



# MPEG



# Audio digitale – MPEG

**Motion Picture Coding Experts Group (MPEG)** è un gruppo di lavoro che si occupa di produrre standard di rappresentazione, codifica, elaborazione ecc... per audio, video e in generale tecnologie multimediali. Ha prodotto molti standard, oggi ampiamente utilizzati.

**MPEG-1** è lo standard più vecchio cronologicamente per la codifica audiovisiva. Definisce tre diversi livelli di compressione per l'audio:

- **Layer I** – Più semplice, il bitrate è superiore a 128 kbps.
- **Layer II** – Complesso, il bitrate è esattamente 128 kbps.
- **Layer III (MP3)** – Molto complesso, il bitrate è di 64 kbps.

Dopo l'MPEG-1 si sono susseguiti nel tempo MPEG-2, MPEG-4, MPEG-7 e molti altri.



# 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), a partire da MPEG-21

- *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:
  - Obbligatori (***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.

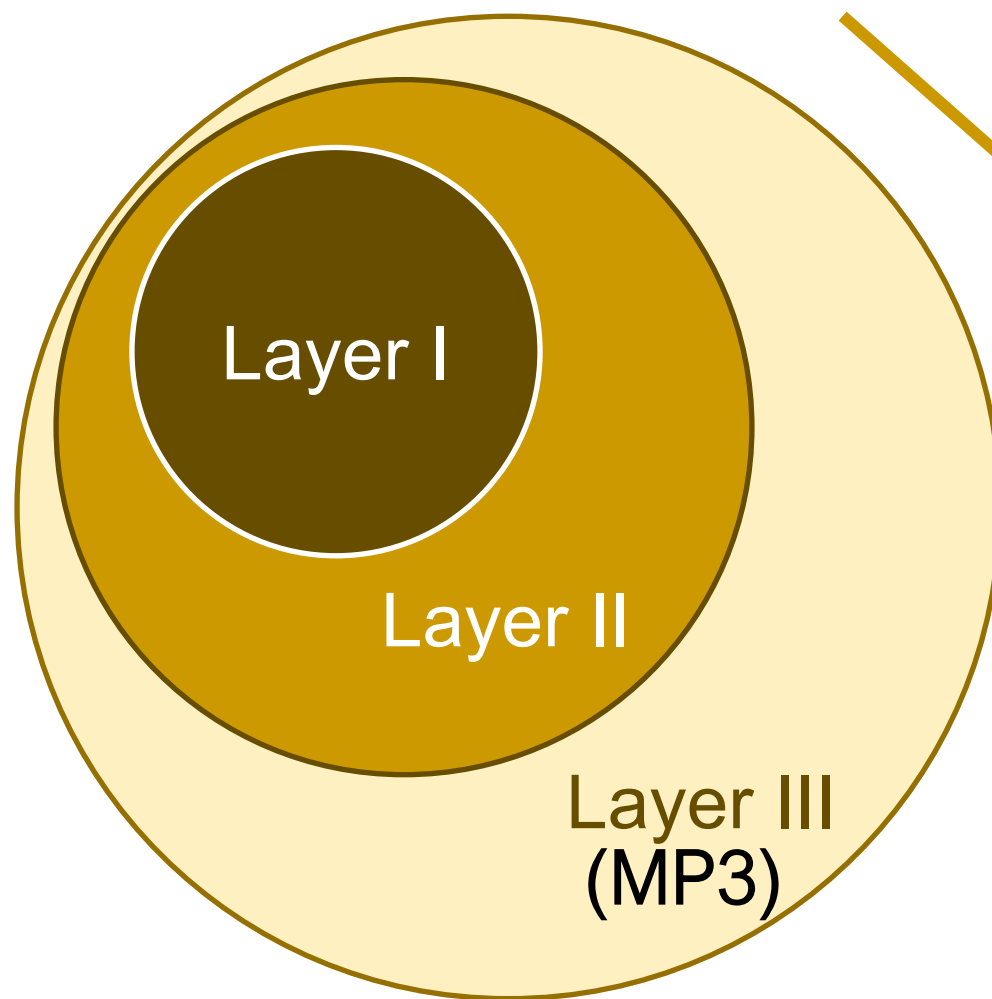
Ciò che si paga è la libera implementazione  
della compressione. Uno sviluppatore più bravo  
di un altro può richiedere di venire pagato di più.





# MPEG-1

## 3 Livelli di compressione



Da Layer I a Layer III:  
migliore compressione  
al prezzo di maggiore  
complessità in encoder e decoder

Vediamo i layer nel dettaglio,  
partendo dall'algoritmo di base  
nel Layer I →



# MPEG-1 | Layer I (MP1)

Dal dominio del tempo al dominio delle frequenze

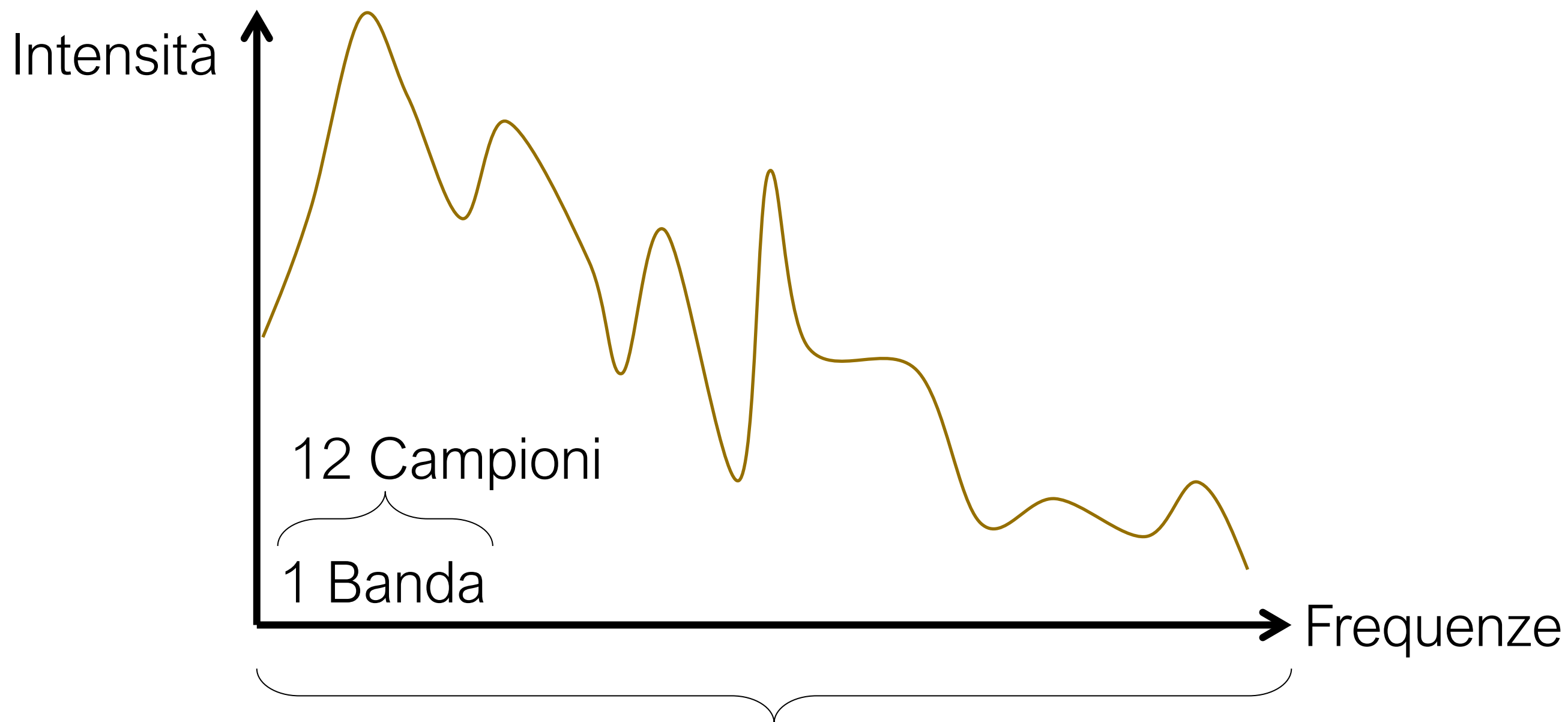
384 Campioni (Dominio del tempo)

→ *Banco di filtri (con DCT)* che producono 32 bande di frequenza  
(12 Campioni / Banda)

→ 384 Campioni (Dominio delle frequenze)



# MPEG-1 | Layer I (MP1)



32 Bande → 384 Campioni Tot.



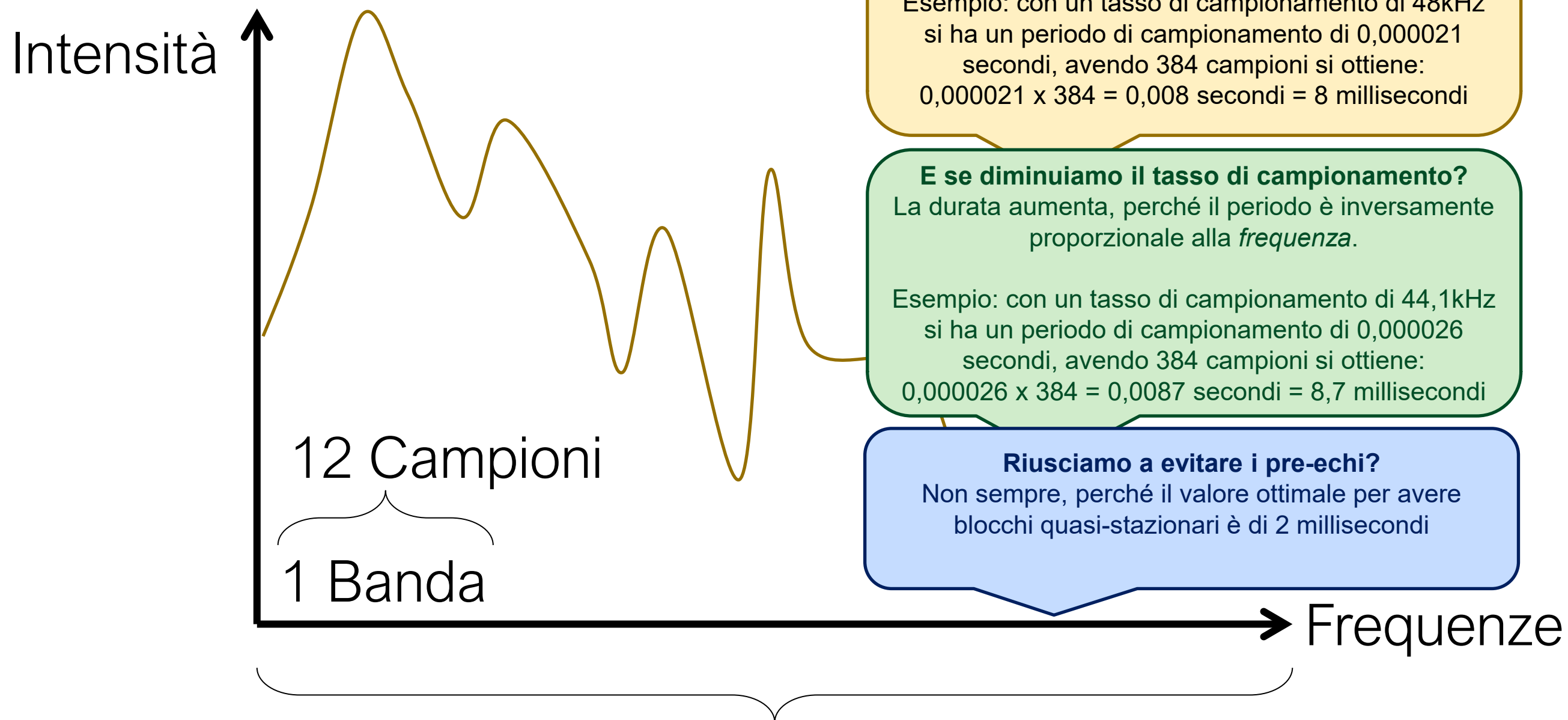
# MPEG-1 | Layer I (MP1)



32 Bande → 384 Campioni Tot.



# MPEG-1 | Layer I (MP1)



**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

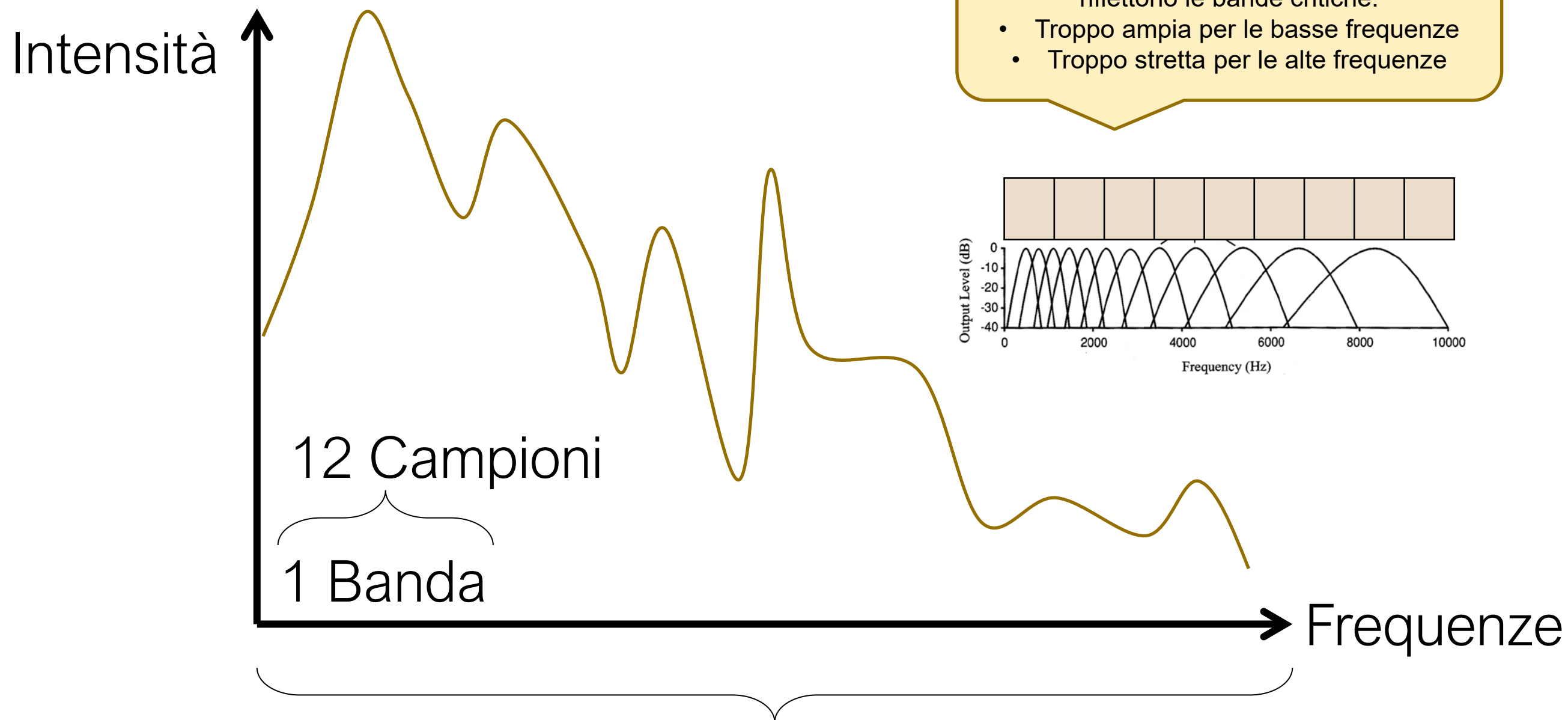
**Riusciamo a evitare i pre-echi?**

Non sempre, perché il valore ottimale per avere blocchi quasi-stazionari è di 2 millisecondi

32 Bande → 384 Campioni Tot.



# MPEG-1 | Layer I (MP1)



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)





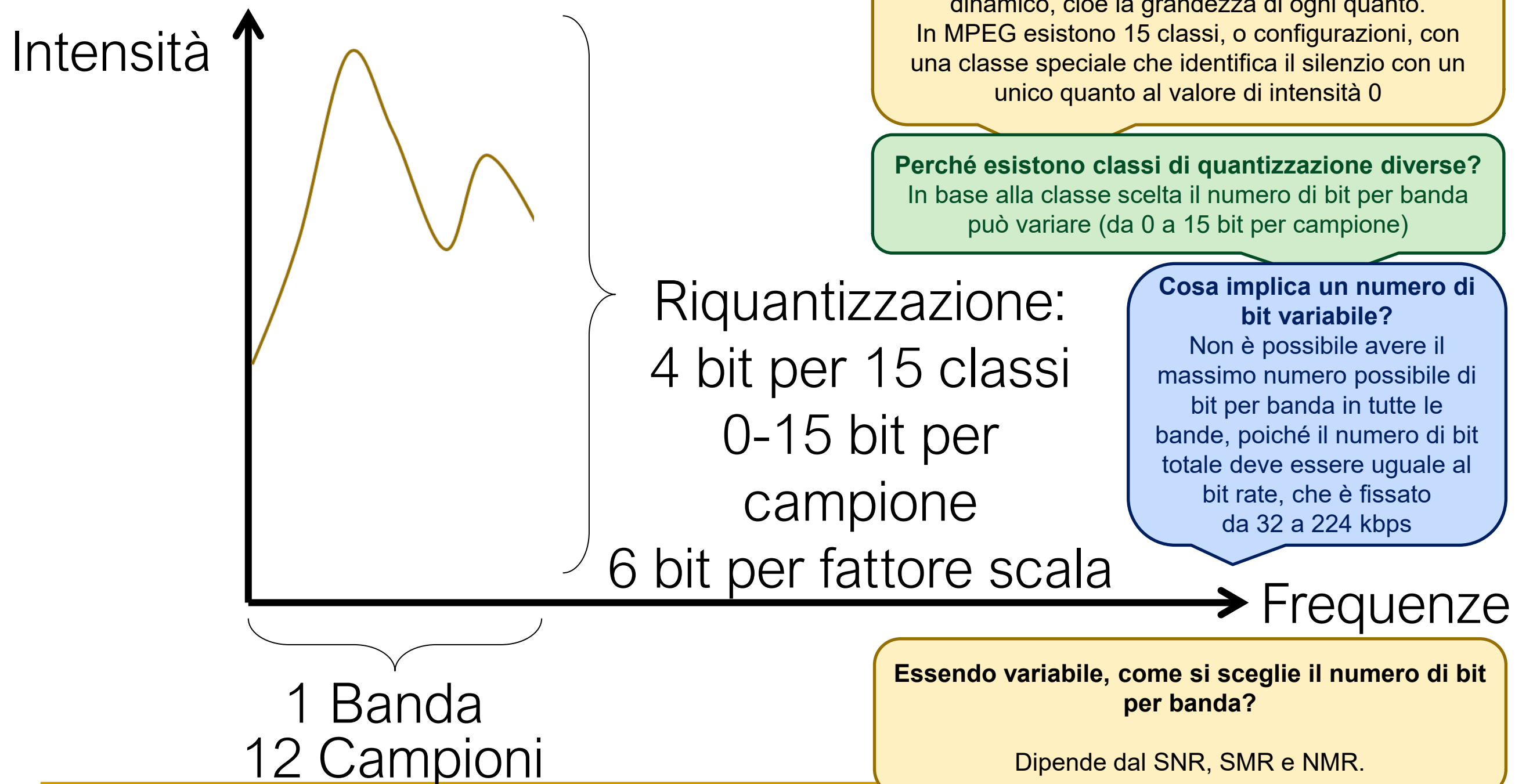
# MPEG-1 | Layer I (MP1)







# MPEG-1 | Layer I (MP1)





# 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: 15 classi (configurazioni) di ri-quantizzazione uniforme
- 0-15 bit per campione



# MPEG-1 | Layer I (MP1)

Header	CRC (ECC)	Allocazione Bit	Fattori di Scala	Campioni per Sottobanda	Dati Aux
variabile	variabile	32 x 4 bit	32 x 6 bit	32 x 12 (=384) da 0 a 15 bit	variabile

Avendo un numero di bit per banda variabile, è necessario utilizzare degli header per guidare il processo di decodifica

- I filtri adiacenti sul banco soffrono di sovrapposizione delle bande in modo significativo
  - *Analisi Psicoacustica*: Il segnale viene 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)
- ❑ Frequenze divise in 3 regioni, riquantizzate con:
  - Se frequenza bassa: 15 classi (4 bit)
  - Se frequenza media: 7 classi (3 bit)
  - Se frequenza alta: 3 classi (2 bit)
- ❑ Si possono risparmiare bit se regioni di frequenza condividono lo stesso ordine di grandezza

Esiste per ogni classe una configurazione 0 per indicare il silenzio per una data regione di frequenza

Regime stazionario: 1 solo ordine di grandezza  
Regime transitorio: potrebbero bastare 2 ordini di grandezza



# MPEG-1 | Layer III (MP3)

- Miglioramenti al Layer II:
  - ❑ Utilizzo di una trasformata del coseno modificata (MDCT). Permette un partizionamento delle bande che assomiglia a quello delle bande critiche.
  - ❑ Quantizzazione non uniforme. Si elevano i valori a  $3/4$ .
  - ❑ Si raggruppano sotto-bande con lo stesso fattore di scala.
  - ❑ Codifica di Huffman - con pattern di sincronizzazione per gestire la codifica a lunghezza variabile.
  - ❑ Bitrate variabile tra blocchi diversi.



# 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