RESTAURO AUDIO DEL SEGNALE VOCALE MEDIANTE PROIEZIONE LOCALE

**Sergio Canazza, **Gian Antonio Mian*

Università di Udine, Dip. Scienze Storiche e Documentarie
Università di Padova, Dip. Ingegneria dell'Informazione

RIASSUNTO

La voce umana, se si prendono in considerazione intere parole, consiste in un fenomeno complesso e altamente non stazionario. Tuttavia è possibile rintracciare numerosi frammenti del segnale vocale, in particolare in corrispondenza dei fonemi con presenza di vocali, in cui risulta evidente una periodicità locale.

Utilizzando gli strumenti tipici dello studio degli "embedding spaces", in particolare mediante il ricorso a vettori di ritardo, si può ricostruire un attrattore locale che evidenzia come esistano strutture riconoscibili anche in presenza di segnale rumoroso. Si intendono sfruttare queste caratteristiche di periodicità locale per sviluppare un algoritmo che porti alla riduzione del rumore sovrapposto al segnale vocale.

Supposto che si sia individuata una zona periodica nel segnale, la prima parte dell'algoritmo proposto consiste nella stima della lunghezza del periodo e nel successivo frazionamento del segnale in blocchi corrispondenti al periodo stimato.

Una volta che si dispone di diverse realizzazioni tra loro simili della forma d'onda periodica, generalmente una ventina all'interno dello stesso fonema, viene calcolato per ognuna di esse lo scarto rispetto alla media e la matrice di covarianza dello scarto.

Tale matrice viene studiata con una scomposizione ai valori singolari, ciascuna realizzazione viene proiettata sui vettori ortonormali corrispondenti ai valori singolari più piccoli, ottenendo così quella che si suppone essere la componente rumorosa del segnale. Con una semplice sottrazione è così possibile recuperare la parte originale del segnale vocale.

Il problema principale consiste nella località della periodicità nella voce umana, caratteristica che obbliga ad una scelta dei parametri (dimensione della finestra e passo di avanzamento per la ricerca della periodicità, raggio di vicinanza entro cui due forme d'onda vengono ritenute simili, numero di valori singolari da considerare per la proiezione, ...) strettamente dipendente dalle caratteristiche temporali del segnale e del rumore.

Per evitare un'ispezione visiva del segnale si fa ricorso a funzioni di costo relative a misure del periodo e raggio di vicinanza, che rendono più semplice la taratura dei parametri.

I risultati ottenuti operando con una realizzazione in Matlab dell'algoritmo evidenziano che il rumore bianco a larga banda sovrapposto al segnale vocale viene eliminato particolarmente bene nelle zone del segnale in cui sono presenti vocali (partendo da segnale reale rumoroso con rapporto segnale-disturbo (SNR) di 5 dB, si arriva ad un restauro che presenta SNR di 12 dB circa), mentre nelle zone del segnale in cui sono presenti consonanti non si nota alcun miglioramento.

Sessione 16: Restauro del segnale VOCE

Venerdì!