

Modele neurale și statistice pentru recunoașterea vorbirii

Neural and Statistical Models for Speech Recognition

As. Ing. DUMITRU Corneliu Octavian

Teza are ca obiectiv elaborarea unui sistem de recunoaștere a vorbirii în limba română (**ASRS_RL** = **A**utomatic **S**peech **R**ecognition **S**ystem for **R**omanian **L**anguage) și crearea bazelor de date necesare pentru testarea sistemului (**CDRL** = **C**ontinuous **D**atabase **R**omanian **L**anguage, **VDRL** = **V**owel **D**atabase **R**omanian **L**anguage, **DDRL** = **D**igit **D**atabase **R**omanian **L**anguage). Pentru aceasta au fost analizate, aplicate și evaluate metode actuale din domeniul recunoașterii vorbirii. S-au folosit următoarele metode de parametrizare a semnalului vocal: metoda cepstrală, metoda LPC (Linear Predictiv Coding) și metoda PLP (Perceptual Linear Predictiv), extrăgându-se parametrii MFCC (Mel Frequency Cepstral Coefficient), variațiile de ordinul întâi și doi ai acestora, coeficienții LPC și coeficienții PLP.

Am realizat și studii comparative pentru coeficienții utilizați (MFCC, LPC, PLP), pentru timpii necesari testării fiecărui tip de parametru și pentru diferite tipuri de baze de antrenare (cu vorbitori masculini, cu vorbitori feminini și cu vorbitori masculini și feminini) pentru a vedea care set de parametri va conduce la cele mai bune rezultate.

Pentru modelarea sistemului am utilizat metode statistice (Modelele Markov Ascunse - MMA), metode neurale (Perceptronul Multistrat - MLP) și modele hibride (MMA și MLP).

Am făcut o comparație între rezultatele obținute în cazul antrenării MMA doar cu monofoni (cazul modelelor independente de context) și în cazul antrenării MMA cu trifoni (cazul modelelor dependente de context).

Pentru a reduce dezavantajul MMA și anume lipsa de discriminare între modele, am utilizat rețelele neuronale artificiale (MLP). Această metodă neuronală a fost aplicată pentru recunoașterea cifrelor. Rezultatele obținute în recunoașterea vocalelor în limba română cu metoda neuronală le-am comparat cu cele obținute cu metoda statistică pentru a vedea care din cele două metode conduc la rezultate mai bune.

Pentru a crește performanțele sistemului de recunoaștere am încercat utilizarea unui sistem hibrid (combinand MMA și MLP) pentru recunoașterea cifrelor. Eliminându-se astfel pe de o parte lipsa de discriminare dintre modele în cazul MMA și pe de altă parte problema acomodării în timp a sevenței de vorbire în cazul MLP.