



# Università degli Studi di Bergamo

---



**DIP. DI INGEGNERIA DELL'INFORMAZIONE E  
METODI MATEMATICI**

## **RETI INTERNET MULTIMEDIALI**

---

### **MPEG**

# Introduzione

---

- Il gruppo di esperti MPEG (Motion Picture Experts Group) si occupa della definizione dei seguenti standard
  - MPEG-1, MPEG-2, MPEG-4: sistemi e tecniche di codifica audio/video
  - MPEG-7: sistemi di ricerca di contenuti audio/video in database multimediali e descrittori base (forme, colori, etc.)
  - MPEG-21: sistemi per lo scambio di contenuti multimediali attraverso diverse tecnologie di rete/dispositivi e gestione della protezione dei diritti

# Introduzione

---

- MPEG definisce un sistema di codifica asimmetrico
  - Il codificatore richiede molte più risorse (computazionali e temporali) del decodificatore
  - Sviluppato per applicazioni non real-time (es. video storage).
- Gli standard definiti in MPEG sono organizzati in tre parti principali
  - Sistema: multiplexing di diversi flussi audio/video codificati per lo storage e la trasmissione
  - Codifica Video
  - Codifica Audio

# **MPEG**

---

## **Codifica Video**

# Introduzione

- I diversi standard di *codifica video* definiti da MPEG si rivolgono a diverse tipologie di applicazioni
  - Esigenze di risorse (banda, complessità, etc.)
  - Qualità del segnale video riprodotto

Standard	MPEG-1	MPEG-2 MP@ML
Risoluzione (Y)	352 x 240	720 x 480
Frame rate	30 fps	30 fps
Bit rate	1.5 Mbps	15 Mbps
Applicazioni	Storage Video	DVD, Digital Broadcast TV, HDTV

# MPEG - Group of Pictures

---

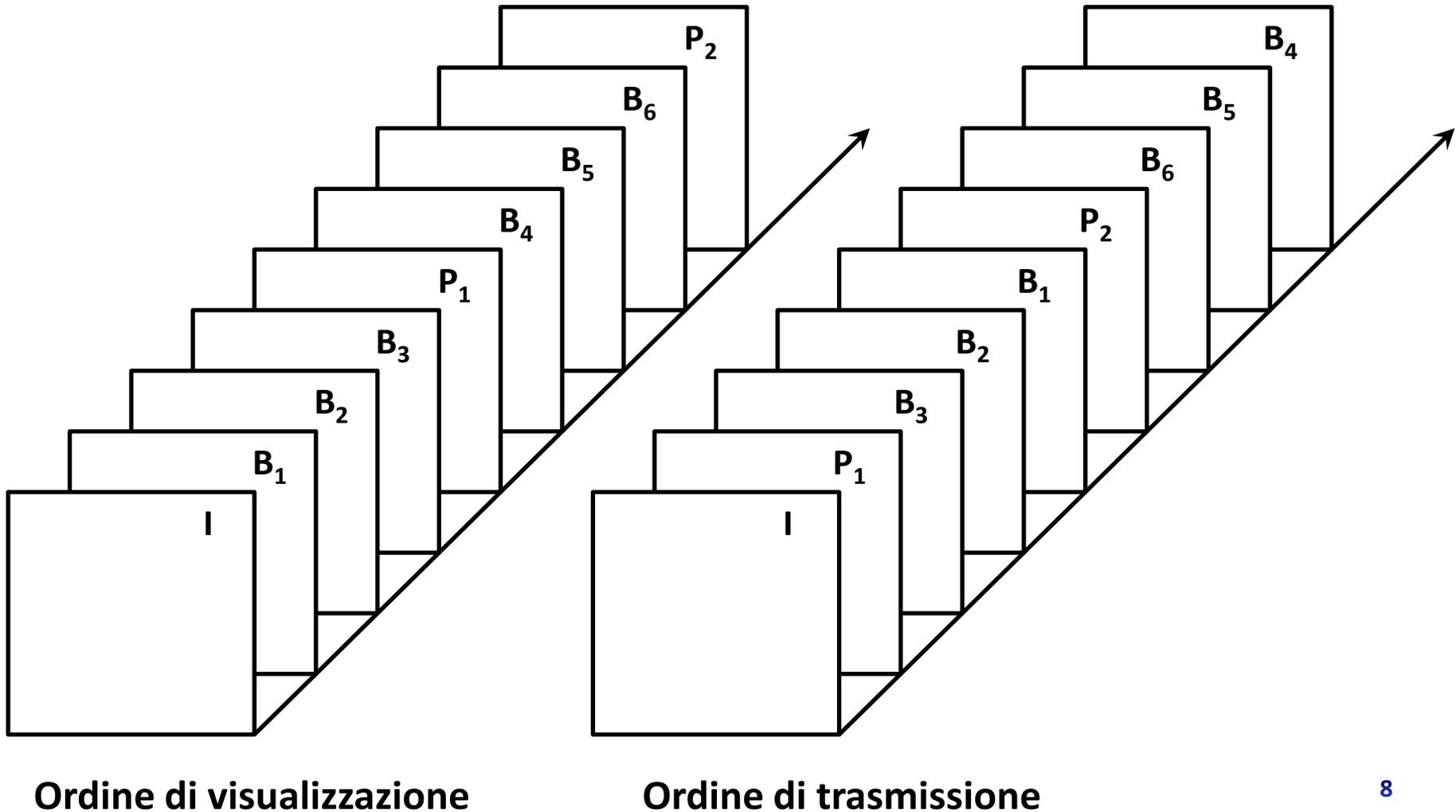
- Il **flusso video** viene codificato come una **sequenza di frames o pictures**
- Esistono tre tipologie di pictures:
  - I: Intra-Coded Pictures
  - P: Predictive Coded Pictures (forward prediction)
  - B: Bi-Directional Predictive Coded Pictures (bi-directional prediction)
- L'insieme dei frames viene suddiviso in sottoinsiemi detti Group of Pictures(GOP)
  - Ogni GOP è composta da una I picture, e diverse P e B pictures

# MPEG - Group of Pictures

---

- A causa delle tecniche di predizione utilizzate, al ricevitore la sequenza di frames decodificati risulterà diversa rispetto a quella di visualizzazione
  - Un P frame dipende da un I frame precedente
  - Un B frame dipende sia da un frame precedente sia frame successivo (siano essi I o P)
- E' necessario riordinare la sequenza di frames prima di visualizzarli

# MPEG - Group of Pictures



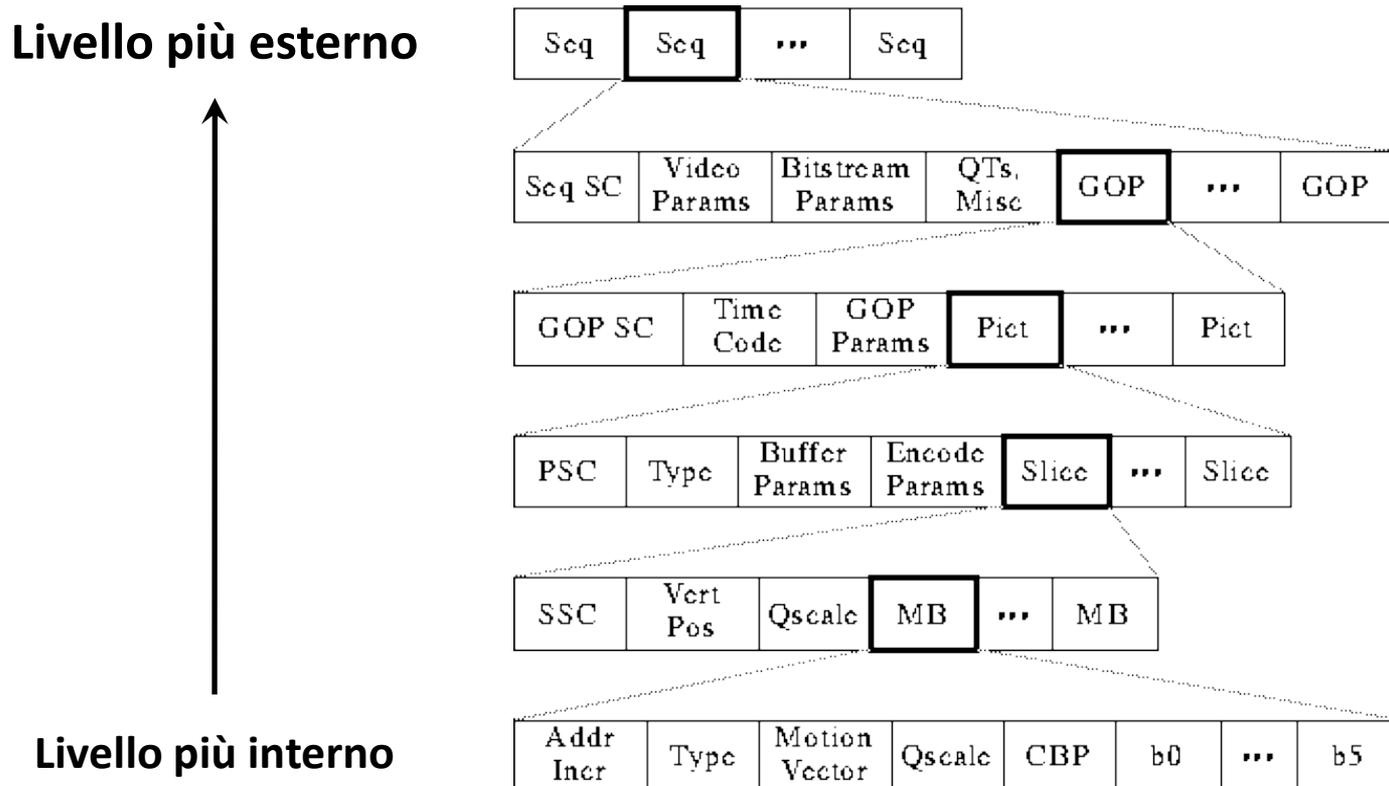
# MPEG – Formato Video

---

- Il flusso video è codificato in una struttura gerarchica formata da diversi livelli
  - Blocco 8x8
  - Macroblocco
    - 4 blocchi Y (una matrice 16x16)
    - 2 blocchi C (un  $C_b$  e un  $C_r$ )
  - Slice: sequenza contigua di macroblocchi (stessa tonalità di grigi)
  - Picture (I, P o B)
  - Group of Picture (GOP)

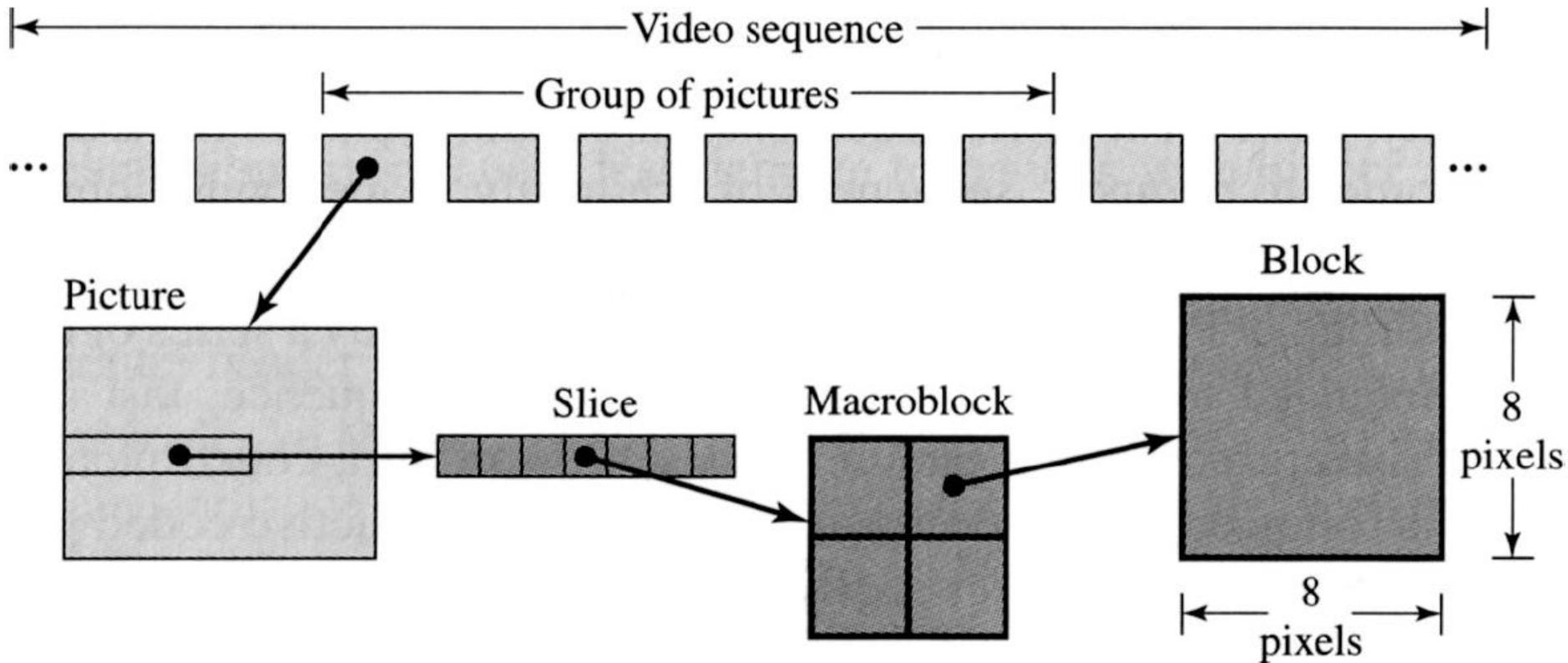
# MPEG – Formato Video

## ■ MPEG Bitstream: struttura gerarchica



# MPEG – Formato Video

## ■ MPEG Bitstream: struttura gerarchica



# Codifica Intra-frame (I-frame)

---

- Un frame viene codificato utilizzando solo le informazioni contenute nel frame stesso
  - Utilizzato per **I pictures**
- Ognuno dei blocchi 8x8 di luminanza e crominanza
  - DCT: coefficienti trasformati con DCT
  - Quantizzazione: output della DCT quantizzato
  - Entropy coding: coefficienti quantizzati vengono codificati utilizzando uno schema lossless
- Il processo è del tutto simile a quello eseguito da JPEG sequenziale

# Codifica Intra-frame (I-frame)

---

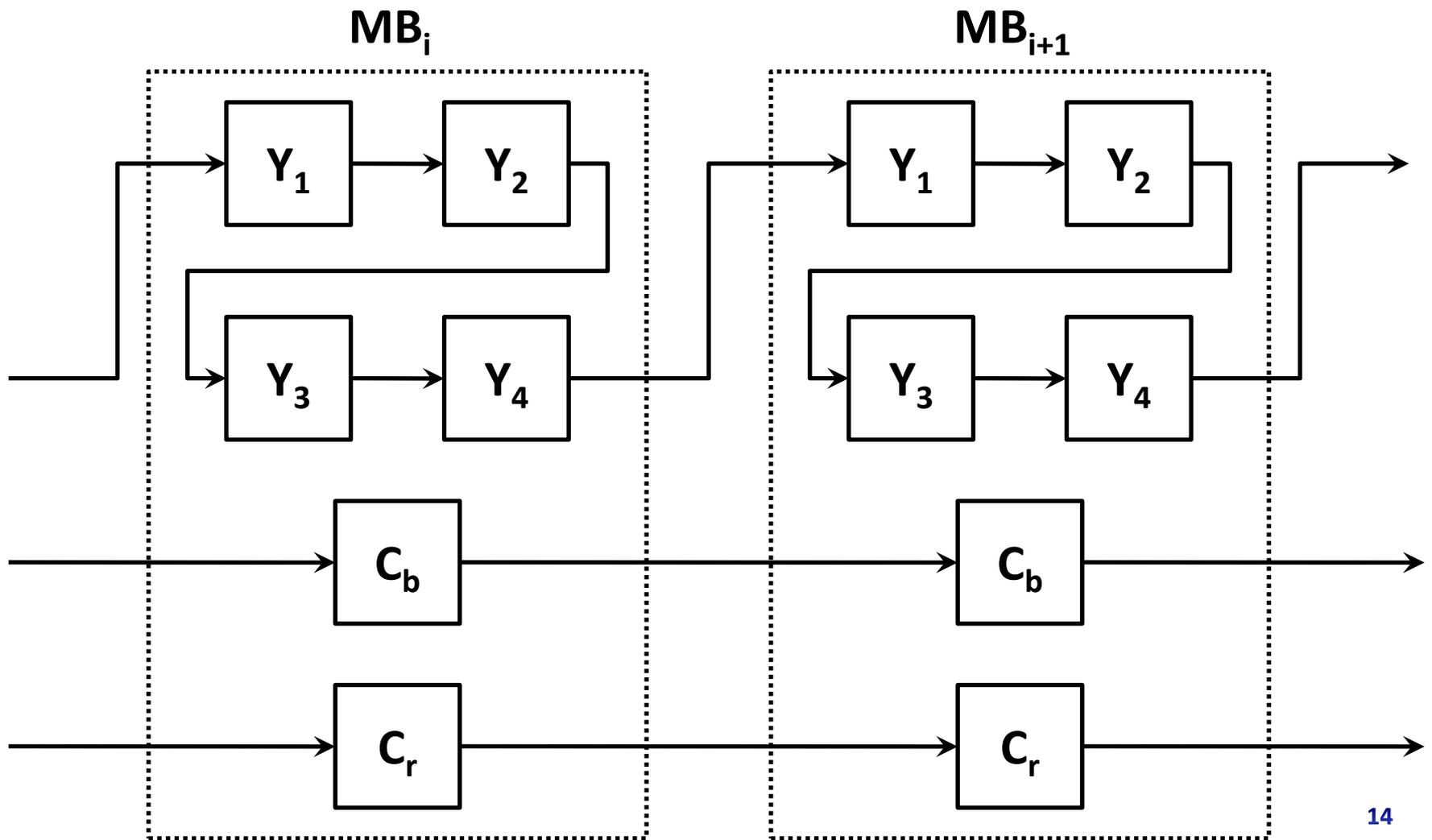
## ■ Codifica dei coefficienti DC di un I-frame

- La ridondanza dei coefficienti DC dei blocchi 8x8 adiacenti è ridotta utilizzando una codifica DPCM (Differential Pulse Coded Modulation)

$$\Delta DC_k^i = DC_k^i - P_k^i \quad P_k^i \rightarrow \text{Predittore del valore DC del blocco } i\text{-esimo del macroblocco } j\text{-esimo}$$

- Il segnale risultante dalla differenza (varia nel range [-255; 255]) è codificato utilizzando una compressione lossless VLC (Variable Length Codes)
  - Lo standard definisce il codice VLC
- L'ordine di predizione dei coefficienti DC dipende dall'ordine di codifica dei blocchi 8x8

# Codifica Intra-frame (I-frame)



# Codifica Intra-frame (I-frame)

---

- Codifica dei coefficienti AC di un I-frame
  - Molti dei coefficienti AC ad alta frequenza sono nulli
  - E' più efficiente codificare la posizione e il valore dei (pochi) coefficienti non nulli
- La sequenza di coefficienti AC quantizzati (zig-zag) viene codificata utilizzando una sequenza di simboli **<run:level>**
  - **Run:** indica il numero di elementi nulli che precedono un elemento non nullo
  - **Level:** indica il valore dell'elemento non nullo (successivo alla sequenza di elementi nulli)
- Ogni simbolo **<run:level>** viene codificato con un codice Variable Length Code (VLC) standardizzato

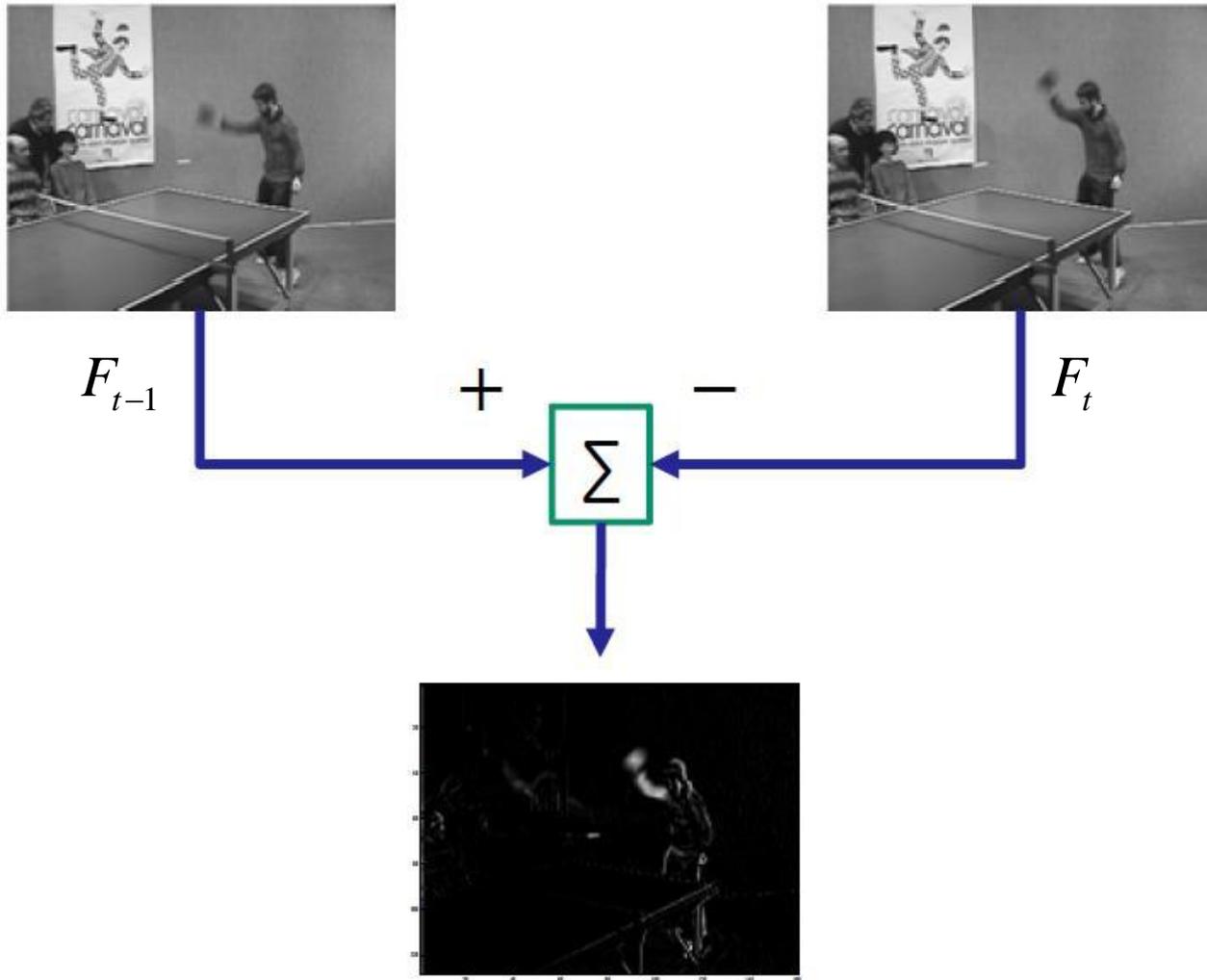


# Codifica Inter-frame (P-frame)

---

- In generale due frames consecutivi sono molto simili tranne per alcuni dettagli dovuti
  - Movimento (della camera e degli oggetti)
  - Variazione della luce (movimento delle ombre)
- Se si codifica il frame risultante dalle differenza tra il frame corrente e il frame precedente si aumenta la compressione
  - Poca variazione nel frame risultante (pochi coefficienti DCT)

# Codifica Inter-frame (P-frame)

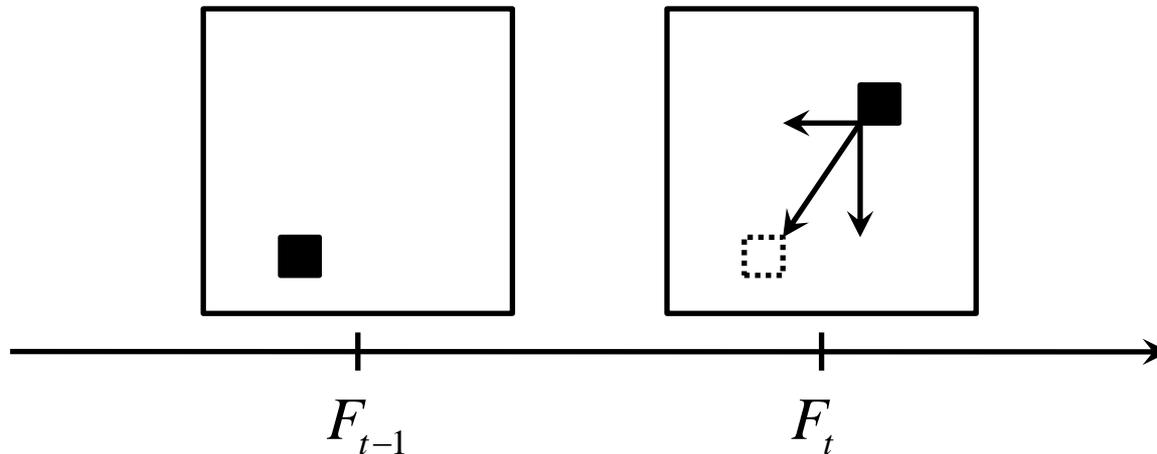


# Codifica Inter-frame (P-frame)

---

- Per migliorare ulteriormente la codifica si usa la tecnica di **block-based motion compensated prediction error**
- E' necessario eseguire la stima del movimento dei blocchi tra due frame consecutivi
  - **Motion estimation**
- Prima di eseguire la differenza si genera un nuovo frame utilizzando il frame di riferimento e l'informazione ottenuta dall'algoritmo di motion estimation (vettore di movimento)
  - **Errore di predizione:** differenza tra il frame corrente e la versione ottenuta dalla predizione (motion compensated frame)

# Codifica Inter-frame (P-frame)



**Codifica**

$$F_t^M = mc(\tilde{F}_{t-1}, \vec{v})$$

$$D_t = F_t - F_t^M$$

$$P_t = enc(D_t)$$

$$\vec{v}' = enc(\vec{v})$$

**Decodifica**

$$\vec{v} = dec(\vec{v}')$$

$$D_t = dec(P_t)$$

$$F_t^M = mc(\tilde{F}_{t-1}, \vec{v})$$

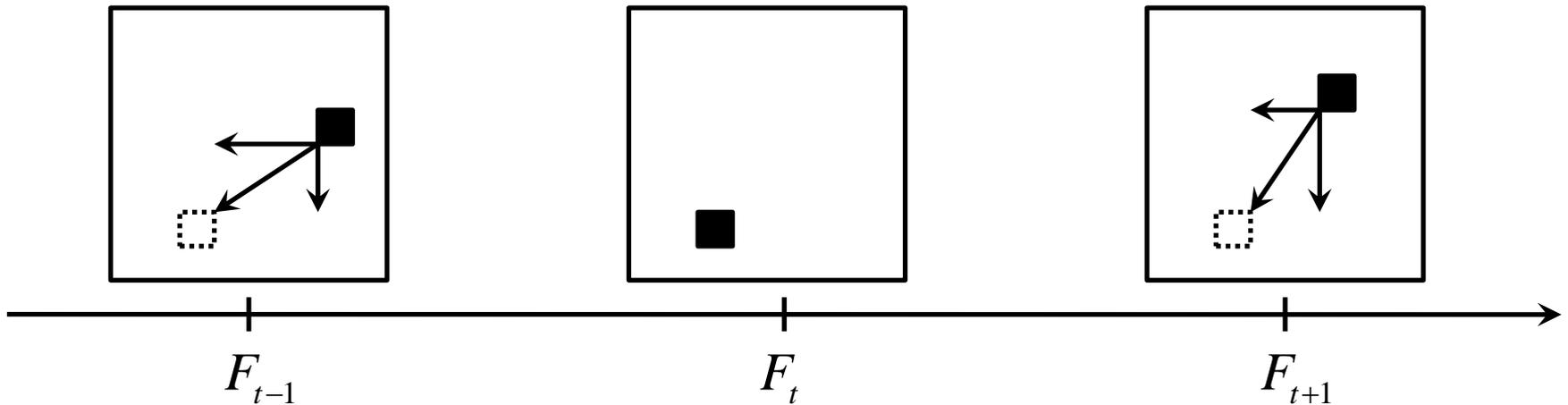
$$F_t = F_t^M + D_t$$

# Codifica Inter-frame (B-frame)

---

- La predizione Bi-direzionale utilizza due frame di riferimento
- Per ogni macroblocco del frame predetto vengono generati due vettori di movimento
  - Un vettore relativo a un frame nel passato
  - Un vettore relativo a un frame nel futuro
- La predizione bidirezionale è utile per la compressione di aree libere (non coperte da altri oggetti), ottenendo in un miglior rapporto di compressione
- I B-frame non vengono utilizzati come frame di riferimento per la predizione di altri frame

# Codifica Inter-frame (B-frame)



$$\vec{v}_t = \alpha_1 v_{t-1} + \alpha_2 v_{t+1}$$

$$\alpha_1 = 0.5 \wedge \alpha_2 = 0.5$$

**Bi-directional Prediction**

$$\alpha_1 = 1 \wedge \alpha_2 = 0 \longrightarrow$$

**Forward Prediction**

$$\alpha_1 = 0 \wedge \alpha_2 = 1 \longrightarrow$$

**Backward Prediction**

# Codifica Inter-frame (B-frame)

---

- Non tutti i macroblocchi dei B-frame derivano da predizione bi-direzionale
- I macroblocchi che compongono i B-frame sono codificati utilizzando predizione
  - Bidirezionale
  - Backward
  - Forward
  - Nessuna (Intra-coded)
- Viene utilizzata la tipologia di predizione che minimizza la varianza dell'errore di predizione della componente di luminanza

# Codifica Inter-frame

---

- L'informazione necessaria alla ricostruzione del macroblocco generato dalla codifica Inter-frame (P o B) è formata
  - Vettore di movimento
  - Errore di predizione
- La trasformazione DCT è applicata all'errore di predizione, prima della quantizzazione e della codifica lossless

# MPEG - Codifica

---

- Il codificatore MPEG è composto dai seguenti componenti
  - Controller: controllo e sincronizzazione del flusso video
  - FDCT (Forward DCT)
  - Quantizzatore
  - VLC encoder (Entropy coding)
  - Reconstruction Module
  - Motion Estimator (predizione del movimento)
  - Motion Compensation (compensazione del movimento)
- Il Reconstruction Module il Motion Compensation sono necessari per generare i frame di riferimento con cui eseguire la predizione del movimento (stima sugli stessi dati per codificatore e decodificatore)

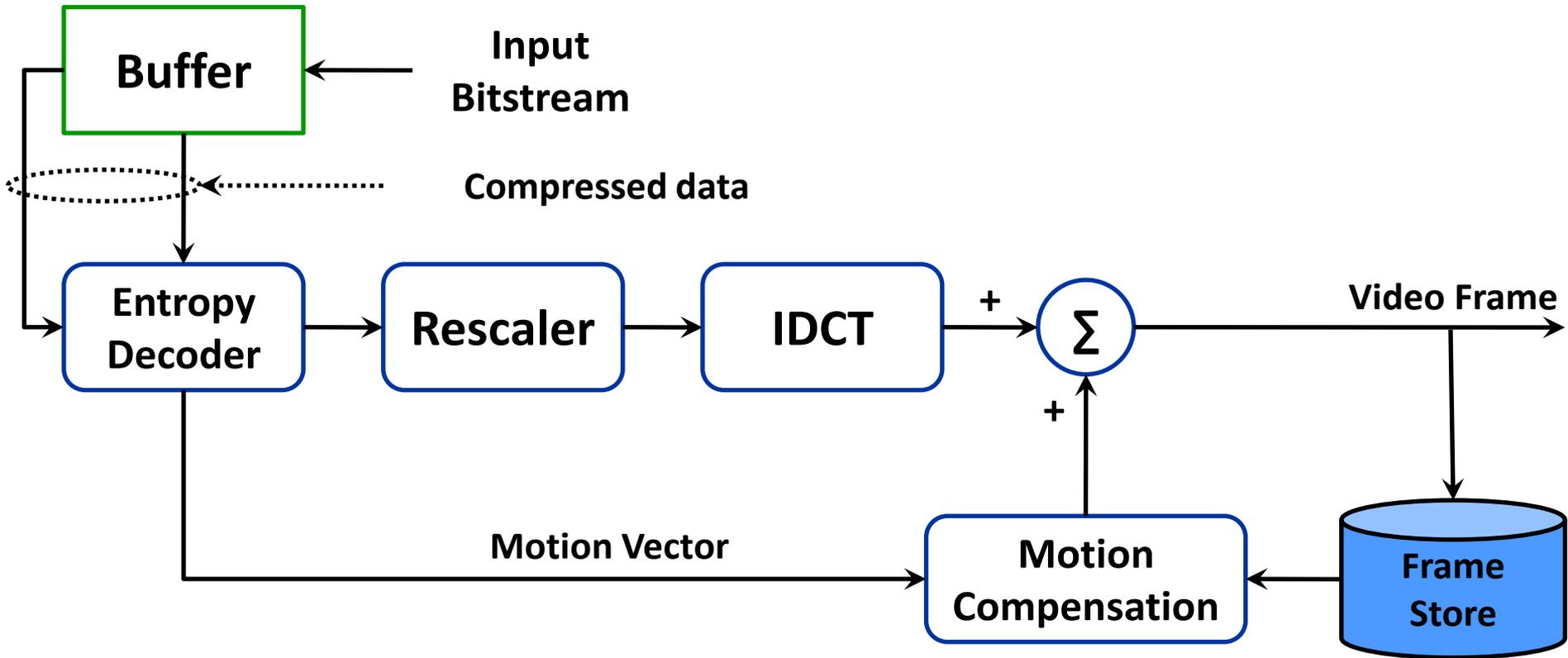


# MPEG - Decodifica

---

- Il decodificatore MPEG è più semplice del codificatore (codifica asimmetrica)
  - La ricostruzione è guidata dalle informazioni inviate dal codificatore (Coefficienti della DCT e motion vector)
- E' composto essenzialmente da due componenti
  - Reconstruction Module
  - Motion Compensation (compensazione del movimento)

# MPEG - Decodificatore



# MPEG

---

## IDCT Mismatch

# IDCT Mismatch

---

- L'efficienza di compressione dipende generalmente dalla codifica inter-frame
- Nella codifica inter-frame i dati di un frame predetto (P o B) sono aggiunti ai dati decodificati da precedenti frame
- Minime variazioni prodotte dalle IDCT del codificatore e del decodificatore si propagano fino a un I-frame
  - Se non controllate diventano percettibili come distorsione del segnale
- La differenza prodotta dalle IDCT del codificatore e del decodificatore è nota come IDCT mismatch error

# IDCT Mismatch

---

- L'errore di IDCT mismatch non è da confondere con l'errore dovuto alla quantizzazione dei coefficienti della DCT
  - Entrambe le IDCT operano sugli stessi coefficienti quantizzati
- Lo standard non definisce l'algoritmo per il calcolo della IDCT
  - Indica solo delle tolleranze espresse come errori
- La precisione finita con cui può essere rappresentata l'output della IDCT contribuisce all'errore di mismatch

# IDCT Mismatch

---

- Lo standard definisce una procedura per la misura dell'accuratezza della IDCT
- Un set di valori di riferimento (reference data set) composto da 10000 blocchi 8x8 viene fornito per il calcolo dell'accuratezza
- La propria implementazione della IDCT deve essere eseguita sullo stesso data set
- Due insiemi di valori (pel: picture element)
  - Reference pel  $\rightarrow f_k(y, x)$
  - Reconstructed pel  $\rightarrow \hat{f}_k(y, x)$

# IDCT Mismatch

---

- L'errore di misura tra i dati di riferimento e quelli calcolati dalla propria IDCT deve rispettare certi requisiti

$$e_k(y, x) = \hat{f}_k(y, x) - f_k(y, x)$$

- Requisiti:

$$ppe(y, x) = \max \{ |e_k(y, x)| \} \leq 1 \quad k = 1, 2, \dots, 10000$$

$$pmse(y, x) = \frac{\sum_{k=1}^{10000} e_k^2(y, x)}{10000} \leq 0.06$$

# IDCT Mismatch

---

## ■ Requisiti:

$$omse = \frac{\sum_{x=0}^7 \sum_{y=0}^7 \sum_{k=1}^{10000} e_k^2(y, x)}{640000} \leq 0.02$$

$$pme(y, x) = \frac{\sum_{k=1}^{10000} e_k(y, x)}{10000} \leq 0.015$$

$$ome = \frac{\sum_{x=0}^7 \sum_{y=0}^7 \sum_{k=1}^{10000} e_k(y, x)}{640000} \leq 0.0015$$

# IDCT Mismatch

---

- Per limitare la propagazione degli errori di IDCT Mismatch, deve essere eseguita una codifica intra-frame (I-Frame) ogni 132 P-frame
  - Processo noto come **force updating**
- Nella pratica il force updating è raro, poiché gli I-frame appaiono con una frequenza di 1 ogni 10-20 frames
  - Semplifica le procedure di image editing, fast forward e fast reverse playback, semi-random access

# MPEG

---

## Motion Estimation

# Motion Estimation

---

- A differenza dell'algoritmo di Motion Compensation, lo standard non definisce un algoritmo di Motion Estimation
  - Motion Compensation: utilizzo del vettore di movimento durante la codifica/decodifica
  - Motion Estimation: stima del vettore di movimento
- L'algoritmo di Motion Estimation è il componente più importante nella codifica MPEG
  - Dalla sua implementazione dipendono l'efficacia (rapporto di compressione) e l'efficienza (qualità) di codifica

# Motion Estimation

---

- L'algoritmo di motion estimation determina un vettore di movimento per ogni Macro-Blocco in cui è suddiviso il frame
- Assunzione: i vettori di movimento sono costanti in un MacroBlocco
  - Un unico vettore di movimento è assegnato al macro-blocco composto dai 4 blocchi 8x8 di luminanza e i 2 blocchi 8x8 di crominanza
- La ridondanza tra Macro-Blocchi adiacenti è ridotta codificando le differenze tra i rispettivi vettori (si applica DPCM)

# Motion Estimation

---

- Algoritmo di motion estimation
  - Suddividi il frame corrente in macro-blocchi
  - Stima un vettore di movimento per ogni macro-blocco utilizzando un frame di riferimento
  - All'interno di una finestra di ricerca del frame di riferimento (area  $W^2$ ), cerca il macro-blocco che minimizza una misura dell'errore (best match MB)
  - Genera il vettore di movimento da applicare al MB del frame di riferimento per produrre il MB del frame corrente
- Per minimizzare la complessità computazionale si utilizzano strategie di ricerca efficienti

# Criterio di Matching

---

- Esistono diverse misure dell'errore che si vuole minimizzare
  - Mean Absolute Distortions (Error)
  - Sum of Absolute Distortions (Error)
  - Mean Square Distortion (Error)

$$MAD(x, y) = \frac{1}{256} \sum_{i=0}^{15} \sum_{j=0}^{15} |F_t(x+i, y+j) - F_{t-1}(x+dx+i, y+dy+j)|$$

$$SAD(x, y) = \sum_{i=0}^{15} \sum_{j=0}^{15} |F_t(x+i, y+j) - F_{t-1}(x+dx+i, y+dy+j)|$$

$$MSD(x, y) = \frac{1}{256} \sum_{i=0}^{15} \sum_{j=0}^{15} (F_t(x+i, y+j) - F_{t-1}(x+dx+i, y+dy+j))^2$$

# Criterio di Matching

- Generalmente si utilizza  $MAD(x,y)$  come misura dell'errore, poiché fornisce una rappresentazione intera dell'errore più facilmente rappresentabile
  - $x$  e  $y$  sono riferiti al upper-left corner: punto  $(0,0)$
  - $i$  e  $j$  indicano spostamenti laterali a destra e in basso all'interno dello stesso Macro-Blocco

MB del Frame Corrente

MB del Frame di Riferimento

$$SAD(x, y) = \sum_{i=0}^{15} \sum_{j=0}^{15} |F_t(x+i, y+j) - F_{t-1}(x+dx+i, y+dy+j)|$$

Spostamento Orizzontale

Spostamento Verticale

# Strategie di Ricerca

---

- Strategie di ricerca del vettore che minimizza l'errore
  - Full search
  - 2-D Logarithmic search
  - Hierarchical search
  - Diamond search
- Strategia Full Search
  - Tutti i possibili vettori di spostamento dell'area di ricerca vengono considerati dall'algoritmo
  - Alta complessità computazionale, ma data la sua regolarità è facilmente parallelizzabile
  - Trova la soluzione ottima

# Strategia Full search

## ■ Strategia Full Search

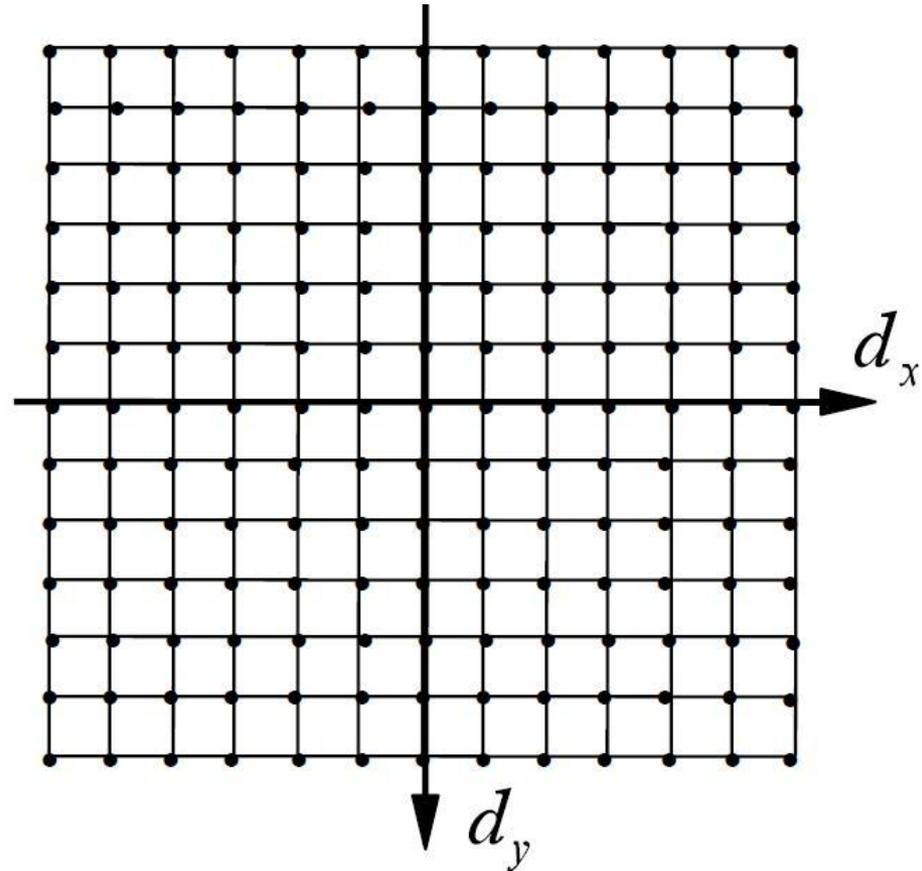
- Prova tutte le combinazioni ottenute variando  $dx$  e  $dy$
- Seleziona la combinazione che minimizza la misura dell'errore

## ■ Complessità

- Assunzione: 3 operazioni per ogni confronto (sottrazione, valore assoluto, addizione)



$$W^2 \cdot N^2 \cdot 3 \rightarrow O(W^2 \cdot N^2) \quad N = 16$$



# Strategia 2-D Logarithmic

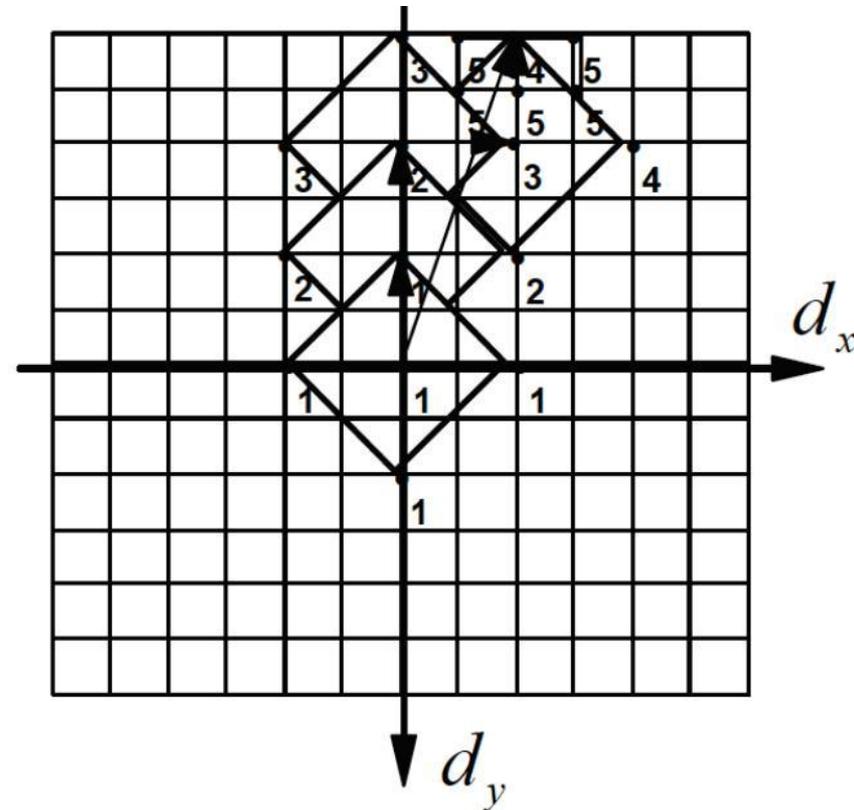
---

## ■ Strategia 2-D Logarithmic

- Trova una soluzione sub-ottima, ma richiede meno risorse della Full Search
- Versione 2-D della ricerca binaria
- Il confronto della misura dell'errore (MAD) è effettuata solo tra i MB di una area che contiene 5 punti vicini
- L'area è spostata iterativamente, centrandola nel punto che minimizza la misura dell'errore

# Strategia 2-D Logarithmic

- La misura MAD è calcolata solo per i 5 punti all'interno dell'area quadrata disposta in diagonale
- L'area viene spostata ad ogni iterazione
- Il lato dell'area viene dimezzato quando
  - Il punto selezionato si trova sul bordo dell'area di ricerca
  - Il punto selezionato si trova nel mezzo dell'area di ricerca corrente

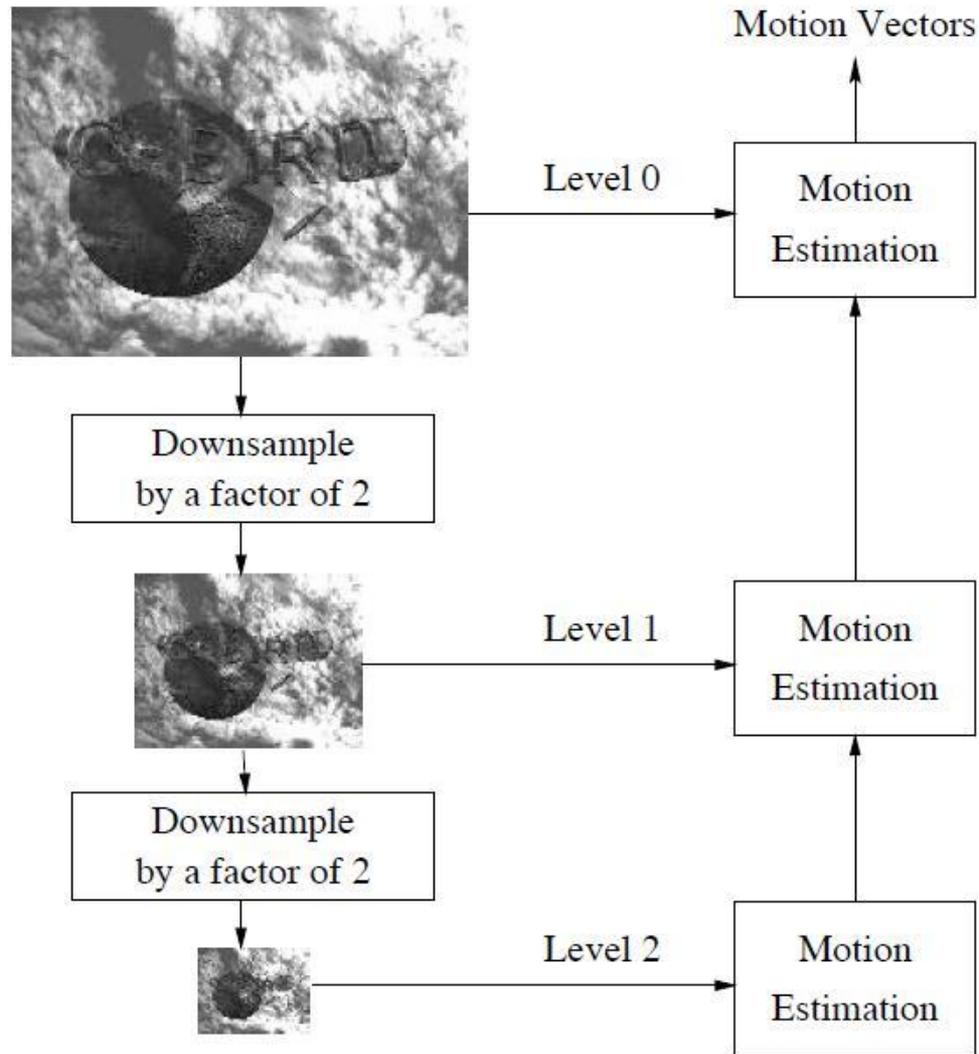


# Strategia Hierarchical Logarithmic

---

- Strategia Hierarchical search
  - La procedura di ricerca è eseguita utilizzando diverse versioni dei frame a risoluzioni diverse
  - Trova il vettore che minimizza l'errore nel frame a più bassa risoluzione
  - Si migliora la stima modificando il vettore iterativamente incrementando la risoluzione dei frame

# Strategia Hierarchical Logarithmic



# **MPEG-2**

---

**Algoritmo e differenze rispetto MPEG-1**

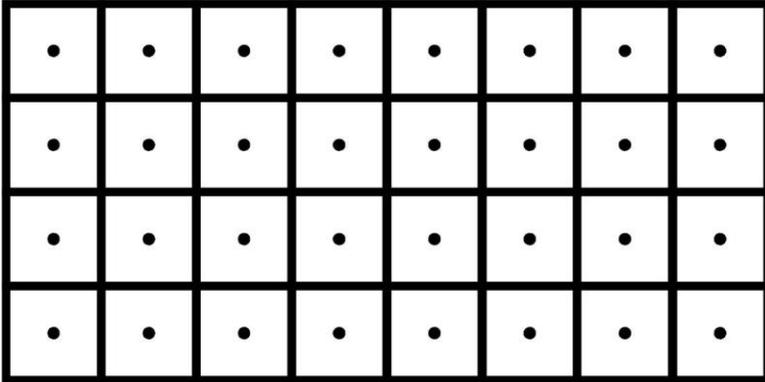
# MPEG-2 – Introduzione

---

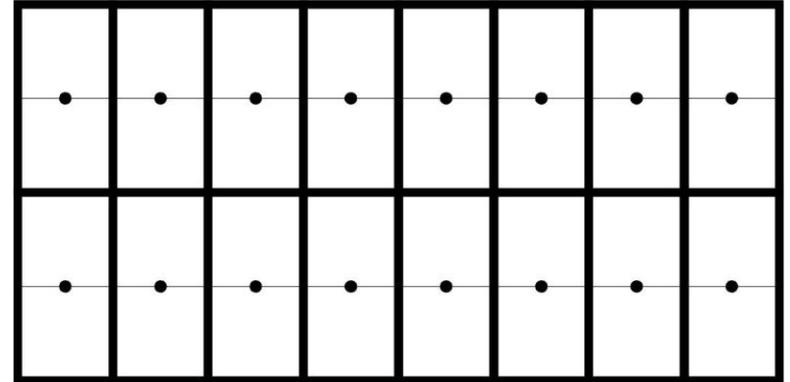
- MPEG-2 è stato sviluppato per supportare applicazioni con data rate maggiori di quelle che utilizzano la codifica MPEG-1
- MPEG-2 condivide molti degli aspetti della codifica di MPEG-1
- MPEG-2 definisce un meccanismo di codifica più efficiente per segnali video interlacciati
- MPEG-2 definisce tre formati di campionamento delle tre componenti  $YC_bC_r$ 
  - 4:2:0 → componenti  $C_b$  e  $C_r$  sottocampionati nella direzione orizzontale e verticale
  - 4:2:2 → componenti  $C_b$  e  $C_r$  sottocampionati nella sola direzione orizzontale
  - 4:4:4 → nessun sottocampionamento

# MPEG-2 – Introduzione

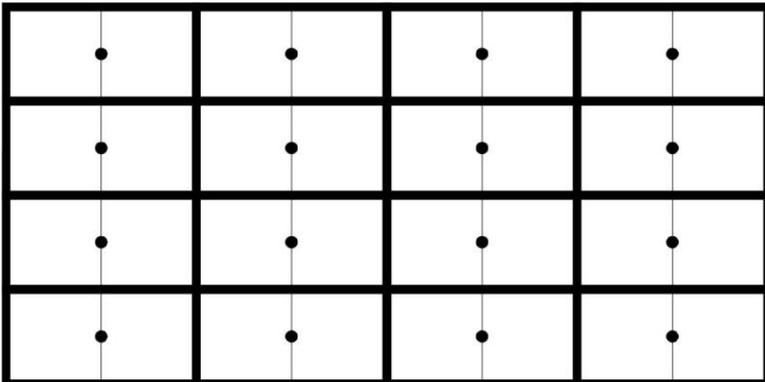
4:4:4



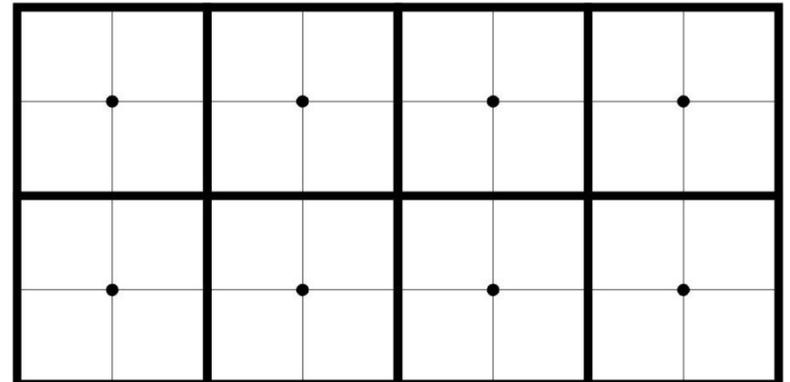
4:4:0



4:2:2



4:2:0



# MPEG-2 – Formato Video

---

- MPEG-2 specifica diversi formati video, definiti dalla combinazione di due parametri
  - **Profile:** frame rate e color resolution
  - **Level:** image resolution e maximum bit rate
- MPEG-2 definisce 5 diversi profili:
  - Simple Profile (SP)
  - Main Profile (MP)
  - SNR Scalable Profile (SNR)
  - Spatially Scalable (Spt)
  - High Profile (HP)

# MPEG-2 – Formato Video

- Esempi di profili definiti da MPEG-2
  - In tutti i profili tutti macroblocco di una slice devono essere codificati (non possono esserci buchi nel pattern di codifica)
- In MPEG-2 una slice deve essere composta da macroblocchi appartenenti alla stessa linea (la slice deve essere contigua)

Profilo	CPB	MPEG-1	SP	MP
Formato Macro-Blocco	4:2:0	4:2:0	4:2:0	4:2:0
Pictures	I, P, B, D	I, P, B, D	I, P	I, P, B
Slices	All MB	All MB	All MB	All MB
Scalabile	No	No	No	No
Precisione coeff. DC	8 bit	8 bit	8 bit	{8, 9, 10} bit

# MPEG-2 – Formato Video

- MPEG-2 definisce 4 diversi livelli:
  - Low Level (LL)
  - Main Level (ML)
  - High Level 1440 (HL-1440)
  - High Level (HL)

Level	Parametro	Limite
Main (SP@ML)	Campioni/linea	720
	Linee/frame	576
	Frame/secondo	30
	Bitrate	15 Mbps

Level	Parametro	Limite
High (MP@HL)	Campioni/linea	1920
	Linee/frame	1152
	Frame/secondo	60
	Bitrate	80 Mbos
High - 1440 (MP@HL-1440)	Campioni/linea	1440
	Linee/frame	1152
	Frame/secondo	60
	Bitrate	60 Mbps
Main (MP@ML)	Campioni/linea	720
	Linee/frame	576
	Frame/secondo	30
	Bitrate	15 Mbps
Low (MP@LL)	Campioni/linea	352
	Linee/frame	288
	Frame/secondo	30
	Bitrate	4 Mbps

# MPEG-2

---

## System Layer

# MPEG-2 – System Layer

---

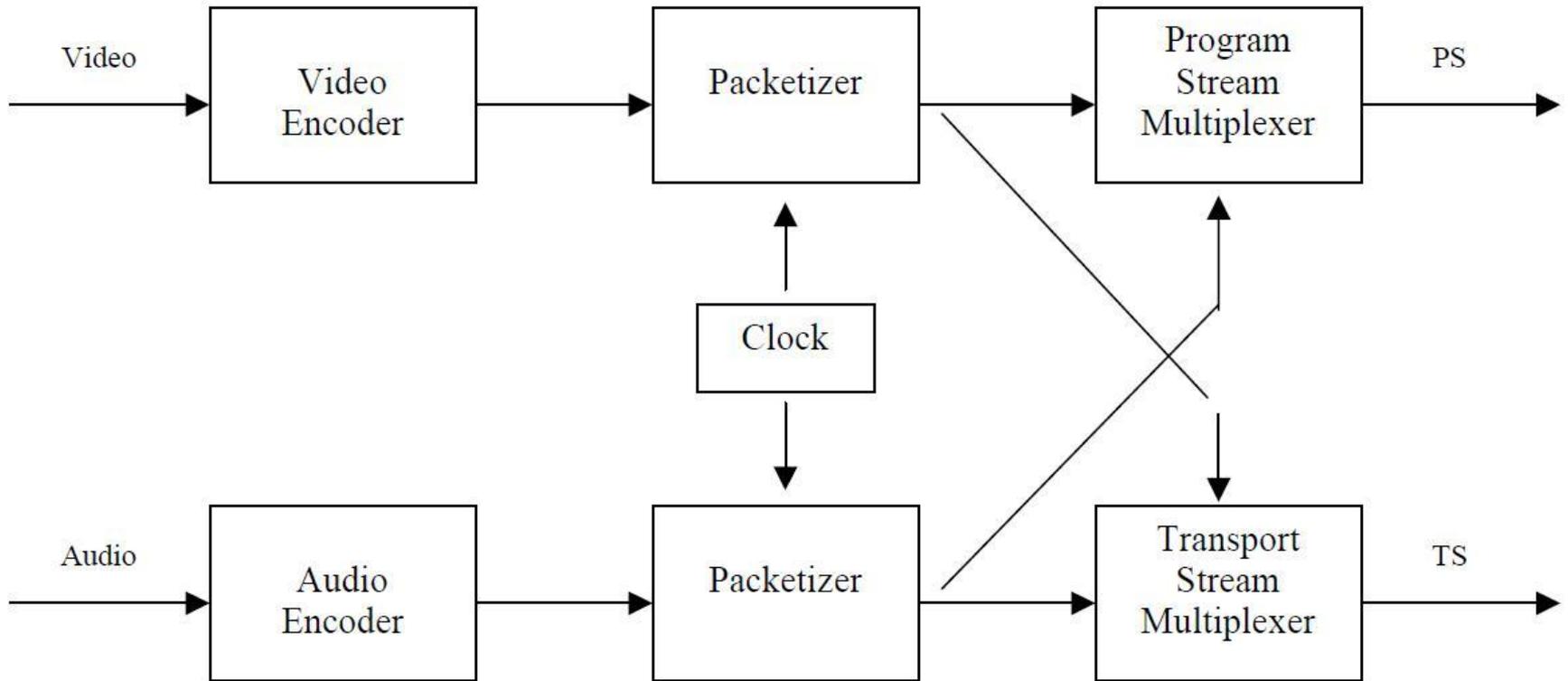
- Oltre a definire la codifica video e audio, MPEG specifica
  - la struttura dei pacchetti utilizzati per il trasporto del flusso multimediale
  - i meccanismi di sincronizzazione
  - un protocollo di trasporto che garantisce l'integrità della trasmissione del flusso multimediale
  - Indicazioni relative allo strage del contenuto multimediale
- Il livello di sistema di MPEG definisce come integrare e sincronizzare diversi flussi dati elementari (Elementary Streams)
- Elementary Stream: singolo flusso multimediale
  - Audio
  - Video
  - Dati e controllo (es. sottotitoli)

# MPEG-2 – System Layer

---

- I flussi ES vengono dapprima strutturati in pacchetti per formare il Packetized Elementary Stream (PES)
- I diversi PES contengono timestamps per la sincronizzazione dei diversi flussi
- I diversi PES vengono successivamente multiplexati per formare un unico flusso di trasporto
- MPEG definisce due alternative modalità per la generazione del flusso di trasporto
  - Program Stream (PS)
  - Transport Stream (TS)

# MPEG-2 – System Layer



# System Layer – ES

## Formato Elementary Stream

SH	SE <sub>1</sub>	EUD <sub>0</sub>	GOPH	EUD <sub>1</sub>	PH	PCE	EUD <sub>2</sub>	PDS	SDMB	SE <sub>2</sub>
----	-----------------	------------------	------	------------------	----	-----	------------------	-----	------	-----------------

Parametro	Funzione
SH	Sequence Header
SE <sub>1</sub>	Sequence Extension
EUD <sub>0</sub>	Extension and User Data 0
GOPH	Group Of Picture Header
EUD <sub>1</sub>	Extension and User Data 1
PH	Picture Header
PCE	Picture Coding Extension
EUD <sub>2</sub>	Extension and User Data 2
PDS	Picture Data containing Slices
SDMB	Slices Data containing Macro-Blocks
SE <sub>2</sub>	Sequence End

# System Layer – PES Header

## Primary PES Header



Parametro	Funzione	Descrizione
PSCP	Packet Start Code Prefix	Preambolo
SID	Stream Identification	Utilizzato per identificare il tipo di flusso
PESPL	PES Packet Length	Numero di byte del pacchetto PES
OPESH	Optional PES Header	Header opzionale del pacchetto PES
SB	Stuffing Bytes	Bit di riempimento per soddisfare i requisiti di rete

# System Layer – PES Header

## Optional PES Header (OPESH)



Parametro	Funzione	Descrizione
SC	Scrambling Control	Tipologia di scrambling utilizzata
P	Priority	Priorità assegnata al PES
DAI	Data Alignment Indicator	Utilizzato per l'allineamento dei byte
C	Copyright	Informazioni relative ai diritti di proprietà intellettuale
O/C	Original/Copy	Contenuto originale
Flags	7 Flags	Flags di controllo
PESHDL	PES Header Data Length	Numero di byte che compongono l'header opzionale
OF <sub>1</sub>	Optional Fields 1	Ulteriori campi opzionali

# System Layer – PES Header

## Optional Fields 1 (OF<sub>1</sub>)



Parametro	Funzione	Descrizione
PTS/DTS	Presentation Time-Stamps / Decoding Time-Stamps	Timestam di decodifica e di visualizzazione dell'unita codificata
ESCR	Elementary Stream Clock Reference	Tempo di arrivo previsto al sistema di decodifica (STD: System Target Decoder)
ESR	Elementary Stream Rate	Rate al quale il decodificatore STD riceve i pacchetti PES
DSMTM	DSM Trick Mode	
ACI	Additional Copy Information	
PESCRC	PES Cyclic Redundancy Check	Codice di controllo integrità (Checksum)
PESE	PES Extension	Ulteriori Estensioni

# System Layer – PES Header

## Packetized Elementary Stream Extension (PESE)



Parametro	Funzione	Descrizione
Flags	5 Flags	Flags di controllo
OF <sub>2</sub>	Optional Fields 2	Ulteriori campi opzionali

# System Layer – PES Header

## Optional Fields 2 (OF<sub>2</sub>)

OF <sub>2</sub>	PHF	PPSC	PSTDB	PESEF
-----------------	-----	------	-------	-------

Parametro	Funzione	Descrizione
OF <sub>2</sub>	Optional Fields 2	Identificatore
PHF	Pack Header Field	Header di un Program Stream
PPSC	Program Packet Sequence Counter	Numero totale di flussi multimediali
PSTDB	P-STD Buffer	Dimensione del buffer del decodificatore STD
PESEF	PES Extension Field	Ulteriori estensioni

# System Layer – Program Stream

---

- MPEG PS è stato progettato per la trasmissione/memorizzazione di contenuti multimediali attraverso canali di comunicazione ad alta affidabilità, ossia con errori di trasmissione trascurabili (es. CD-ROM)
- MPEG PS è composto da più flussi audio e/o video pacchettizzati (PES) multiplexati assieme e trasmessi attraverso pacchetti di dimensione variabile
- I flussi audio e video sono sincronizzati e possono essere decodificati e presentati in modo sincrono

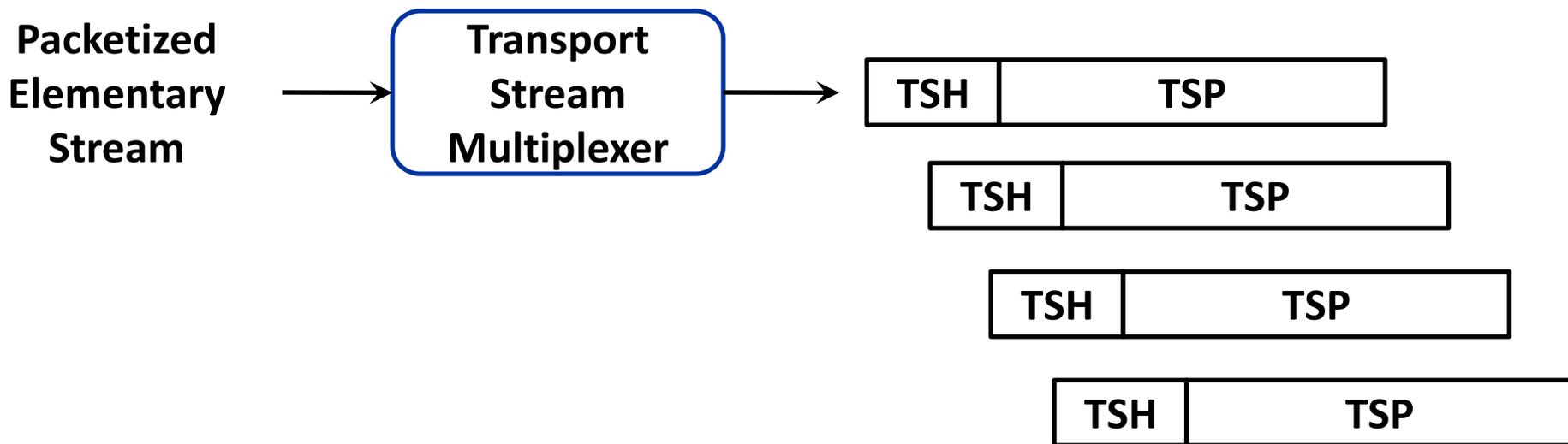
# System Layer – Transport Stream

---

- MPEG TS è pensato per la trasmissione di contenuti multimediali attraverso reti poco affidabili (errori di trasmissione non trascurabili)
- MPEG TS è composto da più flussi audio e/o video (PES), non necessariamente sincronizzati, multiplexati assieme e trasmessi attraverso pacchetti di dimensione fissa (188 bytes)
- Il flusso definito dalla modalità TS è composto da
  - Un Header lungo 4 bytes
  - Un Payload lungo 184 bytes

# System Layer – Transport Stream

---



# System Layer – TS Header



Parametro	Funzione	Descrizione
SB	Synchronization Byte	Preambolo che identifica l'inizio del pacchetto TS
TEI	Transport Error Indicator	Indica un errore nel bitstream del pacchetto
PUSI	Payload Unit Start Indicator	Utilizzato per identificare se il contenuto del pacchetto TS sia un pacchetto PES
TSC	Transport Scrambling Control	Informazione di controllo utilizzata per lo scrambling del pacchetto TS
PID	Packet Identifier	Identifica univocamente la sorgente che ha prodotto il contenuto trasportato dal pacchetto TS
AFC	Adaptation Field Control	Indica se AF è presente o meno
CC	Continuity Counter	Identificatore di pacchetto all'interno dello stesso flusso identificato da PID
AF	Adaptation Field (Optional)	Informazione opzionale non necessariamente contenuta in tutti i pacchetti dello stesso flusso identificato dal valore di PID

# MPEG-4

---

## Overview

# MPEG-4 – Introduzione

---

- MPEG-4 è uno degli standard più recenti di compressione di contenuti multimediali
- Oltre ai meccanismi di compressione pone molta enfasi sull'interattività sperimentata dall'utente
- MPEG-4 estende gli standard che lo precedono adottando una codifica basata su oggetti (**object-based coding**)
  - Garantisce maggiore efficienza (alti rapporti di compressione)
  - Il flusso video può essere codificato manipolando gli oggetti visivi che la scena
- MPEG-4 è stato definito per un ampio spettro di applicazioni con bitrate compresi tra 5 kbps e 10 Mbps

# MPEG-4 – Introduzione

---

- Funzionalità definite da MPEG-4
  - Interattività Content-Based
  - Compressione
  - Universal Access
  - Sincronizzazione
  - Flessibilità di Allocazione di Canali Virtuali
- Caratteristiche MPEG-4
  - Modalità low end-to-end delay
  - User Control
  - Trasmissione attraverso internetwork diverse
  - Sicurezza (autenticazione, cifratura, key management)
  - Trasmissione Multipoint

# MPEG-4 – Codifica VOP-based

---

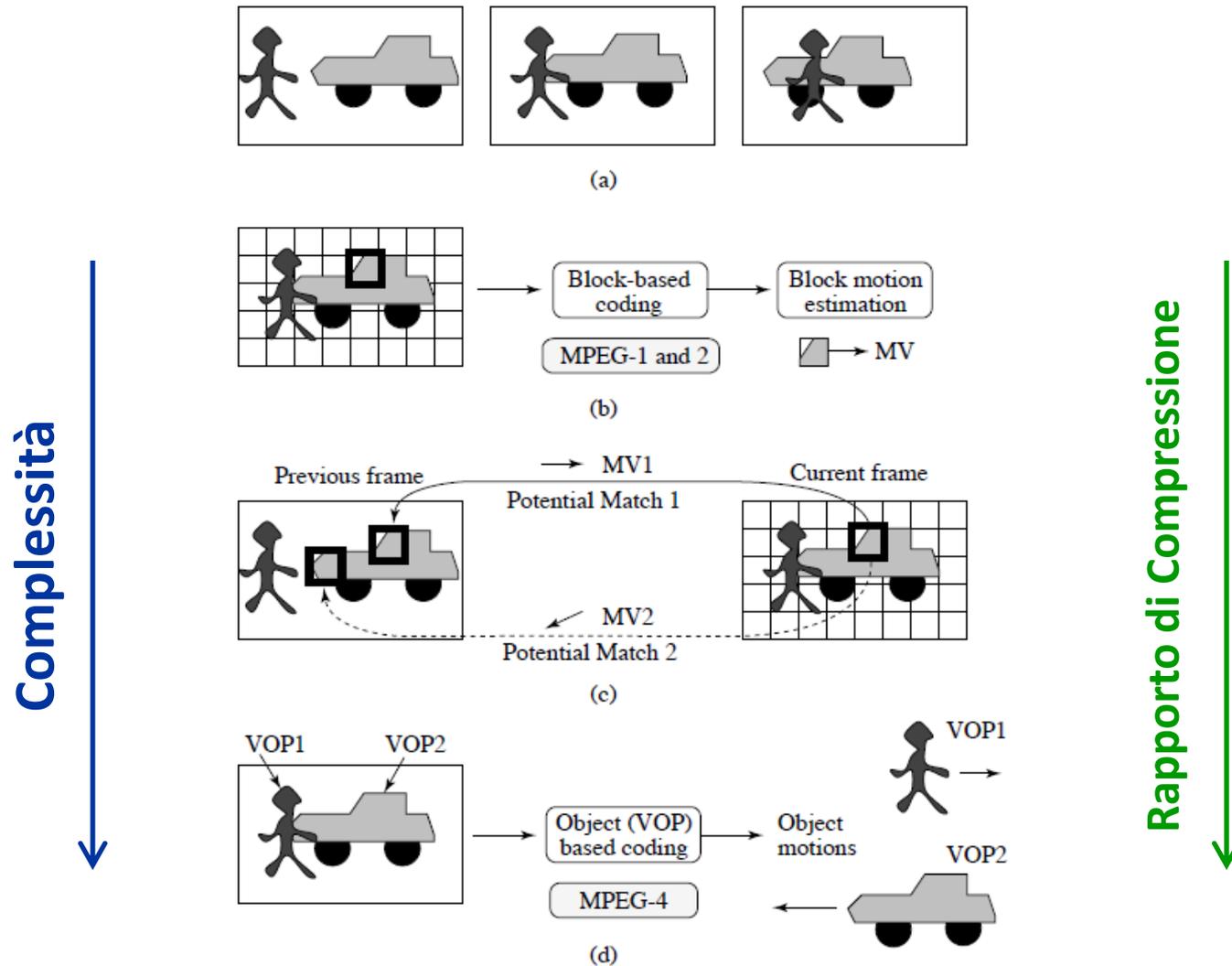
- Il formato del flusso video codificato (bitstream) definito da MPEG-4 è molto diverso da quello definito dai suoi predecessori
- MPEG-4 definisce una struttura gerarchica di rappresentazione della scena visualizzata orientata agli oggetti:
  - Video-object Sequence (VS)
  - Video Object (VO)
  - Video Object Layer (VOL)
  - Video Object Plane (VOP)
  - Group of VOPs (GOV)
- In MPEG-1 e MPEG-2 la codifica è eseguita considerando tutto il frame (**codifica frame-based**)

# MPEG-4 – Codifica VOP-based

---

- Video-object Sequence (VS): trasporta la scena completa codificata con MPEG-4, che può contenere oggetti 2-D e 3-D
- Video Object (VO): identifica un particolare oggetto della scena, che può essere una qualsiasi forma non rettangolare della scena (oggetto, sfondo, etc.)
- Video Object Layer (VOL): semplifica la codifica scalabile di un VO (codifica multi-layered)
- Video Object Plane (VOP): identifica un intero frame o un VO che deve essere codificato o con cui si vuole interagire
- Group of VOPs (GOV): istantanea di un VO in un determinato istante temporale

# MPEG-4 – Codifica VOP-based



# MPEG-4 – Codifica VOP-based

---

- La codifica VOP-based definita da MPEG-4 utilizza la tecnica di motion compensation
  - I-VOP: codifica Intra-frame applicata a un VOP
  - P-VOP: codifica Inter-frame con forward prediction applicata a un VOP
  - B-VOP: codifica Inter-frame con bi-directional prediction applicata a un VOP
- La codifica VOP-based risulta più complessa della frame-based poiché un VOP può avere una forma arbitraria
  - Si deve codificare anche l'informazione che identifica la forma oltre al vettore di movimento e i coefficienti della DCT

# Approfondimenti

---

## ■ Libri

- Chad Fogg; Didier J. LeGall; Joan L. Mitchell; William B. Pennebaker, “MPEG Video Compression Standard”. Springer 1996.
- Iain Richardson, “H.264 and MPEG-4 Video Compression: Video Coding for Next Generation Multimedia”. Wiley 2003.

## ■ Articoli

- D. LeGall, “MPEG: a Video Compression Standard for Multimedia Applications”. Communications of the ACM, vol. 34, no. 4, pp. 46-58, April 1991.
- IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on the H.264/JVC Video Coding Standard, July 2003.

## ■ Links

- <http://www.itu.int/rec/T-REC-H.264-201003-I/en>