

基于 MATLAB 的语谱图显示与分析

The Display and Analysis of Sonogram Based on MATLAB

(1.郑州大学; 2.郑州解放军信息工程大学) 李富强¹ 万红^{1,2} 黄俊杰¹

Li, Fuqiang Wan, Hong Huang, Junjie

摘要: 文章给出了一种基于 MATLAB 的语谱图显示原理及实现方法。该法可对语音信号(或其它类型的似平稳信号)进行频谱图的伪彩色映射及显示。语谱图的类型有宽带和窄带两种, 可按需求设置伪彩色显示的映射。文中介绍了语谱图显示的原理, MATLAB 的相关功能函数, 并给出了一个程序实例及其实验结果。在 MATLAB 中进行了仿真, 结果表明宽带及窄带语谱图对时间分辨率和频率分辨率的不同反映, 并且该法对设备要求低, 编程简单, 可以在普通通用计算机上实现。

关键词: 语谱图; 宽带语谱图; 窄带语谱图; 伪彩色显示; MATLAB

中图分类号: TN911.72

文献标识码: A

文章编号: 1008-0570(2005)10-3-0172-03

Abstract: This article introduces the display principle and realization of spectrogram based on MATLAB. The spectrogram of speech signals and other similar quasi-stationary ones can be done by this method. There are two kinds of spectrogram: wide-band and narrow-band spectrograms. In addition, different pseudo-color-mapping displays are optional. It covers the display principle of spectrogram, the correlative functions of MATLAB and the source program, as well as the experimental results. The simulation in MATLAB turns out that wide-band and narrow-band spectrograms represent different degrees of time and frequency resolution, respectively. This method can be implemented easily without complicated programming based on universal PC.

Keywords: spectrogram; wide-band spectrogram; narrow-band spectrogram; pseudo-color display; MATLAB

1 语谱图

语音的发音过程中, 声道通常都是处于运动状态的, 因此它的共振峰特性也是时变的。不过这个时变过程比起振动过程来说要缓慢得多, 因此一般可以假定它是短时平稳的, 每一时刻我们都可以用这时刻附近的一短段(例如 15ms)语音信号分析得到一种频谱。对语音信号连续地进行频谱分析就可以得到一种二维图谱, 其横坐标表示时间, 纵坐标表示频率, 而每像素的灰度值大小反映相应时刻和相应频率的信号能量密度。这种时频图称为语谱图 (Sonogram 或 Spectrogram), 这种反映语音信号动态频谱特性的时频图在语音分析中有重要实用价值, 被视为可视语言。从语

谱图上不仅能看出任一时刻发音器官的共振峰特征, 而且可以看出语音的基音频率, 是否清音、爆破音等。语谱分析在语音识别、合成及编码中很有意义。

1.1 语谱图的产生机理

语音信号是一种典型的非平稳信号, 但是其非平稳性是由发音器官的物理运动过程而产生的, 此过程与声波振动的速度相比较缓慢, 可以假定在 10~30ms 这样的短时间内是平稳的。傅立叶分析是分析线性系统和平稳信号稳态特性的强有力的手段, 而短时傅里叶分析, 也叫时间依赖傅立叶变换, 就是在短时平稳的假定下, 用稳态分析方法处理非平稳信号的一种方法。

设离散时域采样信号为 $x(n)$, $n=0, 1, \dots, N-1$, 其中 n 为时域采样点序号, N 是信号长度。然后对信号进行分帧处理, 则 $x(n)$ 表示为 $x_n(m)$, $n=0, 1, \dots, N-1$, 其中 n 是帧序号, m 是帧同步的时间序号, N 为帧长 (一帧内的采样点数)。信号 $\{x(n)\}$ 短时傅里叶变换为:

$$X_n(e^{j\omega}) = \sum_{m=-\infty}^{\infty} x(m) \cdot w(n-m) \cdot e^{-j\omega m} \quad (1)$$

其中 $\{w(n)\}$ 为窗序列, 则信号 $x(n)$ 的离散时域傅里叶变换 (DTFT) 为:

$$X(n, e^{j\omega}) = \sum_{m=0}^{N-1} x_n(m) e^{-j\omega m} \quad (2)$$

采用离散傅立叶变换 (DFT) 得:

$$X(n, k) = \sum_{m=0}^{N-1} x_n(m) e^{-j \frac{2\pi k m}{N}} \quad (3)$$

其中 $0 \leq k \leq N-1$, 则 $|X(n, k)|$ 就是的 $x(n)$ 短时幅度谱估计, 而时间 m 处频谱能量密度函数 (或功率谱函数) $P(n, k)$ 为:

$$P(n, k) = |X(n, k)|^2 = (X(n, k)) \times (\text{conj}(X(n, k))) \quad (4)$$

则 $P(n, k)$ 是二维的非负实值函数, 并且不难证明它是信号 $x(n)$ 的短时自相关函数的傅里叶变换。用时间 n 作为横坐标, k 作纵坐标, 将 $P(n, k)$ 的值表示为灰度级所构成的二维图像就是语谱图。通过变换 $10 \log_{10}(P(n, k))$ 能得到语谱图的 dB 表示。

1.2 语谱图的伪彩色映射

对 $P(n, k)$ 实施伪彩色映射, 可以得到具有较高分辨率和较好视觉效果伪彩色语谱图。把 $P(n, k)$ 的最

李富强: 硕士研究生

基金资助: 河南省骨干教师资金资助

中国自控网: <http://www.autocontrol.com.cn>

电话: 010-62132436, 62192616 (T/F)

大值 $P_{\max}(n,k)$ 映射为归一化 1 电平, 而把最小值 $P_{\min}(n,k)$ 映射为归一化 0 电平。再将 $P(n,k)$ 线性映射为 0~1 间的电平 L_i , 计算机监视器再根据 L_i 的值把语谱图以伪彩色模式显示出来。为了得到更好的显示效果, 可以选择适当的基准值 Base, 把小于 Base 值限定在此基准电平上, 把大于 Base 的值线性映射到 0~1 的归一化彩色值。则彩色值矩阵 $L=\{l(n,k)\}$ 的数学表示如下:

$$l(n,k) = \frac{B(n,k) - Base}{\max(B(n,k) - Base)} \quad (5)$$

$$\text{且 } B(n,k) = \begin{cases} P(n,k), & P(n,k) > Base \\ Base, & P(n,k) \leq Base \end{cases} \quad (6)$$

2 相关 MATLAB 功能函数

1967 年美国 MathWorks 公司推出了 MATLAB (Matrix laboratory 的简称)。MATLAB 自推出之后不断更新和扩充, 目前已经推出了 7.x 版本。它的交互式集成界面能够帮助用户快速地完成数值分析、矩阵运算、数字信号处理、仿真建模、系统控制和优化等功能。由于 MATLAB 的这些特性, 它已经成为科研工作 and 工程仿真中的高效助手。

2.1 用于短时频谱分析的功能函数

1) 分帧加窗函数: 下面程序中采用的是 hanning 窗。根据要求还可选用 hamming(N), blackman(N) 及 bartlett(N), 其中 N 为窗长(帧长)。

2) 快速傅里叶变换函数: fft(s)。s 为加窗后的一帧信号。因为实时信号 FFT 的频域样值关于中点(即采样频率的 1/2)对称, 所以 fft(s) 给出的矩阵数据只有前半部分有用。

3) 复数取共扼函数: conj(z)。

2.2 伪彩色映射和语谱图显示的功能函数

1) 伪彩色映射函数: colormap(MAP)。其中 MAP 是所采用的伪彩色映射矩阵, 默认值为 JET, 可以通过 MAP=colormap 获得当前的伪彩色映射矩阵, 它可以是一个任意行的矩阵, 但其必须有且只有三列, 并分别表示红色、绿色和蓝色的饱和度。

2) 频谱图显示函数: imagesc(t, f, L)。当数据比例映射使用整个色域时使用该函数, 其它情况下与 image() 函数作用相同。其中 t 是时间坐标, f 是频率坐标, L 则是从功率谱值经伪彩色映射后的彩色电平值。

3 流程图及实现程序

3.1 流程示意图

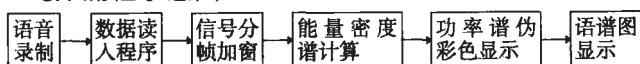


图 1 语谱图产生流程图

3.2 程序

```
function sogram(Winsiz,Shift,Base,Mode,Gray);
[Sg,Fs]=wavread(Wavnam);n=floor((length(Sg)-Winsiz)/Shift)+1;A=zeros(Winsiz/2+1,n);
```

```
for i=1:n n1=(i-1)*Shift+1;n2=n1+(Winsiz-1);s=Sg(n1:n2);s=s.*hanning(Winsiz);
z=fft(s);z=z(1:(Winsiz/2)+1);z=z.*conj(z);z=10*log10(z);A(:,i)=z end
L0=(A>Base);L1=(A<Base);B=A.*L0+Base.*L1;L=(B-Base)./(max(max(B))-Base);
y=[0:Winsiz/2]*Fs/Winsiz;x=[0:n-1]*Shift; if
Mode==1 colormap('default'); else
mymode=gray;mymode=mymode (Gray:-1:1,:);colormap(mymode);
end imagesc(x,y,L);axis xy;
3.3 程序中参数说明
```

Wavnam: 定义待分析信号的路径及文件名。语音信号由计算机声卡录制, 以 .wav 格式存储。程序中 wavread(Wavnam) 读入的语音信号样值赋给矩阵 Sg, 也即待分析信号 x(n), 采样频率赋给 Fs。Winsiz: 定义帧长, 一般应取 2 的幂次, 其目的是适合 FFT 的要求。通过对 Winsiz 的选值可实现宽带频谱或窄带频谱的选择。Shift: 定义帧移值。一般此值小于或等于 Winsiz。Shift 值越小, 时域分辨率越高。Base: 基准电平值。注意, 此值的设定需根据实际经验, 可以通过在多次运行此程序中给出不同的 Base 值, 观察所获得的频谱图的视觉和分辨率效果, 选择一个合适的 Base 值, 如果没有特别要求, 可取默认值 Base=0。Mode: 定义显示模式。1 伪彩色映射, 0 为灰度映射。上述程序中, 伪彩色映射采用默认值 'default' 也即 jet, 可选的其它映射还有: bone, cool, copper, flag, hot, hsv, pink, prism, Gray: 灰度显示层数, 当 Mode=0 时有效, 且 Gray 只能在 1~64 间取值, 为了获得较好灰度显示效果, 一般取值 64。

3 语谱图显示与分析

语谱图中的花纹有横杠、乱纹和竖直条等。横杠是与时间轴平行的几条深黑色带纹, 它们是共振峰。从横杠对应的频率和宽度可以确定相应的共振峰频率和带宽。在一个语音段的语谱图中, 有没有横杠出现是判断它是否是浊音的重要标志。竖直条(又叫冲直条)是语谱图中出现与时间轴垂直的一条窄黑条。每个竖直条相当于一个基音, 条纹的起点相当于声门脉冲的起点, 条纹之间的距离表示基音周期。条纹越密表示基音频率越高。

图 2 是“你好”的灰度语谱图显示。Winsiz=256, Shift=32, Base=0, Mode=0, Gray=64。此图为灰度映射图, 且窗长较短, 可以获得较高的时间分辨率。在此宽带语谱图上, 功率谱的谐波结构在频率轴方向被平滑掉了, 谐波间隙被填平了, 因此在频域看不出基音频率和频谱的谐波形式结构, 但可以明显地看出共振峰结构和语谱包络, 特别是时间分辨率高, 可以清楚地看到浊音共振峰时变过程。

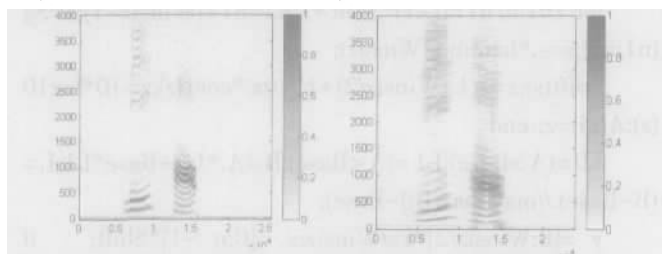


图2 “你好”宽带语谱图
灰度显示图

图3 “你好”窄带语谱图
灰度显示图

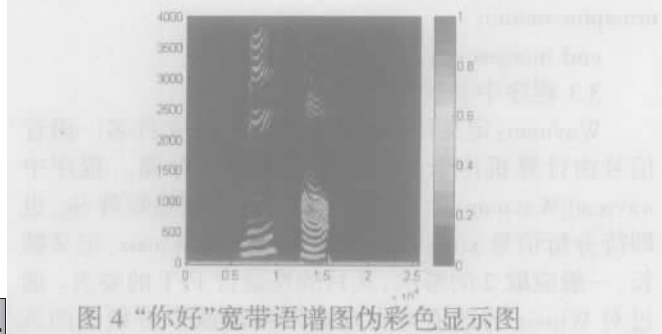


图4 “你好”宽带语谱图伪彩色显示图

图3参数为: Winsiz=2048, shift=128, Base=0, Mode=0, Gray=64。此图仍为灰度映射图,但窗长较长,可以获得较高的频率分辨率。在此窄带语谱图上,滤波器的带宽能在浊音区分辨出每个谐波。语谱的谐波形式的结构在频率方向清晰可见,并能看到基音频率的时变过程。

图4是“你好”的伪彩色宽带语谱图。参数为: Winsiz=256, shift=32, Base=0, Mode=1。Gray值任意。

4 结论

文章阐述了语谱图的产生,以及如何在MATLAB中编程绘出它的伪彩色图及灰度图的方法。对相关的MATLAB函数作了简要介绍,通过对录制语音信号的分析,分别得到它的宽带语谱图及窄带语谱图,并分析了窗长对时间分辨率和频率分辨率的影响,并对宽带语谱图分别进行了灰度与伪彩色显示。通过计算机声卡,可以很方便的完成语音录制,又由于MATLAB具有很丰富的功能函数和方便的绘图功能,通过编程很容易在普通PC机上快速实现语谱图显示,从而为进一步语音信号处理提供便利条件。

参考文献:

- [1] 易克初等. 语音信号处理. 北京: 国防工业出版社, 2000. 15~16.
- [2] 胡航. 语音信号处理. 哈尔滨: 哈尔滨工业大学出版社, 2000. 34~35.
- [3] 赵力. 语音信号处理. 北京: 机械工业出版社, 2003. 42~44.
- [4] 邓华. MATLAB通信仿真及应用实例详解. 北京: 人民邮电出版社, 2003. 1~2.
- [5] 王华等. Matlab在电信工程中的应用. 北京: 中国水利水电出版社, 2001. 226~227.
- [6] 黄忠霖. 控制系统MATLAB计算与仿真. 北京: 国防工业出版社, 2001. 495~502.

作者简介: 李富强, 男, 1982-, 汉族, 郑州大学电气工程学院, 硕士研究生, 研究方向: 数字信号处理, E-mail: lifuqiang@gs.zzu.edu.cn; 万红, 女, 1964-, 汉族,

郑州大学电气工程学院, 副教授, 硕士生导师。郑州解放军信息工程大学, 博士生。研究方向: 数字信号处理; 黄俊杰, 男, 郑州大学电气工程学院, 工程师。

Author brief introduction: Li Fuqiang, male, 1982-, Han, master. School of Electric Engineering, Zhengzhou University. Major: Digital Signal Processing.

(郑州大学) 李富强 万红 黄俊杰

(郑州解放军信息工程大学) 万红

(School of Electronic Engineering, Zhengzhou University, Zhengzhou 450002 China) Li, Fuqiang Wan, Hong Huang, Junjie

联系方式:

(450002 河南省郑州市文化路97号, 郑州大学北区电气工程学院) 李富强

(投稿日期: 2005.3.25) (修稿日期: 2005.4.6)

(接178页) 总之, 本文旨在帮助读者从多个角度深入理解补码的概念, 并形成直观的认识。解决了补码的问题, 也就扫清了大家学习计算机道路上的一个拦路虎。笔者愿与大家一起在科学技术的海洋中尽情遨游, 共同分享成功的喜悦。

参考文献:

- [1] 李文兵. 计算机组成原理. 清华大学出版社, 2002. (2).
- [2] 李明慧. 计算机组成原理. 中央广播电视大学出版社, 2001.
- [3] 沈美明. 汇编语言程序设计. 清华大学出版社, 2000.

作者简介: 张德伟(1972-), 女, 山东临沂人, 毕业于山东科技大学工业自动化专业, 临沂师范学院讲师, 主要从事计算机信息工程和单片机方面的研究。Email: zdwsp@163.com. 沈培锋(1973-), 男, 山东龙口人, 临沂市职业技能学校讲师, 研究方向: 工业自动化。张德珍(1970-), 女, 山东临沂人, 临沂师范学院讲师, 研究方向: 计算机CAD设计。吕月娥(1977-), 女, 河北衡水人, 临沂师范学院助教, 研究方向: 计算机应用技术。Author brief introduction: Zhang Dewei (1972-), female, graduate of Shandong Scientific University, Major research fields: computer information project and single chip computer. Email: zdwsp@163.com. Shen Peifeng (1973-), male, Study Field: Industry Automation.

(276005 临沂师范学院 信息学院) 张德伟 吕月娥

(276001 临沂市职业技能学校) 沈培锋

(276005 临沂师范学院 工程学院) 张德珍

(Department of Computer and Information Science, Linyi Teachers College, 276005) Zhang, Dewei Lv, Yue e

(Linyi Professional Technique School, 276001) Shen, Peifeng

(Engineering Institute, Linyi Teachers' College, 276005) Zhang, Dezhen

通信地址: (276005 临沂师范学院信息学院) 张德伟

(投稿日期: 2005.4.12) (修稿日期: 2005.4.25)