White Paper

iSIZE BitSave: High-quality video at lower bitrates via Intel DL Boost

- Fast and easily integrated with any encoder including AVC, HEVC, VP9 and AV1, without breaking standards or requiring any changes on the client device
- Single-pass through the content with frame latency
- Supports multi-codec, multi-recipe, multi-bitrate and multi-resolution ABR ladders
- Deep neural network models provide 10 to 25% bitrate savings on top of the codec
- Next generation sustainable results and a significant reduction in costs for video delivery

Authors

This whitepaper is jointly developed by iSIZE and Intel Corporation.

iSIZE

Aaron Chadha

Ilya Fadeev

Yiannis Andreopoulos

Intel

Vasavee Vijayaraghavan

Pradeep B. G.

Raul Diaz

Streaming high-resolution video typically comes with an inevitable trade-off between available bandwidth and quality of experience for the end user. Delivering uncompromising video quality typically requires excessively high bitrates, which can result in slow starts, video buffering and high content delivery network (CDN) and storage costs. As the percentage of IP traffic attributed to video increases (estimated to already surpass 82% [1]), these problems are only exacerbated, driving greater urgency for new innovations to address these challenges.

Traditional solutions that attempt to minimize bandwidth without compromising quality are centered around the development of more intelligent video encoders; either by replacing rate control, quantization and prediction strategies within them, or the entirety of a standard video coding pipeline. The latter, however, is a particularly risky proposition for video encoding services, since it requires the creation of bespoke transport mechanisms and decoders across multiple client device types. Likewise, improvements generated by a standards-based codec remain severely constrained by its inherent compliance needs.

iSIZE's innovative BitSave technology comprises an AI-based perceptual optimization that adaptively filters the video content **before it reaches the codec**, such that the output after compressing with any standard video codec is perceptually optimized (e.g., less motion artifacts or blurring), for the same bitrate. BitSave neural networks are able to isolate areas of perceptual importance, such as those with high motion or detailed texture, and optimize their structure to deliver a much more compelling user experience at lower bitrate. In this way, BitSave is able to **reduce bitrate requirements by up to 25% on top of any existing codec, as shown by a recent paper [2]**. That translates to significant total cost of ownership (TCO) savings for video distribution at similar video quality, as well as a strong economic argument for commercial deployment of enhanced end-user experience. iSIZE accomplished this feat through working closely with Intel to optimize its AI models for Intel Xeon Scalable processors with integrated Deep Learning Boost (Intel DL Boost) and the Intel Distribution of OpenVINO toolkit, powered by oneAPI.

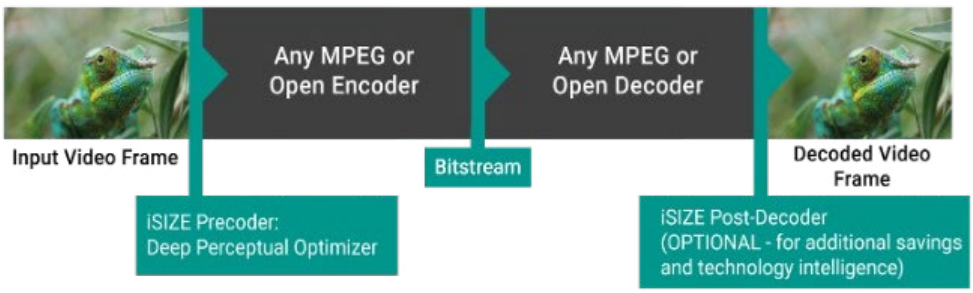

Figure 1. *iSIZE's BitSave preprocessing pipeline*

Importantly, given that BitSave pre-processes video content and does not touch the codec at all, these CPU-accelerated gains are attainable without any of the disadvantages associated with replacing components of a standard codec, as detailed above. This means no changes in encoding streaming infrastructure, or to the client device.

BitSave preprocessing is compatible with any existing codec (including MPEG AVC/H.264, HEVC/H.265, EVC, VVC and AOMedia VP9, AV1, AV2), and supports all resolutions up to 8K. The preprocessing accepts all video formats as input and outputs same resolution pixel-streams in Y4M or MP4, which can be optionally downscaled, before passing to any standard video encoder. Only a single pass processing is required (with single frame latency), prior to any number of subsequent encodings with the video encoder.

iSIZE's proprietary deep learning-based models can run on all Intel hardware and are fully optimized for Intel Xeon Scalable processors, using OpenVINO and advanced features and AI accelerators like AVX-512, VNNI and Intel DL Boost. This was accomplished by partnering with Intel to optimize iSIZE's AI models for Intel Xeon Scalable processors with Deep Learning Boost (Intel DL Boost) and the Intel Distribution of OpenVINO toolkit, powered by oneAPI. Cloud service providers and communications service providers can now realize even better cost savings using iSIZE BitSave technology on servers powered by 2nd generation Intel Xeon Scalable processors with the built-in AI acceleration of Intel DL Boost. The speedup factor for the optimized models is on the order of up to 5x, when compared to non-optimized models running on other CPU instruction sets.[1] The current generation of BitSave models is able to run on 4 CPU cores or higher, achieving **real-time performance on 8 cores of an Intel Xeon Platinum 8259CL 3.1GHz CPU for 1080p 60fps video, and 150 fps for 720p video. In addition, 2160p 60fps video can be handled in real time using a 24-core dual-socket Intel Xeon Platinum 8275CL 3.0GHz CPU.**

Figure 2. *BitSave preprocessing performs a perceptual and rate-constrained decomposition of each frame to ensure that the output of any standard codec is perceptually optimized for a given bitrate*

On typical sequences for VOD and LIVE, involving both static and dynamic content and over multiple QPs/CRFs, BitSave preprocessing is able to offer consistent BD rate gains over multiple presets (ranging from very slow to fast) for a range of standard codecs. Average savings for multiple quality metrics (SSIM, VMAF_NEG, VMAF [2]) are reported in the tables below. More negative values indicate higher bitrate saving when using iSIZE

Sequence	**AVC** (all presets)	**HEVC** (all presets)	**AV1** (cpu=5)
mobcal 720p	-7.9	-8.3	-6.2
parkrun 720p	-17.5	-20.1	-14.6
crowdrun (1080p)	-8.0	-9.0	-12.8
DOTA2 (1080p)	-5.3	-6.5	-12.1
ducks (1080p)	-18.0	-20.1	-23.3
GTAV (1080p)	-5.1	-0.1	-11.4
KristenSara (720p)	0.0	0.0	0.0
Pedestrian (1080p)	-5.1	-4.2	-8.2
Average	**-8.4**	**-8.5**	**-11.1**

Table 1. Representative BD-rates (%) for LIVE video on 8 XIPH sequences. AVC (x264) and HEVC (x265): QPs=[18,22,26,…,42], GOP=150, presets: fast, medium, slow, veryslow and AV1 (aomenc) CPU=5. All BD-rates are average BD-rates of SSIM, VMAF_NEG and VMAF.
Higher negative values indicate higher bitrate saving.

BitSave + codec versus using only the codec with the same recipe. These gains transfer from lower complexity codecs such as AVC, to more intelligent but computationally-demanding codecs such as AV1, and from fast encoding settings like AVC x264 fast to very complex encoding settings like AV1 aomenc cpu=3. On more dynamic content like 'ducks' and 'parkrun', bitrate savings are estimated to range between 20% to 25%, while BitSave is auto- disabled for content that is easy to encode ('KristenSara') and does not need perceptual preprocessing in order to achieve high quality at low bitrate. Average runtimes for BitSave preprocessing on these sequences **range from 64fps for 1080p to 151FPS for 720p video on 8 cores of an Intel Xeon Platinum 8259CL CPU**.

Codec	**AVC** (slow)	**AVC** (very slow)	**AV1** (cpu=3)	**AV1** (cpu=5)
Average BD-rate (%)	-14.35	-14.03	-13.33	-11.62

Table 2. Average BD-rates for ABR VOD ladders, measured on 93 XIPH and CDVL sequences and YouTube UGC 1080p 'LiveMusic' and 'Sport'. All BD-rates are average BD-rates of SSIM, VMAF_NEG and VMAF. The full details of the experimental setup are included in a recent paper [2]. ***Higher negative values indicate higher bitrate saving.***

References

[1] CISCO VNI: https://www.cisco.com/c/dam/m/en_us/solutions/service-provider/vni-forecast-highlights/pdf/Global_2022_Forecast_Highlights.pdf

[2] Chadha, A., Anam, R., Fadeev, I., Giotsas, V., & Andreopoulos, Y. (2020, August). Escaping the complexity-bitrate-quality barriers of video encoders via deep perceptual optimization. In Applications of Digital Image Processing XLIII (Vol. 11510, p. 115100C). International Society for Optics and Photonics, preprint available online at https://www.isize.co/SPIE2020.pdf.

[3] iSIZE at Network Builders: https://networkbuilders.intel.com/ecosystem/isize-technologies

[4] Intel Visual Cloud: www.intel.com/visualcloud

Notices & Disclaimers

Performance varies by use, configuration and other factors. Learn more at www.Intel.com/PerformanceIndex.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Results have been estimated or simulated.

Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Configurations

[1] Performance compared the standard iSize BitSave tool to iSize BitSave optimized with OpenVINO and the OpenVINO Post Optimization Tool. Tested by Intel from May 1, 2020 to September 30, 2020. The optimized BitSave tool had up to 1.8x gains using the 32-bit floating point version and up to 5x gains using the 8-bit integer version. Intel's test used 9 sequences, consisting of a mix of 720p,1080p and 4k resolutions running on Dual socket 2nd generation Intel Xeon Scalable Processor 8280 running at 2.7GHz base frequency with 205W TDP with 2933MHz DIMMS with 1DPC. The performance is calculated as the average number of frames-per-second that the BitSave tool could process during run time of all the frames across all the test clips.

For More Information

iSIZE Technologies: https://www.isize.co **Intel Network Builders:** https://networkbuilders.intel.com