

Tecniche di riconoscimento vocale un ulteriore passo avanti: l'AUDIO MINING

di Gian Paolo Trivulzio

La tecnologia di riconoscimento della voce ha fatto negli ultimi quattro anni notevoli passi in avanti: le aumentate logiche dei programmi informatici, accoppiate alla velocità di elaborazione ed alle notevoli dimensioni di memoria consentono di ottenere risultati interessanti.

La tecnologia è stata adattata a molteplici applicazioni, basti pensare alla possibilità di pronunciare un nome ed ottenere la composizione automatica del numero telefonico corrispondente incluso nella rubrica del proprio telefonino, oppure di richiedere informazioni via telefono (il servizio 12 di Telecom Italia) dando al computer l'indicazione della località e del nome dell'abbonato desiderato. Al termine il computer ci pronuncia il numero telefonico, tale pronuncia è sempre derivata per via software dal testo dell'elenco (tecnologia text-to-speech, dal testo al parlato in contrapposto alla precedente speech-to-text ossia dal parlato al testo).

Dall'insieme di tutte queste possibilità che tendono a svilupparsi nel filone del riconoscimento indipendente dal parlatore (al telefono può parlare chiunque con varie inflessioni, purché in modo chiaro) si aprono nuove orizzonti per risolvere problematiche pratiche finora difficilmente gestibili.

E' il caso della ricerca, all'interno di un testo sonoro, di una determinata frase o parola. Questa ricerca è noiosissima e lunghissima se si utilizzano i tradizionali nastri o cassette, occorre infatti ascoltare il suono, prestare bene attenzione e segnarsi poi i punti che ci interessano se vogliamo successivamente ritrovare quanto abbiamo ascoltato. Se il contenuto di una registrazione è a noi ignoto, questo significa molte volte ascoltare per lungo tempo prima di trovare ciò che ci interessa, un po' più facile è la ricerca se abbiamo già ascoltato il testo recentemente (passando del tempo la memoria si affievolisce, specie se l'informazione non è per noi particolarmente importante).

Nelle attività professionali (studi radiofonici e televisivi, registrazione di eventi, discorsi e simili) la memoria serve, ma non può ricordarsi tutto.

Ecco allora che la tecnologia di riconoscimento della voce, accoppiata alla registrazione digitale (vedi una mia precedente nota al riguardo) hanno consentito di escogitare un software avanzato che prende il nome di audio-mining, letteralmente miniera audio (o di suoni, se vogliamo). In pratica il contenuto del file sonoro contenente la registrazione che ci interessa, viene esaminato e 'decriptato' nelle sue componenti sonore. In questa attività viene anche mantenuta l'informazione del cosiddetto 'time code', ossia del momento esatto (minuto e millesimi di minuto) in cui il suono è stato pronunciato ed il tutto viene memorizzato in appositi file. Terminata questa operazione è disponibile una

speciale banca dati che accoppia testi, suoni e 'time code', per cui è possibile, utilizzando la comune tastiera, ricercare l'informazione desiderata (ad esempio quando l'oratore ha pronunciato la parola 'salame') ed il programma effettuerà tale ricerca come oggi noi ricerchiamo all'interno di un documento in formato elettronico. Ci indicherà i vari punti dove è contenuta questa informazione e sarà quindi possibile riascoltare in pochi istanti quello che ci interessa., semplicemente cliccando sull'informazione che ci viene restituita.

Questo tipo di ricerca può essere utile, come sopra detto, per gli studi professionali di audioregistrazione, ma può avere ulteriori sviluppi applicativi ad esempio nel settore della giustizia potrebbe essere possibile rintracciare informazioni importanti nelle tradizionali intercettazioni telefoniche, evitando così la completa trascrizione di decine o centinaia di telefonate (come oggi avviene) alla ricerca del classico ago nel pagliaio.

Il procedimento per raggiungere il risultato è operativamente facile, ma impegna notevoli risorse di ogni tipo, oltre a quelle già impegnate per lo studio e l'elaborazione di questo software. Non è quindi oggi possibile acquistarlo in macelleria (data l'emergenza mucca pazza anche i macellai si stanno riconvertendo all'informatica), ma per ora è possibile usufruirne quale service a costi abbastanza sostenuti. Ma il passato ci ha insegnato che nell'informatica il tempo passa velocemente e normalmente a vantaggio dell'utente.