 

**实 验 报 告**

|  |  |
| --- | --- |
| **课程名称：** | **数字信号处理实验** |
| **学生姓名：** | **江泽群** |
| **学生学号：** | **201530371299** |
| **学生专业：** | **电子科学与技术** |
| **开课学期：** | **5** |

**电子与信息学院**

**2016年5月**

# 实验七 设计性实验 --语音信号中的基音估计

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **地 点：** | 31 楼 | 312 房； | **实验台号：** | 37 |
| **实验日期与时间：** | 2017/11/29 | | **评 分：** |  |
| **预习检查纪录：** |  | | **实验教师：** |  |

## 实验目的

1、 掌握信号处理的常用方法；

2、 培养通过查阅文献解决问题的能力。

## 实验原理

基音顾名思义就是声音的基础。这里我们主要讨论人的发声，根据声带震动的方式的不同，将声音信号分为清音和浊音。其中浊音需要声带周期性震动，所以具有明显的周期性，这种声带振动的频率称为基音频率，相应的周期就成为基音周期。但清音没有明显的周期性。

通常，基音频率与个人的声带的结构有很大的关系，所以基因频率也能用于识别发音源。一般来说，男性说话者的基音频率较低，而女性说话者和小孩的基音频率相对较高，就是通常说的“男的声音高，女的声音低”。基音周期的估计称谓基音检测，基音检测的最终目的是为了找出和声带振动频率完全一致或尽可能相吻合的轨迹曲线。

基因周期作为语音信号处理中描述激励源的重要参数之一，在语音合成、语音压缩编码、语音识别和说话人确认等领域都有着广泛而重要的问题，尤其对汉语更是如此。此处省略对汉语识别的突出贡献。

## 实验内容

（1）至少用两种方法实现语音（元音字母）的基音频率估计；

（2）估计出一个人发不同汉语语音字母发音的基音频率；

（3）估计出一个人对同一个汉语语音字母的不同声调发音的基音；

（4）估计出不同人对同一个汉语语音字母发音的基音频率

（5）总结出人的发音的一些基本的特点及规律；

## 实验主要程序

自相关法测量元音 A 基音频率的程序：

close all;clc;clear;

[data,fs] = audioread('a.mp3');

data = midValFilter(data);

frame\_len = 1000; %20\*44.1=882

frame\_num = floor(length(data)/frame\_len);

periods = zeros(1, frame\_num);

frames = zeros(frame\_num, frame\_len);

for i=1:frame\_num

frames(i, :) = data(1+(i-1)\*frame\_len : i\*frame\_len)';

periods(i) = getPeriod(myAutoCorr(frames(i, :)));

end

periods = periods ./ 44.1;

periods\_filtered = medfilt1(periods, 20);

fundamental\_freqencies = 1000 ./ periods\_filtered;

stem(fundamental\_freqencies(10:end))

title('A for autocorrelation method')

xlabel('Frame')

ylabel('Fundametal Frequency(Hz)')

程序中调用的 getPeriod 函数，根据自相关函数结果求基音周期，代码如下：

function period = getPeriod( x )

len = length(x);

period = 1000;

peaks = [];

step = 25;

up\_len = 0;

reached = false;

down\_len=0;

for i = 2:len-1

if x(i-1)<x(i) && up\_len<step

up\_len = up\_len + 1;

if up\_len==step

reached = true;

end

end

if x(i+1)<x(i) && reached && down\_len<step

down\_len = down\_len + 1;

if down\_len==step

reached = false;

up\_len = 0;

down\_len = 0;

peaks = [peaks, i-step];

end

end

end

period = max(peaks(2:end)-peaks(1:end-1));

end

谱峰法估计元音 A 基音频率的程序：

close all; clc; clear;

[data,fs] = audioread('a.mp3');

data = midValFilter(data);

N = 2^floor((log2(length(data))));

samples = data(1:N);

y=abs(fft(data,N));

[m,n]=max(y);

f=(n-1)\*fs/N;

frame\_len = 1024; %20\*44.1=882

frame\_num = floor(length(data)/frame\_len);

freqs = zeros(1, frame\_num);

frames = zeros(frame\_num, frame\_len);

for i=1:frame\_num

frames(i, :) = data(1+(i-1)\*frame\_len : i\*frame\_len)';

y=abs(fft(frames(i,:), 1024));

[~, n] = max(y);

freqs(i) = (n - 1) \* fs / 1024;

end

freqs\_filtered = medfilt1(freqs, 25);

stem(freqs\_filtered(10:end-20)) %drop first 10 frame

title('A for peak value method')

xlabel('Frame')

ylabel('Fundametal Frequency(Hz)')

## 实验结果及讨论

自相关法与谱峰法绘制元音 a的基音频率分别如图6-1、图6-2所示：



图6-1



图6-2

自相关法与谱峰法绘制元音 o的基音频率分别如图6-3、图6-4所示：



图6-3



图6-4

还有其他元音的估计方法均使用相同的程序，全部列出来篇幅较长，在此不一一列出。

自相关法与谱峰法绘制绘制改变频率的元音a的基音频率分别如图6-5、图6-6所示：



图6-5



图6-6

自相关法与谱峰法绘制绘制不同人的元音a的基音频率分别如图6-7、图6-8所示：



图6-7



图6-8

实验结论：

1. 自相关法求得的帧与帧之间的基音频率连续性较好，因此自相关法求出来的基音频率精度要比谱峰法高，但自相关法计算时间大大长于谱峰法；
2. 一个人发不同汉语语音字母大致的发音的基音频率：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 元音 | A | O | E | I | U |
| 基音频率(HZ) | 180 | 230 | 200 | 170 | 150 |

1. 提高 A 的发音音调后估计出的基音频率大概为250HZ。
2. 另一个人 A 的发音频率为大概270HZ。
3. 人发音的基音频率因人而异；不同元音之间基音频率相差较大；同一个音，采用不同发音方式可能导致基音频率相差较大。