

DV 高手

金鼎图书工作室 编著

提供从拍摄—编辑—输出—烧录成碟的全流程DV解决方案！

DV(数字视频)专业术语和原理解析，为你开启一个全新的视听境界
数码摄像机、数码相机、摄像头…… 数字视频硬件全搜罗。
用会声会影和Movie Maker，轻松制作个性十足的家庭电影和个人影
Nero Burning ROM与Easy CD Creator的实战演练。



四川电子音像出版中心

第一章 数字影音基础知识

1.1 关于数字视频

1.1.1 视频原理

所谓视频，就是指构成电影和电视的动态影像。电影和电视相同，都是利用人眼对运动物体产生视觉残像这一生理现象，在一秒钟内快速播放 24 或 30 个静态画面，由这些快速播放的静态画面在人的视觉神经中形成动态画面，这其中每一个静态的画面便是在视频领域中经常提及的“帧”。下面的两幅画面就是从同一个视频档案素材中采集的两帧。

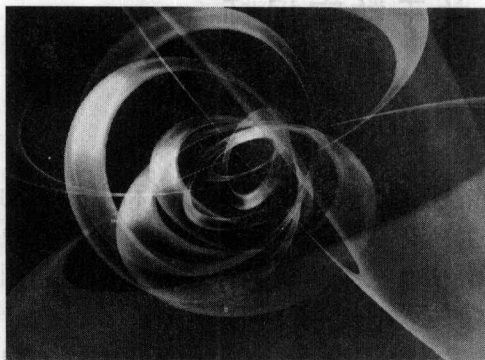
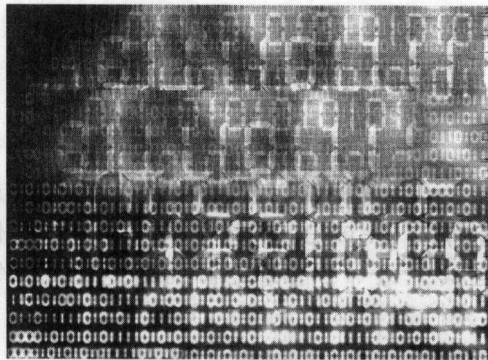


1.1.2 数字视频的概念

所谓数字视频，就是视频在产生、传播、存储的过程中以数字形式存在的全新的视频方

式。数字视频把视频帧的每一个像素以不连续的颜色数值表示，并且按照电脑使用的二进制文件格式进行传送和储存。

模拟视频通过在视频讯号上建立变化来支持图片和声音信息。这类讯号用不同的电压数值表示信息，例如：0V 表示黑，0.7V 表示白，其它灰阶介于二者之间。模拟讯号是连续变化的。



通常有两种方法可以将模拟视频信号传送到计算机中（也就是数字化）：

- （1）使用影片扫描仪把每个模拟影片帧直接转换为数字影片帧。
- （2）把影片转成录像带，然后数字化这些录像带。

可以看出，视频构成的不同并不影响数字视频与模拟视频相互间的交流。将一段模拟视频讯号转换为数字视频文件：首先要透过视频采集卡将一系列的电信号转换成多个 0 与 1 的组合，建立一个视频档案并存储在计算机中。然后，再由专门的数字视频处理软件对视频文件进行必要的数字化处理，最后呈现在你面前的就是一个完整的数字视频成品文件。由数字视频到模拟视频的转换过程就是上面所描述过程的逆过程。

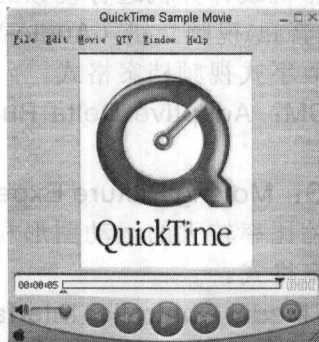
1.1.3 数字视频的应用

依据所涉及领域的不同,数字视频的应用主要分为三大类:专业数字视频处理领域、多媒体数字处理领域、Internet 数字处理领域。

专业数字处理领域要求达到 5M/秒连续读取数据量的能力。因此,这一领域所运用的视频制作设备也是具有专业水准的,包括:更高速的 CPU,更多的内存,更高容量的数据硬盘,专业用的高阶视频采集卡及高容量的影音硬盘(AVHD)。

多媒体数字视频处理领域要求达到 630KB/秒的连续读取资料量的能力,现代的个人计算机基本上都能够达到这种要求;也就是说,现阶段的个人电脑都具备了多媒体级数位视频应用的能力。

5KB/秒读取数据量的能力是 Internet 数字视频应用的基本要求。Internet 视频应用实际上是将视频流量进行愈小化处理,以符合网络带宽限制及网络数据流读取方式的要求。前面我们曾经提及 Internet 采用的 Streaming 视频压缩方式,它的处理过程是先分割视频数据为一节一节的(封包),然后再分段传送;当传送至客户浏览端后,由所设定的特殊的播放方式将它们立即播放,边传边播。几种重要的 Streaming 视频压缩方式包括:Real Networks 的 Real Video, Microsoft 的 Net show 以及 Apple 的 Quick Time。



1.1.4 专业名词解释

- ◆ **RGB**: 是由红、绿、蓝三种原色构成的色彩模式。
- ◆ **像素**: 像素是构成图形的基本元素, 它是位图图形的最小单位。
- ◆ **分辨率**: 分辨率是指图像单位面积内像素的多少。
- ◆ **Alpha 通道**: 一些影像或视频档案的一个灰阶图层, 它被用来将影像或视频分成几个部分, 以便进行编辑或增强。在数字视频编辑中, **Alpha** 信道主要被用于覆迭和镂空罩。
- ◆ **Direct Show**: 它的前身为 **Microsoft** 的 **Video For Windows** 和 **Active Movie**。 **Direct Show** 能够以网络上的数据流格式来采集视频和音讯文件, 支持许多视频/音讯格式, 包括 **AVI**、**DV AVI**、**MPEG**、**Apple Quick Time**、**WAV** 等等。
- ◆ **IEEE 1394**: 此标准可以为 **DV** 摄像机、录像机或任何一种数字音讯/视频装置提供与电脑的高速串行连接。遵循此标准的装置, 可以以每秒至少 **100MB** 的速度来传送数字数据。
- ◆ **DV**: 数字式视频 (**Digital Video**) 代表了一种特别的视频格式, 就和 **VHS** 或 **High.8** 一样。只要你拥有适当的硬件 (采集卡) 和软件 (**DV codec**), 此格式就可透过摄录放影机或计算机进行视频的播放或录制。
- ◆ **AVI**: 音讯\视频交错 (**Audio-Video Interleave**) 是一种专门为微软 **Windows** 环境设计的数字式视频档案格式。
- ◆ **ADPCM**: **Adaptive Delta Pulse Code Modulation** 是用于压缩音效文件的一种压缩方式。
- ◆ **MPEG**: **Moving Picture Experts Group** 的简称, 运动图像专家组, 该组织定义了一种压缩比率较大的运动图形和声音的压缩标准。通常也指由此组织定义的压缩标准的档案格式。
- ◆ **MCI**: **Media Controller Interface** 的简称, 是由微软公司设计的一种软件驱动程序以开启音讯和视频档案能够在 **windows** 中播放。

◆ SMPTE: Society of Motion Picture and Television Engineers 的简称。是在录相磁带上识别帧信号的一种标准,用于对视频影像的编辑控制。SMPTE 以 Hour: Minute: Second: Frame 的形成确定帧的位置。

1.2 数字视频的编辑合成

传统的电视在编辑时应用的是 A/B ROLL 方式,它要用到两台放影机(A和B),一台录像机和一台转换机(Switcher)。A和B放影机中的录像带上储存有已经采集好的视频片段,这些片段的每一帧都有时间码。如果现在要把A带上的a视频片段和B带上的b视频片段连接在一起,就必须先设定好a片段要从哪一帧开始,到哪一帧结束,即确定好“In点”和“Out点”。同样,对于b片段也要设定好相应的“开始”和“结束”点。当将两个视频片段连接在一起时,可以使用转换机来设定转场效果,当然也可以透过它来制作更多的特效。

线性编辑

线性编辑又称作在线编辑,前面介绍的传统电视编辑就属于这类编辑。它的特点是直接用母带来剪辑。经常会遇到这种情况,在家中用录像机欣赏影片的时候,如果有一段是你非常喜欢的,想多看几遍,就必须不断地快进或倒退才能找到想看的那一段。传统的线性编辑也是如此:如果在编辑好的录像带上要插入或删除视频片段,那么在插入点或删除点以后的所有视频片段都要重新剪辑一次。

非线性编辑(数字编辑)

相对于线性编辑,我们在计算机中进行的视频编辑就是非线性编辑。要进行非线性编辑,首先必须把录像带上的节目片段采集下来,即把模拟讯号转换为数字讯号,然后存储到硬盘中。这样,在需要调整某段数字视频的位置或时间,只要在计算机上使用鼠标和键盘就可以完成视频编辑的操作。

关于视频合成

素材以数字视频文件的形式存储于计算机中，并不代表它就是我们所需要的数字视频成品文件；我们还要利用一些硬件与软件资源，对素材进行处理，使视频文件能够以我们所希望的效果呈现画面。

视频合成依靠的是数字视频系统的支持，它当然包括硬件和软件两方面：

1、硬件方面主要指数字视频系统的系统硬件架构，包括：

计算机系统（含计算机主机、屏幕、键盘、鼠标、光驱）

视频采集卡（含视频采集及转录功能）

AV 硬盘（影音硬盘）

声卡

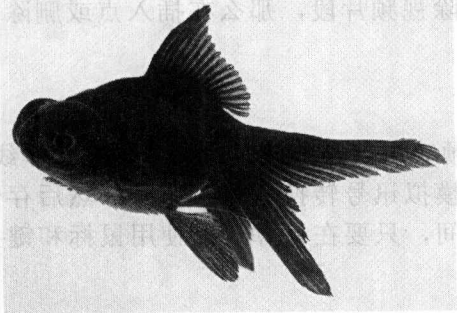
其他辅助设备（监视器、Beta CAM、图形加速卡及摄影机等）

2、软件方面主要是指数字视频系统的系统软件架构，包括：

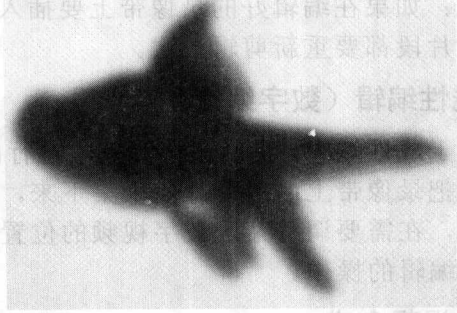
操作系统（本书以 Windows XP 系统进行编辑操作）

视频编辑软件（例如本书所介绍的 Moviemaker、会声会影等）

数字合成实际上就是指对众多不同的数字视频素材的艺术性组合和加工。下面的图像效果就是通过数字化合成制作出来的。



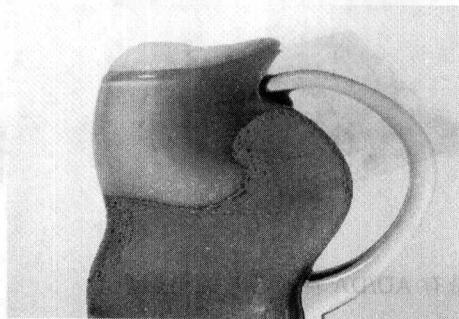
原图



合成图



原图



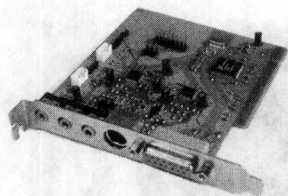
合成图

1.3 关于数字音频

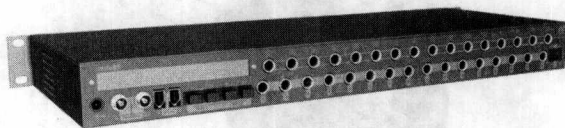
1.3.1 数字音频原理

大家知道，无论现在的多媒体电脑功能如何强大，其内部也只能处理数字信息。而我们听到的声音都是模拟信号，怎样才能让电脑也能处理这些声音数据呢？还有，究竟模拟音频与数字音频有什么不同呢？数字音频究竟有些什么优点呢？这些都是我们下面所要介绍的。

把模拟音频转成数字音频，在电脑音乐里就称作采样，其过程所用到的主要硬件设备便是模拟/数字转换器（Analog to Digital Converter，即 ADC），电脑声卡中集成了模拟/数字转换芯片，其功能相当于模拟/数字转换器。

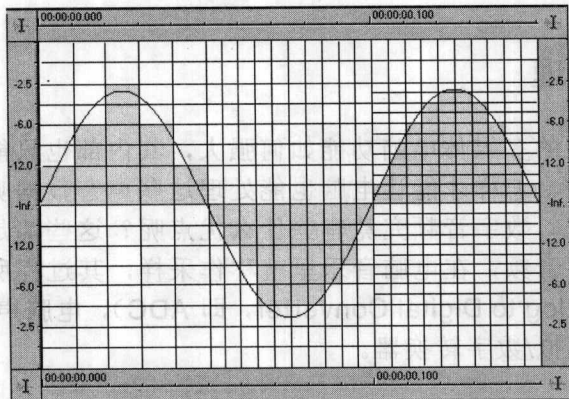


带有 AD/DA 转换芯片的电脑声卡



专业级的 AD/DA 转换器

采样的过程实际上是将通常的模拟音频信号的电信号转换成许多称作“比特 (Bit)”的二进制码 0 和 1，这些 0 和 1 便构成了数字音频文件。如下图，图中的正弦曲线代表原始音频曲线；填了颜色的方格代表采样后得到的结果，二者越吻合说明采样结果越好。

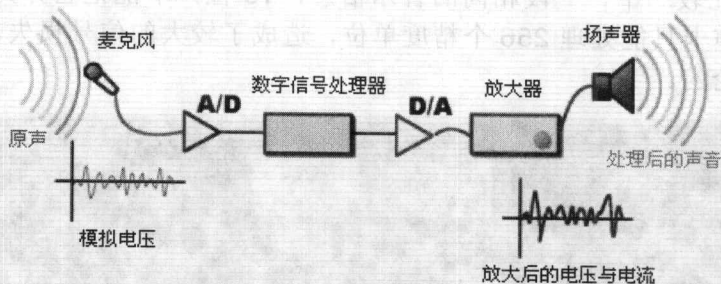


上图中的横坐标便是采样频率；纵坐标便是采样分辨率。图中的格子从左到右，逐渐加密，先是加大横坐标的密度，然后加大纵坐标的密度。显然，当横坐标的单位越小即两个采样时刻的间隔越小，则越有利于保持原始声音的真实情况，换句话说，采样的频率越大则音

质越有保证；同理，当纵坐标的单位越小则越有利于音质的提高。

现在我们来进行一个计算，看看一个数字音频文件的数据量到底有多大。假设我们采用 44.1kHz、16bit 来进行立体声（即两个声道）采样，即采样成标准的 CD 音质（也称作红皮书音频）。那么就是说，一秒钟内采样 44.1 千次，每次的数据量是 $16 \times 2 = 32\text{bit}$ （因为立体声是两个声道）。而大家知道，一个字节（Byte）含有 8 个位（Bit），那么一秒钟内的数据量便是 $44.1\text{k} \times 32\text{bit} / (8\text{bit} / \text{Byte}) = 176.4\text{kByte}$ 。一个汉字在电脑里占用两个字节，那么 176.4kB 的空间可以存储 $176.4\text{k} / 2 = 88200$ 个汉字，也就是说一秒钟的数字音频数据量与近九万个汉字（一部中篇小说）的数据量相当。由此可见，数字音频文件的数据量是十分庞大的。

那么数字音频又是如何播放出来的呢？首先，将这些由大量数字描述而成的音乐送到一个叫做数/模转换器（Digital to Analog Converter，即 DAC）的线路里。它将数字还原成一系列相应的电压值，然后通过有助于稳定的保持线路，最后将信号由低通滤波器输出。这样，比较平缓的具有脉动电压的模拟信号可继续发送至放大器和扬声器，电流经过放大再转变成声音。



模拟音频又是怎样进行录制与播放的呢？首先，声波通过麦克风，空气分子的振动转变为电信号的波动（数字录音也必需经过这一步）。录音磁头的电磁铁根据通过电流的大小而产生大小不同的磁场，磁场的变化情况会相应的记录在磁带上（实际上是磁带上的磁粉排列发

生了变化), 这样便完成录音过程。播放时, 放音磁头读出印记在磁带上的磁场大小变化的情况(即磁粉的排列位置), 并转变为相应的电信号。之后的情况与数字音频的播放完全类似, 即这些波动的电信号(模拟信号)继续传送至放大器和扬声器, 电信号重新转变为声音(即空气分子的振动)。

1.3.2 数字音频术语

采样的位数

采样位数可以理解为声卡处理声音的解析度。这个数值越大解析度就越高, 录制和回放的声音就越真实。

电脑中的声音文件是用数字 0 和 1 来表示的。所以在电脑上录音的本质就是把模拟声音信号转换成数字信号。反之, 在播放时则是把数字信号还原成模拟声音信号输出。声卡的位是指声卡在采集和播放声音文件时所使用数字声音信号的二进制位数。声卡的位客观地反映了数字声音信号对输入声音信号描述的准确程度。8 位代表 2 的 8 次方--256, 16 位则代表 2 的 16 次方--64K。比较一下, 一段相同的音乐信息, 16 位声卡能把它分为 64K 个精度单位进行处理, 而 8 位声卡只能处理 256 个精度单位, 造成了较大的信号损失, 最终的采样效果自然是无法相提并论的。



数字音频采样器

采样的频率

采样频率是指录音设备在一秒钟内对声音信号的采样次数，采样频率越高声音的还原就越真实越自然。在当今的主流声卡上，采样频率一般共分为 22.05KHz、44.1KHz、48KHz 三个等级，22.05 只能达到 FM 广播的声音品质，44.1KHz 则是理论上的 CD 音质界限，48KHz 则更加精确一些。

单声道

单声道是比较原始的声音复制形式，早期的声卡采用的比较普遍。当通过两个扬声器回放单声道信息的时候，我们可以明显感觉到声音是从两个音箱中间传递到我们耳朵里的。这种缺乏位置感的录制方式用现在的眼光看自然是很落后的，但在声卡刚刚起步时，已经是非常先进的技术了。

立体声

单声道缺乏对声音的位置定位，而立体声技术则彻底改变了这一状况。声音在录制过程中被分配到两个独立的声道，从而达到了很好的声音定位效果。这种技术在音乐欣赏中显得尤为有用，听众可以清晰地分辨出各种乐器来自的方向，从而使音乐更富想象力，更加接近于临场感受。立体声技术广泛运用于自 Sound Blaster Pro 以后的大量声卡，成为了影响深远的一个音频标准。时至今日，立体声依然是许多产品遵循的技术标准。



电脑音箱

四声道环绕

三维音效的主旨是为人们带来一个虚拟的声音环境，通过特殊的 **HRTF** 技术营造一个趋于真实的声场，从而获得更好的游戏听觉效果和声场定位。而要达到好的效果，仅仅依靠两个音箱是远远不够的，所以立体声技术在三维音效面前就显得捉襟见肘了，新的四声道环绕音频技术则很好的解决了这一问题。

四声道环绕规定了 4 个发音点：前左、前右，后左、后右，听众则被包围在这中间。同时还增加了一个低音音箱，以加强对低频信号的回放处理（这也就是如今 **4.1** 声道音箱系统广泛流行的原因）。就整体效果而言，四声道系统可以为听众带来来自多个不同方向的声音环绕，可以获得身临各种不同环境的听觉感受，给用户以全新的体验。如今四声道技术已经广泛融入于各类中高档声卡的设计中，成为未来发展的主流趋势。

5.1 声道

5.1 声道已广泛运用于各类传统影院和家庭影院中，一些比较知名的声音录制压缩格式，譬如杜比 **AC-3 (Dolby Digital)**、**DTS** 等都是以 5.1 声音系统为技术蓝本的。其实 5.1 声音系统来源于 4.1 环绕，不同之处在于它增加了一个中置单元。这个中置单元负责传送低于 80Hz 的声音信号，在欣赏影片时有利于加强人声，把对话集中在整个声场的中部，以增加整体效果。

3D 音频

1、Direct Sound 3Dg —— 源自于 Microsoft DirectX 的老牌音频 API。

对不能支持 **DS3D** 的声卡，它的作用是一个需要占用 CPU 的三维音效 **HRTF** 算法，使这些早期产品拥有处理三维音效的能力。但是从实际效果和执行效率看都不能令人满意。所以，此后推出的声卡都拥有了一个所谓的“硬件支持 **DS3D**”能力。**DS3D** 在这类声卡上就成为了 **API** 接口，其实际听觉效果则要看声卡自身采用的 **HRTF** 算法能力的强弱。

2、A3D —— 美国 Aureal 公司所开发，分为 1.0 和 2.0。

1.0 版包括 **A3D Surround** 和 **A3D Interactive** 两大领域，特别强调在立体声硬件环境下

就可以得到真实的声场模拟。2.0 则是在 1.0 基础上加入了声波追踪技术，进一步加强了 3D 音频性能，它是当今定位效果最好的 3D 音频技术。

3、EAX —— “环境音效扩展集”

EAX 是建立在 DS3D 上的，只是在后者的基础上增加了几种独有的声音效果指令。EAX 特点是着重对各种声音在不同环境条件下变化和表现进行渲染，但对声音的定位能力不如 A3D。

MIDI

MIDI 是 Musical Instrument Digital Interface 的简称，意为音乐设备数字接口。它是一种电子乐器之间以及电子乐器与电脑之间的统一交流协议。我们可以从广义上将其理解为电子合成器、电脑音乐的统称，包括协议、设备等等相关的含义。

时下在一些游戏软件和娱乐软件中，我们经常可以发现很多以 MID、RMI 为扩展名的音乐文件，这些就是在电脑上最为常用的 MIDI 格式。MIDI 文件是一种描述性的“音乐语言”，它将所要演奏的乐曲信息用字节表述下来。譬如在某一时刻，使用什么乐器，以什么音符开始，以什么音调结束，加以什么伴奏等等，所以 MIDI 文件非常小巧。

波表合成

波表的英文名称为“WAVE TABLE”，从字面翻译就是“波形表格”的意思。其实它是将各种真实乐器所能发出的所有声音（包括各个音域、声调）录制下来，存贮为一个波表文件。播放时，根据 MIDI 文件纪录的乐曲信息向波表发出指令，从“表格”中逐一找出对应的声音信息，经过合成、加工后回放出来。由于它采用的是真实乐器的采样，所以效果自然要好于 FM。一般波表的乐器声音信息都以 44.1KHz、16Bit 的精度录制，以达到最真实的效果。理论上，波表容量越大合成效果越好。

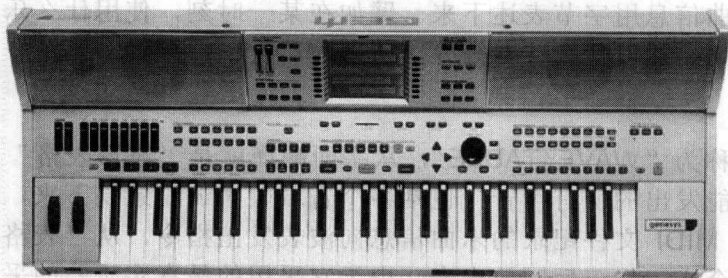
DLS 技术

PCI 声卡的问世和普及带来了波表合成的一次小小的“革命”，其关键在于 DLS 技术的运用。DLS 全称为“Down Loadable Sample”，意为：可供下载的采样音色库。其原理与软

波表颇有异曲同工之处，也是将音色库存贮在硬盘中，待播放时调入系统内存。但不同点在于运用 DLS 技术后，合成 MIDI 时并不利用 CPU 来运算，而依靠声卡自己的音频处理芯片进行合成。其中原因在于 PCI 声卡的数据宽带达到 133Mb/秒，大大加宽了系统内存与声卡之间的传输通道。从而既免去了传统 ISA 波表声卡所要配备的音色库内存，又大大降低了播放 MIDI 时的 CPU 占用率。而且这种波表库可以随时更新，并利用 DLS 音色编辑软件进行修改，这都是传统波表所无法比拟的优势。

FM 合成

既然 MIDI 文件只是一种对乐曲的描述，本身不包含任何可供回放的声音信息，那么一首首动听的电脑音乐又是如何被我们的声卡播放出来的哪？这就要通过形式多样的合成手段了。早先的 ISA 声卡普遍使用的是 FM 合成，既“频率调变”。它运用声音振荡的原理对 MIDI 进行合成处理。但由于技术本身的局限，加上这类声卡采用的大多数为廉价的 YAMAHA OPL 系列芯片，效果自然不好。



专业级的数字音频合成器

复音数的含义

在各类声卡的命名中，我们经常会发现诸如 64、128 之类的数字。有些用户乃至商家将它们误认为是 64 位、128 位声卡。其实就现在的技术发展状况而言，声卡更本没有发展到，也没有必要发展到如此高的数据处理通道，64、128 代表的只是此卡在 MIDI 合成时可以达



到的最大复音数。所谓“复音”是指 MIDI 乐曲在一秒钟内发出的最大声音数目。波表支持的复音值如果太小，一些比较复杂的 MIDI 乐曲在合成时就会出现某些声部被丢失的情况，直接影响到播放效果。好在如今的波表声卡大多提供 64 以上的复音值，而多数 MIDI 的复音数都没有超过 32，所以音色丢失的现象不会发生。

DSP

即 Digital Signal Processing 数字信号处理。DSP 技术在音调控制，失真效果器，Wah-wah 踏板等模拟电子领域有广泛的应用同时，DSP 在模拟均衡和混响等多种效果上也能大显身手。通过电脑 CPU 或专门的 DSP 芯片都可以处理 DSP 动作，不同的是，专门的 DSP 芯片处理要比电脑 CPU 处理更优化，速度更快。

信噪比

以 dB 计算的信号最大保真输出与不可避免的电子噪音的比率。该值越大越好。低于 75dB 这个指标，噪音在寂静时有可能被发现。AWE64 Gold 声卡的信噪比是 80dB，较为合理。SBLIVE 更是宣称超过 120dB 的顶级信噪比。总的说来，由于电脑里的高频干扰太大，所以声卡的信噪比往往不令人满意。但 SBLIVE! 提供了一个数字输出口 SPDIF，可绕过输出时的模拟部分，极大地减少了噪音和失真，同时又极大地提高了动态范围和清晰度。

FM

Synthesis Frequency Modulation 调配合成技术。FM 发声器的原理是先对音色本质进行研究计算之后再通过人工方式“模拟”（或合成）其频率，使用调变波去调变载波，最终获得不同的音色表现。

GM

即通用 MIDI 标准系统第一级（General MIDI system Level1），制定于 1991 年。在 GS 标准基础上，规定了 MIDI 设备的最大同时发音数不得少于 24 个、鼓镲等打击乐器作为一组单独排列、128 种乐器音色有统一的排列方式等。GM 的音色排列方式基本上沿袭了 GS 标准。它一推出便得到了全球 MIDI 厂商的一致支持，实现了 MIDI 设备间的交流。

XG

YAMAHA 于 1994 年提出了自己的音源标准——XG。XG 在兼容 GM 的基础上做了大幅度的扩展，能够进行“音色编辑”和自由发展不同层次的器材，具有强大的可扩充性。

SMPTE (The Society of Motion Picture and Television Engineers)

它是目前在影音工业中得到广泛应用的一个时间码概念。该码用于设备间驱动的时间同步，计数方式，主要参数格式是 Hours: Minutes : Second : Frames。SMPTE 24 Film Sync 以每秒 24 帧的速度播放，通常用于电影工业；SMPTE 30 Non-Drop 适用于音频领域。

音频压缩 (Audio Compression)

是指降低信号动态以滤除噪声和避免动态过大的失真。通过不同的计算方式、忽略人耳不易察觉的频段、或通过制造听觉上的错觉，从而大幅度降低音频数据的数量，却令音质基本不变甚至更好。