8 = The 0th row represents the visual left-hand side of the image, and the 0th column represents the visual bottom.

Default is 1.

Support for orientations other than 1 is not a Baseline TIFF requirement.

PhotometricInterpretation

The color space of the image data.

Tag = 262 (106.H)

Type = SHORT

N = 1

0 = WhiteIsZero. For bilevel and grayscale images: 0 is imaged as white. 2**BitsPerSample-1 is imaged as black. This is the normal value for Compression=2.

1 = BlackIsZero. For bilevel and grayscale images: 0 is imaged as black. 2**BitsPerSample-1 is imaged as white. If this value is specified for Compression=2, the image should display and print reversed.

2 = RGB. In the RGB model, a color is described as a combination of the three primary colors of light (red, green, and blue) in particular concentrations. For each of the three components, 0 represents minimum intensity, and 2**BitsPerSample - 1 represents maximum intensity. Thus an RGB value of (0,0,0) represents black, and (255,255,255) represents white, assuming 8-bit components. For PlanarConfiguration = 1, the components are stored in the indicated order: first Red, then Green, then Blue. For PlanarConfiguration = 2, the StripOffsets for the component planes are stored in the indicated order: first the Red component plane StripOffsets, then the Green plane StripOffsets, then the Blue plane StripOffsets.

3= Palette color. In this model, a color is described with a single component. The value of the component is used as an index into the red, green and blue curves in the ColorMap field to retrieve an RGB triplet that defines the color. When PhotometricInterpretation=3 is used, ColorMap must be present and SamplesPerPixel must be 1.

4 = Transparency Mask.

This means that the image is used to define an irregularly shaped region of another image in the same TIFF file. SamplesPerPixel and BitsPerSample must be 1. PackBits compression is recommended. The 1-bits define the interior of the region; the 0-bits define the exterior of the region.

A reader application can use the mask to determine which parts of the image to display. Main image pixels that correspond to 1-bits in the transparency mask are imaged to the screen or printer, but main image pixels that correspond to 0-bits in the mask are not displayed or printed.

The image mask is typically at a higher resolution than the main image, if the main image is grayscale or color so that the edges can be sharp.

There is no default for PhotometricInterpretation, *and it is required*. Do not rely on applications defaulting to what you want.

PlanarConfiguration

How the components of each pixel are stored.

Tag = 284 (11C.H)

Type = SHORT

N = 1

1 = *Chunky* format. The component values for each pixel are stored contiguously. The order of the components within the pixel is specified by PhotometricInterpretation. For example, for RGB data, the data is stored as RGBRGBRGB…

2 = *Planar* format. The components are stored in separate "component planes." The values in StripOffsets and StripByteCounts are then arranged as a 2-dimensional array, with SamplesPerPixel rows and StripsPerImage columns. (All of the columns for row 0 are stored first, followed by the columns of row 1, and so on.) PhotometricInterpretation describes the type of data stored in each component plane. For example, RGB data is stored with the Red components in one component plane, the Green in another, and the Blue in another.

PlanarConfiguration=2 is not currently in widespread use and it is not recommended for general interchange. It is used as an extension and Baseline TIFF readers are not required to support it.

If SamplesPerPixel is 1, PlanarConfiguration is irrelevant, and need not be included.

If a row interleave effect is desired, a writer might write out the data as PlanarConfiguration=2—separate sample planes—but break up the planes into multiple strips (one row per strip, perhaps) and interleave the strips.

Default is 1. See also BitsPerSample, SamplesPerPixel.

ResolutionUnit

The unit of measurement for XResolution and YResolution.

Tag = 296 (128.H)

Type = SHORT

N = 1

To be used with XResolution and YResolution.

1 = No absolute unit of measurement. Used for images that may have a non-square aspect ratio, but no meaningful absolute dimensions.

The drawback of ResolutionUnit=1 is that different applications will import the image at different sizes. Even if the decision is arbitrary, it might be better to use dots per inch or dots per centimeter, and to pick XResolution and YResolution so that the aspect ratio is correct and the maximum dimension of the image is about four inches (the "four" is arbitrary.)

2 = Inch.

3 = Centimeter.

Default is 2.

RowsPerStrip

The number of rows per strip.

Tag = 278 (116.H)

Type = SHORT or LONG

N = 1

TIFF image data is organized into strips for faster random access and efficient I/O buffering.

RowsPerStrip and ImageLength together tell us the number of strips in the entire image. The equation is:

StripsPerImage = floor ((ImageLength + RowsPerStrip - 1) / RowsPerStrip).

StripsPerImage is *not* a field. It is merely a value that a TIFF reader will want to compute because it specifies the number of StripOffsets and StripByteCounts for the image.

Note that either SHORT or LONG values can be used to specify RowsPerStrip. SHORT values may be used for small TIFF files. It should be noted, however, that earlier TIFF specification revisions required LONG values and that some software may not accept SHORT values.

The default is 2**32 - 1, which is effectively infinity. That is, the entire image is one strip.

Use of a single strip is not recommended. Choose RowsPerStrip such that each strip is about 8K bytes, even if the data is not compressed, since it makes buffering simpler for readers. The "8K" value is fairly arbitrary, but seems to work well.

See also ImageLength, StripOffsets, StripByteCounts, TileWidth, TileLength, TileOffsets, TileByteCounts.

SamplesPerPixel

The number of components per pixel.

Tag = 277 (115.H)

Type = SHORT

N = 1

SamplesPerPixel is *usually* 1 for bilevel, grayscale, and palette-color images. SamplesPerPixel is *usually* 3 for RGB images.

Default = 1. See also BitsPerSample, PhotometricInterpretation, *ExtraSamples*.

Software

Name and version number of the software package(s) used to create the image.

Tag = 305 (131.H)

Type = ASCII

See also Make, Model.

StripByteCounts

For each strip, the number of bytes in the strip after compression.

Tag = 279 (117.H)

Type = SHORT or LONG

N = StripsPerImage for PlanarConfiguration equal to 1.

= SamplesPerPixel * StripsPerImage for PlanarConfiguration equal to 2

This tag is required for Baseline TIFF files.

No default.

See also StripOffsets, RowsPerStrip, TileOffsets, TileByteCounts.

StripOffsets

For each strip, the byte offset of that strip.

Tag = 273 (111.H)

Type = SHORT or LONG

N = StripsPerImage for PlanarConfiguration equal to 1.

= SamplesPerPixel * StripsPerImage for PlanarConfiguration equal to 2

The offset is specified with respect to the beginning of the TIFF file. Note that this implies that each strip has a location independent of the locations of other strips. This feature may be useful for editing applications. This required field is the only way for a reader to find the image data. *(Unless TileOffsets is used; see TileOffsets.)*

Note that either SHORT or LONG values may be used to specify the strip offsets. SHORT values may be used for small TIFF files. It should be noted, however, that earlier TIFF specifications required LONG strip offsets and that some software may not accept SHORT values.

For maximum compatibility with operating systems such as MS-DOS and Windows, the StripOffsets array should be less than or equal to 64K bytes in length, and the strips themselves, in both compressed and uncompressed forms, should not be larger than 64K bytes.

No default. See also StripByteCounts, RowsPerStrip, TileOffsets, TileByteCounts.

SubfileType

A general indication of the kind of data contained in this subfile.

Tag = 255 (FF.H)

Type = SHORT

N = 1

Currently defined values are:

1 = full-resolution image data

2 = reduced-resolution image data

3 = a single page of a multi-page image (see the PageNumber field description).

Note that several image types may be found in a single TIFF file, with each subfile described by its own IFD.

No default.

This field is deprecated. The NewSubfileType field should be used instead.

Threshholding

For black and white TIFF files that represent shades of gray, the technique used to convert from gray to black and white pixels.

Tag = 263 (107.H)

Type = SHORT

N = 1

1 = No dithering or halftoning has been applied to the image data.

2 = An ordered dither or halftone technique has been applied to the image data.

3 = A randomized process such as error diffusion has been applied to the image data.

Default is Threshholding = 1. See also CellWidth, CellLength.

XResolution

The number of pixels per ResolutionUnit in the ImageWidth direction.

Tag = 282 (11A.H)

Type = RATIONAL

N = 1

It is not mandatory that the image be actually displayed or printed at the size implied by this parameter. It is up to the application to use this information as it wishes.

No default. See also YResolution, ResolutionUnit.

YResolution

The number of pixels per ResolutionUnit in the ImageLength direction.

Tag = 283 (11B.H)

Type = RATIONAL

N = 1

No default. See also XResolution, ResolutionUnit.

Section 9: PackBits Compression

This section describes TIFF compression type 32773, a simple byte-oriented run-length scheme.

Description

In choosing a simple byte-oriented run-length compression scheme, we arbitrarily chose the Apple Macintosh PackBits scheme. It has a good worst case behavior (at most 1 extra byte for every 128 input bytes). For Macintosh users, the toolbox utilities PackBits and UnPackBits will do the work for you, but it is easy to implement your own routines.

A pseudo code fragment to unpack might look like this:

```
Loop until you get the number of unpacked bytes you are expecting:
    Read the next source byte into n.
    If n is between 0 and 127 inclusive, copy the next n+1 bytes literally.
    Else if n is between -127 and -1 inclusive, copy the next byte -n+1
    times.
    Else if n is -128, noop.
Endloop
```

In the inverse routine, it is best to encode a 2-byte repeat run as a replicate run except when preceded and followed by a literal run. In that case, it is best to merge the three runs into one literal run. Always encode 3-byte repeats as replicate runs.

That is the essence of the algorithm. Here are some additional rules:

- Pack each row separately. Do not compress across row boundaries.
- The number of uncompressed bytes per row is defined to be (ImageWidth + 7) / 8. If the uncompressed bitmap is required to have an even number of bytes per row, decompress into word-aligned buffers.
- If a run is larger than 128 bytes, encode the remainder of the run as one or more additional replicate runs.

When PackBits data is decompressed, the result should be interpreted as per compression type 1 (no compression).

Section 10: Modified Huffman Compression

This section describes TIFF compression scheme 2, a method for compressing bilevel data based on the CCITT Group 3 1D facsimile compression scheme.

References

- "Standardization of Group 3 facsimile apparatus for document transmission," Recommendation T.4, Volume VII, Fascicle VII.3, Terminal Equipment and Protocols for Telematic Services, The International Telegraph and Telephone Consultative Committee (CCITT), Geneva, 1985, pages 16 through 31.
- "Facsimile Coding Schemes and Coding Control Functions for Group 4 Facsimile Apparatus," Recommendation T.6, Volume VII, Fascicle VII.3, Terminal Equipment and Protocols for Telematic Services, The International Telegraph and Telephone Consultative Committee (CCITT), Geneva, 1985, pages 40 through 48.

We do not believe that these documents are necessary in order to implement Compression=2. We have included (verbatim in most places) all the pertinent information in this section. However, if you wish to order the documents, you can write to ANSI, Attention: Sales, 1430 Broadway, New York, N.Y., 10018. Ask for the publication listed above—it contains both Recommendation T.4 and T.6.

Relationship to the CCITT Specifications

The CCITT Group 3 and Group 4 specifications describe communications protocols for a particular class of devices. They are not by themselves sufficient to describe a disk data format. Fortunately, however, the CCITT coding schemes can be readily adapted to this different environment. The following is one such adaptation. Most of the language is copied directly from the CCITT specifications.

See Section 11 for additional CCITT compression options.

Coding Scheme

A line (row) of data is composed of a series of variable length code words. Each code word represents a run length of all white or all black. (Actually, more than one code word may be required to code a given run, in a manner described below.) White runs and black runs alternate.

To ensure that the receiver (decompressor) maintains color synchronization, all data lines begin with a white run-length code word set. If the actual scan line begins with a black run, a white run-length of zero is sent (written). Black or white run-lengths are defined by the code words in Tables 1 and 2. The code words are of two types: Terminating code words and Make-up code words. Each run-length is represented by zero or more Make-up code words followed by exactly one Terminating code word.

Run lengths in the range of 0 to 63 pels (pixels) are encoded with their appropriate Terminating code word. Note that there is a different list of code words for black and white run-lengths.

Run lengths in the range of 64 to 2623 (2560+63) pels are encoded first by the Make-up code word representing the run-length that is nearest to, not longer than, that required. This is then followed by the Terminating code word representing the difference between the required run-length and the run-length represented by the Make-up code.

Run lengths in the range of lengths longer than or equal to 2624 pels are coded first by the Make-up code of 2560. If the remaining part of the run (after the first Make-up code of 2560) is 2560 pels or greater, additional Make-up code(s) of 2560 are issued until the remaining part of the run becomes less than 2560 pels. Then the remaining part of the run is encoded by Terminating code or by Make-up code plus Terminating code, according to the range mentioned above.

It is considered an unrecoverable error if the sum of the run-lengths for a line does not equal the value of the ImageWidth field.

New rows always begin on the next available byte boundary.

No EOL code words are used. No fill bits are used, except for the ignored bits at the end of the last byte of a row. RTC is not used.

An encoded CCITT string is self-photometric, defined in terms of white and black runs. Yet TIFF defines a tag called PhotometricInterpretation that also purports to define what is white and what is black. Somewhat arbitrarily, we adopt the following convention:

The "normal" PhotometricInterpretation for bilevel CCITT compressed data is WhiteIsZero. In this case, the CCITT "white" runs are to be interpretated as white, and the CCITT "black" runs are to be interpreted as black. However, if the PhotometricInterpretation is BlackIsZero, the TIFF reader must reverse the meaning of white and black when displaying and printing the image.

Table 1/T.4 Terminating codes

White run length	Code word	Black run length	Code word
0	00110101	0	0000110111
1	000111	1	010
2	0111	2	11
3	1000	3	10
4	1011	4	011
5	1100	5	0011
6	1110	6	0010
7	1111	7	00011
8	10011	8	000101
9	10100	9	000100
10	00111	10	0000100
11	01000	11	0000101
12	001000	12	0000111
13	000011	13	00000100
14	110100	14	00000111
15	110101	15	000011000
16	101010	16	0000010111
17	101011	17	0000011000
18	0100111	18	0000001000
19	0001100	19	00001100111
20	0001000	20	00001101000
21	0010111	21	00001101100
22	0000011	22	00000110111
23	0000100	23	00000101000
24	0101000	24	00000010111
25	0101011	25	00000011000
6	0010011	26	000011001010
27	0100100	27	000011001011
28	0011000	28	000011001100
29	00000010	29	000011001101
30	00000011	30	000001101000
31	00011010	31	000001101001
32	00011011	32	000001101010
33	00010010	33	000001101011
34	00010011	34	000011010010
35	00010100	35	000011010011
36	00010101	36	000011010100
37	00010110	37	000011010101
38	00010111	38	000011010110
39	00101000	39	000011010111
40	00101001	40	000001101100
41	00101010	41	000001101101
42	00101011	42	000011011010
43	00101100	43	000011011011
44	00101101	44	000001010100
45	00000100	45	000001010101
46	00000101	46	000001010110
47	00001010	47	000001010111
48	00001011	48	000001100100
49	01010010	49	000001100101
50	01010011	50	000001010010
51	01010100	51	000001010011

White run length	Code word	Black run length	Code word
52	01010101	52	000000100100
53	00100100	53	000000110111
54	00100101	54	000000111000
55	01011000	55	000000100111
56	01011001	56	000000101000
57	01011010	57	000001011000
58	01011011	58	000001011001
59	01001010	59	000000101011
60	01001011	60	000000101100
61	00110010	61	000001011010
62	00110011	62	000001100110
63	00110100	63	000001100111

Table 2/T.4 Make-up codes

White run length	Code word	Black run length	Code word
64	11011	64	0000001111
128	10010	128	000011001000
192	010111	192	000011001001
256	0110111	256	000001011011
320	00110110	320	000000110011
384	00110111	384	000000110100
448	01100100	448	000000110101
512	01100101	512	0000001101100
576	01101000	576	0000001101101
640	01100111	640	0000001001010
704	011001100	704	0000001001011
768	011001101	768	0000001001100
832	011010010	832	0000001001101
896	011010011	896	0000001110010
960	011010100	960	0000001110011
1024	011010101	1024	0000001110100
1088	011010110	1088	0000001110101
1152	011010111	1152	0000001110110
1216	011011000	1216	0000001110111
1280	011011001	1280	0000001010010
1344	011011010	1344	0000001010011
1408	011011011	1408	0000001010100
1472	010011000	1472	0000001010101
1536	010011001	1536	0000001011010
1600	010011010	1600	0000001011011
1664	011000	1664	0000001100100
1728	010011011	1728	0000001100101
EOL	000000000001	EOL	00000000000

Additional make-up codes

White and Black run length	Make-up code word
1792	00000001000
1856	00000001100
1920	00000001101
1984	000000010010
2048	000000010011
2112	000000010100
2176	000000010101
2240	000000010110
2304	000000010111
2368	000000011100
2432	000000011101
2496	000000011110
2560	000000011111

Part 2: TIFF Extensions

Part 2 contains extensions to Baseline TIFF. TIFF Extensions are TIFF features that may not be supported by all TIFF readers. TIFF creators who use these features will have to work closely with TIFF readers in their part of the industry to ensure successful interchange.

The features described in this part were either contained in earlier versions of the specification, or have been approved by the TIFF Advisory Committee.

Section 11: CCITT Bilevel Encodings

The following fields are used when storing binary pixel arrays using one of the encodings adopted for raster-graphic interchange in numerous CCITT and ISO (International Organization for Standards) recommendations and standards. These encodings are often spoken of as “Group III compression” and “Group IV compression” because their application in facsimile transmission is the most widely known.

For the specialized use of these encodings in storing facsimile-transmission images, further guidelines can be obtained from the TIFF Class F document, available on-line in the same locations as this specification. This document is administered by another organization; paper copies are not available from Aldus.

Compression

Tag = 259 (103.H)

Type = SHORT

N = 1

3 = T4-encoding: CCITT T.4 bi-level encoding as specified in section 4, Coding, of CCITT Recommendation T.4: “Standardization of Group 3 Facsimile apparatus for document transmission.” International Telephone and Telegraph Consultative Committee (CCITT, Geneva: 1988).

See the T4Options field for T4-encoding options such as 1D vs 2D coding.

4 = T6-encoding: CCITT T.6 bi-level encoding as specified in section 2 of CCITT Recommendation T.6: “Facsimile coding schemes and coding control functions for Group 4 facsimile apparatus.” International Telephone and Telegraph Consultative Committee (CCITT, Geneva: 1988).

See the T6Options field for T6-encoding options such as escape into uncompressed mode to avoid negative-compression cases.

Application in Image Interchange

CCITT Recommendations T.4 and T.6 are specified in terms of the serial bit-by-bit creation and processing of a variable-length binary string that encodes bi-level (black and white) pixels of a rectangular image array. Generally, the encoding schemes are described in terms of bit-serial communication procedures and the end-to-end coordination that is required to gain reliable delivery over inherently unreliable data links. The Group 4 procedures, with their T6-encoding, represent a significant simplification because it is assumed that a reliable communication medium is employed, whether ISDN or X.25 or some other trustworthy transport vehicle. Because image-storage systems and computers achieve data integrity and communication reliability in other ways, the T6-encoding tends to be prefered for imaging applications. When computer storage and retrieval and interchange of facsimile material are of interest, the T4-encodings provide a better match to the current generation of Group 3 facsimile products and their defenses against data corruption as the result of transmission defects.

Whichever form of encoding is preferable for a given application, there are a number of adjustments that need to be made to account for the capture of the CCITT binary-encoding strings as part of electronically-stored material and digital-image interchange.

PhotometricInterpretation. An encoded CCITT string is self-photometric, defined in terms of white and black runs. Yet TIFF defines a tag called PhotometricInterpretation that also purports to define what is white and what is black. Somewhat arbitrarily, we adopt the following convention:

The "normal" PhotometricInterpretation for bilevel CCITT compressed data is WhiteIsZero. In this case, the CCITT "white" runs are to be interpretated as white, and the CCITT "black" runs are to be interpreted as black. However, if the PhotometricInterpretation is BlackIsZero, the TIFF reader must reverse the meaning of white and black when displaying and printing the image.

FillOrder. When CCITT encodings are used directly over a typical serial communication link, the order of the bits in the encoded string is the sequential order of the string, bit-by-bit, from beginning to end. This poses the following question: In which order should consecutive blocks of eight bits be assembled into octets (standard data bytes) for use within a computer system? The answer differs depending on whether we are concerned about preserving the serial-transmission sequence or preserving only the format of byte-organized sequences in memory and in stored files.

From the perspective of electronic interchange, as long as a receiver's reassembly of bits into bytes properly mirrors the way in which the bytes were disassembled by the transmitter, no one cares which order is seen on the transmission link because each multiple of 8 bits is transparently transmitted.

Common practice is to record arbitrary binary strings into storage sequences such that the first sequential bit of the string is found in the high-order bit of the first octet of the stored byte sequence. This is the standard case specified by TIFF FillOrder = 1, used in most bitmap interchange and the only case required in Baseline TIFF. This is also the approach used for the octets of standard 8-bit character data, with little attention paid to the fact that the most common forms of data communication transmit and reassemble individual 8-bit frames with the low-order-bit first!

For bit-serial transmission to a distant unit whose approach to assembling bits into bytes is unknown and supposed to be irrelevant, it is necessary to satisfy the expected sequencing of bits over the transmission link. This is the normal case for communication between facsimile units and also for computers and modems emulating standard Group 3 facsimile units. In this case, if the CCITT encoding is captured directly off of the link via standard communication adapters, TIFF FillOrder = 2 will usually apply to that stored data form.

Consequently, different TIFF FillOrder cases may arise when CCITT encodings are obtained by synthesis within a computer (including Group 4 transmission, which is treated more like computer data) instead of by capture from a Group 3 facsimile unit.

Because this is such a subtle situation, with surprisingly disruptive consequences for FillOrder mismatches, the following practice is urged whenever CCITT bilevel encodings are used:

a. TIFF FillOrder (tag 266) should always be explicitly specified.

b. FillOrder = 1 should be employed wherever possible in persistent material that is intended for interchange. This is the only reliable case for widespread interchange among computer systems, and it is important to explicitly confirm the honoring of standard assumptions.

c. FillOrder = 2 should occur only in highly-localized and preferably-transient material, as in a facsimile server supporting group 3 facsimile equipment. The tag should be present as a safeguard against the CCITT encoding “leaking” into an unsuspecting application, allowing readers to detect and warn against the occurence.

There are interchange situations where fill order is not distinguished, as when filtering the CCITT encoding into a PostScript level 2 image operation. In this case, as in most other cases of computer-based information interchange, FillOrder=1 is assumed, and any padding to a multiple of 8 bits is accomplished by adding a sufficient number of 0-bits to the end of the sequence.

Strips and Tiles. When CCITT bi-level encoding is employed, interaction with stripping (Section 3) and tiling (Section 15) is as follows:

a. Decompose the image into segments—individual pixel arrays representing the desired strip or tile configuration. The CCITT encoding procedures are applied most flexibly if the segments each have a multiple of 4 lines.

b. Individually encode each segment according to the specified CCITT bi-level encoding, as if each segment is a separate raster-graphic image.

The reason for this general rule is that CCITT bi-level encodings are generally progressive. That is, the initial line of pixels is encoded, and then subsequent lines, according to a variety of options, are encoded in terms of changes that need to be made to the preceding (unencoded) line. For strips and tiles to be individually usable, they must each start as fresh, independent encodings.

Miscellaneous features. There are provisions in CCITT encoding that are mostly meaningful during facsimile-transmission procedures. There is generally no significant application when storing images in TIFF or other data interchange formats, although TIFF applications should be tolerant and flexible in this regard. These features tend to have significance only when facilitating transfer between facsimile and non-facsimile applications of the encoded raster-graphic images. Further considerations for fill sequences, end-of-line flags, return-to-control (end-of-block) sequences and byte padding are introduced in discussion of the individual encoding options.

T4Options

Tag = 292 (124.H)

Type = LONG

N = 1

See Compression=3. This field is made up of a set of 32 flag bits. Unused bits must be set to 0. Bit 0 is the low-order bit.

Bit 0 is 1 for 2-dimensional coding (otherwise 1-dimensional is assumed). For 2-D coding, if more than one strip is specified, each strip must begin with a 1-

dimensionally coded line. That is, RowsPerStrip should be a multiple of “Parameter K,” as documented in the CCITT specification.

Bit 1 is 1 if uncompressed mode is used.

Bit 2 is 1 if fill bits have been added as necessary before EOL codes such that EOL always ends on a byte boundary, thus ensuring an EOL-sequence of 1 byte preceded by a zero nibble: xxxx-0000 0000-0001.

Default is 0, for basic 1-dimensional coding. See also Compression.

T6Options

Tag = 293 (125.H)

Type = LONG

N = 1

See *Compression = 4.* This field is made up of a set of 32 flag bits. Unused bits must be set to 0. Bit 0 is the low-order bit. The default value is 0 (all bits 0).

bit 0 is unused and always 0.

bit 1 is 1 if uncompressed mode is allowed in the encoding.

In earlier versions of TIFF, this tag was named Group4Options. The significance has not changed and the present definition is compatible. The name of the tag has been changed to be consistent with the nomenclature of other T.6-encoding applications.

Readers should honor this option tag, and only this option tag, whenever T.6-Encoding is specified for Compression.

For T.6-Encoding, each segment (strip or tile) is encoded as if it were a separate image. The encoded string from each segment starts a fresh byte.

There are no one-dimensional line encodings in T.6-Encoding. Instead, even the first row of the segment’s pixel array is encoded two-dimensionally by always assuming an invisible preceding row of all-white pixels. The 2-dimensional procedure for encoding the body of individual rows is the same as that used for 2-dimensional T.4-encoding and is described fully in the CCITT specifications.

The beginning of the encoding for each row of a strip or tile is conducted as if there is an imaginary preceding (0-width) white pixel, that is as if a fresh run of white pixels has just commenced. The completion of each line is encoded as if there are imaginary pixels beyond the end of the current line, and of the preceding line, in effect, of colors chosen such that the line is exactly completable by a code word, making the imaginary next pixel a changing element that’s not actually used.

The encodings of successive lines follow contiguously in the binary T.6-Encoding stream with no special initiation or separation codewords. There are no provisions for fill codes or explicit end-of-line indicators. The encoding of the last line of the pixel array is followed immediately, in place of any additional line encodings, by a 24-bit End-of-Facsimile Block (EOFB).

000000000001000000000001.B.

The EOFB sequence is immediately followed by enough 0-bit padding to fit the entire stream into a sequence of 8-bit bytes.

General Application. Because of the single uniform encoding procedure, without disruptions by end-of-line codes and shifts into one-dimensional encodings, T.6-encoding is very popular for compression of bi-level images in document imaging systems. T.6-encoding trades off redundancy for minimum encoded size, relying on the underlying storage and transmission systems for reliable retention and communication of the encoded stream.

TIFF readers will operate most smoothly by always ignoring bits beyond the EOFB. Some writers may produce additional bytes of pad bits beyond the byte containing the final bit of the EOFB. Robust readers will not be disturbed by this prospect.

It is not possible to correctly decode a T.6-Encoding without knowledge of the exact number of pixels in each line of the pixel array. ImageWidth (or TileWidth, if used) must be stated exactly and accurately. If an image or segment is overscanned, producing extraneous pixels at the beginning or ending of lines, these pixels must be counted. Any cropping must be accomplished by other means. It is not possible to recover from a pixel-count deviation, even when one is detected. Failure of any row to be completed as expected is cause for abandoning further decoding of the entire segment. There is no requirement that ImageWidth be a multiple of eight, of course, and readers must be prepared to pad the final octet bytes of decoded bitmap rows with additional bits.

If a TIFF reader encounters EOFB before the expected number of lines has been extracted, it is appropriate to assume that the missing rows consist entirely of white pixels. Cautious readers might produce an unobtrusive warning if such an EOFB is followed by anything other than pad bits.

Readers that successfully decode the RowsPerStrip (or TileLength or residual ImageLength) number of lines are not required to verify that an EOFB follows. That is, it is generally appropriate to stop decoding when the expected lines are decoded or the EOFB is detected, whichever occurs first. Whether error indications or warnings are also appropriate depends upon the application and whether more precise troubleshooting of encoding deviations is important.

TIFF writers should always encode the full, prescribed number of rows, with a proper EOFB immediately following in the encoding. Padding should be by the least number of 0-bits needed for the T.6-encoding to exactly occupy a multiple of 8 bits. Only 0-bits should be used for padding, and StripByteCount (or TileByteCount) should not extend to any bytes not containing properly-formed T.6-encoding. In addition, even though not required by T.6-encoding rules, successful interchange with a large variety of readers and applications will be enhanced if writers can arrange for the number of pixels per line and the number of lines per strip to be multiples of eight.

Uncompressed Mode. Although T.6-encodings of simple bi-level images result in data compressions of 10:1 and better, some pixel-array patterns have T.6-encodings that require more bits than their simple bi-level bitmaps. When such cases are detected by encoding procedures, there is an optional extension for shifting to a form of uncompressed coding within the T.6-encoding string.

Uncompressed mode is not well-specified and many applications discourage its usage, prefering alternatives such as different compressions on a segment-by-segment (strip or tile) basis, or by simply leaving the image uncompressed in its entirety. The main complication for readers is in properly restoring T.6-encoding after the uncompressed sequence is laid down in the current row.

Readers that have no provision for uncompressed mode will generally reject any case in which the flag is set. Readers that are able to process uncompressed-mode content within T.6-encoding strings can safely ignore this flag and simply process any uncompressed-mode occurences correctly.

Writers that are unable to guarantee the absence of uncompressed-mode material in any of the T.6-encoded segments must set the flag. The flag should be cleared (or defaulted) only when absence of uncompressed-mode material is assured. Writers that are able to inhibit the generation of uncompressed-mode extensions are encouraged to do so in order to maximize the acceptability of their T.6-encoding strings in interchange situations.

Because uncompressed-mode is not commonly used, the following description is best taken as suggestive of the general machinery. Interpolation of fine details can easily vary between implementations.

Uncompressed mode is signalled by the occurence of the 10-bit extension code string

0000001111.B

outside of any run-length make-up code or extension. Original unencoded image information follows. In this unencoded information, a 0-bit evidently signifies a white pixel, a 1-bit signifies a black pixel, and the TIFF PhotometricInterpretation will influence how these bits are mapped into any final uncompressed bitmap for use. The only modification made to the unencoded information is insertion of a 1-bit after every block of five consecutive 0-bits from the original image information. This is a transparency device that allows longer sequencences of 0-bits to be reserved for control conditions, especially ending the uncompressed-mode sequence. When it is time to return to compressed mode, the 8-bit exit sequence

0000001t.B

is appended to the material. The 0-bits of the exit sequence are not considered in applying the 1-bit insertion rule; up to four information 0-bits can legally precede the exit sequence. The trailing bit, ‘t,’ specifies the color (via 0 or 1) that is understood in the next run of compressed-mode encoding. This lets a color other than white be assumed for the 0-width pixel on the left of the edge between the last uncompressed pixel and the resumed 2-dimensional scan.

Writers should confine uncompressed-mode sequences to the interiors of individual rows, never attempting to “wrap” from one row to the next. Readers must operate properly when the only encoding for a single row consists of an uncompressed-mode escape, a complete row of (proper 1-inserted) uncompressed information, and the extension exit. Technically, the exit pixel, ‘t,’ should probably then be the opposite color of the last true pixel of the row, but readers should be generous in this case.

In handling these complex encodings, the encounter of material from a defective source or a corrupted file is particularly unsettling and mysterious. Robust readers will do well to defend against falling off the end of the world; e.g., unexpected EOFB sequences should be handled, and attempted access to data bytes that are not within the bounds of the present segment (or the TIFF file itself) should be avoided.

Section 12: Document Storage and Retrieval

These fields may be useful for document storage and retrieval applications. They will very likely be ignored by other applications.

DocumentName

The name of the document from which this image was scanned.

Tag = 269 (10D.H)

Type = ASCII

See also PageName.

PageName

The name of the page from which this image was scanned.

Tag = 285 (11D.H)

Type = ASCII

See also DocumentName.

No default.

PageNumber

The page number of the page from which this image was scanned.

Tag = 297 (129.H)

Type = SHORT

N = 2

This field is used to specify page numbers of a multiple page (e.g. facsimile) document. PageNumber[0] is the page number; PageNumber[1] is the total number of pages in the document. If PageNumber[1] is 0, the total number of pages in the document is not available.

Pages need not appear in numerical order.

The first page is numbered 0 (zero).

No default.

XPosition

X position of the image.

Tag = 286 (11E.H)

Type = RATIONAL

N = 1

The X offset in ResolutionUnits of the left side of the image, with respect to the left side of the page.

No default. See also YPosition.

YPosition

Y position of the image.

Tag = 287 (11F.H)

Type = RATIONAL

N = 1

The Y offset in ResolutionUnits of the top of the image, with respect to the top of the page. In the TIFF coordinate scheme, the positive Y direction is down, so that YPosition is always positive.

No default. See also XPosition.

Section 13: LZW Compression

This section describes TIFF compression scheme 5, an adaptive compression scheme for raster images.

Restrictions

When LZW compression was added to the TIFF specification, in Revision 5.0, it was thought to be public domain. This is, apparently, not the case.

The following paragraph has been approved by the Unisys Corporation:

"The LZW compression method is said to be the subject of United States patent number 4,558,302 and corresponding foreign patents owned by the Unisys Corporation. Software and hardware developers may be required to license this patent in order to develop and market products using the TIFF LZW compression option. Unisys has agreed that developers may obtain such a license on reasonable, nondiscriminatory terms and conditions. Further information can be obtained from: Welch Licensing Department, Office of the General Counsel, M/S C1SW19, Unisys Corporation, Blue Bell, Pennsylvania, 19424."

Reportedly, there are also other companies with patents that may affect LZW implementors.

Reference

Terry A. Welch, "A Technique for High Performance Data Compression", IEEE Computer, vol. 17 no. 6 (June 1984). Describes the basic Lempel-Ziv & Welch (LZW) algorithm in very general terms. The author's goal is to describe a hardware-based compressor that could be built into a disk controller or database engine and used on all types of data. There is no specific discussion of raster images. This section gives sufficient information so that the article is not required reading.

Characteristics

LZW compression has the following characteristics:

- LZW works for images of various bit depths.
- LZW has a reasonable worst-case behavior.
- LZW handles a wide variety of repetitive patterns well.
- LZW is reasonably fast for both compression and decompression.
- LZW does not require floating point software or hardware.

- LZW is lossless. All information is preserved. But if noise or information is removed from an image, perhaps by smoothing or zeroing some low-order bitplanes, LZW compresses images to a smaller size. Thus, 5-bit, 6-bit, or 7-bit data masquerading as 8-bit data compresses better than true 8-bit data. Smooth images also compress better than noisy images, and simple images compress better than complex images.
- LZW works quite well on bilevel images, too. On our test images, it almost always beat PackBits and generally tied CCITT 1D (Modified Huffman) compression. LZW also handles halftoned data better than most bilevel compression schemes.

The Algorithm

Each strip is compressed independently. We strongly recommend that RowsPerStrip be chosen such that each strip contains about 8K bytes before compression. We want to keep the strips small enough so that the compressed and uncompressed versions of the strip can be kept entirely in memory, even on small machines, but are large enough to maintain nearly optimal compression ratios.

The LZW algorithm is based on a translation table, or string table, that maps strings of input characters into codes. The TIFF implementation uses variable-length codes, with a maximum code length of 12 bits. This string table is different for every strip and does not need to be reatained for the decompressor. The trick is to make the decompressor automatically build the same table as is built when the data is compressed. We use a C-like pseudocode to describe the coding scheme:

```
InitializeStringTable();
WriteCode(ClearCode);
  = the empty string;
for each character in the strip {
        K = GetNextCharacter();
        if  +K is in the string table {
                =  +K;  /* string concatenation */
        } else {
               WriteCode (CodeFromString( ));
               AddTableEntry( +K);
                 = K;
        }
} /* end of for loop */
WriteCode (CodeFromString( ));
WriteCode (EndOfInformation);
```

That's it. The scheme is simple, although it is challenging to implement efficiently. But we need a few explanations before we go on to decompression.

The "characters" that make up the LZW strings are bytes containing TIFF uncompressed (Compression=1) image data, in our implementation. For example, if BitsPerSample is 4, each 8-bit LZW character will contain two 4-bit pixels. If BitsPerSample is 16, each 16-bit pixel will span two 8-bit LZW characters.

It is also possible to implement a version of LZW in which the LZW character depth equals BitsPerSample, as described in Draft 2 of Revision 5.0. But there is a major problem with this approach. If BitsPerSample is greater than 11, we can not use 12-bit-maximum codes and the resulting LZW table is unacceptably large. Fortunately, due to the adaptive nature of LZW, we do not pay a significant compression ratio penalty for combining several pixels into one byte before compressing. For example, our 4-bit sample images compressed about 3 percent worse, and our 1-bit images compressed about 5 percent better. And it is easier to write an LZW compressor that always uses the same character depth than it is to write one that handles varying depths.

We can now describe some of the routine and variable references in our pseudocode:

InitializeStringTable() initializes the string table to contain all possible single-character strings. There are 256 of them, numbered 0 through 255, since our characters are bytes.

WriteCode() writes a code to the output stream. The first code written is a ClearCode, which is defined to be code #256.

is our "prefix string."

GetNextCharacter() retrieves the next character value from the input stream. This will be a number between 0 and 255 because our characters are bytes.

The "+" signs indicate string concatenation.

AddTableEntry() adds a table entry. (InitializeStringTable() has already put 256 entries in our table. Each entry consists of a single-character string, and its associated code value, which, in our application, is identical to the character itself. That is, the 0th entry in our table consists of the string <0>, with a corresponding code value of <0>, the 1st entry in the table consists of the string <1>, with a corresponding code value of <1> and the 255th entry in our table consists of the string <255>, with a corresponding code value of <255>.) So, the first entry that added to our string table will be at position 256, right? Well, not quite, because we reserve code #256 for a special "Clear" code. We also reserve code #257 for a special "EndOfInformation" code that we write out at the end of the strip. So the first multiple-character entry added to the string table will be at position 258.

For example, suppose we have input data that looks like this:

Pixel 0:<7>

Pixel 1:<7>

Pixel 2:<7>

Pixel 3:<8>

Pixel 4:<8>

Pixel 5:<7>

Pixel 6:<7>

Pixel 7:<6>

Pixel 8:<6>

First, we read Pixel 0 into K. K is then simply <7>, because is an empty string at this point. Is the string <7> already in the string table? Of course, because all single character strings were put in the table by InitializeStringTable(). So set equal to <7>, and then go to the top of the loop.

Read Pixel 1 into K. Does K (<7><7>) exist in the string table? No, so we write the code associated with to output (write <7> to output) and add K (<7><7>) to the table as entry 258. Store K (<7>) into . Note that although we have added the string consisting of Pixel 0 and Pixel 1 to the table, we “re-use” Pixel 1 as the beginning of the next string.

Back at the top of the loop, we read Pixel 2 into K. Does K (<7><7>) exist in the string table? Yes, the entry we just added, entry 258, contains exactly <7><7>. So we add K to the end of so that is now <7><7>.

Back at the top of the loop, we read Pixel 3 into K. Does K (<7><7><8>) exist in the string table? No, so we write the code associated with (<258>) to output and then add K to the table as entry 259. Store K (<8>) into .

Back at the top of the loop, we read Pixel 4 into K. Does K (<8><8>) exist in the string table? No, so we write the code associated with (<8>) to output and then add K to the table as entry 260. Store K (<8>) into .

Continuing, we get the following results:

After reading:	We write to output:	And add table entry:
Pixel 0		
Pixel 1	<7>	258: <7><7>
Pixel 2		
Pixel 3	<258>	259: <7><7><8>
Pixel 4	<8>	260:<8><8>
Pixel 5	<8>	261: <8><7>
Pixel 6		
Pixel 7	<258>	262: <7><7><6>
Pixel 8	<6>	263: <6><6>

WriteCode() also requires some explanation. In our example, the output code stream, <7><258><8><8><258><6> should be written using as few bits as possible. When we are just starting out, we can use 9-bit codes, since our new string table entries are greater than 255 but less than 512. *After adding table entry 511, switch to 10-bit codes (i.e., entry 512 should be a 10-bit code.) Likewise, switch to 11-bit codes after table entry 1023, and 12-bit codes after table entry 2047.* We will arbitrarily limit ourselves to 12-bit codes, so that our table can have at most 4096 entries. The table should not be any larger.

Whenever you add a code to the output stream, it “counts” toward the decision about bumping the code bit length. This is important when writing the last code word before an EOI code or ClearCode, to avoid code length errors.

What happens if we run out of room in our string table? This is where the ClearCode comes in. As soon as we use entry 4094, we write out a (12-bit) ClearCode. (If we wait any longer to write the ClearCode, the decompressor might try to interpret the ClearCode as a 13-bit code.) At this point, the compressor reinitializes the string table and then writes out 9-bit codes again.

Note that whenever you write a code and add a table entry, is not left empty. It contains exactly one character. Be careful not to lose it when you write an end-of-table ClearCode. You can either write it out as a 12-bit code before writing the ClearCode, in which case you need to do it right after adding table entry 4093, or you can write it as a 9-bit code after the ClearCode . Decompression gives the same result in either case.

To make things a little simpler for the decompressor, we will require that each strip begins with a ClearCode and ends with an EndOfInformation code. Every LZW-compressed strip must begin on a byte boundary. It need not begin on a word boundary. LZW compression codes are stored into bytes in high-to-low-order fashion, i.e., FillOrder is assumed to be 1. The compressed codes are written as bytes (not words) so that the compressed data will be identical whether it is an 'II' or 'MM' file.

Note that the LZW string table is a continuously updated history of the strings that have been encountered in the data. Thus, it reflects the characteristics of the data, providing a high degree of adaptability.

LZW Decoding

The procedure for decompression is a little more complicated:

```
while ((Code = GetNextCode()) != EoiCode) {
        if (Code == ClearCode) {
                InitializeTable();
                Code = GetNextCode();
                if (Code == EoiCode)
                        break;
                WriteString(StringFromCode(Code));
                OldCode = Code;
        }  /* end of ClearCode case */
        else {
                if (IsInTable(Code)) {
                        WriteString(StringFromCode(Code));
                        AddStringToTable(StringFromCode(OldCode
)+FirstChar(StringFromCode(Code)));
                        OldCode = Code;
                } else {
                        OutString = StringFromCode(OldCode) +
FirstChar(StringFromCode(OldCode));
                        WriteString(OutString);
                        AddStringToTable(OutString);
                        OldCode = Code;
                }
        } /* end of not-ClearCode case */
} /* end of while loop */
```

The function GetNextCode() retrieves the next code from the LZW-coded data. It must keep track of bit boundaries. It knows that the first code that it gets will be a 9-bit code. We add a table entry each time we get a code. So, GetNextCode() must switch over to 10-bit codes as soon as string *#510* is stored into the table. *Similarly, the switch is made to 11-bit codes after #1022 and to 12-bit codes after #2046.*

The function StringFromCode() gets the string associated with a particular code from the string table.

The function AddStringToTable() adds a string to the string table. The "+" sign joining the two parts of the argument to AddStringToTable indicates string concatenation.

StringFromCode() looks up the string associated with a given code.

WriteString() adds a string to the output stream.

When SamplesPerPixel Is Greater Than 1

So far, we have described the compression scheme as if SamplesPerPixel were always 1, as is the case with palette-color and grayscale images. But what do we do with RGB image data?

Tests on our sample images indicate that the LZW compression ratio is nearly identical whether PlanarConfiguration=1 or PlanarConfiguration=2, for RGB images. So, use whichever configuration you prefer and simply compress the bytes in the strip.

Note: Compression ratios on our test RGB images were disappointingly low: between 1.1 to 1 and 1.5 to 1, depending on the image. Vendors are urged to do what they can to remove as much noise as possible from their images. Preliminary tests indicate that significantly better compression ratios are possible with less-noisy images. Even something as simple as zeroing-out one or two least-significant bitplanes can be effective, producing little or no perceptible image degradation.

Implementation

The exact structure of the string table and the method used to determine if a string is already in the table are probably the most significant design decisions in the implementation of a LZW compressor and decompressor. Hashing has been suggested as a useful technique for the compressor. We have chosen a tree-based approach, with good results. The decompressor is more straightforward and faster because no search is involved—strings can be accessed directly by code value.

LZW Extensions

Some images compress better using LZW coding if they are first subjected to a process wherein each pixel value is replaced by the difference between the pixel and the preceding pixel. See the following Section.

Acknowledgments

See the first page of this section for the LZW reference.

The use of ClearCode as a technique for handling overflow was borrowed from the compression scheme used by the Graphics Interchange Format (GIF), a small-color-paint-image-file format used by CompuServe that also uses an adaptation of the LZW technique.

Section 14: Differencing Predictor

This section defines a Predictor that greatly improves compression ratios for some images.

Predictor

Tag = 317 (13D.H)

Type = SHORT

N = 1

A predictor is a mathematical operator that is applied to the image data before an encoding scheme is applied. Currently this field is used only with LZW (Compression=5) encoding because LZW is probably the only TIFF encoding scheme that benefits significantly from a predictor step. See Section 13.

The possible values are:

1 = No prediction scheme used before coding.

2 = Horizontal differencing.

Default is 1.

The algorithm

Make use of the fact that many continuous-tone images rarely vary much in pixel value from one pixel to the next. In such images, if we replace the pixel values by differences between consecutive pixels, many of the differences should be 0, plus or minus 1, and so on. This reduces the apparent information content and allows LZW to encode the data more compactly.

Assuming 8-bit grayscale pixels for the moment, a basic C implementation might look something like this:

```
char      image[ ][ ];
int       row, col;

/* take horizontal differences:
 */
for (row = 0; row < nrows; row++)
          for (col = ncols - 1; col >= 1; col--)
                 image[row][col] -= image[row][col-1];
```

If we don't have 8-bit components, we need to work a little harder to make better use of the architecture of most CPUs. Suppose we have 4-bit components packed two per byte in the normal TIFF uncompressed (i.e., Compression=1) fashion. To find differences, we want to first expand each 4-bit component into an 8-bit byte, so that we have one component per byte, low-order justified. We then perform the horizontal differencing illustrated in the example above. Once the differencing has been completed, we then repack the 4-bit differences two to a byte, in the normal TIFF uncompressed fashion.

If the components are greater than 8 bits deep, expanding the components into 16-bit words instead of 8-bit bytes seems like the best way to perform the subtraction on most computers.

Note that we have not lost any data up to this point, nor will we lose any data later on. It might seem at first that our differencing might turn 8-bit components into 9-bit differences, 4-bit components into 5-bit differences, and so on. But it turns out that we can completely ignore the “overflow” bits caused by subtracting a larger number from a smaller number and still reverse the process without error. Normal two’s complement arithmetic does just what we want. Try an example by hand if you need more convincing.

Up to this point we have implicitly assumed that we are compressing bilevel or grayscale images. An additional consideration arises in the case of color images.

If PlanarConfiguration is 2, there is no problem. Differencing works the same as it does for grayscale data.

If PlanarConfiguration is 1, however, things get a little trickier. If we didn’t do anything special, we would subtract red component values from green component values, green component values from blue component values, and blue component values from red component values. This would not give the LZW coding stage much redundancy to work with. So, we will do our horizontal differences with an offset of SamplesPerPixel (3, in the RGB case). In other words, we will subtract red from red, green from green, and blue from blue. The LZW coding stage is identical to the SamplesPerPixel=1 case. We require that BitsPerSample be the same for all 3 components.

Results and Guidelines

LZW without differencing works well for 1-bit images, 4-bit grayscale images, and many palette-color images. But natural 24-bit color images and some 8-bit grayscale images do much better with differencing.

Although the combination of LZW coding with horizontal differencing does not result in any loss of data, it may be worthwhile in some situations to give up some information by removing as much noise as possible from the image data before doing the differencing, especially with 8-bit components. The simplest way to get rid of noise is to mask off one or two low-order bits of each 8-bit component. On our 24-bit test images, LZW with horizontal differencing yielded an average compression ratio of 1.4 to 1. When the low-order bit was masked from each component, the compression ratio climbed to 1.8 to 1; the compression ratio was 2.4 to 1 when masking two bits, and 3.4 to 1 when masking three bits. Of course, the more you mask, the more you risk losing useful information along with the noise. We encourage you to experiment to find the best compromise for your device. For some applications, it may be useful to let the user make the final decision.

Incidentally, we tried taking both horizontal and vertical differences, but the extra complexity of two-dimensional differencing did not appear to pay off for most of our test images. About one third of the images compressed slightly better with two-dimensional differencing, about one third compressed slightly worse, and the rest were about the same.

Section 15: Tiled Images

Introduction

Motivation

This section describes how to organize images into tiles instead of strips.

For low-resolution to medium-resolution images, the standard TIFF method of breaking the image into strips is adequate. However high-resolution images can be accessed more efficiently—and compression tends to work better—if the image is broken into roughly square tiles instead of horizontally-wide but vertically-narrow strips.

Relationship to existing fields

When the tiling fields described below are used, they replace the StripOffsets, StripByteCounts, and RowsPerStrip fields. Use of tiles will therefore cause older TIFF readers to give up because they will have no way of knowing where the image data is or how it is organized. **Do not** use both strip-oriented and tile-oriented fields in the same TIFF file.

Padding

Tile size is defined by TileWidth and TileLength. All tiles in an image are the same size; that is, they have the same pixel dimensions.

Boundary tiles are padded to the tile boundaries. For example, if TileWidth is 64 and ImageWidth is 129, then the image is 3 tiles wide and 63 pixels of padding must be added to fill the rightmost column of tiles. The same holds for TileLength and ImageLength. It doesn't matter what value is used for padding, because good TIFF readers display only the pixels defined by ImageWidth and ImageLength and ignore any padded pixels. Some compression schemes work best if the padding is accomplished by replicating the last column and last row instead of padding with 0's.

The price for padding the image out to tile boundaries is that some space is wasted. But compression usually shrinks the padded areas to almost nothing. Even if data is not compressed, remember that tiling is intended for large images. Large images have lots of comparatively small tiles, so that the percentage of wasted space will be very small, generally on the order of a few percent or less.

The advantages of padding an image to the tile boundaries are that implementations can be simpler and faster and that it is more compatible with tile-oriented compression schemes such as JPEG. See Section 22.

Tiles are compressed individually, just as strips are compressed. *That is, each row of data in a tile is treated as a separate "scanline" when compressing.* Compression includes any padded areas of the rightmost and bottom tiles so that all the tiles in an image are the same size when uncompressed.

All of the following fields are required for tiled images:

Fields

TileWidth

Tag = 322 (142.H)

Type = SHORT or LONG

N = 1

The tile width in pixels. This is the number of columns in each tile.

Assuming integer arithmetic, three computed values that are useful in the following field descriptions are:

TilesAcross = (ImageWidth + TileWidth - 1) / TileWidth

TilesDown = (ImageLength + TileLength - 1) / TileLength

TilesPerImage = TilesAcross * TilesDown

These computed values are not TIFF fields; they are simply values determined by the ImageWidth, TileWidth, ImageLength, and TileLength fields.

TileWidth and ImageWidth together determine the number of tiles that span the width of the image (TilesAcross). TileLength and ImageLength together determine the number of tiles that span the length of the image (TilesDown).

We recommend choosing TileWidth and TileLength such that the resulting tiles are about 4K to 32K bytes before compression. This seems to be a reasonable value for most applications and compression schemes.

TileWidth must be a multiple of 16. This restriction improves performance in some graphics environments and enhances compatibility with compression schemes such as JPEG.

Tiles need not be square.

Note that ImageWidth can be less than TileWidth, although this means that the tiles are too large or that you are using tiling on really small images, neither of which is recommended. The same observation holds for ImageLength and TileLength.

No default. See also TileLength, TileOffsets, TileByteCounts.

TileLength

Tag = 323 (143.H)

Type = SHORT or LONG

N = 1

The tile length (height) in pixels. This is the number of rows in each tile.

TileLength must be a multiple of 16 for compatibility with compression schemes such as JPEG.

Replaces RowsPerStrip in tiled TIFF files.

No default. See also TileWidth, TileOffsets, TileByteCounts.

TileOffsets

Tag = 324 (144.H)

Type = LONG

N = TilesPerImage for PlanarConfiguration = 1

= SamplesPerPixel * TilesPerImage for PlanarConfiguration = 2

For each tile, the byte offset of that tile, as compressed and stored on disk. The offset is specified with respect to the beginning of the TIFF file. Note that this implies that each tile has a location independent of the locations of other tiles.

Offsets are ordered left-to-right and top-to-bottom. For PlanarConfiguration = 2, the offsets for the first component plane are stored first, followed by all the offsets for the second component plane, and so on.

No default. See also TileWidth, TileLength, TileByteCounts.

TileByteCounts

Tag = 325 (145.H)

Type = SHORT or LONG

N = TilesPerImage for PlanarConfiguration = 1

= SamplesPerPixel * TilesPerImage for PlanarConfiguration = 2

For each tile, the number of (compressed) bytes in that tile.

See TileOffsets for a description of how the byte counts are ordered.

No default. See also TileWidth, TileLength, TileOffsets.

Section 16: CMYK Images

Motivation

This section describes how to store separated (usually CMYK) image data in a TIFF file.

In a separated image, each pixel consists of N components. Each component represents the amount of a particular ink that is to be used to represent the image at that location, typically using a halftoning technique.

For example, in a CMYK image, each pixel consists of 4 components. Each component represents the amount of cyan, magenta, yellow, or black process ink that is to be used to represent the image at that location.

The fields described in this section can be used for more than simple 4-color process (CMYK) printing. They can also be used for describing an image made up of more than 4 inks, such an image made up of a cyan, magenta, yellow, red, green, blue, and black inks. Such an image is sometimes called a high-fidelity image and has the advantage of slightly extending the printed color gamut.

Since separated images are quite device-specific and are restricted to color prepress use, they should not be used for general image data interchange. Separated images are to be used only for prepress applications in which the imagesetter, paper, ink, and printing press characteristics are known by the creator of the separated image.

Note: there is no single method of converting RGB data to CMYK data and back. In a perfect world, something close to cyan = 255-red, magenta = 255-green, and yellow = 255-blue should work; but characteristics of printing inks and printing presses, economics, and the fact that the meaning of RGB itself depends on other parameters combine to spoil this simplicity.

Requirements

In addition to satisfying the normal Baseline TIFF requirements, a separated TIFF file must have the following characteristics:

- **SamplesPerPixel = N.** SHORT. The number of inks. (For example, N=4 for CMYK, because we have one component each for cyan, magenta, yellow, and black.)
- **BitsPerSample = 8,8,8,8 (for CMYK).** SHORT. For now, only 8-bit components are recommended. The value “8” is repeated SamplesPerPixel times.
- **PhotometricInterpretation = 5 (Separated - usually CMYK).** SHORT. The components represent the desired percent dot coverage of each ink, where the larger component values represent a higher percentage of ink dot coverage and smaller values represent less coverage.

Fields

In addition, there are some new fields, all of which are optional.

InkSet

Tag = 332 (14C.H)

Type = SHORT

N = 1

The set of inks used in a separated (PhotometricInterpretation=5) image.

1 = CMYK. The order of the components is cyan, magenta, yellow, black. Usually, a value of 0 represents 0% ink coverage and a value of 255 represents 100% ink coverage for that component, but see DotRange below. The InkNames field should not exist when InkSet=1.

2 = not CMYK. See the InkNames field for a description of the inks to be used.

Default is 1 (CMYK).

NumberOfInks

Tag = 334 (14E.H)

Type = SHORT

N = 1

The number of inks. Usually equal to SamplesPerPixel, unless there are extra samples.

See also ExtraSamples.

Default is 4.

InkNames

Tag = 333 (14D.H)

Type = ASCII

N = total number of characters in all the ink name strings, including the NULs.

The name of each ink used in a separated (PhotometricInterpretation=5) image, written as a list of concatenated, NUL-terminated ASCII strings. The number of strings must be equal to NumberOfInks.

The samples are in the same order as the ink names.

See also InkSet, NumberOfInks.

No default.

DotRange

Tag = 336 (150.H)

Type = BYTE or SHORT

N = 2, or 2*SamplesPerPixel

The component values that correspond to a 0% dot and 100% dot. DotRange[0] corresponds to a 0% dot, and DotRange[1] corresponds to a 100% dot.

If a DotRange pair is included for each component, the values for each component are stored together, so that the pair for Cyan would be first, followed by the pair for Magenta, and so on. *Use of multiple dot ranges is, however, strongly discouraged in the interests of simplicity and compatibility with ANSI IT8 standards.*

A number of prepress systems like to keep some "headroom" and "footroom" on both ends of the range. What to do with components that are less than the 0% aim point or greater than the 100% aim point is not specified and is application-dependent.

It is strongly recommended that a CMYK TIFF writer not attempt to use this field to reverse the sense of the pixel values so that smaller values mean more ink instead of less ink. That is, DotRange[0] should be less than DotRange[1].

DotRange[0] and DotRange[1] must be within the range [0, (2**BitsPerSample) - 1].

Default: a component value of 0 corresponds to a 0% dot, and a component value of 255 (assuming 8-bit pixels) corresponds to a 100% dot. That is, DotRange[0] = 0 and DotRange[1] = (2**BitsPerSample) - 1.

TargetPrinter

Tag = 337 (151.H)

Type = ASCII

N = any

A description of the printing environment for which this separation is intended.

History

This Section has been expanded from earlier drafts, with the addition of the **InkSet, InkNames, NumberOfInks, DotRange**, and **TargetPrinter**, but is backward-compatible with earlier draft versions.

Possible future enhancements: definition of the characterization information so that the CMYK data can be retargeted to a different printing environment and so that display on a CRT or proofing device can more accurately represent the color. ANSI IT8 is working on such a proposal.

Section 17: HalftoneHints

This section describes a scheme for properly placing highlights and shadows in halftoned images.

Introduction

The single most easily recognized failing of continuous tone images is the incorrect placement of highlight and shadow. It is critical that a halftone process be capable of printing the lightest areas of the image as the smallest halftone spot capable of the output device, at the specified printer resolution and screen ruling. Specular highlights (small ultra-white areas) as well as the shadow areas should be printable as paper only.

Consistency in highlight and shadow placement allows the user to obtain predictable results on a wide variety of halftone output devices. Proper implementation of theHalftoneHints field will provide a significant step toward device independent imaging, such that low cost printers may to be used as effective proofing devices for images which will later be halftoned on a high-resolution imagesetter.

The HalftoneHints Field

HalftoneHints

Tag = 321 (141.H)

Type = SHORT

N = 2

The purpose of the HalftoneHints field is to convey to the halftone function the range of gray levels within a colorimetrically-specified image that should retain tonal detail. The field contains two values of sixteen bits each and, therefore, is contained wholly within the field itself; no offset is required. The first word specifies the highlight gray level which should be halftoned at the lightest printable tint of the final output device. The second word specifies the shadow gray level which should be halftoned at the darkest printable tint of the final output device. Portions of the image which are whiter than the highlight gray level will quickly, if not immediately, fade to specular highlights. There is no default value specified, since the highlight and shadow gray levels are a function of the subject matter of a particular image.

Appropriate values may be derived algorithmically or may be specified by the user, either directly or indirectly.

The HalftoneHints field, as defined here, defines an achromatic function. It can be used just as effectively with color images as with monochrome images. When used with opponent color spaces such as CIE L*a*b* or YCbCr, it refers to the achromatic component only; L* in the case of CIELab, and Y in the case of

YCbCr. When used with tri-stimulus spaces such as RGB, it suggests to retain tonal detail for all colors with an NTSC gray component within the bounds of the R=G=B=Highlight to R=G=B=Shadow range.

Comments for TIFF Writers

TIFF writers are encouraged to include the HalftoneHints field in all color or grayscale images where BitsPerSample >1. Although no default value is specified, prior to the introduction of this field it has been common practice to implicitly specify the highlight and shadow gray levels as 1 and 2**BitsperSample-2 and manipulate the image data to this definition. There are some disadvantages to this technique, and it is not feasible for a fixed gamut colorimetric image type. Appropriate values may be derived algorithmically or may be specified by the user directly or indirectly. Automatic algorithms exist for analyzing the histogram of the achromatic intensity of an image and defining the minimum and maximum values as the highlight and shadow settings such that tonal detail is retained throughout the image. This kind of algorithm may try to impose a highlight or shadow where none really exists in the image, which may require user controls to override the automatic setting.

It should be noted that the choice of the highlight and shadow values is somewhat output dependent. For instance, in situations where the dynamic range of the output medium is very limited (as in newsprint and, to a lesser degree, laser output), it may be desirable for the user to clip some of the lightest or darkest tones to avoid the reduced contrast resulting from compressing the tone of the entire image. Different settings might be chosen for 150-line halftone printed on coated stock. Keep in mind that these values may be adjusted later (which might not be possible unless the image is stored as a colorimetric, fixed, full-gamut image), and that more sophisticated page-layout applications may be capable of presenting a user interface to consider these decisions at a point where the halftone process is well understood.

It should be noted that although CCDs are linear intensity detectors, TIFF writers may choose to manipulate the image to store gamma-compensated data. Gamma-compensated data is more efficient at encoding an image than is linear intensity data because it requires fewer BitsPerPixel to eliminate banding in the darker tones. It also has the advantage of being closer to the tone response of the display or printer and is, therefore, less likely to produce poor results from applications that are not rigorous about their treatment of images. Be aware that the PhotometricInterpretation value of 0 or 1 (grayscale) implies linear data because no gamma is specified. The PhotometricInterpretation value of 2 (RGB data) specifies the NTSC gamma of 2.2 as a default. If data is written as something other than the default, then a GrayResponseCurve field or a TransferFunction field must be present to define the deviation. For grayscale data, be sure that the densities in the GrayResponseCurve are consistent with the PhotometricInterpretation field and the HalftoneHints field.

Comments for TIFF Readers

TIFF readers that send a grayscale image to a halftone output device, whether it is a binary laser printer or a PostScript imagesetter should make an effort to maintain the highlight and shadow placement. This requires two steps. First, determine the highlight and shadow gray level of a particular image. Second, communicate that information to the halftone engine.

To determine the highlight and shadow gray levels, begin by looking for a HalftoneHints field. If it exists, it takes precedence. The first word represents the gray level of the highlight and the second word represents the gray level of the shadow. If the image is a colorimetric image (i.e. it has a GrayResponseCurve field or a TransferFunction field) but does not contain a HalftoneHints field, then the gamut mapping techniques described earlier should be used to determine the highlight and shadow values. If neither of these conditions are true, then the file should be treated as if a HalftoneHints field had indicated a highlight at gray level 1 and a shadow at gray level 2**BitsPerPixel-2 (or vice-versa depending on the PhotometricInterpretation field). Once the highlight and shadow gray levels have been determined, the next step is to communicate this information to the halftone module. The halftone module may exist within the same application as the TIFF reader, it may exist within a separate printer driver, or it may exist within the Raster Image Processor (RIP) of the printer itself. Whether the halftone process is a simple dither pattern or a general purpose spot function, it has some gray level at which the lightest printable tint will be rendered. The HalftoneHint concept is best implemented in an environment where this lightest printable tint is easily and consistently specified.

There are several ways in which an application can communicate the highlight and shadow to the halftone function. Some environments may allow the application to pass the highlight and shadow to the halftone module explicitly along with the image. This is the best approach, but many environments do not yet provide this capability. Other environments may provide fixed gray levels at which the highlight and shadow will be rendered. For these cases, the application should build a tone map that matches the highlight and shadow specified in the image to the highlight and shadow gray level of the halftone module. This approach requires more work by the application software, but will provide excellent results. Some environments will not have any consistent concept of highlight and shadow at all. In these environments, the best an application can do is characterize each of the supported printers and save the observed highlight and shadow gray levels. The application can then use these values to achieve the desired results, providing the environment doesn't change.

Once the highlight and shadow areas are selected, care should be taken to appropriately map intermediate gray levels to those expected by the halftone engine, which may or may not be linear Reflectance. Note that although CCDs are linear intensity detectors and many TIFF files are stored as linear intensity, most output devices require significant tone compensation (sometimes called gamma correction) to correctly display or print linear data. Be aware that the PhotometricInterpretation value of 0, 1 implies linear data because no gamma is specified. The PhotometricInterpretation value of 2 (RGB data) specifies the NTSC gamma of 2.2 as a default. If a GrayResponseCurve field or a TransferFunction field is present, it may define something other than the default.

Some Background on the Halftone Process

To obtain the best results when printing a continuous-tone raster image, it is seldom desirable to simply reproduce the tones of the original on the printed page. Most often there is some gamut mapping required. Often this is because the tonal range of the original extends beyond the tonal range of the output medium. In some cases, the tone range of the original is within the gamut of the output medium, but it may be more pleasing to expand the tone of the image to fill the range of the output. Given that the tone of the original is to be adjusted, there is a whole range of possibilities for the level of sophistication that may be undertaken by a software application.

Printing monochrome output is far less sophisticated than printing color output. For monochrome output the first priority is to control the placement of the highlight and the shadow. Ideally, a quality halftone will have sufficient levels of gray so that a standard observer cannot distinguish the interface between any two adjacent levels of gray. In practice, however, there is often a significant step between the tone of the paper and the tone of the lightest printable tint. Although usually less severe, the problem is similar between solid ink and the darkest printable tint. Since the dynamic range between the lightest printable tint and the darkest printable tint is usually less than one would like, it is common to maximize the tone of the image within these bounds. Not all images will have a highlight (an area of the image which is desirable to print as light as possible while still retaining tonal detail). If one exists, it should be carefully controlled to print at the lightest printable tint of the output medium. Similarly, the darkest areas of the image to retain tonal detail should be printed as the darkest printable tint of the output medium. Tones lighter or darker than these may be clipped at the limits of the paper and ink. Satisfactory results may be obtained in monochrome work by doing nothing more than a perceptually-linear mapping of the image between these rigorously controlled endpoints. This level of sophistication is sufficient for many mid-range applications, although the results often appear flatter (i.e. lower in contrast) than desired.

The next step is to increase contrast slightly in the tonal range of the image that contains the most important subject matter. To perform this step well requires considerably more information about the image and about the press. To know where to add contrast, the algorithm must have access to first the keyness of the image; the tone range which the user considers most important. To know how much contrast to add, the algorithm must have access to the absolute tone of the original and the dynamic range of the output device so that it may calculate the amount of tone compression to which the image is actually subjected.

Most images are called normal key. The important subject areas of a normal key image are in the midtones. These images do well when a so-called “sympathetic curve” is applied, which increases the contrast in midtones slightly at the expense of contrast in the lighter and darker tones. White china on a white tablecloth is an example of a high key image. High key images benefit from higher contrast in lighter tones, with less contrast needed in the midtones and darker tones. Low key images have important subject matter in the darker tones and benefit from increasing the contrast in the darker tones. Specifying the keyness of an image might be attempted by automatic techniques, but it will likely fail without user input. For example, a photo of a bride in a white wedding dress it may be a high key image if you are selling wedding dresses, but may be a normal key image if you are the parents of the bride and are more interested in her smile.

Sophisticated color reproduction employs all of these principles, and then applies them in three dimensions. The mapping of the highlight and shadow becomes only one small, albeit critical, portion of the total issue of mapping colors that are too saturated for the output medium. Here again, automatic techniques may be employed as a first pass, with the user becoming involved in the clip or compress mapping decision. The HalftoneHints field is still useful in communicating which portions of the intensity of the image must be retained and which may be clipped. Again, a sophisticated application may override these settings if later user input is received.

Section 18: Associated Alpha Handling

This section describes a scheme for handling images with alpha data.

Introduction

A common technique in computer graphics is to assemble an image from one or more elements that are rendered separately. When elements are combined using compositing techniques, matte or coverage information must be present for each pixel to create a properly anti-aliased accumulation of the full image [Porter84]. This matting information is an example of additional per-pixel data that must be maintained with an image. This section describes how to use the ExtraSamples field to store the requisite matting information, commonly called the associated alpha or just alpha. This scheme enables efficient manipulation of image data during compositing operations.

Images with matting information are stored in their natural format but with an additional component per pixel. The ExtraSample field is included with the image to indicate that an extra component of each pixel contains associated alpha data. In addition, when associated alpha data are included with RGB data, the RGB components must be stored premultiplied by the associated alpha component and component values in the range [0,2**BitsPerSample-1] are implicitly mapped onto the [0,1] interval. That is, for each pixel (r,g,b) and opacity A, where r, g, b, and A are in the range [0,1], (A*r,A*g,A*b,A) must be stored in the file. If A is zero, then the color components should be interpreted as zero. Storing data in this pre-multiplied format, allows compositing operations to be implemented most efficiently. In addition, storing pre-multiplied data makes it possible to specify colors with components outside the normal [0,1] interval. The latter is useful for defining certain operations that effect only the luminescence [Porter84].

Fields

ExtraSamples

Tag = 338 (152.H)

Type = SHORT

N = 1

This field must have a value of 1 (associated alpha data with pre-multiplied color components). The associated alpha data stored in component SamplesPerPixel-1 of each pixel contains the opacity of that pixel, and the color information is pre-multiplied by alpha.

Comments

Associated alpha data is just another component added to each pixel. Thus, for example, its size is defined by the value of the BitsPerSample field.

Note that since data is stored with RGB components already multiplied by alpha, naive applications that want to display an RGBA image on a display can do so simply by displaying the RGB component values. This works because it is effectively the same as merging the image with a black background. That is, to merge one image with another, the color of resultant pixels are calculated as:

$$C_r = C_{over} * A_{over} + C_{under} * (1 - A_{over})$$

Since the "under image" is a black background, this equation reduces to

$$C_r = C_{over} * A_{over}$$

which is exactly the pre-multiplied color; i.e. what is stored in the image.

On the other hand, to print an RGBA image, one must composite the image over a suitable background page color. For a white background, this is easily done by adding 1 - A to each color component. For an arbitrary background color C_{back}, the *printed color* of each pixel is

$$C_{print} = C_{image} + C_{back} * (1 - A_{image})$$

(since C_{image} is pre-multiplied).

Since the ExtraSamples field is independent of other fields, this scheme permits alpha information to be stored in whatever organization is appropriate. In particular, components can be stored packed (PlanarConfiguration=1); this is important for good I/O performance and for good memory access performance on machines that are sensitive to data locality. However, if this scheme is used, TIFF readers must not derive the SamplesPerPixel from the value of the PhotometricInterpretation field (e.g., if RGB, then SamplesPerPixel is 3).

In addition to being independent of data storage-related fields, the field is also independent of the PhotometricInterpretation field. This means, for example, that it is easy to use this field to specify grayscale data and associated matte information. Note that a palette-color image with associated alpha will not have the colormap indices pre-multiplied; rather, the RGB colormap values will be pre-multiplied.

Unassociated Alpha and Transparency Masks

Some image manipulation applications support notions of transparency masks and soft-edge masks. The associated alpha information described in this section is different from this *unassociated alpha* information in many ways, most importantly:

- Associated alpha describes opacity or coverage at each pixel, while clipping-related alpha information describes a boolean relationship. That is, associated alpha can specify fractional coverage at a pixel, while masks specify either 0 or 100 percent coverage.
- Once defined, associated alpha is not intended to be removed or edited, except as a result of compositing the image; it is an integral part of an image.

Unassociated alpha, on the other hand, is designed as an ancillary piece of information.

References

[Porter84] "Compositing Digital Images". Thomas Porter, Tom Duff; Lucasfilm Ltd. ACM SIGGRAPH Proceedings Volume 18, Number 3. July, 1984.

Section 19: Data Sample Format

This section describes a scheme for specifying data sample type information.

TIFF implicitly types all data samples as unsigned integer values. Certain applications, however, require the ability to store image-related data in other formats such as floating point. This section presents a scheme for describing a variety of data sample formats.

Fields

SampleFormat

Tag = 339 (153.H)

Type = SHORT

N = SamplesPerPixel

This field specifies how to interpret each data sample in a pixel. Possible values are:

1 = unsigned integer data

2 = two's complement signed integer data

3 = IEEE floating point data [IEEE]

4 = undefined data format

Note that the SampleFormat field does not specify the size of data samples; this is still done by the BitsPerSample field.

A field value of "undefined" is a statement by the writer that it did not know how to interpret the data samples; for example, if it were copying an existing image. A reader would typically treat an image with "undefined" data as if the field were not present (i.e. as unsigned integer data).

Default is 1, unsigned integer data.

SMinSampleValue

Tag = 340 (154.H)

Type = the field type that best matches the sample data

N = SamplesPerPixel

This field specifies the minimum sample value. Note that a value should be given for each data sample. That is, if the image has 3 SamplesPerPixel, 3 values must be specified.

The default for SMinSampleValue and SMaxSampleValue is the full range of the data type.

SMaxSampleValue

Tag = 341 (155.H)

Type = the field type that best matches the sample data

N = SamplesPerPixel

This new field specifies the maximum sample value.

Comments

The SampleFormat field allows more general imaging (such as image processing) applications to employ TIFF as a valid file format.

SMinSampleValue and SMaxSampleValue become more meaningful when image data is typed. The presence of these fields makes it possible for readers to assume that data samples are bound to the range [SMinSampleValue, SMaxSampleValue] without scanning the image data.

References

[IEEE] "IEEE Standard 754 for Binary Floating-point Arithmetic".

Section 20: RGB Image Colorimetry

Without additional information, RGB data is device-specific; that is, without an absolute color meaning. This section describes a scheme for describing and characterizing RGB image data.

Introduction

Color printers, displays, and scanners continue to improve in quality and availability while they drop in price. Now the problem is to display color images so that they appear to be identical on different hardware.

The key to reproducing the same color on different devices is to use the CIE 1931 XYZ color-matching functions, the international standard for color comparison. Using CIE XYZ, an image's colorimetry information can fully describe its color interpretation. The approach taken here is essentially calibrated RGB. It implies a transformation from the RGB color space of the pixels to CIE 1931 XYZ.

The appearance of a color depends not only on its absolute tristimulus values, but also on the conditions under which it is viewed, including the nature of the surround and the adaptation state of the viewer. Colors having the same absolute tristimulus values appear the same in identical viewing conditions. The more complex issue of color appearance under different viewing conditions is addressed by [4]. The colorimetry information presented here plays an important role in color appearance under different viewing conditions.

Assuming identical viewing conditions, an application using the tags described below can display an image on different hardware and achieve colorimetrically identical results. The process of using this colorimetry information for displaying an image is straightforward on a color monitor but it is more complex for color printers. Also, the results will be limited by the color gamut and other characteristics of the display or printing device.

The following fields describe the image colorimetry information of a TIFF image:

WhitePoint	chromaticity of the white point of the image
PrimaryChromaticities	chromaticities of the primaries of the image
TransferFunction	transfer function for the pixel data
TransferRange	extends the range of the transfer function
ReferenceBlackWhite	pixel component headroom and footroom parameters

The TransferFunction, TransferRange, and ReferenceBlackWhite fields have defaults based on industry standards. An image has a colorimetric interpretation if and only if both the WhitePoint and PrimaryChromaticities fields are present. An image without these colorimetry fields will be displayed in an application and hardware dependent manner.

Note: In the following definitions, BitsPerSample is used as if it were a single number when in fact it is an array of SamplesPerPixel numbers. The elements of this array may not always be equal, for example: 5/6/5 16-bit pixels. BitsPerSample should be interpreted as the BitsPerSample value associated with a particular component. In the case of unequal BitsPerSample values, the definitions below can be extended in a straightforward manner.

This section has the following differences with Appendix H in TIFF 5.0:

- removed the use of image colorimetry defaults
- renamed the ColorResponseCurves field as TransferFunction
- optionally allowed a single TransferFunction table to describe all three channels
- described the use of the TransferFunction field for YCbCr, Palette, WhiteIsZero and BlackIsZero PhotometricInterpretation types
- added the TransferRange tag to expand the range of the TransferFunction below black and above white
- added the ReferenceBlackWhite field
- addressed the issue of color appearance

Colorimetry Field Definitions

WhitePoint

Tag = 318 (13E.H)

Type = RATIONAL

N = 2

The chromaticity of the white point of the image. This is the chromaticity when each of the primaries has its ReferenceWhite value. The value is described using the 1931 CIE xy chromaticity diagram and only the chromaticity is specified. This value can correspond to the chromaticity of the alignment white of a monitor, the filter set and light source combination of a scanner or the imaging model of a rendering package. The ordering is white[x], white[y].

For example, the CIE Standard Illuminant D65 used by CCIR Recommendation 709 and Kodak PhotoYCC is:

3127/10000,3290/10000

No default.

PrimaryChromaticities

Tag =319 (13F.H)

Type = RATIONAL

N = 6

The chromaticities of the primaries of the image. This is the chromaticity for each of the primaries when it has its ReferenceWhite value and the other primaries have their ReferenceBlack values. These values are described using the 1931 CIE xy chromaticity diagram and only the chromaticities are specified. These values can correspond to the chromaticities of the phosphors of a monitor, the filter set and light source combination of a scanner or the imaging model of a rendering package. The ordering is red[x], red[y], green[x], green[y], blue[x], and blue[y].

For example the CCIR Recommendation 709 primaries are:

640/1000,330/1000,

300/1000, 600/1000,

150/1000, 60/1000

No default.

TransferFunction

Tag =301 (12D.H)

Type = SHORT

N = {1 or 3} * (1 << BitsPerSample)

Describes a transfer function for the image in tabular style. Pixel components can be gamma-compensated, companded, non-uniformly quantized, or coded in some other way. The TransferFunction maps the pixel components from a non-linear BitsPerSample (e.g. 8-bit) form into a 16-bit linear form without a perceptible loss of accuracy.

If N = 1 << BitsPerSample, the transfer function is the same for each channel and all channels share a single table. Of course, this assumes that each channel has the same BitsPerSample value.

If N = 3 * (1 << BitsPerSample), there are three tables, and the ordering is the same as it is for pixel components of the PhotometricInterpretation field. These tables are separate and not interleaved. For example, with RGB images all red entries come first, followed by all green entries, followed by all blue entries.

The length of each component table is 1 << BitsPerSample. The width of each entry is 16 bits as implied by the type SHORT. Normally the value 0 represents the minimum intensity and 65535 represents the maximum intensity and the values [0, 0, 0] represent black and [65535,65535, 65535] represent white. If the TransferRange tag is present then it is used to determine the minimum and maximum values, and a scaling normalization.

The TransferFunction can be applied to images with a PhotometricInterpretation value of RGB, Palette, YCbCr, WhiteIsZero, and BlackIsZero. The TransferFunction is not used with other PhotometricInterpretation types.

For RGB PhotometricInterpretation, ReferenceBlackWhite expands the coding range, TransferRange expands the range of the TransferFunction, and the TransferFunction tables decompand the RGB value. The WhitePoint and PrimaryChromaticities further describe the RGB colorimetry.

For Palette color PhotometricInterpretation, the Colormap maps the pixel into three 16-bit values that when scaled to BitsPerSample-bits serve as indices into the TransferFunction tables which decompand the RGB value. The WhitePoint and PrimaryChromaticities further describe the underlying RGB colorimetry.

A Palette value can be scaled into a TransferFunction index by:

index= (value * ((1 << BitsPerSample) - 1)) / 65535;

A TransferFunction index can be scaled into a Palette color value by:

value= (index * 65535L) / ((1 << BitsPerSample) - 1);

Be careful if you intend to create Palette images with a TransferFunction. If the Colormap tag is directly converted from a hardware colormap, it may have a device gamma already incorporated into the DAC values.

For YCbCr PhotometricInterpretation, ReferenceBlackWhite expands the coding range, the YCbCrCoefficients describe the decoding matrix to transform YCbCr into RGB, TransferRange expands the range of the TransferFunction, and the TransferFunction tables decompand the RGB value. The WhitePoint and PrimaryChromaticities fields provide further description of the underlying RGB colorimetry.

After coding range expansion by ReferenceBlackWhite and TransferFunction expansion by TransferRange, RGB values may be outside the domain of the TransferFunction. Also, the display device matrix can transform RGB values into display device RGB values outside the domain of the device. These values are handled in an application-dependent manner.

For RGB images with non-default ReferenceBlackWhite coding range expansion and for YCbCr images, the resolution of the TransferFunction may be insufficient. For example, after the YCbCr transformation matrix, the decoded RGB values must be rounded to index into the TransferFunction tables. Applications needing the extra accuracy should interpolate between the elements of the TransferFunction tables. Linear interpolation is recommended.

For WhiteIsZero and BlackIsZero PhotometricInterpretation, the TransferFunction decompands the grayscale pixel value to a linear 16-bit form. Note that a TransferFunction value of 0 represents black and 65535 represents white regardless of whether a grayscale image is WhiteIsZero or BlackIsZero. For example, the zeroth element of a WhiteIsZero TransferFunction table will likely be 65535. This extension of the TransferFunction field for grayscale images is intended to replace the GrayResponseCurve field.

The TransferFunction does not describe a transfer characteristic outside of the range for ReferenceBlackWhite.

Default is a single table corresponding to the NTSC standard gamma value of 2.2. This table is used for each channel. It can be generated by:

```
NValues = 1 << BitsPerSample;
for (TF[0]= 0, i = 1; i < NValues; i++)
        TF[i]= floor(pow(i / (NValues - 1.0), 2.2) * 65535 + 0.5);
```

TransferRange

Tag = 342 (156.H)

Type = SHORT

N = 6

Expands the range of the TransferFunction. The first value within a pair is associated with TransferBlack and the second is associated with TransferWhite. The ordering of pairs is the same as for pixel components of the PhotometricInterpretation type. By default, theTransferFunction is defined over a range from a minimum intensity, 0 or nominal black, to a maximum intensity,(1 << BitsPerSample) - 1 or nominal white. Kodak PhotoYCC uses an extended range TransferFunction in order to describe highlights, saturated colors and shadow detail beyond this range. The TransferRange expands the TransferFunction to support these values. It is defined only for RGB and YCbCr PhotometricInterpretations.

After ReferenceBlackWhite and/or YCbCr decoding has taken place, an RGB value can be represented as a real number. It is then rounded to create an index into the TransferFunctiontable. In the absence of a TransferRange tag, or if the tag has the default values, the rounded value is an index and the normalized intensity value is:

```
index = (int) (value + (value < 0.0? -0.5 : 0.5));
intensity = TF[index] / 65535;
```

If the TransferRange tag is present and has non-default values, it provides an offset to be used with the rounded index. It also describes a scaling. The normalized intensity value is:

```
index = (int) (value + (value < 0.0? -0.5 : 0.5));
intensity = (TF[index + TransferRange[Black]] -
          TF[TransferRange[Black]])
          / (TF[TransferRange[White]] - TF[TransferRange[Black]]);
```

An application can write a TransferFunction with a non-defaultTransferRange as follows:

```
black_offset = scale_factor * Transfer(-TransferRange[Black]ar /
          (TransferRange[White] - TransferRange[Black]));
for (i = 0; i < (1 << BitsPerSample); i++)
          TF[i] = floor(0.5 - black_offset + scale_factor
                    * Transfer((i - TransferRange[Black])
                    / (TransferRange[White] - TransferRange[Black])));
```

The TIFF writer chooses scale_factor such that the TransferFunction fits into a 16-bit unsigned short, and chooses the TransferRange so that the most important part of the TransferFunction fits into the table.

Default is [0, NV, 0, NV, 0, NV] where NV = (1 <<BitsPerSample) - 1.

ReferenceBlackWhite

Tag =532 (214.H)

Type = RATIONAL

N = 6

Specifies a pair of headroom and footroom image data values (codes) for each pixel component. The first component code within a pair is associated with ReferenceBlack, and the second is associated with ReferenceWhite. The ordering of pairs is the same as those for pixel components of the PhotometricInterpretation type. ReferenceBlackWhite can be applied to images with a PhotometricInterpretation value of RGB or YCbCr. ReferenceBlackWhite is not used with other PhotometricInterpretation values.

Computer graphics commonly places black and white at the extremities of the binary representation of image data; for example, black at code 0 and white at code 255. In other disciplines, such as printing, film, and video, there are practical reasons to provide footroom codes below ReferenceBlack and headroom codes above ReferenceWhite.

In film applications, they correspond to the densities Dmax and Dmin. In video applications, ReferenceBlack corresponds to 7.5 IRE and 0 IRE in systems with and without setup respectively, and ReferenceWhite corresponds to 100 IRE units.

Using YCbCr (See Section 21) and the CCIR Recommendation 601.1 video standard as an example, code 16 represents ReferenceBlack, and code 235 represents ReferenceWhite for the luminance component (Y). For the chrominance components, Cb and Cr, code 128 represents ReferenceBlack, and code 240 represents ReferenceWhite. With Cb and Cr, the ReferenceWhite value is used to code reference blue and reference red respectively.

The full range component value is converted from the code by:

```
FullRangeValue = (code - ReferenceBlack) * CodingRange
          / (ReferenceWhite - ReferenceBlack);
```

The code is converted from the full-range component value by:

```
code = (FullRangeValue * (ReferenceWhite - ReferenceBlack)
          / CodingRange) + ReferenceBlack;
```

For RGB images and the Y component of YCbCr images, CodingRange is defined as:

```
CodingRange = 2 ** BitsPerSample - 1;
```

For the Cb and Cr components of YCbCr images, CodingRange is defined as:

```
CodingRange = 127;
```

For RGB images, in the default special case of no headroom or footroom, this conversion can be skipped because the scaling multiplier equals 1.0 and the value equals the code.

For YCbCr images, in the case of no headroom or footroom, the conversion for Y can be skipped because the value equals the code. For Cb and Cr, ReferenceBlack must still be subtracted from the code. In the general case, the scaling multiplication for the Cb and Cr component codes can be factored into the YCbCr transform matrix.

Useful ReferenceBlackWhite values for YCbCr images are:

[0/1, 255/1,128/1, 255/1, 128/1, 255/1]

no headroom/footroom

[15/1, 235/1, 128/1, 240/1, 128/1, 240/1]

CCIR Recommendation 601.1 headroom/footroom

Useful ReferenceBlackWhite values for BitsPerSample = 8,8,8 Class R images are:

[0/1, 255/1,0/1, 255/1, 0/1, 255/1]

no headroom/footroom

[16/1, 235/1, 16/1, 235/1, 16/1, 235/1]

CCIR Recommendation 601.1 headroom/footroom

Default is [0/,NV/1, 0/1, NV/1, 0/1, NV/1] where NV = 2 ** BitsPerSample - 1.

References

[1] *The Reproduction of Colour in Photography, Printing and Television,* R. W. G. Hunt, Fountain Press, Tolworth, England,1987.

[2] *Principles of Color Technology*, Billmeyer and Saltzman, Wiley-Interscience, New York, 1981.

[3] *Colorimetric Properties of Video Displays*, William Cowan, University of Waterloo, Waterloo, Canada, 1989.

[4] *TIFF Color Appearance Guidelines*, Dave Farber, Eastman Kodak Company, Rochester, New York.

Section 21: YC_bC_r Images

Introduction

Digitizers of video sources that create RGB data are becoming more capable and less expensive. The RGB color space is adequate for this purpose. However, for both digital video and image compression applications a color difference color space is needed. The television industry depends on YC_bC_r for digital video. For image compression, subsampling the chrominance components allows for greater compression. TIFF YC_bC_r (which we shall call *Class Y*) supports these images and applications.

Class Y is based on CCIR Recommendation 601-1, "Encoding Parameters of Digital Television for Studios." Class Y also has parameters that allow the description of related standards such as CCIR Recommendation 709 and technological variations such as component-sample positioning.

YC_bC_r is a distinct PhotometricInterpretation type. RGB pixels are converted to and from YC_bC_r for storage and display.

Class Y defines the following fields:

YC_bC_rCoefficients	transformation from RGB to YC_bC_r
YC_bC_rSubSampling	subsampling of the chrominance components
YC_bC_rPositioning	positioning of chrominance component samples relative to the luminance samples

In addition, ReferenceBlackWhite, which specifies coding range expansion, is required by Class Y. See Section 20.

Class Y YC_bC_r images have three components: Y, the luminance component, and C_b and C_r, two chrominance components. Class Y uses the international standard notation YC_bC_r for color-difference component coding. This is often incorrectly called YUV, which properly applies only to composite coding.

The transformations between YC_bC_r and RGB are linear transformations of uninterpreted RGB sample data, typically gamma-corrected values. The YC_bC_rCoefficients field describes the parameters of this transformation.

Another feature of Class Y comes from subsampling the chrominance components. A Class Y image can be compressed by reducing the spatial resolution of chrominance components. This takes advantage of the relative insensitivity of the human visual system to chrominance detail. The YC_bC_rSubSampling field describes the degree of subsampling which has taken place.

When a Class Y image is subsampled, each C_b and C_r sample is associated with a group of luminance samples. The YC_bC_rPositioning field describes the position of the chrominance component samples relative to the group of luminance samples: centered or cosited.

Class Y requires use of the ReferenceBlackWhite field. This field expands the coding range by describing the reference black and white values for the different components that allow headroom and footroom for digital video images. Since the default for ReferenceBlackWhite is inappropriate for Class Y, it must be used explicitly.

At first, it might seem that the information conveyed by Class Y and the RGB Colorimetry section is redundant. However, decoding YC_bC_r to RGB primaries requires the YC_bC_r fields, and interpretation of the resulting RGB primaries requires the colorimetry and transfer function information. See the RGB Colorimetry section for details.

Extensions to Existing Fields

Class Y images use a distinct PhotometricInterpretation Field value:

PhotometricInterpretation

Tag = 262 (106.H)

Type = SHORT

N = 1

This Field indicates the color space of the image. The new value is:

6 = YC_bC_r

A value of 6 indicates that the image data is in the YC_bC_r color space. TIFF uses the international standard notation YC_bC_r for color-difference sample coding. Y is the luminance component. C_b and C_r are the two chrominance components. RGB pixels are converted to and from YC_bC_r form for storage and display.

Fields Defined in Class Y

YC_bC_rCoefficients

Tag = 529 (211.H)

Type = RATIONAL

N = 3

The transformation from RGB to YC_bC_r image data. The transformation is specified as three rational values that represent the coefficients used to compute luminance, Y.

The three rational coefficient values, *LumaRed*, *LumaGreen* and *LumaBlue*, are the proportions of red, green, and blue respectively in luminance, Y.

Y, C_b, and C_r may be computed from RGB using the luminance coefficients specified by this field as follows:

$$Y = (\mathit{LumaRed} * R + \mathit{LumaGreen} * G + \mathit{LumaBlue} * B)$$

$$C_b = (B - Y) / (2 - 2 * \mathit{LumaBlue})$$

$C_r = (R - Y) / (2 - 2 * LumaRed)$

R, G, and B may be computed from YC_bC_r as follows:

$R = C_r * (2 - 2 * LumaRed) + Y$

$G = (Y - LumaBlue * B - LumaRed * R) / LumaGreen$

$B = C_b * (2 - 2 * LumaBlue) + Y$

In disciplines such as printing, film, and video, there are practical reasons to provide footroom codes below the ReferenceBlack code and headroom codes above ReferenceWhite code. In such cases the values of the transformation matrix used to convert from YC_bC_r to RGB must be multiplied by a scale factor to produce full-range RGB values. These scale factors depend on the reference ranges specified by the ReferenceBlackWhite field. See the ReferenceBlackWhite and TransferFunction fields for more details.

The values coded by this field will typically reflect the transformation specified by a standard for YC_bC_r encoding. The following table contains examples of commonly used values.

Standard	*LumaRed*	*LumaGreen*	*LumaBlue*
CCIR Recommendation 601-1	299 / 1000	587 / 1000	114 / 1000
CCIR Recommendation 709	2125 / 10000	7154 / 10000	721 / 10000

The default values for this field are those defined by CCIR Recommendation 601-1: 299/1000, 587/1000 and 114/1000, for *LumaRed*, *LumaGreen* and *LumaBlue*, respectively.

YC_bC_rSubSampling

Tag = 530 (212.H)

Type = SHORT

N = 2

Specifies the subsampling factors used for the chrominance components of a YC_bC_r image. The two fields of this field, *YC_bC_rSubsampleHoriz* and *YC_bC_rSubsampleVert*, specify the horizontal and vertical subsampling factors respectively.

The two fields of this field are defined as follows:

Short 0: *YC_bC_rSubsampleHoriz*:

1 = ImageWidth of this chroma image is equal to the ImageWidth of the associated luma image.

2 = ImageWidth of this chroma image is half the ImageWidth of the associated luma image.

4 = ImageWidth of this chroma image is one-quarter the ImageWidth of the associated luma image.

Short 1: *YC_bC_rSubsampleVert*:

1 = ImageLength (height) of this chroma image is equal to the ImageLength of the associated luma image.

2 = ImageLength (height) of this chroma image is half the ImageLength of the associated luma image.

4 = ImageLength (height) of this chroma image is one-quarter the ImageLength of the associated luma image.

Both C_b and C_r have the same subsampling ratio. Also, *YC_bC_rSubsampleVert* shall always be less than or equal to *YC_bC_rSubsampleHoriz*.

ImageWidth and ImageLength are constrained to be integer multiples of *YC_bC_rSubsampleHoriz* and *YC_bC_rSubsampleVert* respectively. TileWidth and TileLength have the same constraints. RowsPerStrip must be an integer multiple of *YC_bC_rSubsampleVert*.

The default values of this field are [2, 2].

YC_bC_rPositioning

Tag = 531 (213.H)

Type = SHORT

N = 1

Specifies the positioning of subsampled chrominance components relative to luminance samples.

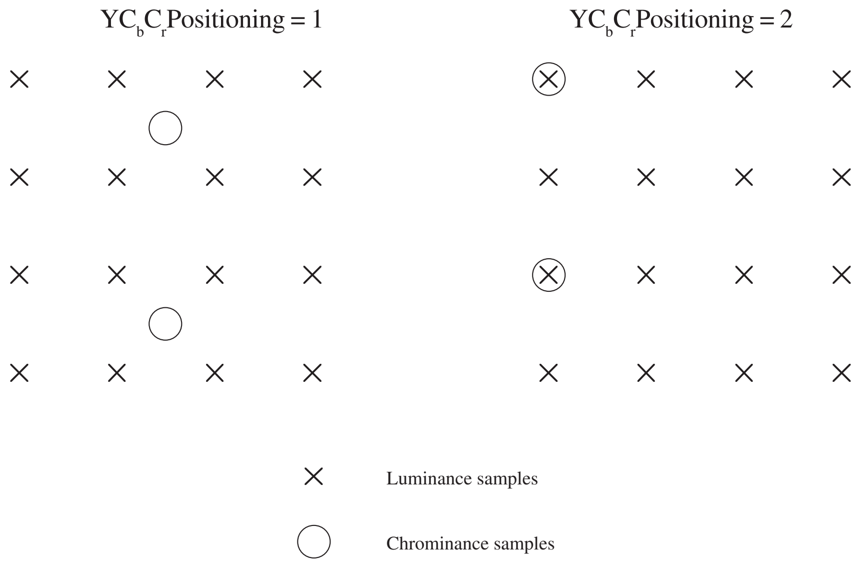
Specification of the spatial positioning of pixel samples relative to the other samples is necessary for proper image post processing and accurate image presentation. In Class Y files, the position of the subsampled chrominance components are defined with respect to the luminance component. Because components must be sampled orthogonally (along rows and columns), the spatial position of the samples in a given subsampled 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 luminance component. The horizontal and vertical offsets of the first chrominance sample are denoted Xoffset[0,0] and Yoffset[0,0] respectively. Xoffset[0,0] and Yoffset[0,0] are defined in terms of the number of samples in the luminance component.

The values for this field are defined as follows:

Tag value	YC_bC_r Positioning	X and Y offsets of first chrominance sample
1	centered	Xoffset[0,0] = *ChromaSubsampleHoriz* / 2 - 0.5 Yoffset[0,0] = *ChromaSubsampleVert* / 2 - 0.5
2	cosited	Xoffset[0,0] = 0 Yoffset[0,0] = 0

Field value 1 (centered) must be specified for compatibility with industry standards such as PostScript Level 2 and QuickTime. Field value 2 (cosited) must be specified for compatibility with most digital video standards, such as CCIR Recommendation 601-1.

As an example, for *ChromaSubsampleHoriz* = 4 and *ChromaSubsampleVert* = 2, the centers of the samples are positioned as illustrated below:



Proper subsampling of the chrominance components incorporates an anti-aliasing filter that reduces the spectral bandwidth of the full-resolution samples. The type of filter used for subsampling determines the value of the YC_bC_rPositioning field.

For YC_bC_rPositioning = 1 (centered), subsampling of the chrominance components can easily be accomplished using a symmetrical digital filter with an even number of taps (coefficients). A commonly used filter for 2:1 subsampling utilizes two taps (1/2,1/2).

For YC_bC_rPositioning = 2 (cosited), subsampling of the chrominance components can easily be accomplished using a symmetrical digital filter with an odd number of taps. A commonly used filter for 2:1 subsampling utilizes three taps (1/4,1/2,1/4).

The default value of this field is 1.

Ordering of Component Samples

This section defines the ordering convention used for Y, C_b, and C_r component samples when the PlanarConfiguration field value = 1 (interleaving). For PlanarConfiguration = 2, component samples are stored as 3 separate planes, and the ordering is the same as that used for other PhotometricInterpretation field values.

For PlanarConfiguration = 1, the component sample order is based on the subsampling factors, *ChromaSubsampleHoriz* and *ChromaSubsampleVert,* defined by the YC_bC_rSubSampling field. The image data within a TIFF file is comprised of one or more “data units”, where a data unit is defined to be a sequence of samples:

- one or more Y samples
- a C_b sample
- a C_r sample

The Y samples within a data unit are specified as a two-dimensional array having *ChromaSubsampleVert* rows of *ChromaSubsampleHoriz* samples.

Expanding on the example in the previous section, consider a YC_bC_r image having *ChromaSubsampleHoriz* = 4 and *ChromaSubsampleVert* = 2:

Y component

Y00	Y01	Y02	Y03	Y04	Y05		
Y10	Y11	Y12	Y13				

Cb component

Cb00	

Cr component

Cr00	

For PlanarConfiguration = 1, the sample order is:

$$Y_{00}, Y_{01}, Y_{02}, Y_{03}, Y_{10}, Y_{11}, Y_{12}, Y_{13}, Cb_{00}, Cr_{00}, Y_{04}, Y_{05} \ldots$$

Minimum Requirements for YCbCr Images

In addition to satisfying the general Baseline TIFF requirements, a YCbCr file must have the following characteristics:

- SamplesPerPixel = 3. SHORT. Three components representing Y, Cb and Cr.
- BitsPerSample = 8,8,8. SHORT.
- Compression = none (1), LZW (5) or JPEG (6). SHORT.
- PhotometricInterpretation = YC_bC_r (6). SHORT.
- ReferenceBlackWhite = 6 RATIONALS. Specify the reference values for black and white.

If the conversion from RGB is not according to CCIR Recommendation 601-1, code YC_bC_rCoefficients.

Section 22: JPEG Compression

Introduction

Image compression reduces the storage requirements of pictorial data. In addition, it reduces the time required for access to, communication with, and display of images. To address the standardization of compression techniques an international standards group was formed: the Joint Photographic Experts Group (JPEG). JPEG has as its objective to create a joint ISO/CCITT standard for continuous tone image compression (color and grayscale).

JPEG decided that because of the broad scope of the standard, no one algorithmic procedure was able to satisfy the requirements of all applications. It was decided to specify different algorithmic processes, where each process is targeted to satisfy the requirements of a class of applications. Thus, the JPEG standard became a “toolkit” whereby the particular algorithmic “tools” are selected according to the needs of the application environment.

The algorithmic processes fall into two classes: lossy and lossless. Those based on the Discrete Cosine Transform (DCT) are lossy and typically provide for substantial compression without significant degradation of the reconstructed image with respect to the source image.

The simplest DCT-based coding process is the baseline process. It provides a capability that is sufficient for most applications. There are additional DCT-based processes that extend the baseline process to a broader range of applications.

The second class of coding processes is targeted for those applications requiring lossless compression. The lossless processes are not DCT-based and are utilized independently of any of the DCT-based processes.

This Section describes the JPEG baseline, the JPEG lossless processes, and the extensions to TIFF defined to support JPEG compression.

JPEG Baseline Process

The baseline process is a DCT-based algorithm that compresses images having 8 bits per component. The baseline process operates only in sequential mode. In sequential mode, the image is processed from left to right and top to bottom in a single pass by compressing the first row of data, followed by the second row, and continuing until the end of image is reached. Sequential operation has minimal buffering requirements and thus permits inexpensive implementations.

The JPEG baseline process is an algorithm which inherently introduces error into the reconstructed image and cannot be utilized for lossless compression. The algorithm accepts as input only those images having 8 bits per component. Images with fewer than 8 bits per component may be compressed using the baseline process algorithm by left justifying each input component within a byte before compression.

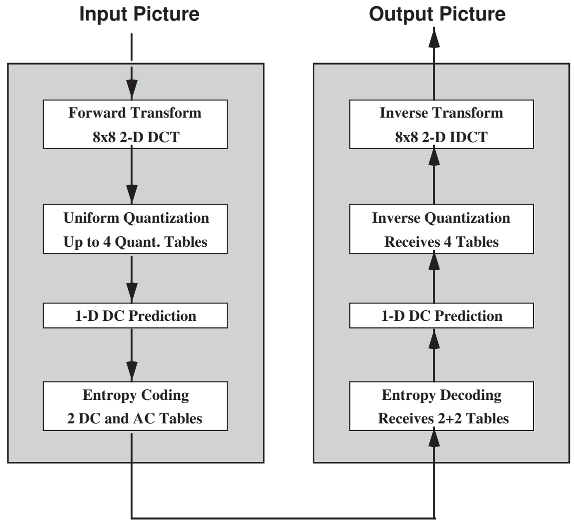


Figure 1. Baseline Process Encoder and Decoder

A functional block diagram of the Baseline encoding and decoding processes is contained in Figure 1. Encoder operation consists of dividing each component of the input image into 8x8 blocks, performing the two-dimensional DCT on each block, quantizing each DCT coefficient uniformly, subtracting the quantized DC coefficient from the corresponding term in the previous block, and then entropy coding the quantized coefficients using variable length codes (VLCs). Decoding is performed by inverting each of the encoder operations in the reverse order.

The DCT

Before performing the foward DCT, input pixels are level-shifted so that they range from -128 to +127. Blocks of 8x8 pixels are transformed with the two-dimensional 8x8 DCT:

$$F(u,v) = \frac{1}{4} C(u)C(v) \sum \sum f(x,y) \cos \frac{\pi(2x+1)u}{16} \cos \frac{\pi(2y+1)v}{16}$$

and blocks are inverse transformed by the decoder with the Inverse DCT:

$$f(x,y) = \frac{1}{4} \sum \sum C(u)C(v) F(u,v) \cos \frac{\pi(2x+1)u}{16} \cos \frac{\pi(2y+1)v}{16}$$

with $u, v, x, y = 0, 1, 2, \ldots 7$

where x, y = spatial coordinates in the pel domain
u, v = coordinates in the transform domain

$$C(u), C(v) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } u, v = 0 \\ 1 & \text{otherwise} \end{cases}$$

Although the exact method for computation of the DCT and IDCT is not subject to standardization and will not be specified by JPEG, it is probable that JPEG will adopt DCT-conformance specifications that designate the accuracy to which the DCT must be computed. The DCT-conformance specifications will assure that any two JPEG implementations will produce visually-similar reconstructed images.

Quantization

The coefficients of the DCT are quantized to reduce their magnitude and increase the number of zero-value coefficients. The DCT coefficients are independently quantized by uniform quantizers. A uniform quantizer divides the real number line into steps of equal size, as shown in Figure 2. The quantization step-size applied to each coefficient is determined from the contents of a 64-element quantization table.

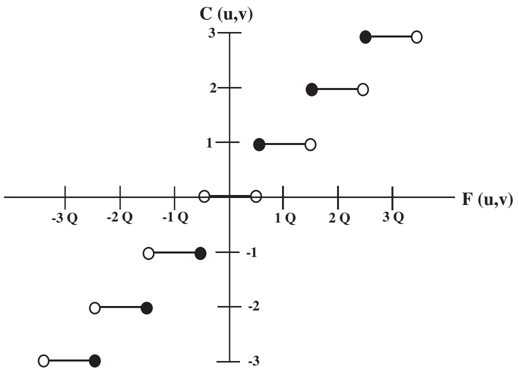


Figure 2. Uniform Quantization

The baseline process provides for up to 4 different quantization tables to be defined and assigned to separate interleaved components within a single scan of the input image. Although the values of each quantization table should ideally be determined through rigorous subjective testing which estimates the human psycho-visual thresholds for each DCT coefficient and for each color component of the input image, JPEG has developed quantization tables which work well for CCIR 601 resolution images and has published these in the informational section of the proposed standard.

DC Prediction

The DCT coefficient located in the upper-left corner of the transformed block represents the average spatial intensity of the block and is referred to as the “DC coefficient”. After the DCT coefficients are quantized, but before they are entropy coded, DC prediction is performed. DC prediction simply means that the DC term of the previous block is subtracted from the DC term of the current block prior to encoding.

Zig-Zag Scan

Prior to entropy coding, the DCT coefficients are ordered into a one-dimensional sequence according to a "zig-zag" scan. The DC coefficient is coded first, followed by AC coefficient coding, proceeding in the order illustrated in Figure 3.

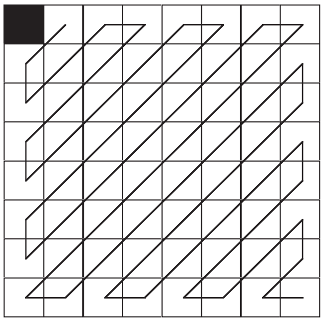


Figure 3. Zig-Zag Scan of DCT Coefficients

Entropy Coding

The quantized DCT coefficients are further compressed using entropy coding. The baseline process performs entropy coding using variable length codes (VLCs) and variable length integers (VLIs).

VLCs, commonly known as Huffman codes, compress data symbols by creating shorter codes to represent frequently-occurring symbols and longer codes for occasionally-occurring symbols. One reason for using VLCs is that they are easily implemented by means of lookup tables.

Separate code tables are provided for the coding of DC and AC coefficients. The following paragraphs describe the respective coding methods used for coding DC and AC coefficients.

DC Coefficient Coding

DC prediction produces a "differential DC coefficient" that is typically small in magnitude due to the high correlation of neighboring DC coefficients. Each differential DC coefficient is encoded by a VLC which represents the number of significant bits in the DC term followed by a VLI representing the value itself. The VLC is coded by first determining the number of significant bits, SSSS, in the differential DC coefficient through the following table:

SSSS	Differential DC Value
0	0
1	-1, 1
2	-3,-2, 2,3
3	-7..-4, 4..7
4	-15..-8, 8..15
5	-31..-16, 16..31

6	-63..-32, 32..63
7	-127..-64, 64..127
8	-255..-128, 128..255
9	-511..-256, 256..511
10	-1023..-512, 512..1023
11	-2047..-1024, 1024..2047
12	-4095..-2048, 2048..4095

SSSS is then coded from the selected DC VLC table. The VLC is followed by a VLI having SSSS bits that represents the value of the differential DC coefficient itself. If the coefficient is positive, the VLI is simply the low-order bits of the coefficient. If the coefficient is negative, then the VLI is the low-order bits of the coefficient-1.

AC Coefficient Coding

In a similar fashion, AC coefficients are coded with alternating VLC and VLI codes. The VLC table, however, is a two-dimensional table that is indexed by a composite 8-bit value. The lower 4 bits of the 8-bit value, i.e. the column index, is the number of significant bits, SSSS, of a non-zero AC coefficient. SSSS is computed through the same table as that used for coding the DC coefficient. The higher-order 4 bits, the row index, is the number of zero coefficients, NNNN, that precede the non-zero AC coefficient. The first column of the two-dimensional coding table contains codes that represent control functions. Figure 4 illustrates the general structure of the AC coding table.

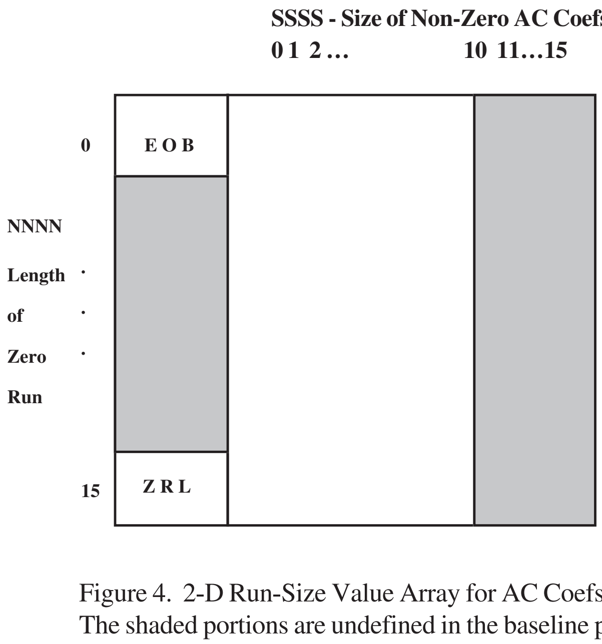


Figure 4. 2-D Run-Size Value Array for AC Coefs
The shaded portions are undefined in the baseline process

The flow chart in Figure 5 specifies the AC coefficient coding procedure. AC coefficients are coded by traversing the block in the zig-zag sequence and counting the number of zero coefficients until a non-zero AC coefficient is encountered. If the count of consecutive zero coefficients exceeds 15, then a ZRL code is coded and the zero run-length count is reset. When a non-zero AC coefficient is found, the number of significant bits in the non-zero coefficient, SSSS, is combined with the zero run-length that precedes the coefficient, NNNN, to form an index into the two-dimensional VLC table. The selected VLC is then coded. The VLC is followed by a VLI that represents the value of the AC coefficient. This process is repeated until the end of the block is reached. If the last AC coefficient is zero, then an End of Block (EOB) VLC is encoded.

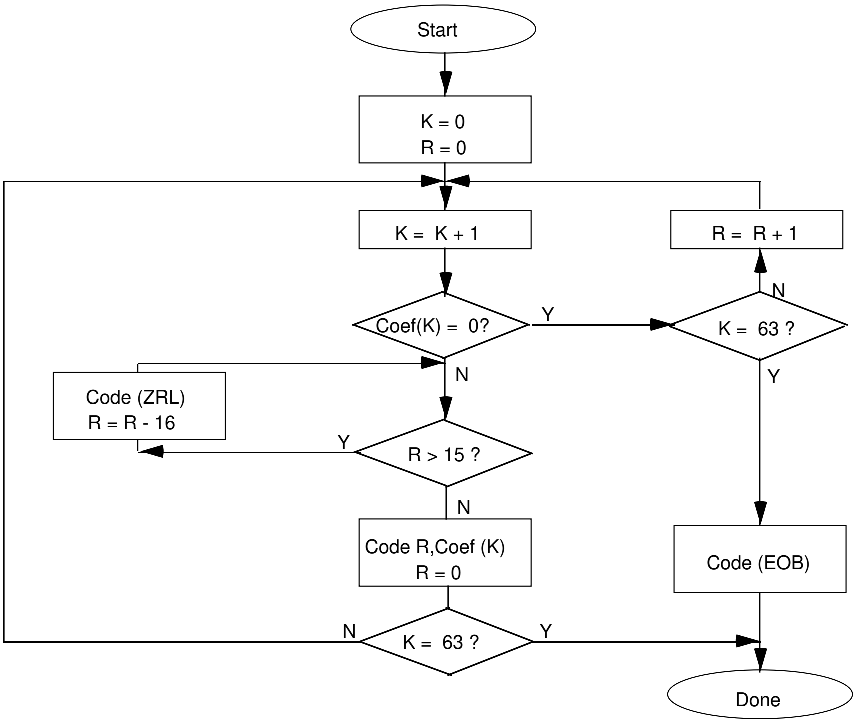


Figure 5. Encoding Procedure for AC Coefs

JPEG Lossless Processes

The JPEG lossless coding processes utilize a spatial-prediction algorithm based upon a two-dimensional Differential Pulse Code Modulation (DPCM) technique. They are compatible with a wider range of input pixel precision than the DCT-based algorithms (2 to 16 bits per component). Although the primary motivation for specifying a spatial algorithm is to provide a method for lossless compression, JPEG allows for quantization of the input data, resulting in lossy compression and higher compression rates.

Although JPEG provides for use of either the Huffman or Arithmetic entropy-coding models by the processes for lossless coding, only the Huffman coding model is supported by this version of TIFF. The following is a brief overview of the lossless process with Huffman coding.

Control Structure

Much of the control structure developed for the sequential DCT procedures is also used for sequential lossless coding. Either interleaved or non-interleaved data ordering may be used.

Coding Model

The coding model developed for coding the DC coefficients of the DCT is extended to allow a number of one-dimensional and two-dimensional predictors for the lossless coding function. Each component uses an independent predictor.

Prediction

Figure 6 shows the relationship between the neighboring values used for prediction and the sample being coded.

```
   |   |   |   |   |
 --+---+---+---+---+--
   |   | C | B |   |
 --+---+---+---+---+--
   |   | A | Y |   |
 --+---+---+---+---+--
   |   |   |   |   |
```

Figure 6. Relationship between sample and prediction samples

Y is the sample to be coded and A, B, and C are the samples immediately to the left, immediately above, and diagonally to the left and above.

The allowed predictors are listed in the following table.

Selection-value	Prediction
0	no prediction (differential coding)
1	A
2	B
3	C
4	A+B-C
5	A+((B-C)/2)
6	B+((A-C)/2)
7	(A+B)/2

Selection-value 0 shall only be used for differential coding in the hierarchical mode. Selections 1, 2 and 3 are one-dimensional predictors and selections 4, 5, 6, and 7 are two dimensional predictors. The divide by 2 in the prediction equations is done by a arithmetic-right-shift of the integer values.

The difference between the prediction value and the input is calculated modulo 2**16. Therefore, the prediction can also be treated as a modulo 2**16 value. In the decoder the difference is decoded and added, modulo 2**16, to the prediction.

Huffman Coding of the Prediction Error

The Huffman coding procedures defined for coding the DC coefficients are used to code the modulo 2**16 differences. The table for DC coding is extended to 17 entries that allows for coding of the modulo 2**16 differences.

Point Transformation Prior to Lossless Coding

For the lossless processes only, the input image data may optionally be scaled (quantized) prior to coding by specifying a nonzero value in the point transformation parameter. Point transformation is defined to be division by a power of 2.

If the point transformation field is nonzero for a component, a point transformation of the input is performed prior to the lossless coding. The input is divided by 2**Pt, where Pt is the value of the point transform signaling field. The output of the decoder is rescaled to the input range by multiplying by 2**Pt. Note that the scaling of input and output can be performed by arithmetic shifts.

Overview of the JPEG Extension to TIFF

In extending the TIFF definition to include JPEG compressed data, it is necessary to note the following:

- JPEG is effective only on continuous-tone color spaces:

Grayscale	(Photometric Interpretation = 1)	
RGB	(Photometric Interpretation = 2)	
CMYK	(Photometric Interpretation = 5)	(See the CMYK Images section.)
YC_bC_r	(Photometric Interpretation = 6)	(See the YCbCr images section.)

- Color conversion to YC_bC_r is often used as part of the compression process because the chrominance components can be subsampled and compressed to a greater degree without significant visual loss of quality. Fields are defined to describe how this conversion has taken place and the degree of subsampling employed (see the YCbCr Images section).

- New fields have been defined to specify the JPEG parameters used for compression and to allow quantization tables and Huffman code tables to be incorporated into the TIFF file.

- TIFF is compatible with compressed image data that conforms to the syntax of the JPEG interchange format for compressed image data. Fields are defined that may be utilized to facilitate conversion from TIFF to interchange format.

- The PlanarConfiguration Field is used to specify whether or not the compressed data is interleaved as defined by JPEG. For any of the JPEG DCT-based processes, the interleaved data units are coded 8x8 blocks rather than component samples.

- Although JPEG codes consecutive image blocks in a single contiguous bitstream, it is extremely useful to employ the concept of tiles in an image. The TIFF Tiles section defines some new fields for tiles. These fields should be stored in place of the older fields for strips. The concept of tiling an image in both dimensions is important because JPEG hardware may be limited in the size of each block that is handled.

- Note that the nomenclature used in the TIFF specification is different from the JPEG Draft International Standardittee Draft (ISO DIS 10918-1) in some respects. The following terms should be equated when reading this Section:

TIFF name	JPEG DIS name
ImageWidth	Number of Pixels
ImageLength	Number of Lines
SamplesPerPixel	Number of Components
JPEGQTable	Quantization Table
JPEGDCTable	Huffman Table for DC coefficients
JPEGACTable	Huffman Table for AC coefficients

Strips and Tiles

The JPEG extension to TIFF has been designed to be consistent with the existing TIFF strip and tile structures and to allow quick conversion to and from the stream-oriented compressed image format defined by JPEG.

Compressed images conforming to the syntax of the JPEG interchange format can be converted to TIFF simply by defining a single strip or tile for the entire image and then concatenating the TIFF image description fields to the JPEG compressed image data. The strip or tile offset field points directly to the start of the entropy coded data (not to a JPEG marker).

Multiple strips or tiles are supported in JPEG compressed images using restart markers. Restart markers, inserted periodically into the compressed image data, delineate image segments known as restart intervals. At the start of each restart interval, the coding state is reset to default values, allowing every restart interval to be decoded independently of previously decoded data. TIFF strip and tile offsets shall always point to the start of a restart interval. Equivalently, each strip or tile contains an integral number of restart intervals. Restart markers need not be present in a TIFF file; they are implicitly coded at the start of every strip or tile.

To maximize interchangeability of TIFF files with other formats, a restriction is placed on tile height for files containing JPEG-compressed image data conforming to the JPEG interchange format syntax. The restriction, imposed only when the tile width is shorter than the image width and when the JPEGInterchangeFormat Field is present and non-zero, states that the tile height must be equal to the height of one JPEG Minimum Coded Unit (MCU). This restriction ensures that TIFF files may be converted to JPEG interchange format without undergoing decompression.

Extensions to Existing Fields

Compression

Tag = 259 (103.H)

Type = SHORT

N = 1

This Field indicates the type of compression used. The new value is:

6 = JPEG

JPEG Fields

JPEGProc

Tag = 512 (200.H)

Type = SHORT

N = 1

This Field indicates the JPEG process used to produce the compressed data. The values for this field are defined to be consistent with the numbering convention used in ISO DIS 10918-2. Two values are defined at this time.

1= Baseline sequential process

14= Lossless process with Huffman coding

When the lossless process with Huffman coding is selected by this Field, the Huffman tables used to encode the image are specified by the JPEGDCTables field, and the JPEGACTables field is not used.

Values indicating JPEG processes other than those specified above will be defined in the future.

Not all of the fields described in this section are relevant to the JPEG process selected by this Field. The following table specifies the fields that are applicable to each value defined by this Field.

Tag Name	JPEGProc =1	JPEGProc =14
JPEGInterchangeFormat	X	X
JPEGInterchangeFormatLength	X	X
JPEGRestart Interval	X	X
JPEGLosslessPredictors		X
JPEGPointTransforms		X
JPEGQTables	X	
JPEGDCTables	X	X
JPEGACTables	X	

This Field is mandatory whenever the Compression Field is JPEG (no default).

JPEGInterchangeFormat

Tag = 513 (201.H)

Type = LONG

N = 1

This Field indicates whether a JPEG interchange format bitstream is present in the TIFF file. If a JPEG interchange format bitstream is present, then this Field points to the Start of Image (SOI) marker code.

If this Field is zero or not present, a JPEG interchange format bitstream is not present.

JPEGInterchangeFormatLength

Tag = 514 (202.H)

Type = LONG

N = 1

This Field indicates the length in bytes of the JPEG interchange format bitstream. This Field is useful for extracting the JPEG interchange format bitstream without parsing the bitstream.

This Field is relevant only if the JPEGInterchangeFormat Field is present and is non-zero.

JPEGRestartInterval

Tag = 515 (203.H)

Type = SHORT

N = 1

This Field indicates the length of the restart interval used in the compressed image data. The restart interval is defined as the number of Minimum Coded Units (MCUs) between restart markers.

Restart intervals are used in JPEG compressed images to provide support for multiple strips or tiles. At the start of each restart interval, the coding state is reset to default values, allowing every restart interval to be decoded independently of previously decoded data. TIFF strip and tile offsets shall always point to the start of a restart interval. Equivalently, each strip or tile contains an integral number of restart intervals. Restart markers need not be present in a TIFF file; they are implicitly coded at the start of every strip or tile.

See the JPEG Draft International Standard (ISO DIS 10918-1) for more information about the restart interval and restart markers.

If this Field is zero or is not present, the compressed data does not contain restart markers.

JPEGLosslessPredictors

Tag = 517 (205.H)

Type = SHORT

N = SamplesPerPixel

This Field points to a list of lossless predictor-selection values, one per component.
The allowed predictors are listed in the following table.

Selection-value	Prediction
1	A
2	B
3	C
4	A+B-C
5	A+((B-C)/2)
6	B+((A-C)/2)
7	(A+B)/2

A, B, and C are the samples immediately to the left, immediately above, and diagonally to the left and above the sample to be coded, respectively.

See the JPEG Draft International Standard (ISO DIS 10918-1) for more details.

This Field is mandatory whenever the JPEGProc Field specifies one of the lossless processes (no default).

JPEGPointTransforms

Tag = 518 (206.H)

Type = SHORT

N = SamplesPerPixel

This Field points to a list of point transform values, one per component. This Field is relevant only for lossless processes.

If the point transformation value is nonzero for a component, a point transformation of the input is performed prior to the lossless coding. The input is divided by 2**Pt, where Pt is the point transform value. The output of the decoder is rescaled to the input range by multiplying by 2**Pt. Note that the scaling of input and output can be performed by arithmetic shifts.

See the JPEG Draft International Standard (ISO DIS 10918-1) for more details. The default value of this Field is 0 for each component (no scaling).

JPEGQTables

Tag = 519 (207.H)

Type = LONG

N = SamplesPerPixel

This Field points to a list of offsets to the quantization tables, one per component. Each table consists of 64 BYTES (one for each DCT coefficient in the 8x8 block). The quantization tables are stored in zigzag order.

See the JPEG Draft International Standard (ISO DIS 10918-1) for more details.

It is strongly recommended that, within the TIFF file, each component be assigned separate tables. This Field is mandatory whenever the JPEGProc Field specifies a DCT-based process (no default).

JPEGDCTables

Tag = 520 (208.H)

Type = LONG

N = SamplesPerPixel

This Field points to a list of offsets to the DC Huffman tables or the lossless Huffman tables, one per component.

The format of each table is as follows:

> 16 BYTES of "BITS", indicating the number of codes of lengths 1 to 16;
>
> Up to 17 BYTES of "VALUES", indicating the values associated with those codes, in order of length.

See the JPEG Draft International Standard (ISO DIS 10918-1) for more details.

It is strongly recommended that, within the TIFF file, each component be assigned separate tables. This Field is mandatory for all JPEG processes (no default).

JPEGACTables

Tag = 521 (209.H)

Type = LONG

N = SamplesPerPixel

This Field points to a list of offsets to the Huffman AC tables, one per component. The format of each table is as follows:

> 16 BYTES of “BITS”, indicating the number of codes of lengths 1 to 16;
>
> Up to 256 BYTES of “VALUES”, indicating the values associated with those codes, in order of length.

See the JPEG Draft International Standard (ISO DIS 10918-1) for more details.

It is strongly recommended that, within the TIFF file, each component be assigned separate tables. This Field is mandatory whenever the JPEGProc Field specifies a DCT-based process (no default).

Minimum Requirements for TIFF with JPEG Compression

The table on the following page shows the minimum requirements of a TIFF file that uses tiling and contains JPEG data compressed with the Baseline process.

Tag = NewSubFileType (254) Type = Long Length = 1 Value = 0	Single image
Tag = ImageWidth (256) Type = Long Length = 1 Value = ?	
Tag = ImageLength (257) Type = Long Length = 1 Value = ?	
Tag = BitsPerSample (258) Type = Short Length = SamplesPerPixel Value = ?	8 : Monochrome 8,8,8 : RGB 8,8,8 : YCbCr 8,8,8,8 : CMYK
Tag = Compression (259) Type = Long Length = 1 Value = 6	6 : JPEG compression
Tag = PhotometricInterpretation (262) Type = Short Length = 1 Value = ?	0,1 : Monochrome 2 : RGB 5 : CMYK 6 : YCbCr
Tag = SamplesPerPixel (277) Type = Short Length = 1 Value = ?	1 : Monochrome 3 : RGB 3 : YCbCr 4 : CMYK
Tag = XResolution (282) Type = Rational Length = 1 Value = ?	
Tag = YResolution (283) Type = Rational Length = 1 Value = ?	
Tag = PlanarConfiguration (284) Type = Short Length = 1 Value = ?	1 : Block Interleaved 2 : Not interleaved
Tag = ResolutionUnit (296) Type = Short Length = 1 Value = ?	
Tag = TileWidth (322) Type = Short Length = 1 Value = ?	Multiple of 8
Tag = TileLength (323) Type = Short Length = 1 Value = ?	Multiple of 8
Tag = TileOffsets (324) Type = Long Length = Number of tiles Value = ?	
Tag = TileByteCounts (325) Type = Long Length = Number of tiles Value = ?	
Tag = JPEGProc (512) Type = Short Length = 1 Value = ?	1 : Baseline process
Tag = JPEGQTables (519) Type = Long Length = SamplesPerPixel Value = ?	Offsets to tables
Tag = JPEGDCTables (520) Type = Long Length = SamplesPerPixel Value = ?	Offsets to tables
Tag = JPEGACTables (521) Type = Long Length = SamplesPerPixel Value = ?	Offsets to tables

References

[1] Wallace, G., "Overview of the JPEG Still Picture Compression Algorithm", Electronic Imaging East '90.

[2] ISO/IEC DIS 10918-1, "Digital Compression and Coding of Continuous-tone Still Images", Sept. 1991.

Section 23: CIE L*a*b* Images

*What is CIE L*a*b*?*

CIE La*b* is a color space that is colorimetric, has separate lightness and chroma channels, and is approximately perceptually uniform. It has excellent applicability for device-independent manipulation of continuous tone images. These attributes make it an excellent choice for many image editing functions.

1976 CIEL*a*b* is represented as a Euclidean space with the following three quantities plotted along axes at right angles: L^* representing lightness, a^* representing the red/green axis, and b^* representing the yellow/blue axis. The formulas for 1976 CIE $L^*a^*b^*$ follow:

$L^*=116(Y/Y_n)^{1/3}-16$ for $Y/Y_n > 0.008856$

$L^*=903.3(Y/Y_n)$ for $Y/Y_n <= 0.008856$ *see note below.

$$a^*=500\left[(X/X_n)^{1/3}-(Y/Y_n)^{1/3}\right]$$

$$b^*=200\left[(Y/Y_n)^{1/3}-(Z/Z_n)^{1/3}\right].$$

where X_n,Y_n, and Z_n are the CIE X, Y, and Z tristimulus values of an *appropriate* reference white. Also, if any of the ratios X/X_n, Y/Y_n, or Z/Z_n is equal to or less than 0.008856, it is replaced in the formulas with

$$7.787F + 16/116,$$

where F is X/X_n, Y/Y_n, or Z/Z_n, as appropriate (note: these low-light conditions are of no relevance for most document-imaging applications). Tiff is defined such that each quantity be encoded with 8 bits. This provides 256 levels of L^* lightness; 256 levels (+/- 127) of a^*, and 256 levels (+/- 127) of b^*. Dividing the 0-100 range of L^* into 256 levels provides lightness steps that are less than half the size of a "just noticeable difference". This eliminates banding, even under conditions of substantial tonal manipulation. Limiting the theoretically unbounded a^* and b^* ranges to +/- 127 allows encoding in 8 bits without eliminating any but the most saturated self-luminous colors. It is anticipated that the rare specialized applications requiring support of these extreme cases would be unlikely to use CIELAB anyway. All object colors, in fact all colors within the theoretical MacAdam limits, fall within the +/- 127 a^*/b^* range.

The TIFF CIELAB Fields

PhotometricInterpretation

Tag = 262 (106.H)

Type = SHORT

N = 1

8 = 1976 CIE *L*a*b**

Usage of other Fields.

BitsPerSample: 8

SamplesPerPixel - ExtraSamples: 3 for *L*a*b**, 1 implies *L** only, for monochrome data.

Compression: same as other multi-bit formats. JPEG compression applies.

PlanarConfiguration: both chunky and planar data could be supported.

WhitePoint: does not apply

PrimaryChromaticities: does not apply.

TransferFunction: does not apply

Alpha Channel information will follow the lead of other data types.

The reference white for this data type is the *perfect reflecting diffuser* (100% diffuse reflectance at all visible wavelengths). The *L** range is from 0 (perfect absorbing black) to 100 (perfect reflecting diffuse white). The *a** and *b** ranges will be represented as signed 8 bit values having the range -127 to +127.

Converting between RGB and CIELAB, a Caveat

The above CIELAB formulae are derived from CIE *XYZ*. Converting from CIELAB to *RGB* requires an additional set of formulae for converting between *RGB* and *XYZ*. For standard NTSC primaries these are:

$$\begin{array}{ccc} 0.6070 & 0.1740 & 0.2000 \\ 0.2990 & 0.5870 & 0.1140 \\ 0.0000 & 0.0660 & 1.1110 \end{array} \quad * \quad \begin{array}{c} R \\ G \\ B \end{array} \quad = \quad \begin{array}{c} X \\ Y \\ Z \end{array}$$

Generally, D65 illumination is used and a perfect reflecting diffuser is used for the reference white.

Since CIELAB is not a directly displayable format, some conversion to RGB will be required. While look-up table accelerated CIELAB to RGB conversion is certainly possible and fast, TIFF writers may choose to include a low resolution RGB subfile as an integral part of TIFF CIELAB.

Color Difference Measurements in CIELAB

The differences between two colors in L^*, a^*, and b^* are denoted by DL^*, Da^*, and Db^*, respectively, with the total (3-dimensional) color difference represented as:

$$\Delta E^*_{ab} = \left[(\Delta E^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2\right]^{1/2}.$$

This color difference can also be expressed in terms of L^*, C^*, and a measure of hue. In this case, h_{ab} is *not* used because it is an angular measure and cannot be combined with L^* and C^* directly. A linear-distance form of hue is used instead:

CIE 1976 *a,b hue-difference,* ΔH^*_{ab}

$$\Delta H^*_{ab} = \left[(\Delta E^*)^2 - (\Delta L^*)^2 - (\Delta C^*)^2\right]^{1/2}.$$

where DC^* is the chroma difference between the two colors. The total color difference expression using this hue-difference is:

$$\Delta E^*_{ab} = \left[(\Delta L^*)^2 + (\Delta H^*)^2 + (\Delta b^*)^2\right]^{1/2}.$$

It is important to remember that color difference is 3-dimensional: much more can be learned from a DL*a*b* triplet than from a single DE value. The DL*C*H* form is often the most useful since it gives the error information in a form that has more familiar perception correlates. Caution is in order, however, when using DH* for large hue differences since it is a straight-line approximation of a curved hue distance.

The Merits of CIELAB

Colorimetric.

First and foremost, CIELAB is colorimetric. It is traceable to the internationally-recognized standard CIE 1931 Standard Observer. This insures that it encodes color in a manner that is accurately modeled after the human vision system. Colors seen as matching are encoded identically, and colors seen as not matching are encoded differently. CIELAB provides an unambiguous definition of color without the necessity of additional information such as with RGB (primary chromaticities, white point, and gamma curves).

Device Independent.

Unlike RGB spaces which associate closely with physical phosphor colors, CIELAB contains no device association. CIELAB is not tailored for one device or device type at the expense of all others.

Full Color Gamut.

Any one image or imaging device usually encounters a very limited subset of the entire range of humanly-perceptible color. Collectively, however, these images and devices span a much larger gamut of color. A truly versatile exchange color space should encompass all of these colors, ideally providing support for all visible color. RGB, PhotoYCC, YCbCr, and other display spaces suffer from gamut limitations that exclude significant regions of easily printable colors. CIELAB is defined for all visible color.

Efficiency

A good exchange space will maximize accuracy of translations between itself and other spaces. It will represent colors compactly for a given accuracy. These attributes are provided through visual uniformity. One of the greatest disadvantages of the classic CIE system (and RGB systems as well) is that colors within it are not equally spaced visually. Encoding full-color images in a linear-intensity space, such as the typical RGB space or, especially, the XYZ space, requires a very large range (greater than 8-bits/primary) to eliminate banding artifacts. Adopting a *non*-linear RGB space improves the efficiency but not nearly to the extent as with a perceptually uniform space where these problems are nearly eliminated. A uniform space is also more efficiently compressed (see below).

Public Domain / Single Standard

CIELAB maintains no preferential attachments to any private organization. Its existence as a single standard leaves no room for ambiguity. Since 1976, CIELAB has continually gained popularity as a widely-accepted and heavily-used standard.

Luminance/Chrominance Separation.

The advantages for image size compression made possible by having a separate lightness or luminance channel are immense. Many such spaces exist. The degree to which the luminance information is fully-isolated into a single channel is an important consideration. Recent studies (Kasson and Plouffe of IBM) support CIELAB as a leading candidate placing it above CIELUV, YIQ, YUV, YCC, and XYZ.

Other advantages support a separate lightness or luminance channel. Tone and contrast editing and detail enhancement are most easily accomplished with such a channel. Conversion to a black and white representation is also easiest with this type of space.

When the chrominance channels are encoded as opponents as with CIELAB, there are other compression, image manipulation, and white point handling advantages.

Compressibility (Data).

Opponent spaces such as CIELAB are inherently more compressible than tristimulus spaces such as RGB. The chroma content of an image can be compressed to a greater extent, without objectionable loss, than can the lightness content. The opponent arrangement of CIELAB allows for spatial subsampling and efficent compression using JPEG.

Compressibility (Gamut).

Adjusting the color range of an image to match the capabilities of the intended output device is a critical function within computational color reproduction. Luminance/chrominance separation, especially when provided in a polar form, is desirable for facilitating gamut compression. Accurate gamut compression in a tri-linear color space is difficult.

CIELAB has a polar form (*metric hue angle,* and *metric chroma*, described below) that serves compression needs fairly well. Because CIELAB is not perfectly uniform, problems can arise when compressing along constant hue lines. Noticeable hue errors are sometimes introduced. This problem is no less severe with other contending color spaces.

This polar form also provides advantages for local color editing of images. The polar form is not proposed as part of the TIFF addition.

Getting the Most from CIELAB

Image Editors

The advantages of image editing within a perceptually uniform polar color space are tremendous. A detailed description of these advantages is beyond the scope of this section. As previously mentioned, many common tonal manipulation tasks are most efficiently performed when only a single channel is affected. Edge enhancement, contrast adjustment, and general tone-curve manipulation all ideally affect only the lightness component of an image.

A perceptual polar space works excellently for specifying a color range for masking purposes. For example, a red shirt can be quickly changed to a green shirt without drawing an outline mask. The operation can be performed with a loosely, quickly-drawn mask region combined with a hue (and perhaps chroma) range that encompasses the shirt's colors. The hue component of the shirt can then be adjusted, leaving the lightness and chroma detail in place.

Color cast adjustment is easily realized by shifting either or both of the chroma channels over the entire image or blending them over the region of interest.

Converting from CIELAB to a device specific space

For fast conversion to an RGB display, CIELAB can be decoded using 3x3 matrixing followed by gamma correction. The computational complexity required for accurate CRT display is the same with CIELAB as with extended luminance-chrominance spaces.

Converting CIELAB for accurate printing on CMYK devices requires computational complexity no greater than with *accurate* conversion from any other colorimetric space. Gamut compression becomes one of the more significant tasks for any such conversion.

Part 3: Appendices

Part 3 contains additional information that is not part of the TIFF specification, but may be of use to developers.

Appendix A: TIFF Tags Sorted by Number

TagName	Decimal	Hex	Type	Number of values
NewSubfileType	254	FE	LONG	1
SubfileType	255	FF	SHORT	1
ImageWidth	256	100	SHORT or LONG	1
ImageLength	257	101	SHORT or LONG	1
BitsPerSample	258	102	SHORT	SamplesPerPixel
Compression	259	103	SHORT	1
Uncompressed	1			
CCITT 1D	2			
Group 3 Fax	3			
Group 4 Fax	4			
LZW	5			
JPEG	6			
PackBits	32773			
PhotometricInterpretation	262	106	SHORT	1
WhiteIsZero	0			
BlackIsZero	1			
RGB	2			
RGB Palette	3			
Transparency mask	4			
CMYK	5			
YCbCr	6			
CIELab	8			
Threshholding	263	107	SHORT	1
CellWidth	264	108	SHORT	1
CellLength	265	109	SHORT	1
FillOrder	266	10A	SHORT	1
DocumentName	269	10D	ASCII	
ImageDescription	270	10E	ASCII	
Make	271	10F	ASCII	
Model	272	110	ASCII	
StripOffsets	273	111	SHORT or LONG	StripsPerImage
Orientation	274	112	SHORT	1
SamplesPerPixel	277	115	SHORT	1
RowsPerStrip	278	116	SHORT or LONG	1
StripByteCounts	279	117	LONG or SHORT	StripsPerImage
MinSampleValue	280	118	SHORT	SamplesPerPixel
MaxSampleValue	281	119	SHORT	SamplesPerPixel
XResolution	282	11A	RATIONAL	1
YResolution	283	11B	RATIONAL	1
PlanarConfiguration	284	11C	SHORT	1
PageName	285	11D	ASCII	
XPosition	286	11E	RATIONAL	
YPosition	287	11F	RATIONAL	
FreeOffsets	288	120	LONG	
FreeByteCounts	289	121	LONG	
GrayResponseUnit	290	122	SHORT	1

GrayResponseCurve	291	123	SHORT	2**BitsPerSample
T4Options	292	124	LONG	1
T6Options	293	125	LONG	1
ResolutionUnit	296	128	SHORT	1
PageNumber	297	129	SHORT	2
TransferFunction	301	12D	SHORT	{1 or SamplesPerPixel}* 2** BitsPerSample
Software	305	131	ASCII	
DateTime	306	132	ASCII	20
Artist	315	13B	ASCII	
HostComputer	316	13C	ASCII	
Predictor	317	13D	SHORT	1
WhitePoint	318	13E	RATIONAL	2
PrimaryChromaticities	319	13F	RATIONAL	6
ColorMap	320	140	SHORT	3 * (2**BitsPerSample)
HalftoneHints	321	141	SHORT	2
TileWidth	322	142	SHORT or LONG	1
TileLength	323	143	SHORT or LONG	1
TileOffsets	324	144	LONG	TilesPerImage
TileByteCounts	325	145	SHORT or LONG	TilesPerImage
InkSet	332	14C	SHORT	1
InkNames	333	14D	ASCII	total number of charac ters in all ink name strings, including zeros
NumberOfInks	334	14E	SHORT	1
DotRange	336	150	BYTE or SHORT	2, or 2* NumberOfInks
TargetPrinter	337	151	ASCII	any
ExtraSamples	338	152	BYTE	number of extra compo- nents per pixel
SampleFormat	339	153	SHORT	SamplesPerPixel
SMinSampleValue	340	154	Any	SamplesPerPixel
SMaxSampleValue	341	155	Any	SamplesPerPixel
TransferRange	342	156	SHORT	6
JPEGProc	512	200	SHORT	1
JPEGInterchangeFormat	513	201	LONG	1
JPEGInterchangeFormatLngth	514	202	LONG	1
JPEGRestartInterval	515	203	SHORT	1
JPEGLosslessPredictors	517	205	SHORT	SamplesPerPixel
JPEGPointTransforms	518	206	SHORT	SamplesPerPixel
JPEGQTables	519	207	LONG	SamplesPerPixel
JPEGDCTables	520	208	LONG	SamplesPerPixel
JPEGACTables	521	209	LONG	SamplesPerPixel
YCbCrCoefficients	529	211	RATIONAL	3
YCbCrSubSampling	530	212	SHORT	2
YCbCrPositioning	531	213	SHORT	1
ReferenceBlackWhite	532	214	LONG	2*SamplesPerPixel
Copyright	33432	8298	ASCII	Any

Appendix B: Operating System Considerations

Extensions and Filetypes

The recommended MS-DOS, UNIX, and OS/2 file extension for TIFF files is ".TIF".

On an Apple Macintosh computer, the recommended Filetype is "TIFF". It is a good idea to also name TIFF files with a ".TIF" extension so that they can easily imported if transferred to a different operating system.

Index