

有限词汇、非特定人关键词识别系统

温艳红(兰州大学)、李维(兰州铁道学院)、张子范、张吉庆、王粤(兰州大学)

a Limited Vocabulary、Speaker – Independent Keyword Recognition System

Wen Yanhong Li Wei Zhan Zifan ZhanJiqing Wang Yue
(Lanzhou University) (Lanzhou railway college) (Lanzhou University)

摘要:本文介绍了一个在 386 微机上实现的有限词汇,非特定人关键词识别系统。利用模糊矢量量化,更精确反映输入语音矢量。该系统采用离散的隐含马尔可夫模型(HMM),为不同的语音集团建立多套 HMM 参数,显著地提高了识别率,孤立词识别率可达 97%。利用分段 K 平均程序进行 HMM 参数重估,最终建立精确的额外语音、关键词 HMM 模型。识别关键词利用分层构筑 HMM 算法进行,关键词的识别率可达到 90% 以上。

关键字:特征提取 模糊矢量量化 马尔可夫模型 分段 K 平均程序 层构筑 HMM 算法

一、引言:目前,有限词汇,非特定人语音识别系统已投入应用,它的算法与任务都适用于一套限定的词汇,在使用中也受到这套词汇的约束。而本文所研究的关键词识别对于很多语声响应系统和数据库查询系统而言,是一项非常实用的技术。识别话语中的重要信息—关键词,而去掉一些次要的话语信息,并不妨碍对于用户意愿的了解,人类的听觉认知系统也是这样工作的。在使用中,系统只识别句中的关键词,对其他非关键词,使用者可任意选择不受约束。

二、特征提取:采用性线预测编码算法(LPC)求得 LPC 系数,再由该系数形成倒谱系数和差分倒谱系数。

三、模糊矢量量化:在语音识别中,通过矢量量化,把语音特征空间量化为有限的点集,从而极大地降低了计算量。模糊矢量量化用特征空间多个点来表征语音特征参数在空间中的位置。因此,可以较精确的反映输入矢量在特征空间的位置,本文采用 kohonen 神经网络设计码书。

四、离散马尔可夫模型:根据不同的语音集团,建立多套 HMM 参数,孤立词的识别率可达到 97%。

五、分段 K 平均程序:由于要把关键词从含其他语音的句子中提取出来,因此,除对关键词建立模型外,必须对额外语音建立相应的 HMM 模型。模型参数提取的正确与否,直接关系到关键词识别率的高低。为了建立精确模型,还必须对一句话中关键词及额外语音进行精确分割。因此采用分段 K 平均程序代替人工实现语音之间科学分割,最终建立精确模型。

六、分层构筑 HMM 算法:关键词识别中,利用(level – building HMM)算法识别句中的关键词。关键词识别率达到 90%。

七、实验:词汇表为“1、2、5、开、关”五个关键词,录音为 30 个人说含关键词的句子,条件是在安静的实验室进行,内容是涉及电视选台及开关的。额外语音就为除该关键词以外的内容。为了节约运算开销,把额外语音模型的数量逐步减少,发现识别率并未剧烈减小,而识别时间大大减小了,最后用一个额外语音模型,仍能获得较好的识别效果。