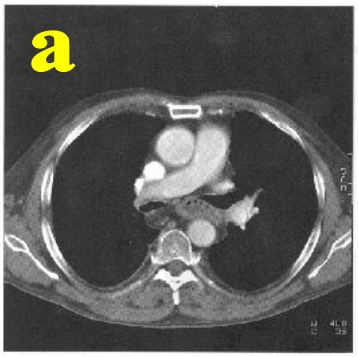
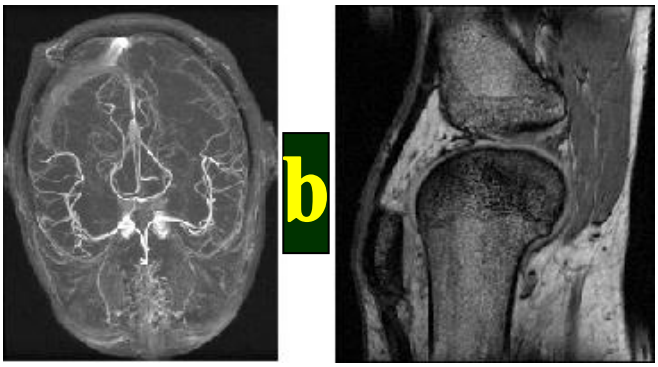
Acquisizione e qualità delle bioimmagini

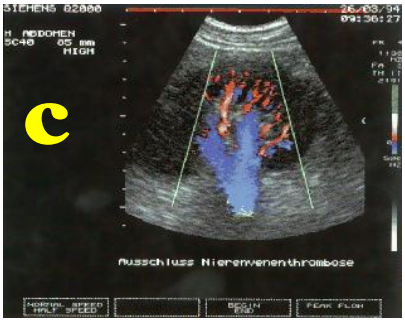
Gabriele Cevenini

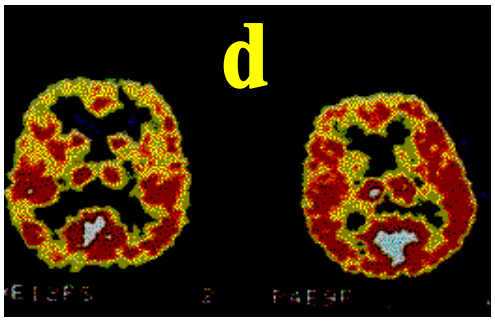
Dipartimento di Biotecnologie Mediche, Università degli Studi di Siena

Primariamente, gli esseri umani sono creature visive. Oltre il 99% delle informazioni dall'ambiente circostante viene recepito tramite la vista. Non tutti gli animali dipendono così largamente dai loro occhi: per esempio i pipistrelli usano gli ultrasuoni; il gatto ha una scarsa visione, ma un forte olfatto; i serpenti localizzano la preda attraverso le emanazioni di calore; i pesci avvertono i campi elettrici; persino gli uccelli, che notoriamente hanno una vista più sviluppata della nostra, hanno diverse modalità di visione.
L'immagine rappresenta la forma esteriore degli oggetti, esaminati direttamente o indirettamente attraverso la vista o altri sistemi sensoriali in grado di percepire la radiazione elettromagnetica emessa dai corpi. Nell'estensione più ampia della parola immagine, s'intende anche la rappresentazione mentale o artistica di una scena reale o fantastica.
Esistono importanti differenze nel paragonare la visione umana, il tipo d'informazione che essa trae dalle immagini e le modalità con cui agisce, con l'analisi d'immagine attraverso l'uso di strumenti ed elaboratori elettronici. Questi ultimi operano in modo quantitativo, mentre la visione è primariamente qualitativa e comparativa.
Le immagini biomediche (bioimmmagini) vengono solitamente usate a scopo diagnostico: sono la rappresentazione grafica, bidimensionale o tridimensionale, di segnali spaziali o di segnali spazio-temporali (filmati) provenienti dal paziente, sottoposto o meno ad opportuni stimoli.
All'osservatore inesperto le immagini biomediche appaiono assimilabili ad una realtà anatomica vista in trasparenza: ciò non è vero perchè il mezzo che le produce non è l'occhio umano. Molti eventi complessi concorrono a formare l'immagine. Essi coinvolgono diversi ambiti: fisico, chimico, ottico, elettronico, informatico. La conoscenza delle modalità connesse a questa successione di eventi è condizione necessaria per una corretta interpretazione delle immagini in chiave clinica e per valutarne le loro potenzialità diagnostiche. Qualità d'immagine e contenuto diagnostico associato sono spesso legati strettamente. Ad esempio lo studio di una lesione meniscale, richiede oltre ad una conoscenza dell'anatomia del ginocchio, la ricerca e l'analisi delle caratteristiche della lesione: dimensioni, contorni e sede. Ciò dipende in gran parte dalle modalità tecniche con cui l'immagine digitale è stata prodotta. Le immagini sono limitate dalla *risoluzione spaziale*, dalla *scala dei livelli di grigio o di colore* e dal *rumore*.









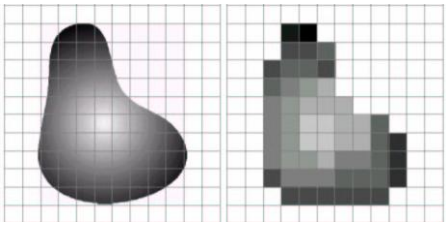
Bioimmagini: a) TAC di una sezione che mostra il cuore e i polmoni; b) RMN del cervello e del ginocchio; c) ecografia di un rene con tecnica eco-color-Doppler; d) PET del cervello.

Acquisizione

La più importante differenza tra la visione umana e gli strumenti per la produzione delle immagini deriva da come esse sono acquisite.
Il più diffuso sensore è la videocamera, ma esistono anche altri dispositivi come le camere allo stato solido (dette anche *scanner*), cioè sistemi costituiti da un *chip* dotato di sequenze ordinate di piccoli elementi d'immagine (picture elements o *pixel*).
In relazione allo schema di misura ed acquisizione dei segnali biomedici, mostrato in una lezione precedente, le camere rappresentano dunque i sensori con cui una bioimmagine è trasformata in una grandezza elettrica a sua volta rappresentabile su un monitor tramite la riproduzione sequenziale nel tempo di tutti i suoi pixel, linea dopo linea (scansione).
Si chiamano camere tutti quei dispositivi fotografici, televisivi, ecografici, a raggi X, a tomografia assiale computerizzata (TAC) a risonanza magnetica nucleare (RMN), a radionuclidi (PET, SPECT, ecc.), ecc., per trasferire immagini reali su apparecchiature elettroniche.
Tutte le bioimmagini a scopo diagnostico sono ottenute in scale di grigi: quelle a colori in realtà sono a falsi colori, cioè creati artificialmente per rappresentare e distinguere sostanze omogenee (sangue, metaboliti, ecc.). Le bioimmagini possono essere analogiche o digitali.

Bioimmagini analogiche. Sono caratterizzate dalla coincidenza tra il rivelatore del segnale proveniente dal paziente e il supporto sul quale si forma e visualizza l'immagine. L'esempio più frequente è l'immagine radiografica, rivelata e visualizzata su pellicola. Altro esempio è l'immagine ecografica che si forma dinamicamente su un monitor televisivo; quest'ultimo è un ottimo supporto per la visualizzazione d'immagini in movimento.

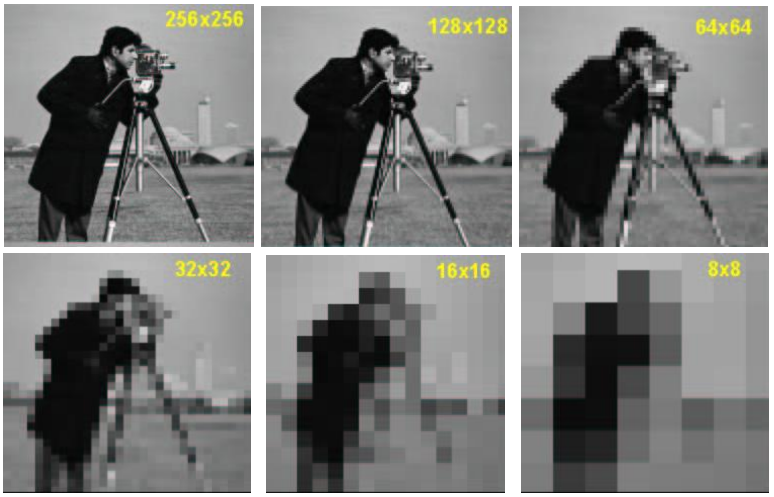
Bioimmagini digitali. Sono caratterizzate dal fatto che il segnale, prima di essere utilizzato per la generazione dell'immagine, subisce un processo di quantizzazione, in modo che solo un numero discreto di valori possa essere rappresentato sull'immagine finale. Le bioimmagini possono risultare digitali direttamente all'origine, se acquisite o ricostruite con sistemi digitali (p.es. TAC e RMN) o essere il risultato di una conversione analogico-digitale (A/D), da immagini su carta (p.es. foto o grafici), pellicole (p.es. immagini radiologiche convenzionali) o video. La conversione A/D consiste nell'applicazione di un campionamento spaziale dell'immagine e di una quantizzazione della scala dei grigi. Ciò può comportare una significativa perdita d'informazione rispetto all'immagine analogica originaria, riscontrabile visivamente come scadimento di *qualità dell'immagine*. Nonostante tutto, a causa degli enormi vantaggi derivati dall'uso di un elaboratore elettronico, ormai la quasi totalità delle bioimmagini è di tipo digitale. Esse possono infatti essere facilmente archiviate, trasmesse elettronicamente ed assai più agevolmente elaborate a fini diagnostici.
L'immagine digitale bidimensionale è formata da righe e colonne (matrice) di quadratini (o rettangolini) detti *pixel*, a ciascuno dei quali è associato un numero che esprime la gradazione di grigio o la tonalità di colore. Visivamente il pixel è quindi la più piccola porzione dell'immagine riprodotta in modo uniforme. Nell'immagine tridimensionale si parla più propriamente di *voxel* (volumetti a forma di cubo o parallepipede).



La dimensione del pixel identifica la *risoluzione spaziale* dell'immagine digitale in termini del più piccolo oggetto o struttura curvilinea, identificabili sulla stessa. In pratica, la risoluzione spaziale coincide con la frequenza spaziale di campionamento; si misura in righe di pixel per cm o più frequentemente in pixel per pollice (*pixels per inch*, ppi).

Nella figura accanto, l'immagine analogica (a sinistra) è digitalizzata con una scarsa risoluzione spaziale. Le dimensioni della matrice definiscono invece il numero dei pixel per ognuno degli assi cartesiani: p.es. 256×256, 512×512, ecc. Questi particolari valori numerici dipendono dalla natura binaria degli elaboratori elettronici e sono pertanto potenze di 2: nell'esempio, rispettivamente, 2^8 e 2^9.

L'immagine digitale può essere espressa attraverso una matrice di I(x,y) di numeri binari, dove x e y identificano la posizione del pixel in termini di riga (x) e colonna (y). In particolare, ad ogni pixel si associa un numero binario costituito da una stringa di bit (byte). Se l'immagine è a colori si hanno tre byte per pixel, corrispondenti al sistema colorimetrico usato per rappresentare l'immagine: p.es. rosso, verde e blu (sistema RGB, red-green-blue). Nelle foto accanto è riportato l'ingrandimento di un particolare relativo all'immagine digitale di una farfalla su un fiore: si notano i pixel di colore uniforme.



Effetto del campionamento spaziale sulla qualità d'immagine: in giallo il numero di pixel.

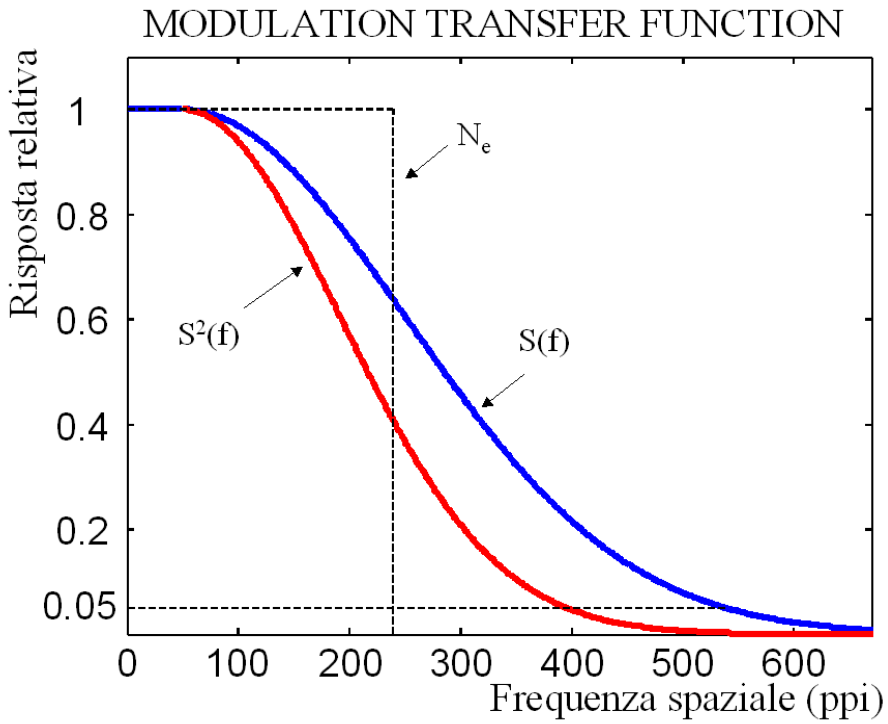
La lunghezza del byte (numero di bit) definisce l'intervallo di valori che il pixel può assumere (*livelli di quantizzazione*): p.es. 0-255 (8 bit), 0-1023 (10 bit), 0-4095 (12 bit). Tanto più alto è il numero di bit usati per la rappresentazione della scala di grigi o di colore, tanto migliore è il range dinamico dell'immagine, detto anche *risoluzione di contrasto*. In realtà, l'occhio umano ha una fisiologia tale per cui anche l'osservatore più raffinato non distingue più di 40 livelli di grigio. Con soli 6 bit (2^6=64 livelli) si ha una risoluzione di contrasto adeguata alle potenzialità della nostra vista. Oltre 32 livelli la qualità resta praticamente immutata (vedi figura sotto). Tuttavia, a fini diagnostici, si preferisce spesso usare un numero maggiore di livelli, p.es. qualora si voglia contrastare maggiormente l'immagine in ristretti intervalli di grigio (*windowing*).

Effetto della quantizzazione sulla qualità d'immagine: in rosso i livelli di grigio

Qualità e contenuto informativo

Il contenuto informativo di un'immagine varia in relazione al meccanismo d'interazione tra sistema sensoriale e mondo fisico esterno. Attraverso il processo di trasduzione si acquisiscono diverse grandezze fisiche che catturano le caratteristiche essenziali degli oggetti osservati. L'immagine acquisita da una macchina fotografica o dai sistemi sensoriali naturali (di un uomo, un gatto o un insetto) non è identica. La qualità d'immagine dipenderà pertanto dalla tipologia e dalla complessità dei sistemi sensoriali utilizzati per ottenerla artificialmente.
Nelle immagini digitali ideali (senza disturbi d'acquisizione), il contenuto informativo può essere molto semplicemente definito come il prodotto tra il numero di pixel e il numero di livelli d'ampiezza associato a ciascun pixel.
Rumore. La qualità delle immagini reali si riduce a causa del *rumore*, ossia dei disturbi d'acquisizione dell'immagine che in pratica limitano il numero di livelli d'ampiezza identificabili univocamente. Il rumore di un'immagine digitale può essere semplicemente definito come il grado d'incertezza sul livello d'ampiezza con cui ogni pixel è codificato. Il contenuto informativo delle immagini digitali reali si definisce allora come il prodotto tra il numero di pixel e il cosiddetto rapporto segnale/rumore (signal-to-noise ratio, SNR). Il SNR esprime quante volte l'intensità del segnale utile è maggiore dell'intensità di rumore. In presenza di rumore di elevata intensità tale rapporto diminuisce, ad indicare un minor numero di livelli di ampiezza distinguibili univocamente.
Anche un'eccessiva risoluzione spaziale può comportare un aumento di rumore sul singolo pixel; infatti, quando si scende nelle dimensioni del pixel, il grigio o il colore ad esso associato viene ottenuto come media di un numero sempre più basso di particelle finite (grani di argento delle foto, fosfori, raggi X, raggi elettronici ecc.). Scendendo fino al livello della dimensione delle particelle stesse (risoluzione massima), l'incertezza dovuta al rumore diventa massima, potendo avere solo due livelli di ampiezza. In immagini particolarmente rumorose, a una minor risoluzione spaziale si associa spesso un miglioramento della qualità, come illustrato in figura.

A sinistra, immagine rumorosa di una rosa con eccessiva risoluzione spaziale (pixel troppo piccoli). A destra, la stessa rosa, campionata con risoluzione più bassa, appare di miglior lettura visiva.

Funzione di trasferimento della modulazione. In base a quanto detto in precedenza si evince come l'informazione contenuta nelle immagini digitali reali dipenda in maniera interdipendente dalla risoluzione spaziale, dal numero di livelli d'ampiezza (contrasto) e dal rumore.
Per meglio caratterizzare il contenuto informativo di un'immagine ottenuta da uno strumento d'acquisizione, si ricorre al dominio della frequenza spaziale. Si riporta un diagramma relativo alla sola ampiezza del segnale (si trascura cioè la fase), detto funzione di trasferimento della modulazione (modulation transfer function, MTF) che esprime la risposta in funzione della frequenza spaziale dell'immagine, espressa in varie unità di misura: si usano di solito le coppie di linee (orizzontali o verticali) per millimetro (o pollice), ma si possono riportare anche altre unità più semplici. Nella figura si usano i ppi, comunemente utilizzati per definire la risoluzione spaziale delle immagini digitali. A frequenza nulla, si assume una risposta massima, cioè uguale a 1 rispetto all'immagine reale originaria. All'aumentare della risoluzione (frequenza di

campionamento), l'ampiezza relativa S(f) (linea blu) e il suo quadrato $S^2(f)$ (linea rossa) che indica più propriamente i livelli di contrasto relativo (quindi anche il SNR), diminuiscono a causa del rumore.
La risoluzione spaziale corrispondente a un livello di ampiezza relativa pari a 0,05 (contrasto del 5%) viene assunta come valore limite, oltre il quale un'ulteriore riduzione del passo di campionamento spaziale (dimensione del pixel) comporterebbe un contrasto non più accettabile. Nel caso in figura è pari a circa 400 ppi.
Un indice ancora più importante che si ricava dalla MTF è la *banda equivalente di rumore* N_e la quale è ottenuta calcolando l'area sottesa dalla curva $S^2(f)$, cioè facendone l'integrale. Il valore di risoluzione per cui l'area del rettangolo di altezza 1 e base N_e è pari all'integrale della curva quadratica, definisce la banda nella quale un sistema equivalente ha il 100% di risposta in ampiezza e lo 0% sopra N_e. Per scopi pratici perciò la risoluzione spaziale corrispondente a N_e (nel nostro esempio, all'incirca pari a 240 ppi) è da preferirsi alla risoluzione spaziale limite.