

MPEG 基础和协议分析指南

（包括 DVB 和 ATSC）

目录

第一章 MPEG 概述

- 1.1 整合
- 1.2 压缩的必要性
- 1.3 压缩原理
- 1.4 压缩在电视中的应用
- 1.5 数字视频压缩简介
- 1.6 音频压缩简介
- 1.7 MPEG 码流
- 1.8 监视和分析的必要性
- 1.9 压缩带来的问题

第二章 视频压缩

- 2.1 是空间域编码还是时间域编码？
- 2.2 空间域编码
- 2.3 加权
- 2.4 扫描读取
- 2.5 熵编码
- 2.6 空间域编码器
- 2.7 时间域编码
- 2.8 运动补偿
- 2.9 双向编码
- 2.10 I P B 图象帧
- 2.11 MPEG 压缩器
- 2.12 预处理器
- 2.13 小波变换

第三章 音频压缩

- 3.1 听觉机理
- 3.2 子带编码
- 3.3 MPEG 层 1
- 3.4 MPEG 层 2
- 3.5 变换编码
- 3.6 MPEG 层 3
- 3.7 MPEG-2 音频
- 3.8 MPEG-4 音频
- 3.9 AC-3

第四章 MPEG 标准

- 4.1 什么是 MPEG
- 4.2 MPGE-1
- 4.3 MPEG-2
 - 4.3.1 MPEG-2 的类和级
- 4.4 MPEG-4
 - 4.4.1 MPEG-4 标准文件
 - 4.4.2 对象编码
 - 4.4.3 视频和音频编码
 - 4.4.4 可分级性
 - 4.4.5 MPEG-4 的其它方面
 - 4.4.6 MPEG-4 的未来
- 4.5 MPEG-7
- 4.6 MPEG-21

第五章 基本码流

- 5.1 视频基本码流句法
- 5.2 音频基本码流

第六章 打包的基本码流 (PES)

- 6.1 PES 包
- 6.2 时间标记
- 6.3 PTS / DTS

第七章 节目流

- 7.1 记录与传输
- 7.2 节目流简介

第八章 传输流

- 8.1 传输流的工作过程
- 8.2 传输流包
- 8.3 节目时钟基准 (PCR)
- 8.4 包识别符 (PID)
- 8.5 节目专用信息 (PSI)

第九章 数字调制

- 9.1 调制原理
- 9.2 模拟调制
- 9.3 正交调制
- 9.4 简单数字调制系统
- 9.5 相移键控
- 9.6 正交幅度调制 (QAM)
- 9.7 残留边带调制 (VSB)
- 9.8 编码正交频分复用 (COFDM)

9.9 综合业务数据广播 (ISDB)

9.9.1 ISDB-S 卫星系统

9.9.2 ISDB-C 有线系统

9.9.3 ISDB-T 地面调制

9.9.4 ISDB 小结

第十章 DVB 和 ATSC 概述

10.1 引言

10.2 再复用

10.3 业务信息 (SI)

10.4 误码校正

10.5 信道编码

10.6 内编码

10.7 数字发射

第十一章 数据广播

11.1 数据广播应用

11.1.1 节目相关数据

11.1.2 机会数据

11.1.3 网络数据

11.1.4 增强电视

11.1.5 交互电视

11.2 内容封装 (Content Encapsulation)

11.2.1 MPEG 数据封装

11.2.1.1 数据管道 (Data Piping)

11.2.1.2 数据流

11.2.1.3 DSMCC (数字存储媒体指令和控制)

11.2.1.4 MPE (多协议封装)

11.2.1.5 传送带

11.2.1.6 数据传送带

11.2.1.7 对象传送带

11.2.1.8 对象传送带的广播

11.2.1.9 MPEG-2 数据同步

11.2.2 DVB 数据封装

11.2.3 ATSC A/90 数据封装

11.2.4 ARIB 数据封装

11.3 数据内容传送

11.3.1 DVB 公告

11.3.2 ATSC 公告

11.4 内容提交

11.4.1 机顶盒中间件

11.4.2 DVB 多媒体家庭平台 (MHP)

11.4.3 ATVEF DASE 高级电视增强论坛和数字电视应用软件环境

11.4.4 DASE 数字电视应用软件环境

第十二章 MPEG 测试

- 12.1 测试要求
- 12.2 传输流分析
- 12.3 分层图示
- 12.4 解释性图示
- 12.5 句法和 CRC 分析
- 12.6 过滤
- 12.7 定时分析
- 12.8 基本码流测试
- 12.9 萨尔诺夫顺应性比特流
- 12.10 基本码流分析
- 12.11 传输流的创建
- 12.12 PCR 不准确度信号的发生

术语汇编

第一章 MPEG 概述

MPEG 是目前最常用的一项音频 / 视频压缩技术，它实际上并不是一个单一的标准，而是包括了适合于各种不同应用的一系列标准，但这些标准是以一些共同理论为基础的。MPEG 是活动图象专家组的缩写词，该专家组是联合技术委员会（Joint Technical Committee, JTC1）的一部分，JTC1 是由 ISO（国际标准化组织）和 IEC（国际电工委员会）建立的。JTC1 负责信息技术，在 JTC1 中，下设有负责“音频、图象编码以及多媒体和超媒体信息”的子组 SG29。在 SG29 子组中，又设有多个工作小组，其中就包括 JPEG（联合图片专家组）和负责活动图象压缩的工作组 WG11。因此，可以认为 MPEG 是 ISO/IEC JTC1/SG29/WG11。

既然把 MPEG 描述为多个词首字母的组合，那么，就其工作过程而言，正如欧洲电信标准学会所说：“条件接收表（CAT）引导综合接收机 / 解码器去寻找与所使用的条件接收（CA）系统相关的授权管理信息（EMM）”。如果您能够理解这句话，您就不需要这本书了！

1.1 整合

数字技术在音频和视频上已经取得了迅速的进展，这是由多种因素决定的。其中一个重要因素就是数字信息的可靠性更高，可以通过编码以消除信息中的错误。这就意味着数字技术能够消除在记录和传输过程中所产生的损耗。激光唱盘（CD）就是采用数字技术的第一个消费类产品。

相对早期的聚乙烯唱盘，CD 的声音质量有了很大的改善。仅仅比较质量是不够的，真正的意义在于：数字记录和数字传输技术的应用使内容的处理达到了模拟技术所不可能达到的程度。一旦音频和视频信号被数字化后，其内容就是数据的形式。这种数据和其它任意类型的数据一样，可以用同样的方式处理。这样，数字视频和数字音频就作为数据而进入了计算机技术领域。

计算机和视音频的融合是数据处理和脉冲编码调制（PCM）应用的必然结果。数字媒介可以存储任何类型的信息，因此，将计算机存储器件应用于数字视频是很容易的。计算机技术和数字视音频相融合的第一个应用实例是非线性工作站，这项应用对于模拟来说是不可能的。另一个例子是将音频、视频、图形、文字和数据存储在同一介质的多媒体技术，这在模拟域中也是无法实现的。

1.2 压缩的必要性

数字视频首先在后期制作应用中获得了成功。在后期制作中，虽然数字视频的代价高昂，但它可以复制无限次而无损伤。然而，后期制作中所使用的标准数字视频，其数据率高达每秒 200 兆比特以上，这就需要大容量的存储器和宽带传输。只有在存储器和带宽需求容易满足的情况下，数字视频才有可能获得广泛的应用。压缩的目的正在于此。

压缩就是使数字视音频具有较低数据率的一种方式。压缩具有以下优点：

- ▶ 对于给定的信源素材，它只需要较少的存储量。
- ▶ 在实时工作时，压缩可降低所需带宽。此处，压缩可以使数据在存储介质间的传输速度更快，例如，在磁带和硬盘之间可以实现数据的快速传输。
- ▶ 采用压缩记录格式可以减少记录密度，这样就可以降低记录设备对环境因素和设备维护的要求。

1.3 压缩原理

在传递信息内容时，为了减少所需数据量，可以采用两种不同的基本技术。在实用的压缩系统中，常常是这两种技术的组合应用，采用了十分复杂的方式。

第一种压缩技术是提高编码效率。对于给定的信息，可以采用许多编码方式。在最简单

的视音频数据中，也包含有一定量的冗余度，这就是下面我们要讨论的“熵”的概念。

许多编码技术可以减少或除去这种冗余度。例如游程编码和可变字长编码系统（如霍夫曼编码）。如果应用适当，上述编码技术完全是可逆的，这就是说，解压缩后的数据与编码系统的输入数据是相同的。这种类型的压缩称为无损压缩。存档的计算机程序如 PKZip 就采用了这种无损压缩。

很明显，无损压缩虽然十分理想，但它却不能提供视音频应用所需要的数据压缩量。然而，正因为它是无损压缩，所以它可用于系统的任意点，通常可用在有损压缩器的数据输出端。

如果除去信息中的冗余度并不能满足所需要的数据压缩量，那就必须要丢弃某些（非冗余的）信息。有损压缩系统就是通过去除不相关的信息或相关性较低的信息来实现所需要的压缩量。不存在对任意数据流均为适用的通用的有损压缩技术；因为对相关性的评价只能就应用内容本身才能确定，在压缩时应当了解数据代表什么，它又是如何使用的。在电视情况下，图象和声音的再现是为人眼系统和听觉系统而提供的，因此，在设计一个有效的压缩系统时，就必须充分考虑人的主观感受因素。

在视频信号中，某些信息是不能被人的视觉系统所察觉的，因此，这样的信息内容就是真正的无关信息。只丢弃无关图象信息的压缩系统可称为视觉无损压缩系统。

1.4 压缩在电视中的应用

电视信号，无论是模拟还是数字，总包含有大量的信息。在电视发展的早期实际上已经在使用带宽压缩技术了。最早的例子或许是隔行扫描。对于给定的行数和给定的图象刷新速率，隔行扫描使信号的所需带宽压缩了一半。隔行扫描是有损处理，在隔行扫描中，由于垂直信息和瞬态信息之间的干扰而有所失真，从而降低了图象中的垂直清晰度。不过，所丢弃的大多数信息主要是无关信息，因此，在电视发展的早期，在清晰度和带宽均可接受的情况下，隔行扫描是一项简单的、成功的技术。然而，在复杂的数字压缩处理过程中，隔行扫描及其失真却带来了新的问题。MPEG-2 技术中的一些复杂性，就是因为需要处理隔行扫描信号而带来的。与处理逐行扫描信号相比较，编码效率要低些，信号损失也要大些。

在电视发展过程中，彩色的出现又带来了新的课题。彩色摄像机产生的是 GBR 信号，这样，彩色信息量就是单色信息的三倍——这就需要三个单色通道传输彩色信息。

为了解决这一问题，第一步就是将 GBR 信号转换为一个亮度信号（通常以 Y 表示）和两个色差信号即 U 和 V 信号，或 I 和 Q 信号。抽取亮度信号是在解决与单色接收机的兼容性问题上前进了一大步，然而，信号带宽的减小主要还是来自于色差信号的处理。

现在来考察人的视觉系统。人眼对亮度敏感，可以“看见”分辨率很高的图象。其它的彩色信息，人眼的分辨率就要低得多。这样的综合效果就是，在受到某种限制的情况下，如果将代表亮度场景的、清晰的单色图象与模糊的（低带宽）彩色信息相重叠，那么仍将出现清晰的彩色图象。这一优点在处理 GBR 信号时是不可能具备的，因为 GBR 中的每一种信号既包含亮度信息，又包含彩色信息。然而，在 YUV 域中，大多数亮度信息被 Y 信号所传送，极少量的信息由色差信号传送。这样，就有可能滤除一部分彩色色差信号以大幅度地降低所需传送的信息。

以上是消除（大多数）无关信息的一个例子。在通常的观看条件下，人的视觉系统不能显著地感受到色差信号中的高频信息，因此，色差信号中的高频信息也可以被放弃。在 NTSC 电视传输中，每路色差信号中只传送大约 500KHz 以内的信号频率分量，但在许多应用中，图象仍然是相当清晰的。

在 NTSC 和 PAL 处理过程中，压缩带宽的最后一步是将色差信号“隐藏”在单色信号频谱中的未用部分。尽管这种处理严格说来不是无损的，但它可以提高信号的编码效率。

与模拟技术相比较，数字域技术有相当大的差别，但有些原理是一样的。例如，MPEG 可将信号转换为不同的域以分离无关信息，但仍然将信号转换为色差域，在数字处理中可以滤除一部分色差信号以降低垂直分辨率。

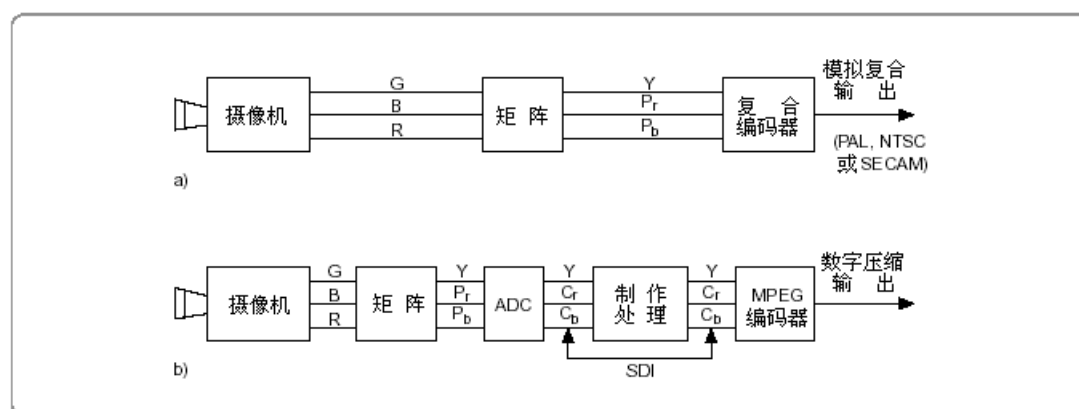


图 1—1

图 1—1a 表示的是传统的电视系统，GBR 摄像机信号变换为 Y、Pb、Pr 分量信号以用于制作，而后编码为模拟复合信号以用于传输。图 1—1b 表示的是目前正在使用的（与传统电视系统对应的）系统。Y、Pb、Pr 信号经数字化后变为 Y、Cb、Cr 信号，并以 SDI 形式通过制作系统，再经 MPEG 编码后用于传输。显然，MPEG 在这里只是作为传输复合信号的一种更有效的替代手段。此外，MPEG 具有更大的灵活性，因为可以按照其应用需要调整到合适的传输比特率。较低的比特率就意味着较低的分辨率，这样的 MPEG 图象可用于视频会议和视频电话。

数字视频广播（DVB）和先进电视制式委员会（ATSC）——即在欧洲和美国提出的数字电视广播标准中，信号必须要经过压缩，否则需要太宽的频带。压缩技术延长了 DVD（数字电视 / 通用光盘）的重放时间，整场电影可以存储在一个单独的光盘上。压缩也降低了电子新闻采集和其它电视制作应用的成本。DVB、ATSC 和数字视频光盘（DVD）等均是基于 MPEG—2 压缩技术。

在磁带记录中，允许采用适度的压缩可以提高数字 Betacam 和 Digital-S 系统的可靠性，然而在 SX、DVB、DVCPRO 和 DVCAM 中，采用压缩技术有利于实现系统设备小型化。在基于硬盘的视频服务器中，压缩可以降低存储器的成本。通过压缩减少了传输带宽，这样就能允许更多的用户去访问一个给定的服务器。这项特性对于 VOD（视频按需点播）应用也是重要的。

1.5 数字视频压缩简介

在所有的实际节目素材中，存在着两种类型的信号分量：即异常的、不可预见的信号分量和可以预见的信号分量。异常分量称为熵，它是信号中的真正信息。其余部分称为冗余，因为它不是必需的信息。冗余可以是空间性的，如在图象的大片区域中，邻近像素几乎具有相同的数值。冗余也可以是时间性的，例如连续图象之间的相似部分。在所有的压缩系统编码器中都是将熵与冗余相分离，只有熵被编码和传输，而在解码器中再从编码器的发送的信号中计算出冗余。图 1—2a 表示了这一过程。

一个理想的编码器可以抽取所有的熵并只将熵传输到解码器。理想的解码器能够从熵中恢复原始信号。然而实际上，这种理想的编码解码器是不可能实现的。因为这种理想的编码

器在技术上很复杂，而且为了使用时间性冗余而造成很长的延时。在某些应用中，例如节目记录或某些广播传输中，有些延时还是可以接受的，但在视频会议中却是不允许的。还有，一个非常复杂的编码器在价格上也是很昂贵的。这也就是说，不存在一个理想的压缩系统。

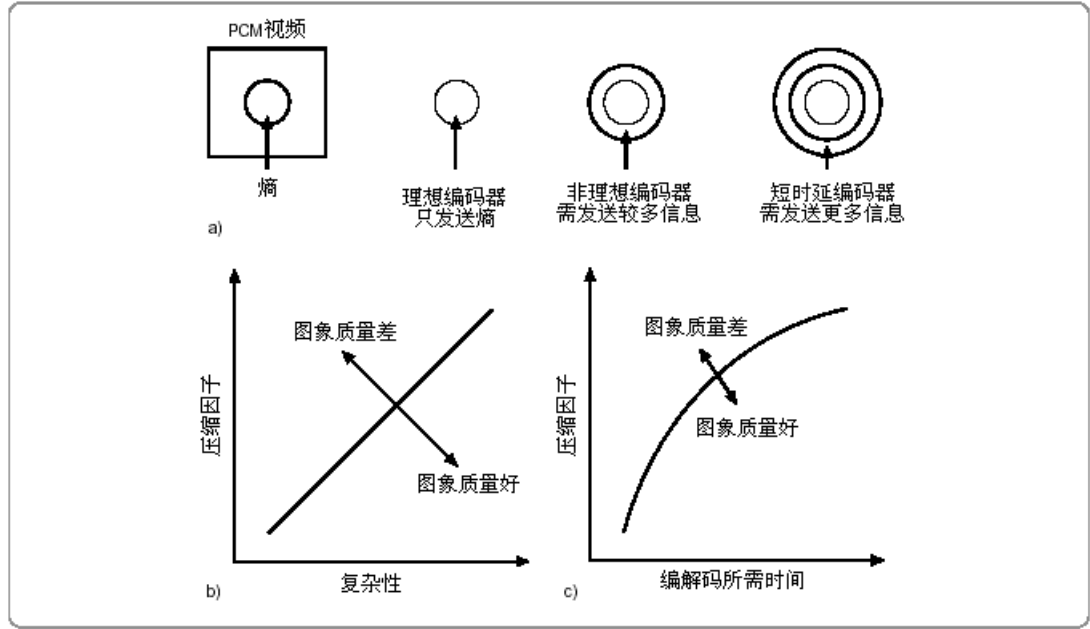


图 1—2

在实际应用中，需要各种各样的编码器，它们有着不同的时延，复杂程度也不一样。MPEG 之所以有着广泛的应用，就在于它不是一个单一的压缩格式，而是可以灵活地组合在一起的各种标准化的编码工具，这样可以满足各种应用需求。在被压缩的数据中，也包含了所采用的编码方式，这样，无论编码器使用什么编码方式，解码器都能自动地进行处理。

在 MPEG—2 和 MPEG—4 编码中，可以按照复杂程度的不同而划分为几类，每一类可以有不同的级，级与输入图象的分辨率相关。第四章将对类和级作详细介绍。

既然有许多不同的数字视频格式，也就有不同的比特率。例如，高清晰度系统的比特率是标准清晰度系统的6倍。仅仅了解编码器的输出比特率是不够的，更重要的是压缩因子，它是编码器的输入比特率与已压缩的比特率之比，例如压缩因子为2:1, 5:1等等。

然而，压缩过程中所包含的各种变化因素使得压缩因子的确定是一件非常困难的事。图 1—2a 表示的是一个理想的编码器，如果全部熵均被发送，那么图象质量很好。不过，如果为了减少编码器的输出比特率而提高压缩因子，将造成有一部分熵不能被发送，图象质量也会下降。注意，在压缩系统中如果出现图象质量下降，说明压缩过大（图 1—2b）。如果可用的比特率不够，较好的方法是降低输入图象的熵以避免这个区域。这可以用滤波器来实现。在主观上，将滤波器所造成的分辨率下降与压缩失真造成的图象损伤相比较，人们宁愿接受前者。

为了识别熵的完整性，就会使理想压缩器变得十分复杂。出于经济上的原因，实际应用的压缩器应当不太复杂但应能发送较多的数据，以保证所有的熵均被发送。图 1—2b 表示出了编码器的复杂度与压缩性能之间的关系。在保证一定的图象质量下，所需的压缩因子愈高，则编码器就愈是复杂。

视频信号中的熵是不断变化的。在记录播音员报导新闻的场景时，信息中有较大的冗余度，压缩也比较容易。与此不同的是，在对风中飘动的树叶或不停奔跑的足球运动员的场景

进行压缩记录时就要困难的多，这是因为此时信息的冗余度较低（信息量较大或熵较大）。在后两种情况下，如果有些熵未被发送，图象的质量就会降低。不过，我们可以在通道比特率保持不变但图象质量可变以及通道中的图象质量保持不变而比特率可变之间作出选择。电信网络工作人员在实际应用中倾向于保持传输比特率不变，可以使用缓冲存储器来平均熵的变化，只要所产生的延迟能被接受即可。但在记录时，可变比特率更便于处理，DVD就是使用可变比特率，它是通过缓存以平均比特率，使数据率保持在光盘系统的可接受范围之内。

帧内编码是利用空间性冗余度的一项技术，即冗余度就在该图象之内；帧间编码是利用时间性冗余度。帧内编码可以单独使用，如用于静止图象的 JPEG 标准，它也可以和帧间编码结合使用，如 MPEG 标准。

帧内编码利用了典型图象的两种特性。首先，并非所有的空间域频率均同时存在；其次，空间域频率愈高，则其幅度可能愈小。帧内编码需要对图象中的空间域频率进行分析，这种分析是便于变换，例如小波变换和 DCT（离散余弦变换）。变换所产生的系数是用于描述每个空间域频率的幅度。在一般情况下，许多系数为零，或者接近于零，这些系数可以忽略，从而可以减少数据率。

帧间编码是利用连续画面之间的相似性。如果某幅画面可被解码器使用，那么解码器只须利用图象差即可得到下一个画面。随着图象中目标的运动，图象差也会增加，但是这种变化可以通过运动补偿来获取，因为一般而言，从一个画面过渡到下一画面时运动目标不会有太大的变化。如果这种运动可以被测量，那么利用前一画面中目标移动到新位置的信息就可以得出与当前画面非常接近的目标位置。这种移动过程由发送至解码器的一对水平位移量和垂直位移量（它们统称为运动矢量）所决定。传送运动矢量的所需数据率要低于传送图象差的数据率。

MPEG 编码既可用于隔行扫描图象也可用于非隔行扫描图象。时间轴上的某点图象称为“画面”，而不管它是一场还是一帧。隔行扫描不是理想的数字压缩信号源，因为它本身就是一种压缩技术。时间域编码更为复杂，因为某一场中的像素会处于下一场的不同位置。

运动补偿能够减少但不能消除连续画面之间的差别。这种图象差本身是空间性的，可以采用上述基于帧内编码的变换编码。运动补偿只是简单地减少差值图象中的数据量。

采用较大的时间跨度可以提高时间域编码器的效率。图 1-2c 表明，如果需要更高的压缩因子，则在处理过程中必须有更长的时间跨度，但这样会产生较长的编码延迟。很明显，时间域编码信号在编辑时也较难于处理，因为某一给定输出图象的内容有可能是基于早先发送的图象数据。要使编辑得以正常进行，制作系统将不得不对时间域编码作出一定的限制，当然，对时间域编码的限制也就限制了压缩因子的可用范围。

1.6 音频压缩简介

PCM 数字音频通道的比特率大约只有每秒 1.5 兆比特，它大约为 4:2:2 数字视频的 0.5%。在适度的视频压缩方案中，如数字 Betacam，音频压缩是不必要的。但是，随着视频压缩因子的提升，音频压缩随之也日益受到关注。

音频压缩可利用它的两个特点。第一，在典型的音频信号中，并非所有的频率均同时存在。第二，由于存在着掩蔽效应，人耳不可能分辨出音频信号中的每个细节。音频压缩是通过滤波或变换将音频频谱分散到各个频带中，在描述较低电平的频带时只使用少量的数据。低电平频带中的掩蔽效应可以消除或降低特定频带的可听度，这样可以只需发送较少的数据。

由于人耳的听力较为灵敏，音频压缩不象视频压缩那样容易实现。当掩蔽和被掩蔽的声音信号在空间上重合时，掩蔽效应才能真正发挥作用。在单声记录的情况下总是空间重合，

但在立体声记录时却不是这样，在立体声记录中，如果低电平信号处于该记录声音的不同部分，仍然可以听见低电平信号。因此，在立体声和环绕声系统中，对于某一给定的音频质量，应当采用较低的压缩因子。音频压缩较为复杂的另一因素是不良扬声器的延迟，它实际上掩盖了压缩失真。使用不良扬声器来测试压缩器，则会给出不真实的结果，那种在测试时感到满意的信号，在用好设备测试时却并不令人满意。

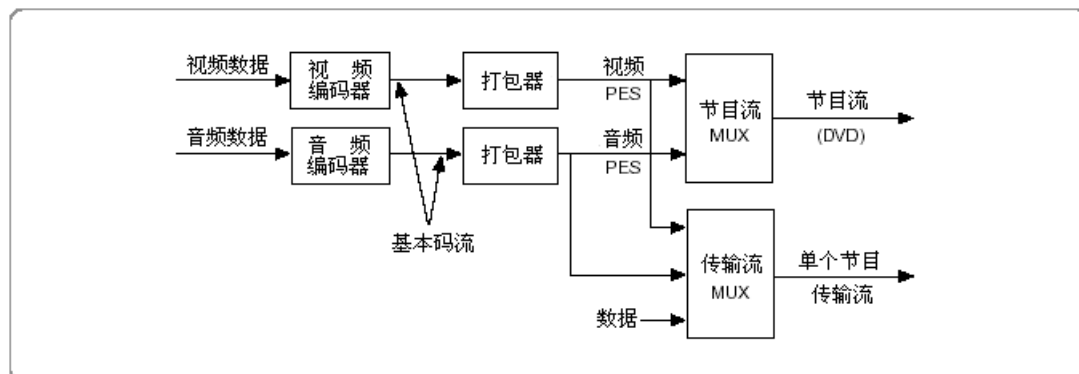


图 1-3

1.7 MPEG 码流

单个 MPEG 音频或视频编码器的输出是基本码流。基本码流是一种无限接近的实时信号。为方便使用，可以将基本码流分解为长度可以改变的数据包，形成打包的基本码流（PES）。为了识别数据包的起始处，数据包还需要有包头信息，并应包含有时间标记，因为打包破坏了时间轴的连续性。

图 1-3 表示的是一个视频 PES 和几个音频 PES 经复用后形成一个节目流，所有的编码器均锁相于一个共有的时钟。各 PES 的时间标记用于确保视音频间的唇音同步。节目流具有长度可变的数据包和包头。节目流用于光盘和硬盘间的数据传输，码流中基本上无误码，可用于任意大小数据量的数据文件。DVD 使用节目流。

为用于信号传输和数字广播，若干个节目及其相关的 PES 可以经复用而成为一个单独的传输流。传输流不同于节目流，在传输流中，PES 包可以进一步分解为长度较短的且字节数固定的数据包，也可以携带使用不同时钟的已编码的多个节目。这是可以实现的，因为传输流具有节目时钟基准（PCR）机制，它允许传送多个时钟，也可从中选择一个时钟并在解码器中再生。单个节目传输码流（SPTS）同样是可能的，它可存在于编码器和复用器之间。由于传输流能够使解码器时钟与编码器时钟相锁定，因此 SPTS 比节目码流更为通用。

传输流不只是音频和视频 PES 的复用，除了已压缩的音频、视频和数据外，传输流中还含有用于描述该比特流的亚数据。在亚数据中，包括有节目关联表（PAT），它列出了该传输流中的每一套节目。PAT 中的每一项均指向一个节目映像表（PMT），它列出了组成每套节目的各个基本码流。某些节目是开放式的，但是有一些节目是用于条件接收（加密）的，这些信息也包含在亚数据中。

传输流由长度固定的传输流包所组成，每个传输流包中含有 188 个字节。每个传输流包均携带有一个节目识别码（PID）。同一基本码流中的所有传输流包，均具有相同的 PID，这样解码器（或解复用器）就可以选择它所需要的基本码流，同时排斥其它基本码流。传输流包应能连续计数，这样，就能保证在被解码的码流中，其中的传输包均能被接收。为使解码器能正确识别每个包的起始处并能将比特流转换为码字，需要一个有效的同步系统。

1.8 监视和分析的必要性

MPEG 传输流有着非常复杂的结构,它通过链接表和包识别码来分离节目以及节目中的基本码流。在每个基本码流中,同样有着复杂的结构,需要用解码器来区分如矢量、系数和量化表等。

传输故障可以划分为两大类。在第一类中,传输系统能够正确地将信息由某一编码器/复用器传送至解码器,既没有比特错误也没有传输中的附加抖动,但是编码器/复用器或解码器却有故障。在第二类故障中,编码器/复用器和解码器均无问题,但数据从一端至另一端的传输过程中却有问题。如果要使问题得到迅速的解决,那么,判断故障是在编码器/复用器中,或是在传输过程中,或是在解码器中,这就显得十分重要了。

传输中的同步问题如同步的丢失或其结构的破坏,均可能影响到整个传输流的接收。传输流协议中的缺陷,可能会妨碍解码器去寻找节目中的所有数据,也许产生只传送声音而不传送图象的故障。如果传送的数据虽然正确但抖动太大,也会引起解码器的定时问题。

如以上所述,MPEG 传输系统失效,故障可能在编码器,也可能在复用器或解码器。那么,怎样来判断故障呢?首先,应当检查传输流是否符合 MPEG 编码标准。如果该码流不符合标准,就不能说是解码器的问题。如果码流符合标准,那就要怀疑解码器了。

传统的视频测试工具,如信号发生器、波形监视器和矢量仪,它们在分析 MPEG 系统时并不适用,除非是将它们用来检查 MPEG 系统的输入信号或输出信号是否正常。在 MPEG 系统中使用的测试工具应当是:例如一种可靠的、能产生有效 MPEG 测试信号的信号源,它是测试接收设备和解码器的基本仪器。还应当使用合适的分析仪来评估编码器、传输系统、复用器和再复用器的性能,这样才具有较高的置信度。作为视频行业高级测试设备的长期供应商,泰克公司能够适应技术的发展不断地推出新的测试和测量解决方案,从而使 MPEG 用户确信,即便是复杂的压缩系统也能实现正常运行,当系统出现问题时,能够迅速地予以解决。

1.9 压缩带来的问题

MPEG 压缩是有损压缩,因此,被解码的内容与原始的信号内容是有区别的。源信号的熵是变化的,当熵变高时,MPEG 系统在解码时可能会出现人眼可以发现的失真。在时间域压缩中,我们假设连续画面之间是有冗余的(时间域压缩是基于连续画面的相似性),如果实际情况不是这样,即它们之间没有冗余,系统就会出现失真。例如,在新闻发布会上,当闪光灯闪烁时,包含闪烁的单个画面完全不同于其相邻画面,这时,编码失真就会十分明显。

如果屏幕上出现无规则的运动,或者有几个不相关的运动对象,在进行编码处理时就需要大量的矢量带宽,为此,需要降低图象数据的可用带宽,才能满足这一要求。此外,所产生的可见失真产物,其电平也是变化的,并与运动相关。这种现象通常发生在体育节目中。

量化的精度不高会造成亮度轮廓失真和彩色斑点,看起来就象是有污点的影象或者大片彩色区上的块状。就人的主观感觉而言,压缩失真比由模拟电视传输系统所造成的相对固定的图象损伤更令人讨厌。

解决上述问题的唯一途径是降低压缩因子。这样,当人们在使用压缩技术时,就不得不在高压缩因子的经济性和它带来的失真度之间作出有价值的判断。

时间域编码除了会增加编码和解码的延迟之外,它还给图象的编辑造成了困难。事实上,MPEG 数据流是不能够任意编辑的。所以有这种限制,是因为在时间域编码中,对某一画面的解码可能需要早先的画面内容,而这早先的画面内容在编辑完成后可能就没法获得了。还有,在 MPEG 中,不一定是按照实际图象出现的先后顺序来发送图象,这将使编辑更为复杂。

如果采用合适的编码技术,还是可以编辑图象的。不过这只能在间隔较宽的接合处进行。

如果需要任意编辑，那么，只能让 MPEG 码流经历解码—修改—记录这一过程，当然，这会造成信号质量的下降。

第二章 视频压缩

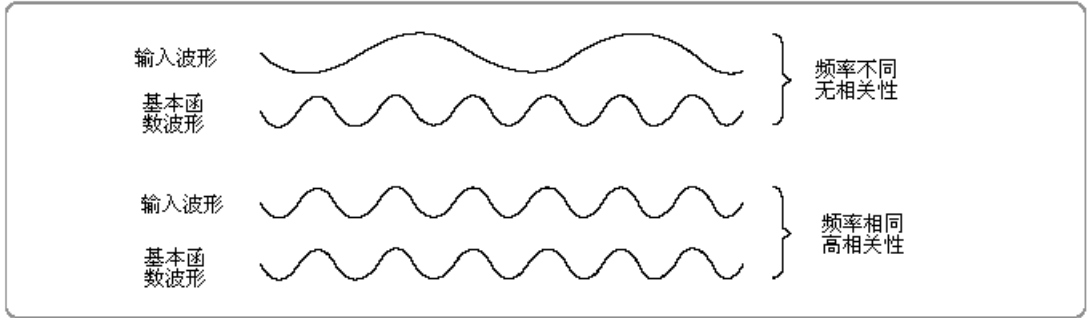


图 2—1

本章将介绍基于人眼感知的视频压缩技术。对于视频压缩中的重要部分，如变换和运动补偿，将作为 MPEG 编码器的结构予以介绍。

2.1 是空间域编码还是时间域编码？

正如第一章所述，视频压缩技术利用了图象的空间性冗余和时间性冗余的特点。在 MPEG 中，首先是利用连续画面之间的相似性来降低时间性冗余，具体作法是尽可能多地利用已经发送的图象信息以生成或“预测”当前图象。在使用这项技术时，只需要发送差值图象，就可以消除实际图象与预测图象之间的差异。而差值图象又需经过空间域压缩。从实用考虑，在解释时间压缩之前先解释空间域压缩要容易些。

空间域压缩是依赖于图象大块区域中相邻象素间的相似性和图案区的主要空间域频率。JPEG 系统只使用了空间域压缩，因为它为传输单个静止图象而设计的。不过，JPEG 也可对连续的单个图象编码以用于视频。不过，在所谓的“运动 JPEG”应用中，与时间域编码应用相比较，不会有那样高的压缩因子。尽管如此，由运动 JPEG 生成的比特流，却可以用作基于逐个画面的自由编辑。

2.2 空间域编码

空间域编码的第一步是通过变换以对空间域频率进行分析。所谓变换，简而言之，就是波形在不同域中的表示方法，在这里指的是频率域。频率变换的输出中含有一系列系数，这些系数用来描述当前有多少个频率。通过反变换就可将原始波形复原。只要在处理这些系数时具有足够高的精度，那么反变换后的输出就应与原始波形相同。

最著名的变换是付立叶变换。通过付立叶变换可以找出输入信号中的每一频率。在付立叶变换中，将输入波形（函数）与某一目标频率的样值（可称为基本函数）相乘，然后对该乘积求积分，这样就可找出每一个频率。由图 2—1 可见，当输入波形不包含目标频率时，则积分值为零；但如果输入波形中包含该频率，则所求得积分值就是描述该分量频率的幅度系数。

在输入信号中，如果频率分量与基本函数同相，则得出该频率的幅度系数；如果频率分量与基本函数正交，则积分值仍然为零。因此，有必要采用这样的方法：利用相互正交的基本函数，对输入信号的每一频率搜索两次，以检测输入信号中的每一频率的相位。

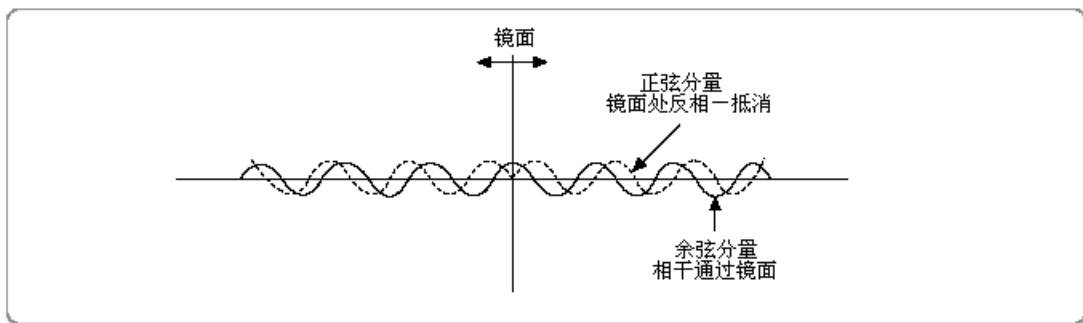


图 2-2

付立叶变换有一个缺点，它需要每一频率的正弦和余弦分量系数。在余弦变换中，输入波形在乘以基本函数之前与其自身在时间上是呈镜象的。如图 2-2 所示，镜面处的相干效应消除了所有的正弦分量而使所有的余弦分量加倍。因此，正弦基本函数是不必要的，每一频率只需要一个系数。

离散余弦变换（DCT）与余弦变换有着某些相似之处，它广泛用于 MPEG 的两维形式。它将图象中的一个 8×8 的像素块变换为一个 8×8 系数块。由于在变换为系数块的过程中，需要乘以分式，这就使得系数具有的字长要超过象素样值的字长。一般而言，一个 8 比特的象素样值会产生一个 11 比特的系数值。表面上看来，DCT 不具备任何压缩作用。但实际上恰相反，DCT 可以将信源象素转换为便于压缩的形式。

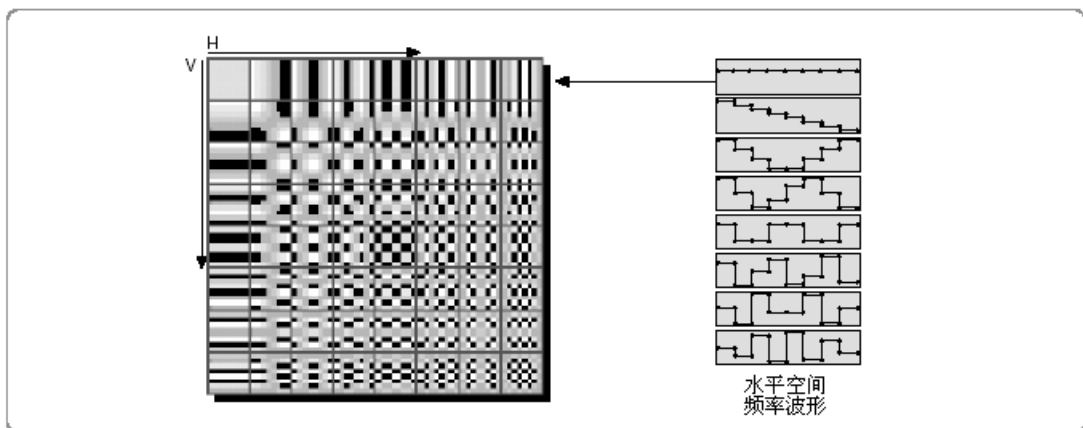


图 2-3

图 2-3 表示的是一个 8×8 DCT 块中各个系数的反变换结果。在亮度信号的情况下，左上方的系数是整个像素块的平均亮度或 DC 分量。沿着像素块的上方箭头移动时，水平空间频率增加；沿着左列箭头移动时，垂直空间频率增加。在实际图象中，会同时出现不同的垂直空间频率和水平空间频率，DCT 块中的某点系数代表着所有可能的水平频率和垂直频率的组合。

图 2-3 也表示了一维水平波形的 8 个系数。将这些波形与各个不同的系数幅度及任一极性组合在一起，就能再现 8 个象素的任意组合。因此，2 维 DCT 的 64 个系数的组合，可以产生原始的 8×8 象素块。很明显，对于彩色图象，也需要对色差样值进行处理。将 Y、Cb、Cr 组合成单独的 8×8 阵列，然后各自独立地进行 DCT 变换。

在大量的实际节目素材中，有许多 DCT 系数均为零或接近零值，因此，这些系数可不必传送。基于这一事实，在进行压缩的情况下也不会造成实际图象质量的损伤。如果需要更

高的压缩因子，还可以减少非零系数的字长。不过，字长的缩短会降低这些系数的精度，从而带来处理过程中的损耗。如果仔细观察的话，也许能够看出损耗对图象质量所产生的影响。

2.3 加权

图 2—4 表明，人眼对图象噪声的感受是空间频率的函数而并非与频率无关。人们也许能够忍受较高空间频率处的噪声。还有，图象中的精密细节会有效地掩蔽视频噪声，但在大片图象区的噪声就十分明显了。读者可能知道，在传统的噪声测量中要按频率进行加权处理，其目的就是使噪声的测量结果更接近于人眼的主观感受。

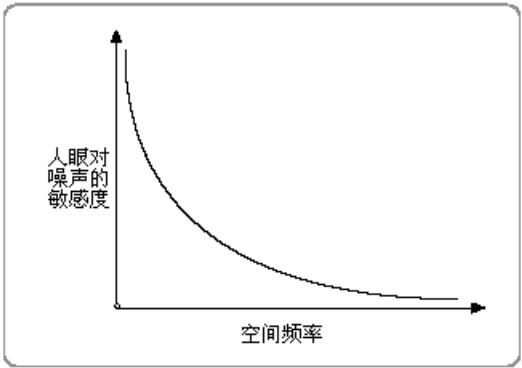


图 2—4

变换编码中的压缩降低了 DCT 系数的精度，它与 PCM 中的缩短样值字长有类似的效果，这就导致了噪声的提升。不过，在 PCM 中，因字长的缩短而使噪声在所有频率上都得到相同的提升。既然 DCT 是将信号分解在不同的频率上，当然它也能控制噪声的频谱。实际的作法是，通过加权处理，能够有效地使低频系数的精度要优于高频系数的精度。

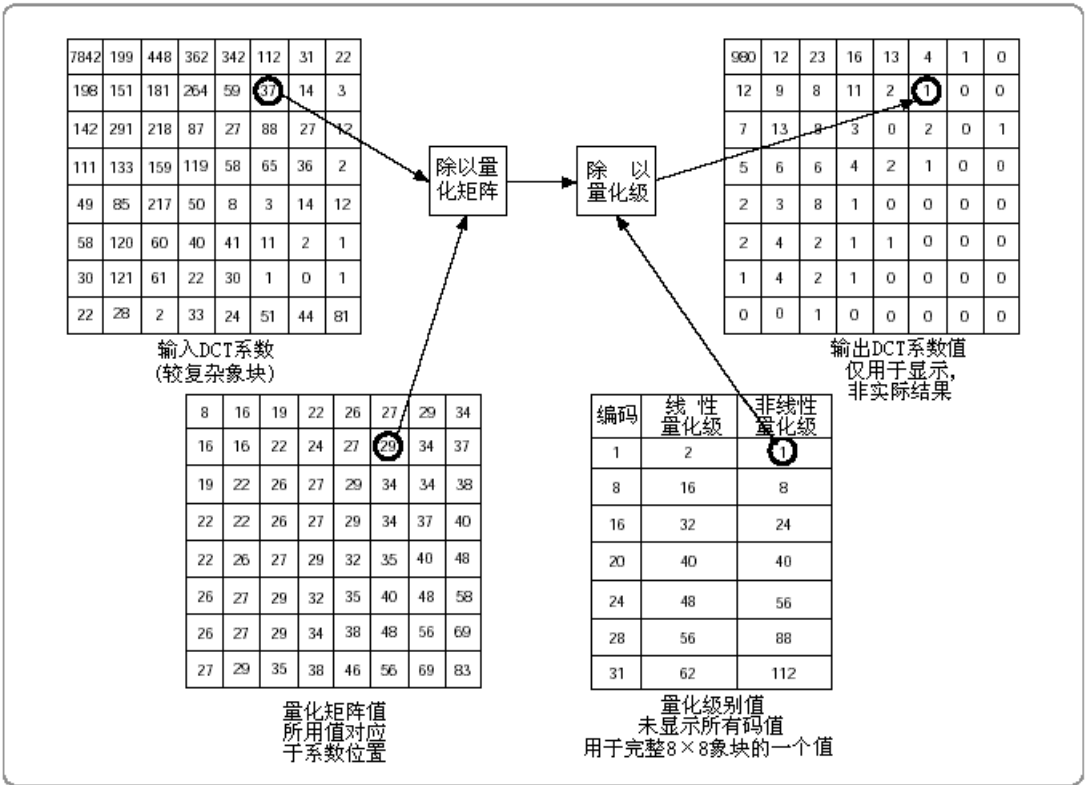


图 2—5

图 2—5 所示为一加权处理过程，DCT 系数除以一个因子，该因子是两维频率的函数，即因子的取值决定于频率的高低。DCT 的低频系数除以小因子，高频系数除以大因子。相除后的结果再去掉尾数，取其最接近的整数。这种舍位（truncation）过程是再量化的一种形式。如果不加权，这种再量化处理将均匀地增加量化步长；如果施以加权，将按照除数因子非均匀地增加量化步长。

这样，代表较低空间频率的 DCT 系数用相对较小的步长来量化，噪声只有少量增加；代表较高空间频率的系数用较大的步长来量化，相应噪声增加较多。较小的量化步长意味着只需用较少的比特位来区分该步长，由此就得到了压缩。

在解码器中，在低位比特处添加零以使已加权的系数恢复到其正确数值。然而这些系数应乘以加权因子的倒数。很明显，高频处的乘积因子较大，因此量化噪声也较大。在去加权处理后，系数将具有其原始的 DCT 输出值和再量化的误差，较高频处的再量化误差也大于低频处的再量化误差。

除了舍位法之外，也可对已加权系数采用非线性再量化处理，使量化步长的大小随系数的大小而增加。采用这种技术可以得到更高的压缩因子，不过失真度也会相应增加。

十分明显，数据的压缩程度，或者说数据输出比特率，决定于再量化的处理过程。不同的比特率需要不同的加权表。在 MPEG 中，有可能使用各种不同的加权表，应将所使用的加权表传送至解码器，以保证解码的正确性。

2.4 扫描读取

在一般节目素材中，最重要的 DCT 系数通常位于变换矩阵的左上角附近。加权后，低值系数再经舍位处理后有可能为零。如果首先传送的是所有非零的系数，而后再以一个代码来表示剩余的全部零值，那么传输效率将会得到显著的提高。为此可采用一种特别的扫描读取的方式，它是按照高值系数出现几率大小的顺序来先后发送系数的。图 2—6a 表示在一非隔行扫描系统中，出现高值系数的几率在左上角为最大，而在右下角的几率为最低。在这里，最好的读出顺序就是沿着 45 度对角线进行之字形扫描读取。

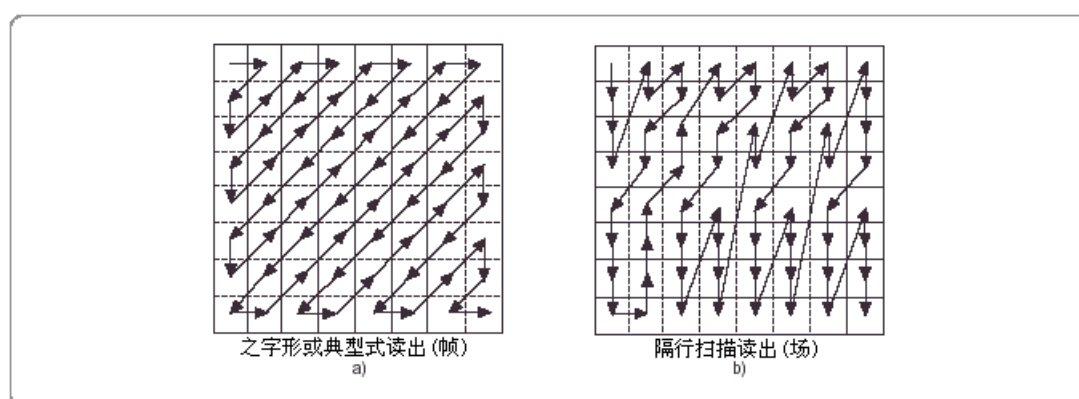


图 2—6

图 2—6b 是另一种扫描读取图示，它可用于隔行扫描系统。在隔行扫描信号源中，一场图象中的一个 8×8 DCT 块延伸至两倍屏幕区，这样，对于给定的图象细节，出现的垂直频率是水平频率的两倍。因此，隔行扫描图象中的理想扫描读取方式也是沿着对角线进行的，但对角线的倾斜度应当是图 2—6a 的两倍。图 2—6b 表示对于一个给定的垂直空间频率，在

扫描相同水平空间频率之前所进行的扫描方式。

2.5 熵编码

在实际视频中，并非所有的空间频率均同时存在。因此，在 DCT 系数矩阵中会存在零值项。在再量化的过程中，去掉了一些小系数值，增加了零值数。尽管使用了扫描读取方式，但在各有效系数之间仍然会出现零系数。对于这些零系数，如果采用游程编码（RLC），就能有效地进行处理。当矩阵中存在着重复值时（如一串零值），RLC 就只传送零值的个数而不是逐个传送单个比特。

在研究实际视频中特定系数值的出现概率后发现，在实用中，某些系数值会经常出现，而某些系数值则不常出现。对于这样的统计信息，如果使用可变字长编码（VLC），则可使码率得到进一步的压缩。经常出现的数值转换为短码字，不常出现的数值转换为长码字。为便于解码，任一码字不能是另一码字的前缀。

2.6 空间域编码器

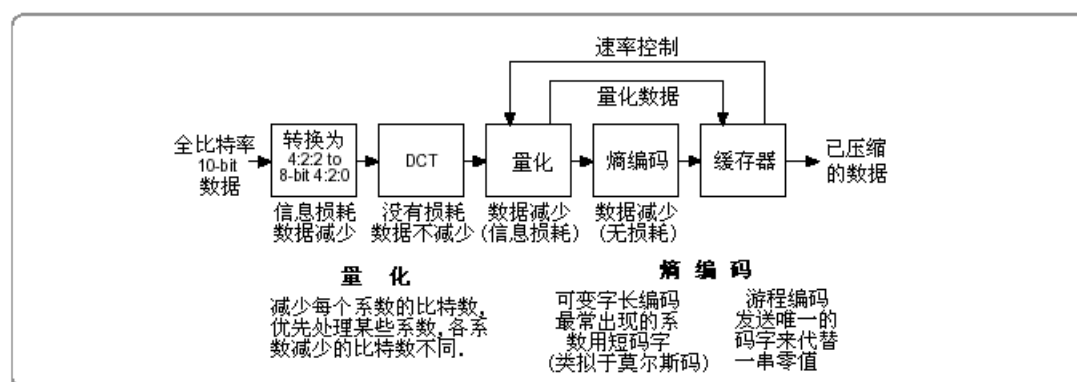


图2.7将前面述及的所有空间域编码方法绘制在一起。假定输入视频信号是4:2:2 SDI（串行数字接口）格式，它可以是8比特或10比特的字长。MPEG只使用8比特分辨率，因此需要对10比特的SDI信号进行舍位处理。大多数MPEG类别使用4:2:0取样，还需要通过低通滤波器 / 内插级处理。在舍位和彩色取样过程中会引入少量的不可恢复的信息损耗，数据率也相应按比例地有所减少。此外，还需保存光栅扫描输入格式，以便于转换为8×8像素块。

在DCT中，它将图象信息变换为频率域信息。DCT本身并不能进行任何压缩。在DCT之后，各系数经加权和舍位处理，这是首次有效压缩。然后利用之字形扫描读取加权后的系数，以增加首先读出有效系数的概率。在读出系数块中最后一个非零系数之后，产生EOB（块结束）码，作为该块系数的输出结束。

此后，再通过游程编码和可变字长编码对系数再次进行压缩。在可变比特率系统中，量化可以是固定的，但在固定比特率的系统中，需要使用缓冲存储器来吸收数据编码过程中的异常变化。高细节图象有可能将缓存器填满，而简单图象则允许它是空的。如果缓存器接近于溢出时，这时不得不加大量化步长，提高压缩因子。

在解码器中，对所接收的比特率进行串并转换，对熵编码进行逆向处理以再现已加权的DCT系数。该系数按照之字形扫描读取方式存放在DCT矩阵中，经去加权处理以再生DCT系数块。而后经DCT反变换，使8×8像素块得以恢复。为了取得光栅扫描输出，将像素块储存在RAM中，每次读出一行即可。为了从4:2:0数据中获取4:2:2输出，则需要按照图2—8进

行垂直内插处理。

在4:2:0格式中，色度样值位于垂直轴中相邻两行亮度样值的中间。这样，在使用隔行扫描信号源时，色度和亮度样值都是均匀分布的。

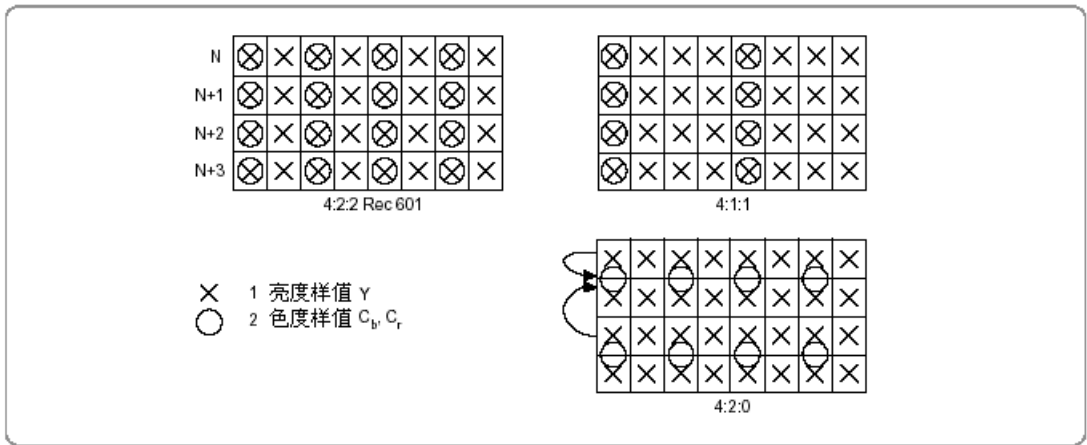


图2—8

2.7 时间域编码

在帧间编码中，仅传送图象间的差值是利用了图象的时间性冗余。图2—9是使用减法器并将图象延迟从而计算出图象差值的示意图。图象差值是从减法器右边输出的图象，然后按照前面所介绍的内容利用空间域编码器对其进行压缩编码。在解码器一端，用反空间域编码得出差值图象，并将差值图象与前一图象相加，这样就得到下一图象。

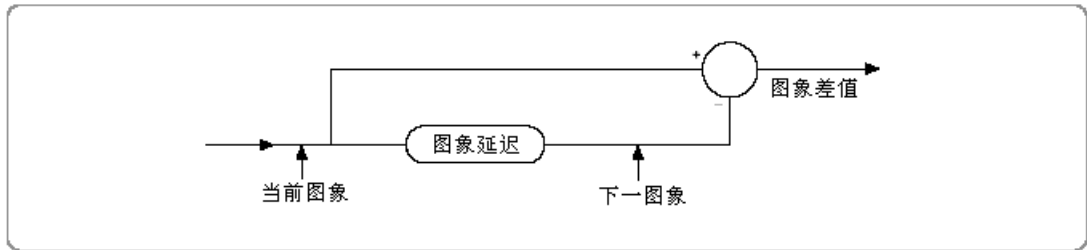


图2—9

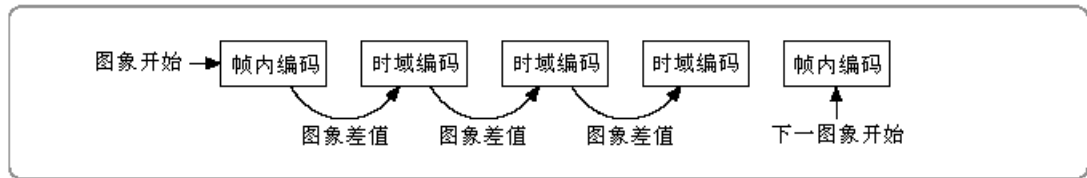


图2—10

这种简单的编码系统有几个缺点。首先，如果只发送图象差值，在传输开始后立即就解码（出完整的图象）是不可能的。这就使得解码器在某一比特流切换为另一比特流（例如观察者改变频道时）之后难于立即给出图象。其次，如果差值数据中的任何一部分有错误，那么这将在图象中造成误码的无限积累。

为了解决上述问题,可以使用改进后的系统。在图2-10中,系统周期性地发送完整的图象,这些完整的图象称为帧内编码图象(或称I图象),它们是只经过空间域压缩而获取的图象。如果有误码发生或者切换了频道,那么在下一I图象上就能恢复正确的解码而不会造成误码的扩散或积累。

2.8 运动补偿

图象中运动的存在降低了图象间的相似性,增加了生成差值图象的所需数据量。而运动补偿正是用来增加图象间的相似性。图2-11为运动补偿的示意图。当某一对象在电视机屏幕中运动时,它在每帧图象中可能处于不同的位置,但就其外观特性而言不会有太大的改变。在编码器中通过对运动的测量,就可以减小图象间的差值。将图象差值作为矢量发送给解码器。解码器使用这个矢量来移动前一图象中的对象,使其在新的图象中处于合适的位置。

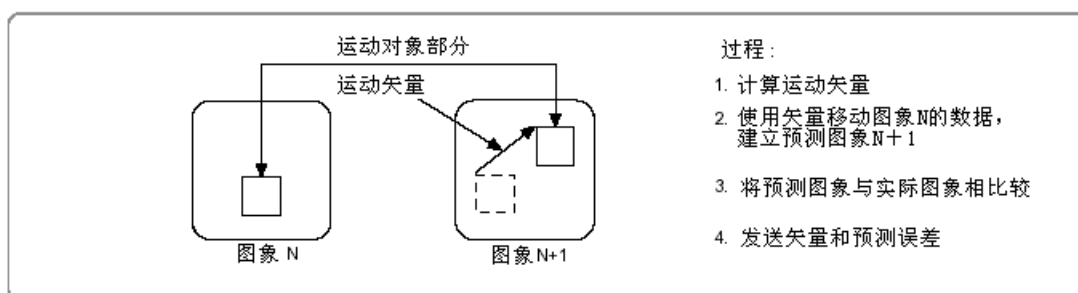


图2-11

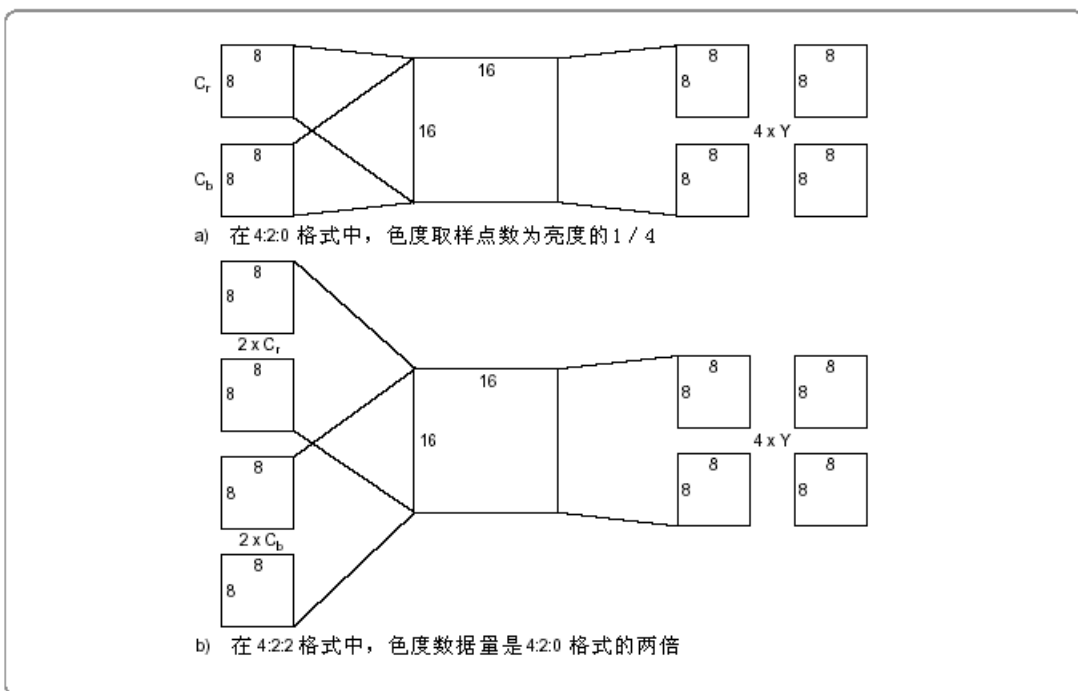


图2-12

一个运动矢量可以控制被称为宏块的整个图象区域的移动。宏块的大小由DCT编码和色度[亚](#)取样的结构所决定。图2-12a表示在4:2:0系统中，色度样值的垂直和水平间隔恰为亮度样值间隔的两倍。一个单独的 8×8 色度样值块，与4个 8×8 亮度块占据同样的区域：这样，

它就是一个运动矢量可以移动的最小图象区。一个4:2:0宏块含有4个亮度块，一个Cb块和一个Cr块。

在4:2:2格式中，色度仅在水平轴上被亚取样。图2—12b表示在4:2:2格式中，一个单独的 8×8 DCT色度样值块占据的区域与两个亮度块相同。一个4:2:2宏块包含有4个亮度块，两个Cb块和两个Cr块。

运动估值用于比较两个连续画面的亮度数据。并将第一个画面中的宏块作为参考。在整个搜索范围内，以半像素的精度在所有可能的位移上测量出参考图象与下一图象间的相关性。找出最大相关性，并用它来代表正确的运动。

运动矢量具有垂直分量和水平分量。在典型的节目内容中，一个对象的运动可以包含许多宏块。如果只是矢量差值被传送，可以得到较大的压缩因子。当图象中一个大的对象运动时，如果邻近宏块具有相同的矢量，则矢量差为零。

运动矢量是与宏块相关联的，而与图象中的实际对象没有关联。但有时宏块中的某一部分运动而另一部分不运动，这时就不可能得到正确的补偿。如果用传送矢量对宏块的运动部分进行补偿，则宏块中的静止部分将会有不正确的位移，这就需要用差值数据来校正静止部分。如果不发送矢量，那么静止部分是正确的，但需要用差值数据来校正运动部分。实际应用的压缩器可能会尝试这两种方式，并从中选择一种所需数据量较少的方式。

2.9 双向编码

图象中的某一对象运动时，其前沿处的背景被遮挡，而其前沿处的背景则被露出。露出的背景则需要传输新的数据，因为该背景区在前一图象中是被遮挡的，因此不可能从前一图象中取得这种新数据。同样地，摄像机在摇镜头时，新的区域进入视区，该区域内的信息在此前是不知道的。MPEG是使用双向编码来解决这一问题的，在双向编码中，它既可以从当前图象的前一图象中获取信息，也可以从当前图象的后一图象中获取信息。背景在露出过程中，它会在后一图象中出现，则可以使信息作后向运动并及时地创建前一图象的相应部分。

图2—13是双向编码的示意图。它是以一个单独的宏块为基础，双向编码图象可以从前一图象中或者从后一图象中获取运动补偿数据，甚至还可使用前向数据和后向数据的平均值。双向编码明显地降低了因改善预测效果而所需要的差值数据量。MPEG并没有对编码器的构建作出规定，而只是制定了顺应性比特流。不过，一种智能性的解码器应能尝试所有的三种编码方式，并从中选择一种具有较少的数据传输的方式。

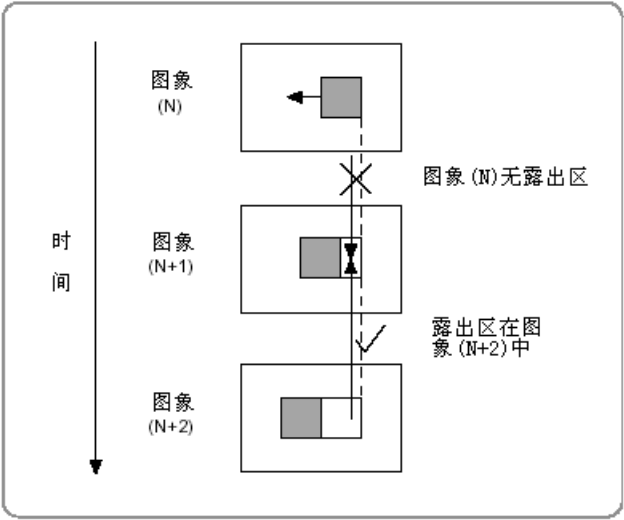


图2—13

2.10 IPB图像帧

在MPEG中，需要三种不同类型的图像以支持差值编码和双向编码，从而使误差的积累减至最小。

I图像即帧内编码图像，在编码时它不需要其它帧的参考信息。与另外两种图像类型相比较，I图像需要大量的数据，因此它不常传送而仅在必要时予以传送。I图像主要由变换系数组成，它不含运动矢量。它在解码时也不需要参考其它图像，因此它允许观察者在I图像时切换频道。I图像用以阻止误差的积累或扩散。

P图像是来自前一图像的前向预测图像，前一图像可以是I图像或P图像。P图像数据则是由前一图像中用以描述每个宏块的矢量和宏块变换系数所组成，后者是用来描述该宏块校正数据和差值数据的变换系数。如果通过运动补偿搜索没有发现合适的匹配宏块，则通过发送的帧内数据对该宏块进行编码。P图像所需要的数据量大致为I图像的一半。

B图像是来自先前的和（或）随后的I或P图像而作出的双向预测图像。B图像数据是由描述先前或随后的图像数据的矢量所组成。它同时含有帧内编码数据以提供必要的校正。此外，当用运动补偿搜索没有发现合适的匹配宏块时，可以使用发送的帧内数据来对该宏块进行编码。双向预测是十分有效的，因此B图像中的大多数宏块主要使用运动矢量来编码。还有，B图像不可作为其它图像的编码参考，这样就消除了误差积累的可能性。但它允许编码器使用更主动的再量化以校正数据。典型的B图像需要的数据量大约为I图像的1 / 4。

需要注意的是B图像并非必须使用来自两个方向的预测，在某些情况下，它可以只使用一个方向的预测。这在构成闭合式图像组（GOP）时是可以选择的。

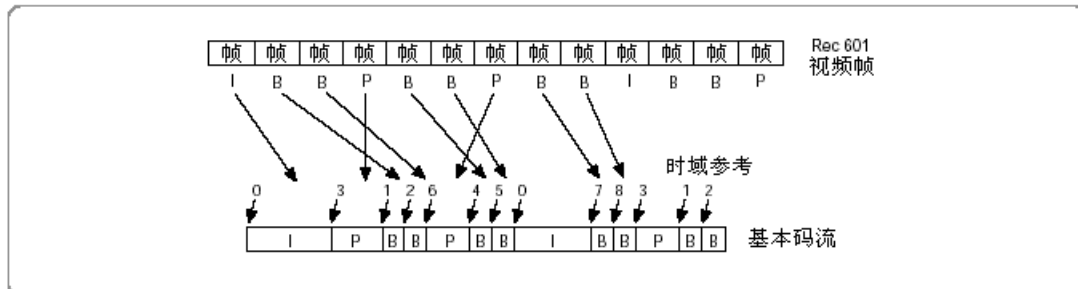


图2—14

图2—14是GOP的示意图。GOP代表着该序列中I、P和B图像。一般而言，GOP结构是通过其序列重复的，但GOP长度和结构则可以随时改变。对GOP的长度并没有明确的限定，但作为传输使用，其长度一般为12或15帧图像。

MPEG时间压缩的特性就意味着图像的传输顺序与图像的显示顺序并不一致。P图像自然要在I图像之后或作为预测参考的P图像之后，这里并没有特殊的要求。不过，在对双向编码的B图像解码之前，必须首先接收和解码它的两个参考图像。在图2—14中，上方是GOP图像的显示顺序，下方是GOP图像的传输顺序。须注意的是，在传输顺序中，B图像总是在它的两个预测参考图像之后。

有两种类型的GOP，即开放式图像组和闭合式图像组。闭合式GOP在其自身外无需参考图像。在显示顺序中，闭合式GOP是以I图像开始而以P图像结束。在传输顺序中，在最后一个P图像之后，通常还有两个B图像。然而在显示时，这两个B图像却在最后一个P图像之前。

在显示顺序中，也可以用B图像作为闭合GOP的起始图像或者为结束图像，但是，在这

种情况下的作为起始或结束的B图象只能使用单向预测编码。位于闭合GOP起始处的B图象只能使用后向预测。而位于闭合GOP结束处的B图象只能使用前向预测——这点与P图象相似，但B图象却可用于再量化等。

开放式GOP在预测矢量上却没有上述限制。例如，GOP结束处的B图象既可以使用以最后一个P图象为参考的前向预测，也可以使用下一个GOP的第一个I图象的后向预测。这种开放式结构的传输效率要略高一些，但是这种预测却越过了图象边界。也使得比特流的接合更为困难，而且在变换频道时可能更易引起图象错误。

在图象场景发生变化时，编码器也可能会改变GOP结构。越过场景变更的预测通常会失效，这是因为位于场景变化两侧处的相邻两帧图象之间存在着较大的熵。编码器可以检测到场景的改变，通过导入场景变化的闭合GOP，并用代表下一场景的第一帧图象即I图象来启动新的GOP（开放式或闭合式）。

如果不按时间先后顺序发送图象数据，那么编码器和解码器均需要额外的存储器，这当然会导致延迟。这时，位于帧内预测图象和前向预测图象之间的双向预测编码图象的数量应当受到限制，一方面是为了降低成本，同时也使传输时的延迟减至最小。

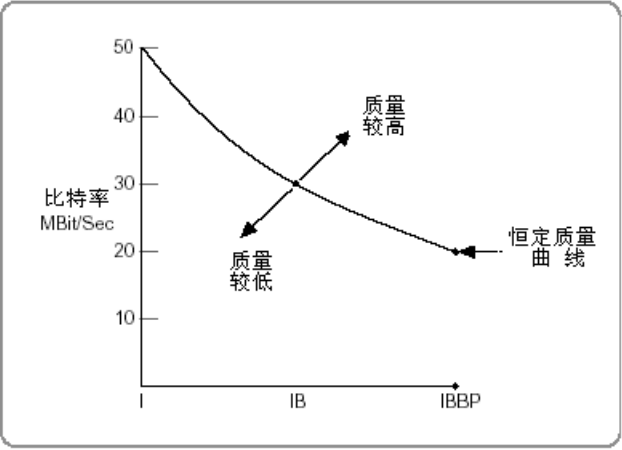


图2-15

在图2-15中，表示在压缩因子和编码延迟之间应当作出的折衷。对于给定的图象质量，只传送I图象的所需比特率是传送IBBP图象序列的两倍多。

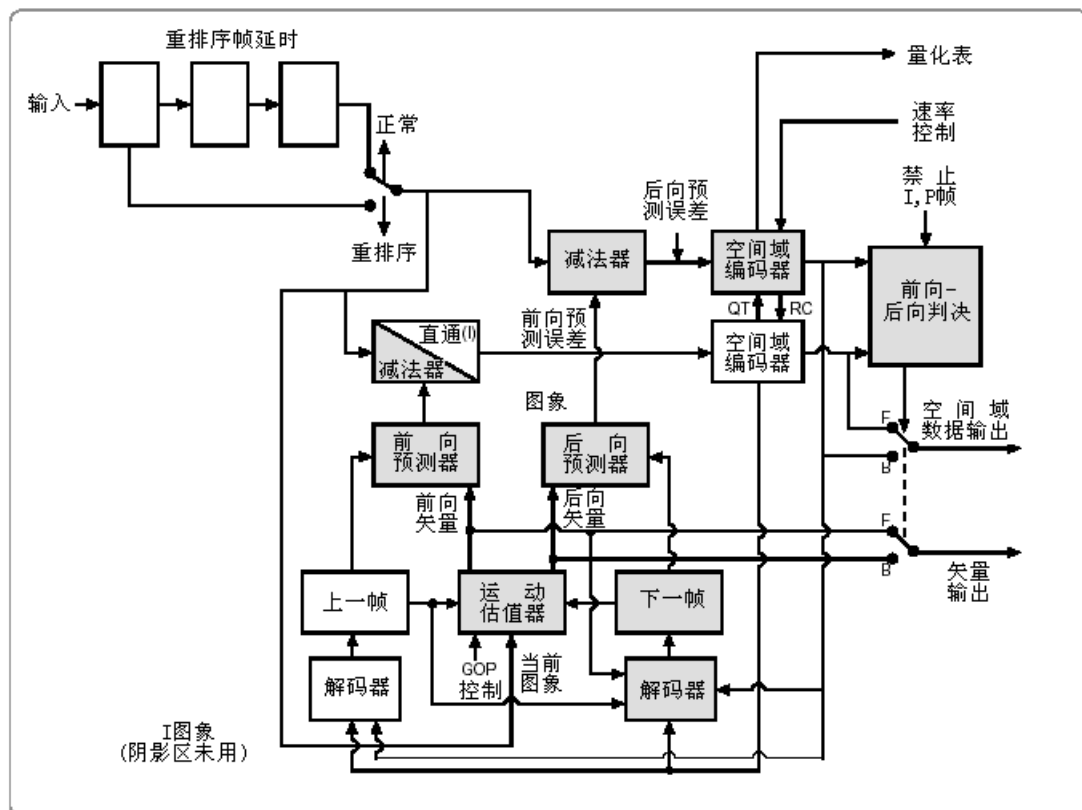


图2-16a

2.11 MPEG 压缩器

图2-16a, b和c给出的是一种典型的双向运动补偿器结构。经过预处理后的输入视频进入一串帧存储器，经旁通以改变图象的先后顺序。重排序后的图象数据进入减法和运动估算法器。为创建I图象，可选取经输入延时后的图象，关闭减法器，使数据直接进入空间域编码器（参见图2-16a）。减法器的输出数据也转入帧存储器，帧存储器里可以保存好几帧图象。这样，I图象就保存在帧存储器中。

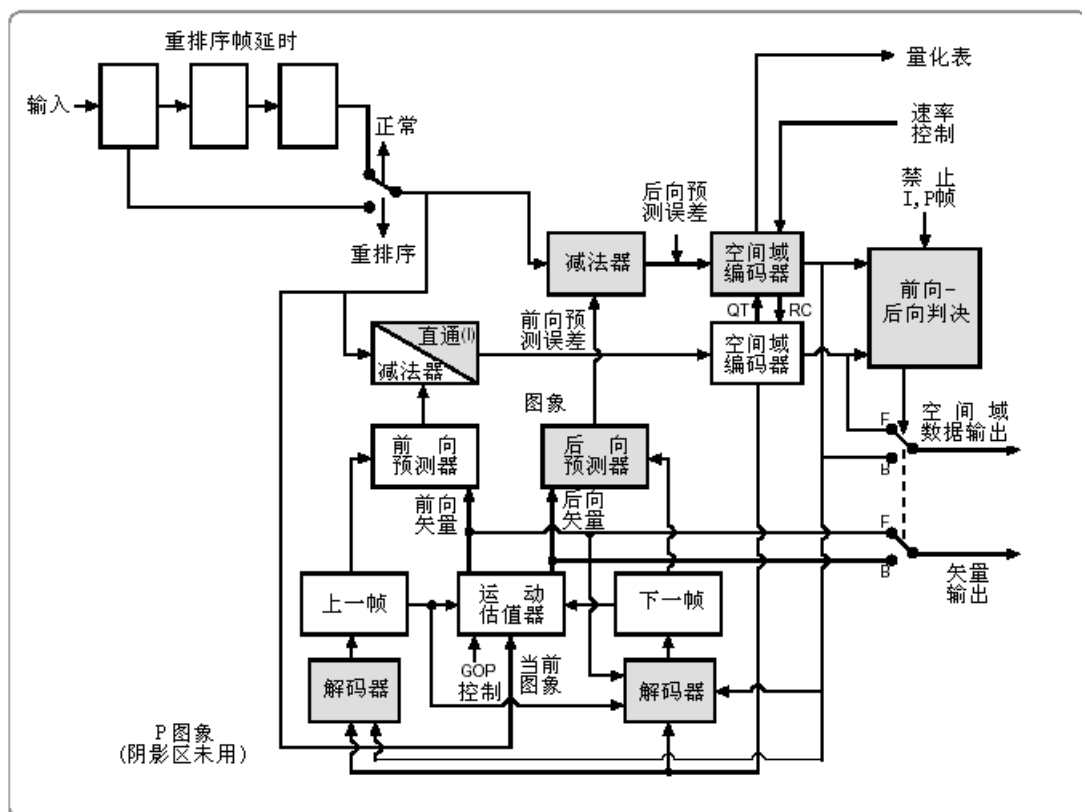


图2—16b

为对P图象编码，可在输入缓存器中将B图象旁路，选取下一图象即P图象（参见图2—16b）。在运动估值器中，将输出存储器中的I图象与输入存储器中的P图象相比较，产生前向运动矢量。利用运动矢量使I图象中的宏块移位以建立预测P图象。再从实际P图象中减去这个预测P图象即可得出预测误差，预测误差经空间域编码后与矢量一起发送。同时将预测误差与预测P图象相加以创建本机已解码的P图象，并将其放入输出存储器中。

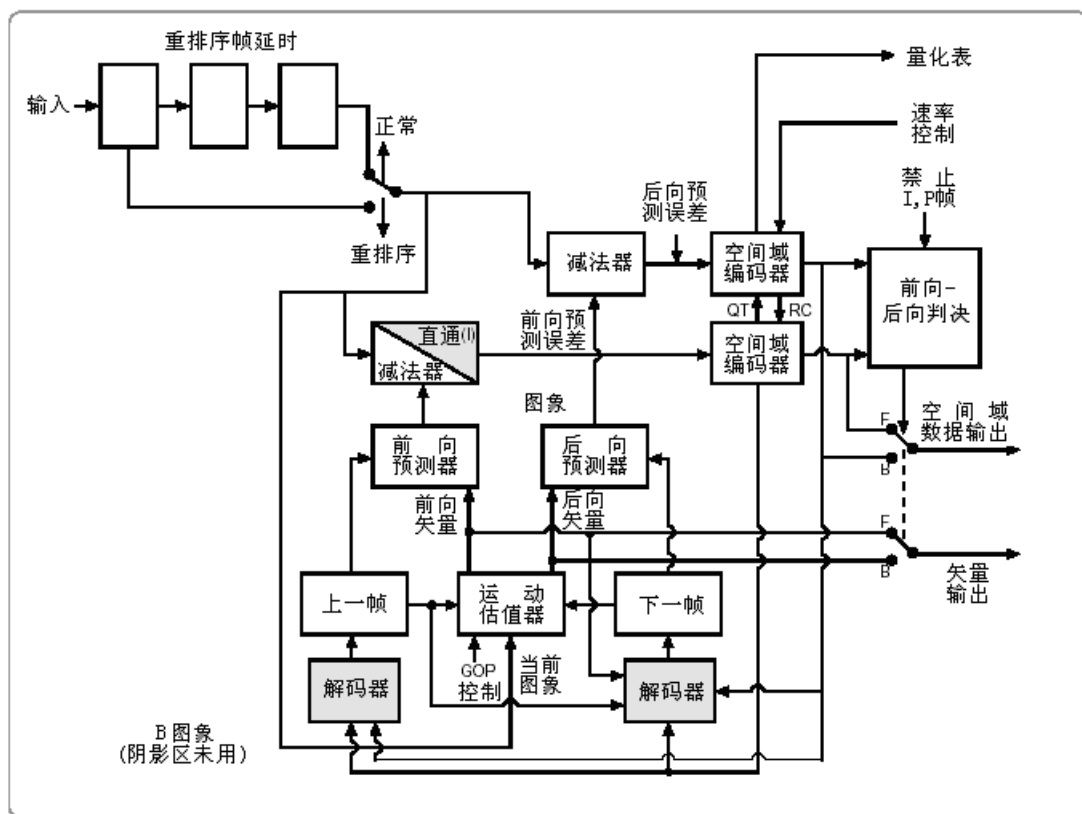


图2-16c

经过上述处理后，在输出存储器里已保存有I图象和P图象。现在可以从输入缓存器中选取B图象了。运动矢量补偿器将选取的B图象与它前一帧的I图象和它后一帧的P图象进行比较，即可获取双向矢量（参见图2-16c），在进行前向和后向运动补偿后，产生了两个预测B图象。然后从当前B图象中分别减去这两个预测B图象即可得出前向和后向数据差。再以宏块为基础，逐个进行比较，以选取代表最小差值的前向或后向数据。这两个图象差值数据经空间域编码后随同矢量一起发送。

在对I、P图象之间的所有B图象进行编码时，输入端的存储器再次被旁通，以前一P图象为参考，创建新的P图象。

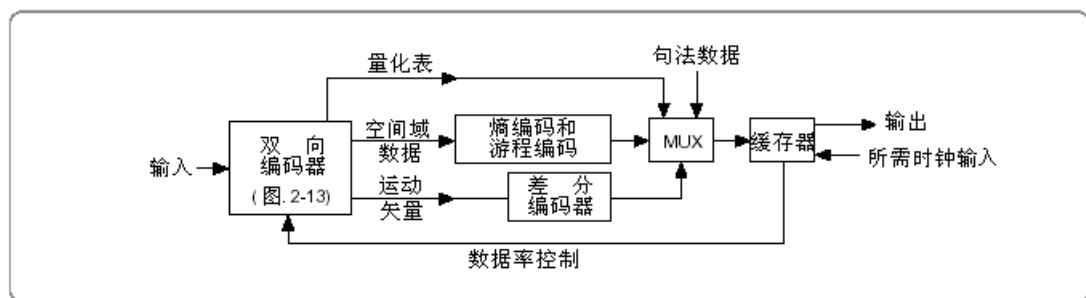


图2-17

图2-17所示为MPEG编码器。运动补偿输出经空间域编码后与运动矢量一起在复用器

中相加。句法数据也送入复用器相加。句法数据用于图象类型的识别（I，P或B图象），它还可以提供解码器所需的其它信息（参见第五章——基本码流）。数据经缓存后输出，从而允许比特率的时间域变动。如果平均比特率太高，则缓存器有可能被填满。为避免缓存器溢出，不得不对量化加以严格的限制。同样地，如果缓存器出现下溢，则量化可以放宽以保持适当的平均比特率。

2.12 预处理

压缩器的作用是尽量降低图象中或图象之间的冗余度。那种使冗余度降低但本身却不是图象信息的任何内容，都不是我们所希望的。噪声和胶片颗粒更带来一些问题，因为它们通常存在于整个图象区。在经过DCT处理后，噪声会导致更多的非零系数，编码器无法判断它们是否来自图象本身的信息。同时，对这些系数的编码也加重了量化的负担，降低了图象的质量。噪声也降低了连续画面之间的相似性，增加了所需的差值数据。

在复合视频信号的解码输出视频中，其中的剩余副载波是一个严重的问题，因为它会产生较高的空间频率，这些频率在分量节目中通常表现为较低的电平。副载波在相邻图象之间的交替相位也会导致差值数据的增加。当然，在送入MPEG的信号中，任何可见的复合视频解码失真都有可能在MPEG解码器中再次出现。

在实际操作过程中，应力求避免不必要的运动。例如，摄像机固定不牢，除了会使图象晃动外，还会增加图象差值和矢量传输的数据量。这种现象也会在电视电影中发生，例如，胶片定位孔损坏使得电影摇晃或抖动。总之，用来压缩的视频应当具有尽可能高的信号质量。如果高质量的视频不能实现，也要尽可能地降低噪声和采用其它稳定技术。

如果压缩因子较高，那么图象的失真度会增加，当输入信号质量较差时更是如此。在这种情形下，最好的办法也许是使用预滤波来降低编码信号的熵。视频信号经过两维低通滤波后，可以降低所需系数的数量并减少失真度。当然，这会损失一些清晰度，但宁愿要较低的清晰度，也不要较大的压缩失真。

在大多数MPEG-2应用中使用的是4:2:0取样格式。这样，当源信号为4:2:2时，就需要对色度进行下取样（downsampling）处理。在MPEG-1中，对亮度和色度采用了下取样处理后，产生的输入图象即CIF（通用图象格式）图象在水平方向上仅有352个像素。这项技术使熵减少一倍以上。如果需要更高的压缩，可以使用QCIF（四分之一通用图象格式）图象，它在水平方向上仅有176个像素。下取样的过程是将空间域低通滤波器与内插器结合使用。对隔行扫描信号进行下取样处理是不合适的，因为隔行扫描信号的垂直细节分散在两场，这样处理可能会影响运动相关度。

如果信源素材是电视电影，这样的视频信号具有与正常视频不一样的特性。在50Hz视频中，成对的两场图象代表同一电影帧图象，在它们之间不存在运动。这样，两场之间的运动或者是零（即没有运动），或者是两帧之间的运动，总之二者相互交替。在60Hz视频中，通过3:2下拉以实现从24Hz电影到60Hz视频的转换。一帧分为两场；下一帧分为三场，如此类推。

然而，五场中的一场完全是多余的。MPEG在处理电影素材时，取消了3:2系统中的第三场。传输中的24Hz代码用以提示解码器重读一场存储信号来再生3:2序列。在50Hz和60Hz电视电影中，取消成对场中的交织以创建帧，这样只在帧与帧之间进行运动测量。通过在帧存储器中读取交替行的方式，解码器就能再次建立隔行扫描格式。

对于压缩器而言，剪辑是难于处理的一件事，因为它通常会造成预测的几乎完全失效，并且需要大量的校正数据。如果能容忍某种编码延迟，在剪辑时，编码器可以提前对剪辑进行检测并对图象组结构进行动态修改，插入I图象并与剪辑保持一致。在这种情况下，剪辑只需处理极少量的额外数据。在I帧之前的B图象几乎肯定需要使用前向预测。在某些非实时

应用中，如DVD制作，编码器可以在输入视频下提取两次：一次用于识别较困难的或较高的熵区并建立编码方式，另一次则对输入视频进行实际压缩。

2.13 小波变换

所有的变换都具有不确定性，这是因为对频率域分辨得愈准确，那么对时间域分辨得就愈不准确（反之亦然）。在大多数变换中，如在离散付立叶变换(DFT)和离散余弦变换(DCT)中，块的长度是固定的，因此时间和频率的分辨率也是固定的。频率系数是代表线性刻度中均匀空间的数值。然而，人类的感官却是呈对数性的，DFT和DCT的平均分度所给出的频率分辨率，则会在某一端显得不足，而在另一端又会超出。

小波变换则不存在这样的问题，因为它的频率分辨率是某一倍频程的固定的一小部分，这就使它具有对数的特性。在进行变换时，是将块长作为频率的函数来改变的。如果频率下降，则块长加大。这就是说，小波变换的特性是使所有的基本函数均含有相同的周期数，这些周期具有简单的比例关系，沿时间轴可以搜索到各个不同的频率。在图2—18中，将DFT / DCT的固定块大小与小波的可变块长作了对比。

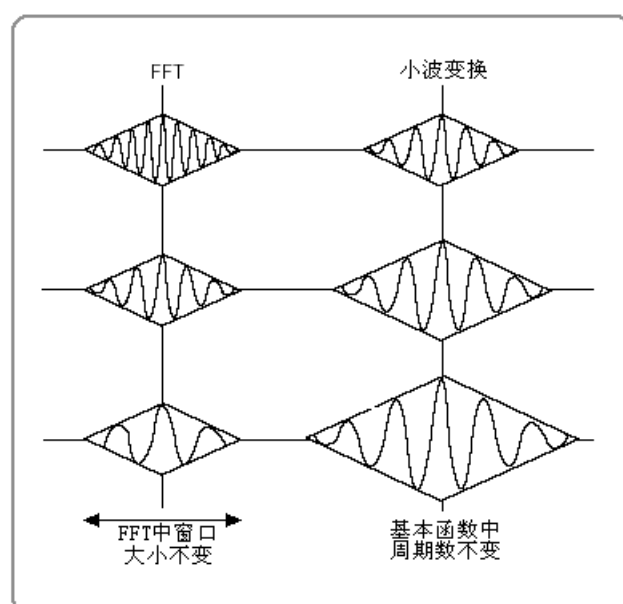


图2—18

小波特别适用于音频编码，因为它能自动地适应相互矛盾的需求，这就是时间中的瞬态准确定位和恒定音调的准确评估。

小波在视频编码中具有的优点是它能很容易地产生分辨率可变的信号。在移动视频中，由于难于给可变大小块指配运动矢量，使小波的应用受到限制，但它在静止图象或I图象编码中却不是什么问题。小波编码在极低比特率应用中具有明显的优势。因小波系数的过度量化而产生的失真一般表现为模糊不清，但与DCT系数的过度量化而造成的块效应失真相比较，这种损伤并不那么令人讨厌。

第三章 音频压缩

有损音频压缩是完全建立在人耳的听觉特性上，因此在介绍音频压缩之前应当先了解人耳的听觉特性。令人惊奇的是，人耳的听觉分辨力（特别是对立体声）实际上远高于人的视

觉，因此在处理音频压缩时应予以注意。与视频压缩一样，按照所需的压缩因子不同，音频也具有多种复杂程度不同的压缩方式。

3.1 听觉机理

人的听觉是人耳与大脑 / 神经共同处理的物理过程，正是这一过程才使我们产生声音的感受。我们对声音的感受与耳道中的实际声波并非完全一致，这是因为丢失了实际信息中的某些熵。在音频压缩系统中，如果丢失的恰好是这一部分熵，那么该系统仍能产生好的效果。

人耳的听觉系统由外耳、中耳和内耳所组成。外耳包括耳道和鼓膜。鼓膜将入射声转换为振动，这一过程与话筒的振膜十分相似。声音的振动在内耳中是通过淋巴液传送的。淋巴液的阻抗要远高于空气的阻抗，这样中耳就好象是一个阻抗匹配变换器，从而提高了声音功率的传送效率。

在图 3-1 中，声音振动通过镫骨传送到内耳，镫骨作用在椭圆窗上，内耳中淋巴液的振动传送到耳蜗，耳蜗是头骨中的螺旋腔（为清楚起见，未将图形展开）。耳底膜沿耳蜗长度方向而伸展，耳底膜的质量和硬度也沿其长度方向而变化。在接近椭圆窗一端，耳底膜硬而轻，因此共振频率较高；而另一端，耳底膜重而软，共振频率较低。耳底膜所能达到的共振频率的范围就决定了人耳的听觉范围，这一范围对大多数人为 20Hz 至大约 15kHz。

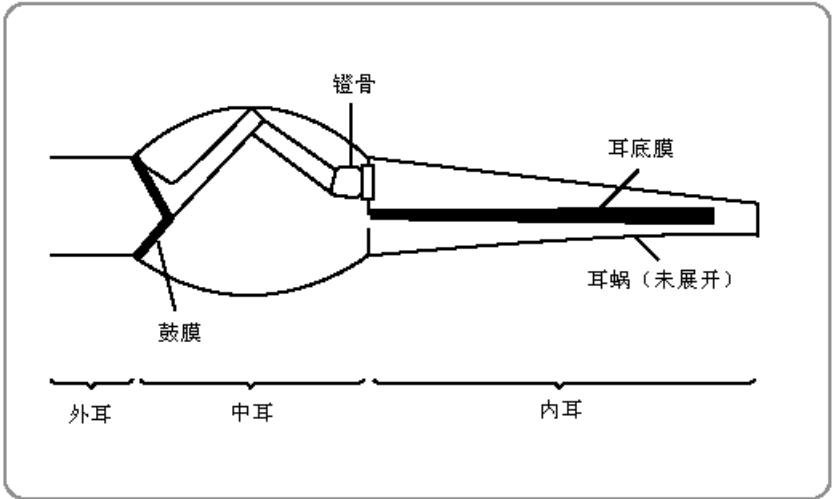


图 3-1

输入声音中的不同频率会引起耳底膜不同部分的振动。每个部分具有不同的神经末梢以分辨音调的高低。耳底膜还有受神经末梢控制的微小肌肉，它们共同作为一种正反馈系统来改善耳底膜共振的 Q 值。

耳底膜的共振效应与变换分析仪的特性十分相似。按照变换的不确定性理论，对某一信号的频率域分辨得愈准确，则对该信号的时间域分辨得愈不准确。因此，如果某种变换将两频率分辨开的能力愈强，则该变换将这两个事件在时间上分辨开的能力就愈弱。人耳的听觉系统在时间上不确定的分辨力与频率的不确定分辨力之间取得了某种平衡，当然，在这种平衡中，这两种能力都不是十分完善的。

由于受频率分辨力的限制，也就不能区分相隔较近的两个频率，这种现象称为听觉掩蔽，它定义为对某种声音的感知灵敏度因另一声音的存在而降低。

图 3-2a 表示人的听觉阈值是频率的函数。显然，感知最灵敏的频率应当是在语音范围内。在单音信号中，该阈值如图 3-2b 所示。注意音调的阈值在高频处和在低频的某个范围

内上升。当人耳接收声音的频谱比较复杂时，例如音乐，该阈值几乎在所有的频率上均会提升。这样，只有在音乐中的寂静期间，才能听见来自模拟音频盒带上的滋滋声。压扩技术正是使用了这一原理，它在录制或传送节目之前先放大低电平音频信号，然后再将其恢复至正确的电平。

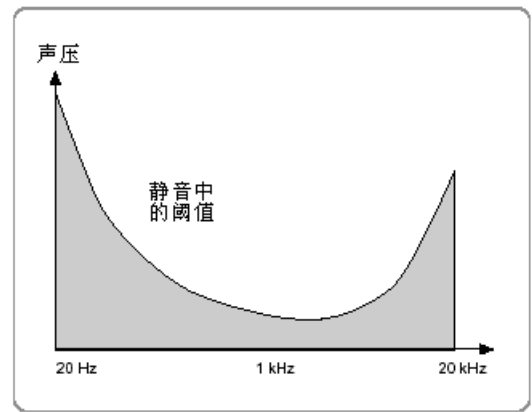


图 3—2a

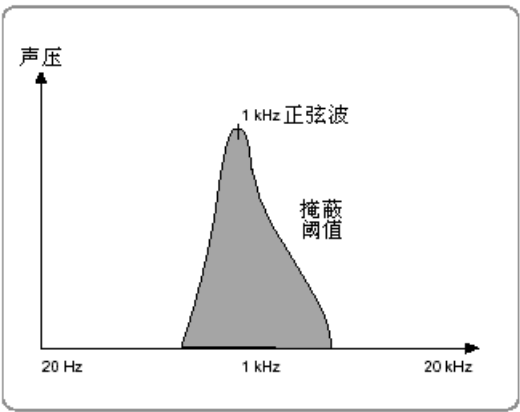


图 3—2b

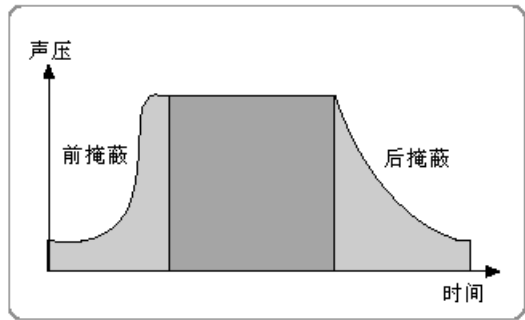


图 3—3

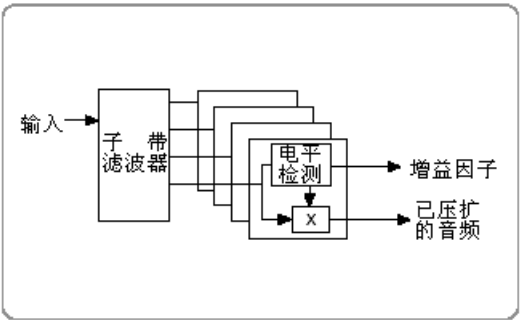


图 3—4

人耳的共振响应使人耳的时间分辨力受到了限制。Q 因子指的是在声音被人听见之前它至少应存在 1 毫秒左右。由于人耳的响应较慢，这样，即使两个信号不是同时出现，也仍然可能会发生掩蔽效应。当掩蔽信号持续作用于电平较低的被掩蔽信号时，在实际掩蔽信号存在区域的前后会发生前掩蔽和后掩蔽。图 3—3 就是这种现象的示意图。

掩蔽提升了听觉阈值，压缩器正是利用了这一特点来提升背景噪声（noise floor），这样可以用较少的比特数来表示音频波形。当然，只能在有效掩蔽的频率区内提升背景噪声。为了最有效地利用掩蔽，有必要将音频频谱分解为一些不同的频率段，从而可以在各个频段中引入不同的压扩量和噪声。

3.2 子带编码

图 3—4 为一频带分裂压扩器的示意图。频带分裂滤波器是一组具有较窄的频带、具有线性相位的滤波器，它们相互交迭在一起并有着相同的带宽。每个频带的输出是由代表波形的样值所组成。在每个频带中，在传送前先将音频输入放大至最大电平，以后再将每个电平恢复至正确电平。传输过程中所拾取的噪声在每个频带中均被降低。将噪声降低与听觉阈值相比较，可以看出，由于掩蔽作用的结果使得某些频带可以容忍更大的噪声。这样，经过压扩之后，每个频带就有可能缩短样值的字长。在这项技术中，分辨率虽然受到损失，但掩蔽

了所引入的量化噪声，由此实现了压缩。

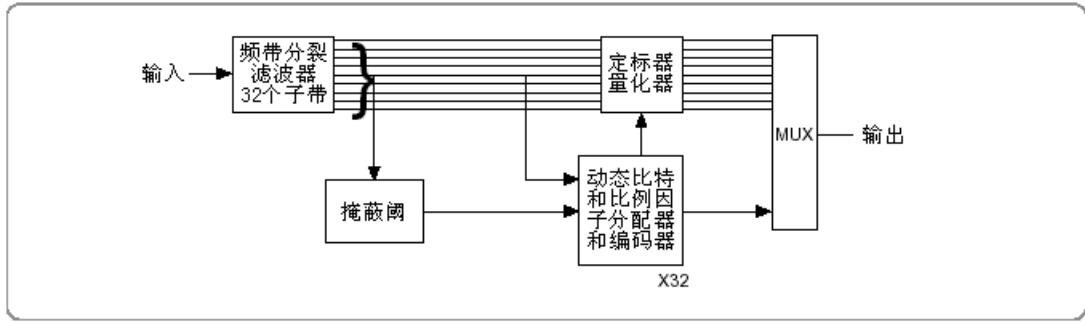


图 3—5

图 3—5 所示为一种简单的频带分裂编码器，它用在 MPEG 第一层。数字音频输入至频带分裂滤波器后，将音频信号频谱划分为多个频带。在 MPEG 中，频带数是 32 个。按照相同的时间长度将时间轴划分为块。在 MPEG 第一层中，共有 384 个输入样值。这样，在滤波器输出的 32 个频带中，每个频带有 12 个样值。在每个频带中，通过倍增将信号电平提升至最高电平。在块期间内，所需增益保持恒定，各个比例因子随同每个频带的各块一起发送，以便解码器进行反处理。

利用 512 点 FFT 对 MPEG 第一层的滤波器组输出进行分析，以确定输入信号的频谱。通过分析以驱动掩蔽模型去确定每个频带所希望达到的信号掩蔽比。可用的掩蔽愈大，则每个频带中样值的精度就愈差。通过再量化，使样值精度降低，字长缩短。这种字长的缩短对频带中的每一个码字都是恒定的，但不同的频带可以使用不同的字长。在传送时，每个频带的字长就作为该频带的比特分配码，以使解码器正确地将串行比特流转换为并行比特流。

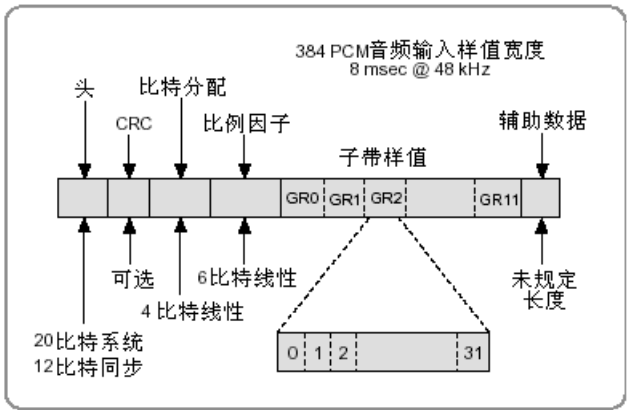


图 3—6

3.3 MPEG 层 1

图 3—6 所示为 MPEG 第一层音频比特流。在同步比特和码流数据头之后，是 32 比特的分配码（每组 4 比特）。分配码用来描述每个子带中的样值字长。随后是各子带中用于压扩的 32 个比例因子。这些比例因子确定了编码器中的所需增益，以便将音频恢复至正确电平。在比例因子之后，即为各子带的音频数据。

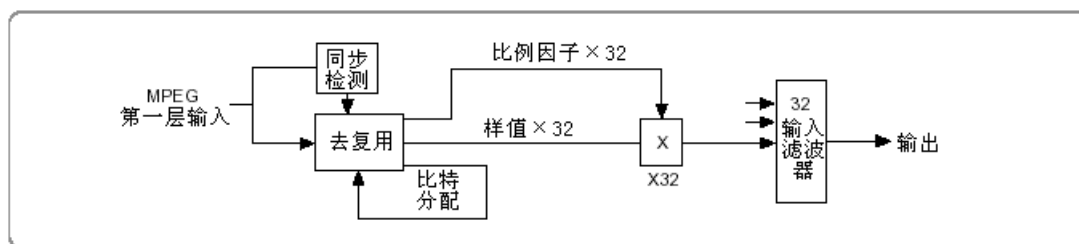


图 3—7

图 3—7 所示为 MPEG 第一层解码器。首先通过定时发生器检测同步比特，比特分配数据和比例因子数据经串并转换后，通过比特分配数据对可变长度样值进行串并转换。再按照比例因子数据进行反量化和解压缩，使每个频带恢复至原来的正确电平。将 32 个单独的子带通过组合滤波器组合在一起，最后形成音频输出。

3.4 MPEG 层 2

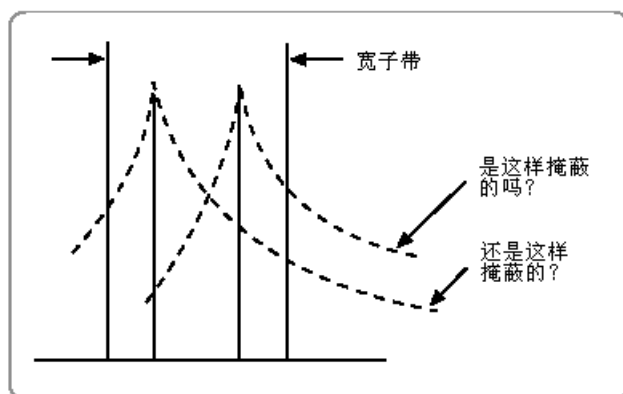


图 3—8

图 3—8 表示，在用频带分裂器去驱动掩蔽模型时，由于在该频带内能量位置的不确定以及只有 32 个频带，因此很难作出准确的频谱分析。在图示这种较差的情况下，背景噪声也不能提升太多，因为掩蔽也许无法进行。只有更准确的频谱分析才能有较高的压缩因子。在 MPEG 第二层中，可以在一个并行的过程中进行频谱分析。将 1024 点 FFT（快速付立叶变换）直接用于信号输入端以驱动掩蔽模型。为使频率更为准确，必须增加变换的时间宽度，这需要将块的大小提高到 1152 个样值。

当层 2 与层 1 的块压扩方案相同时，不需要发送全部比例因子，因为在实际节目素材中含有一定的冗余度。在相同的频带内，连续块比例因子差值为 2dB 的时间要少于 10%，通过对一组三个连续的比例因子的分析结果，以确定发送的比例因子数。当子带内容大致相同时（即所谓静态的或准静态的节目），在三个比例因子中可以只发送一个。当给定子带中的瞬态内容增加时，可发送两个或三个比例因子。比例因子选择码也应发送，这样解码器才能识别每个子带中所包含的有关比例因子的数量和位置等信息。采用这项技术，可以使比例因子比特率有效地降低至一半。

3.5 变换编码

层 1 和层 2 是基于频带分裂滤波器的子带编码，其信号仍然是以波形来表示。然而，层

3 却采用了与视频编码相似的变换编码。正如以前所述,人耳对入射声音的处理也是一种频率变换,不过由于耳底膜的 Q 值,这种响应较为缓慢,不能迅速地提升和下降。这样,如果将音频波形变换为频率域,就不需要频繁地发送这些变换系数。这正是变换编码的基础。对于更高的压缩因子,可对变换系数再次量化以减少其精度。这一过程会产生噪声,但可将它置于掩蔽最大的频率上。通过变换编码,可以准确了解音频信号的输入频谱,从而能够建立精确的掩蔽模型。

3.6 MPEG 层 3

如果要求有更高的压缩因子,这时才需使用这种复杂的编码层次。层 3 与层 2 有某些相似之处。通过离散余弦变换可以使每个变换块有 384 个输出系数。对输入样值直接进行处理,可以得到 384 个输出系数。但在多级编码器中,可以在层 1 和层 2 的基础上结合 32 频带滤波器通过混合变换来进行。此后, QMF (正交镜像滤波器) 输出的 32 个子带再由 12 频带 MDCT (改进型离散余弦变换) 作进一步处理,这样就可以得到 384 个输出系数。

为避免瞬态前回声,可使用两个窗口。窗口的切换是通过人的心理—听觉模型来进行的。已经发现,当音频信号在平均值以上时,前回声与音频中的熵相关。为了得到最高的压缩因子,可对系数进行非均匀量化和霍夫曼编码。在这项技术中,将最短的字长指配给最常用的码值。

3.7 MPEG—2 音频

尽管原先已有 MPEG 音频层 1、2、3 系统的命名,但现在为更准确起见,改称为 MPEG—1 层 1、2、3。MPEG—2 定义为 MPEG—1 音频的扩展,它是一种高级的编码系统。

在 MPEG—2 中,允许使用比 MPEG—1 更低的取样数据率。这是不太严格的逆向兼容,为与 MPEG—1 有互操作性,只需在 MPEG—1 解码器中增加附表。

MPEG—2 BC (后向兼容) 音频提供 5.1 声道 (5 个全带宽声道和 1 个窄带宽的低频效果声道)。MPEG—2 BC 在其核心仍为 MPEG—1 (2 声道) 比特流,仅以某种形式扩展为多通道,这种扩展对 MPEG—1 解码器是可以忽略的。

MPEG—2 AAC (高级音频编码) 是更为复杂的系统,它具有分辨率更高的滤波器组和其它编码工具。它能明显地提高编码效率,但它却不能后向兼容。

3.8 MPEG—4 音频

MPEG—4 是基于对象的编码 (参见 4.4.2 节)。MPEG—4 音频对象可以代表自然声或合成声。在自然声编码中, MPEG—4 的工具箱包括 MPEG—2 AAC 和其它各种工具。其中包括用于极低比特率的参数编码和所谓码激励线性预测 (CELP) 编码技术,这种编码技术可以应用在中间比特率的语音编码中。MPEG—4 支持各种形式的分级编码,其中包括可在传输系统中某些点上使用的分级比特流。

对象编码允许在解码点进行有选择的解码。例如,传送一个包括两个对象的协奏曲——管弦乐队和独奏。正常的解码是一项完整的工作,但您可以只对管弦乐队解码而成为单独的部分“实况”。类似的方法也可用于节目解码,例如,听众可以选择“混合消除”模式以取消体育运动节目中的解说声。

毫无疑问, MPEG—4 的合成音频技术具有广阔的应用前景。其中包括“文本至语音”技术和“被动配乐” (score driven) 技术,利用“结构化的音频交响乐团语言” (SAOL),通过下载工具来合成音乐。

3.9 AC-3

AC-3 是由杜比实验室所发明的音频编码技术，它用于 ATSC 系统中，以代替 MPEG 的一种音频编码方案。AC-3 已成为正式的 ATSC 标准 a/52。这样，它成为 DVB 的一个可选部分，并为摩托罗拉公司 Digicypher II 系统所采用。

PCM 输入音频进入 AC-3 编码器后，分解为多个相互重叠的窗口块，如图 3-9 所示。在这些窗口块中，每块均含有 512 个样值，由于它们是完全重叠的，因此冗余度为 100%。在变换后，每块就有 512 个系数。因为存在着冗余，利用所谓时间混叠消除 (TDAC) 技术对这些系数中取一，这样就得到 256 个系数。

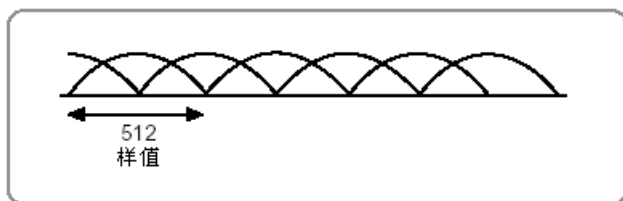


图 3-9

然后对输入波形进行分析，如果在块后一半有明显的波形瞬变，则将波形分裂成两个以阻止前回声。在这种情况下，仍为 256 个系数，但频率分辨率降低至一半，而时间域分辨率则加倍。在比特流中设置一个标志，以通知解码器这一过程已经完成。

输出系数使用浮点记数法，即用尾数和指数来表示。这是一种与科学记数法等效的二进制。其中指数是有效的比例因子。块中指数的集合给出了输入信号的频谱分析，它是按照对数比例以有限的精度作出的，称为频谱包络。它是掩蔽模型输入信号的频谱分析，由此确定噪声在每一频率上的提升幅度。

掩蔽模型驱动再量化处理，用尾数舍入法来降低各系数的精度。传输数据的有效部分由这些尾数值组成。

指数部分也需传送，但并非象它们中间有可以进一步发掘的冗余那样进行直接传送。在一个块中，仅有第一个（最低频）指数以绝对形式传输，其余指数是以差分形式传输，解码器将这些差值加至先前数值。如果输入音频具有平滑的频谱，那么好几个频带的指数都有可能是相同的。指数可以按两个或四个划分为组，并带有描述过程结束的标志。

六个块组合成一个 AC-3 同步帧。帧的第一个块总是具有完整的指数数据，但在静态信号的情况下，帧中的其它块可能会使用同样的指数。

第四章 MPEG 标准

在第二章和第三章中已经就视音频压缩技术以及 MPEG 标准中的应用工具作了说明。本章介绍 MPEG 的历史和结构以及各种 MPEG 标准的进展情况。

4.1 什么是 MPEG

MPEG 即活动图象专家组，它是隶属于国际标准化组织 (ISO) 和国际电工委员会 (IEC) 的一个委员会。IEC 处理有关电学和电子学技术的国际标准；除此之外的其它各种国际标准实际上均由 ISO 处理。在信息技术的早期，ISO 和 IEC 建立了联合技术委员会 (JTC1) 以从事于 IT 方面的工作。JTC1 下面有一些工作组，其中就包括 JPEG (联合图片专家组) 和 WG11，而 WG11 就是 MPEG。

在 MPEG 会议召集人意大利 Leonardo Chiariglione 博士的领导下, 该委员会于 1988 年成立。在正常情况下每年举行四次 MPEG 会议, 会议代表由 1988 年的 15 人左右发展到 2002 年约 300 人。MPEG 确立了一系列标准并得到了普遍的采纳, 例如 MPEG-1、MPEG-2 和 MP3 音频压缩标准 (MPEG-1 音频, 层 3)。这一情形在 MPEG-4 时有些变化, 但并非是因为标准本身有什么缺陷, 而是因为出版准许条款以及对第一款有强烈反响致使长期延误, 该条款最终于 2002 年前期获得出版。

应当注意到, MPEG 本身并不承担颁发许可证的角色, 作为 ISO 和 IEC 下属的一个委员会, 在它制订的标准中包含的技术应当是“合理的和非歧视性的条款”, 但对于什么是“合理的”, 并没有普遍认可的定义。发放许可证是相关专利持有者的责任, 一般而言, 这也意味着全世界许多从事研究开发的组织希望能看到某些回报。

MPEG-2 的一些专利持有者聚集在一起并成立了 MPEG-LA (MPEG 许可证管理局)。所有的基本专利均由该机构认证, 任何希望使用该标准的组织均允许使用其中的一部分。有关 MPEG-2 的这项工作还是正常进行的。但如上所述, MPEG-4 却面临着很大的困难, 出版许可的延迟造成了 MPEG-4 在当前商业上未能获得很大的成功。(当然, 这一情形有些变化, MPEG-4 行业论坛正在努力地工作, 以寻找专利持有者和用户都能接受的解决方案。2002 年中期所发表的修订建议可能会更容易地被接受)。

4.2 MPEG-1

MPEG-1 系统, 即 ISO / IEC 11172, 它是第一个用于运动图象的国际压缩标准, 开发于 1988 至 1992 年间。与 JPEG 一样, 它使用了 DCT 变换, 系数量化和可变字长编码, 但它包含了用于时间域压缩的运动补偿。

MPEG-1 由以下三部分组成:

- ▶ 系统 ISO / IEC 11172-1, 多路复用结构。
- ▶ ISO / IEC 11172-2, 视频编码。
- ▶ ISO / IEC 11172-3, 音频编码。

MPEG-1 在技术上显示了巨大的成功。它设计用于 SIF 图象格式, 即 352×288 (25Hz PAL) 或 352×240 (30Hz NTSC) 的图象码流压缩系统及其相关音频码流, 其压缩后的总数据率大约为 1.5Mbit/秒。这样的数据率适用于 T1 数据电路传送和 CD-ROM 重放系统, 所具有的图象分辨率与家用视频录像机大致相当。只要看一下音频 CD 的数量, 就知道 MPEG-1 获得了多么大的成功。标准的音频 CD, 它携带有双通道音频, 取样率为 44.1kHz, 分辨率为 16 比特。音频数据传输率可达 1.5Mbit/秒。这样, MPEG-1 对视音频的压缩使得视频和音频都具有相同的数据传输率!

CIF 格式是欧洲和美国 SIF (信源输入格式) 的折衷格式, 它的空间域分辨率为 625 SIF (352×288) 和时间域分辨率 525 SIF (29.97Hz)。这一格式是视频会议的基础。

MPEG-1 可以设计用于 CIF 图象, 但没有处理隔行扫描图象的相应工具, 正因为如此, 它对广播电视没有造成太大的影响。

在结束 MPEG-1 的讨论之前, 需要注意的是要看标准中所包含的实际内容以及它是如何实现互操作性的。MPEG-1 标准定义了一整套工具, 比特流的句法以及解码器的工作方式。但它并没有定义编码器的工作方式——它可以是能够提供有效句法比特流的任何设备。还有, 该标准既没有定义图象质量, 也没有定义编码质量, 从而允许开发各种编码技术而无需改变现有的标准, 也不会造成现有解码器的废弃。这种模式沿用在整个 MPEG 标准中。这种模式的采用, 所带来的成功是明显的: 尽管 MPEG-2 是应用于视频的, 但在今天的 DVB 传输系统应用中, 仍然将 MPEG-1 中层 2 音频作为主要的音频压缩系统。

4.3 MPEG-2

MPEG-1 于 1991 年终止修订（即以后的变化仅允许作编辑性的修改）。就在这一年启动了 MPEG-2 进程。MPEG-2 最终于 1994 年成为一个标准。在 MPEG-2 进程中，最初的目标是很简单的，即需要一个能够适应广播视频质量的标准。这种编码应当能够完全满足标准清晰度图象（ 704×480 ，在 29.97Hz 时和 704×576 ，在 25Hz 时）的需求，并能有效地用于隔行扫描视频编码。

MPEG-2 在许多方面都代表着 MPEG 时代的到来。将 MPEG 的巨大灵活性与大规模集成电路日益增长的可用性相结合，这就意味着 MPEG-2 具有非常广泛的应用。MPEG-3 的终结是 MPEG-2 成功的一个精彩场面。MPEG-3 本来是用于高清晰度电视的。MPEG-3 的夭折，是因为 MPEG-2 能够很容易地将其应用调整到高清这一领域，这一事实而且是愈来愈明显。这样，MPEG-2 就理所当然地成为 ATSC 和 DVB 广播标准的基础，同时它也被 DVD 用于压缩系统。

MPEG-2 还可用于图象中的运动对象。通过对下面将要讨论的类别和等级的划分，它能够为一应用建立一项全面的标准，但它又以进取的姿态转向更为急需的应用。MPEG-2 的开拓性工作一直延续到进入 2002 年。

MPEG-2 标准文件即 ISO / IEC 13818，目前有 10 部分。在 MPEG-2 标准中，最重要部分是：

- ▶ ISO / IEC 13818-1 系统（传输流和节目流），PES，T-STD 缓存器模型和基本 PSI 表：CAT、PAT、PMT 和 NIT。
- ▶ ISO / IEC 13818-2 视频编码。
- ▶ ISO / IEC 13818-3 音频编码。
- ▶ ISO / IEC 13818-4 MPEG 测试和一致性。
- ▶ ISO / IEC 13818-6 数据广播和 DSMCC。

MPEG-2 的一个主要文件是在 13818-1 中定义的，即第八章中所描述的传输流。这一设计方案的灵活性和可靠性使它有着广泛的应用，包括 MPEG-4 和 MPEG-7 数据。

注：DVB 和 ATSC 传输流是在“节目分组”中传送视音频 PES，它完全不同于“节目流”（节目流用在 DVD 和 CD 中）。

MPEG 传输流一般具有恒定的比特率，但节目流的比特率一般是可变的。

4.3.1 MPEG-2 的类和级

人们可能会感到一些意外，在设计 MPEG-1 时只有一个任务：即仅对固定大小（像素数）的图象和相关音频进行编码，且音频码率是已知的 1.5Mbit / 秒。MPEG-1 工具和句法可能并且已经有许多其它的应用，但是这种应用是在标准之外的应用，它需要有专利的编码器和解码器。符合 MPEG-1 标准的只有一种类型的解码器。

MPEG-2 在开始启动时也是这样，它本来是用于广播图象和声音的编码，即能用于 525 / 60 和 625 / 50 格式的隔行扫描电视系统。然而，当设计工作取得进展时，发现当时正被开发的工具能够处理许多格式的图象，而且能够在各种数据率下使用。此外，已经开发出用于可分级编码系统的更为复杂的工具。这意味着在实用中不会只是某一种解码器。如果有某种符合标准的解码器，它不得不处理使用所有可能的工具而编码的高速比特流，那么它就不会是作为主要码流应用的经济性解码器。举一个简单例子，某一解码设备能够解码高清晰度信号，譬如 20Mbit / 秒数据率，那么它肯定比只能对标准清晰度即码流约为 5Mbit / 秒信号

解码的解码器更为昂贵。如果需要使用一件昂贵的设备来作简单应用，这样的标准决不会是个好标准。

MPEG 使用一种两维结构即类和级来划分比特流和解码器。定义“类”为所应用的工具，例如，双向编码（B 帧）这一工具可以用在主类中，但简单类却不使用。“级”只与参数分级（scale）相关。高级解码器应能接收高速比特流，和主级解码器相比较，它就应当有较多的解码器缓存和容量较大的帧存储器。然而，高级中的主类（MP@HL）和主级中的主类（MP@ML）却使用完全相同的编码 / 解码工具和句法要素。

图 4—1 表示的是由 MPEG—2 所定义的和级的配对图示。图中以水平轴表示类，以垂直轴表示级。注意并非所有的组合都是有效的；只有在该标准中定义的完整对才是允许的。按照标准要求，任一类 / 级的解码器必须能在较低的和级上使用。例如，一个 MP@ML 解码器必须能对低级主类（MP@LL）的比特流解码和对主级简单类（SP@ML）的比特流解码。

高级		4:2:0 1920x1152 80 Mbps I, P, B	4:2:2 1920x1080 300 Mbps I, P, B			4:2:0, 4:2:2 1920x1152 100 Mbps I, P, B
高级—1440		4:2:0 1440x1152 60 Mbps I, P, B			4:2:0 1440x1152 60 Mbps I, P, B	4:2:0, 4:2:2 1440x1152 80 Mbps I, P, B
主级	4:2:0 720x576 15 Mbps I, P	4:2:0 720x576 15 Mbps I, P, B	4:2:2 720x608 50 Mbps I, P, B	4:2:0 720x576 15 Mbps I, P, B		4:2:0, 4:2:2 720x576 20 Mbps I, P, B
低级		4:2:0 352x288 4 Mbps I, P, B		4:2:0 352x288 4 Mbps I, P, B		
级别 类别	简单类	主类	4:2:2 类	SNR	空间类	高类

图 4—1

简单类不支持双向编码，因此它只有 I 图象和 P 图象输出。这样可以降低编码和解码时延，并可使用较简单的硬件。简单类只在主级上有定义。

主类是设计用于大多数应用。低级使用低分辨率输入信号，它每行只有 352 个像素。大多数广播应用需要使用 MPEG 的子集，即 MP@ML，它支持 SDTV（标准清晰度电视）。

高级—1440 是一种高清晰度方案，与主级相比较，它的清晰度加倍。高级（HP@ML）的清晰度不再加倍，仍具有和高级 1440 一样的分辨率，但使用 16:9 宽高比格式，这样它的水平样值数由 1440 增加到 1920。

如果在压缩系统中使用空间变换和再量化，就有可能产生可分级信号。可分级处理就是将输入信号划分为一个主信号和一个“助手（helper）”信号。主信号可以被单独解码以给出具有一定质量的图象，但如果再附加上助手信号信息，则图象质量将会得到改善。

例如，在一常规 MPEG 解码器中，采用较粗的再量化系数，对具有适当信噪比的图象

进行编码。而后就在发送端对其解码，并从原始图象信号中减去解码后的图象信号（应当是以像素为单位，逐个相减），这样就得到量化噪声图象。该图象经压缩后就作为助手比特流信号传送。简单的解码器只能对主比特流和噪声比特流解码，但较复杂的解码器还可对助手比特流解码，将主信号与助手信号结合起来，就可以产生接收效果更好的低噪声的图象。这就是 SNR（信噪比）可分级性的工作原理。

除了 SNR 可分级之外，还可以有另外的分级方式。如果只对 HDTV 图象中的低空间频率进行编码，可以产生一个能够用 SDTV 接收机解码的主比特流。将这个主比特流就地解码后即为低清晰度图象，再从原始的 HDTV 信号中减去低清晰度图象信号，所得到的信号即为清晰度增强图象信号。同样，对它编码后以形成助手信号。在接收端用一适当的解码器将主信号和助手信号结合起来，就能重建 HDTV 图象。这就是空间可分级性的工作原理。

图4-1中的高类支持SNR和空间可分级性，同时还允许采用4:2:2取样。

开发 4:2:2 类的目的是用来改善与数字制作设备的兼容性。在本类中，不需要对高类作复杂的改动就能按 4:2:2 方式运行。例如，某一 HP@ML 解码器应能支持 SNR 可分级性，但它在制作中是不需要的。4:2:2 类和其它类一样，均具有 GOP 结构的自由度，但在实际工作中，4:2:2 类却经常使用较短的 GOP，这样可使编辑更容易些。与 4:2:0 相比较，4:2:2 运行需要较高的比特率，对于给定的图象质量，使用短 GOP 又需要进一步增加比特率。

类和级的概念是由 MPEG-2 开发的，它提供了应用的灵活性和可靠性。在 MPEG-4 中，使用了更为复杂的类和级阵列，这将在下面讨论。

4.4 MPEG-4

国际化是一个相当缓慢的进程。新技术又在不断地出现，人们通常希望，能将新技术与正在研发中的标准相结合。然而，这种持续的改进意味着这种标准永无止境而且也无法进入实用。为了确保标准的制定能够最终地完成，应当有严格的规则，即在标准化过程中的某一阶段之后，禁止对标准作实质性的改动。这样，在标准被正式接受之时，通常也遗留下一些尚待进一步提高和扩展的问题。MPEG-2 就是这样的。如前所述，MPEG-3 在启动之后又遭终止，因此，以后的工作就由 MPEG-4 来进行。MPEG-4 的两个版本已经完成，有关进一步扩展的工作仍在继续。

MPEG-4 最初主要关注的是极低码率的视音频编码。实际上，MPEG-4 对以下三种比特率范围作了明确的最佳化处理：

- ▶ 64kbit / 秒以下。
- ▶ 64 至 384kbit / 秒。
- ▶ 384kbit / 秒至 4Mbit / 秒。

低比特率仍然是一个主要的目标，已经有了一些开创性的意见。有关误码的校正，也给予了极大的关注，这将使 MPEG-4 非常适合于易产生误码的环境中使用，例如个人手持式传输设备。不过，在 MPEG-4 中，还有使用比特率高达 38.4Mbit / 秒的类和级，以及使用数据率高达 1.2Gbit / 秒的演播室质量的类和级，这些工作仍在进行中。

更重要的是，MPEG-4 已成为更加广泛的压缩系统——它利用强大的工具开发出了全新的多媒体编码概念，使之适用于互操作性和更大范围内的应用。在 MPEG-4 标准中，仅正式的“概述”部分就长达 67 页，因此这里只能就该系统作一扼要的介绍。

4.4.1 MPEG-4 标准文件

MPEG-4 标准的主要部分如下：

- ▶ ISO / IEC 14496-1 系统。

- ▶ ISO / IEC 14496-2 图象。
- ▶ ISO / IEC 14496-3 音频。
- ▶ ISO / IEC 14496-4 一致性测试。
- ▶ ISO / IEC 14496-6 传送多媒体综合框架 (DMIF)。

4.4.2 对象编码

与传统传输系统的最大区别是有关对象的概念。在 MPEG-4 中，可以对某一确定场景的不同部分作为视频对象和音频对象分别予以编码和传送，在接收端由解码器集中在一起或作重新组合。每个不同的对象类型可以使用它最适合的工具对其进行编码。对象也可以各自单独产生，对某一场景可以分别进行分析，例如，可将前景对象和背景对象分开处理。在一次有趣的演示中，对足球比赛的视频场面进行处理，曾将足球从其余场景中分离出去，只传送背景（它是没有足球的场景）以戏弄那些没付费收看的观众。这样处理后，可以使任何人都能观看运动员和场地，但只有付费的人才能看见足球！

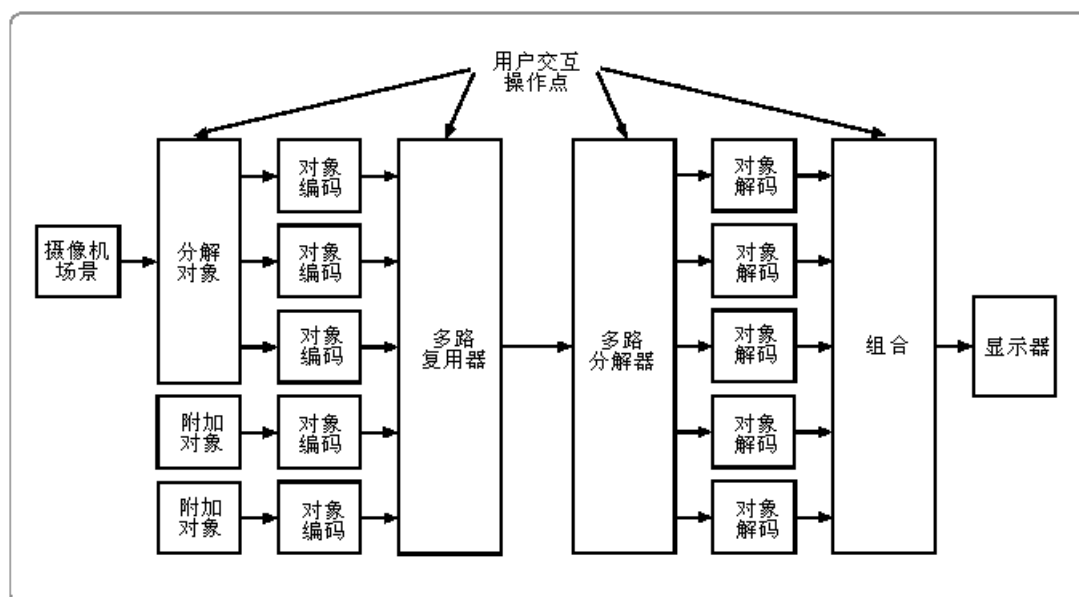


图 4-2

MPEG-4 的这种面向对象的使用方法赋予 MPEG-4 码流的三项关键特性是：

- ▶ 可以使用不同的技术分别对多个对象进行编码，而后由解码器加以组合。
- ▶ MPEG-4 中的对象既可以是自然的原始的场景，如摄像机中的场景，也可以是合成的内容，如文本。
- ▶ 比特流中的指令以及（或者）用户的选择，在同一比特流中可以具有不同的表现形式。

上述功能并非是必须使用的——MPEG-4 也提供了传统的视音频编码，通过提高编码效率以及对误码的校正能力来对 MPEG-2 作出改进。当然，MPEG-4 的真正强大之处还是来自于它的上述体系结构。各自独立的对象编码具有很多优点。每个对象的编码都可以用自己最适合的方式来进行。这样，也可以按其需要使用不同的空间分级或者时间域分级编码（参见 4.4.3 节）。

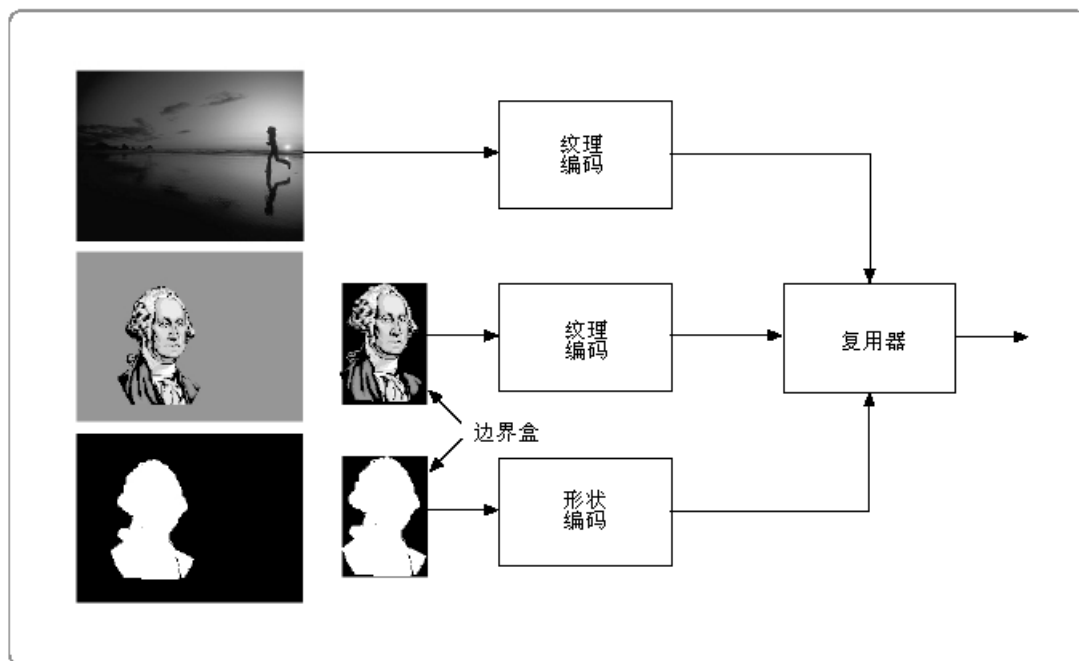


图 4—3

4.4.3 视频和音频编码

在 MPEG—4 中有许多视频编码工具与 MPEG—2 相似，不过预测编码应用得更好，熵编码的效率也更高，因此性能有所增强。然而，如何应用这些工具，却明显不同于以前的标准。

MPEG—4 的视频对象编码，在最简单的模型上是以 MPEG—2 同样的方式对视频进行编码的，不过 MPEG—4 中是用一矩形框来描述单个视频对象的。对于图象部分则使用所谓的纹理编码。如果视频对象在一个以上，其中还有一些对象具有不规则的形状，但一般而言，所有这些对象均小于全屏背景对象，因此只需对该对象的有效区进行编码，同时还应表现对象的形状和位置。在 MPEG—4 标准中，包括了矩形对象和不规则对象的形状编码工具，并以二进制或灰度级来表示（类似于阿尔法通道）。以上概念如图 4—3 所示。

至于音频编码，MPEG—4 所使用的工具也与 MPEG—1 和 MPEG—2 相似，不过 AAC 更为有效。对多个音频“对象”分别编码，然后在解码器端加以组合。和视频信号一样，音频对象既可以是自然的，也可以是合成的。

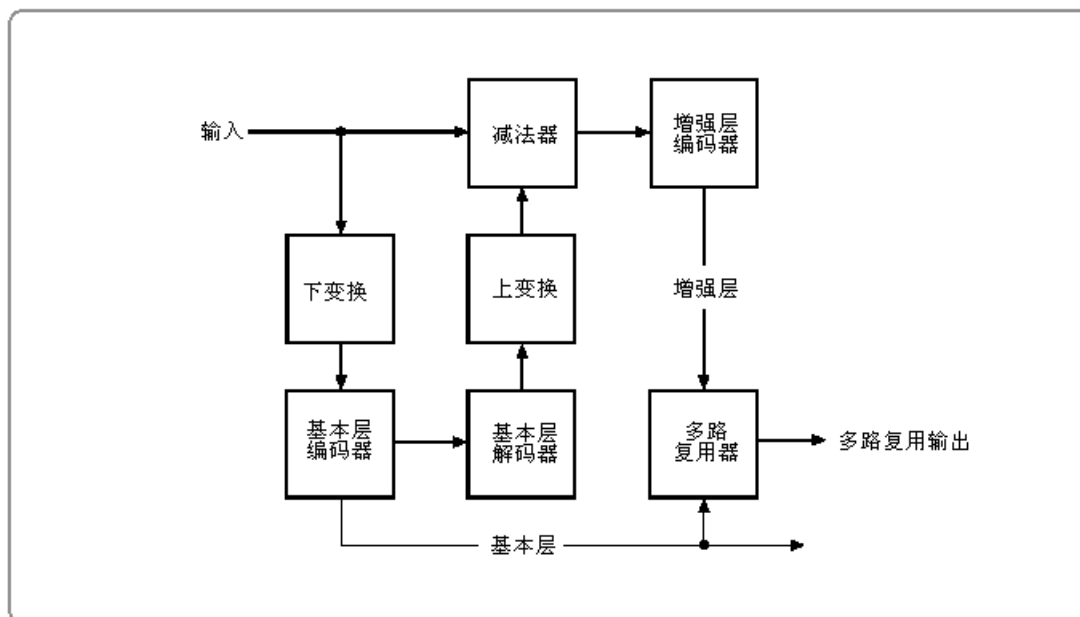


图 4—4

4.4.4 可分级性

在有关媒体的压缩中，可分级性意指在同一比特流中按照质量等级（一个以上）来分配节目内容。在 MPEG—2 和 MPEG—4 中，均使用传统的模式提供了可分级类。编码器产生一个基本层和一个或多个增强层，如图 4—4 所示。在传输或编码时如果可用资源不足，可以取消增强层。可分级类的工作过程要求在编码时必须作出有关质量级别的全部判决，在实际使用中应对增强层的数量严格限制（通常只有一个增强层）。

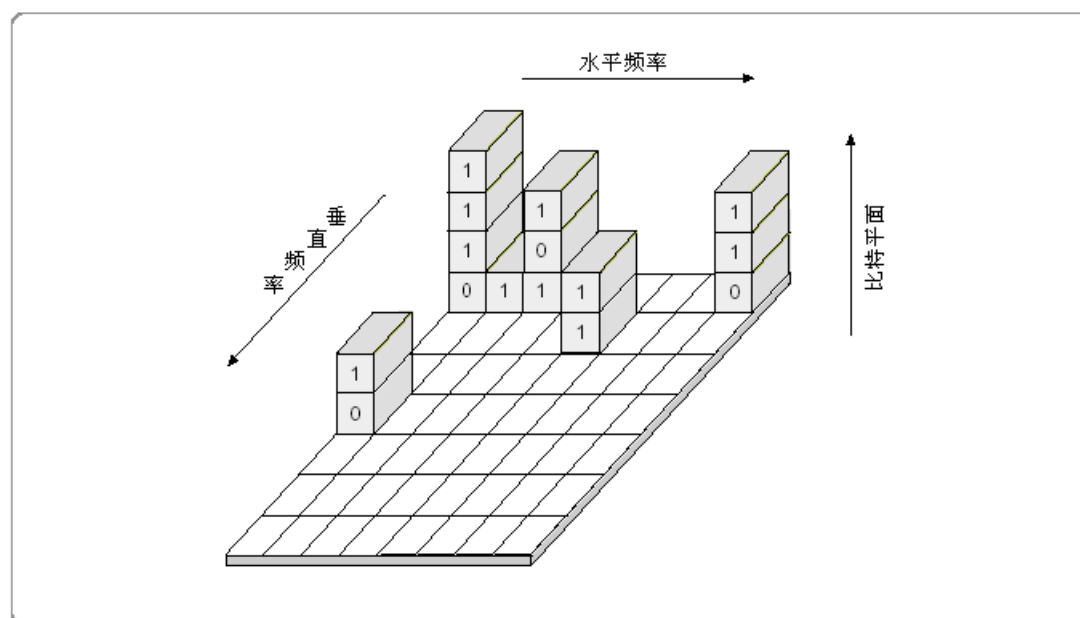


图 4—5

在 MPEG-4 的后续版本中，包括有精细纹理可分级类（fine grain scalability, FGS）。在这种技术中，产生代表最高质量等级的单个比特流，在传输路径的下游，可以抽取较低质量版本的比特流。FGS 使用的是比特平面（bit-plane）编码技术，其示意图见图 4-5。从最高有效位开始，按照每次一比特的方式对量化系数“切片”，以此给出最大（和最高有效）系数的粗略表示。此后继续切片，给出最高有效系数的较准确表示，以及次最高有效系数的近似粗略表示，如此类推。

包括 FGS 在内的空间可分级，也可以用时间域分级来组合，即当系统资源受到限制的情况下，允许在传输和（或）编码时使用较低的帧频。如上所述，对象可以是不同的分级，因此，对某一重要的前景对象而言，保持完整的时间域分辨率可能较为合适；至于背景，可用较低的速率来刷新。

4.4.5 MPEG-4 的其它方面

MPEG-4 的内容相当庞大，以上所介绍的只涉及到该标准的很少一部分。例如，还有与对象编码相联系的、使用高质量编码的演播室类，它允许组成某一视频的各个单独元素使用结构化的存储器。MPEG-4 的进一步扩展甚至还能提供适合于数字影院的高质量等级图象。在图 4-6 中，表示出目前已定义的 MPEG-4 分类（注意图 4-6 只给出了类，在每一类中一般都定义了多个级别）。

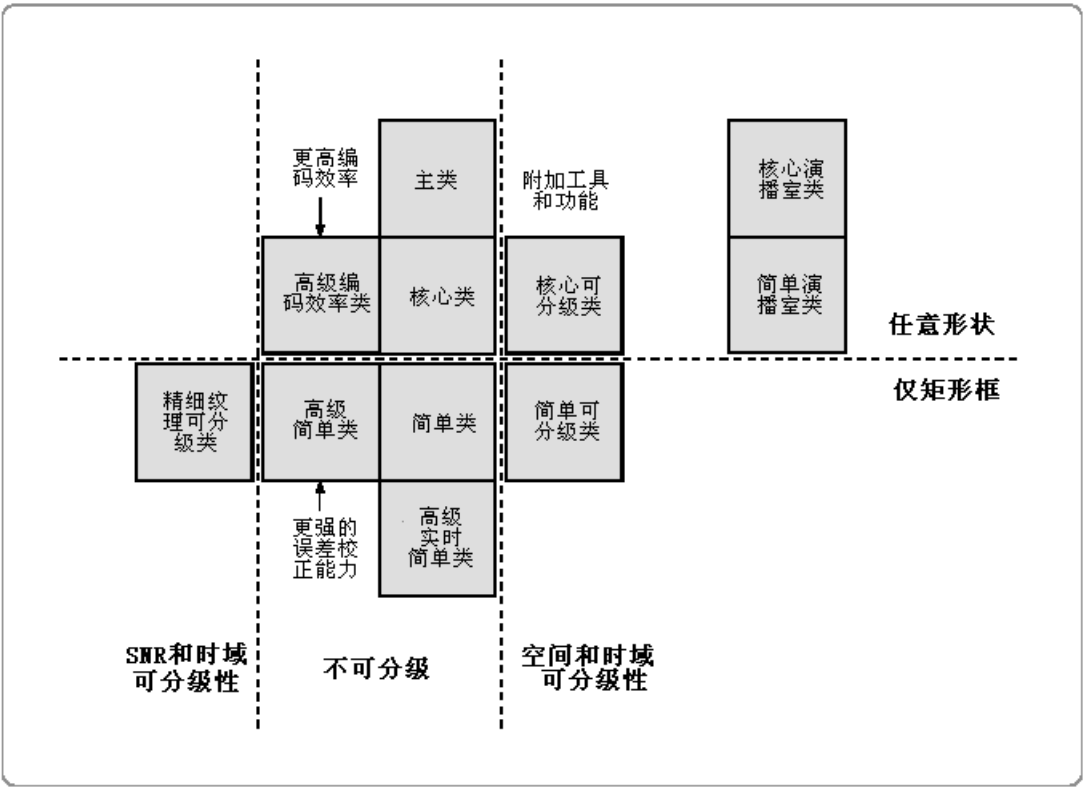


图 4-6

在 MPEG-4 中，有一些对象类型的定义是很有趣的。其中一个例子是子画面（sprite）。一个子画面是一个静止的背景对象，通常它要大于观察区或显示面。例如，视频游戏在某一背景的前面进行。如果使用子画面，可以传送一次大型静止背景，当游戏进行到与将要显示的背景相配合的部位时，就让观察区随之而动。

MPEG-4 还定义了脸部动画类和身体动画类。在每种情况下均可使用默认的脸部动画或身体动画，并可指令发送给动画。此外，还可通过比特流来修改默认对象。例如，先发送一个特定的脸部动画，而后又发送活泼的脸部。与语言相关的复杂动画指令还可以让一存储的脸部动画使用多种语言去“读取”文本。

有人把 MPEG-4 描述为视频游戏标准。当然，MPEG-4 中的许多结构在当前行业的适用性上还是理想性的。无论如何，即便是对 MPEG-4 标准大致了解一下，也可以发现它在各个方面都具有强大的功能和相当的深度，它的应用前景是无止境的。

4.4.6 MPEG-4 的未来

如前所述，MPEG-4 是一种范围广泛的、各种标准的组合，它具有强大的功能，适合于各种应用，就理论而言是这样的。但在实践上，MPEG-4 却没能取得多大的成功。实际上，许多观察家都曾预期 MPEG-4 能在因特网视听素材上成为主要的编码方法，并取代目前正在使用的、享有专利的各种编解码器。但时至今日，这种情况并未出现，即使在今后一个时期，也没有出现的可能。之所以不能成功，有以下两点理由。

首先是技术上的，也是性能上的原因。MPEG-4 所使用的视频压缩技术是基于 ITU 开发的 H.26x，它产生于上世纪九十年代早期。因特网上的音频和视频的传输分配技术是一个竞争非常激烈的行业，一些主要对手，如 Apple, Microsoft 和 RealNetworks 都开发出了一些享有专利的视频编码方案，而且它们都超过了当前的 MPEG-4 编解码器。

第二个原因是有关专利许可证的问题。直至 2002 年早期，那些希望使用 MPEG-4 的公司还不清楚他们应当向专利持有者支付专利权费用。有关 MPEG-4 基本层的许可证方案业已公布，但却受到来自业内的强烈反响。目前还不清楚 MPEG-4 的更复杂级的许可证条款。这样，许可证条款的提出，也就不能在推动 MPEG-4 成为主要的应用标准上有什么作为。

希望在未来。在 ITU 和 MPEG 的共同努力下，成立了联合视频组(joint video team, JVT)，正在为一种编解码器即 H.26L 而开展工作。据 ITU 称，“H.26L 的设计是一种基于块的兼有运动补偿的变换编码器——它与以前的设计在实质上有些相似但在许多细节上却有所不同。H.26L 显著地提高了可用块大小的数量以及增加了运动估值的可用参考图象数。”这种新的编解码器具有高得多的运动估值精度（在一些应用中为 1 / 8 像素），它所基于的主块大小为 4×4，而不是大多数 MPEG 系统所使用的 8×8。

预期 H.26L 在编码效率上有实质性的改善，这也是基本层参与者的目标，它可适用于因特网码流，并且将是免费的。JVT 的第一阶段工作预期在 2002 年完成，作为 MPEG-4 的第 10 部分。

4.5 MPEG-7

因为 MPEG-3 被取消了，按说实际标准的顺序就应当是 MPEG-1, MPEG-2 和 MPEG-4。一些委员希望下一个标准是 MPEG-5，其它有一些委员则对二进制的自然序列感兴趣，因此更希望是 MPEG-8。最后得出的结论是，任何简单的序列都不能表达从 MPEG-1 到 MPEG-4 工作上的基本差别，因此就选定了 MPEG-7。

MPEG-7 与压缩无关；它是关于元数据的，即所谓“关于比特的比特”。元数据是用来描述其它数字数据内容的数字信息。按现代的说法，用来传送信息的节目素材，或实际图象，或视频音频，或数据对象等，这些统称为数据要素（data essence）。元数据要告诉人们：它要了解的一切，全在要素中。

凡是接触过信息存储体如录象带、书籍、音乐等的人都了解，准确地编制目录和索引是多么地重要，同时也是相当困难的。当然，存储的信息只有当它的存在是已知的并且在需要

时又能够及时地取出，这样的存储信息才是有用的。

我们常常要接触到这样的问题，需要对一些标签、目录和索引等类似信息进行整理。近几年来，计算机行业给我们提供了各种有效的、廉价的和相互关联的数据库，并能通过强大的搜索引擎以异乎寻常的方式去访问已被存储的信息。当然，这种信息应当以搜索引擎能够使用的方式而存在。

问题的实质是，在今天的世界上，大量的、新的媒体内容迅速不断地涌现，随着信息量的增大和数字存储媒体价格的降低，能够存储的内容也就愈来愈多。通过本地网络和广域网络，能够存取和传送所能发现的信息。搜索引擎可以发现我们所需要的信息以及与素材本身相链接的数据库，但是我们需要以搜索引擎所适合的形式进入数据库以取得必需的索引信息。

从 MPEG 的早期标准中我们可以猜想到，MPEG 委员会当时并没有对产生数据的机制给予足够的关注。MPEG 的看法是，如果它创建一个标准化的结构，并有市场的需求，则技术上的空白就会被填补。在以往的 MPEG 标准中，对码流句法和解码器作了规定。在 MPEG-7 中，也只是句法被标准化，如图 4-7 所示。至于元数据的产生，则不作规定，正如不针对可能使用元数据的应用作出规定一样。MPEG-7 规定了元数据的表达方式。这意味着它规定了数据库的进入范围，以使搜索引擎的设计者了解在数据库中可能使用了什么样的描述元素以及它们是怎样编码的。

MPEG-7 定义了描述符和描述模式的结构，使它们几乎可以表征所有的事物。现在至少是在理论上，可以将一些基本要素（如彩色直方图和形状）组合成复杂实体（例如用来表示单个脸部）。通过描述符和描述模式，可以自动搜索索引素材，将数据库中的素材用以构成场景显示，例如可将美国总统克林顿和美联储主席格林斯潘组合在一起。这种结构并不局限于图象，例如可通过声音样本以搜索帕瓦罗蒂的录音或录象，或者只需在键盘上敲入几个音符，就可以找到与其匹配的或相接近的旋律。

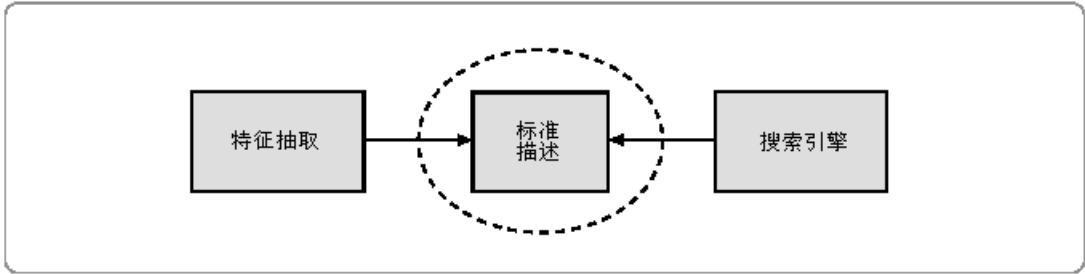


图 4-7

由于存储器件和网络系统的迅猛发展，使得人们有可能访问大量的数字信息。当技术的进步满足了 MPEG-7 的需求时，我们就能以几年前所无法想象的方式去检索和提取信息。我们还要求系统能够控制对保密信息以及与内容相关的商业事务的存取。这正是 MPEG-21 的目标所在。

4.6 MPEG-21

MPEG-21 在分类上再一次不同于 MPEG 委员会的以往工作。它的基本概念是相当简单的一一尽管它涉及面很广。MPEG 力图创建一个用于管理和使用数字资产的完整结构，其中包括伴随该结构的商业事务和权限管理所支持的全部基础设施。图象表述（vision statement）是“能够跨越各种网络和器件以透明地和广泛地使用多媒体资源”。

有关 MPEG-21 的工作范围，可以通过它的技术报告草案中所定义的几种架构来表示。

1. *数字条目说明* 预计是“为定义数字条目而建立的一种统一和灵活的抽象概念并能共同使用的模式”。该模式应当对所有的媒体资源类型和描述模式都是开放的和可以扩展的，并且应当支持分层结构以利于搜索和管理。

2. MPEG-21 的*数字条目表述* 是技术性的，它应当用于内容编码和提供该内容中所有元素同步的所需全部机制。预期该层将至少对 MPEG-4 进行定位。

3. *数字条目标识和描述* 将提供数字条目（与所有内容元素相链接）标识和描述的框架。它可能包括 MPEG-7 的描述模式，还应当包括“新一代的标识系统，以支持有效、准确和自动的事件管理和报告（许可证事务，使用规则，监视和跟踪等）。”它应当满足各类 MPEG-21 用户的需求。

4. *内容管理和使用* 应当对 MPEG-21 数字条目和描述的存储和管理的接口和协议作出定义。在保留使用权限的同时，它应当支持内容的归档和目录的编制，并可以跟踪条目和描述的改变。MPEG-21 中的这一部分也可能支持一种“交易”形式，用户可以交换它们有关访问内容权限的个人信息，以及支持“个人频道”和类似结构机制的形式化。

5. *知识产权管理和保护* 是一个基本组成。当前围绕 MP3 音频文件的争论证明需要一种能被数字世界所识别的新的版权机制。这也说明当内容不被保护时，它是没有价值的。MPEG-21 将建造在进行中的 MPEG-4 和 MPEG-7 工作的基础上，当然它也需要扩展以适应新的数字条目类型和新的传送机制。

6. MPEG-21 *终端和网络* 将对广域网中条目的传送进行编址，并能在各种终端上提交内容。就概念而言，网络应能将具有全数字影院质量的电影提交给电影院，也可以通过速率较慢的网络向家庭用户提供较低质量的电影（按照不同的费用）。在以上两种情况下都应有用户数量或传送类型上的某种限制。至于与传送和提交相关的任何问题，以及过程如何复杂，都是用户不必关心的。

7. 最后，还需要事件报告，以“使所有可报告事件性能的度量（metrics）和接口符合标准。”可以举一个最直观的例子，如果系统允许某一用户访问已受保护的条目，可以相信，相应的报酬也就出来了！

第五章 基本码流

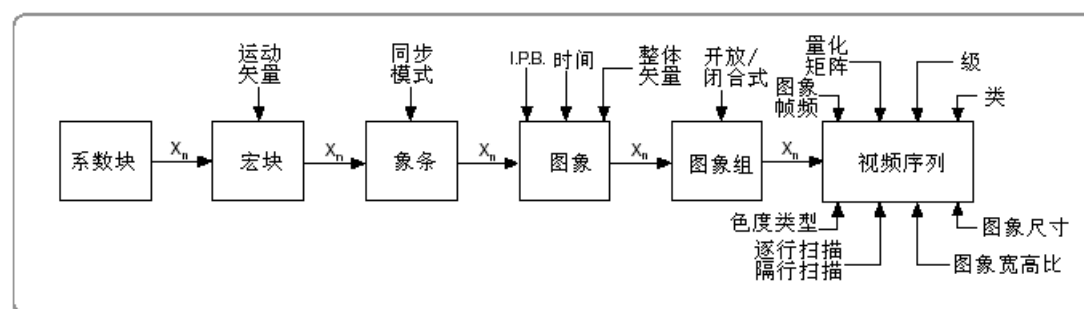


图 5-1

基本码流是由编码器输出的原始基础码流，它只含了解码器所必需的、并与原始图象或原始音频相接近的信息。MPEG 对已压缩信号的句法作了严格的定义，以保证解码器能够正常解码。MPEG 对编码器未作定义，但它必须能够提供句法正确的码流。

MPEG 这样规定的优点是它与实际情况相符合。在世界上，解码器的数量要远多于编码器，对解码器标准化后，可以降低它的生产成本。与之相反，编码器则可以更复杂和更昂贵，对成本的考虑也可以放宽，这样，随着编码器的日益复杂，所提供的图象质量也会更高。当编码器和解码器的复杂程度存在差别时，这种编码系统可称为非对称系统。

MPEG 对编码器不作定义，这就有可能通过改进编码算法来提高图象质量，而且产生的码流也能被早期的解码器所接收。还有，因为未对编码器作定义，也就是在非公共范围内允许使用专用的编码算法。

5.1 视频基本码流句法

图 5—1 给出了基本视频码流的结构。图象信息的基本单元是 DCT（离散余弦变换）像素块，它表示图象中 8×8 像素阵列，这些像素可以是 Y、Cb 或 Cr。首先发送的是像素块中的直流（DC）系数，它比其它系数能更准确地被重现。随后发送的是块中的其它系数，最后发送是块结束（EOB）码。

宏块是由像素块组成的，它是图象中用于运动补偿的基本单元。每个宏块在其数据序列头部都存放有一个两维运动矢量。在 B 图象中，该矢量既可以是前向的，也可以是后向的。可以是基于场的运动补偿，也可以是基于帧的运动补偿，并作出相应的标记。宏块中系数量化的比例也应作出标记。解码器通过运动矢量以获取来自前一图象和后一图象的信息，从而产生一个预测图象。对像素块进行反变换以产生一个校正图象，并将它与预测图象相加以产生解码输出。在 4:2:0 格式编码中，每个宏块具有四个 Y 像素块和两个色差像素块。为了便于识别哪个像素块是描述哪个分量，应当按照规定的顺序来发送像素块。

由宏块组成像素条。像素条应当始终是表示图象中从左到右的水平条。在 MPEG 中，像素条可以从图象中任意部位开始，其大小也可以是任意的，但在 ATSC 中，它必须从图象的左边沿开始。（在同一水平线上的）若干个像素条可以跨越整个屏幕宽度。像素条是用于可变字长编码和差值编码同步的基本单元。像素条中的第一个矢量应当以绝对完整的形式发送，其余矢量则可以用差值形式发送。在 I 图象中，像素条中的第一个 DC 系数也应当以绝对完整的形式发送，其余的 DC 系数可用差值形式发送。在差值图象中，由于上述系数是不作预测使用的，因此也就没有上述规定。

在基本码流中，如果发生了比特错误，无论是可变字长码元的串并转换受到破坏，还是随后的差值编码系数或矢量中出现误码，像素条结构均可以通过在比特流中提供再同步点来加以恢复。

许多像素条组合成图象，它是一场或一帧画面的有效部分。无论图象是已编码的 I 或 P 图象还是已编码的 B 图象，在图象头中均定义有一个时间基准，有了它，图象才能按照正确的时间显示。如果出现摇镜头或作俯仰拍摄的场景，则每个宏块的矢量是相同的。这时可以发送整个图象的整体矢量（global vector），各单个矢量就是这个从整体矢量导出的差值。

若干个图象可以组成一个图象组，按照传送顺序，图象组应当从 I 图象开始。图象组是时间编码的基本单位。在 MPEG 标准中，图象组的使用是一个可选项，但实际上它是必需的。在两个 I 图象之间，可以放置数量不固定的 P 图象和（或）B 图象，具体放置位置如第二章所述。图象组可以是开放式的或闭合式的。在闭合式的图象组中，在对最后一个 B 图象解码时不需要下一个图象组的 I 图象，而且比特流可以在图象组结束处截断。

如果使用图象组，那么若干个 GOP 组合起来就可以产生一个视频序列。视频序列用一个序列起始码作为开始，后面跟着是序列头，并且以序列末端码结尾。在整个视频序列中的某一位置，还可以放置附加的序列头，这样，便于在该视频序列中找到解码的起始部分，例如在重放数字视频光盘或录象带时就会出现这种情况。在序列头中，规定了图象的垂直和水平尺寸，图象的宽高比，色度信号的亚取样格式，图象的帧频，是逐行扫描还是隔行扫描，类别和级别，比特率以及在帧内编码和帧间编码中所使用的量化矩阵等。

如果没有序列头数据，则解码器就不能识别比特流，因此序列头是解码器能够开始正常工作的进入点。进入点的间隔会对正确解码的时延有些影响，例如在观看者变换电视频道时就有可能发生。

5.2 音频基本码流

在 MPEG-2 复用器中可以嵌入各种类型的音频码流。它们是按照 MPEG 层 1、2、3 或 AC-3 进行编码的。所使用的音频编码类型包含在一个描述符中，解码器通过对它的读取以启动相应的解码类型。

音频压缩过程完全不同于视频压缩。它没有相对应的 I、P 和 B 帧类型，各个音频帧中所包含的音频数据量始终都是相同的。也没有相对应的双向编码，因此音频帧的传送是按照时间顺序来进行的。

在 MPEG-2 音频中，序列头中的描述符包含有音频压缩层和使用的压缩类型（例如联合立体声，joint stereo），以及原始的取样率。音频序列是由许多存取单元（AU）所组成的，存取单元是已编码的音频帧。

如果象 ATSC 那样使用 AC-3 编码，则会在序列头中对这一用法给予说明。有关 AC-3 同步帧中的音频存取单元（AU）的内容参见 3.7 节。AC-3 同步帧代表与 1536 个音频样值等效的时间区间，当取样率为 48kHz 时，该区间为 32ms；取样率是 32kHz 时为 48ms。

第六章 打包的基本码流（PES）

在实际应用中，需要将来自压缩器的携带音频或视频的连续基本码流分解成若干个数据包。在数据包头中包含有用于同步的时间标记，通过这些包头就可识别数据包。使用 PES 包可以创建节目流或传输流。

6.1 PES 包

在 PES 中，长度不固定的基本码流可以按其应用划分为大小不同的数据包。其大小可能有数百 KB（千字节），当然，具体长度随应用而变。

在每个包的前面都有一个 PES 包头。图 6-1 给出了一个包头的内容。数据包由一个具有 24 比特的起始码前缀和一个码流 ID 开始，码流 ID 是用来识别包的内容，例如是视频包还是音频包，还可进一步用它来识别音频编码的类型。以上两种参数（即起始码前缀和码流 ID）组成了包起始码，通过它来识别数据包的开始。需要注意的是，不要把 PES 中的数据包与传输流中的小得多的数据包相混淆，虽然它们的名称是一样的。

由于 MPEG 只定义了传输流，而没有对编码器作出规定，设计者可以选择构建一个将基本码流转换为传输流的多路复用器。在这种情况下，PES 包就不再以可识别的形式存在，而是以逻辑形式存在于传输流的有效载荷中。

6.2 时间标记

经过压缩后，由于采用了双向编码，因此图象不再按时间顺序发送。各帧图象的数据量是不同的，同时在经过复用和传输之后，也具有不同的时延。为使码流中的音频和视频一起锁定，就需要在每一图象中定期地插入时间标记。

时间标记是一个具有 33 比特的数字，它是受 90kHz 时钟驱动的计数器的一个取样值。该时钟是由 27MHz 节目时钟经过分频（分频倍率为 300）而获取的。由于显示时间具有相同的间隔，因此没有必要在每个显示单元内插入时间标记。时间标记也可以由解码器插入，但无论是节目流还是传输流，其插入的时间间隔不应大于 700ms。

时间标记用以指示一个特定的存取单元在时间轴上应处于什么位置。通过插入在视频和音频 PES 包头中的时间标记来保证唇音同步。当解码器接收到被选中的 PES 包时，它对每一存取单元进行解码并将它们存入 RAM 以作缓冲。每当时间计数到达规定的时间标记值时，

再从 RAM 中提取。这种运行方式具有两个优点。第一，在每一基本码流中都能得到有效的时基校正。第二，可以保证视频和音频基本码流相互同步以构成节目。

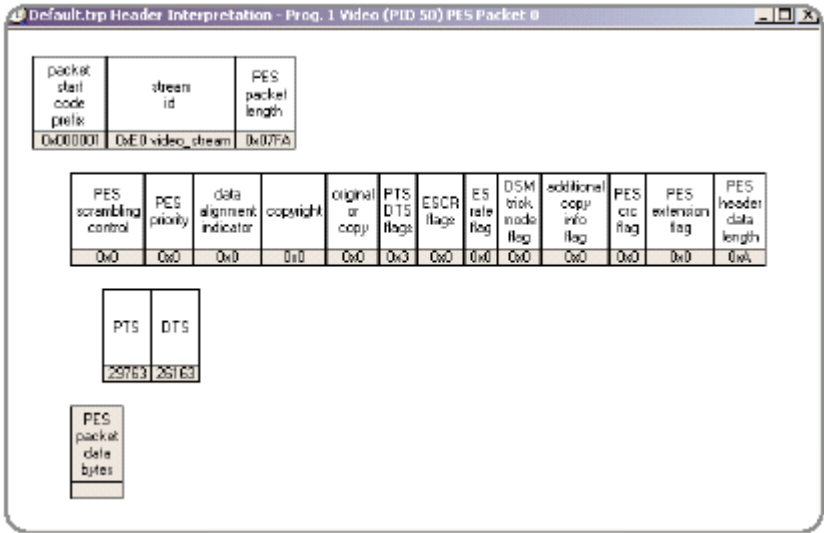


图 6-1

6.3 PTS / DTS

在使用双向编码时，对某一图象的解码必须在其显示之前的一段时间内进行，这样它才能作为解码 B 图象的源数据。例如，图象的显示顺序是 IBBP，但图象的传输顺序则是 IPBB。这样，就存在着两种类型的时间标记。解码时间标记（DTS）用以指示该图象应被解码的时间；而显示时间标记（PTS）则指示的是解码器输出该图象的时间，也即显示时间。

由于 B 图象的解码时间和显示时间是一致的，因此它只含有 PTS。当解码器接收到 IPBB 图象序列时，它必须在解码第一个 B 图象之前先对 I 和 P 图象解码。解码器只能每次解码一帧图象，因此首先对 I 图象解码并将其存储起来，待 P 图象被解码时，即输出已解码的 I 图象，随后才是 B 图象。

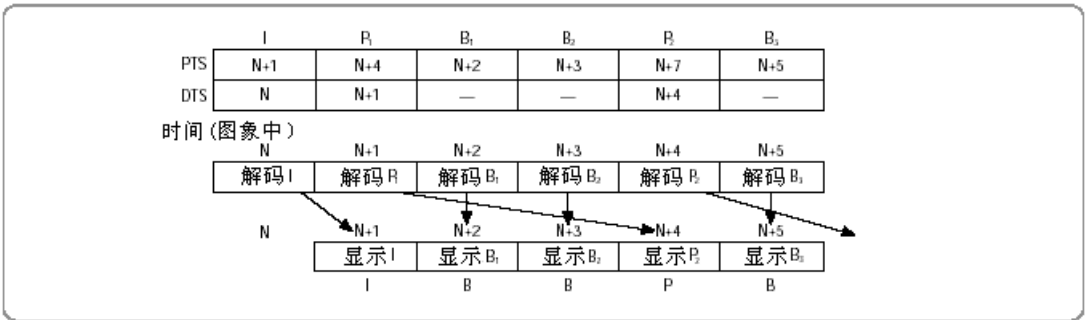


图 6-2

如图 6-2 所示，当接收到某个含有一个 I 图象的存取单元时，在其头文件中应有 DTS 和 PTS，且两个标记之间的时间间隔为一个图象周期。如果使用双向编码，则 P 图象应在 I 图象之后，它的头文件中也应有一个 DTS 和一个 PTS，但这两个标记之间的间隔为三个图象周期，这样才能插入两个 B 图象。如果现在接收的是 IPBB，那么 I 应延迟一个图象周期，P 应延迟三个图象周期，两个 B 图象则不延迟。最后的显示顺序就成为 IBBP。很明显，如果改变了图象组的结构，例如在 I 和 P 之间有更多 B 图象，那么就必須增加 B 图象中 DTS

和 PTS 之间的时间间隔。

在数据包头中，设置有 PTS / DTS 标志，它能指示某图象中是只有一个 PTS 时间标记还是含有 PTS 和 DTS 两个时间标记。音频数据包可以含有多个存取单元，且在包头中只含一个 PTS。由于音频数据包必须按时间顺序发送，所以音频包中不含 DTS。

第七章 节目流

节目流是将若干个 PES 打包码流组合起来的一种方式，它可作为记录应用如 DVD 中。

7.1 记录与传输

对于给定的图象质量，压缩视频的数据率随图象内容而变。可变比特率的频道将会给出最好的图象质量。但在节目传输的实际应用中，大多数频道的传输比特率是固定的，并采用填充无意义数据的方法使总比特率保持不变。

在 DVD 中，采用填充的方法则是对存储器容量的一种浪费。这是因为可以改变存储介质的速度（加快或降低），例如在光驱中就是以物理方式来改变数据的传送速率。这种方式不需要浪费存储器的容量即可改变传输通道的数据率。当介质重放时，可调整其速度使数据缓存器的容量大约有一半填充数据，这时可不必考虑实际比特率的大小，因为它是动态变化的。如果解码器从缓存器中读取数据的速率增加，表示缓存器中的存储数据在减少，这时只需简单地驱动系统提高存取速率以恢复平衡。不过，只有当音频和视频的编码（与解码）是来自于同一时钟时才能使用上述方法，否则将改变记录长度。

为了满足不同的需求，对节目流和传输流的设计在上述方法中各择其一。即在记录环境中，通常是一个节目，则采用可变比特率；在传输环境中，通常是多个节目共用一通道，则使用固定比特率传送。

在 DVD 重放过程中，不会出现与源信号的同步锁定问题。DVD 重放器是通过本机的同步脉冲发生器（内部或外部）来确定视频时基的，它只需简单地从光盘中读取数据即可按该时基提供图象。但在传输中，解码器必须重建时基而且必须与编码器的时基保持一致，否则会使（解码器中的）缓冲器上溢或下溢。这样，传输流就要使用节目时钟基准（PCR），而在节目流中则是不需要的。

7.2 节目流简介

节目流是一种 PES 包的复用，它携带有若干个利用同一主时钟或系统时钟（STC）而编码的基本码流。这种码流也许就是一个视频码流以及与其相关的音频码流，也可以是多通道的音频节目（只含音频）。基本视频码流划分为存取单元（AU），每个存取单元都包含有描述一帧图象的压缩数据。这些图象可以是 I、P 或 B 帧，但每帧图象都带有一个 AU 序号，从而保证正确的显示顺序。一个视频 AU 就是一个节目码流包。在视频中，包的大小并不固定。例如，I 图象包就比 B 图象包大得多。数字音频存取单元的大小一般是相同的，若干个存取单元组成一个节目码流包。注意不要把这里的数据包与传输流数据包相混淆，传输流数据包较小且有固定的大小。视频和音频存取单元的边界在时间轴上一般并不重合，但这不会有什么问题，因为每个存取单元的边界都有自己的时间标记结构。

第八章 传输流

传输流是许多种 PES 包的多路复用。在节目流中，由于音频和视频均锁定于一个公共时钟，因此可利用其时间标记以重建时间轴。但在传输流中，则需为解码器所接收的每个节目重建时钟，这就需要附加一个句法层以提供 PCR 信号。

8.1 传输流的工作过程

传输流携带有许多不同的节目，每个节目都可以使用不同的压缩因子并具有不同的比特率。尽管总比特率是不变的，但每个节目的比特率却是可以动态改变的。这一过程可称为统计复用。这样，当某一个节目素材因比特率不足而难于处理时，它可以从另外一个较容易处理的节目素材中取得一部分带宽。在每个视频 PES 中，都可以有不同数量的并与其相关的音频 PES 和数据 PES。尽管传输流具有这种灵活性，但解码器应当能够从一个节目转换到另一个节目并能正确地选择相应的音频和数据通道。对某些节目可以采取保护措施，以使这些节目只能被付费用户收看。在传输流中，应当含有用以管理这种保护措施的条件接收(CA)信息，同时还应有处理上述任务的节目专用信息 (PSI)。

传输流将 PES 数据转换为更小的数据包，这些数据包具有固定的大小（必要时可使用填充比特）。当这些数据包到达解码器时，有可能出现定时抖动。此外时分复用也会造成时间的延迟，但这种因素是不确定的，因为分配给每个节目的比特流部分是不需要固定的。可以利用时间标记来解决这一问题，但前提是要有稳定的时钟。这样，在传输流中就必须含有重建稳定时钟的附加数据。

数字视频制作设备的正常运行与一个稳定的同步时钟分配系统密切相关。在视频制作中虽然可以使用同步锁定，但在长距离传输应用中，对~~不同~~时钟进行分配则是不可行的。在传输流中，不同的节目也许来自于不同的地方，它们之间也不需要同步。因此，传输流必须为每个节目提供各自独立的同步。

这种同步方法称为 PCR，它可以重建一个稳定的基准时钟，在解码器中，基准时钟经过分频后能够创建时间基线，这样，每个节目的基本码流中的时间标记才是可以使用的。因此，可以把节目的定义理解为共享同一个定时基准的一组基本码流。

在只有一个节目的传输流(SPTS)中，用以重建音频和视频节目时钟的就只有一个 PCR 通道。SPTS 通常作为音频 / 视频编码器和复用器之间的通讯联络使用。

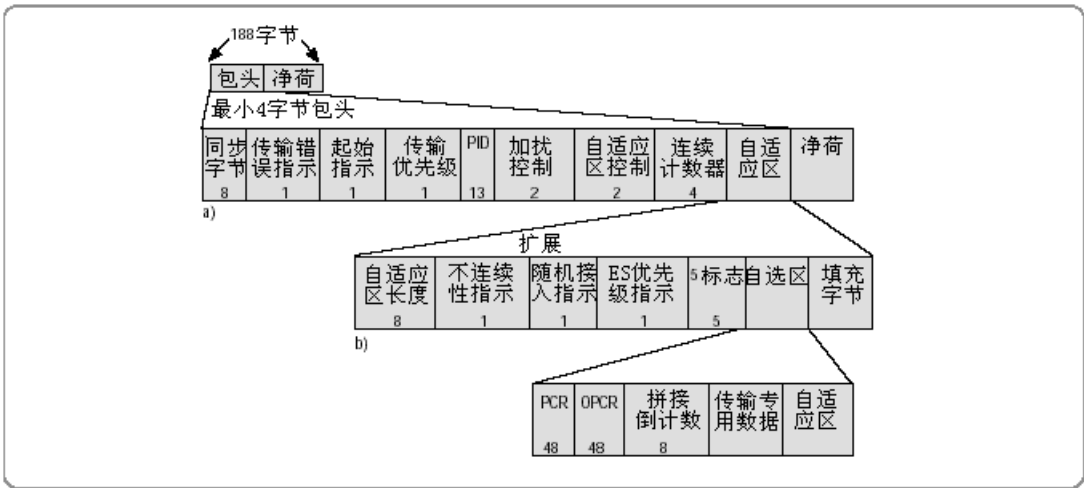


图 8-1

8.2 传输流包

图 8—1 所示为传输流包的结构示意。包的大小为固定的 188 个字节，它总是划分为包

头和有效载荷（净荷）两部分。图 8—1a 为最小的 4 字节包头。在包头中含有以下重要信息：

- ▶ 同步字节。同步字节是作为解码器识别用的，从而可对包头和有效载荷进行串并转换。
- ▶ 传输错误指示。在传输层之上的错误校正层中，如果原始误码率（BER）太高而无法校正时，通过它的设置可以指示传输流包中可能含有错误。有关错误校正层的详细介绍参见第十章—DVB 和 ATSC 概述。
- ▶ 包识别符（PID）。它是 13 比特的代码，用于区分不同类型的传输流包。有关 PID 的细节稍后讨论。
- ▶ 连续性计数器。它是 4 比特数值。每当发送一次新的具有相同 PID 的传输流包时，即可被复用器递增。它用于确定传输流包是否有丢失、重复或顺序错误等现象发生。

在某些情况下可能需要更多的包头信息，如果出现这种情况，就要设置自适应场控制比特来指示这时的包头比特数要多于正常数。图 8—1b 表示当出现上述情形时，通过自适应场长度码来指示附加的包头长度。这样包头便被扩展，相应地要减小有效载荷以维持恒定的传输流包长度。

8.3 节目时钟基准（PCR）

用于一特定节目的编码器具有 27MHz 的节目时钟。当输入信号为 SDI（串行数字接口）视频时，其比特时钟经 10 分频以产生编码器的节目时钟。如果有数个节目均来自于同一制作设备，那么它们可能具有相同的时钟。在输入信号是模拟视频的情况下，需将行同步周期乘以锁相环中的一个常数以得到 27MHz 时钟信号。

传输流包头中的自适应场用以周期性放置 PCR 码，以在解码器产生一个被锁定的时钟。当编码器或复用器切换信号源时，PCR 就会中断。这时连续性计数也会受到干扰。这种情况是由不连续性指示器处理的，它会告知解码器将有干扰发生。如果没有通知解码器，那么不连续性就是一种错误状态。

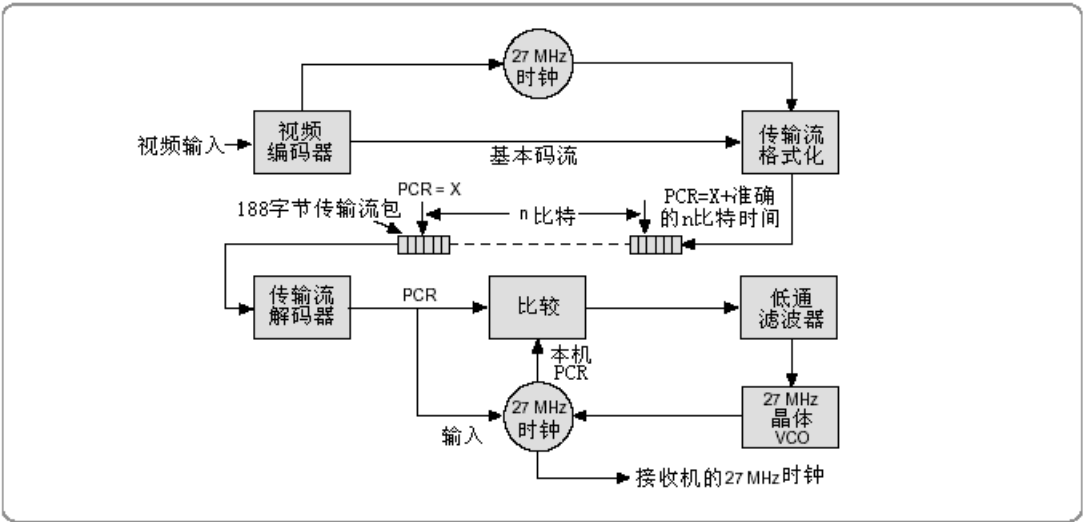


图 8—2

图 8—2 表示解码器是如何利用 PCR 以重建每个节目的 27MHz 时钟。编码器时钟驱动一个稳定运行的二进制计数器，计数器的数值经周期性取样后并作为 PCR 放置在包头自适应场内。PCR 是一个 42 比特的数字，它表示一个 33 比特的 PCR 基础位再加 9 比特的 PCR

扩展位以进一步提高分辨率（PCR 基础位，如同 PTS 一样，是一个 33 比特的数字，它是受 90kHz 时钟驱动的计数器的取样值）。每个编码器产生的传输流包具有不同的 PID。解码器通过被选节目的正确 PID 以识别传输流包，而含有其它 PID 的传输流包则被放弃。在解码器中，一个压控振荡器（VCO）产生标称的 27MHz 时钟信号，并由它驱动本机 PCR 计数器。将本机 PCR 计数器的输出与来自传输流包头的 PCR 相比较，所得到的差值即为 PCR 相位误差。PCR 相位误差经滤波后去控制 VCO，这样最终使本机的 PCR 与传输流包头中的 PCR 相同步。通过加强 VCO 滤波以确保该时钟不被 PCR 传输中的抖动所调制。不连续性指示器将重新设置本机 PCR 计数，还可选用它来减轻滤波，以使系统快速地与新的定时相锁定。

在 MPEG 中，要求 PCR 的传送速率达到每秒内至少发送 10 个 PCR，而按 DVB 规范则需在每秒内最少发送 25 个 PCR。

8.4 包识别符（PID）

传输流包头中的 13 比特场存放有包识别符（PID）。解复用器在解复用时正是利用它来区分含有不同类型信息的传输流包。尽管各码流的比特率的总和可能会有变化，但传输流比特率则应当是恒定的。这就需要使用空值数据包来作填充处理。如果实际有效载荷率下降，则需插入较多的空包。所有的空包均具有相同的 PID，其值为 8191（在二进制中为 13 个 1）。

在给定的传输流中，属于某一给定基本码流的所有数据包均具有相同的 PID。在解复用时，解复用器只需接收含有正确 PID 的传输流包就可以很容易地选择给定基本码流的所有数据。利用视频、音频和数据码流的 PID 就可以选择全部节目的数据，如字幕和图文电视。只有当解复用器在传输流包与其所在的基本码流之间建立了正确的关联，它才能正确地选择这个传输流包。也只有当解复用器能识别正确的 PID 内容时它才能完成这项任务。总之，有关 PID 的信息是十分重要的，这正是节目专用信息（PSI）的功能。

8.5 节目专用信息（PSI）

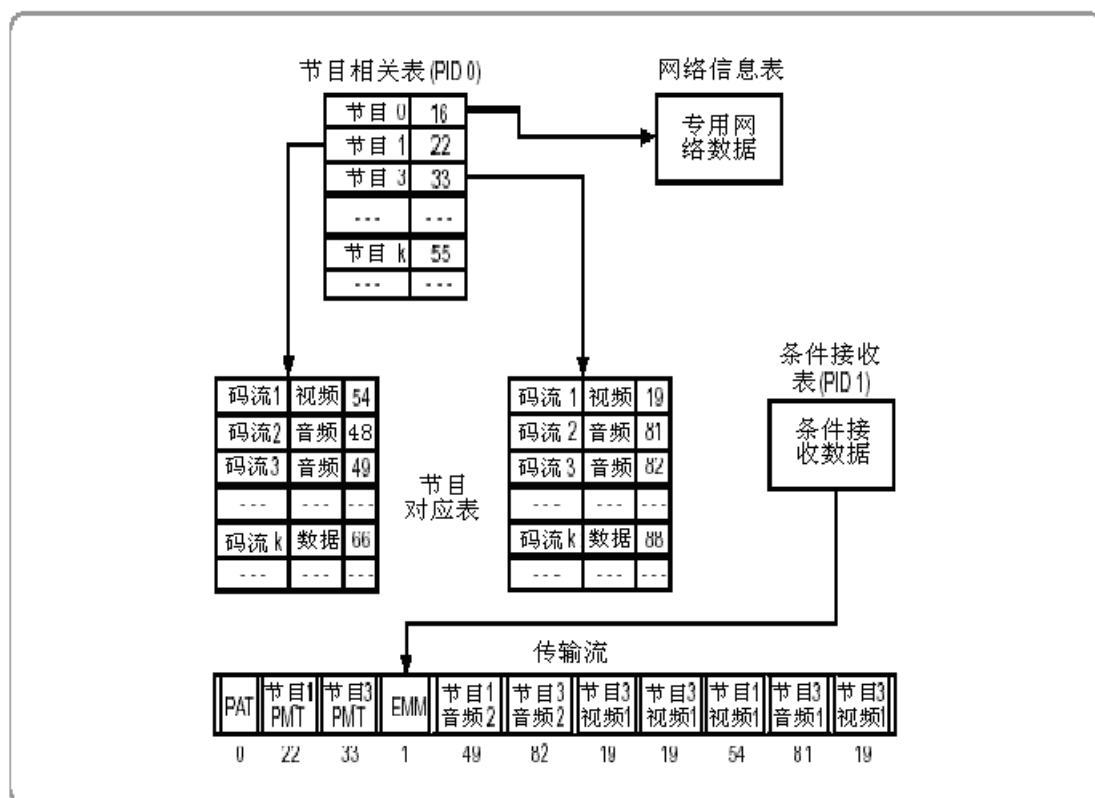


图 8—3

PSI 由具有唯一 PID 的数据包所传送。其中有些 PSI 已经标准化了，有些 PSI 由节目关联表 (PAT)、条件接收表 (CAT) 和传输流描述表 (TSDT) 所规定。这些数据包应当周期性地放置在每个传输流中。PAT 的 PID 总为 0，CAT 的 PID 总为 1，而 TSDT 的 PID 总为 2，以及空包的 PID 为 8191。以上这些数值是唯一的 PID，并由 MPEG 标准所固定。在解复用时，解复用器应当通过访问相对应的表来确定其余的 PID。不过，在使用 ATSC 和 DVB 中的 PID 时有些限制。在这方面（和某些其它方面），MPEG 和 DVB / ATSC 并不是可以完全互换的。所有的 DVB 和 ATSC 传输流都应当符合 MPEG-2 标准 (ISO / IEC 13818-1)，但是并非所有的 MPEG-2 传输流都符合 ATSC (A / 65A) 或 DVB (EN 300 468) 标准。

传输流中的节目都列在节目关联表 (PAT) 数据包 (PID=0) 中，它含有每个 PMT 包的 PID。PAT 中的第一项是节目 0，是为网络数据而保留的，它含有网络信息表 (NIT) 包的 PID。NIT 的使用在 MPEG-2 是可选项，但在 DVB 中则是必须的。

授权控制信息 (ECM) 的 PID 和授权管理信息 (EMM) 的 PID 均包含在条件接收表 (CAT) 数据包 (PID=1) 中。

如图 8-3 所示，属于同一节目的视频、音频和数据基本码流的 PID 均在节目映像表 (PMT) 表的数据包中。每个 PMT 包通常都有自己的 PID，但在 MPEG-2 中对此未作要求。每个 PMT 中的节目数由各个 PMT 唯一地确定。

对于一个给定的网络信息表 (NIT)，不仅含有它所在的传输流本身的信息，而且也含有能被同一解码器接收的其它传输流的信息，例如，通过调谐以接收不同的 RF 频道或者调整天线以对准不同的卫星。在 NIT 中列出了许多其它传输流的相关信息，其中每个传输流都有一个描述符，在描述符中指定了无线电频率和轨道位置等有关内容。在 DVB 中，还包括有附加元数据即 DVT-SI，并将 NIT 作为 DVB-SI 的一部分，相关内容将在第十章——DVB 和 ATSC 概论中加以介绍。在作一般性讨论时，还是使用 PSI / SI 这一术语。

当解复用器首次接收一个传输流时，它应当在传输流包头中寻找数值为 0 和 1 的 PID，在所有 PID 为 0 的传输流包中都含有 PAT，在所有 PID 为 1 的传输流包中都含有 CAT 数据。

解复用器通过读取 PAT，寻找 NIT 的 PID 和各节目关联表（PMT）的 PID。找到 PMT 后，解复用器就能发现各基本码流的 PID。

因此，如果需要对一特定节目解码，必须要首先查询 PAT，然后查询 PMT，以找出该节目中所有基本码流的 PID。如果节目被加密，还必须访问 CAT。没有 PAT 就不可能解复用，而锁定速度决定于 PAT 的发送频率，在 MPEG 中，规定 PAT 包和与其相关的 PMT 包之间的最大发送间隔为 0.5 秒。在 DVB 和 ATSC 中，NIT 保留在具有特定 PID 的传输流包中。

第九章 数字调制

MPEG 系统对视频、音频和其它数据进行编码并对其数据流打包。如果是用于存储，那么只需将数据流记录在硬盘上或 DVD 上既可。但如作为其它应用，则需要将 MPEG 数据以某种传输流的形式从一个地方传送到另一个地方或传送到许多地方。电视就是利用有线、卫星或地面传输等各种形式将信号传送到分配和广播系统。这些传输技术都需要把数据调制在一些载波上。本章将就为实现上述目的而采用的各种数字调制方案作以简要介绍。

本章仅就调制技术进行讨论。要使数据流适合于传输，还必须作其它处理，它们均与所选择的调制方式和通道特性有关，这些内容将在第十章予以讨论。

9.1 调制原理

一个连续的载波信号，只有当它被欲传送的信息以某种方式调制后，它才能传输信息。对载波的调制可以有三种形式：即幅度调制、频率调制或相位调制。当然，频率和相位是密切相关的。尽管一般都把它们看成是不同的调制方式，但它们之间的区别是不太明显的。某些“频率调制”实际上是通过“相位调制”来进行的。

9.2 模拟调制

在模拟调制中，载波的幅度或相位是按照视音频信号或视频信号的幅度而改变（被调制）的，而且通常是按照线性规律而改变。模拟系统中也可以使用相位调制，最明显的例子就是在 NTSC 和 PAL 电视系统中色差信号对彩色副载波的调制。这个例子是很有用的。我们可以设想彩色信息有两种形式。第一种形式是它可用相位角（对应于色调）和幅度（对应于饱和度）来表示，另一种形式是用两个色差信号（即 $B-Y$ 和 $R-Y$ ）的数值来表示。如果我们再进一步把这两个色差分量当作坐标图中的 x 和 y ，那么这两种形式就是密切相关的，如图 9-1 所示。该图形与我们熟知的矢量显示很相似。

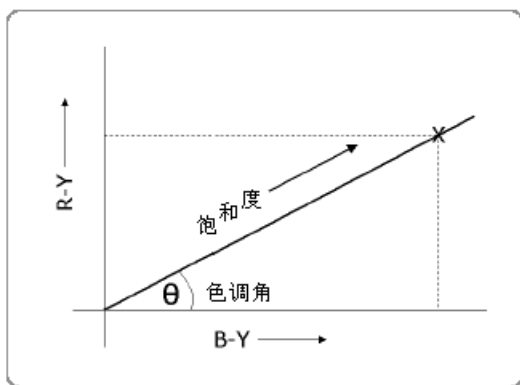


图 9—1

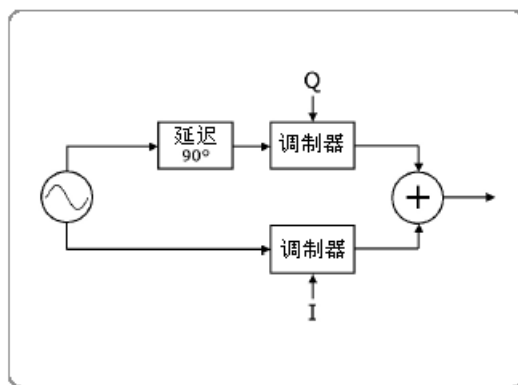


图 9—2

9.3 正交调制

图 9—2 表示我们如何用两个不同的信号来调制同一个载波，这种方法即所谓“正交调制”。它将一个载波划分为两个路径，其中一路载波被延迟，延迟时间为该载波的 $1/4$ 周期。这路载波与另一路载波频率相同，但相位相差 90° 。用两个适当的信号分别对这两路载波进行幅度调制，然后将两路调制后的载波相加，最后得到的是一个已调制的信号，它的幅度和相位由这两个调制信号所决定。

解调几乎是一个相同的过程。被接收的信号送往两个解调器，在每个解调器中，将所接收的信号与本机振荡信号相乘，本机振荡信号应与被接收信号同步，两个振荡信号相位差为 90° ，每个解调器分别恢复正交调制中的两个调制信号。

9.4 简单数字调制系统

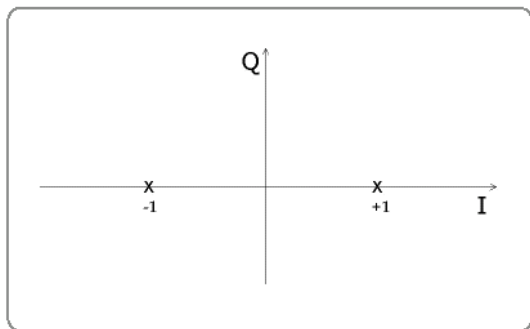


图 9—3

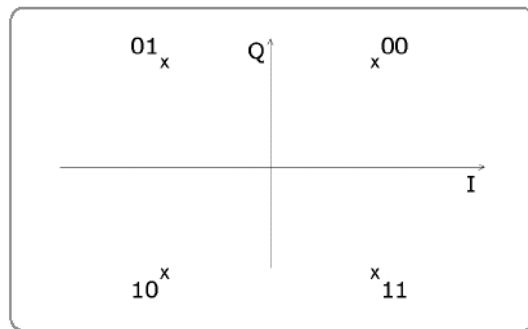


图 9—4

大多数调制系统使用正交调制中的某些形式，通常用 I 和 Q 来表示相互正交的两坐标轴。有时也仅使用一个调制轴。

所有的数字调制方案都兼顾了带宽利用率、传输可靠性和方案复杂程度等方面。符号率 (symbol rate) 是决定传输信号带宽的主要因素。符号率是调制被改变的速率，因此它和调制信号的频率是相同的。

有些简单的数字调制系统每个符号只能携带一个比特信息，换句话说，每个符号在两种可能状态中只代表其中一种状态，即在二进制中或代表零，或代表 1。在这种情况下，系统的比特率和符号率是相同的。不过，在其它系统中，每符号具有多种可能的状态，这样每符号就可以传送一个比特以上的信息。一般而言，状态数是 2 的乘方，这样，系统的比特率就是符号率的整数倍。

通常以调制类型来给数字调制系统命名，同时在前面冠以代表每符号状态数的一个数字。例如，4QAM 表示每符号 4 个可能状态的正交幅度调制。四个状态就可以传送 2 比特信息（00，01，10，11），所以 4QAM 系统的比特率是该系统符号率的两倍。

最简单的数字调制系统每符号传送一比特信息。每个符号具有两种可能的状态，即二进制中的零和二进制中的 1。这种状态可以用调幅、调频或调相来实现，但最常用的是相位调制。

二进制频移键控（BFSK 或 2FSK）是使用一个载波频率代表二进制零，另一不同的载波频率代表二进制 1。有时频率差很小，可以通过相位调制器来实现。

二进制相移键控（BPSK 或 2PSK）是使用一个幅度恒定的载波相位代表二进制零，相反的相位（180 度相移）代表二进制 1。通常用星座图来代表某一符号的几种不同的可能状态，星座图可以用来表示 I 和 Q 调制信号允许取值的几种不同的组合。图 9-3 是 BPSK 星座图示意，它是非常简单的，只使用了一个坐标轴，而且只允许取两个值。

上述调制系统的传输十分可靠，接收机只需凭有无信号（或凭信噪比的大小）就足以在每符号的两种可能状态中确定其中的一种状态。但是，这样的系统却不能使通道频谱得到有效的利用，所使用的传输带宽与比特率相同。不过，它可以用于十分恶劣的传输路径中，例如太空遥测。

9.5 相移键控

有关 BPSK 或 2PSK 如前所述。PSK 调制的另一种形式是同时使用 I、Q 轴，即正交相移键控（QPSK）。QPSK 是最常用的一种调制形式，每个坐标轴使用两个数值。其星座图如图 9-4 所示。在 QPSK 中，每符号具有四种可能的状态，因此每符号可以携带两比特信息。图 9-4 是以二进制数值表示的一种可能的状态映象。QPSK 广泛用于卫星通讯中。

8PSK 不太常见，但它也用于卫星通讯系统，特别是在日本。其星座图如图 9-5 所示。8PSK 每符号传送 3 比特信息，即比特率是符号率的 3 倍。

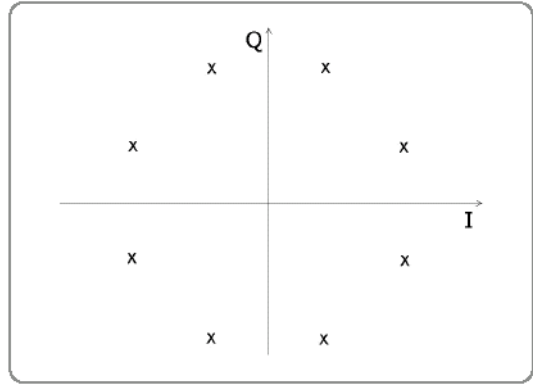


图 9-5

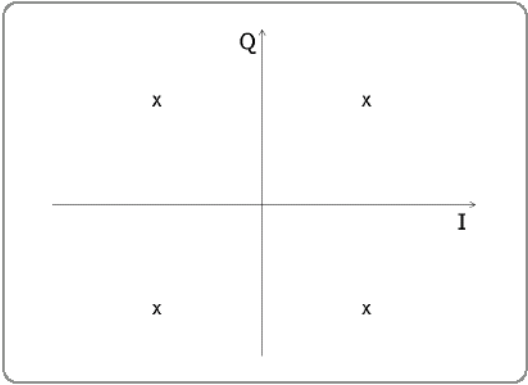


图 9-6

9.6 正交幅度调制（QAM）

正交幅度调制是许多传输系统的基础。它在调制时不仅要使用 I、Q 轴，而且与特定的取值有关。在每个坐标轴上允许有两个或更多的幅度电平。

最简单的正交幅度调制是 4QAM，它在每个坐标轴上只使用两个数值，每符号提供四种可能的状态。其星座图如图 9-6 所示，该图看起来与 QPSK 相同，并且也是每符号传送两比特信息。

16QAM 在每轴上使用四个数值，它可以提供 16 种可能的状态。16QAM 系统每符号传

送 4 比特。如果每轴允许使用 6 个数值，那么就有 36 种可能的状态。实际上，只用 32 个状态就可以传送 5 比特，还有四种状态在 32QAM 中并不使用。16QAM 星座图如图 9—7 所示，32QAM 星座图如图 9—8 所示。在 32QAM 中，位于四角的状态是不使用的，它们代表最高的信号幅度，因此，也就是最高的发送功率。

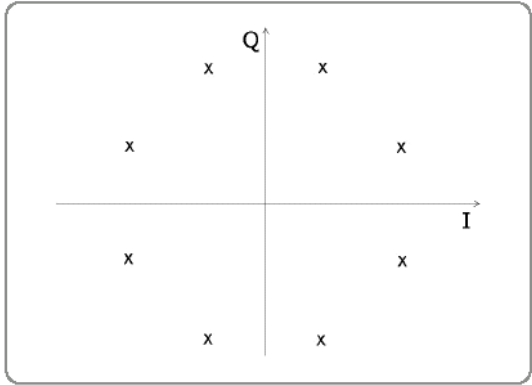


图 9—7

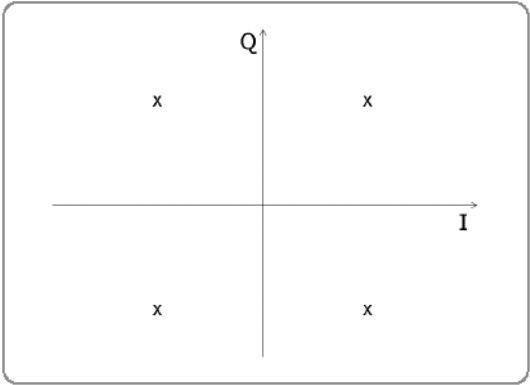


图 9—8

图 9—7 和图 9—8 有助于观察如何在比特率和传输可靠性之间取得折衷。在噪声和抖动存在的情况下，32QAM 中的各个状态（具有相同的发送功率）间的间隔愈小，那么使解码发生错误的可能性就要增加。换句话说，每符号的可能状态数愈多，在给定的误码率的情况下，对信噪比的要求就愈高。

如果可以保证有良好的信噪比，那么就可以进一步增加星座密度。在 64QAM 中，每轴允许使用 8 个数值，这样每符号就可以传送 6 比特。64QAM 是世界上应用最广泛的有线系统调制方案，它在充分利用已有的有线基础设施和传输可靠性之间取得了很好的平衡。256QAM 在一些最近开发的有线电视系统中也获得了应用，它在每个调制轴上允许取 16 个数值，每符号传送 8 比特信息。

9.7 残留边带调制（VSB）

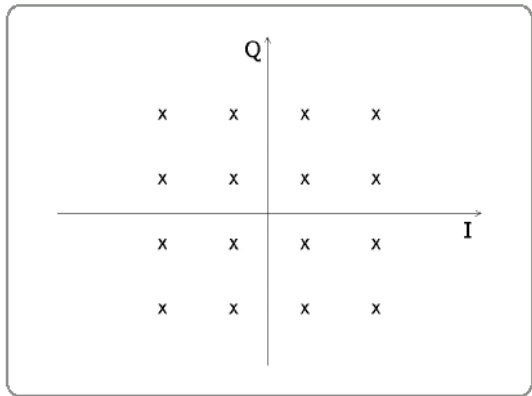


图 9—9

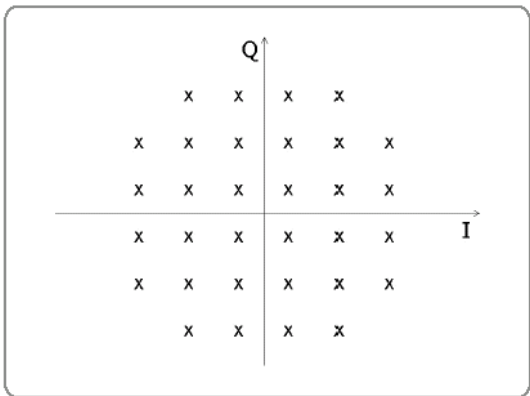


图 9—10

当某一载波被调制时，在这个载波频率的高端和低端会有边带产生。例如，在符号率为每秒 3 兆符号的 QAM 系统中，其高端和低端的边带大约宽 3MHz，这样所需要的正常频道带宽为 6MHz。

为了恢复调制信号的幅度和相位信息（或正交调制的两轴），接收机就应当恢复两个边带。

残留边带系统在传输之前消除了一个边带中的绝大部分频谱，这样就只有一个调制轴可以被恢复（或者说，另一调制轴用来抑制不需要的边带）。2VSB 的星座图与 BPSK 相同。图 9-9 为 4VSB 的星座图，图 9-10 为 8VSB 的星座图，它们每符号分别传送 2 比特和 3 比特。

8VSB 调制方案在美国被 ATSC 数字电视标准所采纳。最初建议是 4VSB，每符号传送 2 比特。不过，后来发现在 8VSB 中，结合使用格形编码（参见第十章）和卷积内码纠错，可以提供相同的数据率，并具有更好的信噪比性能。

9.8 编码正交频分复用（COFDM）

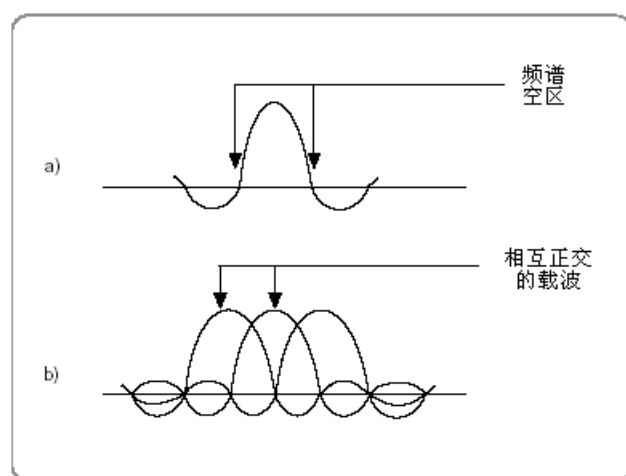


图 9-11

在以上调制系统中，是将基带信号提供给调制器，而调制器是工作在单一载波上以产生传输边带信号。宽带调制系统的另一方案是使用许多个窄带载波，各载波的间隔要严格符合规定。图 9-11a 表示在一数字已调载波中，左右两边各有一个频谱空区。将另一相同的载波可以放置在这里而不会形成干扰，因为这两个载波是相互正交的，如图 9-11b 所示。这就是 OFDM（正交频分复用）的工作原理。在实用中，使用的是所谓编码正交频分复用，它采用了维特比卷积编码，能在非理想通道的状态下明显地提高传输性能。维特比卷积编码将在下一章予以介绍。COFDM 用于数字视频（DVB）广播地面系统中，即 DVB-T。

在 OFDM 系统中，所使用的载波数是非常多的。DVB-T 可选用 1705 或 6817 个载波（即所谓 2K 系统或 8K 系统）。由于分配给每个载波的带宽是相当窄的，相应符号率也会降低，传输每个符号的时间长度就会增加。这正是 OFDM 能够克服多径干扰的主要原因。

在单载波系统如 8VSB 中，较高的数据率意味着符号时间很短。在 ATSC 数字电视系统中，每秒大约要传输 1 千 1 百万个符号，那么每符号的时间间隔会小于 100ns。这意味着即使很短的多径延迟也会产生符号间的干扰，因为代表一个符号的信号经延迟后将在接收下一个符号的时间内到达。

与上述不同，OFDM 系统使用了数千个载波，一个符号的间隔时间大约为几百毫秒（与数据率、载波数和使用的调制方式有关）。同时采取给每个符号加以“保护带”——人为地延长符号的所需时间，从而实际上可以消除符号间的干扰。当然，这会降低符号率，但影响相对较小。例如，如果正常符号间隔是 200μs，保护带是 50μs，符号率仅降低 20%——但符号间干扰的消除使高阶星座的使用成为可能，它增加的符号率也许要超过因设置保护带而

造成的损失。不过，上述技术在宽带单载波系统中是不实用的。作为一个例子， $50\mu\text{s}$ 的保护带加在 100ns 的符号时间上，这将使数据率降低的百分比是多少！

对多径干扰的改善也使 COFDM 系统非常适合于单频网络。在单频网络中，有两个或多个相互同步的发射机发射着同样的信号。接收机所接收的信号可能是在不同的时刻收到来自多个发射机的信号，这与接收机所处的位置和天线系统有关。如果接收信号路径完全不同，那么主信号比次信号可能要强得多，这样的码间干扰是很小的。如果接收信号路径和信号强度的差别很小，那么保护带就可以防止符号间的干扰。

COFDM 是一种很灵活的系统，它可以经“调整”以适应各种传输需求，但是，传输可靠性的增加是以数据率的降低为代价的。有关单载波和多载波哪个更好的问题也存在着很多争议，但是一般认为，在简单的通道条件下，如果以相同的数据率和相同的复盖区域来比较，COFDM 的所需功率要略大于 VSB。同时 COFDM 在发射功率上有着较高的峰值—平均值比，这有可能对其它服务有较大的干扰。

不过，许多人认为，在复杂的多径传输环境中，如城市中心的“都市峡谷”，COFDM 能够提供更可靠的接收质量。

9.9 综合业务数据广播 (ISDB)

综合业务数据广播 (ISDB) 是日本为数字电视业务而开发的系统，它使用了多种调制方案。它支持多种级别的分层系统。例如，它既可为非常困难的移动环境提供低数据率的接收服务，又可边缘地区静止接收提供中等数据率（标准清晰度）的服务，同时还可为良好接收条件提供高数据率（可以用于 HDTV）服务。

当前在日本使用三种 ISDB 调制系统：

9.9.1 ISDB-S 卫星系统

ISDB-S 于 2000 年 12 月投入运行。在 ISDB-S 系统中，两个广播机构可以共享一个卫星转发器。用于广播的卫星也可称为 BS 数字卫星系统，用于通讯的卫星可称为 CS 数字卫星系统。

共享转发器的广播机构之间如达成协议，那么一个转发器总共可使用 8 个传输流。比特率总和决定于转发器的带宽和所使用的调制方式。例如，一个 34.5MHz 的转发器，包括前向纠错在内，最大数据率为 56.610Mbit/s 。

在分层调制中，允许在 48 个数据包帧中按传输流包改变调制模式。每个包均被指派一个调制时间段 (modulation slot)。支持四种调制模式，即 BSPK (1/2)，QPSK (至 7/8) 和 TC8PSK，时间段数按照所使用的模式而改变。

9.9.2 ISDB-C 有线系统

该系统的主要特征是在一单个 64QAM 载波上发送多个传输流。开发该系统的目的是能够有效地转发 ISDB-S 信号上所携带的信息。在一路 BS 数字载波上，所传送的最大数据率典型值为 52.17Mbit/s ，而一路 64QAM / 6MHz 信号的数据率为 29.162Mbit/s 。因此，应当至少使用两个有线电视频道来转发一路 BS 载波信息。完整的 BS 数字业务是由四套广播设置所组成，包括保护带在内，共需带宽约 174MHz 。传送这种信息要使用 ISDB-C 8 个有线频道，如果使用每载波一路传输流的传统有线传输频道来传送，则需多达 29 个频道。

有 52 个调制时间段，另有一个时间段用于同步包头，即 TSMF (传输码流复用帧)。

9.9.3 ISDB-T 地面调制

ISDB-T 频道被分为 13 段 (一般每段带宽为 400 至 500kHz)，每段均可使用单独的

COFDM 传输方式。在分层结构中，每层均可单独选择影响传输可靠性的所有参数（载波数，保护带长度，调制类型，卷积编码）。例如，最可靠的传输段可以使用较长的保护带，采用 QPSK 调制方式和 1 / 2 卷积编码。电平最高段可以使用较短的保护带，采用 64QAM 调制方式和 7 / 8 卷积编码——它的数据率是可靠段的许多倍。

中心段可用于部分接收，它是为窄带接收机设计的，窄带接收机只能接收这个段的信号。

在标准的 OFDM 中，全频道带宽代表一单个层。所使用载波的间隔带宽是某一确定频率的倍数。在 ISDB-T 中，频道带宽为 5.6MHz，最多可划分为 13 段，每段带宽为 429kHz。通过传送具有不同参数的各个 OFDM 段组来实现 ISDB-T 的分层传输。由这些段组构成层。在非分层调制模式中，所有的 13 段均使用相同的调制方案。

9.9.4 ISDB 小结

ISDB-S 提供的方法是将不同的调制模式应用到多个传输流中，并在带宽为 34.5MHz 的频道内以单个载波的形式传送多个传输流。

ISDB-C 提供的方法是，在单个 6MHz 的频道内以单个载波的形式并使用相同的调制模式来传送多个传输流。

ISDB-T 是在一单个 6MHz 频道内涉及到同一传输流的三种传输模式。

第十章 DVB 和 ATSC 概述

MPEG 压缩已经在广播中获得了广泛的应用，并在将来显得日益重要。本章将讨论数字电视广播中的其它要求以及两种主要的 DTV 标准的工作过程。

10.1 引言

ATSC（先进电视制式委员会）是美国制定地面数字广播标准的一个组织机构。DVB 指的是数字视频广播项目和由 DVB 项目所确立的标准和应用。DVB 最初是欧洲的一个项目，但它建立的标准和规范也被世界上其它地区所接受。这些标准和规范包括了所有的传输介质，例如卫星、有线和地面广播。

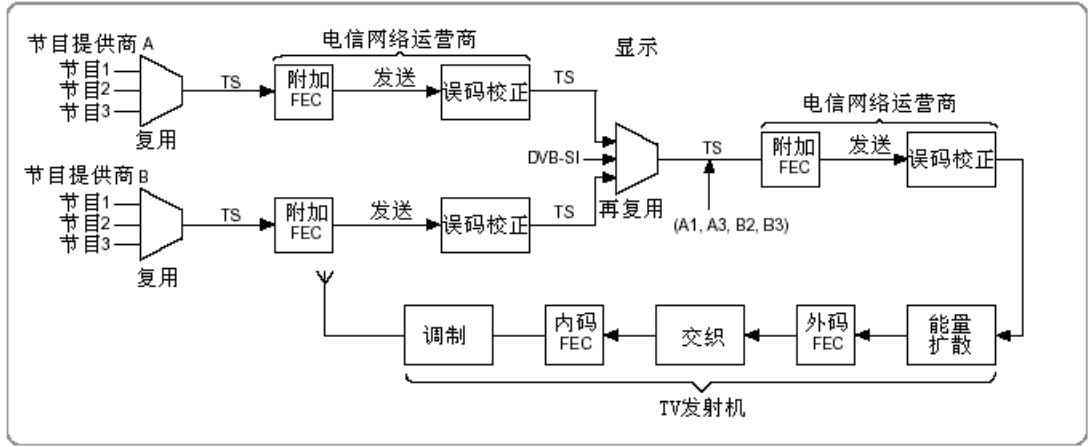


图 10-1

数字广播有着不同的分配和传输需求，如图 10-1 所示。广播业者提供含有若干个电视

节目的传输流。但传输流本身不具备防护误码的能力，特别是已压缩数据中的误码，其后果是十分严重的。传输流应当给发射机、卫星上行链路和有线前端提供无误码的内容。在这里，所谓无误码指的是误码率为不大于 1×10^{-11} 。这项任务通常是由电信运营商来完成的，电信运营商将按需要使用误码校正的附加层（误码校正的具体方法与传输通道有关）。这个误码校正层对接收端应当是透明的。

某一发射机或有线运营商可能并不需要一个传输流中的所有节目。通过再复用可以接收所需要的传输流和频道，并将它们编码到一个单独的输出码流中。这种配置是可以动态地改变的。

数字域的广播是将传输流传送到用户的整个环节组成的。所使用的传输通道无论是有线，还是卫星或地面，其过程十分相似，都是用一种标准的方式将描述传输的元数据编码到传输流中。在 DVB 中，这种元数据称为业务信息（DVB-SI），它包括图文电视和节目细节，不管它们是来自自身的复用还是其它的复用。

在广播中，几乎无法对信号质量和噪声或干扰进行控制。这就需要前向纠错（FEC）层的一些形式。与长途电信网络运营商所使用的 FEC 不同，后者是具有专利的（或按欧洲电信标准学会（ETSI）标准化，ETSI 对 SDH 和 PDH 网络的 DVB 传输作了规定），广播中使用的 FEC 标准则必须使接收机能够对它进行处理。

很明显，误码校正的引入会增加发射机或有线传输的比特率。遗憾的是，可靠的、经济的无线电数据发送和有线数据传输需要的不仅仅是串并转换的数据。一个实用的系统必须要经过通道编码。

10.2 再复用

再复用是一件相当复杂的工作，因为再复用器必须将来自其它复用器某部分的传输流组合起来，重新输出一个顺应性比特流。利用节目关联表和节目映像表，找出所需节目的 PID，这样才能获取某一给定输入传输流的所需数据。在再复用器的两个输入传输流中，有可能有相同的 PID，因此，必须变更一个或多个基本码流的 PID。传输流包头应当继续传送节目时钟基准（PCR），以使终端解码器能够再生 27MHz 时钟。在新的复用器中，含有 PCR 数据包的位置可能会有所不同，再复用器也应当对 PCR 数值进行编辑以在时间轴上反映出它的新位置。

同时还需要对节目对应表和节目相关表进行编辑，以反映新的传输流结构，条件接收表（CAT）也是如此。

如果被选节目流比特率的总和小于输出比特率，那么再复用器还应创建具有相应 PID 的填充数据包。但是，如果再复用器的输入传输流是来自于统计复用器，有可能会使新传输流的瞬时比特率超出通道容量。如果从几个不同的传输流中选中了若干个节目，而这些节目又包含着较大的瞬时熵时，这种情况是可能发生的。在这种情况下，唯一的解决办法是在一个或多个比特流中再次进行压缩并创建新的、较短的系数，从而降低比特率。

10.3 业务信息（SI）

在将来，数字传输会给观众提供非常多的节目、图文电视和用户所需要的业务，这些内容可能分散到大量不同的传输流中。无论是观众还是 IRD（综合接收机 / 解码器）都需要得到帮助，以显示出哪些内容是需要并输出被选择的业务。这就要用到元数据。元数据的功能超出了 MPEG-PSI（节目专用信息）的功能范围，它可称为 DVB-SI（业务信息）。DVB-SI 包括了 NIT，而 NIT 在 MPEG 传输流中是可选项。

DVB-SI 作为附加的传输包嵌入在传输流中，它具有唯一的 PID，携带有 IRD 的技术信息。DVB-SI 也包含有电子节目指南（EPG）信息，例如节目的种类，节目的时间和所

在的频道，哪些地区可以接收。节目也可以是分级的（rated），以让父母亲作出判断。

作为 MPEG-PSI 的补充信息，DVB-SI 应包括以下各表：

- ▶ 网络信息表（NIT）：它是在一个传输流中用以描述许多传输流的信息。NIT 所传送的信息涉及到复用器的物理结构，经由给定网络所携带的传输流和网络本身的特性等。通过最初的网络 ID 和 NIT 中的传输流 ID，就可以识别传输流。
- ▶ 业务描述表（SDT）：DVB 传输流中的每项业务都具有业务描述符，这些描述符组成了业务描述表。传输流中的业务可以是电视、无线电广播或图文电视。业务描述符中还包括有业务提供商的名称。
- ▶ 事件信息表（EIT）：EIT 在 DVB 中用于描述节目名称、起始时间、节目持续时间等。
- ▶ 时间和日期表（TDT）：TDT 将 UTC 时间和日期标记嵌入到传输流中。

DVB-SI 还定义了其它可选表，包括：节目组相关表（BAT），运行状态表（RST），时间偏置表（TOT）和填充表（ST）。

和 DVB 一样，ATSC 使用了 MPEG-2 专用区域表（private section table）以定义几个新表。有关这些新表的设置由 ATSC 在 A / 65A 中定义，它们是必须遵循的，是节目和系统信息协议（PSIP）的一部分。ATSC PSIP 也是对 MPEG-PSI 的补充，它应当包括以下各表：

- ▶ 地面虚拟频道表（TVCT）：定义嵌入在传输流中的 MPEG-2 节目，该传输流携带有 TVCT。
- ▶ 原版指南表（Master Guide Table, MGT）：定义传输流中所有其它 PSIP 表的节目类型、包识别符和版本信息，但系统时间表（STT）除外。
- ▶ 分级区域表（Rating Region Table, RRT）：对由传输流中的内容咨询描述符提交的 TV 双亲准则系统作出定义。
- ▶ 系统时间表（STT）：定义当前日期和时间。
- ▶ 节目信息表（EIT-n）：对节目信息表中的前四个节目（EIT-0, EIT-1, EIT-2, EIT-3）进行描述，共有 12 个小时的电视节目，每个节目 3 小时，包括 TVCT 中列出的所有虚拟频道。

10.4 误码校正

由于长距离传输中有很多因素是无法控制的，因此必须进行误码校正。在某些系统中，误码校正是很充分的，因为在出现误码时可以请求重新发送。显然，这个办法对于实时信号如电视是行不通的。另一个办法是使用前向误码校正（FEC），它是将足够的额外比特（即所谓冗余）添加在数据上，以便解码器进行实时校正。

现代系统中使用的 FEC 通常是基于里德-索罗门（R-S）编码。对它进行全面的讨论不在本书范围之内。简而言之，R-S 编码是在数据上添加冗余以形成编码字，当每个符号作为两个联立方程的最小项使用时，如果没有误码，则代数和（或校验子）总是为零。这种获取零状态而不考虑数据使得校验较为容易。在传输流中，在附加误码校正数据之前的包长总是 188 个字节。附加 16 个字节的 R-S 冗余数据后包长为 204 个字节（在实用中，即使在没有使用 FEC 的情况下，也可使用 204 个字节的包长。当 FEC 被添加或删除时，可以使用 16 个填充字节以避免码流的重新计数）。

当校验子不为零时，求解联立方程会得到误码校正所需的两个数值：误码的位置和误码的性质（误码值）。然而，如果误码的大小超出添加冗余量的一半时，则误码就无法校正。实际上，在典型的传输通道中，信号质量是呈统计性的，这意味着有时可能是因噪声而引起的单个信号比特错误，有时又可能突然发生大量的比特错误，这种情况的出现可能是因闪电或电子设备的干扰所引起。

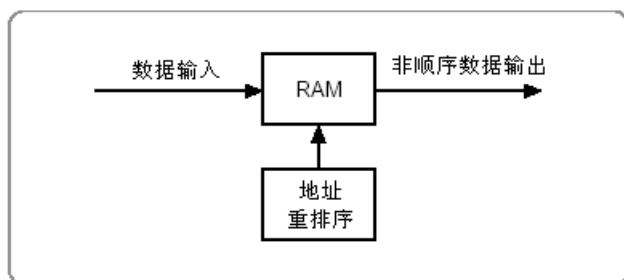


图 10-2

要做到每个编码字不受脉冲串干扰的影响，在经济上是不合算的，因为这类脉冲串干扰并不经常出现。为此，可采用交织技术来加以解决。如图 10-2 所示，在使用交织时，信源数据经过了 FEC 编码，在数据传送前，先将其送入 RAM 缓存器。图 10-3 给出的是一种可能采用的方法，数据按照行序进入 RAM，但读出时却按照列序，然后将重新排序的数据发送出去。在接收端，再利用另一个 RAM 将数据恢复到原先的顺序，这一过程为去交织。通过交织过程，可以使通道中的一串误码在去交织后分散为大量的但却是单个符号的错误，而单个比特误码则是比较容易校正的。

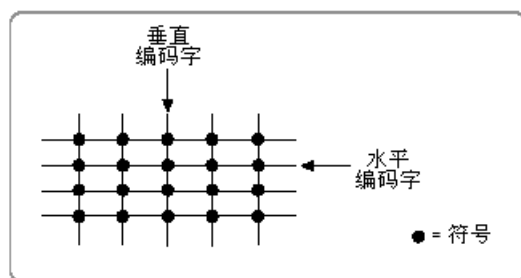


图 10-3

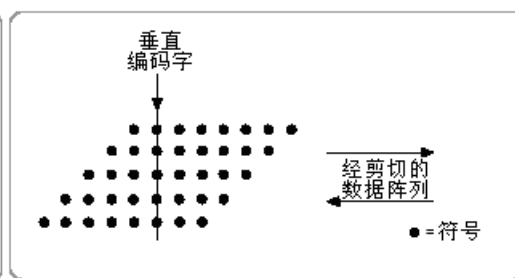


图 10-4

如果一串误码数进入最大可校正的范围，则系统较易发生随机比特错误，使编码字无法校正。解决的办法是在交织之后使用内码并在去交织之前予以校正，这样就可以防止在进入去交织存储器时出现的随机误码。

如图 10-3 所示，在利用块交织结构作交织处理时，其结果是一种乘积编码。图 10-4 表示，交织也可以被卷积，在卷积中，每行给予不同的延时，数据阵列被剪切。卷积交织或交叉交织的优点是在交织和去交织时只需使用少量的存储器。

10.5 信道编码

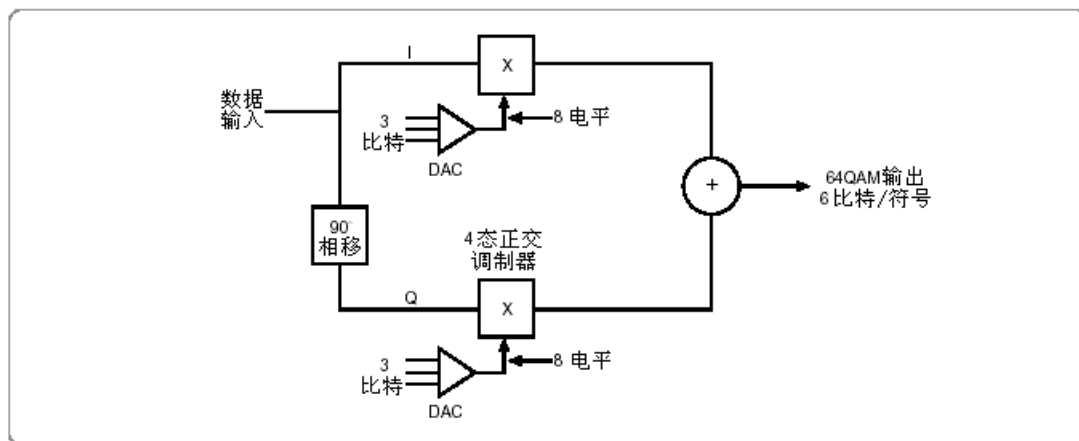


图 10-5

在信号的传输过程中，直接使用原始的串行二进制数据是不合适的。这有几个原因。其一，相同比特的运行容易引起直流偏移和比特时钟的丢失。其二，无法对信号频谱进行控制，传输数据信号的所需频带太宽。在实用无线电通讯和有线系统中，必须要使用被称为信道编码的调制方案。数字调制技术已在第 9 章介绍过。图 10-5 表示的是使用 64QAM 调制器进行信道编码的工作原理。

在上述方案中，传输信号的频谱是信号相关的。信号频谱中的某些部分可能含有较高的能量，它会对其它业务形成干扰，而另一部分频谱却只有较少的能量，但它又容易受到外来的干扰。在实用中，有必要进行随机化处理，以去掉数据传输频谱中的相关性。图 10-6 所示为数据的随机化处理即能量扩散过程。在图中，在串行数据被送入调制器之前，将一个伪随机序列与串行数据相加。这样，数据传输频谱类似于噪声频谱而带有相对平稳的统计特性。很明显，如图所示，在接收端应当减去与发送端相同的伪随机序列以恢复原始数据。随机化处理不能用于同步模式，否则就不能对其进行检测。

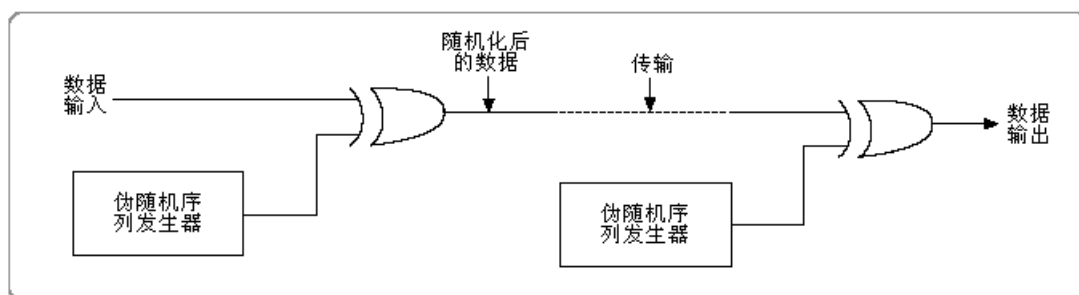


图 10-6

10.6 内码编码

FEC 系统的内编码，它是用来防止因随机误码而造成交织性能的下降。合适的内编码应能明显地增加传输的信噪比，从而防止这类误码的产生。在格子编码中，格子编码和多电平发信一起使用，若干个多电平符号可以形成一个组。由一特定的符号组所形成的波形可称为一个网格 (trellis)。如果每个符号都具有 8 个电平，那么三个符号就可能有 512 个网格。

在格子编码中，对数据的编码应当使只有确定的网格波形才能代表有效的数据。如果只有 64 个网格代表无差错的数据，那么每符号可以发送 2 比特数据。余下的比特是一种冗余的形式，因为除了正确的 64 个网格之外，其余的网格都是错误的。如果在接收到的一个网

格中，符号中的某一电平因受噪声的影响而模糊，但这种模糊是可以解决的，因为正确的电平应当是能够给出有效网格的电平。这种技术即所谓最大似然解码。

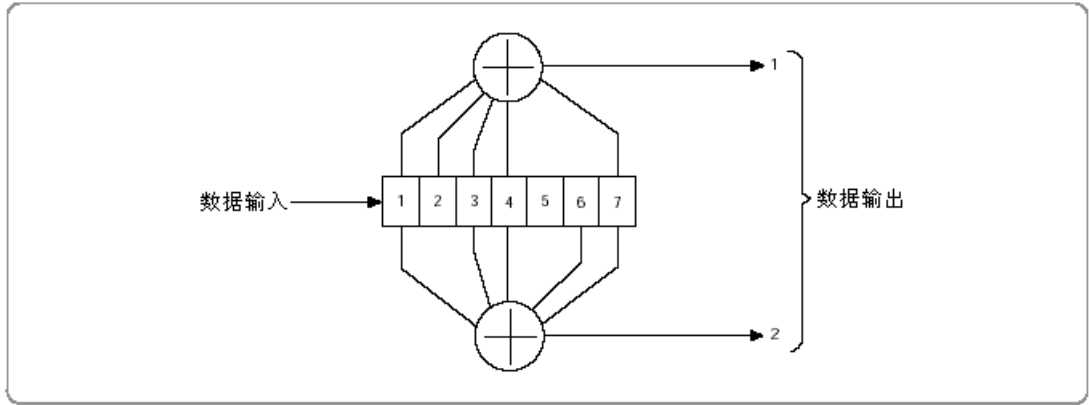


图 10-7

在 64 个有效网格中，其差别应当尽可能地大，这样，在信噪比较差的情况下仍能使系统继续工作。如果格形编码出错，外编码将会予以校正。

在 DVB 中，可以使用维特比卷积编码。如图 10-7 所示，数据经过交织后，即被送入移位寄存器，移位寄存器有两路输出，分别代表输入数据的不同奇偶校验，这样可使比特错误得以校正。很明显，每个输入比特都会有两路输出，因此，可把图 10-7 中所示的编码器称为 1/2 编码器。比例在 1/1 和 1/2 之间的任何编码器仍能发送原始数据，但冗余量会发生变化。不能发送完整的 1/2 输出称为击穿（puncturing），它允许在比特率和校正能力之间取得任何所需要的平衡。

10.7 数字发射

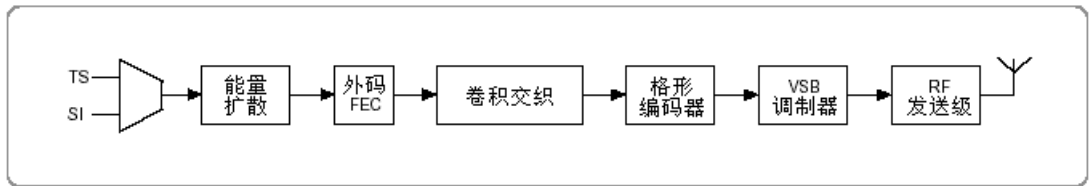


图 10-8

图 10-8 所示为 ATSC 数字发射机的基本组成。描述传输的业务信息添加在传输流中。对传输流进行随机化处理，进入 R-S 外码校正编码器，这时数据上添加了冗余。而后进入卷积交织处理，对数据比特重新排序，以使传输流中的相邻数据在此后的传输流中不再相邻。再进入内码格形编码器，它可为残留边带调制器提供多电平信号。

图 10-9 所示为 DVB-T 发射机。与上述相同，添加业务信息后经随机化处理以使能量扩散。在交织之前添加外码 R-S 校验符。在交织后进行内码处理，将编码后的数据送入 COFDM 调制器，调制器的输出经上变频后产生 RF 输出。

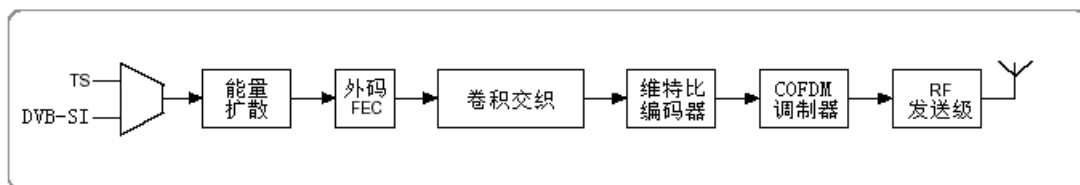


图 10—9

在接收机中，首先要提取比特时钟，用它来控制整个接收系统的定时。而后是信道编码的反过程，以获取原始数据及传输差错信息。用内码来校正随机误码，并识别较大的差错以在去交织后有助于通过外码作进一步的校正。除去随机信息后就可得到原始的传输流。接收机应能识别 PAT、业务信息（SI）以及通过 PAT 找到 PMT，以使用户了解传输流中的可用信息，并可在复用器中选择所需要的节目。

第十一章 数据广播

前面几章是 MPEG—2 传输流的基本部分。与传统的模拟广播相似，MPEG—2 传输流的主要应用是传送已压缩的视频和音频码流。不过，MPEG—2 传输流的一个主要优点是它不仅能够传送视频和音频，而且还能够携带数据。当然，模拟电视系统的确也能携带数据，但将它的最大数据带宽与传输流中的可用数据带宽相比较，那就非常有限了。

本章将介绍 MPEG—2 数据传送的不同方法，以及一些地区将数据压缩到传输流中的各种方式，如 DVB，ATSC 和 ARIB（日本广播工业和商业协会）等。然后我们将讨论如何通过机顶盒把数据提交给用户（虽然严格说来，它并不属于 MPEG—2 的范围）。

11.1 数据广播应用

对使用传输流的数据广播来说，有许多种不同的应用类型，每种应用类型可能需要不同的数据类型并带有不同的定时需求。

例如，在因特网上发送的数据类型就非常不同于提供机顶盒固件刷新的所需信息。报价信息的非实时刷新与实况知识竞赛就有着完全不同的定时需求，在知识竞赛中，问和答都应当使用视频和音频准确地同步发送。

为了发送数据，MPEG—2 提供了大量的各种技术。技术的选择是在满足应用定时需求的同时也要兼顾带宽的优化（以及提供这项业务的成本）。

应用的类型大致可以按照它们的实时需求与交互的层次来进行分类，如以下所述。

11.1.1 节目相关数据

EPG（电子节目指南）能够给出传输流中所携带的有关电视频道或某个节目的信息，但 MPEG—2 的基本规范并没有对 EPG 作出规定。一些区域标准 ATSC、DVB 和 ARIB 使用 MPEG—2 专用表句法（例如 EIT）以提供有关节目的附加信息。这些表需要在规定的时间内发送出去，并给出有关业务或节目的起始时间、内容提要和其它信息。

然而这种级别的信息是很不够的，于是出现了许多专用的 EPG，这些 EPG 扩展了可用的信息量，还能提供诸如增强搜索、喜爱频道和其它增值业务等功能。这些增强的 EPG 通常永久地保留在机顶盒中，结合使用标准表信息和专用数据格式。通常这类节目信息对传输时间要求不高，因此可以使用较为廉价的低比特率技术。在许多情况下进行设计时，就将数据保存在机顶盒中，这样对用户而言是直接可用的，尽管实际总的传输时间相对较长。

11.1.2 机会数据 (Opportunistic Data)

在具有视音频码流的传输流中，它的全部带宽一般是很少能被全部使用，因为还需要处理再复用和表插入的一些复杂限制。机会数据系统可以使用有用数据来替换某些空包，从而使这种多余的容量得到充分的利用。当然，这种数据的带宽是不能保证的，而且数据量可能很小。因此它只能作为非实时应用。

能够使用这类数据的某些应用可能是文件传输，例如价目表等。这种应用的唯一要求是传输应当整夜进行，这样就可以使用较低的带宽，使廉价的机会数据业务得到充分的应用。

11.1.3 网络数据

在网络数据中，一种常见的应用是将广播传输流作为网络数据的载体。在最简单的情况下，卫星链路可以提供很高的带宽，从而将地理上两个不同的点联系在一起。实际上可用这种网络链接来支持任何类型的网络协议。

目前最常见的网络数据应用是使用 IP 协议的高速因特网下载，在这种情况下，通过一个低速的因特网连接向因特网服务器发出指令和页面请求。然后网络服务器经由卫星链路发送被请求的数据，其发送速度远高于 ISDN（综合业务数字网）或者普通的宽带连接。

如果只能使用一个慢速的反向信道，则可以采用另一种形式。在这种情况下，例如，可以按照需要将当前 TV 节目相关的数据通过被发送的信号而传送，观察者无须了解他们正在访问某一因特网的站点。

11.1.4 增强电视

在增强电视应用中没有反向信道，所需要的所有数据应当通过被发送的传输流来传送。尽管这会对应用的设计和范围带来明显的限制，但它仍能提供较为广泛的应用，如游戏、智力竞赛和商用信息。在这种情况下，观察者完全没有必要了解这些信息仍然是通过电视来进行的。特别是在使用预先录制的节目素材中，这种场合下是用音频 / 视频相对直接地对使用数据进行同步，如电视节目中的问答比赛。

11.1.5 交互电视

交互电视与增强电视不同，它们的主要区别是在交互电视中具有发送或接收大量个人数据的反向信道。这将显著地增加交互电视所支持的应用范围，包括能够提供与其它观众的实时交互，如多个参与者的游戏竞赛或投票表决，并能提供真正的个人信息反馈。

当高速因特网的连接与一个真正的反向信道结合在一起时，将是交互电视的进一步扩展，这将提供个人数据与广播电视的无缝集成，同时也使节目或其它素材的 VOD 传送的真正实现成为可能。

11.2 内容封装 (content encapsulation)

数据广播的第一步是将数据以一种适合于传输的形式封装在传输流中。各种不同形式的数据种类繁多，MPEG-2 标准给出了许多不同的数据封装方法。区域性标准如 DVB 和 ATSC，它们在 MPEG-2 的一些基本可选项的基础上又作了进一步的扩展以满足各地的应用需求，但是它们仍然构建在 MPEG-2 标准的核心之上。

11.2.1 MPEG 数据封装

11.2.1.1 数据管道 (data piping)

数据管道用于数据向网络中某一目标机顶盒的简单异步发送。它的主要应用是发送闭合系统中的专用数据。在这个闭合系统中，目标机顶盒已被预先编程，按照确定的 PID 以寻找特定的数据。被传送的数据直接放置在 MPEG-2 传输流包的有效载荷区内，不再需要任何定时信息。其典型应用是向某一区域内的所有销售出口提供产品库存的夜间刷新。

11.2.1.2 数据流

数据流用于数据的端到端传送，其形式可以是异步的或者是同步的。采用的方法与视音频业务相似，也是作为 PES 包传送的。事实上，视频和音频其实就是一种同步数据流业务的特例。由于 PES 包是在业务信息（SI）中发送的，它携带有定时信息，因此数据流比数据管道更灵活，不过它在实用中仍采用专用闭合系统中的类似形式。

11.2.1.3 DSMCC（数字存储媒体指令和控制）

在标准 MPEG-2 专用表中，MPEG-2 DSM-CC 规范（ISO / IEC 13818-6）提供了数据广播的更深层的方式。DSMCC 原先是作为支持 VOD 的一种方式而设计的，它用于网络中节目素材的 VOD 传送。该协议随后得到进一步的扩展，能够解决数据在多个网络提供商之间的按需传送（使用 MPE 模式）和定期传送（使用传送带模式）。

11.2.1.4 MPE（多协议封装）

多协议封装（MPE）允许经由传输流在 DSM-CC 表中传送任何通讯协议的数据报。数据报是一种数据的逻辑结构，它包含有关数据的全部规定信息，如数据量的大小和数据内容，数据的去向和它是如何取得的等等。

数据报的最常见应用是用于因特网业务，在因特网业务中，TCP / IP 数据报携带的信息包括源端和目的端的逻辑（IP）地址以及目标的媒体访问控制（MAC）地址（唯一的网络地址）。MPE 几乎支持任何类型的网络协议，当然它不局限于 TCP / IP 数据。

11.2.1.5 传送带

传送带主要是经由传输流定期传送信息。尽管传送带的内容可以应目标用户的请求而改变，但它更常用于有规律的重复传送，而不管任何对象是否正在收听或当时是否需要这种数据。对那种需要特定数据项目的对象而言，他只需简单地等待直至该数据的重新发送。

有两种不同类型的传送带，即对象传送带和数据传送带，它们之间的主要差别是：

- ▶ 数据传送带只含有未指定内容数据的模块，由对象去确定利用所接收的数据去作什么。
- ▶ 对象传送带包含可识别的数据对象，如图象、文本文件或可执行的应用文件，还含有传送带中所有对象的目录列表。

数据传送带通常用于下载新的系统软件到某一机顶盒，而对象传送带则用于购物服务、EPG 以及发送应用程序和游戏程序等。

在数据和对象传送带中，内容项目是按一定的周期间隔重复发送的。不过，对象传送带可以很容易地改变单个对象的重复速率。例如，下一小时的 EPG 的重复速率要远高于下一月的 EPG。对象的重复速率可以是一种商业决定，它由业务提供商作出，目的是使通道带宽得到充分的利用。

无论是对象传送带还是数据传送带，都是基于 MPEG-2 规范 ISO 13818-6 的 DSM-CC 扩展部分，以及 DVB、ARIB 和 ATSC 系统的特定扩展部分。

11.2.1.6 数据传送带

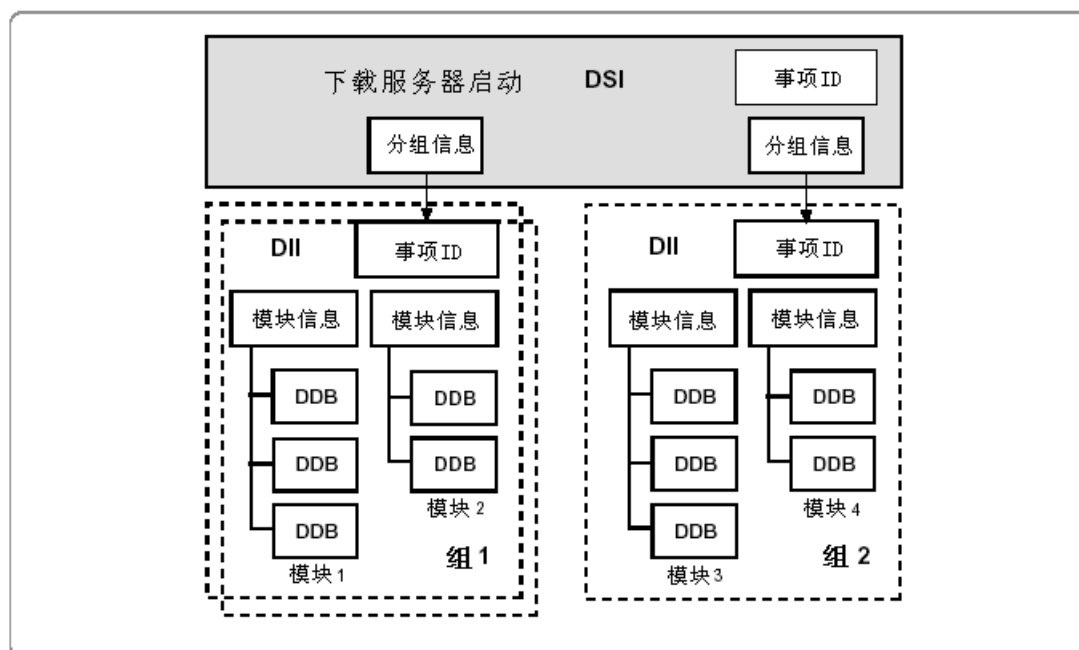


图 11-1

数据传送带不含任何单独的数据项目或目录结构，但它可以有单片数据模块，以使对象用户了解数据的内容和用这些数据来干什么。

数据传送带的结构如图 11-1 所示。将数据的一个完整项目定义为一个“模块”。可将传输模块划分为一个或多个块。每一块均被发送，每一块都作为下载数据块（DDB）信息中有效载荷的一部分，DDB 信息应符合 MPEG 定义的专用表句法。可以按照任何顺序或以任何周期来发送 DDB 信息，因此需要一个机制以识别哪个 DDB 属于哪个模块。

利用下载信息标志（DII）将一个模块中的 DDB 链接在一起。多于一个模块的信息可以在一个 DII 信息中，DII 信息形成一个组。通常一个组是含有逻辑相关的数据模块。

如果相关模块数超过了组合成单个 DII 信息的模块数，那么多个 DII 信息形成一个超组。它们均被下载服务器启动（DownloadServerInitiate，DSI）信息所引用。

一层数据传送带含有在一个 DII 中所引用的少量模块。

二层数据传送带含有参照一个或多个 DII 信息的 DSI 信息。它可能小于或大于单个传送带。

二层传送带的典型应用是用于多语言支持。一组 DII 信息可以将一个模块中的可执行程序与英语文本一起传送到第二个模块。然后第二组只传送含有法语文本的一个模块，但重复应用模块的开销除外。

11.2.1.7 对象传送带

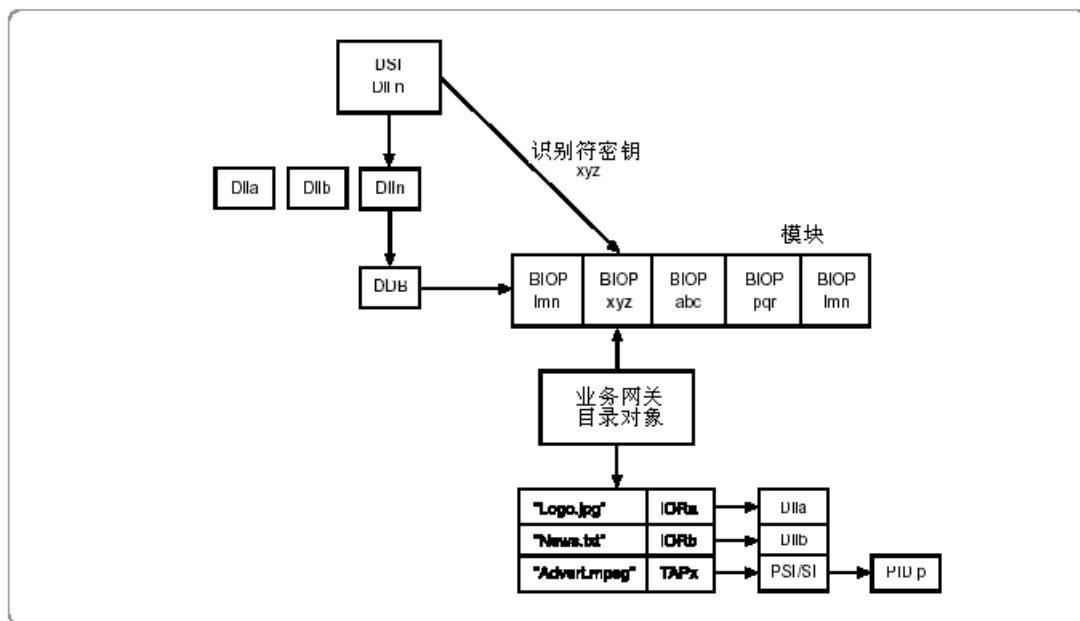


图 11-2

对象传送带是某个服务器至某个接收机的个别广播，广播内容是已标识数据的可识别项目。这些项目即称为对象，它们可以是图象，文本文件，节目，某一视频 PID 的指示符，目录列表或在该传送带中可用的业务网关。相关对象经分组后并作为单个传送带一起被发送以形成一个业务域。可以按照需要随时发送对象，不同的对象可以具有完全不同的重复速率。

对象传送带的一个主要特征是所有对象均采用 BIOP（广播对象请求代理间协议）模式进行传送。传统的软件开发商多年来一直在使用 ORB（对象请求代理）。BIOP 扩展了基本系统，使之能够在广播环境中通过各业务提供商的不同网络支持对象的识别和应用。

就本质而言，BIOP 是用于交换信息的一种方法。这种信息是关于传送带中被广播对象的信息。BIOP 既可以包含对象，也可以只提供对象的指示标记。BIOP 可以指示如何使用对象，包括提供使用该对象时所需应用程序的下载链接地址。

对象传送带和数据传送带也有相似之处，那就是将对象组结合在一起以形成模块。这种基本的数据传送带方法用于发送使用块、模块和 DII 的数据。主要差别在于将 DSI 直接用于指向业务网关的目录对象，通过网关目录对象以寻找传送带中的所有其它对象。整个方案如图 11-2 所示。

11.2.1.8 对象传送带的广播

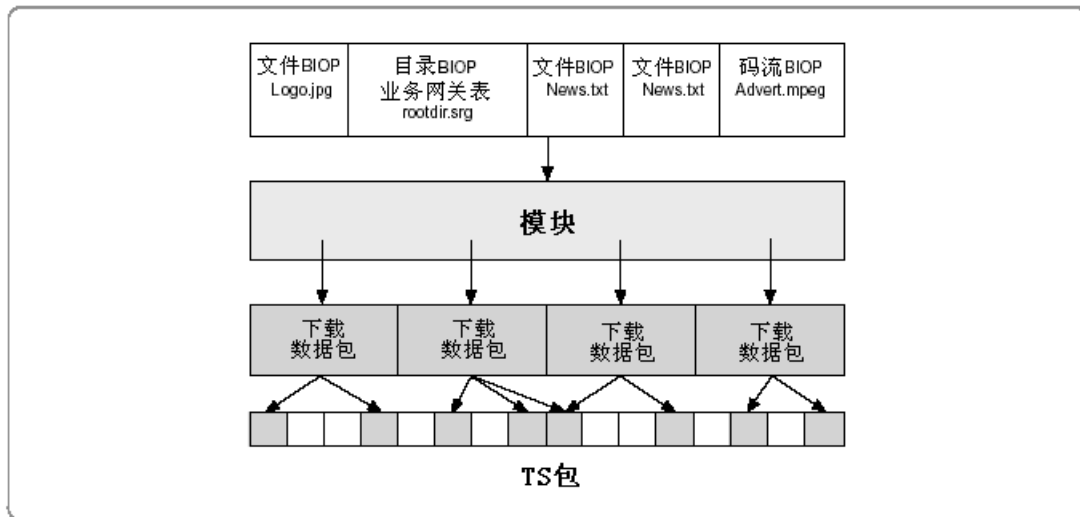


图 11-3

对这一问题的全面解释已经超出了本书的范围。以下仅就这一问题作以简要的介绍（参见图 11-3）。

目录、文件和码流对象以与数据传送带相同的方法进行传送，即将模块拆分为块，将这些块作为下载数据块（DDB）有效载荷区而发送。

下载服务器启动（DSI）信息包含了特定目录对象的位置，这种特定目录对象就称为业务网关。DSI 信息在 SI 中被引用，以形成特定对象传送带的起始点。DDB 包含有业务网关对象的发送模块，下载信息指示（DII）中含有 DDB 的定位信息，而 DSI 与 DII 是有关联的，通过这样的方式，将它们联系在一起。

对象通过 IOR（可互操作的对象基准）在目录对象中被引用。在目录对象中，包含了访问同一业务域中的对象或有关另一对象传送带（包括其它传输流的目标传送带广播）所需的全部信息。

对 IOR 中结构的命名可称为类体（profile body），它用于描述某一对象的位置。类体可划分为两类：

BIOP 类体——仅用于该业务域中的对象。

自选类体（Lite Options Profile Body）——用于其它服务器或传输流的对象

如果该对象可用于多个传送带，IOR 可以包含多个类体，机顶盒可以从中选择以访问最容易的和最快捷的类体。

用于广播 DII、DDB 和视频 / 音频码流的实际 PID 直至发送前是不知道的，分支（tap）正是为了解决这一问题。因此，在传送带中对 PID 的所有查阅只能在分支项中进行，分支与实际 PID 之间的联系是在 SI 中完成的。采用这种方式将大大简化各个网络间码流的再复用过程。

11.2.1.9 MPEG-2 数据同步

数据广播必须以某种方式与正在播出的节目保持同步。如果将数据的实时传送作为它的同步方法是不实用的，除非是在不太严格的实时应用中，例如比赛成绩的刷新，在这种场合中几秒或稍多的误差是没有什么影响的。然而，在有些情况下即使只有一两秒的误差也会有很大的影响，例如在问答比赛中，如果回答过早，就会出现严重的后果。

MPEG-2 为不同类型的数据封装提供了不同的定时机制。除了准实时数据传送之外，数据管道和 MPE 不支持任何形式的定时机制。

数据流中的 PES 包中含有显示时间标记（PTS）和解码时间标记（DTS），这两种时间标记都与由传输流中 PCR 数值中导出的系统时钟同步。这种机制与视音频同步和解码有着相同的准确度。

但在 MPEG-2 数据传送带中却没有定时信息。不过，在对象传送带中可以含有一种被称为“码流事件”的特定对象，它包含了与电视节目中正常重放时间（NPT）相关的定时信息。NPT 与 PCR 不同，因为 NPT 时钟可以暂停，例如某种商业中断。换句话说，某一节目的 NPT 可以与该节目的时间轴完全同步，而不管该节目什么时候发送或者被商业停播而中断。

11.2.2 DVB 数据封装

DVB 采用了标准的 MPEG-2 数据封装方法，仅作了很小的改动，主要是删去了原先规范（EN 301 192）中可能会引起歧义的一些内容，其中包括 DVB 规定了略有不同的 MPE 格式，并在最大数据 PES 包的大小上作了限制。DVB 认为，必须通过提供唯一的网络 ID 指配，以保证在使用 IOR 的数据业务与服务器 ID 之间没有任何歧义。

为使用 DVB 字幕和 DVB 图文电视，DVB 定义了两种特定的数据流 PES 格式，并通过一种标准化的格式，利用数据传送带以发送机顶盒（STB）固件刷新。DVB 还定义了一系列数据描述符和附加表，这些内容将在随后的发信和公告部分里予以讨论。这些附加部分是很有用的，通过使用某一描述符可为 MPE 数据报中的 MAC 地址提供一种简单的掩蔽，以支持窄带的、多路的广播 MPE 业务。

11.2.3 ATSC A/90 数据封装

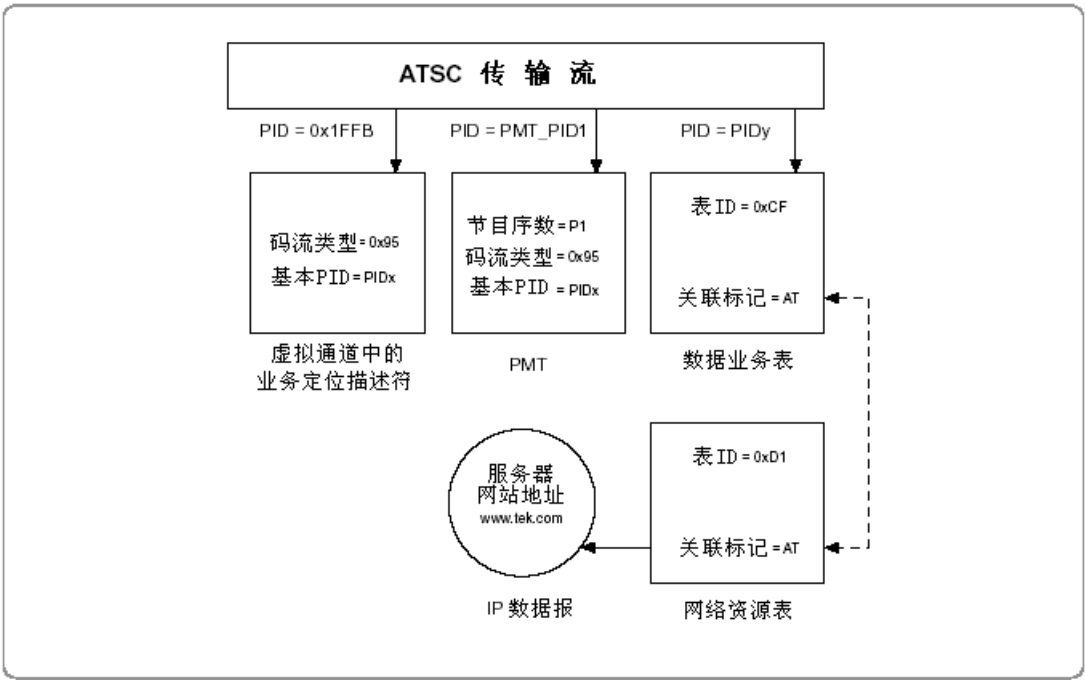


图 11-4

ATSC A/90 规范的完成要比 DVB 系统晚几年，它与 DVB 和 MPEG-2 标准有明显的不同。其一，A/90 完全不使用对象传送带和 CORBA/IOI 系统，而使用了一种单独的表即网络资源表（NRT），以提供链接到一种参考数据的捆绑信息，该参考数据具有自身的实际定

位。除了可以简化系统的运行外，A/90 还能直接支持因特网 URL，而因特网 URL 是不被 DVB 或 MPEG 支持的。A/90 的运行机制参见图 11-4。

其二，A/90 仅使用数据传送带，在数据传送带中，可以有效地只含一单个对象。A/90 又对数据传送带格式进行了扩展，使之可包括一个 DSM-CC 自适应头，自适应头含有显示时间标记，用以同步数据传送。A/90 不支持 MPEG 中有节目 NPT 的概念。

A/90 还定义了附加的数据类型，包括它自己的 DSM-CC MPE 版本，以及允许 IP 数据以 PES 格式传送。与 DVB 相比较，A/90 支持更多的定时模式，以用于各种不同的数据类型。

11.2.4 ARIB 数据封装

日本的 ARIB 标准是在 A/90 标准之后制定的，它是所有数据封装系统中最简单的一种。ARIB 不支持数据管道或对象传送带，它通过数据传送带内容上的特定目录结构，使用数据传送带格式以传送一个或多个实体。ARIB 不允许参考任何实体，实体也不在 PMT 中被直接引用，因此既不需要 CORBA（公共对象请求代理机构），也不需要 NRT 类型的链接。

11.3 数据内容传送

首先要按照有关的国家或国际标准对数据业务进行封装，然后通过传输流发送至 STB。为使 STB 使用某项数据业务，首先，它应当了解该数据业务是否存在以及什么时候可以使用（“公告” announcement）。其次，当数据业务正在广播时，STB 应能发现和识别该项数据业务中的各个部分（“发信” signaling）。MPEG-2 仅定义了 PAT 和 PMT，因此，实际上，只有通过码流类型和某些基本描述符的检验，才有可能实现上述的“公告”和“发信”功能。

尽管只是对 PAT 和 PMT 的使用有许多限制，但对数据业务或数据事件的程序安排也是不可能的。因此在 DVB 和 ATSC 中，都定义了一些附加表，这些附加表大大地增加了提供有关数据业务的信息量。要注意的是，在所有的情况下都必须发送含有 PID 的数据，至少在一个 PMT 中提供一个进入项。

11.3.1 DVB 公告

DVB 认为数据业务既与单个事件（例如一个问答比赛节目）相关，也可以是某项服务（例如体育频道）的一部分。实际上，数据业务与 MPEG-2 规范是一致的，数据业务与常规的音频或视频码流在逻辑上并没有什么区别。

因此，在使用数据业务时，只需在 EIT 或 SDT 表中定义一些附加的描述符，以通知某数据业务或节目是可以使用的，然后使用以上所介绍的方法将数据业务和各部分的实际定位发送出去即可。

唯一的新表是应用信息表，它只是对 MHP 业务应用作出规定。在应用信息表中，含有传送带中的所有应用项和指向其引导类别的指示符，并且可以在改变频道时自动启动或自动取消某项应用。

11.3.2 ATSC 公告

与 DVB 相比较，在 A/90 中，对数据事件和数据业务的处理与视频 / 音频业务是完全分开进行的。因此，它需要添加一些表以支持该项应用。这些表是数据业务表（DST），它列出了正被发送的数据业务项；还有数据事件表（DET），DET 与用于常规节目安排信息的 EIT 十分类似，并且使用相同的句法。第三个表是长期业务表（long term services table），它与 DET 相似，但它提供的是比 DET 为期更长的安排信息，而 DET 最多只能提供 16 天的节

目安排。

11.4 内容提交

数据广播的最后一步是在 STB 上运行该项应用程序并向观察者提交信息。应当注意的是,如果视频、音频或数据是按照正确的协议被封装的话,某一传输流对其所携带的实际内容是完全不可知的。换句话说,后续部分几乎完全脱离了 MPEG-2、DVB 或 ATSC 数据广播标准,应用程序和业务数据只是被传输流所传送并被 STB 所使用。

11.4.1 机顶盒中间件

在某些情况下,例如因特网数据,数据的使用环境是自身定义的,相对而言不太重要。然而在大多数情况下,例如游戏和应用程序则是重要的,数据应用的结构应使其在目标机顶盒中能正常运行。在大多数情况下,STB 具有软件应用编程接口(API),它将应用程序的高级功能调用连接到软件驱动程序上,软件驱动程序控制机顶盒中的实际硬件并与传输流中的数据相链接。这种介于应用程序和驱动程序之间的软件层,即所谓“中间件”。

目前世界上使用着各种各样的 STB。其性能和特性各不相同,一些机顶盒只具有非常基本的功能,只能解码和显示一种基本的标准清晰度图象,而最近生产的机顶盒具有强大的处理功能。一些最近生产的机顶盒包含有高性能的图形芯片,能够在所有的操作系统(如 Linux 和 Windows)下运行,还配备有记录节目的硬盘驱动器,具有 DVD 重放功能的 DVD ROM,并可运行常规程序如微软或 Star Office 程序包。

一些原先的机顶盒也许不能显示为新型高性能机顶盒而设计的简单图象。为此,一些应用程序开发商不得不使用不同的编程语言为各种机顶盒、中间件和操作系统编写完全不同的程序,很明显,效率低下且增加了开发成本。

于是人们试图对机顶盒进行标准化,至少是在一些基本功能上进行标准化,应当明确定义它的功能,并能被广播业者以一种非常容易接受的方式对其进行控制。

解决这一问题的另一途径是给出一种全面的解决方案,通过提供一种中间件和应用程序开发工具以开发顺应性程序,这样的系统是更为开放的一种系统,只要机顶盒满足一些基本的所需功能,该中间件经许可后就能够以不同的性能和特性运行在各种不同的机顶盒上。

目前,作为一种交互系统,一种具有专利的数据传输方案和“有墙公园(walled garden)”交互业务已经获得了广泛的应用,它们经由反向信道而提供真正的交互性。一些开发商提出的方案是使用标准的对象传送带以传送应用程序,这在欧洲被普遍采用。还有许多其它的专利系统。

一种更为开放的标准是由多媒体超媒体专家组(MHEG)所定义的应用环境。由 MHEG 衍生的一种特定方案即 MHEG-5,它为低性能低内存应用特别是为机顶盒作了优化处理。MHEG-5 已经取得了一些成功,特别是在英国,将它用于接收地面电视广播的机顶盒。

当前的动向是通过一种明确的已定义的编程 API,使开放的中间件标准化。这样有利于消费者从各种机顶盒中作出选择,促使应用开发商重用编码。最近出现了一些新的标准,其中包括来自 DVB 的多媒体家庭平台(MHP)。在美国,高级电视增强论坛(ATVEF)和数字 TV 应用软件环境(DASE)也都处于开放标准的前沿。

11.4.2 DVB 多媒体家庭平台(MHP)

DVB 多媒体家庭平台(MHP)定义了一种非常广泛的应用程序接口,它具有几种不同的类别和级别,以适合于不同性能和不同价格的机顶盒和业务应用。它支持增强电视和交互电视(使用反向信道),同时支持因特网连接。在 DVB 的其它标准中,并没有对因特网提供直接的支持。

MHP 应用的传送是通过标准的 DVB 对象传送带和 AIT 表（本文前面已作过介绍）来实现的。它可以自动启动、暂停或停止应用程序的运行，并可采用高速缓存以加快应用程序的启动速度。MHP 标准是基于 java 技术，它包含了一套由现有标准如 java TV、HAVI（家庭音频视频互操作性）和 DAVIC（数据音频视频理事会）导出的应用程序接口。

MHP 的应用既可基于 java，也可基于 HTML，它含有许多来自各供应商的 MHP 软件开发工具包（SDK）。MHP 应用必须经过一致性测试和电子签名，以确保系统的互操作性和安全性。

MHP 标志是受到保护的，带有 MHP 标志的机顶盒在销售前必须经过一致性处理。正规的互操作性测试应当在最近两年内进行，以保证所有的 MHP 应用都能在所有的由不同厂商提供的 MHP 机顶盒上正常运行。

事实上 MHP 是欧洲的标准，该项业务于 2001 年 8 月在芬兰正式启动。目前有大量的测试业务正在全欧洲境内广播。德国将从 2002 年中期开始使用 MHP，其它国家也会在最近几年内采用。

然而，MHP 并不只是欧洲的标准，它已被世界上许多地区所采纳。例如在在在韩国已被指定用于卫星数据广播，在澳大利亚用于地面广播。2002 年二月，美国的开放有线应用平台（OCAP）组织已采用 MHP 作为美国的数据有线广播标准。还有许多其它国家和组织预计在最近几年内会采用 MHP 标准。

不应低估 MHP 对全球产生的影响。在 2002 年期间，当前交互电视业务市场的主导厂商已经提供了 MHP 插入以用于他们的带有自己专利的中间件，从而有效地将他们的专利机顶盒推向全球市场。随着 MHP 在欧洲各国的普遍采用，将近十亿收看者的潜在市场已显现出来，这预示着这个开放标准的美好未来。

11.4.3 ATVEF DASE

有时人们说，ATVEF（高级电视增强论坛）和 DASE（数字电视应用软件环境）的区别在于：前者建议使用 HTML 但允许使用 java，而后者建议使用 java 但允许使用 HTML。尽管如此，它们之间还是存在某些其它基本区别。

ATVEF 规范是由一个跨越行业的制造厂商组制订的，其目的是通过定义标准化的功能，使交互内容能够运行在机顶盒、基于 PC 的接收机和交互电视机上。这样就产生了增强的内容规范（ECS）。

ATVEF ECS 是在 HTML 的基础上创建交互电视内容，不过也允许使用 JavaScript。它还需要其它许多基本功能，例如使用便携式网络图形（.png）格式，并将它作为标准的图象互换格式。因此，ATVEF 与因特网中的网页设计有着基本的通用性，这对应用程序开发商是十分有利的。

ATVEF 的一个显著特征是它在标准中并未对传输机制作严格的定义。它能够跨越模拟和数字传输模式而运行得很好，不管有无视频内容。它之所以能够这样，是因为它定义的是一种独立于传输的内容格式和使用 IP 并将它们作为基准结合。这样，ATVEF 就可以运行在任何使用 IP 的传输模式上。而且，还能使 ATVEF 的运行跨越于 NTSC 模式以及 MPEG 传输流，并能提供一种简单的返回路径机制。ATVEF 在美国被用于模拟和数字电视系统。

11.4.4 DASE

与 ATVEF 相比较，DASE 是由美国 ATSC 所开发的一种相当复杂的中间件规范。由于 ATSC 也开发了 A/90 数据广播标准，DASE 规范提供了与 A/90 传输模式的联系，而且还对 API 作出定义以使 DASE 应用可以继续。DASE-1 标准于 2002 年早期发布，它是一项可选择的标准。

DASE 将发信方案和公告方案以及被称为数据要素的实际应用数据内容结合在一起。DASE 允许使用两种类型的数据要素（包括它们的应用环境和功能）：

- ▶ 基于.xml和.xml 标记文本和脚本的说明性数据要素。
- ▶ 基 javatv xlets 的程序数据要素。

DASE 接收机提供处理两种数据要素和公用核心功能如.jpeg 或.png 编解码器的引擎，它也同时提供与正被发送的音频或视频内容的链接，还定义了许多图形屏幕格式，从而将交互内容（包括 HDTV）显示在屏幕上。

DASE 应用与 MHP 应用收视周期模式相似，它可以被初始化，被激活或暂停，也可以不经初始化，在必要时可使用高速缓存。

尽管 DASE 有着 ATSC 的背景，但由于它进入这个领域相对较晚，所以至今尚未得到广泛的采用。

第十二章 MPEG 测试

MPEG 测试的基本要求是，它能对现有的传输流进行分析，以检测该传输流是否符合标准。为使这种测试完备，应当创建已知符合标准的传输流。

12.1 测试要求

尽管 MPEG 技术与以前的技术有着根本的区别，但对测试的基本要求是相同的。在可操作的基础上，用户需要一种简单的、正规的置信度检测，以确信设备能够完全正常地工作。在出现故障的情况下，需要迅速地确定故障的位置，并需要对故障的性质作深入细致的研究。和所有的信号测试一样，MPEG 测试的方法是：产生已知有效的信号，并将它插入系统，而后在系统各点对已送入的信号进行测量。

MPEG 系统与传统广播电视设备的一个最大区别是它存在着多个信息层，每个信息层对其以下的各层都是透明的。十分重要的是要确定故障发生在哪一层，以避免作无效的故障搜索。

例如，如果在 MPEG 解码器的图象监视器中出现可见的图象缺陷，这些缺陷的出现具有各种可能性。也许是编码器的故障，而传输流本身并无问题，它只能原封不动地提交这种故障信息。另一方面，也许编码器是好的，是传输层破坏了数据。在 DVB 中，数据处理的层次更多，例如能量扩散，误码校正和交织处理等。如此复杂的数据层需要一种结构性的测试方法并使用正确的工具以查找故障。在本书中，已对压缩数据的协议分析作过介绍，它有助于用户从中找到这样的结构性测试方法。另外，阅读有关压缩电视测试的介绍，了解图象质量评估，也是很有帮助的。有关压缩电视测试的讨论可参见泰克公司的应用手册“A Guide to Video Measurements for Compressed Television Systems”（压缩电视系统视频测量指南）。

12.2 传输流分析

MPEG 传输流具有非常复杂的结构，但使用诸如 AD953 之类的分析仪，则可按照一种逻辑方式对 MPEG 结构进行分解，从而可使用户观察到任何所需要的细节。这时可利用许多通用的分析方法对实际传输流进行实时分析，其中包括传输流中节目的分层显示以及每个码流的比特率分配显示。

如果需要进行更详尽的分析，则只有将传输流中某部分记录下来以便随后作深入的分

析。这种分析方法即所谓延时测试，例如，它可用来检查时间标记的内容。

在使用延时测试时，MPEG 传输流分析仪就象一台逻辑分析仪一样，为 MPEG 提供了特有的数据解释工具。如同所有的逻辑分析仪，需要一种实时的触发机制以确定数据俘获发生的时间或条件。在图 12-1 所示的分析仪中，它包含了三个部分，即实时部分、存储部分和延时部分。在实时分析中，仅实时部分工作，还需将该部分与信号源相连接。在俘获时，实时部分用来测定何时触发俘获。分析仪中还包括有被称为滤波器的工具，以对俘获前后作有选择的分析。

一旦俘获完成，延时部分即对所俘获的数据进行分析，而不再需要输入信号。这与存储示波器十分相似，在存储示波器中，可以直接显示实时输入信号，也可保存信号波形以待以后分析。

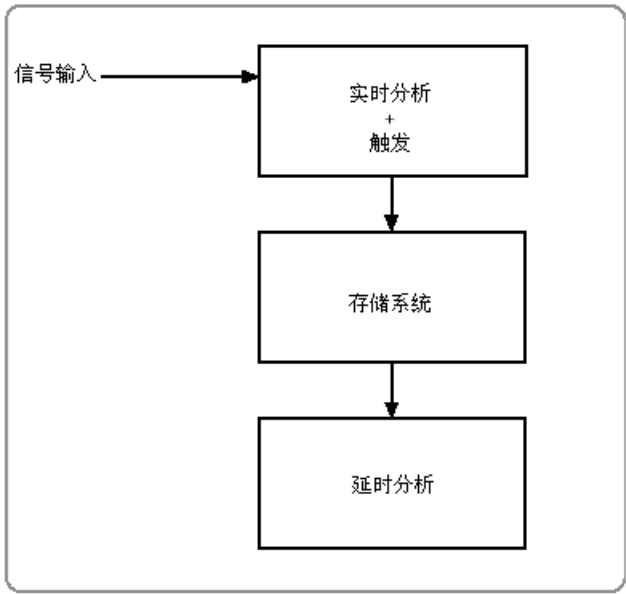


图 12-1

12.3 分层图示

在分析不熟悉的传输流时，从分层图示开始是一种非常好的方法，它能以图表的方式给出传输流中各个组成部分的显示。图 12-2 就是分层显示的一个例子，它是用泰克公司的 MTX100 采集的分层显示。在图 12-2 中，从图形的左上角开始，整个传输流被分解为很多分量，每个图标就代表每个码流分量。表 12-1 给出了分层图示中所使用的各个图标以及它们所代表的含义。用户可以非常方便地观察到有多少个节目流和每个节目流中的视频和音频内容。每个图标都代表编码句法结构的一个顶层，在其下面有许多更低的分析和信息层。

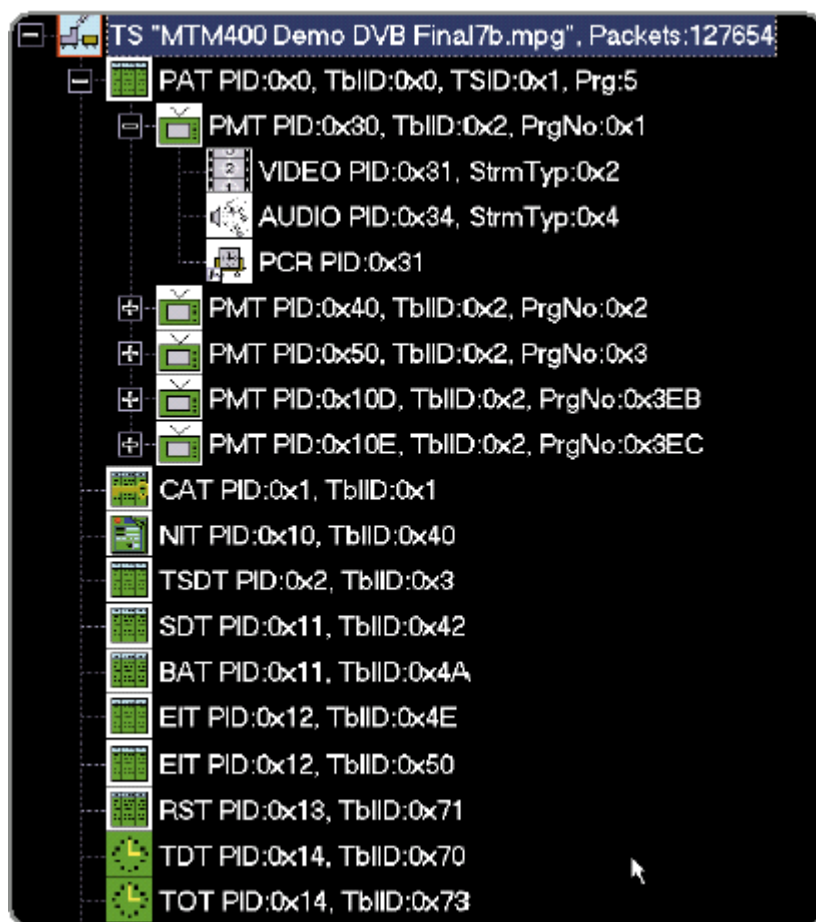



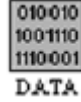


图 12—2

MPEG 分析仪是通过传输流的 PSI 数据中所包含的 PAT 和 PMT 来创建分层图示的。在每个图标后面都显示有这些表的 PID。对任一解复用器或解码器而言，PAT 和 PMT 是它们正常工作的基本数据。如果分析仪不能给出分层图示，或在分层图示中有明显的错误，则被测传输流中就存在着 PAT / PMT 错误。要使设备对这样的码流作出解释也几乎是不可能的。

表 12—1 分层显示图标

图标	含 义
	多路复用的传输包。该图标代表组成该码流的所有传输包（每传输包中含 188 个字节或 204 个字节），如果您把传输流想象为一列火车，那么这个图标就代表该火车中的每个车厢，至于火车的具体配置（例如敞车、闷罐车或底卸式车）和车厢中是什么货物则不必考虑。
	某一特定 PID（节目 PID）的传输包。其它部分（表，时钟，PES 包）是传输包中的“有效载荷”，也可以是具有相同 PID 的若干个传输包中的有效载荷。PID 数值出现在图标之后。在分层图示中，该图标右边的图标代表属于这个 PID 值的传输包的有效载荷。
	含有独立 PCR 时钟的传输包。PID 值在该图标之后。

 PAT	PAT（节目相关表）部分。它总含有 PID 为 0 的传输包。
 PMT	PMT（节目映像表）部分。
 NIT	NIT（网络信息表）。它能从选择菜单中通过 PSI / SI 指令访问 SI 表。它也可用于专用信息部分。在选择 DVB 项（在自选菜单中）时，这个图标也可以代表 SDT、BAT、EIT 和 TDT 部分。
 PES	PES（打包的基本码流）。在该图标代表的所有的传输包中，均包含有一个给定的基本码流。若干个传输包的有效载荷可组成单个 PES 包。
 VIDEO	视频基本码流。
 AUDIO	音频基本码流。
 DATA	数据基本码流。
 ECM	ECM（授权控制信息）部分。
 EMM	EMM（授权管理信息）部分。

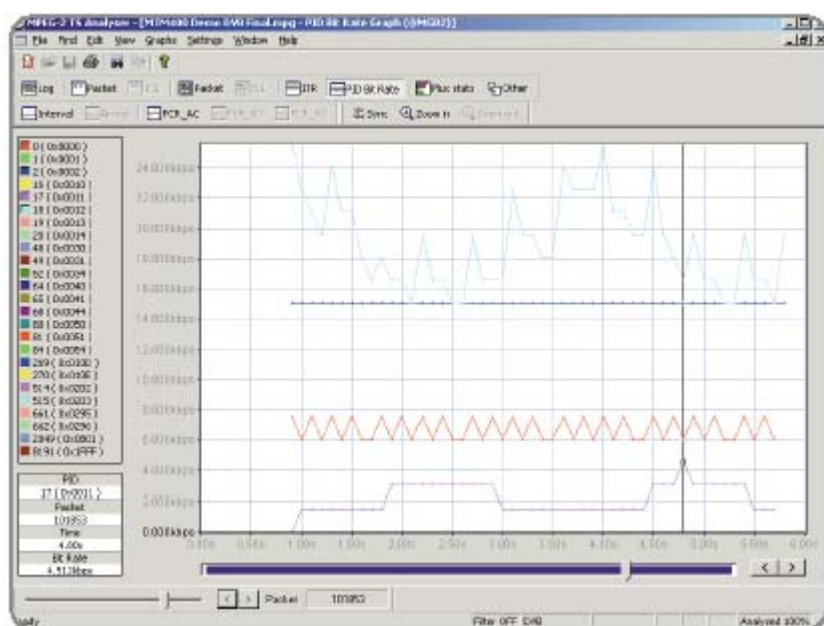


图 12—3

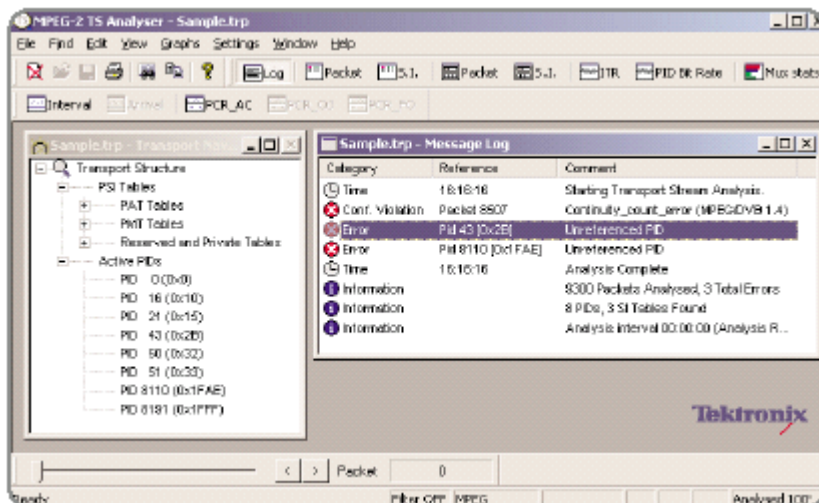


图 12—4

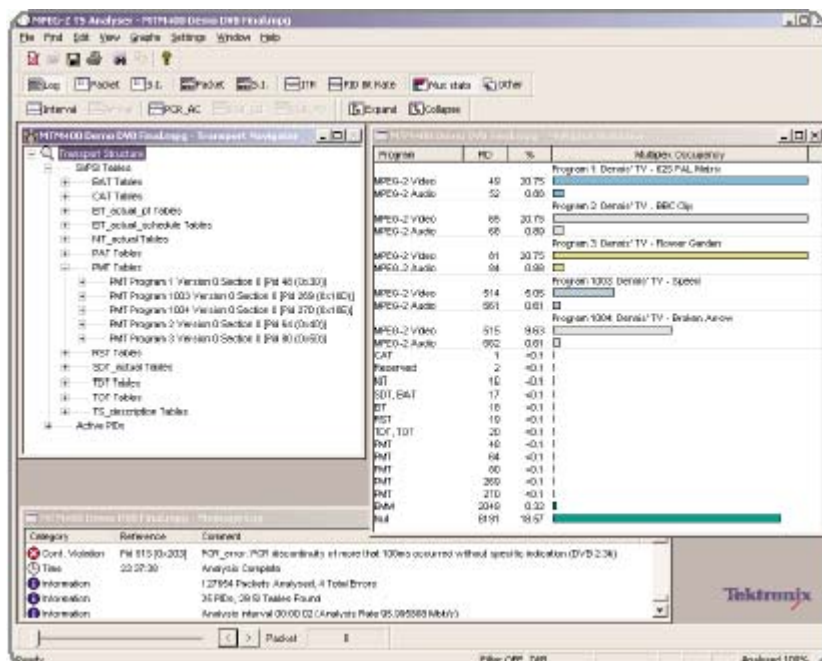


图 12—5

解复用器或解码器能否锁定于传输流决定于 PSI 数据的发送频率。图 12—3 是 PSI / SI 的速率显示，它给出了系统信息的插入频率。PSI / SI 信息应当与传输流中的实际内容一致，例如，假定某个给定的 PID 出现在某个 PMT 中，那么，在该传输流中就应当可以找出这个 PID 的数值。一致性查验功能就是进行这样的比较检查。在图 12—4 中，显示出某一传输流中的一致性错误，该传输流中出现了两个未能识别的数据包。

在 MUX（复用器）分配表中，可以用图表的形式表示分配给每个 PID 或每个节目的传输流的比例。图 12—5 就是一个 MUX 分配表显示的例子。在分层图示和 MUX 分配表中，均显示出传输流中的元素数和各元素被分配的带宽比例。

12.4 解释性图示

为了在未指定的位置查验指定的数据，一种可能的替代方法是在指定的位置分析未指定的数据，这些未指定的数据包含在单个传输流包中，或在表中，或在 PES 包中。这种分析方法即所谓解释性图示，因为分析能够自动地对数据进行分析和解码，而后还会显示其含义。图 12—6 就是这样一个例子，它给出了某一传输流包的十六进制显示和解释性图示。由于被选择的项目是能改变的，因此可以显示出相对该传输流起点的包序号。

图 12—7 是 PAT 解释性图示的另一个例子。

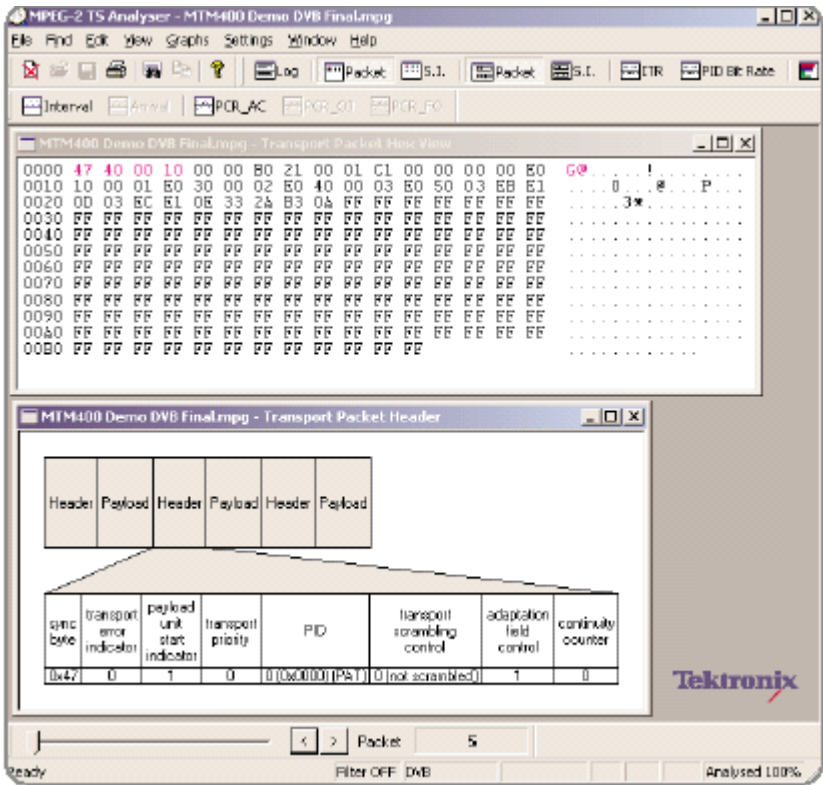


图 12—6

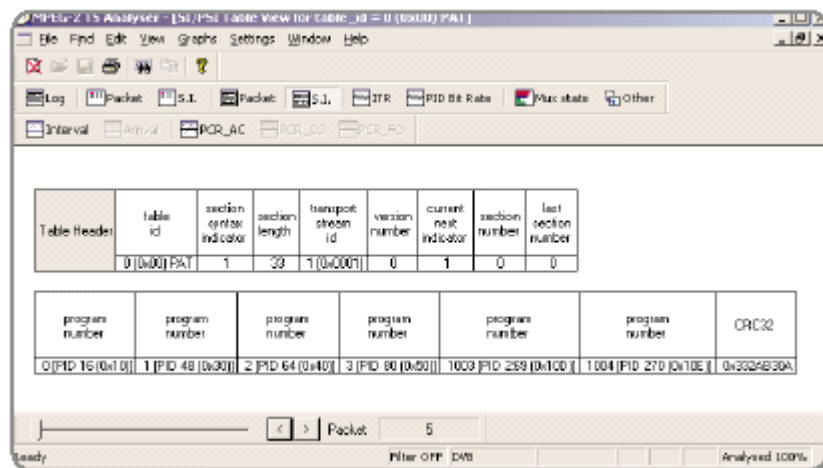


图 12—7

12.5 句法和 CRC 分析

为使传输流传送节目素材，全凭编码器对句法的正确运用。如果不能正确地设置固定标

志比特、同步模式、包起始码和包计数，那么解码器就会对该比特流作出错误的解释。句法查验功能的检查对象是所有非节目素材的数据比特，并显示出所有的句法错误。那些寄生的差错可能是因传输错误而引起；而一致性差错则来自于编码器或复用器故障。图 12—8 给出的是句法错误以及循环冗余校验（CRC）丢失的图示。

许多 MPEG 表都附加有校验和或 CRC 以用于误码检测。MPEG 分析仪可以重新计算校验和并将它与实际的校验和相比较。还有，寄生的 CRC 失配可能是由码流的比特错误所引起，但一致性的 CRC 错误则与硬件故障有关。

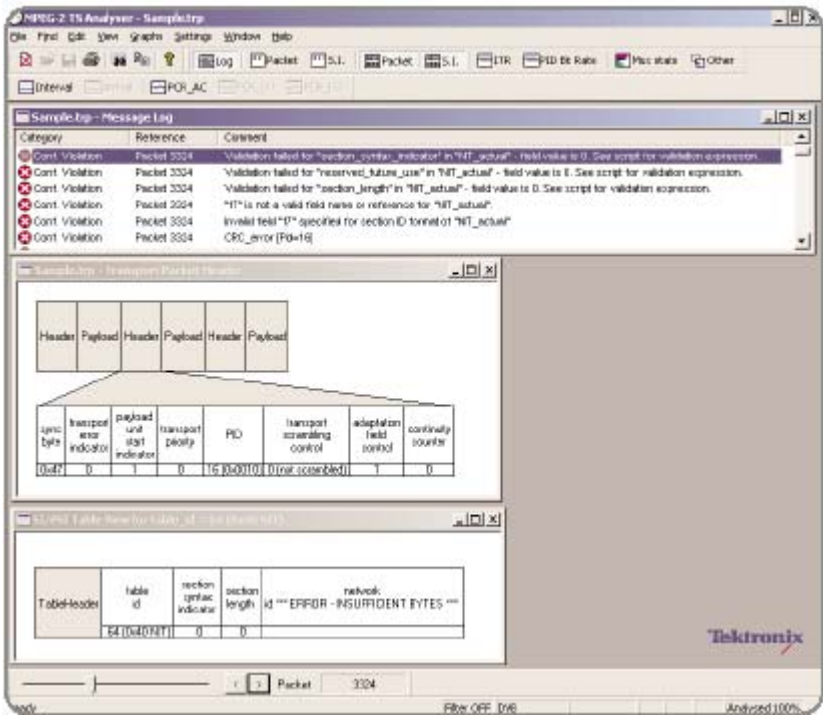


图 12—8

12.6 过滤

在传输流中包含有大量数据，在实际的故障状态中，除了存在严重的问题之外，一般大多数数据是有效的，也许只有一个基本码流或一个节目受到影响。在这种情况下，更有效的方法是有选择性地进行测试，这就是过滤的功能。

就实质而言，过滤就是在对传输流进行检测的过程中，让用户在使用分析仪时具有更多的选择性，用户可以只对满足某种条件的部分数据进行分析，而不必接收每个比特。

其中一个条件是对传输流包头进行过滤，这样可以只对具有给定 PID 的传输流包进行分析。利用这种方法，只需选择 PID 0，就能非常方便地对 PAT 进行查验，并由此出发，可以读出所有其它的 PID。如果某一被怀疑传输流的 PID 是已知的，也许从分层图示中，就能轻而易举地找出那个 PID 来进行分析。

12.7 定时分析

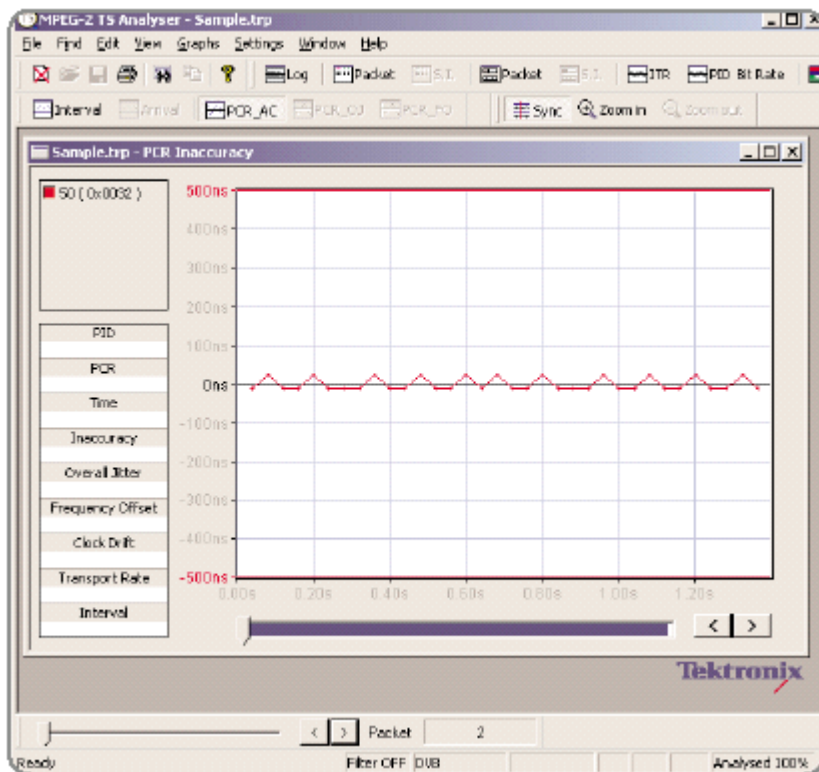


图 12—9

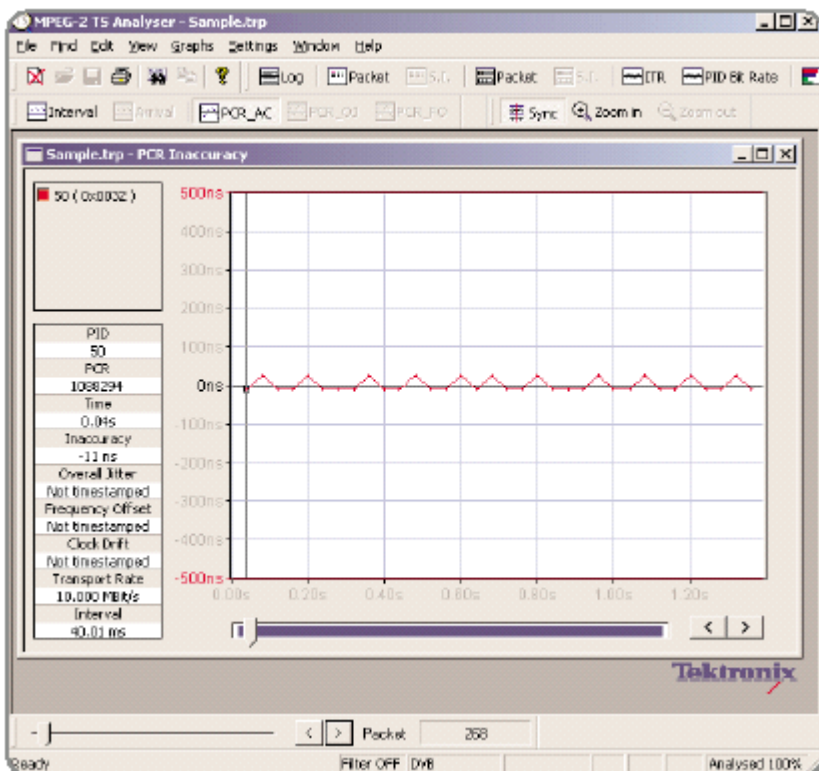


图 12—10

以上所介绍的测试方法是用来查验传输流中的元素和句法是否正确。然而，为了正确地显示实时的音频和视频，传输流还必须为解码器提供准确的定时。这项任务的确认是通过对

PCR 和时间标记数据的分析而实现的。

节目时钟数据的正确传送是至关重要的，因为节目时钟数据控制着整个解码过程的定时。通过分析每个节目中的 PCR 就能表明，PCR 数据的发送是否具有足够的速率，它是否具有足够高的精度。

来自复用器的 PCR 数据一般是准确的，但在再复用过程中，可能要将某一给定节目的数据包放置在时间轴上的不同位置，这就要求再复用器对 PCR 数据进行编辑。因此，重要的是要测试再复用后的 PCR 数据的精度是否合乎要求。

图 12—9 所示是 PCR 的测试显示，它给出了被接收的 PCR 相对一个平均时钟的位置。图 12—10 是对 PCR 进一步分析的图示，在这里，每个 PCR 均可以被打开以显示 PCR 数据。为了测量 PCR 的不准确度，分析仪通过使用先前的 PCR 和比特率对 PCR 作出预测，产生被称为内插的 PCR。从估算的 PCR 中减去实测的 PCR 值，即可得到被测 PCR 不准确度的估算值。

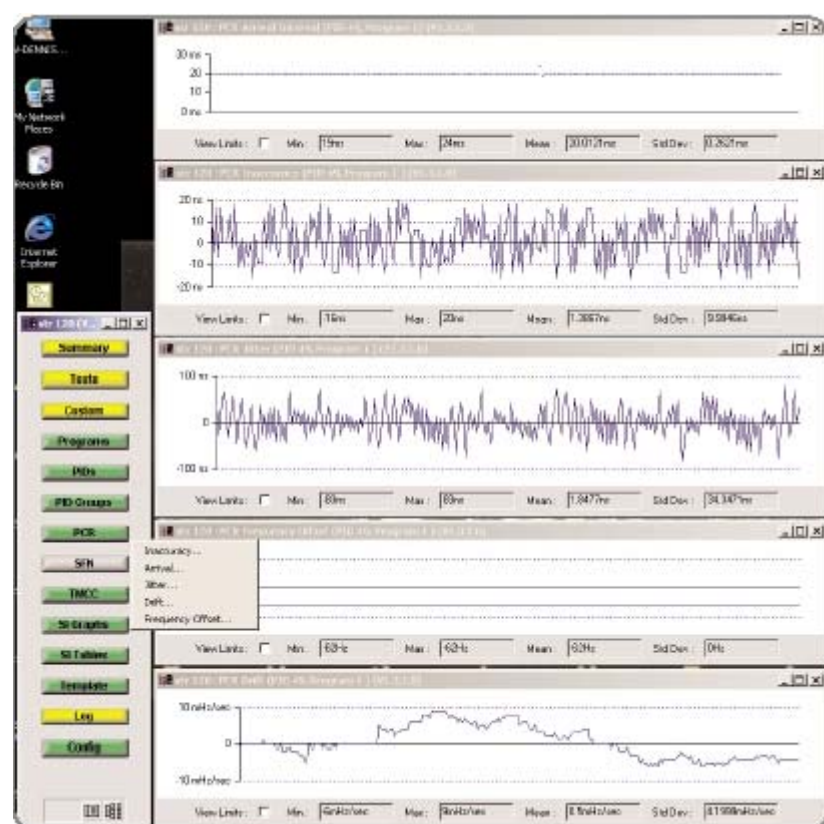


图 12—11

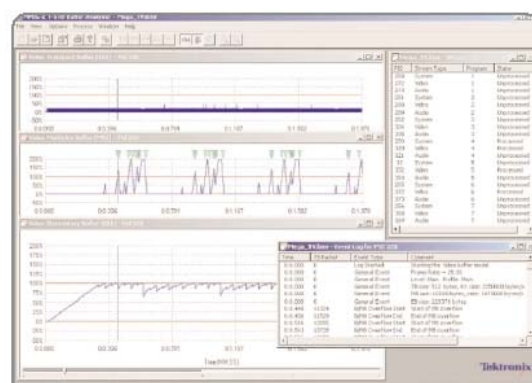
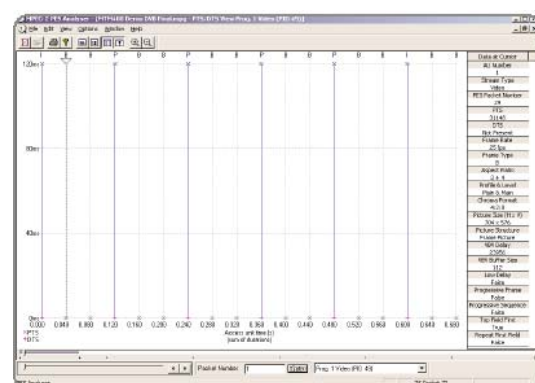


图 12-12

图 12-13

测试 PCR 的另一种方法如图 12-11 所示。在图中，给出了 PCR 周期、不准确度、抖动、频率偏调和频率漂移的图形显示，这种显示还可以实时刷新。

图 12-12 是某一被选基本码流的时间标记显示。它还可以给出访问单元、显示时间（presentation time）和解码时间（如果位置合适的话）等显示。

在 MPEG 处理的过程中，需要对图象重排序，也要使用不同的图象类型，这均会导致时间的延迟，因此在编码器和解码器中均需使用缓存器。对一给定的基本码流，它的编码应当受到解码器中可用缓存的限制。为此，MPEG 定义了一种解码器模型，它称为 T-STD（传输流系统目标解码器）。任一编码器或复用器不应当超出 T-STD 的缓冲能力而使数据流失真。在传输流中，含有被称为 VBV（视频缓存检验）的参数，它对某一给定基本码流所需缓存的大小作出了规定。

在分析仪的 T-STD 分析显示中，能够用图形的方式给出缓存的占有率，这样可以很方便地观察到是否有上溢或下溢发生。图 12-13 即为缓存显示图形。

正常的压缩器 / 复用器输出信号的使用范围是有限的，因为这种输出信号具有不确定性。如果发现解码器的输出有问题，就不能保证在测试时一定会出现同样的故障，因为相同的视频信号不一定会产生相同的传输流。在这种情况下，就需要一种可以绝对重复使用的传输流，这样，在需要进行研究或调整时，通过它就可以使这种故障能够再次出现。

传输流抖动应当限制在一定范围之内，但设计良好的解码器在这个范围之外也应能使节目恢复，这样才能保证运行的可靠性。如果使用现有的传输流，是无法对解码器的抗抖动性能进行测试的。这是因为，假如传输流符合标准，因解码器能正常工作也就无需测试；假如传输流有问题而使解码器不能正常工作，那么既不能重现这种传输流而且结论也不明确，因为这既可能是因抖动而引起，也可能是因为传输流在其它方面不符合标准。较好的解决方案是产生一种在各个方面均符合标准的传输流，再附加上可以控制的 PCR 不准确度，那么我们就知道，在这样的传输流中，唯一不符合标准的就是这种不准确度。AD953 的编辑功能就能创建这样的测试信号。

12.8 基本码流测试

MPEG 比特流有着十分灵活的特性，在 MPEG 比特流中，究竟包含有多少种可能性和多少种组合方式，这几乎是无法计算的。由于未对编码器作出定义，也就不能要求编码器厂商使用每一种可能的方式，的确，就经济而言，这也是不现实的。然而，这一事实使测试变得相当困难，因为某一解码器与某特定编码器之间的工作状况并不能证明它们是否合乎标准。当然，在实际使用解码器时，只需不选择让它无法工作的方式就行了。

一致性测试还有另一方面的问题。由于编码器具有不确定性，即使是在视音频信号重复输入的情况下，也不一定会产生相同的比特流，在 I、P 和 B 图象以及视频帧之间，相同的图象排序一般也不会重复出现。因此，如果解码器在某次给定的测试中出现了问题，也许下一次测试时却不会出现。这就为故障的寻找造成了困难，也无法确定故障是在编码器还是在解码器。编码的难度主要取决于节目素材的性质，但也不需要给对给定节目素材的每个参数在整个编码范围内都要进行试验。

为使测试提供有意义的结果，需要以下两种工具：

- ▶ 一种已知符合标准的测试信号源（顺应性信号源），它能满足整个编码范围内的测试需求。这种信号必须具有可确定性，能够再现解码器中所出现的故障。萨尔诺夫顺应性比特流就是为完成这项任务而设计的。

- ▶ 一种基本码流分析仪，它能对编码器的全部句法进行查验，看其是否合乎标准。



图 12—14

12.9 萨尔诺夫顺应性比特流

符合上述要求的比特流已由萨尔诺夫公司特别设计，它用于解码器的顺应性测试。可以将这种比特流经复用后进入传输流中以馈送至解码器。

在使用萨尔诺夫顺应性比特流时，无需干预解码器的内部工作。为避免对解码器输出进行长时间的分析，可以让它创建一种简单的测试图象，在测试结束时，只需将图象监视器与解码器输出相连接，就可以在屏幕上观察到测试结果。

这样的简单测试图象有许多种。图 12—14 显示的是灰色检验图象，用户可以察看屏幕以查找有无异常出现，在灰色场面中，这种异常对比是十分清楚的。当然，也有一些不是灰色的检验图象。

在进行测试时，如果存在某种故障，或许完全就没有图象出现。在测试结束时会在屏幕上显示“VERIFY”的字样。

进一步的测试则需要观察者检验某运动目标穿越图象作平滑运动的状况。如果存在着定时或图象排序问题，将会观察到有抖动出现。

萨尔诺夫全套测试信号可用来依次检查所有的 MPEG 句法元素。在其中的一种测试中，被测比特流在开始时仅使用 I 图象，接下来加入 P 图象，然后再加进 B 图象，这样可以检验所有的 MPEG 图象类型能否被处理以及图象的重排序是否正确。也可以利用它来查证被测设备与 MPEG—1 的逆向兼容性。在另一种比特流测试中，利用了各种不同的 GOP 结构，这样可以检查运动矢量在整个取值范围内的工作状况，还可用来测试在改变像素条大小或改变缓存大小时的工作状况。

萨尔诺夫比特流除了可提供解码器测试外，它还包含有一种测试序列，该序列能使良好的解码器输出标准的视频测试信号，以检验 DAC（数字 / 模拟转换器）、信号电平以及复合或 Y / C 解码器。这种测试序列能将解码器当作视频测试信号发生器使用，它可以产生如区域板（zone plate）信号、斜坡信号和彩条信号。

12.10 基本码流分析

基本码流部分是传输流应当透明地送交的有效载荷。传输流在传送基本码流时，不会理会传送的基本码流是否合乎标准。换句话说，对传输流顺应性的测试只是意味着检验该传输流能否不失真地传送基本码流，但不意味着基本码流在它构建之初就有着正确的组成。

基本码流的结构或句法是由压缩器决定的。因此对基本码流的测试实质上是压缩器测试的一种形式。应当注意到，某一压缩器能够产生顺应性的句法，但是仍然可能会出现较差的视音频质量。因为压缩器是基于算法规则而不具有确定性，如果压缩器不适当地工作于通用性较差的运行模式，那么基本码流就可能表现为时好时差的不顺应性。

在传输流中通常包含来自不同编码器的几个节目，基本码流问题往往是限于某一个节目，而传输流问题则往往影响到所有节目。如果是在某一特定解码器的输出中发现问题，就应当对这个解码器进行萨尔诺夫顺应性测试。如果测试表明是正常的，那么故障就可能在输入信号上。如果传输流句法已被测试，或者其它节目工作正常，那么基本码流分析也得到证实。

基本码流分析可以从句法的顶层开始并继续向下进行。序列头是非常重要的，因为正是序列头将压缩中所使用的所有相关模式和参数信息传递给解码器。应当按照本书 5.1 和 5.2 节介绍的基本码流句法来检查。图 12—15 所示为一序列头在 AD953 中的扩展显示。更下层的测试如图 12—16 所示，图中为一已解码的 B 帧图象，该图象上还叠加有运动矢量。

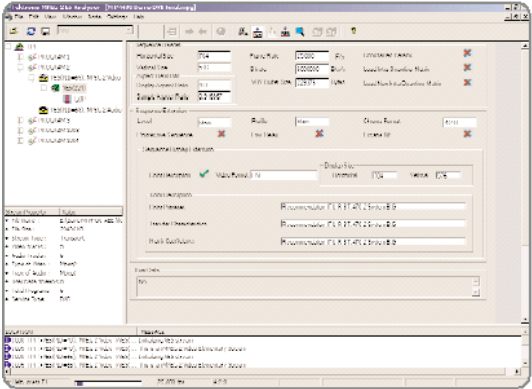


图 12—15

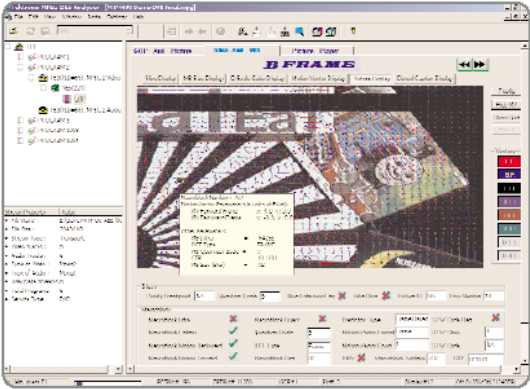


图 12—16

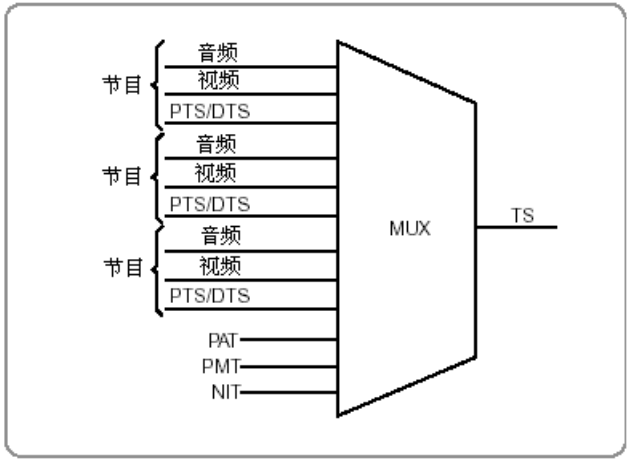


图 12—17

12.11 传输流的创建

如果怀疑解码器有问题，如果能够创建一种质量已知的测试信号将是用得着的。在图 12—17 中，表示一个 MPEG 传输流应当包含的节目专用信息（PSI），例如用于描述一个或

多个节目码流的 PAT, PMT 和 NIT。每个节目码流都应当有自己的 PCR, 而基本码流应有周期性的时间标记。

在 DVB 传输流中还附加有业务信息, 例如 BAT, SDT 和 EIT 表。PSI / SI 编辑器能把 PSI / SI 的任意顺应性组合插入到用户测试码流中。

很明显, 以上每个项目都各自需要可用的传输流速率。复用器提供了一个速率标尺以显示已使用的总比特率。剩余比特率则全部用于插入填充包, 填充包的 PID 均为 1 (在二进制中), 解码器将不予以接收。

12.12 PCR 不准确度信号的发生

在 MPEG 解码器中, 它必须使用 PCR 数据中的时钟样值来再生一个连续的时钟, 并以此时钟来驱动解码器中的锁相环。环路应经过滤波和阻尼处理, 以使 PCR 数据到达时间中的抖动不会造成时钟的不稳定。

为了测试锁相环的性能, 需要一种不准确度为已知的测试信号, 否则, 这种测试将毫无意义。为此, AD953 可以模拟这种不准确度。因为 AD953 是一种基准发生器, 它具有高度稳定的时钟电路, 其实际抖动输出是很小的。为了产生抖动的效果, PCR 数据的定时是不能改变的, 但可以修改 PCR 数值, 以使该数值中所包含的 PCR 计数略微不同于理想值。利用修改后的 PCR 值就会在解码器中产生相位误差, 其效果与实际抖动并无区别。

上述方法有一个优点, 即只需修改 PCR 数值, 就可以很方便地在任何节目流中附加任意所需大小的抖动, 而且不需要改变所有其它的数据。至于传输流中的其它节目流则无需附加抖动, 但实际上, 最好是用一稳定的节目流作为基准。

按照不同的检测要求, 可以用多种方式对时基加以调制, 以此来确定环路相位误差的频谱, 从而对环路的滤波性能进行测试。方波抖动值会在或前或后相等的间隔内交替变化, 正弦波抖动值会引起具有正弦波采样的相位误差, 而随机抖动则会引起类似于噪声的相位误差。

术语汇编

AAC—高级音频编码。

AAU—音频访问单元。参见访问单元。

AC-3—由杜比实验室发明的音频压缩方案, 并指定为 ATSC 数字电视标准。在消费设备领域中称为杜比数字。

Access Unit (访问单元)—图象或声音块的已编码数据以及其后的缓存 (空值)。

A/D—模拟 / 数字转换器。

AES—音频工程学会。

Anchor Frame (锚定帧)—用作预测参考的视频帧。I 帧和 P 帧一般用作锚定帧, 但 B 帧不是锚定帧。

ANSI—美国国家标准学会。

API—应用编程接口。

ARIB—(日本) 广播工业和商业协会。

Asynchronous Transfer Mode (ATM) (异步转移模式)—用于宽带数字网络中恒定速率和突发信息有效传输的一种数字信号协议。ATM 数字码流中含有长度固定的数据包 (称为“信元”), 每个信元含有 53 个字节 (每字节 8 比特)—其中 5 字节的包头和 48 字节的信息净荷。

ATM—参见 Asynchronous Transfer Mode。

ATSC—（美国）高级电视系统制式委员会。

ATVEF—（美国）高级电视增强论坛。

AU—访问单元。

BAT—节目组相关表。

BER—误码率。

BFSK—二进制频移键控。

BIOP—广播 ORB 间协议。

Bit rate（比特率）—由通道至解码器输入端的已压缩比特流的传送速率。

Block（块）—一块是像素值或 DCT 系数的一个阵列，通常为 8×8 块，代表亮度或色度信息。

Bouquet（节目组）—通过网络 ID 和 PID 的组合来识别节目的一组传输流，它是 DVB—SI 的一部分。

BPSK—双（相）相移键控。

CA—条件接收。用于指示节目是否已被加扰的信息。

CAT—条件接收表。PID 为 1（参见第八章—传输流）的传输流包，它包含关于加扰系统的信息。参见 ECM 和 EMM。

CD—激光唱盘。

CELP—码激励线性预测。

Channel Code（通道编码）—一种编码技术，它将原始数据转换为能够记录或能被无线 / 有线传输的信号。

CIF—通用图象格式。一种 352×240 象素格式，用于 30 帧 / 秒视频会议。

Closed GOP（封闭式 GOP）—某一封闭式图象组中，在对双向编码的最后一帧图象进行编码时不需要下一个图象组中的数据。封闭式 GOP 用于形成比特流中的接合点。

Coefficient（系数）—确定变换中某一特定频率或基本函数幅度的数值。

CORBA—公共对象请求代理机构。

COFDM—编码正交频分复用，OFDM 的一种改进型。一种使用许多载波的数字调制方案，每个载波可以使用很低的数据率传送信息。

Compression（压缩）—减少代表数据项的比特数。

CRC—循环冗余校验。

DAC—数字 / 模拟变换器。

DASE—数字电视应用软件环境。

DAVIC—数字音频视频协会。

DCT—离散余弦变换。

DDB—下载数据块（DownloadDataBlock）。

DET—离散付立叶变换。

DII—下载信息指示（DownloadInfoIndication）。

Dolby Digital（杜比数字）—参见 AC-3。

DSI—下载服务器启动（DownloadServerInitiate）。

DSMCC—数字存储媒体指令和控制。

DST—数据业务表。

DTS—解码时间标记。DTS 是 PES 包头的一部分，它指示访问单元能被解码的时间。

DVB—数字视频广播。一般指的是由欧洲广播业主、制造厂商、管理部门等组成的协会，该组织制定了欧洲的数字电视和数据业务的传送标准，包括 DVB-C（有线），DVB-S（卫星）和 DVB-T（地面）版。

DVB-SI—DVB 业务信息。在 DVB 复用中用于描述各个复用内容的信息，包括 NIT，SDT，EIT，TDT，BAT，RST 和 ST（参见第十章—DVB 和 ATSC 概述）。

DVC—数字视频盒式录像机。

DVD—数字通用光盘或数字视频光盘。

Elementary Stream（基本码流）—由压缩器输出的用于传送单路视音频信号的原始码流。

ECM—授权控制信息。用于规定控制字或其它码流特定加扰参数的条件接收信息。

ECS—增强内容规范。

EIT—事件信息表。DVB-SI 的一部分。

EMM—授权管理信息。用于规定特定解码器授权级别或授权业务的条件接收信息。授权对象可以是单个解码器，也可以是一组解码器。

ENG—电子新闻采集。用于描述以视频记录来代替新闻胶片的术语。

熵编码—对信号的数字形式采用降低冗余度的可变字长无损编码。

EOB—块结束。

EPG—电子节目指南。通过数据传输而非印刷纸张的方式传递的节目指南。

ETSI—欧洲电信标准学会。

GOP—图象组。任一个图象组中的传输顺序是以 I 图象开始，至下一 I 图象前的最后一帧图象结束。

HAVI—家庭声象互操作性（系统）。

Huffmam coding（霍夫曼编码）—一种编码形式，它使用不同字长的代码来表示具有不等拟然性的符号。

IEC—国际电工委员会。

Inter-coding（帧间编码）—利用顺序图象间冗余的压缩编码。也称为时间域编码。

Interleaving（交织）—与误码校正一起使用的一种技术，它将突发误码分解为许多小误码。

Intra-coding（帧内编码）—完全在一帧图象内进行的压缩编码。也称为空间域编码。

IOR—可互操作的对象基准。

IP—因特网协议。

I-pictures（I 图象）—帧内编码图象。

IRD—综合接收机 / 解码器。RF 接收机和 MPEG 解码器的一种组合，它使电视机能够接收数字传输信号。

ISDB—综合业务数据广播，由日本开发的数字广播系统。

ISO—国际标准化组织。

ITU—国际电信联盟。

JPEG—联合图片专家组。

JTC1—IEC 的联合技术委员会。

JVT—联合视频组。

Level（级）—用于给定类别中输入图象的尺寸（参见第二章—视频压缩）。

MAC—媒体访问控制。

Macroblock（宏块）—由若干个亮度和色差 DCT 像素块代表的屏幕区域，整个宏块受一个运动矢量的控制。

Masking（掩蔽）—一种心理声学现象，指的是某一声音在其它声音存在的情况下不能被听见的现象。

MDCT—改进的离散余弦变换。

MGT—原版指南表。

MHP—多媒体家庭平台。

Motion Vector（运动矢量）—分别表示预测参考图象中某一区域的垂直位移和水平位移的一对数值。

MP@HL—高级中的主类。

MP@LL—低级中的主类。

MP@ML—主级中的主类。

MPE—多协议封装。

MPEG—活动图象专家组 ISO/IEC JTC1/SC29/WG11，也指由该专家组开发的标准。

MPEG-LA—MPEG 许可证协议。

NIT—网络信息表。在一个传输流中的用来集中描述许多传输流的信息。

NPT—正常重放时间。

NRT—网络资源表。

Null Packets（空包）—不含数据的“填充”包，填充是为了在使用可用净荷时保持恒定的比特率。空包的 PID 总是为 8191（在二进制中各位数均为 1）。

OCAP—开放的有线应用平台。

OFDM—正交频分复用。

PAL—逐行倒相。

PAT—节目关联表。PID 代码为零的传输流包中所包含的数据（参见第八章—传输流），MPEG 解码器通过它来确定传输流中有哪些节目。PAT 指向 PMT，然后通过 PMT 依次指向每个节目中的视频、音频和数据内容。

PCM—脉冲编码调制。将某种模拟源（如音频或视频信号）波形以周期性的数字样值来表示的技术术语。PCM 是一种未压缩的数字信号。

PCR—节目时钟基准。编码器时钟计数的样值，它在节目流包头中发送，用以解码器的时钟同步。

PES—打包的基本码流。

PID—节目识别符。传输包头中 13 比特的代码。PID 0 指示含有 PAT PID 的传输流包（参见第八章—传输流），PID 1 指示含有 CAT 的传输流包。PID 8191（在二进制中为全 1）指示空包（填充包）。属于同一基本码流的所有包均具有相同的 PID。

PMT—节目映像表。在 PAT 中指向传输流的视频、音频和数据内容的表。

Packets（包）—该术语具有两种含义，应结合上下文判断：在节目流中，包是包含一个或多个显示单元的一个单元；在传输流中，包是一种较小的、具有固定大小的数据单元。

Pixel (像素) — 图象元素 (有时简称为 pel)。图象的最小单元, 它代表一个样值或一组样值, 例如 GBR 或 YCrCb。

Preprocessing (预处理) — MPEG 编码之前的视频信号处理。预处理的例子有降噪, 向下采样 (downsampling), 剪接 — 编辑标识 (cut-edit identification) 和 3:2 下拉标识 (pulldown identification) 等。

Profile (类) — 所使用的编码句法分类。

PS — 节目流。

PSI — 节目专用信息。在 MPEG 传输流中和在每个节目的基本码流中用于跟踪各个节目的信息。PSI 包括 PAT、PMT、NIT、CAT、ECM 和 EMM。

PSI / SI — MPEG PSI 和 DVB-SI 组合在一起的一般术语。

PSIP — 节目和系统信息协议。

PTS — 显示时间标记。即显示单元能被观察者使用的时间。

PU — 显示单元。一帧 (场) 被压缩的图象或一个音频块。

QAM — 正交幅度调制, 一种数字调制系统。

QCIF — 四分之一分辨率 (176×144 像素) 通用互换格式。参见 CIF。

QMF — 正交镜像滤波器。

QPSK — 正交相移键控, 一种数字调制系统, 特别适合于卫星链路传输应用。

QSIF — 四分之一分辨率信源输入格式。参见 SIF。

Quantization (量化) — 逼近信号数值的一种分层处理方法, 在一系列预先定义的数值中, 指配其中一个数值。

RLC — 游程编码。计算相似比特的个数的一种编码方案, 而不是将各相似比特逐个发送。

RRT — 分级区域表。

RST — 运行状态表。

R-S — 里德 — 索罗门是 DVB 使用的一种多项式函数, 它可以给每个传输包提供 8 个字节的保护。

SAOL — 结构化的音频交响乐团语言。

Scalability (可分级性) — MPEG 的一种特性, 通过多个视频数据层以提供多个质量级别的图象。这样, 高级解码器可以使用多个数据层以提供较好的图象质量, 而简单的解码器只使用第一层数据但仍然可提供基本图象。

SDI — 串行数字接口。常用于数字视频信号制作的串行同轴电缆接口标准。

SDK — 软件开发工具包。

SDT — 业务描述表。该表列出了传输流中各项业务的供应商。

SDTV — 标准清晰度电视。

SI — 参见 DVB-SI。

SIF — 信源输入格式。MPEG-1 使用的二分之一分辨率输入信号。

Slice (像素条) — 一组连续宏块序列。

SMPTE — (美国) 电影电视工程师学会。

SNR — 信噪比。

SP@ML — 主级中的简单类。

SPTS — 单个节目传输流。

ST — 填充表。

STB—机顶盒。

STC—系统时钟。在同一节目中用于视频 / 音频编码的共用时钟。

STT—系统时间表。

Stuffing（填充）—附加的无意义数据，用以保持比特率的恒定。

Syndrome（校验子）—误码校验计算的初始结果。一般地，如果校验子为零，则认为没有误码。

TCP / IP—传输控制协议 / 因特网协议。

TDAC—时间混叠消除。AC-3 音频压缩中使用的一种编码技术。

TDT—时间和日期表。DVB-SI 中使用。

TOT—时间偏置表。

Transport Stream（TS，传输流）—若干个节目流的复用。通过不同的包 ID（PID）实现解复用（参见 PSI、PAT、PMT 和 PCR）。

Truncation（舍位）—通过删除低位比特以缩短样值或系数的字长。

T-STD—传输流目标解码器。该解码器具有假定由编码器提交的一定数量的缓冲存储器。

TVCT—地面虚拟频道表。

VAU—视频存取单元。节目流中一个已压缩的图象。

VBV—视频缓存检验。

VCO—压控振荡器。

VLC—可变字长编码。它是将短码指配给较常出现的数值，长码指配给不常出现的数值的一种压缩技术。

VOD—视频点播。在视频点播系统中，电视节目或电影可以应某个用户的请求而单独传送。

VSF—残留边带调制。ATSC 使用的一种数字调制系统。

Wavelet（小波）—使用基本函数的一种变换，其块长度不固定，随着频率的降低而加长。

Weighting（加权）—通过预先乘以数值而改变噪声分布的一种方法。

Y/C — 亮度和色度。

有关详细信息

泰克公司备有内容丰富的各种应用手册、技术介绍和其他资料，并不断予以充实，以帮助那些从事前沿技术研究的工程技术人员。请访问www.tektronix.com

版权© 2002, 泰克有限公司, 保留所有权利。泰克公司的产品受美国和国外专利权法保护, 包括已发布和尚未发布的产品。以往出版的相关资料信息由本手册所代替。泰克公司保留更改产品规格和定价的权利。TEKTRONIX和TEK是泰克有限公司的注册商标。所有其他相关商标名称是其各公司的服务商标、商标或者注册商标。

25W-11418-4