

# 语音信号共振峰提取方法的研究分析

青海师范大学物理系 杨丹 姜占才 余莹良 李振起

**[摘要]**目前的提取语音共振峰的方法比较多,常用的方法有倒谱法、LPC谱估计法、LPC倒谱法,但没有一种方法是十分完美的,为了系统的深入的研究共振峰的提取,本文对同一帧语音信号进行了不同方法的共振峰提取实验仿真,给出了具体的共振峰频率数据,同时对不同方法的优缺点进行了科学的、深入的研究分析,用Matlab对算法进行仿真实验,实验结果表明共振峰的提取方案中LPCC倒谱法避免了一般同态处理中对复对数的麻烦,可以较为理想的实现对共振峰的提取。

**[关键词]**共振峰 倒谱法 LPC谱估计法 LPC倒谱法 同态处理

## 0.引言

在语音数字信号处理的研究中,语音信号的共振峰是一个十分重要的性能参数。共振峰是准周期脉冲激励进入声道时产生的一组共振频率。共振峰参数包括共振峰频率和频带宽度,它是区别不同韵母的重要参数。当前共振峰的提取方法有很多种,每种方法各有利弊,本文对多种共振峰提取方法进行实验和比较分析。

与基音提取类似,得到精确的共振峰估值也是很困难的,这些困难是:

### (1)虚假峰值

在正常情况下,频谱包络中的最大值完全是由共振峰引起的。但有时会出现虚假峰值,在利用非线性预测分析方法的频谱包络估计时,出现虚假峰值情况较多,而采用线性预测方法时,出现虚假峰值情况得到了很大的改善。

### (2)共振峰合并

当出现相邻共振峰时,可能会有频率靠得太难以分辨的情况,而寻找一种理想的能对共振峰合并进行识别的共振峰提取算法有不少实际困难。

### (3)高音调语音

传统的频谱包络估值方法是利用由谐波峰值提供的样点,而高音调语音(如女声和童声)的谐波间隔比较宽,因而为频谱包络估值所提供的样点比较少。而利用线性预测进行频谱包络估值可以一定程度上改善这一问题。

## 1.谱包络提取

共振峰信息包含在语音频谱包络中,因此共振峰参数提取的关键是估计自然语音频谱包络,并认为谱包络中的最大值就是共振峰。下图1是对21帧信号的线性预测谱包络提取分析。

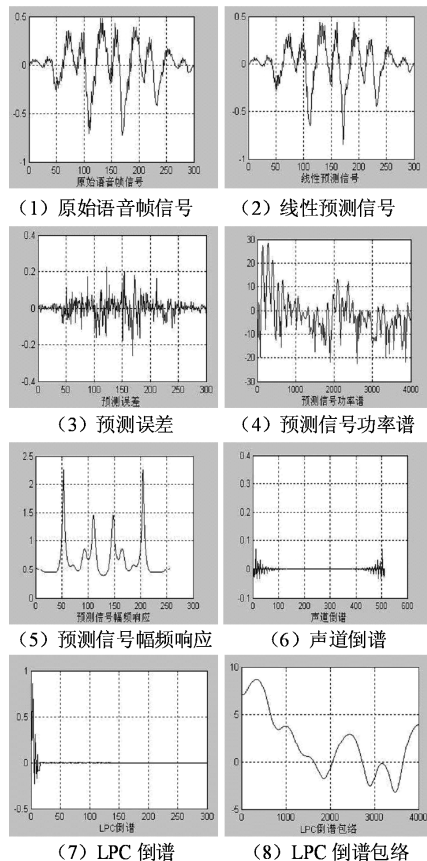


图1

## 2.倒谱法提取共振峰

虽然可以直接对语音信号求离散傅里叶变换(DFT),然后用DFT谱来提取语音信号的共振峰参数,但是,直接DFT的谱要受基频谐波的影响,最大值只能出现在谐波频率上,因而共振峰测定误差较大。为了消除基频谐波的影响,可以采用同态解卷技术,经过同态滤波后得到平滑的谱,这样简单地检测峰值就可以直接提取共振峰参数,因而这种方法更为有效和精确。因为倒谱运用对数运算和二次变换将基音谐波和声道的频谱包络分离开来。因此用低时窗 $l(n)$ 从语音信号倒谱 $c(n)$ 中所截取出来得 $h(n)$ ,能更精确地反映声道响应。这样,由 $h(n)$ 经DFT得到的 $\hat{H}(k)$ ,就是声道的离散谱曲线,用 $\hat{H}(k)$ 代替直接DFT的频谱,因为去除了激励引起的谐波波动,所以可以更精确地得到共振峰参数。

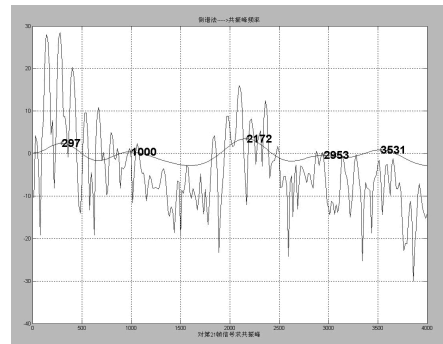


图2 倒谱法求取的共振峰参数图

## 3.LPC法提取共振峰

从线性预测导出的声道滤波器是频谱包络估计器的最新形式,线性预测提供了一个优良的声道模型(条件是语音不含噪声)。尽管线性预测法的频率灵敏度和人耳不相匹配,但它仍是最廉价、最优良的行之有效的办法。

用线性预测可对语音信号进行解卷:即把激励分量归入预测残差中,得到声道响应的全级模型 $H(z)$ 的分量,从而得到这个分量的 $a_i$ 参数。尽管其精度由于存在一定的逼近误差而有所降低,但去除了激励分量的影响。此时求出声道响应分量的谱峰,就可以求出共振峰,这里有两种途径:一是用标准的求取复根的方法计算全级模型分母多项式 $A(z)$ 的根,称为求根法;一是用运算量较少的DFT法,求 $A(z)$ 的离散频率响应 $A(k)$ 的谷点来得到共振峰的位置。因为 $A(z) = 1 - \sum_{i=1}^p a_i z^{-i}$ ,所以若求此多项式系数序列 $(1, a_1, a_2, \dots, a_p)$ 的DFT,就可以得到 $A(k)$ 。但是一般预测阶数 $p$ 不大,这就影响了求其谷点即求其共振峰频率值的精度。为了提高DFT的频率分辨率,可以采用补0的办法增加序列的时间长度,即用 $(1, a_1, a_2, \dots, a_p, 0, 0, \dots, 0)$ 进行DFT,为了能利用FFT,长度一般取为64点、128点、256点、512点等。另外也可以采用抛物线内插技术,解决频率分辨率较低的情况下的共振峰频率值的求取。

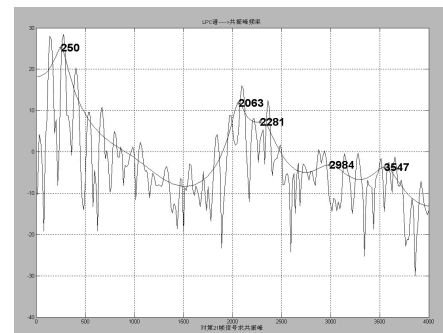


图3 LPC谱估计法求取的共振峰参数图

#### 4. 求根法提取共振峰

找出多项式复根的过程通常采用牛顿-拉夫逊(Newton-Raphson)算法。其方法是一开始先猜测一个根值并就此猜测值计算多项式及其导数的值,然后利用结果再找出一个改进的猜测值。当前后两个猜测值之差小于某门限时结束猜测过程。由上述过程可知,重复运算找出复根的计算量相当可观。然而,假设每一帧的最初猜测值域前一帧的根的位置重合,那么根的帧到帧的移动足够小,经过较少的重复运算后,可使新的根的值会聚在一起。当求根过程初始时,第一帧的猜测值可以在单位圆上等间隔设置。

具体过程是:设  $z_i = r_i e^{j\theta_i}$  为第一个根,则其共轭值  $z_i = r_i e^{-j\theta_i}$  也是一个根。设与  $i$  对应的共振峰频率为  $F_i$ , 3dB 带宽为  $B_i$ , 则它们存在下面的关系:

$$2\pi TF_i = \theta_i \quad e^{-B_i \pi T} = r_i$$

所以:

$$F_i = \frac{\theta_i}{2\pi T} \quad B_i = \frac{-\ln r_i}{\pi T}$$

式中,  $T$  是采样周期。

因为预测器阶数  $p$  是预先选定的,所以复共轭根的数量最多是  $p/2$ 。因而判断某一个极点属于哪一个共振峰的问题就不太复杂。而且,不属于共振峰的额外极点容易排除掉,因为其带宽比共振峰带宽要大得多。

#### 5. LPC 倒谱法提取共振峰

语音信号的倒谱可以通过对信号做傅里叶变换,取模的对数,在求反傅里叶变换得到。由于频率响应  $H(e^{j\omega})$  反应声道的频率响应和被分析信号的谱包络,因此用  $\log |H(e^{j\omega})|$  做傅里叶反变换求出的 LPC 倒谱系数,也可以被认为是包含了信号谱的包络信息,因此可以将其看作对原信号短时倒谱的一种近似。

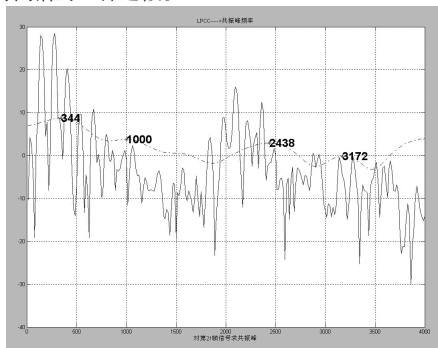


图4 LPC倒谱法求取的共振峰参数图

通过线性预测分析得到的合成滤波器的系统函数为  $H(z) = \frac{1}{1 - \sum_{i=1}^p a_i z^{-i}}$ , 其冲击响应为  $h(n)$ 。下面求  $h(n)$  的倒谱  $\hat{h}(n)$ , 首先

根据同态分析方法, 有  $\hat{H}(z) = \log H(z)$  因为  $H(z)$  是最小相位的, 即在单位圆内是解析的, 所以  $\hat{H}(z)$  一定可以展开成级数形式, 即  $\hat{H}(z) = \sum_{n=1}^{+\infty} \hat{h}(n) z^{-n}$ , 就是说  $\hat{H}(z)$  的逆变换  $\hat{h}(n)$  是存在的。设  $\hat{h}(0) = 0$ , 将式两边同时对  $z^{-1}$  求导, 得

$$\frac{\partial}{\partial z^{-1}} \log \left( \frac{1}{1 - \sum_{i=1}^p a_i z^{-i}} \right) = \frac{\partial}{\partial z^{-1}} \sum_{n=1}^{+\infty} \hat{h}(n) z^{-n}$$

得到

$$\sum_{n=1}^{+\infty} n \hat{h}(n) z^{-n+1} = \frac{\sum_{i=1}^p i a_i z^{-i+1}}{1 - \sum_{i=1}^p a_i z^{-i}}$$

(上接第165页)

- [1] 王宏志. 一本《晚清翻译史》的构思[J]. 中国比较文学, 2001, (2): 104.
- [2] 谭汝谦. 中国译日本书综合目录[M]. 香港: 香港中文大学出版社, 1980.
- [3] (日) 实藤惠秀. 中国人留学日本史[M]. 谭汝谦, 林启彦译. 北京: 生活·读书·新知三联书店, 1983.
- [4] 李喜所. 近代中国的留学生[M]. 北京: 北京人民出版社, 1987. 122, 124, 144, 198-201.
- [5] 费正清. 剑桥中国晚清史下卷[M]. 北京: 中国社会科学出版社, 1993. 405, 406, 333.

有

$$\left( 1 - \sum_{i=1}^p a_i z^{-i} \right) \sum_{n=1}^{+\infty} n \hat{h}(n) z^{-n+1} = \sum_{n=1}^{+\infty} i a_i z^{-i+1}, \text{ 令式左右两边 } z \text{ 的各次}$$

幂前系数分别相等得到

$$\begin{cases} \hat{h}(1) = a_1 \\ \hat{h}(n) = a_n + \sum_{i=1}^{n-1} \left( 1 - \frac{i}{n} \right) a_i \hat{h}(n-i) & 1 \leq n \leq p \\ \hat{h}(n) = \sum_{i=1}^p \left( 1 - \frac{i}{n} \right) a_i \hat{h}(n-i) & n > p \end{cases}$$

按照上式可直接从预测系数  $\{a_i\}$  求得倒谱  $\hat{h}(n)$ 。这个倒谱是根据现行预测模型得到的, 又称为 LPC 倒谱。

#### 6. 分析和总结

以上实验表明, 倒谱法因为其频谱曲线的波动比较小, 所以估计共振峰参数的效果是较好的, 但其运算量太大。

LPC 法的缺点是用一个全极点模型逼近语音谱, 对于含有零点的某些音来说  $A(z)$  的根反映了极零点的复合效应, 无法区分这些根是相对于零点还是极点, 或完全与声道的谐振极点有关。声道的系统函数可以由一组 LPC 预测系数唯一确定, 因此通过分析能估计出声道调制的效果, 即获得共振峰参数。当然, LPC 分析阶  $P$  的选择很重要, 它近似等于语音信号的抽样频率, 这是因为语音谱一般可用每 1 kHz 具有 1 对共振极点的平均密度来表示声道造成的响应, 于是采样频率为  $F_s$  (kHz) 的语音信号的 LPC 谱大约有  $F_s$  个极点。LPC 分析在大多数情况下能成功提取语音的共振峰参数。

求根法求共振峰的优点在于通过对预测多项式系数的分解可以精确地确定共振峰的中心频率和带宽。

LPC 倒谱由于利用线性预测中声道系统函数  $H(z)$  的最小相位特性, 因此避免了一般同态处理中对复对数的麻烦。

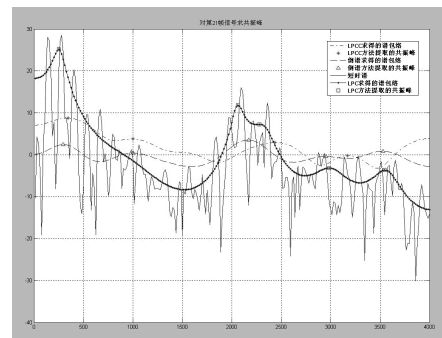


图5 共振峰参数求取三种方法比较分析图

当前还有很多种新的共振峰提取方法, 每个方法都是在传统共振峰提取方法的基础上进行研究和改进提出的, 在一定的基础上较好的实现了共振峰的提取, 效果也较为理想, 但都保留了一定的改进空间供后来者继承和研究创新。

#### 参考文献

- [1] Wat anabe A. Formant estimation method using inverse-filter control [J]. IEEE Transactions on Speech and Audio Processing, 2001, 9(4): 317-326.
- [2] L.R. Rabiner, B.H. Juang. Fundamentals of Speech Recognition. Prentice Hall, Englewood Cliffs, 1993. 清华大学出版社(影印), 2002 年.
- [3] 黄海, 陈祥献. 基于 Hilbert-Huang 变换的语音信号共振峰频率估计[J]. 浙江大学学报: 工学版, 2006, 40(11): 1926-1930.
- [4] 胡航. 语音信号处理(修订版). 哈尔滨工业大学出版社, 2002 年.
- [5] 于萍, 韩冰等. 噪声声学分析和电声门图的比较研究[J]. 听力学及言语疾病杂志, 2005, 13(3): 160-163.
- [6] 韩继庆, 张磊, 郑铁然. 语音信号处理. 清华大学出版社, 2004 年.

1993. 405, 406, 333.

- [6] 劝同乡父老遣子弟航洋游学书. 游学译编(6).
- [7] 陈潮. 近代留学生[M]. 上海: 上海古籍出版社, 1998. 58, 83-84.
- [8] 李亚舒, 黎难秋. 中国科学翻译史[M]. 长沙: 湖南教育出版社, 2002. 231, 209-211.
- [9] 梁启超. 论变法不知本原之害[M]. 饮冰室合集文集之一.
- [10] 潘玉田, 陈永刚. 中西文献交流史[M]. 北京: 北京图书馆出版社, 1997. 72.
- [11] 葛兆光. 1895 的中国: 思想史上的象征意义[J].