

Digital Image File Types Explained

JPG, GIF, TIFF, PNG, BMP. What are they, and how do you choose? These and many other file types are used to encode digital images. The choices are simpler than you might think.

Part of the reason for the plethora of file types is the need for *compression*. Image files can be quite large, and larger file types mean more disk usage and slower downloads. Compression is a term used to describe ways of cutting the size of the file. Compression schemes can be *lossy* or *lossless*.

Another reason for the many file types is that images differ in the number of colors they contain. If an image has few colors, a file type can be designed to exploit this as a way of reducing file size.

Lossy vs. Lossless compression

You will often hear the terms "lossy" and "lossless" compression. A lossless compression algorithm discards no information. It looks for more efficient ways to represent an image, while making no compromises in accuracy. In contrast, lossy algorithms accept some degradation in the image in order to achieve smaller file size.

A lossless algorithm might, for example, look for a recurring pattern in the file, and replace each occurrence with a short abbreviation, thereby cutting the file size. In contrast, a lossy algorithm might store color information at a lower resolution than the image itself, since the eye is not so sensitive to changes in color of a small distance.

Number of colors

Images start with differing numbers of colors in them. The simplest images may contain only two colors, such as black and white, and will need only 1 bit to represent each pixel. Many early PC video cards would support only 16 fixed colors. Later cards would display 256 simultaneously, any of which could be chosen from a pool of 2^{24} , or 16 million colors. New cards devote 24 bits to each pixel, and are therefore capable of displaying 2^{24} , or 16 million colors without restriction. A few display even more. Since the eye has trouble distinguishing between similar colors, 24 bit or 16 million colors is often called TrueColor.

The file types

TIFF is, in principle, a very flexible format that can be lossless or lossy. The details of the image storage algorithm are included as part of the file. In practice, TIFF is used almost exclusively as a lossless image storage format that uses no compression at all. Most graphics programs that use TIFF do not compression. Consequently, file sizes are quite big. (Sometimes a lossless compression algorithm called LZW is used, but it is not universally supported.)

PNG is also a lossless storage format. However, in contrast with common TIFF usage, it looks for patterns in the image that it can use to compress file size. The compression is exactly reversible, so the image is recovered exactly.

GIF creates a table of up to 256 colors from a pool of 16 million. If the image has fewer than 256 colors, GIF can render the image exactly. When the image contains many colors, software that creates the GIF uses any of several algorithms to approximate the colors in the image with the limited palette of 256 colors available. Better algorithms search the image to find an optimum set of 256 colors. Sometimes GIF uses the nearest color to represent each pixel, and sometimes it uses "error diffusion" to adjust the color of nearby pixels to correct for the error in each pixel.

GIF achieves compression in two ways. First, it reduces the number of colors of color-rich images, thereby reducing the number of bits needed per pixel, as just described. Second, it replaces commonly occurring patterns (especially large areas of uniform color) with a short abbreviation: instead of storing "white, white, white, white, white," it stores "5 white."

Thus, GIF is "lossless" only for images with 256 colors or less. For a rich, true color image, GIF may "lose" 99.998% of the colors.

JPG is optimized for photographs and similar continuous tone images that contain many, many colors. It can achieve astounding compression ratios even while maintaining very high image quality. GIF compression is unkind to such images. JPG works by analyzing images and discarding kinds of information that the eye is least likely to notice. It stores information as 24 bit color. Important: the degree of compression of JPG is adjustable. At moderate compression levels of photographic images, it is very difficult for the eye to discern any difference from the original, even at extreme magnification. Compression factors of more than 20 are often quite acceptable. Better graphics programs, such as Paint Shop Pro and Photoshop, allow you to view the image quality and file size as a function of compression level, so that you can conveniently choose the balance between quality and file size.

RAW is an image output option available on some digital cameras. Though lossless, it is a factor of three or four smaller than TIFF files of the same image. The disadvantage is that there is a different RAW format for each manufacturer, and so you may have to use the manufacturer's software to view the images. (Some graphics applications can read some manufacturer's RAW formats.)

BMP is an uncompressed proprietary format invented by Microsoft. There is really no reason to ever use this format.

PSD, PSP, etc. , are proprietary formats used by graphics programs. Photoshop's files have the PSD extension, while Paint Shop Pro files use PSP. These are the preferred working formats as you edit images in the software, because only the proprietary formats retain all the editing power of the programs. These packages use layers, for example, to build complex images, and layer information may be lost in the nonproprietary formats such as TIFF and JPG. However, be sure to save your end result as a standard TIFF or JPG, or you may not be able to view it in a few years when your software has changed.

Currently, GIF and JPG are the formats used for nearly all web images. PNG is supported by most of the latest generation browsers. TIFF is not widely supported by web browsers, and should be avoided for web use. PNG does everything GIF does, and better, so expect to see

PNG replace GIF in the future. PNG will *not* replace JPG, since JPG is capable of much greater compression of photographic images, even when set for quite minimal loss of quality.

File size comparisons

Below are comparisons of the same image saved in several popular file types. (Note that there is no reason to view more than one of the TIFFs or the PNG. Since all are lossless formats, their appearance is identical.)

File type	Size
Tiff, uncompressed	901K
Tiff, LZW lossless compression (yes, its actually bigger)	928K
JPG, High quality	319K
JPG, medium quality	188K
JPG, my usual web quality	105K
JPG, low quality / high compression	50K
JPG, absurdly high compression	18K
PNG, lossless compression	741K
GIF, lossless compression, but only 256 colors	286K

When should you use each?

TIFF

This is usually the best quality output from a digital camera. Digital cameras often offer around three JPG quality settings plus TIFF. Since JPG always means at least some loss of quality, TIFF means better quality. However, the file size is huge compared to even the best JPG setting, and the advantages may not be noticeable.

A more important use of TIFF is as the working storage format as you edit and manipulate digital images. You do not want to go through several load, edit, save cycles with JPG storage, as the degradation accumulates with each new save. One or two JPG saves at high quality may not be noticeable, but the tenth certainly will be. TIFF is lossless, so there is no degradation associated with saving a TIFF file.

Do NOT use TIFF for web images. They produce big files, and more importantly, most web browsers will not display TIFFs.

JPG

This is the format of choice for nearly all photographs on the web. You can achieve excellent quality even at rather high compression settings. I also use JPG as the ultimate format for all my digital photographs. If I edit a photo, I will use my software's proprietary format until finished, and then save the result as a JPG.

Digital cameras save in a JPG format by default. Switching to TIFF or RAW improves quality in principle, but the difference is difficult to see. Shooting in TIFF has two disadvantages compared to JPG: fewer photos per memory card, and a longer wait between photographs as the image transfers to the card. I rarely shoot in TIFF mode.

Never use JPG for line art. On images such as these with areas of uniform color with sharp edges, JPG does a poor job. These are tasks for which GIF and PNG are well suited. See JPG vs. GIF for web images.

GIF

If your image has fewer than 256 colors and contains large areas of uniform color, GIF is your choice. The files will be small yet perfect. Here is an example of an image well-suited for GIF:



Do NOT use GIF for photographic images, since it can contain only 256 colors per image.

PNG

PNG is of principal value in two applications:

1. If you have an image with large areas of exactly uniform color, but contains more than 256 colors, PNG is your choice. Its strategy is similar to that of GIF, but it supports 16 million colors, not just 256.
2. If you want to display a photograph *exactly* without loss on the web, PNG is your choice. Later generation web browsers support PNG, and PNG is the only lossless format that web browsers support.

PNG is superior to GIF. It produces smaller files and allows more colors. PNG also supports *partial transparency*. Partial transparency can be used for many useful purposes, such as fades and antialiasing of text. Unfortunately, Microsoft's Internet Explorer does not properly support PNG transparency, so for now web authors must avoid using transparency in PNG images.

Other formats

When using graphics software such as Photoshop or Paint Shop Pro, working files should be in the proprietary format of the software. Save final results in TIFF, PNG, or JPG.

Use RAW only for in-camera storage, and copy or convert to TIFF, PNG, or JPG as soon as you transfer to your PC. You do not want your image archives to be in a proprietary format.

Although several graphics programs can now read the RAW format for many digital cameras, it

is unwise to rely on any proprietary format for long term storage. Will you be able to read a RAW file in five years? In twenty? JPG is the format most likely to be readable in 50 years. Thus, it is appropriate to use RAW to store images in the camera and perhaps for temporary lossless storage on your PC, but be sure to create a TIFF, or better still a PNG or JPG, for archival storage.