



ROBUST NEURAL NETWORK-BASED ESTIMATION OF ARTICULATORY FEATURES FOR CZECH

*Petr Mizera, Petr Pollak**

Abstract: The article describes a neural network-based articulatory feature (AF) estimation for the Czech speech. First, the relationship between AFs and a Czech phone inventory is defined, and then the estimation based on the MLP neural networks is done. The usage of several speech representations on the input of the MLP classifiers is proposed with the purpose to obtain a robust AF estimation. The realized experiments have proved that an ANN- based AF estimation works very reliably especially in a low noise environment. Moreover, in case the number of neurons in a hidden layer is increased and if the temporal context DCT-TRAP features are used on the input of the MLP network, the AF classification works accurately also for the signals collected in the environments with a high background noise.

Key words: *Speech recognition, articulatory features, robust estimation, neural networks, MLP, temporal patterns, TRAP*

Received: December 6, 2013

DOI: 10.14311/NNW.2014.24.027

Revised and accepted: October 3, 2014

1. Introduction

Speech technology applications are nowadays used in many situations, when a natural voice input is used in communication between humans and machines or in the detection of various phenomena using automated speech analysis. As in other research fields, various subparts of speech technology systems use artificial neural networks for the classification purposes [3, 9, 25]. ANNs can be found, for example, as phone classifiers in the Temporal Patterns (TRAP) speech feature extraction [6, 8], as voice activity detectors, as a subpart of the combined Artificial Neural Network and Hidden Markov Model (ANN/HMM) classifiers, in language modelling for continuous speech recognition, or as Kohonen self-organized maps used for selected classification purposes [33], etc.

To improve the robustness of spontaneous or noisy speech recognition, Articulatory Features (AFs) were used in various tasks in continuous speech recognition.

*Petr Mizera, Petr Pollak – Corresponding author, Czech Technical University in Prague, Faculty of Electrical Engineering, CTU FEE K13131, Technicka 2, 166 27 Praha 6, Czech Republic