

图像压缩技术

数字图像处理 第6讲

内容

- 图像压缩概述
- 变长编码
- 预测编码
- 变换编码
- 图像压缩国际标准

图像压缩

- 动机

- 图像的数据量大，占用较多的存储资源
- 图像的传输，需要占用较大的网络带宽
 - 800万像素照片占用 23MB

Color space: RGB

Dimensions: 1200×1600

Kind: JPEG

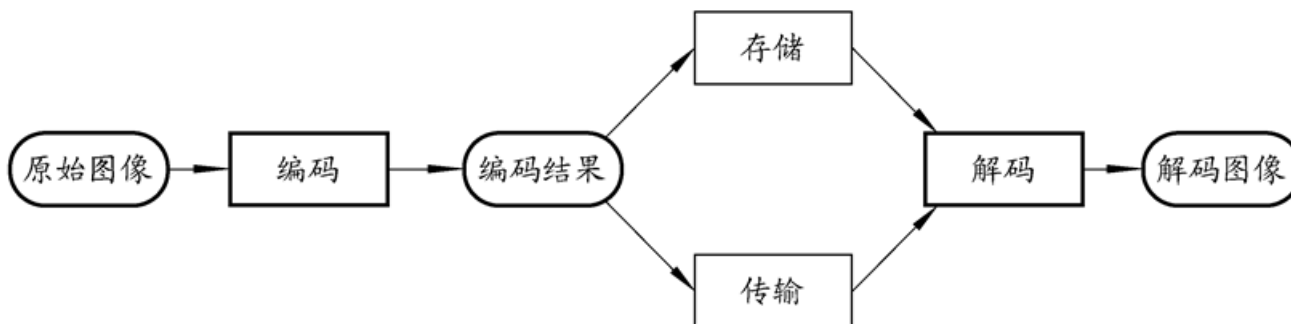
Size: 181,088 bytes

- 图像压缩

- 通过对原始图像进行编码以减少数据量

Raw data: $1200 \times 1600 \times 3 = 5,760,000$ bytes

Compression ratio: $\frac{181,088}{5,760,000} = 3.14\%$



图像压缩率

- 图像原始存储方法下的平均码长为 m
 - 每个像素值占用的二进制位数
 - 灰度图像通常 8 bit
 - 彩色图像通常 $8 \times 3 = 24$ bit
- 图像经过编码后的平均码长为 n
 - 该压缩方法的压缩率
 - 通常大于1

$$C_R = \frac{m}{n}$$

图像压缩类型

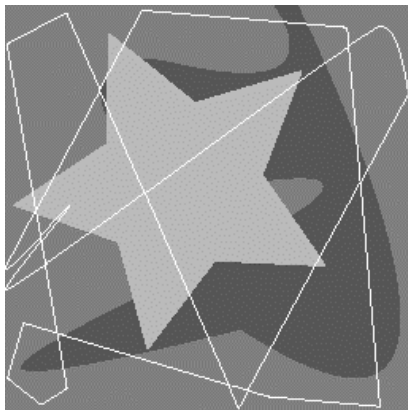
- 信息保存型
 - 在压缩和解压缩过程中没有信息损失
 - 无损压缩
 - 压缩率一般在2 ~ 10之间
- 信息损失型
 - 图像经过压缩后并不能经解压缩恢复原状
 - 有损压缩
 - 常能取得较高的压缩率（几十以上）

图像压缩的可行性

- 数据和信息：数据是信息的载体
 - 对给定量的信息可用不同的数据量来表示
- 图像的原始存储方式存在冗余
 - 冗余：多余的重复或啰嗦内容（信息、语言等）
- 冗余类型
 - 编码冗余
 - 像素相关冗余
 - 心理视觉冗余

编码冗余

- 图像编码是表示图像信息的符号系统
 - 对每个信息赋予的符号长度为码长
- 平均码长越短，编码效率越高
- 编码冗余指选择的编码方法效率不高



r_k	$p_r(r_k)$	Code 1	$l_1(r_k)$	Code 2	$l_2(r_k)$
$r_{87} = 87$	0.25	01010111	8	01	2
$r_{128} = 128$	0.47	10000000	8	1	1
$r_{186} = 186$	0.25	11000100	8	000	3
$r_{255} = 255$	0.03	11111111	8	001	3
r_k for $k \neq 87, 128, 186, 255$	0	—	8	—	0

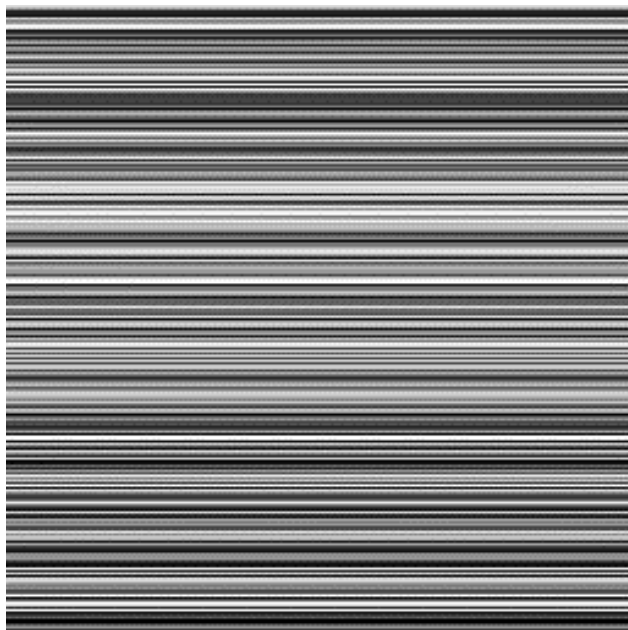
Code1平均码长 8bit/pixel

Code2平均码长: 1.81 bit/pixel

$$L_{avg} = 0.25(2) + 0.47(1) + 0.25(3) + 0.03(3) \approx \mathbf{1.81} \text{ bit/pixel}$$

像素相关冗余

- 像素之间是存在相关性的
- 无视这种相关性带来数据的重复



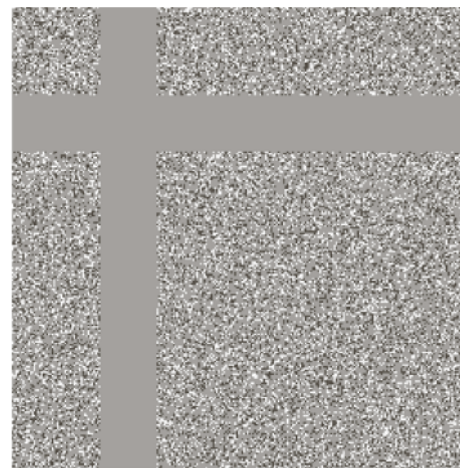
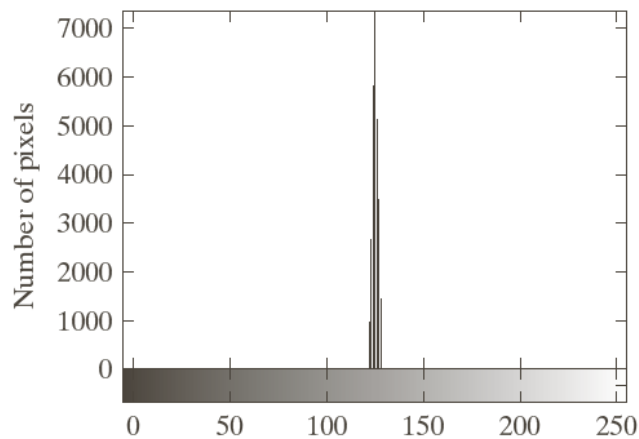
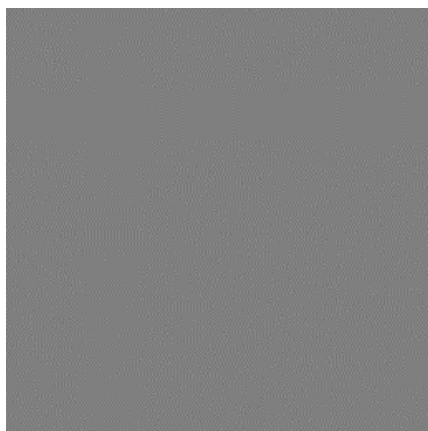
保存每种像素值出现的
(行号, 起点, 终点) 就可以

00000000: (75, 1, 240)
00000001: (36, 1, 240),(49, 1, 240)
00000010: (81, 1, 240)

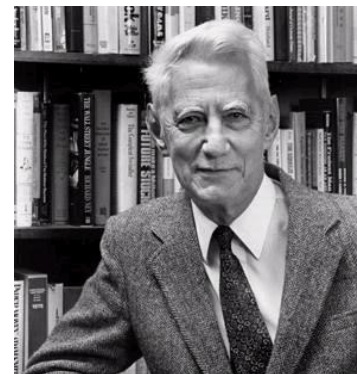
...

心理视觉冗余

- 人对有些视觉信息并不敏感
- 在不削弱图像视觉感知的情况下可以去除
- 通常是有损压缩



图像信息的衡量 - 信息论



- Claude Shannon 香农
 - 1948年《通讯的数学原理》
 - 信息论奠基人
- 发生概率为 $P(E)$ 的随机事件 E 的信息量

$$I(E) = \log_2 \frac{1}{P(E)} = -\log_2 P(E) \quad \text{单位: bit}$$

- 事情发生越频繁, 信息量越少
 - $P(E)=1$ 信息量为0, 比如“太阳从东边升起来”

熵与编码极限

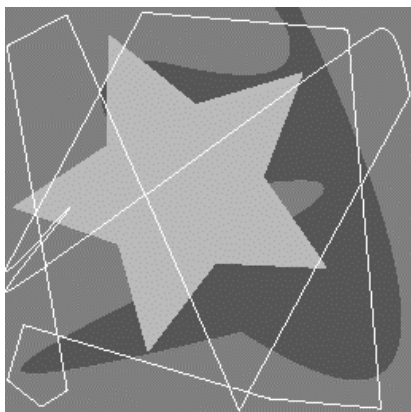
- 信源：产生消息的源头，如交通信号灯
- 某信源有 $\{a_1, a_2, \dots, a_J\}$ 种事件发生，其发生概率 $\{P(a_1), P(a_2), \dots, P(a_J)\}$ 信源平均信息量称为熵

$$H = - \sum_{j=1}^J P(a_j) \log_2 P(a_j)$$

- 香农第1定律
 - 熵给定了无损描述一个信源所需平均码长的最小值
 - 抛硬币（正 0.5，反 0.5） $H = 1$
 - 掷色子（各是1/6 概率） $H = 2.58$
 - 熵反映了信源的不确定性

图像的熵与压缩极限

- 某幅灰度图像可看作由一个灰度信源生成
- 可通过各灰度值的出现概率计算该信源的熵
- 熵决定了描述该信源的所需的最小平均码长
- 则该图像可被该码长完全编码



$$H = -\sum_{k=0}^{L-1} p(r_k) \log_2 p(r_k)$$

r_k	$p_r(r_k)$	Code 1	$l_1(r_k)$	Code 2	$l_2(r_k)$
$r_{87} = 87$	0.25	01010111	8	01	2
$r_{128} = 128$	0.47	10000000	8	1	1
$r_{186} = 186$	0.25	11000100	8	000	3
$r_{255} = 255$	0.03	11111111	8	001	3
r_k for $k \neq 87, 128, 186, 255$	0	—	8	—	0

$$H = -0.25 \log_2 0.25 + 0.47 \log_2 0.47 + 0.25 \log_2 0.25 + 0.03 \log_2 0.03 \\ \approx -0.25(-2) + 0.47(-1.09) + 0.03(-5.06) \approx \mathbf{1.66} \text{ bit/pixel}$$

$$L_{avg} = 0.25(2) + 0.47(1) + 0.25(3) + 0.03(3) \approx \mathbf{1.81} \text{ bit/pixel}$$

保真度准则

- 无损压缩可以完全保真
- 有损压缩存储损失，需要衡量损失程度
- 客观保真度准则
 - 用编码输入图与解码输出图的差值函数来计算
 - 总误差

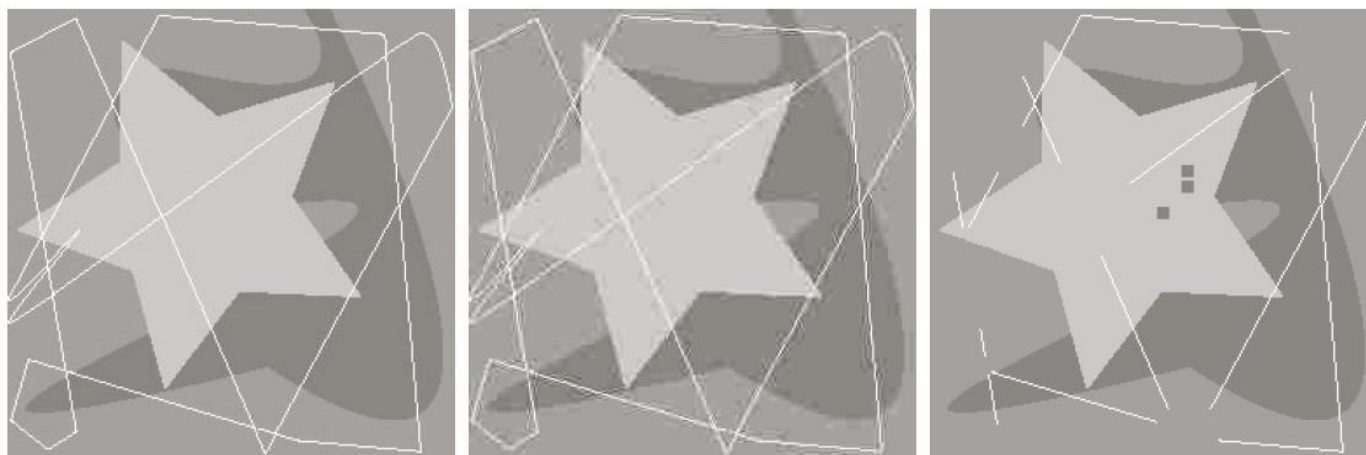
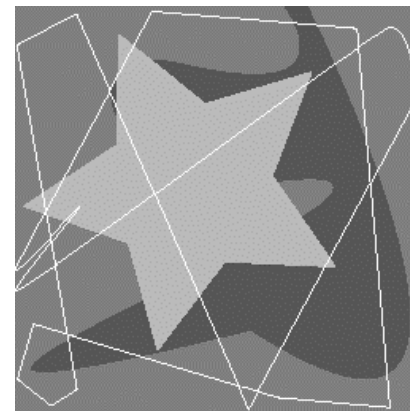
$$\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} |\hat{f}(x, y) - f(x, y)|$$

- 均方根误差

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

保真度准则

- 主观保真度准则
 - 使用人的主观评估来衡量图像的质量



三个图的均方根误差： 5.1 15.6 14.2

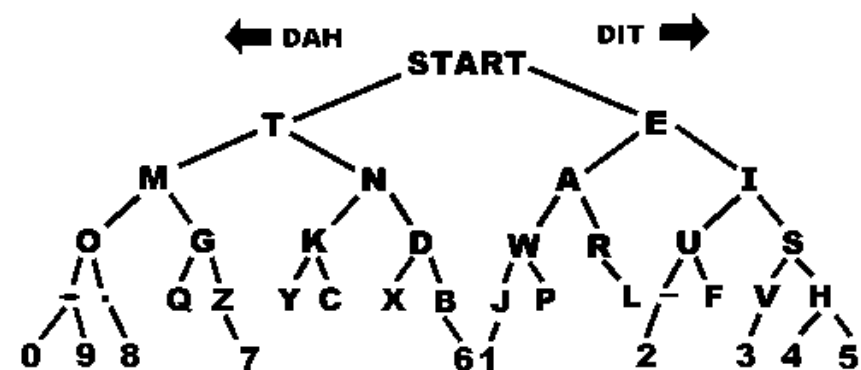
内容

- 图像压缩概述
- 变长编码
 - 哈夫曼编码
- 预测编码
- 变换编码
- 图像压缩国际标准

莫尔斯码

- 美国人 摩尔斯 1837 有线电报发明者

A	• —	U	• • —
B	— • • •	V	• • • —
C	— • — •	W	— • —
D	— • • •	X	— • • —
E	•	Y	— • — —
F	• • — •	Z	— — • •
G	— — • •		
H	• • • •		
I	• •		
J	• — — —		
K	— • — —	1	• — — — —
L	— • • •	2	• • — — —
M	— — —	3	• • • — —
N	— •	4	• • • • —
O	— — — —	5	• • • • •
P	• — — • •	6	— • • • •
Q	— • — • •	7	— — • • •
R	• — • • •	8	— — — • •
S	• • • •	9	— — — — •
T	—	0	— — — — —



变长编码

- 属于统计编码、无损压缩
- 利用像素值的概率分布
 - 出现概率大的像素值用短码长，概率小的用长码长
 - 减少编码冗余
- 也称为熵编码
 - 编码后图像的平均码长接近图像的熵
- 常见方法
 - 哈夫曼编码
 - 算术编码

哈夫曼 (Huffman) 编码

- 编码分两步
- 消减信源
 - 按概率排序
 - 合并概率最低的两个
 - 持续到只剩两个

初始信源		信源的消减步骤	
符号	概率	1	2
b_2	0.38	0.38	0.62
b_4	0.30	0.32	0.38
b_3	0.22	0.30	
b_1	0.10		

- 符号赋值
 - 从后往前递推
 - 每次后面补0/1

初始信源			对消减信源的赋值	
符号	概率	码字	1	2
b_2	0.38	1	0.38 1	0.62 0
b_4	0.30	01	0.32 00	0.38 1
b_3	0.22	000	0.30 01	
b_1	0.10	001		

信源的熵1.86 哈夫曼平均码长1.94

哈夫曼 (Huffman) 编码

- 解码

- 直接查表
- 如01100001 对应 b_4 b_2 b_3 b_4

初始信源	
符号	码字
b_2	1
b_4	01
b_3	000
b_1	001

- 哈夫曼编码用于图像时就是建立灰度值映射表
- 编码特点
 - 即时性, 解码时无需参考后续符号
 - 唯一性, 任何符号串只能以一种方式解码

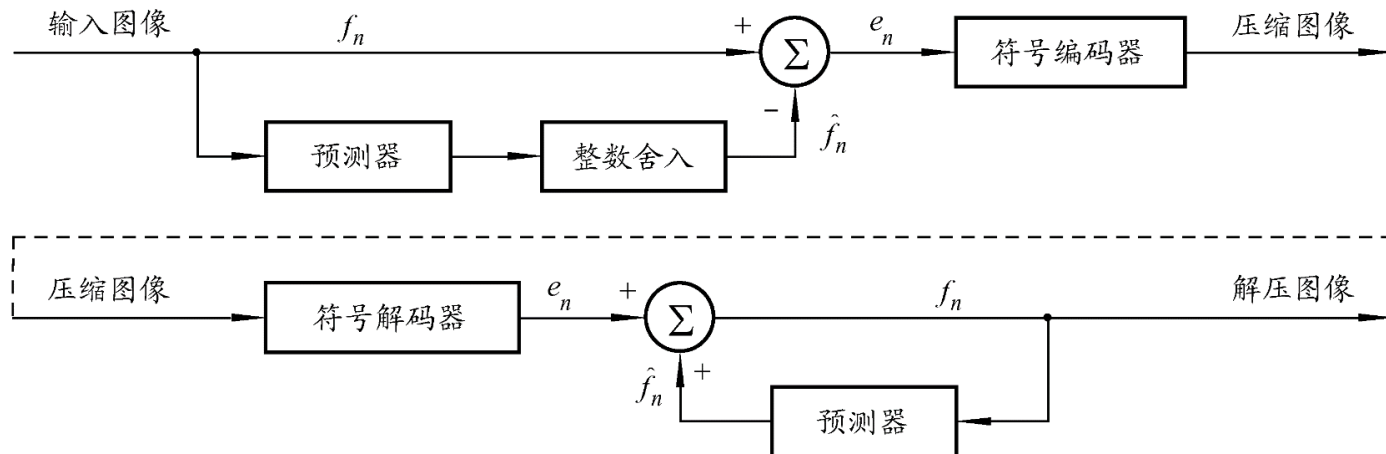
内容

- 图像压缩概述
- 变长编码
- 预测编码
- 变换编码
- 图像压缩国际标准

预测编码

- 利用图像中存在相关冗余来进行压缩
 - 用邻近像素来预测当前像素，求出预测误差
 - 再对误差图像进行编码，如哈夫曼
 - 误差图像的方差比原图像小，所需要码长更短

$$\hat{f}_n = \text{round} \left[\sum_{k=1}^m a_k f_{n-k} \right] \quad e_n = f_n - \hat{f}_n$$



预测编码举例

预测器 $\hat{f}_n = a_{n-2}f_{n-2} + a_{n-1}f_{n-1} \quad a_{n-2} = a_{n-1} = 0.5$

编码过程

f_n			2	4	6	8	8	4	2	10
\hat{f}_n			2	4	3	5	7	8	6	3
e_n			2	4	3	3	1	-4	-4	7

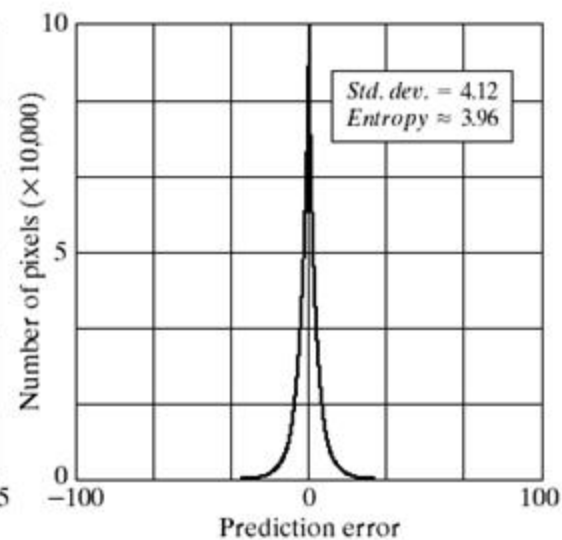
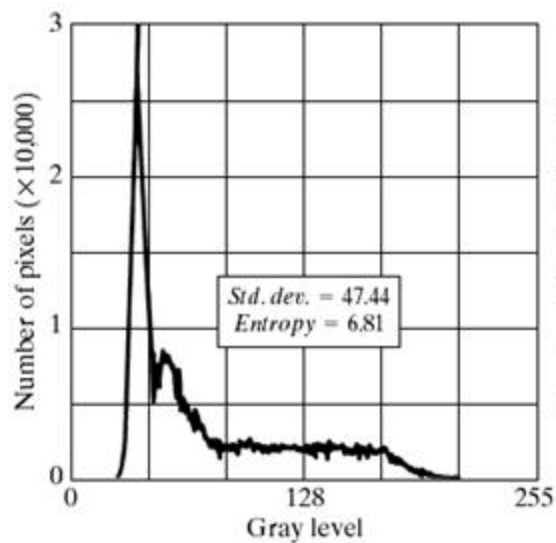
解码过程

e_n			2	4	3	3	1	-4	-4	7
\hat{f}			2	4	6	8	8	4	2	10

预测编码

- 预测器
 - 一阶预测器
 - 熵下降
 - 所需码长缩短

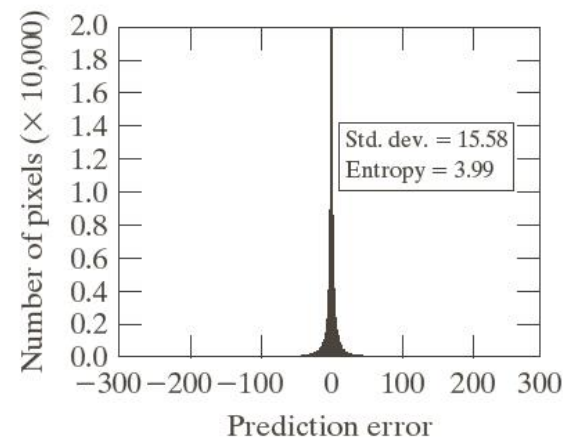
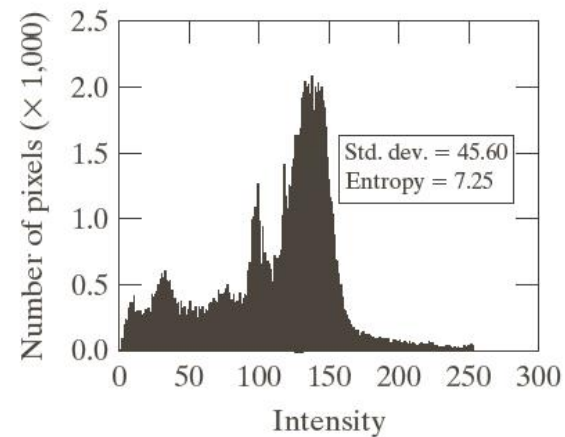
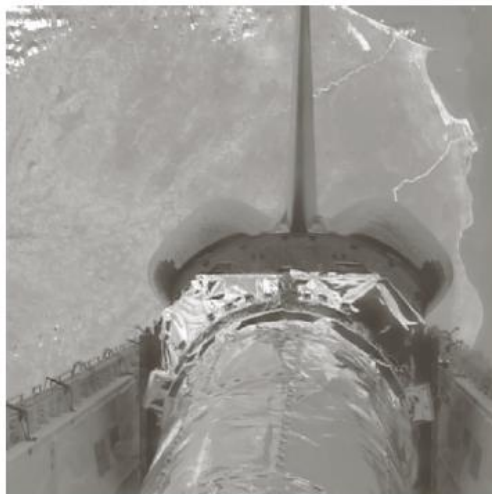
$$\hat{f}_n(x, y) = \text{round}[af(x, y - 1)]$$



预测编码

- 预测器
 - 一阶预测器
 - 熵下降
 - 所需码长缩短

$$\hat{f}_n(x, y) \\ = \text{round}[af(x, y - 1)]$$



内容

- 图像压缩概述
- 变长编码
- 预测编码
- 变换编码
- 图像压缩国际标准

变换编码

- 将图像经过变换（如变到频域）得到系数
- 对系数进行量化和编码
- 变换方法可以选择很多
 - 傅里叶变换
 - 离散余弦变换
 - 小波变换
- 变换编码方法是非信息保持型的（有损压缩）
- 是主流编码方式，JPEG采用此方法

离散余弦变换 DCT

- 一种频域变换

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

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

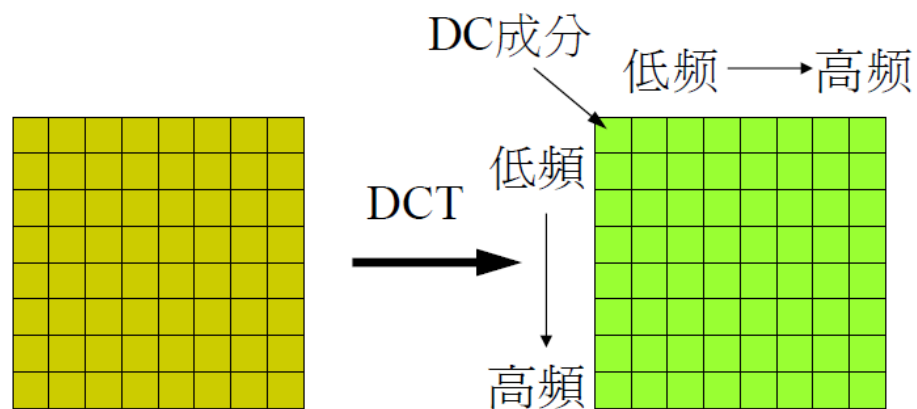
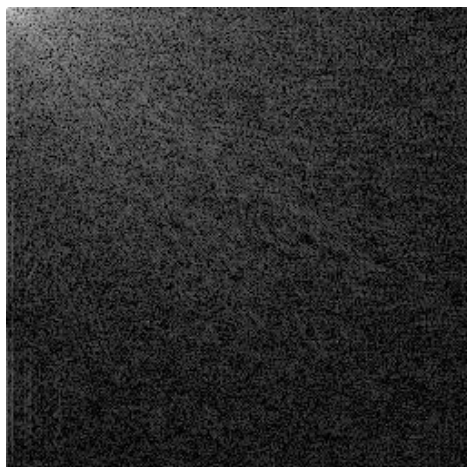
$$u, v = 0, 1, \dots, N-1$$

- $a(u)$ 为归一化加权系数

$$a(u) = \begin{cases} \sqrt{1/N} & u = 0 \\ \sqrt{2/N} & u = 1, 2, \dots, N-1 \end{cases}$$

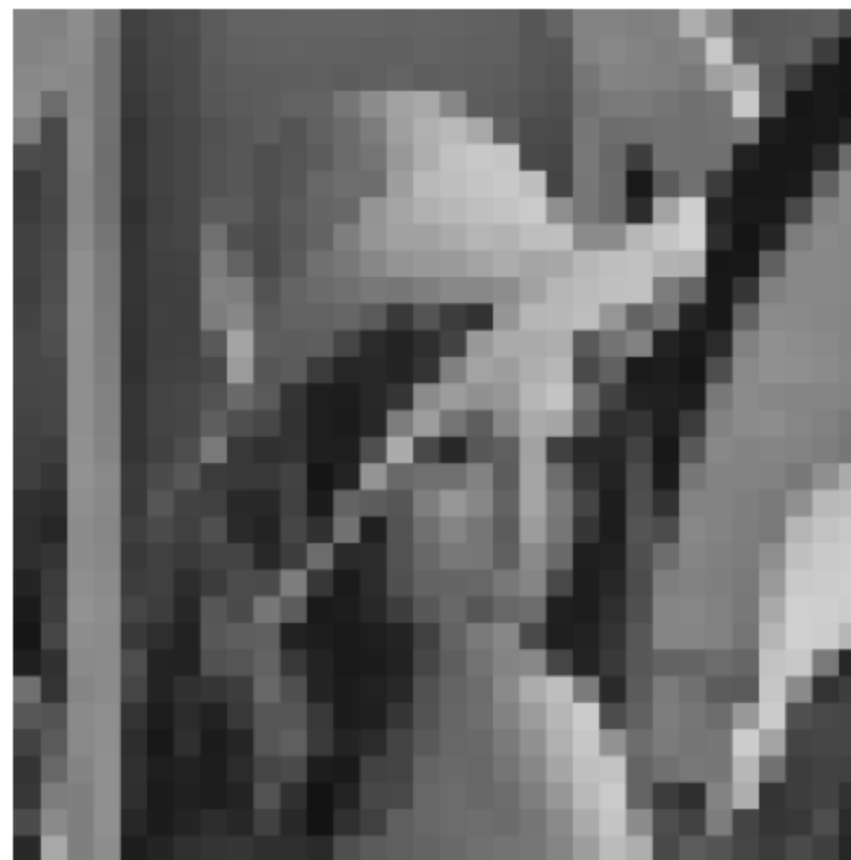
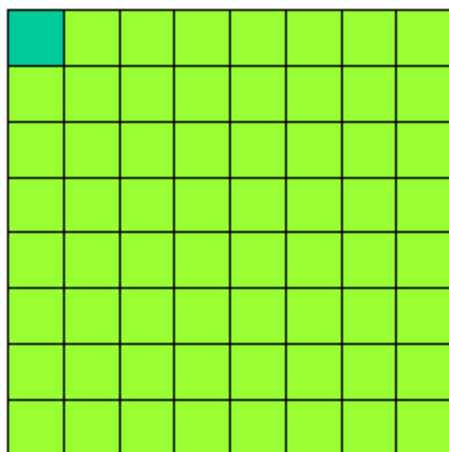
图像的DCT变换

- 图像分成8x8小块，做DCT变换



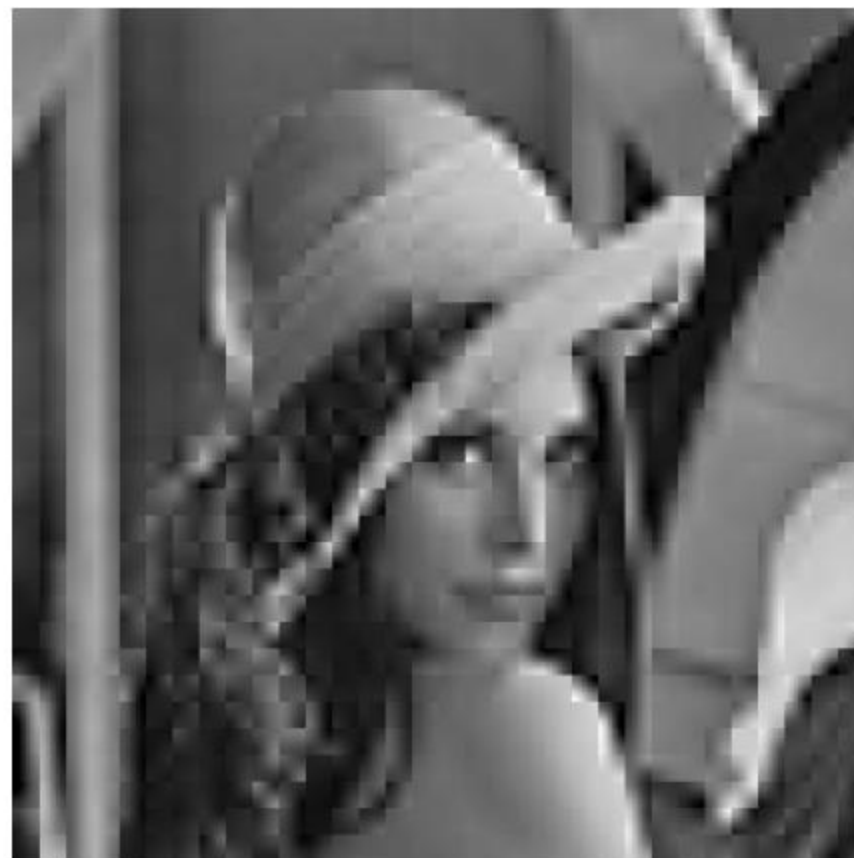
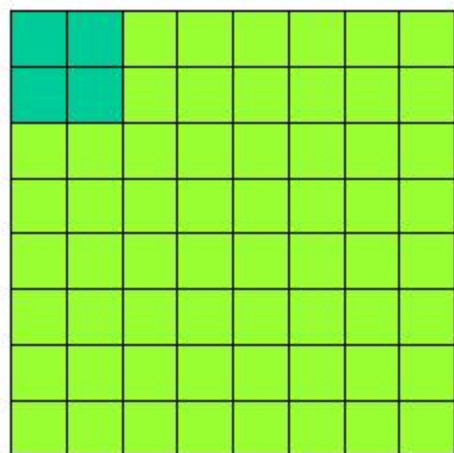
图像的DCT变换

- 仅保留直流分量后重建效果



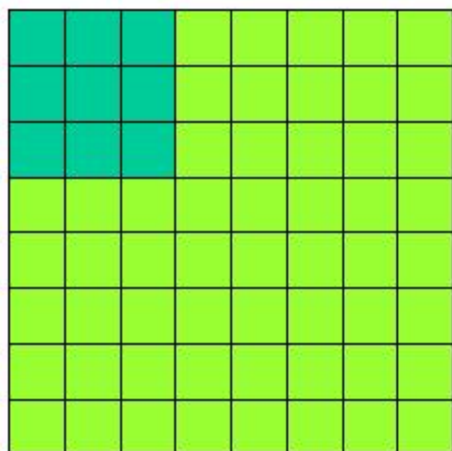
图像的DCT变换

- 保留4个分量后重建效果



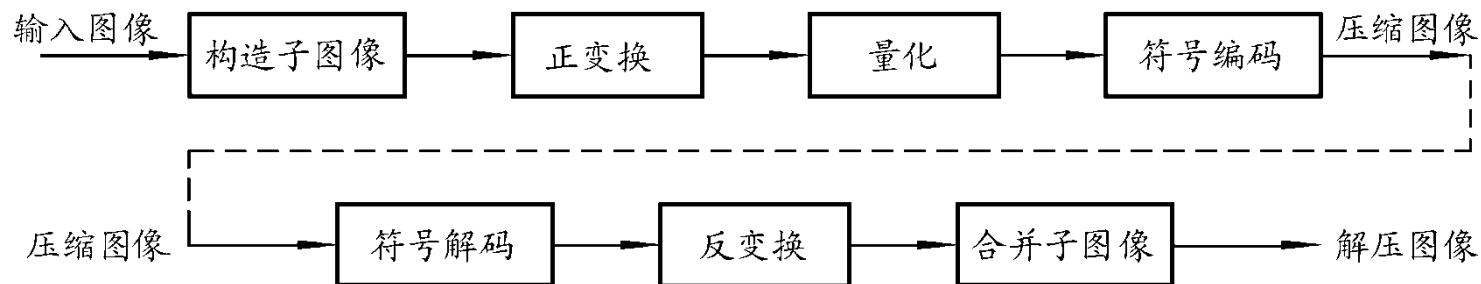
图像的DCT变换

- 保留9个分量后重建效果



基于DCT的变换编码

- 压缩部分
 - 构造子图像：将图像分成小块，如8x8
 - 变换：DCT变换
 - 量化：保留信息量大的系数，并确定其精度
 - 符号编码：哈夫曼编码等
- 解码部分是逆操作，除了量化



内容

- 图像压缩概述
- 变长编码
- 预测编码
- 变换编码
- 图像压缩国际标准

静止图像压缩国际标准

- JPEG

- 1986年成立的联合图像专家组 (joint picture expert group) 所制定的静止图像的压缩标准
- 使用离散余弦变换 DCT
- 通常达到25: 1的压缩率

- JPEG2000

- 由ISO和原CCITT两个组织的联合图像专家组于1997年开始征集提案, 于2000年问世
- 使用离散小波变换 DWT

运动图像压缩国际标准

- Motion JPEG
 - 用JPEG的方法对运动视频/电视信号进行编码，这也称为运动JPEG
- MPEG-1, MPEG-2, MPEG-4
 - 运动图像专家组（MPEG）制定的压缩标准
- H.264/AVC
 - JVT（joint video team）制定的一个面向未来IP和无线环境下的视频压缩的国际标准