

# 一种基于 LabVIEW 和 MATLAB 的语音识别方法

史剑锋, 常国栋, 李志刚

(中国科学技术大学自动化系, 安徽 合肥 230027)

**【摘要】** 论文介绍了在 LabVIEW 环境中用声卡采集语音信号, 利用 MATLAB script 节点进行 HMM 算法的训练和识别。此方法充分利用了 LabVIEW 灵活的图形编程和 MATLAB 数学计算, 实行优势互补, 提高了编程效率, 便于语音采样参数的控制, 是研究语音识别的一条有效途径。

**【关键词】** LabVIEW; MATLAB; HMM; 采样; 语音识别

**【中图分类号】** TP391.42

**【文献标识码】** A

**【文章编号】** 1009-8054(2007) 07-0057-03

## A Voice-Recognition Method Based on LabVIEW and MATLAB

SHI Jian-feng, CHANG Guo-dong, LI Zhi-gang

(Department of Automation, University of Science and Technology of China, Hefei Anhui 230027, China)

**【Abstract】** This article introduces the gathering pronunciation signal with the sound card in the LabVIEW environment, and the training and the recognition of HMM algorithm by using the pitch point of MATLAB script. This method has made full use of the LabVIEW nimble graph programming and the MATLAB mathematical computation, implemented mutually supplementary in superiority, and thus raised the programming efficiency, is convenient for the control of pronunciation sampling parameter, and is also an effective way for the studies of speech recognition.

**【Keywords】** LabVIEW; MATLAB; HMM; Sampling; Speech recognition

## 0 引言

LabVIEW 采用基于流程图的图形化编程方式, 它编程简单、易于理解、效率高, 针对数据采集、仪器控制、信号分析和数据处理等任务, 设计提供了丰富完善的功能图标。LabVIEW 提供了 MATLAB Script 节点, 可以在节点中编辑 MATLAB 程序, 解决复杂的数值运算, 这为语音识别的研究和测试提供了契机。

收稿日期: 2007-02-05

作者简介: 史剑锋, 1975 年生, 男, 新疆人, 硕士研究生, 研究方向: 语音识别技术; 常国栋, 1952 年生, 男, 合肥人, 硕士生导师, 研究方向: 数据库技术、程序设计、系统仿真等; 李志刚, 1976 年生, 男, 硕士研究生, 研究方向: 传感器技术。

## 1 语音识别主要算法

语音识别系统常用的算法有 ANN(神经网络)、DTW(动态时间归整匹配)、HMM(隐马尔可夫)的识别和训练算法。HMM 算法是目前最成功的一种语音识别模型和算法。常用的特征矢量有 LPCC(线性预测倒谱系数)和 MFCC(美尔频标倒谱系数), 其中 MFCC 参数具有良好的识别性能和抗噪能力, 采用滤波器组方法计算<sup>[1]</sup>:

$$c_n = \sum_{m=1}^M \log x^i(k) \cos\left(\frac{\pi m(k-1/2)}{M}\right), n=1, 2, \dots, L$$

HMM 算法需要解决的三个基本问题<sup>[1]</sup>:

(1) 概率计算: 已知观察序列  $O=(o_1, o_2, \dots, o_T)$  和模型  $\lambda=(A, B)$ , 计算由模型  $\lambda$  生成  $O$  的概率  $P(O|\lambda)$ 。主要采用前向算法和后向算法解决:

$$P(O|\lambda) = \sum_{i=1}^N \sum_{j=1}^N \alpha_i(i) A_{ij} b_j(o_{i+1}) \beta_{i+1}(j) = \sum_{i=1}^N \alpha_i(i) \beta_i(j)$$

(2) 最优状态序列搜索(识别过程): 已知观察序列  $O$

和模型,选择最佳的状态序列 $X$ 。采用Viterbi算法,最佳准则: $\delta_i(i) = \max_{q_1, q_2, \dots, q_{i-1}} p[q_1, q_2, \dots, q_{i-1}, q_i = i, o_1, o_2, \dots, o_i | \lambda]$ 。

(3) 参数估计:对应训练过程。根据观察序列不断修正模型参数( $A, B$ ),使 $P(O| \lambda)$ 最大。主要采用Baum - Welch重估算法。

## 2 声卡的设置

从数据采集的角度来看,PC声卡是一种音频范围内的数据采集卡,它同时具有A/D和D/A转换功能,不仅价格低廉,而且兼容性好、性能稳定、灵活通用。如果对测量对象的频率在音频范围(20Hz~20KHz)内,而且对采样频率等指标没有太高的要求,就可以考虑使用声卡。而语音音频范围一般在3400Hz以内,满足声卡采集的要求。在采集语音信号前,要检查声卡的设置,保证已配置的输入功能(录音功能)不处于静音状态。

LabVIEW中提供了一系列使用Windows底层函数编写的与声卡有关的函数。这些函数集中在Functions All Function Graphics & Sound Sound下。由于使用Windows底层函数直接与声卡驱动程序打交道,因而封装层次低,速度快,而且可以访问、采集缓冲区中任意位置的数据,具有很大的灵活性,能够满足实时不间断采集的需要。

## 3 系统实现

### 3.1 LabVIEW编程<sup>[2]</sup>

在LabVIEW中,利用Windows声卡对语音信号进行采样。采样程序如图1所示,SI Config设置声卡中与数据采集相关的硬件参数,将device设置为0,采样位数为单声道16位,采样频率11025Hz,缓冲区长度为默认值8192字节;SI Start通知声卡开始采集外部数据;SI Read将数据缓冲区中的内容读取到用户程序的数组中,一次可读取缓冲区长度的一半(4096字节);SI Stop通知声卡停止采集外部数据;SI Clear完成最终的清理工作,

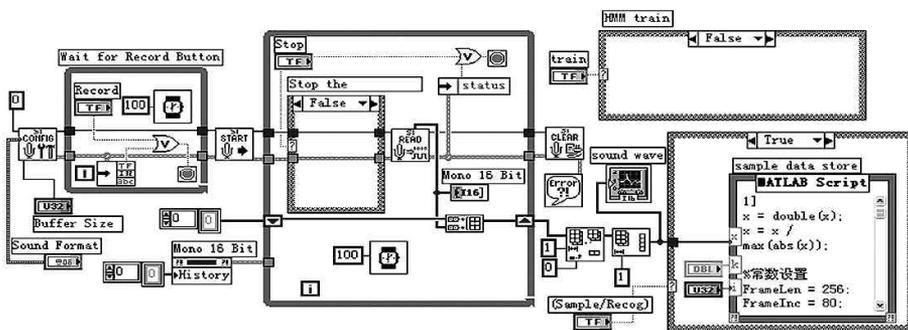


图1 LabVIEW采样语音数据实施语音识别

释放请求的一系列系统资源;While Loop和Case Structure用于采样的开始和结束;循环中利用100ms延时可以降低循环的频度,减少CPU负担;Sound wave显示语音波形;第一个Case结构的True为默认值,执行语音端点检测和mfcc的计算,False执行语音的识别;HMM train分枝结构的False为默认值,True执行对采样样本的HMM模型训练。

### 3.2 MATLAB编程<sup>[3]</sup>

利用MATLAB Script节点,对采样的语音数据进行处理,程序流程如图2所示。

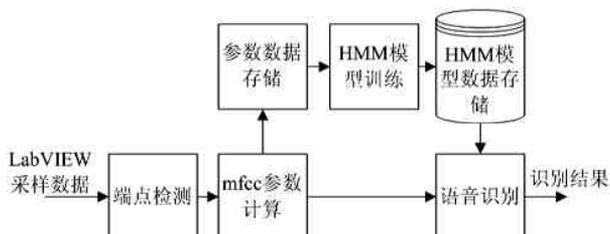


图2 节点内HMM算法程序流程

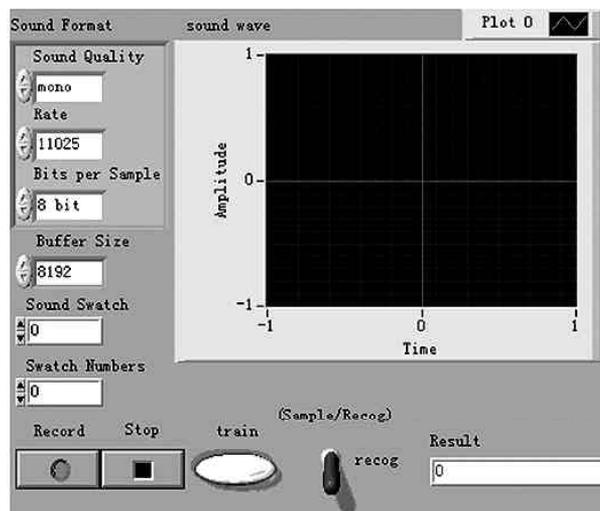


图3 LabVIEW虚拟前面板

采样数据经过帧长256、帧移80的帧分后,通过短时能量和过零率的端点检测,将有效的原语音数据存储于结构数组samples(k).wave中。

利用12阶MFCC参数及其一阶差分作为语音信号的特征参数,存储于结构数组samples(k).data。在HMM模型训练中,先要模型初始化,利用K均值对samples(k).data聚类,然后用Baum-Welch算法进行n次迭代,重估HMM模型

参数,将参数存储于HMM{*i*}的结构数组中。在识别算法中,计算最终概率和最后节点,回溯最佳状态路径。

### 3.3 虚拟前面板

虚拟前面板提供与用户交互的图形界面,如图3所示。通过面板上的各种按钮、开关等控件可以实现对整个系统的操作控制。通过鼠标和键盘即可设定和控制采样频率、缓冲区长度、采样位数、录音、训练等相应的参数。

## 4 结语

利用LabVIEW和MATLAB混合编程,实现了一个简单的语音识别系统,可用于小词汇量的非特定人连续汉

语语音的实时识别。该系统编程方便、计算简单、识别时间短,是研究语音识别的一条有效途径。

### 参考文献

- [1] 韩纪庆,张磊,郑铁然. 语音信号处理[M]. 北京:清华大学出版社,2004,09:2-215.
- [2] 侯国屏,王坤,叶齐鑫. LabVIEW7.1编程与虚拟仪器设计[M]. 北京:清华大学出版社,2005.2.
- [3] 何英,何强. MATLAB扩展编程[M]. 北京:清华大学出版社,2002,02:289-372. 

(上接第56页)

化。这一阶段,利用统一的控制平面实现分布式智能。在GMPLS技术进一步成熟的基础上,特别是NNI信令协议最终实现标准化的前提下,控制平面可扩展到全网范围统一实现,控制智能的实现也更为强大和全面。

### 2.2 现有传送网引入ASON的意义

(1) 超宽带业务、非标准带宽业务和按需带宽业务可以更好地满足通信传送网中日益复杂和繁重的通信保障任务,避免目前网络服务体系中各类繁琐过程,并且随着任务需求的变化,保障服务也会随之发生变化,为用户提供更为方便、快捷、灵活的网络服务。

(2) 自动发现技术可以更好地适应目前通信网中各节点的光纤数量成倍增长现状。自动发现机制处理的数据量大、速度快,可避免手工配置过程中可能发生的错误,并能帮助诊断是否有错误的连接存在,为现有传送网提供更为可靠、安全的网络服务。

(3) 控制平面提供的快速故障恢复能力和多种生存性方式(保护和恢复)可以更好地满足传送网复杂环境的需求,为现有传送网提供更为可靠、高效、灵活的网络服务。

## 3 需要重点考虑的几个问题

——标准化问题。目前ASON标准化工作仍未最后完成,这就需要我们及时跟踪最新进展,在引进产品时注重与国际和国内标准接轨,解决好各厂商设备之间的互联互通问题,并尽量采用一个厂商的设备进行组网。

——光缆资源问题。ASON构建的是Mesh网,需要每个节点做到3个或3个以上的出口路由,因此,应及早完

善现有传送光纤物理网以满足ASON发展的需要,从全网的角度来组织通路。

——与旧网的兼容性。网络性能是网络升级的主要依据,因此在网络规划之前必须对现有网络进行详细的网络性能评估,需要对网络的负载强度,尤其是一些重要节点设备的工作状态有较为确切的了解。

## 4 结语

以ASON为代表的智能化光网络是新一代网络,代表了光通信的发展方向,这种先进的技术和组网思路能给通信网带来的好处是显而易见,但是,它的复杂性和高技术性也是不言而喻的。因此,跟踪ASON最新技术和标准制定发展动态,积极探索其引入策略和方法是摆在我们面前的一项长期而艰巨的任务。

### 参考文献

- [1] 张杰,徐云斌,宋鸿升,等. 自动交换光网络ASON[M]. 第一版,北京:人民邮电出版社,2004. 53.
- [2] 唐雄燕,左鹏. 智能光网络—技术与应用实践[M]. 北京:电子工业出版社,2005. 35.
- [3] 包东智. ASON技术的发展及应关注的热点问题[J]. 通信技术,2004,(05):47.
- [4] 徐云斌. ASON标准化研究进展[J]. 电信网技术,2005,(09):20.
- [5] 刘占霞,杨广强. 现有传送网向ASON演进的基本策略[J]. 电信技术,2006,(08):13.
- [6] 张海懿. ASON标准的进展情况[J]. 电信技术,2006,(08):23. 