INTERNATIONAL STANDARD

ISO/IEC 15444-2

First edition
2004-05-15

Information technology — JPEG 2000 image coding system: Extensions

Technologies de l'information — Système de codage d'image JPEG 2000: Extensions

Reference number
ISO/IEC 15444-2:2004(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) 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 15444-2:2004 was prepared jointly 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.801.

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 9: Interactivity tools, APIs and protocols*

— *Part 12: ISO base media file format*

The following parts are under preparation:

— *Part 8: Secure JPEG 2000*

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

— *Part 11: Wireless JPEG 2000*

INTERNATIONAL STANDARD
ITU-T RECOMMENDATION

Information technology – JPEG 2000 image coding system: Extensions

1 Scope

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

This Recommendation | International Standard:

- specifies extended decoding processes for converting compressed image data to reconstructed image data;
- specifies an extended codestream syntax containing information for interpreting the compressed image data;
- specifies an extended file format;
- specifies a container to store image metadata;
- defines a standard set of image metadata;
- provides guidance on extended encoding processes for converting source image data to compressed image data;
- provides guidance on how to implement these processes in practice.

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

- ITU-T Recommendation T.81 (1992) | ISO/IEC 10918-1:1994, *Information technology – Digital compression and coding of continuous-tone still images: Requirements and guidelines.*
- ITU-T Recommendation T.82 (1993) | ISO/IEC 11544:1993, *Information technology – Coded representation of picture and audio information – Progressive bi-level image compression.*
- ITU-T Recommendation T.83 (1994) | ISO/IEC 10918-2:1995, *Information technology – Digital compression and coding of continuous-tone still images: Compliance testing.*
- ITU-T Recommendation T.84 (1996) | ISO/IEC 10918-3:1997, *Information technology – Digital compression and coding of continuous-tone still images: Extensions.*
- ITU-T Recommendation 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.*
- ITU-T Recommendation 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).*
- ITU-T Recommendation T.87 (1998) | ISO/IEC 14495-1:2000, *Information technology – Lossless and near-lossless compression of continuous-tone still images – Baseline.*

- ITU-T Recommendation T.88 (2000) | ISO/IEC 14492:2001, *Information technology – Lossy/lossless coding of bi-level images.*
- ITU-T Recommendation T.800 (2002) | ISO/IEC 15444-1:2003, *Information technology – JPEG 2000 image coding system: Core coding system.*

2.2 Additional references

- ITU-T Recommendation T.42 (1996), *Continuous-tone colour representation method for facsimile.*
- ISO/IEC 8859-1:1998, *Information technology – 8-bit single-byte coded graphic character sets – Part 1: Latin alphabet No. 1.*
- ISO 8601:2000, *Data elements and interchange formats – Information interchange – Representation of dates and times.*
- ISO 3166-1:1997, *Codes for the representation of names of countries and their subdivisions – Part 1: Country codes.*
- ISO 3166-2:1998, *Codes for the representation of names of countries and their subdivisions – Part 2: Country subdivision code.*
- ISO/IEC 11578:1996, *Information technology – Open Systems Interconnection – Remote Procedure Call (RPC).*
- ISO/IEC 646:1991, *Information technology – ISO 7-bit coded character set for information interchange.*
- ISO 5807:1985, *Information processing – Documentation symbols and conventions for data, program and system flowcharts, program network charts and system resources charts.*
- ISO/IEC 15938, *MPEG-7.*
- ISO 10126-2:1991, *Banking – Procedures for message encipherment (wholesale) – Part 2: DEA algorithm.*
- IEEE Standard 754-1985 R1990, *IEEE Standard for Binary Floating-Point Arithmetic.*
- IETF RFC 1321 (1992), *The MD5 Message-Digest Algorithm.*
- IETF RFC 1766 (1995), *Tags for the Identification of Languages.*
- IETF RFC 2279 (1998), *UTF-8, A transformation format of ISO 10646.*
- IETF RFC 2630 (1999), *Cryptographic Message Syntax.*
- IETF RFC 2313 (1998), *PKCS #1: RSA Encryption, Version 1.5.*
- International Color Consortium, ICC profile format specification. ICC.1.
- IEC 61966-2-1:1999, *Multimedia systems and equipment – Colour measurement and management: Part 2-1: Colour management – Default RGB colour space – sRGB, plus its Amendment 1: 2003.*
- Digital Imaging Group, Flashpix digital image file format. Version 1.0.1, 10 July 1997.
- PIMA 7666. *Photography-Electronics still picture imaging-Reference Output Medium Metric RGB Color encoding: ROMM-RGB*.
- PIMA 7667:2001. *Photography-Electronics still picture imaging-Extended sRGB color encoding e-sRGB*.
- Federal Information Processing Standard Publication (FIPS PUB) 186-2, Digital Signature Standard (DSS). <http://www.itl.nist.gov/fipspubs/fip186-2.pdf>
- ANSI X9.30.2-1997, *Public Key Cryptography for the Financial Services Industry – Part 2: The Secure Hash Algorithm (SHA-1)*. <http://www.itl.nist.gov/fipspubs/fip180-1.htm>
- W3C. Extensible Markup Language (XML 1.0), 2nd edition Rec-xml-2000106, <http://www.w3.org/TR/REC-xml>.
- W3C. Namespaces in XML, Rec-xml-names-19990114, <http://www.w3.org/TR/1999/REC-xml-names>.
- W3C. XML Schema Part 1: Structures, Rec-xmlschema-1-20010502, <http://www.w3.org/TR/xmlschema-1>.
- W3C. XML Schema Part 2: Datatypes, Rec-xmlschema-2-20010502, <http://www.w3.org/TR/xmlschema-2>.

3 Definitions

For the purposes of this Recommendation | International Standard, the following definitions apply. The definitions defined in ITU-T Rec. T.800 | ISO/IEC 15444-1 clause 3 also apply to this Recommendation | International Standard except for the terms decomposition level, sub-band and resolution which are redefined in this clause.

3.1 **attribute**: An XML construct that is a name-value pair extending or qualifying the meaning of an element.

3.2 **cell**: An optional subdivision of a tile used for low-memory encoding and decoding.

3.3 **component**: Compressed data from the codestream representing a single set of two-dimensional data.

3.4 **component collection**: A subset of intermediate components used as inputs to a multiple component transformation stage, and a subset of intermediate components obtained as outputs from a multiple component transformation stage. The subset's constituent components may occur in an arbitrary order, i.e., permuted with respect to their order of appearance in the set of input or output intermediate components.

3.5 **component reconstruction arrays**: A general term that refers to any of the following; decorrelation transformation array, dependency transformation array or offset array.

3.6 **compositing**: The act of combining two compositing layers into a single, non-redundant set of image channels.

3.7 **compositing layer**: A set of non-redundant channels drawn from one or more codestreams that shall be treated as a group. The set of compositing layers within the JPX file may then be combined by compositing or animation instructions to form a rendered result. For example, one layer may be a simple RGBA codestream. Another layer may consist of R, G and B channels generated by the application of a palette to one component from codestream 1, and an opacity channel directly extracted from codestream 2.

3.8 **deadzone**: The interval within which all sub-band coefficients are quantized to 0.

3.9 **decomposition level**: A collection of sub-bands where each coefficient has the same spatial impact or span with respect to the original samples. These include the LL, LH, HL, HH, LX, HX, XL, and XH sub-band splits out of decomposition sublevels.

3.10 **decomposition sub-level**: A collection of sub-bands that result from splits of a sub-band from a lower decomposition sub-level or splits of either LL, LX or XL sub-bands from a higher decomposition level.

3.11 **decorrelation transformation array**: An array of coefficients that maps the input components of a component collection to the output components of the collection via a multiple component decorrelation transformation.

3.12 **dependency transformation array**: An array of coefficients that maps the input components of a component collection to the output components of the collection via a multiple component dependency transformation.

3.13 **element**: An XML construct that consists of a start tag and an end tag with data enclosed within.

3.14 **HX sub-band**: The sub-band obtained by forward horizontal high-pass analysis filtering and no vertical analysis filtering. This sub-band contributes to reconstruction with inverse horizontal high-pass synthesis filtering and no vertical synthesis filtering.

3.15 **intermediate component**: A single two-dimensional array of data involved in a stage of a multiple component transformation.

3.16 **JPX baseline**: A specific subset of the features of the JPX file format.

3.17 **JPX baseline reader**: An application that correctly interprets all files that conform to the definition of a JPX baseline file.

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

3.19 **LX sub-band**: The sub-band obtained by forward horizontal low-pass analysis filtering and no vertical analysis filtering. This sub-band contributes to reconstruction with inverse horizontal low-pass synthesis filtering and no vertical synthesis filtering.

3.20 **metadata**: Additional data associated with the image data beyond the image data**.**

3.21 **namespace**: A collection of names, identified by a URI, that allows XML documents of different sources to use the same element names within a single document to avoid element name conflicts**.**

3.22 **offset array**: An array of coefficients containing offsets which are added to intermediate components during multiple component transformation of a component collection.

3.23 **reconstructed image component**: The set of output intermediate components from the final transformation stage in the inverse multiple component transformation process.

3.24 **rendered result**: The result generated by combining the compositing layers in the JPX file, either by composition or animation.

3.25 **resolution**: The spatial relation of samples to a physical space. In this Recommendation | International Standard, the decomposition levels of the wavelet transformation create resolutions that differ by powers of two in either just horizontal, just vertical or both horizontal and vertical directions. The last (highest) decomposition level includes either an LL, LX or XL sub-band which is considered to be a lower resolution. Therefore, there is one more resolution level than decomposition levels.

3.26 **sub-band**: A group of transformation coefficients resulting from the sequence of low-pass and high-pass filtering operations, either just horizontally, just vertically or both horizontally and vertically.

3.27 **spatially reconstructed component**: A component which has been extracted from the codestream and passed through the decoding and inverse wavelet transformation process as specified by this Recommendation | International Standard. The set of spatially reconstructed components is the set of input components to the first transformation stage in the inverse multiple component transformation process.

3.28 **transformation stage**: A set of component collections and associated multiple component transformations.

3.29 **visual masking**: Visual masking is a mechanism where artefacts are masked by the image acting as a background signal.

3.30 **XH sub-band**: The sub-band obtained by no forward horizontal analysis filtering and vertical high-pass analysis filtering. This sub-band contributes to reconstruction with vertical high-pass synthesis filtering and no inverse horizontal synthesis filtering.

3.31 **XL sub-band**: The sub-band obtained by no forward horizontal analysis filtering and vertical low-pass analysis filtering. This sub-band contributes to reconstruction with vertical low-pass synthesis filtering and no inverse horizontal synthesis filtering.

4 Abbreviations

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

CCITT International Telegraph and Telephone Consultative Committee, now ITU-T

DPI Dots per inch

IPR Intellectual Property Rights

UUID Universal Unique Identifier

5 Symbols

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

ADS Arbitrary decomposition styles marker

ATK Arbitrary transformation kernels marker

CBD Component bit depth definition marker

DCO Variable DC offset marker

DFS Downsample factor styles marker

MCC Multiple component collection transformation marker

MCO Multiple component transformation ordering marker

MCT Multiple component transformation definition marker

NLT Non-linearity point transformation marker

VMS Visual masking marker

6 General description

The purpose of this clause is to give an overview of this Recommendation | International Standard. Terms defined in previous clauses in this Recommendation | International Standard will also be introduced. (Terms defined in clauses 3 and 4 in ITU-T Rec. T.800 | ISO/IEC 15444-1 continue to apply in this Recommendation | International Standard.)

This Recommendation | International Standard defines a set of lossless (bit-preserving) and lossy compression methods for coding continuous-tone, bi-level, grey-scale, colour digital still images, or multi-component images. This set of methods extends the elements in the core coding system described in ITU-T Rec. T.800 | ISO/IEC 15444-1. Extensions which pertain to encoding and decoding are defined as procedures which may be used in combination with the encoding and decoding processes described in ITU-T Rec. T.800 | ISO/IEC 15444-1. Each encoding or decoding extension shall only be used in combination with particular coding processes and only in accordance with the requirements set forth herein. These extensions are backward compatible in the sense that decoders which implement theses extensions will also support configuration subsets that are currently defined by ITU-T Rec. T.800 | ISO/IEC 15444-1. This Recommendation | International Standard also defines extensions to the compressed data format, i.e., interchange format and the abbreviated formats.

6.1 Extensions specified by this Recommendation | International Standard

The following extensions are specified in this Recommendation | International Standard.

6.1.1 Syntax

An extension of the code stream syntax is described in Annex A. This extension provides all the codestream signalling in this Recommendation | International Standard. Further, it anticipates signalling needed for future specifications that include this Recommendation | International Standard as a normative reference. In addition to the codestream syntax defined in ITU-T Rec. T.800 | ISO/IEC 15444-1, the following capabilities are supported: variable DC offset, variable scalar quantization, trellis coded quantization, visual masking, arbitrary decomposition, arbitrary transformation kernels, single sample overlap, multiple component transformations, non-linear transformation, arbitrary regions of interest. These extended markers conform to the same rules as the syntax in ITU-T Rec. T.800 | ISO/IEC 15444-1.

6.1.2 Variable DC offset

An extension which provides for variable DC offset is described in Annex B. Variable DC offset may be used to generate a better data distribution for input to the ICT or RCT multi component transformation, defined in ITU-T Rec. T.800 | ISO/IEC 15444-1, and/or the wavelet transformation. Images with very skewed sample distributions may benefit from a non-default DC offset.

6.1.3 Variable scalar quantization

An extension that provides for variable scalar quantization is described in Annex C. This extension allows smaller or larger deadzones to be used with the scalar quantizer. This technique may improve visual appearance of low level texture.

6.1.4 Trellis coded quantization

An extension of the quantization is described in Annex D. This extension provides for trellis coded quantization (TCQ). The TCQ algorithm applies spatial-varying scalar quantization to its input sequence by choosing one of four scalar quantizers for each sample. Quantizer indices from supersets of these quantizers along with quantizer transitions in the form of a trellis provide all information necessary to reconstruct TCQ encoded wavelet coefficients.

6.1.5 Visual masking

An extension which provides for visual masking is described in Annex E. Visual masking is a mechanism where artefacts are masked by the image acting as a background signal. The main goal is to improve the image quality, especially for displays. The first effect of this technique is to improve the image quality, where the improvement becomes greater as the image becomes more complex. The second main effect of this technique is that for a given fixed bit-rate, the image quality is more robust against variations in image complexity. This is accomplished at the encoder via an extended non-linearity interposed between the transformation stage and the quantization stage.

6.1.6 Arbitrary decomposition

An extension providing for arbitrary decomposition of the tile component is described in Annex F. This extension can control the bandpass extent of wavelet sub-bands and thus provide control over the decorrelation process in order to