

AGENTI VIRTUALI PER L'INSEGNAMENTO E L'APPRENDIMENTO

Piero Cosi*

IFD-CNR

Istituto di Fonetica e Dialettologia – Consiglio Nazionale delle Ricerche

e-mail: cosi@csrf.pd.cnr.it www: <http://www.csrf.pd.cnr.it>
<http://nts.csrf.pd.cnr.it/Ifd>

1. SOMMARIO

Questo lavoro si prefigge semplicemente di introdurre e descrivere lo scenario presente e futuro in cui le nuove tecnologie audio/visive relative alla realizzazione degli “agenti parlanti”, emotivi ed espressivi, sono e saranno sempre più protagoniste nella comunicazione uomo-computer.

2. INTRODUZIONE

Negli ultimi anni la riproduzione di un testo scritto tramite una voce sintetica (*Text-To-Speech, TTS*) ha raggiunto un grado di qualità tale da poter essere sfruttato proficuamente in moltissime applicazioni [1]. Allo stesso modo nel campo della grafica digitale si sono fatti progressi un tempo immaginabili che hanno consentito la realizzazione di veri e propri **Agenti Virtuali** in grado di interagire con l'utente, nel caso ad esempio di sistemi di comunicazione uomo-macchina, in modo semplice e naturale [3]. La naturalezza della “voce” sintetica come pure la verosimiglianza delle espressioni facciali ad essa collegate sono ovviamente il punto focale di questi sistemi e ne determinano la loro efficace applicabilità ed utilizzazione.

Relativamente alla sintesi audio, già da molto tempo, esistono sistemi in grado di “parlare” in modo intelligibile, in cui l'ascoltatore, cioè, riesce a capire il contenuto del messaggio prodotto, a distinguere le parole e a individuarne il significato. In questi sistemi però è indiscutibile riconoscere l'identità sintetica del parlante, perché alcune delle caratteristiche naturali del linguaggio parlato umano, quali ad esempio la corretta coarticolazione della pronuncia dei fonemi, l'intonazione, il ritmo, la fluidità e altri parametri ancora, non sono correttamente simulati. La ricerca in questo settore, grazie soprattutto allo sviluppo di nuovi modelli teorici e di nuove tecniche digitali di sintesi, alla simultanea e vertiginosa crescita delle capacità computazionali dei calcolatori e alla facile reperibilità di enormi

librerie di dati vocali, ha affinato le tecniche di sintesi portando a risultati più che soddisfacenti in ormai quasi tutte le lingue.

Allo stesso modo, le tecniche digitali di animazione sono ormai arrivate a livelli talmente sofisticati per cui è spesso difficile distinguere gli attori “umani” da quelli “sintetici”, cioè animati artificialmente, nelle produzioni cinematografiche o nei programmi interattivi e nei giochi al computer. Pur tuttavia rimane da sottolineare che la riproduzione artificiale di espressioni umane è comunque ben diversa dalla loro generazione automatica a partire da un testo scritto e per questa devono continuamente essere sviluppate ulteriori nuove ricerche e nuovi modelli.

3. APPLICAZIONI

Le possibili applicazioni di facce parlanti espressive e naturali sono intuitivamente numerosissime e l'elenco seguente vuole essere solo un piccolo esempio di quelle che potrebbero maggiormente trarre vantaggio dall'utilizzazione di agenti virtuali animati e parlanti:

- **Servizi di telecomunicazioni.** Sapendo che circa il 70% delle telefonate richiede un livello minimo di interazione si può pensare ad esempio ad un/una “centralinista virtuale” che riesca a gestire il servizio automaticamente. Oppure alcune informazioni possono essere “lette” in tempo reale comunicando, ad esempio, chi sta chiamando il nostro cellulare, prima di prendere effettivamente la linea, oppure i messaggi presenti nel nostro account di posta elettronica o ancora i fax che abbiamo ricevuto, ecc.
- **Insegnamento e apprendimento.** Gli agenti parlanti possono infatti essere abbinati con un sistema computerizzato per l'insegnamento e l'apprendimento in generale come anche del linguaggio e delle lingue.
- **Aiuto alle persone disabili.** I disabili visivi possono trarre enormi vantaggi da un sistema di riproduzione vocale naturale, mentre gli agenti parlanti espressivi e naturali, introducendo la modalità visiva, oltre a quella uditiva, nella comunicazione uomo-macchina, possono essere di grande aiuto soprattutto alle persone che soffrono di disturbi uditivi.
- **Libri e giocattoli animati e parlanti.** Il mercato dei giocattoli ha già sfruttato la sintesi in passato, ma la bassa qualità della riproduzione ha indirizzato le ditte verso prodotti che utilizzassero frasi pre-registrate più naturali. Un'inversione di tendenza, motivata da una migliore qualità acustica della sintesi vocale e soprattutto dall'interesse dimostrato dalle fasce più giovani di utilizzatori, verso l'introduzione di agenti parlanti, è già in atto.
- **Messaggi vocali.** In alcuni casi le informazioni orali sono più efficienti di quelle scritte o visuali perché permettono di essere recepite mantenendo

l'attenzione focalizzata su fonti di informazioni diverse. Da qui l'idea di incorporare i sintetizzatori vocali nei sistemi di controllo o di misura, per esempio all'interno delle cabine degli aeroplani per evitare di sovraccaricare le consolle di strumenti visuali.

- **Comunicazioni uomo-macchina e multimedialità.** La sintesi audio-video di alta qualità dei futuri agenti parlanti è una strada necessaria da intraprendere per consentire una comunicazione più diretta, naturale e diffusa tra l'uomo e le macchine.

- **Ricerca.** Gli agenti virtuali parlanti possiedono infatti una caratteristica peculiare che li rende strumenti molto importanti, ad esempio nelle ricerche dei linguisti, degli psicologi. Poiché sono completamente controllabili, nel senso che possono riprodurre fedelmente la voce e le espressioni per cui sono stati realizzati, si possono ad esempio progettare opportuni esperimenti percettivi per verificare le ipotesi delle proprie teorie. Ad esempio, mediante la sintesi uditiva, si può studiare l'efficienza del proprio modello prosodico/intonativo oppure, mediante l'animazione facciale, verificare le teorie psicologiche di percezione multi-modale del linguaggio.

4. “INSEGNAMENTO E APPRENDIMENTO”

L'introduzione degli agenti virtuali parlanti emotivi, abbinati ad un opportuno sistema computerizzato potrà contribuire sicuramente a rendere più efficace ed interattiva l'utilizzazione, soprattutto da parte di studenti in età prescolare e scolare, di sistemi automatici per l'insegnamento e l'apprendimento.

Nel campo della disabilità, come già accennato, i disabili visivi possono trarre enormi vantaggi da un sistema di riproduzione vocale naturale, ad esempio, mediante la semplice fruizione di materiale scritto o l'interazione con i sistemi operativi dei PC in tutti quei lavori, sempre più diffusi, che richiedano l'uso assiduo dei calcolatori.

Introducendo inoltre la modalità visiva nella comunicazione uomo-macchina gli agenti virtuali parlanti possono consentire alle persone che soffrono di disturbi uditivi di poter apprendere più efficacemente le caratteristiche del linguaggio parlato [3-4] (vedi Figura 1).

Una delle principali novità che sono e saranno sempre più introdotte mediante l'utilizzazione di agenti virtuali parlanti emotivi sarà la possibilità di trasmettere non solo il significato del messaggio, ma anche le emozioni e i possibili sentimenti ad esso collegati. Questa caratteristica rende la multimedialità nella comunicazione un passo fondamentale e necessario per una sempre maggiore naturale interazione uomo-macchina e le recenti apparizioni di insegnanti, giornalisti ed annunciatori/trici virtuali in internet

ed in sistemi di apprendimento ed intrattenimento ne sono una prova significativa.



Figura 1. Bambini della Tucker-Maxon Oral School di Portland Oregon durante le loro lezioni sull'apprendimento del linguaggio

5. CONSIDERAZIONI FINALI

Si può certamente affermare che almeno nel breve periodo, gli agenti virtuali certo non sostituiranno gli insegnanti umani, ma sicuramente li affiancheranno con frequenza e diffusione sempre maggiori soprattutto al fine di rendere più efficaci le lezioni e più produttivo l'apprendimento.

Bibliografia

- [1] Van Santen et al. (editors) (1997). *Progress in Speech Synthesis*, Springer Verlag New York, Inc. 1997.
- [2] D.W.Massaro (1998). *Perceiving Talking Faces. From Speech Perception to a Behavioral Principle*, MIT Press, 1998.
- [3] Cole, R. et. al. New tools for interactive speech and language training: Using animated conversational agents in the classrooms of profoundly deaf children. In *Proceedings of ESCA-MATISSE ESCA/SOCRATES Workshop on Method and Tool Innovations for Speech Science Education*, London, UK, 16-17 April, 1999.
- [4] Cole, R. et al. Intelligent Animated Agents for Interactive Language Training . In *STiLL: ESCA Workshop on Speech Technology in Language Learning*, Stockholm, Sweden, May 1998.