

Reti di Calcolatori

PROGETTO DI FINE CORSO

Progettazione e realizzazione algoritmo di Voice Activity Detection

Andrea Zanella

June 7, 2021

Regole

- Il progetto è opzionale e riservato agli studenti che hanno svolto e consegnato entrambi gli homework proposti durante il corso, conseguendo una votazione sufficiente.
- Il progetto sostituisce la prima prova d'esame 'standard' (quiz) per la sola parte *multimedia*.
- Il progetto deve essere consegnato entro la data prestabilita senza alcuna possibilità di proroga.
- Chi opta per la prova a progetto ma non fosse soddisfatto del voto, può sostenere le prove d'esame standard (quiz) *successive* al primo appello.

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 minore 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 richiede di implementare l'algoritmo di VAD in MATLAB 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 e determinare la presenza o meno di voce per ognuno di questi intervalli.
- Salvare l'output binario del VAD su un file `outputVADN.txt`. Nota bene: per semplificare la correzione, che avverrà in modo automatico, il file di output deve essere **in formato testo**, dove i simboli 1 o 0 sono salvati in formato ASCII.
- 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 del simulatore, 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.
- Report di massimo 2 pagine che descrive l'algoritmo proposto (in formato flow chart) e riporta un esempio del risultato ottenuto tramite simulazione, usando i file di test forniti (il formato con cui riportare questi risultati è a libera scelta dello studente).