

文章编号: 1000-5013(2009)01-0117-02

一种改进的语音识别抗噪算法

刘菁华

(华侨大学 信息科学与工程学院, 福建 泉州 362021)

摘要: 为了提高在噪声环境下的语音识别性能,提出一种融合信号级和特征参数级抗噪的抗噪算法.该算法首先对带噪语音用最小均方误差估计法进行语音增强,后端对原始的带噪语音运用自相关法,以有效抑制加性和卷积噪声.实验结果表明,该算法能有效提高系统在噪声环境下,特别是低信噪比情况下的识别率.

关键词: 语音识别; 抗噪算法; 语音增强; 信噪比

中图分类号: TN 912.3

文献标识码: A

1 常见的抗噪算法

1.1 信号级抗噪算法

信号级抗噪方法主要是指,从带噪语音信号中提取尽可能纯净的原始语音,亦即语音增强.现在语音识别系统中运用得比较多、抗噪效果较好且易于实现的方法,有谱减法^[1]和最小均方误差(MMSE)估计方法等.谱减法是最大似然估计,它完全放弃了对语音频谱的分布假设;而短时谱幅度的MMSE估计做为一种最小均方误差估计,具有对特定的失真准则和后验概率不敏感的性质,因此具有更好的增强效果.这两种方法都属于语音增强,在滤除噪声的同时会给语音信号带来失真,对识别率仍有一定的影响.

1.2 特征参数级抗噪算法

目前,语音识别系统使用较多的特征参数为倒频谱类参数.Viiki等^[2]提出的将Mel的倒谱参数进行归一化,较好地解决卷积噪声问题;You等^[3]提出的基于自相关序列的Mel频率倒谱系数(RAS-MFCC)参数,在不需要噪声先验知识,避免大量计算的情况下,能较好地消除系统的加性噪声.RAS-MFCC比MFCC具有更好的鲁棒性,特别是在低信噪比的情况下,能较大地提高系统的识别率.

1.3 模型级抗噪算法

目前较好的模型级抗噪方法是平行模型结合处理法(PMC)^[4-5].PMC是将噪声和纯净语音的隐Markov模型(HMM)参数经过倒谱运算,在对数谱域联合,然后再还原到倒谱域得到带噪语音的HMM参数.

2 改进的抗噪算法

首先利用最小均方误差估计法对带噪语音进行适当的语音增强,提高语音信号的信噪比,并对增强后的语音进行端点检测,得到语音端点信息.然后,运用自相关法对原始的带噪语音的Mel倒谱特征参数进行特征参数变换,将新得到的特征参数和前面得到的端点信息作为测试模板,与参考模板进行语音识别.该方法使得信号级抗噪算法和特征参数抗噪算法起到相辅相成的作用.语音增强在前端有效抑制背景噪声,而由它带来的乘性语音失真及加性剩余噪声对识别的影响在后续补偿中得到减轻.

文[6]采用PMC基于模型的噪声补偿算法,作为语音增强的后续补偿,对剩余噪声通过调整模型参数互来减少剩余噪声造成的失配来进行补偿.在此基础上,语音增强算法采用最小均方误差估计法,在提高信噪比方面性能更优.在后端采用对特征参数的变换方法,以便同时降低乘性语音失真及加

收稿日期: 2007-12-17

作者简介: 刘菁华(1980-),女,助教,主要从事语音信号处理的研究. E-mail: babysiyu@hqu.edu.cn.

性剩余噪声对系统识别效果的影响.

实验所用系统为基于 DTW 模型,采用特定人、孤立词语音识别系统.参考模板为 25 个短语,采用加性的高斯白噪声.语音和噪声信号经过 4.41 kHz 采样,16 bit 量化为数字信号,并在计算机中按一定比例混合生成不同信噪比(R_{SN})的带噪语音,帧长为 25 ms,帧移为 10 ms.比较单独使用语音增强 MMSE 法和特征参数变换 RAS-MFCC 法,以及联合使用以上两种方法的识别率(),如表 1 所示.

由表 1 可看出,在信噪比较高时,语音增强法也可以达到较好的鲁棒性,但在信噪比较低时,语音增强法的剩余噪声很严重,导致系统识别率较低.单独使用特征参数变换法降噪,在各种信噪比的情况下,系统的识别率都不是太高.只有联合算法能够显著提高系统的识别率,特别是在低信噪比时的识别精度.由此可见,联合算法整体优于单独使用任何一种算法.

表 1 几种抗噪算法的识别率(单位: %)

Tab. 1 The recognition rates of several antinoise algorithms(unit: %)

方法	R_{SN} / dB					
	- 5	0	5	10	15	20
MMSE	0.18	0.36	0.54	0.87	0.95	0.98
RAS-MFCC	0.23	0.68	0.71	0.76	0.80	0.86
联合算法	0.62	0.80	0.92	0.95	0.98	0.99

3 结束语

本文提出一种融合信号级抗噪、特征参数抗噪法及识别模型抗噪法的各自优点的抗噪方法.在识别过程的各个阶段分别提高识别的性能,可以使得整个系统的性能得到进一步的提高.

参考文献:

- [1] DONOHO D L. De-noising by soft-thresholding[J]. IEEE Trans on Information Theory, 1995, 41(3): 613-627.
- [2] VILIKI O, BYE D, LA YRU KA K. A recursive feature vector normalization approach for robust speech recognition in noise[C]. Proceedings of ICASSP 98. Seattle: IEEE Intestinal Conference on Acoustics, Speech and Signal Processing, 1998: 733-736.
- [3] YOU K H, WANG H C. Robust features derived from temporal trajectory filtering for speech recognition under the corruption of additive and convolutional noises[C]. Proceedings of ICASSP 98. Seattle: IEEE Intestinal Conference on Acoustics, Speech and Signal Processing, 1998: 577-580.
- [4] VAN COMPERNOLLE D. Noise adaptation in the hidden markov model speech recognition system[J]. Computer Speech and Language, 1987, 3: 151-168.
- [5] GALES M J F, YOUNG S J. Robust continuous speech recognition using parallel model combination[J]. IEEE Transactions on Speech and Audio Processing, 1996, 4(5): 352-359.
- [6] 丁 沛, 曹志刚. 融合语音增强与后续补偿的抗噪声语音识别方法[J]. 清华大学学报: 自然科学版, 2003, 43(7): 919-922.

A Improved Method for Robust Speech Recognition

LIU Jing-hua

(College of Information Science and Engineering, Huaqiao University, Quanzhou 362021, China)

Abstract: In order to improve the speech recognition accuracy in noisy environments, an algorithm proposed, which combines the signal level with the characteristic parameter level antinoise algorithm. At first, MMSE algorithm is used for the speech enhancement module, and then, new robust features RAS-MFCC effectively suppressed the additive and convolutional noise. The results show that this algorithm effectively improves the speech recognition accuracy in noisy environments, especially in low signal-to-noise ratio (SNR) conditions.

Keywords: speech recognition; antinoise arithmetic; speech enhancement; signal-to-noise ratio

(责任编辑: 鲁 斌 英文审校: 吴逢铁)