

NAME

djvudigital - creates DjVu files from PS or PDF files.

SYNOPSIS

djvudigital [*options*] *inputfile* [*outputfile*]

DESCRIPTION

This program creates a DjVu file from the PostScript (**.ps**), GZipped PostScript (**.ps.gz**), Encapsulated PostScript (**.eps**), or Portable Document Format (**.pdf**) file *inputfile*.

The output file name is either given by argument *outputfile* or generated by replacing the input file name suffixes by the DjVu suffix (**.djvu**).

PREREQUISITES

This program depends on a specific GhostScript driver. If your GhostScript program does not provide this driver, please check <http://djvu.sourceforge.net/gsdjvu.html>.

OPTIONS**--verbose, -v**

Displays more informational messages while converting the file.

--quiet, -q

Do not display informational messages while converting the file.

--dpi=*resolution*

Specify the desired resolution to *resolution* dots per inch. The default is 300 dpi.

--psrotate=*angle*

Rotate the PostScript file by *angle* degrees clockwise. Only the values **0**, **90**, **180**, and **270** are supported. This option only applies to PostScript files. PDF files are always converted according to their native orientation.

--epsf=*disposition*

Specify how to handle Encapsulated PostScript files. Argument *disposition* can take the values **crop**, **fit**, and **ignore**. The default disposition **crop** creates a DjVu file whose size matches the bounding box of the Encapsulated PostScript file. Value **fit** rescales the graphics to the default

page size. Value **ignore** disables all Encapsulated PostScript specific code. This option requires Ghostscript 7.07 or better.

--exact-color

Enables a more accurate rendering of the colors. This option requires GhostScript 6.52 or better.

--threshold=*thres*

Specify a threshold for the foreground/background separation code. Acceptable values of *thres* range from 0 to 100. Larger values place more information into the foreground layer. The default threshold value is 80.

--bg-subsample=*sub*

Specify the background subsampling ratio. Argument *sub* must be an integer between 1 and 6. The default value is 3.

--bg-slices=*n+...+n*

Specify the encoding quality of the background layer. The syntax for the argument is similar to that described for the **-slice** option of command **c44**. The default is **72+11+10+10**.

--fg-colors=*ncolors*

Specify the maximum number of distinct colors in the foreground layer. Argument *ncolors* can take integer values between 1 and 4000. The default value is 256.

--fg-image-colors=*ncolors*

Specify the maximum number of distinct colors in an image for considering encoding it into the foreground layer. Argument *ncolors* can take integer values between 1 and 4000. The default value is 256.

--words

Extract the text from the PostScript code and incorporates this information into the DjVu file. This option records the location of every word.

--lines

Extract the text from the PostScript code and incorporates this information into the DjVu file. This option saves a few bytes by only recording the location of each line.

--gsarg=*arg1[,arg2,...,argN]*

Insert extra arguments on the GhostScript command line.

--cseparg=*arg1[,arg2,...,argN]*

Insert extra arguments on the command line of program **csepdjvu** or **msepdjvu**.

--poppler=keywords

This option causes **djvudigital** to extract additional information from PDF files using the tool **pdftotext** that comes bundled with the Poppler library. Selected information is then added to the djvu file as a postprocessing step. This option is ignored when the input file is not a PDF file. Argument *keywords* is a comma separated list of keywords. When this list contains keyword **meta**, the metadata extracted by **pdftotext** is inserted into the djvu file. When this list contains keyword **text**, the textual information extracted by **pdftotext** is inserted into the djvu file, possibly replacing the information gathered using the options **--words** or **--lines**. This is useful for instance when a scanned PDF file contains a hidden text layer that is not recognized by Ghostscript and therefore not passed to the djvudigital backend.

--sepfile

Produces a separated data file instead of a DjVu file. Program **csepdjvu** can then convert the separated data file into a DjVu file.

--check

Display the names of the two auxiliary programs found by **djvudigital**, namely a suitable ghostscript interpreter and a suitable backend encoder. See the next two section for details.

--dryrun

Simply display the ghostscript command line generated by **djvudigital** without running it. No output file is produced

--help

Display the manual page for **djvudigital**.

GHOSTSCRIPT ISSUES

Program **djvudigital** internally relies on a specific Ghostscript driver named **djvusep**. This driver analyzes the logical structure of the sequence of PostScript rendering commands and decides to execute each command into either the foreground or the background layer. The GhostScript driver produces a separated data file that is then compressed using the DjVuLibre program **csepdjvu**.

Before processing the input file, program **djvudigital** searches a Ghostscript executable providing the **djvusep** driver. The search starts with the file specified by the environment variable **GSDJVU** and continues with command line executables named **gs** and **gsdjvu**.

The DjVuLibre source code contains instruction to compile such a GhostScript executable. More information can be obtained from <http://djvu.sourceforge.net/gsdjvu.html>.

CSEPDJVU ISSUES

The output of the **djvusep** GhostScript driver must be processed by the DjVuLibre program **csepdjvu**. This program can also be replaced by the the proprietary Lizardtech program **msepdjvu**. Before processing the input file, program **djvudigital** searches such an executable. The search starts with the file specified by the environment variable **CSEPDJVU** and continues with command line executables named **msepdjvu** and **csepdjvu**.

OTHER PROGRAMS

The option **--poppler=keywords** relies on the tool **pdftotext** that comes with the Poppler library and the tool **djvused** that comes with djvulibre. Only recent versions of **pdftotext** that accept the option **-bbox** are supported. Both tools are searched by first trying the files specified by the environment variables **PDFTOTEXT** and **DJVUSED**, and then trying executables named **pdftotext** or **djvused** found along the shell executable path.

CREDITS

The first version of this converter was written by L on Bottou <leonb@users.sourceforge.net> in AT&T Labs. The DjVuLibre version is derived from code graciously released by Lizardtech in January 2004.

BUGS

Program **djvudigital** can only process input files that GhostScript can process properly.

SEE ALSO

djvu(1), **csepdjvu(1)**, **c44(1)**, **gs(1)**, **gzip(1)**