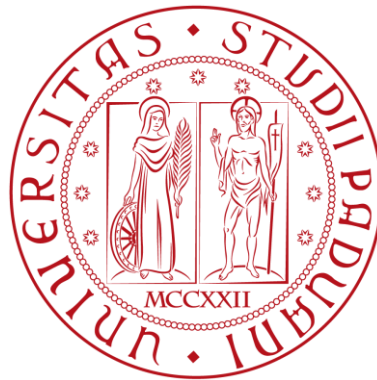


**UNIVERSITÀ DEGLI STUDI DI PADOVA**

**FACOLTÀ DI INGEGNERIA**

**LAUREA TRIENNALE IN INGEGNERIA DELLE  
TELECOMUNICAZIONI**



**SISTEMI DI TELEVISIONE  
DIGITALE – DVB**

**RELATORE: Ch.mo Prof. ROBERTO CORVAJA**

**LAUREANDO: ROMANI DANIELE**

**ANNO ACCADEMICO 2009-2010**







*A tutti coloro che mi hanno sostenuto in questo lungo percorso,*



# Indice

<b>SOMMARIO</b>	<b>XI</b>
<b>INTRODUZIONE</b>	<b>XIII</b>
La digitalizzazione del segnale televisivo	XIII
<b>CAPITOLO 1</b>	<b>17</b>
Le prime trasmissioni	17
Le trasmissioni a colori	17
Sistema PAL	20
Sistema NTSC	23
Sistema SECAM	25
<b>CAPITOLO 2</b>	<b>27</b>
Digitalizzazione segnali audio - video	27
MPEG: caratteristiche generali	29
<b>Principi della codifica video MPEG-1</b>	<b>31</b>
Tipi di immagini MPEG	32
Riordino delle immagini	33
Livelli di una sequenza video MPEG	34
Previsione, stima del movimento e compensazione	35
Controllo del bit-rate di uscita	37
<b>Codifica video MPEG-2 (applicazione broadcast)</b>	<b>38</b>
Livelli e profili dell'MPEG-2	38
MPEG-2: specifiche di previsione (immagini interlacciate)	40
<b>MPEG audio: principi di funzionamento</b>	<b>42</b>
I layer dell'MPEG audio	44
MPEG audio: formato del pacchetto	45
<b>CAPITOLO 3</b>	<b>47</b>
DVB - Digital Video Broadcasting	47
Sistemi analogici	47

<b>DVB-S: caratteristiche della modulazione</b>	<b>49</b>
Encoder: elaborazione del segnale	50
Scrambling	51
Codifica error-protection	51
Filtraggio	51
Modulazione	52
Decoder: decodifica del segnale	53
Demodulatore	53
Filtraggio e recupero del clock	54
Decoder Viterbi	54
Rilevatore del sync-byte	54
De-interleaver, RS decoder e Energy-dispersal Remover	54
Interfaccia in banda base	55
DVB-S: caratteristiche	55
<b>DVB-T: caratteristiche</b>	<b>56</b>
Encoder: elaborazione del segnale	57
Interleaver interno e Symbol Mapping	57
Adattamento trama	59
OFDM: scelta dei parametri	59
DAC e front-end	59
Decoder: decodifica del segnale	60
Prestazioni dello standard	61
<b>Caratteristiche degli standard: tabella riassuntiva</b>	<b>63</b>
<b>Conclusioni</b>	<b>63</b>
<b>Bibliografia</b>	<b>64</b>







# Sommario

Questo lavoro inizia con una breve spiegazione dei motivi e delle modalità che hanno portato allo sviluppo di un sistema televisivo digitale, definendo il contesto temporale e spaziale che ha permesso la definizione di uno standard accettato a livello mondiale: il Digital Video Broadcasting (DVB).

Nel primo capitolo viene fatto un rapido excursus storico dove viene descritto lo sviluppo della televisione, dalla sua nascita in bianco e nero, alla definizione, negli Stati Uniti, del primo standard a colori (NTSC) e la successiva diffusione nel continente europeo di sistemi a colori simili a quello statunitense (PAL, SECAM).

Il secondo capitolo inizia con il richiamo delle caratteristiche di campionamento e quantizzazione degli standard digitali dei segnali televisivi, definiti dall'ITU-R. Viene poi introdotto il gruppo di riferimento per la standardizzazione della codifica video e audio: l'MPEG (Motion Pictures Expert Group) dal quale vennero definiti prima lo standard MPEG-1, ma non adatto alle applicazioni broadcast, e successivamente l'MPEG-2, derivato dal precedente e adeguato alle trasmissioni televisive. Nel seguito del capitolo ne vengono richiamate le strutture e le funzioni principali per concludere poi con la definizione dell'MPEG-audio e delle sue caratteristiche.

Nel terzo capitolo si entra nello specifico dello standard DVB e in modo particolare di quello satellitare (DVB-S) e quello terrestre (DVB-T). Vengono per tutti e due definite le caratteristiche di modulazione, richiamato uno schema base dell'encoder e del decoder per concludere con le prestazioni.

In fine è posta una tabella riassuntiva dei principali sistemi DVB.



# Introduzione

## La digitalizzazione del segnale televisivo

La digitalizzazione del segnale televisivo e la disponibilità di nuovi mezzi trasmissivi ha creato, rispetto al passato, un'estesa gamma di servizi per la diffusione di programmi televisivi all'utente finale; in particolare tra i sistemi diffusivi oggi disponibili vale la pena di ricordare quelli via satellite, via cavo (con scarso interesse per l'Italia ma usati in molti altri Paesi), i recenti sistemi di broadcasting terrestre e i più recenti sistemi in ADSL e fibra che si appoggiano sul protocollo IP, ma che non sono stati finora completamente standardizzati. L'ottimizzazione di un sistema diffusivo si basa essenzialmente sul compromesso di garantire un'elevata qualità di servizio (bassa probabilità di errore sul segnale al ricevitore) pur utilizzando bit-rate e potenze limitate. Nella teoria delle comunicazioni le tecniche per la riduzione del bit-rate sono identificate come codifiche di sorgente: il loro scopo è eliminare nel modo più efficiente possibile tutte le ridondanze presenti nel segnale audio-video originale e vengono spesso denominate standard di compressione.

L'ottimizzazione del sistema trasmissivo propriamente detto si basa, invece, sulla scelta di un'opportuna codifica di canale (con codici correttori di errori che aumentano leggermente la ridondanza per ridurre drasticamente la probabilità di errore) e del sistema di modulazione più appropriato alle caratteristiche di propagazione del mezzo trasmissivo. Gli studi condotti nell'ambito della codifica di canale e di sorgente hanno originato e continuano a dar vita a standard caratterizzati da prestazioni sempre migliori che non sempre però sono adottati a livello universale e tanto meno garantiscono la compatibilità con terminali di utente già esistenti. Va sottolineato che, nella pratica corrente, s'identifica col nome di standard di codifica tutta la famiglia degli standard di codifica di sorgente e con il nome di standard di trasmissione la famiglia degli standard per i diversi mezzi trasmissivi (e le loro evoluzioni) con riferimento alle scelte effettuate sui sistemi di modulazione e di codifica di canale.

Per i sistemi di codifica si fa riferimento essenzialmente a quelli studiati dall'Organismo di standardizzazione MPEG (*Motion Picture Expert Group*) che sono riusciti ad ottenere un consenso a carattere universale sia per quanto riguarda diversi settori merceologici (informatica, telecomunicazioni, broadcasting e consumer) sia a livello di copertura geografica. Per i sistemi trasmissivi si considerano quelli studiati in ambito dell'organismo DVB (*Digital Video Broadcasting*) che, pur avendo caratteristiche essenzialmente europee, è riuscito a diffondere i propri standard in tutto il mondo.

Per comprendere gli sviluppi attuali della televisione digitale, è utile ripercorrere brevemente la storia, dagli esordi ad oggi, dei cambiamenti conseguenti alla

digitalizzazione. All'inizio degli anni '80 sembrava che l'evoluzione tecnologica del sistema radiotelevisivo negli Stati Uniti, in Europa e Giappone dovesse avviarsi verso un modello che privilegiava il miglioramento qualitativo dell'immagine, in termini di definizione e di resa cromatica.

Notevoli investimenti venivano effettuati già da qualche anno in Europa e Giappone nella direzione dell'EDTV (*Enhanced Definition TV*); il Giappone, inoltre, da circa un decennio aveva impegnato enormi risorse nello sviluppo della TV ad alta definizione HDTV (*High Definition TV*) che sembrava lo strumento ideale per offrire all'utente la sensazione di partecipare all'evento combinando una visione perfetta su uno schermo panoramico con l'audio stereofonico surround. Gli Stati Uniti, dal canto loro, avevano intravisto nel lancio dell'HDTV una favorevole occasione per rivitalizzare l'industria nazionale dell'elettronica di consumo, mentre l'Europa, rimasta pressoché assente dal comparto dell'informatica, non intendeva lasciarsi sfuggire di mano anche quello della televisione. Le soluzioni full digital erano allora ai primi passi, con la standardizzazione dei formati video di studio (è degli anni '80 la raccomandazione ITU-R BT.601) e dei sistemi di codifica video di tipo PCM, richiedenti velocità di trasmissione molto elevate (dell'ordine di 200 Mbit/s) anche per segnali a definizione convenzionale.

La diffusione digitale all'utente sembrava ancora un'utopia e le proposte, al cui sviluppo lavorava l'industria - MUSE (*MULTiple Subsampling Encoding*) in Giappone, MAC (*Media Access Control*) e HDMAC (*High Definition MAC*) in Europa - erano, di fatto, sistemi analogici assistiti da flussi di informazione digitale. La svolta radicale si ebbe nel 1990 quando un consorzio formato da RAI, Telettra, RTVE (l'allora Ente Pubblico radiotelevisivo spagnolo) e Politecnico di Madrid nell'ambito del progetto europeo EUREKA 256 dimostrò la fattibilità di un sistema di trasmissione interamente digitale di TV ad alta definizione, durante i campionati mondiali di calcio in Italia del 1990. Oltre 16 partite furono trasmesse in diretta, attraverso il satellite sperimentale italiano Olympus, in salette appositamente attrezzate presso le sedi RAI, con circa 800 spettatori ciascuna. Il rivoluzionario sistema era basato sull'utilizzo della trasformata DCT (*Discrete Cosine Transform*) che avrebbe costituito, quattro anni più tardi, la base essenziale dello standard MPEG-2, accettato anche dagli Stati Uniti. Il sistema studiato permetteva di trasmettere i segnali HDTV sulla stessa banda satellitare (70 Mbit/s ridotti l'anno dopo a 34 Mbit/s) richiesta dai sistemi analogici MUSE e HDMAC, ma offriva una qualità video di gran lunga superiore. Inoltre, a parità di risoluzione dell'immagine, permetteva di ridurre di un fattore dieci la potenza trasmessa dal satellite, mandando di fatto in pensione sul nascere i grandi e costosi satelliti nazionali per la radiodiffusione. Se, nel 1990, l'evento sembrò ignorato dalla comunità dei radiodiffusori e dell'industria elettronica di consumo europea, tuttavia in breve tempo la televisione digitale per l'utente domestico divenne una realtà. Il passo successivo fu compiuto con la standardizzazione del sistema di codifica video ISO-MPEG-2, sotto la guida di un altro Centro Ricerche Italiano, lo CSELT. Questo sistema, orientato al mercato di massa, permetteva ulteriori riduzioni della

banda trasmissiva (circa 5 Mbit/s per programma televisivo a definizione convenzionale e circa 19 Mbit/s per programma HDTV) e concentrava la complessità sul codificatore per ridurre i costi dei ricevitori. In tal modo la compressione del segnale video nata per trasmettere programmi ad alta qualità, si avviò rapidamente ad essere impiegato per moltiplicare il numero di programmi trasmissibili nella larghezza di banda di un canale in cui era allocato, in precedenza, un solo programma analogico. L'Europa reagì rapidamente, creando nel 1992 il Digital Louncing Group, che diventò in breve il progetto DVB (*Digital Video Broadcasting*). Dopo aver speso alcuni mesi per studiare un sistema di TV/HDTV terrestre, il gruppo per lo studio del DVB comprese la grande opportunità di business della TV digitale via satellite, dietro la spinta degli operatori della TV a pagamento: l'idea dell'alta definizione fu abbandonata anche a causa della mancanza di televisori commerciali di grandi dimensioni, a favore della possibilità di trasmettere su un unico canale molti programmi a definizione convenzionale.

Fu ancora il Centro Ricerche RAI a svolgere un ruolo fondamentale ed ad assumere il coordinamento dell'attività di definizione del primo standard di trasmissione per TV digitale: ingegneri del Centro presiedettero il gruppo di specialisti del consorzio DVB che in sei mesi, da giugno a dicembre 1993, definì lo standard di trasmissione DVB-S, una pietra miliare per la diffusione satellitare su base mondiale. Il sistema per la televisione digitale terrestre DTT (*Digital Terrestrial TV*) europea fu introdotto con lo standard denominato DVB-T e nacque circa due anni dopo, in diretta concorrenza con il sistema americano ATSC (*Advanced Television Systems Committee*): il primo a definizione normale e multicanale, particolarmente adatto alla ricezione in condizioni critiche (anche portatile), il secondo ad alta definizione, a programma singolo e meno robusto dal punto di vista trasmissivo. È dal trasmettitore Rai di Torino Eremo che fu diffuso nel 1998 il primo segnale DTT in Italia. La televisione digitale terrestre può oggi rappresentare un'importante frontiera per i broadcaster italiani di servizio pubblico e commerciale. Il 2004 ha visto la partenza dei servizi dei grandi operatori nazionali e di alcune emittenti locali e molti nuovi programmi digitali appositamente studiati per tale applicazione cominciano a raggiungere gli utenti. Una novità, oltre all'aumento del numero dei programmi, è costituita dalle applicazioni di TV interattiva basate sulla piattaforma a standard MHP (*Multimedia Home Platform*): servizi di pubblica utilità per il cittadino, super-teletext, votazione a distanza all'interno dei programmi televisivi, pubblicità interattiva, giochi, servizi bancari. La televisione potrà gradualmente offrire all'utente la navigazione Internet, senza perdere tuttavia le proprie caratteristiche di elettrodomestico di facile utilizzo per tutti.





# Capitolo 1

## Le prime trasmissioni

All'inizio del diciannovesimo secolo, numerosi scienziati e ricercatori pensarono alla possibilità di trasmettere immagini a lunga distanza e già alla fine del secolo, precisamente nel 1892, Alexander Bain inventò il primo sistema telegrafico per la trasmissione di immagini.

Nel 1924, John Logie Baird riuscì ad effettuare la prima trasmissione video a distanza dove si potevano distinguere le forme degli oggetti trasmessi.

In quegli anni anche gli ingegneri della Bell Telephone svilupparono un sistema per la sincronizzazione di suoni con immagini in movimento, che segnarono l'inizio dei filmati sonori che poco dopo, il 6 Agosto del 1926, portò alla produzione del primo film sonoro commerciale dal titolo "Don Juan".

Nel 1923, il russo Vladimir Kosma Zworykin brevettò il primo iconoscopio, strumento per la proiezione di immagini su di una sorta di piatto ricoperto da una speciale sostanza e basato sul tubo a raggi catodici (CRT), ossia lo stesso sistema utilizzato dall'ormai non più moderno cinescopio televisivo. Successivamente Zworykin collaborò con la RCA di David Sarnoff per la produzione del primo televisore CRT, chiamato orticon, che iniziò ad esser prodotto in massa nel 1945.

Nel marzo del 1935, la prima trasmissione ufficiale fu diffusa in una rete di locali pubblici tedeschi e un anno più tardi il primo importante evento sportivo mondiale, i Giochi Olimpici di Berlino, vennero trasmessi in diretta.

Durante la seconda Guerra Mondiale la Germania fu l'unico stato europeo a tenere in onda le trasmissioni televisive. Parigi iniziò nell'Ottobre del 1944, Mosca nel Dicembre del 1945, e la BBC nel Giugno del 1946 quando ormai la televisione era divenuta un successo mondiale.

## Le trasmissioni a colori

Già alla fine del 1940, negli Stati Uniti fabbricanti di televisori e società di radiodiffusione si unirono allo scopo di definire degli standard per un nuovo sistema TV a colori. La proposta venne ufficialmente riconosciuta nel 1952 da parte della FCC (*Federal Communications Commission*), conosciuta anche come NTSC (*National Television Standard Committee*), e definita RCA.

Questa era l'unica costruita sulle basi della compatibilità bi-direzionale dello standard monocromatico già esistente ossia, un ricevitore monocromatico doveva essere in grado di visualizzare le trasmissioni a colori in bianco e nero e, allo stesso modo, un ricevitore a colore, visualizzare le trasmissioni in bianco e nero.

In Europa, le trasmissioni a colori iniziarono ufficialmente una decina di anni più tardi, nel 1967, con i sistemi SECAM (*Séquentiel Couleur à Mémoire*) e PAL (*Phase Alternating-Line*).

Tali standard richiesero ampi studi preliminari sulla percezione del colore e un grande ingegno, a tal punto che a 40 anni dall'uscita del primo standard NTSC, e nonostante alcune imperfezioni, viene tutt'ora utilizzato.

I tre segnali trasmessi dalla telecamera, rosso, giallo e blu (RGB) dovevano essere da un lato visualizzabili in bianco e nero, senza molti artefatti sul ricevitore televisivo in uso e dall'altro poter essere trasmessi sulla banda di un canale già utilizzato per le trasmissioni in colore.

L'idea di base era quella di trasformare, mediante una combinazione lineare, le tre componenti (R, G, B) in altre tre equivalenti (Y, Cb, Cr) o (Y, U, V):

$$\begin{array}{ll} Y = 0.587G + 0.299R + 0.1145B & \text{segnale di luminanza} \\ Cb = 0.564(B-Y) \text{ o } U = 0.493(B-Y) & \text{crominanza blu o differenza di colore} \\ Cr = 0.713(R-Y) \text{ o } V = 0.877(R-Y) & \text{crominanza rossa o differenza di colore} \end{array}$$

La combinazione utilizzata per la luminanza (o "luma") del segnale è stata scelta per essere il più possibile simile al segnale di uscita di una telecamera monocromatica, in modo da permettere al ricevitore televisivo di trattarlo come un semplice segnale in bianco e nero. I valori dei segnali di crominanza (o "chroma") rappresentano la "colorazione" dell'immagine del segnale Y, e consentono, mediante una ricombinazione lineare, il recupero da parte del ricevitore dei segnali RGB originali.

Studi sulla percezione visiva hanno dimostrato che la risoluzione dell'occhio umano è minore per la componente di colore rispetto alla componente di luminanza. Questo significa, almeno per immagini naturali, che i segnali di crominanza possono occupare una ridotta larghezza di banda (un ottavo di quella riservata alla luminanza), che permise quindi di inserire i segnali di crominanza all'interno dello spettro video già esistente.

La combinazione Y, Cb, Cr è utilizzata da tutti i sistemi TV a colori, tra cui i più recenti standard digitali.

Per essere in grado di trasportare i tre segnali su di un canale Tv esistente (6 MHz negli Stati Uniti, 7 o 8 MHz in Europa) è stata aggiunta una sottoportante all'interno dello spettro video, modulata dai segnali di crominanza a larghezza di banda limitata, dando così vita ad un nuovo segnale chiamato **CVBS** (*Color Video Baseband Signal*).

La portante deve essere posta nella parte superiore dello spettro video e deve rimanere entro i limiti della larghezza di banda video esistenti per non interferire con il segnale di luminanza e quindi con i ricevitori in bianco e nero visti precedentemente .

Il sistema descritto fino a questo punto è comune a tutti gli standard mondiali (PAL, SECAM, NTSC) mentre le caratteristiche che li differenziano sono legati alle frequenze utilizzate e al modo di modulare la sottoportante.

# Sistema PAL

La soluzione adottata al problema di aggiungere due segnali indipendenti di crominanza in un canale che è già quasi interamente occupato dal segnale di luminanza Y, prevede di combinare i due segnali aggiunti in uno solo e di inserirlo nello stesso canale utilizzando una tecnica nota come *frequency interleaving*. Si parte considerando lo spettro del segnale video monocromatico, mostrato in Fig. 1.1, dove si nota che esso è formato da raggruppamenti di frequenze separate fra di loro e centrate su multipli interi della frequenza di riga  $f_H = 15625$  Hz. Ciascuno è poi formato da picchi verticali distanziati fra loro di multipli interi della frequenza di semiquadro, ossia 50 Hz. Le ampiezze delle portanti, di intensità decrescente al crescere della frequenza, sono separate fra di loro da intervalli di frequenza vuoti, tanto più estesi quanto più si sale in frequenza.

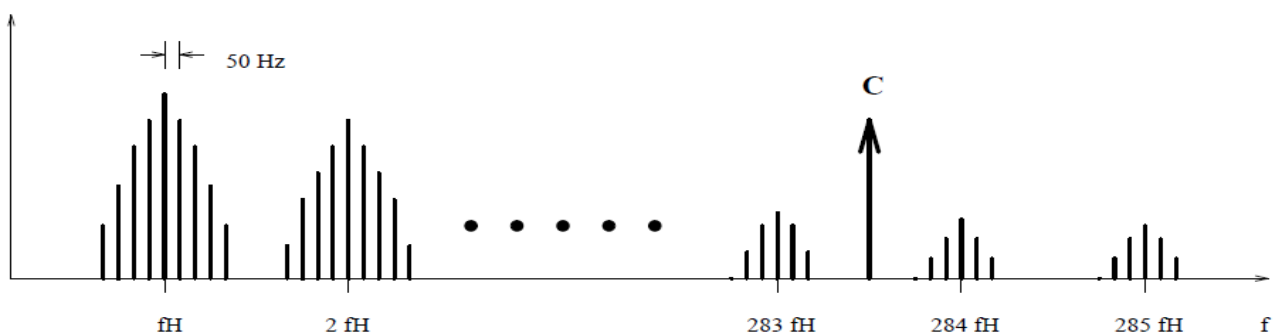


Fig 1.1 Spettro del segnale video di luminanza (Y).  $f_H$  vale 15.625 kHz (standard europeo a 625 linee, 50 Hz)

Questo è il motivo per il quale l'informazione del colore può essere inserita in quella della luminanza: il meccanismo è quello di inserire lo spettro del segnale colore (che è simile allo spettro della luminanza, solo meno esteso) negli spazi vuoti, a partire da una frequenza (sottoportante colore) che sia esattamente a metà strada fra due armoniche di  $f_H$  (Fig. 1.2).

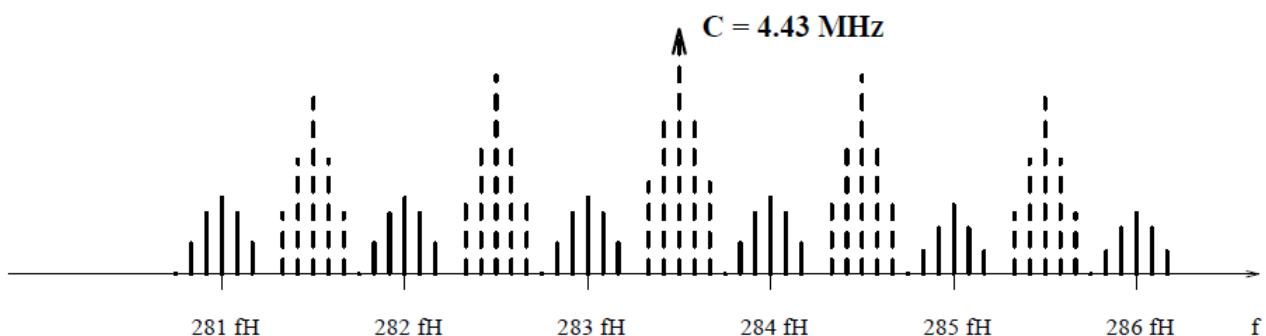


Fig 1.2 linee tratteggiate = crominanza, linee continue = segnale Y

Nel tempo si può notare come il segnale colore produca delle piccole variazioni sinusoidali sul segnale di luminanza: siccome tutte le sinusoidi sono multiple di mezza frequenza di linea invertono la loro fase da una riga alla successiva e da un semiquadro all'altro.

Questo effetto viene facilmente mediato via dall'osservatore, col che si è ottenuta la compatibilità con il segnale monocromatico.

Nel sistema PAL la sottoportante colore è situata fra le armoniche 283 e 284, ovvero in  $f_{scPAL} = 283.5 \times f_H = 283.5 \times 15625 = 4.43 \text{ MHz}$ .

Il segnale monocromatico di luminanza Y viene trasmesso a piena banda, cioè 5 MHz, per non perdere dettagli di immagine. Il segnale colore non necessita di una banda così larga, in quanto l'occhio umano distingue il colore solo per aree piuttosto grandi. Da studi effettuati pare che l'occhio distingua bene i colori per aree di larghezza maggiore o uguale ad un venticinquesimo della larghezza dello schermo: aree così grosse hanno componenti spettrali solo fino a circa 0.5 MHz. Inoltre, per aree di dimensione media, con componenti spettrali fra 0.5 e 1.5 MHz, la visione del colore è ridotta in modo tale che un solo segnale di cromaticità è sufficiente a portare l'intera informazione visibile. Si può asserire che le frequenze fra 1.5 e 5 MHz sulle componenti di cromaticità non portano informazione che possa essere riconosciuta dal sistema visivo umano e che, quindi, la banda necessaria per trasmettere il segnale di colore è circa di 3 MHz ( $\pm 1.5$  MHz attorno alla portante colore) per ciascuno dei due segnali di cromaticità.

I due segnali (R-Y) e (B-Y) vengono trasmessi simultaneamente sulla stessa portante utilizzando una modulazione QAM. In pratica vengono entrambi modulati con la stessa frequenza portante, ma con una differenza di fase di 90 gradi. In questo modo i due segnali sono in quadratura sulla stessa portante e possono essere demodulati separatamente senza interferenze (purché il ricevitore sia in grado di ricostruire frequenza e fase della portante).

Chiamando C il segnale ottenuto combinando insieme (R-Y) e (B-Y) si può notare come l'ampiezza (cioè il modulo) di C, il cui valore istantaneo corrisponde alla differenza fra (R-Y) e (B-Y), non sia altro che la saturazione istantanea, mentre la sua fase porta l'informazione di tinta (hue).

Per finire resta da spiegare come il ricevitore possa conoscere frequenza e fase della sottoportante colore con precisione sufficiente per demodulare correttamente la QAM, considerando che, per risparmiare potenza ed evitare interferenze con il segnale di luminanza, il segnale colore viene modulato DSB-SC, cioè a doppia banda laterale con portante soppressa. L'informazione sulla portante viene fornita al ricevitore inserendo alcuni cicli di essa (fra 8 e 11 in generale, 10 nel PAL) insieme ai sincronismi.

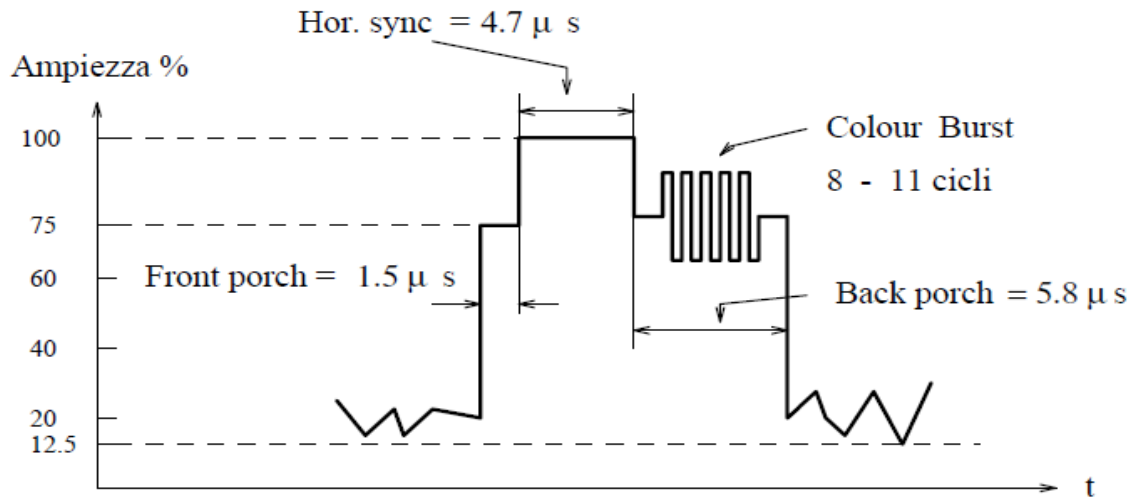


Fig 1.3 Posizione del Colour Burst (fra 8 e 11 cicli della sottoportante colore) nel Back Porch di ciascun impulso di sincronismo orizzontale.

Questo esempio di portante, detto *Colour Burst Signal*, è posizionato nella parte finale o *Back Porch* di ciascun sincronismo di linea e non interferisce con i sincronismi perché è di ampiezza inferiore e posizionato dopo l'impulso di sincronismo orizzontale vero e proprio (Fig. 1.3). La fase del *Colour Burst* varia di  $90 (\pm 45)$  gradi fra una linea e la successiva, fornendo quindi sempre il riferimento corretto per la decodifica.

# Sistema NTSC

Il sistema NTSC è ovviamente compatibile con lo standard a 525 linee americano, ed è stato il primo standard colore adottato nel mondo. Rispetto al sistema PAL, nell'NTSC, è disponibile una banda inferiore: per questo motivo l'informazione sul colore è stata ridotta. Si è notato infatti come la risposta del sistema visivo umano al colore non sia uniforme, ma vari con la tinta. In particolare la risoluzione al colore ha un massimo sui colori arancio (rosso + verde) e ciano (blu + verde) ed un minimo dalle parti dei colori viola (blu + rosso) e verde-giallo. Di conseguenza sono stati generati due segnali colore nuovi, detti I e Q (Fig. 1.4), approssimativamente posti nelle regioni di massima e minima risoluzione al colore.

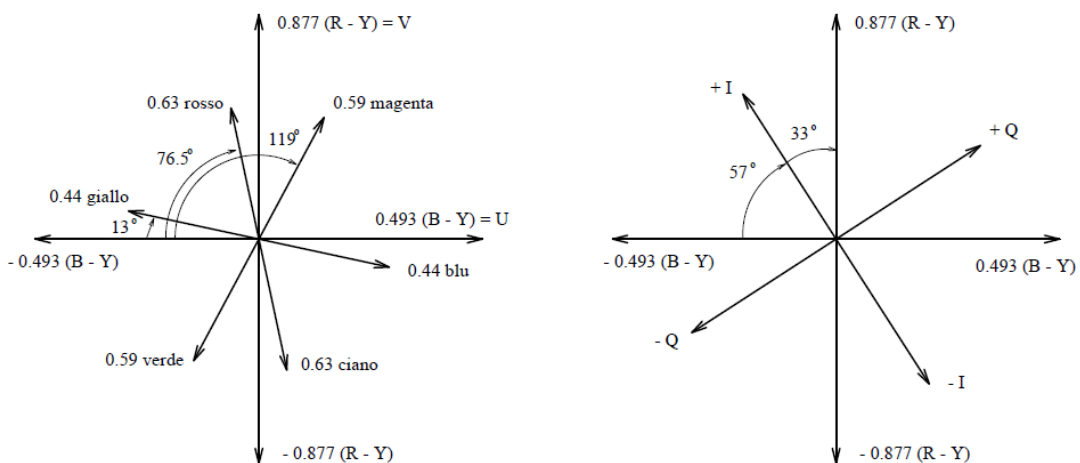


Fig. 1.4 Sinistra: ampiezza e fase dei colori primari e complementari dopo lo scalamento di  $(R - Y)$  e  $(B - Y)$ . Destra: posizione degli assi I e Q del sistema NTSC.

Questi segnali sono generati ruotando di 33 gradi in senso antiorario gli assi U e V:

$$I = \cos(33)V - \sin(33)U = 0.74 (R - Y) - 0.27 (B - Y) = 0.60 R - 0.28 G - 0.32 B$$

$$Q = \sin(33)V + \cos(33)U = 0.48 (R - Y) + 0.41 (B - Y) = 0.21 R - 0.52 G - 0.31 B$$

A questo punto si è deciso di assegnare a I una banda di  $\pm 1.5$  MHz e a Q di  $\pm 0.5$  MHz, in accordo con la diversa sensibilità dell'occhio ai diversi colori. Q viene trasmessa usando una modulazione a doppia banda laterale e quindi occupa 1 MHz attorno alla sottoportante colore. I viene invece trasmesso tenendo l'intera banda inferiore (1.5 MHz) ma solo 0.5 MHz della superiore, utilizzando una modulazione vestigiale che occupa 2 MHz in tutto. Siccome poi I e Q sono poste in quadratura sulla stessa sottoportante si ha che l'intero segnale colore occupa 2 MHz, cioè da -1.5 a +0.5 MHz.

La frequenza esatta della sottoportante colore nel sistema NTSC è stata fissata pari a 3.579545 MHz. Infatti per evitare l'interferenza fra colore e luminanza (che apparirebbe come una serie di puntini luminosi sullo schermo) si è scelto di usare un multiplo dispari di metà della frequenza di riga, ovvero metà frequenza di riga dopo la 227 armonica. Questo porterebbe la sottoportante colore in  $227,5 \times f_H = 227,5 \times 15750 = 3.583125$  MHz.

Purtroppo tale frequenza, pur non interferendo con il segnale di luminanza, interferisce con la portante audio producendo una nota di 0.92 MHz che, a sua volta, interferisce con la luminanza. Il problema potrebbe essere evitato se la sottoportante audio, posta a 4.5 MHz, fosse un multiplo dispari esatto di metà frequenza di riga. Siccome non lo è, e non può essere spostata (il segnale deve essere compatibile con quello in bianco e nero), la soluzione adottata è stata quella di spostare la frequenza di riga. Infatti se si sceglie  $f_{hNTSC} = 4,5 \text{ MHz} / 286 = 15734,26 \text{ Hz}$  si ottiene quanto voluto. La variazione fra  $f_{hNTSC}$  per il segnale in bianco e nero (pari a 15750 Hz) e quella per il colore è sufficientemente piccola (15.74 Hz) da non creare problemi di aggancio agli oscillatori dei televisori in bianco e nero e quindi la compatibilità è mantenuta. Ovviamente, siccome il numero di linee non varia, cambia anche la frequenza di scansione verticale, da 60 a 59.94 Hz ( $15734,26 / 262,5$ ). Anche in questo caso la differenza è talmente piccola (0.06 Hz) da non creare problemi di aggancio. La frequenza della sottoportante colore si ottiene quindi come  $227,5 \times 15734,26 = 3.579545$  MHz.

La trasmissione di I e Q avviene poi con una modulazione QAM a portante soppressa. Siccome non è stata prevista l'inversione della fase ad ogni linea (come nel PAL), il sistema NTSC è sensibile agli errori di fase sul segnale ricevuto. Tali errori sono visibili come variazioni di tinta per cui i televisori americani dispongono di un circuito di controllo automatico di tinta (ATC).



# Sistema SECAM

Come detto prima, il sistema SECAM, sviluppato in Francia, differisce da PAL ed NTSC nel modo con cui viene trasmessa l'informazione sul colore. Il principio base è quello di inviare una sola componente di colore per ogni riga di immagine. Il ricevitore la immagazzina poi in una linea di ritardo (a ultrasuoni) e la combina con l'informazione della riga precedente per ottenere i colori. Il nome, *sequential a memorie*, deriva appunto da questo meccanismo di trasmissione. Se una riga porta l'informazione su (R-Y) nella successiva c'è (B-Y): poiché il numero di linee per immagine è dispari, se la riga  $n$  di un quadro contiene (R-Y) la riga  $n$  del quadro dopo contiene (B-Y). E' quindi necessario un impulso di identificazione che permetta al ricevitore di associare la componente colore giusta ad ogni riga: questo impulso viene trasmesso durante la ritraccia verticale. Le componenti di crominanza, opportunamente filtrate, vengono alternativamente trasmesse mediante una modulazione di frequenza (con preenfasi). Vi sono stati diversi sistemi SECAM, anche incompatibili fra loro (certi con 819 linee ed altri con 625).

Il sistema SECAM-III prevede 625 linee a 50 Hz con un canale di larghezza di banda 8 MHz: le sottoportanti colore sono due, una per (R-Y) posta in  $282 f_H = 4,40625$  MHz ed una per (B-Y) in  $272 f_H = 4,250$  MHz (questo minimizza le interferenze con il segnale monocromatico).

Le due componenti di crominanza sono filtrate a 1.5 MHz e pesate in modo da ottenere

$$D_R = -1.9 (R-Y)$$

$$D_B = 1.5 (B-Y)$$

La trasmissione FM prevede una deviazione di  $280 D_R$  kHz e  $230 D_B$  kHz per i due segnali, con una deviazione massima ammessa di 500 kHz verso il basso e 350 kHz verso le alte frequenze. In questo modo certi colori sono meglio definiti di altri: la scelta di  $D_R$  e  $D_B$  ne tiene ovviamente conto.



# Capitolo 2

## Digitalizzazione segnali audio - video

Fino a qualche anno fa negli studi di registrazione e editing televisivi venivano utilizzati diversi formati digitali come il D1 (components) e il D2 (composite). Allo scopo di facilitare l'interoperabilità delle apparecchiature, l'ex CCIR<sup>1</sup> (*Comité Consultatif International du Télégraphe et du Téléphone*) ha standardizzato le caratteristiche di digitalizzazione (Raccomandazione CCIR-601) e interfacciamento (raccomandazione CCIR-656) di segnali video digitali in formato component (Y, Cr, Cb nel formato 4:2:2).

La frequenza di campionamento standard, indipendentemente dal segnale d'ingresso, è di 13.5 MHz per i segnali in banda base R, G, B ed Y (ITU 601). La quantizzazione delle trasmissioni broadcast utilizza 256 livelli (8 bits) che quindi portano ad un data rate di  $H_0 = 108$  Mbit/s. Per i segnali di cromaticità, Cb e Cr, la frequenza di campionamento è limitata a 6.75 MHz e il data rate diviene quindi di 54 Mbit/s.

La raccomandazione CCIR-601 definisce i parametri per i segnali video Y, Cb e Cr nel formato 4:2:2 con 8 bit per campione (quattro campioni Y, due Cb e Cr, con una disposizione finale di 10 bit). La scelta della frequenza di 13.5 MHz è indipendente dal tipo di sistema (625 linee o 525 linee) ma il numero di campioni attivi per riga è di 720 per la luminanza e 360 per i segnali di cromaticità, in entrambi i sistemi di riferimento.

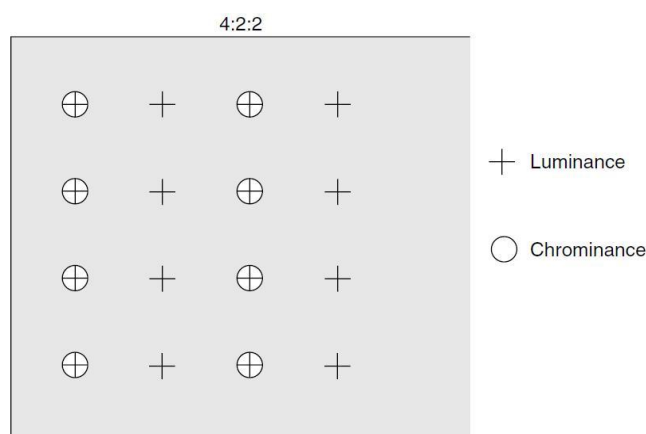


Fig 2.1 Posizione dei campioni nel formato 4:2:2

[1] La CCIR è stata il ramo della radiocomunicazione dell'ex CCITT (Comité Consultatif International du Télégraphe et du Téléphone), successivamente rinominata ITU (International Telecommunications Union). La CCIR è ora chiamata ITU-R

I segnali di crominanza, Cr e Cb, sono presenti simultaneamente ad ogni riga, la risoluzione verticale per la crominanza è la stessa per la luminanza (480 linee per i sistemi a 525 linee, 576 linee per sistemi a 625 linee). Il bit-rate risultante da questo processo è  $13.5 \times 8 + 2 \times 6.75 \times 8 = 216$  Mbit/s. Con una quantizzazione di 10 bit, il bit-rate diventa 270 Mb/s, notevolmente minore di una trasmissione RGB (formato 4:4:4) che utilizza un data rate di  $H_0 = 324$  Mbit/s. Tuttavia, se si tiene conto della ridondanza relativa alla digitalizzazione del periodo inattivo del segnale video (blanking periods), il bit-rate utile scende a 166 Mbit/s. Questi periodi di blanking possono essere riempiti con altri dati utili, come il suono digitale, segnali utili alla sincronizzazione, e altre informazioni.

Per quanto riguarda l'audio, la codifica MPEG-1 utilizza le frequenze di campionamento a 32, 44.1, 48 kHz. (Tabella 2.2)

Signals	Clock [MHz]	b [bit]	$H_0$ [Mbit/s]	$H_{0Total}$ [Mbit/s]	Format
R	13.5	8	108	324	4:4:4 ITU 601
G	13.5	8	108		
B	13.5	8	108		
Y	13.5	8	108	216	4:2:2 ITU601
Cb	6.75	8	54		
Cr	6.75	8	54		

Tabella 2.1 Caratteristiche della digitalizzazione e del bit-rate del segnali video (ITU 601)

Standard	Clock [kHz]	$H_{0Mono}$ [kbit/s]	$H_{0Stereo}$ [kbit/s]	Uses
DSR	32	512	1024	Digital satellite radio
G	44.1	706	1412	Audio CD
B	48	768	1536	Professional audio studio

Tabella 2.2 Caratteristiche della digitalizzazione e del bit-rate del segnale audio

## MPEG: caratteristiche generali

Nel 1990, numerose applicazioni multimediali su diverse piattaforme presentarono la necessità di memorizzare e riprodurre immagini in movimento e suoni associati in formato digitale. L'ISO decise quindi di formare un gruppo di esperti sulla stessa linea di quello fatto poco tempo prima con JPEG<sup>2</sup> (*Joint Photographic Experts Group*), composto da membri provenienti da diversi rami dell'elettronica, dell'informatica e delle telecomunicazioni, chiamato MPEG (*Motion Pictures Expert Group*). Il primo risultato del lavoro fu lo standard internazionale ISO / IEC 11.172, ampiamente conosciuto come MPEG-1. L'obiettivo principale era quello di permettere la memorizzazione su CD-ROM o CD-I (singola velocità in quel momento) di video e audio stereo, il che implicava un massimo bit-rate di 1,5 Mb/s. Oltre alla ridondanza intrinseca derivata da JPEG per le immagini fisse, la codifica di immagini in movimento consente lo sfruttamento della ridondanza temporale molto importante tra immagini successive che compongono una sequenza video. Per codificare le immagini venne scelto il formato SIF che permetteva di sfruttare un minor bit-rate a scapito di una minor risoluzione. Questo formato si ottiene dimezzando la risoluzione spaziale così come la risoluzione temporale, che diventa 25Hz per i sistemi a 625 linee e 29.97 Hz per i sistemi a 525 linee.

A seconda degli standard originari, le risoluzioni spaziali diventano quindi:

- Risoluzione di luminanza:  $360 \times 288$  (625 linee) o  $360 \times 240$  (525 righe);
- Risoluzione di cromaticanza:  $180 \times 144$  (625 linee) o  $180 \times 120$  (525 linee).

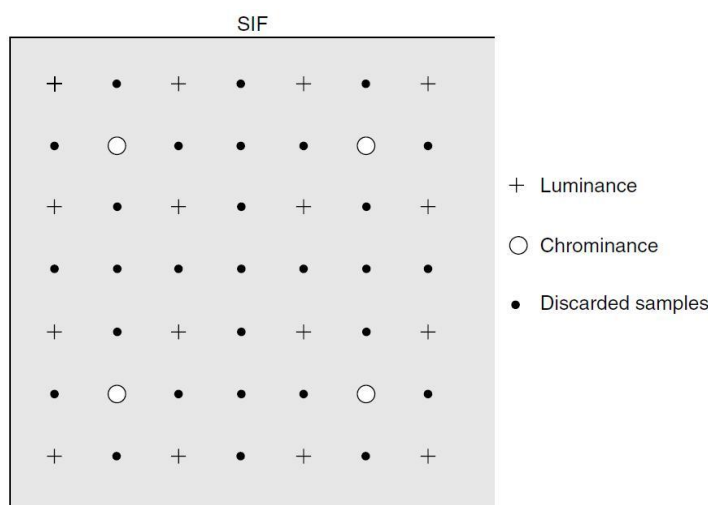


Fig 2.2 Posizione dei campioni nel formato SIF

L'algoritmo di compressione audio utilizzato per i canali audio di accompagnamento è noto come MUSICAM, utilizzato anche nel sistema europeo radio digitale DAB (*Digital Audio Broadcasting*).

Lo standard MPEG-1, pubblicato nel Novembre del '92, è definito in tre sezioni:

- MPEG-1 system (ISO/IEC 11172-1): definisce MPEG-1 multiplex structure;
- MPEG-1 video (ISO/IEC 13818-2): definisce MPEG-2 video coding;
- MPEG-1 audio (ISO/IEC 13818-3): definisce MPEG-2 audio coding.

Tuttavia, la qualità delle immagini ottenibile da MPEG-1 non era adatto alle applicazioni broadcast, dato che non si era presa in considerazione la codifica di immagini interlacciate o una possibile evoluzione verso l'HDTV.

Il gruppo MPEG lavorò quindi verso uno sviluppo di uno standard più flessibile, ottimizzato per le trasmissioni broadcast.

Nel Novembre del '94, uscì MPEG-2, che come il predecessore venne specificato in tre parti distinte:

- MPEG-2 system (ISO/IEC 13818-1): definisce l'MPEG-2 streams;
- MPEG-2 video (ISO/IEC 13818-2): definisce l'MPEG-2 video coding;
- MPEG-2 audio (ISO/IEC 13818-3): definisce l'MPEG-2 audio coding.

Il sistema MPEG-2 divenne poi lo standard codifica di sorgente dell'European DVB (*Digital Video Broadcasting*), sistema di trasmissione TV digitale, che è il risultato del lavoro iniziato nel 1991 da ELG (*European Launching Group*), in seguito diventato DVB committee.

## Principi della codifica video MPEG-1

Come indicato in precedenza, l'obiettivo principale per MPEG-1 è stato quello di raggiungere video (con audio) di qualità media con un di bit-rate di 1,5 Mb/s per la memorizzazione su CD-ROM. La parte video utilizza 1,15 Mb/s e i restanti 350 kb/s viene utilizzata per dati audio e altre informazioni. Tuttavia, le specifiche di MPEG-1 sono molto flessibili e permettono di impostare diversi parametri basandosi sul compromesso tra complessità dell'encoder, tasso di compressione e qualità generale finale.

La codifica video utilizza gli stessi principi con perdita di JPEG, a cui vengono aggiunte nuove tecniche per formare l'MPEG-1 "toolbox"; queste tecniche permettono di sfruttare la forte correlazione tra immagini successive in modo da ridurre notevolmente la quantità di informazioni necessarie da trasmettere o memorizzare. Queste tecniche, note come "*prediction with movement compensation*", sono basate sul dedurre le immagini di una sequenza dalle precedenti e dalle successive, con un minimo di informazione aggiuntiva rappresentata dalle differenze tra le immagini. Ciò richiede la presenza nel codificatore MPEG di uno *stimatore di movimento*, che è la funzione più complessa e ne influenza maggiormente le prestazioni.

Poiché stiamo parlando di immagini in movimento, la decodifica deve essere realizzata in tempo reale (questo significa un ritardo accettabile), il che implica, un hardware dedicato a questo. La codifica, che quindi è molto più complessa, può essere fatta in più di un passaggio per le applicazioni dove la qualità è di primaria importanza a differenza delle trasmissioni in real time dove il tempo diventa la caratteristica principale da soddisfare a scapito di una minor qualità. La realizzazione pratica del codificatore è dunque un compromesso tra velocità, tasso di compressione, complessità e qualità dell'immagine. Inoltre, il tempo di sincronizzazione e tempo di accesso random a una sequenza devono essere mantenute entro un limite accettabile (non superiore a 0.5 s).

## Tipi di immagini MPEG

MPEG definisce tre tipi di immagini, disposte come mostrato in Figura 2.3:

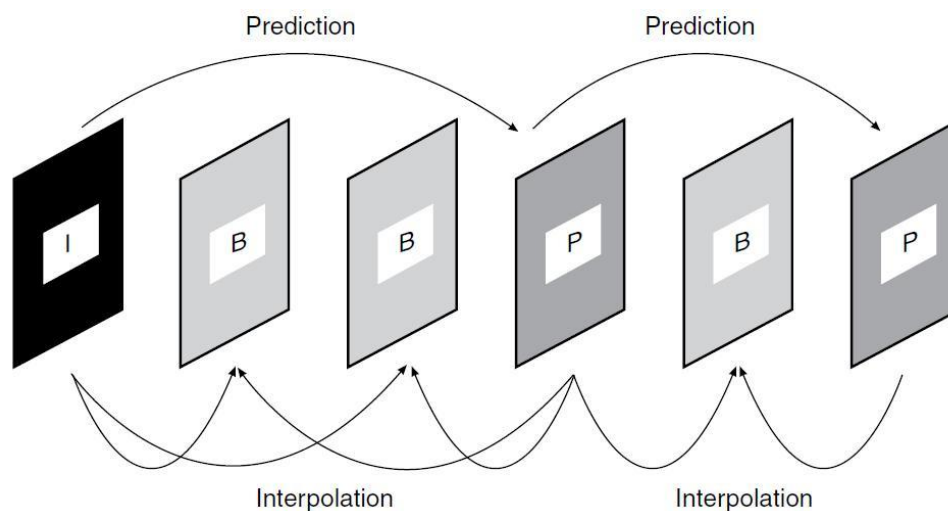


Fig 2.3 Concatenazione dei tre tipi di immagini MPEG

- **I (intra)**: le immagini sono codificate in un modo simile al formato JPEG, senza riferimento ad altre immagini, il che significa che contengono tutte le informazioni necessarie per la loro ricostruzione da parte decoder e per questo motivo sono il punto di partenza per l'accesso ad una sequenza video. Il tasso di compressione delle immagini tipo I è relativamente basso, ed è paragonabile a un'immagine JPEG.

- **P (predicted)**: le immagini sono codificate dalla precedente immagine I o P, utilizzando le tecniche di previsione del movimento con compensazione. Le immagini di tipo P possono essere utilizzate come base per la predizione delle successive immagini, ma poiché la compensazione del movimento non è perfetta, non è possibile estendere di molto il numero di immagini tipo P tra due immagini tipo I. Il tasso di compressione delle immagini P è sensibilmente più alto del precedente tipo I.

- **B (bi-directional or bi-directionally predicted)**: le immagini sono codificate mediante interpolazione bi-direzionale tra le immagini tipo I e P che precedono e seguono quella considerata. Dato che non sono utilizzate per la codifica delle immagini successive, le immagini tipo B non propagano errori di codifica ed inoltre offrono il più alto tasso di compressione.

A seconda della complessità del codificatore utilizzato, è possibile codificare o solamente I, o la coppia I e P, oppure la terna I, P, B, con risultati molto diversi per quanto riguarda il tasso di compressione e la risoluzione di accesso random ma anche in termini di tempo di codifica e qualità finale percepita.



Due parametri descrivono la successione di I,P e B (Fig. 2.4):

$M$  è la distanza (in numero di immagini) tra due successive immagini P

$N$  invece indica la distanza tra due immagini tipo I, e delimita quindi un “gruppo di immagini” (GOP).

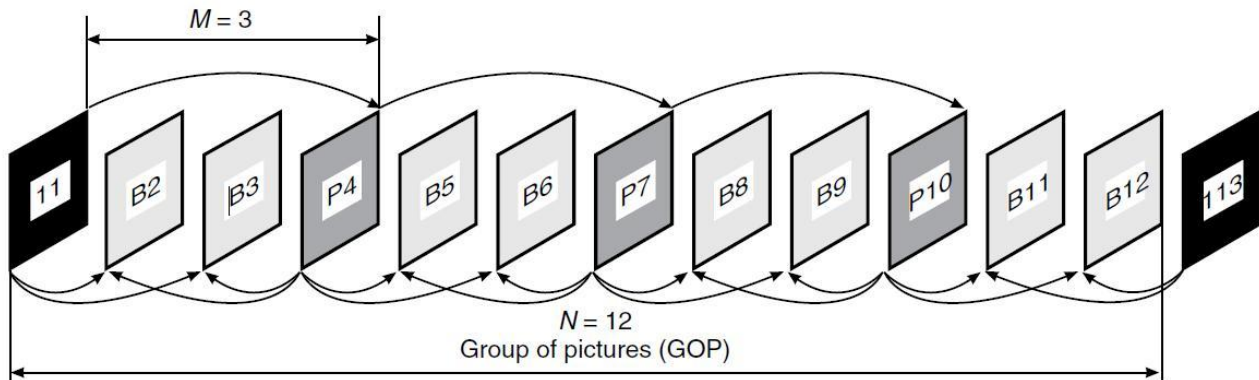


Fig 2.4 Esempio di un gruppo immagini MPEG ( $M=3$ ,  $N=12$ )

I parametri utilizzati sono in genere  $M = 3$  e  $N = 12$ , al fine di ottenere una qualità video soddisfacente con un tempo di accesso random accettabile ( $<0,5$  s) (bit rate di 1,15 Mb/s). Con questi parametri, una sequenza video è composta come segue: 1/12 delle sue immagini sono di tipo I (8,33%), 1/4 di tipo P (25%), e 2/3 sono tipo B (66%); il tasso di compressione globale è massimizzata dal fatto che le immagini più frequenti sono caratterizzate dal più alto tasso di compressione.

## Riordino delle immagini

La sequenza delle immagini dopo la decodifica deve avere lo stesso ordine della sequenza originale prima della codifica. Con i parametri sopra ( $M = 3$ ,  $N = 12$ ), la corrispondenza tra la posizione dell'immagine nella sequenza e il tipo è la seguente:

1(I) 2(B) 3(B) 4(P) 5(B) 6(B) 7(P) 8(B) 9(B) 10(P) 11(B) 12(B) 13(I)

dove il numero indica la posizione nella sequenza e la lettera il tipo di immagine.

Tuttavia, al fine di codificare o decodificare un'immagine tipo B, sia l'encoder che il decoder avranno bisogno dell'immagine I o P precedente e successiva. Ciò richiede riordino della sequenza delle immagini originali, che nel nostro esempio porterebbe alla sequenza:

1(I) 4(P) 2(B) 3(B) 7(P) 5(B) 6(B) 10(P) 8(B) 9(B) 13(I) 11(B) 12(B)

L'aumento del tasso di compressione che si ottiene dall'utilizzo delle immagini tipo B porta però ad un aumento del ritardo di codifica (due immagini extra) e necessita di una dimensione di memoria maggiore sia per la codifica che per la decodifica (un'immagine in più da memorizzare).

## Livelli di una sequenza video MPEG

La sequenza video MPEG è definita da una gerarchia a livelli con diverse funzioni (Figura 2.4):

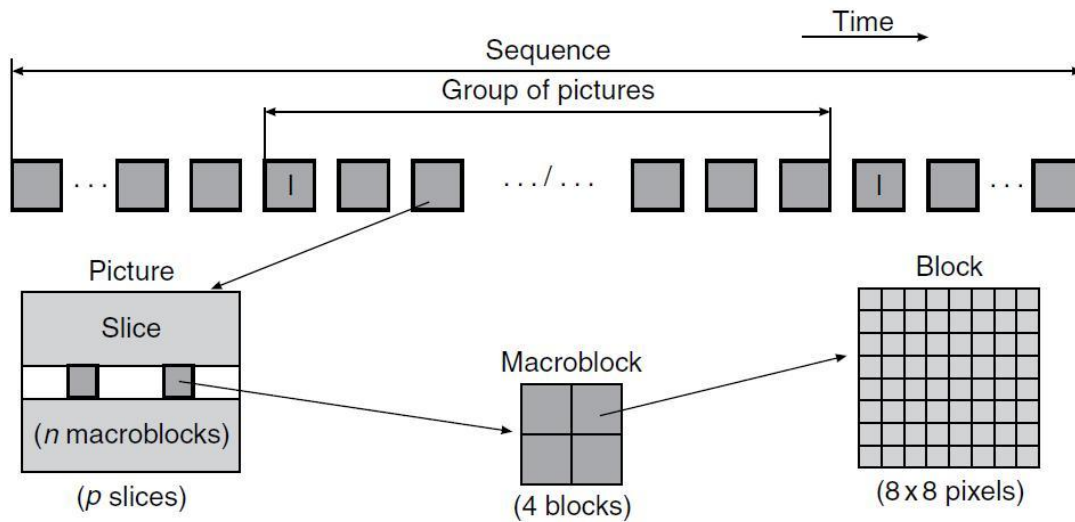


Fig 2.4 Gerarchia livelli MPEG

**Sequence:** è il livello più alto che definisce il contesto valido dell'intera sequenza (parametri video di base, ecc.)

**Group of pictures (GOP):** livello che determina l'accesso casuale alla sequenza, che inizia sempre con un'immagine tipo I. Nell'esempio di cui sopra ( $M = 3$ ,  $N = 12$ ) il GOP è costituito da 12 immagini.

**Picture:** Questa è l'unità elementare di visualizzazione già vista precedentemente; i formati possibili sono I, P e B.

**Slice:** Questo è il livello per l'indirizzamento intra frame e (ri)sincronizzazione, utilizzato per il correzione di errori. Viene definito come un insieme di macroblocchi contigui. L'insieme delle slices copre l'intera immagine senza alcuna sovrapposizione. La dimensione di ognuna può teoricamente variare da un macroblocco di un'immagine ad una fila completa di macroblocchi.

**Macroblock:** Questo layer è utilizzato per la stima di movimento/compensazione. Un macroblocco ha una dimensione di  $16 \times 16$  pixel ed è composta da quattro blocchi di luminanza e due blocchi di cromaticanza (uno per Cr ed uno per Cb) (Figura 2.5).

**Block:** Come in JPEG, l'immagine viene divisa in blocchi di  $8 \times 8$  pixel. Il blocco è lo strato in cui ha luogo la DCT.

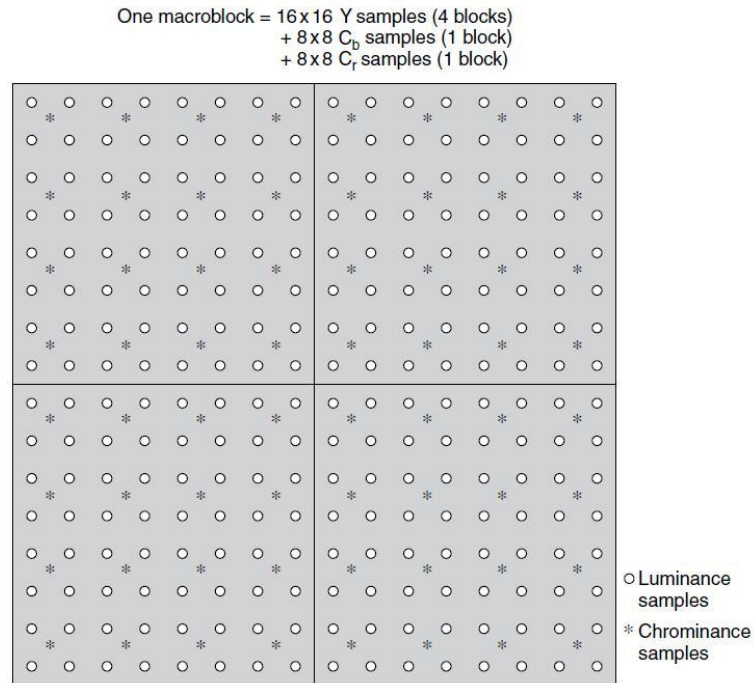


Fig 2.5 Composizione di un 4:2:0 macroblock (campioni Y, \* = campioni Cb,Cr)

A causa della divisione dell'immagine in un numero intero di macroblocchi, la risoluzione orizzontale di luminanza dell' MPEG-1/SIF passa da 360 pixel dell'originale SIF a 352 pixel (22 macroblocchi), dal momento che 360 non è un multiplo di 16. La risoluzione effettiva è quindi  $352 \times 288$  a 25 Hz ( $22 \times 18 = 396$  macroblocchi) per le immagini provenienti da sistemi a 625 linee, e  $352 \times 240$  a 30 Hz ( $22 \times 15 = 330$  macroblocchi) per le immagini provenienti da sistemi a 525 linee.

## Previsione, stima del movimento e compensazione

Si è visto che le immagini di tipo P e B sono sempre precedute e/o seguite da altre immagini. In una sequenza di immagini in movimento, gli spostamenti di un oggetto si traducono in differenti posizioni dell'oggetto stesso nelle immagini immediatamente successive, senza che vi sia alcuna correlazione evidente tra le zone di spostamento.

La *stima del movimento* consiste nel definire un "vettore movimento" che garantisce la correlazione tra una zona di partenza sulla prima immagine e una zona di arrivo sulla seconda, utilizzando una tecnica conosciuta come "block matching" (accoppiamento di

blocco). Tutto questo viene fatto a livello macroblock ( $16 \times 16$  pixel) spostando un macroblocco dall'immagine corrente all'interno di una piccola finestra di ricerca sull'immagine precedente, e quindi confrontando tutti i macroblocchi possibili della finestra, al fine di trovare quello con le minori differenze. La differenza nella posizione dei due macroblocchi corrispondenti dà un vettore di movimento (*motion vector*, Fig.2.6) che sarà applicato a tutti e tre i componenti del macroblocco (Y, Cb, Cr).

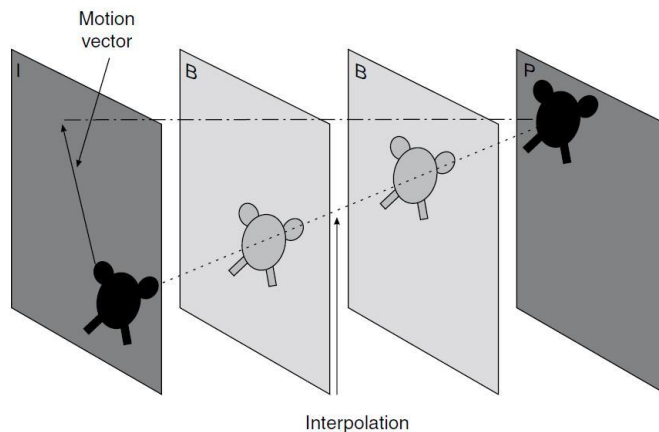


Fig 2.6 Illustrazione di una compensazione di movimento

Nel confrontare un'immagine P e una I, o due immagini P, a causa della distanza temporale tra queste (distanza di tre immagini in caso di  $M = 3$ ,  $N = 12$ ), generalmente il *block matching* non sarà perfetto e i vettori di movimento potranno essere di ampiezze relativamente elevate. Per questo motivo la differenza (*prediction error*) tra il blocco effettivo che deve essere codificato e il matching block viene calcolato e codificato in un modo simile ai blocchi delle immagini I (DCT, quantizzazione, RLC / VLC). Questo processo è chiamato *motion compensation*.

Per immagini tipo B, i vettori di movimento vengono calcolati per interpolazione temporale dai vettori dell'immagine P successiva in tre modi diversi (avanti, indietro, e bi-direzionale); il risultato porta il più piccolo errore previsto e l'errore stesso viene codificato allo stesso modo delle immagini P. Solo i macroblocchi diversi dall'immagine (o immagini) utilizzata per la previsione dovranno essere codificati, il che riduce notevolmente la quantità di informazioni necessarie per la codifica B e P.

Poiché la dimensione degli oggetti in movimento è generalmente più grande di un macroblocco, vi è una forte correlazione tra i vettori di movimento di blocchi consecutivi, e quindi per i vettori viene utilizzato un metodo di codifica differenziale (DPCM), riducendo così il numero di bit necessari. Quando la previsione non dà un risultato utile (per esempio nel caso di una telecamera in movimento dove compaiono

sempre nuove zone nell'immagine), le corrispondenti parti dell'immagine sono "intra" codificate nello stesso modo fatto per le immagini tipo I.

## Controllo del bit-rate di uscita

Il flusso di bit generato dal video (o audio) encoder, chiamato *elementary stream* (ES), al fine di soddisfare i vincoli del canale (trasmissione o registrazione/riproduzione) e le specifiche del buffer di input del decoder MPEG di riferimento, deve essere mantenuto il più possibile costante. Questa caratteristica non è garantita dal processo di codifica descritto precedentemente, causa le caratteristiche estremamente diverse di dettaglio e movimento delle immagini da codificare.

Per controllare il tasso di bit all'uscita del codificatore, l'encoder è dotato di uscita con un buffer *FIFO*; la quantità di informazioni contenute nel buffer è controllata e mantenuta entro limiti predeterminati mediante un ciclo di feedback che modifica i parametri di quantizzazione, la quale ha grande influenza sul bit-rate. In questo modo, è possibile ottenere un bit-rate costante, con una risoluzione che dipende dal contenuto e dalla quantità di movimento nell'immagine (più movimento, minore è la risoluzione). Un diagramma a blocchi schematico di un codificatore MPEG, che dà soltanto un'idea della sua reale complessità, è mostrato nella Figura 2.7.

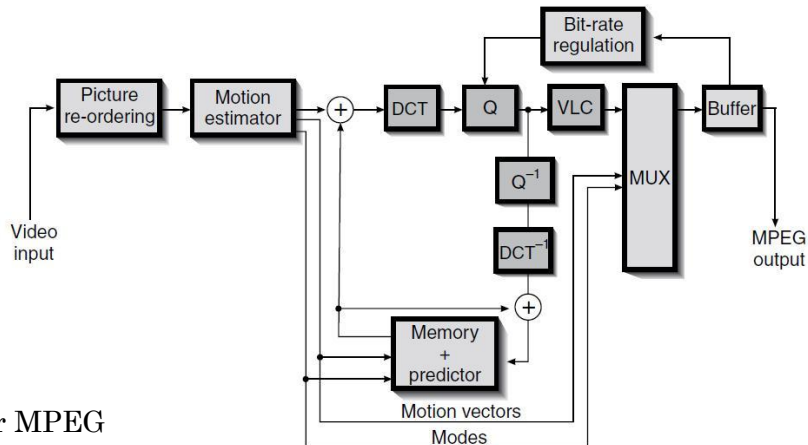


Fig 2.7 Encoder MPEG

Il decoder (Fig. 2.8) non dovendo effettuare la stima del movimento risulta formato da un'architettura molto più semplice; inoltre questo fatto è positivo anche in ambito di diffusioni di decoder MPEG, che risulteranno in numero notevolmente maggiore rispetto agli encoder.

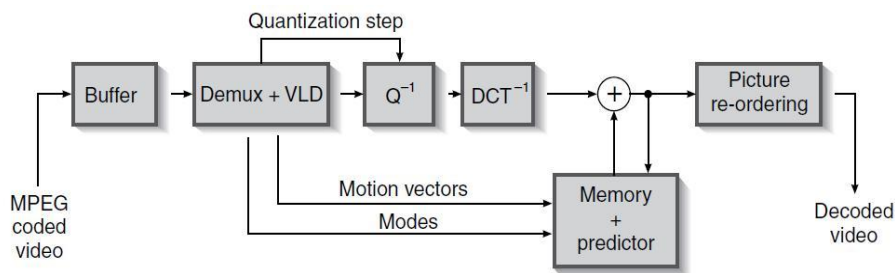


Fig 2.8 Decoder MPEG

## Codifica video MPEG-2 (applicazione broadcast)

MPEG-2 può essere descritto come un “compression toolbox”, cioè un superset dell’MPEG-1 più complesso, che ne utilizza tutti gli strumenti e ne aggiunge altri. Inoltre MPEG-2 è compatibile con MPEG-1, il che vuol dire che un decodificatore MPEG-2 è in grado di decodificare tutti i flussi provenienti da codifica MPEG-1.

### Livelli e profili dell’MPEG-2

Lo standard MPEG-2 è costituito da quattro livelli e cinque profili. I livelli definiscono le diverse risoluzioni possibili delle immagini, che partendo dal formato SIF arrivano al moderno HDTV, mentre i profili determinano gli strumenti di compressione utilizzati; inoltre non tutte le combinazioni tra livelli e profili sono possibili.

I livelli vengono descritti come segue:

- *low level*: corrisponde alla risoluzione SIF utilizzata dall’MPEG-1 (fino a 360×288)
- *main level*: corrisponde alla risoluzione standard 4:2:0 (fino a 720×576)
- *high-1440 level*: standard destinato alla risoluzione HDTV (fino a 1440×1152)
- *high level*: ottimizzato per la risoluzione HDTV wide-screen (fino a 1920×1152)

Per i profili la descrizione è più complessa:

- *simple profile*: definito con il fine di semplificare il più possibile encoder e decoder, a scapito di un più alto bit-rate in quanto non fa uso della previsione bi-directional (immagini tipo B).
- *main profile*: miglior compromesso tra tasso di compressione e costo, in quanto utilizza tutti e tre i tipi di immagini (I, P, B), ma porta una complessità maggiore nell’encoder e nel decoder.
- *scalable profile*: pensato per un uso futuro. Permette la trasmissione di un’immagine di qualità base (base layer) in termini di risoluzione spaziale (spatially scalable profile) o accuratezza

di quantizzazione (SNR scalable profile), e di informazioni supplementari che permettono di valorizzare le caratteristiche dell'immagine. Questo potrebbe ad esempio servire per trasmettere un'immagine in modo compatibile sia per la ricezione con decoder base sia con decoder HD.

◦ *high profile*: progettato per trasmissioni HDTV in formato 4:2:0 e 4:2:2.

In questo momento, la combinazione più importante per i consumer broadcast in Europa è definita “*main profile at main level (MP@ML)*” che corrisponde alla codifica MPEG-2 di immagini interlacciate nel formato 4:2:0 con una risoluzione di  $720 \times 480$  @ 30 Hz o  $720 \times 576$  @ 25 Hz, sfruttando tra l'altro la codifica delle immagini I, P, B. In base al compromesso raggiunto tra bit-rate, qualità e natura delle immagini da trasmettere, il bit rate è generalmente compreso tra 4 Mb/s (qualità simile a PAL o SECAM), e 9 Mb/s (qualità quasi come CCIR-601 Studio).

L'intero processo di codifica e decodifica descritto precedentemente per l'MPEG-1 è valido pure per il formato MPEG-2 (MP@ML), ed in particolare la gerarchia a livelli (dal blocco alla sequenza), mostrato nella Figura 2.4. Vi è però una piccola differenza nella definizione di *slice*, in quanto non necessariamente copre l'intera immagine ed è solamente composta da blocchi continui della stessa riga orizzontale (Fig 2.9).

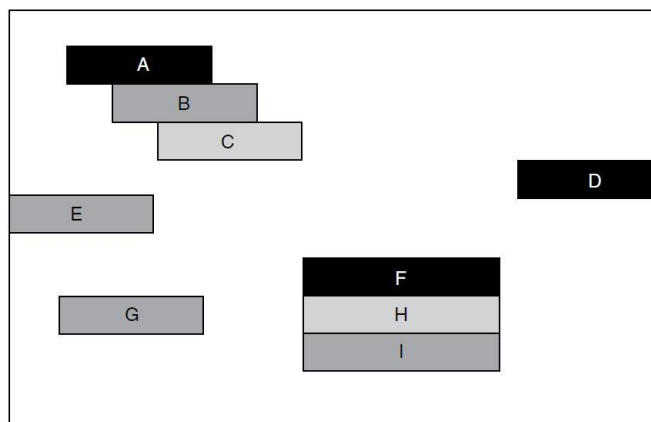


Fig 2.9 Slice MPEG-2 nel caso più generale

La principale novità di MPEG-2 rispetto a MPEG-1 è l'elaborazione di immagini interlacciate, che introduce alcuni metodi specifici e una certa complessità per ottenerla. Per conseguire risultati ottimali, le immagini interlacciate dovranno essere trattate in modo diverso a seconda dell'importanza dei movimenti tra due campi di un'immagine: i casi estremi sono, da un lato, le immagini provenienti da film, dove i due campi provengono da l'immagine stessa (almeno in sistemi a 50 Hz), e dall'altro, le immagini televisive di eventi sportivi dove le differenze dovute al movimento tra i due campi di un'immagine può essere considerevole.

## MPEG-2: specifiche di previsione (immagini interlacciate)

La sequenza temporale della posizione verticale delle linee appartenenti a campi successivi in un sistema interlacciato è mostrato nella Figura 2.10.

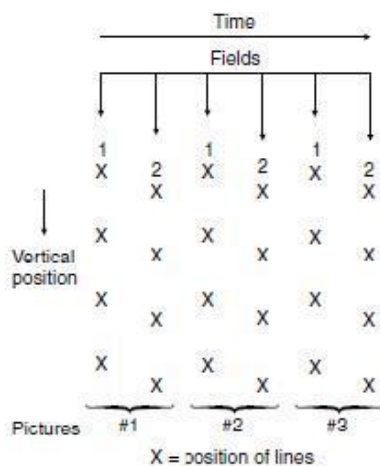


Fig 2.10 Posizione delle linee di successivi campi in sistemi interlacciati

Per la codifica delle immagini interlacciate, MPEG-2 permette di scegliere due strutture possibili:

- ◇ La struttura *frame* (detto anche progressiva) è più adatta per i casi dove c'è poco movimento tra due campi successivi. Macroblocchi e blocchi vengono tagliati fuori dal frame completo (Fig. 2.11), e così la DCT viene applicata a punti verticali consecutivi separati uno dall'altro da un periodo di 20 ms (durata di un campo in sistemi a 50 Hz), che non portano nessun problema alle parti di un'immagine con poco movimento. In questa modalità, tuttavia, è possibile codificare i blocchi più animati nella modalità field, il che significa posizionare i blocchi in un solo campo.

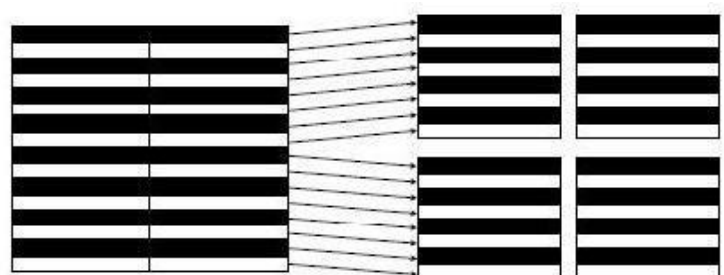
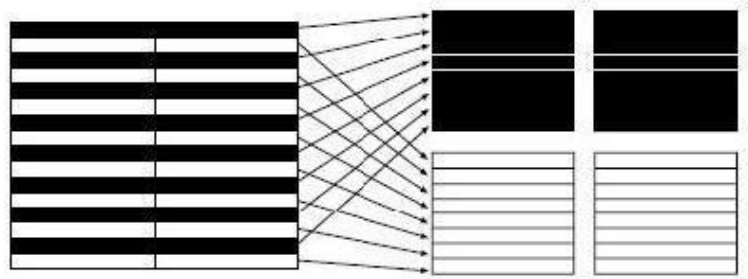


Fig 2.11  
Frame mode: ritaglio blocchi da macroblocchi



- ◇ La struttura *field* (detta anche interlacciata) è preferibile quando avvengono importanti movimenti tra campi successivi; in questo caso, al fine di evitare una consistente componente a frequenza verticale che ridurrebbe l'efficienza della successiva fase relativa alla DCT, macroblocchi vengono tagliati da un campo (Figura 2.12), che viene così considerato come immagine indipendente. Nel caso di stima del movimento sono possibili diversi modi; un macroblocco può essere previsto in frame, field o un attraverso l'unione dei due.

Fig 2.12  
Field mode: ritaglio blocchi da macroblocchi



- ◇ Nella modalità frame, un macroblocco preso da un campo dispari è utilizzato per prevedere il corrispondente campo dispari nel macroblocco successivo e lo stesso succede per i campi pari. Il vettore di movimento in questo caso risulta pari alla durata di due campi (40 ms).
- ◇ In modalità field, la previsione di un macroblocco è fatta utilizzando il campo precedente e i vettori di movimento corrispondono alla durata di un campo (20 ms)
- ◇ In modalità mista, la previsione è fatta da macroblocchi appartenenti a due frame.

Gli schemi di encoder e decoder MPEG-1 delle figure 2.7 e 2.8 valgono anche per l'MPEG-2.

## MPEG audio: principi di funzionamento

L'obiettivo di base per la codifica audio consiste nell'eliminare quella parte di informazione che risulta superflua all'udito umano, e senza la quale il segnale audio rimane comunque comprensibile.

È ben noto che la sensibilità dell'orecchio umano è maggiore per le frequenze che vanno da 1 a 5 kHz. La curva di sensibilità, che rappresenta la soglia di udibilità o percezione in funzione della frequenza, in assenza di qualsiasi "segnale disturbante", è mostrato in Figura 2.13; si può osservare che un segnale è udibile se supera la soglia di udibilità.

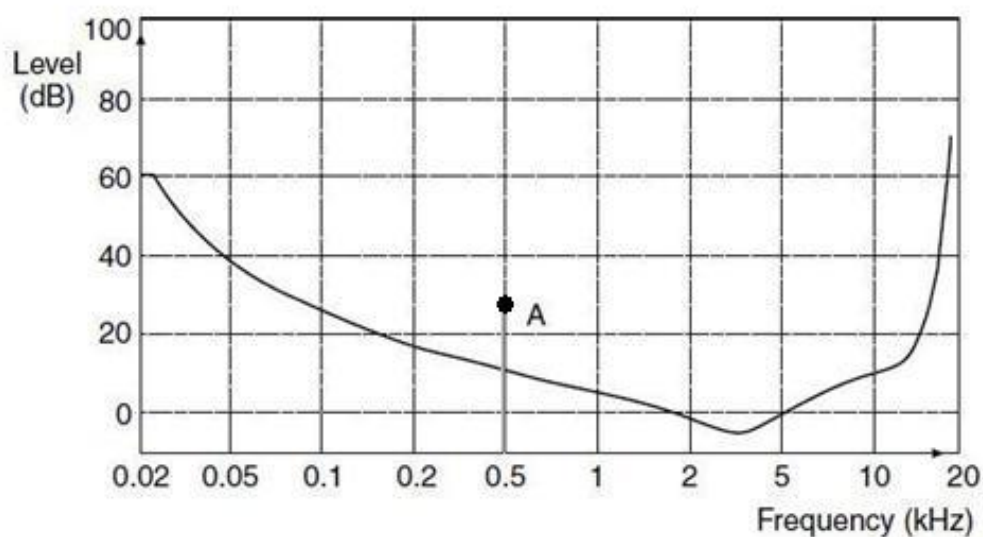


Fig 2.13 Sensibilità dell'orecchio in funzione della frequenza (A è udibile)

Questo grafico non tiene conto della possibile presenza di più segnali con frequenze relativamente vicine tra loro. Infatti, in una situazione come quella appena descritta, l'orecchio risulterebbe meno sensibile nell'intorno di quella frequenza (fenomeno della *frequency masking*).

Inoltre c'è un altro effetto, chiamato *temporal masking*, in cui un suono di ampiezza elevata disturba i segnali che lo precedono e lo seguono.

Al fine di limitare questi effetti si è studiato un *modello psico-acustico* dell'udito umano.

Questo modello è stato utilizzato come base per la realizzazione di un codificatore "percettivo", caratterizzato da una curva di mascheramento e livelli di quantizzazione che variano in funzione dei segnali da codificare. Il principio del processo di codifica consiste per primo nel dividere la banda di frequenza in 32 sottobande di larghezza uguale per mezzo di un banco di filtri polifase. Il segnale di uscita da un filtro sotto-banda corrisponde ad una durata di 32 campioni PCM (Figura 2.14)

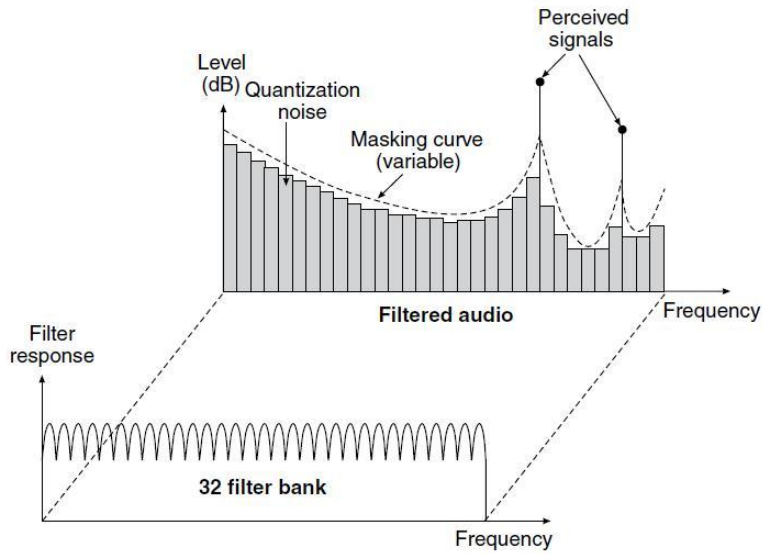


Fig 2.14 Principio di codifica percettiva audio

Le regioni di frequenza dove l'orecchio è più sensibile possono essere quantificate con maggiore precisione rispetto ad altre regioni. Schemi semplificati di encoder e decoder audio MPEG sono mostrati nelle figure 2.15 e 2.16.

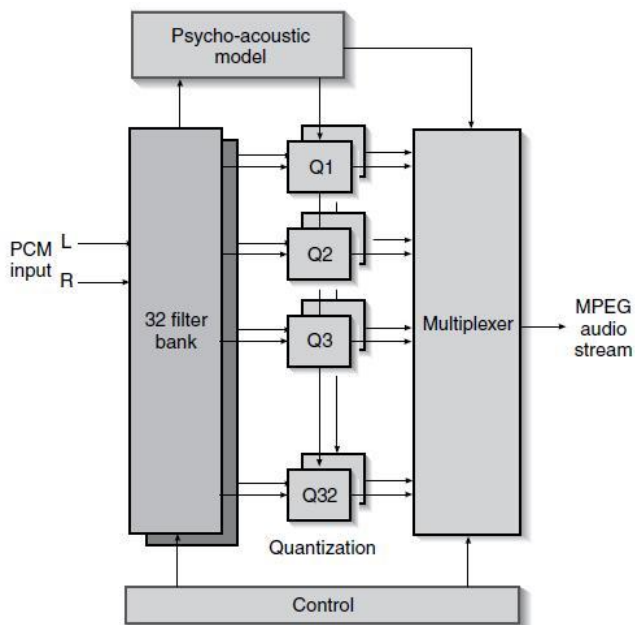


Fig 2.15 MPEG audio encoder

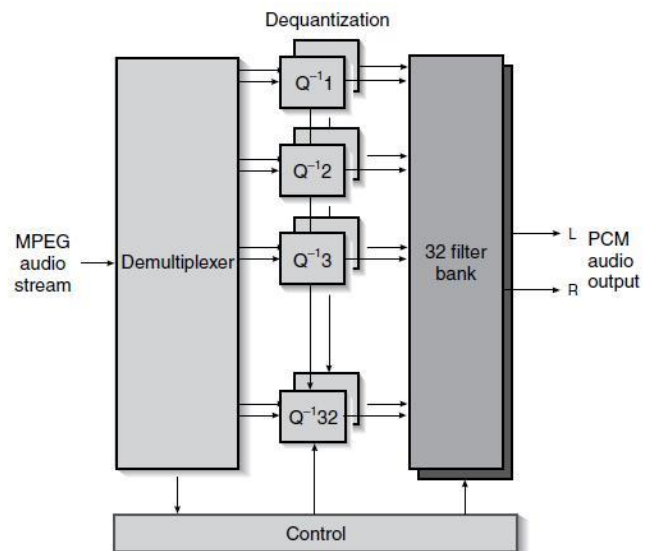


Fig 2.16 MPEG audio decoder

## I layer dell'MPEG audio

Lo standard audio MPEG definisce tre livelli di codifica, che offrono tassi di compressione molto diversi per una data qualità audio percepita:

- *Layer I*: (o pre-MUSICAM) utilizza un bit-rate fisso scelto tra 14 possibili che variano da 32 a 448 kb/s. Il vantaggio principale del layer I è la relativa semplicità del encoder e decoder.

Il modello psico-acustico utilizzato è conosciuto come modello 1. La precisione dei coefficienti di quantizzazione dei filtri sotto-banda è definita per l'intera durata del frame da un numero a 4 bit, che permettono quindi una codifica da 0 a 15 bit per ogni sottobanda.

- *Layer II*: Questa è la modalità principale utilizzata nei sistemi DVB e utilizza l'algoritmo noto come MUSICAM che è stato sviluppato per la radio digitale europea (DAB, *Digital Audio Broadcasting*). Per un audio di qualità equivalente, lo strato II richiede un bit-rate più piccolo dal 30 al 50 % del livello I, a scapito di un moderato aumento di complessità dell'encoder e del decoder. Il bit rate scelto è fissato tra 32-192 kb/s per canale.

Il modello psico-acustico utilizzato è lo stesso dell'layer I (modello 1), ma la durata dei fotogrammi è tre volte più lunga. Al fine di ridurre il tasso di bit, la precisione dei coefficienti di quantizzazione dei filtri sotto-banda diminuisce con la frequenza (definita sulla quantizzazione 4 bit per le bande basse, 3 bit per le bande medie e 2 bit per le bande alte).

- *Layer III*: Questa è l'ormai famoso formato MP3, che è stato sviluppato con un modello psico-acustico diverso (modello 2), una codifica di Huffman e l'analisi del segnale basata sulla DCT. Per una determinata qualità, il rapporto di compressione ottenibile con layer III è circa il doppio di quello del layer II, ma l'encoder e decoder sono sostanzialmente più complessi e il tempo per la codifica/decodifica è molto maggiore.

Il Layer III è utilizzato per applicazioni unidirezionali a basso bit-rate e non viene utilizzato per i sistemi di televisione digitale mentre i sistemi DVB utilizzano i layer I e II dell'audio MPEG-1 in quattro modi possibili:

- stereo: il canale destro e sinistro sono codificati in modo completamente indipendente;
- joint-stereo: sfrutta la ridondanza tra i canali destro e sinistro al fine di ridurre il bit-rate audio
- dual-channel: due canali indipendenti che trasportano suoni non correlati
- mono: un solo canale audio

## MPEG audio: formato del pacchetto

Il frame audio è l'unità elementare di accesso ad una sequenza dell'MPEG audio; è formato da quattro campi principali:

- header a 32 bit,
- parità (CRC) maggiore di 16 bits
- dati audio di lunghezza variabile
- dati ausiliari (AD) di lunghezza variabile

Header	CRC	AUDIO data			AD
System 32 bits	Parity 16 bits	Allocation bits/SBS	Scaling factors	Sub-band samples (SBS) (12 × 32 sub-band samples)	Ancillary data

Fig. 2.17  
MPEG audio  
frame layer I

- The header carries the synchronization and system information (see Table 3.1).
- The use of the parity (CRC) is optional.
- The "bit allocation per SBS" field contains 32 integers coded on 4 bits, each of them defining the resolution for coding the samples of one of the 32 sub-bands.
- The "scaling factors" field contains 32 integers coded on 6 bits, each of them giving the multiplication factor of the samples of one sub-band.

Header	CRC	AUDIO data				AD
System 32 bits	Parity 16 bits	Allocation bits/SBS	Selection SCFSI	Scaling factors	Sub-band samples (SBS) (three portions of 12 sub-band samples each)	Ancillary data

Fig. 2.18  
MPEG audio  
frame layer II

- The header carries the synchronization and system information (see Table 3.1).
- The use of the parity (CRC) is optional.
- The “bit allocation per SBS” field contains 32 integers coded on 2–4 bits depending on the sub-band, each of them defining the resolution for coding the samples of one of the 32 sub-bands and whether or not these sub-band samples are grouped in threes.
- The SCFSI (scale factor selection information) field indicates whether the scale factor is valid for the whole frame duration or whether there are two or three different scale factors.
- The “scaling factors” field contains 32 integers coded on 6 bits, each of them giving the multiplication factor of the samples of one sub-band for the portion of the frame defined by the SCFSI.

# Capitolo 3

## DVB - Digital Video Broadcasting

Effettuata la codifica di sorgente (MPEG audio e video, multiplexing, ed eventuale scrambling) e di canale, si ottiene un flusso di dati da trasmettere all'utente finale mediante la modulazione di una portante.

A seconda del tipo di mezzo trasmissivo utilizzato (satellite, cavo, rete terrestre), la larghezza di banda disponibile per la trasmissione dipende da numerosi fattori tecnici e legislativi ed in particolare i fattori tecnici variano molto a seconda della provenienza del segnale (soprattutto considerando l'SNR e l'eco). In genere:

- ricezione satellitare: il rapporto segnale-rumore è molto piccolo ( $\leq 10$  dB) ma soffre di una presenza consistente di echi.
- ricezione terrestre: le condizioni sono più complicate, specialmente se è richiesta una ricezione mobile con semplici antenne (echi variabili a causa di cammini multipli, interferenze, ampie variazioni dei livelli di segnale)

## Sistemi analogici

Per quanto riguarda i sistemi di diffusione analogici vengono utilizzate tecniche di modulazione diverse; nei sistemi digitali si devono ottimizzare le specifiche sui vincoli del canale di trasmissione per ottenere la compatibilità con i sistemi precedenti:

- satellite:
  - la larghezza di banda del canale è generalmente compresa tra 27 e 36 MHz;
  - viene utilizzata una modulazione di frequenza (FM) per la trasmissione di programmi TV analogici (larghezza di banda 6-8 MHz con relativi supporti sonori);
- reti terrestri:
  - la larghezza di banda del canale varia da 6 (Stati Uniti) a 7 o 8 MHz (Europa);

- viene utilizzata una modulazione di ampiezza (AM) con banda laterale vestigiale (VSB) per il video e una o più portanti audio;

Le trasmissioni digitali devono generalmente sfruttare le caratteristiche appena definite, quindi utilizzare la stessa larghezza di banda di canale relativa alla propria controparte analogica in modo che possano coesistere ed essere compatibili trasmissioni analogiche e digitali.



## DVB-S: caratteristiche della modulazione

Al fine di determinare la miglior modulazione possibile, una serie di studi teorici e prove pratiche sono state eseguite via satellite. In Figura 3.1 si può vedere, in condizioni ideali, il tasso teorico di errore sul bit (*Bit Error Rate, BER*) come funzione del rapporto SNR per modulazioni in quadratura, partendo dalla 4-QAM (QPSK) fino alla 64-QAM. Si può osservare che per un dato BER, la QPSK è migliore, in termini di SNR, della 64-QAM di quasi 12 dB.

Considerando il rapporto segnale-rumore a livello ricevitore, il massimo, quindi la migliore efficienza spettrale, è stato trovato utilizzando 2 bit/simbolo cioè una modulazione QPSK.

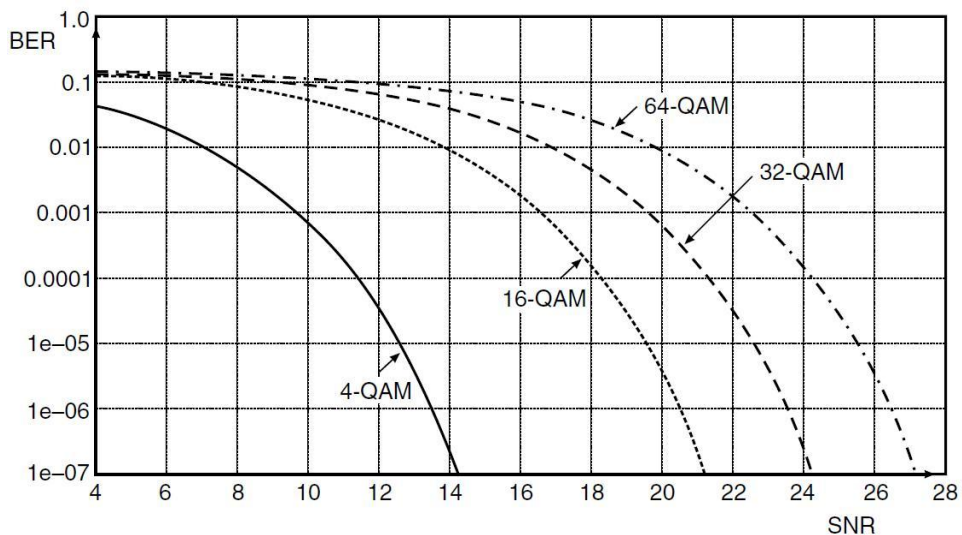


Fig 3.1 BER come funzione del SNR per modulazione QAM

In figura 3.2 si può notare l'effetto del rumore sulla costellazione QPSK in uscita dal demodulatore di una trasmissione satellitare rumorosa.

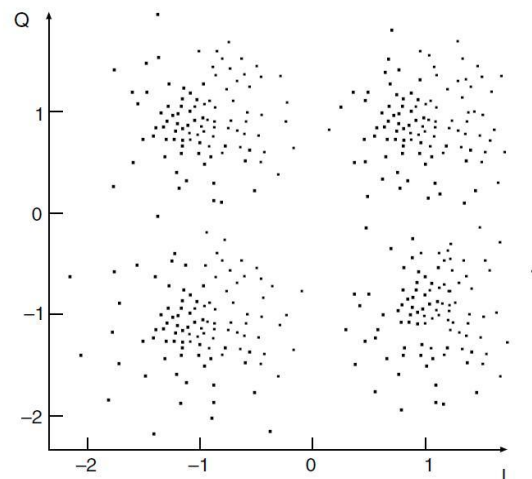


Fig 3.2 Costellazione QPSK affetta da rumore

Si può capire che, sopra un certo livello di rumore, il demodulatore non sarà in grado di distinguere con certezza un punto della costellazione dai suoi vicini. Maggiore è il numero di punti della costellazione, minore sarà il livello di rumore massimo accettabile.

Esistono altri tipi di disturbi in trasmissione e in ricezione (risposta in frequenza imperfetta, interferenze, intermodulazione, ecc...) che accrescono l'interferenza di intersimbolo (*ISI*) e appaiono come rumore sulla costellazione.

Un altro problema a cui il ricevitore deve far fronte nel caso di modulazione digitale QAM è che non esiste un riferimento assoluto di fase per la demodulazione della costellazione (in contrasto con la demodulazione della sottoportante NTSC o PAL, dove si invia un burst di riferimento all'inizio di ogni riga). Per questo motivo si ha un'ambiguità di fase di 90° che impedirà la sincronizzazione del ricevitore fino a che la fase di demodulazione sarà corretta.

## Encoder: elaborazione del segnale

In seguito verranno definite le caratteristiche dei processi di modulazione e codifica per lo standard satellitare seguendo lo schema di figura 3.3

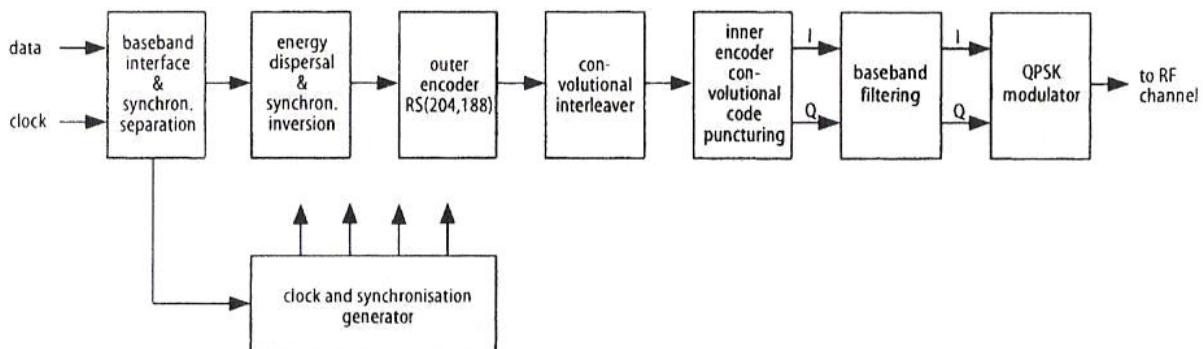


Figura 3.3 Diagramma a blocchi dell'elaborazione del segnale da trasmettere

Le seguenti caratteristiche del flusso di trasporto sono particolarmente importanti per codifica e modulazione nella trasmissione satellitare:

- Il flusso di trasporto MPEG-2 è composto da singoli pacchetti (frames) con una lunghezza di 188 bytes ognuno. I primi quattro byte formano l'header e il primo è il byte di sincronizzazione (sync byte).
- Il "Transport-error indicator bit" è definito nell'header. Se il pacchetto non è più decodificabile per i troppi errori nel canale, allora questo bit è utilizzato per indicare un pacchetto errato.

## Scrambling

Il flusso di trasporto MPEG-2 viene identificato come una sequenza di pacchetti dati di lunghezza fissa pari a 188 byte. Con una tecnica chiamata scrambling (*dispersione d'energia*), la sequenza di byte viene quindi scorrelata per generare un flusso di dati con densità spettrale di potenza il più possibile uniformemente distribuito.

I segnali sono combinati bit a bit con il flusso di uscita di un generatore pseudo-random, attraverso un'operazione di "or-esclusivo" (addizione in modulo 2). Il generatore è implementato con un registro a scorrimento retroazionato, che viene re-inizializzato all'inizio di ogni ottavo frame in accordo con un predeterminato bit-pattern.

## Codifica error-protection

E' il passo che segue lo scrambling, che prevede una codifica concatenata da tre funzioni: codificatore codice di blocco, interleaver e codificatore convoluzionale.

Il primo livello di protezione viene fornito ai dati trasmessi, utilizzando un codice a blocchi non binario di tipo Reed-Solomon RS(204,188), che permette la correzione di un massimo di 8 byte errati per ogni pacchetto di 188 byte.

L'interleaver utilizza una tecnica di *interleaving convoluzionale* per "mescolare" la sequenza di dati trasmessa, in modo da renderla più robusta in caso di lunghe sequenze di errori.

Il codificatore convoluzionale offre un secondo livello di protezione sfruttando un codice convoluzionale binario con puncturing, che spesso viene indicato con il termine FEC (*Forward Error Correction*). I valori di codifica ammessi sono cinque: 1/2, 2/3, 3/4, 5/6, 7/8.

## Filtraggio

Successivamente alla codifica per la protezione d'errore, i dati sono pre-filtrati e poi inviati al modulatore.

Il filtraggio è utilizzato per soddisfare questi due obiettivi:

- il segnale deve avere una banda limitata per evitare la possibilità di *crosstalk* con i canali adiacenti
- inoltre deve soddisfare il primo criterio di Nyquist

Il filtro di Nyquist a coseno rialzato e fattore di roll-off  $\alpha=0.35$  è stato scelto come filtro caratteristico per gli standard satellitari.

## Modulazione

I segnali filtrati arrivano all'ingresso del modulatore QPSK come due componenti: I (in fase) e Q (in quadratura), (vedi figura 3.4).

La codifica Gray alloca ad ogni punto della costellazione i relativi bit di codifica, e se nella decodifica il segnale oltrepassa una sola delle soglie di decisione allora si avrà soltanto un bit errato.

Prima della trasmissione, il segnale IF è convertito in radio frequenza.

In questo modo due bit per simbolo sono trasmessi; quindi in circostanze ideali larghezza di banda utilizzata è  $B=2 \text{ bit}/(\text{s} \times \text{Hz})$ , ma considerando la non idealità del filtro è sicuramente minore.

A seconda della configurazione del sistema, il rapporto tra larghezza di banda e symbol rate ( $BW/R_s$ ) è circa di:

$$BW / R_s = 1.27 \text{ Hz} / (\text{symbols/s})$$

Quindi la larghezza di banda utilizzata è:

$$B = (\text{bit per symbol}) / (BW/R_s) = (2 \text{ bits} / \text{symbol}) / [1.27 / (\text{symbols/s})] = 1.57 \text{ bit} / (\text{s Hz})$$

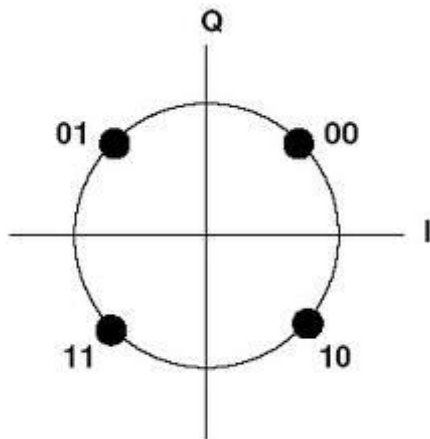


Fig 3.4 Costellazione QPSK

## Decoder: decodifica del segnale

Per la ricezione dei segnali da satellite, il processo di decoding in generale si potrebbe descrivere come il processo inverso effettuato dall'encoder. In più, gli errori che avvengono nel canale devono essere corretti e inoltre, è necessario recuperare l'informazione per il sincronismo.

In figura 3.5 si può vedere lo schema a blocchi del ricevitore.

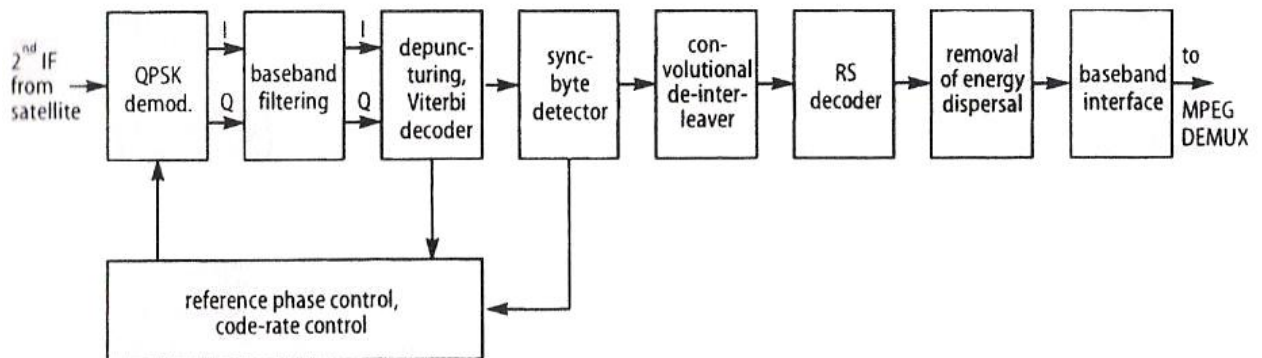


Fig 3.5 Diagramma a blocchi per decodifica lato ricevitore

## Demodulatore

Il demodulatore QPSK prima di tutto deve recuperare la frequenza della portante dal segnale d'ingresso. Un particolare problema è l'incertezza di fase della portante, che deve essere aggiustata per permettere una demodulazione corretta di I e Q. Il segnale d'ingresso può assumere una di quattro possibili posizioni di fase formanti un angolo retto tra loro, ognuna delle quali rappresenta due bit (Fig 3.4).

Il demodulatore può essere implementato insieme al sistema di recupero della fase. La frequenza portante per la demodulazione è creata da un voltage-controlled oscillator (VCO). Il VCO è gestito da un segnale di controllo generato da i due segnali in banda base demodulati in un loop di controllo.

La decisione sulla correttezza della posizione di fase può essere presa solamente dal decoder, in due passaggi successivi. Un'incertezza di fase di  $\pm 90^\circ$  può essere risolta in un primo step, e in un secondo l'incertezza residua di  $180^\circ$  può essere rimossa.

Una correzione di  $90^\circ$  può essere ottenuta scambiando le componenti I e Q dopo la demodulazione e successivamente invertendo una di loro.

Un errore di fase di  $180^\circ$  è compensato da un'inversione del flusso di bit nella posizione nel decoder al quale è stato individuato.

## Filtraggio e recupero del clock

In seguito alla demodulazione, i segnali in banda base vengono filtrati con un filtro a radice quadrata del coseno-rialzato, come nel caso della trasmissione. Il filtro può essere implementato come un filtro digitale dopo un sovracampionamento del segnale demodulato in banda base oppure con un filtro passa-banda analogico prima del demodulatore.

Il clock del segnale filtrato è determinato con l'aiuto di un circuito PLL (*Phase locked loop*) in modo che il segnale possa essere campionato.

## Decoder Viterbi

Un decoder Viterbi utilizza l'algoritmo di Viterbi per la decodifica di un flusso di bit codificato sulla base di un codice convoluzionale utilizzando il metodo FEC per la correzione degli errori.

Durante il processo, il decoder Viterbi conta il numero degli errori che son stati identificati e trasmette questo numero all'unità di controllo che ha il compito di determinare i parametri di trasmissione sconosciuti:

- fase di riferimento per la demodulazione (incertezza possibile di 90°)
- schema di puncturing
- sincronismo per il depuncturing; questo processo è periodico, orientato al blocco e simile al puncturing

All'uscita del decoder i dati poi vengono ricodificati e ripristinato il puncturing, in accordo con i valori trovati nella decodifica.

## Rilevatore del sync-byte

La decodifica richiede la suddivisione del flusso di dati in pacchetti MPEG-2 TS, e poi la delimitazione delle strutture di 8-pacchetti per la rilevazione dell'energia da rilevare. Il sync-byte all'inizio di ogni frame è utilizzato per questo.

Lo scopo del rilevatore di sync-byte è proprio quello di determinarne la posizione per poi ricavarne il sincronismo di segnale per il decoder RS, il de-interleaver e per la rimozione della dispersione di energia.

## De-interleaver, RS decoder e Energy-dispersal Remover

Il canale di trasmissione satellitare è essenzialmente un canale Gaussiano, il quale introduce un rumore Gaussiano al segnale. Il de-interleaver opera la funzione inversa dell'interleaver, cioè ridispone in maniera ordinata la sequenza arrivata. Il decoder RS

opera anch'esso la funzione inversa dell'encoder e restituisce l'informazione decodificata per poi passarla all'energy-dispersal remover che effettua l'operazione inversa di quella che avviene nel trasmettitore utilizzando un circuito analogo.

### Interfaccia in banda base

Dopo che l'MPEG-2 TS è stato ricostruito, viene trasferito al TS demultiplexer. Per la segnalazione di errori che sono avvenuti nel canale di trasmissione e che sono stati riconosciuti dall'*error protection* ma non son è stato possibile correggerli, MPEG-2 fornisce un elemento di segnalazione chiamato "*transport-error indicator bit*" che si trova al primo posto appena dopo il sync-bye. Questo bit permette al demultiplexer di identificare il rispettivo pacchetto come errato e di prendere le necessarie precauzioni.

### DVB-S: caratteristiche

Tenendo conto di tutte le considerazioni viste precedentemente, le principali caratteristiche adottate per le trasmissioni TV digitali sono riportate in Tabella 3.1 Tabella 3.2 dà il massimo bit-rate su un canale satellitare DVB in funzione della larghezza di canale e code rate per una modulazione QPSK con  $\alpha=0,35$ .

Parametri	DVB-S satellite
Channel width	26–54 MHz
Modulation type	QPSK (= 4-QAM)
Roll-off factor ( $\alpha$ )	0.35
Frequency range	10.7–12.75 GHz

Tabella 3.1 Caratteristiche trasmissioni TV digitali DVB-S

Channel width (MHz)	Maximum symbol rate (MHz)	Maximum useful bit-rate (Mb/s)				
		Rc=1/2	Rc=2/3	Rc=3/4	Rc=5/6	Rc=7/8
54	42.2	38.9	51.8	58.3	64.8	68.0
46	35.9	33.1	44.2	49.7	55.2	58.0
40	31.2	28.8	38.4	43.2	48.0	50.4
36	28.1	25.9	34.6	38.9	43.2	45.4
33	25.8	23.8	31.7	35.6	39.6	41.6
30	23.4	21.6	28.8	32.4	36.0	37.8
27	21.1	19.4	25.0	29.2	32.1	34.0
26	20.3	18.7	25.0	28.1	31.2	32.8

Tabella 3.2 Massimo bit-rates in funzione dell'ampiezza di canale e di code rate (DVB-S) (da prETS 300 421, © European Telecommunication Standards Institute 1995)

## DVB-T: caratteristiche

All'inizio del DVB project la priorità della televisione digitale era quella di poter trasmettere via satellite e via cavo; solo al completamento e alla messa in funzione di questi due standard si iniziò ad elaborare delle specifiche per un possibile standard per la trasmissione terrestre, che venne adottato nel 1995.

La trasmissione terrestre, dal punto di vista dei requisiti d'utente e per quanto riguarda le caratteristiche tecniche e della linea di trasmissione, è molto più complessa delle trasmissioni via cavo e satellite. Inoltre, finché i due sistemi erano in fase di sviluppo è stato possibile ampliare le conoscenze tramite l'esperienza pratica acquisita nel campo professionale, per esempio riguardo alle tecniche di modulazione, e successivamente, furono stabilite le condizioni sugli standard per le trasmissioni digitali terrestri, che dovevano anche soddisfare il cliente finale:

- Il sistema per la trasmissione terrestre che si adottò ebbe caratteristiche molto simili ai sistemi via satellite e via cavo. Questo assicurò una tecnologia del ricevitore di casa molto simile a quella utilizzata per sistemi satellitari e via cavo.
- I programmi DVB dovevano essere trasmessi in pacchetti dati con una capacità la più grande possibile. La larghezza di banda di canale scelta in Europa dovrà supportare un canale di 8 MHz. I canali che richiedevano 7 MHz non furono utilizzati, il che vuol dire non furono sfruttate le bande I e III (VHF).
- Il sistema doveva permettere una copertura d'area ottima per la ricezione stazionaria con antenne a tetto. La ricezione mobile non era ancora un obiettivo di sviluppo.
- Possibilità per segnali DVB di essere trasmessi in reti terrestri a singola frequenza. Le reti a singola frequenza consistono di trasmettitori che trasmettono esattamente un identico flusso di dati in modo sincronizzato con ogni altro flusso, sfruttando la stessa frequenza di trasmissione.

Il requisito per l'utilizzo di reti a singola frequenza porta automaticamente alla scelta della tecnica di modulazione definita come OFDM (*orthogonal frequency division multiplex*) e la combinazione di questa tecnica di modulazione e il metodo di protezione degli errori utilizzata prende il nome di COFDM (*coded orthogonal frequency division multiplex*).

In seguito alle richieste di stati non europei e alla crescita della ricerca sullo sviluppo della ricezione mobile, dove evidentemente si richiedono utilizzi di grandezze di canale



anche di 6 e 7 MHz, lo standard DVB-T ha considerato nel suo standard la possibilità di utilizzo di queste.

## Encoder: elaborazione del segnale

Numerose componenti del sistema per la trasmissione terrestre dei segnali DVB sono identiche a quelle utilizzate dal sistema DVB-S precedentemente descritto. In figura 3.6 si può vedere il diagramma a blocchi dell'encoder DVB-T. Le componenti in grigio sono già state discusse nella sezione del DVB-S.

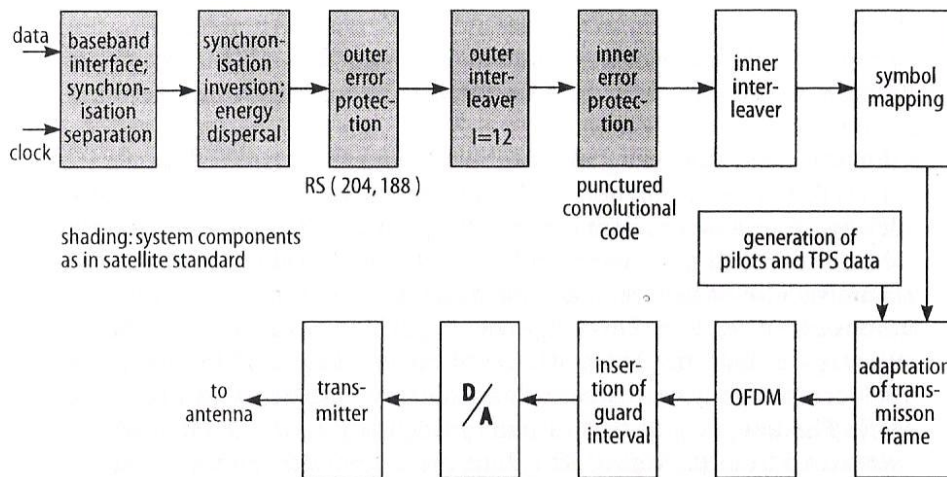


Fig 3.6 Diagramma a blocchi di un encoder DVB terrestre

## Interleaver interno e Symbol Mapping

Dato che l'OFDM è una tecnica di modulazione multi portante, che consiste nel modulare simboli di durata  $T_s$  un elevato numero di portanti  $N$ , con una distanza tra portanti adiacenti di  $1/T_s$ .

L'interleaving opera in due passaggi. In un primo step, 126 bit vengono combinati in un blocco unico e successivamente si opera l'interleaving bit a bit e quindi un grande numero di blocchi vengono racchiusi da un solo blocco, all'interno del quale gruppi di bit (simboli) vengono "mescolati" tra loro (simbol interleaving).

A seconda della scelta del metodo di modulazione utilizzata per ogni singola portante, vengono utilizzati un numero determinato di bit. Le possibili scelte sono: QPSK, 16-QAM o 64-QAM.

Per la QPSK sono richiesti 2 bit per la modulazione di ogni portante (uno per la parte reale e uno per quella immaginaria, in accordo con i 4 punti possibili della costellazione), invece per la 16-QAM ne sono richiesti 4 (2 per ogni asse, in modo da soddisfare i 16 punti possibili della costellazione) e infine la 64-QAM ne richiede 6 (3 per la parte reale e 3 per quella immaginaria, 64 punti possibili della costellazione). Conseguentemente, per la creazione di portanti OFDM modulate QPSK, si utilizzano due funzioni di interleaving bit a bit connesse in parallelo. Per la 16-QAM se ne utilizzano quattro e per la 64-QAM sei. In figura 3.7 si può vedere un esempio di una struttura di interleaving interno per la generazione di 16-QAM simboli.

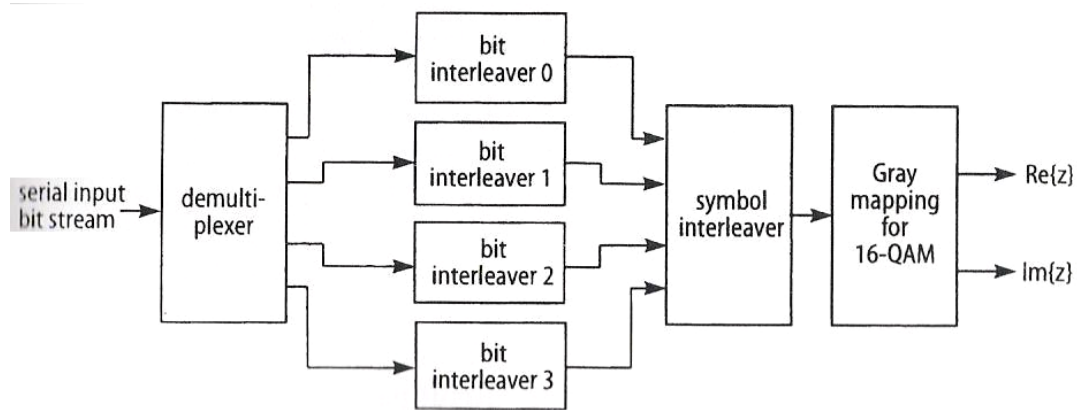
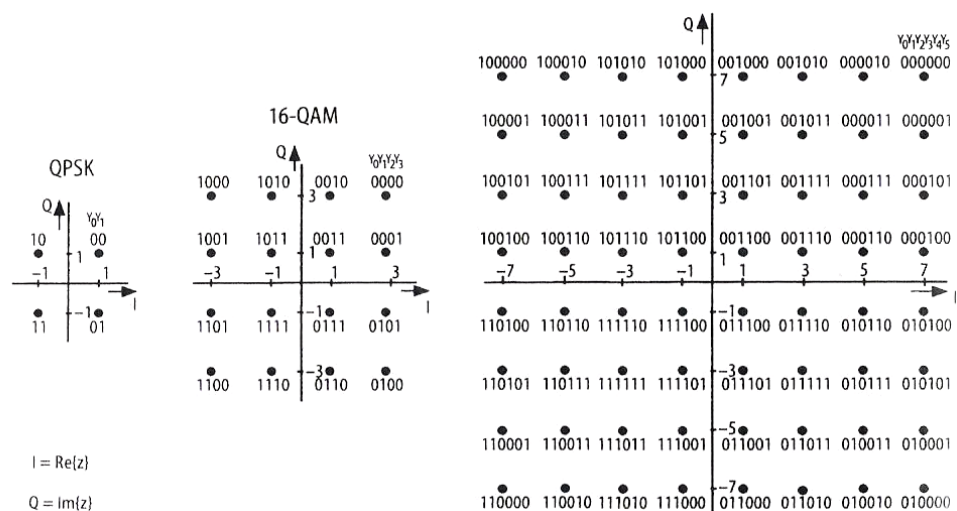


Fig 3.7 Inner interleaver 16-QAM

Il secondo step di operazioni dell'interleaving consiste nell'operare sui simboli uscenti dall'operazione bit a bit descritta precedentemente e dipende quindi dalla tecnica di modulazione delle portanti utilizzate. L'uscita non porterà ad un flusso di dati continuo ma bensì un susseguirsi di blocchi ad intermittenza.

Successivamente i simboli vengono mappati in strutture di bit secondo il metodo Gray, che differenzia parole di codice adiacenti al massimo di un bit e tutto questo per aumentare l'affidabilità del sistema. Il risultato di un "Gray mapping" si può vedere in figura 3.8.

Fig. 3.8 Costellazioni possibili delle portanti DVB-T



## Adattamento trama

I simboli generati dal processo di mappatura vengono raccolti in blocchi di lunghezza costante pari a 1512, 3024 o 6048 simboli per blocco. Si genera una trama (frame) di lunghezza pari a 68 blocchi e quattro trame (frame) costituiscono una supertrama (superframe).

## OFDM: scelta dei parametri

Nell'OFDM molti sono i parametri che vengono scelti, in modo indipendentemente l'uno dall'altro, e che quindi determinano le performance di tutto il sistema.

Uno dei parametri cruciali è la lunghezza dell'intervallo di guardia. Se reti a singola frequenza in cui la distanza tra trasmettitori e locazioni riceventi corrispondenti è la massima possibile, per esempio 60 km, allora la lunghezza dell'intervallo di guardia non può più essere scelta liberamente. Deve essere almeno di 200  $\mu\text{s}$  ( $200 \mu\text{s} \times 300,000 \text{ km/s} = 60 \text{ km}$ ). Così la durata dell'intervallo di guardia implica una riduzione in termini di tempo utile per il trasferimento di dati, che riduce quindi la capacità utile del canale. La lunghezza dell'intervallo di guardia va tenuto relativamente piccolo rispetto alla durata dei simboli  $T_s$ . In molte applicazioni, come ad esempio la trasmissione su di un'area metropolitana invece che su tutto il territorio nazionale, l'utilizzo di un intervallo di guardia di 200  $\mu\text{s}$  non è appropriato e porta solo ad una inutile riduzione della capacità di canale, quindi questo parametro va gestito caso per caso. Ad esempio, per trasmissioni con distanza tra trasmettitori e ricevitori di 15 km si può considerare come buono un intervallo di guardia di 50  $\mu\text{s}$ .

Nella modulazione OFDM di sistemi DVB-T esistono due possibili scelte del numero di portanti: modo a 2K e modo a 8K, la prima composta da 1512 portanti e la seconda da 6048, separate rispettivamente di 4 kHz o di 1 kHz.

## DAC e front-end

Il segnale digitale viene opportunamente trasformato in un segnale analogico da un convertitore digitale-analogico (*DAC, Digital to Analog convert*), ed infine modulato a frequenza radio (UHF o VHF). La larghezza di banda occupata da ciascun singolo segnale DVB-T è adatta per la trasmissione sul canale predefinito che generalmente ha larghezza di banda di 8 MHz (si stanno diffondendo altre larghezze di banda di 6 e 7 MHz, come detto precedentemente).

## Decoder: decodifica del segnale

In ricezione, il decoder contiene tutti i circuiti necessari ad effettuare la demodulazione e la decodifica del segnale ricevuto, utilizzando tecniche duali, ovvero complementari o inverse, a quelle descritte per la trasmissione. In figura 3.9 si può vedere uno schema base dei processi per la decodifica.

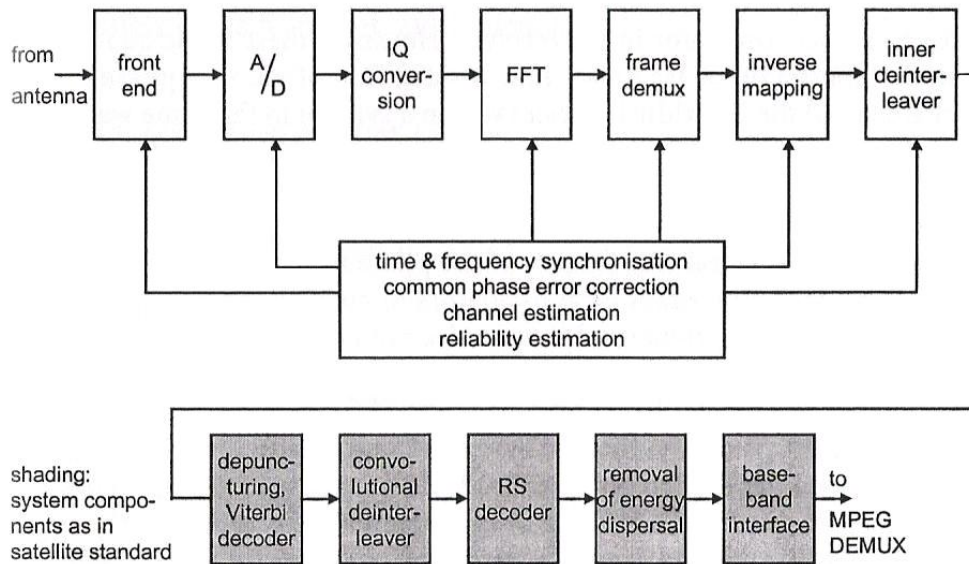


Fig 3.9 Diagramma a blocchi dei processi di un segnale nel ricevitore

Un ricevitore DVB-T deve far fronte a notevoli degradazioni che avvengono durante le trasmissioni terrestri. In più, il ricevitore introduce ulteriori errori (anche nella fase) nel segnale da demodulare. Ricevitori possono differire nei modi di gestire i sincronismi di tempo e frequenza nonché della stima di canale. Le varie tecniche di stima del canale hanno una notevole importanza nelle prestazioni del ricevitore come ha grande importanza la stabilità dell'oscillatore; infatti, da un lato si può evitare il fenomeno del crosstalk, dato che le portanti sono approssimativamente separate da 1.1 kHz o 4.4 kHz, campionando esattamente al valore di zero della funzione  $\sin(x)/x$ , dato che la frequenza centrale della portante considerata non viene attraversata dalle portanti che le circondano (condizione di ortogonalità, figura 3.10); dall'altro lato, bisogna mantenere la stabilità per recuperare l'informazione di fase contenuta nei segnali modulati QAM o QPSK.

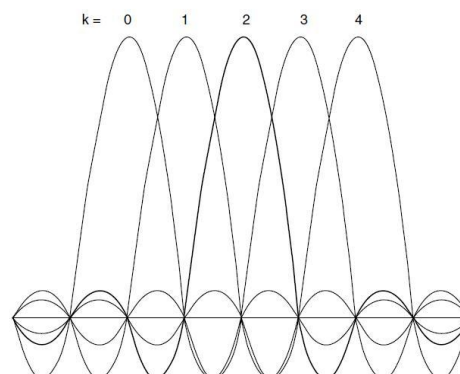


Fig. 3.10 Spettro di portanti adiacenti con modulazione OFDM

Esistono diversi contesti sul modo di ricezione dei segnali: stazionario, portatile e mobile. Da qui si può ben capire che i componenti dei sistemi di ricezione, dai ricevitori alle antenne, saranno ben diversi da un caso all'altro.

Per un primo generale approccio si può iniziare con una versione classica dei sistemi di ricezione. Seguendo l'immagine di figura 3.11 si può subito notare la presenza di un sintonizzatore (*tuner*) per la conversione del segnale RF, scelto tra le frequenze da 47 a 230 e da 470 a 860 MHz, in un primo segnale IF ( $F_{IF}=36.15\text{MHz}$ ), di un elaboratore di segnali IF che restituisce un segnale in banda base ( $F_S$  nell'ordine di 4.5MHz) adatto per la demodulazione e la decodifica al decoder DVB-T che a sua volta pone in uscita un segnale MPEG-2.

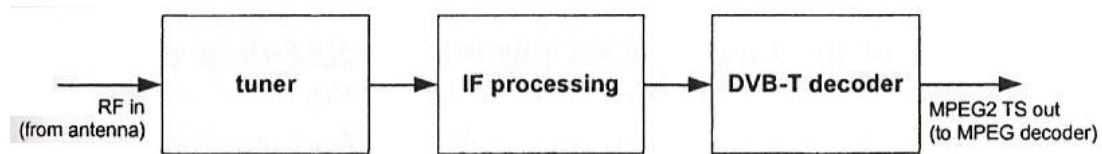


Fig 3.11 Diagramma a blocchi top-level per la ricezione di segnali DVB-T

I passi d'interesse per la descrizione sono quelli che portano alla demodulazione COFDM dove i segnali I e Q sono ricostruiti dal segnale digitale IF, e il segnale è demodulato attraverso una *fast Fourier transform* (FFT) di 2k o 8k punti a seconda della modalità utilizzata (numero di portanti diverse).

## Prestazioni dello standard

Le prestazioni dello standard DVB-T è analizzato secondo due criteri: (1) velocità di trasmissione utile disponibile, (2) rapporto carrier-to-noise nel canale di trasmissione richiesto per ricezione QEF (quasi error free, mezzo con basso tasso d'errore).

Nella pagina seguente sono riportate due tabelle: la prima di valori caratteristici del bit-rate utile, in funzione del tipo di modulazione, del code rate e dell'intervallo di guardia e nella seconda del rapporto C/N nel canale di trasmissione relativo a tre tipi di canale (Gaussiano, Ricean, Rayleigh).

Tabelle delle prestazioni:

Modulazione	Code rate	Relative length of guard interval ( $\Delta$ )			
		1/4	1/8	1/16	1/32
QPSK	1/2	4.98	5.53	5.85	6.03
	2/3	6.64	7.37	7.81	8.04
	3/4	7.46	8.29	8.78	9.05
	5/6	8.29	9.22	9.76	10.05
	7/8	8.71	9.68	10.25	10.56
16-QAM	1/2	9.95	11.06	11.71	12.06
	2/3	13.27	14.75	15.61	16.09
	3/4	14.93	16.59	17.56	18.10
	5/6	16.59	18.43	19.52	20.11
	7/8	17.42	19.35	20.49	21.11
64-QAM	1/2	14.93	16.59	17.56	18.10
	2/3	19.91	22.12	23.42	24.13
	3/4	22.39	24.88	26.35	27.14
	5/6	24.88	27.65	29.27	30.16
	7/8	26.13	29.03	30.74	31.67

Tabella 3.3 Bit rate utile relativo a diversi intervalli di guardia in Mbit/s

Modulazione	Code rate	Gaussian channel [dB]	Ricean channel [dB]	Rayleigh channel [dB]
QPSK	1/2	3.1	3.6	5.4
	2/3	4.9	5.7	8.4
	3/4	5.9	6.8	10.7
	5/6	6.9	8.0	13.1
	7/8	7.7	8.7	16.3
16-QAM	1/2	8.8	9.6	11.2
	2/3	11.1	11.6	14.2
	3/4	12.5	13.0	16.7
	5/6	13.5	14.4	19.3
	7/8	13.9	15.0	22.8
64-QAM	1/2	14.4	14.7	16.0
	2/3	16.5	17.1	19.3
	3/4	18.0	18.6	21.7
	5/6	19.3	20.0	25.3
	7/8	20.1	21.0	27.9

Tabella 3.4 Rapporto C/N minimo in canali trasmissivi con modulazioni per ricezioni QEF

## Caratteristiche degli standard: tabella riassuntiva

Parameter	Cable (DVB-C)	Satellite (DVB-S)	Sat. (DVB-S2)	Terr. (DVB-T)	Mobile (DVB-H)
Video coding	MPEG-2 or MPEG-4 AVC (H264)			H264, H263 ...	
Audio coding	MPEG-1 (layer II), Dolby Digital or AAC			G7xx, AAC	
Scrambling	DVB-CSA (Common Scrambling Algorithm)			tbd	
Link layer FEC	N.A.	N.A.	N.A.	N.A.	MPE-FEC (optional)
Transport packet	188 bytes	188 bytes	188 bytes (1)	188 bytes	188 bytes
External channel coding	Reed-Solomon (204,188, $T = 8$ )	Reed-Solomon (204,188, $T = 8$ )	BCH	Reed-Solomon (204,188, $T = 8$ )	Reed-Solomon (204,188, $T = 8$ )
Byte interleaving	Forney, depth = 12	Forney, depth = 12	N.A.	Forney, depth = 12	Forney, depth = 12
Internal channel coding none	Convolutional, $R_c = 1/2, 2/3, 3/4, 5/6,$ or $7/8$	LDPC	1/4 to 9/10	Convolutional, $R_c = 1/2, 2/3, 3/4, 5/6,$ or $7/8$	Convolutional, $R_c = 1/2, 2/3, 3/4, 5/6,$ or $7/8$
Bit interleaving	N.A.	N.A.	By blocks (8PSK, 16 and 32 APSK)	N.A.	N.A.
Roll-off factor	15%	35%	20, 25, or 35%	N.A.	N.A.
Modulation	16 to 256 QAM	QPSK	QPSK to 32 APSK	OFDM 2K/8K	OFDM 2K/4K/8K
Channel width	6, 7, or 8 MHz	27 ~ 36 MHz	27 ~ 36 MHz	6, 7, or 8 MHz	5, 6, 7, or 8 MHz

## Conclusioni

In questo lavoro si è descritta l'evoluzione del sistema televisivo, partendo dalla nascita della tv analogica, prima in bianco e nero, poi a colori, fino ad arrivare ai sistemi digitali, ed in particolare il sistema europeo DVB, accettato successivamente a livello mondiale.

Si è quindi introdotta la codifica MPEG e i suoi successivi sviluppi che hanno permesso la diffusione di un segnale audio-video digitale di qualità sempre maggiore. Ci si è poi soffermati sulla descrizione delle specifiche dei due standard digitali più diffusi: satellitare DVB-S e terrestre DVB-T.

In futuro, si cercheranno di sviluppare gli standard DVB per sfruttare al meglio l'alta definizione e per passare da una ricezione non più solo stazionaria ma bensì anche mobile.

## Bibliografia

- [1] Hervé Benoit, “Digital Television, Satellite, Cable, Terrestrial, IPTV, Mobile TV in the DVB Framework”, Terza edizione, *Elsevier* 2008.
- [2] U. Reimers, “DVB, The Family of International Standards for Digital Video Broadcasting”, Seconda edizione, Springer 2005.
- [3] Marcelo S.Alencar, “Digital Television Systems”, *Cambridge university press* 2009.
- [4] U. Reimers, “DVB—The Family of International Standards for Digital Video Broadcasting”, *articolo IEEE presso* [ieeexplore.ieee.org](http://ieeexplore.ieee.org) 2006.









