

Codifica Video MPEG

Ing. Francesco Benedetto - Prof. Gaetano Giunta

Laboratorio di Telecomunicazioni (COMLAB)
Università degli Studi Roma Tre

Perché comprimere...



Si supponga di voler trasmettere 30 fotografie in bianco e nero, digitalizzate e quantizzate con 8 bit/pixel di dimensione CIF (352x288) nel caso si utilizzi una codifica naturale senza compressione (es. BMP o TIFF).

Ogni fotografia risulta composta da $352 \times 288 = 101376$ pixel. Ogni foto viene poi digitalizzata e quantizzata con 8 bit per ogni pixel, ottenendo un numero di bit per ogni foto pari a $8 \times 101376 = 811008$ bit. Per 30 foto il conto totale in bit risulta pari a $811008 \times 30 = 24330240$ bit (ovvero più di 3MB di informazione).

Supponendo di utilizzare un canale GSM a 9.6 kbit/s, il tempo necessario alla trasmissione risulta pari a circa 42.2 minuti, con un modem a 64 kbit/s si riduce a 6.3 minuti.

Gruppo MPEG



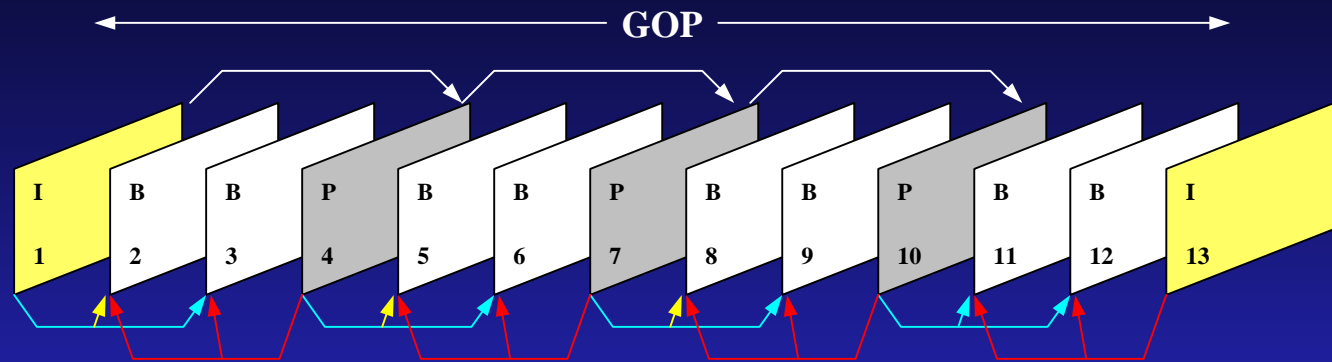
- Le direzioni di lavoro originariamente seguite dal gruppo MPEG sono 3:
- codifica digitale del segnale fino a 1,5 Mbit/s, codifica a 10 Mbit/s, codifica a 40 Mbit/s.
- Queste tre diverse direzioni di lavoro sono battezzate: MPEG-1, MPEG-2, MPEG-3, ma l'attività sul filone MPEG-3 viene abbandonata, 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*)

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.

Flusso video MPEG 2



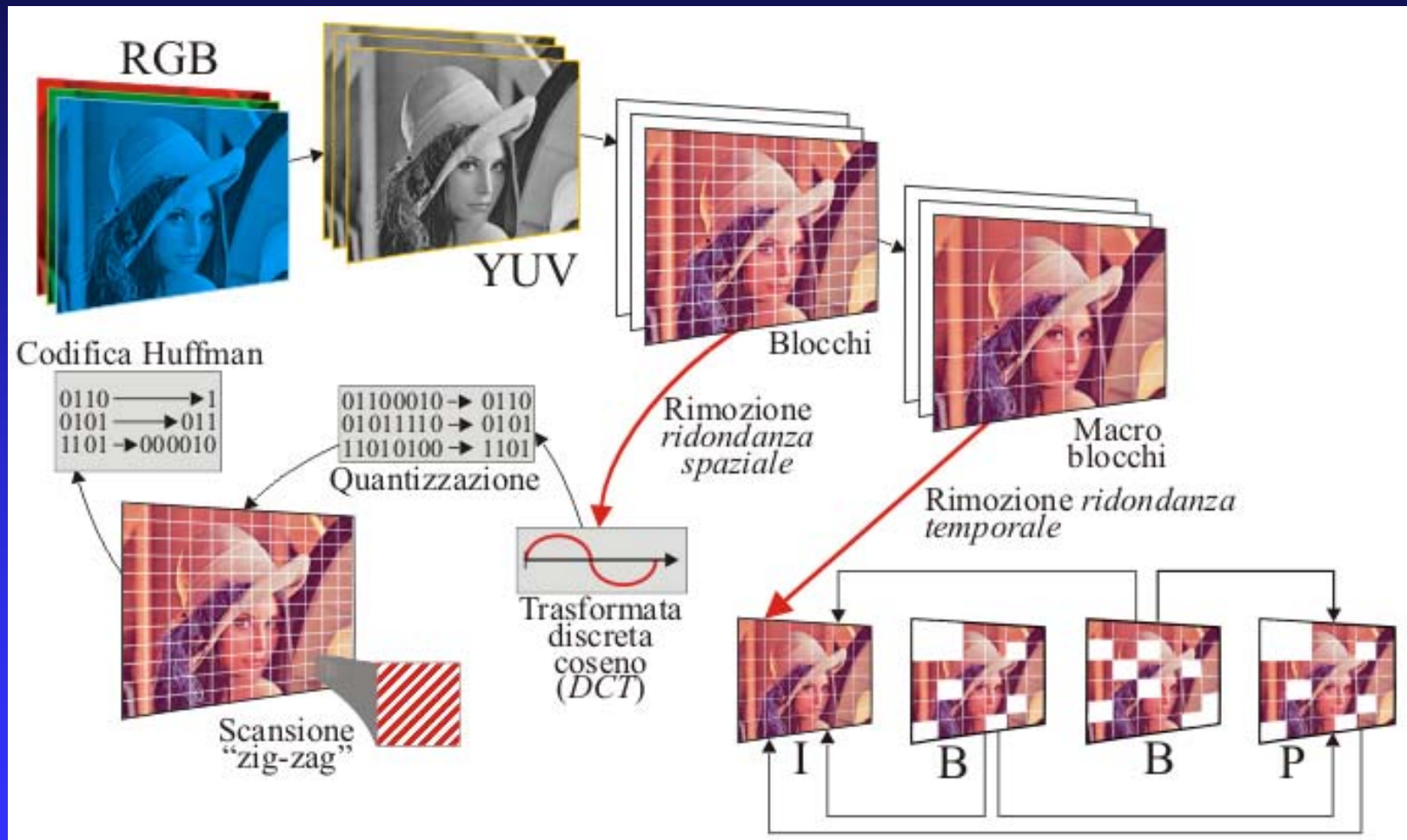
- Quindi, ogni porzione contiene uno o più macro blocchi, consistenti di quattro blocchi di luminanza (Y) e di due blocchi di crominanza (U,V). infine, il blocco è l'unità di base di codifica di dimensione 8x8 pixel.
- Per ottenere un alto rapporto di compressione, ci si serve sia delle ridondanze spaziali che di quelle temporali. La ridondanza spaziale viene ridotta dall'utilizzo di un sottocampionamento delle componenti di crominanza (U,V), in accordo con la sensibilità dell'apparato visivo umano.

Flusso video MPEG 2

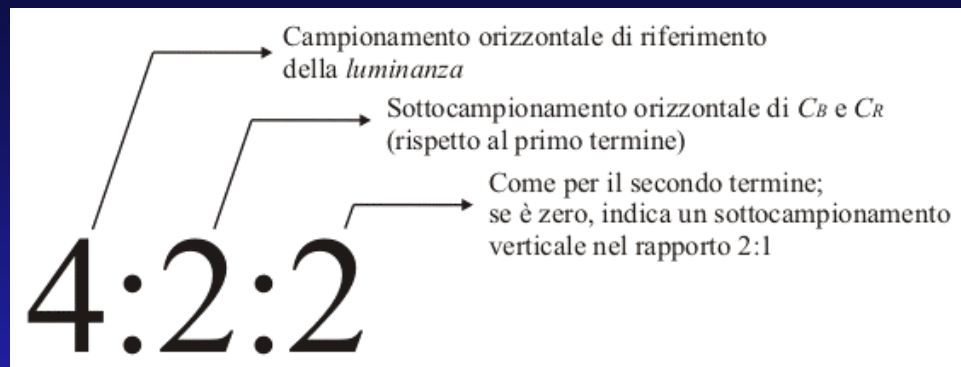


- Successivamente, viene utilizzata la trasformata coseno discreta sui blocchi con componenti Y e U,V. i coefficienti DCT sono quantizzati e infine codificati utilizzando un codice a lunghezza variabile. La ridondanza temporale viene ridotta da una predizione temporale di alcune trame derivate da altre trame moto-compensate. L'errore di predizione viene quindi codificato.

Flusso video MPEG 2

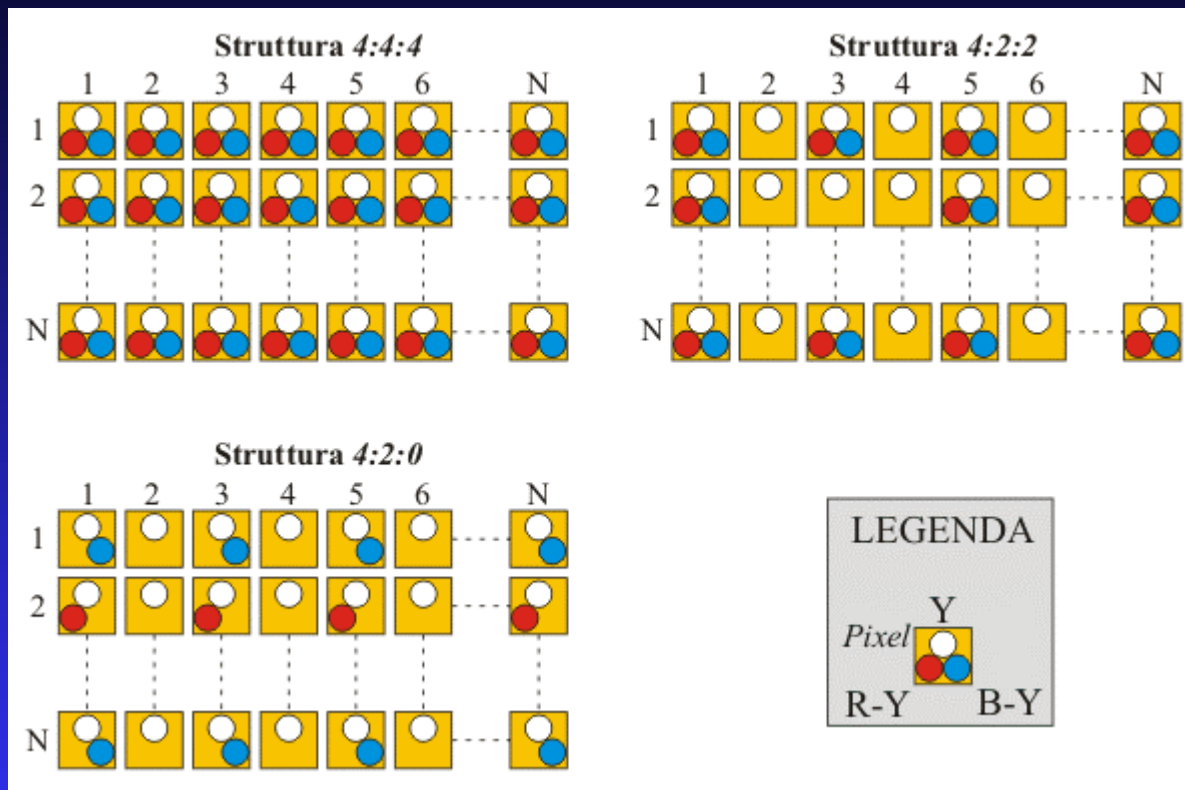


Sottocampionamento



- 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



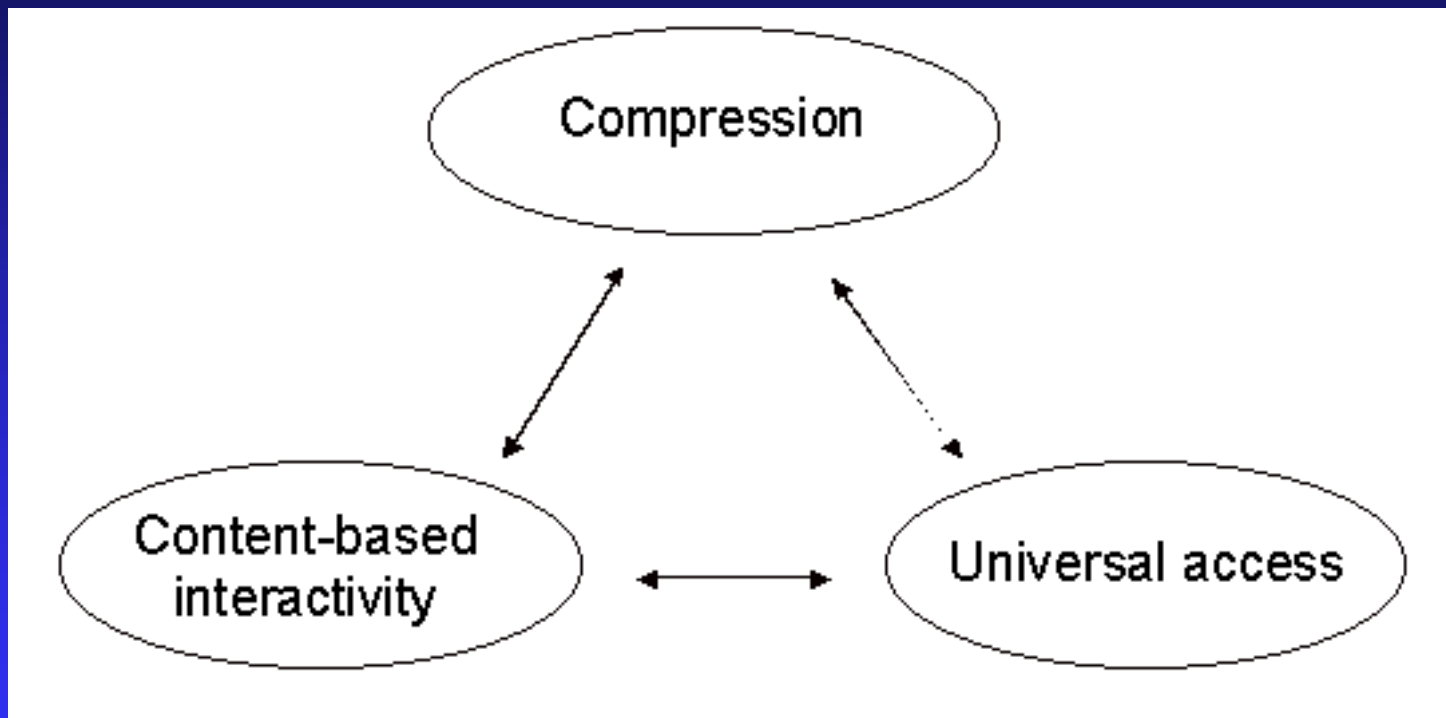
- La struttura 4:2:0 è generalmente adoperata nel JPEG, JFIF, H.261 e nel MPEG-1, mentre la struttura 4:2:2 è sfruttata nella codifica video MPEG-2.

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

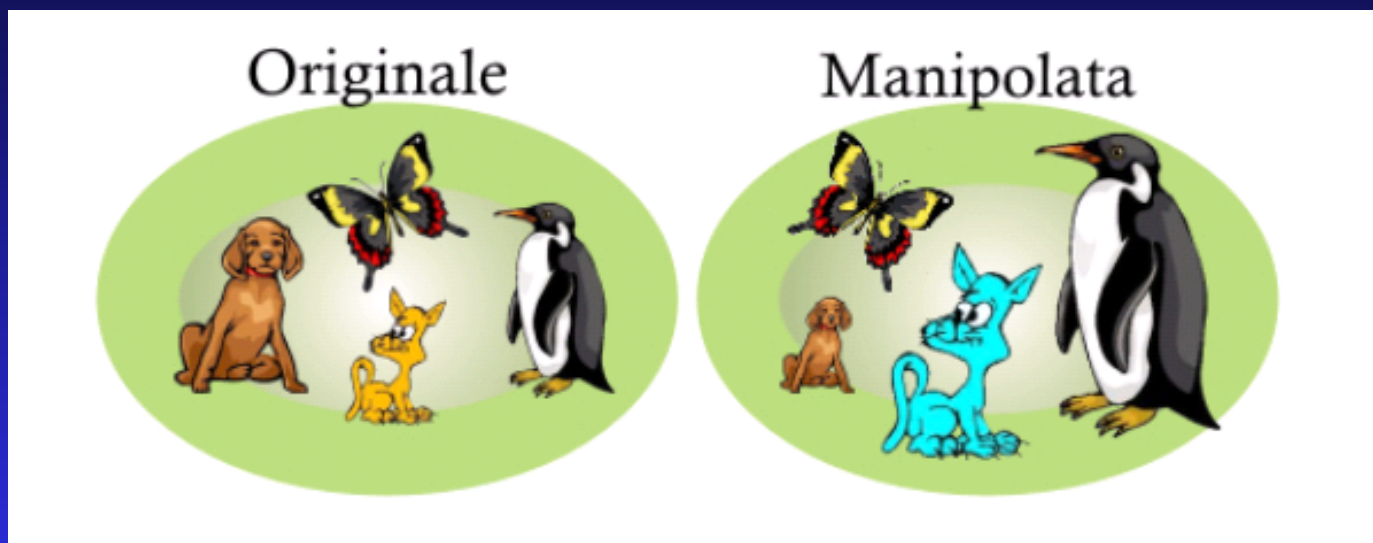


- **Compressione dati:**
 - Metodi efficienti per immagazzinamento e compressione dati.
 - Miglioramento efficienza di codifica.
 - Codifica di più flussi di dati concorrenti.

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.



- 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.

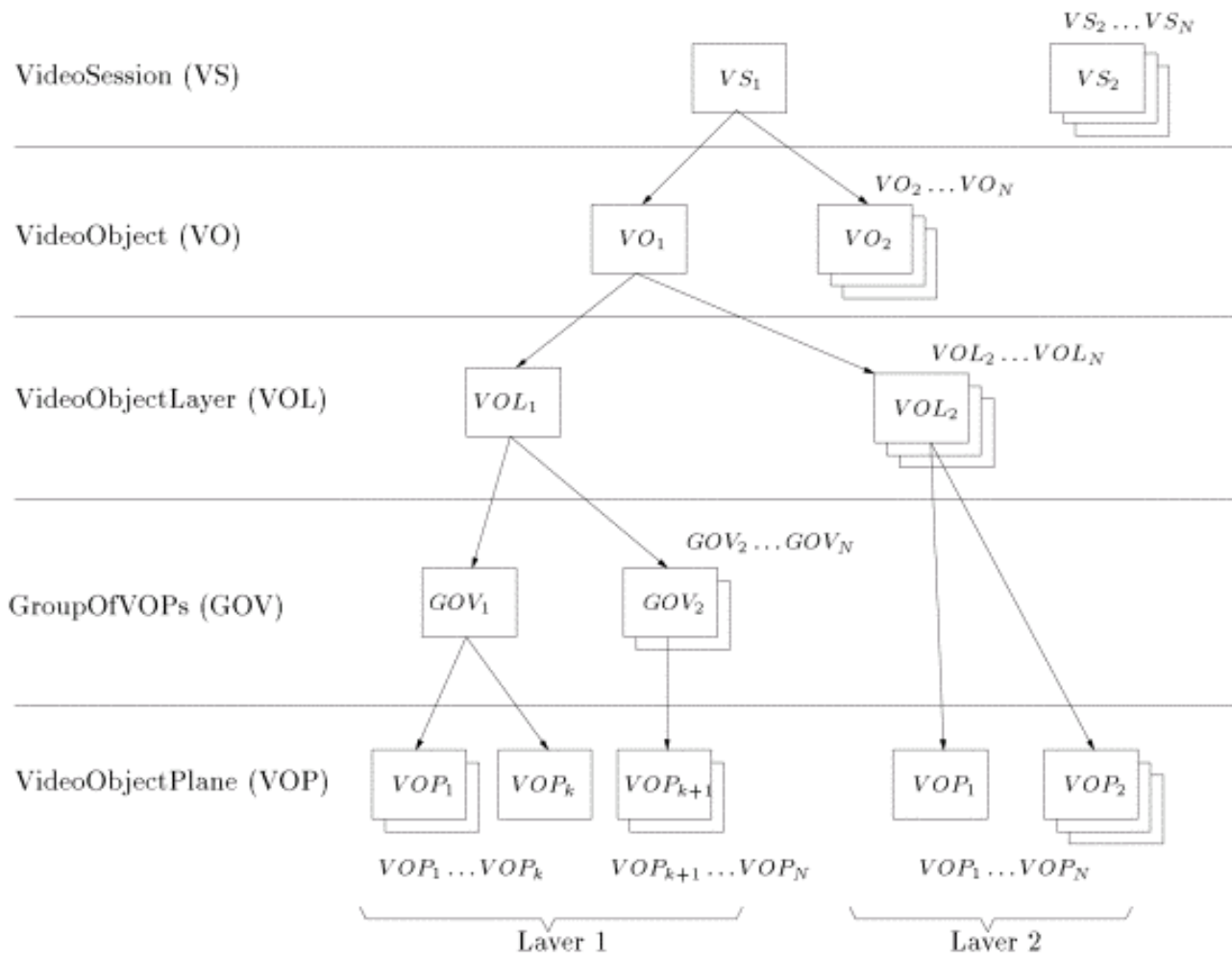
Classi MPEG4



- **Video Session (VS)**: compone le sequenze video incorporando oggetti dalle altre 3 classi.
- **Video Object (VO)**: oggetto all'interno di una scena.
- **Video Object Layer (VOL)**: esalta la risoluzione spaziale e temporale di ciascun VO.
- **Video Object Plane (VOP)**: è un'occorrenza di VO ad un determinato istante.

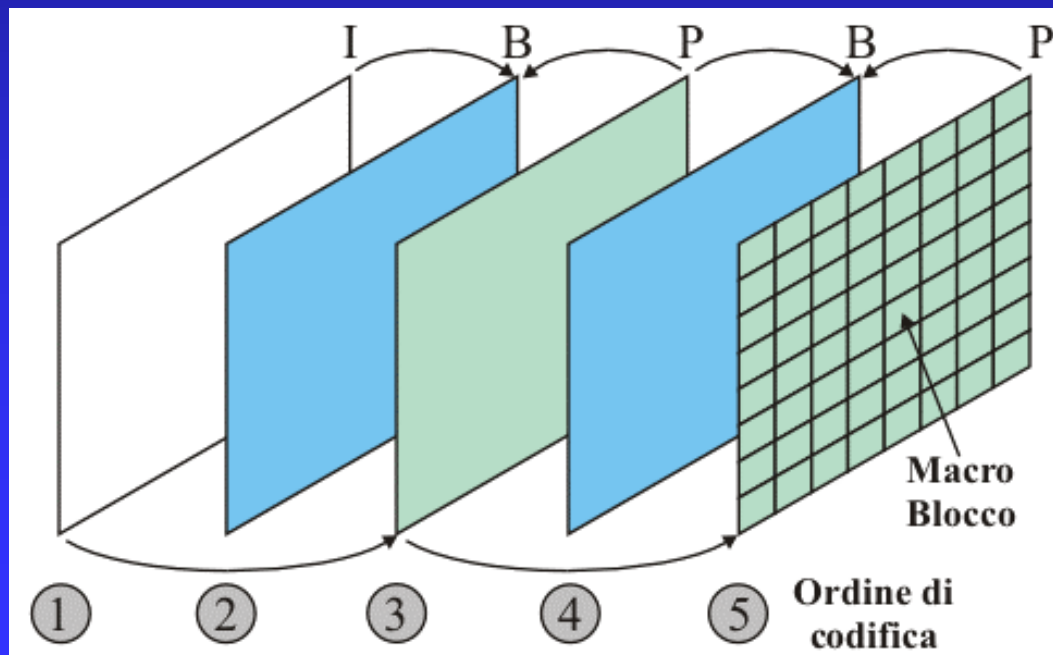
Una *VS* contiene uno o più *VO*, ciascuno dei quali possiede uno o più *VOL* costituiti da una sequenza di *VOP*.

Classi MPEG4



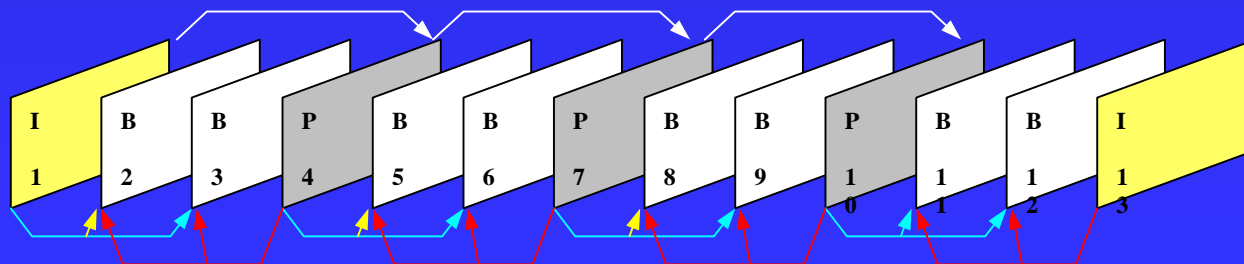
Codifica video

- Codifica diversa a seconda dei frames
 - *Interframe* con riferimento alle immagini precedenti
 - *Intraframe* indipendente dagli altri frames
- Codifica Interframe a doppia predizione una nel passato e una nel futuro

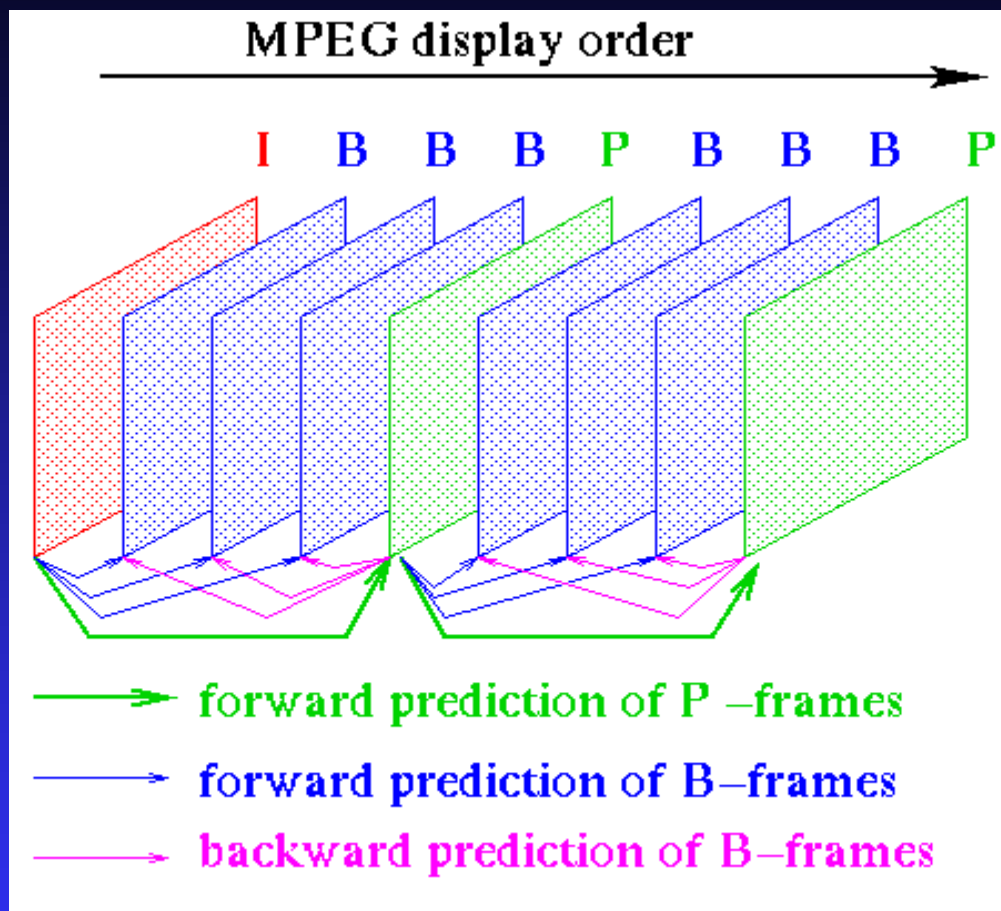


Codifica video

- Codifica intra (I), predizione temporale (P), predizione bidirezionale (B). In MPEG-2 ci riferiamo alle trame, in MPEG-4 ai VOPs.
- I - trama (I – VOPs): sono codificate senza nessun riferimento alle altre trame.
- P - trama (P - VOPs): sono codificate riferendosi al precedente I o P.
- B - trama (B - VOPs): sono codificate riferendosi sia alle trame (o VOPs) precedenti che a quelle successive.

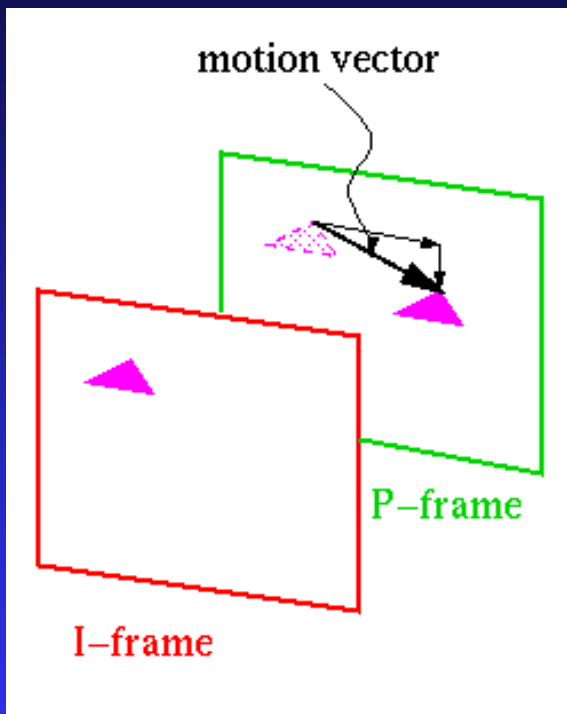


Codifica video



- 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.

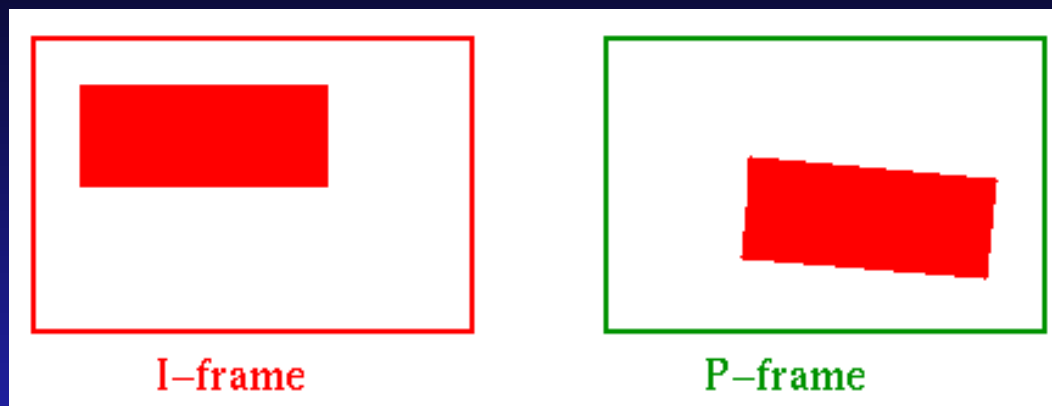
La predizione



- La predizione consiste nel fornire un vettore di movimento (*motion vector*) che dichiara 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.

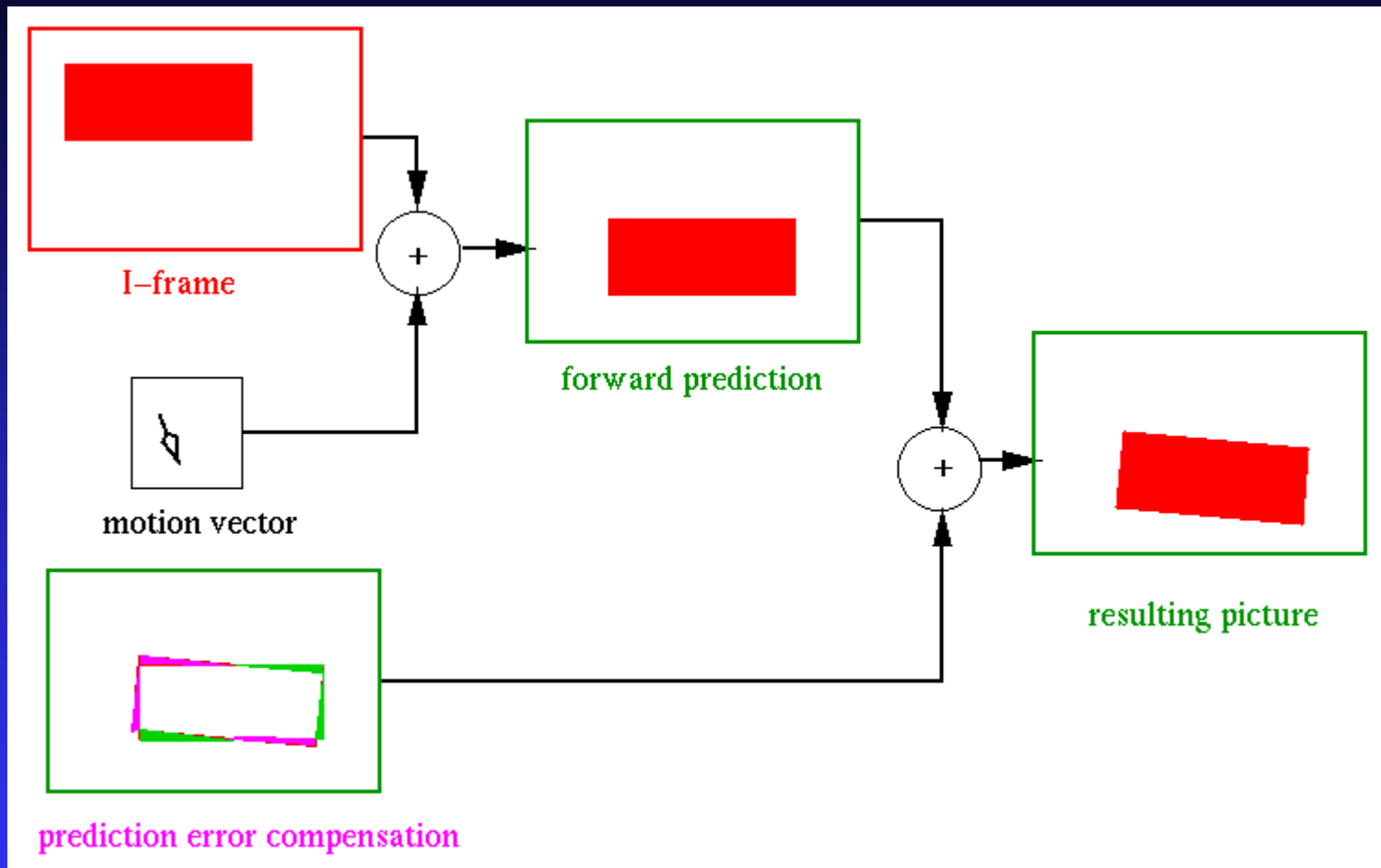
- Un valore positivo corrisponde ad un movimento a destra o verso il basso; uno negativo ad uno spostamento a sinistra o in alto rispettivamente.

Errore di predizione



- Il modello precedente assume che ogni differenza tra oggetti può essere resa disponendo i pixel in zone differenti. In realtà ciò non è sempre vero.
- Per ovviare a tale inconveniente si utilizza una matrice di compensazione dell'errore di predizione.

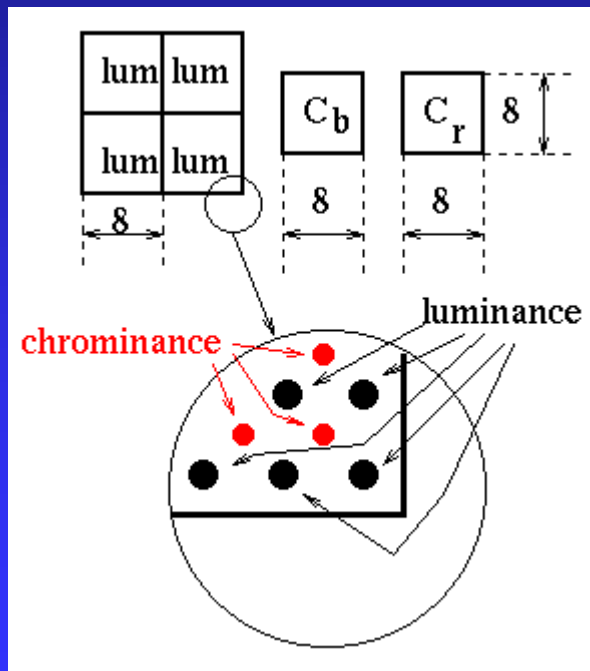
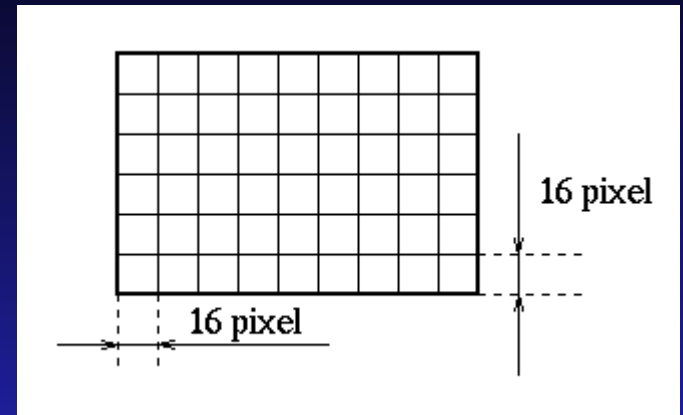
Errore di predizione



1. Si applica il motion vector.
2. Si compensa l'errore di predizione.

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.

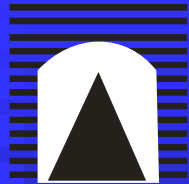


- Ogni blocchetto 8x8 è poi codificato attraverso la Discrete Cosine Transform (DCT).

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

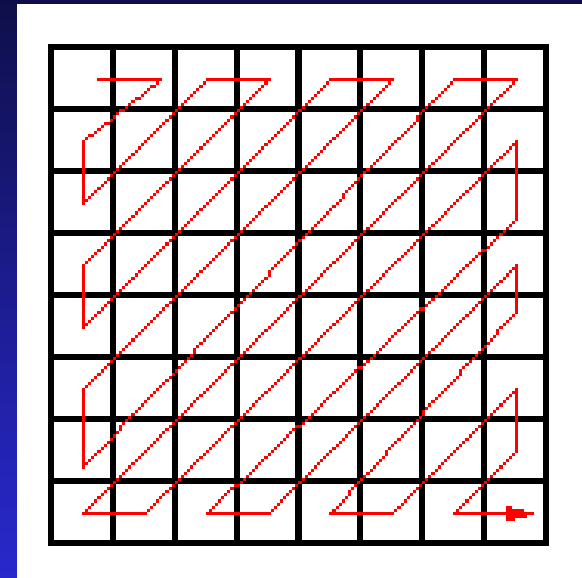
$$C_u = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } u = 0, \\ 1 & \text{if } u > 0 \end{cases}; C_v = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } v = 0, \\ 1 & \text{if } v > 0 \end{cases}$$

- 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) .

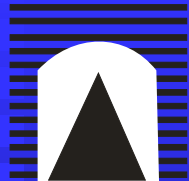


DCT

- 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.



IDCT

- Il decoder può ricostruire il valore dei pixel tramite trasformazione inversa:

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

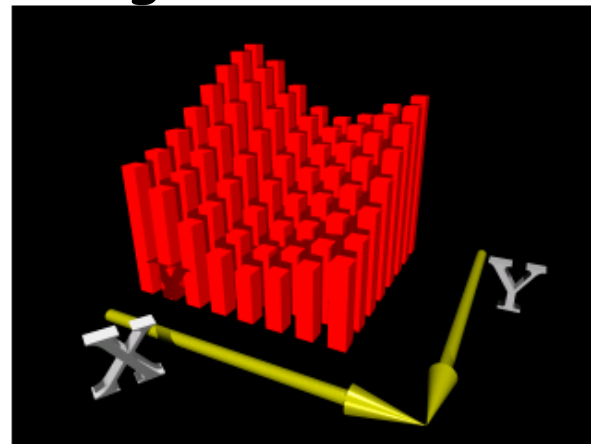
DCT

```
700 100 100 0 0 0 0 0
200  0  0 0 0 0 0 0
  0  0  0 0 0 0 0 0
  0  0  0 0 0 0 0 0
  0  0  0 0 0 0 0 0
  0  0  0 0 0 0 0 0
  0  0  0 0 0 0 0 0
  0  0  0 0 0 0 0 0
```

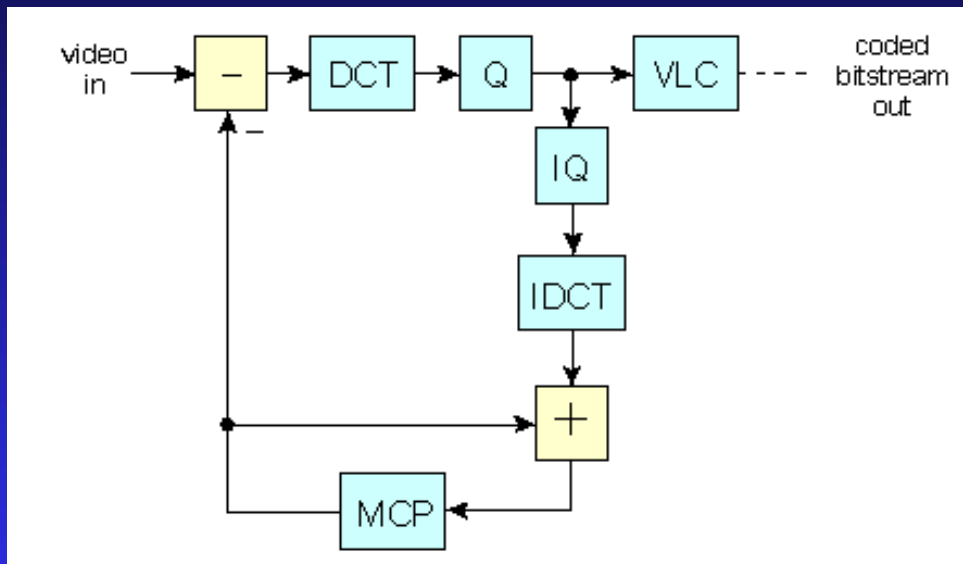
IDCT

```
156 144 125 109 102 106 114 121
151 138 120 104  97 100 109 116
141 129 110  94  87  91  99 106
128 116  97  82  75  78  86  93
114 102  84  68  61  64  73  80
102  89  71  55  48  51  60  67
 92  80  61  45  38  42  50  57
 86  74  56  40  33  36  45  52
```

Diagramma a barre



Codificatore MPEG2



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

(I)DCT = (inverse) discrete cosine transform

VLC = variable-length coder

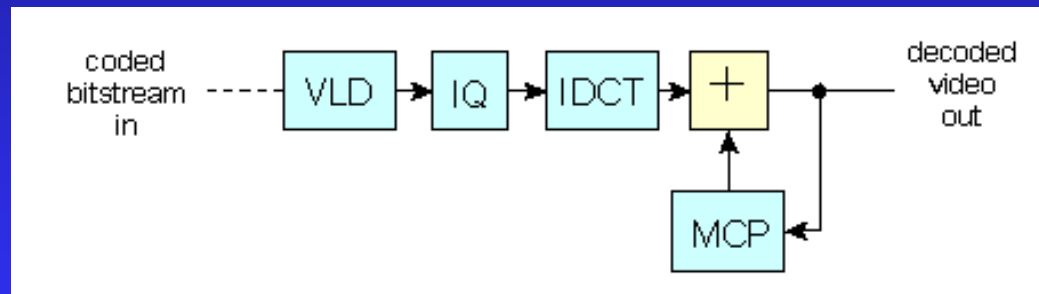
(I)Q = (inverse) quantisation

VLD = variable-length decoder

MCP = motion-compensated prediction

De-codificatore MPEG2

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



(I)DCT = (inverse) discrete cosine transform

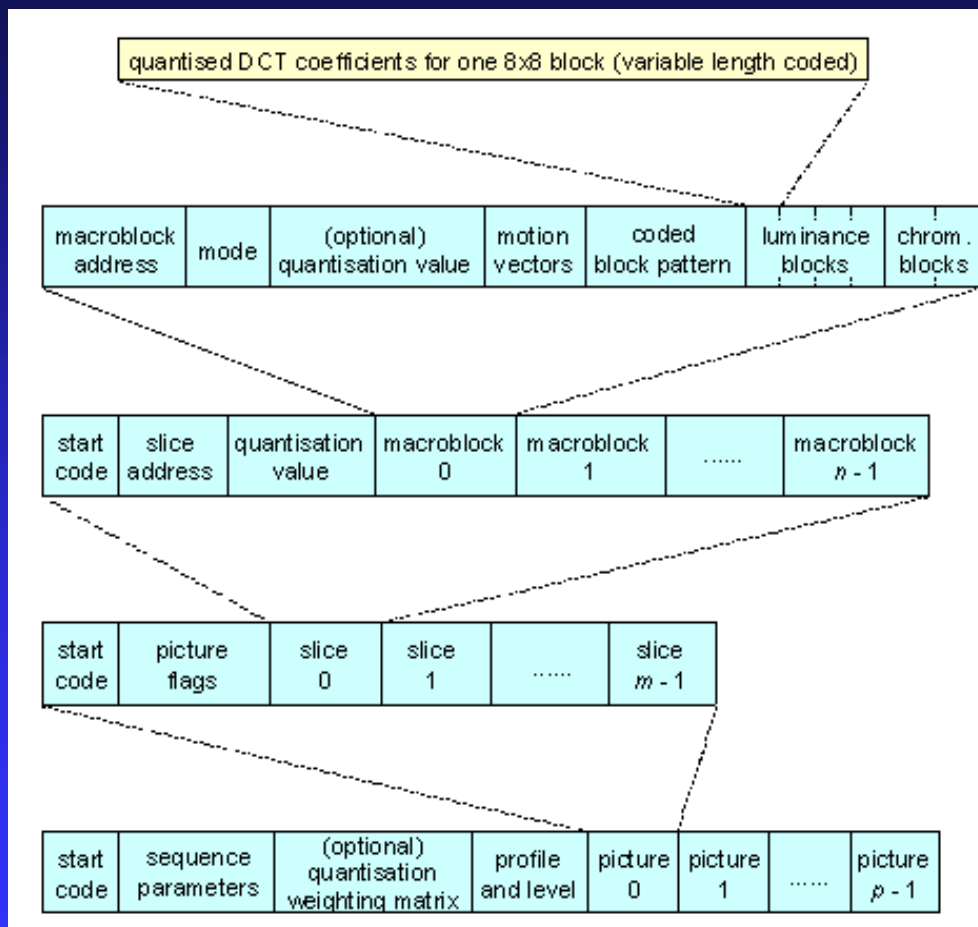
(I)Q = (inverse) quantisation

MCP = motion-compensated prediction

VLC = variable-length coder

VLD = variable-length decoder

MPEG2 bit-stream



- Block layer

- Macroblock layer
(4 Y, 2 C_BC_R)

- Slice layer
(n macroblocks)

- Picture layer
(m slices)

- Sequence layer
(p pictures)

Livelli MPEG2



Level	Max. frame, width, pixels	Max. frame, height, lines	Max. frame, rate, Hz	Max. bit rate, Mbit/s	Buffer size, bits
Low	352	288	30	4	475136
Main	720	576	30	15	1835008
High-1440	1440	1152	60	60	7340032
High	1920	1152	60	80	9781248

- MPEG2 definisce 4 livelli di codifica, ciascuno identificato da requisiti particolari sui parametri da utilizzare dal codificatore.

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



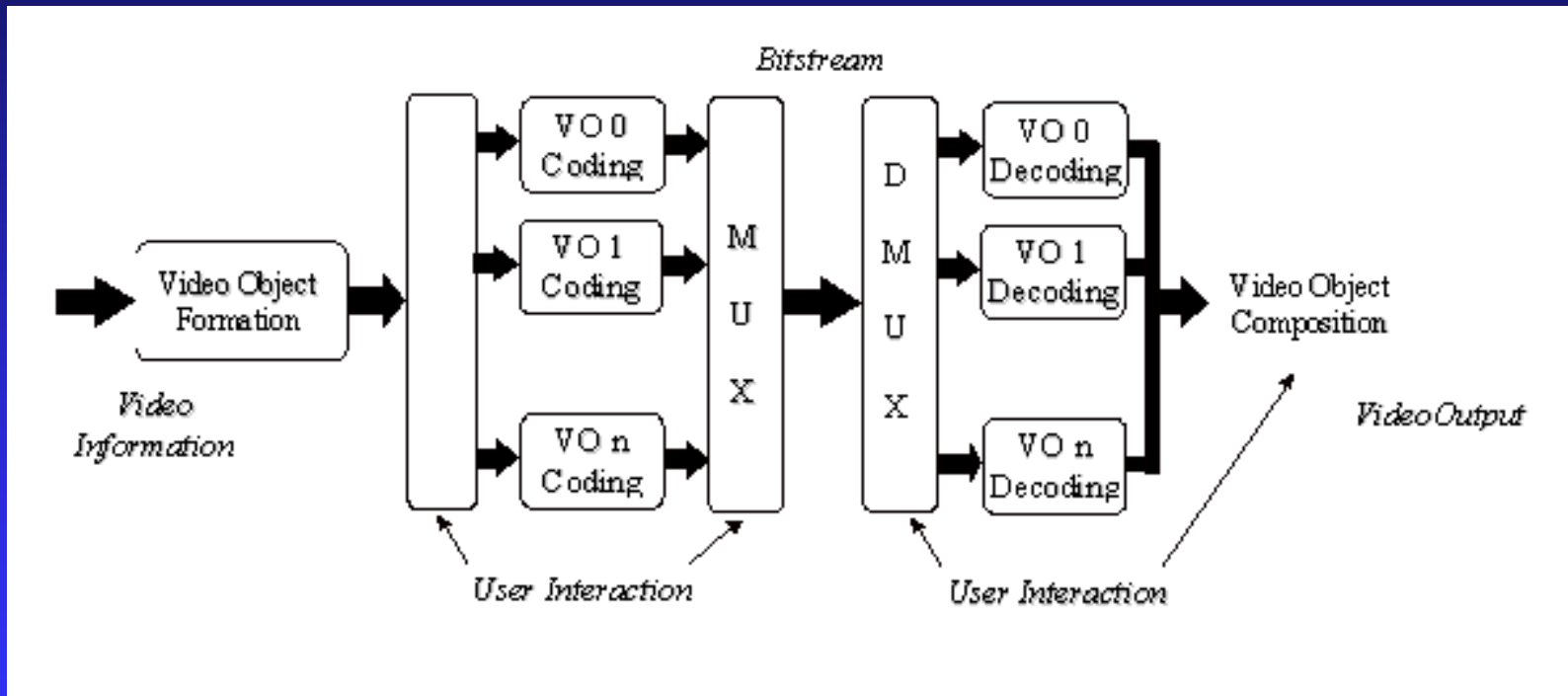
- Codifica: l'utente decide la distribuzione del bit-rate disponibile tra i vari VO.
- Multiplazione: l'utente può alterare la descrizione della scena.
- Demultiplazione: l'utente può richiedere la ricezione di solo una parte del bit-stream.
- Decodifica: l'utente può agire sulla composizione degli oggetti.

Architettura MPEG4

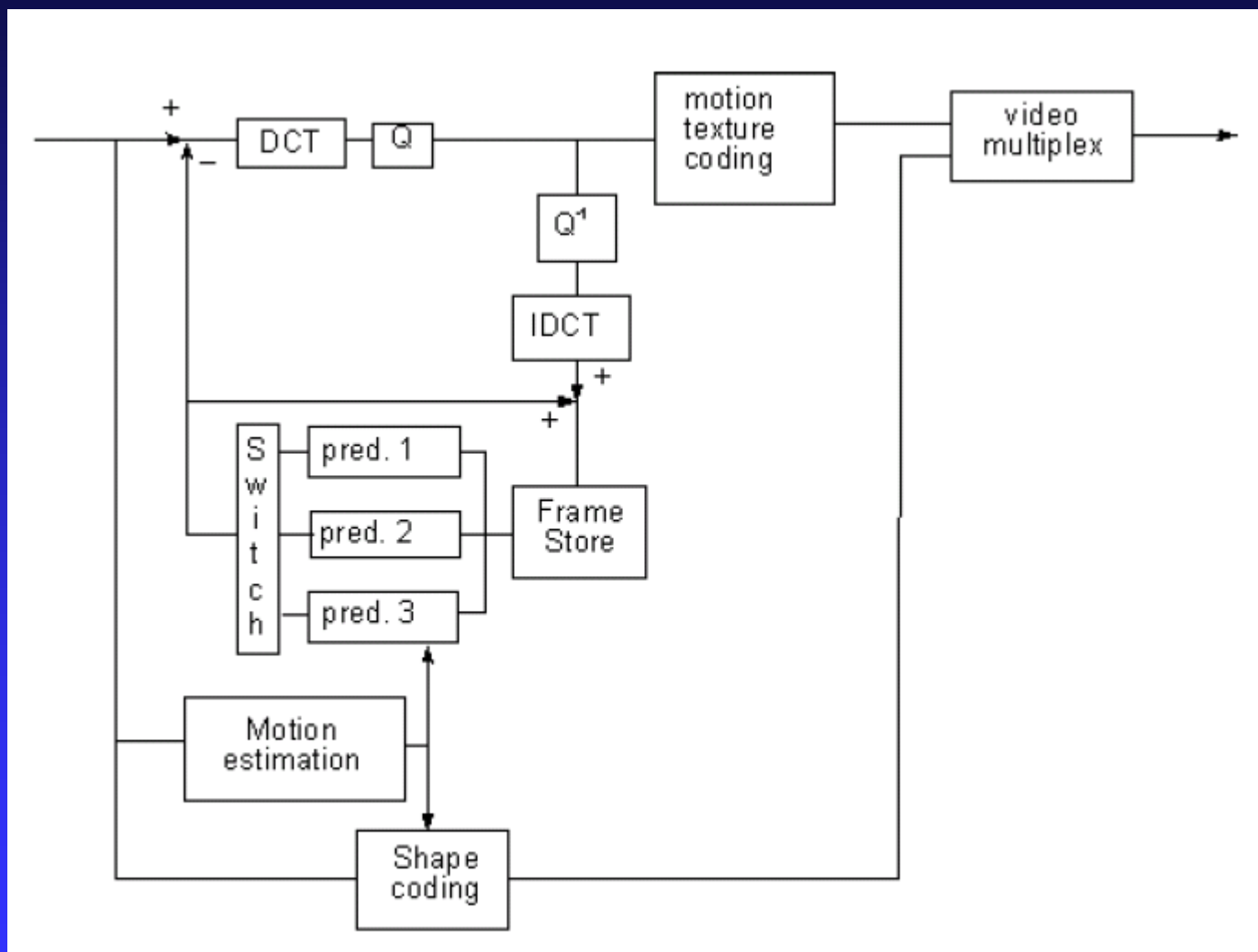


- Ogni VO è poi codificato e trasmesso al decoder. Dopo la decodifica l'oggetto è rappresentato nello spazio di colore YUV secondo il formato 4:2:0.
- Codifica della forma di ciascun VOP utilizzando una codifica binaria e una basata sulla scala dei grigi.
- Compensazione e stima del movimento mediante la suddivisione dei VOP in blocchi 8x8 o in macroblocchi 16x16.
- Codifica della tessitura con una DCT 8x8. Predizione efficiente dei coefficienti della DCT, scalabilità temporale e spaziale dei VOP.

MPEG 4 coder/decoder



Codificatore MPEG4



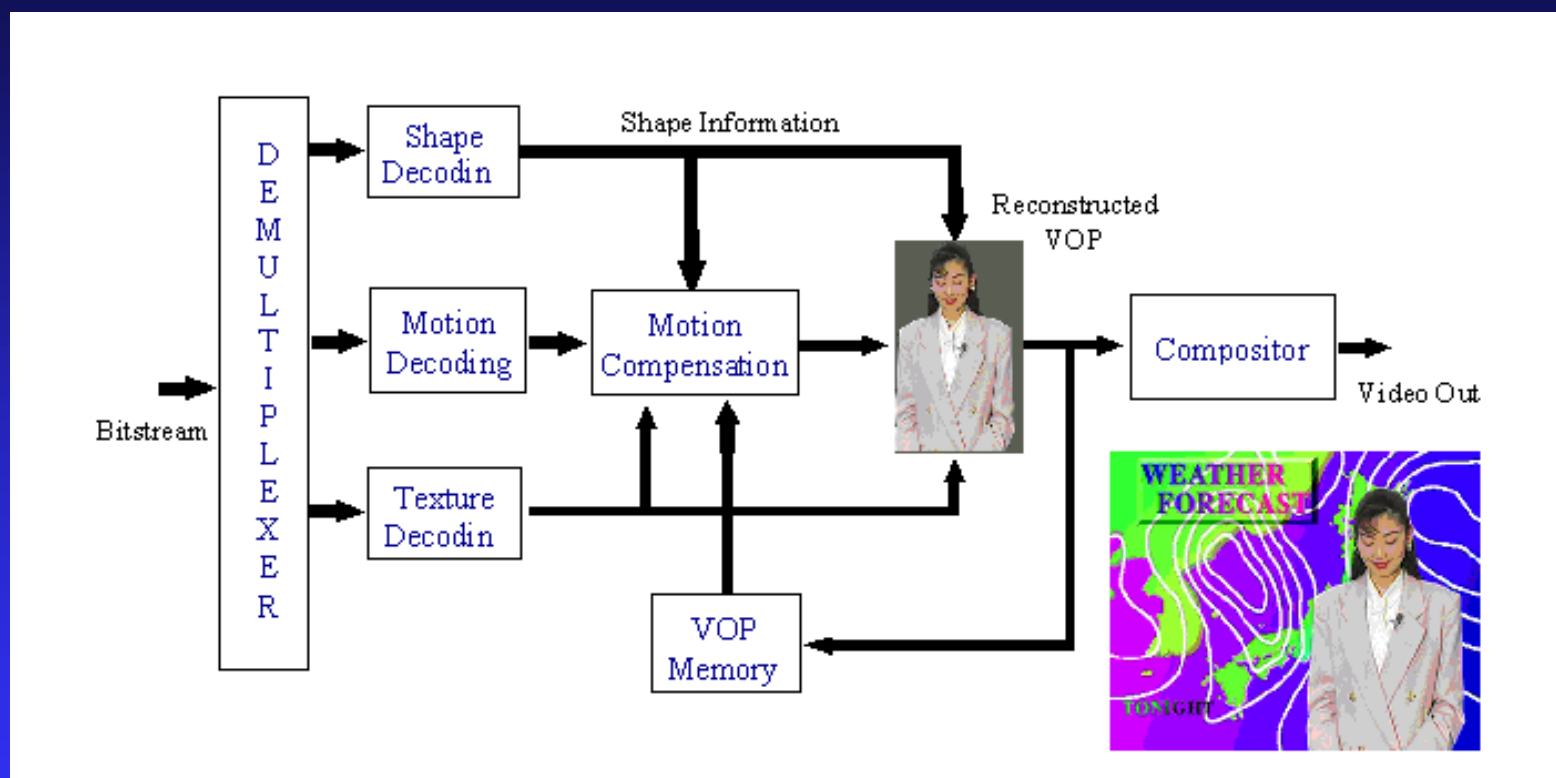


Estrazione di VO





Manipolazione VOP



Codifica della forma

- Codifica della forma binaria: si intende l'informazione che definisce quali pixel del supporto dell'oggetto appartengono all'oggetto in un dato istante.
- Rappresentazione con una matrice delle stesse dimensioni del blocco: ogni elemento può essere 0 o 255 a seconda che sia interno o esterno all'oggetto.
- Codifica della forma a scala di grigi: ogni elemento della matrice indica il grado di trasparenza, variabile tra 0 e 255.

Compressione e bitrate

➤ Codifica MPEG-1

➤ 50 kbit/s

➤ 500 kbit/s

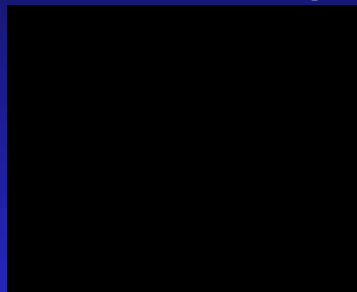
➤ 75 kbit/s

➤ 100 kbit/s

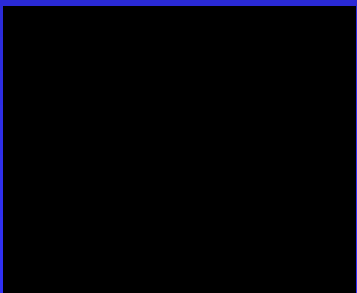
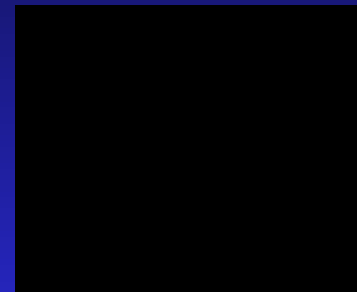
Compressione e bitrate

➤ Codifica MPEG-1

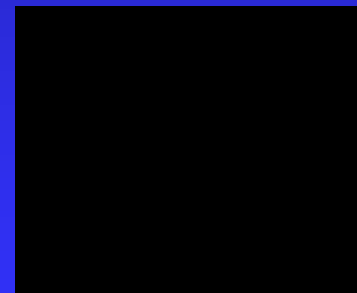
➤ 50 kbit/s



➤ 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

➤ Codifica MPEG-2 200 kbit/s

➤ 1 fps

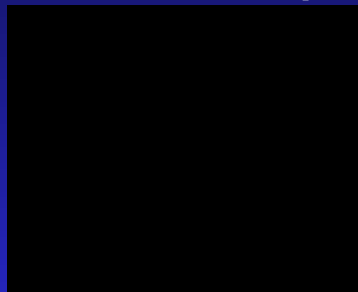
➤ 4 fps

➤ 7.5 fps

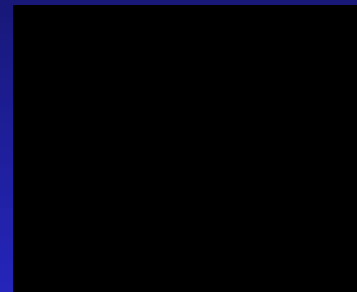
Compressione e bitrate

➤ Codifica MPEG-2

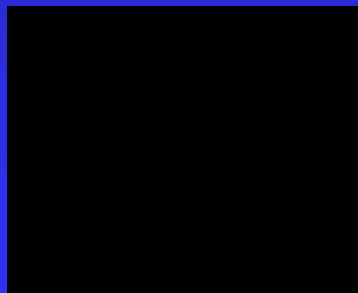
➤ 50 kbit/s



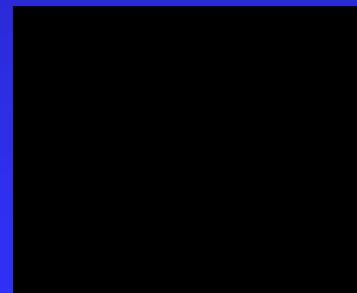
➤ 75 kbit/s



➤ 500 kbit/s



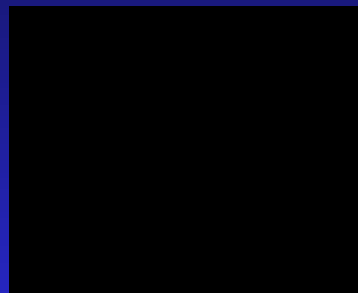
➤ 100 kbit/s



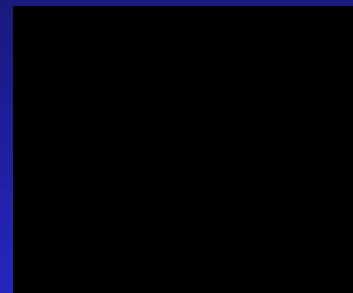
Compressione e bitrate

➤ Codifica MPEG-2

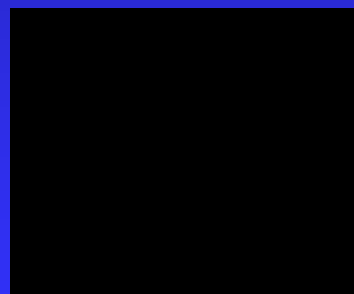
➤ 1 fps



➤ 4 fps



➤ 7.5 fps



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

Principali Riferimenti



G. Giunta: “*Appunti sulla DCT*”, Università degli Studi “ROMA TRE”, Facoltà di Ingegneria (scaricabile gratuitamente dal sito www.comlab.uniroma3.it/ens.html).

G. Giunta: “*Appunti sulla codifica MPEG-4*”, Università degli Studi “ROMA TRE”, Facoltà di Ingegneria (scaricabile gratuitamente dal sito www.comlab.uniroma3.it/ens.html).

R. Schafer: “*MPEG-4: a multimedia compression standard for interactive applications and services*”, Electronics & Communication Engineering Journal, Dicembre 1998.

P.N. Tudor: “*MPEG-2 video compression*”, Electronics & Communication Engineering Journal, Dicembre 1995.