

第3讲 视频数据的压缩与编码

2008 硕士研究生课程



主讲：于筱清



内容提要

<http://media.hust.edu.cn>

- ☐ 多媒体数据压缩的必要性和可能性
- ☐ 多媒体数据压缩编码的分类
- ☐ PCM与预测编码
- ☐ 无损压缩编码
- ☐ 变换编码
- ☐ JPEG
- ☐ MPEG

数字媒体实验室

2

§ 3.1 多媒体数据压缩的必要性和可能性

<http://media.hust.edu.cn>

- ☐ 从目前计算机的软硬件发展水平及发展趋势来看，可以断言：
 - 在将来很长的一段时期内，数字化的媒体信息数据以压缩形式存储和传播仍将是唯一的选择
 - 数据压缩的必要性
 - 数据压缩的可能性

数字媒体实验室

3

数据压缩的必要性

<http://media.hust.edu.cn>

- ☐ 信息时代的重要特征是信息的数字化，数字化了的信息带来了“信息爆炸”
- ☐ 数字化了的视频和音频信号的数量之大是非常惊人
- ☐ 举例说明

数字媒体实验室

4

举例一：

<http://media.hust.edu.cn>

- ☐ 双通道立体声激光唱盘 (CD-DA)，采样频率为44.1kHz，采样精度16位/样本，其一秒钟时间内的采样位数为 $44.1 \times 103 \times 16 \times 2 / 8 = 0.176 \text{ MB/S}$ 。一个650MB的CD-ROM，可存约1小时的音乐。

数字媒体实验室

5

举例二：

<http://media.hust.edu.cn>

- ☐ 数字电视图像 SIF (source input format) 格式，NTSC制、彩色、4:4:4采样每帧数据量：
 $352 \times 240 \times 3 = 253 \text{ KB}$
 每秒数据量 (比特率)： $253 \times 30 = 7.59 \text{ MB/S}$
 一片CD-ROM可存帧数： $650 / 0.253 = 2.569 \text{ K帧/片}$ ，一片CD-ROM节目时间：
 $(650 / 7.59) / 60 = 1.42 \text{ 分/片}$

数字媒体实验室

6

举例三：

<http://media.hust.edu.cn>

- 一个陆地卫星（LandSat-3）的例子（其水平、垂直分辨率分别为2340和3240，四波段、采样精度7位），它的一幅图像的数据量为
 $2340 \times 3240 \times 7 \times 4 = 212 \text{ Mbit}$ ，按每天30幅计，每天数据量为 $212 \times 30 = 6.36 \text{ Gbit}$ ，每年的数据量高达2300Gbit。

数字媒体实验室

7

数据压缩的可能性

<http://media.hust.edu.cn>

- 人们研究发现，多媒体数据中存在着大量的冗余
- 通过去除那些冗余数据可以使原始数据极大地减少，从而解决多媒体数据量巨大的问题
- 多媒体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
- 因此，进行图像压缩研究的起点是研究多媒体数据的冗余性

数字媒体实验室

8

(1) 空间冗余

<http://media.hust.edu.cn>

- 这是静态图像存在的最主要的一种数据冗余，一幅图像记录了画面上可见景物的颜色
- 同一景物表面上各采样点的颜色之间往往存在着空间连贯性
- 可以通过改变物体表面颜色的像素存储方式来利用空间连贯性，达到减少数据量的目的

数字媒体实验室

9



(2) 时间冗余

<http://media.hust.edu.cn>

- 这是序列图像（电视图像、运动图像）表示中经常包含的冗余
- 序列图像一般为位于一时间轴区间内的一组连续画面，其中的相邻帧往往包含相同的背景和移动物体
- 只不过移动物体所在的中间位置略有不同，所以后一帧的数据与前一帧的数据有许多共同的地方

数字媒体实验室

11



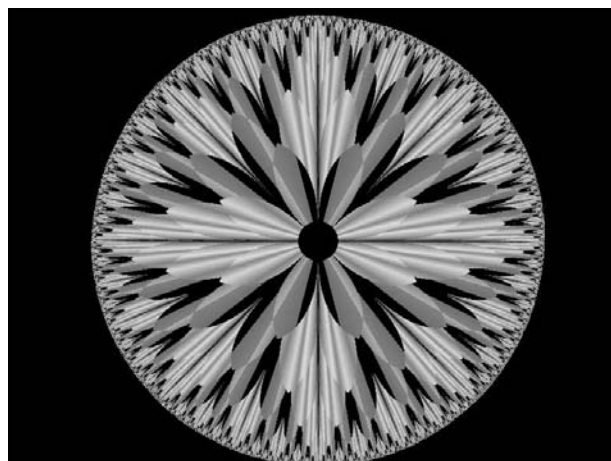
(3) 结构冗余

<http://media.hust.edu.cn>

- 有些图像的纹理区，图像的像素值存在着明显的分布模式
- 例如，方格状的地板图案等，我们称此为结构冗余
- 已知分布模式，可以通过某一过程来生成图像

数字媒体实验室

13



(4) 知识冗余

<http://media.hust.edu.cn>

- 有些图像的理解与某些基础知识有相当大的相关性
- 例如：人脸的图像有固定的结构，比方说嘴的上方有鼻子，鼻子的上方有眼睛，鼻子位于脸的中线上等等
- 这类规律性的结构可由先验知识和背景知识得到，称此类冗余为知识冗余

数字媒体实验室

15

- 根据已有的知识，对某些图像中所包含的物体，可以构造其基本模型，并创建对应各种特征的图像库，进而图像的存储只需要保存一些特征参数，从而可以大大减少数据量。
- 知识冗余是模型编码主要利用的特性

(5) 视觉冗余

<http://media.hust.edu.cn>

- 视觉系统对图像的亮度和色彩度的敏感性相差很大
- 随着亮度的增加，视觉系统对量化误差的敏感度降低
- 人眼的视觉系统把图像的边缘和非边缘区域分开来处理
- 人类的视觉系统总是把视网膜上的图像分解成若干个空间有向的频率通道后再进一步处理

数字媒体实验室

17

(6) 图像区域的相同性冗余

<http://media.hust.edu.cn>

- 在图像中的两个或多个区域所对应的所有像素值相同或相近而产生的数据重复性存储
- 记录了一个区域中各像素的颜色值，则与其相同或相近的其他区域就不再需记录其中各像素的值
- 向量量化(vector quantization)方法就是针对这种冗余性的图像压缩编码方法

数字媒体实验室

18

随着对人类视觉、听觉系统和图像模型的进一步研究，人们可能会发现更多的冗余性，使视频数据压缩编码的可能性越来越大，从而推动视频压缩技术的进一步发展。

§ 3.2 多媒体数据压缩编码的分类

<http://media.hust.edu.cn>

- 多媒体数据压缩方法根据不同的依据可产生不同的分类
- 第一种，根据质量有无损失可分为：
 - 损失编码
 - 无损失编码

数字媒体实验室

20

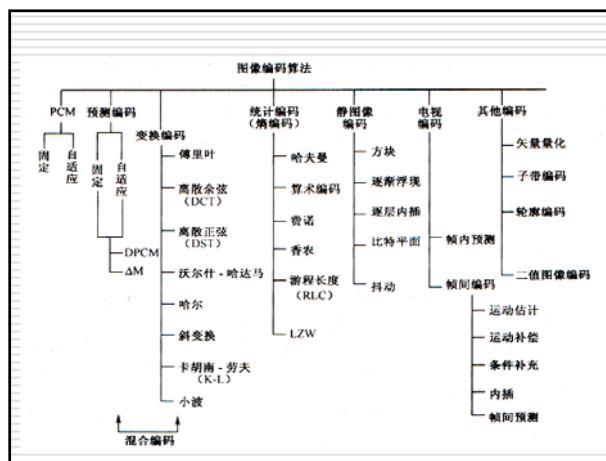
- 第二种，按照其作用域在空间域或频率域上分为：

- 空间方法
- 变换方法
- 混合方法

- 第三种，根据是否自适应分为：

- 自适应性编码
- 非自适应性编码

- 一般来说一个编码方法都有其相应的自适应方法



§ 3.3 PCM与预测编码

<http://media.hust.edu.cn>

- 脉冲编码调制(PCM)
- 增量调制(DM)
- 自适应增量调制 (ADM)
- 自适应脉冲编码 (APCM)
- 差分脉冲编码调制 (DPCM)
- 子带编码 (SBC)

数字媒体实验室

23

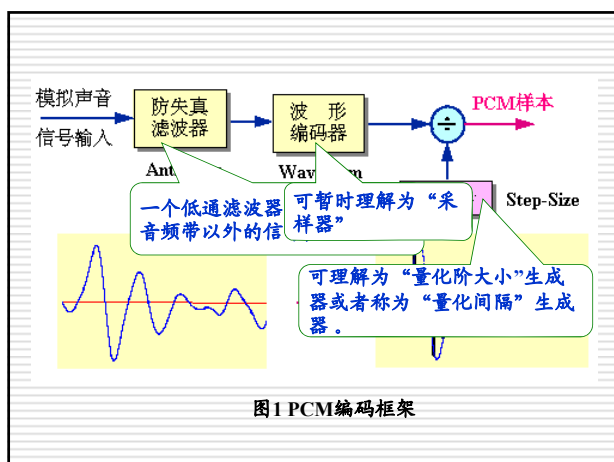
§ 3.3.1 脉冲编码调制

<http://media.hust.edu.cn>

- 脉冲编码调制(pulse code modulation, PCM)是概念上最简单、理论上最完善的编码系统
- PCM是最早研制成功、使用最为广泛的编码系统，但也是数据量最大的编码系统

数字媒体实验室

24



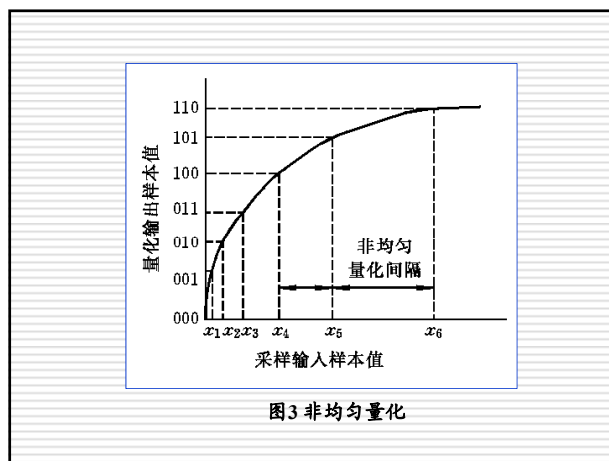
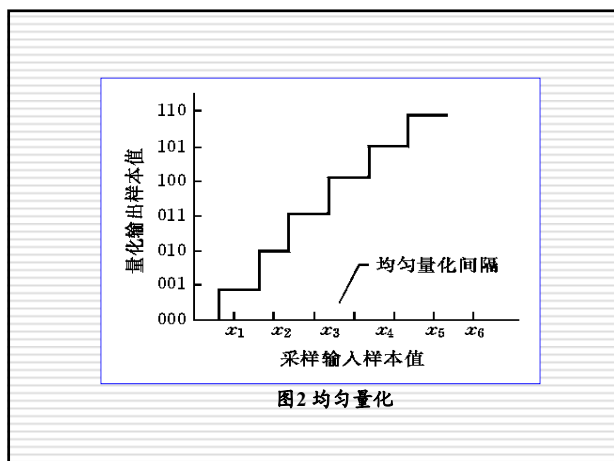
均匀量化与非均匀量化

<http://media.hust.edu.cn>

- 如果采用相等的量化间隔对采样得到的信号作量化，那么这种量化称为**均匀量化**，也称为线性量化
- **非均匀量化**也称为非线性量化，对输入信号进行量化时，大的输入信号采用大的量化间隔，小的输入信号采用小的量化间隔

数字媒体实验室

26



§ 5.3.2 增量调制(DM)

<http://media.hust.edu.cn>

- 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
- PCM是对每个采样信号的整个幅度进行量化编码，因此它具有对任意波形进行编码的能力

数字媒体实验室

29

- 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“非负”，则用“1”表示；相反则用“0”表示，或者相反
- 由于DM编码只须用1位对语音信号进行编码，所以DM编码系统又称为“1位系统”

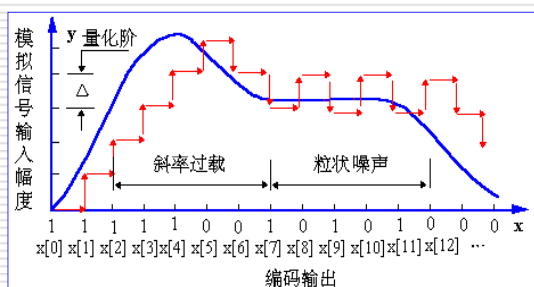


图4 DM波形编码示意图

DM的缺点:

<http://media.hust.edu.cn>

- “斜率过载” (slope overload)
- “粒状噪声” (granular noise)

数字媒体实验室

32

§ 3.3.3 自适应增量调制 (ADM)

<http://media.hust.edu.cn>

- 为了使增量调制器的量化阶 Δ 能自适应,也就是根据输入信号斜率的变化自动调整量化阶 Δ 的大小,以使斜率过载和粒状噪声都减到最小,许多研究人员研究了各种各样的方法
- 算法依据: 检测到斜率过载时开始增大量化阶 Δ ,而在输入信号的斜率减小时降低量化阶 Δ

数字媒体实验室

33

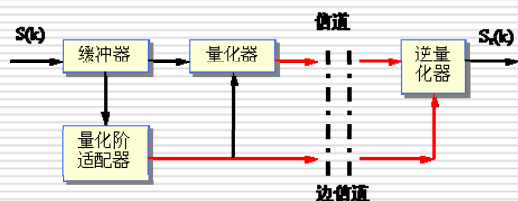
§ 3.3.4 自适应脉冲编码调制 (APCM)

<http://media.hust.edu.cn>

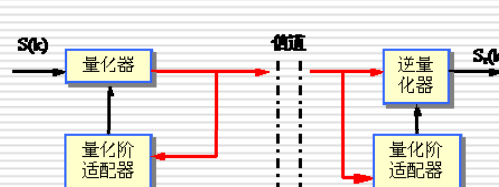
- 自适应脉冲编码调制(Adaptive Pulse Code Modulation, APCM)是根据输入信号幅度大小来改变量化阶大小的一种波形编码技术
- 这种自适应可以是瞬时自适应,即量化阶的大小每隔几个样本就改变,也可以在较长时间周期里发生变化

数字媒体实验室

34



(a) 前向自适应



(b) 后向自适应

§ 3.3.5 差分脉冲编码调制 (DPCM)

<http://media.hust.edu.cn>

- 差分脉冲编码调制DPCM (Differential Pulse Code Modulation) 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
- 基本思想: 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

数字媒体实验室

37

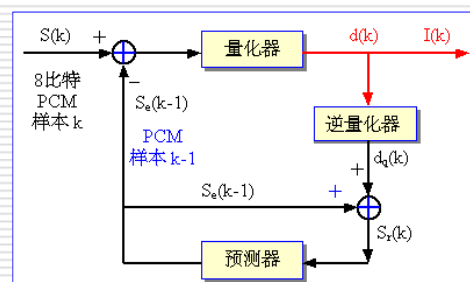


图5 DPCM方块图

§ 3.3.6 自适应差分脉冲编码调制 (ADPCM)

<http://media.hust.edu.cn>

- 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
- 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

数字媒体实验室

39

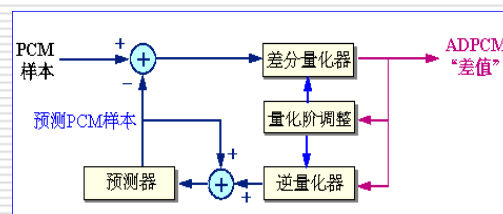


图6 ADPCM方块图

§ 3.3.7 子带编码 (SBC)

<http://media.hust.edu.cn>

- 使用一组带通滤波器(band-pass filter, BPF)把输入信号的频带分成若干个连续的频段, 每个频段称为子带
- 对每个子带中的信号采用单独的编码方案去编码。在信道上传送时, 将每个子带的代码复合起来
- 在接收端译码时, 将每个子带的代码单独译码, 然后把它们组合起来, 还原成原来的信号

数字媒体实验室

41

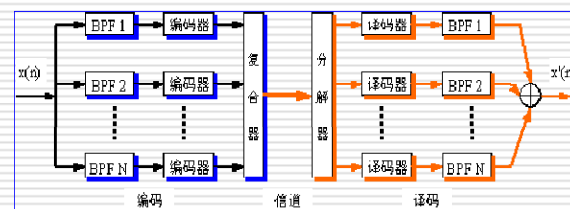


图7 子带编码方块图

子带编码的好处:

<http://media.hust.edu.cn>

- 对每个子带信号分别进行自适应控制, 量化阶 (Quantization Step) 的大小可以按照每个子带的能量电平加以调节
- 具有较高能量电平的子带用大的量化阶去量化, 以减少总的量化噪声
- 可根据每个子带信号在感觉上的重要性, 对每个子带分配不同的位数, 用来表示每个样本值



数字媒体实验室

43

§ 3.4 无损压缩编码

<http://media.hust.edu.cn>

- 香农-范诺编码与霍夫曼编码 (Huffman)
- 算术编码
- 行程长度编码 (RLE)
- 词典编码



数字媒体实验室

44

§ 3.4.1 香农-范诺编码与霍夫曼编码

<http://media.hust.edu.cn>

- 统计编码的理论基础是**信息论**
- 根据信息论的原理, 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**
- 如果要求在编码过程中不丢失信息量, 即要求保存信息熵, 这种信息保持编码又叫做**熵保持编码**, 或者叫**熵编码**



数字媒体实验室

45

熵 (Entropy) 的概念

<http://media.hust.edu.cn>

- **熵**是信息量的度量方法, 它表示某一事件包含的消息越多, 事件发生的可能性就越小, 数学上就是概率越小 (如爆炸性新闻)
- 某个事件的信息量用 $I_i = -\log_2 p_i$ 表示, 其中 p_i 为第 i 个事件的概率

$$0 < p_i \leq 1$$



数字媒体实验室

46

信源S的熵的定义

<http://media.hust.edu.cn>

- 按照香农编码的最小长度, 信源S的熵定义为:

$$H(s) = \eta = \sum p_i \log_2 (1/p_i) = \sum p_i I_i$$

p_i 是符号 s_i 在 S 中出现的概率。

表示包含在 s_i 中的信息量, 也就是编码所需要的位数。



数字媒体实验室

47

[例1]

<http://media.hust.edu.cn>

- 一幅用256级灰度表示的图像, 如果每一个象素点灰度的概率均为 $p_i = 1/256$, 编码每一个象素点就需要8位。



数字媒体实验室

48

[例2]
<http://media.hust.edu.cn>

- 有一幅40个象素组成的灰度图像，灰度共有5级，分别用符号A、B、C、D和E表示
- 40个象素中出现灰度A的象素数有15个，出现灰度B的象素数有7个，出现灰度C的象素数有7个等等，如表1所示
- 如果用3位表示5个等级的灰度值，也就是每个象素用3位表示，编码这幅图像总共需要120位

数字媒体实验室

49

表1 符号在图像中出现的次数

符号	出现的次数
A	15
B	7
C	7
D	6
E	5

按照香农理论，这幅图像的熵为：

$$H(S) = (15/40) \times \log_2(40/15) + (7/40) \times \log_2(40/7) + \dots + (5/40) \times \log_2(40/5) = 2.196$$

这就是说每个符号用2.196位表示，40个象素需用87.84位。

香农-范诺编码
<http://media.hust.edu.cn>

- 最早阐述和实现这种编码的是Shannon（1948年）和Fano（1949年），因此被称为**香农-范诺（Shannon-Fano）算法**
- 这种方法采用从上到下的方法进行编码
- 首先按照符号出现的频度或概率排序，例A、B、C、D和E
- 然后使用递归方法分成两个部分，每一部分具有近似相同的次数

数字媒体实验室

52

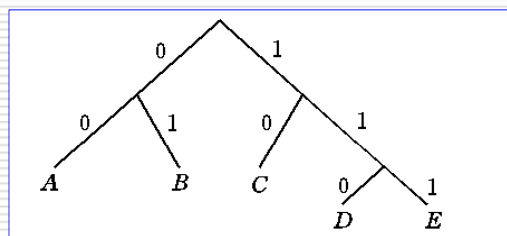


图8 香农-范诺算法编码举例

表2 香农-范诺算法举例表

符号	出现的次数(p_i)	$\log_2(1/p_i)$	分配的代码	需要的位数
A	15 (0.375)	1.4150	00	30
B	7 (0.175)	2.5145	01	14
C	7 (0.175)	2.5145	10	14
D	6 (0.150)	2.7369	110	18
E	5 (0.125)	3.0000	111	15

- 按照这种方法进行编码得到的总位数为91，压缩比约为1.3 : 1

霍夫曼编码

<http://media.hust.edu.cn>

- 霍夫曼编码方法于1952年问世
- 广泛应用与各种数据压缩技术中，且仍不失为熵编码中的最佳编码方法
- 霍夫曼编码的理论依据是**变字长编码理论**
- 编码器的编码输出码字是字长不等的码字
- 按编码输入信息符号出现的统计概率，给输出码字分配以不同的字长

数字媒体实验室

55

- 对于编码输入中，出现**大**概率的信息符号，赋以**短**字长的输出码字
- 对于编码输入中出现**小**概率的信息符号，赋以**长**字长的输出码字
- 可以证明，按照概率出现大小的顺序，对输出码字分配不同码字长度的变字长编码方法，其输出码字的平均码长最短，与信息熵值接近，编码方法最佳

霍夫曼编码的具体步骤归纳如下：

<http://media.hust.edu.cn>

- (1) 概率统计（如对一幅图像或m幅同种类型图像作灰度信号统计）得到n个不同概率的信息符号
- (2) 将n个信源信息符号的n个概率按概率大小排序
- (3) 将n个概率中最后两个小概率相加，这时概率个数减为n-1个

数字媒体实验室

57

(4) 将n-1个概率按大小重新排序

(5) 重复(3)，将新排序后的最后两个小概率再相加，相加和与其余概率再排序

(6) 如此反复重复n-2次，得到只剩两个概率序列

(7) 以二进制码元(0, 1)赋值，构成霍夫曼哈夫曼码字

编码结束

[例3]

输入信息符号	输入概率	第一步	第二步	第三步	第四步	第五步
x_1	0.35	0.35	0.35	0.35	0.40	0.60
x_2	0.20	0.20	0.20	0.25	0.35	0.40
x_3	0.15	0.15	0.20	0.20	0.25	
x_4	0.10	0.10	0.15	0.20		
x_5	0.10	0.10	0.10			
x_6	0.06	0.10				
x_7	0.04					

图9 霍夫曼编码步骤

码长	输入信息符号	霍夫曼码	第一步	第二步	第三步	第四步	第五步
2	x_1		0.35 ←	0.35 ←	0.35 ←	0.35 ←	0.40 ←
2	x_2		0.20 ←	0.20 ←	0.20 ←	0.25 ←	0.35 ←
3	x_3		0.15 ←	0.15 ←	0.20 ←	0.20 ←	0.25 ←
3	x_4		0.10 ←	0.10 ←	0.15 ←	0.20 ←	
3	x_5		0.10 ←	0.10 ←	0.10 ←		
4	x_6		0.06 ←	0.10 ←			
4	x_7		0.04 ←				

图10 霍夫曼码字构成

采用霍夫曼编码时有两个问题值得注意：

<http://media.hust.edu.cn>

- 霍夫曼码没有错误保护功能 在译码时，如果码串中没有错误，那么就能一个接一个地正确译出代码
- 霍夫曼码是可变长度码 很难随意查找或调用压缩文件中间的内容，然后再译码，这就需要在存储代码之前加以考虑

课堂作业

<http://media.hust.edu.cn>

信息符号及其概率如下：

a	a1	a2	a3	a4	a5
p(a)	0.5	0.25	0.125	0.0625	0.0625

求其Huffman编码、信息熵和平均码长。

§ 3.4.2 算术编码

<http://media.hust.edu.cn>

- 算术编码在图像数据压缩标准(如JPEG)中扮演了重要的角色
- 在算术编码中，信号用0到1之间的实数进行编码，算术编码用到两个基本的参数：符号的概率和它的编码间隔

算术编码

<http://media.hust.edu.cn>

- 信源符号的概率决定压缩编码的效率，也决定编码过程中信源符号的间隔，而这些间隔包含在0到1之间
- 编码过程中的间隔决定了符号压缩后的输出
- 算术编码器的编码过程可用下面的例子加以解释

[例4]

<http://media.hust.edu.cn>

- 假设信源符号为{00, 01, 10, 11}，这些符号的概率分别为{0.1, 0.4, 0.2, 0.3}，根据这些概率可把间隔[0, 1)分成4个子间隔：[0, 0.1), [0.1, 0.5), [0.5, 0.7), [0.7, 1)。

表3 信源符号、概率和初始编码间隔

符号	概率	初始编码间隔
00	0.1	[0, 0.1)
01	0.4	[0.1, 0.5)
10	0.2	[0.5, 0.7)
11	0.3	[0.7, 1)

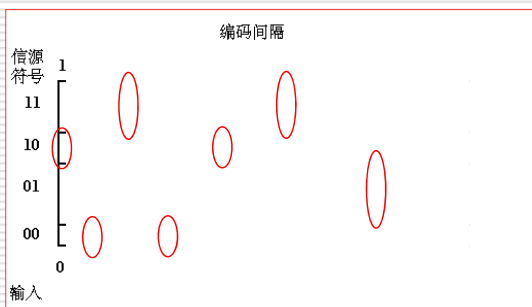


图11 算术编码过程举例

表4 编码过程

步骤	输入符号	编码间隔	编码判决
1	10	[0.5, 0.7)	符号的间隔范围[0.5, 0.7)
2	00	[0.5, 0.52)	[0.5, 0.7)间隔的第一个1/10
3	11	[0.514, 0.52)	[0.5, 0.52)间隔的最后一个1/10
4	00	[0.514, 0.5146)	[0.514, 0.52)间隔的第一个1/10
5	10	[0.5143, 0.51442)	[0.514, 0.5146)间隔的第五个1/10开始, 二个1/10
6	11	[0.514384, 0.51442)	[0.5143, 0.51442)间隔的最后3个1/10
7	01	[0.5143836, 0.514402)	[0.514384, 0.51442)间隔的4个1/10, 从第1个1/10开始
8	从[0.5143876, 0.514402]中选择一个数作为输出: 0.5143876		


表5 译码过程

步骤	间隔	译码符号	译码判决
1	[0.5, 0.7)	10	0.51439在间隔 [0.5, 0.7)
2	[0.5, 0.52)	00	0.51439在间隔 [0.5, 0.7)的第1个1/10
3	[0.514, 0.52)	11	0.51439在间隔 [0.5, 0.52)的第7个1/10
4	[0.514, 0.5146)	00	0.51439在间隔 [0.514, 0.52)的第1个1/10
5	[0.5143, 0.51442)	10	0.51439在间隔 [0.514, 0.5146)的第5个1/10
6	[0.514384, 0.51442)	11	0.51439在间隔 [0.5143, 0.51442)的第7个1/10
7	[0.51439, 0.5143948)	01	0.51439在间隔 [0.51439, 0.5143948)的第1个1/10
7	译码的消息: 10 00 11 00 10 11 01		

算术编码需要注意的几个问题:

<http://media.hust.edu.cn>

- 由于实际的计算机的精度不可能无限长, 运算中出现溢出是一个明显的问题, 但多数机器都有16位、32位或者64位的精度, 因此这个问题可使用比例缩放方法解决
- 算术编码器对整个消息只产生一个码字, 这个码字是在间隔 $[0, 1)$ 中的一个实数, 因此译码器在接收到表示这个实数的所有位之前不能进行译码;
- 算术编码也是一种对错误很敏感的编码方法, 如果有一位发生错误就会导致整个消息译错

 数字媒体实验室

70

课堂作业

<http://media.hust.edu.cn>

假设信源符号为{00, 01, 10, 11}, 这些符号的概率分别为{0.2, 0.1, 0.3, 0.4}, 求输入信号{10 11 00 01 10 01 00}的算术编码, 请写出详细的解题步骤。

数字媒体实验室

71

§ 3.4.3 行程编码 (RLE-Run Length Encoding)

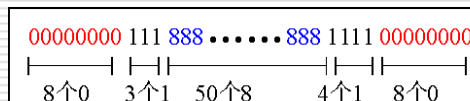



图12 RLE编码的概念

 数字媒体实验室

72

行程编码 (RLE-Run Length Encoding)

- 这种压缩编码称为**行程编码**，具有相同颜色并且是连续的像素数目称为**行程长度**
- RLE所能获得的压缩比有多大，这主要是取决于图像本身的特点
- 如果图像中具有相同颜色的图像块越大，图像块数目越少，获得的压缩比就越高
- 反之，压缩比就越小

数字媒体实验室

73

§ 3.4.4 词典编码 (Dictionary Encoding)

<http://media.hust.edu.cn>

- 词典编码 (Dictionary Encoding) 的根据是数据本身包含有重复代码这个特性
- 例如，文本文件和光栅图像就具有这种特性
- 词典编码法的种类很多，归纳起来大致有两类

数字媒体实验室

74

第一类词典法(基本思想)

<http://media.hust.edu.cn>

- 企图查找正在压缩的字符序列是否在以前输入的数据中出现过
- 用已经出现过的字符串替代重复的部分，它的输出仅仅是指向早期出现过的字符串的“指针”

数字媒体实验室

75

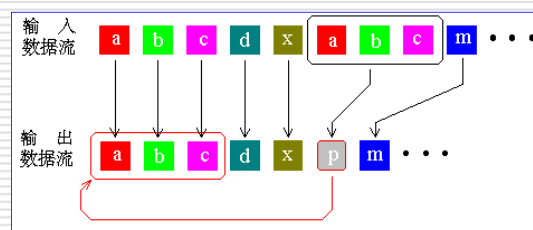


图13 第一类词典法编码的概念

第二类词典法(基本思想)

<http://media.hust.edu.cn>

- 企图从输入的数据中创建一个“短语词典 (dictionary of the phrases)”，这种短语不一定具有具体含义，可以是任意字符的组合
- 编码数据过程中当遇到已经在词典中出现的“短语”时，编码器就输出这个词典中的短语的“索引号”，而不是短语本身

数字媒体实验室

77

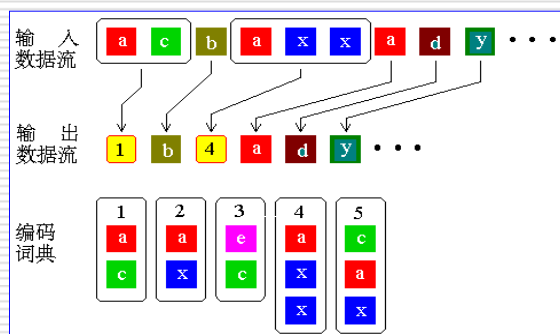


图14 第二类词典法编码的概念

§ 3.5 变换编码

<http://media.hust.edu.cn>

- 变换编码的基本原理
- 离散余弦变换—DCT变换
- 小波变换—WT变换

数字媒体实验室

79

§ 3.5.1 变换编码的基本原理

<http://media.hust.edu.cn>

- 变换编码不是直接对空域图像信号编码，而是首先将空域图像信号映射变换到另一个正交矢量空间（变换域或频域）
- 产生一批变换系数，然后对这些变换系数进行编码处理
- 图15示出一个变换编解码进行过程的示意图

数字媒体实验室

80

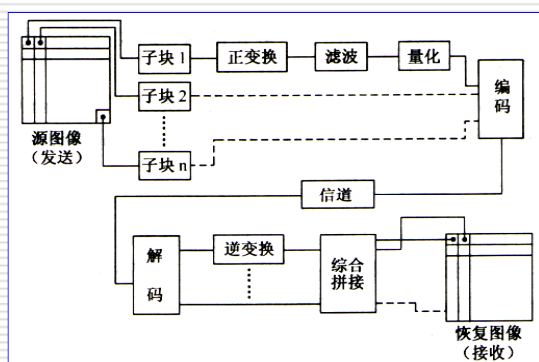


图15 变换编、解码过程示意图

变换编码的基本原理（续）

<http://media.hust.edu.cn>

- 在发送端将原始图像分割成1到n个子图像块，每个子图像块送入正交变换器作正交变换
- 变换器输出变换系数经滤波、量化、编码后送信道传输到达接收端
- 接收端作解码、逆变换、综合拼接，恢复出空域图像

数字媒体实验室

82

提出问题：

<http://media.hust.edu.cn>

- 数字图像信号经过正交变换为什么能够压缩数据量呢？

数字媒体实验室

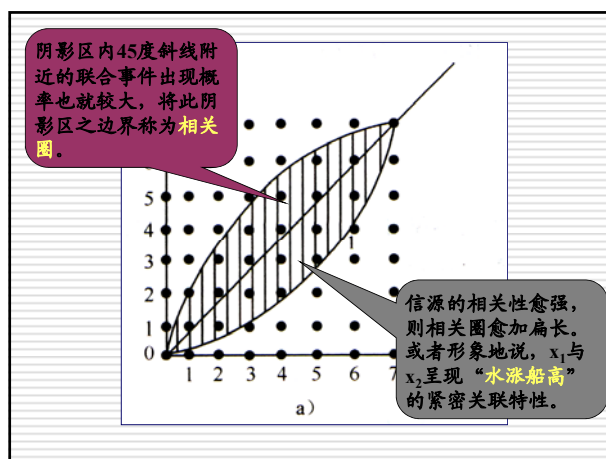
83

先让我们看一个最简单的时域三角函数的例子

$$y(t) = A \sin 2\pi f t,$$
 当 f 从 $-\infty$ 到 $+\infty$ 改变时， $y(t)$ 是一个正弦波。

- 假如将其变换到频域表示，只需幅值 A 和频率 f 两个参数就足够了，可见 $y(t)$ 在时域描述，数据之间的相关性大，数据冗余度大
- 而转换到频域描述，数据相关性大大减少，数据冗余量减少，参数独立，数据量减少

- 现再举一个例子，设有两个相邻的数据样本 x_1 与 x_2 ，每个样本采用3位编码，因此各有 $2^3=8$ 个幅度等级
- 而两个样本的联合事件，共有 $8 \times 8=64$ 种可能性，可用二维平面坐标表示
- 其中 x_1 轴与 x_2 轴分别表示相邻两样本可能的幅度等级



- 为了要对圈内各点的位置进行编码，就要对两个差不多大的坐标值分别进行编码
- 当相关性愈弱时，此相关圈就显方圆形状，说明 x_1 处于某一幅度等级时， x_2 可能出现在不相同的任意幅度等级上

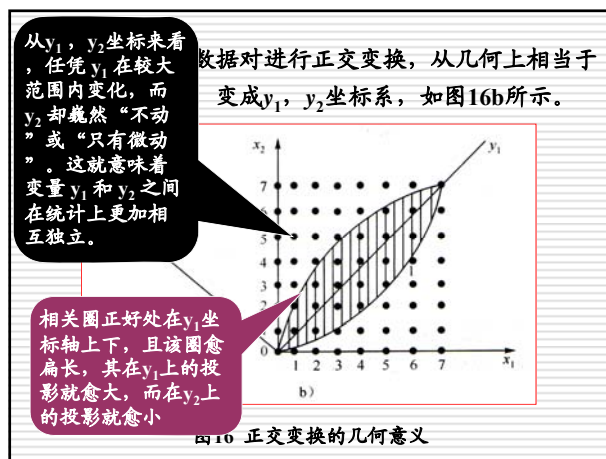


图16 正交变换的几何意义

变换编码的基本原理

<http://media.hust.edu.cn>

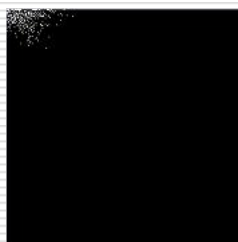
- 利用视频图像块像素值之间的相关性，把视频图像变换到一组新的基上
- 在变换域里视频图像要比空间域里简单
- 视频图像的**相关性明显下降**，信号的能量主要集中在**少数几个低频变换系数上**
- 采用量化和熵编码可有效地压缩其数据

变换编码的基本原理

- 一帧图像内容以不同的亮度和色度像素分布体现出来
- 这些像素的分布依图像内容而变，毫无规律可言
- 但是通过变换（如DCT变换），像素分布就有了规律
- 代表低频成份的量分布于左上角，而越高频率成份越向右下角分布
- 根据人眼视觉特性，去掉一些不影响图像基本内容的细节（高频分量），达到压缩码率的目的



原始图像



频域图像

91

变换编码的历史及应用

<http://media.hust.edu.cn>

- 变换编码技术已有近30年的历史，技术上比较成熟，理论也比较完备
- 广泛应用于各种视频图像数据压缩
 - 单色图像
 - 彩色图像
 - 静止图像
 - 运动图像



数字媒体实验室

92

变换编码的种类

<http://media.hust.edu.cn>

- 正交变换的种类很多
 - 傅立叶(Fourier)变换
 - 沃尔什(Walsh)变换
 - 哈尔(Haar)变换
 - 斜(slant)变换
 - 余弦变换
 - 正弦变换
 - K-L(Karhunen-Loeve)变换
 -



数字媒体实验室

93

§ 3.5.2 离散余弦变换-DCT变换

<http://media.hust.edu.cn>

- 余弦变换是傅里叶变换的一种特殊情况
- 在傅里叶级数展开式中，如果被展开的函数是实偶函数，那么，其傅里叶级数中只包含余弦项
- 再将其离散化便可以导出余弦变换，称之为离散余弦变换DCT (Discrete Cosine Transform)



数字媒体实验室

94

DCT变换的特点

<http://media.hust.edu.cn>

- 离散余弦变换在数字图像数据压缩编码技术中与最佳变换K-L变换媲美
- DCT与K-L变换压缩性能和误差很接近
- DCT计算复杂度适中，具有可分离特性，速度快
- 具有较强的抗干扰能力
- 在图像数据压缩中广泛应用
 - JPEG、MPEG、H.261



数字媒体实验室

95

(1) 二维DCT公式

<http://media.hust.edu.cn>

对每个单独的彩色图像分量，把整个分量图像空间分成一个 8×8 的图像块，并作为二维离散余弦变换DCT的输入，通过DCT变换，把能量集中在少数几个系数上。

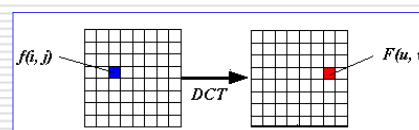


图17 离散余弦变换



数字媒体实验室

96

变换的公式如下:

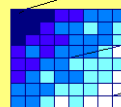
正变换 (FDCT)

$$F(u, v) = \frac{1}{4} C(u) C(v) \left[\sum_{i=0}^7 \sum_{j=0}^7 f(i, j) \cos \frac{(2i+1)u\pi}{16} \cos \frac{(2j+1)v\pi}{16} \right] \quad (3-1)$$

逆变换 (IDCT)

$$f(i, j) = \frac{1}{4} C(u) C(v) \left[\sum_{u=0}^7 \sum_{v=0}^7 F(u, v) \cos \frac{(2i+1)u\pi}{16} \cos \frac{(2j+1)v\pi}{16} \right] \quad (3-2)$$

DC 分量 $F(0,0)$ 图象平均灰度值, AC 分量 63 个系数
高频部分, 图象细节



其中,

$$\begin{cases} C(u), C(v) = 1/\sqrt{2}, & \text{当 } u, v = 0; \\ C(u), C(v) = 1, & \text{其它} \end{cases}$$

$f(i, j)$ 经过变换后, $F(0, 0)$ 是直流系数, 其它为交流系数。

(2) 在计算二维的DCT变换时, 可使用下面的计算公式把两维的DCT变换变成一维的DCT变换, 如图18所示。

$$F(u, v) = \frac{1}{2} C(u) \left[\sum_{i=0}^7 G(i, v) \cos \frac{(2i+1)u\pi}{16} \right] \quad (3-3)$$

$$G(i, v) = \frac{1}{2} C(v) \left[\sum_{j=0}^7 f(i, j) \cos \frac{(2j+1)v\pi}{16} \right] \quad (3-4)$$

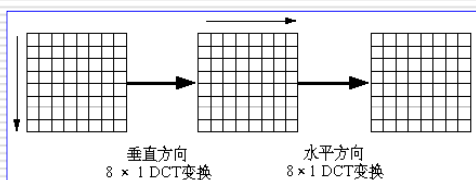


图18 二维DCT变换方法

§ 3.5.3 离散小波变换-WT变换

<http://media.hust.edu.cn>

- 小波 (Wavelet): “小波”就是小的波形
- 所谓“小”是指它具有衰减性, 是有限宽度的波, 主要是相对于无限宽度的正、余弦波而言
- 小波变换 (Wavelet Transform): 以有限宽度的波为基函数的变换
- 小波变换是时间 (空间) 频率的局部化分析, 有“数学显微镜”的美誉

小波变换的特点

<http://media.hust.edu.cn>

- 与Fourier变换相比,小波变换是空间(时间)和频率的局部变换
- 能有效地从信号中提取信息
- 通过伸缩和平移等运算功能可对函数或信号进行多尺度的细化分析,解决了Fourier变换不能解决的许多困难问题

数字媒体实验室

103

小波变换的应用

<http://media.hust.edu.cn>

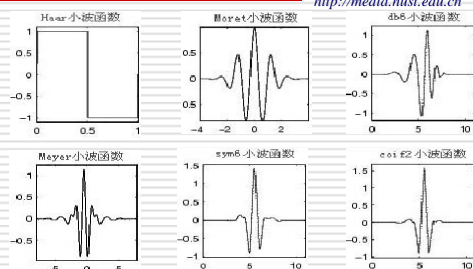
- 信号和信息处理专家认为,小波分析是时间-尺度分析和多分辨率分析的一种新技术
- 它在信号分析、语音合成、图像识别、计算机视觉、数据压缩、地震勘探、大气与海洋波分析等方面的研究都取得了有科学意义和应用价值的成果

数字媒体实验室

104

小波变换的种类

<http://media.hust.edu.cn>



数字媒体实验室

105

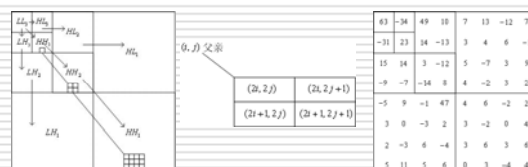
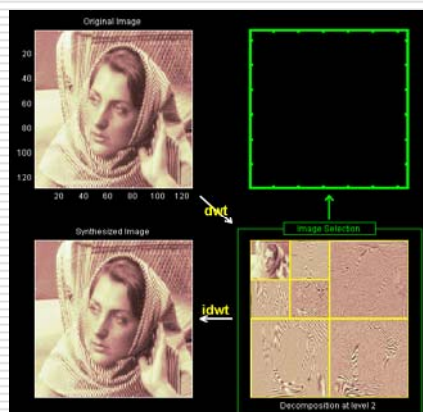
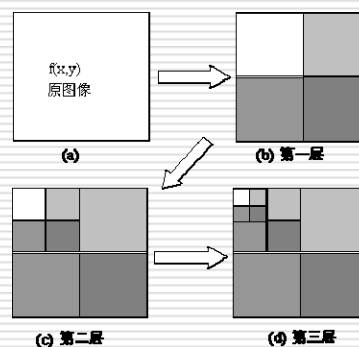


表5 哈尔小波变换过程

分辨率	平均值	细节系数
4	[9 7 3 5]	
2	[8 4]	[1 -1]
1	[6]	[2]