

基于 MATLAB 的语音信号时频域参数分析

the Character Analysis of Speech Signal with Time and Frequency Based on MATLAB

李春泉 伍军云 熊 殷

Li Chunquan Wu Junyun Xiong Yin

(南昌大学信息工程学院, 江西 南昌 330031)

(School of Information Engineering, Nanchang University, Jiangxi Nanchang 330031)

摘 要: 本文针对语音信号时域、频域参数进行了系统详尽的分析, 提出了一种语音信号时频域参数模块化的思路, 并利用 SIMULINK 实现语音信号频域实时参数分析仿真。

关键词: 语音信号; MATLAB; 过零率; 短时傅立叶

中图分类号: TN912.3

文献标识码: A

文章编号: 1671-4792-(2007)9-0106-03

Abstract: This paper analyzes the characters of speech signals in time and frequency, and give some thoughts with regard to blocking the characters of speech signals. It has been realized based on MATLAB by using SIMULINK in frequency.

Keywords: Speech Signal; MATLAB; Zero Crossing Rate; STFT

0 引言

语音信号参数分析是语音信号处理的前提和基础。语音信号处理包括语音通信、语音增强、语音合成、语音识别和说话人识别等方面。只有通过语音信号的分析才能获得语音本质特性的参数, 才能利用这些参数进行高效的语音通信, 才能建立语音合成的语音库, 也才可能建立用与语音识别的模板和知识库。此外, 语音合成音质的好坏、语音识别率的高低, 都取决于语音信号参数分析的准确性和精度。因此, 语音信号参数分析是语音信号处理研究中一项非常有意义的工作^[1]。

MATLAB 是一种功能强大、效率高、交互性好的数值计算和可视化计算机高级语言, 它将数值分析、信号处理和图形显示有机地融合为一体, 形成了一个极其方便、用户界面友好的操作环境。本文就是在 MATLAB 基础上来进行语音信号参数的分析^{[2][3]}。

1 参数分析

语音信号是一种典型的非平稳信号。但是, 由于语音的形成过程是与发音器官的运动密切相关的, 这种物理运动比起声音振动速度来讲要缓慢得多, 因此语音信号常常可被假定为短时平稳的, 即在 10~20ms 这样的时间段内, 其频谱特性和某些物理特征参量可被近似地看作不变。这样, 我们

就可以采用平稳过程的分析处理方法来处理, 一般而言语音信号处理的方法都是基于这种短时平稳的假设的。根据语音信号所分析参数的不同, 语音信号参数分析可以分为时域、频域、倒谱域分析等^[4]。本文仅涉及时域及频域参数分析。

2.1 时域分析

进行语音信号最为直观的分析方法就是时域分析。语音信号本身就是时域信号, 因而时域分析是最早使用, 也是应用最广泛的一种方法, 这种方法直接利用语音信号的时域波形。时域分析通常用于最基本的参数分析以及语音的分割、预处理和大分类等。时域分析方法的特点是: 第一, 表示语音信号比较直观, 物理意义明确; 第二, 实现起来比较简单, 运算量少; 第三, 可以得到语音的一些重要参数; 第四, 采用示波器等通用设备, 使用简单^[5]。

2.1.1 短时能量分析

短时能量定义: $\text{volume} = 10 \cdot \log(\sum_{i=1}^n s_i^2)$ 。

短时能量分析用途: 第一, 可以区分清音段和浊音段, 因为浊音时的短时平均能量值比清音时大得多; 第二, 可以用来区分声母与韵母的分界、无声与有声的分界、连字的分界等。如对于高信噪比的语音信号, 短时平均能量用来区分有无语音。无语音信号噪声的短时平均能量很小, 而有语音信号的能量则显著增大到某一个数值, 由此可以区分语音信号的开始点或者终止点。下面给出了两种实现短时能量分析的 MATLAB 代码, 图一中显示了分析结果。

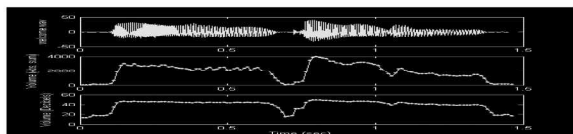
`function [y, fs, nbits]=wavReadInt(waveFile)`

基金资助: 南昌大学自然科学基金资助项目, 基金代码 (Z-03709)

```

[y, fs, nbits]=wavread(waveFile);y=y*2^nbits/2;
% waveReadInt: Same as wavread, but return integer value of y
% Usage: [y, fs, nbits]=wavReadInt(waveFile)
% If nbits=8, then -128<=y<=127
% If nbits=16, then -32768<=y<=32767
waveFile='lcq.wav'; frameSize=256; overlap=128;
[y, fs, nbits]=wavReadInt(waveFile);
fprintf('Length of %s is %g sec.\n', waveFile, length(y)/fs);
frameMat=buffer(y, frameSize, overlap);
frameNum=size(frameMat, 2);volume1=zeros(frameNum, 1);
volume2=zeros(frameNum, 1);
for i=1:frameNum
    frame=frameMat(:,i);
    frame=frame-mean(frame);
% zero-justified
    volume1(i)=sum(abs(frame));
% method 1
    volume2(i)=10*log10(sum(frame.^2));
% method 2
end
time=(1:length(y))/fs;
frameTime=((0:frameNum-1)*(frameSize-overlap)+0.5*frameSize)/fs;
subplot(3,1,1); plot(time, y); ylabel(waveFile);
subplot(3,1,2); plot(frameTime, volume1, '-');
ylabel('Volume (Abs. sum)');
subplot(3,1,3); plot(frameTime, volume2, '-');
ylabel('Volume (Decibels)');
xlabel('Time (sec)');

```



图一 两种短时能量分析结果

2.1.2 短时过零率分析

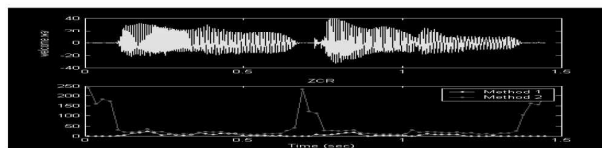
过零就是信号通过零值。对于连续语音信号,可以考察其时域波形通过时间轴的情况。对于离散时间信号,如果相邻的取样值改变符号则称为过零。由此可以计算过零数,过零数就是样本改变符号的次数。单位时间内的过零数称为平均过零数。语音信号 $x(n)$ 的短时平均过零数定义为: $Z_n = S\{ \text{abs}\{\text{sgn}[x(n)] - \text{sgn}[x(n-1)]\} \cdot w(n-m)\}$ 。短时过零分析通常用在端点检测,特别是用来估计清音的起始位置和结束位置。

下面给出了计算过零率两种方式的MATLAB代码,图二显示了分析结果。

```

waveFile='lcq'; frameSize=256; overlap=0;
[y, fs, nbits]=wavReadInt(waveFile);
frameMat=buffer(y, frameSize, overlap);
zcr1=sum(frameMat(1:end-1,:).*frameMat(2:end,:))<0); % Method 1
zcr2=sum(frameMat(1:end-1,:).*frameMat(2:end,:))<=0); % Method 2
time=(1:length(y))/fs; frameNum=size(frameMat, 2);
frameTime=((0:frameNum-1)*(frameSize-overlap)+0.5*frameSize)/fs;
subplot(2,1,1); plot(time, y); ylabel(waveFile);
subplot(2,1,2); plot(frameTime, zcr1, '-');
frameTime, zcr2, '-');
title('ZCR'); xlabel('Time (sec)'); legend('Method 1', 'Method 2');

```



图二 两种短时零率分析结果

从图二可以清楚地看出,清音的过零率要大于浊音的过零率。

2.1.3 基音的估计

浊音信号的周期称为基音周期,它是声带振动频率的倒数,基音周期的估计称为基音检测。基音检测是语音处理中的一项重要技术,它在有调语音辨意、低速率语音编码、说话人识别等方面起着非常关键的作用。但在实现过程中,由于声门激励波形不是一个完全的周期脉冲串,再加上声道影响去除不易、基音周期定位困难、背景噪声影响强烈等一系列因素,基音检测面临着很大的困难。现在已有很多性能优越的基音检测算法,自相关基音检测算法就是一种基于语音时域分析理论较好的算法^[6]。在这里基于声音文件比较稳定的基础上,使用观察法获取基音周期,下面给出了实现代码,图三给出了直观的结果。

```

waveFile='lcq01.wav'; [y, fs, nbits]=wavread(waveFile);

```

```

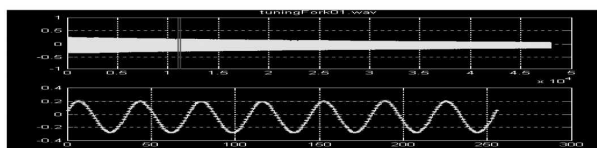
index1=11000; frameSize=256;
index2=index1+frameSize-1;
frame=y(index1:index2); subplot(2,1,1); plot(y);
grid on title(waveFile); line(index1*[1 1], [-1 1], 'color', 'r');
line(index2*[1 1], [-1 1], 'color', 'r'); subplot(2,1,2); plot(frame, '-');
grid on point=[7, 189];

```

2.2 频域分析

短时傅立叶分析

在运用离散时间傅立叶变换分析语音信号的变化时,会



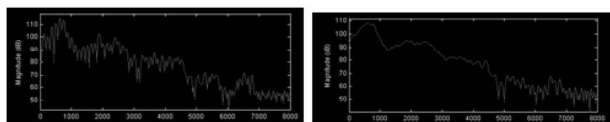
图三 基音波形

遇到这样的问题,即单一的傅立叶变换并不能反映时间变化的频谱信息,诸如时变共振峰和谐波。具体而言,通常将信号的每一时刻与其相邻时刻信号的傅立叶变换相联系,这样就可以及时跟踪信号的频谱变化。信号 $\{x(n)\}$ 的短时傅立叶变换如下:

$$X_n(\omega) = \sum_{m=-\infty}^{\infty} x(m)w(m-n) e^{-j\omega m}$$

可以验证,在短时傅立叶分析中对于同一种窗函数而言,其通带宽度与窗长成反比。如果希望频率分辨率高,则窗长应尽量取长一些;如果希望时间分辨率高,则窗长尽量取短一些。由此可见,傅立叶分析的时间分辨率和频率分辨率是相互矛盾的,这是短时傅立叶本身所固有的弱点。短时傅立叶分析一般采用汉明窗作为分析窗^[7]。

图四中从左往右分别是使用20ms的汉明窗和10ms的汉明窗的短时傅立叶分析图。

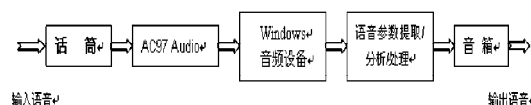


图四 短时傅立叶分析图

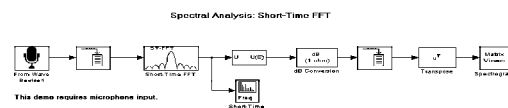
通过基于MATLAB和短时频域分析,并结合图四,能够得出^[8]:第一,长窗具有较高的频率分辨率,但具有较低的时间分辨率。从一个周期到另一个周期,共振峰是要发生变化的,这一点即使从语音波形上也能够看出来。然而,如果采用较长的窗,这种变化就模糊了,因为长窗起到了时间上的平均作用。第二,短窗的频率分辨率低,但具有较高的时间分辨率。采用短窗时,能够从短时频谱中提取出共振峰从一个周期到另一个周期所发生的变化。当然,激励源的谐波结构也从短时频谱上消失了。第三,在对语音信号进行短时傅里叶分析时,窗长需要折衷考虑。一方面,短窗具有较好的时间分辨率因而能够提取出语音信号中的短时变化;但另一方面,损失了频率分辨率。还应当注意到,语音信号的基音周期是有一个覆盖范围的(例如,女性和儿童的基音周期就比成年男性短得多),因此,窗宽的选择应当考虑到这个因素。第四,汉明窗都具有低通的性质,且在截止频率处比较尖锐,当其通带较窄时(窗越宽,通带越窄),加窗后的频谱更能够较好反映短时语音信号的频谱,窗越宽这种逼近越好。

2 语音信号时频参数提取平台设计(图五)

通过对上述语音信号时频参数的分析,可以把常用的语音信号特征提取的技术方法模块化,综合形成一个语音参数提取分析处理的平台。这样,便于今后进一步语音信号处理各种层面的研究。下面给出一个由SIMULINK实现的语音实时平台模块(图六)。



图五 语音参数提取分析处理平台



图六 语音实时平台模块

3 结束语

语音处理是目前信息学领域的研究热点和难点,其任务是研究如何利用信号处理技术研究语音信号,使未来的计算机“能听会说”。语音信号的处理是对语音信号进行分析,只有通过分析得到参数才能做进一步的研究。本文从语音信号时域和频域角度出发着重分析了语音信号时频域特征参数及其提取实现方法,并提出了参数平台设计模块化的思路。

参考文献

- [1]王炳锡.语音编码[M].西安:西安电子科技大学出版社,2002.
- [2]谢斌,蔡虔,钟文涛.基于MATLAB/SIMULINK的QDPSK通信系统仿真[J].科技广场,2006,11.
- [3]陈家焱,陈冬娇,张达响.基于Matlab的声音信号采集与分析处理[J].计算机与现代化,2005,5:91~92.
- [4]易克初,等.语音信号处理[M].北京:国防工业出版社,2006,6.
- [5]胡航.语音信号处理[M].哈尔滨:哈尔滨工业大学出版社,2000,5.
- [6]付青青,吴爱平.基于MATLAB的语音信号自相关基音检测[J].长江大学学报,2006,03(4):99-101.
- [7]胡广书.数字信号处理理论、算法与实现[M].北京:清华大学出版社,1997.
- [8]王炳锡,等.实用语音识别基础[M].北京:国防工业出版社,2005.

作者简介

李春泉(1980—),男,汉族,助教,硕士,江西南昌大学信息工程学院电子系(前湖校区),主要研究方向:信号与信息处理。