

Il riconoscimento della voce (III)

Nell'ambito dell'applicazione del riconoscimento della voce, possiamo trovare molte tecniche di classificazione dei campioni, quindi in queste pagine vi spiegheremo in modo dettagliato quali sono quelli che offrono i migliori risultati e sono più promettenti per il futuro.

Comparazione di campioni utilizzando tecniche di Programmazione Dinamica (DTW)

Questo metodo consiste nel comparare ognuno dei segmenti in cui si scompone un segnale strutturato con campioni o strutture memorizzate in un database (dizionario).

Ogni campione non è altro che un insieme di caratteristiche acustiche ordinate linearmente nel tempo. Questa tecnica, che si utilizza sia per risolvere problemi di riconoscimento di parlata continua che di parole isolate, possiede anche una certa indipendenza dalla persona che parla ed è conosciuta come DTW (Dynamic Time Warping).

Modelli occulti di Markov (HMM)

Il modello stocastico del segnale della parola, anche conosciuto come creazione di modelli occulti di Markov, fornisce i

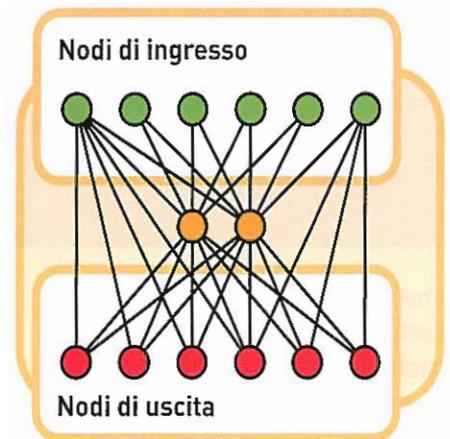
maggiori risultati avuti sinora sia per il riconoscimento della parola isolata che di quella continua, indipendentemente dall'interlocutore.

Utilizza un algoritmo di allineamento non lineare, conosciuto come algoritmo di Viterbi, capace di ordinare direttamente una sequenza di segmenti di ingresso con un insieme di campioni stocastici che rappresentano le parole del dizionario.

Reti neurali (NN)

Le reti neurali sono strutture di elaborazione parallela delle informazioni. Sono formate da nodi semplici collegati fra loro e raggruppati in differenti livelli.

Il livello è noto come livello di ingresso, quello finale come livello di uscita. Ogni nodo rappresenta un fonema. Man mano



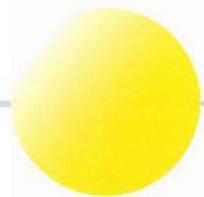
Reti neurali a tre livelli.

che i fonemi vengono classificati, si avanza per i differenti livelli che compongono la rete. Per sua natura l'organizzazione non è lineare e possiede una grande capacità di classificazione.

Questo metodo è diventato uno degli strumenti più attrattivi per la soluzione del problema di riconoscimento della parola, dato



Le caratteristiche della voce possono variare da una persona all'altra.



I rumori dell'ambiente possono rendere difficile il riconoscimento.

della parola, dobbiamo ricordare che queste applicazioni non sono esenti dal commettere errori, quindi non possiamo utilizzare queste tecnologie in applicazioni in cui il buon funzionamento può essere critico o in sistemi in tempo reale.

Tuttavia, è possibile utilizzare il riconoscimento della voce per comunità di portatori di handicap, o utilizzandolo come sistema di ingresso dei dati.

Inoltre possono anche essere controllate mediante comandi vocali applicazioni come navigatori di Internet, giochi, accessi a database, ecc. Altri campi in cui è utilizzata la tecnologia di riconoscimento della voce sono i servizi di telefonia senza operatore, in cui una centralina interagisce con l'utente, e nella telefonia mobile, dove i terminali compongono il numero del contatto desiderato quando si pronuncia il suo nome. Però questi sono solo pochi esempi del riconoscimento automatico della voce nella vita quotidiana. Attualmente sono in corso diversi studi su questa tecnologia per il suo utilizzo in diversi settori, come ad esempio la domotica e soprattutto la medicina.

che ha ottenuto risultati comparabili a quelli ottenuti con i metodi classici, che presentano diversi inconvenienti come per esempio l'eccessivo tempo di elaborazione di cui hanno bisogno.

Il problema del riconoscimento automatico della parola

Esistono molti fattori che influenzano il processo di riconoscimento della voce, e quindi il suo rendimento. Possiamo differenziare due tipi di fattori, esterni e interni. Per quanto riguarda questi ultimi il principale è la differenza nei suoni a causa fondamentalmente dei diversi accenti o modi di parlare di ogni persona. Costituiscono un problema anche le diverse velocità del parlato, la coartazione, i rumori prodotti dall'interlocutore (respirazione), il contesto della conversazione, il suo stato d'animo, ecc. Fra i fattori esterni ricordiamo la qualità della conversione e della trasmissione del segnale elettrico, che può generare problemi per le differenze fra le caratteristiche dei cavi nei

microfoni, delle linee telefoniche, ecc., e la quantità di rumore captato con il segnale della voce, la vicinanza di altre sorgenti sonore. A questi fattori ne dobbiamo aggiungere anche altri come i diversi dialetti che esistono per una lingua, l'utilizzo di parole inesistenti nel dizionario dell'applicazione, la costruzione di frasi non permesse dalla grammatica del linguaggio, o l'utilizzo di abbreviazioni.

Applicazioni del riconoscimento automatico della parola

Quando utilizziamo applicazioni del riconoscimento automatico

Esistono videogiochi controllati mediante ordini vocali.

