

A Time-Domain Speech Signal Analysis System Based on MATLAB

Jun Li

Department of Electronic Engineering, Hubei University of Economic, Wuhan
Email: lijun@hbue.edu.cn

Received: Nov. 1st, 2013; revised: Nov. 15th, 2013; accepted: Nov. 21st, 2013

Copyright © 2013 Jun Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract: A speech signal analysis and processing system is realized in this paper under the Matlab environment. Signal acquisition and playback are also achieved in the system and a lot of time-domain parameters can be obtained while signal analysis is processing, such as short-term signal zero-crossing rate, short-time energy, short-time auto-correlation function, and so on. Powerful numerical calculation, signal processing, and graphics display capabilities of Matlab are fully used in this system. It is easy to use, and it has the versatility and other advantages.

Keywords: Matlab; Speech Signal Analysis; Signal Processing

基于 Matlab 的语音信号时域分析系统

李俊

湖北经济学院电子工程系, 武汉
Email: lijun@hbue.edu.cn

收稿日期: 2013 年 11 月 1 日; 修回日期: 2013 年 11 月 15 日; 录用日期: 2013 年 11 月 21 日

摘要: 本文设计了一个语音信号分析处理系统, 并在 Matlab 环境下进行了实现。该系统能够实现信号的采集与播放, 同时对语音信号进行时域分析与处理, 获得信号的短时过零率, 短时能量、短时自相关函数等参数。该系统充分利用了 Matlab 强大的数值计算、信号处理、图形显示等能力, 具有使用方便, 通用性强等优点。

关键词: Matlab; 语音信号分析; 信号处理

1. 引言

语音信号分析是一门新兴的涉及到很多领域的交叉学科^[1,2]。伴随着通信技术的发展, 语音采集和分析仪器的数字化、智能化、小型化和多功能化的发展越来越快, 其分析速度也有了大幅度的提高^[3]。但现在市场上的语音分析系统价格昂贵, 且操作复杂, 大多只能用于特定的测量分析项目^[4-6]。基于上述不足, 本文设计了一款基于 Matlab 的语音信号分析系统, 能够方便地对语音信号进行时域分析, 具有价格便宜、

通用性强等优点。

2. 语音处理系统的设计

2.1. 语音信号的采集过程与预处理

语音信号作为一维模拟信号, 利用计算机来分析处理时, 应先进行采样和量化, 转换成数字信号。Matlab 中提供了强大的数据采集工具箱, 可满足控制声卡对数据进行采集的要求, 通过如下方法可以驱动声卡来进行: 一是将声卡作为对象处理采集语音信

号；二是调用 Wavrecord 功能函数采集语音信号；三是运用 Audiorecorder 对象采集语音信号。在实际的语音信号处理时，为了减少信息的丢失，采样频率一般根据采样定理选取为 8 kHz~10 kHz。采样后的信号在幅度上仍然保持着连续的特点，所以要进行量化处理。量化的目的是将信号波形的幅度值离散化。

作为短时平稳信号，语音信号在时域频域分析时，对数据要进行分帧加窗处理。一个好的窗函数的标准是：在时域因为是语音波形乘以窗函数，所以要减小时间窗两端的坡度，使窗口边缘两端不引起急剧变化而平滑过渡到零，这样可以使截出的语音波形缓慢降为零，减小语音帧的阶段效应；在频域要有较宽的 3 dB 带宽以及较小的边带最大值。另外一方面，不同人的基音周期变化很大，主要集中在 70 Hz~1000 Hz。所以窗一般选择在 10 ms~20 ms 之间。此外，为了避免在加窗时加入多余的高频信号，应该使用平滑窗，例如汉明窗。

2.2. 语音信号的时域分析

语音信号典型的时域特征包括短时能量、短时平均过零率、短时自相关系数和短时平均幅度差等参数。

2.2.1. 短时能量分析

语音信号的能量分析是基于语音信号能量随时间有相当大的变化，特别是清音段的能量一般比浊音段的小得多。对语音信号采用短时分析时，信号流的处理用分段或分帧来实现。用可移动的有限长度窗口进行加权的方法来实现。窗口可采用矩形窗。信号 $\{x(n)\}$ 的短时能量定义为：

$$E_n = \sum_{m=-\infty}^{\infty} [x(m) \cdot w(n-m)]^2$$

令则有 $h(n) = w^2(n)$ ，则有

$$E_n = \sum_{m=-\infty}^{\infty} x(m)^2 \cdot h(n-m)$$

此式的含义可以用低通滤波作用来解释， $h(n)$ 是低通滤波器的单位冲击响应语音信号的短时平均幅度定义为：

$$M_n = \sum_{m=-\infty}^{\infty} |x(m)| \cdot w(n-m)$$

2.2.2. 短时平均过零率

信号的过零率是表征信号变化快慢的一个标量，是频率的一种简单的度量。过零率主要用于粗略地描述信号的频谱特性，这就是用多带滤波器将信号分为若干个通道，对各通道进行短时平均过零率和短时能量的计算，即可粗略地估计频谱特性。另外还能用于判别清音和浊音、有话与无话，其中清音的短时过零率高，而浊音的短时过零率低；有语音的短时过零率高，无语音的短时过零率低。信号 $\{x(n)\}$ 的短时平均过零率定义为：

$$Z_n = \sum_{m=-\infty}^{\infty} |\text{sgn}[x(n)] - \text{sgn}[x(n-1)]| \cdot w(n-m)$$

一般取

$$w(n) = \begin{cases} \frac{1}{2N} & 0 \leq n \leq N-1 \\ 0 & \text{其他} \end{cases}$$

2.2.3. 短时自相关函数

信号 $\{x(n)\}$ 的短时自相关函数 $R_n(k)$ 就是在信号的第 n 个样本附近用短时窗截取一段信号做自相关计算所得结果，定义为：

$$R_n(k) = \sum_{m=-\infty}^{\infty} x(m) \cdot x(m+k) \cdot h_k(n-m)$$

其中 $h_k(n) = w(n) \cdot w(n-k)$ 。

3. 语音处理系统的实现

3.1. 语音信号的采集与播放

Matlab 提供了强大的信号处理工具箱，可以非常方便地用来进行语音信号分析的采集与播放。本系统直接调用了直接调用函数的方法来进行信号实时的采集与播放。部分源代码如下：

```
fs=11025; % 取样频率
duration=3; % 录音时间
fprintf('Press any key to start %g seconds of recording...\n',duration);
pause;
fprintf('Recording...\n');
y=wavrecord(duration*fs,fs); %duration*fs 是总的采样点数
fprintf('Finished recording.\n');
```

```
fprintf('Press any key to play the recording...\n');
pause;
```

3.2. 短时能量

语音信号的能量随着时间变化比较明显，一般清音部分的能量比浊音的能量小得多。语音信号的短时能量分析给出了反应这些幅度变化的一个合适的描述方法。短时能量随时间分布的结果如图 1 所示。

由图 1 可见，在用短时能量反映语音信号的幅度变化时，不同的窗函数以及相应窗的长短均有影响。窗的长短影响起决定性作用。窗过大如 $N = 512$ 时，等效于很窄的低通滤波器，不能反映幅度 En 的变化；窗过小如 $N = 32$ 时，短时能量随时间急剧变化，不能得到平滑的能量函数。

短时能量函数的应用：1) 可用于区分清音段与浊音段。 En 值大对应于浊音段，如图 1 中采样点 10,000~20,000 段； En 值小对应于清音段，如图 1 中采样点

20,000~40,000 段。2) 可用于区分浊音变为清音或清音变为浊音的时间，如图 1 中采样点 15,000 附近能量开始逐渐减小。

3.3. 短时平均过零率

图 2 给出了一小段语音信号的短时平均过零率随着时间的分布。

短时过零率可以粗略估计语音的频谱特性。由语音的产生模型可知，发浊音时，声带振动，尽管声道有多个共振峰，但由于声门波引起了频谱的高频衰落，因此浊音能量集中于 3 kHz 以下。而清音由于声带不振动，声道的某些部位阻塞气流产生类白噪声，多数能量集中在较高频率上。高频率对应着高过零率，低频率对应着低过零率，那么过零率与语音的清音浊音就存在着对应关系，如图 2 中样本的过零率结果可以看出，在采样点的 0~50 段过零率较高，是清音段，而从 50 开始过零率锐减，可以看出开始为浊音段。

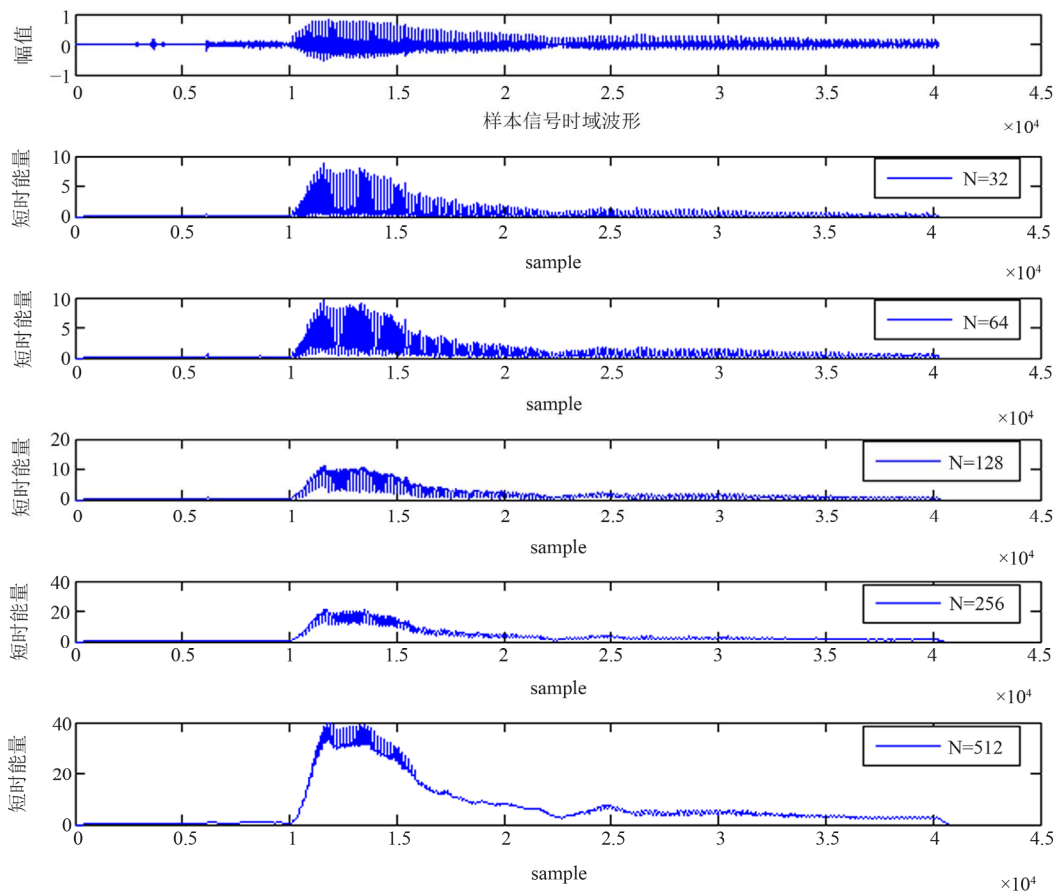


Figure 1. Distribution of temporal energy when N at different values
图 1. N 为不同取值时短时能量分布

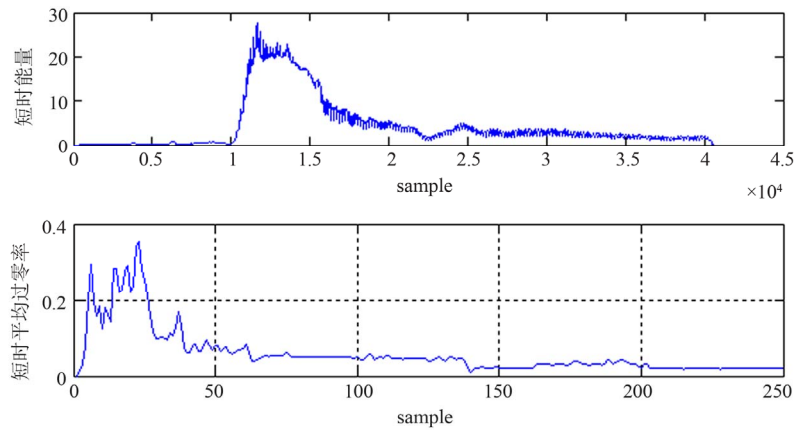


Figure 2. Realization of the short time zero cross rate
图 2. 短时平均过零率实现

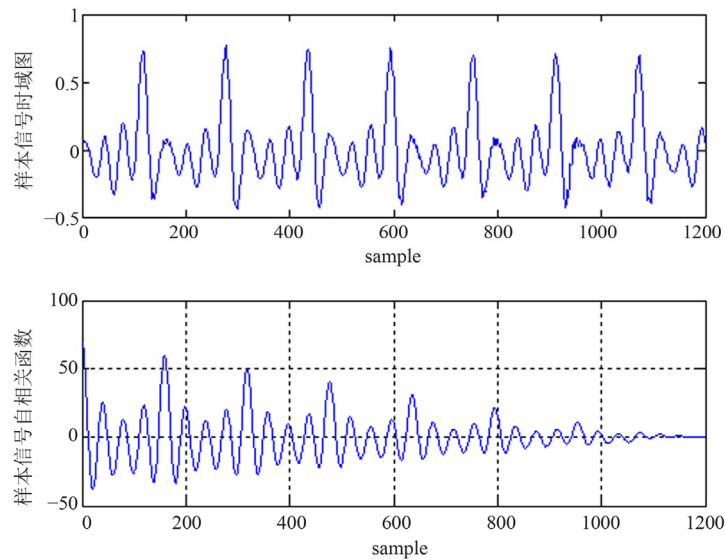


Figure 3. Realization of the Short time autocorrelation function
图 3. 短时自相关函数实现

另外，可以将短时平均过零率与短时能量结合起来判断语音起止点的位置，即进行断点检测。在背景噪声较小的情况下，短时能量比较准确，但当背景噪声较大时，短时平均过零率可以获得较好的检测效果。因此一般的识别系统，其前端的端点检测都是将这两个参数结合用于检测语音是否真的开始。

短时平均过零率的另一个用途是作为语音频域分析的一个中间步骤。方法是不用窗口型的低通滤波器来处理过零，而改用多通道的带通滤波器，这是的输出就是频域的短时平均过零率，如果再加上用带通滤波器的短时能量的输出，就可以得到语音信号的频域分析结果。

3.4. 短时自相关函数

一般情况下，相关函数用于测定两个信号在时域内的相似程度，自相关函数主要研究信号本身的同步性，周期性。短时自相关函数是在信号的第 n 个样本点附近用短时窗截取一段信号，做自相关计算所得的结果。具体实现结果如图 3。

自相关函数可以用于衡量信号自身时间波形的相似性，常用来区分语音是清音还是浊音，或估计浊音语音信号的基音周期。清音和浊音的发声机理不同，因而在波形上也存在着较大的差异。浊音的时间波形呈现出一定的周期性，波形之间相似性较好；清音的时间波形呈现出随机噪声的特性，样点间的相似

性较差。从上图选择的 1200 个样本(约 1 帧)的自相关函数曲线连续性较好,可以判定是浊音段。

4. 结论

本系统运用 Matlab 软件实现了对语音信号的分析 and 处理,分别对语音信号的时域与频域进行了分析,能够获得信号的短时过零率,短时能量、短时自相关函数、频谱强度等参数。充分利用了 Matlab 强大的数值计算功能、方便的声卡控制功能和信号分析能力,完成了设计的目的。该系统在语音信号的采集与分析的实际应用中取得了良好的效果,具有较好的应用前景和发展潜力。

参考文献 (References)

- [1] 张雪英 (2010)数字语音处理及 MATLAB 仿真. 电子工业出版社, 北京.
- [2] 赵力 (2011) 语音信号处理. 机械工业出版社, 北京.
- [3] 陈家焱, 陈冬娇, 张达响 (2005) 基于 Matlab 的声音信号采集与分析处理. *计算机与现代化*, **6**, 91-96.
- [4] 李春泉, 伍军云, 熊殷 (2007) 基于 MATLAB 的语音信号时频域参数分析. *科技广场*, **9**, 22-24.
- [5] 周渊, 王炳和, 刘斌胜 (2004) 基于 MATLAB 的噪声信号采集与分析系统的设计. *电声技术*, **7**, 52-54.
- [6] 陈宇峰 (2006) 基于声卡和 Matlab 的语音信号采集和处理. *中国现代教育装备*, **5**, 49-56.