INTERNATIONAL STANDARD

ISO/IEC
15444-3

First edition
2002-09-01

AMENDMENT 2
2003-12-01

Information technology — JPEG 2000 image coding system —

Part 3:
Motion JPEG 2000

AMENDMENT 2: Motion JPEG 2000 derived from ISO base media file format

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

Partie 3: Motion JPEG 2000

AMENDEMENT 2: Motion JPEG 2000 dérivée du format de fichier de média de base de l'ISO

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO/IEC 15444-3:2002/Amd 2:2003
https://standards.iteh.ai/catalog/standards/sist/a37ab345-f6a3-4abe-8f27-
25d66715601b/iso-iec-15444-3-2002-amd-2-2003

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 2 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

This document is the consequent revision of the Motion JPEG 2000 specification, based on the common text of the MP4 and MJ2 formats, which is called the ISO Base Media File Format; it will be consolidated subsequently into a new edition of ISO/IEC 15444-3:2002.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO/IEC 15444-3:2002/Amd 2:2003
https://standards.iteh.ai/catalog/standards/sist/a37ab345-f6a3-4abe-8f27-
25d66715601b/iso-iec-15444-3-2002-amd-2-2003

iTeh STANDARD PREVIEW
(standards.iteh.ai)

Information technology — JPEG 2000 image coding system —

Part 3:
Motion JPEG 2000

AMENDMENT 2: Motion JPEG 2000 derived from ISO base media file format

1 Scope

This Amendment specifies the use of the wavelet-based JPEG 2000 codec for the coding and display of timed sequences of images (motion sequences), possibly combined with audio, and composed into an overall presentation. In this specification, a file format is defined, and guidelines for the use of the JPEG 2000 codec for motion sequences are supplied.

iTeh STANDARD PREVIEW

2 Normative references (standards.iteh.ai)

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

ISO 639-2:1998, *Codes for the representation of names of languages — Part 2: Alpha-3 code*

ISO/IEC 14496-1:2001, *Information technology — Coding of audio-visual objects — Part 1: Systems* [particularly the syntax description language (SDL), clause 14]

ISO/IEC 15444-12, *Information technology — JPEG 2000 image coding system — Part 12: ISO base media file format* (technically identical to ISO/IEC 14496-12)

ITU-T Rec.T.800 | ISO/IEC 15444-1, *Information technology — JPEG 2000 image coding system: Core coding system*

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

3.1
motion sequence
timed sequence of JPEG 2000 images

4 Compatibility and Technology derivation

4.1 Family Members

This is a stand-alone specification; it defines the file format for MJ2. However, it stands as a member of a family of specifications with common formatting.

The other family members include:

- The JPEG 2000 single image format, JP2.

- The ISO Base Media File Format, on which the MP4 format is based.

- The QuickTime file format, on which the ISO Base Media format is based.

These specifications share a common definition for the structure of a file (a sequence of objects, called boxes here and atoms in QuickTime), and a common definition of the general structure of an object (the size and type).

All these specifications require that readers ignore objects that are unrecognizable to them.

This specification takes precedence over those from which it inherits, in any case where there are differences or conflicts; however no such conflicts are known to exist.

4.2 ISO Media file Inheritance and Compatibility

The Motion JPEG 2000 file format is defined as derived from the ISO Base Media file format. Notwithstanding anything in that base specification, hint tracks are not a normative part of this specification. Sections 7 (Streaming Support) and 10 (RTP Hint Track Format), though compatible with this specification, do not form normative parts of this specification. They may be used as a compatible, optional, extension, but are not required for compatibility with this specification. There may be license implications in the use of this or other compatible extensions to this format.

4.3 JP2 Inheritance and Compatibility

The still image format, JP2, defines a number of boxes. The signature box from that specification shall be present. If the JP2 specification requires a particular position (e.g. first in the file), that positioning shall be followed here.

It is permissible under this specification to make a file that adheres to both this specification and the JP2 specification. In that case:

1) The compatibility list shall include all the compatible brands.

2) The objects (boxes or atoms) required by the JP2 specification shall also be present.

3) The objects (boxes or atoms) optional in the JP2 specification may also be present.

A still image reader, reading a file which contains both a presentation (conformant to this specification) and a still image, would 'see' only the still image. Likewise a motion reader would 'see' only the presentation. A more powerful reader may display both, or offer the user a choice.

The JP2 specification includes an optional IPR (Intellectual Property Rights) box which is therefore also optional in this specification. Among other issues this addresses unique identification and protection of content.

4.4 Conformance

Implementations of Motion JPEG 2000 decoders shall support JPEG 2000 image sequences, as well as raw and twos-complement audio if audio output is available. They may also support compressed audio, using MP4 formats, or other track types from MPEG-4. The support of such MPEG-4 tracks is not required; however,

readers shall not fail if they are present. If MPEG-4 composition (BIFS) is used, then the simple composition used in this specification should also be set up in such a way that a reader not implementing BIFS will display a suitable result.

Files conformant with this specification shall contain at least one Motion JPEG 2000 video track. They may contain more video tracks, uncompressed audio, or compressed MP4 audio.

4.5 Profiles and Levels

There are two tools for profiling Motion JPEG 2000 files.

The first consists of the optional specification of tools and levels of the JPEG 2000 coding system (codestream features). These are indicated in the optional sample description extension JP2 Profile Box (see below 5).

The second tool allows a file overall to be identified as belonging to a definition which forms a proper subset of the general specification. Such definitions might restrict such features as:

- the use of data references, and multiple files;

- the layout order of the boxes, and the data within the boxes (e.g. that data is in time order and interleaved);

- the use of profiles of the JPEG 2000 codestream;

- the existence of other tracks, and their format (e.g. audio, MPEG-7, etc.).

The conformance to these restricted profiles is indicated in the file type box by the addition of the compatible profiles as brands within the compatibility list. Annex A defines the available profiles in this specification.

4.6 Visual Composition

Composition of multiple image sequences in a 2D environment can be achieved by using multiple video tracks which overlap in time. Their composition is defined by the following structures:

- The matrix in the track header specifies their positioning and scaling.

- The layer field in the track header specifies the front-to-back ordering of the tracks.

- The graphics mode and opcolor fields in the video media header are used to specify the ways in which each track is composited onto the existing image (this compositing is performed from back to front).

Applications requiring more complex compositing may use the BIFS system from MPEG-4, optionally. The matrix, graphics mode, and layers should be setup so that a reader not implementing BIFS displays the desired result. Matrix values which occur in the headers specify a transformation of video images for presentation. The point (p,q) is transformed into (p', q') using the matrix as follows:

```
(p q 1) *    | a   b   u |    = (m n z)
             | c   d   v |
             | x   y   w |

m = ap + cq + x;  n = bp + dq + y;  z = up + vq + w;

p' = m/z;  q' = n/z
```

The coordinates {p,q} are on the decompressed frame, and {p', q'} are at the rendering output. Therefore, for example, the matrix {2,0,0, 0,2,0, 0,0,1} exactly doubles the pixel dimension of an image. The co-ordinates transformed by the matrix are not normalized in any way, and represent actual sample locations. Therefore {x,y} can, for example, be considered a translation vector for the image.

The co-ordinate origin is located at the upper left corner, and X values increase to the right, and Y values increase downwards. {p,q} and {p',q'} are to be taken as absolute pixel locations relative to the upper left hand corner of the original image (after scaling to the size determined by the track header's width and height) and the transformed (rendering) surface, respectively.

Each track is composed using its matrix as specified into an overall image; this is then transformed and composed according to the matrix at the movie level in the MovieHeaderBox. It is application-dependent whether the resulting image is 'clipped' to eliminate pixels, which have no display, to a vertical rectangular region within a window, for example. So for example, if only one video track is displayed and it has a translation to {20,30}, and a unity matrix is in the MovieHeaderBox, an application may choose not to display the empty "L" shaped region between the image and the origin.

All the values in a matrix are stored as 16.16 fixed-point values, except for u, v and w, which are stored as 2.30 fixed-point values. For upwards compatibility into the MPEG-4 BIFS (scene composition) system, matrices used here restrict (u,v,w) to be (0,0,1), for which the hex values are (0,0,0x40000000). This permits the simple composition used here to be mapped into BIFS if a scene later requires full scene management.

The values in the matrix are stored in the order {a,b,u, c,d,v, x,y,w}.

Tracks are composed to the presentation surface from back (highest layer number) to front (lowest layer number), against an indeterminate initial colour. There are various composition modes available; the backmost (first-rendered) track would normally use 'copy' as the initial image is indeterminate. Subsequent layers can then be composed on top in a variety of ways. The following table details the composition modes available. Note that (currently) only the 'transparent' mode uses the opcolor field.

Table 1 — Graphics Composition Modes

Mode	Code	Description
Copy	0x0	Copy the source image over the destination
Transparent	0x24	Replace the destination pixel with the source pixel if the source pixel isn't equal to the opcolor. (Also known as 'blue-screen').
Alpha	0x100	Replace the destination pixel with a blend of the source and destination pixels, with the proportion controlled by the alpha channel. The alpha channel is applied to all channels.
Pre-multiplied black alpha	0x102	Pre-multiplied with black means that the colour components of each pixel have already been blended with a black pixel, based on their alpha channel value. Effectively, this means that the image has already been combined with a black background, which must be removed before composition.
Component alpha	0x110	One or more alpha channels are present, which are applied to individual colour channels, and the image must be composed channel-by-channel

Images are only alpha-composed if both the graphics composition mode requests alpha composition, and the images contain alpha channels, as declared by the Channel Definition Box inside the JP2 Header Box. Therefore the graphics mode can be used to prevent alpha composition of an image with alpha channels, if that is desired.

If there is a single alpha channel applied to the entire image, then the value of the graphics must be 'Alpha' if that channel is a straight 'Opacity' channel, and must be 'Pre-multiplied black alpha' if that channel is a 'Pre-multiplied' opacity channel. If there are one or more alpha channels in the image which are applied to individual channels and not to the whole image, and alpha composition is desired, then the 'Component alpha'

value must be used for the graphics mode. Support of 'Component alpha' composition is optional in Part 3 of this specification.

The alpha blending formulas are defined in Part 1 of this specification.

NOTE   Use of the "transparent" opcode may be yield unexpected results when the image codestreams are compressed in a non-reversible fashion, or are subject to scaling in quality or resolution, either during or after content production. Such operations are not guaranteed to preserve individual sample values precisely.

4.7   Box Order

All JPEG 2000 files start with a signature box, therefore this rule is added to the Box Order rules:

   1)   The JP2 Signature Box and File Type Box **shall** occur first and second in the file.

5   File Identification

In the file type compatibility box, the brand shall be 'mjp2' for files conforming to this specification, and 'mjp2' shall be a member of the compatibility list.

See Annex A for a complete list of all profile brand names.

The preferred file extension is '.mj2'. The MIME type video/mjp2 is used, as defined in the appropriate RFC.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

6   Required Additions

6.1   Sample Description Box

6.1.1   Definition

Box Types: 'mjp2', 'raw ', 'twos'
Container: Sample Table Box ('stbl')
Mandatory: Yes
Quantity: Exactly one

A Motion JPEG 2000 visual sample entry shall contain a JP2 Header Box from the JPEG 2000 Part 1 specification; however, the sRGB YCC enumerated colour space (codepoint 18) from Part 2 may also be used to identify the colour space used, in addition to the Part 1 enumerated colour spaces (such as sRGB and greyscale). If the JP2 Header Box indicates the presence of alpha channels, then the 'depth' field in the VisualSampleEntry must also indicate their presence, with the value 0x20. Similarly if the JP2 Header Box defines a monochrome image with no alpha, the 'depth' field must contain the value indicating grayscale (0x28). Otherwise, the 'depth' field must declare colour images (0x18).

If two fields are present in the samples, the JP2 Header Box applies to the complete image, not to each field individually. Therefore the height as declared in the JP2 Header Box and the VisualSampleEntry applies to the entire de-interlaced image.

NOTE   This means that you may not be able to construct a legitimate JP2 file by composing a single field's codestream together with the JP2 header box found in the sample description.

The sample format for Motion JPEG 2000 data is a set of boxes. Currently this specification permits only JP2 Codestream Boxes ('jp2c') as defined in the JP2 specification. If there is no Field Coding Box present, or the field count is 1, the sample shall contain precisely one codestream box. If the field count is 2 then there shall be two codestream boxes. Other boxes, if present in the sample, shall be ignored. The last (or only) box in the sample may have a value of 0 for its length field, indicating it extends to the end of the sample, as indicated by the sample size given in the sample size table. The actual codestreams presented to the decoder are formed