

ICS 35.040
L 71



中华人民共和国国家标准

GB/T 17975.2—2000
idt ITU-T H. 262:1995

信息技术 运动图像及其伴音信号的 通用编码 第2部分:视频

Information technology—Generic coding of moving picture
and associated audio information—Part 2: Video

2000-07-14 发布

2001-03-01 实施

国家质量技术监督局 发布

目 次

前言	III
ITU 前言	IV
I 引言	V
I.1 目的	V
I.2 应用	V
I.3 型和级	V
I.4 可分级和不可分级句法	VI
1 范围	1
2 引用标准	1
3 定义	2
4 符号和缩略语	9
4.1 算术运算符	9
4.2 逻辑运算符	10
4.3 关系运算符	10
4.4 位运算符	10
4.5 赋值	10
4.6 助记符	10
4.7 常数	11
5 协定	11
5.1 描述比特流句法的方法	11
5.2 函数的定义	12
5.3 预留、禁用和标记位	12
5.4 算术精度	12
6 视频比特流句法和语义	13
6.1 编码视频数据结构	13
6.2 视频比特流句法	22
6.3 视频比特流语义	37
7 视频解码过程	59
7.1 高层句法结构	60
7.2 变字长解码	60
7.3 反扫描	63
7.4 反量化	64
7.5 反 DCT	67
7.6 运动补偿	68
7.7 空间分级	79

7.8	SNR 分级	90
7.9	时间分级	94
7.10	数据划分	96
7.11	混合分级	97
7.12	解码过程的输出	98
8	型和级	101
8.1	GB/T 17191.2 的兼容性	103
8.2	定义的型之间的关系	103
8.3	定义的级之间的关系	105
8.4	分级层	105
8.5	规定的型、级和层的参数值	107
8.6	解码器兼容性要求	110
附录 A(标准的附录)	离散余弦变换	111
附录 B(标准的附录)	变长码表	112
B.1	宏块寻址	112
B.2	宏块类型	112
B.3	宏块模式	117
B.4	运动矢量	118
B.5	DCT 系数	119
附录 C(标准的附录)	视频缓存校验器	127
附录 D(提示的附录)	算法支持的特性	131
D.1	概述	131
D.2	视频格式	131
D.3	图像质量	132
D.4	数据率控制	132
D.5	短延时模式	133
D.6	随机选取/信道跳换	133
D.7	可分级性	133
D.8	兼容性	138
D.9	本标准与 GB/T 17191.2 的区别	139
D.10	复杂性	141
D.11	编辑编码比特流	141
D.12	特控模式	141
D.13	误码掩盖	142
D.14	级联序列	149
附录 E(提示的附录)	型和级的限定	149
E.1	型中句法元素的限定	149
E.2	允许的层组合	161

前 言

1988年,国际标准化组织(ISO)和国际电工委员会(IEC)联合成立了运动图像专家小组(MPEG)。MPEG小组针对运动图像及其声音的压缩编码研究制定了若干个标准,如MPEG-2标准。MPEG-2标准适用于高质量高码率的图像和声音的压缩编码应用,MPEG-2标准的第二部分涉及运动图像的视频编码,称为《信息技术:运动图像及其伴音信号的通用编码方法——视频部分》。ISO和IEC对MPEG-2视频标准的标准号为ISO/IEC 13818,相应的视频部分的标准号为ISO/IEC 13818-2。ITU等同地采用了ISO/IEC 13818-2,其采标号为ITU-T H. 262。

MPEG-2标准是数字电视广播、数字视频存储和传送的基础标准,运动图像压缩编码是数字电视及相关领域的关键技术之一。在MPEG-2视频标准中,规定了视频数据的编码方法以及重建图像所需的解码处理。它给出了应用于广泛领域内的通用视频编码方案、比特率、图像分辨率和质量。它的基本编码算法是运动补偿预测与DCT的混合。编码的图像可以是隔行扫描,也可以是逐行扫描。必需的算法要素综合在单一句法中,并依据“型”(功能度)和“级”(参数值)定义了有限数目的子集,以便于此种通用编码标准的实际应用。

国际上,MPEG-2标准已在与图像与声音存储和传输相关的领域得到广泛的应用。例如:数字电视广播(标准清晰度和高清晰度)和数字声音广播、数字视频光盘(DVD)、电视节目和电影资料的数字保存、在计算机中运动图像序列的压缩、存储和传送等。MPEG-2标准已成为许多行业的国际通用标准。如在数字电视广播方面,美国、欧洲、日本等已提出了数字电视广播标准的国家和地区,其数字电视广播标准中的图像编码标准都采用MPEG-2标准。

为使我国运动图像的通用编码标准与国际标准接轨,方便电视节目和图像序列的国际交换,方便相应领域的技术交流,利于相关设备和产品的开发和生产,在标准制定过程中决定等同采用国际标准ITU-T H. 262:1995.7《信息技术:运动图像及其伴音信号的通用编码方法——视频部分》,并包括了1996年11月的两个修订本H. 262 Corrigendum 1、H. 262 Corrigendum 2及两个补充本H. 262 Amendment 1、H. 262 Amendment 2,还包括了1998年2月的两个补充本H. 262 Amendment 3、H. 262 Amendment 4。在标准起草过程中,起草人把国家标准草案与相应的国际标准进行了详细的对比,力争达到使该国家标准既忠实于国际标准原文,又符合中文的语言和技术名词的行业习惯用法。本标准在层次划分,章、条设立,编写格式和方法与等同采用的国际标准一致。本标准采用的术语、定义,其全部技术内容与ITU-T H. 262一致。但本标准与等同采用的国际标准存在细微差别。一是在“引用标准”中,对应我国已有国家标准的国际标准,已改成我国的国家标准。二是本标准取消了国际标准中的附录F、附录G、附录H。这些变动对本标准无影响。

本标准的附录A、附录B和附录C是标准的附录。

本标准的附录D和附录E是提示的附录。

本标准由广播电影电视总局提出。

本标准由全国信息标准化技术委员会归口。

本标准起草单位:国家广播电影电视总局广播科学研究院、北京广播学院。

本标准主要起草人:杨杰、张颖辉、张永辉。

ITU 前言

ITU(国际电信联盟)是联合国在电信领域的一个专门机构。ITU 电信标准部门(ITU-T)是国际电信联盟的一个常设机构。ITU-T 制定世界电信标准(建议),179 个成员国、84 个电信运营机构、145 个科研和工业组织以及 38 个国际组织加入 ITU-T。

WTSC 第一号决议(赫尔辛基,1993 年)确定了 ITU-T 成员对建议批准的程序。另外,每四年召开一次的世界电信标准大会,对提交的建议进行讨论并确定以后的研究计划。

在 ITU-T 范围内的某些信息技术领域中,在与 ISO 和 IEC 合作的基础上制定一些必要的标准。1995 年 7 月 10 日通过了 ITU-T 建议 H.262。此建议与 IEC/ISO 13818 国际标准等同。

I 引言

I.1 目的

本标准这部分的制定是为了响应于对各种应用(如数字存贮媒体、电视广播和通信)中活动图像和伴音的通用编码方法的日益增长的要求。使用本标准意味着活动视频可以作为一种计算机数据形式进行处理,并且可以存贮在各种存贮媒体上,可以在现有的和未来的网络中传输和接收,可以在现有的和将来的广播信道中分配。

I.2 应用

本标准的应用包括下列领域,但不限于此。

- BSS 广播卫星业务(到户)
- CATV 有线电视分配(通过光纤网络、铜线等)
- CDAD 有线数字音频分配
- DSB 数字声音广播(地面和卫星广播)
- DTTB 数字地面电视广播
- EC 电子影院
- ENG 电子新闻采集(包括 SNG,卫星新闻采集)
- FSS 固定卫星业务(例如到前端)
- HTT 家庭电视影院
- IPC 个人间通讯(会议电视,可视电话等)
- ISM 交互式存贮媒体(光盘等)
- MMM 多媒体邮件
- NCA 新闻和时事
- NDB 网络数据库业务(通过 ATM 等)
- RVS 远端视频监视
- SSM 串行存贮媒体(数字 VTR 等)

I.3 型和级

本标准拟成为一个通用的应用于较广范围的比特率、分辨率、质量和服务的标准。应用包括数字存贮媒体、电视广播和通信。在制定本标准过程中,已经考虑了典型应用的各种要求,建立了必要的算法要素,并且把它们集合成单一句法。因此,本标准便于不同应用之间的比特流交换。

考虑到实现本标准全部句法的实用性,我们通过“型”和“级”来约定句法的有限子集。在本标准的第3章中将正式定义这些术语和其他相关术语。

“型”是本标准定义的整个比特流句法的一个规定子集。在一个“型”规定的句法范围内,由于比特流中参数的取值的不同对编码器和解码器的性能要求仍然可能有很大的差异。例如,可以规定帧幅的宽(大约)为 2^{14} 样点,高为 2^{14} 行。目前要求能实现处理所有可能帧幅的解码器是不现实和不经济的。

为了解决这个问题,在每一型中又定义了“级”。级是对比特流中各参数进行限定的集合。这些限定可以是对数量的简单限制。也可以采用对参数算法组合加以约束的形式(例如,帧宽乘帧高乘帧频)。

遵循本标准的比特流是使用通用的句法。为获得完整句法的子集,比特流中包含标志和参数,它们指示后面的比特流中是否出现句法元素。为了说明对句法的约束(因此要规定一种型),只需要约束用来

说明后面出现的句法元素作出规定的标志和参数的值。

I.4 可分级和不可分级句法

全部句法可以分成两大类：一是不可分级句法，其结构为 GB/T 17191.2 定义的句法的一个超级集合。不可分级句法的主要特点是有隔行视频信号用的附加压缩工具。第二类为可分级句法，其主要特性为能够从整个比特流的各部分中重建有用视频。这一特性的获得是借助于以两层或多层构建比特流，并从独立的低层开始，再增加若干增强层。低层可以用不可分级句法，或者在某些情况下遵循 GB/T 17191.2 句法。

I.4.1 不可分级句法概述

在不可分级句法中所定义的编码表示方法能获得高压缩比，同时保持好的图像质量。算法不是无损的，因为在编码中不能保留样点的精确值。在特定的比特率下要获得好的图像质量，需要非常高的压缩比，而仅利用帧内图像编码是不能实现的。然而纯帧内图像编码可最佳地满足随机存取的要求。技术的选择要考虑兼顾高质量的图像和高压缩比，以及对编码比特流进行随机存取方面的要求，需要在帧内编码和帧间编码之间，在减少时间冗余度的因果和非因果方法之间仔细地加以平衡。

有许多种技术方法用于获得高压缩比。算法中首先使用基于块的运动补偿，以减少时间冗余。运动补偿，既应用过去的图像对当前图像作因果预测，又应用过去和未来的图像对当前图像作非因果内插预测。为每个 16 样点乘 16 行的图像块定义一个运动矢量。在量化之前对预测误差做进一步压缩，利用离散余弦变换(DCT)去除空间相关性。量化是去除不重要信息的一个不可逆过程。最后，运动矢量与预测误差的 DCT 信息组合在一起，并且使用变字长码进行编码。

I.4.1.1 时间处理

因为随机存取和高效率压缩之间的要求相互冲突，所以定义了三种主要图像类型。帧内编码图像(I 图像)的编码不需参考其他图像。这种编码的图像在编码序列中提供解码开始的存取点，它们的编码只有中等程度的压缩。预测编码图像(P 图像)编码的效率比较高，它使用过去的帧内编码图像或预测编码图像作为参考进行运动补偿的预测。通常，它作为进一步预测的参考。双向预测编码图像(B 图像)提供最高程度的压缩，但运动补偿需要过去的和未来的参考图像。双向预测编码图像永远不作为别的预测的参考(除非在空间分级增强层中将结果图作为参考)。三种类型图像在序列中的组织是非常灵活的。由编码器根据应用的要求来选择。图 I-1 示明三种不同图像类型之间关系的一个例子。

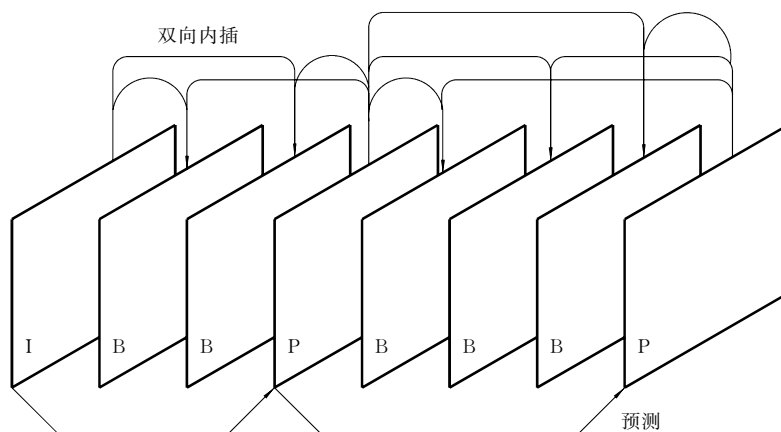


图 I-1 时间上图像结构示例

I.4.1.2 对隔行视频的编码

隔行视频信号的每帧由两场组成,彼此相隔一个场周期。本标准既允许一个帧按一幅图像进行编码,也允许按两幅图像对两场进行编码。可以在逐帧的基础上,自适应地选择帧编码或场编码。帧编码是特别适合于运动不快但细节丰富的视频场景。能够从第一场中预测第二场的场编码,对有快速运动时会更有效。

I.4.1.3 运动表示——宏块

如同在 GB/T 17191.2 中选择 16×16 宏块作为运动补偿单元,它是从运动信息得到编码增益和为表示它所需要开销两者之间平衡的结果。对每个宏块可以以许多不同的方法进行时间预测。例如,在帧编码中,从前面参考图像进行预测既可以基于帧,也可以基于场。依据宏块的类型,对运动矢量信息和其他辅助信息与压缩过的预测误差在每一个宏块中是一起进行编码的。将前面的编码运动矢量作基准,对运动矢量进行差分编码,并使用变字长码在逐幅图像的基础上,能表示运动矢量的最大长度是可以逐帧编程的,所以可以满足大多数需要的应用,而不牺牲大多数正常情况下的系统性能。

计算合适的运动矢量是编码器的任务。本标准未规定对此应如何做。

I.4.1.4 空间冗余降低

原始图像和预测误差信号都有很高的空间冗余度。本标准使用基于块的 DCT 方法,并配合具有视觉加权的量化和游程编码。经运动补偿的预测或内插之后,得到的预测误差图像分成 8×8 的块。对这些块进行 DCT 变换,并在量化之前进行加权。量化之后许多 DCT 系数的值为零,所以可使用二维游程编码和变字长编码对剩下的 DCT 系数进行高效编码。

I.4.1.5 色度格式

除了支持 GB/T 17191.2 的 $4:2:0$ 格式外,本标准还支持 $4:2:2$ 和 $4:4:4$ 色度格式。

I.4.2 可分级扩展

在本标准中的可分级工具被设计用来支持那些单层视频所不能支持的应用系统。在值得注意的应用中,所涉及的领域有视频通信,异步转移模式(ATM)网络中视频的传送,视频标准的交互工作,有多种空间、时间和质量分辨率的分层视频业务,含有常规电视的 HDTV,允许向更高时间分辨率 HDTV 过渡的系统等。虽然,解决可分级视频的简单办法是同播技术,这项技术基于传输/存贮多路独立的视频编码数据,但另一种更有效的是分级视频编码方法,是将分配给一定视频的带宽部分地再用于下一个视频的编码。在可分级视频编码中假设给定一个编码比特流,不同复杂性的解码器都能够进行解码,并显示出合适的重建图像。与单层编码器相比,可分级编码器可能会增加复杂性。不过,本标准提供几种不同形式的可分级方法,它们以相应的复杂性作不重叠的应用。所提供的基本可分级工具有:数据分割、SNR 可分级、空间分级和时间分级。而且,也支持这些基本可分级工具的组合,称为混合可分级。在基本可分级的情况下,允许有称之为较低层和增强层的两层视频,而在混合可分级中可以支持多达三层。下面的表 I-1 至 I-3 给出各种可分级的一些实例应用。

表 I-1 SNR 可分级的应用

较低层	增强层	应用
GB/T 14857	与较低层相同的分解力和格式	两路标准质量电视(SDTV)业务
高清晰度	与较低层相同的分解力和格式	两路 HDTV 质量业务
$4:2:0$ 高清晰度	$4:2:2$ 色度同播	视频制作/分配

表 I-2 空间分级的应用

低 层	增强层	应 用
逐行(30 Hz)	逐行(30 Hz)	CIF/SCIF 兼容或分级
隔行(30 Hz)	隔行(30 Hz)	HDTV/SDTV 可分级
逐行(30 Hz)	逐行(30 Hz)	GB/T 17191.2 与本标准的兼容性
隔行(30 Hz)	逐行(60 Hz)	向高分解力逐行 HDTV 过渡

表 I-3 时间分级的应用

低 层	增强层	更高层	应 用
逐行(30 Hz)	逐行(30 Hz)	逐行(60 Hz)	向高分辨率逐行 HDTV 过渡
隔行(30 Hz)	隔行(30 Hz)	逐行(60 Hz)	向高分辨率逐行 HDTV 过渡

I.4.2.1 空间分级扩展

空间分级是在下面的视频应用中使用的工具,这些应用包括:电信、视频标准的交互,视频数据库浏览,标准电视和高清晰度电视的交互工作等;也就是,具有基本共性的视频系统必需最少有两种空间分解力的视频层。空间分级包括从单个视频源产生两个空间分解力的视频层。低层本身编码来提供基本空间分解力,而增强层应用经空间内插的低层,载有输入视频源的全空间分解力。较低层和增强层可以都使用本标准,或者较低层使用 GB/T 17191.2 标准,增强层使用本标准。后一种情况更有利于视频编码标准之间方便的交互工作。而且,空间分级提供了对每层使用的视频格式选择的灵活性。空间分级的另外一个优点是能够提供对传输误码的抵御,因为可以用误码特性较好的信道传输较低层上较重要的数据,而可以用误码特性较差的信道传输较不重要的增强层数据。

I.4.2.2 SNR 分级扩展

SNR 分级是在下面的视频应用中使用的工具,这些应用包括:电信,多种质量的视频业务,标准电视和高清晰度电视;也就是,具有基本共性的视频系统必须具有最少两种视频质量层。SNR 分级包括从单个视频源产生两个空间分解力相同但视频质量不同的视频层。低层本身编码来提供基本视频质量,而增强层的编码是为了增强低层。在将增强层加到较低层上时,重新产生输入视频的较高质量数据。低层和增强层可以都使用本标准,或者较低层使用 GB/T 17191.2 标准,增强层使用本标准。SNR 分级的另外一个优点是能够提供对传输误码的高度抵御能力,因为可以用误码特性较好的信道传输较低层上较重要的数据,而可以用误码特性较差的信道传输较不重要的增强层数据。

I.4.2.3 时间分级扩展

时间分级是在从电信到 HDTV 的各种视频应用中使用的工具。对于这些应用,从较低时间分辨率系统向较高时间分辨率系统的过渡是必要的。在许多情况中,较低时间分辨率视频系统可以是现存系统或低廉的早期系统,逐渐地推入更复杂的系统。时间分级包括将视频帧进行分层,低层本身编码以提供基本的时间率,增强层依据对较低层的预测进行编码,这些层在解码和时间复用后产生视频源的全时间分辨率。较低时间分辨率系统只能对低层进行解码,提供基本时间分辨率;而将来更复杂的系统可以对这两层进行解码,提供高时间分辨率视频,而且保持与早期系统交互工作。时间分级的另外一个优点是能够提供对传输误码的抵御能力,因为低层上较重要的数据可以通过误码特性较好的许多信道传输,而较不重要的增强层可以通过误码特性较差的信道传输。

I.4.2.4 数据划分扩展

数据划分是在两个信道用于传输或存储视频比特流情况下使用的工具,例如是 ATM 网络、地面广播、磁媒体等情况。比特流划分给两个信道,比特流的较重要部分(诸如首标、运动矢量、低频 DCT 系数)在误码特性较好的信道中传输,较不重要的数据(诸如高频 DCT 系数)在误码特性较差的信道中传输。这样,由于比特流的重要部分得到了较好的保护,使信道误码引起的劣化减到最小。不是专用于对数据划分的比特流进行解码的解码器,不能对任一信道来的数据进行解码。

中华人民共和国国家标准

信息技术 运动图像及其伴音信号的 通用编码 第2部分:视频

GB/T 17975.2—2000
idt ITU-T H.262:1995

Information technology—Generic coding of moving picture and associated audio information—Part 2: Video

1 范围

本标准规定了数字存储媒体和数字视频通信用的图像信息的编码表示,并且规定了解码过程。这种表示支持恒定码率传输、可变码率传输、随机存取、信道跳换、可分级解码、比特流编辑、以及诸如快速正放、快速倒放、慢放、暂停和静像等特殊功能。本标准与 GB/T 17191.2 前向兼容,并与 EDTV、HDTV、SDTV 格式上兼容或下兼容。

本标准主要应用于数字存储媒体、视频广播与通信。存储媒体可以与解码器直接联接,或是通过诸如总线、LAN 或电信链路等通信设施与解码器联接。

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为本标准的条文。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 14857—1993 演播室数字电视编码参数规范(eqv CCIR 601-3)

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 17191.2—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2部分:视频(idt ISO/IEC 11172-2:1993)

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 17576—1998 CD 数字音频系统(idt IEC 908—1987)

GB/T 17975.1—2000 信息技术 运动图像及其伴音信号的通用编码 第1部分:系统
(idt ISO/IEC 13818-1:1996)

IEC 461—1986 视频磁带记录的时间和控制码

ITU-T T.81(JPEG) ISO/IEC 10918-1 连续影调静止图像的数字压缩和编码

ITU-T H.261 p×64 kbit/s 数据率的音视频业务的视频编解码

ITU-T H.320(1997) 建议,窄带可视电话系统及终端设备

ITU-R BR.648 建议书 音频信号的数字记录

ITU-R 955-2 建议书 对(500~3 000) MHz 范围内的移动、便携和固定接收机的卫星声音广播

IEEE 1180—1990 实现 8×8 离散余弦反变换的规范