



A REAL-TIME 8K-60-FPS-HDR VVC/H.266 SOFTWARE ENCODER FOR NEXT-GENERATION LIVE APPLICATIONS

M. Alvarez-Mesa¹, C. C. Chi¹, S. Sanz-Rodriguez¹, D. F. de Souza¹,
R. Velhal²

¹ Spin Digital Video Technologies GmbH, Germany
² Intel Corporation, USA

ABSTRACT

VVC/H.266 is the newest video coding standard designed to significantly improve compression efficiency over HEVC/H.265 and provide efficient coding for a broader range of video formats. Live VVC encoding will facilitate the adoption of next-generation applications such as 8K terrestrial broadcasting or 8K adaptive bitrate streaming. This paper aims to provide an analysis of the capabilities of VVC for live applications by describing a real-time 8K VVC software encoder. The encoder has been extensively optimised with advanced mode decision and partitioning algorithms, SIMD instructions, and a scalable multithreading architecture. The proposed encoder achieves 24% bitrate reduction at the same quality, over a highly optimised real-time HEVC encoder that was deployed during the Tokyo 2020 Olympics games. By combining all optimizations and Intel's 4th generation Xeon® Scalable Processors the new VVC encoder reaches the performance for 8K 60 fps, 10-bit, HDR live encoding at broadcast-grade quality.

INTRODUCTION

The 8K-UHD TV format has been designed to provide a stronger sensation of realness and, at the same time, a much more immersive experience in which the user is fully absorbed by the audiovisual content, Sugawara and Masaoka (1). Recent studies have also proven that 8K brings in additional quality benefits, such as the perception of a higher depth of field and 3-dimensionally, resulting in an increased sense of realness and immersion, Park et al. (2), Masaoka et al. (3).

Data storage, transmission, and playback of uncompressed 8K video are complex and expensive. Uncompressed live 8Kp60 (7680x4320px, 10-bit, at 60 fps) video results in a data rate of 48 Gbps. To achieve practical distribution data rates, advanced video coding is needed. The main challenge of an 8K video encoder is, therefore, to compress the video signal in a *very high quality* to ensure the expected Quality of Experience (QoE) and, at the same time, deliver a *relatively low bitrate* so that the encoded bitstream can be transmitted over practical distribution networks, for example over the open Internet. Moreover, for live media applications, the encoder must operate in *real time* and with *low latency*.

8K live broadcasting and streaming has been tested and deployed in multiple environments and applications. Some highlights include: the 8K satellite channel launched by NHK in Japan, Hara et al. (4); the live 8K OTT streaming pilot over the Internet of the Tokyo 2020 Olympic Games, Maglitta and Velhal (5), and the 8K VR live streaming service



of the Beijing 2022 Winter Olympic Games, Koenen et al. (6). In all these cases the High Efficiency Video Coding (HEVC)/H.265 codec has been used for final distribution. NHK published in 2015 the first recommendation for 8K live services to produce broadcast-grade quality and defined a bitrate of 85 Mbps using HEVC, Sugito et al (7). Although lower bitrates have been enabled by more recent HEVC encoders, HEVC is reaching a point of saturation for real-time conditions, which means that adding more computation will result in marginal compression gains.

To further reduce the bandwidth beyond the HEVC capabilities, the latest video coding standard, called Versatile Video Coding (VVC)/H.266, is needed, Bross et al. (8). VVC includes features to provide better compression efficiency for ultra-high-resolution content with extended colour gamut and higher bit-depth support, which make it more suitable for encoding 8K 10-bit HDR video compared to HEVC. VVC also includes coding tools for more efficient Adaptive Bitrate (ABR) streaming and scalable encoding.

Optimised real-time VVC encoders will be essential in the near future to facilitate the deployment of next generation 8K live applications such as 8K terrestrial broadcasting or 8K ABR streaming. Although some 8K VVC live encoders have already been announced, KDDI (9) and Spin Digital (10), these only reach the performance required for live 8K at maximum 30 fps. The first real-time 8Kp60 VVC encoding implementations are also expected to be CPU-based, running on the latest-generation CPU architectures. However, to the best of our knowledge, none has yet been released or demonstrated.

IMPLEMENTING A LIVE 8K VVC ENCODER

A VVC software encoder for 4K and 8K HDR live broadcasting and streaming has been developed by Spin Digital. The encoder has been extensively optimised for latest-generation Intel CPU architectures to achieve the performance and compression levels required for 8K video. In this section we describe the target architectures and the optimizations applied to the encoder for reaching the required performance, quality, and application requirements.

Platform Description

The encoder is designed to run on a standard CPU architecture. The target platform consists of a dual-socket CPU based on the 4th Generation Intel® Xeon® Scalable processor, also known as Sapphire Rapids or SPR. Two important characteristics of the latest generation of CPU architectures, including SPR, is the inclusion of many CPU cores and wide Single Instruction Multiple Data (SIMD) vector units, Nassif et al (12)

The SPR architecture allows up to 60 cores (120 threads with Hyper-Threading Technology) per socket, with up to 120 cores in dual-socket configuration, Mulnix (13). The CPU architecture also supports Intel AVX-512 instructions with 512-bit wide vector units, Intel DL Boost Vector Neural Network Instructions (VNNI), and Intel Advanced Matrix Extensions (Intel AMX), all of which can be used to accelerate media applications including video codecs, Chi et al (14). Table 1 summarises the main features of the Sapphire Rapids (SPR) architecture compared to its predecessor the 3rd Gen Intel Xeon Scalable, also called Ice Lake (ICL).

Mapping the Live Encoder Application to the Target CPU Platform

The encoder should be able to make efficient use of the current target platform as well as those of the foreseeable future. This entails being able to make use of wide SIMD instructions and many cores (e.g., more than 100 cores or 200 threads). But using more



cores should not increase the latency of the encoder. The target encoding latency of 1-3 seconds should be maintained even when more computing resources are used. This implies that coarse-grain parallelization strategies (e.g., GOP, Intra-Period or segment level parallelism) are not appropriate. The challenge is to utilise as many cores as possible in an efficient way while keeping the encoding latency under a few seconds.

	3rd Gen Intel Xeon Scalable Ice Lake (ICL)	4th Gen Intel Xeon Scalable Sapphire Rapids (SPR)
SIMD instructions	AVX-512 + DL-Boost (VNNI)	AVX-512 + DL Boost (VNNI) + AMX
Core count / Hyper Threading	Up to 40 cores Up to 80 threads	Up to 60 cores Up to 120 threads
DRAM Memory	8 channels DDR-4 3200	8 channels DDR-5 4800
PCI bus	PCI Express v4.0	PCI Express v5.0
Inter-processor communication: UPI	3 links per CPU - 20x wide 11.2 GT/s	4 links per CPU - 24x wide 16 GT/s

Table 1 – Main features of the target platform: a dual-socket 4th Gen Intel Xeon Scalable Sapphire Rapids (SPR) processor with high core counts and improved memory and I/O.

Two main types of optimizations have been performed for mapping the encoder to the target platform and reaching the target performance: a low-level SIMD optimization and a high-level multithreading architecture.

- **SIMD optimizations:** SIMD has been applied to most of the encoder modules using SSE4.2 (128-bit), AVX2 (256-bit), and AVX512 with DL- Boost (512-bit) instructions. The overall application performance when using AVX-512 (and DL-Boost) in the SPR platform is up to 21% higher compared to AVX2.
- **Multithreaded architecture:** the encoder includes a scalable parallelization framework that combines multiple levels of parallelism including wavefronts, picture partitions, and frame-level, which allows the performance to scale to systems with large numbers of CPU cores.

On the high-level (See picture 1) the encoder application is divided into four main processing stages: input capture, lookahead, VVC encoding, and output muxing and streaming. The input capture collects the input frames from an SDI source, in the lookahead stage the complexity of the input frames is estimated, and a rate control model is used to allocate the bits to the video frames. The VVC encoding stage performs the actual encoding, including among others, block partitioning, mode decision, and final bitstream creation. The model is updated after the frames finish encoding. The Muxing and Streaming stage takes the resulting bitstream and produces the output in the required format.

The parallelism that can be exploited is a combination of wavefront and frame-level parallelism. Wavefront parallelism allows processing blocks of the same frame in parallel following dependencies from neighbouring blocks. To increase the parallelism, additional frames can be processed in parallel, Chi et al (15).

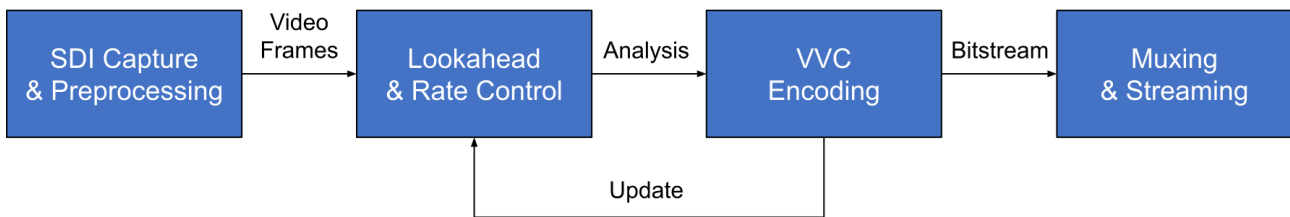


Figure 1 – High-level architecture of the VVC live encoder

As the target platform includes up to 120 cores and 240 threads, it is expected that more parallelism must be extracted to fully use the CPU resources. Considering that the encoder is already exploiting wavefronts and frame-level parallelism there are basically two ways of increasing the parallel processing without significantly impacting the latency or quality: more spatial parallelism in the form of picture partitions (slices, tiles, or sub-pictures) or more temporal parallelism with more frames-in-flight by increasing the encoding buffer (HRD).

When increasing the buffer size from the default setting of 1 second to, for example, 2 seconds, the number of frames in flight is also duplicated. The use of a 2-second buffer has negligible impact on quality compared to a 1-second buffer, but the overall latency is increased, which will be 1 second higher. In some use cases, such as HLS and DASH streaming, with typical latencies of several seconds this is not an issue; for broadcasting, with latencies ranging from 3 to 11 seconds, Amazon Web Services (16), this increase in latency could still be acceptable.

When adding picture partitions (horizontal x vertical partitions), more intra-frame parallelism can be extracted without dependencies and with the added benefit of reduced latency. If the number of partitions is small there is a minimal objective quality impact, but some subjective quality degradations might appear which require careful consideration.

Integration into a Live Encoding Framework

Finally, the VVC encoder has been integrated into a complete live framework that includes input capture based on 12G SDI, pre-processing (scaling, colour conversion, tone mapping), core VVC (and HEVC) encoding, audio encoding (MPEG-H Audio, AAC), lookahead analysis, Constant Bitrate (CBR) control with HRD model and Variable Bitrate (VBR) control, perceptually optimised encoding, and streaming for HTTP (HLS, DASH) or TS-over-IP delivery (RTP, SRT, RIST, Zixi).

ENCODER ASSESSMENT FOR 8K-UHD BROADCASTING AND STREAMING

The proposed real-time VVC encoder has been assessed in terms of compression efficiency, encoding complexity and multithreaded encoding speed.

The video encoder was configured assuming an **8K-UHD live broadcasting and streaming application**. This use case requires that the rate control algorithm is enabled and that long Group-of-Picture (GOP) structures are used for maximum compression efficiency with frequent random-access points (e.g., an intra period of 1 to 3 seconds long).

The results are also compared with those of four open-source CPU-based software encoders and two GPU-based hardware encoders of different coding standards, including HEVC, AV1 and VVC.



Video Sequences

Seven video sequences in 8K format (7680x4320 pixels, 60 fps, 4:2:0, 10-bit, SDR/HDR, and BT.709/BT.2020) were used to evaluate the compression efficiency and complexity of the VVC encoder. Each sequence has a 1-minute duration which is long enough to stabilise the rate control. The test sequences include: *BerlinSeqs* from Fraunhofer HHI (17) *FollowCar* and *MC2* from Poznan Supercomputing and Networking Center (PSNC) (18); *Superposition* from Unigine (19); two clips from NHK Technologies called *Fuji* and *Haru*; and one nature content from The Explorers (20) called *Teaser 1*.

Table 2 presents detailed information of the test sequences, including Spatial Information (SI) and Temporal Information (TI), ITU-R (21). Two clips are in High Dynamic Range (HDR) PQ with a BT.2020 colour gamut, one in HDR HLG also with BT.2020, whereas the remaining are in SDR BT.709. The 8K clips present medium to high spatial complexities and low to medium temporal complexities.

	Producer	Type	Format	SI	TI
BerlinSeqs	Fraunhofer HHI	Footage	8Kp60 PQ	100.8 (med)	59.3 (low)
FollowCar	PSNC	Footage	8Kp59.94 SDR	150.8 (med)	113.3 (med)
Fuji	NHK Tech.	Timelapse	8Kp59.94 SDR	133.6 (med)	20.5 (low)
Haru	NHK Tech.	Footage	8Kp59.94 HLG	186.9 (high)	65.5 (low)
MC2	PSNC	Footage	8Kp59.94 SDR	187.2 (high)	86.1 (low)
Superposition	Unigine	CGI	8Kp60 SDR	153.4 (med)	92.1 (low)
Teaser 1	The Explorers	Footage	8Kp50 PQ	101.6 (med)	52.7 (low)

Table 2 – Technical information of the 8K clips: producer, type of content, format, SI, TI

Video Encoders

The proposed VVC encoder (SpinVVC) has been compared to a state-of-the art HEVC 8K live encoder (SpinHEVC). In addition, other open-source and GPU-based encoders have been used for comparison and reference, including: two HEVC software implementations (x265, SVT-HEVC), one HEVC GPU-based implementation (NVENC-HEVC); one GPU-based AV1 implementation (OneVPL-AV1); and one VVC software encoder (VVenC). Table 3 provides a description of all the video encoders under assessment.

	Type	Standard	Version	Release date	Developer
x265	Software	HEVC	3.5	Apr. 2022	FFmpeg (22)
SVT-HEVC	Software	HEVC	1.5.1	June 2021	Intel (23)
NVENC-HEVC	Hardware	HEVC	12.0	Nov. 2022	Nvidia (24)
SpinHEVC	Software	HEVC	2.0	Feb. 2023	Spin Digital (25)
OneVPL-AV1	Hardware	AV1	2.8	Nov. 2022	Intel (26)
VVenC	Software	VVC	1.8.0	Apr. 2023	Wieckowski et al (27)
SpinVVC	Software	VVC	2.0	Feb. 2023	Spin Digital (28)

Table 3 – Technical information about the assessed encoders

Comparison Metrics

The video encoders were compared in terms of compression efficiency, encoding complexity and encoding performance.



Compression efficiency: BD-rate

The Bjøntegaard Delta (BD)-rate method, Bjøntegaard (29) and (30), was used to compute compression efficiency. It computes the average bitrate increase produced by a test encoder referred to a baseline encoder at the same quality. SpinHEVC was selected as the baseline encoder.

To have a robust assessment of the quality gains (or losses), the BD-rate method was calculated using following four quality metrics: Peak Signal-to-Noise Ratio (PSNR), Perceptually Weighted PSNR (XPSNR), Helmrich (31), Multi-Scale Structural Similarity (MS-SSIM), Wang et al. (32), and Video Multi-method Assessment Function (VMAF), Netflix (33). The PSNR, XPSNR, and MS-SSIM metrics were calculated using the luma and chroma components (note that VMAF only considers the luma component).

Encoding complexity: CPU time

Encoding complexity was measured as average CPU time (including both user-level and system-level CPU utilisation) over the target bitrates during the encoding process, relative to a reference encoder (i.e., SpinHEVC). CPU time is the accumulated time across all CPU cores and can therefore be considered as *single-threaded encoding time*. For CPU time measurements, a server with 4x Intel Xeon Platinum 8176 CPU (4x 28 cores) and an Ubuntu 20.4 OS was used. It is worth mentioning that the CPU time metric is only applicable to software encoders.

Encoding performance: average encoding speed

The maximum performance of the encoders was measured in terms of frames per second; for the software CPU-based encoders this was obtained running the encoders on the multi-core platform and enabling the maximum encoding threads; for the hardware GPU-based encoders the fps was obtained by using the selected GPUs. This metric allows us to determine if an encoder fulfils the minimum requirements for real-time encoding at a target frame rate.

Computing Platforms

Two servers were used to measure encoding performance. One included a 3rd Generation Intel Xeon Scalable processor (Ice Lake or ICX), the other one included a 4th Generation Intel Xeon Scalable Processor (Sapphire Rapids or SPR), which were used for testing the CPU-based encoders. Two GPUs were used for testing the GPU-based encoders: an Nvidia RTX 3070 GPU to encode with NVENC-HEVC and an Intel ARC A770 GPU to encode with OneVPL-AV1. Table 4 shows the specifications of these two platforms.

	Ice Lake (ICX)	Sapphire Rapids (SPR)
CPU	2x Intel Xeon Platinum 8368 @ 2.4 GHz (2x 38 cores)	2x Intel Xeon Platinum 8480+ @ 2.00GHz (2x 56 cores)
DRAM	16x 16 GB DDR4 3200 MHz	16x 32 GB DDR5 4800 MHz
GPU	Intel ARC A770, Nvidia RTX 3070	None
OS	Red Hat 8.5	Ubuntu 22.04

Table 4 – Specifications of the computing platforms: CPU, DRAM, GPU, and OS

Encoding Settings

The video encoders were configured assuming the 8K-UHD broadcast scenario as follows: random-access encoding mode (long GOP), open GOP, 1-second intra period, and CBR with a 1-second HRD buffer, except for VVenC which only supports unconstrained VBR



when using random-access mode. Other encoding parameters, such as GOP size, GOP structure, and lookahead window, were kept at their default values.

The encoders were tuned to maximise PSNR. Their perceptually optimised encoding (POE) modes were disabled, as none of the objective quality metrics used for the analysis are able to detect the subjective improvements that these modes can provide, Sanz-Rodríguez and Alvarez-Mesa (34).

As for the encoding bitrates, a range from 20 Mbps to 80 Mbps was used. 80 Mbps corresponds to the current recommended bitrate for 8K live using HEVC (7), lower bitrates are expected to be feasible using newer HEVC encoders and the latest VVC encoders, Sugito et al. (35).

For each encoder, several presets were selected to analyse different trade-offs between quality and speed. In addition, in the case of NVENC-HEVC, the high-quality (hq) tuning mode was chosen, as it is the best suited for a broadcast scenario.

Compression Efficiency and Encoding Complexity

Figure 2 shows the results produced by the encoders in terms of BD-rate based on PSNR, XPSNR, MS-SSIM, and VMAF, and CPU time. The baseline encoder is *SpinHEVC - fast*, where *fast* is the preset that provides a good trade-off between quality and performance and reaches the required speed for 8K 60 fps live applications when running on the Ice Lake server (25). The BD-rate results for the hardware encoders are plotted on the left side of the figures without CPU time information since CPU time is not meaningful for GPU-based hardware encoders.

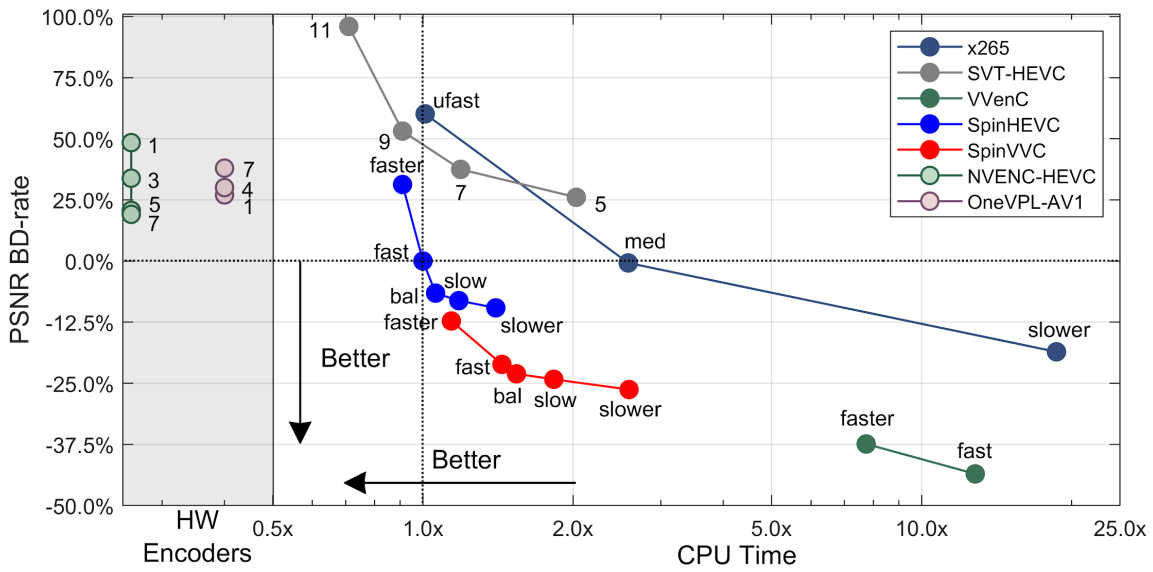
The proposed VVC encoder achieves around 20% compression gains compared to the HEVC baseline at the cost of 44% more computation. The BD-rate gains when comparing *SpinVVC - fast* to *SpinHEVC - fast* are: 21.1% PSNR, 21.6% XPSNR, 17.6% MS-SSIM, and 24.6% VMAF.

Other VVC encoders, designed for offline applications, such as VVenC, achieve higher compression efficiency (*VVenC - fast* obtains 43.5% PSNR BD-rate savings compared to the baseline), but at the cost of 12.8x more computation, which makes it not suitable for live applications; in addition, it does not include the required features for live encoders such as CBR rate control with HRD buffer model.

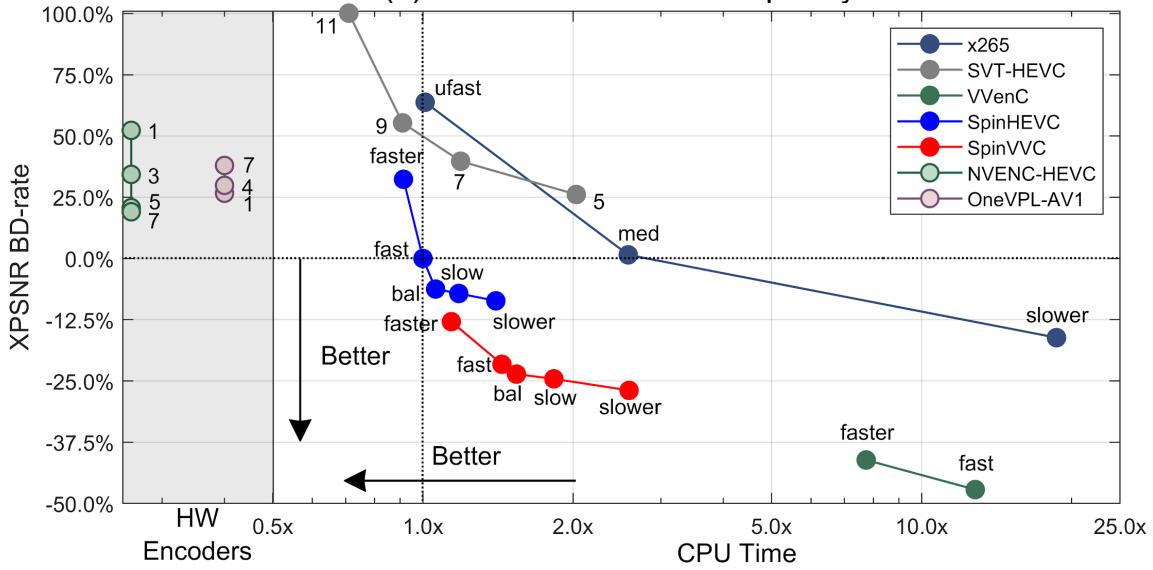
We estimate that the range of real-time operation lies between 1x to 3x more computation than the baseline, which is a highly optimised HEVC encoder that can run at 8K 60 fps.

SpinVVC achieves a compression efficiency similar or slightly higher than HEVC encoders operating at their slow presets. For example, *SpinVVC - fast* PSNR BD-rate gains are 21.14% compared to 18.57% of *x265 - slow*, but *SpinVVC* requires about 13 times lower complexity (a CPU time of 1.44x versus 18.64x).

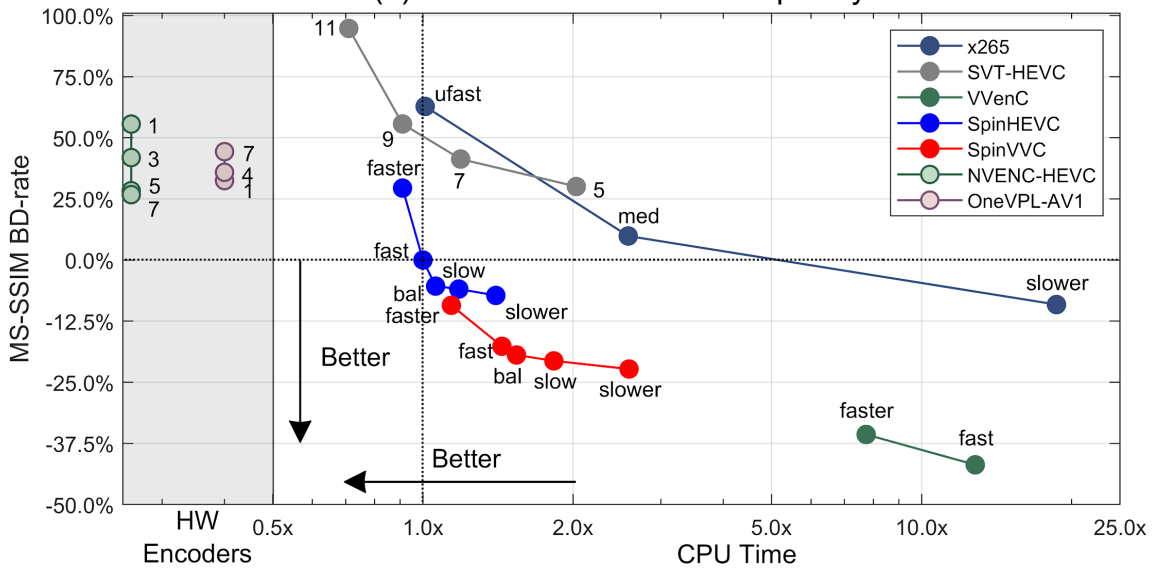
It is also observed that the GPU-based encoders exhibit PSNR BD-rate increases with respect to the HEVC baseline: for NVENC-HEVC, from 19.09% (preset 7) to 48.43% (preset 1), and for OneVPL-AV1 from 27.05% (preset 1) to 37.98% (preset 7). When compared to the proposed VVC encoder (*SpinVVC - fast*), the PSNR BD-rate losses are much higher: between 50% and 89% (NVENC-HEVC) and between 59% and 74% (OneVPL-AV1).



(a) PSNR BD-rate vs Complexity



(b) XPSNR BD-rate vs Complexity



(c) MS-SSIM BD-rate vs Complexity

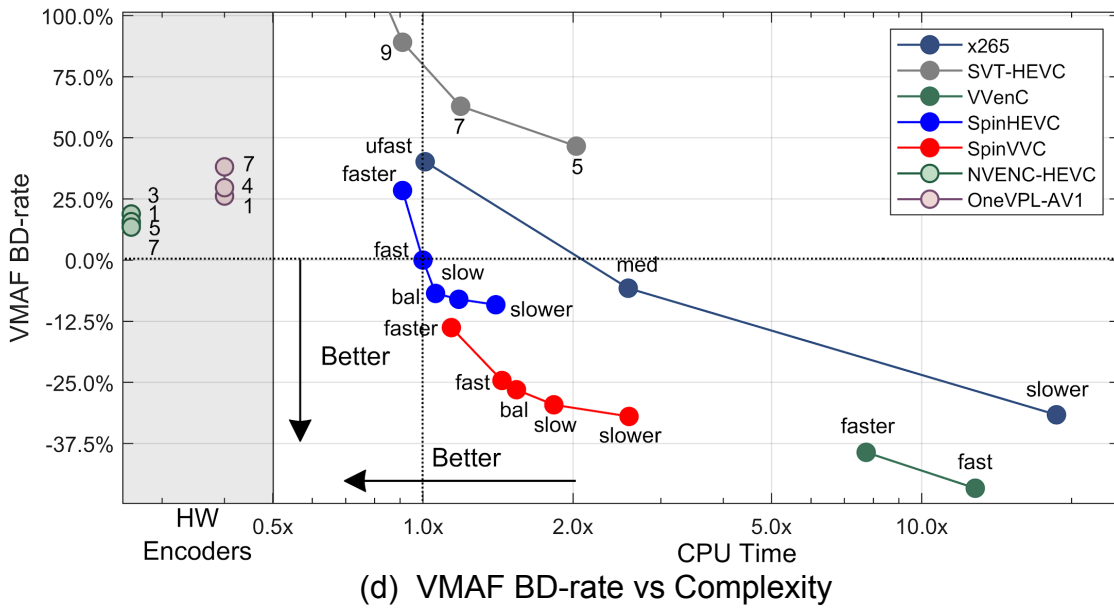
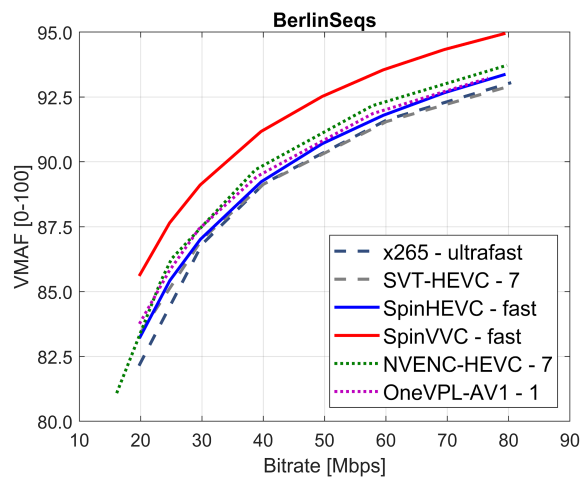
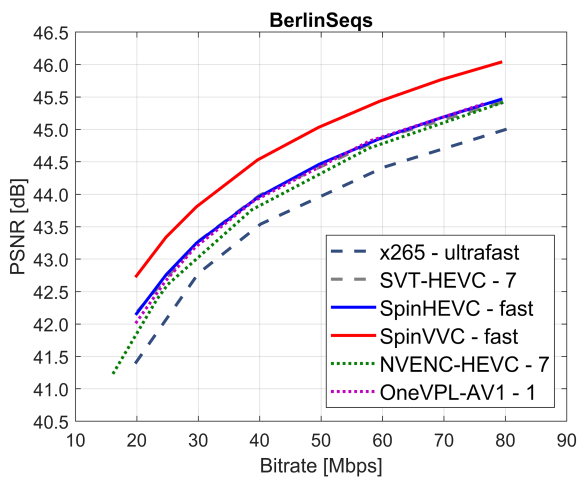


Figure 2 – BD-Rate based on PSNR (a), XPSNR (b), MS-SSIM (c), and VMAF (d), and CPU utilisation time for 8K video relative to the HEVC baseline (*SpinHEVC - fast*)

Selected Quality-Bitrate Plots

Figure 3 depicts the quality-bitrate curves based on PSNR and VMAF for three representative video sequences: *BerlinSeqs*, *FollowCar* and *MC2*. Only the encoders that produce similar complexities, from 1.00x to 1.51x, were included: x265 - ultrafast (1.01x), SVT-HEVC - 7 (1.19x), SpinHEVC - fast (1.00x, the baseline), and SpinVVC - fast (1.44x). As for the hardware encoders, the presets that produce the highest quality were selected. As can be observed, SpinVVC achieves the highest quality at the same bitrate followed by SpinHEVC.



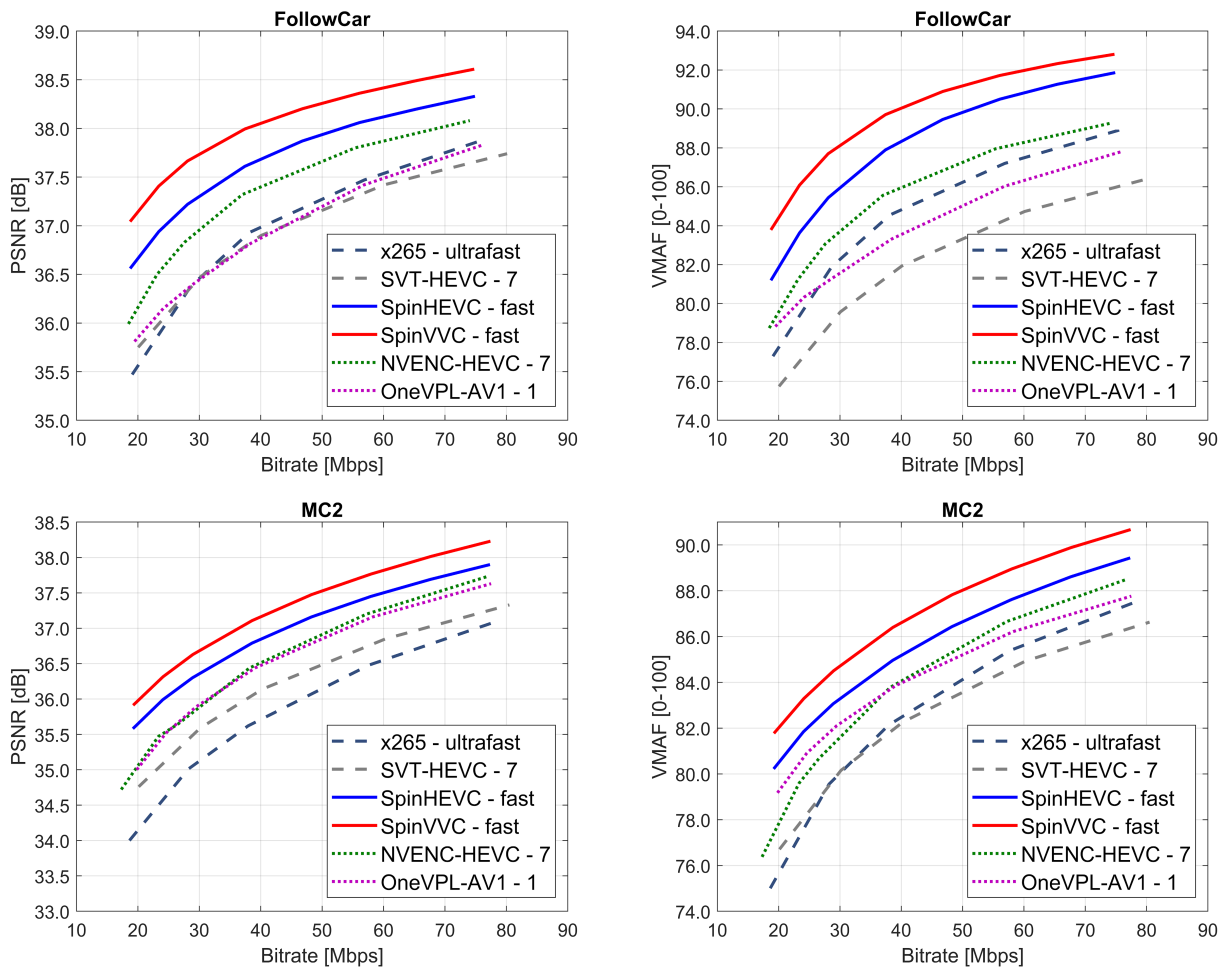


Figure 3 – PSNR-bitrate (left) and VMAF-bitrate (right) curves for BerlinSeqs (top), FollowCar (centre), and MC2 (bottom)

Multithreaded Encoding Speed

The encoding performance, in terms of frames per second (fps), was measured at three different bitrates (20, 40, and 80 Mbps) for the proposed VVC encoder on the two target platforms (ICX and SPR, see Table 3) and compared to the baseline HEVC encoder. *BerlinSeqs* was used in these experiments as it is one of the most challenging videos in terms of encoding performance. The results are presented in Table 6.

The baseline encoder, *SpinHEVC - fast*, produces on the ICX server the encoding speed needed for live 8Kp60 for all the bitrates under analysis (a performance higher than 70 fps has been identified as required for 8Kp60 real-time encoding). The same encoder on the SPR server achieves 96 fps, which corresponds to a 1.2x speedup. These extra CPU resources can also be used to generate higher-quality HEVC video by selecting a more computationally demanding preset. For example, *SpinHEVC - slow* achieves the required performance for 8Kp60 up to 80 Mbps and results in about 8.0% BD-rate savings relative to the baseline.

When running *SpinVVC - fast* on the ICX server, the encoding speed is below the target real-time frame rate. By using the additional resource of the SPR server it achieves a speed of 70 fps at around 30 Mbps (1.3x speedup).



A detailed analysis showed that the VVC encoder on the SPR server is not using the CPU cores at its maximum. To achieve more performance, as mentioned before in the section on encoder parallelism, more parallelism can be enabled using picture partitions or more frames-in-flight.

When adding picture partitions (2x1 vertical partition), the performance increased to 65.6 fps to 69.5 fps at 40 Mbit/s for the VVC - fast preset. When increasing the HRD buffer size from 1 second to 2 seconds, the performance increases from 65.6 fps to 69.7 fps at 40 Mbps (similar to adding picture partitions). The decision of using picture partitions or frames-in-flight is application and platform dependent.

Encoder - preset	Ice Lake (ICX)			Sapphire Rapids (SPR)		
	20 Mbps	40 Mbps	80 Mbps	20 Mbps	40 Mbps	80 Mbps
SpinHEVC - fast	91.10	86.16	78.55	107.21	103.58	96.01
SpinHEVC - slow	72.61	64.20	54.02	88.22	78.91	66.97
SpinVVC - faster	73.61	62.34	49.16	92.72	79.79	65.01
SpinVVC - fast	59.65	51.31	41.14	75.90	65.58	53.78
SpinVVC - fast - 2x1	-	-	-	79.28	69.53	57.35
SpinVVC - fast - 2s buffer	-	-	-	79.36	69.67	58.17

Table 6 – Multithreading encoding performance (in FPS) for the SpinHEVC and SpinVVC encoders with multiple presets in the two platforms under analysis, also including 2x1 picture partitions and larger buffer (default is 1s)

As for the other evaluated encoders, the only one that produces encoding speeds above 60 fps for 8K is *SVT-HEVC - 11* (72 to 79 fps), but at the cost of doubling approximately the bitrate to achieve the same level of quality as SpinHEVC - fast (BD-rate losses: PSNR 95.95%, VMAF 139.65%).

CONCLUSIONS

In this paper, a VVC/H.266 encoder for 8Kp60 10-bit HDR live applications has been described. The encoder is based on a flexible CPU-based software solution that has been highly optimised for the latest generation of CPU systems, with enhancements including SIMD instructions, a scalable multithreaded implementation for systems with many CPU cores, and mode decision and partitioning algorithms to reduce core VVC complexity.

The VVC encoder achieves compression efficiency gains of 24% (BD-rate VMAF) at the cost of 44% more CPU resources when compared to a highly optimised 8K HEVC baseline.

The encoder, running on a 4th Generation Intel Xeon Scalable CPU server, achieves the performance required for 8Kp60 10-bit HDR live encoding at 40 Mbps, making it a viable choice for live 8K applications over constrained bandwidths, such as terrestrial broadcasting or internet streaming. For other applications where higher bitrates are allowed, optimised HEVC encoders have proven to still be a good live encoding option for delivering high-quality 8Kp60 10-bit HDR video.

As future work lines, quality and performance enhancements will be implemented in the 8K VVC encoder and tested on state-of-the-art computing platforms including, among others, perceptually-optimised encoding, live ABR streaming, and content-aware live encoding.



REFERENCES

1. M. Sugawara and K. Masaoka, "UHDTV Image Format for Better Visual Experience," in Proceedings of the IEEE, vol. 101, no. 1, pp. 8-17, Jan. 2013.
2. D. Park, Y. Kim, and Y. Park, "Hyperrealism in Full Ultra High-Definition 8K Display," SID Symposium Digest of Technical Papers, 50: pp. 1138-1141, June 2019.
3. K. Masaoka, Y. Nishida, M. Sugawara, E. Nakasu, and Y. Nojiri, "Sensation of Realness from High-Resolution Images of Real Objects," in IEEE Transactions on Broadcasting, vol. 59, no. 1, pp. 72-83, Mar. 2013.
4. S. Hara, A. Hanada, I. Masuhara, and T. Yamashita, "Celebrating the Launch of 8K/4K UHDTV Satellite Broadcasting and Progress on Full-Featured 8K UHDTV in Japan," in SMPTE Motion Imaging Journal, vol. 127, no. 2, pp. 1-8, Mar. 2018.
5. J. Maglitta, R. Velhal, "Intel's 8K HDR Internet Livestreams: The Future of OTT Sports, Entertainment, and Gaming Broadcasts," VentureBeat, Oct. 12, 2021: <https://venturebeat.com/media/intels-8k-hdr-internet-livestreams-the-future-of-ott-sports-entertainment-and-gaming-broadcasts/>
6. R. Koenen, C. Willock, R. Velhal, R. Cole, F. Lorenceau, D. Poolman, J. Umansky, M. van Eck, "Trailblazing with 8K Live VR at Beijing 2022," IBC2022 Technical Papers, Nov. 2022.
7. Y. Sugito, K. Iguchi, A. Ichigaya, K. Chida, S. Sakaida, H. Sakate, Y. Matsuda, Y. Kawahata, N. Motoyama. "HEVC/H.265 Codec System and Transmission Experiments aimed at 8K Broadcasting," IBC Conference 2015. Sep. 2015.
8. B. Bross, Y-K. Wang, Y. Ye, S. Liu, J. Chen, G. J. Sullivan, J-R. Ohm, "Overview of the Versatile Video Coding (VVC) Standard and Its Applications," IEEE Transactions on Circuits and Systems for Video Technologies, vol. 31, no. 10, pp: 3736-3764, Oct. 2021.
9. KDDI Research, "World's First 8K Video Live Transmission Experiment Using Real-time H.266/VVC Codec," KDDI Research Website, April 19, 2021: <https://www.kddi-research.jp/english/newsrelease/2021/041901.html>
10. Spin Digital, "Spin Digital Releases VVC/H.266 Encoder and SDK for 4K and 8K Live Video," Spin Digital Website, Feb. 20, 2023: <https://spin-digital.com/announcements/vvc-encoder-sdk-4k-8k-live-video/>
11. T. Schierl, M. M. Hannuksela, Y. -K. Wang and S. Wenger, "System Layer Integration of High Efficiency Video Coding," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1871-1884, Dec. 2012
12. N. Nassif et al., "Sapphire Rapids: The Next-Generation Intel Xeon Scalable Processor," 2022 IEEE International Solid- State Circuits Conference (ISSCC), San Francisco, CA, USA, 2022, pp. 44-46
13. David Mulnix. "Technical Overview Of The 4th Gen Intel® Xeon® Scalable processor family". Intel website. July 2022. <https://www.intel.com/content/www/us/en/developer/articles/technical/fourth-generation-xeon-scalable-family-overview.htm>
14. C. C. Chi, M. Alvarez-Mesa, B. Bross, B. Juurlink and T. Schierl, "SIMD Acceleration for HEVC Decoding," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 25, no. 5, pp. 841-855, May 2015



15. C. C. Chi et al., "Parallel Scalability and Efficiency of HEVC Parallelization Approaches," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1827-1838, Dec. 2012
16. Amazon Web Services, "Achieving Broadcast-grade Low Latency in Live Video Streaming," AWS Whitepaper, 2019:
https://pages.awscloud.com/rs/112-TZM-766/images/GEN_elemental-wp-achieving-low-latency.pdf
17. Fraunhofer HHI, "Reference Sequences," 3GPP, 2022:
<https://dash-large-files.akamaized.net/WAVE/3GPP/5GVideo/ReferenceSequences/>
18. Immersify, "Content & Demos - Immersify," Immersify Project Website, 2018:
<https://immersify.eu/content-demos/>
19. Unigine, "Superposition benchmark," UNIGINE Benchmarks, 2017:
<https://benchmark.unigine.com/superposition>
20. The Explorers, "The Earth's first Inventory in High-Definition (4K/8K HDR)," The Explorers Website, 2021: <https://theexplorers.com/svod>
21. ITU-R, "Methodologies for the Subjective Assessment of the Quality of Television Images," Recommendation ITU-R BT.500-14, Oct. 2019.
22. FFmpeg, "Download FFmpeg," FFmpeg Website, 2022. Online [Available]:
<https://ffmpeg.org/download.html#build-linux>
23. Intel, "Scalable Video Technology for HEVC Encoder (SVT-HEVC Encoder)," Open Visual Cloud Github, 2022: <https://github.com/OpenVisualCloud/SVT-HEVC>
24. Nvidia, "NVIDIA Video Codec SDK Documentation," NVIDIA Documentation Center, 2022: <https://docs.nvidia.com/video-technologies/video-codec-sdk/index.html>
25. S. Sanz-Rodriguez, M. Alvarez-Mesa, C. C. Chi, "HEVC Real-time Software Encoder for 8K Live Video Applications". Spin Digital White Paper. Dec. 2020.
<https://spin-digital.com/tech-blog/whitepaper-spin-enc-live/>
26. Intel, "oneAPI Video Processing Library (oneVPL) Dispatcher, Tools, and Examples," OneVPL Github, 2022: <https://github.com/oneapi-src/oneVPL>
27. A. Wieckowski, J. Brandenburg, T. Hinz, C. Bartnik, V. George, G. Hege, C. Helmrich, A. Henkel, C. Lehmann, C. Stoffers, I. Zupancic, B. Bross, D. Marpe, "VvenC: An Open And Optimized VVC Encoder Implementation," 2021 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Shenzhen, China, 2021, pp. 1-2
28. S. Sanz-Rodriguez, M. Alvarez-Mesa, and C. C. Chi, "A VVC/H.266 real-time software encoder for UHD live video applications," JVET-AB0043, October 2022.
29. G. Bjontegaard, "Calculation of Average PSNR Differences between RD curves," ITU-T SG16/Q6 VCEG 13th meeting, Austin, Texas, USA, April 2001, Doc. VCEG-M33.
30. G. Bjontegaard, "Improvements of the BD-PSNR model," ITU-T SG16/Q6 VCEG 35th meeting, Berlin, Germany, 16–18 July, 2008, Doc. VCEG-AI11.
31. C. R. Helmrich, M. Siekmann, S. Becker, S. Bosse, D. Marpe and T. Wiegand, "Xpsnr: A Low-Complexity Extension of The Perceptually Weighted Peak Signal-To-Noise Ratio For High-Resolution Video Quality Assessment," ICASSP 2020 - 2020 IEEE International



Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020, pp. 2727-2731.

32. Z. Wang, E.P. Simoncelli, and A.C. Bovik, "Multi-scale Structural Similarity for Image Quality Assessment," Proceedings of the 37th IEEE Asilomar Conference on Signals, Systems and Computers 2, pp 1398-1402, Nov. 2003.

33. Netflix, "Netflix/vmaf: Perceptual Video Quality Assessment Based on Multi-method Fusion." Netflix GitHub, 2021: <https://github.com/Netflix/vmaf>

34. S. Sanz-Rodriguez, M. Alvarez-Mesa, "Perceptually Optimized Encoding: Benefits, Possibilities, and Difficulties of Assessment with Current Quality Metrics". Spin Digital Technical Blog. May 2023:

<https://spin-digital.com/tech-blog/perceptually-optimized-encoding/>

35. Y. Sugito, Y. Kondo, D. Arai and Y. Kusakabe, "Modeling Perceived Quality on 8K VVC Video Under Various Screen Sizes and Viewing Distances," in IEEE Access, vol. 10, pp. 97237-97247, Sep. 2022.