JPEG XS COMPRESSION FOR VIDEO TELEMETRY

Stephen Schaphorst
Delta Digital Video
Horsham, PA, USA

George Nelson
Delta Digital Video
Horsham, PA, USA

ABSTRACT

The Joint Photographic Experts Group (JPEG) XS video compression standard provides visually lossless imagery at typical compression ratios of 10:1, allowing transmission of multiple channels of high-definition video over GbE networks. As video bandwidths continue to increase with higher definition and higher frame rate sources, visually lossless encoding with sub-millisecond encode/decode latencies, a networked infrastructure, and a reliably constant bitrate transport stream, becomes a compelling alternative to uncompressed video. An overview of JPEG XS and transport methods is presented along with targeted test platform applications, including ground network video distribution and AI-enabled object detection-based Region of Interest (ROI) use cases for bandwidth-limited data links. Comparisons and tradeoffs with other light and higher compression technologies are discussed.

KEYWORDS

JPEG XS, Video Compression, Visually Lossless, Latency

INTRODUCTION

An overview of the Joint Photographic Experts Group (JPEG) XS video compression standard [1-3] is presented along with potential use cases for the mezzanine compression system in video telemetry. Use cases and requirements for JPEG XS are described, as well as targeted processing platforms. Efforts by the JPEG XS committee to create a viable alternative to uncompressed video include provisions for such features as visually lossless imagery, low codec complexity, precise constant bitrate control, and sub-millisecond latency. Comparative studies of JPEG XS with other codecs are reviewed. An overview of JPEG XS transport protocols is presented, including MPEG-2 Transport Stream (TS) [4], as well as SMPTE Standards 2022-2 [5] and 2110-22 [6] for transport over Internet Protocol (IP) networks. File formats for the storage of JPEG XS video and single images are listed. Potential use cases for JPEG XS for video telemetry are described including ground video distribution networks, video recording and transmission applications, and region-of-interest streaming for remote vehicles or platforms with SWaP-C constraints.

OVERVIEW OF JPEG XS

This overview focuses on Parts 1, 2, and 3 of the JPEG XS specification: the core coding system [1], profiles and buffer models [2], and transport and container formats [3]. The working group for the JPEG XS core coding system identified key use cases and features [7,8], including (1) visually lossless compression, (2) ultra-low latency, with a maximum of 32 video lines of end-to-end algorithmic latency, (3) low complexity implementations, allowing 4Kp60 encoding on a standard i7 x86 processor, or on a low-cost FPGA without requiring external memory, and (4) encoder robustness to seven or more encode-decode cycles to support intermediate processing workflows such as cropping, editing, and overlay insertion. Initial image formats to be supported included VGA (640x480) to 8K (7680x4320), with RGB and YCbCr component types at 4:2:2 and 4:4:4 component subsampling, component bit depths from 8 to 12 bits per component (bpc), and frame rates from 24fps to 120fps.

In 2022 the second edition of the JPEG XS specification added several profiles, including Mathematically Lossless and raw Bayer profiles, support for 4:2:0 subsampling, and bit depths of 16 bpc. A third edition is planned for release in 2024 and is expected to include Temporal Differential Coding (TDC) profile, targeting gaming and remote desktop use cases, to allow compression ratios of 20:1 by adding a frame buffer with allowances for variable bitrate.

The JPEG XS compression system consists of the classical rate-distortion tool chain of color and spatial decorrelation, transform quantization, and entropy coding of quantization indices. The methods employed for these steps were carefully selected to allow for ultra-low latency and low complexity while allowing visually lossless encoding at significant compression ratios [9]. The JPEG XS color transform is the same Reversible Color Transform as used for JPEG 2000 compression. JPEG XS spatial decorrelation uses an asymmetric Discrete Wavelet Transform (DWT), typically operating on a narrow number of video lines at a time. The low latencies achievable by the algorithm come from an architecture offering a high degree of parallelism for both the data paths with vector instructions and execution paths with multi-core processing. Rate allocation proceeds over a window of the incoming video lines, using an output smoothing buffer to monitor the buffer fill level and provide available capacity to the compression engine, maintaining precise bitrate control and a maximum end-to-end latency of 32 lines.

Performance Evaluation

JPEG XS has been optimized for visually lossless compression as specified in ISO/IEC standard 29170-2:2015, which details a subjective evaluation procedure [10]. The procedure defines a "flicker test" method to measure whether an observer can distinguish uncompressed original reference imagery from compressed or processed imagery. The procedure interleaves original and processed images to measure whether the observers detect a flickering in the video. JPEG XS was tested with a variety of imagery over compression ratios from 3:1 to 10:1. JPEG XS was found to pass the flicker testing for a medium compression ratio of 6:1 for all imagery tested except for the most complex [10].

The overview of Descampe et al [11] includes an objective assessment of JPEG XS compared to an equivalent low-latency extension of JPEG 2000 (referred to as J2K LL) [12], VC-2 [13], and ProRes [14] based on PSNR measurements, along with a complexity analysis. The objective analysis compares the JPEG XS Main 444.12 and High 444.12 profiles with J2K LL configurations for bitrates from 2 to 8 bits per pixel (bpp), or compression ratios from 12:1 to 4:1. Similar results were reported for other video formats, such as 4:2:2 10-bit source imagery.

J2K LL configurations were chosen to maintain high quality with similar latencies as JPEG XS. The JPEG XS and J2K LL PSNR results are similar, with J2K LL slightly outperforming JPEG XS at the expense of higher latency and significantly higher complexity. The comparison of JPEG XS with ProRes and VC-2 shows ProRes with lower quality at medium compression ratios (3 to 6 bpp), and similar quality levels at lower and higher compression ratios. VC-2, configured for low latency and low complexity, does not reach the same quality levels as JPEG XS and J2K LL. A lower complexity version of JPEG 2000 called High Throughput JPEG 2000 (HTJ2K) was standardized in 2019 [15]. A comparison between HTJ2K and several codecs, including a non-optimized version of JPEG XS has been presented [16], indicating in most cases JPEG XS PSNR values between HTJ2K 9/7, 1 VWT (vertical wavelet transform levels) values and 2 VWT values. One advantage of JPEG XS compared to HTJ2K as a replacement for uncompressed video is the method for precise output bitrate control. JPEG XS was also compared for robustness to encode/decode cycling with J2K LL and an HEVC at a similar profile. At a bit rate of 4 bpp J2K LL and JPEG XS maintained a constant quality over 10 cycles, while HEVC decreased in quality by over 2dB over 8 cycles.

JPEG XS Transport and Containers

Specifications for the transport of JPEG XS include the carriage of JPEG XS in an MPEG-2 transport stream [4] as well as specifications for transport over IP networks using SMPTE ST 2022-2 [5] and SMPTE ST 2110-22 [6].

SMPTE ST 2022-2 is a specification for the transport of constant bitrate MPEG-2 transport streams over RTP. The SMPTE 2022 suite of media-over-IP standards includes specifications for forward error correction as well as seamless protection switching. The Video Services Forum (VSF) technical recommendation TR-07 [17] calls out SMPTE 2022-2 and further defines profiles for the transport of JPEG XS video and associated audio and metadata, with wide area network (WAN) applications as the primary use case.

SMPTE 2110-22 is a specification for the transport of constant bitrate compressed video elementary streams in RTP packets. The SMPTE 2110 suite of standards specifies the transport of compressed or uncompressed video and associated audio and metadata as a system of RTP-based essence streams with a common reference clock. VSF TR-10, the Internet Protocol Media Experience (IPMX) set of technical recommendations, is based on the SMPTE 2110 standards and the AMWA Networked Media Open Specifications (NMOS) [18] which support network management, discovery, and registration. VSF TR-10-11 [19] is the technical recommendation that describes the transport of constant bitrate compressed video in the IPMX protocol suite.

JPEG XS video streams and single images can be stored using file formats described in the table below.

File Format	**Imagery Content**	**Specification**
Material Format MXF (.mxf)	Video	SMPTE ST 2124 [20]
ISO Base Media File Format (.mp4)	Video	ISO/IEC 21122-3 [3], and references therein
MPEG-2 Transport Stream (.ts)	Video, single images	ISO/IEC 13818-1 Ed. 8 Annex W [4]
High Efficiency Image Format (.heif)	Video, single images	ISO/IEC 21122-3 [3], and references therein
JPEG XS (.jxs)	Single images	ISO/IEC 21122-3 [3]

Note that the MPEG-2 TS specification includes the transport of JPEG XS still images which could be used for the real-time object detection use case discussed below.

JPEG XS USE CASES FOR VIDEO TELEMETRY

This section focuses on three potential use cases of JPEG XS for video telemetry systems. The first use case presents JPEG XS as an alternative to uncompressed or H.264/H.265 video for ground video distribution. The second case highlights the benefits of JPEG XS compression for recording and subsequent transmission on the remote vehicle or platform, particularly for transmitting systems with significant size, weight, power, and cost (SWaP-C) constraints. The third case explores options for real-time region-of-interest monitoring and video processing at the telemetry link receiving system.

Ground Network Video Distribution

The first use case addresses ground video distribution requirements for ultra-low latency visually lossless video, historically fulfilled by the transmission of uncompressed video over SDI. JPEG XS presents an option to replace SDI video with GbE networked distribution of multi-channel high-definition video or single-channel 4Kp60 video stream. JPEG XS 10:1 compression of 4Kp60 8 bps 4:2:2 video has a bitrate less than 800Mbps. Low-power JPEG XS video encoders co-located with sensors allow video transmission over GbE links and networked switching to multiple range real-time monitoring and recording stations.

Video Recording, Post-Mission Transmission

The second use case concerns missions involving high bitrate video recording followed by post-mission transmission. The benefits of a visually lossless lightweight 10:1 compression technology like JPEG XS for this use case compared to the recording of uncompressed video are described in the RCC SR-22-002 Image Compression report [21]. The report details significant savings in time and resources for Kineto Tracking Mount-based workflows compared to recording and downloading uncompressed video, as well as diminishing returns for further compression and complexity from H.264 or H.265 compression.

For video recording and playback applications, vehicles or platforms with SWaP-C constraints can benefit from JPEG XS compared to uncompressed or more complex compression technologies in several ways:

- The low complexity of JPEG XS compression reduces the power consumed and cost compared to more complex compression technologies
- Compression ratios up to 10:1 permit longer recording times, shorter download times compared to uncompressed video, as well as the advantage of lower speed interfaces with recording media
- Vehicles or platforms with raw Bayer sensors can take advantage of JPEG XS raw Bayer profile that saves the extra video processing step in converting raw Bayer to YCrCb or RGB color formats

Real-Time Visually Lossless Video Telemetry

There has been a rapid growth in the use of unmanned vehicles and object detection for safety, security, and surveillance. With the advent of fast object detection algorithms such as You Only Look Once (YOLO) and Single Shot Detection (SSD), as well as deep learning algorithms with light-weight architectures, object detection is increasingly performed on remote vehicles and platforms. However, for low-power vehicles or scenarios where the accuracy and determination of objects are critical there can be an advantage to offloading processing to the ground station and connected servers. An area of study has developed for best methods and architectures for edge-assisted video processing and object detection with respect to power consumption of the vehicle and the latency and reliability of detection results [22].

The visually lossless, low complexity, and low latency attributes of JPEG XS compression make it a good fit for these use cases. The limited datalink bandwidth prevents the transmission of full frame rate high-definition video at a compression ratio of 10:1. By narrowing the pixel range or decimating the frame rate, JPEG XS can be used for encoding and transmission of high-quality imagery from remote sensors to the receiving system. The table below lists sample bitrates, ranging from 20.7Mbps for 1080p video at 5fps to 2.3Mbps for 320x240 region-of-interest video at 15fps at a compression ratio of 10:1, a component bit depth of 10 bits, and 4:2:2 chroma sampling.

Video Resolution	Frame Rate (fps)	JPEG XS Bitrate (kbps) at 10:1 Compression
1920x1080	5	20,736
1280x720	5	9,216
640x480	10	6,114
640x480	5	3,072
320x240	15	2,304

Real-time transmission of high-quality video imagery permits receiving systems to perform high-power video processing object detection methods, geolocation, and unique object identification with access to large databases. Significant uncertainties in object classification or identification could elicit snapshot requests from the ground station for particular regions of interest.

CONCLUSION

An overview of the JPEG XS video compression standard is presented, including a description of MPEG-2 TS and RTP transport methods as well as container formats. JPEG XS performs favorably in a review of comparisons with other similar compression technologies with respect to image quality and algorithm complexity. The JPEG XS video compression standard, offering visually lossless imagery at compression ratios up to 10:1, ultra-low latency, and low complexity, can be a useful technology for video telemetry systems. Three use cases are described: (1) a ground GbE video distribution application as a networked replacement of SDI video coax or fiber distribution, (2) a video on-board recording and download application highlighting the benefits of 10:1 compression for download time compared to uncompressed video, and (3) a real-time JPEG XS video transmission case to support reliable object detection. The three use cases can benefit from the low complexity of the JPEG XS algorithm for remote platforms with SWaP-C constraints.

REFERENCES

[1] ISO/IEC 21122-1:2022, "Information technology — JPEG XS Low-latency Lightweight Image Coding System", Part 1: Core coding system," 2022.

[2] ISO/IEC 21122-2:2022 "Information technology — JPEG XS Low-latency Lightweight Image Coding System — Part 2: Profiles and buffer models," 2022.

[3] ISO/IEC 21122-3:2022 "Information technology JPEG XS Low-latency Lightweight Image Coding System — Part 3: Transport and buffer formats," 2022.

[4] ISO/IEC 13818-1:2021: "Information Technology - Generic Coding of moving pictures and associated audio information - Part 1: Systems," Ed. 8, Annex W, 2021.

[5] SMPTE ST 2022-2:2007 "Unidirectional Transport of Constant Bit Rate MPEG-2 Transport Streams on IP Networks," 2007.

[6] SMPTE ST 2110-22:2022, "Professional Media Over Managed IP Networks: Constant Bit-Rate Compressed Video," 2022.

[7] ISO/IEC JTC 1/SC 29/WG1, "Use cases and requirements for ISO/IEC 21122-1 (JPEG XS Part-1, Core Coding System) v1.3," February 2018.

[8] ISO/IEC JTC 1/SC 29/WG1, "Use cases and requirements for ISO/IEC 21122-1 (JPEG XS Part-1, Core Coding System) v2.1," Oct. 2020.

[9] T. Richter, S. Foessel, J. Keinert, A. Descampe, G. Rouvroy, J.B. Lorent, "JPEG-XS - A high quality mezzanine image codec for video over IP," in SMPTE 2017 Annual Technical Conference and Exhibition (pp. 1-15), October 2017.

[10] ISO/IEC 29170-2: 2015, "Information technology — Advanced image coding and evaluation, Part 2: Evaluation procedure for near lossless coding," 2015.

[11] A. Descampe, T. Richter, T. Ebrahimi, S. Foessel, J. Keinert, T. Bruylants, P. Pellegrin, C. Buysschaert, G. Rouvroy, "JPEG XS - A New Standard for Visually Lossless Low-Latency Lightweight Image Coding," Proceedings of the IEEE, Volume: 109, Issue: 9, September 2021.

[12] ISO/IEC 15444-2:2004, ITU-T Rec. T.801, "JPEG 2000 image coding system: extensions," 2004.

[13] SMPTE ST 2042-1:2017, "VC-2 Video Compression," 2017.

[14] SMPTE Engineering Project (ANSI), "Apple ProRes Bitstream Syntax and Decoding Process. SMPTE registered disclosure document RDD 36," 2015.

[15] ISO/IEC 15444-15:2019, ITU-T Rec. T.814, "JPEG 2000 image coding system - part 15: High-throughput jpeg 2000," 2019.

[16] T. Edwards and M. Smith, "High Throughput JPEG 2000 for Broadcast and IP-based Applications," in SMPTE 2020 Annual Technical Conference and Exhibition, 2020.

[17] Video Services Forum (VSF) TR-07:2022, "Transport of JPEG XS Video in MPEG-2 Transport Stream over IP," 2022.

[18] AMWA Networked Open Media Specifications, https://specs.amwa.tv/nmos/

[19] Video Services Forum (VSF) TR-10-11:2022, "Internet Protocol Media Experience (IPMX): Constant Bit-Rate Compressed Video," February 23, 2024.

[20] SMPTE ST 2124:2020, "Material Exchange Format – Mapping JPEG XS Codestreams into the MXF Generic Container," 2020.

[21] Range Commanders Council, Optical Systems Group, SR-22-002, "Image Compression," September 2023.

[22] P. Amanatidis, D. Karampatzakis, G. Iosifidis, T. Lagkas, A. Nikitas, "Cooperative Task Execution for Object Detection in Edge Computing: An Internet of Things Application," Applied Sciences, 13, 4982, April 2023.