



Formati Audio

Parte 1

Prof. Filippo Milotta
milotta@dmi.unict.it



Standard MPEG

- **MPEG:** *Motion Picture-Coding Experts Group*

ISO: International Organization for Standards
IEC: International Electrotechnical Commission

- Avviato nel 1988 dalla **ISO / IEC**
 - Standard di:
 - Compressione, Decompressione, Elaborazione, Codifica
 - Per video, audio e contenuto multimediale



MPEG-1 (1992)

Rapida carrellata storica,
poi vedremo in dettaglio MPEG-1

- 300 kbps per audio stereofonico (1-2 canali)
- Tasso di campionamento: 33, 44.1, 48 kHz
- Compressione con bit-rate da 32 a 224 kbps
 - Bit-rate = 192 kbps per una buona qualità
- Tre *Livelli* di compressione:
 - Layer I : b.r.>128 kbps
 - Layer II : b.r.=128 kbps
 - Layer III : b.r.=64 kbps



MPEG-2 (1994)

- B.r.=6 Mbps
- Audio con canali surround
 - 5: Sinistro, centrale, destro, sx-dx surround
 - .1: Subwoofer per le basse frequenze
- **MPEG-3**
 - Pensato per la TV-HD, è stato assorbito nel 2



Altre versioni di MPEG

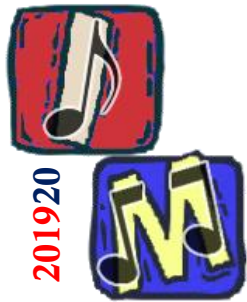
■ MPEG-4 (1999)

- L'audio è composto da diversi oggetti indipendenti

Coincide con l'inizio dell'era di Internet,
ora è importante gestire bene gli archivi condivisi

■ MPEG-7 (2001)

- Standard per la ricerca, il filtraggio e la gestione delle informazioni (e non della codifica, come i precedenti)
- Usa XML
- Insieme a MPEG-4 viene spesso denominato **MPEG-47** per codifica e descrizione



Altre versioni di MPEG

...forse si rischia di condividere troppo?
Si pone l'attenzione sul *Digital Rights Management (DRM)*, riguardo ai diritti d'autore

■ MPEG-21 (2001)

- ❑ Standard per la definizione di un framework per lo sviluppo di applicazioni multimediali
- ❑ Definisce la tecnologia per lo scambio, il consumo e il commercio degli elementi digitali

■ MPEG-D (2007)

- ❑ *Parte 1*: MPEG Surround
- ❑ *Parte 2*: Spatial Audio Object Coding
- ❑ *Parte 3*: Unified Speech and Audio Coding



Proprietà delle codifiche MPEG

- Retrocompatibilità (sempre)
- Libertà nella implementazione:
 - Obbligatorie (***Normativa***)
 - Formato dell'audio compresso
 - Algoritmo di decodifica
 - Liberi (***Informativa***)
 - Algoritmo di compressione
 - Chi usa l'algoritmo di compressione sviluppato originariamente per MPEG dall'istituto Fraunhofer deve pagare una royalty

Può sembrare un controsenso,
ma la parte obbligatoria non si paga,
mentre quella libera sì.

Ciò che si paga è la libera implementazione
della compressione. Uno sviluppatore più bravo
di un altro può richiedere di venire pagato di più.



MPEG-1 | Layer I (MP1)

Dal dominio del tempo al dominio delle frequenze

384 Campioni (Dominio del tempo)

→ Trasformata di Fourier Veloce (FFT)

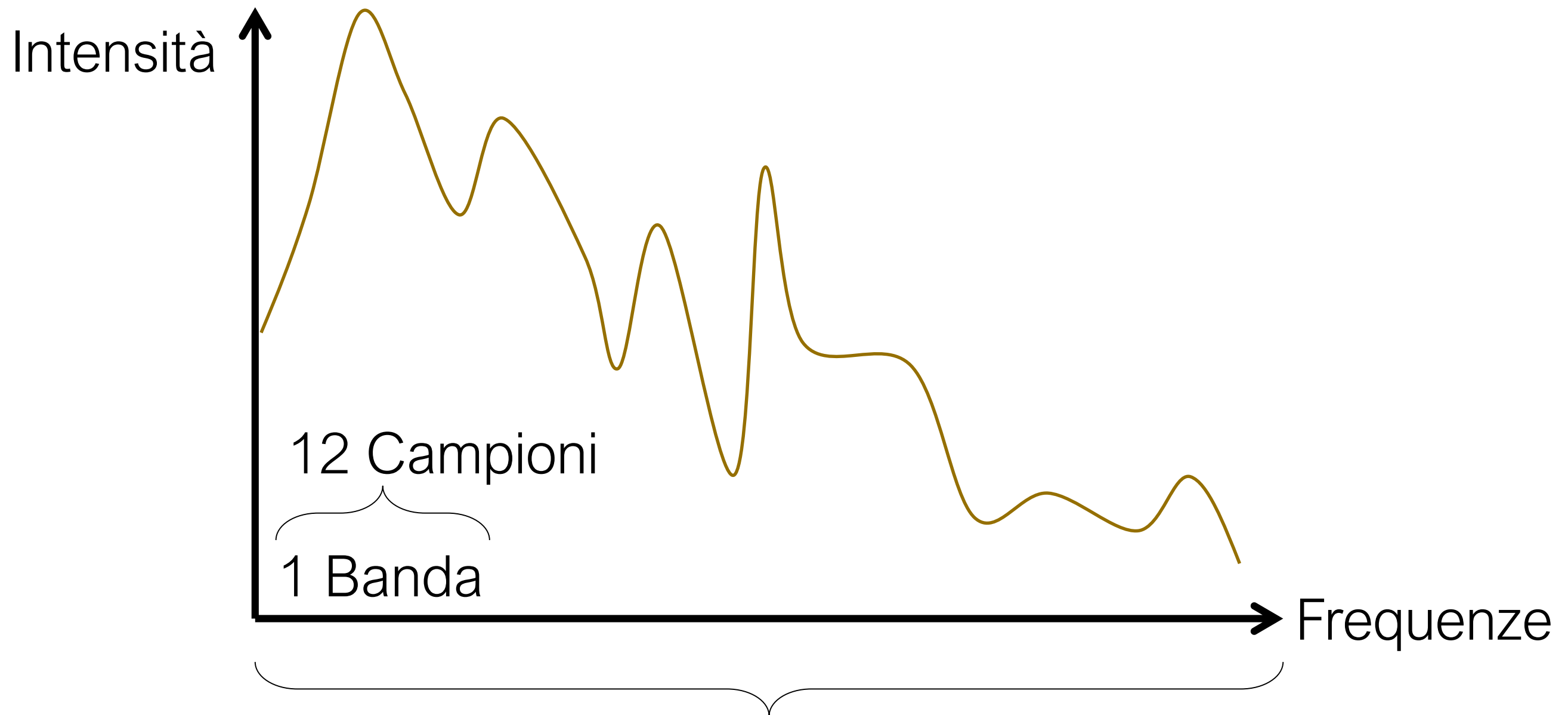
→ 384 Campioni (Dominio delle frequenze)

→ *Filtri Polifase*: 32 bande di frequenza
(12 Campioni / Banda)



MPEG-1 | Layer I (MP1)

Filtri Polifase

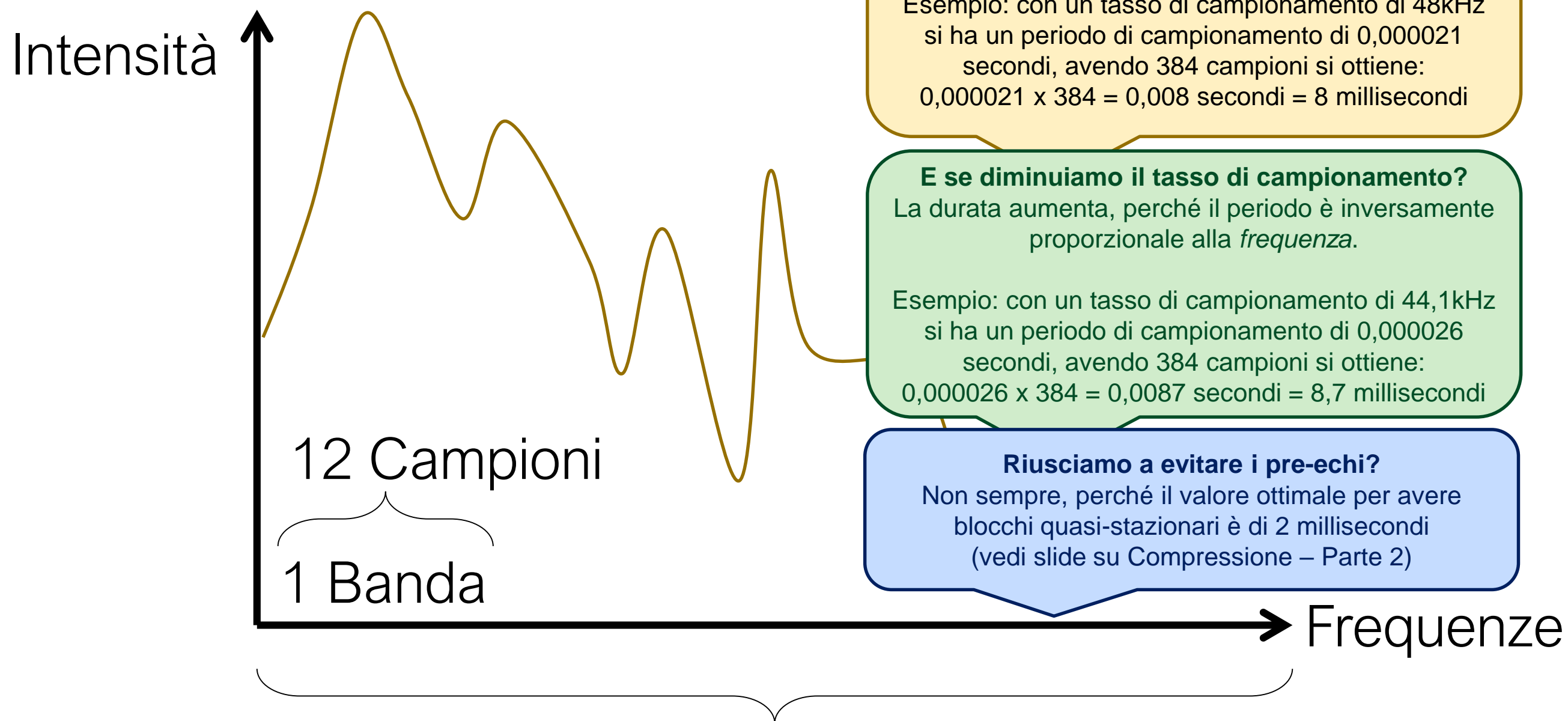


32 Bande → 384 Campioni Tot.



MPEG-1 | Layer I (MP1)

Filtri Polifase



A quanto tempo corrisponde questo spettro?

Di nuovo, dipende dal tasso di campionamento.

Esempio: con un tasso di campionamento di 48kHz si ha un periodo di campionamento di 0,000021 secondi, avendo 384 campioni si ottiene:
 $0,000021 \times 384 = 0,008$ secondi = 8 millisecondi

E se diminuiamo il tasso di campionamento?

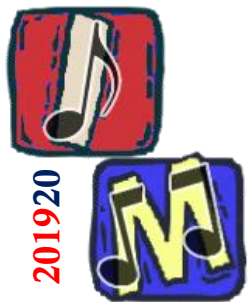
La durata aumenta, perché il periodo è inversamente proporzionale alla *frequenza*.

Esempio: con un tasso di campionamento di 44,1kHz si ha un periodo di campionamento di 0,000026 secondi, avendo 384 campioni si ottiene:
 $0,000026 \times 384 = 0,0087$ secondi = 8,7 millisecondi

Riusciamo a evitare i pre-echi?

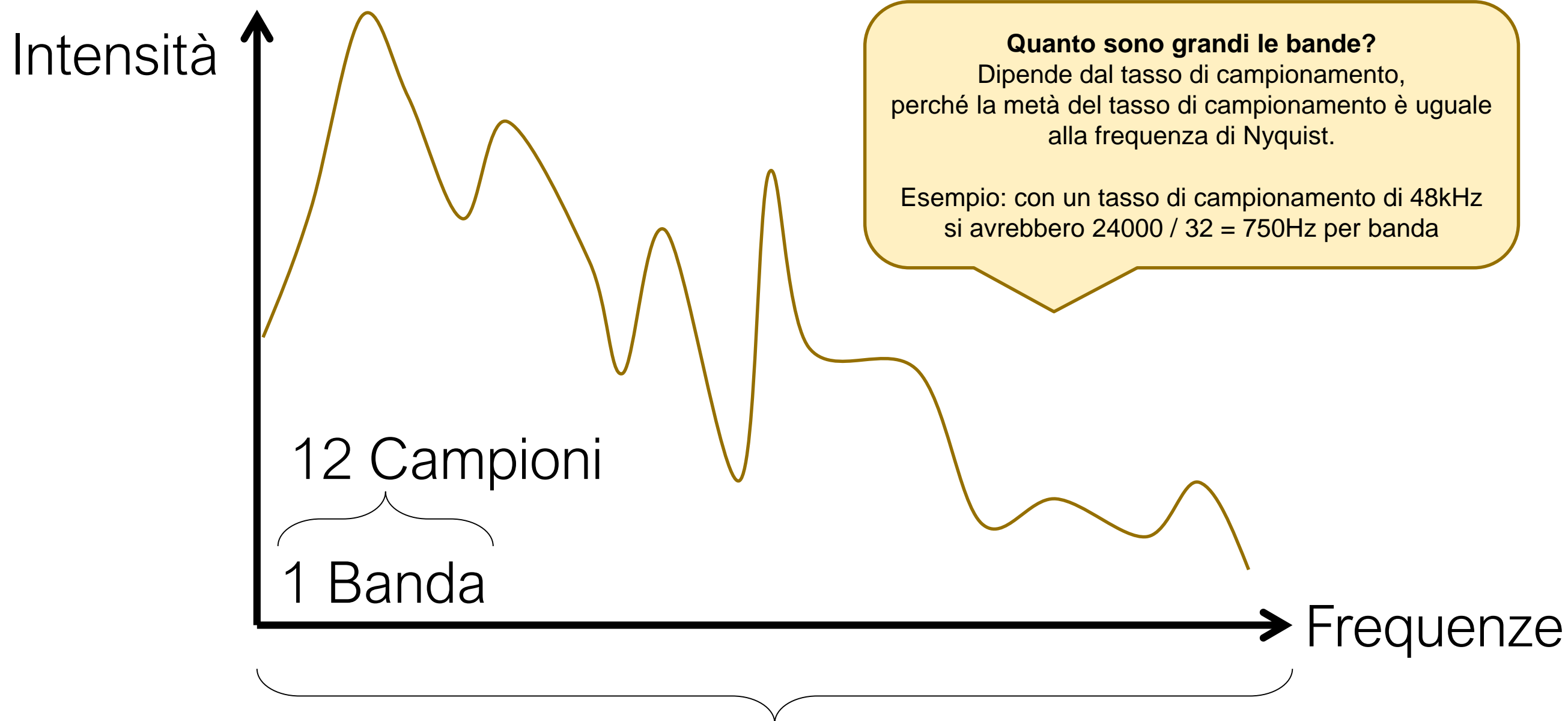
Non sempre, perché il valore ottimale per avere blocchi quasi-stazionari è di 2 millisecondi (vedi slide su Compressione – Parte 2)

32 Bande → 384 Campioni Tot.



MPEG-1 | Layer I (MP1)

Filtri Polifase



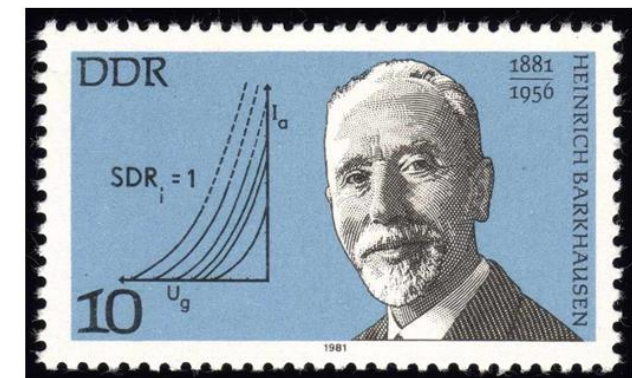
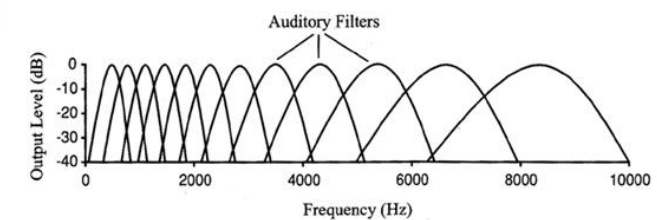
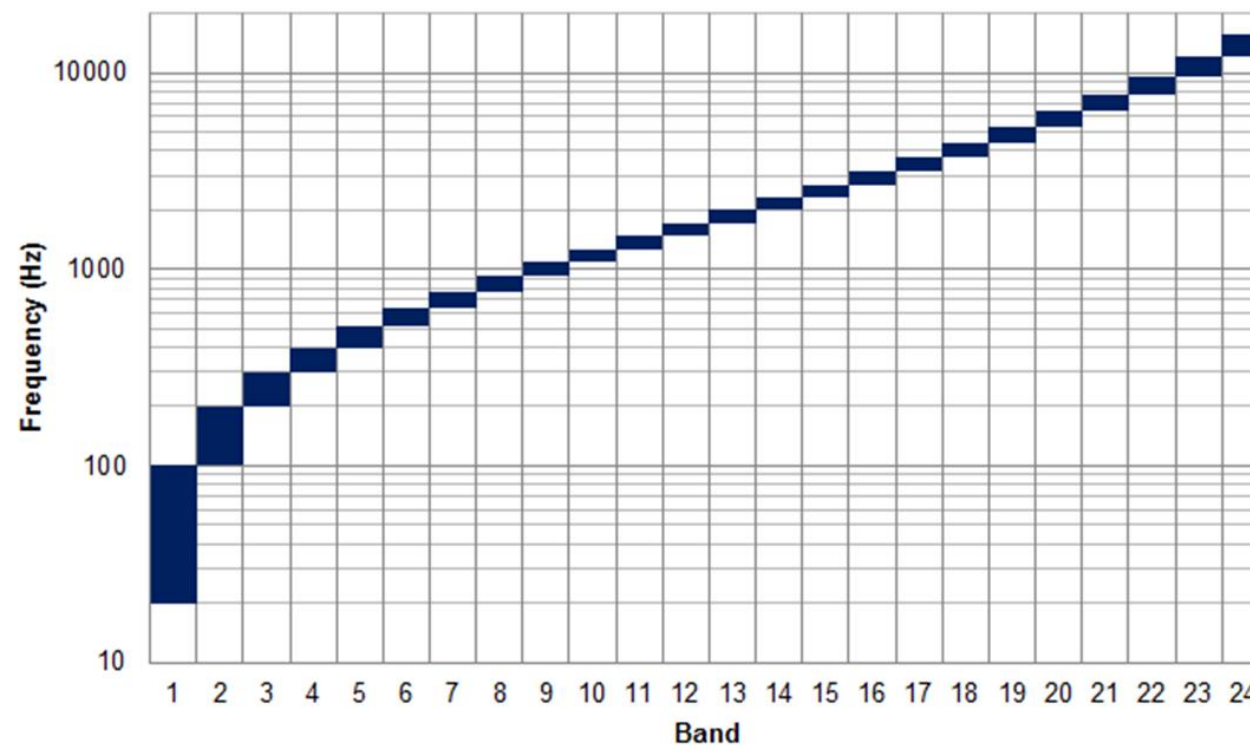
32 Bande → 384 Campioni Tot.

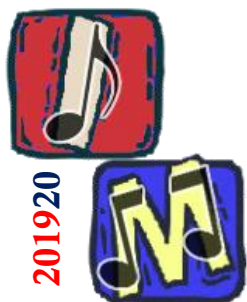
Ripasso – Psicoacustica Parte 3



Mascheramento e Banda Critica Scala di Bark

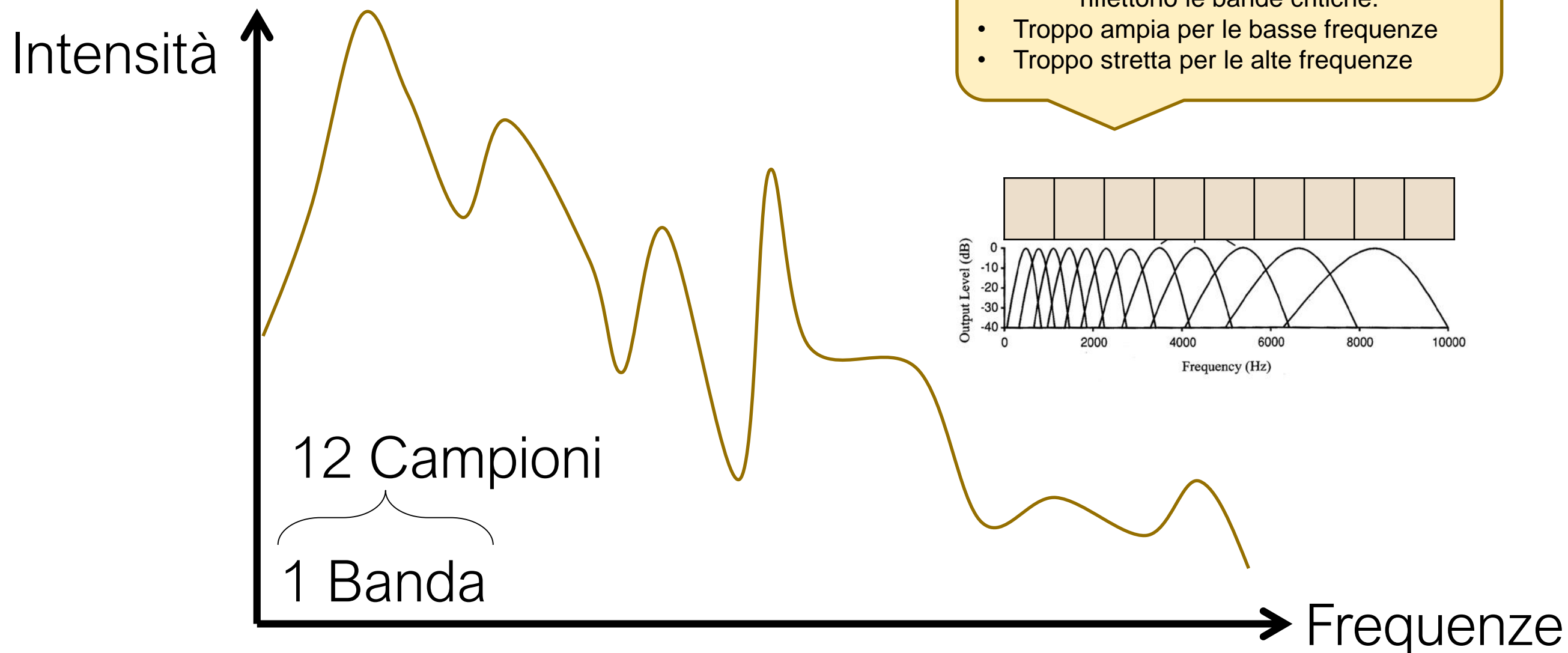
- L'intera gamma delle frequenze udibili viene ripartita in **24 bande critiche** (o filtri auditori, o filtri cocleari)





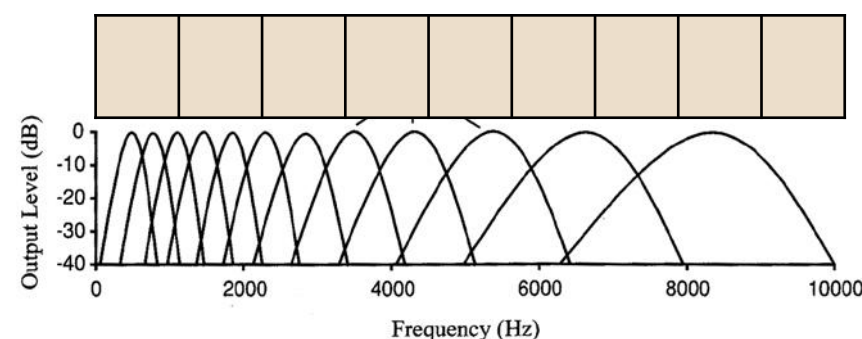
MPEG-1 | Layer I (MP1)

Filtri Polifase

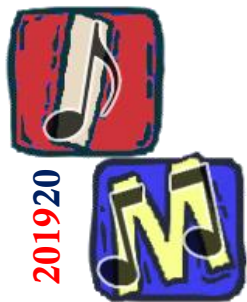


Le 32 bande di larghezza costante non riflettono le bande critiche:

- Troppo ampia per le basse frequenze
- Troppo stretta per le alte frequenze



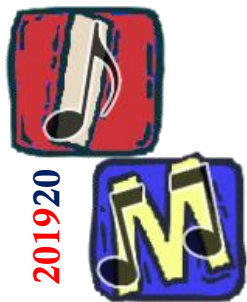
32 Bande → 384 Campioni Tot.



MPEG-1 | Layer I (MP1)

Filtri Polifase

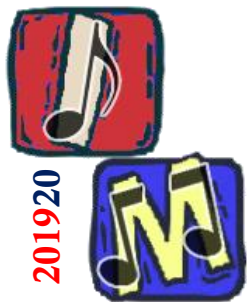




MPEG-1 | Layer I (MP1)

Filtri Polifase





MPEG-1 | Layer I (MP1)

Filtri Polifase



32 Bande → 384 Campioni Tot.



MPEG-1 | Layer I (MP1)

Riassumendo

1. Divisione delle frequenze in 32 bande

- *Le bande sono uguali, a differenza delle bande critiche*
- Esempio: Campionamento a 48kHz
 - Nyquist: 24kHz
 - Banda = $24k / 32 = 750\text{Hz}$

2. 12 campioni per banda (sotto-bande)

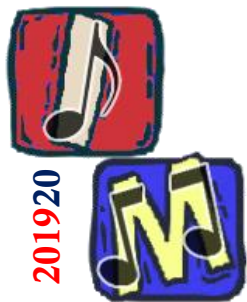
- 6 bit: ogni campione è normalizzato rispetto al picco della intera banda (fattore di scala fissato) – Effetto Compansion [**Ripasso!**]
- 4 bit: 14 classi di ri-quantizzazione uniforme



MPEG-1 | Layer I (MP1)

Header	CRC	Allocazione Bit	Fattori di Scala	Campioni per Sottobanda	Dati Aux
variabile	variabile	32 x 4 bit	32 x 6 bit	32 x 12 (=384)	variabile

- I filtri adiacenti sul banco soffrono di sovrapposizione delle bande in modo significativo
 - *Analisi Psicoacustica*: Il segnale viene anche analizzato mediante una FFT (con finestra a 512 punti)
- In fase di decoding: si applica il banco di filtri in modalità di sintesi con codifica PCM



MPEG-1 | Layer II (MP2)

Analisi Psicoacustica:
viene effettuata ogni 384
campioni

■ Miglioramenti al Layer I:

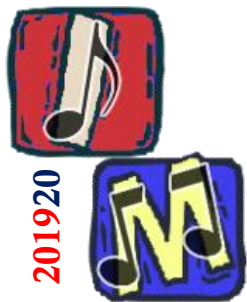
- ❑ La risoluzione della FFT: 512 → 1024 punti
- ❑ Tassi di campionamento: 32/44.1/48 → 16/22.05/24 kHz
- ❑ Blocchi più grandi: 384 → 1152 (=3 blocchi MP1)
- ❑ 3 categorie di classi di riquantizzazione:
 - Frequenza bassa: 15 classi (4 bit)
 - Frequenza media: 7 classi (3 bit)
 - Frequenza alta: 3 classi (2 bit)
- ❑ Si può usare più di un fattore di scala per banda (uno per categoria di classe di riquantizzazione)

Esiste per ogni classe una configurazione 0 per indicare l'assenza di un range di frequenze in una banda



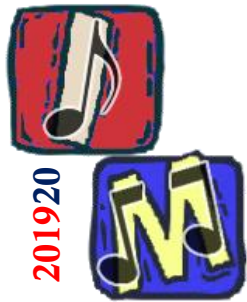
MPEG-1 | Layer III (MP3)

- Miglioramenti al Layer II:
 - Trasformata Discreta del Coseno Modificata (MDCT): permette un partizionamento delle frequenze simile a quello delle bande critiche
 - Quantizzazione non-uniforme
 - Codifiche μ -Law e A-Law
 - Fattori di scala applicabili anche alle sotto-bande
 - Codifica di Huffman
 - A causa della codifica a lunghezza variabile, è necessario un pattern di sincronizzazione temporizzato
 - B.r. variabile fra i blocchi (encoder complessi)



Prestazioni di MP3

Qualità	Compressione	Bandwidth	Canale	Bit-Rate
Telefonica	96:1	2.5 kHz	mono	8 kbps
> Radio AM	24:1	7.5 kHz	mono	32 kbps
Radio FM	24:1	11 kHz	stereo	64 kbps
Quasi CD	16:1	15 kHz	stereo	96 kbps
CD	12:1	>15 kHz	stereo	128 kbps



Formati Audio Avanzati

■ **MPEG–AAC : Advanced Audio Coding**

- ❑ Incluso in MPEG-4
- ❑ Supporta fino a 48 canali audio (contro i 2 di MP3)
- ❑ Un B.R.=128 kbps in AAC
è comparabile a un B.R.=192 kbps in MP3

■ **Dolby AC-3 (Audio Coding)**

- ❑ Compressione di tipo percettivo

La Dolby ha il brevetto sulla parte libera di AAC



Formati Audio Avanzati

■ WMA: Windows Media Audio

Nato per non pagare le royalties su MP3

- ❑ Formato proprietario di Microsoft (blackbox)
- ❑ Prestazioni migliori di MP3
- ❑ Buone prestazioni sulla musica ma non sulla voce

■ FLAC: Free Lossless Audio Codec

- ❑ Compressione lossless (~50%)
- ❑ Confrontato con ZIP (~10-20%), è ottimizzato per la compressione della voce



FFmpeg

- Uno strumento utilissimo:
 - <https://www.ffmpeg.org/>

