# JPEG: New Enhancements and Future Prospects

*Daniel T. Lee*
*Hewlett-Packard Laboratories, Palo Alto, California*

## Abstract

This paper describes the recent work on still image coding standard being developed by ISO/IEC JTC2 SC29/WG1, Coding of Still Picture. New extensions to JPEG in the areas of variable quantization, selective refinement, tiling and an image interchange file format are described. An overview of the new ITU Color FAX standard which is based on JPEG is also presented.

## Introduction

The JPEG (Joint Photographic Experts Group) committee effort has, after seven years of work, culminated in the publication in 1993 the JPEG international standard for compression of continuous-tone color images.[1,2] In parallel with JPEG, the JBIG (Joint Bi-level Image Experts Group) committee has also published its result as in international standard for compression of bi-level images.[3] Both of these standards were developed under the joint auspices of ISO and ITU-T (International Telecommunication Union-Telecommunication Standardization Sector, formerly known as the CCITT) under ISO/IEC JTC1/SC29/WG9 and WG10, respectively. In November 1993, The WG9 and WG10 committees merged to form a single committee, the ISO/IEC JTC1/SC29/ WG1, Coding of Still Pictures.

The JPEG standard, *Digital Compression and Coding ofContinuous-tone Still Images*, is published as three parts, of which the last part (Part 3) is still under development.

- ITU-T Rec. T.81 / ISO/IEC 10918-1: *Requirements and guidelines*,
- ITU-T Rec. T.83 / ISO/IEC 10918-2: *Compliance testing,*
- ITU-T Rec. T.84 / ISO/IEC 10918-3: *Extensions*.

Part 1, what is simply known as JPEG, specifies requirements and implementation guidelines for continuous-tone still image encoding and decoding processes, and for the coded representation of compressed image data. These processes and representations are intended to be generic, that is, to be applicable to a broad range of applications for color and grayscale still images within communications and computer systems.

Part 2 specifies tests for determining whether implementations comply with the requirements for the various encoding and decoding processes specified in Part 1 and was published in 1994.

The current work, Part 3, specifies requirements and guidelines for encoding and decoding extensions to the processes defined in Part 1, and for the coded representation of compressed image data of these extensions. It was promoted to a Committee Draft status in November 1994 and is expected to become a standard by the end of 1995.[4] In this paper, the key features of JPFG Part 3 are presented.

The standardization of facsimile is under the auspices of ITU-T (formerly the CCITT). In particular, the Group 3 facsimile standard, a digital protocol designed to operate over analog data channels, has received much success since its approval in 1981. It has become one of the most popular communication standards that has ever been deployed and is used throughout the world over general switched telephone networks. The Group 4 standard, designed to operate over digital channels, is used in a more limited area. In November 1994, the color extensions to both Group 3 and Group 4, the new Color FAX standard that was based on JPFG, was formally approved by ITU. In this paper, an overview of this stan- dard is presented.

## JPEG Part 3 Extensions

There are four major extensions proposed in JPEG Part 3: variable quantization, selective refinement, tiling and an image interchange file format called Still Picture Interchange File Format (SPIFF). Other minor extensions include addition of a version number to the compressed data format and increasing the limit on the number of data units in a minimum coded unit to 20. Descriptions of the four major extensions are given below.[4]

### Variable Quantization Extension

The variable quantization extension is an enhancement to the quantization procedure of DCT-based processes which provides for changes to the quantization table values within a scan at the $8 \times 8$ block level. This extension may be used in conjunction with any of the DCT-based processes with the exception of the Baseline Process. The quantization procedure as defined in JPEG is the step in the encoding process where each of the 64 DCT coefficients are quantized using one of 64 corresponding values from a quantization table. JPEG permits quantization tables to be redefined prior to the start of a scan but does not allow quantizations table values to be changed within a scan. The variable quantization extension defined by this Specification provides for scaling of quantization values at the $8 \times 8$ block level.

The variable quantization extension introduces a quantizer scale factor which may be coded in the compressed data stream at the start of any $8 \times 8$ block. The quantizer scale factor is used to scale the quantization table values which correspond to the AC coefficients in the quantization procedure. All defined quantization tables.are scaled by the same quantizer scale factor.

This extension provides the following capabilities:

- the ability to compress an image to less than a bounded size with a single sequential pass over the image. The capability is valuable to applications which utilize a fixed-size compressed picture memory;
- the ability to better exploit the masking properties of the human visual system and thereby achieve greater compression rates for the same subjective quality;
- the ability to transcode, i.e. entropy decoding followed by entropy encoding, between coded data representations defined by JPEG Part 3 and those defined by some other standards, such as MPEG.

**Selective Refinement Extension**

The selective refinement extension is used to select a sub-part of an image for further refinement. The different types of selective refinement are described below.

*Hierarchical selective refinement* is used in the hierarchical mode of operation for refining a sub-part of an image. The location of the sub-part of the image to be selectively refined is specified immediately prior to a differential frame within an hierarchical sequence. The size of the sub-part is specified in the differential frame header. The differential frame is then added to only the specified sub-part. One of the primary uses for this type of selective refinement is for coding a particular region of interest with greater detail than the remainder of the image.

*Progressive selective refinement*, the second type of selective refinement, is used in the DCT-based progressive mode of operation. This type of selective refinement is used for similar reasons as hierarchical selective refinement. Progressive selective refinement may be applied to the DCT-based processes which use spectral selection, successive approximation, or both procedures in combination.

*Component selective refinement*, the third type of selective refinement, may be used in all modes of operation for specifying a sub-part frame which contains fewer than the total number of color components defined in the frame header. The most common use for this type of selective refinement is for representing images which are mixed grayscale and color images.

**Tiling Extension**

The tiling extension is used to associate a number of sub-images, also called tiles, in order to form a single tiled image. The tiling extension is also used to repre- sent images which have dimensions larger than 65535 on a side. Three types of tiling are proposed.

Simple tiling can be described as simply breaking up a larger image into smaller rectangular “tiles”. Simple tiling is useful for dividing a large image into pieces where random access into the middle of a compressed image can be handled. The tiles are fixed size except possibly for the tiles on the right and bottom sides of the image when either tiled image dimension is not an integer multiple of the corresponding tile dimension. Tiles are contiguous, non-overlapping and are coded sequentially from left to right and top to bottom. They all must have the same component identifiers, sampling factors, and entropy coders.

*Pyramidal tiling*, the second type of tiling, offers a method of storing multiple resolution versions of an image within the same compressed data stream. Pyramidal tiling is useful for providing access to lower resolution versions of a larger image. For example, pyramidal tiling provides the capability to view large images on a soft display, as required for browsing through “thumbnail” versions of images in a database or for side-by-side image comparison. Pyramidal tiling allows for tiles of one resolution level to overlap those of other resolution levels. The lowest resolution level must be coded first in the data stream and be followed by levels of increasing resolution. Tiles within a single resolution layer must conform to the rules for simple tiling.

Composite tiling, the third type of tiling, is one where there are no restrictions except that all tiles shall have the same component identifiers. Composite tiling is useful for relating diverse sub-images into an image collage, i.e. a single composite image.

**Still Picture Interchange File Format Extension (SPIFF)**

When JPEG was first designed, it did not define an image interchange file format that can be used by other applications. The original intent of JPEG was to allow other applications to encapsulate JPEG data within their own composite file formats. To some extent this has been accomplished in both standardized file formats (e.g., TIFF ODA, CGM) and proprietary or *de facto* standards (e.g., JFIF RTF PostScript II, etc.). The JPEG committee felt that it was possible to design a very simple file format which can specify the vast majority of images created. SPIFF is a file format developed in Part 3 for the interchange of files containing image data between application environments. It is intended to be a generic file format that is simple in nature and does not include many of the features found in more comprehensive file formats.

The primary design requirements of SPIFF was backwards compatibility.[4] The definition of the header and directory of a SPIFF file is such that when a SPIFF file is supplied to most of the currently known commercial and public domain applications that read JPEG compressed image data, they will decode the compressed data stream properly without using any other information in the SPIFF file. SPIFF files may contain compressed image data for bi-level or continuous-tone (grayscale or colour) images. Several different standard compression algorithms are supported: MH, MR, MMR, JBIG, and JPEG. In addition to the compressed image data, SPIFF includes all information necessary to completely decode the data and render it on a given output device, within the constraints imposed by that device.

Other requirements considered in the design of SPIFF are that parameter size shall not be constrained by any anticipated limits, e.g., 64K bytes, there shall be minimal buffering requirements and the file shall be compatible with applications which use serial access, e.g., point-to-point transmission, networks, or Unix pipes. SPIFF also defined optional fields which allow creators of images to include an image title, image description, creator identification, copyright message, contact information, time stamp, and other useful information.

**Overview of SPIFF Specifications**

The *file header* is the first data that appears in the file and serves to identify the file's contents as SPIFF data. The header also contains vital information about the image such as the application profile, number of components, and image dimensions.

The *directory* is a sequence of directory entries. The directory contains, or contains references to, information necessary to accurately render decoded image data.

Directory entries may contain *direct data*, or may refer to *indirect data*. Direct data is typically used if the amount of data is small and fits within the directory entry (less than 65528 bytes). If the data for a particular directory entry is too large to fit as direct data, the entry will contain a reference to the indirect data.

Every SPIFF conformant file shall contain at least one compressed image data stream. This compressed image data stream contains all data necessary to decode the components of the image contained in this file. This data, in combination with some of the information contained in the directories, is what is necessary to accurately render the image on any given output device.

Ancillary data in the file may contain one or more "thumbnail" image representations, each of which may optionally be represented by a compressed image data stream and, consequently, more than one of these data streams may be present in any given SPIFF conformant file.

### High-level Syntax

Figure 1 specifies the order of high level constituent parts of the interchange file format for all files conformant to SPIFF.

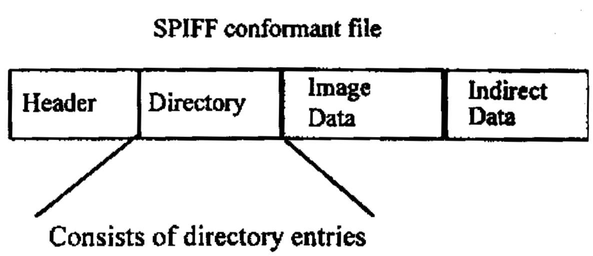


*Figure 1. High-level Syntax for SPIFF*

The block labeled "directory" is optional. The block Labeled "indirect data" is also optional and, if present, consists of one or more individual indirect data items corresponding to directory entries in the directory.

In summary, SPIFF was developed both to maintain backwards compatibility and to have similar structure as TIFF so that SPIFF files may be easily converted to TIFF and vice versa.

## ITU Color FAX Standard

In November 1994, the continuous-tone color image extensions to Group 3 and Group 4 facsimile were formally approved by ITU and a new international Color FAX standard has hence been established.[5,6] In the new Color FAX standard, the CIE (1976) L* a* b"—CIELAB color space is standardized as the interchange color space. The standard implements JPEG in a fully compliant manner, encouraging interchange between facsimile and non-facsimile applications. In a Color FAX data stream, the JPEG-encoded image data consist of a series of markers, parameters, and scan data that specify the image coding parameters, image size, bit-resolution, and entropy-encoded block-interleaved data. The data stream is encoded for facsimile transfer using the error correction mode (ECM) specified in the Group 3 protocol, the ITU-T Rec. T.30 Annex A.[7]

### Color Representation

The choice of color representation for Color FAX was a subject of intense study and discussion at ITU during 1992 and 1993. In the end, after an exhaustive study of different color representations, the CIE (1976) L*a*b* space (CIELAB) was chosen as a flexible, relatively uniform, and device independent color specification.

Since CIELAB is a relative color metric, the choice of illuminant, white point, and measurement conditions is necessary to define the representation precisely. The CIE D50 illuminant was chosen in agreement with com- mon practice in the graphic arts industry, along with a perfectly diffuse, 100% reflecting white point. A mea- surement geometry of 45-0 illuminant to measurement angle is also specified. The choice of gamut range is as follows:

$$L^* = [0, 100]$$
$$a^* = [-85, 85]$$
$$b^* = [-75, 125].$$

It was chosen to serve several goals. The default gamut range is sufficiently wide to span existing hard copy output devices. The range is narrow enough to avoid excessive quantization error when the data is represented in 8 bits/component. The particular choices of gamut range are believed to represent existing hard copy devices, as well as facilitating effective implementation. The conversions from real values in CIELAB to the 8-bit integer representations are performed as show:

$$L = (L^*) * (255/100)$$
$$a = (a^*) * (255/170) + 128$$
$$b = (b^*) * (255/200) + 96,$$

where L, a, and b represent 8 bit integers, and L*, a*, and b* represent real numbers.

Following successful negotiation, any alternative gamut range may be specified by the transmitter. This is intended to allow for soft-copy device gamuts or for more accurate specification of colors within a narrower gamut range. In addition, 12 bits/channel of data may be transmitted as an option.

### Spatial Representation

The spatial resolution chosen as basic for color facsimile is $200 \times 200$ pels/25.4mm. This spatial resolution is familiar in most fax machines as the "fine mode." The chrominance channels a* and b* are subsampled to $100 \times 100$ pels/25.4mm using a symmetric 4-tap filter. This reduces the number of DCT calculations needed for image coding, and takes advantage of the lower visual sensitivity to chroma modulation. In addition, optional spatial resolutions of $300 \times 300$ and $400 \times 400$ pels/25.4 mm are available upon successful negotiation, as is a non-subsampled chroma mode.[6]

### Optional Features

The basic values of the Color FAX standard are designed to allow the efficient transmission of hard-copy continuous-tone image data. The optional features included in the standard are chosen to provide additional flexibility for applications. These features include higher spatial resolution, no spatial sub-sampling, 12 bits/plane bit precision, and custom gamut range. These options are only available following successful negotiation between terminals. In addition, the JPEG restart marker is supported. This marker permits re-synchronization of the entropy encoded data in the event of data loss. This option may be used without prior negotiation. The optional use of a custom illuminant is under study.

## Future Prospects of JPEG

Image compression is an absolute necessity for applications which rely on the communication of images across networks or other transmission media, e.g., facsimile, remote image databases, teleconferencing. It is also essential for reduction of memory usage in storage and archival applications. The benefits of compression and the availability oflow cost compression implementations either in software or hardware have resulted in the proliferation of compressed images to the extent that today most images are compressed immediately after their creation. The new Color FAX standard based on JPEG will enable the efficient communication of continuous-tone color images over telephone lines and other networks where it was not easily done before. The work being advanced by the JPEG committee in JPEG Part 3 will enhance the use of JPEG to even larger variety of digital imaging applications on the information superhighway. The immediate benefits of the variable quantization would be seen in applications like transcoding from MPEG (moving image) data streams, constant bandwidth (rate-control) compression and fixed file size compression. The new tiling structure would provide applications such as hypertext with selected parts of an image in which an application can process independently. It can be applied effectively to address system specific limitations such as the availability of display memory and networked systems. Tiling can also allow a number of images, possibly from dissimilar sources, to be linked together in a single composite image. Lastly, SPIFF provides a file form that can be reliably used for the interchange of com- pressed images between application environments which was something missing before. In summary, JPEG has established itself as the key choice of compression for digital image and the additional enhancements will provide more capabilities to meet the ever needs of digital imaging applications on the information superhighway.

## References


1. ITU-T Rec. T.81 / ISO/IEC 10918-1, Information Technology — Digital compression and coding of continuous tone still images, Part 1: Requirements and Guidelines, 1993.
2. William B. Pennebaker and Joan L. Mitchell, *JPEG Still Image Data Compression Standard*, Van Nostrand Reinhold, 1993.
3. ITU-T Rec. T.82 / ISO/IEC 11544, Information Technology—Coded Representation of Pictures and Audio Information—Progressive Bi-level Image Compression, 1993.
4. ITU-T Draft Rec. T.84 / ISO/IEC CD 10918-3, Information Technology—Digital compression and coding of continuous-tone still images, Part 3: Extensions. ISO/IEC JTC1/SC29 Document N973 rev, November 13, 1994.
5. M. Matsuki and D. T. Lee, "Development of Color Facsimile Standards: An Overview," *SID Digest of Tech. Papers,* Vol. **25**, 1994, pp. 883-886.
6. A. Mutz and D. T. Lee, "An International Standard for Color Facsimile," *Proc. 2nd IS&T/SID Color Imaging Conference*, Nov, 15-18, 1994, pp. 52-54.
7. ITU-T Rec. T. 30, Procedures for Document Facsimile Transmission in the General Switched Telephone Network, 1994.



☆ This paper was previously published in *IS&T's 1995 Annual Conference Proc.,* p. 58 (1995).