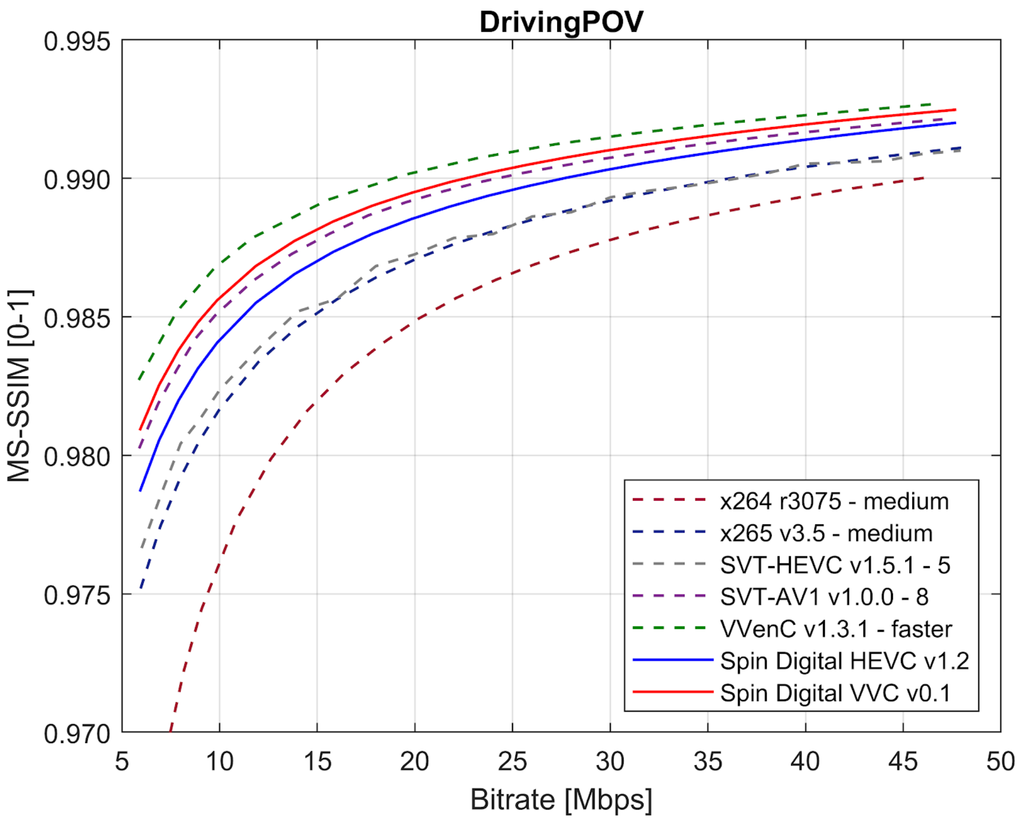


Figure 10: MS-SSIM-bitrate curves for DrivingPOV

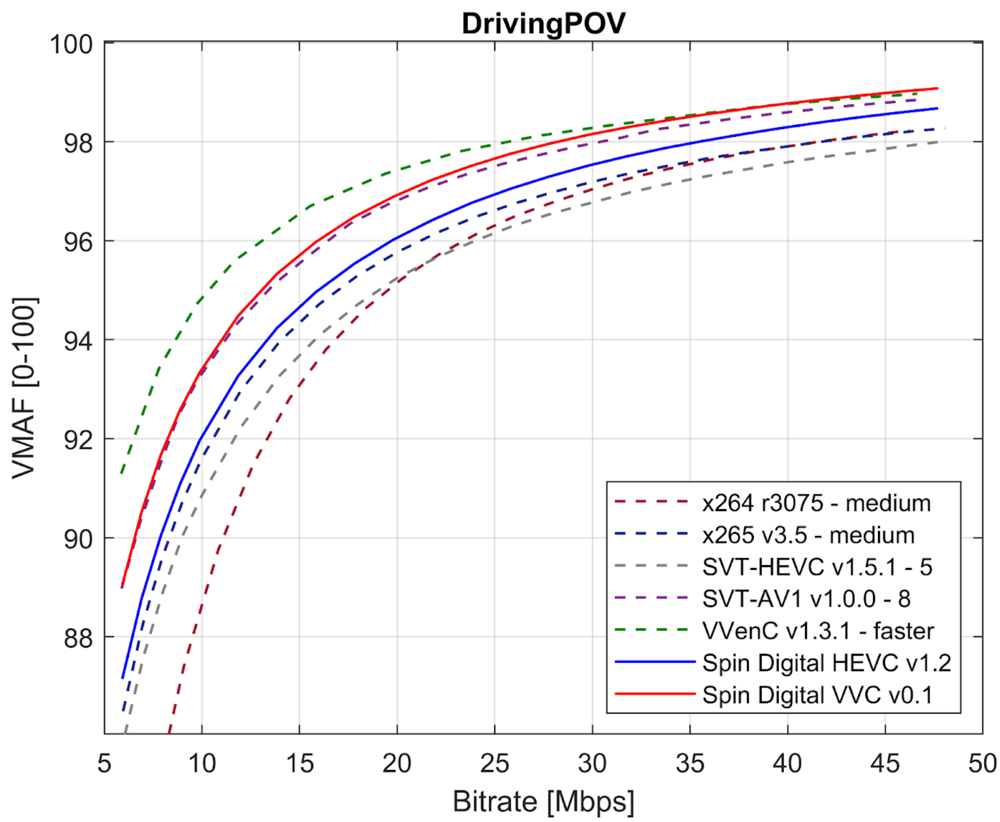


Figure 11: VMAF-bitrate curves for DrivingPOV

5.8. MULTITHREADED ENCODING SPEED

The performance of the encoders for 4K UHD 10-bit video was measured on a high-end server with two Intel Xeon Platinum 8368 CPUs (2x 38 cores). ***DrivingPOV*** was used in these experiments, being the most difficult sequence in the test set from an encoding speed point of view.

The criterion chosen to configure the encoders for performance tests was to select the bitrate such that a constant quality is achieved. A target VMAF score of 95 was chosen to produce a very high quality that in turn results in a sufficiently high bitrate for benchmark purposes.

The video encoders were also configured to enable the maximum number of threads for the target server. In order to reduce the impact of uncompressed input YUV file disk reading and output bitstream writing, the encoders were configured to use fast file read modes or encoding from memory and not to write the resulting bitstreams to the disk.

Table 7 shows the obtained performance results for each encoder and preset. In addition it includes the actual bitrate that produces a VMAF score of 95 and the overall CPU utilization in CPU cores. The system has a maximum of 76 CPU cores each one with 2-way Simultaneous Multithreading (SMT), also known as HyperThreading, for a theoretical maximum of 152 cores. Figure 12 shows the bitrate and encoding speed generated by all the encoders.

Table 7: Performance results corresponding to DrivingPOV generated by the encoders and presets in terms of actual bitrate for a VMAF of 95, encoding speed, and CPU utilization

Encoder - preset	Bitrate [Mbps]	Encoding speed [fps]	Encoding time [spf]	CPU utilization [CPU cores]
x264 r3075 - slower	18.50	21.00	0.048	27.02
x264 r3075 - med	20.00	58.28	0.017	31.37
x264 r3075 - vfast	31.60	80.26	0.012	27.04
x265 v3.5 - slower	12.29	2.72	0.368	23.12
x265 v3.5 - medium	17.01	19.46	0.051	22.92
x265 v3.5 - ultrafast	19.01	63.44	0.016	28.38
SVT-HEVC v1.5.1 - 3	17.39	27.49	0.036	144.60
SVT-HEVC v1.5.1 - 5	19.08	79.42	0.013	126.97
SVT-HEVC v1.5.1 - 7	20.89	115.31	0.009	66.45
SVT-HEVC v1.5.1 - 9	29.48	133.20	0.008	44.18
SVT-HEVC v1.5.1 - 11	34.98	146.54	0.007	31.34
SVT-AV1 v1.0.0 - 4	10.26	3.81	0.262	31.88
SVT-AV1 v1.0.0 - 6	10.36	14.20	0.070	40.98
SVT-AV1 v1.0.0 - 7	10.99	21.80	0.046	37.02
SVT-AV1 v1.0.0 - 8	13.38	37.35	0.026	37.2
SVT-AV1 v1.0.0 - 9	14.76	57.55	0.017	41.69
SVT-AV1 v1.0.0 - 10	16.43	82.38	0.012	33.25
SVT-AV1 v1.0.0 - 12	18.31	118.98	0.008	31.6
VVenC v1.3.1- fast	9.07	1.94	0.515	15.10
VVenC v1.3.1 - faster	10.37	2.77	0.361	14.11
Spin Digital HEVC v1.2	15.90	169.01	0.006	69.29
Spin Digital VVC v0.1	12.94	96.17	0.010	76.85

According to these results, the HEVC and VVC real-time encoders developed by Spin Digital are able to achieve a performance beyond real-time (60 fps) and, at the same time, produce the lowest bitrates under real-time conditions. The HEVC and VVC encoders run, respectively, at 169 fps and 96 fps producing average bitrates of 15.90 Mbps and 12.94 Mbps. Except VVenC, which is not designed for live applications, the rest of encoders require very fast presets to reach encoding speeds above 60 fps, but at the expense of a noticeable increase in the bitrate for achieving the same quality.

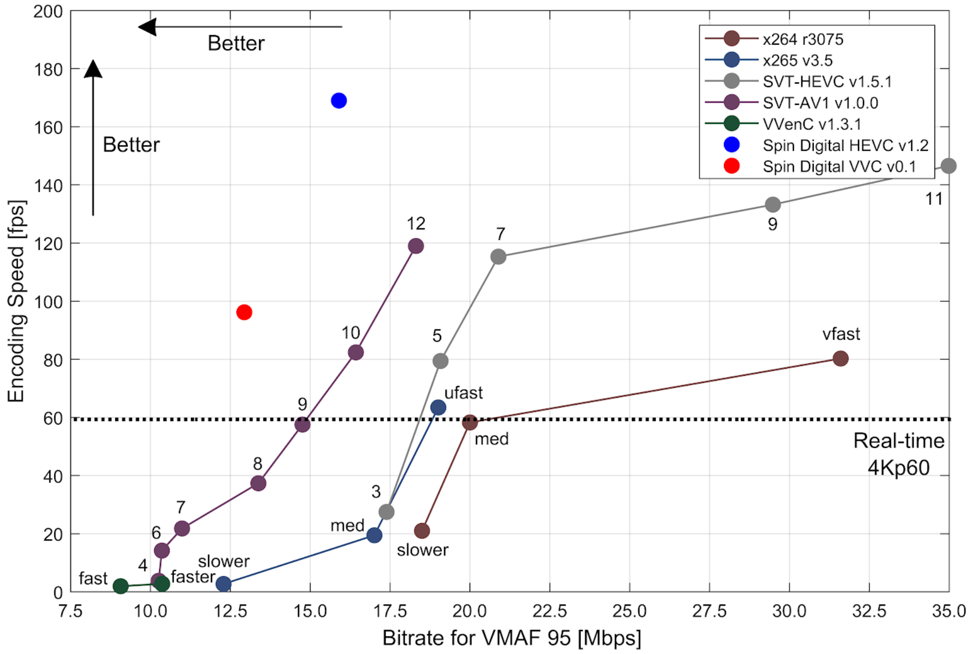


Figure 12: Actual bitrate for a VMAF of 95 and encoding speed produced by the encoders and presets when encoding DrivingPOV using two Intel Xeon Platinum 8368 CPUs (2x 38 cores)

Spin Digital's VVC encoder achieves a performance beyond real-time video and, at the same time, produces the lowest bitrate for a given target quality under real-time conditions.

It is also noteworthy that the complexity when running at maximum speed in a multicore system is very different from the single-threaded complexity reported in Table 4. In particular, it can be noticed that the parallel processing performance of SVT-AV1 is less efficient than that of Spin Digital VVC. Although in single threading Spin Digital VVC and SVT-AV1 - 8 have similar complexity, in multithreading SVT-AV1 - 8 is 2.5 to 3.0 times slower than Spin Digital VVC. In order to achieve real-time performance for 4Kp60 video, SVT-AV1 requires preset 10 which results in a higher bitrate than Spin Digital VVC (27% higher for DrivingPOV) for the same VMAF quality.

It is also observed in Table 7 that SVT-HEVC and Spin Digital's HEVC and VVC encoders are the ones that make the best use of CPU resources, which explains their very high encoding speed. The lower CPU utilization achieved by the other encoders (e.g. up to 47 CPU cores for SVT-AV1) indicates that their parallel architectures are noticeably less efficient than those implemented in the Spin Digital's encoders.

Note that these performance results have been obtained considering a quality criterion based on VMAF. In order to verify that the selected bitrates are also representative for high-quality video measured with other metrics, the PSNR, XPSNR, MS-SSIM, and VMAF scores generated by each encoder and preset are reported in Table 8.

As can be seen, the selected bitrates generate a VMAF score of around 95, MS-SSIM values between 0.985 and 0.988, and PSNR and XPSNR values exceeding 40.9 dB up to 41.7 dB. Although a non-negligible variation is observed in PSNR (and XPSNR), at these high quality levels, the differences in distortion are assumed to be not significant in perceptual terms. Therefore, constant visual quality can be assumed. A more detailed objective and subjective validation is beyond the scope of this paper. The results obtained in terms of bitrate and encoding performance are conclusive for the purposes of comparing different encoders.

Table 8: PSNR, XPSNR, MS-SSIM, and VMAF scores produced by the encoders and presets under assessment when encoding DrivingPOV at the selected bitrates

Encoder - preset	Bitrate [Mbps]	PSNR [dB]	XPSNR [dB]	MS-SSIM [0-1]	VMAF [0-100]
x264 r3075 - slower	27.52	40.70	40.57	0.984	94.30
x264 r3075 - med	30.06	40.73	40.57	0.984	94.34
x264 r3075- vfast	43.09	41.24	40.89	0.986	94.90
x265 v3.5 - slower	17.23	41.01	40.93	0.985	95.00
x265 v3.5 - medium	24.02	41.28	41.17	0.986	95.04
x265 v3.5 - ultrafast	36.15	41.14	41.06	0.986	95.03
SVT-HEVC v1.5.1 - 3	29.98	41.41	41.39	0.987	94.73
SVT-HEVC v1.5.1 - 5	40.04	41.52	41.48	0.987	94.91
SVT-HEVC v1.5.1 - 7	46.02	41.59	41.54	0.988	95.00
SVT-HEVC v1.5.1 - 9	57.02	41.73	41.67	0.988	95.00
SVT-HEVC v1.5.1 - 11	69.07	41.68	41.63	0.988	95.03
SVT-AV1 v1.0.0 - 4	16.44	41.30	41.38	0.987	95.03
SVT-AV1 v1.0.0 - 6	17.11	41.13	41.22	0.986	95.02
SVT-AV1 v1.0.0 - 7	21.04	41.14	41.21	0.986	95.03
SVT-AV1 v1.0.0 - 8	24.97	41.44	41.46	0.987	95.01
SVT-AV1 v1.0.0 - 9	23.83	41.41	41.46	0.987	95.02
SVT-AV1 v1.0.0 - 10	25.69	41.34	41.34	0.987	95.07
SVT-AV1 v1.0.0 - 12	32.83	41.16	41.15	0.986	95.01
VVenC v1.3.1 - fast	11.85	41.28	41.49	0.987	94.96
VVenC v1.3.1 - faster	15.65	41.32	41.42	0.987	94.92
Spin Digital HEVC v1.2	19.65	41.39	41.45	0.987	94.99
Spin Digital VVC v0.1	14.90	41.35	41.43	0.987	94.96

6. RECOMMENDED BITRATE FOR 4K LIVE VIDEO

For creating new generation streaming services based on VVC it is fundamental to estimate a recommended bitrate. This depends on several factors such as:

1. Resolution and frame rate: 1080p, 4K-UHD, 8K-UHD, and 25/30 or 50/60 fps
2. Coding standard: AVC/H.264, HEVC/H.265, AV1, VVC/H.266
3. Video encoder implementation
4. Use case: offline (VoD), live (live OTT, broadcasting)
5. Type of content: TV studio, TV series, live sports, UGC, gaming, screen content
6. A quality criterion, which can be based on an objective metric (e.g. PSNR, VMAF) or a subjective experiment (e.g. MOS).

This section aims to answer this open question for the use case of UHD-TV (4Kp60 10-bit HDR) live applications using two state-of-the-art real-time encoders: Spin Digital HEVC and Spin Digital VVC.

The video sequences are those described in Section 5.1, which can be considered representative for this use case as they encompass different types of content and a wide range of spatio-temporal complexities. In addition, 9 more 4K video footage have been added to this analysis: Aerial, DinnerScene and WindNature (Netflix 2022), and 6 excerpts of *The Explorers* nature content (The Explorers 2021). In all, the set of videos to determine the recommended range is composed of 20 1-minute clips.

VMAF-based criterion

The quality criterion selected is based on the VMAF metric, as follows. The recommended bitrate is the one that results in 1) a minimum VMAF score of 90 for most of the test video sequences (i.e. 75% of all sequences in the test set) and, 2) a minimum VMAF score of 70 for any test video sequence (see Table 9).

Table 9: VMAF-based quality criterion to determine the recommended bitrate for a video encoder

VMAF score	Equivalent MOS	Quality description	Percentage of videos
≥90	≥4.5	“good” - “excellent”	≥75%
≥70	≥3.5	“fair” - “good”	100%

VMAF 90 corresponds to a Mean Opinion Score (MOS) of around 4.5 (between “good” and “excellent” quality) and VMAF 70 to a MOS of 3.5 (between “fair” and “good” quality) (Li et al. 2018). This criterion is based on a similar recommendation for 8K HEVC content to achieve broadcast-grade video quality by NHK TSRL (Ichigaya and Nishida 2016). The first condition guarantees high quality of service in most cases, while the second ensures that highly complex videos are still viewed with acceptable quality.

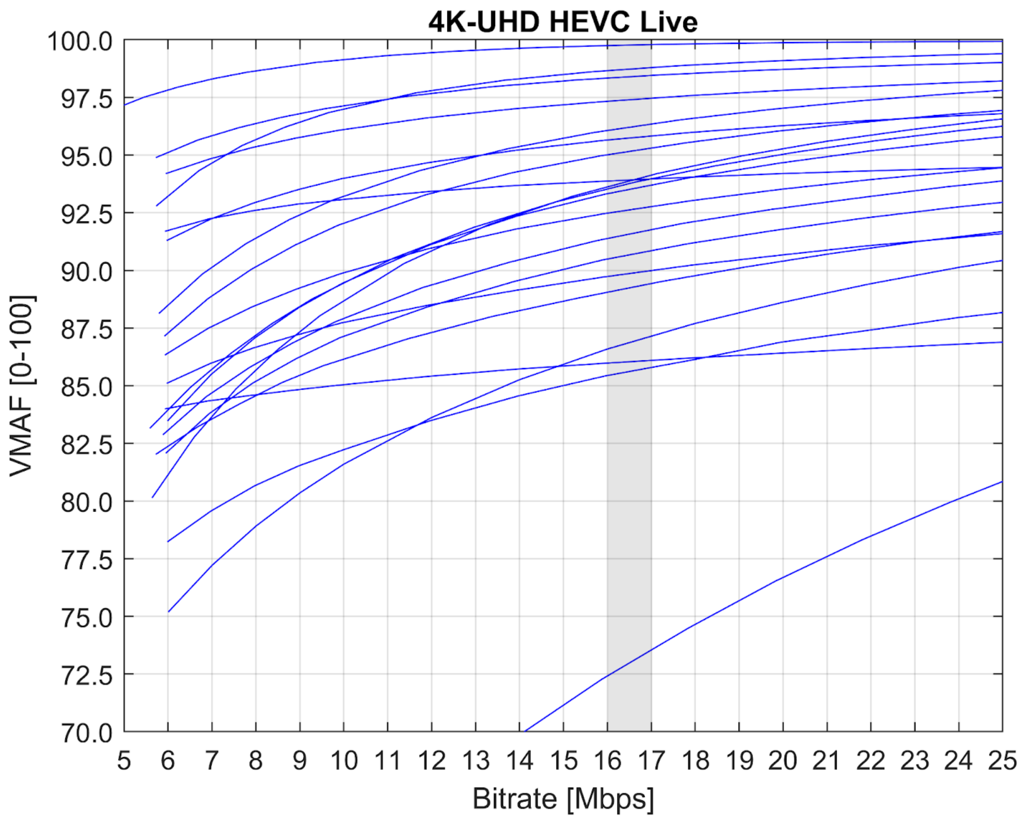


Figure 14: VMAF-bitrate curves produced by Spin Digital HEVC and recommended range of bitrates

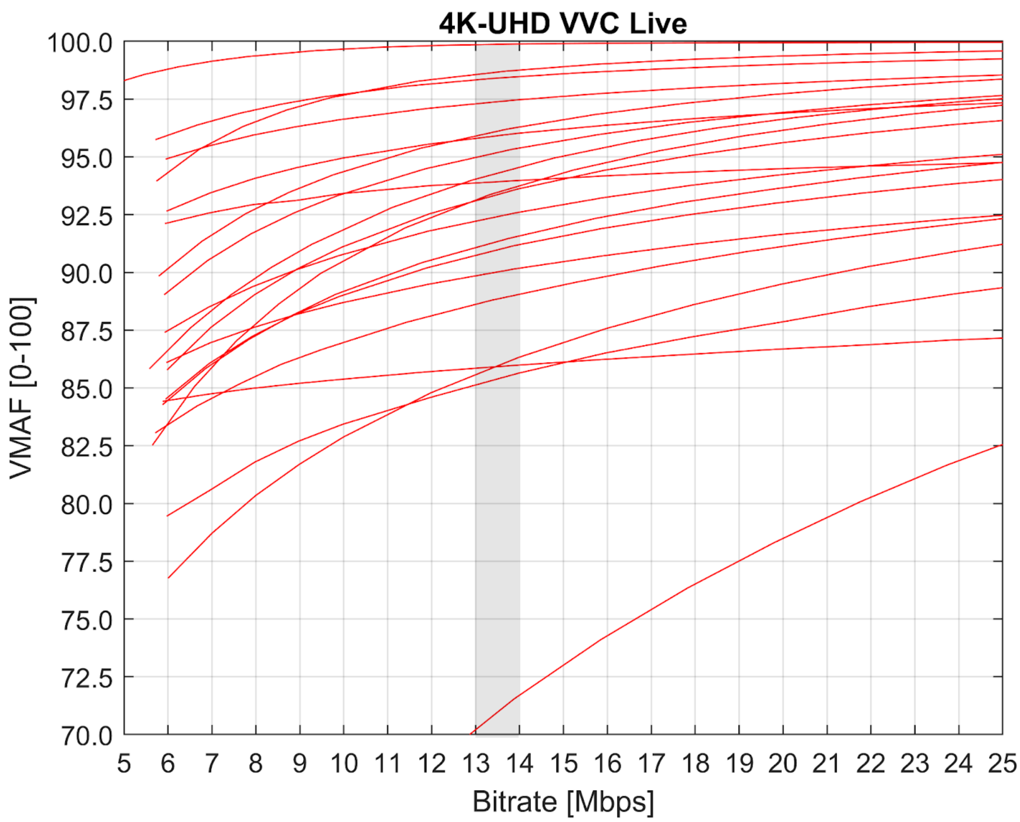


Figure 15: VMAF-bitrate curves produced by Spin Digital VVC and recommended range of bitrates

Figures 14 and 15 show the VMAF-bitrate curves generated by the HEVC and VVC real-time encoders, respectively. The figures illustrate in lightgray the regions where the quality criterion of Table 9 is met for both encoders. As can be seen, the VMAF-based criterion for high-quality 4K live video with Spin Digital HEVC is met at 16-17 Mbps, whereas the bitrate recommendation using Spin Digital VVC can go down to 13-14 Mbps.

The recommended bitrate range for 4Kp60 live applications with Spin Digital VVC goes from 13 to 14 Mbps, a significant reduction from the 16 to 17 Mbps required by Spin Digital HEVC

The sequence that produces the lowest quality score is ToddlerFountain. Since this video mostly contains water jets and splashes, Spin Digital HEVC - as well as the other encoders under evaluation - need to allocate many bits to this type of high-frequency texture in order to encode it with a certain level of quality.

Subjective test

In order to verify the visual quality of the videos encoded at the recommended bitrates, an informal subjective test was conducted by Spin Digital. The videos described in Section 5.1 were encoded with Spin Digital HEVC at 17 Mbps as well as with Spin Digital VVC at 14 Mbps. To conduct this test, a playback PC running Spin Digital's media player – Spin Player– which supports both HEVC (Spin Digital 2 2022) and VVC (Spin Digital 3 2022) was used. The videos were watched on a 4K HDR TV at a viewing distance of about 1.5 times the height of the TV (ITU-R 2019).

According to the subjective test, the perceived quality was generally very high. The video clips were rated as "excellent", except BerlinSeqs (VMAF 92), SolLevante (VMAF 93), and ToddlerFountain (VMAF 72-74). The first two sequences were rated "good" to "excellent", since the visible coding artifacts - mainly in the form of blocking effect - were not perceived bothersome. ToddlerFountain, which is the most problematic clip in the test set according to the objective metrics, was scored as "good" due to the continuous but not annoying blocking artifacts in the water jets and splashes.

7. 4K/8K VVC LIVE STREAMING AND BROADCASTING

For a video encoder to work reliably in a live streaming environment, several requirements in terms of performance and features should be met. These requirements are summarized next:

- **High average encoding speed:** The performance of the encoder should be on average higher than the target video frame rate (e.g. 60 fps), with a sufficient tolerance margin so that the instantaneous speed does not fall below the target frame rate in case of complex scenes. This requirement has been proven in the previous section for Spin Digital VVC encoder for 4Kp60 video.

- **Stable performance over time:** The encoder should also exhibit a stable encoding speed over time in order to prevent frame dropping that can degrade the temporal continuity of the encoded video.

- **Real-time operation mode:** The encoder should process the content at the target video resolution and frame rate, but if the content is too complex and the underlying hardware does not have the required computing capacity, the encoder should drop input frames instead of adding encoding delay.

- **HRD compliance:** The synchronization of encoders and decoders in live applications is defined by the HRD model, which specifies constraints on picture timing, buffer size and buffer handling. To ensure high-quality and smooth playback, appropriate encoding decisions should be made to prevent buffer overflows and underflows while providing consistent video quality over time.

- **I/O modules:** The live encoder should be equipped with I/O modules that can receive an uncompressed video and audio from the source (e.g. a camera) via SDI or a contribution stream over IP, and deliver the compressed output stream either over HTTP (e.g HLS, DASH) or TSoIP (e.g. UDP, RTP, SRT), or both simultaneously.

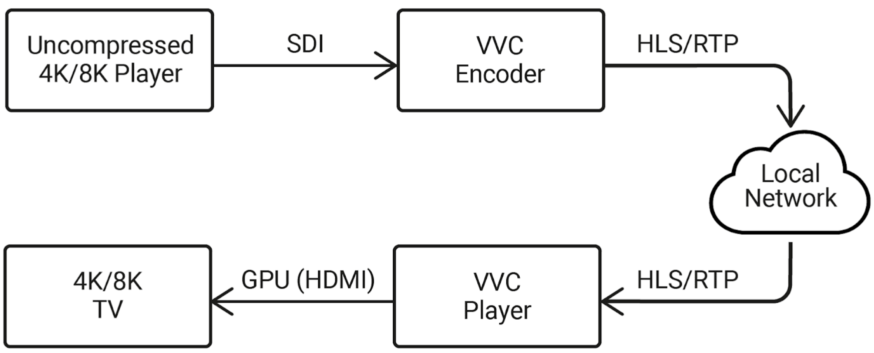


Figure 16: 4K/8K HDR live encoding and streaming workflow based on Spin Digital's VVC encoding and playback solutions

The new VVC encoder has been integrated into a framework for live applications designed to fulfill all the above requirements. Using the enhanced framework a complete VVC live encoding, streaming, playback workflow has been validated for 4Kp60 and 8Kp30 HDR 10-bit video (see Figure 16). The workflow is composed of three components:

- **Uncompressed media player:** A real-time software media player developed by Spin Digital that plays uncompressed YUV video over 12G-SDI in 4K and 8K resolutions together with PCM audio.
- **VVC real-time encoder:** An encoding server with 12G-SDI video capture, VVC video and AAC audio encoding, and HTTP and TSoIP streaming modules.
- **VVC media player:** A playback PC running Spin Digital's VVC media player, which handles VVC and AAC decoding, video and audio rendering, and 4K and 8K video output over GPU with HDMI interface.

A series of 4Kp60 video sequences with associated audio were live encoded in VVC at 25 Mbps and streamed over a local network using two transmission protocols: HLS for HTTP-based streaming, and RTP for TSoIP broadcast. Although 13-14 Mbit/s were determined to be the recommended rate, 25 Mbit/s was used for this test to stress the encoder performance.

On the receiving side, the VVC media player decoded and played back the reconstructed pictures on the UHD display. The same experiment was also performed successfully using videos in 8Kp30 10-bit format. Below are some images of the real-time encoder and player during the test session.

This proof-of-concept demonstrates that VVC UHD live encoding and streaming is now possible with current computing technologies and highly optimized VVC encoder and decoder software solutions.

> VVC UHD live encoding and streaming is now possible with Spin Digital's VVC encoder and decoder software solutions running on current computing technologies.

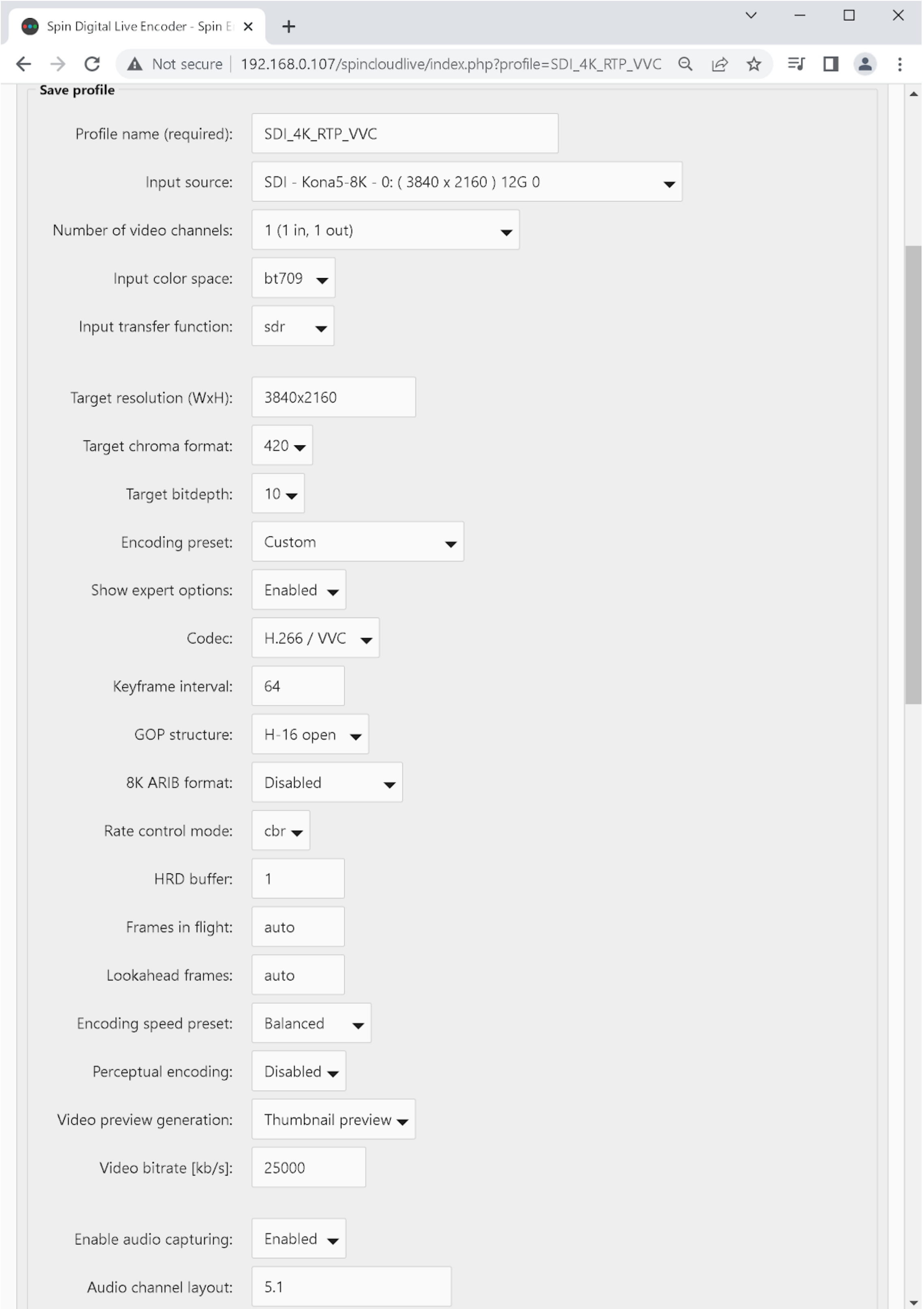


Figure 17: Configuration of encoding parameters

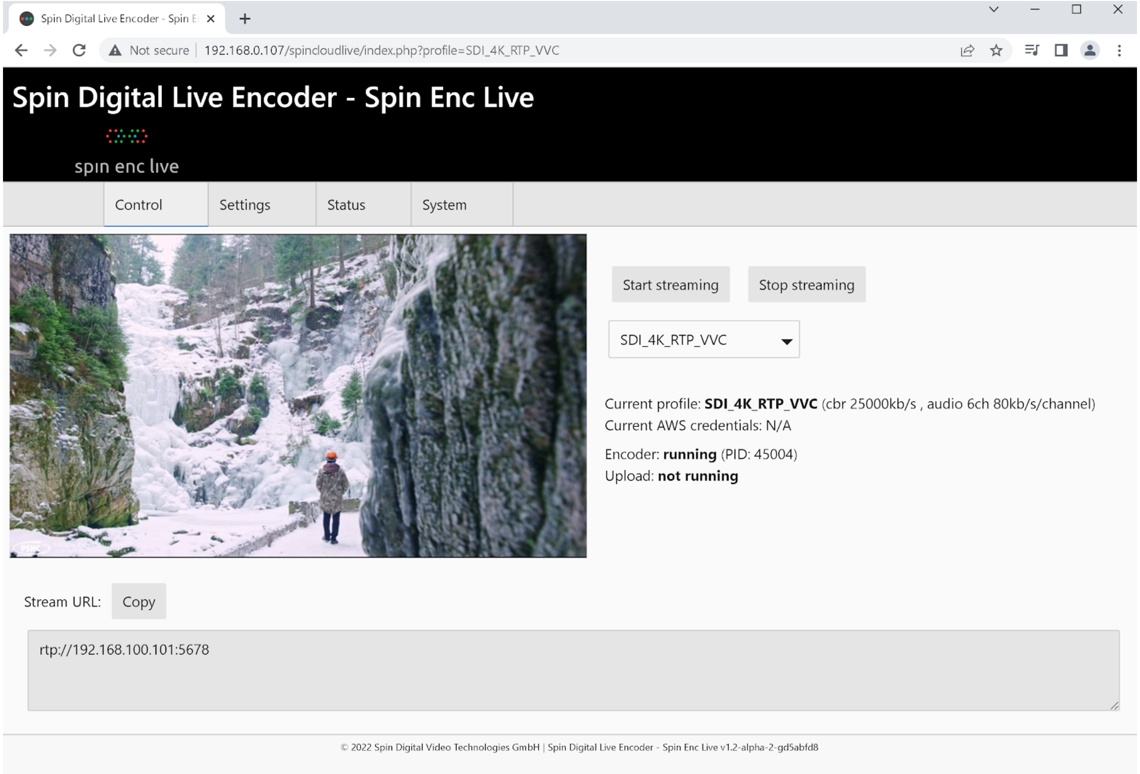


Figure 18: Encoding and streaming control

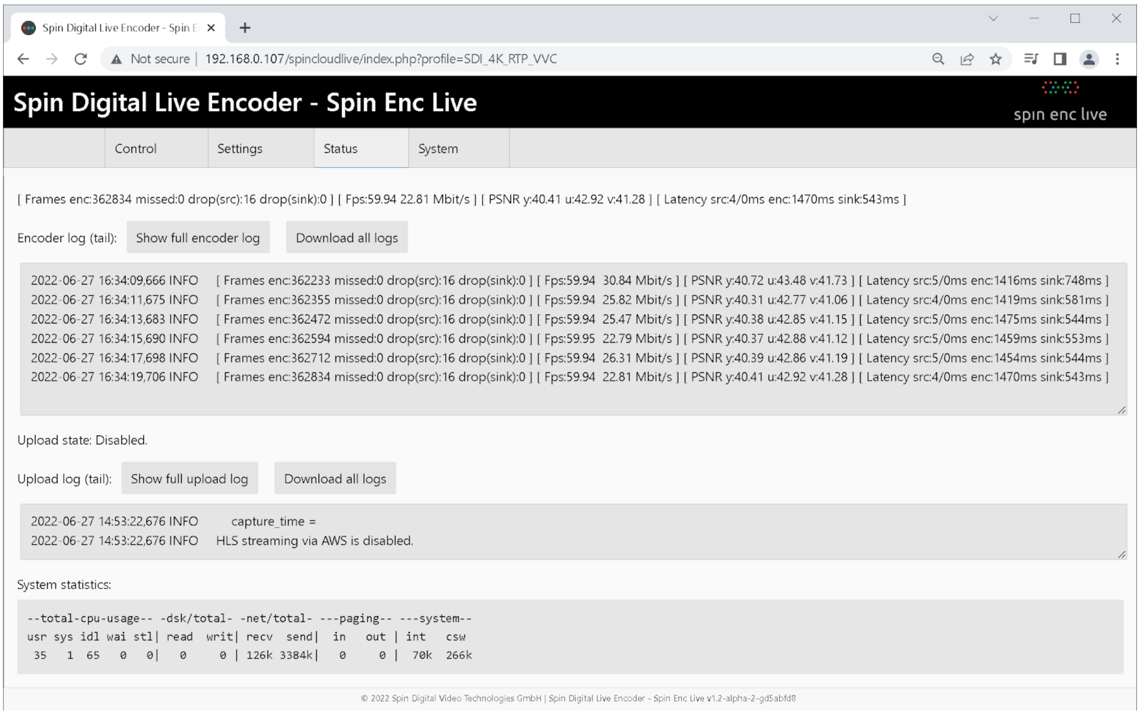


Figure 19: Status tab displaying encoding statistics in real-time

Figure 20: Live playback using Spin Player VVC

8. CONCLUSIONS

In this paper a VVC/H.266 real-time software encoder has been presented. The new encoder is designed to produce the compression, quality, and performance levels required for high-end UHD live streaming and broadcasting applications.

The main highlights of the VVC real-time encoder are summarized next:

- **Higher compression efficiency than state-of-the-art HEVC live encoder:** The first version of the VVC encoder achieves 18% bitrate savings, ranging from 9% to 34% depending on the video content, for the same quality with respect to a highly optimized HEVC real-time encoder. This has been possible at the cost of 2 times higher computational complexity, which is considered reasonable and has been proven feasible with current generation of CPU architectures. The VVC encoder has significantly lower complexity than other VVC implementations that require between 10 and 50 times more computation than HEVC. For 8K video, initial results have shown that the new VVC encoder achieves 26% bitrate savings at equal quality relative to the HEVC live encoder with 1.5 times greater complexity.

- **Real-time UHD video encoding on a single dual-socket server:** The new VVC encoder is a highly-optimized CPU-based software solution that can process 4K video at 60 fps as well as 8K video at 30 fps both in 10-bit HDR in real-time on a single server with two Intel Xeon Platinum processors with a total of 76 CPU cores.

- **Outperforming optimized open-source HEVC and AV1 encoders:** When running on a multi-core CPU system targeting 4K 60 fps real-time operation, the new VVC encoder produces the required encoding performance and results in the highest compression efficiency when compared to open-source HEVC and AV1 encoders such as x265, SVT-HEVC, and SVT-AV1.

- **Lower bandwidth for UHD live video:** Using an objective quality criterion based on VMAF and informal subjective tests, the recommended range of bitrates for 4Kp60 live applications using Spin Digital's VVC encoder ranges from 13 to 14 Mbps, compared to 16 to 17 Mbps needed by Spin Digital's HEVC to achieve the same level of quality. For 8Kp60 live applications the recommended bitrate with the new VVC encoder is 40 Mbps instead of the 50 Mbps required by the HEVC encoder.

- **Ready for 4K/8K live streaming and broadcasting:** The VVC encoder has been integrated into a live streaming framework developed by Spin Digital that includes: input capture via SDI and IP, pre-processing, pre-analysis, advanced rate control, audio and video encoding, and broadcast over IP and Internet streaming. Together with Spin Digital's VVC decoder and media player (Spin

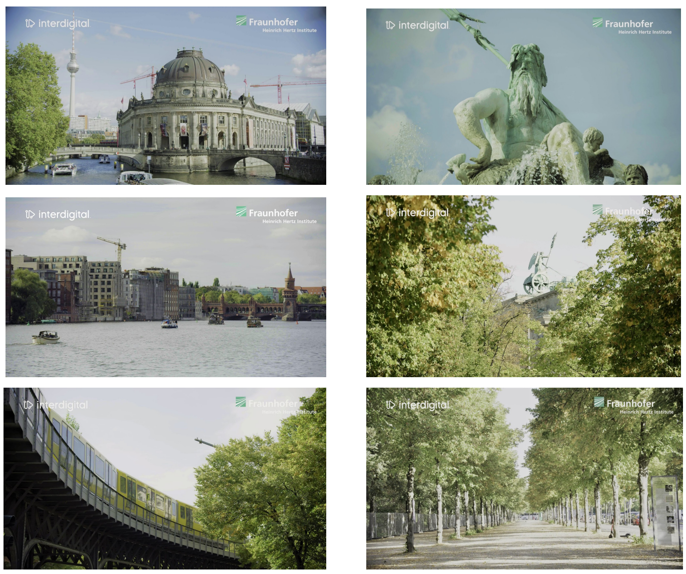
Player VVC), the end-to-end VVC encoding, streaming, and playback workflow has been successfully validated for 4Kp60 and 8Kp30 videos.

The new VVC encoder is the first generation of a VVC CPU-based software implementation, and it is expected that the encoder will improve over time for delivering UHD live video at higher quality with lower bitrates. VVC live encoding improvements will emerge with the use of new and more advanced encoding algorithms combined with the performance increases of next-generation CPU architectures. As a result, it is expected that the VVC encoder will be able to compress 8Kp60 10-bit HDR video in real-time for live applications with increased compression and quality.

APPENDIX: TECHNICAL DATA SHEET OF THE TEST VIDEO SEQUENCES

This appendix presents technical information about the video sequences used for codec comparison, including: sequence name, producer, source, content type, master file format, preprocessing operations, preprocessed file format and duration. Each table corresponds to a test video sequence.

Sequence	BasketballGame
Producer	Netflix
Source	Chimera (Netflix 2022)
Type	Footage
Master file format	- Format: TIF - Resolution: 4096x2160 pixels - Frame rate: 59.94 fps - Color model: RGB - Bit-depth: 16 bits - Transfer function: PQ - Color space: DCI-P3
Preprocessing	Center cropping from 4096x2160 pixels to 3840x2160 pixels, RGB to YUV-4:2:0, 16 bits to 10 bits, DCI-P3 to BT.2020
Preprocessed file format	- Format: Uncompressed YUV - Resolution: 3840x2160 pixels - Frame rate: 59.94 fps - Color model: YUV-4:2:0 - Bit-depth: 10 bits - Transfer function: PQ - Color space: BT.2020
Duration	1 minute: from 29’25” to 30’25”
Thumbnail	

Sequence	Berlin Sequences
Producer	Fraunhofer HHI, InterDigital
Source	(Fraunhofer HHI, InterDigital 2022)
Type	Footage
Master file format	- Codec: Uncompressed YUV - Resolution: 7680x4320 pixels - Frame rate: 60 fps - Color model: YUV-4:2:0 - Bit-depth: 10 bits - Transfer function: PQ - Color space: BT.2020
Preprocessing	Downsampling from 7680x4320 pixels to 3840x2160 pixels
Preprocessed file format	- Codec: Uncompressed YUV - Resolution: 3840x2160 pixels - Frame rate: 60 fps - Color model: YUV-4:2:0 - Bit-depth: 10 bits - Transfer function: PQ - Color space: BT.2020
Duration	1 minute: concatenation of six 10-second clips: BodeMuseum, NeptuneFountain2, OberbaumSpree, QuadrigaTree, SubwayTree, TiergartenParkway
Thumbnail	

Sequence	Driving POV
Producer	Netflix
Source	Chimera (Netflix 2022)
Type	Footage
Master file format	- Codec: TIF - Resolution: 4096x2160 pixels - Frame rate: 59.94 fps - Color model: RGB - Bit-depth: 16 bits - Transfer function: PQ - Color space: DCI-P3
Preprocessing	Center cropping from 4096x2160 pixels to 3840x2160 pixels, RGB to YUV-4:2:0, 16 bits to 10 bits, DCI-P3 to BT.2020
Preprocessed file format	- Codec: Uncompressed YUV - Resolution: 3840x2160 pixels - Frame rate: 59.94 fps - Color model: YUV-4:2:0 - Bit-depth: 10 bits - Transfer function: PQ - Color space: BT.2020
Duration	1 minute: from 23’00” to 24’00”
Thumbnail	

Sequence	FollowCar
Producer	Poznan Supercomputing and Networking Center - Immersify Project
Source	(Immersify 2018)
Type	Footage
Master file format	- Format: DPX - Resolution: 7680x4320 pixels - Frame rate: 59.94 fps - Color model: RGB - Bit-depth: 12 bits - Transfer function: SDR - Color space: BT.709
Preprocessing	Downsampling from 7680x4320 pixels to 3840x2160 pixels, RGB to YUV-4:2:0, 12 bits to 10 bits
Preprocessed file format	- Format: Uncompressed YUV - Resolution: 3840x2160 pixels - Frame rate: 59.94 fps - Color model: YUV-4:2:0 - Bit-depth: 10 bits - Transfer function: SDR - Color space: BT.709
Duration	1 minute: from 0’00” to 01’00”
Thumbnail	

Sequence	Karkonosze Mountains - Second clip (MC2)
Producer	Poznan Supercomputing and Networking Center - Immersify Project
Source	(Immersify 2018)
Type	Footage
Master file format	- Format: DPX - Resolution: 7680x4320 pixels - Frame rate: 59.94 fps - Color model: RGB - Bit-depth: 12 bits - Transfer function: SDR - Color space: BT.709
Preprocessing	Downsampling from 7680x4320 pixels to 3840x2160 pixels, RGB to YUV-4:2:0, 12 bits to 10 bits
Preprocessed file format	- Format: Uncompressed YUV - Resolution: 3840x2160 pixels - Frame rate: 59.94 fps - Color model: YUV-4:2:0 - Bit-depth: 10 bits - Transfer function: SDR - Color space: BT.709
Duration	1 minute: from 0’00” to 01’00”
Thumbnail	

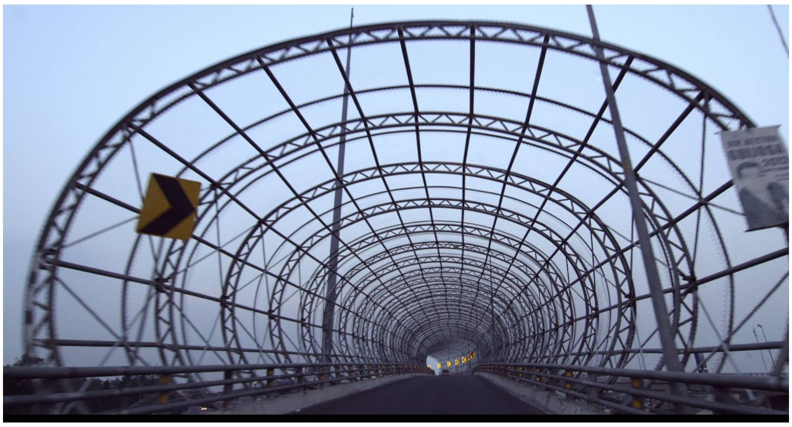
Sequence	Meridian
Producer	Netflix
Source	Meridian (Netflix 2022)
Type	Footage and CGI
Master file format	- Format: Jpeg 2000 - Resolution: 3840x2160 pixels - Frame rate: 59.94 fps - Color model: RGB - Bit-depth: 16 bits - Transfer function: PQ - Color space: P3-D65
Preprocessing	RGB to YUV-4:2:0, 16 bits to 10 bits, P3-D65 to BT.2020
Preprocessed file format	- Format: Uncompressed YUV - Resolution: 3840x2160 pixels - Frame rate: 59.94 fps - Color model: YUV-4:2:0 - Bit-depth: 10 bits - Transfer function: PQ - Color space: BT.2020
Duration	1 minute: from 6’18” to 07’18”
Thumbnail	

Sequence	RollerCoaster
Producer	Netflix
Source	Chimera (Netflix 2022)
Type	Footage
Master file format	- Format: TIF - Resolution: 4096x2160 pixels - Frame rate: 59.94 fps - Color model: RGB - Bit-depth: 16 bits - Transfer function: PQ - Color space: DCI-P3
Preprocessing	Center cropping from 4096x2160 pixels to 3840x2160 pixels, RGB to YUV-4:2:0, 16 bits to 10 bits, DCI-P3 to BT.2020
Preprocessed file format	- Format: Uncompressed YUV - Resolution: 3840x2160 pixels - Frame rate: 59.94 fps - Color model: YUV-4:2:0 - Bit-depth: 10 bits - Transfer function: PQ - Color space: BT.2020
Duration	1 minute: from 9’00” to 10’00”
Thumbnail	

Sequence	SolLevante
Producer	Netflix
Source	SolLevante (Netflix 2022)
Type	Animation
Master file format	- Format: ProRes - Resolution: 3840x2160 pixels - Frame rate: 24 fps - Color model: YUV-4:4:4 - Bit-depth: 10 bits - Transfer function: PQ - Color space: BT.2020
Preprocessing	YUV-4:4:4 to YUV-4:2:0
Preprocessed file format	- Format: Uncompressed YUV - Resolution: 3840x2160 pixels - Frame rate: 24 fps - Color model: YUV-4:2:0 - Bit-depth: 10 bits - Transfer function: PQ - Color space: BT.2020
Duration	2 minutes: from 0’15” to 02’15”
Thumbnail	

Sequence	Superposition
Producer	Unigine
Source	(Unigine 2017)
Type	Animation
Master file format	- Format: ProRes - Resolution: 7680x4320 pixels - Frame rate: 60 fps - Color model: YUV-4:2:2 - Bit-depth: 10 bits - Transfer function: SDR - Color space: BT.709
Preprocessing	Downsampling from 7680x4320 pixels to 3840x2160 pixels, YUV-4:2:2 to YUV-4:2:0
Preprocessed file format	- Format: Uncompressed YUV - Resolution: 3840x2160 pixels - Frame rate: 60 fps - Color model: YUV-4:2:0 - Bit-depth: 10 bits - Transfer function: SDR - Color space: BT.709
Duration	1 minute: from 0’00” to 01’00”
Thumbnail	

Sequence	ToddlerFountain
Producer	Netflix
Source	Chimera (Netflix 2022)
Type	Footage
Master file format	- Format: TIF - Resolution: 4096x2160 pixels - Frame rate: 59.94 fps - Color model: RGB - Bit-depth: 16 bits - Transfer function: PQ - Color space: DCI-P3
Preprocessing	Center cropping from 4096x2160 pixels to 3840x2160 pixels, RGB to YUV-4:2:0, 16 bits to 10 bits, DCI-P3 to BT.2020
Preprocessed file format	- Format: Uncompressed YUV - Resolution: 3840x2160 pixels - Frame rate: 59.94 fps - Color model: YUV-4:2:0 - Bit-depth: 10 bits - Transfer function: PQ - Color space: BT.2020
Duration	1 minute: from 21’40” to 22’40”
Thumbnail	

Sequence	TunnelFlag
Producer	Netflix
Source	El Fuente (Netflix 2015)
Type	Footage
Master file format	- Format: Uncompressed YUV - Resolution: 4096x2160 pixels - Frame rate: 59.94 fps - Color model: YUV-4:2:0 - Bit-depth: 10 bits - Transfer function: SDR - Color space: BT.709
Preprocessing	Center cropping from 4096x2160 pixels to 3840x2160 pixels
Preprocessed file format	- Format: Uncompressed YUV - Resolution: 3840x2160 pixels - Frame rate: 59.94 fps - Color model: YUV-4:2:0 - Bit-depth: 10 bits - Transfer function: SDR - Color space: BT.709
Duration	1 minute: 10-second long - concatenated 6 times
Thumbnail	

REFERENCES

Bern, T. 2019. “The Socialbakers Social Media Trends Report.” Socialbakers. https://www.socialbakers.com/blog/social-media-trends-report-key-insights-you-need-to-know.

Bjøntegaard, G. 2001. “Calculation of Average PSNR Differences between RD Curves.” *Doc. VCEG-M33. ITU-T SG16/Q6 VCEG.*

Bjøntegaard, G. 2008. ““Improvements of the BD-PSNR model.” *Doc. VCEG-AI11, ITU-T SG16/Q6 VCEG*.

Bross, B., J. Chen, S. Liu, and Y.-K. Wang. 2020. “Versatile Video Coding (Draft 10).” *Doc. JVET-S2001 of ITU-T/ISO/IEC Joint Video Exploration Team (JVET), 19th meeting by Teleconference*, (June).

Careless, J. 2021. “How Broadcasters Thrive In an OTT World.” TVTechnology. https://www.tvtechnology.com/news/how-broadcasters-thrive-in-an-ott-world.

Cisco Systems. 2018. “Cisco Visual Networking Index: Forecast and Trends, 2017–2022.” Cisco Systems White Paper. http://web.archive.org/web/20181213105003/https:/www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.pdf.

Elsevier. 2020. “OTT and live streaming services: Past, Present and Future.” Research.com. https://research.com/special-issue/ott-and-live-streaming-services-past-present-and-future.

The Explorers. 2021. “The Earth' s first Inventory in High-Definition (4K/8K HDR).” The Explorers Website. https://theexplorers.com/svod.

Fraunhofer HHI, InterDigital. 2022. “Reference Sequences.” 3GPP. https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/.

Grois, D., K. Giladi, K. Choi, M.W. Park, Y. Piao, M. Park, and K.P. Choi. 2021. “Performance Comparison of Emerging EVC and VVC Video Coding Standards with HEVC and AV1.” *in SMPTE Motion Imaging Journal* 130 (May): 1-12. 10.5594/JMI.2021.3065442.

Helmrich, C. 2021. “fraunhoferhhi/xpsnr: A Filter Plug-in Adding Support for Extended Perceptually Weighted Peak Signal-to-Noise Ratio (XPSNR) Measurements in FFmpeg.” Fraunhofer HHI GitHub. https://github.com/fraunhoferhhi/xpsnr.

Helmrich, C., S. Bosse, H. Schwarz, D. Marpe, and T. Wiegand. 2020. “A Study of the Extended Perceptually Weighted Peak Signal-to-Noise Ratio (XPSNR) for Video Compression with Different Resolutions and Bit-depths.” *ITU Journal: ICT Discoveries Articles* 3, no. 1 (May): 1-8. 10.1109/ICASSP40776.2020.9054089.

Ichigaya, A., and Y. Nishida. 2016. “Required Bit Rates Analysis for a New Broadcasting Service Using HEVC/H.265.” *IEEE Transactions on Broadcasting* 62, no. 2 (June): 417-425. 10.1109/TBC.2016.2550778.

Immersify. 2018. “Content & Demos – Immersify.” Immersify Website.

https://immersify.eu/content-demos/.

ITU-R. 2019. “BT.500 : Methodologies for the Subjective Assessment of the Quality of Television Images.” ITU Website. https://www.itu.int/rec/R-REC-BT.500-14-201910-I/en.

Li, Z., J. Bampis, J. Novak, A. Aaron, K. Swanson, A. Moorthy, and J. De Cock. 2018. “VMAF: The Journey Continues.” Netflix TechBlog. https://netflixtechblog.com/vmaf-the-journey-continues-44b51ee9ed12.

Linkedin. 2022. “LinkedIn Live | Overview.” LinkedIn Business. https://business.linkedin.com/marketing-solutions/linkedin-live.

Netflix. 2021. “Netflix/vmaf: Perceptual Video Quality Assessment Based on Multi-method Fusion.” Netflix GitHub. https://github.com/Netflix/vmaf.

Netflix. 2022. “Open Source Content.” Netflix Open Content. https://opencontent.netflix.com/.

Schwarz, B. 2022. “Brazil's TV 3.0 used Ateme's Titan to Prove the Viability of its Live 8K Use Case.” 8K Association Newsletter. https://8kassociation.com/industry-info/8k-news/brazils-tv-3-0-used-atemes-titan-to-prove-the-viability-of-its-live-8k-use-case/.

Schwarz 2, B. 2022. “Live Production 8K Video is Down to 48 Mbps With 40 Mbps Already in Spin Digital's Sights.” 8K Association Newsletter. https://8kassociation.com/industry-info/8k-news/live-production-8k-video-is-down-to-48-mpbs-with-40-mpbs-already-in-spin-digitals-sights/.

SES. 2022. “Innovative hybrid broadcast/OTT services.” SES. https://www.ses.com/find-service/broadcasters/ott-broadcasters.

Spin Digital. 2020. “Whitepaper: 8K HEVC Real-time Encoder (Spin Enc Live).” Spin Digital Tech Blog. https://spin-digital.com/tech-blog/whitepaper-spin-enc-live/.

Spin Digital. 2022. “8K HEVC Real-time Encoder (Spin Enc Live).” Spin Digital Products. https://spin-digital.com/products/spin_enc_live/.

Spin Digital 2. 2022. “8K HEVC Media Player (Spin Player HEVC).” Spin Digital Products. https://spin-digital.com/products/spin_player/.

Spin Digital 3. 2022. “8K VVC Media Player (Spin Player VVC).” Spin Digital Products. https://spin-digital.com/products/spin_player_vvc/.

Tange, O. 2018. “GNU Parallel 2018.” Zenodo. https://zenodo.org/record/1146014#.Yf0NNPgo8UE.

Ultra HD Forum. 2021. “Guidelines – Ultra HD Forum.” Ultra HD Forum. https://ultrahdforum.org/guidelines/.

Unigine. 2017. “Superposition benchmark.” UNIGINE Benchmarks. https://benchmark.unigine.com/superposition.

Wang, Z., E.P. Simoncelli, and A.C. Bovik. 2003. “Multi-scale Structural Similarity for Image Quality Assessment.” *Proceedings of the 37th EEEE Asilomar Conference on Signals, Systems and*

Computers 2 (November): 1398-1402. 10.1109/ACSSC.2003.1292216.

Wyman, O. 2021. “5 Reasons why Live Streaming will still be relevant after COVID.” StriveCast. https://strivecast.com/5-reasons-why-live-streaming-will-still-be-relevant-after-covid/#.

Xiph. 2015. “Xiph.org :: Derf's Test Media Collection.” Xiph.org :: Test Media. https://media.xiph.org/video/derf/.

spin digital

Spin Digital Video Technologies GmbH.
Helmholtzstraße 2-9, 10587 Berlin. Germany
www.spin-digital.com | info@spin-digital.com