

# 基于matlab变声器的设计

青海民族大学物理与电子信息工程学院 刘志城 陈 超

**【摘要】**为了实现对男声变换到女声，在语音信号参数分析过程采用短时自相关法提取语音信号的基音周期，同时用LPC倒谱分析法分析共振峰的范围，通过matlab编写程序修改语音参数并接近于女声的范围，构造GUI界面。在实验中，输入一段语音信号，输出时即实现了由男声到女声的变换效果。因此对于语音信号参数的修改能够实现男声女声之间的变换。

**【关键词】**短时自相关法；LPC倒谱；语音信号；matlab；GUI

随着生活水平的提高，科技的不断进步，很多人为了娱乐，从而希望改变自己的声音；还有如今的许多的访问节目为了保护被访问者，都对声音进行了相应的处理。本设计通过编写matlab程序，修改相关声音参数，使其频率发生相应的变化，在输出时达到变声。

## 1. 变声原理

在进行性别变声时，主要考虑基音周期、基频和共振峰频率的变化。其中男生、女生和童声的基频、共振峰的关系如图1所示：基音周期改变时，基频、共振峰同时变化，若伸展既有男变女、女变童，反之亦可。本实验是基于男生录制的声音进行相关参数提取，修改接近于女声，实现男声到女声的变换。

## 2. 提取参数

### 2.1 基于短时自相关法的基音周期估值

进行自相关的计算可采用两种方法，一种是对语音信号进行低通滤波，另一种是对语音信号进行中心削波处理。本实验采用第一种方法，通过matlab编程采用自相关算法

人群	基频分布 (Hz)	共振峰频率分布
男声	[50-180]	偏低
女声	[160-380]	中
童声	[400-1000]	偏高

图1 人群中的基频和共振峰关系

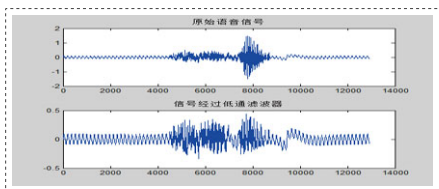


图2 原始语音信号和低通滤波后信号

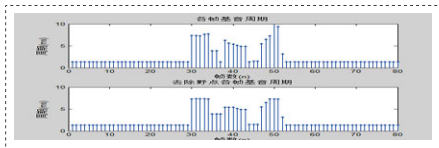


图3 基音周期和去除野点基音周期

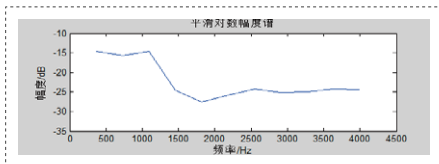


图4 幅度谱

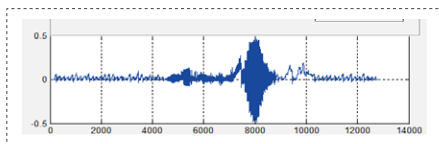


图5 合成变声语音信号波形

可以实现基音周期的估值，即对语音信号进行低通滤波，然后进行自相关计算。在低通滤波时，采用巴特沃斯滤波器。

### 2.1.1 构建巴特沃斯低通滤波器

根据人的说话特征设定相应指标参数，对本段语音设计计算出巴特沃斯模拟滤波器的阶数N为5，3dB截止频率 $\Omega_c$ ， $\Omega_c = \Omega_p(10^{0.1\alpha_p} - 1)^{-1/(N-1)}$ ，算出 $\Omega_c$ 为0.175，归一化低通原型系统函数为：

$$G_a(p) = \frac{1}{p^5 + b_4 p^4 + b_3 p^3 + b_2 p^2 + b_1 p + b_0}$$

其中  $b_0 = 1.0000$ ,  $b_1 = 3.2361$ ,  $b_2 = 5.2361$ ,  $b_3 = 5.2361$ ,  $b_4 = 3.2361$

将  $p = s/\Omega_c$  代入  $G_a(p)$  中，得到低通滤波器<sup>②</sup>

$$H_a(s) = \frac{\Omega_c^5}{s^5 + b_4 \Omega_c^4 s^4 + b_3 \Omega_c^3 s^3 + b_2 \Omega_c^2 s^2 + b_1 \Omega_c s + b_0 \Omega_c^5}$$

根据设定的滤波器编写matlab程序，当信号经过低通滤波器后，对原始信号滤波产生结果如图2所示，低通滤波后，保留基音频率，然后再用2kHz采样频率进行采样，采样序列为 $x(n)$ ，后进行下一步的自相关计算。

### 2.1.2 语音信号的短时自相关函数<sup>①</sup>

定义语音信号自相关函数如下：

$$R_x(k) = \sum_{n=0}^{N-k-1} x_n(n)x_n(n+k)$$

其中k为信号延迟点数； $x_n(n)$ 为语音信号；N为语音帧长度。经过低通滤波之后，取160个样点数，帧长取10ms，对每帧语音求短时自相关，取得自相关最大点数，自相关函数在基音周期处表现为峰值，这些峰值点之间的间隔的平均值就是基音周期，从而估计出基音周期，但是由于图中存在野点，编写matlab程序除去野点算出对应基音周期如图3所示，设基音周期值为PT，调动PT，接近女声，设新的为PT1；

### 2.2 LPC倒谱法提取共振峰

通过线性预测分析得到合成滤波器的系统函数为：

$$H(z) = \frac{1}{1 - \sum_{i=1}^p a_i z^{-i}}$$



图6 GUI界面设计

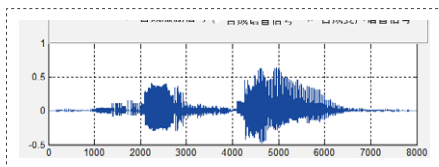


图7 仿真合成变声图

其冲激响应为 $h(n)$ ， $a_i$ 为预测系数。

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 $\hat{h}(n)$ 可以展开级数形式，即 $\hat{h}(n) = \sum_{i=1}^p a_i \hat{h}(n-i)$ ，就是说 $\hat{h}(n)$ 的逆变换 $\hat{h}(n)$ 是存在的。设 $\hat{h}(0)=0$ ，将式两端同时对 $z^{-1}$ 求导，得到：

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

有 $(1 - \sum_{i=1}^p a_i z^{-i}) \sum_{n=1}^{\infty} n \hat{h}(n) z^{-n-1} = \sum_{n=1}^{\infty} a_i \hat{h}(n-i) z^{-n-1}$ ，令式左右两边Z

的各次幂前系数分别相等得到：

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

在本实验中，取预测阶数为10，语音的样点数为160，按照上式可直接从预测系数 $\{a_i\}$ 求得倒谱 $\hat{h}(n)$ <sup>④</sup>，这个倒谱是根据线性预测模型得到，即称为LPC倒谱。编写matlab程序，实现如图4所示，谱的峰值基本上对应于共振峰频率，用GT表示共振峰；修改共振峰，使共振峰接近女声范围，新的为GT1；

## 3. 线性预测语音信号合成

根据线性预测<sup>①</sup>的基本思想，用过去M个样点值来预测<sup>⑥</sup>现在或未来的样点值：

$$\hat{y}(n) = \sum_{i=1}^M a_i y(n-i)$$

$\hat{y}(n)$ 为预测信号， $a_i$ 为预测系数， $y(n)$ 为取样信号，预测误差 $\varepsilon(n)$ ：

$$\varepsilon(n) = y(n) - \hat{y}(n) = y(n) - \sum_{i=1}^M a_i y(n-i)$$

这样就可以通过在某一个准则下使预测误差 $\varepsilon(n)$ 达到最小值的方法来决策唯一的一组线性预测参数 $a_i$  ( $i=1, 2, \dots, M$ )

把语音信号的某一帧内的短时均方预测误差定义为：

$$E\{\varepsilon^2(n)\} = E\{[y(n) - \sum_{i=1}^M a_i y(n-i)]^2\}$$

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

$$E\{[y(n) - \sum_{i=1}^M a_i y(n-i)][y(n-j) - \sum_{i=1}^M a_i y(n-i)]\} = 0 \quad j=1, \dots, M$$

上式表明采用最佳预测系数时，预测误差 $\varepsilon(n)$ 与过去的样点值正交。由于语音信号的短时平稳性，要分帧处理，对于每一个样点值记为 $y_n(n) \dots y_n(n+N-1)$ ，这段语音记为 $Y_n$ 。

对于语音段 $Y_n$ ，并记 $\Phi_n(j,i) = E\{y_n(m-j)y_n(m-i)\}$ ，对于语音段 $Y_n$ ，它的自相关函数为：

$$R_n(j) = \sum_{m=j}^{N-1} y_n(m)y_n(m-j) \quad 1, \dots, M$$

根据Yule—walker方程，可以解出样值，用这种方法定期地改变激励参数 $u(n)$ 和预测系数 $a_i$ ，并使用修改过后的基音周期和共振峰参数，就能合成出语音，合成语音样本

# 一种理解和设计传输门逻辑和静态CMOS门逻辑电路的方法

南通大学理学院 闵 意 方靖淮 仲崇贵 纪宪明 董正超

**【摘要】**在充分了解NMOS管和PMOS管电学特性的基础上,我们总结了一个理解和设计传输门逻辑和静态CMOS门逻辑电路的方法,这种方法能够简单易懂的去理解基于MOS器件的传输门和静态CMOS门逻辑电路。运用这种方法,我们也可以方便的去设计传输门和静态CMOS门逻辑电路。我们的方法将静态逻辑门电路和传输门逻辑电路有机的统一起来,便于理解学生的理解和记忆。

**【关键词】**传输门逻辑; CMOS门逻辑; NMOS管; PMOS管

## 1. 引言

21世纪是信息科学的世纪,电子科学与技术是信息科学发展的基础学科。半导体集成电路作为电子科学技术的核心,是电子类相关专业的重要基础课程。而半导体集成电路这门课程中,静态逻辑电路<sup>[1-3]</sup>这一块又是比较重要的一部分。静态逻辑电路分为静态CMOS逻辑电路和传输门逻辑电路。现有的绝大多数教材表明传输门逻辑电路的理解设计方法和静态CMOS逻辑电路的理解设计方法是不一样的。人们总结出了一套设计静态CMOS复合逻辑门电路的通用方法<sup>[3]</sup>,其步骤如下:

(1) 调整布尔代数式(也叫逻辑关系式),使得输出为负逻辑。

(2) 当逻辑关系式为“或”时,PMOS管串联,NMOS管并联。

(3) 当逻辑关系式为“与”时,PMOS管并联,NMOS管串联。

(4) 改变尺寸可调整速度或输入阈值。而对于传输门逻辑<sup>[3]</sup>,其理解和设计的方法是二义判决图BBD。这两类方法不统一,给学生在理解和设计逻辑电路造成很大的麻烦。本论文提出一种能够简单地理解传输门逻辑和静态CMOS门逻辑的方法。运用这种方法,学生也可以简单地设计传输门和静态CMOS逻辑电路。

## 2. 方法

下面我们介绍一下这个方法。我们这个方法分为如下三个部分:

(1) 对于单个NMOS管而言,漏极输出C等于源极输入A和栅极输入B的“与”。逻辑表达式为 $C = A \cdot B$ 。图1显示了单个NMOS管。根据NMOS管高电压通、低电压阻的特性,我们可以得到 $C = AB$ 。

(2) 对于单个PMOS管而言,漏极输出C等于源极输入A和栅极输入B的“非”的“与”。逻辑表达式为 $C = \bar{A}\bar{B}$ 。图2显示了单个PMOS管。根据PMOS管高电压阻、低电压通的特性,我们可以得到 $C = \bar{A}\bar{B}$ 。

(3) 对于两个MOS管并联而言,总的输出等于各个MOS输出的“或”。我们以两个NMOS管并联说明这种情况。图3显示了两个NMOS管并联的情形。我们可以得到:

如下:

$$S(n) = \sum_{i=1}^n a_i s(n-i) + G \cdot u(n)$$

$a_i$ 为预测系数; $G$ 为模型增益; $u(n)$ 为激励;合成样本 $s(n)$ ;  $p$ 为预测器阶数;

编写matlab程序,合成效果如图5所示:

## 4. GUI界面设计

在matlab新建GUI<sup>®</sup>,添加各部件,完成如图6所示。

打开一段男生录制的声音,进行仿真实验,效果如图7所示。

## 5. 结论

在实验中,用短时自相关法及LPC倒谱分析方法提取语音参数,用线性预测合成语

$$C = C1 + C2 = A1 \cdot B1 + A2 \cdot B2$$

## 3. 结果和讨论

(1) 我们先用上述的方法来理解静态CMOS逻辑电路和传输门逻辑电路。

首先看CMOS反相器,图4显示了CMOS反相器的示意图。从图4中我们可以看出对于输出 $V_{OUT}$ , PMOS管和NMOS管是并联的关系。利用上述的方法,我们可以得到PMOS管和NMOS管的漏极分别为 $V_{DD} \cdot \bar{V}_{IN}$ 和 $GND \cdot V_{IN}$ ,所以 $V_{OUT} = V_{DD} \cdot \bar{V}_{IN} + GND \cdot V_{IN} = \bar{V}_{IN}$ 。可以看出我们的方法对于理解简单的CMOS反相器逻辑是适合的。

我们再来理解一个复杂一点的两输入的异或门静态CMOS逻辑电路。图5显示了静态CMOS异或门逻辑电路示意图。运用我们的方法,我们可以得出:对于P网,有 $C1 = V_{DD} \cdot \bar{B} = \bar{B}$ 、 $C2 = V_{DD} \cdot \bar{B} = \bar{B}$ 、 $C3 = C1 \cdot \bar{A} = \bar{B} \cdot \bar{A}$ 、 $C4 = C2 \cdot \bar{A} = \bar{B} \cdot \bar{A}$ 。对于N网,有 $C7 = GND \cdot A = 0$ 、 $C8 = GND \cdot \bar{B} = 0$ 、 $C9 = C7 + C8 = 0$ 、 $C5 = C9 \cdot \bar{A} = 0$ 、 $C6 = C9 \cdot B = 0$ 。所以, $O = C3 + C4 + C5 + C6 = \bar{B} \cdot \bar{A} + \bar{B} \cdot A + 0 + 0 = \bar{B} \cdot \bar{A} + \bar{B} \cdot A$ 。可以看出我们的方法能够简单的理解较为复杂的两输入静态CMOS异或门逻辑电路。我们有理由相信对于理解更为复杂的静态CMOS逻辑电路,我们的方法同样适用。

理解完了静态CMOS逻辑电路,我们在来看看传输门逻辑电路。图6显示了一个基于CMOS传输门构成的同或门逻辑电路。运用我们的方法,我们可以得出: $C1 = \bar{B} \cdot \bar{A} = \bar{B} \cdot A$ 、 $C2 = \bar{B} \cdot A = \bar{B} \cdot A$ 、 $C3 = B \cdot \bar{A} = B \cdot \bar{A}$ 、 $C4 = B \cdot \bar{A} = B \cdot \bar{A}$ 、 $C5 = C1 + C2 = \bar{B} \cdot A + \bar{B} \cdot A = \bar{B} \cdot A$ 、 $C6 = C3 + C4 = B \cdot \bar{A} + B \cdot \bar{A} = B \cdot \bar{A}$ 、 $O = C5 + C6 = \bar{B} \cdot A + B \cdot \bar{A} = A \cdot B + \bar{A} \cdot \bar{B}$ 。因为前面我们已经理解了CMOS反相器,所以图6中我们直接用反相器逻辑功能。可以看出我们的方法可以很简单的理解传输门逻辑电路。从图6我们也可以看出,对与CMOS传输门,其逻辑功能跟随NMOS传输门。

我们再来理解一个全加器中静态的曼彻斯特进位电路。图6显示了全加器中静态的曼彻斯特进位电路示意图。运用我们的方法,可以得到:

$$OUT1 = C \cdot \bar{P} = C \cdot P, OUT2 = C \cdot P, OUT3 = V_{DD} \cdot \bar{G} = G, OUT4 = GND \cdot D = 0,$$

$$C_5 = OUT1 + OUT2 + OUT3 + OUT4 = C \cdot P + C \cdot P + G + 0 = C \cdot P + G$$

可以看出,按照我们的方法,正确的理解了全加器中静态的曼彻斯特进位电路的逻辑表达式。

音能够实现由男声到女声的变换。这种分析方法简单易行,且对参数有一个较好的把握,预测误差也会受到相应的变化,整个过程不会对合成效果产生较大的偏差,具有可行性;本实验最大不同点在于新建的LPC倒谱法,用于共振峰的提取,能够对参数比较有效的控制,减去比较复杂计算以及激励问题;整个过程成功实现由男声到女声的变换。

注释:

①张雪英.数字语音信号处理及matlab仿真[M].电子工业出版社,2011.

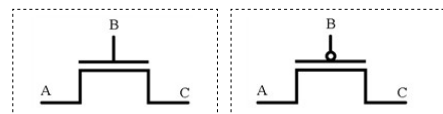


图1 NMOS管

图2 PMOS管

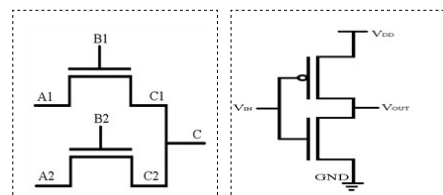


图3 两个NMOS管并联

图4 CMOS反相器的示意图

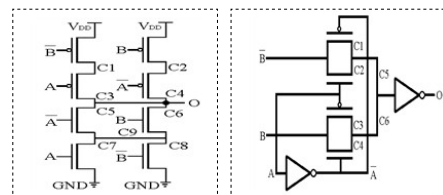


图5

图6

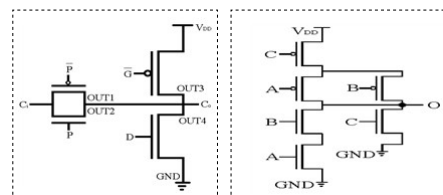


图7

图8

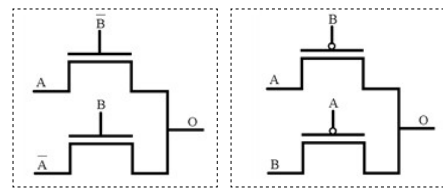


图9

图10

(2) 运用我们的方法来设计静态CMOS逻辑电路和传输门逻辑电路。

我们首先来设计一个静态CMOS逻辑电路。设计静态CMOS逻辑电路就是理解静态CMOS逻辑电路逆过程。我们以逻辑表达式

②高西全,丁玉美.数字信号处理(第三版)[M].西安电子科技大学出版社,2011.

③陈杰.Matlab宝典.电子工业出版社[M].2007.

④杨丹,姜占才,余荃良,李振超.语音信号共振峰提取方法的研究分析[C].[s.n.],2012.

⑤赵力.语音信号处理[M].机械工业出版社,2008.

⑥贺艳平.基于线性预测下的语音合成[C].西北民族大学学报,2010.

⑦陈生潭,郭宝龙,李学武,高建宁.信号与系统(第三版)[M].西安电子科技大学出版社,2008.