INTERNATIONAL STANDARD

ISO/IEC 10918-5

First edition
2013-05-01

Information technology — Digital compression and coding of continuous-tone still images: JPEG File Interchange Format (JFIF)

Technologies de l'information — Compression numérique et codage des images fixes à modelé continu: Format d'échange de fichiers JPEG (JFIF)

Reference number
ISO/IEC 10918-5:2013(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

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 10918-5 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 T.871 (05/2011).

ISO/IEC 10918 consists of the following parts, under the general title *Information technology — Digital compression and coding of continuous-tone still images*:

— *Part 1: Requirements and guidelines*

— *Part 2: Compliance testing*

— *Part 3: Extensions*

— *Part 4: Registration of JPEG profiles, SPIFF profiles, SPIFF tags, SPIFF colour spaces, APPn markers, SPIFF compression types and Registration Authorities (REGAUT)*

— *Part 5: JPEG File Interchange Format (JFIF)*

— *Part 6: Application to printing systems*

Table of Contents

INTERNATIONAL STANDARD
RECOMMENDATION ITU-T

Information technology – Digital compression and coding of continuous-tone still images: JPEG File Interchange Format (JFIF)

1 Scope

This Recommendation | International Standard specifies the JPEG File Interchange Format (JFIF).

The JPEG File Interchange Format (JFIF) is a minimal file format which enables the exchange of JPEG encoded images (according to Rec. ITU-T T.81 | ISO/IEC 10918-1) having 1 or 3 colour channels and 8 bits per colour channel between a wide variety of platforms and applications. This minimal format does not include some advanced features found in various other specified file formats. The purpose of this format is to provide for a basic form of exchange of JPEG images. The optional inclusion of thumbnail images for rapid browsing is also supported.

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.

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*.

2.2 Paired Recommendations | International Standards equivalent in technical content

None.

2.3 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*.

3 Definitions

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

3.1 JPEG File Interchange Format (JFIF): The interchange format specified in this Recommendation | International Standard for exchange of images encoded according to the JPEG standard (Rec. ITU-T T.81 | ISO/IEC 10918-1) having 1 or 3 colour channels and 8 bits per colour channel.

3.2 thumbnail: Reduced resolution representation of the main JPEG (Rec. ITU-T T.81 | ISO/IEC 10918-1) coded image that can be used to identify the image by its content.

NOTE – Thumbnails are commonly used to browse multiple images quickly using a low resolution visual representation of the images, rather than using file names or other metadata.

4 Abbreviations

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

Ap_i Byte i of application data

APP_0 Application data marker, type 0

EOI	End of Image
ICC	International Color Consortium
JFIF	JPEG File Interchange Format
JPEG	Joint Photographic Experts Group
Lf	Length of frame header
Lp	Length of application data segment
Nf	Number of components in frame
RGB	Red, Green, and Blue (colour component values)
SOF	Start of Frame
SOI	Start of Image
YC_BC_R	Luminance (denoted as Y), Chrominance toward Blue, and Chrominance toward Red (colour component values)

5 Conformance

Some requirements in this Recommendation | International Standard are expressed as format or syntax requirements rather than as software or hardware implementation requirements. Implementations fall into two categories: JFIF decoders and JFIF encoders.

In order for a JFIF decoder to be considered conforming, the decoder shall not report errors when processing conforming instances of the specified format, except when forced to do so by resource exhaustion or when the encoded image data uses non-baseline features of Rec. ITU-T T.81 | ISO/IEC 10918-1 that are not supported by the decoder.

NOTE – The decoder should report errors when processing non-conforming deviations from the specified format.

In order for a JFIF encoder to be considered conforming, the files produced by the encoder shall be formatted as specified.

6 JPEG File Interchange Format (JFIF) overview

6.1 JPEG compression

Any JPEG (Rec. ITU-T T.81 | ISO/IEC 10918-1) process is supported by the syntax of the JPEG File Interchange Format (JFIF), provided the encoded image has 1 or 3 colour channels and 8 bits per colour channel.

NOTE 1 – It is strongly recommended that the JPEG baseline process, as defined in Rec. ITU-T T.81 | ISO/IEC 10918-1, be used for the purposes of file interchange. This ensures maximum compatibility among applications supporting JPEG coded images.

Files conforming to the JPEG File Interchange Format shall conform to the interchange format specified in Rec. ITU-T T.81 | ISO/IEC 10918-1. The encoded image in the JPEG File Interchange Format shall have 1 or 3 colour channels and 8 bits per colour channel. Additionally, the JFIF APP_0 marker (see 6.4) shall be present immediately following the SOI marker (specified in Rec. ITU-T T.81 | ISO/IEC 10918-1).

NOTE 2 – The interchange format specified in Rec. ITU-T T.81 | ISO/IEC 10918-1 requires that all table specifications used in the encoding process are coded in the image data prior to their use.

6.2 Colour space

The colour space to be used is YC_BC_R as defined by Rec. ITU-R BT.601 (256 levels) but with a different scaling as specified below. If only one component is used, that component shall be the Y component channel as specified below.

NOTE – The colour space specification herein can provide only a basic level of colour fidelity. The use of supplemental metadata such as an ICC profile (e.g., as specified in ISO 15076-1) may be necessary to provide a more accurate colour characterization.

If three components are used, they shall be present in the image with the ordering of the components such that the first component is the Y channel, the second component is the C_B channel, and the third component is the C_R channel.

6.3 JFIF APP_0 marker segment

The JFIF APP_0 marker segment shall immediately follow the SOI marker. The JFIF APP_0 marker segment is defined as an APP_0 marker (specified in Rec. ITU-T T.81 | ISO/IEC 10918-1) containing the null-terminated string: "JFIF" encoded as specified in 10.1 in the first five application data bytes of the marker segment (Ap_i, for i = 1 to 5). Additional APP_0 marker segments may also be present, provided the associated application data bytes do not begin with this string. The JFIF APP_0 marker segment provides some information that is not contained in the JPEG (Rec. ITU-T T.81 | ISO/IEC 10918-1) stream, such as: version number, horizontal and vertical pixel densities (expressed in dots per inch or dots per cm), pixel aspect ratio (derived from the horizontal and vertical pixel densities), and an optional thumbnail encoded as 24-bit *RGB* image data.

NOTE – The "JFIF" string as specified in 10.1 is encoded according to Rec. ITU-T T.50 | ISO 646.

6.4 APP_0 marker used to specify JFIF extensions

Additional APP_0 marker segment(s) can optionally be used for the JFIF extensions specified in clause 10. If used, these segment(s) must immediately follow the JFIF APP_0 marker segment. Decoders shall skip any unsupported JFIF extension segments and continue decoding. The JFIF extension APP_0 marker is identified by the zero-terminated string "JFXX" encoded as specified in 10.2 in the first five application data bytes of the marker segment (Ap_i, for i = 1 to 5). The JFIF extension APP_0 marker segment contains a 1-byte code with a value specified in clause 10 that identifies the particular extension type.

NOTE – The encoding of the "JFXX" string as specified in 10.2 is encoded according to Rec. ITU-T T.50 | ISO 646.

6.5 Application marker segments used for application-specific information

Additional APP_0 and other application marker segments may be used to hold application-specific information that does not affect the ability to decode or display the JFIF file segment. Application-specific marker segments must appear after the JFIF APP_0 and any JFIF extension APP_0 marker segments. Application-specific APP_0 marker segments shall contain application data bytes such that the initial application data bytes contain a zero-terminated string value that identifies the application. For application-specific APP_0 marker segments, this string shall not be the zero-terminated string values "JFIF" as specified in 10.1 or "JFXX" as specified in 10.2, to avoid conflict with this Recommendation | International Standard. It is recommended that this string represent an organization name or company trademark. Generic strings such as "dog", "cat", "tree", etc., should not be used.

7 Conversion to and from *RGB*

The interpretations of Y, C_B, and C_R are derived from the E'_Y, E'_{C_B}, and E'_{C_R} signals defined in the 625-line specification of Rec. ITU-R BT.601, but these signals are normalized so as to permit the usage of the full range of 256 levels of the 8-bit binary encoding of the Y component. More precisely, they are specified by the following relationships:

$$Y = \text{Min}(\text{Max}(0, \text{Round}(255 * E'_Y)), 255)$$

$$C_B = \text{Min}(\text{Max}(0, \text{Round}(255 * E'_{C_B} + 128)), 255)$$

$$C_R = \text{Min}(\text{Max}(0, \text{Round}(255 * E'_{C_R} + 128)), 255)$$

using the following mathematical definitions:

$$\text{Round}(x) = \lfloor x + 0.5 \rfloor$$

$$\text{Min}(x, y) = \begin{cases} x & ; \quad x \le y \\ y & ; \quad x > y \end{cases}$$

$$\text{Max}(x, y) = \begin{cases} x & ; \quad x \ge y \\ y & ; \quad x < y \end{cases}$$

where the E'_Y, E'_{C_B} and E'_{C_R} are defined as in Rec. ITU-R BT.601. Values of E'_Y have a nominal range of 0.0 to 1.0 and those for E'_{C_B} and E'_{C_R} have a nominal range of –0.5 to +0.5 for conventional colorimetry. The values of Y, C_B, and C_R must be clamped to the range from 0 to 255 as shown above. YC_BC_R colours (with 256 levels per component) can alternatively be computed directly from full scale 8-bit per colour channel *RGB* colours in which black is represented by (0, 0, 0) and white is represented by (255, 255, 255) by using the following formulae:

$$
\begin{aligned}
Y &= \mathrm{Min}(\mathrm{Max}(0, \mathrm{Round}(\quad 0.299 * R + 0.587 * G + 0.114 * B \quad)), 255)\\
C_B &= \mathrm{Min}(\mathrm{Max}(0, \mathrm{Round}((-0.299 * R - 0.587 * G + 0.886 * B) / 1.772 + 128)), 255)\\
C_R &= \mathrm{Min}(\mathrm{Max}(0, \mathrm{Round}((0.701 * R - 0.587 * G - 0.114 * B) / 1.402 + 128)), 255)
\end{aligned}
$$

which, to four decimal position accuracy, can be approximated by:

$$
\begin{aligned}
Y &= \mathrm{Min}(\mathrm{Max}(0, \mathrm{Round}(0.299 * R + 0.587 * G + 0.114 * B)), 255)\\
C_B &= \mathrm{Min}(\mathrm{Max}(0, \mathrm{Round}(-0.1687 * R - 0.3313 * G + 0.5 * B + 128)), 255)\\
C_R &= \mathrm{Min}(\mathrm{Max}(0, \mathrm{Round}(0.5 * R - 0.4187 * G - 0.0813 * B + 128)), 255)
\end{aligned}
$$

NOTE 1 – Not all image file formats store image samples in the order R_0, G_0, B_0, ... R_n, G_n, B_n. The sample order must be verified before converting an *RGB* file to JFIF.

The inverse relationship for computing full scale 8-bit per colour channel gamma pre-corrected *RGB* values (following Rec. ITU-R BT.601 gamma pre-correction and colour primary specifications) from YC_BC_R colours (with 256 levels per component) can be computed as follows:

$$
\begin{aligned}
R &= \mathrm{Min}(\mathrm{Max}(0, \mathrm{Round}(Y + 1.402 * (C_R - 128))), 255)\\
G &= \mathrm{Min}(\mathrm{Max}(0, \mathrm{Round}(Y - (0.114 * 1.772 * (C_B - 128) + 0.299 * 1.402 * (C_R - 128)) / 0.587)), 255)\\
B &= \mathrm{Min}(\mathrm{Max}(0, \mathrm{Round}(Y + 1.772 * (C_B - 128))), 255)
\end{aligned}
$$

which, to four decimal position accuracy, can be approximated by:

$$
\begin{aligned}
R &= \mathrm{Min}(\mathrm{Max}(0, \mathrm{Round}(Y + 1.402 * (C_R - 128))), 255)\\
G &= \mathrm{Min}(\mathrm{Max}(0, \mathrm{Round}(Y - 0.3441 * (C_B - 128) - 0.7141 * (C_R - 128))), 255)\\
B &= \mathrm{Min}(\mathrm{Max}(0, \mathrm{Round}(Y + 1.772 * (C_B - 128))), 255)
\end{aligned}
$$

The *RGB* values used for representation of thumbnails stored in JFIF files as specified in clauses 10.1, 10.4, or 10.5 are interpreted according to this full-scale (256 levels per component) convention, in which black is represented by (0, 0, 0) and white is represented by (255, 255, 255).

NOTE 2 – The definition of 8-bit per colour channel *RGB* values in digital form as used in these equations differs from the one used in Rec. ITU-R BT.601, as the definition used here assumes full-scale 256-level usage in which black is represented by (0, 0, 0) and white is represented by (255, 255, 255), whereas in Rec. ITU-R BT.601, nominal *RGB* black is represented by (16, 16, 16) and nominal white is represented by (235, 235, 235).

NOTE 3 – As this Recommendation | International Standard is based on the prior informally-circulated JFIF version 1.02 specification that was produced in 1992, which referenced Rec. ITU-R BT.601 (formerly CCIR 601), it references that specification for definition of the E'_Y, E'_{C_B}, and E'_{C_R} signals that correspond to the YC_BC_R values specified herein. However, since the development of the prior JFIF version 1.02 specification, additional industry specifications have been developed, Rec. ITU-R BT.601 has been updated, and common industry practice has emerged which often follows the sYCC specification in IEC 61966-2-1/Amd.1. The difference between the use of the colour interpretation specification in this Recommendation | International Standard and that of the sYCC specification may be considered negligible in practice. Moreover, as previously noted, the colour space specification herein can provide only a basic level of colour fidelity. The use of supplemental metadata such as an ICC profile (e.g., as specified in ISO 15076-1) may be necessary to provide a more accurate colour characterization.

8 Image orientation

In JFIF files, the image orientation is always top-down (in terms of human viewing intent). This means that the first image samples encoded in a JFIF file are located in the upper left hand corner of the image and encoding proceeds from left to right and top to bottom. Top-down orientation is used for both full resolution images and thumbnails. The process of converting an image file having bottom-up orientation to JFIF should include inverting the order of all image lines before Rec. ITU-T T.81 | ISO/IEC 10918-1 encoding.

9 Spatial relationship of components

Specification of the spatial positioning of pixel samples within components relative to the samples of other components is necessary for proper image post processing and accurate image presentation. In JFIF files, the position of the pixels in sub-sampled components is defined with respect to the highest resolution component. Since components must be sampled orthogonally (along rows and columns), the spatial position of the samples in a given sub-sampled component may be determined by specifying the horizontal and vertical offsets of the first sample, i.e., the sample in the upper left corner, with respect to the highest-resolution component.

The horizontal and vertical offsets of the position of the first sample in a sub-sampled component, $Hoffset_i[0,0]$ and $Voffset_i[0,0]$, relative to the position of the upper left sample for the largest component, are defined as follows:

$$Hoffset_i[0,0] = (Nsamples_{ref} / Nsamples_i) / 2 - 0.5$$

$$Voffset_i[0,0] = (Nlines_{ref} / Nlines_i) / 2 - 0.5$$

where:

$Nsamples_{ref}$ is the number of samples per line in the largest component;

$Nsamples_i$ is the number of samples per line in the i-*th* component;

$Nlines_{ref}$ is the number of lines in the largest component;

$Nlines_i$ is the number of lines in the i-*th* component.

As used here, the division operator "/" produces a real-valued result without truncation or rounding.

NOTE 1 – Proper sub-sampling of components incorporates an anti-aliasing filter, which reduces the spectral bandwidth of a full resolution component prior to sub-sampling of that component. Sub-sampling can easily be accomplished using a symmetrical digital filter with an even number of taps (coefficients). A commonly used filter for 2:1 sub sampling utilizes two taps (1/2, 1/2).

As an example, consider a three-component image which is comprised of components having the following dimensions:

– Component 1: 256 samples, 288 lines;

– Component 2: 128 samples, 144 lines;

– Component 3: 64 samples, 96 lines.

For a JFIF file, the centres of the samples of such an image would be positioned as illustrated below in Figure 1:

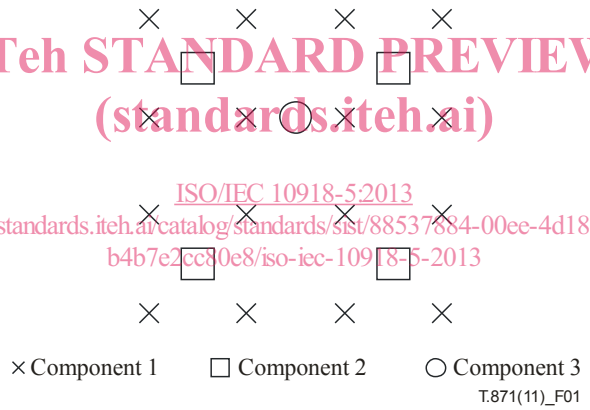

Figure 1 – Centres of the samples of three components

NOTE 2 – While this definition matches some industry specifications, it differs from the convention used for C_B and C_R colour component sub-sampling in Rec. ITU-R BT.601 and a number of other digital video formats. In practice, the difference may be considered negligible, or pre-processing of the chrominance components may be performed to produce a more accurate reconstruction of the compressed image.

NOTE 3 – In common industry usage, only two forms of colour component sub-sampling are typically encountered in practice for three-component images. The most common of these forms is known as 4:2:0, in which the C_B and C_R colour components are sub-sampled by a factor of two in both the horizontal and vertical dimensions. The other relatively-common form is known as 4:2:2, in which the C_B and C_R colour components are sub-sampled by a factor of two in the horizontal dimension only. The 4:2:2 form of sub-sampling is used primarily for video-related applications (especially with the use of interlaced-scan systems). For JFIF usage, the use of sub-sampling formats other than 4:2:0 is therefore discouraged, as such other formats may not be supported in some applications.