INTERNATIONAL STANDARD

ISO/IEC 15444-1

Third edition
2016-10-15

Information technology — JPEG 2000 image coding system: Core coding system

Technologies de l'information — Système de codage d'images JPEG 2000: Système de codage de noyau

Reference number
ISO/IEC 15444-1:2016(E)

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: Foreword - Supplementary information

This third edition cancels and replaces the second edition of ISO/IEC 15444-1:2004 which has been technically revised. It also incorporates ISO/IEC 15444-1:2004/Cor.1:2007, ISO/IEC 15444-1:2004/Cor.2:2008, ISO/IEC 15444-1:2004/Cor.3:2015, ISO/IEC 15444-1:2004/Cor.4:2015, ISO/IEC 15444-1:2004/Amd.1:2006, ISO/IEC 15444-1:2004/Amd.2:2009, ISO/IEC 15444-1:2004/Amd.3:2010, ISO/IEC 15444-1:2004/Amd.4:2013, ISO/IEC 15444-1:2004/Amd.5:2013, ISO/IEC 15444-1:2004/Amd.6:2013, ISO/IEC 15444-1:2004/Amd.7:2015 and ISO/IEC 15444-1:2004/Amd.8:2015.

ISO/IEC 15444-1:2016 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 Rec. ITU-T 800 (11/2015).

CONTENTS

INTERNATIONAL STANDARD
ITU-T RECOMMENDATION

Information technology – JPEG 2000 image coding system: Core coding system

1 Scope

This Recommendation | International Standard defines a set of lossless (bit-preserving) and lossy compression methods for coding bi-level, continuous-tone grey-scale, palletized colour, or continuous-tone colour digital still images.

This Recommendation | International Standard:

- specifies decoding processes for converting compressed image data to reconstructed image data;
- specifies a codestream syntax containing information for interpreting the compressed image data;
- specifies a file format;
- provides guidance on encoding processes for converting source image data to compressed image data;
- provides guidance on how to implement these processes in practice.

NOTE – As this specification was first published as common text only after ISO/IEC JTC1 had approved the first edition in 2000, edition numbers in the ITU and ISO/IEC versions are offset by one. This is the second edition of ITU-T T.800 and the third edition of ISO/IEC 15444-1.

2 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.

2.1 Identical Recommendations | International Standards

- Recommendation ITU-T T.81 (1992) | ISO/IEC 10918-1:1994, *Information technology – Digital compression and coding of continuous-tone still images: Requirements and guidelines*.
- Recommendation ITU-T T.84 (1996) | ISO/IEC 10918-3:1997, *Information technology – Digital compression and coding of continuous-tone still images: Extensions*.
- Recommendation ITU-T T.84 (1996)/Amd.1 (1999) | ISO/IEC 10918-3:1997/Amd.1:1999, *Information technology – Digital compression and coding of continuous-tone still images: Extensions – Amendment 1: Provisions to allow registration of new compression types and versions in the SPIFF header*.
- Recommendation ITU-T T.86 (1998) | ISO/IEC 10918-4:1999, *Information technology – Digital compression and coding of continuous-tone still images: Registration of JPEG Profiles, SPIFF Profiles, SPIFF Tags, SPIFF colour Spaces, APPn Markers, SPIFF Compression types and Registration Authorities (REGAUT)*.
- Recommendation ITU-T T.87 (1998) | ISO/IEC 14495-1:2000, *Lossless and near-lossless compression of continuous-tone still images – Baseline*.
- Recommendation ITU-T T.88 (2000) | ISO/IEC 14492:2001, *Information technology – Lossy/lossless coding of bi-level images*.
- Recommendation ITU-T T.810 (2006) | ISO/IEC 15444-11:2007, *Information technology – JPEG 2000 image coding system: Wireless*.
- ISO/IEC 646:1991, *Information technology – ISO 7-bit coded character set for information interchange*.
- ISO 8859-15:1999, *Information technology – 8-bit single-byte coded graphic character sets – Part 15: Latin alphabet No. 9*.

2.2 Additional references

- Recommendation ITU-R BT.601-6 (2007), *Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios*.

- Recommendation ITU-R BT.709-5 (2002), *Parameter values for the HDTV standards for production and international programme exchange*.
- IEC 61966-2-1:1999, *Multimedia systems and equipment – Colour measurement and management – Part 2-1: Colour management – Default RGB colour space – sRGB*.
- IEC 61966-2-1:1999/Amd.1:2003, *Multimedia systems and equipment – Colour measurement and management – Part 2-1: Colour management – Default RGB colour space – sRGB*.
- IETF RFC 2279 (1998), UTF-8, *a transformation format of ISO 10646*.
- ISO 11664-1:2007 (CIE S 014-1/E:2006), *Colorimetry – Part 1: CIE standard colorimetric observers*.
- ISO 14721, *Space data and information transfer systems – Open archival information system – Reference model*.
- ISO 15076-1, *Image technology colour management – Architecture, profile format and data structure – Part 1: Based on ICC.1:2010*.
- ISO 26428-1:2008, *Digital cinema (D-cinema) distribution master – Part 1: Image characteristics*.
- ISO/IEC 11578:1996, *Information technology – Open Systems Interconnection – Remote Procedure Call*.

3 Definitions

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

3.1 $\lfloor x \rfloor$, **floor function**: This indicates the largest integer not exceeding x.

3.2 $\lceil x \rceil$, **ceiling function**: This indicates the smallest integer not exceeded by x.

3.3 **5-3 reversible filter**: A particular filter pair used in the wavelet transformation. This reversible filter pair has 5 taps in the low-pass and 3 taps in the high-pass.

3.4 **9-7 irreversible filter**: A particular filter pair used in the wavelet transformation. This irreversible filter pair has 9 taps in the low-pass and 7 taps in the high-pass.

3.5 **access unit**: A coded representation of one video frame.

3.6 **AND**: Bit wise AND logical operator.

3.7 **arithmetic coder**: An entropy coder that converts variable length strings to variable length codes (encoding) and vice versa (decoding).

3.8 **auxiliary channel**: A channel that is used by the application outside the scope of colourspace conversion. For example, an opacity channel or a depth channel would be an auxiliary channel.

3.9 **bit**: A contraction of the term "binary digit"; a unit of information represented by a zero or a one.

3.10 **bit-plane**: A two dimensional array of bits. In this Recommendation | International Standard a bit-plane refers to all the bits of the same magnitude in all coefficients or samples. This could refer to a bit-plane in a component, tile-component, code-block, region of interest, or other.

3.11 **bit stream**: The actual sequence of bits resulting from the coding of a sequence of symbols. It does not include the markers or marker segments in the main and tile-part headers or the EOC marker. It does include any packet headers and in-stream markers and marker segments not found within the main or tile-part headers.

3.12 **big-endian**: The bits of a value representation occur in order from the most significant to the least significant.

3.13 **box**: A portion of the file format defined by a length and unique box type. Boxes of some types may contain other boxes.

3.14 **box contents**: Refers to the data wrapped within the box structure. The contents of a particular box are stored within the DBox field within the box data structure.

3.15 **box type**: Specifies the kind of information that shall be stored with the box. The type of a particular box is stored within the TBox field within the box data structure.

3.16 **byte**: Eight bits.

3.17 **channel**: One logical component of the image. A channel may be a direct representation of one component from the codestream, or may be generated by the application of a palette to a component from the codestream.

3.18 cleanup pass: A coding pass performed on a single bit-plane of a code-block of coefficients. The first pass and only coding pass for the first significant bit-plane is a cleanup pass; the third and the last pass of every remaining bit-plane is a cleanup pass.

3.19 codestream: A collection of one or more bit streams and the main header, tile-part headers, and the EOC required for their decoding and expansion into image data. This is the image data in a compressed form with all of the signalling needed to decode.

3.20 code-block: A rectangular grouping of coefficients from the same sub-band of a tile-component.

3.21 code-block scan: The order in which the coefficients within a code-block are visited during a coding pass. The code-block is processed in stripes, each consisting of four rows (or all remaining rows if less than four) and spanning the width of the code-block. Each stripe is processed column by column from top to bottom and from left to right.

3.22 coder: An embodiment of either an encoding or decoding process.

3.23 coding pass: A complete pass through a code-block where the appropriate coefficient values and context are applied. There are three types of coding passes: significance propagation pass, magnitude refinement pass and cleanup pass. The result of each pass (after arithmetic coding, if selective arithmetic coding bypass is not used) is a stream of compressed image data.

3.24 coefficient: The values that are the result of a transformation.

3.25 colour channel: A channel that functions as an input to a colour transformation system. For example, a red channel or a greyscale channel would be a colour channel.

3.26 component: A two-dimensional array of samples. An image typically consists of several components, for instance, representing red, green, and blue.

3.27 compressed image data: Part or all of a bit stream. It can also refer to a collection of bit streams in part or all of a codestream.

3.28 conforming reader: An application that reads and interprets a JP2 file correctly.

3.29 context: Function of coefficients previously decoded and used to condition the decoding of the present coefficient.

3.30 context label: The arbitrary index used to distinguish different context values. The labels are used as a convenience of notation rather than being normative.

3.31 context vector: The binary vector consisting of the significance states of the coefficients included in a context.

3.32 decoder: An embodiment of a decoding process, and optionally a colour transformation process.

3.33 decoding process: A process which takes as its input all or part of a codestream and outputs all or part of a reconstructed image.

3.34 decomposition level: A collection of wavelet sub-bands where each coefficient has the same spatial impact or span with respect to the source component samples. These include the HL, LH, and HH sub-bands of the same two dimensional sub-band decomposition. For the last decomposition level, the LL sub-band is also included.

3.35 delimiting markers and marker segments: Markers and marker segments that give information about beginning and ending points of structures in the codestream.

3.36 discrete wavelet transformation (DWT): A transformation that iteratively transforms one signal into two or more filtered and decimated signals corresponding to different frequency bands. This transformation operates on spatially discrete samples.

3.37 encoder: An embodiment of an encoding process.

3.38 encoding process: A process that takes as its input all or part of the source image data and outputs a codestream.

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

3.40 fixed information markers and fixed information marker segments: Markers and marker segments that offer information about the original image.

3.41 functional markers and functional marker segments: Markers and marker segments that offer information about coding procedures.

3.42 **grid resolution**: The spatial resolution of the reference grid, specifying the distance between neighbouring points on the reference grid.

3.43 **guard bits**: Additional most significant bits that have been added to sample data.

3.44 **header**: Either a part of the codestream that contains only markers and marker segments (main header and tile-part header) or the signalling part of a packet (packet header).

3.45 **HH sub-band**: The sub-band obtained by forward horizontal high-pass filtering and vertical high-pass filtering. This sub-band contributes to reconstruction with inverse vertical high-pass filtering and horizontal high-pass filtering.

3.46 **HL sub-band**: The sub-band obtained by forward horizontal high-pass filtering and vertical low-pass filtering. This sub-band contributes to reconstruction with inverse vertical low-pass filtering and horizontal high-pass filtering.

3.47 **image**: The set of all components.

3.48 **image area**: A rectangular part of the reference grid, registered by offsets from the origin and the extent of the reference grid.

3.49 **image area offset**: The number of reference grid points down and to the right of the reference grid origin where the origin of the image area can be found.

3.50 **image data**: The components and component samples making up an image. Image data can refer to either the source image data or the reconstructed image data.

3.51 **in-bit-stream markers and in-bit-stream marker segments**: Markers and marker segments that provide error resilience functionality.

3.52 **informational markers and informational marker segments**: Markers and marker segments that offer ancillary information.

3.53 **instantaneous bit rate**: For each frame, this corresponds to the size of the contiguous codestream for the frame in bits multiplied by the frame rate.

3.54 **irreversible**: A transformation, progression, system, quantization, or other process that, due to a systemic or quantization error, disallows lossless recovery. An irreversible process can only lead to lossy compression.

3.55 **JP2 file**: The name of a file in the file format described in this Recommendation | International Standard. Structurally, a JP2 file is a contiguous sequence of boxes.

3.56 **JPEG**: Used to refer globally to the encoding and decoding process of the following Recommendations | International Standards:

- Rec. ITU-T T.81 | ISO/IEC 10918-1;
- Rec. ITU-T T.83 | ISO/IEC 10918-2;
- Rec. ITU-T T.84 | ISO/IEC 10918-3;
- Rec. ITU-T T.86 | ISO/IEC 10918-4.

3.57 **JPEG 2000**: Used to refer globally to the encoding and decoding processes in this Recommendation | International Standard and their embodiment in applications.

3.58 **LH sub-band**: The sub-band obtained by forward horizontal low-pass filtering and vertical high-pass filtering. This sub-band contributes to reconstruction with inverse vertical high-pass filtering and horizontal low-pass filtering.

3.59 **LL sub-band**: The sub-band obtained by forward horizontal low-pass filtering and vertical low-pass filtering. This sub-band contributes to reconstruction with inverse vertical low-pass filtering and horizontal low-pass filtering.

3.60 **layer**: A collection of compressed image data from coding passes of one or more code-blocks of a tile-component. Layers have an order for encoding and decoding that must be preserved.

3.61 **lossless**: A descriptive term for the effect of the overall encoding and decoding processes in which the output of the decoding process is identical to the input to the encoding process. Distortion-free restoration can be assured. All of the coding processes or steps used for encoding and decoding are reversible.

3.62 **lossy**: A descriptive term for the effect of the overall encoding and decoding processes in which the output of the decoding process is not identical to the input to the encoding process. There is distortion (measured mathematically). At least one of the coding processes or steps used for encoding and decoding is irreversible.

3.63 **magnitude refinement pass**: A type of coding pass.

3.64 **main header**: A group of markers and marker segments at the beginning of the codestream that describe the image parameters and coding parameters that can apply to every tile and tile-component.

3.65 **marker**: A two-byte code in which the first byte is hexadecimal FF (0xFF) and the second byte is a value between 1 (0x01) and hexadecimal FE (0xFE).

3.66 **marker segment**: A marker and associated (not empty) set of parameters.

3.67 **mod**: mod(y,x) = z, where z is such that $0 \leq z < x$, and such that y – z is a multiple of x.

3.68 **packet**: A part of the bit stream comprising a packet header and the compressed image data from one layer of one precinct of one resolution level of one tile-component.

3.69 **packet header**: Portion of the packet that contains the signalling necessary for decoding that packet.

3.70 **pointer markers and pointer marker segments**: Markers and marker segments that offer information about the location of structures in the codestream.

3.71 **precinct**: A one rectangular region of a transformed tile-component, within each resolution level, used for limiting the size of packets.

3.72 **precision**: Number of bits allocated to a particular sample, coefficient or other binary numerical representation.

3.73 **progression**: The order of a codestream where the decoding of each successive bit contributes to a "better" reconstruction of the image. What metrics make the reconstruction "better" is a function of the application. Some examples of progression are increasing resolution or improved sample fidelity.

3.74 **quantization**: A method of reducing the precision of the individual coefficients to reduce the number of bits used to entropy-code them. This is equivalent to division while compressing and multiplying while decompressing. Quantization can be achieved by an explicit operation with a given quantization value or by dropping (truncating) coding passes from the codestream.

3.75 **raster order**: A particular sequential order of data of any type within an array. The raster order starts with the top left data point and moves to the immediate right data point, and so on, to the end of the row. After the end of the row is reached the next data point in the sequence is the left-most data point immediately below the current row. This order is continued to the end of the array.

3.76 **reconstructed image**: An image that is the output of a decoder.

3.77 **reconstructed sample**: A sample reconstructed by the decoder. This always equals the original sample value in lossless coding but may differ from the original sample value in lossy coding.

3.78 **reference grid**: A regular rectangular array of points used as a reference for other rectangular arrays of data. Examples include components and tiles.

3.79 **reference tile**: A rectangular sub-grid of any size associated with the reference grid.

3.80 **region of interest (ROI)**: A collections of coefficients that are considered of particular relevance by some user-defined measure.

3.81 **resolution level**: Equivalent to the decomposition level with one exception: the LL sub-band is also a separate resolution level.

3.82 **reversible**: A transformation, progression, system or other process that does not suffer a systemic or quantization error and therefore, allows lossless signal recovery.

3.83 **sample**: One element in the two-dimensional array that comprises a component.

3.84 **segmentation symbol**: A special symbol coded with a uniform context at the end of each coding pass for error resilience.

3.85 **selective arithmetic coding bypass**: A coding style where some of the code-block passes are not coded by the arithmetic coder. Instead the bits to be coded are appended directly to the bit stream without coding.

3.86 **shift**: Multiplication or division of a number by powers of two.

3.87 **sign bit**: A bit that indicates whether a number is positive (zero value) or negative (one value).

3.88 **sign-magnitude notation**: A binary representation of an integer where the distance from the origin is expressed with a positive number and the direction from the origin (positive or negative) is expressed with a separate single sign bit.

3.89 **significance propagation pass**: A coding pass performed on a single bit-plane of a code-block of coefficients.

3.90 **significance state**: State of a coefficient at a particular bit-plane. If a coefficient, in sign-magnitude notation, has the first magnitude 1 bit at or before the given bit-plane, it is considered "significant". If not, it is considered "insignificant".

3.91 **source image**: An image used as input to an encoder.

3.92 **sub-band**: A group of transform coefficients resulting from the same sequence of low-pass and high-pass filtering operations, both vertically and horizontally.

3.93 **sub-band coefficient**: A transform coefficient within a given sub-band.

3.94 **sub-band decomposition**: A transformation of an image tile-component into sub-bands.

3.95 **superbox**: A box that itself contains a contiguous sequence of boxes (and only a contiguous sequence of boxes). As the JP2 file contains only a contiguous sequence of boxes, the JP2 file is itself considered a superbox. When used as part of a relationship between two boxes, the term "superbox" refers to the box which directly contains the other box.

3.96 **tile**: A rectangular array of points on the reference grid, registered with and offset from the reference grid origin and defined by a width and height. The tiles which overlap are used to define tile-components.

3.97 **tile-component**: All the samples of a given component in a tile.

3.98 **tile index**: The index of the current tile ranging from zero to the number of tiles minus one.

3.99 **tile-part**: A portion of the codestream with compressed image data for some, or all of a tile. The tile-part includes at least one, and up to all of the packets that make up the coded tile.

3.100 **tile-part header**: A group of markers and marker segments at the beginning of each tile-part in the codestream that describe the tile-part coding parameters.

3.101 **tile-part index**: The index of the current tile-part ranging from zero to the number of tile-parts minus one in a given tile.

3.102 **transformation**: A mathematical mapping from one signal space to another.

3.103 **transform coefficient**: A value that is the result of a transformation.

3.104 **XOR**: Exclusive OR logical operator.

4 Abbreviations and symbols

4.1 Abbreviations

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

1D-DWT	One-dimensional Discrete Wavelet Transformation
CCITT	International Telegraph and Telephone Consultative Committee, now ITU-T
CPRL	Component, Position, Resolution and Layer data packet progression order within a tile
CSF	Contrast Sensitivity Function
DCP	Digital Cinema Package
FDWT	Forward Discrete Wavelet Transformation
FEC	Forward Error Correction
ICC	International Color Consortium
ICT	Irreversible Component Transform
IDWT	Inverse Discrete Wavelet Transformation
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
ITTF	Information Technology Task Force
ITU	International Telecommunication Union
ITU-T	International Telecommunication Union – Telecommunication Standardization Sector (formerly the CCITT)