

MATLAB在数字语音分析中的应用

张 勇, 柳玉海

(河北省广播电视技术中心, 河北 石家庄 050012)

摘 要: 在数字语音信号分析中使用MATLAB可以简化程序编写和实验构建的过程,文章重点研究了MATLAB语言应用于数字语音信号分析的优越性,分析了MATLAB的语言功能,并对基于MATLAB的语音时域分析、频域分析、时频分析的结合进行了应用分析。

关键词: 语音处理; LPC分析; MATLAB; 交互式图形用户界面

Application of MATLAB in Digital Speech Analysis

ZHANG Yong, LIU Yu-hai

(Radio & TV Technology Center of Hebei Province, Shijiazhuang, Hebei 050012, China)

Abstract: In digital speech signal analysis, using MATLAB can simplify programming and the process of constructing experiment. The advantages of MATLAB language applied in digital speech signal analysis are emphatically studied, the language functions of MATLAB are analyzed, and the time-domain analysis, frequency domain analysis, and joint time-frequency analysis of speech based on MATLAB are applied and analyzed.

Key words: speech processing; LPC analysis; MATLAB; interactive GUI

0 引言

MATLAB是Matrix Laboratory的缩写,由MathWorks公司于1984年正式推出,内核采用C语言编写。MATLAB是一个包括数值计算、高级图形和可视化的集成科技计算环境,也是一种高级程序设计语言。灵活的MATLAB语言可使工程师和科学家简练地表达他们的思想,其强有力的数值计算方法和图形便于测试和探索新的思想,而集成的计算环境便于产生快速的实时结果。MATLAB得到了各个领域专家学者的广泛关注,其强大的扩展功能为用户提供了强有力的支持;它集数学计算、图形计算、语言设计和神经网络等30多个工具箱于一体,具有极高的编程效率^[1],极大地方便了科学研究和工程应用。

语音处理中往往把数字化的语音信号表示为一维或二维(对应于双声道立体声数据)矩阵,因此基于矩阵运算的MATLAB就很自然地应用到语音处理领域。

1 MATLAB在语音处理中的应用

1.1 基本的编程运算

MATLAB提供了语音文件的读写函数以及录音和放音功能,如表1所示,使用时只需按照函数的语法规则正确输入参数即可。通过这些函数可以得到语音的采样频率、量化精度和通道数等参数。同时,MATLAB提供了语音的和、差等线性运算,以及卷积、相关等非线性运算。对于语音处理中常用到的各种窗函数,MATLAB也都提供了相应的函数,比如hamming(n)即长度为n点的汉明窗。

MATLAB一个重要的特点是易扩展性。近几年来,有许多科学家、数学家、工程师等开发了一些新的、有价值的应用程序,这些应用程序都可以被纳入MATLAB工具箱。比如voice box工具箱,其中包含了很多与语音信号处理相关的函数,可以

在有关网站上下载,将其加入到MATLAB的搜索路径,就可以作为MATLAB的库函数来方便地调用了。

表1 与语音输入输出相关的函数

函数名	功能描述
wavread auread	读语音文件
wavwrite auwrite	写语音文件
sound wavplay	放音
soundsc	归一化放音
soundview	可视化语音输出
wavrecord	录音

下面以语音特征参数提取为例来说明MATLAB的应用。

首先,读入语音文件,并返回采样频率fs:

```
[x,fs]=wavread('example.wav');
```

为了提升信号频谱的高频部分,在计算语音参数之前,一般要将信号通过一个预加重滤波器 $H(z)=1-\mu z^{-1}$ ($0.9 < \mu < 1.0$)的滤波,即: $xx=filter([1-\mu],1,x)$;

然后,采用voice box工具箱中的enframe函数:

```
f=enframe(xx,len,inc);
```

其中len指定帧长,inc指定帧移,矩阵f的每一行是一帧语音数据。

接下来,可以针对每一个短时语音帧来提取特征参数。线性预测分析技术^[2]是目前广泛应用的特征参数提取技术。MATLAB中提供了计算LPC(Linear Predictive coding,线性预测编码)系数的函数,名为lpc: $a=lpc(x,n)$ 。这里x为一帧语音信号,n为计算LPC参数的阶数。在语音识别系统中,很少直接使用LPC系数,而是由LPC系数推导出另一种参数:线性预测倒谱系数。在MATLAB中也提供了用于倒谱分析的函数,如cceps,rceps等。近年来,人们模拟人耳对不同频率语音的感知

特性,提出了 Mel 频率倒谱参数^[2](MFCC),MATLAB 中的函数 melcepst 就可直接计算语音信号的 MFCC 参数。

使用 MATLAB 编程与人们进行科学计算的思路和表达方式相近,不需要大量原始而传统的编程过程,要解决很多复杂的问题只需要短短的几条命令就能实现;研究人员可直观、方便地进行分析、计算及设计工作,从而能节省大量时间。

1.2 时域分析

MATLAB 具有很强的绘图功能,它有一系列的绘图函数,可以方便地实现函数或数据的可视化显示。直接使用 MATLAB 中的曲线绘图命令 plot 就可将读入的语音信号的波形表示出来。同时,通过计算语音的能量、过零率等一系列参数可以对语音信号进行时域分析。图 1 和图 2 给出了用 MATLAB 画出的时域波形和能量、过零率的对比,语句内容为女性讲话“同舟共济”。

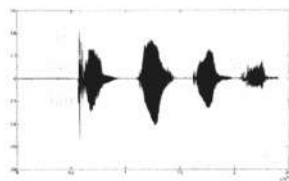


图1 语音时域波形

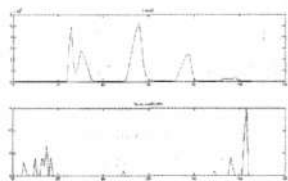


图2 能量、过零率对比

1.3 频域分析

如果利用 fft(离散傅立叶变换)或者 rfft(实数的离散傅立叶变换)函数先计算出语音信号的傅立叶变换,再将其显示出来,便得到了相应的语音频谱。可以针对每一个短时语音帧来提取特征参数以对语音进行频域分析,利用 LPC 分析技术对语音特征进行表述^[2]。图 3 给出了语音的频谱包络以及 LPC 分析的极点(12 点),语句内容为女性讲话“同舟共济”。

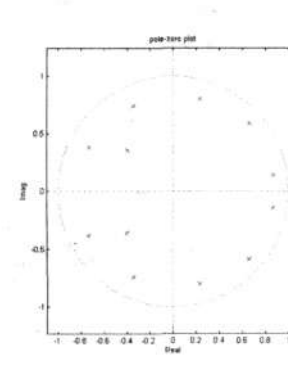
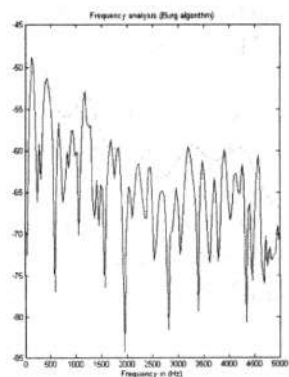


图3 语音频谱及LPC分析

1.4 时频分析的结合

语音的时域分析和频域分析是语音分析的两种重要方法,但是这两种单独分析的方法均有局限性。人们致力于研究语音的时频分析特性,把和时序相关的傅立叶分析的显示图形称为语谱图。借助于 MATLAB 中的 specgram 函数(时频分析函数)和 imagesc 函数(图像显示函数),我们可以方便地得到语音信号的语谱图。具体使用时可在 specgram 函数中设定语音分析的帧长、帧移、窗函数和 FFT 运算点数等参数^[3]。图 4 和图 5 给出了用 MATLAB 画出的宽带和窄带语谱图,语句内容为男

性讲话“同舟共济”。图中横轴为时间,纵轴为频率,任一给定频率成分在给定时刻的强弱用相应点的颜色表示(原图为彩色图像)。窄带语谱图具有良好的频率分辨率,有利于观察基音频率及其各次谐波,但它的时间分辨率较差;而宽带语谱图正相反,具有良好的时间分辨率及较差的频率分辨率,能观察出语音的共振峰及清辅音的能量汇集区。另外,我们还可以把语谱图看作是一幅图像^[4],利用 MATLAB 中提供的图像处理函数来对其进行研究。语谱图上的纹路,称之为“声纹”,它因人而异,所以可以利用声纹来鉴别不同的讲话人。虽然对基于语谱图的识别技术的可靠性还存在相当大的怀疑,但目前这一技术已在司法领域得到某些认可并被采用^[5]。

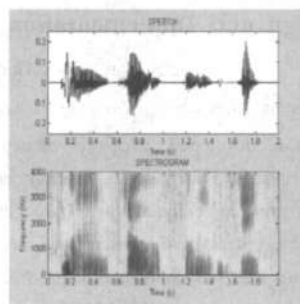


图4 宽带语谱图

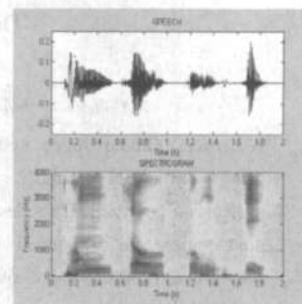


图5 窄带语谱图

2 交互式 GUI 工具

MATLAB 提供了多功能的交互式 GUI(图形用户界面)工具, SPTool^[6]就是其中之一。在 MATLAB 命令窗口中键入 sptool,立即弹出一个 SPTool 窗口。SPTool 窗口是交互式信号分析处理环境,通过这一窗口,用户无须详细了解 MATLAB 信号处理工具箱函数的语法规则,就可以可视化地进行大部分信号分析处理。窗口有信号(Signal)、滤波器(Filter)、频谱(Spectra)三栏,用户能够使用鼠标在计算机屏幕上控制数字语音信号的观察、测量和时域分析,观测滤波器的时域和频域特性、设计和实现数字滤波器以及观察信号的频谱,分析信号的频率成分等。

除此之外, MATLAB 中还提供了数字滤波器设计分析工具 FDAtool、窗函数设计工具 Wintool 和随机噪声生成器 RANDTool 等交互式 GUI 工具。

3 结束语

由以上的讨论可以得知,在语音信号处理中, MATLAB 的确是一个得力的科研助手,它提供了方便高效的编程环境。MATLAB 语言的设计思想可以说代表了当前计算机的发展方向,在今后的使用中,我们相信会发现它更大的潜力。

参考文献:

- [1] 何强,何英. MATLAB 扩展编程[M]. 清华大学出版社, 2002.
- [2] 罗宾纳. 语音识别基本原理(影印版)[M]. 清华大学出版社, 1999.
- [3] 赵红怡,张常年. 数字信号处理及其 MATLAB 实现[M]. 化学工业出版社, 2002.
- [4] M. Ahmadi, N. J. Bailey, B. S. Hoyle. Phoneme recognition using speech image[J]. IEEE Trans Signal Processing, 1996. 1:675
- [5] 赵力. 语音信号处理[M]. 机械工业出版社, 2003.
- [6] 黄文梅,熊桂林,杨勇. 信号分析与处理: MATLAB 语言及应用[M]. 国防科技大学出版社, 2000.

