



The information in this tutorial is specific to CONTENTdm version 4.3 and may have some outdated information. Please check back at this link location soon for the new version 5 tutorial.

## Understanding Image File Types

This document describes a variety of file types for images and discusses the implications for use. In the first two sections, the two standard formats used for Web images are discussed. In Section III, the archival standard is discussed. In Section IV, three other formats are discussed.

### Learn About

**Section I: JPEG**

**Section II: GIF**

**Normal**

**Interlaced**

**Animated GIF**

**Section III: TIFF**

**Uncompressed**

**LZW Compressed**

**Group IV Compression**

**JPEG Compressed**

**Section IV: Other File Types**

**MrSID®**

**PNG**

**JPEG2000**

### Section 1: JPEG

JPEG (Joint Pictures Expert Group) is designed for compressing either full-color or gray-scale images of natural, real-world scenes. It works well on photographs, naturalistic artwork, and similar material. It may not be the best format to use for lettering, simple cartoons, or line drawings. Web browsers support this format natively. Along with GIF, JPEG is the standard format for Web images.

JPEG stores full color information: it stores 24 bits/pixel, which means it can store up to 16 million colors. JPEG images display very well on monitors that support more than 256 colors.

The JPEG algorithm rearranges the image information into color and detail information, compressing color more than detail because our eyes are more sensitive to detail than to color, making the compression less visible to the naked eye. It sorts the detail information into fine and coarse detail and discards the fine detail first because our eyes are more sensitive to coarse detail than to fine detail.

JPEG is a standard lossy image compression algorithm. Lossy compression means that only a part of the original information is still there when the file is uncompressed, although the user may not notice any change.

The degree of lossiness can be varied by adjusting compression parameters. This means you can trade off file size against output image quality. Below are two sample JPEG images created with Adobe Photoshop®. Figure 1 is a high-quality image with a size of 35kb, while Figure 2 is a low-quality image with a size of 4.7kb.



Figure 1. High Quality JPEG with File Size of 35kb



Figure 2. Lower Quality JPEG with File Size of 4.7kb

Notice the boxing effect in the low quality image in Figure 2. This is a characteristic of a JPEG compression algorithm for high compression ratios. The low-quality representation is still very faithful to the original photograph despite being compressed by an additional factor of four.

## Section II: GIF

GIF (Graphic Interchange Format) works best for images with only a few distinct colors, such as line drawings and simple cartoons. GIF is useful for cartoon images that have less than  $256 \cdot (2^8)$  colors, grayscale images, and black and white text.

GIF, like JPEG, is a standard format for Web images. The primary limitation of a GIF is that it only works on images with 8 bits per pixel or less, which means 256 or fewer colors. Most color images are 24 bits per pixel. To store these in GIF format you must first convert the image from 24 bits to 8 bits. The conversion will result in a loss of data and a considerable degradation in quality. Computer monitors that display only 256 colors or less display GIFs well.

GIF is a lossless image file format. With lossless compression, all of the data that was originally in the file remains after the file is uncompressed. GIF compresses images using LZW compression. LZW compression is named for the individuals who invented it, Lempel-Zev, Welch. It was originally invented to compress text for transmission over telephone lines. This form of compression is a lossless compression form with ratios varying depending on the color complexity of an image. The compression rate decreases as the amount of information (number of bits) increases.

There are three primary types of GIF images.

### Normal

The GIF image data is stored sequentially. Notice in Figure 3 that the full color palette is not used. The image was created in PhotoShop using the default palette of the 216 colors used by all Web browsers. You can achieve much better results using GIF with a custom or adaptive palette similar to the one used in Figure 4.



Figure 3. GIF With A Web Safe Palette



Figure 4. GIF with an Adaptive Palette

### Interlaced

With interlaced GIFs, the lines of the image are not stored sequentially, but are interlaced. For example, instead of storing lines 1 through 10 in order, it stores line 1 and then lines 3,5,7,9,2,4,6,8,10. This allows applications to display part of the image first and then fill in the missing lines to complete the image.

### Animated GIF

This allows you to store multiple GIF images in the same image file. Usually the images are displayed sequentially over time, creating a small animation.

## Section III: TIFF

TIFF (Tagged Image File Format) has emerged as the standard archiving image file format for library use. Its strengths are that the format is extensible -- new image types can be introduced without invalidating older types—and portable -- it is independent of hardware and operating system types.

There are many types of TIFF files; the most common ones are described below.

### Uncompressed

Uncompressed TIFF images are stored in an uncompressed raw format. This is the basic format for archival images because the data is left unchanged.

### **LZW Compressed**

This lossless format works the same as GIF; however, LZW compression in TIFF supports more than 8 bits per pixel. This is a good choice for archival images because it offers compression without any loss of data. Each bit is identical to the uncompressed version.

### **Group IV Compression**

Group IV Compression is used to compress black and white images. It is commonly used to compress documents for applications such as fax machines.

### **JPEG Compressed**

This type uses JPEG compression and stores it in a TIFF file.

## **Section IV: Other File Types**

Other electronic file types are used within Web pages and in archiving environments.

### **MrSID®**

MrSID (Multi-resolution Seamless Image Database) is used in photography, mapping/GIS, document management, medical imaging, and games. It is an image type that uses a lossy, wavelet-based compression algorithm. It supports large files, tiling, and scalability in quality and resolution. It works by putting together hundreds of small image tiles into one large seamless image that can be compressed and decompressed with little or no degradation. For more information, read the description on the Lizard Tech Web site, [www.lizardtech.com](http://www.lizardtech.com).

### **PNG**

PNG (Portable Network Graphics) is a file format for image compression. It was developed as a patent free replacement for GIF (Unisys owns the GIF format). It provides a number of improvements over the GIF format.

Like a GIF, a PNG file uses lossless compression. It allows you to make a trade-off between file size and image quality when the image is compressed. Typically, an image in a PNG file can be 10 to 30% more compressed than in a GIF format. Like GIFs, you can make one color transparent, but you can control the degree of transparency (this is also called "opacity"). Interlacing is supported and is faster in developing than in the GIF format. Images can be saved using true color as well as in the palette and gray-scale formats.

### **JPEG2000**

JPEG2000 format is emerging as a standard for image compression. It provides much better image quality at smaller file sizes than JPEG does. Based on wavelet compression, JPEG2000 offers both lossless and lossy compression. JPEG2000 formats provide good image quality, even at very high compression ratios such as 80:1. JPEG2000 creates scalable image files, which means that no decompression is needed for reformatting.

Other new functionalities include region of interest coding, improved error resilience, resolution scalability, random access or spatial scalability, and quality scalability. For more information, visit [www.jpeg2000info.com](http://www.jpeg2000info.com).

