多媒体编码及其信息安全应用

Multimedia Coding and Its Application to Information Security

第二讲
多媒体编码基础

授课时间：2022年2月28日

内 容 提 纲

1. 信源编码基础理论
2. 预测编码、变换编码、统计编码
3. 波形编码、参数编码、混合编码
4. 小结

内 容 提 纲

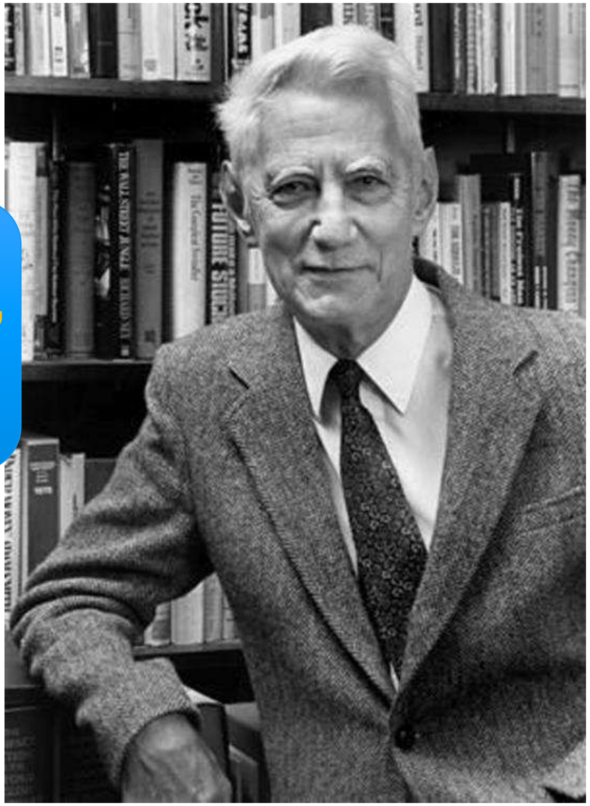
1. **信源编码基础理论**
2. **预测编码、变换编码、统计编码**
3. **波形编码、参数编码、混合编码**
4. **小结**

1.1 概述

- 香农对信息理论、编码理论的主要**贡献**
  - **1948年，发表《通信的数学理论》**

1948年，香农在论文《通信的数学理论》（A mathematical theory of communication）中，用概率测度和数理统计的方法系统地讨论了通信的基本问题，得出了一些重要而带有普遍意义的结论。

  - 1949年，发表《噪声下的通信》
  - 1956年，发表《噪声信道的零差错容量》
  - 1959年，发表《在保真度准则下的离散信源编码定理》
  - 1961年，发表《双路通信系统》
- 香农理论的**核心**：
  - 在通信系统中**采用适当的编码**后能够实现高效率和高可靠性的信息传输
  - 给出了信源编码定理和信道编码定理



克劳德·艾尔伍德·**香农**（Claude Elwood **Shannon**，1916年4月30日—2001年2月24日）
美国数学家、**信息论的创始人**

1.1 概述

香农信息论对信源编码的指导意义

- 针对系统有效性问题，香农信息论研究在保证信息传输可靠性或传输错误概率小于某一给定值的条件下，如何最有效地利用信道的传输能力
- 研究在给定信源及信源编码有一定失真的条件下，信源编码的最低速率是多少
- 各种信源中普遍存在的统计多余度的消除，如何在无失真或限定失真的条件下对信源进行高效压缩编码。香农第一定理和香农第三定理分别从理论上给出了无失真信源编码与限失真信源编码的压缩极限，对于压缩编码的研究具有重要的理论指导意义
- 香农信息论为我们寻找信源有效编码方法提供了理论依据和有价值的改进方向

1.2 信源的分类

- 单符号信源：输出是单个符号（代码）的消息
  - 离散信源
  - 连续信源
- 平稳随机序列信源：信源输出的消息由一系列符号序列所组成，可用N维随机矢量 $\boldsymbol{X}=(X_1,X_2,\ldots,X_N)$来描述，且随机矢量$\boldsymbol{X}$的各维概率分布都与时间起点无关----平稳！！
  - 离散平稳信源
  - 连续平稳信源
  - 无记忆（独立）离散平稳信源
  - 有记忆信源
  - m阶马尔可夫信源
- 随机波形信源

1.3 信源的数学模型

❍ 离散信源(单符号)

- **特点：** 输出是单个符号（代码）的消息，符号集的取值A:$\{a_1, a_2,\ldots, a_q\}$是有限的或可数的，可用一维离散型随机变量X来描述。
  - 例：投硬币、书信、电报符号等等。
- **数学模型：** 设每个信源符号a_i出现的(先验)概率$p(a_i)$ (i=1,2,…,q) 满足：

$$\sum_{i=1}^{q} p(a_i) = 1$$

$$\begin{bmatrix} X \\ p(x) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & \ldots & \ldots & a_q \\ p(a_1) & p(a_2) & p(a_3) & \ldots & \ldots & p(a_q) \end{bmatrix}$$

- 概率空间能表征离散信源的统计特性，因此又称概率空间为信源空间。

1.3 信源的数学模型

连续信源(单符号)

- **特点：** 输出是单个符号（代码）的消息，输出消息的符号集A的取值是连续的，可用一维的连续型随机变量X来描述。
  - 例：语音信号、热噪声信号、测控系统中有关电压、温度、压力等测得的连续数据等。
- **数学模型：** 连续型的概率空间。即：

$$\begin{bmatrix} X \\ p(x) \end{bmatrix} = \begin{bmatrix} (a,b) \\ p(x) \end{bmatrix} \quad \text{或} \quad \begin{bmatrix} R \\ p(x) \end{bmatrix}$$

满足 $$\int_a^b p(x)dx = 1 \quad \text{或} \quad \int_R p(x)dx = 1$$

1.3 信源的数学模型

平稳随机序列信源

- **总体特点**：

  信源输出的消息由一系列符号序列所组成，可用N维随机矢量 $\boldsymbol{X}=(X_1,X_2,\ldots,X_N)$来描述，且随机矢量 $\boldsymbol{X}$ 的各维概率分布都与时间起点无关----平稳！

- 离散平稳信源：每个随机变量X_i $(i=1,2,\ldots,N)$都是离散型随机变量

  例如，投掷硬币(多个同时)

- 连续平稳信源：每个随机变量X_i $(i=1,2,\ldots,N)$都是取值连续的随机变量

  例如，多信源语音信号{X(t)}、热噪声信号{n(t)}

离散无记忆平稳信源

- 离散平稳信源的特例，信源发出的符号相互统计独立，即各随机变量X_i (i＝1,2,…,N)之间统计独立
- 性质：
  - 独立－>$p(\boldsymbol{X}) = p(X_1, X_2, \ldots, X_N) = p_1(X_1) \cdot p_2(X_2) \cdots p_N(X_N)$
  - 平稳－> $p_1(X_i) = p_2(X_i) = \cdots = p_N(X_i)$ －> $p(X_i)$

$$p(\mathbf{X}) = p(X_1 X_2 \ldots X_N) = \prod_{i=1}^{N} p(X_i)$$

- 设各随机变量X_i取值同一符号集A:$\{a_1, a_2, \ldots, a_q\}$，则

$$p(\mathbf{x} = \alpha_i) = p(a_{i1} a_{i2}, \ldots, a_{iN}) = \prod_{k=1}^{N} p(a_{ik}), i_k = (1, 2, \ldots, q)$$

N维随机矢量的一个取值，
$\alpha_i = (a_{i1}\ a_{i2} \ldots a_{iN})$

$p(a_{ik})$是符号集A的一维概率分布

1.3 信源的数学模型

若信源空间：$\begin{bmatrix} X \\ p(x) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & \cdots & a_q \\ p(a_1) & p(a_2) & p(a_3) & \cdots & \cdots & p(a_q) \end{bmatrix}$

- 描述的信源X的各输出X_i间统计独立、且取值于同一符号集A，则X为离散无记忆信源，称该信源输出的N维随机矢量***X***为离散无记忆信源X的N次扩展信源。
- 若***X***取值为符号集$\boldsymbol{\alpha_i} = (a_{i1}a_{i2}\ldots a_{iN})$，其中$(i_1, i_2, \ldots, i_N = 1, 2, \ldots, q)$，则离散无记忆信源的N次扩展信源的数学模型是X信源空间的N重空间：

$$\begin{bmatrix} X^N \\ p(\alpha_i) \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{q^N} \\ p(\alpha_1) & p(\alpha_2) & \cdots & p(\alpha_{q^N}) \end{bmatrix}$$

满足，$p(\alpha_i) = \prod_{k=1}^{N} p(a_{i_k}), i_k = (1, 2, \ldots, q)$

1.3 信源的数学模型

❍有记忆信源

- 信源在不同时刻发出的符号之间是相互依赖的，即信源输出的平稳随机序列**X**中，各随机变量X_i之间相互依赖。
- 需在**N**维随机矢量的联合概率分布中，引入条件概率分布来说明它们之间的关联。
  - **例**：汉字组成的中文序列中，只有根据中文的语法、习惯用语、修辞制约和表达实际意义的制约所构成的中文序列才是有意义的中文句子或文章。
  - 所以，在汉字序列中前后文字的出现是有依赖的，不能认为是彼此不相关的。

其他自然语言如英文等也都是如此。

m阶马尔可夫信源

- 不同时刻发出的符号间的依赖关系

$$p(x_i \mid \cdots x_{i+2}x_{i+1}x_{i-1}x_{i-2}\cdots x_{i-m}\cdots x_1)$$
$$= p(x_i \mid x_{i-1}x_{i-2}\cdots x_{i-m})\ (i=1,2,\cdots,N)$$

- 当记忆信源的记忆长度为m+1时，称这种有记忆信源为m阶马尔可夫信源。

- 若上述条件概率与时间起点 i 无关，信源输出的符号序列可看成为时齐马尔可夫链，则此信源称为时齐马尔可夫信源。

1.4 信源的熵

- 讨论基本的离散信源(即输出为单个符号的消息，且前后消息之间互不相关)
- 基本的离散信源可用一维随机变量X来描述信源的输出，信源的数学模型可抽象为:

$$\begin{bmatrix} X \\ p(x) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & \cdots & a_q \\ p(a_1) & p(a_2) & p(a_3) & \cdots & \cdots & p(a_q) \end{bmatrix} \quad \sum_{i=1}^{q} p(a_i) = 1$$

问题：这样的信源能输出多少信息?

每个消息的出现携带多少信息量?

❍信息的测度

- 考虑：
  - 信息的测度（信息量）与不确定性消除的程度有关：消除的不确定性＝获得的信息量；
  - 不确定性是随机的，我们用概率论和随机过程理论来测量，概率小-->不确定性大；
- 推论：
  - 概率小 -->信息量大，即信息量应该是概率的单调递减函数；
  - 信息量应该具有可加性。

自信息

- 设离散信源X的概率空间为：

$$\begin{bmatrix} X \\ p(x) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & \cdots & a_q \\ p(a_1) & p(a_2) & p(a_3) & \cdots & \cdots & p(a_q) \end{bmatrix} \quad \sum_{i=1}^{q} p(a_i) = 1$$

- 称事件a_i发生所含有的信息量为 a_i 的**自信息量**。定义为：

$$I(a_i) = f[p(a_i)]$$

$I(a_i)$代表两种含义：

（1）当事件a_i发生之前，表示事件a_i发生的不确定性

（2）当事件a_i发生之后，表示事件a_i所提供的信息量

1.4 信源的熵

- 某事件a_i发生所含有的信息量，应该是该事件发生的先验概率$p(a_i)$的函数。即：

$$I(a_i) = f[p(a_i)]$$

- 根据客观事实与人们的习惯，函数 $f[p(a_i)]$ 应该满足以下条件：

（1）它应是先验概率 $p(a_i)$的单调递减函数，即

当$p(a_1) > p(a_2)$ 时，有 $f[p(a_1)] < f[p(a_2)]$；

（2）当$p(a_i) =1$时， $f[p(a_i)] = 0$

（3）当$p(a_i) =0$时， $f[p(a_i)] = \infty$

（4）两个独立事件的联合信息量应等于它们各自的信息量之和，即统计独立信源的信息量等于它们分别的信息量之和。

- 对数函数满足上述条件：

$$\boxed{I(a_i) = f[p(a_i)] = \log_r \frac{1}{p(a_i)}} \quad （r>1）$$

说明：

- 计算自信息量时要正确计算有关事件发生的概率；
- 自信息量的单位取决于对数的底；
  - 底为“2”，则单位为“比特（bit, binary unit）”；
  - 底为“e”，则单位为“奈特（nat, nature unit）”；
  - 底为“10”，则单位为“哈特（hat, Hartley）”；
  - 可以用换底公式进行换算：

$$\log_a X = \frac{\log_b X}{\log_b a}$$

1 nat = 1.44bit , 1 hat = 3.32 bit；

- 通常采用以“2”为底的对数，为了书写简洁，常把底数“2”略去不写。

例：某地某月份的气象资料如下表所示，求相应事件的发生所带来的自信息量。

x_i	x_1(晴)	x_2(阴)	x_3(雨)	x_4(雪)
$p(x_i)$	1/2	1/4	1/8	1/8

解：由上表可求出各事件发生所带来的自信息量为：

$$I(x_1) = -\log_2 p(x_1) = 1(bit)$$

$$I(x_2) = -\log_2 p(x_2) = 2(bit)$$

$$I(x_3) = -\log_2 p(x_3) = 3(bit)$$

$$I(x_4) = -\log_2 p(x_4) = 3(bit)$$

◆本例说明：一个随机事件发生的概率越小，它的不确定性就越大，那么当它发生时对外提供的自信息量也就越大。

1.4 信源的熵

信息熵

- 对一个信源发出不同的消息所含有的信息量也不同，不能用它来作为**整个信源**的信息度量。
- **定义**自信息的数学期望-----**平均自信息量**$H_r(X)$，为**信息熵**：

$$H_r(X)=E\left(\log_r \frac{1}{p(a_i)}\right)=-\sum_{i=1}^{q} p(a_i)\log_r p(a_i)$$

当r＝2时：$\boxed{H(X)=E\left(\log \frac{1}{p(a_i)}\right)=-\sum_{i=1}^{q} p(a_i)\log p(a_i)}$

$$H_r(X)=H(X)/\log r$$

1.4 信源的熵

- 由于这个表达式和统计物理学中热熵的表达式相似，且在概念上也有相似之处，因此借用“熵”这个词，把**H(X)**称为“信息熵”；
- 信息熵的单位由自信息量的单位决定，即取决于对数的底。

$H_r(X)$的单位：r 进制单位 / 符号 （r>1）

（bit / 符号）

熵的计算例子：

一布袋内放入100个球，其中80个球是红色的，20个球是白色的。随便摸出一个球，猜是什么颜色。该事件的概率空间为：

$$\begin{pmatrix} X \\ p(X) \end{pmatrix} = \begin{bmatrix} a_1 & a_2 \\ 0.8 & 0.2 \end{bmatrix} \qquad H_r(X) = -\sum_{i=1}^{q} p(a_i)\log_r p(a_i)$$

如果被告知摸出的是红球，那么获得的信息量是：

$$I(a_1) = -\log p(a_1) = -\log 0.8 = 0.32 \text{（比特）}$$

如被告知摸出来的是白球，所获得的信息量应为：

$$I(a_2) = -\log p(a_2) = -\log 0.2 = 2.32 \text{（比特）}$$

平均摸取一次所能获得的信息量为：

$$H(X) = p(a_1)\,I(a_1) + p(a_2)\,I(a_2) = 0.72 \text{（比特/符号）}$$

1.4 信源的熵

❍熵的含义

- 熵----考虑的是整个集合的统计特性，它从平均意义上来表征信源的总体特征。
- 在信源输出前，信息熵H(X) 表示信源的平均不确定性；
- 在信源输出后，信息熵H(X)表示每个消息提供的平均信息量；
- 信息熵H(X) 表征了变量X的随机特性。
- 例如，有两信源X、Y，其概率空间分别为：

$$\begin{pmatrix} X \\ p(x) \end{pmatrix} = \begin{bmatrix} a_1 & a_2 \\ 0.99 & 0.01 \end{bmatrix} \quad \begin{pmatrix} Y \\ p(y) \end{pmatrix} = \begin{bmatrix} a_1 & a_2 \\ 0.5 & 0.5 \end{bmatrix}$$

$$\boxed{H_r(X) = -\sum_{i=1}^{q} p(a_i) \log_r p(a_i)}$$

计算其熵，得：$H(X) = -p(a_1)\log p(a_1) - p(a_2)\log p(a_2)$

$= 0.08$（bit /符号）

$H(Y)=1$（bit / 符号）

H(Y)＞H(X)，表明信源Y比信源X的平均不确定性要大。

1.4 信源的冗余度

- 对离散信源，信源符号等概率分布时熵最大，其平均信息量记为：**$H_0 = \log q$**
- 由于信源符号间的依赖关系使信源的熵减小，即：

$$\log q = H_0 \geq H_1 \geq H_2 \geq \ldots \geq H_{m+1} \geq \ldots \geq H_\infty$$

- 信源符号之间依赖关系越强，每个符号提供的平均信息量越小。
- 为此，引入信源的冗余度来衡量信源的相关程度(有时也称为多余度)。

1.4 信源的冗余度

- 定义：
- 信源效率———一个信源实际的信息熵与具有同样符号集的最大熵的比值。

$$\eta = \frac{H_\infty}{H_0}$$

- 信源冗余度———针对最大熵而言，信源中无用信息所占的比例。

$$\gamma = 1 - \eta = 1 - \frac{H_\infty}{H_0}$$

式中：$\mathbf{H_0}$为熵的最大值，$\mathbf{H_\infty}$为熵的实际值。

1.4 信源的冗余度

- 关于冗余度
  - 冗余度γ 越大，实际熵H_∞ 越小。说明信源符号之间的依赖关系越强，即符号之间的记忆长度越长。
  - 冗余度γ 越小，表明信源符号之间依赖关系越弱，即符号之间的记忆长度越短。
  - 当冗余度等于零时，信源的信息熵就等于最大值H_0，这表明信源符号之间不但统计独立无记忆，而且各符号还是等概率分布。

- 从提高信息传输效率的角度出发，总是希望减少冗余度（压缩），这是信源编码的作用；
- 从提高信息抗干扰能力来看，总是希望增加或保留冗余度，这是信道编码要达到的目的。

1.4 信源的冗余度

例：英语----字母表

以英文字母组成的信源为例，信源的输出是英文字母组成的序列。英文字母共**26**个加上空格共**27**个符号。所以由英文字母组成的信源的最大熵：

$$H_0 = \log 27 = 4.76 \text{（比特／符号）}$$

考虑到字母之间的依赖关系，可以把英文信源作进一步的近似，看作为**M**维平稳信源。这样可以求得：

$H_1 = 4.03$ 比特／符号

$H_2 = 3.32$ 比特／符号

$H_3 = 3.1$ 比特／符号

……

$H_\infty = 1.4$ 比特／符号

$$\eta = \frac{H_\infty}{H_0}$$

$$\gamma = 1 - \eta = 1 - \frac{H_\infty}{H_0}$$

信源效率： $\eta = 0.294$

信源冗余度：$\gamma = 0.706$

1.4 信源的冗余度

- 英文信源的冗余度说明：
  - 文中有70.6%是由语言结构定好的；
  - 只有29.4%是写文字的人可以自由选择的。
  - 在传递或存储英语信息时，那些有关联的字母可进行大幅度地压缩。
  - 例如100页的书，大约只要存储29页就可以了，其中约71页可以压缩掉。
- 信源的剩余度表示这种信源可压缩的程度。

本例的启示：

→可以通过去除实际信源中符号间的相关性，来达到减少信源中符号个数的目的，从而使实际信源最佳化，前提条件是没有改变信源对外提供的平均信息量。

→对最佳信源来说，用来对信源符号进行编码所需的总比特数可以降低。

这一思想正是数据压缩的基本思想，它在语音压缩编码、图像压缩编码等诸多的数据压缩研究领域中得到了广泛的应用。

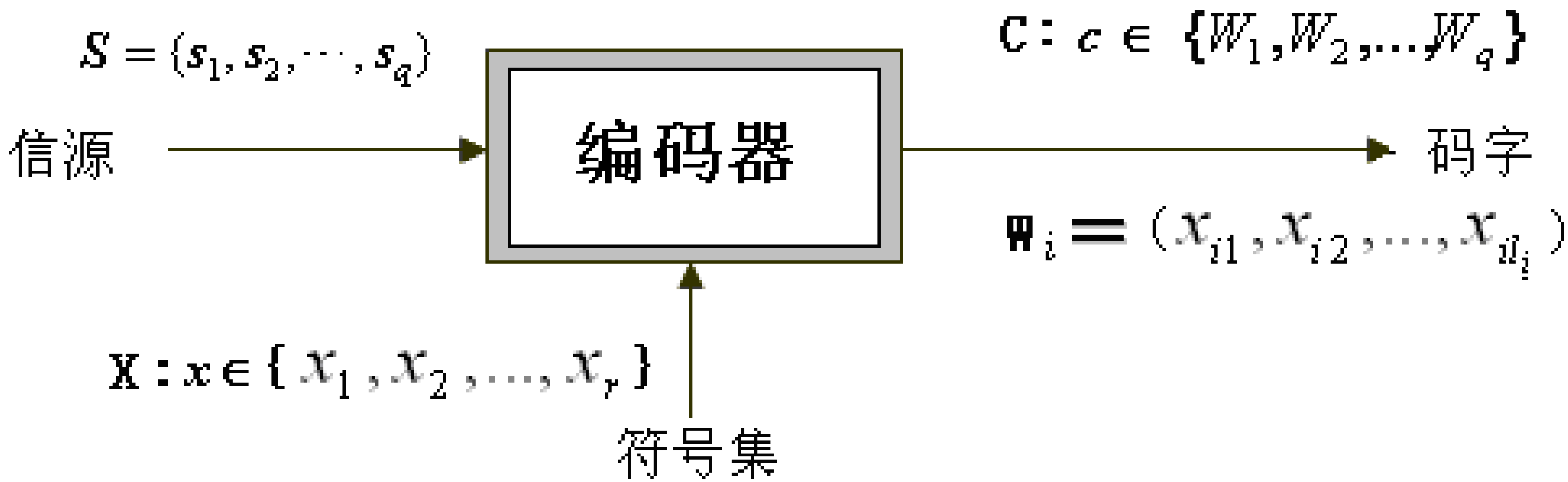
1.4 信源的冗余度

- 信源编码的基本途径有两个:
  - 使编码后各个符号出现的概率尽可能相等，即均匀化分布----方法主要是统计编码（熵编码）。
  - 使编码后使序列中的各个符号之间尽可能地互相独立，即解除相关性----方法包括预测编码和变换编码。

信源编码常分为无失真信源编码和限失真信源编码，前者主要用于文字、数据信源的压缩，后者主要用于图像、语音信源的压缩。

1.5 编码器模型

- 由于信源编码可以不考虑抗干扰问题，所以它的数学模型比较简单。
- 下图为一个编码器模型：



1.6 码的分类

[例] 设信源S的概率空间为：

$$\begin{bmatrix} S \\ p(s) \end{bmatrix} = \begin{bmatrix} s_1 & s_2 & s_3 & \cdots & \cdots & s_q \\ p(s_1) & p(s_2) & p(s_3) & \cdots & \cdots & p(s_q) \end{bmatrix} \qquad \sum_{i=1}^{q} p(s_i) = 1$$

编码成 二元码：

信源符号s_i	码字
s_1	00
s_2	01
s_3	10
s_4	11

二次扩展码：

信源符号	码字	信源符号	码字
α_1	0000	α_5	0100
α_2	0001	:	:
α_3	0010	:	:
α_4	0011	α_{16}	1111

信源S的二次扩展信源：

符号序列与码字一一对应

$$S^2 = [\alpha_1 = s_1 s_1, \alpha_2 = s_1 s_2, \alpha_3 = s_1 s_3, \ldots, \alpha_{16} = s_4 s_4]$$

1.6 码的分类

❍唯一可译码

若码的任意一串有限长的码符号序列只能被**唯一地译成所对应的信源符号序列**，则此码称为唯一可译码（单义可译码）。否则就称为非唯一可译码或非单义可译码。

- **例**：
- C_1=(1,01,00)　是唯一可译码，

如码字序列10001101只可唯一划分为1,00,01,1,01；

- C_2=(0,10,01) 为非唯一可译码，

如码字序列01001可划分为0,10,01或01,0,01。

1.6 码的分类

- 唯一可译码又分为即时码和非即时码：
- **即时码**-----如果在接收端收到一个完整的码字后，就能立即进行译码，这样的码叫做即时码；
- 即时码又称为**非延长码**，对即时码而言，在码本中任意一个码字都不是其它码字的前缀部分。
- **非即时码**-----在接收端收到一个完整的码字后，还需等下一个符合码接收后才能判断是否可以译码，这样的码叫做非即时码。
- 对非即时码来说，有的码是唯一可译的，有的码是非唯一可译的，主要取决于码的总体结构。

内 容 提 纲

1. 信源编码基础理论
2. 预测编码、变换编码、统计编码
3. 波形编码、参数编码、混合编码
4. 小结

2.1 预测编码

- 预测编码是数据压缩三大经典技术（统计编码、预测编码、变换编码）之一，它是建立在信源数据**相关性**之上的
- 由信息理论可知，对于**相关性**很强的信源，**条件熵可远小于无条件熵**，因此人们常采用尽量解除相关性的办法，使信源输出转化为独立序列，以利于进一步压缩码率
- 常用的**解除相关性的主要措施**是预测和变换，其实质都是进行序列的一种**映射**
  - 一般来说，**预测编码**有可能完全解除序列的相关性，但必需确知序列的概率特性；
  - **变换编码**一般只解除矢量内部的相关性，但它可有许多可供选择的变换方法，以适应不同的信源特性。

2.1 预测编码

- **预测编码的基本思想**是通过提取与每个信源符号有关的新信息，并对这些新信息进行编码来消除信源符号之间的相关性
- 实际中常用的**新信息**为信源符号的当前值与预测值的差值，这里正是由于信源符号间存在相关性，所以才使预测成为可能，对于独立信源，预测就没有意义

$$
\begin{bmatrix}
1 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & 0 & 0 & \cdots & 0 & 0 \\
1(0) & 0 & 0 & \cdots & 0 & 0 \\
2(1) & 0 & 0 & \cdots & 0 & \vdots \\
3(1) & 0 & 0 & \cdots & 0 & 0 \\
4(1) & 0 & 0 & & 0 & 0 \\
5(1) & 0 & 0 & & 0 & 0
\end{bmatrix}
\begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\
3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 \\
4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 \\
5 & 5 & 5 & 5 & 5 & 5 & 5 & 5 & 5 & 5
\end{pmatrix}
$$

2.1 预测编码

- 预测的理论基础主要是估计理论
- 所谓估计就是**用实验数据组成一个统计量**作为某一物理量的估计值或预测值
  - 若估值的数学期望等于原来的物理量，就称这种估计为**无偏估计**；
  - 若估值与原物理量之间的**均方误差最小**，就称之为**最佳估计**，基于这种方法进行预测，就称为最小均方误差预测，所以也就认为这种预测是最佳的。
- 要实现**最佳预测**就是要找到计算预测值的**预测函数**

2.1 预测编码

- 设有信源序列 $s_{r-k},\dots,s_{r-2},s_{r-1},\cdots$，$k$阶预测就是由$s_r$的前$k$个数据来预测 s_r。

  令预测值为： $s'_r = f(s_{r-1},s_{r-2},\dots,s_{r-k})$

  式中$f(*)$是待定的预测函数。

  要使预测值具有最小均方误差，必须确知k个变量的联合概率密度函数，这在一般情况下较难得到，实际中常用比较简单的线性预测方法。

- 线性预测是取预测函数为各已知信源符号的线性函数，即取 s_r 的预测值为：

  $$s'_r = \sum_{i=1}^{k} a_i s_{r-i}$$

  其中a_i为预测系数。

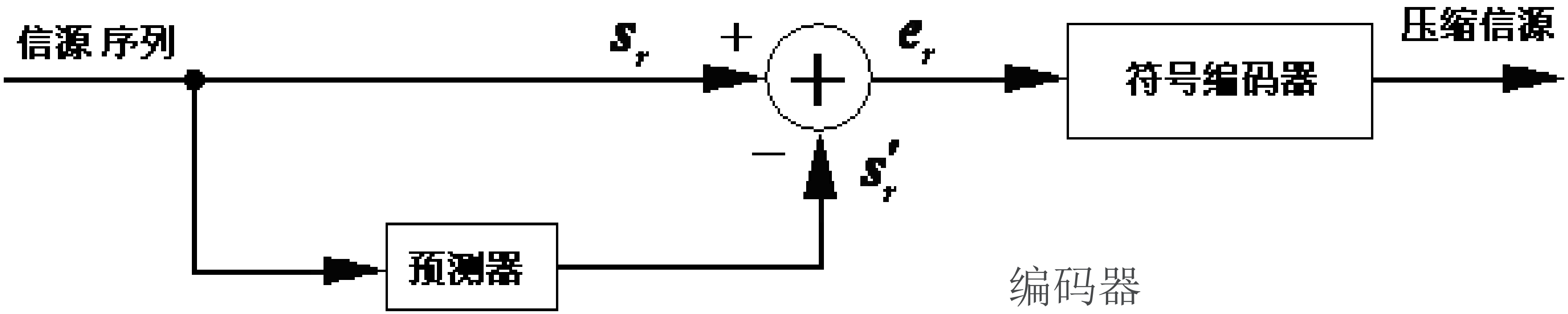
- 最简单的预测是令 $s'_r = s_{r-1}$

  称为前值预测，常用的差值预测就属于这类。

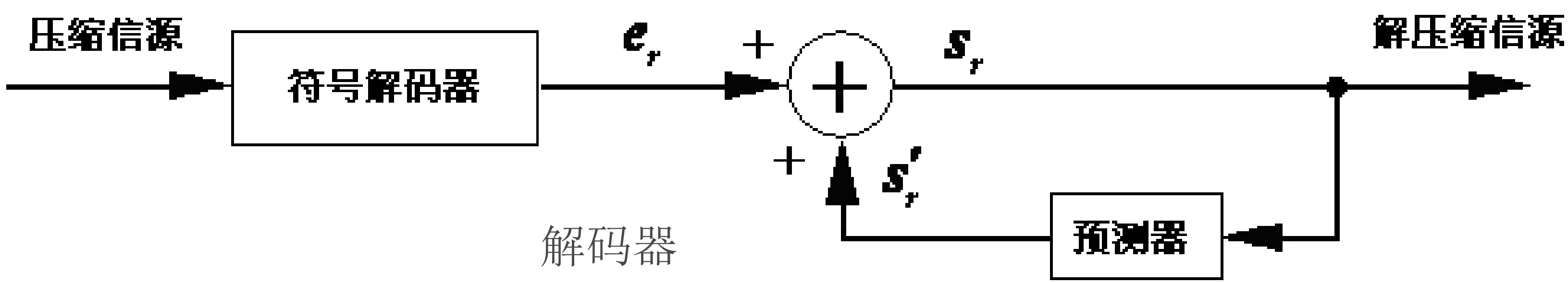
2.1 预测编码

差值预测编码系统

- 由于相关性很强的信源可较精确地预测待编码的值，使得这个差值的**方差**将远小于原来的信源取值，所以在同样失真要求下，**量化级数可大大减少**，从而较显著地压缩码率。
- 差值预测编码系统的框图如下，在编码端主要由一个符号编码器和一个预测器组成，在解码端主要由一个符号解码器和一个预测器组成。



编码器



解码器

2.1 预测编码

- 当输入信源序列逐个进入编码器时，预测器根据若干个过去的输入符号产生对当前输入符号的估计值。同时计算预测误差：

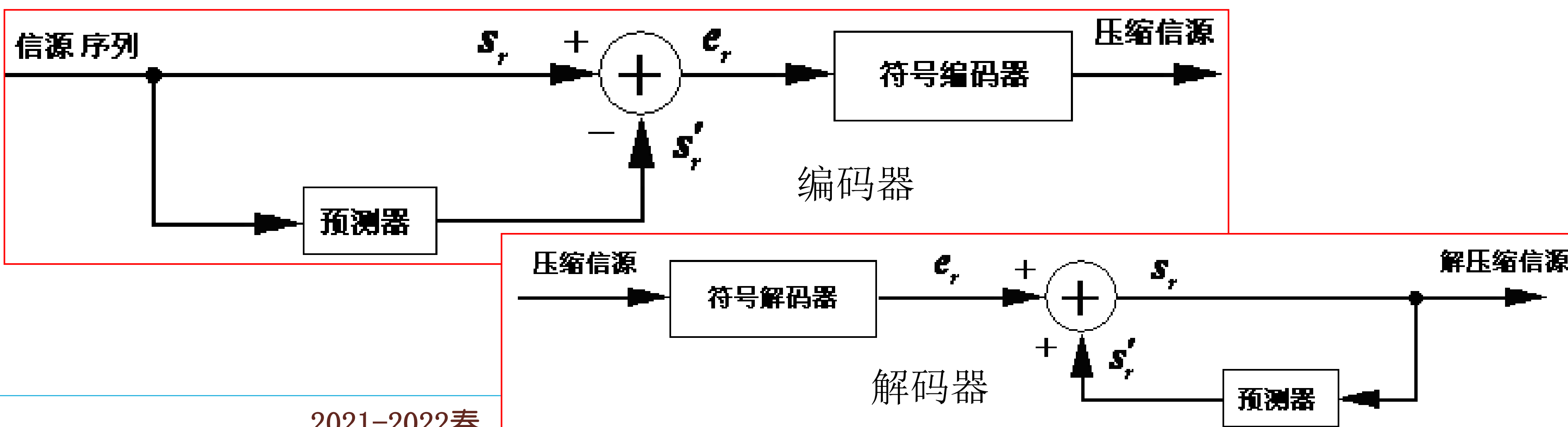
$$e_r = s_r - s_r'$$

- 在解码器中根据接收到的变长码字重建 e_r ，并执行下列操作：

$$s_r = e_r + s_r'$$

线性预测编码

- 而 s_r' 可通过式 $s_r' = \sum_{i=1}^{k} a_i s_{r-i}$ 进行预测得到。

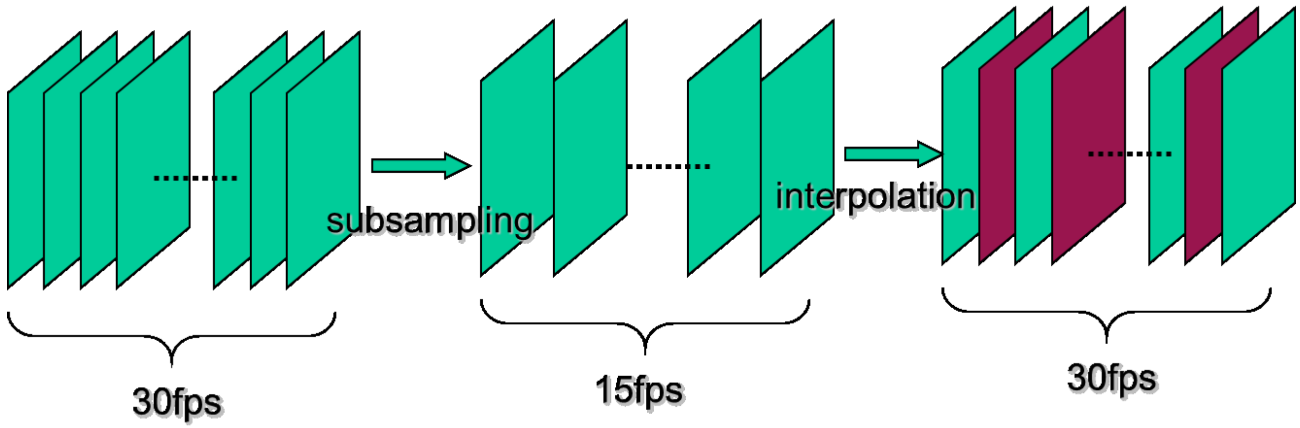


2.1 预测编码

- 媒体（如声音，图像等）数据的**相邻采样值**都非常接近，即具有很强的**相关性**。利用相邻采样值间的相关性进行预测，可以实现很高的编码效率
- 预测编码是一种有损编码。因为它的实现是基于声音或图像数据的物理特性——当采样间隔很密时，相邻采样点之间取值非常接近
- 预测编码利用数据间存在的**时间相关性**（声音数据）或**空间相关性**（图像数据）进行数据压缩

2.1 预测编码

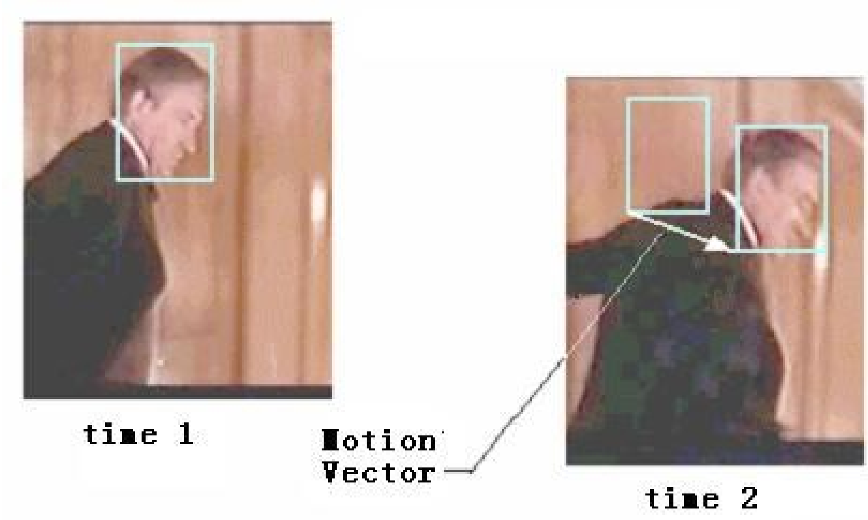
○举例：



subsampling
interpolation
30fps
15fps
30fps

时域

当前帧
下一帧
差分帧



time 1
Motion Vector
time 2

2.2 变换编码

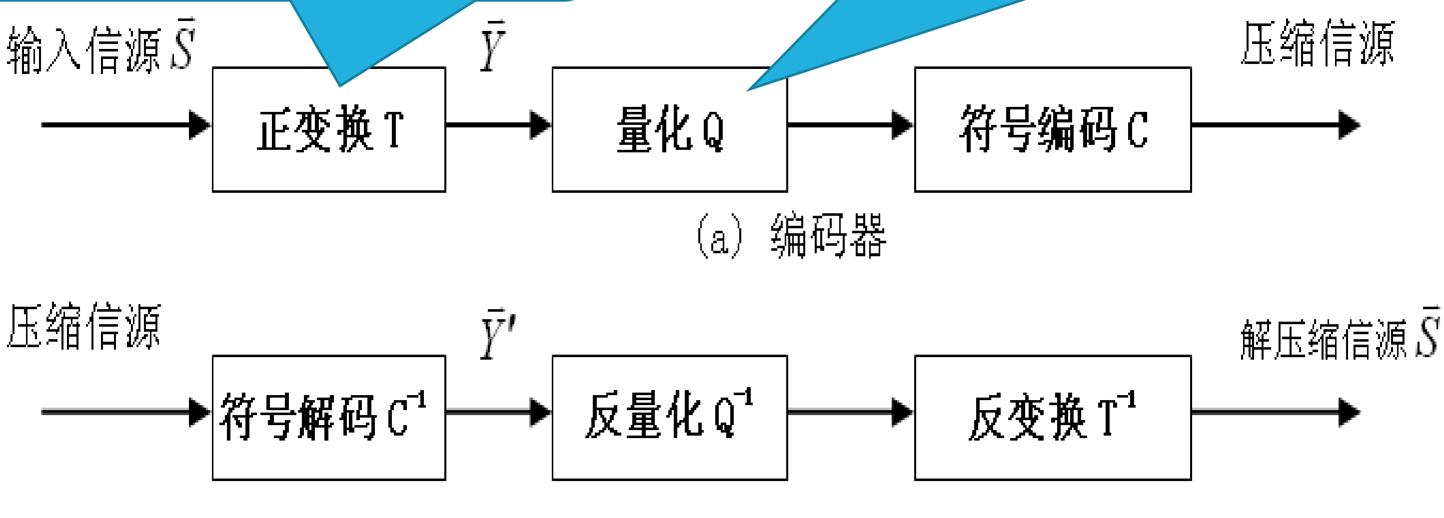
- 对于有记忆信源，由于信源前后符号之间具有较强相关性，要提高信息传输的效率首先需要解除信源符号之间的相关性，解除相关性可以在**时域上进行**（如预测编码），也可以在**变换域上进行**，这就是本节要介绍的变换编码方法
- 对于图像信源等相关性较强的信源，常采用基于正交变换的变换编码方法进行数据压缩
- 变换编码的**基本手段**是将原来在空间（时间）域上描述的信号，通过一种数学变换（例如傅里叶变换等），将信号变到变换域（如频域等）中进行描述
- **在变换域中，变换系数之间的相关性常常显著下降，并常有能量集中于低频或低序系数区域的特点，**这样就容易实现码率的压缩，并还可大大降低数据压缩的难度

2.2 变换编码

- **变换编码要可用，必须要求其反变换存在**
- 严格来说，当反变换存在时，变换编码是一种无损编码。但是在实现时，基于数字系统的精度有限，无法无失真地表示每个变换系数，所以它其实是有损的
- **一般的变换编码系统框图**

使输出的压缩信源矢量中各分量之间的**相关性大大减弱**，而且使**能量集中**到少数几个分量上

量化后等于“0”的系数可以不传送，因此在一定保真度准则下可**达到压缩数据率的目的**，量化参数的选取主要根据保真度要求或恢复信号的主观评价效果来确定



(b) 解码器

2.2 变换编码

- 在变换编码方法中最关键的是正交变换的选择。
- **最佳的正交变换是KL变换**，由于KL变换使变换后随机矢量的各分量之间完全独立，因而它常作为衡量正交变换性能的标准，在评价其它变换的性能时，常与KL变换的结果进行比较。
- KL变换的最大**缺点**是计算复杂，而且其变换矩阵与信源有关，实用性不强。

- 为此人们又找出了各种较实用的变换，如:

  离散傅里叶变换（DFT）

  离散余弦变换（DCT）

  沃尔什-哈达玛变换（WHT）等等

  其中性能较接近KL变换的是离散余弦变换（DCT），在某些情况下，DCT能获得与KL变换相同的性能，因此**DCT也被称为准最佳变换**。

2.2 变换编码

- 离散余弦变换（DCT）
  - 离散余弦变换是DFT的一种特殊形式，当基函数为离散余弦函数时，DFT就成为了DCT。
  - 离散余弦变换主要用于图像压缩编码，基本方法是先将整体图像分成N×N像素块，然后对N×N像素块逐一进行DCT变换。变换后得到像素块的频率系数矩阵，其中对应的高频成分系数通常为零，而低频成分系数有较小的值。
- 离散余弦变换分一维DCT和二维DCT：
  - （1）一维DCT为：

$$C(u)=\begin{cases}\frac{1}{\sqrt{2}}, u=0\\ 1, u\neq 0\end{cases}$$

C(u)为变换系数函数

$$F(u)=C(u)\sqrt{\frac{2}{N}}\sum_{x=0}^{N-1}f(x)\cos\frac{(2x+1)u\pi}{2N} \qquad u=0,1,\cdots$$

$$f(x)=C(u)\sqrt{\frac{2}{N}}\sum_{u=0}^{N-1}F(u)\cos\frac{(2x+1)u\pi}{2N} \qquad x=0,1,\cdots$$

f(x)为长度为N的序列，F(u)也为长度为N的序列。

2.2 变换编码

- **（2）二维DCT为：**

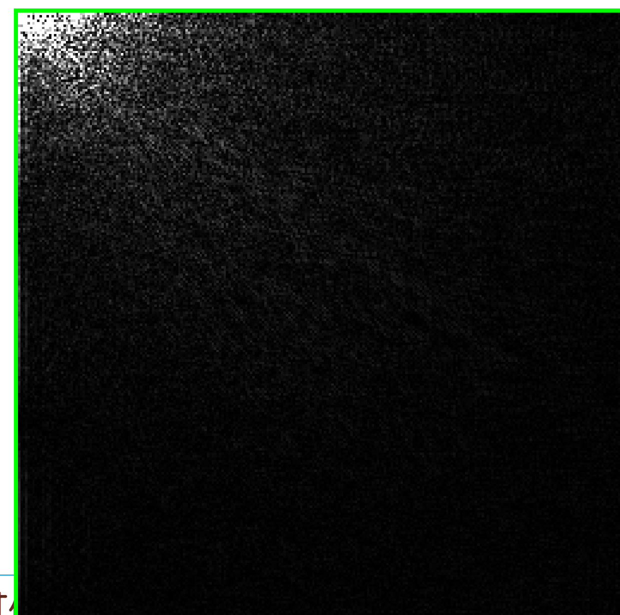
$$F(u,v)=\frac{2C(u)C(v)}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)\cos\frac{(2x+1)u\pi}{2N}\cos\frac{(2y+1)v\pi}{2N}$$

$$C(u)C(v)=\begin{cases}\frac{1}{\sqrt{2}},u,v=0\\ 1,u,v\neq 0\end{cases}$$

$$f(x,y)=\frac{2}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}C(u)C(v)F(u,v)\cos\frac{(2x+1)u\pi}{2N}\cos\frac{(2y+1)v\pi}{2N}$$

$u,v=0,1,\cdots \quad \cdots$

- **例：图像信源如图所示，对它们进行DCT，则其DCT系数图所示，图中亮度越大表示对应的DCT系数的值也越大**



2.3 统计编码

- 统计编码的基本思想：
  - 对于无记忆的信源，根据码字出现的概率分布特性寻找概率与码字长度间的最优化匹配，据此对信息进行压缩。
  - 根据消息出现概率的分布特性而进行的压缩编码，其目的在于在消息和码字之间找到明确的一一对应关系，以便在恢复时能够准确无误地再现出来，或者至少是极相似的找到相当的对应关系。
- 统计编码是一种无损编码，常用于图像、文档等要求无损失的压缩中
- 实现原理：
  - 有些媒体（如图像，文档）数据中各样点值的出现概率在编码前可以统计出，结合其出现概率进行的编码可以充分降低数据量，同时又保证了媒体的质量。

2.3 统计编码——霍夫曼编码

- 1952年，霍夫曼（Huffman）提出了一种构造最佳码的方法，这是一种最佳的逐个符号的编码方法，一般称作**霍夫曼码**
- 基本思想：
  - 对于**出现概率较大的符号取较短的码长**，而对概率较小的符号则取较长的码长。
- 霍夫曼编码（Huffman Encoding）又称变长度编码，或最优编码，即遵照霍夫曼编码原则的结果一定是**平均码长最短**
- 霍夫曼编码的特点：只适用于有限个离散信源；且实现起来相当复杂

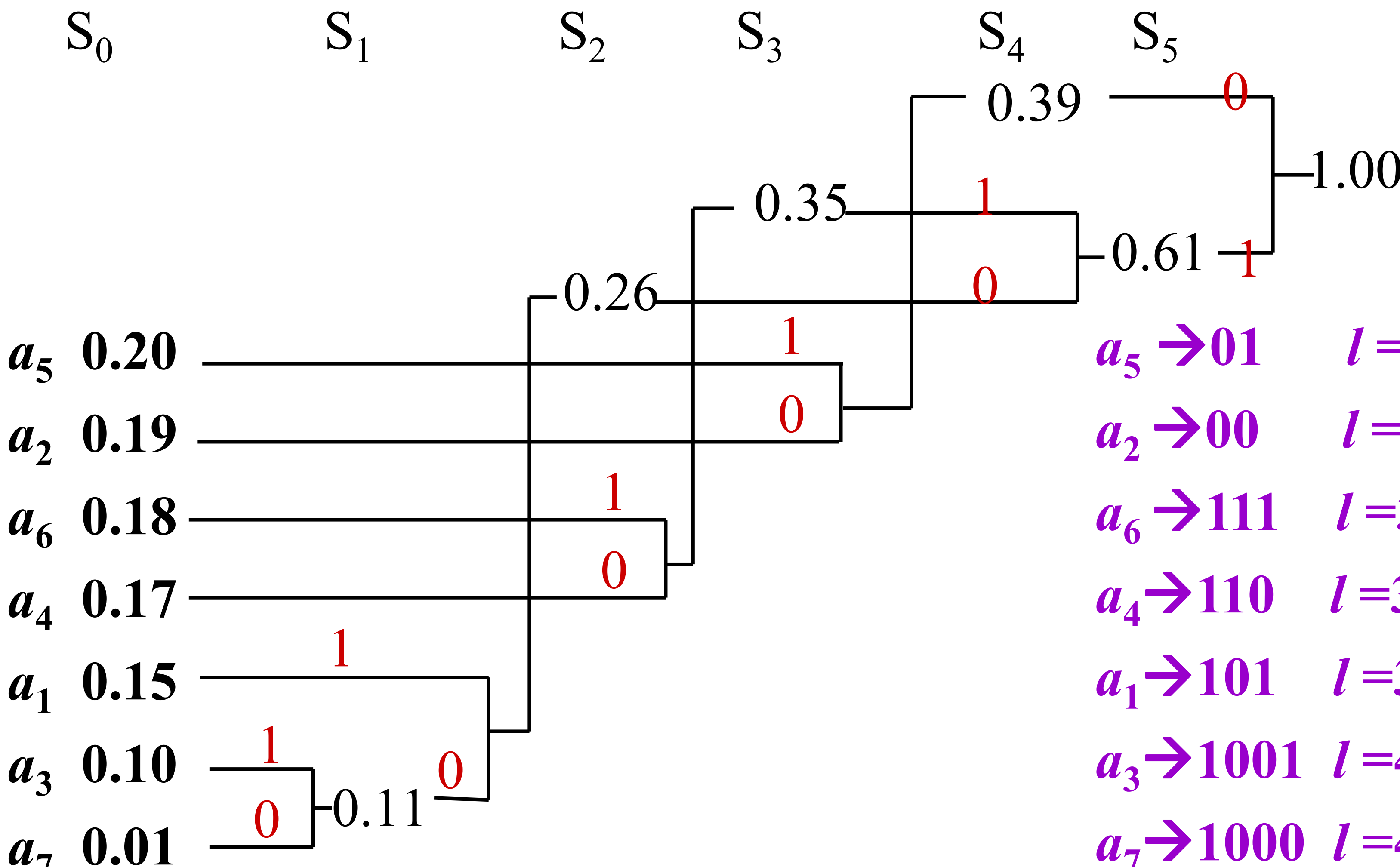
2.3 统计编码——霍夫曼编码

编码步骤（对二元系统）：

1. 将q个信源符号按**概率递减**的方式排列起来；
2. 用“0”,“1”码符号分别表示概率最小的两个信源符号，并将这两个概率最小的信源符号合并成一个符号，从而得到只包含 q-1个符号的新信源，称之为**S信源的缩减信源S_1**；
3. 将缩减信源S_1的符号仍按概率大小以递减次序排列，再将其最后两个概率最小的符号合并成一个符号，并分别用“0”，“1”码符号表示，这样又形成了q-2个符号的**缩减信源S_2**；
4. 依次下去，直到信源**最后只剩下两个符号**为止，将这最后两个符号分别用“0”，“1” 码符号表示;
5. 从最后一级缩减信源开始，**向前返回读码符号**，得出各信源符号所对应的码符号序列，即得出对应的码字。

[例] 将下列消息按二元霍夫曼方法编码

$$\begin{bmatrix} S \\ p(a_i) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 \\ 0.15 & 0.19 & 0.1 & 0.17 & 0.2 & 0.18 & 0.01 \end{bmatrix}$$



$a_5 \rightarrow 01 \quad l=2$

$a_2 \rightarrow 00 \quad l=2$

$a_6 \rightarrow 111 \quad l=3$

$a_4 \rightarrow 110 \quad l=3$

$a_1 \rightarrow 101 \quad l=3$

$a_3 \rightarrow 1001 \quad l=4$

$a_7 \rightarrow 1000 \quad l=4$

2.3 统计编码——霍夫曼编码

从：$\begin{bmatrix} S \\ p(a_i) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 \\ 0.15 & 0.19 & 0.1 & 0.17 & 0.2 & 0.18 & 0.01 \end{bmatrix}$

可求出信源的熵为：

$$H(S) = -\sum_{i=1}^{7} p(a_i)\log p(a_i) = 2.61(\text{比特/符号})$$

从：***l*** **=** (2,2,3,3,3,4,4) 可求得平均码长为：

$$\overline{L} = \sum_{i=1}^{7} p(a_i) l_i = 2.72\ (\text{二元码符号/信源符号})$$

编码效率为：$\eta = \dfrac{H(S)}{\overline{L}} = 0.96$

$a_5 \rightarrow 01$ l=2

$a_2 \rightarrow 00$ l=2

$a_6 \rightarrow 111$ l=3

$a_4 \rightarrow 110$ l=3

$a_1 \rightarrow 101$ l=3

$a_3 \rightarrow 1001$ l=4

$a_7 \rightarrow 1000$ l=4

2.3 统计编码——霍夫曼编码

按霍夫曼码的编码方法，可知这种码**有如下特征：**

①它是一种**分组码**：各个信源符号都被映射成一组固定次序的码符号；

②它是一种**即时码**：一串码符号中的每个码字都可不考虑其后的符号解出来；

③它是一种**唯一可解的码**：任何码符号串只能以一种方式解。

$a_5 \rightarrow 01$	$l=2$
$a_2 \rightarrow 00$	$l=2$
$a_6 \rightarrow 111$	$l=3$
$a_4 \rightarrow 110$	$l=3$
$a_1 \rightarrow 101$	$l=3$
$a_3 \rightarrow 1001$	$l=4$
$a_7 \rightarrow 1000$	$l=4$

解码:霍夫曼码串可通过从左到右检查各个符号进行解码。本例中码串：1001000111000111可译为：

$$\underbrace{1001}_{a_3}\underbrace{00}_{a_2}\underbrace{01}_{a_5}\underbrace{110}_{a_4}\underbrace{00}_{a_2}\underbrace{111}_{a_6}$$

说明：

- 霍夫曼码是一种即时码，可用码树形式来表示。
- 每次对缩减信源最后两个概率最小的符号，用“0”和“1”码是可以任意的，所以可得到不同的码，码长 l_i 不变，平均码长也不变。
- 若当缩减信源中缩减合并后的符号的概率与其他信源符号概率相同时，从编码方法上来说，它们概率次序的排列没有限制，因此可得到不同的码。

➔对给定信源，霍夫曼编码方法得到的码**并非是唯一**，**但平均码长不变**。

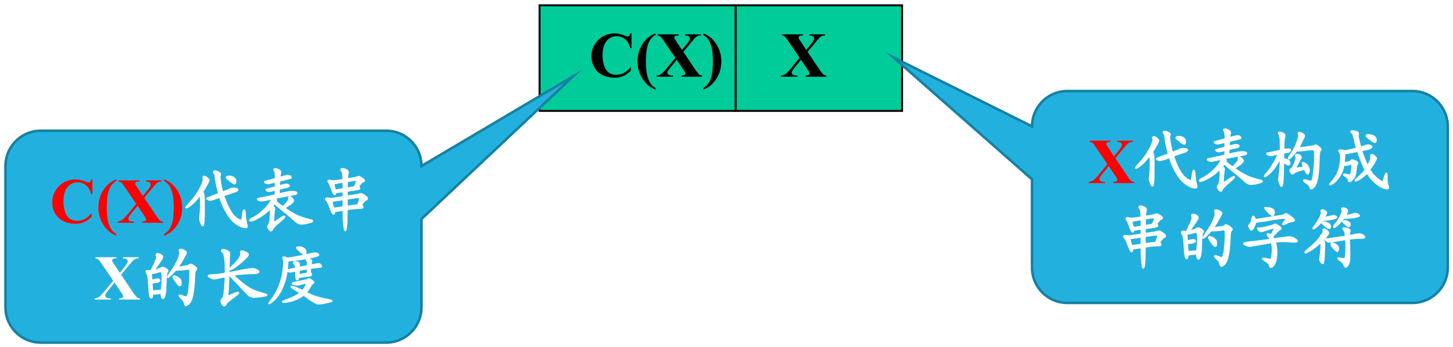
- **基本思想：**

将数据流中连续出现的字符或像素值用该符号加出现次数来表示。

- **游程编码又称为行程编码（Run-length Encoding），也是一种常用于文档和图像中的编码方式。**
- **RLE编码简单直观，编码/解码速度快，因此许多图形和视频文件，如BMP、TIFF及AVI等格式文件的压缩均采用此方法。**

编码方法：

基本的游程编码就是在数据流中直接用2个字符来表示，其数据结构如图所示：



例：字符串AAABCDDDDDDDDBBBBB，游程编码结果为3A1B1C8D5B.

2.3 统计编码——游程编码

- 优点：

1、编码方式简单，直观。

2、译码和编码采用相同的规则进行，得到与压缩完全相同的数据，从而可以实现无损压缩。

- 缺点：

1、在特定情况下，如单一颜色背景下的图形图像压缩中可以取得较高的压缩比；而对于复杂颜色的图形图像，则压缩比较低。

2、由于其在压缩实现过程中未考虑被压缩数据的物理含义，所以获得的压缩比也较低。

2.3 统计编码——算数编码

❑基本思想：

类似于Huffman编码，对概率较大的符号采用短码，对概率较小的符号采用长码，但Huffman编码只能使用整数比特，而它可以利用分数比特逼近于信源。

❑基本原理：

1、根据信源中出现不同符号序列的概率不同，把[0,1)区间划分为互不重叠、宽度恰好是各符号序列的概率的子区间。

2、信源中的各符号序列将可用各子区间中的任意一个实数表示，这个数就是该符号所对应的码。

2.3 统计编码——算数编码

❑算术编码步骤：

1、建立信源概率表。

2、扫描信源发出的符号序列，对其进行编码。

❑分类：

1、静态算术编码：信源符号概率是固定的算术编码。

2、自适应算术编码：信源符号概率是动态变化的算术编码。

❑举例说明静态算术编码：

假设信源符号为{00,01,10,11}，这些符号概率分别为{0.1,0.4,0.2,0.3}。

根据这些概率，可把间隔[0,1)分成4个子间隔：

[0,0.1),[0.1,0.5),[0.5,0.7),[0.7,1)。

如果二进制消息序列的输入为：10 00 11 00 10 11 01

编码过程：第一个符号为10，它的编码范围对应[0.5,0.7)而第二个符号00对应的编码范围是[0,0.1)。因此，它的间隔就取[0.5,0.7)的第一个1/10作为新间隔[0.5,0.52)依此类推……

❑自适应算术编码：

首先假定各符号概率的初始值相同，然后其概率根据出现的情况做相应的改变。

自适应模式可以不预先定义概率模型，但要求编码器和译码器使用相同的概率模型。

自适应算术编码的编码效率很高，当信源符号概率比较接近时，可优于Huffman编码。

具体编码过程略。

内 容 提 纲

1. 信源编码基础理论
2. 预测编码、变换编码、统计编码
3. 波形编码、参数编码、混合编码
4. 小结

3.1 概述

- 语音编码（Speech Coding） 从信息论角度看，信源编码是要以最少的数码表示信源所发的信号，语音编码属于信源编码的范畴。语音编码通过减少传输码率（或存储量），来达到提高传输（或存储）效率的目的。作为传输语音的压缩技术，语音编码在通信史上一直都扮演着极为重要的角色。
- 语音编码分为三类：

  * 波形编码：重建后的语音时域信号的波形与原语音信号保持一致。

  * 参数编码：通过建立语音信号的产生模型，提取其特征参数来编码，波形上不要求与原信号匹配，又称声码器技术。

  * 混合编码：有机结合以上两种编码方式，基于语音产生模型的假定并采用分析合并技术。

三种编码方式的比较

	波形编码	参数编码	混合编码
编码信息	波形	模型参数	综合
比特率	9.6~64Kbps	2.4~9.6Kbps	16~24Kbps
优点	适应能力强，语音质量好	有效降低了编码比特率	语音质量明显提高
缺点	随着量化粗糙语音质量下降	合成语音质量低，处理复杂度高	编码速率明显上升
典型代表	自适应差分编码调制（ADPCM）	LPC- 10、LPC-10E	多脉冲激励线性预测编码（MPLPC） 规则脉冲激励线性预测编码（RPE-LPC）

语音编码性能的评价指标

编码速率

直接反映了语音编码对语音信息的压缩程度

现在大部分编码标准都是固定速率编码，其范围为0.8kbit/s~64kbit/s

语音质量评价

是语音编码性能的最根本指标，分为两类：主观评价方法和客观评价方法

编解码延时

一般用单次编解码所需时间表示。
公用电话网:不超过5~10ms
移动蜂窝通信系统: 不超过100ms

算法复杂度

影响到语音编解码器的硬件实现，它决定了硬件实现的复杂程度、体积、功耗及成本等

3.2 波形编码

- **波形编码**　波形编码是语音编码系统在早期所广泛采用的方法，它把语音信号当成普通的波形信号来处理从而保持原波形形状。

　　波形编码适应能力强，合成语音质量好，但比特率过高，编码的效率也不尽如人意。

- 几种典型的波形编码：
  - 脉冲编码调制（PCM）
  - 自适应预测编码（APC）
  - 自适应增量调制（ADM）
  - 自适应差分脉冲编码调制（ADPCM）
  - 子带编码（SBC）

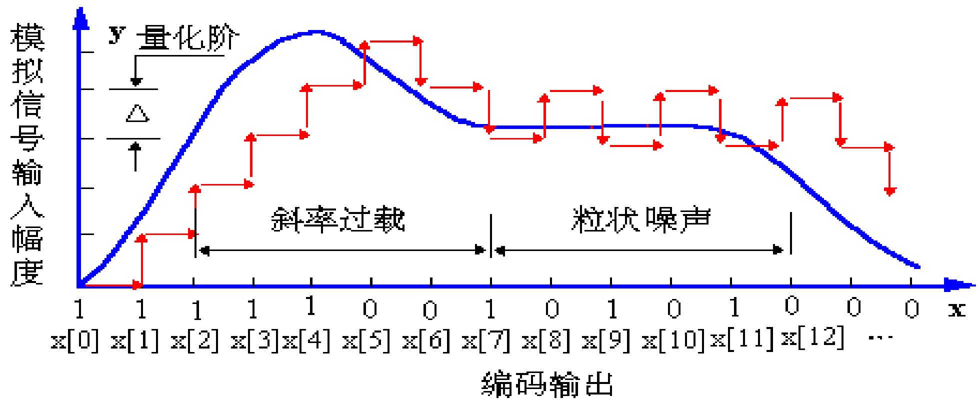
3.2 波形编码

- **脉冲编码调制（PCM）**
  - **形式一：均匀PCM**
    - 最简单最原始的波形编码方式，没有运用压缩技术，产生的比特率也极高，故在当今运用极少。
  - **形式二：非均匀PCM**
    - 将信号进行非线性变换后再均匀量化，变换后信号具有均匀概率密度分布。编码时常采用对数变换压缩（译码时指数扩展）。
  - **形式三：自适应PCM**
    - 无论是均匀或是非均匀PCM，量化间隔总是随着量化器的确定而固定。而自适应PCM引入的自适应幅值变化概念使得量化误差可以匹配于输入信号方差，或是量化器增益G可以随着幅值而变化，从而使信号能量在量化前恒定。
- **自适应预测编码（APC）**
  - **利用线性预测可以改进编码的量化器性能。因为预测误差的动态范围和平均能量均比输入信号小。如果对预测误差进行量化和编码，则量化比特数将减少。在接收端，只要使用与发送端相同的预测器，就可恢复原信号。**

自适应差分脉冲编码调制（ADPCM）

增量调制（Delta modulation，DM）

- DM的主要思路为：首先判别下一个语音信号值比当前的信号值是高还是低，如果高则编码为“1”；如果低则编码为“0”。该方法虽然编码简单，但是会产生斜率过载和颗粒噪声



差分脉冲编码调制（Differential PCM，DPCM）

- DPCM采用多位量化来量化两个采样之间的差分信号，因为差分信号比原语音信号的动态范围和平均能量都小，因此对相邻样本间的差信号（差分）进行编码，可压缩信息量，从而大大降低信道负载。

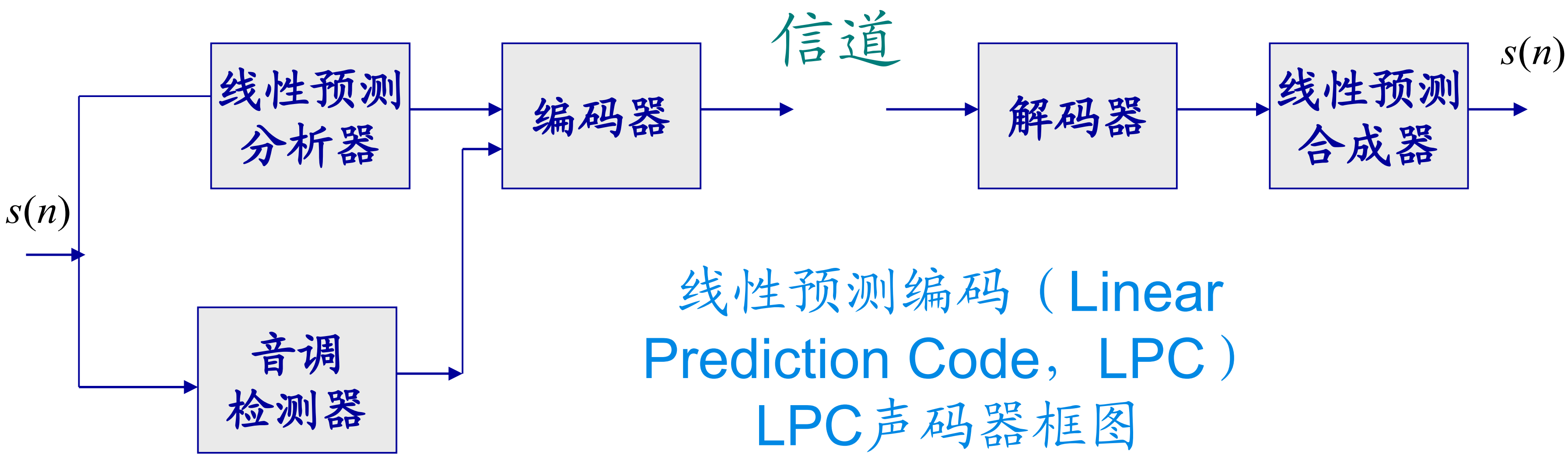
自适应差分脉冲编码调制（ADPCM）

3.3 参数编码

- 参数编码在**提取语音特征参数**时，往往会利用某种**语音生成模型**在幅度谱上**逼近**原语音，以使重建语音信号有尽可能高的可懂度，即力图保持语音的原意，但**重建语音的波形**与原语音信号的波形却**有相当大的区别**。利用参数编码实现语音通信的设备通常称为**声码器**
- 例如通道声码器、共振峰声码器、同态声码器以及广泛应用的线性预测声码器等都是典型的语音参数编码器。其中，比较有实用价值的是**线性预测声码器**，这是因为它较好地解决了编码速率和编码语音质量的问题

3.3 参数编码

- 实际应用中一般采用的实现设备是**线性预测声码器**



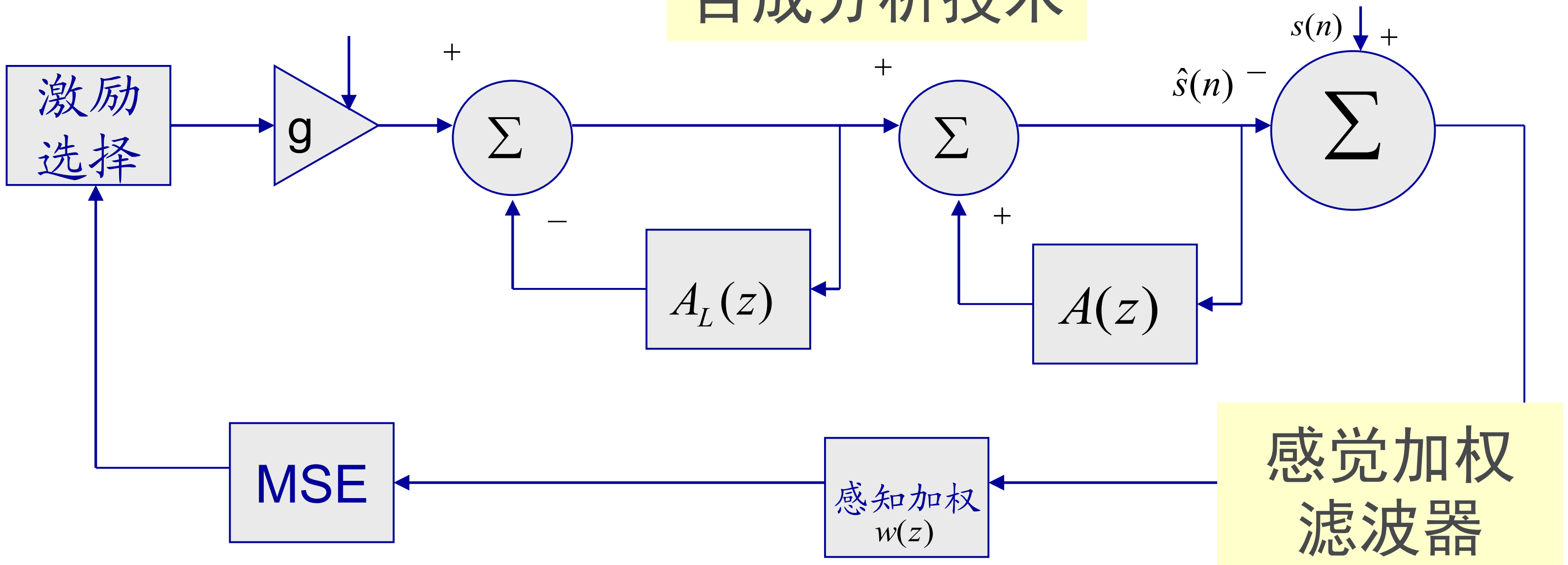
线性预测编码（Linear Prediction Code，LPC）

LPC声码器框图

3.4 混合编码

- **波形编码**能保持较高的语音音质，抗干扰性较好，硬件上也容易实现，但比特速率较高，而且时延较大。**参数编码**的比特速率大大降低，最大可压缩到2kbps左右，但自然度差，语音质量难以提高。80年代后期，综合上述两种方式的**混合编码技术**被广泛使用
- 合成编码方法（LPAS）是目前得到最广泛研究**闭环或分析-合成方法**的语音编码算法，它通过线性预测确定系统参数，并通过确定激励序列

合成分析技术



内 容 提 纲

4 小结

- 信源编码基础理论
  - 信源的分类
  - 信源的数学模型
  - 信源的熵
  - 信源的冗余度
  - 信源编码器模型
  - 码的分类
- 预测编码、变换编码、统计编码
  - 差值预测编码系统
  - 离散余弦变换（DCT）
  - 霍夫曼编码、游程编码、算数编码
- 波形编码、参数编码、混合编码

谢 谢
Q&A

欢迎电子邮件、QQ与微信交流问题！