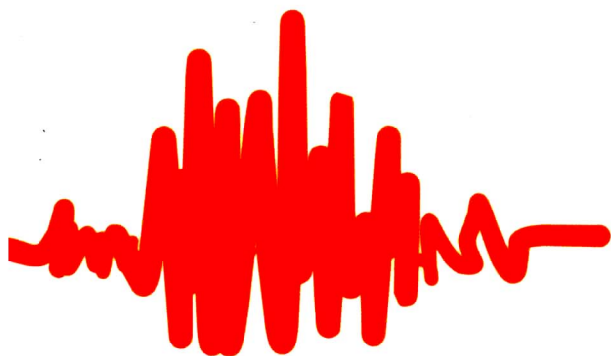




现代传播·广播电视传播 MODERN COMMUNICATION

丛书主编 王文科 陈少波

Digital Audio Production



数字音频 制作技艺

周忠成 王春明 编著

《数字音频制作技艺》一书强调理论知识和实践技能的有机统一。书中的内容包含两大部分：其一详述了数字音频的基本知识、基本概念和基本技能。同时，作者力求囊括行业发展的先进理念、工艺流程和处理技术。其二，从实战的角度介绍了各种数字音频制作平台的制作技巧。

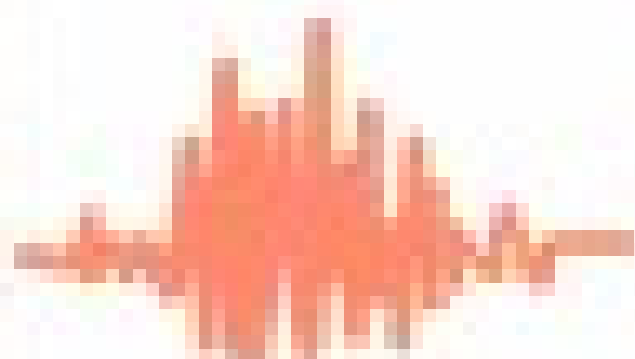
书中采用了作者实际工作中的制作实例，从实战的角度一步一步地分解数字音频的制作工艺和技巧运用，由浅入深，直观易懂。数字音频制作平台包括：数字一体机VS-890、COOLEDT PRO、CAKEWALK、AMPLITUDE 2496、PRO TOOLS、NUENDO等，基本涵盖了数字音频制作中普通应用和高级应用的整个层面。



ZHEJIANG UNIVERSITY PRESS

浙江大學出版社

Chapter 10: The Role of the Teacher



Chapter 10: The Role of the Teacher

Chapter 10: The Role of the Teacher

Chapter 10: The Role of the Teacher





Digital Audio Production



数字音频 制作技艺

周忠成 王春明 编著



ZHEJIANG UNIVERSITY PRESS

浙江大学出版社

图书在版编目(CIP)数据

数字音频制作技艺/周忠成, 王春明编著. —杭州: 浙江大学出版社, 2008. 6

ISBN 978-7-308-05978-7

I. 数… II. ①周…②王… III. 音频设备—数字技术
IV. TN912.2

中国版本图书馆 CIP 数据核字(2008)第 073489 号

数字音频制作技艺

周忠成 王春明 编著

丛书策划 李海燕

责任编辑 李玲如

封面设计 俞亚彤

出版发行 浙江大学出版社

(杭州天目山路 148 号 邮政编码 310028)

(E-mail: zupress@mail. hz. zj. cn)

(网址: <http://www.zjupress.com>

<http://www.press.zju.edu.cn>)

电话: 0571-88925592, 88273066(传真)

排 版 浙江大学出版社电脑排版中心

印 刷 杭州浙大同心教育彩印有限公司

开 本 787mm×960mm 1/16

印 张 22.75

字 数 421 千

版 印 次 2008 年 6 月第 1 版 2008 年 6 月第 1 次印刷

书 号 ISBN 978-7-308-05978-7

定 价 38.00 元

版权所有 翻印必究 印装差错 负责调换

浙江大学出版社发行部邮购电话 (0571)88072522

目 录

CONTENTS

第一章 数字音频基础	001
1.1 模拟音频和数字音频	002
1.2 数字音频的种类	004
1.3 WAV 音频	008
1.4 MIDI 音频	035
1.5 3D 音频技术.....	041
1.6 音频信号的特点	042
1.7 计算机语音识别技术	043
思考与练习.....	048
第二章 数字音频的压缩编码标准	050
2.1 多媒体数据压缩的基本原理	050
2.2 音频信息编码基础	056
2.3 音频信息的压缩编码算法	060
2.4 音频信息压缩编码标准及质量评估	064
思考与练习.....	074
第三章 一体化专用数字音频工作站	075
3.1 数字音频处理设备	075
3.2 声音节目制作系统	076
3.3 VS-890 数字音频工作站总体概述.....	078
3.4 VS-890 的面板、接口和显示	083
3.5 VS-890 的系统设置.....	088
3.6 VS-890 的基本操作技巧	090
3.7 多轨录音	095



3.8 音轨编辑	101
3.9 缩混输出	109
思考与练习	112
第四章 电脑音乐制作技艺	113
4.1 MIDI 技术概述	114
4.2 Cakewalk 9.0 的主要功能介绍	116
4.3 Cakewalk 9.0 的主要工作窗口	121
4.4 播放 MIDI 样本文件	130
4.5 用五线谱写入法输入音乐的主旋律	133
4.6 录制和编辑由 CD 播放的音频节目	141
4.7 综合制作实例	147
4.8 制作 MIDI 效果	166
思考与练习	170
第五章 数字音频剪辑技艺	171
5.1 数字音频技术概述	171
5.2 音频编辑软件 CoolEdit Pro2.0 的简单介绍	172
5.3 音频素材的播放及基本操作	177
5.4 基本的音频编辑功能	184
5.5 从视频素材中捕捉音频素材	193
5.6 基本音效的使用	194
4.8 综合制作示例	206
思考与练习	216
第六章 Samplitude2496	217
6.1 Samplitude 概述	217
6.2 Samplitude 的基本操作	218
6.3 Samplitude 的编辑功能	223
6.4 Samplitude 的声音处理	229
6.5 效果器的使用	234
思考与练习	251



第七章 Pro Tools 数字音频工作站	252
7.1 Pro Tools 系统集成	253
7.2 Pro Tools 系统应用过程中涉及的几个专有名词	260
7.3 Pro Tools 系统的工作窗口	264
7.4 Pro Tools 系统主要的菜单命令及其运用	278
7.5 Pro Tools 系统基本功能的运用	283
7.6 综合实例——电视剧的后期音频编辑	308
思考与练习	317
第八章 NUENDO 应用技巧	318
8.1 概 述	318
8.2 NUENDO 的设置	318
8.3 音频录制	327
8.4 音频编辑与音频处理技巧	333
8.5 样本编辑器	342
8.6 音频效果器	348
思考与练习	357



第一章 数字音频基础



本章要点

- 模拟音频与数字音频。
- 数字音频采样、量化的基本原理；数字音频的文件格式和音频信号的特点。
- 音频卡的工作原理以及音频卡的功能和分类。
- 计算机的语言输出技术、语音合成技术及语音识别系统的类型。

声音属于听觉类媒体,是人们用来传递、交流信息最方便、最直接的方式之一。音频一般指的是频率范围在 20Hz~20kHz 之间的声音,但实际上“音频”通常被作为“音频信号”或“声音”的同义词。音频可以分电话语音、宽带音频、调频广播和高保真音频四种。随着多媒体信息处理技术的快速发展,计算机的数据处理能力不断增强,因而以计算机为平台的音频处理技术也越来越受到重视,并得到了广泛的应用。如视频图像的配音、配乐,静态图像的解说、背景音乐,可视电话,电视会议中的语音,游戏中的音响效果;虚拟现实中的声音模拟,用声音控制 Web,电子读物的有声输出等等,都离不开数字音频及数字音频的处理技巧。



1.1 模拟音频和数字音频

1.1.1 模拟音频

声音是一种波,是机械振动在介质(空气)中传播形成的。机械振动越强,声音越大。模拟音频指的是通过拾音器(话筒)将机械振动信号转换成在时间上连续变化的电信号,电信号以模拟电压的幅度来表示声音的强弱,以模拟电压变化的频率来表示声音声调的高低。模拟音频记录主要采用传统的磁记录方式。模拟音频信号经功放大后送到扬声器,通过扬声器中的电磁体将电信号还原成机械振动,发出人耳能听到的声音。目前,模拟录音的动态范围可达 80dB。但模拟磁性录音性能受电磁性能的影响较大,磁带频率特性的微小变化也会对声音质量的优劣产生较大的影响。若想进一步提高录音、放音的音质,并对声音进行编辑处理,则必须采用数字音频技术。

1.1.2 数字音频

数字音频指的是用一系列不连续的二进制数字来表示声音信号的音频,准确地说,数字音频分为两大类:一类是纯数字音频,也可以称为原创数字音频,是数字音频制作的前期产物,音频的基本单元是数字音符,核心是音源;另一类是数字化音频,是数字音频制作中后期处理的基本音频形式。数字音频具有保真度好、动态范围大的特点。以计算机为平台的音频必须是以数字形式存在的,因此,如果要用计算机来记录、存储、处理和传输模拟音频信号,则必须先把模拟音频信号转换成用有限个二进制数字表示的离散序列,即实现模拟音频的数字化。

在数字音频技术中,把表示声音强弱的模拟电压用数字来表示。如 0.5V 电压用数字 20 表示,2V 电压用 80 表示。但是模拟电压幅度即使在某个电平范围内,仍然可以有无穷多个电压值,如 1.2V,1.21V,1.215V,⋯在用数字来表示音频幅度时,只能把无穷多个电压幅度值用有限个数字表示出来,即把某一幅度范围内的电压用一个数字来表示。这种处理称为量化,因此,很显然量化是有误差的。模拟音频的数字化我们在后面的章节中介绍。



数字音频按质量可以分为如表 1.1 所示的几个等级。

表 1.1 数字音频的质量等级

信号类型	频率范围(Hz)	采样率(kHz)	量化精度(位)
电话语音	200~3400	8	8
宽带音频	50~7000	16	16
调频广播	20~15k	37.8	16
高质量音频	20~20k	44.1	16

1.1.3 模拟音频和数字音频的关系

需要首先说明的是,这里谈的模拟音频和数字音频的关系指的是模拟音频和数字化音频的关系。模拟音频在时间上是连续的,而以数字表示的音频则是一个数据序列,在时间上只能是断续离散的。因此,当把模拟音频转变成数字音频时,需要每隔一段时间间隔在模拟声音波形上取一个幅度值。这个过程称为采样,实现这个步骤使用的设备是模/数转换器(A/D Converter)。以采样频率 44.1kHz 为例,转换器以每秒 44100 次的速率对声波进行采样,每一次采样都记录下原始模拟声波在某一时刻的样本值,将这一串样本值连接起来,同样可以真实地描述这一段模拟声波。每一秒钟采样的次数称为采样频率或采样率,单位为 Hz(赫兹)。采样频率越高则所能描述的声波频率也越高,得到的数字音频的高保真度也越高,即越能真实地反映原先的模拟音频的声音品质。

总之,数字化音频是一个离散的二进制数据序列,是模拟音频经抽样、量化后得到的。可以通过计算机系统存储、处理、传输的各种音频,光盘记录的 Audio CD 和数字磁带(DAT)中存储的音频信息都是数字音频。模/数转换器(A/D Converter)可以将模拟音频转化成数字音频;而数/模转换器(D/A Converter)则可以将数字音频转换成模拟音频,然后通过音响系统将声音回放出来。

需要指出的是,Audio CD 唱片虽然是数字音频产品,但根据唱片在制作工艺中的差异一般可以分为 AAD、ADD、DDD 三种。这三种分类分别代表唱片在录音、编辑、成品过程中所使用的方法是模拟(Analog)的还是数字(Digital)的。A 代表模拟,D 代表数字,AAD 就说明其录音和编辑是用模拟方式,最后灌片是用数字方式,这类唱片大多是将过去录制的音乐转换成 CD 唱片;ADD 则是有一个数字编辑和修改过程,许多古典音乐大师的演奏作品



多录制于模拟时代,我们现在听到的 CD 就是经过编辑修改后重新灌录的;而 DDD 的唱片则是现代全数字的录音制作产品。

1.2 数字音频的种类

在多媒体计算机中可以存储处理的数字音频按音频的文件格式的不同分很多种,主要有:WAV 格式、mp3 格式、MIDI 格式、RA 格式、WMA 格式、AIF 格式、VOC 格式、SNO 格式及 RMI 格式等。

1.2.1 波形音频

波形音频是多媒体计算机获得声音最直接、最简便的方式。在这种方式中,通常以麦克风、立体声录音机、CD 激光唱盘、MD 等作为声音信号的输入源,音频卡以一定的采样频率和量化位数对输入声音进行实时的模数转化,将模拟声音信号转换为数字信号(A/D 转换),然后以某种格式保存在硬盘上。声音文件重放时,音频卡将文件中的数字信号再实时的还原成模拟信号(D/A 转换),经混音器混合功放放大后由扬声器输出声音。

1. WAV 音频

WAV 波形文件是 Windows 所使用的标准数字音频文件,也是目前应用最广泛的音频格式,文件的扩展名是 .WAV。在适当的硬件及计算机控制下,使用波形文件能够重现各种声音,包括不规则的噪声、CD 音质的音乐、单声道及立体声声音。多数音频卡都能以 16 位量化级、44.1kHz 的采样率录制和重放立体声声音。

波形文件的主要缺点是文件太大,所需要的存储容量很大,每分钟 CD 音质的 WAV 音频要占用 10.09MB 的存储空间,所以不适合用于长时间的声音记录。由于原始声音数据量太大,有必要采用硬件或软件的方法对声音数据进行压缩处理。常用的软件压缩编码方法主要有 ACM(Microsoft's Audio Compression Manager)和 PCM(脉冲编码调制)等。当然,如果对声音质量要求不高,则可以通过降低采样频率和量化位数或利用单声道来录制 WAV 文件,此时的 WAV 文件大小可以大大地减小。由于人的发声频率范围在中频段,因此,使用 8 位量化级、11.025kHz 采样率就能高保真地记录和还原一般人讲话的声音信息。

通过 Windows 的对象连接与嵌入(OLE)技术,波形文件可以嵌在其他 Windows 应用系统中使用。由于波形文件记录的是数字化音频信号,因此,



可由计算机对音频信号进行编辑和处理,如零点剪辑、变调、加各种厅堂混响、多轨声音叠加及缩混输出,等等。以计算机为平台的音频编辑工具很多,常用的主要有:Cooleedit Pro、Samplitude2496、Sonar、Nurando、Pro Tools,等等。Windows 中的“录音机”(Sound Recorder)也是一个具备部分声音简单剪辑功能的小型工具。

2. CD 音频

符合 MPC2 标准的 CD-ROM 驱动器不仅可以读取 CD-ROM 光盘的信息,还能播放数字 CD 唱盘(CD-DA),这样多媒体计算机就能够利用已经非常成熟的数字音响技术来获得高质量的音频——CD 音频。CD 音频也是一种数字化音频,是采用 16 位量化级、44.1kHz 采样率的双声道 WAV 音频。对人的听觉器官而言,CD 音频可以高保真的记录和重现 20Hz~20kHz 的原始声音信息,每张 CD 唱盘能记录约 74 分钟的高保真声音信息。

在多媒体计算机上输出 CD 音频信号一般有两种途径:一种是通过 CD-ROM 驱动器前端的耳机插孔直接输出,另一种是使用音频卡由计算机的多媒体音响系统输出声音。前者的输出音质不受音频卡质量的影响,但不能使用音频卡的混声功能;而后者虽然可以与波形或 MIDI 音频进行混音输出,但音频卡的放大功率都比较小(一般不超过 4W),通常需要有源扬声器或配置外部声音放大器来获得足够的声压。

3. HDCD

HDCD 是一种先进的音响编码技术,以企业自主开发的专利技术来记录音响数据,其深度及细致程度,就如在专业录音室的水平一样,逼真而细腻。HDCD 是以 20 位的编码速度编制的 CD,比一般以 16 位编制的 CD 更能表现音响效果的层次感,声音表达不但有远近距离感,同时也令人声及纯音乐显得更自然和清澈。

4. DVD-Audio

DVD-Audio 是由 DVD Forum Audio Working Group 与 International Steering Committee 共同制定的规格,也是 DVD 家族中重要的一员。相较于 DVD-Video 压缩过的 16 位、48kHz 采样频率的音乐,DVD-Audio 的 6 声道更清脆,可达到 24 位、96kHz 的采样音乐的潜力,甚至听力不大好的人也可区分出两者的差别,但是这两种格式都远远超越了 CD 格式,即 16 位、44.1kHz 的声音。由于 DVD-Audio 所支持的格式显然比 DVD-Video 和 CD 的 PCM(Pulse Coded Modulation,脉冲编码调制)音乐有更高的品质,因此,DVD-Audio 可以表现更丰富的 3D 环场音效,其动态频率的范围比 CD 还要大 4 倍。



5. SACD

SACD 即 Super Audio CD 的缩写,是索尼和飞利浦在它们联合开发的 MMCD(单面双层结构的高密度光碟)基础上研制推出的新数字音频格式。SACD 采用了名为 DSD(Direct Stream Digital,直接数字流编码)的新编码方式,信息储存量为普通 CD 的 6 倍。SACD 以高达 2.8224MHz 的采样频率(为 CD44.1kHz 的 6 倍)把原始的模拟音频信号量化为 1bit 的数字音频信号,当还原为模拟音频信号重播时,所还原的波形与原先音乐的模拟波形几乎毫无二致,比 CD(44.1kHz/16bit)或 DVD Audio(96kHz/24bit)的波形更为完整。因此,其声音的清晰度和信噪比都很高,在 20Hz~20kHz 频率范围内的动态范围达 120dB。SACD 容量与 DVD-Audio 相同,均为 4.7GB。

6. VOC 音频

VOC 文件是 Creative 公司推出的波形音频文件格式,也是声霸卡(Sound Blaster)使用的音频文件格式(CMF 文件(Creative Music File)也是随音频卡一起诞生的文件格式),因此也可以说是一种企业标准的文件格式。每个 VOC 文件由文件头块(header block)和音频数据块(data block)组成。文件头包含一个标识、版本号和一个指向数据块起始的指针。数据块分成各种类型的子块,如声音数据、静音、标记、ASCII 码文件,重复以及终止标志、扩展块等。利用音频卡提供的软件可以实现 VOC 和 WAV 文件的转换。

1.2.2 MIDI 音频

MIDI 音频是多媒体计算机中应用比较广泛的另一种数字音频。与 WAV 音频不同,MIDI 音频可以满足对音乐长时间记录和播放的需要。

由于 MIDI 文件相当于是符号化了的乐谱,记录的不是声音本身,而是发声的控制信息,如音色、音符、音程、有无颤音和弯音轮效果、有无混响或合唱、音量等等,因此所需的存储空间很小,每分钟的立体声 MIDI 音乐,其文件大小只有 7kB 左右。MIDI 文件的扩展名为.MID。

MIDI 文件要通过音源将声音回放出来,因此,MIDI 音频不能真实重现人的说话声、自然界的某些音效以及某些民族乐器如埙的音乐旋律。此外,MIDI 的回放质量受音频卡上合成芯片音源质量的严重限制,用一般的音频卡难以产生真实的音乐演奏效果。近年国外流行的音频卡普遍采用波表法进行音乐合成,音源的专业性有所提高,从而使 MIDI 音乐的质量大大提高(效果接近 CD 音质),但波表卡比较昂贵,在我国还未完全普及。



1.2.3 mp3 音频

mp3 由德国一家研究机构推出,目前一般指在计算机上压缩和存储音乐文件的一种格式。mp3 是技术术语的缩写,也就是 MPEG-1 Audio Layer-3 的缩写。MPEG(Moving Picture Experts Group)在汉语中译为活动图像专家组,特指活动影音压缩标准,MPEG 音频文件是 MPEG1 标准中的声音部分,也叫 MPEG 音频层。根据压缩质量和编码复杂程度,MPEG 音频层被划分为三层,即 Layer-1、Layer-2、Layer-3,分别对应 mp1、mp2、mp3 这三种声音文件,并根据不同的用途,使用不同层次的编码方式。MPEG 音频编码的层次越高,编码器越复杂,压缩率也越高,mp1 和 mp2 的压缩率分别为 4:1 和 6:1~8:1,而 mp3 的压缩率则高达 10:1~12:1,也就是说,一分钟 CD 音质的音乐,未经压缩需要 10.09MB 的存储空间,而经过 mp3 压缩编码后只需要 1MB 左右的存储空间。不过,mp3 对音频信号采用的是有损压缩方式,为了降低声音失真度,mp3 采取了“感官编码技术”,即编码时先对音频文件进行频谱分析,然后用过滤器滤掉噪音电平,接着通过量化的方式将剩下的每一位打散排列,最后形成具有较高压缩比的 mp3 文件,这种压缩方式使压缩后的音频文件在回放时能够达到比较接近音源的声音回放效果。

SRSWOW 是 SRS Labs 公司针对 mp3 等新一代的音频格式而开发的一种音质增强技术。这种技术基于人耳的生理声学和心理声学原理,深入挖掘了立体声音响中的三维环绕信息,可显著扩展声场的宽度、高度和深度,使声音变得更加自然、悦耳,久听不易疲劳。而且,它还将 SRS 的 Tru Bass 低音增强技术融入其中,可以产生更加深邃、浑厚、富有弹性的低音,在用容积较小的多媒体音箱重放时所产生的低音效果几乎可以和大音箱相媲美。它所产生的虚拟声场的范围要比扬声器自身产生的声场范围大许多,相应地,最佳听音区的范围也要大许多。

mp3PRO 利用的是由 Coding Technologies 公司开发的 codes enhancement 技术和 Thomson 与 Fraunhofer 共同开发的 mp3 技术。当制作 mp3PRO 文件时,编码器将音频分为两部分:一部分是将音频数据中的低频段部分分离出来,通过传统的 mp3 技术编码得出的正常的 mp3 音频流。此举可令 mp3 编码器专注于低频段信号从而获得更好的压缩质量,而且原来的 mp3 播放器也可播放 mp3PRO 文件。另一部分则是将分离出来的高频段信号进行编码并嵌入到 mp3 流中,传统的 mp3 播放器会将其忽略掉,而新的 mp3PRO 播放器则可从中还还原出高频信号,并将两者进行组合,得到高质量的全带宽的声音。通过这样的技术,使得 mp3PRO 能在 64kbps 的编码率便



可提供与 128kbps 的 mp3 相同的质量。

1.2.4 RA 音频

RA、RAM 和 RM 都是 Real 公司成熟的网络音频格式,采用了“音频流”技术,所以非常适合网络广播。在制作时可以加入版权、演唱者、制作者、Mail 和歌曲的 Title 等信息。RA 可以称为互联网上多媒体传播的霸主,适合于网络上进行实时播放,是目前在线收听网络音乐最好的一种格式。

1.2.5 WMA 音频

微软在开发自己的网络多媒体服务平台上主推 ASF(Audio Steaming Format),这是一个开放式的并支持各种各样的网络和协议的数据传输的标准,支持音频、视频以及其他一系列的多媒体类型。而 WMA 是 Windows Media Audio 的缩写,相当于只包含音频的 ASF 文件。WMA 文件在 80kbps、44kHz 的模式下压缩比可达 18:1,而且压缩速度比 mp3 快一倍。

WMA 可以说是微软公司推出的与 mp3 格式齐名的一种新的音频格式,这种音频格式采用微软公司开发的新一代网上流式数字音频压缩技术,这种压缩技术的特点是同时兼顾了保真度和网络传输需求,所以具有一定的先进性。由于 WMA 在压缩比和音质方面都超过了 mp3,更是远胜于 RA(Real Audio),即使在较低的采样频率下也能产生较好的音质,再加上 WMA 有微软的 Windows Media Player 作其强大的后盾,所以一经推出就赢得一片喝彩。网上的许多音乐纷纷转向 WMA,许多播放器软件也纷纷开发出支持 WMA 格式的插件程序,估计用不了多长时间,WMA 就会成为网络音频的主要格式。

1.3 WAV 音频

1.3.1 WAV 音频的获得

WAV 音频是通过音频卡(声卡)的模/数转换器将模拟音频数字化以后得到的。我们知道计算机内的信息都是以二进制的数字形式记录的,因此必须把随时间连续变化的模拟音频信号转换成离散的以 0,1 组成的二进制数字形式的数字信号,即转换为数字音频信号,才能用计算机进行存储、处理和传输,这个过程称为声音信号的数字化过程。在计算机里,将模拟的电信号转换



为数字信息,是通过模数转换器 A/D Converter 完成的。相反,声音播放时,数字化声音转换成模拟信号是由数模转换器 D/A Converter 完成的。声音信号的整个模数转换过程如图 1-1 的方框图所示。

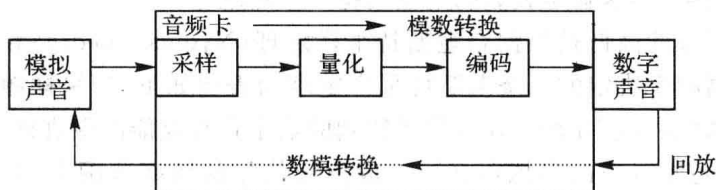


图 1-1 声音的模数转化过程

声音的数字化主要分三个步骤,即采样、量化和编码。采样就是每间隔一个时间段采集一次声音信号的幅度样本,每秒钟采集样本的次数称为采样频率。量化就是把采样得到的声音信号幅度值转换为二进制数字值。时间上的离散叫采样,幅度上的离散称为量化。编码则是按某种格式、标准和规范将采集与量化得到的与音频相关的所有二进制信息排列和组合起来,以便计算机能对这些信息进行存储、处理和传输。

1. 采样和采样频率(Sampling Rate)

采样频率是指将模拟声音波形数字化时,每秒钟所采集的声波幅度样本的次数,采样频率的计算单位是 kHz。采样频率越高声音的保真程度越高,但用于存储音频的数据量也越大。一般来讲,可以根据音频素材的种类如语言、音乐等选择合适的采样频率。

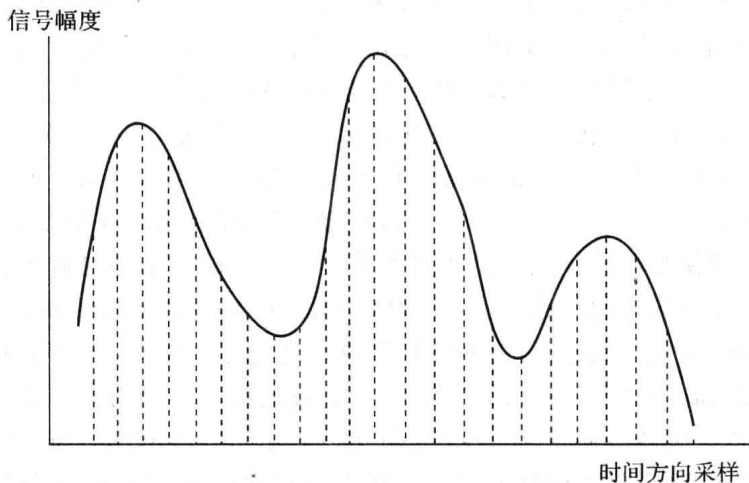


图 1-2 波形采样量化



采样过程如图 1-2 所示,横轴代表采样的间隔和次数即采样频率,纵轴代表采样点的幅度值,因此,采样得到的波形实际上是方波,也就是说,采样前后波形有失真,但是,如果采样点足够多的话,我们会发现采样前后的波形会无限接近乃至重合,这就是著名的奈奎斯特采样定理。

采样频率的高低是根据奈奎斯特采样定理(Nyquist theory)和声音信号本身的最高频率决定的。奈奎斯特采样定理的表述如下:当采样频率不低于声音信号的最高频率的两倍时,采样得到的数字音频就能高保真地(也是相对的)记录和还原原来的模拟声音。一般声音信号的频率范围大约在 20Hz~20kHz 之间,根据采样理论,如果采样频率大于 40kHz,那么数字化以后得到的数字音频就可以高保真的记录模拟音频。常用的音频采样频率有:8kHz、11.025kHz、22.05kHz、16kHz、44.1kHz、48kHz、96kHz 等,也就是说对一般的声音采样用 44.1kHz 的采样频率就可以达到高保真了,当然要根据音频对象的不同选用不同的采样频率,如制作高品质 DVD,可以采用 96kHz 的采样频率。

2. 量化、量化位数和分辨率

量化的过程是把采样得到的每一个样本值从模拟量转换为二进制的数字值的过程。量化的过程需先将整个幅度划分成有限个小幅度的集合,如果用一位来量化采样得到的幅度值,那么可以将幅度分为两个等级;如果用 2 位来量化采样得到的幅度值,那么可以将幅度分为 2^2 ,即 4 个等级;用 8 位二进制码字来表示幅度,就可以把整个幅度划分成 2^8 ,即 256 个量化等级。1,2,...,8 称为量化位数或叫量化精度。常用的量化位数有 8 位、12 位和 16 位。例如,8 位量化级表示每个采样点可以表示 256 个(0~255)不同的量化值,而 16 位量化则可表示 65536 个不同的量化值。又比如,要对全班的考试成绩进行量化,如果用 1 位来量化,那么就可以用 2^1 ,即两种状态(0,1)来表示,即可以用合格和不合格两种状态来表示学生的成绩,100 分是合格,61 分也是合格,所以,如果采用一位量化是远远无法真实反映学生的考试结果。如果采用两位量化,就有 2^2 ,即四种(00、01、10、11)状态,对应学生成绩中的优、良、及格、不及格,但即使是两位量化很显然离准确表示学生的成绩还是有差距的。依此类推,如果增加量化位数,用 100 种状态就可以相当真实的反映学生的考试成绩。因此,量化位数越高,越能体现被数字化信息的细节。对声音信息来说量化位数的高低决定了声音的动态范围,即被记录和重放的声音最高与最低之间的差值。实践表明,16 位的量化级足以表示从人耳刚刚听得见的极细微的声音到感觉难以忍受的巨大噪声这样大的声音范围。同样,量化位数越高音质越好,数据量也越大。所以,量化位数也叫分辨率。