

Speech recognizer model-building experiments at the level of acoustics and phonetics, on behalf of developing a speech recognizer for medical reporting

Szabolcs Velkei, Klára Vicsi

Budapest University of Technology and Economics,
Department of Telecommunications and Media Informatics,
Laboratory of Speech Acoustics
1117 Budapest, Magyar tudósok krt. 2.

E-mail: vicsi@tmit.bme.hu, velkei@tmit.bme.hu

In this article a HMM based speech recognition system is introduced, which has been developed at the Laboratory of Speech Acoustics. The final aim is the development of a middle sized continuous speech recognizer. New methods have been developed in the acoustical preprocessing, in the statistical model-building, moreover phonetically, phonological and morphemic levels have been involved for the recognition process. In the first year the optimization of the acoustic, phonetic level was prepared.

The developed systems, called MKBP 0.8 were compared with the well known HMM toolkit, HTK. Different evaluation methods were examined and explained, how the results depend on the methods of the evaluation.

The comparison research shows that by the optimization of the acoustic preprocessing and the development of the acoustics-phonetics models we can increase the recognition accuracy and decrease the time of the processing. Of course, the involvement of the higher linguistic levels will increase significantly the recognition accuracy, but with a better starting it is possible to obtain better ending results.