

LINUX

Di Marco Fioretti

Fotografia digitale: *tutti i vantaggi dei metadati*

Con archivi di migliaia di foto è indispensabile catalogare correttamente. Con quali programmi?

I metadati sono tutte quelle informazioni di contorno associate ai dati o file che manipoliamo direttamente. Sono indispensabili per catalogare e soprattutto correlare correttamente i file o le informazioni che ci interessano. Negli ultimi anni anche moltissime persone che fanno soltanto un uso elementare del computer hanno cominciato, magari senza nemmeno rendersene conto, a fare i conti con questo problema. Stiamo parlando dei metadati associati alle fotografie digitali. Ormai qualunque fotocamera digitale inserisce automaticamente in ogni fotografia, come minimo, tutte le informazioni tecniche su quel particolare scatto: velocità, apertura del diaframma, sensibilità ISO e via dicendo. Molti modelli consentono di aggiungere commenti o etichette di vario tipo a ogni immagine. La stes-

sa cosa fanno praticamente tutti i software di gestione delle stesse fotografie.

Gestire correttamente i metadati delle fotografie digitali è indispensabile per almeno due motivi. In primo luogo, visto il costo pressoché nullo di ogni scatto e la capienza delle schede attuali, è facilissimo ritrovarsi con centinaia o migliaia di foto stivate più o meno disordinatamente dentro il computer. La seconda ragione è che, a differenza delle foto stampate, quelle digitali sono di fatto più difficili da ritrovare senza un indice fatto correttamente, proprio perché non sono oggetti fisici.

Nel mondo Open Source esistono diversi strumenti per aggiungere o gestire molto rapidamente molti metadati a grandi quantità di foto digitali. Questo mese presenteremo

proprio questi strumenti e i casi in cui sono più utili, insieme ad alcuni consigli su come usarli. Per saperne di più sugli standard citati nel resto dell'articolo potete leggere l'altro articolo di questa rubrica oppure i vari link (in inglese) citati nel box "Risorse".

Prima di iniziare è opportuno porsi un paio di domande di natura generale. La prima è: dove conviene conservare i metadati, all'interno dello stesso file a cui si riferiscono oppure in un database esterno? La risposta è, a meno di non avere spazio a sufficienza o altri problemi particolari, "in tutti e due i posti". Avendo un database centralizzato, vale a dire un indice comune, si possono fare ricerche di foto molto più rapidamente e semplicemente che se si dovessero esaminare ogni volta tutte le foto. Oltre a questo, funzioni come la Fuzzy Search di digiKam citata più avanti sarebbero impossibili o quasi da realizzare senza indici. Conservare tutti i metadati possibili anche all'interno delle foto garantisce invece di non perderli anche se le immagini stesse vengono copiate su Dvd o se si cambia software di gestione delle fotografie.

Passiamo alla seconda domanda: un utente finale ha solo bisogno di aggiungere metadati o anche di modificare quelli già esistenti? Entrambe le cose. La possibilità di catalogare grandi quantità nel minor tempo possibile è di ovvia utilità, soprattutto quando si acquisiscono con uno scanner anni e anni di vecchi album di famiglia. D'altra parte, dover tornare sui propri passi per correggere metadati già presenti po-



I primi standard a livello mondiale per catalogare fotografie in maniera professionale (Iptc) sono nati prima dei personal computer ma sono ancora utilizzati e accessibili da programmi Open Source come digiKam.

Breve storia dei formati per metadati *nelle fotografie digitali*

Fino agli anni 70 del secolo scorso le agenzie di stampa distribuivano le fotografie alle redazioni dei giornali di tutto il mondo tramite fax o telescrivente. Oltre alla fotografia, era indispensabile trasmettere anche tutte le informazioni necessarie per farne un uso corretto e permettere di ritrovarla facilmente in futuro negli archivi. Questo avveniva semplicemente scrivendo a macchina, in un codice più o meno formale, tutti quei dati su un foglio trasmesso insieme alla foto.

Quei metadati erano tanti e anche piuttosto importanti, visto che in ambito professionale è assolutamente necessario evitare sia la diffusione di informazioni non accurate sia problemi legali di vario tipo. Sbagliare di un giorno la data della fotografia di un incontro riservato fra uomini politici oppure non riconoscere il copyright di un'immagine può avere conseguenze assai gravi per una testata giornalistica.

Alla fine degli anni 70, proprio per ridurre tutti questi rischi e facilitare la collaborazione fra agenzie diverse, un organismo internazionale chiamato International Press Telecommunications Council (Iptc) pubblicò uno standard che specificava come scrivere, ovviamente a mano, tutte queste informazioni. Oltre a tutti i metadati che potrebbero essere sufficienti a un privato per indicizzare il suo album di famiglia, il formato Iptc comprende campi come didascalie, fonte della foto, attribuzione del copyright, località e categoria della notizia, quest'ultima codificata secondo le direttive dell'associazione americana degli editori di giornali. Nel decennio successivo alla sua pubblicazione lo standard venne poi aggiornato per essere utilizzabile anche da stazioni radio e Tv o tramite computer, per poter fare finalmente a meno di carta e macchine da scrivere. Già dal 1994 era possibile aggiungere o modificare campi Iptc nelle fotografie digitali con Photoshop. Alla fine degli anni 90, con l'esplosione dei personal computer, di Internet e poi delle tecnologie Xml, il formato Iptc si è ulteriormente trasformato. Oggi, in parte grazie ad Adobe, questi metadati sono utilizzabili anche nel formato Xml descritto dalla tecnologia Xmp (eXtensible Metadata Platform). La versione 1.0 del nuovo standard "IPTC Core Schema for XMP" è stata pubblicata a marzo 2005.

Exif

Mentre il mondo della stampa tradizionale si adeguava alle nuove tecnologie arrivavano le prime fotocamere digitali per il mercato consumer e iniziava anche in questo campo un lavoro di standardizzazione, anche se con scopi profondamente diversi. Il risultato è l'Exchangeable Image File Format (Exif, www.exif.org).

Dal punto di vista dell'utente finale, Exif è semplicemente un contenitore standard di metadati inseribile all'interno di qualsiasi file Jpeg. Oggi tutte le fotocamere digitali salvano automaticamente proprio in questo formato tutte le informazioni possibili su come è stata scattata ogni fotografia: tempo di apertura, diaframma, configurazione del flash, sensibilità Iso e molto altro. Molti modelli permettono all'utente di aggiungere subito dopo lo scatto anche commenti o etichette di vario tipo. È evidente quindi che lo scopo di Exif e quello della famiglia Iptc/Xmp sono profondamente diversi. Uno è nato per informazioni puramente tecniche e, da questo punto di vista, è e rimane un ottimo aiuto per il fotografo autodidatta. Grazie a Exif è diventato molto più facile confrontare la qualità di una foto con il modo in cui è stata scattata e quindi migliorare la propria tecnica. Iptc e Xmp, invece, sono stati creati per risolvere problemi completamente diversi, di natura gestionale, comuni solo in grandi organizzazioni.

Qual è dunque lo standard migliore? In pratica, almeno dal punto di vista dell'utente privato che vuole semplicemente catalogare le foto di famiglia su qualunque sistema operativo, la differenza è minima e probabilmente conviene usare solo Exif, per tre motivi. Prima di tutto, come spieghiamo nell'articolo principale, esistono parecchi programmi Open Source capaci di lavorare indifferenteemente con tutti e tre gli standard. In secondo luogo molti dei campi assenti da Exif e presenti negli altri formati sono davvero necessari soltanto ai professionisti. Infine, e questa è forse la ragione più importante, categorie e commenti si possono di fatto inserire anche in formato Exif.

trebbe capitare a tutti. Il caso più comune è quello in cui, al ritorno da una vacanza con centinaia di foto digitali, ci si accorge che data e ora della macchina fotografica non erano impostate correttamente o che il nome di qualche località di villeggiatura è stato scritto male nei commenti. Modificare a mano i metadati di tutte le fotografie non varrebbe la pena, ma visto che con i programmi di cui parleremo fra poco basta un solo comando per sistemare un'intera collezione perché non farlo?

Esiste infine un altro tipo di uso dei metadati che potrebbe essere utilissimo per iniziare a riordinare album digitali ed è possibile in maniera assai semplice e rapida con tutti i programmi da linea di comando di cui parleremo fra poco: l'assegnazione automatica a ogni file Jpeg di un nome significativo, contenente

ad esempio luogo e data esatta di quella fotografia. La sintassi varia da programma a programma e non è possibile in questo spazio fornire esempi completi, ma la procedura di base è davvero semplice:

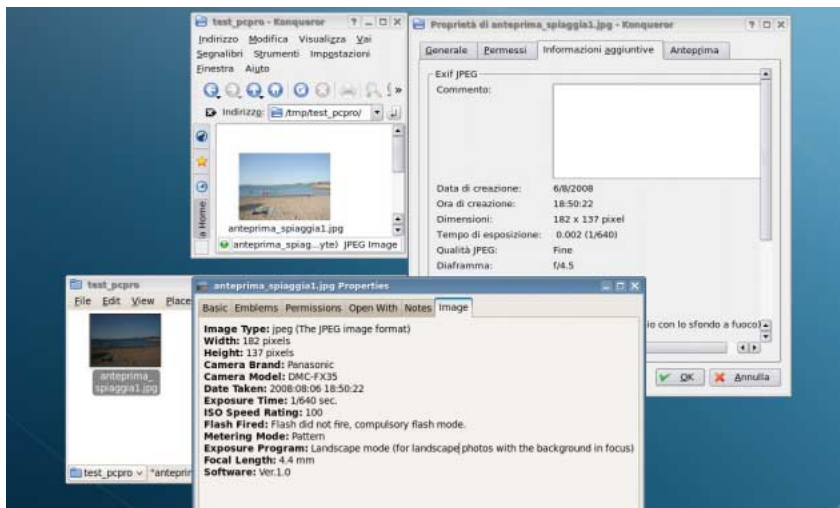
- 1) estrai da un file Jpeg il suo timestamp in formato AAAAMMDD HHDD**
- 2) salva il timestamp nella variabile DATA**
- 3) estrai dallo stesso file, se disponibile, il campo "Luogo"**
- 4) salva quel campo nella variabile LUOGO**
- 5) copia il file originale nel file \$LUOGO_\$DATA.jpg**

ovviamente questo sistema può essere facilmente adattato alle preferenze di ogni utente, per esempio salvando ogni file in una cartella col nome della località corrispondente.

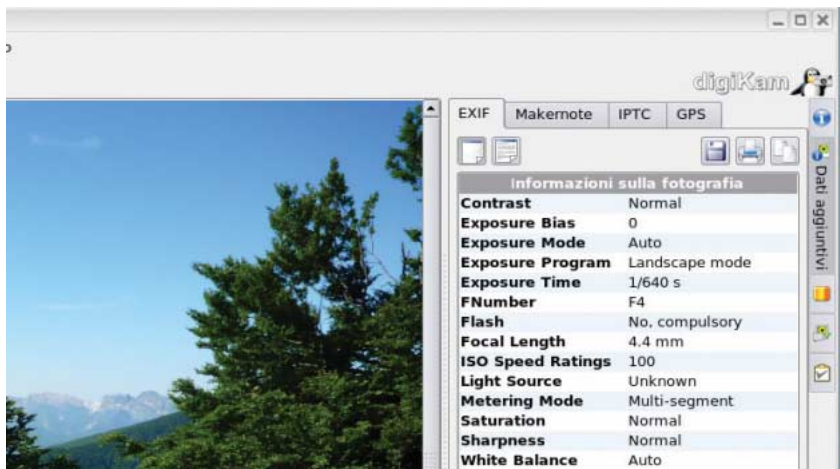
Supporto per metadati fotografici a livello desktop

Il popolare programma di fotoritocco Gimp, invece, non permette di visualizzare o modificare i metadati. Se è stato compilato con il supporto per la libreria libexif, come avviene in tutte le distribuzioni Gnu/Linux più popolari, Gimp è comunque capace di non alterare la sezione Exif di un file Jpeg quando lo si modifica e salva di nuovo in questo formato. Attenzione però! Quanto abbiamo appena detto significa, in pratica, che dopo aver modificato un file con Gimp i valori di alcuni suoi campi Exif potrebbero non essere più consistenti con le caratteristiche del file modificato.

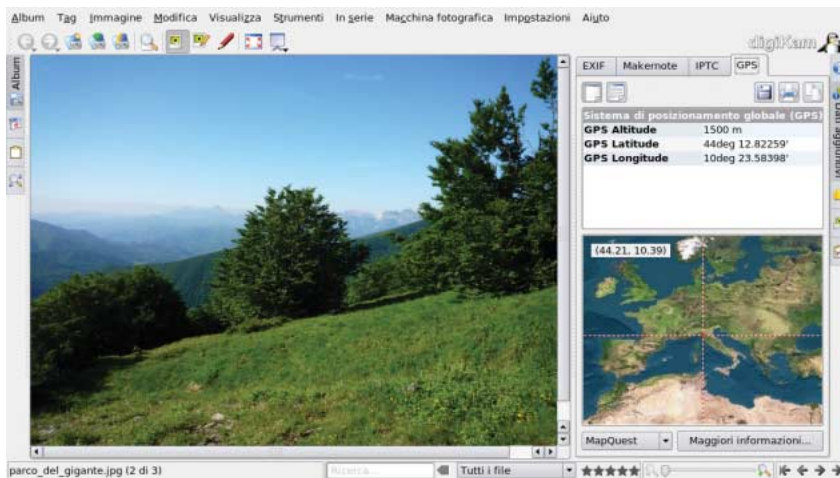
I file manager che ormai vanno per la maggiore sui desktop Gnu/Linux sono Nautilus per Gnome, Konque-



I metadati Exif che descrivono come è stata scattata una fotografia sono visibili su Linux in apposite schede dei file manager Konqueror (sopra) e Nautilus.



digiKam conosce gli standard principali per i metadati nelle fotografie digitali (Exif e Iptc) oltre a quelli non standard di vari costruttori di fotocamere e ai dati di geolocalizzazione forniti dai navigatori Gps.



L'uso dei metadati geografici è facilissimo sotto Linux: in questa schermata, digiKam mostra automaticamente su un planisfero il punto in cui è stata scattata una foto, leggendo le coordinate inserite nella foto stessa.

ror e il suo successore Dolphin per Kde. Tutti questi programmi sono già in grado di visualizzare i metadati Exif anche se non possono, almeno per il momento, modificarli. Nel caso di Konqueror e Dolphin occorre selezionare l'immagine, fare clic su di essa con il tasto destro, aprire la finestra "Proprietà" e selezionare la scheda "Informazioni aggiuntive". Konqueror consente anche di inserire dei commenti. Con Nautilus, invece, dopo aver aperto con la stessa procedura la finestra "Proprietà" si deve passare alla scheda chiamata "Immagine".

Le capacità di questi file manager sono più che sufficienti per sfruttare i metadati Exif a scopo didattico: visualizzare a fianco di una fotografia, o meglio ancora di una serie di scatti dello stesso soggetto, i relativi valori di diaframma, tempo di esposizione o configurazione del flash è un sistema assai efficace per migliorare la propria tecnica fotografica. Se si passa ai programmi specializzati nella gestione di fotografie digitali le possibilità di gestione dei metadati aumentano notevolmente, anche se l'efficienza a volte lascia a desiderare, ma occorre fare attenzione. Programmi come f-spot (<http://f-spot.org/>), gthumb (<http://gthumb.sourceforge.net/>) o digiKam (<http://digiKam.org>) non hanno alcun problema a visualizzare il contenuto dei campi Exif o Iptc, anche se non possono modificarli tutti. A parte la geolocalizzazione di cui parleremo fra poco l'eccezione, nel caso di digiKam, è la possibilità di scrivere nelle immagini alcuni parametri configurabili dall'utente (Autore, Riconoscimenti, Fonte e Detentore del copyright) nel formato Iptc.

La conoscenza di Exif viene usata in questi programmi anche per visualizzare le fotografie nella maniera ottimale o modificarle automaticamente, per esempio ruotando di novanta gradi tutte e sole le immagini con orientamento *landscape*. Le interfacce a finestre come quelle di f-spot o digiKam sono davvero facili e gradevoli da usare finché ci si limita a *visualizzare* collezioni o a fare ricerche al loro interno. Catalogare una foto alla volta invece, magari passando attraverso due o tre finestre, non è certo una procedura

velocissima. A parte la lentezza, che comunque è un problema serio solo se si devono indicizzare parecchie decine di foto a ogni sessione, il rischio maggiore è quello della portabilità.

Non tutti questi programmi trattano tutti i marcatori Exif nello stesso identico modo, soprattutto quelli non standard aggiunti ai propri modelli dai vari fabbricanti di macchine digitali. Almeno parte dei dati impiegati per indicizzare le foto viene conservata in database interni al programma, che non è detto rimangano completamente utilizzabili quando si cambia software. In alcuni casi questo è inevitabile.

DigiKam, per esempio, ha una interfaccia per la ricerca Fuzzy (approssimata), cioè di foto *simili* a quella che si sta guardando, che funziona comparando le *fingerprints* ("impronte digitali", cioè una sorta di firme digitali) delle varie foto. Le *fingerprints* sono appunto metadati conservati in un database interno a digiKam, ma per farne effettivamente uso occorre anche un programma con le stesse capacità di digiKam.

Scrittura e modifica automatica dei metadati Exif

Lo scenario di gran lunga più comune di gestione dei metadati fotografici è quello in cui le foto sono nel formato Jpeg e i metadati in formato Exif, i due standard supportati da tutte le fotocamere oggi sul mercato. Tutti i programmi da linea di comando di cui stiamo parlando supportano comunque anche gli standard Iptc e Xmp, nonché i formati nativi non compresi di parecchi fabbricanti, quindi le considerazioni che seguono valgono anche in questi casi.

Exiftools (www.sno.phy.queensu.ca/~phil/exiftool/) è un insieme di vari moduli Perl per la gestione Exif e da un programma d'interfaccia per la linea di comando. Questa architettura e il supporto per le espressioni regolari descritto più avanti gli garantisce una grande flessibilità.

I singoli moduli sono infatti utilizzabili da qualsiasi programma Perl (utilizzabili anche su Windows!), mentre l'utility già citata, chiamata

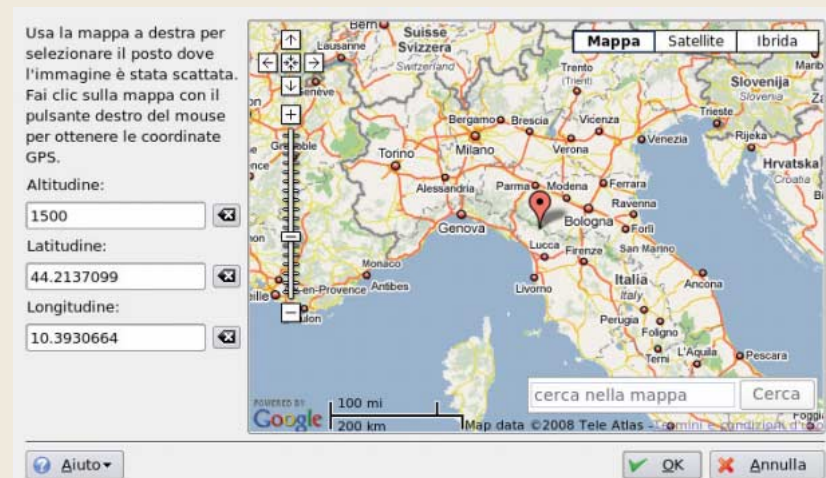
La geolocalizzazione

Il termine geolocalizzazione (a volte tradotto in inglese con *geocoding* o *geotagging*) indica l'associazione di una posizione geografica precisa a un generico documento digitale e il salvataggio in formato standard di questa informazione, normalmente all'interno del file stesso che corrisponde a quel documento. Una pagina di blog che descrive una vacanza a Parigi può essere geolocalizzata tanto quanto le singole fotografie Jpeg di quello stesso viaggio.

La geolocalizzazione è un campo vastissimo, piuttosto complesso e ancora in pieno sviluppo su Internet. Per quanto riguarda le fotografie digitali, comunque, ricavare le coordinate geografiche del punto in cui sono state scattate e aggiungerle a una fotografia è facilissimo se non si hanno pretese di precisione militare, ovvero se ci si accontenta di approssimazioni di qualche decina di metri. Sotto Linux, a livello di linea di comando si può usare *exiftool* anche per questo scopo.

Lo standard Exif tratta latitudine, longitudine e quota come qualsiasi altro campo, quindi per geolocalizzare completamente una fotografia basta aggiungervi sei campi con la tecnica spiegata in questo articolo, cioè i tre appena citati più i rispettivi riferimenti. Il nome e il formato preciso di tutti i campi da inserire sono descritti nella sezione Gps della documentazione di *exiftool*, all'indirizzo www.sno.phy.queensu.ca/~phil/exiftool/TagNames/GPS.html.

Geolocalizzare una fotografia da un'interfaccia a finestre è ancora più semplice. Il programma digiKam citato nel resto dell'articolo ha due funzioni dedicate proprio a queste operazioni. La prima è utile solo a chi usa, quando scatta foto digitali, anche un navigatore Gps capace di salvare su file in formato Gpx le coordinate dei luoghi visitati: selezionando dal menu principale Immagine->Geolocalizzazione->Correlazione si apre una finestra da cui si può caricare un file Gpx. Fatto questo, penserà digiKam ad associare a ogni foto dell'album corrente la posizione registrata dal navigatore nel momento più vicino a quello in cui quella particolare foto è stata scattata. Su Windows si può effettuare la stessa operazione con Geotag (<http://geotag.sourceforge.net/>).



L'interfaccia grafica per la geolocalizzazione di digiKam è estremamente intuitiva, anche se la precisione non è la massima possibile. Basta trovare tramite Google Maps il punto in cui è stata scattata una fotografia, aggiungere a mano la sua quota se lo si desidera e fare clic!

appunto *exiftool*, permette di sfruttare tutte le funzioni di questo pacchetto senza conoscere affatto Perl, interattivamente oppure in script shell più o meno complessi. *Exiftool* include varie funzioni per il cambiamento della data di un file e l'aggiunta, modifica o cancellazione

di qualsiasi metadato Exif, Iptc o Xmp. Il formato dei comandi è molto semplice. Per aggiungere un commento a un file Jpeg, ad esempio, basta scrivere:

```
exiftool -Comment="questo è un commento" fotografia.jpg
```

La stessa procedura vale per qualsiasi altro metadato. A volte non c'è nemmeno bisogno di scrivere a mano i valori da inserire nei vari campi Exif: l'opzione `-TagsFromFile`, seguita da un nome di file, copia direttamente i metadati da quella foto digitale a tutte le altre ancora da catalogare. L'opzione più interessante di `exiftool` è comunque quella chiamata `-if`, come si vede in questo esempio tratto direttamente dalla documentazione:

```
exiftool -alldates+=1 -if '$CreateDate ge "2006:04:02"' dir
```

che significa "aumenta di un'ora il timestamp di ogni immagine, ma *soltanto* nelle foto scattate dopo il 2 aprile 2006".

In altre parole è possibile, senza scrivere un intero script Perl o shell, dichiarare quali condizioni devono essere soddisfatte perché un qualsiasi comando `exiftool`, come avviene nell'esempio per `—alldates`, venga effettivamente eseguito.

Altre soluzioni Open Source per scrivere metadati Exif da script o interfaccia a finestre sono `Exiv2` (www.exiv2.org), `jhead` (www.sen-tex.net/~mwandel/jhead/) e `Jbrout` (<http://jbrout.python-hosting.com/>). `Jhead`, disponibile anche per Windows, è abbastanza simile a `exiftool`. `Exiv2` è in grado, come `exiftool`, di usare qualsiasi foto digitale come sorgente di metadati da inserire, senza far scrivere nulla all'utente, ed è compatibile anche con i formati nativi delle fotocamere Canon e lo standard `Dng` (Digital Negative Format) di Adobe. `Jbrout` contiene plugin (come `digiKam` e `F-spot`) per caricare foto sul proprio account Flickr o PicasaWeb.

Concludiamo citando un programma a metà, per capacità e architettura, fra quelli da linea di comando appena menzionati e i gestori di foto più popolari come `F-spot` e `digiKam`. `Mapivi` (<http://mapivi.sourceforge.net/>) è un programma a finestre scritto in Perl con librerie grafiche Tk e utilizzabile anche su Windows.

La prima caratteristica che distingue `Mapivi` dalle altre interfacce grafiche è la sua preferenza, per così dire, per lo standard `Iptc` rispetto a quello `Exif`, tanto da essere stato definito "il miglior modo di lavorare con dati IPTC su Linux senza ricorrere a un database relazionale". Mentre nel caso di `Exif` l'unico campo modificabile con `Mapivi` è il timestamp, in campo `Iptc` è possibile gestire gerarchie di etichette, categorie e modelli, oppure estrarre da una galleria le foto marcate, secondo questo standard, come "importanti".

Oltre a convertire i file `Jpeg` nei formati `Gif` o `Png` `Mapivi` è anche capace di ruotarle o ritagliarle (*cropping*) senza alcuna degradazione. Un'altra caratteristica interessante da questo punto di vista è la capacità di correggere i pixel spenti o caldi interpolando da quelli adiacenti.

Per quanto riguarda la gestione degli album sono inoltre presenti le funzioni necessarie per individuare le differenze fra varie *album* fotografici e creare gallerie con pagine `Html` statiche per Internet o `Cd-rom` a partire da un modello definibile dall'utente. `Mapivi` è anche in grado di convertire automaticamente al formato "AAA MM GG-OO MM.jpg" i nomi di tutte le foto di un album, usando come parametro il timestamp `Exif` inserito direttamente nel file `Jpeg` o la data del file stesso. •

Programmi più robusti con GlobalGcc

Il progetto `GlobalGcc` (www.ggcc.info) è stato creato a settembre 2006 da un consorzio di 15 aziende e organizzazioni internazionali guidato dal distributore Linux Mandriva (www.mandriva.com). L'obiettivo di `GlobalGcc` è sviluppare varie estensioni al compilatore `Gcc` del progetto `Gnu` per aumentarne le prestazioni e le capacità di analisi della qualità del software. La natura di queste estensioni è descritta in dettaglio nell'annuncio leggibile su <http://gcc.gnu.org/ml/gcc/2006-10/msg00676.html>. In questo contesto i coordinatori del progetto hanno recentemente annunciato la disponibilità di una utility capace di analizzare automaticamente codice sorgente C o C++ (www.ggcc.info/?q=codingrules) per verificarne il rispetto di varie regole di codifica. Lo scopo di queste regole è migliorare non tanto la leggibilità del codice quanto la sua affidabilità e facilità di manutenzione. Le basi teoriche del metodo applicato da questo programma, disponibile con licenza `Gpl` come `Gcc`, sono descritte nell'articolo www.ggcc.info/files/coding-rules-submitted-padr.pdf. L'Italia partecipa a `GlobalGcc` tramite il dipartimento di dell'Università di Parma.

Strumenti per fotografia HDR

Nel mondo della fotografia digitale la sigla `Hdr` (High Dynamic Range, in italiano Alta gamma dinamica) indica una tecnica di fotoritocco dagli effetti particolari. Il trattamento `Hdr` permette di rappresentare in un'unica fotografia un insieme di valori di luminosità, da quelli molto alti a quelli estremamente bassi, molto più esteso di quello altrimenti ottenibile su monitor o carta, con risultati notevoli anche se non proprio realistici. Per avere un'idea dell'aspetto che ha una foto `Hdr` si può visitare la galleria www.flickr.com/groups/qtpfsgui su Flickr. Concettualmente, per ottenere un'immagine `Hdr` è necessario sovrapporre diverse fotografie a gamma dinamica ristretta della stessa scena (cioè normali foto `Jpeg`), ognuna scattata con un'esposizione diversa. Su Linux esistono almeno due strumenti specializzati proprio per questo tipo di lavoro. `Qtpfsgui` (<http://qtpfsgui.sourceforge.net/about.php>) è un ambiente grafico per creare immagini `Hdr` in vari formati fra cui `Tiff` e lo standard aperto `OpenExr` (www.openexr.com). `Psftools` (<http://psftools.sourceforge.net>) è invece un gruppo di programmi per elaborazioni `Hdr` da linea di comando, utilizzabile tramite appositi plugin anche da programmi come `Matlab` o il suo concorrente Open Source `Octave`. Il pacchetto contiene comunque anche dei visori di immagini `Hdr`, per verificare immediatamente la qualità delle immagini che è in grado di creare.

Risorse

Il blog www.sitepoint.com/blogs/2006/08/24/open-source-image-archiving-exif-lptc-xmp-and-all-that/ contiene una buona introduzione generale ai problemi legati all'archiviazione di foto digitali a livello semiprofessionale con software Open Source.

Un'introduzione abbastanza sintetica alla struttura interna del formato `Exif` è la pagina www.media.mit.edu/pia/Research/deepview/exif.html. Una buona presentazione degli standard `Iptc`, adatta a non specialisti e interessante anche dal punto di vista storico, si trova su www.controlledvocabulary.com/imagetdatabases/lptc_naa.html#lptcchart.

L'articolo www.sphoto.com/techinfo/rawconverters/pages/exiflptc.htm contiene diversi commenti rilevanti sulle differenze, anche dal punto di vista pratico, fra i metadati `Exif` e quelli `Iptc`. La storia di `Exiftools` si può invece leggere nell'intervista al suo autore, Phil Harvey, su www.schnebeck.de/an-interview-with-phil-harvey.