

学校编码: 10384
学号: 31520131153294

分类号 _____ 密级 _____
UDC _____

厦 门 大 学

硕 士 学 位 论 文

基于随机数字串的文本提示型声纹识别系统
The Text-prompted Speaker Recognition System With
Random Digit Strings

刘志坚

指导教师姓名: 洪青阳 副教授
专业名称: 计算机技术
论文提交日期: 2016 年 月
论文答辩时间: 2016 年 月
学位授予日期: 2016 年 月

答辩委员会主席: _____
评 阅 人: _____

2016 年 月

厦门大学博硕士学位论文摘要库

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学博硕士学位论文摘要库

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1. 经厦门大学保密委员会审查核定的保密学位论文，于 年 月 日解密，解密后适用上述授权。

2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

厦门大学博硕士学位论文摘要库

摘要

声纹识别技术作为模式识别与人工智能的一个重要分支，在生物特征识别中占据举足轻重的作用，近年来，在科研和工业领域都得到了快速发展和广泛关注。本文主要是对基于随机数字串的文本提示型声纹识别系统进行研究，在此基础上引入了 DNN-HMM 算法以及在 HTK3.5 上进行数字识别相关实验。

本文从传统的声纹识别算法的介绍入手，然后从理论和实际应用中分析了传统算法存在的不足，由此引入了基于绑定混合模型的提示型声纹识别算法，阐述了算法的相关原理和执行流程，并通过实验验证其性能。本文在 Android 嵌入式平台上实现了基于随机数字串的提示型声纹识别系统，对相关算法进行了实际应用，从工程上验证了文本提示型声纹识别算法。

在实际应用中，文本相关的声纹识别技术一直存在着录音回放这一大安全隐患，而与此同时基于随机文本的文本提示型声纹识别技术又存在语法文件过于复杂，训练较为耗时，生成模型庞大。基于这样的瓶颈，本文重点研究了基于随机数字串的提示型声纹识别技术。我们以真实录制的数据（TSDIGIT 数据集）为出发点，首先介绍了绑定混合模型应用于说话人识别系统中对系统识别性能的提升，然后阐述了基于 DNN-HMM 的语音识别算法，并利用 HTK3.5 对 TSDIGIT 数据集进行实验，验证了 DNN-HMM 对语音识别性能的明显提升。

最后，结合实际项目，构建特定声纹识别引擎，在 Android 嵌入式平台上实现了基于随机数字串的提示型声纹识别系统。

关键词： 文本提示型声纹识别； 绑定混合模型； 深度神经网络.

厦门大学博硕士学位论文摘要库

Abstract

Speaker recognition, as one important branch of pattern recognition and artificial intelligence, is playing an important role in biometric authentication. In recent years, it has developed rapidly and become more popular in the industry. This study focuses on the text-prompted speaker recognition system with random digit strings. The DNN-HMM algorithm and the digit recognition experiment based on HTK3.5 are also studied.

This study starts with the basic introductions of classic speaker recognition algorithms and the analysis of their defects. After that, we have a deep study of tided mixture model (TMM), including its principle and working process, and the experiment proved its superior performance. In order to apply the algorithm into practice, the text-prompted speaker recognition system with random digit strings is developed based on Android platform.

In practice, there has been security issue from playback in the text-dependent speaker recognition while the ordinary text-prompted speaker recognition technology has encountered the difficulties of complex syntax, time-consuming training, and large models. Due to the performance bottleneck, we have a further study of the text-prompted speaker recognition technology with random digit strings. Based on TSDIGIT data set, we firstly introduce the performance improvements of speaker recognition system based on TMM. Then, we describe the DNN-HMM algorithm and conduct an experiment based on TSDIGIT data set using HTK3.5 tools. Experimental results have demonstrated the excellent performance of DNN-HMM.

In the end, we design and realize a text-prompted speaker recognition system with random digit strings, which is running on the Android platform.

Keywords: Text-prompted speaker recognition; TMM; DNN.

厦门大学博硕士学位论文摘要库

目 录

摘 要.....	I
Abstract.....	III
目 录.....	V
Table of Contents.....	IX
第一章 绪论	1
1.1 研究背景及意义.....	1
1.2 文本提示型声纹识别技术研究及现状.....	2
1.2.1 文本提示型声纹识别技术介绍.....	3
1.2.2 文本提示型声纹识别技术研究现状.....	4
1.3 本文主要工作.....	5
1.4 本文结构安排.....	5
第二章 声纹识别系统概述	7
2.1 引言.....	7
2.2 声纹识别系统的基本结构.....	7
2.2.1 特征提取.....	8
2.2.2 阈值选取.....	11
2.2.3 得分规整.....	11
2.3 GMM-SVM 声纹识别系统.....	13
2.3.1 GMM-SVM 声纹识别系统.....	14
2.3.2 GMM-SVM 在动态口令声纹识别系统中的问题.....	15
2.4 基于 i-vector 特征的 PLDA 声纹识别系统.....	16

2.4.1 i-vector/PLDA 声纹识别系统.....	16
2.4.2 i-vector/PLDA 在动态口令声纹识别系统中的问题.....	18
2.5 本章小结.....	19
第三章 绑定混合模型和说话人识别	21
3.1 引言.....	21
3.2 GMM-UBM 说话人识别系统	21
3.3 TMM-UBM 说话人识别系统	23
3.4 实验对比与分析.....	24
3.5 本章小结.....	27
第四章 深度神经网络和语音识别	29
4.1 引言.....	29
4.2 GMM-HMM 语音识别系统	29
4.2.1 隐马尔可夫模型技术.....	30
4.2.2 GMM-HMM 语音识别框架.....	36
4.3 DNN-HMM 语音识别系统	37
4.3.1 深度神经网络技术.....	37
4.3.2 DNN-HMM 语音识别框架.....	41
4.4 基于 HTK3.5 的语音识别实验.....	42
4.4.1 实验数据和实验环境配置.....	42
4.4.2 实验流程.....	43
4.4.3 实验结论.....	49
4.5 本章小结.....	51
第五章 基于 Android 的动态口令声纹识别系统.....	53

5.1 项目意义.....	53
5.2 项目方案.....	53
5.2.1 系统开发环境.....	56
5.2.2 动态口令声纹识别 JNI.....	56
5.2.3 动态口令声纹识别 SDK.....	61
5.2.4 演示系统与工程总结.....	63
5.3 本章小结.....	64
第六章 全文总结及工作展望.....	65
参 考 文 献.....	67
攻读硕士期间的科研成果.....	71
致 谢.....	73

厦门大学博硕士学位论文摘要库

Contents

Abstract in Chinese	I
Abstract in English	III
Contents in Chinese	V
Contents in English	IX
Chapter 1 Introduction	1
1.1 Research Background	1
1.2 Current Issues of Speaker Recognition	2
1.2.1 Summary of Text-prompted Speaker Recognition	3
1.2.2 Research Progress of Text-prompted Speaker Recognition	4
1.3 Main Works of This Study	5
1.4 Structure of This Study	5
Chapter 2 Overview of Speaker Recognition System	7
2.1 Introduction	7
2.2 Basic Structure of Speaker Recognition System	7
2.2.1 Feature Extraction	8
2.2.2 Threshold Selection	11
2.2.3 Score Normalization	11
2.3 Speaker Recognition System Based on GMM-SVM	13
2.3.1 GMM-SVM System	14
2.3.2 Issues of GMM-SVM in Dynamic Password System	15
2.4 Speaker Recognition System Based on i-vector/PLDA	16

2.4.1 i-vector/PLDA System.....	16
2.4.2 Issues of i-vector/PLDA in Dynamic Password System.....	18
2.5 Summary.....	19
Chapter 3 Speaker Recognition System Based on TMM.....	21
3.1 Introduction.....	21
3.2 Speaker Recognition System Based on GMM-UBM.....	21
3.3 Speaker Recognition System Based on TMM-UBM.....	23
3.4 Experimental Result and Analysis.....	24
3.5 Summary.....	27
Chapter 4 Deep Neural Network and Speech Recognition.....	29
4.1 Introduction.....	29
4.2 Speech Recognition System Based on GMM-HMM.....	29
4.2.1 Hidden Markov Model.....	30
4.2.2 Speech Recognition Based on GMM-HMM.....	36
4.3 Speech Recognition System Based on DNN-HMM.....	37
4.3.1 Deep Neural Network.....	37
4.3.2 Speech Recognition Based on DNN-HMM.....	41
4.4 Experiments of Speech Recognition Based on HTK3.5 Tools.....	42
4.4.1 Experimental Data and Configuration.....	42
4.4.2 Experimental Procedures.....	43
4.4.3 Experimental Result.....	49
4.5 Summary.....	51
Chapter 5 Speaker Recognition System of Dynamic Password.....	53

Degree papers are in the “[Xiamen University Electronic Theses and Dissertations Database](#)”.

Fulltexts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.