



Università degli Studi di Bergamo



DIPARTIMENTO DI INGEGNERIA E SCIENZE APPLICATE

RETI INTERNET MULTIMEDIALI

Compressive Video

Motion JPEG, Digital Video, H.261, H.263

Il documento è adattato da materiale cortesemente messo a disposizione dal Prof. Stefano Paris e dal Prof. Vittorio Trecordi

Introduzione

- I formati di compressione per contenuti audio e video introducono nuovi elementi per aumentare il rapporto di codifica
 - Codifica nello spazio e nel tempo
 - Sfruttamento delle caratteristiche dell'occhio umano nella percezione del movimento
- Esistono numerosi standard specifici per diverse applicazioni multimediali
 - Internet streaming (Digital TV broadcasting, VoD, etc.)
 - Video conference
 - ...

Introduzione

- I segnali audio e video digitalizzati sono in genere limitatamente sensibili alle perdite
 - La compressione riduce la robustezza
- Tre tipologie di traffico audio/video
 - *Traffico conversazionale interattivo*: più parti contestualmente presenti e tempi di risposta necessariamente limitati (es. video conference)
 - Requisiti di ritardo molto stringenti
 - Qualità audio/video non eccessivamente elevata
 - *Traffico streaming non real-time*: comunicazione asimmetrica punto-multi-punto (es. video broadcasting)
 - Requisiti di ritardo non eccessivamente stringenti
 - Qualità audio/video generalmente elevata
 - *Traffico streaming real-time*: comunicazione asimmetrica punto-multi-punto (es. video broadcasting di eventi live)
 - Requisiti di ritardo stringenti
 - Qualità audio/video generalmente elevata

Introduzione

■ Standard di compressione video

Standard	Applicazione	Rate	Raster
Motion JPEG	Camcoders	30 Mbps	
Digital Video	Camcoders	25 Mbps	720 x 480
H.261	Video Interattivo su ISDN	64 kbps	
H.263	Video Interattivo su reti a pacchetto	64-384 kbps	
MPEG-1	Streaming Video (VCR quality)	Fino 1.5 Mbps	352 x 240
MPEG-2	Streaming Video (Broadcast quality)	4-15 Mbps	704 x 480
MPEG-4 (H.264)	Video su dispositivi low-rate	> 32 kbps	
HDTV	Televisione ad alta qualità	20 Mbps	1920 x 1080

Introduzione

- Standard delle dimensioni delle immagini
 - Common Intermediate Format (CIF) formato standard
 - Gli altri formati standard sono ottenuti come multipli del CIF
- Facilmente convertibili negli standard analogici PAL e NTSC
- Le componenti di cromaticità U e V hanno una risoluzione spaziale dimezzata lungo entrambe le direzioni rispetto alla luminanza Y (sotto-campionamento 4:2:0)

Formato Immagine	Matrice Luminanza	Bit Rate a 30 fps
SQCIF	128 x 96	4.4 Mbps
QCIF	176 x 144	9.1 Mbps
CIF	352 x 288	36.5 Mbps
4CIF	704 x 576	146 Mbps
16CIF	1408 x 1152	584 Mbps

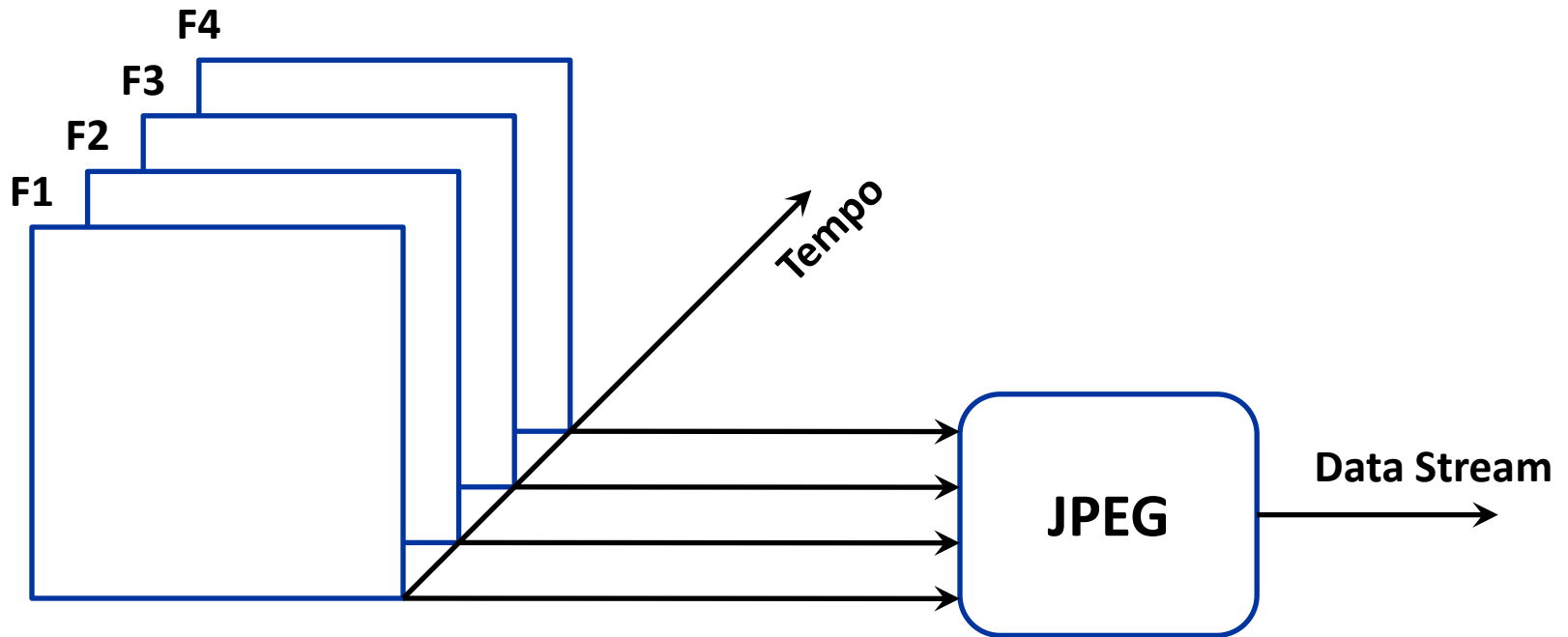
VIDEO COMPRESSO

Motion JPEG, Digital Video

Motion JPEG

- Motion JPEG è una semplice applicazione ed estensione di JPEG per realizzare video
 - Non è standardizzata
- Si codifica ogni frame della successione che rappresenta il video con uno degli algoritmi JPEG
- Il rapporto di compressione è molto contenuto in quanto non si sfrutta il contenuto informativo ridondante presente in frame successivi (codifica iter-frame)
 - Ad esempio, se ogni frame occupa in media 1 Mb e si hanno 30 fps, il bitrate risultante è 30 Mbps
- L'implementazione è molto semplice in quanto si possono riutilizzare algoritmi e strutture dati di JPEG
 - Semplifica editing, memorizzazione, reverse play

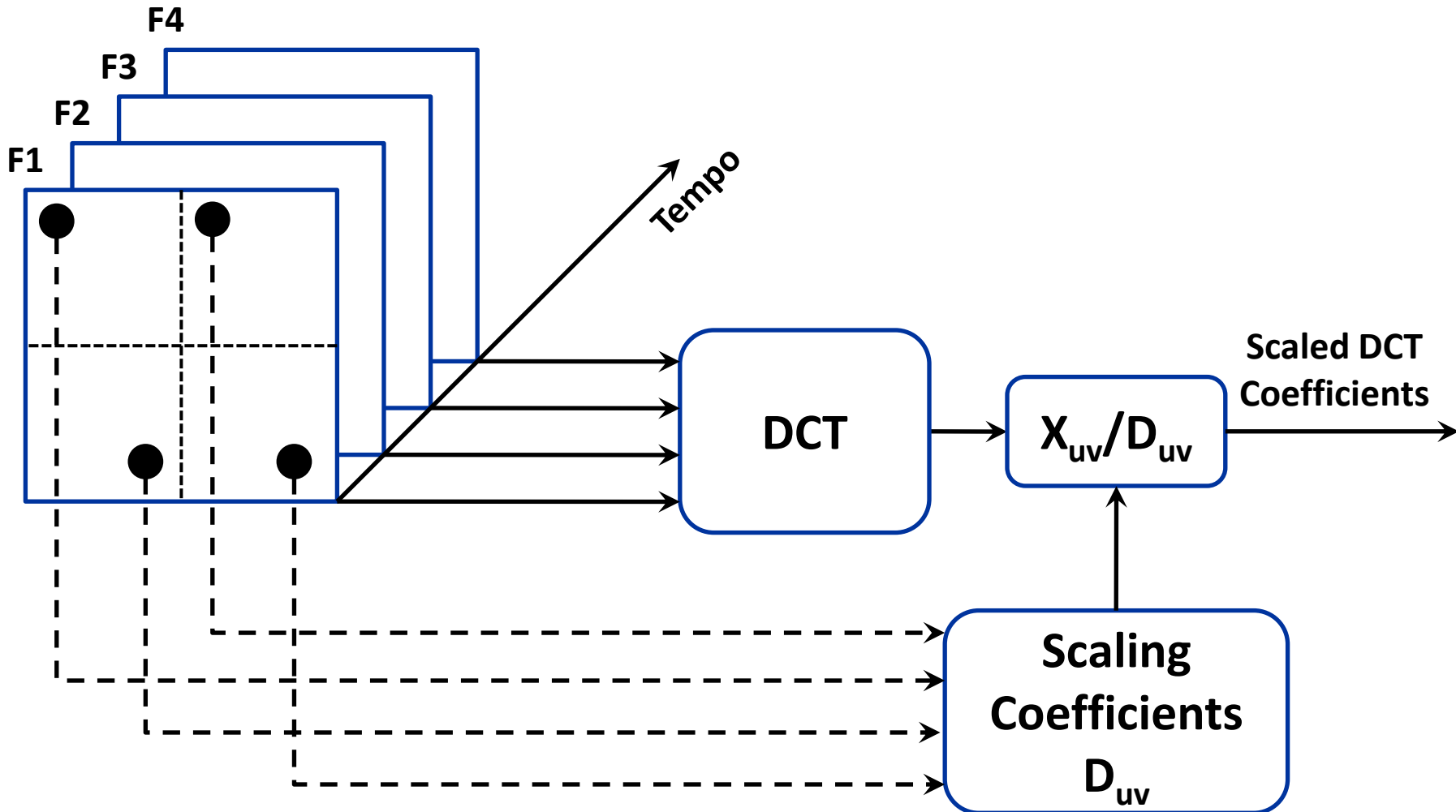
Motion JPEG



Digital Video

- Digital Video (DV) è una versione migliorata ed estesa di Motion JPEG
 - Come Motion JPEG è una delle compressioni utilizzate da videocamere
- Viene eseguita un'ottimizzazione locale della tabella di quantizzazione utilizzata per i coefficienti della DCT per codificare più efficacemente parti dei frame
- Può utilizzare frame composti da linee interlacciate
- Produce un flusso dati con un rate di 25 Mbps, che aumenta fino a 29 Mbps con l'aggiunta dell'audio e di tecniche di correzione degli errori

Digital Video



Interlacciamento

- In un video a *scansione interlacciata* si ha la trasmissione di *campi*, non di frame
 - Ogni campo è composto da metà delle linee che comporrebbero un frame
 - Si hanno alternativamente campi relativi alle linee pari e campi relativi alle linee dispari
 - Ogni campo è catturato a un istante di tempo diverso
- Ad esempio, invece di trasmettere frame a una frequenza di 50 fps, si trasmettono campi a una frequenza di 50 campi al secondo (25 campi pari e 25 campi dispari)
 - Viene trasmessa metà dell'informazione rispetto a una scansione standard (*scansione progressiva*)

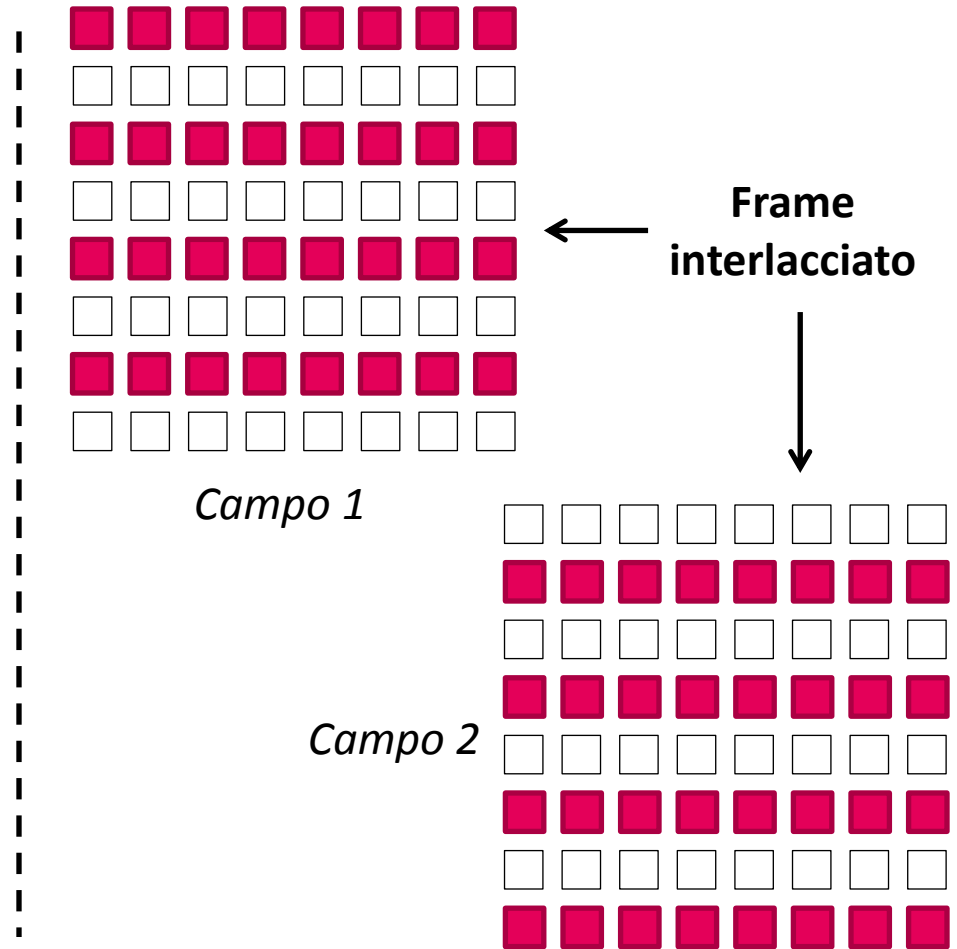
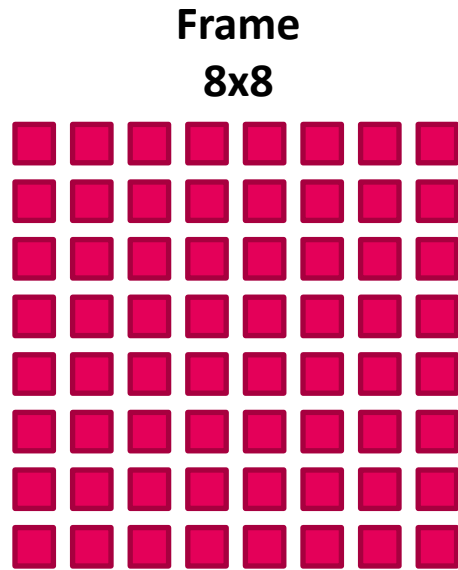
Interlacciamento

- Il cervello mantiene la percezione di un'immagine per circa 1/50 di secondo (*fenomeno phi*)
 - Una scansione interlacciata non altera la percezione del movimento presente nel video
 - Si risparmia la trasmissione di metà dell'informazione (si trasmette l'equivalente di 25 fps)
- L'operazione di *deinterlacciamento* permette di passare dalla modalità interlacciata alla modalità progressiva
- In caso di movimenti rapidi possono essere introdotti degli artefatti (*interlacing effect*)

Interlacing effect

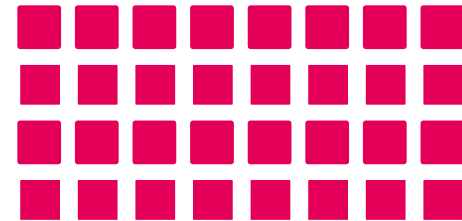
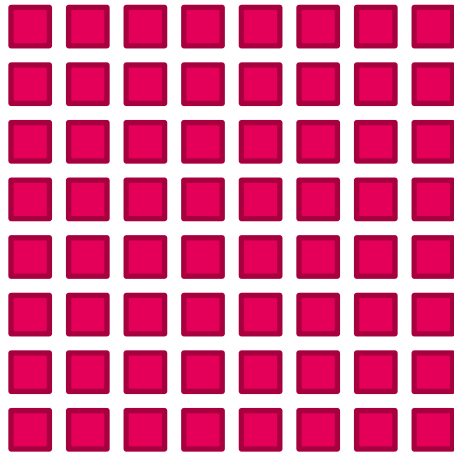


Interlacciamento

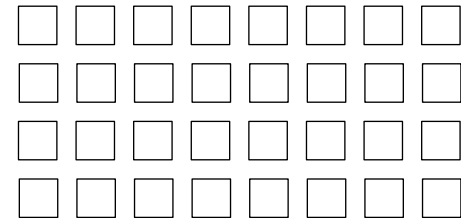


Interlacciamento

Frame
8x8



Campo 1



Campo 2

VIDEO CONFERENZA

H.261, H.263

Introduzione

- H.261 e H.263 sono due standard ITU-T per applicazioni di videoconferenza real-time in reti con rate di trasmissione ridotti
 - H.261 per connessioni a rate costante ISDN (64 kbps)
 - H.263 per reti a pacchetto
- Entrambi utilizzano risoluzioni e banda minori rispetto a MPEG-1 e MPEG-2
- Entrambi i formati di codifica sono specificati nelle raccomandazioni H.320 e H.323 come codifiche per multimedia conference

Introduzione

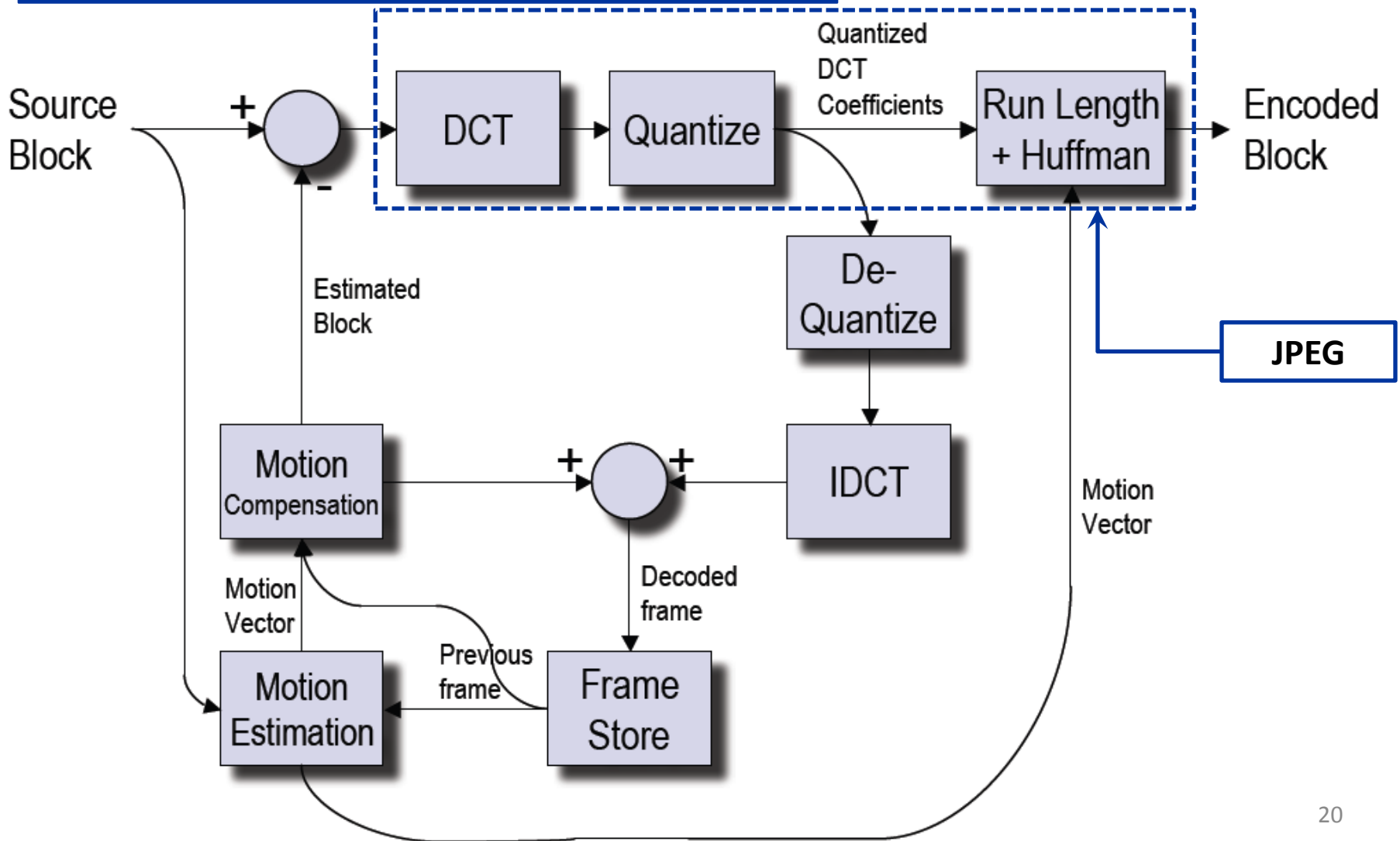
- H.261 e H.263 sono stati sviluppati per applicazioni interattive bidirezionali (o multidirezionali)
 - I ritardi in entrambe le direzioni devono essere comparabili e ridotti
 - Non si fa alcuna assunzione sulla tipologia di canale di trasmissione (full-duplex, half-duplex, etc.)
 - La qualità è generalmente inferiore rispetto alla qualità offerta dallo standard MPEG
- La complessità implementativa per la codifica e la decodifica deve rimanere contenuta
 - Codifica e decodifica simmetriche
 - Utilizzo della trasformata DCT per compressione dei frame e introduzione di una codifica inter-frame

H.261

- H.261 supporta due tipi di risoluzione: CIF e Quarter CIF (QCIF)
- Il frame rate di H.261 è pari a 30 fps
 - E' possibile ridurlo fino a un fattore 4

Resolution	Component	Horizontal	Vertical
CIF	Luminanza	352	288
	Crominanza	176	144
QCIF	Luminanza	176	144
	Crominanza	88	72

H.261 – Schema Codifica



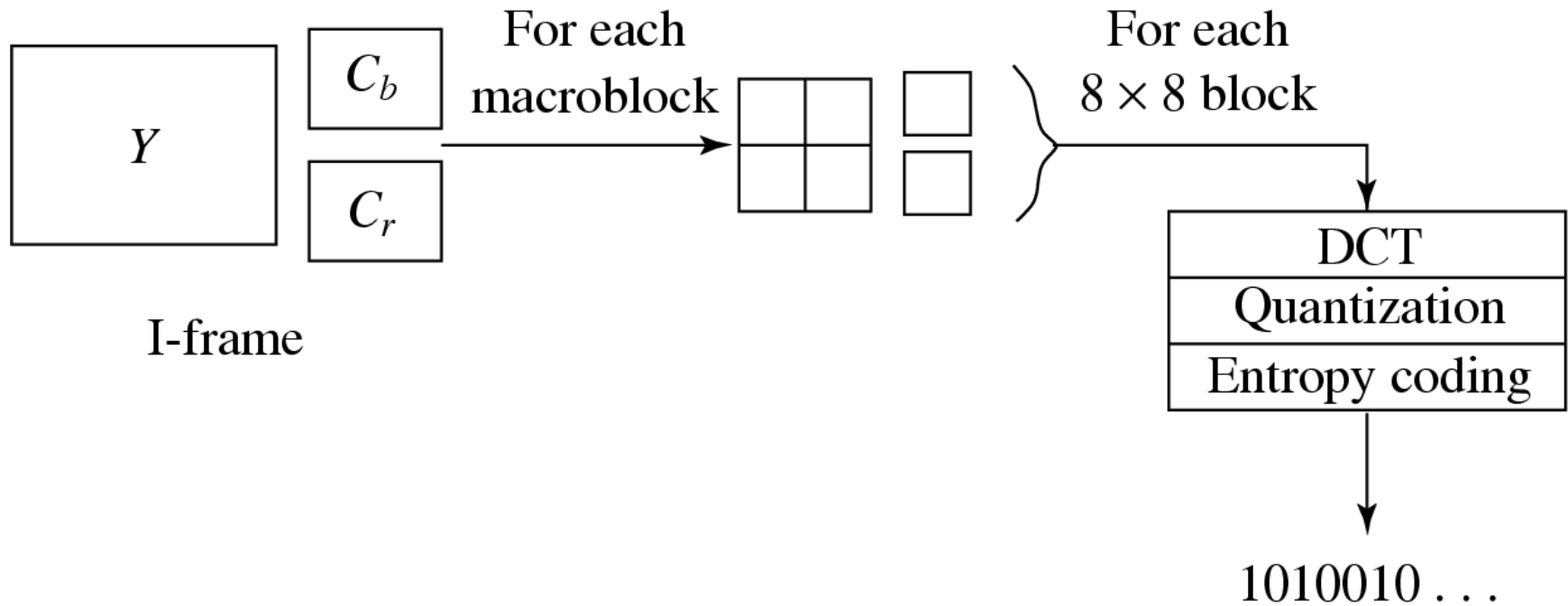
H.261 - Codifica

- I dati che compongono l'immagine raster di partenza sono suddivisi in *macro-blocchi*
 - 4 blocchi 8x8 dei coefficienti di luminanza Y
 - 1 blocco 8x8 dei coefficienti di cromaticanza U
 - 1 blocco 8x8 dei coefficienti di cromaticanza V
- I macro-blocchi sono soggetti a due tipi di codifica
 - Intra-frame: codifica spaziale
 - Inter-frame: codifica temporale (motion estimation/motion compensation)
- Tutti i macro-blocchi sono codificati Intra-frame, alcuni sono codificati sia Intra-frame che Inter-frame
 - Ad esempio, tutti i macro-blocchi del primo frame di un video sono codificati solo intra-frame

H.261 - Codifica Intra-frame

- Per ognuno dei blocchi 8x8 di luminanza e crominanza si effettua
 1. DCT: coefficienti trasformati con DCT
 2. Quantizzazione: output della DCT quantizzato
 3. Entropy coding: coefficienti quantizzati vengono codificati utilizzando uno schema lossless
- In fase di decodifica si effettuano le operazioni di
 1. Entropy decoding: inverso dell'entropy encoding
 2. Rescaler: inverso della quantizzazione
 3. IDCT: inverso della DCT
- Il processo è del tutto simile a quello eseguito da JPEG sequenziale

H.261 - Codifica Intra-frame



H.261 - Quantizzazione

- La quantizzazione in H.261 utilizza lo stesso $step_{size}$ per tutti i coefficienti DCT di posizione corrispondente tra diversi blocchi dello stesso macro-blocco
- Essendo DCT e QDCT i coefficienti DCT prima e dopo la quantizzazione si ha che

- I coefficienti DC valgono

$$QDCT = \text{round} \left(\frac{DCT}{step_{size}} \right) = \text{round} \left(\frac{DCT}{8} \right)$$

- I coefficienti AC valgono

$$QDCT = \left\lfloor \frac{DCT}{step_{size}} \right\rfloor$$

- In cui $step_{size} \in [1,31]$ intero

H.261 - Codifica Inter-frame

- Una replica di ogni blocco ricostruita a partire dai coefficienti quantizzati è inserita nel *frame store*
- Dalle repliche dei blocchi si ricrea una replica del frame originale
 - Il codificatore effettua una operazione di decodifica
- La codifica Inter-frame è una codifica predittiva
 - La replica del frame originale viene utilizzata come riferimento dall'algoritmo di predizione per predire il successivo frame

H.261 - Codifica Inter-frame

- La codifica Inter-frame si basa sulla codifica Intra-frame
 - Si effettua una codifica spaziale della differenza tra il frame corrente e quello precedente piuttosto che del frame stesso
- Generalmente l'informazione da codificare nel frame ottenuto dalla differenza di due frame successivi è consistentemente inferiore a quella contenuta in un frame considerato individualmente

$$F_D = F_2 - F_1$$

- Due frame consecutivi sono con alta probabilità simili (alta correlazione), ed è possibile codificare la loro differenza mediante un numero inferiore di bit

H.261 - Codifica Inter-frame

- Il *movimento* aumenta le differenze tra due macro-blocchi corrispondenti appartenenti a frame successivi
 - Si assume che le differenze tra due frame successivi siano dovute principalmente a movimenti della telecamera o di oggetti nella scena
- Si cerca di stimare e compensare il movimento per mezzo delle tecniche di *motion estimation* e *motion compensation*
- Viene quindi codificata con lo schema DCT - Quantizzazione - Entropy Coding (codifica JPEG sequenziale) la differenza delle due frame *dopo* la compensazione del movimento

H.261 - Codifica Inter-frame

Motion estimation

- Per ogni macro-blocco viene generato un *vettore di movimento* (o *motion vector*)
- L'algoritmo di motion estimation effettua la ricerca tra (parte o tutti) i macro-blocchi del frame precedente alla ricerca del macro-blocco più simile
 - Non è detto che il macro-blocco precedente scelto per codificare il vettore indichi del reale movimento
 - Lo standard non specifica come eseguire la ricerca
- Il macro-blocco più simile (*miglior predittore*) del frame precedente è il macro-blocco che *minimizza la differenza* (o *l'errore di predizione*) tra il macro-blocco del frame corrente considerato e quelli del frame precedente
 - La ricerca del miglior predittore è un'operazione complessa (richiede molta potenza di calcolo)

H.261 - Codifica Inter-frame

Motion estimation

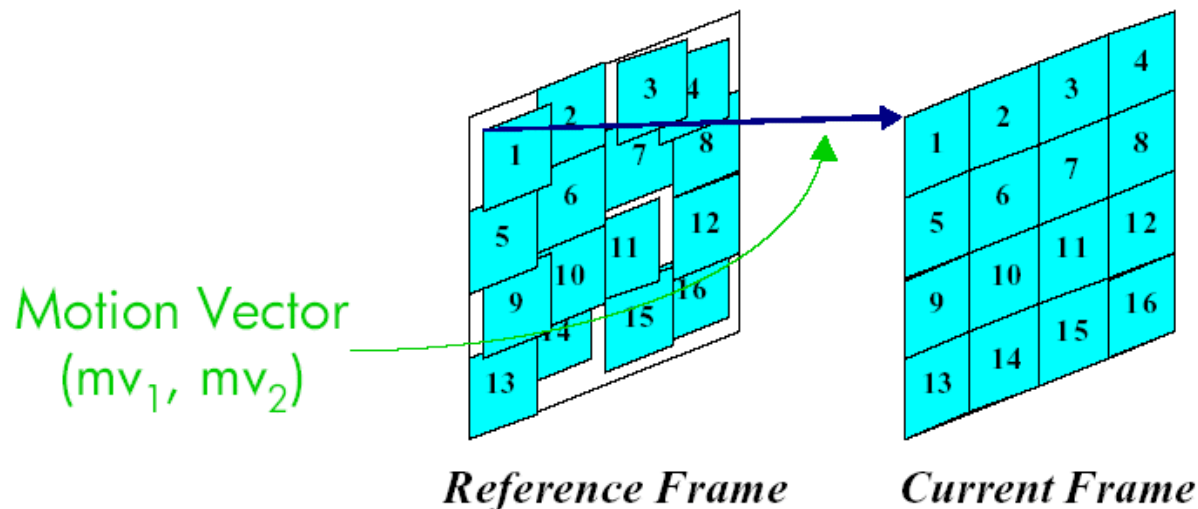
- L'algoritmo di motion estimation risponde alla domanda: qual è il vettore di movimento ottimo?
 - Selezionare una finestra di ricerca per il macro-blocco da codificare
 - Investigare tutte (o parte) delle posizioni all'interno della finestra di ricerca
 - Selezionare la posizione con il minimo *Mean Square Error*
- Il vettore di movimento indica la traslazione in termini di coordinate tra il macro-blocco del frame corrente e il macro-blocco usato per la predizione
- L'operazione di motion estimation non è standardizzata
- Il vettore di movimento viene codificato per mezzo di una codifica entropica e trasmesso



H.261 - Codifica Inter-frame

Motion compensation

- Predizione del movimento e moto-compensazione riducono l'errore di predizione, aumentando la compressione
- L'algoritmo di motion compensation utilizza i vettori di movimento per andare a sintetizzare la trasformazione tra la frame che si usa per la predizione e la frame che si sta codificando



H.261 - Codifica Inter-frame

Motion compensation

- L'algoritmo di *motion compensation* commette un *errore di predizione*
 - L'errore è la differenza nei valori di luminanza o crominanza tra il macro-blocco corrente e quello predetto
 - Errore codificato utilizzando lo schema DCT -> Quantizzazione -> Entropy Coding
- Il valore dell'errore ε determina l'informazione trasmessa per ogni macro-blocco
 - Basso ($\varepsilon < th_1$): solo vettore di movimento
 - Medio ($th_1 < \varepsilon < th_2$): vettore di movimento ed errore di predizione
 - Alto ($\varepsilon > th_2$): solo intra-frame (motion compensation controproducente)

H.261 - Codifica Inter-frame

Motion compensation



Frame(n) da codificare



Frame(n-1) utilizzata per la predizione

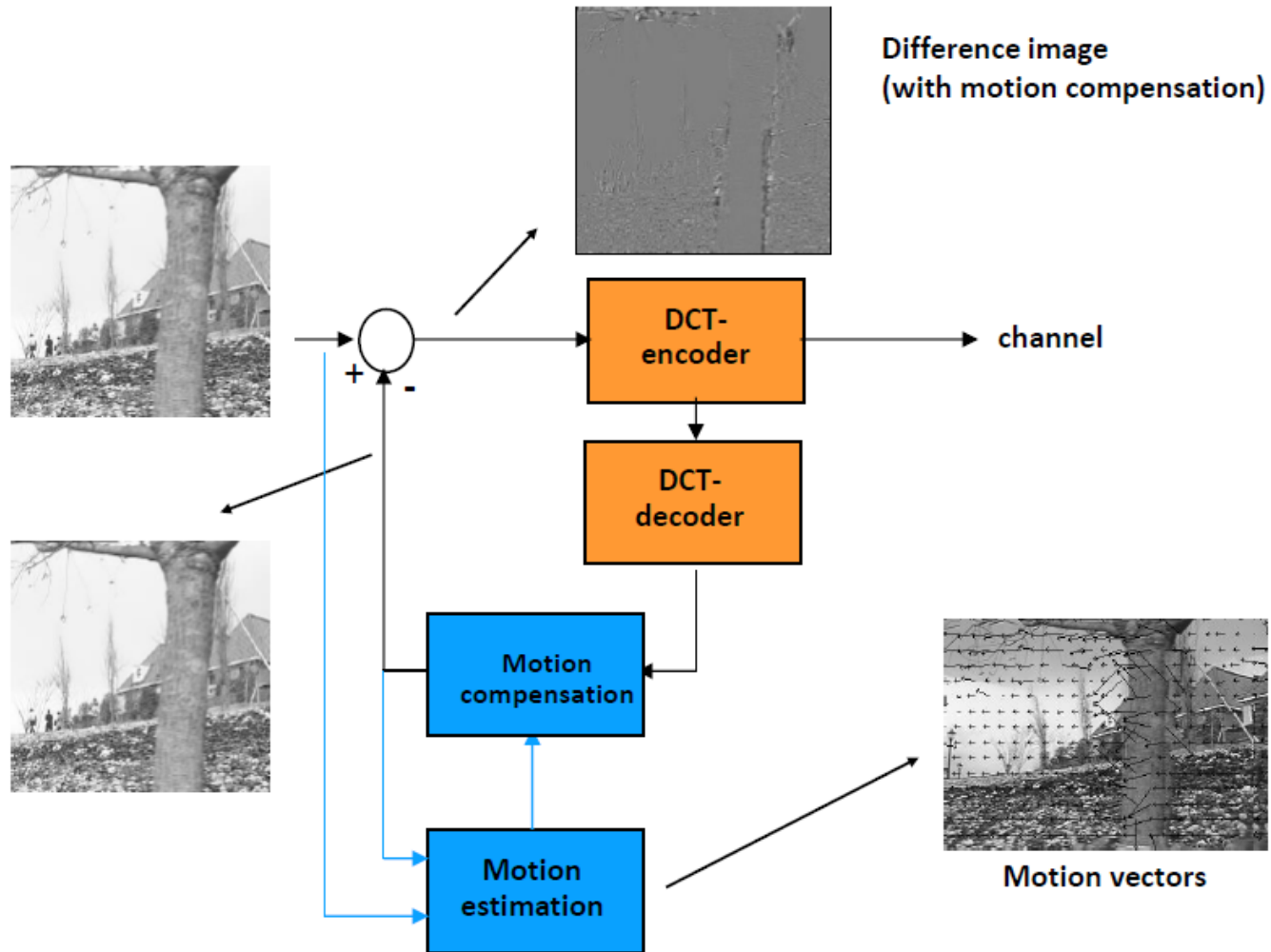


*Errore di predizione senza utilizzare
l'informazione di moto*

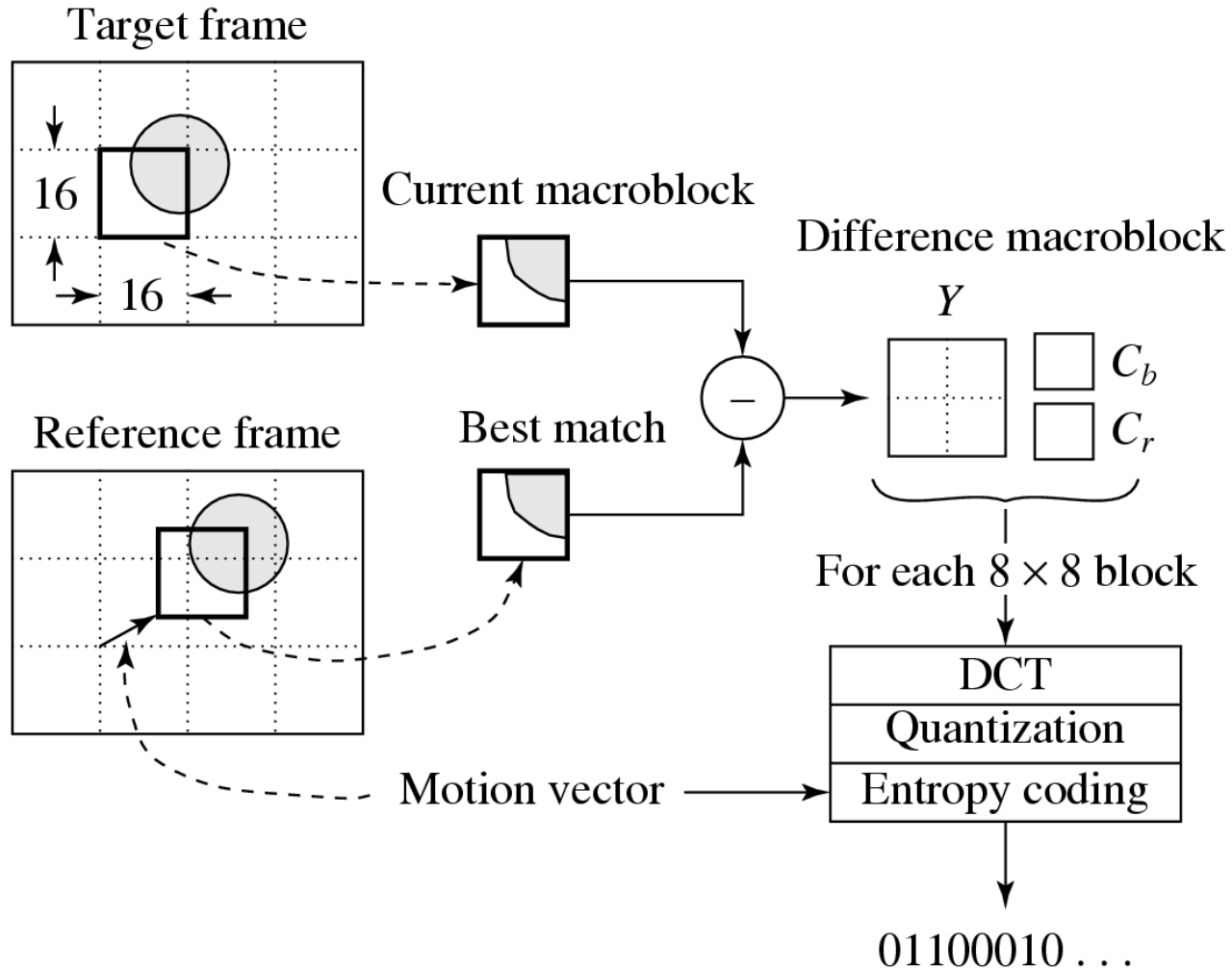


Errore di predizione utilizzando i vettori di moto

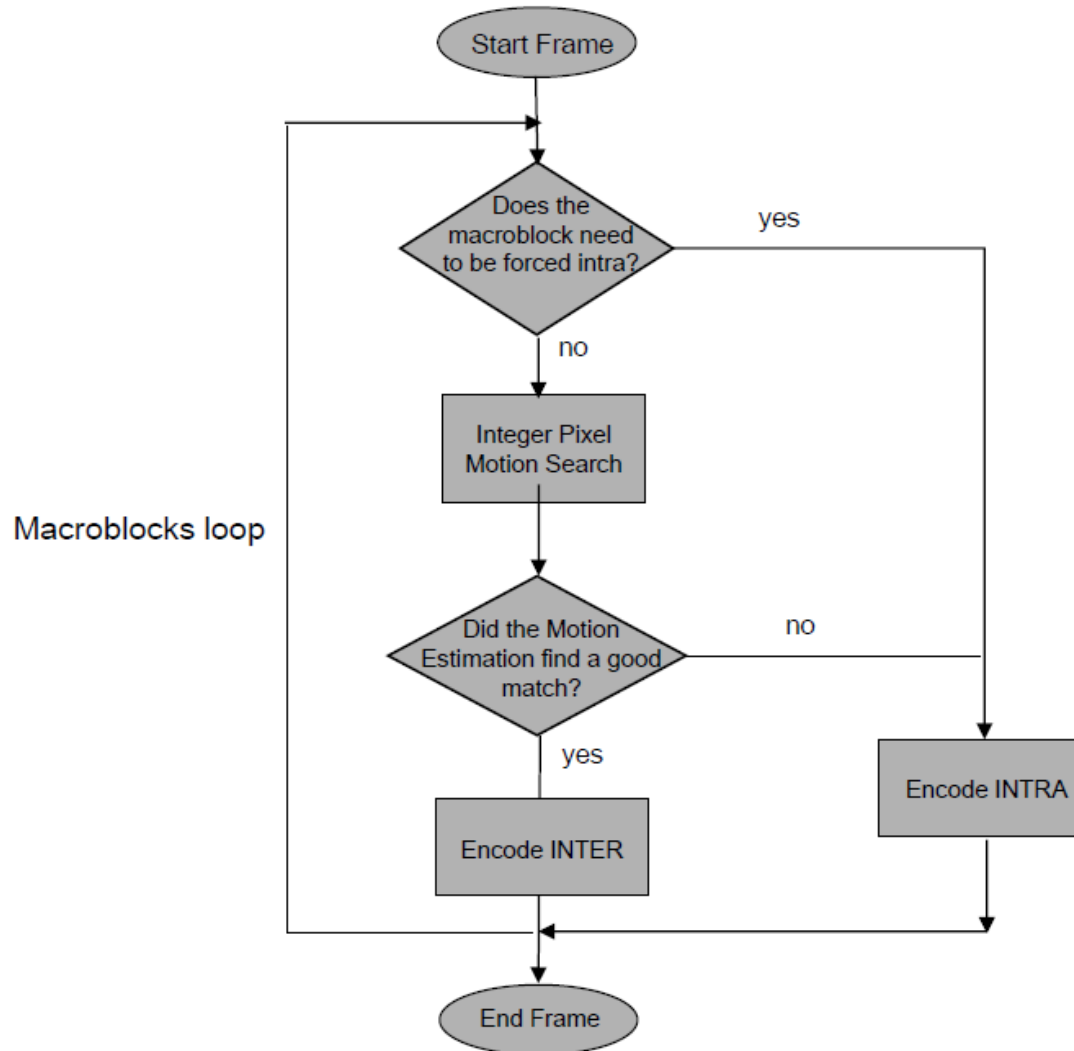
H.261 - Codifica Inter-frame



H.261 - Codifica Inter-frame



Schema a blocchi della codifica H.261



H.261 - Codifica Inter-frame

- I dati trasmessi alla destinazione sono quindi composti da
 1. Coefficienti della DCT quantizzati e codificati con una codifica lossless relativi al frame delle differenze tra due frame successivi
 2. Dati di movimento generati dall'algoritmo di predizione (*motion vector*)
 3. Tabelle di quantizzazione per luminanza e crominanza

H.261 - Buffer Management

- La quantità di informazione del flusso video codificato varia nel tempo (tra frame successivi) ed è funzione di diverse componenti
 - Complessità dei frame
 - Movimento
- Il canale su cui viene trasmesso il flusso video ha un rate di trasmissione costante (CBR)
 - E' necessario immagazzinare l'informazione codificata in un buffer prima dell'invio sul canale
 - Il buffer può essere soggetto a overflow se non gestito correttamente

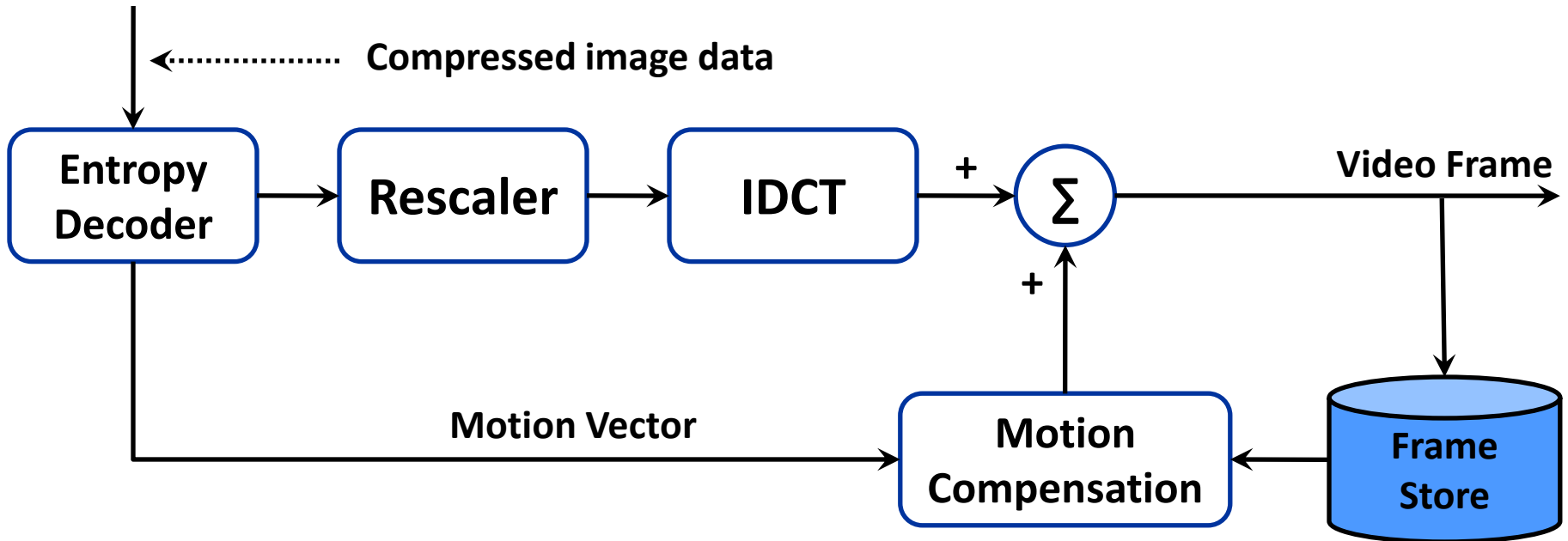
H.261 - Buffer Management

- Un algoritmo di gestione del buffer evita l'overflow
- L'algoritmo agisce sulla fase di quantizzazione modificando opportunamente i coefficienti di quantizzazione
 - Se l'occupazione del buffer supera una determinata soglia t_1 , vengono utilizzati per i coefficienti alti valori
 - Aumenta il rapporto di compressione
 - Diminuisce la qualità del video
 - Se l'occupazione del buffer scende sotto una soglia t_2 , vengono utilizzati per i coefficienti bassi valori
 - Diminuisce il rapporto di compressione
 - Aumenta la qualità del video

H.261 - Decodifica

- Durante la decodifica vengono eseguite le operazioni inverse della codifica in ordine inverso
- Vengono generati 2 frame
 - I macro-blocchi degli errori sono decodificati utilizzando lo schema Entropy Decoding – Rescaler – IDCT per generare la **replica del frame delle differenze**
 - Utilizzando l'ultimo frame decodificato e i vettori di movimento contenuti nel frame corrente si genera un **frame di riferimento** (motion-compensated reference frame)
- La somma dei due frame precedenti permette di generare una replica del frame originale

H.261 - Decodifica

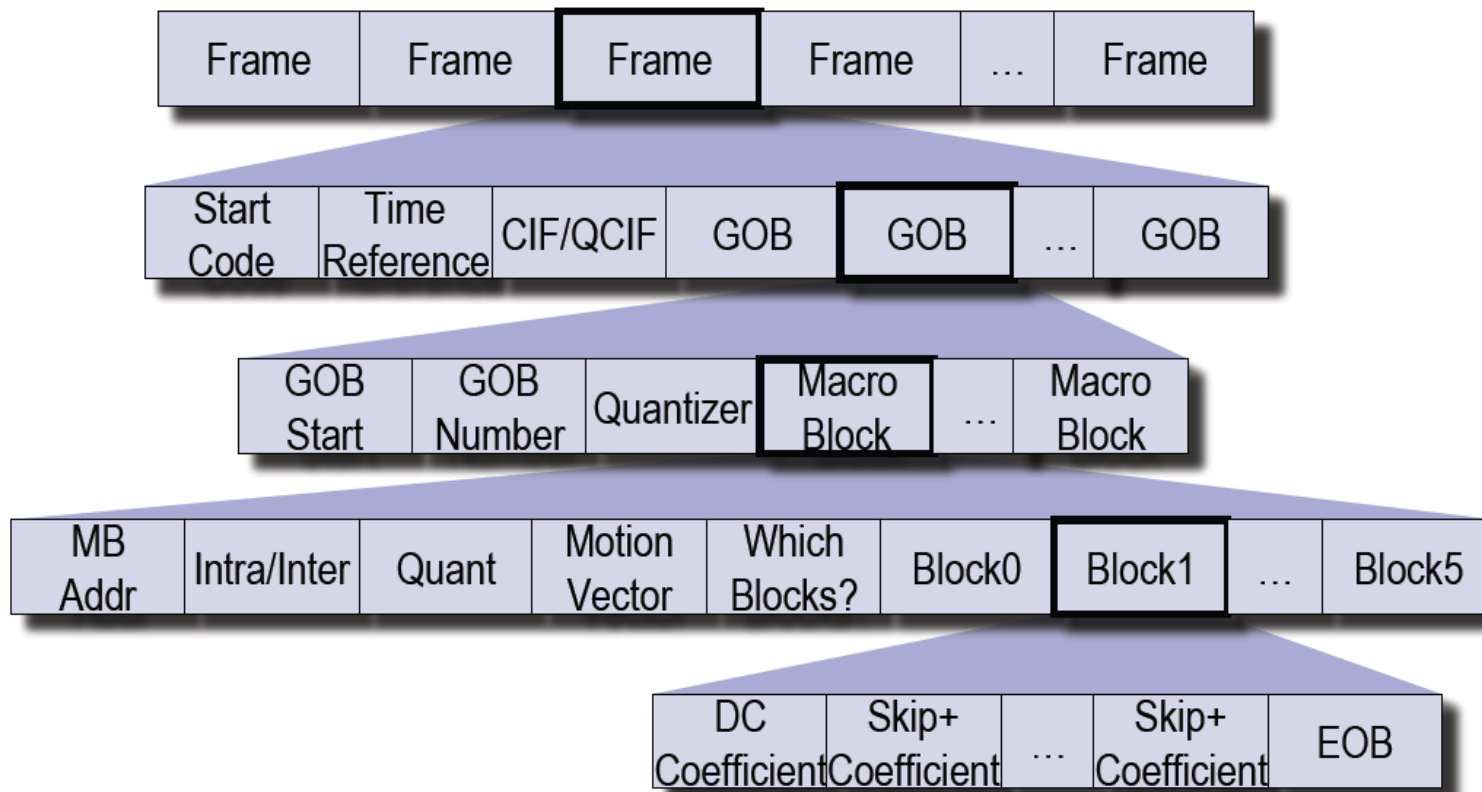


H.261 – Codifica Bidirezionale

- Finora si è considerata solo un tipo di codifica unidirezionale
 - La predizione è eseguita rispetto al frame precedente
- Si può migliorare la predizione?
 - Predizione bidirezionale
 - Utile se nel frame successivo un macro-blocco viene nascosto da un oggetto in movimento
- Compressione Bidirezionale
 - Trova il macro-blocco best-match tra i macroblocchi del frame precedente e successivo
 - Migliora il rapporto di compressione
 - Aumenta il ritardo e la complessità degli algoritmi

H.261 - Bitstream

- Formato del flusso video codificato (bitstream)



H.261 - Bitstream

■ Struttura gerarchica

- Ogni frame contiene molti GOBs
- Ogni GOB contiene a sua volta molti macro-blocchi
- Ogni macro-blocco contiene i 6 blocchi 8x8 di luminanza e crominanza (4:2:0) e il vettore di movimento
- Ogni blocco contiene i coefficienti della DCT codificati

H.263

- Evoluzione di H.261
 - Definito per applicazioni con bassi rate di trasmissione
 - Utilizzato in applicazioni Internet-based con vincoli stringenti sul rate di trasmissione (es. Modem 28.8-56 kbps)
- Il progetto è derivato da H.261, ma sono introdotte nuove tecniche per migliorare il rapporto di compressione

H.263

- H.263 utilizza *half-pixel prediction* per i vettori di movimento (a differenza della full-pixel utilizzata da H.261)
 - Il vettore di movimento può cadere tra due pixel (il cui valore è ottenuto dalla loro interpolazione)
- Sono definiti i seguenti 4 miglioramenti
 - Unrestricted Motion Vector (vettori di movimento che possono cadere oltre la fine del frame)
 - Codifica aritmetica
 - Vettori di movimento separati per ogni blocco di luminanza all'interno di un macro-blocco
 - Codifica bidirezionale
- Utilizza tutti i 5 formati CIF

Confronto H.261 – H.263

