Introduction to digital images

Digital Image File Types

JPG, GIF, TIFF, PNG, BMP. What are they, and how do you choose? The technical details are in the appendix to this presentation, if you are interested enough to read on. In simple terms there are different methods of compressing very large digital images to a much more manageable size, depending on the number of colours in the image.

Some of these algorithms manage to retain the accuracy of the original: these are called lossless compression. Others allow some degradation to make the file size smaller and are called lossy compression.

File size comparisons

Below are comparisons of the same image saved in several popular file types.

File type	Size
Tiff, uncompressed	901K
Tiff, LZW lossless compression (yes, its 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 colours	286K

JPG vs GIF

The two most popular image formats for the web are JPG and GIF. These formats represent very different approaches to the problem of delivering an image with minimum file size and download time. While GIF is great for computer generated images with limited palettes, JPG is far better for photographs. It gives better quality images for the same file size. Below is a comparison. The image on the left is a blowup of a very small region of a large image compressed with JPG.

The image on the right is the same region compressed using GIF. The GIF file is 2.4 times bigger than the JPG, but is clearly of lower quality.

The best quality output from a digital camera is normally 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 "open, 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.

Image editing

There are literally hundreds of image editing programs available, from free to expensive and with capabilities ranging from practically useless to unbelievably clever (and complicated).

Before investing a lot of money in one of the more complicated image editing programs you should look at the editing abilities of the Office applications in which you are likely to use the end product. Outlook, Word, PowerPoint and Publisher are all excellent when it comes to basic editing such as cropping and can also add interesting effects to your images.

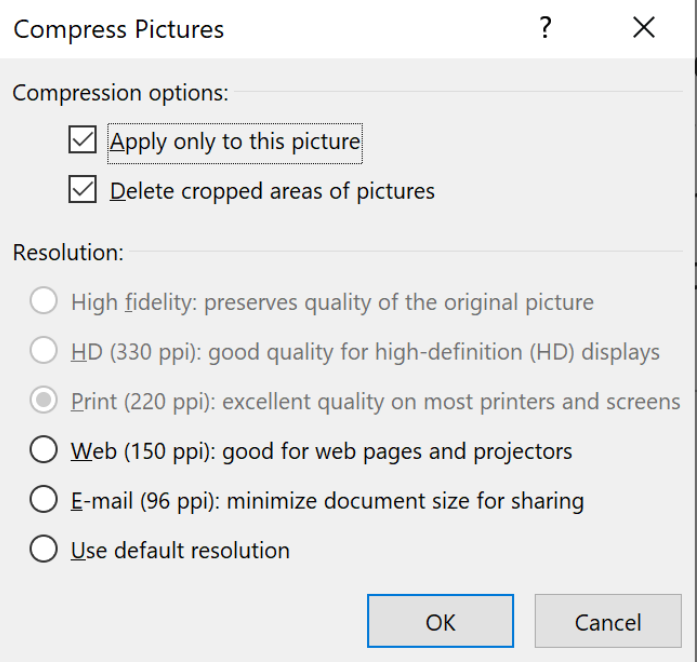Reducing the size of a PowerPoint presentation. (Or indeed Word or Publisher document, or Outlook email.)

If you open any file containing images it is possible to compress the pictures: this reduces the file size without making the images physically smaller. Click on any image, then click on "Picture Tools" (highlighted in green), then on "Compress Pictures" (highlighted in purple).
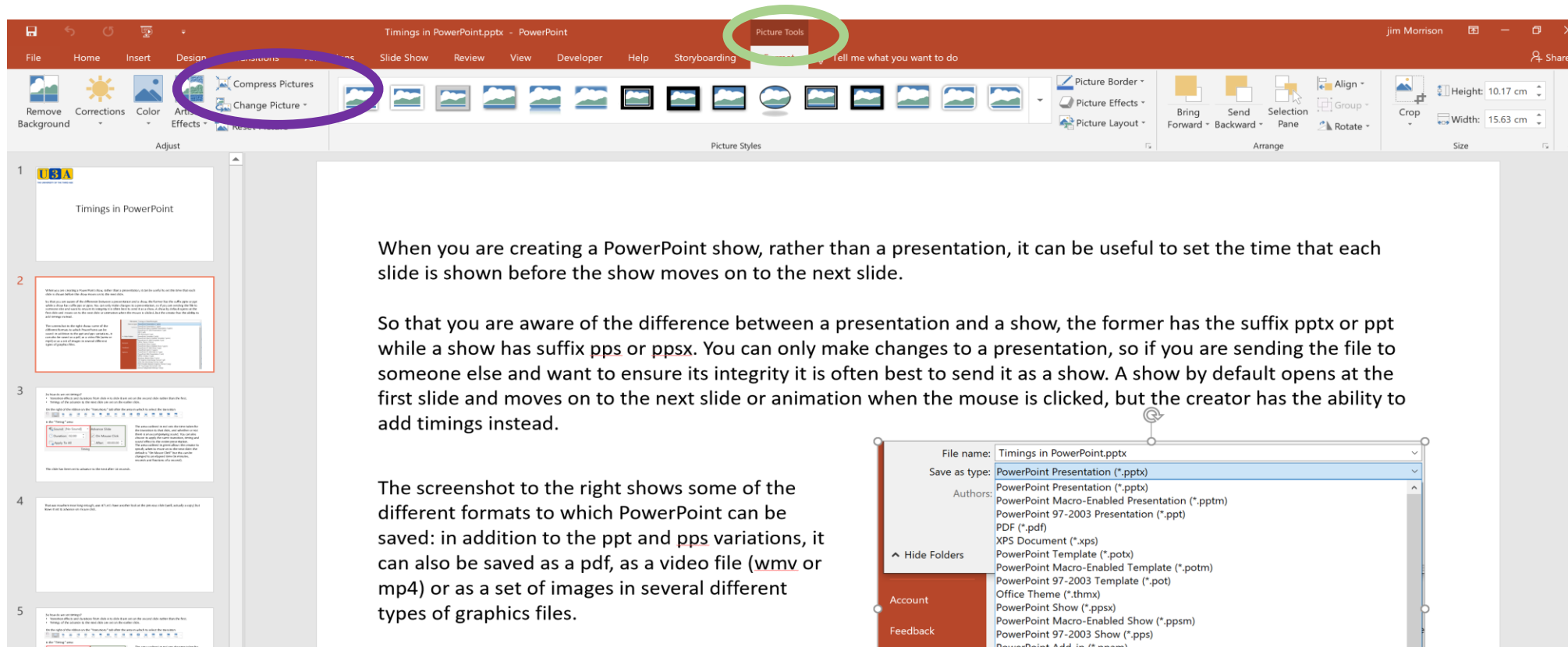
This will open a new window showing the different image qualities available. (In the screenshot below, the image is already less than print quality so the first 3 options are not available.)

It is important to note that although you are almost certain to want to compress all the pictures in your show or document, PowerPoint annoyingly defaults to reducing the size of only the selected picture.

Ensure that you untick this box as well as selecting the resolution.

The screenshot below shows the file sizes which can be achieved with each of the different settings when applied to a very large (or even ridiculously large) presentation.

Compress Pictures   ?   ✕

Compression options:
☑ Apply only to this picture
☑ Delete cropped areas of pictures

Resolution:
○ High fidelity: preserves quality of the original picture
○ HD (330 ppi): good quality for high-definition (HD) displays
◉ Print (220 ppi): excellent quality on most printers and screens
○ Web (150 ppi): good for web pages and projectors
○ E-mail (96 ppi): minimize document size for sharing
○ Use default resolution

OK   Cancel

| File	Type	Size	Description
America_base.pptx	Microsoft PowerPoint...	1,145,667 KB	Original
America_fidelity.pptx	Microsoft PowerPoint...	403,513 KB	High fidelity (35%)
America_high.pptx	Microsoft PowerPoint...	349,510 KB	HD (30%)
America_print.pptx	Microsoft PowerPoint...	222,457 KB	Print (19%)
America_web.pptx	Microsoft PowerPoint...	122,948 KB	Web (10%)
America_xmail.pptx	Microsoft PowerPoint...	56,310 KB	E-mail (5%)

Vector formats

As opposed to all the previous image formats, which are called raster formats (where the data describes the characteristics of each individual pixel), vector image formats contain a geometric description which can be rendered smoothly at any desired display size.

It is possible to print a vector logo on a small sheet of copy paper, and then enlarge the same vector logo to size A0 or even much bigger and keep the same crisp quality. The picture on the right illustrates the difference between vector and raster images under extreme magnification.

Specialised software such as Adobe Illustrator (£20 per month), CorelDraw (full version £600) or Inkscape (free) is needed to work with vector graphics. At some point, all vector graphics must be converted to raster images in order to be displayed on digital monitors.

7x Magnification

Vector

Raster

Ice Cream
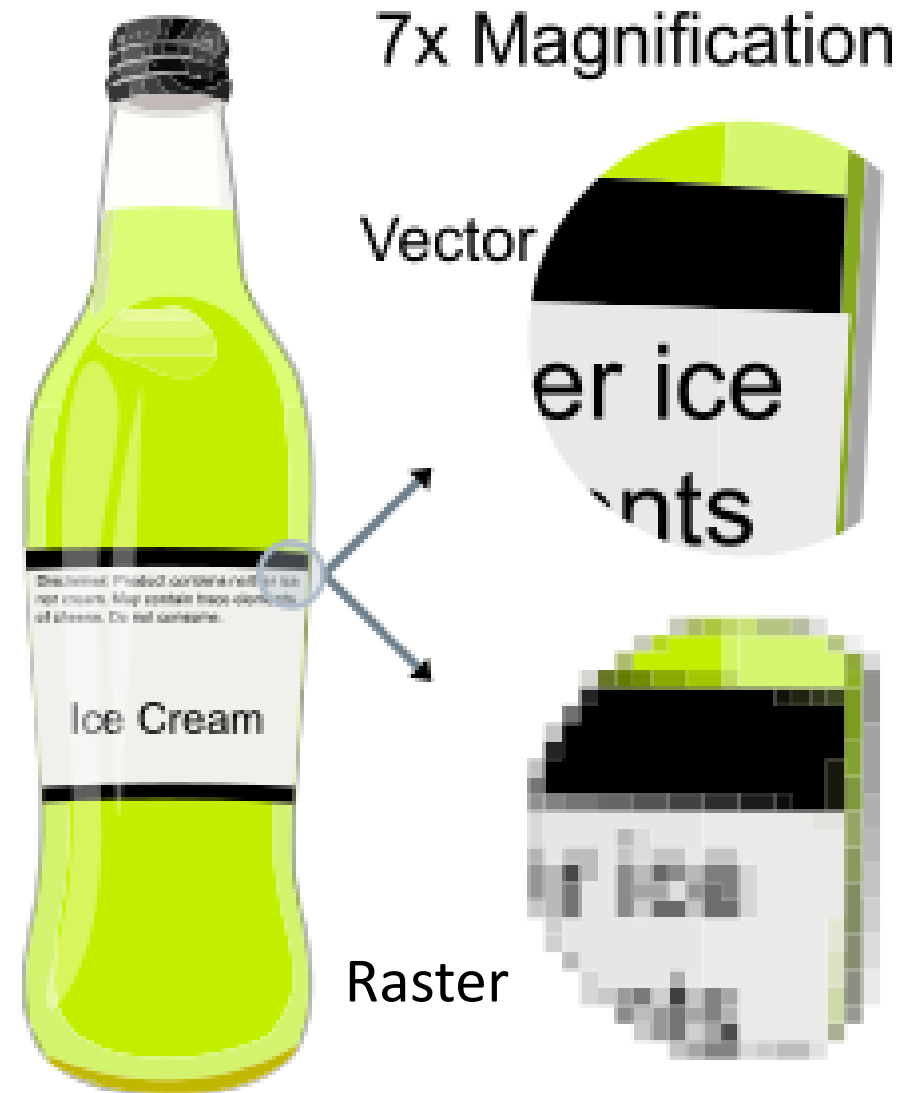
THE END

But please continue if the really technical stuff interests you.

The technical stuff

Digital Image File Types

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. One reason for the many file types is that images differ in the number of colours they contain. If an image has few colours, a file type can be designed to exploit this as a way of reducing file size. Another 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 they take up more disk space and are slower to transfer and download. Compression is a term used to describe ways of cutting the size of the file. Compression schemes can be *lossy* or *lossless*: horrible Americanisms explained on the next slide.

Lossy vs. 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. (Still awake?) 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 colour information at a lower resolution than the image itself, since the eye is not so sensitive to small changes in colour.

Number of colours

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

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 use 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 colours from a pool of 16 million. If the image has fewer than 256 colours, GIF can render the image exactly. When the image contains many colours, software that creates the GIF uses any of several algorithms to approximate the colours in the image with the limited palette of 256 colours available. Better algorithms search the image to find an optimum set of 256 colours. Sometimes GIF uses the nearest colour to represent each pixel, and sometimes it uses "error diffusion" to adjust the colour of nearby pixels to correct for the error in each pixel. GIF achieves compression in two ways. First, it reduces the number of colours of colour-rich images. Second, it replaces commonly occurring patterns (especially large areas of uniform colour) 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 colours or fewer. For a rich, true colour image, GIF may "lose" 99.998% of the colours.

JPG is optimized for photographs and similar continuous tone images that contain many, many colours. 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 colour. 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. Some 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 those programs, 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 non-proprietary formats such as TIFF and JPG. However, be sure to save your end result as a standard TIFF or JPG, even if you use a proprietary format as well, or you may not be able to view it in a few years when your software has changed.

When to use each file type?

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 "open, 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.

PNG is of principal value in two applications:

1.If you have an image with large areas of exactly uniform colour, but which contains more than 256 colours, PNG is your choice. Its strategy is similar to that of GIF, but it supports 16 million colours, 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, producing smaller files and allowing more colours.

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.

A new format WEBP is popular for high-quality web images with smaller file size.