

16th July 2015, SLP conf. night session@Suwako, Nagano

言語識別

presented in ICASSP 2015

岡本 拓磨

情報通信研究機構

Improved language identification using deep bottleneck network

pp. 4200–4204

■ Y. Song, R. Cui, X. Hong, I. Mcloughlin, J. Shi and L. Dai

■ USTC and Anhui Post and Telecommunication College

■ 目的

■ 短い発話や方言における精度改善

■ 従来法

■ 入力層を音響特徴量, 出力層を音素ラベルとしたDNNのBN特徴量をi-vector表現

■ 提案法の特徴

■ BNだけでなく出力層(Senones proposed in INTERSPEECH 2014)を使用

■ 実験結果

■ NIST LRE2011 アラビア語方言, NIST LRE2009の23言語にて従来法よりも高精度

Improved language identification using deep bottleneck network

■ 提案法のブロック図

本文のFig. 1 参照

■ 実験結果

本文のTable. 1 参照

本文のTable. 2 参照

Nearest neighbor discriminant analysis for language recognition

pp. 4205–4209

■ S. O. Sadjadi, J. W. Pelecanos and S. Ganapathy

■ IBM Research, Watson Group

■ 目的

■ i-vector型ノイズロバスト言語識別のポスト処理

■ 従来法の問題点

■ linear discriminant analysis(LDA)を使用しi-vectorをガウス分布で仮定

■ 言語数が少ない場合効果が低い

■ 提案法

■ LDAの一般型であるnearest neighbor discriminant analysis(NDA)を導入

■ 実験結果

■ DARPA RATS プログラムのノイズを含む音声を用いた実験により高精度を確認

Nearest neighbor discriminant analysis for language recognition

■ 提案法のブロック図

本文のFig. 1参照

■ i-vectorの比較

本文のFig. 3参照