实验二　线性预测分析

【实验目的】

1、了解线性预测分析在语音信号处理中的重要性和必要性。

2、掌握线性预测分析的基本思想。

3、掌握 MATLAB 进行线性预测分析的流程。

【实验要求】

1、编程要求：编写一段 MATLAB 程序。

2、实现功能：实现语音的线性预测分析（LPC）。

【实验原理】

1、语音信号线性预测分析

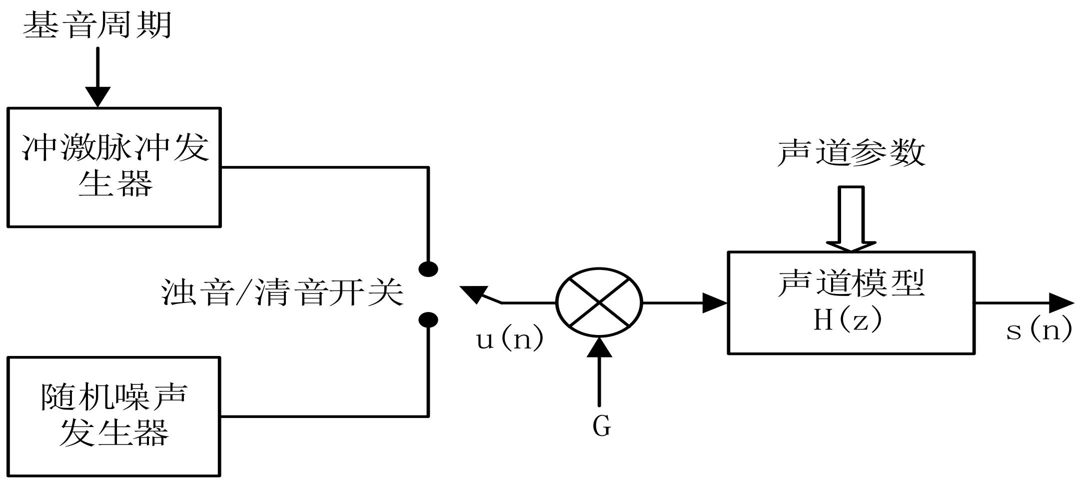

图 2-1 简化的语音产生模型

图 2-1 是简化的语音产生模型，将辐射、声道以及声门激励的全部效应简化为一个时变的数字滤波器来等效，其传递函数为

$$H(z)=\frac{S(z)}{U(z)}=\frac{G}{1-\sum_{i=1}^{p}a_i z^{-i}} \qquad (2\text{-}1)$$

这种表现形式称为 p 阶线性预测模型，这是一个全极点模型。

此时，$s(n)$ 和 $u(n)$ 间的关系可以用差分方程

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

表示，称系统

$$\hat{s}(n)=\sum_{i=1}^{p}a_i s(n-i) \qquad (2\text{-}3)$$

为线性预测器。$\hat{s}(n)$ 是 $s(n)$ 的估计值，它由过去 p 个值线性组合得到的，即由 $s(n)$ 过去的值来预测或估计当前值 $s(n)$。式中 $a_i\,(i=1,2,\cdots,p)$ 是线性预测系数。线性预测系数可

以通过在某个准则下使预测误差$e(n)$达到最小值的方法来决定，预测误差的表示形式如下：

$$e(n)=s(n)-\hat{s}(n)=s(n)-\sum_{i=1}^{p}a_i s(n-i) \qquad (2\text{-}4)$$

预测的二次方误差为

$$E=\sum_n e^2(n)=\sum_n[s(n)-\hat{s}(n)]^2=\sum_n[s(n)-\sum_{i=1}^{p}a_i s(n-i)]^2 \qquad (2\text{-}5)$$

为使 E 最小，求 E 对a_i的偏导为 0，即

$$\frac{\partial E}{\partial a_j}=0 \quad (1\le j\le p) \qquad (2\text{-}6)$$

则有，

$$\frac{\partial E}{\partial a_j}=2\sum_n s(n)s(n-j)-2\sum_{i=1}^{p}a_i\sum_n s(n-i)s(n-j)=0 \quad (1\le j\le p) \qquad (2\text{-}7)$$

定义$\phi(j,i)=\sum_n s(n-i)s(n-j)$，则式（2-7）可简化为

$$\phi(j,0)=\sum_{i=1}^{p}a_i\phi(j,i) \quad (1\le j\le p) \qquad (2\text{-}8)$$

联立式（2-5）、式（2-7）和式（2-8），可得最小均方误差表示为

$$E=\phi(0,0)-\sum_{i=1}^{p}a_i\phi(0,i) \qquad (2\text{-}9)$$

因此，最小误差有一个固定分量$\phi(0,0)$和一个依赖于预测系数的分量$\sum_{i=1}^{p}a_i\phi(0,i)$构成。为求解最佳预测器系数，必须首先求出$\phi(j,i)(i,j\in[1,p])$，然后可按照式（2-8）进行求解。很显然，$\phi(j,i)$的计算及方程组的求解都是十分复杂的。

2、线性预测分析的自相关解法

为了有效地进行线性预测分析，求得线性预测系数有必要用一种高效的方法来求解线性方程组。虽然可以用各种各样的方法来解包含p个未知数的p个线性方程，但是系数矩阵的特殊性质使得解方程的效率比普通解法的效率要高得多。自相关法是经典解法之一，其原理是在整个时间范围内使误差最小，即设$s(n)$在$0\le n\le N-1$以外等于 0，等同于假设$s(n)$经过有限长度的窗（如矩形窗、海宁窗或汉明窗）的处理。

通常，$s(n)$的加窗自相关函数定义为

$$r(j)=\sum_{n=0}^{N-1}s(n)s(n-j) \quad 1\le j\le p \qquad (2\text{-}10)$$

同式（2-8）比较可知，$\phi(j,i)$等效为$r(j-i)$。但是由于$r(j)$为偶函数，因此$\phi(j,i)$可表示为

$$\phi(j,i)=r(|j-i|) \qquad (2\text{-}11)$$

此时式（2-8）可表示为

$$\sum_{i=1}^{p} a_i r(|j-i|) = r(j) \qquad 1 \le j \le p \tag{2-12}$$

则最小均方误差改写为

$$E = r(0) - \sum_{i=1}^{p} a_i r(i) \tag{2-13}$$

展开式（2-11），可得方程组为

$$\begin{bmatrix} r(0) & r(1) & r(2) & \cdots & r(p-1) \\ r(1) & r(0) & r(1) & \cdots & r(p-2) \\ r(2) & r(1) & r(0) & \cdots & r(p-3) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ r(p-1) & r(p-2) & r(p-3) & \cdots & r(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} r(1) \\ r(2) \\ r(3) \\ \vdots \\ r(p) \end{bmatrix} \tag{2-14}$$

式（2-14）左边为相关函数的矩阵，以对角线为对称，其主对角线以及和主对角线平行的任何一条斜线上所有的元素相等。这种矩阵称为托普利兹（Toeplitz）矩阵，而这种方程称为 Yule-Walker 方程。对于式（2-14）的矩阵方程无需像求解一般矩阵方程那样进行大量的计算，利用托普利兹矩阵的性质可以得到求解这种方程的一种高效方法。

这种矩阵方程组可以采用递归方法求解，其基本思想是递归解法分布进行。在递推算法中，最常用的是莱文逊-杜宾（Levinson-Durbin）算法（如图 2-2 所示）。

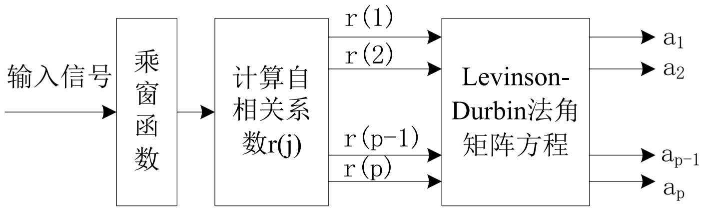

图 2-2 自相关解法

算法的过程和步骤为：

① 当 i=0 时，$E_0 = r(0), a_0 = 1$ （2-15）

② 对于第 i 次递归 $(i = 1, 2, \cdots, p)$：

i $$k_i = \frac{1}{E_{i-1}} \left[r(i) - \sum_{j=1}^{i-1} a_j^{i-1} r(j-i) \right] \tag{2-16}$$

ii $$a_i^{(i)} = k_i \tag{2-17}$$

iii 对于 j=1 到 i-1 $a_j^{(i)} = a_j^{(i-1)} - k_i a_{i-j}^{(i-1)}$ （2-18）

iv $$E_i = (1 - k_i^2) E_{i-1} \tag{2-19}$$

③ 增益 G 为

$$G=\sqrt{E_p} \tag{2-20}$$

通过对式（2-16）~式（2-18）进行递推求解，可获得最终解为

$$a_i=a_j^{(p)} \quad 1\le j\le p \tag{2-21}$$

由式（2-19）可得

$$E_p=r(0)\prod_{i=1}^{p}(1-k_i^2) \tag{2-22}$$

由式（2-22）可知，最小均方误差E_p一定要大于 0，且随着预测器阶数的增加而减小。因此每一步算出的预测误差总是小于前一步的预测误差。这就表明，虽然预测器的精度会随着阶数的增加而提高，但误差永远不会消除。由式（2-22）还可知，参数k_i一定满足

$$|k_i|<1, \quad 1\le i\le p \tag{2-23}$$

由递归算法可知，每一步计算都与k_i有关，说明这个系数具有特殊的意义，通常称之为反射系数或偏相关系数。可以证明，它就是多项式$A(z)$的根在单位圆内的充分必要条件，因此它可以保证系统$H(k)$的稳定性。

3、线性预测的其它参数

用线性预测分析法求得的是一个全极点模型的传递函数。在语音产生模型中，这一全极点模型与声道滤波器的假设相符合，而形式上是一自回归滤波器。用全极点模型所表征的声道滤波器，除预测系数$\{a_i\}$外，还有其他不同形式的滤波器参数。这些参数一般可由线性预测系数推导得到，但各有不同的物理意义和特性。在对语音信号做进一步处理时，为了达到不同的应用目的时，往往按照这些特性来选择某种合适的参数来描述语音信号。

1）预测误差及其自相关函数

由式（2-4）可知，预测误差为

$$e(n)=s(n)-\sum_{i=1}^{p}a_i s(n-i) \tag{2-24}$$

而预测误差的自相关函数为

$$R_e(m)=\sum_{n=0}^{N-1-m}e(n)e(n+m) \tag{2-25}$$

2）反射系数和声道面积

反射系数$\{k_i\}$在低速率语音编码、语音合成、语音识别和说话人识别等许多领域都是非常重要的特征参数。由式（2-18）可得：

$$\begin{cases} a_j^{(i)}=a_j^{(i-1)}-k_i a_{i-j}^{(i-1)} \\ a_{i-j}^{(i)}=a_{i-j}^{(i-1)}-k_i a_j^{(i-1)} \end{cases} \quad j=1,\cdots,i-1 \tag{2-26}$$

进一步推导，可得：

$$a_j^{(i-1)} = (a_j^{(i)} + a_j^{(i)} a_{i-j}^{(i)}) \big/ (1-k_i^2) \qquad j=1,\cdots,i-1 \tag{2-27}$$

由线性预测系数$\{a_i\}$可递推出反射系数$\{k_i\}$，即

$$\begin{cases} a_j^{(p)} = a_j & j=1,2,\cdots,p \\ k_i = a_i^{(i)} \\ a_j^{(i-1)} = (a_j^{(i)} + a_j^{(i)} a_{i-j}^{(i)}) \big/ (1-k_i^2) & j=1,\cdots,i-1 \end{cases} \tag{2-28}$$

反射系数的取值范围为$[-1,1]$，这是保证相应的系统函数稳定的充分必要条件。从声学理论可知，声道可以被模拟成一系列截面积不等的无损声道的级联。反射系数$\{k_i\}$反映了声波在各管道边界处的反射量，有：

$$k_i = \frac{A_{i+1} - A_i}{A_{i+1} + A_i} \tag{2-29}$$

式中，A_i是第i节声管的面积函数。式（7-29））经变换后，可得声管模型各节的面积比为：

$$\frac{A_i}{A_{i+1}} = \frac{(1-k_i)}{(1+k_i)} \tag{2-30}$$

3）线性预测的频谱

由式（2-1）可知，一帧语音信号$x(n)$模型可化为一个p阶的线性预测模型。当$z=e^{j\omega}$时，能得到线性预测系数的频谱（令$G=1$）：

$$H(e^{j\omega}) = \frac{G}{1-\sum_{n=1}^{p} a_n e^{-j\omega n}} \tag{2-31}$$

根据语音信号的数字模型，在不考虑激励和辐射时，$H(e^{j\omega})$即$X(e^{j\omega})$的频谱的包络谱。线性预测系数的频谱勾画出了 FFT 频谱的包络，反映了声道的共振峰的结构。

【实验步骤】

1、 根据莱文逊-杜宾自相关法求线性预测系数的原理，编写 MATLAB 函数，并与 MATLAB 自带的 LPC 函数进行比较，测试效果如图 2-3 所示。

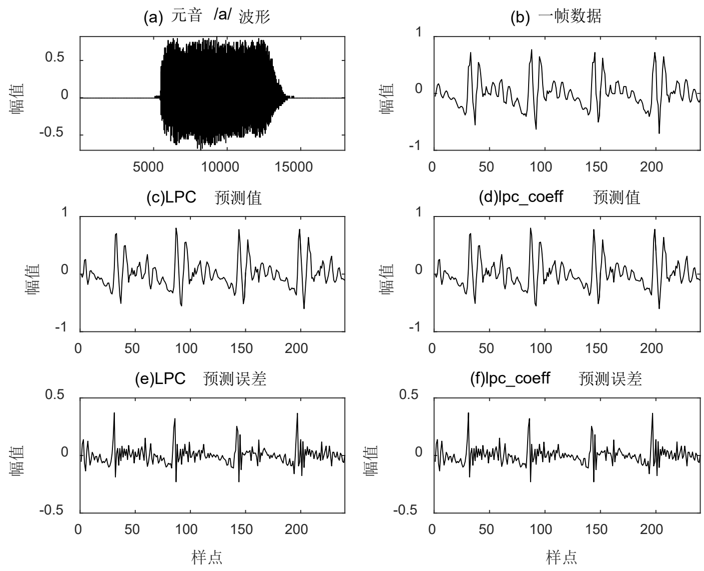

图 2-3 线性预测系数对比

函数定义如下：

名称：lpc_coeff

功能：用莱文逊-杜宾自相关法计算线性预测系数。

调用格式：

[ar,G]=lpc_coeff(s,p)

说明：输入参数 s 是一帧数据；P 是线性预测阶数。输出参数 ar 是按公式$E_i=(1-k_i^2)E_{i-1}$计算得到的预测系数$\{a_i\}$（i=1，2，•••，p），共得 p 个预测系数；G 是按公式 G=$\sqrt{E_p}$计算得到的增益系数。

2、编写求取 LPC 预测系数的复频谱函数，并与 FFT 频谱进行对比，测试效果如图 2-4 所示。

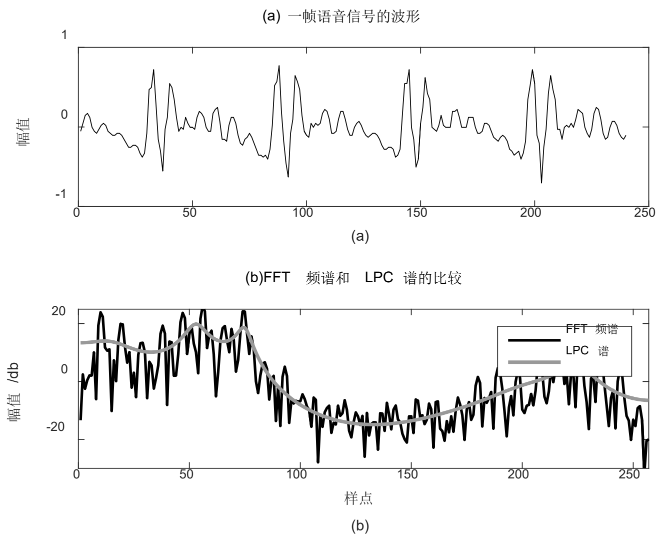

图 2-4 LPC 预测系数的复频谱与 FFT 频谱

LPC 预测系数的复频谱函数的定义如下：

名称：lpcff

功能：计算线性预测系数的复频谱。

调用格式：

```
ff=lpcff(ar,np)
```

说明：ar 是线性预测系数，np 是 FFT 阶数；输出 ff 是线性预测系数的复倒谱。

【思考题】

1、从信号对数幅度谱和 LPC 谱分别可以得到信号的那些参数？

2、对比信号对数幅度谱和 LPC 谱的差异，并说明其原因？