

Informazione multimediale

Lettere e numeri non costituiscono le uniche informazioni utilizzate dai computer ma sempre piu' applicazioni utilizzano ed elaborano anche altri tipi di informazione:

diagrammi, immagini, suoni.

In questi casi si parla di applicazioni di tipo **multimediale**

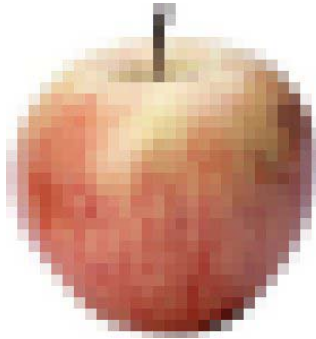
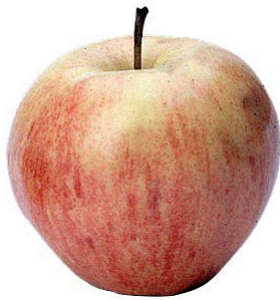
Codifica di immagini e suoni

Due fasi (logicamente) distinte:

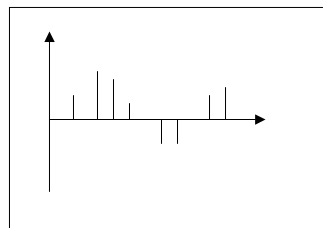
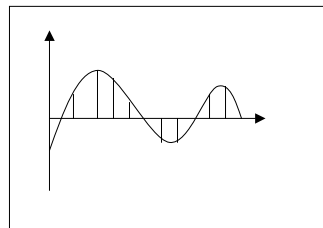
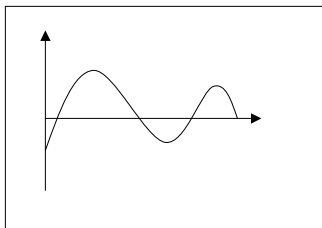
- Trasformazione di un'informazione continua in una collezione di informazioni discrete
- Codifica approssimata delle informazioni discrete

Rappresentazione discreta di informazioni continue: immagini

...

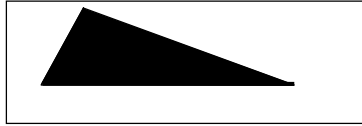


... suoni

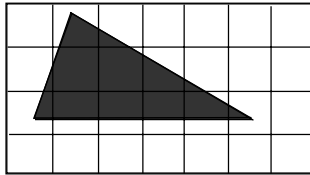


La codifica delle immagini

caso più semplice: immagini in bianco e nero senza livelli di grigio



Suddividiamo l'immagine mediante una griglia formata da righe orizzontali e verticali a distanza costante

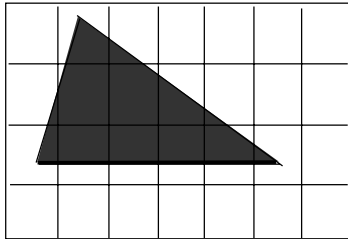


I pixel

- I quadratini della griglia sono chiamati *pixel* (picture elements) e sono intesi come unità costituenti dell'immagine
- La codifica di un'immagine consiste nella codifica dei pixel in cui viene scomposta l'immagine
- Assumiamo che un pixel sia codificato con un singolo bit che vale
 - **0** se nel pixel il *bianco* è predominante
 - **1** se nel pixel il *nero* è predominante

La codifica delle immagini

- Poiché una sequenza di bit è lineare, è necessario definire delle convenzioni per ordinare la griglia dei pixel in una sequenza. Assumiamo che i pixel siano ordinati dal basso verso l'alto e da sinistra verso destra

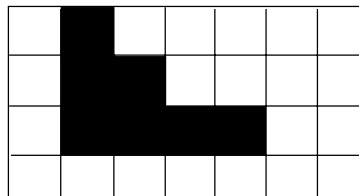


0 ₂₂	1 ₂₃	0 ₂₄	0 ₂₅	0 ₂₆	0 ₂₇	0 ₂₈
0 ₁₅	1 ₁₆	1 ₁₇	0 ₁₈	0 ₁₉	0 ₂₀	0 ₂₁
0 ₈	1 ₉	1 ₁₀	1 ₁₁	1 ₁₂	0 ₁₃	0 ₁₄
0 ₁	0 ₂	0 ₃	0 ₄	0 ₅	0 ₆	0 ₇

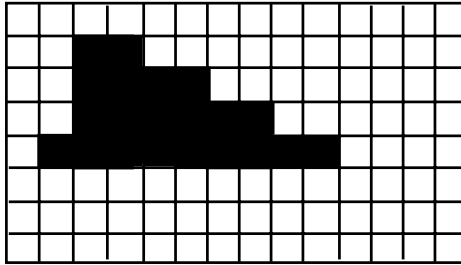
Con questa convenzione la rappresentazione della figura sarà data dalla stringa binaria

000000 011100 011000 010000

- Non sempre il contorno della figura coincide con le linee della griglia. Quella che si ottiene nella codifica è un'approssimazione della figura originaria
- Se riconvertiamo la stringa **000000011110001100000100000** in immagine otteniamo



- La rappresentazione sarà più fedele all'aumentare del numero di pixel, ossia al diminuire delle dimensioni dei quadratini della griglia in cui è suddivisa l'immagine



Risoluzione

Il numero di pixel in cui è suddivisa un immagine si chiama **risoluzione** e si esprime con una coppia di numeri ad es. 640×480 pixel (orizzontali \times verticali)

Immagini con chiaroscuro

- Assegnando un bit ad ogni pixel è possibile codificare solo immagini senza livelli di chiaroscuro
- Le immagini in bianco e nero hanno diversi livelli di intensità di grigio
- Per codificarle, si usa la stessa tecnica: per ogni pixel si stabilisce il suo *livello medio di grigio*. A tale livello viene assegnata convenzionalmente una rappresentazione binaria

- ogni pixel è codificato con un numero di bit > 1
ad esempio:
- se utilizziamo quattro bit possiamo rappresentare $2^4=16$ livelli di grigio,
- se utilizziamo otto bit ne possiamo distinguere $2^8=256$, ecc.

Esercizio

- Quanti byte occupa un'immagine in bianco e nero di 300x200 pixel con 16 livelli di grigio?
- Soluzione 30000 byte ~ 30KB

Codifica di immagini a colori

- Analogamente possiamo codificare le immagini a colori. In questo caso si tratta di individuare un certo numero di sfumature di colore differenti e di codificare ogni sfumatura mediante un'opportuna sequenza di bit
- Ad esempio, i monitor utilizzano un numero di colori per pixel che va da 256 fino a sedici milioni di colori

- La rappresentazione di un'immagine mediante la codifica dei pixel, viene chiamata:
codifica bitmap
- Il numero di bit destinati alla codifica dei colori si chiama **profondità** del colore
- Dimensione di un'immagine:
dimensione = profondità x risoluzione
- Esempio: Un'immagine con profondità di otto bit e risoluzione 640X480 richiederà:
2.457.600 bit (307.200 byte)
- I monitor utilizzano *risoluzioni* di 640X480, 1024X768, oppure 1280X1024

Tecniche di compressione

Esistono tecniche di compressione per *ridurre* lo spazio occupato dalle immagini. Le tecniche si basano su i seguenti fatti:

- le *regolarità delle immagini*: non tutti i pixel sono diversi, di solito pixel vicini sono simili (hanno colore simile)
- la *percezione umana ha dei limiti*: non siamo in grado di percepire differenze cromatiche al di sotto di una certa soglia
- *Non tutti i singoli colori* dello spettro *sono presenti* in ciascuna immagine

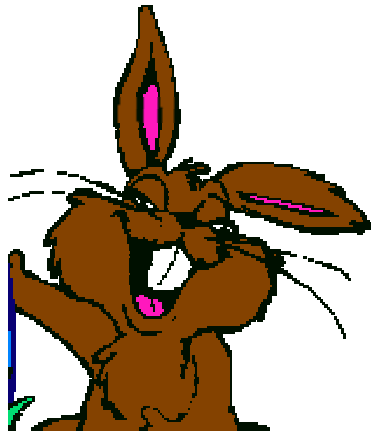
Tipi di compressione

- **compressione senza perdita di informazione:** si memorizzano pixel vicini identici una volta sola e si ricorda quante volte occorrono nell'immagine (Compressione Run Length Encoding= RLE)
- **compressione con perdita di informazione:** non si memorizzano tutti i pixel, ma solo una *frazione* di essi. Si usano funzioni matematiche di interpolazione per ricostruire i pixel mancanti

Formati standard

- **BMP** formato bitmap di Window, formato non compresso
- **GIF** (Graphic Interchange format) utilizza 8 bit per pixel e quindi distingue 256 colori. Usa una tecnica di *compressione senza perdita* (nel senso che i valori di tutti i pixel sono memorizzati)
- **JPEG** (Joint Photographic Expert Group) utilizza la codifica RGB (24bit). Usa una tecnica sofisticata di *compressione con perdita*.
- **TIFF** (Tag Image File Format) formato strutturato usato per la stampa, supporta vari metodi di compressione

Immagine GIF:2K



Formato gif 93K



Formato jpeg 30K

Immagini in movimento

- Memorizzazione mediante sequenze di fotogrammi
- La qualità della memorizzazione dipende dal numero di fotogrammi (o frame) al secondo (**fps**)
- Esempio: le immagini televisive vengono trasmesse con 25/30 fotogrammi al secondo, con una risoluzione di 576×720, e colori a 16 bit
- Supponi di avere una risoluzione di 200x100, che vengano memorizzati 24 frame al secondo e i colori siano codificati con profondità di 16 bit. Quanti byte occupa un filmato di 3 minuti?
- $20000 \times 2 \times 24 \times 180 \text{ byte} = \sim 170\text{MB}$

Codifica differenziale

- Per ottimizzare lo spazio non si memorizzano tutti i fotogrammi.
- I fotogrammi variano in modo continuo: si memorizza un primo fotogramma (chiamato *reference frame*) in modo completo, e per i successivi N solo le differenze con il primo.
- Si tratta di compressione **temporale**: interframe
- Inoltre si applica una compressione **spaziale**: intraframe, compressione di ogni singolo frame

Formati standard

- **MPEG**: memorizza in modo completo solo un fotogramma ogni 12, degli altri solo le differenze
- **AVI**: (Microsoft)
- **QuickTime**: (Apple)
- **Real Video**: (Real)

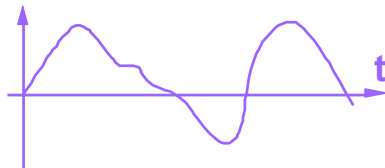
Elaborazione dell'immagine

- Una volta digitalizzate, le immagini possono essere elaborate facilmente
- Elaborare un'immagine digitalizzata vuol dire applicare una trasformazione alla sequenza di bit che codifica l'immagine
- Esempio: cambiare/neutralizzare il colore

La codifica dei suoni

- Anche i suoni possono essere rappresentati in forma digitale
- Dal punto di vista fisico un suono è un'alterazione della pressione dell'aria che, quando rilevata, ad esempio dall'orecchio umano, viene trasformata in un stimolo auditivo al cervello
- L'alterazione della pressione deve avere le caratteristiche di una *vibrazione*
- La *durata*, *l'intensità* e la *frequenza* della vibrazione sono le quantità fisiche che rendono un suono diverso da ogni altro
- Mediante un microfono le variazioni della pressione dell'aria (vibrazioni) vengono trasformate in un segnale elettrico

- Fisicamente un suono è rappresentato come un'onda (onda sonora) che descrive la variazione della pressione dell'aria nel tempo



Sull'asse delle ascisse viene rappresentato il tempo e sull'asse delle ordinate viene rappresentata la variazione di pressione corrispondente al suono stesso

Intensita'/frequenza

- L'intensita' del suono descrive l'ampiezza delle variazioni dell'onda sonora e si misura in **decibel** (DB)
- L'altezza del suono (grave/acuto) dipende dalle **frequenze** di variazione delle onde elementari che compongono l'onda sonora
- Le frequenze si misurano in **Hertz** (Hz)
- L'orecchio umano e' sensibile alle frequenze tra 20 e le 20000Hz

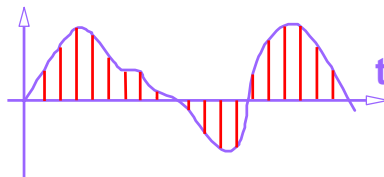
- La rappresentazione in forma d'onda è **analogica** e fornisce una descrizione *continua* dell'onda sonora
- Le rappresentazioni di tipo analogico non sono utilizzabili in informatica, data l'impossibilità di trattare informazioni di tipo continuo
- È necessario trovare un modo per codificare in forma **digitale (numerica)** un'onda sonora

Digitalizzazione del suono

- Rappresentazione digitale: Assegnare numeri che rappresentano il valore dell'ampiezza della curva in istanti successivi
- Fasi logiche della digitalizzazione:
 - *Campionamento*
 - *Quantizzazione* e *Codifica binaria*

Campionamento

- Si effettuano ***campionamenti*** dell'onda (cioè si misura il valore dell'onda a intervalli costanti di tempo)



- Quanto più frequentemente il valore di intensità dell'onda viene campionato, tanto più precisa sarà la sua rappresentazione

Tasso di Nyquist

- Tasso di **Nyquist**: per ricostruire il segnale è necessario e sufficiente un numero di campioni almeno **due volte** la frequenza dell'onda (o della sua componente di massima frequenza)

Quindi:

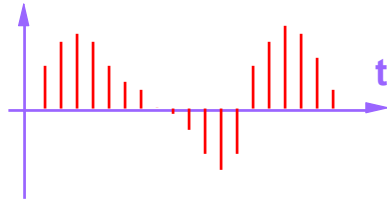
- Fissato il tasso di campionamento e' possibile catturare solo frequenze non piu' grandi della meta' del tasso di campionamento

Tasso di campionamento

- La voce umana ha come frequenza massima 4000 Hz: richiede un tasso di campionamento di 8000 campioni al secondo
- Per catturare frequenze fino a 20.000Hz e' necessario campionare il segnale almeno 40.000 volte al secondo
- Il tasso di campionamento dei **CD** è 44.100 campioni al secondo

Digitalizzazione dei campioni

- La sequenza dei valori numerici ottenuta dai campioni può essere facilmente digitalizzata



Quantizzazione e codifica

- Un dispositivo (detto ADC Analog-Digital converter): trasforma impulsi elettrici (i campioni) in una sequenza di numeri
- Il numero di bit destinati alla rappresentazione dei valori numerici ripartisce i valori di tensione in un insieme di livelli discreti:
- Esempio con 8bit (1byte) si distinguono 256 livelli di tensione
- Esempio con 16 bit (2byte) si distinguono 65536 livelli di tensione

Approssimazione della quantizzazione

- Ogni numero non rappresenta un valore preciso di voltaggio, ma un intervallo. Tanto piu' piccolo e' tale intervallo, tanto e' accurata la quantizzazione
- Esempio se l'ampiezza varia da **-10 a +10** volt una codifica con 1byte rappresenta un intervallo di 0,078125 V dell'ampiezza totale **(20/256)**.
- I CD musicali usano una codifica a 16bit perfettamente adeguata rispetto alla percezione umana

Spazio occupato

- **Esempio:** 1 secondo di voce umana a 8000 campioni al secondo, campionata a 8 bit richiede
 $64.000\text{bit} = \sim 8\text{KB}$
- **Esempio:** 1 secondo di musica qualità CD richiedono (2 canali stereofonia):
 $2 \times 44.100 \times 16 = 1.411.200\text{bit} = \sim 175\text{KB}$
1 minuto di musica richiede quindi $10.584.000 = \sim 10\text{MB}$

Formati di codifica

- Codifiche standard:
 - formato WAV (MS-WINDOWS)
 - formato RA (Real Networks)
- **MIDI**
 - codifica note e strumenti, non voce
 - molto efficiente
- **MP3** (variante MPEG per i suoni)
 - compressione *con perdita* (si fonda su caratteristiche psico-acustiche umane)
 - molto efficiente (compressione fino a 1:22)
 - molto diffuso

Riproduzione e elaborazione del suono

- Riproduzione:
 - processo inverso: un dispositivo (DAC digital-analog converter) trasforma la sequenza di numeri in impulsi elettrici
 - Il segnale elettrico viene inviato ad un dispositivo sonoro (altoparlante)
- Elaborazione del suono: come per le immagini una volta digitalizzati i suoni possono essere manipolati mediante trasformazioni matematiche (es. Moltiplicazione dell'ampiezza = aumento volume)

Riassumendo cosa posso codificare con 8.000.000 di bit (~1MB)?

- ~1 milione di caratteri: un libro di 350 pagine
- ~1000x1000pixel a 256 livelli di grigio (Immagine in bianco e nero ad alta risoluzione)
- ~2 minuti di voce
- ~5,6 secondi di musica qualità CD
- ~0,135 secondi di un filmato a 24 frame al secondo con colori a 8 bit

- *Le codifiche riducono qualsiasi informazione a sequenze di 01, sono i programmi che interpretano in modo diverso queste sequenze*