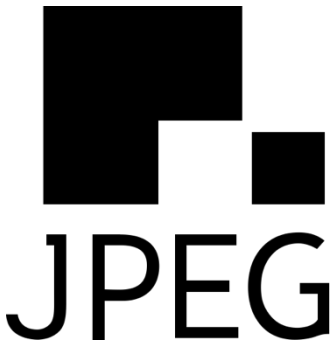

JPEG White paper:
JPEG XS, a new standard for visually lossless low-latency lightweight image coding system

Version 2.0

ISO/IEC JT1/SC29/WG1
WG1N83038

Editors (*in alphabetical order*):
Charles Buysschaert – IntoPIX SA, Belgium
Antonin Descampe – IntoPIX SA, Belgium
Siegfried Fößel– Fraunhofer IIS, Germany
Joachim Keinert – Fraunhofer IIS, Germany
Pascal Pellegrin – IntoPIX SA, Belgium
Thomas Richter – Fraunhofer IIS, Germany
Gaël Rouvroy – IntoPIX SA, Belgium

Table of Contents

Executive Summary

JPEG XS is a new International Standard from the JPEG Committee (formally known as ISO/IEC SC29 WG1). It provides visually lossless image compression thanks to an interoperable low-latency lightweight coding system that can be used as a mezzanine codec within any AV market. Among the targeted use cases are video transport over professional video links (SDI, IP, Ethernet), real-time video storage, memory buffers, omnidirectional video capture and rendering, and sensor compression (for example in cameras and in the automotive industry). This implies optimal support of a wide range of implementation technologies such as ASICs, FPGAs, CPUs and GPUs. Moreover, the codec allows to accurately set the targeted bitrate to match the available bandwidth, e.g. Gigabit Ethernet, and it offers an end-to-end delay equal to a fraction of a frame. Typical compression ratios are up to 10:1 for both 4:4:4 and 4:2:2 images but can also be higher depending on the nature of the image or the requirements of the targeted application. This white paper presents the main use cases and key features of this new standard.

1 Introduction

Video bandwidth requirements are growing fast, as video resolutions, frame rates and numbers of streams to manage are constantly increasing. Capacity of video links and communication channels are growing too, though slower and, more importantly, through huge investment that needs to be amortized over several years. Transport of video streams in the broadcast industry illustrates this situation very well: manufacturers are currently switching from HD to UHD requiring 12G-SDI or 10 Gbps Ethernet, and 8k is already being looked at. As a consequence, uncompressed storage and live video transmission becomes unaffordable and unmanageable within current systems and infrastructures, while next generation channels are still being tested or not yet affordable. Facing this reality, the use of lightweight compression is very attractive as it allows to smooth the everlasting transition between successive generations of infrastructures and protocols.

Such lightweight compression scheme should allow increasing resolution, frame rate and number of streams while safeguarding all advantages of an uncompressed stream, i.e. interoperability, visually lossless quality, low power consumption, low latency in coding and decoding, ease of implementation, small size on chip and fast software running on general purpose CPU. In this context, the JPEG Committee (formally known as ISO/IEC SC29 WG1) has standardized a lightweight codec called JPEG XS.

2 Use cases and targeted applications

In a nutshell, JPEG XS is a candidate technology wherever uncompressed video is used today. However, it has been specifically designed to meet the requirements of live production, broadcast and digital cinema workflows, Pro-AV markets, Virtual Reality (VR) gaming, and sensor compression [1].

2.1 Transport over video links and IP networks

In the Broadcast studios, SDI infrastructures are massively deployed, mainly HD-SDI (1.5 Gbit/s) or 3G-SDI (3 Gbit/s). However, broadcasters are now moving progressively to IP. Preference currently goes to 1 Gigabit Ethernet [GE] links for remote production or 10 GE infrastructures for in-house studios. 25, 40 or 100 GE links are usually not yet affordable. Given the available bandwidth, truly uncompressed video is therefore not an option anymore, both in the studio and in remote production, as 4K, 8K, and higher framerate [fps] need to be supported.

A light compression that could visually preserve the quality of an uncompressed stream therefore appears to be a good solution to transport upcoming video streams with infrastructures (some soon-to-be) deployed, as shown in Table 1.

Table 1: Typical JPEG XS use cases for video transport over IP

	1GE IP	**10GE IP**	**1GE IP + JPEG XS**	**10GE IP + JPEG XS**
2K 60P 4:2:2 10-bit	✖	✔ up to **3** streams	✔ up to **2** streams with 6 :1 compression	✔ up to **12** streams with 4 :1 compression
4K 60P 4:2:2 10-bit	✖	✖	✔ **1** stream with 12 :1 compression	✔ up to **3** streams with 4 :1 compression
8K 60P 4:2:2 10-bit	✖	✖	✖	✔ **1** stream with 5 :1 compression

Besides visually lossless quality, robustness to multiple encoding and decoding cycles is also critical, as it allows chaining of multiple devices that recompress the signal. The JPEG Committee tested the robustness of JPEG XS for at least 10 encoding-decoding cycles. Moreover, the additional latency introduced by one coding and decoding cycle should be below a couple of lines in order to avoid any human-perceptible delay between signals processed by different processing chains.

2.2 Real-time video storage

Similar to transport of video streams, storage of high-resolution streams requires lightweight compression to allow them to be written in real-time on lower cost storage devices. For instance, cameras use internal storage like SSD drives or SD cards to store large streams of images, but access rates on these devices are limited. Moreover, as for the video transport case, multiple encoding-decoding cycles must deliver the same quality as a single compression and decompression operation.

2.3 Frame buffer compression

Enabling lightweight compression for buffers inside video processing devices can drastically reduce the system's form factor, decrease the number of interconnect wires and extend the battery life for battery powered systems. For instance, JPEG XS could be used in the following contexts:

- Buffer for high refresh rate displays (120 to 600 Hz)
- Storage and replay buffer for high speed cameras
- Reference frame buffer inside AVC/H.264 or HEVC/H.265 hardware codecs.

This last use case (reference frame buffer compression) is of particular interest as portable video devices use external SDRAM chips to store the reference frames involved in the inter-frame prediction mechanisms. Power consumption of such SDRAM is proportional to the required access bandwidth: upcoming UHD resolutions (4K or 8K) will therefore increase the overall power consumption of such codec in a way unaffordable for power-aware designs. Very low complexity compression appears to be an elegant solution to keep consumption inside reasonable limits.

2.4 Omnidirectional video capture and rendering

JPEG XS is also intended to be used in head-mounted displays for Virtual or Augmented Reality (VR/AR). To get an immersive experience, displays with resolutions above 8 Megapixels and 90 frames per second and per eye are necessary. Such applications require a very low latency coding scheme so as to ensure a tight synchronization between movement and display.

2.5 Sensor compression

More and more image sensors are used in industrial environments with increased resolution. In this context, JPEG XS offers a convenient way to ensure transport of image sequences within industrial networks.

For instance, recent developments of the automotive sector involve the use of an increasing number of image sensors with high resolution and framerates. These sensors typically capture the image by using a Bayer pattern. Data from these sensors need to be processed with a maximum responsiveness, therefore implying a very low latency along the whole dataflow. In terms of implementation and given the number of sensors, power consumption needs to be constrained as much as possible because of thermal considerations and the necessary operation in all kind of climatic conditions.

3 Key features of the JPEG XS codec

Based on the above-described use cases, the following requirements have been identified, and are now the key features of this new standard [1].

- **Visually lossless quality** with imperceptible flickering between original and compressed image (as defined in [2]). Typical compression ratios are up to 10:1 for both 4:4:4 and 4:2:2 images with up to 12-bit component precision but can also be higher depending on the nature of the image or the requirements of the targeted application.
- **Multi-generation robustness**, i.e. no significant quality degradation for up to 10 encoding-decoding cycles.
- **Multi-platform interoperability**. The JPEG XS use cases require real-time implementations on several different platforms: CPU, GPU, FPGA and ASIC. Each of these platforms is best exploited when a specific degree of parallelism is available in the implemented codec. For instance, a multi-core CPU implementation will benefit from a coarse-grained parallelism while GPU or FPGA will better take advantage of a fine-grained parallelism. Therefore, to optimally support the different target platforms, the JPEG XS codec needs to allow for different kinds of end-to-end parallelization. More importantly, real-time encoding on a given platform (a FPGA for instance, exploiting a fine-grained parallelism) shall allow real-time decoding of the generated codestream on any other platform (including for instance a multi-core CPU exploiting a different kind of parallelism), without sacrificing neither the low complexity nor the low latency features described below.
- **Low complexity**, both in hardware and software. For JPEG XS to be a legitimate candidate to replace uncompressed video transport, very low complexity implementations need to be achievable. In practice, on the software side, JPEG XS has been designed so that an i7 processor is able to process 4k 4:4:4 60p content in real time. On the hardware side, FPGA implementations should not require any external memory and should not occupy more than 50% of Artix7 XC7A200T or 25% of a Cyclon5 5CEA9 when applied to 4k 4:4:4 60p content.
- **Low latency**. As indicated above, whether it be in video transport applications (especially live production), in AR/VR applications, or in any other use case requiring a tight synchronization between the signal and a human interaction, the cumulated delay required by all processing steps the signal has to go through has to be kept below the human perception threshold. To this end, and based on inputs from the different application fields, JPEG XS offers a scalable algorithmic latency, ranging from a small number of lines down to below a single line for a combined encoder-decoder suite.

4 Comparison to state-of-the-art codecs

Based on the requirements above, it is easy to see that existing standards do not comply with the needs of film and broadcast applications. JPEG-LS (ISO/IEC 14495-1 | ITU-T Rec. T.87) and JPEG (ISO/IEC 10918-1 | ITU-T Rec. T.81) as well as its successor JPEG-XT (ISO/IEC 18477-1), which provides backward compatible support of higher bit depths, make a precise rate control difficult, and typical implementations show a latency of one frame. JPEG 2000 (ISO/IEC 15444-1:2004 | ITU-T Rec. T.800) uses a complex entropy coder, implying many hardware and software resources for real-time implementations. HEVC (ISO/IEC 23008-2 | ITU-T Rec. H.265), as a distribution codec needs a huge encoding complexity without ensuring multi-generation robustness. VC-2 (SMPTE ST 2042-1) on the other hand is of low complexity, but the applied technology only delivers limited image quality. ProRes, as documented by a SMPTE disclosure document (SMPTE RDD 36), is based on macro blocks of 16x16 pixels, making a low latency implementation below 32 lines impossible. Moreover, the symbol wise entropy coding makes fast CPU implementations challenging. DSC (from VESA) finally targets ASIC-based display compression, making efficient implementations on FPGAs and GPUs hard to achieve.

Considering these shortcomings of the existing codecs, the JPEG committee has elaborated JPEG XS, that provides a precise rate control with a latency below 32 lines and that fits in a low cost FPGA. The compression quality was requested to be superior to VC-2 while supporting implementation on different platforms.

5 JPEG XS profiles and formats

The JPEG XS Core Coding System uses a Discrete Wavelet Transform (DWT) followed by the entropy coding of the amplitude level of groups of 4 consecutive coefficients, while actual coefficients values are raw-coded. To allow different levels of latency and complexity, several profiles, each with maximum vertical wavelet decompositions, are defined. Each profile targets specific applications, as described in Table 2.

Table 2: JPEG XS profiles

Main	Light	Light-subline	High
• Default profile • 1 vertical DWT • Natural / CGI / Screen content • Broadcast, Pro-AV, Frame Buffers, Display links	• Less complex • Slightly less efficient • 1 vertical DWT • Natural content • Broadcast, industrial cameras, in-camera compression	• Less buffering • Less efficient • No vertical DWT • Natural content • Cost-sensitive applications	• More quality • 2 vertical DWT • Natural / CGI / Screen content • Same applications as « Main » but for high-end devices, contribution, cinema remote production

Beside the Core Coding System and profiles, the JPEG Committee has also standardized different transport and container formats for a JPEG XS codestream, to allow storage and transport of JPEG XS images within different frameworks and using different protocols, as described in Table 3.

Table 3: JPEG XS transport and container formats

Format	Type	Description – Main purpose	extension	Standard document
JXS	JPEG XS Fileformat	For storing of single images	.jxs	ISO/IEC 21122-3
MP4	ISO Base Media File format (ISOBMFF)	For storing of video	.mp4	ISO/IEC 21122-3
HEIF	High Efficiency Image File Format	For storing of mixed image and video content	.heif	ISO/IEC 21122-3
MPEG-2 TS	Transport stream	MPEG-2 Transport stream for JPEG XS	n/a	ISO/IEC 13818-1 Ed. 7th AMD1
RTP	RTP payload format	IP transport for JPEG XS	n/a	Upcoming RFC at IETF
SMPTE 2110	System stream	Encapsulation of compressed video stream in SMPTE 2110	n/a	Upcoming SMPTE 2110-22

6 Performance

In order to validate that this newly developed codec meets all requirements, the JPEG Committee has analyzed different technologies submitted after the Call for Proposals. Complexity and latency have been evaluated [3], as well as quality, and several rounds of Core Experiments have further improved the specification. For quality, both objective [4] and subjective [5][6] evaluations have been made with various kinds of content, including natural, CGI, screen or “pathological” images. Results have shown

that the visual transparency requirement is met within the targeted compression range. Figure XX shows a representative example of how JPEG XS behaves compared to VC-2 or a low-latency flavour of JPEG 2000 (using tiles with a height of 8 lines and a tile-based rate allocation). As seen there, JPEG XS outperforms VC-2, and its 2 vertical wavelet transforms profile is usually on par with JPEG 2000, with sometimes one ahead and sometimes the other.

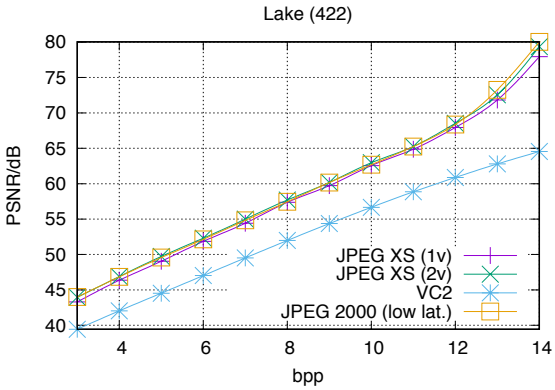

Figure 1: Sequence-PSNR for the 4:2:2 10-bit sequence “Lake”. Experiments are described in more details in [4].

7 Conclusion

JPEG XS is a new International Standard for visually lossless low-latency lightweight image coding, designed to compensate for continuously increasing bandwidth requirements in video transport links. It is a candidate technology wherever uncompressed video is used today. A simple yet efficient coding scheme allows to keep latency and complexity very low and at the same time achieve visually lossless quality at compression ratios up to 10:1. Quality evaluations show very good performance compared to other existing codecs, especially for multi-generation applications.

Beyond the JPEG XS Core Coding System, multiple profiles and formats have been defined allowing usage of this new codec within many applications. Extensions and improvements are now being considered and investigated. Among these extensions, one can cite HDR support, Raw-Bayer improved support and multi-spectral imaging.

For more references, information and detailed description of the Standard, the reader is invited to consult the official JPEG XS webpage, available at https://jpeg.org/jpegxs.

References

[1] 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,” output document wg1n78028, Rio de Janeiro, BR (February 2018), https://jpeg.org/downloads/jpegxs/wg1n78028.pdf. Accessed: 2019-01-25.

[2] ISO/IEC JTC 1/SC 29/WG1, “Information technology -- Advanced image coding and evaluation methodologies -- Part 2: Evaluation Procedure for nearly lossless coding,” ISO/IEC 29170-2 (2015).

[3] Pellegrin, P., Keinert, J., Richter, T., Martinez, M., and Descampe, A., “JPEG XS Complexity and latency analysis report for Core Experiments CE1,” 74th SC29/WG1 Meeting, input document wg1m74032, Geneva, CH (January 2017).

[4] Willème, A., Richter, T., Rosewarne, C., and Macq, B., “Overview of the JPEG XS objective evaluation procedures,” Proc. of Appl. of Digital Image Proc. XL, SPIE (2017).

[5] McNally, D., Bruylants, T., Willème, A., Schelkens, P., Ebrahimi, T., and Macq, B., “JPEG XS call for proposals subjective evaluations,” Proc. of Appl. of Digital Image Proc. XL, SPIE (2017).

[6] Alexandre Willème, Saeed Mahmoudpour, Irene Viola, Karel Fliegel, Jakub Pospisil, Touradj Ebrahimi, Peter Schelkens, Antonin Descampe, & Benoit Macq (2018, September). Overview of the JPEG XS core coding system subjective evaluations. In Applications of Digital Image Processing XL (to appear). International Society for Optics and Photonics.