JPEG XS in-depth series

Raw image compression

ISO/IEC JTC1 SC29/WG1
WG1N100275

Editors (*in alphabetical order*):
Christophe Biernaux – IntoPIX SA, Belgium
Tim Bruylants – IntoPIX SA, Belgium
Siegfried Foessel– 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 an International Standard from the JPEG Committee (formally known as ISO/IEC SC29/WG1). It was initially standardized in 2019 and a revised 2nd edition was published in 2022. This new edition brings further improvements to JPEG XS such as new coding tool support for the efficient high-quality compression of raw Bayer still image and video content. JPEG XS does so by relying on its famous visually lossless, low-latency and lightweight image coding system at the core [2]. In this context, JPEG XS can transparently replace uncompressed raw to reduce bandwidth and power requirements.

This document discusses the application domains where raw content matters and presents how JPEG XS addresses the challenges that inherently come with raw image and video workflows.

1 Introduction

Colored digital image content is typically acquired by image sensor chips that employ some type of color filter array (CFA) pattern on top of the luminance-sensitive pixel elements. One very commonly used CFA pattern is the Bayer filter mosaic comprised of two-by-two-pixel groups with half of the color filters in green, one quarter in red and the other quarter in blue (see Figure 1). As such, each pixel's color is captured in only for one of the three color channels. For visualization, the spatially disjoint color information must be first transformed into a full red, green and blue (RGB) color format where missing color information is calculated by means of interpolation. This conversion process is often referred to as debayering or demosaicing, effectively rendering the image for visualization. High-quality demosaicing can be a complex process as it needs to avoid producing any false artifacts – like chromatic aberration or aliasing – while also maximally preserving the image resolution and considering various properties such as white balance, optical black level and sensor noise. On the other hand, the demosaicing process is also constrained by the limitations of the camera hardware and must be of low enough complexity to allow for fast and efficient hardware implementations. In addition, the process is mathematically irreversible and relies on permanent choices that influence the visual output of the final image. And finally, demosaicing also increases the amount of data trifold.

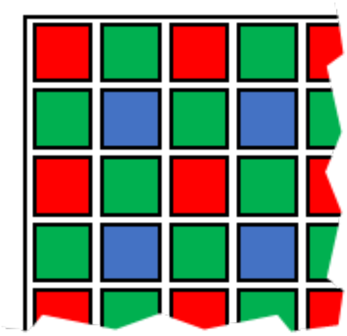

Figure 1 – Illustration of an RGGB Bayer CFA pattern.

For these reasons, professional (still image and video) cameras typically employ a raw workflow where the camera stores the original linear sensor information, with or without compression. Doing so allows shifting complex processing operations, like the demosaicing step, away from the camera. However, in many cases these workflows rely on proprietary raw formats with typically a straightforward compression step that offers modest compression factors at around 2:1. Such limited compression performance is usually a consequence of conservative choices made in the design of proprietary raw storage formats (e.g. the requirement for mathematically lossless compression, or a limited power budget), and is a strong indicator that the applied algorithms do not optimally exploit the data statistics of the raw content.

This document discusses the benefits that JPEG XS can bring as an open ISO standard to improve the raw workflow by reducing the required bandwidth and storage of still image and video cameras while at the same time providing visually lossless image quality. In fact, this paper explains that JPEG XS can even reduce the overall complexity and help to reduce camera power consumption [2][3][4].

2 Improving the raw workflow

As explained, direct image sensor compression in still image and raw movie cameras brings several advantages compared to the usual compression in the YUV or RGB domains. To start, raw image content that comes from an imaging sensor in its native pixel format requires less data to be processed. For example, a Bayer CFA pattern uses only one third of the data of a subsequent RGB or YUV image obtained after demosaicing. Thus, by applying direct compression of the raw image sensor data, this expansion can be avoided and additional savings in the amount of data can be achieved, reducing the bandwidth and storage requirements even further. Moreover, the steady transition towards high dynamic range (HDR) content production further pushes the need to keep full sensor information to better support post-processing and image analysis workflows. Thus, it is evident that, depending on the application or use case, new and more specific requirements exist today.

In **broadcast and high-end cameras**, the camera control unit (CCU) forms an essential component in a live television broadcast chain. This unit is responsible for powering the professional video camera, handling the signals sent over the camera cable to and from the camera, and can be used to control various camera parameters remotely. The CCU usually also converts the raw data stream into YUV 4:2:2 signals, managing various aspects such as black level, denoising, color conversion, white balance, debayering, gamma curve and chroma subsampling. The video link between the camera and the CCU usually relies on uncompressed raw data to guarantee minimal latency (<< 1ms). However, with the increasing resolutions and frame rates, and long distance remote CCU, there exists a need for compression with extreme low latency at ratios from 2:1 to 4:1 or better while maintaining visually lossless quality.

In **prosumer and consumer cameras**, including **mobile phones**, low power and low cost are of utmost importance to assure affordable mobility. At the same time, image sensor resolution and frame rate continuously increase, causing an opposite trend by significantly raising the amounts of produced data and thus also bandwidth and power usage. Reducing the data transfers by applying compression to the image sensor data is a practical solution to counteract this trend. However, this is only possible by means of a compression system that has low memory and limited computational demands, in order to be able to significantly reduce fabrication cost, power consumption and heat production. Moreover, in applications where uncompressed is currently the norm, image quality and low latency are the key features, thus any proposed compression needs to be transparent.

Machine vision systems typically rely on powerful image processing systems to process the captured image information. The cameras are typically separated from their respective processing system due to various restrictions like mounting and sizing constraints (think hazardous or extreme environmental conditions). Because of this, it is beneficial to use cameras as pure image sensor devices with as few computational capabilities as possible, while safeguarding the analytical properties of the image data. For this, only low latency compression is applicable, and all other operations should be left to the machine vision processing system.

A last category of major applications can be found in the **automotive industry** regarding autonomous driving and driving assistance systems (ADAS) – in a sense, these are also machine vision systems. For such applications multiple sensors are usually combined and their data is transmitted to the electronic central unit (ECU) for further processing and joint analysis. The ECU must be able to process the data as fast as possible, thus ultra-low latency is extremely important. Furthermore, given the large number of sensors and due to several constraints regarding the cabling inside a vehicle, compression will help to control aspects like bandwidth, power consumption, cost and thermal characteristics.

In each of these mentioned application domains, a raw workflow with efficient compression can bring many advantages by reducing bandwidth, power and storage requirements, on the condition that complexity and latency can be kept under control, combined with high visual quality and access to the original sensor data.

3 JPEG XS

3.1 Brief technical overview

The second edition of JPEG XS [2][3] brings support for efficient and direct compression of raw Bayer CFA images [4] while maintaining its famous low complexity and low latency properties. In other words, JPEG XS can now be employed wherever uncompressed raw content is the norm, to save bandwidth and power.

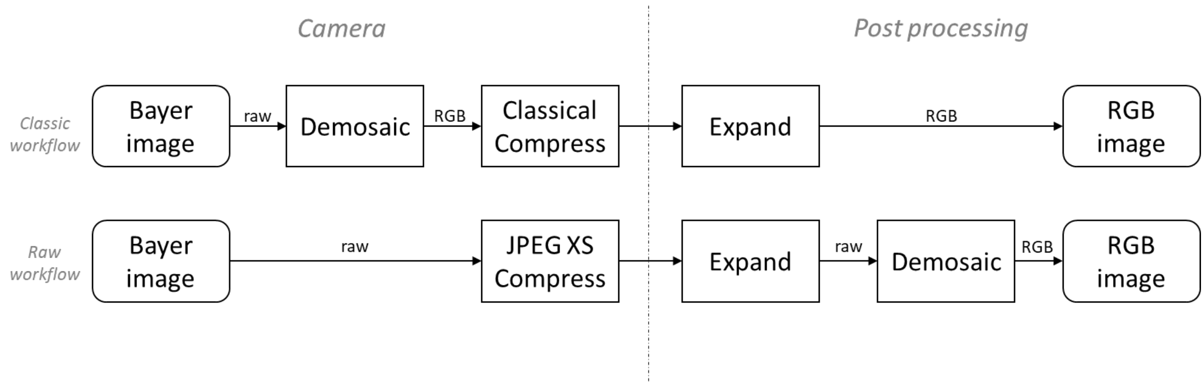

Figure 2 – High-level overview of a classical (top) and a JPEG XS raw Bayer (bottom) compression workflow. JPEG XS defers the irreversible and high resource demosaicing step to post processing (outside of the camera).

JPEG XS uses a specially designed color transform, called the Star Tetrix transform, in combination with a set of configurable non-linear transforms (NLTs) to compensate for sensor gains, noise, optical black level and white balance. Apart from these two additions, the overall JPEG XS compression scheme remains fundamentally unchanged – i.e., the exact same wavelet transformation, rate-distortion optimizer, entropy coder and code stream format are used (see Figure 3). This means that implementers can rely on well-established knowledge about the performance behavior and build on top of existing implementations to facilitate adoption. It should be emphasized that, with the Star Tetrix transform, JPEG XS takes CFA pattern data directly, avoiding the need to apply irreversible choices regarding the conversion to RGB that many other compression schemes suffer from.

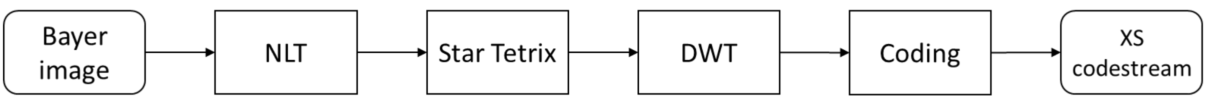

Figure 3 – Encoding steps inside JPEG XS to compress a raw Bayer CFA image, using the new NLT and Star Tetrix operations.

When encoding, the first (optional) step is the application of a non-linear transform (NLT) that allows JPEG XS to better adjust for the specific transfer curve characteristics that will eventually be applied at the display device, significantly improving the visual quality at the observer side. Usually this is handled as an integral part of the demosaicing step when transforming linear raw Bayer samples into one of the typical non-linear RGB color spaces. Thus, the NLT enables JPEG XS to correct the sensor samples while still relying on its uniform or dead zone quantizers to achieve optimal compression results, especially at the higher quality levels. JPEG XS currently standardizes two NLTs, a quadratic NLT and an extended NLT with improved handling of low-exposure sensor output. The quadratic NLT implements the standard gamma correction with an approximated gamma value of 2 instead of 2.2 to simplify implementations. The extended NLT on the other hand implements an S-like formed non-linearity, with configurable parameters that allow adjusting it to the sensor at hand. Important to notice is that JPEG XS still compresses and stores all sample values below the optical black level, since these are often crucial in the further development of the image (e.g., for denoising operations). Secondly, the Star Tetrix transform is a specially designed reversible multi-component color transform. Star Tetrix works directly on the raw Bayer samples to optimize further compression with XS in the exact same manner as for the other supported image types (like RGB and YUV).

As such, JPEG XS presents itself as the solution to the ever-increasing bandwidth requirements and power limitations for raw image content. It offers compression factors between 6:1 and 12:1 with visually lossless quality and can do around 2.5:1 mathematically lossless compression. JPEG XS can be deployed transparently into existing workflows without any drawbacks, because it takes raw Bayer CFA data as input at the encoder and delivers raw Bayer CFA data as output at the decoder.

3.2 Raw profiles and levels

For the compression of raw Bayer CFA images, three Bayer profiles are defined, as shown in Table 1, each targeting a different complexity level. The three profiles also offer different latency guarantees, with the inherent latency related to the number of vertical wavelet transforms. The lowest latency can be achieved with the LightBayer profile, especially in combination with the inline Star-Tetrix mode. All XS Bayer profiles support sample bit depths of 10, 12, 14 and 16 bits, and support the optional quadratic and extended NLTs.

Table 1 – XS Bayer profiles.

LightBayer	**MainBayer**	**HighBayer**
• Reduced complexity, slightly lower compression performance • No vertical DWT • Ultra-low latency • Inline Star-Tetrix only • Low power applications	• Default profile • 1 vertical DWT • Low latency • Full and inline Star Tetrix	• Best quality • 2 vertical DWT • Low latency • Full and inline Star Tetrix • Same applications as MainBayer but for high-end equipment

4 Performance results

This section presents some compression performance results obtained by applying JPEG XS in a raw workflow. It is evident that what matters is the perceived visual quality of the result. Yet, any visual quality assessment requires rendering the raw Bayer image to generate one out of many visually acceptable RGB image versions that can be displayed. Recall that the process to convert raw sensor data to an RGB image involves making irreversible choices for the demosaicing process, application of non-linear transforms, and other corrections like filtering, noise reduction, etc. All of this will visually affect the rendered result. In this paper, the conversion from raw Bayer content to RGB was done using the Adaptive Homogeneity-Directed (AHD) algorithm [5], combined with the standard corrections (white balance, black level, and color conversion) using the camera provided values. For each result, demosaicing and corrections were applied either before or after the actual compression step, depending on whether the codec compresses RGB or raw Bayer data (see Figure 2).

The results were generated for a selected set of 20 raw Bayer still images, taken with modern digital cameras of various brands and models, having 12-bit and 14-bit sensor depths (images were captured without compression or with mathematically lossless compression [6]).

To start, a comparison baseline is established by providing the following compression-decompression results in the 24-bit RGB domain at various target bit rates – being 1.00, 1.50, 2.00, 3.00 and 4.00 bits per pixel (bpp), using:

- JPEG compression: Render the raw input image to 24-bit RGB. Then compress this 24-bit RGB with JPEG [7]. The decompressed result is a 24-bit RGB image.
- XS HighBayer: Compress the raw input image at native bit-depth using XS with HighBayer profile. Decompress and render to 24-bit RGB.

All codecs were tuned to prioritize for optimizing the Peak-Signal-to-Noise (PSNR) output – as opposed to optimizing for visually perceived quality – to provide a fair comparison in the PSNR domain. It is also important to note that JPEG has the demosaicing step before compression, which is exactly what a raw workflow wants to avoid. In other words, even in the case that PSNR values indicate good quality, there is still the huge functional drawback of rendering to RGB before applying compression.

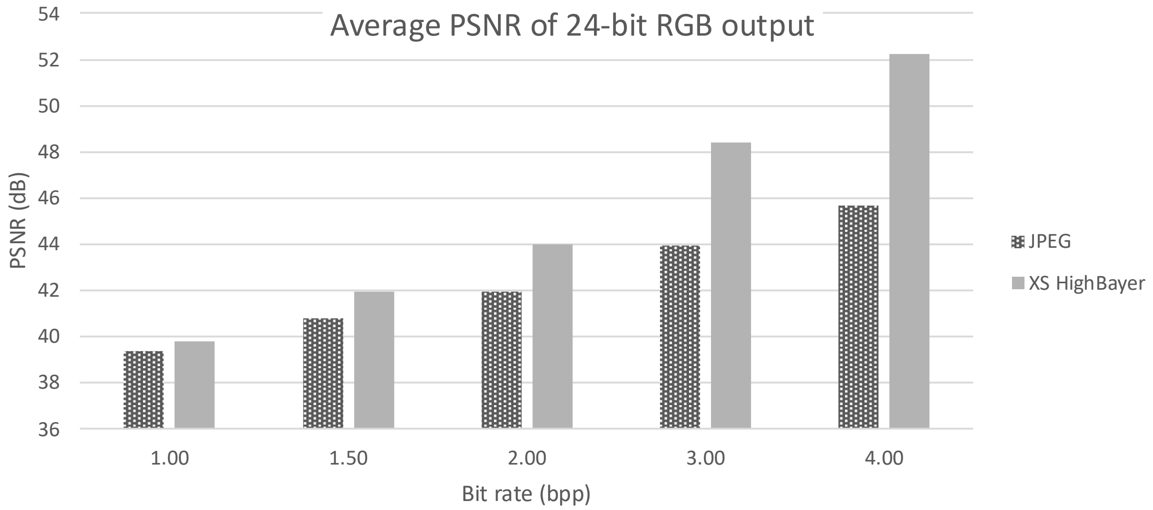

Figure 4 – Compression performance of JPEG (8-bit RGB input) against JPEG XS using HighBayer profile (native raw input, 8-bit RGB output).

Figure 4 shows the objective quality comparison of JPEG against JPEG XS (using the HighBayer profile) when decoding image content into the 24-bit RGB domain. The bars in this graph represent PSNR values, measured in dB, calculated on the 24-bit RGB images. In this scenario and at low bit rates, JPEG XS with the raw coding tools (HighBayer profile) performs like JPEG. However, at higher bit rates JPEG XS significantly outperforms JPEG. This shows the benefit that the raw coding tools in JPEG XS bring.

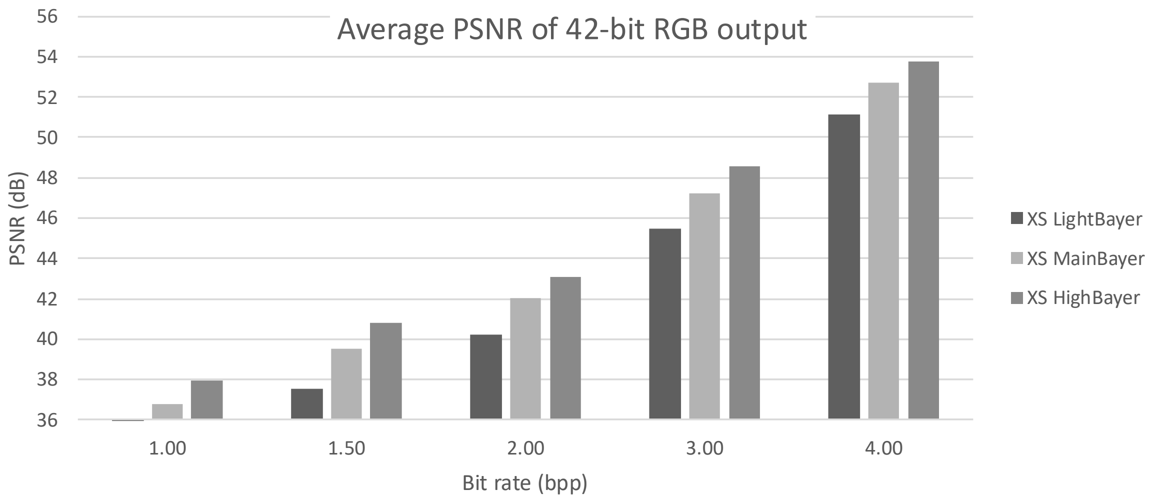

Figure 5 – Comparison between the three Bayer profiles in JPEG XS in the 42-bit RGB domain for 14-bit images. This demonstrates the trade-off that can be made with JPEG XS between compression performance and complexity.

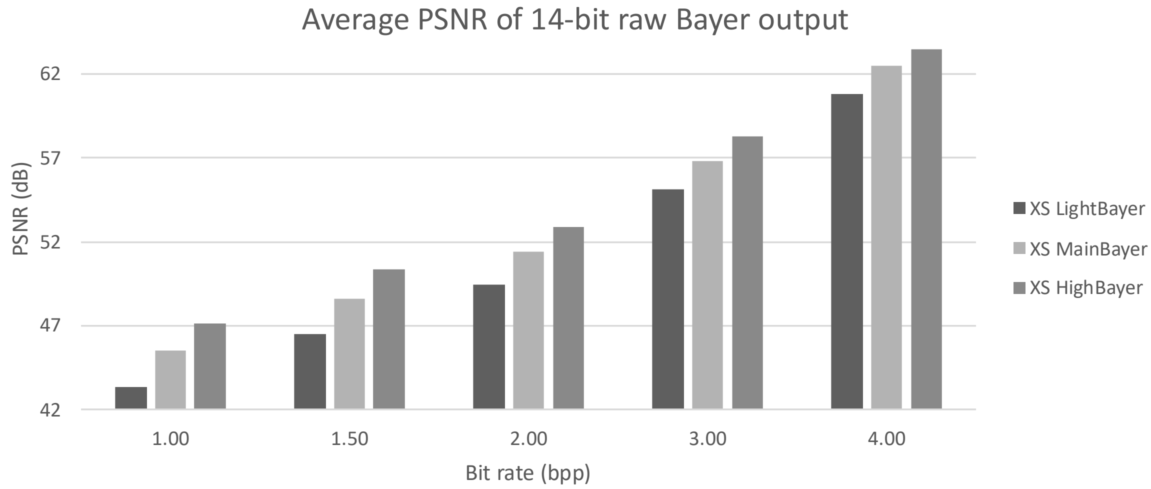

Figure 6 – The same comparison as in Figure 5 for 14-bit images but showing raw Bayer domain PSNR values (without demosaicing).

Figure 5 and Figure 6 show the impact of the three raw Bayer profiles that JPEG XS offers. The differences between LightBayer, MainBayer and HighBayer are listed in Table 1. Basically, the profiles differ in the applied number of vertical wavelet decomposition levels – being 0, 1 and 2 respectively – and the way that the Star-Tetrix transform is applied – being Inline or Full. These results show that slightly increasing the complexity – and inherently also the latency – via the profiles also increases the compression performance, and hence also the delivered image quality.

The results above were generated for raw Bayer still images, which is exactly what the raw coding tools in the 2nd edition of JPEG XS were originally designed for. However, JPEG XS is also actively being deployed in professional video and broadcast applications for the compression of video sequences. In such scenarios, compression at 3 bpp for 4K@60 content has proven to deliver visually lossless quality. This means that in combination with the raw Bayer coding tools, JPEG XS can now also be used to compress raw Bayer video sequences at even better quality (or lower bit rates).

Moreover, ISO/IEC SC29 WG1 has already begun the work on the 3rd edition of JPEG XS to standardize further improvements and additional coding tools that will help with new features such as improved compression of raw video sequences and desktop graphics content [1]. The target is to allow JPEG XS to deliver visually lossless compression of 4K and 8K at 60 frames per second at bit rates as low as 1.5 bpp.

5 Conclusions

The 2nd edition of JPEG XS, published in 2022, standardizes additional coding tools that focus on the efficient compression of raw image and video content in combination with its visually lossless, low-latency and lightweight image coding system. The major benefit is that it is now possible to apply high-quality compression in any raw workflow where before uncompressed used to be the norm, allowing to save on bandwidth and storage without any drawbacks. JPEG XS is known for its ultra-low latency and low-complexity, and this remains true for raw image and video content. Moreover, performance tests indicate that JPEG XS outperforms alternative technologies.

For more information and a detailed description of the Standard, the reader is invited to consult the official JPEG XS webpage, available at http://www.jpeg.org/jpeg-xs.

References

[1] ISO/IEC JTC1 SC29 WG1, "Use cases and requirements for JPEG XS v3.1," Online, https://ds.jpeg.org/documents/jpegxs/wg1n100090-094-REQ-Use_Cases_and_Requirements_for_JPEG_XS_v3.1.pdf, Jan 2022.
[2] ISO/IEC JTC1 SC29 WG1, "JPEG XS, a new standard for visually lossless low-latency lightweight image coding system, v2.0," Online, https://ds.jpeg.org/whitepapers/jpeg-xs-whitepaper.pdf.
[3] 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.
[4] Richter, T., Foessel, S., Descampe, A., Rouvroy, G., "Bayer CFA Pattern Compression with JPEG XS," In IEEE Transactions on Image Processing, 30, (pp. 6557-6569), July 2021.
[5] K. Hirakawa and T. W. Parks, "Adaptive homogeneity-directed demosaicing algorithm," In IEEE Transactions on Image Processing, vol. 14, no. 3, pp. 360-369, March 2005.
[6] "Raw Pixls Us," Online, https://raw.pixls.us.
[7] ISO/IEC IS 10918-7, Rec. ITU-T T.873, "Information technology - Digital compression and coding of continuous-tone still images: Reference software".