

第 20 讲: DCT 图像压缩

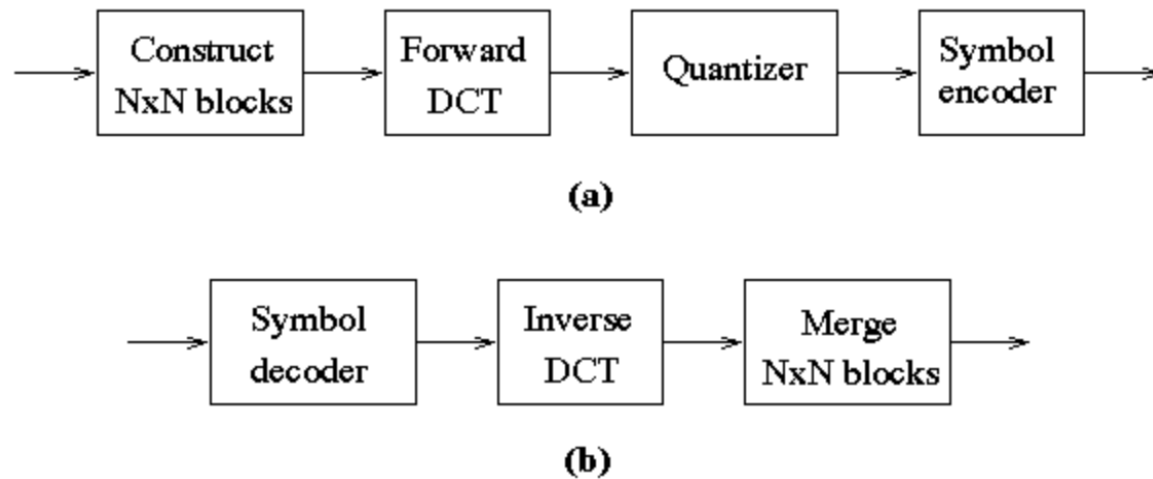
- 变换编码原理

- DCT
- 均匀量化
- 变换系数编码

- ISO JPEG标准

- 基本模式
 - 量化矩阵
 - 直流系数的预测编码
 - 交流系数的“之”字形扫描和编码
- 渐进编码模式
- 分层编码模式

变换编码原理



- 分为 **8×8 的块**以利用像素间的**局部相关性**。
- **变换编码**旨在将**最多的信息**压缩到**最小数量**的变换系数中。
- **量化器**可消除那些携带了无关紧要信息的系数。
- **符号编码器**通过熵编码（VLC）来压缩由量化的变换系数构成的符号。

为什么使用离散余弦变换？

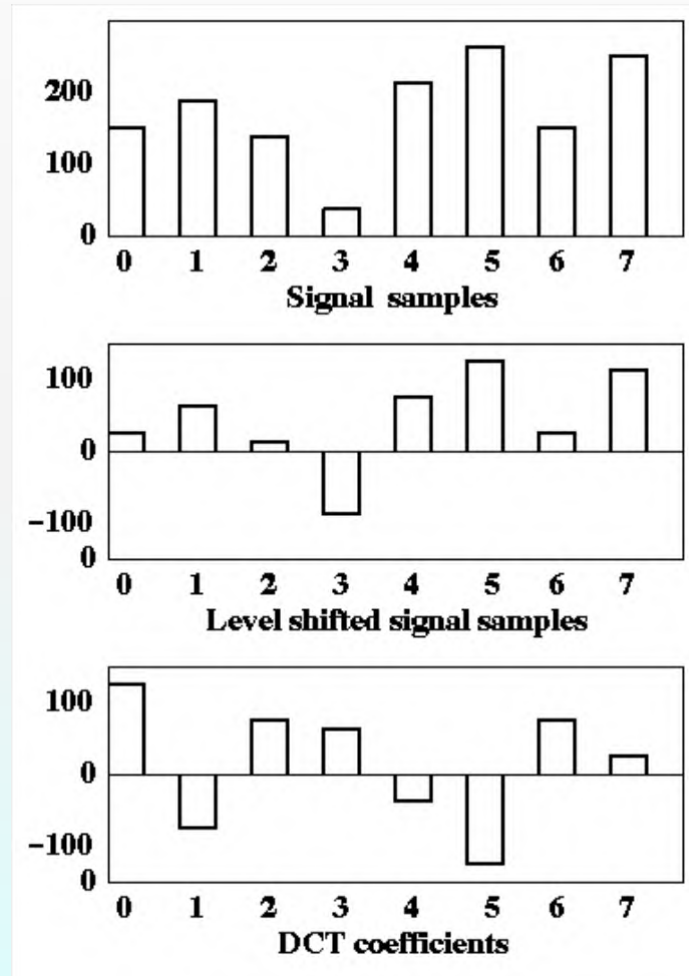
- DCT是**正交**的。变换域中的均方量化误差等于图像域中的均方误差。
- 可使用**FFT快速实现**。
- 2D DCT是**可分离**的。
- 接近**最佳能量压缩特性**，且其**基函数与图像无关**。
 - 对于具有高相关系数的平稳1阶马尔可夫信源，DCT的性能接近Karhunen-Loeve变换（KLT）。
- 有**专用硬件**。

演示: 1-D DCT

8 点信号

电平移位信号

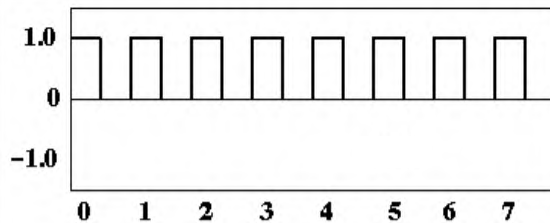
8 点 DCT



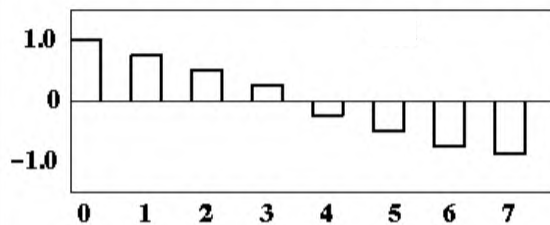
DCT 基函数

8 点 DCT

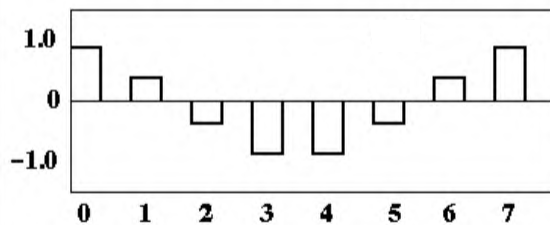
k=0, DC



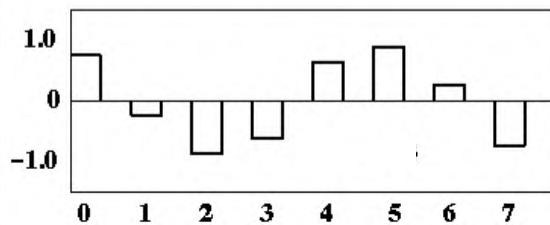
k=1



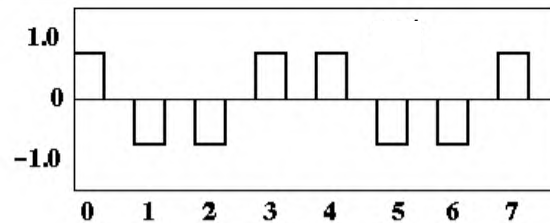
k=2



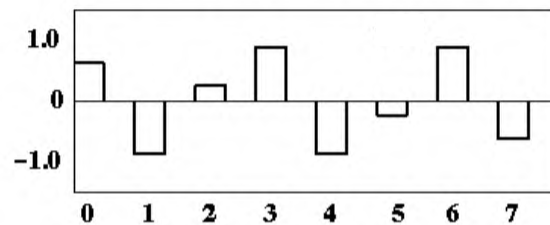
k=3



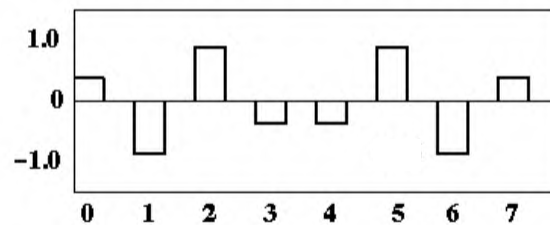
k=4



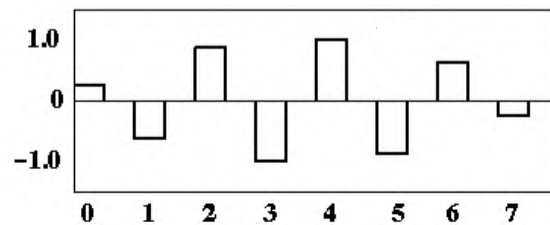
k=5



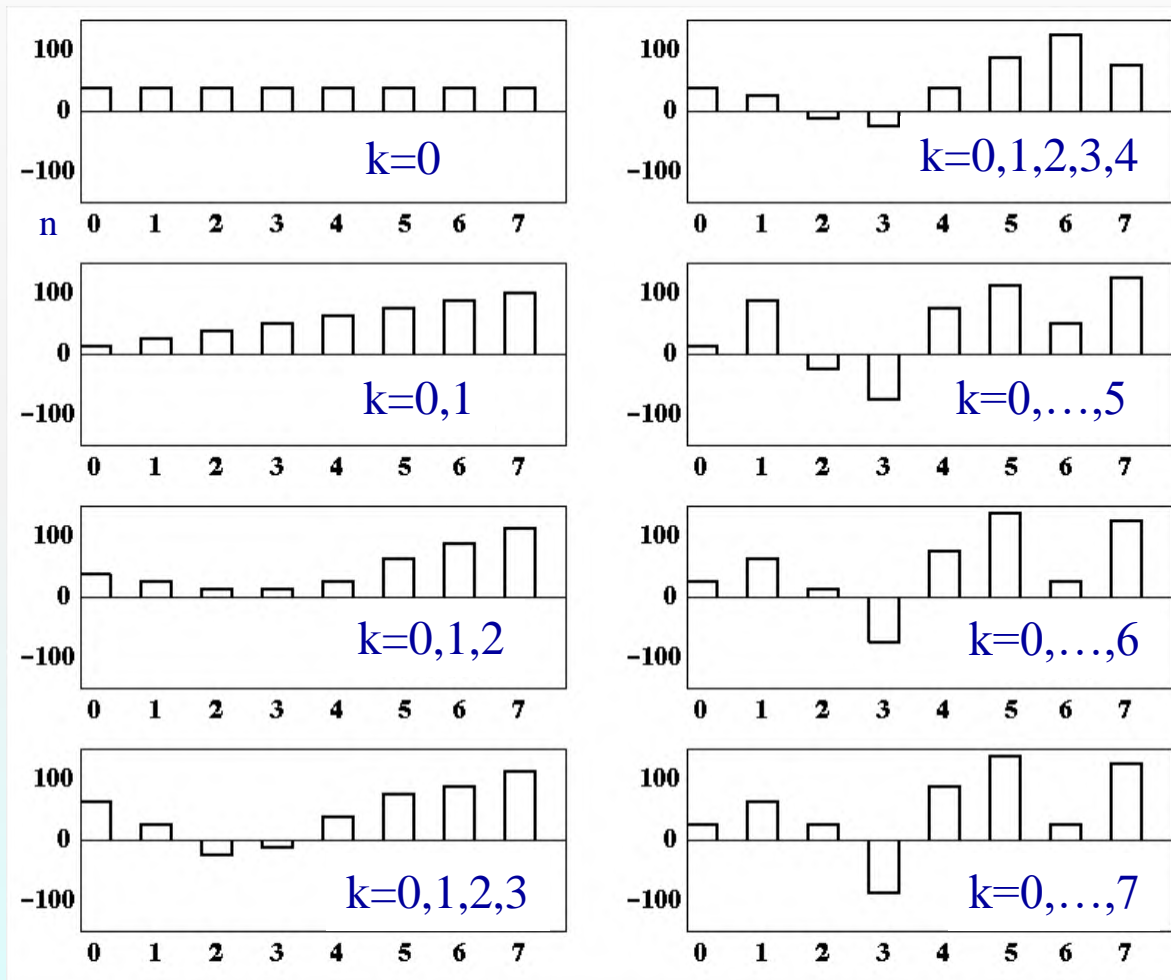
k=6



k=7



信号逐次相加



2D DCT

- **II 型 DCT 变换**施加到 8×8 的块:

$$S(k_1, k_2) = \sqrt{\frac{4}{N^2}} C(k_1) C(k_2) \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} s(n_1, n_2) \cos\left(\frac{\pi(2n_1+1)k_1}{2N}\right) \cos\left(\frac{\pi(2n_2+1)k_2}{2N}\right)$$

其中

$$C(k) = \begin{cases} 1/\sqrt{2} & \text{if } k = 0 \\ 1 & \text{otherwise} \end{cases}$$

- II型DCT的**对称扩展**由下式给出:

$$g(n) = s(n) + s(2N-1-n), \quad 0 \leq n \leq 2N-1$$

- 计算**N点信号** $s(n)$ 的N点DCT $S(k)$ 的**2N点 FFT**算法如下:

- 1) 构造 2N点**对称扩展信号** $g(n)$ 。
- 2) 计算 $g(n)$ 的**2N点DFT** $G(k)$, $k = 0, \dots, 2N-1$.
- 3) $S(k) = W_{2N}^{k/2} G(k)$, $k = 0, \dots, N-1$, 其中 $W_{2N}^k = e^{-j\frac{2\pi k}{2N}}$

示例: 通过DCT进行能量压缩

139	144	149	153	155	155	155	155
144	151	153	156	159	156	156	156
150	155	160	163	158	156	156	156
159	161	162	160	159	159	158	159
159	160	161	162	162	155	155	155
161	161	161	161	160	157	157	157
162	162	161	163	162	157	157	157
162	162	161	161	163	158	158	158

图像块 (Lena)

1260	-1	-12	-5	2	-2	-3	-1
-23	-17	-6	-3	-3	0	0	-1
-11	-9	-2	2	0	-1	-1	0
-7	-2	0	1	1	0	0	0
-1	-1	1	2	0	-1	1	1
2	0	2	0	-1	1	1	-1
-1	0	0	-1	0	2	1	-1
-3	2	-4	-2	2	1	-1	0

NINT[DCT 块]

- Ref. M. Rabbani and P. Jones, Digital Image Compression Techniques, SPIE Press, 1991.

均匀量化

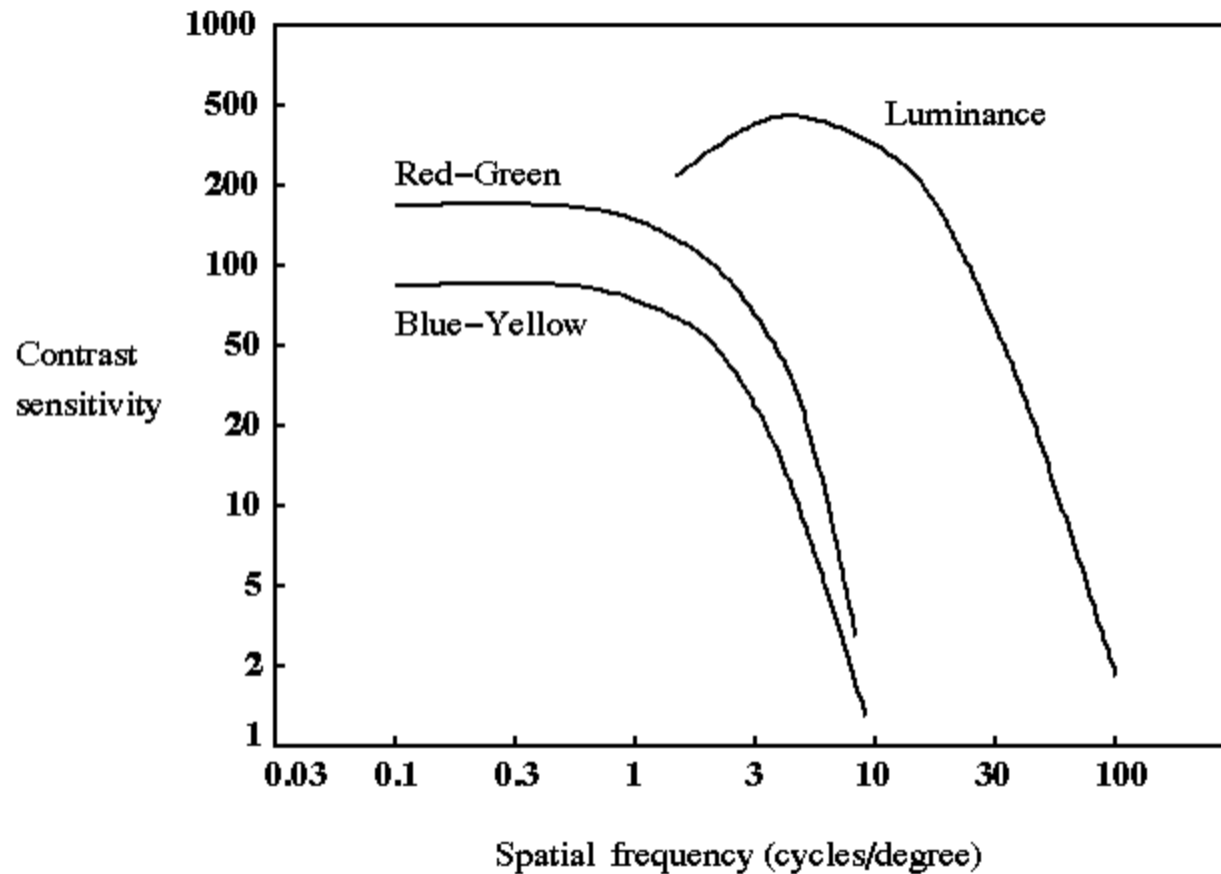
- 均匀量化可由下式来实现：

$$\hat{S}(k_1, k_2) = NINT \left\{ \frac{S(k_1, k_2)}{T(k_1, k_2)} \right\}$$

其中 $\hat{S}(k_1, k_2)$ 是 $S(k_1, k_2)$ 的量化近似， $T(k_1, k_2)$ 是**量化步长**或阈值。

- 具有**非零** $\hat{S}(k_1, k_2)$ 的系数被**保留**。
- 根据**相对视觉重要性**，在 8×8 块的DCT系数中选择 $T(k_1, k_2)$ 。（较粗的量化会使较大的 $T(k_1, k_2)$ 用于高频系数。）

人类视觉加权量化



变换系数编码

- 哪些变换系数是重要的？

- 区域编码：对固定数量和位置的变换系数进行编码，
例如低频系数
- 内容自适应：一维变换系数排序和游程编码

- 直流系数编码

- 直流系数的块间预测

静止图像压缩标准

ITU-T G3/G4	二进制图像 (非自适应) -发送传真文件 (1980)
ISO JBIG	二进制和位平面编码 -处理逐行式传输和半色调 (1994)
ISO JPEG	静止灰度和彩色图像 - 基于块的 DCT (1993)
ISO JPEG-LS	新的无损编码标准 -非线性预测, 基于上下文的Rice-Goulomb 编码 (1997)
ISO JPEG 2000	新的小波编码标准 (正在进行中)

JPEG 标准

- JPEG (ISO / IEC 10918-1 | ITU 建议T.81) 描述了**连续色调** (灰度或彩色) **静止图像**的压缩方法。
- 1988年1月的会议上, JPEG对自适应离散余弦变换方法达成共识。JPEG草案国际标准 (DIS) 文件于1991年10月提交。**JPEG于1993年成为国际标准。**
- JPEG**不是用于图像交换的完整架构**。没有指定特定的文件格式、空间分辨率或颜色空间模型。

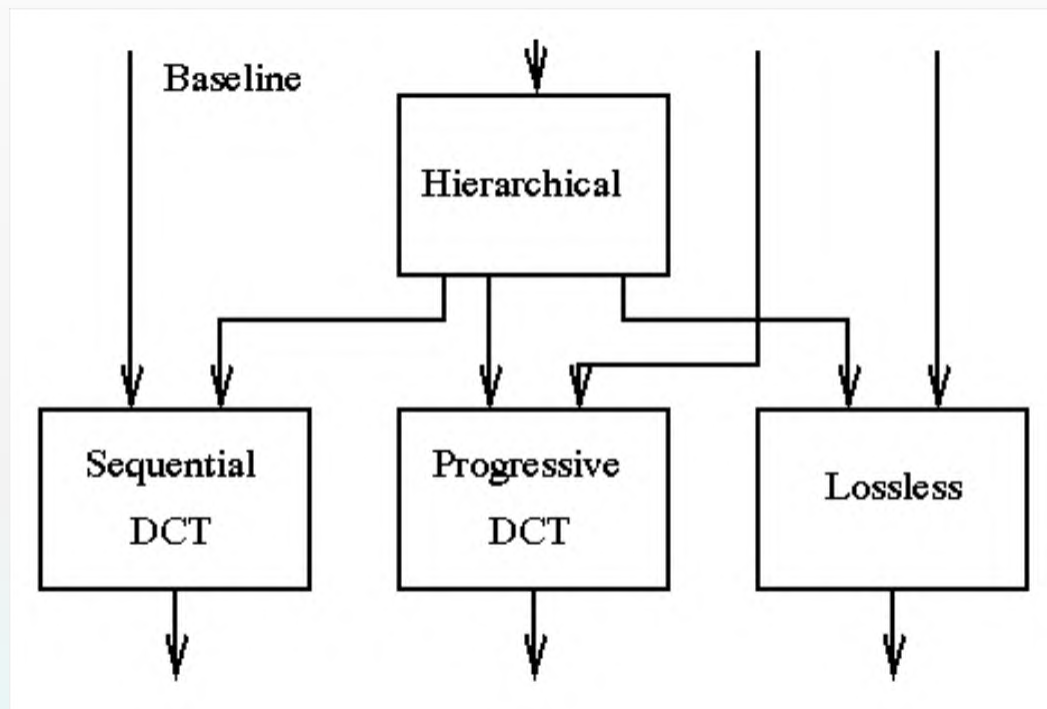
允许JPEG压缩图像的图像文件格式示例:

- JPEG文件交换格式 (**JFIF**) 是一种最小的文件格式, 可以进行JPEG比特流的交换。它使用标准色彩空间 (ITU-R 601-1) 。
- 标签图像文件格式 (**TIFF™**) 版本6.0支持JPEG基本和霍夫曼无损编码模式, 可以是灰度, RGB, CMYK和YUV色彩空间。

JPEG 特征

- **独立分辨率**：任意信源分辨率可以通过将图像内部填充到8像素的倍数以适应DCT操作模式。
- **精度**：8比特和12比特/像素的DCT操作模式。
在无损模式下，精度可以在2到16比特/像素之间，尽管JBIG在**4-5比特/像素以下**表现更好。
- **无绝对比特率目标**：比特率/质量权衡主要由量化矩阵控制。
- **亮度 - 色度可分离性**：从编码的彩色图像中恢复仅亮度图像而无需解码色度的能力。
- **可扩展性**：渐进式或分级的层数没有界限。

JPEG 操作模式



- 要兼容JPEG，系统必须支持基本算法。

Ref. W. B. Pennebaker and J. L. Mitchell, **JPEG Still Image Data Compression Standard**, Van Nostrand Reinhold, 1993.

JPEG 基本算法

- **DCT计算**：图像分为 8×8 块。块的2-D DCT在电平移位后计算。
- **DCT系数和之字形扫描的量化**：默认的对色度和亮度的量化矩阵可被调整以提供各种压缩级别。
- **可变长编码 (VLC)**：使用一种定义了系数值和连零个数的VLC码字对AC系数进行编码。指定了标准VLC码表。DC系数是使用DPCM编码的。

在基本模式下，输入和输出数据精度为8比特/像素，量化后DCT值限制为11比特/样本。

量化矩阵

- **平坦量化**：可以对所有DCT系数使用一个量化阶（阈值）。
- **HVS加权量化**：根据人类视觉系统响应，阈值随频率而变化。所有块都使用单个阈值矩阵。
- **自适应量化**：通过简单的缩放，量化矩阵可以被允许从块到块改变。缩放参数称为“mquant”。
 - 在所有情况下，保留系数的位置因块而异。

JPEG 缺省量化矩阵

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

亮度量化表

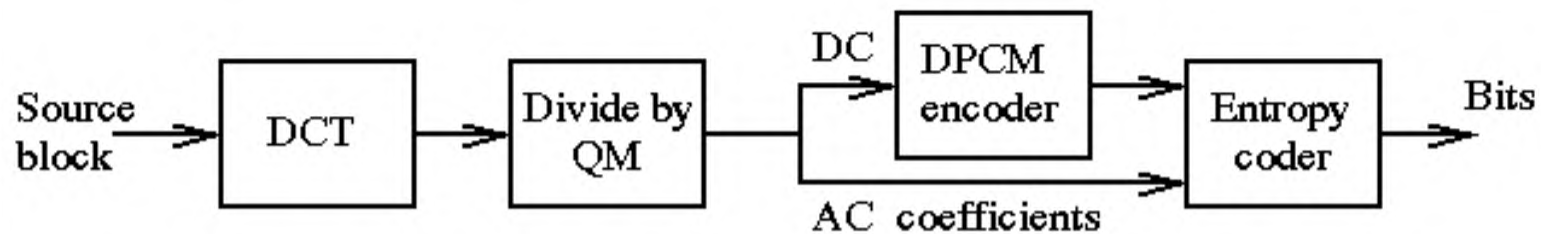
17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

色度量化表

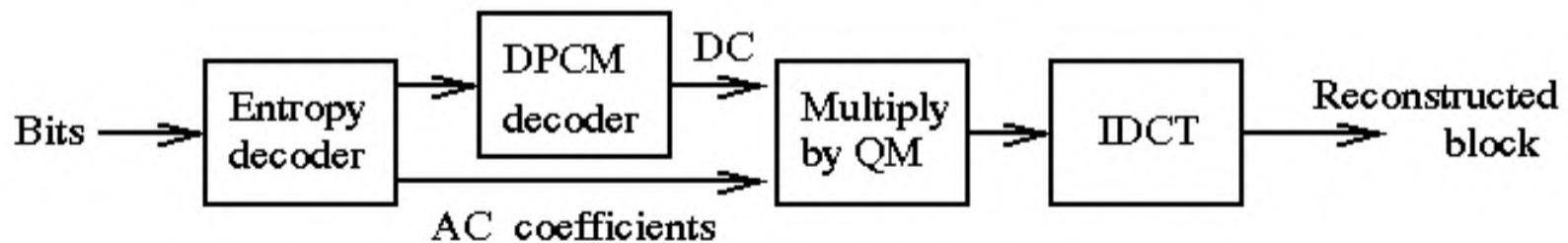
JPEG 自适应量化

- 允许**空间自适应量化**，其中量化矩阵可以在**块与块之间缩放**。
 - 例如，基于对亮度方差的测量，在高活动（边缘），中等活动（纹理）和统一块之间进行区分。
 - 与非自适应量化相比，性能提高30%。
- ISO DIS 10918-3, JPEG Extension, August 1994.

系数的编/解码



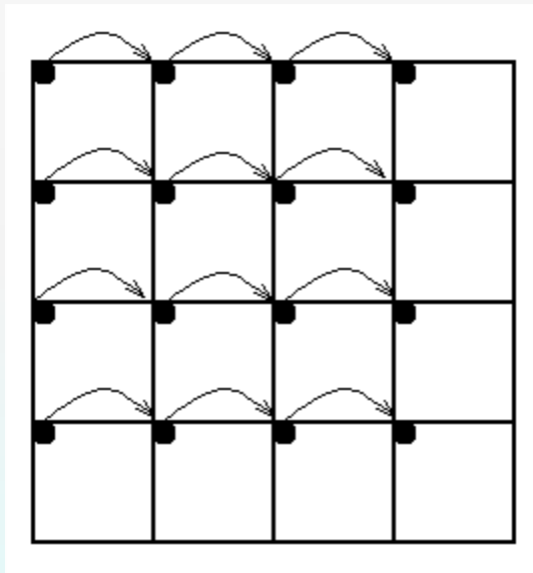
(a)



(b)

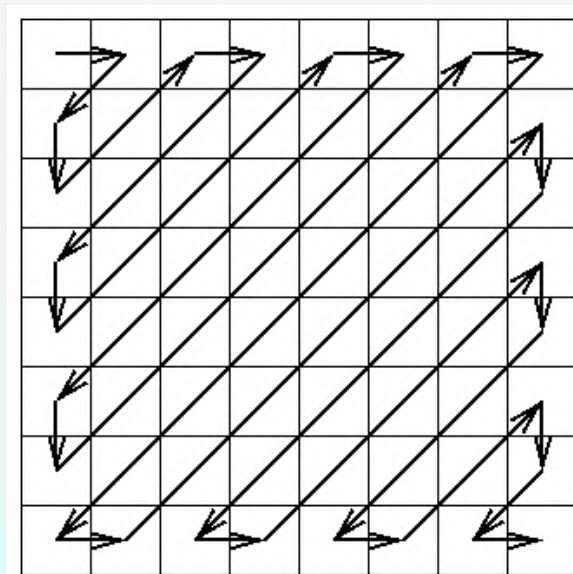
DC系数编码

DC系数的DPCM编码（也称为DC预测）



AC系数编码

- 由于保留系数的位置因块而异，所以量化后的AC系数被**之字形扫描**并形成（**游程**，**量化电平**）对。
- **量化电平** = 非零系数的值；
游程 = 在非零系数之前的零的个数。
- 对这些(游程，量化电平)对进行**熵编码**，即对出现频率较低的对使用长码，反之则使用短码)。



JPEG基本算法 — 示例

-76	-73	-67	-62	-58	-67	-64	-55
-65	-69	-62	-38	-19	-43	-59	-56
-66	-69	-60	-15	16	-24	-62	-55
-65	-70	-57	-6	26	-22	-58	-59
-61	-67	-60	-24	-2	-40	-60	-58
-49	-63	-68	-58	-51	-65	-70	-53
-43	-57	-64	-69	-73	-67	-63	-45
-41	-49	-59	-60	-63	-52	-50	-34

电平移位后的 8×8 原始图
像块 (移位了 128)

-415	-29	-62	25	55	-20	-1	3
7	-21	-62	9	11	-7	-6	6
-46	8	77	-25	-30	10	7	-5
-50	13	35	-15	-9	6	0	3
11	-8	-13	-2	-1	1	-4	1
-10	1	3	-3	-1	0	2	-1
-4	-1	2	-1	2	-3	1	-2
-1	-1	-1	-2	-1	-1	0	-1

DCT变换结果

示例 (cont'd)

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

JPEG 建议的量化矩阵

-26	-3	-6	2	2	0	0	0
1	-2	-4	0	0	0	0	0
-3	1	5	-1	-1	0	0	0
-4	1	2	-1	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

量化后系数矩阵

示例 (cont'd)

- 之字形扫描后的**1-D**系数序列:

-26 -3 1 -3 -2 -6 2 -4 1 -4 1 1 5 0 2 0 0 -1 2 0 0 0 0 0 -1 -1 EOB
其中EOB表示块的结尾。

- **编码DC系数**: 对当前块和先前块的**DC**系数之间的差进行编码。
- **编码AC系数**: 定义 (RUN, LEVEL) 为符号; 例如(0,-3); (0,1); (0,-3)...
- 对这些符号进行**VLC (霍夫曼或算术) 编码**。

示例: JPEG 解码器

- 解码器是**实现逆运算**。也就是说，首先将接收到的系数乘以相同的量化矩阵以获得：

320	55	-30	16	72	-80	51	0
-36	-24	14	38	26	0	0	0
-14	-13	16	24	40	0	0	0
-14	0	0	29	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

- 执行**逆DCT**，并将每个元素**加上128**：

195	140	119	148	197	171	120	170
186	153	143	158	193	168	124	175
174	169	172	168	187	166	128	177
169	180	187	171	185	170	131	170
172	182	186	168	186	175	131	161
177	180	181	168	189	175	129	161
181	178	181	174	191	169	128	173
183	178	184	181	192	162	127	185

- 重建误差**最多为 ± 25 灰度级

编码器控制和压缩损伤

- 在JPEG中，解码图像质量和比特率（文件大小）之间的折中可以通过对默认量化矩阵进行缩放的**全局缩放因子**进行控制。
- 以**大于1的因子**对量化矩阵进行缩放导致较粗略的量化，可降低比特率，但代价是更高的压缩误差。
- **压缩比（CR）** 被定义为原始图像与压缩图像的码率的比率。对于8比特单色图像， $CR = 8$ （即，**1比特/像素**）可实现大多数图像的视觉无损压缩。在 $CR = 15$ 或更高时则会观察到严重的振铃和块效应。
- 大多数JPEG的实现为用户提供了用于折中图像质量和图像尺寸的参数，例如**缩放量化矩阵的品质因子（QF）**参数。QF的值的范围通常在1-100之间，其中 $QF = 75$ 对应于未缩放的默认矩阵，但是也可能因具体实现而有所不同。

彩色图像压缩

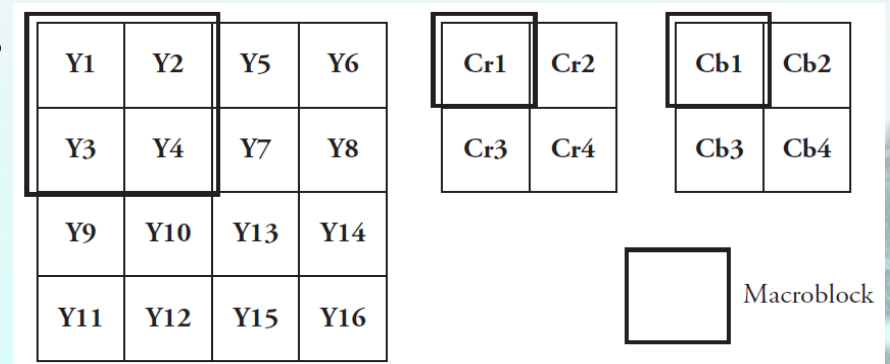
- JPEG使用**标准色彩空间**（ITU-R 601-1）。它将RGB图像转换为亮度色度空间，称为**Y-Cr-Cb空间**，由下式定义：

$$Y = 0.3 R + 0.6 G + 0.1 B$$

$$Cr = \frac{B - Y}{2} + 0.5$$

$$Cb = \frac{R - Y}{1.6} + 0.5$$

- 因为人眼对色度分量**Cr和Cb的高频分量相对不敏感**，所以它们在两个方向上被**二倍下采样**。
- **宏块**：四个8×8亮度块和相关联的色度块。比特流以**宏块顺序组织**。



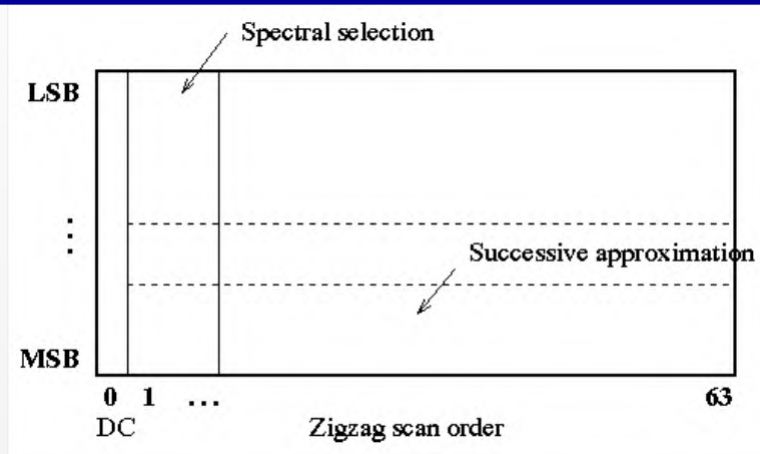
交错与非交错扫描

- 一幅彩色图像可以被编码为 **非交错** (3 个独立的扫描) 或 **交错** (一个独立的扫描).
- **非交错:**
 - 扫描 1: $Y_1, Y_2, Y_3, \dots, Y_{16}$
 - 扫描 2: Cr_1, Cr_2, Cr_3, Cr_4
 - 扫描 3: Cb_1, Cb_2, Cb_3, Cb_4
- **交错:**

$Y_1, Y_2, Y_3, Y_4, Cr_1, Cb_1, Y_5, Y_6, Y_7, Y_8, Cr_2, Cb_2, \dots$

对于交错扫描, 块根据帧和扫描头数据中指定的参数进行排序。

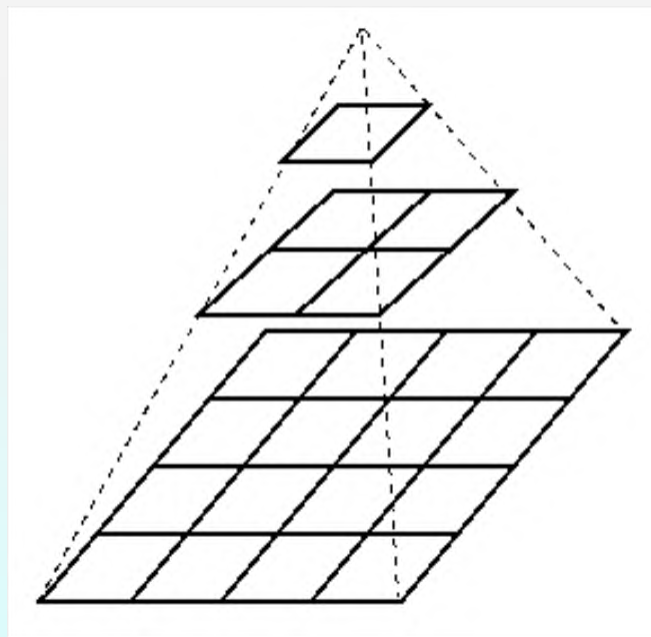
JPEG – 渐进

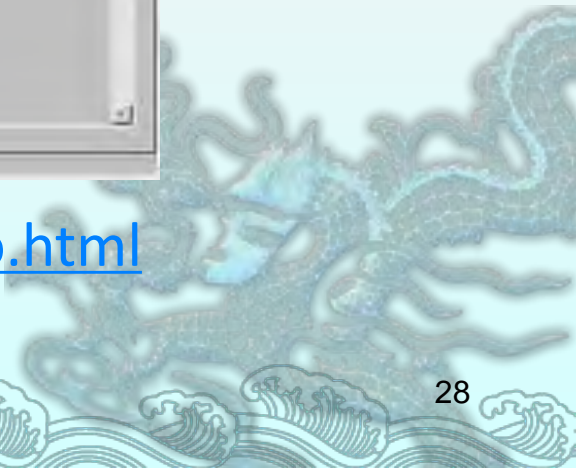


- **渐进模式**由“扫描”序列组成，每个扫描序列都对一部分量化的DCT系数进行编码。
- **频谱选择**：DCT系数被分组成各个频带。较低的频带通常首先被编码（发送）。
- **连续近似**：首先以较低的精度发送信息，然后在稍后的扫描中进行细化。
- 可以**组合两个过程**以提供平滑变化的质量。

JPEG -分层

- **第一级（最低分辨率）** 使用顺序或逐行JPEG模式之一进行编码。然后对每个分级阶段的输出进行上采样（内插），并用作下一阶段的预测。
- **极低比特率**的图像质量超过了任何其他JPEG模式，但这是以描述所有级数的较高比特率为代价的。





<http://homepage.tudelft.nl/c7c8y/VcDemo.html>

习题

- 7.1
- 7.2
- 7.6
- 7.8 $E\{(X_n - \rho X_{n-1})X_n\} = 0$ 改为: $E\{(X_n - \rho X_{n-1})X_{n-1}\} = 0$