FINAL
DRAFT

AMENDMENT

ISO/IEC 15444-3 FDAM 3

ISO/IEC JTC **1**

Secretariat: **JISC**

Voting begins on:
2005-01-27

Voting terminates on:
2005-03-27

Information technology — JPEG 2000 image coding system —

Part 3: **Motion JPEG 2000**

AMENDMENT 3: Definition of compliance points and testing for Motion JPEG 2000

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

Partie 3: Motion JPEG 2000

AMENDEMENT 3: Définition des points et des essais de conformité pour Motion JPEG 2000

Please see the administrative notes on page iii

Reference number
ISO/IEC 15444-3:2002/FDAM 3:2005(E)

In accordance with the provisions of Council Resolution 21/1986, this document is **circulated in the English language only**.

Contents

Page

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.

Amendment 3 to ISO/IEC 15444-3:2002 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

Introduction

ITU-T Rec.T.800 | ISO/IEC 15444-1 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. ITU-T Rec.T.802 | ISO/IEC 15444-3 specifies the use of the wavelet-based JPEG 2000 codec for the coding and display of timed sequences of images. The Motion JPEG 2000 file format (MJ2) is designed to contain one or more motion sequences of JPEG 2000 images, with their timing, and also optional audio annotations, all composed into an overall presentation. ITU-T Rec.T.803 | ISO/IEC 15444-4 provides the framework, concepts, and methodology for testing and the criteria to be achieved to claim compliance to JPEG 2000 standard, i.e. to still images only.

This document makes use of the latter framework of conformance testing to apply it to the mentioned motion sequences. The objective of standardization in this field is to promote interoperability between MJ2 encoders and decoders and to test these systems for compliance to these specifications. Compliance testing is the testing of a candidate product for the existence of specific characteristics required by a standard. It involves testing the capabilities of an implementation against both the compliance requirements in the relevant standard and the statement of the implementation's capability.

With this document the framework, concepts, methodology for testing and the criteria to be achieved to claim compliance to ITU-T Recommendation T.802 | ISO/IEC 15444-3 are specified. Whilst extending the specifications of ITU-T Rec.T.803 | ISO/IEC 15444-4 to timed sequences of images, the mind was on widely used applications. This document describes compliance points and testing procedures for Motion JPEG 2000 decoders only.

Information technology — JPEG 2000 image coding system —

Part 3:
Motion JPEG 2000

AMENDMENT 3: Definition of compliance points and testing for Motion JPEG 2000

Add the following reference to clause 2:

ITU-T Rec.T.803 | ISO/IEC 15444-4: *Information technology — JPEG 2000 image coding system: Conformance testing*

Add the following clauses after clause 7:

8 Definition of compliance points

This section describes a number of compliance points (Cpoints) for ITU-T Rec.T.802 | ISO/IEC 15444-3. The points and parameters are described to provide assistance in designing a compliant decoder. Actual compliance is determined by the test methods in clause 10 and the codestreams, reference images, and tolerances in clause 11. The definitions of compliance points in this section are useful for the design of an encoder. The parameters may correspond to particular parts of an implementation.

Because of resource limitations, implementations of Motion JPEG 2000 sometimes will not be able to decode a codestream in its entirety. This section defines various parameters for which a specific implementation might be limited. A set of values for every parameter defines a compliance point. Thus an implementation of a particular Cpoint must guarantee resources as defined in all the parameters.

8.1 H, W, C: Image size guarantees

Decoders may be limited in the size of the output image that they are capable of producing, due to physical display characteristics or memory limitations. H, W, and C are respectively the largest height, width, and number of components that are required to be decoded for a decoder in the compliance point. Codestreams containing more samples than the H, W, and C for a Cpoint shall still be decoded, provided they contain a resolution equal to or less than HxW. Compliance for these codestreams is based on the ability to decode at the largest size smaller than or equal to that specified by the decoder's Cpoint, while preserving aspect ratio. The requested image size is defined by the height and width fields in the applicable 'VisualSampleEntry' from the MJ2 file.

Equation 8.1 and Equation 8.2 express these restrictions. The maximum $r \le T_L$ that satisfies both conditions is the number of levels that must be decoded. T_L is defined in subclause 8.8. The variables w_r and h_r denote requested with and height from the VisualSampleEntry segment defined in subclause 6.1. The term $\min_{\forall i}(N_L(i))$ denotes the minimum number of decomposition levels declared in any COD or COC marker segments whether used in main header or in tile-part headers as defined in Annex A of ITU-T Rec.T.800 | ISO/IEC 15444-1.

If a non-negative r does not exist to satisfy both conditions for any tile or for the whole image, then no decoder obligation exists. A decoder claiming compliance at some Cpoint with image dimensions H x W and number of components C, must also be capable of decoding any sequences with width less than or equal to W, height less than or equal to H, and number of components less than or equal to C. For each Cpoint, the minimum values for H, W, and C are specified in table AMD3-1.

$$\left\lceil \frac{w_r}{2^{\min_{\forall i}(N_L(i))-r}} \right\rceil \leq W \qquad \textbf{8.1}$$

$$\left\lceil \frac{h_r}{2^{\min_{\forall i}(N_L(i))-r}} \right\rceil \leq H \qquad \textbf{8.2}$$

8.2 N_{cb}: Code-block parsing guarantee

Decoders need not decode compressed bits that cannot be recovered from the codestream due to excessive parser memory being required. An upper bound for the parser state memory required to reach a point x in the codestream may be determined from the total number of code-blocks for which state information must be kept, the total number of precincts for which a packet has been encountered, and the total number of components of the codestream.

At position x in the codestream, $N_{cb}(x)$ is defined as the total number of code-blocks in every precinct where the first header byte of at least one received packet for the precinct lies outside the range 0x80 to 0x8F.

Decoders are permitted to stop parsing the codestream at the point, x, once $N_{cb}(x) > N_{cb}$, where N_{cb} is defined for each compliance point. Decoders are permitted to stop parsing the codestream once packet headers with more than N_{cb} code-blocks have been encountered. Code-blocks in packets prior to the packet with the N_{cb}^{th} code-block shall be decoded up to the limits of other parameters in the compliance point.

NOTE — Packets headers with the first bit set to 0 are defined as empty. The above definition adds all the code-blocks associated with such precincts to N_{cb} for these empty packets because a decoder requires more memory for these packets than for packets starting in the listed range.

8.3 N_{comp}: Component parsing guarantee

Decoders could be required to buffer information about each component for many thousands of components just to parse a codestream. To limit the required memory, decoders are permitted to stop parsing the codestream at a point, x, once the following condition is reached:

$C_{max}(x) > N_{comp}$

where $C_{max}(x)$ is defined as the largest component index for which a packet has been encountered up to point x regardless of the emptiness or the relevance of the packet.

Code-blocks in packets prior to the above stop condition shall be decoded up to the limits of other parameters in the compliance point.

8.4 L_{body}: Coded data buffering guarantee

The parser state memory described in subclause 8.2 is required to parse packets regardless of whether their code-blocks are relevant to the dimensions and number of components for which compliance is being claimed. For those code-blocks that are relevant, the implementation is required to store the recovered packet bytes. These are the code bytes that are processed by the block decoder (Annexes C and D, ITU-T Rec.T.800 | ISO/IEC 15444-1).

After a given number of decoded codestream bytes, x, the quantity $L_{body}(x)$ is defined as the total number of packet bytes that have been encountered so far in packets whose precincts are relevant to the dimensions and components for which compliance is being claimed. Although some implementations may be able to decode some of these packet bytes incrementally, L_{body} represents an upper bound on the number of packet bytes that must be stored by the decoder prior to decoding. If the number of relevant packet bytes exceeds L_{body}, then the Implementation Under Test (IUT) is allowed to stop reading the codestream and to decode the code-blocks obtained up to the limits of other parameters in the compliance point.

8.5 M: Decoded Bit-plane guarantee

The decoder shall decode all of the packet bytes recovered by the parser in accordance with the requirements described above. This obligation is limited to the most significant *M* bit-planes of each code-block. Specifically, the block decoder must correctly decode the first $3(M - P_B) - 2$ coding passes, if available, of any relevant code-block, *b*, where P_b is the number of zero-valued most significant bit-planes signaled in the relevant packet header as described in Annex B of ITU-T Rec.T.800 | ISO/IEC 15444-1. The decoder is free to decode any number of additional coding passes for any code-block. Codestreams with large values for the number of guard bits will have a larger number of zero-valued most significant bit planes, and thus a decoder of any given Cpoint will decode fewer useful bit-planes. Likewise, codestreams with large values for the shift in the RGN marker segment may have fewer bit-planes decoded.

8.6 P: 9-7I Precision guarantee

Codestreams that make use of the irreversible 9-7 discrete wavelet transform will require dequantization, the 9-7 inverse discrete wavelet transform, and potentially the inverse irreversible component transform (ICT). The precision values for the wavelet transform are chosen to allow high quality imagery at various bit-depths, e.g. 8, 12, or 16 bits per sample. However, for Cpoint-0, the accuracy of the 9-7I filter required is set such that it is possible to be compliant by decoding and inverse quantizing and performing a 5-3I (irreversible 5-3) inverse wavelet transform. This allows lower cost decoders to be used for the lowest compliance point only. For higher compliance points, using the 5-3 filter in place of the 9-7 filter will not be sufficient to pass the compliance tests.

Using the 5-3 inverse wavelet transform to decode imagery compressed with the 9-7 wavelet introduces signal dependent noise. For example errors are highest around edges in the imagery. Because induced errors are signal dependent, there is no "precision" specified for the implementation of the wavelet transform for Cpoint-0. Instead, the bounds on accuracy of the 9-7 transform have been set for each Cpoint-0 reference image to allow an implementation to use the 5-3I inverse wavelet filter. Using the 5-3I inverse wavelet transform instead of a 9-7I filter does not relieve a decoder of the requirement to perform inverse quantization.

For compliance points other than Cpoint-0, the precision guarantee in Table 1 refers to the implementation's minimum word size that will achieve the target MSE values for the test streams.

To facilitate end-to-end testing for compliance, dequantization may be performed using mid-point rounding. That is, the value of *r* in Equation G.6 of ITU-T Rec.T.800 | ISO/IEC 15444-1 can be $r = 1/2$. Implementations under test may provide the option of using different values for the reconstruction parameter, *r*; however, if the value $r = 1/2$ is supported and employed for compliance testing this will typically increase the ease of passing.

8.7 B: 5-3R Precision guarantee

A decoder is expected to implement the reversible 5-3R IDWT exactly, for component bit-depths of *B* bits/sample or less, as specified in the SIZ marker segment (see Annex A of ITU-T Rec.T.800 | ISO/IEC 15444-1). If a codestream employs the reversible component transform (RCT) and the IUT claims compliance at 3 or more components, it must be able to perform both the 5-3R IDWT and the inverse RCT exactly for bit-depths of *B* bits/sample or less.

8.8 T_L: Transform level guarantee

For each Cpoint, a decoder is expected to be able to synthesize a minimum number of levels of the IDWT, T_L. For codestreams that contains more than T_L decomposition levels, the decoded image from a compliant decoder in a given Cpoint may include only the top resolution levels.

8.9 L: Layer guarantee

For each Cpoint, a decoder is expected to decode a minimum number of layers, *L*, in a codestream. For codestreams that contains more than *L* layers, the decoded image from a compliant decode in a given Cpoint may include only the top *L* layers. This relieves compliant decoders from the burden of decoding inefficient codestreams with an excessive number of layers.

8.10 Progressions

For all Cpoints, a decoder is expected to decode all possible progressions as specified in the COD marker segment. If a POC marker segment is used in a codestream, Cpoint-0 to Cpoint-3 decoders shall decode packets associated with the first progression order specified in the POC marker segment for that tile. Additional packets in the tile may be skipped.

8.11 Tiles

If an image is divided into tiles the following restrictions apply to tile dimensions:

$$128 \le XTsiz / \min(XRsiz^i, YRsiz^i) \le 1024 \quad \textbf{8.3}$$

$$XTsiz = YTsiz \quad \textbf{8.4}$$

8.12 Tile-parts

Codestreams may contain multiple tile-parts for each tile. Profile-0 codestreams require all initial tile-parts to appear in spatial order in the codestream before other tile parts. Cpoint-0 to Cpoint-3 decoders may ignore tile-parts beyond the first even if N_{cb} or L_{body} has not been reached.

8.13 Precincts

Tiles may contain several precincts. Cpoint-0 decoders need only decode the first precinct in each subband of each tile.

8.14 Frame-rate and bit-rate

Frame-rate: A compliant real-time decoder must report the lowest frame rate that it can sustain when decoding all frames, as well as the number of skipped frames (fields) when it achieves real-time.

Bit-rate: A compliant real-time decoder must also report the highest bit-rate which can always be fully decoded in real-time (bit-rate guarantee).

8.15 Profile: codestream guarantee

Profiles provide limits on the codestream syntax parameters. Two profiles are defined in ITU-T Rec.T.802 | ISO/IEC 15444-3, labeled 'unrestricted' and 'simple'. Conformance testing of the rich feature set of unrestricted codestreams is not targeted in this document. Thus, for all Cpoints compliant decoders need only to handle motion representations in MJ2 simple profile, indicated by the brand 'mj2s' in top-level file-type box.

NOTE — Conforming to simple profile means that restrictions of Profile-0 defined in Annex A.10 of ITU-T Rec.T.800 | ISO/IEC 15444-1 apply to embedded codestreams, with the exception of tile dimensions where YTsiz = XTsiz = 128 is replaced by definitions of subclause 8.11

9 Compliance point definitions

Table AMD3-1 defines four compliance points in terms of the parameters.

Table AMD3-1 — Definitions of compliance points (Cpoint) for Part-3

Parameter	**Cpoint-0**	**Cpoint-1**	**Cpoint-2**	**Cpoint-3**
WxH(Size)	360x288	720x576	1920x1080	4096x3112
C(Components)	3	3	4	4
N_{cb}	399	1371	8428	50656
N_{comp}	4	4	4	4
L_{body}	2^{17} bytes	2^{20} bytes	2^{23} bytes	2^{26} bytes
M	11	13	15	19
P	Low enough to allow 5x3 I decoding of 9x7 I data	16 bit fixed point implementation	16-bit fixed point implementation	20 bit fixed point implementation
B	8	10	12	16
T_L	3	4	5	5
L	15	15	15	15
Progressions	For all Cpoints, a decoder is expected to decode all possible progressions as specified in the COD marker segment. If a POC marker segment is used in a codestream, a Cpoint-0 decoder shall decode packets associated with the first progression order specified in the POC marker segment for that tile. Additional packets in the tile may be skipped. For all other Cpoints, packets may be skipped only due to other limitations (e.g. N_{cb} and L_{body}) and there is no explicit limitation on the number of progression order changes that may occur.			
Tiles	Single tile image or square tiles with dimensions ranging from 128 to 1024	Single tile image or square tiles with dimensions ranging from 128 to 1024	Single tile image or square tiles with dimensions ranging from 128 to 1024	Single tile image or square tiles with dimensions ranging from 128 to 1024
Tile-parts	Decode only first tile-part per tile	Decode only first tile-part per tile	Decode only first tile-part per tile	Decode only first tile-part per tile
Precincts	Decode first precinct per sub-band	Decode all precincts	Decode all precincts	Decode all precincts
File format	MJ2 simple profile	MJ2 simple profile	MJ2 simple profile	MJ2 simple profile

10 Definition of test methods

Compliance testing procedures apply as defined in Annex B of ITU-T Rec.T.803 | ISO/IEC 15444-4, with following extensions:

1. A particular executable test suite (ETS) defines the test codestreams (TCS), output images and error tolerances. This is done in clause 11 for the four defined compliance points by taking specified frames from MJ2 sequences. Implementations under test (IUT) must therefore be able to output decoded visual samples in a format (see Annex G) that they can be compared as defined in ITU-T Rec.T.803 | ISO/IEC 15444-4. Any visual composition transformations do not apply to this test method.

2. In addition to the testing of single visual samples IUT will be evaluated by using MJ2 reference sequences defined in clause 11. The test procedure is: