

基于 MATLAB 的语音信号 LPC 技术分析研究

THE RESEARCH OF THE LPC ANALYSIS OF THE SPEECH SIGNAL PROCESSING BASED ON MATLAB

(南京工业大学)王 静

WANG JING

摘要:线性预测技术作为一种基于全极点模型假定和均方预测误差最小准则下的波形逼近技术。本文简要介绍了 LPC 技术的基本原理,并利用 MATLAB 这一有力工具对语音信号进行了 LPC 分析,并对阶数的选取做了比较分析。

关键词:语音信号;LPC;MATLAB 仿真

中图分类号:TN912.3 文献标识码:A

Abstract:The LPC(linear prediction coding) analysis is a theory based on the all-pole model assumption and minimum mean square rule to approach a waveform, therefore, it is independent of the assumption of the speech production model. It is utilized in many field like speech waveform coding. This paper introduces the principle of the PLC analysis and chooses the proper of with MATLAB. Key words:speech signal processing, LPC, MATLAB simulation.

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。线性预测是语音处理中的核心技术,它在语音识别、合成、编码、说话人识别等方面都得到了成功的应用。

1 概述

根据语音信号的产生模型,语音信号 $S(z)$ 是一个线性非因果稳定系统 $V(z)$ 受到信号 $E(z)$ 激励所产生的输出。在时域中,语音 $s(n)$ 信号是单位取样响应 $v(n)$ 和激励信号 $e(n)$ 的卷积。在语音信号数字处理所涉及的各个领域需要根据 $s(n)$ 来求 $v(n)$ 和 $e(n)$ 。例如语音信号的共振峰频率等(下面的 MATLAB 实验中可见)。在实现各种语音编码、识别、合成等算法时需要用到卷积运算。解卷积算法因可减少计算代价,其研究也是十分重要的。解卷积算法可分为两大类。一种是为线性系统 $V(z)$ 建立一个线性模型,然后对模型参数按某种最佳准则进行估计。第二种是无需线性系统建立一个模型的非参数解卷算法,其中重要的一种是同态信号处理方法。本文是第一种方法并采用最小均方准则(LMS)对全极点 AR 模型参数进行估计的 LPC 算法。

2 线性预测(LPC)

2.1 LPC 基本原理

假设一个随机过程用一个阶全极点系统受白噪声激励产生的输出来模拟,则传递函数:

$$H(z) = S(z)/U(z) = G/(1 - \sum_{k=1}^p a_k z^{-k})$$

其中 G 为常数, $S(z)$ 和 $U(z)$ 分别为输出 $S(n)$ 信号和输入信号 $u(n)$ 的 Z 变换,那么可以表示为差分方程:

$$S(n) = \sum_{k=1}^p a_k S(n-k) + Gu(n)$$

$\{a_k\}$ 有关部分为信号前 P 个样本来预测当前样本,定义预

测器和线性预测器的系统函数为:

$$\tilde{S}(n) = \sum_{k=1}^p a_k S(n-k), \quad P(z) = \sum_{k=1}^p a_k z^{-k}$$

此是表明预测误差序列是 $S(n)$ 通过具有如下系统函数产生的:

$$A(z) = 1 - \sum_{k=1}^p a_k z^{-k}$$

要使预测误差在某以短时总能量最小可求出此准则下最佳预测系数 $\{a_k\}$:

$$E_e = \sum_m e_e^2(m) = \sum_m [S_e(m) - \tilde{S}(m)]^2 = \sum_m [S_e(m) - \sum_{k=1}^p a_k S(m-k)]^2$$

使 E_e 达到最小的 $\{a_k\}$ 必定满足 $\partial E_e / \partial a_i = 0 (i=1,2,\dots,p)$, 由此得到以 $\{a_k\}$ 为变量的线性方程组:

$$\tilde{E}_n = \sum_m S^2(m) - \sum_{k=1}^p a_k \sum_m S_e(m) S(m-k) = \Phi(0,0) - \sum_{k=1}^p a_k \Phi(0,k)$$

$\tilde{E}_n = \sum_m e_e^2(m) = \tilde{G}^2 \sum_{m=0}^{N-1} u^2(m)$ 得到 $\tilde{G} = \tilde{E}^{1/2}$ 。预测残差序列 $e(n)$ 将接近于白噪声序列(清音)或 δ 脉冲串(浊音)时,此估算出 \tilde{G} 的及最佳线性预测系数 $\{a_k\}$ 来重构原信号 $S(n)$ 可获得很好的效果。

2.2 模型讨论

我们是基于 $s(n)$ 由理想全极点模型产生的。当 $s(n)$ 由非理想模型产生,而用理想模型来估计参数时,只能说是用此理想模型来逼近实际模型。在此情况下无所谓原产生模型的实际阶数应等于何值的问题,此时可提高 P 值可改善逼近效果。在语音处理领域中,模型 P 一般选在 8~12 之间。在下面的 MATLAB 实验中可看到效果。

3 卷积同态系统简介

与线性模型相对应,无需线性系统建立一个模型的非参数解卷算法,其中重要的一种是同态信号处理方法。

3.1 卷积同态系统

如果一个系统具有上式所表示的性质,则称之为“卷积同态系统”。同态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:

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

3.2 倒频谱:

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

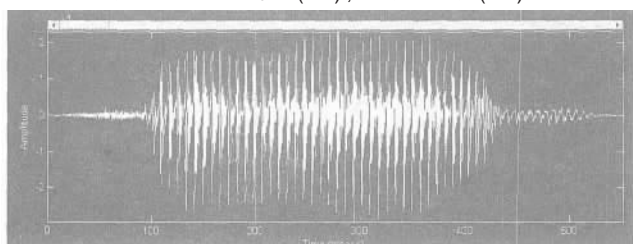
示量

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

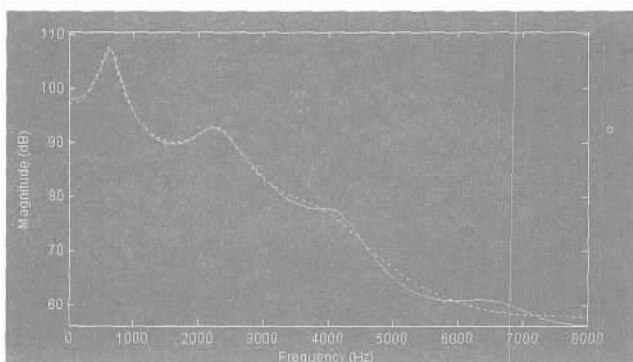
4. MATLAB 仿真过程

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和 PLC 频域分析。

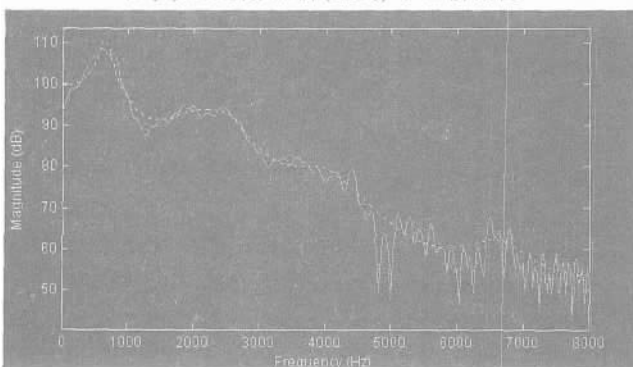
图(a)为英文单词“had”的时域波形图,采样频率为 16kHz。使用预加重处理,即通过传递函数形为 $H(z) = 1 - 0.97z^{-1}$ 的 FIR 滤波器,窗函数选择 Hamming 窗,在 200ms 处我们点击可以看到“had”中元音/ae/的 PLC 频域图。图中我们可以看到前三个明显的共振峰频率的估计值(Hz),共振峰幅度(dB)



图(a) 时域波形图



图(b) LPC 阶数 8 阶(虚线)与 12 阶比较



图(c) FFT 和 LPC(虚线)分析谱

图(b)中可见,要准确得到语音信号特征信息,适当地选取 LPC 阶数是必要的。一般情况下 12 阶的 LPC 分析已经达到电

话语音信号要求。但在特殊情况下是不够的,例如 16Hz 的语音采样信号。下图我们将看到 LPC 阶数对共振峰估计准确性的影响。图中叠加的虚线为 8 阶 LPC 共振峰估计。

图(c)为元音/ae/的 FFT 和 LPC 分析谱(虚线部分为 LPC 分析)在同一图上的对比。可以明显看到前三个共振峰

4 结束语

以上主要介绍了 LPC 技术的基本原理,经 MATLAB 的仿真实验,我们看到了使用 LPC 技术得到的一些特征参数,并对阶数的选取做了对比,最后对比了 FFT 的频谱。可见 LPC 技术在语音信号的处理和语音识别等领域都有广泛的应用。

本文作者创新点:理论分析研究了语音信号的 LPC 方法,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理论研究进行了分析,得到了直观的结果并进行了模型讨论。

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]韦江维,廖义奎,农建波,班世炳.TMS320C5402 在语音信号数字化及无线传输中的应用. [J]微计算机信息.2001.11.3:38
- 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
- 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:2007.2.3)(修稿日期:2007.3.5)

(上接第 308 页)

作者简介:尹士伟(1982-),男,北京,硕士研究生,研究方向:虚拟现实与三维可视化, E-mail:shiwei_yin@163.com.; 张光年(1950-),男,副教授,研究方向:数字图像处理和三维可视化;郭新宇(1973-),男,博士,副研,主要研究方向:农业信息技术。

Biography:Yin ShiWei(1978-),male, Major graduate student, Major in Virtual Reality And 3D visualization.

(100037 北京 首都师范大学信息工程学院)尹士伟 张光年

(100089 北京 国家农业信息化工程技术研究中心)郭新宇

(College of Information Engineering, Capital Normal University, Beijing, 100037, China)Yin ShiWei1, Zhang GuangNian1

(National Engineering Research Center for Information Technology in Agriculture, 100089, Beijing)Guo XinYu

通讯地址:(100037 北京 首都师范大学信息工程学院 2004 级研究生)尹士伟

(收稿日期:2007.2.25)(修稿日期:2007.3.28)