

Capitolo 17

Teoria dell'Informazione e Codifica

17.1 Codifica di sorgente

17.1.1 Codifica audio

17.1.2 Codifica video

In accordo al metodo di realizzazione dei segnali video analogici, in cui i singoli quadri sono codificati indipendentemente gli uni dagli altri, la codifica video digitale può essere realizzata semplicemente applicando tecniche di codifica di immagine (come jpeg) ad ognuno dei quadri che costituiscono la sequenza video: un approccio di questo tipo prende il nome di *moving JPEG* o MJPEG.

D'altra parte, i quadri relativi ad istanti temporali vicini sono spesso molto simili tra loro, anche se ovviamente quanto simili, e per che estensione temporale, dipende dal tipo di filmato. Questa presenza di *memoria* nella sorgente apre la possibilità al tentativo di ridurre il tasso informativo prodotto ricorrendo a tecniche predittive, tentando quindi di *stimare il movimento* presente in quadri contigui, e trasmettere solo l'informazione necessaria a *compensare* l'errore di predizione.

Tipo di quadro Come abbiamo fatto notare al § 17.1.1.3, le tecniche di codifica predittiva sono particolarmente sensibili agli errori di trasmissione, che possono causare una perdita di sincronismo tra i predittori di trasmissione e ricezione, e quindi l'impossibilità di ricostruire la restante parte di segnale. Per questo, nella codifica video sono presenti con cadenza fissa dei quadri di riferimento in corrispondenza biunivoca con un unico quadro di partenza, detti *intracoded frames* o **I-frames**, che permettono al ricevitore di ri-partire da una condizione nota. Tra due quadri **I** sono poi presenti un certo numero di quadri **P** (*predicted*) come in fig. 4.11a, oltre che quadri **B** (*bidirectional*) come in fig. 4.11b, e che corrispondono rispettivamente alla codifica della compensazione del movimento calcolato a partire da un unico quadro precedente, o da una coppia di quadri passato e futuro.

I quadri **I** sono codificati mediante l'algoritmo JPEG, usando lo stesso coefficiente di quantizzazione per tutti i pixel delle DCT, conseguendo un rapporto di compressione relativamente basso, e sono inseriti con cadenza fissa con un

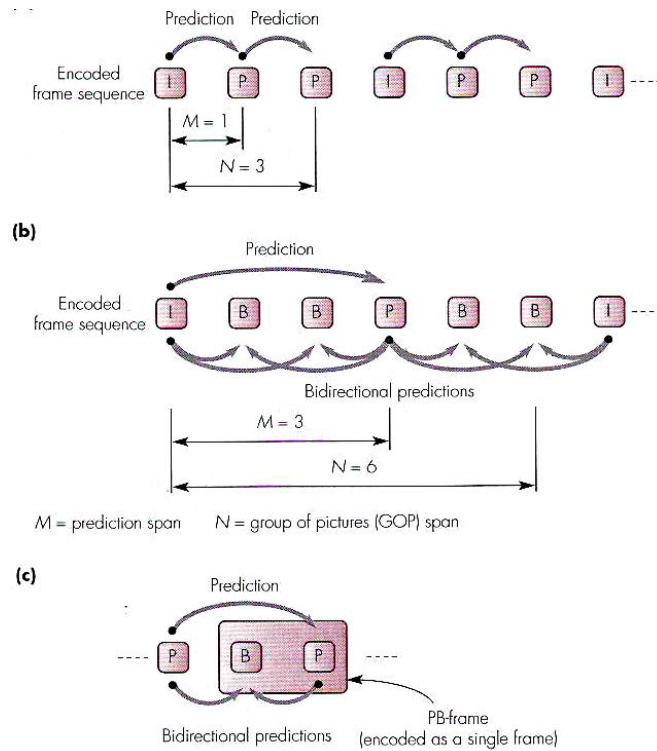


Figure 4.11 Example frame sequences with: (a) I- and P-frames only; (b) I-, P- and B-frames; (c) PB-frames.

periodo N tipicamente compreso tra 3 e 12: la sequenza di quadri compresi tra due quadri **I** è detta *group of pictures* o GOP.

Come mostrato in figura, la codifica di quadri **P** può dipendere dal quadro **I** immediatamente precedente, o dalla ricostruzione di un precedente quadro **P**, ottenendo un fattore di compressione maggiore che per i quadri **I**; la massima distanza temporale tra **P** e l'originale **I** è detta *intervallo di predizione*, indicato con M . Per realizzare la compensazione del movimento, ogni regione del nuovo quadro è confrontato con regioni *limitrofe* (riducendo così la complessità di ricerca) del quadro precedente, ma nel caso di contenuti in rapido movimento questo può non essere sufficiente. Nel caso dei quadri **B** la ricerca delle regioni simili è invece svolta rispetto ai quadri **I** (o **P**) situati sia nel passato che al futuro, migliorando la precisione della stima di movimento, e conseguendo rapporti di compressione ancora maggiori, a patto di subire un aumento del ritardo di codifica, legato al dover attendere un quadro futuro. Allo scopo di ridurre il ritardo di decodifica, la sequenza di quadri viene trasmessa con un ordine diverso da quello dei quadri originali, consentendo ai quadri **B** di essere riprodotti non appena ricevuti, e non dopo la ricezione del quadro futuro da cui dipendono. Pertanto, se la sequenza originale è ad esempio

IBBPBBPBBIBBP...

questa verrà trasmessa nell'ordine

IPBBPBBIBBPBB...

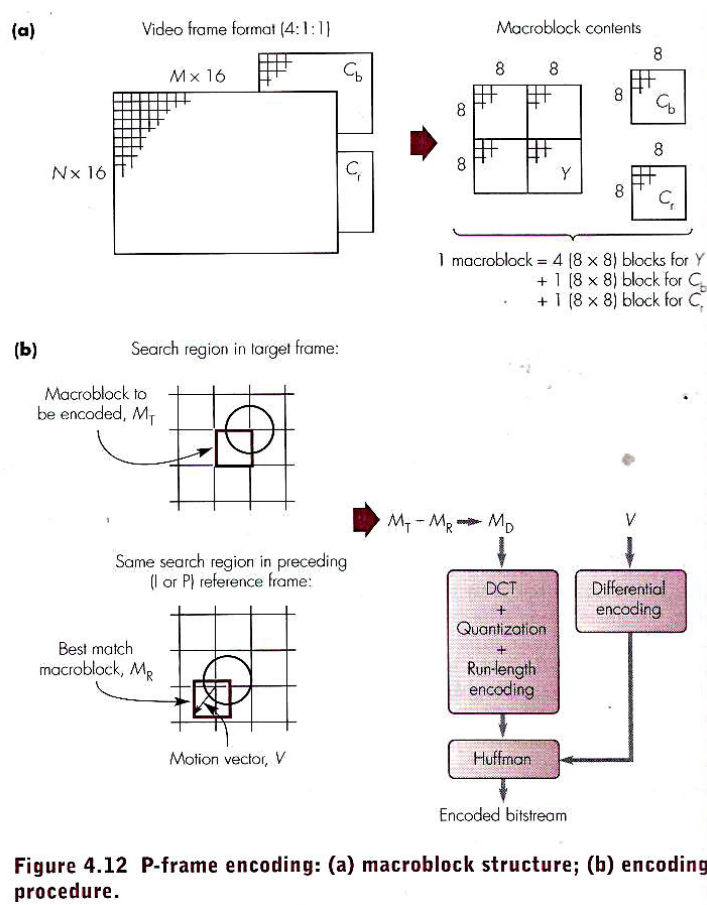


Figure 4.12 P-frame encoding: (a) macroblock structure; (b) encoding procedure.

Stima del movimento e compensazione Specifichiamo innanzitutto cosa intendere con il termine regione usato sopra nel definire il dominio dell'operazione di confronto necessario alla stima di movimento. Come mostrato in fig. 4.12a, considerando una suddivisione in componenti $Y C_b C_r$, ed un sottocampionamento 4:1:1, il quadro originale è suddiviso in macroblocchi corrispondenti a 16×16 pixel, ciascuno corrispondente a quattro blocchi di 8×8 per la luminanza, e un blocco 8×8 per ciascuna componente di cromaticità; ogni blocco 8×8 corrisponde ad un equivalente numero di coefficienti DCT, ed è individuato all'interno del quadro in base al suo indirizzo.

Nella codifica dei quadri P, ogni macroblocco M_T del quadro corrente (*target*) è confrontato pixel per pixel con il corrispondente macroblocco M_R del quadro di riferimento, e nel caso sia riscontrata una sufficiente similitudine¹ complessiva, viene trasmesso solo l'indirizzo del blocco. Altrimenti, il confronto viene ripetuto per tutti i possibili spostamenti del macroblocco target nell'ambito dei macroblocchi contigui², e qualora sia individuata una buona corrispondenza, il macroblocco viene codificato dal *vettore di movimento* V e

¹Il confronto è svolto considerando i soli valori di luminanza, e la similitudine valutata come media tra i valori assoluti delle differenze di luminanza.

²l'effettiva estensione dell'area di ricerca non è oggetto di standardizzazione, mentre lo è la rappresentazione del risultato della ricerca.

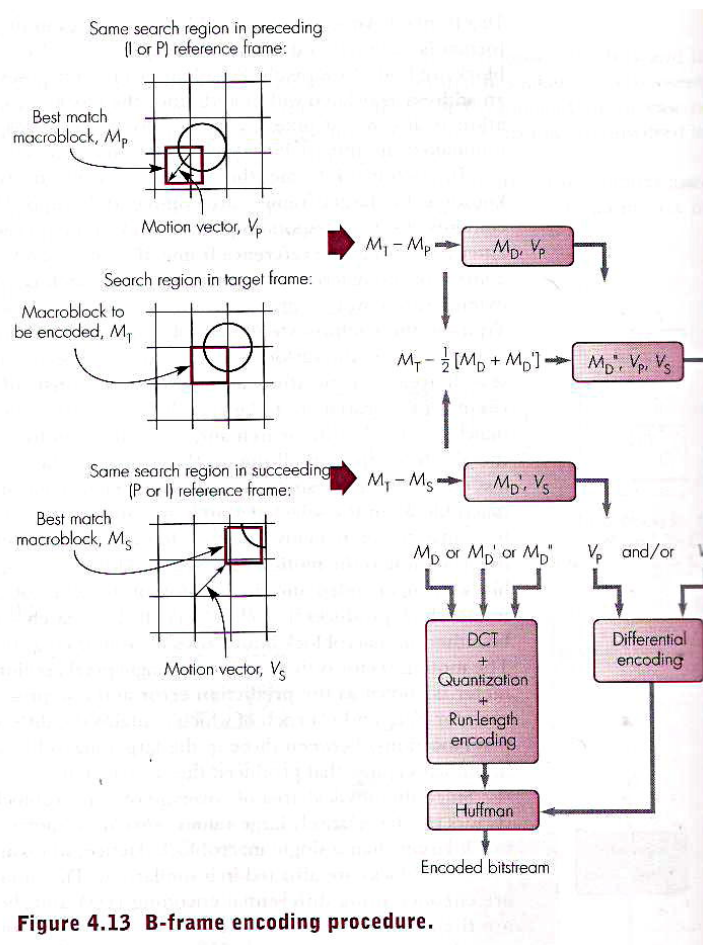


Figure 4.13 B-frame encoding procedure.

dall'errore di predizione M_D . Con riferimento alla fig. 4.12b in cui l'immagine è simboleggiata da un cerchio, V rappresenta lo spostamento da applicare a M_R per portarlo a coincidere al meglio con il quadro precedente, ed è codificato come una coppia (x, y) corrispondente ad una *risoluzione di un pixel*. Al contrario M_D è composto da tre matrici ($Y C_b C_r$) dei valori differenza tra quelli di M_T spostato di V , ed M_R . I valori di V e di M_D relativi ai diversi macroblocchi di un quadro seguono quindi due diversi percorsi di codifica.

Nel caso in cui la regione di ricerca sia estesa, i valori V possono risultare relativamente grandi; d'altra parte è probabile che macroblocchi vicini esibiscano vettori di spostamento molto simili tra loro. Per questi motivi, la sequenza dei V calcolati per macroblocchi contigui viene prima sottoposta ad un processo di codifica differenziale, e quindi i valori di differenza sono rappresentati da codeword a lunghezza variabile di Huffman. D'altra parte le tre matrici differenza sono invece sottoposte alla stessa sequenza di operazioni dei quadri **I** (DCT, quantizzazione, codifica entropica), conseguendo però un fattore di compressione più elevato, essendo il macroblocco differenza con valori quasi tutti molto piccoli.

Nel caso in cui la stima di movimento fallisca (o a causa di una estensione

di ricerca insufficiente, oppure per un reale cambio di scena), il macroblocco è codificato in modo indipendente come avviene per i quadri **I**.

I macroblocchi dei quadri **B** (vedi fig. 4.13) sono invece confrontati sia con il precedente quadro M_P che con il successivo M_S , ottenendo due possibili insiemi di matrici differenza M'_D e M''_D ed associati vettori V_p e V_S ; viene inoltre calcolato un ulteriore insieme M'''_D come differenza tra M_R e la media dei macroblocchi (spostati) di riferimento, e determinata infine quale delle tre possibilità fornisca il minimo errore di predizione. In base a questa scelta, si determina quale insieme di matrici differenza codificare, assieme ai rispettivi vettori di movimento. Nel caso prevalga la predizione basata sulla media tra macroblocchi di riferimento, il vettore di movimento complessivo può determinare un potere di risoluzione a livello di *sub-pixel*.

Questioni realizzative La fig. 4.14 riassume la sequenza di operazioni applicate alle tre tipologie di quadro **I**, **P** e **B**. Mentre nel primo caso queste seguono lo schema previsto dalla codifica JPEG, i quadri **P** possono meritare qualche commento: allo scopo di alimentare correttamente il componente di stima di movimento, il codificatore mantiene memoria del quadro di riferimento, all'inizio posto pari ad un quadro **I**, e quindi sostituito da una copia dell'ultimo quadro **P**, ottenuto risommando il quadro differenza al precedente quadro di riferimento. Lo stesso schema di calcolo è svolto nel caso di quadri **B**, tenendo ora conto anche del quadro successivo.

Rimarchiamo ora il fatto che, in funzione dell'esito del processo di stima di movimento, esistono tre diverse possibilità di rappresentazione per ogni macroblocco dei quadri **P** e **B**:

- se non vi è movimento, viene trasmessa solo la sua posizione;
- se vi è movimento e si trova un riferimento abbastanza simile, sono trasmessi il vettore di movimento e le matrici differenza;
- se non si è trovato un riferimento abbastanza simile, viene effettuata una codifica *inter* come per il caso dei quadri **I**.

Ciò determina l'esigenza di strutturare un formato di trama di dimensione (e velocità) variabile in una forma del tipo mostrato in fig. 4.14d, in cui ad ogni macroblocco è associato un tipo (**I**, **P** o **B**), il suo indirizzo nell'ambito del quadro, il coefficiente di quantizzazione dei termini di DCT, ed il vettore di movimento (se c'è). Quindi si dichiara l'identità dei blocchi presenti (che potrebbero non esserci in caso di immagini statiche), e per questi viene infine prodotta la sequenza di informazioni previste dalla codifica JPEG.

H.261

E' lo standard di codifica video definito da ITU-T a fine anni '80 per le applicazioni di videotelefonata su ISDN, ed anche se oggi tecnicamente superato, resta comunque un valido sistema di riferimento che consente la retro-compatibilità tra apparati³. Il principale vincolo al suo progetto è la necessità di produrre una velocità ridotta e limitata a multipli di 64 kbps.

³Vedi ad es. la presentazione di M.Handley, o la trattazione di Wikipedia

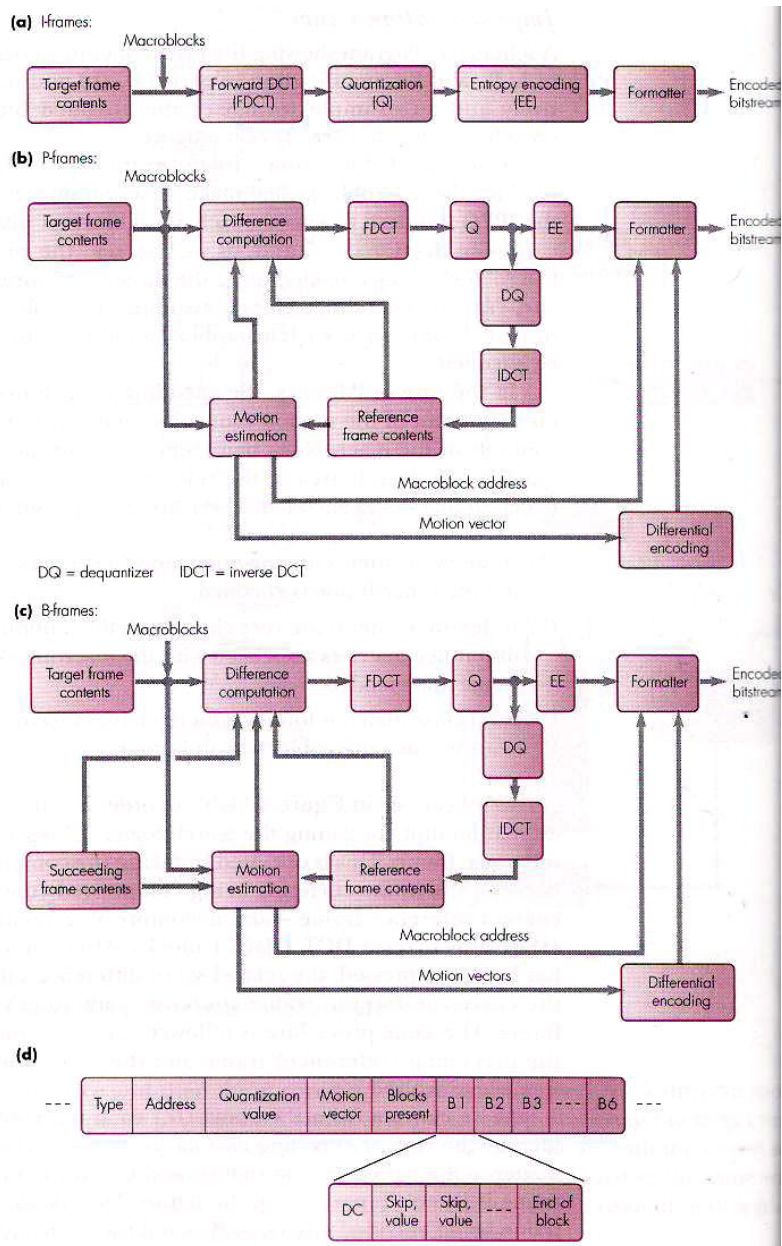


Figure 4.14 Implementation schematics: (a) I-frames; (b) P-frames; (c) B-frames; (d) example macroblock encoded bitstream format.

	Y	C_b, C_r
CIF	352 x 288	176 x 144
QCIF	176 x 144	88 x 72

La scelta del formato di immagine è limitata a quanto mostrato in tabella, mentre la scansione è non interlacciata e la velocità di rinfresco di 30 quadri/secondo per CIF oppure 15 o 7.5 per QCIF. Sono usati solo quadri di tipo I e P, con un GOP di 4 (ossia 3 P ogni I), e sono usate le procedure descritte

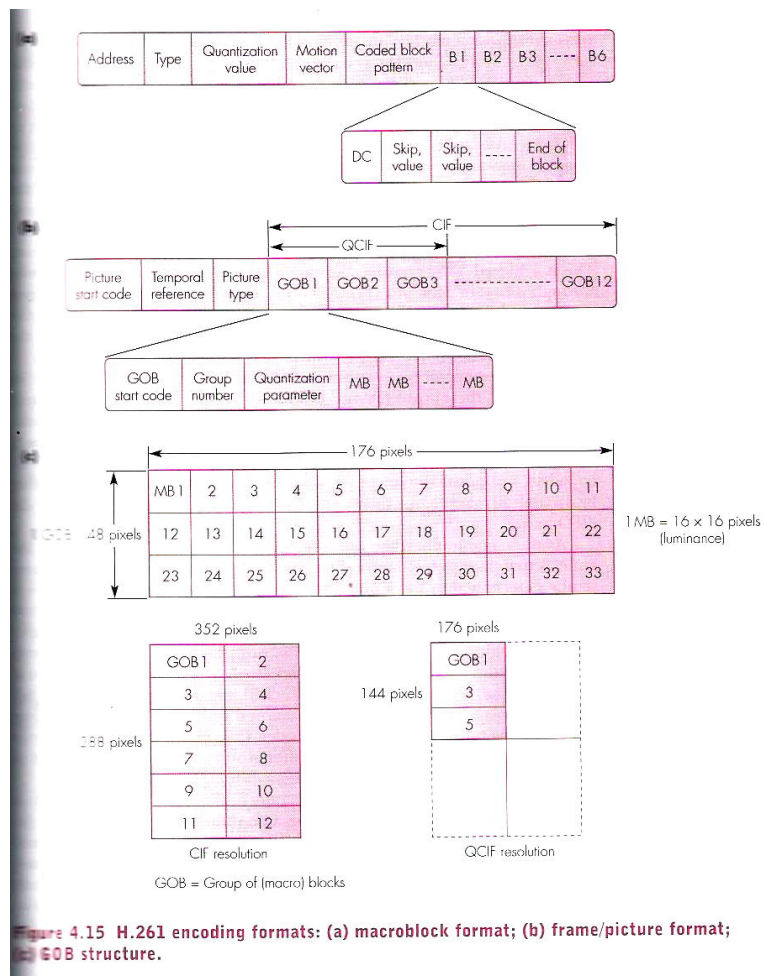
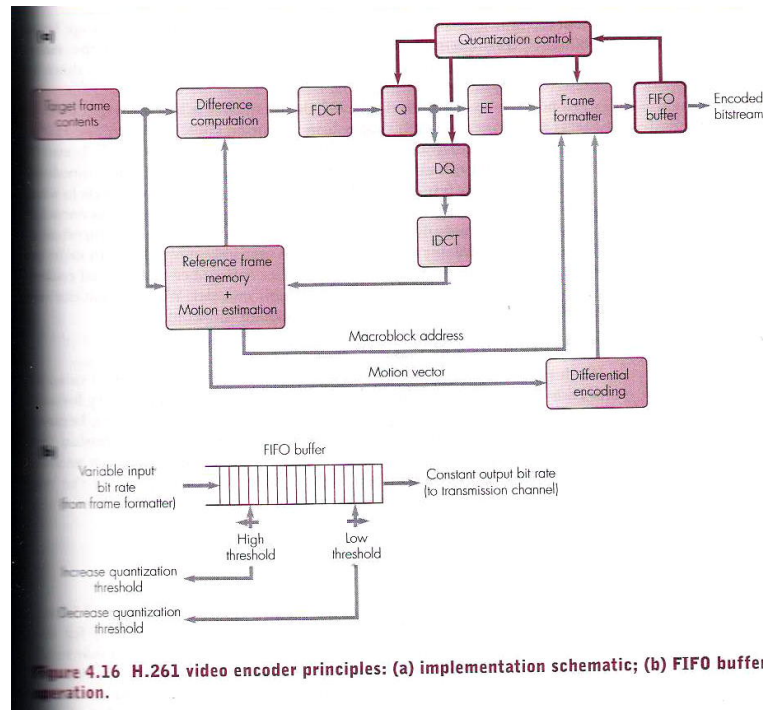


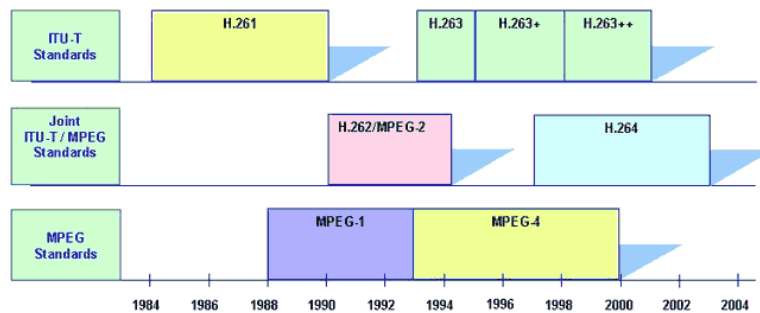
Figure 4.15 H.261 encoding formats: (a) macroblock format; (b) frame/picture format; (c) GOB structure.

alla sezione precedente per rappresentare ogni quadro nei termini di macroblocchi composti da 16x16 pixel (4 blocchi di 8x8) di luminanza e 2 blocchi 8x8 per ogni componente di colore C_b , C_r . Ogni macroblocco segue la tipica formattazione mostrata in fig. 4.15a; tre file di 11 macroblocchi sono poi raggruppati in una nuova struttura sintattica detta GOB (*Group of (macro)Blocks*), in modo da permettere l'interoperabilità tra formati come mostrato in fig. 4.15c, che mostra come la risoluzione CIF sia composta da 12 GOB, mentre la QCIF da soli 3 GOB.

Il bitstream risultante per un quadro di segnale video h.261 ha la struttura mostrata in fig. 4.15b, che evidenzia la presenza di un *codice di inizio quadro*, un *riferimento temporale* necessario alla sincronizzazione con la traccia audio, e l'indicazione del tipo di quadro (**I** o **P**); segue quindi la sequenza dei GOB, ognuno dei quali a sua volta segue la sintassi mostrata sempre in fig. 4.15b. Ad inizio GOB è presente uno *start code* scelto in modo da non poter essere presente nella sequenza di codici di Huffman che seguono, e che permette la risincronizzazione nel caso di GOB *mancanti* (vedi appresso), in modo da poter tornare a riprodurre un quadro in corrispondenza del primo GOB disponibile.



Controllo di velocità Dato che la codifica video produce una velocità di trasmissione variabile, questa può eccedere la capacità del canale a disposizione, ed un modo *drastico* per risolvere il problema è di scartare alcuni GOB. Il campo *Group number* dell'intestazione dei GOB permette quindi di collocare il nuovo GOB anche in mancanza dei suoi predecessori. Un approccio più articolato è quello mostrato dalla figura 4.16, che ripercorre le tappe già discusse e relative al calcolo del vettore di movimento ed alla codifica degli errori di predizione, ma pone in evidenza il campo di intervento di un componente di *controllo quantizzazione*, che variando l'entità dei coefficienti di quantizzazione della DCT, permette di ridurre e/o aumentare la velocità di codifica complessiva. In particolare, il controllo di quantizzazione opera in base allo stato di riempimento del FIFO *buffer* mostrato in fig. 4.16b, alimentato dal risultato del processo di codifica e formattazione video, e da cui sono prelevati i dati da inviare a velocità costante. Nel caso in cui la velocità media di codifica ecceda quella disponibile, l'aumento della occupazione del buffer determina l'aumento del coefficiente di quantizzazione, e quindi una riduzione di qualità ma anche della velocità media di codifica; ovviamente, anche l'inverso è possibile, ossia un miglioramento di qualità mediante riduzione dei coefficienti di quantizzazione, nel caso in cui la scena sia statica, e la codifica produca un basso bit rate che tende a svuotare la FIFO. Nel caso di un aumento improvviso di velocità, si possono addirittura *scartare* alcuni GOB, mentre per i successivi si adottano coefficienti di quantizzazione ridotti, comunicati anche al lato ricevente per mezzo dell'apposito campo della intestazione GOB, come mostrato in fig. 4.15b.



H.263

Anche questo definito da ITU-T a partire dal 1995, nasce per risolvere i problemi di bassa qualità dell'H.261 a velocità molto ridotte, come quelle offerte dai collegamenti modem *dial-up* precedenti all'introduzione dell'ADSL, ovvero per migliorare la gestione delle condizioni di errore possibili sia sul canale *dial-up* che *wireless*. Le specifiche originarie si sono in seguito arricchite⁴ di estensioni, favorendo l'adozione del codec da parte di altre applicazioni (inclusi i filmati di *youtube*), ed aggiungendo il supporto oltre che ai formati nativi QCIF e S-QCIF, anche a CIF, 4CIF e 16CIF. A partire dal 2003 si è formato un gruppo di lavoro congiunto tra ITU-T VCEG (*Video Coding Expert Group*) e ISO/IEC MPEG (*Moving Pictures Experts Group*), che segue la definizione del suo successore, l'H.264 detto anche AVC (Advanced Video Coding) o MPEG-4 *part 10*, determinando l'arresto dello sviluppo di H.263, che resta comunque (assieme all'h.261) supportato da un gran numero di applicazioni multimediali. Sebbene la struttura generale del codificatore e del bitstream ricalchi quella vista per l'H.261, sono state introdotte alcune novità significative, che tentiamo di elencare appresso.

Tipi di quadro In h.263 sono usati, oltre ai quadri di tipo **I** e **P**, anche quelli bidirezionali **B**, consentendo di ottenere fattori di compressione maggiori a parità di qualità percepita.

Slice I GOB sono ridefiniti come singole *strisce* di macroblocchi, quindi ad esempio per i formati CIF e QCIF un GOB è ora formato da 11 macroblocchi in fila, anziché 33 come avveniva per l'H.261.

Vettori di movimento estesi La stima di movimento dell'H.261 si arresta in corrispondenza dei bordi del quadro, per cui anche se un oggetto è solo parzialmente uscito di scena, il macroblocco corrispondente viene codificato in modalità *intra*. Al contrario H.263 permette di estendere la ricerca anche a vettori di spostamento che cadono al di fuori del quadro, alla ricerca di una corrispondenza parziale, consentendo al contempo maggiore efficienza e minor distorsione.

⁴Nel 1998 viene rilasciato l'H.263v2, noto anche come H.263+ o H.263 1998, e nel 2000 è emesso l'H.263v3 noto anche come H.263++ o H.263 2000; inoltre l'MPEG-4 Part 2 è compatibile con l'H.263, in quanto un bitstream H.263 di base viene correttamente riprodotto da un decodificatore MPEG-4.

Predizione avanzata Anziché determinare il vettore di movimento in base al confronto di un intero macroblocco, i 4 blocchi 8x8 che lo costituiscono sono confrontati in modo indipendente con il quadro di riferimento, permettendo una migliore compensazione del movimento anche per l'immagine di oggetti che non solo traslano, ma si deformano. In definitiva, sono prodotti 4 diversi vettori di movimento per ogni macroblocco.

Resistenza agli errori La presenza di un errore nella ricezione⁵ di un GOB, oltre ad impedire la corretta riproduzione dello stesso, ostacola la riproduzione anche dei quadri successivi che dipendono dai pixel presenti nel GOB, e peggio ancora l'errore finisce per estendersi anche ad altri GOB, in virtù degli effetti dell'errore sulla ricostruzione dei macroblocchi predetti in presenza di movimento. Per ridurre l'estensione temporale dell'effetto dell'errore, e non dover attendere fino alla ricezione del successivo quadro **I**, si fa uso del canale di ritorno presente nel tipo di collegamenti per cui è stato definito l'h.263, che consente al decodificatore di inviare dei NACK che notificano al mittente la coppia (quadro, GOB) per la quale si è rilevato un errore. Il codificatore è quindi in grado di valutare esso stesso le conseguenze sui quadri successivi, e può provvedere a fornire una codifica *intra* per tutti i blocchi che necessitano di essere rapidamente risincronizzati.

MPEG-1

Il *Moving Pictures Expert Group* di ISO emette una serie di standard ognuno orientato ad un particolare dominio applicativo di segnali multimediali, come

- MPEG-1 adotta un formato SIF di 352x288 pixel inteso per la *memorizzazione* audio-video a qualità VHS su CDROM, a velocità fino a 1.5 Mbps;
- MPEG-2 è orientato alla *memorizzazione e trasmissione* audio-video secondo quattro livelli di risoluzione, per ognuno dei quali diversi profili individuano tecniche alternative di codifica;

⁵Qualcuno potrebbe aver notato che nella definizione degli standard fin qui discussi, non sono previsti controlli di tipo *checksum* nel bitstream prodotto. D'altra parte essendo le informazioni codificate di natura auto-sincronizzante, la presenza di errori determina presto presso il ricevitore una condizione di disallineamento, e la decodifica di valori non previsti, come ad esempio la ricezione di vettori di movimento o coefficienti DCT fuori dinamica, o codeword di Huffman non valide, od un numero eccessivo di coefficienti. Per tale via, il ricevitore diviene in grado di accorgersi dell'errore che si è verificato.

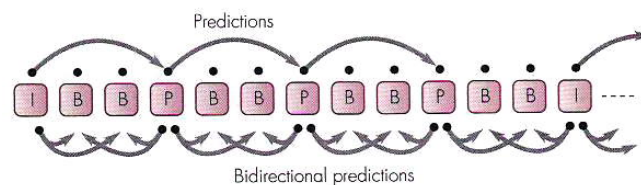


Figure 4.20 MPEG-1 example frame sequence.

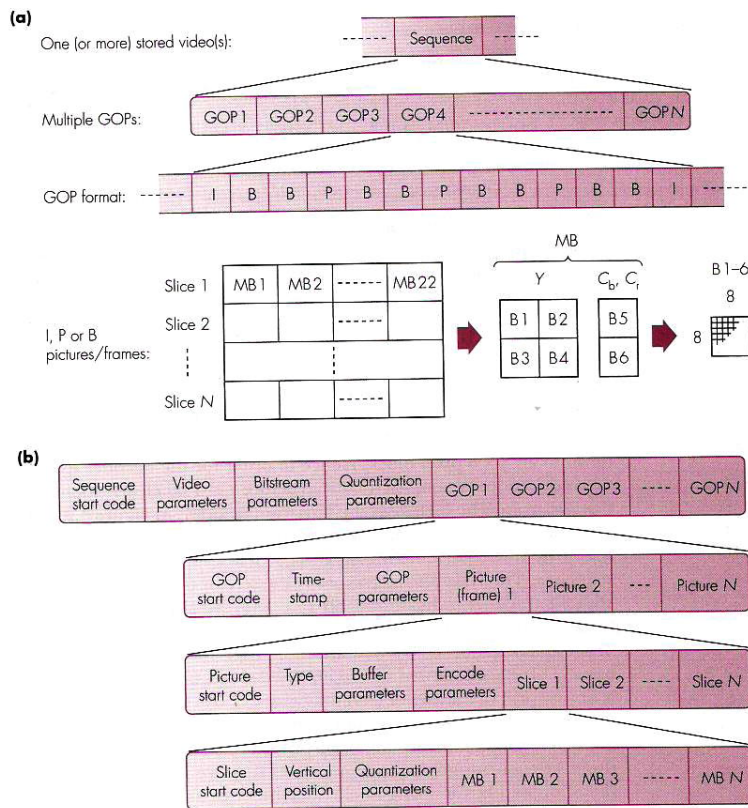


Figure 4.21 MPEG-1 video bitstream structure: (a) composition; (b) format.

- MPEG-4 è stato inizialmente concepito per applicazioni simili a quelle dell'H.263, ma il suo uso si è successivamente esteso ad un'ampia gamma di applicazioni Internet.

Anche MPEG-1 adotta tecniche del tutto simili a quelle dell'H.261, con una scansione dell'immagine progressiva ed un sottocampionamento delle componenti di colore 4:1:1, una frequenza di quadro di 25 Hz, l'adozione di quadri di tipo I, P e B, e la rappresentazione dei quadri in termini di macroblocchi composti da 16x16 pixel di luminanza, e due blocchi 8x8 per ciascuna componente di colore. Le principali differenze sono che

- possono essere inseriti riferimenti temporali *all'interno* di un quadro, permettendo al decodificatore di sincronizzarsi più rapidamente. L'intervallo tra due marche temporali è chiamato *slice* e comprende una sequenza orizzontale di macroblocchi, tipicamente che copre una intera riga, ovvero 22 macroblocchi;
- l'uso dei quadri di tipo B aumenta la distanza temporale tra i quadri di tipo P ed il loro riferimento, e quindi determina una maggiore distanza coperta dalle porzioni di immagine in movimento, cosicché l'ampiezza della finestra di ricerca adottata dal componente di detezione di movimento è stata estesa.

Livello MPEG-2	formato	bit rate (Mbps)	applicazione
Low	SIF	< 1.5	registrazione qualità VHS
Main	4:2:0	< 15	DVB
	4:2:2	< 20	
High 1440	4:2:0	< 60	HDTV 4/3
	4:2:2	< 80	
High	4:2:0	< 80	HDTV 16/9
	4:2:2	< 100	

La figura 4.21a illustra la *struttura gerarchica* del bitstream risultante, secondo il quale l'intero filmato (*Sequence*) è costituito da una sequenza di GOP, ed ogni GOP da una sequenza di quadri, ognuno costituito da una sequenza di *slice* che comprendono ognuno 22 macroblocchi, ognuno con 6 blocchi. La sezione b di fig. 4.21 entra più nel dettaglio del formato del bitstream.

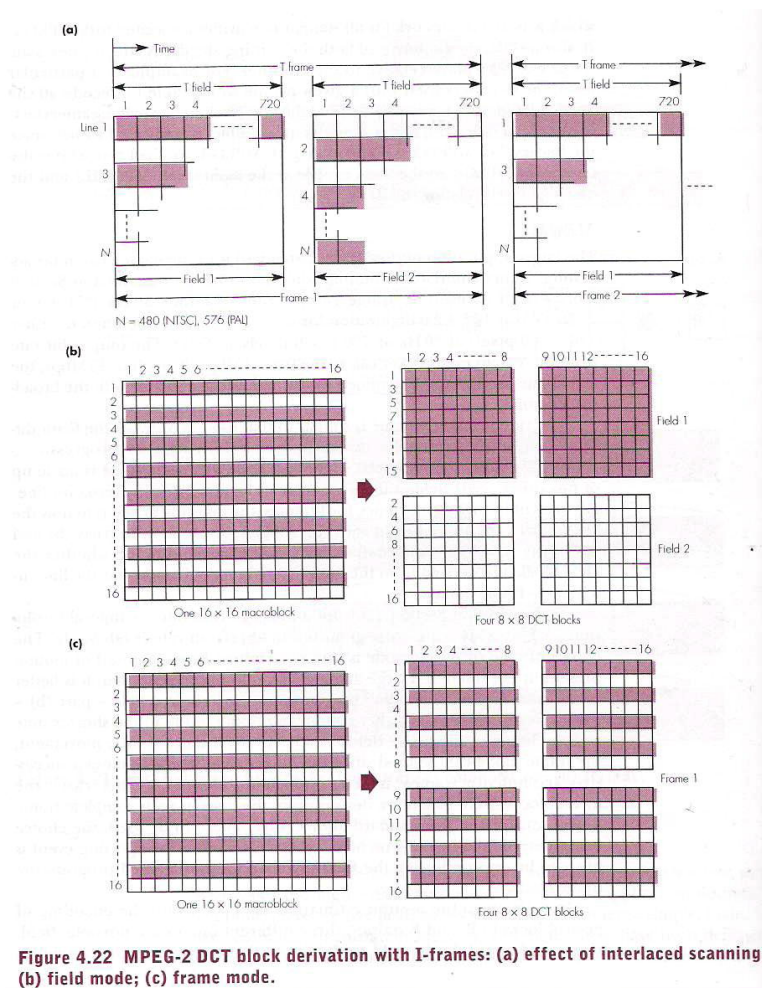
MPEG-2

Allo scopo di poter usare questo stesso standard per diversi contesti applicativi, sono stati definiti i *quattro livelli* mostrati in tabella, e per ogni livello sono quindi definiti cinque profili (*simple, main, spatial resolution, quantization accuracy, high*) in modo da permettere lo sviluppo di nuove tecnologie. Il livello *Low* è compatibile con MPEG-1. Affrontiamo ora la descrizione di ciò che è offerto dal profilo *Main* al livello *Main* (MP@ML).

MP@ML L'obiettivo è la diffusione televisiva DVB, con scansione interlacciata, risoluzione 720x576 a 25 quadri/secondo (PAL), sottocampionamento 4:2:0 per una velocità risultante tra i 4 ed i 15 Mbps. La principale differenza rispetto all'MPEG-1 è legata alla modalità di scansione *interlacciata*, in modo che come mostrato in fig. 4.22a ogni quadro è costituito da due sottoquadri (o *campi*) con le righe rispettivamente dispari e pari, ponendo la questione: come comporre i blocchi da 8x8 pixel su cui eseguire la DCT? Sono possibili due alternative:

- la *modalità campo* (fig. 4.22b) in cui i 16x16 pixel di un macroblocco sono ripartiti tenendo assieme prima le sole righe dispari del primo campo, e quindi le sole righe pari del secondo campo, oppure
- la *modalità quadro* (fig. 4.22c) in cui si usa la stessa suddivisione già vista per il caso non interlacciato, mescolando i due campi in ognuno dei blocchi.

La scelta migliore su quale tra le due modalità adottare dipende dal tipo di scena che si sta rappresentando. Se è presente molto movimento, è meglio adottare la modalità campo: infatti in tal caso i pixel di uno stesso blocco sono collezionati in un tempo pari a metà dell'intervallo di quadro, mentre nella seconda metà si collezionano i pixel della seconda serie di blocchi, ottenendo in definitiva un risultato *meno mosso*; viceversa in presenza di una scena con poco movimento, può essere adottata la modalità quadro.



Per quanto riguarda la stima di movimento, sono ora previste tre possibilità: la *modalità campo* prevede che i campi dispari usino come riferimento i campi pari del quadro precedente, ed i campi pari quelli dispari dello stesso quadro: in tal modo l'intervallo temporale su cui è valutato il movimento è metà del periodo di quadro. Nella *modalità quadro* invece, i campi pari e dispari usano come riferimento i rispettivi campi pari e dispari del quadro precedente, ed è più idoneo nel caso di movimenti lenti. Il meglio di entrambi i modi si ottiene con la *modalità mista*, in cui sono attuati entrambi gli approcci, e viene scelto quello in grado di dar luogo alla distorsione minore.

HDTV Sono definiti tre standard, ATV, DVB e MUSE, rispettivamente per il Nord America, l'Europa ed il Giappone, a cui si aggiunge la specifica HDTV di ITU-R relativa a studi di produzione e scambio internazionale, e che definisce un rapporto di aspetto 16/9 con 1920 colonne per 1152 righe (di cui solo 1080 visibili), con scansione interlacciata e sottocampionamento 4:2:2. ATV include le specifiche di ITU-R, oltre che un formato ridotto sempre con aspetto 16/9 ma risoluzione 1280x720, e che adotta la codifica video MPEG-2 MP@ML e quella

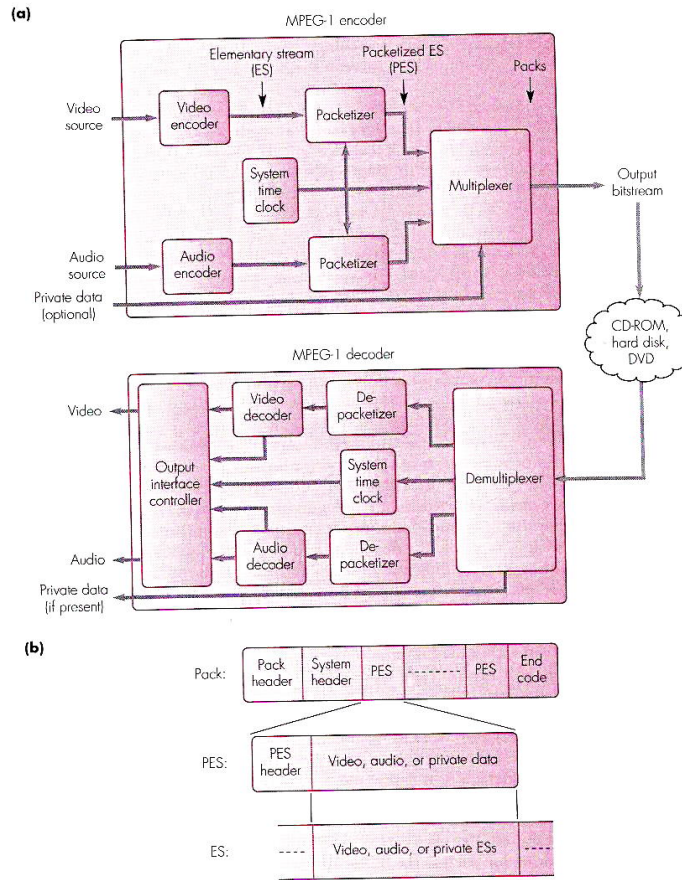


Figure 5.18 MPEG-1: (a) encoder/decoder; (b) output bitstream format.

audio AC-3. Il DVB è basato su di un rapporto di aspetto 4/3 con 1440x1152 pixel (di cui 1080 visibili), pari cioè al doppio della risoluzione PAL di 720x576. La codifica video è MPEG-2 SSP@H1440 (*Spatially Scaleable Profile at High 1440*), simile all'MP@ML, mentre la codifica audio è MPEG audio layer 2.

Contenitori

I flussi binari prodotti dai diversi codificatori (audio, video) vengono indicati come *Elementary Stream* (ES), e possono essere multiplati assieme per produrre un nuovo formato idoneo alla registrazione di un contenuto multimediale completo, eventualmente arricchito da un flusso di *dati privati*, come mostrato in fig. 5.18a per il caso di MPEG-1. Prima di effettuare la moltiplicazione, gli ES sono suddivisi in *pacchetti* di dimensione variabile, a cui viene aggiunta una *intestazione* contenente un riferimento temporale con risoluzione 33 bit prodotta da un clock a 90 MHz, il tipo di dati del pacchetto (audio, video o dati) e la dimensione del payload (tipicamente, 2048 byte). Il risultato di questa fase è un *Packetized ES* (PES), che può essere inserito assieme ad altri PES in una

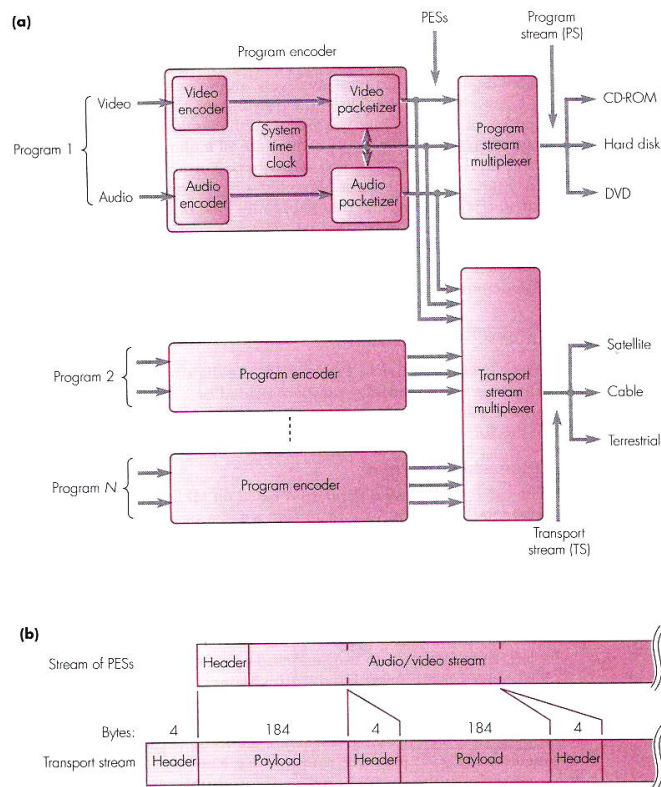


Figure 5.20 TV program multiplexing: (a) PS and TS generation; (b) TS format.

struttura di trama detta *Pack*, che può infine essere memorizzata ai fini di una successiva riproduzione.

Nel caso della trasmissione dei contenuti mediante un mezzo broadcast, tipicamente il singolo Program Stream PS (equivalente a quello prima indicato come *Pack*) viene ulteriormente multiplexato assieme ad altri, in modo da realizzare un *Transport Stream* (TS), come mostrato in fig. 5.20. In particolare, la parte b) della figura mostra come i PES siano ora suddivisi in segmenti di lunghezza fissa e pari a 184 byte, intestati con 4 byte, producendo in definitiva una struttura di trama con pacchetti di 188 byte.