



Inserting figures into T_EX documents

Szabó Péter
<pts@inf.bme.hu>

April 2002
presented at EuroBachoT_EX 2002

Motivation
Sampled image model
Transparency
The image as a file
Compression in . . .
sam2p
Two sample runs
GUI screen shot
Implementation
Further work
Conclusion



Page 1 of 12

Home Page

Go Back

Full Screen

Close Quit



✧ Motivation ✧

The problem: Sampled images (such as scanned pictures and screen shots), especially those with high resolution, occupy much space on disk, and thus PostScript documents containing such images are printed extremely slowly, and PDF files with such images require much time to download.

Solution: create and embed images in compressed form adequate for the image type.

New problems:

- 🌀 the printer may not support the compression used
- 🌀 (de)compression may take significant time
- 🌀 currently no sophisticated converters exist
- 🌀 software patents

Motivation

Sampled image model

Transparency

The image as a file

Compression in . . .

sam2p

Two sample runs

GUI screen shot

Implementation

Further work

Conclusion



Page 2 of 12

Home Page

Go Back

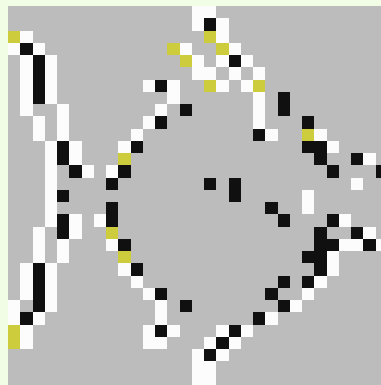
Full Screen

Close

Quit



✧ Sampled image model ✧



- ☞ a *sampled image* is a rectangular array of *pixel* samples
- ☞ the color of each pixel is the member of a *color space*
- ☞ each color has one or more *components* (such as red, green and blue), depending on a color space

sam2p supports:

- ☞ the *8-bit Gray* color space (0 is black, 255 is white)
- ☞ the *8-bit RGB* color space
- ☞ *indexed* (paletted) images up to 256 color in 8-bit RGB

Motivation

Sampled image model

Transparency

The image as a file

Compression in . . .

sam2p

Two sample runs

GUI screen shot

Implementation

Further work

Conclusion



Page 3 of 12

Home Page

Go Back

Full Screen

Close

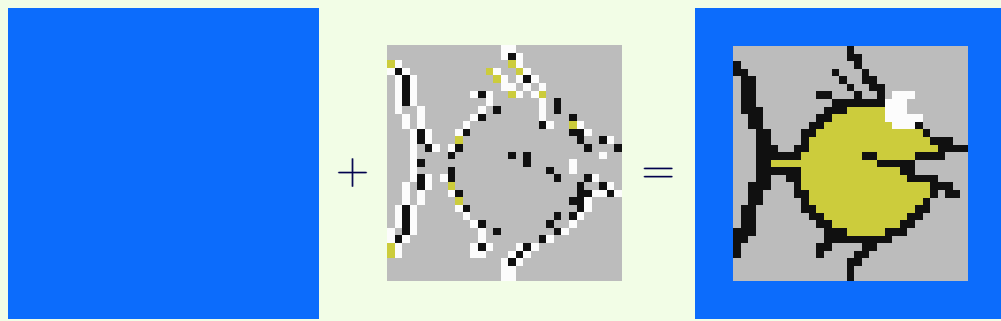
Quit



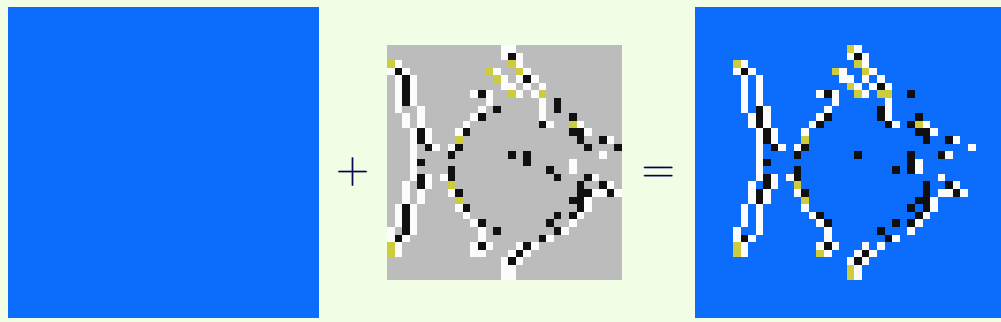
sam2p

✧ Transparency ✧

Image composition with *opaque* colors:



Making the gray background *transparent*:



In some image formats it is possible to specify multiple (≥ 3) levels of transparency (this feature is called *alpha channel*). sam2p supports only simple (1-bit) transparency, and only for Indexed SampleFormats.

- Motivation
- Sampled image model
- Transparency**
- The image as a file
- Compression in ...
- sam2p
- Two sample runs
- GUI screen shot
- Implementation
- Further work
- Conclusion



Page 4 of 12

Home Page

Go Back

Full Screen

Close Quit



✧ The image as a file ✧

- ✧ *FileFormat*: format of the output file. Examples: PostScript Level1 ... 3, PDF, PNM, GIF, PNG etc.
- ✧ *SampleFormat*: color space and bit depth specification. Examples: Indexed1, Transparent2, Rgb4, Gray8.
- ✧ *TransferEncoding*: ASCII, Binary or binary in ASCII hexadecimal or ASCII85 encoding. EPS files should not be Binary, PDF files should.
- ✧ *Compression*. None, LZW (patented till 19 June 2003), ZIP, JPEG, Fax or RLE.
- ✧ *Predictor*: additional transformation to increase compression ratio of LZW or ZIP
- ✧ *Hints*: hints and parameters of the above
- ✧ other meta-information such as the name of the photographer

Motivation
Sampled image model
Transparency
The image as a file
Compression in ...
sam2p
Two sample runs
GUI screen shot
Implementation
Further work
Conclusion



Page 5 of 12

Home Page

Go Back

Full Screen

Close

Quit



✂ Compression in PDF and PS ✂

PSL1 provides no decompression, but it is possible to implement decompression filters in PostScript. The author has implemented ZIP in a few weeks of hard work. RLE should be fairly easy. PDF and PSL2 have a the following predefined decompression filters:

- ✂ *RLE*: encoding the length of runs of consecutive same bytes. Simple, but effective only for a very limited class of images.
- ✂ *Fax*: modified Huffman encoding used in fax sheet transfers. About twice as effective as RLE, but only for 1-bit images.
- ✂ *LZW*: generic compressor that recognizes repeating strings. Fairly effective, but patented till 19 June 2003.
- ✂ *ZIP*: generic compressor that recognizes repeating strings better than LZW. Effective, but supported only in PSL3 and PDF1.2, thus most PostScript printers don't support it.
- ✂ *JPEG*: lossy compression, uses DCT and frequency-domain methods. Super for screen-resolution (70–100 DPI), continuous-tone images such as photographs.

Motivation
Sampled image model
Transparency
The image as a file
Compression in . . .
sam2p
Two sample runs
GUI screen shot
Implementation
Further work
Conclusion



Page 6 of 12

Home Page

Go Back

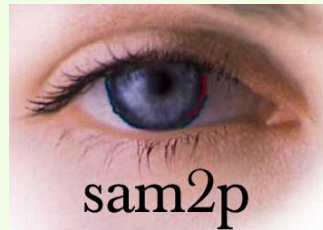
Full Screen

Close

Quit

✂ sam2p ✂

sam2p is a UNIX command line utility written in ANSI C++ that converts many raster (bitmap) image formats into Adobe PostScript or PDF files. The images are not vectorized. sam2p gives full control to the user to specify standards-compliance, compression, and bit depths. sam2p provides some Level3 compression filters even on Level1 devices.



<http://www.inf.bme.hu/~pts/sam2p/>

The operation of sam2p:

1. recognizes and loads the specified sampled image
2. converts the image to the specified SampleFormat
3. applies Compression and writes the OutputFile
4. for rules not matching the image, proceeds with the next rule



Motivation

Sampled image model

Transparency

The image as a file

Compression in ...

sam2p

Two sample runs

GUI screen shot

Implementation

Further work

Conclusion



Page 7 of 12

Home Page

Go Back

Full Screen

Close

Quit

✂ Two sample runs ✂



```
~$ ./sam2p examples/ptsbanner_tralzw.job
This is sam2p v0.34.
Available Loaders: JAI PNG JPEG TIFF PNM BMP GIF LBM XPM PCX TGA.
Available Appliers: Meta Empty PNG TIFF6 JPEG-JAI JPEG PNM GIF89a PSL1-Indexed1-
Hex PSL1-Indexed1 PSL1-Mask-Hex PSL1-Indexed1-Hex PSL1-Indexed1 PSL1-Mask PS+PDF
PDF-JAI l2jbin l1fa85g P-Tr P-Op.
./sam2p: Notice: job: read InputFile: examples/ptsbanner.gif
./sam2p: Warning: SampleFormat: Transparent8 separates too many colors
./sam2p: Warning: SampleFormat: Indexed8 would be better than Transparent8
./sam2p: Notice: job: written OutputFile: test.eps
Success.
~$ _
```

```
~$ ./sam2p examples/fusi_tiff_bad.pdf.job
This is sam2p v0.34.
Available Loaders: JAI PNG JPEG TIFF PNM BMP GIF LBM XPM PCX TGA.
Available Appliers: Meta Empty PNG TIFF6 JPEG-JAI JPEG PNM GIF89a ...
./sam2p: Warning: buildProfile: don't know how, ignoring OutputRule #0
./sam2p: Warning: buildProfile: all OutputRules are useless
notice: 3 colorimetric sample(s), EXTRASAMPLE_ASSOCALPHA
notice: writing PPM file
notice: writing PBM alpha
./sam2p: Notice: PNM: loading alpha after PNM
./sam2p: Notice: job: read InputFile: examples/fusi.tiff
./sam2p: Error: applyProfile: no applicable OutputRule
~$ _
```

Motivation

Sampled image model

Transparency

The image as a file

Compression in ...

sam2p

Two sample runs

GUI screen shot

Implementation

Further work

Conclusion



Page 8 of 12

Home Page

Go Back

Full Screen

Close

Quit

✧ GUI screen shot ✧



sam2p Job Editor

FileFormat	SampleFormat	Compression
<input type="checkbox"/> PS L1	<input type="checkbox"/> Opaque	<input type="checkbox"/> None
<input type="checkbox"/> PS LC	<input type="checkbox"/> Transparent	<input type="checkbox"/> LZW
<input type="checkbox"/> PS L2	<input type="checkbox"/> Gray 1	<input checked="" type="checkbox"/> ZIP Effort <input type="text" value="5"/>
<input type="checkbox"/> PS L3	<input type="checkbox"/> Indexed 1	<input type="checkbox"/> RLE R.S. <input type="text" value=""/>
<input type="checkbox"/> PDF B 1.0	<input type="checkbox"/> Mask	<input type="checkbox"/> Fax K <input type="text" value=""/>
<input checked="" type="checkbox"/> PDF B 1.2	<input type="checkbox"/> Transparent 2	<input type="checkbox"/> DCT
<input type="checkbox"/> PDF 1.0	<input type="checkbox"/> Gray 2	<input type="checkbox"/> JG Q'ity <input type="text" value=""/>
<input type="checkbox"/> PDF 1.2	<input type="checkbox"/> Indexed 2	<input type="checkbox"/> JAI
<input type="checkbox"/> GIF 89a	<input type="checkbox"/> Transparent 4	Predictor <input type="text" value="15"/>
<input type="checkbox"/> Empty	<input type="checkbox"/> RGB 1	<input checked="" type="checkbox"/> WarningOK
<input type="checkbox"/> Meta	<input type="checkbox"/> Gray 4	TransferEncoding
<input type="checkbox"/> PNM	<input type="checkbox"/> Indexed 4	<input type="checkbox"/> Binary
<input type="checkbox"/> PAM	<input type="checkbox"/> Transparent 8	<input type="checkbox"/> ASCII
<input type="checkbox"/> PIP	<input type="checkbox"/> RGB 2	<input type="checkbox"/> Hex
<input type="checkbox"/> JPEG	<input type="checkbox"/> Gray 8	<input checked="" type="checkbox"/> A85
<input type="checkbox"/> TIFF	<input checked="" type="checkbox"/> Indexed 8	TransferCPL <input type="text" value="65"/>
<input type="checkbox"/> PNG	<input type="checkbox"/> RGB 4	<input checked="" type="checkbox"/> Tmp Remove
	<input type="checkbox"/> RGB 8	

```

/InputFile (ScreenShot.png)
/OutputFile (ScreenShot.eps)
/TmpRemove true /LoadHints pop /Templates pop
/Profile [ <<
  /FileFormat /PDFB1.2
  /SampleFormat /Indexed8
  /WarningOK true
  /TransferEncoding /A85
  /Compression /ZIP
  /Predictor 15
  /Transparent null
  /Hints <<
    /Effort 5
    /RecordSize 0
    /K 0
    /Quality 75
    /ColorTransform pop
    /TransferCPL 65
  >>
]
  
```

Debug messages, sam2p output:

```

This is sam2p v0.34.
Available Loaders: JAI PNG JPEG TIFF PNM BMP GIF LBM XP
M PCX TGA.
Available Appliers: Meta Empty PNG TIFF6 JPEG-JAI JPEG
PNM GIF89a PSL1-Indexed1-Hex PSL1-Indexed1 PSL1-Mask-He
x PSL1-Indexed1-Hex PSL1-Indexed1 PSL1-Mask PS+PDF PDF-
JAI 12jbin 11fa85g P-Tr P-Op.
Usage: sam2p <filename.job>
  
```

Current dir /home/guests/pts/prg/pshack./jpeg2pdf/sam2p-0.34

JobFile ScreenShot.job ++ Load Job Run

InputFile ScreenShot.png !e Save Job Quit

OutputFile ScreenShot.eps ++

- Motivation
- Sampled image model
- Transparency
- The image as a file
- Compression in ...
- sam2p
- Two sample runs
- GUI screen shot**
- Implementation
- Further work
- Conclusion



Page 9 of 12

Home Page

Go Back

Full Screen

Close Quit



✧ Implementation ✧

- ⌚ ANSI C++ for speed
- ⌚ only a minimal set of libraries for portability
- ⌚ works on UNIX, but can be ported to other platforms
- ⌚ invokes external programs for reading and writing some image files (e.g. JPEG), and for some compression (e.g. Fax)
- ⌚ modular source code makes adding new input and output formats easy
- ⌚ new color spaces would be hard to add
- ⌚ sparing memory handling
- ⌚ it would be hard to convert it to a library, because there are global variables, and errors are fatal (the process exits)

Motivation
Sampled image model
Transparency
The image as a file
Compression in . . .
sam2p
Two sample runs
GUI screen shot
Implementation
Further work
Conclusion



Page 10 of 12

Home Page

Go Back

Full Screen

Close

Quit



✂ Further work ✂

🌀 *image model*

- ★ add real alpha channel, output it in PSL3 and PDF1.3
- ★ load and save images with bit depth of 12 and 16 (PDF supports 8 bits, PostScript supports 12 bits, TIFF supports 16 bits)
- ★ add more color spaces (DeviceCMYK and CIE-based)
- ★ indexed images should have > 256 colors

🌀 *features*

- ★ add operations: crop, draw border, rotate, translate, scale
- ★ add color transformations: quantize, brightness–contrast, to grayscale
- ★ respect input gamma correction and transfer functions

🌀 *file formats and filters*

- ★ provide alternative implementations of compression filters
- ★ add TIFF output, add more EPS and PDF output
- ★ support some more image formats for both load and save

Motivation
Sampled image model
Transparency
The image as a file
Compression in . . .
sam2p
Two sample runs
GUI screen shot
Implementation
Further work
Conclusion



Page 11 of 12

Home Page

Go Back

Full Screen

Close

Quit



✧ Conclusion ✧

- ☞ sam2p provides all compression filters, and it provides some filters even on PSL1
- ☞ it supports a wide range of SampleFormats
- ☞ it can load most widely used image formats
- ☞ sam2p seems to work fine, but it has not been tested exhaustively
- ☞ the image model could be extended on demand
- ☞ new EPS and PDF writers can be added to support the rare cases
- ☞ it would be hard to convert sam2p to a library
- ☞ sam2p is already the best and most general sampled PDF and EPS generator available that the author knows of

Motivation

Sampled image model

Transparency

The image as a file

Compression in ...

sam2p

Two sample runs

GUI screen shot

Implementation

Further work

Conclusion



Page 12 of 12

Home Page

Go Back

Full Screen

Close

Quit