FEATURE SELECTION AND STACKING FOR ROBUST DISCRIMINATION OF SPEECH, MONOPHONIC SINGING, AND POLYPHONIC MUSIC (WedPmOR5)

Author(s): Björn Schuller (Technische Universität München, Germany)
Bernardo José Brüning Schmitt (Technische Universität München, Germany)
Dejan Arsic (Technische Universität München, Germany)
Stephan Reiter (Technische Universität München, Germany)
Manfred Lang (Technische Universität München, Germany)
Gerhard Rigoll (Technische Universität München, Germany)

Abstract: In this work we strive to find an optimal set of acoustic features for the discrimination of speech, monophonic singing, and polyphonic music to robustly segment acoustic media streams for annotation and interaction purposes. Furthermore we introduce a novel ensemble-based classification approach within this task. From a basis of 276 attributes we select the most efficient set by SVM-SFFS. Additionally relevance of single features by calculation of information gain ratio is presented. As a basis of comparison we reduce dimensionality by PCA. We show extensive analysis of different classifiers within the named task. Among these are Kernel Machines, Decision Trees, Neural Networks, and Bayesian Classifiers. Moreover we improve single classifier performance by Bagging and Boosting, and finally combine strengths of classifiers by StackingC. The database is formed by 2,114 samples of speech, and singing of 58 persons. 1,000 Music clips have been taken from the MTV-Europe-Top-20 1980–2000.

[continued in the next page]
FEATURE SELECTION AND STACKING FOR ROBUST DISCRIMINATION OF SPEECH, MONOPHONIC SINGING, AND POLYPHONIC MUSIC (WedPmOR5)

Author(s):
Björn Schuller
Bernardo José Brüning Schmitt
Dejan Arsic
Stephan Reiter
Manfred Lang
Gerhard Rigoll
(Technische Universität München, Germany)

Abstract:
(cont.)
The outstanding discrimination results of a working real–time capable implementation stress the practicability of the proposed novel ideas.