

Chp13 变换编码

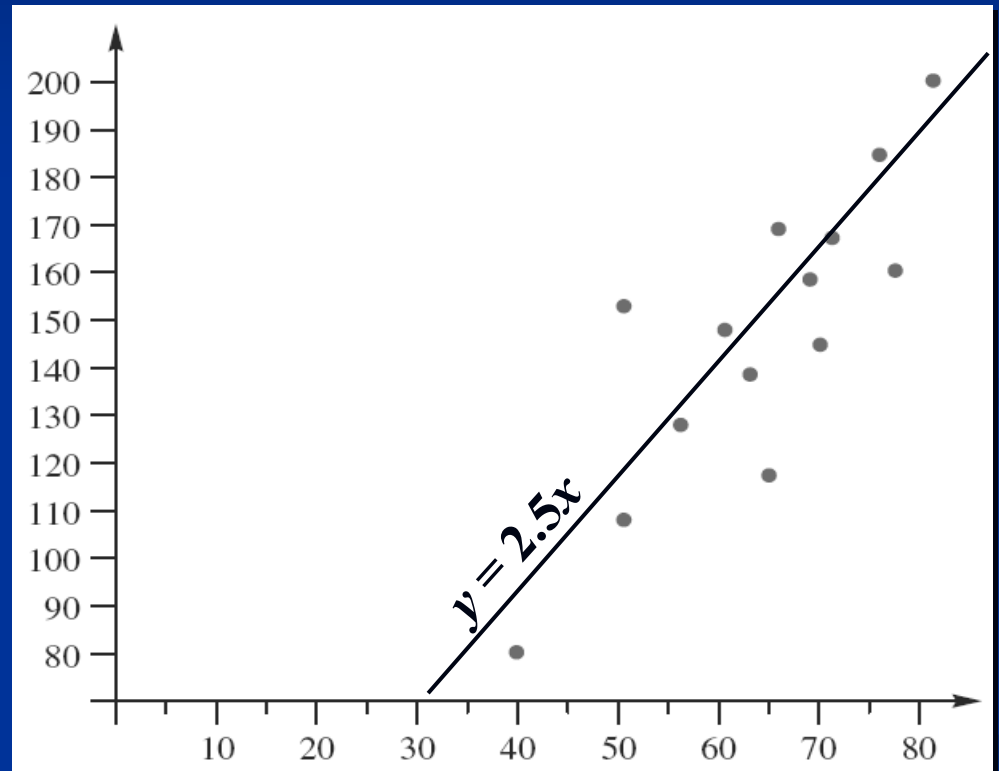
- 变换编码(Transform Coding): 将信源输出分解/变换为其组成部分, 然后根据每个成分的特性分别进行编码
 - 如果取得一个输入序列, 并将其转换为另一序列, 其中大多数信息都包含在少数几个元素中, 那就可以对这些元素及其在新序列中的位置进行编码和传送, 从而实现数据压缩。

主要内容

- Karhunen-Loève变换(KLT)
- 离散余弦变换(discrete cosine transform, DCT)
- 变换系数的量化编码
- JPEG

例：动机

Height	Weight
65	170
75	188
60	150
70	170
56	130
80	203
68	160
50	110
40	80
50	153
69	148
62	140
76	164
64	120



例：动机 ——旋转

■ 考虑（可逆）变换

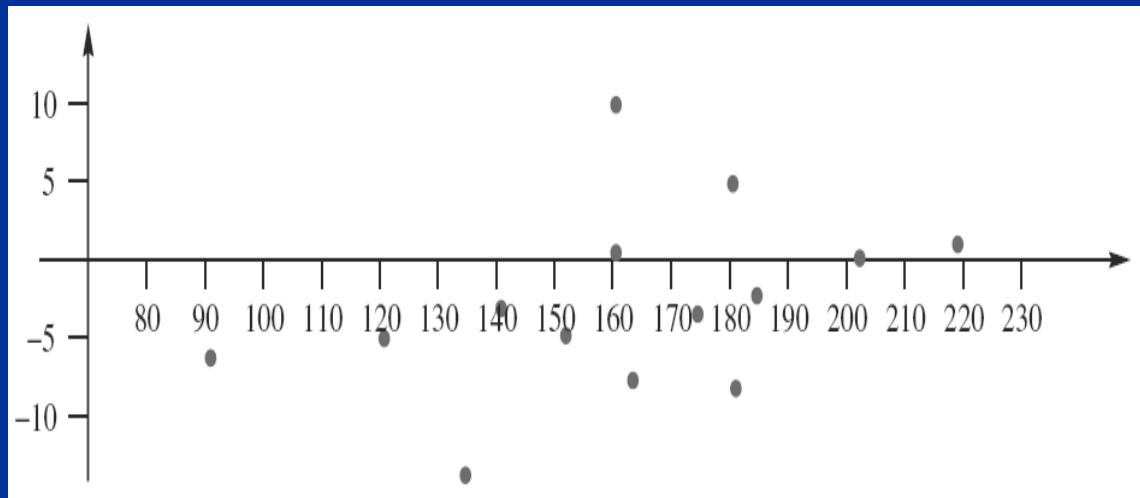
$$\boldsymbol{\theta} = \mathbf{A}\mathbf{x}$$

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}, \mathbf{A} = \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix}, \phi = \arctan 2.5$$

$$\mathbf{A} = \begin{bmatrix} 0.37139068 & 0.92847669 \\ -0.92847669 & 0.37139068 \end{bmatrix}$$

例：动机 ——变换后的序列

First Coordinate	Second Coordinate
182	3
202	0
162	0
184	-2
141	-4
218	1
174	-4
121	-6
90	-7
161	10
163	-9
153	-6
181	-9
135	-15



例：动机 ——压缩

- 抛弃坐标第二维...
- 用定长码编码，可降低50%！

例：动机 ——重构序列

Height	Weight
65	170
75	188
60	150
70	170
56	130
80	203
68	160
50	110
40	80
50	153
69	148
62	140
76	164
64	120

原始值

Height	Weight
68	169
75	188
60	150
68	171
53	131
81	203
65	162
45	112
34	84
60	150
61	151
57	142
67	168
50	125

重构值

$$\mathbf{x} = \mathbf{A}^{-1}\boldsymbol{\theta}, \quad \mathbf{A}^{-1} = \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix}$$

例：动机 ——误差分析

$\{\hat{x}_n\} \equiv$ 重构序列

$$\hat{\theta}_i = \begin{cases} \theta_i & i = 0, 2, 4, \dots \\ 0 & \text{otherwise} \end{cases}$$

$$\sum_{i=0}^{N-1} (x_i - \hat{x}_i)^2 = \sum_{i=0}^{N-1} (\theta_i - \hat{\theta}_i)^2$$

- 误差取决于被置为0的那些 θ_n 的幅值
 - 如果幅值很小，则误差也很小
 - 即大多数信息在每个数据对的第一个元素中

例：动机

——从统计分析的角度

- 上面我们从几何的角度分析了变换
- 我们也可以从统计分析的角度来审视变换：
 - 当变换对序列去相关时，能得到最大压缩
 - 主成分方法(Principle Component Method)
 - 将多个变量通过线性变换以选出较少个重要变量的多元统计方法
 - 数学上的降维方法。当数据的维度很大时，对高维数据的（距离）计算可能会耗费大量的计算资源，通过**PCA**可以降低计算的复杂性避免过拟合现象的发生。
 - 即样本—样本之间的相关性为0

变换编码

■ 变换

- 将原始序列 $\{x_n\}$ 分成大小为 N 的块
- 将每个块映射成变换序列 $\{\theta_n\}$
 - 可逆映射
 - 每块的不同元素通常有不同的统计特性（如前例）

■ 量化，根据

- 目标平均码率
- 变换序列各元素的统计特性
 - 可能对不同的子序列采取不同的技术
- 变换后系数的失真对重构序列的影响

■ 熵编码

- 定长码、Huffman编码、算术编码、RLE+算术编码 ...

变换

- 我们主要考虑线性变换:

$$\theta_n = \sum_{i=0}^{N-1} x_i a_{n,i}$$

- $\{\theta_n\}$ 每个元素的特性与其位置有关
 - 如: 在上述例子中元素的位置是奇数还是偶数

- 设计

- 变换序列中每个元素的方差决定编码策略
- N 与特定领域有关, 基于实践考虑
 - 计算代价、延迟、信源统计特性的稳定性

- 重构:

$$x_n = \sum_{i=0}^{N-1} \theta_i b_{n,i}$$

变换(2)

- 写成矩阵形式 (1-D) :

$$\boldsymbol{\theta} = \mathbf{A} \mathbf{x} \quad [\mathbf{A}]_{i,j} = a_{i,j}$$

$$\mathbf{x} = \mathbf{B} \boldsymbol{\theta} \quad [\mathbf{B}]_{i,j} = b_{i,j}$$

$$\mathbf{A} \mathbf{B} = \mathbf{B} \mathbf{A} = \mathbf{I}$$

- 2-D变换:

$$\Theta_{k,l} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} X_{i,j} a_{i,j,k,l}$$

- 可分解的2-D变换:

$$\Theta_{k,l} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} a_{k,i} X_{i,j} a_{i,j}$$

变换(3)

- 矩阵形式的可分解2-D变换:

$$\Theta = \mathbf{A}\mathbf{X}\mathbf{A}^T$$

$$\mathbf{X} = \mathbf{B}\Theta\mathbf{B}^T$$

- 我们要讨论的变换都是正交变换(Orthonormal transforms)

正交变换

$$\Theta = \mathbf{A} \mathbf{X} \mathbf{A}^T$$

$$\mathbf{X} = \mathbf{B} \Theta \mathbf{B}^T$$

- 正交矩阵：如果一个矩阵满足

$$\mathbf{A} \mathbf{A}^T = \mathbf{A}^T \mathbf{A} = \mathbf{I}$$

- 其中 \mathbf{I} 为单位阵，则称 \mathbf{A} 为正交矩阵。

- 每行/列点积为1，不同行/列的点积为0（归一化正交）

- 正交矩阵的逆矩阵等于其转置： $\mathbf{B} = \mathbf{A}^{-1} = \mathbf{A}^T$

- 正交反变换：

$$\mathbf{X} = \mathbf{A}^T \Theta \mathbf{A}$$

正交变换(2)

$$\boldsymbol{\theta} = \mathbf{A}\mathbf{x}$$

- 能量守恒性质:

$$\sum_{i=0}^{N-1} \theta_i^2 = \boldsymbol{\theta}^T \boldsymbol{\theta} = (\mathbf{A}\mathbf{x})^T \mathbf{A}\mathbf{x} = \mathbf{x}^T \mathbf{A}^T \mathbf{A}\mathbf{x} = \mathbf{x}^T \mathbf{x} = \sum_{i=0}^{N-1} x_i^2$$

- 总能量守恒，但通常能量在各系数上分布并不均匀

- 变换编码增益为
$$G_{TC} = \frac{\sigma_X^2}{\sqrt[N]{\prod_{k=0}^{N-1} \sigma_{\theta_k}^2}} = \frac{\frac{1}{N} \sum_{k=0}^{N-1} \sigma_{\theta_k}^2}{\sqrt[N]{\prod_{k=0}^{N-1} \sigma_{\theta_k}^2}}$$

变换系数方差的算术均值

变换系数方差的几何均值

- 增益与系数方差的集中程度有关
- 若每个系数的方差相等，则没有增益

从信号分解的角度

$$\mathbf{x} = \mathbf{A}^T \boldsymbol{\theta}$$

- 考虑变换（从反变换观察）：反变换为

$$\begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} a_{00} & a_{10} \\ a_{01} & a_{11} \end{bmatrix} \times \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix} = \theta_0 \begin{bmatrix} a_{00} \\ a_{01} \end{bmatrix} + \theta_1 \begin{bmatrix} a_{10} \\ a_{11} \end{bmatrix}$$

- 变换行 = 基向量

- $\theta_i = \langle a_i^T, x \rangle$ 表示 x 与 a_i 之间的相似性
- 相似性越高，变换系数越大

$$\boldsymbol{\theta} = \mathbf{A} \mathbf{x}$$

- 例：

- 第一行：低通信号， x_0 和 x_1 的均值
- 第二行：高通信号， x_0 和 x_1 的差值
- \mathbf{A} 为正交矩阵

$$\mathbf{A} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \times \begin{bmatrix} \alpha \\ \alpha \end{bmatrix} = \begin{bmatrix} \sqrt{2}\alpha \\ 0 \end{bmatrix}$$

从信号分解的角度(2)

■ 例(续): 考虑两个序列:

■ 低通: (3, 1)

■ 高通: (3, -1)

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \times \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} \\ \sqrt{2} \end{bmatrix}$$

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \times \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} \sqrt{2} \\ 2\sqrt{2} \end{bmatrix}$$

矩阵的角度

- 1-D变换：变换矩阵的行展开
- 类似的，2-D变换可视为变换矩阵行向量外积构成的矩阵的展开

$$\mathbf{A} = \begin{bmatrix} a_{00} & a_{01} & \cdots & a_{0N-1} \\ a_{10} & a_{11} & \cdots & a_{1N-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N-10} & a_{N-11} & \cdots & a_{N-1N-1} \end{bmatrix}$$

外积：

$$\alpha_{i,j} = \begin{bmatrix} a_{i0} \\ a_{i1} \\ \vdots \\ a_{iN-1} \end{bmatrix} \begin{bmatrix} a_{j0} & a_{j1} & \cdots & a_{jN-1} \end{bmatrix} = \begin{bmatrix} a_{i0}a_{j0} & a_{i0}a_{j1} & \cdots & a_{i0}a_{jN-1} \\ a_{i1}a_{j0} & a_{i1}a_{j1} & \cdots & a_{i1}a_{jN-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{iN-1}a_{j0} & a_{iN-1}a_{j1} & \cdots & a_{iN-1}a_{jN-1} \end{bmatrix}$$

矩阵的角度(2)

- 例: $\mathbf{A} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
- 外积为:

$$\alpha_{0,0} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \times \frac{1}{\sqrt{2}} [1 \quad 1] = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

$$a_{0,1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \times \frac{1}{\sqrt{2}} [1 \quad -1] = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$$

$$\alpha_{1,0} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \times \frac{1}{\sqrt{2}} [1 \quad 1] = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix},$$

$$a_{0,1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \times \frac{1}{\sqrt{2}} [1 \quad -1] = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

矩阵的角度(3)

- 例（续）： $\mathbf{A} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
- 逆变换为：

$$\mathbf{B} = \mathbf{A}^{-1} = \mathbf{A}^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} x_{00} & x_{01} \\ x_{10} & x_{11} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \theta_{00} & \theta_{01} \\ \theta_{10} & \theta_{11} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$= \theta_{00} \alpha_{0,0} + \theta_{01} \alpha_{0,1} + \theta_{10} \alpha_{1,0} + \theta_{11} \alpha_{1,1}$$

DC系数

AC 系数

基矩阵

例： 正交变换

已知：

$$f(x, y) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

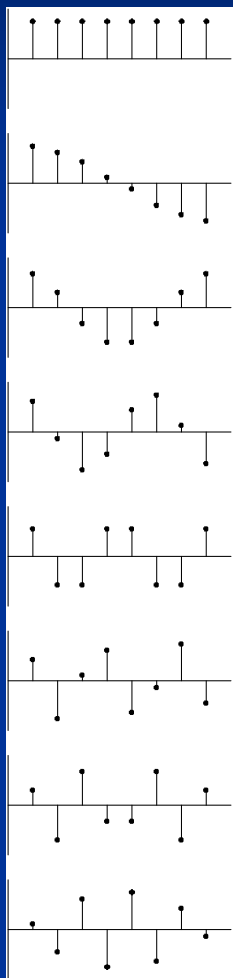
用矩阵算法求其变换后系数。

$$F(u, v) = CfC^T$$

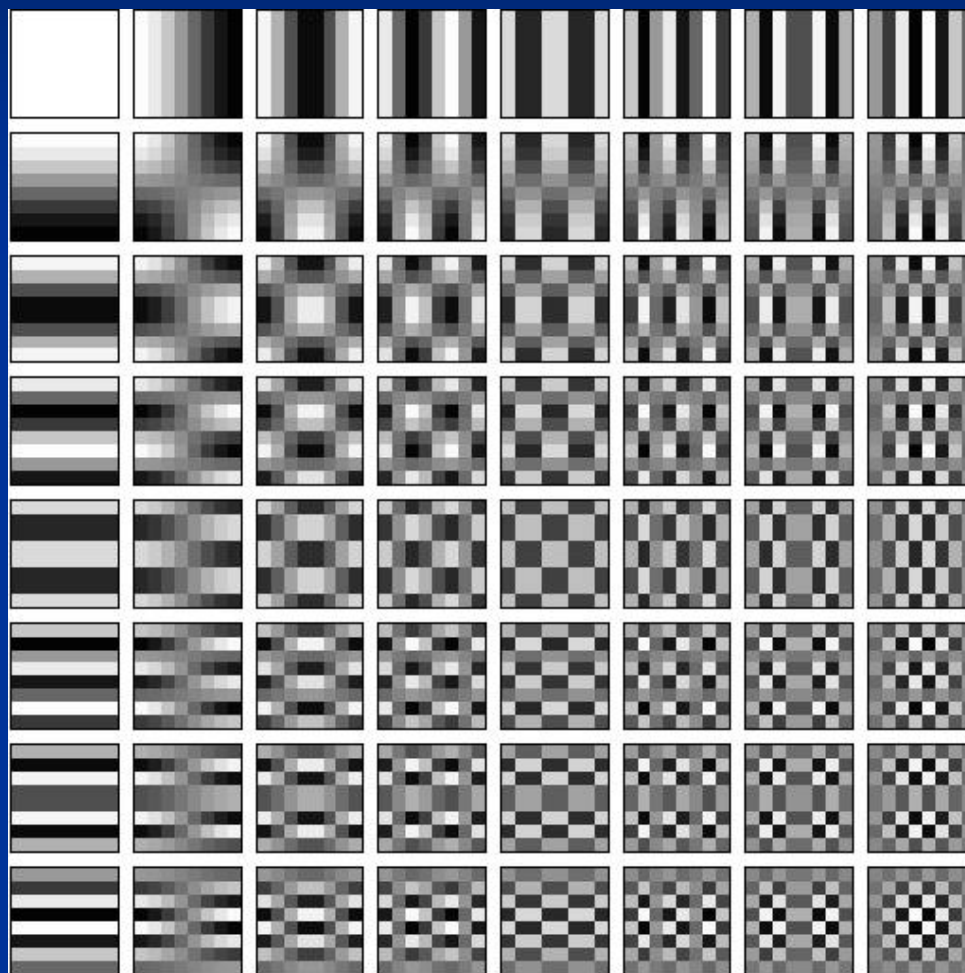
$$= \begin{bmatrix} 0.5 & 0.5 & 0.5 & 0.5 \\ 0.65 & 0.27 & -0.27 & -0.65 \\ 0.5 & -0.5 & -0.5 & 0.5 \\ 0.27 & -0.65 & 0.65 & -0.27 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.5 & 0.65 & 0.5 & 0.27 \\ 0.5 & 0.27 & -0.5 & -0.65 \\ 0.5 & -0.27 & -0.5 & 0.65 \\ 0.5 & -0.65 & 0.5 & -0.27 \end{bmatrix}$$

变换的核

1D变换核矩阵元素



2D变换核矩阵元素



Karhunen-Loève Transform (KLT)

- 亦称为Hotelling Transform
 - Hotelling于1933年用于离散数据去相关
 - Karhunen、Loève 分别于1947年和1948用于连续函数分析
 - Kramer和Mathews、Huang和Schultheiss分别于1956年、1963年用于数据压缩（变换编码）
- 在统计分析中被称为主成份分析
(Principal Components Analysis, PCA)

KLT正交变换的基本概念

- 对正交变换总的能量保持：

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

$$\|\mathbf{y}\|^2 = \mathbf{y}\mathbf{y}^T = \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} = \|\mathbf{x}\|^2$$

- 总的能量保持，但是通常能量在各系数上分布并不均匀
- X 的协方差矩阵：

$$\phi_X = E\left\{[X - E(X)][X - E(X)]^T\right\} = \begin{bmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1N} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2N} \\ \vdots & & & \\ \phi_{N1} & \phi_{N2} & \cdots & \phi_{NN} \end{bmatrix}$$

$$\text{其中, } \phi_{ij} = E\left\{[x_i - E(x_i)][x_j - E(x_j)]\right\} = \phi_{ji}$$

KLT正交变换的基本概念

■ Y 的协方差矩阵:

$$\begin{aligned}\phi_Y &= E\left\{[Y - E(Y)][Y - E(Y)]^T\right\} \\ &= E\left\{A[X - E(X)][X - E(X)]^T A^T\right\} \\ &= A\phi_X A^T\end{aligned}$$

为使得变换后系数各不相关，即要令 Y 的协方差矩阵中只存在对角线元素。

对于一个实对称矩阵 Φ ，必存在一个正交矩阵 Q ，使得

$$Q\Phi Q^{-1} = \text{diag}[\lambda_1 \lambda_2 \cdots \lambda_N] = \Lambda$$

式中, 对角阵 $\Lambda = \text{diag}[\lambda_1 \lambda_2 \cdots \lambda_N]$ 的 N 个对角元 $\lambda_1 \lambda_2 \cdots \lambda_N$ 是 ϕ 的特征根

Q 中的第 i 个行向量是 ϕ 的第 i 个特征根 λ_i 所对应的满足归一化正交条件的特征向量

$$\text{即: } \phi q_i = \lambda_i q_i, \quad q_i^T q_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

KLT正交变换的基本概念

- 只要选择正交矩阵 Q 作为变换核矩阵，其行向量是 X 的协方差矩阵的特征向量，则变换后的矢量 Y 的协方差矩阵为：

$$\begin{aligned}\phi_Y &= E\left\{[Y - E(Y)][Y - E(Y)]^T\right\} \\ &= E\left\{[QX - E(QX)][QX - E(QX)]^T\right\} \\ &= QE\left\{[X - E(X)][X - E(X)]^T\right\}Q^T \\ &= Q\phi_X Q^T = \Lambda\end{aligned}$$

则变换后各分量之间的相关性被全部去除

KLT正交变换的基本概念

- 以矢量 X 的协方差矩阵的归一化正交特征向量(q_i)所构成的正交矩阵 Q ，对该矢量所做的正交变换 $Y=QX$ 称作Karhunen-Loeve变换，即KL变换
- 即：要实现KLT，必须先知道输入信号的协方差矩阵，然后再求出协方差矩阵的特征值和特征向量。用特征向量构成该输入信号矢量的正交变换核矩阵。

A 为 n 阶方阵, 若存在数 λ 和 n 维非0列向量 x 使

$$Ax = \lambda x$$

则 λ 为 x 的特征值, x 为 A 对应于特征值 λ 的特征向量

方程组形式: $(A - \lambda I)x = 0$

如果特征值和特征向量存在, 则上述方程必有非0解 x , 即要求系数矩阵 $A - \lambda I$ 的秩 $< n$, 所以系数行列式 $|A - \lambda I| = 0$, 该式是关于 λ 的 n 次多项式, 即矩阵 A 的特征多项式

一个特征值可对应无穷多个不同的特征向量。某一特征向量只对应某一特征值。所以方阵 A 的不同特征值所对应的特征向量线性无关。

求法:

(1) $|A - \lambda I| = 0$ 的多项式, 求出特征值

(2) 把每个特征值代入 $(A - \lambda I)x = 0$, 求出齐次线性方程组的一个基础解系

的全部解 $\xi_1, \xi_2, \dots, \xi_s$, 则 $\sum_{i=1}^s k_i \xi_i$ 是 A 对应于特征值 λ 的全部特征向量, 其中

k_1, k_2, \dots, k_s 不全为0。

- 例：若已知随机信号 X 的协方差矩阵为求正交矩阵 Q 。

$$\phi_X = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

解：(1)求特征值

$$\text{令 } \begin{vmatrix} 1-\lambda & 1 & 0 \\ 1 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = 0, \text{按 } \lambda_1 > \lambda_2 > \lambda_3 \text{ 次序可解出: } \lambda_1 = 2, \lambda_2 = 1, \lambda_3 = 0$$

(2)求特征向量

将 $q_i = [q_{i1}, q_{i2}, q_{i3}]^T$ 代入 $\phi q_i = \lambda_i q_i$, 解3个方程组:

$$\text{由 } \begin{bmatrix} q_{11} + q_{12} \\ q_{11} + q_{12} \\ q_{13} \end{bmatrix} = 2 \begin{bmatrix} q_{11} \\ q_{12} \\ q_{13} \end{bmatrix}, \quad \text{得 } q_{11} = q_{12} = a, q_{13} = 0, \text{ 即 } q_1 = \begin{bmatrix} a \\ a \\ 0 \end{bmatrix}$$

$$\text{由} \begin{bmatrix} q_{21} + q_{22} \\ q_{21} + q_{22} \\ q_{23} \end{bmatrix} = 2 \begin{bmatrix} q_{21} \\ q_{22} \\ q_{23} \end{bmatrix}, \quad \text{得} q_{21} = q_{22} = 0, q_{23} = b, \text{即} q_2 = \begin{bmatrix} 0 \\ 0 \\ b \end{bmatrix}$$

$$\text{由} \begin{bmatrix} q_{31} + q_{32} \\ q_{31} + q_{32} \\ q_{33} \end{bmatrix} = 2 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \text{得} q_{31} = q_{32} = c, q_{33} = 0, \text{即} q_3 = \begin{bmatrix} c \\ -c \\ 0 \end{bmatrix}$$

- 待定实常数可由归一化正交条件解得：

$$a = \frac{\sqrt{2}}{2}, b = 1, c = \frac{\sqrt{2}}{2}$$

$$Q = [q_1 \quad q_2 \quad q_3]^T = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -1 & 0 \end{bmatrix}$$

- 得到归一化正交矩阵：

$$Q\phi_X Q^T = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -1 & 0 \end{bmatrix}^T = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

KLT的性质(2)

- 任何正交变换的行列式的值: $\det(\mathbf{A}) = \pm 1$
- 任何正交变换的协方差的行列式的值:

$$\det(\mathbf{R}_{\theta\theta}) = \det(\mathbf{A}) \det(\mathbf{R}_{xx}) \det(\mathbf{A}^T) = \det(\mathbf{R}_{xx})$$

- KLT变换后（对角）协方差矩阵的行列式

$$\det(\mathbf{R}_{\theta\theta}) = \prod_{k=0}^{N-1} \sigma_{\theta_k}^2$$

- Hadamard不等式: 任何对称、半正定矩阵的行列式小于等于其对角线元素的乘积

$$\prod_{k=0}^{N-1} \sigma_{\theta_k}^2 (\text{KLT}) = \det(\mathbf{R}_{\theta\theta}) \leq \prod_{k=0}^{N-1} \sigma_{\theta_k}^2 (\mathbf{A})$$

- 变换系数方差的几何均值最小 \rightarrow 编码增益 G_{TC} 最大

KLT(3)

- 既然KLT是最佳的，为什么还会有其他的变换呢？
 - 没有快速计算方法
 - KLT取决于信号的统计性质
 - 对平稳过程，KLT比较合适
 - 但对大多数输入，需重复计算/传输变换矩阵
- →寻找结构化的变换，使得其性能接近KLT
 - DCT

离散余弦变换

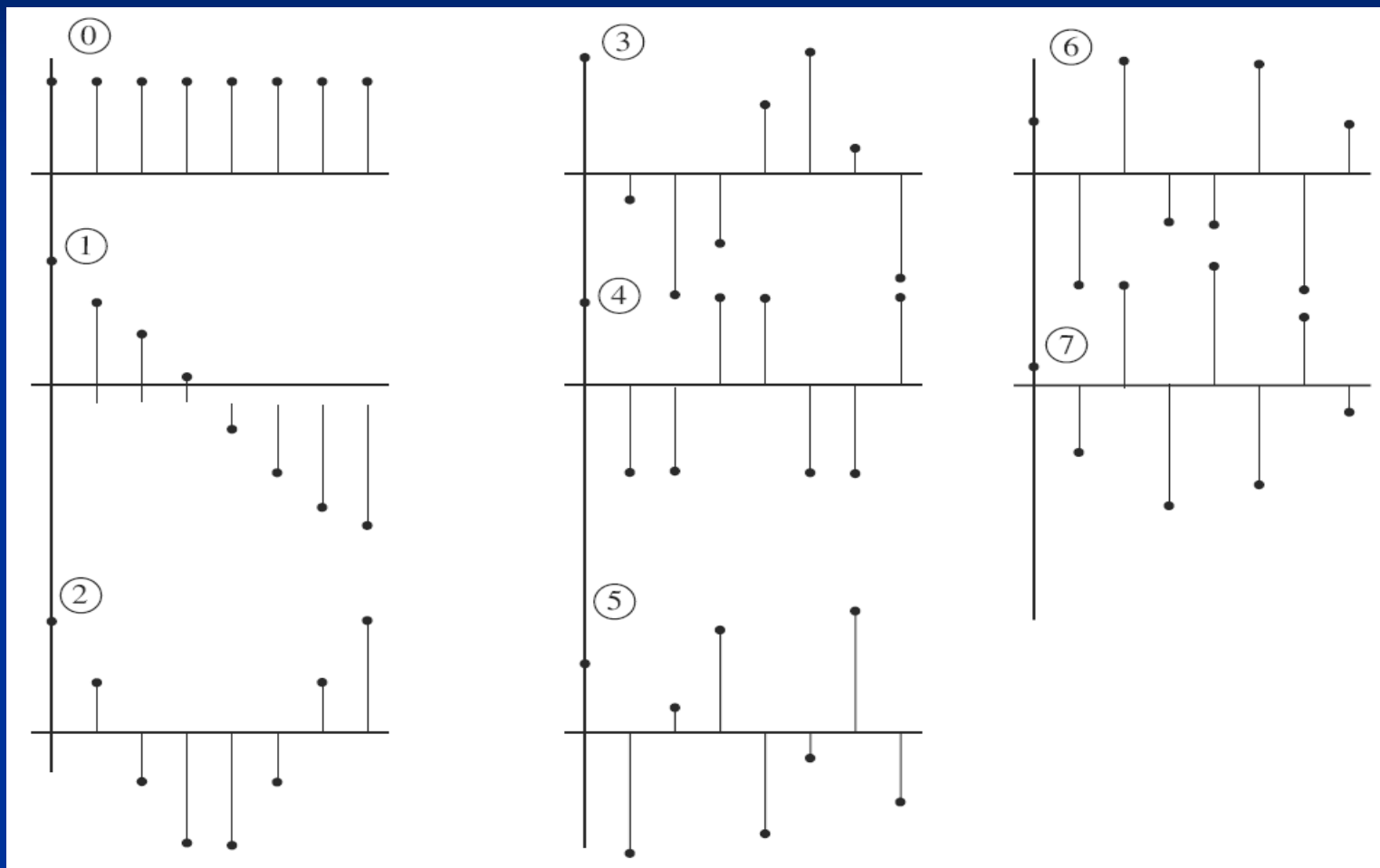
(Discrete Cosine Transform, DCT)

- 变换矩阵为cosine的函数

$$[C]_{i,j} = \begin{cases} \sqrt{\frac{1}{N}} \cos \frac{(2j+1)i\pi}{2N} & i = 0, j = 0, 1, \dots, N-1 \\ \sqrt{\frac{2}{N}} \cos \frac{(2j+1)i\pi}{2N} & i = 1, 2, \dots, N-1, j = 0, 1, \dots, N-1 \end{cases}$$

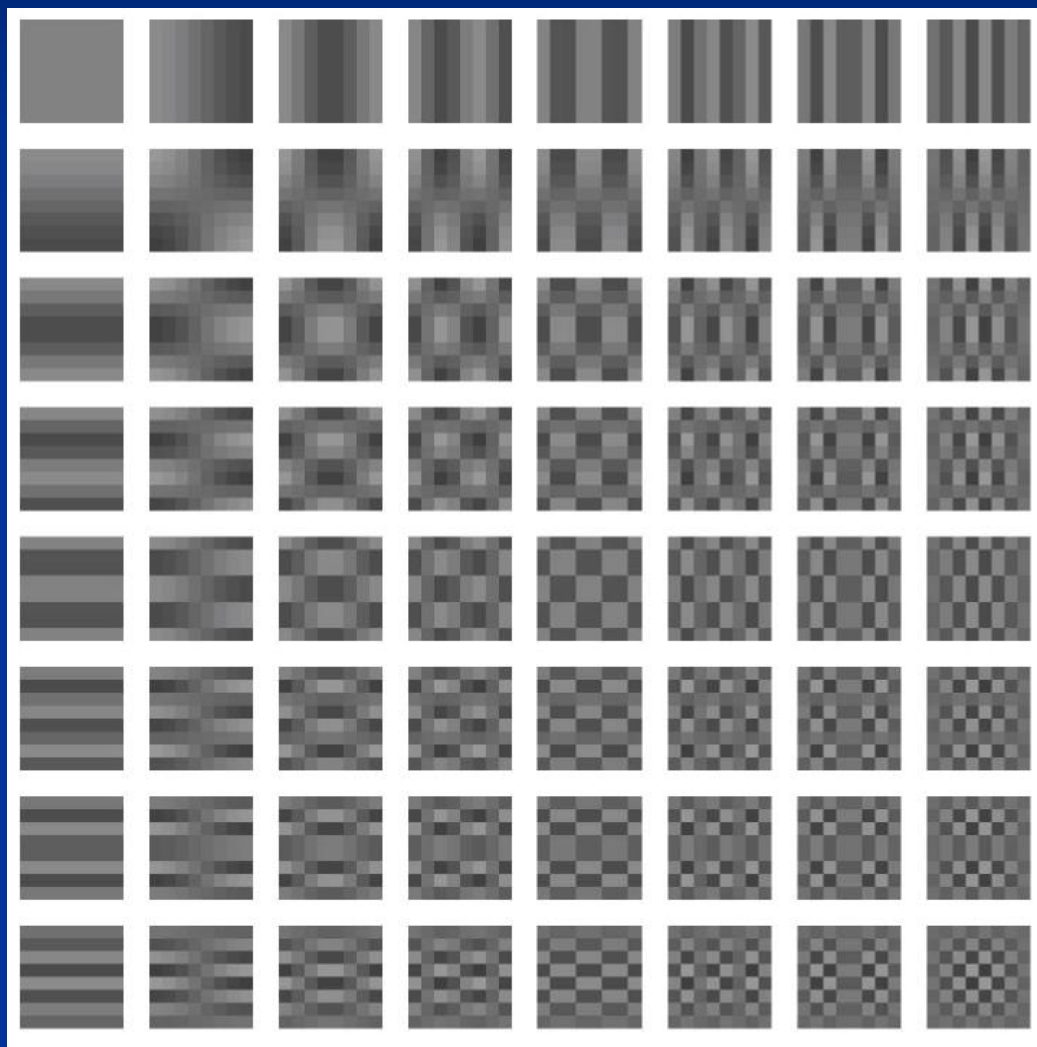
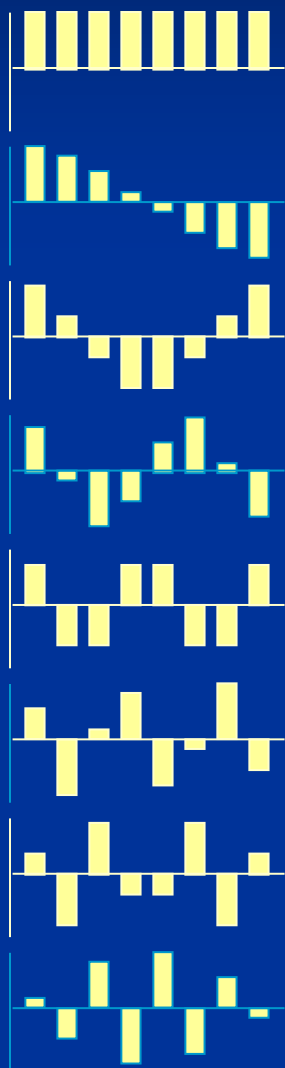
- 源自DFT
 - 但更适合压缩

DCT基向量



变化随行号增大而增大（频率越来越高）

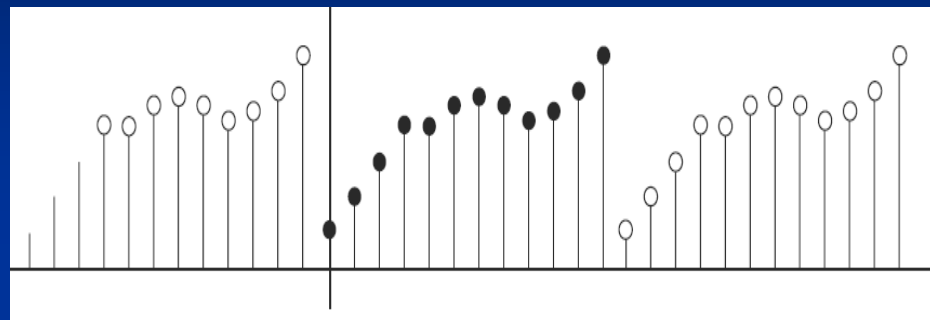
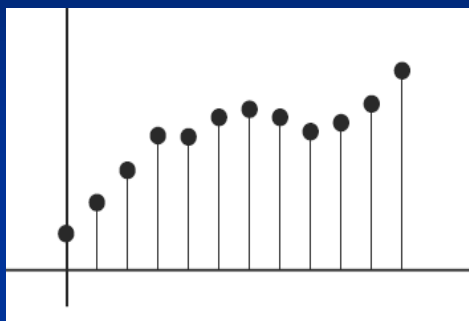
DCT基矩阵



DFT vs. DCT

DFT在边界不连续

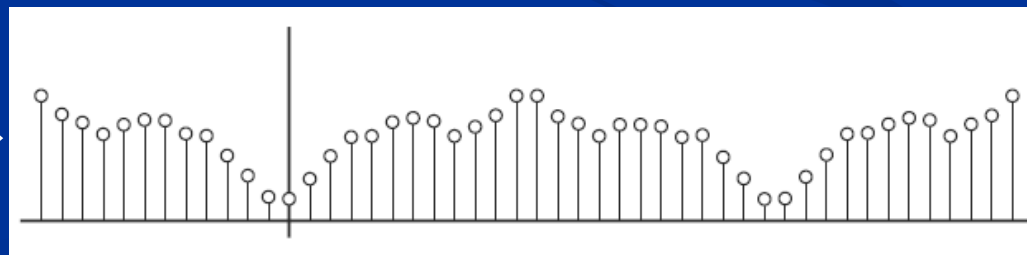
DFT:



DCT:



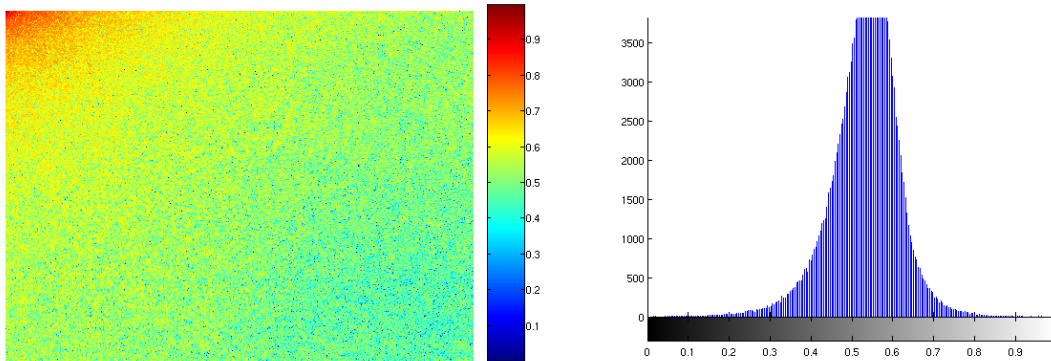
DCT更连续



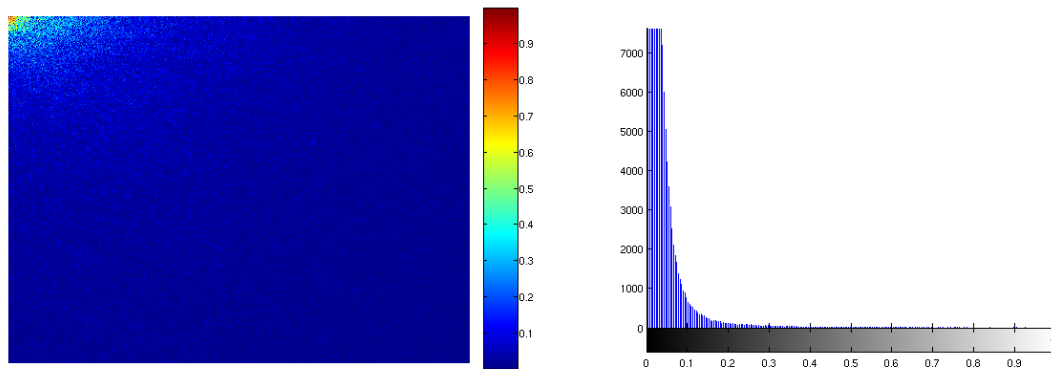
函数的不连续影响**Fourier**级数的收敛，从而需要更多基函数，影响压缩



DFT



DCT



T(2)

DCT变换后的能量更集中
更适合压缩

DCT的性质

- 对Markov过程:

$$\rho = \frac{E[x_n x_{n+1}]}{E[x_n^2]}$$

- 当 ρ 很大时, DCT的性能接近KLT的性能
- 实际上, 很多信源都是 ρ 很大的Markov过程
- \rightarrow DCT是一个最流行的变换
 - JPEG
 - MPEG
 - H.261
 - ...

离散正弦变换

(Discrete Sine Transform, DST)

- 变换矩阵为sine的函数：

$$[S]_{i,j} = \sqrt{\frac{2}{N+1}} \sin \frac{\pi(i+1)(j+1)}{N+1}, \quad i, j = 0, 1, \dots, N-1$$

- 类似DCT，有很好的性质
 - 当 ρ 很小时，DST的性能接近KLT的性能
 - 在图像/语音编码应用中，与DCT变换互补

离散Discrete Walsh-Hadamard变换

- N 阶Hadamard 矩阵

$$\mathbf{H}\mathbf{H}^T = N\mathbf{I}$$

- $N = 2^k$ 的构造规则:

$$\mathbf{H}_1 = [1], \quad \mathbf{H}_{2N} = \begin{bmatrix} \mathbf{H}_N & \mathbf{H}_N \\ \mathbf{H}_N & -\mathbf{H}_N \end{bmatrix}$$

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$\mathbf{H}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

...

DWH变换

- 从Hadamard 矩阵 \mathbf{H}_N 推导得到变换矩阵 \mathbf{H}
 - 归一：乘以 $1/\sqrt{N}$
 - 将每行以Wash序（*sequency order*）排列

■ 如：

$$\mathbf{H} = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}$$

■ 性能：

- 计算简单：+/-，在有限硬件上容易实现
- 但压缩性能比采用DCT稍差，用于H.264亮度和色度DC系数的变换

各种变换的比较 (1)

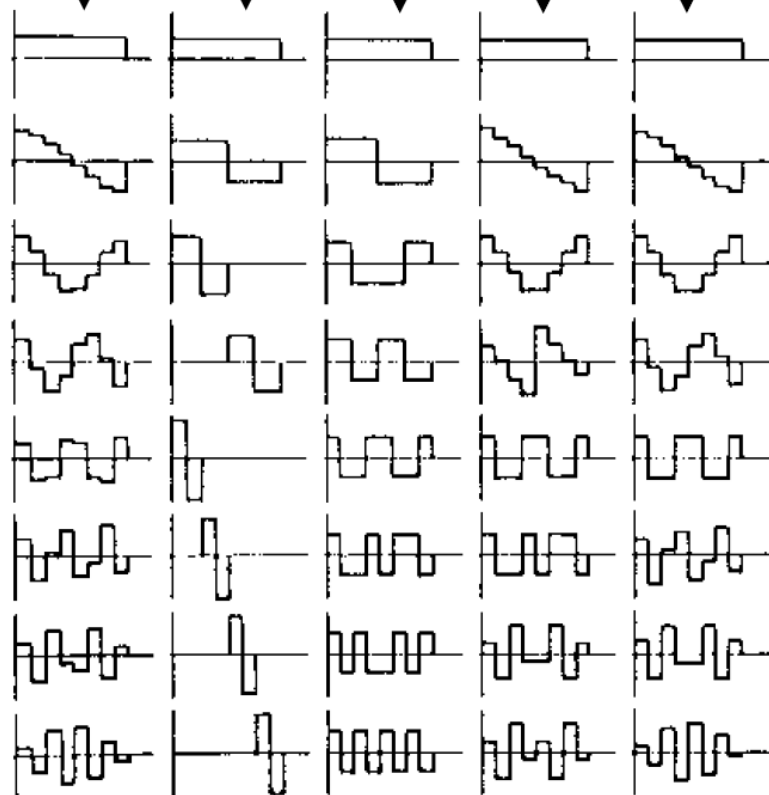
Karhunen Loève transform [1948/1960]

Haar transform [1910]

Walsh-Hadamard transform [1923]

Slant transform [Enomoto, Shibata, 1971]

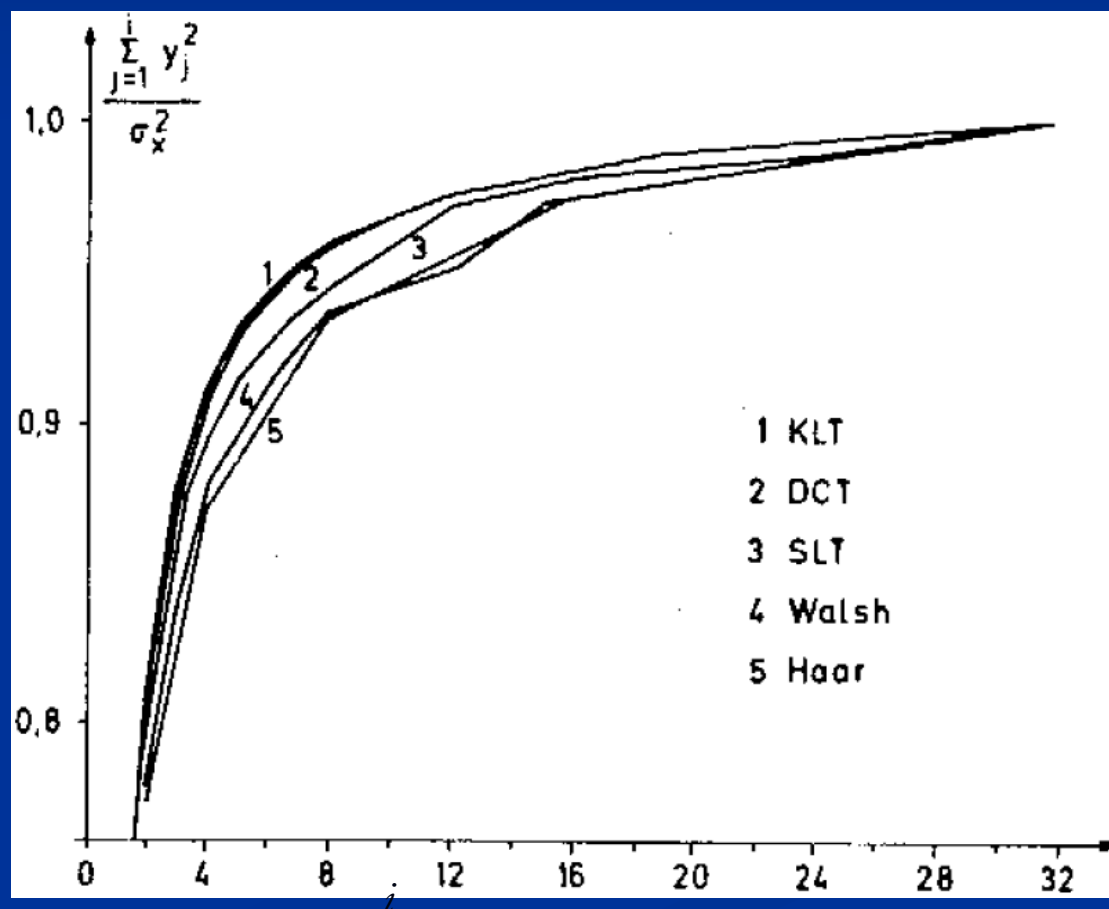
Discrete Cosine Transform (DCT)
[Ahmet, Natarajan, Rao, 1974]



Comparison of 1-d
basis functions for
block size $N=8$

各种变换的比较 (2)

- 对块大小为 1×32 的典型自然图像，各种变换的能量集中比较：



主要内容

- Karhunen-Loève变换(KLT)
- 离散余弦变换(discrete cosine transform, DCT)
- 变换系数的量化编码
- JPEG

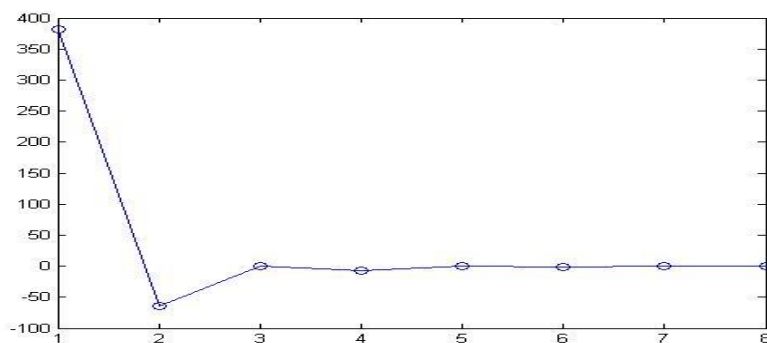
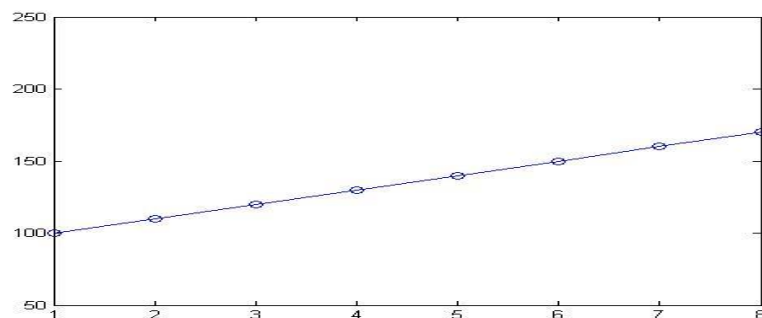
例1：空域量化 vs. 变换域量化

- $x = [100 \ 110 \ 120 \ 130 \ 140 \ 150 \ 160 \ 170]^T$

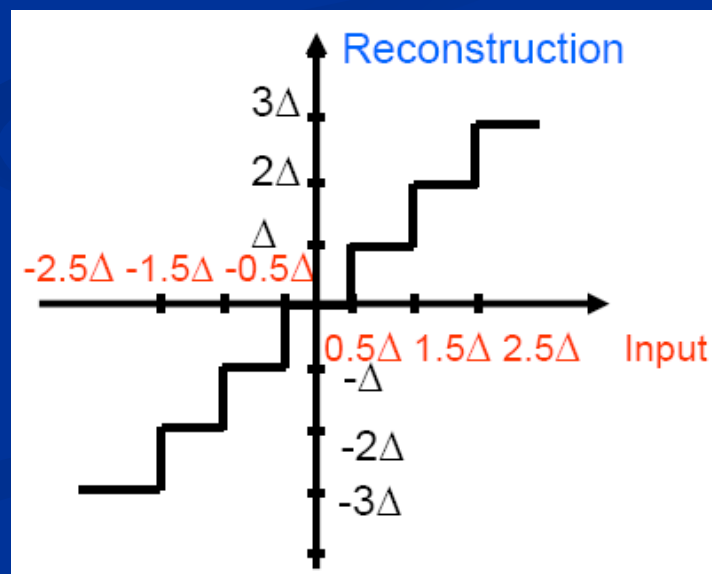
- 8点DCT变换：

$$y = [381.84 \ -64.42 \ 0 \ -6.73 \ 0 \ -2.01 \ 0 \ -0.5070]$$

能量主要集中在前两个系数



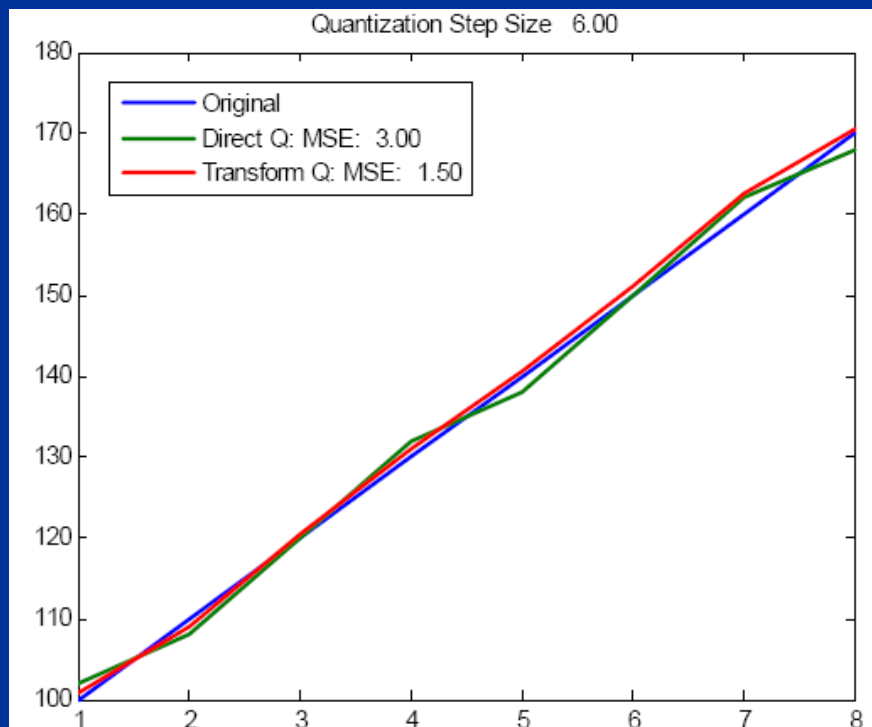
7水平的中平量化器



例1：空域量化 vs. 变换域量化(2)

- 方案1：直接对原始数据进行量化
- 方案2：对DCT系数进行量化
 - $\Delta=6$ ，量化后的DCT系数：[64 -11 0 -1 0 0 0 0]
 - 3个非0 DCT系数

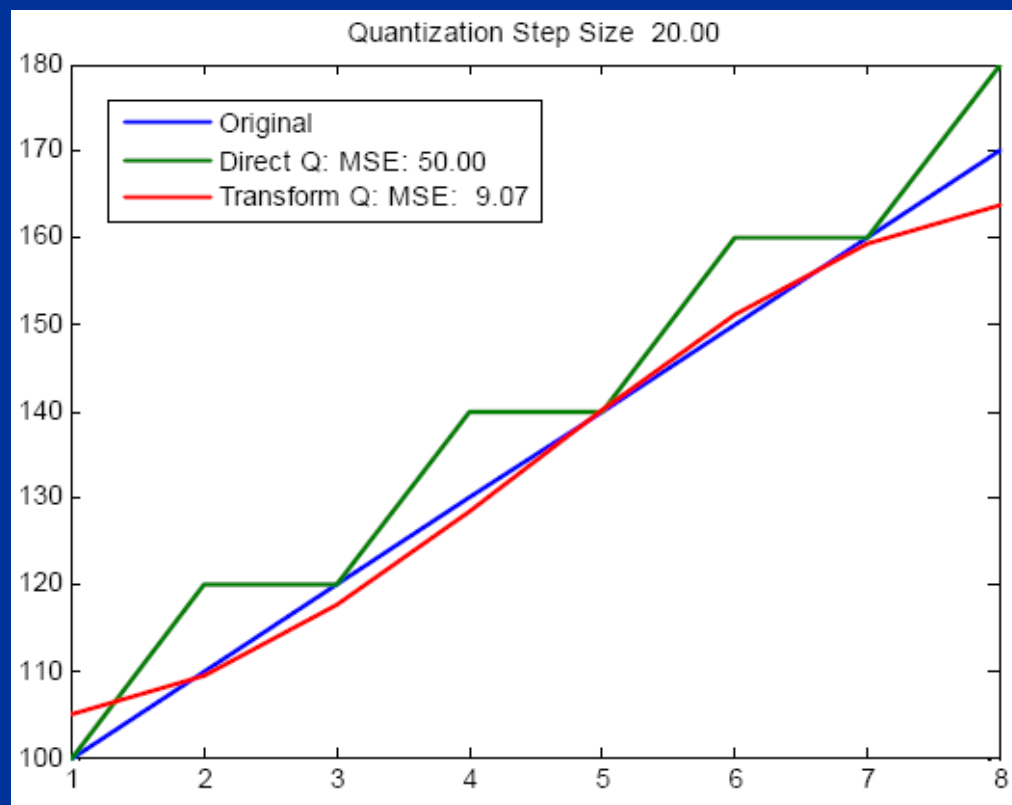
MSE:
w/o DCT: 3.0
w/ DCT: 1.5



例1：空域量化 vs. 变换域量化(3)

- $\Delta=20$, 2个非0 DCT系数: [19 -3 0 0 0 0 0 0]
 - DCT系数重构效果仍然很平滑
- 直接方法开始产生块/mosaic效应

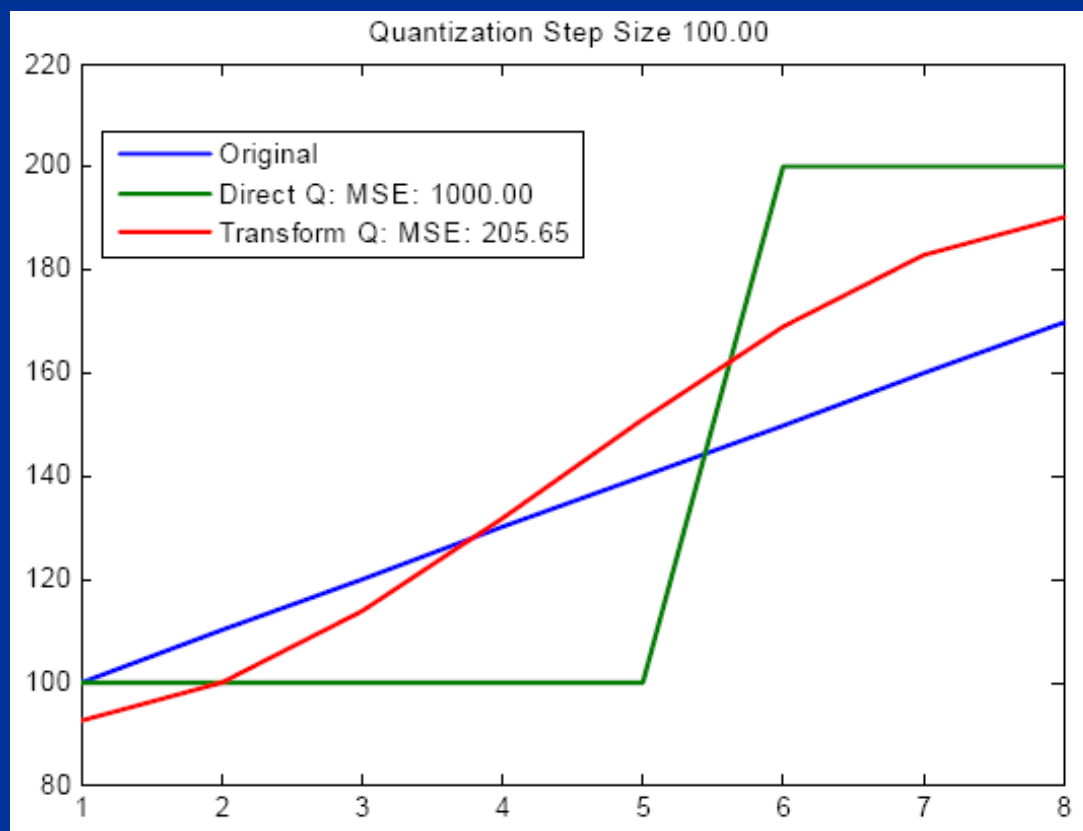
MSE:
w/o DCT: 50.0
w/ DCT: 9.07



例1：空域量化 vs. 变换域量化(4)

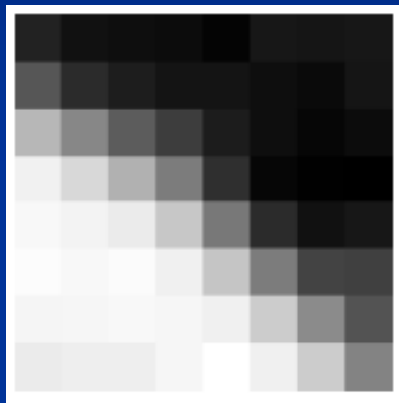
- $\Delta=100$, 2个非0 DCT系数: [4 -1 0 0 0 0 0 0]
 - DCT系数重构效果仍然平滑
- 直接方法产生的块/mosaic效应更多

MSE:
w/o DCT: 1000
w/ DCT: 205



例2：空域量化 vs. 变换域量化

■ 输入数据：



89	78	76	75	70	82	81	82
122	95	86	80	80	76	74	81
184	153	126	106	85	76	71	75
221	205	180	146	97	71	68	67
225	222	217	194	144	95	78	82
228	225	227	220	193	146	110	108
223	224	225	224	220	197	156	120
217	219	219	224	230	220	197	151

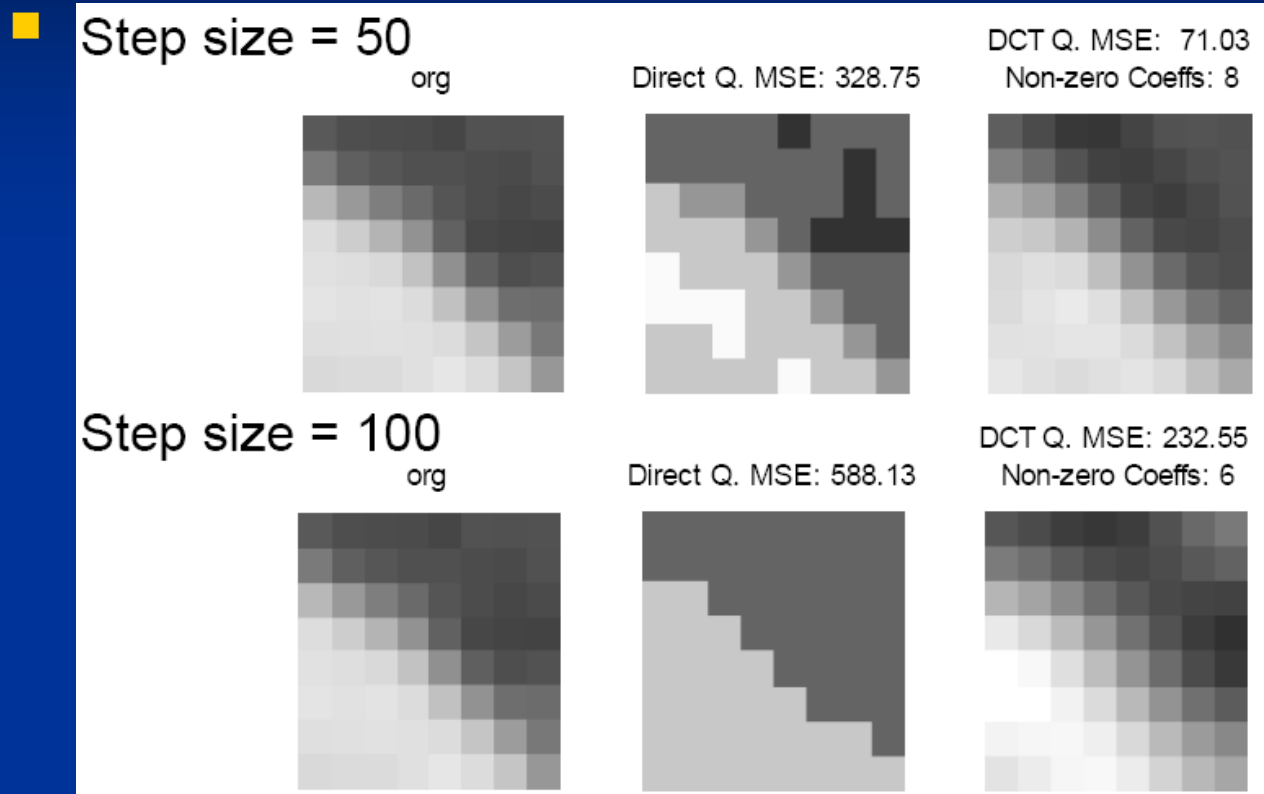
■ 2-D DCT变换系数（取整）：



1155	259	-23	6	11	7	3	0
-377	-50	85	-10	10	4	7	-3
-4	-158	-24	42	-15	1	0	1
-2	3	-34	-19	9	-5	4	-1
1	9	6	-15	-10	6	-5	-1
3	13	3	6	-9	2	0	-3
8	-2	4	-1	3	-1	0	-2
2	0	-3	2	-2	0	0	-1

大多数能量集中在左上角

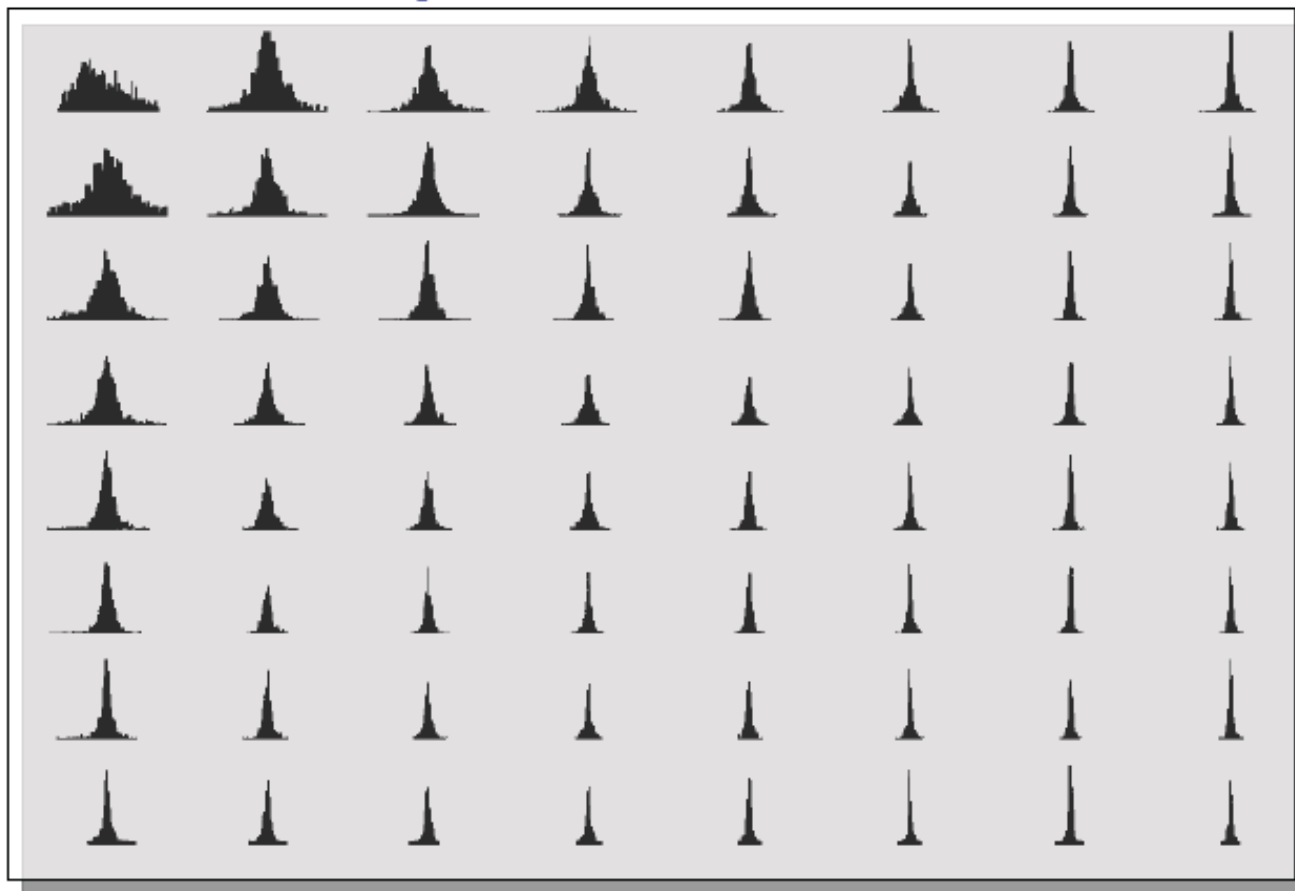
例2：空域量化 vs. 变换域量化(2)



- 在变换域量化通常能得到更好的结果
- 我们还可以做得更好
 - 对不同的子带采取不同的量化步长

DCT系数的分布

Histograms for 8x8 DCT coefficient amplitudes measured for test image
[Lam, Goodman, 2000]



Test image
Bridge

AC coefficients: Laplacian PDF

DC coefficient distribution similar to the original image

变换系数的比特分配

- 我们讨论的变换为正交变换
 - 变换过程为能量守恒，所以总的误差=量化误差
- 两种比特分配方法
 - 变换系数的平均特性
 - 根据各个变换系数的需要分配比特数
- 每个变换系数的能量（输入方差）： $\sigma_{\theta_k}^2$
- 每个变换系数的量化误差： $D_k(R_k) = \varepsilon^2 \sigma_{\theta_k}^2 2^{-2R_k}$
- 总的量化误差： $D^T(R) \cong \sum_{k=0}^{N-1} D_k(R_k) = \sum_{k=0}^{N-1} \varepsilon^2 \sigma_{\theta_k}^2 2^{-2R_k}$

变换系数的比特分配(2)

■ 比特分配问题:

- 计算 $\{R_k\}$ 使得 $D^T(R) = \sum_{k=0}^{N-1} \varepsilon^2 \sigma_{\theta_k}^2 2^{-2R_k}$ 最小
- 并满足比特率: $R = \frac{1}{N} \sum_{k=0}^{N-1} R_k$

■ 变换系数能量:

$$\theta = \mathbf{A} \mathbf{x}$$

$$\mathbf{R}_{\theta\theta} = E[\theta\theta^T] = E[\mathbf{A} \mathbf{x} \mathbf{x}^T \mathbf{A}^T] = \mathbf{A} \mathbf{R}_{\mathbf{x}\mathbf{x}} \mathbf{A}^T$$

- $\sigma_{\theta_k}^2$ 为 $\mathbf{R}_{\theta\theta}$ 对角线上第 k 个元素

变换系数的比特分配(3)

- 用Lagrangian乘子法得到最佳的比特分配

$$J = D^T(R) + \lambda \left(R - \frac{1}{N} \sum_{k=0}^{N-1} R_k \right)$$
$$= \sum_{k=0}^{N-1} \varepsilon^2 \sigma_{\theta_k}^2 2^{-2R_k} + \lambda \left(R - \frac{1}{N} \sum_{k=0}^{N-1} R_k \right)$$

- 对所有的 k , $\frac{\partial J}{\partial R_k} = 0$

$$\frac{\partial J}{\partial R_k} = 2(\ln 2) \varepsilon^2 2^{-2R_k} \sigma_{\theta_k}^2 - \frac{\lambda}{N} = 0$$

$$\Rightarrow D_k(R_k) = \varepsilon^2 2^{-2R_k} \sigma_{\theta_k}^2 = \frac{\lambda}{2(\ln 2) N} = \lambda_0$$

- \rightarrow 每个变换系数的量化误差的方差尽可能相等

变换系数的比特分配(4)

$$\blacksquare \varepsilon^2 2^{-2R_k} \sigma_{\theta_k}^2 = \frac{\lambda}{2(\ln 2)N} = \lambda_0$$

$$\Rightarrow R_k = \frac{1}{2} \log_2 \frac{\varepsilon^2 \sigma_{\theta_k}^2}{\lambda_0}$$

■ →方差越大，比特数越多

■ 代入比特率约束

$$R = \frac{1}{N} \sum_{k=0}^{N-1} R_k = \frac{1}{N} \sum_{k=0}^{N-1} \frac{1}{2} \log_2 \frac{\varepsilon^2 \sigma_{\theta_k}^2}{\lambda_0} \Rightarrow \lambda_0 = \varepsilon^2 2^{-2R} \left(\prod_{k=1}^N \sigma_{\theta_k}^2 \right)^{\frac{1}{N}}$$

■ 最佳失真和每个系数的码率分别为

$$R_k = R + \frac{1}{2} \log_2 \left[\left(\sigma_{\theta_k}^2 \right) / \left(\prod_{k=1}^N \sigma_{\theta_k}^2 \right)^{\frac{1}{N}} \right], D_k(R_k) = \varepsilon^2 \sigma_{\theta_k}^2 2^{-2R_k}$$

变换系数的比特分配(5)

- 变换编码的最佳失真为

$$D^T(R) = \sum_{k=0}^{N-1} D_k(R_k) = N\lambda_0 = N\varepsilon^2 2^{-2R} \left(\prod_{k=0}^{N-1} \sigma_{\theta_k}^2 \right)^{\frac{1}{N}}$$

- 假设对原始信号的码率失真函数为

$$D(R) \cong \varepsilon^2 \sigma_x^2 2^{-2R}$$

- 则变换编码增益为

$$G_T = 10 \log_{10} \frac{D(R)}{D^T(R)} = 10 \log_{10} \frac{\varepsilon^2 \sigma_x^2 2^{-2R}}{N \varepsilon^2 2^{-2R} \left(\prod_{k=0}^{N-1} \sigma_{\theta_k}^2 \right)^{\frac{1}{N}}} = 10 \log_{10} \frac{\sigma_x^2}{N \left(\prod_{k=0}^{N-1} \sigma_{\theta_k}^2 \right)^{\frac{1}{N}}}$$

变换系数的比特分配(6)

■ 变换编码增益为

$$G_T = 10 \log_{10} \frac{\sigma_x^2}{N \left(\prod_{k=0}^{N-1} \sigma_{\theta_k}^2 \right)^{\frac{1}{N}}} = \frac{\frac{1}{N} \sum_{k=0}^{N-1} \sigma_{\theta_k}^2}{\sqrt[N]{\prod_{k=0}^{N-1} \sigma_{\theta_k}^2}}$$

算术均值

几何均值

- σ_x^2 为 R_{xx} 对角线的元素，对平稳过程， R_{xx} 每个 (i, i) 相等
- 增益与系数方差的集中程度有关
- 若每个系数的方差相等，则没有增益

变换系数的比特分配(7)

- 上述最佳 R_k 不一定为整数，甚至不能保证为正数
 - $R_k < 0 \Rightarrow R_k = 0$
 - 但增大了平均码率，还需均匀减小非0 R_k
- 上述推导的过程中，有两个重要结论
 - ϵ 因子取决于输入分布和量化器，如果假定该因子对于所有系数和量化器都相等，则每个量化器对总重构误差方差的贡献相同
 - R_k 值每增加1比特，重构误差的标准差将降低一半

$$D_k(R_k) = \epsilon^2 \sigma_{\theta k}^2 2^{-2R_k}$$

递归比特分配

- 满足约束： $R_k \geq 0$ 且 R_k 为整数
- 所以码率分配算法为：
 - 1、计算每个成分的方差 $\sigma_{\theta_k}^2$
 - 2、对所有的 $k = 0, 1, \dots, N - 1, R_k = 0$
 - 3、对所有的方差 $\sigma_{\theta_k}^2$ 排序，对最大的方差的成分 k' 分配1比特
$$R_{k'} = R_{k'} + 1, \sigma_{\theta_{k'}}^2 = 1/4 \sigma_{\theta_{k'}}^2$$
 - 若比特数用尽，停止；否则转第3步
- 上述算法称为zonal sampling
 - 最终为较高方差的系数分配更多的比特
 - 希望使每个系数导致的失真量相等

递归比特分配(2)

8*8 变换的比特分配

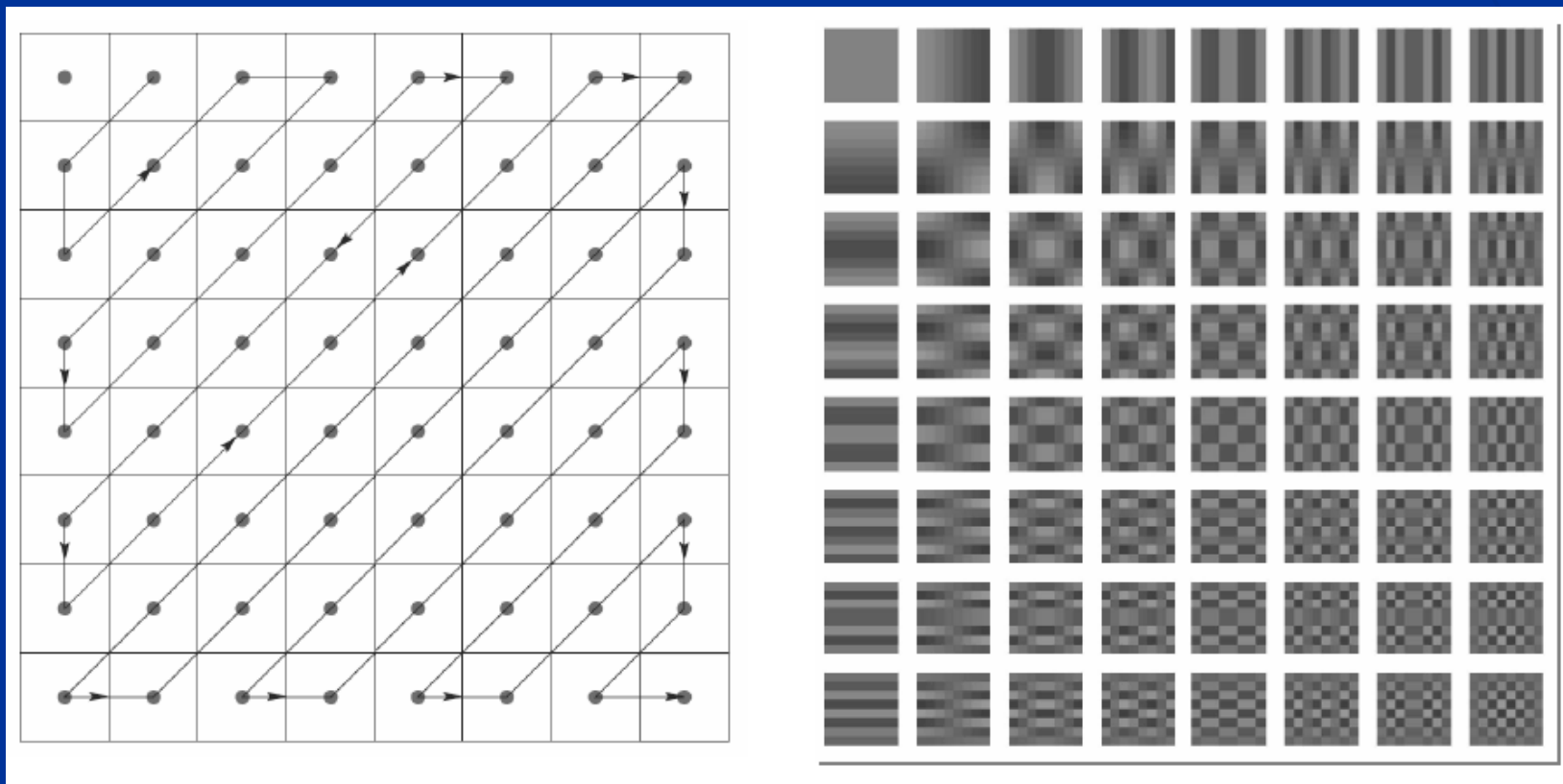
8	7	5	3	1	1	0	0
7	5	3	2	1	0	0	0
4	3	2	1	1	0	0	0
3	3	2	1	1	0	0	0
2	1	1	1	0	0	0	0
1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

阈值编码

- zonal sampling
 - 重构误差方差可由公式获得
 - ϵ 因子取决于输入分布和量化器，且假定该因子对于所有系数和量化器都相等，理想化条件！
 - 基于平均值进行比特分配，无法正确重构局部变化
 - 简单背景中具有明显边缘的物体
 - 边缘像素数目相对于总像素数很少
- 阈值编码：对所有大于阈值的系数进行编码，而丢弃其他较小的系数

阈值编码(2)

- 对2-D图像编码可采用Zigzag方式编码
 - 可得到很多高频系数为0，发送EOB（End of Block）即可



总结

- 正交变换：对信号空间的坐标进行旋转
- 变换的目的：
 - **去相关**：使得在量化时，对各系数单独量化（标量量化），而不会损伤过多效率（与矢量量化相比）
→ 变换本身并不会产生压缩，但变换会使得量化更容易实现，真正实现压缩的是变换后的量化
 - **能量集中**：稀疏，将原始信号的能量压缩到尽可能少的系数→对原始信号只用少数幅值较大的系数表示
- 比特分配与方差的log成正比
- KLT是最佳的，但其是信号相关的，没有快速算法
- DCT性能接近KLT，有快速计算算法
 - 通常采用DCT：对图像采用8x8的块大小、均匀量化、Z字扫描（JPEG, MPEG, ITU-T H.261, H.263...）

下节课内容

- 下节课内容：JPEG
- 作业： Sayood 3rd, pp.421-422
 - 2, 4