

# 高噪声环境下基于倒谱距离的语音端点检测算法的实现

王媛 刘玢

(中国农业大学 信息与电气工程学院, 北京 100083)

**摘要** 为了在高噪声环境下得到较清晰的语音信息,对倒谱距离算法进行了改进,利用语音信号的倒谱向量判断语音端点信息。MATLAB 仿真实验结果表明,在白噪声和汽车噪声环境下,信噪比分别为 10 dB 和 15 dB 时,得到的语音单字清晰度分别达到 93% 和 95%,实现了 90~100 dB 高噪声环境下可靠的语音通信。

**关键词** 倒谱距离; 语音端点检测; 语音信号; 高噪声

中图分类号 TN 912.35

文章编号 1007-4333(2006)02-0082-03

文献标识码 A

## Implement of endpoint detection method of speech based on cepstrum distance under high noise

Wang Yuan, Liu Heng

(College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China)

**Abstract** In order to get speech communication as clear as possible under a high noise environment, a research on the improved algorithm about cepstrum distance was conducted to estimate the message of endpoint voice with speech cepstrum vectors. The experiments show that cepstrum distance method has high performance in background with high noise. Via MATLAB, under the environment of white noise and car noise, the accuracies of the single-word definition in the speech experiments are 93% and 95% for SNRs 10 dB and 15 dB, respectively.

**Key words** cepstrum distance; voice endpoint detection; SNR; high noise

语音端点检测技术<sup>[1]</sup>是数字化语音降噪的核心技术。语音信号端点检测的目的是从连续采样得到的数字信号中检测出语音信号段和噪声信号段。准确的语音端点检测不仅提高了系统处理效率,同时也能够提高系统的识别率。传统的端点检测方法如短时能量、过零率、时频域的基音检测等算法在平稳噪声或高信噪比时性能较好,广泛应用于 90 dB 以下的噪声环境中,随着环境噪声的增强,当噪声强度高于 90 dB 时易发生漏检或虚检情况。基于倒谱距离的语音端点检测算法,可用于 90~100 dB 高噪声环境,提高语音信号的清晰度。目前,在大多数语音识别系统中,倒谱系数(如线性预测倒谱系数(LPC-CEP)或 Mel 刻度倒谱系数(MFCC))被选为代表语音信号的特征参数<sup>[2]</sup>,本文中采用倒谱向量取代这些简单特征参数作为语音端点检测的判决特征向量。

### 1 倒谱距离测量法

利用数字信号处理理论,把信号倒谱看成信号能量谱密度函数  $S(\omega)$  的对数的傅里叶级数<sup>[3]</sup>展开,即

$$\log S(\omega) = \sum_{n=-\infty}^{\infty} c_n e^{-jn\omega} \quad (1)$$

式中:  $c_n$  为倒谱系数,  $c_n = c_{-n}$  为实数,且

$$c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log S(\omega) d\omega \quad (2)$$

对于一对谱密度函数  $S_1(\omega)$  和  $S_2(\omega)$ ,应用 Parsavel 定理可用倒谱距离来表示其对数谱的均方距离

$$d^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\log S_1(\omega) - \log S_2(\omega)|^2 d\omega = \sum_{n=-\infty}^{\infty} (c_{1n} - c_{2n})^2 \quad (3)$$

收稿日期: 2005-07-08

基金项目: 辽宁省自然科学基金资助项目(2001102067)

作者简介: 王媛,硕士研究生;刘玢,副教授,主要从事高噪声环境抗噪通信技术的研究, E-mail:cauzd@163.com

式中： $c_n$  和  $c_n$  分别为对应于谱密度函数  $S(\ )$  和  $S(\ )$  的倒谱系数。对数谱的均方距离<sup>[4]</sup>可以表示 2 个信号谱的区别，故可作为一个判决参数。本文中用倒谱距离测量方法来判定各个信号帧是语音帧还是噪声帧。实际上，基于能量的检测方法可以看作倒谱距离测量方法的特例，因为  $c_0$  含有信号的能量信息。

倒谱距离测量法的步骤如下：

1) 假定抽样信号的前几帧信号是背景噪声。利用这些帧的倒谱系数的平均值作为背景噪声倒谱系数的估计值，用向量  $C$  表示。

2) 计算每帧信号的倒谱系数，以及每帧信号的倒谱系数与噪声倒谱系数估计值的倒谱距离。式

(3) 可近似为

$$d = 4.3429 \sqrt{c_0 - c_0)^2 + 2 \sum_{n=1}^p (c_n - c_n)^2} \quad (4)$$

式中： $c_n$  为对应于  $C$  的倒谱系数； $p$  为倒谱系数的阶数。

3) 由式(4) 计算的各帧倒谱距离得到倒谱距离轨迹，然后利用近似能量方法中门限判决方法检测语音段和噪声段。

4) 为使背景噪声倒谱系数的估计值  $C$  能够随噪声变化，采用自适应处理过程<sup>[5]</sup>，背景噪声倒谱估计值  $C$  利用已经检测过的前一信号帧倒谱向量，按照

$$C = aC + (1 - a) C_i \quad (5)$$

规则更新。式中： $a$  为时间调整因子； $i$  为上一信号帧帧号； $C_i$  为倒谱向量。

## 2 实验结果与分析

### 2.1 倒谱距离算法仿真实验

实验采用一段语音信号“中国农业大学”，经 8 kHz 采样和 16 bit 量化处理，分别加汽车噪声和白噪声作为样本，算法的软件流程见图 1。用 MATLAB 软件仿真<sup>[6]</sup>这段语音信号，结果见图 2~ 4。

图 2(a) 为纯语音信号归一化后的波形，通过前 20 帧语音信号得到此时的阈值  $C$ ；图 2(b) 为语音端点检测结果，中间有值区间表示语音信号。90 dB 汽车噪声环境下的语音波形见图 3(a)，此时的语音信号信噪比已达到 15 dB。由图 3(b) 可以看出，经倒谱距离处理后被测语音特征已被明显地区分出，

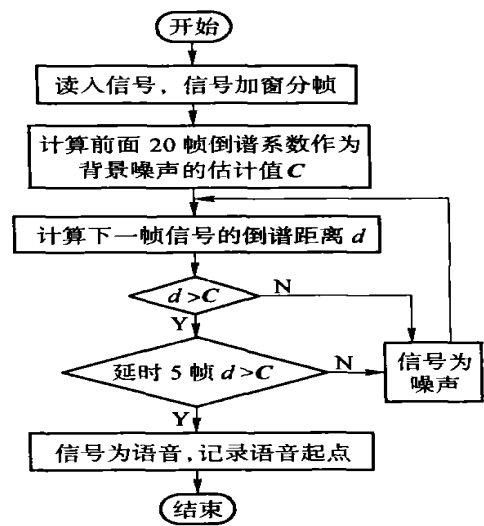


图 1 倒谱距离算法软件流程图

Fig. 1 Software flow chart of cepstrum distance

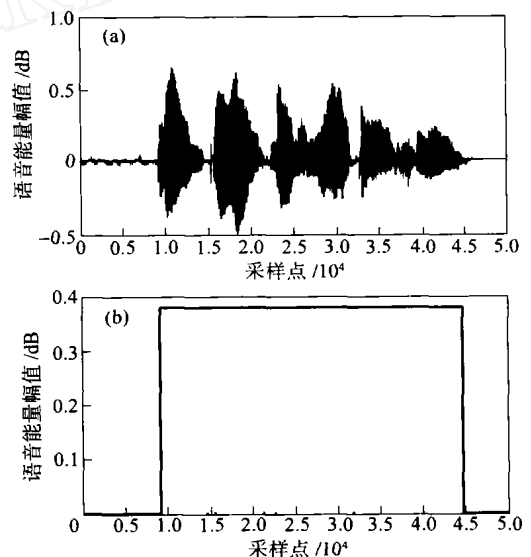


图 2 纯净语音波形图(a)及语音端点检测图(b)

Fig. 2 Clear speech waveform (a) and endpoint detection (b) charts

在无语音信号处，倒谱距离值在 0 dB 附近微弱波动，即经倒谱距离处理后，汽车噪声强度被削弱，语音信号特征变得明显起来，处理效果十分理想。从图 3(c) 可以看出，语音端点检测效果良好。图 4 示出 100 dB 白噪声环境下语音信号信噪比 10 dB 时的实验结果。从图 4(a) 可以看出，语音信号被噪声污染得十分严重，有些语音信息基本无法识别。经倒谱距离处理后，虽本底噪声有所增强，无语音信号处倒谱距离值提升到 1 dB 左右，且波动范围较大，使

语音信号轻音部分特征的提取受到影响(图4(b)),但语音端点检测仍较准确地找到了语音起止点(图4(c)),实现了可靠的语音端点检测。

从 MATLAB 仿真试验<sup>[7]</sup>结果可以看出,在 90 ~ 100 dB 高噪声环境下,倒谱距离法对于语音特征参数的提取是较准确的。

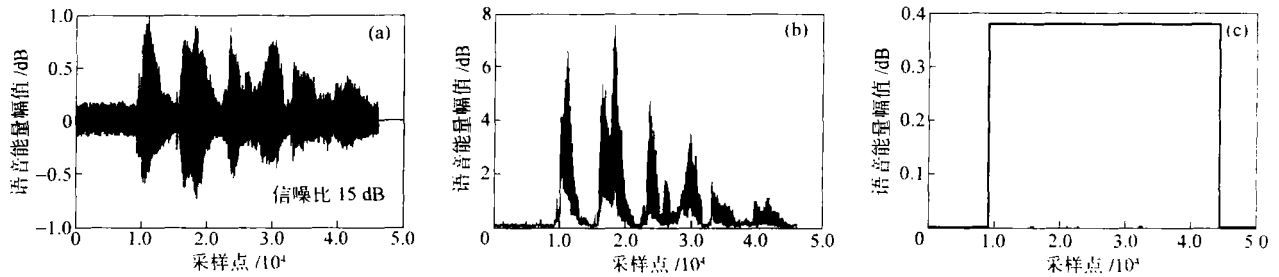


图3 带汽车噪声语音波形图(a)及其倒谱距离图(b)和语音端点检测图(c)

Fig. 3 Car noise speech waveform (a), cepstrum distance (b) and endpoint detection (c) chart

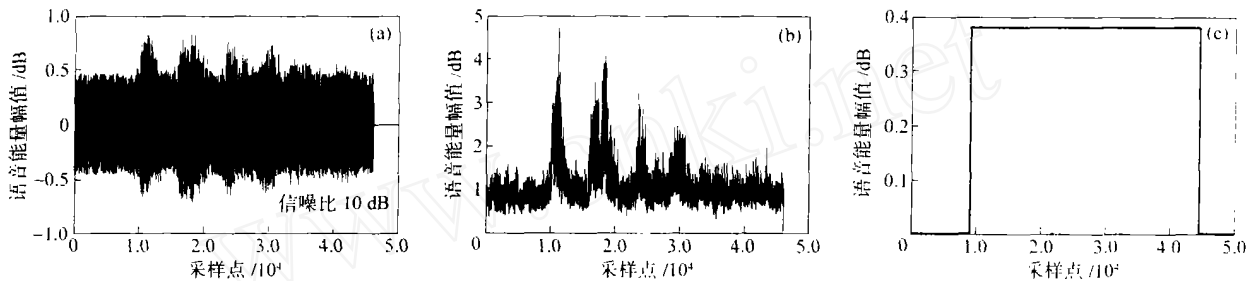


图4 带白噪声语音波形图(a)及其倒谱距离图(b)和语音端点检测图(c)

Fig. 4 White noise speech waveform (a), cepstrum distance (b) and endpoint detection (c) chart

## 2.2 语音测听实验

实验采用 TI 公司(美国)的 TMS320VC5402 平台,系统主要外围芯片包括:FLASH(AMD29LV800)、CODEC(AIC23)和 DSP(TMS320VC5402)。

以标准词可懂度测试词表为依据,以硬件电路为平台,20个参加测试者分为10人1组,分别进行倒谱距离处理后的语音测听实验。统计结果表明,经倒谱距离处理后,90 dB 汽车噪声下的语音单字清晰度超过 95%,100 dB 白噪声下的语音单字清晰度达到 93%。

## 3 结束语

倒谱距离法提高了语音信号的处理速度和精度,并尽可能地保持了原始信号的特征参数,在语音信号处理上有很好的发展空间。实验结果表明,在 90 dB 汽车噪声下,语音的单字清晰度超过 95%,而在 100 dB 白噪声下,语音单字清晰度也达到了 93% 的水平。

对大于 110 dB 高噪声环境下的语音信号,倒谱距离波形仍然会受到噪声的干扰,强度较弱的语音

或清音会丢失,此问题有待进一步研究。

## 参 考 文 献

- [1] 蔡莲红. 现代语音技术基础与应用[M]. 北京:清华大学出版社,2003:27~29
- [2] 张雄伟. 现代语音处理技术及应用[M]. 北京:机械工业出版社,2003:47~58
- [3] Quatieri F T. 离散时间语音信号处理-原理与应用[M]. 赵胜辉,刘家康,谢湘,等. 北京:电子工业出版社,2004:240~247
- [4] Chang H Y, Soo N K, Susanto R. An invertible frequency eigendomain transformation for masking-based subspace speech enhancement [J]. IEEE Signal Processing Letters, 2005,12:461~464
- [5] 于迎霞,史家茂. 一种改进的基于倒谱特征的带噪端点检测方法[J]. 计算机工程,2004,30(19):85~87
- [6] 王宏. MATLAB6.5 及其在信号处理中的应用[M]. 北京:清华大学出版社,2004:192~235
- [7] Childers D G. MATLAB 之语音处理与合成工具箱[M]. 北京:清华大学出版社,2004:95~128