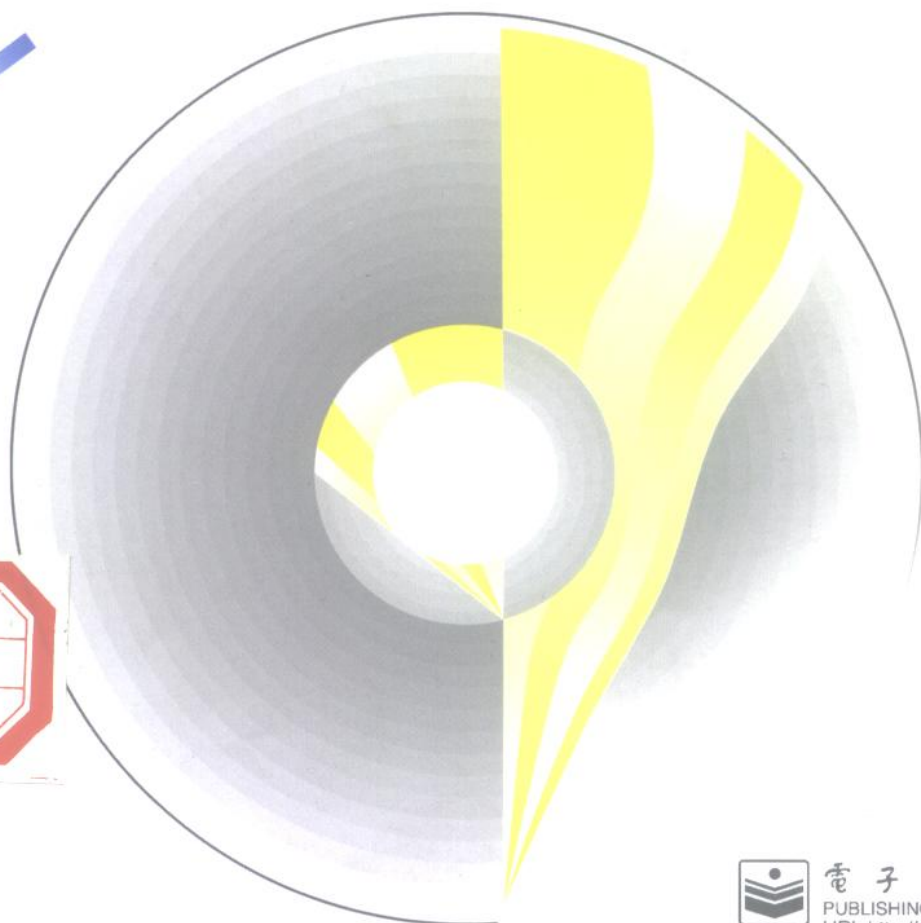


多媒体计算机系列丛书

多媒体计算机声卡 技术及应用

周敬利 余胜生 等 编著



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS
URL: <http://www.phei.com.cn>

TP334
Z81

416011

多媒体计算机系列丛书

多媒体计算机声卡技术及应用

周敬利 余胜生 等编著



00416011

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

声卡是多媒体计算机最基本的部件。本书从数字音频技术与计算机音乐、声卡的结构、音频文件格式及应用软件等方面讲述声卡技术及应用。

全书共 10 章,包括:多媒体计算机音乐;声卡及其芯片的工作原理;MIDI 基础知识及合成器;光盘存储技术与光盘驱动器;音频文件格式及应用软件;声卡驱动程序开发以及声卡的选择与安装。

本书是作者在科研和教学方面的总结,是国内作者较系统全面介绍声卡技术的书籍。它既可供从事多媒体计算机软硬件开发人员阅读,也可供广大师生及 MPC 使用者参阅。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,翻版必究。

图书在版编目(CIP)数据

多媒体计算机声卡技术与应用/周敬利,余胜生编著.北京:电子工业出版社,1998

(多媒体计算机系列丛书)

ISBN 7-5053-5018-8

I.多… II.①周… ②余… III.多媒体电子计算机-声音输出装置 IV.TP334

中国版本图书馆 CIP 数据核字(98)第 25323 号

JS186/12

丛 书 名:多媒体计算机系列丛书

书 名:多媒体计算机声卡技术及应用

编 著 者:周敬利 余胜生 等

责任编辑:邓又强

排版制作:电子工业出版社计算机排版室

印 刷 者:

北京科报印刷厂

装 订 者:

出版发行:电子工业出版社 URL:<http://www.phei.com.cn>

北京市海淀区万寿路 173 信箱 邮编 100036

经 销:各地新华书店

开 本:787×1092 1/16 印张:19.75 字数:480 千字

版 次:1998 年 12 月第 1 版 1998 年 12 月第 1 次印刷

书 号:ISBN 7-5053-5018-8
TP·2487

定 价:28.00 元

凡购买电子工业出版社的图书,如有缺页、倒页、脱页、所附磁盘或光盘有问题者,请向购买书店调换。

若书店售缺,请与本社发行部联系调换。电话 68279077

前 言

随着计算机的普及、高速信息公路的兴起,多媒体技术已开始进入千家万户。它使枯燥乏味的计算机屏幕变成丰富多彩、声像并茂的显示屏,改变了人们的学习与工作环境,为人类的文明开辟了一个崭新的时代。

一台基本的多媒体微型计算机,除 PC 机外,至少应包含声卡、CD-ROM 驱动器。如果需要更好的质量,则还可配上图像卡、扫描仪、磁带机等其它硬件设备,因此声卡及 CD-ROM 驱动器是多媒体计算机的最基本部件。故书中以此为主题进行阐述。

本书是笔者近几年来在科研与教学方面的总结,由周敬利、余胜生、王有成、袁双庆、何颖醛、韩琪、余嵘峰、李榕编写,其目的是希望本书能够对从事多媒体计算机软、硬件开发与使用的人员、广大师生以及从事相关领域(如通讯、管理、销售等)的人员,在学习和工作中有所裨益。

全书共分 10 章,第 1 章描述了多媒体计算机的基本组成和多媒体计算机音乐;第 2 章讲述了声卡的基本工作原理、性能、指标及分类;第 3 章描述了声卡硬件中关键芯片的各种类型和芯片编程;第 4 章为声卡应用的硬件环境及有关测试内容;第 5 章为乐器数字接口(MIDI)阐述有关的定义、标准及指令;第 6 章主要对多媒体计算机的 CD-ROM 进行介绍,阐述 CD-ROM 硬件组成的工作原理、类型、标准及相关接口;第 7 章是音频文件格式,包括 RIFF 文件和波形文件及 MIDI 音频文件等有关格式;第 8 章则重点介绍声卡的应用软件;第 9 章描述了声卡的驱动程序开发过程;第 10 章对各种声卡的选购及安装作了详细的说明。书中第 1 章 1.1 至 1.4 节由余胜生和韩琪编写,1.5 至 1.6 节由李榕编写;第 2 章、第 10 章由袁双庆编写;第 3 章由韩琪编写;第 4 章由余嵘峰编写;第 5 章和第 7 章由何颖醛编写,第 6 章由周敬利编写,第 8 章、第 9 章由王有成编写。全书由周敬利统编。

本书的录入与排版工作由张琴完成。

在编写此书过程中,参考和引用了许多文献及资料,均为国际标准、国外公司所馈赠以及国内公开发行的书刊及杂志上发表的文献。在此向这些提供文献资料的公司、作者、编者表示衷心的感谢。

在此书写作的过程中受到作者单位的领导、老师及朋友的帮助和支持,在此表示衷心的感谢。

在本书出版的过程中,得到出版社的领导和编辑的帮助与大力支持,为此书他们付出了辛勤的劳动,在此表示深深的谢意。

由于作者的学术水平及实际能力有限,错误与不妥之处,敬请读者指正。

周敬利于华工园

1997 年 12 月 28 日

目 录

第 1 章 多媒体计算机及多媒体计算机音乐	(1)
1.1 多媒体计算机的基本组成	(1)
1.1.1 MPC 1.0、MPC 2.0 标准	(1)
1.1.2 MPC 3.0 标准	(2)
1.2 数字音频技术	(5)
1.3 器乐声的计算机合成及演奏	(8)
1.4 音乐效果处理以及环境效果处理	(9)
1.5 数字滤波器及 FFT 频谱分析	(10)
1.6 数字音频处理中的特殊效果	(11)
1.6.1 数字延时	(11)
1.6.2 数字混响	(12)
1.6.3 音调变换	(14)
第 2 章 声卡的基本工作原理	(17)
2.1 声音信号的产生和处理	(17)
2.1.1 声音信号的产生与表达	(17)
2.1.2 声音信号的数字化	(18)
2.1.3 声音信号的重构	(21)
2.1.4 声音信号编码调制技术	(23)
2.1.5 数字化声音信号的容量与质量之间的矛盾	(27)
2.2 声卡的硬件结构	(28)
2.2.1 声卡的工作原理	(28)
2.2.2 声卡硬件结构举例	(28)
2.3 声卡的软件结构	(30)
2.4 声卡的应用及编程接口	(30)
2.5 声卡的主要性能指标和分类	(35)
第 3 章 声卡芯片工作原理	(37)
3.1 DSP 芯片	(37)
3.1.1 数字化声音处理	(37)
3.1.2 DSP 芯片的功能概述	(38)
3.1.3 DSP 编程简介	(38)
3.1.4 数字化声音 I/O 编程	(40)
3.1.5 DSP 命令	(54)
3.2 Codec 芯片	(65)
3.2.1 概述	(65)
3.2.2 模拟硬件描述	(66)

3.2.3 数字硬件描述	(68)
3.2.4 寄存器映射	(71)
3.2.5 编程考虑	(72)
3.3 OPTI 芯片	(77)
3.3.1 概述	(77)
3.3.2 特征	(78)
3.3.3 应用	(78)
3.3.4 寄存器映射	(78)
3.3.5 寄存器定义	(79)
第 4 章 声卡应用、测试及硬件环境	(85)
4.1 DOS 环境	(85)
4.1.1 DOS 下声卡的应用使用环境	(85)
4.1.2 DOS 下声卡的应用编程环境	(89)
4.2 Windows 环境	(97)
4.2.1 Windows 下声卡应用使用环境	(97)
4.2.2 Windows 下声卡兼容性测试环境	(97)
4.2.3 Windows 下声卡应用开发环境	(104)
4.3 声卡的硬件设计环境	(108)
第 5 章 MIDI 基础知识与合成器	(110)
5.1 MIDI 基本概念	(110)
5.2 MIDI 术语	(110)
5.3 MIDI 硬件接口与连接硬件	(111)
5.4 MIDI 规范	(113)
5.4.1 Note On —— 键按下	(113)
5.4.2 Note Off —— 键释放	(114)
5.4.3 Polyphonic Key Pressure Aftertouch —— 复音键按键力度	(115)
5.4.4 Channel Pressure Aftertouch —— 通道按键力度	(115)
5.4.5 Pitch Wheel Change —— 改变音调调节轮	(115)
5.4.6 Control Change —— 控制切换	(115)
5.4.7 Program Change —— 程序切换 (音色切换)	(117)
5.4.8 MIDI 系统指令	(118)
5.4.9 元事件 (Meta Event)	(121)
5.5 MIDI 采样数据传送标准	(123)
5.5.1 采样堆格式	(124)
5.5.2 采样数据传送	(125)
5.6 MIDI 合成器原理	(126)
5.6.1 声音分解	(127)
5.6.2 声音 (乐音) 的要素	(127)
5.6.3 合成器制作声音的方法	(128)
5.7 FM 合成芯片 —— YMF262	(130)

5.7.1 综述	(130)
5.7.2 FM 合成	(130)
5.7.3 硬件操作	(131)
5.7.4 音频输出接口	(131)
5.7.5 寄存器功能说明及用法	(132)
第 6 章 光盘存储技术与光盘驱动器	(141)
6.1 概述	(141)
6.1.1 光盘驱动器的分类	(141)
6.1.2 光盘的记录媒体	(146)
6.2 CD-ROM 的国际标准	(147)
6.3 CD 盘片和光盘驱动器	(148)
6.3.1 光盘存储器的读/写原理	(148)
6.3.2 光盘驱动器的工作原理	(152)
6.3.3 CD-ROM 的卷和文件结构	(164)
6.3.4 CD-ROM 的数据格式	(167)
6.3.5 CD-ROM 的错误校正	(168)
6.3.6 CD-ROM 驱动器的接口	(168)
6.4 VCD 播放系统及工作原理	(173)
6.4.1 VCD 结构及组成	(174)
6.4.2 VCD 的解码芯片——CI480	(176)
6.5 CD-ROM 在多媒体技术中的应用	(177)
第 7 章 音频文件格式	(179)
7.1 RIFF 文件格式	(179)
7.1.1 块(Chunk)	(179)
7.1.2 带嵌套块的“RIFF”块	(180)
7.1.3 带嵌套块的“LIST”块	(180)
7.1.4 RIFF 文件表示方法	(181)
7.2 RIFF 数据块的创建与访问	(183)
7.2.1 RIFF 数据块标准信息结构	(183)
7.2.2 创建 RIFF 型数据/信息块	(183)
7.2.3 游历 RIFF 文件	(184)
7.3 音频波形文件格式	(190)
7.3.1 波形文件 WAVE 的格式定义	(190)
7.3.2 多媒体 I/O 管理系统中波形文件标准形式	(191)
7.4 MIDI 音频文件的格式	(194)
7.5 创通音频文件格式(.VOC)	(197)
7.5.1 .VOC 文件头	(197)
7.5.2 .VOC 文件数据区	(198)
7.6 Adlib 音乐文件格式(.ROL)	(201)
7.6.1 .ROL 文件头	(201)

7.6.2	.ROL 文件数据区	(202)
7.6.3	AdLib 乐器库格式	(204)
第 8 章	声卡的应用软件	(206)
8.1	Sound Blaster 声卡应用软件	(206)
8.2	Windows Sound System 声卡应用软件	(209)
8.2.1	音乐 CD 播放器 CD player	(209)
8.2.2	媒体播放器 Media player	(211)
8.2.3	声音记录器 Sound Recorder	(213)
8.2.4	音量控制器 Volume Control	(216)
8.3	虚拟 Hi-Fi 音响 —— 音乐台 Audiostation	(217)
8.4	MIDI 作曲演奏软件 —— MIDI Session	(221)
8.4.1	MIDI Session 的组成	(221)
8.4.2	MIDI Session 的菜单 Menus	(222)
8.4.3	MIDI Session 的放音、录音功能	(228)
8.4.4	MIDI Session 的编辑功能	(232)
8.4.5	MIDI Session 的热键功能	(244)
第 9 章	声卡驱动程序开发简介	(246)
9.1	预备知识	(246)
9.1.1	一个典型的即插即用声卡	(246)
9.1.2	逻辑划分声卡各功能部件	(247)
9.1.3	驱动程序的开发平台简述	(248)
9.2	声卡驱动程序的总体结构设计	(253)
9.2.1	设计目标	(253)
9.2.2	开发环境	(254)
9.2.3	声卡驱动程序的总体结构	(255)
9.3	声卡虚拟设备驱动程序的开发	(258)
9.3.1	声卡虚拟设备驱动程序的对外接口	(259)
9.3.2	即插即用的实现	(262)
9.3.3	端口跟踪与声卡增强功能的实现	(267)
9.4	声卡驱动程序的开发	(274)
9.4.1	Windows 3.X 环境下的音频驱动消息定义	(274)
9.4.2	Windows 95 环境下的音频驱动消息定义	(276)
9.4.3	同时回放和录音	(278)
9.4.4	管道通讯	(281)
第 10 章	声卡的选购与安装	(285)
10.1	声卡的应用与选购	(285)
10.2	最新声卡介绍	(295)
10.3	SOUND BLASTER 声卡的安装	(297)
10.4	即插即用声卡及其安装	(303)
	参考资料	(306)

第 1 章 多媒体计算机及多媒体计算机音乐

1.1 多媒体计算机的基本组成

1.1.1 MPC 1.0、MPC 2.0 标准

多媒体计算机简称为 MPC，是指具有多媒体功能、符合多媒体 PC 机规范的一类 PC 机。早在 1990 年 11 月，Microsoft 公司就召集了多媒体开发人员会议，制定了 MPC 规范 1.0 版，并成立了 MPC 市场协会。1992 年 5 月又对其 1.0 版进行了修订。1993 年 5 月，多媒体微机市场协会发布了 MPC 规范 2.0 版。“MPC”和“MPC2”商标表明 PC 硬件系统符合 MPC 标准规范。MPC 1.0 规范对硬件的最低要求如表 1-1 所示。

表 1-1 MPC 1.0 规范对硬件的最低要求

组成部分	描 述
CPU	Intel 386SX 或兼容微处理器
RAM	2MB
磁盘存储	3.5 英寸的软盘驱动器，30MB 的硬盘
光盘存储	带 CD-DA 输出的 CD-ROM
音频	8 位 DAC，线性 PCM 采样，22.05kHz 和 11.025kHz 速率，DMA/FIFO 中断方式 8 位 ADC，线性 PCM 采样，11.025kHz 速率，话筒电平输入，音乐合成器板上 带有模拟音频混音功能
显示适配器	VGA
输入	101 键盘，两键鼠标
I/O	串口、并口、操纵杆端口，MIDI I/O 端口

其中 CD-ROM 具有 150KB/s 的持续传输速率，平均查找时间小于 1s，10,000 小时的 MTBF。驱动器以 150KB/s 持续传送数据时，占用 CPU 带宽应少于 40%，这要求读取数据块大小不小于 16K。CD-ROM 具有红皮书规定的 CD-DA 输出能力，具有前面板音量控制功能，CD-ROM/XA 音频可以作为选项提供。

对于 MPC 的音频能力来说，8 位（建议 16 位）的 ADC 和 DAC(Digital Analog Converter)使用线性 PCM 采样，采样率为 11.025kHz，22.05kHz 和 44.1kHz。对于数模转换器，当缓冲器空时，采样 DMA 或 FIFO 缓冲器中断传送方式输出数据，立体声可选。输出 11.025kHz 或 22.05kHz 的音频数据时，占用小于 CPU 百分之十的带宽；输出 44.1kHz 的音频数据时，占用小于 CPU 百分之十五的带宽。对于模数转换器，当缓冲器满时，采用 DMA 或 FIFO 缓冲器中断传送方式输入数据。

内部混音器组合三种（建议四种）声源，并有立体声、线路音频信号输出的能力。四种声源是 CD 红皮书数字音频、合成器、DAC（波形音频）和辅助设备输入源（建议有，但不是必须的）。每个输入源至少有 3 位（8 级）音量控制（建议有 4 位或更强的音量控制）。

多媒体 Windows 操作系统定义了 MIDI(Musical Instrument Digital Interface)文件写作标

准，支持设备无关的按标准格式创作的 MIDI 文件播放。多媒体 Windows 包含一个用户可配置的 MIDI 映射程序，它可以改变 MIDI 数据流中的 MIDI 通道和合成音色号。MIDI 制造商协会定义了合成音色和通道映射标准 GMM(General MIDI Mode)，即通用 MIDI 模式，MPC 规范中的音频卡上的合成器采用这个标准。

MPC 的显示适配器是与 VGA 兼容的显示适配器，基本的 MPC 使用显示模式的 12H (640×480、16 色)。建议用增强的显示适配器，达到 640×480、256 色。像素操作时，对于每像素 1 位、4 位、8 位的与设备无关的位图 (DIBs)。当占用 CPU 全部时间时，每秒钟可处理 350K 个像素，当占用 40% 的 CPU 时间时，每秒可处理 140K 个像素。

MPC 的系统软件平台是带多媒体扩展的 Windows 3.0 或 Windows 3.1。其中的软件开发包 SDK(Software Development Kit)和多媒体开发包 MDK(Multimedia Development Kit)提供了多媒体应用程序接口 API(Application Program Interface)、多媒体设备驱动程序和媒体控制接口。

MPC 规范 2.0 提供了硬件平台的规格，微处理器提高到 486SX，时钟频率 25MHz；RAM 为 4MB；硬盘为 160MB；显示适配器要求为 640×480、64K 彩色；CD-ROM 驱动器的持续传输速率为 300KB/s，平均存取时间为 400ms，支持 CD-ROM/XA 格式，音频的采样精度为 16 位。

1.1.2 MPC 3.0 标准

一、最小系统硬件要求

1. 中央处理器

中央处理器 CPU 必须通过一系列的多媒体计算机标准测试。测试的基准是只带一级 Cache 的 75MHz Pentium 处理器。请注意，这并不意味本规范排除其他品牌的处理器。只要这些处理器能够通过测试，则同样满足三级多媒体计算机规范。

2. 内存 RAM

具有 8MB 以上的内存 RAM

3. 软盘驱动器

3-1/2 英寸，1.44MB

4. 硬盘驱动器

(1) 最小 540MB

(2) 15ms 的平均存取时间

(3) 1.5MB 每秒的保持吞吐率

5. CD-ROM 光盘驱动器

(1) 600KB 每秒的保持传输率。

(2) 250ms 的平均存取时间，4 倍速方式。

(3) 在 600KB 每秒的保持传输率下，所占用的 CPU 带宽不大于 40%；在 300KB 每秒的保持传输率下，所占用的 CPU 带宽不大于 20%。

(4) 请求 CPU 的数据块不小于 16KB，引导时间不大于读取一个 CD-ROM 数据块的时间。

(5) 必须能够读取 CD 音乐碟，即符合红皮书的 CD 碟，格式一和格式二的格式化 CD 碟，包括混合模式碟，多块式光碟，以及 CD-ROM 数据碟，CD-ROM XA，PHOTO CD，可记录

式 CD(第二部分), 视频 VCD 和 CD-I 格式的 CD 碟。数据必须以 2048, 2336 及/或 2352 字节大小的数据块传送到主机。具体采用何种大小的数据块视不同格式的 CD 而定。另外, CD-ROM 驱动器及驱动软件必须与 MICROSOFT 公司的 MSCDEX.EXE 2.2 或者更新版本兼容, 以实现音频 API 接口, 并且可以读取 Q 通道的信息。

(6) CD-ROM 驱动器在播放 CD 音乐时, 有音频输出和音量控制功能。

(7) CD-ROM 驱动器必须具有自己的缓冲区, 而且要实现预先读缓冲。

(8) 顺序访问时间: 应用程序通过标准的操作系统访问方法, 能够读取顺序的、无错的、33.3ms 长的 16KB 数据块。而且, 每次读操作所花时间不超过 13.3ms。

(9) 后台 CPU 利用率: 驱动程序正常情况下, 不能够占用 CPU 的周期, 除非是响应主机的系统请求。

6. 音频

(1) 16 位的数模转换器(DAC), 并带有: 线性 PCM 采样, DMA 或 FIFO 缓冲输出能力, 且缓冲区空后, 具有中断能力。可以支持 44.1k, 22.05k, 11.025kHz 的采样率, 立体声。传输 22.05k 和 11.025kHz 的采样数据时, 占用 CPU 的带宽不超过 10%, 而传输 44.1kHz 的采样数据时, 占用 CPU 的带宽不超过 15%。

(2) 16 位的模数转换器(ADC), 并带有: 线性 PCM 采样, DMA 或 FIFO 缓冲输出能力, 且缓冲区满后, 具有中断能力。可以支持 44.1k, 22.05k, 11.025kHz 的采样率、立体声。传输 22.05k 和 11.025kHz 的采样数据时, 占用 CPU 的带宽不超过 10%, 而传输 44.1kHz 的采样数据时, 占用 CPU 的带宽不超过 15%。支持麦克风输入。

(3) 具有波形表合成能力。

(4) 16 位立体声和波形表合成所占用的 CPU 带宽不超过 10%。

(5) CD-ROM 驱动器可以播放音乐 CD(红皮书), 并有音频输出和音量调节。

(6) 内部合成器应具有多种声音、多种音色合成的能力。可以同时合成六种旋律加两种鼓声/ 撞击声的能力。

(7) 内部混合器具有混合三种(推荐为四种)音频输入信号的能力, 以及混合当前立体声输出信号的能力。四个输入音源为: CD 红皮书标准, 合成器, DAC(波形)以及(推荐但不要求的)辅助输入。每个输入信号至少有位(八级)音量控制, 而且是按照对数规律调整(四位或者更多位的音量控制更好)。如果所有的输入源信号电平为 10dB(消费电子标准: 1mV 的信号加载到 600 Ω 的负载上时, 定义为 0dB), 而且没有衰减, 混合器应不裁剪输入信号, 即混合器不应达到饱和, 混合器的输出介于 0dB 与 3dB 之间。强烈推荐每个输入音源有自己的音量控制寄存器, 而且还有一个主数字音量控制寄存器, 也推荐支持更多的音源输入。

(8) 扬声器。如果系统配有外部扬声器, 则对扬声器的要求为:

A. 至少为双音箱系统。

B. 频率响应范围为 120Hz-17.5kHz。

C. 最大功率必须标定。而且测试时, 输入信号为 100Hz, 1kHz 和 10kHz。单通道输入功率至少为 3W, 此时, 失真不超过 1%。1kHz 的信号以 6W 的平均功率(3W+3W)加载到 4 Ω 的负载上时, 失真也不超过 1%。

D. 双通道驱动方式。

(9) 在双扬声器的对称轴上 0.5m 处, 所测到的声压级从 250Hz 到 7.5kHz, 应有 92dB。

(10) 输入连接器规定如下: 3.5mm 立体声耳机插头, 顶点为左声道, 中间为右声道, 载

体为地。或者输入用 3.5mm 工业标准的立体声耳机插头，至少有六条腿的引脚，支持三对立体声信号的输入。

(11) 扬声器必须采用单声道耳机插头。顶点为正，载体为地。或者左扬声器输出 3.5mm 单声道插头，其顶点为正，音箱的箱体为地。如果还有立体声耳机输出插头，则当使用耳机时，扬声器必须被静音。

(12) 信噪比至少为 65dB。

(13) 在适当的输出功率下，输入灵敏度不小于 300mV。

(14) 必须有音量控制，高音控制和低音控制。

(15) 输入阻抗必须大于 5000 Ω 。

(16) 在三音箱系统中，对卫星音箱的要求同上。

(17) 对低音箱的要求：

A. 频率响应至少为 40Hz-250Hz(+/-3dB)

B. 功率至少为 15W，失真度小于 1%

C. 功率测量频率为 40Hz 和 100Hz 的信号，失真度小于 1%

D. 为了得到最大的输出功率，输入灵敏度应可调整，范围为 300mV 到 1V

E. 左右声道的输入均应混合到低音箱的电路中

F. 输入阻抗必须大于 1000 Ω

7. 图形性能

(1) 要求具有彩色空间变换能力和标定能力。

(2) 视频子系统被允许的情况下，要求具有直接访问帧缓冲区的能力。帧的分辨率为 352(288×30)帧每秒，15 位的像素。物缩放，无裁减。

(3) 必须设定一个测试集，以测试系统的图形性能。图形性能的测试结果应可接受。

8. 视频回放

(1) 要求具有兼容 MPEG1 的能力。

(2) 要求能够直接访问帧缓冲区。帧的分辨率为 352(288×30)帧每秒。15 位的像素。物缩放，不丢失任何一帧。

(3) 回放性能必须经过测试集的测试。回放效果应可以接受。

9. 用户输入

(1) 标准的 IBM 式 101 键盘和 DIN 连接器，或采用键组合方式而有同样功能的键盘。

(2) 带有总线式或者串行式的双键鼠标。且系统至少有一个通讯口可以使用。

10. I/O 端口

(1) 标准的 9 针或者 25 针异步串行口，可以编程到 57.6K 波特率。可切换的串行口中断号。

(2) 标准的 25 针，具有中断能力的并行双向端口。

(3) 一个具有输入，输出，直通能力的 MIDI 端口。且支持中断方式，支持输入与 FIFO 缓冲方式传输。

(4) IBM 式的模拟或者数字游戏操纵杆接口。

二、系统软件

(1) 多媒体计算机的系统软件必须与 Windows 3.11 在二进制级上兼容。

(2) 系统必须提供在二进制级上的与 DOS 6.0 或者更高版本的 DOS 兼容。

MPC 作为一种技术规格, 指导技术开发人员设计多媒体 PC 机; MPC 作为一种 PC 机和升级套件的商标, 表明它们符合 MPC 标准。

从 MPC 规范中看出, 除普通 PC 系统外, 其基本系统的配置主要有 CD-ROM 驱动器, 它为台式机带来了价格便宜的 600 多兆字节的只读存储设备, 存有图形、动画、图像、声音、文字、数字音频、程序、资源的 CD 已广泛使用, 因此 CD-ROM 对 MPC 来说是必须的。数字音频处理的支持也是需要的。音频卡具有 A/D 和 D/A 音频转换的能力, 并且带有内部合成器、混音器和 MIDI 接口, 可以合成音乐、混合多种声源、外接 MIDI 电子音乐设备。图文并茂的多媒体表现需要分辨率高、同屏显示色彩丰富的显示卡支持, 并具有 Windows 的显示驱动程序, 在 Windows 下的像素运算速度要快。带有图形用户接口 GUI(Graph User Interface) 加速器的局部总线显示适配器使 Windows 的显示速度大大加快。

MPC 规范中没有引入视频板的标准。一方面是因为具有实用价值的动态视频板价格昂贵, 它适合作为一个选项引入到 MPC 系统中, 而不是基本(最低)配置。另一方面, 在 PC 适配器支持下, 可以表现各种动画和影片媒体。

MPC 必须要有很高性能的硬件支持, 否则在存储容量、音质和视觉效果以及实时性等方面就达不到较高的要求。音频和视频的专用处理芯片以及新型的体系结构是 MPC 硬件支持的关键。多媒体计算机系统的体系结构只要考虑以下内容: (1)总线带宽: 设计专有的视频总线, 增加系统总线带宽, 达到多路动态数字音频视频设计传输的要求; (2)接口: 连接 CD-ROM 等大容量存储设备的 SCSI 接口, 具有各种音频、视频输入输出端口, 容易扩展宽带通信网络接口和各种交互设备(如触摸屏、操纵杆、数据手套等)接口; (3)集成: 随着大规模集成电路和专有 ASIC 芯片设计的发展, 视频和音频处理子系统将集成到计算机的主板上。视频音频处理芯片将是大部分计算机的基本配置。

1.2 数字音频技术

一、听觉类媒体

凡是通过声音形式以听觉传递信息的媒体, 都属于听觉类媒体, 主要有以下几类:

1. 波形声音

所谓波形声音, 实际上已经包含了所有的声音形式, 因为可以把任何声音都按采样量化, 并恰当恢复出来。事实上, 波形声音就是自然界中所有声音的“第一印象”, 或称为“数字化拷贝”。

2. 语音

因为人的说话声不仅是一种波形, 而且还具有内在的语言、语音学内涵, 可以经由特殊的方法而提取, 即进行一次抽象。所以常把它作为一种特殊的媒体。

3. 音乐

音乐与语音相比, 形式更为规范。事实上, 音乐就是符号化了的声音, 这种符号就是乐曲, 乐谱则是转变为符号媒体形式的声音。就计算机媒体而言, MIDI 是十分规范的一种形式。

二、听觉特性

首先，从听觉角度来讨论声音媒体的有关特性、声学原理以及质量标准。

1. 声音三要素

声音三要素即为音调、音强、音色。音调与声音的频率有关，频率快则声音高，频率慢则声音低。人的听觉范围是 20Hz~20kHz。音强又称为响度，取决于声音的幅度，亦即振幅的大小和强弱。而音色则是由混入基音的泛音所决定的。每个基音又都有其固有的频率和不同音强的泛音，从而使得每种声音具有特殊的音色效果。

2. 声音的连续谱

声音是一种弹性波，声音信号可以分成周期信号与非周期信号两类。周期信号即是单一频率音调的信号，其频谱是线性谱；而非周期信号包含一定频带的所有频率分量，其频谱是连续谱。真正的线性谱仅可从计算机或类似的专门声音设备中才能产生和听到。这种声音听起来十分单调。而其它声音信号或者属于完全的连续谱，如电路中的平滑噪声，听起来完全无音调，或者属于线性谱中混有一段段的连续谱成分，只不过这些一段段的连续谱成分比起那些线性谱成分来说要弱，以致使整个声音还是表现出线性谱的有调特性。也正是这些连续谱成分使声音听起来饱满、生动。自然界的声音大多属于这一种。

3. 声音的方向感

声音的传播是以声波形式进行的。由于人类的耳朵能够判别出声波到达左右耳的相对时差、声音强度，所以能够判别出声音的来源方向。同时也由于空间作用使声音来回反射，而造成声音的特殊的空间效果。例如我们在剧场中聆听到的声音和在公园中聆听到的声音效果是不一样的。因此，现在的音响设备等都在竭力模拟这种立体声效果和空间感效果。在虚拟现实环境中会需要这种声音空间效果，所以作为媒体应能描述这种声音空间要求。

4. 声音的质量

声音的质量与声音的频率范围有关，一般来说，频率范围越宽，声音的质量也就越高。衡量声音的质量有一个标准。对语音来说，常用可懂度、清晰度、自然度来衡量；而对音乐来说，保真度、空间感、音响效果都是重要的指标。

5. 声音是时基类媒体

可以这样说，没有时间就没有声音。声音具有过程性，适合在一个时间段中表现。声音常常处于一种伴随状态，如伴音、伴奏等，起一种气氛渲染、感觉相乘的作用。由于时间性，声音数据具有很强的前后相关性，数据量要大得多，实时性要求也比较高。由于声音是连续的，所以又称之为连续型时基媒体类型。

三、波形声音处理的方法

1. 采样量化方法

波形声音就是对外界声音进行采样并量化的结果。根据奈奎斯特采样定律，只要采样频率高于信号最高频率的两倍，就可以完全从采样中恢复原始信号波形。这就是在实际的采样过程中，采用 40.1kHz 作为高质量声音标准的原因。

对声音波形的采样，就是按采样的频率间隔、不断地获取幅度的量值，使离散的声音波形转变为离散的数字量。当需要时，可以再将这些离散的数字量转变成为连续的波形。如果采样频率足够高，恢复出的声音应与原始声音没有什么差别。无论什么样的声音，音乐还是

语音，都能按波形声音采样、存储并且再现。

2. 有关的技术参数

(1) 频率

采样的频率等于波形被等分的份数。频率越高，声音质量就越接近原始声音，所需的存储量也越大。标准的采样频率有三个，即 44.1kHz、22.05kHz 和 11.025kHz。

(2) 信息度

每个采样点存放的信息数量是指采样点测量的精度。采样的信息量是通过将每个波形采样垂直等分而形成的，8 位采样指的是将采样幅度划分为 256 等分，16 位采样就可以划分为 65536 等分。很显然，用来描述波形特性的垂直单位数量越多，采样越接近原始的模拟波形，存储量也就要求越大。波形的基线与波形上限间的单位被称为波形的动态范围。8 位采样将波形分为 256 个单位，该波形的动态范围一定覆盖所有的 256 个单位的大部分。如果波形的动态范围只覆盖 128 个单位，则说明声音的精确度（质量）降低了，因为每个采样只使用了 7 位。

(3) 声道数

使用通道的个数表明声音记录是只产生一个波形(单声道)还是产生两个波形(立体声双声道)。立体声听起来比单声道的声音丰满且有一定的空间感，但需要两倍的存储空间。

(4) 数据量

无论质量如何，声音的数据量都非常大。如果不经压缩，声音的数据量是巨大的，其值可由下式推算：

数据量 = (采样频率 × 每个采样位数 × 声道数) / 8 (B/s)

例如，一个小时的 HIFI 立体声音乐要占有 6.35×10^8 B，即 635MB。

3. 波形声音的采集、处理与输出

波形声音的获取是通过声音数字化接口进行的，输入的声音经数字化后存入计算机中，在需要时，再将其恢复成原始波形输出。

对于声音的处理主要集中在压缩、编辑和效果处理上。压缩常在硬件或低层软件中完成，以求降低数据量。对声音的编辑常常是进行分段、组合、首尾处理等，以求单一的声音片段能以干净、准确的形式出现。效果处理也常常放在编辑操作中，常用的处理有回声处理、倒叙处理、音色效果处理等。一般情况下，用户更关心录入声音的效果。所以常常在录入时用最高频率采样，在存储时则通过计算机进行变换后转为较低频率的声音数据文件。这样做的效果要比直接用低频率采样好得多。

四、数字音频技术的基础

数字音频技术的基础是 PCM(Pulse Code Modulation) (脉冲代码调制)。PCM 的基本工作过程为：

(1) 把原音频信号用一固定频率来取样，如果取样频率选取恰当的话，则原信号可用取样值来加以准确地恢复。

(2) 把每一取样值转换成二进制数。此时的分辨率，以二进制数的位数来决定，这个位数叫做量化比特数，通常为 14~16b。

(3) 表示各取样值的二进制数，构成一串“1”和“0”的比特数列。这样就完成了 PCM 调制。

所得的 PCM 信号具有三个主要优点:

(1) 对一定电平以下的噪声可以完全抑制而还原出原来的信号。从理论上说, 对噪声的抑制能力很强。即使某些噪声使得不能再现原来的 PCM 序列, 也可以用纠错码来加以解决。所以, 一个 PCM 系统总是能采用这样或那样的方法, 既能保证正常工作, 又可避免由于噪声引起的信号劣化。

(2) 利用时间分享法把几个不同的信号编成一系列编码信号, 则可实现时间间隔多路传输制。

(3) 上述量化系统所用的量化比特数决定了还原音质的信噪比(S/N)。如果量化比特数较大, 就可得到比其它调制方式高的信噪比。

1.3 器乐声的计算机合成及演奏

由于音乐是完全可用符号来表示的, 所以音乐可看作是符号化的声音媒体。有相对音乐符号化的形式, 其中应用最广泛的是 MIDI。

MIDI 是指乐器数字接口, 是数字音乐的国际标准。任何电子乐器, 只要有处理 MIDI 消息的微处理器, 并有合适的硬件接口, 都可以成为一个 MIDI 设备。MIDI 消息, 实际上就是乐谱的数字描述。在这里, 乐谱完全由音符序列、定时以及被称为合成音色的乐器定义组成。

标准的 MPC 平台通过内部合成器 (Synthesizer) 或通过外接到计算机 MIDI 的外部合成器播放 MIDI 文件。MIDI 合成器的类型目前有两种: 频率调制 FM 合成和波形表 (Wave Table) 合成。FM 方式最早应用于音频卡上, 现在使用得较为普遍。利用波形表方式合成音乐的效果更逼真。因为在各种合成方式中, 每种乐器对应一种波形或多种波形。合成音乐时以查表方式获取乐器波形, 通过声音生成器和扬声器输出。这里主要介绍 FM 合成器。

音频卡利用频率调制产生各种乐器的音色。一种声音的音色在很大程度上依赖于其谐波的频率和振幅。FM 合成器对声音的音色提供简捷而有效的控制。

频率调制 FM 合成器包含 18 个操作单元。调频合成需要两个操作单元: 调制单元和载波单元。

每个单元有三个功能模块: 脉冲产生器、包络产生器和正弦表。每个单元具有生成一个正弦波的能力。正弦波的频率依赖于脉冲产生器, 振幅依赖于包络产生器, 如图 1-1 所示:

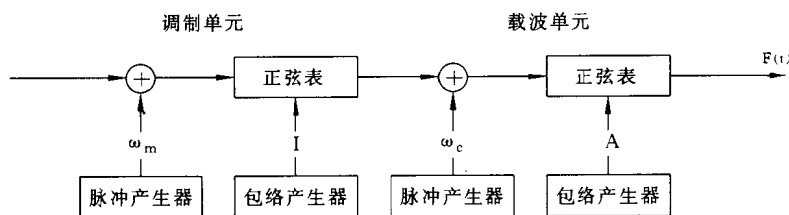


图 1-1 FM 合成器

调制单元的输出用于调制载波单元。载波单元的结果输出波形包含谐波, 其频率等于载波频率与调制频率各种倍数的和值和差值。谐波的振幅依赖于调制波形的振幅。通过改变调制单元的频率和振幅, 就可以改变 FM 合成器产生的声音音质。

以下公式表示 FM 声音产生与图 1-1 中四个参数的关系:

$$F(t) = A \sin(\omega_c t + I \sin \omega_m t)$$

这里, A 是输出信号振幅, I 是调制单元输出信号的振幅。 ω_c 和 ω_m 分别是载波信号频率和调制信号的频率。

18 个操作单元同时可以产生 9 个通道的 FM 声音, 每个 FM 声音使用 2 个操作单元, 或用操作单元 1-12 产生 6 个通道的 FM 声音, 其它用于产生打击乐器声音。也可以让每个操作单元都产生一种正弦波, 多个不同振幅和周期的正弦波合成语音。

另外, FM 合成器中的操作单元包含内置的颤音振荡器和振幅调制振荡器, 分别产生颤音(vibrato)效果和震音(tremolo)效果。这些效果用于产生更接近于自然乐器的声音。

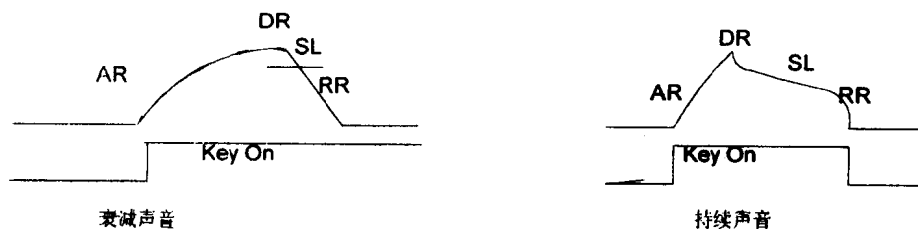


图 1-2 波形包络

图 1-2 示出合成器使用的两种包络类型。在图中, AR(Attack Rate)为上升率, 表示声音波形前沿上升时间; DR(Decay Rate) 为衰减率, 表示上升之后的衰减时间; SL(Sustain Level) 为持续电平, 是声音持续电平, 对于持续声音, 它表示释放模式前的一段不变的播放电平, 对于衰减声音, 它表示衰减模式与释放模式的交汇点。对于持续声音, 释放率 RR(Release Rate) 定义键释放(Key-off) 后声音的消失时间, 对于衰减声音, 释放率表示持续电平点之后的衰减时间。

1.4 音乐效果处理以及环境效果处理

应用残响和混响(重复的回声但振幅渐小)信号的不同组合可令大脑错误地判断周围环境。例如: 残响时间为 2s, 回声为 50ms, 大脑可解释为置身于一个 15m 长的山洞中或类似的坚硬表面的建筑物中; 如残响的时间为 150ms, 大脑则感觉到此环境为有柔软吸音材料布置的空间。因此, 利用延时时间不同可令大脑对周围环境产生错觉, 从而制造出逼真的临场效果、广阔空间及特殊的音响效果。

混响是声音在时间和空间上的延续和扩展, 除了能增加声音的丰满浑厚的感觉外, 还能传递环境空间尺寸的信息, 通过反射声波的时序和方位给人以立体感觉。因此, 混响是立体声效果不可缺少的重要组成部分。

环绕立体声是指与普通的两声道平面立体声相比而言的, 这种音响的听音效果是好像把听音者用声音包围起来一样。这种放音方式所产生的重放声场除了保留原信号的声源方向感