Informatik-Oberseminar

Fachgruppe Informatik

Lehrstuhl für Informatik 6

RHEINISCH-WESTFÄLISCHE TECHNISCHE HOCHSCHULE AACHEN

EINLADUNG

Zeit: 14. August 2006, 14:00 Uhr

Ort: Raum 5056, Ahornstr. 55

Referent: Dipl.-Ing. András Zolnay

Titel: Acoustic Feature Combination for Speech Recognition

Abstract:

In this presentation, the use of multiple acoustic features of the speech signal is considered for speech recognition. The goals are twofold: on the one hand, new acoustic features are presented, on the other hand, feature combination methods are investigated in order to find an effective integration of the newly developed features into state-of-the-art speech recognition systems.
The most commonly used feature extraction methods are the Mel Frequency Cepstrum Coefficients (MFCC), Perceptual Linear Prediction (PLP), and variations of these techniques. These methods are mainly based on the models of the human auditory system. There have also been attempts at using articulatory motivated acoustic features for speech recognition which are motivated by models of the human speech production system. This work focuses partially on the development of new articulatory motivated acoustic features. On the one hand, the use of the voicing feature is investigated. On the other hand, the novel spectrum derivative feature is introduced which aims to capture the differences between magnitude spectra produced by obstruent and sonant consonants.
The articulatory motivated features are tested in combinations with state-of-the-art acoustic features based on auditory models mainly. The features are combined both directly using Linear Discriminant Analysis (LDA) as well as indirectly on model level using Discriminative Model Combination (DMC). Both methods have already been used successfully in automatic speech recognition systems. In this talk, a comparative study is presented which describes and analyzes the application of these methods to feature combination. Robustness issues of the LDA based method are addressed which are induced by increasing the amount of acoustic features coefficients. An application of DMC to feature combination is introduced based on the splitting of the acoustic model into separate scalable knowledge sources. After the analysis of the individual methods, a comparison is carried out on the basis of the underlying acoustic emission models.
Experimental results are presented for small- and large-vocabulary tasks. The results show that the accuracy of automatic speech recognition systems can be significantly improved by the combination of auditory and articulatory motivated features.

Es laden ein: Die Dozenten der Informatik