JPEG XT: A Compression Standard for HDR and WCG Images [Standards in a Nutshell]

Alessandro Artusi, Rafał K. Mantiuk, Thomas Richter, Pavel Korshunov, Philippe Hanhart, Touradj Ebrahimi, Massimiliano Agostinelli

I. BACKGROUND

A. Motivations

HIGH BIT DEPTH data acquisition and manipulation have been largely studied at the academic level in the last 15 years and are rapidly attracting interest at the industrial level. An example of the high interest for High Dynamic Range (HDR) imaging is the use of 32-bit floating point data for video and image acquisition and manipulation that allows a variety of visual effects that closely mimic the real world visual experience of the end-user [1] (see Fig. 1 for an illustration). At industrial level, we are witnessing increasing traction towards supporting HDR and Wide Color Gamut (WCG). WCG leverages HDR for each color channel to display a wider range of colors. Consumer cameras are currently available with 14 or 16 bit A/D converter. Rendering devices are also appearing with the capability to display HDR images and video with a peak brightness of up to 4000 nits and to support wide color gamut (ITU-R Rec. BT.2020 [2]) rather than the historical ITU-R Rec. BT.709 [3]. This trend calls for a widely accepted standard for higher bit depth support that can be seamlessly integrated into existing products and applications.

While standard formats such as JPEG 2000 [5] and JPEG XR [6] offer support for high bit depth image representations, their adoption requires a non-negligible investment that may not always be affordable in existing imaging ecosystems, and induces a difficult transition, as they are not backward-compatible with the popular JPEG image format.

Instead, most digital camera and mobile phone manufacturers either store images in proprietary *RAW* formats or, more commonly, offer an *HDR mode*, which produces a traditional low dynamic range image with improved details. The former solution creates a *vendor lock-in* problem for consumers, making it difficult to efficiently use images produced by such

Fig. 1. A typical LDR image taken with default settings of a Nikon D7100 camera (left) and an HDR image fused from 5 exposures and tone-mapped with `drago03` [4] for display (right).

cameras in practice due to a lack of interoperability between proprietary formats. The latter solution generates a low dynamic range (LDR) 8-bit JPEG version from the captured high bit depth image. In other words, visual information contained in the original high bit depth *digital negative* is irremediably lost, which is not optimal for editing, creative enhancements, or even viewing on HDR-capable display devices.

The JPEG XT standard aims to overcome all these drawbacks and to lower the entry barriers to the market. While offering new features such as lossy or lossless representation of WCG and HDR images, JPEG XT remains backwards compatible with the legacy JPEG standard. As a result, legacy applications can reconstruct an 8-bit/sample LDR image from any JPEG XT codestream. This LDR version of the image and the original HDR image are related by a tone-mapping process that is not constrained by the standard and can be freely defined by the encoder.

B. The Standards

The Joint Photographic Experts Group (JPEG) formally known as ISO/IEC JTC1/SC29/WG1 is universally recognized as the leading committee for compressed image formats. The JPEG committee began the standardization of JPEG XT technology in 2012. A call for proposals was issued in June 2012, at the Vienna meeting, to which 6 organizations responded, namely, Dolby, Ecole Polytechnique Fédérale de Lausanne (EPFL), University of Stuttgart, Trellis Management, Vrije Universiteit Brussel (VUB), and University of Warwick. As a result, JPEG XT was initiated as a new work item and a set of requirements for its potential applications was identified. The

A. Artusi is with the GiLab of the University of Girona, Spain; e-mail: artusialessandro4@gmail.com.

Rafal K. Mantiuk is with the Computer Laboratory, University of Cambridge (UK); email: rkm38@cam.ac.uk.

T. Ricther is with the Computing Center of the University of Stuttgart, Germany; email: richter@tik.uni-stuttgart.de.

P. Korshunov is with MMSPG, EPFL, Lausanne, Switzerland; email: pavel.korshunov@idiap.ch.

P. Hanhart is with the MMSPG, EPFL, Lausanne, Switzerland; email: philippe.hanhart@epfl.ch.

T. Ebrahimi is with the MMSPG, EPFL, Lausanne, Switzerland; email: Touradj.Ebrahimi@epfl.ch.

M. Agostinelli is with Trellis Manageament Co Ltd.; email: max@xdepth.com.

Manuscript received ....

JPEG XT image coding system is organized into nine parts that hierarchically define the baseline coding architecture, known from the legacy JPEG standard, an extensible file format specifying a common syntax for extending the legacy JPEG codestream, and application of this syntax for coding integer or floating point samples between 8 and 16 bits precision. This coding architecture is then further refined to enable lossless and near-lossless coding, and is complemented by an orthogonal extension for representing opacity data, commonly known as *alpha-channels* (see Table I).

II. TECHNOLOGY

HDR images require more than the typical 8-bits per sample, e.g. an integer value in [0,255] for a component of a pixel in the legacy JPEG standard, for faithful image representation. The original JPEG specifications do include a 12-bit mode and the lossless JPEG coding mode supports up to 16-bits per sample. Unfortunately, these two JPEG variants are incompatible with the popular 8-bit mode and hence are rarely used in practical applications, such as digital photography.

JPEG XT builds on top of the widely adopted 8-bit mode of JPEG and extends it both in a forward- and backward-compatible way. It is a superset of the 8-bit mode, i.e. JPEG XT re-uses existing JPEG technology whenever possible. As a result, legacy JPEG implementations shall be able to decode a LDR image from a JPEG XT stream.

A. Components

Typically a standard specifies only the decoder side but, for the sake of clarity, we will briefly introduce how the HDR image is preprocessed at the encoder level in order to take advantage of the existing JPEG 8-bit mode. The input of the encoder is typically a pair of LDR/HDR images. Prior to encoding, the representation of the HDR image is preprocessed using a combination of four elementary operations: (i) scalar non-linear functions that can be described either by a parametrized curve or a look-up table, (ii) 3×3 matrix multiplication, (iii) vector addition of three-dimensional vectors, and (iv) scalar multiplication of a vector by a scalar number. These operations are applied to each pixel independently, without taking the coding history or the neighborhood of the pixel into account. Pre-processing is therefore straightforward to parallelize. This pre-processing step yeilds two layers, a LDR image and an extension image, that are both encoded with existing 8-bit mode of JPEG. While JPEG XT defines a unified decoder design that arranges these components into a workflow, a typical decoder or encoder would not implement all of these components. In real life, a codec is likely to only implement a subset of these operations. As will be detailed later, JPEG XT defines *profiles* that specify a subset of the full configuration space and hence simplify the design of codecs.

As in JPEG, the pre-processed input is then decorrelated with a discrete cosine transform (DCT), quantized and entropy coded. Since the bit-precision of the legacy 8-bit mode is limited, JPEG XT defines two alternate mechanisms to improve it: *refinement coding* and *residual coding*.

Fig. 2. Simplified version of the JPEG XT decoder. $\mathcal{B}$ is the base layer and is always represented as a JPEG codestream with 8-bit per sample. $\mathcal{E}$ is the extension layer that used in conjunction with $\mathcal{B}$ allows the reconstruction of the HDR image.

Refinement coding extends the coding precision in the DCT domain thanks to a coding mechanism that is closely related to the progressive coding mode of the legacy JPEG standard. It extends the coding precision to 12-bits in the spatial domain. The most significant bits of the quantized DCT coefficients are encoded by a regular JPEG coding mode, forming the codestream that legacy applications can interpret. On the other hand, the least significant bits are encoded with the so-called *successive approximation scan*, which is part of the progressive coding mode also defined in the legacy JPEG standard. However, the encoded coefficients are not included in the regular codestream. They become part of a side-channel (an *extension layer*) that is hidden from legacy applications. The transport of this side-channel is discussed below in more details.Refinement coding cannot represent an arbitrary LDR/HDR image pair on its own. The LDR image is indeed implicitly defined by the most significant bits of the HDR stream, making refinement coding alone only suitable for simple applications. The capability to encode an HDR image with an independently defined LDR layer is granted by residual coding that operates entirely in the spatial domain. Using the four elementary operations available at preprocessing, it computes from the LDR/HDR image pair a LDR layer that represents the base codestream that is visible to legacy applications, and an extension layer for the remaining information required to reconstruct the HDR image. The extension layer is also coded by a second regular JPEG mode, and the resulting extension codestream, similarly to the refinement scans, becomes part of a side channel that is hidden for legacy applications. Both mechanisms, residual and refinement coding, can be combined. For example, the bit-precision of the extension layer from residual coding could be increased by using refinement scans.

B. Profiles

While preprocessing offers a variety of methods to generate an extension layer from a given LDR/HDR image pair, we restrict, for the sake of simplicity, the discussion to the three profiles currently defined in the JPEG XT standard, whose decoding workflow is depicted in Fig. 2.

	Title	Description
Part 1	Core Coding System Specification	Definition of the core coding technology, which is the legacy JPEG specifications. Other parts of JPEG XT builds on top of this baseline coding system in a backwards compatible way [7].
Part 2	Extensions for HDR Images	It supplies a legacy syntax for a subset of the tools specified in Part 7 [7].
Part 3	Box File Format	Definition of an extensible and flexible container format, called *boxes*, extending legacy JPEG and the ISO-based media format [7].
Part 4	Conformance Testing and Evaluation	Definition of the methodology to verify that the various Parts of the standard are meeting the normative requirements [7].
Part 5	Reference Software Implementation	Reference software for Part 6-7-8-9 making use of the file box format specified in Part 3 [7].
Part 6	Intermediate Dynamic Range (IDR) Integer Coding	Definition of a scalable coding engine supporting all bit depths between 9 and 16 bits per sample that remains compatible with legacy applications [7].
Part 7	HDR Floating-Point Coding	Definition of a coding engine for images in a *HDR representation*, e.g. using floating point samples [7].
Part 8	Lossless and Near-Lossless Coding	Definition of a lossless and near-lossless coding engine for IDR and HDR image representations using coding technologies from Parts 6 and 7 [7].
Part 9	Encoding of Alpha Channels	Extension of the other parts of the ISO/IEC 18477 standard to support opacity information for LDR, IDR and HDR images [7].

TABLE I
DESCRIPTION OF THE JPEG XT STANDARD.

Two layers, $\mathcal{B}$ and $\mathcal{E}$, are used for the reconstruction of the HDR image $\mathcal{I}$. $\mathcal{B}$ is the base layer, which represent the LDR image as a JPEG image with 8-bits per sample in the ITU BT.601 RGB colorspace. $\mathcal{E}$ is the extension layer, which includes the additional information to reconstruct the HDR image $\mathcal{I}$ starting from the base layer $\mathcal{B}$. The coding tools of the overall JPEG XT infrastructure used to merge $\mathcal{B}$ and $\mathcal{E}$ together are then profile dependent.

Profile A reconstructs the HDR image $\mathcal{I}$ by multiplying a luminance scale μ with the base image $\mathcal{B}$ after inverse gamma correction using the 1st base nonlinearity Φ_A

$$\mathcal{I}(x,y) = \mu\big(\mathcal{E}_0(x,y)\big) \cdot \Big[C\, \Phi_A\big(\mathcal{B}(x,y)\big) + \nu\Big(S\,C\,\Phi_A\big(\mathcal{B}(x,y)\big)\Big) \cdot R\,\mathcal{E}^{\perp}(x,y)\Big] \quad (1)$$

where C and R are 3x3 matrices implementing color transformations, $\mu(.)$ is a scalar function of the luma component of the extension layer $\mathcal{E}_l$ (*Post-scaling nonlinearity* block) and $\mathcal{E}^{\perp}$ the extension layer projected onto the chroma-subspace, i.e. $\mathcal{E}$ with its luma component set to zero. The matrix C transforms from ITU-R BT.601 to the target colorspace in the extension layer. R is an inverse color decorrelation transformation from YCbCr to RGB in the extension layer to clearly separate the luminance component from the chromaticities (YCbCr) at encoding level. These matrices are also commonly used in the other two profiles. S is a row-vector transforming color into luminance, and ν is a scalar function taking in input luminance values. Typically, $\nu(x) = x + \epsilon$ where ϵ is a "noise floor" that avoids instability in the encoder for very dark image regions.

Profile B reconstructs the HDR image $\mathcal{I}$ by computing the quotient that can be expressed as a difference in the logarithmic scale:

$$\mathcal{I}(x,y)_i = \sigma \exp\Big(\log\Big(\big[C\,\Phi_B\big(\mathcal{B}(x,y)\big)\big]_i\Big) - \log\Big(\Psi_B\big(\big[R\,\mathcal{E}(x,y)\big]_i\big) + \epsilon\Big)\Big)$$
$$= \sigma \frac{\big[C\,\Phi_B\big(\mathcal{B}(x,y)\big)\big]_i}{\Psi_B\big(\big[R\,\mathcal{E}(x,y)\big]_i\big)+\epsilon} \quad (i=0,1,2) \tag{2}$$

where i is the index of the RGB color channels. Φ_B and Ψ_B are two inverse gamma applied to the base and extension layers respectively. Φ_B has the objective to linerise the base layer, while Ψ_B intends to better distribute values closer to zero in the extension layer. The scalar σ is an exposure parameter that scales the luminance of the output image to optimize the split between base and extension layers.

Profile C also employs a sum to merge base and extension images, but here Φ_C not only approximates an inverse gamma transformation, but implements a global inverse tone-mapping procedure that approximates the (possibly local) tone mapping operator (TMO) that was used to create the LDR image. The extension layer is encoded in the logarithmic domain directly, avoiding an additional transformation. Finally, $\log$ and $\exp$ are substituted by piecewise linear approximations that are implicitly defined by re-interpreting the bit-pattern of the half-logarithmic IEEE representation of floating-point numbers as integers. It is then easily seen that this simple "casting" between number formats implements two functions $\psi\log$ and $\psi\exp$ that behave approximately like their precise mathematical counterparts, though they provide the additional advantage of being exactly invertible. The reconstruction algorithm for profile C can then be written:

$$\mathcal{I}(x,y) = \psi\exp\Big(\hat{\Phi}_C\big(C\,\mathcal{B}(x,y)\big) + R\,\mathcal{E}(x,y) - 2^{15}(1,1,1)^T\Big) \tag{3}$$

where $\hat{\Phi}_C(x) = \psi \log\big(\Phi_C(x)\big)$, in which Φ_C is the global inverse tone-mapping approximation. 2^{15} is an offset shift to make the extension image symmetric around zero. The codestream never specifies Φ_C directly, but rather includes a representation of $\hat{\Phi}_C$ in the form of a lookup-table, allowing to skip the time-consuming computation of the logarithm.

C. Lossless Coding

An important feature of profile C is that it allows implementations operating entirely with integers until the very last step, where the exponential generates floating-point output. All of these operations, including the exponential mapping, are exactly invertible. Part 8 of JPEG XT defines now on this basis a lossless coding engine by fully specifying the DCT in the base layer, and by bypassing the DCT entirely in the extension layer. The reader may now verify that the entire operation chain has, indeed, an exact inverse as the coding residual the decoder requires for a given DCT and a given base image is exactly predictable by the encoder, and hence can be computed ahead to generate exactly the required sample values.

III. TRANSPORT

Residual coding and refinement coding create additional codestreams that need to be incorporated into the legacy JPEG syntax such that current decoders are able to see only the legacy LDR image and skip over the extension layers. A JPEG XT codestream may thus contain up to three side-channels for image information: (i) a refinement codestream, (ii) a residual codestream, and (iii) a residual-refinement codestream. Accounting for the potential presence of an opacity layer, as defined in Part 9 of the standard, up to four additional codestreams may be further included: (i) an alpha channel, (ii) a residual alpha channel, (iii) an alpha channel refinement, and (iv) a residual alpha channel refinement. This information is added to the metadata that configures the post-processing chain of the decoder.

The legacy JPEG syntax already includes a generic extension mechanism by using so-called *APP-markers*. JPEG XT reserves one of them. However, APP-markers do not carry the data directly. Instead, their payload consists of so-called *boxes* that yield a better and cleaner structure of its contents. Boxes are not new to JPEG XT; they were previously introduced by MPEG and JPEG 2000. The payload data of a box is prefixed by a type and a size such that decoders unaware of a specific box type may simply ignore it. In summary, the JPEG XT file format is a JPEG codestream with APP markers whose contents, when re-assembled at the decoder, make up a single box, or a superbox containing multiple other boxes (cf. Fig. 3). This is necessary because the capacity of a single APP-marker is limited to 64Kbytes, whereas a box may be larger and hence span across several APP markers. Instructions describing how to assemble markers into boxes is included in the first bytes of the marker data itself. Legacy decoders will simply skip over the markers, and hence will also ignore all boxes and their data.
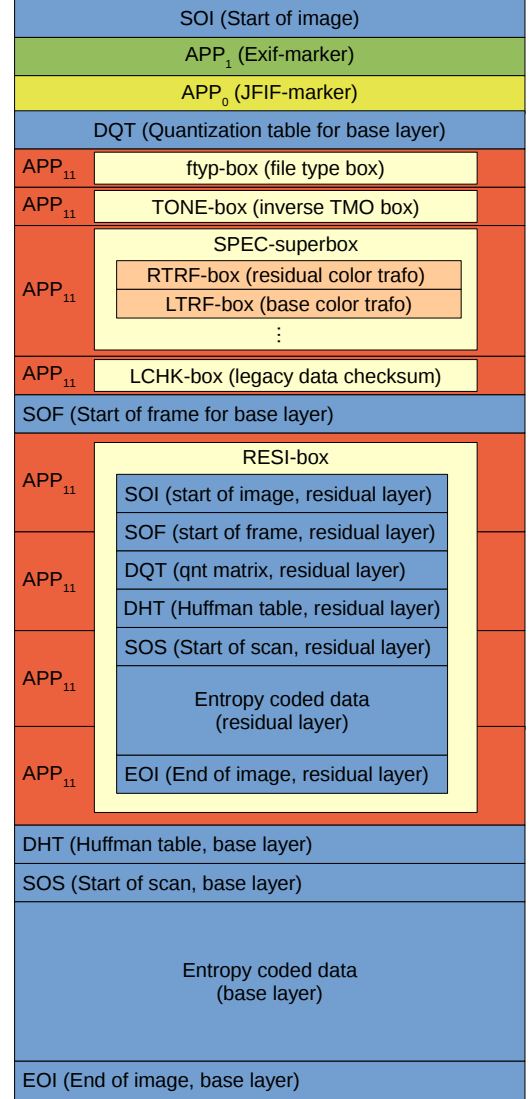
Fig. 3. JPEG XT file format: APP markers contain the boxes structure. Blue syntax elements were defined in the legacy JPEG standard, yellow elements were defined in later parts. The green Exif marker is defined outside of JPEG. The red APP_{11} markers are not part of the legacy JPEG standard and will thus be ignored by legacy JPEG readers. In contrast, JPEG XT readers will be able to interpret the payload data in the boxes that contain relevant JPEG XT information. Since the size of an APP marker is limited to 64 kB, a single box may extend over several markers; the RESI box carrying the residual codestream is a typical example. Otherwise, an APP marker or a series of APP markers contains a single box. This box may be, however, a superbox whose payload again consists of boxes. The SPEC box defining the instruction how to re-assemble the HDR image from base and exension image is a typical example.

The payload data of the boxes includes the metadata defining the post-processor in the decoder or the entropy-coded data of refinement and residual codestreams. The decoder pick out the data it requires for its operation based on the type signaled in the header of the box.

IV. Testing and Performances

The challenge of testing backward-compatible HDR compression is that the compression performance does not depend only on a single quality control parameter, but also on the quality settings for the base layer and on the choice of tone-mapping operator (TMO), which produces this layer. Therefore, the performance of JPEG XT needs to be evaluated using a comprehensive set of varying parameters and on a dataset that covers a large set of standard's use cases.

The JPEG committee has carried out a large number of experiments, using both subjective and objective methodologies, to asses the capability of the JPEG XT. A set of 12 objective quality metrics were tested on 106 HDR images (resolutions range from full HD to 4K) covering a high range of scenes typically captured in HDR images, including indoor, outdoor scenes, architecture, landscapes, portraits, frames from HDR video, and computer generated images. All images were carefully selected by experts in HDR imaging from the following publicly available datasets: Fairchild's HDR Photographic Survey[8]and EPFL's dataset of HDR images[9].

Since a TMO can be freely selected for encoding and its selection is not part of JPEG XT specifications, we tested 5 different commonly used operators: a simple gamma-based operator *gamma*, a global logarithmic tone-mapping operator [4] *drago03*, a global version of the photographic operator [10] *reihard02*, an operator optimized for encoding [11] *mai11*, and a local operator with strong contrast enhancement [12] *mantiuk06*. To fully understand the implications of the tone-mapping operators and JPEG XT parameters, all possible combinations of these parameters were tested. We used the combination of 10 base quality levels × 10 extension layer quality levels × 5 TMO's × 3 profiles, which results in a total of 1 500 conditions for each of the 106 images resulting in 159 000 tests. However, such a large number of conditions clearly cannot be tested in a subjective experiment. Therefore, from the total 106 HDR images, a subset of 20 images was selected by experts for subjective evaluations and these images were adjusted for viewing on SIM2 HDR monitor. Please refer to [13] for more details on the subjective evaluations.

The results of subjective experiments are crucial to select the right image quality metric and to provide ground truth reference. However, a subjective experiment alone cannot cover the entire space of parameters. Due the tedious nature of those experiments, only a limited number of images can be tested, which makes the findings of such studies difficult to generalize. For that reason, we analyzed compression performances with respect to HDR-VDP-2 [14], the best performing objective quality metric according to a set of evaluations. The image quality computed for a range of base and extension layer quality settings may result in arbitrary bit rates, making the results difficult to aggregate. Therefore, the predicted quality values were linearly interpolated to find the HDR-VDP-2 Q-scores for each desired bit rate. This step was necessary to determine average performance and confidence intervals for all tested profiles.

In Figure 4, we compare the performance of the three profiles with popular HDR image formats, including lossless
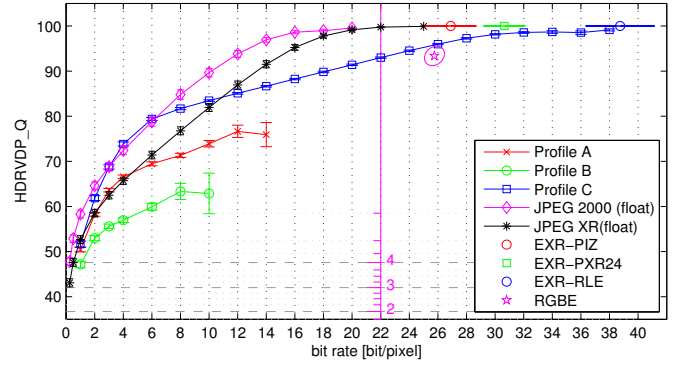

Fig. 4. The mean compression performance of each profile compared with popular near-lossless HDR image formats: OpenEXR using its three compression algorithms, and Radiance RGBE (.hdr). The base layer quality was fixed at 80. The ellipses denote 95% confidence interval. The magenta scale in the middle of the plot shows mean-opinion-score (MOS) values corresponding to HDRVDP_Q predictions.

OpenEXR, Radiance RGBE, lossy JPEG 2000, and JPEG-XR (floating encoding). OpenEXR and Radiance offer lossless compression, however the loss happens when converting images to their internal pixel formats: 8-bit RGB channels and shared 8-bit mantissa (E) for Radiance RGBE; and 16-bit half-float (sign, 5-bit exponent, 10-bit mantissa) for OpenEXR. Note that our reference images were stored in 32-bit per-color channel, uncompressed PFM files. JPEG 2000 employs a lossy wavelet-based compression while JPEG-XR uses a two-stage frequency transform, combining the features of both DCT and wavelet transforms. We can notice differences in quality performances between profiles i.e. C vs. A and B; however, these differences are above the predicted mean-opinion-score (MOS) value of 4.6 and are unlikely to be noticeable [15].

HDR-VDP-2 did not detect any degradation in quality for all OpenEXR compression formats (HDRVDP_Q 100 is the highest quality), while small losses in quality were detected for Radiance RGBE. All those lossless formats preserve very high quality but require at least 27 bits per pixel. JPEG XT performs unexpectedly well when compared with other lossy compression methods. Below 10 bit/pixel, JPEG XT performs better than JPEG XR. Below 6 bit/pixel, the performance of JPEG XT is comparable to JPEG 2000, even though the former encodes an additional tone-mapped image and employs a standard DCT-based JPEG codec, rather than a more advanced compression algorithms found in both JPEG 2000 and JPEG XR, which are newer.

The additional precision of these formats may be needed, however, if the content has to be edited, tone-mapped or further processed. Only profile C offers encoding at precisions matching those offered by OpenEXR format. The bit rate of profile C for the same quality is slightly higher. However, profile C encodes additionally a backward-compatible base layer, which is missing in OpenEXR images.

V. Conclusion

The lack of an High-Dynamic-Range (HDR) image coding standard has brought the HDR imaging community to rely on

specific vendor formats that are unsuitable for the exchange and storage of such images. This has clearly hampered the development of the HDR imaging technology so far. The new upcoming standard called JPEG XT is backwards compatible to the popular 8-bit mode of ISO/IEC 10918/ ITU Rec. T.81 (also known as JPEG), is the response to this situation. In this column we have presented the design philosophy and the technical details of JPEG XT, followed by an extensive evaluation of its performances. Objective evaluation demonstrate the robustness of the upcoming standard to the influence of its parameters: the quality for the base and extension layers and the tone mapping used for the base layer. Comparison to near-lossless and lossless existing formats has shown that the upcoming standard is capable to encode HDR imaging with high MOS of 4.5 already with a bit-rates varying from 1.1 to 1.5 bit-per-pixels, providing 23 times size reduction.

Interestingly, some of the tools developed for JPEG XT to compress HDR images may also prove useful in other application use cases. In the future, JPEG will explore how to leverage on these new mechanisms in other context. For instance, the layered structure of JPEG XT is very appealing to provide images with privacy features. Sensitive parts could be obfuscated, e.g. blurred or pixelated, in the base layer accessible to everybody whereas the extension material would contain these parts that only individual with necessary credentials could have access to. Alternately, the layered structured of JPEG XT could also provide a means to record the editing history of a particular image file. The base layer would contain the latest version of the image, whereas the extension layers would enable the ability to travel back in time and get access to earlier versions of the image. Finally, the box structure of JPEG XT makes it a natural candidate to become part of JPEG 2000 Interactive Protocol, known as JPIP, an interactive image browsing protocol similar in essence to the proprietary Google Maps.

VI. ACKNOWLEDGMENTS

REFERENCES

[1] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers, *Advanced High Dynamic Range Imaging: Theory and Practice*. CRC Press, (AK Peters Ltd), 2011.

[2] ITU, "ITU-R Recommendation BT.2020: Parameter values for ultra-high definition television systems for production and international programe exchange," International Telecommunications Union, 2012.

[3] *ITU-R Recommendation BT.709-5: Parameter Values for the HDTV Standards for Production and for International Programme Exchange*, ITU (International Telecommunication Union), Geneva, 2002.

[4] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," *Computer Graphics Forum*, vol. 22, no. 3, pp. 419–426, Sep. 2003.

[5] ITU, "Recommendation ITU-T Rec. T.800, Information technology: JPEG 2000 image coding system: Core coding system," International Telecommunications Union, 2002.

[6] ——, "Recommendation ITU-T Rec. T.832 ISO/IEC 29199-2, Information technology: JPEG XR image coding system Image coding specification," International Telecommunications Union, 2012.

[7] T. Richter, W. Husak, , N. Ajit, T. Arkady, P. Korshunov, T. Ebrahimi, A. Artusi, M. Agostinelli, S. Ogawa, P. Schelkens, T. Ishikawa, and T. Bruylants, "JPEG XT information technology: Scalable compression and coding of continuous-tone still images." [Online]. Available: http://jpeg.org/jpegxt/index.html

[8] M. Fairchild, "Fairchild's hdr photographic survey," 2008. [Online]. Available: http://rit-mcsl.org/fairchild/HDR.html

[9] H. Nemoto, P. Korshunov, and T. Hanhart, Philippe ad Ebrahimi, "Epfl's dataset of hdr images," 2015. [Online]. Available: http://mmspg.epfl.ch/hdr-eye

[10] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Trans. Graph.*, vol. 21, no. 3, p. 267, Jul. 2002.

[11] Z. Mai, H. Mansour, R. Mantiuk, P. Nasiopoulos, R. Ward, and W. Heidrich, "Optimizing a tone curve for backward-compatible high dynamic range image and video compression," *IEEE Trans. Image Processing*, vol. 20, no. 6, pp. 1558–1571, 2011.

[12] R. Mantiuk, K. Myszkowski, and H. Seidel, "A perceptual framework for contrast processing of high dynamic range images," *ACM Trans. Applied Perception*, vol. 3, no. 3, pp. 286–308, 2006.

[13] P. Korshunov, P. Hanhart, T. Richter, A. Artusi, R. Mantiuk, and T. Ebrahimi, "Subjective quality assessment database of HDR images compressed with JPEG XT," in *7th International Workshop on Quality of Multimedia Experience (QoMEX)*, 2015.

[14] M. Narwaria, R. K. Mantiuk, M. P. Da Silva, and P. Le Callet, "HDR-VDP-2.2: A calibrated method for objective quality prediction of high dynamic range and standard images," *Journal of Electronic Imaging*, vol. in print, 2015.

[15] A. Artusi, R. Mantiuk, T. Richter, P. Hanhart, P. Korshunov, M. Agostinelli, A. Ten, and T. Ebrahimi, "Overview and evaluation of the JPEG XT HDR image compression standard," *Real Time Image Processing Journal*, 2015.

Alessandro Artusi (artusialessandro4@gmail.com) is a Ramon y Cajal fellow at the University of Girona, Spain. He is a member of the ISO/IEC/JCTC1/SC29/ WG1 Committee (also known as JPEG), coeditor of the JPEG XT standard, and recipient of the Emerging Standards Awards from the British Standard Institute. He has contributed to more then 40 papers, books, courses, and patents on the high-dynamic range imaging field.

Rafał K. Mantiuk (rkm38@cam.ac.uk) is a senior lecturer at the Computer Laboratory, University of Cambridge, United Kingdom. He is the author of a popular HDR image quality metric, HDR-VDP-2; the coauthor of pfstools, software for high-dynamic range image processing; and the coauthor of more than 40 papers and book chapters on tone mapping, high dynamic range-quality assessment, and video compression.

Thomas Richter (richter@tik.uni-stuttgart.de) is a researcher at the Computing Center of the University of Stuttgart, Germany. He is one of the editors of the JPEG XT standard and the chair of the test and quality group of JPEG.

Pavel Korshunov (pavel.korshunov@epfl.ch) is a postdoctoral researcher at the Multimedia Signal Processing group at Ecole Polytechnique Fédérale de Lausanne. He has more than 50 research publications and is a coeditor of the JPEG XT standard.

Philippe Hanart (philippe.hanhart@epfl.ch) is a Ph.D. student working on quality of experience in immersive video technologies at Ecole Polytechnique Fédérale de Lausanne. He is involved in MPEG and JPEG to evaluate new high dynamic range coding formats.

Touradj Ebrahimi (Touradj.Ebrahimi@epfl.ch) is a professor at Ecole Polytechnique Fédérale de Lausanne heading its Multimedia Signal Processing group. He is also the convenor (chair) of the JPEG Standardization Committee.

Massimiliano Agostinelli (max@xdepth.com) is the chief technical officer at XDepth. He is member of the ISO/ IEC/JCTC1/SC29/WG1 Committee (also known as JPEG) and a coeditor of the JPEG XT standard.