



中华人民共和国国家标准

GB/T 17191.3—1997
idt ISO/IEC 11172-3:1993

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 3 部分:音频

Information technology—Coding of moving pictures
and associated audio for digital storage media
at up to about 1.5 Mbit/s—Part 3:Audio

1997-12-26 发布

1998-08-01 实施

国家技术监督局 发布

前 言

本标准等同采用国际标准 ISO/IEC 11172-3:1993《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 3 部分:音频》。

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 是标准的附录;附录 C、附录 D、附录 E、附录 F、附录 G、附录 H 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 11172-3:1993 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象及其伴音的编码》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 构成标准的一部分。附录 C、附录 D、附录 E、附录 F、附录 G、附录 H 仅提供参考信息。

引言

注：对运动图像编码标准(MPEG)中音频部分感兴趣的读者，在阅读正文第 1 章和第 2 章之前，应该先阅读本引言、附录 A(图)和附录 C(编码过程)。

为了帮助读者理解存储的压缩数据流和它的译码，下面依次介绍：编码、存储和解码。

0.1 编码

编码器处理数字音频信号，并产生存储所需的数据流。编码器的算法没有标准化，可以使用多种算法进行编码：如音频掩蔽阈值估计、量化和缩放。然而编码器的输出应使 2.4 说明的解码器能产生适合实际应用的音频信号。

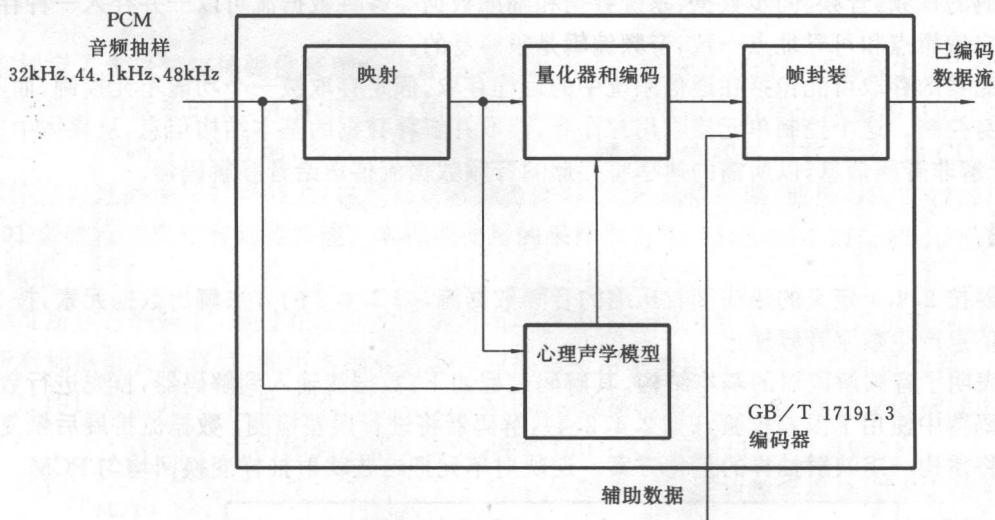


图 1 音频编码器基本结构框图

图 1 表明了音频编码器的基本结构。编码过程如下：输入音频抽样被读入编码器。映射建立输入音频数据流的经滤波的子抽样表示。经映射的抽样可能是子带抽样(如层 I 和层 II，见下面)或变换子带抽样(如层 III)。心理声学模型建立一组控制量化器和编码的数据。这些数据随实际编码器实现的不同而不同。一种可能性是利用音频掩蔽阈值估计来实现量化器的控制。量化器和编码部分从已映射的输入抽样中生成一组编码符号，该部分依赖于编码系统。帧封装将来自其他部分的输出数据汇集成实际数据流，如果需要的话，再加上其他信息(如误差校正)。

有四种不同的编码模式：单声道模式、双声道模式(两个独立的音频信号编在一个数据流内)、立体声模式(立体声的左和右声道编在一个数据流内)和联合立体声模式(带有与立体声不相关或冗余信息的立体声左、右声道编在一个数据流内)。

0.2 层

根据应用需求，可以使用不同层次的编码系统，编码器的复杂性和性能随之增高。本标准音频 N 层解码器能够解出 N 层内或低于 N 层的所有层的数据流。

层 I

这一层包括：将数字音频输入信号映射成 32 个子带。固定分段以便按格式把数据分成块。心理声学模型用以确定自适应位分配，利用块压扩量化并格式化。理论上，层 I 编码/解码的最少延时约为 19 ms。

层 II

这一层提供位分配、缩放因子和抽样的附加编码。使用了不同的帧格式。这层理论上的最小编码/解码延时约为 35 ms。

层 III

这一层采用混合带通滤波器来提高频率分辨率。它还增加了一个差值量化量(非均匀)、自适应分段和量化值的熵编码。这层理论上的最小编码/解码延时为 59 ms。

联合立体声编码能够作为一个附加的特性加入到任何层中。

0.3 存储

已编码的视频、音频、同步数据、系统数据和辅助数据等各种数据流可以一并存入一种存储媒体中。如果将限定编辑点和可寻址点一致，音频编辑是很容易的。

对存储器的存取可能包括在通信系统中的远程存取。假定存取被一个功能单元控制，而不是被音频解码器本身控制。这个控制单元接收用户命令，读取并解释数据的基本结构信息，从媒体中读取已存储的信息，分解非音频信息，以所需的速率将存储的音频数据流传送给音频解码器。

0.4 解码

解码器按 2.4.1 定义的语法接收压缩的音频数据流，按 2.4.2 的方法解出数据元素，按 2.4.3 的规定用这些信息产生数字音频输出。

图 2 表明了音频解码器的基本结构。其解码过程如下：数据流输入到解码器，首先进行数据流扩展，如果在编码器中使用了误差校验，(见 2.4.2.4)，解码器将进行误差检测。数据流扩展后恢复各种信息。重构单元将重构一组映射抽样的量化方案。逆映射单元把这些映射抽样变换回均匀 PCM。

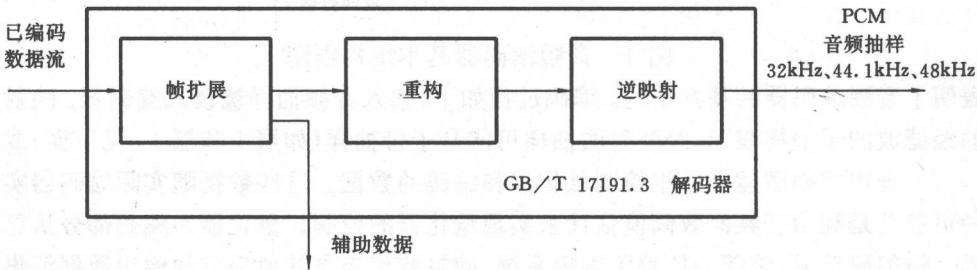


图 2 音频解码器结构框图

目次

前言 Ⅲ

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 1

2.1 定义 1

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 12

附录 A(标准的附录) 图 35

附录 B(标准的附录) 表 39

附录 C(提示的附录) 编码过程 69

附录 D(提示的附录) 心理声学模型 104

附录 E(提示的附录) 错误的位敏感性 131

附录 F(提示的附录) 差错屏蔽 133

附录 G(提示的附录) 联合立体声编码 133

附录 H(提示的附录) 专利持有者名单 136



C9900583

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 3 部分:音频

GB/T 17191.3—1997
idt ISO/IEC 11172-3:1993

Information technology—Coding of moving pictures
and associated audio for digital storage media
at up to about 1.5 Mbit/s—Part 3:Audio

1 概述

1.1 范围

本标准规定了为存储媒体提供高质量的音频编码表示和高质量音频信号的解码方法。编码器的输入和解码器的输出与现在的 PCM 标准兼容,如标准的压缩光盘(CD)和数字音频磁带(DAT)。

本标准以 1.5 Mbit/s 的数据,为应用存储媒体提供连续的音频和视频数据流,如 CD、DAT 和硬盘。存储媒体可以直接连到解码器上,也可以用其他方式,如通过通信线路,或按 GB/T 17191.1 定义的 GB/T 17191 多数据方式与解码器相连。本标准使用的采样率为 32 kHz、44.1 kHz 和 48 kHz。

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 17191.2—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) |
| 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详细说明 | | |
| IEC 908:1987 | CD 数字音频系统 | | |

2 技术要素

2.1 定义

GB/T 17191 采用下列定义。如果仅针对某一部分,则在方括号内注明。

2.1.1 AC 系数[视频] ac coefficient [video]

在一个或两个方向上为非零的所有 DCT 系数。

2.1.2 存取单元〔系统〕 access unit〔system〕

对压缩音频,存取单元是音频存取单元。对压缩视频,存取单元是一幅图象的编码表示。

2.1.3 自适应段〔音频〕 adaptive segmentation〔audio〕

音频信号数字表示的一分段,其时间是可变的。

2.1.4 自适应位分配〔音频〕 adaptive bit allocation〔audio〕

根据心理声学模型,以时间和频率变化的方式对子带进行位分配。

2.1.5 自适应噪声分配〔音频〕 adaptive noise allocation〔audio〕

按照心理声学模型,以时间和频率变化的方式对频带赋以编码噪声。

2.1.6 混迭〔音频〕 alias〔audio〕

奈奎斯特子采样所产生的镜像信号分量。

2.1.7 分析滤波器组〔音频〕 analysis filterbank〔audio〕

在编码器中把宽带 PCM 音频信号变换成一组子采样的子带样值。

2.1.8 音频存取单元〔音频〕 audio access unit〔audio〕

对层 I 和 II,音频存取单元定义为可由自身进行解码的编码位流的最小部分。其中解码是指“完全重构的声音”。对层 III,音频存取单元是指利用以前获得的主要信息可以解码的一部分位流。

2.1.9 音频缓冲器〔音频〕 audio buffer〔audio〕

系统目标解码器中用于存储压缩音频数据的缓冲器。

2.1.10 音频序列〔音频〕 audio sequence〔audio〕

一个不中断的音频帧序列,其中以下的参数保持不变:

——ID(标识符)

——层

——采样频率

——对层 I 和 II:位速率索引

2.1.11 向后运动矢量〔视频〕 backward motionvector〔video〕

按照显示顺序,利用后继图作为运动补偿的参考图中的运动矢量。

2.1.12 巴〔音频〕 Bark〔audio〕

临界频带速率单位。巴尺度是一个从频率尺度到与人耳的频率选择性相对应的音频范围的非线性映射。

2.1.13 双向预测编码图;B图〔视频〕 bidirectionally predictive-coded picture;B-picture〔video〕

使用根据过去和/或将来的参考图的运动补偿预测进行编码的图。

2.1.14 位速率 bitrate

压缩位流从存储媒体输入到解码器的速率。

2.1.15 块压扩〔音频〕 block companding〔audio〕

某一时间内音频信号数字表示的归一化。

2.1.16 块〔视频〕 block〔video〕

一个 8 行 8 列的正交像素块。

2.1.17 边界〔音频〕 bound〔audio〕

利用强度立体声编码的最低子带。

2.1.18 字节对齐 byte aligned

在编码位流中,如果一位的位置(从位流的第一位开始)是 8 的倍数,则该位就称为字节对齐的。

2.1.19 字节 byte

8 位的序列。

2.1.20 信道 channel

存储和传输 GB/T 17191 流的数字媒体。

2.1.21 声道〔音频〕 channel [audio]

立体声信号的左、右信道。

2.1.22 色度(分量)〔视频〕 chrominance (component) [Video]

按照 CCIR 建议 601 中定义的方式,用于表示与主色彩相关的两色差信号之一的矩阵、块或单个像素。

2.1.23 编码音频位流〔音频〕 coded audio bitstream [audio]

按照本标准中规定的音频信号的编码表示。

2.1.24 编码视频位流〔视频〕 coded video bitstream [video]

按照 GB/T 17191.2 的规定,一幅或多幅图像序列的编码表示。

2.1.25 编码顺序〔视频〕 coded order [video]

图像存储和解码的顺序。该顺序未必与显示的顺序一样。

2.1.26 编码表示 coded representation

以编码形式表示的数据元素。

2.1.27 编码参数〔视频〕 coding parameters [video]

表现编码视频位流特征的、可以由用户定义的参数的集合。由编码参数表征位流。由能够解码的位流表征解码器。

2.1.28 分量〔视频〕 component [video]

组成一幅图的 3 个矩阵(亮度和两个色度)之一的矩阵、块或单个像素。

2.1.29 压缩 compression [video]

缩减表示数据项的位数。

2.1.30 恒位速率编码视频〔视频〕 constant bitrate coded video [video]

具有恒定的平均位速率的压缩视频位流。

2.1.31 恒位速率 constant bitrate

压缩位流时从开始到结束的位速率是常数的操作。

2.1.32 约束参数〔视频〕 constrained system [video]

2.4.3.2 中定义的编码参数集中的值。

2.1.33 约束系统参数流(CSPS)〔系统〕 constrained system parameter stream (CSPS) [system]

符合 GB/T 17191.1 的 2.4.6 中定义的约束的 GB/T 17191 多路流。

2.1.34 CRC

循环冗余码。

2.1.35 临界带速率〔音频〕 critical band rate [audio]

频率的心理声学函数。对一给定的可听频率,它正比于低于该频率的临界带的个数。

2.1.36 临界带〔音频〕 critical band [audio]

与人耳的频率选择性相对应的谱域的心理声学度量。

2.1.37 数据元素 data element

编码前或解码后的数据项的表示。

2.1.38 DC 系数〔视频〕 dc-coefficient [video]

在两个方向上频率均为零的 DCT 系数。

2.1.39 DC 编码图;D 图〔视频〕 dc-coded picture;D-picture [video]

仅用其自身信息进行编码的图。在编码表示的 DCT 系数中,仅有 DC 系数。

2.1.40 DCT 系数 DCT coefficient

—给定的余弦基函数的幅值。

2.1.41 解码流 decoded stream

压缩位流的解码重构。

2.1.42 解码器输入缓冲器〔视频〕 decoder input buffer〔video〕

在视频缓冲检验器中规定的先进先出(FIFO)缓冲器。

2.1.43 解码器输入速率〔视频〕 decoder input rate〔video〕

在视频缓冲检验器中规定的且在编码视频位流中编码的数据速率。

2.1.44 解码器 decoder

解码处理的实体。

2.1.45 解码(处理) decoding(process)

GB/T 17191 中定义的处理,即读入编码的位流并产生解码图像或音频采样值。

2.1.46 解码时标;DTS〔系统〕 decoding time-stamp〔system〕

可在数据包头中出现的一个域,它用于表明系统目标解码器解码一个存取单元的时间。

2.1.47 去矫〔音频〕 de-emphasis〔audio〕

对存储或传输的音频信号所采取的一种滤波以消除因加重而引起的线性失真。

2.1.48 逆量化〔视频〕 dequantization〔video〕

量化 DCT 系数在对其位流中的表示进行解码之后和作 DCT 逆变换之前所进行的比例变换处理。

2.1.49 数字存储媒体;DSM digital storage media

数字存储或传输的器件或系统。

2.1.50 离散余弦变换;DCT〔视频〕 discrete cosine transform〔video〕

离散余弦正变换或离散余弦逆变换。DCT 是一种可逆的正交变换。GB/T 17191.2 的附录 A 中定义了 DCT 逆变换。

2.1.51 显示顺序〔视频〕 display order〔video〕

解码图像进行显示的顺序。通常与输入编码器编码的顺序相同。

2.1.52 双声道模式〔音频〕 dual channel model〔audio〕

在一个位流中编码有两个相互独立节目内容(如两种语言)的音频信道模式。编码过程与立体声的相同。

2.1.53 编辑 editing

对一个或多个压缩位流进行操作,以产生一个新的位流。合格的编辑位流必须满足 GB/T 17191 规定的要求。

2.1.54 基本流〔系统〕 elementary stream〔system〕

编码视频、音频或其他编码位流的统称。

2.1.55 加重〔音频〕 emphasis〔audio〕

音频信号在进行存储或传输前所进行的滤波,以提高高频下的信噪比。

2.1.56 编码器 encoder

编码处理的实体。

2.1.57 编码(处理) encoding (process)

在 GB/T 17191 中没有规定的一种处理,读入输入的图像或音频样值流,以产生一个符合 GB/T 17191 中定义的有效位流。

2.1.58 熵编码 entropy coding

信号数字表示的一种变长无损编码,用以减少冗余。

2.1.59 快速正放〔视频〕 fast forward playback〔video〕

按显示顺序,以比真实时间快的速度播放一图像序列或图像序列的一部分。

2.1.60 FFT

快速傅立叶变换。离散傅立叶变换的快速算法(一种正交变换)。

2.1.61 滤波器组〔音频〕 filterbank [audio]

覆盖全部音频范围的一组带通滤波器。

2.1.62 固定分段〔音频〕 fixed segmentation [audio]

分成固定时间段的音频信号数字表示的一段。

2.1.63 禁止 forbidden

定义编码位流的条款中用到的“禁止”表示某值决不能被使用。这常用于避免起始码的误判。

2.1.64 强迫更新〔视频〕 forced updating [video]

不时地对宏块进行帧内编码处理,以确保编码器和解码器间的 DCT 逆变换不匹配误差不至于过分积累。

2.1.65 向前运动矢量〔视频〕 forward motion vector [video]

用于运动补偿的运动矢量,它以显示顺序上在其前的图为其参考图。

2.1.66 帧〔音频〕 frame [audio]

与从音频存取单元的音频 PCM 样值相对应的部分音频信号。

2.1.67 自由格式〔音频〕 free format [audio]

对每层小于最大有效位速率且与所定的位速率不同的所有位速率。

2.1.68 将来参考图〔视频〕 future reference picture [video]

依显示顺序在当前图之后的参考图。

2.1.69 区组〔层Ⅰ〕〔音频〕 granules [Layer I] [audio]

量化前被一起考虑的 32 个子带中连续 3 个子带样值的集合。它们对应于 96 个 PCM 样本。

2.1.70 区组〔层Ⅱ〕〔音频〕 granules [Layer II] [audio]

载有自身辅助信息的 576 频率线。

2.1.71 图组〔视频〕 group of pictures [video]

支持随机操作的一幅或多幅编码图的序列。

2.1.72 汉宁窗口〔音频〕 Hann window [audio]

傅立叶变换之前将逐个样值应用于音频样值块的时间函数。

2.1.73 哈夫曼编码 Huffman coding

一种熵编码方法。

2.1.74 混合滤波器组〔音频〕 hybrid filterbank [audio]

子带滤波器组和 MDCT 的串行组合。

2.1.75 IMDCT〔音频〕

改进的离散余弦逆变换。

2.1.76 强度立体声〔音频〕 intensity stereo [audio]

基于在高频下仅保留左、右声道的能量包络以缩减立体声音频节目中的立体声不相关或冗余的一种方法。

2.1.77 隔行扫描〔视频〕 interlace [video]

常规电视图像的一种性质,其中相间各行在时间上连续。

2.1.78 帧内编码〔视频〕 intra coding [video]

宏块或图像仅用其自身的信息进行编码。

2.1.79 帧内编码图; I 图〔视频〕 intra-coded picture; I-picture [video]

仅用其自身信息进行编码的图。

2.1.80 GB/T 17191(多路复用)流〔系统〕 ISO/IEC 11172(multiplexed) stream [system]

由零个或零个以上的基本流(按 GB/T 17191.1 的定义组成的)组成的位流。

2.1.81 联合立体声编码〔音频〕 joint stereo [audio]

缩减立体声不相关或立体声冗余的所有方法。

2.1.82 联合立体声模式〔音频〕 joint stereo mode [audio]

使用联合立体声编码的音频编码算法的一种模式。

2.1.83 层〔音频〕 layer [audio]

本标准中定义的音频系统编码层次的一个层次。

2.1.84 层〔视频和系统〕 layer [video and systems]

GB/T 17191.2 以及 GB/T 17191.1 中定义的视频和系统规范中数据层次中的一个层次。

2.1.85 亮度(分量)〔视频〕 luminance (component)[video]

表示单色信号的矩阵、块或单个像素,符合 CCIR 建议 601 定义的与基色差信号相联系。表示亮度的符号是 Y。

2.1.86 宏块〔视频〕 macroblock [video]

从图像的亮度分量的一个 16×16 区域得到的 4 个 8×8 的亮度数据块和与之相对应的 2 个色度数据块。宏块有时指像素数据,有时指像素值的编码表示,以及 GB/T 17191.1 所定义的其他宏块层语法中的数据元素。从上下文可明确具体用法。

2.1.87 映射〔音频〕 mapping [audio]

通过子带滤波和/或 MDCT,音频信号从时域到频域的转换。

2.1.88 屏蔽现象〔音频〕 .masking [audio]

人类听觉系统的一种性质,当存在另外一种音频信号时,一种音频信号不能被感知。

2.1.89 屏蔽阈值〔音频〕 masking [audio]

是一个频率和时间的函数,在该值以下,音频信号不能被人类听觉系统所感知。

2.1.90 MDCT〔音频〕

改进离散余弦变换。

2.1.91 运动补偿〔视频〕 motion compensation [video]

利用运动矢量来改善像素值预测的效率。预测利用运动矢量提供相对过去或/和将来参考图的偏移。参考图中包含有用于形成预测差值信号的以前解码的像素值。

2.1.92 运动估值〔视频〕 motion estimation [video]

编码过程中估算运动矢量的处理过程。

2.1.93 运动矢量〔视频〕 motion vector [video]

用于运动补偿的二维矢量,它提供从当前图像的坐标位置到参考图像坐标的偏移。

2.1.94 MS 立体声〔音频〕 MS stereo [audio]

基于对和信号与差信号而不是左右声道进行编码来减小立体声音频节目中出现的立体声不相关和/或冗余的一种方法。

2.1.95 帧间编码〔视频〕 non-intra coding [video]

宏块或图像的编码不仅用到其自身的信息而且还用到其他时间的宏块或图像的信息。

2.1.96 非音调成分〔音频〕 non-tonal component [audio]

音频信号中的一种类似噪声的分量。

2.1.97 奈奎斯特采样 Nyquist sampling

以大于或等于信号带宽的两倍的频率进行采样。

2.1.98 包〔系统〕 pack [system]

一个包由一个包头和跟随其后的一个或多个分组组成。它是 GB/T 17191.1 中描述的系统编码语法中的一层。

- 2.1.99 分组数据〔系统〕 packet data〔system〕
分组中基本流的连续字节。
- 2.1.100 分组头〔系统〕 packet header〔system〕
包含于分组数据中的载有基础流数据信息的数据结构。
- 2.1.101 分组〔系统〕 packet〔system〕
一个分组由一个头和其后面的一些基础数据流中的连续数据字节组成。
- 2.1.102 填充〔音频〕 padding〔audio〕
通过有条件地在音频帧中加入一槽来调整音频帧的平均时间长度,使其与对应的 PCM 采样值的持续期相适应。
- 2.1.103 过去参考图〔视频〕 past reference picture〔video〕
按显示次序出现在当前图之前的参考图。
- 2.1.104 像素纵横比〔视频〕 pel aspect ratio〔video〕
显示像素的标称垂直高度与标称水平宽度之比。
- 2.1.105 像素〔视频〕 pel〔video〕
图元素。
- 2.1.106 图周期〔视频〕 picture period〔video〕
图速率的倒数。
- 2.1.107 图速率〔视频〕 picture rate〔video〕
解码输出图的标称速率。
- 2.1.108 图〔视频〕 picture〔video〕
源、编码的或重构的图像数据。1 个源图或重构图由 3 个 8 Bit 的矩阵组成。3 个矩阵分别代表一个亮度信号和两个色度信号。图层是 GB/T 17191.2 中定义的编码语法的一个层。注意,GB/T 17191 总是用“图”这个词,而不是场或帧。
- 2.1.109 多相滤波器组〔音频〕 polyphase filterbank〔audio〕
一组具有相等带宽的且相位之间有特殊关系的滤波器,它使滤波器组能有效地实现。
- 2.1.110 预测〔视频〕 prediction〔video〕
利用预测值提供当前解码图像的像素值或数据元素的估计值。
- 2.1.111 预测编码图;P 图〔视频〕 predictive-coded picture;P-picture〔video〕
利用从过去的参考图得到的运动补偿预测进行编码的图。
- 2.1.112 预测差值〔视频〕 prediction error〔video〕
像素或数据元素的实际值同预测值的差。
- 2.1.113 预测值〔视频〕 predictor〔video〕
以前解码的像素值或数据元素的线性组合。
- 2.1.114 演示时标;PTS〔系统〕 presentation time-stamp〔system〕
可出现在分组头中的一个域,它指示演示单元出现在系统目标解码器中的时间。
- 2.1.115 演示单元;PU〔系统〕 presentation unit〔system〕
解码的音频操作单元或解码的图。
- 2.1.116 心理声学模型〔音频〕 psychoacoustic model〔audio〕
人类听觉系统屏蔽行为的一个数学模型。
- 2.1.117 量化矩阵〔视频〕 quantization matrix〔video〕
逆量化器使用的一组 8 位的 64 个数。
- 2.1.118 量化 DCT 系数〔视频〕 quantized DCT coefficients〔video〕
逆量化之前的 DCT 系数。量化 DCT 系数的变长编码表示作为压缩视频的一部分进行存储。

- 2.1.119 量化器比例因子〔视频〕 quantizer scalefactor [video]
位流中的一个数据元素,解码过程中用其确定逆量化尺度。
- 2.1.120 随机操作 random access
在任一点上进行读取并对编码的数据流进行解码的处理过程。
- 2.1.121 参考图〔视频〕 reference picture [video]
按显示顺序与当前图最邻近的 I 图或 P 图。
- 2.1.122 顺序调整缓冲器〔视频〕 reorder buffer [video]
系统目标解码器中的缓冲器,用来存储重构的 I 图或 P 图。
- 2.1.123 重量化〔音频〕 requantization [audio]
对编码子带样值进行解码以恢复原始量化值。
- 2.1.124 保留 reserved
当在定义编码位流的条文中使用“保留”一词时,表明其值可用于国家标准将来定义的扩充中。
- 2.1.125 反向播放〔视频〕 reverse playback [video]
以与正常播放顺序相反的顺序显示图像序列的过程。
- 2.1.126 比例因子带〔音频〕 scalefactor band [audio]
层Ⅲ中的一组频率线,它可通过一个比例因子进行缩放。
- 2.1.127 比例因子索引〔音频〕 scalefactor index [audio]
比例因子的一个数值码。
- 2.1.128 比例因子〔音频〕 scalefactor [audio]
量化前对一组数值进行缩放的因子。
- 2.1.129 序列头〔视频〕 sequence header [video]
编码位流中的一种数据块,包含有一些数据元素的编码表示。
- 2.1.130 辅助信息 side information
位流中的控制解码器所需的信息。
- 2.1.131 空白宏块〔视频〕 skipped macroblock [video]
没有存储数据的宏块。
- 2.1.132 子图〔视频〕 slice [video]
一系列的宏块。它是 GB/T 17191.2 中定义的编码语法的一层。
- 2.1.133 槽〔音频〕 slot [audio]
槽是位流的一个基本部分。在层Ⅰ中,一个槽为 4 个字节;在层Ⅱ、Ⅲ中,一个槽为 1 个字节。
- 2.1.134 源流 source stream
压缩编码前,一个单一的非多路复用的采样值流。
- 2.1.135 传播函数〔音频〕 spreading function [audio]
描述屏蔽频率传播的函数。
- 2.1.136 起始码〔系统和视频〕 start codes [system and video]
嵌入编码位流中的唯一的 32 位码。它可用于包括标识编码语法中的层在内的多种目的。
- 2.1.137 STD 输入缓冲器〔系统〕 STD input buffer [system]
系统目标解码器输入时的一种先进先出缓冲器,用于存储解码前基础流的压缩数据。
- 2.1.138 立体声模式〔音频〕 stereo mode [audio]
是这样一种模式:形成一个立体声对(左和右)的两个音频信道编码在一个位流中。编码过程与双通道模式相同。
- 2.1.139 填充(位);填充(字节) stuffing(bit);stuffing(bytes)
插入到压缩位流中的码字,但在解码过程中被丢弃。它用于增加流的位速率。

2.1.140 子带〔音频〕 subband [audio]

音频频带的一部分。

2.1.141 子带滤波器组〔音频〕 subband filterbank [audio]

覆盖全部音频频率范围的一组带滤波器。本标准中子带滤波器组是一个多相滤波器组。

2.1.142 子带样值〔音频〕 subband samples [audio]

音频编码器中的子带滤波器组产生一输入音频流的滤波、子采样表示。其中滤波样值叫子带样值。从输入的 384 个时间连续的音频样值中,在 32 个子带中每个子带产生 12 个时间连续的子带样值。

2.1.143 同步字〔音频〕 syncword [audio]

音频流中嵌入的一个 12 位码,它用于标识一帧的开始。

2.1.144 合成滤波器组〔音频〕 synthesis filterbank [audio]

解码器中的用于从子带样值重构 PCM 音频信号的滤波器组。

2.1.145 系统头〔系统〕 system header [system]

GB/T 17191.1 中定义的一种数据结构,它载有概括 GB/T 17191 多路流系统特征的信息。

2.1.146 系统目标解码器;STD〔系统〕 system target decoder [system]

解码过程的一个假设的参考模型,它用于描述 GB/T 17191 多路流的语义。

2.1.147 时标〔系统〕 time-stamp [system]

用来标明一个事件发生的时间。

2.1.148 三重组〔音频〕 triplet [audio]

一个子带中的三个连续的子带样值集合。32 个子带的三重组形成一个区组。

2.1.149 音调分量〔音频〕 tonal component [audio]

音频信号的一个正弦波状的分量。

2.1.150 可变位速率 variable bitrate

解码压缩位流过程中位速率随时间而变化的操作。

2.1.151 变长编码;VLC variable length coding

一种可逆的编码过程,对较常出现的事件赋予较短的码字,不常出现的事件赋予较长的码字。

2.1.152 视频缓冲检验器;VBV〔视频〕 video buffering verifier [video]

概念上与编码器的输出相连接的一种假设的解码器。它用于约束编码器或编辑过程中生成的数据速率的变化量。

2.1.153 视频序列〔视频〕 video sequence [video]

一个或多个图组组成的序列。它是 GB/T 17191.2 所定义的编码语法的一个层。

2.1.154 “Z”字扫描顺序〔视频〕 zig-zag scanning order [video]

DCT 系数的一种从(大约)最低空域频率到最高空域频率的顺序。

2.2 符号与缩略语

用于描述本标准的数学运算符号与 C 程序设计语言中的类似。然而,关于整数除法的取位和舍入又有专门的规定。位运算符的定义假设整数是用补码表示的。数字和计数的循环一般从零开始。

2.2.1 算术运算符

+	加。
-	减(作为二元运算符)或负(作为一元运算符)。
++	增 1。
--	减 1。
*	乘。
^	乘方。

/	整数除法(结果取整)。例如:7/4 和 -7/-4 取整结果为 1, -7/4 和 7/-4 取整结果为 -1。
//	整数除法(结果四舍五入)。若不特殊说明小数部分的 5 作进位处理。如 3//2=2, -3//2=-2。
DIV	整数除法(往负向取整,即取小于正常除法运算结果的最大整数)。
	绝对值。 $ x = \begin{cases} x, & x \geq 0 \\ 0, & x = 0 \\ -x, & x < 0 \end{cases}$
%	取模运算,仅对正数有定义。
Sign()	$\text{Sign}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}$
NINT()	取整运算符,将其实参舍入成与它最接近的整数。小数部分的 5 作进位处理。
sin	正弦。
cos	余弦。
exp	指数。
$\sqrt{\quad}$	平方根。
log ₁₀	以 10 为底的对数。
log _e	以 e 为底的对数。
log ₂	以 2 为底的对数。
2.2.2 逻辑运算符	
	逻辑“或”。
&&	逻辑“与”。
!	逻辑“非”。
2.2.3 关系运算符	
>	大于。
>=	大于或等于。
<	小于。
<=	小于或等于。
=	等于。
!=	不等于。
max[,...]	取参数中的最大值。
min[,...]	取参数中的最小值。
2.2.4 位运算	
使用按位运算时,操作数均假设采用补码表示。	
&	“与”。
	“或”。
>>	右移(带符号扩展)。
<<	左移(移后填 0)。
2.2.5 赋值运算	
=	赋值运算符。
2.2.6 助记符	

下列助记符用于描述编码位串流中的不同数据类型。

bslbf	位串, 左边位开始, GB/T 17191 规定位串从左开始书写。位串是单引号内的 1 或 0 组成的串, 如 '1000 0001'。位串当中的空白只是为了阅读方便, 没有其他意义。
ch	通道。如果 ch 值是 0, 是指立体声信号的左通道或两个独立信号中的第一个信号(音频)。
nch	通道号; 等于 1 表示单通道模式, 等于 2 表示其他模式(音频)。
gr	音频层 II (Layer II) 3 * 32 子带采样的粒度; 音频层 III (Layer III) 18 * 32 子带采样中的粒度。
main_data	位串流的 main_data 部分包含比例因子、哈夫曼编码数据以及辅助信息(音频)。
main_data_beg	位于位串流中帧的 main_data 的开始处, 具体位置是前帧 main_data 末端的位置再加上 1 位。它由前帧的 main_data_end 值计算出来(音频)。
part 2_length	main_data 中比例因子所占的位数(音频)。
rpchof	余数多项式的系数, 最高幂在前(音频)。
sb	子波段(音频)。
sblimit	没有分配位的最低子波段的个数(音频)。
scfsi	选择比例因子的信息(音频)。
switch_point_l	使用了窗口开关上的点的长块比例因子波段的个数(音频)。
switch_point_s	使用了窗口开关上的点的短块比例因子波段的个数(音频)。
uimsbf	无符号整数, 最高位在前。
vlclbf	变长码, 左边位开始。这里“左”指 VLC 码的书写顺序。
window	当 block_type=2 时, 实际时间槽的个数, $0 \leq \text{window} \leq 2$ (音频)。

2.2.7 常量

π	3.14159265358.....。
e	2.71828182845.....。

2.3 描述位流的语法规则

2.4.3 中描述了解码器得到的位流。位流中的每一数据项用黑体字印刷, 并描述了它们的名称、位长度、类型助记符以及传送顺序等。在位流中由解码数据元素所引起的动作与该数据元素的值以及前面解码的数据元素有关。2.4.4 中描述了数据元素的解码以及解码中所用的状态变量的定义。以下结构用于表述数据元素何时出现以及何时为正常类型的条件。

注意, 本语法中使用了这样的“C”代码约定: 一个非零的变量或表达式相当于一个为真的条件。

while(条件){	如条件为真, 数据流中将出现数据元素组, 直到条件为假为止
data_element	
...	
}	
do{	
data_element	数据元素至少出现一次。
...	
}while(条件)	数据元素不停重复直到条件为假。
if(条件){	
data_element	如果条件为真, 则接下来数据流中出现第一组数据元素。
...	