

说话人识别中 LPCCEP 倒谱分量的相对重要性

张 华, 裘雪红

(西安电子科技大学 计算机学院, 陕西 西安 710071)

摘 要:文中研究表明,反映说话人特征信息的特征参数矢量的各个分量通常具有不同的分布,对正确识别说话人身份的有效性是有差别的。文中通过增减分量的方法对 LPCCEP 的各维分量进行分析,得到一个关于 LPCCEP 参数各维分量的平均贡献序列,将此序列运用于 WDMVQ 有助于系统识别率的提高。

关键词:LPCCEP;说话人识别;欧式距离;WDMVQ

中图分类号:TP391

文献标识码:A

文章编号:1005-3751(2006)04-0067-02

On the Importance of Components of LPCCEP
in Speaker Recognition

ZHANG Hua, QIU Xue-hong

(College of Computer, Xidian University, Xi'an 710071, China)

Abstract: This article indicates that each component of the speaker's characteristic parameter vector usually has the different distribution, and has different validity. This article analysis each component of LPCCEP parameter, find out an average contribution sequence, and the WDMVQ system that utilizes this sequence has recognition rate enhancement.

Key words: LPCCEP; speaker recognition; Euclidean distance; WDMVQ

0 引言

说话人识别系统中特征提取过程就是抽取保持语音最重要特征,并消除与语音无关信号的干扰,其性能对识别系统的性能有直接影响^[1]。寻找具有良好性能的特征及其提取算法是提高识别系统性能的根本途径之一。

LPC是语音信号分析的重要方法之一,考察多种 LPC 语音特征量, LPC 倒谱系数(LPCCEP, 简称为 LPCC)对于说话人识别的效果最好^[2,3]。语音信号的倒谱可以反映声道响应的状况,如果能对 LPCC 的每一维特征分量的识别能力作定量化分析,将是十分有意义的。

文中采用增减特征分量的方法对 16 阶 LPCC 的每一维特征份量做了定量分析,以此为依据,对 LPCC 各维特征分量对说话人识别性能的影响进行了评价。由于文中的目的是比较 LPCC 各分量对识别的贡献,而不是刻意追求最高的识别率,因此文中采用传统 VQ 识别算法^[2,4]。

1 LPCCEP 特征参数

通过高效线性预测系数 LPC 求解算法,可得到语音信号声道响应特性的谱估计,对应的倒谱特征系数可由

LPC 推出,即 LPCC。其具体推导为:

$$\begin{cases} c_1 = a_1 \\ c_n = \sum_{k=1}^{n-1} (1 - \frac{k}{n}) a_k c_{n-k} + a_n, \quad (1 < n \leq p) \end{cases} \quad (1)$$

式中,线性预测系数 a_n ($1 < n \leq p$) 可从数字语音信号 S_n 中利用高效的 LPC 算法求得,再用式(1)求得对应倒谱系数 c_n ($1 < n \leq p$)。

2 倒谱分量相对重要性的评价方法

倒谱分量相对重要性评价方法有两种方法:一是通过定义各分量的 F 比可以得到各特征分量的区分能力;二是通过增减分量的方法考察每个特征分量的贡献^[2]。文中采用第二种方法,每个 LPCCEP 倒谱分量的平均贡献 $R(i)$ 由下式计算^[5]:

$$R(i) = \frac{1}{n} \left(\sum_{j>i} (p(i, j) - p(i+1, j)) + \sum_{j<i} (p(j, i) - p(j, i-1)) \right) \quad (2)$$

式中, n 是倒谱阶数, $p(i, j)$ 是以 i 阶到 j 阶倒谱系数为特征的识别率。

图 1 是式(2)的图示说明,比如,以第 0 至 2 阶倒谱系数为特征的识别率 $p(0, 2)$ 减去以第 0 至 1 阶倒谱系数为特征的识别率 $p(0, 1)$, 就得到在以 $C_0 \sim C_1$ 为特征时倒谱分量 C_2 对识别的贡献。将在所有可能的顺序组合的 C_2 贡献平均就得到文中定义的 C_2 的平均贡献(相对重要

收稿日期:2005-08-05

作者简介:张 华(1982-),女,四川人,硕士研究生,研究方向为说话人识别;裘雪红,教授,硕士研究生导师,研究方向为多媒体信息处理。

性)。正值的平均贡献 $R(i)$ 表明由于添加该特征往往会
使识别率增加,负的平均贡献 $R(i)$ 则相反。由于文中实
验仅顺序添加或舍弃特征分量,因此平均贡献 $R(i)$ 仅表
示该分量的相对重要性,而不表示各分量之间的相互依赖
关系。

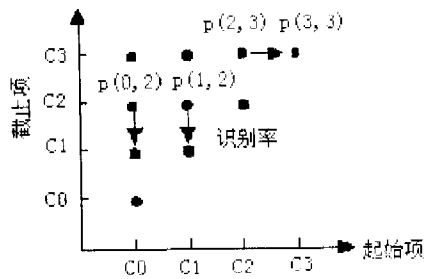


图 1 倒谱分量平均贡献评价方法

3 实验所用语音数据库

文中所用语音库采用自建数据库。数据的录音环境
为普通机房,录音软件为 CoolEdit2. 1,采样频率为 11.
025kHz,量化精度 16bit。

数据库包含 10 男 10 女,样本内容为 0~9 共 10 个
数字的连续发音,一共发 8 遍。识别时,取其中一次录音作
为训练样本,剩下的 7 次录音作为测试样本。

4 LPCC 各分量的相对重要性

表 1 是纯净语音条件下 LPCC 各分量顺序组合的
说话人识别率,行方向为起始 LPCC 分量,列方向为截止
LPCC 分量。由表 1 按式(2)计算出各倒谱分量的平均
贡献,如图 2 所示。图中填充块的高低表示识别时包含该
LPCC 特征分量而增加的平均识别率,横坐标为 LPCC 系
数序号。

正值的平均贡献表明添加该特征往往会使识别率
增加,数值越大说明对于提高识别率的贡献越大。相反,
负值的平均贡献表明添加该特征会使识别率降低。从图 2
中可以看到,16 阶的 LPCC 特征矢量中每一分量的平均

贡献值均为正值,但各个分量的平均贡献是各不相同的。
其中 C10 的平均贡献值最大,为 11.08。按平均贡献的大
小为各维分量做一个排序,将得到一个平均贡献序列:
|C10,C2,C6,C14,C8,C9,C4,C16,C12,C15,C5,C1,C13,
C3,C11,C7|。

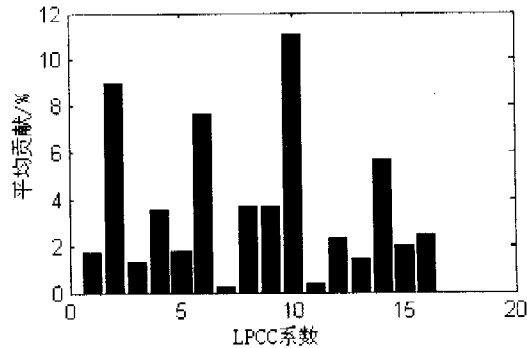


图 2 纯净语音下 LPCC 各分量的平均贡献

5 结 论

在 VQ 定义的欧式距离失真测度下。LPCC 各分量
对于说话人识别的平均贡献是各不相同的。分布也没有
很明显的规律,其中平均贡献最大的是分量 C10。分析实
验得到的平均贡献序列,大部分偶阶分量的平均贡献要高
于奇阶分量的平均贡献。

传统 VQ 定义下的欧式距离对于特征参数的每一分
量所赋予的权值都是相同的。根据 WDMVQ(Weight Dis-
tortion Measure VQ)的思想,不同码字对说话人识别系统
有不同的识别能力,识别能力的大小用权值来衡量,大的
权值有大的识别能力,反之亦然。

平均贡献大的码字识别能力大,则平均贡献大的码字
应得到大的权值。基于这一思想,本文尝试赋予不同平均
贡献的码字以不同的权值,其中权值仅是按照平均贡献大
小的序列作从 1 到 0 的简单递减。此时的识别率由原来
的 91.91% 提高到了 92.56%。这一结果表明将实验中所
得到的平均贡献序列作为 WDMVQ 权值分配依据对于识
别率的提高是有明显帮助的。

表 1 纯净语音条件下,LPCC 相邻分量顺序组合的说话人识别率

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
C1	60.29	84.56	86.03	88.97	89.71	91.18	91.18	91.18	91.18	91.91	91.91	91.91	91.91	91.91	91.91	91.91
C2		78.68	81.62	83.09	85.29	88.97	88.97	90.44	91.18	91.91	91.91	91.18	91.18	91.91	91.91	91.91
C3			55.15	61.76	63.97	80.88	80.88	81.62	82.35	88.97	88.97	88.23	88.97	88.24	89.71	90.44
C4				59.56	63.97	78.68	78.68	79.41	80.88	88.24	88.24	87.5	87.5	88.97	88.97	89.71
C5					54.41	66.91	66.91	75.74	78.68	86.76	86.76	87.5	87.5	87.5	87.5	88.24
C6						59.56	60.29	71.32	77.94	86.76	86.76	88.24	87.5	87.5	87.5	88.24
C7							40.44	61.76	66.18	81.62	81.62	81.62	81.62	84.56	84.56	84.56
C8								58.09	65.44	81.62	81.62	81.62	82.35	84.56	84.56	84.56
C9									57.35	76.47	77.94	81.62	82.35	85.29	84.56	85.29
C10										66.91	72.06	76.47	78.68	81.62	81.62	81.62
C11											48.53	60.29	63.97	67.65	68.38	69.12
C12												60.29	63.97	67.65	68.38	69.12
C13													57.35	63.24	63.24	67.65
C14														59.56	59.56	63.24
C15															16.92	42.65
C16																11.76

如说一个起草文件的流程,可能的流程是:草拟文件 -> 项目组长意见 -> 项目经理审核 -> 部门经理审核 -> 总经理审核,在总经理要求重新起草文件时,他只需选择将文件回退至起草员,而不需要一步一步地依次按原审核路线退回。

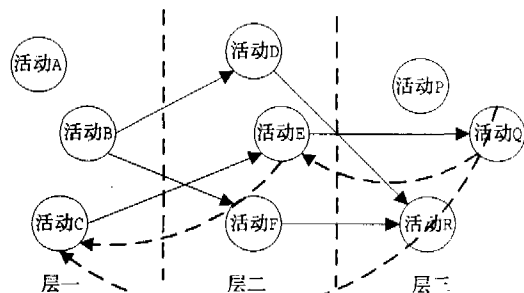


图 3 流程回退图例

1.2.2 模型的优缺点

●该模型具有以下优点:

(1)业务流程描述简单、清晰,可以通过建模更好地理解企业的业务需求。

(2)过程描述具有高度的可配置性,易维护,扩展性强,可以很好地适应中小企业业务流程多变性的特点。

(3)为流程定义中的一些难以描述、难以建模的过程给出了一个行之有效的解决方案。增强了流程定义的使用范围。

●当然也由一些缺点,具体表现在:

(1)用户角色等组织结构只是作为活动节点的属性出现,弱化了流程角色间的相互作用关系。

(2)将子流程节点当作一个完整的流程来定义,忽略了两者的区别。对子流程节点的定义有待改进。

2 有向图 workflow 模型的应用

下面以笔者参与的电子档案管理系统中电子文件的申请浏览流程为例,来说明基于活动节点和有向连接弧的有向图模型在实际中的应用。

●电子文件浏览申请流程有以下 4 个步骤:

(1)项目设计人员填写申请表,提出申请某一电子文件的查阅权限。

(2)项目审批人给出审批意见,审批同意则该申请单继续提交,否则返回(1)并附有审批意见。

(上接第 68 页)

文中 LPCC 分量的相对重要性是通过相邻若干分量的组合得到的,不能推广到不相邻分量组合的情况。对于 LPCC 的各个分量做出了定量的分析,得到了一个平均贡献序列,此序列可以作为 WDMVQ 权值分配的依据。

参考文献:

[1] Rabiner L, Juang B H. Fundamental of Speech Recognition [M]. New York: Prentice Hall, 1993.

(3)项目经理给出审批,过程同(2)。

(4)部门主管审批,审批同意则流程流转结束,赋予申请用户浏览电子文件的权限,否则退回。

●按照上面的步骤进行流程定义,分以下几步:

1)建立流程活动节点。

这里建立 4 个活动节点,名称依次为:项目设计人员申请单,项目审批人审批,项目经理审批,项目主管审批。

2)建立有向连接弧。

按照实际的流程流过程,建立相应的有向连接弧,这里的流转条件也比较简单,仅仅是同意或者不同意。有一点需要注意,就是在项目经理审批时,可能会出现联合审批的情况,这时还要定义存储审批意见的全局变量。

3)定义用户的权限。

4)将用户的权限加到流程的活动节点中。

这样,一个流程定义就基本上定义完成了。当然这只是针对于相对简单的流程而言,稍微复杂的流程还需要考虑到诸如节点资源库和流程执行的一系列的定义过程,还有子流程的一些流转处理问题等等。

3 结束语

通过对基于有向图的建模方式的研究,提出了一些改进,并提出了基于活动节点和有向连接弧的有向图建模的方法。该过程建模对诸如联合审批和流程回退机制都有很好的支持。并很好地应用于项目开发中,取得了一定的效果。工作流建模是工作流管理系统中非常重要的一步。模型建立的好坏直接关系到工作流管理系统的实施的成败,因此还需要做更深层次的研究。

参考文献:

[1] Workflow Management Coalition. Workflow Standard - Interoperability Abstract Specification. Version 1.0[S]. 1996.
[2] Workflow Management Coalition. The Workflow Reference Model. Version 1.1[Z]. 1995.
[3] 罗海滨,范玉顺,吴澄.一种面向企业用户的工作流模型[J].计算机集成制造系统-CIMS,2000,6(3):55-61.
[4] 金字,周良,丁秋林.基于活动网络图的工作流建模工具的设计与实现[J].航空计算技术,2003,33(1):48-51.
[5] 谈贤芳,刘鲁源.一种基于 XML 的工作流过程模型的实现[J].控制工程,2003,10(3):198-200.

[2] 蔡莲红,黄德智,蔡锐.现代语音技术基础与应用[M].北京:清华大学出版社,2003.

[3] 杨行峻,迟惠生.数字语音信号处理[M].北京:电子工业出版社,1995.

[4] Knedera N, Arai T, Hermansky H, et al. On the Importance of Various Modulation Frequencies for Speech Recognition[A]. In: Proceedings of EUROSPEECH[C]. Rodos, Greece: [s. n.], 1997.

[5] 王吉林.利用矢量量化的说话人识别系统的研究[J].安徽工业大学学报,2005,22(3):282-285.