AUTOMATIC LEARNING OF SPECTRAL FEATURES EXTRACTED FROM CONTINUOUS SPEECH

R. De Mori, S. Rivoira and A. Serra, Instituto Electrotecnica Generale Politecnica Di Torino Torino, (Italy)

The paper describes a system based on stochastic finite state automata for learning how a spoken syllable, extracted from a sentence, gives different features when pronounced in different moments, in different contexts and by different speakers. The outline of the project in which such knowledge is used is described. The purpose is to design a speech understanding system implemented on a computer. The main sources of knowledge to be used are at the following levels: parametric, acoustic, syllabic, lexical, syntactic and semantic.

One of the most difficult problems is that of obtaining syllables from the acoustic data. The solution proposed in this paper uses formant evolutions in time as acoustic forms, a descriptive language of such evolutions and stochastic finite-state automate for the attribution of pieces of descriptions to syllables. The paper describes such automata.