Frédéric Dufaux,
Gary J. Sullivan,
and Touradj Ebrahimi

The JPEG XR Image Coding Standard

JPEG XR is the newest image coding standard from the JPEG committee. It primarily targets the representation of continuous-tone still images such as photographic images and achieves high image quality, on par with JPEG 2000, while requiring low computational resources and storage capacity. Moreover, it effectively addresses the needs of emerging high dynamic range imagery applications by including support for a wide range of image representation formats.

BACKGROUND

MOTIVATION

Fast-evolving digital imaging technologies have made digital photography a tremendous success with consumers and professionals. At the heart of this success story, the baseline JPEG image coding standard (ITU-T T.81 | ISO/IEC 10918-1) has played a key enabling role. However, while it has become one of the most widely used standards in the world, the nearly 20-year-old JPEG technology is reaching its limits and has begun to hinder the development of innovative features and performance enhancements in digital photography. More recently, the JPEG committee produced the JPEG 2000 standard (ITU-T T.800 | ISO/IEC 15444-1), introducing a number of novelties and new functionalities. However, it has notably higher computational resource requirements and has not made a notable impact in mobile and embedded environments such as the digital photography market. Serious photographers have been switching to using "raw" image encoding for many purposes to avoid the technology limitations of

baseline JPEG, but raw encoding has very high storage capacity requirements, is generally camera specific, and typically lacks interoperability and published format documentation.

JPEG XR (ITU-T T.832 | ISO/IEC 29199-2) is a new image coding system primarily targeting the representation of continuous-tone still images such as photographic images. It is designed to address the limitations of today's formats and to simultaneously achieve high image quality while limiting computational resource and storage capacity requirements. Moreover, it aims at providing a new set of useful image coding capabilities focused around high/extended dynamic range imagery. The "XR" part of the JPEG XR name evokes the intent of the design to apply to an "extended range" of applications beyond the capabilities of the original baseline JPEG standard. JPEG XR is based on the "HD Photo" technology developed by Microsoft to address the demands of consumer and professional digital photography. It is expected that JPEG XR will help pioneer inventive and groundbreaking products and services in the marketplace, and will bring widespread added value for consumers and professionals alike.

OBJECTIVES

JPEG XR is a new compression format supporting high dynamic range (HDR) formats for a new generation of digital cameras and other imaging applications to bring a new level of successful user experience. It is designed to give cost and compression benefits for consumer, "pro-sumer," and professional digital photography. On top of that, it also has core code stream features that support interactive Web

imaging and other flexible image interaction usage scenarios.

JPEG XR's architecture reflects the new requirements specific to high/extended dynamic range functionalities. The traditional baseline JPEG coding format uses a bit depth of eight for each of the three red green blue (RGB) color channels, resulting in 256 values per channel or 16,777,216 color values. However, more demanding applications may require a bit depth of 16, providing 65,536 representable values for each channel or over $2.8 * 10^{14}$ color values for a three-channel RGB image. Additional scenarios may necessitate even greater bit depths and sample representation formats. Conversely, when memory or processing power is at a premium, as few as five or six bits per channel may be used. To fulfill these needs and others for a very broad variety of applications, JPEG XR has been designed to include support for a wide range of image representation pixel formats, even including support for the floating-point and "radiance" formats sometimes used in the most demanding HDR applications.

Besides this special focus on HDR imagery applications, other major JPEG XR design objectives include better compression for enhanced quality, a unified system for lossless and lossy compression, cost-effective computational performance, and new progressive decoding features for more powerful image access and manipulation.

Finally, the JPEG committee has encouraged patent holders to allow JPEG XR to be implementable free of royalty and license fees to foster widespread adoption of the specification and help ensure that it can be implemented by the widest number of organizations.

ISSUING BODY, STRUCTURE OF THE STANDARD AND SCHEDULE

The Joint Photographic Experts Group (JPEG) is a working group that produces joint standards of three major international standardization organizations: the International Organization for Standardization (ISO), the International Electrotechnical Commission (IEC), and the International Telecommunication Union Telecom Sector (ITU-T). JPEG is universally recognized as the leading committee for compressed image formats, and it is responsible for the popular JPEG, JBIG, JPEG-LS, and JPEG 2000 families of imaging standards.

The JPEG committee began the standardization of JPEG XR technology in July 2007. The initial design proposal was submitted by Microsoft, based on its HD photo technology, while the subsequent development and future evolution of JPEG XR as a standard has been the responsibility of the JPEG committee.

Under the general title "Information Technology – JPEG XR Image Coding System," the JPEG XR image coding system suite of standards (ISO/IEC 29199) consists of the following current and planned parts:

■ Part 1: System architecture: This part is a nonnormative technical report (TR), describing an overview of different parts of the specifications and providing some guidelines on best encoding and decoding practices. It is now a working draft (WD), and it is expected to become a formally published TR in 2010.

■ Part 2: Image coding specification: This part specifies the JPEG XR image coding format. It was approved as a

final standard in both ITU-T (where it is known as ITU-T Recommendation T.832) in March 2009 and in ISO/IEC (as ISO/IEC 29199-2) in June 2009.

■ Part 3: Motion JPEG XR: This part specifies the use of JPEG XR encoding for stored sequences of moving images with associated timing information. The Motion JPEG XR file format is based on the ISO Base Media File Format standard. It is currently a committee draft (CD) in the ISO/IEC approval process, and it is scheduled to reach final approval status in 2010.

■ Part 4: Conformance testing: This part specifies a set of tests designed to verify whether code streams, files, encoders, and decoders meet the normative requirements specified in Part 2. The tests specified provide methods to (nonexhaustively) verify whether encoders and decoders meet these requirements. This part has currently reached the status of final committee draft (FCD) in the ISO/IEC approval process and is on target to reach final approval in 2010.

■ Part 5: Reference software: This part provides reference software for Part 2. The reference software can aid adoption of standard by providing an example implementation that can be used as a basis for making encoder and decoder products and can be used to test conformance and interoperability as well as to demonstrate the capabilities of the associated standard. It includes both encoder and decoder functionality. This part is now at FCD status and is scheduled to reach final approval in 2010.

TARGET APPLICATIONS

JPEG XR is intended for broad use in a very wide range of digital image handling and digital photography applications. Key application target areas include the following:

■ Robust and high-fidelity image acquisition technologies, such as a wide range of camera applications. Using JPEG XR, a more accurate representation of the full range of captured image signal fidelity can be retained, avoiding the bottleneck introduced by the older baseline JPEG standard and avoiding the high storage capacity and interoperability difficulties associated with camera raw image formats.

■ High dynamic range imaging workflows, including editing software suites and high-capability image processing pipelines. The ability of JPEG XR to retain an extended dynamic range with a high degree of signal fidelity can help prevent loss of quality in end-to-end workflow environments.

■ Computationally constrained signal processing environments, such as mobile and embedded applications. As the computational demands associated with JPEG XR images are substantially lower than with other coding systems such as JPEG 2000 that also provide high quality capability.

The use of JPEG XR for HDR imaging is illustrated in Figure 1, as well as a comparison with JPEG and "raw" encoding. When using JPEG, tone mapping or other image adjustments must properly be applied prior to encoding, as the encoding results in a substantial irreversible loss of information. If image adjustments are applied to an image after encoding and decoding it using the baseline JPEG format, serious quality degradation will often be evident. Raw encoding effectively supports HDR, but at the cost of high storage requirements and interoperability problems. JPEG XR successfully overcomes these shortcomings, enabling HDR while keeping efficient performance.
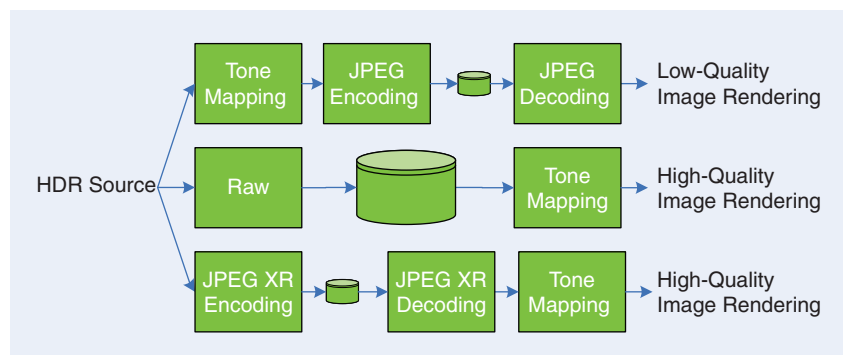
[FIG1] JPEG XR for high dynamic range imaging.

TECHNOLOGY

FUNCTIONALITIES

In this section, we discuss the technology and design features of the JPEG XR standard. Resources for further information about the standard and its design and capabilities are listed in "Resources."

The JPEG XR specification enables greater effective use of compressed imagery with broad and diverse application requirements. JPEG XR supports a wide range of color encoding formats including monochrome, RGB, CMYK, and n-component encodings using a variety of unsigned integer, fixed point, and floating-point decoded numerical representations with a variety of bit depths. The primary goal is to provide a compressed format specification appropriate for a wide range of applications while minimizing the implementation requirements for encoders and decoders. A special focus of the design is support for emerging HDR imagery applications.

JPEG XR also provides an extensive set of additional functionalities, including lossless or lossy compression support using the same signal processing operations, image tile segmentation for random access and large image formats, support for low-complexity compressed-domain image manipulations, support for embedded thumbnail images and progressive resolution refinement, embedded code-stream scalability for both image resolution and fidelity, alpha plane support, and bit-exact decoder results for both fixed and floating point image formats. These features are all achieved together with high compression capability and low computational and memory resource requirements. In fact, only simple integer signal processing operates are required even when compressing or decompressing images that use floating point representations in the decompressed domain.

The standard specifies a syntax for representing compressed image data, and it specifies the associated decoding process. These processes and representations are generic, that is, they are applicable to a broad range of applications using compressed color and gray-scale images in communications and computer systems and within mobile and embedded devices.

ARCHITECTURE

The standard specifies the syntax and semantics of JPEG XR coded images and the associated decoding process that produces an output image from the coded image. An input image is operated on by an encoder that creates a JPEG XR coded image. The decoder then operates on the coded image, producing an output image that is either an exact or approximate reconstruction of the input image

RESOURCES

The JPEG XR Standard and Technical Overview
- ITU-T Rec. T.832 | ISO/IEC 29199-2, "Information technology – JPEG XR image coding system – Image coding specification," [Online]. Available: http://www.itu.int/rec/T-REC-T.832.
- S. Srinivasan, C. Tu, S. L. Regunathan, and G. J. Sullivan, "HD Photo: A new image coding technology for digital photography," in *SPIE Appl. Dig. Image Proc. XXX*, vol. 6696, paper 6696-0A, Aug. 2007.

Draft Conformance Test Set and Reference Software
- ISO/IEC FCD 29199-4: Information technology—JPEG XR image coding system—Part 4: Conformance testing [ISO/IEC JTC 1/SC 29/WG 1 N 5017] [Online]. Available: http://www.itscj.ipsj.or.jp/sc29/open/29view/29n10428c.htm.
- ISO/IEC FCD 29199-5: Information technology—JPEG XR image coding system—Part 5: Reference software [ISO/IEC JTC 1/SC 29/WG 1 N 5020] [Online]. Available: http://www.itscj.ipsj.or.jp/sc29/open/29view/29n10430c.htm.

JPEG XR High Dynamic Range Capabilities
- S. Srinivasan, Z. Zhou, G. J. Sullivan, R. Rossi, S. L. Regunathan, C. Tu, and A. Roy, "Coding of high dynamic range images in JPEG XR/HD Photo," in *SPIE Appl. Dig. Image Proc. XXXI*, vol. 7073, paper 7073-42, Aug. 2008.

JPEG XR Signal Processing Analysis
- C. Tu, S. Srinivasan, G. J. Sullivan, S. L. Regunathan, and H. S. Malvar, "Low-complexity hierarchical lapped transform for lossy-to-lossless image coding in JPEG XR / HD Photo," in *SPIE Appl. Dig. Image Proc. XXXI*, vol. 7073, paper 7073-12, Aug. 2008.
- H. S. Malvar, G. J. Sullivan, and S. Srinivasan, "Lifting-based reversible color transformations for image compression," in *SPIE Appl. Dig. Image Proc. XXXI*, vol. 7073, paper 7073-07, Aug. 2008.

JPEG XR Performance Analysis
- F. De Simone, L. Goldmann, V. Baroncini, and T. Ebrahimi, "Subjective evaluation of JPEG XR image compression," in *SPIE. Appl. Dig.Image Proc. XXXII*, vol. 7443, paper 7443-21, Aug. 2009.
- D. Schonberg, G. J. Sullivan, S. Sun, and Z. Zhou, "Perceptual encoding optimization for JPEG XR image coding using spatially adaptive quantization step size control," in *SPIE Appl. Dig. Image Process. XXXII*, vol. 7443, paper 7443-22, Aug. 2009.

Signal Processing Background Concepts
- H. S. Malvar, "Fast progressive image coding without wavelets," in *Proc. IEEE Data Compression Conf.*, Snowbird, UT, Mar. 2000, pp. 243–252.
- H. S. Malvar, *Signal Processing with Lapped Transforms*. Boston, MA: Artech House, 1992.
- H. S. Malvar, "Biorthogonal and nonuniform lapped transforms for transform coding with reduced blocking and ringing artifacts," *IEEE Trans. Signal Processing*, pp. 1043–1053, Apr. 1998.

(depending on the fidelity represented by the encoder). The image coding specification also includes a tag-based file storage format for storage and interchange of such coded images. Furthermore, profiles and levels, which determine conformance requirements for classes of encoders and decoders, are specified. Aspects of color imagery representations and color management are also addressed in the image coding standard.

JPEG XR is designed such as to lead to a cost-effective, high-performance, embedded-system-friendly compression architecture. In particular, it requires a small memory footprint, and it is based on integer-only operations without using divides. Moreover, the signal processing structure is highly amenable to parallel processing. Thanks to its lightweight processing, JPEG XR can be effectively implemented in low-end digital signal processors (DSPs) as well as in low cost application-specific integrated circuits (ASICs), making it suitable for consumer device implementation. Furthermore, the same signal processing operations are used for both lossless and lossy compression modes.

JPEG XR handles a wide range of image formats, with many of them supporting HDR imagery. Supported color image representations include monochrome, RGB, CMYK, and n-component. The supported pixel formats cover 8- and 16-b unsigned integer, 16- and 32-b fixed point, 16- and 32-b floating point, and several packed bit layout formats such as 5-5-5, 5-6-5, and 10-10-10 bit layouts for R-G-B.

TOOLS

JPEG XR is based on a block transform design, and it uses some of the same high-level building blocks as in most image compression schemes, such as color conversion, spatial transformation, scalar quantization, coefficient scanning, and entropy coding.

The algorithm provides native support for both RGB and CMYK color types by converting these color formats to an internal luma-dominant format through the use of a reversible color transform. In addition, YUV, mono-

chrome, and arbitrary n-component color formats are supported.

A wide range of numerical encodings are supported at multiple bit depths. Formats such as 8 and 16 b, as well as additional specialized packed bit formats, are supported for both lossy and lossless compression. The 32-b formats are supported using lossy compression. Up to 24 b of source image precision are retained through the various transforms. While only integer arithmetic is used for internal processing, lossless and lossy coding are also supported for fixed-point and floating-point image data as well as for integer image formats.

The spatial transform converts the image data to a frequency domain representation. A lifting-based reversible hierarchical lapped biorthogonal transform (LBT) is used. The transform requires only a small number of integer processing operations for both encoding and decoding. It is exactly invertible in integer arithmetic and hence supports lossless image representation. The transform is based on two basic operators: the core transform and the optional overlap filtering. The core transform is similar to the widely used discrete cosine transform (DCT) and can exploit spatial correlation within a block-shaped region. Conversely, the overlap filtering is designed to exploit the correlation across block boundaries as well as mitigate blocking artifacts. It can be switched on or off by the encoder.

To enable the optimization of the quantization fidelity based on the sensitivity of the human visual system and the coefficient statistics, JPEG XR uses a flexible coefficient quantization approach controlled by quantization parameters (QPs), where the selection of a QP can vary in a flexible way across different spatial regions, frequency bands, and color channels. To further improve compression capability, interblock coefficient prediction is also applied on the transform coefficients to remove interblock redundancy. Adaptive coefficient scanning is used to convert the two-dimensional array of transform coefficients within a block into a one-dimensional vector to be encoded. The scan patterns

are adapted dynamically based on the local statistics of coded coefficients. Finally, the transform coefficients are entropy encoded. For this purpose, a variable-length coding (VLC) look-up table approach is used in which a VLC table is selected among a small set of fixed predefined tables, with the table selection being controlled adaptively based on the local statistics.

Note, however, that only the decoding process and the coding syntax are actually normatively specified in the standard. Figure 2 illustrates the various stages of the specified JPEG XR decoding process, which is basically a step-by-step inversion of the expected encoding process.

COMPARISON WITH OTHER STANDARDS

The JPEG committee finalized the baseline JPEG standard (ITU-T T.81 | ISO/IEC 10918-1) back in 1992. It has become the universal predominant standard for digital imaging, including digital photography and imagery on the Internet. Its ubiquitous adoption has made it one of the most successful international standards ever established. However, its limited functionalities are becoming an obstacle to further innovation of products and services for some key application domains.

The more recent JPEG 2000 standard (ITU-T T.800 | ISO/IEC 15444-1) has gained market adoption in some important areas, notably including digital cinema, geographic information systems (GIS), medical imaging, surveillance and identification, and the archiving of cultural heritage and photographic materials. However, JPEG 2000 has not made a notable impact in mobile and embedded environments, such as the digital photography market.

As JPEG XR is better tailored to the needs of digital photography and particularly supports emerging HDR applications and has a variety of features to provide for flexible use while retaining very practical implementation characteristics, it is expected that it will make inroads and foster innovation in the photographic application domain, leading to
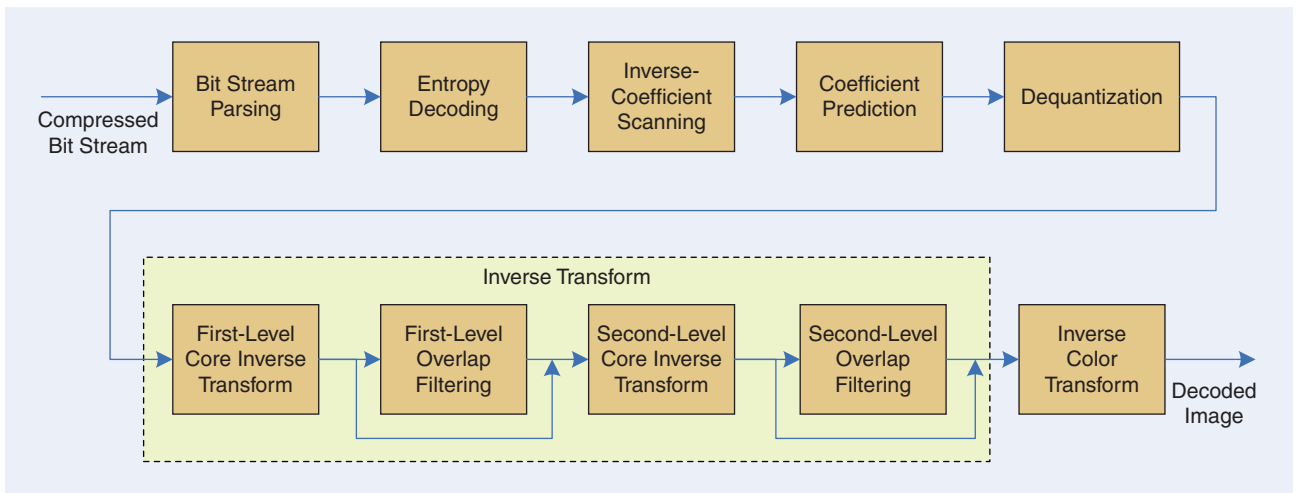
[FIG2] Basic block diagram of the normative decoding process.

TESTING AND PERFORMANCE

The JPEG committee carried out a number of experiments to thoroughly assess the performance of JPEG XR when compared to JPEG and JPEG 2000. The coding schemes have been analyzed using several objective metrics and subjective viewing tests. Test image data sets were selected for these experiments, representative of a variety of content relevant to today's digital photography. These high resolution images cover a broad gamut of content with distinct characteristics.

As JPEG, JPEG 2000 and JPEG XR are based on very different technologies, the artifacts induced by the compression systems are also very different. Reliable quality assessment is therefore especially challenging. Moreover, it is not easy to properly weight impairments of different color channels to obtain an overall assessment of the full color data. Peak signal-to-noise ratio (PSNR) has been one useful objective measurement for image quality measurement. However, PSNR is often poorly correlated with subjective image quality. For this reason, other perceptually driven objective metrics have also been used, including single-scale and multi-scale mean structural SIMilarity (SSIM), visual information fidelity – pixel (VIF-P), PSNR-HVS-M, and DC Tune measurements.

Based on extensive experimentation by several research groups, it was observed that the compression performance of JPEG XR and JPEG 2000 implementations is typically very close to each other, depending on various implementation details. Most of the time, JPEG 2000 is slightly superior, but the difference is marginal, especially in the range of bit rates which are meaningful for digital photography. In addition, both JPEG XR and JPEG 2000 consistently outperform JPEG with a significant performance gap.

To validate the results of the objective metrics, subjective testing has also been carried out. For this purpose, a double stimulus continuous quality scale (DSCQS) method has been used, adapted to deal with the evaluation of still images. A procedure from the Video Quality Experts Group (VQEG) has been used to benchmark the results in terms of prediction accuracy, monotonicity and consistency.

Finally, it is important to emphasis that JPEG XR is not standardizing the encoding process. As one of the attributes of the design is to offer a great flexibility for the encoder, it allows for future improvements and optimizations as well as competition in the marketplace. Therefore, considering that JPEG XR technology has just been emerging very recently, it is expected that future JPEG XR implementations will achieve superior performance.

FURTHER TECHNICAL DEVELOPMENTS

The JPEG committee has initiated an exploratory activity, referred to as advanced image coding (AIC), to study potential technologies for a further next generation of image compression systems. Note that the term "image" should be understood in a broad sense in this context, encompassing not only still picture but also other forms of visual content such as image sequences, stereoscopic, multiview, volumetric, and animated volumetric representations.

Closely linked to AIC, another important activity is to define superior evaluation methodologies that are appropriate for the demands of current and future imagery applications. Indeed, comprehensive and robust assessment of different imaging technologies is a critical prerequisite step towards developing new image compression systems. This proves to be a very challenging issue, due to the complexity to reliably measure image quality, either by subjective testing or objective metrics.

The JPEG Committee feels that the definition and validation of evaluation methodologies can be useful for a wide technical and scientific audience and hence foresees that this activity should lead to new standard specification, for instance in the form of a TR. In this way, all stakeholders in the digital imaging

Much work has been done and continues to be done in this area. One tool in this vein is Impulse Accelerated Technologies' (http://www.impulseaccelerated.com) Impulse CoDeveloper, which can partition a problem between an FPGA's embedded processor and custom hardware implemented in the FPGA, effectively using the FPGA logic as an accelerator for computationally expensive parts of the code.

Raising the layer of abstraction even higher, graphical DSP development environments try to eliminate the coding requirements altogether. Here, the software development environment is replaced by a drag-and-drop block diagram, where the data flow is indicated by interconnections between functional blocks. These environments can target hardware or software implementations. In a simple software targeted systems, the blocks might represent subroutines that accept data, perform a task, and return some result that is used in the next block. In a hardware targeted system, the blocks might represent logic circuits with inputs and outputs.

Two very popular graphical DSP development packages are the Mathworks Simulink (http://www.mathworks.com/products/simulink/, http://www.mathworks.com/applications/dsp_comm/), and National Instruments LabView DSP (http://www.ni.com/dsp/, ftp://ftp.ni.com/pub/devzone/pdf/tut_7006.pdf). The Mathworks MATLAB product is the classic command-line design and analysis tool for DSP, composed of numerous toolboxes that range from very generic (i.e., the signal processing toolbox) to quite specific (i.e., the filter design HDL coder). Simulink builds on these capabilities to create a graphical model-based design paradigm that can target hardware and software targets. National Instruments LabView provides similar capabilities but is more flavored toward development and implementation than analysis. National Instruments also supports LabView directly with an extensive line of hardware for data acquisition and control. In a relatively recent development, LabView now has graphical blocks that will run MATLAB m-file scripts, and there is even

support for converting Simulink models for use in LabView.

CONCLUSION

There are many pitfalls, some of them quite subtle, that can derail a DSP project as it moves into the hardware implementation stage. Since many signal processing experts may have little experience in computer engineering, we've tried to provide in this article some recommendations and Web resources to help the reader avoid as many of these problems as possible.

AUTHORS

Michael Morrow (morrow@engr.wisc.edu) is a faculty associate in the Electrical and Computer Engineering Department, University of Wisconsin-Madison.

Cameron Wright (ChgW@uwyo.edu) is an associate professor in the Electrical and Computer Engineering Department, University of Wyoming.

Thad Welch (thadwelch@boisestate.edu) is a professor and chair of the Electrical and Computer Engineering Department, Boise State University. [SP]

---

value chain could benefit from this knowledge to assess their image compression implementations as well as the associated capture and display devices.

PRODUCTS

JPEG XR (or its compatible HD Photo predecessor) is supported in the Microsoft Windows Vista and Windows 7 operating systems and in some configurations of Windows XP. This includes support of various features such as image viewing in Windows Photo Gallery and Windows Live Photo Gallery, thumbnail preview in the Windows Explorer shell, and image management in the Windows Imaging Component (WIC) system. Various other Microsoft-led efforts have included support for the feature, such as the SeaDragon, Photosynth, Image Composite Editor, and HD View initiatives.

Support for JPEG XR images is a key feature of the ECMA 388 Open XML

Paper Specification (OpenXPS), which specifies a set of conventions for the use of XML and other widely available technologies to describe the content and appearance of paginated documents. OpenXPS readers and printers thus include support for JPEG XR images.

Corel Paint Shop Pro X2 includes integrated support for HD Photo / JPEG XR images, and preliminary plug-in support has been distributed for the Adobe Photoshop environment (for Windows and Mac operating systems).

Megachips Corporation of Japan has announced JPEG XR support in its embedded systems for mobile applications such as camera signal processing. Pegasus Imaging Corp. has included support for the new standard in its digital imaging toolkit products.

Now that JPEG XR has been approved as a final standard by ITU-T and ISO/IEC, support for JPEG XR imagery is begin-

ning to emerge in a very broad variety of other product environments.

AUTHORS

Frédéric Dufaux (frederic.dufaux@epfl.ch) is a senior researcher at the Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. He is an active member of the JPEG committee.

Gary J. Sullivan (garysull@microsoft.com) is a video/image technology architect with Microsoft Corporation, Redmond, Washington. He is the lead editor of the JPEG XR standard and the rapporteur of the ITU-T visual coding experts group (VCEG) that manages the ITU-T side of the JPEG partnership. He is an active member of the JPEG committee.

Touradj Ebrahimi (touradj.ebrahimi@epfl.ch) is a professor at the Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. He is an active member of the JPEG committee. [SP]