

名词解释:

1、数字媒体

数字媒体是指以二进制数的形式记录、处理、传播、获取过程的信息载体，这些载体包括数字化的文字、图形、图像、声音、视频影像和动画等感觉媒体，和表示这些感觉媒体的表示媒体（编码）等，通称为逻辑媒体，以及存储、传输、显示逻辑媒体的实物媒体。但通常意义下所称的数字媒体常常指感觉媒体。

2、互联网

互联网，即广域网、局域网及单机按照一定的通讯协议组成的国际计算机网络。互联网是指将两台计算机或者是两台以上的计算机终端、客户端、服务端通过计算机信息技术的手段互相联系起来的结果，人们可以与远在千里之外的朋友相互发送邮件、共同完成一项工作、共同娱乐。

3、非线性编辑系统

非线性编辑是相对于线性编辑而言的，非线性编辑直接从计算机的硬盘中以帧或文件的方式迅速、准确地存取素材，进行编辑的方式。它是以计算机为平台的专用设备，可以实现多种传统电视制作设备的功能。编辑时，素材的长短和顺序可以不按照制作的长短和顺序的先后进行。

4、CMYK 色彩模式

CMYK 也称作印刷色彩模式，是一种依靠反光的色彩模式，和 RGB 类似，CMY 是 3 种印刷油墨名称的首字母：青色 Cyan、品红色 Magenta、黄色 Yellow。而 K 取的是 black 最后一个字母，之所以不取首字母，是为了避免与蓝色(Blue)混淆。从理论上来说，只需要 CMY 三种油墨就足够了，它们三个加在一起就应该得到黑色。但是由于目前制造工艺还不能造出高纯度的油墨，CMY 相加的结果实际是一种暗红色。

5、虚拟现实

虚拟现实(Virtual Reality, 简称 VR,)是利用电脑模拟产生一个三维空间的虚拟世界，提供使用者关于视觉、听觉、触觉等感官的模拟，让使用者如同身历其境一般，可以及时、没有限制地观察三度空间内的事物。

6、显示适配器

显示器配置卡简称为显卡，是个人电脑最基本组成部分之一。显卡的用途是将计算机系统所需要的显示信息进行转换驱动，并向显示器提供行扫描信号，控制显示器的正确显示，是连接显示器和个人电脑主板的重要元件，是“人机对话”的重要设备之一。显卡作为电脑主机里的一个重要组成部分，承担输出显示图形的任务，对于从事专业图形设计的人来说显卡非常重要。

7、三维动画

三维动画又称 3D 动画，是近年来随着计算机软硬件技术的发展而产生的一新兴技术。三维动画软件在计算机中首先建立一个虚拟的世界，设计师在这个虚拟的三维世界中按照要表现的对象的形状尺寸建立模型以及场景，再根据要求设定模型的运动轨迹、虚拟摄影机的运动和其它动画参数，最后按要求为模型赋上特定的材质，并打上灯光。当这一切完成后就可以让计算机自动运算，生成最后的画面。

8、RGB 色彩模式

RGB 色彩模式（也翻译为“红绿蓝”，比较少用）是工业界的一种颜色标准，是通过对红(R)、绿(G)、蓝(B)三个颜色通道的变化以及它们相互之间的叠加来得到各式各样的颜色的，RGB 即是代表红、绿、蓝三个通道的颜色，这个标准几乎包括了人类视力所能感知的所有颜色，是目前运用最广的颜色系统之一。

9、NPC

NPC 是（英文 Non-Player-Controlled Character ）“非玩家控制角色”的缩写。一般是指游戏中由程序控制的角色。

填空题

- 1、虚拟现实是人们通过计算机对复杂数据进行可视化操作与交互的一种全新方式。
- 2、根据拉斯韦尔的“五 W”模式，数字媒体的主要特点分别是传播者多样化、传播内容海量、传播渠道交互化、受传者个性化和传播效果智能化。
- 3、计算机图形图像处理软件分为位图软件和矢量软件两种。
- 4、数字媒体是指以二进制数的形式记录、处理、传播、获取过程的信息载体。
- 5、主流三维动画制作软件主要有 Maya 、 3D Max 和 SoftImage XSI 。

简答题

- 1、简述 AVI 格式优点与缺点。

AVI 格式的优点是图像质量好，可以跨多个平台使用，其缺点是体积过于庞大，而且更加糟糕的是压缩标准不统一，最普遍的现象就是高版本 Windows 媒体播放器播放不了采用早期编码编辑的 AVI 格式视频，而低版本 Windows 媒体播放器又播放不了采用最新编码编辑的 AVI 格式视频，所以我们在进行一些 AVI 格式的视频播放时常会出现由于视频编码问题而造成的视频不能播放或即使能够播放，但存在不能调节播放进度和播放时只有声音没有图像等一些莫名其妙的问题。

- 2、简述位图与矢量图的差异与特点。

计算机中显示的图形一般可以分为两大类——矢量图和位图。

矢量图，也称为面向对象的图像或绘图图像，在数学上定义为一系列由线连接的点。矢量文件中的图形元素称为对象。每个对象都是一个自成一体的实体，它具有颜色、形状、轮廓、大小和屏幕位置等属性。矢量图可以在维持它原有清晰度和弯曲度的同时，多次移动和改变它的属性，而不会影响图例中的其它对象。这些特征使基于矢量的程序特别适用于图例和三维建模，因为它们通常要求能创建和操作单个对象。基于矢量的绘图同分辨率无关。

位图是有一定数量的方形像素构成的。当放大位图时，可以看见赖以构成整个图像的无数单个方块。扩大位图尺寸的效果是增大单个像素，从而使线条和形状显得参差不齐。然而，如果从稍远的位置观

看它，位图图像的颜色和形状又显得是连续的。

矢量图与位图最大的区别是，它不受分辨率的影响。因此在印刷时，可以任意放大或缩小图形而不会影响出图的清晰度，可以按最高分辨率显示到输出设备上。

3、简述 Flash 动画技术的特点和在网络中的应用情况。

Flash 是一种创作工具，被大量应用于互联网网页的矢量动画文件格式。使用向量运算（Vector Graphics）的方式，产生出来的影片占用存储空间较小。使用 Flash 创作出的影片有自己的特殊档案格式（swf）。Flash 的设计公司声称全世界 97% 的网络浏览器都内建 Flash 播放器（Flash Player）是 Adobe 提出的“富因特网应用”（RIA）概念的实现平台。设计人员和开发人员可使用它来创建演示文稿、应用程序和其它允许用户交互的内容。Flash 可以包含简单的

动画、视频内容、复杂演示文稿和应用程序以及介于它们之间的任何内容。通常，使用 Flash 创作的各个内容单元称为应用程序，即使它们可能只是很简单的动画。您可以通过添加图片、声音、视频和特殊效果，构建包含丰富媒体的 Flash 应用程序。

Flash 特别适用于创建通过 Internet 提供的内容，因为它的文件非常小。Flash 是通过广泛使用矢量图形做到这一点的。与位图图形相比，矢量图形需要的内存和存储空间小很多，因为它们是以数学公式而不是大型数据集来表示的。

4、简述三维动画的制作流程。

三维动画的制作流程为建模、材质、灯光、动画、摄影机控制、渲染等。

建模，是动画师根据前期的造型设计，通过三维建模软件在计算机中绘制出角色模型。

材质贴图，材质即材料的质地，就是把模型赋予生动的表面特性，具体体现在物体的颜色、透明度、反光度、反光强度、自发光及粗糙程度等特性上。贴图是指把二维图片通过软件的计算贴到三维模型上，形成表面细节和结构。

灯光，目的是最大限度地模拟自然界的光线类型和人工光

线类型。灯光起着照明场景、投射阴影及增添氛围的作用。

摄影机控制，依照摄影原理在三维动画软件中使用摄影机工具，实现分镜头剧本设计的镜头效果。

动画，根据分镜头剧本与动作设计，运用已设计的造型在三维动画制作软件中制作出一个个动画片段。

渲染，是指根据场景的设置、赋予物体的材质和贴图、灯光等，由程序计算出一幅完整的画面或一段动画。

5、简述虚拟现实的基本特征。

多感知性（Multi-Sensory）

所谓多感知是指除了一般计算机技术所具有的视觉感知之外，还有听觉感知、力觉感知、触觉感知、运动感知，甚至包括味觉感知、嗅觉感知等。理想的虚拟现实技术应该具有一切人所具有的感知功能。

沉浸感（Immersion）

又称临场感，指用户感到作为主角存在于模拟环境中的真实程度。理想的模拟环境应该使用户难以分辨真假，使用户全身心地投入到计算机创建的三维虚拟环境中，如同在现实世界中的感觉一样。

交互性 (Interactivity)

指用户对模拟环境内物体的可操作程度和从环境得到反馈的自然程度 (包括实时性)。

构想性 (Imagination)

强调虚拟现实技术应具有广阔的可想像空间, 可拓宽人类认知范围, 不仅可再现真实存在的环境, 也可以随意构想客观不存在的甚至是不可能发生的环境。

第一章 数字媒体技术概论

媒体的分类*5

- **感觉媒体**: 用户接触信息的感觉形式
 - 视觉、听觉、触觉等
- **表示媒体**: 信息的表示形式 (传送、存储感觉媒体)
 - 文字、图像、声音、视频、动画等
- **显示媒体**: **表现和获取信息的** 物理设备
 - 显示器、打印机、扬声器、键盘等
- **存储媒体**: 存储表示媒体 (数字化后的感觉媒体的代码) 的物理设备
 - 磁带、磁盘、光盘等
- **传输媒体**: 传输数据的物理设备
 - 光缆、电缆、电磁波、交换设备等
- **数字媒体概念**
 - **数字媒体是指最终以二进制数的形式记录、处理、传播、获取的信息媒体。**包括**数字化的文字、图形、图像、声音、视频和动画及其编码等逻辑媒体**和存储、传输、显示逻辑媒体的**物理媒体**。但常常指逻辑媒体。
 - 数字媒体是数字化的**内容作品以现代网络为主要传播载体**, 通过完善的**服务体系**, 分发到终端和用户进行**消费**的全过程。

数字媒体的特性*4

- **数字化**

我们过去熟悉的媒体几乎都是以**模拟**的方式进行存储和传播的, 而数字媒体却是以**比特**的形式通过**计算机进行存储、处理和传播**。
- **多样性**

信息载体的多样性 (多媒体)

 - **融合**两种或两种以上媒体。**连续媒体 (音频和视频)**是人机交互最自然的媒体, 是对计算机更人性化的要求
- **交互性**

计算机的**“人机交互作用”**, 数字媒体是**以网络或信息终端为介质**的互动传播媒介。

 - 传统媒体 (电视、广播): **被动接受信息**
 - 多媒体: 是一种人—机交互式媒体, 向用户提供交互式的使用、加工和控制信息的手段

- 简单的交互式应用：如视频点播（VOD）、从数据库中检索出某人的照片等
- 高级的交互式应用：虚拟现实（Virtual Reality，VR）
- **集成性**
 - 多种**信息媒体**的集成
 - 处理这些媒体的**设备与设施**的集成
- //
- 趣味性
- 技术与艺术的融合

多媒体是技术与应用发展的必然

- 在计算机发展的初期，人们只能用**数值**这种媒体承载信息
 - 只能通过“0”和“1”表示信息
 - 纸带机和卡片机是主要的 I/O 设备
 - **机器语言**时代
 - 计算机的应用只限于极少数计算机专业人员
- 1950's~1970's，出现了**高级**程序设计语言，开始用文字作为信息的载体
 - 用**文字**（如英文）编程，输入计算机，处理结果也可用文字表示输出
 - 这时的 I/O 设备主要是打字机、键盘和显示终端
 - 计算机的应用扩大到具有**一般文化程度的科技人员**
- 1980's 开始，人们致力于研究将**声音、图形和图像作为新的信息媒体**输入输出计算机，这将使计算机的应用更为直观、容易
 - 1984 年 Apple 公司的 Macintosh 个人计算机首次引进“位映射”的图形机理，用户接口开始使用 Mouse 驱动的窗口技术和图符（Windows and Icon）
 - 这使得文化水平较低的公众，包括儿童在内都能够使用计算机

传播模式

大众传播模式：一对多

媒体信息传播模式（P6）

数字媒体传输模式：信源和信宿都是计算机；信源和信宿的位置是**可以随时互换**的；理想信道是**具有足够带宽的、可以传输比特流的高速网络信道**。网络可能由电话线、光缆或卫星通信构成。数字媒体可以是**多点之间**的传播。

超媒体传播模式：双向互动，非线性，多途径

视觉类媒体

- 视觉是人类感知信息**最重要的途径**
- 视觉类媒体元素
 - 文本(Text)、图像(Image)、图形(Graphic)、视频(Video)、动画(Animation)
- 文本是**使用最早的计算机媒体信息**
 - 英文文本：由 ASCII 码表示，每个字母、数字、标点或数学符号各占一个字节
 - 中文文本：由**国标 GB2312-80 的双字节编码**表示，每个字符占**两个字节**
 - **WPS (.wps)、Word (.doc)、Notepad (.txt) ...**
 - 非格式化文本文件、格式化文本文件

图像（点位图）

- 一幅图像可定义为一个二维函数 $f(x,y)$
 - (x,y) 为空间坐标； $f(x,y)$ 为图像在点 (x,y) 处的强度
 - 数字图像：当 x, y 和 f 都是有限的、离散的

图形（矢量图）

- 图形一般是指计算机绘制的画面，如图表等
- 不直接描述数据的每一点，而是描述产生这些点的过程和方法。图形的格式是一组描述点、线、面等几何图形的大小、形状、位置的指令
- 被操作时不会失真，清晰度与分辨率无关
- 虽然对用户而言是一样的，但是两者的存储结构和表示方法完全不同
- 图形是矢量结构的画面存储形式，抽象，数据量小，但显示成本高
- 图像是栅格结构的画面形式，大小与分辨率有关，基本元素是像素，逼真，数据量大
- 图形是更加抽象化的图像

一幅幅有联系的静态图像连续播放，利用了人眼的视觉暂留性，即每秒连续播放 20-30 帧 (frame)，产生运动画面的效果

- 视频：单帧是采集的真实图像
- 动画：单帧是由计算机产生或人工画出的图像或图形

听觉类媒体

- 波形声音 (Wave)
 - 20~20kHz，波形声音是对自然界声音进行采样和量化的结果，是自然界所有声音的拷贝
- 语音 (Speech)
 - 300~3.4kHz，语音（或称话音）是指人的说话声，它除了是声音的载体外，还包含语意和情感等信息
- 音乐 (Music)
 - 符号化了的声音，多媒体中专指 MIDI 音乐

研究领域

- 数字媒体表示与操作，包括数字声音及处理、数字图像及处理、数字视频及处理、数字动画技术等。
- 数字媒体压缩，包括通用压缩编码、专门压缩编码（声音、图像、视频）技术等。
- 数字媒体存储与管理，包括光盘存储（CD 技术、DVD 技术等）、媒体数据管理、数字媒体版权保护等。
- 数字媒体传输，包括流媒体技术、P2P 技术等。

研究方向

- 数字声音处理：包括音频及其传统技术（记录、编辑技术）、音频的数字化技术（采样、量化、编码）、数字音频编辑技术、语音编码技术（如 PCM、DA、ADM）。
- 数字图像处理：包括数字图像的计算机表示方法（位图、矢量图等）、数字图像的获

取技术、图像的编辑与创意设计。

- 数字**视频**处理：包括数字视频及其基本**编辑**技术、**后期特效处理**技术。常用的视频处理软件有 Premiere 等。
- 数字**动画设计**：包括动画的基本原理、动画设计基础（包括环节：构思、剧本、情节链图片、模板与角色、背景、配乐）、数字二维动画技术、数字三维动画技术、数字动画的设计与创意。常用的动画设计软件有 3DMAX、Flash 等。
- 数字**游戏设计**：包括游戏设计相关软件技术（DirectX、OpenGL 等）、游戏设计与创意。
- 数字媒体**压缩**：包括数字媒体压缩技术及分类、通用的数据压缩技术（行程编码、字典编码、熵编码等）、数字媒体压缩标准，如用于声音的 MP3、用于图像的 JPEG、用于运动图像的 MPEG。
- 数字媒体**存储**：包括内存存储器、外存储器、光盘存储器等
- 数字媒体**管理与保护**：包括数字媒体的数据管理、媒体存储模型及应用、数字媒体版权保护概念及框架、数字版权保护技术，如加密技术、数字水印技术、权利描述语言等。
- 数字媒体**传输**技术：包括流媒体传输技术、P2P 技术、IPTV 技术等。

应用领域

- 教育培训
- 电子商务
- 信息发布
- 游戏娱乐
- 电子出版
- 创意设计等

一. 一些常识

1. 目前我国采用的电视制式是 PAL 25 帧/s

2. 图像和视频编码的国际标准是 JPEG, MPEG, H.26X

一、填空题

- 1、对于媒体的含义，可以从_____和_____两个范畴理解。传递信息的载体、存储信息的实体
- 2、国际电信联盟定义了五种媒体，它们分别是_____、_____、_____、_____和_____。感觉媒体,表示媒体，显示媒体，存储媒体，传输媒体。
- 3、计算机记录和传播的信息媒体的一个共同的重要特点就是信息的最小单元是_____。比特
- 4、数字媒体的特点有_____、_____、_____、_____和_____。数字化，交互性，趣味性，集成性，技术与艺术的融合
- 5、数字媒体概念

第二章 数字音频技术基础

音频的概念及特性

- 音频与声音

- 物理学上，声音：波动的能量，声波
声音的三个基本特征：**频率、振幅和波形**
- 生理学上，声音是指声波作用于听觉器官所引起的一种主观感觉。
人对声音的主观描述：**音调、响度、音色**和音长等

音调 声音的高低，与**频率**（振动速率，Hz）有关

音色 与**波形**有关，取决于声波的**频谱**，即由混入基因的**泛音**所决定的。日常生活中人们听到的多是**复合音**，单纯的纯音（实验室的音频发生器、医院音叉）是很少的。

音强 声音的**响度**，与声音信号的**振幅**成正比，由**声能或声压**决定，以分贝（dB）为单位。

- 语音信号（speech）：人说话的信号频率通常为 **300~3400Hz**
- 亚音信号（subsonic）：小于 20Hz 的信号
- 超声波信号（ultrasonic）：高于 20KHz 的信号

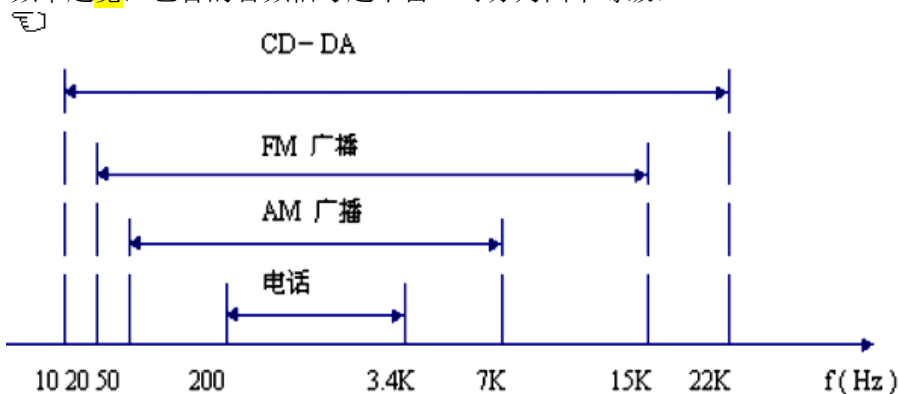
客观质量评价

- **频带宽度**

声音信号是由许多**频率不同的分量信号**组成的**复合**信号。

复合信号的频率范围称为**频带宽度（带宽）**。（单位 Hz）

频率越**宽**，包含的音频信号越丰富。可分为四个等级：



- **动态范围**

音频信号的**最大强度**和**最小强度之比**称为声音的动态范围。（单位 dB）

动态范围越**大**，说明音频信号的**相对变化范围大**，则音响效果越好。

记录声音

机械留声机 Edison→钢丝录音机 出现磁性记录→磁带录音机 电磁感应→数字音频 Philips
唱片→磁带→ CD 光盘→磁盘文件→流媒体

音频的数字化

- 模拟音频：时间和幅度上都是**连续**的（无限个点）
模拟磁性录音技术，受电磁性能影响较大
- 数字音频：时间和幅度上都是**离散**的（有限个点）
(模拟音频（采集）→A/D 转换→数字音频（存储、传输）→A/D 转换→模拟音频（还原）)
计算机、数字 CD、数字磁带（DAT）

数字化处理过程

- **采样**：**时间**上的离散化
- **量化**：**幅度**上的离散化（各个等级值均为整数→多对一→信息损失）

- 编码：将各个等级值转换成 0 和 1

数字化音频的质量取决于两个重要参数（带宽+动态范围）

1.采样频率：每秒钟采集多少个声音样本 $f_s = 1/\Delta t$ (Hz)

- 采样定理（Nyquist theory）

奈奎斯特理论指出：为不产生**低频失真**，**采样频率不应低于录制的最高频率的两倍。**

（ $f_s \geq 2F_{\max}$ ）

- 人类听觉的频率范围大约为：20~20kHz，为保证不失真，采样频率应在 **40kHz 左右**
- 常用的采样频率有：8kHz, 11.025kHz, 22.05kHz, **44.1kHz, (CD 音质), 48**

2.量化位数（量化深度/精度）。位数越多，声音质量越高，而**存储空间**也越多(b)

对每个采样点幅度进行量化所需要的二进制位数，它决定了模拟信号数字化以后的动态范围。常用的有 8 位、12 位和 16 位（8 位：256 个量化级（0~255）16 位：65 536 个量化级，16 位足矣）

<均匀量化>（线性量化）：相等的量化间隔对采样样本作量化。

缺点：为适应幅度大的输入信号，同时又要满足精度高的要求，就需要增加样本的位数

<非均匀量化>（非线性量化）：对大的输入采用大的量化间隔，小的输入采用小的量化间隔。做到在满足精度要求的情况下使用较少的位数

数字音频的存储

WAV 文件每秒存储量(字节，不经压缩)=采样频率 (Hz) × 量化位数 (位) × 通道数 / 8

如一分钟 CD-DA 音乐（CD 音质）所需的存储量为：(44.1×1000×16×2×60/8)B = 10 584 000B

常见声音文件

WAV 格式：波形声音文件；最早；无损（44.1kHz，16 位量化位数）；用于**存储**；存储空间需求太大

MIDI 格式（乐器数字接口的缩写，泛指数字音乐的标准）：**存储指令**，**合成**声音；易编辑；存储空间需求小；可做背景音乐

- 产生 MIDI 乐音的方法主要有两种：

<频率调制（FM）合成法> 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。采用不同的**调制波频率和调制指数**，就可以方便地合成不同频谱分布的波形，再现某些音色

<波表（wavetable）合成法> 把真实**乐器**发出的声音以数字的形式记录下来，播放时**改变播放速度**，从而改变音调周期，生成各种音阶的音符

CDA 格式

MP3 格式：**压缩**

RA、RM、WMA 格式：用于**传输**

声卡功能

音频录放、编辑

音乐合成

文语转换

CD-ROM 接口

MIDI 接口

游戏接口

模拟音频处理设备

- 话筒
 - 声音的收集，完成声能向电能的转化
- 音箱
 - Speaker, 扬声器
 - 还原声音，将音频电流信号转化成声音信号
- 模拟调音台
 - 可以拾取多路信号，每路声音信号可单独进行处理。还可以进行各种声音的混合。拥有多种输出。

数字音频处理设备

- (1) 专用数字音频设备
- (2) 非专为处理音频而设计的多媒体计算机
- 数字音频设备
 - (1) 数字调音台
 - (2) 数字录音机
 - (3) 数字音频工作站

数字音频编辑方式

- (1) 数字录音
- (2) 数字音乐创作 (MIDI)
- (3) 声音剪辑
- (4) 声音合成 (混音)
- (5) 增加特效
- (6) 文件操作

不管是专用设备，还是多媒体计算机，在处理数字音频时，其关键的硬件技术内核包括：

- 1) 模数转换器 (A/DC)
- 2) 数模转换器 (D/AC)
- 3) 数字信号处理器 (DSP)
 - 如：模拟和产生声场，控制声效等。

数字音频编辑软件分类

- 音源软件 (音序器软件)：数字音乐创作
- 编辑软件：声音的录音、剪辑、混音合成和特效处理等。

几个关键的技术概念：

- 1、声道
 - 声道是指声音在录制或播放时在不同空间位置采集或回放的相互独立的音频信号
 - 声道数也就是声音录制时的音源数量或回放时相应的扬声器数量
 - 声道就是不同位置发出的声音 (立体声即双声道)
- 2、音轨
 - 音轨就是在音频处理软件中看到的一条一条的平行“轨道”。每条音轨分别定义了该条音轨的属性
 - 在音序器软件中，一条音轨对应于音乐的一个声部或者对于一种乐器，它把 MIDI 或

者音频数据记录在特定的时间位置。

- 所有的音频处理软件都可以允许多音轨操作，也就是在某一段时间内，可以同时让多个音频素材同时播放，产生混音效果。
- 3、时序 (先后顺序)

二、填空题

- 2、按照声音的来源及作用，可分为_____、_____和_____。人声、乐音、响音
- 3、人耳可感受声音频率的范围为 20—20000Hz。声音高于 20000Hz 为_____，低于 20Hz 为_____。超声波、次声波
- 4、话筒的主要功能就是_____。进行声音能量的收集
- 5、音箱的主要功能就是_____，将音频电流信号变换成声音信号。还原声音
- 6、一般而言,音频数字化通常经过三个阶段,即_____,_____,_____。采样、量化、编码
- 7、数字化过程中,有两个非常重要的指标,一是_____,二是_____。量化深度、采样频率
- 8、采样频率越高,量化深度_____,声音质量_____。越大、越好
- 9、_____也称为比特率,即每秒种音频的二进制数据量。音频流码率
- 10、数字音频的常用格式有_____,_____,_____(任意列举三种)。WAV 格式、MIDI 格式、MP3 格式、CDA、MP3 Pro、WMA、MP4
- 11、数字音频的技术操作具体可以归纳为六点,即_____,_____,_____,_____,_____,_____。数字录音、数字音乐创作、声音剪辑、声音合成、增加特效、文件操作
- 12、_____是将模拟的音频电压流信号转换成数字脉冲电压信号。模数转换器
- 13、数字音频编辑软件可分为两种,一种是_____,另一种是_____。音源软件(音序器软件)、编辑软件
- 14、_____就是声音录制时的音源数量或回放时相应的扬声器数量。声道数
- 15、_____是指用一连串二进制数据来保存的声音信号。数字音频
- 17、_____可以分为两类:一类是专用数字音频设备,另一类是非专为处理音频而设计的多媒体计算机。数字音频处理设备

三、判断题

- 2、音频数字化的三个阶段是采样、量化、编码。
- 4、MAC 声音格式是 Apple 计算机公司开发的声音文件格式,其扩展名为.snd。
- 5、音频流码率也称比特率。

四、简答题

- 2、什么是数字音频?

答:数字音频是指用一连串二进制数据来保存的声音信号。这种声音信号在储存和电路传输及处理过程中,不在是连续的信号,而是离散的信号。

五、计算题

- 1、选择采样频率为 44.1KHZ、样本精度为 16bit 的声音数字化参数,在不采用压缩技术的情况下,录制 2min 的双声道的音频信号需要的存储空间为多少(MB)?

解: $(44.1 \times 1000 \times 16 \times 2 \times 60 \times 2) / 8 / 1024 / 1024 = 20.19 \text{ MB}$

六、论述题

- 1、数字音频处理过程中最关键的问题是音频的数字化问题。

答:这句话是正确的。

数字音频处理过程中最关键的问题：音频的数字化问题，也就是如何获得数字音频问题。
 要获得数字化音频信号，可以考虑两种途径：一是将现场声源的模拟信号或已存储的模拟声音信号通过某中方法转换成数字音频;二是在数字化设备中创作出数字音频，比如电子作曲。一般而言，数字化通常经过三个阶段，即采样—量化—编码。

第三章 数字图像的处理技术

图像可以定义为二维函数 $f(x, y)$ ，
 其中 x 和 y 是空间坐标， f 的振幅称为该点图像的强度或灰度
 数字图像：当 x, y 和 f 的振幅值都是有限的离散量时

数字图像最初应用之一在报业
 X 射线图像增强（工业、医学）
 从航空和卫星图像研究污染模式
 用于处理退化图像的图像增强和恢复过程
 机器感知
 光学字符识别（OCR）
 生物测定学 脸，指纹，虹膜
 自动目标识别（军用）
 用于产品装配和检验的工业机器视觉

可见光：380-780nm

颜色匹配是线性的

为了产生一幅数字图像，我们需要将连续的传感数据转换成数字形式

采样：数字化坐标值 量化：数字化振幅值

图像质量：采样→空间分辨率（图像中可辨别的最小细节，对于数字图像： $M \times N$ ）；

量化：灰度级分辨率（灰度变化最小，对于数字图像： L ）（颜色深度）

数字图像的种类与特点

矢量图 小 不会失真 大小取决于图的复杂程度 难以表现色彩层次丰富的逼真图像效果

点位图 大 会失真 大小取决于分辨率 像素数量固定 每个像素用若干个二进制位来指定属性（PSD 格式：分离图层-ps；BMP 格式：Windows 无损；JPEG 格式：压缩失真，适用于网页图像中；GIF 格式）

矢量图和点位图之间可以用软件进行转换，由矢量图转换成点位图采用光栅化(rasterizing)技术，这种转换也相对容易；由点位图转换成矢量图用跟踪(tracing)技术，这种技术在理论上说是容易，但在实际中很难实现，对复杂的彩色图像尤其如此。

颜色

①颜色是光与人类视觉系统交互产生的结果。

②颜色是物体对于人类的生理学特征，而非物体本身的物理学特征。

一．图像的数字化

1. 一种图像公式模型

显然，对于一副图像，图像任意一处强度满足以下条件

$$0 < f(x, y) < \infty$$

$f(x,y)$ 可以分解为两项

$$f(x, y) = i(x, y)r(x, y)$$

$i(x,y)$ 为该点的**照度**， $r(x,y)$ 为该点的**反射率**。

二．灰度直方图

1. 灰度直方图是一种直方图，其横坐标为灰度，纵坐标为每个灰度所对应的像素的数目。如果对于一副 $M \times N$ 的图像，我们设横轴变量为 r ，纵轴变量为 n ， r 、 n 之间对应的函数为 h ，其中 r 的范围是 0 到 $L-1$ 。那么有如下推论

$$\textcircled{1} h(r) = n$$

$$\textcircled{2} \sum_{k=0}^{L-1} n_k = M \times N$$

$$\textcircled{3} \sum_{k=0}^{L-1} n_k / n = 1$$

颜色模型（颜色空间）

机器：

RGB 颜色模型 计算机系统彩色显示器 **加色模型** $F=r[R]+g[G]+b[B]$ ，三个分量各占 1B，共 2^{24} 种颜色

CMYK 颜色模型 打印机 描述非发光体 **减色模型** 青色（Y）品红（M）黄色（Y）与黑色（K）油墨混合

人类：

HSB（HSI）颜色模型 人 色调（色相，H）饱和度（H，混入白色的程度）亮度（B，混入黑色的程度），色调和饱和度统称为色度

YUV（YCrCb）颜色模型 网络电视 所需存储容量小（亮色分离，注重亮度）

分辨率

显示分辨率 显示屏横向和侧向最大像素点个数（平均点距）。

图像分辨率 若采样像素点距固定，则原始图像尺寸越大，所得图像分辨率越大。

扫描分辨率 打印分辨率

颜色深度 一幅图像中可使用颜色数的最大值，单位 bit， $L=2^K$

一位二进制颜色编码的图像颜色深度为 1，也称二值图像（0.1）

• **24 位** 颜色可称之为真彩色，位深度是 24，它能组合成 2 的 2^4 次幂种颜色，即：16,777,216 种颜色（或称千万种颜色），超过了人眼能够分辨的颜色数量。当用 24 位来记录颜色时，实际上是以 $2^8 \times 2^8 \times 2^8 = 2^{24}$ ，即红、绿、蓝（**RGB**）三基色各以 2 的 8 次幂，256 种颜色而存在的，三色组合形成了一千六百多万种颜色。

• **32 位** 颜色的位深度是 32，实际上是 $2^8 \times 2^8 \times 2^8 \times 2^8 = 2^{32}$ ，即青、洋红、黄、黑（**CMYK**）四种颜色各以 2 的 8 次幂，256 种颜色而存在，四色的组合就形成 4,294,967,296 种颜色，或称为超千万种颜色。

真彩色 位深度是 24，有 2^{24} 种颜色

伪彩色 找一个最接近的颜色匹配（查找 R、B、G 强度值）

直接色

图像的大小计算 像素*颜色深度/8 单位为 B

(eg.分辨率 768*576, 颜色深度 24 图像大小为 $768*576*3/1024/1024=1.26\text{MB}$)

像素空间关系

- 像素间联系

像素 p 的邻域:

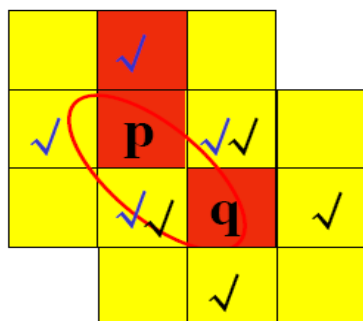
- 4 近邻 (4-neighbors): $N_4(p)$
- 对角近邻 (D-neighbors) : $N_D(p)$ (也是 4 个)
- 8 近邻 (8-neighbors) : $N_8(p)$

- 连通性(connectivity)是描述区域(region)和边界(boundary)的基本概念

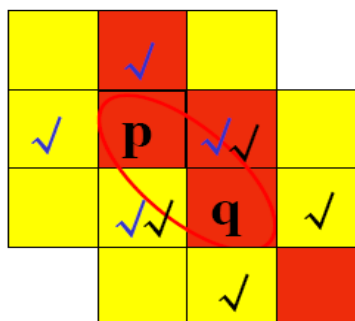
两个像素具有连通性的两个必要条件

- 两个像素是否相邻
- 它们的灰度级是否满足相似性准则
 - 设 V 是具有相似灰度的集合。对于二进制图像, 可令 $V=\{1\}$; 对于具有 256 灰度级的灰度图像, V 是这 256 数值中的一个子集
- 4 连通 (4-adjacency) 4 近邻
- 8 连通 (8-adjacency) 8 近邻
- m 连通 (m-adjacency, 混合连通) 像素 p 和 q 的象素值都属于集合 V , 如果
 - (i) q 属于集合 $N_4(p)$, 或者
 - (ii) q 属于集合 $N_D(p)$, 且 $N_4(p) \cap N_4(q)$ 中没有像素值属于集合 V 的像素
 则称 p 和 q 是 m 连通

是m连通



不是m连通



- 几何失真校正

- 空间变换: 对图像平面上的像素进行重新排列以恢复原空间关系
- 灰度插值: 对空间变换后的像素赋予相应的灰度值以恢复原位置的灰度值
- 失真: 线性\非线性失真

- 缩放图像

- 1. 创建新的像素位置

- 2. 灰度赋值

- 最近邻插值 (零阶插值) 将变换后的图像中的原像素点最邻近像素的灰度值赋给原像素点

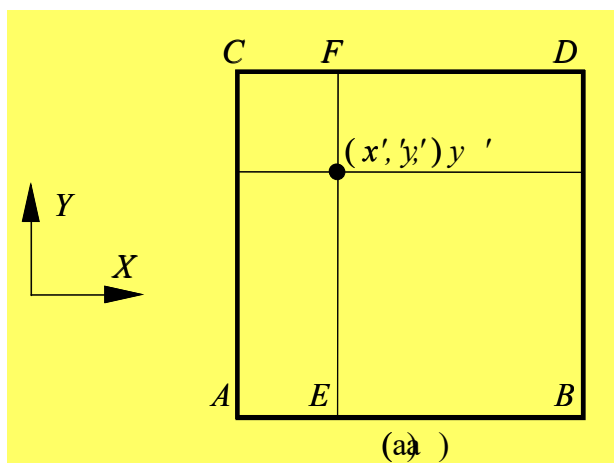
- 双线性插值 如果选择一个坐标系使得 的四个已知点坐标分别为 $(0, 0)$ 、 $(0, 1)$ 、 $(1, 0)$ 和 $(1, 1)$, 那么插值公式就可以化简为

- $$f(x,y)=f(0,0)(1-x)(1-y)+f(1,0)x(1-y)+f(0,1)(1-x)y+f(1,1)xy$$

$$g(E) = (x'-i)[g(B) - g(A)] + g(A)$$

$$g(F) = (x'-i)[g(D) - g(C)] + g(C)$$

$$g(x', y') = (y'-j)[g(F) - g(E)] + g(E)$$



一、选择题

3. 一幅 320×240 的真彩色图像,未压缩的图像数据量是 A 24 位真彩色
A 225KB B 230.4KB C 900KB D 921.6KB
4. 下列哪个是 photoshop 图像最基本的组成单元: (C)
A. 节点 B. 色彩空间
C. 像素 D. 路径

二、填空题

1、常见的颜色模型包括_____、_____和_____等。RGB、CMYK、HSB、YUV、CIE Lab 等。

四、问答题

1、图像的基本属性包含哪些?

答:图像的属性包含分辨率、颜色深度、真/伪彩色、图像的大小及种类等。

2、什么是图层

五、论述题

1、简述位图和矢量图的区别,并分别说明位图和矢量图的获取方法。

第四章 数字视频及编辑

扫描的机制

隔行扫描 (电视) 扫描两遍 (奇数场和偶数场, 合称一帧) 总行数必须是奇数

逐行扫描 (计算机显示图像)

帧速率和场速率 (行频、场频、帧频)

先有场后有帧

彩色全电视信号

- 黑白电视中只需传送一个**亮度信号**，而在彩色电视机中，则需要在满足与黑白电视兼容、而且**不增加为黑白电视所规定的信道带宽**（如 6 MHz）的条件下，同时传送亮度信号和**2 个色差信号**
- 如何找到一个可取的方案去实现这一要求，是彩色电视制式所解决的问题。不同的**色差信号传送方案**就形成了不同的**彩色电视制式**
- NTSC 制式系统对信道微分相位敏感，导致图像**色调失真**。
- PAL 制式是为克服这一缺点而提出的

PAL 制式：分别隔一场抽掉奇数行和偶数行

美/加/日/韩/台——29.97 帧/秒——NTSC(525 线-480 可视)——720×480

澳/中/欧/南美——25 帧/秒——PAL(625 线-576 可视)——720×576

法国/部分非洲——25 帧/秒——SECAM(625 线-576 可视)——720×576

- 复合电视信号（CVBS）
或称全电视信号，亮度信号和色度信号**采用频谱间置方法复合**在一起。这种方法易导致亮色串扰、清晰度降低等问题
- 分离电视信号（S-Video,（S 端子视频信号））
是**亮度和色差分离**的一种电视信号(Y/C),
- 分量电视信号
是指**每个基色分量作为独立的电视信号**，如：RGB 或 YUV，使用分量视频信号是表示颜色的最好方法，但需要比较宽的带宽和同步信号。

- **数字电视图像**有很多优点

例如，可直接进行**随机存储**使电视图像的**检索**变得很方便

复制数字电视图像和在网络上**传输**数字电视图像都**不会造成质量下降**

很容易进行**非线性电视编辑**等等

分量电视信号的数字化

- 先从复合彩色电视图像中**分离出彩色分量**，然后数字化

复合电视信号的数字化

- 对**色度信号和亮度信号**共频带所形成的复合电视信号**直接**进行数字化
- 早在 20 世纪 80 年代初，国际无线电咨询委员会 CCIR (International Radio Consultative Committee)就制定了彩色电视图像数字化标准，称为 CCIR 601 标准，现改为 ITU-R BT.601 标准
- 为了保证信号的同步，采样频率必须是电视信号行频的倍数。CCIR 为 NTSC、PAL 和 SECAM 制式制定的共同的电视图像采样标准：

$$f_s = 13.5\text{MHz}$$

- 人眼对彩色细节的分辨能力远比对亮度细节的分辨能力低

- **数字视频文件格式**

MPEG（Moving Pictures Experts Group）

（去时间冗余）运动补偿、帧间预测

MPEG-1、 MPEG-2、 MPEG-4

MOV

ASF (Advanced Streaming Format), MPEG-4 算法

WMV (Windows Media Video)

RM (Real Media)

RMVB

- (1) 本地影像视频
- **AVI** 格式, 这种视频格式的优点是**图像质量好**, 可以跨多个平台使用, 其缺点是体积过于庞大。**压缩标准不统一**是其主要问题。
- **DV-AVI** 格式, 它可以通过电脑的 IEEE 1394 端口传输视频数据到电脑, 也可以将电脑中编辑好的视频数据回录到数码摄像机中。这种视频格式的文件扩展名一般是.avi, 所以也叫 DV-AVI 格式。
- **MPEG** 格式, MPEG 文件格式是**运动图像压缩算法**的**国际标准**, 它采用了**有损压缩**方法减少运动图像中的冗余信息, 从而达到压缩的目的(其最大压缩比可达到 200:1)。
- **DivX** 格式, 是由 MPEG-4 衍生出的另一种视频编码(压缩)标准, 也即 DVDrip 格式, 它采用了 DivX 压缩技术对 DVD 盘片的视频图像进行高质量压缩, 同时用 MP3 或 AC3 对音频进行压缩, 然后再将视频与音频合成并加上相应的外挂字幕文件而形成的视频格式。其画质直逼 DVD 并且体积只有 DVD 的数分之一。MOV 格式, 美国 Apple 公司开发的一种视频格式, 具有较高的压缩比率和较完美的视频清晰度等特点, 但是其最大的特点还是**跨平台性**, 即不仅能支持 MacOS, 同样也能支持 Windows 系列。
- (2) 网络影像视频
- **ASF** 格式, 是微软为了和 Real Player 竞争而推出的一种视频格式。用户可以直接使用 Windows 自带的 Windows Media Player 对其进行播放。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。
- **WMV** 格式, 也是微软推出的一种采用独立编码方式并且可以直接在网上实时观看视频节目的文件压缩格式。WMV 格式的主要优点包括: 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。
- **RM** 格式, 这种格式的另一个特点是用户使用 Real Player 播放器可以在不下载音频/视频内容的条件下实现**在线播放**。另外, RM 作为目前**主流网络视频格式**, 可以通过其 Real Server 服务器将其它格式的视频转换成 RM 视频并由 Real Server 服务器负责对外发布和播放。
- **RMVB** 格式, 是一种由 RM 视频格式升级延伸出的新视频格式。RMVB 视频格式打破了原先 RM 格式那种平均压缩采样的方式, 在保证平均压缩比的基础上合理利用比特率资源, 就是说静止和动作场面少的画面场景采用较低的编码速率, 这样可以留出更多的带宽空间, 而这些带宽会在出现快速运动的画面场景时被利用。这样在保证静止画面质量的前提下, 大幅地提高了运动图像的画面质量, 从而图像质量和文件大小之间就达到了微妙的平衡。
- **数码摄像机就是 DV**, 是指摄像机的图像处理及信号的记录全部使用**数码信号**完成的摄像机。此种摄像机的最大的特征是**磁带上记录的信号为数码信号, 而非模拟信号**。
- **DVD 数码摄像机**由于其使用的存储媒介是 **DVD 刻录盘**, 所以与普通磁带摄像机相比, 在简便易用性上取得了突破性的进步:

DVD 数码摄像机可以**随机地进行回放**, 免去了倒带、快进等繁琐程序;

省却了上传到电脑后再制作成光盘的步骤, 拍摄后可直接在 DVD 播放机或 PC 上播放, 不必另外购置刻录机和压缩卡。

- 视频编辑

- 视频编辑包括了两个层面的操作含义：其一是传统意义上简单的画面拼接；其二是当前在影视界技术含量高的后期节目包装——影视特效制作。
- 就技术形式而言，视频可以分为两种形式：线性编辑和非线性编辑（是相对于传统上以时间顺序进行线性编辑而言，非线性编辑借助计算机来进行数字化制作。）。

一、选择题

- 1、电影的拍摄是以电影胶片为载体,借助透镜组(物镜)的光学成像,并根据视觉的生理与心理特征,以()摄取被摄对象的一系列姿态渐次变化而活动连贯的静止画面的过程.非曲直。B
A. 25 幅/秒 B. 24 幅/秒 C. 28 幅/秒 D. 3028 幅/秒
- 2、下列不是目前世界上现行的彩色电视制式的是()。D
A. PAL 制 B. NTSC 制 C. SECAM 制 D. SECNM 制
- 3、数字视频的属性有()。D
①视频分辨率 ②图像深度 ③帧率 ④压缩质量
A. ①② B. ①②③ C. ②③④ D. 全部
- 4、下列属于网络影像视频的是()。C
A. AVI B. MPEG C. ASF D. MOV
- 5、摄像机主要由()组成.D
①镜头系统 ②主机 ③寻像机 ④附件
A. ①②④ B. ②③ C. ①②③ D. ①②③
- 6、数字视频主要的获取途径包括().D
(1) 从现成的数字视频中截取 (2) 利用计算机软件作视频
(3) 用数字摄像机直接摄录 (4) 视频数字化
A. (4) B. (1) (2) C. (2) (3) D. (1) (2) (3) (4)
- 7、下列 DVD 数码摄像机与普通磁带摄像机相比优点叙述错误的为()。C
A. DVD 数码摄像机可以随机进行回放,免去了倒带、快进等繁琐程序
B. 省去了上传到计算机后再制作成光盘的步骤
C. 价格便宜适合大多数人使用
D. 拍摄后不必另外购置刻录机和压缩卡就可在 DVD 播放机或 PC 机上播放

二、填空题

- 1、视频编辑就技术而言可以分为两种形式：____和____。线性编辑、非线性编辑
- 2、数字视频编辑的一般步骤有：准备素材文件,进行素材的剪切,进行画面的粗略编辑，添加画面的过渡效果，____、____、____。添加字幕、处理声音效果、生成视频文件
- 3、我国采用的电视制式为____,帧频为____帧/秒.PAL、25
- 4、数字视频编辑后期特效处理的三种关键技术：____、____、其他视频特效.抠像、动画特效

三、判断题

- 2、视频采集卡的作用是将模拟视频信号输入计算机，并转换成数字信号。
- 3、计算机采用的是非隔行扫描,电视机采用的是隔行扫描。

三、简答题

- 1、请简要概括电视的工作原理。

答:电视是根据人眼视觉特性以一定的信号形式实时传送活动景物的技术。在发送端,用电视摄像机把景物转变成相应的电信号,电信号通过一定的途径传输到接收端,再由显示设备显示出原景物。

- 2、电视图像数字化常用的方法有哪些?

答:电视图像数字化常用的方法有以下两种:(1)先从复合彩色电视图像中分离出彩色分量,然后

数字化。(2) 首先用一个高速 A—D 转换器对彩色全电视信号进行数字化, 然后在数字域中进行分离, 以获得所希望的 YCBCR、YUV、YIQ 或 RGB 分量数据。

3、视频编辑包括哪两个层面的操作含义?

答: 其一是传统意义上简单的画面拼接; 其二是当前在影视界技术含量高的后期节目包装——影视特效制作。

4、什么是后期特效处理?

答: 后期特效处理是指: 通过跟随、抠像、校色、合成等操作分开各层的影像, 在影像上加特殊效果。

5、镜头: 从不同的角度、以不同的焦距、用不同的时间一次拍摄下来, 并经过不同处理的一段胶片, 它是一部影片的最小单位。

四、问答题

1、数字视频的获取方式有哪几种?

答: 数字视频的获取方式主要包括以下四种: 第一, 数字视频资源可以从现成的数字视频库中截取。包括从已有的视频文件中截取和从网站上下载。 第二, 利用计算机软件制作视频。比较常见的就是利用 Flash 软件制作二维动画和利用 3DSMAX 制作三维动画。 第三, 通过数字摄像机直接记录成数字文件格式, 然后存储到硬盘或者光盘之中。 第四, 通过一些重要设备和技术来实现从模拟视频到数字视频信号的转化。

2、后期特效处理所包含的内容有哪些? 后期特效处理的作用是什么?

答: 影视后期制作包括三大方面: 一是组接镜头, 也就是剪辑; 二是特效的制作; 三是在立体声进入到电影以后, 产生的后期声音制作。

后期特效处理不仅可以使平淡的视频片段化腐朽为神奇, 而且可以产生包括画面本身的变化、旋转、模仿旧电影胶片、色调变化等视频效果。

三、简答题

1、写出你所了解的游戏编程式语言 (至少写出 5 种)。

答: C 语言、C++、汇编语言、Pascal 语言、Visual Basic、Java 等。

第五章 数字媒体压缩技术

压缩的重要性

- 多媒体信息的数据海量性, 与当前硬件技术所能提供的计算机存储资源和网络带宽之间有很大差距
- 目前, 数字化的媒体信息数据以压缩形式存储和传输仍将是唯一选择
- 庞大的数据量是多媒体技术发展中的一个非常棘手的**瓶颈问题**
- 通过数据压缩手段把信息数据量压下来, 以压缩形式存储和传输
 - 紧缩节约了**存储空间**
 - 提高了通信干线的**传输效率**
 - 使计算机**实时处理**音频、视频信息, 以保证播放出高质量的视频、音频节目成为可能

图像数据压缩技术就是研究如何利用图像**数据的冗余性**来减少图像数据量的方法

空间冗余: 图像在空间上存在很大相关性, 如: 相邻像素值重复或非常接近。

时间冗余: 视频图像在时间上存在很大相关性。(二者合称**统计冗余**)

视觉冗余 (感知冗余): 人眼感受不到图像中的一些复杂细节信息, 丢弃这些信息不会影响视觉感受。视觉系统对图像的亮度和色彩度的敏感性相差很大 (YUV、YIQ 颜色空间)

人眼对高频信息不太敏感，对纹理丰富的内容区域（结构冗余）进行较粗的量化，以提升编码效率

知识冗余

信息熵冗余

听觉：人的听觉具有掩蔽效应

同时掩蔽、异时掩蔽

- 人耳对不同频段的声音敏感程度不同，通常对低频更敏感
- 对相位变化不敏感
- 人耳听不到或感知极不灵敏的声音分量都可视为冗余的

压缩方法分类

- 质量有无损失
 - 有损编码、无损编码（熵编码）
- 作用域
 - 空间方法、变换方法、混合方法
- 是否自适应
 - 自适应性编码、非自适应性编码

• 脉冲编码调制（PCM）

它实际上是连续模拟信号的数字采样表示。PCM 编码器和解码器位于一个图像编码系统的起点和终点，它们实际上分别是 A/D 和 D/A 转换器

其它编码方法通常都是在多媒体数据模拟信号经过 PCM 编码后再进行的

数据压缩的**理论基础是信息论**。根据信息论原理，可以找到最佳数据压缩编码方法，**数据压缩的理论极限是信息熵**

A. 信息量的定义

信息是用不确定性的度量来定义的。

信息论定义了一种度量信息量的方法

$$I(x_j) = -\log_2 P(x_j)$$

P 是信源发出 x 的先验概率

B. 信源的熵

- 信源的熵定义为信源 X 发出的 n 个随机事件 x_j (j=1,2,...,n) 的自信息量的统计平均。

$$H(X) = -\sum_{j=1}^n P(x_j) \log_2 P(x_j)$$

• 公式为

$$0 \leq H(X) \leq \log_2 n$$

• 信源的熵的范围是

C. 最佳编码判据

- 在编码中用熵衡量是否是最佳编码

- 对一种编码定义 N_0 为平均码长

$$\sum_{j=1}^n p_j N_j$$

计算公式为

- 当 $N_0 > H(X)$, 此时的压缩编码有冗余, 不是最佳编码
- 相等时, 是最佳编码。

- 熵的大小与信源的概率分布有着密切联系
- **等概率事件的熵最大**: 当信源中各事件为等概率分布时, 熵具有极大值 $\log_2 n$ 。n 为信源中事件个数

- 音频质量评价方法分为两类:

客观测量法: 信噪比 (SNR)

主观测量法: 主观平均判分法 (mean opinion score, MOS), 一般采用 5 分制

脉冲编码调制 (PCM)

差分脉冲编码调制 (DPCM)

自适应差分脉冲编码调制 (ADPCM) 自适应改变量化阶的大小: 小的量化阶去编码小的差值, 大的量化阶去编码大的差值

①. 无损编码方法

1. 霍夫曼编码 (P246)

2. 行程编码 (P247)

3. 词典编码 (P248)

1) LZ77 算法

2) LZW 算法

② 有损编码方法

1. 脉冲编码调制

2. 增量调制

3. 差分脉冲编码调制

按数据压缩编码的原理和方法进行划分

- **统计编码**, 主要针对**无记忆信源**, 根据信息码字**出现概率的分布特征**而进行压缩编码, 寻找概率与码字长度间的最优匹配, 对于出现频率大的符号用较少的位数表示, 而对于出现频率小的符号用较多的位数来表示。(霍夫曼编码、算术编码、香农编码、词典编码)
- **预测编码**是利用空间中**相邻数据的相关性**来进行压缩数据的, 编码器记录与传输的不是样本的真实值, 而是它与预测值的差。(PCM, DM, DPCM)
- **变换编码**是将图像**时域**信号转换为**频域**信号进行处理。(DCT)
- **分析—合成编码**是指通过对源数据的分析, 将其分解成一系列更适合于表示的“基元”或从中提取若干更为本质意义的参数, 编码仅对这些基本单元或特征参数进行。

按照媒体的类型进行压缩划分

图像压缩标准 声音压缩标准 运动图象压缩标准

8.2.2 霍夫曼编码

霍夫曼编码(Huffman)是运用信息熵原理的一种无损编码方法,这种编码方法根据源数据各信号发生的概率进行编码。在源数据中出现概率大的信号,分配的码字越短;出现概率越小的信号,其码字越长,从而达到用尽可能少的码表示源数据。

霍夫曼编码的算法

初始化,根据符号概率的大小顺序对符号进行排序。

把概率最小的两个符号组成一个新符号(节点),即新符号的概率等于这两个符号概率之和。

重复第2步,直到形成一个符号为止(树),其概率和等于1。

分配码字。码字分配从最后一步开始反向进行,即从最后两个概率开始逐渐向前进行编码,对于每次相加的两个概率,给概率大的赋“0”,概率小的赋“1”(也可以全部相反,如果两个概率相等,则从中任选一个赋“0”,另一个赋“1”)。

霍夫曼编码的特点

霍夫曼编码构造出来的编码值不是唯一的。

对不同信号源的编码效率不同

由于编码长度可变,因此译码时间较长;没有保护功能,可能会导致错误传播(码串中一个码译错导致其他代码出错);编码长度的不统一,也使得硬件实现有难度。

行程编码

行程编码又称行程长度编码(Run Length Encoding, RLE),是一种熵编码。这种编码方法广泛地应用于各种图像格式的数据压缩处理中。

行程编码的原理是在给定的图像数据中寻找连续重复的数值,然后用两个字符取代这些连续值。即将具有相同值的连续串用其串长和一个代表值来代替,该连续串就称为行程,串长称为行程长度。

行程编码分类

定长(表示串长的二进制位数)编码 不定长编码

算术编码

算术编码不是将单个信源符号映射成一个码字,而是把整个信源表示为实数线上的0到1之间的一个区间,其长度等于该序列的概率,再在该区间内选择一个代表性的小数,转化为二进制作为实际的编码输出

8.2.4 词典编码

词典编码(dictionary encoding)技术属于无损压缩技术,主要是利用数据本身包含许多重复的字符串的特性。可以用一些简单的代号代替这些字符串,就可以实现压缩,实际上就是利用了信源符号之间的相关性。字符串与代号的对应表就是词典。

8.2.4.1 LZ77 算法

LZ77 算法在某种意义上又可以称为“滑动窗口压缩”,该算法将一个虚拟的、可以跟随压缩进程滑动的窗口作为词典,要压缩的字符串如果在该窗口中出现,则输出其出现位置和长度。

LZ77 算法具体步骤

(1)把编码位置设置到输入数据流的开始位置。(2)找窗口中最长的匹配串(3)以“(Pointer, Length) Characters”的格式输出,其中Pointer是指向窗口中匹配串的指针,Length表示匹配字符的长度,Characters是前向缓冲存储器中的不匹配的第1个符。(4)如果前向缓冲存储器不是空的,则把编码位置和窗口向前移(Length+1)个字符,然后返回到步骤。

PCM 编码包括如下三个过程: (有损)

- 一 采样, 将模拟信号转换为时间离散的样本脉冲序列。 量化, 将离散时间连续幅度的抽样信号转换成为离散时间离散幅度的数字信号。编码, 用一定位数的脉冲码组表示量化采样值。

PCM 编码的优点

有很强的抗干扰性

能方便的利用计算机编程，实现各种智能化设计。

8.2.6 增量调制 (DM)

增量调制也称 Δ 调制(delta modulation, DM)，它是一种预测编码技术，是 PCM 编码的一种变形。

DM 是对实际的采样信号与预测的采样信号之差的极性进行编码，将极性变成“0”和“1”这两种可能的取值之一。如果实际的采样信号与预测的采样信号之差的极性为“正”，则用“1”表示；相反则用“0”表示，或者相反。(一位系统)输入信号变化快：斜率过载；输入信号变化慢：粒状噪声

8.2.7 差分脉冲编码调制

差分脉冲编码调制(Differential Pulse Code Modulation, DPCM)是利用样本与样本之间存在的信息冗余度来进行编码的一种数据压缩技术。

差值脉冲编码调制是利用信号的相关性找出可以反映信号变化特征的一个差值量进行编码。

DPCM 的基本工作原理

根据过去的样本来估算(estimate)下一个样本信号的幅度大小，这个值称为预测值，然后对实际信号值与预测值之差进行量化编码，从而就减少了表示每个样本信号的位数。

它与脉冲编码调制(PCM)不同之处在于，PCM 是直接对采样信号进行量化编码，而 DPCM 是对实际信号值与预测值之差进行量化编码。

声音压缩标准——MP3

MP3 是 MPEG-1 的标准草案中音频编码的 Layer 3。MP3 最大特点是能以较小的比特率、较大压缩比达到近乎完美的 CD 音质，制作简单，交流方便。MP3 压缩编码是一个国际性全开放的编码方案，其编码算法流程大致分为时频映射、心理声学模型、量化编码三大功能模块，这三个功能模块是实现 MP3 编码的关键。

MP4

MP4 是 MPEG-2 AAC(ISO/IEC 13818-7)技术(Advanced Audio Coding)。

MP4 的特点是音质更加完美而压缩比更大。它增加了诸如对立体声的完美再现、比特流效果音扫描、多媒体控制、降噪等 MP3 没有的特性，使得在音频压缩后仍能完美的再现 CD 的音质。

8.3.2 图像压缩标准

JPEG 包含两种基本的压缩算法：无损压缩算法(基于差分脉冲调制)和有损压缩算法(基于离散余弦变换)。

JPEG 算法压缩编码步骤

(1) 使用正向离散余弦变换(Forward Discrete Cosine Transform, FDCT)把信息从空间域变换成频率域的数据，并利用数据的频率特性进行处理；(2) 使用加权函数对 DCT 系数进行量化(信息损失)，这个加权函数对于人的视觉系统是最佳的；(3) 使用霍夫曼可变字长熵编码器对量化系数进行编码。

JPEG 压缩编码算法的主要计算步骤

(1) 正向离散余弦变换(FDCT)。(2) 量化(quantization)。(3) Z 字形编码(zigzag scan)。(4) 使用差分脉冲编码调制(differential pulse code modulation, DPCM)对直流系数(DC)进行编码。(5) 使用行程长度编码(Run Length Encoding, RLE)对交流系数(AC)进行编码。(6) 熵编码(entropy coding)。

JPEG2000

JPEG2000 弥补了传统 JPEG 压缩技术缺陷，有以下优点

高压压缩率 支持无损压缩和有损压缩 渐进传输 对感兴趣区域压缩

8.3.3.1 MPEG 标准

运动图像专家组 MPEG 是由国际标准化组织 ISO 和国际电工委员会 IEC 联合成立的，负责开发电视图像数据和声音数据的编码、解码和它们的同步标准。这个专家组开发的标准称为 MPEG 标准。(视频、音频、系统)

MPEG-1 和 MPEG-2 是 MPEG 组织制定的第一代视音频压缩标准，为 VCD、DVD 及数字电视和高清晰度电视等产业的飞速发展打下了牢固的基础

MPEG-4 是基于第二代视音频编码技术制定的压缩标准,以视听媒体对象为基本单元,实现数字视音频和图形合成应用、交互式多媒体的集成,目前已经在流式媒体服务等领域开始得到应用

MPEG-1 标准

- 1、采用 JPEG 压缩算法去掉空间冗余
- 2、采用运动补偿算法去掉时间冗余
- 引入了心理声学模型(根据人耳的听觉特性分配编码比特位数)

MPEG-2 标准

主要特点是:

(1) MPEG-2 解码器兼容 MPEG-1 和 MPEG-2 标准。4) 允许在画面质量、存储容量和带宽之间选择,在一定范围内改变压缩比。

MPEG-4 标准

MPEG-4 标准于 1998 年 11 月公布,是各种音频/视频对象的编码,包括了系统、电视图像、音频、一致性测试和参考软件、传输多媒体集成框架等,采用基于对象的编码、基于模型的编码等第二代编码技术,提高编码效率

MPEG-4 的目标是支持多种多媒体的应用,特别是多媒体信息基于内容的检索和访问,可以根据不同的应用需求现场配置解码器。其编码系统也是开放的,可以随时加入新的有效的算法模块。

MPEG-7 标准

MPEG-7 标准于 2001 年公布,称为多媒体内容描述接口,包括系统、描述定义语言、电视图像、音频、多媒体描述框架、参考软件以及一致性测试七个部分。

MPEG-7 标准的目的是产生一个描述多媒体内容的标准,支持对多媒体信息在不同程度层面上的解释和理解,从而使其可以根据用户的需要进行传递和存取。

MPEG-7 注重的是提供视听信息内容的描述方案,并不包括针对不同应用的特征提取方法和搜索引擎。

MPEG-21 标准

- MPEG-21 基于两个基本概念:分布和处理基本单元(Digital Item, DI)以及 DI 与用户间的互操作。MPEG-21 可表述为以一种高效、透明和可以互操作的方式支持用户交换、接入、使用甚至操作 DI 的技术。

H. 26X 系列视频标准

H. 26X 系列视频标准是国际电信联盟 ITU 的视频编码专家组 (ITU-T) 制定的系列图像压缩标准,主要有 H. 261、H. 263、H264 等。这些视频标准主要应用于实时视频通信领域,如会议电视、可视电话等。

8.3.3.3 AVS 标准

一、选择题:

- 1、下面哪一项不是数字媒体压缩标准 () B
 - A. 声音压缩标准
 - B. 文件压缩标准
 - C. 图像压缩标准
 - D. 视频压缩标准
- 5、Mp3 压缩编码是一国际性全开放的编码方案,其编码算法流程大致分为时频映射、心理声学模型、() 等三大功能块。 A
 - A. 量化编码
 - B. 多项滤波
 - C. 采样
 - D. 词典编码
- 6、视频编码标准 ITU—T 中视频会议标准有 H.261、H. 262 和 () B
 - A. H.264
 - B. H.263
 - C. H.265
 - D. H.266
- 7、增加容错功能,以支持丢包情况下的恢复,不包括 () D
 - A. 片断重同步
 - B. 数据分割
 - C. 可逆 VLC
 - D. 旧预测

二、填空题

- 1、一幅大小为 480*360 的黑白图像,每像素用 8bit 表示, 其大小为_____KB。168.75KB
- 2、一段采样频率为 44.1kHz,采样精度为 16bit/样本, 双通道立体声的数字音频, 其 1 秒钟的音频数据量为_____KB。172.3KB
- 3、按照媒体的类型进行压缩, 并进行标准化, 数据压缩分类可以分为_____, _____、_____, _____三种。图像、声音、运动图像
- 4、JPEG 包含_____, _____两种基本的压缩算法。无损压缩、有损压缩
- 5、MPEG2 压缩标准的文件格式是_____。DAT
- 6、根据压缩数据的原理把数据压缩方法分为四种:_____, _____、_____, _____。预测编码、预测编码、变换编码、分析-合成编码
- 7、数据在冗余主要体现在_____, _____、_____, _____、_____等。空间冗余、结构冗余、时间冗余、视觉冗余、知识冗余、信息冗余、信息熵冗余
- 8、数据压缩的理论极限是_____。信息熵
- 11、常用的无损压缩算法是_____。霍夫曼算法

三、判断题

- 1、在计算机系统的音频数据存储和传输中, 数据压缩会造成音频质量的下降。(对)
- 2、多媒体技术最大难题是海量数据存储与传送电视信号数字化后的数据量。(对)
- 3、数字压缩的英文缩写是 Data compression. (对)
- 7、在数字视频信息获取与处理过程中, 正确的顺序是采样、D/A 变换、压缩、存储、解压缩、A/D 变换。(错)
- 8、动态图像数据量的压缩比与压缩后数据的存储容量成正比, 与图像质量成反比。(错)
- 9、多媒体数据压缩按压缩方法分为有失真压缩和无失真压缩。(对)
- 10、数据压缩的原理是: 将原有的数据经过物理压缩,使其占地空间减少, 但不会影响文件应用性。(对)

四、简答题

- 1、简述 JPEG 和 MPEG 的主要差别。

答: JPEG 是适用于连续色调、多级灰度、彩色或单色静止图像的数据压缩标准.MPEG 视频压缩技术是针对运动图像的数据压缩技术。为了提高压缩比,MPEG 中帧内图像数据和帧间图像数据压缩。

- 2、数据压缩技术的三个主要指标是什么?

答:数据压缩技术有三个主要指标, 一是压缩前后所需的信息存储量之比要大; 二是实现压缩的算法要简单, 压缩、解压缩速度快, 尽可能地做到实时压缩和解压缩; 三是恢复效果要好, 要尽可能地完全恢复原始数据。

五、论述题

1. 什么是无损压缩、有损压缩.并列出你所知道的哪些格式是无损压缩,哪些格式是有损压缩?(每项至少写两个)

答: 无损压缩指使用压缩后的数据进行重构,重后的数据与原来的数据完全相同。无损压缩用于要求重构的信号与原始的信号完全一致的场合。常用的无损压缩算法有霍夫曼编码和 LZW 片算法, 无损压缩也称为可逆编码; 有损压缩是指压缩后的数据进行重构后与原来的数据有所不同, 但不影响人们对原始资料表达的信息造成误解.有损压缩适用于重构信号不一定非要和原始信号完全相同的场合。有损压缩也称为不可逆编码。

例如无损压缩的有: wav、ape、flac CD ape BMP、TIFF PSD PNG mp3

有损压缩: JPEG TGA EXIF gif ASF n AVI MPEG DIVX MOV RMVB rm (不是很全面)

第九章

一、选择题

- 1、将主存储器分为主存储器、高速缓冲存储器和 BIOS 存储器，这是按（ ）标准来划分。A
A. 工作原理 B. 封装形式 C. 功能 D. 结构
- 2、现在市场上流行的内存条是 _____ 线的。D
A. 30 线 B. 72 线 C. 128 线 D. 168 线
- 3、计算机上的内存包括随机存储器和 _____，内存条通常指的是_____。AB
A. ROM B. DRAM C. SDRAM
- 4、目前使用的内存主要是 _____。ABCD
A. SDRAM B. DDR SDRAM
C. DDR II SDRAM D. Super SDRAM
- 5、DDR SDRAM 内存的金手指位置有 _____ 个引脚。A
A. 184 B. 168 C. 220 D. 240
- 6、通常衡量内存速度的单位是 _____。A
A. 纳秒 B. 秒 C. 十分之一秒 D. 百分之一秒
- 7、下列哪一项不是光盘的主要特点。B
A. 记录密度高 B. 采用接触方式读写信息 C. 信息保存时间长 D. 价格低廉
- 8、下列哪项不属于光盘的分类 A
A. 防水性光盘存储器 B. 只读性光盘存储器 C. 多次可写光盘 D. 可擦写光盘存储器
- 9、可重复刻录光盘又称 C
A. VCD B. CD—R C. CD—RW D. CD—DA
- 10、下列哪些不属于 CD—ROM 驱动器的主要技术参数 C
A 速度 B. 数据传输速率 C. 光导厚度 D. 纠错能力
- 11、DVD 盘的标准直径为多少 A
A. 120mm B. 360mm C. 480mm D. 60mm
- 12、根据 DVD 的容量和格式可以将 DVD 大致分为几种。C
A. DVD—5 DVD—3 DVD—9 DVD—18
B. DVD—5 DVD—8 DVD—10 DVD—18
C. DVD—5 DVD—9 DVD—10 DVD—18
D. DVD—3 DVD—5 DVD—8 DVD—18
- 13、DVD 光盘的 4 种结构中容量最大的是 B
A. 单面双层 B. 双面双层 C. 单面单层 D. 双面单层
- 14、根据 DVD 的用途, 可以将 DVD 盘片分为五大类, 其中不包括 D
A. DVD Audio B. DVD—R C. DVD Video D. DVD Film
- 15、下列哪一个不是组成软盘存储器的一部分 (C)
A. 软盘 B. 软盘驱动器 C. 光盘 D. 软驱适配器
- 17、以下不属于外存储器的是 (C)
A. 移动硬盘 B. U 盘 C. 显示存储器 D. 软盘
- 18、不是衡量硬盘常用指标的是 (C)
A. 容量 B. 转速 C. 交错因子 D. 高速缓存容量

二、填空题

- 1、内存又称为 _____、_____；外存称为 _____。内部存储器、辅助存储器
- 2、_____ 是计算机系统的记忆部件，是构成计算机硬件系统的必不可少的一个部件。通常，根据

存储器的位置和所起的作用不同,可以将存储器分为两大类: _____ 和 _____. 存储器、内存、外存

3、内存的主要性能指标有: _____ 、 _____ 、 _____ 、 _____ 、和 _____. 内存容量、工作频率、数据带宽、工作电压、CAS 延迟时间

4 、计算机的内存是由 _____ 、 _____ 和 _____ 三个部分构成. RAM、ROM、Cache

5 、内存存在广义上的概念泛指计算机系统中存放数据与指令的半导体存储单元,它主要表现为三种形式: _____ 、 _____ 和 _____. RAM、ROM、Cache

6 、 RAM 一般又可分为两大类型: _____ 和 _____. SRAM、DRAM

7 、内存的数据带宽的计算公式是: 数据带宽 = _____ \times _____. 内存的数据传输频率、内存数据总线位数/8

8 、内存的工作频率表示的是内存的传输数据的频率,一般使 _____ 为计量单位. MHZ

9 、奇偶校验方法只能从一定程度上检测出内存错误,但 _____ ,而且不能 _____. 不能进行纠错、检测出双位错误

10 、只读存储器 (Read Only Memory) 的重要特点是只能 _____ ,不能 _____. 其刷新原理与 SRAM 类似,但消耗能量_____, 所以通常关闭计算机电源之后,其中数据还被保留。读取、写、较大

10 光盘具有_____, _____, _____等特点, 光盘的应用越来越广泛。容量大 存取速度快 不易受干扰

11 按光盘的读写性能,可将光盘分为_____, _____, _____三种类型. 只读型光盘存取器 多次可写光盘存取器 可擦写光盘存取器

三、问答题

2、什么是交错因子?

答案:每两个连续逻辑扇区之间所间隔的物理扇区数。

3、反应内存性能的技术指标有哪些?

答案:1) 存储周期 2) 数据宽带和带宽 3) 内存的“线”数 4) 容量 5) 内存的电压

第十二章 数字媒体技术发展趋势

一、填空题

1、内容分发包含从_____ Cache 的过程;从实现上看,有两种主流的内容分发技术_____和_____。

2、CDN 负载均衡系统实现_____。

3、整个 CDN 的核心是_____。

4、负载均衡的_____和_____直接决定了整个 CDN 的效率和性能。

5、三种传输网络、三大传输标准体系,组成了一个_____;不同的数字电视系统都会占据一个_____。

6、数字电视传输网络主要指术有_____、_____和_____。

7、ATSC 标准特点_____、_____。

8、10.目前世界上有三种数字电视标准,分别是: _____、_____和_____。

9、数字媒体压缩包括_____、_____。

10、在编目标准上,国际上为了方便广电行业各个单位之间的媒体资产交换,SMPTE 制定了完善的元数据模式(编目标准),称之为_____。

11、 SMEF 元数据模型包含_____个实体和_____个属性用来描述实体。

12、 根据学习的信息源,可将用记跟踪的方法分为两种: _____ 和_____。

13、 虚拟现实系统有_____、_____、_____、_____、_____5 个关键成分。虚拟现实技术的实现的三大技术分别为_____、_____、_____。

14、在数字版权保护技术领域中, 国外推出的 DRM 中具体的实现涉及的关键技术主要有_____、_____、_____。

二、简答:

16、虚拟现实的含义.

虚拟现实技术 (Virtual Reality), 又称灵境技术, 是指用立体眼镜和传感手套等一系列传感辅助设施来实现的一种三维现实。人们通过这些设施以自然的方式向计算机送入各种动作信息, 并且通过视觉、听觉及触觉设施使人们得到三维的视觉、听觉及触觉等感觉世界

19、请简述数字媒体内容处理包含哪些典型技术?

答案: 数字媒体内容处理技术包括: 音/视频编/解码技术, 内容条目技术, 内容聚合技术, 虚拟现实技术, 数字版权保护技术.

三、名词解释(写出它们的中文含义)

(1) VR (2) CDN (3) DRM

(1) 虚拟现实 (或灵境技术)

(3) 数字版权保护技术

答案

填空题

- 1、
- 2、CDN 的内容路由功能
- 3、负载均衡系统
- 4、准确性、效率
- 5、数字电视“网格”、 网格
- 6、地面、有线、卫星三大数字电视广播传输网络
- 7、频谱效率高、功率峰均比低
- 8、ATSC、DVB、ISDB
- 9、压缩编码、压缩编码
- 10、 (DCMI)
- 11、 (142)、(500)
- 12、 显式跟踪、隐式跟踪
- 13、 虚拟世界、虚拟现实软件、计算机、输入设备、输出设备;建模技术、显示技术、三维场景中的交互技术
- 14、 密码术、数字水印、身份认证、信任计算