

学校编码: 10384

分类号 _____ 密级 _____

学号: 20051302334

UDC _____

厦 门 大 学

硕 士 学 位 论 文

嵌入式声纹识别系统的研究与实现

The Research and Implementation of Embedded Speaker
Recognition System

张 彩 红

指导教师姓名: 洪 青 阳 博士

删除的内容: 讲师

专 业 名 称: 计算机应用技术

论文提交日期: 2008 年 月

论文答辩时间: 2008 年 月

学位授予日期: 2008 年 月

答辩委员会主席: _____

评 阅 人: _____

2008 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

兹提交的学位论文，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的其他个人或集体的研究成果，均在文中以
明确方式标明。本人依法享有和承担由此论文产生的权利和责任。

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

本人完全了解厦门大学有关保留、使用学位论文的规定。厦门大学有权保留并向国家主管部门或其指定机构送交论文的纸质版和电子版，有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅，有权将学位论文的内容编入有关数据库进行检索，有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

本学位论文属于

1. 保密 ()，在年解密后适用本授权书。
2. 不保密 ()

(请在以上相应括号内打“√”)

作者签名： 日期： 年 月 日

导师签名： 日期： 年 月 日

厦门大学博硕士学位论文摘要库

摘要

说话人识别技术是指通过对说话人的语音进行信号处理自动鉴别说话人身份的一门技术，是生物认证技术的重要手段之一。它在很多领域都有重要的应用价值。本文介绍的说话人确认系统是建立在 ARM9 开发板硬件基础之上，采用高斯混合全局背景模型算法实现的嵌入式系统。它主要分为三部分：特征向量美尔频率倒谱系数（MFCC）的提取、说话人确认算法实现、硬件平台及其实现。

提取特征向量 MFCC 就是从说话人的语音信号中提取能够体现该说话人声学个特征的参数，并对提取算法以速度为优先指标进行运算优化。本文在语音信号时域处理方法研究的基础上，介绍了特征参数的提取过程及其计算优化。

在说话人识别算法方面，首先介绍了说话人识别的主要模式匹配方法：隐马尔可夫模型（HMM）、高斯混合模型（GMM）、人工神经网络（ANN）等。研究表明，GMM 方法具有独特的优越性在语文本无关的说话人确认系统中取得良好的识别效果。本文重点阐述高斯混合全局背景模型算法，它主要由三个模块构成：全局背景模型(UBM)的建立、目标说话人模型的建立及其自适应、说话人确认决策。

在硬件平台及其实现方面，我们首先介绍嵌入式开发板的选择依据，并最终决定选择基于 ARM9 的 H2410EB 开发板。本文重点阐述了基于 ARM9 的说话人确认系统的设计和实现。由于嵌入式开发板处理能力和存储能力的局限性，采用浮点运算转定点运算、计算预处理、数学快速近似计算、最大值机制等降低系统的运算负荷，提高运算效率，保证嵌入式系统实时性的要求。

关键词： GMM-UBM； ARM9； 嵌入式系统.

厦门大学博硕士学位论文摘要库

Abstract

Speaker recognition is to find out the speaker's identity based on the spoken utterances. As one of the most important biology authentication technologies, it can be widely used in many security areas. In this work, we'll build a speaker recognition system based on the platform of ARM9, using Gaussian mixture models-universal background model. We focus on three key parts: feature extraction, speaker recognition's algorithm, and hardware platforms.

Feature extraction is generally carried out to transform speech waveform into a sequence of observation vectors. Currently, the predominant choice of parameters is *mel* frequency cepstral coefficient (MFCC). In this thesis, we will describe the extraction of MFCC and its optimization.

For the speaker recognition algorithm, we firstly introduce some widely used algorithms: hidden Markov model (HMM), Gaussian mixture models (GMM), and artificial neural networks (ANN). In recent years, it is more common to represent speakers with the Gaussian mixture model (GMM). We'll address in detail three issues: uniform background model (UBM), the adaptation of speaker model, and the verification strategy.

For the hardware platform, we firstly explain the reason to select ARM9 H2410EB. Then we describe the design and implementation of speaker verification system based on this embedded platform. Several new calculational methods are applied to improve the verification performance, such as the transformation from floating-point operation to pointing operation, pretreatment of the calculation, fast and approximate mathematic calculation, and MAX mechanism etc. These methods guarantee the real-time performance of embedded system.

Key Words: GMM-UBM; ARM9; Embedded System.

厦门大学博硕士学位论文摘要库

目 录

第一章 绪论	1
1.1 生物认证.....	1
1.2 说话人识别技术概论	2
1.3 嵌入式说话人识别研究的现状	3
1.4 课题技术难点与需解决的问题	4
1.5 论文的主要工作和章节安排	5
第二章 语音信号处理	6
2.1 语音信号的产生	6
2.1.1 语音的发音器官.....	6
2.1.2 语音的声学特征.....	7
2.1.3 语言信号在时域和频域的表达.....	8
2.1.4 语谱图.....	8
2.2 语音信号的数字模型	9
2.2.1 激励模型	9
2.2.2 声管模型	10
2.2.3 共振峰模型	10
2.2.4 辐射模型	11
2.3 语音信号的数字化和预处理	11
2.4 语音信号线性预测倒谱系数 (Linear Prediction Cepstral Coding, LPCC) 特征参数的提取.....	15
2.5 语音信号 MFCC 特征参数的提取	15
第三章 说话人识别算法	19
3.1 说话人识别算法概述	19
3.2 说话人识别系统特征的选取与评价.....	21
3.3 基于 GMM-UBM 的说话人确认系统	23
3.3.1 高斯混合模型 GMM	23

带格式的：分散对齐

删除的内容: 1

删除的内容: 1

删除的内容: 1

删除的内容: 1

删除的内容: 2

删除的内容: 2

删除的内容: 3

删除的内容: 3

删除的内容: 4

删除的内容: 4

删除的内容: 5

删除的内容: 5

带格式的：分散对齐

删除的内容: 6

删除的内容: 7

删除的内容: 6

删除的内容: 6

删除的内容: 7

删除的内容: 6

带格式的：分散对齐

删除的内容: 15

删除的内容: 16

删除的内容: 15

删除的内容: 15

删除的内容: 16

删除的内容: 15

带格式的：分散对齐

删除的内容: 19

删除的内容: 20

删除的内容: 19

删除的内容: 19

删除的内容: 20

删除的内容: 19

3.3.2GMM-UBM 模型自适应.....	24
3.3.3 实验结果	26
第四章 嵌入式说话人识别系统的硬件平台	28
4.1 硬件的选择	28
4.1.1 凌阳 61 板.....	28
4.1.2 EasyARM2131 开发板.....	31
4.2 基于 ARM9 的 H2410EB 开发板	35
第五章 嵌入式说话人识别系统的算法实现	44
5.1 浮点运算转换为定点运算.....	44
5.1.1 四则运算的转换	45
5.1.2 计算的预处理	46
5.1.3 数学的快速近似计算.....	47
5.1.4 计算方法的改进.....	47
5.2 通用背景模型 UBM 的建立.....	48
5.3 说话人模型的优化	49
5.4 实验结果和实验分析.....	50
第六章 课题总结和展望	52
6.1 课题总结	52
6.2 课题展望	52
参考文献	54
致谢.....	57

- 带格式的：分散对齐
- 删除的内容: 45
- 带格式的：分散对齐
- 删除的内容: 44
- 删除的内容: 44
- 删除的内容: 44
- 删除的内容: 45
- 删除的内容: 44
- 删除的内容: 45
- 删除的内容: 46
- 删除的内容: 45
- 删除的内容: 46
- 删除的内容: 47
- 删除的内容: 46
- 删除的内容: 47
- 删除的内容: 48
- 删除的内容: 47
- 删除的内容: 47
- 删除的内容: 48
- 删除的内容: 47
- 删除的内容: 48
- 删除的内容: 47
- 删除的内容: 48
- 删除的内容: 48
- 删除的内容: 49
- 删除的内容: 48
- 删除的内容: 49
- 删除的内容: 48
- 删除的内容: 49
- 删除的内容: 50
- 删除的内容: 49
- 删除的内容: 50
- 删除的内容: 51
- 删除的内容: 49
- 带格式的：分散对齐
- 删除的内容: 52
- 删除的内容: 53
- 删除的内容: 51
- 删除的内容: 52
- 删除的内容: 53
- 删除的内容: 51
- 删除的内容: 52
- 删除的内容: 53
- 删除的内容: 51
- 带格式的：分散对齐
- 删除的内容: 54
-
-
-
-

Contents

Chapter 1: Introduction	1
1.1 Biological Authentication.....	1
1.2 Introduction of Speaker Recognition	2
1.3 The present research of Embedded Speaker Recognition	3
1.4 The present research difficulties and problems	4
1.5 Research contents and chapter arrangement.....	5
Chapter 2: Speech signal processing	6
2.1 The production of speech signal.....	6
2.1.1 speech organs	6
2.1.2 acoustics character	7
2.1.3 time domain and frequency domain conversion of speech	8
2.1.4 speech spectrogram	8
2.2 The digital models of speech signal.....	9
2.2.1 motivational model	9
2.2.2 vocal tract model.....	10
2.2.3 formant model.....	10
2.2.4 radiant model	11
2.3 The digital and pretreatment of speech signal	11
2.4 The extracting of LPCC	15
2.5 The extracting of MFCC	15
Chaper 3: The algorithm of speaker recongnition	19
3.1 The introduction of speaker recognition algorithm.....	19
3.2 The choice and maluation of characteristic quantity in the system	21
3.3 The speaker verification system based on GMM-UBM	23
3.3.1 Gaussian Mixture Model.....	23
3.3.2 Universal Background Model	24

删除的内容: 4

删除的内容: 6

3.3.3	Experimental results.....	26	
Chapter 4: The hardware platform of embedded speaker			
recognition		28	
4.1	The choice of hardware	28	
4.1.1	Sunplus MCU.....	28	
4.2.2	EasyARM2131.....	31	
4.3	H2410EB-ARM9.....	35	
Chapter 5: The implementation of embedded speaker verification			
system.....		44	
5.1	The transformation from floating to pointing operation	44	
5.1.1	The transformation of arithmetic	45	删除的内容: 4
5.1.2	The pretreatment of operation	46	删除的内容: 5
5.1.3	Fast and approximate operation	47	
5.1.4	Improvement of mathematical operation.....	47	
5.2	Modeling of UBM.....	48	
5.3	The optimization of speaker models.....	49	
5.4	Experiment results and analysis.....	50	删除的内容: 49
Chaper 6: Conclusion and Future work.....			
		52	删除的内容: 1
6.1	Conclusion.....	52	删除的内容: 1
6.2	Future work.....	52	删除的内容: 1
References.....			
		54	删除的内容: 3
Acknowledgement.....			
		57	删除的内容: 5

第一章 绪论

1.1 生物认证

随着人工智能和生物科学技术的不断发展，生物认证技术越来越受到人们的重视。生物认证也叫生物识别，是通过计算机利用人体所固有的生理特征或行为特征来进行个人身份的鉴定。

在我们的日常生活中，很多场合都需要我们进行身份的认证。目前国内外个人身份认证多采用“身份证件”+“密码”等静态口令形式，如 ATM 机用户必须同时提供银行卡和密码。为了便于记忆，使用者经常采用生日、电话号码等容易被猜测的信息作为密码，或者把密码抄在某处，这样很容易造成密码泄漏，存在着很大的安全隐患。因此必须采用更为先进有效的身份验证技术来代替这种传统的身份验证方法。生物认证技术可利用人的生理特征、行为特征来进行个人身份的鉴定：生物特征可以唯一的标识一个人。生物认证技术被公认是目前最为方便与安全的识别技术。

相比较于传统身份验证方法，生物特征识别的认证具有不被遗忘、防伪性能好、不易伪造或被盗、随身携带和随时随地可用等优点，是密码时代的终结者（盖茨语）。说话人识别是从说话人发出的语音信号中自动提取说话人信息，并对说话人进行识别的研究领域。它是交叉运用心理学、生理学、语音信号处理、模式识别、统计学习理论和人工智能的综合性研究课题^[1]。不同人的指纹不同，与此类似，每个人由于发音器官的生理差异以及后天形成的行为差异，使得他们语音中都带有强烈的个人色彩。这为说话人识别提供了技术可能性。

人的语音可以非常自然的产生，这使得说话人识别具有其他生物验证不具备的优势：

(1) 数据易取性：说话者的语音携带着自身的声纹特征且说话人语音随处可在获取非常方便，使用者的接受程度很高；

(2) 采集成本低廉性：无需特殊的语音采集设备，一个普通的麦克风即可做输入设备；

(3) 远程的可操作性：通过手机、网络等通讯手段可进行用户的远程身份

认证;

(4) 嵌入式系统的可行性: 说话人识别的算法复杂度相对要低;

(5) 身份认证的隐蔽性: 语音随处可见, 说话人识别可以在不知不觉中完成身份认证。

这些优势使得说话人识别的应用越来越受到系统开发者和用户的青睐。据统计, 说话人识别的世界市场占有率为 15.8%^[21], 并有不断上升的趋势, 具有广泛的市场应用前景。

1.2 说话人识别技术概论

在计算机研究领域中, 说话人识别是人工智能方向语音识别的一个分支。它是指通过计算机自动识别测试语音说话人身份的一门技术。

据说说话人识别的历史可追溯到 17 世纪, 在 1660 年查尔斯一世谋杀案中, 人们首次把语音作为推断犯人作案的线索。对说话人识别的研究始于 20 世纪 30 年代, 1937 年 C. A. Linnbdehgr 先生的儿子拐骗事件成为语音说话人个性研究的开端。1945 年, 美国 Bell 实验室的 L. G. Kesta 利用语音频谱图提出“声纹”(voiceprint) 的概念。其后, 随着计算机技术和电子技术的发展, Bell 实验室的 S. Pruzansky 提出了基于模式匹配和概率统计方差分析的说话人识别方法, 引起信号处理领域许多学者的注意, 从而引发说话人识别研究的高潮。1966 年, 美国法院第一次采用此方法进行了取证。

说话人识别按其最终完成的任务可分为两类: 说话人辨认 (Speaker Identification) 和说话人确认 (Speaker Verification), 前者是根据说话人语音确定为 N 个参考说话人中的某一个, 是一个选择问题; 后者是证实说话人的身份与其声明的是否一致, 是一个二选一的判定问题^[21]。话人辨认又可分为“开集”和“闭集”两种。开集是假定待识别的说话人可以在设定的集合内, 也可以在集合外。闭集是假定待识别的说话人一定在集合内。例如, 一个说话人集中注册的说话人的个数为 N , 则在识别时, 说话人辨认需要进行 N 次比较和判决: 测试语音和说话人集合中的每个说话人的参考模型都要进行一次匹配计算。如果是开集的情况, 还要对这 N 个人以外的语音作出拒绝的判别。说话人辨认系统的识别率会随着集合人数的增加而降低。而说话人确认系统只涉及

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts 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.

厦门大学博硕士论文摘要库