

第3章 通信信源与信源编码

3.1 信源特性及信源分类

3.1.1 信源的基本特性

通信信源是通信系统的源端，它所产生的输出信号是一种随机过程。

信源可分为有记忆和无记忆两大类，也可以分为连续和离散两大类。

无记忆信源在不同时刻所产生的输出是相互独立的，它在任意不同时刻的两个样点都是两个互不相关的随机变量。

有记忆信源在不同时刻所产生的输出值是相互关联的，当前输出值的概率分布与前面出现的值有关，例如：那些可用马尔可夫过程、短时平稳过程或循环平稳过程描述的信号，都可以看作是由有记忆信源所产生的输出。

通信信源所产生的输出是用来携带信息进行通信的；因此无论是离散信源还是连续信源，一般都是利用收发两端共知的一个由有限个信源字符(letter 或 symbol)组成的字符集(alphabet)来表示信息的；按表达信息内容的需要，一次又一次地从字符集中选择一个适当的字符作为输出来携带信息进行传输。这些信源字符的传输和存储，一般采用某种物理量(例如：声、光、电)的一组数值或一种波形来表现，因此有些表现为离散信源，有些表现为连续信源。

3.1.2 离散信源

一个离散信源所发出的信源字符，一般属于一个字符数目有限的字符集，例如汉字的国标一级字表，共有 3763 个常用汉字。一个通信系统的发端和收端都利用同一个字符集，以便进行编码和译码。

无记忆离散信源的统计特性可以用字符集中各种字符出现的概率分布来描述；当它为均匀分布时携带信息的效率最高。

有记忆离散信源当前输出什么字符，与过去或将来输出的字符统计相关。这种相关性在时域表现为马尔可夫性，在频域表现为功率密度谱不平坦，即自相关函数不是 δ 函数。有记忆信源的统计特性的描述，需要采用多维随机变量的联合概率分布及自相关函数来描述。信源的记忆性或相关性反映信息的冗余性，因此同样大小信源字符集的信源，无记忆信源携带信息的效率高于有记忆信源。

3.1.3 连续信源

连续信源产生的输出是一种连续随机过程，如果它是频带有限的，那么可以以适当的采样间隔采样为离散随机过程。

3.1.4 采样定理

(1) 随机过程的采样

如果随机过程 $X(t)$ 是频带有限的，设带宽小于等于 W ，那么 $X(t)$ 可以用一个以随机变量序列 $\{X(n/2W), n = -\infty, \dots, -1, 0, 1, \dots, \infty\}$ 为系数的级数表示：

$$\begin{aligned} X(t) &= \sum_{n=-\infty}^{\infty} X(n/2W) \text{Sinc}(2\pi W(t - n/2W)) \\ &= \sum_{n=-\infty}^{\infty} X(n/2W) \frac{\sin[2\pi W(t - n/2W)]}{2\pi W(t - n/2W)} \end{aligned} \quad (3-1-1)$$

其中 $\{X(n/2W)\}$ 为 $X(t)$ 以时间间隔 $T = 1/(2W)$ 采样得到的样点序列；它是 $X(t)$ 在正交基函数族 $\{\text{Sinc}(2\pi W(t - n/2W))\}$ ； $n = -\infty, \dots, -1, 0, 1, \dots, \infty$ 上的投影。

(2) 确知信号的采样

【采样定理 1】 设连续信号 $x(t)$ 的频谱 $X(f)$ 是频带有限的，即当 $|f| \geq f_c$ 时 $X(f) = 0$ ；那么对它进行均匀间隔采样得到离散信号 $\{x(n)\}$ ，如果采样间隔 T 满足 $T < 1/(2f_c)$ ，则由 $\{x(n)\}$ 可以精确地重构 $x(t)$ ，即：

$$x(t) = \sum_{n=-\infty}^{+\infty} x(n) \frac{\sin \frac{\pi}{T}(t - nT)}{\frac{\pi}{T}(t - nT)} \quad (3-1-2a)$$

并且其频谱 $X(f)$ 也可由离散信号 $\{x(n)\}$ 或其频谱 $\check{X}(f)$ 唯一确定，即

$$X(f) = T \sum_{n=-\infty}^{+\infty} x(n) e^{-j2\pi fnT} = \check{X}(f) \quad f \in \left[\frac{-1}{2T}, \frac{1}{2T} \right] \quad (3-1-2b)$$

【采样定理 2】 以 T 为采样间隔采样连续信号 $x(t)$ 得到的离散信号 $\{x(n)\}$ ，无论是否满足采样定理 1 的条件，二者的频谱都存在如下关系：

$$\check{X}(f) = \sum_{m=-\infty}^{+\infty} X(f + \frac{m}{T}) \quad (3-1-3)$$

这个定理的(3-1-3)式给出了因不满足采样定理 1 而产生频谱混叠失真 (aliasing) 的规律。

【推论】设采样率为 f_T 的离散信号 $\{\ddot{x}(n)\}$ 的频谱为 $\ddot{X}(f)$ ，对 $\{\ddot{x}(n)\}$ 进行 $q:1$ 的下采样后得到采样率为 f_T/q 离散信号 $\{\dot{x}(m)\}$ ，其频谱为 $\dot{X}(f)$ ，则这两个频谱总存在以下关系：

$$\dot{X}(f) = \sum_{m=0}^{q-1} \ddot{X}(f + mf_T/q) \quad (3-1-4)$$

$\ddot{X}(f)$ 的周期为 f_T ，而 $\dot{X}(f)$ 的周期为 f_T/q 。

【采样定理 3】设 $x(t)$ 是一个下界频率为 f_L 和上界频率为 f_H 的带通实信号，且满足 $f_L > f_H/2$ ；若采样频率 $f_T = 1/T$ 满足

$$\frac{2f_H}{m+1} < f_T < \frac{2f_L}{m} \quad m = 1, 2, 3, \dots \quad (3-1-5)$$

其中 m 为某个合适的正整数，那么 $x(t)$ 可用其采样信号 $\{\ddot{x}(n)\}$ 和 f_L 精确地重构而无混叠失真；如果 $x(t)$ 是一个频带为 (f_L, f_H) 的带通解析信号，那么重构信号不出现混叠失真的充分必要条件是： $f_T > (f_H - f_L)$ 。

定理证明的要点：

1) 对 $x(t)$ 以时间间隔 T 进行采样，相当于用一个周期为 T 周期性的 Kroneker δ 函数与它相乘，频域为二者之卷积；而周期性的 Kroneker δ 函数的频谱函数是一个周期为 $1/T$ 的 δ 函数。要想使二者的卷积不出现频谱混叠现象，只要在卷积过程中任何时刻都没有两个 δ 脉冲同时对准 $x(t)$ 的正频率频谱和它的镜像频谱就行。

2) 如果对于 $x(t)$ 的复解析信号 $\mathbf{x}(t)$ 进行采样，即正交采样，则只要采样频率大于 $1/T$ ，在卷积过程中就没有两个 δ 脉冲同时对准频谱的现象，因为它没有镜像频谱。

3) 当带通实信号的边界频率不满足 $f_L > f_H/2$ 时，就找不到比 $2f_H$ 还低的采样频率，因此不能采用上述定理。

采样信号的频谱是周期为 f_T 的周期性频谱，每个周期中包含有一个正频率频谱 $X(f)$ 的复本和它的镜像的复本，如图 3-1-1 所示。

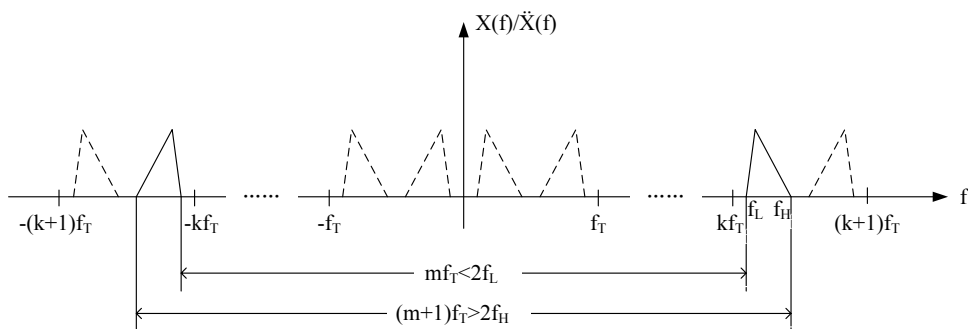


图 3-1-1 带通信号及其采样信号的频谱

3.2 信息的度量和信源熵

3.2.1 信息量的度量单位

● 信息的度量和单位:

因此用 $\text{Log}(1/P_i)$ 来衡量信源字符集中第 i 个信源字符所携带的信息量。

当 $\text{Log}_2(1/P_i)=1$ 时, 该字符所携带的信息量定义为 **1 比特(bit)**;

当 $\text{Ln}(1/P_i)=1$ 时, 该字符所携带的信息量定义为 **1 奈特(Nat)**;

当 $\text{Log}_{10}(1/P_i)=1$ 时, 该字符所携带的信息量定义为 **1 哈特(Hart)**;

$1\text{nat}\approx 1.44\text{bit}$, $1\text{Hart}\approx 3.32\text{bit}$; 数字通信中通常以二进制为基础, 因此采用比特作为信息量单位最为常用。

● 互信息:

设两个离散随机变量 X 和 Y 的取值范围分别为 $\{x_1, x_2, \dots, x_n\}$ 和 $\{y_1, y_2, \dots, y_m\}$, 其概率分布分别为 $\{P(x_1), P(x_2), \dots, P(x_n)\}$ 和 $\{P(y_1), P(y_2), \dots, P(y_m)\}$; 如果二者存在相关性, 定义 $P(x_i|y_j)$ 为在出现事件 $Y=y_j$ 的条件下事件 $X=x_i$ 出现的概率, 那么事件 $Y=y_j$ 的出现提供关于事件 $X=x_i$ 的信息量定义为:

$$I(x_i; y_j) = \text{Log}[P(x_i | y_j) / P(x_i)] \quad (3-2-1)$$

$I(x_i; y_j)$ 称为 $X=x_i$ 和 $Y=y_j$ 之间的互信息。

● 平均互信息:

离散随机变量 X 和 Y 之间的平均互信息为:

$$I(X; Y) = \sum_{i=1}^n \sum_{j=1}^m P(x_i, y_j) I(x_i; y_j) = \sum_{i=1}^n \sum_{j=1}^m P(x_i, y_j) \cdot \text{Log} \frac{P(x_i, y_j)}{P(x_i)P(y_j)} \quad (3-2-2)$$

X 和 Y 之间的平均互信息可以用条件熵来表示:

$$I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \quad (3-2-3)$$

3.2.2 信源熵

一个信源的熵定义为该信源每发送一个字符所携带信息量的平均值。

(1) 无记忆离散信源的熵

如果一个信源共有 n 个信源符号 $\{x_1, x_2, \dots, x_n\}$, 各个符号的概率分布为

$P(X = x_i) = p_i \quad i = 1, 2, \dots, n$ ，并且 $p_1 + p_2 + \dots + p_n = 1$ ；那么其熵为：

$$H(X) = -\sum_{i=1}^n p_i \log p_i \quad (3-2-4)$$

(2) 有记忆信源的熵

有记忆随机过程的统计特性要用联合概率来描述，因此记忆长度为 L 的信源，其熵也应该用 L 维联合概率分布函数 $P(X_1, X_2, \dots, X_L)$ 来计算，则相邻 L 个字符的熵为：

$$H(X_1, X_2, \dots, X_L) = -\sum P(X_1, X_2, \dots, X_L) \log P(X_1, X_2, \dots, X_L) \quad (3-2-5)$$

平均每个字符的熵为：

$$H_L(X) = H(X_1, X_2, \dots, X_L) / L \quad (3-2-6)$$

如果把前面 $L-1$ 个随机变量作为条件，则可求得最后一个变量的熵，称为条件熵 $H_L(X_L | X_1, X_2, \dots, X_{L-1})$ 。那么， $H_L(X)$ 可表示为：

$$H_L(X) = [H(X_1) + H(X_2 | X_1) + H(X_3 | X_1, X_2) + \dots + H_L(X_L | X_1, X_2, \dots, X_{L-1})] / L \quad (3-2-7)$$

条件概率空间的概率分布一般要比无条件概率空间的概率分布更集中，因此条件熵小于无条件熵；条件越多其熵值越小，再考虑到平稳性，则有如下递增关系：

$$\begin{aligned} H_L(X_L | X_1, X_2, \dots, X_{L-1}) &\leq H(X_L | X_2, X_3, \dots, X_{L-1}) = H(X_{L-1} | X_1, X_2, \dots, X_{L-2}) \\ &\leq H(X_{L-1} | X_2, X_3, \dots, X_{L-2}) = H(X_{L-2} | X_1, X_2, \dots, X_{L-3}) \\ &\leq H(X_{L-2} | X_2, X_3, \dots, X_{L-3}) = H(X_{L-3} | X_1, X_2, \dots, X_{L-4}) \\ &\dots\dots \\ &\leq H(X_2 | X_1) \leq H(X_1) \end{aligned} \quad (3-2-8)$$

由此可见(3-2-7)式给出的平均每个字符的熵 $H_L(X)$ 的值，介于无条件熵 $H(X_1)$ 与条件熵 $H_L(X_L | X_1, X_2, \dots, X_{L-1})$ 之间。

(3) 离散马尔可夫信源的熵

如果一个信源字符集大小为 n 的一阶马尔可夫信源所产生的输出是一个各态历经的一阶平稳马尔可夫链，那么其极限分布(即各个信源字符出现的概率)可由下述 K-C 方程求得：

$$p_j = \sum_{i=1}^n p_i p_{ij} \quad j = 1, 2, \dots, n \quad (3-2-9)$$

其中 $\{p_{ij}, i, j = 1, 2, \dots, n\}$ 为状态转移概率矩阵， p_{ij} 为前一步处于状态 i 而当前一步转移到状态

j 的概率。于是，这个一阶马尔可夫信源的熵为：

$$H(X) = H(X_2 | X_1) \quad (3-2-10)$$

(4) 连续信源的差熵

连续随机变量不能像离散随机变量那样定义自信息，因而也不能相应地定义熵，但可类似地定义一种熵，称为差熵：

$$H(X) = - \int_{-\infty}^{\infty} p(x) \text{Log} p(x) dx \quad (3-2-11)$$

(5) 连续信源的平均互信息

设连续随机变量 X 和 Y 的联合 PDF 为 $p(x, y)$ ，它们的边沿 PDF 分别为 $p(x)$ 、 $p(y)$ ，那么 X 和 Y 之间的平均互信息定义为：

$$I(X; Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x) p(y|x) \text{Log} \frac{p(x)p(y|x)}{p(x)p(y)} dx dy \quad (3-2-12)$$

Y 已知时 X 的平均条件熵定义为：

$$H(X|Y) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) \text{Log} p(x|y) dx dy \quad (3-2-13)$$

于是有

$$I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \quad (3-2-14)$$

3.3 离散信源信息的编码

3.3.1 定长编码

一个字符集大小为 L 的无记忆离散信源 X ，当各字符的概率相等时其熵的值最大，即：

$$H(X) \leq \text{Log} L \quad (3-3-1)$$

采用固定长度码字，对于字符集大小为 L 的无记忆离散信源所发出的字符序列逐个地进行无失真编码，则每个码字的长度需要 r 比特：

$$r = \text{Int}[\log_2(L-1)] + 1 \quad (3-3-2)$$

这里 Int 表示舍尾取整，因此应取 r 比特才能确保 L 种字符都有对应的码字，即 $2^r \geq L$ 。

【离散信源定长编码定理】对于一个信源熵等于 $H(X)$ 比特的无记忆离散平稳信源 X ，可以将信源发出的字符序列每 J 个字符分为一组，用 N 比特长的码字进行编码，对于任意小的正数 ϵ ，只要 J 选择的足够大，并使码率 R 满足

$$R \equiv \frac{N}{J} \geq H(X) + \varepsilon \quad (3-3-3)$$

就可以保证编码失败的码组出现的概率 P_e 可以任意小；反之，若码率

$$R \equiv \frac{N}{J} \leq H(X) - \varepsilon \quad (3-3-4)$$

则编码失败码组出现的概率 P_e 将随 J 的增大而趋于 100%。

此定理的含义是：对一个信源发出的符号序列进行定长分组编码，只要码率大于信源熵，码组又足够长，则无失真编码失败的概率可以无穷小；而若码率小于信源熵，无失真编码失败的可能性非常大。

【例】将一本中文书看作一个信源，如果所用汉字都包含在国标一级字表的 3763 个汉字之中，则国标一级字表就是信源字符集。将书中的汉字用代码表示，本来需要用 12 比特长的代码才能表示一个汉字。但是，假如根据该书中各种汉字出现的概率求得的信源熵 $H(X)=7.9$ 比特/字符，而将书中每 J 个汉字看作一个码组用一个 N 比特的代码表示，那么只要 $N/J > 7.9$ ，即平均每个汉字 8 比特以上，就有可能成功表示各种可能的码组，但要使表示失败的概率任意小，则要求 J 足够大。

3.3.2 变长编码

对于非等概率分布的无记忆离散信源的编码，采用可变长度的码字进行编码可获得更高的编码效率，但其码字必须满足前缀条件，即任意一个码字不能是任何其它码字的前缀。

● Kraft 不等式

一个包括有 L 个码字且都满足前缀条件的码字集是否存在，其充分和必要条件是：

$$\sum_{k=1}^L 2^{-n_k} \leq 1 \quad (3-3-5)$$

其中 n_1, n_2, \dots, n_L 分别为各个码字的长度。

● 离散信源变长编码定理

【定理 1】若一离散无记忆信源的熵为 $H(X)$ ，每个信源字符用 m 进制码字进行变长编码，则一定存在一种无失真编码方法其码字平均长度 \bar{K} 满足不等式：

$$\frac{H(X)}{\log m} \leq \bar{K} < \frac{H(X)}{\log m} + 1 \quad (3-3-6)$$

【定理 2】对于一个字符熵为 $H(X)$ 的离散无记忆信源进行变长编码，必定存在一种无失真编码方法，其平均信息率 R 满足不等式：

$$H(X) \leq R < H(X) + \varepsilon H(X) \quad (3-3-7)$$

其中 ε 为任意小的正数。

此两个定理的含义是：采用不定长的 m 种码字对于每 J 个字符构成的一个码组进行编码，只要平均信息率（即平均每个字符的比特数） $R > H(X)$ ，就可以确保编码成功。

【例】仍以一本中文书中汉字编码为例，因为信源熵是 $H(X)=7.9$ 比特/字符；但因包含 J 个汉字的各种可能码组出现的概率不同，我们采用 m 种比特数不同的码字表示各个码组，概率高的用短码，概率低的用长码；那么只要码率（平均每个汉字编码的比特数）大于 $H(X)$ ，就可以确保编码成功。

3.3.3 Huffman 编码算法

设一个离散信源共有 n 个信源字符 $\{x_1, x_2, \dots, x_n\}$ ，字符的概率分布为： $P(X=x_i)=p_i \quad i=1, 2, \dots, n$ ，并且 $p_1+p_2+\dots+p_n=1$ ；则采用 Huffman 树形结构算法可以生成一个非等长的二进制码字集 $\{c_1, c_2, \dots, c_n\}$ ，即码书，使码字集中各码字分别与信源字符集相对应，即 c_i 对应于 x_i 。该码书生成的步骤如下：

① 比较 n 个概率值 $\{p_1, p_2, \dots, p_n\}$ 的大小，将概率最小的两个字符，例如 x_i 和 x_j 所对应码字 c_i 和 c_j 的最后一个比特分别指定为 0 和 1；

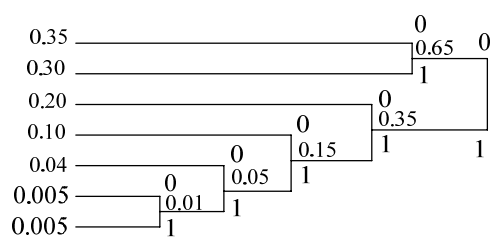
② 将这两个最小概率值合并，即 $p_{ij}=p_i+p_j$ ，并用 p_{ij} 取代 p_i 和 p_j ，得到树中第 1 个节点。

③ 将 p_{ij} 与其它 $n-2$ 个概率值进行比较，重新找出两个概率最小的，如果这两个概率值所对应的信源字符，例如是 x_i 和 x_j ，而不是前述的 x_i 和 x_j ，则将这两个信源字符所对应的码字 c_i 和 c_j 的最后 1 比特分别指定为 0 和 1；如果 x_i 与前述两个字符中的一个相同，例如 $x_i=x_i$ ，即曾经分配过 1 比特，则将 c_i 的倒数第 2 比特指定为 0 或 1。

④ 再次将两个新的最小概率值合并，得到树中一个新的节点，将此概率值与其它的 $n-3$ 个（也可能还有 $n-2$ ）概率值进行比较，重新找出两个概率最小的，如果这两个概率值所对应的信源字符不包括在前面已分配过比特的码字，则将这两个码字的最后 1 比特分别指定为 0 和 1；如果这两个码字中有一个（或两个）曾经分配过比特，则将它所涉及的码字的前一比特指定为 0 或 1。

如此重复进行，直到所有的概率值合并为 1，由此得到长短不一的 n 个码字，即构成一个 Huffman 树形码书。

【例】由 7 个字符构成的信源字符集，各字符出现概率分别为 0.005, 0.005, 0.04, 0.1, 0.2, 0.3, 0.35；那么按照上述编码方法设计 Huffman 码书的 7 个码字分别为：11111, 11110, 1110, 110, 10, 01, 00。



字	概率	自信	码
x_2	0.30	1.7370	01
x_3	0.20	2.3219	10
x_5	0.04	4.6439	1110
x_6	0.005	7.6439	1111
x_7	0.005	7.6439	1111
		$H(X)=2.11$	$\bar{R}=2.21$

图 3-3-1 Huffman 编码举例

3.3.4 Lempel-Ziv 编码算法

$L-Z$ 算法^[6,7]是一种不涉及信源统计特性的通用编码算法,它是一种可变—固定码长的无失真编码方法,广泛用于计算机文件压缩。

● 算法简述

$L-V$ 算法是一种由变长到定长的编码,待编码的信源符号序列被划分为许多长短不同的片段(phrases),并同时自动产生一个字典,然后用这个字典中的码字对于各个片段进行编码,接收端采用同样的字典进行译码。

设片段长度的最大值预定为 L_s , 缓存器的长度为 n ; 则编码器采用固定长度 L_c 的码字对各个片段进行编码时,码长 L_c 应满足下式才能完全成功编码而又不产生缓存器溢出:

$$L_c = 1 + \text{Int}(\text{Log}(n - L_s)) + \text{Int}(\text{Log}L_s) \quad (3-3-8)$$

3.4 无记忆序列的压缩编码

一个连续随机过程,在按照采样定理离散化为样点序列或矢量序列后,各个样点或矢量的分量都还是模拟量,要表示为有限字长的数字需要进行量化和编码;如何进行量化和压缩编码才能效率高而代价小,需要定义失真测度和率失真函数来衡量失真代价的大小。

3.4.1 率失真函数

● 率失真函数的定义

一个无记忆信源 X , 设 x 表示该信源输出的某个样点的值, x' 表示 x 经编码后恢复的值, $d(x, x')$ 表示二者的相对失真量, $E[d(x, x')]$ 为编码大量样点时每个样点所产生的平均失真量。为了以小于或等于平均失真量 D 对这个信源的输出进行编码,平均每个信源符号所需的最低比特率 $R(D)$, 称为率失真函数(Rate-distortion Function):

$$R(D) = \min_{p(x'|x): E[d(x, x')] \leq D} I(X, X') \quad (3-4-1)$$

其中 $I(X, X')$ 是信源原始输出和信源编码结果之间的互信息。

对于无记忆矢量序列中矢量的编码，同样可以定义率失真函数，矢量 \mathbf{x} 经编码再恢复后的矢量为 \mathbf{x}' ，其失真量为 $d(\mathbf{x}, \mathbf{x}')$ 。

● 无记忆高斯信源的率失真函数

如果用每个符号的均方误差来度量失真，那么表示一个时间离散、幅度连续、无记忆高斯信源的输出所需的最低信息速率应为：

$$R_g(D) = \begin{cases} \frac{1}{2} \log_2(\sigma_x^2 / D) & \text{当 } 0 \leq D \leq \sigma_x^2 \\ 0 & \text{当 } D \geq \sigma_x^2 \end{cases} \quad (3-4-2)$$

【限失真的信源编码定理】 对于任意给定的失真 D ，一定存在一种编码方式，其最小信息速率为 $R(D)$ 比特/符号，而译码恢复信源输出时所产生平均每个符号的失真值可以任意逼近 D 。

【定理— $R(D)$ 的上界】 对于零均值、有限方差 σ_x^2 （采用均方差失真测度）、幅度连续的无记忆信源，其率失真函数的上界为：

$$R(D) \leq \frac{1}{2} \log_2(\sigma_x^2 / D) \quad (0 \leq D \leq \sigma_x^2) \quad (3-4-3)$$

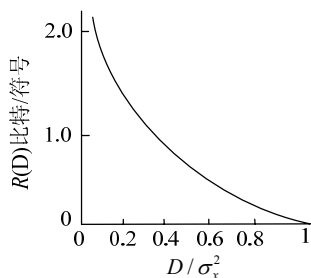


图 3-4-1 连续幅度无记忆高斯信源的率失真函数

这个上界就是高斯信源的率失真函数，如图 3-4-1 所示。本定理表明：在指定均方误差的限制下，表示一个高斯信源输出信息所需的信息速率是最高的，比其它任何信源的都高。

3.4.2 采样信号的量化编码

连续信号 $x(t)$ 的采样通常采用固定的时间间隔，采样值被量化为 $L = 2^K$ 种电平 $\{x_1, x_2, \dots, x_L\}$ ，那么每一个样点值可以用一个 K 比特的二进制数表示；如果输入样点值 $x(t)$ 与 x_k 之间失真测度值 $d(x(t), x_k)$ 最小，就将 $x(t)$ 量化为 x_k ，而将其下标 k 作为其编码。

这里 $x(t)$ 的采样点值都是标量，因此这种量化称为标量量化。设量化器的输入、输出样点值分别为 x 、 x' ，而 x 的 PDF 为 $p(x)$ ，那么信号值量化导致的平均失真量为：

$$D = \int_{-\infty}^{\infty} d(x, x') \cdot p(x) dx \quad (3-4-4)$$

(1) 均匀量化

如果量化器的 L 个量化电平 $\{x_1, x_2, \dots, x_L\}$ 是均匀分布的，即相邻两个量化值之差 Δ 固定不变，相应的各量化区间也是等长的；第 k 个量化值所对应的量化区间为 $((x_{k-1} + x_k)/2, (x_k + x_{k+1})/2)$ ，那么落在第 k 个区间的信号样点值都量化为 x_k ，并编码为 k ，这就是均匀量化。

● 均匀量化器输出信号的量化信噪比

既然无论 x 落在量化区间的哪一点，量化误差 $q = x' - x$ 都限制在 $\pm \Delta/2$ ，当 Δ 足够小时， q 在区间 $[-\Delta/2, \Delta/2]$ 中的分布近似为均匀分布，可由下式计算因量化产生的噪声的均方误差：

$$\sigma_q^2 = \int q^2 \frac{1}{\Delta} dq = \Delta^2 / 12 \quad (3-4-5)$$

若设量化器的量化范围为 $(-V, V)$ ，那么量化间距可表示为：

$$\Delta = 2V / L = 2V / 2^K \quad (3-4-6)$$

其中 K 为量化编码的比特数， $L = 2^K$ 。设信号的均方差值为 σ_x^2 ，那么量化信噪比为：

$$SNR = \frac{\sigma_q^2}{\sigma_x^2} = 3L^2 \frac{V^2}{\sigma_x^2} = 3 \times 2^{2K} \frac{V^2}{\sigma_x^2} \quad (3-4-7)$$

表示为分贝：

$$SNR(dB) = 6.02K + 4.77 - 10 \log(V^2 / \sigma_x^2) \quad (3-4-8)$$

由此可见量化信噪比与输入信号的 PDF 无关，而只与其峰平功率比 V^2 / σ_x^2 以及 K 有关。

(2) 最佳标量量化

【定理】最佳量化器的必要条件是：①最佳划分遵循最邻近规则；②各个分划所对应的最佳量化值就是各个分划的质心。

当量化器的 L 个量化值 $\{x_1, x_2, \dots, x_L\}$ 不是等阶距，而是根据输入信号值 $x(t)$ 的 PDF 设计，满足以下两个必要条件：

① L 个量化值所对应的量化区间为 $(\frac{x_{k-1} + x_k}{2}, \frac{x_k + x_{k+1}}{2})$ ， $k = 1, 2, \dots, L$ ；

② 每个量化值等于相应量化区间的质心值： $x_k = \int_{(x_{k-1} + x_k)/2}^{(x_k + x_{k+1})/2} x \cdot p(x) dx$ ； $k = 1, 2, \dots, L$

则其平均量化失真最小，而称为最佳量化器。

最佳量化器输出信号的平均量化失真值为：

$$D = \sum_{k=1}^L \int_{x_{k-1}}^{x_k} f(x'_k - x) p(x) dx \quad (3-4-9)$$

其中 $f()$ 表示某种失真测度，例如 $f(x'_k - x) = |x'_k - x|^2$ 。

例如：语音信号采用 8 比特对数 PCM 量化结果，其量化信噪比相当于 13 比特线性 A/D 变换所得结果。

● 设计最佳量化器的迭代算法步骤

① 初始化：将所有的训练样本按 L 电平均匀量化规则，以 $\{x_1, x_2, \dots, x_{L+1}\}$ 为分界点划分为 L 类；

② 每类求出一个平均值，即得到各区间的质心 $\{q_1, q_2, \dots, q_L\}$ ；

③ 以 $\{q_1, q_2, \dots, q_L\}$ 作为各类的质心，按照最小平方误差准则对所有训练样本进行重新分类，并累计总的量化失真量（误差平方和）。如果本次迭代的总量化失真比前一次迭代的总量化失真显著减小或者尚未达到预定的精度或迭代次数，则转向②继续迭代；否则停止迭代，所得到的 L 个量化值 $\{q_1, q_2, \dots, q_L\}$ 就是最终结果。

(3) 脉冲编码调制 (PCM)

PCM 应该包括抽样、量化和编码成为数据比特流，它是一种将信源看成无记忆信源逐个样点进行量化编码的方法。

以语音信号的抽样量化为例，采用非均匀量化，其信噪比可显著提高，例如：8 比特量化精度的对数规律非均匀量化器，其量化信噪比与 13 比特均匀量化相似；国际标准对数 PCM 有 μ -律和 A -律两种稍微不同的规范， μ -律 PCM 中输入和输出的关系为：

$$y = x_{\max} \cdot \frac{\text{Ln}(1 + \mu \cdot |x| / x_{\max})}{\text{Ln}(1 + \mu)} \cdot \text{sgn}(x) \quad (3-4-10)$$

其中 μ 值的取值范围可从 10 至 100， μ 值决定它坐标轴的压缩程度。

3.4.3 矢量量化

(1) 矢量量化的基本概念

矢量量化，简称 VQ，是将每个矢量看作一个整体，基于矢量间失真测度值最小准则来进行近似表示矢量的一种压缩编码的方法，其框图如图 3-4-2 所示。

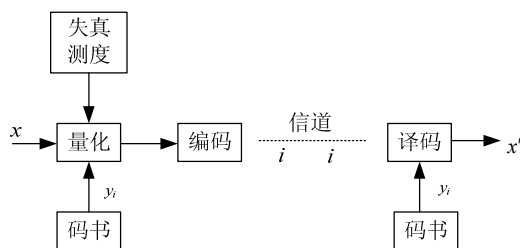


图 3-4-2 矢量量化编码/译码器

对于 N 维的矢量 $\mathbf{x} = [x_1, x_2, \dots, x_N]^T \in \mathcal{R}^N$ 进行矢量量化, 欲将每个矢量编码为 K 比特, 则先设计一个含有 $L (=2^K)$ 个码字的码书 $\mathcal{C}: \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_i, \dots, \mathbf{y}_L\} \in \mathcal{R}^N$, 然后根据使某种失真测度 $d(\mathbf{x}, \mathbf{y}_i)$ 值最小的准则, 从码书中选择一个码字 \mathbf{y}_i 来近似表示 \mathbf{x} , 而将 \mathbf{y}_i 的下标 i 作为 \mathbf{x} 的编码, 即:

$$c(\mathbf{x}) = \hat{i} = \underset{i = 1, 2, \dots, N}{\text{Arg min}} d(\mathbf{x}, \mathbf{y}_i) \quad (3-4-11)$$

接收端根据收到的编码 \hat{i} , 从码书 \mathcal{C} 中找出 \mathbf{y}_i 用于近似代替和恢复原矢量 \mathbf{x} , 即 $\hat{\mathbf{x}} = \mathbf{y}_i$ 。

(2) VQ 编码效率很高的原因

与对矢量的每一维分量分别进行标量量化编码相比, 采用矢量量化编码的效率一般都要高得多, 即达到同样的平均失真量所需的编码比特率要低得多; 这是因为 VQ 在进行压缩编码时能利用矢量的以下特征:

- ① 矢量信源多维概率密度函数的形状或概率分布;
- ② 矢量各分量之间的线性相关性和统计不独立性;
- ③ 多维空间中胞腔的形状及其多变性;
- ④ 便于选择一种主观感觉最好的多维空间失真测度来进行 VQ。

(3) VQ 码书设计算法

关于最佳量化器两个必要条件可推广到矢量量化码书的设计。对于一个给定训练矢量序列或矢量的概率密度函数, 交替地运用这两个必要条件, 就可以得到一个准最佳的 VQ 码书。

设给定一个训练矢量集 \mathcal{Z} 中共有 L_z 个 N 维矢量, $\mathcal{Z}: \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{L_z} \in \mathcal{R}^N\}$, 用它来设计一个含有 L 个 N 维码字矢量的码书 $\mathcal{C}: \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_i, \dots, \mathbf{y}_L\} \in \mathcal{R}^N$, 一般要求 $L_z \geq (50 \sim 100) \cdot L$, 码书大小 $L = 2^K$, K 是正整数。

LBG 算法码书设计步骤如下:

- ① 先求 \mathcal{Z} 中所有矢量的质心, 例如求均值矢量 $\bar{\mathbf{v}} = \sum_{n=1}^{L_z} \mathbf{v}_n / L_z$ 看作为质心, 然后采用

扰动法将 $\bar{\mathbf{v}}$ 分裂成两个码字: $\mathbf{y}_1^{(1)} = \bar{\mathbf{v}} + \boldsymbol{\delta}$, $\mathbf{y}_2^{(1)} = \bar{\mathbf{v}} - \boldsymbol{\delta}$; 其中 N 维矢量 $\boldsymbol{\delta}$ 的 N 个分量都是很小的随机数; 将这两个码字矢量看作是一个大小为 \tilde{L} ($=2$) 的初始码书 $\mathbb{C}^{(m)}: \{\mathbf{y}_i^{(m)}, i=1, \dots, \tilde{L}\}$, $m=1$; 令 $D^{(m)}=0$;

② 用上一步得到的 \tilde{L} 大小的码书 $\mathbb{C}^{(m)}$, 对 \mathbb{Z} 中的所有矢量进行矢量量化, 即基于失真度值 $d(\mathbf{x}, \mathbf{y}_i)$ 最小准则, 将 \mathbb{Z} 中的矢量划分为 \tilde{L} 类; 同时累计量化失真, 得到本轮(第 m 轮)量化的总失真量 $D^{(m)}$;

③ 求出每类矢量的质心矢量, 并用它们构成新码书, 即 $\mathbb{C}^{(m)}: \{\mathbf{y}_i^{(m)}, i=1, 2, \dots, \tilde{L}\}$;

④ 如果总失真的相对变化量 $\gamma \triangleq |D^{(m)} - D^{(m-1)}| / D^{(m)} > 0.005$, 令 $m = m + 1$, 然后转向②; 否则转向⑤;

⑤ 如果 $\tilde{L} = L$ 转向⑦; 否则转向⑥;

⑥ 采用同上的扰动法将上一步的码书 $\mathbb{C}^{(m)}$ 中每个码字分裂成两个, 从而构成一个大小为 $2\tilde{L}$ 的新码书, 并令 $\tilde{L} = 2\tilde{L}$, $m = m + 1$, $D^{(1)} = 0$, 然后转向②;

⑦ 停止迭代, 输出最近第③步产生的最新码书, $\mathbb{C}^{(m)}: \{\mathbf{y}_i^{(m)}, i=1, 2, \dots, L\}$ 。

在 LBG 算法中, 码书的大小从 1 开始逐步加倍增大, 直至希望大小; 增大的办法是分裂法; 码书中的码字是采用迭代法逐步优化的, 直至迭代优化能引起的量化总误差的相对减小很不明显时为止。

(4) 矢量量化的应用

矢量量化在语音信号识别、语音合成、语音编码、图像编码和信道信息压缩编码等方面都得到了很好的应用。

① 用于语音或图像数据的压缩编码:

将一个 N 维矢量当作一个整体看待来定义失真测度, 有时与人的主观感觉能很好地相吻合, 因而使矢量量化在达到很高压缩比的情况下所恢复的信号的主观感觉失真还很小。例如: 描述语音信号频谱的一组参数一般需要十几个实数, 共约 400 比特左右, 每种频谱对应于一个类音素, 而人类听觉系统能区分的类音素种类只有几百中, 用 8 比特或 9 比特就足以表示。因此, 如果所定义的失真测度与主观感觉的差异大小基本吻合的话, 采用 VQ 方法表示一种语音信号的频谱所需数据率可压缩 50 倍以上。目前语音、图像和视频的压缩编码, 其所以效率那么高, VQ 在其中发挥了非常重要的作用。

② 用于通信信道参数的压缩编码:

在基于多天线的多输入/多输出系统或多载波调制通信系统中, 常常需要由接收端向发射端反馈信道状态信息, 描述信道状态的矢量也具有很好的空间集聚性, 可以采用 VQ 技术大幅度降低传输它们时所需的数据率。

3.5 有记忆序列的压缩编码

采用有记忆序列的压缩编码方法可以获得更高的编码效益。

3.5.1 线性预测技术

(1) 线性预测的基本原理

将连续信号看成是有记忆信源所产生的输出，按照奈奎斯特准则采样离散化为样点序列 $\{x(n)\}$ ，那么这种随机序列是有记忆的，如果这种相互关联性可假定为是 p 阶线性相关的，任意时刻的样点 $x(n)$ 可以用其前 p 个样点值的加权和近似表示：

$$x(n) \approx \sum_{i=1}^p a_i x(n-i) \triangleq \hat{x}(n) \quad n = -\infty, \dots, -1, 0, 1, \dots, \infty \quad (3-5-1)$$

这就是线性预测；其中预测系数 $\{a_i\}$ 可以根据最小均方误差准则求得；先写出预测误差表达式：

$$e(n) = x(n) - \hat{x}(n) = x(n) - \sum_{i=1}^p a_i x(n-i) \quad (3-5-2a)$$

此式也可看作是 $\{x(n)\}$ 通过一个预测误差滤波器 $A(z)$ 所产生的输出为 $\{e(n)\}$ ，这里

$$A(z) = 1 - \sum_{i=1}^p a_i z^{-i} \quad (3-5-2b)$$

采用大量的样本对于这个预测误差能量进行统计，得到总的预测误差能量为：

$$E = \sum_n [e(n)]^2 = \sum_n [x(n) - \sum_{i=1}^p a_i x(n-i)]^2 \quad (3-5-3)$$

为求误差能量 E 的极小值，令 $\partial E / \partial a_i = 0$ ， $i = 1, 2, 3, \dots, p$ ；即可导出一个以 $\{a_i\}$ 为未知变量的线性方程组(尤拉-瓦尔克方程)：

$$\sum_{k=1}^p a_k R_i(k) = R_i(0) \quad i = 1, 2, \dots, p \quad (3-5-4)$$

其中 $\{R_0(k) \quad k = 0, 1, 2, \dots, p\}$ 为 $\{x(n)\}$ 的自相关函数，而 $R_i(k)$ 是 $R_0(k)$ 循环移位 i 个样点所得复本，即 $R_i(k) = R_0(k')$ ，这里 $k' = (k+i)_{\text{mod } p} + 1$ ； $R_0(k)$ 的 $p+1$ 个值的求法是，通常可以采用一个适当长的窗函数序列，从希望分析时段的 $\{x(n)\}$ 中截取 $2L+1$ 个样点，求其时间自相关值：

$$R_0(k) = \sum_{n=-L+k}^{L-k} [x(n+n_0) \cdot w(n)] \cdot [x(n+n_0+k) \cdot w(n+k)] \quad k = 0, 1, 2, \dots, p \quad (3-5-5)$$

$\{R_0(k) \quad k = 0, 1, 2, \dots, p\}$ 得到后，即可由方程组(3-5-4)解出最佳线性预测系数 $\{\hat{a}_i\}$ ；再将它们用于(3-5-1)式进行线性预测，则其平均预测误差将是最小的；而且由(3-5-2a)式给出的预测误差序列 $\{e(n)\}$ 将变为高斯白噪声序列，其功率比原信号序列 $\{x(n)\}$ 的功率大幅度降低，可以采用较低的比特率量化编码，即使再加上 $\{\hat{a}_i\}$ 的编码数据，所需总的编码比特率比直接对 $\{x(n)\}$ 进行编码所需的比特率要低得多。

由(3-5-2a)式可以直接推出一个能近似地重构原信号样点序列的公式:

$$\hat{x}(n) = \hat{e}(n) + \sum_{i=1}^p \hat{a}_i \hat{x}(n-i) \quad (3-5-6)$$

可用于接收端的信号重构。

(3-5-6)式本质上是一个全极点线性系统 $H(z) = 1 / (1 - \sum_{i=1}^p \hat{a}_i z^{-i})$ 的输入输出关系的差分表达式, 即 $\{\hat{x}(n)\}$ 是以零均值高斯白噪声序列 $\{\hat{e}(n)\}$ 作为输入, 激励线性系统 $H(z)$ 时所产生的输出。显然 $H(z) = 1 / A(z)$ 。

由此得到一个**重要结论**: 如果一个有记忆随机序列 $\{x(n)\}$ 的记忆长度为 p , 其相关性可假定为线性的, 那么 $\{x(n)\}$ 可以看作是由一个零均值高斯白噪声序列激励一个 p 阶全极点线性系统 $H(z) = 1 / (1 - \sum_{i=1}^p \hat{a}_i z^{-i})$ 所产生的输出; 显然 $\{x(n)\}$ 是一种平稳高斯随机过程, 但不是白色的。

这个随机过程的记忆长度 p , 可理解为自相关函数 $R(\tau)$ 具有非零值的最大延迟为 p , 当 $\tau > p$ 时 $R(\tau)$ 几乎等于 0。只要 $R(\tau)$ 已知, 便可直接给出这个随机过程中相邻 $p+1$ 个样点的联合 PDF; 这是因为 $\{x(n)\}$ 的均值也应为 0, 因而协方差矩阵就是自相关矩阵, 所以可利用 $p+1$ 维高斯随机变量的标准表达式给出其 PDF。

由于预测误差序列 $\{e(n)\}$ 是高斯白噪声序列, 其频谱幅度是平坦的, 设其方差为 σ^2 ; 那么所分析的信号段 $\{x(n+n_0).w(n)\}$ 的能量密度谱函数应为

$$|X(\omega)|^2 = \sigma^2 \cdot |H(e^{j\omega})|^2 = \sigma^2 / |A(e^{j\omega})|^2 \quad (3-5-7)$$

因此线性预测技术也可以用作作为一种功率谱估计方法。

(2) 线性预测系数的几种推演参数

定义 $R(k) \triangleq R_0(k) / R_0(0)$; $k = 0, 1, 2, \dots, p$ 为归一化自相关函数; 那么从(3-5-4)所示的方程组可见, 若 $\{R(k) \quad k = 0, 1, 2, \dots, p\}$ 为已知参数, 则可以解得 $\{a_i \quad i = 1, 2, \dots, p\}$, 反之亦然; 这说明这两种参数是等价的。

线性预测系数还可推演出其它几种等价的或近似等价的参数, 例如: 反射系数、倒谱系数、预测器多项式的根、线性预测系数的自相关函数、线谱对参数等。这几种参数虽然近似等价, 都可以用来进行编码和信号重构, 但它们的特性和失真测度可能存在较大的差别。

(3-5-4)式所示的线性方程组有多种解法, 其中最简单的解法是 Levinson-Durbin 递推算法; 给定自相关函数 $\{R(k) \quad k = 0, 1, 2, \dots, p\}$, 递推线性预测系数 $\{a_1, a_2, \dots, a_p\}$ 的程序为:


```

K(0)=R(0);
for(i=1; i<p+1; i++)
{ K(i)=R(i);
  for(j=1; j<i; j++) { K(i)=K(i)-a_j(i-1).R(i-j);}
  K(i)=K(i)/E(i-1);a_i(i)=K(i);
  for(l=1; l<i; l++) { a_l(i)=a_l(i-1)-K(i).a_{i-l}(i-l);}
  E(i)=[1-(K(i))^2].E(i-1);
}
    
```

最终得到的 $\{a_p(i); i=1,2,\dots,p\}$ 就是 $\{a_1,a_2,\dots,a_p\}$; 同时得到 $\{K(i); i=1,2,\dots,p\}$ 就是反射系数 $\{K_1,K_2,\dots,K_p\}$ 。此外还得到 p 种不同阶数线性预测的预测误差能量 $\{E(i); i=1,2,\dots,p\}$ 。

(3) 反射系数与格型滤波器

前述预测误差滤波器 $A(z)=1-\sum_{i=1}^p \hat{a}_i z^{-i}$ 本质上是一个 p 阶的 FIR 滤波器；它可以采用 (3-5-2a) 式那样的直接结构实现，但这种结构的参数灵敏度高，不仅参数的表示或编码要求的比特率较高，而且滤波时的有限字长效应比较明显，甚至可能出现不稳定而发生数据溢出。采用格型网络结构实现 FIR 滤波器可以避免这些缺点，其基本结构如图 3-5-1 所示。

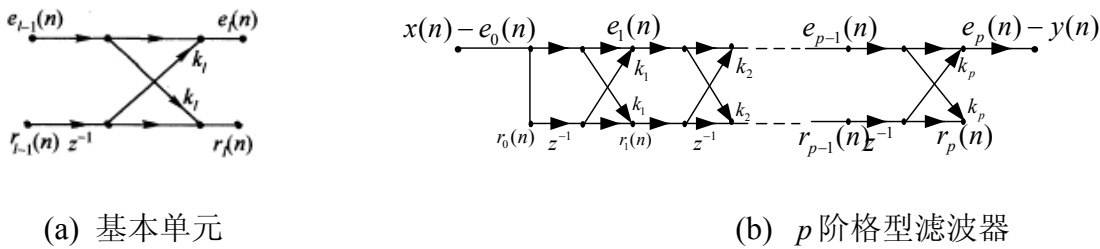


图 3-5-1 实现 p 阶 FIR 滤波器的格型网络

p 阶 FIR 滤波器的格型滤波器由 p 个基本单元级联而成；每个基本单元由 4 个前向支路构成，其差分方程为：

$$e_i(n) = e_{i-1}(n) + r_{i-1}(n).K_i \quad i = 1, 2, \dots, p \quad (3-5-8a)$$

$$r_i(n) = e_{i-1}(n).K_i + r_{i-1}(n) \quad i = 1, 2, \dots, p \quad (3-5-8b)$$

其中 $\{K_1, K_2, \dots, K_p\}$ 是反射系数，它是由线性预测系数 $\{a_1, a_2, \dots, a_p\}$ 推演出的一种等价参数。

p 个基本单元级联网络的最终输出 $\{e_p(n)\}$ ，等于(3-5-2a)式得到预测误差序列 $\{e(n)\}$ 。

由线性预测系数 $\{a_1, a_2, \dots, a_p\}$ 递推得到相应的反射系数 $\{K_1, K_2, \dots, K_p\}$ 的程序：

```

for(j=p; j>0; j--) { a_j^{(p)} = -a_j; }
K_p = a_p^{(p)};
for(i=p; i>0; i--)
{ for(n=1; n<i-1; n++) { a_n^{(i-1)} = (a_n^{(i)} + a_i^{(i)} a_{i-n}^{(i)}) / [1 - (K_i)^2]; }
  K_{i-1} = a_{i-1}^{(i-1)};
}
    
```

保证系统稳定的充分必要条件是： $-1 \leq K_i \leq 1 (\forall i)$ 。反射系数的内插特性和量化特性都优于线性预测系数，因此更适于压缩编码中的应用。

当反射系数 $\{K_1, K_2, \dots, K_p\}$ 已知时，递推得到相应线性预测系数 $\{a_1, a_2, \dots, a_p\}$ 的程序：

```
for(i=1; i < p+1; i++)
{
  a_i^{(i)} = K_i;
  for(j=1; j < i; j++) { a_j^{(i)} = a_j^{(i-1)} - K_i a_{i-j}^{(i-1)}; }
}
for(n=1; n < i; n++) { a_n = -a_n^{(p)}; }
```

采用格型滤波器实现线性预测有一个重要特点，就是它的各级的预测误差是相互独立的，因此允许去掉最后一级或多级，虽然预测误差增大，但余下的前几级的参数仍然是正确的；这就便于自适应调整级数和预测精度。

(3-5-6)式所示的信号重构全极点滤波器 $\hat{x}(n) = \hat{e}(n) + \sum_{i=1}^p \hat{a}_i \hat{x}(n-i)$ ，其输入、输出分别为 $\hat{e}(n)$ 和 $\hat{x}(n)$ ，等价的格型滤波器表达式为：

$$e_p(n) = \hat{e}(n) \quad (3-5-9a)$$

$$e_{i-1}(n) = e_i(n) - r_{i-1}(n-1)K_i \quad i = p, p-1, \dots, 1 \quad (3-5-9b)$$

$$r_i(n) = e_{i-1}(n)K_i + r_{i-1}(n-1) \quad i = p, p-1, \dots, 1 \quad (3-5-9c)$$

$$\hat{x}(n) = r_0(n) = e_0(n) \quad (3-5-9d)$$

其中 $\{K_1, K_2, \dots, K_p\}$ 是与线性预测系数 $\{\hat{a}_1, \hat{a}_2, \dots, \hat{a}_p\}$ 对应的反射系数。与前述全零点滤波器格型网络不同之处是，其基本单元中4条支路有两条是反馈支路。

当 $p=2$ 时的全极点格型滤波器结构如图3-5-2所示。

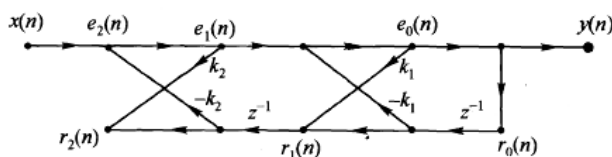


图 3-5-2 二阶全极点滤波器的格型网络实现

(4) 基于线性预测的几种失真测度

基于线性预测的失真测度有许多种，例如：板仓-斋田测度、归一化似然比测度、倒谱测度、线谱对测度等。但在数字通信中最有实用价值可能是前两种：

① 板仓-斋田(Itakura-Satio)失真测度

板仓-斋田失真测度又称似然比失真测度。采用线性预测技术对于一个 p 阶线性相关的有

记忆随机序列进行功率谱估计，每组线性预测系数对应于一种功率谱，而其估计误差 $\{e(n)\}$ 是一个能量最小化的高斯白噪声序列，因此(3-5-7)式给出的功率谱估计 $\sigma^2/|A(e^{j\omega})|^2$ 是这个随机序列功率谱的最大似然估计，可以采用似然比来衡量两个不同随机序列功率谱的差异大小。

设 $\{x^{(1)}(n)\}$ 和 $\{x^{(2)}(n)\}$ 都是 p 阶线性相关的有记忆随机序列，其归一化自相关函数分别表示为 $\mathbf{R}^{(1)}=[\tilde{R}^{(1)}(1) \ \tilde{R}^{(1)}(2) \ \dots \ \tilde{R}^{(1)}(p)]^T$ 和 $\mathbf{R}^{(2)}=[\tilde{R}^{(2)}(1) \ \tilde{R}^{(2)}(2) \ \dots \ \tilde{R}^{(2)}(p)]^T$ ，其线性预测系数分别表示为 $\mathbf{a}^{(1)}=[a_1^{(1)} \ a_2^{(1)} \ \dots \ a_p^{(1)}]^T$ 和 $\mathbf{a}^{(2)}=[a_1^{(2)} \ a_2^{(2)} \ \dots \ a_p^{(2)}]^T$ ，其功率谱分别为 $\sigma_1^2/|A^{(1)}(e^{j\omega})|^2$ 和 $\sigma_2^2/|A^{(2)}(e^{j\omega})|^2$ ，那么这两种功率谱之间的**似然比失真测度定义为**

$$d_{IS}(X^{(1)}; X^{(2)}) \triangleq \int_{-\pi}^{\pi} \frac{|A^{(2)}(e^{j\omega})|^2}{2\pi |A^{(1)}(e^{j\omega})|^2} d\omega - \text{Ln} \frac{\sigma_1^2}{\sigma_2^2} - 1 \quad (3-5-10a)$$

$$=[(\mathbf{a}^{(2)})^t \cdot \mathbf{R}^{(1)} \cdot \mathbf{a}^{(2)}] / \sigma_2^2 - \text{Ln}(\sigma_1^2) + \text{Ln}(\sigma_2^2) - 1$$

此式还可等价地表示为：

$$d_{IS}(X^{(1)}; X^{(2)}) = r_2(0) \cdot R^{(1)}(0) + 2 \sum_{k=1}^p r_2(k) \cdot R^{(1)}(k) - \text{Ln}(\sigma_1^2) + \text{Ln}(\sigma_2^2) - 1 \quad (3-5-10b)$$

其中 $r_2(k) = \sum_{i=0}^{p-k} a^{(2)}(i)a^{(2)}(i+k)$ ， $k=0,1,2,\dots,p$ 为线性预测系数 $\mathbf{a}^{(2)}$ 的自相关函数。

② 归一化似然比失真测度

如果只考虑两种谱形之间的差异大小，而不考虑二者能量大小，由似然比失真测度可得到

$$d_{GL}(X^{(1)}; X^{(2)}) = [(\mathbf{a}^{(2)})^t \cdot \mathbf{R}^{(1)} \cdot \mathbf{a}^{(2)}] / \sigma_2^2 - 1 \quad (3-5-11)$$

$$= r_2(0) \cdot R^{(1)}(0) + 2 \sum_{k=1}^p r_2(k) \cdot R^{(1)}(k) - 1$$

这两种失真测度都不满足对称性和三角不等式两个条件，因而不能用作距离测度。但它们能很好地反映两种谱的差异大小，也就是能很好地衡量一组线性预测参数相对于另一组线性预测参数之间总体失真值大小，因而适用于线性预测系数的矢量量化。

(5) 线性预测原理用于有记忆序列的压缩编码

设 $\{x(n)\}$ 可假定为一个记忆长度为 p 随机序列，其样本之间的相关性是线性的，那么我们可以采用如上所述的 p 阶线性预测原理，在每一短时段中估计一组最佳线性预测系数，并将这组系数进行压缩编码；设由此编码恢复得到的线性预测系数为 $\{\hat{a}_i\}$ ，那么以 $\{\hat{a}_i\}$ 为滤波系数，采用如(3-5-2)所示的预测误差滤波器求出预测误差序列 $\{e(n)\}$ ，并将它逐个样点地进行压缩编码，然后将它们与线性预测系数的编码一起发送。

设 $\{\hat{e}(n)\}$ 是接收端由 $\{e(n)\}$ 的编码恢复得到的序列， $\{\hat{a}_i\}$ 为接收端恢复的线性预测系数，那么采用(3-5-6)式即可近似重构原信号，得到 $\{\hat{x}(n)\}$ 。

(6)几种实用的线性预测编码方法

①自适应预测编码(APCM)：前述线性预测器的系数 $\{a_i\}$ 在整个编码过程中可以是固定不变的；但若让它们能自适应地改变，即每隔一定时间重新估计并修改一次，这就是APCM。

②差分脉冲编码调制(DPCM)：当上述线性预测编码方法中的系数 $\{a_i\}$ 只取一个系数 $a_1 = -1$ 时，线性预测简化为差分运算 $e(n) = x(n) - x(n-1)$ ；这就是DPCM。

③自适应差分脉冲编码调制(ADPCM)：对于差分运算的输出进行量化时，自适应改变阶距，并且将量化误差通过一个自适应预测器反馈到输入端，使编码误差进一步减小，就构成ADPCM。

④增量调制(ΔM)：对输入的连续信号以远远高于奈奎斯特临界采样频率进行采样，那么相邻两个样点之间的差分值得变得很小，以至于可以用1bit进行量化编码；这种DPCM编码器就变成了增量调制编码器。

3.5.2 子带编码

将连续信源产生的输出用多个带通滤波器进行滤波，只要这些滤波器是按照镜像滤波器成对设计的，这些滤波器的输出相加就可以无失真恢复原信号。信号经带通滤波后再分别进行编码，各个子带的频谱特性相对平坦一些，样点之间的相关性减小，加之不同子带在进行编码时可以根据它的重要性分配不同的比特数，因此可以压缩编码数据。

3.5.3 变换域编码

将连续信源产生的输出采用某种正交变换进行变换，变换域中各个信号分量的相关性大大减小，于是可以根据各个分量的重要性进行量化编码，可以大幅度减小比特率。常用的正交变换有离散傅立叶变换、离散余弦变换、小波变换等。这些变换都是可逆的，接收端采用逆变换对于在变换域量化编码的结果进行逆变换，就可近似恢复原信号。这些编码算法在图像压缩编码等方面取得了很大的成功。

3.5.4 基于信源模型假定的信源编码方法

如果信源特性可以采用某种数学模型描述，那么基于这个模型进行压缩编码，可以获得更高的编码效率。例如：语音信号可以用语音产生的声道模型进行描述，因而只要对估计出的声道模型参数和激励信号进行编码，接收端就可以重新合成出与原信号感觉上相类似的信

号，这就是声码器技术。

由于全极点声道模型抓住了语音信号共振峰特性的本质，而且有前述线性预测编码方法作为基础，计算上十分简捷，因此以线性预测为中心的一大类声码器在语音信号压缩编码方面取得了很大成功。包括：LPC 声码器、CELP 声码器、多带激励（MBE）声码器等。基于声码器技术，压缩语音编码数据率的效率是非常高的，例如将 64Kbps 的 PCM 语音信号压缩到 8Kbps 左右时，恢复语音的音质还可以是透明的；压缩到 600bps 时，恢复语音的可懂度还能接近 100%。

第3章 习题

3-1 根据带通信号采样定理（推论），将一个带宽为 2MHz、中心频率为 70MHz 的带通信号以采样频率 f_T 进行采样量化，并采用本地振荡的频率为 f_0 数字正交下变频法求它的零中频复信号；选择一个不引起采样混叠失真的采样频率值 f_T ，并给出相应的 f_0 值；画出采样前后信号的频谱示意图。

3-2 某个离散无记忆信源的字符集共有 8 个字符，其概率分布分别为 0.25, 0.2, 0.15, 0.12, 0.10, 0.08, 0.05 和 0.05。

- 对信源的输出设计一个 Huffman 二进制编码；
- 计算平均每个信源字符的编码的二进制数字个数 \bar{R} ；
- 计算信源熵，并与 \bar{R} 作比较。

3-3 两个二进制随机变量的联合分布为：

$$P(X=0, Y=0) = P(X=0, Y=1) = P(X=1, Y=1) = 1/3, \text{ 计算} \\ H(X), H(Y), H(X|Y), H(Y|X) \text{ 和 } H(X, Y)。$$

3-4 关于语音信号的抽样量化，为什么采用 8 比特 A 率非均匀量化所得数字语音信号，其语音质量与 13 比特均匀量化的结果相当？如果信号峰值功率与平均功率之比等于 10，那么 8 比特均匀量化所得数字信号的量化信噪比是多少？

3-5 假设 U 是在 [-1 1] 区间服从均匀分布的随机变量。（1）设计一具有最小均方误差的 3bit 量化器；（2）证明满足要求的 3bit 量化器是唯一的。

3-6 假如计算机键盘上有 110 个字符，每个字符用二进制码元来表示。（1）每个字符需要用多少位二进制比特来表示？（2）如果以相同的概率发送每个字符，试求每个字符包含的信息量。

3-7 假设一模拟信号频率范围为 300Hz~3000Hz，利用 7KHz 的抽样频率生成 PCM 信号。（1）画出 PCM 系统的方框图，包含发送机、信道和接收机；（2）讨论如何采用非均匀量化以提高系统的性能。

3-8 试对随机变量 x 进行量化，已知 $p(x) = e^{-x^2/2\sigma^2} / \sqrt{2\pi}$ ， $-\infty < x < \infty$ ，失真函数为 $(x-y)^2$ ，量化级数 $m=3$ 。求最佳量化值、量化区间以及最小平均失真 D_{\min} 。

3-9 证明：对于广义正态平稳过程所取得的平稳序列，线性最佳预测就是条件期望预测。

3-10 两个随机变量 X、Y，其和为 $Z=X+Y$ ，若 X、Y 相互独立，试证：（1） $H(X) \leq H(Z)$, $H(Y) \leq H(Z)$ ；（2） $H(XY) \geq H(Z)$ 。

3-11 对于任意三个离散随机变量 X、Y、Z，证明：

$$(1) H(XYZ) = H(XZ) + H(Y|X) - I(Z, Y|X); \quad (2) H(XYZ) - H(XY) \leq H(ZX) - H(X)$$

3-12 在数字化自然采样中，模拟信号 $g(t)$ 与一周期方波脉冲串相乘，脉冲串的重复频率为 f_s ，每一脉冲的宽度为 T ，并有 $f_s T \ll 1$ 。(1) 试求通过自然采样后信号的频谱；(2) 证明如果自然采样满足奈奎斯特采样定理，则可以通过采样后的信号完全恢复原信号。

3-13 一个离散信源有 3 个独立码元 A、B 和 C，概率分别为 0.9、0.07、0.03。为该信源设计一个二进制霍夫曼编码，求编码的平均码长。