

多媒体技术基本概念



第6章 多媒体技术及其应用

根据考试大纲的规定，本章要求考生掌握以下知识：

多媒体系统基础知识；

简单图形的绘制，图像文件的处理方法；

音频和视频信息的应用；

多媒体应用开发过程。

但从历年考试试题来看，主要集中在音频、视频、图形和图像等方面。考生复习时应掌握基本概念，熟悉有关的多媒体文件容量、量化方面的计算。

在多媒体应用开发过程方面，考生要注意多媒体应用系统具有以下特点：

增强了计算机的友好性；

涉及技术领域广、技术层次高；

多媒体技术的标准化；

多媒体技术的集成化和工具化。

多媒体应用系统开发组需要应用系统组长、多媒体设计师、音频专家、视频专家、写作专家、多媒体程序员等人员，其具体开发过程与非多媒体项目是类似的，请读者参考本书有关软件工程知识的章节。

6.1 多媒体技术基本概念

多媒体主要是指文字、声音和图像等多种表达信息的形式和媒体，它强调多媒体信息的综合和集成处理。多媒体技术依赖于计算机的数字化和交互处理能力，它的关键是信息压缩技术和光盘存储技术。

1.亮度、色调和饱和度

视觉上的彩色可用亮度、色调和饱和度来描述，任意一种彩色光都是这3个特征的综合效果。

亮度是光作用于人眼时所引起的明亮程度的感觉，它与被观察物体的发光强度有关；由于其强度不同，看起来可能亮一些或暗一些。对于同一物体照射的光越强，反射光也越强，感觉越亮，对于不同物体在相同照射情况下，反射性越强者看起来越亮。显然，如果彩色光的强度降至使人看不清了，在亮度等级上它应与黑色对应；同样如果其强度变得很大，那么亮度等级应与白色对应。此外，亮度感还与人类视觉系统的视敏功能有关，即使强度相同，颜色不同的光进入视觉系统，也可能会产生不同的亮度。

色调是当人眼看到一种或多种波长的光时所产生的彩色感觉，它反映颜色的种类，是决定颜色的基本特性。如红色、绿色等都是指色调。不透明物体的色调是指该物体在日光照射下，所反射的各光谱成分作用于人眼的综合效果；透明物体的色调则是透过该物体的光谱综合作用的效果。

饱和度是指颜色的纯度，即掺入白光的程度，或者说是指颜色的深浅程度。对于同一色调的彩光，饱和度越深，颜色越鲜明，或者说越纯。例如，当红色加进白光之后冲淡为粉红色，其基本色调还是红色，但饱和度降低；换句话说，淡色的饱和度比深色要低一些。饱和度还和亮度有关，

因为在饱和的彩色光中增加白光的成分，由于增加了光能，因而变得更亮了，但是它的饱和度却降低了。如果在某色调的彩色光中掺入别的彩色光，会引起色调的变化，掺入白光时仅引起饱和度的变化。

2.三原色原理

三原色原理是色度学中最基本的原理，是指自然界常见的各种颜色光，都可由红（R）、绿（G）、蓝（B）3种颜色按不同比例相配制而成；同样绝大多数颜色光也可以分解成红、绿、蓝三种色光。当然三原色的选择并不是唯一的，也可以选择其他3种颜色为三原色，但是，3种颜色必须是相互独立的，即任何一种颜色都不能由其他两种颜色合成。由于人眼对红、绿、蓝3种色光最敏感，因此，由这3种颜色相配制所得的彩色范围也最广，所以一般都选用这3种颜色作为基色。

3.彩色空间

RGB彩色空间：在多媒体计算机技术中，用得最多的是RGB彩色空间表示。因为计算机的彩色监视器的输入需要R、G、B 3个彩色分量，通过3个分量的不同比例，在显示屏幕上可以合成所需要的任意颜色，所以不管多媒体系统采用什么形式的彩色空间表示，最后的输出一定要转换成RGB彩色空间表示。

YUV彩色空间：在现代彩色电视系统中，通常采用三管彩色摄像机或彩色CCD摄像机，把摄得的彩色图像信号经分色棱镜分成R0、G0、B0 3个分量的信号；分别经放大和校正得到三基色，再经过矩阵变换电路得到亮度信号Y、色差信号R-Y和B-Y，最后发送端将Y、R-Y和B-Y 3个信号进行编码，用同一信道发送出去，这就是我们常用的YUV彩色空间。

YUV彩色空间与RGB彩色空间的换算关系如下：

$$Y = 0.3 \times R + 0.59 \times G + 0.11 \times B;$$

$$U = (B - Y) \times 0.493;$$

$$V = (R - Y) \times 0.877.$$

其他彩色空间表示：彩色空间表示还有很多种如CIE（国际照明委员会）制定的CIE XYZ、CIE AB彩色空间和CCIR（Consultative Committee International Radio）制定的CCIR601-2YCC彩色空间，以及HIS（Hue,Saturation,Intensity）等。

版权方授权希赛网发布，侵权必究

[上一节](#) [本书简介](#) [下一节](#)

数据压缩标准

6.2 数据压缩标准

1.H.261

H.261是国际电联ITU-T的一个标准草案，H.261又称为P*64,其中P为64Bb/s的取值范围，是1到30的可变参数，它最初是针对在ISDN上实现电信会议应用特别是面对面的可视电话和视频会议而设计的。实际的编码算法类似于MPEG算法，但不能与后者兼容。H.261在实时编码时比MPEG所占用的CPU运算量少得多，此算法为了优化带宽占用量，引进了在图像质量与运动幅度之间的平衡折中机制，也就是说，剧烈运动的图像比相对静止的图像质量要差。因此这种方法是属于恒定码流可变

质量编码而非恒定质量可变码流编码。

2.H.263

H.263是国际电联ITU-T的一个标准草案，是为低码流通信而设计的。但实际上这个标准可用在很宽的码流范围，而非只用于低码流应用，它在许多应用中可以认为被用于取代H.261。H.263的编码算法与H.261一样，但做了一些改善和改变，以提高性能和纠错能力。H.263标准在低码率下能够提供比H.261更好的图像效果。H.263支持五种分辨率，即除了支持H.261中所支持的QCIF和CIF外，还支持SQCIF、4CIF和16CIF，SQCIF相当于QCIF一半的分辨率，而4CIF和16CIF分别为CIF的4倍和16倍。

1998年IUT-T推出的H.263+是H.263建议的第二版，它提供了12个新的可协商模式和其他特征，允许使用更多的源格式，图像时钟频率也有多种选择，拓宽应用范围；另一重要的改进是可扩展性，它允许多显示率、多速率及多分辨率，增强了视频信息在易误码、易丢包、异构网络环境下的传输。H.263已经基本上取代了H.261。

3.M-JPEG

M-JPEG (Motion-Join Photographic Experts Group) 技术即运动静止图像（或逐帧）压缩技术，广泛应用于非线性编辑领域，可精确到帧编辑和多层图像处理，把运动的视频序列作为连续的静止图像来处理，这种压缩方式单独完整地压缩每一帧，在编辑过程中可随机存储每一帧，可进行精确到帧的编辑，此外M-JPEG的压缩和解压缩是对称的，可由相同的硬件和软件实现。但M-JPEG只对帧内的空间冗余进行压缩。不对帧间的时间冗余进行压缩，故压缩效率不高。

M-JPEG标准所根据的算法是基于DCT（离散余弦变换）和可变长编码。M-JPEG的关键技术有变换编码、量化、差分编码、运动补偿、霍夫曼编码和游程编码等。M-JPEG的优点是可以很容易做到精确到帧的编辑，设备比较成熟；缺点是压缩效率不高。

4.MPEG-1

MPEG是活动图像专家组 (Moving Picture Experts Group) 的缩写，MPEG组织最初得到的授权是制定用于“活动图像”编码的各种标准，随后扩充为“伴随的音频”及其组合编码。后来针对不同的应用需求，解除了“用于数字存储媒体”的限制，成为现在制定“活动图像和音频编码”标准的组织。

MPEG-1标准于1993年8月公布，用于传输1.5Mb/s数据传输率的数字存储媒体运动图像及其伴音的编码。该标准包括5个部分：第一部分说明了如何根据第二部分（视频）及第三部分（音频）的规定，对音频和视频进行复合编码。第四部分说明了检验解码器或编码器的输出比特流符合前三部分规定的过程。第五部分是一个用完整的C语言实现的编码和解码器。

5.MPEG-2

MPEG组织于1994年推出MPEG-2压缩标准，以实现视/音频服务与应用互操作的可能性。MPEG-2标准是针对标准数字电视和高清晰度电视在各种应用下的压缩方案和系统层的详细规定，编码率从每秒3Mb/s~100Mb/s，标准的正式规范在ISO/IEC13818中。MPEG-2不是MPEG-1的简单升级，在系统和传送方面做了更加详细的规定和进一步的完善，特别适用于广播级的数字电视的编码和传送，被认定为SDTV和HDTV的编码标准。

MPEG-2图像压缩的原理是利用了图像中的两种特性：空间相关性和时间相关性。这两种相关性使得图像中存在大量的冗余信息。如果我们能将这些冗余信息去除，只保留少量非相关信息进行传输，就可以大大节省传输频带。而接收机利用这些非相关信息，按照一定的解码算法，可以在保证一定的图像质量的前提下恢复原始图像。一个好的压缩编码方案就是能够最大限度地去除图像中

的冗余信息。

MPEG-2的编码图像被分为3类，分别称为I帧、P帧和B帧。I帧图像采用帧内编码方式，即只利用了单帧图像内的空间相关性，而没有利用时间相关性。P帧和B帧图像采用帧间编码方式，即同时利用了空间和时间上的相关性。P帧图像只采用前向时间预测，可以提高压缩效率和图像质量。P帧图像中可以包含帧内编码的部分，即P帧中的每一个宏块可以是前向预测，也可以是帧内编码。B帧图像采用双向时间预测，可以大大提高压缩倍数。

为更好地表示编码数据，MPEG-2用句法规定了一个层次性结构。它分为6层，自上到下分别是：图像序列层、图像组（GOP）、图像、宏块条、宏块、块。

6.MPEG-4

MPEG组织于1999年2月正式公布了MPEG-4（ISO/IEC14496）标准第一版本。同年年底发布MPEG-4第二版，且于2000年年初正式成为国际标准。

MPEG-4与MPEG-1和MPEG-2有很大的不同。MPEG-4不只是具体压缩算法，它是针对数字电视、交互式绘图应用（影音合成内容）、交互式多媒体（WWW、资料撷取与分散）等整合及压缩技术的需求而制定的国际标准。MPEG-4标准将众多的多媒体应用集成于一个完整的框架内，旨在为多媒体通信及应用环境提供标准的算法及工具，从而建立起一种能被多媒体传输、存储、检索等应用领域普遍采用的统一数据格式。

MPEG-4标准同以前标准的最显著的差别在于它是采用基于对象的编码理念，即在编码时将一幅景物分成若干在时间和空间上相互联系的视频音频对象，分别编码后，再经过复用传输到接收端，然后再对不同的对象分别解码，从而组合成所需要的视频和音频。MPEG-4系统的一般框架是：对自然或合成的视听内容的表示；对视听内容数据流的管理，如多点、同步、缓冲管理等；对灵活性的支持和对系统不同部分的配置。

与MPEG-1、MPEG-2相比，MPEG-4具有如下独特的优点：基于内容的交互性；高效的压缩性；通用的访问性。MPEG-4提供了易出错环境的鲁棒性，来保证其在许多无线和有线网络及存储介质中的应用，此外，MPEG-4还支持基于内容的可分级性，即把内容、质量、复杂性分成许多小块来满足不同用户的不同需求，支持具有不同带宽、不同存储容量的传输信道和接收端。

MPEG-4的主要应用领域有：因特网多媒体应用；广播电视；交互式视频游戏；实时可视通信；交互式存储媒体应用；演播室技术及电视后期制作；采用面部动画技术的虚拟会议；多媒体邮件；移动通信条件下的多媒体应用；远程视频监控；通过ATM网络等进行的远程数据库业务等。

7.MPEG-7

MPEG-7标准被称为“多媒体内容描述接口”，为各类多媒体信息提供一种标准化的描述，这种描述将与内容本身有关，允许快速和有效地查询用户感兴趣的资料。它将扩展现有内容识别专用解决方案的有限能力，特别是它还包括了更多的数据类型。换言之，MPEG-7规定一个用于描述各种不同类型多媒体信息的描述符的标准集合，该标准于1998年10月提出。

MPEG-7的目标是支持多种音频和视觉的描述，包括自由文本、N维时空结构、统计信息、客观属性、主观属性、生产属性和组合信息。对于视觉信息，描述将包括颜色、视觉对象、纹理、草图、形状、体积、空间关系、运动及变形等。

MPEG-7的目标是根据信息的抽象层次，提供一种描述多媒体材料的方法，以便表示不同层次上的用户对信息的需求。以视觉内容为例，较低抽象层将包括形状、尺寸、纹理、颜色、运动（轨道）和位置的描述。对于音频的较低抽象层包括音调、调试、音速、音速变化、音响空间位置。最

高层将给出语义信息：如“这是一个场景：一只鸭子正躲藏在树后并有一辆汽车正在幕后通过。”抽象层与提取特征的方式有关：许多低层特征能以完全自动的方式提取，而高层特征需要更多人的交互作用。MPEG-7还允许依据视觉描述的查询去检索声音数据，反之也一样。

MPEG-7的目标是支持数据管理的灵活性、数据资源的全球化和互操作性。

MPEG-7标准化的范围包括：一系列的描述子（描述子是特征的表示法，一个描述子就是定义特征的语法和语义学）；一系列的描述结构（详细说明成员之间的结构和语义）；一种详细说明描述结构的语言、描述定义语言（DDL）；一种或多种编码描述方法。

MPEG-7标准可以支持非常广泛的应用，具体如下：音视数据库的存储和检索；广播媒体的选择（广播、电视节目）；因特网上的个性化新闻服务；智能多媒体、多媒体编辑；教育领域的应用（如数字多媒体图书馆等）；远程购物；社会和文化服务（历史博物馆、艺术走廊等）；调查服务（人的特征的识别、辩论等）；遥感；监视（交通控制、地面交通等）；生物医学应用；建筑、不动产及内部设计；多媒体目录服务（如黄页、旅游信息、地理信息系统等）；家庭娱乐（个人的多媒体收集管理系统等）。

8. MPEG-21

制定MPEG-21标准的目的是：（1）将不同的协议、标准、技术等有机地融合在一起；（2）制定新的标准；（3）将这些不同的标准集成在一起。MPEG-21标准其实就是一些关键技术的集成，通过这种集成环境对全球数字媒体资源进行透明和增强管理，实现内容描述、创建、发布、使用、识别、收费管理、产权保护、用户隐私权保护、终端和网络资源抽取、事件报告等功能。

任何与MPEG-21多媒体框架标准环境交互或使用MPEG-21数字项实体的个人或团体都可以看做是用户。从纯技术角度来看，MPEG-21对于“内容供应商”和“消费者”没有任何区别。MPEG-21多媒体框架标准包括如下用户需求：内容传送和价值交换的安全性；数字项的理解；内容的个性化；价值链中的商业规则；兼容实体的操作；其他多媒体框架的引入；对MPEG之外标准的兼容和支持；一般规则的遵从；MPEG-21标准功能及各个部分通信性能的测试；价值链中媒体数据的增强使用；用户隐私的保护；数据项完整性的保证；内容与交易的跟踪；商业处理过程视图的提供；通用商业内容处理库标准的提供；长线投资时商业与技术独立发展的考虑；用户权利的保护，包括：服务的可靠性、债务与保险、损失与破坏、付费处理与风险防范等；新商业模式的建立和使用。

9. DVI

DVI视频图像压缩法是Intel公司推出的一个压缩算法，其性能与MPEG-1相当，即图像质量可达到VHS的水平。压缩后的图像数据率约为1.5Mb/s.应用Intel公司生产的i750芯片组，即82750PB和82750DB可实时完成DVI视频图像的编码和解码算法。

为了扩大DVI技术的应用，Intel公司又推出了DVI算法软件解码算法，称为Indeo技术。它能将未压缩的数字视频文件压缩为1/5~1/10.Indeo技术已被附加在某些产品中，如微软公司Video for Windows和苹果公司的Quicktime.

Indeo技术使用多类有损和无损压缩技术。Indeo技术在视频捕获卡记录的同时实时地对它进行压缩，因此未压缩的数据不需存在盘上。从视频摄像机、VCR或激光盘上接收到的任何标准格式（如NTSC存在的视频）都由视频捕获卡（如Intel的Smart Video Recorder Board）转换为数字格式。

图形图像

6.3 图形图像

在计算机科学中，图形和图像这两个概念是有区别的：图形一般指用计算机绘制的画面，如直线、圆、圆弧、任意曲线和图表等；图像则是指由输入设备捕捉的实际场景画面或以数字化形式存储的任意画面。

图像是由一些排成行列的像素组成的，在计算机中的存储格式有BMP、PCX、TIF、GIF等，一般数据量都较大。它除了可以表达真实的照片外，也可以表现复杂绘画的某些细节，并具有灵活和富有创造力等特点。

与图像文件不同，在图形文件中只记录生成图的算法和图上的某些特征点，也称矢量图。在计算机还原输出时，相邻的特征点之间用特定的很多段小直线连接就形成曲线，若曲线是一条封闭的图形，也可靠着色算法来填充颜色。它的最大优点是容易进行移动、缩放、旋转和扭曲等变换，主要用于表示线框型的图画、工程制图、美术字等。常用的矢量图形文件有3DS（用于3D造型）、DXF（用于CAD）、WMF（用于桌面出版）等。图形只保存算法和特征点，所以相对于位图的数据量来说，它占用的存储空间也较小。但由于每次屏幕显示时都需重新计算，故显示速度没有图像快。另外，在打印输出和放大时，图形的质量较高而点阵图常会发生失真。

下面为了叙述的方便，我们不再区分图形和图像。

图形的主要指标为分辨率、色彩数与灰度。分辨率一般有屏幕分辨率和输出分辨率两种，前者用每英寸行数与列数表示，数值越大，图形质量越好；后者衡量输出设备的精度，以每英寸的像素点数表示，数值越大越好。如果一个图形是16位图像，则颜色数为2的16次方，共可表现65 536种颜色。当图形达到24位时，可表现1 677万种颜色，即真彩。常见的色彩位表示一般有2位、4位、8位、16位、24位、32位、64位等。

常见的图形有以下几种。

BMP:BMP是PC机上最常见的位图格式，有压缩和不压缩两种形式。BMP格式可表现从2位到24位的色彩，分辨率也可从480像素×320像素至1024像素×768像素。该格式在Windows环境下相当稳定，所以在对文件大小没有限制的场合中运用最为广泛。

DIB:DIB描述图像的能力基本与BMP相同，并且能运行于多种硬件平台，只是文件较大。

PCX:PCX是由Zsoft公司创建的一种经过压缩且节约磁盘空间的PC位图格式，它最高可表现24位图形。过去有一定的市场，但随着JPEG的兴起，其地位已逐渐降低了。

DIF:DIF是Auto CAD中的图形文件，它以ASCII方式存储图形，表现图形在尺寸大小方面十分精确，可以被CoreDraw、3DS等软件调用编辑。

WMF:WMF是Microsoft Windows图元文件，具有文件短小、图案造型化的特点，整个图形内容常由各独立组成部分拼接而成。但该类图形比较粗糙，并只能在Microsoft Office中调用编辑。

GIF:GIF是在各种平台的各种图形处理软件上均可处理的经过压缩的图形格式，该格式存储色彩最高只能达到256种。由于存在这种限制，除了Web网页还在使用它外，其他场合已很少使用了。

JPEG:JPEG格式可以大幅度地压缩图形文件。同样一幅画面，用JPEG格式存储的文件是其他类

型图形文件的1/10~1/20,而色彩数最高可达到24位,所以它被广泛运用于Internet上,以节约网络传输资源。JPEG文件之所以较小,是以损失图像质量为代价的。

PSD:PSD是Photoshop中的标准文件格式,专门为Photoshop而优化。

CDR:CDR是CoreDraw的文件格式。

PCD:Photo CD格式,由Kodak公司开发,其他软件系统对其只能读取。

版权方授权希赛网发布,侵权必究

[上一节](#) [本书简介](#) [下一节](#)

第6章:多媒体技术及其应用

作者:希赛教育软考学院 来源:希赛网 2014年01月26日

音频

6.4 音频

用计算机处理声音归结为语音合成、存储和输出等技术。

语音合成技术可分为发音参数合成、声道模型参数合成和波形编辑合成,语音合成策略可分为频谱逼近和波形逼近。

发音参数合成对人的发音过程进行直接模拟,定义了唇、舌、声带的相关参数,由这些发音参数估计声道截面积函数,进而计算声波。但由于人发音生理过程的复杂性,理论计算与物理模拟之间存在差异,合成语音的质量暂时还不理想。声道模型参数语音合成方法基于声道截面积函数或声道谐振特性合成语音,这类合成器的比特率低,音质适中。波形编辑语音合成技术基于时域波形修改的语音合成技术,直接把语音波表数据库中的波形级联起来,输出连续语流。这种语音合成技术用原始语音波形替代参数,而且这些语音波形取自自然语音的词或句子,它隐含了声调、重音、发音速度的影响,合成的语音清晰自然。其质量普遍高于参数合成。

推动喇叭发声的电信号是连续的模拟信号。计算机只能存储数字信号,模拟信号转换成数字信号包括采样和量化两个过程。采样是在一系列离散的时间点上测量模拟信号的大小,而量化则是用数字量来表示该大小。

实现计算机语音输出有两种方法:一是录音/重放,二是文-语转换。若采用第一种方法,首先要把模拟语音信号转换成数字序列,编码后暂存于存储设备中(录音),需要时再经解码,重建声音信号(重放)。录音/重放可获得高音质声音,并能保持特定人或乐器的音色。但所需的存储容量随发音时间线性增长。

第二种方法是基于声音合成技术的一种声音产生技术,它可用于语音合成和音乐合成。文-语转换是语音合成技术的延伸,它能把计算机内的文本转换成连续自然的语声流。若采用这种方法输出语音,应预先建立语音参数数据库、发音规则库等。需要输出语音时,系统按需求先合成出语音基元,再按语音学规则或语言学规则,连接成自然的语声流。文-语转换的参数库不随发音时间增长而加大,而规则库却随语音质量的要求而增大。

常见的音频格式如下。

WAVE:WAVE格式的声音文件的扩展名为WAV,这种格式记录了声音的波形,即模拟信号的采样数值。WAV文件所记录的声音文件能够和原声基本一致。在播放WAV文件时,只需进行数字模拟转换,将数字量转换成相应的电信号值并构成模拟信号即可推动喇叭发音。从理论上说,采样率达

44kHz(每秒采样44 000次)、采样字节长度达16位的音质已能和常规CD唱片相当。因为WAVE格式要把声音的每个细节都记录下来，而且不压缩，所以它的文件很大。例如，如果采样率为44kHz，那么每一秒钟就有 $44K \times 16 \times 2$ (立体声) $= 1\ 441\ 792$ 位产生，那么，一张650MB的空白光盘最多也只能容纳五六十分钟的节目。

MOD:MOD格式的声音文件的扩展名可为MOD、ST3、XT、S3M和FAR的任意一种。MOD及播放器大约起源于20世纪80年代初，原先是作为软声卡问世的，MOD只是这类音乐文件的总称。MOD格式的文件里不仅存放了乐谱(最初只能支持4个声道，到现在已有16,甚至32个声道的文件及播放器了)，而且存放了乐曲使用的各种音色样本。由于制作人创作歌曲使用的音色样本同听众回放文件时使用的音乐样本完全相同，所以这样的文件有几个显著的优点：回放效果明确；音色种类永无止境。

MPEG-3:MPEG-3格式的声音文件的扩展名为MP3,MPEG-3记录了音乐经数字比压缩的编码，压缩较大，在网络、可视电话通信方面，大有用武之地。但MPEG-3的失真较大。在播放MP3文件时，需要相应的解码器将它转换成模拟信号的数字序列，再经数字模拟转换推动喇叭发音。

Rea Audio:Rea Audio格式的声音文件的扩展名为RA,Rea Audio也是为了解决网络传输带宽资源而设计的，因此主要目标是压缩比和容错性，其次才是音质。Rea Audio压缩比很大，相对而言，Rea Audio的音质量比MPEG-3好。

CD Audio:CD Audio格式的声音文件的扩展名为CDA,回放和采样字节都是16位，现在有些厂家在录制CD时采用20位录音，这样就产生了一些耳朵听不到但大脑感觉得到的波形，可谓CD中的精品。CDA的缺点是：无法编辑，文件太大。

MIDI:MIDI格式的声音文件的扩展名是MID.MIDI (Musica Instrument Digital Interface,乐器数字接口)泛指数字音乐的国际标准，它始创于1982年。MIDI描述了音乐演奏过程的指令，利用MIDI文件演奏音乐，所需的存储量最少。MIDI标准规定了不同厂家的电子乐器与计算机连接的电缆和硬件。作为音乐工业的数据通信标准，MIDI是一种非常专业的语言，它能指挥各音乐设备的运转，而且具有统一的标准格式，能够模仿原始乐器的各种演奏技巧甚至是无法演奏的效果。MIDI依赖于回放设备，为了避免这种缺点，网络上出现了“软波表”之类的软音源。采用专业音源的波表，利用CPU对网络上传来的短短的MIDI数据进行回收，其效果能够被制作者预测。MIDI的另一个缺点就是不能记录人声等声音。

版权方授权希赛网发布，侵权必究

[上一节](#) [本书简介](#) [下一节](#)

第6章：多媒体技术及其应用

作者：希赛教育软考学院 来源：希赛网 2014年01月26日

视频

6.5 视频

动态图像，包括动画和视频信息，是连续渐变的静态图像或图形序列沿时间轴顺次更换显示，从而构成运动视感的媒体。当序列中每帧图像是由人工或计算机产生的图像时，常称为动画；当序列中每帧图像是通过实时摄取自然景象或活动对象时，常称为影像视频，或简称视频。

视频信息在计算机中存放具体格式有很多，常见的有下列几种。

Quicktime:Quicktime是苹果公司的产品，采用了面向最终用户桌面系统的低成本、全运动视频的方式，在软件压缩和解压缩中也开始采用这种方式。向量量化是Quicktime的软件压缩技术之一，它在最高为30帧/s下提供的视频分辨率是320像素×240像素，而且不用硬件帮助。向量量化预计可成为全运动视频的主要技术，向量量化方法达到的压缩比例为25:1~200:1.其视频信息采用MOV或QT文件格式。

AVI:AVI是微软公司的视频格式。音频视频交错（AVI）也是桌面系统上低成本低分辨率的视频格式，AVI可在160×120的视窗中以15帧/s回收视频并可带有8位的声音，也可以在VGA或超级VGA监视器上回收。与超过320线的VCR分辨率相比，这一分辨率明显低于正常电视信号的分辨率。AVI很重要的一个特点是可伸缩性，使用AVI算法的性能依赖于它一起使用的基础硬件。AVI包括了几种基于软件的压缩和解压缩算法，其中某些算法被优化用于运动视频，其他算法则被优化用于静止视频。

ReaMedia:ReaMedia是ReaNetworks公司所制定的音频/视频压缩规范，采用了流的方式播放，使用户可以边下载边播放，而且其极高的影像压缩率虽然牺牲了一些画质与音质，但却能在较慢的网速上流畅地播放ReaMedia格式的音乐和视频。ReaMedia是目前Internet上最流行的跨平台的客户/服务器结构多媒体应用标准，其采用音频/视频流和同步回放技术实现了网上全带宽的多媒体回放。在ReaMedia规范中主要包括3类文件：ReaAudio（用以传输接近CD音质的音频数据）、ReaVideo（用来传输连续视频数据）和ReaFash（ReaNetworks公司与Macromedia公司合作推出的新一代高压缩比动画格式）。其文件格式通常为RA或RM,一张用RM格式压缩的光盘上可以存放4部电影。ReaPayer是ReaMedia的播放工具，利用Internet资源对这些符合ReaMedia技术规范的音频/视频进行实况转播。

ASF:ASF是Advanced Streaming Format（高级流格式）的缩写，是微软公司为了和ReaMedia竞争而发展出来的一种可以直接在网上观看视频节目的文件压缩格式。由于它使用了MPEG-4的压缩算法，所以压缩率和图像的质量都很不错。因为ASF是以一个可以在网上即时观赏的视频流格式存在的，所以它的图像质量比VCD差，但比同是视频流格式的ReaMedia格式要好。

WMV:WMV是一种独立于编码方式的、在Internet上实时传播多媒体的技术标准，微软公司希望用其取代QuickTime之类的技术标准，以及WAV、AVI之类的文件扩展名。WMV的主要优点包括：本地或网络回放、可扩充的媒体类型、部件下载、可伸缩的媒体类型、流的优先级化、多语言支持、环境独立性、丰富的流间关系，以及扩展性等。

版权方授权希赛网发布，侵权必究

[上一节](#) [本书简介](#) [下一节](#)

例题分析

6.6 例题分析

D

例题1（2011年5月试题12）

以下编码方法中，（12）属于熵编码。

- (12) A.哈夫曼编码 B.小波变换编码 C.线性预测编码 D.PCM

例题分析：

本题主要考查编码方法。

根据解码后数据与原始数据是否完全一致可以分为两大类：一类是熵编码、冗余压缩法，也称无损压缩法、无失真压缩法；二是熵压缩法，也称有损压缩法、有失真压缩法。其中熵编码的一个重要特征就是解码后数据与原始数据完全一致，要达到这种效果，显然，那么压缩比就较低。这种方法一般适用于文本、数据的压缩，当然，如果要求有高质量的视音频的话，也可以采用这种方式压缩。常见的熵编码方法有LZ编码（Lempel-Zev编码）、行程编码、哈夫曼编码（Huffman编码）。而小波变换编码、线性预测编码、PCM、DPCM都是有损压缩。

哈夫曼编码是一种可变长的编码方式，该方法完全依据字符出现的概率来构造异字头的平均长度最短的码字。

小波变换编码的原理是因为大多数语音和图像信号倾向于非均匀频谱，即只集中在某一时间段的某一频段，小波包基表示信号时正好有这种自适应性，可通过选定一个阈值将此时间段和频段的相应系数保留编码而丢弃其他时间段和频段上的小幅值的系数；另外，小波包基的每次分解都把信号分成低频和高频两部分，而多数的语音和图像信号的统计特性表明大幅值的系数往往集中在低频区内，这样可给那些小幅值系数分配很少的比特数，甚至可以不传输或存储，从而压缩了数据。

线性预测编码中被预测的信号各样值都是此前样值的线性组合的预测编码。

PCM（脉冲编码调制）是把一个时间连续，取值连续的模拟信号转换成时间离散，取值离散的数字信号后在信道中传输。脉冲编码调制就是对模拟信号先抽样，再对样值幅度量化，编码的过程。

例题答案：(12) B

例题2（2011年5月试题13）

CIF视频格式的图像分辨率为（13）。

- (13) A.352x240 B.352x288 C.640x480 D.320x240

例题分析：

CIF是Common Intermediate Format的简称，即常用的标准化图像格式。在H.323协议簇中，规定了视频采集设备的标准采集分辨率CIF = 352×288像素。

例题答案：(13) B

例题3（2011年5月试题14）

由ISO制定的MPEG系列标准中，（14）是多媒体内容描述接口标准。

- (14) A.MPEG-1 B.MPEG-2 C.MPEG-4 D.MPEG-7

例题分析：

MPEG即活动图像专家组，MPEG组织制定的各个标准都有不同的目标和应用，目前已提出的标准有MPEG-1、MPEG-2、MPEG-4和MPEG-7。

MPEG-1标准用于传输1.5Mbps数据传输率的数字存储媒体运动图像及其伴音的编码。

MPEG-2主要用于实现视/音频服务与应用互操作的可能性，是针对标准数字电视和高清晰度电视在各种应用下的压缩方案和系统层的详细规定。

MPEG-4将众多的多媒体应用集成于一个完整的框架内，旨在为多媒体通信及应用环境提供标准的算法及工具，从而建立起一种能被多媒体传输、存储、检索等应用领域普遍采用的统一数据格

式。

MPEG-7 标准被称为“多媒体内容描述接口”，为各类多媒体信息提供一种标准化的描述，这种描述将与内容本身有关，允许快速和有效的查询用户感兴趣的资料。MPEG-7的目标是支持多种音频和视觉的描述，包括自由文本、N维时空结构、统计信息、客观属性、主观属性、生产属性和组合信息。对于视觉信息，描述将包括颜色、视觉对象、纹理、草图、形状、体积、空间关系、运动及变形等。

例题答案：(14) D

版权方授权希赛网发布，侵权必究

[上一节](#) [本书简介](#) [下一节](#)

第 7 章：计算机的体系结构和主要部件

作者：希赛教育软考学院 来源：希赛网 2014年01月26日

机内代码及运算

第7章 计算机的体系结构和主要部件

也许有人认为系统设计师似乎不需要什么硬件的知识，计算机的硬件理论知识并非是空中楼阁，不是好看或者用来考试的，它是确确实实的每台计算机设计和制造的基础，而且在学习的过程中，我们能够发现许多在硬件上使用的原则在软件上使用同样有益。很多时候，我们能发现，计算机软件和硬件并非截然分开，而是有个此消彼长、相互促进的发展过程。比如早期的中央处理器是没有浮点运算功能，浮点运算需要使用软件实现，而后来，许多处理器都内置了浮点运算功能。另外，嵌入式系统的软件设计师必须考虑硬件的问题。

7.1 机内代码及运算

人人都知道计算机只处理二进制数据，二进制是最简单的进制方式，只有0和1两个基数，也就是说，计算机底层硬件只要能保持两个状态即可，这样使得计算机的底层设计变得简单，出错的概率也大为减小。当然二进制数据使得表示和保存数据的长度大大增加，但是大规模和超大规模的集成电路使得这成为次要的问题，人们能在越来越小的芯片空间里容纳越来越多的电路。

另一方面计算机为了使得处理方便，其内部存储数据的格式和我们看见的是不同的。

版权方授权希赛网发布，侵权必究

[上一节](#) [本书简介](#) [下一节](#)

第 7 章：计算机的体系结构和主要部件

作者：希赛教育软考学院 来源：希赛网 2014年03月13日

数的进制

7.1.1 数的进制

1. 进制的表示法

R进制，通常说法就是逢R进1.可以用的数为R个，分别是0,1,2,..., R-1.例如十进制数的基数为10,即可以用到的数码个数为10,它们是0,1,2,3,4,5,6,7,8,9.二进制数的基数为2,可用的数码个数为2,它

们是0和1。

为了把不同的进制数分开表示，避免造成混淆，采用下标的方式来表示一个数的进制，如十进制数56表示为：(56)₁₀,八进制数42表示为：(42)₈。

对于任意一个R进制数，它的每一位数值等于该位的数码乘以该位的权数。权数由一个幂表示，即幂的底数是R,指数为k,k与该位和小数点之间的距离有关。当该位位于小数点左边，k值是该位和小数点之间数码的个数，而当该位位于小数点右边，k值是负值，其绝对值是该位和小数点之间数码的个数加1.

例如十进制数1234.56,其数值可计算如下：

$$1234.56 = 1 \times 10^3 + 2 \times 10^2 + 3 \times 10^1 + 4 \times 10^0 + 5 \times 10^{-1} + 6 \times 10^{-2}$$

例如二进制数10100.01的值可计算如下：

$$10100.01 = 1 \times 2^4 + 1 \times 2^2 + 1 \times 2^{-2}$$

2.进制的转换

1) R进制数转换成十进制数

按照上面的表示法，即可计算出R进制数十进制的值。

2) 十进制数转换为R进制数

最常用的是"除以R取余法".例如将十进制数94转换为二进制数：

2 94	余 0
2 47	1
2 23	1
2 11	1
2 5	1
2 2	0
1	1

将所得的余数从低位到高位排列 (1011110)₂就是94的二进制数。

3) 二进制数与八进制数、十六进制数之间的转换

将二进制数转换为八进制数，只有将每3个二进制数转换为八进制数即可，将二进制数转换为十六进制数，只要将每4个二进制数转换为八进制数即可。将八进制数转换为二进制数，只要将每个八进制数转换为3位二进制数即可，将十六进制数转换为二进制数，只要将每个十六进制数转换为4位二进制数即可。上面的转换都是以小数点作为计算数码个数的起点。八进制数和十六进制数转换可先转换为二进制数，然后再转换为目标进制。

版权方授权希赛网发布，侵权必究

[上一节](#) [本书简介](#) [下一节](#)

原码、反码、补码、移码

7.1.2 原码、反码、补码、移码

一个正数的原码、补码、反码是相同的，负数则不同。先提一个问题，为什么在计算机中要使用这些编码方式呢？

1.原码

将最高位用做符号位（0表示正数，1表示负数），其余各位代表数值本身的绝对值的表示形式。这种方式是最容易理解的。

例如，+11的原码是00001011，-11的原码是10001011。

但是直接使用原码在计算时却会有麻烦，比如 $(1)_{10} + (-1)_{10} = 0$ ，如果直接使用原码则：

$$(00000001)_2 + (1000001)_2 = (10000010)_2$$

这样计算的结果是-2，也就是说，使用原码直接参与计算可能会出现错误的结果。所以，原码的符号位不能直接参与计算，必须和其他位分开，这样会增加硬件的开销和复杂性。

2. 反码

正数的反码与原码相同。负数的反码符号位为1，其余各位为该数绝对值的原码按位取反。这个取反的过程使得这种编码称为“反码”。

例如，-11的反码：11110100

同样对上面的加法，使用反码的结果是：

$$(00000001)_2 + (11111110)_2 = (11111111)_2$$

这样的结果是负0，在人们普遍的观念中，0是不分正负的。反码的符号位可以直接参与计算，而且减法也可以转换为加法计算。

3. 补码

正数的补码与原码相同。负数的补码是该数的反码加1，这个加1就是“补”。

例如，-11的补码： $11110100 + 1 = 11110101$

再次做加法是这样的：

$$(00000001)_2 + (11111111)_2 = (00000000)_2$$

直接使用补码进行计算的结果是正确的。注意到我们这里只是举例，并非证明。

对一个补码表示的数，要计算其原码，只要对它再次求补，可得该数的原码。

由于补码能使符号位与有效值部分一起参加运算，从而简化运算规则，同时它也使减法运算转换为加法运算，进一步简化计算机中运算器的电路，这使得在大部分计算机系统中，数据都使用补码表示。

版权方授权希赛网发布，侵权必究

[上一节](#) [本书简介](#) [下一节](#)