

多媒体数据压缩技术

学习目标：

- (1) 了解：音频数据压缩编码标准和图像、视频压缩编码标准。
- (2) 理解：多媒体数据压缩编码的必要性和分类。
- (3) 掌握：压缩方法的分类、常用数据压缩方法的基本原理。

多媒体数据具有海量性特征,如果不对它们进行压缩,计算机系统难以对它们进行存储和处理。本章将介绍数据压缩方法和压缩标准。

3.1 多媒体数据压缩的概述

3.1.1 多媒体数据压缩的必要性

多媒体信息包括文本、声音、图形、图像、动画和视频等多种媒体信息,经过数字化处理后其数据量非常大,下面列举几个未经过压缩的数字化信息的例子。

(1) 声音

1 分钟立体声音乐采样频率为 44.1kHz,16 位量化位数的数据量为 $44.1 \times 1000 \times 16 \times 2 \times 60 / 8 = 10.09\text{MB}$,那么存储一首 4 分钟的歌曲约需 40MB。

(2) 图像

一张分辨率为 1024×768 ,颜色深度为 24 的图像所占的存储空间是 $1024 \times 768 \times 24 / 8 = 2.25\text{MB}$ 。

(3) 视频

国际无线电咨询委员会(International Consultative Committee for Radio,ICCR)格式、PAL 制式、4:4:4 采样,每帧数据量为 $720 \times 576 \times 3 = 1.19\text{MB}$ 。1 秒钟(25 帧/秒)的视频数据量为 $25 \times 1.19\text{MB} = 29.75\text{MB}$,那么 1 张 650MB 的 CD-ROM 光盘只能存储约 $650 / 29.75 = 21.85$ 秒的视频。

从以上列举的数据可以看出多媒体数据的海量性,这为数据的存储、信息的传输以及计算机的运行速度都带来了巨大的压力,为了达到令人满意的图像、视频画面质量和听觉效果,必须解决音频、图像、视频数据的大容量存储和实时传输问题。因此,使用数据压缩技术减少信息的数据量,将信息以压缩的形式进行存储和传输,既节省存储空间又能提高

传输速度。

3.1.2 多媒体数据压缩的可能性

多媒体数据中的数据量并不完全等于它们所携带的信息量,多媒体信息数字化后存在着大量的冗余信息,冗余是指信息存在的各种性质的多余度。减少数据冗余可以节省存储空间,有效利用网络带宽。数据的冗余主要包括以下内容。

1. 空间冗余

空间冗余是图像中存在的最主要的数据冗余。在同一幅图像中,规则物体和规则背景表面的采样点的颜色往往具有空间连贯性,这些具有空间连贯性的采样点在数字化后表现为空间数据冗余。例如,图像中的一块颜色相同的区域中所有像素点的色彩、光强和饱和度都是相同的,这时数据就存在较大的空间冗余。

2. 时间冗余

时间冗余是音频和视频数据经常包含的冗余。时间冗余反映在视频图像序列中,相邻图像之间有较大的相关性,后一帧的数据与前一帧的数据有许多共同之处,这种共同之处是由于相邻帧记录了相邻时刻的同一场景画面,所以称为时间冗余。

3. 信息熵冗余

信息熵冗余也称编码冗余,如果图像中平均每个像素使用的比特数大于该图像的信息熵,则图像中存在冗余,这种冗余就称为信息熵冗余。

4. 视觉冗余

由于受生理特性的限制,人类的视觉系统对图像场的敏感性是非均匀和非线性的,人眼并不能察觉图像场的所有变化。但在记录原始的图像数据时,通常假定视觉系统是线性和均匀的,对视觉敏感和不敏感的部分同等对待,从而产生了比理想编码更多的数据,这就是视觉冗余。例如,人类视觉系统一般分辨能力约为 2^6 灰度等级,而一般图像的量化采用的是 2^8 灰度等级,即存在着视觉冗余。

5. 听觉冗余

人耳对不同频率的声音的敏感性是不同的,不能察觉所有频率的变化,对某些频率不必特别关注,因此存在听觉冗余。

6. 结构冗余

结构冗余是指图像在部分结构上的类似性所产生的冗余,如图像中存在很强的纹理结构或自相似性,称为结构冗余。当一幅图像有很强的结构特性,纹理和影像色调等与物体表面结构有一定的规则时,其结构冗余很大。例如,方格状的地板图案等。

7. 知识冗余

知识冗余指对图像的理解与某些先验知识有相当大的相关性。例如,人脸的图像有固定的结构,嘴的上方有鼻子,鼻子上方有眼睛,鼻子位于正脸图像的中线上;再比如,当接到一个成语的前3个字“大惊小”时,立刻就会知道下一个字肯定是“怪”。这种规律性的结构可由先验知识和基础知识得到,这类冗余称为知识冗余。在图像和声音中都存在

这种冗余。

随着对人类视觉系统的进一步研究,人们可能会发现更多的冗余性,使得数据压缩编码的可能性越来越大,进而推动数据压缩技术的发展。

3.1.3 压缩方法分类

由于大量的多媒体数据存在数据冗余,因此可以采用某种算法对数据进行重新整理,以减小数据量而不损失其中的信息,达到数据压缩的目的,这些算法就是压缩方法,也叫编码方法。

多媒体数据冗余类型不同,就有不同的压缩方法。根据解码后数据质量有无损失,压缩方法可分为有损压缩和无损压缩两大类。在这个基础上,根据编码原理进行分类,可分为预测编码、变换编码、统计编码、PCM 编码和混合编码,如图 3-1 所示。

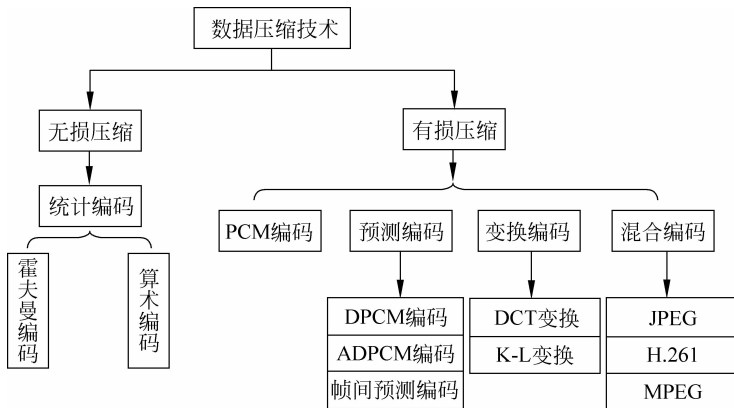


图 3-1 数据压缩方法分类

有损压缩方法压缩了熵,会减少信息量,因为熵定义为平均信息量,而损失的信息是不能再恢复的,因此这种压缩方法是不可逆的。有损压缩方法允许一定程度的失真,所以可用于对图像、声音和视频等数据的压缩。如采用混合编码的 JPEG 标准,它对自然景物的灰度图像,一般压缩比可达到几倍至十几倍,而对自然景物的彩色图像,压缩比可达到几十倍甚至上百倍。采用自适应差分脉冲编码调制(ADPCM)编码的声音数据,压缩比一般能达到 $(4:1) \sim (8:1)$ 。压缩比最大的是视频数据,采用混合编码的 DVI 多媒体系统,压缩比通常能达到 $(50:1) \sim (100:1)$ 。

无损压缩方法去掉或减少了数据中的冗余,但这些冗余值是可以重新插入到数据中的,因此它是可逆的过程。无损压缩方法由于不会产生失真,在多媒体技术中一般用于文本数据的压缩,它能保证百分之百地恢复原始数据。

3.1.4 压缩方法的衡量指标

衡量数据压缩方法的性能好坏有 3 个重要的指标:压缩比、实现压缩的算法、恢复效果。

第一个指标是压缩比,就是压缩过程中输入数据量和输出数据量之比,压缩比要大,即压缩前后所需要的信息存储量之比要大。

第二个指标是实现压缩的算法,压缩、解压缩速度要快,尽可能做到实时压缩和解压缩。尤其是对于动态视频的压缩和解压缩,速度是至关重要的。

第三个指标是恢复效果,这与压缩的类型有关。要考虑数据的恢复效果,恢复效果要好,要尽可能地恢复原始数据。

3.2 音频压缩编码的方法

音频信号的压缩方法有多种,无损压缩法包括不引入任何数据失真的各种熵编码;有损压缩法又可分为波形编码、参数编码和混合编码。

1. 波形编码

波形编码利用采样和量化过程来表示音频信号的波形,将时间域信号直接变换为数字代码,力图使重建语音波形保持原语音信号的波形形状。波形编码的基本原理是在时间轴上对模拟语音按一定的速率采样,然后将幅度样本分层量化,并用代码表示。解码是其逆过程,将收到的数字序列经过译码和滤波恢复成模拟信号。脉冲编码调制(PCM)是最简单、最基本的编码方法,它直接赋予采样点一个代码,没有进行压缩,所需的存储空间较大。为了减少存储空间,利用音频采样的幅度分布规律和相邻样本值具有相关性的特点,提出了差分脉冲编码调制(DPCM)和自适应差分脉冲编码调制(ADPCM)等算法,实现了数据的压缩。波形编码的特点是适应性强、音频质量好、高码率,适合高保真语音和音乐信号,但其压缩比不大。

2. 参数编码

参数编码又称为声源编码,是将音频信号表示成某种模型的输出,利用特征提取的方法抽取必要的模型参数和激励信号的信息,并对这些信息编码进行传输。解码为其逆过程,将收到的数字序列经变换恢复特征参量,再根据特征参量重建音频信号。参数编码是通过对音频信号特征参数的提取和编码,力图使重建音频信号保持原始音频的特性。常用的音频参数有线性预测系数、滤波器组、共振峰等。参数编码的特点是压缩比大、计算量也大、还原信号的质量较差、自然度低。

3. 混合编码

混合编码使用参数编码技术和波形编码技术,结合了这两种方法的优点,力图保持波形编码的高质量 and 参数编码的低速率。例如,多脉冲激励线性预测编码(MPLPC)、规划脉冲激励线性预测编码(KPELPC)和码本激励线性预测编码(CELP)等都属于混合编码技术。

3.3 图像和视频数据压缩方法

图像和视频数据压缩方法分成两种类型:有损压缩和无损压缩。典型的无损压缩编码有霍夫曼编码和算术编码,典型的有损压缩编码有预测编码和变换编码等。

3.3.1 预测编码

1. 预测编码的基本原理

预测编码是根据某一种模型,利用以前的(已收到)一个或几个样本值,对当前的(正在接收的)样本值进行预测,将样本实际值和预测值之差(差值)进行编码,这种编码方法称为预测编码方法。预测越准确,误差值就越小,那编码所需的位数就可以减少,以达到压缩的目的。

预测编码方法的优点是算法简单、速度快、易于硬件实现;缺点是编码的压缩比不高、误码易于扩散、抗干扰能力差。

2. 预测编码的分类

预测编码方法分为线性预测编码和非线性预测编码两种,其中线性预测编码又称为差分脉冲编码调制(DPCM)。预测编码里介绍以下3种编码。

(1) 差分脉冲编码调制

模拟量经A/D变换,得到二进制码的过程,这就是著名的脉冲编码调制(PCM)。差分脉冲编码调制不是对每一样本值都进行量化,而是预测下一个样本值,并量化实际值和预测值之间的差,达到压缩的目的。解压时,仍然使用同样的预测器,并将这一预测值和存储的已量化的差值相加,产生出近似的原始信号,基本恢复原始数据。在DPCM中,特殊的“1bit 量化”情况称为 Δ 调制。

DPCM的关键点在于预测器和量化器的设计。一般情况下,一个好的预测器可以使许多实际值和预测值之间的差值很小,或为零。因此,误差信号量化器所需的量化间隔通常比原信号所需的量化间隔少,可以用较少的比特来表示量化的误差信号,得到数据压缩。

(2) 自适应差分脉冲编码调制 ADPCM

DPCM是预测器和量化器设计好后,整幅图都用它,不再变化。为了进一步改善量化性能或压缩数据率,可以采用自适应的方法,这种方法称为自适应差分脉冲编码调制ADPCM。

ADPCM具有自适应特性,该编码包括自适应量化和自适应预测两种形式,主要适用于对中等质量的音频信号进行高效率压缩,如对语音信号的压缩、调幅广播音质信号的压缩等。

① 自适应量化。自适应量化是在一定的量化级数下,减少量化误差或在相同误差情况下压缩数据。当信号分布不均匀时,能随输入信号的变化改变量化区间的大小,自适应量化必须对输入信号的幅值进行估计,有了估计才能确定相应的改变量。若估计在信号的输入端进行,称前馈自适应;若估计在量化输出端进行,称反馈自适应。对于能量分布较大的系数分配较多的比特数,采用较小的量化步长;反之分配较少的比特数,采用较大的量化步长,从而达到压缩的目的。

② 自适应预测。自适应预测是根据常见的信息源求得多组固定的预测参数,将预测参数提供给编码使用。在实际编码时,根据信息源的特性,以实际值与预测值的均方误差最小为原则,随着编码区间的不同,预测参数自适应地变化,以达到准最佳预测。

(3) 帧间预测编码

帧间预测编码是利用视频图像各帧之间的相关性,即时间相关性,减少帧内图像信号的冗余,采用预测编码的方法消除序列图像在时间上的相关性,即不直接传送当前帧的像素值,而是传送 x 和其前一帧或后一帧对应像素 x 之间的差值,来达到图像压缩的目的,其广泛用于普通电视、会议电视、视频电话、高清晰度电视的压缩编码。

帧间预测编码分为运动补偿的帧间预测和帧间内插法两种类型。

运动补偿的帧间预测包括以下几个步骤:首先,将图像分解成相对静止的背景和若干运动的物体,各个物体可能有不同的位移,但构成每个物体的所有像素的位移相同,通过运动估值得到每个物体的位移矢量;然后,利用位移矢量计算经运动补偿后的预测值;最后对预测误差进行量化、编码、传输,同时将位移矢量和图像分解方式等信息送到接收端。

帧间内插法,活动图像的帧间内插编码是在系统发送端每隔一段时间丢弃一帧或几帧图像,在接收端再利用图像的帧间相关性将丢弃的帧通过内插恢复出来,以防止帧率下降引起闪烁和动作不连续。

3.3.2 变换编码

1. 变换编码的基本原理

变换编码不是直接对空域图像信号编码,而是首先将空域图像信号映射变换到另一个正交矢量空间(变换域或频域),获得一批变换系数,然后对这些变换系数进行编码处理。该变换过程是逆过程,使用逆变换可以恢复原始数据。

例如,有两个相邻的数据样本 x_1 和 x_2 ,每个样本采用 3b 编码,因此各自都有 8 个幅度等级,两个样本的联合事件共有 64 种可能性,用 64 个点表示。横坐标表示 x_1 的 8 种等级,纵坐标表示 x_2 的 8 种等级,考虑到样本值的相关性, x_1 和 x_2 同时出现相近幅度等级的可能性最大,即很可能出现在 $x_1 = x_2$ 直线附近,如图 3-2(a)所示。如果对该数据进行正交变换,将坐标系逆时针旋转 45° ,如图 3-2(b)所示,在新坐标系中 y_1 对应到 $x_1 = x_2$ 这条直线,那么变换后的数据样本集中在 y_1 轴上,对这部分数据进行量化、编码和传输,其他数据不做处理,这样就达到了压缩数据的目的。

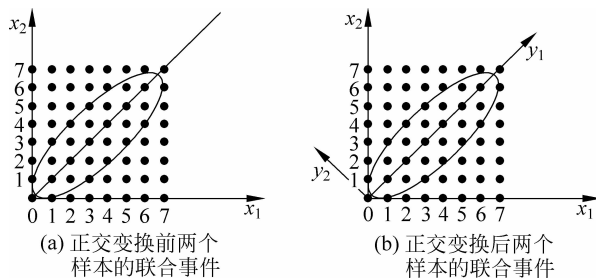


图 3-2 正交变换前后两个样本的联合事件

变换编码主要在变换域上进行,而预测编码主要在时(空)域上进行。

变换编码系统中压缩数据有 3 个步骤,即变换、变换域采样和量化。由于变换是可逆

的,变换本身不进行数据压缩,它只把信号映射到另一个域,使信号在变换域里容易进行压缩,变换后的样本值更独立和有序。为了取得满意的结果,某些重要系数的编码位数比其他的要多,某些系数就被忽略了。在对量化后的变换样本值进行比特分配时,要考虑使整个量化失真最小。

2. 变换编码的分类

变换编码技术比较成熟,理论也比较完善,已广泛用于各种图像数据压缩。变换编码的种类也很多,常用的变换编码方法有傅里叶变换、K-L 变换、哈尔变换、离散余弦变换等。下面介绍 K-L 变换和离散余弦变换。

(1) 最佳的正交变换——K-L 变换

数据压缩主要是去除信源的相关性。设信源序列为一个 n 行 n 列的矩阵 $\mathbf{X}$,对一幅图像进行扫描,从上到下共 n 行,每行从左到右取 n 个值。若考虑信号存在于无限区间上,而变换区域又是有限的,那么表征相关性的统计特性就是协方差矩阵,协方差矩阵主对角线上各元素就是变量的方差,其余元素就是变量的协方差,且为一对称矩阵。

当协方差矩阵中除主对角线上元素之外的各元素都为零时,就等效于相关性为零。所以,为了有效地进行数据压缩,常常希望变换后的协方差矩阵为一对角矩阵,同时也希望主对角线上各元素随第 i 行第 j 个量值的增加很快衰减。因此,变换编码的关键在于:在已知输入信号矩阵 $\mathbf{X}$ 的条件下,根据它的协方差矩阵去寻找一种正交变换 T ,使变换后的协方差矩阵满足或接近为一对角矩阵。

当经过正交变换后的协方差矩阵为一对角矩阵,且具有最小均方误差时,该变换称最佳变换,也称 K-L 变换。K-L 变换虽然具有均方误差意义上的最佳性能,但需要预先知道信源的协方差矩阵并求出特征值。求特征值与特征向量并不是一件容易的事,维数较高时甚至求不出来,即使能借助计算机求解,也很难满足实时处理的要求,而且从编码应用来看还需要将这些信息传输给解码端,这些原因导致 K-L 变换在工程实践中不能广泛应用。人们一方面继续寻求特征值与特征向量的快速算法,另一方面则寻找虽不是“最佳”但较容易实现的变换方法,而 K-L 变换成为评价这些变换性能的标准。

(2) 离散余弦变换——DCT 变换

余弦变换是傅里叶变换的一种特殊情况。在傅里叶级数展开式中,如果被展开的函数是实偶函数,那么,其傅里叶级数中包含余弦项,再将其离散化,由此可导出余弦变换,或称为离散余弦变换(DCT)。

DCT 被认为是性能接近 K-L 变换的准最佳变换,因为 DCT 与 K-L 变换压缩性能和误差很接近,而 DCT 计算复杂度适中,产生的系数容易被量化,又具备可分离性,算法快速,而且 DCT 算法是对称的,利用逆 DCT 算法可以用来解压缩。由于 DCT 的种种优点,在图像数据压缩中,常使用离散余弦变换编码。

3.3.3 统计编码

香农定理的要点是:信源中含有自然冗余度,这些冗余度既来自信源本身的相关性,又来自信源概率分布的不均匀性,只要找到去除相关性或改变概率分布不均匀性的手段和方法,也就找到了信息熵编码的方法。但信源所含的平均信息量是进行无失真编码的

理论的极限,只要不低于此极限,就能找到某种适宜的编码方法,去逼近信息熵,实现数据压缩。

1. 统计编码原理——信息量和信息熵

在日常生活中,如听到一则新闻,看到一篇文章,从信息论的观点看,称为“消息”。在消息中,有些内容是人们事先不知道的,这些不确定的内容称为信息。一个消息的可能性越小,其信息越多;消息的可能性越大,则信息越少。

香农定理应用概率来描述不确定性。事件出现的概率越小,不确定性越多,信息量越大,反之则少。在数学上,所传输的消息是其出现概率的单调下降函数。信息量是指从 N 个相等可能事件中选出一个事件所需要的信息度量或含量,也就是在辨识 N 个事件中特定的一个事件的过程中所需要提问“是”或“否”的最少次数。例如,要从 32 个数中选定某一个数,可以先提问“是否大于 16”,不论回答是或否都消去了半数的可能事件,这样继续问下去,只要提问 5 次这类问题,就能从 32 个数中选定某一个数。这是因为每提问一次都会得到 1 比特的信息量。在 32 个数中选定某一个数所需要的信息量是 $\log_2 32 = 5(\text{b})$ 。设从 N 个数中选定任一个数 x 的概率为 $P(x)$,假定选定任意一个数的概率都相等,即 $P(x) = \frac{1}{N}$,因此信息量为

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

熵,来源于 20 世纪 40 年代由 Claude Shannon 创立的信息论中的一条定理,这一定理借用了热力学中的名词“熵”(Entropy)来表示一条信息中真正需要编码的信息量。如果将信息源所有可能事件的信息量进行平均,即可得到信息的熵。一个事件发生的概率越小,其信息熵越高,所含的信息量越大。

图像编码中,信源 X 的熵为

$$H(X) = E\{I(x_j)\} = -\sum_{j=1}^n P(x_j) \cdot \log_2 P(x_j) \quad (j = 1, 2, 3, \dots, n)$$

其中,等概率事件的熵最大,为

$$H(X) = -\sum_{j=1}^n \frac{1}{N} \log_2 \frac{1}{N} = \log_2 N$$

当 $P(x_1) = 1$ 时, $P(x_2) = P(x_3) = \dots = P(x_j) = 0$, 这时熵:

$$H(X) = -P(x_1) \log_2 P(x_1) = 0$$

由上可得熵的范围为

$$0 \leq H(X) \leq \log_2 N$$

在编码中用熵值来衡量是否为最佳编码。若以 L_c 表示编码器输出码字的平均码长,其计算公式为

$$L_c = \sum_{j=1}^n P(x_j) L(x_j) \quad (j = 1, 2, 3, \dots, n)$$

其中, $P(x_j)$ 是信源 X 发出 x_j 的概率, $L(x_j)$ 为 x_j 的编码长。

当 $L_c > H(X)$ 有冗余,不是最佳;

$L_c < H(X)$ 不可能;

$L_c = H(X)$ 最佳编码 (L_c 稍大于 $H(X)$)

熵值为平均码长 L_c 的下限。

2. 统计编码的类型

统计编码的基本思想是: 根据信息出现概率的分布特征而进行压缩编码, 寻找概率与码字长度间的最优匹配。常见的统计编码有霍夫曼编码和算术编码。

(1) 霍夫曼(Huffman)编码

霍夫曼编码是霍夫曼在 1952 年提出来的一种从下到上的编码方法, 利用变字长最佳编码, 实现信源符号按概率大小顺序排列。现已广泛应用于 JPEG 和 MPEG 等各种信息编码标准中。它的算法步骤如下。

- ① 按照符号出现的概率大小进行排序。
- ② 把最小的两个概率值相加, 得到一个新的概率序列。
- ③ 重复上述两个步骤, 直到概率值为 1。
- ④ 从后往前进行编码, 概率大的赋予 1, 概率小的赋予 0 (反过来也可以)。
- ⑤ 写出每个符号的码字。

下面以一个具体例子, 说明霍夫曼编码过程。

例如, 已知信源:

$$X = \left\{ \begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ 0.35 & 0.25 & 0.20 & 0.10 & 0.05 & 0.05 \end{array} \right\}$$

对其进行 Huffman 编码, 并计算平均码长。

Huffman 编码步骤如下。

信源符号	概率	编码过程	编码	码长
x_1	0.35	1	11	2
x_2	0.25	0	10	2
x_3	0.20	1	01	2
x_4	0.10	1	001	3
x_5	0.05	1	0001	4
x_6	0.05	0	0000	4

码字的平均码长 L_c 以下的公式计算:

$$L_c = \sum_{j=1}^n P(x_j) L(x_j)$$

$$= (0.35 + 0.25 + 0.20) \times 2 + 0.10 \times 3 + (0.05 + 0.05) \times 4 = 2.3(\text{b/pel})$$

通过例子, 归纳霍夫曼编码的特点: 霍夫曼编码构造出的码不唯一; 霍夫曼编码字长参差不齐; 霍夫曼编码对不同信源的编码效率是不同的; 对信源进行霍夫曼编码后, 形成一个霍夫曼表。

(2) 算术编码

算术编码比霍夫曼编码复杂,但是它不需要传送像霍夫曼编码那样的霍夫曼码表,而且算术编码还有自适应能力,因此算术编码是实现高效压缩数据中很有前途的编码方法。

算术编码的原理是,将被编码信源表示为 $[0,1)$ 区间的一个实数,根据各符号出现的概率构造其所在区间,随着信息字符的不断出现,其所在区间越来越小,对应表示的实数也越来越小,那么表示这一消息所需的二进制位数就越多。信源中连续符号根据某一模式生成概率的大小来缩小间隔,可能出现的符号要比不太可能出现的符号缩小范围少,只增加了较少的比特。

对二进制编码来说,信源符号只有两个。因此在算术编码初始阶段可预置一个大概率 P_e 和小概率 Q_e ,然后对被编码比特流符号进行判断。设编码初始化子区间为 $[0,1)$, Q_e 从0算起,则 $P_e=1-Q_e$,随着被编码数据流符号的输入,子区间逐渐缩小。

区间计算方法如下。

新子区间左端=前子区间左端+当前子区间左端 $\times$ 前子区间长度

新子区间长度=前子区间长度 $\times$ 当前子区间的长度

最后得到的子区间长度决定了表示该区域内的某一个数所需的位数。

解码过程是逆过程,首先将区间 $[0,1)$ 按 Q_e 靠近零侧、 P_e 靠近1侧分割成两个子区间,判断被解码的码字落在哪个子区间,然后赋予对应符号。

算术编码的特点:算术编码的模式选择直接影响编码效率;算术编码的自适应模式无须先定义概率模型,对无法进行概率统计的信源来说比较合适;在信源符号概率接近时,算术编码比霍夫曼编码效率高;算术编码的硬件实现比霍夫曼编码要复杂;算术编码在JPEG的扩展系统中被推荐代替霍夫曼编码。

算术编码和霍夫曼编码相比,有如下几个异同点。

- ① 算术编码的编码效率更高。
- ② 它们都是对错误很敏感的编码方法,如果有一位发生错误就会导致整个消息译错。
- ③ 它们的信源概率都是固定的,而且要事先统计确定。

3.4 多媒体数据的压缩标准

3.4.1 音频数据的压缩标准

关于音频数据的压缩标准问题,国际电话电报咨询委员会 CCITT、国际电信联盟 ITU-T 和国际标准化组织 ISO 先后提出了一系列的音频压缩标准的建议,针对不同的质量要求,制定了不同的标准。

1. G.711

本建议公布于1972年,它给出话音信号编码的推荐特性。话音的采样率为8kHz,允许偏差是 $\pm 50 \times 10^{-6}$,每个样本值采用8位二进制编码,对应的比特流速率为64Kbps,使用A律和 μ 律非线性量化技术。本建议中分别给出了A律和 μ 律的定义,它是将13位的PCM按A律、14位的PCM按 μ 律转换为8位编码。该建议还规定,在物理介质上连