第4章 图像编码与压缩

本章重点:

- > 图像编码与压缩的基本概念、理论及其编码分类
- > 常用的无损压缩方法
- > 常用的有损压缩方法

第四章 图像编码与压缩

- >4.1 图像编码必要性与可能性
- >4.2 图像编码分类
- > 4.3 图像编码评价准则
- > 4.4 图像编码模型
- > 4.5 无损压缩
- ▶4.6 有损压缩
- > 4.7 JPEG图像编码压缩标准
- > 4.8 MPEG视频编码压缩标准
- > 4.9 小结

4.1 图像编码的必要性与可能性

4.1.1图像编码的必要性

- ▶数字图像庞大数据要求计算机处理速度、存储容量。
- > 数据量需要压缩。
- >媒体信息(特别是动态视频)数据量非常大。

例:一幅1024×1024分辨率24位真彩色图像,数据量为:

$$1024 \times 1024 \times 8 \times 3/8 = 3MB$$
;

若以30帧/秒播放,每秒数据量为:

$$3 \times 30 = 90 \text{MB}$$

例: 陆地卫星LandSat-3分辨率为2340×3240, 四波段, 采样精度7位, 则一幅图像数据量为:

2340 ×3240 ×7 ×4=212Mb

按每天传输30幅计,每天数据量为:

- ▶没有图像编码与压缩技术,大容量图像信息存储与传输是 难以实现。
- ▶传送图像,要求数据量压缩。信道带宽、通信链路容量一定,采用编码压缩技术,减少传输数据量,提高通信速度。

4.1.2图像编码的可能性

- > 图像、声音具有很大压缩潜力。
- ▶ 位图图像存储方式为例,像素之间行和列方向具有相关性;
- > 允许一定限度失真,图像数据很大程度压缩。
- > 失真在人眼允许误差范围内, 前后图像难觉察差别。
- 数据压缩技术是多媒体系统关键技术。
- 数据压缩是基于原始信源的数据的冗余度。

图像数据冗余

- > 空间冗余 、时间冗余
- > 结构冗余、知识冗余
- > 视觉冗余、 图像区域的相同性冗余
- > 纹理的统计冗余

空间冗余





18 17 19 17 21 29 45 59 65 59 58 66 67 61 69 60 22 20 20 17 19 25 51 65 82 90 84 74 73 78 57 56 27 23 23 18 17 21 42 47 66 90 97 90 84 86 58 61 28 25 24 21 19 21 24 24 30 50 77 95 93 84 79 77 26 24 24 23 22 23 26 38 37 28 43 77 93 88 10291 24 20 20 21 22 23 40 68 75 47 29 48 80 97 10997 23 16 15 17 19 19 36 55 73 68 44 33 58 92 108103 23 14 11 13 15 15 16 12 36 69 64 35 42 77 108110 18 21 20 19 16 7 8 19 18 13 13 18 17 5 11 23 48 57 38 45 84 122128 18 10 13 28 35 29 42 51 53 46 40 63 104140137 22 24 15 18 35 46 58 77 82 60 35 42 90 140152140 27 19 21 35 44 46 53 52 38 36 72 131 172 164 146 26 24 31 46 54 28 14 13 31 70 128174187180156 20 26 36 60 88 10174 55 63 99 138178196186190163 22 28 50 91 133152149140160189197201198182192165



大部分区域所有像 素值相同。

▶物体采样点

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

か

の

か

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

の

时间冗余

> 主要指视频相邻帧之间的冗余。





视觉冗余

- >人视觉系统对图像场敏感性是非均匀和非线性;
- ▶视觉亮度和色彩敏感度不同,产生视觉冗余。
- >原始图像的记录数据,是线性和均匀的。

结构冗余

- > 纹理区的像素值间存在明显分布模式, 称为结构冗余。
- > 图像理解与知识的相关性,如人脸固定结构。

图像区域的相同性冗余

▶图像两个或多个区域像素相同或相近,产生数据重复性存储,即图像区域的相似性冗余。





4.2图像编码分类

- 解压重建图像和原图像是否具有误差,图像编码与压缩方法分为无误差(亦称无失真、无损、信息保持)编码和有误差(有失真或有损)编码两大类。
- ▶据编码作用域划分,图像编码分为空间域编码和变换域编码两大类。
- ▶編码技术,分为预测编码、变换编码、统计编码、轮廓编码、模型编码等。

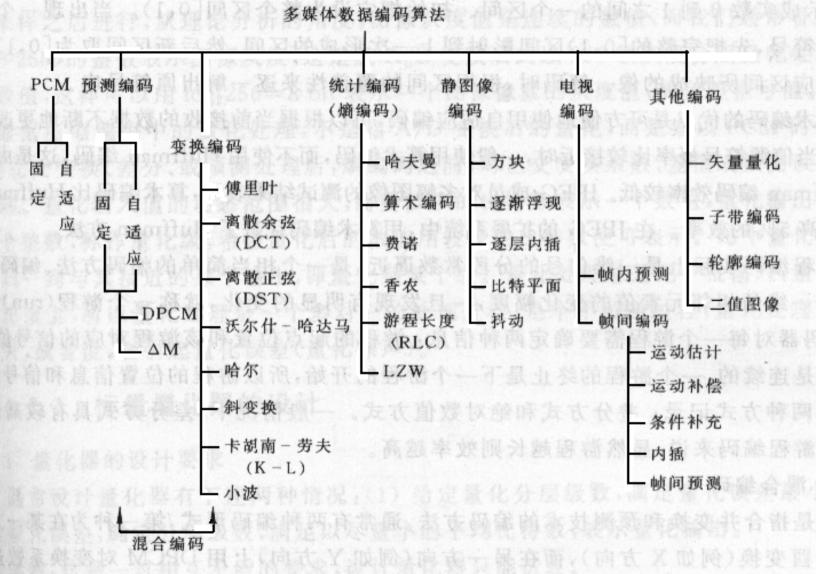


图 4.1 多媒体数据编码分类

多媒体数据编码

- ➤ 脉冲编码调制(PCM)
- > 预测编码
- > 变换编码
- > 统计编码(熵编码)
- > 混合编码
- > 静态图像编码
- > 动态图像编码
- > 其他编码



4.3 图像编码评价准则

- ▶压缩图像质量评价,由解码图像与原始图像的差异度表达。
- >两幅图像偏离度的测度,称为保真度(逼真度)准则。
- 准则可分为两大类:客观保真度准则和主观保真度准则。

4.3.1 客观保真度准则

- > 客观保真度准则:
- ▶原图像和解码图像之间的均方根误差和均方根信噪比两种。 均方根误差:

$$e_{rms} = \left\{ \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left[\hat{f}(x,y) - f(x,y) \right]^2 \right\}^{1/2}$$
 (4-2)

均方信噪比:

$$SNR_{ms} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \hat{f}(x,y)^2 / \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left[\hat{f}(x,y) - f(x,y) \right]^2$$
 (4-3)

上式求平方根,得到均方根信噪比。

4.3.2主观保真度准则

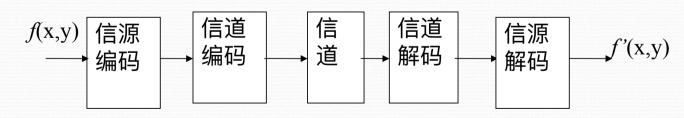
- >相同客观保真度的不同图像,人的视觉效果可能不同。
- ▶主观保真度,统计平均意义下的度量准则,图像细节无法 反映。
- ▶常用方法是对一组(不少于20人)观察者显示图像,并将该 图像的评分取平均,用来评价一幅图像的主观质量。

例如可用{-3, -2, -1, 0 , 1, 2, 3}来代表主观评价{很差, 较差, 稍差, 相同, 稍好, 较好, 很好}。 表4.1 电视图像质量评价尺度

评分	评价	说明
1	优秀	图像质量非常好,如同人能 想象出的最好质量
2	良好	图像质量高,观看舒服,有 干扰但不影响观看
3	可用	图像质量可以接受,有干扰 但不太影响观看
4	刚可看	图像质量差,干扰有些妨碍 观看,观察者希望改进
5	差	图像质量很差,几乎无法观 看
6	不能用	图像质量极差,不能使用

4.4 图像编码模型

- > 图像压缩系统包括: 编码器和解码器。
- ▶图像f(x,y)输入编码器,输入数据生成一组符号。
- >信道传输后,编码表达符号由解码器重构,生成输出图像。

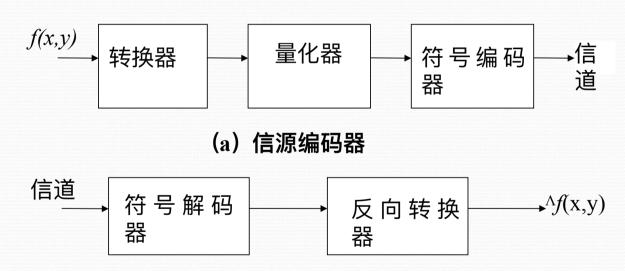


一个常用于图像压缩系统模型

4.4.1信源编码器和信源解码器

- > 信源编码器减少(或消除)图像冗余(像素间或视觉)。
- 从原理来看主要分为三个阶段:
 - 1) 输入数据转换为降低像素间冗余的数据集合。
 - 2) 去除原图像信号相关性。
 - 3) 找一种近于熵,又利于计算机处理编码方式。
- > 信源解码器包含两部分:符号解码器和反向转换器。

> 编码器模型



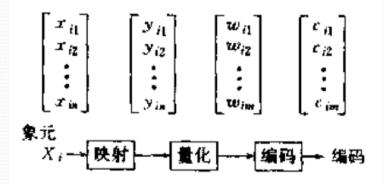
(b) 信源解码器

4.4.2信道编码器和解码器

- ▶信道带噪声或易于出错误时,信道编码器和解码器在整个 译码解码中扮演重要角色。
- ▶信道编码器和解码器通过向信源编码数据中插入预制的冗余数据来减少信道噪声的影响。
- ➤信道编码技术是由R. w. Hamming提出, 即向被编码数据中加入足够的位数以确保可用码字间变化位数最小。

信源编码器与信源解码器

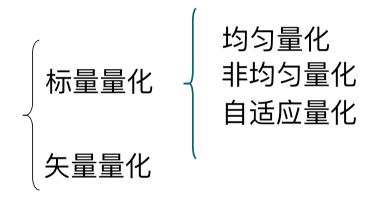
> 信源编码器减少或消除编码、像素间或心理视觉冗余。



- > 映射: 信号映射后可用较少比特编码, 如差分。
- >量化:映射后值,均匀量化或非均匀量化。
- ▶ 多对一映射量化,一批输入量化为一个输出,过程不可逆, 信息丢失引起量化误差(量化噪声)。

量化方法和量化特性

> 量化方法



- >标量量化:数据逐个进行量化。
- ▶矢量量化:数据分组,每组K个数构成矢量,以矢量为单位量化,提高压缩比。

矢量量化的编码与解码

- ➤输入量: 待编码K维矢量(尺寸n×n图象块中的n²个像素)
- ➤ 码本C: 一个具有L个 K维矢量集合(一个长度为L的表,表每一个分量是一个K维矢量y, 称其为码字)。
- >矢量量化编码过程:

码本C中搜索一个与输入矢量最接近码字 y_i (i=1,2,...,L);传输时并不传送码字 y_i 本身,只传送其下标号 i 。下标所需比特数仅 log_i L;

编码器

- \rightarrow 编码器输入 $w_{i,}$ w_{i} 取M个值 $w_{1,}$ $w_{2,}$..., w_{m} 之一,输出二进制码 c_{i} 。
- 编码器不会引入误差。
- ▶ 编码器使M个可能输入,能分配一个唯一二进制码字。
- > 对 w_1 , w_2 , w_3 分别赋予码字, $c_1=0, c_2=1, c_3=01$ 。
- **>** 比特流0011,可译: $c_1c_1c_2c_2$,也可译 $c_1c_3c_2$,不唯一。
- ➤ 赋予码字c1=0, c2=10, c3=11, 比特流0011, 唯一码c1c1c3。

> 信源解码器包含两个部分:符号解码器和反向映射器。

信道
$$\longrightarrow$$
 符号解码 \longrightarrow 反向映射 $\longrightarrow_{\widehat{f}(x,y)}$

4.5无损压缩

- > 无损压缩可以精确无误地从压缩数据中恢复出原始数据。
- 常见无损压缩技术:基于统计概率方法和基于字典技术。
- > 霍夫曼(Huffman)编码, 基于统计概率的方法,
- ▶据信息论的变长编码定理和信息熵知识,较短码表示概率 大符号,长码代表概率小符号,实现数据压缩。
- > 统计编码方法利用概率分布特性码。

- ◆ 基于字典技术的数据压缩技术有两种:
- ➤游程编码(RLC, Running Length Coding),适用灰度级少、数据相关性强的图像压缩。
- ▶但最不适用于每个像素都与周围像素不同的情况。
- ▶LZW编码,基于字典T,将输入字符串映射成定长(通常为12位)的码字。
- Lempel-Ziv-Welch Encoding

4.5.1霍夫曼编码

◆ 1.理论基础

一个事件集合 $x_1, x_2,...,x_n$,一个基本概率空间,其相应概率为 $p_1, p_2,...,p_n$,且 $p_1+p_2+...+p_n=1$ 。每个信息的信息量为:

$$I(x_k) = -\log_a(p_k) \tag{4-9}$$

在概率空间中每一事件的概率不相等时,平均<mark>不肯定程</mark> 度或平均信息量叫作熵 H,则:

$$H = E\{I(x_k)\} = \sum_{k=1}^{n} p_k I(x_k) = \sum_{k=1}^{n} -p_k \log_a p_k$$
 (4-10)

> 图象熵:

设数字图像像素<mark>灰度级集合</mark>为($W_1,W_2,...,W_M$),其对应概率分别为 $P_1,P_2,...,P_M$,则图像信息熵H为:

$$H = -\sum_{k=1}^{n} P_k \log_a P_k$$

- ▶a取2时,H单位为比特。a取e时,H单位为奈特。图像编码中a取2。
- ▶图像信息熵是图像平均信息量,表示图像各灰度级比特数的统计平均值。
- > 等概率事件熵最大。
- ▶信息熵是无失真编码理论极限。低于此极限无失真编码方法不存在。

例:设8个随机变量具有同等概率为1/8,计算信息熵H。

解:根据公式可得:

$$H=8*[-1/8*(log2(1/8))=8*[-1/8*(-3)]=3$$

> 编码效率

一般情况下, 编码效率公式:

η**=H/R%**

H为信息熵,R为平均码字长度。

> 平均码字长度

设 β_k 为图像第k个码字 C_k 的长度(二进制代码位数), 其出现概率为 P_k ,该图像赋予码字的平均码长R为:

$$R = \sum_{k=1}^{n} \beta_k P_k$$

- ▶根据信息熵编码理论,证明在R≥H条件下,总可设计出 某种无失真编码方法。
- >编码结果远大于H,表明编码效率很低,比特数太多。
- >编码结果R等于或接近于H,编码方法称为最佳编码。
- ➤编码结果使R<H,必然丢失信息而引起图像失真。允许 失真的编码方法。

图像熵编码方法

- > 熵编码使编码后图像平均比特数R尽可能接近图像熵H。
- ▶ 灰度级数出现概率赋予码字长度,概率大的灰度级用短码字,反之,用长码字。
- > 编码获得的平均码长最短。
- > 常用熵编码方法:
- > Huffman编码
- > Fano-shannon编码
- > 游程编码
- > 算术编码

◆ 2. Huffman编码

- > Huffman编码是1952年由Huffman提出一种编码方法。
- ▶根据信源数据符号发生的概率进行编码。信源数据出现概率越大的符号,码越短;概率越小符号,码长越长。
- > 源数据用尽可能少的码符号表示。
- > 最佳的变长编码方法。

Huffman编码具体方法:

>设信源A的信源空间为:

$$[A \bullet P]: \begin{cases} A: a_1 & a_2 & \cdots & a_N \\ P(A): P(a_1) & P(a_2) & \cdots & P(a_N) \end{cases}$$

其中 $\sum_{i=1}^{N} P(a_i) = 1$,用r个码符号集 $X: \{x_1, x_2, \dots, x_r\}$ 编码信源A中符号($i = 1, 2, \dots, N$)。

编码方法:

- (1) 信源符号按概率大小排序;
- (2) 两个最小概率的元素之概率加起来;
- (3) 该概率之和同其余概率由大到小排队,再把两个最小概率加起来,再<u>重新排队</u>;
- (4) 重复到最后只剩下两个概率为止。

例1:设有编码输入 $X = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ 。其频率分布分别为 $(x_1) = 0.4$, $P(x_2) = 0.3$, $P(x_3) = 0.1$, $P(x_4) = 0.1$, $P(x_5) = 0.06$, $P(x_6) = 0.04$,现求其最佳霍夫曼编码 $W = \{w_1, w_2, w_3, w_4, w_5, w_6\}$ 解:Huffman编码过程:

符号	概率	1	2		3	4
						 0.6
x_1	0.4	0.4	0.4		0.4	0.4
x_2	0.3	0.3	0.3		0.3	
x_3	0.1	0.1	0.2	•	0.3	
\mathcal{X}_4	0.1	0.1	0.1			
x_5	0.06	 0.1				
x_6	0.04					

0.6赋予0, 0.4赋予1, 0.4传递到x1, x1编码便是1。

0.6来自0.3+0.3, x2的0.3是单独概率, x2赋予0。

0.2和0.1求和的0.3赋予1。

x2编码为00,剩余元素的前两个码为01。0.1赋予1,0.2赋予0。以此类推,最后得到诸元素的编码如下:

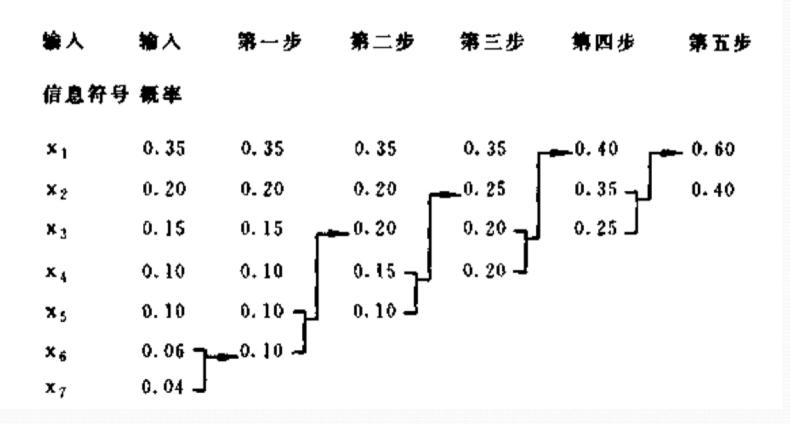
元素 <i>x1</i>	x1	x2	<i>x</i> 3	x4	<i>x</i> 5	<i>x</i> 6
概 率 P(x1)	0.4	0.3	0.1	0.1	0.06	0.04
编码w1	1	00	011	0100	01010	01011

经霍夫曼编码后,平均码长为:

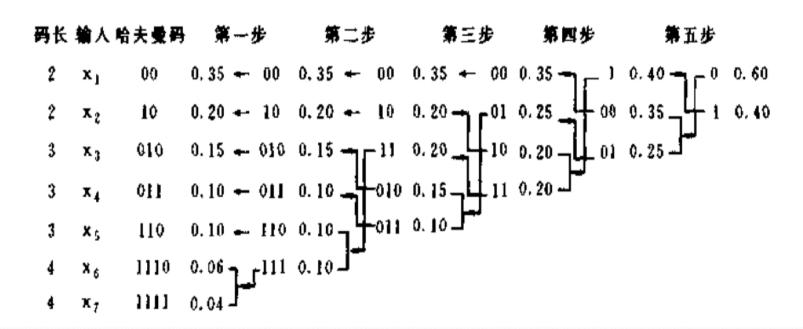
$$\overline{B} = \sum_{i=1}^{6} P(\omega_i) n_i$$
=0.4*1+0.30 * 2+0.1 * 3+0.1 * 4+0.06 * 5+0.04 * 5
=2.20(bit)

信源熵为H=2.14 bit,编码后计算平均码长为2.2bit,非常接近于熵。Huffman编码是较好编码。

例2:Huffman编码举例



Huffman码字的构成



码字的平均码长 7 以下面公式计算

$$\overline{N} = \sum_{j=1}^{n} P_{j} L_{j}$$

$$= \sum_{j=1}^{n} (P_{j} L_{j})$$

$$= (0.35 + 0.20) \times 2 + (0.15 + 0.10 + 0.10) \times 3 + (0.06 + 0.04) \times$$

$$= 2.55 \text{bits/pel}$$

뼭

$$H = -\sum_{j=1}^{n} P(x_{j}) \cdot \log_{2} P(x_{j})$$

$$= -\sum_{j=1}^{l} P(x_{j}) \cdot \log_{2} P(x_{j})$$

$$= -\left[0.35 \cdot \log_{2} 0.35 + 0.20 \cdot \log_{2} 0.20 + 0.15 \cdot \log_{2} 0.15 + (0.10 \cdot \log_{2} 0.10) \times 2 + 0.06 \cdot \log_{2} 0.06 + 0.04 \cdot \log_{2} 0.04\right]$$

$$= 2.13 \text{bits/pel}$$

通过这个例子,总结如下特点:

(1) 平均码长 N>H(熵);

- (2) 平均码长 \overline{N} <3 bits(等长码需要的比特数);
- (3) 保证解码的唯一性,短码字不构成长码字的前缀;
- (4) 在接收端需保存一个与发送端相同的哈夫曼码表

- > 二叉树方法实现Huffman编码方法较便利;
- ➤ Huffman编码最佳,其构造码不唯一,但平均码长相同, 不影响编码效率和数据压缩性能。
- ➤ Huffman码的码长参差不齐,存在一个输入、输出速率匹配问题,解决办法设置一定容量的缓冲存储器
- ➤ Huffman码在存储或传输过程中,如出现误码,可能会引起误码的连续传播
- > Huffman编码对不同信源的编码效率也不尽相同。

- ▶ Huffman编码应用时,需要与其他编码结合,进一步提高数据压缩比。
- > Huffman编码方法传输时,需要同时传输Huffman编码表。
- > Huffman编码需要多次排序

4.5.2 香农一费诺编码

- > 霍夫曼编码法需多次排序, 十分不便;
- ▶ 费诺(Fano)和香农(Shannon)分别提出简单编码类似方法。

①把 $x_1,...,x_n$ 按概率大小排列,分成两组 $x_1,...,x_k,x_{k+1},...,x_n$

$$\sum_{i=1}^k P(x_i) \approx \sum_{j=k+1}^n P(x_j)$$

- ②把两组分别按0,1赋值。
- ▶分组、赋值,不断反复,直到每组只有一个输入为止。每 个赋值排列起来,费诺—香农编码。

以前面哈夫曼编码的例子进行香农 - 费诺编码:

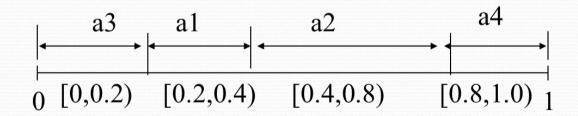
输入	概率								
x_1	0.4	0	0 0						
x_2	0.3		0	10					
x_3	0.1			0	0	1100			
x_4	0.1	1	1	1	1	1	0	1	1101
x_5	0.06			1	0	1110			
x_6	0.04			1	1	1111			

- >理论上, Huffman方法对源数据流编码可达最佳编码效果。
- ➤ 计算机中存储的最小单位是"bit",在一些情况下,实际压缩比大于理论压缩比。
- ➤ 例如源数据流由X和Y两个符号构成,出现概率为2/3和1/3。 字符X熵确定的最优码长为:
- \rightarrow H(X)= log(2/3)=0.585bit
- \rightarrow H(Y)= log(1/3)= 1.58bit

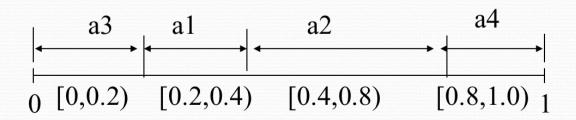
- ▶最佳编码效果,字符X码长为0.58位;字符Y码长为1.58位。
- ➤ 计算机中不可能有非整数位出现,硬件限制使编码只能按"位"进行。Huffman方法编码两个字符,x、y码分别为0和1。
- ▶ 概率大的字符 x不能给予较短的码,实际编码效果不能达到理论压缩比。

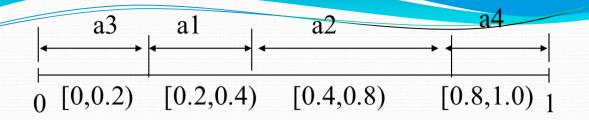
4.5.3 算术编码

- > 算术编码没有延用一个特定代码代替一个输入符号做法;
- ➤ 把待压缩整段数据映射到<mark>半开</mark>区间[0, 1]内某一区段,构造出小于1且大于或等于0的数值。
- > 数值是输入数据流的唯一可译代码。



- > 算术编码把一个信源集合表示为实数轴上区间[0,1]的一段。
- ▶信源每个元素都用一段区间。信源集合元素越多,得到区间就越小;
- ▶区间变小,需更多的数位表示,即区间作为代码原理。
- ▶算术编码先假设一个信源的概率模型,概率缩小表示信源 集区间。





> F: previous; N: new; s: start; e:end

$$N_S = F_S + C_l L_F$$
, N_S 新子区间起始位置,

Fs 前子区间起始位,

C₁ 当前符号区间左端,

 L_F 为前区间长度。

 $Ne = Ns + L_F L_N$,Ne 新子区间结束位置,

Ns 新子区间开始位置,

L_N当前符号区间长度

例子: $a_1a_2a_3$

a1 取值在[<mark>0.2</mark>, 0.4),长度0.2

a2 取值在[0.4, 0.8), 长度0.4

$$a_1 a_2$$
: $0.2 + 0.4 * 0.2 = 0.28$

$$0.28 + 0.4*0.2=0.36$$

$$a_1a_2a_3$$
: $0.28+0*0.08 = 0.28$

$$0.28 + 0.2*0.08 = 0.296$$

• 输入字符	区间长度	范围
• a1	0.2	[0.2, 0.4)
• a2	0.08	[0.28, 0.36)
• a3	0.016	[0.28, 0.296)
• a2	0.0064	[0.2864, 0.2928)
• a4	0.00128	[0.2915, 0.2928]

- 字符串A={a1 a2 a3 a2 a4}全部编码, 范围在[0.2915, 0.2928]
- 可取区间下线进行编码,浮点数表示字符串,减小存储空间

4.5.4游程编码

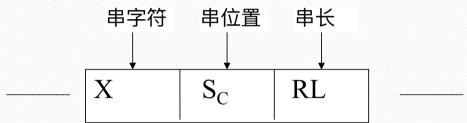
- ▶游程编码(RLC)利用空间冗余度压缩图像,属于统计编码。
- ▶ 图像某块像素采样为: $(x_1, x_2, \cdots x_M)$
- ▶像素值可分为k段,长度为l_i的连续串,每串灰度值相同

$$(x_1, x_2, \cdots x_M) \rightarrow (g_1, l_1), (g_2, l_2), \cdots, (g_k, l_k)$$

g 为每个串代表值, 为串长度。

- 串长是灰度值构成的数据流中,各灰度值重复出现长度。
- 游程长度(Run—length),简写为RL。
- 给出灰度值、对应长度及位置,容易恢复原数据流。
- 二值图像序列0000011100011111111000111
- 游程序列 533733

> RL的基本结构



- > 游程编码:定长游程编码和变长游程编码。
- >定长游程编码:编码游程位数固定,RL位数固定。
- > 灰度相同的像素数超过固定位数,进入下一轮游程编码。
- > 变长游程编码:不同范围的游程用不同位数的编码
- >表示RL位数不固定。

- > 游程编码一般不应用于多灰度图像,适合二值图像编码。
- > 只有"0"和"1"两种符号;连续出现,形成"0" 游程: L(0), "1"游程: L(1)。
- > 达到较好压缩效果,游程编码和其他编码混合使用。
- ▶ 定义游程和长度后,任何二元序列变换成游程长度序列, 简称游程序列。变换可逆,一一对应。

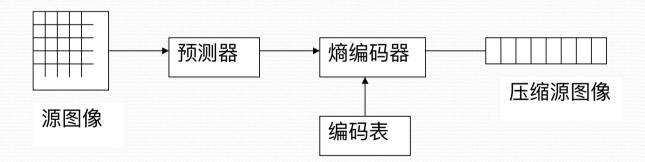
4.5.5 无损预测编码

- \rightarrow 一幅二维静止图像,空间坐标(i, j)像素点灰度为f(i, j)
- $\rightarrow \hat{f}(i,j)$ 是该点的预测灰度,也称预测值或估计值;
- ▶ 计算像素预测值,由同一扫描行的前几个像素,或是前几 行像素,前几帧邻近像素决定。
- > 实际值和预测值之间差值:

$$e(i,j) = f(i,j) - \hat{f}(i,j)$$

- > 图像统计特性,相邻像素之间有强相关性。
- > 像素值可据已知几个像素来估计,即预测。
- ▶ 预测编码:利用以往的样本值预测新样本值,样本的实际值与其预测值的误差值,对误差值进行编码。
- ▶模型理想,图像样本序列时间相关性较强,则误差信号的幅度远小于原始信号。
- > 差值信号不进行量化而直接编码就称之为无损预测编码。

> 无损预测编码器的工作原理图:



 $\hat{f}(i,j)$ 由先前三点预测定义为:

$$f(i,j) = a_1 f(i,j-1) + a_2 f(i-1,j-1) + a_3 f(i-1,j)$$

其中a1, a2, a3称预测系数, 都是待定参数。

如预测器中的预测系数是固定常数,称为线性预测。

预测误差:

$$\begin{split} e(i,j) &= f(i,j) - f(i,j) \\ &= f(i,j) - \left[a_1 f(i,j-1) + a_2 f(i-1,j-1) + a_3 f(i-1,j) \right] \end{split}$$

设a=f(i, j-1),b=f(i-1, j), c=f(i-1, j-1), $\hat{f}(i,j)$ 预测方法如下,有8种选择方法:

			选择方法	预测值 $\hat{f}(x,y)$
			0	非预测
			1	а
С	b		2	b
а	x		3	С
			4	<i>a+b-c</i>
			5	a+(b-c)/2 $b+(a-c)/2$
			6	b+(a-c)/2
			7	(a+b)/2

例:设有一幅图像, f(i-1, j-1),f(i-1, j), f(i, j-1), f(i, j)的灰度值分别为253, 252, 253, 255。

第四种选择方法预测 f(i,j)的灰度值, 计算预测误差。

解:
$$\hat{f}(i,j) = a+b-c= f(i,j-1)+ f(i-1,j)- f(i-1,j-1)$$

=252+253-253=252

预测误差

$$e(i, j) = f(i, j) - \hat{f}(i, j) = 255 - 252 = 3$$

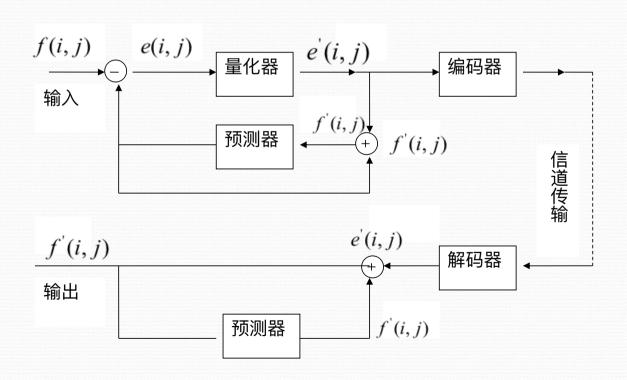
编码预测误差e(i,j)=3, 比直接编码 255 占用少的比特位。

4.6有损压缩

- > 有损编码是以丢失部分信息为代价来换取高压缩比。
- > 有损压缩方法主要有有损预测编码方法、变换编码方法等。
- ▶在预测编码中,对差值信号进行量化后再进行编码就称之 为有损预测编码。
- ➤有损预测方法有多种,差分脉冲编码调制(Differential Pulse Code Modulation, 简称DPCM), 是代表性编码方法。

DPCM(差值编码)系统,编码器和解码器组成:

- ◆ 包括发送、接收和信道传输三个部分
- ◆ 发送端由编码器、量化器、预测器和加减法器组成
- 接收端包括解码器和预测器

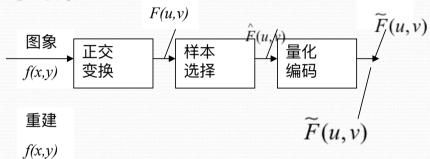


- DPCM包含量化器,量化器导致不可逆信息损失;
- 接收端解码恢复的灰度信号不是真正 f(i,j), 而是重建信号;
- 引入量化器会引起一定程度信息损失,使图像质量受损;
- 利用人眼的视觉特性,丢失不易觉察的图像信息,不会引起明显失真。

4.6.2 变换编码

- > 变换编码不直接对空域图像信号编码;
- ▶ 图像数据经某种正交变换 (如傅立叶变换DFT、离散余弦变换DCT, K-L变换等, 称之为变换域), 产生一批变换系数;
- > 这些变换系数编码处理, 达到压缩图像数据的目的。

变换编码的原理:



- > 图像经正交变换,空域总能量在变换域中得到保持,
- 像素间相关性下降,能量重新分布,并集中在变换域少数的变换系数上。
- ▶ 选择少数F(u,v)重建图像,压缩数据,重建图像误差小。

▶ 变换多采用正交函数为基础变换

矩阵 $T_{N\times N}$,有N个标量 λ_i ,i=1,2,...,N,使 $|T-\lambda_i I|=0$ 的 λ_i 是矩阵T的特征值。

N个满足 $TV_i = \lambda_i V_i$ 向量 V_i 是T的特征向量,构成一个正交基集。

X是一个 $N \times 1$ 的随机向量,X每个分量xi 是随机变量。 X的均值 M_x 由L个样本向量计算:

$$M_x \approx \frac{1}{L} \sum_{l=1}^{L} X_l$$

▶ M_x协方差矩阵

$$\Phi_{Mx} = E\{(X - M_x)(X - M_x)^T\} \approx \frac{1}{L} \sum_{l=1}^{L} X_l X_l^T - M_l M_l^T$$

协方差矩阵实对称,对角元素是随机变量的方差, 非对角元素是它们的协方差。

定义一个线性变换T,任何X向量产生一个新向量Y:

$$Y = T(X - M_x)$$

T的行向量是Mx的特征向量。

协方差矩阵是一个矩阵,其每个元素是各矢量元素间的协方差;

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} \qquad \Sigma_{ij} = \operatorname{cov}(X_i, X_j) = \operatorname{E}[(X_i - \mu_i)(X_j - \mu_j)]$$

$$\Sigma = E \left[(\mathbf{X} - E[\mathbf{X}]) (\mathbf{X} - E[\mathbf{X}])^{\top} \right]$$

$$= \begin{bmatrix} E[(X_1 - \mu_1)(X_1 - \mu_1)] & E[(X_1 - \mu_1)(X_2 - \mu_2)] & \cdots & E[(X_1 - \mu_1)(X_n - \mu_n)] \\ E[(X_2 - \mu_2)(X_1 - \mu_1)] & E[(X_2 - \mu_2)(X_2 - \mu_2)] & \cdots & E[(X_2 - \mu_2)(X_n - \mu_n)] \\ \vdots & \vdots & \ddots & \vdots \\ E[(X_n - \mu_n)(X_1 - \mu_1)] & E[(X_n - \mu_n)(X_2 - \mu_2)] & \cdots & E[(X_n - \mu_n)(X_n - \mu_n)] \end{bmatrix}$$

- > 变换得到Y, 是期望为零的随机向量。
- > Y的协方差矩阵 由 X的协方差矩阵决定:

$$\Phi_{Y} = T\Phi_{X}T^{T}$$

T的各行是 Φ x的特征向量, Φ y是一个对角阵,对角元素是 Φ x 的 特征值。

$$\Phi_{x} =
\begin{bmatrix}
\lambda_{1} & \cdots & \mathbf{0} \\
\vdots & \ddots & \vdots \\
\mathbf{0} & \cdots & \lambda_{N}
\end{bmatrix}$$

随机向量Y由互不相关的随机变量组成, 线性变换T起到消除变量间相关性的作用。

- 特征向量变换是可逆的,可以反变换。
- > 实现信号的K—L变换:
- 1) 要求矢量x的协方差矩阵Φx;
- 2) 求协方差矩阵Φx的特征值λi;
- 3) 求λ对应的Φx的特征向量;
- > 4) 用Φx的特征向量构成正交矩阵T。

例:若已知随机矢量x的协方差矩阵为

$$\Phi_{x} = \begin{bmatrix} 6 & 2 & 0 \\ 2 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

求其正交矩阵T?

1) 接 $\mathcal{M} - \Phi_x \models 0$,求 Φx 的特征值 λi :

$$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} - \begin{bmatrix} 6 & 2 & 0 \\ 2 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} = 0$$
 得:
$$\begin{bmatrix} \lambda - 6 & -2 & 0 \\ -2 & \lambda - 2 & 1 \\ 0 & 1 & \lambda - 1 \end{bmatrix} = 0$$
 可解得: $\lambda_1 = 6.854$ $\lambda_2 = 2$ $\lambda_3 = 0.146$

2) 求λi的特征向量, 将λ1, λ2, λ3代入中求得三个特征向量:

$$V_{1} = \begin{bmatrix} 0.918 \\ 0.392 \\ -0.067 \end{bmatrix} \qquad V_{2} = \begin{bmatrix} 0.333 \\ -0.667 \\ 0.667 \end{bmatrix} \qquad V_{3} = \begin{bmatrix} -0.217 \\ 0.\overline{6}34 \\ 0.742 \end{bmatrix}$$

用V1, V2, V3的转置向量作为正交矩阵T的行向量, 任一向量X=(2, 1, -0.1)的K-L变换为:

$$Y=TX=\begin{bmatrix} 0.918 & 0.329 & -0.067 \\ 0.333 & -0.667 & 0.667 \\ -0.217 & 0.634 & 0.742 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ -0.1 \end{bmatrix} = \begin{bmatrix} 2.234 \\ -0.067 \\ 0.127 \end{bmatrix}$$

则Y的协方差矩阵 Φy 为:

$$\boldsymbol{\phi}_{Y} = T\boldsymbol{\phi}_{X}T^{T} = \begin{bmatrix} 6.854 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0.146 \end{bmatrix}$$

主成分分析法概念

实际问题经常是多指标问题,多数情况下,不同指标间有一定的相关性,增加问题的复杂性。

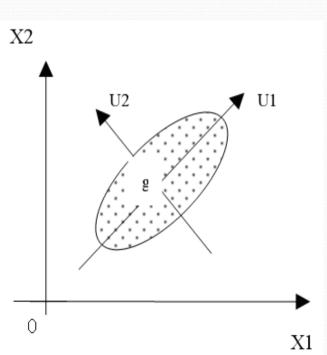
主成分分析法将原指标重新组合成一组 新的综合指标,从中取几个较少的指标,尽可能多地反映原信息量的模式识别统计方法。

主成分分析法概念

- 主成分分析是一种统计学方法,识别数据内在的特征。
- 主成分分析是数学过程,提取少量的主成分达到降低 维数的目的;
- 提取主成分能最大限度包含原数据的特征,最小限度 损失原数据的信息量。



主成分分析法的基本原理

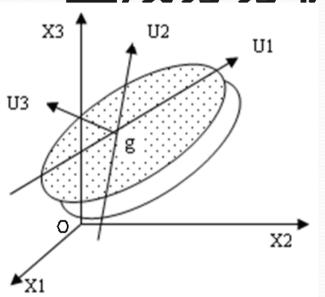


二维数据的重心在g。在方向u1 上,数据的差别最大,反映的信息也 最多,称为数据变异的最大方向。

将原点平移到g,做旋转变换,得到一正交坐标系。将数据点轴上投影,得到一个简化的一维数据系统。

降维处理的核心思想,是省却变 异不大的变量方向。

主成分分析法的基本原理



三维数据的分布是饼状球,在三个方向上差别很小。
 新的坐标系分析数据,原
 三维空间的数据可在二维
 平面得以显示。



主成分分析法设计步骤

- 原始数据标准化为Z
- 求Z的相关系数矩阵R 求R的特征向量和特征值 求计算主成分的贡献率

主成分分析的算法步骤及设

有n个样品,每个样品观测P个指标,将原始数据写成 矩阵:

$$\mathbf{X} = \begin{cases} \mathbf{x}_{11} & \mathbf{x}_{12} & \cdots & \mathbf{x}_{1p} \\ \mathbf{x}_{21} & \mathbf{x}_{22} & \cdots & \mathbf{x}_{2p} \\ \cdots & \cdots & \cdots \\ \mathbf{x}_{n1} & \mathbf{x}_{n2} & \cdots & \mathbf{x}_{np} \end{cases}_{n \times p}$$

天然气水合物测井数据包含2056个样本,9个属性。表中GAMMA是自然伽玛,AC是声波时差,LDEN、SDEN分别是长、短源距离密度,CAL是井径,RES是电阻率。

DEP	GAMMA	AC	LDEN	SDEN	CAL1	CAL2	CAL	RES1	RES2
39.3	26.16	646.59	1.82	1.80	319.04	246.83	263.74	0.24	0.40
39.4	25.50	646.79	1.82	1.81	333.91	246.97	282.94	0.45	0.49
39.5	25.57	647.41	1.83	1.80	339.49	246.06	290.44	0.78	0.58
39.6	26.20	648.04	1.83	1.80	339.44	245.73	292.77	0.94	0.61
39.7	26.43	648.40	1.83	1.78	340.43	246.70	292.59	0.85	0.53
39.8	26.75	648.42	1.84	1.78	342.10	248.15	293.57	0.54	0.75
39.9	26.72	648.81	1.84	1.78	343.26	249.27	295.12	0.42	1.05
40	26.56	649.34	1.83	1.78	344.20	249.73	296.27	0.41	1.42
40.1	27.21	649.25	1.80	1.77	346.52	251.59	296.96	0.47	1.49

原始数据初始化

每个主分量依赖于测量变量的尺度,尺度改变,得到不同的特征值。需对初始变量进行标准化处理,使其方差为1。标准化变换:

$$Z_{ij} = \frac{x_{ij} - \overline{x_j}}{s_j}, i = 1, 2, \dots, n; j = 1, 2, \dots, p$$
其中,
$$\frac{\sum_{i=1}^{n} x_{ij}}{n}, s_j^2 = \frac{\sum_{i=1}^{n} (x_{ij} - \overline{x_j})^2}{n - 1}$$

天然气水合物测井数据标准

化结果

DEP	GAMM A	AC	LDEN	SDEN	CAL1	CAL2	CAL	RES1	RES2
39.3	-2.992	2.123	-0.982	-1.732	0.678	-1.121	-0.571	-4.063	-3.834
39.4	-3.135	2.127	-1.978	-1.667	1.028	-1.115	0.109	-3.624	-3.603
39.5	-3.119	2.143	-1.913	-1.739	1.159	-1.159	0.375	-2.924	-3.385
39.6	-2.985	2.158	-1.874	-1.831	1.158	-1.176	0.458	-2.599	-3.308
39.7	-2.934	2.167	-1.856	-1.998	1.181	-1.128	0.451	-2.793	-3.503
39.8	-2.866	2.168	-1.784	-1.990	1.221	-1.056	0.48	-3.427	-2.949
39.9	-2.873	2.177	1.780	-1.970	1.248	-1.000	0.541	-3.689	-2.194
40	-2.906	2.191	-1.964	-1.980	1.270	-0.978	0.582	-3.717	-1.27

对标准化阵求相关系数矩阵R

$$R = \begin{cases} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ r_{p1} & r_{p2} & \cdots & r_{pp} \end{cases} = \frac{Z^T Z}{n-1}$$

$$\begin{cases} R \\ R \end{cases}$$

```
for(i = 0; i < p; i ++) \\ for(j = 0; j < p; j ++) \\ \{ \\ R[i][j] = 0; \\ for(m = 0; m < n; m ++) \\ R[i][j] += Z[m][i] * Z[m][j]; \\ \}
```

其中,

$$r_{ij} = \frac{\sum z_{mi} \cdot z_{mj}}{n-1}, i, j = 1, 2, \dots, p$$

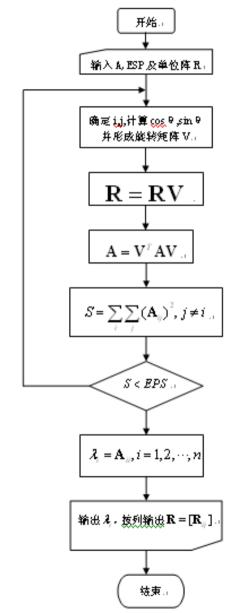
```
for(i = 0; i < p; i ++)
for(j = 0; j < p; j++)
R[i][j]=R[i][j]/(n-1);
```

天然气水合物测井属性间的相关系数

属性	GAM	AC	LDEN	SDEN	CAL1	CAL2	CAL	RES1	RES2
GAM	1.000	-0.318	0.292	0.252	-0.613	-0.535	-0.648	0.400	0.391
AC	-0.318	1.000	-0.398	-0.340	0.475	0.134	0.395	-0.833	-0.828
LDEN	0.292	-0.398	1.000	0.952	-0.162	-0.122	-0.150	0.258	0.222
SDEN	0.252	-0.340	0.952	1.000	-0.142	-0.077	-0.118	0.213	0.172
CAL1	-0.613	0.475	-0.162	-0.142	1.000	0.563	0.947	-0.607	-0.583
CAL2	-0.535	0.134	-0.122	-0.077	0.563	1.000	0.760	-0.195	-0.215
CAL	-0.648	0.395	-0.150	-0.118	0.947	0.760	1.000	-0.523	-0.513
RES1	0.400	-0.833	0.258	0.213	-0.607	-0.195	-0.523	1.000	0.985
RES2	0.391	-0.828	0.222	0.172	-0.583	-0.215	-0.513	0.985	1.000

解样本矩阵的特征向量、特征值

雅可比方法解相关矩阵R 的特征值和特征向量。



雅克比方法的理论依据

·若V是正交矩阵,则正交变换

$$\mathbf{B} = \mathbf{V}^{-1} \mathbf{A} \mathbf{V} = \mathbf{V}^T \mathbf{A} \mathbf{V}$$

● B与A有相同的特征值;若A对称,则存在一正交矩阵V使

• 对角元是A的特征值, V的列向量是特征向量

雅可比算法

• 对称矩阵通过正交变换对角化,得到特征值及特征向量。

设**A**为二阶对称矩阵,
$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$
 $a_{12} = a_{21}$

由线性代数知道, A阵对应于二次型

$$f(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

其系数是A阵的元素,二次型是一条二次曲线:

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = C$$

雅可比算法推导

• 其矩阵形式是

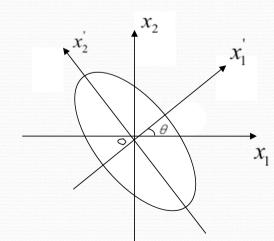
$$\{x_1 \quad x_2\} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} = C$$

• 坐标轴如图旋转,得到平面上的标准二次曲线

$$b_{11}x_1^{'2} + b_{22}x_2^{'2} = C$$

其矩阵形式为

$$\left\{ x_1 \quad x_2 \right\} \begin{bmatrix} b_{11} & 0 \\ 0 & b_{22} \end{bmatrix} \left\{ x_1 \\ x_2 \right\} = C$$



• 新旧坐标的变换关系 $\begin{cases} x_1 \\ x_2 \end{cases} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{cases} x_1 \\ x_2 \end{cases}$

• 轴旋转角度, 其变换矩阵V是正交矩阵, 上式是正交变换

$$\mathbf{V} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

求出特征值,从大到小排序

方差贡献率=
$$\frac{\lambda_i}{\sum\limits_{k=1}^p \lambda_k}$$
 $(i=1,2,\wedge,p)$
累计方差贡献率= $\frac{\sum\limits_{k=1}^i \lambda_i}{\sum\limits_{k=1}^p \lambda_k}$ $(i=1,2,\wedge,p)$

• 累计方差贡献率85%~95%的特征值,对应的特征向量是第一,第二, ..., 第m个主成分。

R的特征恒及贞献率、累计贡献率 献率

主成分	特征值	贡献率	累计贡献率
1	4.571978	%50.7998	%50.7998
2	1.872737	%20.8082	%71.6079
3	1.496850	%16.6317	%88.2396
4	0.448587	%4.9843	%93.2239
5	0.355224	%3.9469	%97.1708
6	0.184553	%2.0506	%99.2214
7	0.043096	%0.4788	%99.7003
8	0.015257	%0.1695	%99.8698
9	0.011719	%0.1302	%100

结果分析

$$F_1 = 0.3278X_1 - 0.3586X_2 + 0.2136X_3 + 0.1909X_4 - 0.3959X_5 -0.2648X_6 - 0.3888X_7 + 0.3930X_8 + 0.3866X_9$$

第一主成分中,指标的系数相近,第一主成分视为九指标共同反应 测井综合指标。

$$\begin{split} F_2 &= -0.1262X_1 - 0.2115X_2 + 0.5712X_3 + 0.5808X_4 + 0.2526X_5 \\ &+ 0.3263X_6 + 0.3154X_7 + 0.0644X_8 + 0.0424X_9 \end{split}$$

第二主成分中,第三、四项指标的影响大,视为距离密度的影响

$$F_3 = -0.2909X_1 - 0.3580X_2 - 0.3219X_3 - 0.3367X_4 + 0.0861X_5 + 0.4171X_6 + 0.2056X_7 + 0.4088X_8 + 0.4236X_9$$

第三主成分,第六、八、九项指标的影响大,反应CAL2和两个电阻率的综合指标。

离散余弦变换 (DCT)

 离散余弦变换(Discrete Cosine Transform, DCT), 变 换核为余弦函数。

一维DCT变换核:
$$g(x,u) = C(u)\sqrt{\frac{2}{N}}\cos\frac{(2x+1)u\pi}{2N}$$

式中, x, u=0, 1, 2, ..., N-1;

$$C(u) = \begin{cases} \frac{1}{\sqrt{2}} & u = 0\\ 1 & 其他 \end{cases}$$

- DCT具正交变换性质 阵向量描述语音和图 的相关特征。DCT实 音、图像最佳变换。
- 视频压缩编码的国际 ,把DCT作为一个基 模块。

一维DCT定义: $\{f(x)|x=0, 1, ..., N-1\}$, 为离散信号列

$$F(u) = C(u)\sqrt{\frac{2}{N}} \sum_{x=0}^{N-1} f(x) \cos \frac{(2x+1)u\pi}{2N}$$

式中,
$$u, x = 0, 1, 2, ..., N-1$$
。

- DCT具正交变换性质,变换阵向量描述语音和图像信号的相关特征。DCT变换是语音、图像最佳变换
- 视频压缩编码国际标准,把DCT作为一个基本处理模块。

写成矩阵的形式F = Gf

$$\cos \pi (i*(2j-1)/2N$$

$$G = \begin{bmatrix} 1/\sqrt{N} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \sqrt{2/N} \begin{bmatrix} \cos(\pi/2N) & \cos(3\pi/2N) & \cdots & \cos((2N-1)\pi/2N) \end{bmatrix} \\ \sqrt{2/N} \begin{bmatrix} \cos(2\pi/2N) & \cos(6\pi/2N) & \cdots & \cos(2(2N-1)\pi/2N) \end{bmatrix} \\ \vdots \\ \sqrt{2/N} \begin{bmatrix} \cos((N-1)\pi/2N) & \cos((N-1)(3\pi/2N) & \cdots & \cos((N-1)(2N-1)\pi/2N) \end{bmatrix} \end{bmatrix}$$

- 一维DCT 逆变换IDCT
- 一维DCT逆变换核与正变换核相同

$$f(x) = \sqrt{\frac{2}{N}} \sum_{u=0}^{N-1} C(u) F(u) \cos \frac{(2x+1)u\pi}{2N}$$

二维离散余弦变换

• 正变换核为

$$g(x, y, u, v) = \frac{2}{\sqrt{MN}} C(u)C(v) \cos \frac{(2x+1)u\pi}{2M} \cos \frac{(2y+1)v\pi}{2N}$$

f(x, y): M×N数字图像矩阵

$$F(u,v) = \frac{2}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)C(u)C(v) \cos \frac{(2x+1)u\pi}{2M} \cos \frac{(2y+1)v\pi}{2N}$$

- 二维DCT矩阵形式:
- F=GfG^T
- 二维DCT逆变换核与正变换核相同,行列可分离,即

$$g(x, y, u, v) = g_1(x, u)g_2(y, v)$$

$$= \frac{2}{\sqrt{M}}C(u)\cos\frac{(2x+1)u\pi}{2M} \cdot \frac{2}{\sqrt{N}}C(v)\cos\frac{(2y+1)v\pi}{2N}$$

• 二维DCT可由两次一维DCT完成, 算法流程与DFT类似

$$f(x,y) \rightarrow F_{f\bar{j}}[f(x,y)] = F(x,v)$$

转置
 $\rightarrow F(x,v)^T \rightarrow F_{\bar{j}|}[F(x,v)^T] = F(u,v)^T$
转置
 $\rightarrow F(u,v)$

- 快速离散余弦变换
- 离散余弦变换计算量相当大,已有多种快速DCT

$$f_e(x) = \begin{cases} f(x) & x=0, 1, 2, ..., N-1, \cos(2N\pi)=1\\ 0 & x=N, N+1, ..., 2N-1, \cos((N+0.5)\pi)=0 \end{cases}$$

$$F(0) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} f_e(x)$$

长度为N的f(x), 延拓为2N点的序列 $f_e(x)$,对 $f_e(x)$ 作 傅里叶变换,取的实部得到 DCT结果。 \square 对于离散余弦逆变换IDCT,首先将F(u)延拓为

$$F_{e}(u) = \begin{cases} F(u) & u=0, 1, 2, ..., N-1 \\ 0 & u=N, N+1, ..., 2N-1 \end{cases}$$
(7-62)

- > 离散余弦变换(DCT)
- > 数字图像压缩编码,最佳变换K-L计算复杂,一般不采用。
- DCT计算复杂度适中,具有可分离、快速算法等特点,图像数据压缩采用离散余弦变换编码很多。
- ➤ JPEG、MPEG、H.261等压缩标准,都用到离散余弦变 换编码进行数据压缩。

- > 余弦变换是傅立叶变换的一种特殊情况。
- > 傅立叶级数展开式中,被展开的函数是实偶函数;
- > 傅立叶级数中只包含余弦项,其离散化可导出余弦变换;
- ▶ 称为离散余弦变换DCT (Discrete Cosine Transform)。

二维离散偶余弦正变换公式为:

$$C(u,v) = E(u)E(v)\frac{2}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)\cdot\cos\left(\frac{2x+1}{2N}u\pi\right)\cdot\cos\left(\frac{2y+1}{2N}v\pi\right)$$
(4-38)

式中, x,y,u,v=0, 1....., N-1。

$$E(u), E(v) = 1/\sqrt{2}$$
 当 $u=v=0$ 时。
 $E(u), E(v) = 1$ 当 $u=1, 2, ..., N-1; v=1, 2, ..., N-1$ 时。

> 二维离散偶余弦逆变换公式:

$$f(x,y) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} E(u)E(v)C(u,v) \cdot \cos\left(\frac{2x+1}{2N}u\pi\right) \cdot \cos\left(\frac{2y+1}{2N}v\pi\right)$$

式中 x, y, u, v = 0,1,...,N-1

当
$$u=v=0$$
时, $E(u), E(v) = 1/\sqrt{2}$

$$\exists u=1, 2,..., N-1; v=1, 2,..., N-1$$
 $E(u), E(v) = 1$

- >二维离散余弦变换核具有可分离特性;
- ▶正、逆变换均可将二维变换分解成一维变换(行、列)进行
- ▶DCT变换编码,不直接变换整个图像,图像分块,M×N图像分成(M/K)×(N/K)子块。

> 优点:

- 1) 降低运算量,一幅512×512图像,分块变换仅需约1/3 运算量;
- 2) 量化和扫描处理简化;
- 3) 传输错误控制在一个块内,不是整个图像扩散。分块通常8×8和16×16。

4.7 JPEG图像编码压缩标准

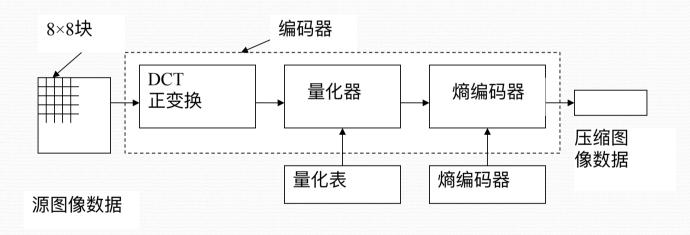
- > JPEG(Joint Photographic Expert Group, 简称JPEG)是联合图像专家小组。
- ▶JPEG算法被确定为JPEG国际标准,国际上彩色、灰度图像 第一个国际标准。
- ▶JPEG标准适于图像压缩,电视图像序列的帧压缩编码也采用JPEG压缩标准。

4.7.1 JPEG的工作模式

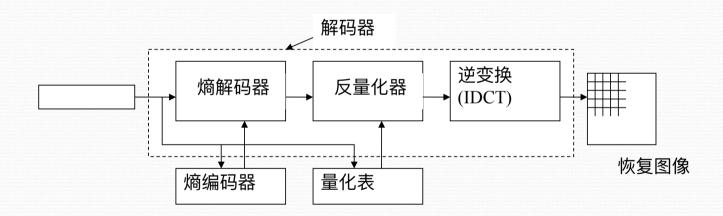
- ▶JPEG对图像分量(R,G,B) 单独编码,采用不同量化和熵 编码方法;
- ➤ 图像分量,JPEG提供4种工作模式。
- ?顺序编码:各分量左上至右下扫描,一次扫描完成编码。
- 冥累进编码: 图像编码多次扫描中完成。
- ?分层编码:图像在多个空间分辨率进行编码。

4.7.2 基本工作模式

基于DCT的 JPEG编码过程框图



解码过程框图



- > JPEG采用8×8子块的二维离散余弦变换(DCT)。
- 原图像采样精度P位,无符号整数。离散余弦正变换输入, 把[0,2P]无符号整数变成[-2P-1,2P-1-1]有符号整数。
- 解码器输出端经(IDCT)后,得到多个8×8图像数据块,数值范围由[-2P-12P-1-1]变到[0,2P]的无符号整数,获得重构图像。

- DCT系数量化处理实现压缩,量化处理是多到一的映射, DCT编解码信息损失。
- ➤ JPEG中采用线性均匀量化器,64个DCT系数除以量化步长,四舍五入取整。
- 量化在主观保真度图像质量前提,丢掉视觉效果影响不大的信息。

例: 给定Lena图像的一个平坦区域(8×8子块)

69	71	75	79	84	89	91
69	70	73	76	83	90	95
77	74	76	74	85	89	95
71	73	76	79	86	91	93
74	77	77	82	88	91	93
78	76	80	84	88	92	95
76	78	80	85	93	94	95
74	79	81	85	86	94	94

给出DCT变换系数量化过程

用JPEG亮度量化矩阵式对每个系数进行均匀量化,量化器输出为:

> 反量化后,进行DCT反变换,解码图像为:

 80
 75
 71
 72
 78
 85
 89
 90

 80
 75
 71
 72
 78
 85
 89
 90

 80
 76
 72
 73
 79
 86
 90
 91

 81
 77
 72
 74
 80
 87
 91
 92

 82
 77
 73
 74
 81
 87
 91
 93

 83
 78
 74
 75
 81
 88
 92
 93

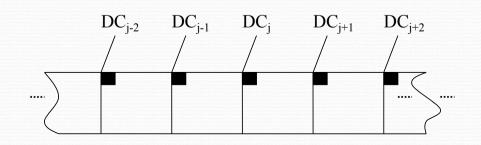
 83
 79
 75
 76
 82
 89
 93
 94

 84
 79
 75
 76
 82
 89
 93
 94

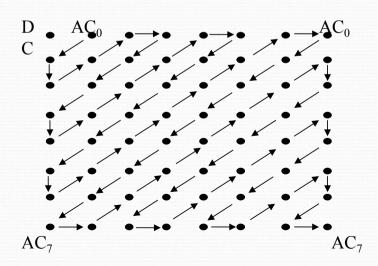
> DCT变换系数,能量集中在少数低频系数:

```
660 1250
         -470496259980
                         10.3993 7.8750 8.4866 5.6025 1.3176
-17.3267 -2.6749
                  5.2236 - 1.3234
                                 0.5222 0.2914 0.2800 -2.281
0.0280
         -0.6463 -0.9545 0.9620
                                 2 4730 1 9783 -0 316
                                                      2.1741
2.3003
         0.4542
                 -2.2403
                        3.5559 1.2907 -1.0024 0.1580 0.9747
-2.3750
         0.1038
                 -3.2220
                        0.9653 1.3750 2.2258 0.3875 3.5236
0 9294
        -13282
                 -2425609828-19317-0697201253-1856
0.3943
        2.6640
                -0.5669 -3.4168 -0.8891 -1.6182 -2.545 -1.732
2.1666
        1.7238 -0.3335 -0.4808 -2.6253 -0.9699 1.4854 -1.183
```

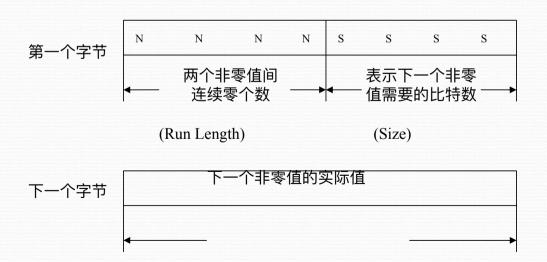
- > 8×8子块的64个变换系数量化后,分成两类:一个直流系数DC和63个交流系数AC。
- ➤ 坐标u=v=0直流系数DC,为空域图像64个像素平均值。
- ▶ 相邻8×8子块间的DC系数强相关,JPEG对DC系数采用 DPCM编码,即相邻块间的DC系数差值DIFF=DC_i-DC_{i-1} 编码。



▶ 其余63个系数,交流系数AC采用行程编码。低频分量多 呈圆环形辐射状向高频率衰减,按Z字形衰减。AC系数按 Z字形扫描读数。



- > 63AC系数用行程编码,采用两个字节。
- > JPEG第一字节的高4位表示连续"0"的个数,低四位表示 后面非"0"系数的位数。
- ➤ 下个字节是AC系数量化值。AC系数行程编码如下:



- > 为进一步压缩数据,对DPCM编码后的DC码和RLE编码 后的AC码的码字再作熵编码。
- ➤ JPEG建议使用两种熵编码方法: 哈夫曼(Huffman)编码和自适应二进制算术编码。

例: Lena测试图像(256×256)的一个8×8块,前一块量化DC系数为-10,此块取值如下:

107 105 104 114 100 112 111 108 104 99 107 108 112 115 117 115 104 101 108 110 109 114 117 114 105 105 105 106 110 109 96 113 102 107 102 113 105 104 107 115 107 106 102 103 106 115 106 121 114 107 87 98 110 102 116 120 114 99 98 95 93 111 115 112 说明JPEG编解码过程。

解:每个像素减128,输入图像取值范围为-2⁷~ 2⁷-1。DCT变换,进行量化,量化器输出为:

- 前块的量化DC系数为-10,该8×8块的DC系数为-1,DIFF=-10-1,SIZE=4,AMP=-11,编码为101,0100。Z字扫描为-2,1,0,0,1,0,-1,0,0,0,0,1,EOB。[RUN,SIZE][AMP]串为[0,2][-2],[0,1][1],[2,1][1],[2,1][1],[1,1][-1],[4,1][1],[EOB]。对[RUN,SIZE]查Huffman码表,对AMP直接编码,得到码字为[01][01],[00][1],[11100][1],[1100][0],[111011][1],[1010]。
- DC编码需7位,AC编码需29位,共需36位。原8×8块共 64个像素,每个像素8位,因此,压缩比为: 64×8/36=14.2

> 用解码器解码后,此块重构图像为:

```
      108
      108
      107
      107
      109
      111
      114
      116

      106
      107
      107
      109
      110
      112
      113
      114

      104
      105
      108
      110
      112
      112
      112
      111

      102
      104
      107
      110
      112
      112
      111
      110

      103
      104
      106
      108
      109
      110
      110
      110

      105
      104
      104
      106
      108
      111
      113

      108
      105
      102
      100
      101
      106
      112
      116

      110
      106
      100
      97
      98
      105
      113
      118
```

4.7.3 JPEG文件格式

- ▶制定JPEG标准时,已定义标记区分和识别图像数据及相关信息。JFIF (JPEG文件交换格式——JPEG File Interchange Format) 1.02版。
- > JPEG文件中的字节格式按照正序排列,高位字节在前,低位字节在后。
- > JPEG文件上分成:标记码(tag)和压缩数据。
- ➤ 标记码给出JPEG图像的宽、高、Huffman码表、量化表等。 绝大多数的JPEG文件只包含的几种标记码。

- ▶ 标记码由两个字节组成,高字节是固定值0xFF。
- ▶ 每个标记前还可添加数目不限的0xFF填充字节。
- > JPEG文件组成:
- ? SOI标记:图像开始(Start of Image)标记。
- APPO标记: JPEG保留给应用程序(Application)的标记码, 而JFIF将文件的相关信息定义在此标记中。
- ? APPn标记,其中n=1~15(任选): 代表其他应用数据块,它的结构包括两部分: APPn长度(Length)和应用详细信息 (application specific information)
- ? 一个或多个量化表DQT(define quantization table)
- ? 一个或多个Huffman表DHT(Define Huffman Table)

4.8 MPEG视频编码压缩标准

- ▶从时间的观点看,数字图像分为静态图像和运动图像,视频信号就是典型的运动图像。
- ▶视频压缩的目标是在尽可能保证视觉效果的前提下减少视频数据率。
- ▶根据压缩前和解压缩后的数据是否完全一致,视频压缩可 分为有损压缩和无损压缩。
- 无损压缩意味着解压缩后的数据与压缩前的数据完全一致。 有损压缩则意味着解压缩后的数据与压缩前的数据不一致。

>视频编解码过程: 外部控制 编码控制 视频 信号 信源编码 多路视频编 发送缓冲 码器 区 a) 视频编码 编码比 特流 信源解码 多路视频解 接收缓冲 器 码器 X b) 视频解码

- > 视频信号的压缩包括两个主要方面: 帧内压缩与帧间压缩。
- > 帧内(Intra frame)压缩也称为空间压缩(Spatial compression)。当压缩一帧图像时,仅考虑本帧的数据而不考虑相邻帧之间的冗余信息。
- ➤ 帧间(Inter frame)压缩是基于许多视频或动画的连续前后两帧具有很大的相关性。即连续的视频其相邻帧之间具有冗余信息。根据这一特性,压缩相邻帧之间的冗余量就可以进一步提高压缩量,减小压缩比。帧间压缩也称为时间压缩,它通过比较时间轴上不同帧之间的数据进行压缩。

- ➤ MPEG (Moving Picture Expert Group)是运动图像专家组的简称。该小组于1991年底提出了用于数字存储媒介的、速率约1.5MB / s的运动图像及其伴音的压缩编码,并于1992年正式通过,通常被称为MPEG标准,此标准后来被定名为MPEG-1。
- ▶到目前为止,MPEG标准己不再是一个单一的标准,而是一个用于全运动视频和相关音频压缩的标准系列,包括MPEG-1、MPEG-2、MPEG-3、MPEG-4和MPEG-7共5个标准,每一个标准都有其特定的应用范围。其中,MPEG-1和MPEG-2标准的应用范围最广。

- >MPEG-1用于加速CD-ROM中图像的传输
- ➤ MPEG-2用于宽带传输的图像,图像质量达到电视广播甚至HDTV的标准。和MPEG-1相比,MPEG-2支持更广的分辨率和比特率范围,将成为数字图像盘(DVD)和数字广播电视的压缩方式
- ➤MPEG-4标准支持非常低的比特率的数据流的应用,如电视电话,视频邮件和电子报刊等

- > MPEG视频压缩分为空间域压缩与时间域压缩。
- ➤ MPEG标准在空间域的压缩,每一帧被作为独立的图像获取,且压缩步骤与JPEG标准的步骤一样。时间域压缩,即帧间编码的基本思想是仅存储运动图像从一帧到下一帧的变化部分,而不是存储全部图像数据,这样做能极大地减少运动图像数据的存储量。
- ➤ 这是通过把帧序列划分成I帧、P帧、B帧,使用参照帧及运动补偿技术来实现的。

4.9 小结

- ▶本章介绍图像编码与压缩的基本概念、理论及其编码分类。 无损压缩和有损压缩技术。
- ➤ 无损压缩是指可以精确无误地从压缩数据中恢复出原始数据的图像压缩方法。常见的无损压缩技术包括基于统计概率的方法和基于字典的技术。
- ▶有损编码是以丢失部分信息为代价来换取高压缩比的。有 损压缩方法主要**有损预测编码方法、变换编码方法**等。

- ➤ 图像压缩的国际标准可分成三个部分:静止图像压缩标准、 运动图像压缩标准和二值图像压缩标准。常用的编码标准 有: JPEG、MPEG、JBIG及H.26x等标准。
- ➤ JPEG是联合图像专家小组开发研制的连续色调、多级灰度、 静止图像的数字图像压缩编码方法。JPEG中的核心算法是 DCT变换编码。
- > MPEG视频压缩分为空间域压缩与时间域压缩。