

Public Abstract

First Name:Chanmann

Middle Name:

Last Name:Lim

Adviser's First Name:Yunxin

Adviser's Last Name:Zhao

Co-Adviser's First Name:

Co-Adviser's Last Name:

Graduation Term:SP 2016

Department:Computer Science

Degree:MS

Title:Towards Deep Learning on Speech Recognition for Khmer Language

In order to perform speech recognition well, a huge amount of transcribed speech and textual data in the target language must be available for system training. The high demand for language resources constrains the development of speech recognition systems for new languages. In this thesis the development of a low-resourced isolated-word recognition system for "Khmer" language is investigated. Speech data, collected via mobile phone, containing 194 vocabulary words is used in our experiments. Data pre-processing based on Voice Activity Detection (VAD) is discussed. As by-products of this work, phoneme based pronunciation lexicon and state tying questions set for Khmer speech recognizer are built from scratch. In addition to the conventional statistical acoustic modeling using Gaussian Mixture Model and hidden Markov Model (GMM-HMM), a hybrid acoustic model based on Deep Neural Network (DNN-HMM) trained to predict context-dependent triphone states is evaluated. Dropout is used to improve the robustness of the DNN, and cross-lingual transfer learning that makes use of auxiliary training data in English is also investigated. As the first effort in using DNN-HMM for low-resourced isolated-word recognition for Khmer language, the system currently performs at 93.31% word accuracy in speaker-independent mode on our test set.