

Scans and bitmaps in T_EX

Siep Kroonenberg*

n.s.kroonenberg@eco.rug.nl

Abstract

Guidelines are given for working with scans and bitmaps, with the emphasis on PostScript printing from T_EX.

Keywords: Scan, image, bitmap, resolution, halftone, PostScript

Not so many years ago, the only graphics the average computer user got his hands on were graphs from spreadsheets and presentation programs and such, and bitmaps were a last-resort format for these graphics. Nowadays, computing has turned graphical with a vengeance, and many users are in a position to incorporate scanned photographs and other bitmapped pictures into their documents. Time to take a closer look at bitmaps and images.

The emphasis in this paper will be on PostScript printing. When I started using PostScript and Ghostscript, I was delighted at the simple logic of PostScript-based graphics. BM2FONT and emT_EX specials became a bad dream quickly to be forgotten.

Colour models

A bitmap is a grid of pixels which each have colour value. There are bilevel or monochrome bitmaps, grayscale and RGB bitmaps, which require resp. 1, 8 and 24 bits of storage per pixel. There are also palette-based or indexed-colour schemes: the image contains a palette of typically 256 (sometimes fewer, never more) RGB values and the colour of a pixel is represented by a palette index, therefore requiring no more than eight bits.

Tip. Although a scanner may be able to save scans in indexed colour format, it is better to scan in RGB mode and do colour reduction afterwards, e.g. in Photoshop or PaintShop Pro.

Other three-dimensional colour models are HSB (hue/saturation/brightness) and CIE-Lab. CIE-Lab is intended to be device-independent, capable of bridging the differences in colour gamut between different devices (monitors, scanners, colour printers, the human eye). PostScript level 1 does not support the CIE-Lab model.

For print purposes, there is the CMYK colour model, in which colours are defined by percentages of the process colours cyan, magenta, yellow and black. There are many issues involved in RGB-to-CMYK conversion, but when

simply printing to a colour printer, the conversion can usually be left to the printer.

File formats

File formats differ in their coding of pixel information, in compression schemes, in colour models supported and in the additional information they can hold.

TIFF is popular for professional print publishing. It supports all colour models described above and various compression schemes. It may require some trial and error to find out which options work best between two applications.

JPEG supports grayscale and RGB. Its distinguishing feature is its *lossy* compression scheme; in fact, JPEG is both the name of a bitmap format and of this compression scheme. JPEG compression is highly effective even at high quality settings. JPEG compression is designed for (scans of) photographs: such pictures don't compress well with conventional methods but suffer little from JPEG compression. Don't use it for other types of pictures such as graphs, diagrams, logos or screenshots. Also keep in mind that degradation may get worse with successive cycles of editing, saving and reopening: apply JPEG compression once, at the end. JPEG is the preferred choice for publishing photographs on the web.

GIF only supports indexed colour and has a very effective, lossless compression scheme. Its principal use is web pictures. A new alternative is *PNG*, which supports RGB in addition to indexed colour (and nobody claims licencing fees for it). Both GIF and PNG support transparency and interlacing.

For T_EX however, the most important format is probably *EPS* or *Encapsulated PostScript*. The basic idea of EPS is that applications can pass the EPS file to the printer as-is: as long as the originating application can explain its meaning to the printer, the graphic capabilities of the 'host' application, i.e. T_EX, are irrelevant. T_EX can still add wrapper code for positioning, scaling, rotating and cropping.

*This paper is an adaptation of 'Werken met scans and bitmaps' that appeared in our departmental newsletter EZ, 1996/5.

Note that the ‘natural’ dimensions of an EPS bitmap don’t depend on printer resolution.

The EPS format is bulky and has two limitations: it can only be printed to a PostScript printer, and as a rule EPS pictures cannot be read by image-processing software. One important exception is that Photoshop can read back the eps files that it writes.

And then there is a multitude of special-purpose and niche formats such as IMG, BMP, PCX, XMB etc. Use these if you have to.

Resolution

For screen representation, we measure pictures usually in pixels, and picture pixels are preferably mapped one-to-one to screen pixels. For PostScript printing, it makes more sense to talk about physical dimensions on the one hand and resolution on the other, although of course a 1-by-1-inch 600 dpi picture contains the same amount of pixel data as a 2-by-2-inch 300 dpi picture.

Black-and-white bitmapped line art should ideally have printer resolution (after any scaling of course): we want to map image pixels one-to-one to printer dots. Unfortunately this often means a very large number of dots, which may compress to a small file but still have to be expanded inside the printer’s memory. Unlike most or all non-PostScript printers, a PostScript printer will do a lot of interpolating to calculate which dots should be on and which should be off, and might choke on a large monochrome bitmap. Only if it is very clever it might recognize the special case where image dots match up exactly with printer dots. So check with your printshop or service bureau which resolutions they recommend. Below, you see a 300 dpi- and a 600 dpi version of the same picture.

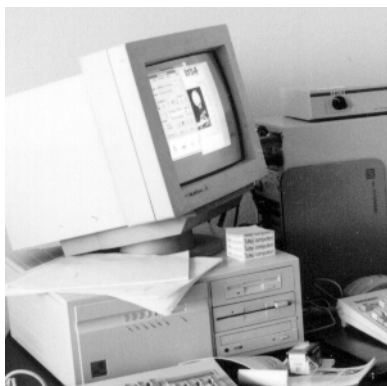


If your black-and-white image is really a halftoned or dithered grayscale picture, make sure that each image pixel maps to an integral number of printer dots, in order to avoid moire effects; see below. Or get hold of the original grayscale image.

Resolution



75 DPI, 48 KB



150 DPI, 183 KB



300 DPI, 725 KB



For screendumps it is also a good idea to map each image pixel to an integral number of printer dots: screendumps frequently contain dithered areas, and text will also look more even.

Tip. You can set the desired resolution in Photoshop, export the image in EPS format and then place the image in your T_EX document without scaling:
`\includegraphics{picture.eps}`

For photographs, choice of resolution is a different story. Normally you won’t need resolutions exceeding 300 dpi even for high-end output: photographs will be rasterized, and resolution is mainly determined by the screen- or ‘line’ frequency. Pick a resolution of 1.2 to 2 times the line frequency: the higher number if there are sharp transitions, the lower for e.g. cloud pictures:

Device	device dpi	lpi	image dpi
screen	96	n/a	96
laserprinter	300	53	60–100
laserprinter	600	85	100–150
imagesetter	1270	128	150–250
imagesetter	2540	150	200–300

Some line frequencies match up better with the printer resolution than others, and the printer or imagesetter might make adjustments accordingly. Or the printer may support FM (Frequency Modulated) screens or variable dot sizes, which features make the concept of line frequency meaningless.

In the resolution example, the difference in quality between the 150 dpi- and 300 dpi versions is exceedingly insignificant; the difference in file size isn’t.

Gray levels

Higher line frequencies make for sharper pictures. They also mean fewer grays. A few possible screen settings for a 600 dpi printer:

screen	halftone cell	grays
60.6	98	99
70.7	72	73
84.9	50	51 (default)
106.1	32	33
141.4	18	19

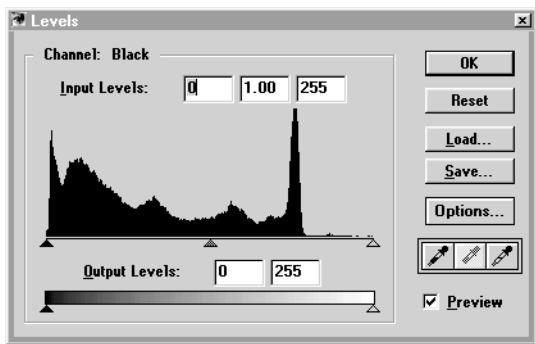
The halftone screens example at the end illustrates some of these settings. These are for 45° screens; for a 0° screen we would have a different series of frequency settings. A full range of 256 grays and a line frequency of 150 lpi requires an imagesetter resolution of at least 2400 dpi. Better halftones are the principal reason for high imagesetter resolutions.

This is not the whole story. Some printers can vary dot size and produce much better grays than the table indicates; a LaserJet 4 printout showed none of the expected stairstepping. On the other hand, the reproduction process imposes limitations of its own on screen frequencies.

You can set a document-wide screen by using Dvips' custom header files. The following code should set a 53 lpi 45° screen:

```
%!
/bop-hook {53 45{dup mul exch dup mul
  add 1. exch sub}setscreen} def
```

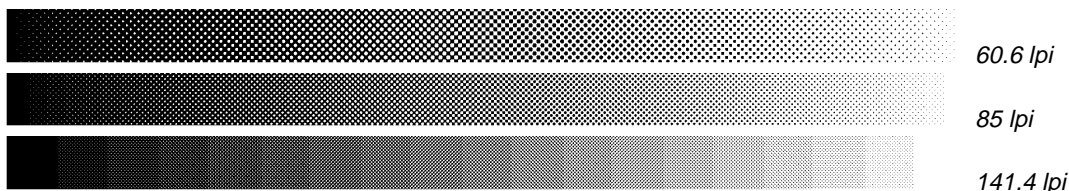
Use this code at your own risk: I have only run a quick test on it and have not used it for production work. Also, Postscript printers and -interpreters may ignore screen settings altogether.



Tuning

Quick-and-easy tonal adjustment: if the tonal range of your picture is e.g. 20–70% black, the contrast will much improve if you stretch this to the full 0–100

Halftone screens



range. The screenshot above displays the Photoshop 3 *image/adjust/levels* dialogue, in which this can be accomplished: move the outer upper two sliders in, ignoring the very thin tail at the right, until they fit snugly around the histogram.

Actually, you might prefer a 5–100 or 5–95 range because patches of pure white or black might look blotchy: slide the lower sliders in to settings of 12 and 243. You can also experiment with the middle of the upper set of sliders to lighten or darken midtones. Make all these adjustments in one visit to this dialog box, in order to minimize quantization errors.

Other programs offer similar facilities although not necessarily as well implemented.

You can also sharpen your scans, but this requires some care, since it easily introduces ugly artefacts.

Printed originals

Scans from printed 'originals' often suffer from fuzziness and moire effects which may be hard to get rid of. Avoid them when you can.

Printing

PostScript and non-PostScript printing each have their own challenges: without PostScript, incorporating graphics is messy at best. With PostScript, it is simple and elegant because you leave the hard work to the printer. If you are careless the printer might choke on the hard work. It helps to restrain yourself with fonts, limit the resolution of images to what is needed and proof your document on a PostScript printer before handing it to someone else.

Tip: if the PostScript 'printer' happens to be GhostScript under Windows, you may get better halftones if you pick the mswinpr2 driver, which lets the default Windows printer driver do the halftoning. You can also try to proof your PostScript with Acrobat Distiller.

References

- A collection of graphics and other file format descriptions can be found at www.wotsit.demon.co.uk.
- For everything related to PostScript, Acrobat and Photoshop, visit Adobe's site: www.adobe.com.
- The Ghostscript site is www.cs.wisc.edu/~ghost.
- A good book on working with scans is *Real World Photoshop* by David Blatner and Bruce Fraser, Peachpit Press, 1996, ISBN 1-56609-169-1.