INTERNATIONAL STANDARD

ISO/IEC 15444-9

First edition
2005-12-01

Information technology — JPEG 2000 image coding system: Interactivity tools, APIs and protocols

Technologies de l'information — Système de codage d'image JPEG 2000: Outils d'interactivité, API et protocoles

Reference number
ISO/IEC 15444-9:2005(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

FIGURES

Page

TABLES

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

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

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

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

ISO/IEC 15444-9 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information,* in collaboration with ITU-T. The identical text is published as ITU-T Rec. T.808.

ISO/IEC 15444 consists of the following parts, under the general title *Information technology — JPEG 2000 image coding system*:

— *Part 1: Core coding system*

— *Part 2: Extensions*

— *Part 3: Motion JPEG 2000*

— *Part 4: Conformance testing*

— *Part 5: Reference software*

— *Part 6: Compound image file format*

— *Part 8: Secure JPEG 2000*

— *Part 9: Interactivity tools, APIs and protocols*

— *Part 11: Wireless JPEG 2000*

— *Part 12: ISO base media file format*

The following parts are under preparation:

— *Part 10: Extensions for three-dimensional data and floating point data*

— *Part 13: An entry level JPEG 2000 encoder*

Introduction

ITU-T Rec. T.800 | ISO/IEC 15444-1 (JPEG 2000) is a specification that describes an image compression system that allows great flexibility, not only for the compression of images but also for access into the codestream. The codestream provides a number of mechanisms for locating and extracting portions of the compressed image data for the purpose of retransmission, storage, display, or editing. This access allows storage and retrieval of compressed image data appropriate for a given application without decoding.

The purpose of this Recommendation | International Standard is to provide a network protocol that allows for the interactive and progressive transmission of JPEG 2000 coded data and files from a server to a client. This protocol allows a client to request only the portions of an image (by region, quality or resolution level) that are applicable to the client's needs. The protocol also allows the client to access metadata or other content from the file.

Any organization contemplating the use of this Recommendation | International Standard should carefully consider its applicability.

The International Telecommunication Union (ITU), the International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) draw attention to the fact that it is claimed that compliance with this Recommendation | International Standard may involve the use of a patent.

The ITU, ISO and IEC take no position concerning the evidence, validity and scope of this patent right.

The holder of this patent right has assured the ITU, ISO and IEC that he is willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statement of the holder of this patent right is registered with the ITU, ISO and IEC. Information may be obtained from the companies listed in Annex M.

Attention is drawn to the possibility that some of the elements of this Recommendation | International Standard may be the subject of patent rights other than those identified in Annex M. ITU, ISO and IEC shall not be held responsible for identifying any or all such patent rights.

INTERNATIONAL STANDARD
ITU-T RECOMMENDATION

Information technology – JPEG 2000 image coding system: Interactivity tools, APIs and protocols

1 Scope

This Recommendation | International Standard defines, in an extensible manner, syntaxes and methods for the remote interrogation and optional modification of JPEG 2000 codestreams and files in accordance with their definition in the following parts of ISO/IEC 15444:

- ITU-T Rec. T.800 | ISO/IEC 15444-1:2004 and its definition of a JPEG 2000 codestream and JP2 file format.
- the JPEG 2000 family of file formats as defined in further parts of ISO/IEC 15444.

In this Recommendation | International Standard, the defined syntaxes and methods are referred to as the JPEG 2000 Interactive Protocol, "JPIP", and interactive applications using JPIP are referred to as "JPIP systems."

JPIP specifies a protocol consisting of a structured series of interactions between a client and a server by means of which image file metadata, structure and partial or whole image codestreams may be exchanged in a communications efficient manner. This Recommendation | International Standard includes definitions of the semantics and values to be exchanged, and suggests how these may be passed using a variety of existing network transports.

With JPIP, the following tasks may be accomplished in varying, compatible ways:

- the exchange of capabilities;
- the negotiation of capabilities to use in a session;
- the request and transfer of the following elements from a variety of containers, such as JPEG 2000 family files, JPEG 2000 codestreams and other container files:
  - selective data segments;
  - selective and defined structures;
  - parts of an image or its related metadata.

2 Normative references

The following Recommendations and International Standards contain provisions which, through reference in this text, constitute provisions of this Recommendation | International Standard. At the time of publication, the editions indicated were valid. All Recommendations and Standards are subject to revision, and parties to agreements based on this Recommendation | International Standard are encouraged to investigate the possibility of applying the most recent edition of the Recommendations and Standards listed below. Members of IEC and ISO maintain registers of currently valid International Standards. The Telecommunication Standardization Bureau of the ITU maintains a list of currently valid ITU-T Recommendations.

- ITU-T Recommendation T.800 (2002) | ISO/IEC 15444-1:2004, *Information technology – JPEG 2000 image coding system: Core coding system.*
- ITU-T Recommendation T.801 (2002) | ISO/IEC 15444-2:2004, *Information technology – JPEG 2000 image coding system: Extensions.*
- ITU-T Recommendation T.802 (2005) | ISO/IEC 15444-3:2005, *Information technology – JPEG 2000 image coding system: Motion JPEG 2000.*
- ISO/IEC 15444-6:2003, *Information technology – JPEG 2000 image coding system – Part 6: Compound image file format.*
- IETF RFC 768 (1980), *User Datagram Protocol*. Available from World Wide Web: <http://www.ietf.org/rfc/rfc0768.txt>.

- IETF RFC 793 (1981), *Transmission Control Protocol*. Available from World Wide Web: <http://www.ietf.org/rfc/rfc0793.txt>.
- IETF RFC 2046 (1996), *Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types*. Available from World Wide Web: <http://www.ietf.org/rfc/rfc2046.txt>.
- IETF RFC 2234 (1997), *Augmented BNF for Syntax Specifications: ABNF*. Available from World Wide Web: <http://www.ietf.org/rfc/rfc2234.txt>.
- IETF RFC 2396 (1998), *Uniform Resource Identifiers (URI): Generic Syntax*. Available from World Wide Web: <http://www.ietf.org/rfc/rfc2396.txt>.
- IETF RFC 2616 (1999), *Hypertext Transfer Protocol – HTTP/1.1*. Available from World Wide Web: <http://www.ietf.org/rfc/rfc2616.txt>.

3 Definitions

For the purposes of this Recommendation | International Standard, the following definitions apply.

3.1 JPEG 2000 Part 1 definitions

The definitions defined in clause 3 of ITU-T Rec. T.800 | ISO/IEC 15444-1:2004 and clause 3 of ITU-T Rec. T.801 | ISO/IEC 15444-2:2004 also apply to this Recommendation | International Standard.

3.2 HTTP definitions

The following definitions are intended to match HTTP/1.1. In the case of any difference, these definitions shall be used.

3.2.1 Connection: A transport layer virtual circuit established between two programs for the purpose of communication.

3.2.2 Entity: The information transferred as the payload of a request or response. An entity consists of metainformation in the form of entity-header fields and content in the form of an entity-body.

3.2.3 Proxy: An intermediary program which acts as both a server and a client for the purpose of making requests on behalf of other clients. Requests are serviced internally or by passing them on, with possible translation, to other servers.

3.3 JPIP definitions

The following definitions are used within this Recommendation | International Standard. In some cases, these definitions differ from those used in other standards and/or Recommendations.

3.3.1 cache (client-side): The cache on the Client is the storage of the JPIP data-bins. The Client may have a limited cache and may have to purge cached JPIP data-bins from time to time.

3.3.2 cacheable: A response is cacheable if a cache is allowed to store a copy of the response message for use in answering subsequent requests. Even if a resource is cacheable, there may be additional constraints on whether a cache can use the cached copy for a particular request.

3.3.3 cache-model (server-side): The server's estimation of the portions of the data-bins available in the client's cache. The server may add items to its estimation of the client's cache because it assumes successfully delivery, or because it has received acknowledgements of transmitted data, or because of cache-model update statements.

3.3.4 channel: A mechanism for grouping requests and responses such that only one request/response is active at a time within the group. Multiple simultaneous requests and responses require multiple channels.

3.3.5 client: A program that establishes connections for the purpose of sending requests.

3.3.6 codestream image region: The codestream image region is the intersection between the image and the region defined by the Offset and Region Size. The codestream image region may be empty (no area).

3.3.7 data-bin: A set of bytes of the same type of data which may be partially delivered.

3.3.8 incremental-codestream: The representation of the codestream as a collection of data-bins (main header, tile header, precinct or tile data-bins) having the same codestream identifier.

3.3.9 JPIP index table: A file format box which provides information about the location of portions of a file or codestream.

3.3.10 logical target: A specific representation of specific original named resource, or a byte range from that specific original named resource, to which the JPIP request is directed. This specific representation might be transcoded from the original named resource.

3.3.11 message: A set of bytes from a single data-bin and the header identifying those bytes and the data-bin.

3.3.12 raw-codestream: The representation of the codestream as a single metadata-bin.

3.3.13 request: A group of fields and values sent from the client to the server to obtain portions of an image or metadata.

3.3.14 resource: A network data object or service that can be identified by a URI. A HTTP target.

3.3.15 response: The bytes sent from the server to the client after receiving a request.

3.3.16 server: An application program that accepts connections in order to service requests by sending back responses. Any given program may be capable of being both a client and a server; use of these terms refers only to the role being performed by the program for a particular connection, rather than to the program's capabilities in general.

3.3.17 session: A collection of requests and responses applying to the same resource for which the server maintains a cache model.

3.3.18 session-based: Where the server maintains a cache model.

3.3.19 stateless: A single request where the server does not make use of a cache-model in determining the response.

3.3.20 target: The logical identification of JPIP data. The name of the main target (often the name of a file on the server).

NOTE – JPEG 2000 files or codestreams may be available in multiple representations (e.g., return type, precinct size) or vary in other ways, each identified as a unique logical target.

3.3.21 tile header: All tile-part headers for a specific tile.

3.3.22 view-window: The portion of the image data the client desires, as expressed by the combination of the following fields that appear in the request: Region Size, Offset, Frame Size, Codestream, Codestream Context, Sampling Rate, ROI and Layers. The view-window is often smaller than the whole image data. If a view-window is implied but not specified, then it shall be taken as a view-window on the entire imagery data of the logical target.

3.4 Symbols

For the purposes of this Recommendation | International Standard, the following symbols apply. The symbols defined in clause 4 of ITU-T Rec. T.800 | ISO/IEC 15444-1:2004 and clause 4 of ITU-T Rec. T.801 | ISO/IEC 15444-2:2004 also apply to this Recommendation | International Standard.

c	An index (starting from 0) of the image component to which the precinct belongs
fx	x-axis frame size for client request view-window
fy	y-axis frame size for client request view-window
fx'	x-axis frame size for suitable codestream resolution
fy'	y-axis frame size for suitable codestream resolution
$fx"$	Modified jpx x-axis frame size for suitable resolution
$fy"$	Modified jpx y-axis frame size for suitable resolution
H_{cod}	The codestream height as recorded in the Image Header (ihdr) box (see Annex I.5.3.1 of ITU-T Rec. T.800 | ISO/IEC 15444-1:2004)
H_{comp}	The height of the composited result, supplied in the JPX composition options box (see Annex M.11.10.1 of ITU-T Rec. T.801 | ISO/IEC 15444-2:2004)
H_{reg}	The height of the compositing layer, as it appears on the compositing layer registration grid
Hs_{inst}	The cropped height
Ht_{inst}	The composited height
l	A unique identifier of the precinct within its codestream

N_L	Is the number of decomposition levels
num_components	The number of components encoded
num_tiles	The number of tiles in the codestream
ox	x-axis offset for client request view-window
ox'	x-axis offset for suitable codestream region
ox"	Modified jpx x-axis offset for suitable region
oy	y-axis offset for client request view-window
oy'	y-axis offset for suitable codestream region
oy"	Modified jpx y-axis offset for suitable region
r	Resolution level
s	A sequence number which identifies the precinct within its tile-component
sx	x-axis size of client request view-window
sx'	x-axis size for suitable codestream region
sx"	Modified jpx x-axis size for suitable region
sy	y-axis size of client request view-window
sy'	y-axis size for suitable codestream region
sy"	Modified jpx y-axis size for suitable region
t	An index (starting from 0) of the tile to which the precinct belongs
W_{cod}	The codestream width as recorded in the Image Header (ihdr) box (see Annex I.5.3.1 of ITU-T Rec. T.800 | ISO/IEC 15444-1:2004)
W_{comp}	The width of the composited result, supplied in the JPX composition options box (see Annex M.11.10.1 of ITU-T Rec. T.801 | ISO/IEC 15444-2:2004)
W_{reg}	The width of the compositing layer, as it appears on the compositing layer registration grid
Ws_{inst}	The cropped width
Wt_{inst}	The composited width
XC_{inst}	The x-axis cropping offset supplied via the relevant instruction (see Annex M.11.10.2.1 of ITU-T Rec. T.801 | ISO/IEC 15444-2:2004)
XO_{inst}	The x-axis compositing offset, described via the relevant compositing instruction (see Annex M.11.10.2.1 of ITU-T Rec. T.801 | ISO/IEC 15444-2:2004)
XO_{reg}	The x-axis codestream registration offset
XOsiz	The horizontal offset from the origin of the reference grid of the relevant codestream's SIZ marker segment
XR_{reg}	The x-axis codestream registration sampling factor, described at the beginning of any codestream registration box (see Annex M.11.7.7 of ITU-T Rec. T.801 | ISO/IEC 15444-2:2004)
Xsiz	The width of the reference grid of the relevant codestream's SIZ marker segment
XS_{reg}	The x-axis registration precision described at the beginning of any codestream registration box (see Annex M.11.7.7 of ITU-T Rec. T.801 | ISO/IEC 15444-2:2004)
YC_{inst}	The y-axis cropping offset supplied via the relevant instruction (see Annex M.11.10.2.1 of ITU-T Rec. T.801 | ISO/IEC 15444-2:2004)

YO_{inst}	The y-axis compositing offset, described via the relevant compositing instruction (see Annex M.11.10.2.1 of ITU-T Rec. T.801 | ISO/IEC 15444-2:2004)
YO_{reg}	The y-axis codestream registration offset
YOsiz	The vertical offset from the origin of the reference grid of the relevant codestream's SIZ marker segment
YR_{reg}	The y-axis codestream registration sampling factor, described at the beginning of any codestream registration box (see Annex M.11.7.7 of ITU-T Rec. T.801 | ISO/IEC 15444-2:2004)
Ysiz	The height of the reference grid of the relevant codestream's SIZ marker segment
YS_{reg}	The y-axis registration precision described at the beginning of any codestream registration box (see Annex M.11.7.7 of ITU-T Rec. T.801 | ISO/IEC 15444-2:2004)

4 Abbreviations

For the purposes of this Recommendation | International Standard, the following abbreviations apply.

ABNF	Augmented Backus-Naur Form
DICOM	Digital Imaging and Communications in Medicine
DWT	Discrete Wavelet Transformation
EOR	End of Response
HTML	HyperText Markup Language
IP	Internet Protocol
JP3D	JPEG 2000 Part 10: 3-D and floating point data
JPIP	JPEG 2000 Interactive Protocol
JPP	JPIP Precinct
JPSEC	JPEG 2000 Part 8: Secure JPEG 2000
JPT	JPIP Tile-part
JPWL	JPEG 2000 Part 11: Wireless
JTC 1	Joint Technical Committee 1
MTF	Modulation Transfer Function
PDF	Portable Document Format
SC 29	Sub-Committee 29
SVG	Scalable Vector Graphics
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UUID	Universal Unique Identifier
VBAS	Variable-length Byte Aligned Segment
WG 1	Working Group 1
XHTML	Extensible HyperText Markup Language
XML	Extensible Markup Language

5 Conventions

5.1 ABNF rules

This Recommendation | International Standard uses the ABNF notation defined in RFC 2234, including the core ABNF syntax rules: ALPHA (letters), CR (carriage return), CRLF (Internet standard newline), CTL (control characters), DIGIT (decimal digits), HEXDIG (hexadecimal digits), LF (line feed), LWSP (linear white space) and SP (space). For the purposes of this Recommendation | International Standard, the following ABNF rules also apply.

```
NZDIGIT = %x31-39              ; 1-9
UPPER = %x41-5A                ; A-Z
LOWER = %x61-7A                ; a-z
UINT    = 1*DIGIT
NONZERO = *"0" NZDIGIT *DIGIT
UINT-RANGE = UINT ["-" [UINT]]
UFLOAT = 1*DIGIT ["." 1*DIGIT]
ENCODED-CHAR = "%" HEXDIG HEXDIG
UUID = 16(HEXDIG)
TOKEN = 1*(ALPHA / DIGIT / "." / "_")
```

This Recommendation | International Standard also defines `PATH`, representing a file or pathname. In the general case, `PATH` values may contain any character, although for a given server architecture, the server shall reject any characters that are not legal on that particular server. In addition, `PATH` shall be properly encoded as specified by the transport technology.

`UINT-RANGE` specifies a range of integer values. The first integer in the range specifies the beginning of the range. If two values are specified, the first and second values specify the inclusive beginning and ending limits to the range. If only the first value and the "`-`" character are specified, the range includes all values greater than or equal to the first value.

A numerical value immediately preceding an ABNF element refers to a repetition of the parameter that follows the number, for the number of times given by the numerical value, with no intervening spaces between each occurrence.

The construct "`1#`" refers to one or more repetitions of the parameter that follows, each occurrence of which is separated by a comma.

The construct "`1$`" refers to one or more repetitions of the parameter that follows, each occurrence of which is separated by a semicolon.

5.2 File format ABNF rules


```
compatibility-code = 4(ALPHA / DIGIT / "_" / ENCODED-CHAR)

box-type = 4(ALPHA / DIGIT / "_" / ENCODED-CHAR)

box-type-list = "*" / 1#(box-type)
```

`box-type` specifies the four characters of the box type. For each character in the box type, if the character is alpha-numeric (A..Z, a..z or 0..9), the character is written directly into the string. If the character is a space (0x20), then that character shall be encoded as the underscore character ("_"). For any other character, a 3-character string is written in its place, consisting of a percent character ("%") followed by two hexadecimal digits representing the value of the character from the box type in hexadecimal. The `compatibility-code` is encoded the same way that a `box-type` is encoded.

`box-type-list` specifies a list of box types. If the value of a `box-type-list` field is "*", then the field refers to all box types.

5.3 Key to graphical descriptions of boxes (informative)

The description of each box is followed by a figure that shows the order and relationship of the parameters in the box. Figure 1 shows an example of this type of figure. A rectangle is used to indicate the parameters in the box. The width of the rectangle is proportional to the number of bytes in the parameter. A shaded rectangle (diagonal stripes) indicates that the parameter is of varying size. Two parameters with superscripts and a grey area between indicate a run of several of these parameters. A sequence of two groups of multiple parameters with superscripts separated by a grey area indicates a run of that group of parameters (one set of each parameter in the group, followed by the next set of each parameter in the group). Optional parameters or boxes will be shown with a dashed rectangle.

The figure is followed by a list that describes the meaning of each parameter in the box. If parameters are repeated, the length and nature of the run of parameters is defined. As an example, in Figure 1, parameters A, B, C and D are 8, 16, 32 bit and variable length respectively. The notation E^0 and E^{N-1} implies that there are N different parameters, E^i, in a row. The group of parameters F^0 and F^{M-1}, and G^0 and G^{M-1} specify that the box will contain F^0, followed by G^0,

followed by F^1 and G^1, continuing to F^{M-1} and G^{M-1} (M instances of each parameter in total). Also, the field D is optional and may not be found in this box.

In addition, in a figure describing the contents of a superbox, an ellipsis (…) will be used to indicate that the contents of the file between two boxes is not specifically defined. Any box (or sequence of boxes), unless otherwise specified by the definition of that box, may be found in place of the ellipsis.

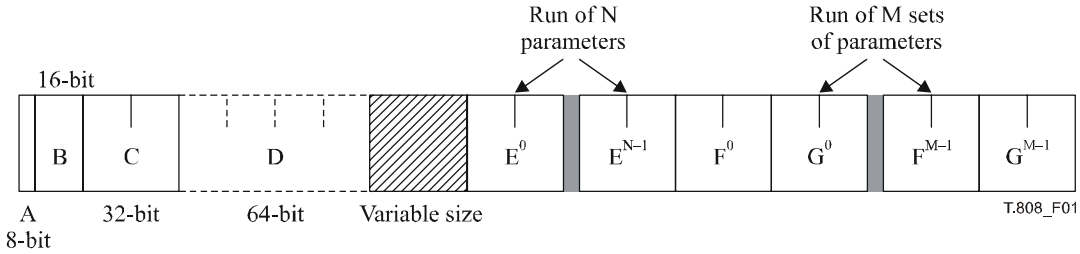

Figure 1 – Example of the box description figures

For example, the superbox shown in Figure 2 must contain an AA box and a BB box, and the BB box must follow the AA box. However, there may be other boxes found between boxes AA and BB. Dealing with unknown boxes is discussed in Annex I.8 of ITU-T Rec. T.800 | ISO/IEC 15444-1:2004.

Figure 2 – Example of the superbox description figures

6 General description

6.1 JPIP protocol

This Recommendation | International Standard describes the syntaxes and methods that are used when a client is accessing JPEG 2000 compressed imagery and imagery related data residing on a JPIP-enabled server. This Recommendation | International Standard enables the flexibility and functionality intended in ITU-T Rec. T.800 | ISO/IEC 15444-1:2004 to be realized across multiple client/server transports.

JPIP defines the interactive protocol to achieve the efficient exchange of JPEG 2000 imagery and imagery-related data. The protocol defines the Client-Server interactions based on a client request and server response as shown in Figure 3. This Recommendation | International Standard defines the JPIP client requests and the JPIP server responses. HTTP/1.1 (RFC 2616), TCP (RFC 793) and UDP (RFC 768) are shown as examples of possible transports for JPIP. The client uses a View-Window request to define the resolution, size, location, components, layers, and other parameters for the image and imagery related data that is requested by the client. The server responds by delivering imagery and imagery-related data with precinct-based streams, tile-based streams, or whole images. The protocol also allows for the negotiation of client and server capabilities and limitations. The client may request information about an image as defined in JPIP index tables from the server, which enables the client to refine its View-Window request to image specific parameters (e.g., byte range requests). The server's cache model is based on the capabilities defined by the client and the statefulness of the session.