JPEG 2000標準ファミリー：次世代画像処理技術の礎石

The JPEG 2000 family of standards : the cornerstone of next generation image processing

マーチン ボリック* マイケル ゴーミッシュ*
Martin BOLIEK Michael J. GORMISH

要　旨

ISO標準であるJPEG2000は画像圧縮方式を規定するだけでなく，これまでの圧縮画像フォーマットでは前例のない多彩な機能を有する．用途に応じたフレキシブルな画像アクセス，データの双方向性，交換性に優れ，豊富なメタデータ，画像の特徴解析情報などを持っている　またその応用用途も静止画写真はもちろんのこと，文字/写真混在文書，動画，地図，衛星写真，マルチスペクトル画像，医療画像などに及ぶ．本文ではJPEG2000標準群（12のパートで構成）のそれぞれのユニークな機能紹介とそれらの組み合わせにより可能な応用例を紹介する．特に基本デコーダー（Part1），動画を扱う（Part3），文字／写真混在文書（Part 6），双方向性プロトコル（Part 9），についてはアプリケーション事例を交えて詳しく紹介する．

ABSTRACT

JPEG 2000 is a new family of compression standards that allow unprecedented access, interactivity, interchange, metadata, and analysis with a variety of image types including still images, compound documents (text and imagery), motion images, geographic, satellite, and multi-spectral images, and medical images. This paper outlines the unique functionality of the twelve standards in this family and shows how they can work together. Special emphasis is given to the base decoder (Part 1), the motion images (Part 3), the compound document images (Part 6), and the interactive protocol (Part 9). Examples of applications that are enabled by these standards are given.

* California Research Center, Ricoh Innovations, Inc.

1. Introduction to JPEG 2000

The JPEG 2000 family of standards is a new type of image compression. While the base compression system (Part 1) achieves better rate-distortion than JPEG and other continuous-tone compressors over a range of rates and image types, the real advantage is in the ability to access just the image data needed *in the compressed form* for specific applications.

Wavelet transformation provides both energy compaction and separate resolutions descriptions. Bit-plane entropy coding provides compression and progressive bit-rate control. Tiling and precincts (tiling in the coefficient domain) enable specific regions to be independently coded. The syntax exposes the codestream segments and the relation to a reconstructed image.

Sophisticated file format options enable rich metadata, multiple codestreams for video frames, compound masking and layering, etc. The JPIP interactive protocol standardizes the request and response conversation needed for a "client/server" or other interactive reconstruction and progressive "deepening" of partial images.

JPEG 2000's rich image analysis (tiling, transformation, entropy coding) exposed with syntax and coupled with state-of-the-art file format and interactive protocols, enable a number of applications that simply were not possible before. This paper introduces the entire family of JPEG 2000 standards, both completed and emerging, and offers three examples of image applications that take advantage of JPEG 2000's unique capabilities.

2. JPEG 2000 in Products Today

The JPEG 2000 family of standards is now being used in a number of commercial products. The base codec is implemented in Apple Computer's Quicktime multimedia resources, enabling many programs to use JPEG 2000. The latest Adobe products, such as Acrobat and Photoshop, support JPEG 2000. Kodak has developed a video system that allows streaming data at the desired resolution, region-of-interest, and bit-rate. Geographic data is represented in JPEG 2000 form for the pan and zoom advantages. JPEG 2000 has been added to the DICOM format for medical images [1]. Even Japanese driver's license pictures now use JPEG 2000. JPEG 2000 is also being considered as the transfer format for digital theaters by the Digital Cinema Initiatives [2].

Overall, JPEG 2000's adoption rate is quite good relative to most standards. It is being used primarily for high quality and large imagery — areas where the original JPEG is unsatisfactory. Of course, the trajectory of most areas of the imaging market is toward this high end.

3. JPEG 2000 Family of Standards

JPEG 2000 has a number of parts that may be used together to achieve new functionality.

Part 1: Minimal decoder — A complete codestream, file format (jp2), and decoder specification, this part of the standard is commonly referred to as JPEG 2000. Fig.1 shows the basic

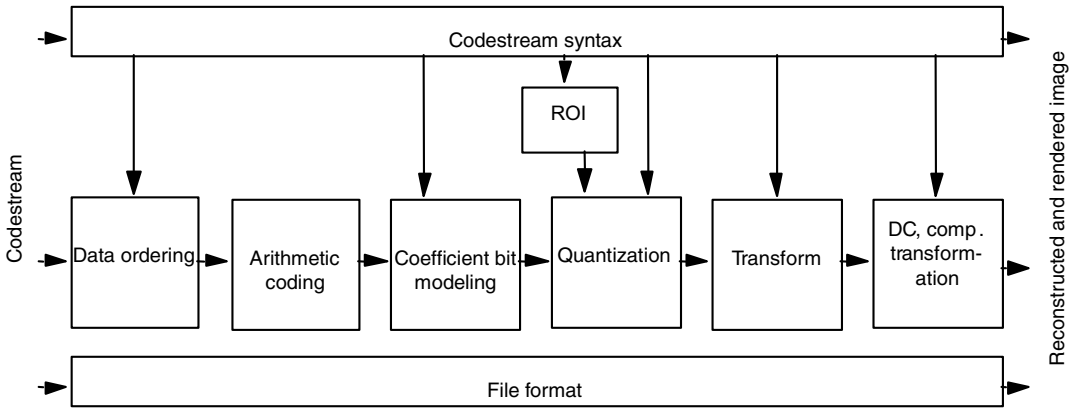

Fig.1 JPEG 2000 Part 1 decoder block diagram

architecture. The wavelet transform offers progression by resolution, the bit-plane entropy coder enables progression by bit-rate, and tiles and precincts permit regional access. Also included is a sophisticated syntax that maps the compressed data segments to the uncompressed image and the jp2 file format [3][4].

Part 2: Extensions — Included in this part are significant file format extensions (.jpx), arbitrary wavelet definitions, trellis coded quantization, wavelet packets, three-dimensional transformation, and other extensions [5].

Part 3: Motion — File format for motion and audio data file format using the structure of Apple's Quicktime and MPEG-4. Enables the same resolution and bit-rate control for video [6].

Part 4: Compliance — Tools and methods for determining the conformance of codestreams and decoders [7].

Part 5: Reference software — Two separate versions of source code implementations, in C (Jasper) and in Java (JJ2000) [8].

Part 6: JPM mixed raster content — File format (.jpm) that allows the masking and blending of image objects. This part is particularly useful for document imagery [9].

Part 7: Hardware profile — Profile intended to enable the manufacture of early VLSI implementations. This profile was moved to an amendment of Part 1.

Part 8: JPSEC security — Enables a number of security tools, such as encryption, watermarking, hashing, etc. to be executed on all or part of a JPEG 2000 bitstream. Tools are signaled at the codestream syntax (not file format) level [10].

Part 9: JPIP interactive protocol — Protocol for partial and interactive transmission of image data. Enables client/sever communication and progressive build up of images [11].

Part 10: JP3D — Isotropic approach to compressing three-dimensional data [12]. This is particularly useful for spatial data.

Part 11: JPWL wireless — Methods for error protection on noisy channels [13].

Part 12: Common text — This text provides for an ISO base media format for both JPEG Part 3 and MPEG standards [14].

Part 13: Encoder — A specification of an encoder [15].

JP Search — New work item to standardize metadata to aid in image and data retrieval [16].

4. Interactive Image Example

Using JPEG 2000 (Part 1) and JPIP (Part 9) together creates a powerful client/server system that interacts with the compressed codestream on the server to send only the appropriate compressed data to the client. In an office environment, this ability to repurpose or reuse high quality, high-resolution imagery by changing resolution, region-of-interest, and bit-rate for display on computer monitors and cell phone displays is critical.

Fig.2 shows an image server, such as an MFP, on the right hand side, which would typically have a disk with high resolution images stored in the JPEG 2000 Part 1 file format (image/jp2 mime type). A client device, shown at the left side of Fig.2, can issue a series of requests for a particular image using syntax defined by JPIP. These requests indicate a region-of-interest for the image, a maximum size of interest, and possibly other information as well. A typical first request is for a version of the whole image at the resolution of the display device.

The server, upon receiving a request is able to select those parts of the JPEG 2000 file that are necessary for the type of viewing being done by the client. In the case of a 600 dpi image being displayed on a 75 dpi monitor, it may be possible to send less than 1/100 of the data for the image. The bandwidth savings obtained by only sending the necessary part of the data is typically much higher than the compression obtained by using the visual system models and statistical models of JPEG and JPEG 2000.

The server assembles (by simple extraction and concatenation) the data requested by the client and returns it in a format specified by JPIP. This format is mainly a copy of a particular part of the JPEG 2000 file, along with header information indicating what part of the JPEG 2000 file has been copied.

In Fig.2, after the client device has received the data and constructed the image, suppose the user decides to zoom in on part of the image. This leads to a second request for more high-resolution data, but only for a small spatial region of the image.

When the server receives this second request, relevant data is again selected from the JPEG 2000 file. With the second and successive requests, the data that has been previously sent need

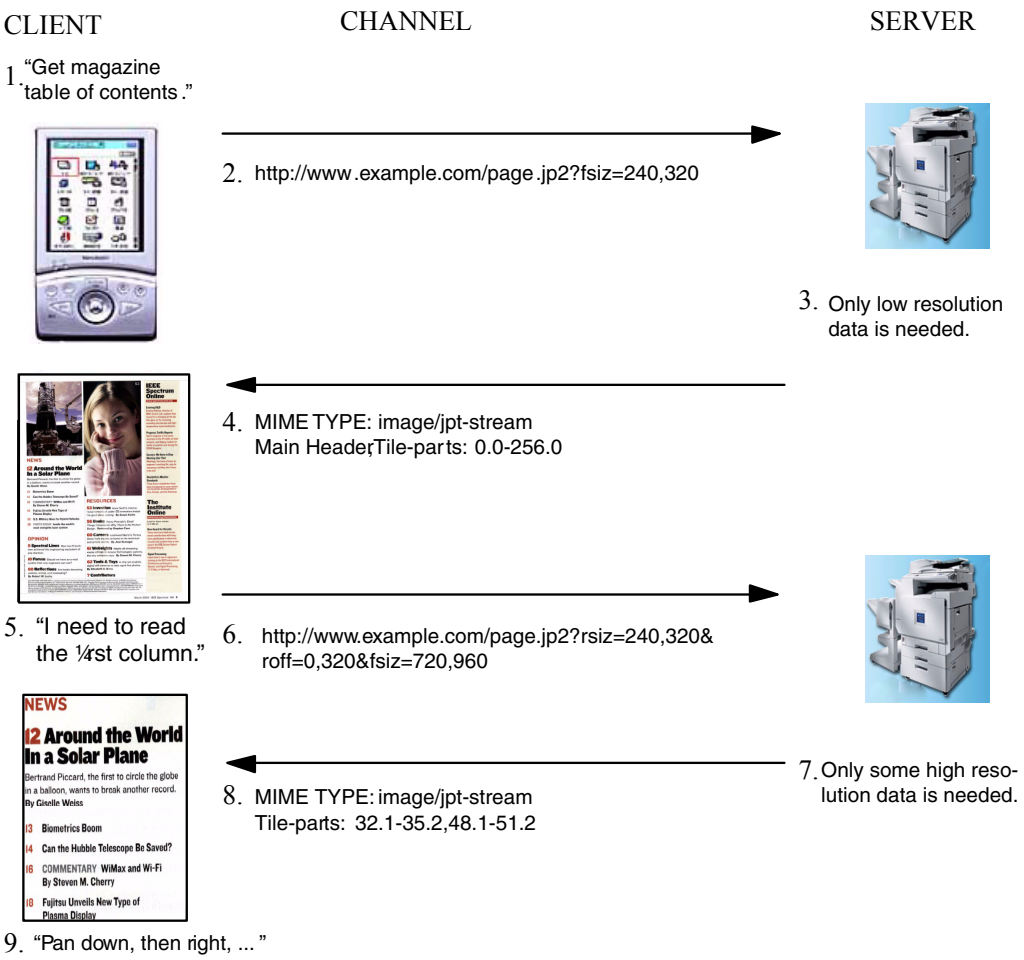

Fig.2 JPIP client/server interaction example

not be transmitted again. This provides additional bandwidth savings.

In order to obtain the savings from successive requests two things are necessary. First the client must save the data it received from the previous requests. Second the server must know what data the client already has when it forms the response to the new request. The JPIP standard provides two methods to for the server to know what data the client has. In one case, the server keeps track of all the data sent to a client, and assumes the client keeps it, unless the client sends a message saying some data was discarded. In the second case, the client simply sends a list of the data that it already has with each request made to the server. In this case, the server can operate without any state.

5. Compound Document Example

While JPEG 2000 Part 1 provides a state-of-the-art continuous-tone wavelet-based still image compression system, there are better compression techniques for line art, some graphics, and half-toned images. JPM (Part 6) allows such binary compressors to be used for appropriate parts of a document and JPEG 2000 to be used in the continuous-tone regions. Fig.3 shows a document image segmented into a foreground image, a background image, and a binary mask image that controls the remixing of the foreground and background.

Perhaps most importantly, however, the same segmentation used for compression advantage can also be used to divide the document image into objects with greater semantic meaning. For example, two different text paragraphs can be stored as separate objects. Perhaps these paragraphs have a specific reading order,

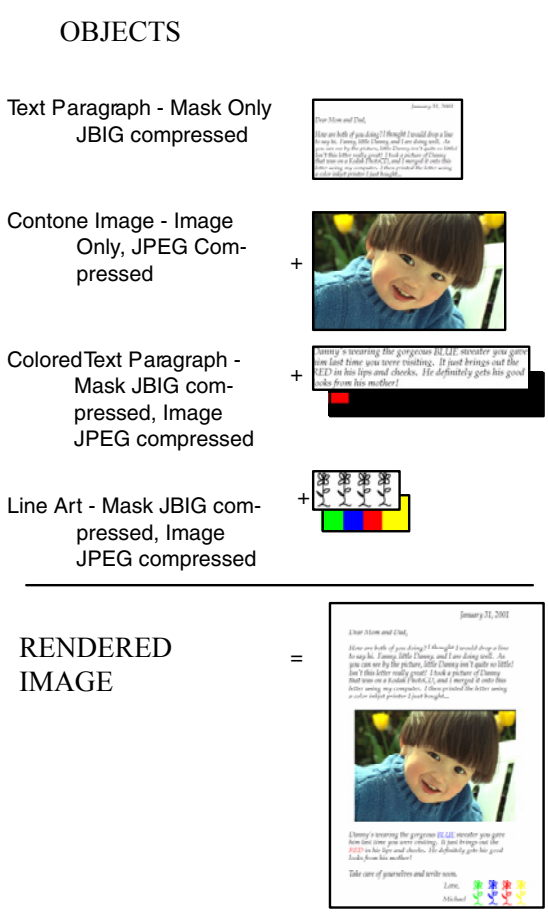

Fig.3 JPM example

but perhaps they might be "reformatted" to allow sorting or truncation or other application-dependent purpose.

In another example, simple edits can be made, e.g. "punch holes" can be removed by adding an opaque object of the background color on top. This object could later be removed from the image, allowing the original to be obtained ("undo"). Thus, it is possible to edit with undo, even after saving the file.

Objects can also be used to change the layout of a document, whether or not the contents of the objects are understood.

The JPM file format also allows multiple pages to be stored in a file. Indeed with all features of JPM, it is usable in most of the applications where PDF is used, but with substantially less computational complexity.

6. Motion Image Example

Motion JPEG 2000 (JPEG 2000 Part 3) defines a file format, which allows multiple JPEG 2000 codestreams to be stored for video. Because multiple resolutions and bit-rates can be extracted from a JPEG 2000 file, Motion JPEG 2000 provides a scalable video compression system. Different resolutions, bit-rates, and frame rates can be selected from the same file, depending on the display size and computational capabilities of a decoder. Fig.4 illuminates the use of Motion JPEG 2000.

Unlike most video compression systems, e.g. MPEG-1, MPEG-2, and H.263, Motion JPEG 2000 does not allow the use of information from one frame to be used when compressing another frame (no motion compensation). This limits the ability to use Motion JPEG 2000 for extremely high compression. However, since every frame is coded independently, and the compression performance of JPEG 2000 on a single frame can be excellent (even lossless), Motion JPEG 2000 is ideal for video editing, medical video applications, security surveillance, any low-latency applications, or even use over noisy channels where some frames may be lost.

The technical advantages of Motion JPEG 2000, along with the lack of "per play" charges, has led the Digital Cinema Initiative to recommend JPEG 2000 to seven major movie companies [2].

7. Ricoh and JPEG 2000

Ricoh's involvement was key to the creation of the JPEG 2000 standard. The Ricoh algorithm, "Compression with Reversible Embedded Wavelets (CREW)," showed the possibilities for this technology [17]. With this example the JPEG committee requested that Ricoh write a proposal for a new work item. The proposal was accepted and the call for proposals was answered in November 1997. Ricoh contributed a number of key technologies to many of the standard's parts and has taken on the role of editor, or co-editor, for parts 1, 2, 4, 7, and 13.

Appropriate video resolution, frame bit-rate, region-of-interest, and frame rate for each device and application.

Fig.4 Motion JPEG 2000 example

8. Conclusions

JPEG 2000 is the first image compression system to offer high-quality performance with the features necessary for a number of applications. JPEG 2000, for example, is the first compression system to enable interactive extraction, transmission, and reconstruction of compressed codestream segments. Thus, bandwidth is truly minimized for a given application.

High quality video imaging, compound document masking and layering, are among the many features enabled by the sophisticated file formats.

References

1. "Supplement 61: JPEG 2000 Transfer Syntaxes," Digital Imaging and Communications in Medicine (DICOM), January 14, 2002.
2. "Digital Cinema System Specification v4.2," Digital Cinema Initiatives, LLC, August 23, 2004.
3. ITU-T Rec. T.800 | ISO/IEC 15444-1:2002, "Information technology -- JPEG 2000 image coding system -- Part 1: Core coding system."
4. D. S. Taubman, M. W. Marcellin, *JPEG 2000 Image Compression Fundamentals, Standards, and Practice*, Kluwer Academic Publishers, Boston, MA, 2002.
5. ITU-T Rec. T.801 | ISO/IEC 15444-2:2004, "Information technology -- JPEG 2000 image coding system: Extensions."
6. ISO/IEC 15444-3:2002, "Information technology -- JPEG 2000 image coding system -- Part 3: Motion JPEG 2000."
7. ISO/IEC 15444-4:2002, "Information technology -- JPEG 2000 image coding system -- Part 4: Conformance testing."
8. ISO/IEC 15444-5:2003, "Information technology -- JPEG 2000 image coding system: Reference software."
9. ISO/IEC 15444-6:2003, "Information technology -- JPEG 2000 image coding system -- Part 6: Compound image file format."
10. JPSEC CD V.2.0, "Information technology -- JPEG 2000 image coding system -- Part 8: Secure JPEG 2000," ISO/IEC JTC1/SC29/WG1 N3397, July 22, 2004.
11. JPIP FDIS, "Information technology -- JPEG 2000 image coding system -- Part 9: Interactivity tools, APIs and protocols," ISO/IEC JTC1/SC29/WG1 N3358, July 19, 2004.
12. JP3D WD V4.0, "Information technology -- JPEG 2000 image coding system -- Part 10: "Extensions for Three-Dimensional Data and Floating Point Data," ISO/IEC JTC1/SC29/WG1 N3369, July 22, 2004.
13. JPWL CD, "Information technology — JPEG 2000 image coding system — Part 11: Wireless JPEG 2000," ISO/IEC JTC1/SC29/WG1 N3386, July 22, 2004.
14. ISO/IEC 15444-12:2004, "Information technology -- JPEG 2000 image coding system -- Part 12: ISO base media file format."
15. JPEG 2000 Encoder WD 1.1," "Information technology — JPEG 2000 image coding system — Part 13: An entry-level JPEG 2000 encoder," ISO/IEC JTC1/SC29/WG1 N3381, July 22, 2004.
16. JPSearch, "Proposal for a new work item, information technology – JPSearch," ISO/IEC JTC1/SC29/WG1 N3374 and N3405, July 22, 2004.
17. M. Boliek, M. Gormish, E. Schwartz, A. Keith, "Decoding compression with reversible embedded wavelets (CREW) codestreams," *Electronic Imaging*, Vol. 7, No. 3, July 1998.