

# 基于 MATLAB 语音信号基频估计方法研究

## RESEARCH OF THE PITCH ESTIMATION OF THE SPEECH SIGNAL BASED ON MATLAB

(南京工业大学)王 静

WANG JING

摘要:不同的语音分析和综合方法就有不同的基音估计的方法,本文就简单的自相关基音估计入手分析讨论,并讨论了倒谱法,小波分析的基音估计。借助 MATLAB 语音工具箱 VOICEBOX 对种不同分析综合方法的基音估计的频域图做了分析比较。

关键词:语音信号;基音估计;MATLAB

中图分类号:TN912.3 文献标识码:A

Abstract: Different approaches of the analysis and synthesis of the speech signal lead to different approaches of the estimation of the pitch contour. This paper discusses the autocorrelation method and cepstrum and wavelet method of the pitch contour estimation. The pitch contour is computed using both of the methods with the help of the VOICEBOX of the MATLAB speech signal processing tool. Key words: speech signal processing pitch contour estimation MATLAB.

### 引言

语音是人们交流思想和进行社会活动的最基本手段,我们要对语音信号进行测定并将其转变为另一种形式,以提高我们的通信能力。基音是语音信号的一个重要特征参数,基音周期估计的准确度对合成语音的质量有决定性作用。

不同的语音分析和综合方法就具有不同的基音估计的方法,在同太分析中,倒谱中的倒谱峰位置提供了基音估计。线性预测误差中基本峰值之间的间隔导出另一种基音周期估计。我们将就自相关和倒谱法对基音估计做分析比较。

### 1 自相关的基音估计法

考虑一离散时间短序列,由下式给定:

$$s_n[m] = s[m]w[n-m]$$

其中  $w[n]$  是长度为  $N_w$  的分析窗。短时自相关函数  $r_n(\tau)$  定义为:

$$r_n[\tau] = s_n[\tau] * s_n[-\tau] = \sum_{m=-\infty}^{\infty} s_n[m]s_n[m+\tau]$$

当  $s[m]$  是周期性的,周期为  $P$ ,则  $r_n[\tau]$  在基音周期  $P$  或附近包含峰值。相关基音估计的获得可以通过可能的基音周期 ( $P > 0$ ) 范围内对误差最小化,误差准则为:

$$E[P] = \sum_{m=-\infty}^{\infty} (s_n[m] - s_n[m+P])^2$$

$E[P]$  相对于  $P$  最小化得到:

$$\hat{P} = \max_P \left( \sum_{m=-\infty}^{\infty} s_n[m]s_n[m+P] \right)$$

其中  $P > \varepsilon$ , 即  $P$  离零足够远。

自相关函数是测量“自相似性”的尺度,所以我们期望对周期性序列的自相关函数在  $P$  附近出现峰值。周期波形的短时自相关函数的包络随着  $P$  的增大而大致线性的减小。

### 2 倒谱基音估计的基本概念和原理

#### 2.1 卷积同太系统:

若一系统的系统函数为  $H(z)$ , 输入序列  $x_1(n)$  和  $x_2(n)$  对应输出序列  $y_1(n)$  和  $y_2(n)$  有:

$$H[x(n)] = H[x_1(n) * x_2(n)] = H[x_1(n)] * H[x_2(n)] = y_1(n) * y_2(n) = y(n)$$

如果一个系统具有上式所表示的性质,则称之为“卷积同太系统”。同太解卷系统服从广义叠加原理,其中输入式卷积运算而输出是普通相加运算。特征系统定义如下:

$$D_s[x(n)] = D_s[x_1(n) * x_2(n)] = D_s[x_1(n)] + D_s[x_2(n)] = \hat{x}_1(n) + \hat{x}_2(n)$$

对于复对数的定义表述如下:

$$\hat{X}(e^{j\omega}) = \log[X(e^{j\omega})] + j \arg[X(e^{j\omega})]$$

#### 2.2 倒频谱:

如果复对数满足上式讨论,那么对于一个傅氏变换取复对数,再做傅氏反变换就能得到卷积特征系统输出,即:

$$\hat{x}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{X}(e^{j\omega}) e^{j\omega n} d\omega$$

特征系统输出  $\hat{x}(n)$  称为“复倒频谱”。我们用“倒频谱”来表示量

$$c(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{X}(e^{j\omega}) e^{j\omega n} d\omega$$

#### 2.3 倒谱的基音估计:

以倒谱为基础的基音估计,其要点为先计算倒谱,然后在预期的基音周期附近寻找一个峰值,如果倒谱的峰值超出了预先设定的门限,则输入的语音段应该定为浊音,而峰的位置就是基音周期的良好估值。如果不存在超出门限的峰值,则输入语音为清音。如果计算的是一个依赖于时间的倒谱(它以时间依赖傅氏变换为基础),则可估计出激励模型以及基音周期随时间的变化。

### 3 小波变换的基音估计方法

#### 3.1 小波分析

小波分析是当前应用数学中迅速发展起来的新领域,它是

时间和频率的局部变换,能有效地从信号中提取信息。小波变换具有恒 Q 性质,在信号的高频域部分,可以取得较好的时间分辨率,在时间的低频域部分,可以取得较好的频率分辨率。因此可以预见小波分析在语音处理系统中具有广泛的应用前景。

### 3.1.1 离散小波变换

离散小波变换即为尺度  $a$  与平移  $\tau$  均离散化的小波变换。

尺度  $a$  按幂级数作离散化,即:

$$a_0^j = 1, a_0^1, a_0^2, \dots, a_0^m \quad \tau = k\tau_0 a_0^j \quad (j=0,1,2,\dots,m)$$

此时相应的小波变换定义为:

$$WT_x(a_0^j, k\tau_0) = \int x(t) \Psi_{a_0^j, k\tau_0}^*(t) dt$$

$$\text{式中 } \Psi_{a_0^j, k\tau_0}^*(t) = a_0^{-j/2} \Psi(a_0^{-j}t - k\tau_0)$$

为  $a, \tau$  离散下的子波基。一般  $a_0$  取 2,  $\tau_0$  取 1, 称为二进正交子波基。

则其相应的二进子波变换为:

$$WT_x(j, k) = \int x(t) \Psi_{j,k}^*(t) dt$$

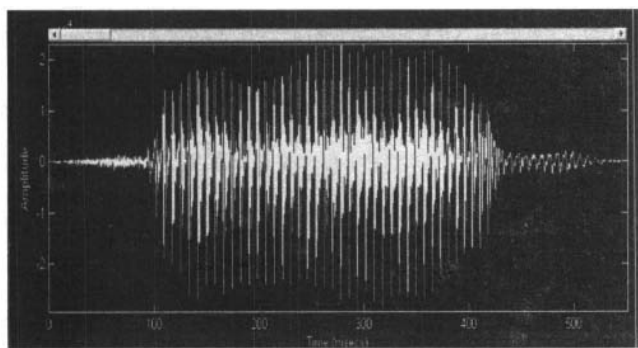
$$\text{式中 } \Psi_{j,k}^*(t) = 2^{-j/2} \Psi(2^{-j}t - k\tau_0)$$

### 3.2 利用小波变换进行基音提取

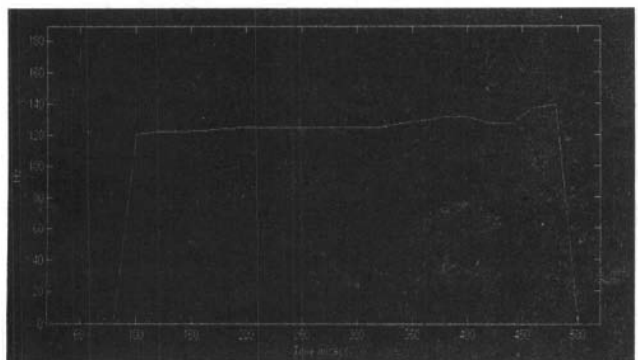
考虑到语音信号基音频率分布在 50-500 Hz 之间,而多分辨分析又具有把频率逐级对分的特性,所以采用 Mallat 算法直接对语音信号进行分解,并利用其低频小波系数对语音信号低频部分进行重构,与传统的 FFT 算法相结合,可有效地对基音频率进行估计。

## 4 MTALAB 分析过程

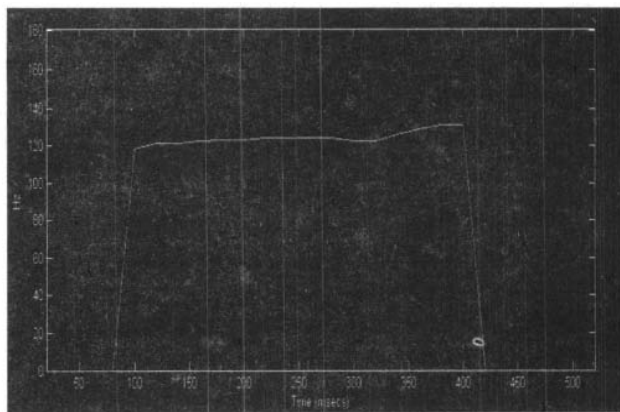
MATLAB 作为 mathworks 公司的高性能的数值计算和可视化软件,集成数值分析、矩阵计算、信号运算、信号处理和图形显示于一体,构成了一个方便且界面友好的用户环境。一下我们在 voicebox 工具箱的帮助下完成对英文 'had' 语音信号的时域波形。



图(1) 英文单词“had”的时域波形图,采样频率为 16kHz。



图(2)



图(3)

预处理使用预加重处理,即通过传递函数形为的 FIR 滤波器,窗函数选择 Hamming 窗,按自相关函数绘出的基音估计图谱如图(2),窗函数选择 Hamming 窗,按设定步幅移动,计算倒谱绘出的基音估计图谱如图(3)。

## 5 结束语

以上主要介绍了几种基音估计的方法,经 MATLAB 的仿真实验,我们看到各自的基音估计图谱,除了上诉方法外还有基于正弦函数的基音估计等方法。本文对两种方法提出各自思路,在其他方法中也有类似理论的分析。

本文作者创新点: 本文在理论分析比较了几种常用的语音信号基音估计的方法基础上,利用新的 MATLAB 语音信号处理工具,对于实际的英文单词语音采集信号进行了预加重,加窗等预处理后根据各自理论进行了分析比较,得到了直观的结果。

参考文献:

[1](美)Thomas F.Quatieri.离散时间语音信号处理--原理于应用[M]. 北京:电子工业出版社,2004.

[2](美)L.R.Rabiner.语音数字信号处理[M]. 北京:电子工业出版社,1993.

[3]Philip Loizou. COLEA MANUAL [M].U of Arkansas at LR. 1998.

[4]周爱军,马海瑞.基于声卡的 LABVIEW 数据采集与分析系统设计,微计算机信息[J].2005.21.9- 1:108.

作者简介:王静(1982.1-),女,汉族,江苏南京人,硕士研究生,现就读于南京工业大学信息科学与工程学院,信号与信号处理专业,E-mail:wj1982@126.com.

Biography:Wang Jing, Female, born in Jan 1982, Jiangsu Nanjing, Han nationality, postgraduate student in College of Information Science and Engineering, Nanjing University of Technology, major in signal and signal processing.

(210009 南京 南京工业大学)王静

Nanjing University of Technology Wang Jing

通讯地址:(210009 南京 南京新模范马路 5 号 152 信箱)王静

(收稿日期:2007.1.12)(修稿日期:2007.2.15)

欢迎订阅 欢迎刊登广告