MODELLI PROSODICI EMOTIVI DI FESTIVAL IN ITALIANO

**Fabio Tesser, **Piero Cosi, **Carlo Drioli, **Graziano Tisato*

* ISTC-irst, Pantè di Povo (Trento)

** ISTC-SPFD CNR, Padova

RIASSUNTO

Utilizzare la sola informazione ricavabile dal testo in ingresso ad un Text To Speech, per predire l'andamento dei parametri prosodici all'interno della frase è un compito molto laborioso se deve essere basato sulla implementazione delle regole che si conoscono sulla prosodia Italiana.

La generazione data-driven della prosodia facilita questo compito soprattutto se si vuole creare differenti set di stili prosodici.

In questa comunicazione si illustrerà l'utilizzo della teoria degli alberi di decisione per la creazione di differenti moduli prosodici utilizzati per generare l'intonazione e il ritmo relativi all'espressione di alcune emozioni.

Per quanto riguarda l'intonazione sarà illustrata una estensione a questa teoria nello spazio VQ-PaIntE (Vector Quantization- Parametric Intonation Events), questo nuovo approccio utilizza una misura di impurità che tiene in considerazione la distanza nello spazio dei vettori quantizzati, fornendo maggiori informazioni sulla forma degli eventi intonativi durante la fase di costruzione dell'albero di decisione.

Le emozioni prese in considerazione sono 6: collera, disgusto, paura, gioia, tristezza, sorpresa.

Un primo modello di prosodia è stato generato utilizzando un database senza emozioni, successivamente i modelli prosodici per le emozioni sono stati modellati utilizzando un database emotivo.

Il database contiene la registrazione di uno speaker professionista che legge una novella di Buzzati sei volte simulando di volta in volta una diversa emozione.

La modellazione prosodica delle emozioni è creata utilizzando un approccio differenziale: il modulo prosodico emotivo cerca di imparare le differenze tra la prosodia neutra (senza emozioni) e i corrispondenti dati prosodici con le emozioni.

Un'attenzione particolare deve essere posta nella scelta del dominio nel quale effettuare questa differenza, in particolare per l'intonazione verrà mostrato come l'applicazione della PCA (Principal Components Analysis) semplifica di molto gli alberi di decisione risultanti.

I risultati dei precedenti algoritmi sono stati implementati nella nuova versione di Festival per l'italiano, mentre sono in fase di sviluppo alcuni esperimenti soggettivi di accettabilità e riconoscimento delle emozioni.