

基于 LPC 距离和变窗长的端点检测

龚 雪 沈天飞 王浩南

(上海大学机电工程与自动化学院 上海 200072)

摘要: 基于 LPC 距离的端点检测首先会对信号进行分帧处理,然后根据 LPC 距离设置阈值依次判断每一帧是否为语音的端点。而在对信号进行分帧时会有以下问题:如果采用较大的窗长,则总帧数会变少,但是这样端点检测的误差也会增加,尤其是对于语音起点的检测。而如果帧长选取过小,则计算量增加,识别速度也会增加。针对这个问题,将变窗长思想和基于 LPC 距离端点检测算法结合起来,提出了一种改进的端点检测算法。该方法在分帧时不采用统一的窗长,而是在静音段采用大窗长快速向过渡段逼近,进入过渡段后采用小窗长慢慢寻找语音信号的起点,其他阶段则采用常规窗长。研究表明,所提出的改进的端点检测方法检测性能要优于其他两种端点检测方法。

关键词: LPC 距离;端点检测;变窗长

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 510.40

Endpoint detection based on LPC distance and variable window length

Gong Xue Shen Tianfei Wang Haonan

(School of Mechatronics Engineering and Automation, Shanghai University, Shanghai 200072, China)

Abstract: Endpoint detection based on LPC distance will first frame the signal processing, and then according to the LPC distance set threshold to determine whether each frame is the endpoint of the voice. However, it will have the following questions in the signal frame: if a larger window length is used, the total number of frames will be reduced, but the error in the endpoint detection will also increase, especially for the detection of the origin of the speech. If the frame length selection is too small, the calculation amount and the recognition speed are increased. Aiming at this problem, this paper proposes an improved endpoint detection algorithm combining the variable window length idea with the LPC distance endpoint detection algorithm. The method does not use the uniform window length in the frame, but in the quiet section uses a large window length quickly approaching the transition section. Then after the transition, the small window length is used slowly to look for the starting point of the voice signal, the other stage is the use of conventional windows long. Simulation results show that the improved endpoint detection method proposed in this paper is superior to the other two endpoint detection methods.

Keywords: LPC distance; endpoint detection; variable window length

0 引 言

在对语音信号进行训练和识别之前,首先需要检测出语音段的起点和终点,即端点检测。端点检测也是语音识别预处理操作中的重要一环,语音识别中一半以上的错误率都源于端点检测^[1]。不正确的端点检测会降低语音识别率,还会增加计算量^[2-3]。因此,对于端点检测的研究是很有必要的。

双门限端点检测是一种比较传统的检测语音起止点的方法,通过对语音信号的短时平均能量和过零率分别设定两个门限,利用短时平均能量检测语音浊音段,利用短时过零率检测清音段。

通过实验发现,当信噪比为 10 dB 时,已经无法有效检测出语音信号的端点^[4]。这是因为过零率对噪声干扰比较敏感,在低信噪比情况下,短时过零率会首先失效。这是因为,短时平均幅度波形看起来尚可,但由于噪声分量的加强,非语音段的能量也明显变高。进一步实验表明,当信噪比低于 0 dB 时,整个含噪语音信号波形类似于白噪声,此时短时平均幅度也会彻底失效,双门限端点检测将不再适用。因此,如何寻找一个应对加性噪声信号干扰鲁棒性更强的特征参数变得尤为重要。基于 LPC 距离的端点检测法对加性噪声不太敏感,并能较好地实现在低信噪比下的端点检测^[5]。

由实验结果可以看出,基于 LPC 距离的端点检测方法

能够检测出双门限法无法检测到的信噪比为 10 dB 的含噪语音信号的起止点。这说明基于 LPC 距离的端点检测方法比双门限端点检测方法具有更好的鲁棒性^[6]。因此本文将采用基于 LPC 距离的端点检测算法。但是,在采用基于 LPC 距离的端点检测算法进行端点检测时,其对语音起始点的检测并不是十分理想,而语音起始点检测的准确性对于识别率是至关重要的^[7-9]。基于 LPC 距离的端点检测首先会对信号进行分帧处理,然后根据 LPC 距离设置阈值依次判断每一帧是否为语音的端点。而在对信号进行分帧时会有以下问题:如果采用较大的窗长,则总帧数会变少,这样做固然会降低计算量,但是这样端点检测的误差也会增加,尤其是对于语音起点的检测。而如果帧长选取过小,则计算量增加,识别速度也会增加。针对这个问题,设想能否在分帧时不采用统一的窗长,而是在静音段采用大窗长快速向过渡段逼近,进入过渡段后采用小窗长慢慢寻找语音信号的起点,其他阶段则采用常规窗长。本文对该方法的实现进行了 MATLAB 仿真研究。实验结果表明该方法不仅可以减少计算量,而且在过渡段使用较小的窗长,能比较精确地检测出端点。

1 端点检测算法的改进

实验发现,在采用基于 LPC 距离的端点检测算法进行端点检测时,其对语音起始点的检测并不是十分理想,而语音起始点检测的准确性对于识别率是至关重要的。基于 LPC 距离的端点检测首先会对信号进行分帧处理,然后根据 LPC 距离设置阈值依次判断每一帧是否为语音的端点^[10]。而在对信号进行分帧时会有以下问题:如采用较大的窗长,则总帧数会变少,这样做固然会降低计算量,但是这样端点检测的误差也会增加,尤其是对于语音起点的检测。而如果帧长选取过小,则计算量增加,识别速度也会增加。针对这个问题,设想能否在分帧时不采用统一的窗长,而是在静音段采用大窗长快速向过渡段逼近,进入过渡段后采用小窗长慢慢寻找语音信号的起点,其他阶段则采用常规窗长。本文将这种变窗长思想和基于 LPC 距离端点检测算法结合起来,提出了一种改进的端点检测算法。各状态下窗长的划分如表 1 所示。

表 1 不同窗长(窗移)参数大小

	窗型	常规窗	大窗	小窗
参数	窗长	1 320	2 640	630
	窗移	440	880	310

该算法的步骤和基于 LPC 距离的端点检测算法类似,唯一不同的是在第二步分帧时采用动态分帧,分帧过程如下:

1) 对整个信号进行分帧(大窗),向后探测,直到到达过渡段;

2) 对信号剩余部分进行分帧(小窗),向后探测,直到到达语音段;

3) 对信号剩余部分进行分帧(常规窗)。

图 1 更加形象地显示了整个分帧过程。可以看出,在静音段使用大窗快速向语音起点逼近;当 LPC 距离超过低门限时,即进入过渡段时使用小窗更加精确地寻找语音真正的起点;当 LPC 距离超过高门限时,即进入语音段时使用常规窗长即可。

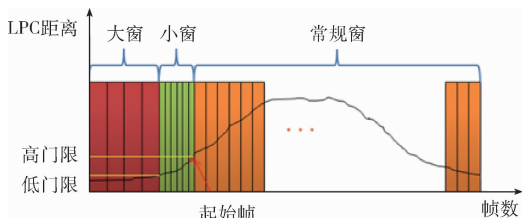


图 1 分帧过程示意图

图 2 所示为信噪比为 50 dB 下改进算法的端点检测效果。对比原基于 LPC 距离的端点检测算法效果图可以看出,本改进算法对于语音信号的起点检测更为准确。仔细观察可以发现相较于改进前的算法,语音段的 LPC 距离并未发生特别大的变化,但在过渡段本改进算法更能准确地反映语音信号。另外,总帧数的减少对于降低运算量也有帮助。

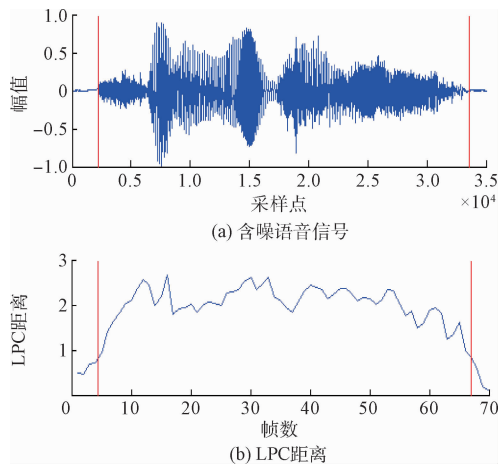


图 2 信噪比为 50 dB 时的检测效果

2 实验与分析

本文的研究目标应用对象是智能轮椅的声控系统,以本课题组开发的基于矢量量化语音识别系统作为应用平台,将双门限、基于 LPC 距离和变窗长的 3 种端点检测算法分别应用到识别系统中对整个系统进行测试实验,给出实验结果并比较 3 者的检测性能。

本文的语音信号为实验室安静环境下 10 个成年男性和 10 个成年女性分别录制的 9 个词条:“前进”、“后退”、“左转”、“右转”、“加速”、“减速”、“打开”、“关闭”、“停止”,共 180

个原始语音信号,声卡采用PC自带的集成声卡,采样率采用44.1 kHz,量化位数为16 bits。噪声信号采用MATLAB生成的高斯白噪声信号。将原始语音信号与高斯白噪声按一定信噪比混合形成含噪语音信号。信噪比分别设定为:0、5、10、15、20、25、30、35、40、45、50 dB。比较在各端点检测算法在不同信噪比下检测后语音识别率,如图3所示。

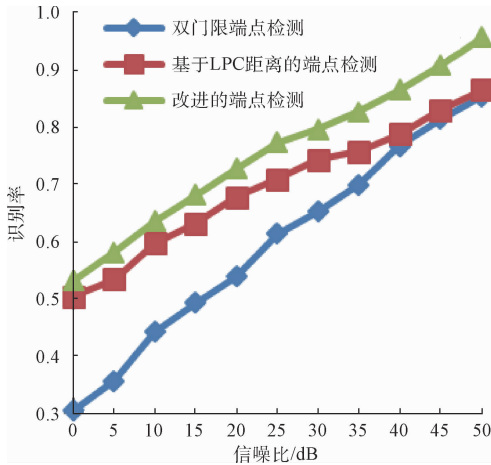


图3 3种端点检测方法性能比较

从图3可以看出:

1) 在低信噪比情况下,双门限端点检测效果要远低于其他两种方法,这主要是因为短时平均幅度和短时过零率抗干扰能力弱,鲁棒性差导致的^[11-12]。而LPC距离的抗干扰能力要高于以上两者,所以在低信噪比情况下基于LPC距离的端点检测方法和改进的端点检测方法性能要优于双门限端点检测^[13]。而在高信噪比情况下,双门限也能够有较好的检测效果,因此其检测后语音识别率不输基于LPC距离的端点检测方法。

2) 改进的端点检测方法的性能整体上要优于基于LPC距离的端点检测方法。随着信噪比的增强,这种差距逐渐加大^[14-15]。由于汉语语言的特点,汉语语音均以元音结尾,因此对于终止点的检测多一点或少一点只是“拖音”的长短而已,并不会对语音语义产生大的影响。改进的端点检测方法能够更加有效的检测语音的起始点,尤其在信噪比较高的情况下。

3 结 论

基于LPC距离端点检测方法,本文提出了一种窗长可变的端点检测方法,该方法在对语音信号分帧时不是采用固定的窗长,而是动态的改变窗长,该方法不仅可以减少计算量,而且在过渡段使用较小的窗长,能比较精确的检测出端点。通过对3种端点检测方法进行实验,发现本文提出的改进的端点检测方法检测性能要优于其他两种端点检测方法。

参考文献

[1] 路青起,白燕燕.基于双门限两级判决的语音端点检

测方法[J].电子科技,2012,25(1):13-15,19.

- [2] 汤霖,姜世芬.多类噪声环境下的语音端点检测[J].计算机工程与应用,2012,48(29):115-116.
- [3] 邓艳容,景新幸,杨海燕.语音端点检测研究[J].计算机系统应用,2012,21(6):240-242.
- [4] 纪振发,杨晖,李然.基于短时自相关及过零率的语音端点检测算法[J].电子科技,2016,29(9):52.
- [5] 张珠瑾,王剑宇.计算机光盘软件与应用[J].计算机光盘软件与应用,2014,26(3):148-149.
- [6] 孙震,张江鑫.关于线性预测滤波器阶数的分析研究[J].杭州电子科技大学学报,2010,30(5):153-156.
- [7] 翁磊,沈天飞,张贺.应用于智能轮椅控制的头部姿态识别[J].电子测量技术,2013,36(12):45-48,59.
- [8] 孙林慧,杨震,季云云,等.基于过完备线性预测字典的压缩感知语音重构[J].仪器仪表学报,2012,33(4):743-749.
- [9] 黎林,朱军.基于小波分析与神经网络的语音端点检测研究[J].电子测量与仪器学报,2013,27(6):528-534.
- [10] 孙科伟,李建海,杨海东.一种强噪声环境中的语音端点检测算法[J].计算机技术与应用,2013,39(12):136-137.
- [11] 汪清泉,黄明红.语音识别的软件实现[J].大众科技,2010,8(13):13-15.
- [12] 刘钰,马艳丽,董培培.语音识别技术概述[J].计算机光盘软件与应用,2010,7(5):98-99.
- [13] 韩芳,靳宗信.低信噪比下的端点检测算法研究[J].西北师范大学学报,2016,52(5):55-58.
- [14] 鲁远耀,周妮,肖珂.强噪声环境下改进的语音端点检测算法[J].计算机应用,2014,34(5):1386-1390.
- [15] 罗元,黄璜,张毅.一种新的语音端点检测方法及其在智能轮椅人机交互中的应用[J].重庆邮电大学学报,2011,23(4):488-490.

作者简介

龚雪,在读硕士研究生,主要研究方向为信号处理、语音处理。

E-mail: 18862611583@163.com

沈天飞,副教授,主要研究方向为先进信号处理技术及其应用,包括小波变换、神经网络等理论在流量计、人脸姿态辨识、语音信号处理中的应用。

E-mail: tfshen@shu.edu.cn

王浩南,在读硕士研究生,主要研究方向为信号处理、图像处理。

E-mail: 478586630@qq.com