基于循环神经网络的音素识别研究

黎长江，胡　燕

(武汉理工大学 计算机科学与技术学院，湖北 武汉 470070)

摘　要：基于隐马尔科夫模型（HMM）和循环神经网络（RNN）的 HMM-RNN 混合模型在语音识别中取得了很大的成功.由于 RNN 的计算过程本身就是上下文相关的，相邻帧的重合部分增加了整个系统的训练时间.针对上述问题，使用连接时序分类（CTC）来代替 HMM 跟 RNN 结合，并在语音分帧过程中去除相邻帧之间的重合部分，使用 TIMIT 语音数据集，进行音素上的识别任务，并且实验结果表明 CTC-BLSTM 模型在音素上的识别率要高于 HMM-BLSTM 混合模型，CTC-BSLTM 在去除帧重合后能够大幅提高系统的训练效率并且保证识别率大致相同.

关键词：语音识别；连接时序分类；循环神经网络

Research of Phoneme Recognition Based on Recurrent Neural Network

LI Chang-jiang，HU Yan

(School of Computer Science and Technology，Wuhan University of Technology，Wuhan 430070，China)

Abstract：Recently HMM-RNN hybrid system has been proved to be successful in speech recognition, since the calculation of RNN is context-dependent， the overlapped part increases the training time. This paper combines CTC with RNN instead of HMM， and remove the overlapped part during framing modeling TIMIT dataset on phone recognition tasks. The experiments show that CTC-BLSTM performs better than HMM-BLSTM on phone recognition， and removing the overlapped part of frames can make system more efficient and ensure the accuracy at a certain degree.

Key words： speech recognition； CTC； RNN

作者简介：

黎长江　男，(1991-)，硕士研究生.研究方向为语音识别.E-mail:jonathan.l@qq.com.

胡　燕　女，(1965-)，教授，硕士生导师.研究方向为智能信息检索.