无失真信源编码

- **信源编码方法**
  - 无失真
    - 匹配编码
  - 有失真
    - 变换编码
    - 识别编码
- **码的分类**
  - 分组/非分组
  - 奇异码/非奇异码（一种分组码中所有码字都不同）
  - 唯一可译码/非唯一可译码
  - 即时码/非即时码
- **定理**
  - 定长码
    - 判断唯一可译码是即时码
      - 任何一个码字都不是其他码字的前缀
    - 存在唯一可译码的条件
      - $q \le r^l \quad r$ 表示码符号集中的码元数. l 定长码长
    - 提高信息传输率的方法
      - 对S扩展信源编码
      - 对概率非常小的不编码
    - 定长编码定理
      - $DMC, H(S)$, 若对长为 N 的序列编码, X 中有 r 个码符号, 码长为 $\ell, \forall \epsilon > 0, \frac{\ell}{N} \ge \frac{H(S)+\epsilon}{log_2 r}$ 满足，当 N 足够大时，几乎可以无失真编码
      - 给出了每个信源符号最少所需要的码符号的理论极限
    - 编码效率
      - $\eta = \frac{H(S)}{R'} = \frac{H(S)}{\frac{\ell}{N} log_2 r}$
      - $\eta = \frac{H(S)}{H(S)+\epsilon}$
      - $\epsilon = \frac{1-\eta}{\eta} H(S)$
    - 对于方差和信源熵，N，η，δ已知
      - $N \ge \frac{D[I(s_i)]}{\epsilon^2 \delta} = \frac{D[I(s_i)]}{H^2(S)} \frac{\eta^2}{(1-\eta)^2 \delta}$
      - $D[I(s_i)] = \sum p(s_i)[-log_2 p(s_i)]^2 - H^2(S)$
      - 表示允许错误概率δ越小，编码效率越高，N就必须长
  - 变长码
    - 即时码
      - Kraft inequality
    - 唯一可译码
      - McMillan inequality
      - 判别准则
    - 紧致码平均码长界限定理
      - $\frac{H(S)}{log_2 r} \le \bar{L} < 1 + \frac{H(S)}{log_2 r}$
    - Shannon First Therom
      - 无失真变长信源编码定理
      - $\frac{H(S)}{log_2 r} \le \frac{\bar{L}_N}{N} < \frac{1}{N} + \frac{H(S)}{log_2 r}$
    - 编码方法
      - Shannon编码
      - Shannon-Fano编码
      - Fano编码
      - Huffman编码
      - LZW编码
      - 游程编码
    - 编码效率
    - 剩余度