

Dig *Italia*

Numero 2 - 2006

Rivista del digitale nei beni culturali

ICCU-ROMA

La qualità delle immagini nei progetti di digitalizzazione

Franco Lotti

Istituto di Fisica Applicata "Nello Carrara" di Firenze

L'articolo esamina le principali cause che influenzano la qualità delle immagini numeriche generate in un progetto di digitalizzazione. Il processo di formazione e acquisizione dell'immagine dalla sua forma analogica a quella numerica è determinante per la qualità dei file destinati alla conservazione (master). Esso dipende in gran parte dalle limitazioni delle prestazioni della strumentazione utilizzata e dalle modalità di acquisizione.

Un secondo aspetto non meno importante riguarda la qualità delle immagini destinate alla fruizione. Recenti sviluppi delle metodologie di compressione delle immagini e di nuovi standard consentono oggi di definire nuovi paradigmi per una fruizione dei contenuti culturali digitali orientata a una molteplicità di servizi e a un'utenza sempre più differenziata.

Introduzione

Si prenderanno in considerazione due momenti del processo che va dalla formazione alla fruizione del prodotto digitale: quello dell'acquisizione, che conduce alla produzione del cosiddetto file master destinato alla conservazione, e quello dell'immagine elaborata per essere fruita dall'utente attraverso un accesso locale o remoto.

Il flusso dei dati in un progetto di digitalizzazione può essere schematizzato nei suoi elementi essenziali come nella Figura 1. Il blocco relativo all'acquisizione comprende sia il processo di formazione dell'immagine a partire da un "originale", sia la sua conversione in forma numerica e la relativa codifica in un formato di file. Esso riassume quindi tutti i processi (ottici, elettronici, numerici) che concorrono alla produzione di un insieme di dati costituenti l'immagine numerica (digitale). Per prima cosa vengono esaminate le varie cause che influenzano la qualità dell'immagine generata, ponendoci a valle di questo blocco (punto A); in un secondo momento si prendono in considerazione gli aspetti qualitativi a valle del blocco successivo in Figura 1 (punto B), che provvede a generare e formattare i dati in funzione degli obiettivi di presentazione, trasmissione, distribuzione e in generale di fruizione, e che è rivolto quindi all'utenza.

La qualità dei master digitali

Il sistema di acquisizione comprende la sorgente di luce, l'ottica, il sensore e infine il software di acquisizione e di calibrazione. Ciascuna di queste componenti lascia una sua impronta sull'informazione che si vuole catturare e contribuisce a modificarne

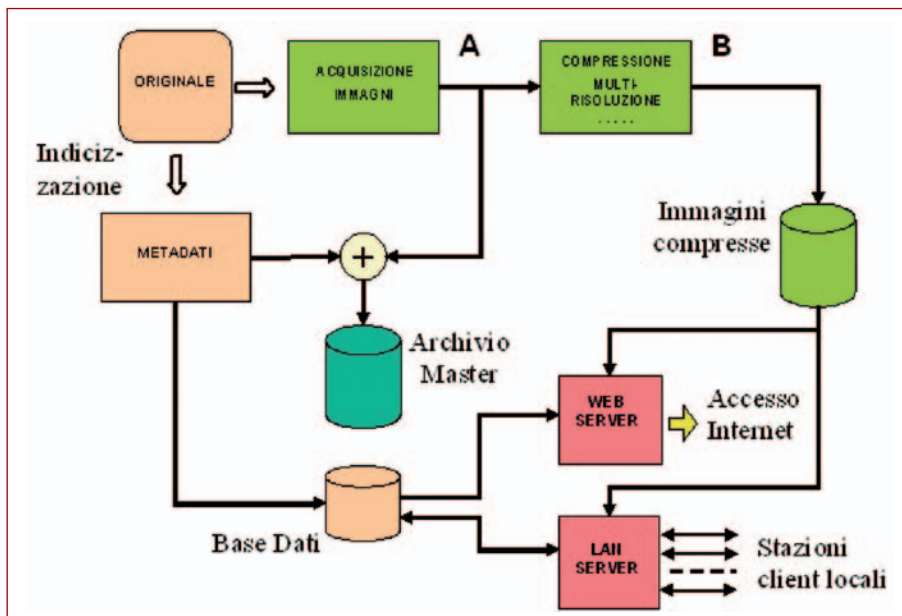


Figura 1. Esempio di schema del flusso dei dati in un progetto di digitalizzazione.

il processo di trasferimento dall'oggetto da riprodurre alla sua rappresentazione numerica finale.

Il livello di qualità ottenibile alla fine di tale processo è dunque il risultato dell'integrazione di tutte le limitazioni e distorsioni presenti nella catena di formazione e acquisizione.

Alcune limitazioni provengono dalle caratteristiche della strumentazione e sono migliorabili con l'impiego di tecnologia più sofisticata, altre invece sono intrinseche al processo stesso di digitalizzazione, come la rappresentazione di un oggetto "continuo" per mezzo di un numero finito di elementi (campionamento spaziale) o la successiva rappresentazione di questi elementi con un numero limitato di valori (quantizzazione). Queste ultime restrizioni impongono il rispetto di determinate regole per ridurre o minimizzare gli errori relativi.

Tutti questi fattori, comunque, dovrebbero essere oggetto di esame nella formulazione dettagliata del progetto di digitalizzazione, sebbene, spesso le specifiche relative alla qualità delle immagini da digitalizzare non vadano oltre la richiesta di due parametri, espressi in termini della cosiddetta *risoluzione* (termine a dire il vero inesatto, ma ormai entrato nell'uso; meglio sarebbe: *densità* o *frequenza di campionamento spaziale*), espressa in pixel per pollice (ppi) e di *profondità di colore* (bit per pixel, bpp). Con ciò si intenderebbe assolvere all'esigenza di garantire un certo livello di qualità delle immagini, commisurato agli obiettivi del progetto e alle disponibilità tecnologiche ed economiche.

Invero questi due parametri sono determinanti per generare correttamente il file numerico ottenuto dal processo di acquisizione, tuttavia non sono sufficienti a garantire un determinato livello di qualità dell'immagine. Vediamone più da vicino il significato.

Risoluzione e campionamento

Il concetto di risoluzione è legato al massimo dettaglio percepibile o rappresentabile. Riferiamoci ad esempio a una configurazione di linee alternate bianche e nere come quelle riportate sui test di risoluzione (*test charts*); ogni coppia di linee (*line pair*) corrisponde a un ciclo bianco-nero, la cui dimensione determina la frequenza di tale alternanza e si esprime con il numero di cicli per millimetro o coppie di linee per mm (lp/mm). Linee più fitte corrispondono a maggiore dettaglio e a frequenze più elevate. La risoluzione è quindi da riferirsi alla massima frequenza riproducibile con un ammesso livello di distorsione.

La teoria dei segnali campionati ci indica la relazione teorica tra la risoluzione e la densità spaziale, o frequenza di campionamento, espressa in numero di campioni per unità di lunghezza sull'originale (ad es. ppi). Essa impone infatti che il campionamento avvenga a una frequenza superiore al doppio della massima frequenza (massimo dettaglio) presente sull'originale: la cosiddetta *frequenza di Nyquist*.

Con riferimento all'esempio della *test chart*, per campionare correttamente un *pattern* di 5 cicli/mm (o lp/mm) secondo il criterio di Nyquist occorre una frequenza di campionamento superiore a 10 campioni/mm, che equivale a 254 ppi. Attenzione però: mentre si può affermare che un campionamento più rado porterebbe all'impossibilità di ricostruire correttamente le oscillazioni di 5 o più cicli/mm, ciò non significa che con 254 ppi riusciremo sicuramente ad avere una perfetta ricostruzione del *pattern* originale. E ciò non solo perché intervengono altri fattori di approssimazione e altri fenomeni di degrado dell'informazione, ma anche per ragioni teoriche, poiché la frequenza di Nyquist si riferisce non tanto al massimo dettaglio che ci interessa riprodurre, quanto al massimo dettaglio (frequenza) che è presente sull'originale.

In altre parole, se nell'area campionata esistono rapide variazioni di intensità luminosa, cioè, nell'esempio precedente, frequenze spaziali superiori a quei 5 cicli/mm che ci interessano, l'energia di queste rientrerà nell'immagine campionata a 254 ppi sotto forma di disturbo "alieno" (da cui il termine *aliasing*), caratteristico del sottocampionamento. Tale disturbo è spesso evidente in forma di *moiré* quando si sottocampionano ad esempio oggetti a trama fitta come i tessuti, riproduzioni a rotocalco o con retinature molto fitte e quindi di frequenza superiore a quella di Nyquist.

Determinare la frequenza massima presente nelle immagini, per regolare in base a essa il campionamento da effettuare, non è generalmente cosa facile. Tanto meno se non è possibile analizzare ciascuna immagine prima di fissare i parametri dell'acquisizione. Nel caso della digitalizzazione di materiale fotografico ci si può riferire alle

caratteristiche della pellicola, considerando come frequenza massima quella potenzialmente rappresentabile con quel tipo di supporto¹.

Più complesso è il caso di materiale documentario originale. Si può ricorrere in tali casi a una selezione di esempi significativi dell'insieme da digitalizzare, analizzando alcuni originali scelti a rappresentare classi omogenee per tipologia, con particolare riferimento alla presenza dei dettagli più sottili.

In teoria anche un solo tratto grafico a inchiostro nero su fondo chiaro, presentando alto contrasto e rapida variazione di livello, ha componenti di frequenza illimitata, perciò non sarebbe possibile evitare che una parte delle sue alte frequenze divenga energia di *aliasing*.

Un modo per risolvere in parte il problema è quello di eliminare, con un filtro passa-basso (*anti-aliasing*) prima del campionamento, le frequenze che superano quella di Nyquist e che comunque non saremmo in grado di rappresentare con il campionamento fissato. Spesso non è necessario ricorrere al filtro *anti-aliasing*, dato che altre parti del sistema (ottica, elettronica), avendo una risposta limitata nelle alte frequenze, funzionano esse stesse da filtro passa-basso. In entrambi i casi l'attenuazione delle alte frequenze si paga con una diminuzione della nitidezza (*sharpness*) dei contorni dell'immagine².

Una volta determinato il campionamento in funzione della risoluzione spaziale desiderata dobbiamo verificare come questa sia effettivamente raggiungibile, tenendo conto di tutte le componenti del sistema, quali le caratteristiche dell'ottica, del sensore e dell'elettronica a esso associata, che dovranno essere tali da consentire il prelievo dei campioni con la prevista densità e precisione. Anche il sistema di illuminazione, specialmente quando non sia solidale con il sensore come negli scanner piani, può introdurre disomogeneità di illuminamento che vanno ottimizzate e di cui occorre tener conto in fase di post-elaborazione.

L'ottica introduce limitazioni soprattutto legate all'apertura dell'obiettivo (luminosità), alla profondità di campo (importante per la messa a fuoco di tutte le parti dell'originale), alle eventuali distorsioni geometriche e cromatiche. Il potere risolutivo delle ottiche associate a un buon sistema di acquisizione non rappresenta tuttavia un limite significativo alle prestazioni del sistema rispetto ad altre cause di errore, come quelle esibite dal sensore. Nella catena dei processi che intervengono sulla formazione dell'immagine il sensore è infatti per la maggior parte dei casi il principale responsabile delle più severe limitazioni alla qualità dell'immagine ottenibile.

Oltre al numero e alla dimensione degli elementi sensibili, che determinano, insieme alla geometria di ripresa e all'ottica, il campionamento spaziale, il sensore "soffre" di limiti intrinseci legati alla sua tecnologia, che ne fissa le caratteristiche di dinamica, efficienza, rumorosità.

Senza entrare nel dettaglio tecnico, tuttavia è utile distinguere tra i sensori che

¹ http://www.iuav.it/CNBA/Giornate-d/2003-Le-Ot/Abstract--/Testo-degl/Buonora_relazione1.pdf_asc1.pdf.

² Un esempio di filtraggio *anti-aliasing* si trova in: <http://www.photo.net/sigma/sd9>.

forniscono dati “non interpolati” (generalmente sensori a cortina lineari o tri-lineari) e quelli che forniscono dati “interpolati” (generalmente sensori a matrice rettangolare di elementi). Infatti, per la definizione dei criteri di campionamento, quello che conta è l’effettiva densità di informazione prelevata dall’originale (legata al numero di campioni per unità di superficie e quindi al numero di ppi).

Nelle acquisizioni a colori tale informazione riguarda le tre componenti primarie (ad es. rosso, verde e blu), perciò la definizione di un pixel a colori richiede la misura dei tre valori relativi. Questa operazione viene realizzata ponendo davanti al sensore opportuni filtri che separano le tre componenti cromatiche e operando in uno dei modi sotto descritti.

Utilizzando sensori lineari (come negli scanner piani o nei dorsi a scansione): con una cortina di sensori allineati e tre scansioni dell’originale (una per colore), oppure con scansione singola e una cortina tripla di sensori (tri-lineare). Utilizzando sensori a matrice: con tre esposizioni separate, oppure, come nella maggior parte delle attuali fotocamere digitali, con un solo scatto e interpolazione. Quest’ultimo metodo consiste nel disporre su ogni elemento sensibile della matrice un micro-filtro che lo rende sensibile esclusivamente a una sola delle tre componenti cromatiche. La matrice risulta così composta da un mosaico di elementi (*Bayer pattern*), alcuni dei quali sensibili al rosso (circa il 25%), altri al verde (circa il 50%) e i restanti al blu (circa il 25%). Dopo l’esposizione, un software provvede a ricostruire, a partire da quest’unico mosaico, tre matrici distinte, una per colore, riempiendo gli elementi mancanti di ciascuna di esse con valori interpolati sulla base dei pixel più vicini. Si passa così da una a tre matrici, interpolando in totale un numero di pixel pari al doppio di quelli acquisiti.

È evidente che il risultato, soprattutto cromatico, non è equivalente ad avere acquisito realmente tutti i pixel su ciascun colore, come avviene nei sensori “non interpolati”. Non tener conto di quanti sono i pixel realmente misurati e non interpolati porta a una sottostima della densità di campionamento e quindi a un’applicazione errata del criterio di Nyquist con possibile comparsa di artefatti.

Tra i sensori a matrice, interessante è infine il caso, a mia conoscenza l’unico a tutt’oggi, del sensore sviluppato da Foveon³, costituito da tre matrici sovrapposte di sensori, sensibili al rosso, verde e blu rispettivamente, che sfruttano la diversa trasparenza ai raggi luminosi di diversa lunghezza d’onda per operare con un solo scatto non interpolato.

La quantizzazione

Il secondo livello di approssimazione che viene operato dal processo di conversione analogico/numerica dopo il campionamento è rappresentato dalla quantizzazione. Si intende con questo termine la discretizzazione (dopo quella spaziale avvenuta con il campionamento) della rappresentazione dei valori di tono corrispondenti alla

³ <http://www.abc-fotografia.com/artic/foveon-x3.htm>.

luminosità di ciascun pixel. In altre parole del passaggio da valori che variano con continuità a un insieme finito di numeri rappresentabili con le tecnologie del calcolo numerico, che, come è noto, utilizzano l'aritmetica binaria. Si parlerà quindi di numero di bit per pixel (bpp), cioè di quante cifre binarie sono necessarie per rappresentare tutte le gradazioni di intensità luminosa del segnale acquisito (profondità di colore). Tipicamente: con 8 bit si possono rappresentare $2^8 = 256$ diversi livelli di grigio o di una componente di colore; con 12 bit 4096 livelli, con 16 bit 65536 livelli. Per i pixel a colori si parlerà rispettivamente di $3 \times 8 = 24$ bpp, 36 bpp e 48 bpp.

Anche nel processo di quantizzazione si perde dell'informazione, riducendo a un insieme di "gradini" la continuità delle sfumature. Questa perdita va sotto il nome di rumore di quantizzazione.

Così come il numero dei livelli, anche la precisione di rappresentazione dell'intensità luminosa, e quindi delle varie tonalità di colore, crescerà (in modo esponenziale) all'aumentare del numero di bit per pixel. Ad esempio con 8 bit si raggiunge una precisione nella rappresentazione della gamma dei grigi pari a circa lo 0,4% ($1/256$) dell'escursione massima tra nero e bianco, mentre con 12 bit (4096 livelli) essa diventa circa lo 0,025%.

Il numero di bpp nativi esibiti da un sensore è generalmente commisurato alla effettiva capacità di quest'ultimo di distinguere due livelli vicini di luminosità o due tonalità vicine di colore. Viene determinato infatti dal costruttore in base alla effettiva dinamica e al rumore esibiti dal sensore. È quindi un parametro indirettamente indicativo, ma non affidabile, della sua qualità.

Generalmente nei sistemi di qualità medio-alta il numero di bit per pixel interni (utilizzati dal sensore), non è inferiore a 12×3 e si arriva a 48 bpp (16×3), anche se in quest'ultimo caso non sempre tutti e 16 bit per colore sono realmente utilizzati dalla sua dinamica.

È il software associato al sensore (generalmente proprietario) che consente poi di ottenere in uscita un file finale rielaborato e corretto (secondo criteri di "ottimizzazione" o "gradevolezza" delle immagini, stabiliti dal costruttore) in formato TIFF non compresso a 3×8 bpp, oppure compresso JPEG o altro. I sistemi di classe alta forniscono opzionalmente in uscita il formato nativo (RAW), non ancora elaborato, a 48 bpp. In questi casi, soprattutto nelle fotocamere digitali, si consiglia di estrarre il file in formato RAW e di effettuare fuori linea le elaborazioni necessarie per generare il formato definitivo per il master, utilizzando procedure non proprietarie o comunque note, che consentano sia di sfruttare tutte le variazioni tonali che offre il sensore quantizzato a 16 bit, sia di effettuare una migliore calibrazione del colore. Traccia di tali procedure dovrà restare nel file format stesso oppure nei metadati associati all'immagine.

Nella determinazione dei parametri di acquisizione, che dipendono dai dati di progetto, si cerca di tenersi sempre al di sopra degli stretti requisiti teorici, per compensare alcune delle inevitabili approssimazioni insite nel passaggio dalla teoria alla pratica.

È bene comunque sottolineare che l'intero sistema è affetto da numerose altre cause di errore, sulle quali i parametri suddetti non hanno influenza diretta. Tra queste vi sono le aperture finite degli obiettivi, i tempi finiti di esposizione, il rumore termico ed elettronico che è sempre presente nei sensori, l'efficienza di questi in funzione dell'intensità e della lunghezza d'onda della luce incidente, le fluttuazioni termiche degli illuminatori, le vibrazioni meccaniche, etc.

Da questo elenco sommario emerge un quadro abbastanza complesso delle cause di degrado del segnale "ideale" che influiscono sulla qualità dell'immagine numerica. Tener conto di tutti i fattori non è semplice e può non essere sempre fattibile. Va notato comunque che la qualità dell'immagine digitalizzata è solo in parte garantita da una corretta definizione dei parametri di "risoluzione" in termini di ppi e di "profondità di colore" (bpp).

MTF: una stima della nitidezza

Un modo per caratterizzare il comportamento di un sensore è quello di descrivere come vengono riprodotte le variazioni di intensità luminosa (contrasto) via via che esse si fanno più fitte, descrivendo dettagli sempre più minuti. Tale metodo, utilizzato ampiamente nel mondo della fotografia per caratterizzare sia il materiale sensibile che l'ottica, conduce alla definizione di una funzione, detta "funzione di trasferimento della modulazione", MTF (*modulation transfer function*)⁴, che riassume in sé molte delle prestazioni del sistema. Come accennato sopra, essa descrive ciò che più ci interessa, e cioè come il sistema riproduce le variazioni di contrasto al crescere della loro frequenza e quindi del dettaglio.

L'andamento della MTF del sistema può essere valutato sottoponendo ad acquisizione apposite *test charts* standard, contenenti alternanze di chiari e scuri sempre più ravvicinati, (con profilo a barre o sinusoidale), come rappresentato nella Figura 2, che descrive un esempio virtuale di un sistema di buon livello: il valore dell'MTF raggiunge il 50% a 38 lp/mm e il 10% a 69 lp/mm. Una buona stima della MTF si può ottenere, con semplici calcoli o con l'ausilio di software proprietari⁵, anche valutando l'andamento di una linea che attraversa un brusco passaggio dal bianco al nero. Un esempio reale è riportato in Figura 3 per un dorso a scansione. La differenza di forma con la curva precedente è dovuta alla differente scala, che qui è lineare. Dalla curva riportata si vede che a 0,177 cicli/pixel (corrispondenti a 10 coppie di linee/mm nel caso specifico di campionamento a 1450 ppi) l'MTF scende al 50%, indicando che tale frequenza viene restituita dal sistema con la metà del contrasto che ha nell'originale.

L'MTF al 50% è riconducibile alla percezione della nitidezza di un'immagine da parte del sistema visivo umano, che dipende appunto dalla MTF e dal livello di rumore;

⁴ <http://www.rlg.org/preserv/diginews/diginews21.html#technical>.

⁵ <http://www.imatest.com/>.

mentre alla massima frequenza percepibile dall'occhio, l'MTF scende a valori al di sotto del 5% e variano abbastanza con la persona e le condizioni di illuminazione. La densità di campionamento spaziale, cioè il numero di ppi, anche se calcolata in modo da rispettare ampiamente il criterio di Nyquist, ci dà un'idea di quale sarà la frequenza massima rappresentabile senza *aliasing*, ma niente dice sulla fedeltà e nitidezza dell'immagine, legata al comportamento del sistema a tutte le altre frequenze. Viceversa, la conoscenza dell'MTF ci dà una descrizione più completa del comportamento globale del sistema, consentendo di stimarne quantitativamente le prestazioni e le caratteristiche essenziali e di effettuare *benchmark* per confrontare i vari apparati. Due sistemi diversi infatti possono esibire la stessa risoluzione massima, ma comportarsi assai diversamente alle frequenze più basse. L'MTF ci consente inoltre di stimare le aberrazioni cromatiche, se valutata separatamente per i tre canali.

Le immagini destinate alla fruizione

Fino alla metà dell'ultimo decennio del secolo scorso la consultazione in linea di grandi banche di immagini generate dai progetti di digitalizzazione presentava problemi economici e gestionali a causa della disponibilità di supporti ancora limitati nella loro rapidità di accesso o nella capacità e affidabilità. Si pensi ad esempio ai juke-box di CD-ROM. Per questo motivo in molti casi si sceglieva di generare una versione compressa delle immagini destinate all'accesso, sia dai terminali delle reti locali, che da Internet. In alcuni casi anche i file master sono stati conservati in forma compressa, a causa dei problemi suddetti. Uno dei compiti del progetto di digitalizzazione era quindi quello di definire i parametri relativi alla compressione, stabilendo un giusto compromesso tra qualità della fruizione da un lato e dimensione dell'archivio di immagini e quindi di impegno della rete dall'altro. Tipicamente, si generavano fino a 3 formati della stessa immagine: la versione non compressa, cioè il master destinato alla conservazione, la versione compressa JPEG per la consultazione locale ad alta qualità e in alcuni casi anche un terzo formato assai più compresso, destinato alla consultazione libera via Internet.

La tecnologia dell'ultimo decennio ha rapidamente modificato lo scenario: nastri magnetici e DVD di consistente affidabilità e *disk arrays* a costi medi inferiori a 10 euro per gigabyte, hanno consentito l'impiego di nuove soluzioni per la memorizzazione dei master e dei dati in linea e indotto a una revisione dei precedenti criteri di progetto. Oggi per esempio nei file master si conservano solo immagini in formato non compresso o compresso senza perdita.

Di pari passo si è evoluto anche il software per la creazione di algoritmi di compressione di immagini, più efficienti e orientati a una fruizione flessibile, adatta alle esigenze di un'utenza differenziata.

Un terzo fattore che sta entrando in giuoco nel modellare lo sviluppo di nuove

modalità di fruizione è rappresentato dall'ingresso dell'*e-commerce* nella fornitura di servizi di tipo culturale.

L'integrazione armonica di queste opportunità, anche se stenta a decollare per una serie di motivazioni sia culturali che economiche, rappresenta un notevole potenziale per ridisegnare nuovi paradigmi di fruizione, in un quadro di maggiore efficienza e sostenibilità.

Le tecniche multirisoluzione

Nello schema di Figura 1 il blocco relativo alla compressione fa riferimento alle tecniche multirisoluzione. Con questo termine vengono indicate genericamente varie metodologie di trattamento delle immagini, utilizzate sia nella compressione che nella visualizzazione, basate su una decomposizione dell'originale in una serie di "sotto-immagini" che contengono caratteristiche di informazione differenziate e ordinate gerarchicamente. Un esempio semplice può essere la cosiddetta "piramidizzazione", che consiste nel ricavare dall'immagine originale alcune versioni ridotte, formando una piramide di vari livelli a risoluzione progressivamente più bassa, che vengono utilizzate in alcuni algoritmi di visualizzazione rapida a risoluzione variabile (*zooming*)⁶.

Tali concetti vengono utilizzati ed estesi nel sistema di compressione JPEG2000, che si basa su una particolare decomposizione in sottobande dell'immagine originale, strutturata in multirisoluzione, cioè la trasformata *wavelet* discreta⁷.

Compressione e fruizione

Ogni procedimento di compressione d'immagini genera, a partire da un file definito come originale, una sua versione compressa di minor "peso", cioè contenente un numero di byte inferiore a quello di partenza. Il "rapporto di compressione" (quoziente tra l'ingombro del file originale e quello compresso) è un parametro che misura l'efficacia di tale compressione. Questa operazione si dirà reversibile (*loss/less*) se dal file compresso si può risalire esattamente all'originale. Negli altri casi si parlerà di compressione con perdita di informazione (*lossy*).

Esistono molti modi per misurare tale perdita, collegati ai vari metodi di stima della qualità dell'immagine ricostruita. Escludendo i metodi soggettivi basati su gruppi di osservatori, i più diffusi misurano la differenza tra l'immagine originale e quella compressa (errore di ricostruzione) e ne valutano alcune funzioni statistiche come l'errore quadratico medio (*root mean square error*, RMSE), la deviazione standard o l'errore massimo.

⁶ <http://www.xlimage.it>.

⁷ Stephane Georges Mallat, *A theory for multiresolution signal decomposition: the wavelet representation*, «IEEE Transactions on Pattern Analysis and Machine Intelligence», 11 (1989), n. 7, p. 674.

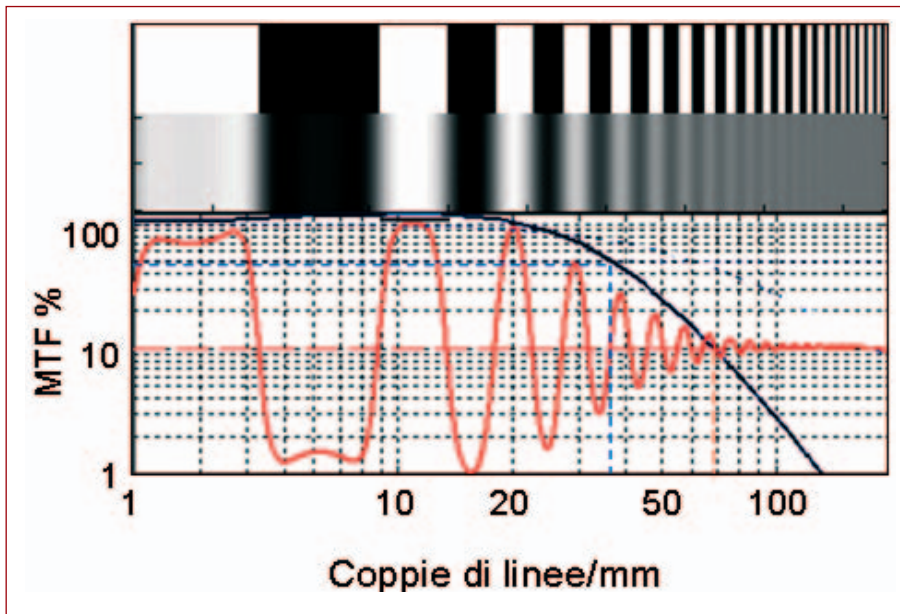


Figura 2. Valutazione della modulation transfer function (MTF). In alto: pattern originale a barre bianche e nere; al centro: immagine acquisita; sotto: in rosso il profilo del livello di grigio acquisito; in nero la funzione MTF.

Alcuni autori hanno proposto indici di qualità formati da combinazioni di parametri legati a considerazioni matematiche e psicovisive⁸. Negli esempi riportati nel seguito li useremo entrambi.

Il formato di compressione *lossy* più diffuso a tutt'oggi è certamente il JPEG. Esso consente di fissare un livello indicativo di qualità del file ricostruibile, da un minimo a un massimo secondo una scala arbitraria. Qualità più bassa produce un file compresso più piccolo, quindi un maggiore rapporto di compressione, il cui valore esatto non è peraltro prevedibile a priori, dipendendo dalle caratteristiche di ogni singola immagine. Il JPEG prevede varie modalità di compressione, invero poco utilizzate nella pratica, salvo quella di poter essere fruito progressivamente con risoluzione crescente, usata soprattutto nelle applicazioni Web⁹.

I limiti più gravi del JPEG risiedono nella comparsa di artefatti pesanti a valori di compressione medio-alta e nella scarsa flessibilità di fruizione: una volta compressa l'immagine con un certo livello di qualità, non è più possibile modificarla, se non con una nuova operazione di compressione.

⁸ Zhou Wang – Alan C. Bovik, *A Universal Image Quality Index*, «IEEE Signal Processing Letters», 9 (2002), n. 3, p. 81-84.

⁹ William B. Pennebaker – Joan L. Mitchell, *JPEG: Still Image Data Compression Standard*, New York: Van Nostrand Reinhold, 1992.

Da ricordare anche il JPEG-LS, che fornisce una versione *lossless*, comunque non integrabile nella versione *lossy* di JPEG.

Il JPEG2000, che ha ormai iniziato da tempo la sua, invero lenta, penetrazione nel mondo delle applicazioni, è da considerarsi molto di più che una nuova versione dell'ormai diffusissimo JPEG. Basato sui nuovi concetti di multirisoluzione e con un formato di file concepito per una fruizione remota e differenziata, esso consente al contempo di gestire la codifica dei file compressi e di regolamentarne la presentazione in una varietà di modalità che ne fanno uno strumento di grande flessibilità e potenzialità¹⁰.

Commentiamo solo alcune delle sue caratteristiche che riteniamo più importanti per le applicazioni di nostro interesse.

- 1 Apertura e interoperabilità, derivanti dalla sua aderenza allo standard ISO/IEC¹¹.
- 2 Efficienza. Confrontato con il JPEG a parità di errori di ricostruzione, JPEG2000 raggiunge rapporti di compressione molto più elevati; oppure, viceversa, fissato il "peso" dell'immagine compressa, fornisce livelli di qualità molto più alti. Tale effetto si evidenzia soprattutto a elevati rapporti di compressione, come mostra il semplice confronto riportato nel seguito.
- 3 Gestione della profondità di colore: dal bianco e nero puro (bi-level: 1 bpp), fino ai 16 bit per ognuna delle componenti del pixel (ad es. 48 bpp per il normale colore RGB).
- 4 Gestione contemporanea della compressione *lossy* e *lossless* (v. successivo punto 7).
- 5 Possibilità di inglobare nel formato del file compresso varie tipologie di metadati, in speciali box in vari formati: XML, come ad es. quelli del *mag schema*, oppure IPR, o altri per la gestione dei diritti di sfruttamento e di proprietà intellettuale; tutti utilizzabili ai fini di un'efficiente consultazione della base dati.
- 6 Gestione di immagini multiple. Utile per effettuare la mosaicatura di grandi immagini, o la selezione di regioni di interesse su cui operare con differente qualità, come ad es. testo e grafica.
- 7 Progressività, disponibile nelle quattro modalità seguenti:
 - Progressività di risoluzione. Consente di trasmettere immagini a scala differente cioè a risoluzione crescente (*zooming*).
 - Progressività di qualità. L'immagine presentata a una determinata scala (dimensione) può aumentare progressivamente la sua qualità fino al livello desiderato, addirittura, fatto assai interessante, fino alla completa

¹⁰ David S. Taubman – Michael W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*, London: Kluwer Academic Publishers, 2002.

¹¹ JPEG2000 è pubblicato come standard ISO/IEC e come raccomandazione ITU-T (ex Consultative Committee of the International Telephone and Telegraph, CCITT).

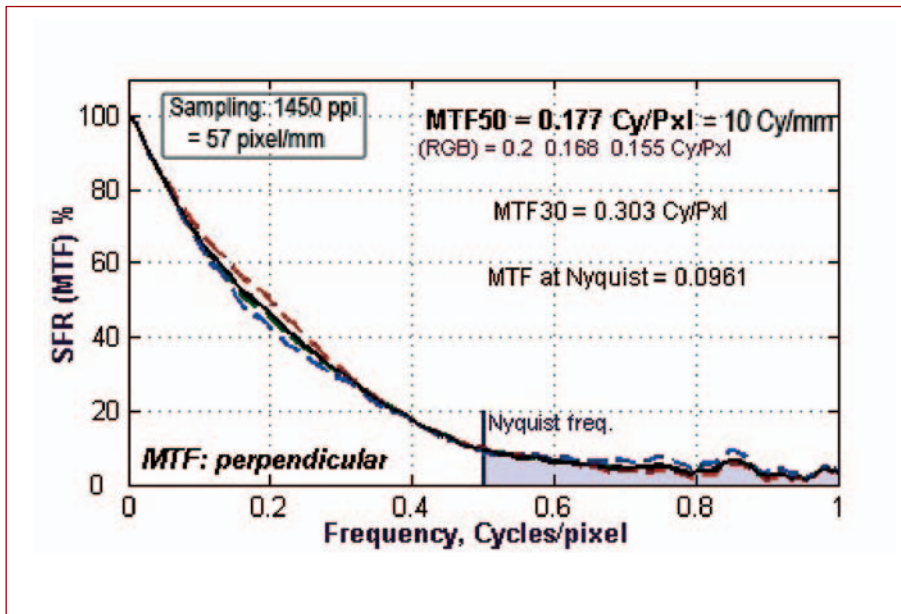


Figura 3. Esempio di Modulation transfer Function (MTF) di un dorso a scansione valutata acquisendo una test chart campionata a 1450 ppi.

- reversibilità (*lossless*).
- Sequenzialità. È possibile decodificare l'immagine compressa via via che viene ricevuta. Cosa non consentita da altri metodi di compressione, che possono presentare l'immagine solo dopo averla ricevuta interamente e poi decodificata. Ciò permette di ridurre i tempi di attesa e la memoria necessaria in ricezione sia dalla rete Internet sia anche da parte di una stampante.
- Infine la quarta modalità consente di gestire separatamente le varie componenti dell'immagine (ad es. per una foto in bianco e nero basta una componente, 3 per il colore RGB, 4 per CMYK; altre componenti possono riguardare informazioni in *overlay*, come note o grafica), fino a un numero notevolissimo di componenti per le immagini multi-spettrali.

Comprimi una volta: decomprimi in molti modi

Con questa frase Taubman¹² sintetizza la caratteristica a mio avviso più promettente del JPEG2000: la possibilità di combinare le varie modalità di fruizione progressiva secondo un'architettura assai flessibile, adatta alle esigenze di un'utenza caratterizzata da profili diversi (utente generico, utente abilitato, pagante, etc).

¹²David S. Taubman, *JPEG2000* cit., p. 410.

Infatti, una volta generato il file compresso con certe caratteristiche di massima, si può progettare a piacere la sequenza di decompressione secondo un preciso ordine di risoluzione, qualità, scala cromatica, etc., così da poter stabilire diversi livelli di fruizione. Per esempio si può trasmettere prima un'ícona in livelli di grigio, poi aggiungervi le componenti di colore con un livello medio di qualità, poi zoomare fino a 8x, poi aumentare la qualità di una sola regione di interesse prefissata, poi infine restituire tutta l'immagine ad alta qualità oppure nella versione *lossless*. In tutti i passi descritti si dovrà trasmettere solo l'informazione aggiuntiva necessaria per raggiungere il successivo livello desiderato, senza dover comprimere di nuovo l'immagine.

Ecco dunque che, avendo la possibilità di conoscere le caratteristiche dell'utente (persona con certi privilegi di accesso o dispositivo con un certo grado di risoluzione), si può conseguentemente differenziare il tipo di ricostruzione e arrestare al livello voluto il processo di trasmissione dell'informazione.

È evidente come queste caratteristiche di scalabilità sia in qualità che in risoluzione possano avere un ruolo decisivo nel progetto di digitalizzazione: una volta scelto il massimo livello di qualità della fruizione e le eventuali regioni di interesse da trattare in maniera differenziata (per es. le parti dipinte in un testo miniato), possiamo generare una sola versione di file compressi. Il progetto di fruizione potrà essere definito successivamente, nel rispetto di un ampio e flessibile paradigma di modalità distinte, e modificato in qualsiasi momento senza necessità di alterare la banca dati delle immagini già compresse.

In conseguenza di queste considerazioni, nello schema di flusso dei dati presentato in Figura 1, avendo supposto il ricorso a tecniche di multirisoluzione, abbiamo rappresentato con un unico blocco l'archivio delle immagini compresse destinate sia alla consultazione locale (LAN), sia al mondo Internet, abbandonando lo schema di due o più livelli di qualità distinti per la fruizione di alta qualità e via Web. Addirittura, scegliendo il *lossless* come massimo livello di qualità, si potrebbe anche eleggere questo archivio a livello di master, se alcuni "ragionevoli dubbi" sulla sicurezza non sconsigliassero al momento questa soluzione.

Un breve confronto

Per dare corpo agli enunciati del paragrafo precedente, riportiamo alcuni risultati tipici ottenuti nel confronto tra compressione JPEG e JPEG2000. Una misura oggettiva, anche se molto sommaria, delle differenze tra i pixel dell'originale e quelli ricostruiti può essere data dall'errore quadratico medio, mediato a sua volta sulle tre bande cromatiche RGB (*average root mean square error, A-RMSE*).

La Figura 4 mostra l'andamento di tale errore, nel caso di una pergamena acquisita a 300 ppi, e compressa con i due algoritmi. Le prestazioni di JPEG2000 superano largamente quelle di JPEG, soprattutto a valori del rapporto di compressione

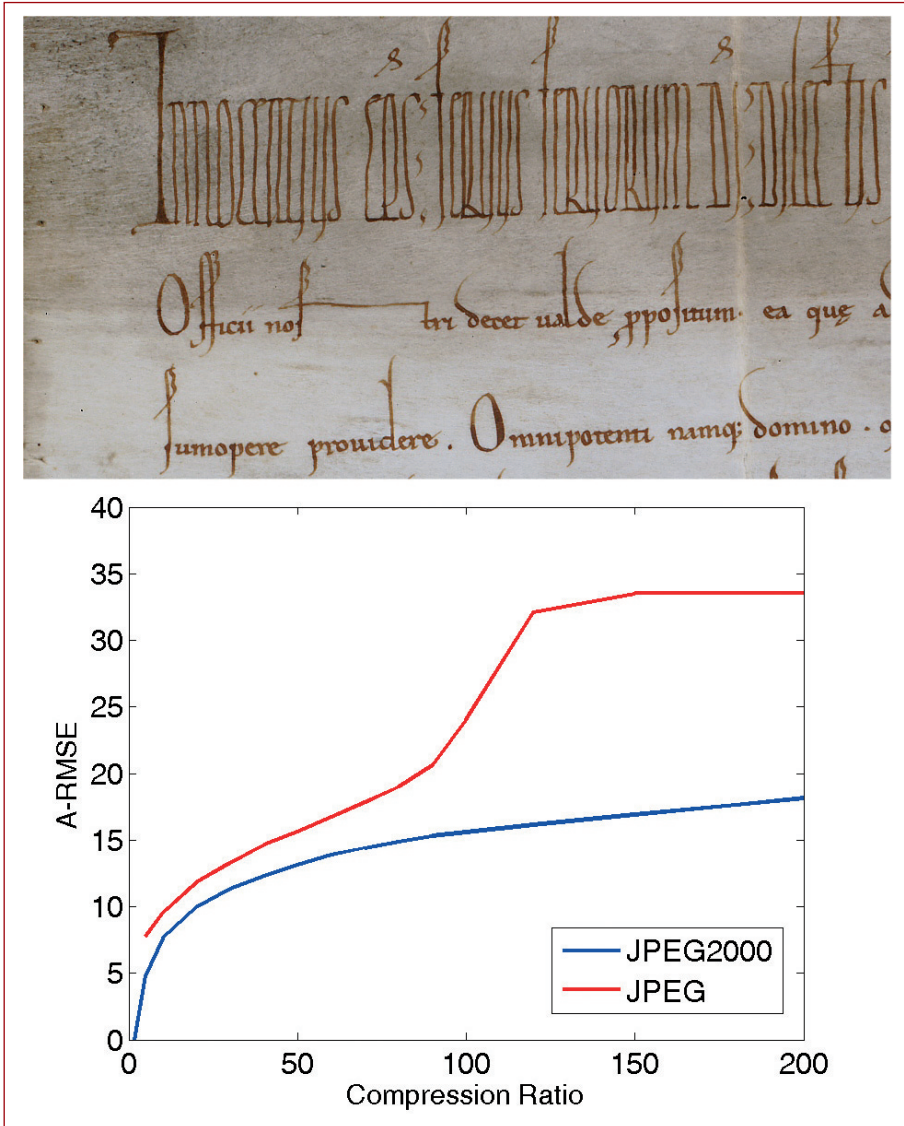


Figura 4 - Confronto tra JPEG e JPEG2000. Errore quadratico medio (A-RMSE) in funzione del rapporto di compressione, per la pergamena riportata in alto, acquisita a 300 ppi (media sulle tre componenti RGB).

superiori a 50. La curva del JPEG oltre CR = 110 diviene quasi costante, ma perde di significato per l'eccessiva distorsione, mentre JPEG2000 continua a dare risultati accettabili anche oltre CR = 200. Ciò è evidente nella Figura 5, in cui sono riportati anche i valori di un indice di qualità proposto da Bovik¹³, che varia tra 0 e 1,

¹³ Zhou Wang – Alan C. Bovik, A Universal Image Quality Index, cit., p.81.

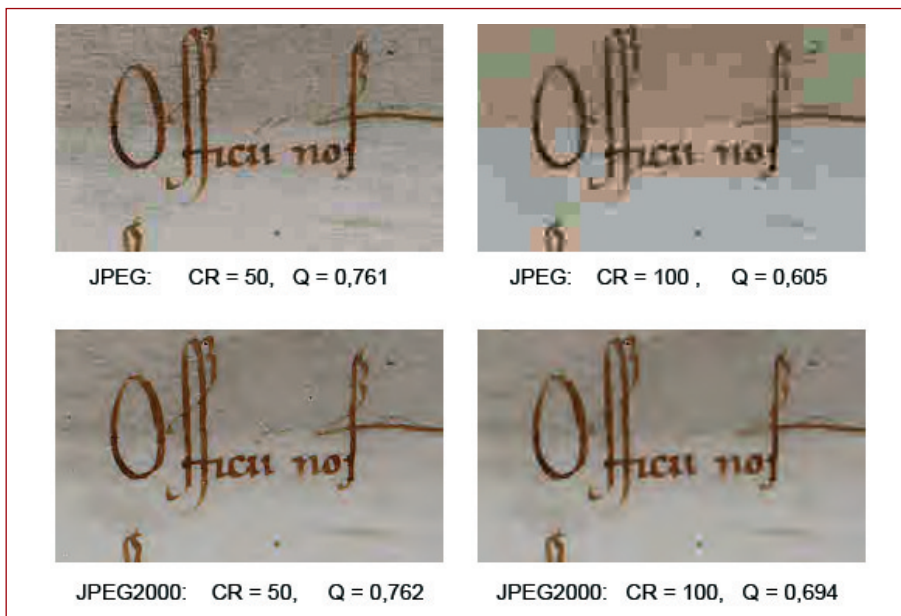


Figura 5 - Dettaglio dell'immagine di Fig. 4 ricostruita con i due algoritmi, per 2 valori medio-alti del rapporto di compressione ($CR = 50$ e $= 100$). Q è il valore dell'indice di qualità proposto in Zhou Wang -Alan C. Bovik, A universal image quality index, 2002 cit., p. 81.

attribuendo valore = 1 nel caso di identità tra immagine originale e decompressa. Il grafico di Figura 4 mostra anche che la curva di JPEG200 arriva fino a errore nullo. Infatti la ricostruzione *lossless* fornisce in questo caso un rapporto di compressione = 1,66, indicando un risparmio netto del 60% senza perdita di informazione.

Conclusioni

Una breve descrizione dei processi tipici della formazione dell'immagine numerica e degli artefatti che essi possono introdurre, seguita da semplici considerazioni sulle altre cause fisiche di errore, evidenzia l'insufficienza dei soli due classici parametri di risoluzione e profondità di colore a garantire un preciso livello di qualità dell'immagine. Risulta perciò utile completare il criterio di scelta dei requisiti di ripresa nelle specifiche tecniche di un progetto di digitalizzazione. Potrebbe essere utile inserire tra i requisiti tecnici i valori minimi della funzione di trasferimento della modulazione (MTF) a frequenze spaziali ben definite, determinati in base al dettaglio e alla nitidezza che vogliamo riprodurre, nonché i valori massimi ammessi per l'*aliasing* e la distorsione cromatica. Ciò potrebbe evitare di fornire alle terze parti esecutrici del progetto l'alibi per trincerarsi dietro al numero di ppi richiesto e fornito, senza badare ad altri aspetti qualitativi del prodotto digitale generato.

Per quanto riguarda la qualità delle immagini destinate alla fruizione, abbiamo infine passato in rassegna alcune importanti caratteristiche del JPEG2000, concludendo che i metodi di compressione come questo, basati sui concetti della multirisoluzione e sulla trasformata *wavelet*, non solo offrono nell'immediato le prestazioni migliori, ma aprono un mondo di possibilità di gestione della presentazione e dell'accesso che meritano di essere meglio conosciute e sfruttate per dare nuovo valore aggiunto al progetto di fruizione.

The main factors affecting digital image quality in a digitization project are outlined and discussed. The imaging and acquisition process has the main impact on the quality of the images to be stored as a master file for preservation. It mainly depends on the equipment features and acquisition methodology. A second not less important quality aspect concerns the images for presentation. Recent developments of image compression techniques and standards allow nowadays the design of new services and the access to digital cultural contents by ever more differentiated classes of users.

Les causes principales qui influencent la qualité des images numériques produites par un projet de numérisation sont ici examinées. Le procédé de formation et d'acquisition de l'image, de sa forme analogique à sa forme numérique, est fondamental pour la qualité des fichiers destinés à la conservation (master). Il dépend surtout des limites des prestations des instruments employés et des modes d'acquisition.

Un deuxième aspect tout aussi important concerne la qualité des images destinées à l'utilisation. Les développements récents de méthodes de compression d'images et de nouveaux standards permettent aujourd'hui de définir de nouveaux paradigmes afin d'orienter l'utilisation des contenus culturels numériques vers des services multiples et des utilisateurs de plus en plus diversifiés.