第四章　　量化

主要内容

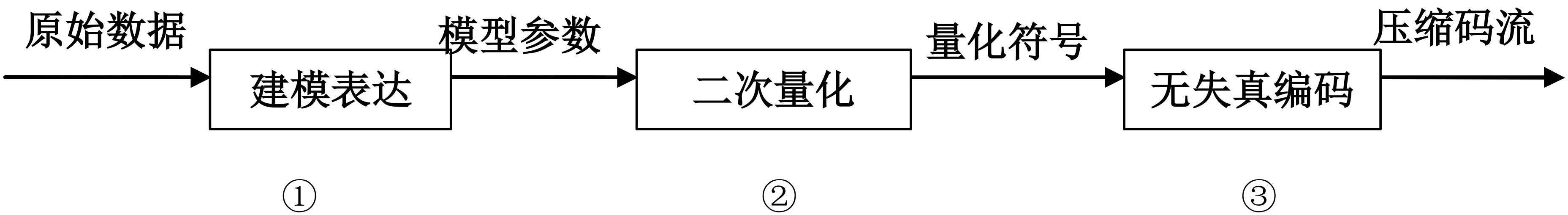
- 4.1 标量量化
  - 4.1.1 量化器的描述
  - 4.1.2 均匀量化
  - 4.1.3 Lloyd-Max算法
  - 4.1.4 熵约束量化*
  - 4.1.5 Deadzone Midtread 量化器
  - 4.1.6 嵌入式量化器
- 4.2 矢量量化
  - 4.2.1 矢量量化的基本思想
  - 4.2.2 LBG算法

4.1 标量量化

- **4.1.1 量化器的描述**
- **4.1.2 均匀量化**
- **4.1.3 Lloyd-Max算法**
- **4.1.4 熵约束量化***
- **4.1.5 Deadzone Midtread 量化器**
- **4.1.6 嵌入式量化器**

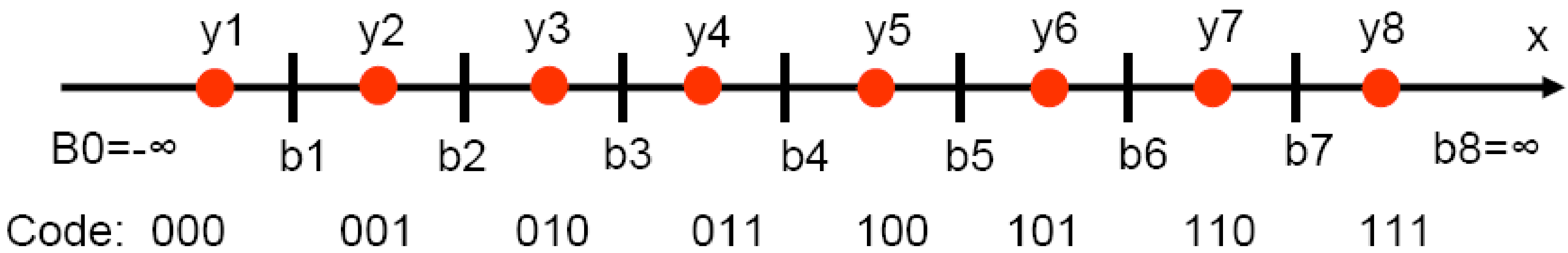
4.1.1 量化器的描述

- 量化：用一个很小的集合表示一个大集合（可能是无限大）的值
  - **如A/D转换**
- 量化是有失真压缩的一个有效工具



  - 对模拟信号，量化还包括A/D转换中的一次量化

4.1.1 量化器的描述



- 将实数线分成M个不相连的区间

$$I_i = [b_i, b_{i+1}), i = 0,1,\ldots,M-1$$

$$b_0 < b_1 < \ldots < b_M$$

I_i： 量化区间（bin）

i： 量化区间的索引

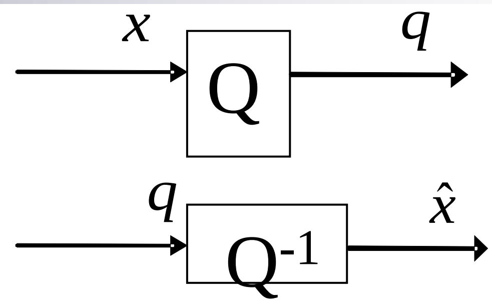
b_i： 决策边界

y_i： 重构（重建）水平

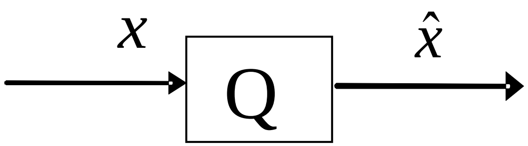
- 编码器将每个区间/bin的**索引**发给解码器
- 解码器用重构水平表示该区间内所有的值

4.1.1 量化器的描述

有时使用下列表示：

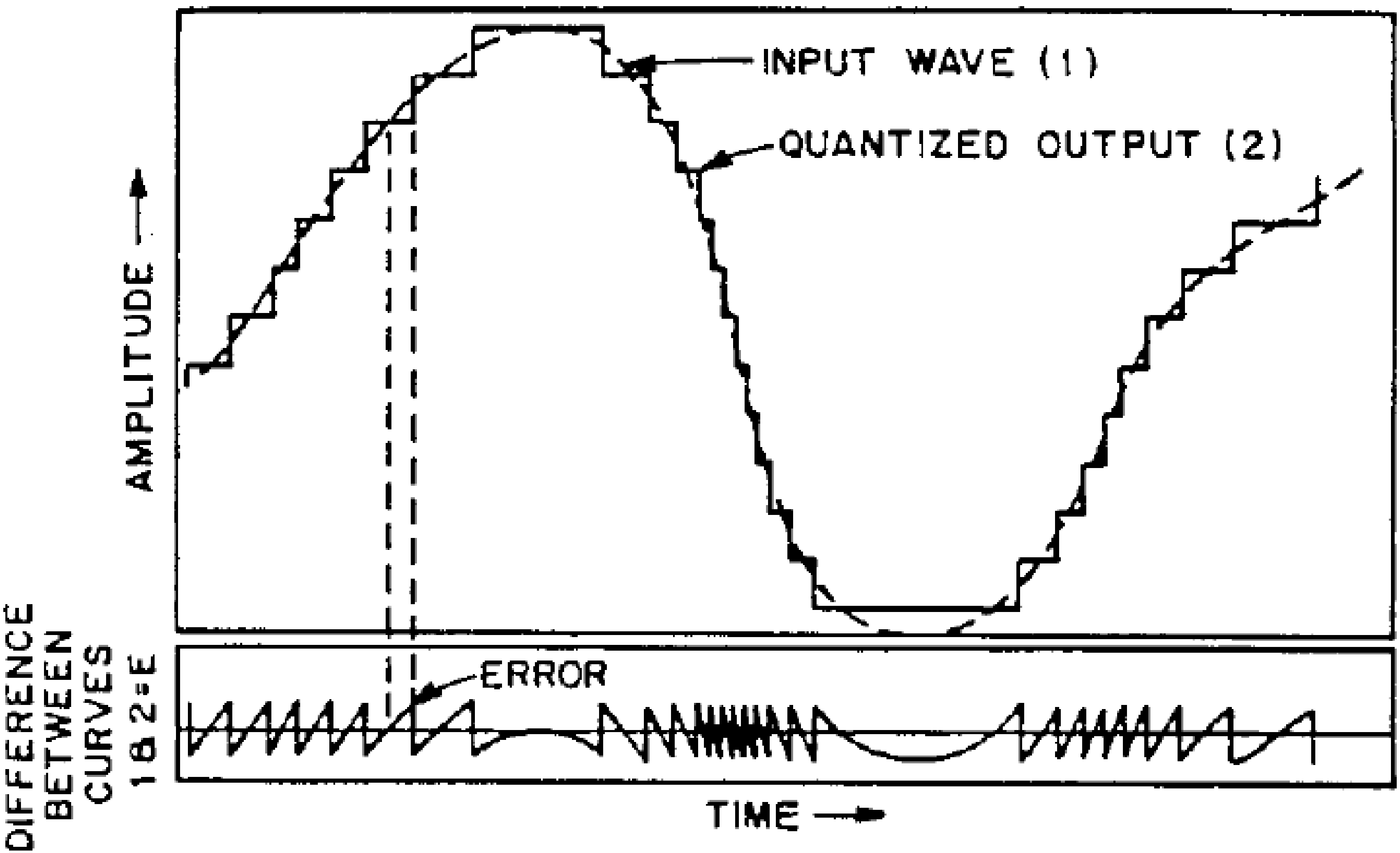


- 标量量化器的输入输出：



- 量化：$q = A(x)$， 将输入 x 用索引 q 表示
- 反量化：$\hat{x} = B(q)$， 将索引 q 映射为输入重构
  - 通常 $B(x)$ 不是 $A(x)$ 的反函数，$\hat{x} \neq x$
- 量化误差： $e(x) = x - \hat{x}$

4.1.1 量化器的描述

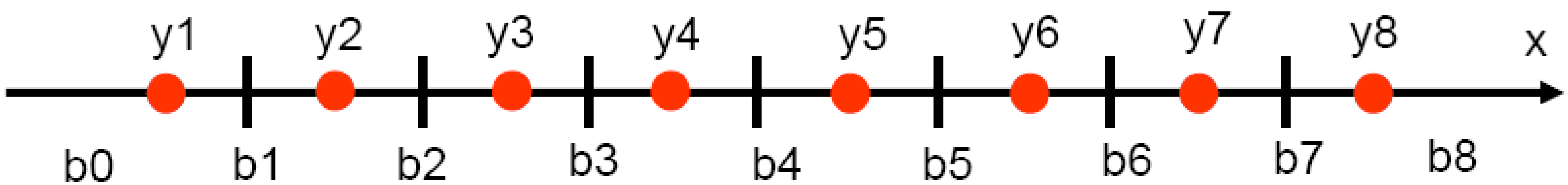


例：量化后的波形图

4.1.1 量化器的描述

- 失真的度量 - 量化误差 $e(x)=x-\hat{x}$

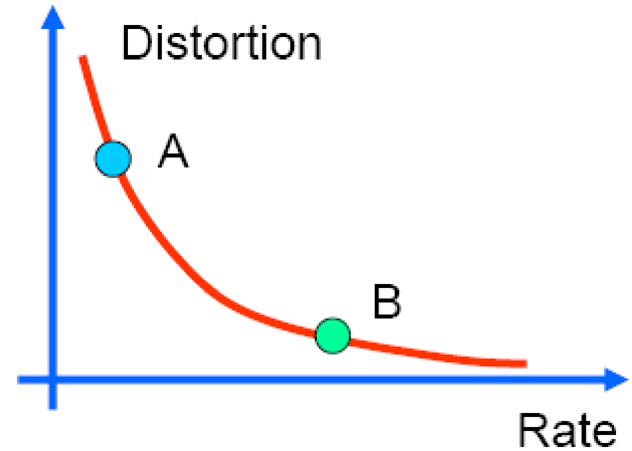
  量化的均方误差（Mean Squared Error, MSE）
  - 所有输入值的平均量化误差
  - 需要知道输入的概率分布



- 量化区间的数目：*M*
- 决策边界： $b_i, i=0,1,\ldots,M$
- 重构水平： $y_i, i=0,1,\ldots,M$
- 重构： $\hat{x}=y_i,\ if\ b_{i-1}<x\le b_i$
- 失真： $$D=MSE=\int_{-\infty}^{\infty}(x-\hat{x})^2 f(x)dx=\sum_{i=1}^{M}\int_{b_{i-1}}^{b_i}(x-y_i)^2 f(x)dx$$

4.1.1 量化器的描述

码率—失真折中



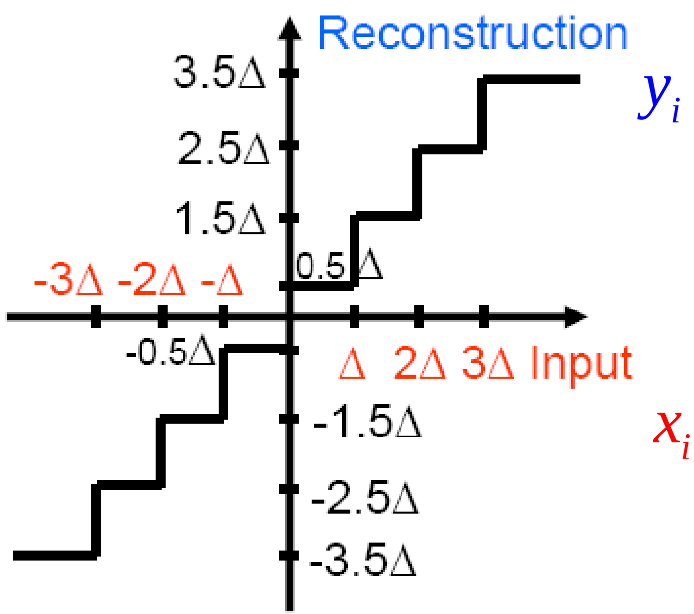
- 需要决定的参数
  - 量化区间的数目
  - 决策边界（判决门限）
  - 重构水平
  - 量化区间索引的码字（量化码字）
- 量化器的设计是码率与失真之间的折中
  - 为了降低编码的比特数，需要减低量化区间的数目→更大的误差
- 性能受率失真理论控制
  - 给定允许失真，求最小码率的量化器
  - 给定码率，求最小失真的量化器

4.1.2 均匀量化（Midrise,中升型）

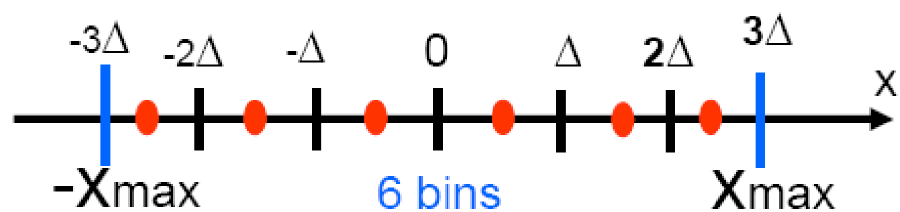
- **均匀**：每个量化区间的大小相同，除了最外面两个区间
  - b_i, y_i 在空间上均匀分布，空间均为 Δ
  - 对内部区间，$y_i = 1/2\left(b_{i-1} + b_i\right)$

过载
颗粒噪声

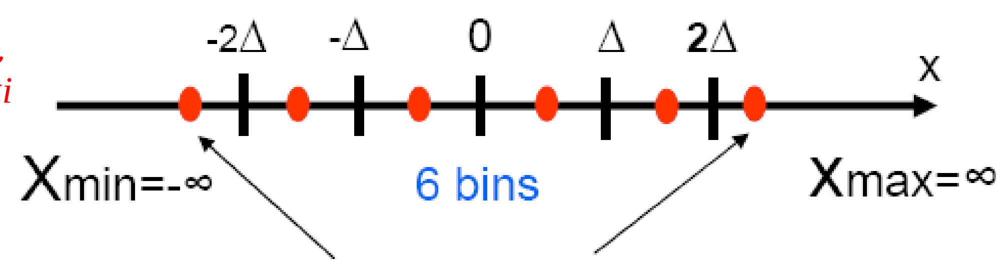
Uniform Midrise Quantizer



For finite Xmax and Xmin:



For infinite Xmax and Xmin:

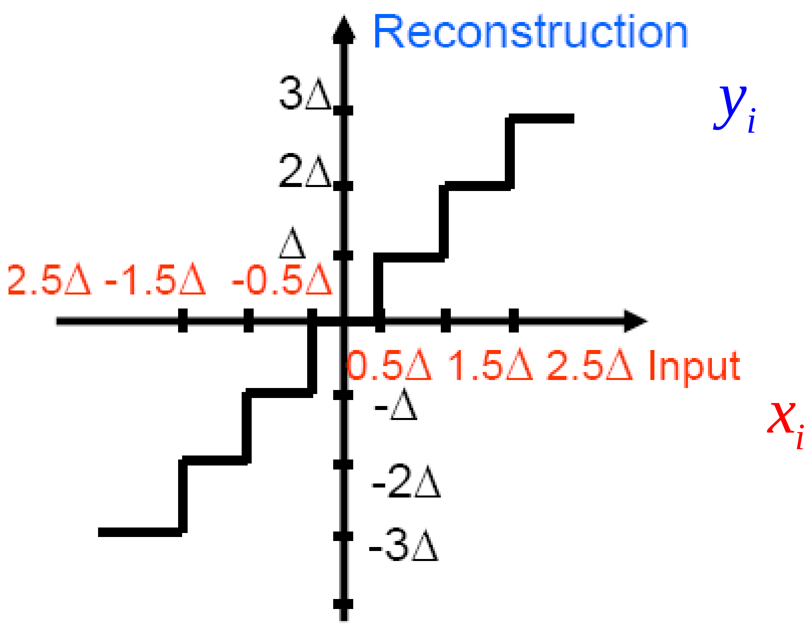


共有偶数个重构水平
0不是一个重构水平

最外部的两个重构水平离内
部仍是一个步长大小

4.1.2 均匀量化（Midtread）

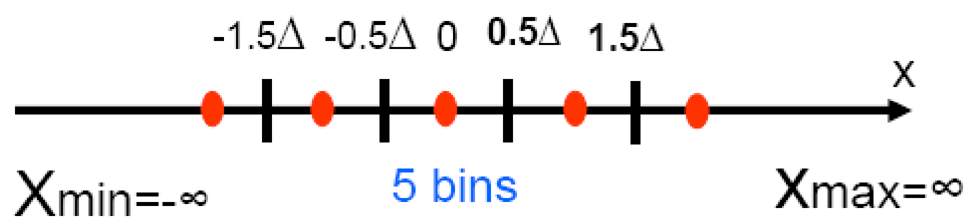
Uniform Midtread Quantizer



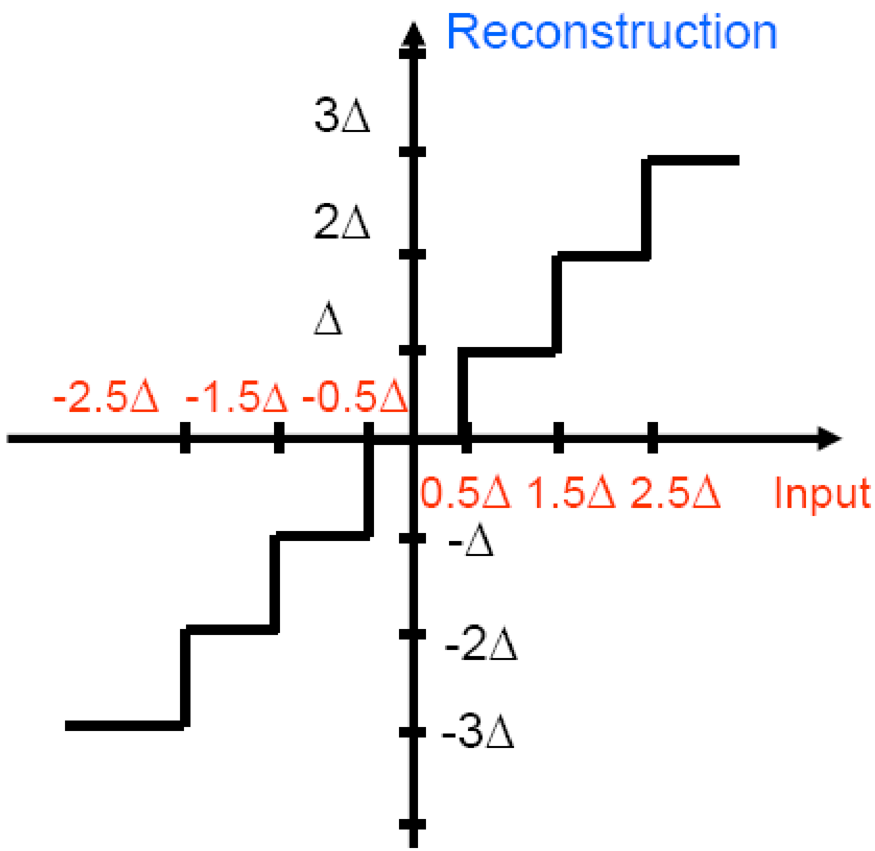
- Odd number of recon levels
- 0 is a recon level
- Desired in image/video coding

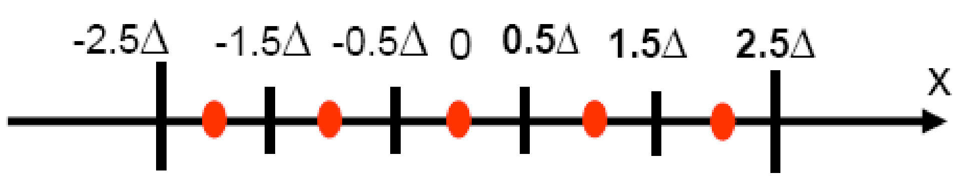
For finite Xmax and Xmin:

For infinite Xmax and Xmin:



4.1.2 均匀量化（Midtread）





- 量化映射：输出索引

$$q = A(x) = sign(x)\left\lfloor |x|/\Delta + 0.5 \right\rfloor$$

  - 例：$x = 1.8\Delta, q = 2$

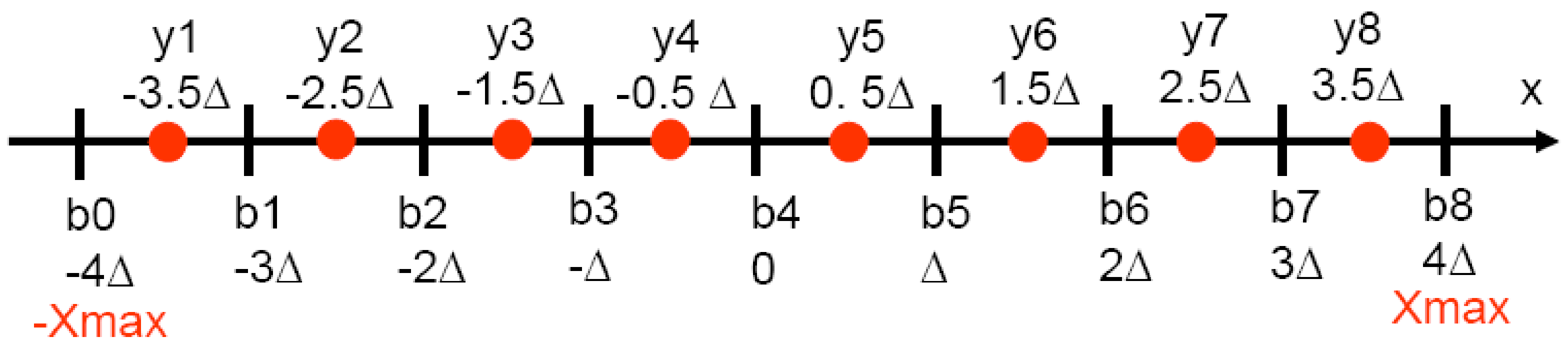
- 反量化映射：

$$\hat{x} = B(q) = q\Delta$$

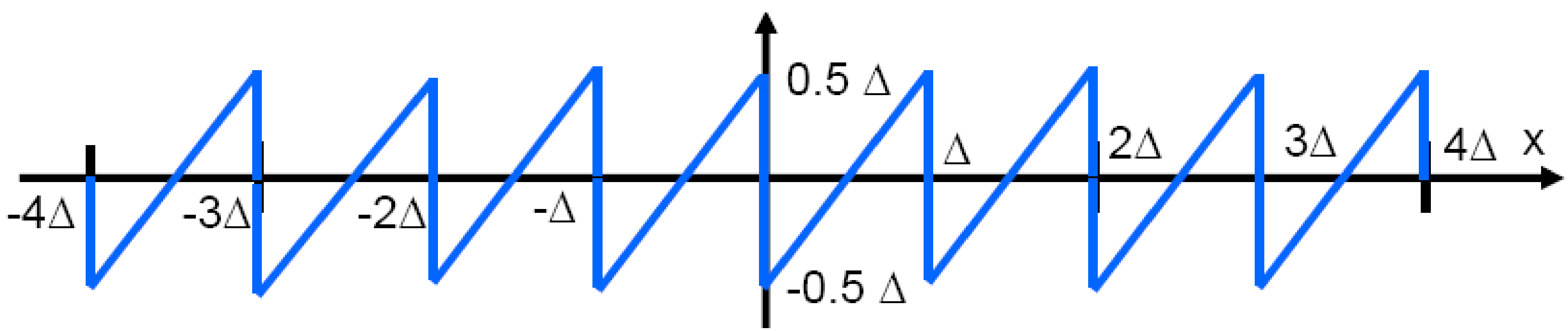
  - 例：$q = 2, \hat{x} = 2\Delta$

4.1.2 均匀量化（Midrise,失真分析）

- 假设输入信源为<u>均匀分布</u>：$[-X_{max}, X_{max}]$：$f(x) = 1/2X_{max}$
- 量化区间的数目为M（对Midrise量化，M为偶数）
- 步长：$\Delta = 2X_{max}/M$



- 量化误差：$e = x - \hat{x}$ 在区间 $[-\Delta/2, \Delta/2]$ 上均匀分布



4.1.2 均匀量化（Midrise,失真分析）

$$D = MSE = \int_{-\infty}^{\infty} \left(x - \hat{x}\right)^2 f\left(x\right) dx = \sum_{i=1}^{M} \int_{b_{i-1}}^{b_i} \left(x - y_i\right)^2 f\left(x\right) dx$$

证明：pdf为 $f(x) = \frac{1}{M\Delta}$

$$D = M \frac{1}{M\Delta} \int_0^{\Delta} \left(x - \Delta/2\right)^2 dx = \frac{1}{\Delta} \frac{1}{12} \Delta^3 = \frac{1}{12} \Delta^2$$

$$D = MSE = \frac{1}{12} \Delta^2$$

- 选择量化区间的数目*M*，使得失真小于允许的水平*D*

$$\frac{1}{12} \Delta^2 \le D \Rightarrow \frac{1}{12} \left(\frac{2X_{\max}}{M}\right)^2 \le D \Rightarrow M \ge X_{\max} \sqrt{\frac{1}{3D}}$$

4.1.2 均匀量化（Midrise,失真分析）

均匀分布在区间 $\left[-X_{\max}, X_{\max}\right]$ 的随机变量的方差：

$$\sigma_X^2 = \int_{-X_{\max}}^{X_{\max}} (x-0)^2 \frac{1}{2X_{\max}} dx = \frac{1}{3} X_{\max}^2$$

- 令 $M = 2^n$ ，即每个量化区间的索引用n个比特表示，则信噪比为

$$SNR = 10\log_{10} \frac{\sigma_X^2}{D(R)} = 10\log_{10} \frac{\sigma_X^2}{\Delta^2/12} = 10\log_{10} \frac{X_{\max}^2/3}{4X_{\max}^2/12M^2}$$

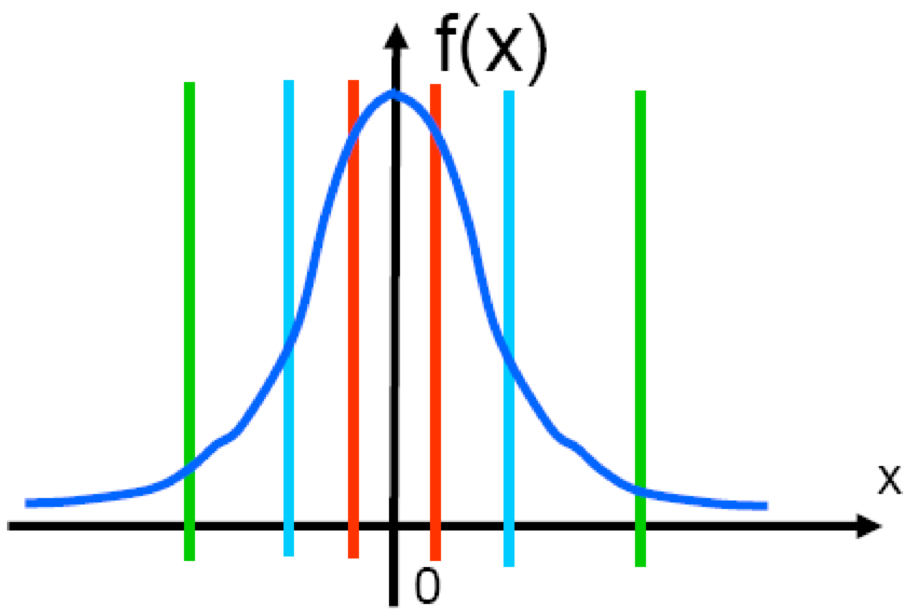
$$= 10\log_{10} M^2 = 10\log_{10} 2^{2n} = \left(20\log_{10} 2\right) n = 6.02n\ dB$$

- 若n—>n+1，则步长减为一半，噪声方差减为1/4，SNR增加6 dB。

4.1.3 Lloyd-Max标量量化器

- 均匀量化器只对均匀分布信源是最佳的
- 对给定的*M*，为了减小*MSE*，我们应该在概率$f(x)$ 较大时缩小量化区间，而在$f(x)$较小时增大量化区间

$$D=\int_{-\infty}^{\infty}(x-\hat{x})^2 f(x)\,dx=\sum_{i=1}^{M}\int_{b_{i-1}}^{b_i}(x-y_i)^2 f(x)\,dx$$



4.1.3 Lloyd-Max标量量化器

- 亦称为pdf-最佳量化器

$$D=\int_{-\infty}^{\infty}(x-\hat{x})^2 f(x)\,dx=\sum_{i=1}^{M}\int_{b_{i-1}}^{b_i}(x-y_i)^2 f(x)\,dx$$

- 给定M，求最佳b_i, y_i使得MSE最小，满足

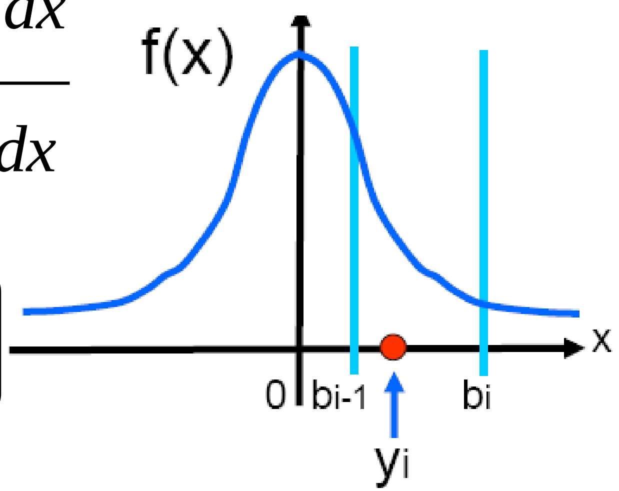
$$\frac{\partial D}{\partial y_i}=0,\quad \frac{\partial D}{\partial b_i}=0$$ 分别固定b_i和y_i

- $$\frac{\partial D}{\partial y_i}=0 \Rightarrow y_i=\mathrm{E}\left(X \mid X\in I_i\right)=\frac{\int_{b_{i-1}}^{b_i} xf(x)dx}{\int_{b_{i-1}}^{b_i} f(x)dx}$$

x在$[b_{i-1},b_i]$上的数学期望

x在$[b_{i-1},b_i]$上的概率

- 即y_i为区间$[b_{i-1},b_i]$的质心。



4.1.3 Lloyd-Max标量量化器

$$\frac{\partial D}{\partial b_i} = 0 \ \Rightarrow \ b_i = \frac{y_i + y_{i+1}}{2}$$

- 即 b_i 为 y_i 和 y_{i+1} 的中点→最近邻量化器
- Lloyd-Max条件总结：

$$\boxed{y_i = \frac{\int_{b_{i-1}}^{b_i} xf(x)dx}{\int_{b_{i-1}}^{b_i} f(x)dx}, \ b_i = \frac{y_i + y_{i+1}}{2}}$$

  - 第一个结果表明，门限（判决）电平应取在相邻量化输出电平的中点
  - 第二个结果表明，量化电平（重建电平）应取在量化间隔的质心上

4.1.3 Lloyd-Max标量量化器

- 与均匀量化器的关系：
  - 当量化器的输入为均匀分布时，$f(x)=c$，Lloyd-Max量化器变为均匀量化器

$$y_i=\frac{\int_{b_{i-1}}^{b_i} xf(x)dx}{\int_{b_{i-1}}^{b_i} f(x)dx}=\frac{c\int_{b_{i-1}}^{b_i} xdx}{c(b_i-b_{i-1})}=\frac{\frac{1}{2}(b_i^2-b_{i-1}^2)}{(b_i-b_{i-1})}=\frac{1}{2}(b_i+b_{i-1})$$

4.1.3 Lloyd-Max标量量化器

- 最佳量化器的条件：

$$\boxed{y_i = \frac{\int_{b_{i-1}}^{b_i} xf(x)dx}{\int_{b_{i-1}}^{b_i} f(x)dx},\ b_i = \frac{y_i + y_{i+1}}{2}}$$

- 给定 b_i，可以计算对应的最佳 y_i
- 给定 y_i 可以计算对应的最佳 b_i

- 问题：如何同时计算最佳的 b_i 和 y_i？
- 答案：迭代方法

4.1.3 Lloyd-Max标量量化器

迭代Lloyd-Max算法（已知$f(x)$）

1. 初始化所有的 $y_i, j=1, D_0=\infty$
2. 更新所有的决策边界： $b_i=\dfrac{y_i+y_{i+1}}{2}$
3. 更新所有的 y_i ：

$$y_i=\frac{\int_{b_{i-1}}^{b_i} xf(x)dx}{\int_{b_{i-1}}^{b_i} f(x)dx}$$

4. 计算*MSE*： $D_j=\sum_{k=1}^{M}\int_{b_{i-1}}^{b_i}\left(x-y_k\right)^2 f\left(x\right)dx$

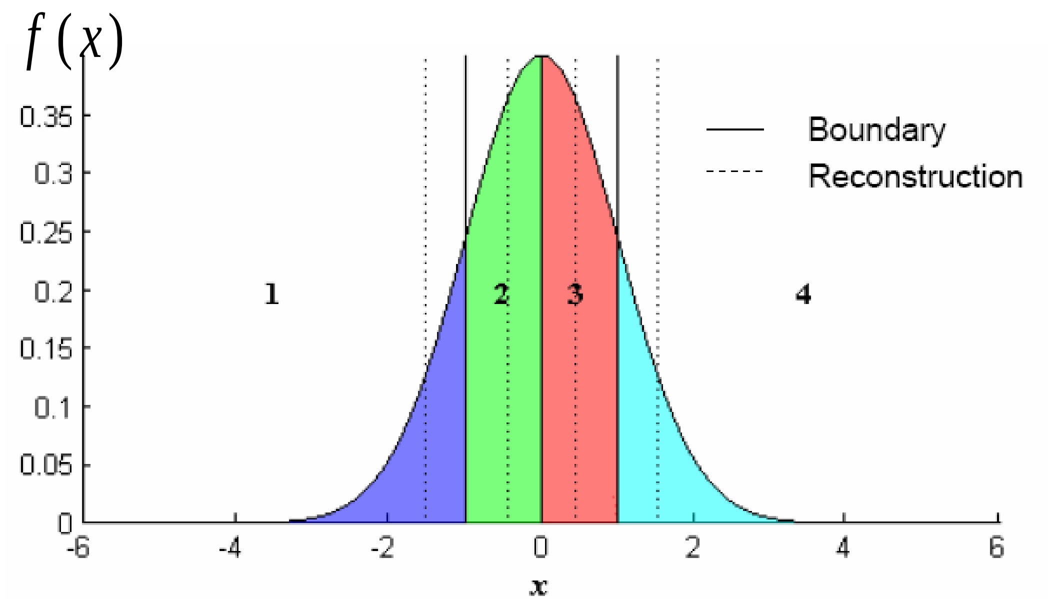
5. 如果 $\left(D_{j-1}-D_j\right)/D_{j-1}<\varepsilon$ 停止；否则 $j=j+1$
，转第2步

4.1.3 Lloyd-Max标量量化器

例：Lloyd-Max算法的应用(I)

- *X*为0均值，1方差的高斯分布，即$X \sim N(0,1)$
- 设计一个4个索引的量化器，使得期望失真D^*最小
- 用Lloyd-Max算法得到最佳量化器
  - 决策边界：-0.98, 0, 0.98
  - 重构水平：–1.51, -0.45, 0.45, 1.51
  - $\boldsymbol{D} = 0.111775$

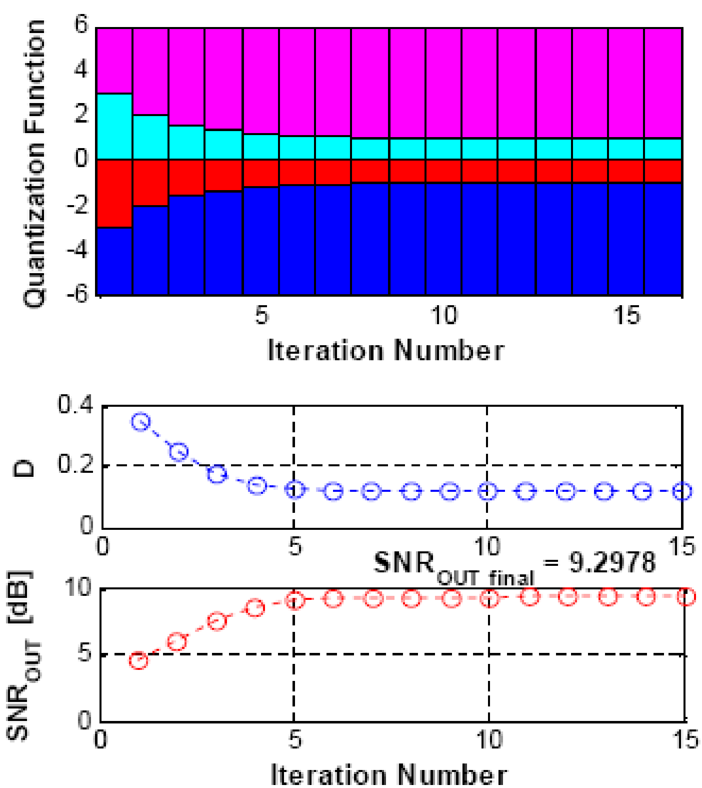
    $\boldsymbol{SNR} = 9.30\boldsymbol{dB}$



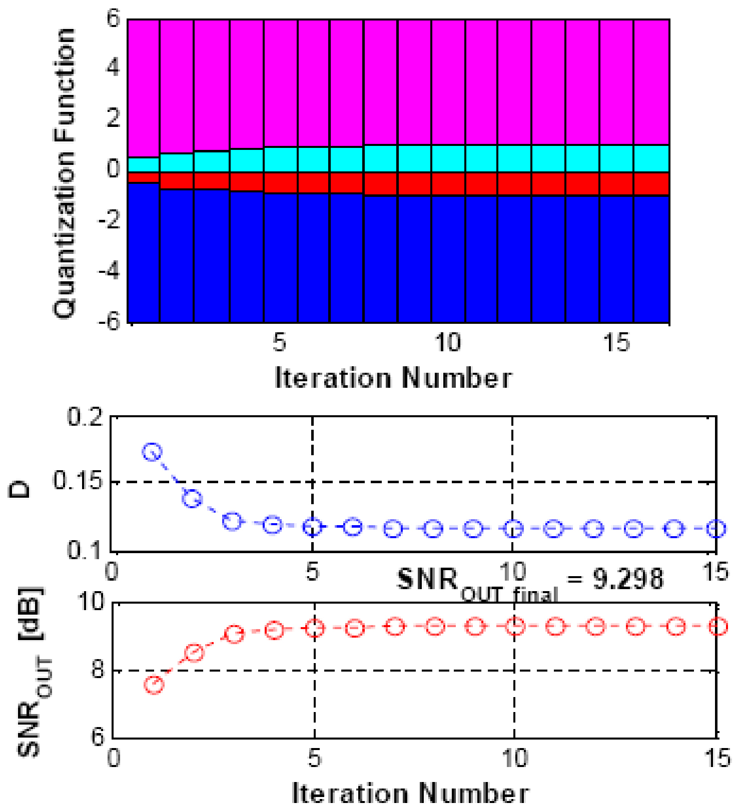
例： Lloyd-Max算法的应用(I)

- 收敛

初始化A：决策边界为：–3, 0, 3



初始化A：决策边界为：–1/2, 0, 1/2

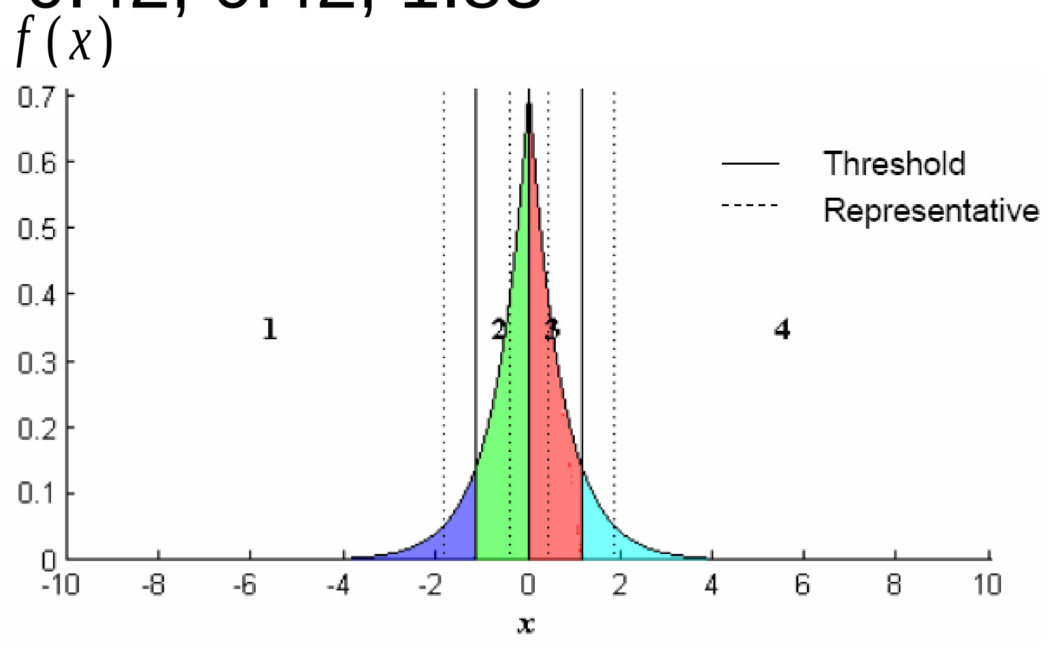


- 在两种情况下，经过6次迭代后，$\left(D-D^{*}\right)/D^{*}<1\%$

例：Lloyd-Max算法的应用(II)

- X为0均值，1方差的Laplacian分布
- 设计一个4个索引的量化器，使得期望失真D^*最小
- 用Lioyd-Max算法得到最佳量化器
  - 决策边界：-1.13, 0, 1.13
  - 重构水平：-1.83, -0.42, 0.42, 1.83
  - $\boldsymbol{D} = 0.18$

    $\boldsymbol{SNR} = 7.54\boldsymbol{dB}$

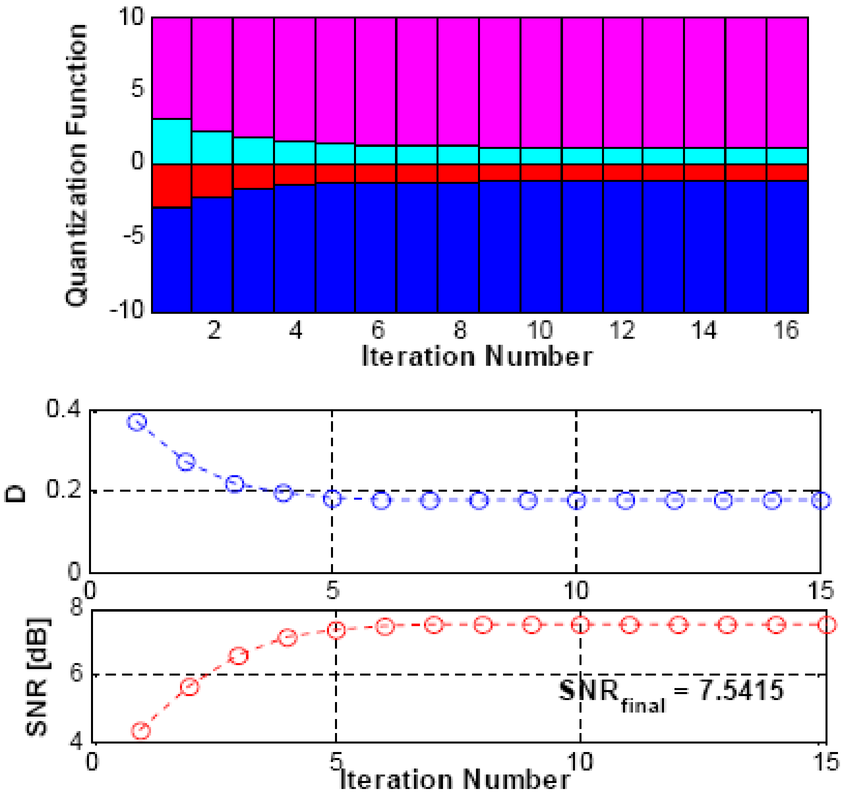


一个好的预测器输出的预测误差通常满足0周围高峰值的分布，如Laplacian分布

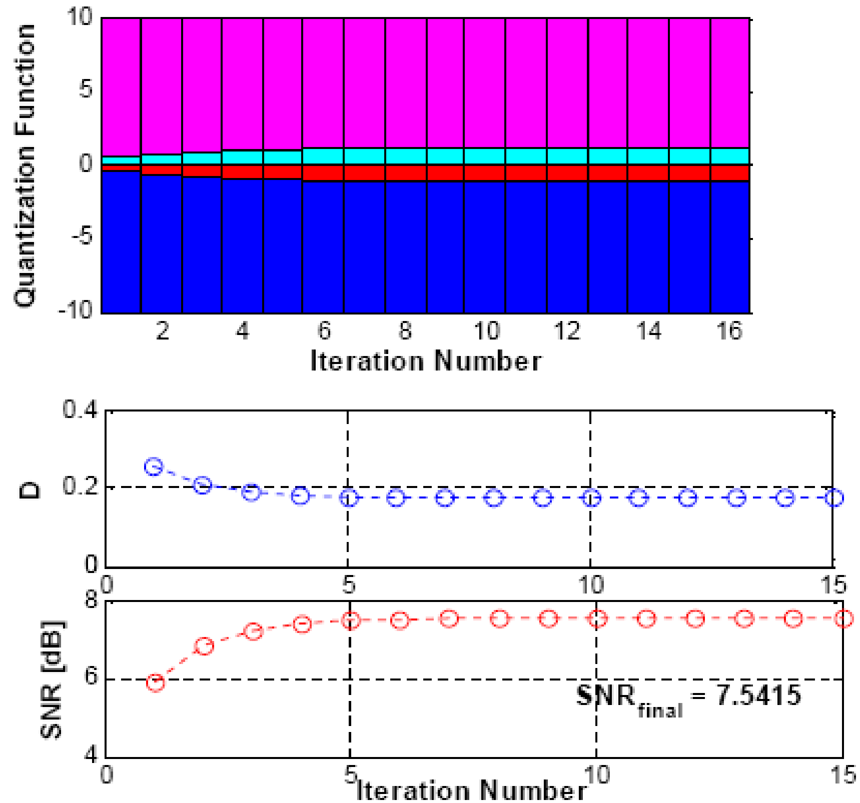
例：Lloyd-Max算法的应用(II)

- 收敛

初始化A：决策边界为：–3, 0, 3



初始化A：决策边界为：–1/2, 0, 1/2



- 在两种情况下，经过6次迭代后，$\left(D-D^{*}\right)/D^{*}<1\%$

高码率近似

- 假设码率很高（***R***很大）， **Lloyd-Max量化器的MSE为**

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

其中 $\varepsilon^2 \sigma_X^2 = \frac{1}{12}\left[\int \sqrt[3]{f(x)}dx\right]^3$

- ε^2依赖于分布，对均匀分布、**Laplacian**分布和高斯分布 ε^2，分别为 $\varepsilon^2 = 1,\ 9/2,\ \sqrt{3\pi}/2 = 2.721$

- 信噪比***SNR*:**$10\log_{10}\frac{\sigma_X^2}{D(R)} = 6.02R - 10\log_{10}\varepsilon^2 \ dB$

  - 对均匀分布、**Laplacian**分布和高斯分布，$10\log_{10}\varepsilon^2$ 分别为：

$$10\log_{10}\varepsilon^2 = 0,\ 6.53,\ 4.35\ dB$$

4.1.4 熵约束标量量化器
(Entropy-constrained scalar quantizer, ECSQ) *

- **Lloyd-Max量化器：**
  - 对索引用固定码率编码：$\log_2 M$ **(*R*)**比特
- **熵约束标量量化器**
  - 对量化索引用变长码编码：
    - 对量化索引用熵编码技术编码
    - 平均码率~重构水平的熵 $\leq \log_2 M$

$$H\left(\hat{X}\right) = -\sum_{k=1}^{M} p_k \log p_k \leq R$$

    - 比**Llyod-Max**量化器的性能更好

P. A. Chou, T. Lookabaugh, R. M. Gray, "Entropy-constrained vector quantization," IEEE Trans. Signal Processing, vol. 37, no. 1, pp. 31-42, Jan 1989

4.1.4 熵约束标量量化器
(Entropy-constrained scalar quantizer, ECSQ) *

- 问题的形式化描述：

最小化 $D=\mathrm{E}\left(\left(X-\hat{X}\right)^2\right)=\sum_{k=1}^{M}\int_{b_{k-1}}^{b_k}\left(x-y_k\right)^2 f\left(x\right)dx$

满足 $H\left(\hat{X}\right)=-\sum_{k=1}^{M}p_k\log p_k\le R$

其中 $p_k=\int_{b_{k-1}}^{b_k}f\left(x\right)dx$

- 用**Lagrange**费用函数：

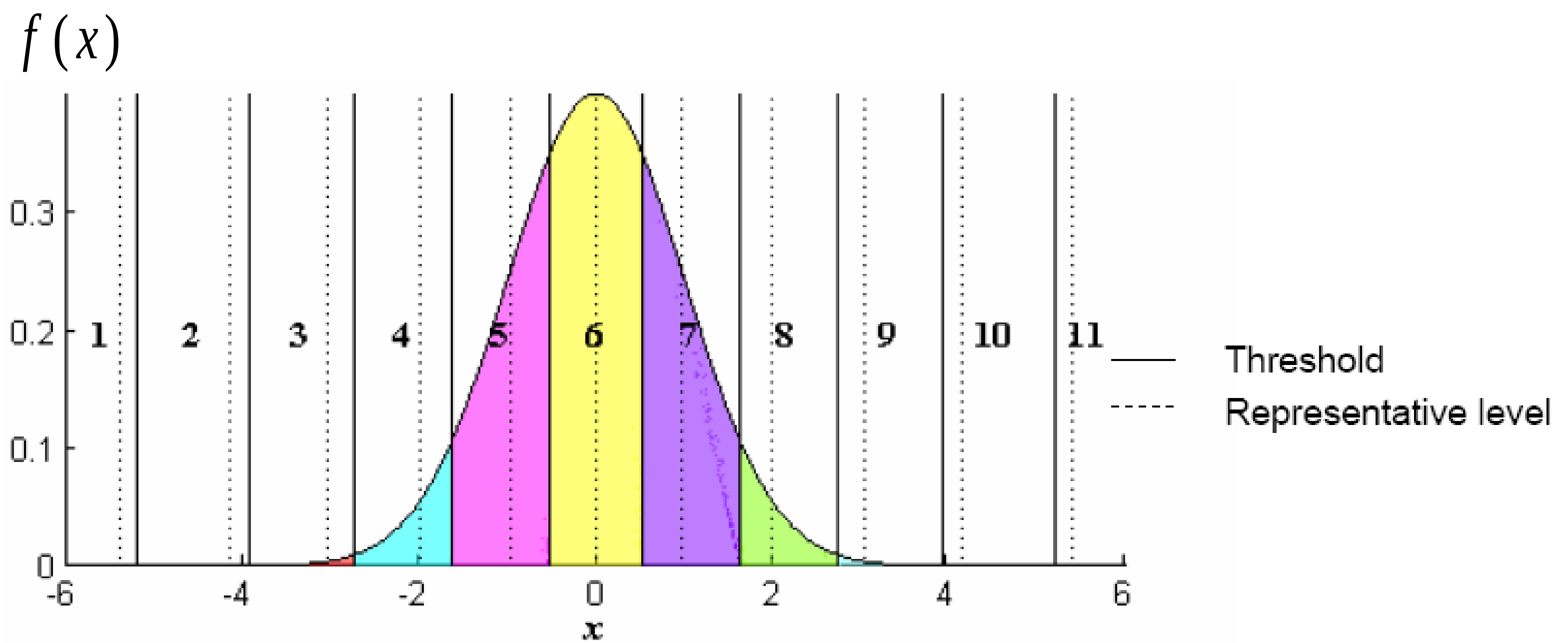
$$J\left(\lambda\right)=\mathrm{E}\left(\left(X-\hat{X}\right)^2\right)+\lambda H\left(\hat{X}\right)$$

  - 太复杂，不能直接求解
  - 用迭代法求解

例： ECSQ算法的应用(I)

- ***X*为0均值，1方差的高斯分布**，即 $X \sim N(0,1)$
- 设计一个 $R \cong 2$ 的**ECSQ**，使得期望失真 D^*最小
  - **11**个区间（**[-6，6]**内）：几乎是均匀
  - $D^* = 0.09 = 10.53dB,\ R = 2.0035$
    - 定长编码：

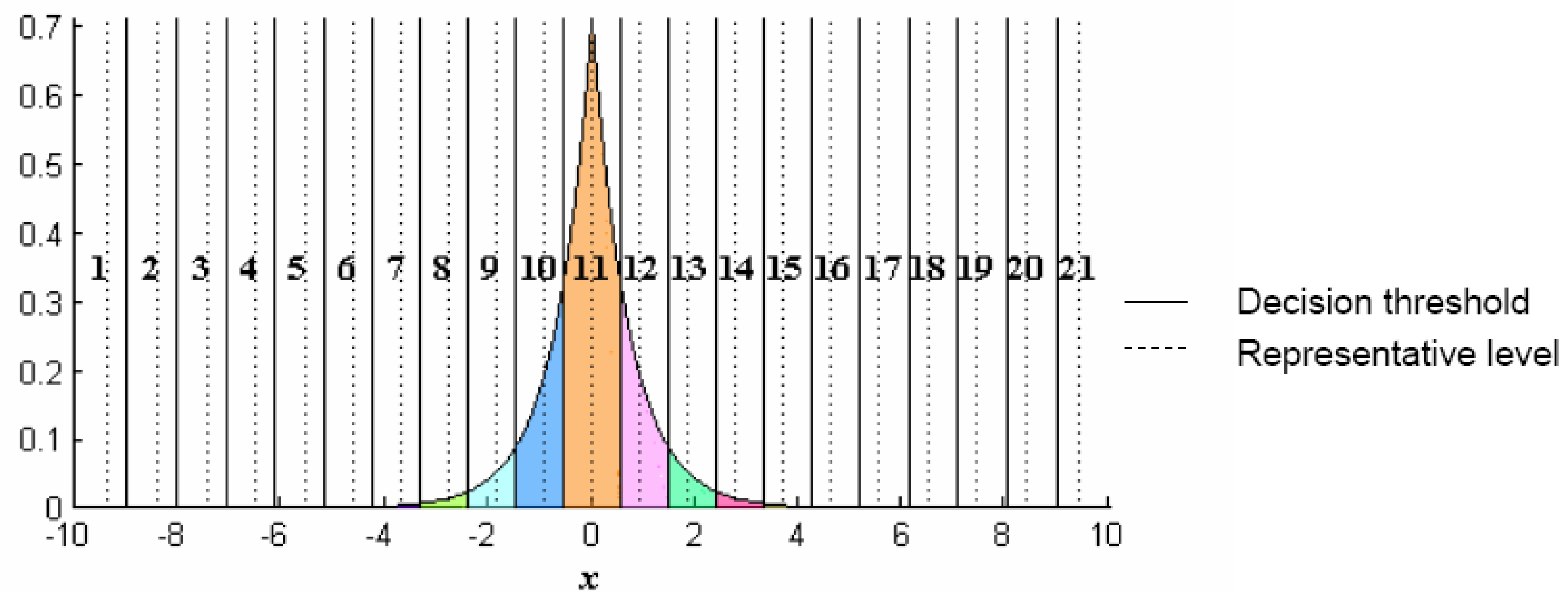
$$D^* = 0.12 = 9.30\ dB$$



例： ECSQ算法的应用(II)

- *X*为0均值，1方差的Laplacian分布
- 设计一个 $R \cong 2$ 的ECSQ，使得期望失真 D^*最小
  - 21个区间（[-10，10]内），几乎是均匀的
  - $D^* = 0.07 = 11.38dB$

$f(x)$



高码率下各种标量量化器的性能比较

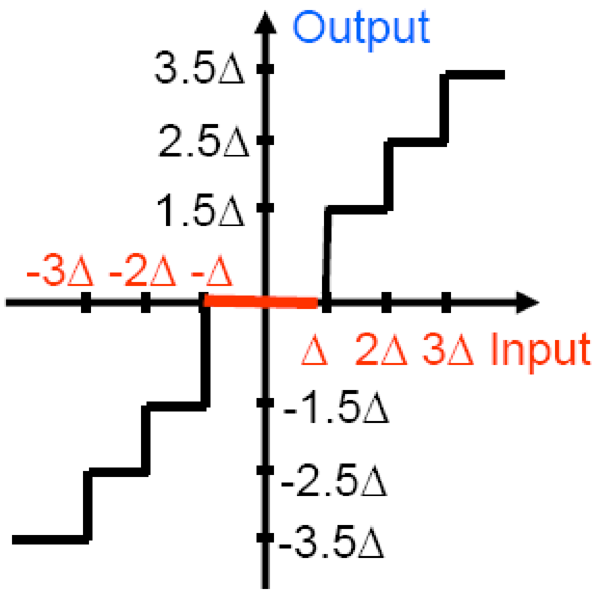
- 高码率下的失真-码率函数：$D(R) \cong \varepsilon^2 \sigma_X^2 2^{-2R}$
- 缩放因子 ε^2

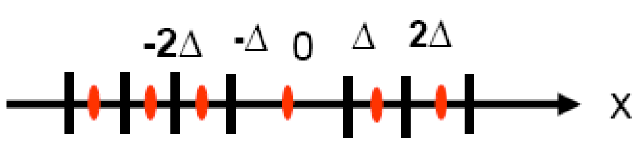
	Shannon LowBd	Lloyd-Max	Entropy-coded
Uniform	$\frac{6}{\pi e} \cong 0.703$	1	1
Laplacian	$\frac{e}{\pi} \cong 0.865$	$\frac{9}{2} = 4.5$	$\frac{e^2}{6} \cong 1.232$
Gaussian	1	$\frac{\sqrt{3}\pi}{2} \cong 2.721$	$\frac{\pi e}{6} \cong 1.423$

相同码率R下，**ECSQ的失真比Lloyd-Max量化器更小**

4.1.5 Deadzone Midtread Quantizer

- 0附近的量化区间大小为其余量化区间的两倍，其他量化区间仍是均匀的
- 产生更多的0
- 对图像/视频很有用





- 量化映射：

$$q = A(x) = sign(x)\left\lfloor |x|/\Delta \right\rfloor$$

- 反量化映射：

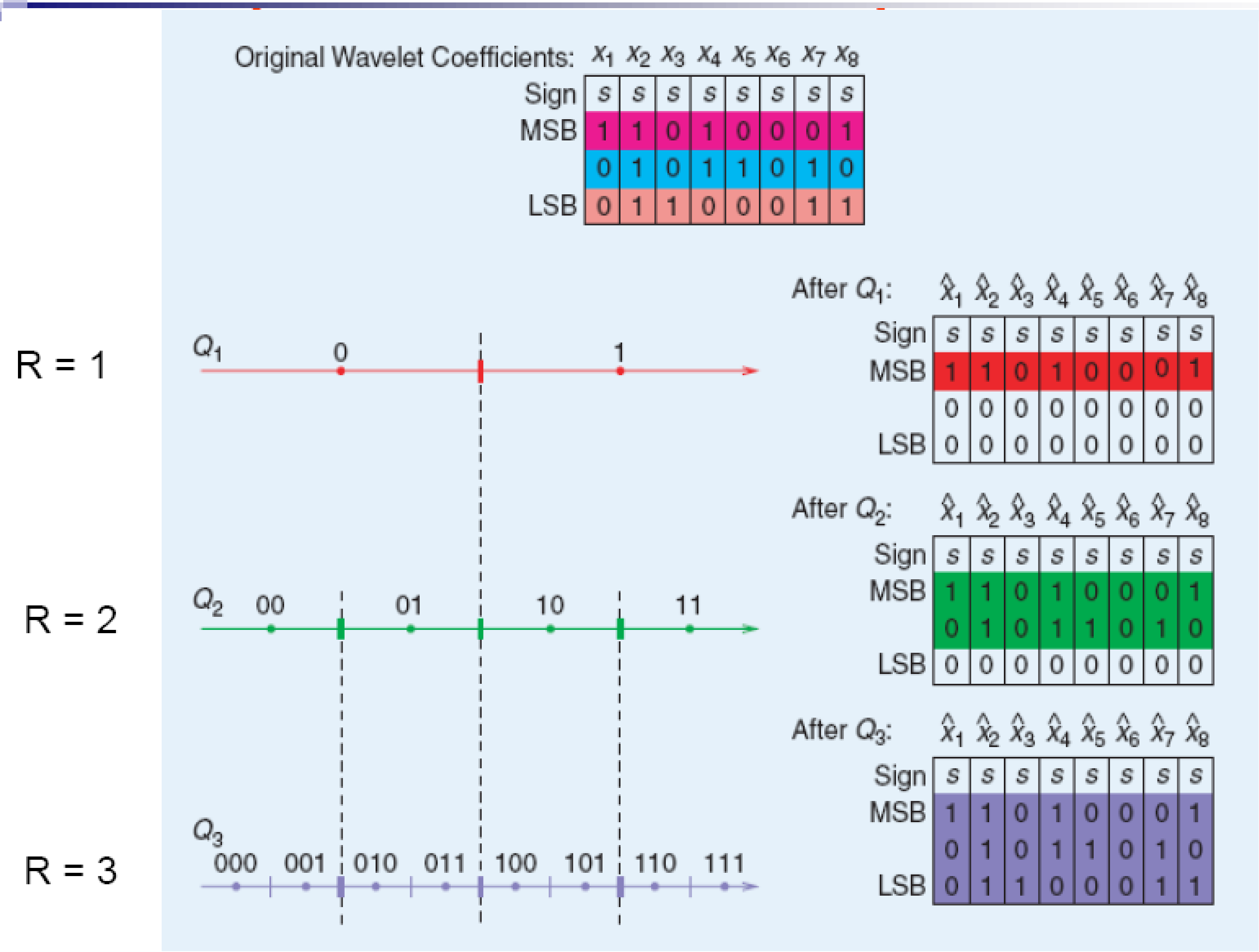
$$\hat{x} = B(q) = \begin{cases} 0 & q = 0 \\ sign(q)\left(\left|q + \frac{1}{2}\right|\right) & q \neq 0 \end{cases}$$

  - 例： $q = 2, \hat{x} = 2\Delta$

4.1.6 嵌入式量化器

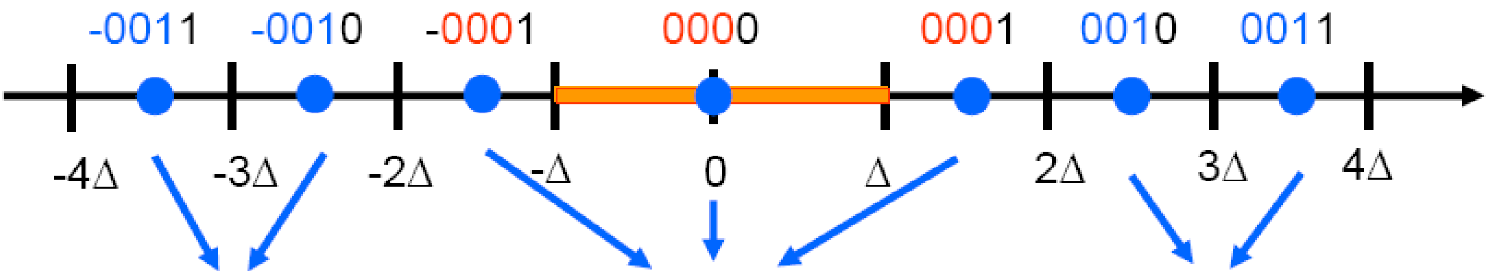
- 动机：可伸缩（scalable）解码
  - 随着比特流的解码，渐近地精化重构数据
  - 对低带宽连接有用
  - 是JPEG2000的一个关键特征
- 嵌套量化：低码率器的区间被再分割，以产生更高码率的量化器
- 可以通过截断量化索引获得较粗燥的量化

例1：均匀量化器

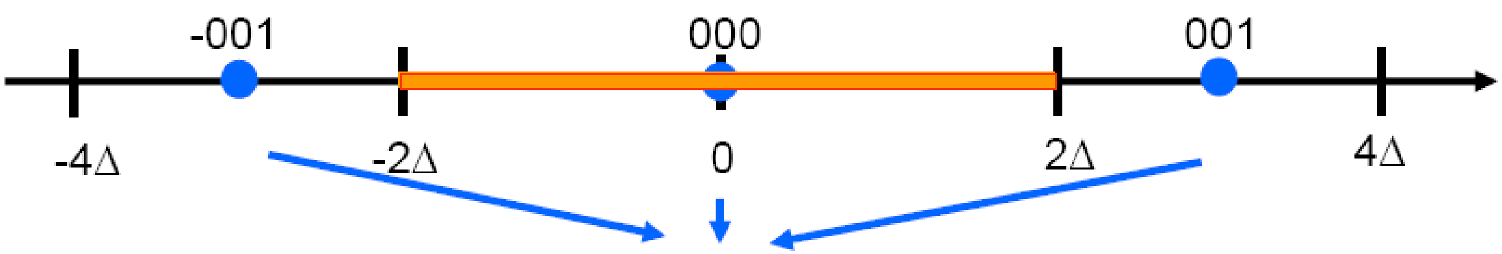

Original Wavelet Coefficients: x_1 x_2 x_3 x_4 x_5 x_6 x_7 x_8
Sign
MSB
LSB
R = 1
R = 2
R = 3
Q_1
Q_2
Q_3
0
1
00
01
10
11
000
001
010
011
100
101
110
111
After Q_1: $\hat{x}_1$ $\hat{x}_2$ $\hat{x}_3$ $\hat{x}_4$ $\hat{x}_5$ $\hat{x}_6$ $\hat{x}_7$ $\hat{x}_8$
After Q_2: $\hat{x}_1$ $\hat{x}_2$ $\hat{x}_3$ $\hat{x}_4$ $\hat{x}_5$ $\hat{x}_6$ $\hat{x}_7$ $\hat{x}_8$
After Q_3: $\hat{x}_1$ $\hat{x}_2$ $\hat{x}_3$ $\hat{x}_4$ $\hat{x}_5$ $\hat{x}_6$ $\hat{x}_7$ $\hat{x}_8$
s s s s s s s s
1 1 0 1 0 0 0 1
0 1 0 1 1 0 1 0
0 1 1 0 0 0 1 1
0 0 0 0 0 0 0 0

例2：Deadzone quantizer

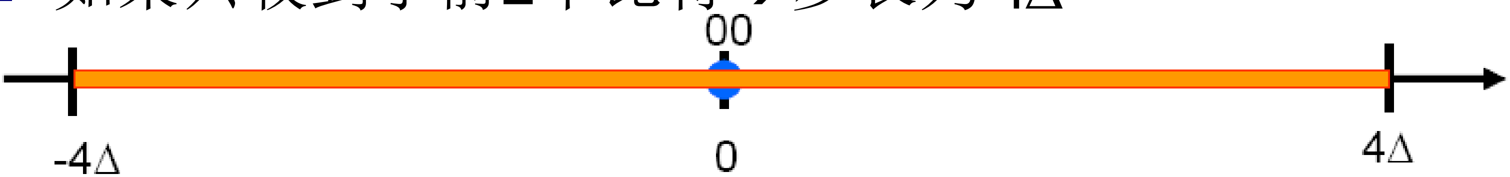
- 假设deadzone量化器的量化区间的索引用4个比特表示
- 如果收到了所有4个比特→步长为Δ



- 如果只收到了前3个比特→步长为2Δ



- 如果只收到了前2个比特→步长为4Δ



标量量化总结

- 对于已知概率模型及其数字特征的随机过程，比较容易根据概率分布安排量化器的决策边界，以得到最小量化失真的优化量化器。
- 如果概率分布是均匀的，则采用均匀量化比较理想
- 对于分布概率模型未知的随机过程，其优化量化器的设计较为困难，可以采用Lloyd-Max算法来解决，但实现时还有一定困难，不宜硬件实现，执行时间也因初始值选取的不同而不同
- 考虑视觉特性的量化器设计：如何根据主观评定规则，设法令压缩编码中产生的各类量化误差在主观上难以察觉

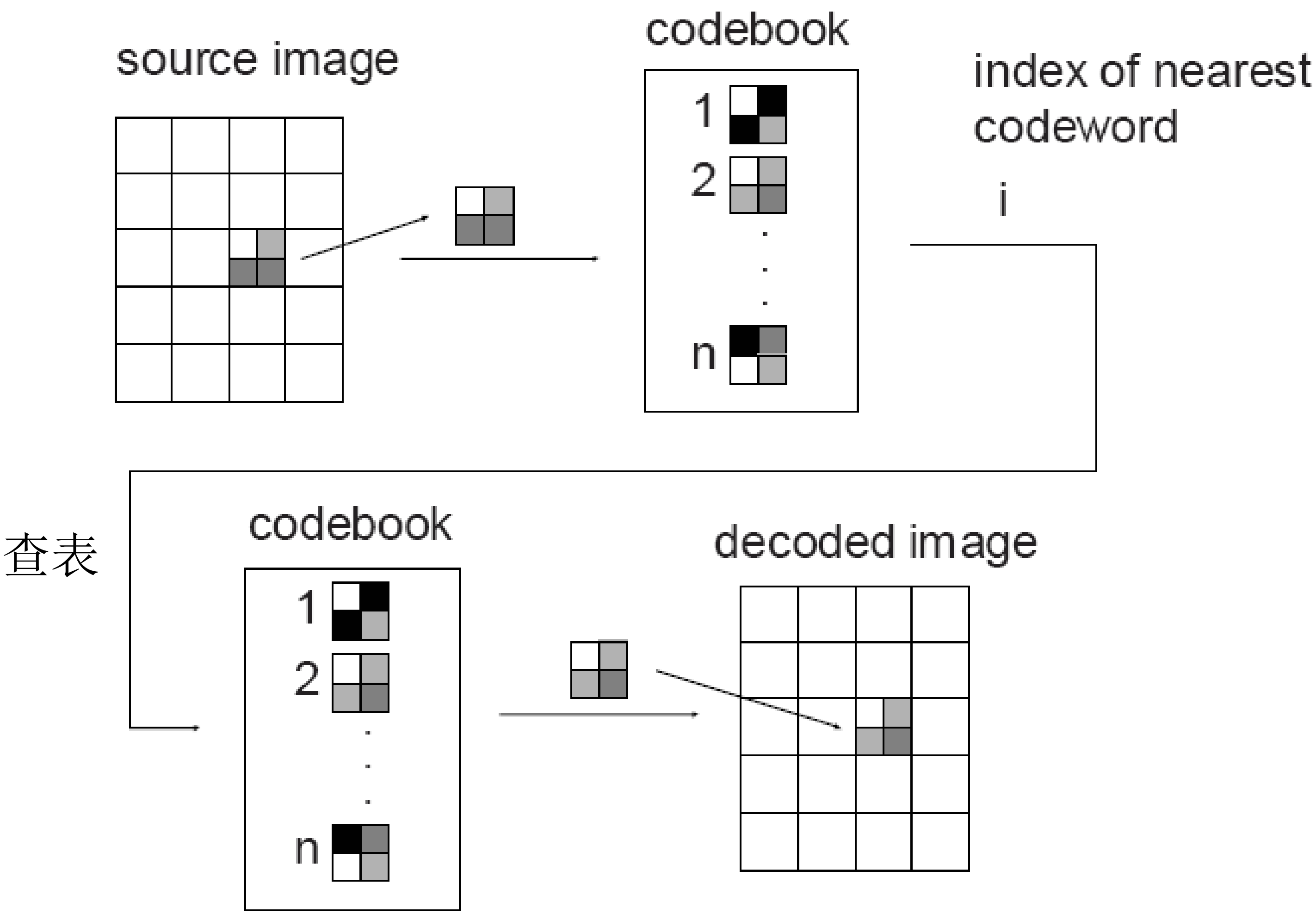
4.2 矢量量化 (Vector Quantization,VQ)

- 压缩符号串比压缩单独符号在原理上可产生更好的效果
  - 如图像和声音的相邻数据项都是相关的
- 矢量量化的思路：
  - 量化时不是处理单个符号，而是一次处理一组符号（矢量）

4.2.1 矢量量化的基本思想

- 以图像编码为例

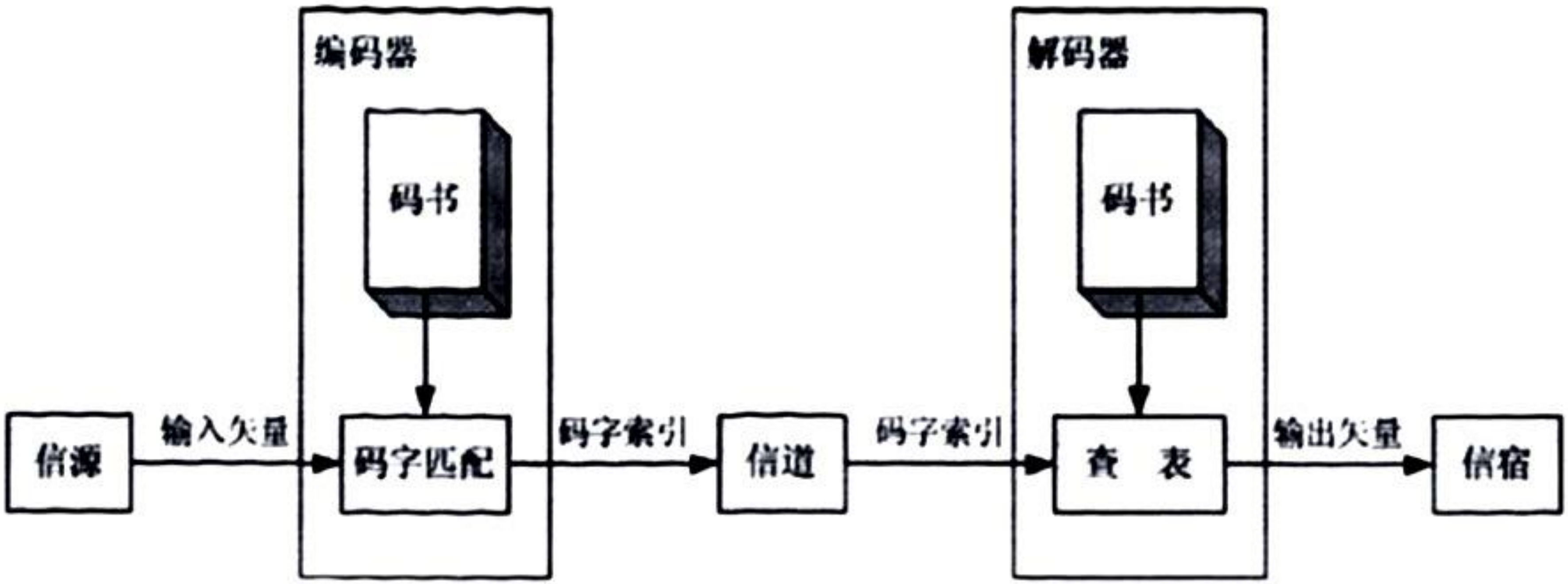
搜索距离最近的码字



4.2.1 矢量量化的基本思想

- 假设块大小为 $a \times b$ ，码书中共有M个码字（码字也是长度为 $a \times b$ 的矢量）
- 则码率 $R = \dfrac{\log M}{a \times b}\ bpp$
- 例：$a = b = 4, M = 1024$
  - 则码率为 $R = 10/16 = 0.63\ bpp$
  - 压缩比为：$8 : 0.63 = 128 : 1$
- 可以通过对索引用熵编码技术得到更高的压缩比

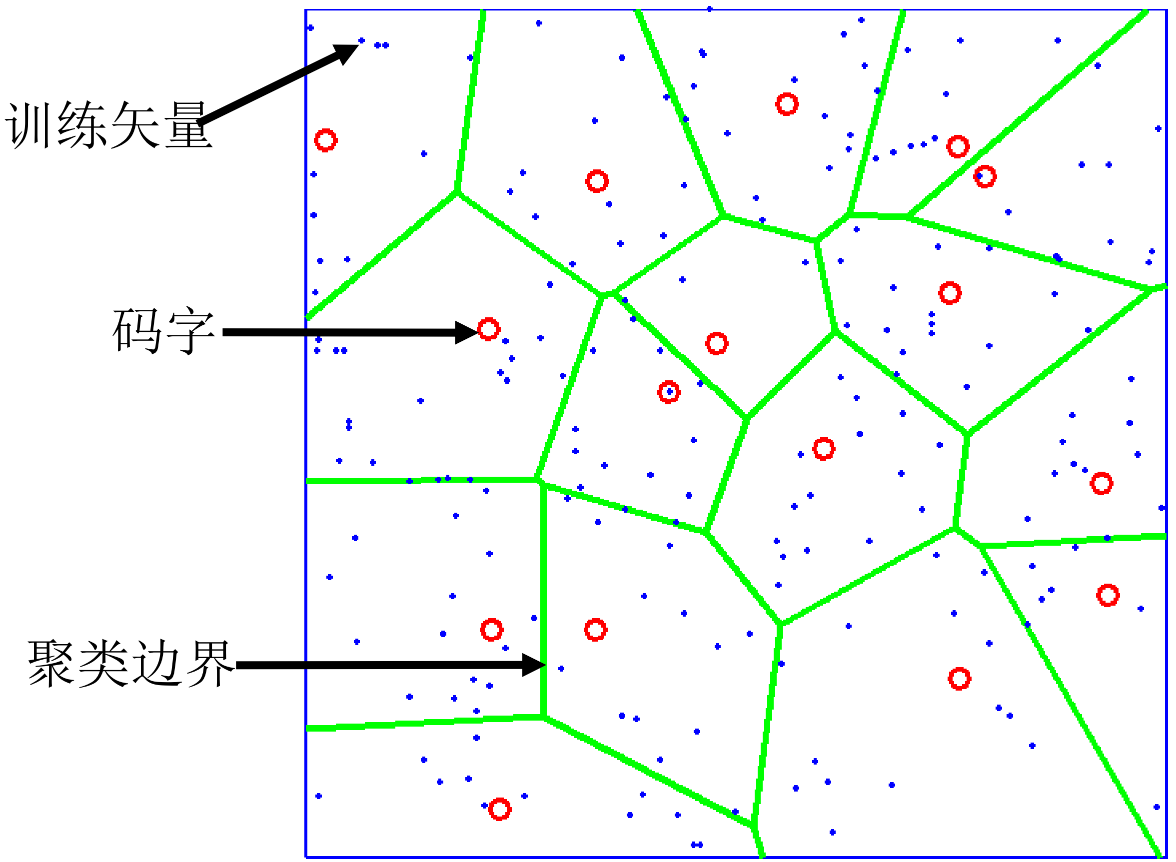
4.2.1 矢量量化的基本思想



4.2.2 LBG算法

- 将Lloyd 算法推广到矢量量化，所以亦称为推广的Lloyd 算法(Generalized Lloyd Algorithm, GLA) [Linde, Buzo, Gray, 1980]

- 给定训练集：
  - 收集的训练集需有代表性



4.2.2 LBG算法

给定训练集：$T = \{ \boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_N \}$

- 1. 初始化所有的 $\boldsymbol{y}_i,\ i = 1, ..., M$
- 2. 对训练集中所有的训练样本 $\boldsymbol{x}_n, n = 1, 2, ..., N$ ，找到距离最近的码字：

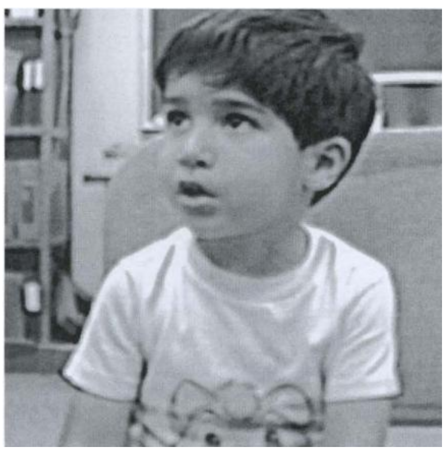
$$Q(\boldsymbol{x_n}) = \boldsymbol{y}_i,\ iff\ \ d(\boldsymbol{x} - \boldsymbol{y}_i) \leq d(\boldsymbol{x} - \boldsymbol{y}_k),\ for\ all\ \ k \neq i$$

- 3. 计算平均失真
- 4. 如果平均失真足够小，停止；否则转第5步
- 5. 用每个量化区域内所有矢量的平均值替代 y_i ，转第2步

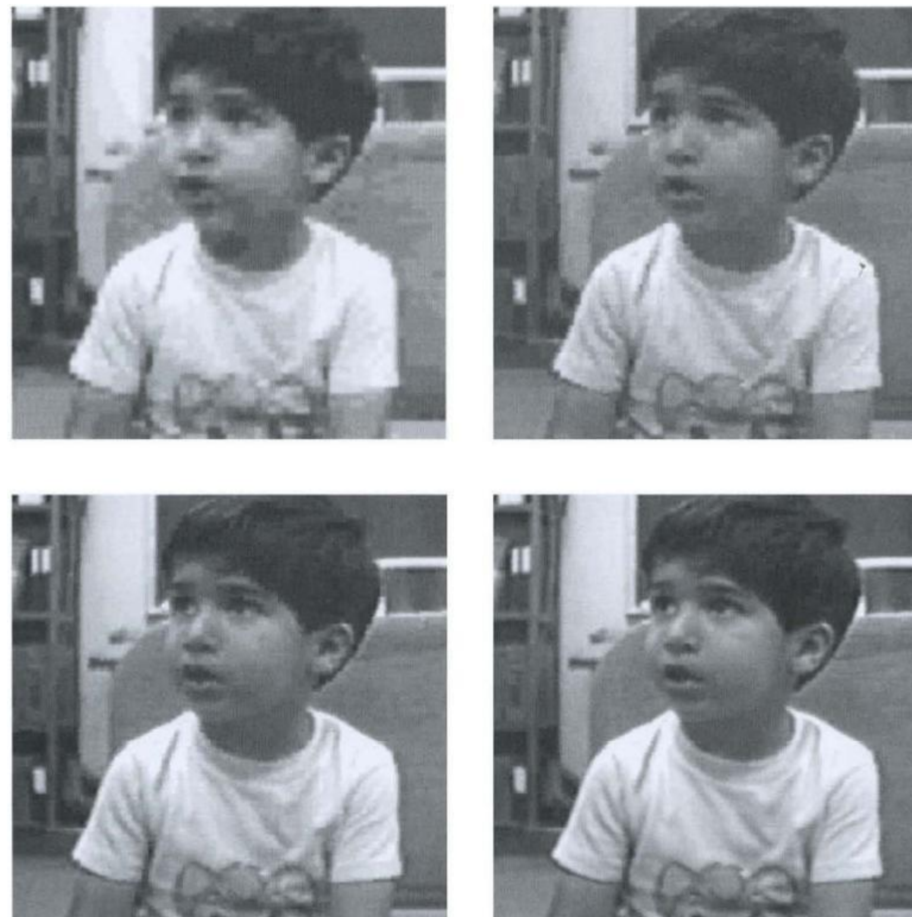
同聚类中的K-means聚类

LBG用于图像压缩

- 每块大小为L=4×4， 用K=16,64,256,1024个码字的码书对图像进行矢量量化
- 码率R=?
- Sinan图像训练得到码书



原始图像



4.2.2 LBG算法

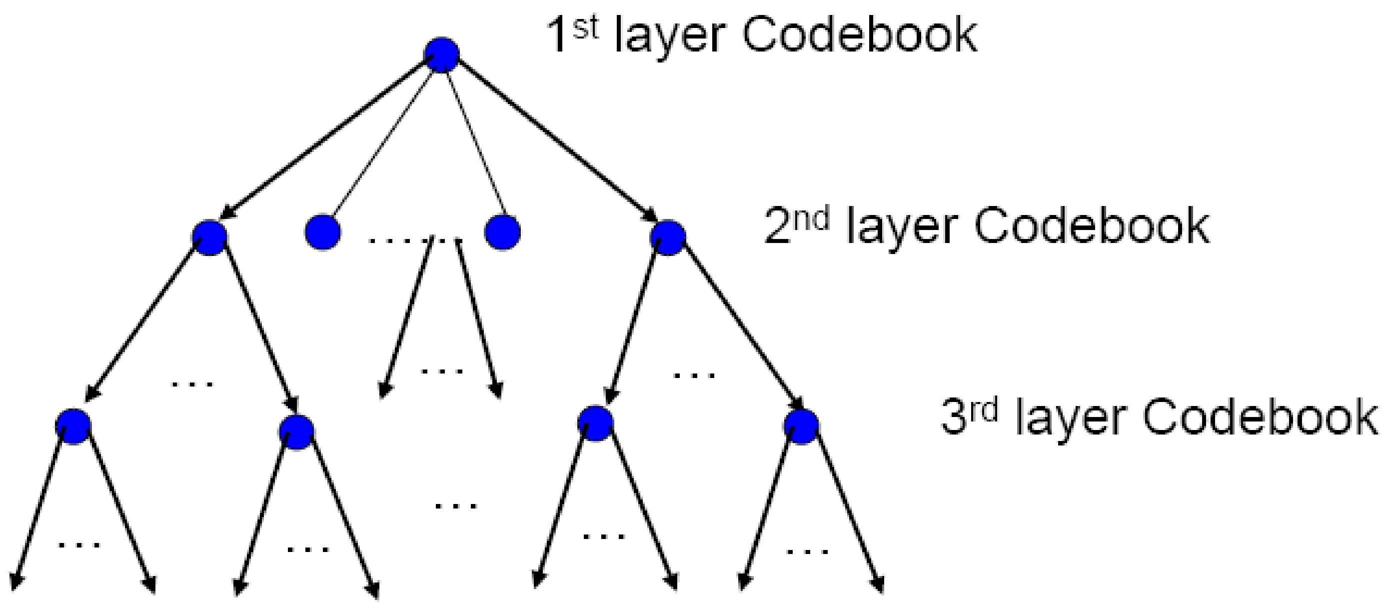
- 优化过程中可能限于局部最小值
  - 依赖于初始码书的选取
- 初始码书的选取
  - 随机选择：重复多次，取失真最小的结果
  - 分裂：从一个类开始，每次将失真最大/数量最多的类分裂成两个
  - 合并：从N个类开始，每次将两个失真最小的类合并
- 码字中缺少结构
  - 编码复杂性高：需要全搜索
  - 存储要求高，码书指数增长

矢量量化的改进

- 产生码书—矢量编码—矢量解码
  - 在矢量编码阶段，编码器将输入图像划分成图像块，并在码书中搜索同输入图像块最接近的码字，用该码字在码书中的索引号表示该图像块．在解码端，解码器可以根据收到的索引号恢复图像．
- 如何快速准确的找到最佳匹配码字是矢量量化中的一个重要问题
  - 穷尽搜索法计算输入矢量同所有码字的欧氏距离，找出其中最小的即为最佳匹配码字，这是一种能保证精度但最耗时的方法

矢量量化的改进

- 树结构的矢量量化
  - 降低搜索复杂性
  - 但需要更多存储
-



矢量量化的改进

- 结构化的矢量量化
  - 金字塔VQ
  - Lattice VQ
- Trellis VQ
- 自适应矢量量化
  - 矢量大小可变
  - 码书项的数目可变

Reading

- J. Max, “Quantizing for Minimum Distortion,” IEEE Trans. Information Theory, vol. 6, no. 1, pp. 7-12, March 1960.
- S. P. Lloyd, “Least Squares Quantization in PCM,” IEEE Trans. Information Theory, vol. 28, no. 2, pp. 129-137, March 1982.
- P. A. Chou, T. Lookabaugh, R. M. Gray, “Entropy-constrained vector quantization,” IEEE Trans. Signal Processing, vol. 37, no. 1, pp. 31-42, January 1989.