

DJVUSED

Руководство пользователя

Оглавление

ОГЛАВЛЕНИЕ	1
1. О ПРОГРАММЕ	2
2. ЗАПУСК DJVUSED	2
2.1 Ключи запуска DJVUSED	3
3. ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ DJVUSED.....	4
3.1 ПОЛУЧЕНИЕ РАЗМЕРА ВЫБРАННОЙ СТРАНИЦЫ	4
3.2 ИЗВЛЕЧЕНИЕ ТЕКСТОВОГО СЛОЯ ВСЕГО ДОКУМЕНТА ИЛИ ОДНОЙ ВЫБРАННОЙ СТРАНИЦЫ	4
3.3 ИЗВЛЕЧЕНИЕ АННОТАЦИЙ ОДНОЙ ВЫБРАННОЙ СТРАНИЦЫ.....	4
3.4 ИЗВЛЕЧЕНИЕ/ВНЕДРЕНИЕ АННОТАЦИЙ И ТЕКСТА ВСЕГО DJVU-ДОКУМЕНТА	4
3.5 ИЗВЛЕЧЕНИЕ/ВНЕДРЕНИЕ ОГЛАВЛЕНИЯ DJVU-ДОКУМЕНТА	5
3.6 ИЗВЛЕЧЕНИЕ СТРАНИЦЫ	5
3.7 СОЗДАНИЕ ВНЕДРЕННЫХ В ДОКУМЕНТ МИНИАТЮР	5
4. КОМАНДЫ DJVUSED	6
4.1 КОМАНДЫ ПРОСМОТРА ИНФОРМАЦИИ О ДОКУМЕНТЕ	6
4.2 КОМАНДЫ ВЫБОРА ЧАНКА И ПРОСМОТРА ВЫБРАННОГО	7
4.3 КОМАНДЫ РАБОТЫ С РАЗДЕЛЯЕМЫМИ АННОТАЦИЯМИ	7
4.4 КОМАНДЫ ДЛЯ ТЕКСТА И АННОТАЦИЙ	7
4.5 КОМАНДЫ ДЛЯ РАБОТЫ С ОГЛАВЛЕНИЕМ	10
4.6 КОМАНДЫ ДЛЯ РАБОТЫ С МИНИАТЮРАМИ	10
4.7 КОМАНДЫ СОХРАНЕНИЯ	11
4.8 ПОМОЩЬ И ИМЯ СТРАНИЦЫ	11
5. ФОРМАТ СКРИПТОВ DJVUSED	12
5.1 СИНТАКСИС СКРИПТОВ DJVUSED	12
5.2 СИНТАКСИС ТЕКСТОВОГО СЛОЯ.....	13
5.3 СИНТАКСИС ОГЛАВЛЕНИЯ.....	14
5.4 СИНТАКСИС АННОТАЦИЙ.....	14
6. ОГРАНИЧЕНИЯ DJVUSED.....	16
7. АВТОРЫ DJVUSED	16
8. ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ	16

1. О программе

Djvused - многоцелевой редактор [DjVu](#). Программа djvused мощная утилита командной строки для управления многостраничными документами, создания или редактирования чанков аннотаций, создания или редактирования скрытых текстовых слоев, подготовки миниатюр и т.д.

Входит в состав [DjVuLibre](#). Англоязычная страница, по которой написано это руководство - [здесь](#).

2. Запуск DJVUSED

Для использования djvused необходимо скачать и установить [дистрибутив](#) DjvuLibre или DjVuLibre+DjView. Djvused.exe будет в папке установки (по умолчанию для Windows - Program Files\DjVuZone\DjVuLibre\).

Поскольку утилита капризно ведет себя в ОС Windows рекомендую для работы с конкретным файлом:

- Создать новый каталог.
- Скопировать туда обрабатываемый файл djvu.
- Скопировать в этот каталог djvused.exe и все lib*.dll msvc*.dll из каталога Program Files\DjVuZone\DjVuLibre\.
- Запустить консоль cmd.exe и сделать текущим созданный каталог.
- Далее можно исполнять команды djvused.exe.

Пример bat-файла djvused.bat:

```
REM Удобная работа с DJVUSED
```

```
REM Выбираем диск - в примере это C и создаем каталог test
```

```
C:
```

```
Md c:\test
```

```
Cd c:\test
```

```
Copy "%ProgramFiles%\DjVuZone\DjVuLibre\djvused.exe" *.*
```

```
Copy "%ProgramFiles%\DjVuZone\DjVuLibre\lib*.dll" *.*
```

```
Copy "%ProgramFiles%\DjVuZone\DjVuLibre\msvc*.dll" *.*
```

```
REM Конец bat-файла
```

Запуск djvused на редактирование файла (файл - обязательный параметр):

djvused [ключи] <djvufile>

Программа считывает документ DjVu djvufile и выполняет набор команд djvused.

Команды djvused могут быть считаны из:

- файла (при запуске с ключом -f),
- командной строки (при запуске с ключом -e),
- стандартного ввода (stdin по умолчанию).

Вывод в файл возможен (см. примеры по тексту), однако есть одна [неприятность](#) - если файл существует, то программа без каких либо сообщений ничего в него не записывает.

2.1 Ключи запуска djpeg

[ключи] могут иметь следующее значение:

-v

Указывает djpeg вывести подсказку командной строки перед считыванием команд и краткое описание действия команд. Опция полезна для отладки скриптов djpeg и для интерактивного ввода команд из стандартного устройства ввода.

-f <scriptfile>

Указывает djpeg считать команды из файла **scriptfile** ([скрипт djvused](#)).

-e <commands>

Указывает djpeg исполнить команды, указанные в **commands**. Рекомендуется команды djpeg вводить в одинарных кавычках (актуально для ОС Linux, в ОС Windows рекомендую вводить в обычных двойных кавычках), иначе оболочка командной строки может их воспринять как собственные команды.

-s

Указывает djpeg сохранить файл **djvufile** после исполнения команд. Аналогично команде **save** после завершения работы с программой.

-u

Указывает djpeg печатать скрытый текстовый слой и аннотации в кодировке UTF-8 вместо кодирования не-ASCII символов **восьмеричной escape-последовательностью**. Эта опция необходима для ручного редактирования или просмотра результатов команд djpeg (**вводить до ключа -e: -u-e**). В ОС Windows вывод будет в кодировке **UTF-8 BOM**.

-n

Указывает djpeg игнорировать команды сохранения файла. Полезно для отладки [скриптов djpeg](#) без внесения изменений в документ DjVu.

3. Примеры использования DJVUSED

Возможно много способов применения `djvused`. Приведем наиболее часто используемые.

3.1 Получение размера выбранной страницы

Команда **`size`** выводит ширину и высоту указанной страницы в синтаксисе HTML. Например, команда вывода размера страницы 3 документа `myfile.djvu`.

```
djvused myfile.djvu -e "select 3; size"
```

3.2 Извлечение текстового слоя всего документа или одной выбранной страницы

Команда **`print-pure-txt`** выводит текст, ассоциированный со страницей или документом. Например, следующая команда выводит весь текст документа. Строки и страницы разделены обычными управляющими символами.

Пример извлечения всего текстового слоя документа в текстовый файл:

```
djvused myfile.djvu -u -e "print-pure-txt" > myfile.txt
```

Команда **`print-txt`** производит более продвинутый вывод текста с описанием структуры и местоположения текста на странице. Синтаксис вывода описан в [разделе 5.2](#). Например, следующая команда выводит текст в расширенном виде для страницы 3 документа `myfile.djvu` (обратите внимание, что ключ вывода в Unicode **`-u`** введен **ДО** ключа **`-e`**):

```
djvused myfile.djvu -u -e "select 3; print-txt" > myfile.txt
```

3.3 Извлечение аннотаций одной выбранной страницы

Аннотации **одной** страницы (страница должна быть выбрана командой «`select`») можно извлечь командой **`print-ant`**. Синтаксис аннотации описан в [разделе 5.4](#). Например, следующая команда выводит аннотацию первой страницы документа `myfile.djvu`.

```
djvused myfile.djvu -u -e "select 1; print-ant" > myfile.txt
```

Команда `print-ant` выводит аннотации, сохраненные в выбранном файле документа (компонентном файле или чанке).

Команда **`print-merged-ant`** выводит аннотации, ассоциированные с текущей страницей, но расположенные во всех файлах (с применением чанка INCL).

3.4 Извлечение/внедрение аннотаций и текста всего djvu-документа

Три команды, **`output-txt`**, **`output-ant`** и **`output-all`** извлекают текст, аннотации и текст+аннотации, соответственно, в [djvused-скрипты](#) `dsed`, из которых их можно внедрить в файл обратно (при переносе текста и/или аннотаций в другой документ - см. [ограничения скрипта](#)). Например, следующая команда производит файл `myfile.dsed`, который содержит все аннотации и текстовый слой документа `myfile.djvu`:

```
djvused myfile.djvu -u -e "output-all" > myfile.dsed
```

Файл `myfile.dsed` является текстовым файлом (с управляющими командами - см. [скрипты](#)), который очень просто отредактировать и внедрить обратно в документ.

Например, следующая команда внедряет все аннотации и текстовый слой из скрипта `myfile.dsed` в документ `myfile.djvu`:

```
djvused myfile.djvu -f myfile.dsed -s
```

Внимание! Обнаружена особенность `djvused` - при извлечении текстового слоя он удаляет пробелы в конце слов (`word`), при вставке этих слов пробелы вставляются опять. Если пробелов несколько, то удаляется только последний. Это приводит к проблемам, если пробелы заданы явно (как это делает ABBYY Finereader 11):

Word (`xxx yyy www hhh "Word"`) - после такого элемента вставляется пробел.

Char (`xxx yyy www hhh " "`) - такой пробел `djvused` сохраняет.

Дополнительный пробел испортит такой текст. См. [решение проблемы](#).

3.5 Извлечение/внедрение оглавления djvu-документа

Например, следующая команда производит файл `myfile.txt`, который содержит оглавление документа `myfile.djvu`:

```
djvused myfile.djvu -u -e "print-outline" > myfile.txt
```

Файл `myfile.txt` является текстовым файлом (см. [скрипты](#)), который очень просто отредактировать и внедрить обратно в документ.

Например, следующая команда внедряет оглавление из файла `myfile.txt` в документ `myfile.djvu`:

```
djvused myfile.djvu -e "set-outline myfile.txt" -s
```

3.6 Извлечение страницы

Команды [save-page](#) и [save-page-with](#) создают документ DjVu, содержащий выбранный файл документа (компонентный файл). Например, следующая команда создает документ `p05.djvu` содержащий страницу 5 документа `myfile.djvu`.

```
djvused myfile.djvu -e "select 5; save-page p05.djvu"
```

Каждая страница документа может содержать разделяемую информацию из других компонентных файлов (при помощи чанка INCL). Команда **`save-page`** производит документ с неразрешенными ссылками на разделяемую информацию. Такой документ не может использоваться отдельно и его необходимо потом включить в многостраничный документ, содержащий **необходимую разделяемую информацию** в других компонентных файлах (чанках).

Команда **`save-page-with`** создает одностраничный документ, содержащий всю связанную разделяемую информацию. Полученный в результате выполнения команды документ можно использовать отдельно. Последующая сборка таких одностраничных документов обратно в многостраничный документ вызовет **ненужное дублирование разделяемой информации** (все общие данные будут продублированы для каждой страницы).

3.7 Создание внедренных в документ миниатюр

Команда [set-thumbnails](#) создает миниатюры, которые потом могут быть просмотрены в просмотрщике DjVu-документов. Например, следующая

команда создает миниатюры размера 64x64 пикс. Для всех страниц документа myfile.djvu.

djvused myfile.djvu -e "set-thumbnails 64" -s

4. Команды DJVUSED

Командные строки могут содержать или не содержать команды `djvused` и могут содержать комментарий. Если команд несколько, то они должны быть разделены точкой с запятой «;». Комментарии начинаются с символа «#» и весь текст после него до конца командной строки считается комментарием.

4.1 Команды просмотра информации о документе

Многостраничные документы DjVu состоят из набора компонентных файлов (чанков). Большинство чанков описывают информацию для какой-то отдельной страницы. Некоторые чанки содержат информацию, которая одинакова для нескольких страниц, такая как общая информация об изображении, общие аннотации или миниатюры.

n	Вывод количества страниц в документе.
ls	Вывод всех чанков документа. Каждая линия содержит необязательный номер страницы, букву, описывающую тип чанка, размер чанка и идентификатор чанка. Буквы типа чанка P, I, A, и T обозначают, соответственно, данные о странице, разделяемую информацию об изображении, аннотацию и данные миниатюр. Номера страниц выводятся только для чанков, содержащих информацию о странице. Когда задано, выводится необязательный заголовок страницы (см. команду <i>set-page-title</i> в разделе 5.4) выводится после идентификатора чанка.
dump	Выводит EA IFF 85 структуру документа или выбранного чанка. Такие же возможности имеет утилита djvudump .
size	Выводит ширину (width) и высоту (height) изображения выбранной страницы (перед выполнением команды необходимо выбрать страницу командой «select»). Размеры страниц выводятся с синтаксисом, позволяющим использовать эти данные в тэге <code><EMBED...></EMBED></code> .

4.2 Команды выбора чанка и просмотра выбранного

Многостраничные документы DjVu состоят из набора компонентных файлов (чанков). Большинство чанков описывают информацию для какой-то отдельной страницы. Некоторые чанки содержат информацию, которая одинакова для нескольких страниц, такая как общая информация об изображении, общие аннотации или миниатюры. **Часть команд `djvused` работают только с выбранным чанком. По умолчанию выбраны все чанки.** Следующие команды используются для выбора конкретных чанков:

select [fileid]	Выбор чанка, заданного аргументом <i>fileid</i> , который должен быть номером страницы (нумерация страниц в документе идет от 1) или идентификатором чанка (имя страницы вида xxx.djvu или заданное имя). Команда <i>select</i> выбирает все чанки, если аргумент <i>fileid</i> пустой. Выбор в текущей версии ограничен - см. раздел 6 .
showsel	Выводит содержимое выбранных чанков в виде, аналогичном выводу команды <i>ls</i> .

4.3 Команды работы с разделяемыми аннотациями

select-shared-ant	Выбирает чанк с разделяемой аннотацией. В настоящее время ПО DjVu поддерживает только один такой чанк на документ. Обычно чанк содержит информацию, на которую ссылаются страницы документа. При отсутствии чанка в документе будет сообщение об ошибке.
create-shared-ant	Создает чанк разделяемых аннотаций и делает его выбранным. Если чанк уже существует, то команда только выбирает существующий чанк. Если чанка аннотаций в документе нет, то создается чанк аннотаций и проверяется, что вся информация со страниц импортирована в него.

4.4 Команды для текста и аннотаций

print-pure-txt

Печатает скрытый текстовый слой выбранных страниц. Такие же возможности имеет команда [djvutxt](#).

Структурная информация иногда представлена управляющими символами. Страницы разделены символом подачи бумаги (form feed- "\f"). Строки заканчиваются символом разрыва строки ("\n"). Колонки, области и параграфы иногда разделяются вертикальным табулятором ("\013"), разделителем групп ("\035") и разделителем объединений ("\037") соответственно. См. [пример](#).

Для вывода текста в файл, пригодный для внедрения обратно, используйте команду ***output-txt***.

print-txt

Печатает полную информацию текстового слоя. Эта информация содержит структуру текста и расположение структурных элементов на изображении страницы. Синтаксис lisp-вывода описан в [разделе 5.2](#). См. [пример](#).

Для вывода текста в файл, пригодный для внедрения обратно, используйте команду ***output-txt***.

remove-txt

Удаляет текстовый слой из выбранных страниц документа. Например, исполнение команды ***select*** **<номер или имя страницы>** и, затем, ***remove-txt*** удаляет текстовый слой выбранной страницы документа DjVu.

Пример удаления текстового слоя из **всего** документа ***myfile.djvu***:
djvused myfile.djvu -e "remove-txt" -s

set-txt [djvusedtxtfile]

Вставляет текстовый слой в выбранные страницы (**перед выполнением команды необходимо выбрать страницу командой «*select*»**). Необязательный аргумент ***djvusedtxtfile*** должен содержать имя файла, содержащего текст для вставки. Этот файл должен содержать информацию в том же виде, что получается при выводе командой ***print-txt***. Если аргумент отсутствует, программа считывает информацию для текстового слоя из **скрипта *djvused***. Считывание идет до конца файла или до строки, содержащей только точку.

output-txt

Печатает в **скрипт *djvused*** полный текстовый слой для выбранных страниц. Этот файл можно редактировать и внедрять в документ командой ***djvused*** с ключом ***-f***. См. [пример](#).

print-ant

Печатает аннотации для выбранного чанка (**перед выполнением команды необходимо выбрать страницу командой «*select*»**). Выводятся аннотации, сохраненные в чанке выбранной страницы. Разделяемые аннотации для выбранной страницы не выводятся. Для вывода всех аннотаций используйте команду ***print-merged-ant***. Синтаксис аннотации описан в [разделе 5.4](#).

Для вывода аннотации в файл, пригодный для внедрения обратно, используйте команду ***output-ant***. См. [пример](#).

print-merged-ant

Печатает все аннотации страницы, **в том числе и из разделяемых аннотаций** (**перед выполнением команды необходимо выбрать страницу командой «*select*»**). Синтаксис аннотации описан в [разделе 5.4](#).

Для вывода аннотации в файл, пригодный для внедрения обратно, используйте команду ***output-ant***. См. [пример](#).

remove-ant

Удаляет аннотацию на выбранных страницах. Например, выполнение команды ***select*** **<номер или имя страницы>** и, затем, ***remove-ant*** удаляет аннотации выбранной страницы документа DjVu.

Пример удаления аннотаций из **всего** документа ***myfile.djvu***:
djvused myfile.djvu -e "remove-ant" -s

set-ant [djvusedantfile]

Вставляет аннотации на выбранные страницы (**перед выполнением команды необходимо выбрать страницу командой «*select*»**). В необязательном аргументе ***djvusedantfile*** можно указать имя файла с аннотациями. Этот файл должен содержать аннотации в том же виде, как они выводятся командой ***print-ant***. Если файл не указан, программа считывает аннотации из **скрипта *djvused***. Считывание идет до конца файла или до строки, содержащей только точку.

output-ant

Печатает аннотации для выбранных страниц документа в [скрипт djvused](#). Этот файл можно редактировать и внедрять обратно в документ командой `djvused` с ключом `-f`. См. [пример](#).

print-meta

Печатает метаданные из аннотаций выбранных страниц (**перед выполнением команды необходимо выбрать страницу командой «[select](#)»**). Эта команда показывает набор информации, выводимой командой ***print-ant*** используя иной синтаксис. Метаданные выводятся в виде пар переменная-значение. Каждая выводимая линия содержит имя переменной, например `author`, `title`, и т.п., затем символ табуляции ("`\t`") и значение переменной в виде строки в двойных кавычках и кодировке UTF-8.

Пример вывода метаданных страницы 1 документа `myfile.djvu`:
djvused myfile.djvu -u -e "select 1; print-meta" > myfile.txt

remove-meta

Удаляет метаданные выбранных страниц. Например, выполнение команды ***select*** **<номер или имя страницы>** и, затем, ***remove-meta*** удаляет аннотации выбранной страницы документа DjVu.

Пример удаления аннотаций из **всего** документа `myfile.djvu`:
djvused myfile.djvu -e "remove-meta" -s

set-meta [djvusedmetafile]

Записывает метаданные в выбранные страницы (**перед выполнением команды необходимо выбрать страницу командой «[select](#)»**). Остальная часть аннотаций остается неизменной. Необязательный аргумент ***djvusedmetafile*** может содержать имя файла с метаданными. Файл должен содержать метаданные в том же виде, как они выводятся командой ***print-meta***. Если файл не указан, то метаданные будут считаны из [скрипта djvused](#). Считывание идет до конца файла или до строки, содержащей только точку.

print-xmp

Печатает данные XMP метаданных, содержащихся в чанке аннотаций указанной страницы (перед выполнением команды необходимо выбрать страницу командой «[select](#)»).

Эта команда выводит часть информации,, которую выводит команда ***print-ant***.

Пример вывода XMP метаданных страницы 1 документа `myfile.djvu`:
djvused myfile.djvu -u -e "select 1; print-xmp" > myfile.txt

remove-xmp

Удаляет тэг XMP метаданных, содержащихся в чанке аннотаций указанной страницы (перед выполнением команды необходимо выбрать страницу командой «[select](#)»).

set-xmp [xmpfile]

Записывает XMP метаданных, содержащихся в чанке аннотаций указанной страницы (**перед выполнением команды необходимо выбрать страницу командой «[select](#)»**). Остальная часть аннотаций остается неизменной. Необязательный аргумент ***xmpfile*** может содержать имя файла с XMP метадан-

ных в том же формате, что они выводятся командой **print-xmp**. Если файл не указан, то XMP метаданных будут считаны из [скрипта djvused](#). Считывание идет до конца файла или до строки, содержащей только точку.

output-all

Печатает в [скрипт djvused](#) полную информацию текстового слоя и аннотаций для всего документа или выбранной страницы. Этот файл можно редактировать и исполнять командой **djvused** с ключом **-f**. См. [пример](#).

4.5 Команды для работы с оглавлением

print-outline

Печатает оглавление документа. Если оглавления нет, то ничего не печатается.

Пример вывода оглавления документа **myfile.djvu**:

```
djvused myfile.djvu -u -e "print-outline" > myfile.txt
```

remove-outline

Удаляет оглавление из документа.

Пример удаления оглавления документа **myfile.djvu**:

```
djvused myfile.djvu -e "remove-outline" -s
```

У меня, почему то, выдает ошибку:

```
djvused: Error (remove-outline): unrecognized command
```

Т.е. команда не узнана?

[Comment](#) By: Leon Bottou (leonb) Date: 2012-10-02 19:05

Message:

Oops. Command "remove-outline" has been added in the git repository

Вместо этой команды для удаления оглавления используйте:

```
djvused -e set-outline djvufile -s
```

где djvufile – имя файла, из которого будет удаляться оглавление.

set-outline [djvusedoutlinefile]

Вставляет оглавление в документ. **djvusedoutlinefile** - имя файла с оглавлением. Оглавление должно быть записано так же, как оно выводится командой **print-outline**. Если файл не указан, то оглавление будет считано из [скрипта djvused](#) (???). Считывание идет до конца файла или до строки, содержащей только точку.

Пример ввода оглавления из файла **myfile.txt** в документ **myfile.djvu**:

```
djvused myfile.djvu -e "set-outline myfile.txt" -s
```

4.6 Команды для работы с миниатюрами

set-thumbnails [sz]

Производится формирование миниатюр размера **szxsz** пикселей и они вставляются в документ. Просмотрщики DjVu документов могут показывать эти миниатюры без загрузки всего файла. Типичный размер миниатюр 48-128 пикселей.

remove-thumbnails

Удаляет внедренные миниатюры из документа DjVu. При необходимости вновь внедрить миниатюры в документ, их нужно сгенерировать командой **set-thumbnails**.

4.7 Команды сохранения

Вышеприведенные команды только модифицируют копию документа DjVu в памяти. Следующие команды записывают модифицированную информацию в документ на диске.

save

Сохраняет модифицированный документ DjVu в редактируемый файл **djvufile**, который был указан в [команде запуска djvused](#). Если модификаций в документе DjVu не было, то сохранение не производится. Если при запуске **djvused** указан ключ **-s**, то до выхода из программы выполняется команда **save**.

save-bundled <filename>

Сохраняет открытый на редактирование документ DjVu как многостраничный документ-подшивку (**bundled**) с именем **filename**. Такую же функцию выполняет утилита [djvmcvt](#).

save-indirect <filename>

Сохраняет открытый на редактирование документ DjVu как многостраничный отдельный (**indirect**) документ с именем **filename**. Все вспомогательные файлы документа будут сохранены в тот же каталог как индексный файл.

Такую же функцию выполняет утилита [djvmcvt](#).

save-page <filename>

Сохраняет выбранную страницу в документ DjVu с именем **filename** (**перед выполнением команды необходимо выбрать страницу командой «[select](#)»**). Выбранная страница может содержать информацию из других чанков, т.н. чанков INCL. Эта команда производит файл с непроверенными ссылками. Этот файл надо сохранять в многостраничный документ, в котором есть чанки, на которые ссылается сохраняемая страница.

save-page-with <filename>

Сохраняет выбранную страницу в документ DjVu с именем **filename** (**перед выполнением команды необходимо выбрать страницу командой «[select](#)»**). Вся информация из разделяемых чанков также копируется в него. Эта команда всегда дает на выходе самостоятельный документ DjVu. С другой стороны, если собрать экспортированные этой командой страницы обратно в многостраничный файл, то разделяемая информация будет дублироваться на каждой странице, что увеличит объем файла.

4.8 Помощь и имя страницы

help

Выводит список команд djvused.

set-page-title <title>

Задаёт имя для страницы в документе (**перед выполнением команды необходимо выбрать страницу командой «[select](#)»**). Если страница имеет заголовки, то последние версии просмотрщиков DjVuLibre выводят эти имена вместо номеров страниц и также используют их при выборе страниц. Команда [ls](#) может быть использована для просмотра имен страниц и их идентификаторов.

ров одновременно. Для сброса имени страницы просто сделайте его идентичным идентификатору страницы.

5. Формат скриптов DJVUSED

Djvused может выводить информацию текстового слоя и/или аннотаций в текстовый файл в кодировке UTF8-ВОМ. В скрипт можно записать свои команды и выполнить их пакетно. См. [пример](#).

Djvused использует в своих скриптах для вывода текстового слоя и аннотаций простой синтаксис в кавычках.

В таком же синтаксисе аннотации хранятся в документе DjVu. Утилита djvused просто сжимает текст по алгоритму [bzz](#).

Этот синтаксис отличается от синтаксиса текстового слоя в документе DjVu. Утилита djvused производит трансляцию компактного двоичного представления, применяемого в документе DjVu, в легко редактируемый формат текста в кавычках.

Вместо номеров страниц в командах select вставляются имена страниц вида "имя_страницы.djvu". При переносе текстового слоя и/или аннотаций необходимо, чтобы имена страниц в скрипте соответствовали именам страниц в документе, в который внедряете текст и/или аннотации! Или необходимо поменять имена страниц на номера страниц вручную. (Можно при помощи поиска-замены, если имена содержат соответствующий номер страницы - страница 1 имеет имя "0001.djvu" и т.д. Помните, что нумерация страниц в djvu-книге начинается с 1).

5.1 Синтаксис скриптов DJVUSED

Скрипты Djvused являются текстовыми файлами ASCII. Допустимы печатаемые знаки ASCII, пробел (space), табулятор (tab), возврат каретки (cr) и символы nl. Использование других символов может привести к непредсказуемым результатам.

Скрипты Djvused состоят из выражений, разделенных пустыми символами (space, tab, cr или nl). Есть четыре типа выражений, именуемых целое (integer), символ (symbol), строка (string) и список (list).

Integer - целое число из одной и более цифр, интерпретируемых как обыкновенное число.

Symbol - символы или идентификаторы - последовательность печатаемых ascii-символов, представляющих собой имя или ключевое слово. Допустимы алфавитно-цифровые знаки, подчеркивание "_", минус "-" и решетка "#". Имена не могут начинаться с минуса или цифры.

String - строка означает случайную последовательность байтов, обычно интерпретируемых как последовательность символов в кодировке UTF-8. String в скриптах djvused аналогичны string в языке C. Они отделяются двойными кавычками.

Некоторые последовательности символов, начинающиеся с обратной косой черты ("\"), имеют особое назначение:

Последовательность символов	ASCII-символ
\a	BEL(007) Bell (Звонок). Используется для управления устройствами сигнализации.
\b	BS(008) Backspace (Возврат на один символ). Указывает на движение механизма печати или курсора дисплея назад на одну позицию.
\t	HT(009) Horizontal Tabulation (Горизонтальное Табулирование). Указывает на движение механизма печати или курсора дисплея до следующей предписанной 'позиции табуляции'.
\n	LF(010) Line Feed (Перевод строки). Указывает на движение механизма печати или курсора дисплея к началу следующей строки (на одну строку вниз).
\v	VT(011) Vertical Tabulation (Вертикальное Табулирование). Указывает на движение механизма печати или курсора дисплея к следующей группе строк (конец строки).
\f	FF(012) Form Feed (Перевод страницы). Указывает на движение механизма печати или курсора дисплея к исходной позиции следующей страницы, формы или экрана.
\r	CR(013) Carriage Return (Перевод каретки). Указывает на движение механизма печати или курсора дисплея к исходной (крайней левой) позиции текущей строки.
\\	BACKSLASH(134) Обратная косая черта "\"
\"	DOUBLEQUOTE(042) Двойные кавычки

Обратная косая черта и три цифры после нее означают байт, чье восьмеричное значение задано цифрами. Остальные последовательности неправильны. Все непечатаемые ascii-символы должны быть исключены.

List - список является последовательностью выражений, разделенных пустыми символами и должны быть заключены в скобки. Все типы выражений допустимы. Включая вложенные списки.

5.2 Синтаксис текстового слоя

Текстовый слой выводится в виде списков, элементами которых являются структурные блоки текста. Структурные блоки имеют вид:

(type xmin ymin xmax ymax ...)

Символ **type** должен быть один на страницу (page), колонку (column), область (region), параграф (para), строку (line), слово (word) или символ (char), приведенных в виде списка по убыванию старшинства. Целые числа (integer) **xmin**, **ymin**, **xmax** и **ymax** являются координатами прямоугольника, ограничивающего расположение текста на изображении страницы. Координаты измеряются в пикселах, начало координат левый нижний угол страницы. Оставшееся выражение в списке является одной строкой, в которой находится текст, ассоциированный с этим структурным компонентом или последовательность структурных компонентов ниже рангом.

Текстовый слой страницы представляется в виде структурного элемента page. Допустимы различные способы структурирования. Например, элемент

структуры уровня страницы может содержать только весь текст страницы, или список строк, или быть структурирован вплоть до отдельного символа.

5.3 Синтаксис оглавления

Оглавление представлено в виде одного списка вида

(bookmarks ...)

Первый элемент списка начинается с **bookmarks**. Последующие элементы списка представляют собой пункты оглавления верхнего уровня. Каждый пункт состоит из списка следующего вида:

(title url ...)

title - имя пункта. **url** - %-кодированный URL, или решетка ("#") после которой идет номер или имя страницы, или знак вопроса ("?") после которого идет команда в cgi-стиле которая будет интерпретироваться просмотрщиком djvu. Остальные выражения в списке описывают пункты оглавления второго и более уровней.

5.4 Синтаксис аннотаций

Аннотации представлены в виде последовательности выражений аннотаций. Допустимы следующие аннотации:

(background color)

Задает цвет фона вокруг изображения в просмотрщике DjVu. Цвет записан в шестнадцатеричном X11 синтаксисе #RRGGBB (RR - красный, GG - зеленый, BB - синий). Например, #000000 - черный и #FFFFFF - белый, #00FF00 - зеленый, #C0C0C0 - темно-серый.

(zoom zoomvalue)

Задает начальный масштаб просмотра. Аргумент **zoomvalue** может иметь значение **stretch** (растянуть), **one2one** (один к одному), **width** (по ширине), **page** (страница целиком), или масштаб в виде множителя - буква «d» и затем **число в диапазоне 1-999** (например, d300 или d150.)

(mode modevalue)

Задает начальный режим просмотра.

Аргумент **modevalue** может иметь значение:

color - цветной,

bw - черно-белый,

fore - передний план,

back - задний план.

(align horzalign vertalign)

Задает **выравнивание изображения в окне просмотрщика**. По умолчанию выравнивание идет по центру.

Аргумент **horzalign** (выравнивание по горизонтали) может принимать значение:

left - по левому краю окна,

center - по центру окна,

right - по правому краю окна.

Аргумент **vertalign** (выравнивание по вертикали) может иметь значение:

top - по верху окна,
center - по центру окна,
bottom - по низу окна.

(maparea URL comment area ...)

Определяет область с гиперссылкой.

URL может иметь вид:

href

(url href target)

href - строка, задающая целевой фрейм аналогично HTML тэгу <A>. Строка **href** может содержать URL в %-й кодировке или состоять из решетки ("#"), после которой идет имя страницы или номер страницы или состоять из символа ("?") после которого идет аргумент в cgi-стиле, который интерпретирует просмотрщик djvu. Номер страницы может содержать необязательный знак, который будет означать, что задан не номер страницы, а относительное смещение относительно текущей страницы. Например, **#-1** и **#+1** можно использовать для задания перемещения на предыдущую и следующую страницу соответственно.

comment - содержит сообщение, которое увидит пользователь, если курсор мыши окажется над областью гиперссылки.

area - определяет форму и местоположение области. Допустимые виды area:

- **(rect xmin ymin width height)** - прямоугольник, **xy**-координаты нижнего левого угла, ширина, высота;
- **(oval xmin ymin width height)** - овал, **xy**-координаты нижнего левого угла, ширина, высота;
- **(poly x0 y0 x1 y1 ...)** - полилиния, **xy**-координаты нижнего левого угла, **x1y1** - координаты следующего узла и т.д.;
- **(text xmin ymin width height)** - текст, **xy**-координаты нижнего левого угла, ширина, высота;
- **(line x0 y0 x1 y1)** - линия, **xy** - координаты нижнего левого угла, **x1y1** - координаты конечной точки.

Все числа имеют размерность пиксел.

Остальные значения аннотации **maparea** обозначают визуальное оформление гиперссылки.

Первый набор опций относится к оформлению границы - **rect**, **oval**, **polygon** или **text**.

- (none) - нет;
- (**xor**) - инверсная;
- (**border color**) - цветная линия;
- (**shadow_in [thickness]**)
- (**shadow_out [thickness]**)
- (**shadow_ein [thickness]**)
- (**shadow_eout [thickness]**)

где **color** - цвет в синтаксисе #RRGGBB, **thickness** - толщина, целое в пределах 1-32. **Shadow** - тень, только для **rect**. Граница по умолчанию - тонкая черная линия. Область **line** не имеет опций границы.

При заданном виде границы ее видно при перемещении курсора мыши над областью гиперссылки. Границу можно сделать видимой постоянно опцией:

- (**border_avis**)

Опции для **rect**:

- (**hilite color**) - подсветка с цветом color;
- (**opacity op**) - прозрачность (op = 0-100, default= 50).

Такая область часто применяется с пустым URL для подсвечивания нужной части изображения.

Опции для **line**:

- (**arrow**) - стрелка на конце;
- (**width w**) - ширина линии;
- (**lineclr color**) - цвет линии.

По умолчанию линия без стрелки, черного цвета с толщиной 1.

Опции для **text**:

- (**backclr bkcolor**) - цвет фона;
- (**textclr txtcolor**) - цвет текста;
- (**pushpin**) - текст свернут, вместо текста иконка, щелчок мышью по которой выводит текст.

По умолчанию цвет фона прозрачный, текст черный, текст не свернут.

(metadata ... (key value) ...)

Задаёт метаданные. Каждый элемент состоит из имени поля (**key**) и его данных (**value**). Два набора ключей полезны: метаданные библиографической системы BibTeX и импортированные PDF DocInfo. BibTeX обычно идут с маленькой буквы - year, booktitle, editor, author, и т.д. DocInfo идут с большой буквы Title, Author, Subject, Creator, Produced, Trapped, CreationDate и ModDate. Последние два ключа должны содержать дату в формате по RFC 3339.

6. Ограничения DJVUSED

Текущая версия djpeg может выделять и работать или с одним чанком или со всеми. Нельзя выделить несколько чанков.

Текущая версия при выводе в файл текста, аннотаций и прочего ("`> myfile.txt`") при его наличии в каталоге не выводит ничего. Так что необходимо вручную удалить такой одноименный файл.

7. Авторы DJVUSED

Первоначальная версия написана Léon Bottou <leonb@users.sourceforge.net>, улучшена Yann Le Cun <profshadoko@users.sourceforge.net>, Florin Nicsa, Bill Riemers <docbill@sourceforge.net> и многими другими.

8. Дополнительная информация

[divu](#), [divutxt](#), [divmcut](#), [divudump](#), [bzz](#),

Emacs djpeg front end djpeg.el on GNU Elpa repository.