In this chapter from *Digital Photo Magic: Easy Image Retouching and Restoration for Librarians, Archivists, and Teachers*, Ernest Perez provides an introduction to the technical aspects of digital photo files, formats, and related technology. By giving readers a basic understanding of their choices for using post-processing software, he aims to help them productively use photo editing tools and make informed decisions.

Digital Image Details

Ernest Perez

There's no denying it—you'll encounter some heavy-duty tech-speak when learning about digital photo files, formats, and related technology. You can spend endless hours studying the technical aspects of the subject, if you like, or you can read the short introduction provided in this chapter. I wrote it to prepare the average reader to begin productively using DPM photo editing tools.

Effective digital image post-processing requires that you make informed decisions in regard to the multiple steps involved in editing and saving your output. This chapter will provide a basic understanding of your choices, sufficient at least for your beginning efforts in using post-processing software.

Graphic Image Formats

Graphic image file formats used for storing digital images are defined by precise technical specifications. Industry-wide cooperative and ISO (International Organization for Standardization) definitions delineate formats such as TIFF or JPG.. Numerous proprietary definitions are also issued by vendors, such as for the GIF (CompuServe) and BMP (Microsoft) formats. The proprietary formats may well become established and eventually become widely adopted by software and hardware vendors. For example, Compuserve's patent on the GIF file format expired in the early 2000s, but it is now commonly used. The PNG (Portable Network Graphics) file format

was created by the ad hoc PNG Working Group in 1995-1996 to improve upon and, in some respects, replace GIF.

Precise format definitions are essential for effective standard usage. These enable hardware and software developers to effectively tailor their products for wide cross-platform compatibility. Brand A Camera must be able to accurately create and transfer Image B, which must be able to be saved by Computer C to Hard Disk D. That file must also be flawlessly loadable and editable by Photo Editor E, and it must be able to be printed by Printer F. No problem—you hope.

Each community standard file definition offers numerous content and coding options to cover the allowable variations in image representation within a single format. These options allow needed flexibility in photo image representation and descriptions within the single format, including innumerable image description variations or alternatives, such as choice of color or black-and-white images, gamma values, color hues, image size, image resolution, brightness, contrast, and screen display methods.

Understandably, then, because you'll need to be transporting the photo images you work with through the maze of photo-editing tools and your own unique mix of hardware, you absolutely, positively need basic knowledge about digital image file formats.

Common Photo Image Formats

The most widely used digital photo image formats are TIFF (Tagged Image File Format), JPEG (Joint Photographic Experts Group), and PNG (Portable Network Graphics), but numerous other specialized formats are also in use, including the following:

- RAW: The original file format created and saved by digital camera hardware
- GIF: A compact and less detailed image format, used mostly for simple line art or graphic illustrations (described near the end of this subsection)

Proprietary formats: Working file formats defined by software vendors for their own applications—for example, Adobe's PSB and PSD, ULEAD PhotoImpact's UFO and UPI, and GIMP's XCF

TIFF files require the most disk space, by a significant margin. The reason TIFF files are so large is that they represent a "lossless

compression" format. TIFF files use file compression algorithms that lose no detail while saving successive versions of files. This means that when you decompress a TIFF file for display or printing, you get exactly the same image as the original one. That holds true after all subsequent editing and post-processing manipulation. TIFF files use "raster" image depiction, essentially describing the color values of every pixel or dot needed to represent an image. As this results in both higher-quality image depiction and larger file sizes, there is both an advantage and a disadvantage. TIFF files are also not editable by as many file-editing tools as some other formats: Many simpler photo-editing applications are restricted to operating on the smaller and more popular JPEG, PNG, and GIF formats.

JPEG uses a "lossy compression" algorithm, applying a much higher compression ratio for image storage. It typically achieves a 10:1 compression ratio, with little perceptible loss in image quality. But using the lossy compression method means that a small amount of the image detail is lost during each successive file save. The human eye can't easily distinguish the small amount of image quality loss, but using JPEG for storage means that you will suffer cumulative quality loss after multiple generations of file saves. JPEG has generally been the most widely used photo image format, although it is now being overtaken by the PNG format for online uses.

The JPEG compression ratio is user-selectable. A JPEG file can be written to disk with a very small amount of lost detail, but this file will be considerably larger than a higher-compression JPEG file.

The newer PNG format—like TIFF, a raster format—uses lossless compression. Its file size, smaller than that of a comparable TIFF file, means faster transfer of images online and faster display time on computer screens. PNG files are thus optimized for transfer speed and online display; they can't quite achieve the print detail of TIFF files. However, the difference is not easily detectable with the naked eye.

Image Types

An image's type refers to one of two principal methods used to actually build or represent an image in a graphic file:

- *Raster images* use a bitmap representation of an image, a dot matrix data structure representing a generally rectangular grid of pixels, or points, of color. Raster images cannot be rescaled without losing quality.

Vector images are built using "graphical primitives" such as squares, circles, and curves, created during processing using geometric formulas to describe the shapes. These files are smaller and are in fact preferable for representing graphic line art or simpler images.

Raster or bitmapped images are better than vector imaging for complex continuous-tone images, such as photographs. However, simpler black-and-white or "line art" images may be efficiently saved as lossless raster images in the GIF file format. These GIF images result in extremely small file sizes, far below even those saved in the (lossy) JPEG format. The highest-quality and most detailed TIFF and BMP format files use raster image processing and are generally, accordingly, much larger in size than vector images.

Until recently, image file size significantly affected choice of format, which often involved a delicate balance between disk storage size (or transmission speed) and image or visual quality. This was a more important consideration "back in the old days," when storage space was much more limited and expensive.

Relative file size of the file formats clearly illustrate the tradeoff of space versus quality. For example, I recently worked with a TIFF image that was 8.5 MB (8,704 KB) in size, yet the same image in a moderate-quality JPEG format was only 482 KB in size. By way of another example, in "A Few Scanning Tips" (Fulton, 2013, "Image File Formats") Wayne Fulton compares the sizes of identical 4×6 images from a 12-megapixel digital camera, saved as RAW, TIFF, JPG, and PNG files. Table 4.1 summarizes his data, depicting the extreme range of file sizes produced.

Fulton concludes, "Large images consume large memory and make our computers struggle. . . . When we double the scan resolution, memory cost goes up 4 times. Multiply resolution by 3 and the memory cost increases 9 times, and more. So this seems a very clear argument to use only the amount of resolution we actually need to improve the image results for the job purpose. More than that is waste" (Fulton, 2013b).

As things now stand, more powerful computers, less expensive computer memory, and enormous decreases in the cost of file storage space actually make this much less of a cost/space-driven decision than ever before. Image quality considerations are much more important now, along with the consideration of image file size effects on online display speeds. However, remember that network bandwidth

speeds are also rapidly increasing—this also is a less important consideration than before.

Table 4.1 Compare the sizes of a 4×6 image saved used different formats. (Data taken from Fulton 2013a)

File Type	Size/Proportion
RAW camera data, 12 megapixel	36 megabyte
TIFF LZW (65–80%)	23.4–28.8 megabyte
PNG (50–65%)	18–23.4 megabyte
JPG (5–20%)	1.8–7.2 megabyte

Table 4.2 Rick Matthews, Wake Forest University, has presented similar comparisons of identical images stored in different formats, as well as at different levels of compression (users.wfu.edu/matthews/misc/graphics/formats/formats.html).

File type	Size
Tiff, uncompressed	901K
Tiff, LZW lossless compression (yes, it's actually bigger)	928K
JPG, High quality	319K
JPG, medium quality	188K
JPG, 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

Preserving Image Quality

Preserving image quality depends greatly on your photo-editing working approach if you do elect to use lossy JPEG files. Accordingly, it's important to think about and plan your editing work flow more carefully when you're using JPEG images than when using other formats. Using a lossy compression file degradation degrades image quality whenever you save a modified file. Unfortunately, the losses are cumulative over repeated generations of file-saving, much as

when you make repeated generations of photocopies or analog audiotapes. It's like compounded interest: A 3% or 5% quality loss is not much, but if you imprudently save multiple versions or generations of the image file, you can reach 20% or 30% degradation, which means you'll pay for multiple lossy file saves in visible loss of final quality.

If you try for higher resolution or photo image detail and don't research methods of doing so, it's almost always going to cost you in much larger file sizes, and you'll unfortunately gain little advantage in resolution quality. Many knowledgeable photographer specialists have investigated the visual improvements of using high-resolution scanning, and they report that most people cannot accurately identify improvements and differences resulting from high-resolution scanning of prints.

Table 4.3 This simplified comparison depicts the positives and negatives of each of the image file formats we've been reviewing.

	TIFF	JPEG	PNG
What you use it for	**Ideal for editing and making large prints** because it contains the largest amount of image information.	**For most printing jobs and sharing through email and over the Internet**, JPEG is the perfect file format.	**Excellent for website display and Internet transfer** due to smaller size and good quality.
Benefits	**TIFF supports layered image files**, good for use with software programs like Photoshop. **TIFF files can be saved with very little compression** making it ideal for printing large sized high-resolution images.	**JPEG has the highest compression** of the older formats and therefore offers the smallest file size. Also the most common file format in use. **Supported by largest number of photo-editing applications.**	**PNG uses lossless compression**, along with high compression and image quality. **Lossless compression and smaller file size** also suit it for archival master use. **Extremely rapid adoption for web and Internet usage**, with increasing compatibility among web browser applications.
Drawbacks	**TIFF files are large.** Depending upon image resolution, you can easily find yourself working with files in the 5–15MB range. **TIFFs are not widely supported by Web browsers**, making them a poor choice for online use.	**A JPEG file degrades each time it is saved**, due to "lossy" compression. Best to save a high-quality copy of the original, then edit copies of that file.	**PNG is not supported as widely by photo-editing applications as the other formats**, although this is changing for the better.

Common-Sense Scanning Decisions

Common-sense scanning decisions should always include a judgment of just how much quality you *really* need, as well as the practical arts and skills of scanning methodology. I recommend the following practical guides to photo print scanning.

- Ed Hamrick's "Batch Scanning Tips," to which a link appears on the DPM companion website (www.update4dpm.com), is a must-read on its topic.
- Luisa Simone presented a practical and still useful guide to scanning in her May 7, 2002, *PC Magazine* article "Solutions: Tools and Tips for the Internet Age—Boost Your Scanning Skills" (www.pcmag.com/article2/0,2817,1163604,00.asp). She covers the practical aspects of pixel resolution setting, brightness and tone control, and the mysterious gamma setting.
- Eric Goodnight's article "How To Properly Scan a Photograph (and Get An Even Better Image)" (www.howtogeek.com/109409/how-to-properly-scan-a-photograph-and-get-an-even-better-image) is an easily understandable guide to scanning prints. Goodnight clearly explains histogram preview charts, as well as scanner software controls and settings you can use to get the best-quality scans possible.

Steve Hoffman's Photography website (www.sphoto.com) provides an immense amount of information covering both digital photography and photo print scanning. Offering more than just Hoffman's gorgeous nature photography, this is a valuable source for anyone who needs to learn about the digital photography and photo print scanning topics. His "Tech Articles and Tutorials" page (www.sphoto.com/techinfo) makes his technical tutorials directly available.

Minimizing Image Quality Losses

Minimizing image quality losses by adhering to good working practices will pay off in both the short and the long run. I highly recommend that you choose the PNG digital format for standard use in your routine post-processing work.

There are several good reasons for this: The PNG format is widely used in post-processing tools. It uses lossless compression, which does not degrade image quality after repeated editing sessions. PNG files are significantly smaller than comparable TIFF files, requiring less storage space and speeding file transmission times. And file conversion utilities are widely available for speedy conversion of PNG files to all other commonly used digital image formats.

When using PNG files, remember the following in your post-processing work:

1. Begin with good-quality images. If you need to start from hardcopy original photo images, perform your original digital scans to PNG format files at an appropriate DPI resolution setting. Resolution of 300 dpi is widely recommended as an excellent default setting for scanning photo print images. Initial scanning to a PNG digital file starts you off with a high-quality image. PNG's lossless compression prevents image degradation if you use multiple editing sessions during your post-processing work.

2. You'll have no difficulty finding quality photo-editing tools compatible with PNG format. When you've completed post-processing work, you can easily convert PNG files to essentially all commonly used digital file formats as needed for your final output.

   More detail on techniques for productive use of scanning software hardware and software is provided in Chapter 7.

3. *If—and only if—they use lossless compression formats*, use the proprietary file format of your photo-editing applications during the editing phase. Some of the better applications that can load and save standard image formats nevertheless use their own lossless compression formats for intermediate working files during the editing process. These applications apparently do this to avoid the problem of accidentally introducing image degradation from using lossy file compression during multiple editing sessions. Some examples of such "safe" proprietary formats are PhotoShop's PSB and PSD files, PhotoImpact's UFO and UPI files, and GIMP's XCF files.

   At the end of your post-processing editing work, you can then save the proprietary intermediate file format to a

high-quality PNG file or a file in whatever other format is required for delivery of your output.

4. Use what you need to use. Be practical: If your standard photo-editing toolkit can't handle some unusual digital photo format you happen to receive, go ahead and use whatever you can find that will get the job done.

Standardize on PNG

Standardize on PNG for your digital image post-processing work. There are simply too many advantages to doing so, and no valid reason to avoid it.

Digital Image Expert Opinion

Digital image expert opinion from Sue Chastain bears out the advantages of using PNG as your preferred format. Chastain is the guide (topic editor) for About.com's widely consulted Graphics Software page (graphicssoft.about.com). She offers a good many practical facts and suggestions for preserving image quality during your post-processing work:

- If you have to use JPEG files, they do not lose quality *every time* they are saved. "Repeated saving *within the same editing session* won't introduce additional damage. It is only when the image is closed, re-opened, edited and saved again" (Chastain, 2013a). In other words, you can minimize quality loss by avoiding the use of your photo-editor's "Save As" function for the purpose of copying, duplicating, or renaming JPEG files. It's preferable to use your file manager or directory utility to perform these functions.
- Get your image files into a lossless format as early as you can during your working process. This minimizes quality loss during editing processes.
- Use a file conversion utility to transform an original JPEG file directly to PNG. A conversion utility generally will not add extraneous space during a conversion to a new PNG file. *Above all, don't use the "Save As" operation in your photo editor for file conversion.* "So with a JPEG that has been converted to PNG you will often get a larger file in kbytes than if you would have gone to PNG to begin with. In other

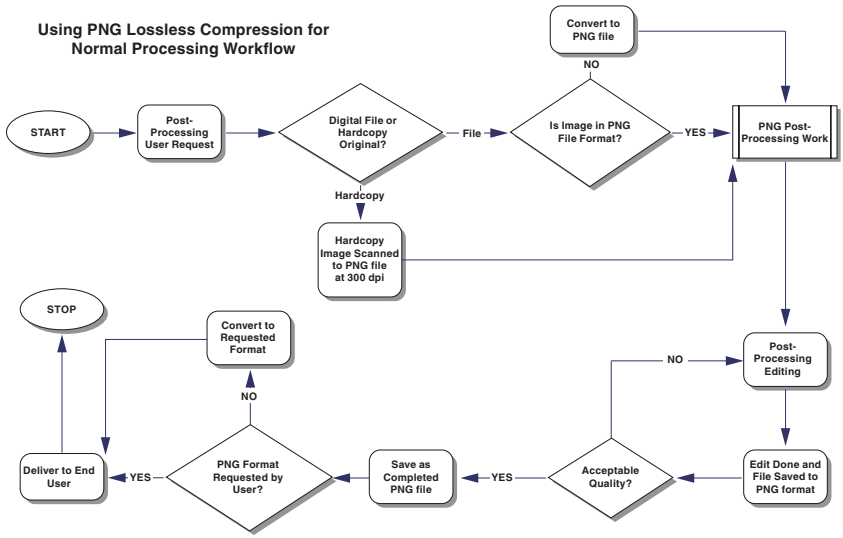

Figure 4.1 illustrates a suggested post-processing workflow based on standardized use of PNG lossless compression digital image files to minimize image editing degradation. At the beginning of the processing, non-PNG files are converted to PNG format, and hardcopy original photo prints or graphic images are scanned to PNG files. These are then directed into the same main post-processing workflow. At the end, PNG files can be converted to another digital image format that is required or preferred.

words, it lowers the efficiency of the PNG compression." (Although Chastain doesn't mention it, I like NCH Software's freeware Pixillon Image Converter, a link to which is available on the DPM website.)

- Chastain states that repeated JPEG file saves don't really add all that much cumulative image degradation if you consistently use a high-quality setting of JPEG compression, such as 95% or 100%. However, if you ever save any intermediate version of your work at a lower quality compression—for example, at 70%—you'll permanently lose image detail. Saving it later at 95% or 100% *does not* recapture the lost detail (Chastain, 2013a).
- She recommends using only a lossless format master copy in a format such as PNG or TIFF for any important archival storage, especially for files that you expect to edit again for future use. JPEG images lose quality each time they

are opened, *edited*, and saved again, so you should avoid archiving with this format. Always keep a lossless master copy of any image you expect to edit again. JPEG, she says, should only be used for archival storage when disk space is your primary consideration. (Given the low cost of today's storage space I don't think it should be: my Hitachi SimpleTech 500 GB external disk drive, for example, cost just $70 during the Christmas sale season of 2012.)

Chastain recommends PNG as a good archival alternative, because PNG files are considerably smaller than TIFF or BMP files. She writes, "PNG is an excellent choice for archiving. In fact, I use PNG for almost everything, especially when I need to port images between various programs" (Chastain, 2013b).

Digital Image Format Geek Knowledge Summary

At this point, I've pretty much conveyed all my geek knowledge about digital image formats. That's about all the technical stuff you need to understand to use the DPM post-processing work approach. As you can see, it's not really all that difficult; the practical guidance in this chapter mostly boils down to the following:

- TIFF and PNG formats give you the highest-quality images.
- Begin with the highest-quality image file (in whatever format) you can get or produce.
- Convert lossy format files to lossless format image files to preserve image quality during post-processing.
- If you must use the lossy JPEG format, consistently save intermediate file versions using the option for the highest-quality, lowest-compression setting that you can use.
- Save as many intermediate versions as you wish during a single editing session. An editing session *ends* when you close the file and exit the program, *not when you just save an intermediate file version for security backup.*
- Conduct as few discrete editing sessions as possible.

Don't use your photo-editing software's Save As function to convert file formats; it will add extraneous space within the file. Instead, use a good image format conversion utility.

Finally, for most purposes keep in mind that you really don't need to worry all that much about image quality so long as it "looks good" to you. Major quality differences aren't all that obvious to the naked eye except in very large photographic prints. For an idea of how difficult photo image quality is to assess, try out the JPEG quality perceptual test by Marcus Ranum (www.ranum.com/fun/lens_work/papers/jpegquality/index.html). You may be surprised at how you do on the test. (I certainly was, even after more than 50 years of experience with silver-based and digital photography.) Ranum's visual judgment exercise bears out Sue Chastain's conclusions that incremental differences in image quality are quite hard to identify using only simple visual examination.

References

Chastain, Sue. 2013a. "JPEG Myths and Facts,"*About.com Graphics Software*, graphicssoft.about.com/od/formatsjpeg/a/jpegmythsfacts.htm.

Chastain, Sue. 2013b. "png format vs jpg," *About.com Graphics Software. User Forum,* Message 1609.1, April 5, 2001. forums.about.com/ab-graphicssoft/messages?lgnF=y&tid=1609.

Fulton, Wayne. 2013a. "Image File Formats—JPG, TIF, PNG, GIF—Which to Use?" *A Few Scanning Tips*. www.scantips.com/basics09.html.

Fulton, Wayne. 2013b. "Color Bit-Depth, and Memory Cost of Images," *A Few Scanning Tips*. www.scantips.com/basics1d.html.

About the Author

Ernest Perez, Ph.D., is a native of San Antonio, TX, whose academic library experience includes working at the University of Texas at Austin Libraries, Southwest Texas State University Library, and the Texas Woman's University Library. In special libraries, he was Library Director at both the *Houston Chronicle* and the *Chicago Sun-Times*. He also managed the Circulation Department Automation Task Force for the *Chicago Sun-Times*, and was Circulation Department Administrator.

Ernest was a leader in the area of full-text database development at both newspapers, and was a recipient of the Henebry Award from the SLA Newspaper Division for his contributions in the automation area. A frequent contributor to the library professional literature, he has

authored over 100 articles and book chapters, acted as contributing and associate editor for numerous journals, and currently serves as Editor of *Library Computing*.

A serious longtime amateur photographer with a deep interest in digital image-processing tools and techniques, Ernest resides in Roseville, California. *Digital Photo Magic* is his first book.

This chapter is from Digital Photo Magic: Easy Image Retouching and Restoration for Librarians, Archivists, and Teachers*, by Ernest Perez, coming in January 2016 from Information Today, Inc.. For more information, visit http://books.infotoday.com.*