

声音识别微移动支付系统

姓名	学号	专业	日期
蔡跃区	3140102240	软件工程	2017.07.15

【摘要】受支付宝Kungfu的启示，我们设想可以设计一款声音识别的微移动支付系统——声付ShengFu。随着指纹认证、人脸识别等技术的发展和应用，声纹作为一项重要的生物特征识别技术，越来越受到关注。相比指纹、人脸等生物特征，声纹有着一些独特的特性，通过研究分析，可以为相关产品在今后引入生物识别技术时提供参考。一种新型的生物识别支付方案使人们仅仅通过声音即可授权基于在线与移动电话的购物支付。

【关键词】手机支付 声纹识别

一、概述

声纹识别(Voiceprint Recognition, VPR)，也称为说话人识别(Speaker Recognition)，有两类，即说话人辨认(Speaker Identification)和说话人确认(Speaker Verification)。前者用以判断某段语音是若干人中的哪一个所说的，是“多选一”问题；而后者用以确认某段语音是否是指定的某个人所说的，是“一对一判别”问题。不同的任务和应用会使用不同的声纹识别技术，如缩小刑侦范围时可能需要辨认技术，而银行交易时则需要确认技术。声纹的唯一性可以作为人体的身份特征，是人体的一张“身份证”，且具有长期稳定的特征信号。声纹识别技术正是利用这一点，将未知人语音与已知人语音分别通过声纹图谱的语音声学特征进行比较和综合分析，以得出两者是否同一的判断过程。

声纹识别广义上分为语音识别和说话人识别两种。语音识别是根据说话人的发音辨认其所说的内容；说话人识别是根据语音来辨认说话人，而并不考虑声音的内容和意义。目前，普遍意义上声纹识别的概念是指说话人识别，即通过语音信号提取代表说话人身份的相关特征(如反映声门开合频率的基频特征、反映口腔大小形状及声道长度的频谱特征等)，进而识别出说话人身份等工作方面的技术。

二、需求分析

1. 总体需求

目前手机支付市场已经初具规模，并且已经基本上被广大手机用户所接受，本文所讨论的语音支付系统，主要目的是解决广大消费用户的使用安全问题，并且在使用说话人识别技术解决安全问题的同时，也要保证用户使用支付软件支付交易过程中的方便简捷性。其次，该系统的UI设计与使用界面也要设计得友好美观，提升用户的服务体验水平。

2. 系统功能需求

主要将需求分成处理前置功能、语音处理功能、支付交易功能、系统管理功能这四个模块。下面简单罗列各模块的功能：

(1) 前置功能

- 接入网关
- 支付网关
- 管理员界面

(2) 语音处理功能模块

- 语音录入
- 语音处理
- 识别结果反馈

(3) 系统管理功能模块

- 商户管理
- 合作银行管理
- 内部管理

(4) 支付处理功能模块

- 受理转发
- 数据下载

三、系统架构设计

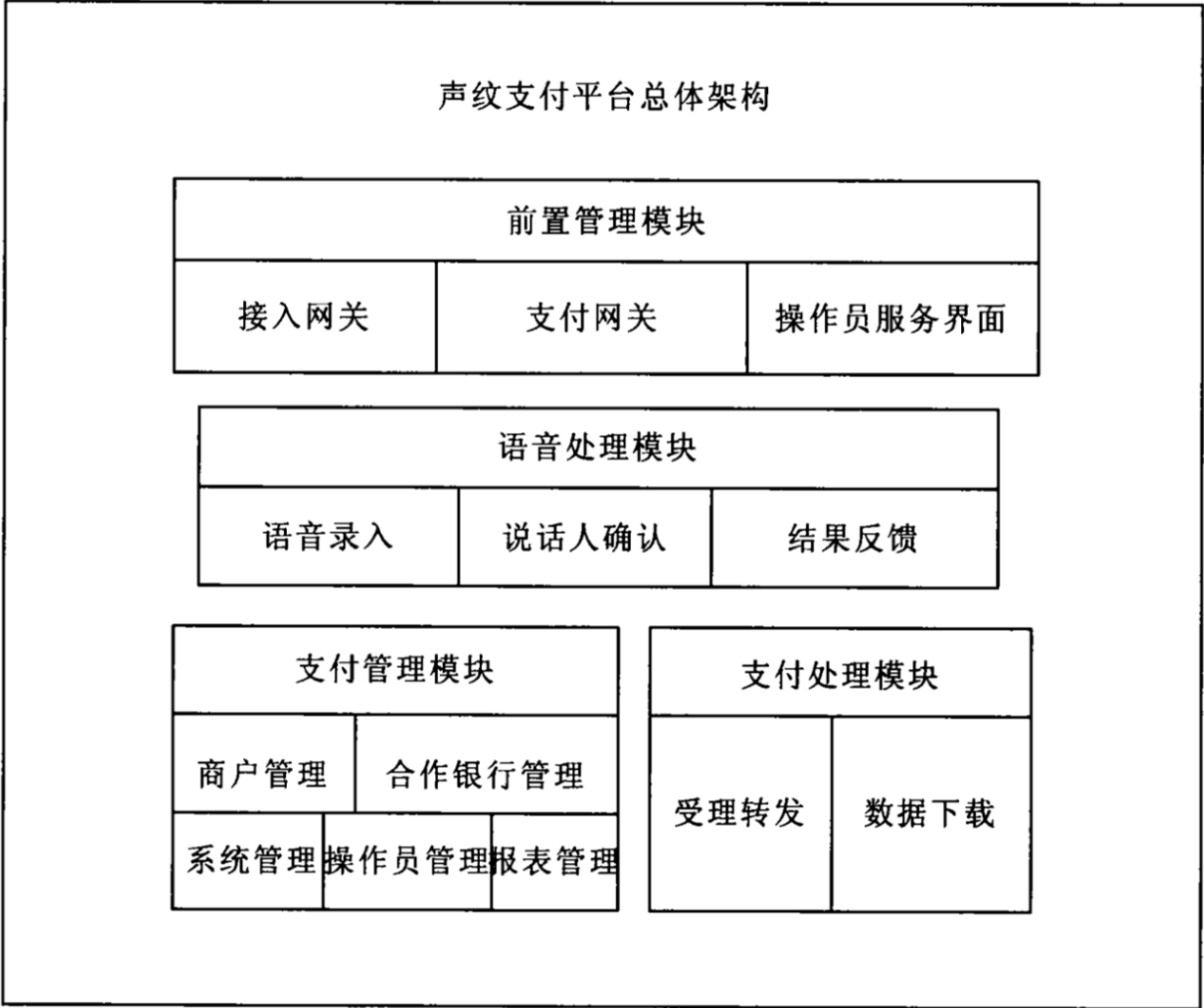


图1 手机语音支付系统架构设计

1. 前置功能模块设计

前置模块功能的作用主要是提供移动手机支付平台的接入与外连通信服务，与此同时，该模块还必须联接到客户端的在线交易平台，以从前端来实现移动手机支付平台的业务前置和数据整合集成接入。除此之外，这个模块还可以进行各种协议，比如:socket机制、hap协议、文件传输等模式的信号接入、输出操作。此模块还能完成协议之间的相互转化。

2. 语音处理功能模块设计

该模块提示用户使用手机自带麦克风录入与注册语音相同内容的语音信息。并将语音信息打包封装成二进制代码，使用TCP协议，通过语音处理接口传送到语音处理服务器。

支付客户端传送过来的语音信息将在这里进行预处理，美尔频标倒谱系数的特征提取，与注册时录入的语音样本进行特征匹配，设定一个阈值。若匹配的结果符合此阈值的范围则返回准许通过信号，反之则返回禁止通过信号。

为防止保证用户的语音具有时效性，系统会将每次录入成功的语音信息作为下次用户进行说话人确认的语音样本。同时为了防止不法分子使用用户录入成功的语音的录音或其它副本进行冒名登录，系统会保存每次用户录入成功的语音信息，并且在进行说话人确认的时候，将当前语音信息与之前录入成功的语音信息进行匹配，若发现匹配结果完全一致的情况，我们则认为当前语音信息为录音或其它副本，从而不准许该用户进行登录或交易等业务办理。

接收处理模块的匹配结果。如果为准许通过信号则准许用户进行付款等相关操作；若接收的为禁止通过信号，则提示用户重新录入语音信息，或提示用户联系管理员等操作。

3. 系统管理功能模块设计

手机支付平台的支付管理模块是平台设计的核心，也是关系到平台能否顺利应用实施的关键。该模块主要包括移动手机支付平台商户管理、手机支付平台合作银行管理、手机支付平台操作员管理、手机支付平台系统内部信息管理、手机支付平台报表管理等设计工作。

4. 支付处理功能模块设计

主要完成移动手机支付交易，还承担着对批量交易的受理与移动支付交易指令的转发任务。

负责从移动手机支付平台获取数据，这些数据主要是与相关手机支付交易业务相关的数据信息。

四、语音处理服务器设计

1. 技术原理

声音识别流程：声纹识别要求从语音信号中提取个体差异，提取出能够反映说话人是谁的信息，从而进行说话人识别。其基本原理是为每一个说话人建立一个能够描述这一说话人个性特征的模型，作为此说话人个性特征的描述，如图2所示。

实现声纹识别技术需要解决三个问题：语音信号的预处理;语音信号的特征提取;测试语音与目标语音匹配的差距计算。

模式匹配：

声纹识别技术的关键在于对各种声学特征参数进行处理，并确定模式匹配方法，主要的模式匹配方法包括：

(1) 模板匹配方法：利用动态时间弯折(DTW)以对准训练和测试特征序列，主要用于固定词组的应用(通常为文本相关任务);

(2) 最近邻方法：训练时保留所有特征矢量，识别时对每个矢量都找到训练矢量中最近的K个，据此进行识别，通常模型存储和相似计算的量都很大；

(3) 神经网络方法：有很多种形式，如多层感知、径向基函数(RBF)等，可以显式训练以区分说话人和其背景说话人，其训练量很大，且模型的可推广性不好；

(4) VQ聚类方法(如LBG)：效果比较好，算法复杂度也不高，和HMM方法配合起来更可以收到更好的效果；

(5)多项式分类器方法：有较高的精度，但模型存储和计算量都比较大；

(6)概率统计方法。声音信息在短时间内较为平稳，通过对稳态特征的统计分析，利用均值、方差等统计量和概率密度函数进行分类判决。这种方法不用对特征参数在时域上进行调整，适合与文本无关的声纹识别。

(7)动态时间规整方法。说话人信息不仅有稳定因素(发声结构和发音习惯)，也有时间因素(语速、语调、重音和韵律)，将识别模板与参考模板进行时间对比，按照某种距离测定得出两模板间的相似度。

(8)矢量量化方法。把每个人的特定文本编成码本，识别时将测试文本按此码本进行编码，以量化产生的失真度作为判决标准，具有识别精度高、判断速度快的特点。

(9)隐马尔可夫模型方法。隐马尔可夫模型是基于转移概率和传输概率的随机模型，它把语音看成由可观察到的符号序列组成的随机过程，符号序列则是发声系统状态序列的输出。识别时，为每个发音个体建立发声模型，通过训练得到状态转移概率矩阵和符号输出概率矩阵。识别时计算未知语音在状态转移过程中的最大概率，根据最大概率对应的模型进行判决。这种方法不需要时间调整，可节约判决时的计算时间和存储量，目前已被广泛应用，缺点是训练时计算量较大。

(10)人工神经网络方法。人工神经网络在某种程度上模拟了生物的感知的特性，是一种分布式并行处理结构的网络模型，具有自组织和自学习能力、很强的复杂分类边界区分能力以及对不完全信息的鲁棒性，其性能近似理想的分类器；缺点是训练时间长，动态时间规整能力弱，网络模型随说话人数目增加可能达到难以训练的程度。

2. 具体实现

处理流程：用户通过智能手机麦克风录入的语音信息，在客户端中是一系列二进制的代码，客户端系统将语音信息进行封装打包为消息头(Head)和消息体(Body)，其中消息头中包含消息的长度(Len)，消息的类型(Style)，手机号(Num)等信息，消息体包含消息的语音信息。使用TCP通信协议传送到服务器端。语音处理服务器端(以下简称服务器)采用LabView的这一功能强大的语音处理软件。

训练模块：服务器收到客户端的语音信息，将利用手机号这一属性在数据库中查找与之相对应账户信息，从而进行说话人确认。主要完成训练阶段语音信号的预加重、分帧、加窗以及端点检测等预处理内容、特征提取以及特征的存储等工作。

确认模块：模式匹配过程主要也是利用MATLAB节点中的相关函数进行处理，利用第二章介绍的高斯混合模型.通用背景模型算法，其中将GMM.UBM的混合数设置为256，核心混合数设为4。将录入的待测语音信息中提取出的特征参数同数据库中提取出的语音信息的特征参数进行逐一的模式匹配。

五、参考文献

[1] 韩超, 梁泉. Android系统原理及开发要点详解[M]. 电子工业出版社, 2010.

- [2] 霍春宝, 张彩娟, 赵红敏. 基于GMM—UBM的说话人确认系统的研究[J]辽宁工业大学学报:自然科学版, 2012, 32(2):98—101.
- [3] 崔艳秋, 张俊星, 李敏等. 基于LabVIEW的说话人确认系统设计[J]. 大连民族学院学报, 2013, 15(5):543—546.
- [4] 黄梅珍. Android NFC手机支付的实现与应用[D]. 华南理工大学, 2012.
- [5] SM. Witt. Use of Speech Recognition in Computer Assisted Language Learning. PhD Thesis, University of Cambridge, 1999.
- [6] 韩赵魏. 第三方支付平台在手机支付中的应用研究[D]. 湖南大学, 2012
- [7] 房安栋, 刘军万. 复杂背景下声纹识别系统的研究方法综述[J]. 电子世界, 2013(3):97—99.