

多媒体计算机的数字音频处理技术*

陈 锐 生

(华侨大学计算机科学系, 泉州 362011)

摘要 介绍计算机的音频扩充卡的结构和功能, 分析多媒体数字音频处理技术的若干关键技术及其相互关系和发展趋势。

关键词 数字音频, 音频卡, 多媒体

分类号 TP 334. 402

声音是多媒体计算机中不可缺少的要素之一, 音乐、音效和语音使多媒体应用程序富有活力。尽管 APPLE 公司的 MAC 微机早已将声音作为其系统标准配置, 但目前在国内外更为流行的普通 IBM 兼容微机却只有一个蜂鸣器, DOS 系统也缺乏实用的音效功能。随着 WINDOWS 系统的发展, IBM 兼容机已具有包括音效在内的多媒体系统功能, 应运而生的音效卡成了目前 IBM 兼容机多媒体化的必要扩充部件^[1]。众所周知, 声音是空气振动的传播, 用一种模拟的连续波形表示, 它具有振幅和周期等特征。要将模拟的波形转换成数字形式, 必须按一定的采样频率和量化精度将连续的声波在时间上离散化和振幅上量化, 才能将声音存入计算机。按照 Nyquist 定理, 要将数字化的声音完美地恢复成原来的连续波形, 必须将采样频率设定到所要记录声音波形中最高频率分量的一倍。人的听觉范围大约在 20~22 kHz, 因此标准的声音采样频率有 44.1 kHz, 22.05 kHz 和 11.025 kHz, 而波形振幅的量化级有 8 比特(256 级)和 16 比特(65 536 级)两种。采样频率越高、量化级越多, 采样样本就越接近原始的模拟波形, 所需要的处理时间和存储量也越大。

1 音频卡的结构与主要功能

目前市场上的音频卡功能上有很大区别, 但它们的基本结构可以分成以下四个部分。

(1) 音频输入输出部分。因为声音是模拟信号, 计算机能处理的却是数字信号, 所以音频卡必须带有模数及数模转换电路, 接受声音输入时经过模数转换将声音模拟信号转换为数字信号后存储在计算机上^[2], 声音输出时经过数模转换将数字信号转换为模拟信号输出计算机, 作为音响设备的音源。有的音频卡带有小功率音频放大器, 可直接驱动无源音箱。此外, 音频输入输出部分通常还支持 DMA 数据的传输功能, 使声音数字信号直接传送到内存, 减轻 CPU 的负担。

* 本文 1995-10-23 收到

(2) 音乐合成部分. 目前的音频卡可带有频率调制(FM)和乐器波表(Wavetable)两种音乐合成方法. FM 频率调制合成通过调用两个或两个以上的原始波形来产生声音,其中包括正弦波、三角波和方波. 用户可利用频率调制营造出特别的音效,但有关的算术函数与现实生活中的声音却难以有一个实际的关系,频率调制合成出的音色少,音质差,不过由于其成本价格低廉而得到广泛应用. 波表合成是一种较新的发音方法,其实其发音机理是很简单的:首先对各种真正乐器的声音进行数字化采样创造出波形数据,然后将各种乐器的波形数据依次固化在只读存储器中形成乐声记录目录,该目录就称为乐器波表. 当一个程序通知波表合成器要演奏哪种乐器时,有关硬件就会用查表法从波表中挑出对应的乐声记录,重新制造出保真度很高的声音,不过查表法得出的波形数据还不能马上送到数模转换器输出,需要进行一些处理,因为采样波形的数据量是很大的,而存储空间却很有限,不可能将每种乐器的每个音高的波形数据都存储下来,存储器中只保存乐器发音较好的若干音高波形样本,其它音高的波形只能通过对这些样本进行数字信号处理得到.

(3) 信号处理部分. 声音数据的压缩是通过修改在模数转换中间的编码过程,以完成压缩. 解压缩算法则是通过修改在数模转换中间的解码过程来完成解压缩的. 声音数据的压缩解压缩(CODEC)算法有多种,这些算法被结合进应用软件,或者编进专门的集成电路,对 16 位的立体声以 44.1 kHz 采样进行简单的处理就会占去 CPU 的相当一部分资源. 因此,通常采用在音频卡上设立专用的集成电路(IC)或通用的数字信号处理器(DSP)的方法完成声音数据的压缩解压缩. 在设有 DSP 的音频卡上,用户还可附加特别音效给模拟音频或者音乐合成输出. 特别音效包括混音、延时、合唱和特别的立体声空间效果;混音器是每块音频卡都有的,其主要作用是来自音乐合成器、模拟音频输出和 CD-ROM 驱动器的模拟音频,以不同音量大小混合在一起送到音频卡的输出端口.

(4) 接口控制部分. 音频卡上还有许多不同的输入输出接口,包括与计算机的总线接口、CD-ROM 控制接口、CD 音频连接器、波表存储器接口、麦克风输入、外部线上输入、扬声器/耳机输出、线上输出和游戏棒/MIDI 接口等.

2 音频数据压缩技术

在多媒体技术中,虽然声音的数据量远远不如视频图像数据量大,但是如果想保存一分钟的单声道声音片段所需要的存储空间(附表)^[3].

附表 一分钟单声道声音片段存储空间表

量化精度/bit	采样频率/kHz	所需存储量/MB · min ⁻¹
8	11.025	0.661
8	22.050	1.323
16	22.050	2.646
16	44.100	5.292

立体声所需存储空间为声道的两倍,可见数字的声音压缩也是很重要的.

在 WINDOWS 多媒体应用中,至今没有为压缩和解压缩波形音频文件确立最终的标准. INTEL 的数字视频间隔法(DVI)和 MICROSOFT 的音频视频间隔法(AVI),以及 CDROMXA,都使用不同的压缩解压缩技术. AVI 使用软件来压缩和解压缩音频数据,其它

方法需要在音频卡上增加专门的 IC,但它们的基本方法是类似的. 对模拟音频的压缩解压缩技术都是基于适配微分脉冲编码调制(ADPCM)的. ADPCM 或与其紧密相关的压缩方法,能提供实时基础上的最大压缩率.

通过模数转换器(ADC)把音频信号转换成数字信号后,如果采样的频率足够快,在任何两个相继的样品值之间的最大增量不会超过 8 比特转换器的 1 比特范围或者 16 比特转换器的 2 比特范围. 这时,如果只保留相继采样间的增量,就能大量地减少用来传输和存储所需的数据量,而不损失音质. 因此,微分脉冲编码调制也称增量调制,它要求在下一个样本在采样之前预测其数值,预测值通过对最新采集的一定数量的样本值求平均数而得. 在实测值与预测值之间的增量用一个比特保存,增量为正时设置为 1,增量为非负时设置为 0,如果采样频率足够快,相邻样本值之间就只有一比特的增量,这样只保存一比特也不会损失音质,但是前提是要有足够快的采样频率. 因此,增量量化比特数的节省是以采样密度增加为代价的,采样后的音频数据总量并不一定减少,实用上增量调制只是以适当高的采样频率进行一比特增量量化的,必然存在一些失真. 虽然增量调制在电话语音信号的数字传输中被采用,因为它不需要高保真度,但要录音就不合适了. 如果要用有限的存储空间保存尽量多的高保真音频数据,就需要在增量量化级数和采样频率的相互消长关系中寻求最佳组合.

适配微分脉冲编码调制(ADPCM)是微分脉冲编码调制的扩展,它使用了一位来描述两个相邻采样值的增量. PHILIPS 使用 4 比特和 8 比特 ADPCM 来对 CD-I 交互式 CDROM 上的音频进行编码,SONY 的 BOOKMAN 使用的 CDROMXA 标准支持 4 和 8 比特 ADPCM,在为 WINDOWS ADPCM 音频的 RIFF 文件的最终标准建立之后,ADPCM 将作为“建议”的功能被加入多媒体 PC 规范. ADPCM 能够由专门设计的 IC 或者 DSP 来完成,几年前已经有了能够实现 ADPCM 的专门 PC 适配卡. 例如 IBM 为存储在 CDROM 或者硬盘的多媒体教程提供了采用 ADPCM 的音频卡,使用 4 比特压缩的 ADPCM 音频增加了能够放入 CDROM 的课程数量,也减少了教学中 CDROM 的交换.

3 数字信号处理器的应用

数字信号处理器(DSP)是通用的微处理器,在现实生活中进行各种信号处理,以 DSP 为基础的音频正在取代专用芯片的解决办法,专用芯片的解决办法包括用于调频合成、波表合成和数据压缩的各种芯片. 对每个 DSP 功能,不是使用专用芯片,而是提供波表合成和帮助音频数据的 CODEC. 使用带有 DSP 的音频卡来处理音频数据可节省计算机资源,让其可以在显示动画或者视频图象的同时又演奏高保真音乐. 由于针对连续采样过程进行了 DSP 微处理器结构的优化,DSP 使得在 ADPCM 技术中用到的预测编码和其它计算比相同速度的通用微机 80386 和 809486 更为快捷. DSP 拥有片内的内存和自己的 I/O 总线结构以便以更多的方式来增加 RAM 和 ROM. 新的 DSP 还带有浮点数字处理器,与 PC 中的 80287 和 80387 类似,以加快精确的数值运算. 计算机通过 I/O 总线向音频卡上的 DSP 发送数据和简单的指令,音频卡负责音频数据处理过程中的所有复杂操作.

在音频卡上附加 DSP 的另一基本原因是,通过替换 DSP 上的不同 ROM 芯片或者从 WINDOWS 应用软件向音频卡上发出特殊的指令,便可重新对 DSP 编程. 与一个专用的 ADPCM 芯片相比较,如 SOUND BLASTER 所用,DSP 芯片的可重编程能力减少了浪费. 比

如,当活动图象专家组(MPEG)的新音频压缩标准出现时,只需更换一片 ROM 芯片来处理它,如果在卡上增加可选的 RAM 芯片,还可以在启动 WINDOWS 时把程序调入卡中,以执行专门的音频信号处理。

高档音频卡上的数字音频效果处理器,它利用 DSP 芯片来产生各种数字化声音效果。由音乐设备数字接口(MIDI)控制的效果处理器先取得一个模拟信号,进行采样,对样本进行处理,然后把声音转换成模拟形式。在这过程中,DSP 通过对原始数据进行样本复制或者延时,能制造出混响或者合唱等声音效果;通过在 DSP 执行专门的代码,它也完全能够模仿电吉他的“效果处理器”功能。根据计算机的速度和 DSP 代码的效率,多媒体系统能够实时生成效果,如果没有高速 DSP 处理音频数据,就只能采用块读写方式处理存放在波形文件中的数据,不能实时生成音频效果。

以 DSP 为基础的音频卡除了高保真功能外,还可以提供调制解调器和电话应答机功能,而这些多功能卡并不比一般音频的价格贵多少。DSP 省去了处理数模转换和模数转换所需要的许多电路,与当今高质量音频卡上的专门 IC 比较,这种能力部分补偿了 DSP 芯片及其内存芯片的额外费用。

4 MIDI 音乐

音乐设备数字接口(MIDI)是计算机和音乐结合的产物,它是一个通过电缆将电子音乐设备连接起来的国际标准协议,其目的是要使多种不同类型的乐器一起合奏。MIDI 的主要应用就是将乐器连上个人计算机,以便向有关设备传送数字化的指令。与其它音乐设备或者其它数字音频技术相比,MIDI 的特点在于它处理音频的方式,它不是将声音编码,而是将 MIDI 乐器上产生的每一个活动编码记录下来。例如,在给琴键上的演奏录音时,计算机用 MIDI 文件格式记录下按了哪个键,力度有多大,时间有多长。每个 MIDI 文件最多能存放 16 个音乐通道的信息。

用计算机可以对这些 MIDI 文件进行编曲和弹奏,弹奏 MIDI 文件时,音序器将 MIDI 消息发送到合成器中,合成器将这些消息转换成某种乐器的声音、合成音色和持续时间,合成器使用 DSP 或其它芯片生成并修改波形,然后将它送到声音发生器和扬声器。需要注意的是,在弹奏 MIDI 文件的过程中计算机并没有产生任何声响,它只是告诉琴键该弹奏哪个键或音符以及何时弹奏,而真正发声的还是乐器本身。演奏存储在 MIDI 文件中的音乐有两种方式,一是通过音频卡中的音乐合成器,二是通过与计算机连接的外部 MIDI 声音合成器。

与多媒体计算机的其它音频技术相比,MIDI 有许多优点。首先,它可节约大量存储空间,一小时的立体声音乐大约只要 400KB,而用其它音频文件方式则需要数百 MB 的存储空间,由于 MIDI 文件小得多,装载 MIDI 文件也比装载波形文件容易得多,这在开发多媒体应用和指定何时播放音乐带来很大方便。其次,是 MIDI 文件在进行音乐编辑时很灵活,在音序器的帮助下可以不断改变音高、声调和不同乐器声音,直到得到想要的音乐。再次,在音乐表现方面,利用 MIDI 技术,在播放波形文件的同时可在 WINDOWS 的后台播放 MIDI 文件以实现配乐,但是不能同时播放两个波形音频文件。

MIDI 文件对存储空间的低要求,使得它成为多媒体计算机的一种重要合成音乐方法。但是,采用 MIDI 文件播放的合成声音质量则受到合成器中乐器组合的限制,不同的音频卡之间

会存在许多声音的区别,不过还可以通过调用音序程序来修改 MIDI 文件进行调整.而采用波形音频文件的优点是可以从任何音源录制声音,并保证它们在任何一台多媒体计算机上都有大致相当的播送效果.

5 结束语

随着多媒体计算机技术的迅速发展,人们对多媒体计算机的声音效果要求也越来越高,数字音频技术也日趋成熟和规范.在目前采用音频卡的多媒体扩充方案中,有两类数字音频生成方式:一是以声音波形数据文件为基础的,采用微分脉冲编码调制技术,目前市场上的音频卡已经形成了以新加坡 Creative Labs 公司的声霸专家 2.0 为主的实际上的工业标准(8 位立体声),并向声霸 16(16 位立体声)标准过渡;二是以音乐合成为基础,它又有频率调制(FM)和乐器波表(Wavetable)这两种方式.频率调制合成音乐是任何音频卡都具有的基本功能,波表合成音乐是提高音效的发展方向,而音乐设备数字接口(MIDI)是音乐合成的标准.许多高档音频卡已经采用数据信号处理芯片 DSP 和波表音乐合成器,能够提供音色丰满的 16 位立体声.以 WINDOWS 为基础的各种音乐编辑软件是进行多媒体音乐创作的有效工具.

相信不久的将来,结合音频 DSP 和波表音乐合成的数字音频技术,将成为个人计算机的标准功能集成到 CPU 和主板上,以减少由于缺乏严格工业标准的各种扩充卡带来的兼容性问题.

参 考 文 献

- 1 吴伟煜.多媒体技术开发指南.大连:大连理工大学出版社,1994.239~263
- 2 钟玉琢.多媒体计算机技术.北京:清华大学出版社,1993.15~20
- 3 Microsoft P. MS WINDOWS 多媒体制作和工具指南.张 蒴等译.北京:海南出版社,1993.41~52

Digital Audio-Frequency Processing Techniques for a Multi-Media Computer

Chen Duansheng

(Dept. of Computer Science, Huaqiao Univ., 362011, Quanzhou)

Abstract Beginning with a description of structure and function of an audio-frequency extended card for a personal computer, the author analyses the key techniques of multi-media digital audio-frequency processing and their interrelation; and also the tendency of development.

Keywords digital audiofrequency, audiofrequency extended card, multi-media