

音频

制作与编辑

PRELUDE ON DIGITAL
AUDIO PRODUCTION



安栋 杨杰 编著

BY AN DONG & YANG JIE



上海音乐学院出版社

上海市重点学科建设项目资助,项目编号:Y0701

Supported by Shanghai Leading Academic Discipline, Project Number: Y0701

音频制作与编辑

Prelude on Digital Audio Production

安栋 杨杰 编著

By An Dong & Yang Jie

上海音乐学院出版社

图书在版编目(CIP)数据

音频制作与编辑/安栋,杨杰编著.

—上海:上海音乐学院出版社,2011.5

ISBN 978-7-80692-624-6

I. ①音… II. ①安… ②杨… III. ①音乐制作
IV. ①J619.1

中国版本图书馆 CIP 数据核字(2011)第 060068 号

书 名: 音频制作与编辑

编 著: 安 栋 杨 杰

责任编辑: 鲍 晟

封面设计: 梁业礼

出版发行: 上海音乐学院出版社

地 址: 上海市汾阳路 20 号

印 刷: 江苏省南通印刷总厂有限公司

开 本: 718×1000 1/16

字 数: 246 千字

印 张: 21.25

版 次: 2011 年 5 月第 1 版 2011 年 5 月第 1 次印刷

印 数: 1-1,500 册

书 号: ISBN 978-7-80692-624-6/J.598

定 价: 48.00 元

本社图书可通过中国音乐学网站 <http://musicology.cn> 购买

CONTENTS | 目 录

第一章 音频编辑与处理	1
1.1 频谱与分析	1
1.1.1 频谱的基本构成	1
1.1.2 静态频谱分析	3
1.1.3 动态频谱分析	6
1.1.4 相位镜(Phase Scope)	9
1.2 噪声处理	13
1.2.1 基本手段	13
1.2.2 专用滤噪工具	14
1.3 音频切片处理及剪辑技巧	22
1.3.1 Sound Forge 的切片	25
1.3.2 ReCycle 的切片	32
1.3.3 音序软件中的切片	38
1.4 样本的编辑与创造	44
第二章 混音与效果器应用	56
2.1 压缩与动态处理	56
2.1.1 动态	56
2.1.2 压缩	58
2.1.3 频率均衡	65
2.2 常见效果器	71
2.2.1 混响(Reverb)	71

2.2.2 合唱(Chorus)与镶边(Flanger)	82
2.3 语音编码效果器(Vocoder)	88
2.4 混音基本技巧	96
2.4.1 前期准备	96
2.4.2 平衡与效果	99
第三章 软件工具集纳	106
3.1 母带处理	106
3.1.1 母带处理基本目的	106
3.1.2 母带处理基本工具	107
3.1.3 母带制作工具	110
3.2 采样技术与软件采样器	112
3.2.1 采样技术基本原理	112
3.2.2 主要采样器软件	115
3.3 格式转换	130
3.3.1 采样数据格式转换	130
3.3.2 常用音频数据格式转换	136
3.4 虚拟 MIDI 驱动	142
第四章 音序软件/工作平台	148
4.1 音序平台	148
4.2 工作站软件	162
4.2.1 Propellerhead Reason	162
4.2.2 Ableton Live	165
4.2.3 Image Line Fruity Loops	168
4.3 乐谱软件	171
第五章 个人音乐工作室	179
5.1 如何构架小型工作室	179
5.1.1 制定计划	179
5.1.2 控制预算	184

5.1.3 建立工作流程	186
5.2 移动工作站	190
5.2.1 选择硬件	190
5.2.2 选择软件	193
5.3 软插件的运用	194
5.3.1 选择效果器插件组	194
5.3.2 选择软件乐器组	200
5.4 工作站的网络化	209
5.4.1 Steinberg System Link	209
5.4.2 FX-Max 的 FX Teleport	213
5.4.3 传统连接方式	215
5.4.4 其他网络化制作技术	215
第六章 影视与多媒体音乐制作	218
6.1 广告音乐	218
6.1.1 沟通	218
6.1.2 前期制作	220
6.1.3 音乐制作	223
6.2 游戏配乐	226
6.2.1 计划先行	227
6.2.2 合作与沟通	228
6.2.3 充分深入地了解一个游戏产品	230
6.2.4 游戏音乐的后期处理	233
6.3 人声录音技巧	234
6.3.1 选择合适的前级	234
6.3.2 歌手的问题	237
6.3.3 话筒的选择与放置	238
6.4 电子乐器与打击乐器	242
6.4.1 合成器与音源	242
6.4.2 鼓组与打击乐器	246

第七章 电脑音乐人性化处理	249
7.1 人性化的管弦乐制作	249
7.1.1 前言	249
7.1.2 基本设置	251
7.1.3 编辑与功能	252
7.1.4 Legato 模式	254
7.1.5 MIDI 功能键	255
7.1.6 Alternation 模式	257
7.1.7 The Matrix	259
7.1.8 复音(Polyphony)功能的开和关	262
7.1.9 连贯模式(Chain Mode)	262
7.1.10 从音序软件到 Giga Studio 3 的连接	263
7.2 软、硬件互动的实现	268
7.2.1 外部控制器	268
7.2.2 控制器与 MIDI	280
7.2.3 Reason 3 外部控制器功能	285
7.2.4 Rewire 模式下外部控制器的连接	290
第八章 素材提炼与特质萃取	295
8.1 使用素材	295
8.1.1 多样化形式	295
8.1.2 综合使用工具	299
8.1.3 特质萃取	311
8.2 采集素材	320
8.2.1 商业素材	320
8.2.2 录制素材	323
8.2.3 提取素材	328
附录 本书涉及的主要软件产品开发商网站	331

第一章 音频编辑与处理

1.1 频谱与分析

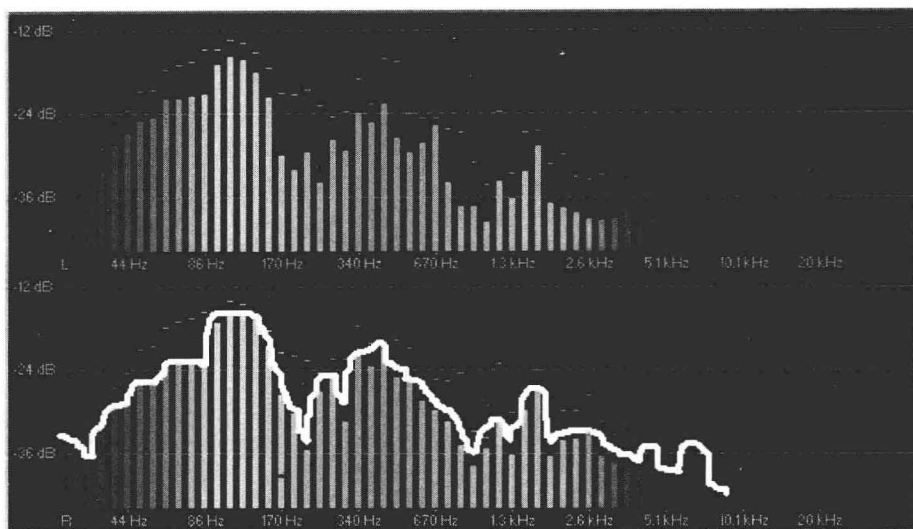
1.1.1 频谱的基本构成

任何一个自然产生的声音都可以被“分解”为一个基波和它的谐波,因此我们能够通过频谱来“勾画”出它的物理特征,通常这一过程采用的是中学物理中提到的快速傅利叶变换(FFT)的计算方法。快速傅利叶变换所反映出的声音波形实际上就是频率和振幅的对应分布,但是它比较复杂,以至我们难以用它来观察一个声音全局的情况,例如音质的变化,所以我们需要一种能够反映更多声音本质特征的描绘方式,它能够迅速地帮助我们了解一个声音在某一刻的频率与能量分布情况,或者一段声音整体的频率与能量分布情况,这就是频谱。

频谱是声音物理特征的一种图形化表现,它的基本组成是频率(频段)和它对应的振幅,它反映了一个声音中不同频率或者频率段上振幅或强度的变化情况,并且从视觉上可以用不同色彩来表示这种变化,这一原理借鉴于光谱分析。例如白噪声的特点是各频段的能量均匀,近似太阳光谱;粉红噪声是在低频段强在高频段弱的噪声,近似偏红的光谱。因此绝大部分我们经常接触的频谱有两根轴线:

横轴是频率轴,表示不同频率或者频段,一般从左向右频率逐渐升高。

纵轴是振幅轴(通常为常用的 dB 电平单位),也有采用响度或者其他方式来表示其能量。



上图是一种最常见的频谱,它反映的是一段立体声某一瞬间的左右声道频谱情况。可以看到,不同频率段上由于不同的电平,组成了很多不规则的竖柱形,这种结构称为离散谱。每个竖柱的顶端如果将它们连起来,就可以组成一条轮廓包络线(如上图右声道的白色轮廓勾勒线),这称为连续谱。在这条连续谱上,鼓起的峰状部分被称为共振峰,因此共振峰就表示这段频率的能量比较强。共振峰的实质就是:在这一段位置上声音的共振(谐波)能量达到一个高点。对于某一个共振峰来说,它有一个中心频率(这是人为的分析方式),它表达了这个共振峰的频率位置,这一频率称为共振峰频率。共振峰最顶部的数值代表这个共振峰的峰值,通常我们近似地使用这个值来描述这个共振峰的特征。共振峰越尖,就说明峰顶两侧的频率振幅下降越快,这个共振峰的特征就越突出;相反,如果共振峰越平,峰顶两侧的频率振幅下降就越慢,这个共振峰的特征则越不突出。以此我们就可以简单地从视觉上看出这段声音的基本变化特征。一个共振峰所涉及到的频率范围称为共振峰带宽,它通常采用在共振峰峰顶下方 3dB 处水平线位置截取连续谱来获得。

我们已经知道,任何一个声音的不同特征是由于这个声音中不同频率的能量分布不同而导致的,这种特征就是我们通常所说的“音质”,即声音的本质情况。因此,频谱的重要性在于它能够让我们看到声音的物理特征。由于人的听

觉系统实际上并非绝对客观,哪怕是一个受过长期专业训练的人也很难对一个声音的特质有准确客观的判断,由于听音环境的变化,主观听觉往往不能进行客观准确的判断。所以频谱的应用就显得非常重要,它能帮助我们对声音的频率和强度分布作出比较客观而准确的判断,特别是对于后期混音、编辑和母带处理,因为这两个制作阶段需要对声音的变化有尽可能准确的观察才能把握和控制。

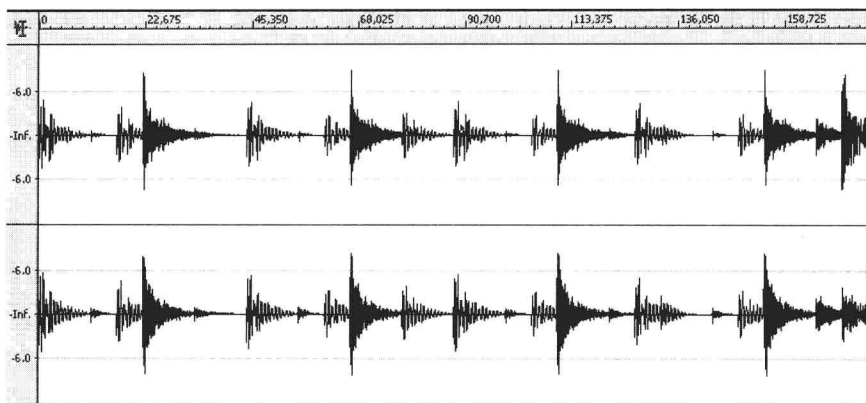
截取频率进行的强度分析方式基本有两种,一种是采用快速傅利叶变换来获得最为完整的 20Hz 到 20kHz 范围内的每个频率上的强度变化,这是一种无级连续分析。另一种方式是采用将频率划分为数个小段,然后对它们每一段的强度进行分析,通常有 4 段、8 段、16 段、32 段,最高一般不超过 128 段。实际上,目前大部分动态频谱分析软件都没有采用完全的快速傅利叶变换计算作为分析,而是采用比分段法更多的频率段来进行分析,例如 256 段或者 512 段。

目前主要使用各种频谱分析软件来获得不同的频谱,有时候甚至会同时需要多种不同频谱来进行综合观察和判断。很多专业软件自己都带有频谱分析工具,也可以根据需要使用另外的频谱分析工具或插件。从应用方式上说,这类软件基本可以分为静态和动态两种。需要注意的是,不同的软件商所开发的频谱分析软件哪怕类型相同,在精度和其他相关方面也有很大的差别,以适合不同的应用要求。

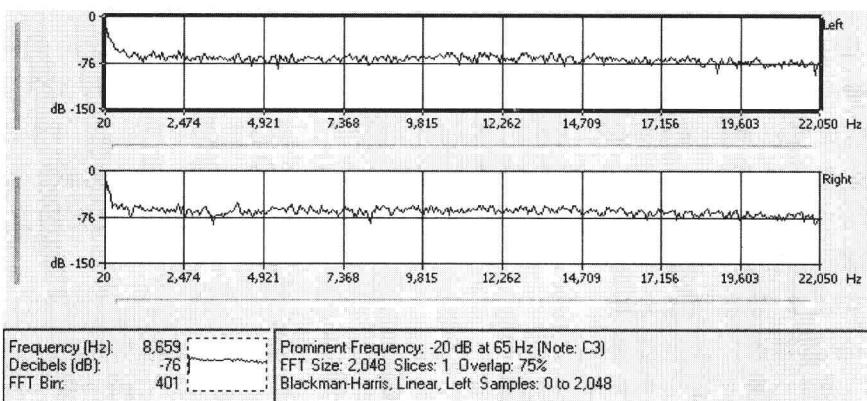
1.1.2 静态频谱分析

静态频谱分析主要是通过统计来获得一段声音整体的频率与其强度的关系。这里的强度通常是用平均峰值(Peak)或者平均共振峰值,不同的软件工具采用不同方式。静态频谱分析实际上是静态音频分析的一种,静态音频分析采用统计方式来获得一段声音的峰值、RMS 响度、主频等基本物理指数,静态频谱分析则采用的图形化的方式将得到的分布情况表现出来,而另一种则直接采取统计表格方式来罗列出这些数字。例如在较早版本的 Sound Forge 4.5 (Sonic Foundry 出品)中,一段立体声的声音采用静态频谱分析统计得到这样的结果:

波形显示:

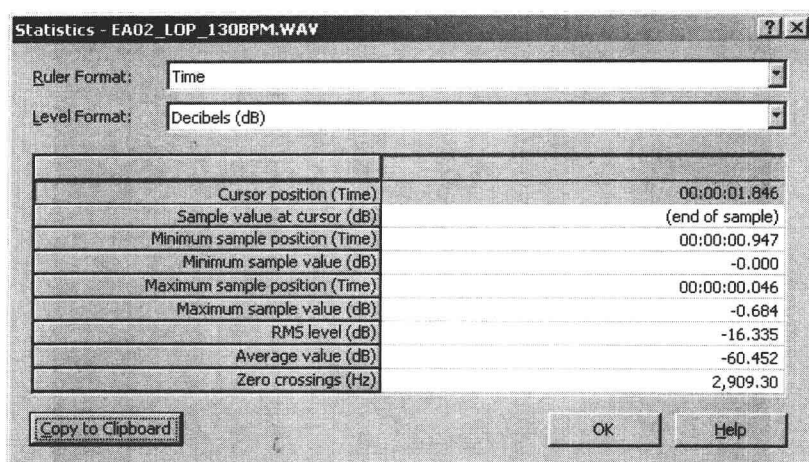


静态频谱分析统计显示:



这种静态频谱分析显示目前已经不多见,它实际就是采用了快速傅利叶变换来进行频谱分析显示的。横轴为频率轴,纵轴为 dB 电平轴。它描述了在这段声音中每一频率的平均峰值情况,并将这些值连接起来行程一条连续谱。

目前大部分专业软件中比较“流行”的是直接采用统计表格,例如 Sound Forge 8.0 中的统计表格(见下图),这种统计是比较灵活的,从下图中我可以看到标尺、电平格式等两个主要单位是允许用户根据需要进行不同的单位与格式:

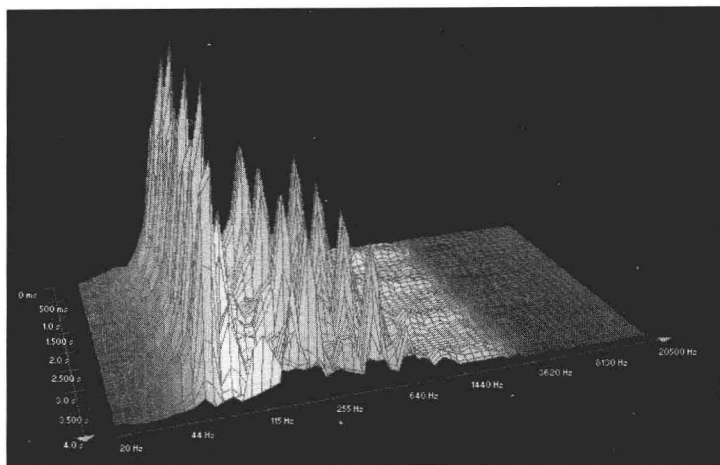


这些主要参数的中文对应含义如下：

光标所在位置	Cursor position
光标所在位置的波形电平	Sample value at cursor
波形段的最低电平值	Minimum Sample position
波形段的最低电平值所处的时间位置	Minimum Sample value
波形段的最高电平值	Maximum Sample position
波形段的最高电平值所处的时间位置	Maximum Sample value
RMS 值	RMS power
平均直流电偏移量	Average value(CD offset)
每秒时间内 0 交叉点所在频率	Zero crossings

静态的频谱分析与波形分析所展示给我们的信息虽然不同,但是两者都能够给出这段声音的基本情况,例如,哪个频率部分最强、整体的 RMS 值大约是多少、最大峰值是多少、最大峰值出现在哪个频率等问题。这些统计结果是做音频编辑与合成、后期处理等工作的重要依据。

静态频谱分析的另一种常用手段,也是最直观的手段是采用 3D 图形绘制技术来描绘频率—电平—时间三者的关系。下图就是 Steinberg Wave lab 提供的一种 3D 静态频谱分析:



在这个 3D 频谱显示中,纵轴表示电平;水平的两根轴分别是频率轴和时间轴。这样的类似“山地模型”的频谱能够直观地显示出在某一时刻频率和它的电平之间的关系。并且 Steinberg Wave lab 这个软件允许用户调整观察角度和色彩关系。

1.1.3 动态频谱分析

对于大部分专业使用者来说,动态频谱分析是最常用的,也是最直观的。所谓的“动态频谱分析”,就是实时地将频率与其电平变化采用图形化手段表现出来。它的表现方式比较多,最常用的一种是采用快速傅利叶变换来实时地描绘,另一种是采用分段电平表方式。以下两图是使用 Steinberg Wave lab 进行动态频谱分析时在同一时刻得到的:

图 A

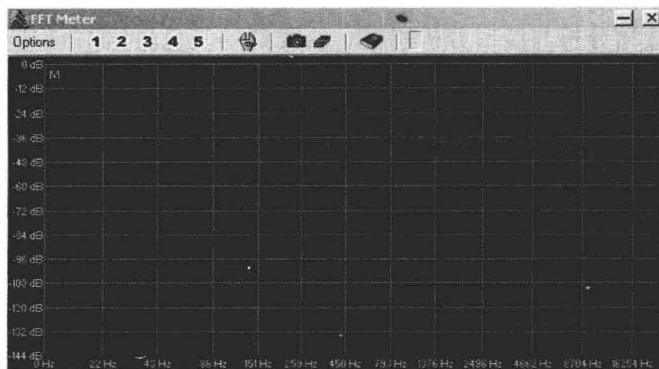


图 B

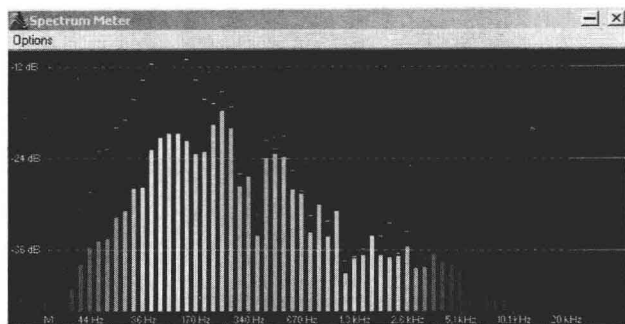
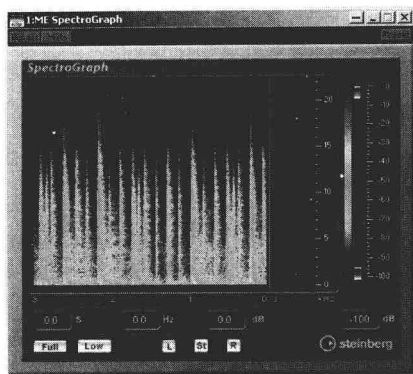


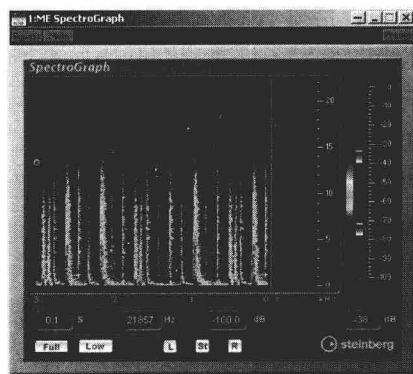
图 A 和图 B 的横轴都是频率轴,纵轴都是 dB 电平轴。两个图的另一个共同点是,它们都采用了峰值滞留来显示。峰值滞留就是说这段声音在某一频率(段)上已经“经历过的”最大峰值,以此我们可以方便地知晓当前这一瞬间的频率峰值状态与已经出现的同一频率上峰值之间的相对关系。图 A 中采用了浅色为连续谱来表示已经出现的最大峰值情况,而图 B 的分段电平图中则采用每一频率电平“柱”上方滞留的点来显示。在图 B 中我们还可以看到,频率越低,色彩越偏向红色,频率越高则越偏向蓝色,这种色彩与频率的对应关系是最常用的标准方式。虽然多数情况下都会采用这样的色系来表现频率关系,但是大部分专业软件依然提供了不同色系选择设置,以适合不同情况的需要和不同的使用习惯。例如要特别关注某一频率的电平变化,那么可以对这一频率(段)设置一个特别亮的色彩,而其他频率(段)则设置较暗的色彩。

在以上两图中,通过对比不难发现一些细节上的不同(尽管它们是同一样本在同一时刻截取的),例如在图 B 中 10kHz 以上显示电平基本消失了,而此时图 A 中 10kHz 以上则可以清晰看到一条呈快速下降(随着频率的升高)的不规则曲线。这就是由于图 A 采用了快速傅利叶变换做精确的“无分段计算”,而图 B 则是采用了多段分析。在实际处理中,或许图 A 给我们的结果又过于细致了,而图 B 则可以很快让我们了解一个基本概况或者说最明显的一些特征部位,我们可以以此为参考做处理。由此可见,在做精确处理的时候,往往一种频谱图形显示是不够的,通过不同的显示与对比可以迅速地对这个声音的表现状态作出更准确地判断。

动态频谱分析中还有一种最为常见的,也是令初学者迷惑的频谱显示方式,如下图:



这种动态频谱分析能够完整地表现出声音在随时间变化的同时,频率与其能量之间的状态。这种显示方式的横轴为时间轴,它能够随时间推移向右滚动,因此我们能够实时地看到一段声音的所有状态。上图中使用的是 Steinberg 的一个图形频谱插件,它的显示框内能够同时容纳 3 秒钟的内容,右端的 0 秒处为当前发生的情况。纵轴为频率轴,位置越高,则表示频率越高。而在显示框的右边则采用快速傅利叶变换作横向的描绘(纵轴最下端为 0Hz、最高处为 20kHz;横轴则表示电平)。这种分析显示的最大优点在于,它使用了色彩来同时描绘频率和电平的分部,如此我们就可以对一段声音进行时间上的前后对比。通常,色彩越红则表示电平越高,色彩越蓝则电平越低。如此我们就可以通过色彩来同时判断频率和它的电平。通过显示框内色彩的变化,我们就能很快知道声音在不同时刻频率与其能量之间的关系,通过前后对比我们就能实时地进行对比观察。通过调节色彩的对比度,我们就能观察更细节的声音能量随时间变化的情况,如图:



如果需要更为具体的数据,可以将鼠标移到这一频谱分析器的显示框内,就能根据鼠标当前所处的位置(在显示框下方)标示出:鼠标所处的样本时间位置、鼠标所处位置的频率、鼠标所处位置的电平。

1.1.4 相位镜(Phase Scope)

人们对声音状态感觉和判断的另一个重要依据是声音位置。而一个声音的位置主要有两个因素在客观上影响了我们双耳的判断:一是声音在左右耳的能量传达;而是这个声音的相位(Phase)。如果一个声音感觉在右边,那么右耳接收到的声音能量就比左耳要大。如果两个完全相同的声音同时向左、右耳等距离传送,那么我们会感觉只有一个声音在头部的正前方出现(即两个相等的单声道感觉);但如果这两个完全相同的声音以 45 度相位差(即相差半个频率周期的时间)向双耳同时、等距传送,那么我们会感觉一个宽阔的立体声存在。听觉的这类现象对人耳来说是非常重要的,然而仅仅采用以上提到的任何一种方式都难以表达出声音的这类特征。这时候我们就需要一种特殊的音频图形显示方式,通常称为相位镜(Phase Scope)。这种分析显示一般会要求提供立体声样本。

下图是两种常见的相位镜,它们都是实时动态显示的:

图 A

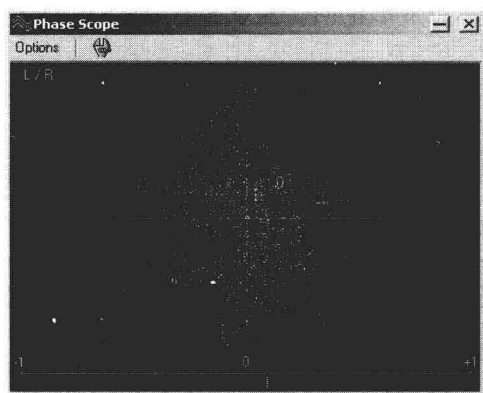
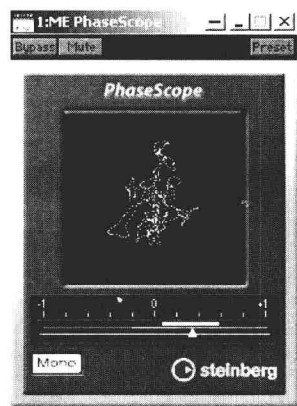


图 B



这两种相位镜的主要界面内容几乎一样,图 B 中的相位显示镜上有 L、M 和 R 三个标记,分别表示左声道(L)、单声道(M)和右声道(R)。虽然图 A 中没有这三个标记,但这并不影响它的显示和使用。

相位镜并不是一种非常容易看懂的显示方式,它是需要一定经验才能依此作出准确判断的,因为相位并不是一个非常直观的概念,它总是在“事实的背

后”影响着声音的一些真实表现,有时候甚至令人难以察觉。因此这种工具主要被后期工程师、混音师与母带工程师使用。需要注意的是,这里的相位是 Phase 而不是声相 Pan,通常我们把 Pan 称为“相位”实际并没有严重错误,只是不够确切而已,因此在音频理论上把 Phase 称为“相位”,把 Pan 称为“声相”。

相位镜所表达的是立体声的左右声道相位和振幅之间的关系,因此我们可以从这上面非常容易看到立体声的声音在左右声道位置上的能量偏移情况,以便于我们控制声音的左右位置平衡。相位镜里的云团状显示点有以下基本规律:

(1) 如果左右声道的内容完全相同,那么相位镜里在中央位置会呈现出一条垂直线,它表示了一个“完美”的听觉上是单声道的信号。这和我们的听觉习惯是完全吻合的。

(2) 如果左右声道的相位正好完全相反,那么相位镜里在中心位置会呈现出一条水平线。这时候我们感受到的是一个比较“舒适”的立体声感觉(感觉不宽也不窄,但是充斥了双耳)。

(3) 如果相位镜里出现的云团状显示点外形一直随机动态变化但是始终保持相对稳定的对称性,那么说明这条立体声信号左右平衡比较良好。

(4) 如果相位镜里出现的云团状显示点外形偏向于左或者右,那么说明这段立体声信号能力向左或向右偏移了。

(5) 如果相位镜里出现的云团状显示点外形变成了一条向左或向右 45 度角倾斜的直线,那么说明这条立体声信号中右声道或者左声道(对应)振幅为 0 或者被弱去(Mute)。

