

王晗

wanghan@bjfu.edu.cn

专题提纲

- →数据压缩基础
- ■图像数据压缩编码
- →动态图像压缩技术和标准
- ►H.26x标准

语音数据:

数字激光唱盘 (CD-DA) 的标准采样频率为44.1 kHz, 量化位数为16位, 立体声。

一分钟 CD-DA 音乐所需的存储量为

 $44.1 \text{ K} \times 16 \times 2 \times 60 \div 8 = 10584 \text{ KB}$

数字电视图象:

标准的VCD图象视频图象(352x240) NTSC制。

每秒30帧, 其数据量为:

 $352x240x3x30Byte = 7.603MB_{\circ}$

8GU盘可存8000/7.603/60=17分钟视频。

- ●由此看出,庞大的数字化信息对存储器的存储容量、通信传输以及计算机处理速度都造成了极大的压力,也因此给我们带来的一个难题,同时给了我们一个机会:如何用软件的手段来解决硬件上的物理极限。
- ▶人们研究发现数字媒体数据中存在大量的冗余信息,因此数据压缩技术就是研究如何利用数据的冗余性来减少数字媒体数据量的方法。

数据冗余

你的妻子, Helen, 将于明天晚上6点零5分在上海的虹桥机场接你。

你的妻子将于明天晚上6点零5分在虹桥机场接你。

Helen将于明晚6点在虹桥接你。

结论:只要接收端不会产生误解,就可以减少承载信息的数据量。

媒 余

指视频或图像中所包含的某些信息与人们的一些先验知识有关。 例如在头肩图像中,头、眼、鼻和最的相对位置等信息就是人类的共性知识。

在视频或图像的纹理区,像素的亮度、色度等信息存在着明显的分布式模式,如果知道了分布模式,就可以通过某种算法来生成图像,即存在结构冗余

研究发现人眼的视觉特性是非均 匀和非线性的。可以对人眼不敏 感的信息少编码甚至编码以压缩 数据量。 空间 冗余 同一帧图像中相邻的像素具有 很强的相关性。

时间 □ 图像序列中相连帧的对应像 □ 素具有很强的相关性

信息熵 冗余

知识

冗余

结构

冗余

视觉

冗余

信息熵是指一组数据所携带的 信息量。信息熵冗余又称为编 码冗余,是指信息熵少于数据 编码的码元长度而形成的冗余。

- ▶ 从数字信号的统计特征方面:空间冗余、时间冗余和信息熵冗余
- ▶ 从更高层次的人类认知层面:视觉冗余、知识冗余和结构冗余
- ▶ 视频能够压缩的根本原因在于视频数据具有较高的冗余度。
- 压缩就是指冗余的消除,主要基于两种技术:统计学和心理视觉。消除统计冗余的基本依据是视频数字化过程在时间和空间上采用了规则的采样过程。视频画面数字化为规则的像素阵列,其密集程度适于表征每点最高的空间频率,而绝大多数画面帧包含非常少甚至不含这种最高频率的细节。



》总结:

〉空间冗余:

数字媒体(如图像、视频、音频)中同类像素重复性较多,则可以用较少的编码信息来表示原数据,它以变换编码、量化和熵编码等技术为基础。

> 时间冗余:

视频相邻帧 (声音) 之间的相似性, 它利用预测和运动补偿对视频图像进行压缩。

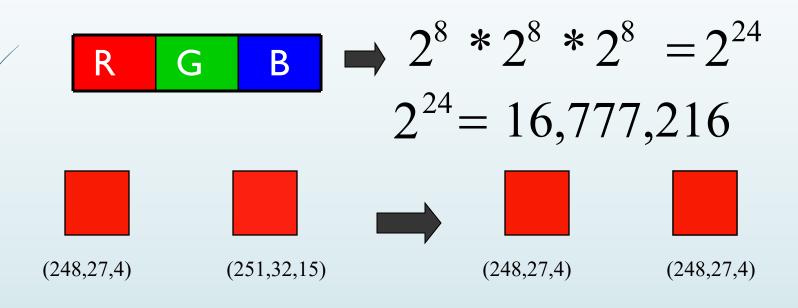
> 结构冗余:

图像的部分区域存在着非常强的纹理结构,或是图像的各个部分之间存在有某种关系,例如自相似性等。

〉 视觉冗余:

人的视觉对某些信号(如颜色)具有不那么敏感的生理特性,例如视觉惰性(对亮度和色度, 蓝色和红绿色, 25帧视频采样等)的遮蔽效应。

■图象的视觉冗余 (彩色)



——数据压缩

设: n1和n2是在两个表达相同信息的数据集中所携带的信息量。

■压缩率 (压缩比):

$$C_R = n_1 / n_2$$

其中,n1是压缩前的数据量,n2是压缩后的数据量。

■相对数据冗余:

$$R_D = (1 - 1/C_R) \times 100\%$$

例: $C_R = 20$; $R_D = 19/20 = 95\%$

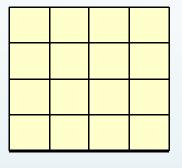
压缩编码的分类

压缩编码的分类方法有很多种,从信息论角度可以分为无损压缩和有损压缩两大类。

- 无损压缩是指数据在压缩或解压缩过程中不会改变或损失原有的信息,解压缩产生的数据是原始对象的完整复制,没有失真。
- ▶ 有损压缩是利用了人类视觉对图像中的某些 频率成分不敏感的特性,允许压缩过程中损 失一定的信息。

压缩编码的分类

——无损压缩





RGB	RGB	RGB	RGB
RGB	RGB	RGB	RGB
RGB	RGB	RGB	RGB
RGB	RGB	RGB	RGB



16 RGB

从原来的16*3*8=284bits压缩为: (1+3)*8=32bits

压缩编码的分类

——有损压缩

36	35	34	34	34
34	34	32	34	34
33	37	30	34	34
34	34	34	34	34
34	35	34	34	31

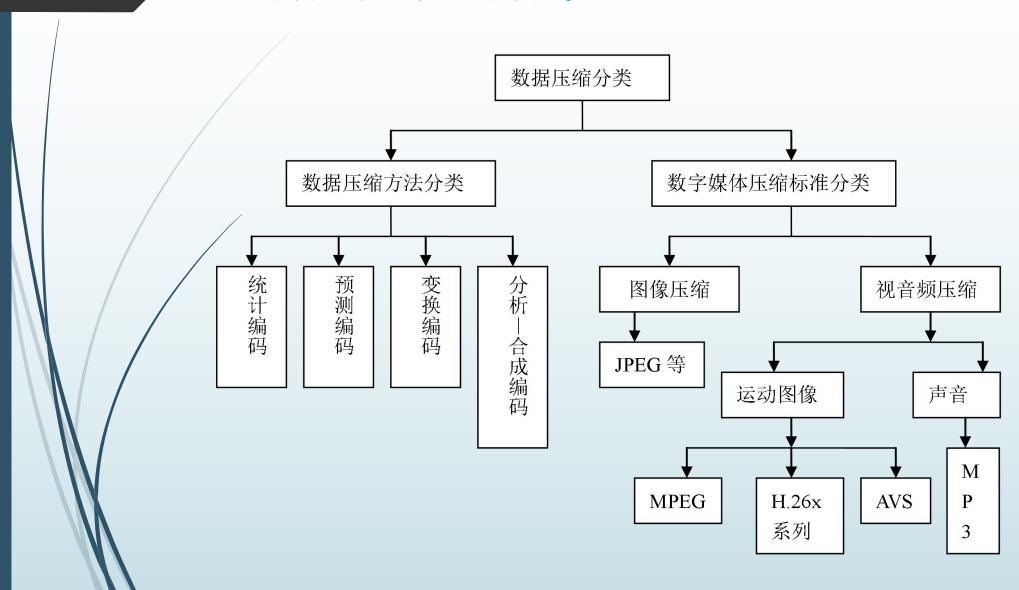


34	34	34	34	34
34	34	34	34	34
34	34	34	34	34
34	34	34	34	34
34	34	34	34	34



25 34

数据压缩的分类



数据压缩的分类

- ★按信息压缩前后比较是否有损失,可以划分有损压缩和无损压缩。
 - ■无损压缩指使用压缩后的数据进行重构,重构后的数据与原来的数据完全相同。常用的无损压缩编码有霍夫曼(Huffman)编码和行程编码。
 - ■有损压缩是指使用压缩后的数据进行重构,重构后的数据与原来的数据有所不同,但不影响人对原始资料表达的信息造成误解。

数据压缩的分类

- ▶按数据压缩编码的原理和方法进行划分
 - ●统计编码(熵编码),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。
 - ■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。
 - ●变换编码是将图像时域信号转换为频域信号进行处理。
 - ●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。

——信息论和熵

- →数据压缩技术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,数据压缩的理论是信息熵。
- 熵是信息量的度量方法,它表示某一事件出现的消息越多,事件发生的可能性就越小,数学上就是概率越小。

---信息量

- ●信息量是指信源中某种事件的信息度量或含量。一个事件出现的可能性愈小,其信息量愈多,反之亦然。
- 若 p_i 为第i个事件的概率为 $0 \le p_i \le 1$,则该事件的信息量为

$$I_i = -\log_2 p_i$$

■一个信源包括的所有数据叫数据量,而数据量中包含有冗余信息。

信息量 = 数据量-冗余量

---信息量

- ●信源X发出的x_j(j=1,2,…,n)共n个随机事件的信息 量的统计平均,即
- $\rightarrow H(X) = E\{I(xj)\} = -\sum_{i=1}^{n} P(x_i) \cdot \log_2 P(x_i)$
 - H(X)称为信源X的"熵",即信源X发出任意一个随机变量的平均信息量。
- ■其中,等概率事件的熵最大,假设有N个事件, 此肘熵为:
- $-H(X) = -\sum_{j=1}^{N} \frac{1}{N} \log_2 \frac{1}{N} = \log_2 N$

——平均码字长度

Bk为第K个码字的长度(二进制代码的长度),其相应出现的概率为Pk,则码字平均长度R:

$$R = \sum_{k=1}^{m} B_{K} P_{k}$$

——平均码字长度

- ■平均码长与信息熵之间的关系为:
 - ■R≥H 有冗余, 不是最佳。
 - ■R < H不可能。
 - ■R = H(X)最佳编码
 - ■熵值H为平均码长R的下限。

---编码效率

编码效率

Y=H/R*100%

其中,H为信息的熵,R为平均码字长度.

——变长最佳编码定理

在变长编码中,对出现概率大的信息符号赋予短码字,而对于出现概率小的信号符号赋予长码字,如果码字长度严格按照所对应符号出现概率大小逆序排列,则编码结果平均码字长度一定小于任何其他排列方式.

变长最佳编码定理是哈夫曼编码方法的理论基础.

——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。
- ■由于Huffman编码所得到的平均码字长度可以接近信源的熵,因此在变长编码中是最佳的编码方法,故也称为熵编码。

——霍夫曼编码

基本思想:

●统计一下符号的出现概率,建立一个概率统计表,将最常出现(概率大的)的符号用最短的编码,最少出现的符号用最长的编码。

——霍夫曼编码

算法实现:

- ■第一步:将符号按出现的概率由大到小排列;
- ■第二步:将最小两个概率相加,形成一个新的概率集合,再按第一步的方法重新排序,如此重复直到只有两个概率为止;
- ■第三步:分配码字:码字分配从最后一步开始反向进行,对最后两个概率分别赋予O、1,如此反向进行到开始的概率排列。

——霍夫曼编码

■ 举例说明:

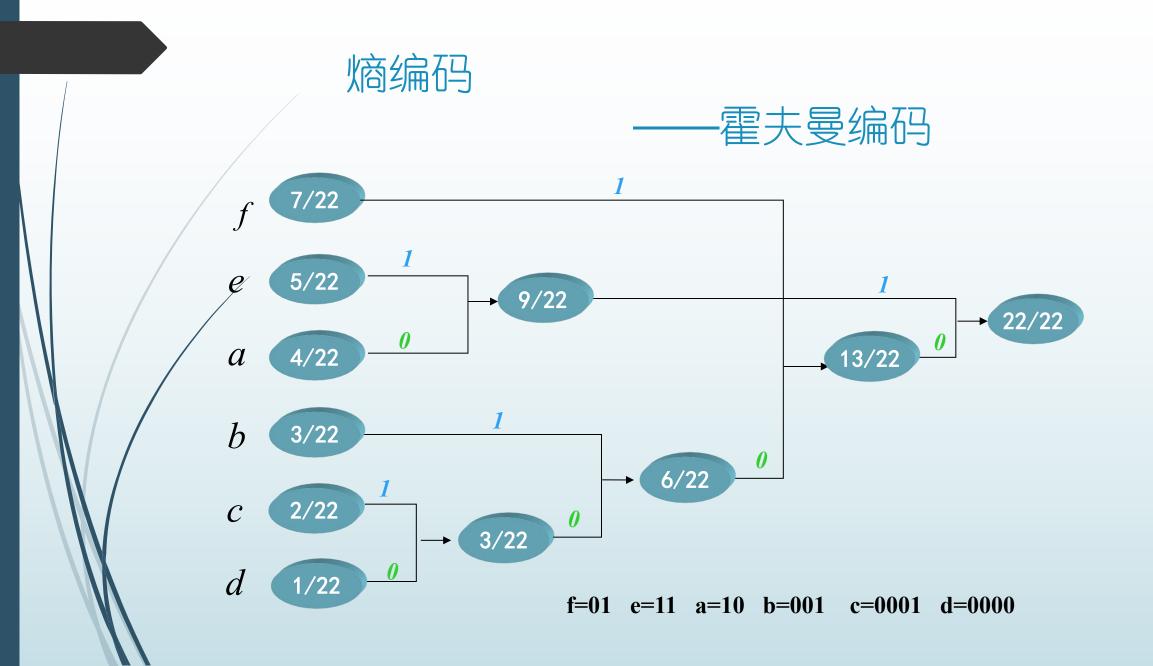
输入数据序列如下:

aaaa bbb cc d eeeee ffffff

3 2

7

(共22*8=176 bits)



编码效率

► Y=H/R*100%=99.2%

H为数字图象的熵,R为平均码字长度.

$$H = -\sum_{k=1}^{M} P_k \log_2 P_k = 2.39$$

$$R = \sum_{k=1}^{m} B_{K} P_{k} = 2.41$$

——霍夫曼编码

经过Huffman编码之后的数据为:

(共
$$7*2+5*2+4*2+3*3+2*4+1*4=53$$
 bits)

——霍夫曼编码

Huffman编码对不同的信源其编码效率不同,适合于对概率分布不均匀的信源编码。

练习

■已知输入图象的灰度级为a,b,...,h,各灰度级出现的概率分别为 0.20,0.19,0.18,0.17.0.15,0.10,0.005,0.005, 要求

0.20,0.19,0.10,0.17.0.13,0.10,0.003,0.003,安水对该图象用Huffman编码方法进行编码,写出各灰度级对应的码字,并求出平均码字长度和编码效率。

-游程编码(RLE-Run Length Encoding)

游程编码也是一种熵编码,通常应用于图像压缩中。通过改变图像的描述方式来实现压缩,即将一行灰度值相同的相邻像素用一个计数值和该灰度值来代替。

基本概念:

●游程:具有相同灰度值的像素序列。

如图象序列<u>aaaa</u> <u>bbb</u> <u>cc</u> <u>d</u> <u>eeeee</u> <u>fffffff</u> 。

■用游程的灰度和行程的长度代替游程本身。

——游程编码

■RLE编码举例:

编码前: aaaa bbb cc d eeeee fffffff

(共22*8=176 bits)

编码后: 4a3b2c1d5e7f

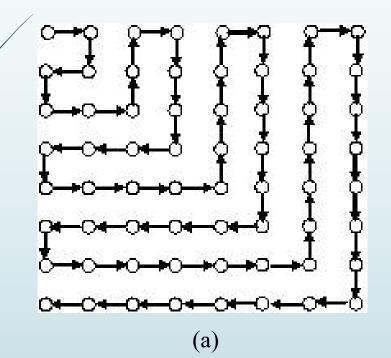
(共12*8=96 bits)

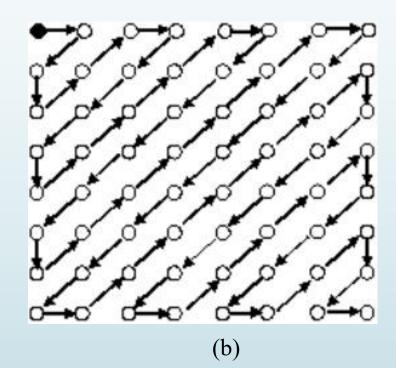
——二维游程编码

将二维排列的像素,采用某种方式转化成一维排列的方式(游程扫描)后按照一维行程编码方式进行编码。

——二维游程编码

下图是两种典型的二维游程编码的扫描方式,也有其他如 自左向右,自上而下等不同方式:





——游程编码

●分析:

- ■对于有大面积色块的图像,压缩效果很好。
- ■对于纷杂的图像,压缩效果不好。

——游程编码的应用

- 传真件中一般都是白色比较多, 而黑色相对比较少。所以可能常常会出现如下的情况:
- **▶** 500w 3b 470w 12b 4w 3b 3000w

- **因汝**: 2048<3000<4096
- ▶ 所以: 计数值必须用12 bit来表示

——游程编码的应用

- 既然已经可以预知白色多黑色少,所以可对白色和黑色的 计数值采用不同的位数。
- ▶ 以这个例子,可以定义:
- 白色: 12 bit, 黑色: 4 bit

——游程编码的应用

■编码为:

500w, 3b, 570w, 12b, 4w, 3b, 3000w

- ▶编码位数为: 12, 4, 12, 4, 12, 4, 12
- ▶所需字节数为: 4*12+3*4=60bit

练习

设图像为:

$$F = \begin{bmatrix} 60 & 60 & 59 & 59 \\ 60 & 60 & 59 & 57 \\ 60 & 59 & 60 & 57 \\ 59 & 60 & 57 & 57 \end{bmatrix}$$

对其进行二维游程编码(注意游程扫描方式)。

—— DPCM压缩技术

差分脉冲编码调制DPCM压缩技术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。如当前像素的灰度颜色的信号数值上与其相邻像素是比较接近的,那么,当前像素的灰度或颜色信号的数值可由前面已出现的像素值,进行预测 (估计) 得到一个预测值,将实际值与预测值求差,对这个差值信号进行编码和传送。

■编码思想

- ■认为相邻像素的信息有冗余。当前像素值可以 用以前的像素值来获得。
- ■用当前像素值fn,通过预测器得到一个预测值^fn,对当前值和预测值求差,对差编码,作为压缩数据流中的下一个元素。由于差比原数据要小,因而编码要小,可用变长编码。大多数情况下,fn的预测是通过m个以前像素的线性组合来生成的。

$$f_n = round[\sum_{i=1}^m \alpha_i f_{n-i}]$$

round为取最近整数,αi为预测系数。

注意: 前m个像素不能用此法编码。

编码步骤:

- ■第一步:对每一个像素灰度f(x,y),由前面的值,通过预测器,求出预测值^f(x,y);
- 第二步: 求出预测误差;e(x,y) = f(x,y) ^f(x,y)
- ■第三步:对误差e(x,y)编码,作为压缩值;
- ■重复上述步骤。

举例:

$$f_n = \text{round}[\sum_{i=1}^m \alpha_i f_{n-i}]$$

 $F = \{154, 159, 151, 149, 139, 121, 112, 109, 129\}$

m = 2
$$\alpha = 1/2$$
预测值
$$f2 = 1/2 * (154 + 159) \approx 156 \qquad e2 = 151 - 156 = -5$$

$$f3 = 1/2 * (159 + 151) = 155 \qquad e3 = 149 - 155 = -6$$

$$f4 = 1/2 * (151 + 149) = 150 \qquad e4 = 139 - 150 = -11$$

$$f5 = 1/2 * (149 + 139) = 144 \qquad e5 = 121 - 144 = -23$$

$$f6 = 1/2 * (139 + 121) = 130 \qquad e6 = 112 - 130 = -18$$

$$f7 = 1/2 * (121 + 112) \approx 116 \qquad e6 = 109 - 116 = -7$$

$$f8 = 1/2 * (112 + 109) \approx 110 \qquad e6 = 129 - 110 = 19$$

差分脉冲编码调制压缩技术小结

差分脉冲编码调制DPCM压缩技术实现比较简单,但压缩能力不高,抗干扰能力差,且对传输中的误差有积累现象。

——离散傅里叶变换

■ 离散傅里叶变换

傅里叶分析是将原始信号分解成不同频率成分的正弦波,将时域信号转变为频域信号的一种数学方法。

离散傅里叶变换 (Discrete Fourier Transform, DFT): 将时域上的波形分解成正弦波的过程;

其逆变换表示为IDFT (Inverse Discrete Fourier Transform, IDFT): 将频域上的波形叠加,映射到时域上。

——离散傅里叶变换

■ 离散傅里叶变换

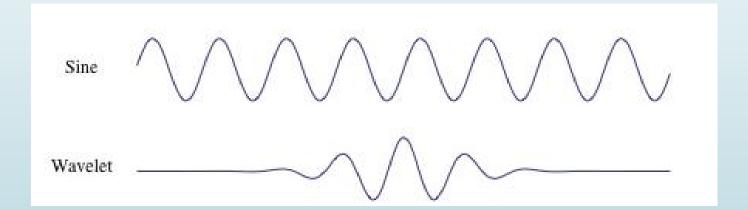
傅立叶变换是一种可逆变换,即它允许原始信号和变换过的信号之间互相转换。

在任意时刻只有一种信息是可用的,也就是说,在傅立叶变换后的频域中不包含时间信息,逆变换后的时域中不包含时间信息。

DFT
$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j n k \frac{2\pi}{N}}$$
 IDFT $x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j n k \frac{2\pi}{N}}$

——小波变换

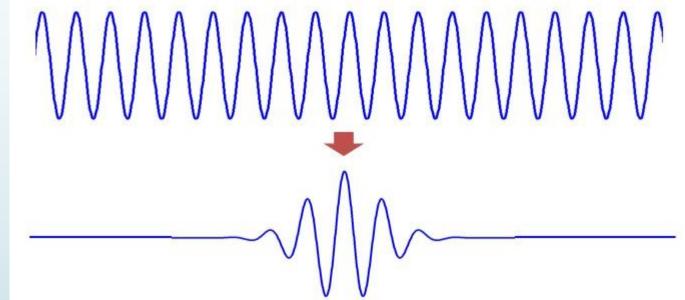
- 多分辨率分析:我们在分析信号的低频部分的时候,只需要较高的频率分辨率和较低的时间分辨率就能够很好的体现低频的信息;对于高频部分,就需要较高的时间分辨率和较低的频率分辨率就能够很好的体现高频的信息。
- 拿图像信号举例,图像的低频部分显示的是图像的基本信息,而高频部分更多的是细节信息。
- 小波变换主要应用于图像编码,其基本思想是,把图像进行多分辨率分析,分解成不同空间、不同频率的子图像,再对子图像进行系数编码。
- ▶ 小波是指一个有限周期内平均值为零的波形。



——小波变换

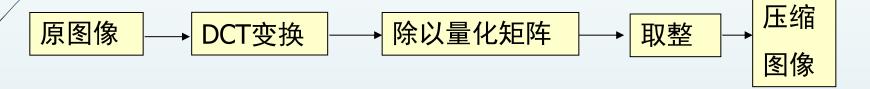
- → 小波变换就是选择适当的基本小波,通过对小波的平移、伸缩而形成的一系列的小波,这簇小波作为基波可以构成一些列嵌套的(信号)子空间。
- 将待分析的图像等信号投影到各个子空间上,以观察相应的特性。
- 在数字图像中使用的小波变换为 离散小波变换 (DWT, Discrete Wavelet Transform)
- 小波图像的各个频带分别对应了原图像在不同尺度和不同分辨率下的细节以及一个由小波变换分解级数决定的最小尺度、最小分辨率下对原始图像的最佳逼近。

$$F(w) = \int_{-\infty}^{\infty} f(t) * e^{-iwt} dt \implies WT(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) * \psi(\frac{t - \tau}{a}) dt$$

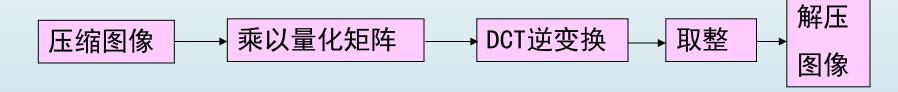


——DCT变换编码

- **DCT**变换编码是有损编码
- 1)编码过程:



2)解码过程:



——DCT变换编码

DCT变换是图象压缩标准中常用的变换方法,如JPEC标准中将图象按照8x8分块,然后利用DCT变换编码实现图象的压缩。

——DCT变换编码

例:

原图像为:
$$F = \begin{bmatrix} 59 & 60 & 58 & 57 \\ 61 & 59 & 59 & 57 \\ 62 & 59 & 60 & 58 \\ 59 & 61 & 60 & 56 \end{bmatrix}$$

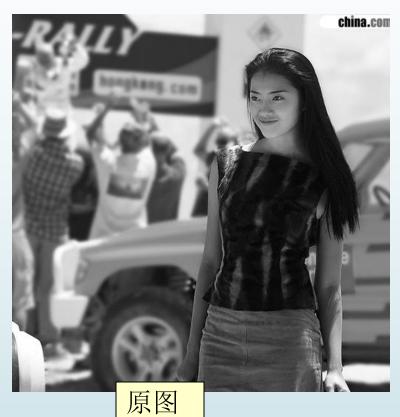


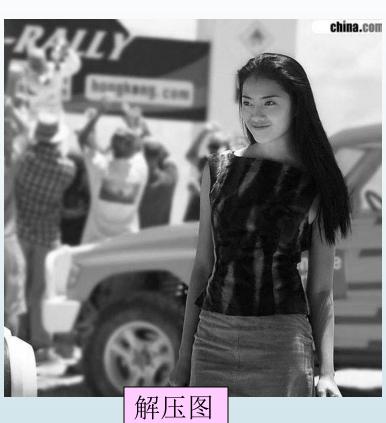
$$D_{1} = \begin{bmatrix} 120.5 & 119.5 & 118.5 & 114.0 \\ -0.27 & -0.65 & -1.58 & 0.38 \\ -2.50 & 1.50 & -0.50 & -1.00 \\ 0.65 & -0.27 & 0.11 & 0.92 \end{bmatrix}$$

除以量化矩阵, 取整

$$C = \begin{bmatrix} 16 & 11 & 11 & 16 \\ 12 & 12 & 14 & 19 \\ 14 & 13 & 16 & 24 \\ 14 & 17 & 22 & 29 \end{bmatrix}$$

——DCT变换编码





混合编码

● 例:

<u>aaaa bbb cc d eeeee fffffff</u> (共22*8=176 bits)

3 2 1 5

游程编码: 4a3b2c1d5e7f

(共6* (8+3) =
$$66Bits$$
)

176



66

混合编码

```
<u>aaaa bbb cc d eeeee fffffff</u> (共22*8=176 bits)

4 3 2 1 5 7
```

Huffman编码:

 $(\pm 7*2+5*2+4*2+3*3+2*4+1*4=53 \text{ bits})$

混合编码

Hufman与行程编码混合:

41030012000110000511701

(共: 3+2+3+3+3+4+3+4+3+2+3+2=35 bits)

练习

设图像为:
$$F = \begin{bmatrix} 59 & 60 & 58 & 57 \\ 61 & 59 & 59 & 57 \\ 62 & 59 & 60 & 58 \\ 59 & 61 & 60 & 56 \end{bmatrix}$$

请对其用混合编码方式进行有损编码。

提示: 1、改变灰度值(有损编码),选择合适的二维行程扫描方式,对一维数据序列进行混合编码(霍夫曼编码+行程编码)。

2、DCT变换编码+行程编码+哈夫曼编码