

# WaveSurfer 语音软件中文使用手册

曹洪林<sup>1,2</sup> 张秋艳<sup>2</sup>

<sup>1</sup> 北京大学中文系 <sup>2</sup> 中国政法大学证据科学研究院

caohonglin@pku.edu.cn; zhangqiuyan\_768@163.com

2012 年 11 月

# WaveSurfer 语音软件中文使用手册

## 目录

1. 软件简介.....	2
2. 界面介绍.....	4
3. 录音采样.....	11
4. 语音编辑.....	13
5. 图形打印.....	15
6. 波形分析.....	17
7. 强度分析.....	22
8. 语图分析.....	25
9. 共振峰分析.....	29
10. 基频分析.....	33
11. 语音标注.....	37
12. 数据分析.....	42
13. 构建自己的分析平台.....	47
14. 有用链接/文献.....	48

## 1. 软件简介

### (1) 基本介绍

WaveSurfer 是一个免费的、开源的语音分析软件，不仅功能强大，而且简单易学，是目前世界范围内最流行的语音分析软件之一。该软件由瑞典皇家理工学院言语技术中心（Centre for Speech Technology, KTH）的 Kåre Sjölander 与 Jonas Beskow 教授设计开发，基本功能主要基于经典的 Snack 声音处理工具包。该软件既能满足初学者进行简单的语音分析和标注，也能实现高级用户的特殊分析需要，具备良好的可视化操控功能。其主要特点有：

- 接口灵活，可以处理多个声音。
- 可以读取并写入多种格式的声音文件，包括：WAV、AU、AIFF、MP3、CSL、SD、Ogg/Vorbis 和 NIST/Sphere 等。
- 可以对语音进行多种格式的标注：读取/写入 HTK(MLF)、TIMIT、ESPS/Waves+ 和 Phondat 等格式，支持编码和统一字符编码标准。

- 播放和录制的文件大小无限制，录音时可以进行实时语图显示。
- 声音的多层次分析：宽带/窄带语图显示，基频、共振峰、FFT 等分析。
- 可以根据需要自由创建分析平台。读者只要能够熟练运用“鼠标右键”+“创建分析窗口 (Create Pane)”基本上就能创建适合自己的分析平台。
- 通过插件 (plugin)，可以添加新的功能。
- 可以被其他应用程序调用。
- 脚本化，内置脚本解释器。

## (2) 下载地址

本使用手册基于 WaveSurfer 软件的最新版本 1.8.8p4 (2011-12-30) 创建，该版本可以在 <http://sourceforge.net/projects/WaveSurfer/> 下载，关于旧版本的详细信息可在 <http://www.speech.kth.se/WaveSurfer/index2.html> 上浏览。关于该软件的其他介绍和参考资料请见本手册第 14 部分“有用链接及文献”。

## 2. 界面介绍

WaveSurfer 软件的界面主要由菜单栏，工具栏，音频控制区等部分组成，界面示意图如图 2-1 所示，基本窗口介绍见图 2-2。

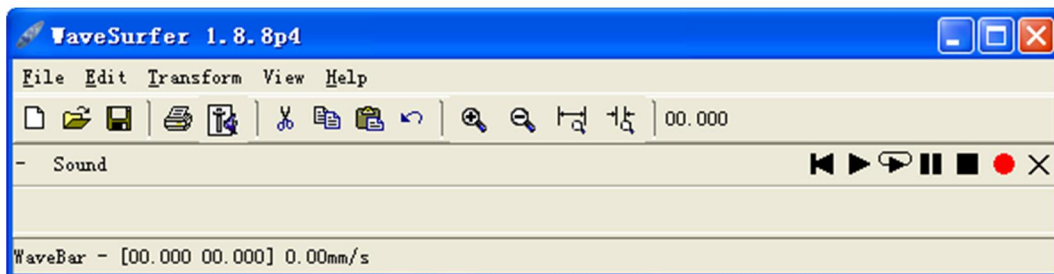


图 2-1: WaveSurfer 软件的界面图

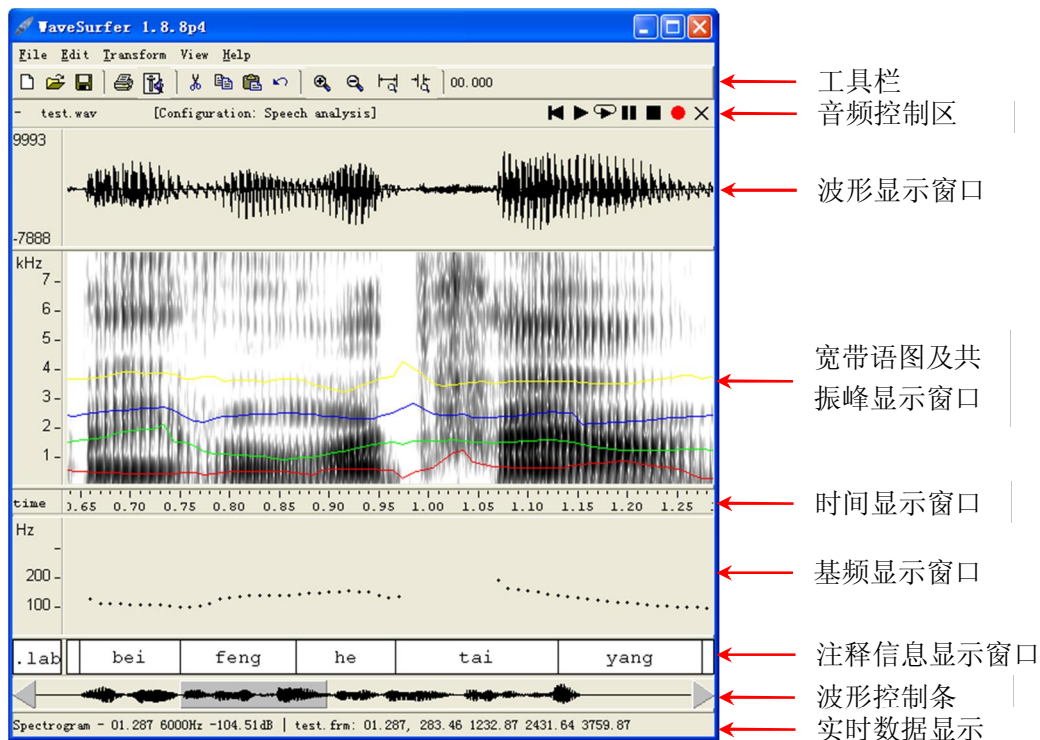


图 2-2: WaveSurfer 软件基本窗口介绍

### (1) 菜单栏

软件的菜单栏主要有 File、Edit、Transform、View 和 Help，具体功能如下：

File 文件

File 的主要功能如表 2-1 所示。

表 2-1: File

New	新建窗口。
Open	打开音频文件。
Revert	经过编辑后，确认不保存时，返回到文件的初始状态。
Save	保存文件。
Save As	将文件另存。
Save Selection	保存选中的音频文件部分。
Show console	打开控制台来使用 Tcl/Tk，用来操纵有关该命令的 GUI 应用程序。
Print	打印/打印预览。
Preferences	属性设置。
Associate File Types	设置音频文件之间的格式关联。
Close	关闭当前打开的音频文件。
Exit	退出软件。

## Edit 编辑

Edit 主要操作选项见表 2-2。

表 2-2: Edit

Undo	撤销编辑操作。注意：只能撤销一步。
Cut	剪切
Copy	复制
Paste	粘贴
Mix Paste	选定一个百分比，复制粘贴缓冲区中的内容。
Seclect All	选中整个音频文件。
Selection to New	将选中的部分置于一个新的窗口中。

## Transform 转换

Transform 的主要功能如表 2-3 所示。

表 2-3: Transform

Convert	转换音频的采样率，编码方式，声道数量。
Amplify	对选中部分进行声音增强。
Fade	设置声音渐变的幅度及函数。
Normalize	设置选中部分的归一化程度。
Echo	设置回声。
Mix Channels	混音。
Invert	波形图以 X 轴为对称轴反转。
Reverse	音频（在时间轴上）反转。
Silence	将选定部分设置为静音。
Remove DC	去除直流分量。

## View 视图

View 的主要功能如表 2-4 所示。

表 2-4: View

Zoom In	放大显示图谱。
Zoom Out	缩小显示图谱。
Zoom Out Full	显示全部音频信息。
Zoom To Selection	将选中的部分放大到整个窗口。

## Help 帮助

Help 的主要功能如表 2-5 所示，其中多数功能都是在线信息，需要连接 HTK 的网站才能阅读相关信息。

表 2-5: Help

Manual	查看手册
FAQ	显示常见的问题
Forum	显示 WaveSurfer 的相关论坛
About Plug-ins	关于插件的说明
About WaveSurfer	软件版本信息等说明

## (2) 工具栏

工具栏主要操作如表 2-6 所示。





表 2-6: 工具栏主要操作

	新建文件
	打开音频文件
	保存
	打印
	控制左右声道音量大小
	剪切
	复制
	粘贴
	撤销前一步操作（只能撤销一步！）
	放大显示图谱
	缩小显示图谱
	将图谱缩小到原始大小
	将选中的部分放大至整个窗口

### (3) 音频控制区

音频控制区主要操作如表 2-7 所示。

表 2-7: 音频控制区的主要操作

	跳转到文件的起始位置
	从光标处开始播放音频文件
	从文件起始位置开始播放音频文件
	循环播放选中的部分
	暂停播放
	停止播放
	跳转到文件的最后位置
	开始录音
	关闭音频文件

### (4) 快捷键

快捷键主要内容如表 2-8 所示。

表 2-8: 快捷键

F1	播放选中部分的音频文件
F2	从光标处开始播放
F3	播放整个音频文件
F5	播放屏幕中显示的视频
F6	开始/停止播放
F8	循环播放
F11	将放大后的文件缩小到原来的状态
F12	将选中的部分放大至整个窗口
Space	开始/暂停播放
Ctrl+r	在未确认保存时恢复到文件打开时的初始状态

上述各功能对应的快捷键是软件默认的设置，读者可以根据自己的需要设置各功能对应的快捷键，步骤为“File→Preferences→Keys”，各功能对应的快捷键如图 2-3 所示。

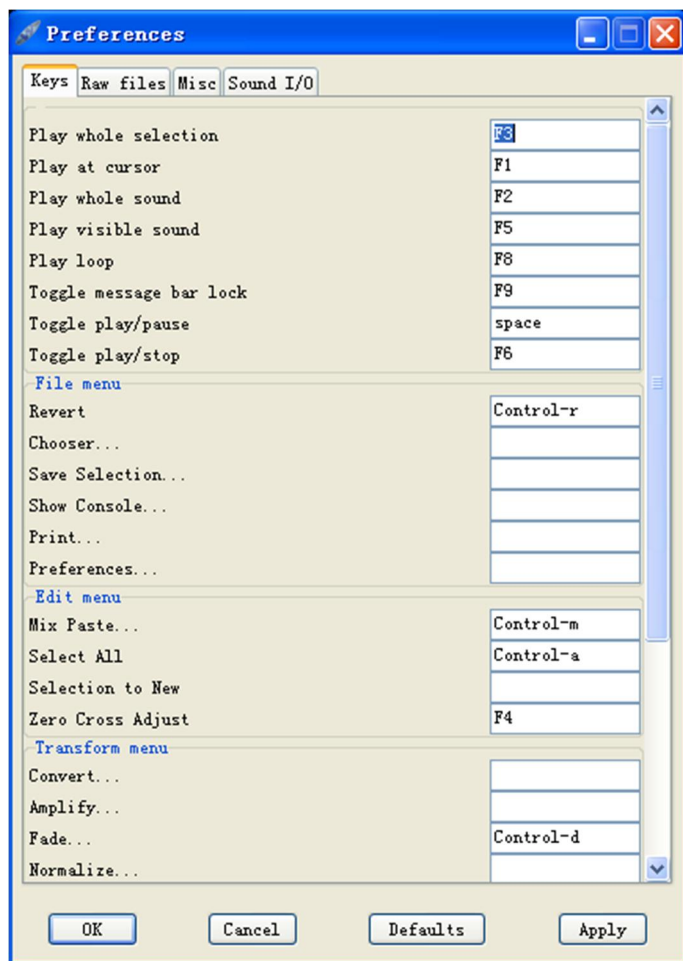


图 2-3: 快捷键设置

读者也可以根据自己的需要, 设置软件与不同格式文件之间的关联, 如图 2-4 所示。

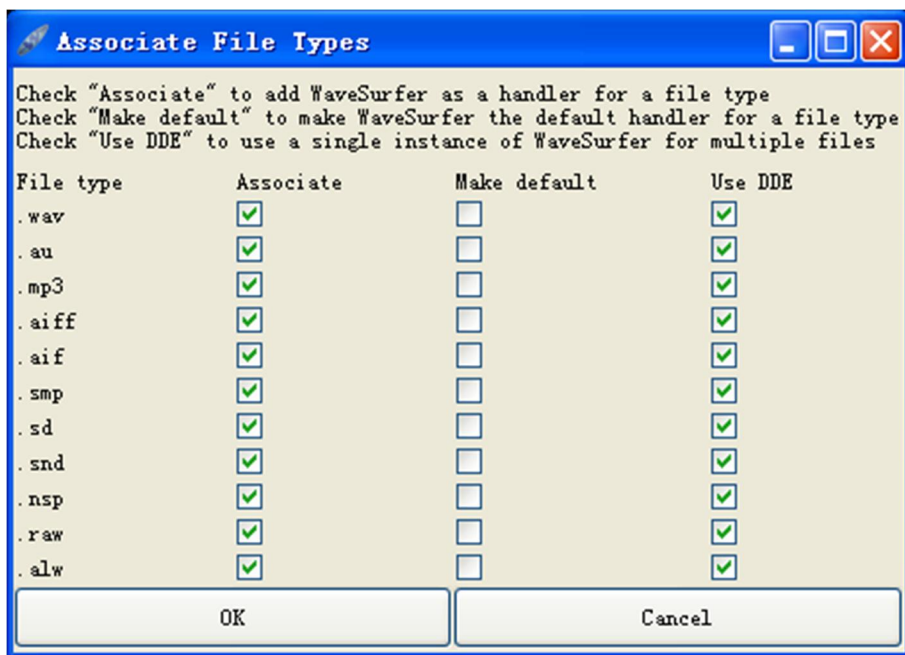


图 2-4: 设置与不同格式文件之间的关联



## (5) 配置选择

打开 WaveSurfer 软件后（如上图 2-1 所示），常需要首先选择一定的配置，该软件提供了多种已经设置好的常用配置模式，如图 2-5 所示。

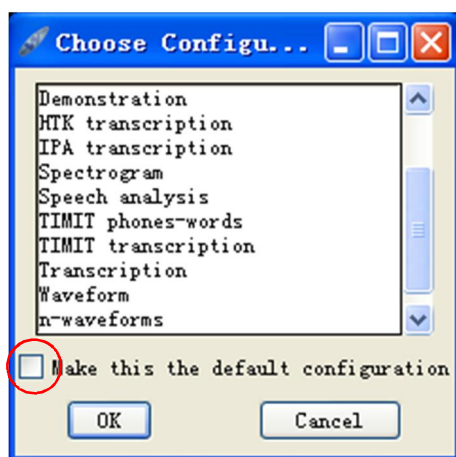


图 2-5: 配置选择

各种常用配置含义见表如下：

表 2-9: 常用配置选项

Demonstration	示例：显示音频的波形图、彩色宽带语图，以及对语图的调整控制选项、对波形振幅的显示范围选项和 FFT 分析窗口。
HTK Transcription	显示音频的宽带语图及 HTK 标注格式。
IPA Transcription	显示音频的宽带语图及 IPA 标注格式。
Spectrogram	只显示音频的宽带语图
Speech analysis	显示音频的波形图、语图，并对共振峰和基频进行分析。
TIMIT phones-words	显示音频的宽带语图及 TIMIT 电话对话标注格式。
TIMIT transcription	显示音频的宽带语图及 TIMIT 标注格式。
Transcription	显示音频的宽带语图及 WaveSurfer 标注格式。
waveform	只显示音频的波形图。
n-waveforms	显示多个波形图。
Standard	打开音频但不显示分析选项。

请注意：一般情况下可以选择图 2-5 中红色圆圈中的标注的复选项，将某配置作为默认选项，但如果选择 Speech analysis（或读者自己创建的其他的包含共振峰、基频等分析的配置模式）作为默认的配置选项，当软件关闭后重新打开时将无法正常使用，将

会弹出“Error in startup script”的报错对话框，如图 2-6 所示。解决方式是删除软件设置过程中形成的系统文件 preferences 即可，在 XP 系统下，preferences 文件的路径为“C:/Documents and Settings/Administrator/.WaveSurfer/1.8/preference”，将其删除后软件即恢复正常。另外一种解决途径是，在退出软件前，依次选择“File→Preferences→Misc→Default configuration”，将其设置为“None（或其他未包含共振峰和基频分析的配置模式）”，如图 2-7 中红色标示部分。

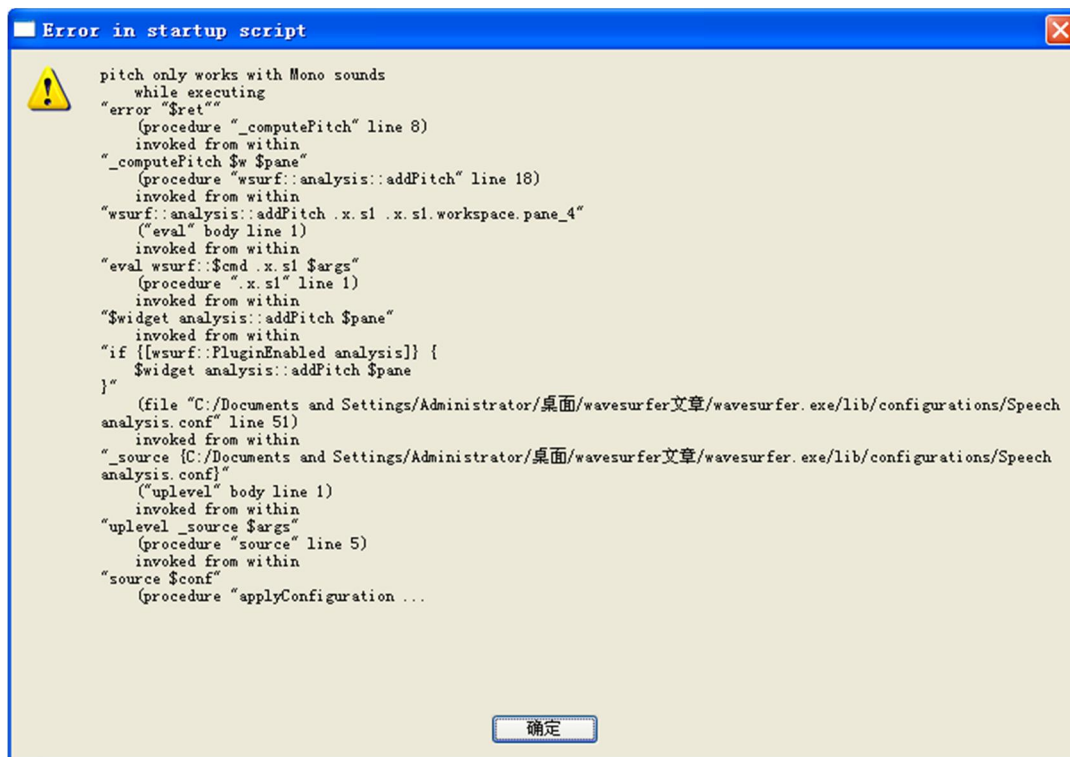


图 2-6: Error in startup script 错误对话框

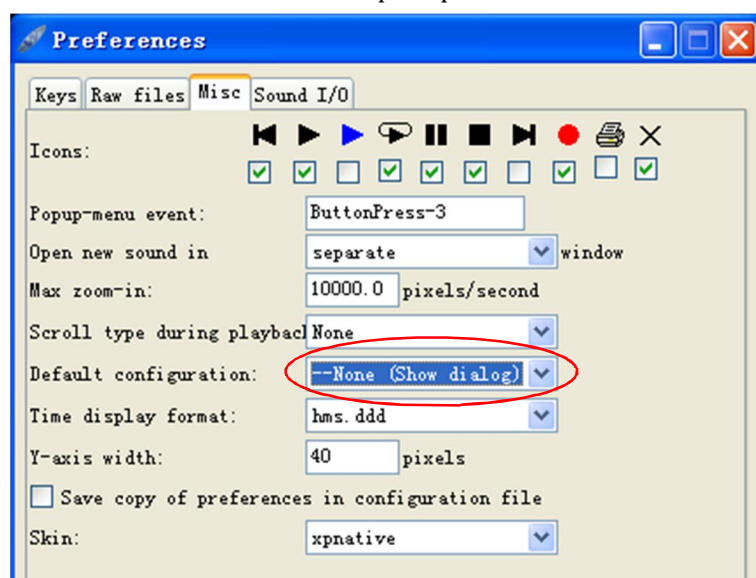


图 2-7: 修改默认的配置选项

### 3. 录音采样

WaveSurfer 软件支持录音功能，并可以实时显示录音的频谱图/波形图，还能够同时进行基频/共振峰分析（不推荐）。查看/修改软件默认的录音设置，可以依次选择“File→Preferences→Sound I/O”选项，如图 3-1 中所示，读者可以根据自己的需要，对图中的设置进行修改。

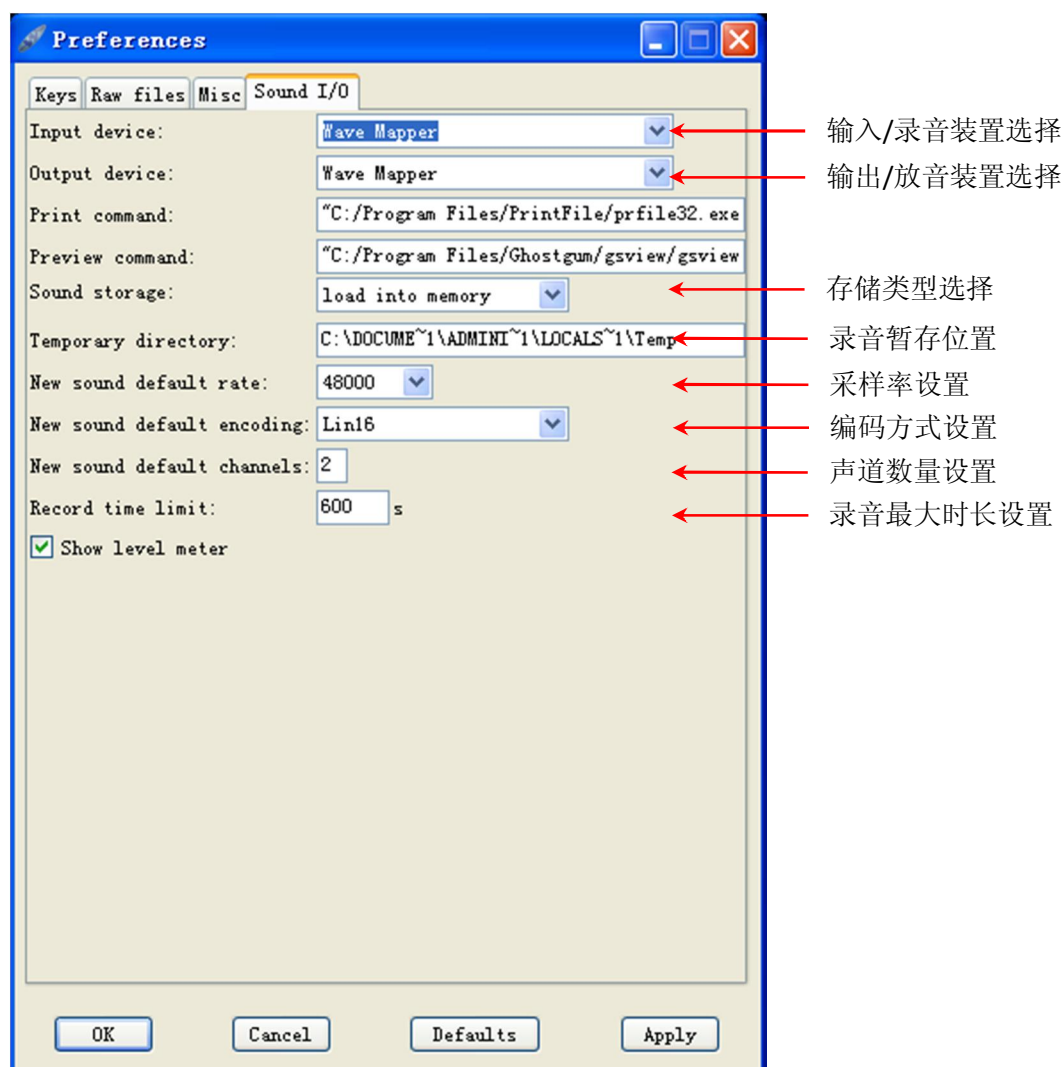




图 3-1: 录音设置

选择完录音设置之后，可以先新建一个录音文件，然后选择自己喜欢的配置模式，在弹出的分析窗口中点击  按钮即开始录音，录音中按空格键即可暂停录音（不推荐），再按下空格键会继续录音，点击  则停止录音，录音结束后保存文件即可。图 3-2 为录音时实时显示频谱图的例子，录音采样率为 16kHz，单声道。

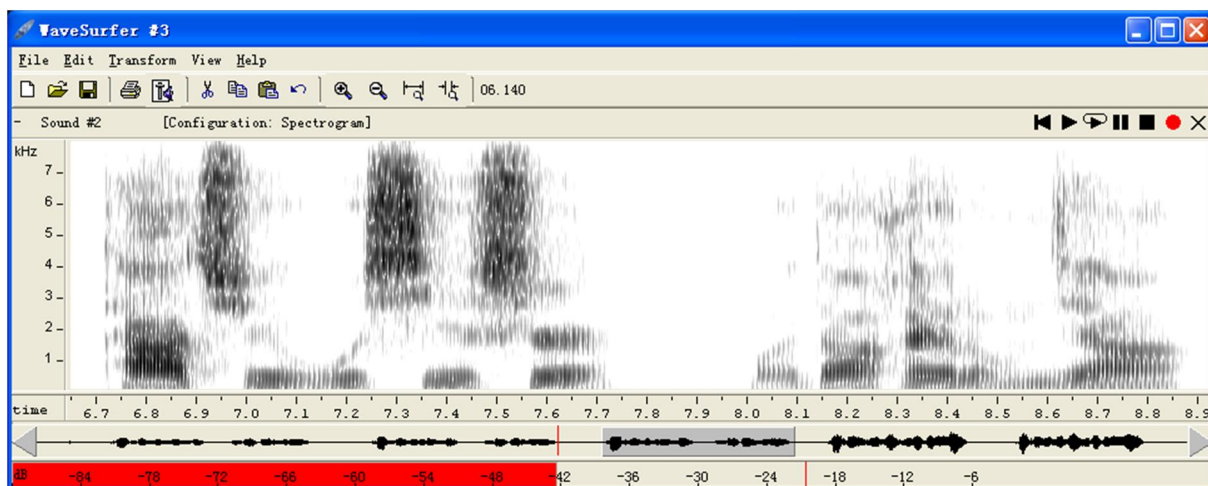


图 3-2: 选择频谱图实时显示录音

## 4. 语音编辑

WaveSurfer 软件实现对音频的快速编辑，如同在 office 软件中操作一样简单易行，可以使用快捷键“ctrl+a”、“ctrl+c”、“ctrl+x”和“ctrl+v”实现全选、复制、剪切、粘贴等操作。举例说，可以将图 3-1 中的[ta]音节选中，将其复制粘贴到该音频的其他位置，再选择“Transform→Reverse”将其翻转编辑（见图 4-2），就会形成一个类似于塞音韵尾的[at]音节。

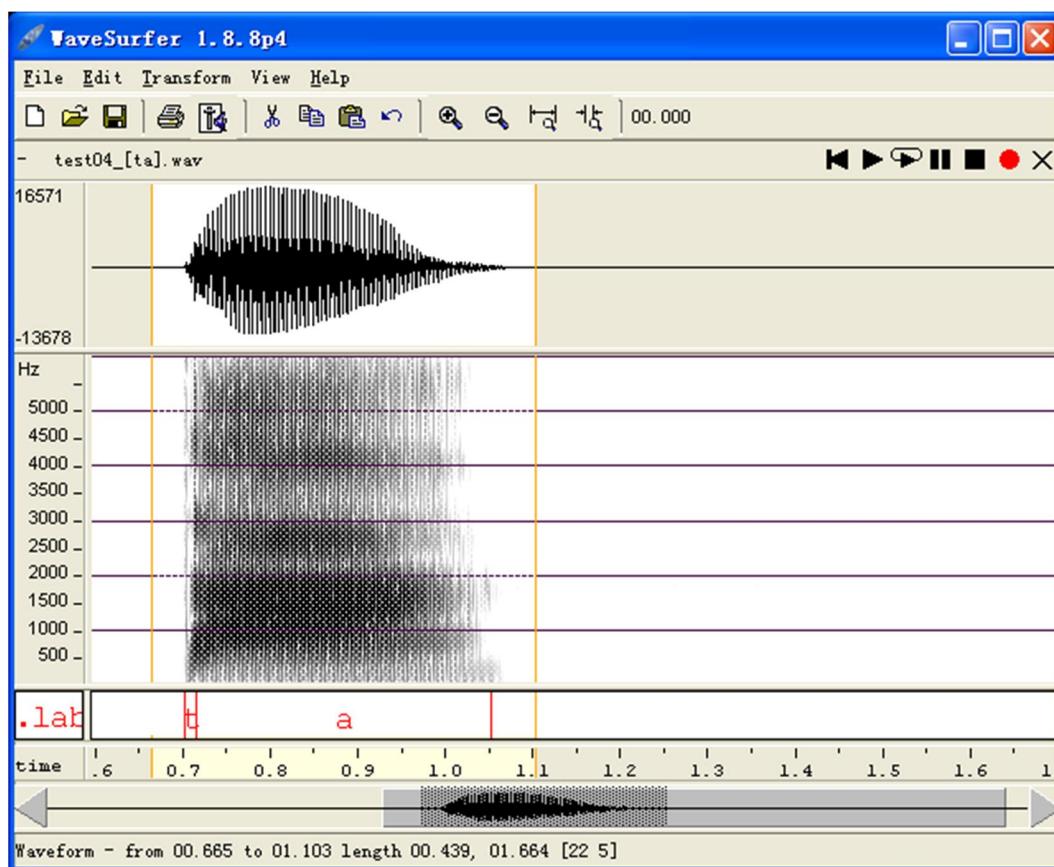


图 4-1: [ta]音节的语图与波形

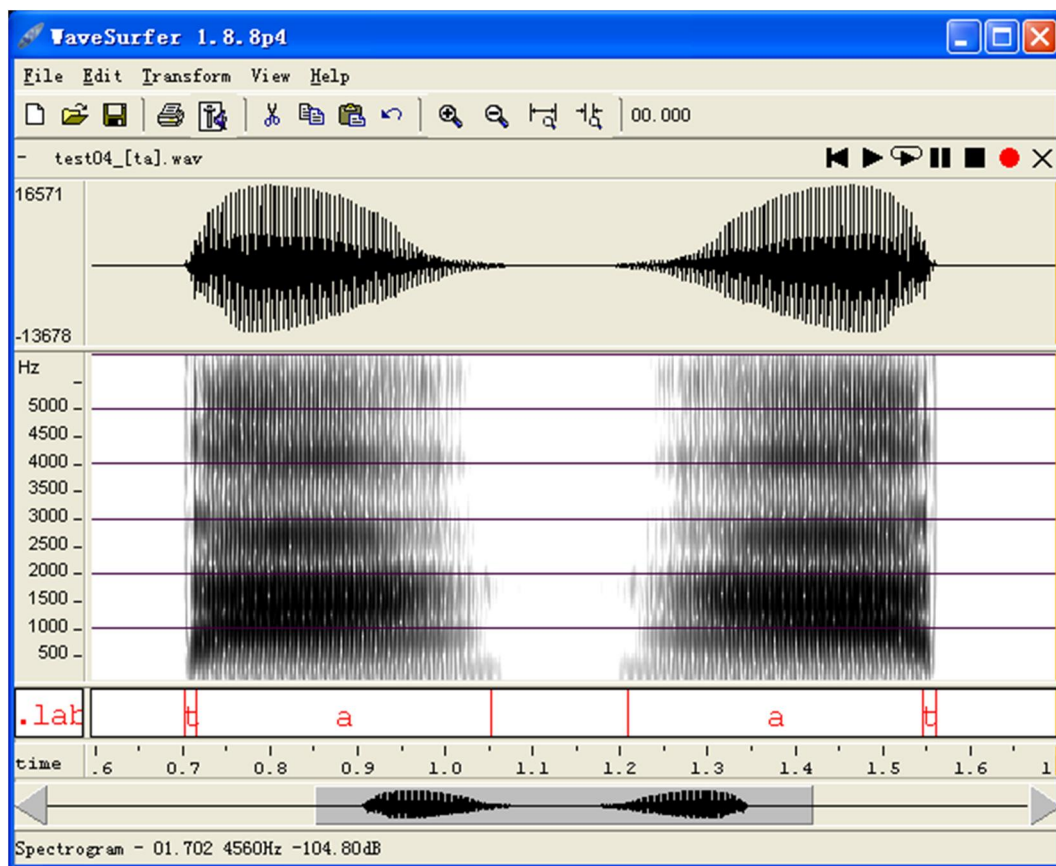


图 4-2: 对[ta]音节的编辑示例

## 5. 图形打印

### (1) 简单截图

在做好分析图形之后，可以选择键盘上的 **Print Screen** 按钮抓取整个屏幕显示的图像，然后直接将粘贴到 office 软件中进行处理，或者利用系统附件中的画图工具进行处理后存为特定格式的图像后再利用。

在此，推荐读者还可以结合使用其他抓图/截图软件作图。本使用手册中所有图形都是作者使用 HyperSnap 抓图软件做出的，该软件可以使用“**ctrl+shift+a**”抓取活动窗口、“**ctrl+shift+b**”抓取按钮、“**ctrl+shift+r**”抓取任意区域的图形等。

### (2) 标准打印

上述简单截图方法方便易行，能够满足大部分的分析需要，但是无法做出高质量的矢量图形，如要达到上述要求，可以选择标准打印模式。WaveSurfer 软件的下载安装包未包含有关打印的相关控件，读者需要自行下载安装。如果读者的打印机支持 ps 文件，可以不安装 Ghostscript 软件（基于 Adobe、PostScript 及可移植文档格式（PDF）的页面描述语言等而编译成的免费软件），如果不支持 ps 打印，则需要安装 Ghostscript 软件。该软件和 GSView 软件（免费 PostScript 文件阅读程序，可以很方便地阅读、打印 PS 文件）可以在互联网上免费下载。

安装完毕后，可以将图形进行全部或部分输出。以图 5-1 为例，如只将显示出的部分（“他去无锡市”）语图输出，可以依次选择“**File→Print**”，则会出现图 5-2 的对话框，将“**print selection only**”复选项选中，即只打印选中部分，如不选中打印全部音频的语图。打印输出效果可见图 5-3 中所示。另外，还可以点击“**preview**”按钮进行效果预览，或者点击“**save**”按钮将图像保存为“**tmp0.ps**”文件（可以修改文件名）。双击该文件，GSView 软件会自动将其打开，直接点击 **Ok** 后便可进行观察（读者可以自行决定是否注册）和其他操作，如图 5-4 中所示。

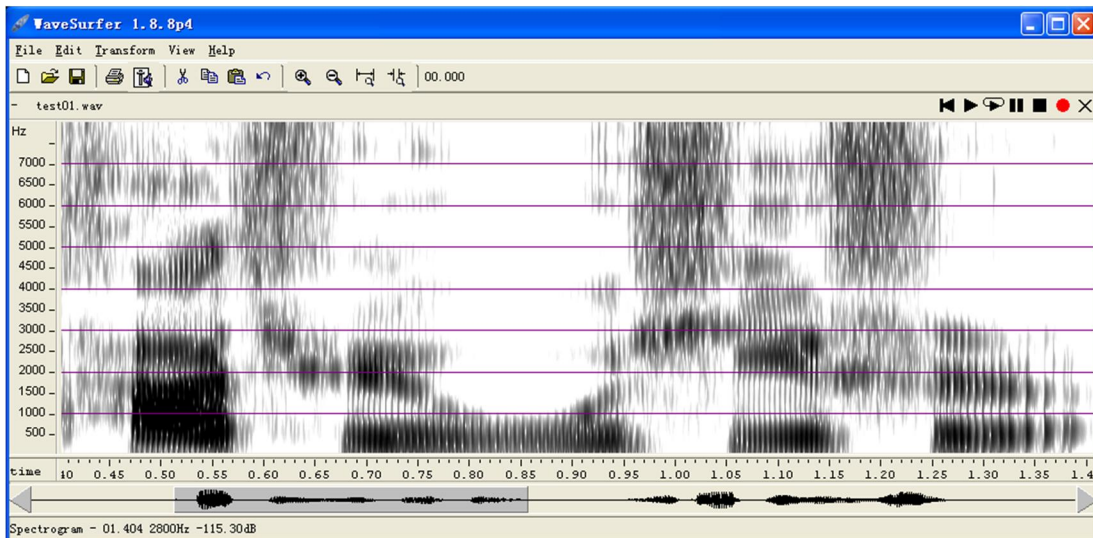


图 5-1: 打印示例语图

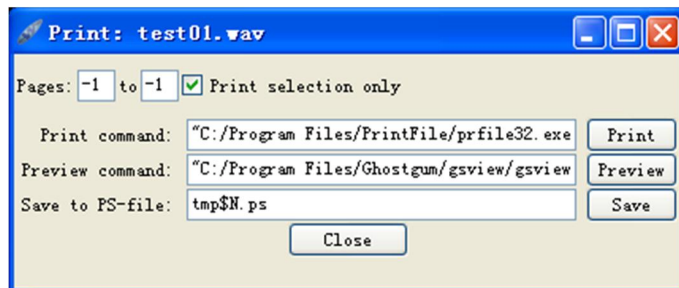


图 5-2: 打印选项

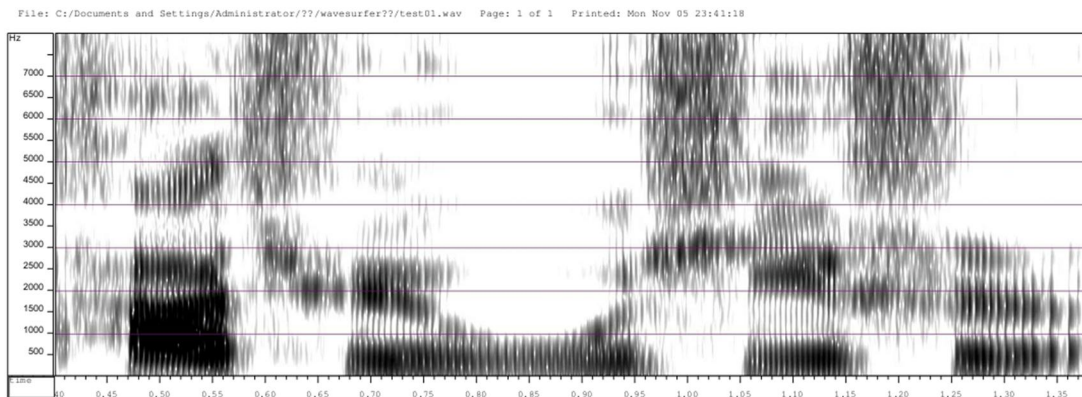



图 5-3: 矢量打印输出效果



图 5-4: GSView 软件注册窗口



## 6. 波形分析

打开一个语音文件（如 test01.wav），选择 Waveform 配置模式（如图 6-1），即可进行波形分析，如图 6-2 中所示。用鼠标选中语音的第一个音节[t<sup>h</sup>a]（他），点击  按钮将此选中的音节在整个窗口范围内显示，如图 6-3。在波形显示窗口中单击鼠标右键，选择“Waveform Blow-Up”（波形细目放大）选项，如图 6-4，即可出现对鼠标所在处前后 0.01s 范围内（可以调整此数据，获得对更小区域的观察）波形信号的放大图像，图 6-5 和图 6-6 分别显示的元音[a]与辅音[t<sup>h</sup>]某处波形的放大图像，很明显可以看出元音[a]的波形具有周期性，而辅音[t<sup>h</sup>]的周期性则不明显。

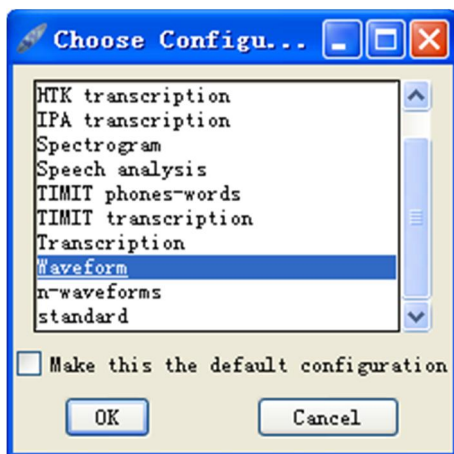


图 6-1: 波形分析配置模式

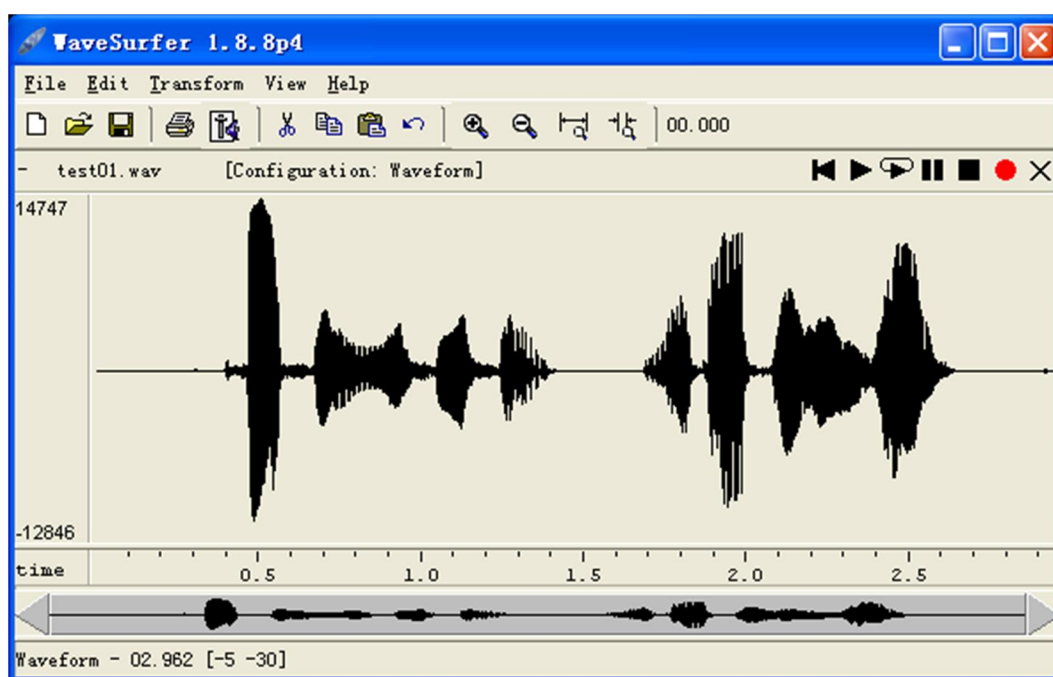


图 6-2: 波形分析窗口示例

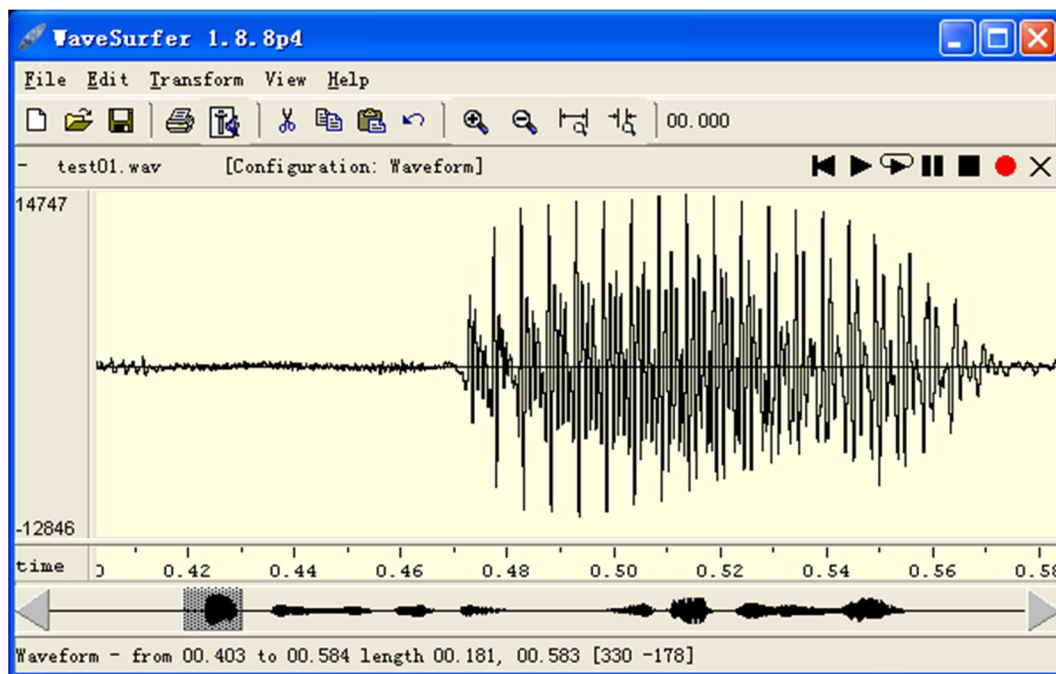


图 6-3: 将选中部分[tʰa]的波形放大至整个窗口

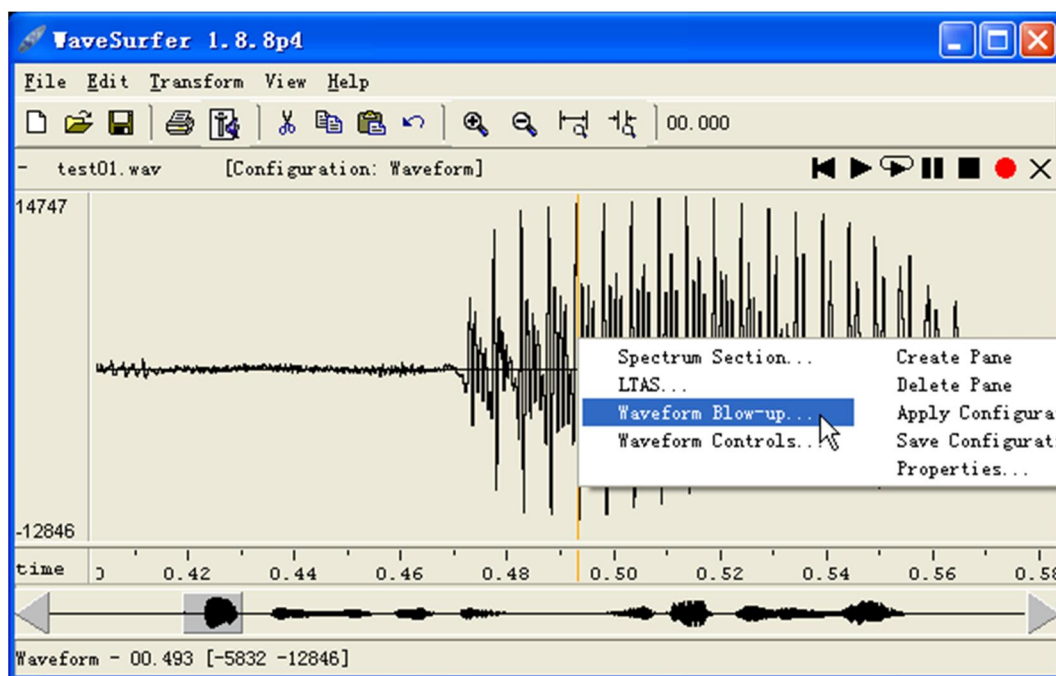


图 6-4: 波形细目放大示例

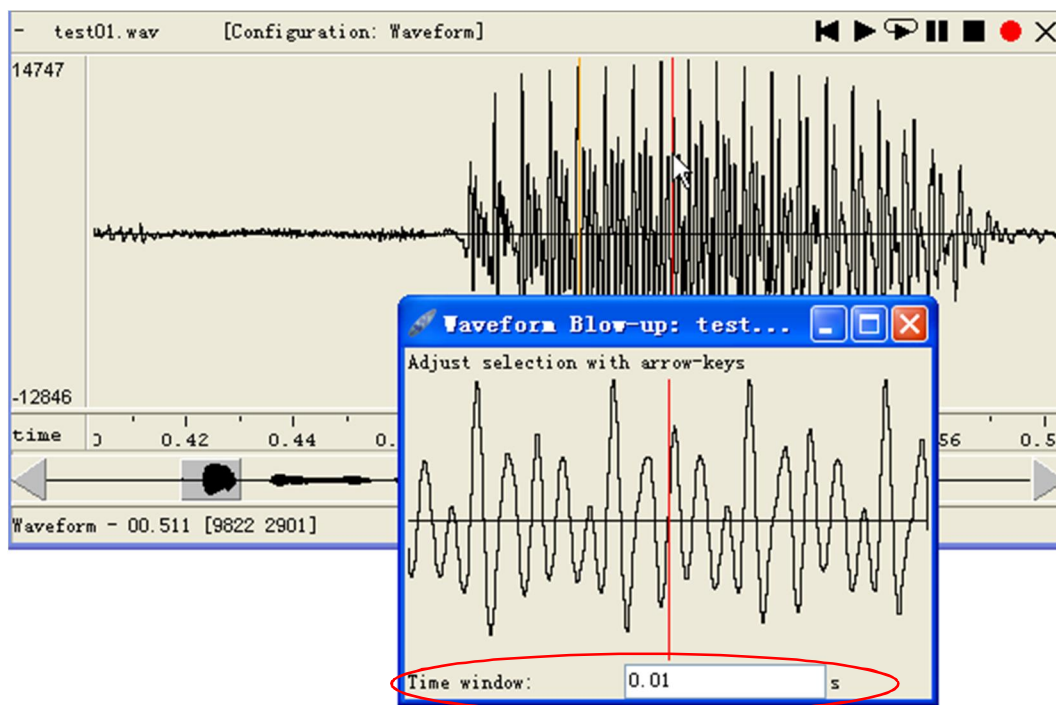


图 6-5: 波形细目放大示例-元音[a]

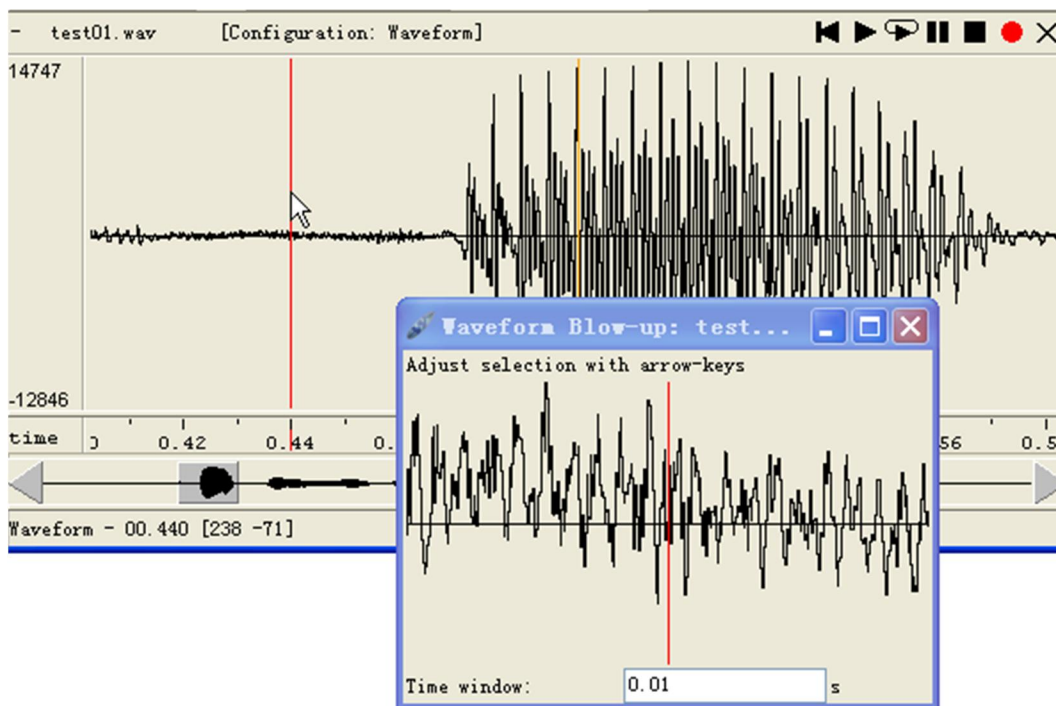


图 6-6: 波形细目放大示例-辅音[tʰ]

为了获得清晰、美观的分析图像，读者需要了解并掌握波形属性设置窗口中的各个选项。选项 Pane（显示栏信息）的属性设置见图 6-7，此处主要是对显示栏各个模块的颜色进行设置（见图 6-8），该设置适用于 WaveSurfer 软件其他任何分析窗口的显示栏设置，因此在下文中不再赘述。

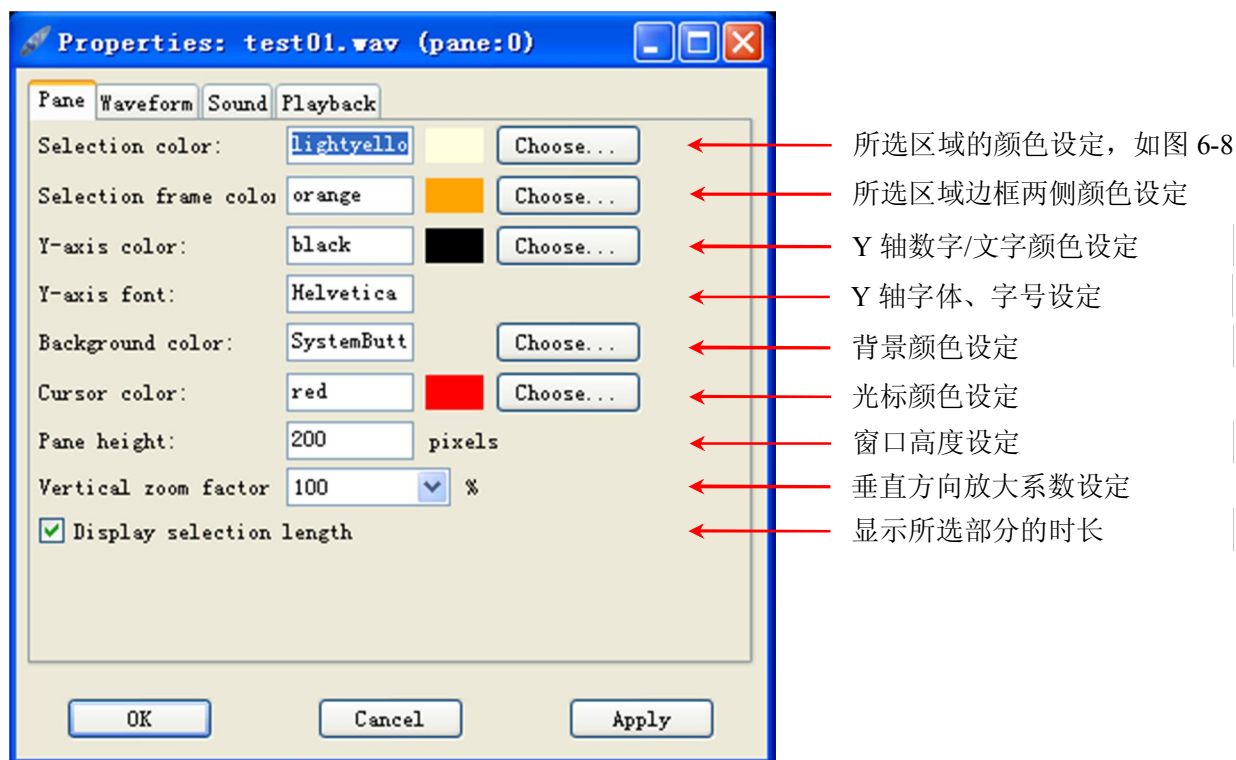


图 6-7: Pane 显示栏信息设置

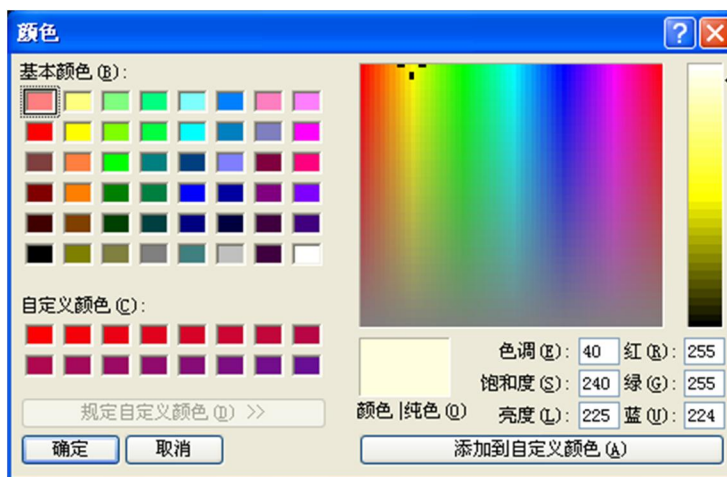


图 6-8: 颜色设置窗口

针对波形的具体参数设置，见图 6-9 中对 Waveform 属性选项的设置。属性选项中 Sound 和 Playback 的描述同样适用于其他分析窗口（下文不再赘述），其中 Sound 选项中包含了音频文件的各种属性，如存放位置、文件格式、采样率、声道数量、编码格式、时长、文件大小、修改时间等信息，见图 6-10。

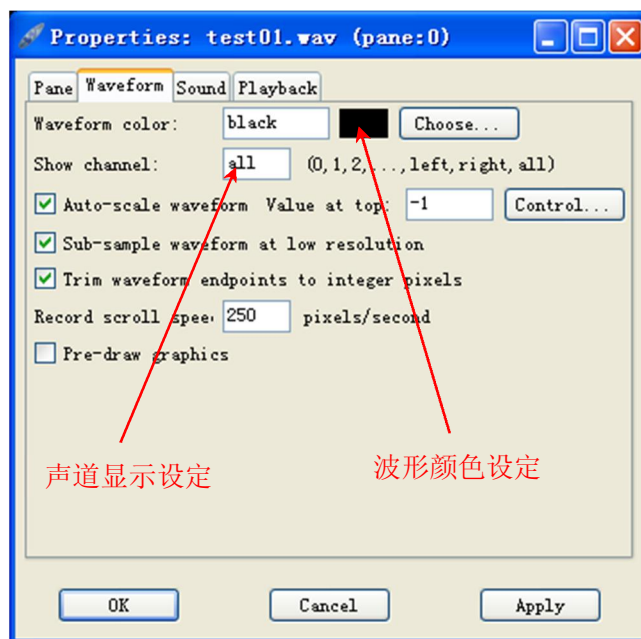


图 6-9: Waveform 属性选项设置

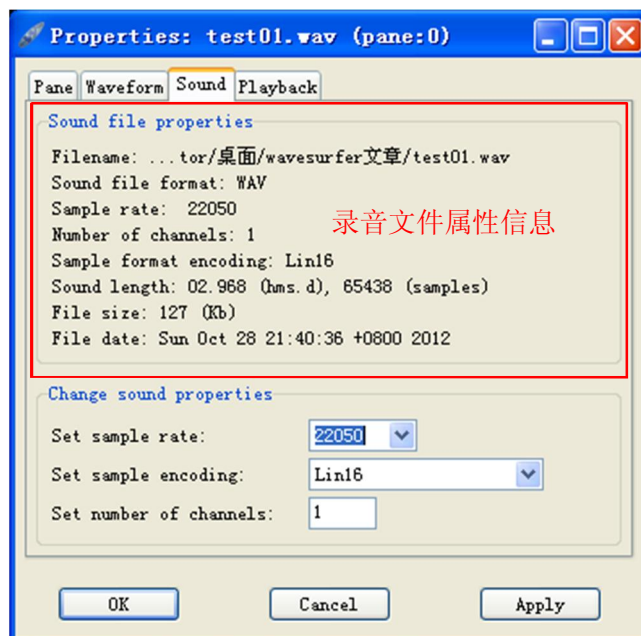


图 6-10: 音频文件属性

## 7. 强度分析

在波形显示窗口（或其他窗口）单击右键，选择“Create Pane→Power Plot”，可以进行强度分析，如图 7-1 和图 7-2 中所示：

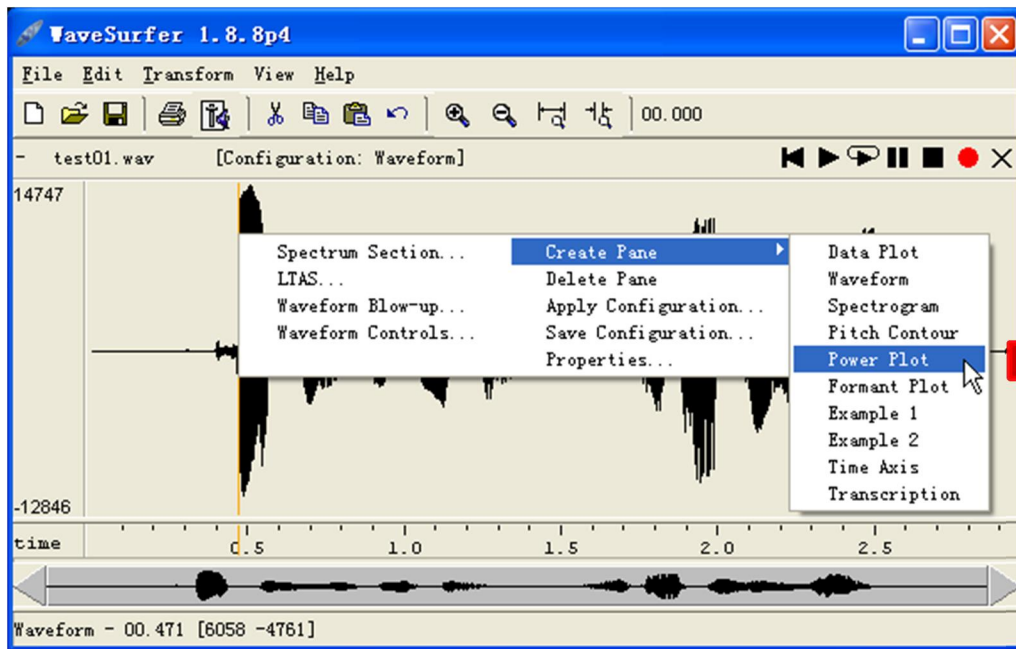


图 7-1: 波形分析窗口增加强度分析

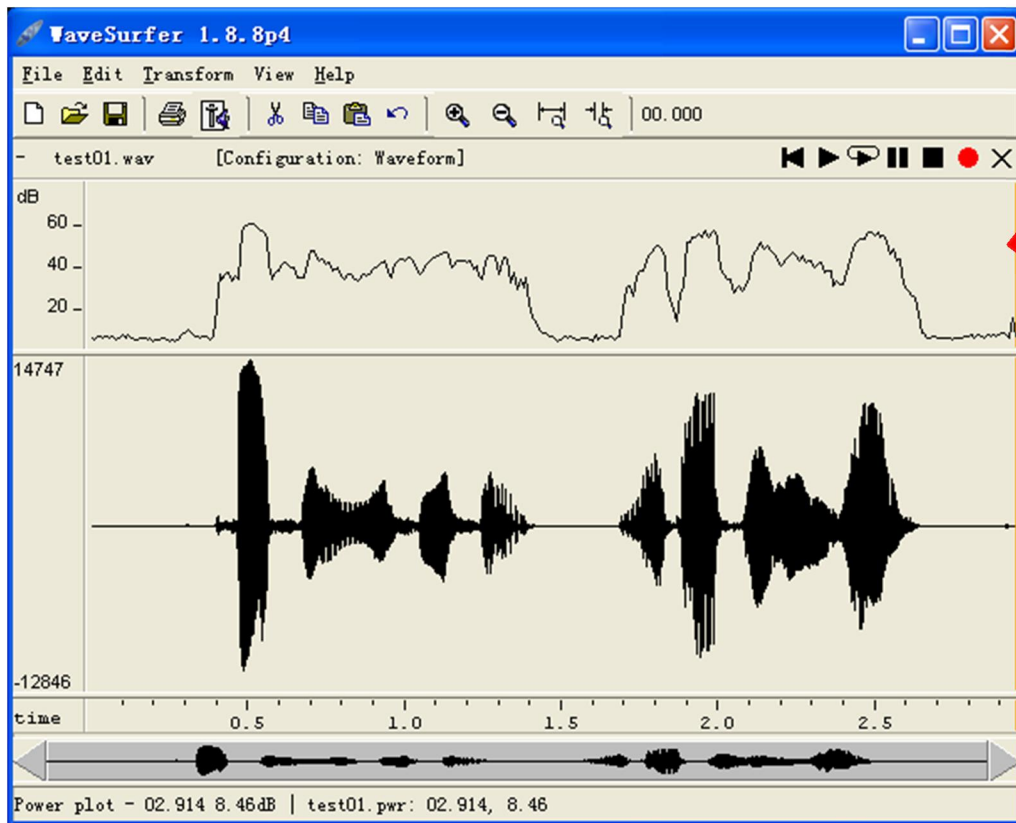
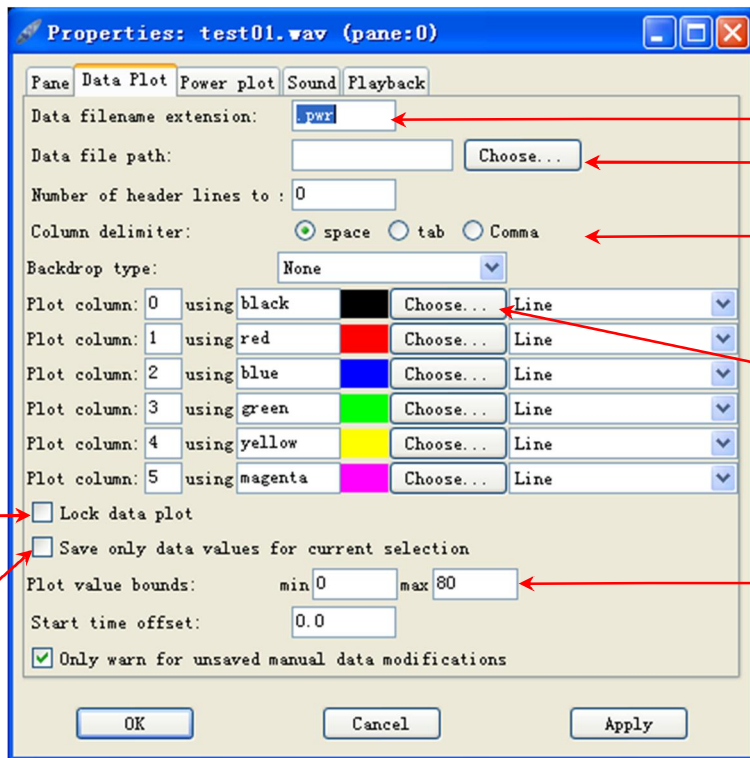


图 7-2: 增加强度分析窗口

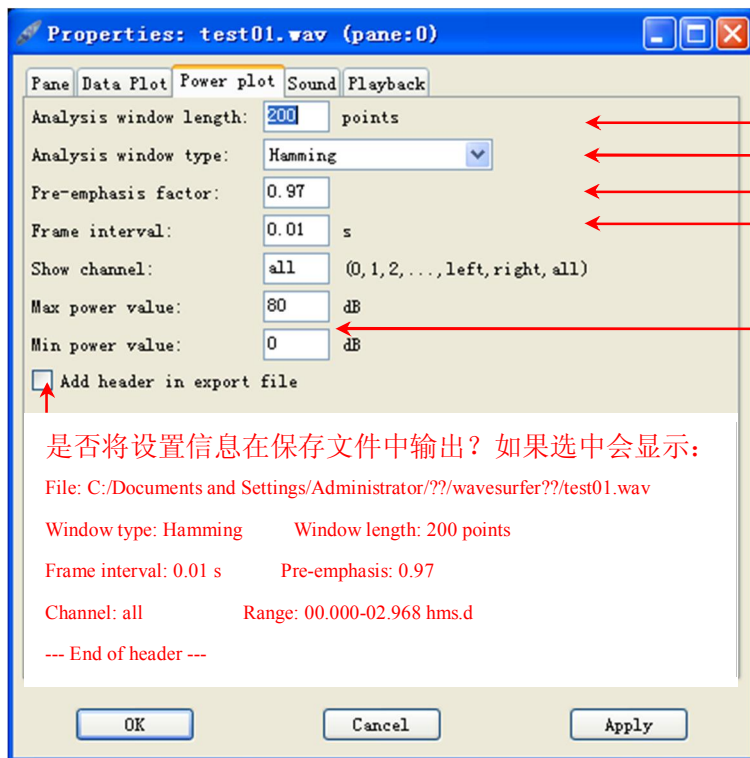
针对强度分析，读者主要需要了解“Data Plot”和“Power Plot”两方面属性的设置情况，此两方面的介绍分别见图 7-3 和图 7-4。



保存数据文件的后缀名  
 保存路径设定  
 多列数据间的分隔符类型：  
 空格、tab、逗号。  
 数据显示颜色设定；形状  
 （线形或点状）设定  
 设定显示数值的范围：  
 最小值与最大值

是否锁定显示的数据，  
 不允许修改？  
 是否只保存选中部分  
 数据的数值？

图 7-3: Data Plot 数据显示属性选项



分析窗长  
 窗口类型：哈明/哈宁等  
 高频提升/预加重系数  
 帧间隔  
 分析数据的最大值与最小值  
 与上图的显示范围不同

是否将设置信息在保存文件中输出？如果选中会显示：  
 File: C:/Documents and Settings/Administrator/??/wavesurfer??/test01.wav  
 Window type: Hamming      Window length: 200 points  
 Frame interval: 0.01 s      Pre-emphasis: 0.97  
 Channel: all                  Range: 00.000-02.968 hms.d  
 --- End of header ---

图 7-4: Power Plot 强度分析属性选项

如果需要将强度数据输出,可以在强度曲线显示区域单击右键,选择“Save Data File” (如图 7-5 所示),可以将强度曲线的数据保存下来,文件名是“test01.pwr”,该文件的后缀名可以进行自主设定,文件可以使用 excel 或记事本软件打开。推荐大家直接设定为.xls 文件,直接使用 excel 打开,便于后续的数据处理和统计分析。

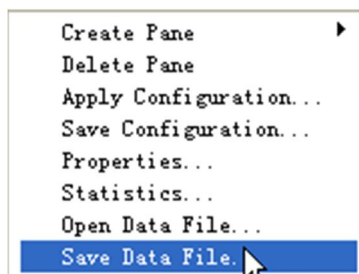


图 7-5: 数据输出



## 8. 语图分析

打开软件后，直接选择 Spectrogram 配置模式（如图 8-1），便可显示语图（默认为 344Hz 的宽带语图），如图 8-2 中所示。或者在波形显示窗口（或其他窗口）单击右键，选择“Create Pane→Spectrogram”，也可进行语图分析，如图 8-3 和图 8-4 中所示。如果不需要显示波形，可以在波形显示窗口单击右键，选择 Delete Pane 即可将波形图删除。

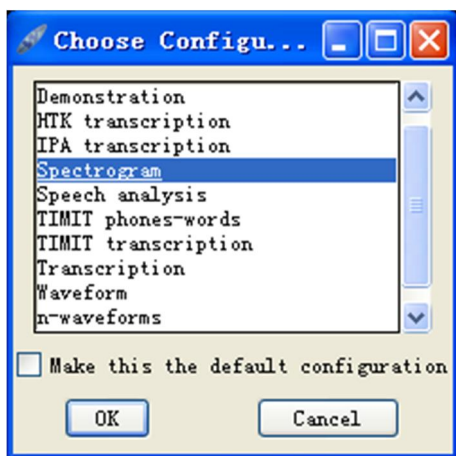


图 8-1: Spectrogram 语图配置模式

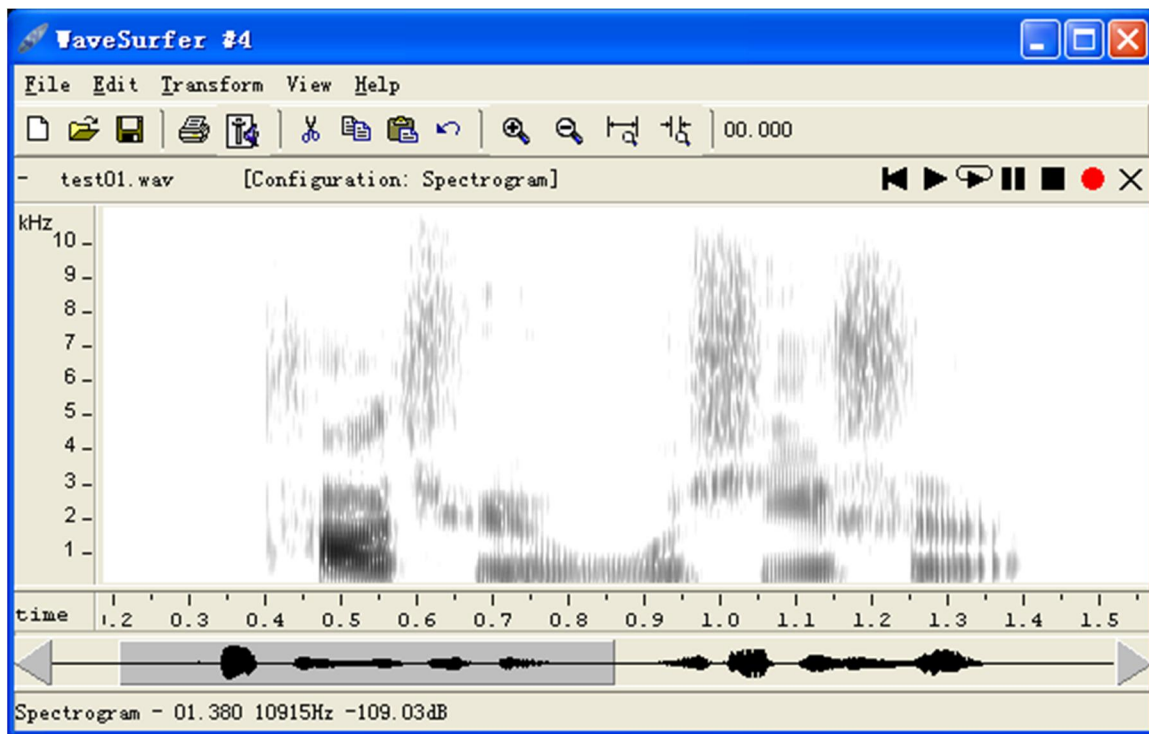


图 8-2: 语图显示示例

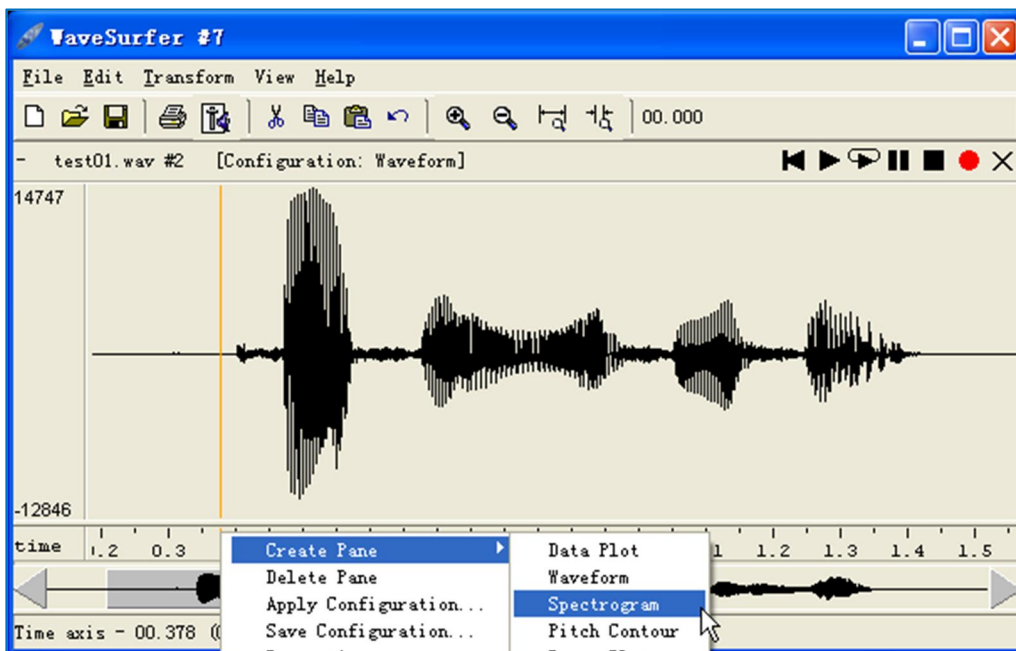


图 8-3: 波形图中增加语图分析选项

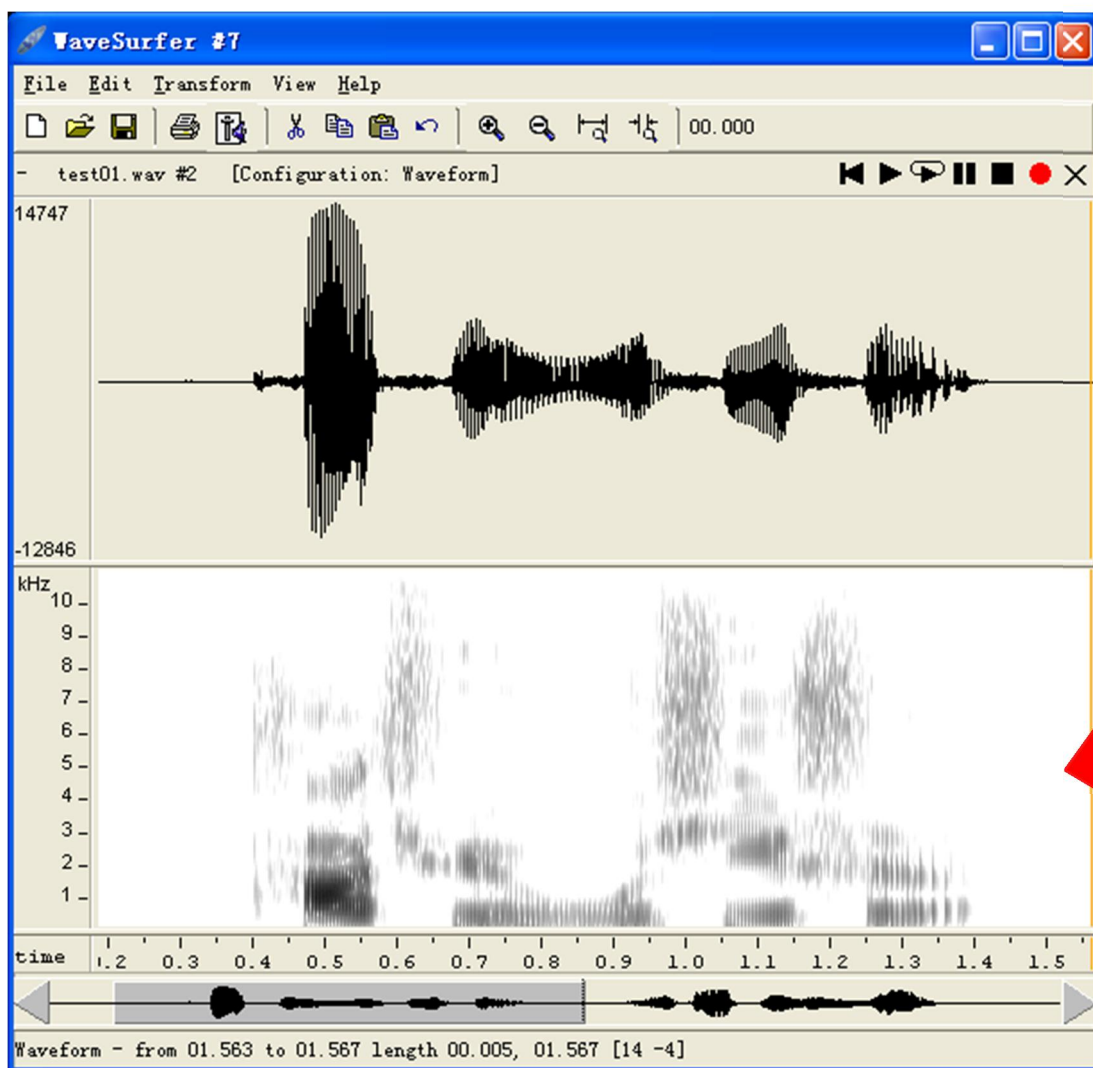


图 8-4: 增加的语图分析窗口

为了得到合适、美观的语图，读者有必要熟悉并掌握语图属性中 Spectrogram 中各个选项的设置情况，具体介绍见图 8-5。

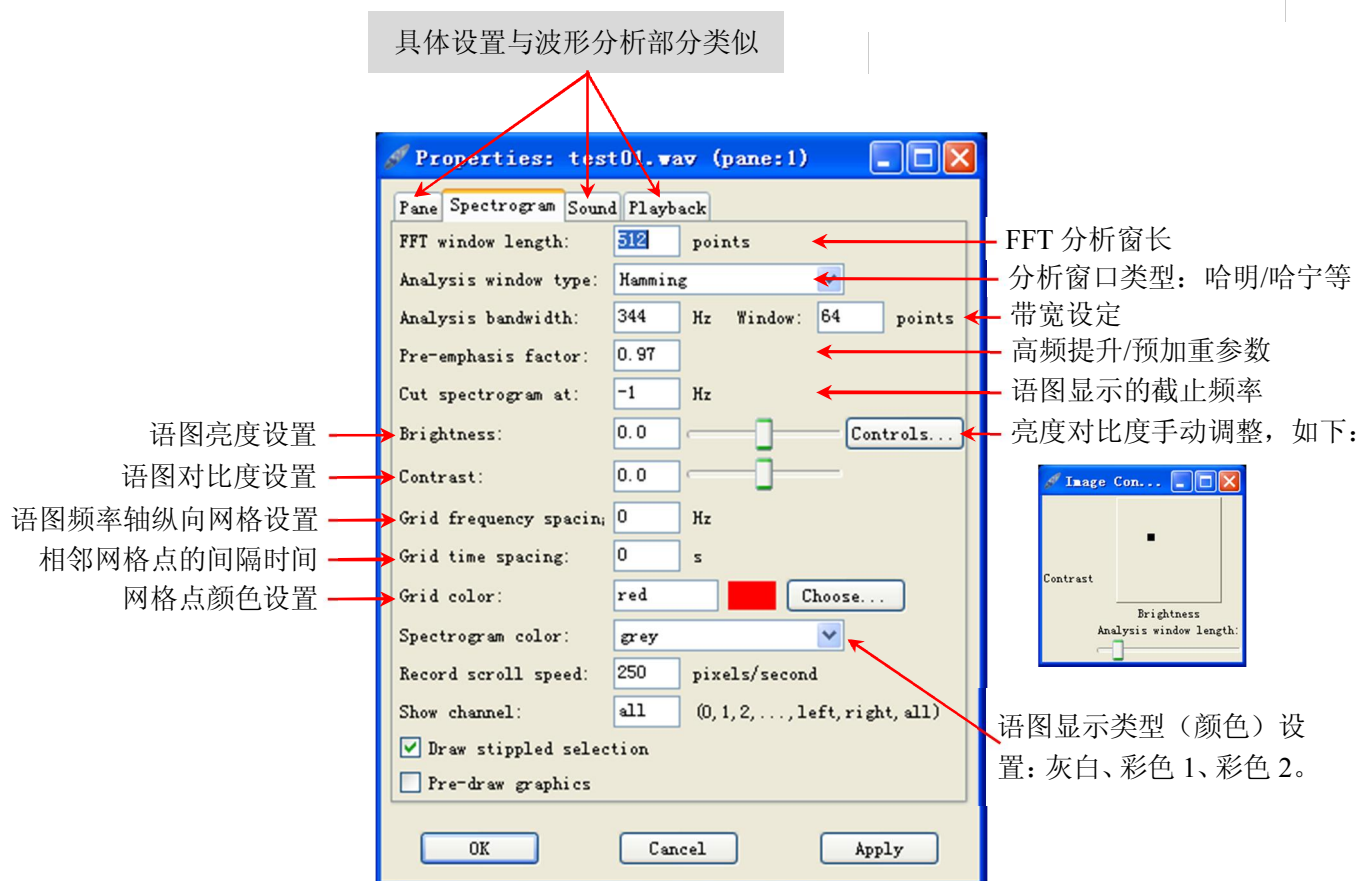


图 8-5: 语图设置选项

#### 提示:

- 1、语图显示的截止频率: 默认设置为“-1Hz”，此时显示语音信号的全部采样频率的一半(根据奈奎斯特定律)，当然，还可以设定为小于上述频率值的任意数值。比如“test01.wav”的采样率为22kHz，图8-2和图8-4中显示语图的频率范围是11kHz，如果只分析元音，可以将“截止频率”设定为5000Hz，语图则会显示为图8-6中的形式。请大家注意，此处设置的只是显示频率范围，并未真正对原始音频进行重新采样(降低采样频率至5000\*2Hz)。
- 2、相邻网格点的间隔时间设置为“0”，则网格显示为直线，如果设置为其他数字，则显示为点“+”，如图8-6中所示，即为“0.1s”的例子。
- 3、WaveSurfer 软件提供了三种语图显示类型，常用的是灰度对比图，另外还设有两种彩色语图，其中之一(color)的样式见图8-6中所示。

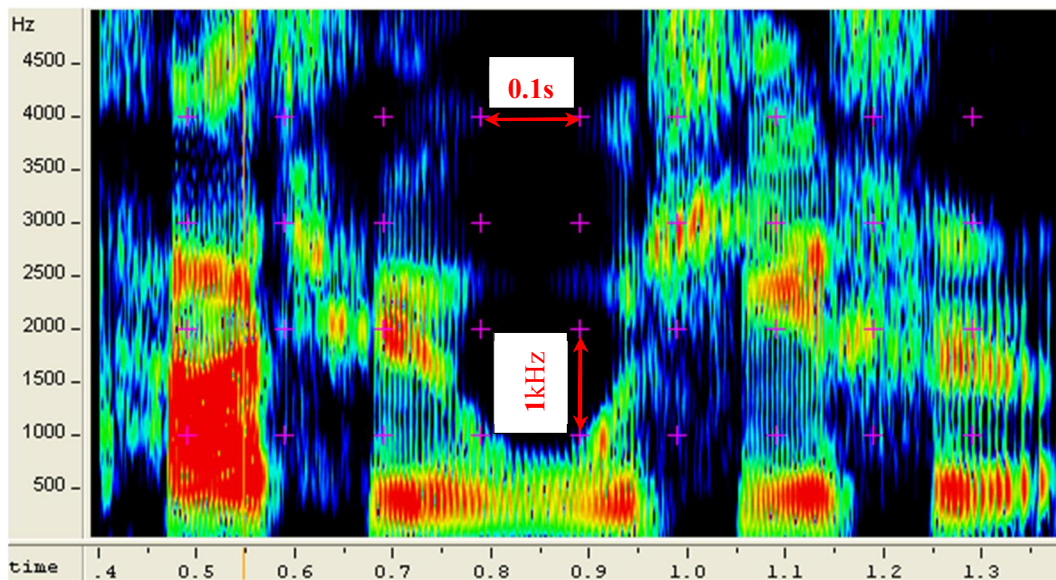


图 8-6: 彩色语图显示示例

4、 如果想将自己的设置保存下来, 便于下次分析时调用, 可以在语图窗口单击右键, 选择“save configuration”进行保存, 可以自设名字, 如“语图\_test01.conf”, 保存在 WaveSurfer 软件默认的路径下 (以 xp 系统为例): C:\Documents and Settings\Administrator\WaveSurfer\1.8\configurations。下次打开文件时即可选择“语图\_test01.conf”, 如下图 8-7 中所示。

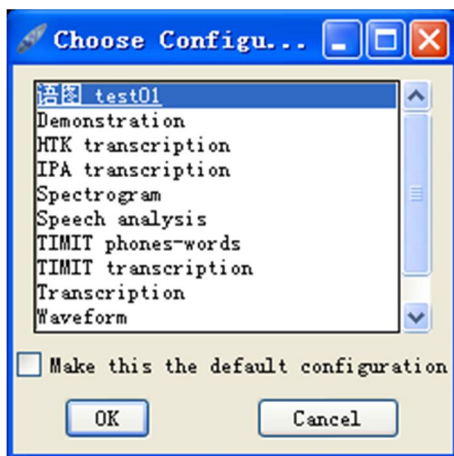


图 8-7: 自行设置的配置模式示例

### 9. 共振峰分析

使用语图模式（或其他任意模式）打开一个语音文件“test05.wav”，并按将其设置为如下图 9-1 中的模式（.lab 标注功能详见本手册第 11 部分），在语图显示窗口单击右键，依次选择 Create Pane→ Formant Plot 选项，即可进行共振峰分析，如图 9-1 中所示。

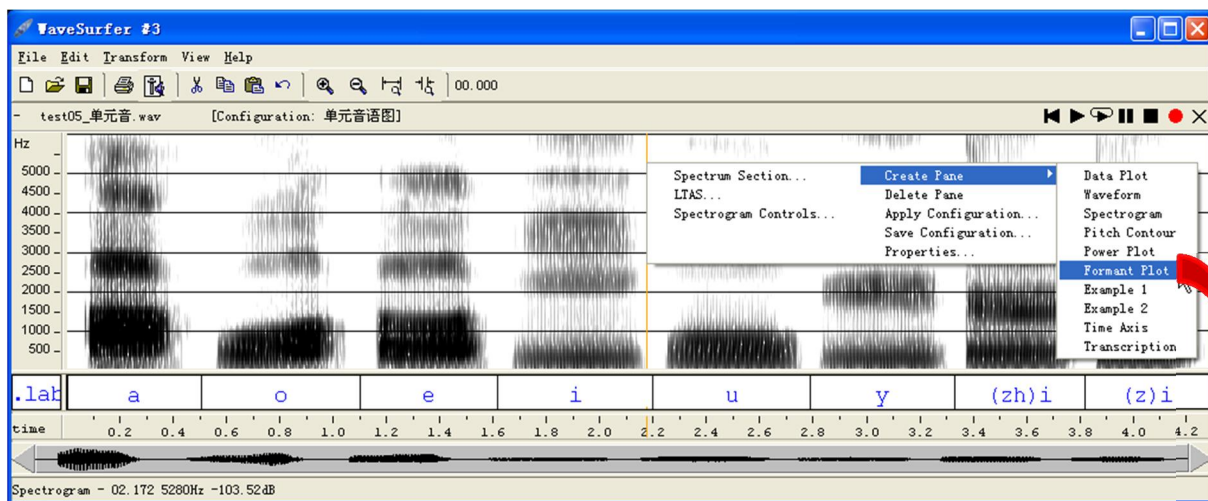


图 9-1: 在语图中创建共振峰分析窗口

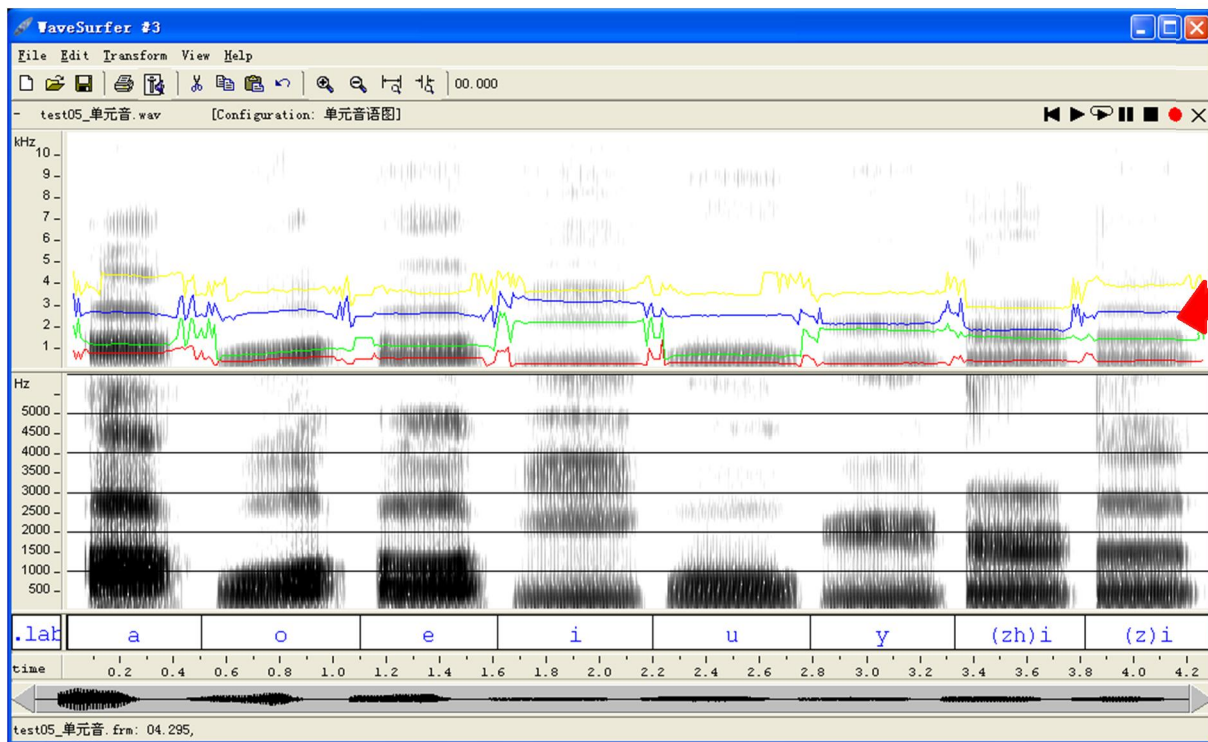


图 9-2: 增加的共振峰分析窗口

右键单击共振峰显示窗口，选择属性（Properties）进行设置。此处（带共振峰的）语图选项的设置方式与只显示语图时的设置方式相同，如果选择下图 9-3（同前）中显示的参数进行设置，则会出现图 9-4 中显示的模式。

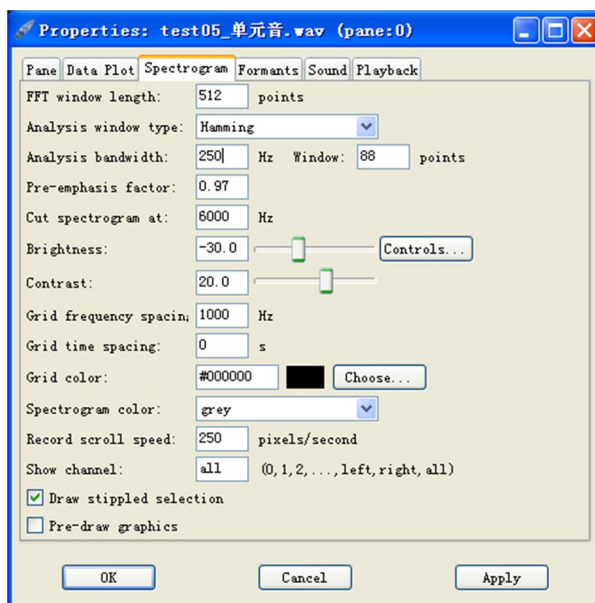


图 9-3: 语图设置

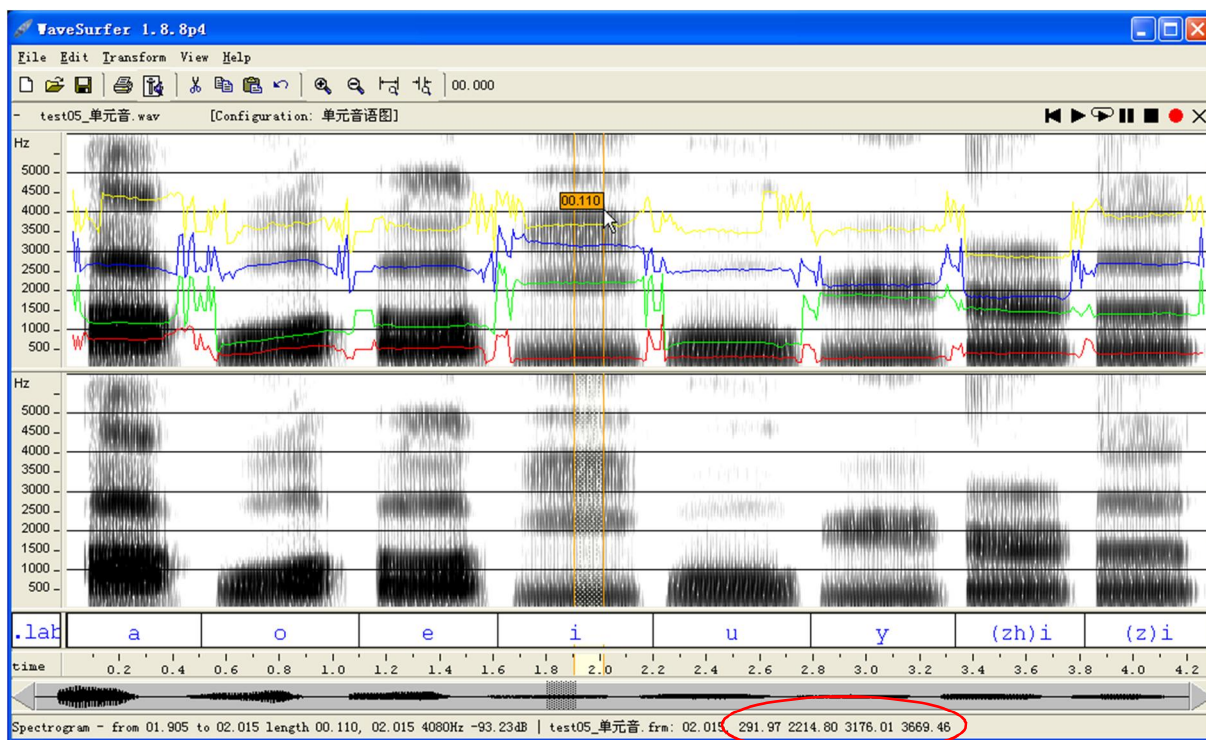


图 9-4: 修改后的语图显示示例

为了准确提取共振峰数据，读者有必要熟悉并掌握共振峰属性中 Data Plot 和 Formants（共振峰）中各个选项的设置情况，具体介绍见图 9-5 和图 9-6。

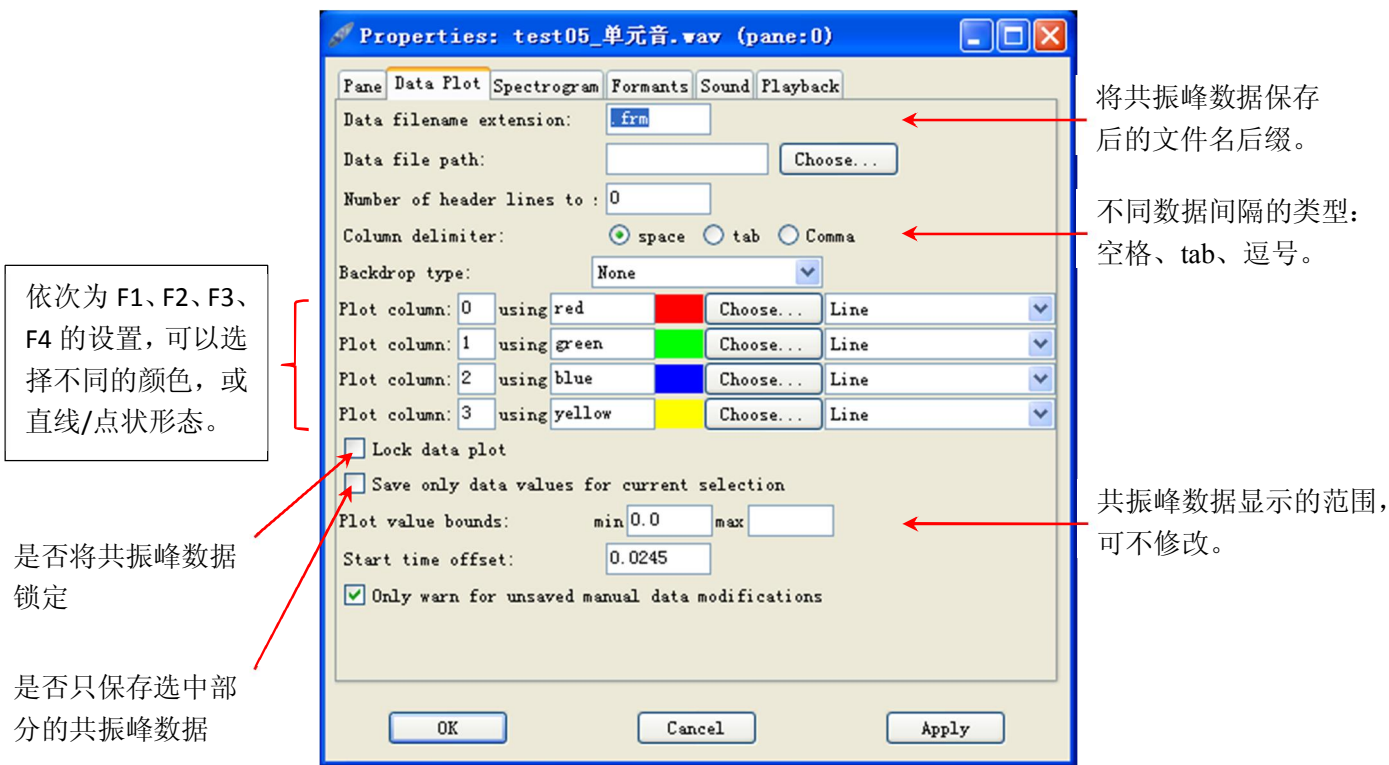


图 9-5: Data Plot 数据输出设置选项

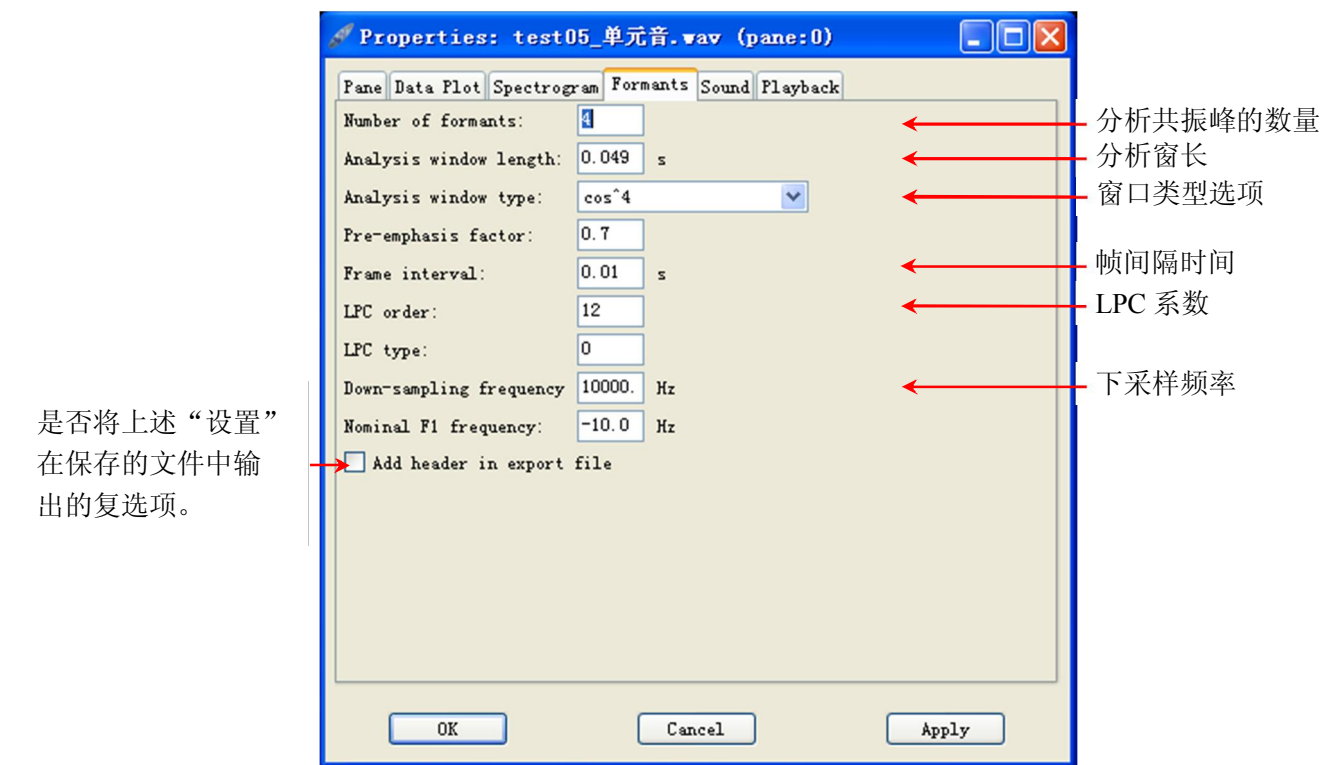


图 9-6: Formants 共振峰设置选项

**提示:**

- 1、为了准确提取共振峰数据，需要针对不同的语音情况（如男声、女声等）做出不同的调整，一般要配合调整共振峰数量、LPC 系数和下采样频率三个选项，一般来讲，“4/12/8000”的组合可以比较准确的提取前 3 条共振峰数据，如果出现提取不准的情况，可以提高 LPC 系数（如 14/16），或改变下采样频率。
- 2、如果想分析更多数量的共振峰，在调整数量选项时，须满足：LPC 系数 $\geq$ 共振峰数量\*2+2 的要求。
- 3、提取共振峰数据

(1) 如果要提取一段语音共振峰的平均值，可以用鼠标选中一段语音（或使用标注功能，选中 Select Label），如图 9-4 中所示，单击鼠标右键选择“Statistics”选项，即会出现如下图 9-7 中显示的对话框：

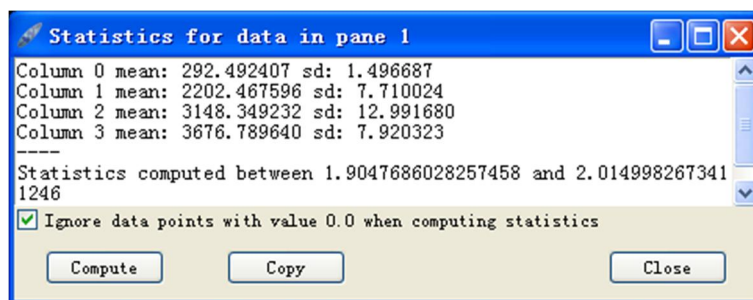


图 9-7: 各共振峰均值统计

该图中给出的是所选（光标间）语音段前 4 条共振峰各自的均值（mean）和标准差（sd），以及所选语音段的起始和结束位置时间点。选择“Copy”选项复制或直接选中其中的数字复制后可以粘贴到 word/excel 文件中进行分析。

(2) 如果要选择光标处的共振峰数据可以直接在语图右下角的信息栏中读取，如图 9-4 中的红色区域所示。

(3) 如果要提取共振峰轨迹的数据，可以单击鼠标右键，选择“Save Data File”，在保存的“test05\_单元音.frm”文件中直接读取和分析（注：可以在图 9-5 的设置中直接设置为 xls 后缀名，保存的文件可以直接用 excel 打开，但该文件并不是真正的 xls 格式，如想保存为 xls 格式，可以打开后另存）。



## 10. 基频分析

使用波形模式打开一个语音文件“test02\_声调”（标注功能详见本手册第 11 部分），在 time 时间轴显示框（或其他任意显示框）单击鼠标右键，依次选择“Create Pane → Pitch Contour”，如图 10-1 中所示，即可显示基频信息，如图 10-2 中所示。

为了获得清晰、美观、准确的基频数据，大家需要熟悉并掌握基频属性中 Data Plot 和 Pitch contour（基频曲线）中各个选项的设置情况，具体介绍见图 10-3 和图 10-4。

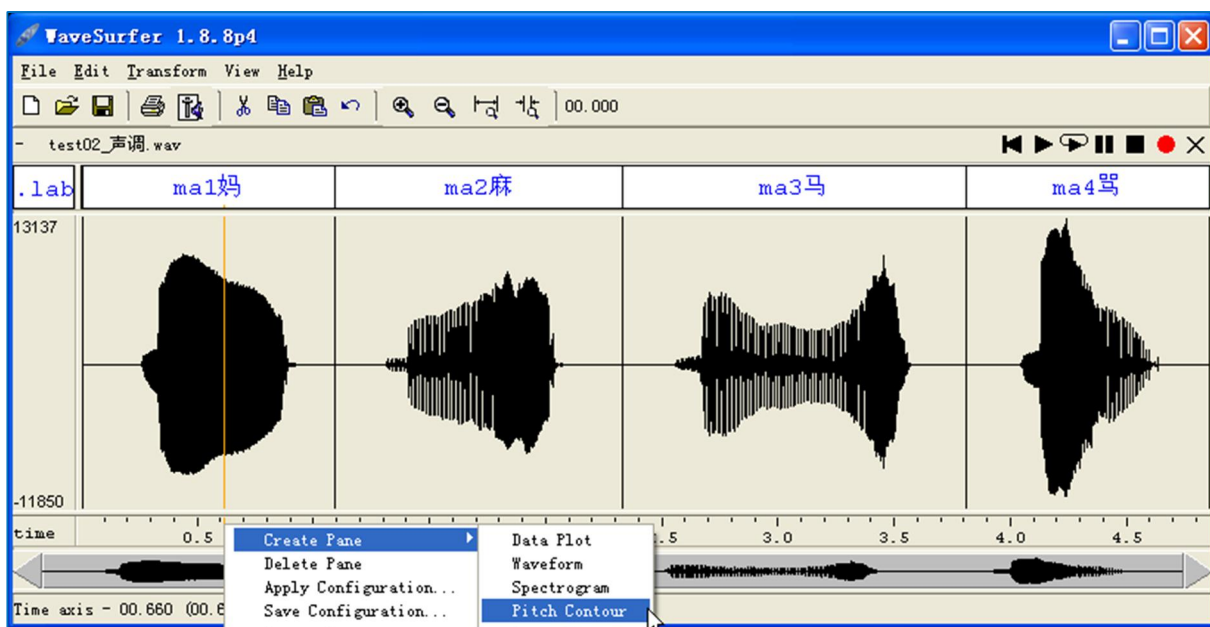


图 10-1：由波形图中创建基频分析窗口

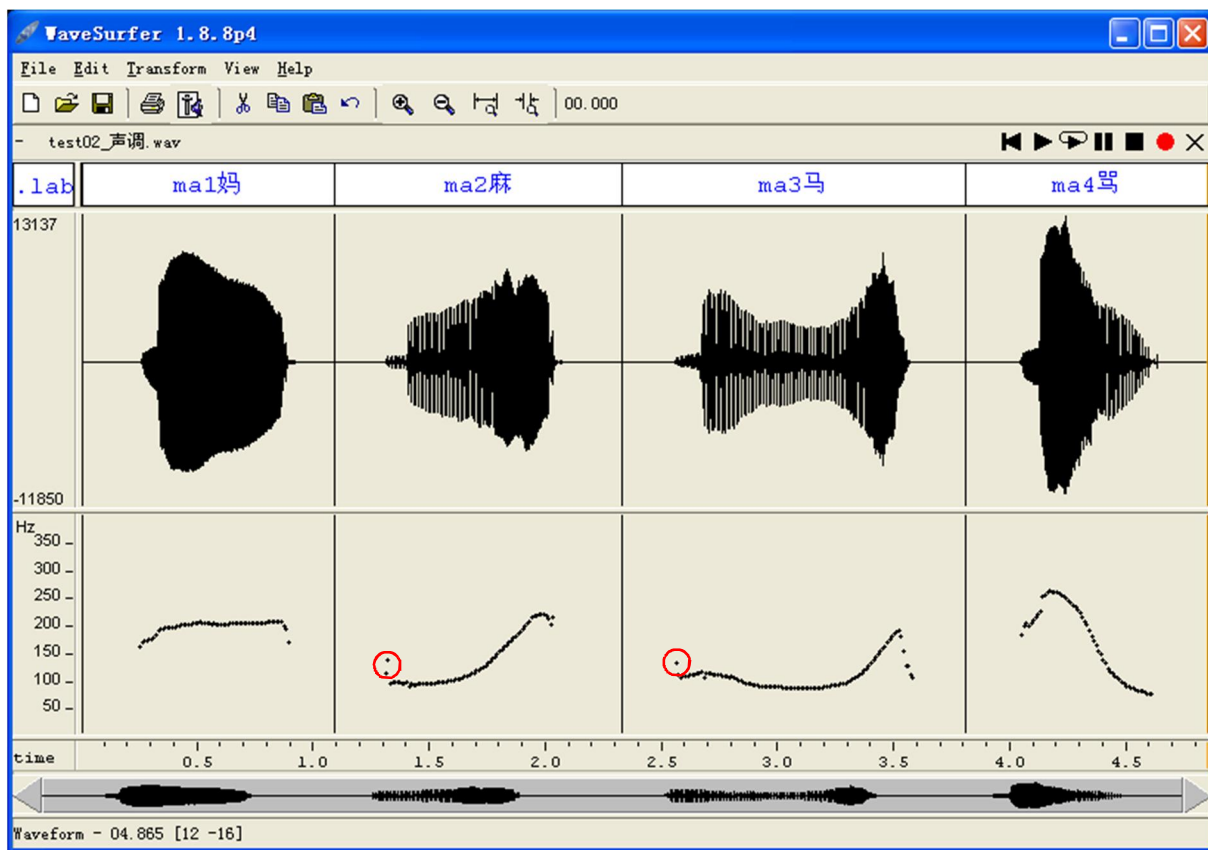
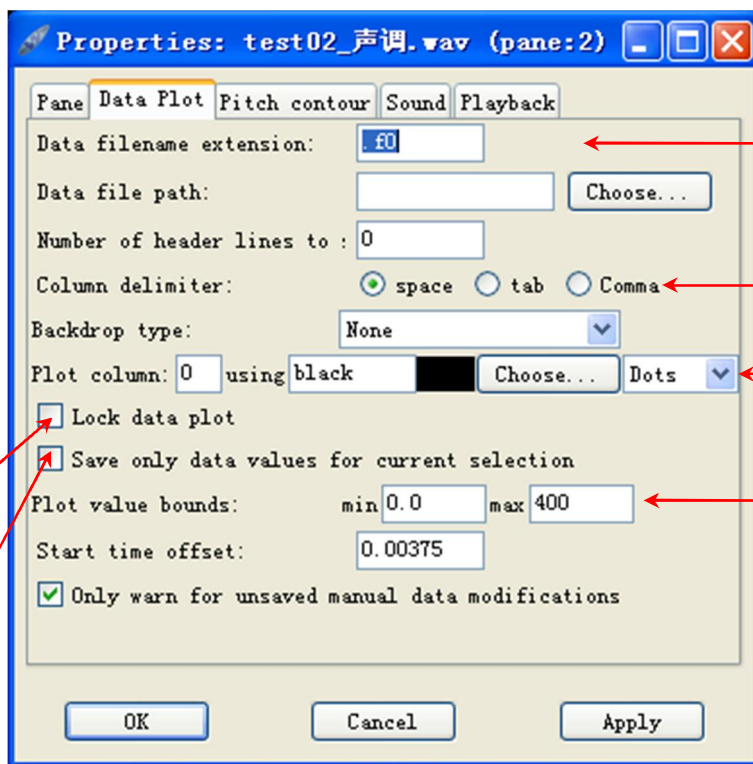


图 10-2: 增加的基频分析窗口



保存基频文件的后缀名

不同数据间隔的类型：  
空格、tab、逗号。

基频数据的颜色设置  
显示类型：点状/线形

显示基频数据的范围

是否将基频数据锁  
定复选项

是否只保存选中部  
分的基频数据

图 10-3: Data Plot 属性设置

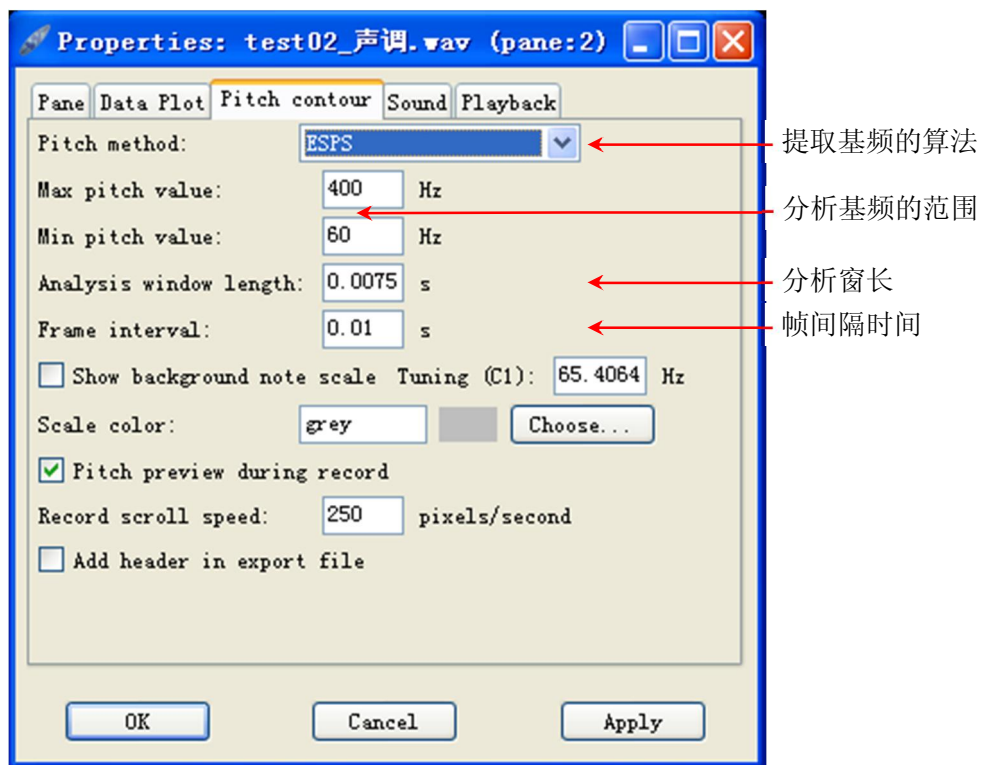


图 10-4: Pitch contour 基频曲线属性设置

提示:

1、WaveSurfer 提供了两种分析基频的方法:

(1) ESPS (Entropic Speech Processing System) 即基于 ESPS/Waves 软件包的自相关算法(ACF:Auto Correlate Function)。

(2) AMDF (Average Magnitude Difference Function) 平均幅度差函数。

读者可以对比两种算法提取基频的差别,但同一次分析最好选择同一种算法,便于比较。

2、提取基频数值 (同共振峰分析)

(1) 在基频显示区域,选中一段语音(光标间),点击鼠标右键,选择“Statistics”,出现图 10-5 中显示的对话框,可以将基频的均值 (mean) 和标准差 (sd) 复制后进行分析。

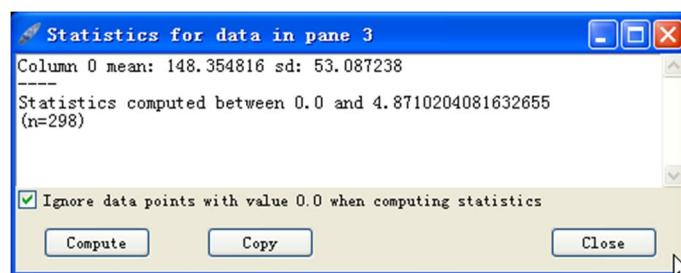


图 10-5: 基频均值统计

(2) 在基频显示区域，亦可以单击鼠标右键，选择“Save Data File”，在保存的“test02\_声调.f0”文件中（用 excel/记事本软件）直接读取和分析。

- 3、对于部分语音的基频数据提取不准确的地方（如图 10-2 中红色圆圈中所示），在未选择“Lock Data Plot”选项时，可以进行手动修改，将改动后数据保存（“test02\_声调.f0”），下次分析时，可以在基频显示区域，单击鼠标右键，选择“Open Data File”读取原保存的文件数据，实现可重复分析。

## 11. 语音标注

WaveSurfer 软件提供了多种标注模式选项：HTK/ISIS、PHONDAT、BOSS、TIMIT、WAVES、WaveSurfer 和 IPA（国际音标）等，其中软件自带的模式就有 4 种（见图 11-1），分别是 HTK transcription、IPA transcription、TIMIT transcription、和（WaveSurfer）Transcription。

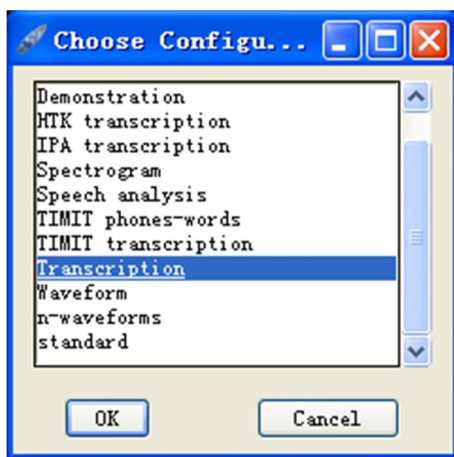


图 11-1:4 种默认的标注配置模式

以（WaveSurfer）Transcription 为例，打开一个语音文件“test03\_标注.wav”，如图 11-2 中所示，显示的是语图（上）和“.lab”标注栏（下）。结合上文中介绍的方法略作调整（添加波形图，修改语图）后即可变为图 11-3 中显示的形式。在“.lab”标注栏中单击鼠标右键，选择“Properties”属性选项，如图 11-4 中所示，根据分析需要查看或修改具体设置。读者可以根据自己的需要设置单（多）层标注体系，如图 11-5 中所示。

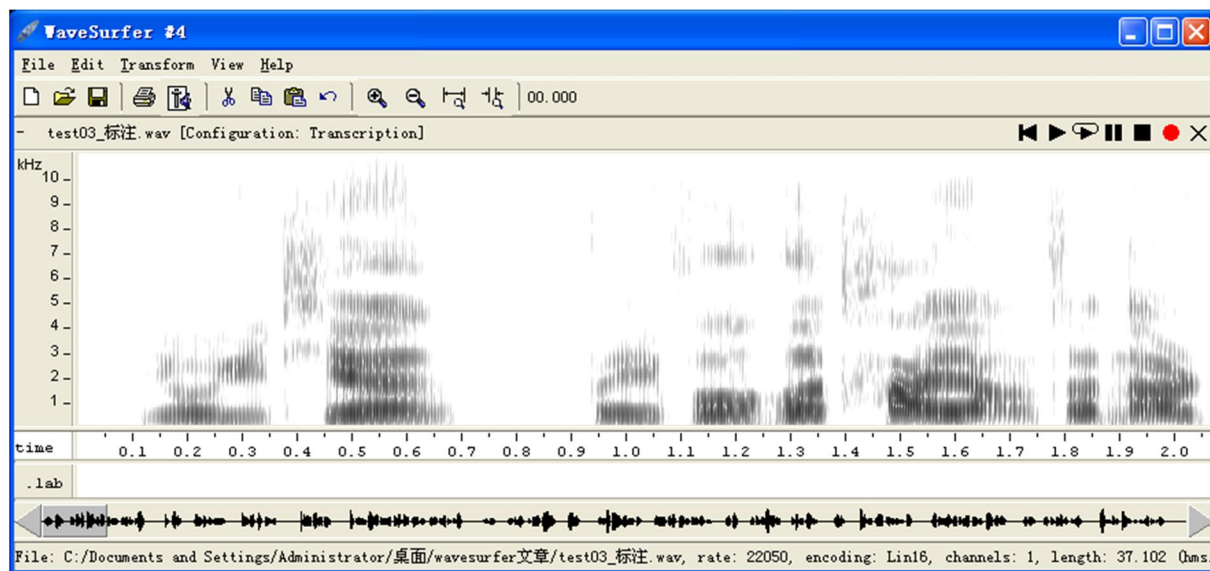


图 11-2:（WaveSurfer）Transcription 标注示例

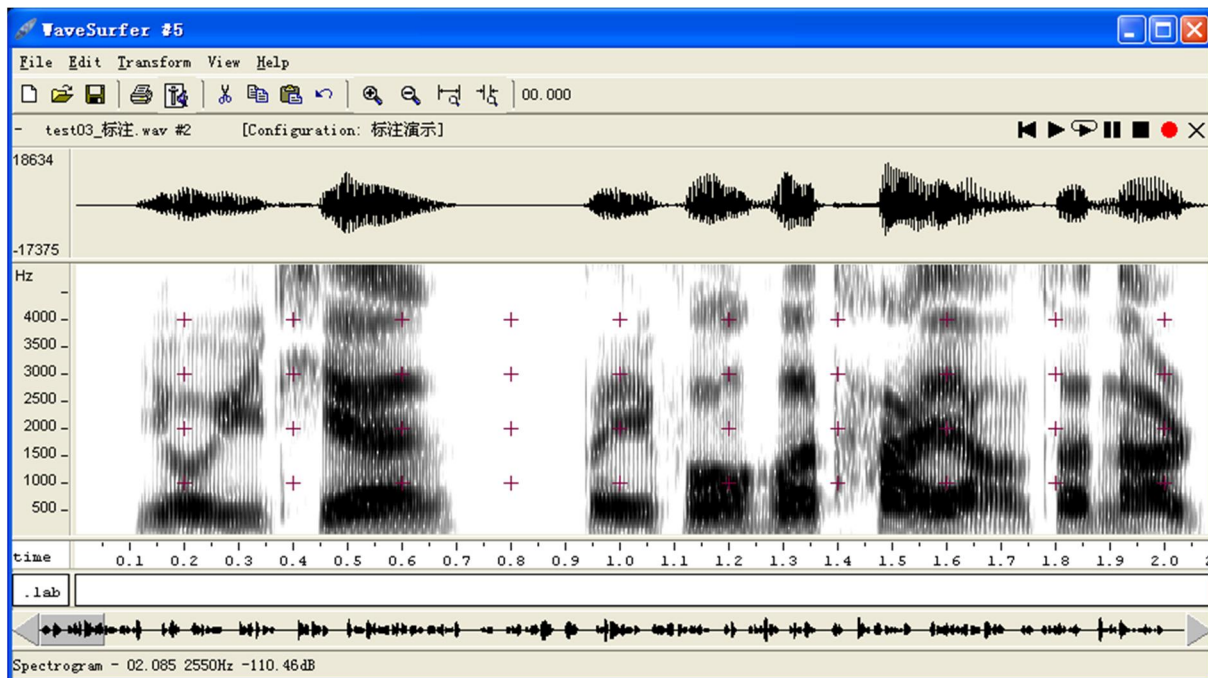


图 11-3: 调整后的标注窗口

标注文件格式

标注文字 3 种排列格式:  
e 靠右/c 居中/w 靠左对齐

保存文件的后缀名

标签颜色

标签之间界限颜色

标签内容背景颜色

是否锁定标签内容

字体/字号设置

在此可以设置快速文字/符号, 通过设置列/行调整输入的信息量。

图 11-4: 标注属性设置

是否将标签的分界扩展到波形/语图/基频/振幅等其他显示栏?

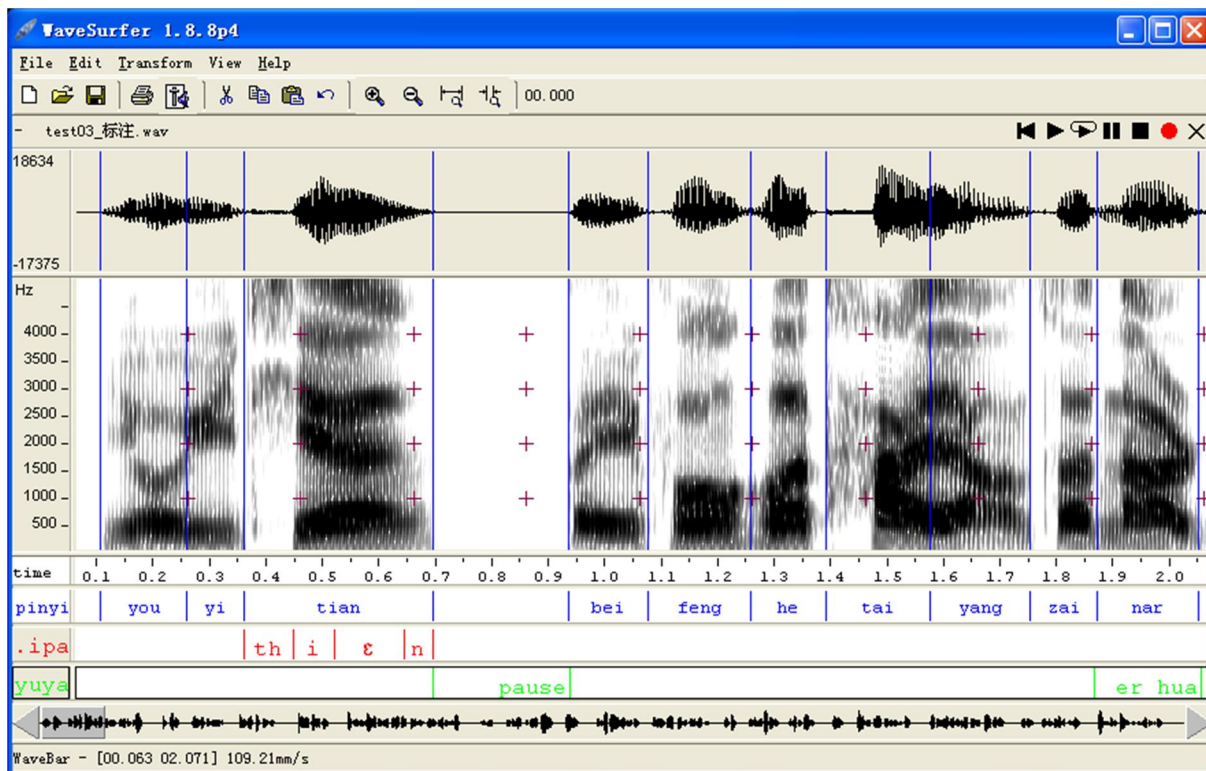


图 11-5: 多层标注体系示例

## 提示:

- 1、在某层标签显示框内，单击鼠标右键，选择“Insert Label”（如图 11-6）即可输入文字/符号，也可直接将光标放到标签显示栏（想输入信息的位置）中直接输入文字/符号，如果想调整位置，可选中标签边界直接进行拖动。如果想删除标签，可在该标签显示的区域内，单击鼠标右键，选择“Delete Label”即可。

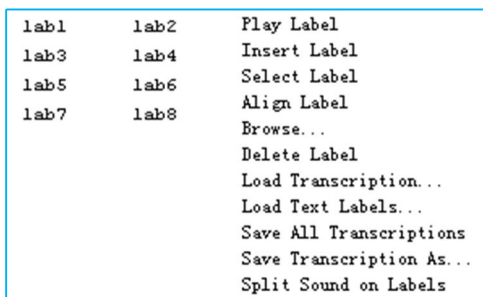


图 11-6

- 2、单击鼠标右键，选择“Select Label”（如图 11-6），选中标签段的语音，即可进行分析：放音/提取数据。
- 3、如果想对某标签中的大量标注内容进行查询，可以选择“Browse”（如图 11-6）选项对特定内容进行查询。如图 11-7 中所示，想查找录音中包含复元音“ai”的出现次数和位置，可以直接输入“ai”点击“Find”按钮，即会在结果栏显示

出两处（标注信息、起止时间）标注“tai”与“zai”。

- (1) 单击“tai”，软件会自动选中该标签，并将其显示到语图的中间位置；
- (2) 双击“tai”，软件同上会自动选中该标签，将其显示到语图的中间位置，同时会自动从标签起始位置放音（Select Start）或者只播放标签区域的语音（Select Word）。

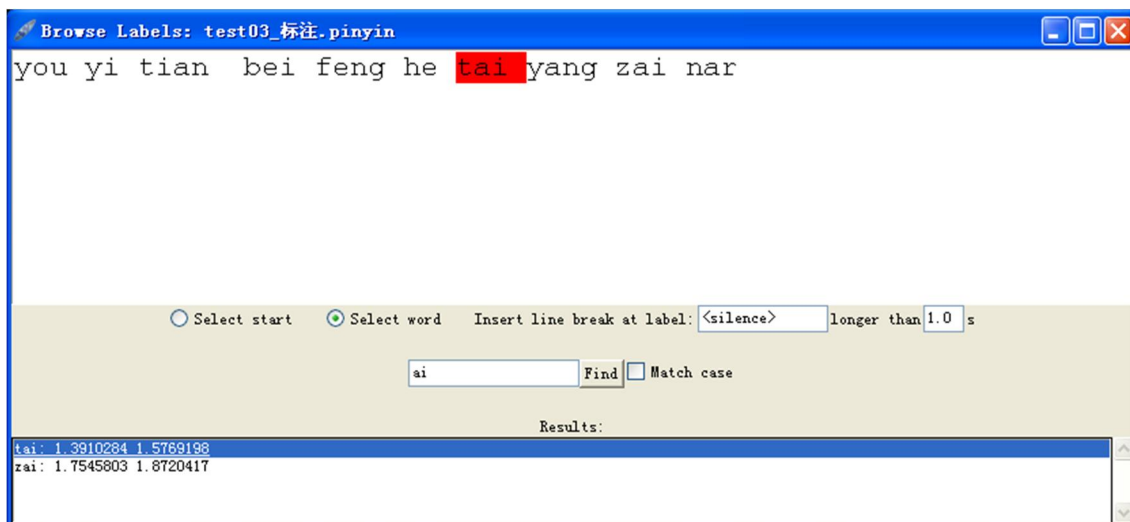


图 11-7: 标注内容查询窗口

- 4、单击鼠标右键，选择“Save (All) Transcription As”对标注信息进行保存。
- 5、如果想读入之前已经标注的信息，可以选择“Load Transcription”选项。
- 6、如果选择“Split Sound on Labels”选项，则可按照标注位置对原始语音进行切分，并自动保存在自动生成的“test03\_标注.wav.split”文件夹内，切分后语音片段将按原标签的顺序进行编号排列，如图 11-8 所示。

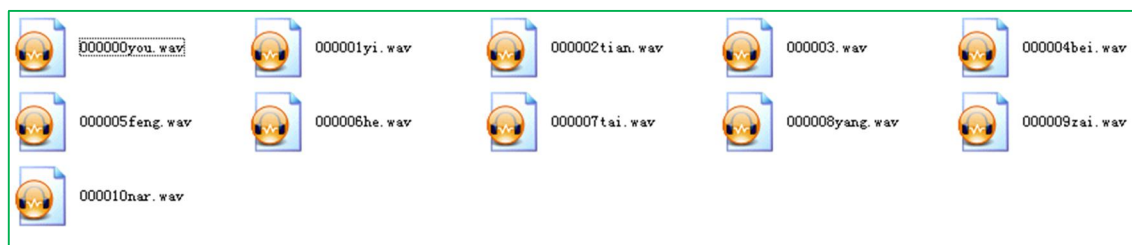


图 11-8: 按标注内容进行语音切分

- 7、如果要想获得标签各个位置的准确数值，可以使用 excel/记事本软件打开保存的标准文件（如“test03\_标注.pinyin”），进行进一步的统计分析，见图 11-9 中所示。



	A	B	C
1	0.1075000	0.2602410	you
2	0.2602410	0.3608675	yi
3	0.3608675	0.6950879	tian
4	0.9350000	1.0777087	bei
5	1.0777087	1.2583057	feng
6	1.2583057	1.3910284	he
7	1.3910284	1.5769198	tai
8	1.5769198	1.7545803	yang
9	1.7545803	1.8720417	zai
10	1.8720417	2.0511704	nar

图 11-9: 保存的标注内容及起止时间点信息

## 12. 数据分析

前面已经对如何提取共振峰数据、基频数据进行了简单的说明，如果需要对其进行进一步的分析，可以使用 excel 软件对保存的 frm 和 f0 文件进行数据提取和分析。

下面介绍 WaveSurfer 软件的 FFT 快速傅里叶变换、LPC 线性预测技术和 LTAS 长时平均功率谱的功能。首先使用语图模式打开“test05\_单元音.wav”文件，如图 12-1 中所示：

### 1、FFT

(1) 在语图中显示的单元音[a]上，点击鼠标右键，选择“Spectrum Section”选项，即可进行光标处的 FFT 分析，如图 12-2 中所示。

(2) FFT 图中显示的频率范围语图中的设置“显示”的范围是一致的，比如此处显示的是 10kHz，而语音信号的频率范围是  $22\text{kHz}/2=11\text{kHz}$ 。

(3) 选中“snapshot”选项下的“test05\_单元音.wav”，即可将刚才的分析数据转换为“红色”图像，如图 12-3 中所示。此时可以对另外的语音进行（FFT 或 LPC）分析，如图 12-4 中红色数据为刚才对元音[a]进行的 FFT，黑色数据为对元音/e/（[ə]）某处的 FFT 分析，可以对两者进行比较分析。

(4) 点击“Export”选项，可以将 FFT 分析的数据输出保存为“spectrum.txt”文件（可以修改文件及后缀名），使用 excel/记事本软件进行统计分析。

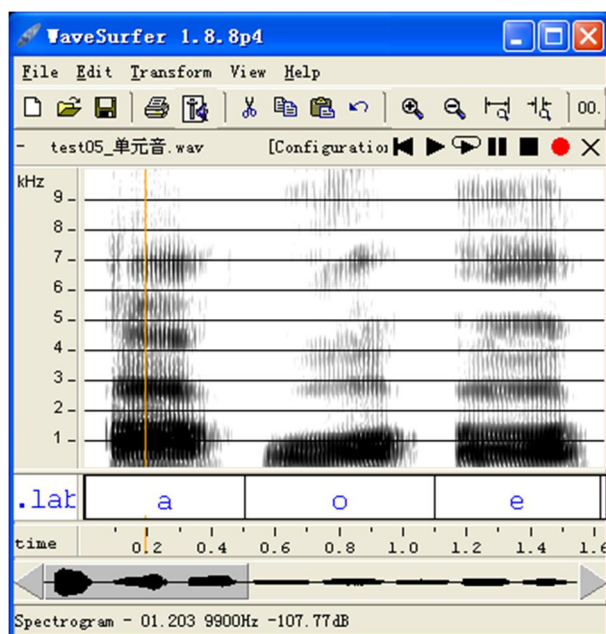


图 12-1：语图模式示例

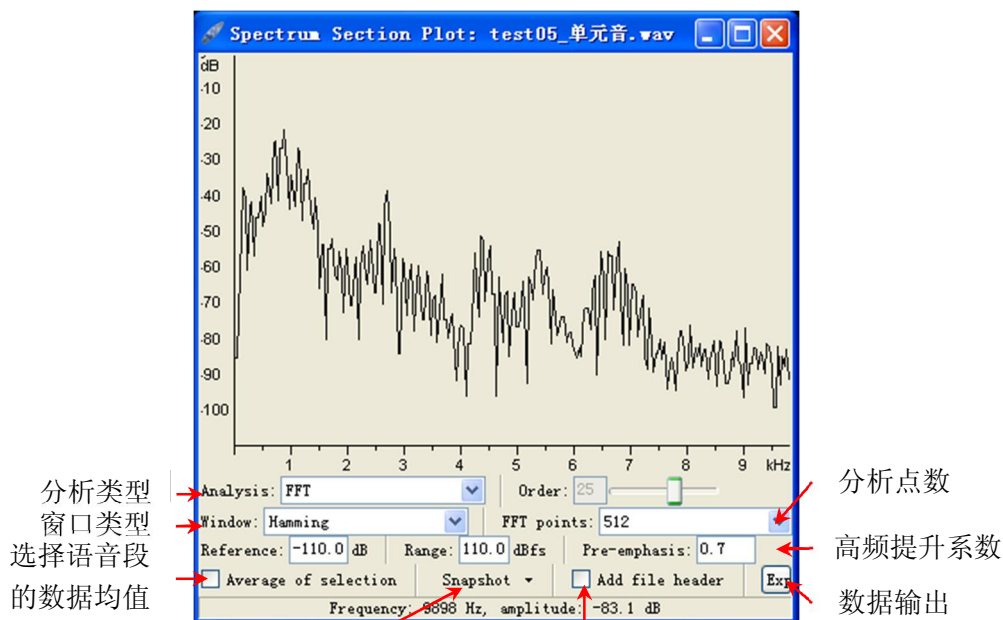


图 12-2: FFT 分析窗口示例

固定数据为图像

是否将设置信息输出?

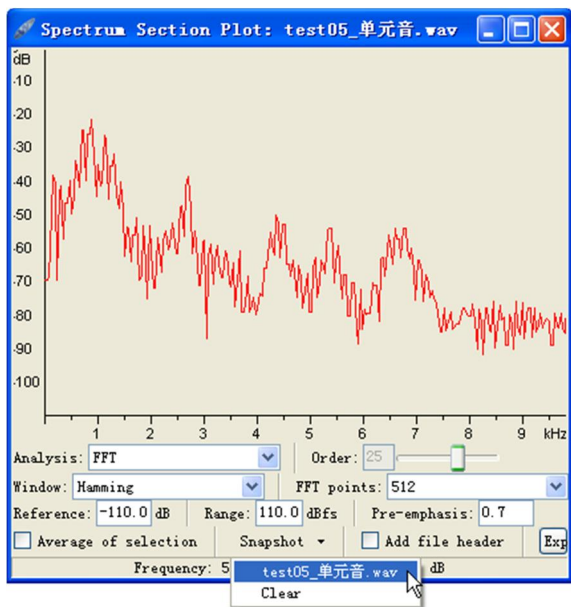


图 12-3: 元音 a 某处的 FFT 分析图

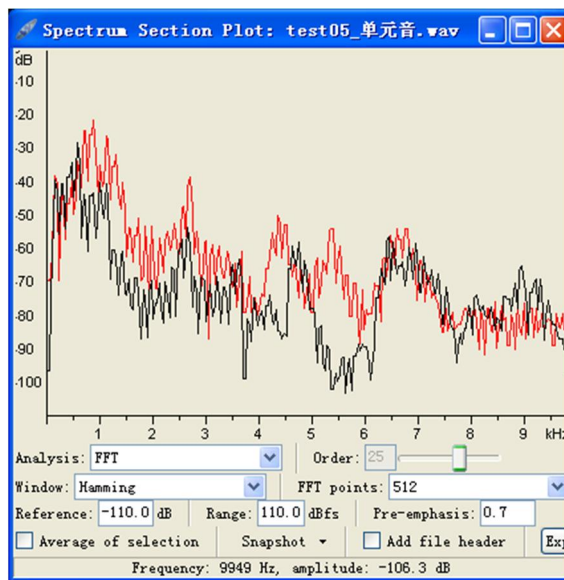


图 12-4: 元音 a 与元音 e 的 FFT 比较

## 2、LPC 分析

将上图 12-2 中分析类型 FFT 改选为 LPC 即可进行 LPC 分析。如下图 12-5。

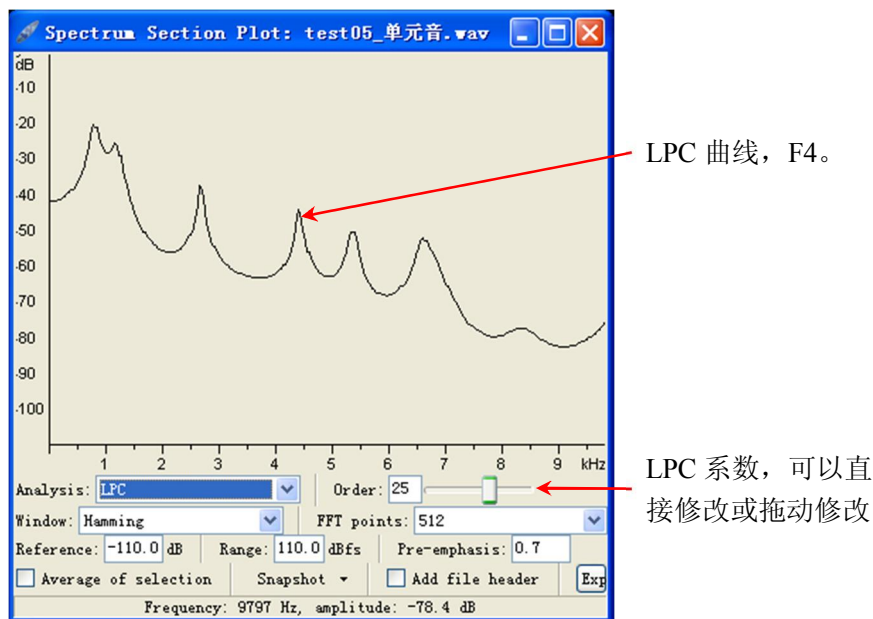


图 12-5: LPC 分析示例

(1) 此 LPC 分析图中显示的频率范围与语图中的频率范围也是一致的，此图为 10kHz。请注意：此处的 LPC 系数设置与共振峰分析中的 LPC 系数设置不同，后者与提取的共振峰数目以及下采样频率有关，比如同时比较 LPC 系数为 25（图 12-6）和 12（图 12-7）时 LPC 曲线与 FFT 的关系，可以看出当 LPC 系数较大时，LPC 分析的极点与 FFT 之间有更好的吻合，即使将语图的频率显示范围降低，如 5kHz，重复上述步骤，如图 12-8 与图 12-9，结果还是当 LPC 系数较大时，LPC 分析的极点与 FFT 之间有更好的吻合。其实不难看出图 12-8 与图 12-9 实际分别是图 12-6 与图 12-7 的一部分（0-5kHz 范围）的放大，WaveSurfer 软件此功能的 FFT 与 LPC 分析都是基于对音频的原始采样频率进行分析的，不同于共振峰分析中与下采样频率（在分析中实际上进行了重新采样）有关的 LPC 系数设定，这一点从输出的“spectrum.txt”文件中也可以清楚地看到。

(2) 如果想从此功能中获得比较准确的 LPC 分析，可以将原始音频信号降低采样率，如 8/10/11kHz，然后再作分析。例如，对 11kHz 的“test05\_单元音 11k.wav”进行分析，如图 12-10 与图 12-11 中所示，可以看出当 LPC 系数为 12 时，比较准确地提取了元音[a]的前 4 条共振峰。

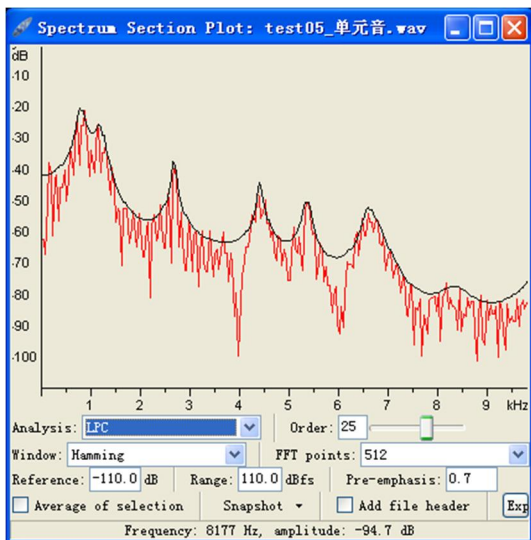


图 12-6: LPC 系数为 25

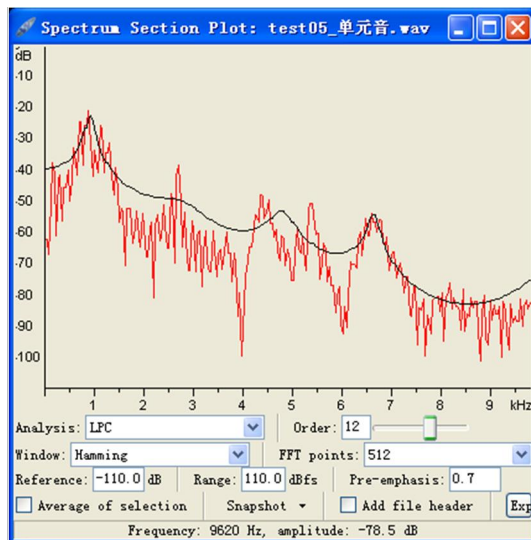


图 12-7: LPC 系数为 12

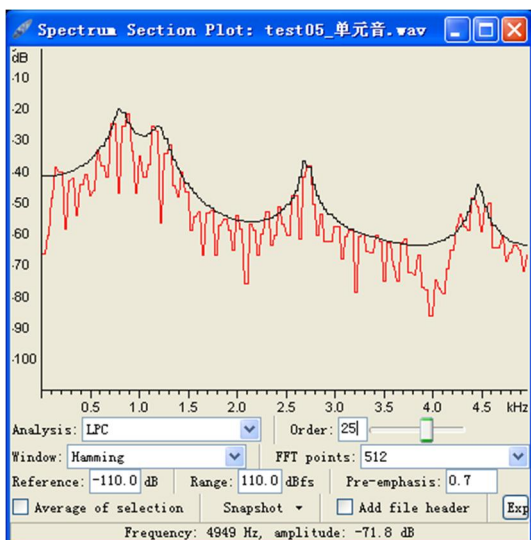


图 12-8: LPC 系数为 25

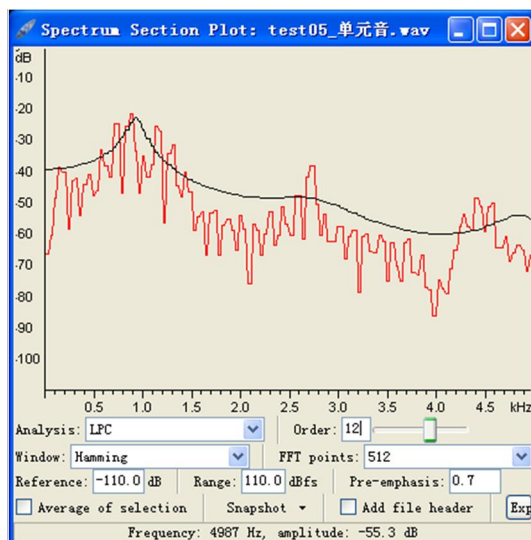


图 12-9: LPC 系数为 12

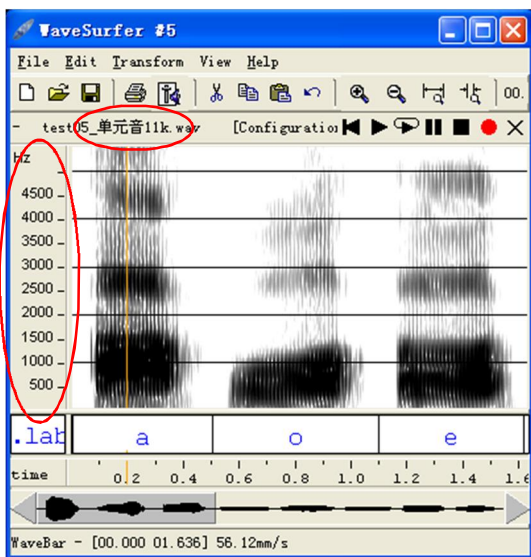


图 12-10: 11kHz 采样时的语图

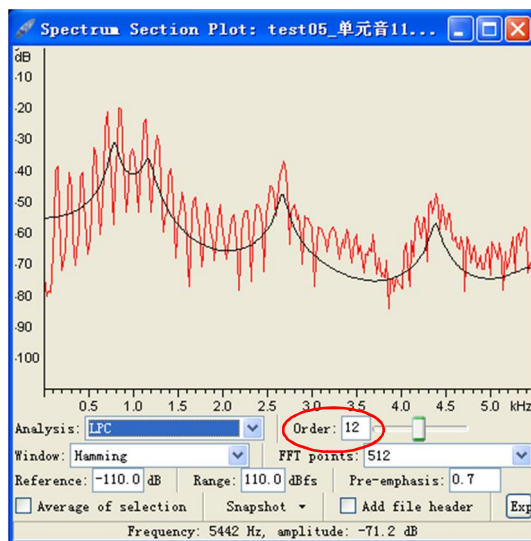


图 12-11: 11kHz 采样时 LPC 与 FFT 分析

### 3、LTAS 分析

该分析操作甚为简单，在语图中单击鼠标右键，选择“LTAS”选项即可，如下图 12-12 中所示。可以看出该图与 FFT 分析基本一致，只是选中了“Average of selection”选项（图 12-12 中红色标示处）。因此，也可以在做 FFT 分析的基础上，选中“Average of selection”，然后再在语图中选择“ctrl+a”也可实现 LTAS 的分析。

LTAS 数据的保存和分析同前 FFT 与 LPC 分析。

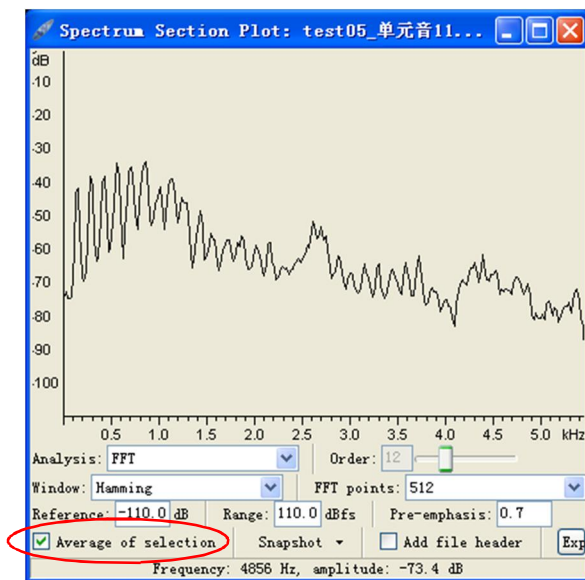


图 12-12: LTAS 分析示例

## 13. 构建自己的分析平台

WaveSurfer 软件的一大优点是读者可以根据自身需要构建自己的分析平台，软件自带的“Speech Analysis”分析模式（如图 13-1）即是一个同时分析共振峰和基频数据的例子，显示效果如图 13-2 中所示。读者完全可以结合上文中对波形、强度、语图、共振峰、基频、标注、时间等显示栏的介绍，在熟练掌握右键中“Create Pane”的功能和“Properties”属性设置的功能的基础上，根据分析需要和个人喜好设置自己的分析平台。

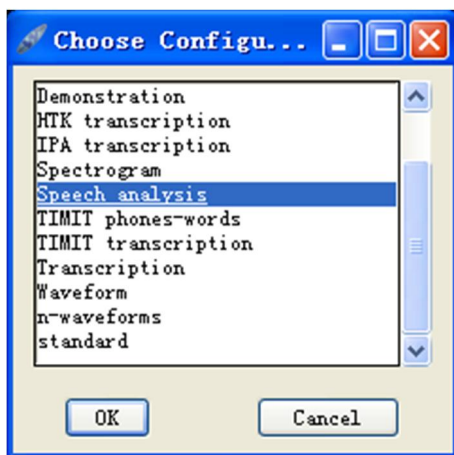


图 13-1: 言语分析配置模式

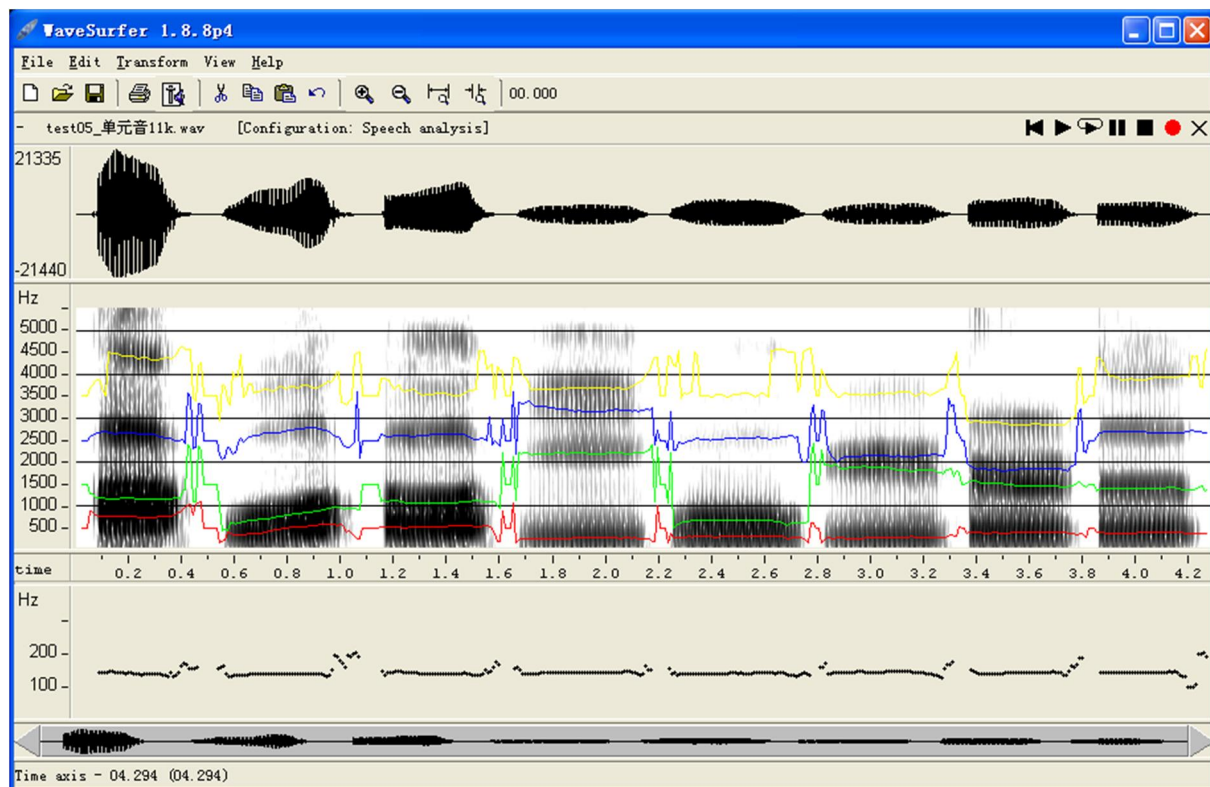


图 13-2: 言语分析模式显示效果

## 14. 有用链接/文献

- [1] <http://www.speech.kth.se/WaveSurfer/>
- [2] <http://www.speech.kth.se/WaveSurfer/man.html> 使用说明简介
- [3] <http://www.speech.kth.se/snack/> snack 软件包说明及下载
- [4] <http://sourceforge.net/projects/WaveSurfer/> 最新版本的下载地址
- [5] <http://www.indiana.edu/~phonlab/manual/wsman157/wsman00.htm> 印地安那大学伯明顿分校语音学课程对 WaveSurfer 的介绍 (2004)
- [6] <http://www.f.waseda.jp/kikuchi/tips/WaveSurfer.html> 早稻田大学菊池英明对 WaveSurfer 的介绍
- [7] <http://cosmos22.cocolog-nifty.com/blog/files/WaveSurfer.pdf> 比较详细的 WaveSurfer 日文说明书
- [8] <http://www.linguistics.ucla.edu/people/hayes/103/WaveSurferHints/> UCLA Bruce Hayes 对 WaveSurfer 的介绍
- [9] <http://www.spectrogramsforspeech.com/tutorials-2/software-download-2/>
- [10] Sjölander K, Beskow J. WaveSurfer—an open source speech tool. In: Yuan B, Huang T, Tang X, editors. Proceedings of ICSLP 2000, 6th Intl Conf on Spoken Language Processing. Beijing: 2000. pp. 464–467.
- [11] Harrison, P.T. (2004). Variability of Formant Measurements. MA Thesis: University of York.
- [12] Catalina (2007). Catalina Forensic Audio Toolbox User’s Manual Version 3.0h  
<http://www.forensicav.ro/download/CatalinaManual3h.pdf>
- [13] Burris, C., Vorperian, H.K., Fourakis, M., et.al., Acoustic analysis software: A quantitative and qualitative comparison of four systems.  
<http://www.waisman.wisc.edu/vocal/ASHA%20Poster-16-Nov-2011-FINAL.pdf>