

Codifica e pacchettizzazione della voce

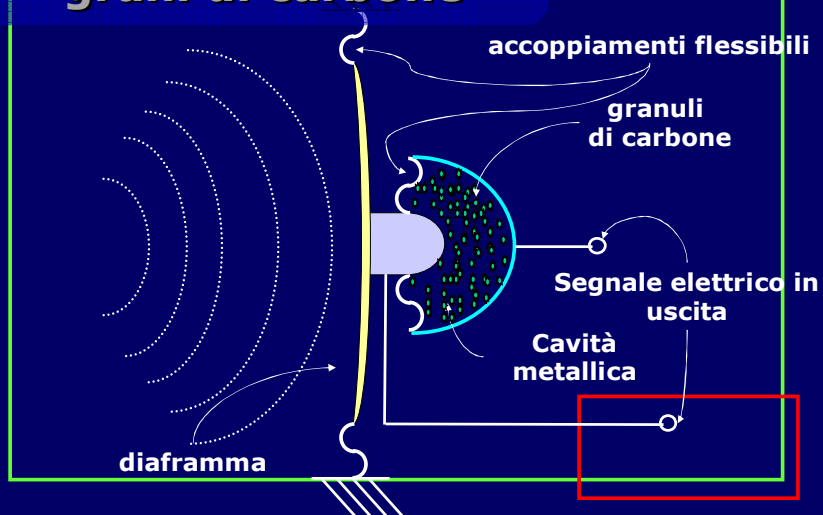
Segnale vocale

- Il segnale vocale è un'onda di pressione in aria
- Un microfono lo trasforma in un segnale elettrico analogico (pressione ↔ volt)

Il trasmettitore a grani di carbone

- **1878: microfono a granuli di carbone**
...con qualche miglioria ancora in uso
- **I granuli di carbone, in una cavità
metallica chiusa da un diaframma,**
forniscono molti percorsi elettrici
possibili
- **Diaframma vibrando fa variare il no.
di possibili percorsi e quindi la
corrente**

Il trasmettitore a grani di carbone



Segnale vocale

- **Processo continuo nel tempo ed in ampiezza**
- **Tramite la codifica si passa da un segnale analogico ad uno numerico**

Valutazione della qualità dei codificatori

Per valutare la qualità di un sistema di codifica della voce si usano tecniche:

- **oggettive (es. SNR, dove N è dato dalla differenza tra i campioni del segnale originale e il segnale codificato)**
- **soggettive**

Valutazione della qualità dei codificatori

**Esempio di tecnica soggettiva:
MOS
(Mean Opinion Score)**

Valutazione della qualità dei codificatori

**Valutazione ottenuta
in modo statistico da un campione
di persone (maschi e femmine)
che ascoltano brevi pezzi di parlato**

Scala MOS

MOS	Qualità	Definizione	Esempio
5	Eccellente	Nessuno sforzo richiesto per la comprensione	Dal vivo CD-audio
4	Buona	Richiede attenzione, ma non uno sforzo apprezzabile	Telefonia fissa

Scala MOS

MOS	Qualità	Definizione	Esempio
3	Discreta	È richiesto un moderato sforzo per la comprensione	GSM
2	Scarsa	Il significato delle frasi è incomprensibile senza un notevole sforzo	

Scala MOS

MOS	Qualità	Definizione	Esempio
1	Insufficiente	La comunicazione è sostanzialmente impossibile	

Codificatori vocali

- Le tecniche di codifica si dividono in:
- **Codificatori della forma d'onda:** usano la sola conoscenza del segnale elettrico istantaneo (es. PCM, DPCM, ADPCM, Mod. Delta)
- **Codificatori di sorgente (vocoder):** sfruttano soprattutto le caratteristiche dell'apparato di fonazione e uditivo (es. LPC)

Codificatori vocali

- In entrambi i casi, il primo stadio del codificatore comprende:
 - Un filtro anti-aliasing
 - Un campionatore
 - Un quantizzatore

Campionamento e quantizzazione

I primi passi per qualsiasi codifica del segnale sono:

Campionamento:

Si rappresenta il segnale continuo tramite un insieme discreto di valori, senza alcun degrado se viene rispettato il teorema di Nyquist

Campionamento e quantizzazione

I primi passi per qualsiasi codifica del segnale sono:

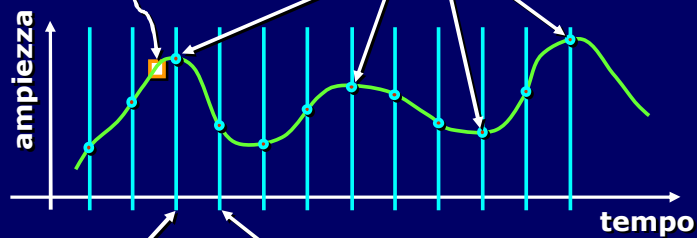
Quantizzazione:

Assegnazione a un insieme discreto di valori dei punti campionati, si ha degrado irreversibile del segnale

Campionamento

segnale analogico

campioni del segnale



intervallo di campionamento T_c

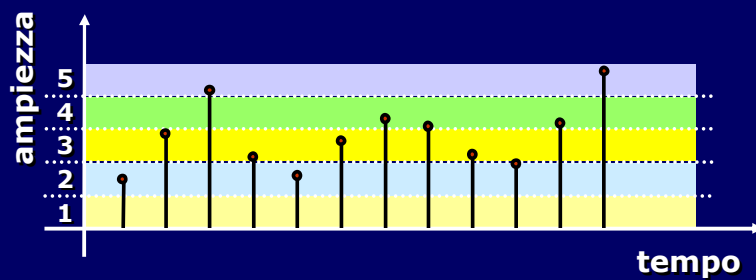
Campionamento

Il teorema di Nyquist definisce il passo di campionamento affinché non si abbia alcun degrado nella ricostruzione del segnale

$$1/T_c > 2B$$

B è il massimo contenuto in frequenza (banda) del segnale da campionare

Quantizzazione



Quantizzazione

- Introduce sempre un degrado noto come *rumore di quantizzazione*
- Il degrado è noto e controllabile e diminuisce con intervalli di quantizzazione più piccoli
- Il degrado può essere ridotto a piacere diminuendo la dimensione dell'intervallo di quantizzazione (ma servono più bit per campione)

PCM uniforme e non

Il PCM (Pulse Code Modulation) è un processo di campionamento e quantizzazione

- La quantizzazione può essere:
 - Uniforme (tutti gli intervalli uguali)
 - Non uniforme (intervalli diversi a seconda dell'ampiezza): compressione prima del quantizzatore uniforme, compensata in ricezione da una espansione (compander)

PCM uniforme e non

PCM uniforme:

CD (~44 kHz, 16 bit)
qualità eccellente (MOS 5)

PCM non uniforme:

Telefonia (8kHz, 8 bit, 64 kbit/s)
qualità buona (MOS 4+),
standard ITU-T G.711

PCM non uniforme

- Oggi si preferisce un campionamento uniforme a 12 bit, seguito da una compressione a 8 bit
- Il campionamento risultante è non uniforme a 8 bit

Modulazione Delta

- È un esempio "particolare"
- Il segnale è campionato a frequenza elevata per ottenere alta correlazione tra i campioni
- La differenza è campionata su un solo bit, che indica se il segnale cresce o decresce

PCM adattativo

- Anche questa tecnica sfrutta la correlazione temporale presente nel segnale vocale
- Modifica nel tempo l'ampiezza degli intervalli di quantizzazione in funzione della dinamica del segnale (adattamento)

PCM adattativo

- Se la variazione nel tempo del segnale è lenta (correlazione) si ottimizza l'ampiezza dell'intervallo in funzione del segnale

PCM adattativo e differenziale (ADPCM)

- Lo standard ITU G.721 usato nei telefoni cordless moderni di tipo DECT è un PCM adattativo e differenziale
- Qualità buona (MOS 4)
- Velocità di trasmissione: 32kbit/s

Codificatori di sorgente

- Il PCM e i suoi derivati codificano il segnale campione per campione
- Le reti di telefonia tradizionale trasmettono campione per campione

Codificatori di sorgente

- Le tecniche di correzione degli errori lavorano bene su blocchi di bit

inoltre...

- In reti a commutazione di pacchetto bisogna accumulare campioni fino a riempire un pacchetto (es. payload 80 byte \Rightarrow 80 campioni PCM \Rightarrow 10ms di voce) prima di iniziare la trasmissione

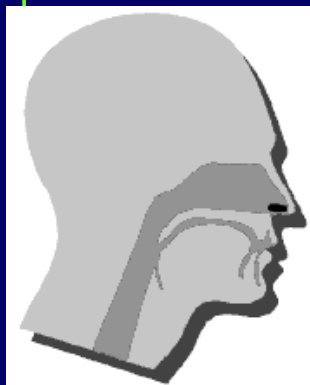
Codificatori di sorgente

- Considerando un segmento vocale (es. un fonema di durata 10-40 ms) è possibile usare algoritmi di codifica e compressione molto efficienti
- Si parte da una codifica PCM uniforme eccellente, si raggruppano da 80 a 320 campioni (10-40 ms) e si lavora sull'insieme (blocco)

Codificatori di sorgente

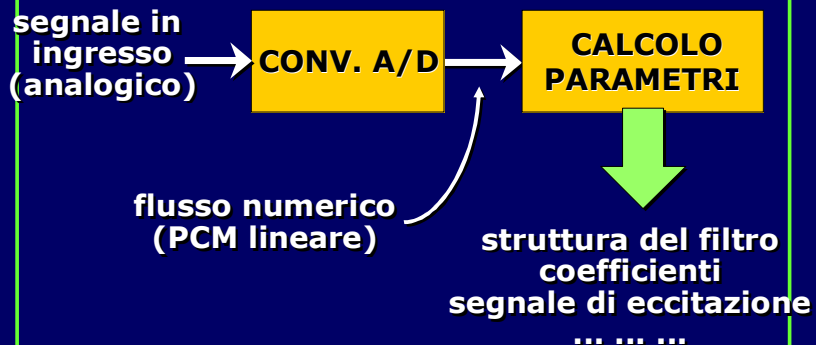
- I codificatori di sorgente si dividono in molte categorie a seconda degli algoritmi usati
- Tutti i codificatori a blocco sono basati su filtri numerici a risposta finita (FIR) di cui il codificatore calcola i parametri
- I parametri vengono trasmessi per la ricostruzione del segnale al ricevitore

Modellizzazione "fisica"



Il filtro usato
rappresenta
i modi vibrazionali
delle corde vocali
e le cavità risonanti
dell'apparato
di fonazione umano

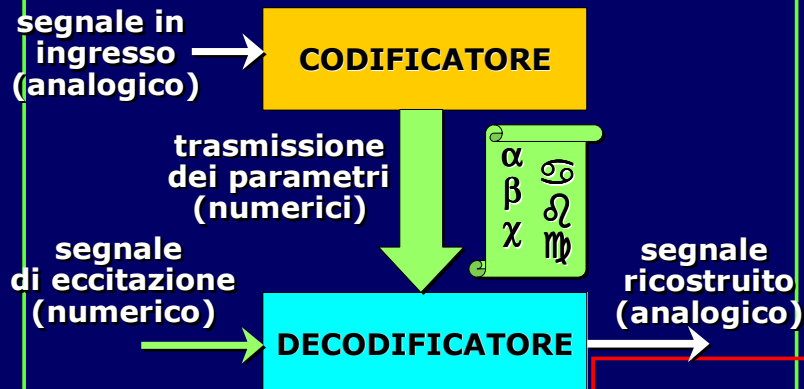
Codificatori di sorgente: codificatore



Codificatori di sorgente: schema al codificatore

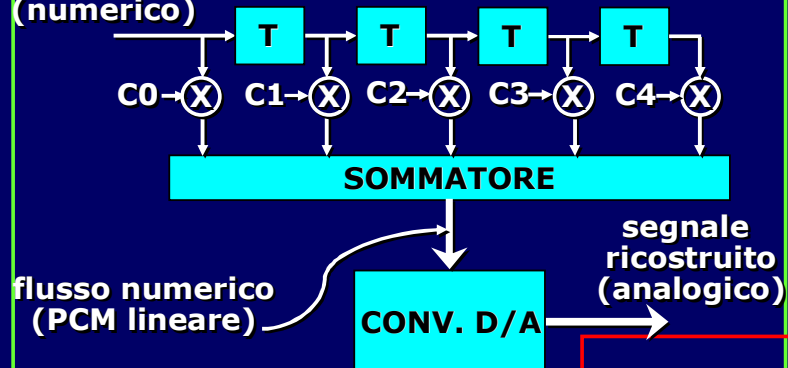


Codificatori di sorgente: schema generale



Codificatori di sorgente: decodificatore

segnale di eccitazione
(numerico)



Soppressione dei silenzi

Con i codificatori a blocco è possibile effettuare la soppressione dei silenzi (VAD – Voice Activity Detection)

Se non vi è attività vocale non vengono generati “pacchetti” di descrizione vocale

Soppressione dei silenzi

Sul canale vengono trasmessi pacchetti di descrizione del silenzio

Risparmio di banda e riduzione dell'interferenza

VAD - criticità

- **I silenzi vanno soppressi senza dare la sensazione di "linea caduta"**
- **Il soppressore deve essere efficiente nell'intervenire, ma soprattutto nel riprendere la codifica, per evitare di "tagliare" l'inizio di nuove lettere/parole**

VAD - criticità

- **I soppressori possono essere:**
 - **Lenti:** intervengono solo durante lunghe pause di parlato
 - **Veloci:** cercano di intervenire anche nei brevissimi silenzi tra parole della stessa frase

Codifica LPC-LTP

- **La codifica Linear Prediction Coding - Long Term Prediction si basa sulla modellizzazione fisica del tratto vocale mediante due filtri, uno a memoria breve (filtro LPC) e uno a memoria lunga (LTP)**

Codifica LPC-LTP con RPE

- Il segnale viene rigenerato eccitando il filtro a memoria breve con rumore gaussiano bianco e il filtro a memoria lunga con un treno di impulsi regolari (RPE - Regular Pulse Excitation)

Codifica LPC-LTP con RPE

La codifica è divisa in:

Predizione di breve periodo (LPC)

Serve a codificare "bene"
i fonemi con consonanti

Codifica LPC-LTP con RPE

La codifica è divisa in:

Predizione di lungo periodo (LTP)

Serve a codificare "bene"
i suoni vocalizzati

Codifica LPC-LTP con RPE

La codifica è divisa in:

Calcolo del miglior segnale
di eccitazione
(RPE)

Codifica CELP

Code Excited Linear Prediction

È un codificatore LPC in cui l'eccitazione per ricostruire il segnale non è rumore bianco ma una sequenza di un "codebook" (catalogo) che minimizza l'errore rispetto al segnale originale

Codifica CELP

Il codificatore è molto complesso perché deve scegliere tra i possibili codici in modo esaustivo

Codificatori GSM

GSM tradizionale:

- Codificatore LPC-LTP con RPE
- 13 kbit/s

GSM Enhanced Full Rate (EFR):

- Telefonini "dual-band" e/o posteriori al 1997
- CELP a 12.2 kbit/s

Codificatori GSM e UMTS

UMTS:

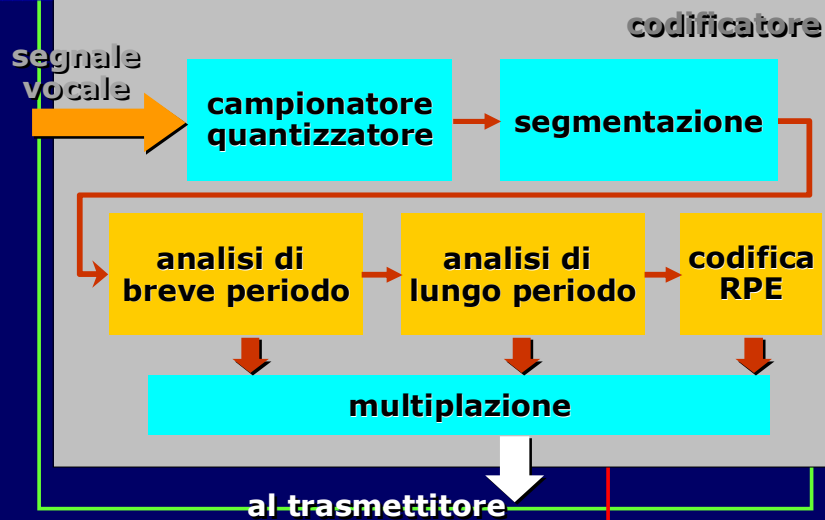
- Codec adattativo a rate variabile (AMR – Adaptive Multi Rate)
- Può essere usato anche con GSM

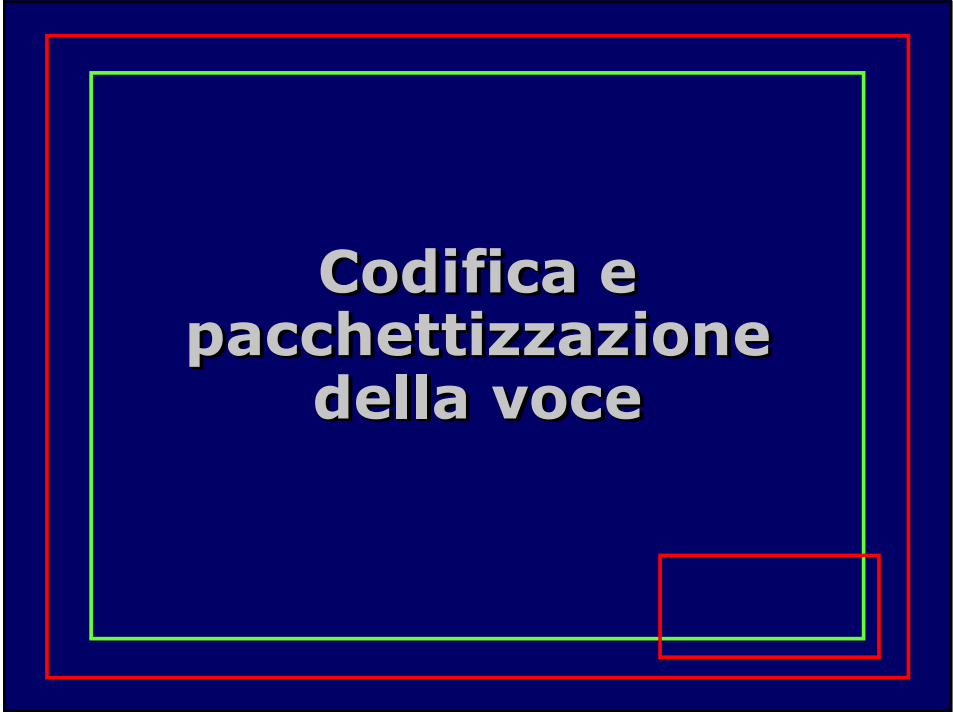
Codificatore GSM "base"

- Codificatore LPC-LTP con RPE (standard ETSI)
- Blocchi da 20 ms che producono 260 bit raggruppati in 3 livelli di importanza
 - 50 ricevono massima protezione
 - 132 ricevono una protezione media
 - 78 non sono protetti per nulla

Totale 13 kbit/s

Codificatore GSM "base"





Codifica e pacchettizzazione della voce