线性预测及其 Matlab 实现

华,李伟,谭艳梅 (广西机电职业技术学院 广西 南宁 530007)

摘 要: 介绍线性预测编码(LPC)及其基本原理, 由语音信号常用的全极点模型的传递函数推导出线性预测方程, 然后 简要讲述了 G. 729 中线性预测分析的基本原理。重点是使用 Matlab 实现 G. 729 中 LPC 算法, 首先阐述了 加窗和自 相关函 数计算,并附上了 Matlab 的程序说明以及加窗前后运行结果图;最后,列出了求解线性预测滤波器系数经典的 Levinson-Durbin 算法, 同时给出了该算法的 Matlab 程序描述。利用 Matlab 程序进行线性预测, 可以很直观地知道分析结果, 为下 一 步将算法在 DSP 上的实现奠定基础。

关键词: 线性预测; Matlab; Levinson Durbin 算法; G 729

中图分类号: TN91

文献标识码:B

文章编号: 1004-373X(2009)07-133-03

Linear Prediction Coding and Its Implementation Based on Matlab

CAO Hua, LI Wei, TAN Yanmei

(Guangxi Technological College of Machinery and Electricity, Nanning, 530007, China)

Abstract: The principle of LPC and its encoding algorithm are mainly discussed. It is deduced the linear prediction equation from the all-pole model which is commonly used in the speech signal. And then it gives a description of the linear prediction analysis in G. 729. Above all, it explains the implementation of Linear Prediction Coding (LPC) algorithm based on Matlab. At first, it tells the window and the computation of auto-correlation, then gives the program of Matlab and the picture of windowed result. At last, it lists the classic Levinson Durbin algorithm for linear prediction coefficients, at the same time it gives the program too. The analytical result can be gained by Matlab linear prediction, and make a foundation for DSP.

Keywords linear prediction; Matlab; Levinson-Durbin algorithm; G. 729

线性预测分析是现代语音信号处理中最核心的技 术之一, 为现代语音信号处理的飞速发展立下了赫赫功 劳,在语音分析、合成、编码、识别等方面都有着广泛的 应用,至今仍是最有效的语音分析技术之一。如 VoIP 和 H.323 网上多媒体通信系统中所使用的 G.729 CS-ACELP 语音编码算法就是一种以共轭代数码本激励 线性预测为基础的高质量的语音编码标准[1]。

线性预测的基本原理

在语音信号处理中最常用的模型是全极点模型[2]。 线性预测所包含的基本概念是,通过使实际语音抽样和 线性预测抽样之间差值的平方和达到最小值,即进行最 小均方误差的逼近, 能够决定惟一的一组预测器系数。 如果一个随机过程用一个 p 阶全极点系统受白噪声激 励产生的输出来模拟,设这个系统的传递函数为,

收稿日期: 2008-09-04

基金项目: 广西省教育厅项目《嵌入式 VoIP 语音网关》 (200608LX232)

$$H(z) = \frac{S(z)}{E(z)} = \frac{G}{1 - \sum_{i=1}^{p} a_{i} z^{-i}} = \frac{G}{A(z)}$$
 (1)

其中: p 是预测器阶数; G 是声道滤波器增益。由此, 语 音抽样s(n) 和激励信号e(n) 之间的关系可用下列差分 方程来表示:

$$s(n) = Ge(n) + \sum_{i=1}^{p} a_i s(n-i)$$
 (2)

即语音样点间有相关性,可以用过去的样点值预测 未来样点值。对于浊音、激励 e(n) 是以基音周期重复 的单位冲激;对于清音 e(n) 是平稳白噪声。

在模型参数估计过程中, 把如下系统称为线性预 测器:

$$\widehat{s}(n) = \sum_{i=1}^{p} a_i s(n-i)$$
 (3)

式中: a_i 称为线性预测系数。从而, p 阶线性预测器的系 统函数具有如下形式:

$$P(z) = \sum_{i=1}^{p} a_i z^{-i} \tag{4}$$

在式(1) 中的 A(z) 称作逆滤波器, 基传输函数为:

$$A(z) = 1 - \sum_{i=1}^{p} a_{i} z^{-i} = \frac{GE(z)}{S(z)}$$
 (5)

预测误差 $\varepsilon(n)$ 为:

$$\varepsilon(n) = s(n) - \sum_{i=1}^{p} a_i s(n-i) = Ge(n) \qquad (6)$$

而线性预测方程可以按如下方法得到: 把某一帧内的短时平均预测误差定义为:

$$E\{\varepsilon^{2}(n)\} = E\{[s(n) - \sum_{i=1}^{p} a_{i}s(n-i)]^{2}\} \quad (7)$$

为使 $E\{\varepsilon^2(n)\}$ 最小, 对 a_i 求偏导, 并令其为零, 有:

$$E_{\{[s(n) - \sum_{i=1}^{p} a_{i}s(n-i)] | s(n-j)\}} = 0,$$

$$j = 1, 2, \dots, p$$
(8)

由此可计算出预测系数。

由于语音信号的短时平稳性, 要分帧处理($10 \sim 30 \text{ ms}$), 对于一帧从n时刻开窗选取的N 个样点的语音段 s(n), 记为 $\Phi_n(j,i)^{[2]}$:

$$\Phi_{n}(j,i) = E\{s_{n}(m-j)s_{n}(m-i)\}$$
 (9)

2 G.729 中线性预测分析的基本原理

用 10 阶线性预测(LP) 作短时分析, LP 合成滤波器定义为:

$$\frac{1}{A(z)} = \frac{1}{1 + \sum_{i=1}^{10} \hat{a_i} z^{-i}}$$
 (10)

这里的 $\hat{a_i}$ ($i=1,2,\cdots,10$) 是量化线性预测系数。用 30 ms 非对称窗加权 s(n),求 80 个样点(10 ms) 自相关系数,用 Levinson 算法求 LP 系数,然后转化为 LSP 系数,在 LSP 域量化和内插。

3 在 Matlab 上实现线性预测

3.1 加窗和自相关函数计算

G.729 语音编码算法大多是用 C 语言编程实现的,程序显得格外冗长,有一定的局限性。由于 M atlab 在数据分析、自动控制、数字信号处理、绘图等方面具有强大的功能,为了直观的表述,就运用 M atlab 来实现线性预测分析这一部分的算法。

图 1 给出了混合窗的示意图, 它由两部分构成, 前半部分是 1/2 个哈明窗; 后半部分是 1/4 余弦函数:

$$\omega_{\text{IP}} = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{399}\right), & n = 0, 1, \dots, 199\\ \cos\frac{2\pi(n - 200)}{159}, & n = 200, 201, \dots, 239 \end{cases}$$
(11)

式中: 有 5 ms(40 个样点)来自下一帧,15 ms(120 个样点)来自上一帧,当前帧 10 ms(80 个样点)加窗过程如

图1所示[3]。

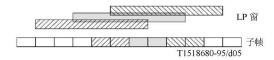


图1 LP分析中的加窗过程

加窗语音 s'(n) 为:

$$s'(n) = \omega_{LP}(n)s(n), \quad n = 0, 1, ..., 239$$
 (12)
对应的自相关函数 $r(k)$ 为:

$$r(k) = \sum_{n=k}^{239} s'(n) s'(n-k), \quad k = 0, 1, \dots, 10(13)$$

其 M atlab 描述如下:

$$\% = == 加窗(混合窗) = = \\ \text{for } n = ((k-1)*240+1): ((k-1)*240+200) \\ m = n - (k-1)*240; \\ w(n) = 0.54-0.46*\cos(2*pi*(m-1)/399); \\ \%1/2 \text{ hamming } \mathbf{\hat{a}} \\ s(n) = sn(n)*w(n); \\ \text{end} \\ \text{for } n = ((k-1)*240+201): (k*240) \\ m = n - (k-1)*240; \\ w(n) = \cos(2*pi*(m-201)/159); \%1/4 余弦窗 \\ s(n) = sn(n)*w(n); \\ \text{end} \\ \text{ord} \\ \end{cases}$$

得到加窗如图 2 所示。

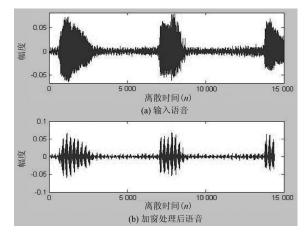


图 2 加窗前后语音比较图

3.2 LP系数的求取

求解线性预测滤波器系数 a_i ,将使用到最经典的 Levinson-Durbin 算法。该算法的步骤为:

(1)
$$E_n^0 = R_n(0)$$
;

(2)
$$k_i = [R_n(i) - \sum_{i=1}^{i-1} a_j^{i-1} R_n(i-j)] / E_n^{i-1};$$

(3)
$$a_i^i = k_i$$
:

(4)
$$a_j^i = a_j^{i-1} - k_i a_{i-j}^{i-1}, \ 1 \le j < i-1;$$

(5)
$$E_n^i = (1 - k_i^2) E_n^{i-1}$$

if $i \le p$ go to (1);

(6)
$$a_i = d_i^p, 1 \le j \le p$$
.

上面各式中括号内的上标表示预测器的阶数。步

为最小均方预测误差;R 为自相关系数;k 表示反射系 数, 取值范围为[-1,1]; d_i 表示i 阶预测器的第j 个系 数。通过运算发现,其实在计算过程中,虽然目标是计算 一个p 阶线性预测器的全部系数, 但实际上在递推过程 中仅算出了所有除数低于p 阶线性预测器的全部系数, 同时计算出了最小预测误差能量(递推式中反射系数的 取值范围是保证系统 H(z) 稳定性的充分必要条件, 即 多项式 A(z) 的全部根都落在单位圆内)。

从推导中得知, $\Phi_n(j,i)$ 的物理意义可以理解为 S_n 的短时自相关函数。因此, 它反映了语音波形的实际情 况, 即波形不同, 值也不同。然而 a_i 的取值由 $\Phi_n(j,i)$ 决 定, 并随 $\Phi_n(j,i)$ 的改变而改变, 因此也可以说 a_i 反映 了语音波形的实际情况^[3] 。

其 Matlab 描述如下:

```
function z = durbin1(y, order)
R = ze ros(1, order + 1);
aa= zeros(order, order);
parcor=zeros(1, order);
% auto correlation
N = size(y, 1);
for h=1: order+1
  R(h) = 0;
    for f= h. N
      R(h) = R(h) + y(f) * y(f-h+1);
    end
  end
parcor(1) = R(2) / R(1);
aa(1, 1) = parcor(1);
E = (1 - par cor(1)^2) * R(1);
for h=2; order
  temp=0;
  for f = 1; h = 1
    temp= temp+ aa(h-1,f) *R(h-f+1);
  end
  parcor(h) = (R(h+1) - temp) / E;
                                               %反射系数
aa(h, h) = par cor(h);
  for f=1: h-1
    aa(h, f) = aa(h-1, f) - par co r(h) * aa(h-1, h-f);
  end
E = E * (1 - parcor(h) 2);
end
```

z=-aa(order,:);

4 结 语

线性预测在语音处理方面的应用很广泛,而用 Matlab 来实现可以很直观地知道分析结果,为下一步 将算法在 DSP 上的实现奠定基础。目前, 语音处理最 普遍的就是使用到网络上, 而基于现代网络 V oIP 使用 到的语音编码中 G. 729 正是其中一种。它的出现, 使 用户能借助两台 PC 传输语音, 也可以把 PC 上的语音 转到手机上, 交互式游戏爱好者们相互之间还可以借助 高科技控制器进行交流。

参考文献

- [1] 吴家安. 现代语音编码技术[M]. 北京: 科学出版社, 2008.
- [2] 柏静, 韦岗. 一种基于线性预测与自相关函数法的语音基音 周期检测新算法[J]. 电声技术, 2005(8): 43-46.
- [3] 王涌,何剑春,刘盛.新型的神经网络线性预测语音编码算
- [4] 薛年喜. Matlab 在数字信号处理中的应用 [M]. 2版. 北京: 清华大学出版社,2008.
- [5] 朱蓉,黄冰. EVRC 语音编码算法研究及仿真[J]. 现代电子 技术, 2006, 29(2): 47-50.
- [6] 范晶,和应民,王桂梅.2 4 Kb/s 混合激励线性预测语音编 码的研究[J]. 牡丹江师范学院学报, 2007(4): 13-14.
- [7] 丛键, 张知易. 一种 600 b/s 极低速率语音编码算法[J]. 电 子与信息学报, 2007, 29(2): 429-433.
- [8] 论: VoIP 语音技术与传统网络的融合发展[EB/OL]. http://voip.microvoip.com/market/m1/200803/61433.html, 2008.
- [9] 语音技术在 Internet 上的新应用[EB/OL]. http://nc.mofcom. gov. cn/ new s/ 1056746. html, 2006.
- [10] Rapidshare Advances in Audio and Speech Signal Processing: Technologies and Applications [DB/OL]. http://rapidshare.com/files/31791068/IGI.rar.html, 2007.

(上接第125页)

参考文献

- [1] 吴仲城. 多维力传感器设计及信号分析方法研究[D]. 北京: 中国科学院等离子体物理研究所,2001.
- [2] Texas Instruments. MSC1210 User's Guide[Z]. 2002.
- [3] Sawada R, Higurashi E. Integrated Micro-displacement Sensor that can be Incorporated into Mini 3-dimensional Actuator Stage. Broadband Optical Networks and Technologies: An Emerging Reality/Optical MEMS/Smart Pixels/Organic
- Optics and Optoelectronics A. IEEE/ LEOS Summer Topical Meetings[C]. 1998.
- [4] 熊幸果,陆德仁. 微力微位移的天平测试方法[3]. 传感技术 学报,1997,10(2):47-52.
- [5] 洪跃,金士良.新型微位移电容式传感器的研制[].上海大 学学报,1995,1(6):652-657.
- [6] 胡永建,王晓梅.基于 MSC1210 的多路高精度温度采集系 统模块[J]. 电子技术应用, 2003, 29(7): 36-38.
- [7] Texas Intruments. MSC1210 Data Sheet Z.

作者简介 沈春山 硕士研究生。主要研究方向为机器人传感器。