

计算机图象处理

(下 册)

容 观 澳

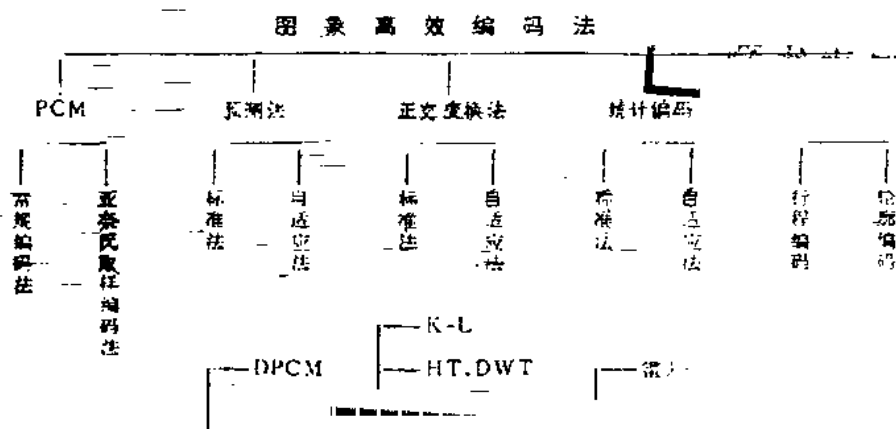
清 华 大 学

1989.8.

1990.9.

数字图象的优点是明显的，这是因为模拟图象的缺点而作的改进，因为模拟图象信号在传输过程中极易受到各种噪声的干扰，而且模拟信号一旦受到污染则很难完全得到恢复。另外，在模拟领域，人与人、人与机器（计算机或智能机），机器与机器之间的信息交换以及对图象的增强、恢复、特征提取和识别等一系列处理都是比较困难的。但是，对于数字图象信号主要的问题是数据量太大，无论是进入计算机，还是保存其数据都是困难的。特别是传输图象首先碰到的困难是图象数字信号占频带太宽——即是通常叫做“信息容量”问题。因此在许多应用领域中，图象信息的传输和存储具有实际的重要性；如数字电视、电视电话、航空侦察、遥感、空间探索、工业中质量控制等应用方面。图象信息的传输和存储要解决的问题是：保持图象的质量或在容许的保真度条件下压缩存储及传输量。数字图象即然需要编码成计算机能处理的信号。通常用标准二进制编码，也叫常规编码 PCM 这种编码也有时叫格雷码或格雷码反码。所谓信息压缩是表示同样图象象元信号所需比特数和常规编码所用比特数的比值表示。能不能压缩呢？由于一般的图象中有大量的、统计性质的多余度或相关性，例如邻近两点几乎有同样的灰度等级、大部分图象中含有生理视觉多余度等情况。这就是说可以去掉一大部分图象的内容，而并不会影响视觉的图象质量，也不会全部丧失图象有用的细节部分。因此，把这些多余的部分略去，就可减少信道容量或存储量。编码除了为解决上述两个问题——数据压缩、图象传输以外，第三个目的是提取图象中的特征，以便快速识别及实时控制。从数据压缩角度看编码主要分为两大类：(1)信息保持的压缩(2)保真度压缩（保持一定的图象质量下的压缩。另外，如果从目前已有的实用方案的角度来分类，可以分为三大类，即预测编码，变换编码及统计编码。而这些方法既适用于固定图象编码，也适用于电视信号编码。就具体编码方法而言可简略地概括见表 7-1 中。

表 7-1



下面我们将讨论如何维持图象的保真度以及为了压缩图象需要了解图象信息的概念以及各种编码方法。应当强调指出，不同的应用目的所采用编码方法有很大区别。

§ 7-1 图象编码中的保真度准则

一、客观保真度准则

通常使用的客观保真度准则有输入图象和输出图象的均方根误差；输入图象和输出图象的均方根信噪比两种。输入图象和输出图象的均方根误差是这样定义的。设输入图象是由 $N \times N$ 个象素组成，令其为 $f(x, y)$ ，其中 $x, y = 0, 1, 2, \dots, N-1$ 。这样一幅图象经过压缩编码处理后，送至受信端，再经译码处理，重建原来图象，这里令重建图象为 $g(x, y)$ 。它同样包含 $N \times N$ 个象素，并且 $x, y = 0, 1, 2, \dots, N-1$ 。在 $0, 1, 2, \dots, N-1$ 范围内 x, y 的任意值，输入象素和对应的输出象素之间的误差可用下式表示

$$e(x, y) = g(x, y) - f(x, y) \quad (7-1-1)$$

而包含 $N \times N$ 象素的图象之均方误差为

$$\begin{aligned} e^2 &= \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} e^2(x, y) \\ &= \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} [g(x, y) - f(x, y)]^2 \end{aligned} \quad (7-1-2)$$

由式(7-1-2)可得到均方根误差为

$$e_{rms} = (e^2)^{\frac{1}{2}} \quad (7-1-3)$$

如果把输入、输出图象间的误差看作是噪声，那么，重建图象 $g(x, y)$ 可由下式表示

$$g(x, y) = f(x, y) + e(x, y) \quad (7-1-4)$$

在这种情况下, 另一个客观保真度准则, 重建图象的均方信噪比如下式表示

$$\begin{aligned} \left(\frac{S}{N}\right)_{ms} &= \frac{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} g^2(x, y)}{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} e^2(x, y)} \\ &= \frac{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} g^2(x, y)}{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} [g(x, y) - f(x, y)]^2} \end{aligned} \quad (7-1-5)$$

均方根信噪比为

$$\left(\frac{S}{N}\right)_{rms} = \left[\frac{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} g^2(x, y)}{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} [g(x, y) - f(x, y)]^2} \right]^{\frac{1}{2}} \quad (7-1-6)$$

二、主观保真度准则

图象处理的结果, 绝大多数场合是给人观看, 由研究人员来解释的, 因此图象质量的好坏与否, 既与图象本身的客观质量有关, 也与人的视觉系统的特性有关。有时候客观保真度完全一样的两幅图象可能会有完全不同的视觉质量, 所以又规定了主观保真度准则。这种方法是把图象显示给观察者, 然后把评价结果加以平均, 以此来评价一幅图象的主观质量。另外一种方法是规定一种绝对尺度, 例如:

- (1) 优秀的: 具有极高质量的图象;
- (2) 好的: 是可供观赏的高质量图象, 干扰并不令人讨厌;
- (3) 可通过的: 图象质量可以接受, 干扰并不讨厌;
- (4) 边缘的: 图象质量较低, 希望能加以改善, 干扰有些讨厌;
- (5) 劣等的: 图象质量很差, 尚能观看, 干扰显著地令人讨厌;
- (6) 不能用: 图象质量非常之差, 无法观看。

另外常用的还有两种准则, 即妨害准则和品质准则

妨害准则如下:

- (1) 没有妨害感觉;
- (2) 有妨害, 但不讨厌;
- (3) 能感到妨害, 但没有干扰;

- (4) 妨害严重, 并有明显干扰;
- (5) 不能接收信息。

品质准则如下:

- (1) 非常好;
- (2) 好;
- (3) 稍好;
- (4) 普通;
- (5) 稍坏;
- (6) 恶劣;
- (7) 非常恶劣。

除此之外, 还可以采用成对比较法, 也就是同时出示两幅图象, 让观察者表示更喜欢哪一幅图象, 借此排出图象质量的等级。也有采用随机抽取法来评定图象质量的。这种方法是把数量相等的原始图象和经编译后的图象混杂在一起, 然后让观察者挑出他认为质量差的图象。质量较差的图象可定义为处理过的图象, 然后统计错挑的概率, 显然错挑概率越大说明图象经处理后的劣化越小。总之, 主观保真度评价方法的准则可不同, 但其基本原理都一样, 当然, 对观察者的视觉条件应有一定的要求。

§ 7-2 PCM编码

一、PCM编码基本原理

脉冲编码调制 (即 PCM 编码) 是将模拟图象信号变为数字信号的基本手段。图象信号的 PCM 编码与语音信号 PCM 编码相比并没有原则上的区别。但是, 图象信号, 特别是电视信号占的频带宽, 要求响应速度快, 因此电路设计与实现上有较大的难度。

图象信号 PCM 编、译码原理方框图如图 7-2-1 所示。图象信号 PCM 编码由前置低通滤波器、取样保持电路、量化器、编码器组成。前置低通滤波器的作用有二个, 一是为满足取样定理的带限要求, 以减少折迭误差; 其二是对杂散噪声也有一定的抑制作用。取样保持电路将完成把时间上是连续的模拟信号进行时间离散化的任务。取样周期由奈奎斯特 (Nyquist) 定理限定。一种适用于视频信号的高速取样保持电路原理图及其输出波形如图 7-2-2 所示。

量化器的任务是把模拟信号的幅值离散化。经取样与量化处理后可产生多值数字信号。编码的任务是把多值的数字信号变成二进制的数字信号, 以便传输或进行后续处理。译码器的原理比较简单, 它包括一个译码电路和一个低通滤波器。译码器把数字信号恢复为模拟信号, 这个模拟信号就是在接收端重建的图象信号。

二、PCM量化噪声

量化是对时间离散的模拟信号进行幅度离散化的过程。这个过程是去

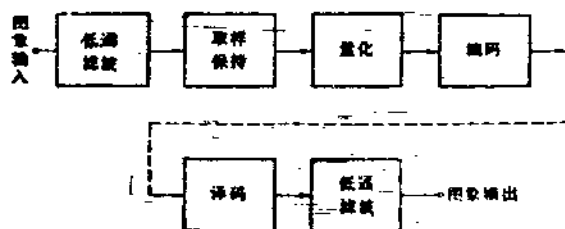


图 7-2-1 PCM编、译码原理方框图

零取整的过程。量化后的样值与原信号相比大部分是近似关系。这样，把连续的数值限制在固定的台阶式的变化之下必然会带来畸变。这种畸变在接收端是无法克服的，只能使其尽量减小。由量化带来的噪声可分为量化噪声和过载噪声。以正弦波输入为例，输入幅度较大和输入幅度较小时的量化噪声如图 7-2-2 所示。图中 (a) 是输入信号超过编码范围时的量化噪声和过载噪声的形成，(b) 是信号未超过编码范围，只有量化噪声的情况。在 PCM 编码中，量化噪声主要取决于码的位数，码位数越多（即量化阶数多）量化噪声的功率越小。一个量化阶的电路可以由下式表示

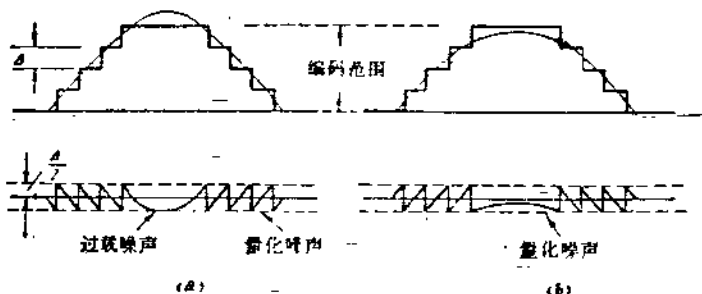


图 7-2-2 量化噪声与过载噪声的形成

$$\Delta = \frac{V}{2^n} \quad (7-2-1)$$

式中 V 为输入信号电压； n 为样值用二进制数表示的比特数。

如果在整个输入幅度内量化阶是一个常数，就称这个量化为均匀量化，否则就是非均匀量化。线性 PCM 编码中均采用均匀量化法。在均匀量化中，设量化阶为 Δ ，量化噪声在 $-\frac{\Delta}{2} \sim +\frac{\Delta}{2}$ 内可看成是均匀分布的，因此其功率可由下式表示

$$P_Q = \int_{-\frac{\Delta}{2}}^{+\frac{\Delta}{2}} \frac{1}{\Delta} x^2 dx = \frac{\Delta^2}{12} \quad (7-2-2)$$

对于过载噪声，当量化特性输入过载点为 V 时，由下式表示

$$N_S = \int_{-\infty}^{-V} (x+V)^2 p(x) dx + \int_V^{\infty} (x-V)^2 p(x) dx \quad (7-2-3)$$

式中 N_S 为过载噪声， x 是输入信号值， $p(x)$ 为输入幅度的概率密度。如果用信噪

比作为客观保真度准则的话，可推得PCM编码在均匀量化下的量化信噪比如下

$$\text{因为 } \Delta = \frac{V}{2^n}$$

$$P_Q = \frac{\Delta^2}{12}$$

$$\frac{V^2}{12}$$

$$\text{所以 } P_Q = \frac{(2^n)^2}{12} \quad (7-2-4)$$

由信噪比的概念，则

$$\begin{aligned} \left(\frac{S}{N}\right)_{dB} &= 10 \lg \frac{V^2}{P_Q} \\ &= 10 \lg \frac{V^2}{\frac{(2^n)^2}{12}} \\ &= 10 \lg \{12 \times (2^n)^2\} \\ &= 20 \lg \sqrt{12} + 20 \lg 2^n \\ &\approx (11 + 6n) \end{aligned} \quad (7-2-5)$$

由式(7-2-5)可见，每增加一位码可得到6 dB的信噪比得益。

值得注意的是量化噪声不同于其他噪声，它的显著特点是仅在信号输入时才出现，所以它是数字化中特有的噪声。一般情况下，直接测量比较困难。

三、编码器、译码器

编码器的任务是把一个多值的数字量用多比特的二进制量来表示。如果量化器输出M个值，那么，对应于M个值中的任何一个值编码器将给定一个二进制码字。

这个码字将由m个二进制数组成。通常情况下 $M = 2^m$ 。编码器的输入与输出关系是一一对应的，其过程是可逆的，因此不会引入任何误差。

线性PCM编码一般采用等长码，也就是说每一个码字都有相同的比特数。其中用得为普遍的是自然二进制码，也有用格雷码的。以 $M = 8$ 为例的自然二进制码和格雷码列入表7-2-1。

对于常规编码来说，减少量化分层数就可以降低比特速率。但是，量化分层数的最小值应满足图象质量的要求。当主观评定图象质量时，为了防止伪轮廓效应，量化分层数必须足够大。实践证明，对于线性PCM编码，黑白图象要6~7 bit，相当于分层数为64到128层，彩色复合编码要8 bit，即256个量化分层，这样恢复的

输入	自然二进制码	格雷码
m_1	0 0 0	0 0 0
m_2	0 0 1	0 0 1
m_3	0 1 0	0 1 1
m_4	0 1 1	0 1 0
m_5	1 0 0	1 1 0
m_6	1 0 1	1 1 1
m_7	1 1 0	1 0 1
m_8	1 1 1	1 0 0

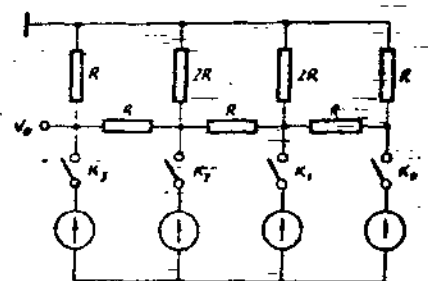


图 7-2-3 电流相加式译码网原理

图象才能与原模拟图象相比拟。如果码位不够就会出现明显的伪轮廓。随着比特数的降低伪轮廓愈加明显。

译码器通常采用电流相加式译码网，原理图如图 7-2-3 所示。在这种方案中，每一位用一个独立的电流源，可以通过分别调整每一位的恒流源来补偿网络误差，因此这种电路对电阻网的精度要求可适当放宽。这种方案的另一个优点是开关的接触电阻不会影响节点上的电压，因此对开关电路的要求不高。电流相加式译码网由三部分组成，这三部分是 T 型电阻网、电子开关、恒流源。总输出电压由下式表示

$$V_0 = \sum_{i=0}^n a_i \cdot 2^i \cdot E_0 \quad (7-2-6)$$

式中 a_i 等于 1 或 0，当第 i 位开关闭合时则为 1，否则为 0。 E_0 是开关闭合时的输出电压。

$$E_0 = \frac{1}{2} R I \left(\frac{1}{2} \right)^n \quad (7-2-7)$$

I 是恒流源输出电流。译码电路的电原理图如图 7-2-4 所示。

四、非线性 PCM 编码

在线性 PCM 编码中，量化阶是均匀的。这样，在小信号输入的情况下信噪比较低，在大信号输入的情况下信噪比较高。为了改善小信号在量化过程中的信噪比，采用一种瞬时压缩扩张技术。这种技术实际上是降低大信号时的信噪比提高小信号时的信噪比，其结果是在不增加数码率的情况下，使信号在整个动态范围内有较均衡的信噪比。

瞬时压扩技术基本有两种方案。一种是如图 7-2-5 所示的方案。在这种方案中，首先将取样后的 PAM 信号进行非线性压缩，然后对压缩后的 PAM 信号进行线性编码。在接收端，首先进行线性译码，然后再送入瞬时扩张器进行非线性扩张，恢复原来的 PAM 信号。这样方案一般采用二极管电压、电流的非线性特性来实现。通常采用图 7-2-6 所示的修正的对数特性。在图 7-2-6 中，对于压缩器来说， X 表示输入， Y 表示输出。对于扩张器来说， Y 表示输入， X 表示输出。参变量 μ 是表示压缩或扩张程度的量。显然，当 $\mu = 0$ 时就是线性编码了。

这种方案电路比较简单，但是，要保证压缩与扩张特性的匹配比较困难。特别是在

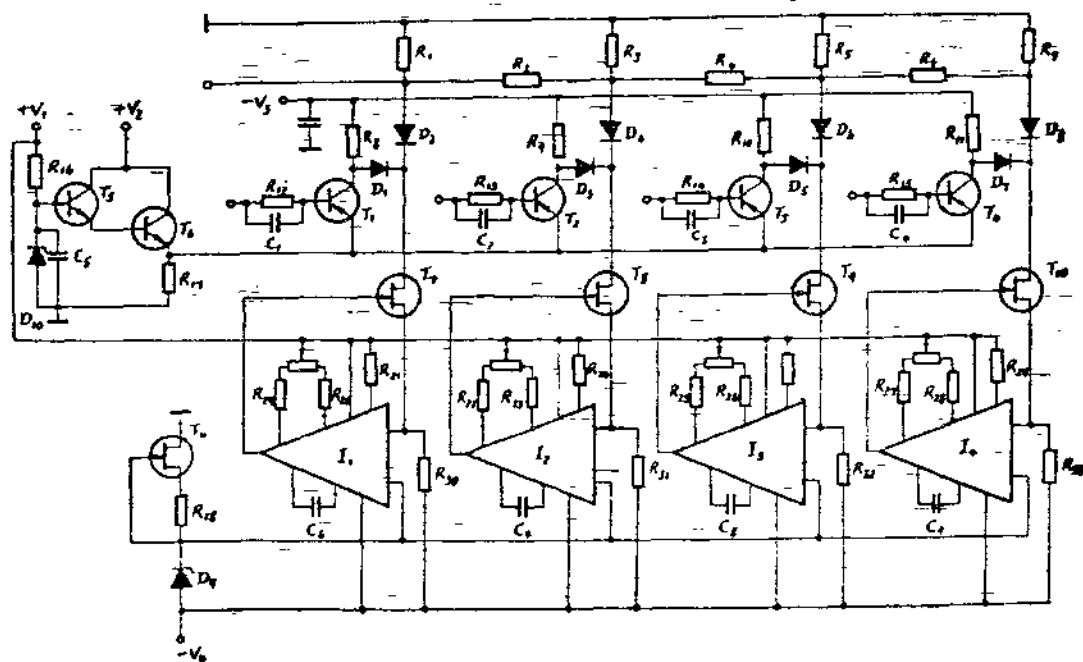


图 7-2-4 译码电路原理图

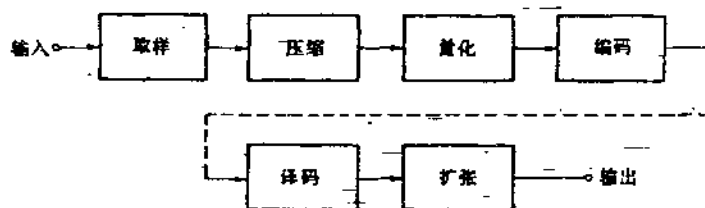


图 7-2-5 一种非线性压扩方案

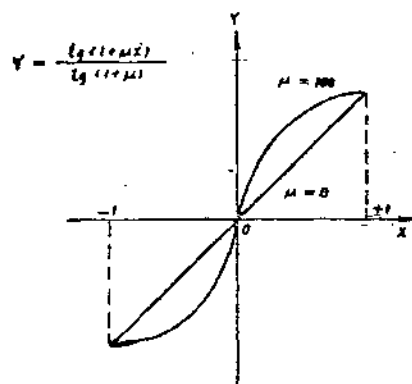


图 7-2-6 理论压扩特性曲线

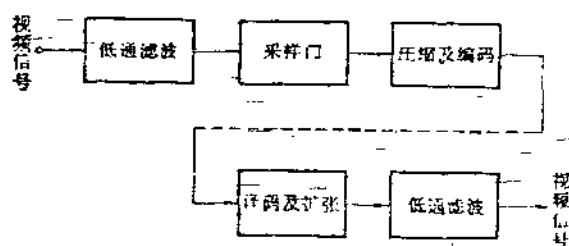


图 7-2-7 数字化非线性压扩技术框图

温度变化范围很宽的情况下必须有恒温设备，因此给生产和调试均增加很大麻烦。

另外一种方案是数字化非线性压扩技术，其原理框图如图 7-2-7 所示。这种方案是把编码与压缩，译码与扩张都分别在编码和译码中一次完成。数字式非线性编码的压扩特性有 μ 特性、A 特性等等。根据 CCITT 1970 年的建议，通常采用 1/3 折

线 ($A = 87.6$) 的压扩特性。13折线压扩特性如图 7-2-8 所示。各折线的斜率列于表 7-2-2 中。由图中可见, 各段折线的斜率是不一样的; 4 至 8 段的小信号区的信噪比都得到了改善。图中的 $u_{\text{入}}$ 表示压缩器的输入, $u_{\text{出}}$ 表示压缩器的输出, V 为过载点电压。图 7-2-8 只画出了信号在正半周时的情况, 负半周时也一样。由于正半周的 7、8 两段和负半周的 7、8 两段斜率都一样, 所以在整个特性中这四段连成一条直线。因此, 总共有 13 条直线段, 简称 13 折线。

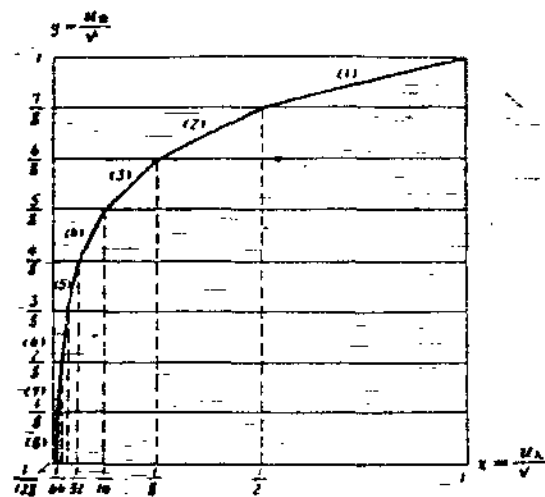


表 7-2-2 各折线段斜率表

折线段	1	2	3	4	5	6	7	8
斜率	$\frac{1}{4}$	$\frac{1}{2}$	1	2	4	8	16	16

图 7-2-8 13折线压扩特性 (信号为正时的八段)

如果令 $x = \frac{u_{\text{入}}}{V}$, $y = \frac{u_{\text{出}}}{V}$, 上述 13 折线可用下式近似表示

$$y = \frac{Ax}{1 + \ln A} \quad 0 < x \leq \frac{1}{A} \quad (7-2-8)$$

$$y = \frac{1 + \ln Ax}{1 + \ln A} \quad \frac{1}{A} < x \leq 1 \quad (7-2-9)$$

式中 A 是一常数, 不同的 A 值可决定一条不同的曲线。在原点处的斜率由下式表示

$$\frac{dy}{dx} = \frac{A}{1 + \ln A} \quad (7-2-10)$$

当原点处的斜率为 16 时, 则

$$\frac{dy}{dx} = \frac{A}{1 + \ln A} = 16$$

可求得 $A = 87.6$ 。图 7-2-8 所示的折线就是 $A = 87.6$ 的 13 折线压缩特性。

总的来说, 如果采用第一种方案, 当 $\mu = 100$ 时, 在输入信号最小情况下信噪比可提高 26 dB, 当采用第二种方案时, 在 $A = 87.6$ 的情况下, 最小信号输入时的信噪比改善大约 24 dB。可见, 两种方案的效果基本相似, 在小信号输入情况下, 信噪比的改善是很显著的。数字化非线性压扩技术的电路实现比第一方案要复杂。但随着数

字集成电路的发展, 电路制做方面的障碍已不复存在。它能较容易地实现压扩特性的匹配, 而且比较稳定可靠, 不用恒温设备, 在生产中也可以保证编、译码的精度, 调试与维护也较简单, 因此这种方案用得愈来愈广泛。

五、亚奈奎斯特取样 PCM 编码

线性 PCM 编码是最基本的数字化手段, 数字化处理时取样速率必需满足奈奎斯特取样定理的要求, 否则, 会产生混叠误差而不能在收端恢复原图象信息。取样定理可由下式表示

$$F = f \times 2 \quad (7-2-11)$$

式中 f 是模拟信号的频带宽度, F 是取样速率。PCM 编码的速率由下式表示

$$R = 2 f n \quad (7-2-12)$$

式中 n 是每样值的比特数。从式 (7-2-12) 可见, 降低 R 的方法有二个, 一是降低 n 值, 其次是降低取样速率。非线性 PCM 方法实质上是用非线性措施尽量减少 n 值。也就是说在尽可能小的 n 值下尽量改善画面的质量。鉴于对画面灰度层次的要求, 非线性 PCM 编码对 R 的减少是有限的。

能否用降低取样率的方法来减少 R 呢? 如果能解决混叠误差问题, 那么, 这种设想也是可行的。经分析, 电视信号的频谱如图 7-2-9 所示。由图可见, 电视图象信号的能量不是均匀连续地分布在带宽范围内, 而是集中在以行频的各次谐波为中心的一束束谱线内。其最高谐波次数为 $352 f_h$ (以 5.5 MHz 带宽计)。这里 f_h 代表行频。如果用低于奈奎斯特取样速率的频率取样, 必然会产生折叠噪声。如果能够使折叠部分如图 7-2-10 所示那样落在原信号频谱的间隙内, 在接收端再采用梳状滤波器将这些折叠噪声滤除, 也可以消除折叠噪声, 正确恢复原来图象。这就是亚奈奎斯特取样 PCM 编码的基本原理。亚奈奎斯特取样 PCM 编码的原理方框图如图 7-2-11 所示。

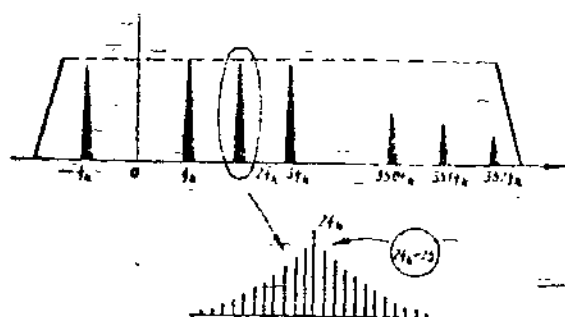


图 7-2-9 电视信号的谱



图 7-2-10 亚奈奎斯特取样频谱折叠部分的设置

这种方案在单色电视信号或彩色电视信号各分量信号编码中, 可选择取样频率等于半行频的奇数倍, 即

$$f_s = (2m + 1) \cdot \frac{f_h}{2} \quad (7-2-13)$$

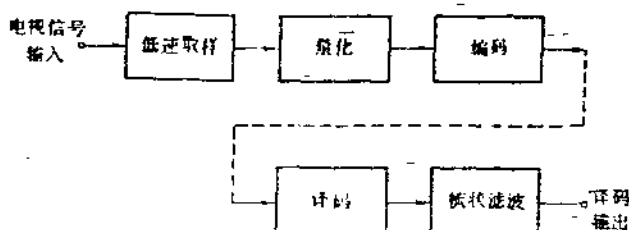


图 7-2-11 亚奈奎斯特取样PCM编码原理方框图

式中 f_s 代表取样频率, f_h 代表行频, m 取正整数。这样, 就可以使频谱折叠部分落入原信号频谱的间隙内。采用这种方案, 大约可节省 30% 的数码率。

如果直接对复合彩色电视信号编码, 由于在频谱中含有接近单纯正弦波的彩色副载波, 它的量化噪声与取样频率关系密切, 如果落入图象频带内会产生辉度和色度的拍频干扰。为了使拍频落入带外, 应将取样频率选为 3 倍或 4 倍的彩色副载波频率, 在这种情况下, 称作常规 PCM 编码的设计。如果选 2 倍的副载波频率为取样频率, 则也属于低速取样范畴。J. P. 罗塞尔曾对 NTSC 制彩色电视信号进行亚奈奎斯特取样 PCM 编码实验。这套系统是在 C R S 技术中心研制的。NTSC 制彩色电视信号的频谱是这样分布的, 亮度信号的频谱能量基本集中在行频 f_h 上, 色度信号的频谱能量集中在半行频的奇次谐波上, 即在 $(n + \frac{1}{2}) f_h$ 上, 而且色度和亮度谱束交错相间。为了采用亚奈奎斯特取样 PCM 编码法, 采取如下取样速率

$$f_s = 2 f_{sc} + \frac{f_h}{4} \quad (7-2-14)$$

或者

$$f_s = 2 f_{sc} - \frac{f_h}{4} \quad (7-2-15)$$

式中 f_{sc} 是 NTSC 彩色副载波, f_h 是行扫描频率, f_s 是取样频率。在译码输出端所使用的梳状滤波器频率在 $f_s - f_v$ 和 f_v 之间, 这样可滤除混叠噪声。 f_v 是 NTSC 制视频信号带宽。这种系统的梳状滤波器设计是很关键的。在这个系统中采用的是横向

梳状滤波器。当取样频率为 $2 f_{sc} + \frac{f_h}{4}$ 时, 基带亮度信号 Y_B 和色度信号 C_B 以及混叠分量 Y_A 和 C_A 的主要谱线如图 7-2-12 所示。显然, 混叠分量 C_A 、 Y_A 是

从 $\frac{f_h}{2}$ 的频率间隔分开的, 因此梳状滤波器在 $\frac{f_h}{2}$ 频率间隔中有最大响应或最小响应。

把视频信号与另一时间序列 TV 行合并起来, 就可构成梳状滤波器。在对 NTSC 彩色

另外,对彩色电视信号进行PCM编码时,可采用分量编码法。因为对R, G, B三个信号进行数字化时,并不需要用相同的精度去量化它们。这三个分量对外加噪声敏感程度并不一样。例如显示亮度恒定的情况下,它们的噪声门限如下:蓝色图象为36 dB,红色象为41 dB,绿色图象为43 dB。对正常图象来讲,三个彩色分量的噪声灵敏度的差别就更大。在蓝色信号中,噪声能见度比红色小10 dB,比绿色小20 dB。为了对全带宽信号量化而不产生明显误差,蓝色信号需4 bit,红色信号需5 bit,绿色信号需6 bit。显然,考虑到这些因素,在分量编码中也有减少数码率的潜力。

7-12

性信源来说, 象素间没有相关性, 可以利用象素灰度值出现概率的不均等性, 采用某种编码方法, 也可以达到压缩数码率的目的。这种根据象素灰度值出现概率的分布特性而进行的压缩编码叫统计编码。

一、信息与信息量

“一幅图象比成百上千个字的含义更丰富”, 这句话可以用信息论的数学方法来加以描述。下面介绍一下信息含量、熵、……等有关信息的一些基本概念。

从 N 个相等可能的事件中选出一个事件来所需要的信息度量、含量或称信息量, 就是在辨识 N 个事件中特定的一个事件的过程中所需要提问“是或否”的最少的次数, 例如要识辨 1 到 32 中选定的某一个数, 可以先提问“是否大于 16?”, 回答出以后就消失了半数的可能事件。这样继续下去只要提问 5 ($= \log_2 32$) 次这类问题, 就能从 32 个数中辨识出所选定的数来。每提问一次, 可以得到 1 位 (比特) 的信息量。因此在 32 个数中辨识一个选定数所需的信息量是

$$\log_2 32 = 5 \text{ (位)}$$

由于 N 个数中选定任一个数 x 的概率 $P(x)$ 设定是相等的即 $\frac{1}{N}$, 因此定义信息量

$I(x)$ 为

$$I(x) = \log_2 N = -\log_2 \frac{1}{N} = -\log_2 P(x) = I[P(x)] \quad (7-3-1)$$

这样定义的信息量具有如下的性质:

$$\left. \begin{aligned} (i) & \text{ 在 } 0 < P < 1 \text{ 时, } I[P(x)] \text{ 是连续的;} \\ (ii) & \text{ 当 } P = 0 \text{ 时, } I[P(x)] = \infty; \\ (iii) & \text{ 当 } P = 1 \text{ 时, } I[P(x)] = 0; \\ (iv) & \text{ 当 } P_1 > P_2 \text{ 时, } I[P_1] < I[P_2]. \end{aligned} \right\} \quad (7-3-2)$$

上面定义的信息量 $I(x)$ 中的 $P(x)$ 是验前概率。假如出现 x_i 的概率是验后概率 $P(x_i | y_j)$, 即在观测到状态 y_j 后, 信源处于 x_i 状态的条件概率。那末可以定义互信息量 $I(x_i, y_j)$ 为

$$I(x_i, y_j) \triangleq \log_2 \frac{P(x_i | y_j)}{P(x_i)} \quad (7-3-3)$$

若观测到 y_j , 可以肯定它是由信源状态 x_i 得出的, 则这时 $P(x_i | y_j) = 1$, 故

$$I(x_i, y_j) = \log_2 \frac{1}{P(x_i)} = -\log_2 P(x_i) = I(x_i) \quad (7-3-4)$$

因此, 一般的信息量 $I(x_i)$ 又称为自信息量。这里定义的信息量有下列一些性质:

(i) 信息量是可加的。若 y_j 及 z_k 表示总共收到的消息的一部分, 则

$$I(x_i, y_j, z_k) = I(x_i, y_j) + I(x_i, z_k | y_j) \quad (7-3-5)$$

这是因为

$$\begin{aligned} I(x_i, y_j, z_k) &= \log_2 \frac{P(x_i | y_j, z_k)}{P(x_i)} \\ &= \log_2 \frac{P(y_j, z_k | x_i) P(x_i)}{P(x_i) P(y_j, z_k)} \\ &= \log_2 \frac{P(z_k | y_j | x_i) P(y_j | x_i)}{P(z_k | y_j) P(y_j)} \\ &= \log_2 \frac{P(y_j | x_i)}{P(y_j)} + \log_2 \frac{P(z_k | y_j | x_i)}{P(z_k | y_j)} \\ &= I(y_j, x_i) + I(z_k | y_j, x_i) \\ &= I(x_i, y_j) + I(x_i, z_k | y_j) \end{aligned}$$

(ii) 互信息量具有对称性质, 即

$$\begin{aligned} I(x_i, y_j) &= \log_2 \frac{P(x_i | y_j)}{P(x_i)} \frac{P(y_j)}{P(y_j)} \\ &= \log_2 \frac{[P(x_i) P(y_j | x_i) / P(y_j)] P(y_j)}{P(x_i) P(y_j)} \\ &= \log_2 \frac{P(y_j | x_i)}{P(y_j)} \quad (7-3-6) \end{aligned}$$

(iii) 互信息量小于或等于自信息量, 即

$$\begin{aligned} I(x_i, y_j) &= \log_2 P(x_i | y_j) - \log_2 P(x_i) \\ &\leq -\log_2 P(x_i) = I(x_i) \quad (7-3-7) \end{aligned}$$

(iv) $-\log_2 P(x_i, y_j) = I(x_i) + I(y_j) - I(x_i, y_j)$, 这是因为

$$\begin{aligned}
-\log_2 P(x_i, y_j) &= -\log_2 [P(x_i)P(y_j|x_i)] \\
&= -\log_2 P(x_i) - \log_2 P(y_j|x_i) \\
&= -\log_2 P(x_i) - \log_2 \frac{P(y_j)P(x_i|y_j)}{P(x_i)} \\
&= -\log_2 P(x_i) - \log_2 P(y_j) \\
&\quad - \log_2 \frac{P(x_i|y_j)}{P(x_i)} \\
&= I(x_i) + I(y_j) - I(x_i, y_j)
\end{aligned}
\tag{7-3-8}$$

二、熵

每个消息或符号的自信息量的统计均值，也即平均信息量被称之为熵。例如一个图象象元有M个灰度等级 x_i ($i=1, 2, \dots, M$)；若每一个灰度等级 x_i 出现的概率为 $P(x_i)$ ，则这个象元可以提供的平均信息量为

$$\begin{aligned}
H(x) &= \sum_{i=1}^M P(x_i) I[P(x_i)] \\
&= -\sum_{i=1}^M P(x_i) \log_2 P(x_i)
\end{aligned}
\tag{7-3-9}$$

当M级灰度的出现概率是均等的，即 $P(x_i) = \frac{1}{M}$ ，则其平均信息量或熵为

$$H(x) = -\sum_{i=1}^M \frac{1}{M} \log_2 \frac{1}{M} = \log_2 M
\tag{7-3-10}$$

可以证明只有在各个 x_i 出现概率相同时，才能得到最大的平均信息量值。例如，若象元的灰度只有2个等级 $x_i = 0, 1$ ，而对应的出现概率为 p 及 $q = (1-p)$ ，则其熵为

$$H(x) = -p \log_2 p - q \log_2 q = \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{(1-p)}
\tag{7-3-11}$$

作出 $H(x)$ 及 p 的关系曲线，可以看出只有在 $p = \frac{1}{2}$ ， $q = 1-p = \frac{1}{2}$ 时，熵值为最大，见图 7-3-1。

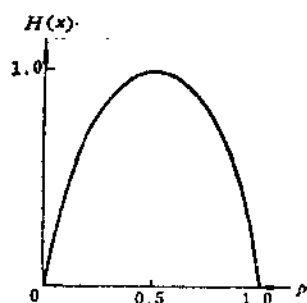
若两个象元 X, Y 分别有灰度等级 $\{x_i\}$ 及 $\{y_j\}$ ，那末可以定义联合熵为

$$H(X, Y) = - \sum_{i=1}^n \sum_{j=1}^m P(x_i, y_j) \log_2 P(x_i, y_j)$$

$$i = 1, 2, \dots, n$$

$$j = 1, 2, \dots, m$$

$$(7-3-12)$$



式中 x_i 有 n 层灰度等级, y_j 有 m 层灰度等级。图 7-3-1 $x_i = 0, 1$ 时熵的曲线

若观测到另一象元 Y 以后, 从观测象元 X 得出的平均信息量可以由条件熵来定义,

即

$$H(H|Y) = - \sum_{i=1}^n \sum_{j=1}^m P(x_i, y_j) \log_2 P(x_i | y_j)$$

$$= - \sum_{i=1}^n \sum_{j=1}^m P(x_i, y_j) [\log_2 P(x_i, y_j) - \log_2 P(y_j)]$$

$$= - \sum_{i=1}^n \sum_{j=1}^m P(x_i, y_j) \log_2 P(x_i, y_j) + \sum_{i=1}^n \sum_{j=1}^m P(x_i, y_j) \log_2 P(y_j)$$

$$= H(X, Y) + \sum_{j=1}^m P(y_j) \log_2 P(y_j)$$

$$(7-3-13)$$

式中 $\sum_{i=1}^n P(x_i, y_j) = P(y_j)$, 因此得条件熵

$$H(X|Y) = H(X, Y) - H(Y) \quad (7-3-14)$$

下面列出熵的一些性质供参考:

(i) 熵是一个非负数, 即 $H(x) \geq 0$; 当 $P(x) = 0$ 或 $P(x) = 1$ 时, $H(x) = 0$

(ii) $H(X) \leq \log_2 M$, M 为 X 的状态数。当 $P(x_i) = \frac{1}{M}$, $i = 1, 2, \dots, M$ 时等式成立。

(iii) 联合熵 $H(X, Y) \leq H(X) + H(Y)$ 。当 X, Y 是统计独立时, 等式成立。