JPEG

JPEG White paper:

JPEG XS, a standard for visually lossless low-latency lightweight image coding system

Version 2.1

ISO/IEC JT1/SC29/WG1 WG1N100328

Editors (*in alphabetical order*): Tim Bruylants – IntoPIX SA, Belgium 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

TABLE OF CONTENTS						
EX	2					
1	INT	3				
2	USE	E CASES AND TARGETED APPLICATIONS	3			
	2.1	TRANSPORT OVER VIDEO LINKS AND IP NETWORKS	3			
	2.2	REAL-TIME VIDEO STORAGE	4			
	2.3	FRAME BUFFER COMPRESSION	4			
2	2.4	OMNIDIRECTIONAL VIDEO CAPTURE AND RENDERING	4			
	2.5	SENSOR COMPRESSION	5			
3	KE١	FEATURES OF THE JPEG XS CODEC	5			
4	COMPARISON TO STATE-OF-THE-ART CODECS 6					
5	5 JPEG XS PROFILES AND FORMATS					
6 PERFORMANCE						
7 CONCLUSION						
REFERENCES						

Executive Summary

JPEG XS is an International Standard from the JPEG Committee formally known as ISO/IEC SC29 WG1. It provides visually lossless image compression by means of an interoperable low-latency lightweight coding system and 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 a targeted bitrate to match the available bandwidth, e.g. Gigabit Ethernet, and it offers an end-to-end delay equal to a tiny fraction of a frame. Typical compression ratios are up to 10:1 for 4:4:4, 4:2:2 and 4:2:0 images, but these can also be higher depending on the nature of the image or the requirements of the targeted application. In addition, JPEG XS can also directly compress raw Bayer sensor data, bypassing the typical conversion to RGB. This white paper presents the main use cases and key features of the JPEG XS standard.



1 Introduction

Video bandwidth requirements are still growing, 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 much 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 moving from HD to UHD, requiring 12G-SDI or 10 Gbps Ethernet, and additionally 8k is slowly emerging. Consequently, uncompressed storage and live video transmission becomes unaffordable and unmanageable within current systems and infrastructures, while next generation transmission channels are not yet deployed or still too expensive. Facing this reality, the use of lightweight visually lossless 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. Currently, the second edition of the standard is available.

2 Use cases and targeted applications

In a nutshell, JPEG XS is the technology to transparently replace uncompressed video. In this respect, 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 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 cost effective. Given the available bandwidth, truly uncompressed video is therefore not an option anymore, both in the studio and in remote production, as 4K and 8K resolutions, and higher frame rates (fps) need to be supported.

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

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 designed the robustness of JPEG XS for a minimum of 10 encoding-decoding cycles. Moreover, the additional latency introduced by one encoding-decoding cycle with JPEG XS is below a couple of lines in order to avoid any human-perceptible delay between signals processed by different processing chains.



	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

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

2.2 Real-time video storage

Similar to transport of video streams, storage of high-resolution streams requires lightweight compression to allow writing the content in real-time to low-cost storage devices. For instance, cameras use internal storage technology like SSD drives or SD cards to store large streams of images, but the access rates on these devices are limited. Moreover, like 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 can be used in the following contexts:

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

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 interframe 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-constrained 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 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 kinds of climatic conditions.

JPEG XS provides specific technology to efficiently handle and compress raw Bayer content. More information on raw compression with JPEG XS is available in a dedicated publication [7] by the JPEG committee.

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 a GPU or an 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 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 codecs that existed in 2016, the JPEG committee 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 coefficient values are copied into the codestream as-is [8]. To allow different levels of latency and complexity, several profiles, each with maximum vertical wavelet decompositions, are defined. Each profile targets specific applications and can be classified into multiple categories. The most important categories are described in Table 2.

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 2: Most relevant JPEG XS profile categories.

Main profile category	High profile category
 Default profile 1 vertical wavelet Natural / CGI / screen content Broadcast, pro-AV, frame buffers, display links 	 More quality 2 vertical wavelets Natural / CGI / screen content Same applications as « Main » but for high-end devices, contribution, cinema remote production

Format	Туре	Description – Main purpose	Extension	Standard document
JXS	JPEG XS File format	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. 7 th AMD1
RTP	RTP payload format	IP transport for JPEG XS	n/a	RFC 9134
SMPTE 2110	System stream	Encapsulation of compressed video stream in SMPTE 2110	n/a	SMPTE ST 2110-22
MXF	Material eXchange Format	For storing of video	.mxf	SMPTE ST 2124

Table 3: JPEG XS transport and container formats

6 Performance

In order to validate that this newly developed codec meets all requirements, the JPEG Committee has analysed 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 1 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 two vertical wavelet transforms profile is usually on par with JPEG 2000.

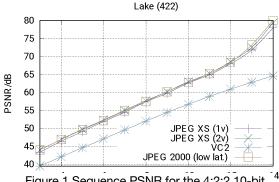


Figure 1 Sequence PSNR for the 4:2:2 10-bit ⁴ sequence "Lake". Experiments are described in more detail in [4].

7 Conclusion

JPEG XS is a stable and excellent International Standard for visually lossless low-latency lightweight image coding, designed to compensate for continuously increasing bandwidth requirements in video transport links. It can replace uncompressed video in any application. 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.

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

References

- [1] ISO/IEC JTC 1/SC 29/WG 1, "Use cases and requirements for JPEG XS v3.1," output document wg1n100090, 2022.
- 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. International Society for Optics and Photonics.
- [7] ISO/IEC JTC 1/SC 29/WG 1, "JPEG XS in-depth series raw image compression", output document wg1n100275, 2022.
- [8] Descampe, A., Richter, T., Ebrahimi, T., Foessel, S., Keinert, J., Bruylants, T., Pellegrin, P., Buysschaert, C., Rouvroy G., "JPEG XS - A New Standard for Visually Lossless Low-Latency Lightweight Image Coding," In Proceedings of the IEEE, May 2021.