

Riconoscimento vocale

Fra le tecnologie di riconoscimento automatico dei dati, quella basata sulla voce è in pieno sviluppo e diffusione. Al momento, il principale settore di applicazione è quello del

picking di magazzino, dove consente un notevole incremento di velocità e di efficienza. Ma non è il solo ambiente dove l'operatività a mani libere, ovvero senza il necessario impiego di lettori di codice a barre o terminali portatili, può rivelarsi utile.

La voce è una tecnologia in notevole crescita nei magazzini e centri di distribuzione: vediamo le sue principali caratteristiche.

La prima e più evidente è ovviamente l'interfaccia utente. Anziché avere una lista stampata, gli operatori che lavorano con sistemi vocali utilizzano il riconoscimento e la sintesi della voce come mezzo di trasmissione delle informazioni. Queste operazioni sono effettuate mediante un terminale che l'utente, in pratica, tiene addosso, alla cintura o anche in tasca; egli dispone altresì di una cuffia, per sentire gli ordini, e di un microfonino, per convalidarli o trasmettere altre informazioni.

Il settore di applicazione più ovvio e adatto alla tecnologia vocale è quello del picking e preparazione degli ordini, ma non è il solo. Nell'industria e nella logistica si può in generale utilizzare la voce per operazioni di prelievo, di "ventila-

HX2 (foto LXE).

zione" dei colli (preparazione per articolo), di cross docking, di inventario o anche di stock replenishment.

L'obiettivo di base del picking vocale è insomma quello di usare la voce come mezzo di comunicazione fra l'applicazione WMS e l'operatore, in modalità bidirezionale. Il fatto è che questa comunicazione è significativa, in quanto è una vera e propria modalità di raccolta dei dati. In pratica essa permette di sostituire la parte "schermo" di un terminale, inviando dei messaggi dall'applicazione all'operatore (speech out) e, anche se non in tutti i casi, la parte "tastiera" o "scanner", pronunciando semplicemente delle parole (speech in). Ne consegue che l'operatore può concentrarsi molto di più sul proprio compito. Può usare entrambe le mani e non togliersi i guanti, cosa nient'affatto secondaria in un ambiente freddo. E sbaglia molto di meno, perché il sistema vocale lo guida passo passo. Infine, pur non essendo questa la voce di costo che ne beneficia maggior-

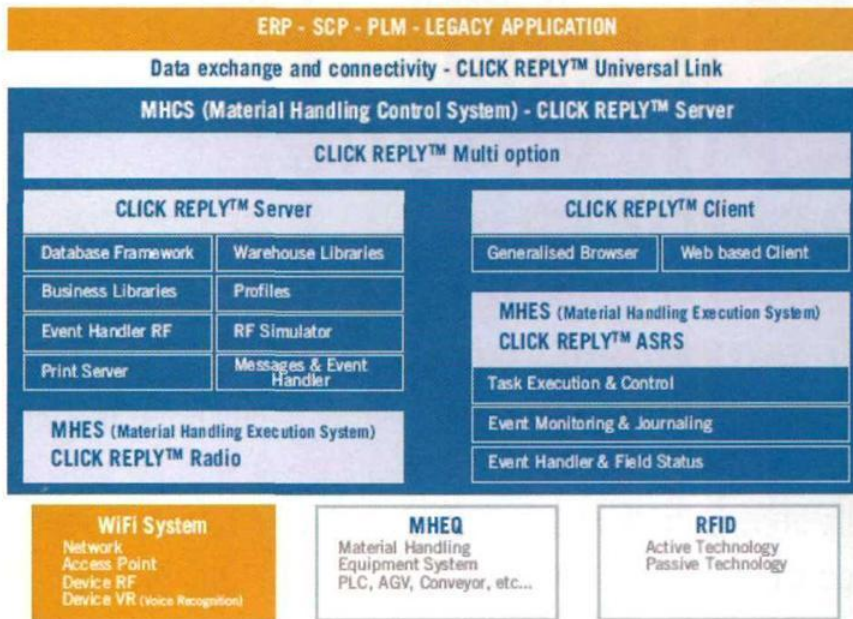


Raccolta dati in magazzino (foto Itworks).

mente, anche la riduzione della documentazione cartacea (liste di picking ecc.) si fa sentire come generale motivo di risparmio.

La base del riconoscimento vocale è il "fonema", ovvero il suono singolo che si ottiene dalla scomposizione delle parole. Ogni lingua ha il proprio set di fonemi, e la tecnologia di base del riconoscimento vocale è proprio quella del riconoscimento dei fonemi: ciò che cambia sono soltanto i requisiti di potenza e di memoria della macchina, diversi se si devono riconoscere poche parole o un intero vocabolario. A cambiare insomma è il tempo di processamento, a seconda della lista di parole che il sistema conosce e deve riconoscere. Il motore vocale classico, necessariamente basato su piattaforme informatiche più potenti, è quello che serve per applicazioni di riconoscimento del parlato, i cosiddetti "dittafoni". Il riconoscimento vocale di poche parole, tipico delle applicazioni voice picking, è nulla più che una semplificazione di questo stesso motore vocale. Inoltre, si parla di "motore" vocale, ma si tratta comunque di prodotti software.





Il sistema VerbaPick di Teleservizi nel magazzino capi di Vestebene, la divisione abbigliamento del gruppo Miroglio, di Pollenzo (Cuneo).

Schema operativo della suite Click Reply, la piattaforma proprietaria Reply per la Supply Chain, con in più il modulo specifico per il Voice Recognition. Tale soluzione presuppone l'utilizzo di hardware Vocollect.

I sistemi vocali possono essere "speaker dependent" o "independent", che è una caratteristica in nuce del motore vocale. La modalità speaker dependent si può scegliere per esempio nel caso in cui vi sia una maggioranza di operatori di lingua straniera, in quanto anche per pronunciare le parole che il sistema si aspetta, non utilizzano il set di fonemi previsto dal sistema stesso. In questo caso si utilizza la modalità speaker dependent, e di conseguenza si deve effettuare una profilazione vocale di tutti gli utenti. Nella modalità speaker independent, al contrario, tale profilazione non è necessaria.

Infine, i sistemi vocali individuano due filoni ben distinti: innanzitutto, si può scegliere fra terminali standard con opzione voce o terminali dedicati alla sola applicazione vocale.

Diversi come prestazioni e come caratteristiche hardware, si devono scegliere in base alle effettive esigenze. L'effetto più evidente della scelta di terminali multimodali è che evita l'acquisto di prodotti dedicati e permette così di ammortizzare più rapidamente il costo di acquisizione dei terminali e di ripartirlo su differenti applicazioni, sia vocali che non. Al contrario i terminali dedicati sono la scelta ottimale in termini di caratteristiche hardware, che riguardano aspetti solo apparentemente di dettaglio, come la qualità delle cuffie auricolari, del segnale audio e della robustezza nel complesso.

Dal lato software e di connettività alla rete, il traffico in aria generato dai sistemi vocali è molto limitato; in aria viaggiano solo caratteri, tutta l'attività di riconoscimento vocale e signal proces-

sing è fatta sul terminale. Ciò che deve essere calibrato è il sistema nel suo complesso, tenendo presente che da un certo punto di vista il sistema vocale può essere definito "sincrono" quindi deve rispondere in tempo reale, altrimenti diventa un collo di bottiglia. Lo stesso vale per la connessione di rete. La copertura Wi-Fi deve essere perfetta, ovunque. Una rete Wi-Fi che andava benissimo con un sistema basato su liste di picking che arrivavano in modalità batch, può rivelarsi insufficiente per un sistema vocale, per un eventuale calo di copertura che, prima, non costituiva un problema. E questo, perché il vocale è sincrono, anziché asincrono. Mentre in qualsiasi sistema asincrono un leggero "buco" di copertura è tollerabile, il sistema vocale necessita di una connessione perfettamente omogenea: pochi byte, ma ovunque.



La voce presso il magazzino ortofrutta del gruppo Conad, sistema realizzato da KFI Trading e Vocollect.