ISO/IEC JTC 1/SC 29/WG 1 **N1105R**

Date: November, 1998

ISO/IEC JTC 1/SC 29/WG 1
(ITU-T SG8)

Coding of Still Pictures

JBIG	**JPEG**
Joint Bi-level Image Experts Group	Joint Photographic Experts Group

TITLE:	JPEG2000 requirements and profiles version 4.0
SOURCE:	Requirements ad hoc group
PROJECT:	JPEG2000
STATUS:	Draft in progress
REQUESTED ACTION:	Provide feed-back and comments to update
DISTRIBUTION:	WG1 Los Angeles Meeting, WG1 Web pages.

Contact:
ISO/IEC JTC 1/SC 29/WG 1 Convener - Dr. Daniel T. Lee
Hewlett-Packard Company, 11000 Wolfe Road, MS42U0, Cupertino, California 95014, USA
Tel: +1 408 447 4160, Fax: +1 408 447 2842, E-mail: Daniel_Lee@hp.com

1 Introduction

This new work item is intended to create a new image coding system that for different types of still images (bi-level, gray-level, color, multi-component) with different characteristics (natural images, scientific, medical, remote sensing imagery, text, rendered graphics, etc.) allowing different imaging models (client/server, real-time transmission, image library archival, limited buffer and bandwidth resources, etc.) preferably within a unified system.

This coding system is intended to provide low bit-rate operation with rate-distortion and subjective image quality performance superior to existing standards, without sacrificing performance at other points in the rate-distortion spectrum. In addition, this system could include many modern features, listed in this document.

T he standard will strive for openness and royalty-free licensing. It should be completed by the end of the millennium and offer state-of-the-art compression for at least ten years and beyond.

This standard will serve still image compression needs that are current not served by the current JPEG standards (IS 10918-1, IS 10918-2, IS 10918-3). (For example, very low bit-rate, progression for the WWW, medical imagery, pre-press, etc.) It is intended to compliment, not replace, the current JPEG standards. Indeed, this standard is expected to include an architectural context that will allow the previous standards to be used as desired on different tiles and/or components *within a single image*. This architecture will allow utilizing all of the current and future work of the JPEG committee to the best advantage.

Vision for JPEG 2000

As digital imagery expands in quality, size, and application there is a greater need for image compression with great flexibility and efficient interchangeability. JPEG 2000 is a new type of image compression system. While offering state-of-the-art compression, JPEG 2000 also offers unprecedented access into the image while still in compressed form. Thus, images can be accessed, manipulated, edited, transmitted, and stored in a minimal information form. JPEG 2000 is a compressed format capable of being the foundation framework for image processing systems. Further, with a capable yet computationally modest decoder, interchangeability between applications is straight forward.

Note to editor: Open standard concept and identification and protection should be included.

2 Purpose and justification

This standard is intended to advance standardized image coding systems to serve applications into the next millennium. It will provide a set of features vital to many high-end and emerging imaging applications by taking advantage of new modern technologies. Specifically, this new standard will address areas where current standards fail to produce the best quality or performance including the following. It will also provide capabilities to markets that currently do not use compression.

- **Low bit-rate compression performance**: Current standards, such as IS 10918-1 (JPEG), offer excellent rate-distortion performance in the mid and high bit-rates. However, at low bit-rates (e.g., below 0.25 bpp for highly detailed gray-level images) the distortion, especially when judged subjectively, becomes unacceptable.
- **Lossless and lossy compression**: There is no current standard that can provide superior lossless compression and lossy compression in a single codestream.

- **Large images:** Currently, the JPEG image compression algorithm does not allow for images greater then 64K by 64K without tiling.

- **Single decompression architecture:** The current JPEG standard has 44 modes, many of which are application specific and not used by the majority of the JPEG decoders. Greater interchange between applications can be achieved if a single common decompression architecture encompasses these features.

- **Transmission in noisy environments**: The current JPEG standard has provision for restart intervals, but image quality suffers dramatically when bit errors are encountered.

- **Computer generated imagery:** The current standard was optimized for natural imagery and does not perform well on computer generated imagery.

- **Compound documents:** Currently, JPEG is seldom used in the compression of compound documents because of its poor performance when applied to bi-level (text) imagery.

3 Objectives for next generation compression

JPEG 2000 will be defined by its features. It will fill a gap in the rate-distortion spectrum (low bit-rate) of current still image compression standards and provide a set of features vital to many high-end and emerging image applications. *Superior low bit-rate performance is primary*. It is desirable to include as many of the other following features as possible.

- **Superior low bit-rate performance**: This standard should offer performance superior to the current standards at low bit-rates (e.g., below 0.25 bpp for highly detailed gray-scale images). This significantly improved low bit-rate performance should be achieved *without sacrificing performance on the rest of the rate-distortion spectrum.* Examples of applications that need this feature include network image transmission and remote sensing. This is the highest priority feature.

- **Continuous-tone and bi-level compression**: It is desired to have a standard coding system that is capable of compressing both continuous-tone and bi-level images. If feasible, this standard should strive to achieve this with similar system resources. The system should compress and decompress images with various dynamic ranges (e.g., 1 bit to 16 bit) for each color component. Examples of applications that can use this feature include compound documents with images and text, medical images with annotation overlays, and graphic and computer generated images with binary and near to binary regions, alpha and transparency planes, and facsimile.

- **Lossless and lossy compression**: It is desired to provide lossless compression naturally in the course of progressive decoding (difference image encoding, or any other technique, which allows for the lossless reconstruction is valid). Examples of applications that can use this feature include medical images where loss is not always tolerated, image archival applications where the highest quality is vital for preservation but not necessary for display, network applications that supply devices with different capabilities and resources, and pre-press imagery.

- **Progressive transmission by pixel accuracy and resolution**: Progressive transmission that allows images to be reconstructed with increasing pixel accuracy or spatial resolution is essential for many applications. This feature allows the reconstruction of images with different resolutions and pixel accuracy, as needed or desired, for different target devices. Examples of applications include the World Wide Web, image archival applications, printers, etc. The image architecture provides for the efficient delivery of image data in many applications such as client/server applications (World Wide Web).

- **Fixed-rate, fixed-size, limited workspace memory**: Fixed-rate (fixed local rate) means that the number of bits for a given number of consecutive pixels equals (or is less than) a certain value. This allows the decoder to run in real-time through channels with limited bandwidth. Examples are remote imaging, motion coding, etc. Fixed-size (fixed global rate) means that the total size of the codestream for a complete image equals a certain value. This allows hardware with a limited memory space to hold the complete codestream regardless of the image. Examples include scanners, printers, etc.

- **Random codestream access and processing**: Often there are parts of an image that are more important than others. This feature allows user defined Regions-Of-Interest (ROI) in the image to be randomly accessed and/or decompressed with less distortion than the rest of the image. Also, random codestream processing could allow operations such as rotation, translation, filtering, feature extraction, scaling, etc.

- **Robustness to bit-errors:** It is desirable to consider robustness to bit-errors while designing the codestream. One application where this is important is wireless communication channels. Portions of the codestream may be more important than others in determining decoded image quality. Proper design of the codestream can aid subsequent error correction systems in alleviating catastrophic decoding failures.

- **Open architecture:** It is desirable to allow open architecture to optimize the system for different image types and applications. This may be done either by the development of a highly flexible coding tool or adoption of a syntactic description language which should allow the dissemination and integration of new compression tools. It is desired to allow the user to select tools appropriate to their application and provide for future growth. With this feature, the decoder is only required to implement the core tool set and a parser that understands the codestream. If necessary, unknown tools are requested by the decoder and sent from the source.

- **Sequential build-up capability (real time coding)**: The standard should be capable of compressing and decompressing images with a single sequential pass. This standard should also be capable of processing an image using component interleave order or non-interleaved order. During compression and decompression, the standard should use context limited to a reasonable number of lines. However, there is no requirement of optimal compression performance during sequential build-up operation.

- **Backwards compatibility with JPEG:** It is desirable to provide for backwards compatibility (or easy transcoding) with the current JPEG standards.

- **Content-based description**: Finding an image in a large database of images is an important problem in image processing. For example, a doctor could request only images from a set that are recognized to have a certain type of tumor. This could have major applicability to the medical, law enforcement and environmental communities, and for image archival applications. Regardless of the techniques used, JPEG 2000 should strive to provide the opportunity for solutions to this problem.

- **Protective image security**: Protection of a digital image can be achieved by means of one or more of methods such as: watermarking, labeling, stamping, fingerprinting, encryption, scrambling, etc. Watermarking and fingerprinting are invisible marks set inside the image content to pass a protection message to the user. Labeling is already implemented in SPIFF and must be easy to transfer back and forth to JPEG 2000 image file. Stamping is a mark set on top of a displayed image that can only be removed by a specific process. Encryption, and scrambling can be applied on the whole image file or limited to part of it (header, directory, image data) to avoid unauthorized use of the image.

? Note: The issue of whether JPEG2000 should cover specific solutions to encryption, watermarking, authentication algorithmsshould be revisited

- **Interface with MPEG-4**: The ongoing standardization process for the "Coding of moving pictures and audio" (ISO/IEC JTC1/SC29/WG11), MPEG-4, is developing a content based coding scheme in which coding tools are chosen from a repertoire so as to address in an optimal way a wide range of functionalities. It is desirable that the coding tool (or tools) developed for the compression of still images in JPEG 2000 are provided with an appropriate interface allowing the interchange and the integration of such tools into the framework of a syntax oriented coding scheme such as MPEG-4. In particular IPR information should be maintained in both insertion and extraction of still pictures to/from moving images.

- **Side channel spatial information (transparency)**: Side channel spatial information, such as alpha planes and transparency planes, are useful for transmitting information for processing the image for display, print, or editing, etc. An example of this is the transparency plane used in World Wide Web applications.

- Object Based Functionality -
  Object based composition: The current standard was designed to encode/decode rectangular framed images, which made it difficult to represent specific regions within the image. Multiple objects with arbitrary shape (or transparency) information will enable users to composite different still images with as much flexibility.
  Object based information embedding: The side information (e.g., description of object) attached to each object will enable the new standard to provide more efficient way to digital image indexing and retrieval

?

It is important to note that, while this standard addresses the needs of this wide variety of images and applications, it desirable that the decoder be able to interpret codestreams with minimal complexity. There must be a cost/benefit assessment of each function that adds complexity to the decoder.

4 Markets and applications

JPEG 2000's desired capabilities shall serve markets and applications such as:
?

- **Internet**
- **Color Facsimile**
- **Printing**
- **Scanning (Consumer, Pre-Press)**
- **Digital Photography**
- **Remote Sensing**
- **Mobile**
- **Medical**
- **Digital Library**
- **E-Commerce**

?

5 Application Requirements

This section contains requirements of specific applications that were envisioned as major users of the JPEG2000 standard. The goal of this section is to identify these requirements and to ensure that appropriate tools and technologies are adopted in JPEG2000 standard to fulfill them.

5.1 Image Type

Definition

Width, height, component bit depth, number of components, and the span of each component (sub-sampling) determine the image type. Also the source of the image should be considered.

Aggregate requirements

The image type should not be limited by the algorithm. The codestream syntax should be capable of the following:

Image width and height	1 to (2^{32} – 1)
Component depth	1 to 32 bits
Number of components	1 to 255 (or more)

Dissimilar component depths (each component can be a different depth)

Dissimilar component spans (each component can have a different coverage)

Changes to the application profiles

Internet: The JPEG2000 should be able to process compound images, with sizes from 32 x 32 up to at least 4K x 4K pixels with 1, 3 (Y, RGB, YUV,…) or 4 components including alpha channel and from 1 to 8 bits/component precision.

Color Facsimile: JPEG2000 should be able to process compound images, with sizes, number of planes and bit depth as specified by the color Fax standard

Printing: The JPEG2000 should be able to process compound images, with typical sizes of 4800 by 6600 pixels (600ppi, 8in by 11in image) with 1, 3, and 4 components, and 8 bits/component precision.

Scanning: The JPEG2000 should be able to process compound images, with typical sizes of 10K x 10K up to at least 20K x 20K pixels with 1, 3 and 4 components, and up to 16 bits/component precision.

Digital Photography: The JPEG2000 should be able to efficiently process natural images, with sizes of at least up to 4K x 4K pixels with 1, 3 components (with spatially correlated components), with a minimum of 8 bits/component and a maximum of 16 bits/component precision. In case of a three component image, JPEG2000 should be able to handle both the full image case, where all 3 components are present for each pixel, and the Color Filter Array (CFA) case, where only 1 of the 3 components is present for each pixel, the components varying in a regular fashion.

[Note: need revision of statement for CFA requirement and addition to glossary for CFA definition.]

Remote Sensing: The JPEG2000 should be able to process infra-red, electro-optical, multi-spectral, hyper-spectral, and SAR images, with sizes from 256 x 256 up to 64K x 64K pixels with 1 up to 500 components, and 8 up to 20 bits/component precision. It should also be able to process one component data with complex pixels (i.e phase information).

Mobile: The JPEG2000 should be able to process compound images, with sizes from 32 x 32 up to at least 4K x 4K pixels with 1 or 3 components (Luminance, RGB, …) and 1 to 8 bits/component.

Medical: The JPEG2000 should be able to process natural images, with sizes from 32 x 32 to at least 10K x 10K pixels with 1 and 3 (Luminance, RGB, …) or 4 components (plus alpha) and up to 16 bits/component.

Digital library: Same as the Internet.

E-commerce: Same as the Internet.

5.2 Uncompressed (no compression)

Definition

The image is stored in the bitstream without compression. This allows applications to take advantage of the JPEG2000 bitstream syntax and its associated features without compressing the image data.

Aggregate requirements

JPEG2000 should provide a mode in which the image is archived as an uncompressed form, leaving other syntax elements (embedding information, etc.) intact.

5.3 Lossless Compression

Definition

The reconstructed image is identical, bit for bit, to the original image. A typical compression ratio for this type of compression is 2:1 using JPEG LS.

Aggregate requirements

Provide performance at least as good as JPEG LS. Performance includes speed, complexity, memory requirements, etc.

5.4 Visually Lossless Compression

Definition

The reconstructed image may differ numerically from the original image, but any such differences are not perceptible under normal viewing conditions. Current baseline JPEG can generally attain visually lossless performance at rates between 1 and 2 bpp for grayscale images and slightly lower rates, in bpp per band, for color images.

Aggregate requirements

Provide 30% improvement over JPEG baseline.

5.5 Visually Lossy Compression

Definition

The reconstructed image contains perceptible differences from the original image under normal viewing conditions. Current baseline JPEG generally exhibits visually lossy performance at rates under 1 bpp for grayscale images and lower rates, in bpp per band, for color images.

Aggregate requirements

Provide 30% improvement over JPEG baseline.

5.6 Progressive Spatial Resolution

Definition

Ability to extract lower resolution images from a codestream without redundant decoding.

Aggregate requirements

At least 9 resolution levels. Should be combinable with progressive quality resolution (see next section).

Note: Should be worded to consider original size and lower limit for image size.

Changes to the application profiles

<u>Color facsimile</u>: Should also require progressive spatial resolution. This would enable facsimile systems where the resolution capability of the destination facsimile is determined and the appropriate resolution sent. This is useful for Internet facsimile, facsimile from a database and "broadcast" facsimile. It would also allow facsimile standards to incorporate generational improvements in performance without altering the standard.

<u>Printing</u>: Should also require progressive spatial resolution. This would enable images to be sent to printers of different resolution. Also quick resolution scaling of images could be performed as called for by the user, page-description language, or other controller.

<u>Scanning</u>: Should also require progressive spatial resolution. This would support interchange with the Internet, Color facsimile, and Printer applications.

5.7 Progressive Quality Resolution

Definition

Ability to extract lower bit-rate images from a codestream without redundant decoding or sacrifice of image quality (at that bit-rate).

Aggregate requirements

At least 8 levels of quality. However, continuous progression has even more flexibility and advantage. Progression for many applications requires lossless to very lossy (low bit-rate). In some applications, moderate compression (near lossless to just noticeable artifacts) is the minimum acceptable. Should be combinable with progressive spatial resolution (see previous section).

Changes to the application profiles

<u>Internet</u>: Progression from lossless to visually lossy.

Color facsimile: Should also require progressive quality resolution. This would enable facsimile systems where the buffering, pixel, or component depth capability of the destination facsimile is determined and the appropriate bits sent. Also, this would allow fixed-time or interrupted transmission of facsimile images. It would also allow facsimile standards to incorporate generational improvements in performance without altering the standard. Lossless to very lossy progression is required.

Printing: Should also require progressive quality resolution. This would enable images to be sent to printers of different sizes, allowing the management of buffers and display lists. Lossless to very lossy progression is required.

Scanning: Should also require progressive quality resolution. This would support interchange with the Internet, Color facsimile, and Printer applications. Lossless to very lossy progression is required.

Digital photography: Moderate lossy to very lossy. Lossless is optional.

Remote sensing Lossless to moderate lossy.

Mobile: Moderate lossy to very lossy.

Medical: Lossless to moderate lossy.

Digital library: Same as the Internet.

E-commerce: Same as the Internet.

5.8 Security

Definition

Three purposes: 1) protect access to the image, 2) identify the image, source or owner in a secure way that cannot be removed or modified by unauthorized parties, 3) indication of the integrity.

Aggregate requirements

To be visually transparent, yet perform the purposes in electronic form and is preserved after image compression and other image processing operations.

Changes to the application profiles

Internet: Read and write.

Color facsimile: Read and write.

Printing: Read only.

Scanning: Write only.

Digital photography: Read and write.

Medical: Read and write.

Digital library: Read and write.

E-commerce: Read and write.

5.9 Error resilience

Definition

To be "robust" (allow complete or acceptable partial decoding) in the presence of errors in the codestream such as random errors, burst errors, and packet or byte loss or insertion errors.

Aggregate requirements

Should have capabilities that make the codestream robust to errors.

- Medical:. Medical imagery does not require more stringent error control within JPEG 2000 than other applications.

? In almost (if not all) medical imaging applications, integrity is maintained at a lower level through the use of reliable, robust, error-correcting media or network communication channels, and to provide additional error robustness within JPEG 2000 at the expense of additional complexity or reduced performance would not be helpful. Neither is it a requirement to "partially" recover some part of an image.

5.10 Complexity Scalability

Definition

Aggregate requirements

JPEG2000 should be scalable in complexity, so that depending on the applications, different levels of complexity can be implemented.

5.11 Strip Processing

Definition

The ability to compress and decompress images with a single sequential pass.

Aggregate requirements

JPEG2000 should support strip processing with 1 or more lines per strip.

5.12 Sensor-Specific Compression Flexibility

Definition

The ability to specify compression algorithm components and/or parameters, such as particular transforms, transform kernels, quantizers or coding strategies, based upon the characteristics of the imaging device and/or data.

Aggregate requirements

JPEG2000 should provide the capability to specify compression algorithm components and/or parameters, such as particular transforms, transform kernels, quantizers or coding strategies, based upon the characteristics of the imaging device and/or data.

5.13 Information embedding

Definition

Efficient embedding of non-image information such as text, voice annotation, web links, and other types of meta-data information into compressed images.

Aggregate requirements

JPEG2000 should allow an efficient embedding of additional information oand ther types of meta-data information into compressed images. For example, text, voice annotation, web links, digitalization tool used; color calibration made; look-up tables to be used; backtrack of all the processing done on the image after it has been digitized and other types of meta-data information into compressed images.

Changes to the application profiles

5.14 Repetitive Encoding/Decoding

Definition

The ability to decode and re-encode iteratively without adding distortion. (Idempotency)

Aggregate requirements

JPEG2000 should be capable of iteratively decoding and re-encoding without adding distortion

Changes to the application profiles

5.15 Object-Based Functionality

Definition

Aggregate requirements

5.16 MPEG4 VTC Compatibility

Definition

Aggregate requirements

JPEG2000 should be able to decode MPEG4 VTC (visual texture coding) bitstream directly.

5.17 Decoder — Backward compatibility/easy transcoding on the client side

Definition

Ability to decode current JPEG baseline images.

Aggregate requirements

5.18 Decoder - ROI Decoding

Definition

Aggregate requirements

The standard should allow the decoder to specify which part of the receiving image (during the stage that the user is receiving an image in progressive transmission schemes) will be decoded at better quality (up to lossless, provided that the image was stored in a lossless mode). The user at the decoder side should be able to decode only specific ROI’s lossless and accept a graceful degradation of the image outside the ROI. Notice that the ROI selected need not be known to the encoder during the encoding.

5.19 Decoder - Fast/Random Data Access

Definition

Aggregate requirements

The JPEG2000 decoder should allow a quick and efficient random access to portions of the compressed image without necessarily decoding the whole image. This should not result in excessive overhead. The speed of the random access should be close to proportional to the size of the image decoded.

5.20 Decoder - Implementation Complexity

Definition

Complexity is related to the handling of three different processes, encoding, transfer and decoding. JPEG2000 has to support the concept of scalable complexity, i.e. terminals or other devices based on different hardware platforms must be able to generate or decode bitstreams according to their computational capabilities. With regard to transfer, in case of server applications the parsing of JPEG2000 bitstreams must be possible for bitrate reduction purposes without the need for transcoding. For battery based applications such as mobile or digital camera, computational complexity and memory costs such as the amount of memory and the memory bandwidth have to be in application typical ranges.

Aggregate requirements

Algorithms should also be considered based on memory usage and accesses.

Note: Need to provide better wording…

5.21 Encoder - ROI Encoding

Definition

Aggregate requirements

The JPEG2000 decoder should allow the user at the encoder side to specify which part (of arbitrary shape) of the image (Region of Interest - ROI) should be coded at better quality, up to lossless, and which lossy. It should allow the user to trade off compression performance with image quality by providing the ability to code an ROI to a higher degree of quality (lossless or lossy).

5.22 Encoder - Fast Encoder

Definition

Aggregate requirements

JPEG 2000 should provide encoding options which support a variety of application processing capabilities, including high-speed applications where low complexity may be critical, and non-real time applications which may accommodate more complex encoding strategies.

5.23 Encoder - Implementation Complexity

Definition

Complexity is related to the handling of three different processes, encoding, transfer and decoding. JPEG2000 has to support the concept of scalable complexity, i.e. terminals or other devices based on different hardware platforms must be able to generate or decode bitstreams according to their computational capabilities. With regard to transfer, in case of server applications the parsing of JPEG2000 bitstreams must be possible for bitrate reduction purposes without the need for transcoding. For battery based applications such as mobile or digital camera, computational complexity and memory costs such as the amount of memory and the memory bandwidth have to be in application typical ranges.

Aggregate requirements

JPEG2000 should have encoder complexity (e.g., memory, speed, small-tile access, power consumption, cost, etc.) no more than 4 times that of the current “Baseline JPEG” under similar implementations and conditions.

5.24 Decoder - Geometric manipulation of compressed images

Definition

Efficient geometric manipulation of JPEG2000 compressed images, limited to rotations of 90, 180, 270 degrees and horizontal and vertical flipping, and any combinations.

Aggregate requirements

Any number of consecutive geometric manipulations should not result in a significant change in the size of the compressed file, nor in any degradation in image quality. The JPEG2000 standard should allow geometric manipulations to support low complexity and low memory applications..

6 Application Profiles

This section contains a table that indicates the mandatory and optional requirements associated with the individual application profiles.

Note: a core profile should be defined. In addition, some of the following profiles should be merged in order to define other more generic profiles (main, simple, etc.) profiles for clusters of applications.

		Overall System																Decoder Specific					Encoder Specific		
		5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	5.10	5.11	5.12	5.13	5.14	5.15	5.16	5.17	5.18	5.19	5.20	5.24	5.21	5.22	5.23
Application Profile		Image Type	Uncompressed	Lossless Compression	Visually Lossless Compression	Visually Lossy Compression	Progressive Spatial	Progressive Quality	Security	Error Resilience	Complexity Scaleability	Strip Processing	Sensor Specific Compression Flexibility	Information Embedding	Repetitive Encoding/Decoding	Object-Based Functionality	MPEG4 VTC Compatability	Backward Compatability Client Side Ease of Transcoding	ROI Decoding	Fast/Random Data Access	Implementation Complexity	Geometric Manipulation	ROI Encoding	Fast Encoder	Implementation Complexity
6.1	Internet	M(1,3) O(4+)			M	M	M	M	O	O				M		M	O	M(Baseline) O(non-baseline)	O	M		~~O~~	O		
6.2	Color Facsimile	M		O	~~M~~O					O		M			M			O		O	M	~~O~~M			M
6.3	Printing	M			M	O												O		O	M	M			M
6.4	Scanning (Consumer, pre-press)	M	O	M	M				~~M~~O											O		~~M~~			M
6.5	Digital Photography	M		O	M	O	O	O	O	O				M	O	O	O	O			O	O		M	M
6.6	Remote Sensing	M(1,3) O(4+)		M	M	M	M	O		O		M	M	O		O		O	O	M		O	M	M	
6.7	Mobile	M(1,3) O(4+)		O	M	M	O	M	O	M	O					O	O	O	O	M	M		O	O	M
6.8	Medical	M(1,3) O(4+)		M	M	M	M	M	~~O~~M	O		M	M	M	O	O		M(Baseline) O(non-baseline)	O	M		~~O~~	M	O	
6.9	Digital Library	M(1,3) O(4+)		O	M	M	M	M	O	O				M		M		M(Baseline) O(non-baseline)	O	M		~~M~~	O		
6.10	E-Commerce	M(1,3) O(4+)		O	M	M	M	M	M	O				M		M		M(Baseline) O(non-baseline)	O	M			O		
Legend: M - Mandatory, O - Optional																									

Appendix A - Terms of references

This appendix contains definitions of different terms used in this document.

Errors

Understood as residual errors in the bitstream input to the decoder after error correction at the system layer.

Robust

Understood as the decoder's ability to continue to decode the image in the presence of errors with a graceful degradation of image quality avoiding a catastrophic breakdown (failure) as the channel conditions deteriorate.

Error types

In environments where JPEG2000 is likely to be used, the transmission noise may be from one of the following 3 major classes:

1. Random bit errors (Binary symmetric channel)
2. Burst errors (Mobile and wireless fading channels)
3. Packet loss (byte deletion/insertion in IP/ATM communication)

The transmission will in some way provide error correction/detection at system layer. When this fails, the JPEG2000 decoder will have to face a bitstream with errors. These again will belong to one of the classes mentioned above.

The following is an attempt to better clarify the term error used in various applications considered in this document

Internet: Packet loss in Internet transmission not having guaranteed packet delivery

Hardcopy Color Facsimile: random errors

Digital photography: Considering hand held devices with transmission equipment this will be very similar to mobile. Considering storage on a device in the camera, random errors at a very small probability.

Remote sensing: Random errors, burst errors (Considering satellite systems and radio based surveillance equipment)

Mobile: Random errors, burst errors, packet loss/insertion (IP)

Medical: None, need the ability for error detection

Region of Interest (ROI)

A *Region Of Interest* (ROI) is defined as an object in the bitstream that the user might access and manipulate (cut, paste, etc.). It can be made of Y,U, and V component plus shape information. The encoder might transmit together with each ROI composition information to indicate where each ROI shall be displayed. At the decoder the user can change the composition of the scene displayed by interacting on the composition information.

The ROI's can have an arbitrary shape and can be of any type (synthetic texture, text, graphics, images,…). The ROI's can be known by construction of the images (for example synthetic images) or can be defined by semi-automatic segmentation. The background is also defined as a ROI. ROI's can be inside one other ROI. Each ROI might have some data, i.e. semantic information linked to it.

Scalability

Realizing that many applications require images to be simultaneously available for decoding at a variety of resolutions or qualities this architecture shall support scalability. In general, scalable coding of still images means the ability to achieve coding of more than one resolution and/or quality simultaneously. Scalable image coding involves generating a coded representation (bitstream) in a manner which facilitates the derivation of images of more than one resolution and/or quality by scalable decoding. Bitstream Scalability is the property of a bitstream that allows decoding of appropriate subsets of a bitstream to generate complete pictures of resolution and/or quality commensurate with the proportion of the bitstream decoded. If a bitstream is truly scalable, decoders of different complexities, from low performance decoders to high performance decoders can coexist, and while low performance decoders may decode only small portions of the bitstream producing basic quality, high performance decoders may decode much more and produce significantly higher quality. The most important types of scalability are *SNR scalability* and *spatial scalability*.

SNR scalability is intended for use in systems with the primary common feature that a minimum of two layers of image quality are necessary. *SNR scalability* involves generating at least two image layers of same spatial resolution but different qualities from a single image source. The lower layer is coded by itself to provide the basic image quality and the enhancement layers are coded to enhance the lower layer. The enhancement layer when added back to the lower layer regenerates a higher quality reproduction of the input image.

Spatial scalability *is intended for use in systems with the primary common feature that a minimum of two layers of spatial resolution are necessary. Spatial scalability involves generating at least two spatial resolution layers from a single source such that the lower layer is coded by itself to provide the basic spatial resolution and the enhancement layer employs the spatially interpolated lower layer and carries the full spatial resolution of the input image source.*

An additional advantage of spatial and SNR scalability types is their ability to provide resilience to transmission errors as the more important data of the lower layer can be sent over channel with better error performance, while the less critical enhancement layer data can be sent over a channel with poor error performance.

Both types of scalability are very important for internet and database access applications, and bandwidth scaling for robust delivery. The *SNR* and *Spatial* scalability types include the progressive and hierarchical coding modes already defined in the current JPEG but they are more general. In addition, the following types of scalability are useful:

Complexity scalability in the decoder allows a given texture or image bitstream to be decoded by decoders of different levels of complexity. The reconstruction quality, in general is related to the complexity of the decoder used. This may involve that less powerful decoders decode only a part of the bitstream. This relates to the spatial scalability, where a decoder might decode only the low resolution image.

Quality scalability allows a bitstream to be parsed into a number of bitstream layers of different bitrate such that the combination of a subset of the layers can still be decoded into a meaningful signal. The bitstream parsing can occur either during transmission or in the decoder. An application of this could be in SNR scalability with two layers, where the decoder selects some ROI's from the base layer and some ROI's from the enhancement layer and composites an image which has a mixed quality.

Notice that scalability applies both to texture as well as in shape.

Composition information

The composition information carried in the composition layer carries information related to the composition rule for each ROI (i.e. how the regions will be displayed, if they will overlay, etc.). One important application is cases where we have text on top of a background image. In that case it should be possible to have the text described as a separate region. The region description header of that regions should specify the text description language (e.g. postscript) and the composition rule (e.g. superposition).

A scene/image can be composed from objects of various types. Composition information is transmitted and at the received site, the user can interact with the compositor and manipulate the objects.

User Interaction

User interaction can have two forms: client-side interaction and server-side interaction.

Client-side interaction involves content manipulation which is handled locally at the end-user's terminal, and can take the following forms:

1. Modification of an attribute of a ROI. This includes changing the position of a ROI, making it visible or invisible, changing the font size of a synthetic text node, etc.

2. Hyperlinking. This involves switching to a new content source after the user selects (e.g., with a mouse click) a particular ROI (or portion of it).

The first form can be easily implemented by applications using JPEG 2000 by providing internal (implementation dependent) mechanisms that link user events (e.g., mouse clicks or keyboard commands) to scene update commands. These scene update commands can be processed by the terminal in exactly the same way as if they originated from the original content source. As a result, this type of interaction does not require standardization.

Hyperlinking requires the presence of location information and the necessary support to access it. This information can be carried by the semantic information data, but it is up to the application to implement the necessary functionality (i.e., switch to the new source when the anchor is selected).

Server-side interaction involves content manipulation that occurs at the transmitting end, initiated by the receiver. For example, the receivers terminal might send information to the transmitter on which ROI's to selectively refine, or to whether he would like to see the image or some ROI's/tiles in higher resolution, etc.

Architecture

A description of a unifying or coherent form or structure. A compression architecture describes the structure to support compression and expansion. An architecture will define a set of tools and their possible associations to allow step-by-step procedures necessary to complete compression or expansion

Backward Compatibility - DecoderA JPEG2000 decoder is backward compatible if it is able to interpret, and expand the compressed imagery information (codestream) produced by a JPEG compliant encoder.

Backward Compatibility - Encoder

A JPEG2000 encoder is backward compatible if it produces a codestream which can be parsed, translated or transcoded by a JPEG compliant decoder and interpreted to the extent that catastrophic decoder failure is avoided. The decoder may skip the expansion process, or expand an image and either flag or not flag its lack of meaning.

Bitstream

A bitstream is the actual sequence of binary digits resulting from the coding of a sequence of symbols (compressed data stream). As such it does not include bits containing information regarding the coder type, its parameters, or other (overhead) information required to decode the sequence.

Codestream

A codestream is a collection of one or more bitstreams and associated (overhead) information required for their decoding and expansion into image data. Such overhead information is restricted to that required for expansion and may include, but is not limited to, markers indicating locations of particular bitstreams, segments indicating transform, quantization and coding types, etc.

File Format

A file format consists of a codestream and additional support data and information not explicitly required for the expansion of image data. Examples of such support data are text fields providing titling, security and historical information, markers to support placement of multiple codestreams within a given data file, and markers to support exchange between platforms or conversion to other file formats.

Compound image

A compound image is a composed of parts generated synthetically and parts coming from natural images. A componf image may degenerate to a fully natural or a fully synthetic image.

?