

# Codifica Video

*Fondamenti di Telecomunicazioni*

*Anno Accademico 2009/2010*

# Contenuti

- Caratteristiche del segnale video
- Codifica video (MPEG-2), standard DVB-T
- Codifica video (MPEG-4), standard 3G

# Perché comprimere...

Si consideri un segnale televisivo analogico standard (PAL) che transita attraverso un sistema di conversione analogico/digitale (campionatore + quantizzatore a 65536 livelli). Si determini l'occupazione di memoria (in bits) di uno stream della durata di un minuto memorizzato senza compressione dei dati.

La banda occupata dal segnale televisivo considerato risulta  $B = 6\text{MHz}$ . Di conseguenza, sceglieremo una frequenza di campionamento almeno pari al doppio della banda occupata, ovvero  $f_c = 12\text{MHz}$ .

Inoltre, il quantizzatore a 65536 livelli corrisponde all'utilizzo di un numero di bit pari a 16 essendo  $\log_2(65536) = 16$ . Si ottiene quindi un bit-rate  $R$  al secondo pari a:

$$R = 16 * f_c = 192\text{Mbit/s}$$

In un minuto sono contenuti 60 secondi, quindi l'occupazione di memoria richiesta per uno stream non compresso della durata di un minuto risulta pari a:

$$X = R * 60\text{s} = 1.44 \text{ GByte}$$

# La compressione

- Ridurre l' occupazione di memoria fisica dei filmati digitali
- Ridurre di conseguenza la banda richiesta per la trasmissione dei filmati multimediali
- Rendere possibile uno streaming di qualità su internet o sulle reti radiomobili

La compressione: soluzione per trasmettere materiali multimediali sulle numerose reti radiomobili presenti e in generale sulle reti a banda limitata

# Le tecniche

- Tecniche di compressione senza perdita di informazione: particolari codifiche riescono a comprimere i dati relativi al filmato senza decurtarne il contenuto informativo
- Tecniche di compressione con perdita: codifiche che producono una riduzione *irreversibile* del contenuto informativo del filmato
- Tipicamente i due tipi di codifica vengono usate contemporaneamente in modo da ottenere significativi rapporti di compressione con una perdita di qualità relativamente bassa

# Gli Standard

## Video

MPEG-1: sistemi a banda larga, video cd (obsoleto)

MPEG-2: sistemi a banda larga, DVD e digitale terrestre (in uso)

MPEG-4: sistemi a banda stretta, telefoni UMTS (in uso)

## Audio

Telefonia fissa (voce: codifiche nel tempo, PCM)

Telefonia mobile (voce: codifiche per modelli, LPC)

Musica (codifiche in frequenza, MPEG o AAC)

# Gruppo MPEG-VIDEO

- Le direzioni di lavoro originariamente seguite dal gruppo MPEG sono 3:
- codifica digitale del segnale fino a 1,5 Mbit/s
- codifica a 10 Mbit/s
- codifica a 40 Mbit/s.

Queste tre diverse direzioni di lavoro sono battezzate: MPEG-1, MPEG-2, MPEG-3, ma l'attività sul filone MPEG-3 viene abbandonata, quando diventa chiaro che tutti gli obiettivi di MPEG-3 possono essere raggiunti attraverso un uso appropriato di MPEG-2.

# Standard Video

- **Mpeg-1** (1992)
- **Mpeg-2** (1994)
- **Mpeg-4** (1998)
- **Mpeg-7** (2001)
- **Mpeg-21** (*in fase di completamento*)



# MPEG 1

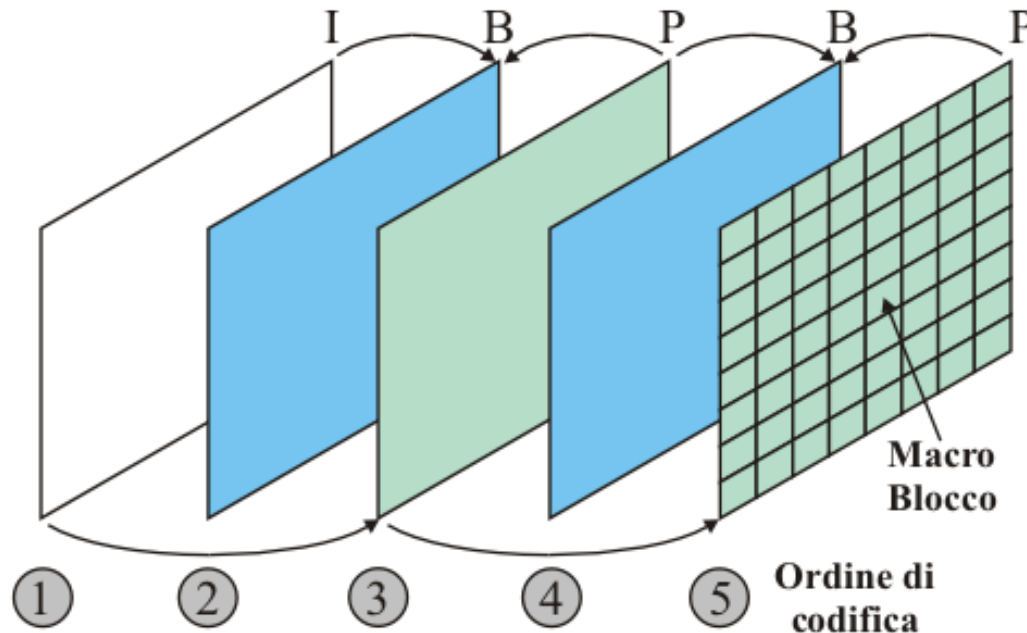
- Composto da 5 parti, le prime 3 parti sono diventate International Standard nel 1993, la parte 4 lo è diventata nel 1994, la parte 5 nel 1995.
- Parte 1: *System*: descrive come sincronizzare diversi flussi audio e video e come trasportarli su canali digitali o supporti di massa.
- Parte 2 - *Video*: descrive la sintassi (header ed elementi del bitstream) e la semantica (il significato dei bit) del bitstream video. La sequenza di immagini è suddivisa in una serie annidata di livelli (sequence, picture, slice, macroblock, block, coefficienti DCT ).
- Parte 3 - *Audio*: descrive la sintassi e la semantica per 3 classi di metodi di compressione chiamati layer I, II e III.
- Parte 4 - *Conformità*: definisce l'insieme dei test di conformità sui bitstream e sui decodificatori.
- Parte 5 - *Simulazione software*: contiene un esempio in linguaggio ANSI C di un codificatore e di un decodificatore software conformi allo standard e relativi alle parti audio, video e system.

# MPEG 2

- Organizzato come il suo predecessore per quanto riguarda le prime 5 parti a cui se ne aggiungono altre 3:
- Parte 6 – *Digital Storage Medium Command and Control (DSM-CC)*: fornisce una sintassi per controllare l'esecuzione e l'accesso casuale come avviene nei videoregistratori (fermo immagine, avanti veloce, riavvolgimento).
- Parte 7 – *Non-backward compatible audio*: descrive nuove estensioni della codifica audio che, a differenza della Parte 3, non sono compatibili con MPEG-1
- Parte 8 – *10-bit video extension*: è una parte molto giovane dello standard che riguarda la codifica su 10 bit dei coefficienti DCT per una migliore qualità da usarsi in ambito professionale e per l'archiviazione.
- Si è espanso il concetto di qualità scalabile tipico di MPEG-1, aggiungendo poi altre caratteristiche quali la trasmissione del flusso multimediale su reti a larga banda, assicurando una buona robustezza nei confronti degli errori della rete, il trasporto parallelo di molteplici canali audio, le funzioni di protezione e di controllo di accesso al flusso...

# Codifica video MPEG-2

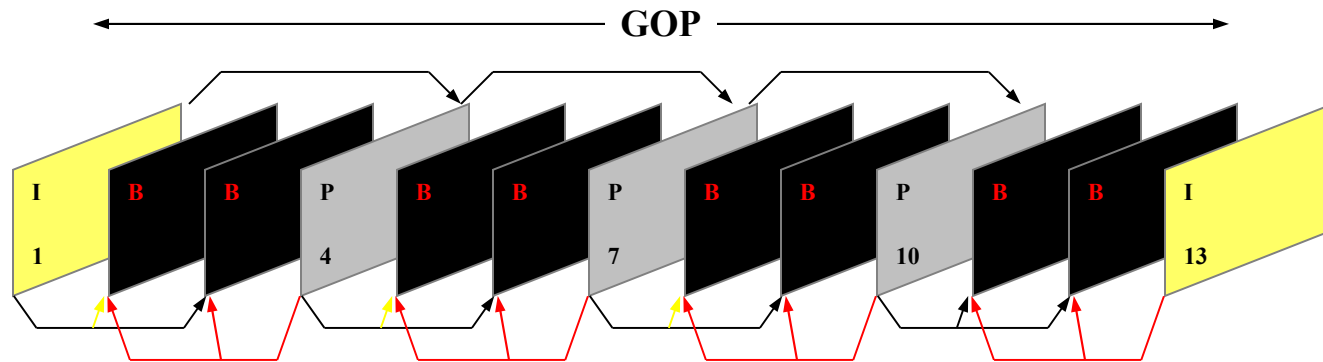
- Codifica diversa a seconda dei frame
  - *Intra-frame* indipendente dagli altri frame (**Frame I**)
  - *Inter-frame* con riferimento alle immagini precedenti (**Frame P**)
  - Codifica Interframe a doppia predizione una nel passato e una nel futuro (**Frame B**)



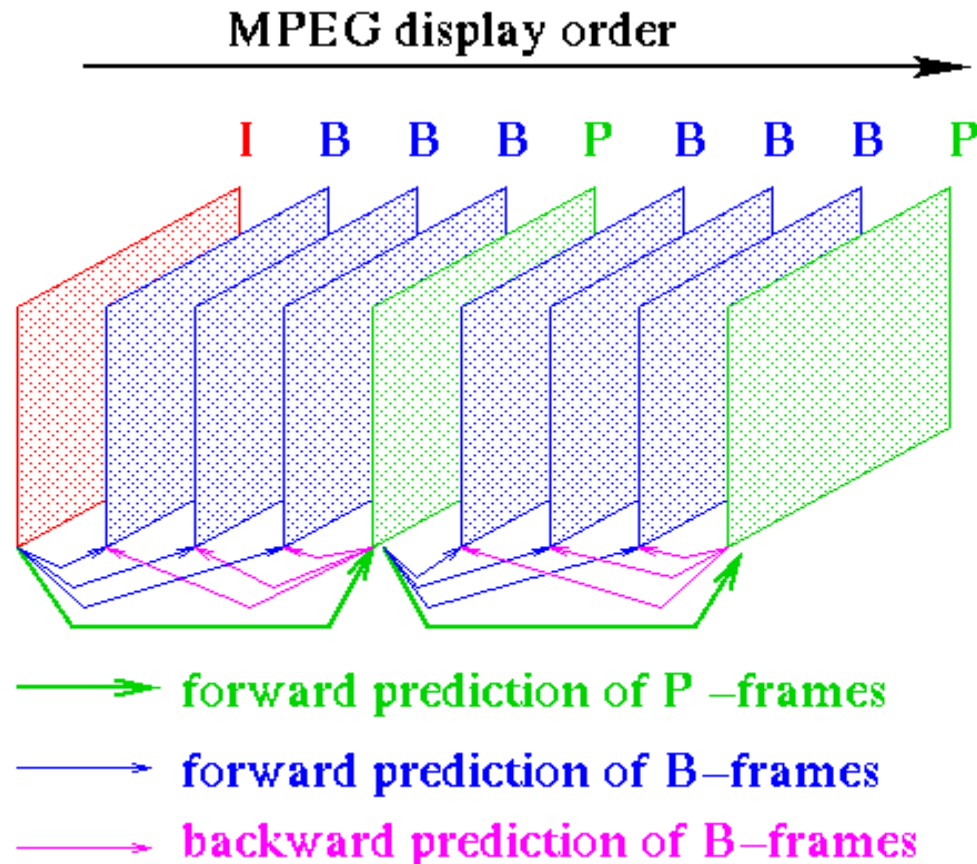
# Codifica video

- Codifica intra (I), predizione temporale (P), predizione bidirezionale (B). In MPEG-2 ci riferiamo alle trame, ovvero i frame.

- I - trama: sono codificate senza nessun riferimento alle altre trame.
- P - trama: sono codificate riferendosi al precedente I o P.
- B - trama: sono codificate riferendosi sia alle trame precedenti che a quelle successive.

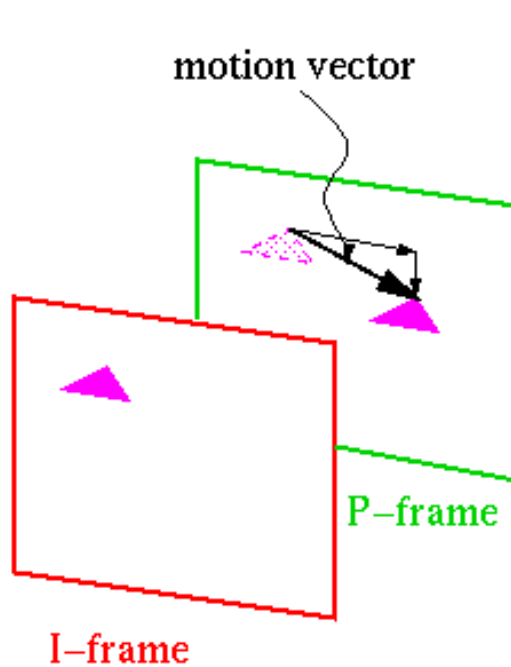


# Codifica video



- L'ordine con cui le trame video vengono trasferite è: I P B B B P B B B. Il decoder deve semplicemente riordinare le trame secondo l'ordine di display.

# La predizione



- La predizione consiste nel fornire un vettore di movimento (*motion vector*) che dichiari come gli oggetti si sono spostati dal quadro I al quadro P. Il motion vector è parte dello stream MPEG ed è suddiviso in una parte orizzontale ed una verticale.
- Il vettore di movimento viene ricercato in ogni macroblocco della componente di luminanza del frame considerato. Avremo tanti motion vectors quanti sono i macroblocchi che compongono il frame.
- Un valore positivo corrisponde ad un movimento a destra o verso il basso; uno negativo ad uno spostamento a sinistra o in alto rispettivamente.

# Flusso video MPEG 2

- Il flusso video MPEG-2 è organizzato con una sintassi stratificata. In una struttura gerarchica "dall'alto in basso", la video sequenza è partizionata in gruppi multipli di immagini (GOPs), che rappresentano degli insiemi di video-trame contigui nell'ordine di trasmissione. Lo strato successivo è costituito da una trama singola, composta da più porzioni.
- Quindi, ogni porzione contiene uno o più macro blocchi, consistenti di quattro blocchi di luminanza (Y) e di due blocchi di cromaticanza (U,V). infine, il blocco è l'unità di base di codifica di dimensione 8x8 pixel.

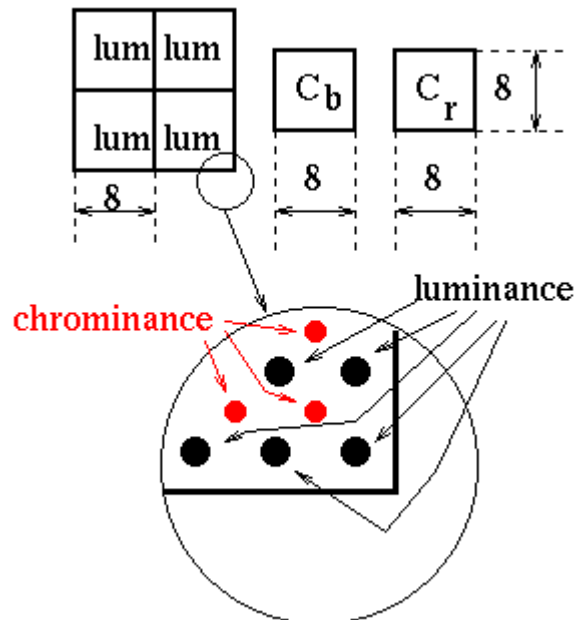
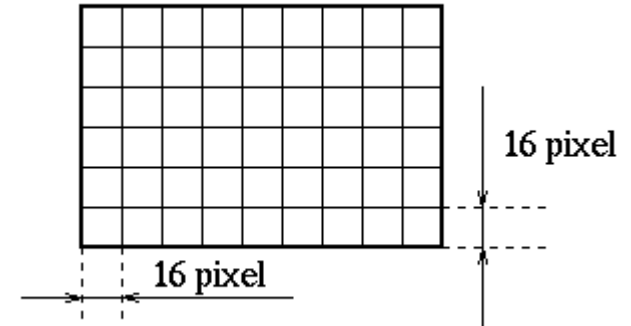
# Flusso video MPEG 2

- Per ottenere un alto rapporto di compressione, si devono eliminare sia le informazioni dovute alle ridondanze spaziali che a quelle temporali. La ridondanza spaziale viene ridotta dall'utilizzo di un sottocampionamento delle componenti di cromaticità  $U$ ,  $V$  (o  $C_B$ ,  $C_R$ ), in accordo con la sensibilità dell'apparato visivo umano.
- Successivamente, viene utilizzata la trasformata coseno discreta sui blocchi con componenti  $Y$  e  $U, V$ . I coefficienti DCT sono quantizzati e infine codificati utilizzando un codice a lunghezza variabile. La ridondanza temporale viene ridotta da una predizione temporale di alcune trame derivate da altre trame moto-compensate. L'errore di predizione viene quindi codificato.



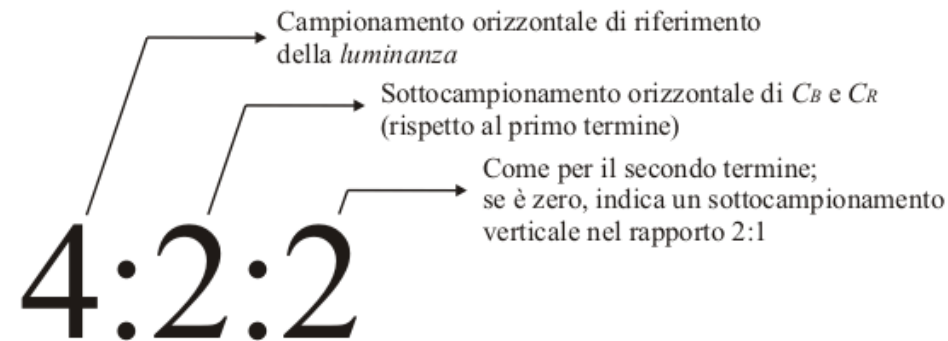
# Macro-blocchi

- Ogni frame è suddiviso in macro-blocchi 16x16. Ogni macro-blocco possiede il proprio motion vector.



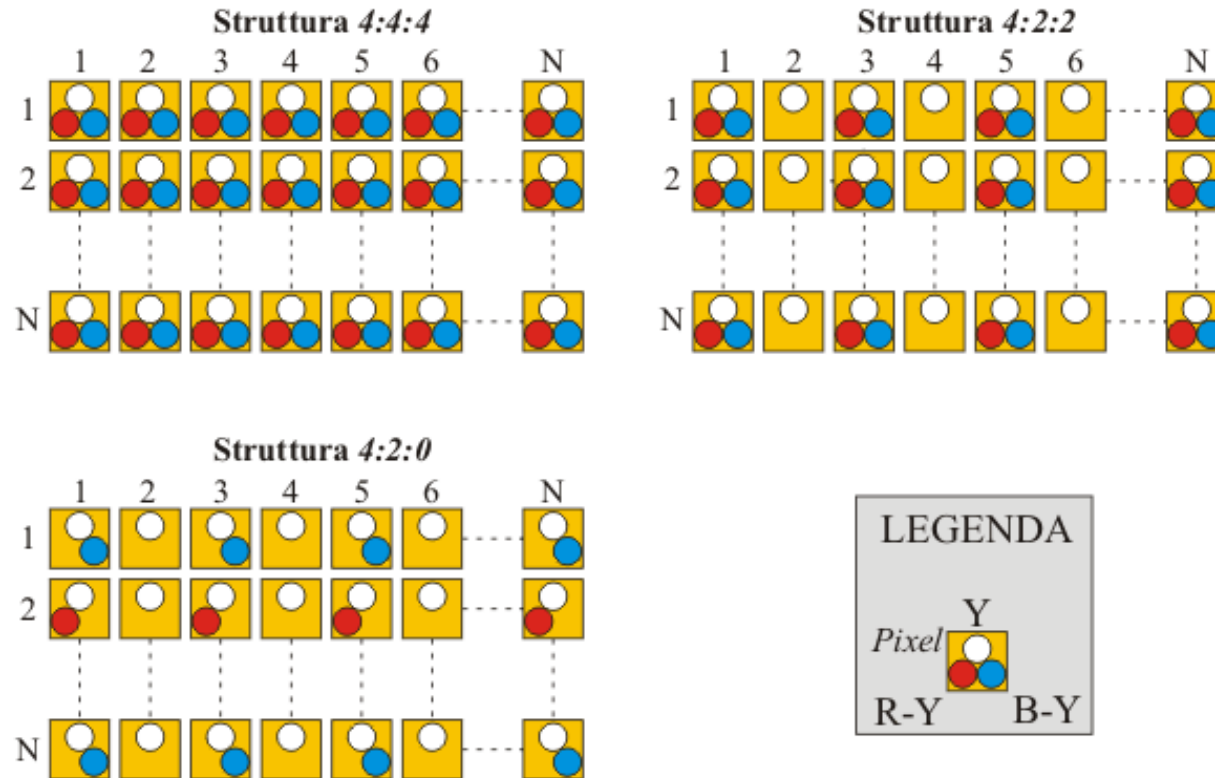
- Ogni macro-blocco è costituito da 4 blocchi di luminanza e da 2 di crominanza. Date le proprietà visive dell'occhio umano, si può applicare un forte sotto-campionamento alle componenti di colore.

# Sottocampionamento



- Il primo termine indica il campionamento orizzontale della componente di luminanza. Il secondo termine specifica il sottocampionamento orizzontale delle componenti  $C_B$  e  $C_R$  rispetto al campionamento operato sulla luminanza. Il terzo termine è utilizzato per indicare il sottocampionamento verticale.

# Sottocampionamento



- La struttura 4:2:0 è generalmente adoperata nel JPEG, JFIF, H.261 e nel MPEG-1, mentre la struttura 4:2:2 è sfruttata nella codifica video MPEG-2.

# DCT



- Ogni blocchetto 8x8 è poi codificato attraverso la Discrete Cosine Transform (DCT).

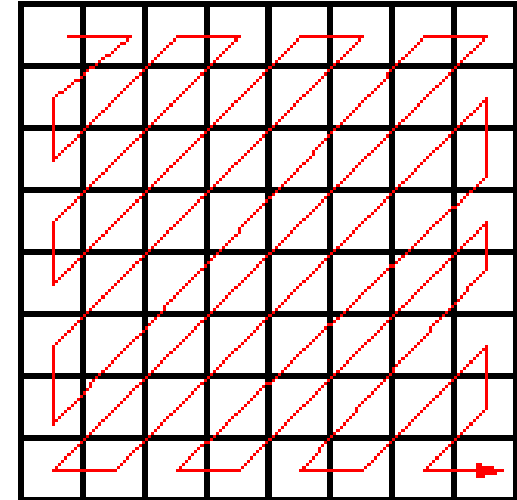
$$F(u, v) = \frac{C_u}{2} \frac{C_v}{2} \sum_{y=0}^7 \sum_{x=0}^7 f(x, y) \cos \left[ \frac{(2x+1)u\pi}{16} \right] \cos \left[ \frac{(2y+1)v\pi}{16} \right]$$

$$C_u = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } u = 0, \\ 1 & \text{if } u > 0 \end{cases}; C_v = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } v = 0, \\ 1 & \text{if } v > 0 \end{cases}$$

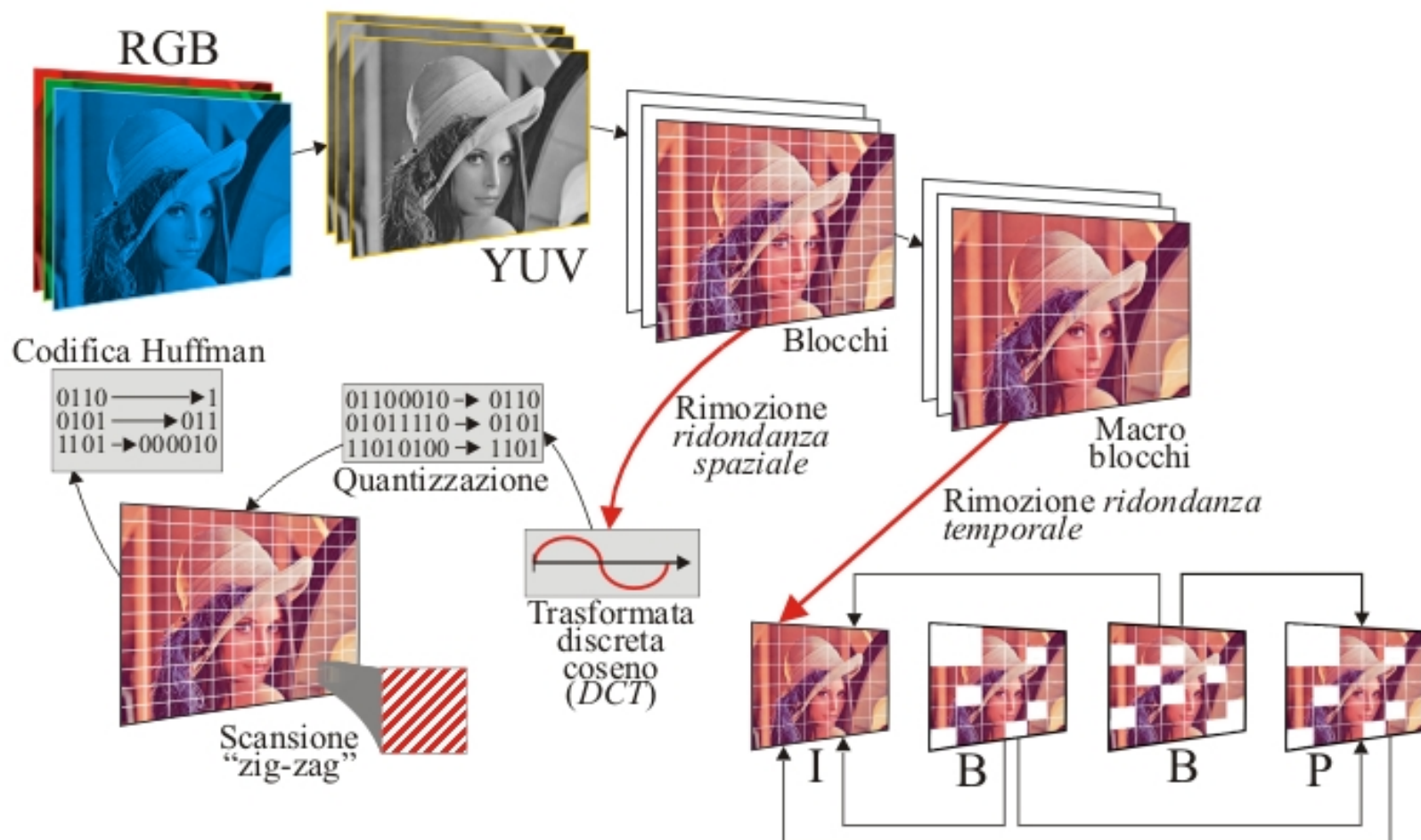
- La DCT può essere vista come un analizzatore armonico: l'input viene suddiviso in 64 segnali base ortogonali, ciascuno rappresentato da una coppia di frequenze spaziali  $(u, v)$ .

# DCT

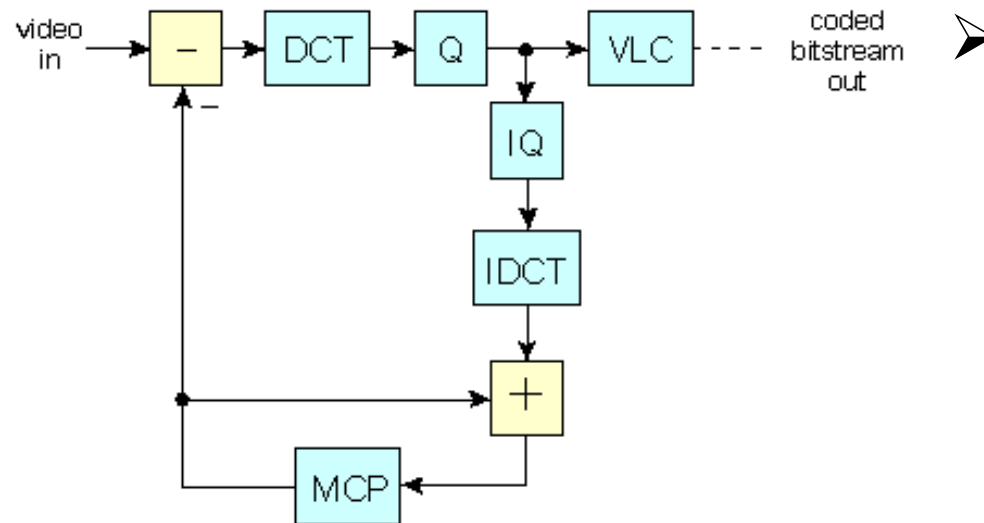
- I coefficienti quantizzati di ogni blocco vengono poi ordinati secondo una scansione a zig-zag.
- Questo ordinamento rende più efficiente la codifica. Infatti, posizionando i coefficienti quantizzati delle basse frequenze prima di quelli alle alte frequenze si aumenta la possibilità di ottenere lunghe sequenze di zeri.



# Flusso video MPEG 2



# Codificatore MPEG2



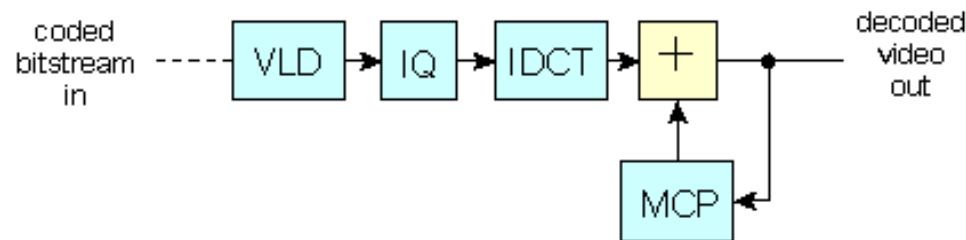
➤ In un codificatore MPEG2 la DCT e la compensazione del moto sono combinate.

(I)DCT = (inverse) discrete cosine transform  
 (I)Q = (inverse) quantisation  
 MCP = motion-compensated prediction

VLC = variable-length coder  
 VLD = variable-length decoder

# De-codificatore MPEG2

- La moto-compensazione si aggiunge al bit-stream decodificato per ottenere il video in uscita.



(I)DCT = (inverse) discrete cosine transform

(I)Q = (inverse) quantisation

MCP = motion-compensated prediction

VLC = variable-length coder

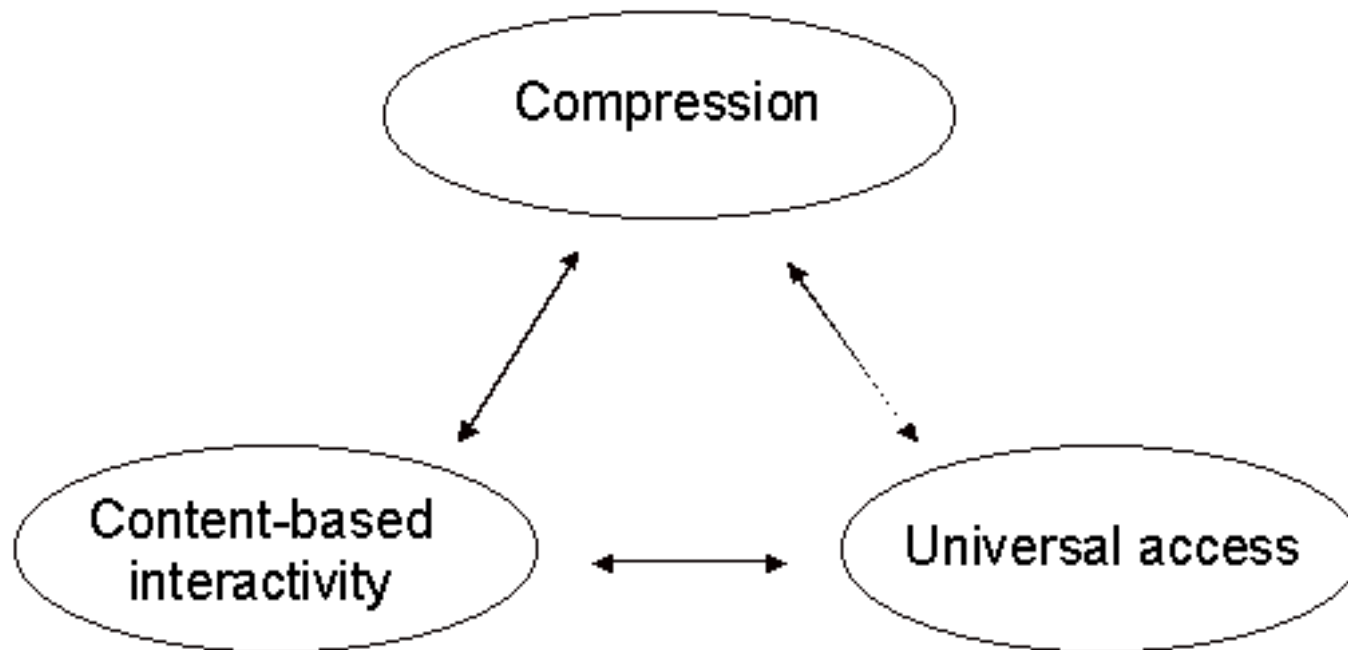
VLD = variable-length decoder



# MPEG-4

- Standard che porta l'audiovisivo digitale sulle reti interattive a basso bit-rate, come Internet o la rete mobile, la cui prima versione è stata promulgata nell'autunno del 1998.
- MPEG-4 è stato scelto come standard di riferimento per la diffusione di contenuti audiovisivi sulle reti mobili di nuova generazione.
- Lo standard MPEG-4 offre una rappresentazione video basata sull'oggetto.

# Funzionalità MPEG 4



# Funzionalità MPEG 4

- *Interattività basata sul contenuto:*
  - Strumenti per l'accesso basato sul contenuto ai dati multimediali.
  - Manipolazione bit-stream in uscita.
  - Codifica ibrida naturale e sintetica.
  - Accesso casuale ai dati.

# Funzionalità MPEG 4

## ➤ *Compressione dati:*

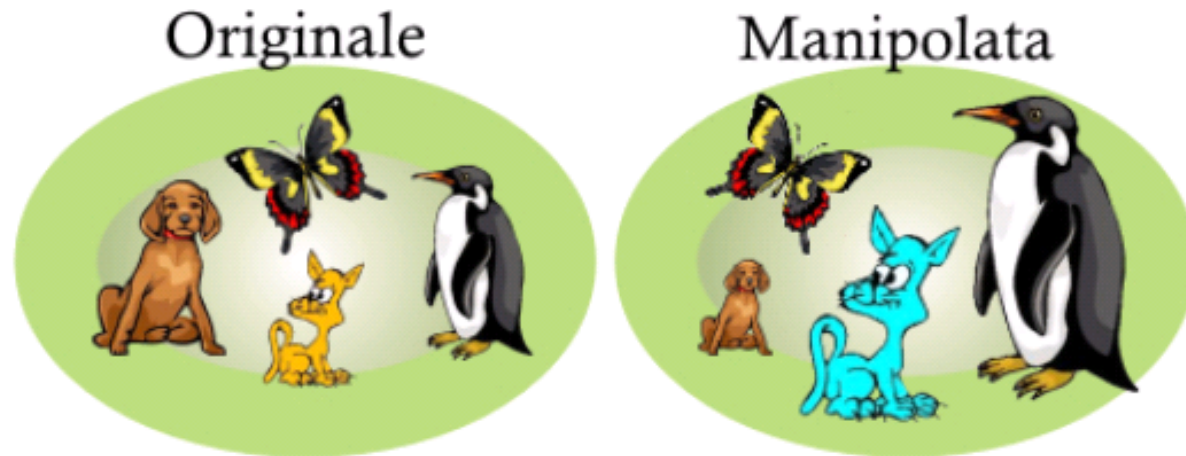
- Metodi efficienti per immagazzinamento e compressione dati.
- Miglioramento efficienza di codifica.
- Codifica di più flussi di dati concorrenti.

# Funzionalità MPEG 4

➤ *Accesso universale:*

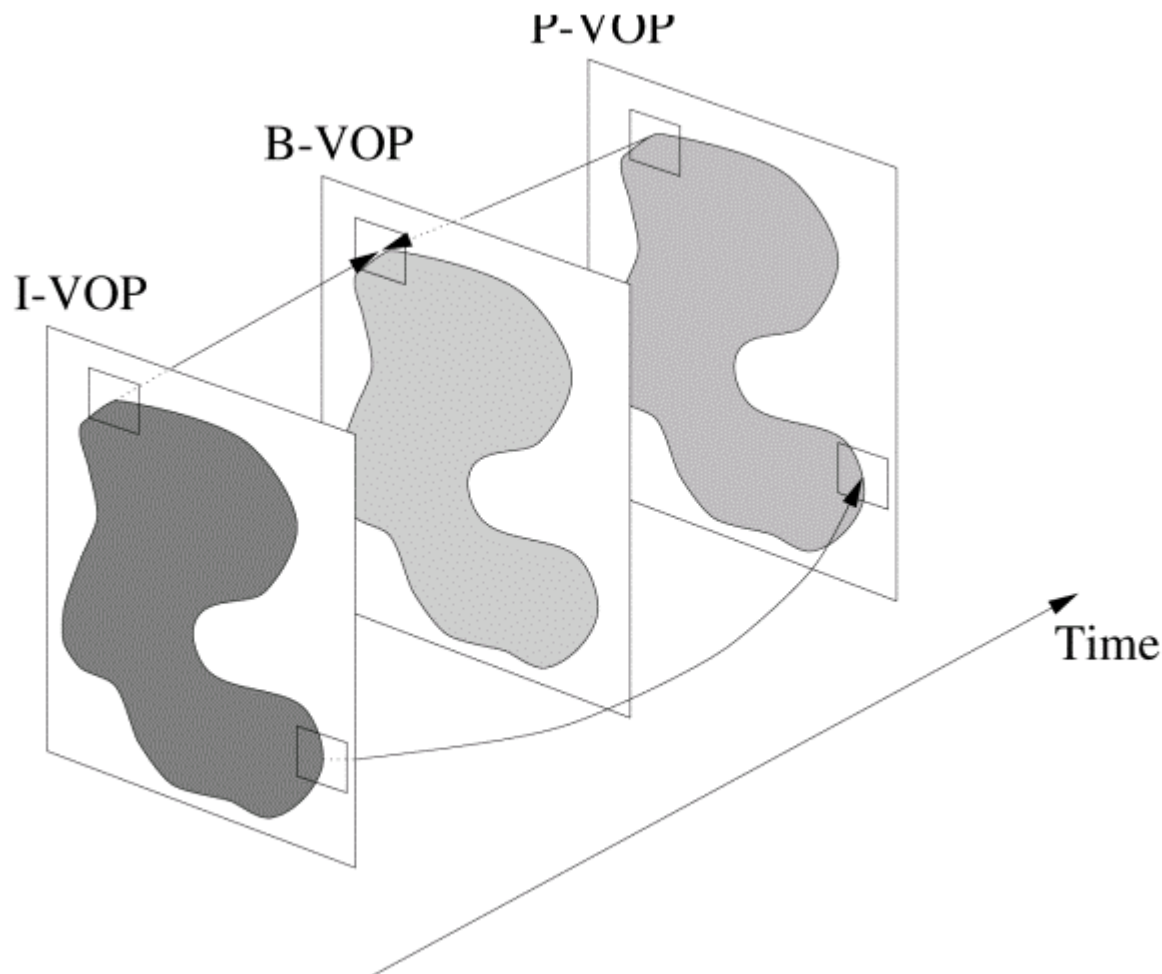
- I dati codificati sono resi accessibili a decodificatori di diversa qualità.
- Robustezza agli errori.
- Scalabilità basata sul contenuto.
- Suddivisione della qualità dell'immagine su più livelli, ognuna con una priorità differente.

# Codifica a oggetti in MPEG 4



- Specificatamente, MPEG-4 considera una scena come fosse composta da video-oggetti (VOs), ognuno descritto dal momento, dalla struttura e dal contorno.
- Ogni oggetto è codificato da una diversa stringa di bit.

# Codifica del movimento



In MPEG-4 si parla di VOP (Video Object Plane), in MPEG-2 di GOP (Group Of Picture).  
Esiste in questo caso un vettore di movimento per ogni oggetto presente nella scena.

# Architettura MPEG4



- Ogni sequenza video è scomposta in VO ognuno con proprietà particolari quali:
  - forma, - movimento, - tessitura (ciò che è all'interno dei confini della forma).
- In questo modo si permette all'utente la manipolazione degli oggetti su 4 livelli:
  - codifica, - multiplazione, - demultiplazione, - decodifica.



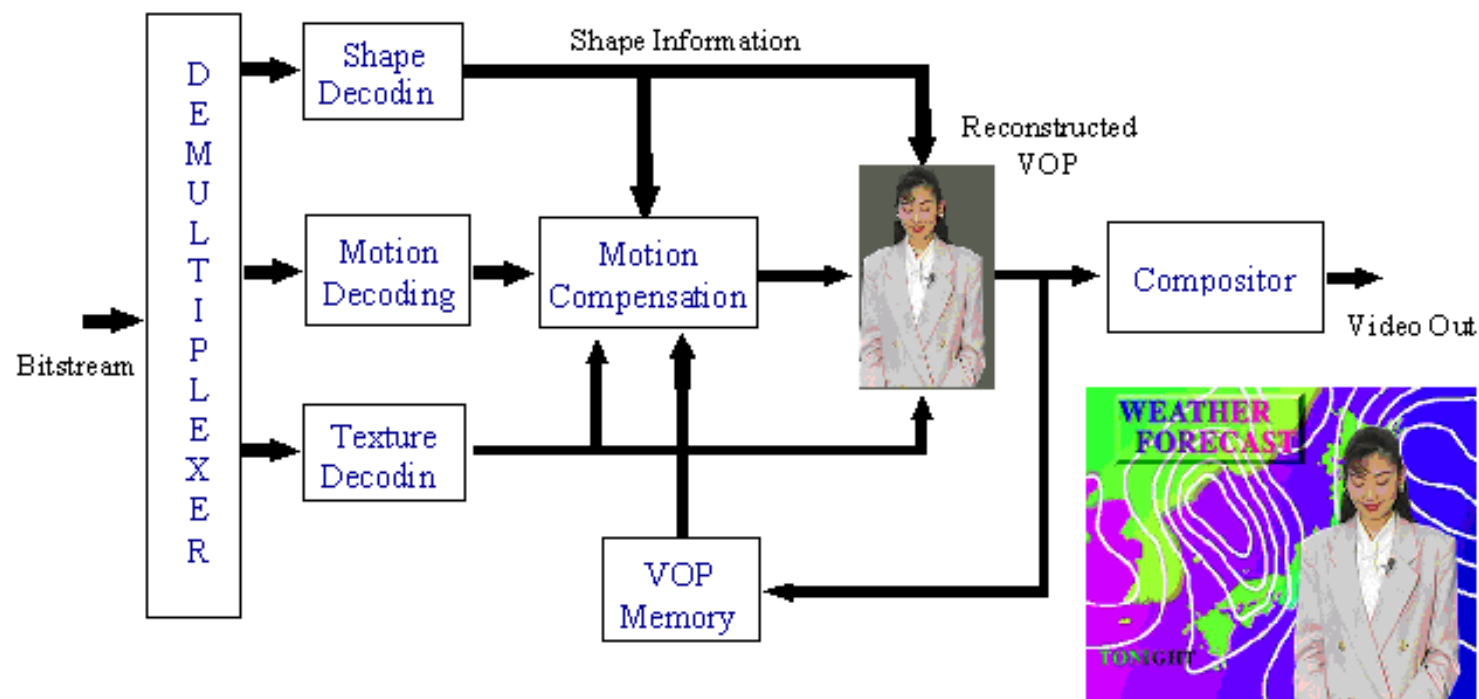
# Architettura MPEG4

- *Codifica*: l'utente decide la distribuzione del bit-rate disponibile tra i vari VO.
- *Multiplicazione*: l'utente può alterare la descrizione della scena.
- *Demultiplicazione*: l'utente può richiedere la ricezione di solo una parte del bit-stream.
- *Decodifica*: l'utente può agire sulla composizione degli oggetti.

# Estrazione di VO



# Manipolazione VOP



# Compressione e bitrate

## ➤ Codifica MPEG-1

➤ 50 kbit/s

➤ 500 kbit/s

➤ 75 kbit/s

➤ 100 kbit/s

# Compressione e bitrate

## ➤ Codifica MPEG-1

➤ 50 kbit/s



➤ 75 kbit/s



➤ 500 kbit/s



➤ 100 kbit/s

# Compressione e bitrate

## ➤ Codifica MPEG-2

➤ 50 kbit/s

➤ 500 kbit/s

➤ 75 kbit/s

➤ 100 kbit/s

# Compressione e bitrate

## ➤ Codifica MPEG-2

➤ 50 kbit/s



➤ 75 kbit/s



➤ 500 kbit/s



➤ 100 kbit/s

# Compressione e bitrate

## ➤ Codifica MPEG-4

➤ 50 kbit/s

➤ 500 kbit/s

➤ 75 kbit/s

➤ 100 kbit/s



# Compressione e bitrate

## ➤ Codifica MPEG-4

➤ 50 kbit/s



➤ 75 kbit/s



➤ 500 kbit/s



➤ 100 kbit/s