Preparing Images for Distribution

Judith A. Birchman & Susan G. Miller
Purdue University

Abstract

In today's publishing environment where online and CD-ROM based materials are as commonplace as printed materials, authors are faced with more options for presenting materials. Educators need to be aware of the demands of creating documents for their classes, publications and presentations. Today, materials are presented in a variety of ways and must be prepared accordingly.

Images such as illustrations, photos and charts are an integral part of most technical papers. For the person compiling the document for print or online venues, images present unique problems. Unlike text documents, which are readily transferable between applications and platforms, images need special attention. In order to create images which are an appropriate quality for distribution, the author must always be mindful of the format of the final output. The final format will affect how the images are created, scanned and submitted.

Familiarity with graphic file types, scanning resolution and color bit-depth is a must for today's authors. This paper will focus on a comparison of the issues involved with preparing and submitting images for print and on-screen venues.

Introduction

Over the course of their careers, most educators become well versed in preparing printed materials which include technical papers, course materials and presentation handouts. With the advent of CD-ROM technology and the Internet, preparing documents for publication has altered this process. Authors already experienced in document preparation now must adapt to the changing standards related to online documents.

As educators, we have grown accustomed to preparing camera-ready papers. In the past, we would cut and paste our text and graphics and mount them on boards to send to the printer. It did not matter what software or platform we used to generate the text and graphics for output. When software applications offered better layout features, we moved on to documents created entirely on the computer. We were able to merge the text and graphics and print a camera-ready paper to submit. Since then, we have expanded to include electronic submission of documents. At this point, factors such as file types and computer platform capability also become a factor. In order to submit papers to the *Engineering Design Graphic Journal*, for example, text and images must be submitted in formats that can be recognized by standard page layout software such as *PageMaker* or *QuarkXpress.*

Since illustrations, charts, and photos are an integral part of most technical papers, it is important to become familiar with the factors that influence the appearance of images in both print and web-based materials. In order to ensure the quality of the images, it is critical that authors consider how the document will be published. The specifications

for images used on the Internet versus those used for a printed document vary and must be appropriate to the publishing venue.

Graphics File Formats

When referring to file formats it is important to understand how the graphic is stored, what programs can read the graphic, and what can be done with it once it is opened. Some file formats are industry standard and are used by many different programs, and others are particular to a given program, like Adobe Photoshop's native file format, which is used almost exclusively by Photoshop itself. EPS (encapsulated PostScript) and TIFF (tagged image file format) are the two primary formats of interest to printing, while GIF, JPEG, and PNG affect those working most of the time on the Web or in multimedia (Blatner, Fleishman, and Roth, 1998).

There are three basic types of graphics file formats: Bitmap files, Vector files, and Metafiles. All of these file formats can move information from one application to another. All programs support their own "native" file format, and many can read or write files in other formats. Some programs can open or import files saved in the native formats of other programs. Whenever you move data between programs or save files in formats other than a publication's native format, a process of translation occurs. And, as in the translation of text from one language to another, the translated version will differ from the original. Sometimes the differences are subtle; sometimes they are more obvious (Kvern, 1997). In order to successfully move files between applications without losing quality, it is necessary to save files in an appropriate format.

The first type of graphic file format is a bitmap file. Bitmap files store pictures as rows and columns or squares known as pixels, with each pixel having a particular gray or color value. Bitmap files are typically created by image editing programs such as Adobe Photoshop, or by the software that is used to run a scanner. TIFF and BMP (Windows bitmap) are examples of bitmap-format graphics files. Continuous tone graphics and photographs with many color variations are examples of bitmap graphics.

The second type of graphic file format is a vector file. Vector files contain sets of instructions for drawing objects, typically geometric shapes such as lines, ellipses, polygons, rectangles, and arcs. Vector files come from two primary sources: drawing programs such as Macromedia FreeHand, Adobe Illustrator, and computer-aided design programs. DXF (drawing exchange format) files are examples of vector-format graphics files. Adobe PostScript paths and type, such as those found in EPS files, are other examples of vector elements, but they are usually contained in a metafile.

The final type of graphic file format is the metafile. Metafiles can contain both vector and bitmap graphics. Macintosh PICT, EPS, CGM (computer graphics metafile), and WMF (Windows metafile) formats are all examples of metafiles.

Digital Images

Digital images fall into two basic categories. The most important difference between print and online images is that they use two different types of digital data. Digital images can be represented as either vector or raster data (Duff, 1997).

As mentioned, vector graphics, also known as object-based graphics, are stored as mathematical descriptions. Vector images include line drawings and images with flat areas of color (*Figure 1*). Since vector images are stored as an object description, the image is independent of the resolution of the output device. In other words, the quality of the image does not vary even if you scale the object. PostScript is a vector-based graphic language. Vector graphics usually have smaller file sizes and the quality doesn't vary when resized (*Figure 2*).

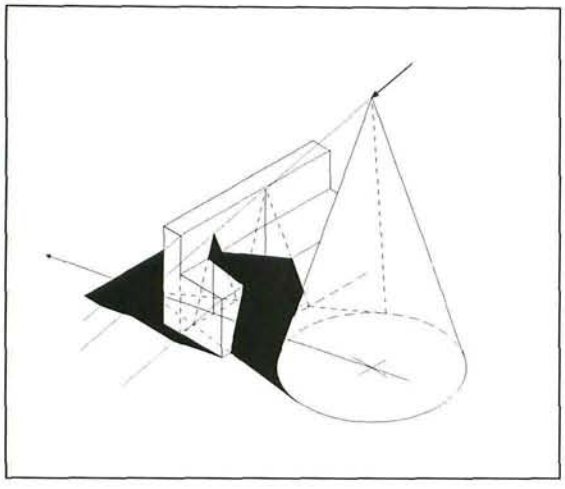

***Figure 1** - Example of a vector graphic.*

Images for online use are raster data. Raster data saves a bitmapped image as a series of dots, which is the same way it is displayed on a computer screen. Whereas, the quality of a raster image appears acceptable on-screen, it usually prints poorly on a high-resolution printer. The reason for this is that raster graphics only look good if printed at the original resolution. When printed on a higher resolution printer, they actually look worse, because as the pixels enlarge, they appear jagged (*Figure 3*). Low-resolution raster images can be used online but high-resolution raster images are required for print. Raster images have larger file sizes and lose quality when enlarged but are necessary for continuous tone and detailed images *(Figure 4)*.

The differences in these two types of digital data impact many factors related to preparing images for print versus the web. These include file formats, image resolution and color systems. In order to create images with an appropriate quality for print or web, you must understand the differences between the two types.

Images for Print

When preparing images for print, image resolution is the most important factor. A quality publication will be output to a high-resolution printer and low-resolution images may fall apart. When submitting files directly from disc or as email attachments for publication in journals, it is important to include everything your editor or printer needs. For example, if your file uses 2 typefaces, 12 pieces of clip art, and 5 scanned images, you will have 20 items to send (1 document file, 2 font folders, 12 clip art files, and 5 scanned files). PostScript files will contain embedded fonts and graphics, obviating the need to send copies of these support files to the printer. But as a precaution it is important to include everything your editor or printer needs.

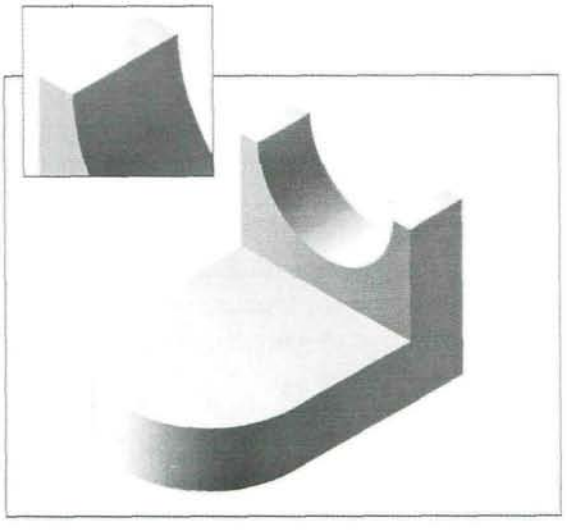

***Figure 2** - EPS at 100% and enlarged to 200%.*

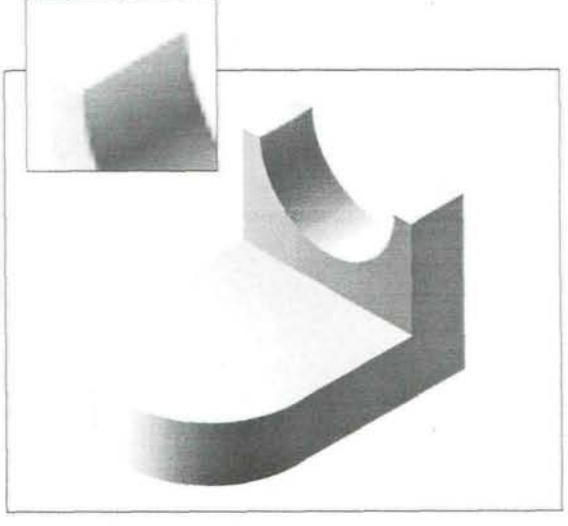

***Figure 3** - TIF at 100% and enlarged to 200%.*

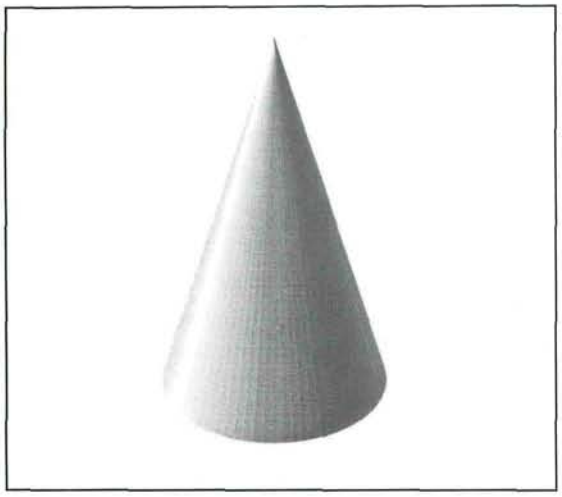

Figure 4. - Example of a raster graphic.

File formats

TIFF stands for *Tagged Image File Format* and is the most widely used, industry-standard bitmapped file format. It was developed by Aldus to transfer color and grayscale images into page-layout programs. TIFF can encode grayscale, RGB, indexed-color, or CMYK color models. It is best used for scanned, continuous tone and photographs. Files saved as TIFFs can be manipulated and changed in programs that import TIFFs. If used for vector graphics, the vectors must be rasterized or turned into dots, so the file size usually increases and you also lose the sizing advantage of a vector image.

Because the TIFF format is the most widely used format for graphic images it can easily move between computer systems; so TIFF files produced on a Mac may be moved to a PC (or vice versa) with few if any problems. TIFF format images may be produced at any resolution or size and may have any number of levels of gray or colors. Furthermore, TIFF images can be easily edited or modified.

The TIFF format does have one major limitation or disadvantage; TIFF files are larger than PCX files because TIFF files contain tags that communicate screen optimization to the printer. A PCX (abbreviation for PC Paintbrush Extension) file was one of the first graphic file formats for the PC.

EPS stands for *Encapsulated PostScript;* it was developed as a generic PostScript file for transferring images into page-layout programs. Similar to the TIFF format, EPS files can be easily moved from one environment to another. Although EPS is a vector format, it can include bitmapped images, making it a metafile format. The data in an EPS file cannot be changed and manipulated in other programs, changes must be made in the native format.

EPS format should primarily be used for line drawings. The biggest disadvantage with EPS files is that they can only be printed to a PostScript printer. Bitmapped images take twice as much disk space in the EPS format than they do in other formats such as TIFF. EPS files can not be easily edited.

Another important consideration in choosing file formats is if you want to place a scanned image into a page layout program and adjust the tonal levels. The tonal values in an EPS file are fixed, and cannot be altered or tweaked unless the image is opened for editing by an image processing program. The image needs to be saved as a TIFF if you want to change the tonal values on the scanned image, this is what makes TIFF so powerful.

Image resolution for printing

Knowing the final destination for a scan will yield an image that contains the right amount of information, has the right amount of color fidelity, and is saved in the most suitable format. It is important to consider how the scan will be used; on the Internet, in a multimedia presentation, or in a printed document.

Images for print tend to be the most demanding; they require the largest file sizes, the highest resolutions, and the greatest number of colors unless printing in grayscale. Screening is a necessity for print output,

because a printing press can't directly reproduce multiple shades of a given ink color (Ihrig,1996).Therefore, it is important to become familiar with several resolution measurements when scanning, such as scanning resolution, screen resolution, printing resolution, and lines-per-inch resolution.

Scanning resolution is normally expressed as pixels-per-inch (ppi) or dots-per-inch. Once the scanning resolution is set and the scan is made it cannot easily be changed. So, line art should be scanned at the resolution of the printer that will be used for final output. As a guideline for scanning it is important to have some idea of the final size the scan will be used, therefore scale your image accordingly before scanning. For example, if the image you need to scan is extremely small it is important to remember to enlarge the image to as close to the final size as possible while scanning. This will help the overall resolution of your image if it is sized at the time it is being scanned rather than after it is placed in an application and scaled. As a rule of thumb, scanning at 150-300 dpi is acceptable in most instances as long as the image is scanned at the final size. Otherwise, if you are not sure what resolution to use, check with the printer or the editor.

Printing resolution is also measured in dots-per-inch and depending on the type of printer the dpi can range anywhere from 72 dpi to 3,000 dpi. So if the final output is a 600 dpi laser printer an image can be scanned at 300 dpi. Although printers can work as high as 3000 dpi they typically work somewhere around 1270 dpi because it still allows for a very high resolution and the visible difference between 1270 dpi to 3000 dpi is hardly noticeable at all.

Lines-per-inch (lpi) resolution (screen frequency) determines how large or small the dots will be that are created by the halftoning process. For example, a 300 dpi laser printer will have a lines-per-inch output of approximately 60 lpi and a 600 dpi laser printer will have from 85 to 100 lpi. As a result it is not necessary to scan grayscale images at the highest possible resolution or to scan at the resolution of your printer. The controlling factor is not the resolution of the scanner or the printer, but the halftone screen frequency (Day, 1993). The screen ruling (in lines per inch) is usually based on the type of paper being used: the rougher the paper, the coarser the screen required. Newspapers printed on newsprint might use 85 screen, on uncoated offset paper 120 or 133 screen, and on matte-coated paper 150 screen.

When scanning images, it will help to know the lpi at which they will be printed. To achieve the best reproduction, a ratio of 2:1, pixel to line screen, is best. For example, an image scanned at 300 ppi would look best if screened no more than 150 lpi. Generally, if you scan images at 1.5 times the line screen, you will get satisfactory results. The goal is to attain the highest number of grays for printing. The halftone line screen and the resolution of the printer together determine the number of gray levels available. As a rule of thumb, the lower the lpi the higher the printer resolution should be in order to attain more levels of gray. At a fixed printer resolution, the lower the lpi, the higher the number of levels of gray. At 300 dpi and 60 lpi,twenty-five grays (plus black & white) would be available. If the line screen were increased to 100 lpi, the grays would drop to nine (plus black & white). This means that if your image needs more grays (many tonal values) to look it would require a lower lpi to print satisfactorily on a 300 dpi printer. To assist you in preparing your images, check with the printer or editor you are working as to the printing specifications.

Print color system

A major difficulty in working with color documents is the difference between the colors seen on the screen and the colors that will actually be printed on the page. Unfortunately, computer monitors and printers use different systems to produce colors.

Monitors use the RGB model, in which varying amounts of red, green, and blue are mixed together to create a given shade. Printers generally use Pantone or spot colors and process colors. Pantone or spot colors are created by blending different inks together to get a required hue and are referred to by a specified number. Pantone colors are typically used for printing one or two color jobs that consist of halftone images and line art. Process colors use the CMYK model which stands for cyan, magenta, yellow, and black (K is used to denote black, because B means blue in the printing industry). Process colors use dot patterns of the 4 process colors (CMYK) to simulate the desired hue. Process colors are used for printing color photographs and illustrations.

Scanning and File-Format Tips

Save your bitmap images in the appropriate color space. Use the right type of TIFF for your publication. When you scan line art (black and white images), you will save hard disk space by saving your TIFF files as Bitmap, rather than saving them as grayscale or color TIFFs. If you are printing only black ink, your images might as well be saved as grayscale TIFFs, rather than color, because you will once again save space on your hard drive.

Another tip to consider is to size and scale the images when scanning to the size they will be when printed or placed in applications. Resizing images after they have been scanned can produce unacceptable results such as "jaggies." Scanning images at the size they will be used also saves disk space. In addition, cropping out any unnecessary white space when scanning will also save disk space. White space creates data and takes up disk space (Day, 1993).

When you want to move artwork from one application to another, don't copy and paste. Instead, export or save the artwork, then place (or import) it in the destination application. There are three reasons for not using the copy and paste commands. First, the formats that most applications put on the Clipboard (Windows or Mac) are usually "stripped-down" versions of their formats which could lead to not getting what you need to print the artwork correctly. Secondly, copying and pasting artwork from one application to another generally doesn't leave you with an externally linked graphic, which is something you might need later to edit the artwork. Finally, you don't always get what you expect from the Clipboard. Sometimes you will get a bitmap format when you expect a vector format, or vice versa.

Images for Web

When preparing images for web use, remember that you are designing for the computer screen. Unlike the printed page which is viewed the same way by everyone, on-screen documents are viewed on a variety of computer systems. The main challenge when creating on-screen images is making them small enough that they download quickly with acceptable quality on a variety of platforms and systems.

As mentioned, web graphics are predominately raster-based images. Since raster images typically have large file sizes, the primary goal is to reduce the file size without sacrificing the quality of the image. By understanding the following issues you can optimize both the size and appearance of your images.

File formats

There are three main file types used for web graphics GIF, JPEG and PNG. All of these file types are usable across platforms and they are compressed files, which is what makes them so popular for use on the web. Each file type has its advantages and disadvantages. It is important to use the file type which is most appropriate for the type of image in order to attain the best quality.

GIF stands for *Graphic Interchange Format* and was developed specifically for online

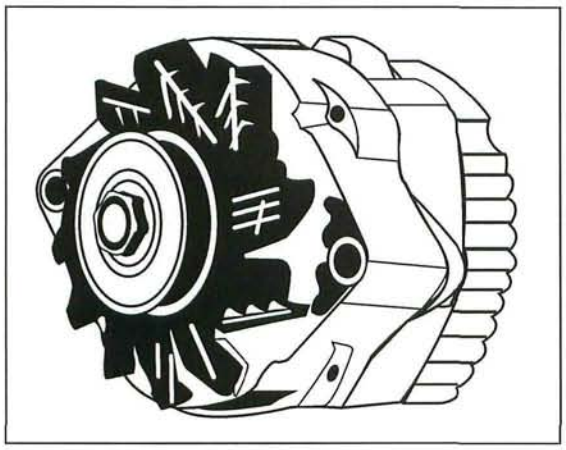

***Figure 5** - Line art is best saved as a GIF.*

use by CompuServe to compress graphics. GIF images are small in file size and use lossless compression schemes, which means that the image does not lose quality when compressed and decompressed.

One of the advantages of GIF files is that they allow you to make select colors transparent which lets you remove the background color for better integrating the graphic on the page. Another advantage is that GIF images can be interlaced which allows the image to gradually display as it loads. They can also be used to produce animations.

The limitation of the GIF format is that it uses indexed color, which limits it to a palette of 256 colors (8-bit). This makes it most suitable for line art, charts and images with areas of solid color (*Figure 5*). Photographs can also be saved in this format but they need to be reduced to 8-bit, indexed color. With a smaller sized photo, the reduction in quality will not be as noticeable. Also, adaptive 8-bit palettes may be specified which predetermine the colors available based on the photo itself.

JPEG stands for *Joint Photographic Experts Group* and was designed to compress continuous tone photographs. JPEG files use a lossy compression scheme, which actually removes data from the image in order to reduce the file size. It is best to retain the file in its native format as you experiment with saving it at various quality levels.

The advantage of JPEG files is that they can contain millions of colors and do not rely on color palettes.

The disadvantage of the JPEG is that they cannot be made transparent or interlaced. In order to use transparency for an edge effect for example, the photo would have to be saved as a GIF. Browsers are also starting to support progressive JPEG which produces an interlaced effect. This can, however, increase the download time due to decompression issues.

Use the JPEG format for photographs, grayscales and artwork that has subtle changes of color (*Figure 6*). It is best not to use it for images with areas of flat color or line drawings, as it does not compress solid colors well.

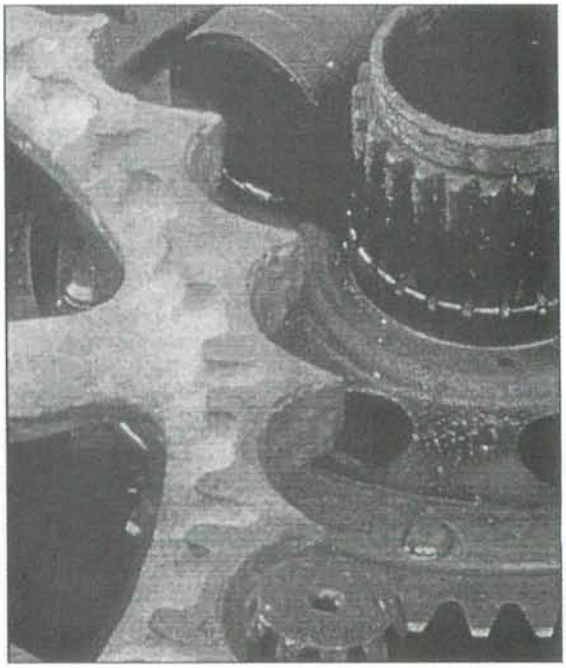

***Figure 6** - Photos for the web are best saved as JPEGS.*

PNG stands for *Portable Network Graphics* and is the newest file format. Like the GIF format that it was intended to replace, it also uses lossless compression and produces even smaller file sizes. PNG is still not fully supported by all browsers although it offers advantages over the GIF format.

PNG files can be saved at a variety if bit-depths—8-bit, 24 bit or 32-bit. It also allows for multiple layers of transparency since it supports alpha channels. This means that problems with halos due to anti-aliasing can be eliminated. It also does a better job of interlacing, as it loads the image faster and displays it sooner. Another advantage is that it adjusts the brightness of the image according to the monitor it is being viewed on. This is due to the fact that that it includes *gamma* or brightness information as part of the image.

As you can see, the only drawback at this time is the lack of browser support for this format. Although plug-ins are available, it is not widely used by web developers at this point in time.

Resolution of the Image

On-screen images do not require a high resolution image like printed documents. Screen resolution, like print resolution, is measured by dots-per-inch or pixels-per-inch. A Macintosh has a screen resolution of 72 dpi and a Windows computer screen has a resolution between 40 to 95 dpi. An image that is scanned or created at a higher resolution than that of the screen will only be able to be viewed at the lower-resolution of the screen. Since all images must be rasterized in order to be displayed on a web page, images at 72 ppi will look fine. An image that looks great on the web at this resolution, however, would not look good if printed on a high-resolution printer.

When preparing images, it's okay to work at a larger size and then sample down. In other words, you can create the image at 150 dpi for print and change it to 72 dpi when you are ready to add it to the web page. Be sure that you size the image **after** you change the resolution to 72 ppi or it will appear smaller when inserted in the web page. Scan images at 150 dpi at the approximate size or at a larger scale and down-sample later. Screen captures used in print materials are a special problem. Since they are low-resolution images and usually color or grayscale, they may not print well. To get a quality print, it would be necessary to have the printer create a halftone as he would for a photograph.

On-screen color

Since monitors use RGB (Red, Green and Blue) color schemes, create images for the web using RGB colors. The most important factor related to color and on-screen images is the bit-depth. A 24-bit image will look better, but take longer to download than an 8-bit image (*Figure 7*). The goal is to reduce the number of colors displayed just enough to reduce the file size of the image without degrading the quality of the image too much. Since the color capabilities of the viewers systems will vary, it is best to plan in advance for these variations.

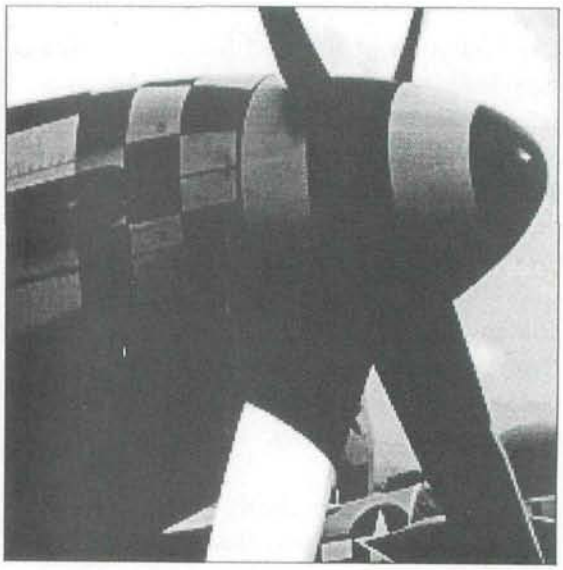

***Figure 7** - 24-bit image.*

Saving images as indexed color is the best way to keep the file size down. Indexed color mode limits the palette to a maximum of 256 colors (*Figure 8*). Images saved as GIF files must be saved in index color mode. When the colors are reduced, *dithering* is used to simulate colors which are not available. In other words, colors are created by the juxtaposition of colors from the existing palette. The image appears to have a pixelated look after switching to index mode, but it is still acceptable quality. If dithering is not specified, and the page is viewed on a monitor with lesser bit-depth, *banding* (*Figure 9*) will occur. Since the image contains more colors than the computer can handle, colors are still substituted but with unpredictable results (Duff, 1997).

Through experimentation, it might be possible to save the image at a lower bit-depth and even further reduce the palette size as well as the file size. When specifying the palette of colors, an adaptive palette can be specified which would pick the best colors based on those used in the image (Weinman, 1997). Additionally, there are browser-safe palettes which contain the 216 colors common to multiple browsers and operating systems (Williams & Tollett, 1998). These browser-safe palettes are recommended if you are creating your own images. In most cases you will probably be scanning images and then your main concern is to reduce the bit-depth and file size when you save it. You can achieve this with a little experimentation with palettes and using the appropriate file format as discussed earlier.

Summary

If you're creating a document that will never travel across platforms and won't need to print to PostScript devices, you may achieve perfectly acceptable results from WMF in Windows or PICT on the Macintosh (these are the platform-specific metafiles for the two operating systems). But if you will be printing to PostScript devices, use EPS for illustrations and TIFF for raster images. For artwork you want to include in an online publication, use GIF or JPEG. Use GIF is the graphic in question contains primarily vector artwork or contains large areas of flat color. Use JPEG if you are dealing with a bitmap graphic images such as scanned photographs that contain subtle, natural transitions between colors. *Table 1* provides guidelines to follow when deciding which file type to use (Kvern, 1997).

***Figure 8** - 8-bit image with dithering.*

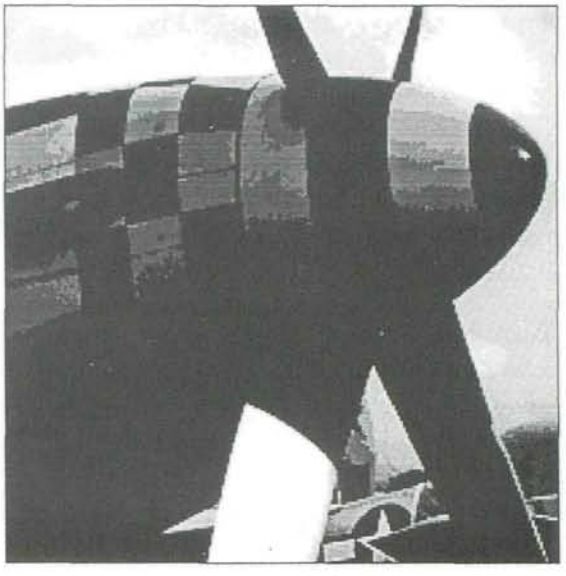

***Figure 9** - 8-bit image without dithering.*

Publish - Print (PostScript)	**File Format to Use**
DXF, vector, metafiles (bitmap & vector)	EPS
PCX, BMP, metafiles (bitmap)	TIFF
Publish - Web	**File Format to Use**
DXF, metafiles (vector)	GIF or PNG
PCX, BMP, TIFF, metafiles (bitmap)	GIF or JPEG or PNG

Table 1 *- File format guide.*

References

Blatner, D. & Fraser, B. (1996). *Real World Photoshop 3*. Berkeley: Peachpit Press.

Blatner, D., Fleishman, G., & Roth, S. (1998). *Real world scanning and halftones*. Berkeley, CA: Peachpit Press.

Day, J. B. (1993). *Super scanning techniques*. New York: Random House.

Duff, J. (1997). *Web page design*. Indianapolis: Sams.net Publishing.

Hamlin, J. (1997). *Photoshop web techniques*. Indianapolis: New Riders Publishing.

Ihrig, S. & Ihrig, E. (1996, November). What goes in must come out: Getting the best scan. *Adobe Magazine*, 66-68.

Kvern, O. M. (1997, Spring). The right match: Don't get burned by the wrong graphics file format. *Adobe Magazine*, 57-60.

Rich, J. & Bozek, S. (1995). *Photoshop in black and white*. Berkeley: Peachpit Press.

Weinman, L. (1997). *Designing web graphics (2nd ed.)*. Indianapolis: New Riders Publishing.

Williams, R. & Tollett, J. (1998). *The non-designer's web book*. Berkeley: Peachpit Press.