CSLT TECHNICAL REPORT-20160003 [Thursday 4th February, 2016]

A study of Similar Word Model for Unfrequent Word Enhancement in Speech Recognition

Xi Ma[1,3] and Dong Wang[1,2*]

*Correspondence: wang-dong99@mails.tsinghua.edu.cn
[1]Center for Speech and Language Technology, Research Institute of Information Technology, Tsinghua University, ROOM 1-303, BLDG FIT, 100084 Beijing, China
Full list of author information is available at the end of the article

Abstract

abstract here..

Keywords: speech recognition; semantic similarity; language model

The popular n-gram language model (LM) is weak for unfrequent words, particularly words that are out of vocabulary (OOV) or out of language (OOL). Conventional approaches such as class-based LMs pre-define some sharing structures (e.g., word classes) to improve the probability estimation for unfrequent words. However, defining such structures requires prior knowledge, and the context sharing based on these structures is generally inaccurate.

This paper presents a novel similar word model to enhance unfrequent words. In principle, this approach enriches the context of an unfrequent word by borrowing context information from some 'similar words'. Compared to conventional class-based methods, this new approach offers a fine-grained context sharing that enhances each word by borrowing context information from words that match the target word best. Additionally, the new approach is highly flexible since no sharing structures need to be defined prior to LM training. We tested the proposed method with a large-scale Chinese speech recognition system. The experimental results demonstrated that the similar word approach can improve performance on unfrequent words significantly, while keeping the performance on general tasks almost unchanged.

1 Introduction

introduction here

2 Related work

related work here

3 Theory

theory here...

4 Experiment

experiments here...

5 Conclusions

conclusions here...

Acknowledgement

Author details
[1]Center for Speech and Language Technology, Research Institute of Information Technology, Tsinghua University, ROOM 1-303, BLDG FIT, 100084 Beijing, China. [2]Center for Speech and Language Technologies, Division of Technical Innovation and Development, Tsinghua National Laboratory for Information Science and Technology, ROOM 1-303, BLDG FIT, 100084 Beijing, China. [3]Department of Computer Science and Technology, Tsinghua University, ROOM 1-303, BLDG FIT, 100084 Beijing, China.

References