

Reti di Calcolatori

Homework 1

Progettazione e realizzazione algoritmo di Voice Activity Detection

Andrea Zanella

March 25, 2022

1 Specifiche

Si consideri un segnale digitale audio mono, in formato PCM, che si assume essere generato in tempo reale. Il trasmettitore esegue una pacchettizzazione del segnale, con pacchetti da 160 campioni audio (corrispondenti a un intervallo audio di 20 ms).

Si vuole progettare un algoritmo di riconoscimento del parlato (Voice Activity Detection - VAD) che determini quali pacchetti hanno contenuto vocale e, quindi, vanno trasmessi, e quali invece possono essere soppressi in quanto privi di contenuto vocale significativo. Da notare che i pacchetti privi di contenuto vocale possono comunque contenere rumore di fondo. Per ogni pacchetto audio, l'algoritmo VAD dovrà generare in uscita il valore 1 nel caso il pacchetto debba essere trasmesso, e 0 altrimenti. La decisione dell'algoritmo VAD relativamente al pacchetto n -esimo può basarsi sui campioni di segnale di tutti i pacchetti precedenti $k \leq n$ e, potenzialmente, anche di campioni successivi, purché il ritardo complessivo di codifica rimanga al di sotto di 50 ms.

L'obiettivo è trasmettere il minor numero possibile di pacchetti che consenta la ricostruzione del segnale vocale senza artefatti (clipping).

2 Simulazione

Ai fini dell'esercitazione, il segnale audio sarà fornito sotto forma di file binari `inputaudioN.data`, dove N è l'indice del test (da 1 a 5). Ogni campione è memorizzato come un intero a 8 bit, con segno. Si può assumere che l'intervallo di quantizzazione sia adattato all'ampiezza del segnale, sicché i valori quantizzati spaziano tendenzialmente sull'intero intervallo.

Si può implementare l'algoritmo di VAD in qualsiasi linguaggio, ma si raccomanda MATLAB, python o C++. Lo script deve eseguire le seguenti operazioni.

- Leggere i campioni audio da un qualsiasi file di tipo `inputaudioN.data`.
- Analizzare il contenuto audio corrispondente a intervalli di 20 ms di segnale in forma sequenziale
- Determinare la presenza o meno di voce per ognuno di questi intervalli e decidere se sopprimere o meno il pacchetto voce.
- Creare un file `outputVADN.data` che rappresenta la versione "ricevuta" del file vocale, dove i pacchetti soppressi dall'algoritmo di VAD sono sostituiti da sequenze di zeri di ugual lunghezza.
- Nota bene: non è necessario simulare lo scambio effettivo di dati tra client e server.

3 Consegna

Si deve consegnare un file ZIP contenente quanto segue:

- Codice sorgente dell'algoritmo, corredato con file `ReadMe.txt` con le istruzioni da lanciare per compilare il codice e applicare l'algoritmo a un file di input specifico.
- I file testo `outputVAD.txt` ottenuti applicando l'algoritmo ai file `inputaudioN.data` forniti in moodle.
- Un report di **massimo 1 pagina** che descrive l'algoritmo proposto (in formato flow chart).