

随着技术的进步,目前电信业务已从电话、数据、传真发展到可提供视觉信息的可视电话、会议电视以及以数字音视频为主要内容的各种类型的多媒体通信。现代图像和多媒体通信业务的发展需要大量地存储、记录和传输各类静止图像和活动图像。它们要求图像质量高,设备稳定可靠,成本低廉,并且能够适应现代网络技术的迅速发展,特别是三大网络——电信网、互联网、广播电视网逐步融合,最终走向三网合一的发展趋势以及移动通信的要求。解决这一问题的必经之路是进行图像和视频信号的数字化。同样,基于计算机技术的多媒体制作也首先要实现数字化才可能进行。

如今,数字可视电话、会议电视已经普遍使用。DVD、数字照相机、数字摄像机已成为热销的消费类电子产品。能够接收数字图像信息的手机已成为流行时尚。以数字音视频为主要内容的多媒体通信在远程教育、远程医疗、图像监视、监控和监测方面也得到广泛应用。标准清晰度电视将很快完成从模拟向数字的过渡,而高清晰度电视广播则从一开始就以数字形式出现。数字化时代已经到来!

图像和视频信号数字化具备许多在模拟信号所不具备的优点,诸如:数字传输质量显著高于模拟传输质量,可经多次中继而不引起噪声严重积累,并且基本不受传输系统非线性影响;易于采用信道编码技术提高传输的可靠性;便于利用时分复用技术与其他电信业务相结合,纳入多种信息业务兼容的数字综合业务网;数字形式易于加密,提高信息安全;数字信号易于借助计算机技术进行处理、存储、分配、管理;数字电路容易大规模集成;数字设备可靠性高、维护简便等等。

但是,任何事物都有两面性,数字化在提供诸多优点的同时也带来一些问题,或者说提出了新的挑战。其中最主要的问题是信号数字化后的数据量太大,数码率太高,用模拟信号的术语来说,就是频带太宽。直接传输对信道利用很不经济,甚至是现有信道难以容纳的;存储则需要巨大的存储容量。同时,过高的数码率也造成多媒体计算机处理方面的困难。

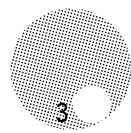
如何妥善解决图像和视频信号数字化后的数据压缩问题,在保证图像质量的前提下,用最少量的数码或最低的数码率实现各类数字图像和视频信息的存储、记录和传输,达到优质、经济、可靠的要求,这就是图像压缩编码这门学科研究的内容。

## 1.2 图像数据压缩

### 1.2.1 图像数据压缩的必要性

以一幅中等分辨率的真彩色图像(分辨率为  $640 \times 480$ , 24 比特/像素)为例,其比特数为  $640 \times 480 \times 24 = 7.37 \text{ Mb}$  (兆比特),需占约 0.9 MB (兆字节)的存储空间。一个 1 GB (1000 MB) 容量的硬盘只能存储这种图像 1000 多幅。而若以每秒 25 帧的速度显示,其传输速率要求达到  $184 \text{ Mb/s}$  (兆比特/秒)。如果作为多媒体信息由 PC 机处理,则无论是总线传输速率、数据存取和交换速率还是网络传输速率,目前都无法达到如此高的要求。图像必须经过数据压缩处理,计算机才有可能具备处理这种信息的能力。

上述问题对数字视频尤为突出。以传输一路带宽为 6 MHz 的黑白电视信号为例,若



采用普通的脉冲编码调制(PCM)方式传输,按照奈奎斯特取样定理的要求,取样频率不应低于  $2 \times 6\text{MHz} = 12\text{MHz}$ 。而主观实验表明,对于图像信号,从模拟形式经过模数(A/D)转换为 PCM 数字形式,这一过程引入的量化误差(失真)只有当量化精度达到或超过 8 比特/取样时,才不会被察觉出来。因此,传输一路 PCM 的黑白电视信号的数码率至少要  $96\text{Mb/s}$ 。从理论上讲,二进制传输信道每  $1\text{Hz}$  带宽能传输的最高数码率是  $2\text{b/s}$ ,这相当于要求信道提供  $48\text{MHz}$  的带宽,是模拟传输的 8 倍。所以,从信息论的观点来看,PCM 传输方式是以带宽为代价换取高的传输质量。

又如,按照 CCIR (国际无线电咨询委员会)1982 年制定的演播室数字电视编码参数,对于标准清晰度彩色电视的亮度分量  $Y$  和色度分量  $R-Y$ 、 $B-Y$  的取样频率分别为  $13.5\text{MHz}$ 、 $6.75\text{MHz}$ 、 $6.75\text{MHz}$ ,每个取样均匀量化为 8 比特,所以,在节目制作过程中,一路标准清晰度彩色电视节目的数码率为

$$(13.5 + 6.75 + 6.75) \times 8 = 216 \text{ Mb/s}$$

相对于演播室节目制作,用于节目传输的数码率可以适当降低一些。按照 CCIR601 建议(1986)的标准清晰度数字电视的传输标准,每帧像素数为  $720 \times 576$ ;帧频为  $25\text{Hz}$ ,隔行扫描;色信号取样格式为  $4:2:0$ (色度分量  $R-Y$ 、 $B-Y$  的取样频率在水平和垂直两个方向均是亮度分量  $Y$  取样频率的  $1/2$ ,且  $R-Y$ 、 $B-Y$  两个色度分量隔行轮送,每行只送这两个分量中的一个);量化为 8 比特/取样。此时的数码率为

$$(720 \times 576) \times 25 \times (1 + 0.25 + 0.25) \times 8 = 124.416 \text{ Mb/s}$$

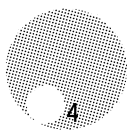
一个容量为  $640\text{MB}$  的 CD-ROM 光盘只能存放约  $40\text{s}$  的这种标准清晰度数字电视节目。而一路高清晰度电视的数码率高达  $1\text{Gb/s}$  以上。如不压缩,DVD 播放机使用的 DVD-ROM 光盘虽有  $4.7\text{GB}$  的容量,也只能存储不到  $40\text{s}$  的 HDTV 节目。数字 HDTV 传输系统要求将一路 HDTV 图像信号的数码率压缩至  $20\text{Mb/s}$  左右,也就是说,要求压缩到原来的  $1/50$ 。

从以上的实例中可以看出,对数字图像和数字视频实行数据压缩势在必行。为了提高多媒体信息传输和存储的效率,降低对信道容量和存储容量的要求,一般不直接采用简单的 PCM 方式传输和存储数字信息,而是对数字化后的信源信号先进行数据压缩,然后再以压缩的形式传输或存储。数据压缩是通过寻求一种有效的信号编码方式来实现的,这种编码称为信源编码,又称高效编码。图像编码是对图像信源进行信源编码,是在保证达到所要求的图像质量前提下,设法降低所必需的数码率而采取的压缩编码技术。通过图像编码达到节省传输带宽或节省所需存储量的目的,同时也为多媒体计算机处理提供可能。

## 1.2.2 图像数据压缩的可能性

图像信号可以压缩的根据有两个方面,一方面是图像信号中存在大量冗余度可供压缩,并且这种冗余度在解码后还可无失真地恢复;另一方面是可以利用人的视觉特性,在不被主观视觉察觉的容限内,通过减少表示信号的精度,以一定的客观失真换取数据压缩。

图像信号的冗余度存在于结构和统计两方面。图像信号结构上的冗余度表现为很强



的空间(帧内的)和时间(帧间的)相关性。统计测量证实,电视信号在相邻像素间、相邻行间、相邻帧间存在着这种强相关性。一般情况下,电视画面中的大部分区域信号变化缓慢,尤其是背景部分几乎不变,正如观看电影胶片,可以发现连续几十张画面变化甚小。电视信号占用 6MHz 带宽,是为了表示画面中突变的轮廓、占画面比例不一定很大的纹理细节以及快速的运动,但在大部分时间内,电视信号并不含这么高的空间和时间频率成分。因此,采取适当的信号处理技术,解除图像和视频信号中由于高度相关性带来的冗余,频带是可以压缩的。

信号统计上的冗余度来源于被编码信号概率密度分布的不均匀。例如,在下面将要讲到的预测编码系统中,需要编码传输的是预测误差信号  $e$ ,它是当前待传像素值  $S$  与它的一个预测值  $\hat{S}$  间的差分信号。预测值是通过在该像素之前已经传出的它的几个近邻像素值预测出来的。由于电视信号相邻像素间的强相关性,这种预测在大部分情况下是很准的,即预测误差很小。因此,预测误差高度集中分布在 0 附近,形成如图 1-1 所示的概率分布  $p(e)$ ,  $p(e)$  近似可用均值为 0 的拉普拉斯分布表示。这种极不均匀的概率分布对采用变字长编码压缩码率极为有利,因为在编码时可以对出现概率高的预测误差(0 及小误差)用短码,对出现概率低的大的预测误差用长码,则总的平均码长要比用固定码长编码短很多。这种编码叫统计编码或概率匹配编码、熵编码,将在本书的第 4 章专门讨论。

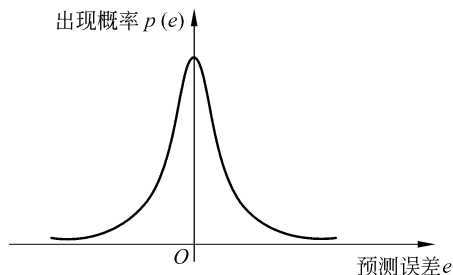


图 1-1 预测误差的概率分布

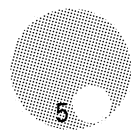
充分利用人的视觉特点,挖掘潜力,是实现数码率压缩的又一重要途径。人眼对图像的细节(空间)分辨率、运动(时间)分辨率和灰度(对比度)分辨率的要求都有一定的限度。对于图像信号在空间、时间以及在幅度方面进行数字化的精细程度只要达到这个限度即可,超出是无意义的。并且视觉心理学和视觉生理学的研究表明,人眼对图像细节、运动和灰度三方面的分辨能力是互相制约的。在人观察景物时,并非对这三方面同时都具备最高的分辨能力。当人眼对图像的某种分辨率要求很高时,对其他方面的分辨率则降低要求。研究表明,人眼对于以上三方面分辨率的要求有如下的制约关系。

### 1. 细节分辨率与运动分辨率

人眼只在观看静止画面时,才需要高达 40 万个以上像素(对于标准清晰度电视而言)的细节分辨率。当画面中的物体出现运动时,所需要的像素数目可显著下降。另外,只有对快速运动的物体,人眼才需要最高的运动分辨率(30Hz 左右的帧频),保持活动画面的连贯性,而对于缓慢运动和静止的场景则帧频可以适当降低:

### 2. 细节分辨率和灰度分辨率

仅在观察图像中的大块面积时,人眼才能分辨出全部 256 个灰度等级。而当观察图像局部的小块面积或精致细节时,并不需要这么多的灰度等级。在灰度突变处,由于视觉



的掩盖效应,人眼不能觉察很细致的灰度差别。

### 3. 灰度分辨率和运动分辨率

当传送一个快速运动图像时,只需采用较少的灰度等级,因为当人眼注意观看运动物体时,对灰度的分辨能力下降。

利用上述视觉特性,在进行图像数字化时可以采用自适应技术,即根据图像的每一局部的特点来决定对它的取样频率和量化的精度,尽量做到与人眼在观看这种图像局部时相关的视觉特点相适配。这种自适应技术不是固定不变地总以最高的精度对图像取样和量化,而是把三种分辨率随时调整到“够用,又不浪费”的程度,因此,可以在不损伤图像主观质量的条件下压缩数码率。在预测编码中,利用受图像局部活动性影响的视觉掩盖效应设计的自适应主观优化量化器;在变换编码中,根据主观视觉特点,对代表不同空间频率的变换系数进行量化时,量化步长采用不同的视觉加权等,都是这方面的典型例子。

## 1.3 图像压缩信源编码过程

图像和视频信源编码的整个过程一般由以下三个步骤完成。

(1) 对表示信号的形式进行某种映射,即变换描写信号的方式。通过这种映射解除或削弱存在于图像信号内部的相关性,降低其结构上存在的冗余度。例如,在预测编码中,取代原始的像素值,用预测误差表示信号;在变换编码中,用一组变换系数表示一组像素值。

(2) 在满足对图像质量一定要求的前提下,减少表示信号的精度。这通过采用符合主观视觉特性的量化来实现。

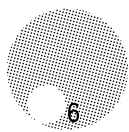
(3) 利用统计编码(例如,赫夫曼编码、算术编码等)消除最终被编码的符号(例如,预测编码的预测误差、变换编码的变换系数、视频编码的运动矢量等)所含的统计冗余度。

图 1-2 所示为包含以上三个步骤的信源编码过程方框图。其中信号映射和统计编码这两个环节是可逆的过程,而量化是不可逆过程。当不加入量化时,通过解码端的反映射和统计解码可以无失真地恢复原始信号;而加入量化后,整个编解码过程造成的失真完全由量化造成。



图 1-2 信源编码过程

PCM 是由模拟形式信号转变为数字形式信号所必需经过的取样、量化过程。本书后面讨论的各种图像压缩编码方法,都假定进入编码器的输入信号是已经经过 A/D 转换的 PCM 信号。PCM 对模拟信号在时间、空间和幅值上的离散处理会不可避免地引入误差。为了保证图像和视频信号从模拟形式转变成 PCM 形式不产生主观上可以察觉的失真,即 PCM 信号再经 D/A 转换,重新回到模拟形式后与原始的信号比较不能被察觉出质量上的差别,必须保证对信号有足够高的取样频率和量化精度。如前所述,取样频率应符合



取样定理的规定,不应低于信号中所含最高频率成分的两倍,以避免频谱混叠失真;而对图像和视频信号的每一分量(如  $R, G, B$  或  $Y, U, V$ )采用均匀量化的量化精度不应低于 8 比特/像素。图 1-2 中,信源编码过程中的量化实际上是对信号进行的第 2 次量化。

为表示图像编码系统的压缩能力,常用的一个术语是压缩比。压缩比  $C$  定义为压缩编码前后平均每像素所需的比特数之比,或者,对于数字视频信号为压缩编码前后数码率之比。编码每像素的比特数用比特/像素(bit per pixel, bpp)表示。

## 1.4 信源编码与信道编码

数字信号在信道中传输时,由于干扰会引起传输误码,特别是当信号表示越有效,即通过信源编码使冗余度压缩越多时,误码影响越严重。就如人们能够容易地发现书刊中的印刷错误,并很可能根据上下文的意思把它纠正过来;但是,如果一封电报,即使仅出现一个错字,一般也很难猜出原意,因为电报的内容经过高度压缩。在图像和视频编码中,如预测编码,其在信道中传输的是预测误差,预测误差在接收端与借助先前接收到的邻近像素值估计出来的预测值相加得到重建像素值,而这个重建像素值又被用来参与估计后续传来像素的预测值。因此,一旦某个像素的预测误差在传输过程中出现差错,则不仅该像素本身不能正确重建,而且这个差错会在其后到达的像素中大面积扩散。如果视频编码的码流在传输过程中出现误码,则差错不仅在一帧画面中扩散,而且很可能在帧间扩散,即时间轴上扩散,造成严重后果。所以,一般在经过信源编码之后,都需要在编码后的码流中有目的地按一定规则加入差错校正码,进行误码防护,这就是信道编码。信源编码是从信号中移去自然存在的冗余度,以达到更有效地表示信号的目的,而信道编码则是重新加入人工控制的冗余度,以获得更好的抗误码传输性能。显然,对于整个通信系统而言,信道编码增加的冗余度应比信源编码移去的冗余度少得多,从而在通信的有效性和可靠性间实现合理的平衡。实际上,设计一个数字图像或视频通信系统,要进行总体的优化设计和经济核算,要平衡图像质量、传输码率、抗误码性能和电路复杂性等各种要求和性能。若需实时处理,还要考虑编、解码算法的简捷、快速。而作为大众传播媒介的广播电视这类点对面的传输系统,为了降低接收机的成本,更需特别注意减少解码器的复杂性。

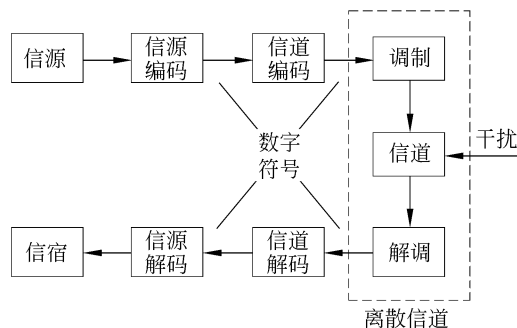


图 1-3 数字通信系统框图

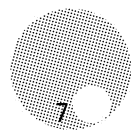


图 1-3 所示为包含有信源编、解码器,信道编、解码器和调制解调器在内的数字通信系统框图。

## 1.5 图像编码算法分类

图像数据压缩编码技术可以有多种分类方法。根据编码对象的不同,可分为静止图像编码、活动图像编码、传真文件编码;二值图像编码和多值(灰度)图像编码;黑白图像编码和彩色图像编码等等。根据压缩过程有无信息损失,可分为有损编码和无损编码(信息保持型编码),或称不可逆编码和可逆编码。根据编码算法中是否采用自适应技术,可分为自适应编码和非自适应编码。最常见的方法是按照算法的原理进行分类。实现图像数据压缩所采用的各类信源编码方法称为图像压缩编码算法,常见的图像编码算法主要可分为如下几类:

### 1. 预测编码

预测编码是利用图像信号在局部空间和时间范围内的高度相关性,以已经传出的近邻像素值作为参考,预测当前像素值,然后量化、编码预测误差。最常用的是差分脉冲编码调制(DPCM)。如果其中的量化器只有两个输出电平,则称为增量调制( $\Delta M$ )。根据图像局部的统计特性可以实现自适应预测,利用人的主观视觉特性可以实现自适应量化。如果预测编码系统中不使用量化器,则可以实现信息保持型预测编码。

与运动估值技术结合的运动补偿帧间预测是目前视频压缩编码系统中解除信号的时间轴冗余最常用和最有效的方法。

### 2. 变换编码

变换编码是将一组像素值经过某种形式的正交变换转换成一组变换系数,然后根据人的主观视觉特性对各变换系数进行不同精度的量化后编码的技术。正交变换的作用是解除像素值间的空间相关性,降低冗余度。用于图像编码的正交变换如离散傅里叶(Fourier)变换(DFT)、沃尔什-哈达玛(Walsh-Hadamard)变换(WHT)、哈尔(Haar)变换(HRT)、斜变换(SLT)、离散余弦变换(DCT)、K-L(Karhunen-Loeve)变换(KLT)等。这些变换中,除了 K-L 变换之外,都有快速算法。K-L 变换是在最小均方误差准则下进行图像压缩的最佳变换,但因其变换矩阵随图像内容而异,因此一般无快速算法。K-L 变换虽不宜用来进行实时编码,但在理论上具有重要意义,可以用来估计变换编码这一编码方式的性能极限。在这些正交变换中,当以自然图像为编码对象时,与 K-L 变换性能最接近的是 DCT。DCT 已被目前的多种静止和活动图像编码的国际标准所采用。

### 3. 统计编码

这是一类根据信息熵原理进行的信息保持型变字长编码,也称熵编码。编码时对出现概率高的事件(被编码的符号)用短码表示,对出现概率低的事件用长码表示。在目前图像编码国际标准中,常见的熵编码有赫夫曼(Huffman)编码和算术编码。如果被编码

事件是连续出现的某一符号的个数(符号串的长度,游程),则对其进行的统计编码称为游程编码。例如,“0”游程编码是指以连续出现的“0”的个数作为一个事件进行统计编码。为了降低编解码器的复杂度,当被编码的符号集比较大时,也可稍微对编码效率作些牺牲,采用准变字长编码,如线性码、对数码等。

#### 4. 子带编码

子带编码属于分析-综合类的编码技术。子带编码的基本思想是,在编码端,将图像信号在频率域分裂成若干子带(subband),而后对各个子带用与其统计特性相适合的编码器及比特分配方案进行数据压缩;在解码端,将分别解码后的各子带信号再综合成重建图像。子带编码有三方面的优点。

(1) 一个子带内的编码噪声(失真)不会扩散到其他子带,所以,图像经过子带编码后,其能量较弱的高频细节仍能较好地保持。

(2) 可以根据视觉特性控制编码噪声的频谱,使之适应人眼对不同频带噪声的敏感程度,将编码比特数在各个子带之间合理分配,有利于提高图像的主观质量。

(3) 子带编码由于其本身具备的频率分裂特性,非常适合于分辨率可分级、质量可分级的图像编码,也非常适合嵌入式的码流结构。

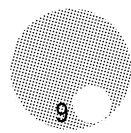
除了通过专门设计的正交镜像滤波器实现的经典子带编码方法之外,小波变换是目前使用最多的子带编码方法。早期的综合高频编码、塔型编码也属于子带编码的范畴。

#### 5. 基于模型的编码

模型基图像编码是一种基于景物三维模型的参数编码方法。相对基于波形的编码而言,对参数编码所需的比特数要少得多,因此可以节省大量的编码数据。依据对图像内容先验知识的了解,在编、解码双方建立一个相同景物的三维模型;基于这个模型,在编码器中,对图像内容进行分析,提取景物参数,例如形状参数、运动参数、纹理参数等;然后将这些参数编码传送给解码端,解码端根据接收到的参数和建立起的景物模型,采用图像合成技术再重建图像。模型基编码也属于分析-综合编码技术。

根据对图像内容先验知识的使用程度,模型基编码可以分为三个层次。最高层次的是“语义基”(semantic-based)图像编码,它是基于限定景物的模型基编码,其中景物里的物体的三维模型是严格已知的。处在中间层次的称为“知识基”(knowledge-based)图像编码,它没有像语义基编码那样掌握编码对象的全面知识,还需要利用物体几何参数来描述对象的变化。最低层次的是“物体基”(object-based 或 object-oriented)图像编码,它是针对未知物体的模型基编码,需要实时构造物体的模型。模型基图像编码目前主要用于以头肩像为对象的低码率可视电话编码。

分形编码实质上也可以认为是一种特殊的基于模型的编码方法,但与上述狭义的模型基编码不同,它所利用的是图像整体与局部间,或局部与局部间的自相似性。例如,基于分块的分形图像编码输出的是被编码像块与作为“模型”被选中的同样处在该图像内部的定义域像块间的仿射变换系数。



## 6. 其他编码算法

图像和视频压缩编码是一个跨学科的研究领域,也是一个十分活跃的研究领域,新的技术和新的方法不断涌现,新的算法层出不穷。图像编码的发展历史表明,应用数学和信号处理领域的新进展很快就会有在图像编码方面得到应用,例如,小波变换、分形几何、神经网络技术等等。

除了上述几大类编码算法外,还可举出很多压缩编码算法,例如,将多值图像转化为二值图像编码的比特平面编码,属于分组编码的矢量量化编码、块截断编码(block truncation coding),神经网络编码,轮廓编码,各种亚取样-内插编码等等。

各类图像编码方法一般都有自适应算法,即编码的参数不是固定不变的,而是根据图像信号的某种局部的或瞬时的统计特性,能自动地进行调整,以求达到更高的压缩效率。

在上述各种图像编码算法中,预测编码和变换编码占有重要的地位,对它们的研究历史长,应用也最广泛。而由这两种基本编码方法联合并与统计编码结合构成的“混合型”编码已成为活动图像编码方案中的主流类型,被很多视频编码国际标准所采纳。

## 1.6 图像编码的标准化

随着世界经济一体化进程的加快,全球性标准化体系的建立已成为规范世界经济发展的重要手段。以网络通信和计算机技术为代表的信息产业得以迅速发展,正是依靠其在产品体系和企业分工上建立的多层次的规范化的发展布局,而各种统一的标准体系的建立则是衔接这些企业和产品的关键性纽带。统一的国际标准中所蕴涵的巨大经济利益正驱使各大企业集团倾力投入其中,“三流企业做产品,二流企业做品牌,一流企业做标准”已成为一个高科技企业发展的必然模式。而当今的国际标准的制定已不再是简单的技术性问题解决过程,而是各大跨国公司集团为未来市场利益进行角逐的商业战场,有时甚至包含更深层次的区域经济和政治层面的内容,因而受到世人的广泛关注。

追溯图像编码技术近半个世纪的研究历史,在 20 世纪 80 年代以前图像编码的应用并不普遍,除了限于当时的技术条件之外,其中一个重要的障碍是缺少通用的图像编码标准。当需要对所传输或存储的图像信息进行高倍压缩时,往往必须采用多项图像压缩技术组成一个复杂的图像编码系统。但由于没有一个共同的标准作基础,不同的系统间数据结构不能兼容,除非各系统所采用的每一编码技术的各个细节完全相同,否则系统间的互接将十分困难。为此,ITU(国际电信联盟)和 ISO/IEC(国际标准化组织)等几大标准化组织自 20 世纪 80 年代后期以来在全世界范围内积极推动,联合各国在相关领域的专家进行共同研究,先后制定了一系列静止和活动图像编码的国际标准,并致力于面向未来应用的多媒体编码标准的研究。

关于静止图像压缩编码,1986 年 ISO 和 ITU 的前身 CCITT(国际电报电话咨询委员会)共同成立了一个联合图像专家组(joint picture experts group),其任务是要制定一个用于连续色调静止图像压缩编码的国际标准,1988 年确定了以自适应离散余弦变换(ADCT)作为该标准的基本算法,并于 1991 年通过了这项标准——JPEG。然而,随着多

媒体技术、互联网和无线通信的发展,人们对数字图像在质量和应用适应性方面提出了更高的要求,要求图像压缩技术不仅有良好的压缩效率,而且还能够根据使用目的和应用环境灵活地处理所生成的编码码流。在此情况下,JPEG 委员会的专家们从 1996 年开始,又致力于新一代的静止图像压缩编码国际标准的研究和测试。2000 年,JPEG 委员会公布的国际标准 JPEG 2000 以小波变换作为基本算法,采用了嵌入式编码技术,在达到更高的图像质量和更高的压缩效率的同时,还能满足在移动和网络环境下对互操作性和可分级性(可伸缩性)的要求。与 JPEG 相比,JPEG 2000 具有更先进的性能和更完善的功能,其应用领域也更为广泛。

在数字视频压缩编码方面,1984 年,ITU-T 的前身 CCITT 针对当时电信业务对会议电视和可视电话等视频技术的需要,成立了一个可视电话编码专家组,其任务是提出一个数码率不高于数字一次群传输速率(2Mb/s)的活动图像编码标准。经过十余个国家众多专家 5 年的倾力合作,1990 年颁布了用于视听业务(audio-visual service)的码率为  $p \times 64\text{Kb/s}$  ( $p=1 \sim 30$ ) 的视频编码国际标准,即 H. 261 建议(ITU-T 及其前身 CCITT 所建立的国际标准通常称为建议(recommendation))。H. 261 是视频压缩编码的第一个国际标准,它的制定集中体现了人们在图像编码方面近 40 年的研究成果,它以图像编码领域中三大经典技术——预测编码、变换编码和统计编码为核心,构造出的混合视频编码系统基本奠定了近代数字视频编码标准的系统框架模型,对其后的各种视频编码标准具有深远影响,因此它被视为图像通信发展史上一个重要的里程碑。随后,ITU-T 针对不同的电信通信网络中对实时视频通信系统的需要,先后完成了 H. 26X 系列中多个视频编码标准,其中包括 H. 261, H. 262, H. 263, H. 263+, H. 263++ , H. 264。

在 H. 261 制定的同时,ISO/IEC 联合技术委员会下属分委员会(SC29)于 1988 年成立了运动图像专家组 MPEG(moving picture experts group),该组织由来自全世界的一些研究机构和工业界的几百位代表组成,旨在制定针对多媒体信息存储(VCD、DVD 等),视频广播、视频流媒体(如基于 Internet 及 DSL 的视频,无线视频)等多种应用系统的音视频多媒体压缩编码标准。该委员会通过代表大会形式先后通过了 MPEG 系列的多个音视频压缩编码标准,包括 MPEG-1, MPEG-2, MPEG-4,以及后来关于多媒体信息系统框架的标准 MPEG-7 和 MPEG-21。

ITU-T 和 ISO/IEC 这两个标准化组织除了独立制定相关标准外,还进行合作,在 1994 年联合开发了 H. 262/ MPEG-2 之后,自 1997 年,ITU-T 的视频编码专家组 VCEG(video coding experts group)和 ISO/IEC 的活动图像专家组 MPEG 再次合作,于 2001 年宣布成立视频联合工作组 JVT(joint video team),配合当前广播电视网、电信网、互联网彼此融合的背景和无线与移动环境的要求,致力于形成新一代的视频图像编码标准。2003 年该工作组正式公布了这一最新的视频编码标准,称之为 H. 264/MPEG-4 AVC(advanced video coding)。由于 H. 264/MPEG-4 AVC 的压缩性能明显超出以前的视频压缩标准并能适应高等级的市场需求,因此,其在视频通信领域、数字广播电视领域、视频存储播放领域都有非常广阔的应用前景。

在我国,关于数字音视频编码标准化工作相对滞后,世纪之交关于第三代移动通信技术和数字视盘 DVD 的知识产权之争,不但使得即将迈出国门的中国企业感受到国际标