

DjVu mini howto



Ver. 1.0 Novembre 2008

Augusto Scatolini (webmaster@comunecampagnano.it)

INTRODUZIONE

ALGORITMO DJVUBITONAL

ALGORITMO DJVUPHOTO

ALGORITMO DJVUDOCUMENT

RISORSE DI LETTURA

RISORSE DI SCRITTURA

WEBGRAFIA

FINE

INTRODUZIONE

DjVu (si pronuncia "déjà vu", è una tecnologia sviluppata alla fine degli anni '90 da un gruppo di ricercatori dei laboratori della AT&T con l'obiettivo di rendere possibile la realizzazione di raccolte digitali di documenti ad alta qualità a colori. DjVu è una tecnologia di compressione d'immagine, un formato file e una piattaforma per la distribuzione via rete di documenti.

Si fonda sul modello MRC, uno standard promosso dall'ITU, che propone la segmentazione dell'immagine in livelli, ciascuno dei quali viene compresso con algoritmi specifici. Separando gli elementi di sfondo, la texture della carta e le foto, e comprimendoli con un algoritmo di tipo wavelet (lo stesso alla base del formato JPEG 2000, Djvu può garantire un'elevata risoluzione del livello del testo.

Tratto da <http://it.wikipedia.org/wiki/DjVu>

La vera novità c'è stata nel dicembre 2000, quando la società LizardTech ha rilasciato sotto licenza GNU GPL la DjVu Reference Library. Adesso, DjVu è un formato aperto; inoltre i suoi creatori sviluppano DjVuLibre, una versione Open Source del software necessario per la sua codifica e decodifica.

Tratto da *GNU Linux Magazine di novembre 2008*

DjVu è inoltre un formato standard: è stato infatti inserito ufficialmente (dicembre 2006) tra i formati descritti dalla nuova versione del set di metadati dello standard Z39.87 detto anche "Data Dictionary-Technical Metadata for Digital Still Images", promosso dal NISO e dall'ANSI.

Il formato immagine DjVu produce file più leggeri dei rispettivi JPEG, GIF, PNG, TIFF, PDF, ecc.

Per esempio nella sezione delle Delibere di Giunta 2008 del sito www.comunecampagnano.it è presente la Delibera n. 8 nel formato DOC di Word, la misura di questo file è 352.256 byte. Lo stesso documento convertito nel formato PDF misura 440.470 byte. Lo stesso documento convertito nel formato DjVu misura solo 49.864 byte. Il fattore di compressione rispetto al file originale DOC è di circa 7 volte e di quasi 9 volte rispetto al PDF.

Quindi il formato DjVu garantisce un'ottima qualità dei documenti acquisiti, minore spazio occupato su disco e una velocità di visualizzazione maggiore.

Ma come è possibile?

Tutto questo è reso possibile grazie alla segmentazione dell'immagine in livelli (maschera del testo, colore del testo e sfondo), ciascuno dei quali viene compresso con algoritmi specifici.

Grazie a questa tecnica Djvu può garantire un elevato dettaglio degli elementi testuali ottenendo file molto "leggeri", veloci da visualizzare anche su computer di vecchia generazione.

ALGORITMO DJVUBITONAL

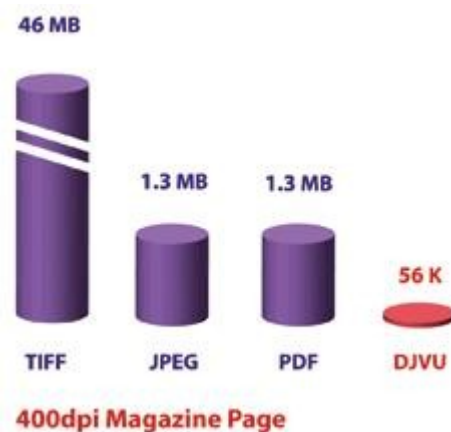
Le immagini bitonali sono quelle in bianco e nero o a due soli colori, la compressione DjVu di questi documenti, acquisiti con risoluzione 300 DPI risulta maggiore da 3 a 10 volte rispetto agli analoghi documenti in formato TIFF o PDF. L'algoritmo utilizzato da DjVu per la compressione è il DjVuBitonal

ALGORITMO DJVUPHOTO

Le immagini fotografiche possono avere anche milioni di colori e sfumature, la compressione DjVu di questi documenti, acquisiti per esempio da una macchina digitale, risulta maggiore di circa 5 volte rispetto agli analoghi documenti in formato JPG. L'algoritmo utilizzato da DjVu per la compressione è il DjVuPhoto.

ALGORITMO DJVUDOCUMENT

Le immagini dei documenti ottenute dall'acquisizione tramite scanner contengono sia testo bitonale che immagini fotografiche con milioni di colori. In questo caso DjVu tramite il processo della **SEGMENTAZIONE** divide il documento in diversi livelli e applica l'algoritmo di compressione più appropriato ad ogni livello



RISORSE DI LETTURA

Analogamente ai lettori di file PDF, per poter leggere i file in formato DjVu si deve scaricare ed installare un lettore (gratuito) presso <http://djvu.sourceforge.net/djview4.html> per gli utenti windows e digitare sulla riga di comando “sudo apt-get install djview4” per gli utenti GNU/Linux Debian e derivate (Ubuntu)

oppure scaricare un lettore portatile da <http://minidjvu.sourceforge.net/> (win)

oppure scaricare un plugin da <http://www.opendoc.it/plugin/index.php?platform=windows> per utenti win e da <http://www.opendoc.it/plugin/index.php?platform=linux> per utenti GNU/Linux

RISORSE DI SCRITTURA

Per produrre un file DjVu si possono usare diversi strumenti:

Programma DjVuSolo scaricabile da <http://djvu.org/files/DjVuSolo3.1-noncom.exe>

Programma Pdf2DjVu (win) scaricabile da <http://www.mediafire.com/?xtj0x1xdlms>

oppure da http://www.4shared.com/file/69271639/1325438a/PdfToDjvuGUI_setup_v11.html

oppure digitare da console “sudo apt-get install djvulibre-bin libdjvulibre21 libdjvulibre-text” per Linux

oppure si possono convertire file direttamente sul web ai seguenti indirizzi:

<http://www.djvu.org/any2djvu/>

<http://any2djvu.djvuzone.org/>

WEBGRAFIA

<http://www.djvu.org/>

<http://www.djvu.org/any2djvu/>

<http://any2djvu.djvuzone.org/>

<http://www.celartem.com/en/download/djvu.asp>

<http://djvu.sourceforge.net/>

<http://djvu.sourceforge.net/djview4.html>

<http://www.lizardtech.com/>

<http://code.google.com/p/pdf2djvu/>

<http://packages.ubuntu.com/intrepid/pdf2djvu>

<http://www.mediafire.com/?xtj0x1xdlms>

http://www.4shared.com/file/69271639/1325438a/PdfToDjvuGUI_setup_v11.html

<http://www.opendoc.it>

<http://www.opendoc.it/plugin/index.php?platform=windows>

<http://www.opendoc.it/plugin/index.php?platform=linux>

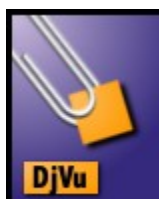
http://sourceforge.net/project/showfiles.php?group_id=32953

<http://javadjvu.foxtrottechnologies.com/>

<http://minidjvu.sourceforge.net/>

<http://www.djvu-to-pdf.info/>

<http://djvuce.sourceforge.net/index.en.xhtml>



FINE

Questo documento è rilasciato con licenza CopyLeft (tutti i rovesci sono riservati) ovvero fatene quello che vi pare! <Http://www.comunecampagnano.it/gnu/miniguide.htm>