



中华人民共和国国家标准

GB/T 17191.2—1997
idt ISO/IEC 11172-2:1993

信息技术 具有 1.5 Mbit/s 数据 传输率的数字存储媒体运动图像 及其伴音的编码 第 2 部分:视频

Information technology—Coding of moving pictures
and associated audio for digital storage media
at up to about 1.5 Mbit/s—Part 2: Video

1997-12-26 发布

1998-08-01 实施

国家技术监督局 发布

目 次

前言	Ⅱ
ISO/IEC 前言	Ⅳ
引言	V
1 概述	1
1.1 范围	1
1.2 引用标准	1
2 技术要素	2
2.1 定义	2
2.2 符号与缩略语	9
2.3 描述位流的语法规则	11
2.4 要求	13
附录 A(标准的附录) 8×8 离散余弦逆变换	33
附录 B(标准的附录) 可变长码表	33
附录 C(标准的附录) 视频缓冲校验器	40
附录 D(提示的附录) 视频编码指南	42
附录 E(提示的附录) 参考文献	87
附录 F(提示的附录) 专利持有者名单	88

前 言

本标准等同采用国际标准 ISO/IEC 11172-2:1993《信息技术——具有 1.5 Mbit/s 数据传输率的数字存储媒体运动图像及其伴音的编码——第 2 部分:视频》。

GB/T 17191 在《信息技术 具有 1.5 Mbit/s 数据传输率的数字存储媒体运动图像及其伴音的编码》总标题下,目前包括以下 4 个部分:

第 1 部分:系统;

第 2 部分:视频;

第 3 部分:音频;

第 4 部分:一致性测试。

本标准的附录 A、附录 B、附录 C 是标准的附录;附录 D、附录 E、附录 F 是提示的附录。

本标准由中华人民共和国电子工业部提出。

本标准由电子工业部标准化研究所归口。

本标准起草单位:清华大学。

本标准主要起草人:钟玉琢、智西湖。

ISO/IEC 前言

ISO(国际标准化组织)和 IEC(国际电工委员会)是世界性的标准化专门机构。国家成员体(它们都是 ISO 或 IEC 的成员国)通过国际组织建立的各个技术委员会参与制定针对特定技术范围的国际标准。ISO 和 IEC 的各技术委员会在共同感兴趣的领域内进行合作。与 ISO 和 IEC 有联系的其他官方和非官方国际组织也可参与国际标准的制定工作。

对于信息技术,ISO 和 IEC 建立了一个联合技术委员会,即 ISO/IEC JTC1。由联合技术委员会提出的国际标准草案需分发给国家成员体进行表决。发布一项国际标准,至少需要 75%的参与表决的国家成员体投票赞成。

国际标准 ISO/IEC 11172-2:1993 是由 ISO/IEC JTC1“信息技术”联合技术委员会制定的。

ISO/IEC 11172 在《信息技术——具有 1.5 Mbit/s 数据传输率的数字存储媒体运动图像及其伴音的编码》总标题下,目前包括以下 4 个部分:

- 第 1 部分:系统;
- 第 2 部分:视频;
- 第 3 部分:音频;
- 第 4 部分:一致性测试。

附录 A、附录 B、附录 C 构成标准的一部分,附录 D、附录 E、附录 F 仅提供参考信息。

引 言

注：对 MPEG 视频部分总貌感兴趣的读者应先读本引言，接着读附录 D，最后再回到第 1 章和第 2 章。

0.1 目的

本标准是为了满足不断增长的对各种存储媒体上压缩视频的一个统一的表达格式的需求而制定的，如 CD、DAT、温盘和光盘驱动等。本标准的规定可用于位流速率在 1.5 Mbit/s 附近的压缩视频序列的编码表示。使用本标准意味着运动视频可作为计算机数据进行操作，并可通过现有的和将来的网络进行传送和接收。编码表示可用于 625 线和 525 线电视系统，也为用于工作站和个人机的显示提供了灵活性。

本标准主要是按照能提供连续的、传输速率约 1.5 Mbit/s 的存储媒体而制定的。但因采用的方法是一般通用的，故可用于更广的范围内。

0.1.1 编码参数

制定本标准的意图是定义一种具有很高灵活性的适用于多种不同应用的源编码算法。为此目的，位流中包含了许多定义编码位流和解码器特性的参数。对此可给出一个例子：算法对于图像来说适用于各种大小和长宽比，以及信道和设备可工作于很大的位速率范围内。

因本标准可表示的位流的特性范围很大，又定义了一个所谓的“约束参数”的这些编码参数的子集。定义约束参数的目标是对一个广泛的有用参数范围提供指导。满足本子集的约束并不是本标准的要求。位流中一个标志位用来指明位流是否是约束参数的。

0.1.2 约束参数汇总

图片宽度	小于或等于 768 像素点
图片高度	小于或等于 576 线
图片范围	小于或等于 396 宏块
像素速率	小于或等于 396×25 宏块/秒
图片速率	小于或等于 30 Hz
运动矢量范围	小于 -64~+63.5 像素(使用半像素矢量) [backward_f_code and forward3-[-]f_code ≤ 4(见表 D7)]
输入缓冲大小(在 VBV 模式下)	小于或等于 327 680 bit
位速率	小于或等于 1 856 000 bit/s(静态位速率)

0.2 算法概述

本标准定义的编码表示可获得高的压缩比，同时可保持良好的图像质量。由于在编码过程中并不是保存精确的像素值，所以算法不是无损的。编码技术的选择是基于要求高质量的图片和高的压缩比与对编码位流的随机操作的需求之间的权衡。以比较满意的位速率达到良好的图片质量需要很高的压缩比，对此仅用帧内编码是无法完成的。然而对随机操作的需求用纯粹的帧内编码却能获得最满意的结果。这

需要在帧内和帧间编码之间,递归和非递归缩减时间冗余之间进行仔细的权衡。

为达到高压缩比采用了一系列的技术。其中第一个几乎与本标准无关,就是为信号选择一个合适的分辨率。再一个就是利用基于块的运动补偿来减少时间冗余的算法。运动补偿用来根据前幅图对当前图进行因果预测,根据后继图对当前图的非因果预测,或者根据前序图和后继图对当前图进行插值预测。运动矢量是针对每一 16×16 (像素 \times 线)的图像区域定义的。差值信号,即预测差,利用离散余弦变换(DCT)作进一步压缩以消除空间相关性,然后进行量化。在量化过程中要丢掉一些不太重要的信息,因而这是不可逆过程。最后,运动矢量与 DCT 信息相结合,用变长码进行编码。

0.2.1 时域处理

因为对随机操作的要求和高效压缩的要求是相互冲突的,因而定义了三种主要的图类型。帧内编码图(I图)编码时无需参考其他图,它给编码序列的解码的起始提供操作点,但仅能获得中等的编码压缩比。预测编码图(P图)利用根据过去的帧内和预测编码图的运动补偿预测,可作更有效的编码。P图还可用作进一步预测的参考。双向预测编码图(B图)能提供最大限度的压缩,然而需要过去和将来的参考图作为运动补偿。双向预测编码图不能用作预测的参考。序列中三种类型图的组织是非常灵活的。这个选择依赖于对应用的要求,留给编码器去作。图1说明了三种不同类型图之间的关系。

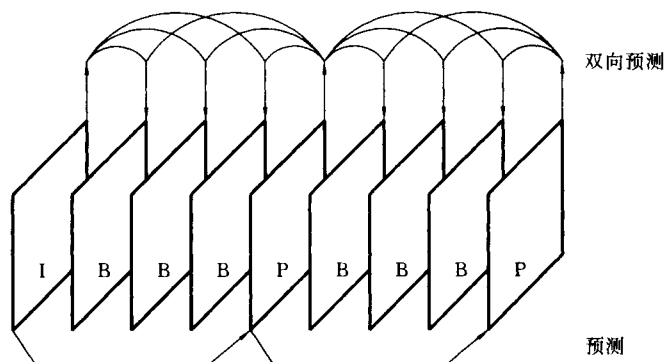


图1 时域图结构示例

本标准中定义的第四种图类型是D图,它提供一种简单而有有限质量的快放模式。

0.2.2 运动补偿——宏块

选择 16×16 的宏块作为运动补偿单位是在使用运动信息所带来的编码效率的提高与其需要的附加存储开销之间折衷的结果。每一宏块可以是多种不同类型中的一种。例如,帧内编码、向前预测编码、向后预测编码以及双向预测编码宏块均允许存在于双向预测编码图中。依照宏块类型,运动矢量信息和其他附加信息同压缩的预测误差信号一样存储于每一宏块中。运动矢量的编码同上一个已编码的运动矢量不同,它使用变长码。矢量可表示的最大长度可以图——图为基础进行编程,以满足最迫切的应用,从而在一般的情况下无需折衷系统的性能。

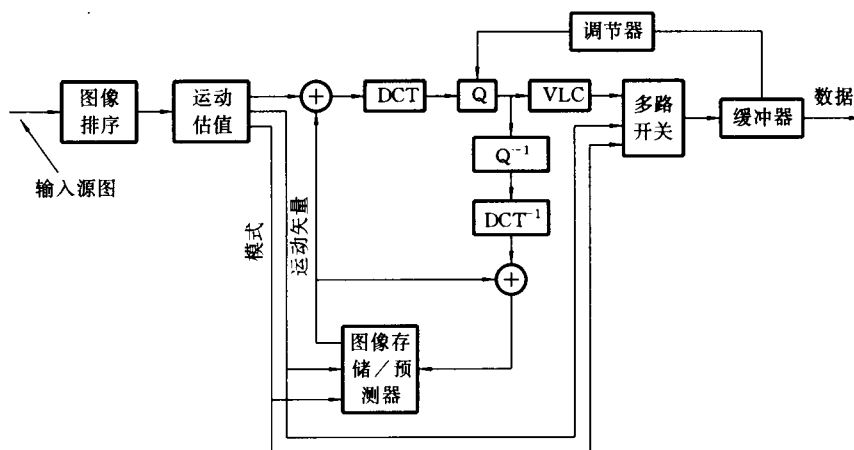
计算合适的运动矢量是编码器的任务。在本标准中没有规定如何去作。

0.2.3 减少空间冗余

原始图像和预测差值信号都有很高的空间冗余。本标准使用一种基于块的视觉加权量化和行程编码的 DCT 方法。作为帧内编码宏块的原始图像以及作为预测编码宏块的预测误差的 8×8 块先变换到 DCT 域并进行比例变换,然后进行量化。量化后许多系数值为 0,于是利用二维游程和变长码来对剩余的系数进行有效化编码。

0.3 编码

本标准中设有规定编码过程。它规定了位流的语法和语义以及解码器中的信号处理。编码器必须在图像质量和编码效率上寻求速度和费用间的权衡上作许多选择。本章简明描述了一个编码器需要完成的功能。图 2 示出了主要的功能块。



其中：DCT —— 离散余弦变换；
 DCT⁻¹ —— 离散余弦逆变换；
 Q —— 量化；
 Q⁻¹ —— 逆量化；
 VLC —— 变长码。

图 2 简化的视频编码器框图

输入的视频信号必须经数字化并表示成一个亮度信号和两个色彩差值信号(Y, C_b, C_r)，然后可进行预处理和格式转换以选择一合适的窗口、分辨率和输入格式。本标准要求色差信号(C_b和 C_r)和亮度信号在垂直和水平方向上按 2:1 进行子采样，并在必要时以非隔行扫描信号进行重格式化。

编码器必须为每幅图选择其类型。确定图类型之后，编码器需估计图中 16×16 宏块的运动矢量。P 图中每一非帧内宏块需一个矢量，而 B 图中则需一个或两个矢量。

如果用 B 图，编码前有必要对图序列顺序做一些调整。因为 B 图的编码用到双向运动补偿预测，它只有在其后继的参考图(I 图或 P 图)解码之后才能够解码。因此，编码器对图像顺序重新做了调整，使得图像按解码器解码的顺序到达解码器。解码器之后又重新恢复正解的顺序。

一幅图内编码的基本单位是宏块。在每幅图中宏块按顺序编码，从左到右，从上到下。每一宏块包含 6 个 8×8 的块：4 个亮度块，1 个 C_b 色度块，一个 C_r 色度块(见图 3)。注意，由 4 个亮度块覆盖的图像区域同每一色度块覆盖的区域相重合。这是为满足人类视觉系统的灵敏度而进行的色度信息子采样造成的。

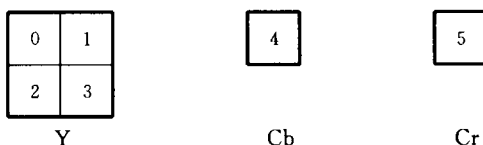


图 3 宏块结构

首先对一给定宏块，选择编码模式。它取决于图像类型，局部的运动补偿预测的有效性，以及块内信号的性质。其次，按照编码模式，形成块内容运动补偿预测要依赖过去和/或将来的参考图。从当前宏块

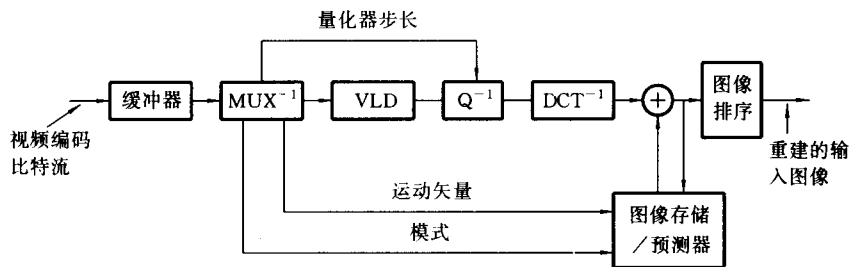
的实际值减去预测值形成差值信号。第三把这些差值信号分成 8×8 的块(4 个亮度块和 2 个色度块), 并对每一块进行离散余弦变换。对每一结果的 8×8 DCT 系数块进行量化, 并对二维的块进行“Z”字扫描, 将其转化成一维量化的 DCT 系数串。第四, 对宏块的附加信息(模式、运动矢量等)和量化的系数数据进行编码。为达最高的编码效率, 对不同数据元素定义一些变长码表。对量化系数数据使用变长编码。

使用不同的图像类型和变长码的一个后果是总的速率是可变的。在固定速率信道的应用中, 可使用一个 FIFO 的缓冲器来缓冲编码器对信道的输出。应监视该缓冲器的状态来控制编码器产生的位数。控制量化过程是控制位速率的最直接方法。本标准中规定了一个缓冲系统的抽象模型(视频缓冲检验器)用以约束用于一幅图的位数最大变化量。这样确保可用一个已知容量的缓冲器对位串流解码。

到此图像的编码表示已经完成。编码器的最后一步是通过解码再次产生 I 图和 P 图以用作后续图像编码的参考图。量化数据经逆量化, 并对每块进行 8×8 DCT 逆变换。产生的预测误差信号再回到到预测信号上, 并限制在要求的范围内, 以产生解码参考图。

0.4 解码

解码是编码的逆操作, 它比编码要简单许多, 因为无需进行运动估值, 所要的选择也很少。在本标准中定义了解码过程。下面是对位流的一种可能解码方法的简要概述。图 4 中示出了主要的功能块。



- 其中: DCT^{-1} ——离散余弦逆变换;
- Q^{-1} ——逆量化;
- MUX^{-1} ——多路分解;
- VLD——可变长解码。

图 4 基本视频解码器框图

对于固定速率的应用, 信道以恒定速率对 FIFO 缓冲器填充位流。解码器读该缓冲器并依据定义的语法对位流中的数据元素进行解码。

解码器读位流过程中识别编码图像起始和图像的类型。它按顺序解码图像中的每一宏块。如果有宏块类型和运动矢量的话, 则用解码器已存储的过去和将来的参考图作为基础构造当前宏块的预测。系数数据经过解码和逆量化, 每一 8×8 的系数块经 DCT 逆变换(见附录 A 所述), 结果加到预测信号上并限制在规定的范围内。

当图中的所有宏块经处理之后, 图像就重构起来了。如果它是 I 图或 P 图则它是后继图的参考图并加以存储, 替代存储最久的参考图。在图像被显示前它们可能需要作从编码的顺序到自然的显示顺序的调整。调整顺序之后, 图像以数字形式作为后处理之用, 并以应用选定的任何方式进行显示。

0.5 编码视频位流的结构

本标准规定了编码视频位流的语法。该语法含有 6 层,每层支持一个信号处理或者一个系统功能:

语义层	功 能
序列层	随机存取单元:上下文
图像组层	随机存取单元:视频
图像层	初始编码单元
片层	重同步单元
宏块层	运动补偿单元
块层	DCT 单元

0.6 算法所支持的特性

使用数字存储媒体上压缩视频的应用程序,除能进行正常的序列播放操作外,还要求能进行一系列的其他操作。编码位流的设计支持这一系列的操作。

0.6.1 随机操作

随机操作是存储媒体上视频的一个极重要的特征,它要求任一图像可在有限时间内进行解码。这意味着位流中存在操作点,即可识别的信息段,它不需参考其他段数据即可解码。每秒的空间内可有两个随机操作点(帧内图)而不会严重损害图像质量。

0.6.2 快速搜索

依赖于存储媒体,有可能从编码位流中扫描操作点(在具体应用目录的帮助下或不在本标准范围内的其他信息的帮助下)以获得快速正放或快速倒放的效果。

0.6.3 反向播放

一些应用可能要求视频信号能进行逆顺序播放,这可通过先用解码器的存储器存储解码后的整个图组,然后再按逆序进行播放。编码器可通过缩短图组长度使这一特征的实现更简单。

0.6.4 鲁棒性

绝大多数的数字存储媒体和信道都是会出错的。应该使用合适的信道编码方案,这已超出本标准的范围了。然而在本标准中定义的压缩方案对残差具有鲁棒性。子图的结构使得解码器可从数据错误恢复过来并对解码重新同步。因而压缩数据的错误会导致解码图像中的错误受到限制。解码器可使用隐含的策略来掩藏这些误差。

0.6.5 编辑

对高效编码和容易编辑的要求是有冲突的。编码结构和语法的设计并没有以简化任何图像的编辑为主要目标。不过,编码中包含了一些特征从而能够对编码数据进行编辑。

中华人民共和国国家标准

信息技术 具有 1.5 Mbit/s 数据 传输率的数字存储媒体运动图像 及其伴音的编码 第 2 部分:视频

GB/T 17191.2—1997
idt ISO/IEC 11172-2:1993

Information technology—Coding of moving pictures
and associated audio for digital storage media
at up to about 1.5 Mbit/s—Part 2: Video

1 概述

1.1 范围

本标准规定了数字存储媒体中视频的编码表示以及解码处理。表示方法除了支持一些特殊功能,如随机操作、快速正放、快速倒放、常速倒放、停顿以及图像静止之外,也支持常速正放。本标准与标准 525 线、625 线电视格式是兼容的,它给个人计算机、工作站的显示使用提供了灵活性。

GB/T 17191 主要适用于支持连续传送速率约 1.5 Mbit/s 的数字存储媒体,比如 CD 盘、数字音频磁带(DAT)以及硬磁盘等。因为其采用了一般化的手段,它可应用于更广的范围内。存储媒体可与解码器直接相连,或者通过通信方法,如总线、LAN 或电信链接。本标准使用约 288 线,每线 352 像素,图速率在约 24 Hz~30 Hz 范围的非间隔扫描视频格式。

1.2 引用标准

下列标准所包含的条文,通过在本标准中引用而构成本标准的条文。本标准出版时,所示版本均为有效。所有标准都会被修订,使用本标准的各方应探讨使用下列标准最新版本的可能性。

- GB/T 17191.1—1997 信息技术 具有 1.5 Mbit/s 数据传输率的数字存储媒体运动图像及其伴音的编码 第 1 部分:系统(idt ISO/IEC 11172-1:1993)
- GB/T 17191.3—1997 信息技术 具有 1.5 Mbit/s 数据传输率的数字存储媒体运动图像及其伴音的编码 第 3 部分:音频(idt ISO/IEC 11172-3:1993)
- GB/T 17191.4¹⁾ 信息技术 具有 1.5 Mbit/s 数据传输率的数字存储媒体运动图像及其伴音的编码 第 4 部分:一致性测试(idt ISO/IEC 11172-4:1995)
- CCIR 建议 601-2 演播室数字电视编码参数
- CCIR 报告 624-4 单色和彩色电视系统特性
- CCIR 建议 648 音频信号的录制
- CCIR 报告 955-2 便携和移动式卫星接收器的声音信号广播 附录 IV 中的先进数字系统 II 的综述
- CCITT 建议 J.17 声音编程电路的预增强
- IEEE 1180:1990 8×8 实现离散余弦逆变换的规范

1) 此标准正在制定中。