

External graphics for LaTeX

Siep Kroonenberg
N.S.Kroonenberg at rug dot nl

Abstract

In this article, we discuss graphics file formats, software to create graphics and procedures to convert them to LaTeX- and pdf_latex-compatible formats.

Keywords

Graphics converting bitmap vector compression eps pdf jpeg lossy lossless resolution

This article is about preparing external graphics for use with LaTeX.

We start out with a quick overview of types of graphics. If you understand what kind of data you are dealing with, you will have a much better chance of getting good results.

Next, we list programs for creating graphics, both free and commercial.

The final part is about programs and procedures for converting graphics into LaTeX-compatible formats.

1 Types of graphics

Graphics can be defined in different ways, depending on the type of information they contain and on the software with which they have been created. Figures 1–6 contain some examples, each together with an enlarged detail.



Figure 1. Bitmapped art: a photograph

A bitmap is built up as a grid of pixels. Figures 1 and 2 show a photograph and a screenshot respectively. The grid structure is obvious in the enlarged detail.

Vector graphics are defined in terms of lines, circles, curves and other geometric shapes. They keep their sharpness at any scale; see figures 3–5.

Some file formats can contain both bitmapped and vector data. In figure 6, the bitmapped background becomes fuzzy when enlarged, but the text on top remains sharp.

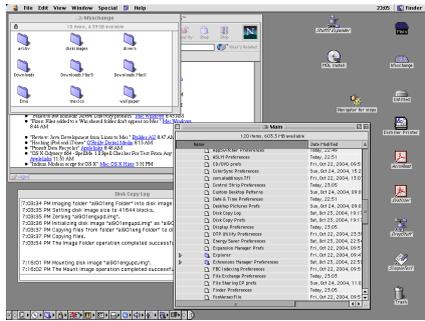


Figure 2. Bitmapped art: screenshot

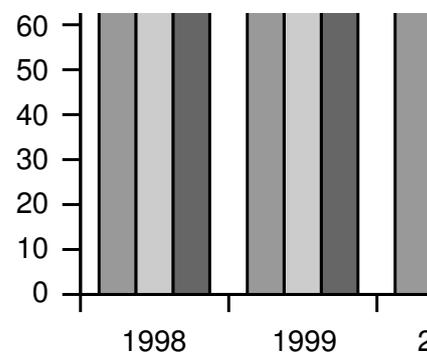
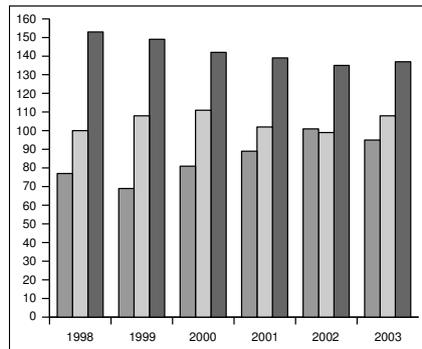


Figure 3. Vector art: OpenOffice.org graph



Figure 4. Vector art: adapted from a Ghostscript example file

1.1 Jpeg compression

High-resolution bitmapped files can get very big. There are various ways to reduce those file sizes.

Lossless compression works by storing information in a more compact way. A very simplified example: instead of enumerating a thousand identical white pixels one by one, you can say at once that the next one thousand pixels are white. Lossless compression can be quite effective when there are large areas of solid colors or regular patterns. Png is a lossless bitmapped format that can be processed directly by pdfflatex.

Lossless compression doesn't work so well with photographic images. When we no longer insist on exactly preserving every bit of information and accept lossy

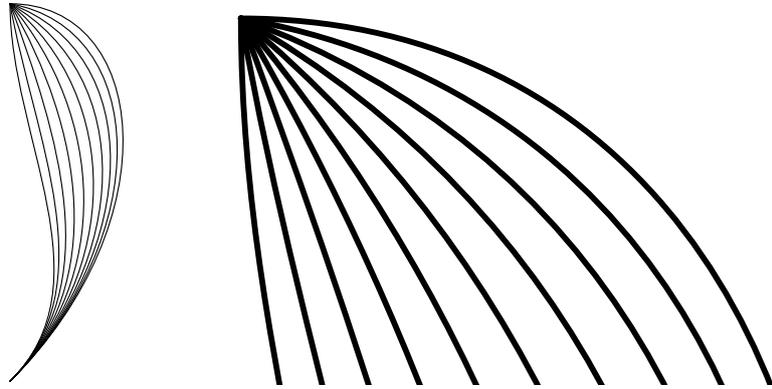


Figure 5. Vector art: generated with MetaPost



Figure 6. Bitmapped and vector combined



Figure 7. Don't use jpeg compression for screenshots.

compression, then very good results can be reached with jpeg. At medium to high quality settings, the loss of information is essentially invisible, but the compression rate is easily 10:1. When you save in jpeg format you can usually make your own tradeoff between file size and image quality.

For non-photographic bitmapped images such as screenshots or logos, jpeg compression produces visible artifacts; see figure 7. To be fair, for the right image quality was set very low in order to make the artifacts more obvious. Here, compression isn't all that good either compared to png. Nevertheless, many people use jpeg compression indiscriminately, even when png would have been much better.¹

1.2 Bitmap resolutions

The resolution of a bitmap should be high enough to look sharp, but, in order to keep file size and loading times within bounds, the resolution should not be higher

than necessary.

For screen viewing, the ideal resolution would be exactly one bitmap pixel per screen pixel, but of course you may not know at what screen resolution and zoom level your document will be viewed.

For printing, good resolutions are:

Photographs. 150–300 DPI (dots per inch) depending on the output device. Most printers and imagesetters simulate grays and other tints with dot patterns or halftone screens. As a consequence, the effective printed resolution of a photograph is much lower than the resolution of the output device, no matter how high the resolution of the original photograph.

Charts and diagrams. 600–1200 DPI. 600 is enough to avoid visible blockiness. Higher resolutions can mean finer detail, if the printer resolution is also high. But vector formats are better for such graphics.

Screenshots. Keep the original resolution.

Of course I am talking about resolution after scaling; if you print a 2" wide, 300 DPI image at a width of 4" then the effective resolution is 150 DPI.

It will do no good to increase the resolution of an existing low-resolution image; it might even make the output fuzzier. So either try to find or create a better original, or use your bad picture as-is.

1.3 Problems with vector graphics

Missing fonts. If some standard fonts (Times, Helvetica, Courier, Symbol, Zapf Dingbat) are not embedded then `epspdfk` (see section 3.2) can help. If the pdf output target is set to prepress then fonts will be embedded during conversion to pdf. `Epspdfk` can convert back and forth between eps and PostScript on the one hand and pdf on the other. You can either apply a pdf-to-eps-to-pdf conversion to individual graphics or a pdf-to-ps-to-pdf conversion to the document as a whole.

If other fonts are missing then you have a real problem.

If you don't mind using Ghostscript from the command-line: the `epswrite` output device replaces characters with little drawings of their shapes: a fool-proof way to get rid of font problems. Don't do this with large amounts of text. I guess it only works if the font is embedded or is known to Ghostscript in some other way.

Zero-width lines. If your graphic has some very thin lines, then check by zooming in whether the lines have some positive width. A line width of zero will be interpreted by the output device as a width of one pixel, which is fine for the screen or for an old 300 DPI laserprinter, but not for a 2400 DPI imagesetter. A line width of e.g. 0.3pt should be safe. You may be able to fix this from within the program with which the graphic was created.

Transparency and fill patterns. Many programs and graphic formats do not support transparency or fill patterns; upon conversion these features might either get lost altogether or simulated with e.g. bitmaps, which might make the file much larger and virtually ineditable. Hang on to the original – which you should do in any case!

General fixes. You may be able to import and fix problems in a draw program: substituting fonts, changing line widths, replacing pattern fills with something else; see section 2 on draw programs.

As a last resort, you can convert your graphic to a bitmapped png file of sufficiently high resolution.

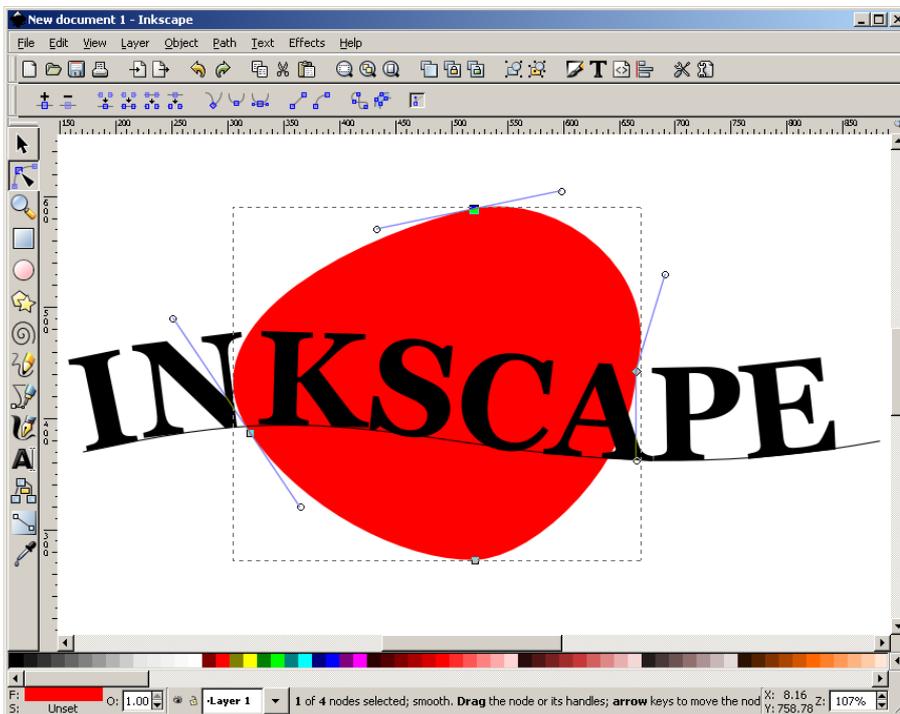


Figure 8. Inkscape, a free vector draw program for Linux, Windows and Mac OS X

2 Creating graphics

In computer graphics, the term drawing means vector art and painting means bitmapped art.

2.1 Drawings and diagrams

You can create drawings and diagrams in Word or PowerPoint. However, it may be difficult to convert graphics from these programs intact to eps or pdf, the formats needed by LaTeX and pdflatex² respectively. Instead, check out some free alternatives such as Inkscape, which is a specialized draw program resembling CorelDRAW, or the Draw module of OpenOffice. OpenOffice.org is fairly good at reading and writing MS Office files, and has a pdf export option.

If you spend a lot of time creating and editing vector graphics, consider buying a professional draw program such as CorelDRAW or Adobe Illustrator.

Under Mac OS X there are several less expensive commercial programs worth checking out, such as OmniGraffle from the Omni Group (www.omnigroup.com), Create from Stone Design (www.stone.com) and Intaglio from Purgatory Design (www.purgatorydesign.com).

In any case, investigate at an early stage how to get your drawings into LaTeX, see section 3.

2.2 Draw programs with LaTeX support

Draw programs which can use TeX for typesetting text include Ipe (cross-platform) and TpX (Windows-only). TpX is also available from CTAN. Both are free.

Section 3.3 has some information on Ipe file formats.

TpX generates LaTeX code, which you have to load into your document with an `\input` command. The various output options of TpX are a bit too complicated to explain here, but you can read all about it in the TpX online help.

2.3 Charts

Charts are normally generated as a byproduct of spreadsheets or mathematical, statistical or econometric software.

2.4 Bitmaps: paint programs and image editors

There exists a large selection of free and inexpensive paint programs and image editors. Paint is a very basic paint program which is included with Windows. IrfanView is a small, simple and free image viewer and converter for Windows. If you need something more substantial, have a look at the GIMP, which was originally developed for Linux. Yet another option is Adobe PhotoShop Elements, which may be easier to work with. It is commercial, but costs a fraction of its professional big brother, PhotoShop, which is the favorite of professional designers. A favorite on the Macintosh is GraphicConverter.

2.5 Screenshots

You can take screenshots without specialized software.

Windows. The PrtScr key will copy the entire screen to the clipboard, and Alt-PrtScr the active window. Most paint- and image-editing programs can retrieve the screenshot from the clipboard, usually with Edit / Paste. They also have tools to crop images.³

Linux. Here, PrtScr will probably take a screenshot as well, either to the clipboard or to a file. I usually take screenshots with the Gimp (File > Acquire > Screenshot).

You can also take a screenshot from the commandline, with `import` from the ImageMagick suite: type `import file.png` and click either a window or the desktop background.

Mac OS X. Cmd-Shift-3 copies the entire screen to a file on your desktop, and Cmd-Shift-4 lets you make a selection.

3 Converting to LaTeX- and pdflatex-compatible formats

I'll only discuss graphics usage for the two most popular output options: generating PostScript with LaTeX and dvips, and generating pdf with pdflatex. The `graphicx` package will automatically detect these two cases. In the LaTeX-plus-dvips case, it will look for graphics in eps format, and in the pdflatex case it will look for graphics in pdf, jpg and png format.⁴

These are stable formats, with little room for ambiguity. Eps and pdf can contain just about any kind of graphics information. This small selection of file formats is therefore not a real limitation, but it can mean extra work.

3.1 Converting bitmaps to png and jpg

Many image editors and paint programs, including IrfanView, can convert to png and jpg. Convert from the ImageMagick suite is a command-line option. ImageMagick also has a GUI viewer and converter: `display` under Unix/X11 and `IMDisplay` under Windows; see Figure 9.

3.2 Converting between PostScript, eps and pdf

Epspdfk and epspdf. These utilities can convert between PostScript, eps and pdf, often with no loss of information. They can also remove unwanted borders (compute tight boundingbox option). `Epspdfk` is a GUI program, and `epspdf` is its command-line counterpart.

epstopdf. An alternative for converting eps to pdf is the command-line program `epstopdf`, which is part of most TeX distributions and is probably already on our system.

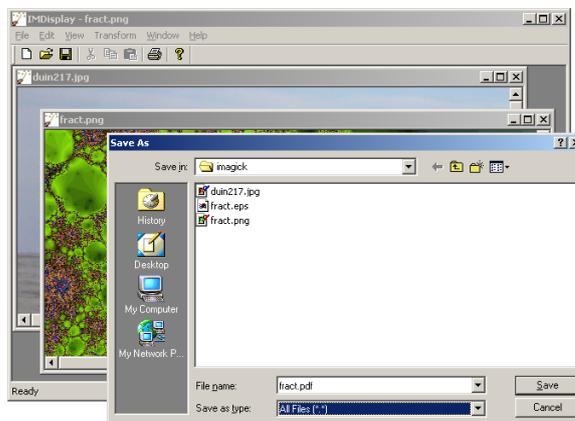


Figure 9. IMDisplay: ImageMagick's GUI component for Windows

convert. On many systems, convert from the ImageMagick suite will use Ghostscript when converting between eps and pdf, and will in that case convert vector graphics to vector graphics. Prefix the target pdf filename with EPDF: if you want to preserve the eps boundingbox. View the result at a high zoom level to make sure that it is still a vector graphic.

Ghostscript and pdftops. You can also use Ghostscript directly for converting eps and PostScript to pdf, and either Ghostscript or pdftops from the xpdf suite for converting in the other direction. But this is not the place for explaining the use of these programs.

3.3 Working with Ipe files

Ipe, see section 2.2, can store drawings in its own private format, with extension ipe or xml, but it can also save them in eps- and pdf format. Ipe drawings in eps- and pdf format have private information hidden inside, which makes it possible for Ipe to read them back in. But if you convert those files to something else then this private information will get lost.

Ipe cannot read arbitrary eps or pdf, but it comes with a command-line import utility pdftoipe. You can put a shortcut to it on your desktop and then simply drop a pdf file on it. The result of the conversion will be in the same directory as the original. You will very likely have to do some cleanup afterwards, especially if there is a lot of text.

3.4 Converting bitmaps to eps

The simpler paint programs and image editors don't convert to eps or pdf, but more advanced ones, such as the GIMP, do. Convert and [IM]Display from the ImageMagick suite also convert to eps and pdf.

You can also convert to eps or pdf with Ipe: create a new, empty document. This happens automatically when you start the program. Select File, Insert image... and load the image that you want to convert. Now save the Ipe document in eps- or pdf format.

3.5 Wmf, emf and the Windows clipboard

Wmf and emf are native vector formats for Windows, and can be read by most Windows graphics programs, including TpX, which I already mentioned in section 2. Another option is Wmf2eps. This shareware program does exactly what its name implies, and its conversions are quite accurate. It uses Windows' native PostScript printerdriver in the background.⁵

If you cannot even save as wmf or emf then again both TpX and Wmf2eps can copy the clipboard content to a file. Note that for TpX you do this with Tools / Capture EMF, not with Paste from the Edit menu.

3.6 Exporting eps and PostScript from Windows programs

If a Windows program doesn't have a usable export option, then you can try to 'print' to a PostScript file. This is approximately what the above wmf2eps program does.⁶

For this, you need to have a PostScript printer driver. If you don't have one installed, go to 'Printers' and start up the Add Printer wizard. Choose Local Printer and uncheck automatic detection. As printer port, you can pick FILE, otherwise you would have to manually check 'Print to File' anytime you print. A good choice for manufacturer and model would be 'Generic' and 'MS Publisher Imagesetter' respectively.

Pay attention to printer settings: in the Print dialog, click 'Properties', then 'Advanced' (on either tab). In the 'Advanced Document Settings' tree, under 'Graphic', 'TrueType Font' should be set to 'Download as Softfont'⁷

Now navigate to first 'Document Options', then 'PostScript Options'. For 'PostScript Output Option' the default setting is 'Optimize for speed'. Change that to 'Optimize for Portability' or 'Archive Format', or, for single pages only, 'Encapsulated PostScript'. These non-default options presumably produce cleaner PostScript code, without printer-specific hacks. Experiment with this and other options if you run into problems (e.g. bad-looking screen output, or part of a graphic getting cut off, or conversion to bitmap).

What works best may depend on your Windows version: under Windows 2000, Archive worked best for me, but I have been warned that this option was unusable in older Windows versions.

Next, the setting 'TrueType Font Downloading Option' should be set to 'Outline', not 'Automatic' or 'Bitmap'.

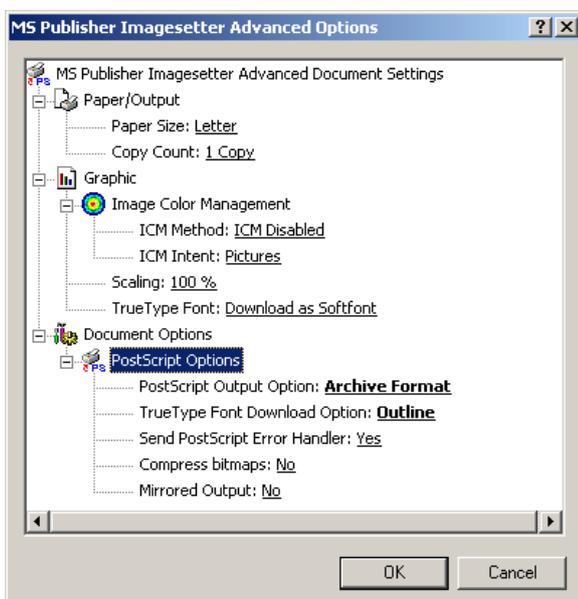


Figure 10. configuring a PostScript printer driver

Notes

1. Jpeg2000 (extension .jp2 or .j2c) is a successor of the jpeg format. This type of compression supports both lossy and lossless compression. Pdflatex can use pdf graphics which internally use jpeg2000 compression, but at the moment you cannot use jpeg2000 images directly.
2. Although pdflatex can also use the bitmapped jpeg and png file formats, we would rather not convert vector graphics to bitmaps.
3. With IrfanView this is not quite obvious, but if you just start to drag the cursor then it changes into a selection tool, and then you can apply Edit / Crop selection.
4. If you specify a non-default output driver as optional parameter to the graphicx package, then different graphics formats may be supported.
5. Most Windows PostScript drivers use the same core. Individual PostScript drivers add to this core a ppd- or PostScript Printer Definition file which is basically an enumeration of the printer's features and properties. Wmf2eps comes with its own ppd.
6. This section is lifted almost literally from my article 'Epspdf, easy conversion between PostScript en Pdf' in Maps 34.
7. The alternative setting is 'Substitute with Device Font'. For prepress use, you should always include all fonts. It is possible to include fonts after the fact, when converting to pdf, but then you run the risk of noticeable discrepancies between the original font and the actually included font.

References

- CTAN, The Comprehensive T_EX Archive Network. <http://www.ctan.org/>.
- Epspdf and epspdfk. <http://tex.aanhet.net/epspdf/>.
- Ghostscript Homepage. <http://www.cs.wisc.edu/~ghost/>.
- ImageMagick Homepage. <http://imagemagick.org/>.
- Inkscape draw program. <http://inkscape.org/>.
- Ipe draw program. <http://tclab.kaist.ac.kr/ipe/>.
- IrfanView image viewer and converter for Windows. <http://www.irfanview.com/>.
- OpenOffice.org homepage. <http://www.openoffice.org/>.
- The Gimp image editor for Windows. <http://gimp-win.sourceforge.net/>.
- TpX draw program. <http://sourceforge.net/projects/tpx/>.
- Wmf2eps Homepage. <http://www.wmf2eps.de.vu/>, wmf2eps can be downloaded from CTAN.
- Xpdf Homepage. <http://www.foolabs.com/xpdf/>.

Siep Kroonenberg
N.S.Kroonenberg at rug dot nl