

第二部分 视频压缩

2.1 空间性或时间性编码?

2.2 空间编码

2.3 加权

2.4 遍历

2.5 熵编码

2.6 空间性编码器

2.7 时间性编码

2.8 运动补偿

2.9 双向编码

2.10 I、P和B帧画面

2.11 MPEG 压缩器

2.12 预处理

2.13 MPEG 的类和级

2.14 小波

第二部分

视频压缩

本部分将说明视频压缩是如何以视觉感受作为基础的。变换和运动补偿等重要关键技术被作为MPEG编码器结构的介绍来加以考虑。

2.1 空间或时间性编码?

如我们在第一部分中看到的那样，视频压缩主要是利用空间性和时间性冗余的特点。在MPEG中，首先通过使用连续画面间的相似之处降低时间性冗余。再通过使用已发送画面中的信息尽可能多地创建和“预见”当前画面。当使用该项技术时，我们仅需要发送差异画面，从而消除了实际画面和预见画面之间的差异。差异画面然后被进行空间压缩。从实际应用的角度出发，在解释时间性压缩这前解释之前性压缩更容易一些。

空间性压缩依赖于图像大块区域中相邻像素间的相似之处和图案区域中的主要空间频率。MPEG系统仅使用空间性压缩，因为它是设计用来传送单独的静止画面。然而，MPEG也可能用来编码连续的单独画面以构成视频节目。在所谓的“运动MPEG”应用中，压缩系数不会象用时间性编码时那样好，但数据流可以逐个画面地自由编辑。

2.2 空间性编码

空间性编码的第一步是用变换方法对空间频率进行分析。转换是表达不同域的波形(在这里是指频域)的一种简单方法。变换的输出是描述一给定频率出现多少的一整套系数。逆向变换则重现原始波形。如果各系数得到精确处理，那么逆向变换的输出就与原始波形完全一致。

我们最了解的变换就是傅里叶变换。该变换可找到输入信号的每个频率，方法是将输入波形乘以目标频率样本找到每个频率，叫作基础函数，然后将乘积求积分。图2.1显示，如果输入波形不含目标频率，则积分为零。而当波形含目标频率时，积分便是描述各频率分量幅度大小的系数。

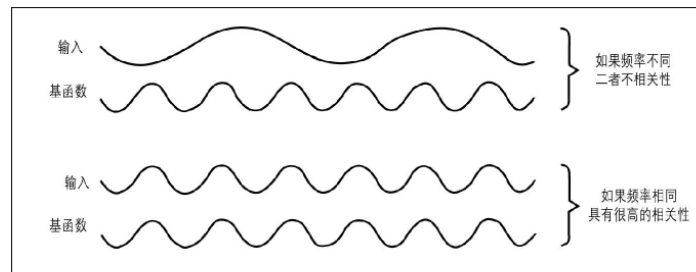


图 2.1

如果频率分量与基础函数同相，将产生一结果值。然而，如果频率分量与基础函数相位相差90度(正交)，则积分将仍为零。所以必需用相互正交的基础函数对每个频率搜寻两次，以保证输入的所有相位都被检测到。

傅里叶变换的缺点是同时需要每个频率的正弦和余弦分量系数。在余弦变换中，输入波形在乘以基础函数之前是与自身呈时间镜像的。图2.2 显示出镜像取消了所有正弦分量。并加倍了所有的余弦分量，正弦基础函数并不需要，所以每个频率只需要一个系数。

离散余弦函数(DCT)是余弦变换的取样形式，在MPEG的二维格式中被广泛应用。一个8 x 8像素块经变换后成为了8 x 8系数块。由于变换需要乘上分数，所以字长会增加，导致系数比像素值有更长的字长。一般来说，8 比特像素块会产生11比特系数块。

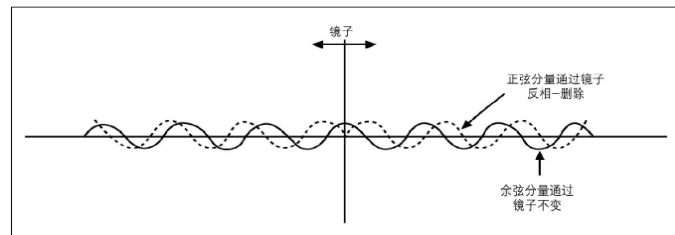


图 2.2

所以，DCT不会导致任何压缩，事实上它会产生相反的效果。然而DCT可以将源像素转变为更容易压缩的形式。

图2.3是每一个单独8x8DCT系数逆向变换后的结果。在亮度信号中，左上方的系数是整块的平均亮度或DC分量。在顶向上移动时(向右)，水平空间频率会增加。在左列上移动时(向下)、垂直空间频率会增加。在实际画面中，不同的垂直和水平空间频率会同时出现，块中一些点的系数将代表所有可能的水平频率和垂直频率的组合。

图2.3也显示了以一维水平波形方式的系数。将这些波形与不同的幅度结合起来，无论水平和垂直方向都能重现8像素的组合。所以二维DCT的64系数的组合将产生原先的8x8像素块。很显然，对彩色画面而言，色差取样也需要处理。Y、Cr和Cb数据被组合成分离的8x8排列，并各自独立变换。

在相当真实的节目内容中，许多系数都是零值或接近零值，所以不会被传送。这就产生了实际无损耗的压缩。如果需要更高的压缩系数，那么非零系数的字长必须缩短。这会导致这些系数精度下降，并将在处理中产生损耗。如果当心的话，这种损耗至少能以可以看到的方式表现出来。

2.3 加权

图2.4显示，人类对画面中噪声的感觉并不是一致的，而是空间频率的函数。在高空间频率的情况下，可以容忍更多的噪声。同时，在画面细节清楚的地方视频噪声能有效地遮蔽掉，而在大块色区域中则很容易看到。读者知道，传统的噪声测量往往被加权，使技术测量与主观结果相关联。压缩降低了系数的精度，并与使用PCM中短字长取样有相似的效果，即噪声电平上升。在PCM中，缩短字长的结果是噪声电平在所有频率上均相等地上升。正如DCT将信号分割到不同频率上那样，我们也可以控制噪声频谱。通过加权处理，可以有效地提高比高频系数精确得多的低频系数精度。

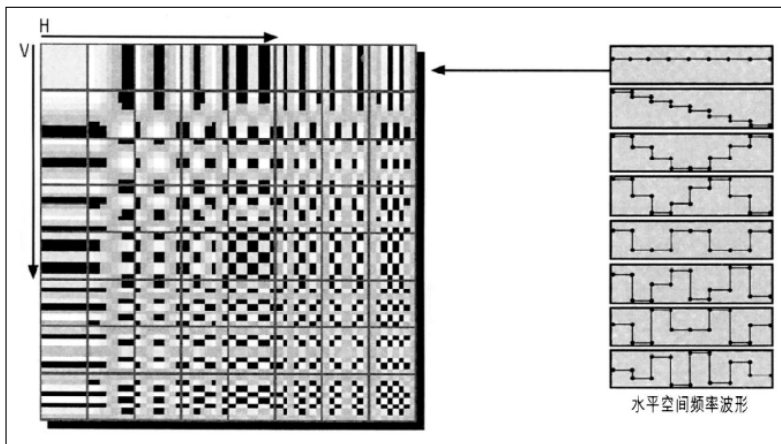


图 2.3

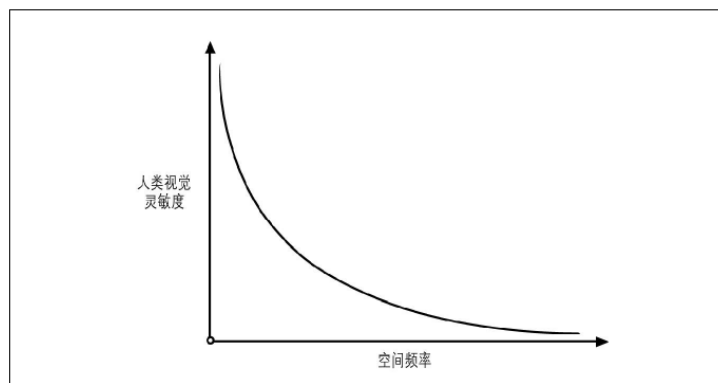


图 2.4

图2.5显示，在加权处理过程中，DCT的系数除以作为二维频率函数的常数。低频率系数将除以小数字，而高频系数将除以大数字。完成除法之后，最没有意义的数位被忽略和截除。截除是再次量化的一种形式。在没有加权的情况下，再次量化将达到两倍于量化步长大小的效果，而在有加权的条件下，步长大小则根据所除的系数增加。

结果，代表低空间频率的系数用相对较小的步长进行再次量化，并伴有稍稍增加的噪声。代表高空间频率的系数用对较大的步长进行再次量化，并伴有较大的噪声。然而，较小的步长意味着需要较少的数位来辨别步长和获得压缩。

在解码器中会加上低次序零来将加权的系数返回到其正确值。然后这些系数将乘以反向加权系数。显然高频系数处乘的系数会较大，从而使再次量化噪声变得较大。在反向加权之后，系数将具有它们原始的DCT输出值，再加上再次量化误差。量化误差在高频率下比在低频率下要大得多。

除了截除的办法外，加权的系数可以被非线性再次量化，从而使量化步长大小随着系数大小而增加。该项技术能获得较高的压缩系数，但却也会得到较差的副产物。

很显然，获得的压缩程度和输出数据率是再次量化处理严谨性的函数。不同的数据率需要不同的加权表。在MPEG中可能使用各种不同的加权表。使用的加权表能够传送给解码器，以自动产生正确的解码。

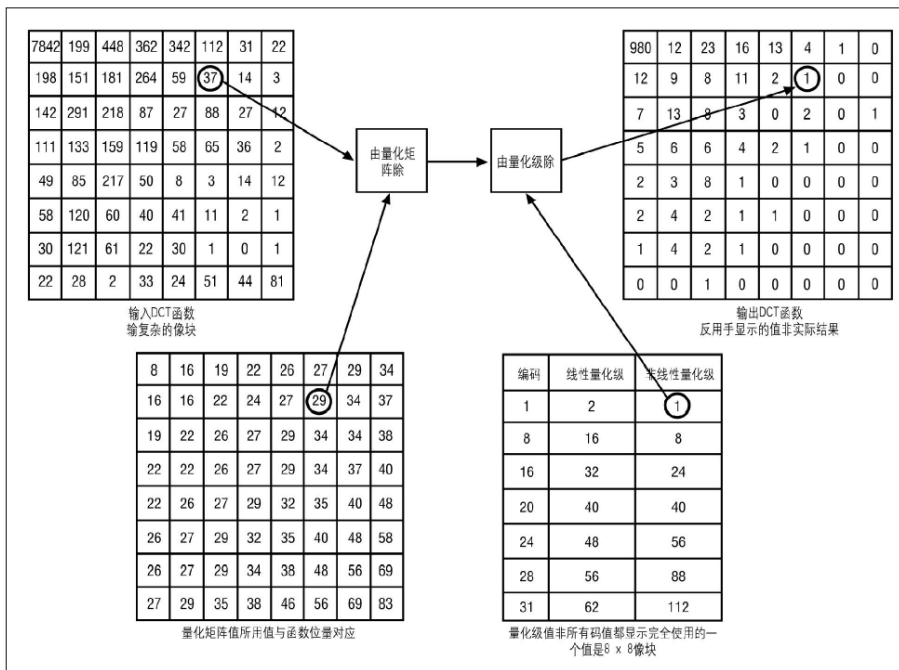


图 2.5

2.4 遍历 (Scanning)

在典型的节目内容中，有效的DCT系数通常位于矩阵的左上角。加权之后，低值系数可能变成零。如果能够先把所有非零系数发送出去，再以一个编码表示所有剩下的部分均为零，那么传送的效率将更大。遍历便是一项能够增加达到该目的可能性的技术，因为它将系数按大小的可能性的降序排列递送。图2.6a 显示了一个非交织系统中，系数的可能性在左上角数值最大，在右下角数值最小。使用时的最佳顺序是按45度对角线以之字形遍历。

在图2.6b 中是一个交织源的遍历。在交织画面中，8 x 8 DCT 块从一场延伸到两倍的垂直屏幕区域，所以对一个给定的画面细节而言，垂直频率则表现为水平频率的两倍。由此可见，交织画面理想的遍历是在两倍陡峭度的对角线上进行。图2.6b显示了一个给定的垂直空间频率在扫描相同水平空间频率之前进行扫描。

2.5 熵编码

在实际视频信号中并非所有空间频率都同时出现，所以DCT系数矩阵中将存在零值。尽管使用了遍历，但零系数仍然会在有效的数值之间出现。运行长度编码(RLC)能够让这些系数得到更有效的处理。当重复数值出现时，比如有一连串零，那么运行长度编码将只传送零值的数量，而不是传送每个单独比特。

在实际视频信号中特定系数出现的可能性是可以研究的。在实际应用中，有些数值经常出现，而有些数值不常出现。这些统计信息可以帮助我们使用可变长度编码(VLC)实现进一步的压缩。经常出现的数值被转换成短码字，而不常出现的数值被转换成成长码字。为了帮助不连续的情况，任何码字均不是另一个码字的字头。

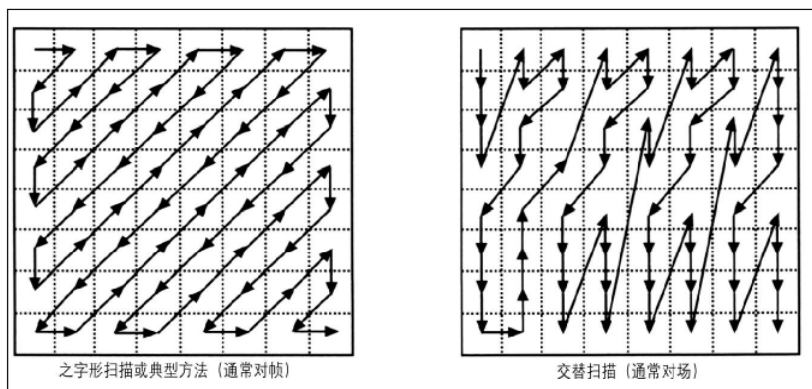


图 2.6

2.6 空间性编码器

图2.7 将前面讨论的所有空间性编码概念联系了起来。假设输入信号为4:2:2SDI(串行数字接口)，字长可能为8位或10位。MPEG只使用8位分辨率，所以在SDI信号含有10位字长时需要进行取舍。大多数MPEG组态用4:2:0的取样方式进行工作，所以就需要一个垂直低通滤波/插入阶段。取舍和色彩亚取样都会带来少量不可改变的信息损耗和数据率的部分降低。因此，我们需要存储光栅扫描输入格式以保证它能转变成8 x 8象素块。

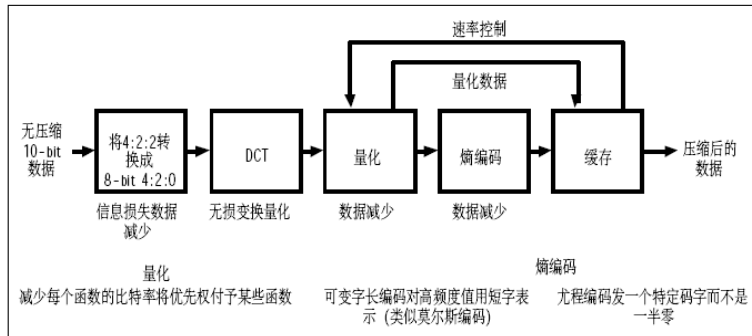


图 2.7

在DCT 阶段是将画面信息变为频域，而DCT本身并没有完成任何压缩。DCT之后，系数被加权和截除，从而实现第一次有效压缩。然后进行之字形遍历以增加有效系数在扫描中率先出现的可能性。在最后一个非零系数后面产生一个EOB(块末) 编码。

在运行长度编码和可变长度编码中系数数据被进一步压缩。在可变数据率系统中，量化是固定的，但在固定数据率系统中，缓冲存储器则用来缓解编码中数据率变化的困难。高度细节化的画面将填补这些缓冲区，而大块画面仍将让缓冲区保持空白。如果缓冲区有过载危险的话，再次量化步长就要变大，以使压缩系数有效增加。

在解码器中，数据流是不连续的，熵编码反向重新产生加权的系数。运用反向加权，根据之字形遍历将系数放置在矩阵中以重建DCT 矩阵。反向变换后，8 x 8 像素块被重新创建。为了获得光栅扫描输出，像素块被存储在RAM 中，每次可读一行。为了从4:2:0 数据中获得4:2:2 输出，需要如图2.8 所显示的那样进行垂直插入处理。

4:2:0中的色度取样被放置在纵轴亮度取样一半的地方，以便在使用交织源时被均匀间隔。

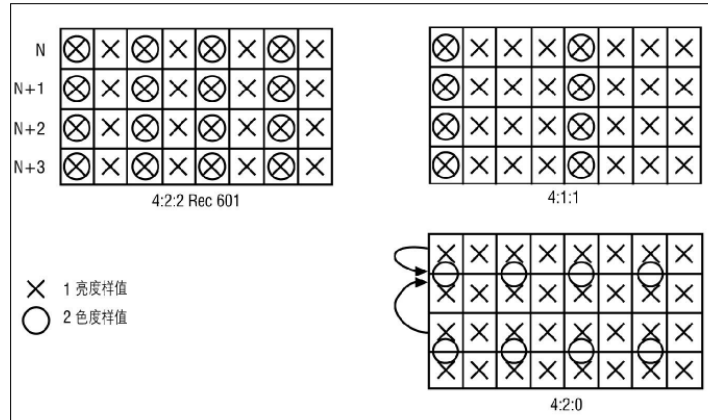


图 2.8

2.7 时间性编码

通过帧间编码或仅传送画面差异可以发掘时间性冗余。图2.9显示，与减法器配合的单幅画面延迟存储可以计算画面差异。画面差异是在其右边的图象，可以如前面讨论过的那样由空间性编码器进一步压缩。解码器将空间性编码反过来，并在上一幅画面上加上差异画面以构成下一幅画面。

这个简单的系统有一些缺点。首先，由于只发送差异，所以在开始发送后不可能马上开始解码。在从一个数据流向另一个数据流转换后(在观众转换频道时会发生)，这一局限性会使解码器提供画面变得困难。其次，如果差异数据的任何部分不正确的话，画面中的误码在传送中将无限地连续。

解决这些问题的办法就是使用并非完全不同的系统。图2.10显示了发送阶段性完整画面的情况。这些画面叫作帧内编码画面(或I-画面)，并且仅能通过空间压缩获得。如果发生误码或转换频道，在下一幅I-画面就能恢复正确解码了。

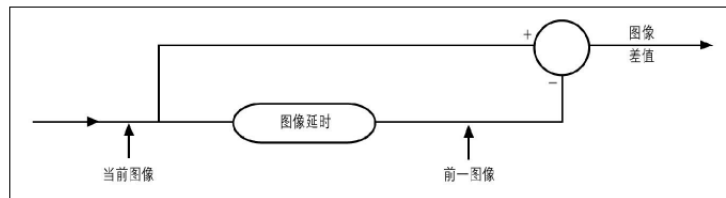


图 2.9

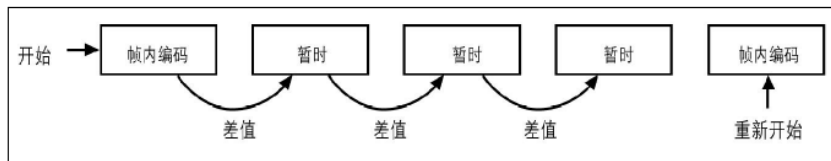


图 2.10

2.8 运动补偿

运动会降低画面之间的相似性，因而增加创建差异画面所需要的数据。运动补偿则是用来增加画面的相似性。图2.11就显示了这一原理。当一个物体在电视屏幕上通过时，它可能在每幅画面上出现在不同的位置，但其本身的外形并没有多大的改变。通过在编码器上测量运动可以降低画面差异。画面差异被以矢量的方式送往解码器。解码器使用该矢量将上一幅画面的一部分移至新画面中更恰当的位置上。

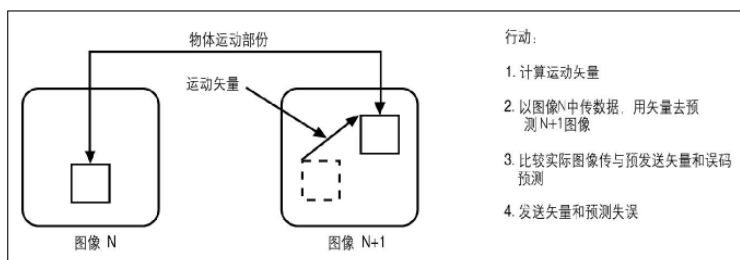


图 2.11

一个矢量控制着被叫作宏块的整个画面区域的移动。宏块的大小由DCT 编码和色彩亚取样结构来决定。图2.12a显示，在4:2:0系统中，彩色取样的垂直和水平空间正好是亮度空间的两倍。单个的8 x 8 DCT彩色取样块在同一区域上延伸为四个8 x 8亮度块；所以这是矢量能够移动的最小画面区域。一个4:2:0的宏块包含四个亮度块、一个Cr块和一个Cb 块。

在4:2:2组中，色彩只是在水平轴上被亚取样。图2.12b 显示，在4:2:2系统中，单个8 x 8 DCT 彩色取样块延伸到两个亮度块。一个4:2:2的宏块包含四个亮度块、两个Cr块和两个Cb 块。

运动估算器是通过比较两个连续画面的亮度数据的方式工作的。第一幅画面中的宏块被用来作为参考。当输入是交织状态时，像素在两个场中的垂直位置是不同的，所以在作比较时需要先插入一场。在整个搜索范围内用半个像素的分辨率测量所有可能的位移，从而得出参考画面和下一幅画面之间的相互关系。当找到最大的相互关系时，该相互关系就被假设为代表正确的运动。

运动矢量有一个垂直分量和一个水平分量。在典型的节目内容中，运动是在一些画面上连续发生的。如果矢量在发送时有变化，那么我们就需要较大的压缩系数。也就是说，如果一个物体以固定速度移动的话，矢量就不发生变化，矢量差异为零。

运动矢量与宏块有关，而与图象中的实际物体无关。但有时宏块的一部分移动，而另一部分不移动。在这种情况下就不可能实现正确补偿。如果移动部分的运动通过发送一个矢量得到补偿的话，那么固定的部分就会被不正确地移动，也就需要差异数据予以校正。如果不发送矢量的话，固定部分将保持正确，也就需要差异数据来校正移动的部分。实际使用的压缩器可能会同时尝试两种方法，并选择其中需要差异数据最少的一种。

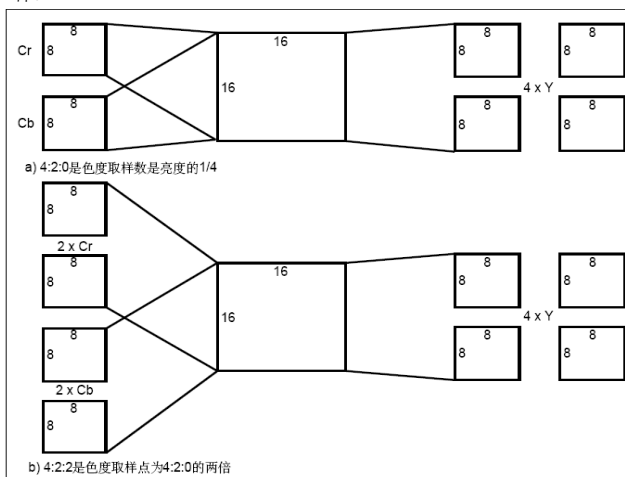


图 2.12

2.9 双向编码

物体移动时，它的前部边缘会遮盖掉背景，而尾部边缘会再现背景。再现的背景就需要传送新的数据，因为该背景区域前面被遮

盖了，从前幅画面上就得不到任何信息。类似的问题在摄像机摇拍时也会发生：新的区域出现时，我们在此之前根本无法知道。MPEG可以通过使用双向编码帮助我们克服这一问题，它允许从当前画面以前或以后的画面中获取信息。如果背景被遮盖时，它仍会在以后的画面中出现。我们可以及时向后移动获取信息，从而创建较早的画面。

图2.13 显示的就是双向编码的概念。在单个宏块的基础上，双向编码画面可以从前面画面、后面画面，以致于前面或后面数据的平均值上获取运动补偿数据。双向编码通过提高预见可能性的程度大大降低了所需的差异数据量。MPEG并不规定如何来建立编码器，而只规定了组成符合要求的数据流的内容。然而，好的压缩器能够尝试所有三种编码方法，并从中选择需要发送数据最少的一种方法。

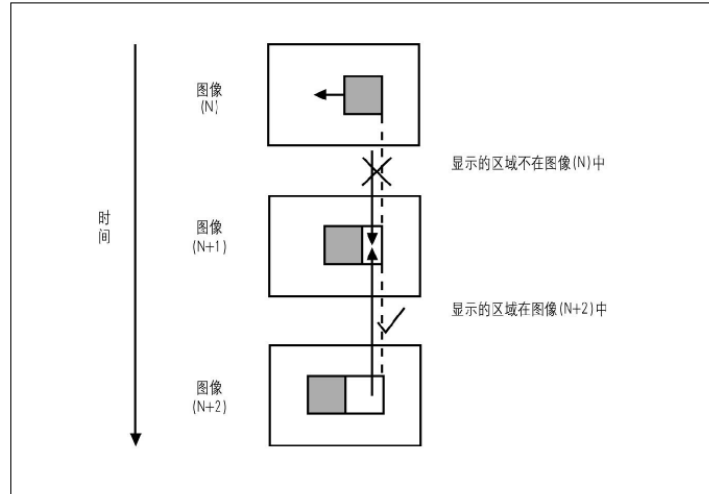


图 2.13

2.10 I、P 和 B 画面

在MPEG中，我们需要三种不同类型的画面来支持差异和双向编码，以减少误码的传递。

I 帧画面是指帧内编码画面，解码时不需要其它额外的信息。与其它画面类型作比较，I 画面需要许多数据，所以只是在需要时才发送I画面。I 画面主要由变换系数组成，不含矢量。它允许观众变换频道，并防止误码的传递。

P 画面是指从前画面中得到的前向预计画面，前面画面可以是I 画面或P 画面。P 画面数据由在前面画面中描述的从每个宏块的矢量所组成，而不是由描述必须加到宏块上的校正或差异数据的变换系数所组成。P画面需要的数据大约是I画面的一半。

B画面是指从前后的I画面或P画面上双向预见的画面。B 画面数据由描述在前面或后面的画面数据中从哪里提取的矢量所组成。B画面包含变换系数，用于校正。由于双向预见非常有效，所以校正数据是最小的。B画面需要的数据大约是I画面的四分之一。

图2.14介绍了GOP的概念，即画面组。画面组由I画面开头。然后P画面间隔排列。余下的画面是B画面。一个GOP的结束被定义为下一个I画面前的最后一个画面。GOP的长度是很灵活的。但普通的值是12或15 幅画面。很显然，如果B画面的数据要从将来的画面中提取的话，这些数据就必须是在解码器中已有的。因此，双向编码需要不按顺序地发送画面数据和暂时存储画面数据。图2.14 也显示了P 画面数据在B 画面数据之前发送。请注意，GOP中最后的B画面在下一个GOP的I 画面出现之前不能被发送，这是因为需要该数据进行双向解码。为了使画面返回到其正确的顺序，每个画面均包含一个时间性参考值。由于画面率也被周期性地放在数据流的头部，所以MPEG文件可在个人电脑上按正确的顺序和时间时标显示。

不按顺序发送画面数据就需要编码器和解码器有额外的存储器，同时也会造成延迟。帧内预见或向前预见画面间的双向编码画面数量必须受到限制，以降低成本和减少延迟。

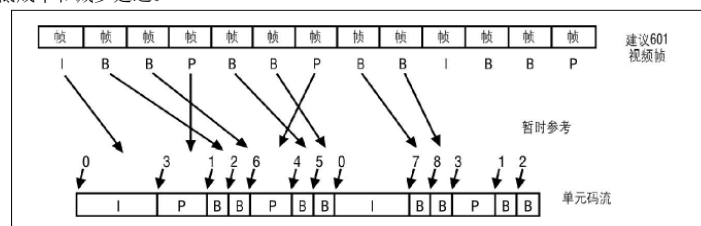


图 2.14

图2.15显示的是必须在压缩系数和编码延迟之间采取一种折衷。对于给定质量而言，仅仅发送I 画面所需的数据率是IBBP序列的两倍。当编辑功能很重要时，IB序列是一种有用的折衷办法。

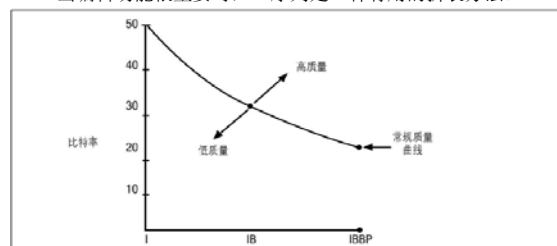


图 2.15

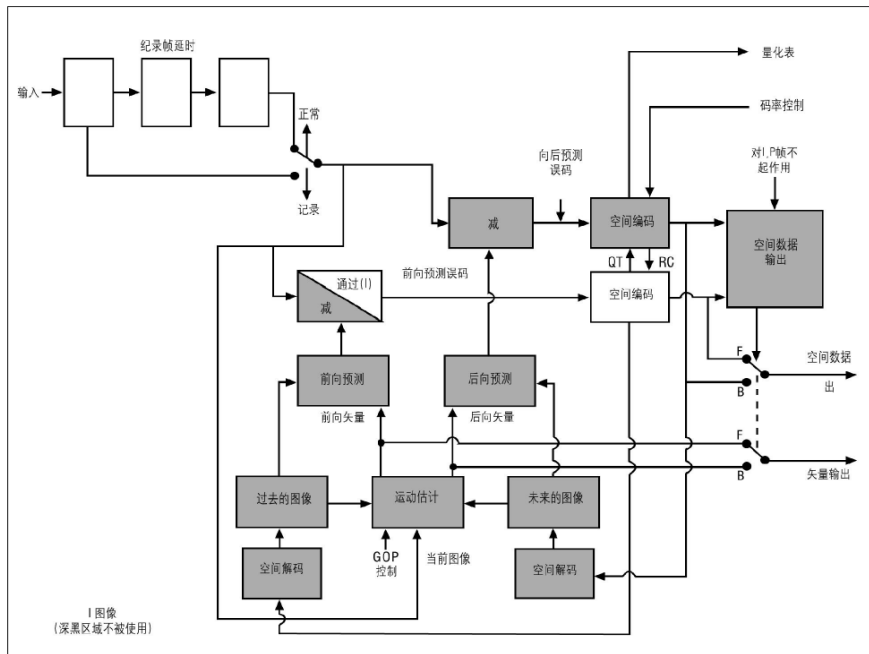


图 2.16a

2.11 MPEG 压缩器

图2.16a、b 和c 是典型的双向运动补偿器结构。预处理的输入视频输入一系列能够旁通以改变画面顺序的帧储存。然后数据再输入到减法器 and 运动估算器中。为了创建I 画面(见图2.16a)，我们选择输入延迟的末端并关闭减法器，使数据直接通过进行空间编码。减法器的输出数据也进入保持多个画面的帧储存。I 画面被保持在存储器中。

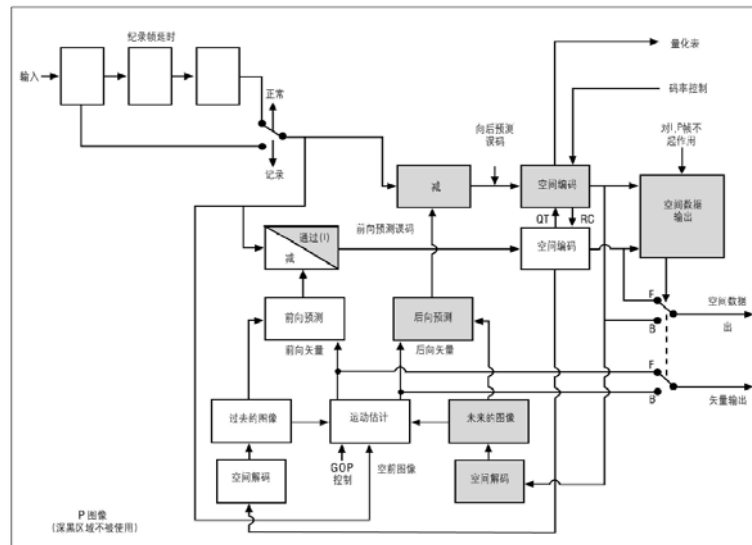


图 2.16b

为了编码P 画面(见图2.16b)，输入缓冲区中的B画面被旁通以便选择将来画面。运动估算器比较输出储存中的I 画面和输入储存中的P 画面，以便创建向前运动矢量。I画面通过这些矢量移动，以制作预见P画面。实际P 画面减去预见P 画面就得到预见误差，该误差经空间性编码后与矢量一起发送。预见误差也被加到预见P画面上以创建本机解码P 画面，它也被输入到输出存储中。

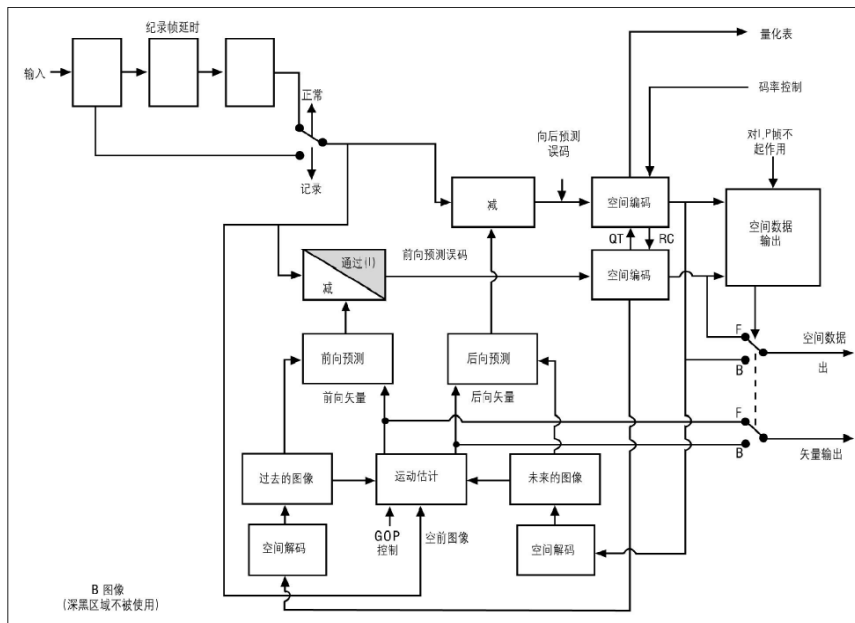


图 2.16c

此时输出存储中含有一个I 画面和一个P 画面。现在可以从输入缓冲区中选择B 画面。运动补偿器(见图2.16c)会将B 画面与前面的I 画面和后面的P 画面进行比较,从而获得双向矢量。然后进行向前和向后的运动补偿,产生两个预见B画面。当前B 画面再减去这些预见B画面。在逐个宏块的基础上,根据哪个数据代表最小的差异来选择向前和向后数据。接着,画面差异被空间编码后与矢量一起发送。

当所有的中间B画面被编码后,输入存储器将再一次被旁通,根据前面的P 画面创建一个新的P 画面。

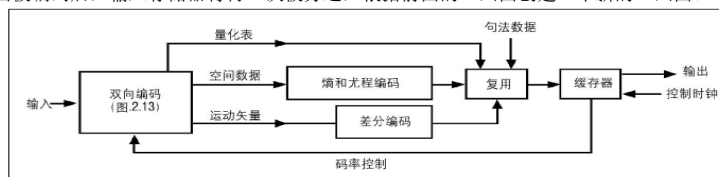


图 2.17

图2.17是MPEG编码器。运动补偿器的输出被空间编码,矢量被加到多路调制器上。句法数据也被加上去以辨别画面类型(I、P 或 B),并向解码器提供其它有助信息(见第四部分)。输出数据经缓冲允许数据率的暂时变化。如果数据率长时间增长,缓冲区将填充以防止过载。量化处理也要更为严谨。同样,如果缓冲区下溢,则量化将松弛以保持平均的数据率。这就意味着编码器中的存储器存储什么内容,解码器中的存储器也将精确地存储什么内容,因此所有以前编码错误的结果也都会表现出来。当实际画面减去预见画面后编码误差会自动减少,这是因为差异数据更加精确了。

2.12 预处理

压缩器力图消除画面内和画面间的冗余。但有时,我们又不希望有任何降低这种冗余的东西存在。噪声和胶片颗粒尤其成问题,因为它们常常在整个画面上出现。在DCT处理之后,噪声会导致更多的非零系数,而编码器却无法将它与真正的画面数据区分开。编码所有的这些系数要求使用更重的量化,从而降低了画面质量。噪声也会降低连续画面间的相似性,从而需要增加差异数据。

从复合视频中解码出来的剩余副载波是一个严重的问题,因为它会导致一般在分量节目低电平处出现的高空间频率。副载波也会在画面与画面之间相位交替,从而引起差异数据的增加。一般来讲,在MPEG编码器输入中可以看到任何复合视频解码的副产物,都会在解码器中重新生成。

我们要避免任何会引起不需要运动的操作。摄像机未固定好不仅会产生晃动的画面,而且会增加画面差异和矢量传送要求。这种情况也会发生在电影胶卷上,因链状孔损坏造成胶片卷起或跳动,产生晃动画面。一般来讲,需要压缩的视频必须尽可能保证高质量。如果不能保证高质量,那么必须具备噪声降低和其它稳定技术。

如果需要高压缩系数,那么副产物的电平会增加,尤其当输入质量很差时更是如此。在这种情况下最好使用预滤波来降低进入编码器的熵。视频信号经过双向低通滤波能降低所需的系数数量和副产物电平。画面轮廓将不很明显,但对副产物高电平来说我们还是愿意作出这一选择。

在大多数MPEG-2应用场合中我们使用4:2:0的取样方式,当源信号为4:2:2 时要求进行色度低取样处理。在MPEG-1中亮度和色度被进一步低取样来产生输入画面或SIF(源输入格式),仅352 像素宽。该项技术通过深度系数降低了熵。对很高的压缩而言,我们使用QSIF(四分之一源输入格式)画面,仅176 像素宽。低取样是将空间低通滤波器和插入器组合起来的处理过程。低取样交织信号时会出现问题,因为垂直细节延续在两场上,会因运动而出现非相关性。

当源内容是电影电视时,视频信号与正常信号相比具有不同的特点。在50Hz视频中,成对的场代表相同的胶片帧,它们之间不存在运动。因此,场之间的运动在无运动与帧间运动之间交替。由于运动矢量是差分传送的,所以它会导致矢量数据的急剧增加。在60Hz 视频中,用3:2下拉方式从24Hz电影软片中获得60Hz 视频。第一帧分成两个场,下一帧分成三个场,依此类推。

然后,一场一分为五是完全多余的。MPEG处理胶片素材的最好方法是忽略3:2系统中的第三场。传送过程中的24Hz 编码让解码器通过重读场存储来重新建立3:2序列。在50 或60Hz 电影电视中,成对的场经解交织来创建帧,再测量帧与帧之间的运动。通过读帧

存储中的交替行，解码器能够重新创建交织。

快切是压缩器难以处理的困难事件，因为它几乎肯定会引起故障，从而需要大量的校正数据。如果编码延迟能够容忍的话，编码器可以预先检测到快切并动态地修改GOP结构，所以让快切与一个I画面的发生相吻合。此时，快切只需少量的额外数据便可处理。I帧前面最后的B画面肯定需要使用前向预测。在一些非实时应用场合中(如DVD)，编码器可在输入视频上通过两次。第一次识别困难或高熵区域，并建立编码方法。第二次则对输入视频进行实际压缩。

2.13 类和级

MPEG适用于大量需要不同性能和复杂性的应用场合。使用MPEG中定义的所有编码工具可以有上百万种可能的组合。在实际应用中，MPEG-2标准被分为各个组态(类)，每个组态又分为不同的层(级)。(见图2.18)。组态基本上是整个具备一定复杂性的编码指令表的子集。层则是与组态一起使用的参数，如画面尺寸或数据率。从理论上讲，一共有24种组合，但并非所有的组合都被定义了。具有一个给定组态和层的MPEG解码器也必须能够解码较低的组态和层。

简单的组态并不支持双向解码，所以只有I画面和P画面可以成为输出。这就降低了编码和解码延迟，允许用较简单的硬件。简单的组态只能在主层上(SP@ML)定义。

主组态有很大的用处。低层使用低分辨率输入，每行只有352个像素。大部分广播应用需要MPEG的MP@ML(主层的主组态)子集，支持SDTV(标准密度电视)。

高1440层次是一个高分辨率方案，与主层次相比分辨率增加一倍。高层不仅使分辨率增加一倍，而且通过增加水平取样的数量(从1440点到1920点之间)来保持16:9格式的分辨率。

在使用空间变换和量化的压缩系统中可产生可度量的信号。可度量过程是指输入产生一个主信号和一个“帮助”信号。主信号可以独自解码给出一个达到一定质量的画面，但如果加上帮助信号中的信息，质量的某些方面就会改善。

例如，普通MPEG编码器通过重量化系数按适当信噪比结果来编码画面。如果画面是本地解码，并与原始画面逐个像素相减，那么就会产生一个量化噪声画面。该画面可以被压缩，并作为帮助信号发送。简单的编码器只能解码主噪声数据流，但复杂的解码器能同时解码两种数据流，并将它们组合起来产生低噪声画面。这就是SNR可度量性理论。

作为一种供选择的方法，仅仅编码HDTV画面中较低的空间频率能够产生一个主数据流，该数据流能够被SDTV接收机解码。如果较低密度画面由本地解码，并与原始画面相减，那么会产生密度提高的画面。该画面可编码到帮助信号中。适当的解码器能够将主信号和帮助信号组合起来重建HDTV画面。这就是空间可度量性理论。

高组态支持SNR可度量性和空间可度量性，同时允许选择4:2:2取样。

4:2:2组态被开发出来改善了数字制作设备的兼容性。该组态允许4:2:2操作，而避免了因使用高组态而带来的复杂性。例如，HM@ML解码器必须支持制作时不要求的SNR可度量性。4:2:2组态具有与其它组态相同的GOP结构自由，但在实际应用中往往使用较短的GOP，以便使编辑变得更为容易。4:2:2操作比4:2:0操作要求有更高的数据率，并且对于给定的画面质量短GOP的使用也要求更高的数据率。

高		4:2:0 1920x1152 80 Mb/s I,PB				4:2:0, 4:2:2 1920x1152 100 Mb/s I,PB
高-1440		4:2:0 1440x1152 60 Mb/s I,PB			4:2:0 1440x1152 60 Mb/s I,PB	4:2:0, 4:2:2 1440x1152 80 Mb/s I,PB
主	4:2:0 720x576 15 Mb/s I,P	4:2:0 720x576 15 Mb/s I,PB	4:2:2 720x608 50 Mb/s I,PB	4:2:0 720x576 15 Mb/s I,PB		4:2:0, 4:2:2 720x576 20 Mb/s I,PB
低		4:2:0 352x288 4 Mb/s I,PB		4:2:0 352x288 4 Mb/s I,PB		
级类	SIMPLE	MAIN	4:2:2 PROFILE	SNR	SPATIAL	HIGH

图 2.18

2.14 小波

所有的变换都有不稳定性，因为知道的频域越精确，则知道的时域越不准确(反之亦然)。在诸如DFT和DCT那样的大多数变换中，块长度是固定的，所以时间和频率分辨率也是固定的。频率系数代表了线性刻度上平均分布的值。然而，因为人类的感觉是对数性的，所以DFT和DCT的平均刻度要么频率分辨率不足，要么分辨率太高。

小波变换并未受这个问题的影响，因为它的频率分辨率是倍频带固定的一部分，所以具有对数特性。块长度改变是频率的函数。随着频率的下降，块将变长。所以，小波变换的特点是基础函数都含有相同数量的周期，这些周期被简单地标在时间轴上以找寻不同的频率。

图2.19显示了DFT/DCT的固定块尺寸与小波的可变尺寸之间的对比情况。

小波对音频编码特别有用，因为它们能自动适应不一致的要求，即准确的瞬态音定位和稳态音中音调的准确评估。

对视频编码而言，小波的优点是不需另作努力便能够产生分辨率可度量的信号。在移动视频中，由于运动矢量被指定到尺寸可变的块很困难，小波的优点就被抵销了。但在静止画面或I画面编码中，这种困难性就不成问题了。

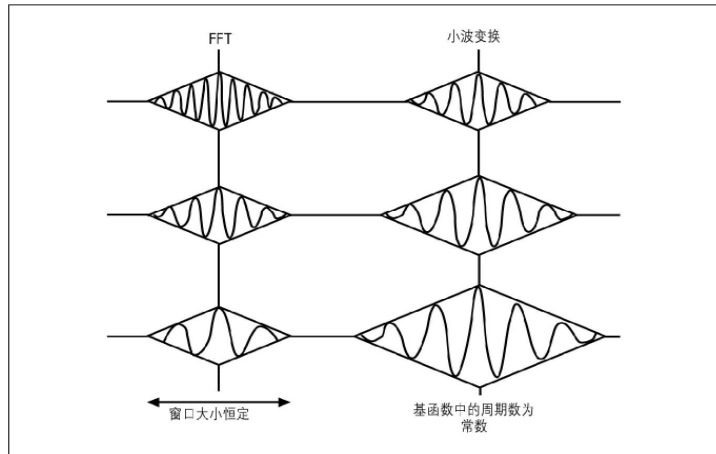


图 2.19