Metadata | Part C

Talking about saving space.

You might have across a number of picture formats, such as bitmap BMP, JPEG aka JPG, and GIF. The main differences between them are in the way they store colour information in a file. Pure colour is stored in something called "raw" images – very large size files that come straight out of expensive professional cameras. We don't usually need that much data in our pictures, so we have different ways to save space.

Bitmap files rely on the fact that we cannot tell the difference between millions of colours. Red, green and blue process allows for thousands of small variations in colour, most of which will never be used. So, if we defined a smaller number of colours, a palette of acceptable colours, we don't have to use as many bits to store them. Bitmap files are the basic and the least efficient of file formats – they just store a grid of pixels, each pixel is assigned a colour from the colour palette – or "mapped" to it, hence bitmap.

We can drastically cut the number of colours to save space. For example, only 256 colours are allowed in the GIF format. Then instead of using 24 bits, 8 for red, 8 for blue, 8 for green, we can only use 8 bit – the pixel will not contain RGB values anymore but only the "colour number" let's say with black being colour 1. We can express it as 0000 0001 and white being colour 255 - 1111 1111 - so 66% of memory space is saved right away.

This is known as lossy compression. You know it's a lossy compression when, as you keep on opening it and resaving in the format the quality gets worse and worse and never looks like the original. The better one is lossless compression, like TIFF or ZIP files which can restore full quality if needed but they take up more disk space and require more powerful computers.

GIF files are better for large areas of uniform colour and when the total number of colours is less than 256 – like cartoons or computer generated pie charts. GIF can make a perfect copy even at high compression giving a small file size. At the same time, JPG is far better for photographs which have many shades of the same colour – that's why it is hard to remove backgrounds from JPEG but GIFs are much easier and are even available as "transparent GIFs" – without any background. But for photographs, JPEG gives better quality images for the same file size. By using "interpolation" – by making similar colour pixels to be actually the same, JPG file can just say "repeat that pixel 5 more times", and save memory. These use different compression techniques so depending on the settings they may degrade the image in different ways.

Most modern cameras generate JPEG files by default, so some quality is lost right away, especially in low light conditions. Our eyes are more sensitive to bright colours which means we will notice if some of the bright colour bits were thrown away.