

# 《多媒体技术基础(第2版)》

## 练习与思考题参考答案

### 第1章多媒体技术概要

#### 1.1 多媒体是什么？

多媒体是融合两种或者两种以上媒体的一种人—机交互式信息交流和传播媒体。使用的媒体包括文字、图形、图像、声音、动画和电视图像(video)。

#### 1.2 超文本的核心思想是什么？超文本系统和超媒体系统有什么差别？

(1) 超文本是通过复杂的、非顺序的关联关系连接在一起的一种文本，其真正含义是“链接”的意思。

(2) 超文本系统是以文本为主并使用超链接构成的信息系统；超媒体系统除文本外还包含图像、音乐、动画或其他元素构成并使用超链接构成的全球信息系统。

#### 1.3 超媒体是什么？多媒体与超媒体之间有什么关系？

(1) 使用文本、图形、图像、声音和电视图像等媒体任意组合的一种交互式信息传播媒体。

(2) 多媒体是超媒体系统中的一个子集。

#### 1.4 SGML是什么语言？HTML是什么语言？它们之间有何关系？

(1) 1986年国际标准化组织(ISO)采用的一个信息管理标准。该标准定义独立于平台和应用的文本文档的格式、索引和链接信息，为用户提供一种类似于语法的机制，用来定义文档的结构和指示文档结构的标签。

(2) HTML是万维网上的文档所用的标记语言。

(3) HTML是SGML的一个子集。SGML使用标签来标志文档中的文本或图形之类的元素，并告诉Web浏览器该如何向用户显示这些元素，以及应该如何响应用户的行为，例如当用户通过按键或鼠标单击某个链接时该如何响应。

#### 1.5 有人认为“因特网就是万维网”，这种看法对不对？为什么？

(1) 不对。

(2) 因特网是专指全球范围内最大的、由众多网络相互连接而成的、基于TCP/IP协议的计算机网络；万维网是指分布在全世界所有HTTP服务器上互相连接的超媒体文档的集合。

#### 1.6 多媒体、万维网和因特网之间有何关系？

多媒体构成了超媒体系统，超媒体系统构成了万维网，万维网是因特网上使用TCP/IP协议和UDP/IP协议的应用系统。

### 第2章数字声音及MIDI简介

#### 2.1 音频信号的频率范围大约多少？语音信号频率范围大约多少？

(1) Audio: 20~20000 Hz

(2) Speech: 300~3400 Hz

#### 2.2 什么叫做模拟信号？什么叫做数字信号？

(1) 幅度或频率发生连续变化的一种信号。

(2) 以二进制代码形式表示有无或者高低的一种信号。

#### 2.3 什么叫做采样？什么叫做量化？什么叫做线性量化？什么叫做非线性量化？

(1) 采样：在某些特定的时刻对模拟信号进行测量的过程。

(2) 量化：幅值连续的模拟信号转化成为幅值离散的数字信号的过程。

- (3) 线性量化：在量化时，信号幅度的划分是等间隔的量化。
- (4) 非线性量化：在量化时，信号幅度的划分是非等间隔的量化。

#### 2.4 采样频率根据什么原则来确定？

**奈奎斯特理论和声音信号本身的最高频率。**

#### 2.5 样本精度为8位的信噪比等于多少分贝？

48 分贝

#### 2.6 声音有哪几种等级？它们的频率范围分别是什么？

见表 2-01。

#### 2.7 选择采样频率为22.050 kHz和样本精度为16位的录音参数。在不采用压缩技术的情况下，计算录制2分钟的立体声需要多少MB(兆字节)的存储空间(1MB=1024×1024B)

$(22050 \times 2 \times 2 \times 2 \times 60) / (1024 \times 1024) = 10.09 \text{ MB}$

#### 2.8 什么叫做MIDI？它有什么特点？

- (1) 音乐合成器、乐器和计算机之间交换音乐信息的一种标准协议。
- (2) 文件比较小；容易编辑等

#### 2.9 用自己的语言说明FM合成声音和乐音样本合成声音的思想。

- (1) 把几种乐音的波形用数字表达，用计算机把它们组合起来，通过数模转换器(DAC)来生成乐音。
- (2) 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。

### 第3章 话音编码

#### 3.1 用自己的语言说出下面3种话音编译码器的基本想法。① 波形编译码器，②音源编译码器，③混合编译码器

- (1) 波形编译码器：不利用生成话音信号的任何知识而企图产生一种重构信号，它的波形与原始话音波形尽可能地一致。
- (2) 企图从话音波形信号中提取生成话音的参数，使用这些参数通过话音生成模型重构出话音。
- (3) 企图填补波形编译码和音源编译码之间的间隔。波形编译码器虽然可提供高话音的质量，但数据率低于 16 kb/s 的情况下，在技术上还没有解决音质的问题；声码器的数据率虽然可降到 2.4 kb/s 甚至更低，但它的音质根本不能与自然话音相提并论。

#### 3.2 列出你所知道的话音编译码器的主要指标(至少2个)

**音质，数据速率**

#### 3.3 试说混合编译码器的发展过程

为了得到音质高而数据率又低的编译码器，历史上出现过很多形式的混合编译码器，但最成功并且普遍使用的编译码器是时域合成-分析(analysis-by-synthesis, AbS)编译码器。这种编译码器使用的声道线性预测滤波器模型与线性预测编码(linear predictive coding, LPC)使用的模型相同，不使用两个状态(有声/无声)的模型来寻找滤波器的输入激励信号，而是企图寻找这样一种激励信号，使用这种信号激励产生的波形尽可能接近于原始话音的波形。AbS 编译码器由 Atal 和 Remde 在 1982 年首次提出，并命名为多脉冲激励(multi-pulse excited, MPE)编译码器，在此基础上随后出现的是等间隔脉冲激励(regular-pulse excited, RPE)编译码器、码激励线性预测 CELP(code excited linear predictive)编译码器和混合激励线性预测 (mixed excitation linear prediction, MELP)等编译码器。

### 3.4 什么叫做均匀量化？什么叫做非均匀量化？

(1) 均匀量化：采用相等的量化间隔对采样得到的信号进行量化。它是线性量化的另一种说法。

(2) 非均匀量化：采用非相等的量化间隔对采样得到的信号进行量化。例如，对大的输入信号采用大的量化间隔，对小的输入信号采用小的量化间隔。它是非线性量化的另一种说法。

### 3.5 什么叫做 $\mu$ 率压扩？什么叫做A率压扩？

(1) 在脉冲编码调制(PCM)系统中，一种模拟信号和数字信号之间进行转换的 CCITT 压(缩)扩(展)标准。在北美 PCM 电话网中，使用 $\mu$ 率压扩算法，详见“3.2.4  $\mu$ 律压扩”。

(2) 在脉冲编码调制(PCM)系统中，一种模拟信号和数字信号之间进行转换的 CCITT 压(缩)扩(展)标准。在欧洲电话网，使用A率压扩算法，详见“3.2.5 A律压扩”。

对于采样频率为 8 kHz，样本精度为 13 比特、14 比特或 16 比特的输入信号，使用 $\mu$ 律压扩编码或使用A律压扩编码，经过 PCM 编码器之后每个样本的精度为 8 比特，输出的数据率为 64 kb/s。

### 3.6 G.711标准定义的输出数据率是多少？T1的数据率是多少？T2的数据率是多少？

(1) G.711 使用 $\mu$ 率和A率压缩算法，信号带宽为 3.4 kHz，压缩后的数据率为 64 kb/s。

(2) T1 总传输率：1.544 Mb/s。

(2) T2 总传输率：6.312 Mb/s。

3.7 图3-23是DM编码器的原理图，如果你已经学过模拟电路和数字电路技术基础，请分析该电路是如何完成增量调制编码的。

(略)

### 3.8 自适应脉冲编码调制(APCM)的基本思想是什么？

根据输入信号幅度大小来改变量化阶大小的一种波形编码技术。这种自适应可以是瞬时自适应，即量化阶的大小每隔几个样本就改变，也可以是音节自适应，即量化阶的大小在较长时间周期里发生变化。

### 3.9 差分脉冲编码调制(DPCM)的基本思想是什么？

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。这种技术是根据过去的样本去估算(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)不同的是，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，存储或者传送的是差值而不是幅度绝对值，这就降低了传送或存储的数据量。此外，它还能适应大范围变化的输入信号。

### 3.10 自适应差分脉冲编码调制(ADPCM)的两个基本思想是什么？

ADPCM 综合了APCM 的自适应特性和DPCM 系统的差分特性，是一种性能比较好的波形编码。它的核心想法是：

- ① 利用自适应的思想改变量化阶的大小，即使用小的量化阶(step-size)去编码小的差值，使用大的量化阶去编码大的差值。
- ② 使用过去的样本值估算下一个输入样本的预测值，使实际样本值和预测值之间的差值总是最小。

## 第4章无损数据压缩

4.1 现有8个待编码的符号 $m_0, \dots, m_7$ 它们的概率如表练习\_表1所示。使用哈夫曼编码算法求出这8个符号的所分配的代码，并填入表中。(答案不惟一。参考答案：1, 000, 001, 011, 0101, 01000, 010010, 010011)

练习\_表1

待编码的符号	概率	分配的代码	代码长度 (比特数)
$m_0$	0.4	1	1
$m_1$	0.2	000	3
$m_2$	0.15	001	3
$m_3$	0.10	011	3
$m_4$	0.07	0101	4
$m_5$	0.04	01000	5
$m_6$	0.03	010010	6
$m_7$	0.01	010011	6

4.2 字符流的输入如练习\_表2所示，使用LZW算法计算输出的码字流。如果对本章介绍的LZW算法不打算进行改进，并且使用练习\_表3进行计算，请核对计算的输出码字流是否为：(1)(2)(4)(3)(5)(8)(1)(10)(11)。并将码字流中的码字填入练习\_表2对应的位置。

练习\_表2

输入位置	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	...
输入字符流	A	b	a	b	c	b	a	b	a	b	a	a	a	a	a	a	a	...
输出码字	A	b	-	ab	c	-	ba			bab	a	-	aa	-	-	aaa		

练习\_表3

步骤	位置	词典		输出码字
		(1)	a	
		(2)	b	
		(3)	c	
1	1	(4)	ab	(1)
2	2	(5)	ba	(2)
3	4	(6)	Abc	(4)
4	5	(7)	Cb	(3)
5	7	(8)	Bab	(5)
6	10	(9)	Baba	(8)
7	11	(10)	Aa	(1)
8	13	(11)	Aaa	(10)
9	16	(12)	Aaa	(11)
...	...	...	...	...

#### 4.3 LZ78算法和LZ77算法的差别在哪里？

(1) LZ77 编码算法的核心是查找从前向缓冲存储器开始的最长的匹配串(4.4.2 LZ77 算法)

(2) LZ78 的编码思想是不断地从字符流中提取新的缀-字符串(**String**)，通俗地理解为新“词条”，然后用“代号”也就是码字(**Code word**)表示这个“词条”。这样一来，对字符流的编码就变成了用码字(**Code word**)去替换字符流(**Charstream**)，生成码字流(**Codestream**)，从而达到压缩数据的目的。(4.4.4 LZ78 算法)

#### 4.4 LZSS算法和LZ77算法的核心思想是什么？它们之间有什么差别？

(1) LZ77 通过输出真实字符解决了在窗口中出现没有匹配串的问题，但这个解决方案包含有冗余信息。(4.4.3 LZSS 算法)

(2) LZ77 编码算法的核心是查找从前向缓冲存储器开始的最长的匹配串(4.4.2 LZ77 算法)

#### 4.5 LZW算法和LZ78算法的核心思想是什么？它们之间有什么差别？

(1) LZW 算法和 LZ78 算法的核心思想都是不断地从字符流中提取新的缀-字符串(**String**)，通俗地理解为新“词条”，然后用“代号”也就是码字(**Code word**)表示这个“词条”。这样一来，对字符流的编码就变成了用码字(**Code word**)去替换字符流(**Charstream**)，生成码字流(**Codestream**)，从而达到压缩数据的目的。(4.4.4 LZ78 算法)

(2) 在 LZW 算法中使用的术语与 LZ78 使用的相同，仅增加了一个术语——前缀根(**Root**)，它是由单个字符串组成的缀-字符串(**String**)。在编码原理上，LZW 与 LZ78 相比有如下差别：①LZW 只输出代表词典中的缀-字符串(**String**)的码字(**code word**)。这就意味在开始时词典不能是空的，它必须包含可能在字符流出现中的所有单个字符，即前缀根(**Root**)。②由于所有可能出现的单个字符都事先包含在词典中，每个编码步骤开始时都使用一字符前缀(**one-character prefix**)，因此在词典中搜索的第 1 个缀-字符串有两个字符。(4.4.5 LZW 算法)

## 第5章彩色数字图像基础

#### 5.1 什么叫做真彩色和伪彩色？

在一幅彩色图像中，每个像素值有 R，G，B 三个基色分量，每个基色分量直接决定显示设备的基色强度，这样产生的彩色称为真彩色。例如用 RGB 5：5：5 表示的彩色图像，R，G，B 各用 5 位，用 R，G，B 分量大小的值直接确定三个基色的强度，这样得到的彩色是真实的原图彩色。

现在，通常把每个像素的颜色值用 24 位表示的颜色叫做真彩色。

24 位表示的颜色总数： $2^{24} = 16\,777\,216$  种颜色。

#### 5.2 分别用3，3和2位表示的一幅图像，问该幅图像的颜色数目最多是多少？

256 种

#### 5.3 如果有一幅256色的图像，问该图的颜色深度是多少？

8 位

#### 5.4 按照JPEG标准的要求，一幅彩色图像经过JPEG压缩后还原得到的图像与原始图像相比较，非图像专家难于找出它们之间的区别，问此时的最大压缩比是多少？

24:1

#### 5.5 JPEG压缩编码算法的主要计算步骤是：①DCT变换，②量化，③Z字形编码，④使用DPCM对直流系数(DC)进行编码，⑤使用RLE对交流系数(AC)进行编码，⑥熵编码。假设计算机的精度足够高，问在上述计算方法中，哪些计算对图像的质量是有损的？哪些计算对图像的质量是无损的？

(1) DCT 变换： 无损

(2) 量化： 有损

- (3) Z 字形编码： 无损
- (4) 使用 DPCM 对直流系数(DC)进行编码： 无损
- (5) 使用 RLE 对交流系数(AC)进行编码： 无损
- (6) 熵编码： 无损

5.6 什么叫做  $\gamma$  校正？在计算机中找一幅彩色图像，使用Microsoft Office 97/2000中的Microsoft Photo Editor或者其他图像处理软件显示该图像，然后使用  $\gamma$  校正功能修改  $\gamma$  值，观察图像有什么变化。

(1) 在屏幕上显示用离散量表示的色彩时采用的一种色彩调整技术。计算机显示器和电视采用的阴极射线管产生的光亮度与输入的电压不成正比，而是等于以某个常数为底，以输入电压为指数的数，这个常数称为  $\gamma$ ，它的值随显示器的不同而改变，一般在 2.5 左右。

(2) (略)

5.7 什么叫做  $\alpha$  通道？它的作用是什么？

在每个像素用 32 位表示的图像表示法中的高 8 位，用于表示像素在一个对象中的透明度。例如，用两幅图 A 和 B 混合成一幅新图 New，新图的像素为：New pixel = (alpha)(pixel A color) + (1-alpha)(pixel B color)。在计算机环境下， $\alpha$  通道可存储在帧缓冲存储器中的附加位平面上。对于 32 位帧缓冲存储器，除高 8 位是  $\alpha$  通道外，其余 24 位是颜色位，红、绿和蓝各占 8 位。

5.8 在计算机中找一幅像素深度为24的彩色图像，使用Office 97/2000中的Microsoft Photo Editor或者其他图像处理软件显示该图像，然后用GIF格式存储，再显示GIF图像。观察图像有什么变化，并分析其原因。

(略)

5.9 PNG图像文件格式的主要特点是什么？

PNG 用来存储灰度图像时，灰度图像的深度可多到 16 位，存储彩色图像时，彩色图像的深度可多到 48 位，并且还可存储多到 16 位的  $\alpha$  通道数据。详见“5.7.4 PNG 格式”。

5.10 通过调查、试验和分析，把BMP，GIF，JFIF和PNG格式的一些特性填入下表。

图像格式名称	是不是有损压缩	支持的最大颜色数
BMP	无	16 777 216
GIF	无	256
JFIF	有	16 777 216
PNG	无	16 777 216

## 第6章颜色的度量体系

6.1 在开拓颜色科学方面, Newton, Thomas Young, Maxwell, Munsell, Ostwald和CIE分别做出了哪些重要贡献？

(1) Newton：发明了颜色圆，用于度量颜色

(2) Thomas ：认为人的眼睛有三种不同类型的颜色感知接收器，大体上相当于红、绿和蓝三种基色的接收器。

(3) Maxwell：探索了三种基色的关系，并且认识到三种基色相加产生的色调不能覆盖整个感知色调的色域，而使用相减混色产生的色调却可以。他认识到彩色表面的色调和饱和度对眼睛的敏感度比明度低。Maxwell 的工作可被认为是现代色度学的基础。

(4) Munsell：开发了第一个广泛被接受的颜色次序制，称为 Munsell color-order system 或者叫 Munsell color system，对颜色作了精确的描述并用在他的教学中。Munsell 颜色次序制也是其



他颜色体系的基础。

(5) Ostwald: 开发了 Ostwald 颜色体系, 根据对颜色起决定作用的波长、纯度和亮度来映射色调、饱和度和明度的值。

(6) CIE 定义了许多度量颜色的标准

## 6.2 什么是颜色空间? 对人、显示设备和打印设备, 通常采用什么颜色参数来定义颜色?

(1) 颜色空间: 表示颜色的一种数学方法, 人们用它来指定和产生颜色, 使颜色形象化。

(2) 对于人来说, 可以通过色调、饱和度和明度来定义颜色; 对于显示设备来说, 人们使用红、绿和蓝磷光体的发光量来描述颜色; 对于打印或者印刷设备来说, 人们使用青色、品红色、黄色和黑色的反射和吸收来产生指定的颜色。

## 6.3 什么叫做颜色系统(即颜色体系)? 简要说明组织和表示颜色的两种方法。

(1) 组织和表示颜色的方法。

(2) 两种方法: 颜色模型(color model), 编目系统(cataloging system)。

注:

(1) 颜色模型: 在台式机排版和图形艺术中, 表示颜色的任何一种方法或约定。在图形艺术和印刷领域, 颜色常用 *Pantone*(公司)颜色匹配系统; 在计算机图形学方面, 以下任何一种不同的色彩系统都可以描述色彩: *HSB*(色调, 饱和度和亮度), *CMY*(青, 品红, 黄)和 *RGB*(红, 绿, 蓝)。

(2) 颜色空间是颜色模型最普通的例子, *RGB*, *HSB*, *CMY*, *CIE XYZ*, *CIE LAB*, *CMYK* 和颜色的光谱描述方法都是颜色模型。

## 6.4 使用你能够找到的工具和资料, 探讨本章介绍的CIE度量体系是否有错误, 哪些地方需要修改和补充。CIE度量体系包括: ① CIE 1931 RGB ②CIE 1931 XYZ ③CIE 1931 xyY ④CIE 1960 YUV和CIE YU'V' ⑤CIE 1976 LUV ⑥CIE 1976 LAB ⑦CIE LUV LCh和⑧CIE LAB LCh (略)

# 第7章颜色空间变换

## 7.1 PAL制彩色电视使用什么颜色模型? NTSC制彩色电视使用什么颜色模型? 计算机图像显示使用什么颜色模型?

(1) PAL 制彩色电视: YUV

(2) NTSC 制彩色电视: YIQ

(3) 计算机图像显示设备: RGB

## 7.2 用YUV或YIQ模型来表示彩色图像的优点是什么? 为什么黑白电视机可看彩色电视图像?

(1) YUV 表示法的一个优点: 它的亮度信号(Y)和色度信号(U, V)是相互独立的, 因此可以对这些单色图分别进行编码; 另一个优点: 可以利用人眼的特性来降低数字彩色图像所需要的存储容量。YIQ 同样具有 YUV 的这两个优点。

(2) 黑白电视能接收彩色电视信号的道理是利用了 YUV/YIQ 分量之间的独立性。

## 7.3 在RGB颜色空间中, 当R=G=B, 且为任意数值, 问计算机显示器显示的颜色是什么颜色? 灰色

## 7.4 在HSL颜色空间中, 当H为任意值, S=L=0时, R, G和B的值是多少? 当H=0, S=1, L=0.5, R, G和B的值是多少?

当 H 为任意值, S=L=0 时,  $R = G = B = 0$

当 H=0, S=1, L=0.5,  $R = 1, G = B = 0$

## 7.5 打开Windows 95/98/2000中的“画图”→“颜色编辑”→程序, 在RGB和HSL转换栏中, 如果R=G=B=255, 问H, S和L的值分别为多少? 分别改变R, G和B的值, 观察H, S和L的值的变化。

如果 R=G=B=255, H=160, S=0, L=240。

(2) (略)

7.6 用MATLAB编写RGB到HSL和HSL到RGB颜色空间的转换程序: rgb2hsl.m和hsl2rgb.m。(略)

7.7 用MATLAB编写Y'CbCr和R'G'B'[0,219]颜色空间的转换程序: RGB2YCbCr和YCbCr2RGB.m。(略)

## 第8章小波与小波变换

8.1 写出矢量空间  $\mathcal{W}^3$  的哈尔小波并画出它的波形。

$$\begin{aligned}\psi_0^3(x) &= \begin{cases} 1 & 0 \leq x < 1/16 \\ -1 & 1/16 \leq x < 2/16 \\ 0 & \text{其他} \end{cases} & \psi_1^3(x) &= \begin{cases} 1 & 2/16 \leq x < 3/16 \\ -1 & 3/16 \leq x < 4/16 \\ 0 & \text{其他} \end{cases} \\ \psi_2^3(x) &= \begin{cases} 1 & 4/16 \leq x < 5/16 \\ -1 & 5/16 \leq x < 6/16 \\ 0 & \text{其他} \end{cases} & \psi_3^3(x) &= \begin{cases} 1 & 6/16 \leq x < 7/16 \\ -1 & 7/16 \leq x < 8/16 \\ 0 & \text{其他} \end{cases} \\ \psi_4^3(x) &= \begin{cases} 1 & 8/16 \leq x < 9/16 \\ -1 & 9/16 \leq x < 10/16 \\ 0 & \text{其他} \end{cases} & \psi_5^3(x) &= \begin{cases} 1 & 10/16 \leq x < 11/16 \\ -1 & 11/16 \leq x < 12/16 \\ 0 & \text{其他} \end{cases} \\ \psi_6^3(x) &= \begin{cases} 1 & 12/16 \leq x < 13/16 \\ -1 & 13/16 \leq x < 14/16 \\ 0 & \text{其他} \end{cases} & \psi_7^3(x) &= \begin{cases} 1 & 14/16 \leq x < 15/16 \\ -1 & 15/16 \leq x < 1 \\ 0 & \text{其他} \end{cases}\end{aligned}$$

8.2 写出4×4哈尔小波变换矩阵。

先写出:  $M_1 = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix}, M_2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$



$$(1) 4 \times 4 \text{哈尔小波变换矩阵: } W = M_1 M_2 = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{2} & 0 \\ \frac{1}{4} & -\frac{1}{4} & 0 & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} & 0 & -\frac{1}{2} \end{bmatrix}$$

(2) 规范化的4×4哈尔小波变换矩阵:

$$W = M_1 M_2 = \begin{bmatrix} \frac{1}{(\sqrt{2})^2} & \frac{1}{(\sqrt{2})^2} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{(\sqrt{2})^2} & \frac{1}{(\sqrt{2})^2} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{(\sqrt{2})^2} & -\frac{1}{(\sqrt{2})^2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{(\sqrt{2})^2} & -\frac{1}{(\sqrt{2})^2} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

8.3 使用MATLAB中的多级一维小波分解函数例程(function)wavedec，对例8.2所示的函数作小波变换。(略)

8.4 使用规范化的小波变换算法，用MATLAB编写一个M文件，重新计算的哈尔小波变换。(略)

$$f(x) = [2, 5, 8, 9, 7, 4, -1, -1]$$

## 第9章小波图像编码

### 9.1 什么叫做零树?

“零树”是指小波变换系数之间的一种数据结构。因为离散小波变换是一种多分辨率的分解方法，每一级分解都会产生表示图像比较粗糙(低频图像)和比较精细(高频图像)的小波系数，在同一方向和相同空间位置上的所有小波系数之间的关系可用一棵树的形式表示，如果树根和它的子孙的小波系数的绝对值小于某个给定的阈值  $T(\text{threshold})$ ，那么这棵树就叫做零树。

### 9.2 解释EZW的含义。

嵌入零树小波编码(embedded zerotree wavelet, EZW)是 Shapiro, J. M 在 1993 年开发的一种编码算法。“小波”表示该算法以离散小波变换为基础，以变换后的大系数比小系数更重要以及高频子带中的小系数可以忽略为背景；“零树”表示小波变换系数之间的一种数据结构，用同一方向和相同空间位置上的所有小波系数构成一棵树，如果树根及其子孙的小波系数的绝对值小于某个给定域值，则这棵树就称为零树；“嵌入”表示一幅图像可以分解成一幅低分辨率图像和分辨率由低到高(表示图像细节)的许多子图像，图像合成过程是通过子图像生成许多分辨率不同的图像。EZW 编码就是按照用户对图像分辨率的不同要求，编码器进行多次编码，每进行一次编码，域值降低 1/2，水平和垂直方向上的图像分辨率各提高 1 倍。编码从最低分辨率图像开始扫描，每当遇到幅度大于域值的正系数就用符号 P 表示，幅度小于域值的负系数用符号 N 表示，树根节点上的系数幅度小于域值而树枝中有大于域值的非零树用符号 Z 表示，零树用符号 T 表示，编码的输出是符号集 {P, N, T, Z, 0, 1} 中的一系列符号。

9.3 如果条件允许，用MATLAB或者其他语言编写执行EZW算法的编码和解码程序(略)

9.4 解释SPIHT的含义。

层树分集算法(set partitioning in hierarchical trees, SPIHT)的根据是, 图像经过小波变换之后, 大部分能量都集中在低频子带。从这个事实出发, 最先传送幅度大的系数, 这样解码器即使在低速率应用环境下也可得到图像的大部分信息。编码树的结构与 EZW 算法的结构类似, 每一个节点要么没有子节点, 要么有 4 个子节点。在编码过程中, 使用三个列表变量存储重要系数和不重要系数。该算法适用于图像的渐进传输, 具有比较高的 PSNR, 复杂度比较低, 计算量比较少, 位速率容易控制等优点。

9.5 如果条件允许, 用MATLAB或者其他语言编写执行SPIHT算法的编码和解码程序。(略)

9.6 请用因特网搜索工具, 查找并阅读EBCOT的详细说明。(略)

9.7 如果条件允许, 用MATLAB或者其他语言编写执行EBCOT算法的编码和解码程序。(略)

9.8 JPEG 2000有许多功能, 请用因特网搜索工具调查和描述它的详细功能。(略)

## 第10章彩色数字电视基础

10.1 世界上主要的彩色电视制式是哪几种?

目前世界上使用的彩色电视制式主要有 PAL, NTSC, SECAM 三种, 都是模拟彩色电视制式。

10.2 隔行扫描是什么意思? 非隔行扫描是什么意思?

(1) 在隔行扫描中, 一帧画面分两场, 第一场扫描总行数的一半, 第二场扫描总行数的另一半。电子束扫完第 1 行后回到第 3 行开始的位置接着扫, 如图 10-01(b)所示, 然后在第 5、7、……, 行上扫, 直到最后一行。奇数行扫完后接着扫偶数行, 这样就完成了一帧(frame)的扫描。隔行扫描要求第一场结束于最后一行的一半, 不管电子束如何折回, 它必须回到显示屏顶部的中央, 这样就可以保证相邻的第二场扫描恰好嵌在第一场各扫描线的中间。正是这个原因, 才要求总的行数必须是奇数。

(2) 在非隔行扫描中, 电子束从显示屏的左上角一行接一行地扫到右下角, 在显示屏上扫一遍就显示一幅完整的图像。

10.3 电视机和计算机的显示器各使用什么扫描方式?

电视机通常使用隔行扫描; 计算机的显示器通常使用非隔行扫描。

10.4 ITU-R BT.601标准规定PAL和NTSC彩色电视的每一条扫描线的有效显示像素是多少?

720

10.5 什么叫做S-Video? 它的连接器结构是什么样?

质量较高的一种电视技术, 用这种技术录制和处理电视时亮度信号(Y)和色差信号(C)是分开进行的。S-Video 是分量模拟电视信号和复合模拟电视信号的一种折中方案。使用 S-Video 有两个优点: (1) 减少亮度信号和色差信号之间的交叉干扰。(2) 不须要使用梳状滤波器来分离亮度信号和色差信号, 这样可提高亮度信号的带宽。

S-VHS 和 Hi-8 摄像机以及盒式录像机(video cassette recorder, VCR)使用这种技术, 它比标准的 VHS 和 8 毫米格式的设备提供质量比较高的彩色图像。在使用 S-video 技术的设备上, 电视信号的输出除含有分离的亮度(Y)和色差信号(C)输出外, 还有复合彩色全电视信号。S-video 设备与其他设备连接的连接器使用 S-video 连接器, 而不是普通的 RCA 连接器。有些电视图像卡和高档家用录像机(VCR)可以支持这种信号。

10.6 对彩色图像进行子采样的理论根据是什么?

人的视觉系统所具有的两种特性。一是人眼对色度信号的敏感程度比对亮度信号的敏感程度低, 利用这个特性可以把图像中表达颜色的信号去掉一些而使人不察觉; 二是人眼对图像细节的分辨能力有一定的限度, 利用这个特性可以把图像中的高频信号去掉而使人不易察觉。子采样是压缩彩色电视信号的一种技术。

### 10.7 图像子采样是在哪个彩色空间进行的?

YCbCr

### 10.8 一幅YUV彩色图像的分辨率为720×576。分别计算采用4:2:2、4:1:1和4:2:0子采样格式采样时的样本数。

(1) 4:4:4 这种采样格式不是子采样格式，它是指在每条扫描线上每 4 个连续的采样点取 4 个亮度 Y 样本、4 个红色差 Cr 样本和 4 个蓝色差 Cb 样本，这就相当于每个像素用 3 个样本表示。

$$720 \times 576 \times 3 = 1\,244\,160$$

(2) 4:2:2 这种子采样格式是指在每条扫描线上每 4 个连续的采样点取 4 个亮度 Y 样本、2 个红色差 Cr 样本和 2 个蓝色差 Cb 样本，平均每个像素用 2 个样本表示。

$$720 \times 576 \times 2 = 829\,440$$

(3) 4:1:1 这种子采样格式是指在每条扫描线上每 4 个连续的采样点取 4 个亮度 Y 样本、1 个红色差 Cr 样本和 1 个蓝色差 Cb 样本，平均每个像素用 1.5 个样本表示。

$$720 \times 576 \times 1.5 = 622\,080$$

(4) 4:2:0 这种子采样格式是指在水平和垂直方向上每 2 个连续的采样点上取 2 个亮度 Y 样本、1 个红色差 Cr 样本和 1 个蓝色差 Cb 样本，平均每个像素用 1.5 个样本表示。

$$720 \times 576 \times 1.5 = 622\,080$$

## 第11章MPEG简介

### 10.1 回顾制定MPEG标准的4个阶段和各阶段的提交的文件类型。

(1) 工作文件(Working Draft, WD): 工作组(Working Group, WG)准备的工作文件

(2) 委员会草案(Committee Draft, CD): 从工作组 WG 准备好的工作文件 WD 提升上来的文件。这是 ISO 文档的最初形式，它由 ISO 内部正式调查研究和投票表决。

(3) 国际标准草案(Draft International Standard, DIS): 投票成员国对 CD 的内容和说明满意之后由委员会草案 CD 提升上来的文件。

(4) 国际标准(International Standard, IS): 由投票成员国、ISO 的其他部门和其他委员会投票通过之后出版发布的文件。

### 10.2 MPEG-1, -2, -4和-7的目标是什么?

MPEG-1 处理的是标准图像交换格式的电视，即 NTSC 制为 352 像素×240 行/帧×30 帧/秒，PAL 制为 352 像素×288 行/帧×25 帧/秒，压缩的输出速率定义在 1.5 Mb/s 以下。这个标准主要是针对当时具有这种数据传输率的 CD-ROM 和网络而开发的，用于在 CD-ROM 上存储数字影视和在网络上传输数字影视。

MPEG-2 标准是一个直接与数字电视广播有关的高质量图像和声音编码标准，是 MPEG-1 的扩充。MPEG-2 提供位速率的可变性能功能，其最基本目标是：位速率为 4~9 Mb/s，最高达 15 Mb/s。

MPEG-4 是为视听数据的编码和交互播放开发算法和工具，是一个数据速率很低的多媒体通信标准。MPEG-4 的目标是要在异构网络环境下能够高度可靠地工作，并且具有很强的交互功能。

MPEG-7 的名称叫做多媒体内容描述接口，目的是制定一套描述符标准，用来描述各种类型的多媒体信息及它们之间的关系，以便更快更有效地检索信息。这些媒体材料可包括静态图像、图形、3D 模型、声音、语音、电视以及在多媒体演示中它们之间的组合关系。在某些情况下，数据类型还可包括面部特性和个人特性的表达。

## 第12章MPEG声音

### 12.1 列出你所知道的听觉系统的特性。

响度感知, 音高感知, 掩蔽效应

## 12.2 什么叫做听阈? 什么叫做痛阈?

(1) 当声音弱到人的耳朵刚刚可以听见时, 称此时的声音强度为“听阈”

(2) 当声音强到人的耳朵刚刚感到疼痛时, 称此时的声音强度为“听阈”。实验表明, 如果频率为 1 kHz 的纯音的声强级达到 120 dB 左右时, 人的耳朵就感到疼痛, 这个阈值称为“痛阈”。

## 12.3 什么叫做频域掩蔽? 什么叫做时域掩蔽?

强纯音掩蔽在其附近同时发声的弱纯音, 这种特性称为频域掩蔽。

(2) 在时间方向上相邻声音之间的掩蔽, 这种特性称为时域掩蔽。

## 12.4 MPEG-1的层1、2和3编码分别使用了听觉系统的什么特性?

层 1: 频域掩蔽特性

层 2: 频域掩蔽特性, 时间掩蔽特性

层 3: 频域掩蔽特性, 时间掩蔽特性

## 12.5 MPEG-1的层1、2和3编码器的声音输出速率范围分别是多少?

层 1: 384 kb/s

层 2: 256~192 kb/s

层 3: 128~112 kb/s

## 12.6 MPEG-1的声音质量是: AM FM 电话 near-CD CD-DA

☐ AM ☐ FM ☐ 电话 ☒ near-CD ☐ CD-DA

## 12.7 什么叫做5.1声道立体环绕声? 什么叫做7.1声道立体环绕声?

(1) “5.1 环绕声”也称为“3/2-立体声加 LFE”, 其中的“.1”是指 LFE 声道。它的含义是播音现场的前面可有 3 个喇叭声道(左、中、右), 后面可有 2 个环绕声喇叭声道, LFE 是低频音效的加强声道。

7.1 声道环绕立体声与 5.1 类似。

参见图 12-19。

## 12.8 简述MPEG-2 AAC的特性。

MPEG-2 AAC 是 MPEG-2 标准中的一种非常灵活的声音感知编码标准。它使用听觉系统的掩蔽特性来减少声音的数据量, 并且通过把量化噪声分散到各个子带中, 用全局信号把噪声掩蔽掉。

AAC 支持的采用频率可从 8 kHz 到 96 kHz, AAC 编码器的音源可以是单声道的、立体声的和多声道的声音。AAC 标准可支持 48 个主声道、16 个低频音效加强通道 LFE、16 个配音声道或者叫做多语言声道和 16 个数据流。

MPEG-2 AAC 在压缩比为 11:1, 即在每个声道的数据率为 $(44.1 \times 16)/11=64$  kb/s, 而 5 个声道的总数据率为 320 kb/s 的情况下, 很难区分还原后的声音与原始声音之间的差别。与 MPEG-Audio 层 2 相比, MPEG-2 AAC 的压缩率可提高 1 倍, 而且质量更高, 与 MPEG 的层 3 相比, 在质量相同的条件下数据率是它的 70%。

## 12.9 什么叫做自然声音? 什么叫做合成声音?

自然声音通常是指人、动物等发出的声音、弹奏乐器或其他音源自然发出的声音。

合成声音通常指合成器生成的声音, 如 MIDI。

## 12.10 什么叫做TTS? 至少列举TTS的3个潜在应用例子。

通过话音合成设备自动地把基于文字的数据转换为声音输出的一种技术。文语转换的最终目标是要使计算机像人一样输出清晰而又自然的声音, 也就是说, 根据文本的内容可以不同的情调来朗读任意的文本。TTS 是一个十分复杂的系统, 涉及到语言学、语音学、信号处理、人

人工智能等诸多的学科。尽管现有的 TTS 系统结构各异，转换方法不同，但是基本上可以分成两个相对独立的部分：①文本分析，通过对输入文本进行词法分析、语法分析，甚至语义分析，从文本中抽取音素和韵律等发音信息。②语音合成，使用从文本分析得到的发音信息去控制合成单元的谱特征(音色)和韵律特征(基频、时长和幅度)，送入声音合成器(软件或硬件)产生相应的语音输出。

通过电话访问信息，朗读文本，盲人计算机等。

## 第13章 MPEG 电视

### 13.1 电视图像数据压缩的依据是什么？

根据人的视觉特性和电视图像数据自身的冗余特性。

注：人的视觉系统具有的两种特性可以用来压缩电视图像数据。一是人眼对色度信号的敏感程度比对亮度信号的敏感程度低，利用这个特性可以把图像中表达颜色的信号去掉一些而使人不易察觉；二是人眼对图像细节的分辨能力有一定的限度，利用这个特性可以把图像中的高频信号去掉而使人不易察觉。

### 13.2 MPEG-1 编码器输出的电视图像的数据率大约是多少？

1.15 Mb/s

### 13.3 MPEG 专家组在制定 MPEG-1/-2 Video 标准时定义了哪几种图像？哪种图像的压缩率最高？哪种图像的压缩率最低？

(1) MPEG 专家组定义了三种图像：帧内图像 I(intra)，预测图像 P(predicted) 和双向预测图像 B(bi-directionally interpolated)。

(2) 双向预测图的压缩率最高，帧内图像的压缩率最低。

### 13.4 有人认为“图像压缩比越高越好”。你对这种说法有何看法？

在图像压缩算法中，为了获得比较高的压缩率，通常要采用有损压缩。这就意味压缩率越高，图像的质量损失越大，重构图像的质量将会越低。

### 13.5 有人说“MPEG-1 编码器的压缩比大约是 200:1”。这种说法对不对？为什么？

不对。理由：(1)电视图像的子采样(4: 1: 1)不是 MPEG 编码器的功能，(2)电视图像画面尺寸(如 PAL 制  $720 \times 576$ )的降采样(变成  $360 \times 288$ )也不属于 MPEG 编码器的功能。

### 13.6 说明电视规格 MP@ML 和 HP@HL 各自的含义。

MP@ML (Main Profile, Main Level)指的是具有这种特性的电视：分辨率为  $720 \times 576 \times 30$ ，子采样格式为 4:2:0，位速率可高达 15 Mb/s。

HP@HL (High Profile, High Level)指的是具有这种特性的电视：分辨率为  $1920 \times 1152 \times 60$ ，子采样格式为 4:2:0 或者 4:2:2，位速率可高达 80 Mb/s。

### 13.7 电视图像的空间分辨率和时间分辨率是什么意思？

空间分辨率：图像的尺寸。例如，PAL 制图像尺寸： $720 \times 576$ ， $360 \times 288$ 。

时间分辨率：图像的帧数/每秒钟。例如，PAL 制电视图像的时间分辨率为 30 帧/秒，NTSC 电视图像的时间分辨率为 25 帧/秒，电影图像的时间分辨率为 24 帧/秒。

## 第14章 只读光盘存储器

### 14.1 只读光盘是如何记录“0”和“1”的？

只读光盘通常是指 CD/DVD 系列只读光盘。在盘上压制凹坑的机械办法来记录“0”和“1”。凹坑的边缘代表“1”，凹坑和非凹坑的平坦部分代表“0”，凹坑的长度和非凹坑的长度都代表有多少个“0”。



#### 14.2 CD-DA的音乐信号的采样频率为什么选择44.1 kHz?

人耳朵(因人而异)能听到的声音信号频率范围是 20~20 000 Hz, 为了避免高于 20 000 Hz 的高频信号干扰采样, 在进行采样之前, 需要对输入的声音信号进行滤波。考虑到滤波器在 20 000 Hz 的地方大约有 10%的衰减, 所以可以用 22 000 Hz 的 2 倍频率作为声音信号的采样频率。但是, 为了能够与电视信号同步, PAL 电视的场扫描为 50 Hz, NTSC 电视的场扫描为 60 Hz, 所以取 50 和 60 的整数倍, 选用了 44 100 Hz 作为激光唱盘声音的采样标准。

#### 14.3 激光唱盘音乐信号的样本位数是16, 它的信噪比是多少? 如果样本位数提高到20, 它的信噪比是多少?

- (1) 96dB
- (2) 120 dB

#### 14.4 为什么物理线路上传输的数字信号都需要采用通道编码?

主要原因有两个, 一是为了改善读出信号的质量, 二是为了在记录信号中提取同步信号。

(例如, 有连续多个字节的全“0”信号或者全“1”信号要记录到盘上, 如果不作通道编码就把它们记录到盘上, 读出时的输出信号就是一条直线, 电子线路就很难区分有多少个“0”或者多少个“1”信号。而对于没有规律的数字信号, 读出时的信号幅度和频率的变化范围都很大, 电子线路很难把“0”和“1”区分开, 读出的信息就很不可靠。因此通俗说来, 通道编码实际上就是要在连续的“0”插入若干个“1”, 而在连续的“1”之间插入若干个“0”, 并对“0”和“1”的连续长度数目即“行[游]程长度”加以限制。)

#### 14.5 CD盘中的EFM是什么意思?

EFM (eight-to-fourteen modulation)是指由 8 比特的代码调制成 14 比特的代码。

(解释: 它是一种物理通道编码技术, 使用这种技术可增加 CD 盘上信息凹坑和非凹坑的长度, 降低读出信号的频率带宽。CD 使用脉冲调宽技术, 把 8 比特的用户数据变换成 14 比特的通道码。在红皮书和 IEC 908 标准中, 在 14 比特的码之间增加 3 个比特的合并位, 构成 8 比特到 17 比特调制编码。在物理存储器上存储数据或者在物理线路上传输数字信号都需要对用户数据进行通道编码。物理盘上记录的数据和真正的用户数据之间需要做变换处理, 这种处理统称为通道编码。采用通道编码的目的主要是两个, 第一是为了改善信号质量, 使得读出信号的频带变窄; 第二是为了在接收端能够从信号本身提取自同步信号。)

#### 14.6 激光唱盘播放机的声音数据传输率是多少?

$$44100 \times 16 \times 2 = 1.4112 \text{ Mb/s}$$

#### 14.7 从CD渡到DVD, 科学家和工程技术人员采取了那些主要技术?

详见“表 14-06 HDCD 技术摘要”。

#### 14.8 目前定义了几种DVD?

详见表“14-03 DVD 和 CD 系列”。

## 第15章光盘存储器的格式

#### 15.1 试论CD标准的重要性。

(略)

#### 15.2 什么叫做CD盘的物理格式?

存储媒体的存储结构, 它规定数据如何放在光盘上。这些数据包括物理扇区的地址、数据的类型、数据块的大小、错误检测和校正码等。

#### 15.3 CD-ROM的扇区地址“00:10:65”表示什么含义?

CD-ROM 的扇区地址用“分、秒, 分秒(1/75 秒)”来表示。00:10:65 表示地址为 00 分 10

秒 65 分秒。

#### 15.4 CD-ROM Mode 1和CD-ROM Mode 2有什么差别？

CD-ROM Mode 2 与 CD-ROM Mode 1 相比，存储的用户数据多 14%，但是由于没有错误检测和错误校正码，因此在这种方式中，用户数据的误码率比 Mode 1 中的误码率要高。在 Mode 2 的扇区地址中，方式 (Mode) 字节域中的值设置成 02，在 Mode 1 的扇区地址中，方式 (Mode) 字节域中的值设置成 01。

#### 15.5 CD-ROM/XA Mode 2 Form 1和CD-ROM/XA Mode 2 Form 2有什么差别？

CD-ROM/XA Mode 2 Form 2 与 CD-ROM/XA Mode 2 Form 1 相比，去掉了 276 字节的 ECC 码，因此存储的用户数据增加了 276 字节。

#### 15.6 CD-Bridge Mode 2 Form 1和CD-Bridge Mode 2 Form 2有什么差别？

D-Bridge Mode 2 Form 2 和 CD-Bridge Mode 2 Form 1 相比，去掉了 276 字节的 ECC 码，因此存储的用户数据增加了 276 字节。

#### 15.7 计算单速CD-ROM的用户数据传输率是多少KB/s(1 KB = 1024 bytes)。

$$(2048 \times 75) / 1024 = 153600 / 1024 = 150 \text{ KB/s}$$

#### 15.8 CD-ROM文件系统的主要组成部分是什么？

一个完整的 CD-ROM 文件系统有三个主要部分组成：

(1) 逻辑格式(logical format): 它是文件格式的同义词。逻辑格式是确定盘上的数据应该如何组织，以及存放在什么地方。说得具体一点就是基本的识别信息放在何处，文件目录应该如何构造，到何处去找盘上的目录，一个应用软件存放在几张光盘上等等。由此也可以看到，逻辑格式与物理格式是不同的。

(2) 源软件(origination software): 它是把数据写到逻辑格式的软件，按逻辑格式把要存到盘上的文件进行装配，所以源软件又称“写”软件。

(3) 目的软件(destination software): 它是把数据从逻辑格式读出来，并且把数据转换成文件，因此目的软件又称为“读”软件。它在终端用户的机器上能够理解逻辑格式，并且使用逻辑格式来访问盘上的文件。

#### 15.9 CD-ROM的逻辑格式是什么意思？

文件格式的同义词。它指定数据文件在盘上的组织和位置，定义文件大小、所有文件在盘上的目录结构以及所需盘片数目等事项。

#### 15.10 CD-ROM的物理扇区、逻辑扇区和逻辑块之间有什么关系？

详见图 15-05。

#### 15.11 MS-DOS和MS-Windows环境下都要有MSCDEX.EXE文件，请问它的功能是什么？

把 ISO 9660 文件结构转变成 MS-DOS 能识别的文件结构。

#### 15.12 试论CD-ROM的文件目录结构与磁盘的文件结构有何差别？

CD-ROM 是只读存储器，而磁盘是可读/写存储器，这就不需要添加、删除目录等功能。

CD-ROM 的写入和读出是不对称的。

寻找数据的时间相对于 20 ms 左右的磁盘大得多。

#### 15.13 用计算机查看VCD盘上根目录和子目录下有什么文件？这些文件的含义是什么？

(略)

## 第16章错误检测和校正

#### 16.1 CRC用于检测错误还是校正错误？

用于检测错误。



16.2 用自己的语言说明错误检测的思想。

(略)

16.3 什么叫做突发错误?

连续多位或连续多个符号都出错的一种错误模式。如盘片的划伤、沾污或盘本身的缺陷都可能出现这种错误,一错就错一大片。

16.4 码块长度为 $n$ ,码块中的信息长度为 $k$ ,问 $(n, k)$ RS码本身能纠正多少个错误?

$(n-k)/2$

16.5 要纠正1个符号的错误,至少需要附加多少个校验符?

2个,用于表示错误的位置和错误的值

16.6 目前CD存储器中使用的CIRC编码技术能够纠正突发错误的最大长度是多少(按汉字字数估算)?

大约连续 224 个汉字的长度。

## 第17章多媒体网络应用及交换技术

17.1 网络上的多媒体通信应用和数据通信应用有什么主要差别?

多媒体通信应用要求在客户端播放声音和图像时要流畅,声音和图像要同步,因此对网络的时延和带宽要求很高。而数据通信应用则把可靠性放在第一位,对网络的时延和带宽的要求不那么苛刻。

17.2 举例说明多媒体网络应用的特点。

例如 IP 电话,要求时延短。

17.3 因特网电话和实时电视会议是多媒体网络应用中频繁交互的应用例子,根据人的听觉系统,对延迟小于150 ms的声音感觉不到有时延,在150 ~ 400 ms之间的时延可以接受,时延超过400 ms的会话就令人甚感别扭。

17.4 T1数字网络的总数据传输率,美国定义为1.544 Mb/s,欧洲的T1称为E1,定义为2.048 Mb/s。

17.5 借助工具书或者技术参考书的帮助理解信息包(packet)的含义。

(1) 信息传输的单位,在网络中作为一个整体从一个设备传向另一个设备。

(2) 在信息包交换网络中,固定的最大长度的传输单位,由二进制数位表示的数据及信息包头组成。信息包头中含有识别号、源地址、目的地址及一些错误控制信息。有时作为通用词用来描述协议堆栈各层的信息传输单位。

(3) 将信息分成小块,并加上路由及控制信息之后的数据。这些信息包能独立地传送,在目的地重组。几乎所有的高速数字网都采用了信息包技术。它同时也是因特网 TCP/IP 协议的基础。

(4) 在 DVD 系统中,每个数据小包由 2048 字节组成。

17.6 总结线路交换、信息包交换和消息交换各有什么特点。

线路交换(circuit switching)的特点是,在开始通信之前通信双方由线路交换中心建立物理连接,维持连接的时间长短取决于消息交换的需要

信息包交换不需要在收发双方建立物理连接,每个信息包都包含有目的地址,因此一个消息分装成的许多信息包不必都沿着同一条线路到达目的地,也不必同时到达目的地,到达目的地的次序也不必按照发送的次序,那条信道有空就往那里传送。

消息交换把整个原始消息经有网络传送到接收方。它的时延比信息包交换长,错误处理时间也比较长。

17.7 面向连接服务和无连接服务的差别是什么?

面向连接服务：在开始发送信息包之前发送端和接收端要进行沟通，建立直接连接，并提醒对方准备接收信息包，然后才开始进入信息包的传送过程。

无连接服务：使用用户数据包协议(UDP)来调用无连接服务。发送端简单地把信息包送到网络上，在传送信息包之前发送端和接收端没有沟通的过程，也没有对方来的确认，因而也不知道目的地是否接收到。无连接服务既没有拥挤控制功能，也没有流程控制功能。

## 第18章因特网与TCP/IP

### 18.1 查阅一些文献资料(如工具书等)，分析和比较网关(gateway)、路由器(router)和网桥(bridge)在网络中所起的作用。

(1) 网关：路由器的早期的名称，许多人仍然使用这个术语表示 OSI 模型第三层的设备。现在，网关一般指 OSI 模型第四层的设备。网关是连接使用不同通信协议的一种功能强大的计算机或工作站。它担负线路交换网络(如电话网络)和信息包交换网络(如因特网)之间进行实时双向通信，提供异种网络之间的连通性，是传统线路交换网络 and 现代 IP 网络之间的桥梁。网关的基本功能可归纳为三种。①转换协议：网关作为一个解释器能够使不同的网络建立联系，②转换信息格式：不同的网络使用不同的编码方法，网关对信息进行转换，使异种网络之间能够自由地交换信息；③传输信息：负责在不同网络之间传输信息。

网关的主要部件包括交换电路网络接口卡、数字信号处理器卡、网络接口和控制处理器；网关的主要软件包括执行所有网关基本功能和选择功能的网关软件，以及特定网关的应用软件，它执行自定义的功能以及管理和控制功能。下面的网关基本结构图表示如何使简易老式电话服务(POTS)系统上的电话与现代的因特网电话之间进行会话。

(2) 路由器：在通信系统中，在网络间转发信息包的一种中介设备。根据路由表和路由协议，路由器读每一个传输帧中的网络地址，然后按照最合适的路径(包括交通负荷、线路成本、传送速度和线路质量等)来确定转发信息包的路径。路由器工作在协议堆的层 3，而网桥和交换机工作在层 2。路由器可用来把局域网分成若干网段以便平衡工作组内的交通，也可用于过滤信息包以提高网络安全和政策性的管理，还可用于在网络边沿用来连接远端的办公室。大多数路由器是对通信功能进行了优化的专门计算机，但是路由器的功能也可通过把路由软件添加到文件服务器(file server)来完成。

(3) 网桥：连接两个局域网(不论它们是否使用相同的通信协议)的一种设备。网桥工作在数据链路层(ISO/OSI 模型的第二层)；网桥是协议独立的；网桥的速度比路由器的速度快；网桥有两个以上端口执行交换功能。

### 18.2 什么叫做IP(Internet Protocol)地址，什么叫做域名地址(domain name address)，它们之间有什么关系。

IP address (Internet Protocol address)的中文译名是 IP 地址，或者叫做网际协议地址。IP 地址是网际协议(Internet Protocol, RFC 791)定义的、唯一标识连接在 TCP/IP 网络上的主机的地址。主机是指能够为其他计算机提供服务的计算机，也称为服务机。在因特网上，每台连网的客户机和服务机都必须要有个唯一的地址，这样才能在计算机之间进行通信。客户机可以申请永久性的地址，也可以动态分配给它的地址。因特网定义了两种形式的地址：一种是机器可识别的用数字表示的地址，称为网际协议地址，简称为 IP 地址；另一种是人比较容易看懂的用字母表示的地址，称为域名地址(domain name address)。IP 地址由美国国家科学基金会于 1993 年组成的因特网信息中心注册服务部门(InterNIC Registration Services)进行分配和注册。IP 地址在扩充之前是一个 32 位(4 字节)的二进制数。它几乎能标识  $2^{32}$ (大约 40 亿)台主机。

### 18.3 IP地址有哪几类？总结它们各有什么特点？

IP 地址分成 5 类：A 类(Class A), B 类(Class B), C 类(Class C), D 类(Class D)和 E 类(Class E)。其中 A、B 和 C 类地址是基本的因特网地址，是用户使用的地址，D 类地址用于多目标广播的广播地址，E 类地址为保留地址。详见图 18-04。

18.4 分别说出IP地址为167.216.130.62和202.98.36.120是哪一类网络？它们的网络数目和连网机器数目各有什么特点。

167.216.130.62:  $167(10)=10100111(B)$ , B 类网络, 网络地址的数目为中等, 主机地址数目也属中等。

202.98.36.120:  $202(10)=11001010(B)$ , C 类网络, 网络地址的数目较多, 主机地址数目较少。

18.5 域名服务(Domain Name Service, DNS)的功能是什么？

可让用户通过域名来查找 UNIX 网络或因特网(TCP/IP 网络)上的计算机。这种功能是通过 DNS 服务器(DNS server)软件来实现, 它维护一个含有域名(主机名)和它们相应 IP 地址的数据库。实质上是维护一个两列的查找表, 其中一列是帮助记忆的计算机名字(如 [www.tsinghua.edu.cn](http://www.tsinghua.edu.cn)), 另一列是用数字表示的 IP 地址(如 166.111.9.2), 计算机名字和它的 IP 地址是相对应的。DNS 服务器也称为名字服务器(name servers)。对于使用 TCP/IP 协议堆的 Windows 网络, 域名服务软件是 Windows 因特网命名服务软件(WINS)。在 Windows 网络中, 只需使用 WINS, 而对混合的 Windows/UNIX 网络, 则需要使用可在 Windows 和 UNIX 下运行的微软 DNS 服务器

18.6 举例说明统一资源地址(Uniform Resource Locator, URL)的含义。

例如, 清华大学图书馆里的数据库目录文件, 用 URL 表示成:

<http://www.lib.tsinghua.edu.cn/NEW/database.html>

其中, [www.lib.tsinghua.edu.cn](http://www.lib.tsinghua.edu.cn) 表示图书馆的服务机, 而 /NEW/database.html 表示 database.html(数据库文件)是在服务机上的 NEW 文件夹下。

18.7 用简洁的语言说出“<http://www.mpeg.org/~tristan/MPEG/MPEG-content.html>”的意思。

MPEG-content.html(文件)在 [www.mpeg.org](http://www.mpeg.org)(服务机)的 ~tristan/MPEG(文件夹)下。

18.8 在网络上下载详细描述HTTP/1.0协议的RFC 1945文件, 试一试看自己读懂其中的多少内容。

(略)

18.9 TCP和UDP是哪一层的协议？这两种协议的有什么主要差别？

(1) TCP 和 UDP 都是在传输层上的协议。不同的网络应用使用不同的协议

(2) TCP 是面向连接服务的协议, 在开始发送信息包之前发送端和接收端要进行沟通, 建立直接连接, 并提醒对方准备接收信息包, 然后才开始进入信息包的传送过程。

UDP 是无连接服务的协议, 发送信息时, 发送端简单地把信息包送到网络上, 在传送信息包之前发送端和接收端没有沟通的过程, 也没有对方来的确认, 因而也不知道目的地是否接收到。无连接服务既没有拥挤控制功能, 也没有流程控制功能。

18.10 网络技术人员对因特网的层次结构和各层使用的协议了如指掌, 网络用户对此也需要了解。若因特网粗略划分成5层, 请写出各层的名称, 各层使用的协议(写出1至2个协议即可)

请参考图 18-08 和 18-09。

18.11 写一篇关于以太网的性能调查报告, 网络类型包含10Base-F, 10BASE-T, 10Base2(也称thin Ethernet), 10Base5(也称Thick Ethernet, ThickNet, ThickWire), 100BaseT以及千兆以太网, 内容包含使用的标准、使用的电缆和电缆连接器、传输的距离、数据传输率、应用场合等, 并把要点综合成表格。

(略)

## 第19章网际多目标广播简介

18.1 什么叫做单目标广播(unicast)和多目标广播(multicast)? 什么叫做网际多目标广播(IP

### Multicast)?

(1) 目标地址只有一个的一种广播技术。如果使用单目标广播服务把相同内容的信息传输给 N 个目标站点,就须要传输 N 个拷贝,即要传输 N 次。

(2) 在数据通信网络中,将信息包同时拷贝传送到一组选定的网络地址,然后发送给每个接收终端的一种广播技术。

(3) 在 TCP/IP 网络(互联网、内联网或因特网)上,把单一数据流同时向多个接收者发送的一种广播技术。用于广播声音和电视,也用于向多个用户下载同一个文件。使用 IP 多目标广播可节省网络带宽,因为文件作为一个数据在主干网上传输,仅在线路终端分开向目标站点传输。

### 18.2 多目标广播使用哪类地址?

D 类地址

18.3 对有兴趣的读者,试阅读并比较下列两个协议: ① 1997年的RFC 2236: B. Fenner, "Internet Group Management Protocol, Version 2". ② 1989年的RFC 1112: S. Deering, "Host extensions for IP multicasting"

(略)

18.4 对有兴趣的读者,阅读声音-电视传输工作组(Audio-Video Transport Working Group)编写的实时传输协议RTP和实时传输控制协议RTCP文件: RFC 1889: H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", 01/25/1996. 并写出摘要。(略)

## 第20章多媒体通信系统技术

### 20.1 ISDN网络和PSTN网络属于线路交换网络还是属于信息包交换网络?

ISDN 网络和 PSTN 网络属于线路交换网络。

### 20.2 目前的多媒体通信系统中的两个重要标准系列是什么? 它们各自要达到什么目标?

(1) H.323(V1/V2)和 H.324。

(2) H.323: 在局域网上可召开桌面电视会议(desktop video conferencing); H.324: 使用调制解调器在公众交换电话网(PSTN)上召开网桌面电视会议。

### 20.3 分别列出H.323和H.324标准系列涉及的重要标准,并分别说明这些标准的用途。

详见表 20-02, 20-03 和 20-04。

### 20.4 什么叫做网关? 网关的主要功能是什么?

详见“20.1.2 网关的功能和结构”

### 20.5 什么叫做会务器? 会务器的主要功能是什么?

为国际电信联盟(ITU)制定的 H.323 多媒体通信标准开发的一种服务器,用于连接 IP 网络上的 H.323 电视会议端点,包括终端、网关和多点控制单元(MCU),为注册的端点提供呼叫控制服务。

### 20.6 H.323定义了哪几种形式的多点多媒体会议?

H.323 标准可支持的会议形式包括: ①由 **D**、**E** 和 **F** 终端参加的集中式电视会议, ②由 **A**、**B** 和 **C** 终端参加的分散式电视会议, ③声像集散混合式多点电视会议, ④会议集散混合式多点电视会议。详见“20.2.6 H.323 多点电视会议”

20.7 如果物质条件允许,上网调查可视电话的技术和发展趋势,然后写一篇调查报告。内容包括: ①使用的标准, ②各种标准的用途, ③列表比较你比较喜欢的三种可视电话终端的外形图、性能指标和价格, ④发表你对可视电话的评论。

(略)

20.8 如果物质条件允许，上网找一个IP电话试用软件，安装在两台多媒体计算机上，体验一下IP电话的质量，并分析有可能出现的问题，并发表你对IP电话的评论。

(略)

## 第21章超文本标记语言(HTML)

### 21.1 谁在制定Web标准?

- (1) 网景公司(Netscape)
- √ (2) 万维网协会(The World Wide Web Consortium)
- (3) 微软公司

### 21.2 哪一个HTML标签用于插入断行?

- (1) <lb>
- (2) <break>
- √ (3) <br>

### 21.3 哪一个HTML标签用于添加背景颜色?

- (1) <background>yellow</background>
- √ (2) <body bgcolor="yellow">
- (3) <body color="yellow">

### 21.4 哪一个HTML标签用于粗体字?

- (1) <bold>
- (2) <bb>
- (3) <b>
- √ (4) <bld>

### 21.5 哪一个HTML标签用于链接“第29届奥林匹克运动会组织委员会”?

- (1) <a>http://www.beijing-2008.org, 第 29 届奥林匹克运动会组织委员会</a>
- (2) <a url="http://www.beijing-2008.org">第 29 届奥林匹克运动会组织委员会</a>
- √ (3) <a href="http://www.beijing-2008.org">第 29 届奥林匹克运动会组织委员会</a>
- (4) <a name="http://www.beijing-2008.org">第 29 届奥林匹克运动会组织委员会</a>

### 21.6 哪一个代码用于链接电子邮件程序?

- (1) <mail href="xxx@yyy">
- (2) <a href="xxx@yyy">
- √ (3) <a href="mailto:xxx@yyy">
- (4) <mail>xxx@yyy</mail>

### 21.7 如何把被链接的文档显示在新的浏览器窗口中?

- (1) <a href="url" new>
- (2) <a href="url" target="new">
- √ (3) <a href="url" target="\_blank">

### 21.8 哪一个标签用于插入图像?

- (1) <image src="image.gif">
- (2) <img href="image.gif">
- √ (3) 
- (4) <img>image.gif</img>

### 21.9 哪一个标签用于插入背景图像?



- (1) 
- (2) <background img="background.gif">
- √(3) <body background="background.gif">

## 21.10 为什么要定义字体的逻辑样式和物理样式？

因为在 SGML 语言中“内容”和“内容的表达方法”是分开的，HTML 是 SGML 中的子集，它继承了 SGML 的特性。例如，SGML 标记“HTML 概要”为一级标题时，并没有指定一级标题中的“HTML 概要”的字体大小(如四号字)和对齐方式(如居中)。这样做的好处是：如果想改变一级标题中的字体(如改为三号字体)和对齐方式(如改为左对齐)时，你所要做的仅仅是改变一级标题的定义。

物理样式和逻辑样式各有它们自己的优点。使用逻辑标签的优点是可强制文档中同一元素的类型的一致性。例如，标记某标题为<h1>比你去记住一级标题的字号、对齐方式等要容易得多，这样就不容易记错，保证了一级标题的一致性；使用物理样式的好处是浏览器会严格遵照你标记的样式显示文本，因此在不希望浏览器改变样式的情况下就应该使用物理格式。

## 21.11 指出“<h1 align="center">html编辑器</h1>”中哪些是元素、标签和属性。

- “<h1>..... </h1>”是标签
- “align=.....”是属性
- “center”是属性值
- “html 编辑器”是元素

## 21.12 一个URL为<http://www.tsinghua.edu.cn/docs/cindex.htm>，说出各部分的含义。

详见“18.1.5 统一资源地址”。

## 21.13 电子邮件的地址为：[liudf@public.bta.net.cn](mailto:liudf@public.bta.net.cn)和[bookshelf@mcp.com](mailto:bookshelf@mcp.com)，分别说出各部分的含义。

Liudf 和 bookshelf 表示电子邮件的账户名，public.bta.net 和 mcp.com 表示电子邮件服务器地址。

## 21.14 把图像/影视文件链接到和嵌入到HTML文档中，在语法结构和工作方式上都有什么差别？分别说出各自的优缺点。

详见“21.4.6 链接内联图像”和“21.4.9 链接外联图像”。

提示：内联图像(inline image)是指与 Web 网页中的文本一起下载和显示的图像，而不需要在你选择之后才下载和显示，在浏览器上表现为文本和图像显示在同一网页上，这种图像也称为内嵌图像或者叫内插图像。

在 HTML 文档中，如果内联图像文件太大，在浏览主文档时就需要很长的时间来等待图像下载。为了改善这种状况，一种切实可行的方法是把大的图像当作外联(external image)图像来处理。在编写 HTML 文档时，用文字或者小的内联图标来代表大图像，而把大图像文件当作一个单独的文档，然后把文字或图标与大图像文件链接在一起。当在浏览器中阅读网页时，读者可用鼠标器激活这个链接，然后下载并显示在另一个窗口上。用这种方式链接的图像称为外联图像。

## 21.15 把声音文件链接到和嵌入到HTML文档中，在语法结构和工作方式上都有什么差别？分别说出各自的优缺点。

详见：“21.4.10 在文档中链接和嵌入声音文件”。

## 21.16 列出你和你的同学所知道的HTML编辑器，并讨论它们的性能。

Microsoft FrontPage, Macromedia Dreamweaver, 任何文字编辑器。

# 第22章可扩展标记语言(XML)

## 22.1 XML的目标是要取代HTML

- (1) 对
- √(2) 错

## 22.2 用XML可定义多少种不同的数据类型?

没有上限

## 22.3 说出下面的XML元素由哪几部分组成、各部分的名称和元素的值。

`<name>Your Name</name>`

元素名称: name

元素: `<name>Your Name</name>`

标签: `<name> </name>`

元素值: YourName

## 22.4 属性值必要用双引号

- (1) 对
- √(2) 错

## 22.5 声明XML版本的正确语法是

- (1) `<xml version="1.0" />`
- (2) `<?xml version="1.0" />`
- √(3) `<?xml version="1.0"?>`

## 22.6 XML使用哪些方法描述XML数据

- (1) 使用 XSL 描述数据。
- √(2) 使用 DTD 描述数据。
- √(3) 使用 XML Schema 描述数据。

## 22.7 下面的陈述哪一个是真的:

- (1) 所有 XML 文档必需有一个 DTD
- (2) 所有 XML 元素必需是小写字母
- √(3) 所有 XML 元素必需要有结束标签

## 22.8 正确引用"mystyle.xml"的方法是

- (1) `<?xml-stylesheet type="text/xsl" href="mystyle.xml" ?>`
- (2) `<link type="text/xsl" href="mystyle.xml" />`
- √(3) `<stylesheet type="text/xsl" href="mystyle.xml" />`

## 22.9 仿照例22.9-1、例22.8-2和例22.9-3的代码,用IE 5.0以上浏览器显示下列数据(至少5名学生的数据)。

通讯录

学号	姓名	性别	电子邮件	籍贯
----	----	----	------	----

(略)

## 第23章可扩展超文本标记语言(XHTML)

### 23.1 XHTML是什么?

XHTML 是用 XML 格式重新改写的 HTML, XHTML 1.0 是 HTML4.01 的改写版本。XHTML 使用的词汇是 HTML 的词汇,使用的语法是 XML 的语法。



## 22.2 XHTML是Web标准?

是

## 22.3 XML和HTML要被XHTML取代吗?

不

## 22.4 HTML要被XHTML取代吗?

是

## 23.5 在下列标签中, 哪一个是正确的XHTML段落标签?

- √ (1) `<p></p>`
- (2) `<P></P>`
- (3) `<P></p>`
- (4) `</p><p>`

## 23.6 在下列标签中, 哪一个是正确的XHTML断行标签?

- √ (1) `<br />`
- (2) `<br>`
- (3) `<break/>`

## 23.7 在下列属性和属性值中, 哪一个是正确的XHTML 的属性和属性值?

- √ (1) `width="80"`
- (2) `WIDTH="80"`
- (3) `WIDTH=80`
- (4) `width=80`

## 23.8 XHTML中的所有元素必需是闭合的吗?

- √ (1) 对
- (2) 错

## 23.9 DOCTYPE声明不用结束标签, 对吗?

- √ (1) 对
- (2) 错

## 23.10 在XHTML文档中, 下列元素中哪一个是强制性的?

- √ (1) doctype, html, head, body 和 title
- (2) doctype, html, head, 和 body
- (3) doctype, html 和 body

## 23.11 下列代码是不是合格的XHTML代码?

- (1) `<p>A <b><i>short</b></i> paragraph</p>`
- (2) `<p>A <b><i>short</i></b> paragraph`
- √ (3) `<p>A <b><i>short</i></b> paragraph</p>`

## 23.12 在下列代码中, 哪一个是正确的语言引用?

- √ (1) `<div lang="en" xml:lang="en">Hello World!</div>`
- (2) `<div xml:lang="en">Hello World!</div>`
- (3) `<div lang="en">Hello World!</div>`

## 23.13 在XHTML文档中允许属性最小化?

- (1) 是
- √ (2) 否

#### 23.14 所有XHTML文档都需要doctype?

- √(1) 是
- (2) 否

### 第24章使用JavaScript

24.1 使用HTML编辑器重新输入本章所示的所有源程序，或者根据读者自己的水平有选择地输入比较难理解的源程序，自己设计程序中没有提供的HTML文档、图和AVI等文件，然后使用Web浏览器检查程序的准确性。

(略)