

DAGA2008/357

Hierarchical multi-class class classification of sound data

J.-H. Bach und J. Anemüller
Universität Oldenburg, Medizinische Physik
j.bach@uni-oldenburg.de

The automatic identification of acoustic objects, particularly speech, in adverse backgrounds can pose a difficult problem. In this work, the subproblem of speech detection in two different backgrounds is tackled using a hierarchical decision tree: The root node provides a coarse classification of the noise background, while the next nodes are optimised to detect speech in one particular background each. The features employed are amplitude modulation spectra of the signals (Kollmeier & Koch, 1993). They represent the modulation content of a signal as a two-dimensional (modulation-frequency vs centre frequency) spectral pattern. The best features for classification at each node are selected independently node-specific) using a sequential floating forward search algorithm. The actual statistical learner is an open source support vector machine (Chang & Lin, 2001, <http://www.csie.ntu.edu.tw/~cjlin/libsvm>). The performance of speech detection in two types of realistic backgrounds, namely pedestrian zone noise and road traffic noise, is presented and compared to non-hierarchical approaches; furthermore, discrimination of male vs female speakers in the same background is evaluated.