



Università degli Studi di Palermo

***Dipartimento di Ingegneria Chimica,
Gestionale, Informatica, Meccanica***

Informatica per la Storia dell'Arte

Anno Accademico 2014/2015

Docente: ing. Salvatore Sorce

La compressione

sulla base di materiale didattico originale del prof. Vincenzo Lombardo – MultiD@MS Torino

Sull'uso dell'e-mail...

- **SUBJECT:** <*testo significativo*>
- Testo chiaro e conciso
- Usare l'italiano *esteso* e grammaticalmente corretto
- Inserire i dati utili per ottenere l'informazione desiderata
- **Firmare** sempre il messaggio
 - In mancanza di **subject** e di **firma**, il messaggio rimarrà **senza risposta!!**
- Altri dettagli nella **sezione F.A.Q.** del mio sito

Notizie

Docente:

Ing. Salvatore Sorce, Ph.D.

salvatore.sorce@unipa.it, 09123862609

Lezioni:

Lunedì, 14-17, aula Multimediale A del Polo Didattico

Ricevimento:

Lunedì, 11-12, @ ex-Dip. Ing. Nucleare, edificio 6, II piano

Dopo il corso: per appuntamento

Sito web:

<http://www.unipa.it/sorce> (LEGGERE LA SEZIONE F.A.Q.)

La compressione

Arte di minimizzare il numero di bit necessari per rappresentare le informazioni

Obiettivi:

- ridurre la memoria occupata
- ridurre i costi (tempo) di trasmissione
- mantenere una buona qualità in confronto alla versione non compressa (al limite, la *stessa* qualità)



Obiettivo: *riproduzione trasparente*

- Riproduzione dei contenuti “uguale” all’originale
 - Audio che anche orecchie sensibili non riescono a distinguere dall’originale (non compresso)
 - Immagini che anche occhi sensibili non riescono a distinguere dall’originale (non compresso)
 - Sequenze audio/video che “si vedono” e “suonano” come l’originale (non compresso)

- Numero minimo di bit mantenendo una riproduzione *trasparente* del segnale



Esempio: l'audio digitale

- Audio digitale non compresso:
 - alta fedeltà delle copie del segnale
 - robustezza
 - gamma dinamica estesa
- Svantaggio: alto tasso di trasferimento dati
 - campionamento: 44.1 kHz (CD), 48 kHz (DAT)
 - quantizzazione: PCM lineare a 16 bit
- L'audio digitale non compresso occupa ...
 - molto spazio sull'hard-disk per la memorizzazione
 - una banda ampia sul canale di trasmissione

Esempio: brano di 1 min con qualità CD

- Parametri
 - campionamento = 44.100 campioni/sec
 - quantizzazione = 16 bit/campione = 2 byte/campione
 - canali = 2
- Memoria = $44.100 \text{ c/s} * 2 \text{ can} * 2 \text{ byte/c} * 60 \text{ s}$
 $\sim 10 \text{ Mbyte per ogni minuto di musica stereo.}$
- Tempo necessario per il trasferimento di un minuto di musica stereo su un collegamento a 200 kbps =

$$\frac{10 \text{ Mbyte} * 8 \text{ bit/byte}}{(200 \text{ kbps} * 60 \text{ s})} = \frac{10 * 1024 * 1024 * 8}{200 * 1024 * 60} \sim 7 \text{ minuti}$$

... per scaricare un solo minuto di musica stereo!
 (sarebbe impossibile qualsiasi applicazione *live*, come la web-radio)

Esempio: immagine a colori 1600x1200 px

- Parametri:
 - $1600 \times 1200 = 1.920.000$ pixel
 - Quantizzazione a 24 bit/px
- Memoria = $19200000 \text{px} \times 24 \text{ bit/px} = 46.080.000 \text{ bit}$
 - In Byte -> $46.080.000 \text{ bit} / 8 \text{ bit/byte} = 5.760.000 \text{ Byte}$
 $\sim 5,5 \text{ Mbyte}$
- Tempo necessario per il trasferimento di un'immagine 1600x1200 a 24 bit/px su un collegamento a 200 kbps =

$$\frac{46.080.000 \text{ bit}}{200 \text{ kbps}} = 230 \text{ secondi } (\sim 4 \text{ minuti})$$

(sarebbe impossibile qualsiasi applicazione *live*, come lo streaming A/V)

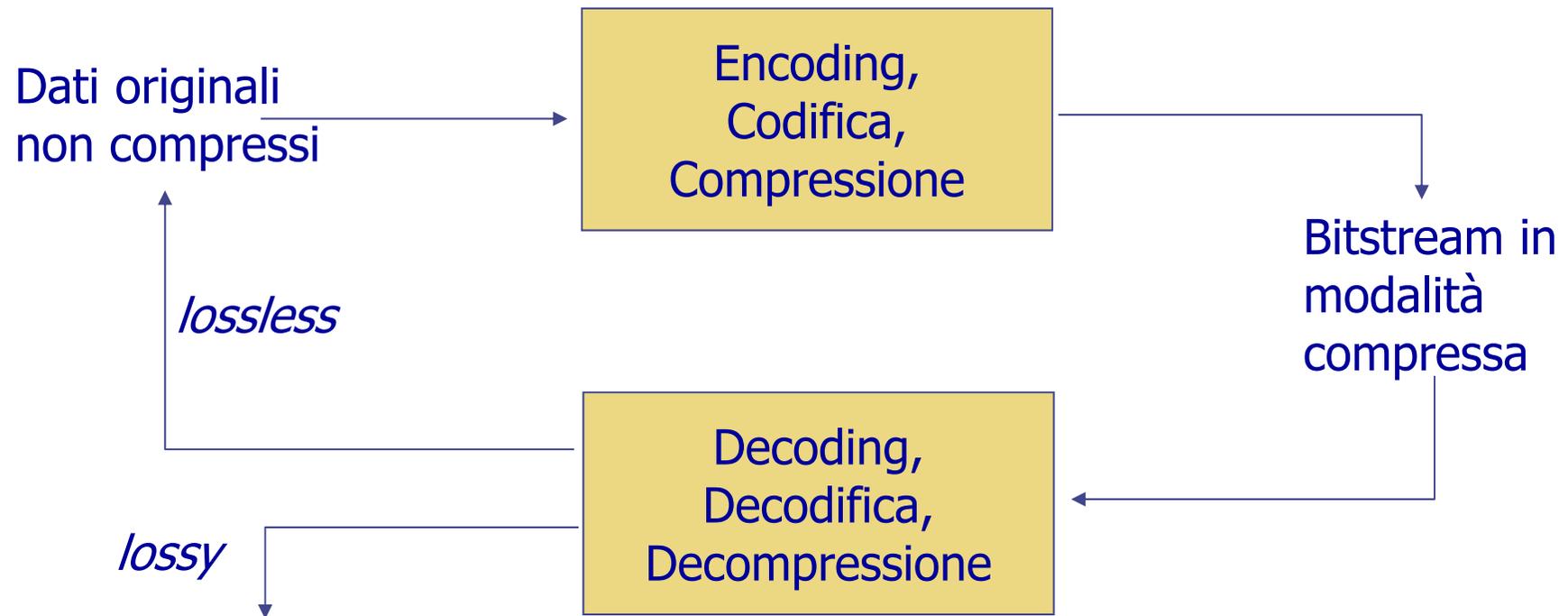
Tipi di compressione

- Classificazione in base al principio:
 - Compressione percettiva
 - ◆ Si basa su considerazioni fisiologiche, cioè su come l'informazione viene percepita (es.: il suono dalle orecchie, le immagini dagli occhi)
 - ◆ Il risultato è (o tende ad essere) percettivamente uguale all'originale
 - ◆ Alti tassi di compressione possibili, a costo di un risultato percettivamente peggiore
 - Compressione non-percettiva
 - ◆ Si basa su considerazioni esclusivamente numeriche
 - ◆ Il risultato può essere identico all'originale (non compresso)
 - ◆ Il tasso di compressione dipende da come è strutturata l'informazione non compressa (può anche peggiorare!)

Tipi di compressione

- Classificazione in base al risultato:
 - Compressione con perdita
 - ◆ La sequenza di bit ricostruita *non è* uguale all'originale
 - ◆ Nonostante ciò, il segnale finale può essere *percepito* come uguale all'originale
 - Compressione senza perdita
 - ◆ La sequenza di bit ricostruita *è uguale* all'originale
 - ◆ Il segnale finale è identico all'originale

Compressione e decompressione



Segnale non compresso *diverso*
ma percepito come l'originale



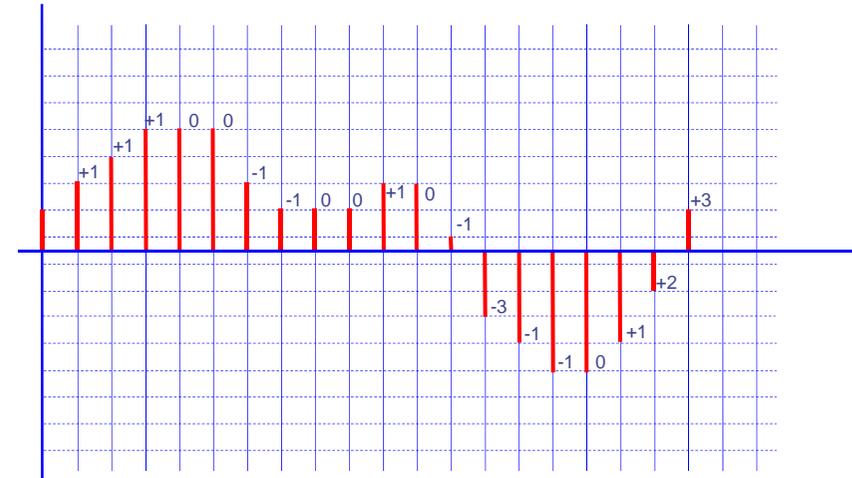
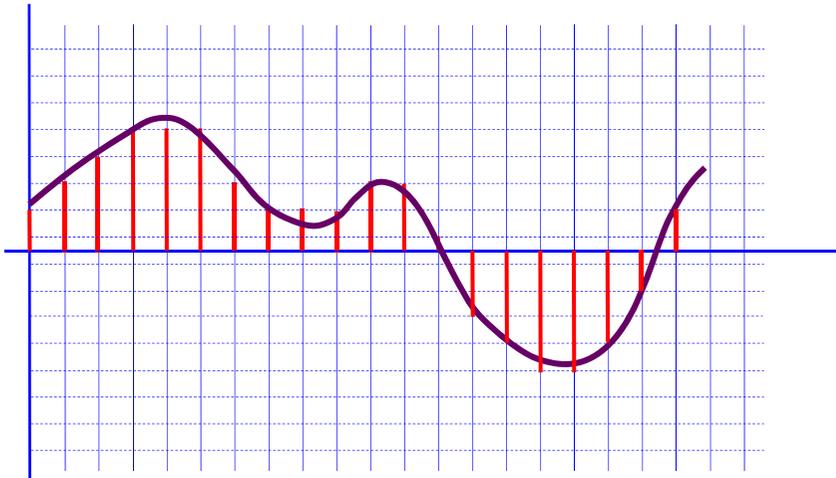
Compressione non-percettiva



Compressione non-percettiva

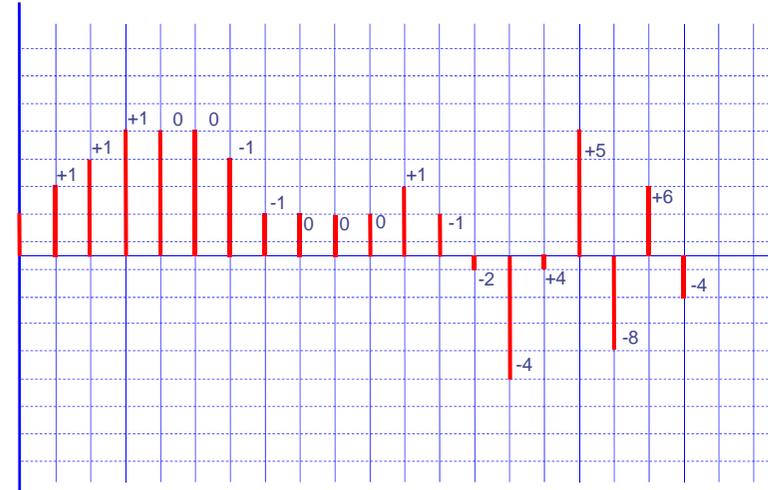
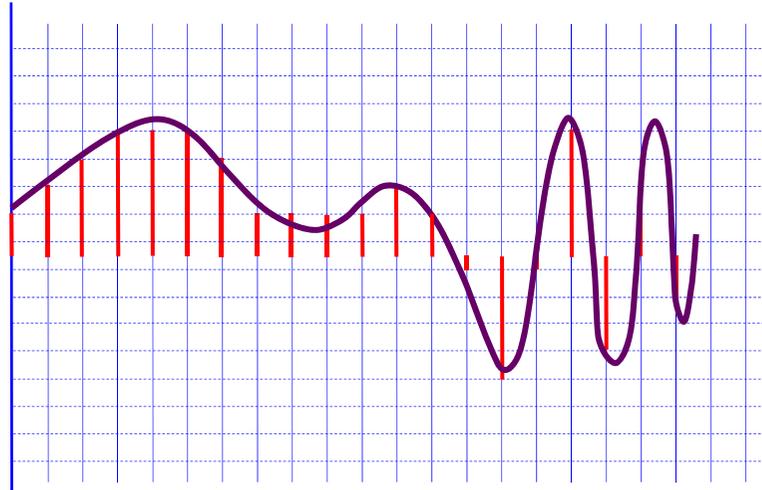
- Basata su sole considerazioni numeriche
- Run-Length Encoding, RLE
 - Codifica una stringa di bit *contando* le sequenze di bit 0 e 1 che si susseguono
 - Esempio: i fax sono lunghe sequenze di 0 e 1 che codificano il contenuto di una pagina in bianco e nero. La codifica RLE è utilizzata per specificare la lunghezza della prima sequenza di 0, seguita dalla sequenza di 1, ecc.
 - la compressione RLE è *lossless*, cioè senza perdita – la rappresentazione originale può essere ricostruita perfettamente
 - Per le immagini, memorizzazione di zone di colore uniforme tramite stringhe speciali
 - Buona per immagini che hanno un numero limitato di colori
 - Se l'immagine è "complessa" la memorizzazione compressa occupa più spazio dell'originale!

Audio: Codifica della differenza (DPCM)



- Ridondanze temporali tra campioni (è molto probabile che campioni "vicini" nel tempo siano abbastanza simili, cioè differiscano di poco)
- La differenza tra due campioni a n bit si può rappresentare con meno bit di n
- Si memorizza la differenza rispetto al campione precedente

Audio: Codifica della differenza (DPCM)



In questa zona bastano
2 bit per codificare le
differenze

In questa
zona servono
4 bit

- Differenze elevate (alte frequenze) non si possono rappresentare con un numero piccolo di bit
- Gli errori introdotti porterebbero a distorsioni sulle alte frequenze
- La compressione DPCM può essere *con perdita*

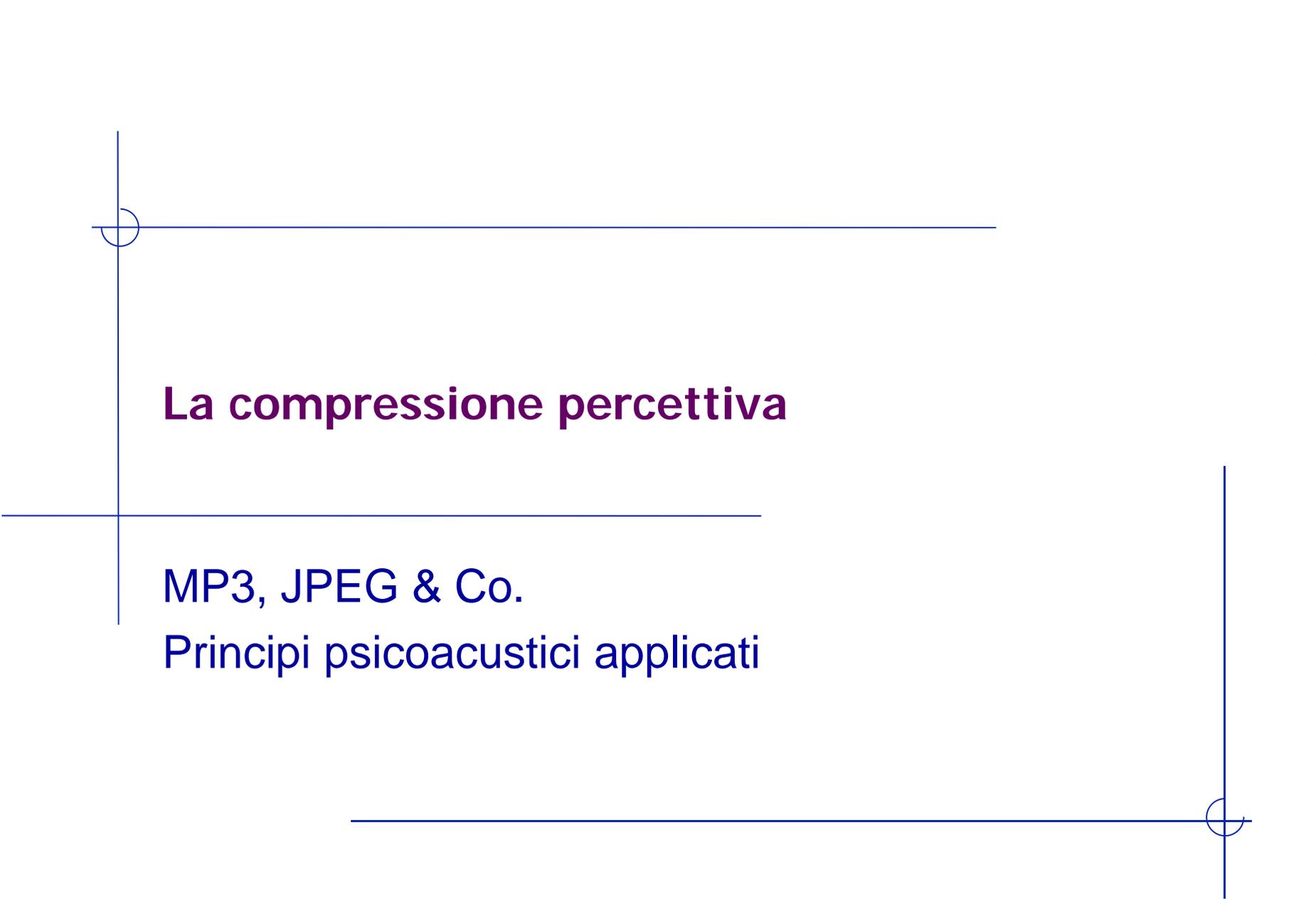
Immagini: Palette, Tavolozza, Look-Up-Table

Non sempre in un'immagine sono presenti tutti i colori rappresentabili con 24 bit (che sono $2^{24} = 16.777.216$)

- Si conta il numero di colori effettivamente presenti
- Si stabilisce il numero di bit necessario a rappresentarli (ricorda: $2^n = \text{nr di rappresentazioni possibili con } n \text{ bit}$)
- Si crea una tabella (codice) che associa ai colori presenti le relative sequenze di bit
- Può essere senza perdita, a prezzo di allegare all'immagine la palette utilizzata
- Non sempre è percorribile (immagini con elevato numero di colori presenti)

Esempio con 256 colori (quindi bastano 8 bit, $2^8 = 256$):

0	R_0	G_0	B_0
1	R_1	G_1	B_1
2	R_2	G_2	B_2
...
...
255	R_{255}	G_{255}	B_{255}

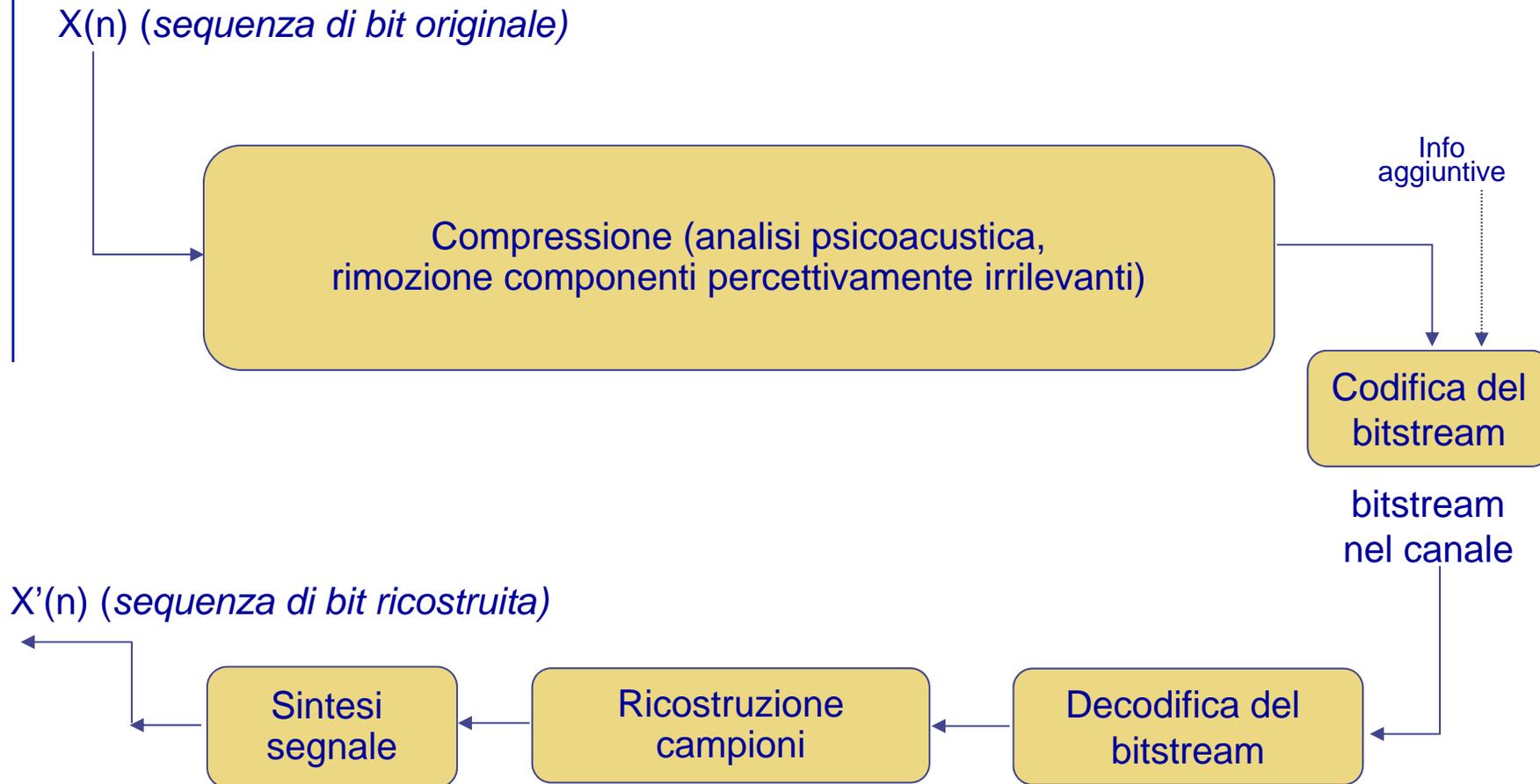
The slide features a white background with several thin blue lines. A vertical line on the left and a horizontal line at the top intersect at a small circle. Another horizontal line is positioned below the top one. A vertical line on the right and a horizontal line at the bottom intersect at another small circle. The text is centered between these lines.

La compressione percettiva

MP3, JPEG & Co.

Principi psicoacustici applicati

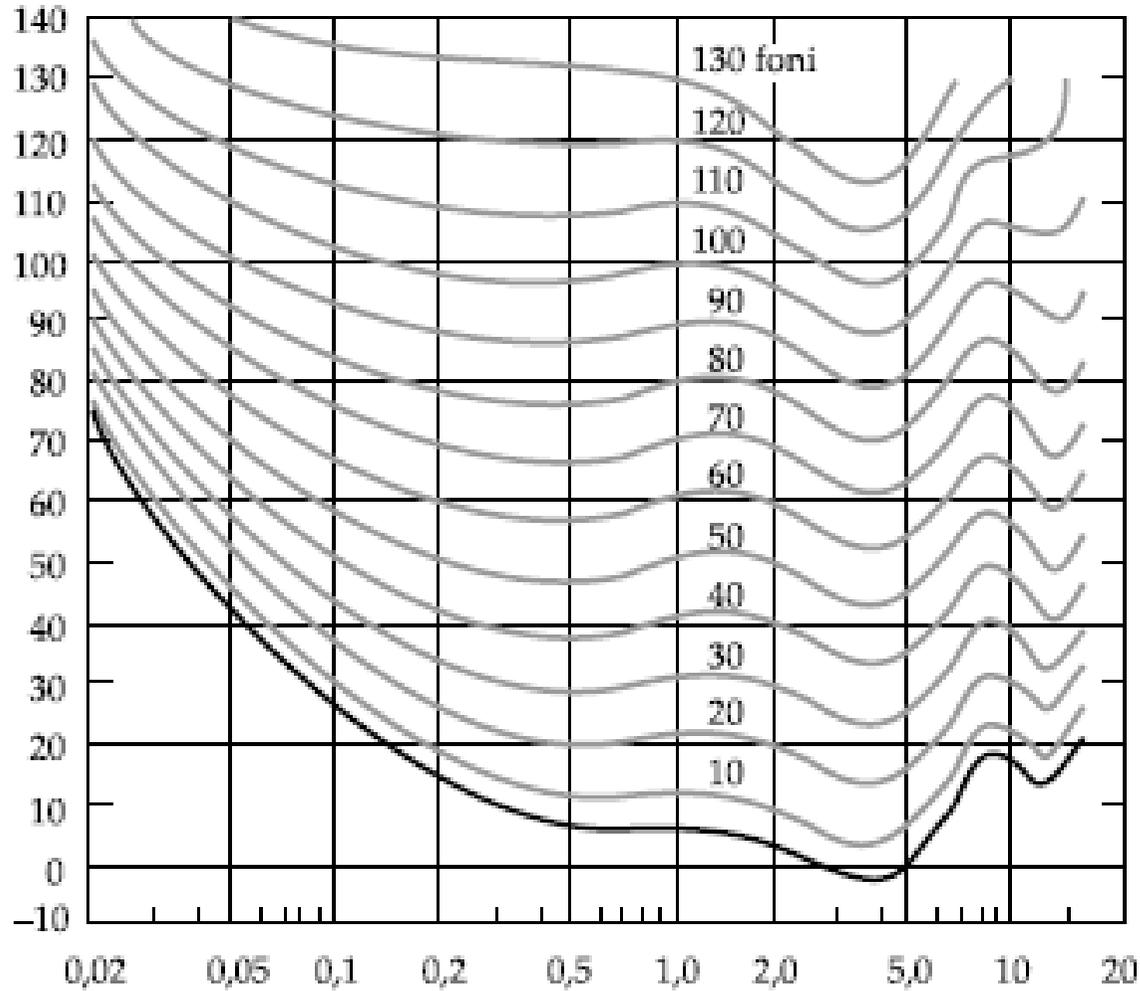
Schema generale



$X'(n) \neq X(n)$, ma, nei limiti della larghezza di banda disponibile, deve "essere percepito" $X(n)$



Audio: Soglia assoluta dell'udito





Audio: MPEG (Motion Picture Experts Group)

Gruppo di lavoro di ISO/IEC per lo sviluppo di standard internazionali per:

- compressione, decompressione, elaborazione, e rappresentazione codificata di
 - Video
 - Audio
 - combinazioni A/V



Audio: Alcuni standard MPEG

- MPEG-1: memorizzazione/recupero video/audio (11/92)
- MPEG-2, standard per la TV digitale (11/94)
- MPEG-4
 - v. 1, standard applicazioni multimediali (10/98)
 - v. 2, standard audio/video HDTV (12/99)
- MPEG-7: standard per la rappresentazione dei contenuti
 - ricerca, filtraggio, gestione di info multimediale
 - rilasciato a luglio 2001



Audio: MPEG-1/2 Layer III (MP3)

- 1987 progetto Eureka (Digital Audio Broadcasting)
- lavoro su audio percettivo in collaborazione tra Fraunhofer IIS e l'Università di Erlangen
- risultato: algoritmo molto potente ISO-MPEG Audio Layer-III

- <http://it.wikipedia.org/wiki/Mp3>



Audio: Lo schema di compressione MPEG-1

- **Struttura di base comune su più layer**
 - calcolo delle sottobande
 - calcolo del modello psicoacustico
- **3 Layer**
 - Layer I: semplice, poco efficace
 - Layer III: complesso, efficace
- **Scelta del layer - rapporto qualità/compressione**



Audio: I tre livelli di compressione

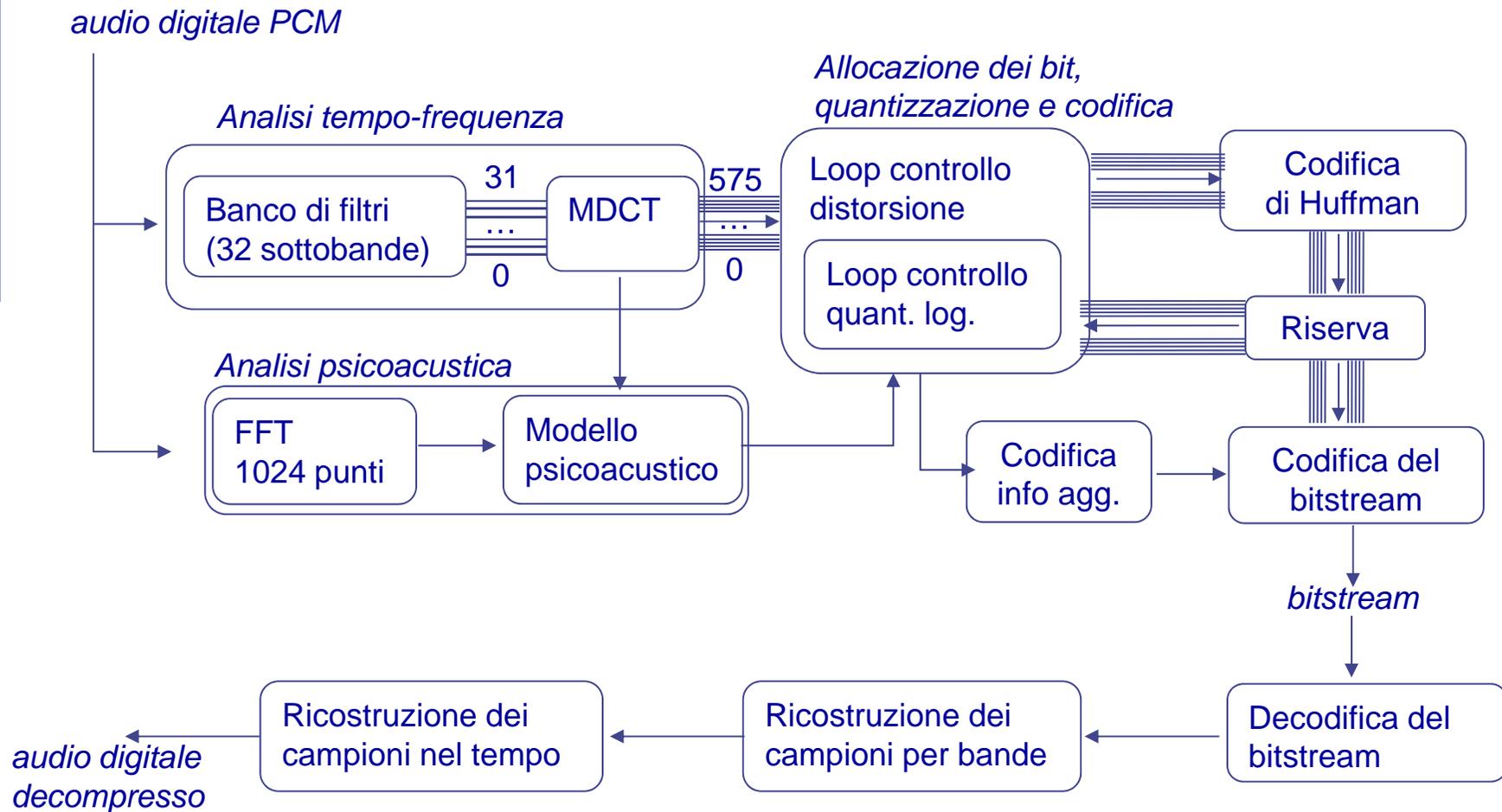
- Layer I
 - il più semplice (bitrate oltre 128 kbps a canale)
 - DCC di Philips usa la compressione di Layer I a 192
- Layer II
 - complessità media (bitrate circa 128 kbps a canale)
 - applicazioni in DAB
- Layer III
 - il più complesso, migliore qualità
 - 64 kbps adatto per trasmissione audio su ISDN



Audio: Modalità di compressione

- Più frequenze di campionamento (32/44.1/48)
- Bitstream compresso supporta mono, dual mono, stereo, joint stereo
- Bit rate da 32 a 224 kbps (compressione da 24 a 2,7 volte) – tassi fissi e variabili
- Supporta il controllo e la correzione degli errori
- Informazioni supplementari

Audio: MPEG 1 - Layer III (complessità)





Audio: Tipiche riduzioni di MPEG

Riferimento standard (CD):

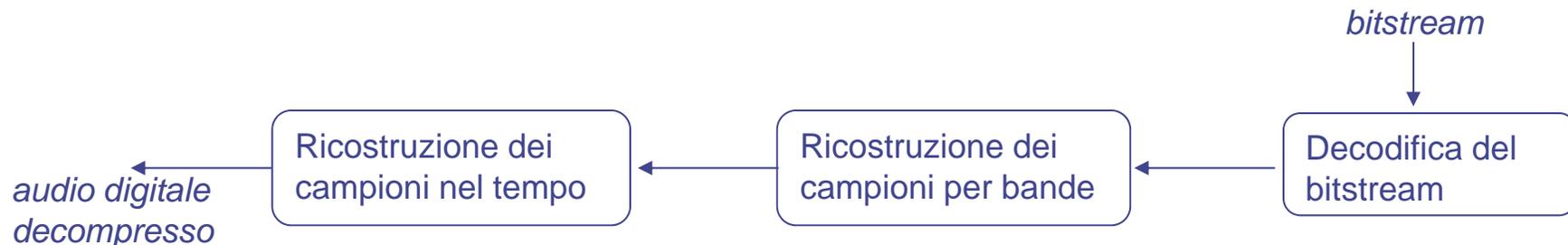
$$16 \text{ bit} * 2 \text{ canali} * 44100 \text{ sr} = 1.411 \text{ kbps}$$

- 1:4 con Layer I
 - corrisponde a 384 kbps per un segnale stereo
- 1:6...1:8 con Layer II
 - corrisponde a 256..192 kbps per un segnale stereo
- 1:10...1:12 con Layer III (MP3)
 - corrisponde a 128..112 kbps per un segnale stereo

sempre mantenendo *percettivamente* la qualità audio CD

Audio: Decodifica

- Sintetizza un segnale a partire dalle componenti spettrali codificate
- Non si ha più lo stesso segnale!!!
- Tutto dipende dal bit-rate



Audio: Esempi di bitrate

- *bit-rate* = numero medio di bit consumati da un secondo di dati audio (kbps)
- bit-rate per il CD = 1411.2 kbps
- bit-rate per MP3 per qualità CD = 128 kbps



Audio: Le performance di MP3

- qualità telefonica: 96:1 (2.5 kHz / mono / 8 kbps)
- meglio di AM radio: 24:1 (7.5 kHz / mono / 32 kbps)
- simile a FM radio: 26...24:1 (11 kHz / stereo / 56...64 kbps)
- quasi-CD: 16:1 (15 kHz / stereo / 96 kbps)
- CD: 14..12:1 (>15 kHz / stereo / 112..128kbps)
 - prende approx. 1Mb/minute di spazio hard-disk
- Oltre: 8...4:1 per la musica acustica

Audio: MP3 Codec

- programma del tipo di una libreria di sistema (collezione di funzioni)
- vengono lanciate dai programmi di frontend
- codec in distribuzione con frontend MP3



Audio: L'encoder migliore

- Non ha senso chiedersi quale sia l'encoder migliore: la risposta dipende dalle esigenze
 - encoder veloci/lenti (> velocità, < fedeltà audio)
 - confrontare mp3 ottenuti da encoder diversi a parità di bit-rate
- Consiglio pratico: creare MP3 con basso (ma adeguato) bitrate con encoder 'lenti'
- Le cose potrebbero cambiare se c'è l'esigenza del "tempo reale" (codifica in diretta)



Audio: I Frontend MP3

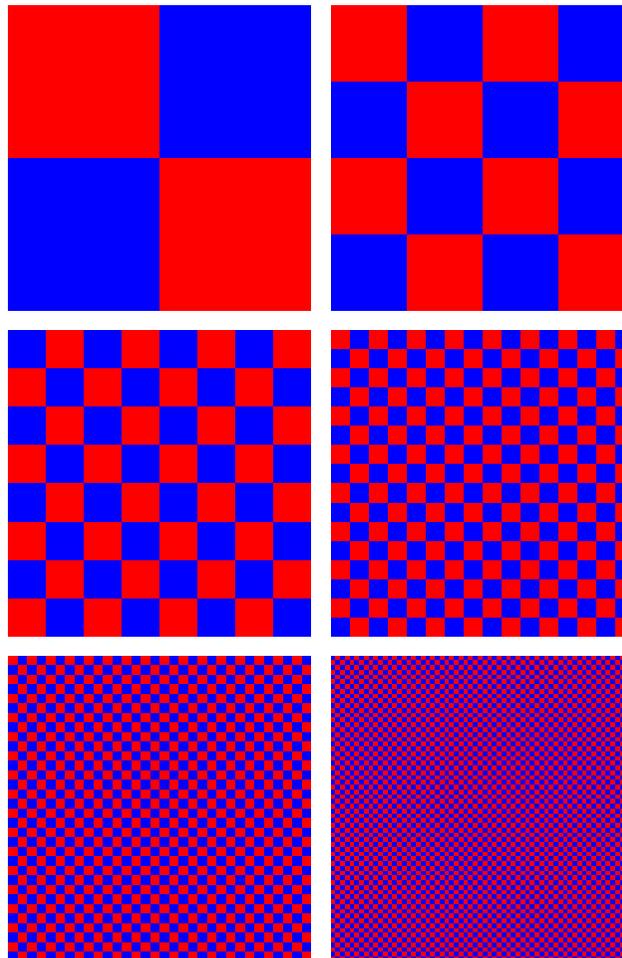
- interfaccia ai codec
- Alcuni frontend implementano funzioni per “normalizzare” il volume, o realizzano ID TAGS
- IDTAG è informazione (testuale, in genere) nel file di layer III (autore, titolo, etc ...)
- Esempio: MP3Editor, Audacity, Nero WaveEditor, ...
- Player MP3
 - suona mentre decomprime
 - i dati audio (campioni) vengono inviati alla scheda audio per la conversione D/A
 - Esempi: winamp, windows media player, videolan, ecc.

Immagini: Dithering

- Il dithering viene usato nelle immagini digitali per creare l'*illusione* di una profondità di colore (*nr di bit per pixel*) maggiore di quella disponibile
- In un'immagine sottoposta a dithering, i colori non disponibili vengono approssimati dalla distribuzione dei pixel colorati con le tinte disponibili.
- L'occhio umano percepisce la diffusione come un amalgama dei colori
- Si può usare come forma particolare di compressione percettiva (con perdita)



Immagini: Dithering in rosso e blu





Immagini: Dithering a 32 colori



Immagini: Dithering

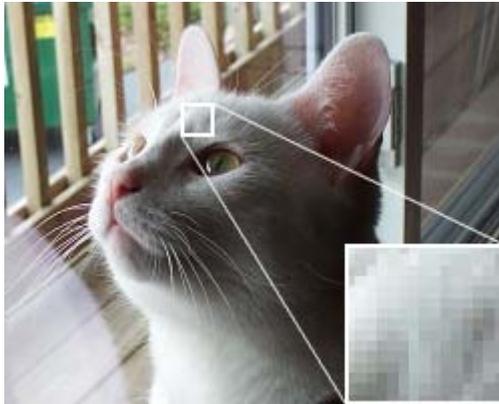


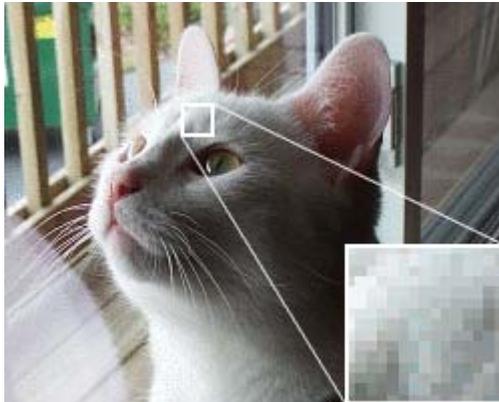
Immagine originale



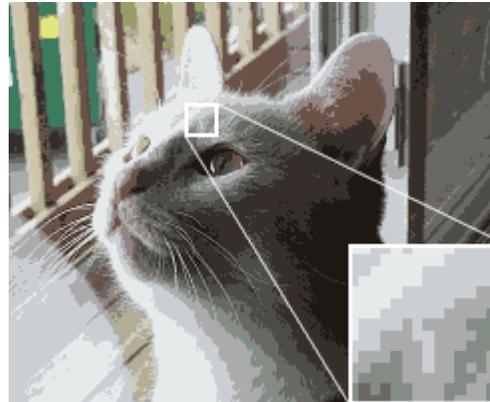
Tavolozza std 8 bit – no dithering



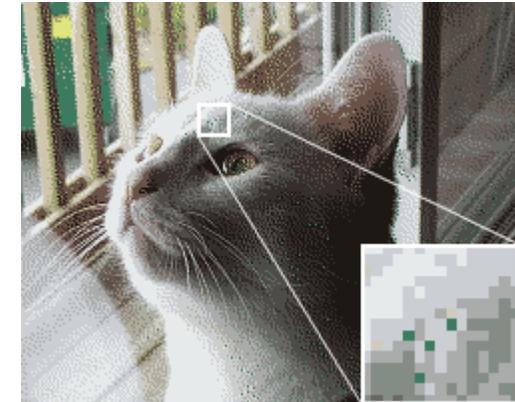
Tavolozza std 8 bit– dithering



Tavolozza opt 8 bit – no dithering



Tavolozza opt 4 bit – no dithering



Tavolozza opt 4 bit – dithering

Fonte: <http://it.wikipedia.org/wiki/Dithering>



Immagini: Compressione JPEG

- Osservazione 1: disponibilità di immagini con un elevatissimo numero di colori e di computer in grado di elaborarle
- Osservazione 2: I nostri occhi non sono sensibili alle piccole variazioni di tonalità, ma sono sensibili alle piccole variazioni di luminosità
- La compressione JPEG provvede a salvare le componenti a bassa frequenza (zone di colore uniformi) e una parte delle componenti ad alta frequenza (particolari dell'immagine)
- fornisce una descrizione meno accurata delle sfumature di colore
 - Rapporto di compressione variabile sulla base della qualità residua desiderata
 - Compressione fino a 20:1 senza variazioni percepibili ai nostri occhi
 - È una compressione *percettiva e con perdita (sempre)*



Immagini: Compressione JPEG



Un dettaglio da un'immagine compressa mediante JPEG. (a) compressione 14:1; (b) compressione 140:1.



Immagini: Compressione JPEG



Qualità 10% - 3,2 KB



Qualità 50% - 6,7 KB



Qualità 90% - 30,2 KB



Qualità 100% - 87,7 KB

Fonte: http://it.wikipedia.org/wiki/Joint_Photoshopping_Experts_Group



Sequenze A/V: Compressione MPEG

- La stessa idea alla base del JPEG, ma applicata ai filmati
- Osservazione: i filmati sono costituiti in genere da lunghe sequenze di immagini consecutive molto simili
 - Il primo frame si comprime con JPEG
 - I successivi si codificano per differenza dal precedente
 - la compressione MPEG memorizza solo "la differenza" tra fotogrammi consecutivi
 - ciò permette livelli di compressione elevatissimi (finché non si avvera uno "stacco" di scena)
- Per la traccia audio: MP3