TECHNICAL REPORT

ISO/TR 12033

First edition
2009-12-01

Document management — Electronic imaging — Guidance for the selection of document image compression methods

Gestion de documents — Imagerie électronique — Directives pour le choix des méthodes de compression d'image

Reference number
ISO/TR 12033:2009(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

In exceptional circumstances, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example), it may decide by a simple majority vote of its participating members to publish a Technical Report. A Technical Report is entirely informative in nature and does not have to be reviewed until the data it provides are considered to be no longer valid or useful.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO/TR 12033 was prepared by Technical Committee ISO/TC 171, *Document management applications*, Subcommittee SC 2, *Application issues*.

This first edition of ISO/TR 12033 cancels and replaces ISO/TS 12033:2001, which has been technically revised.

Introduction

With respect to the rapid increase of applications using digitization techniques, the role of compression methods has become a factor of growing importance for the management of the volumes of stored data.

The effects of the available compression methods vary greatly, depending on the source documents. For example, an electronic image management (EIM) system configured for scanning and storing continuous tone images will have different image compression requirements as compared to an application involving only text.

Practical methods for analysing user requirements for image compression in order to select accurate and optimal image compression schemes are complex. This Technical Report was issued in order to guide users and system developers in their selection of these methods.

Document management — Electronic imaging — Guidance for the selection of document image compression methods

1 Scope

This Technical Report gives information to enable a user or electronic image management (EIM) integrator to make an informed decision on selecting compression methods for digital images of business documents. It provides technical guidance to analyse the type of documents and which compression methods are most suitable for particular documents in order to optimize their storage and use.

For the user, this Technical Report provides information on image compression methods incorporated in hardware or software in order to help the user during the selection of equipment in which the methods are embedded.

For the equipment or software designer, this Technical Report provides planning information.

This Technical Report is applicable only to still images in bit map mode. It only takes into account compression algorithms based on well-tested mathematical work.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 12651:1999, *Electronic imaging — Vocabulary*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 12651 and the following apply.

3.1
compression
process of removing redundancies in digital data to reduce the amount that should be stored or transmitted

NOTE Lossless compression removes only enough redundancy so that the original data can be recreated exactly as it was. Lossy compression sacrifices additional data to achieve greater compression. This is typically useful for greyscale or colour image compression, where details that are not perceptible, or are minimally perceptible, to the human eye can be eliminated, normally with a dramatic increase in compression. It is advisable that lossy compression not be used for documents containing textual information and not be used for long term archival of any type of documents.

3.2
resolution
number of pixels per unit of length

3.3
dots per inch
dpi
number of dots that a scanner (printer) can scan (print) per inch both horizontally or vertically

3.4
brightness
visual sensation that enables an observer to detect luminance

3.5
contrast
ratio of on pixel brightness to off pixel brightness

3.6
bit level
number of bits used to define a pixel

3.7
luminance
Y
luminous flux emitted from a surface

NOTE The former term was photometric brightness.

3.8
chrominance
Cr
Cb
colour portion of the video signal including hue and saturation but not brightness

NOTE Low chroma means the colour picture looks pale or washed out; high chroma means intense colour; black, grey and white have a chrominance equal to zero.

3.9
ITU-T Group 3 and Group 4
compression algorithms standards defined by the ITU-T in Recommendations T.4 and T.6

3.10
Joint Photographic Experts Group
JPEG
name of the committee that developed the ISO/IEC 10918 series which shares the same popular name

NOTE The "J" refers to the joint development with the ITU-T.

3.11
Comité Consultatif International Télégraphique et Téléphonique
former name of the International Telecommunication Union (ITU) standardization body

3.12
compression ratio
relationship of the total bits used to represent the original to the total number of encoded bits

3.13
Joint Bi-level Image Experts Group
JBIG
name of the sub committee that developed ISO/IEC 11544

NOTE The joint committee is with ITU-T. JBIG and JPEG are managed by ISO/IEC JTC1/SC 29/Working Group 1.

4 General

In a document imaging system, users are concerned about the quality of archived images, for two reasons:

a) it can affect the imaging system's future in the medium or even long-term;

b) it is necessary to choose the imaging tools based on an evolving technology.

The digitization process, which by nature transforms an image conveying comprehensible information into a dematerialized one, changes the observer's perception of that image. The observer may consider the image as being improved, though more frequently he considers it degraded. In fact, images undergo a number of successive transformations at different points during the digitization process. At each of these stages, attempts are made to keep the image within acceptable legibility limits, but also to restrict its size to within acceptable economic limits.

The specific role of one of the digitization stages — compression — is to reduce the size of the image. Some compression methods are reversible in that the decompression algorithm restores the initial digital information. These methods are lossless and have no impact on the quality of the image as it is perceived by the human eye. Other methods are lossy, and may cause degradation perceptible to the eye. By adjusting certain parameters, the user can bring a lossy method within acceptable limits; because the acceptance of a lossy method is a subjective judgement. Any image or document, on which a computerized treatment may be applied, should not be compressed with such a method. This is one of the major reasons not to use lossy compression for long-term archiving, as future usage of the image or document is unknown.

While numerous compression methods are described in technical literature, few are stable according to industrial standards. These are based on a limited number of principles:

— dominance of certain patterns,

— pattern repetition, and

— noticeable mathematical properties.

In any individual method, the number of parameters the user can modify is small.

The choice of a method and compression parameters are in large part determined by two considerations:

a) the characteristics of the document;

b) the period of time the document is to be retained (retention time).

Obviously, the graphical contents of a document play a key role in determining the method and its parameters. However, other factors characterizing the application context are also very important (see Table 1).

The graphical content of the document is important to the compression process. A business document that can be copied or faxed as “pure black and pure white” (even if the original was blue ink on yellow paper) are probably best compressed with the technologies developed by the ITU-T for a facsimile. Colour or grey scale photos are probably best compressed using one of the JPEG technologies. But if the photo has been converted to variable size black dots (like many “half-tone” newspaper photos), then JBIG is a superior compression technology.

5 Type of document and digitization parameters

5.1 General

A document is a set of organized information intended for presentation to a human user. Documents can be a single page or a set of pages, and can contain arbitrary content types, such as character content, graphical content, and various types of image content.

The following document content may be found in various types of documents. The classification list hereafter is somewhat arbitrary, but for a given application, these distinctions may be used to understand how to handle a given document.

5.2 Type of documents

This clause focuses on only those documents that are most likely to be archived electronically. These documents include:

— black text on white background, or more technically, dark text on light background (even if the ink happens to be blue or red or other single colour, on whatever colour paper);

— photographs, i.e. black and white or colour;

— mixed documents containing both text and photographs reproduced by a printing process, i.e. black and white or colour.

5.3 Document classification and digitization

5.3.1 General

For the purpose of determining a compression scheme, documents may be described in the following five ways. For each type of document, digitization methods are briefly described.

5.3.2 Black and white documents

Digitizing pages printed in black and white or more generally in bi-tonal mode (primarily text with a unique foreground on a unique background) generates bi-level images where each pixel is represented by a bit.

The most important digitization parameter is resolution.

Resolution should be determined according to visual perception needs and on the limits of the complete imaging process. Human eyes will not see noticeable differences on documents digitized at more than 300 dpi. This is the most commonly used resolution to keep quality unaltered. Any resolution under 300 dpi will have visible effects on the digitized document. A resolution over 300 dpi may be needed when computerized treatment is done on the document. Also, 300 dpi is the resolution limit of the human eye and should be considered as the needed resolution at the visual size, i.e. if the zooming factor to visualize is 4, a resolution of 1 200 dpi on the original size will provide 300 dpi on the visual size.

There are also other parameters, related to image processing, which vary according to the kind of image. If, for example, the images to be digitized are text, then it is advisable to produce black characters that are sharply defined against a white background. The brightness (adjusting the colour of a pixel against a threshold) and contrast parameters (adjusting the colour of a pixel against that of the surrounding pixels) should be adjusted for this purpose.

5.3.3 Grey scale documents

This form of representation is applied to photographic documents, printed on paper from a black and white film.

Digitization changes an initially continuous document into a matrix of pixels whose intensity is encoded in a range of levels. Thus, 8-bit encoding produces 256 grey scales.

The number of grey scales or the bit level should be determined according to visual perception needs and the limits of the complete imaging process. Quality tests have demonstrated that human eyes will not see noticeable differences on grey scale images coded with more than 8 bits. Therefore 8 bits encoding is the most common value.

5.3.4 Pseudo-grey or halftone documents

This category includes images that simulate grey using a variable arrangement of black and white pixels. There are two possibilities:

— the source document is a photographic reproduction in a text; it was produced using a printing technique and is, itself, a pseudo-grey document (a raster image can simulate grey by a pattern of black and white pixels);

— the source document is a photographic original, but was digitized in pseudo-grey for performance reasons, for example to reduce the storage volume or transmission times on a network.

5.3.5 Colour documents

This form of representation is applied to photographic documents, printed on paper from a colour photographic original. Another application is digital colour capture of business documents where yellow highlights, colour boxes, pencil, red pen, etc., is a part of the information capture integrity.

Colour documents are intended to be restored in colour, but may also be reproduced in grey scale.

Colour representation is based on the neuro-physiological properties of the human eye, notably the “visual trivariance” principle, which states that all colours can be produced by combining the three primary colours. Thus, a colour can be represented by three coordinates in a vector space based on primary colours, or by linear combinations of these coordinates.

The colour space most frequently adopted for electronic displays uses an additive of red, green and blue colours. These colours are differentiated by the retinal cones in the eye. Another colour space decouples the variables into one “luminance” variable, and two “chrominance” variables. This colour space is used to transmit television signal.

The most frequently used colour space in printing is of cyan, magenta, and yellow colours. A printed digital image emits light indirectly by reflecting light that falls upon it. For example, a page printed in yellow absorbs (hence this is called a subtractive colour space) the blue component of white light and reflects the remaining red and green components, thereby creating a similar effect as a monitor emitting red and green light. Hence the printing industry mixes cyan, magenta, and yellow inks to create all other colours. Combining these subtractive primary colours will generate black, but in practice black ink is used, hence the term “CMYK” colour space, the last character “K” standing for black.

In a digitized colour image, each pixel is represented by assembling three components corresponding to the primary colours. The bit level adopted for a component determines the quality of hues; the standard of 8 bits per component can represent $256^3 = 16$ million different colours. Representations on a total of 8 bits sent by data communication networks are also fairly frequent.

5.3.6 Mixed documents

Many documents to be archived as pure raster/bitmap images are composed of pages of text containing graphic elements and/or photographic images. There is no completely satisfactory way of representing this type of document:

— a bi-level representation would make illustrations illegible;

— a grey scale or colour representation to preserve illustrations would provide the best quality, but would make storage volumes disproportionately large with respect to the importance of the illustrations (note that there are possible trade-offs between resolution and bit level of grey scale or colour image files).

In many mixed documents, text is considered more important, so a bi-level representation would be used to draw black characters on a white background. The photos would either be lost, or would have to be separated from the text for appropriate representation. In most cases, text and photos can be automatically and successfully separated using segmentation algorithms. Sometimes, segmentation can lead to loss of information (such as captions under photos, or unusual typographic arrangements).

6 Compression methods and standards

6.1 LZW compression (Lempel Ziv Welch)

This method was a patented Unisys method until June 20, 2003 when the patent expired. It is commonly used to compress black and white images, and is part of the GIF implementation. ISO 19005 (PDF/A) forbids the use of this algorithm for long term preservation of data.

LZ77 and FLATE compressions are derived from or preliminary to LZW. As such, they are both contained in the above description.

6.2 RLE compression (run-length encoding)

This method takes into account runs of identical symbols inside data streams (such as characters in an ASCII text). Each data stream is encoded with the number of occurrences of the repetitive elements and the length of the stream.

An RLE algorithm can operate at the bit, byte or pixel level. The basic algorithm works one line at a time, but some variations can also work vertically, taking into account repeating characters in adjacent lines. The RLE method is normally lossless, although to improve efficiency, some variations drop lower-order bits, resulting in loss.

This method is not very efficient for texts and complex photos, because there are few long sequences. It is most efficient for images with large areas of uniform colour.

6.3 ITU-T algorithms

6.3.1 General

ITU-T has defined a series of protocols for transmitting images via facsimile. These protocols are officially named T.4 and T.6, but are popularly known as the Group 3 and Group 4 methods. The compression methods used in archiving are variations of ITU-T. ITU-T may contain end-of-line and end-of-message codes to simplify fax transmissions. These codes are superfluous when these methods are used for archiving.

ITU-T compression types are based on run length encoding using variations of the Huffman algorithm.

ITU-T defines three fax standards, which are used for compressing bi-level images:

— Group 3 modified Huffman (MH) — a one-dimensional compression method (G3 1D);

— Group 3 modified Read (MR) — a two-dimensional compression method (G3 2D);

— Group 4 modified MR (MMR) — a two-dimensional compression method (G4).

6.3.2 Group 3 one-dimensional method (G3 1D)

The Group 3 one-dimensional method (G3 1D) is a variation of the Huffman algorithm. In a bi-level image, each scanned line alternates variable-length zones, composed of black or white pixels. The Group 3 encoder determines the length of each black or white zone, called the run length, and looks up the corresponding code words in the Huffmann table.

Compression occurs because the code words are shorter than the zones they represent. Each code word represents a zone length corresponding to either white or black.

Group 3 is the basic compression algorithm used in Group 3 fax transmission.

The length of the code words was determined when the method was created, based on static observations of typed and hand-written documents. Run lengths with a high probability of occurrence were assigned the shortest code words.

NOTE Although ITU-T compression was initially designed for text documents, it can also be applied to raster photos, although it is less efficient.

Sequences of pixels are represented by two types of code words:

a) configuration code words;

b) termination code words.

Configuration code words represent long zones and termination code words represent short zones. A zone with a length of between 0 and 63 bits is encoded in a termination code word. A stream of between 64 and 2 623 bits is encoded in a configuration code word corresponding to the quotient of the length divided by 64; a termination code word can be added for the remainder. A stream with a length of over 2 623 bits is encoded as a series of configuration code words to which a termination code word can be added.

This one-dimensional encoding scheme eliminates redundancy only within each scan line, left to right. It does not reduce the redundancy between scan lines, up and down.

6.3.3 Group 3 two-dimensional method (G3 2D) and Group 4 method

Where the Group 3 one-dimensional method deals with each scan line of an image individually, the Group 3 two-dimensional method takes advantage of frequently occurring similarities between two successive lines in the same image.

G3 2D is defined as an option of Group 3, which restricts itself to a small number of lines inserted between "one-dimensional" lines. Group 4 uses the same algorithm.

Like G3 1D, the G3 2D algorithm uses breakpoints that separate different colours in a single line ("mutant elements"). In creating an encoded representation of the image, the algorithm takes into account the mutant elements not only in a single line, but also in two adjacent lines. Thus, in addition to the code words used in G3 1D, the G3 2D and G4 methods use code words representing the distance and relative arrangement of mutant elements in two or more adjacent lines.

6.4 JBIG compression

JBIG is the abbreviation of Joint Bi-level Image Group. As its name indicates, this method is used for bi-level images. It is used primarily for text (like T.4 and T.6), though it can also be used for raster photos in printed documents (unlike T.4 and T.6). According to its authors, JBIG is as efficient as T.4 and T.6 for pure text, and 2 to 30 times more efficient for raster photos. Like T.4 and T.6, JBIG is lossless.

The method uses progressive encoding, which manipulates resolution. This encoding system initially transmits images in low resolution (e.g. 25 dpi). Then the resolution is progressively doubled until the resolution of the original image is obtained. There are two advantages to this progressive method:

a) it analyses images with just the necessary degree of detail;

b) it can adapt the resolution of an image according to the characteristics of the output peripheral or the perception needs of the human observer (e.g. transmission may be interrupted as soon as the image is recognized).

ISO/IEC 11544 describes the JBIG method as an assembly of procedure blocks. Most of these blocks have the dual purpose of saving encoding and accelerating processing. Encoding is saved when regions of uniform colour, repetitiveness, and similarities between the low and high-resolution image are detected.

6.5 JBIG2 compression

Designed to overcome the poor compression power of the lossless JBIG method, JBIG2 aims at getting the best lossless compression ratio for bi-tonal documents. And when the final image size is critical, JBIG2 will allow better lossy compression ratio keeping a better quality than JBIG.

However it keeps functionalities such as progressive encoding with a choice of progression based on quality or on content. This allows use of different compression ratios on different parts of a document, following the requirements of quality and/or content of those parts.

JBIG2 is a very efficient method of compression for mixed documents, composed of blocks of texts and pseudo-grey or half-tone pictures.

The JBIG2 standard contains a segmentation method for a page of documents to allow this adaptative coding scheme. It separates text regions, half-tone regions and so-called generic regions (which include line-art pictures, noisy area, graphical drawings, …)

Being an ITU-T approved standard, with included fonts and metadata, it is a very efficient way to encode bi-tonal documents, with a compressed size close to an ASCII coded text document.

6.6 Discrete cosine transform (DCT)

In principle, DCT images are highly correlated from one pixel to the next. That is, if one pixel is a certain shade of grey, odds are that its neighbours are approximately that same shade. This means that some of the redundant information can be lost without a serious impact on the image's legibility. JPEG standards use this method.

The DCT is applied to each element of each block (a block is an 8×8 pixel extracted from the digitized image). DCT transforms the block into 64 frequency coefficients.

This transform has the interesting property of concentrating information into a small number of coefficients, with most other coefficients having a value close to zero.

6.7 Fractal compression

This method is based on fractal geometry, attributed to the French mathematician Benoît Mandelbrot. Unlike analytical geometry which approaches complex curves using infinitesimals composed of line intervals, fractal geometry considers that the division of a curve (or a surface) into smaller elements does not reduce its complexity. A special category of fractals is created by the infinite repetition of the same pattern inside elements subdivided to infinity in a given curve.

Image compression using the fractal method is the opposite of generating fractal images. Instead of creating an image from a given transform, the purpose of compression is to find a series of transformations with which to approach a given digital image.

By analogy with fractal geometry, the basic purpose is to find similarities between the image elements obtained at different levels of subdivision, i.e. if all the small elements can be considered similar to large elements, they do not have to be described in detail, and compression is achieved.

In the fractal method, the algorithm seeks to compare image elements using matrix transforms. The algorithm includes an image subdivision principle and a metric definition which allows it to determine how similar the elements in an image are.

There are no current standards using fractal compression.

6.8 Wavelet compression

This method is based on research in mathematics expanding on Fourier's work in the 19th century.

A Fourier series can represent any periodic function as a linear combination of sinusoidal functions with variable frequencies. Modifications to the Fourier transform are needed to adapt to the representation of discrete valued functions such as those generated by digitization, and to deal with non-periodic functions. This leads to the DCT, the basis of JPEG. However, a Fourier transform is ill-suited for pronounced discontinuities in an image.

Wavelets form families of mathematical non-sinusoidal functions. These functions can have peaks and be rapidly decreasing, which means that they can be linked to non-periodic local characteristics of the image.

As in the DCT method, wavelet compression involves replacing the values generated by digitization with linear combination coefficients. The compression effect is achieved through the presence of near-zero coefficients.

Wavelet compression is used in the JPEG 2000 standard.

6.9 JPEG compression

6.9.1 General

JPEG covers a family of algorithms and two types of compression. The compression commonly referred to as JPEG is lossy compression of which JPEG-LS is the lossless variant. Mutual decoding is not possible.

The lossless type is not very efficient. Lossy compression is the more frequently used type. It contains several steps, that can be invoked individually, some of which are the source of loss. The heart of this system is a mathematical transform known as the discrete cosine transform (DCT).

6.9.2 JPEG steps

The JPEG method consists of several steps:

- apply the discrete cosine transform;
- perform a truncation induced by the digital representation of the data (very low-order values are rounded to zero);

— apply the "quantization" transform;

— perform "post-entropic encoding", which superimposes the Huffman coding technique or arithmetic compression to finalize the JPEG compression.

Except for DCT, the most important step in JPEG is quantization.

Quantization involves applying a linear transform to the coefficients obtained by the DCT. This quantization is used to eliminate frequencies with small contribution (mostly high frequencies) as well as to represent coarsely remaining frequencies.

A quantization matrix controls both the compression ratio obtained and image degradation. It is possible to correct for a "level of loss".

The JPEG experts did not define the quantization matrix strictly in the standard. They simply gave an example adapted to a 720×576-pixel television screen. This example quantization matrix is based on psycho-visual testing with photographs.

6.9.3 Components of JPEG

JPEG primarily describes the principles of compression. Some, like DCT, are specific to JPEG, while others, such as the Huffman coding technique, are not.

Although JPEG does not impose a colour space, it can take advantage of the eye's most acute sensitivity to chromatic and luminance variations. Thus, in a Y, Cr, Cb encoding, it can sub-sample chrominance information with respect to luminance.

JPEG also introduces aspects related to image display dynamics, which are not directly linked to a compression principle. These subsidiary aspects affect the visual perception of the digitized image making them key criteria for JPEG.

— Interlacing: JPEG can transmit Y, Cr and Cb components in three sets, or interlaced.

— Progressive DCT encoding: the image is transmitted several times, with quality improved with each transmission.

The second principle can be used in interactive search systems; most images, which are not considered pertinent, are transmitted with minimal quality, while only those specifically sought are sent with maximum quality.

6.10 JPEG 2000

JPEG 2000 is an image compression system standard based on wavelet technology. JPEG 2000 offers a basic compression mode which is shared for both image and video compression together with extensions which target special purpose applications. For a given image quality, JPEG 2000 files will be 30 % smaller than JPEG files. It offers several file formats:

— two for image data,

— one for video, and

— one for compound data types.

It also standardizes compliance and client-server interoperability. Future work includes security, wireless, 3D modelling, scientific floating point compression, and other extensions for specific market verticals.

The central advantage of JPEG 2000 wavelet compression is the scalability of the wavelet transform selected. This allows decompression of selected spatial areas, resolutions, and accuracies of a larger image or video stream. This means that for a 128×128 thumbnail of a 2 GB image, it will be necessary only to decompress 128×128 of data rather than the 2 GB image followed by downsampling. This scalability extends to colour space specification and other aspects of the file format.

Another advantage of JPEG 2000 is a combined lossy lossless capability. For instance, JPEG 2000 compression could be lossless. But the selected subset chosen for decompression could result in a lossy representation of the image that was compressed.

The central disadvantage of JPEG 2000 is an increase in computational complexity. For a given file size, JPEG 2000 compression and decompression will be slower than JPEG.

But the scalability advantage can overtake the computational complexity disadvantage where large image sizes are desirable or required when the image is compressed, but where smaller subsets or thumbnails are typically needed after decompression. In this example, the decompression complexity of JPEG 2000 is a function of the smaller subset or thumbnail while the decompression complexity of JPEG is a function of the compressed image file size.

6.10.1 Progressive transmission

Progressive transmission allows images to be reconstructed with different resolutions and pixel accuracy, as needed or desired, for different target devices. The image architecture of JPEG 2000 provides efficient delivery of image data in internet/extranet and/or client/server applications.

6.10.2 Transmission in noisy environments

Wireless and internet network are prone to random bit errors and loss of data during transmission. JPEG 2000 is built with error resilience tools and robustness to bit-errors, which allows for error detections and concealment, guaranteeing more reliable, better image transmission in noisy environments.

6.10.3 Region-of-interest

JPEG 2000 includes a regions-of-interest (ROI) coding feature so that certain areas of an image can be coded with better quality than the rest of the image. The ROI area is placed at the beginning of the file so that it will be decompressed before the rest of the image, allowing faster access to this ROI.

6.10.4 Inclusion of metadata and ICC profiles

The JPEG 2000 format has a mechanism by which metadata or ICC profile (International Colour Consortium profile format defined in ISO 15076-1:2005) can be directly included in the image file.

6.10.5 Compound documents

JPEG 2000 contains a mechanism to efficiently compress compound documents (e.g. a mix of colour images and bi-tonal text as in newspapers).

6.10.6 Lossless and lossy compression

JPEG 2000 provides lossless and lossy compression in a single stream of code.