

第3章 声音与影视的编辑

本章学习目标

- 理解数字声音、影视的基本概念及其数字化过程。
- 了解数字声音的常用编码技术以及编码标准。
- 了解合成声音基本概念与 MIDI 规范。
- 了解常用的声音、影视文件格式。
- 了解电视信号及其标准。
- 了解影视编码的国际标准 MPEG 与 H.261 标准。
- 了解流媒体技术原理及其应用领域。
- 掌握用 Audition 3.0 进行声音处理的基本方法。
- 掌握用 Premiere Pro CS4 进行影视处理的基本方法。

3.1 数字声音基础知识

声音是人们用来传递信息最方便、最熟悉的方式,它可以携带大量精细、准确的信息。早期的 PC 不能发音,多媒体技术的发展使计算机能够处理声音信息。音乐和解说可使静态图像变得更加丰富多彩,声音和影视图像的同步播放,可使影视图像更具真实性。随着多媒体信息处理技术的发展,计算机数据处理能力的增强,声音被用做输入或输出,声音处理技术得到了广泛的应用。

3.1.1 数字声音的基本概念

1. 声音信号

人们之所以能听到各种声音,是因为不同频率的声波通过空气产生振动,对人耳刺激的结果。规则声音是一种连续变化的模拟信号,可用一条连续的曲线来表示,称为声波。因声波是在时间和幅度上都连续变化的量,所以称为模拟量。

模拟声音信号有两个基本参数:频率和振幅。

1) 频率

频率是指声音每秒钟振动的次数,一个声源每秒钟可产生成百上千个波峰,每秒钟所产生的波峰数目就是声音信号的频率。声音的频率体现音调的高低,单位用赫兹(Hz)或千赫兹(kHz)表示。例如一个声波信号在一秒钟内有 5000 个波峰,则可将它的频率表示为 5000Hz 或 5kHz。人的耳朵能听到声音的频率范围约在 20Hz~20kHz。

2) 幅度

幅度是从声音信号的基线到当前波峰的距离。幅度决定了信号音量的强弱程度,幅度越大,声音越强。对于声音信号,它的强度用分贝(dB)表示。分贝的幅度就是音量。

2. 模拟声音的数字化

如果要用计算机对声音信息进行处理,则首先要通过 A/D(模/数)转换将模拟声音信号变成数字信号,实现声音信号的数字化。数字化的声音易于用计算机软件处理,现在几乎所有的专业化声音录制器、编辑器都是数字的。对模拟声音的数字化过程涉及声音的采样、量化和编码,如图 3-1 所示。

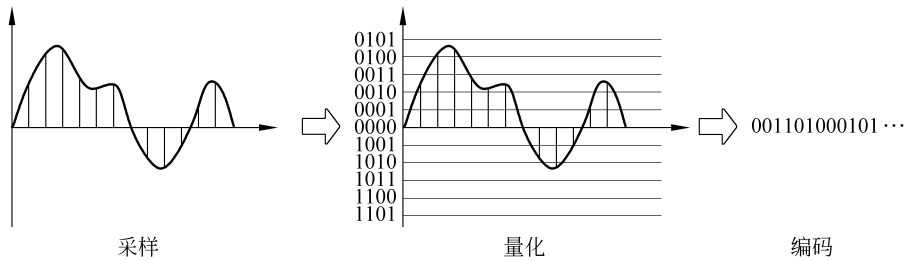


图 3-1 模拟声音信号的数字化过程

1) 采样

为实现 A/D 转换,需要把模拟声音信号波形进行分割,以转换成数字信号,这种方法称为采样(Sampling)。

采样的过程是每隔一个时间间隔在模拟声音的波形上取一个幅度值,把时间上的连续信号变成时间上的离散信号。该时间间隔称为采样周期,其倒数为采样频率。

采样频率是指计算机每秒钟采集多少个声音样本。采样频率越高,即采样的间隔时间越短,则在单位时间内计算机得到的声音样本数据就越多,对声音波形的表示也越精确。

采样频率的选择与声音信号本身的频率有关,根据奈奎斯特(Nyquist)理论,只有采样频率高于声音信号最高频率的两倍时,才能把数字信号表示的声音较好地还原为原来的声音。最常用的采样频率有: 11.025kHz、22.05kHz、44.1kHz 等。

2) 量化

采样所得到的声波上的幅度值,即某一瞬间声波幅度的电压值,影响音量的高低,该值的大小需要用某种数字化的方法来表示。通常把对声波波形幅度的数字化表示称为量化(Quantization)。

量化的过程是先将采样后的信号按整个声波的幅度划分成有限个区段的集合,把落入某个区段内的采样值归为一类,并赋予相同的量化值。采样信号的量化值采用二进制表示,表示采样信号的幅度二进制的位数称为量化位数。如果以 8b 为记录模式,则将其纵轴划分为 2^8 个量化等级,它的量化位数为 8。

在相同的采样频率之下,量化位数越高,声音的质量越好。同样,在相同量化位数的情况下,采样频率越高,声音效果也就越好。这就好比是量一个人的身高,若是以 mm 为单位来测量,会比以 cm 为单位量更加准确。

3) 编码

模拟信号经过采样和量化以后,形成一系列的离散信号——脉冲数字信号。这种脉冲数字信号可以用一定的方式进行编码,形成计算机内部运行的数据。所谓编码,就是按照一定的格式把经过采样和量化得到的离散数据记录下来,并在有效的数据中加入一些用于纠

错同步和控制的数据。在数据回放时,可以根据所记录的纠错数据判别读出的声音数据是否有错,如果有错,可加以纠正。

3.1.2 数字声音的编码技术

数字声音的编码技术分为三类:波形编码、参数编码以及混合编码。

1. 波形编码

波形编码是声音信号常用的编码方法,它直接对波形采样、量化和编码,算法简单,易于实现。而且,声音恢复时能保持原有的特点,因此被广泛应用。常用的波形编码方法有 PCM、DPCM 和 ADPCM 等。

1) PCM(Pulse Code Modulation,脉冲编码调制)

PCM 简称脉码调制,可以直接对声音信号做 A/D 转换,用一组二进制数字编码表示,得到的是未经压缩的声音数据。这是一种最常用、最简单的编码方法。

PCM 编码方法不需要复杂的信号处理技术就能实现瞬时的数据的量化和还原,而且信噪比高。在解码后恢复的声音,只要采样频率足够高,量化位数足够多,就会有很好的质量。但是,这种对声音信号直接量化的方法编码数据量很大,需要很高的传输速率。

在 MPC 中,声卡都具有 PCM 编码和解码的功能。激光唱盘(CD-DA)记录声音时就采用这种方法,存储未经压缩的数字声音信号。

PCM 编码是波形压缩编码的基础,波形压缩编码把 PCM 编码作为输入,并对其进行压缩。

2) DPCM(Differential Pulse Code Modulation,差分脉冲编码调制)

DPCM 是利用声音信号的相关性,通过只传输声音的预测值和样本值的差值来降低声音数据的编码率的一种方法。它采用预测编码技术,实现声音数据的压缩编码。

因为声音信号一般不会发生突然变化,相邻的语音采样值之间存在很大的相关性,从一个采样值到相邻的另一个采样值的差值要比样本值本身小得多。利用预测编码方法建立预测模型,通过预测器对未来的样本进行预测,然后对样本值与预测器得到的预测值之差进行量化和传输。由于这个差值的幅度远远小于样本值本身,需要较少的比特数来表示,这样可以降低数据的编码率,从而使编码数据得到压缩。

3) ADPCM(Adaptive Differential Pulse Code Modulation,自适应差分编码调制)

在实际使用中,由于输入信号的不稳定性,造成 DPCM 方法的信噪比大大降低。因此在 DPCM 编码中加入自适应的方法,就形成了自适应差分编码调制(ADPCM)方案。所以,ADPCM 是对 DPCM 方法的改进,通过调整量化位数,对不同的频段设置不同的量化位数,可使数据得到进一步压缩。

ADPCM 压缩方案压缩倍率可达 2~5 倍,信噪比高,性能优越,因此,多媒体计算机所获得的数字化的声音信息大都采用此压缩方法。MPC 的声卡也提供有 ADPCM 算法,如将 16 位的采样值压缩成 4 位,将 8 位的采样值压缩成 4 位、3 位或 2 位。

2. 参数编码

波形编码方法的编码率比较高,但可以获得较好的音质。除此之外,还有一类编码方式是通过建立声音的产生模型,将声音信号以模型参数表示,再对参数进行编码,这种编码方式称为参数编码。声音重放时,再根据这些参数通过合成各种声音元码来产生声音。参数

编码压缩倍数很高,但计算量大,而且保真度不高,合成声音的质量不如波形编码,所以,它适合于语音信号的编码。

3. 混合编码

将波形编码和参数编码的方法结合起来就称为混合编码。这是一种吸取波形和参数编码的优点,进行综合编码的方法,在降低数据率的同时,能够得到较高的声音质量。典型的混合编码方法有码本激励线性预测编码(CELP)和多脉冲激励线性预测编码(MRLPC)等。

3.1.3 数字声音编码标准

当前编码技术发展的一个重要方向就是综合现有的编码技术,制定全球的统一标准,使信息管理系统具有普遍的互操作性,并确保了未来的兼容性。

1. ITU-T G 系列声音压缩标准

CCITT 和国际标准化组织(ISO)先后提出一系列有关声音编码的建议,对语音信号压缩编码的审议在国际电报电话咨询委员会(CCITT)下设的第十五研究组进行,相应的建议为 G 系列,多由 ITU 发表。

1) G. 711

本建议公布于 1972 年,它给出了语音信号的编码的推荐特性。语音的采样率为 8kHz,每个样值采用 8 位二进制编码,推荐使用 A 律和 μ 律编码。本建议中分别给出了 A 律和 μ 律的定义,它是将 13 位的 PCM 按 A 律,14 位的 PCM 按 μ 律转换为 8 位编码。

2) G. 721

该建议公布于 1984 年,1986 年做了进一步修订。采用自适应差值量化的算法对声音波形编码,数据率为 32kb/s,用于把 64kb/s 的 A 律或 μ 律的 PCM 编码转换成 32kb/s 的 ADPCM 编码,实现对 PCM 信道的扩容。

G. 721 和 G. 711 标准都适用于 200~3400Hz 窄带语音信号,可用于公共电话网。

3) G. 722

这个建议公布于 1988 年,它是针对宽带语音制定的标准,给出了 50~7000Hz 声音编码系统的特性,可用于各种高质量语音的应用。它的编码系统采用子带自适应差分脉冲编码(SB-ADPCM)技术,整个频带分成高和低两个子带,用 ADPCM 分别对每个子带进行编码。系统的比特率为 64kb/s,所以称为 64kb/s(7kHz)声音编码。

4) G. 728

为了进一步降低数据速率,实现低码率、短延时、高质量的目标。在 AT&T Bell 实验室研究的 16kb/s 短延时码激励编码方案(LD-CELP)的基础上,CCITT 于 1992 年和 1993 年分别公布了浮点和定点算法的 G. 728 语音编码标准,该算法编码延时小于 2ms。这个标准可用于可视电话、无绳电话、数字卫星通信、公共电话网、ISDN、数字电路倍增设备(DCME)、声音存储和传输系统、声音信息的记录和发布、地面数字移动雷达、分组化语音等。

2. MPEG 中的声音编码

国际标准化组织/国际电工委员会(ISO/IEC)所属 WG11 工作组,制定推荐了 MPEG 标准。已公布和正在讨论的标准有 MPEG I, MPEG E, MPEG N, MPEG U。其中 MPEG I 标准对应于 ISO/IEC 11172-3(MPEG 音频)。这部分规定了高质量声音编码方法、存储表

示和解码方法。编码器的输入和解码器的输出与现存的 PCM 标准兼容。ISO/IEC 11172 视频、音频的总数据率为 1.5Mb/s。音频使用的采样率为 32kHz, 44.1kHz 和 48kHz。编码输出的数据率有许多种,由相关的参数决定。

MP3 作为目前最为普及的声音压缩格式,为大家所普遍接受,各种与 MP3 相关的软件产品层出不穷,而且更多的硬件产品也开始支持 MP3。MP3 是 MPEG Audio Layer-3 的简称,是 MPEG-1 的衍生编码方案,可以做到 12:1 的惊人压缩比并保持基本可听的音质,随着网络的普及,MP3 被数以亿计的用户接受。

3. AC-3 编码和解码

AC-3 声音编码标准起源于由美国的杜比(DOLBY)公司推出的 DOLBY AC-1。AC-1 应用的编码技术是自适应增量调制(ADM),它把 20kHz 的宽带立体声声音信号编码成 512kb/s 的数据流。AC-1 曾在卫星电视和调频广播中得到广泛应用。1990 年,DOLBY 实验室推出了立体声编码标准 AC-2,它采用类似 MDCT 的重叠窗口的快速傅立叶变换(FFT)编码技术,其数据率在 256kb/s 以下。AC-2 被应用在 PC 声卡和综合业务数字网等方面。

1992 年,DOLBY 实验室在 AC-2 的基础上,又开发了 DOLBY AC-3 的数字声音编码技术。AC-3 提供了 5 个声道从 20Hz 到 20kHz 的全通带频,即正前方的左(L)、中(C)和右(R),后边的两个独立的环绕声通道左后(LS)和右后(RS)。AC-3 同时还提供了一个 100Hz 以下的超低音声道供用户选用,以弥补低音的不足,此声道仅为辅助而已,故定为 0.1 声道。所以 AC-3 被称为 5.1 声道。AC-3 将这 6 个声道进行数字编码,并将它们压缩成一个通道,而它的比特率仅是 320kb/s。

3.1.4 数字声音信息的质量与数据量

采样、量化和编码技术是声音数字化的关键技术。而采样频率、每个采样值的量化位数以及声音信息的声道数目,是影响数字化声音信息质量和存储量的三个重要因素。采样频率越高、量化位数越大、声道数目越多,声音的质量就越高,但存储量就越大。

1. 声音质量的评价

声音质量的评价是一个很困难的问题,也是一个值得研究的课题。目前声音质量的度量有两种基本方法,一种是客观质量度量,另一种是主观质量的度量。

1) 客观质量的度量

对声波的测量包括评价值的测量、声源的测量和音质的测量,其测量与分析工作,是使用带计算机处理系统的高级声学测量仪器来完成的。度量声音客观质量的一个主要指标是信噪比(Signal to Noise Ratio, SNR),信噪比是有用信号与噪声之比的简称,其单位是分贝(dB)。信噪比越大,声音质量越好。

2) 主观质量的度量

采用客观标准方法很难真正评定编码器的质量,在实际评价中,主观的质量度量比客观质量的度量更为恰当和合理。主观的质量度量通常是对某编码器的输出的声音质量进行评价。例如播放一段音乐,记录一段话,然后重放给一批实验者听,再由实验者进行综合评定,得出平均判分(Mean Opinion Score, MOS)。这种判分采用 5 级分制,如表 3-1 所示为不同的 MOS 对应不同的质量级别和失真级别。

表 3-1 MOS 标准

MOS	质量级别	失真级别
5	优(Excellent)	不察觉
4	良(Good)	刚察觉但不可厌
3	中(Fair)	察觉及稍微可厌
2	差(Poor)	可厌但不令人反感
1	劣(Unacceptable)	极可厌(令人反感)

3) 常用的数字化声音技术指标及音质

常用的数字化声音技术指标及声音的质量如表 3-2 所示几种。

表 3-2 常用的数字化声音技术指标及声音的质量表

采样频率 kHz	量化位数 b	每分钟数据量(无压缩)MB		常用编码 方法	质量与应用
		单声道	双声道		
44.1	16	5.05	10.09	PCM	相当于 CD 质量,应用于超高保真质量要求
22.05	16	2.52	5.05	ADPCM	相当于收音质量,可应用于伴音及各种音响效果
	8	1.76	2.52	ADPCM	
11.025	16	1.76	2.52	ADPCM	相当于电话质量,可应用于伴音或解说词的录制
	8	0.63	1.26	ADPCM	

2. 声音信息的数据量

确定了数字声音的采样频率、量化位数和声道数就可以计算出声音信息的数据量,其计算公式为:

$$S = R \times r \times N \times D / 8$$

其中: S 表示文件的大小,单位是 B;

R 表示采样频率,单位是 Hz;

r 表示量化位数,单位是 b;

N 表示声道数;

D 表示录音时间,单位是 s。

3. 应用举例

下面使用 Windows 系统中的录音机录制一段声音,根据不同参数的选择,体验声音的质量,并观察声音文件数据量的变化。

(1) 在 Windows“附件”的“娱乐”组中启动“录音机”,如图 3-2 所示。“录音机”实际上就是一个非常简单实用的声音编辑应用软件。

(2) 单击“录音”按钮,开始录音,此时,声波窗口中出现声音波形,并在声波窗口左侧记录当前录音的位置,而声波窗口右侧则显示所录制声音文件的长度,如图 3-3 所示。如果要结束录音,则单击“停止”按钮.



图 3-2 启动“录音机”

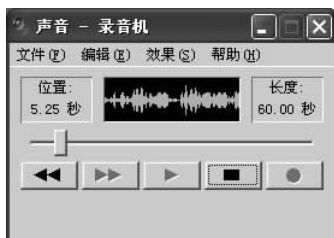


图 3-3 录制声音

(3) 执行“文件”|“属性”命令,打开“声音的属性”对话框,如图 3-4 所示。可以看出,采用 PCM 编码,采样频率为 22 050Hz,量化位数为 16 位,单声道,录音 1 分钟(60 秒)的数据量为: $22\,050 \times 16 \times 1 \times 60 / 8 = 2\,646\,000$ (字节)。

(4) 如果要转换声音的格式,单击“声音的属性”对话框中的“立即转换”按钮,打开“声音选定”对话框,如图 3-5 所示。

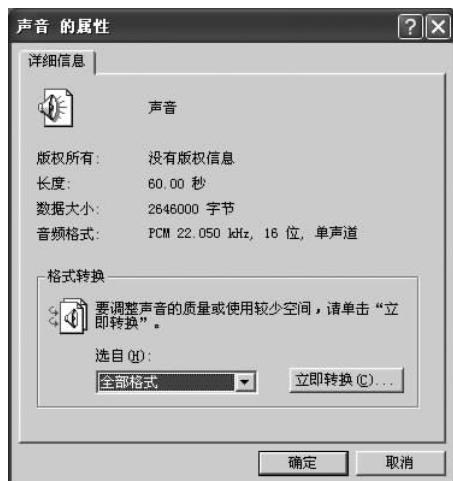


图 3-4 “声音的属性”对话框



图 3-5 “声音选定”对话框

(5) 单击“名称”下拉列表,可以选择一种预定的声音质量,单击“格式”下拉列表可以选择声音的压缩格式,如图 3-6 所示。同样也可以在“属性”下拉列表框里选择采样频率、量化位数和声道数。

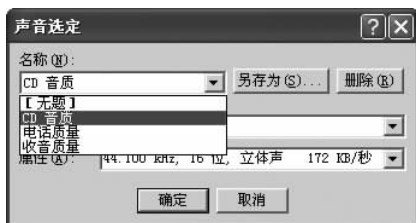


图 3-6 声音格式的选择

(6) 单击“确定”按钮,即可进行格式的转换。

3.1.5 合成声音与 MIDI 规范

1. 音乐合成

一个乐音必备的三要素是：音高、音色和音强。运动的旋律中的乐音还应具备时值，即持续时间。这些要素的理想配合是产生优美动听的旋律的必要条件。

1) 音高

音高指声音的基频。人们通常所说的歌唱演员有女高音、男低音，这里的高和低就是指声音的基频。声音的基频越高，给人的感觉就越激越；相反，声音的基频越低，给人的感觉就越低沉。在处理声音的时候，可以通过改变声音的频率来提高或降低音调。

2) 音色

音色是由声音的频谱决定的，各阶谐波的比例不同，随时间衰减的程度不同，音色就不同。人们能够分辨具有相同音高的钢琴和小号的声音，就是因为它们的音色不同。“小号”的声音之所以具有极强的穿透力和明亮感，是因为“小号”的声音中高次谐波非常丰富。各种乐器的音色是由其自身结构特点决定的。要用计算机模拟具有强烈真实感的旋律，音色的变化是非常重要的。

3) 音强

音强也叫响度，是指声音信号的强弱程度，是由声波振动的振幅决定的。人耳对于声音细节的分辨与音强有关，只有在音强适中时，人耳辨音才最灵敏。如果一个声音响度太低，便难以正确判别它的音高和音色；而响度过高，也会影响判别的准确性。

4) 时值

声波振动的持续时间称为时值，它具有明显的相对性，一个音只有在包含比它更短的音的旋律时才会显得长。时值的变化导致旋律是平缓、均匀或是跳跃、颠簸，它可以表达不同的情感。

5) 音乐合成

任何一种波形信号都可以被分解成若干个频率不同的正弦波，一个乐器的声音也可以由若干个正弦波合成得到。调频(FM)是使高频振荡波的频率按调制信号规律变化的一种调制方式。采用不同调制波频率和调制指数，就可以方便地合成具有不同频谱分布的波形，再现某些乐器的音色。可以采用这种方法得到具有独特效果的“电子模拟声”，创造出丰富多彩的、真实乐器所不具备的音色，这也是FM音乐合成方法特有的魅力。为了使音乐更加真实，人们开发出波形表(wavetable)音乐合成技术。波表合成是把真实音乐声音的数字信号录制后，保存在存储器中，当选择某个乐器时，将所录制的样本信号回放。目前这两种音乐合成技术都应用于多媒体计算机的声卡中。

2. 什么是 MIDI

MIDI 是英文乐器数字接口(Musical Instrument Digital Interface)的缩写，它是一种技术规范，定义了为把电子乐器连接到计算机所需要的电缆和端口的硬件标准，计算机和具有MIDI接口的设备之间进行信息交换的规则，电子乐器之间传送数据的通信协议。它于1988年正式成为数字式音乐的一个国际标准。

凡具有处理MIDI信息的处理器和适当的硬件接口都能构成MIDI设备，如电子钢琴、电子键盘、电子吹奏乐器、电吉他、电子打击乐器，以及计算机的声卡等都是MIDI设备。

MIDI 声音和数字化波形声音完全不同。在记录时,它不像数字化波形声音那样,对声音的波形进行采样、量化和编码,而是记录电子乐器键盘的弹奏过程,例如记录按了哪一个键、力的大小和时间的长短等,实际是将乐曲进行一种数字化的描述,这种描述称为 MIDI 消息(MIDI Message)。当需要播放这段音乐时,从相应的 MIDI 文件中读出 MIDI 消息,由合成器来解释这些消息中的符号,并生成所需要的乐器的声音波形,经放大后由扬声器输出。

由于 MIDI 文件并不记录任何声音的波形,只是一系列指令符号的集合,播放时需要通过音乐合成器的芯片来解释这些符号并产生声音的波形,所以在计算机中播放 MIDI 信息必须使用带有合成器的声卡。

与其他声音相比,MIDI 声音有其自己的优点。

(1) 存储容量小。因为 MIDI 文件中记录的是一系列指令的集合,它并不记录任何声音,所以,对同一首乐曲来讲,MIDI 文件所占的存储量非常小。因此,在需要播放长时间的高质量音乐时,往往采用 MIDI 文件。

(2) 可以提供背景音乐或音响效果的配音功能。当多媒体计算机显示图像、文字、图表时,或者播放波形声音、语音(如解说词)时,可以同时播放 MIDI 音乐作为背景音响效果。

(3) 便于编辑和修改。与波形声音相比,MIDI 声音记录的是符号,是乐谱的数字化表示。因此,它可以在计算机中很方便地任意修改乐曲的速度、音调,甚至可以更换乐器,从而得到不同的效果。在此基础上,人们可以用作曲软件设计乐曲,生成乐谱文件。

(4) 可以在 MIDI 合成器中完全重现原来的演奏。MIDI 音乐的产生是把 MIDI 设备上产生的每个活动记录下来,形成 MIDI 文件,因此,把它再传送到 MIDI 合成器时,就可以完全重现原来的演奏效果。

3. MIDI 的有关术语

(1) 通道(Channels): MIDI 可为 16 个通道提供数据,每个通道可以访问一个独立的逻辑合成器。微软公司使用 1~10 通道作为扩展合成器,13~16 通道作为基本合成器。

(2) 音序器(Sequencer): 是为了 MIDI 作曲而设计的计算机程序或电子装置,用于记录、编辑、播放 MIDI 文件。

(3) 合成器(Synthesizer): 是利用数字信号处理器或其他芯片产生音乐或声音的电子设备。它可以产生并修改波形,然后通过声音产生器和扬声器发出声音。

(4) 乐器(Instrument): 合成器能产生的特定声音称为乐器。每种乐器都有自己的波形,合成器按音色和音调的要求,由不同的乐器组合成最终的声音组合。不同的合成器,乐器音色号不同时,声音的质量也不同。例如,多数乐器都能合成钢琴的声音,但是不同乐器使用的音色号不同时,输出的声音也会有差异。

(5) 复音(Polyphony): 复音是合成器同时支持的最多音符数。如一个能以 6 个复音合成 4 种乐器声音的合成器,可以同时演奏分布于 4 种乐器的 6 个音符。例如,它可以是 4 个音符的钢琴和弦、一个长笛和一个小提琴的声音。

(6) 音色(Timbre): 音色指的是声音的音质,它取决于声音频谱。小提琴、钢琴、长号等都有各自的音色。

(7) 音轨(Track): 一种用通道把 MIDI 数据分隔成单独组、并行组的文件的概念。0 号格式的 MIDI 文件把这些音轨合并成一个。1 号格式 MIDI 文件维持不同的音轨。

4. MIDI 规范

MIDI 规范是一个国际的标准,主要包括以下三个方面的内容。

(1) MIDI 的硬件规范。指的是各种 MIDI 设备之间连接的硬件接口标准和信号传输机制,包括输入/输出通道的类型,连接电缆样式及插座形式等。

(2) MIDI 声音信息的规范。指的是使音乐信息互相交换的一种编码标准。它包括有关音乐成分的信息,如音符、音量、音调、音符时间长短等,是一种表达各种声音的作曲系统。

(3) MIDI 声音合成的规范。指的是各种声音的表达方式,即真实声音信号的规范,它可以采用 FM 合成技术和波形表合成技术的标准。

3.1.6 声音文件的格式

1. WAV 文件

WAV 文件又称波形文件,是 Microsoft 公司的声音文件格式。WAV 格式作为 Microsoft 的标准文件格式,用于保存 Windows 的声音信息资源,被 Windows 平台及其应用程序所广泛支持。WAV 文件来源于对声音的模拟波形的采样,并以不同的量化位数把这些采样点的值转换成二进制数,然后存入磁盘,这就产生了波形文件。

WAV 格式支持 MSADPCM、A 律、 μ 律和其他压缩算法,支持多种音频位数、采样频率和声道,是 PC 上最为流行的声音文件格式,但其文件尺寸较大,多用于存储简短的声音片断。

2. VOC 文件

VOC 文件是 Creative 公司所使用的标准声音文件格式,也是声霸卡(Sound Blaster)所使用的声音文件格式,其文件结构与 WAV 文件类似。

3. MIDI 文件

MIDI 文件是存放 MIDI 信息的标准文件,文件名后缀为. MID。MIDI 文件中包含音符、定时和多达 16 个通道的演奏定义。文件包括每个通道的演奏音符信息:键、通道号、音长、音量和力度。

4. MP3 文件

符合 MPEG 标准中的 MPEG 音频格式的文件称为 MPEG 音频文件。MPEG 音频文件的压缩是一种有损压缩,根据压缩质量和编码复杂程度的不同可分为三层(MPEG Audio Layer1/2/3),分别对应 MP1、MP2 和 MP3 这三种声音文件。MPEG 音频编码具有很高的压缩率,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。CD 音质的音乐,未经压缩需要 10MB 存储空间,而经过 MP3 压缩编码后只有 1MB 左右,其音质基本保持不失真。

Internet 的发展和普及,促进了 MP3 的流行,网络代替了传统唱片的传播途径,扩大了数字音乐的流传范围,加速了数字音乐的传播速度,在一些常用搜索引擎中,MP3 的使用率已名列首位。MP3 凭借其优美的音质和高压缩比而成为最流行的音乐格式。

5. RealAudio 文件

RealAudio 文件是 RealNetworks 公司开发的一种新型流式声音(Streaming Audio)格式的文件;它包含在 RealNetworks 所制定的音频、视频压缩规范 RealMedia 中,主要用于在低速率的广域网上实时传输声音信息;网络连接速率不同,客户端所获得的声音质量也不尽