

Capisci quello che dico?

Il riconoscimento vocale ha prevalentemente poche e precise funzionalità: digitalizzazione della dettatura, esecuzione di comandi vocali, autenticazione e identificazione di soggetti umani attraverso il timbro vocale

DANIELE CATTANEO

Siete già in possesso di un PC con sistema operativo Windows Vista? Bene, avete allora la possibilità di pronunciare la frase “mostra desktop” e ‘miracolosamente’ tutte le finestre aperte dovrebbero ridursi a icona e mostrare il desktop. Se invece intendete visualizzare la finestra di dialogo delle proprietà di sistema sarà sufficiente pronunciare “click con il pulsante destro del mouse su computer” e quindi “click su proprietà”. Quella che state utilizzando non è altro che la funzione di riconoscimento vocale del sistema operativo Vista. Il riconoscimento vocale, sempre più diffuso tra i pacchetti software, infatti, è un processo mediante il quale il linguaggio umano orale prima è riconosciuto e poi elaborato dal computer o da un sistema dedicato.

Applicazioni

Le applicazioni più diffuse del riconoscimento vocale sono in ambito call center, come avviene nel settore della telefonia. Altri esempi piuttosto comuni sono rappresentati dai sistemi di dettatura, dal controllo del sistema di navigazione satellitare oppure dalla gestione della telefonia mobile sulle automobili attraverso comandi vocali. Ma il settore telefonico e dei servizi con call center non è l'unico settore industriale ad avvalersi di questa tecnologia; in diverse realtà del settore della logistica, infatti, la



Fonte: Vocollect

Il riconoscimento vocale è sempre più diffuso nella logistica di magazzino perché aumenta l'efficienza del personale e la qualità del lavoro

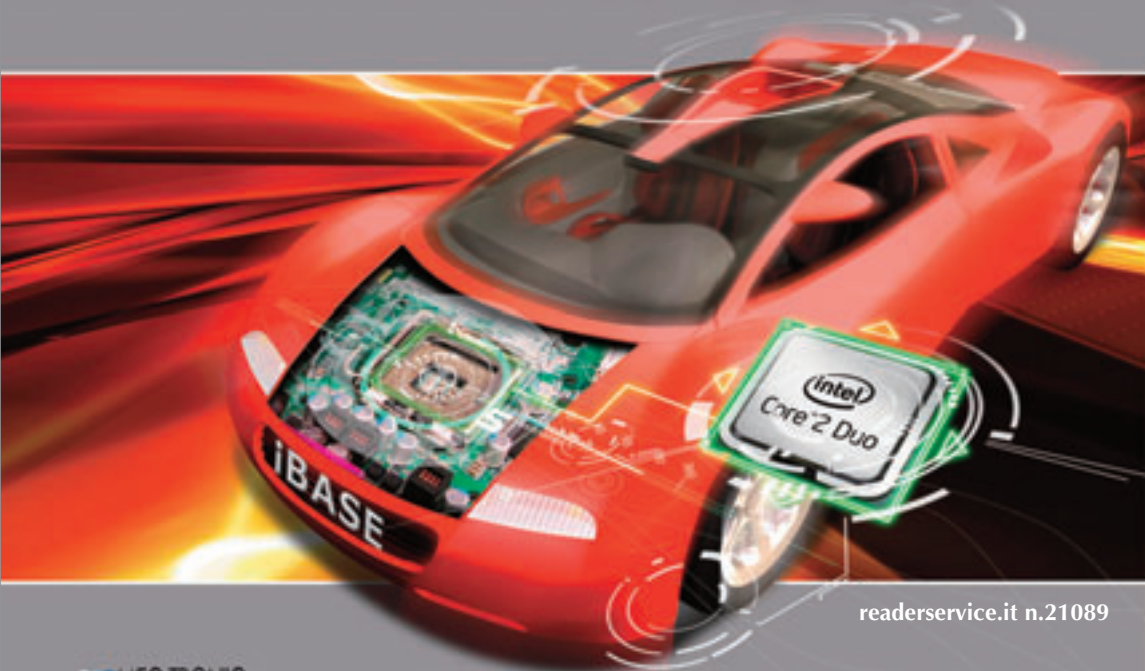
gestione del magazzino sfrutta gli innovativi sistemi vocali: i dati presenti nei sistemi informativi sono tradotti in istruzioni fonetiche per gli operatori di magazzino e la risposta vocale di questi ultimi è a sua volta tradotta e digitalizzata per l'elaborazione del sistema informativo. L'implementazione corretta di questi sistemi permette di migliorare la precisione, incrementare significativamente la produttività, ridurre i tempi di formazione del personale e ridurre notevolmente la quantità di documenti cartacei - aspetto non trascurabile in un periodo in cui le tendenze confermano una digitalizzazione sempre più spinta dei processi aziendali. La soluzione vocale, infatti, consente di comunicare mentre si maneggiano i prodotti, per cui la stessa quantità di lavoro può essere ultimata in meno tempo. Poiché il riconoscimento vocale comporta miglioramenti significativi sia nella produttività sia nell'accuratezza dei dati per operazioni quali il prelievo a magazzino, che rappresenta più del 40 per cento dei costi di manodopera e più del 45 per cento degli errori nelle operazioni di magazzino, l'impatto nel settore logistico non è trascurabile e l'effetto rivoluzionario è paragonato da qualcuno addirittura al codice a barre. Anche il settore sanitario è molto sensibile alle problematiche di digitalizzazione essendo un campo di applicazione ben noto per la documentazione da gestire (immagini, cartelle cliniche

ecc.). L'American Medical Transcription Association ha stimato che la spesa annua per la dettatura di documenti si aggira intorno ai dodici miliardi di dollari; se a questo dato si aggiunge anche la necessità di adottare sistemi di comunicazione efficaci per il lavoro di équipe allora si può immaginare quanto possa essere il valore aggiunto di una soluzione tecnologica con funzioni automatiche di acquisizione, registrazione e distribuzione delle informazioni in tempi rapidi. Tutto questo si traduce ovviamente in un servizio sanitario di qualità più elevata e uno strumento con cui raggiungere questo obiettivo è certamente un sistema di riconoscimento vocale che possa eliminare le note scritte a mano e convertire in tempo reale le informazioni in formato digitale.

Attenti alla biometria

Il riconoscimento della voce è una delle tecnologie utilizzate dalla biometria per l'autenticazione delle persone. La sua trasposizione nell'ambito dei processi automatici incontra un grado di accettabilità da parte degli utenti superiore rispetto alla cattura della geometria del volto e ad altre tecnologie biometriche più intrusive, come il riconoscimento dell'iride. La voce umana è anche l'unica delle caratteristiche biometriche a presentare una connotazione tipicamente fisiologica ma anche una sensibile

IBASE's Full Lineup of Intel® Core™2 Duo Solutions


MEC TRONIC


readerservice.it n.21089

Mini-ITX Motherboard

MI910

Intel® GM965 Express-based



MI900

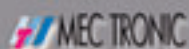
Intel® Q965 Express-based



ECX SBC

ECX800

Mobile Intel® 945GM Express-based

C.so Milano n.180 - Gravellona Toce (VB)
Tel. +39 0323 86931 - Fax +39 0323 869322
Web: <http://www.mectronic.it> - Email: info@mectronic.it






Fonte: Philips

In ambito medico è possibile digitalizzare direttamente sotto dettatura note, diagnosi e consultazioni cliniche

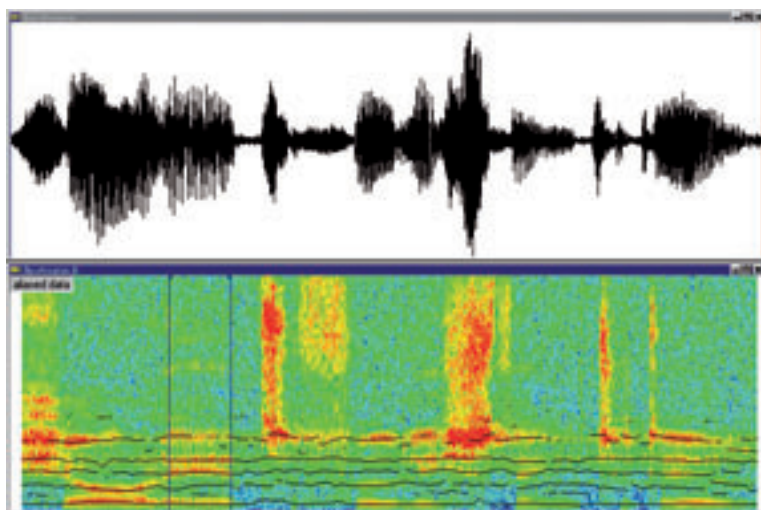
influenza comportamentale legata allo stato psicologico dell'individuo tale da compromettere persino l'unicità del timbro vocale. La principale metodologia utilizzata per individuare l'impronta vocale di una persona consiste nell'analisi del contenuto in frequenze delle onde acustiche risultanti dal flusso d'aria generato nei polmoni, propagato attraverso il condotto tracheale e portato in risonanza dalle corde vocali. Un limite del riconoscimento vocale come tecnica biometrica di verifica e di autenticazione di utenti 'registrati' nel database è proprio il timbro vocale: nel lungo termine è modificabile con l'età o con il degrado fisico, nel breve termine è sensibile a stress, fenomeni influenzali e allergie.

Tecnologia

I programmi di riconoscimento vocale si basano generalmente su algoritmi adattativi di tipo statistico inferenziale che ricostruiscono il linguaggio in base alla frequenza delle parole. Di solito è presente un vocabolario fonetico con la pronuncia base della maggior parte delle parole della lingua selezionata. Altri sistemi per l'analisi del linguaggio, invece, utilizzano un vocabolario costruito dai programmatori e dall'utente, non dal software. Questi pacchetti applicativi non devono avere un'intelligenza propria perché hanno semplicemente il compito di trascrivere quan-

to pronunciato da una persona umana; per questa ragione è di notevole aiuto nel migliorare la qualità del riconoscimento l'utilizzo di un vocabolario non solo di singole parole ma anche di associazioni tipiche tra le parole, inserito dall'uomo oppure ricostruito con mezzi statistici. Nel campo del riconoscimento vocale esistono prevalentemente due tecnologie: quella più recente, denominata Assf (Auditory Spectrum-based Speech Feature), poco diffusa e messa a punto qualche anno fa da ricercatori universitari della Hong Kong Polytechnic University, utilizza una frazione della potenza di calcolo richiesta da applicazioni dello stesso genere per cui è una tecnologia che si presta bene per l'utilizzo su dispositivi come cellulari e PDA. La tecnologia più diffusa è invece la Mfcc (Mel Function Cochlear Coefficient), più pesante per i processori perché utilizza numerosi parametri per decodificare la voce

(o, meglio, la forma d'onda del parlato), mentre la Assf sfrutta un insieme sofisticato di regole decisionali che possono essere eseguite direttamente in memoria, non sul



Spettrogramma della voce umana

Fonte: www.gazzaro.it

processore. La Mfcc funziona molto bene in ambienti silenziosi, mentre la Assf è insuperabile in ambienti disturbati (uffici, stazioni ecc.). Le regole decisionali della tecnologia Assf, inoltre, sembrano favorire il riconoscimento dei singoli comandi e di frasi complesse per cui essa può essere una tecnologia affidabile che permette l'utilizzo di un linguaggio più naturale per la gestione dei dispositivi che la implementeranno. ■