

多媒体关键技术规范

MPEG运动图象 压缩编码标准

(ISO/IEC 11172)

清华大学计算机系 杨品 钟玉琢 蔡莲红 译



机械工业出版社

73-542
661

多媒体关键技术规范

MPEG 运动图象压缩编码标准

(ISO / IEC 11172)

清华大学计算机系 杨 品 钟玉琢 蔡莲红 译



机械工业出版社

9510212

数字式视频和音频数据的压缩编码和解码技术是多媒体计算机、多媒体通信(可视电话、视频会议)、常规电视数字化、高清晰度电视(HDTV)以及交互式电视系统中最核心、最关键的技术。近年来,各国对这一技术的研究和开发已取得了卓有成效的进展。在此基础上,国际标准化组织于1993年8月正式通过并编辑出版了ISO/IEC11172《信息技术——具有1.5Mbit/s数据传输率的数字存储媒体运动图象及其伴音的编码标准》。本标准内容涉及有关系统、视频和音频的定义、符号、语法规则、要求及适用范围等,叙述详尽而具体,附录中还给出了编码的描述、算法、原理、模型以及防止出错的方法等。本书对于从事多媒体技术研究、开发和应用的人员颇有裨益。

INTERNATIONAL STANDARD

ISO/IEC 11172

First edition 1993-08-01

Information technology—Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s—

Part1: Systems Part2: Video Part3: Audio

图书在版编目(CIP)数据

MPEG 运动图象压缩编码标准 (ISO/IEC 11172) / ISO, IEC; 杨品等译.—北京: 机械工业出版社, 1995.10

书名原文: ISO/IEC 11172 Information technology—coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s—part1, part2, part3

ISBN 7-111-04763-X

I.M… II.① I… ② I… ③ 杨… III.图象处理—数据压缩—图象编码—国际标准 IV.TN919.8-65

中国版本图书馆 CIP 数据核字(95)第 06310 号

出 版 人: 马九荣 (北京市百万庄南街1号 邮政编码 100037)

责任编辑: 张蔼玲 版式设计: 霍永明 责任校对: 肖新民

封面设计: 肖 晴 责任印制: 路 琳

机械工业出版社印刷厂印刷·新华书店北京发行所发行

1995年9月第1版第1次印刷

787mm×1092mm¹/16·18印张·438千字

0 001—3 000 册

定价: 34.00 元

9510120

译者序

数字式视频和音频数据的压缩编码和解码技术是多媒体计算机、多媒体通讯（可视电话、视频会议）、常规电视数字化、高清晰度电视（HDTV）以及交互式电视（Interactive Television）系统中的关键技术。从 1948 年 Oliver 提出 PCM 编码理论开始，迄今已有 40 年历史。很多科学家、工程师已先后提出了：预测编码（Predictive Coding）、变换编码（Transform Coding）、矢量量化编码（Vector Quantization Coding）、信息熵编码（Entropy Coding）、子带编码（Subband Coding）、结构编码（Structure Coding）以及基于知识的编码（Knowledge-Based Coding）等，其目的是为了从原始数据中去除信息的时间冗余、空间冗余（诸如信息熵冗余、结构冗余、知识冗余以及视觉冗余等无用信息），保留有用信息，以便提高信息处理、传输和存储效率。

前几年世界上很多国际性的大公司已经研制出不少种多媒体计算机系统，其中卓有成效的公司和系统有：Commodore 公司的 Amiga 系统、Philips/Sony 公司的 CD-I 系统、Apple 公司的 Quick Time 系统以及 Intel/IBM 公司的 DVI 系统，它们都有自己的视频音频信息压缩编码和解码算法，如 DVI 系统中的 PLV 算法。为了提高速度，这些公司还投资了大量研究费用设计了专用的硬件芯片（如 DVI 系统中的 82750 芯片），在当时的国际市场上取得了良好的经济效益和社会信誉，如 Intel/IBM 公司的 DVI 系统荣获了 1991 年秋季 Comdex 博览会最佳多媒体产品奖（Best Multimedia Product）以及最佳展示奖（Best of show）。但是随着最近几年国际标准化组织 ISO（The International Organization for Standardization）先后制定了：ISO/IEC10918 连续色调静态图象的数字压缩和编码、ISO/IEC 11172 具有 1.5Mbit/s 数据传输率的数字存储媒体运动图象及其伴音的编码以及 CCITT 制定的 H.261 P×64 kbps 视声服务用视象编码方式等国际标准，使得上述产品必须逐步走向标准化。

标准化是产业化活动成功的前提，国际标准化组织 ISO 于 1988 年就成立了一个致力于制定有关运动图象压缩编码的组织——MPEG(Moving Picture Experts Group)。MPEG 组织在工作一开始就考虑到一些相关标准化组织的研究成果，这不仅仅是为了避免重复性的工作，更重要的是为 MPEG 提供了重要的背景和技术。对 MPEG 组织影响最大的是 JPEG (Joint Photographics Coding Experts Group) 组织。这两个组织同属于 ISO 的一个工作组。JPEG 的工作主要集中于静止图象的压缩编码（帧内的压缩编码），而 MPEG 既包含了帧内静止图象的编码，也包含了运动图象帧间运动图象的编码。因此，可以认为 MPEG 的工作是 JPEG 工作的延续。

MPEG 工作组制定的 ISO/IEC 11172 具有 1.5Mbit/s 数据传输率的数字存储媒体运动图象及其伴音的编码国际标准，已于 1993 年 8 月正式通过并第一次编辑出版。该标准由系统、视频、音频以及一次性测试四部分组成，每部分都给出了适用范围、标准参考文献、定义、符号和缩写词、描述比特流（位串流）的语法规则以及要求。每部分的附录还给出了

详尽而具体的编码描述、算法、原理、模型以及防止出错的方法。

尽管 JPEG, H.261, MPEG 三种图象的压缩/解压缩标准是针对不同的应用领域,各自有各自的对象,但是它们之间也是相互关联的。从层次上来说, MPEG 层次最高,应用范围也最广,而且随着多媒体和网络技术的进一步发展,今后会日益显示出它的潜力。

本书第一部分——系统由钟玉琢教授主译,第二部分——视频由杨品教授主译,第三部分——音频由蔡莲红教授主译。三位译者又互相进行了校订。在全书的翻译过程中,王凯、周红、王榕、吴怀军、徐宏坤、赵世霞、刘红梅、陈俊华、何恩光、吴卫等同志提供了部分译文初稿并参加了部分复审工作,在此一并表示感谢。

视频和音频压缩编解码技术是一个正在研究的课题,我们希望本书的出版,对广大从事多媒体技术的研究人员和科学工作者有所裨益。必须指出,目前这一技术正处在蓬勃发展阶段,因此新的国际标准以后还会相继推出和出版,以适应新的发展和需要。

限于译者学识水平,再加上时间比较仓促,书中不足和错误之处,恳请读者批评指正。

译 者

1995年4月

目 录

译者序

信息技术——具有 1.5Mbit/s 数据传输率的数字存储媒体运动图象及其伴音的

编码 ISO/IEC 11172-1 系统	1
前言	2
导论	3
1 总论	6
1.1 范围	6
1.2 标准参考资料	6
2 技术要素	7
2.1 定义	7
2.2 符号与缩写词	13
2.3 描述位串流的语法规则	15
2.4 要求	17
附录	29
A 系统编码层的描述	29
B 专利持有者名单	52

信息技术——具有 1.5Mbit/s 数据传输率的数字存储媒体运动图象及其伴音的

编码 ISO/IEC 11172-2 视频	59
前言	60
导论	61
1 总论	67
1.1 范围	67
1.2 标准参考资料	67
2. 技术要素	7
2.1 定义	7
2.2 符号与缩写词	13
2.3 描述位串流的语法规则	15
2.4 要求	67
附录	90
A 8×8 离散余弦逆变换	90
B 可变长码表	90
C 视频缓冲校验器	97
D 视频编码指南	98
E 参考文献	151

F 专利持有者名单	151
信息技术——具有 1.5Mbit/s 数据传输率的数字存储媒体运动图象及其伴音的	
编码 ISO/IEC 11172-3 音频	153
前言	154
导论	155
1 总论	157
1.1 范围	157
1.2 标准参考资料	157
2 技术要素	7
2.1 定义	7
2.2 符号与缩写词	13
2.3 描述位串流的语法规则	15
2.4 要求	157
附录	183
A 图	183
B 表	186
C 编码过程	211
D 心理声学模型	250
E 错误的位敏感性	276
F 错误隐蔽	278
G 联合立体声编码	278
H 专利持有者名单	281

**信息技术——具有 1.5Mbit / s 数据传输率的数字存储媒体运动图
象及其伴音的编码 ISO / IEC 11172-1 系统**

前言

ISO (The International Organization for Standardization)和 IEC (The International Electrotechnical Commission) 是制定国际标准的组织机构。ISO 和 IEC 的国际成员通过相应组织建立的技术委员会开展特殊领域的技术活动, 参与国际标准的开发工作。ISO 和 IEC 技术委员会在共同感兴趣的领域协同工作。与 ISO 和 IEC 有联系的其它国际组织、政府和非政府组织也参加了这项工作。

在信息技术领域, ISO 和 IEC 已建立了一个联合技术委员会, 即 ISO / IEC JTC1, 由联合技术委员会批准的国际标准草案被提交给各国委员投票表决。作为一个国际标准而公开发表需要至少 75% 的国家成员投票批准。

国际标准 ISO / IEC 11172-1 是由联合技术委员会 ISO / IEC JTC1 的信息技术分会 SC29 制定的。该标准包括音频、图象、多媒体和超媒体信息的编码表示。

ISO / IEC 11172 总的标题是: 信息技术——具有 1.5Mbit / s 数据传输率的数字存储媒体运动图象及其伴音的编码, 它由下述几部分组成:

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

此外, 还有 ISO / IEC 11172 的附录 A 和 B。

导论

系统的技术规范提出了把一个或多个符合国际标准的视频和音频数据流与时间信息结合起来形成单一数据流问题。一旦形成了单一数据流，数据将适合于数字式存储和传输，系统技术条件中规定的语法和语义规则使同步播放成为可能，这种播放在各种条件下数据流的恢复和接受都不会引起解码缓冲区的上溢和下溢。语法和语义规则以不同形式载于系统技术规范中，语法规则仅适用于系统层编码，并不能扩展到视频和音频压缩层编码，而语义规则完全适用于混合数据流。

系统的技术规范并没有指定编码器和解码器的体系结构和具体实现，而比特流（也称位流）的性质把需要的功能和特性施加于编码器和解码器。例如，编码器必须满足最小的时钟允许误差。不仅如此，还有其它方面的要求，在设计和实现一个编码器和解码器的过程中，要考虑留有一定的自由度。

为了说明 ISO / IEC 11172 解码器的功能，图 1-1 描绘出音频 / 视频解码系统原型样机的框图。这一系统结构并不是唯一的，系统解码功能（包括时序控制）可以分成基本的数据流解码和媒体规定的解码，但是解码器原理样机图对讨论问题是非常有用的，这还意味着它是设计 ISO / IEC 11172 解码器的标准。实际上，非音频 / 视频数据也是允许的，但这一点并未在图中表示出来。

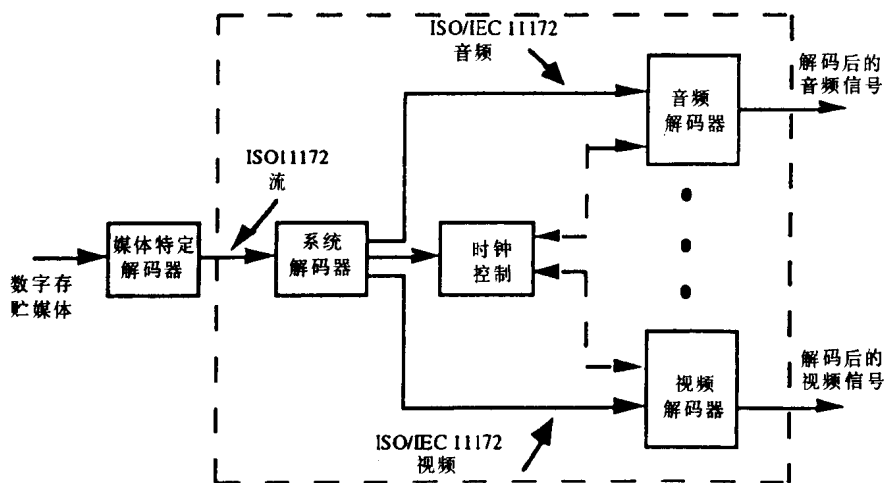


图 1-1 ISO / IEC 11172 解码器样机原理图

图 1-1 所示原型 ISO / IEC 11172 解码器由系统、视频和音频三部分组成，它们分别符合本标准的一、二、三部分。在这个解码器中，一路或多路声音或图象数据流的复合编码表示，假定以某种媒体的特定格式存储在 DSM（数字存储媒体）或网络中，存储的格式在本标准中未做规定。同样本 ISO / IEC 11172 原型解码器也未规定特定媒体的解码部分。

这个原型解码器把 ISO / IEC 11172 数据流作为输入，利用系统解码器可以从数据流中

把时间信息分离出来。系统解码器把数据流分成多路，生成的每一路基本数据流可作为声音或图象解码器的输入，视频、音频解码器的输出形成为了解码后的视频、音频信号。在设计中包含而在原型样机原理图中并没有表示出来的，是系统解码器、视频音频解码器中以及媒体特殊解码器中的时间信息流。视频音频解码器中的视频和音频信息以及同 DSM 之间的同步问题就使用这些时间信息解决。

ISO / IEC 11172 数据流被分成两层：系统层和压缩层。输入到系统解码器的数据流既有系统层又包含了压缩层；输入到视频音频解码器的数据流仅有压缩层。

系统解码器的操作既适应于 ISO / IEC 11172 多路数据流（多路操作），又适应单独的基本数据流（特定的数据流操作）。ISO / IEC 11172 系统层可分成两个子层：一个是用于多路操作（包层），另一个是用于特殊数据流操作的（分组层）。

（1）多路操作（包层）

多路操作包括 DSM 数据恢复的调整，时钟的调整以及缓冲区的管理，这些任务是密切相关的。如果 DSM 传送数据率是可控的，那么可以把 DSM 传送数据率调到使解码器的缓冲区既不上溢也不下溢；但是如果 DSM 的速率不能控制时，那么基本数据流解码器必须能够跟踪 DSM 的时序以避免上溢或下溢。

ISO / IEC 11172 数据流由一系列包组成，这些包的头部即可解决上述任务。每个包的头都规定了每个字节从 DSM 输入到系统解码器的预期时间，这个目标达到时间表可以作为时钟调整和缓冲区管理的参考。解码器无须完全依照达到时间表，但是必须根据它校正偏差。

附加的多路操作是解码器为了解码 ISO / IEC 11172 多路数据流形成所需资源的能力。每个 ISO / IEC 11172 多路数据流的第一个包将传送参数帮助解码器解决这一任务，例如，它包括数据流的最大数据率和同步视频通道的最多个数。

（2）单一数据流操作（分组层）

特定数据流最基本的操作是：a. 多路数据的分离；b. 多个基本数据流的同步重放，下面将阐述这些问题。

① 多路数据的分离

在编码过程中，ISO / IEC 11172 多路数据流是由多路基本数据流构成的，基本数据流不仅包括 ISO / IEC 11172 音频视频数据流，还包括专用的、保留的或填充的数据流。这些数据流暂时分成分组包，这些分组包是串行的。这些分组包都由一个且仅有一个基本数据流编码字节组成。

根据 2.4.3.3 节以及 2.4.5 和 2.4.6 节中规定的约束条件，允许固定和可变分组包的长度。

在解码过程中，多路 ISO / IEC 11172 数据流重新构成基本数据流，因此需要进行多路数据的分离。分组包头中 stream-id 码使多路数据的分离成为可能。

② 同步

在 ISO / IEC 11172 多路数据流中，多个数据流的同步是用演示时间戳来实现。时间戳的基本频率是 90kHz。在 N 个数据流重放过程中，调整所有的重放数据流同步于一个基本的时间基准，而不是调整重放的一个数据流去匹配其它的数据流。这个基本的时间基准可以是 N 个解码器时钟之一，DSM 的时钟，通道时钟或者是外部时钟。

由于演示时间戳依赖于单一基本数据流的解码过程，因此它们是在分组层中。在捕获时编码器记录时间戳，时间戳与相应的编码数据一起传送到解码器时以及当解码器利用这些时间安排演示时，头尾相连的同步得以实现。

在多路数据流中用 DSM 时基的时间戳也能够实现同步。

③ 和压缩层的关系

在某种意义上讲分组层与压缩层是独立的，但也不完全如此。如果分组层不需要从压缩层起始码开始，从这个意义上讲，正如第 2（视频）和 3 部分（音频）中定义的那样，它是独立的。例如，一个视频分组包可在视频数据流任一处开始。但是，位于分组包头中的时间戳编码时用压缩层结构中的演示时间表示（即演示单位）。

（3）系统参考解码器

ISO / IEC 11172 第一部分使用了“系统目标解码器（STD—System Target Decoder）”给出了时序和缓冲关系的表达式。由于 STD 可以根据 ISO / IEC 11172 定义域的值规定参数（例如，缓冲区大小），因此每个多路数据流都可得出一组 STD 自己的参数。它能够确保比特流能够以正常速度播放以及与 STD 保持一致的向前播放。实际的解码器可以假定数据流完全按照 STD 模型播放，但由于它与 STD 设计方法不同，实际解码器必须进行补偿。

1 总论

1.1 范围

ISO / IEC 11172 的这部分详述了系统层的编码过程。经过改进原则上支持在 ISO / IEC 11172-2 和 ISO / IEC 11172-3 中定义的视频和音频混合编码方法。系统层有五个基本功能:

- a) 回放过程中多路压缩数据流的同步;
- b) 把多路压缩数据流分解为单一数据流;
- c) 回放开始时缓冲区的初始化;
- d) 连续缓冲区的管理;
- e) 时间的标识。

ISO / IEC 11172 多路比特流结构可以分成两层: 最外层是系统层; 最内层是压缩层。系统层提供在一个系统由使用一个或多个压缩数据流必要的的能力。视频和音频的技术条件定义了音频视频数据的压缩编码层, 其它类型的数据的编码在这部分技术规定中没有定义, 但遵守这部分 2.4 节定义约束的其它类型数据将在系统层得到支持。

1.2 标准参考资料

下列国际标准中包含的条款, 通过本文的引用, 构成本部分 ISO / IEC 11172 条款的一部分。在本标准出版之时, 所标明的版本是有效的。所有标准都可能进行修订, 基于本部分 ISO / IEC 11172 的成员要尽可能应用以下标准的最新版本。IEC 和 ISO 的成员始终保持有当前有效国际标准的记录。

以下系国际标准, 本不应翻译, 但为便于查阅, 给出参考译文, 但不代表标准译文。

ISO / IEC 11172-2: 1993 信息技术——具有 1.5Mbit / s 数据传输率的数字存储媒体运动图象及其伴音的编码 第 2 部分: 视频。

ISO / IEC 11172-3: 1993 信息技术——具有 1.5Mbit / s 数据传输率的数字存储媒体运动图象及其伴音的编码 第 3 部分: 音频

CCIR 601-2 建议 演播室数字电视编码参数。

CCIR 624-4 报告 单色和彩色电视系统特性。

CCIR 648 建议 音频信号的录制。

CCIR 955-2 报告 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 J.17 建议 声音编程电路的预增强。

IEEE 草拟标准 P1180 / D2 1990 8×8 实现离散余弦逆变换的详细说明。

IEC 公告 908 1987 CD 数字音频系统。

2 技术要素

2.1 定义

下列定义适用于 ISO / IEC 11172, 如果仅针对其一部分, 则在方括号内注明。

2.2.1 AC 系数〔视频〕: 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〕: 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〕: 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〕: 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〕: 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〕: 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〕: 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〕: 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比特流的最小部分。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比特流。

2.1.9 音频缓冲器〔音频〕: 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〕: 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〕: 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 Bark〔音频〕: 临界频带速率单位。Bark 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〕 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: 压缩比特流从存储介质输入到解码器的速率。

2.1.15 块压扩〔音频〕: 某一时间内音频信号数字表示的归一化。

2.1.16 块〔视频〕: 一个 8 行 8 列的正交像素块。

2.1.17 边界〔音频〕: 利用强度立体声编码的最低子带。

2.1.18 字节对齐: 在编码比特流中, 如果一比特的位位置 (从比特流的第一位开始) 是

8 的倍数, 则该比特就称为字节对齐的。

2.1.19 字节: 8 个比特的序列。

2.1.20 信道: 存储和传输 ISO / IEC 11172 流的数字介质。

2.1.21 声道 (音频): 立体声信号的左、右信道。

2.1.22 色度 (分量) (视频): 按照 CCIR Rec 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、块或单个象素。

2.1.23 编码音频比特流 (音频): 按照 ISO / IEC 11172-3 中规定的音频信号的编码表示。

2.1.24 编码视频比特流 (视频): 按照本部分 ISO / IEC 11172 的规定, 一幅或多幅图象序列的编码表示。

2.1.25 编码顺序 (视频): 图象存储和解码的顺序。该顺序未必与显示的顺序一样。

2.1.26 编码表示: 以编码形式表示的数据元素。

2.1.27 编码参数 (视频): 表现编码视频比特流特征的、可以由用户定义的参数的集合。

2.1.28 分量 (视频): 组成一幅图的 3 个矩阵 (亮度和两个色度) 之一的矩阵、块或单个象素。

2.1.29 压缩: 缩减表示数据项的比特数。

2.1.30 恒位速率编码视频 (视频): 具有恒定的平均位速率的压缩视频比特流。

2.1.31 恒位速率: 压缩比特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): 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): 符合 ISO / IEC 11172-1 的 2.4.6 中定义的约束的 ISO / IEC 11172 多重混合流。

2.1.34 CRC: 循环冗余码。

2.1.35 临界带速率 (音频): 频率的心理声学函数。对一给定的可闻频率, 它正比于低于该频率的临界带的个数。

2.1.36 临界带 (音频): 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: 编码前或解码后的数据项的表示。

2.1.38 DC 系数 (视频): 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): 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: 一给定的余弦基函数的幅值。

2.1.41 解码流: 压缩比特流的解码重构。

2.1.42 解码器输入缓冲器 (视频): 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): 在视频缓冲检验器中规定的的数据速率, 它被编码进入视频比特流中。

2.1.44 解码器: 解码处理的实体。

2.1.45 解码 (处理): ISO / IEC 11172 中定义的处理, 即读入编码的比特流并产生解码图象或音频样值。

2.1.46 解码时标; DTS〔系统〕: 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〕: 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〕: 比特中的量化 DCT 系数在解码之后和作 DCT 逆变换之前所进行的比例变换处理。

2.1.49 数字存储介质; DSM: 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〕: 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。附录 A 中定义了 DCT 逆变换。

2.1.51 显示顺序〔视频〕: 解码图象进行显示的顺序。通常这与输入编码器编码的顺序相同。

2.1.52 双声道模式〔音频〕: 在一个比特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: 对一个或多个压缩比特流进行操作, 以产生一个新的比特流。合格的编辑比特流必须满足本部分 ISO / IEC 11172 规定的要求。

2.1.54 基本流〔系统〕: 编码视频、音频或其它编码比特流的统称。

2.1.55 加重〔音频〕: 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: 编码处理的实体。

2.1.57 编码(处理): 在 ISO / IEC 11172 中没有规定的一种处理, 读入输入的图象或音频样值流, 以产生一个符合 ISO / IEC 11172 中定义的有效比特流。

2.1.58 熵编码: 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〕: 按显示顺序, 以比真实时间快的速度播放一图象序列或图象序列的一部分。

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〕: 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〕: 分成固定时间段的音频信号数字表示的一段。

2.1.63 禁止: 定义编码比特流的条款中用到的“禁止”表示某值决不能被使用。这常用于避免起始码的仿真。

2.1.64 强迫更新〔视频〕: 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〕: 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〕: 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〕: 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〕: 依显示顺序在当前图之后的参考图。

2.1.69 区组〔层 II〕〔音频〕: 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。它们构成 PCM 样值的 96%。

- 2.1.70 区组〔层Ⅲ〕〔音频〕**: 载有自身辅助信息的 576 频率线。
- 2.1.71 图组〔视频〕**: 支持随机操作的一幅或多幅编码图的序列。
- 2.1.72 Hann 窗口〔音频〕**: 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 Huffman 编码**: 一种熵编码方法。
- 2.1.74 混合滤波器组〔音频〕**: 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〕**: 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〕**: 常规电视图象的一种性质, 其中相间各行在时间上连续。
- 2.1.78 帧内编码〔视频〕**: 宏块或图象仅用其自身的信息进行编码。
- 2.1.79 帧内编码图; I 图〔视频〕**: 仅用其自身信息进行编码的图。
- 2.1.80 ISO/IEC 11172 (多路复用) 流〔系统〕**: 由零个或零个以上的基础流 (按 ISO/IEC 11172-1 的定义组成的) 组成的比特流。
- 2.1.81 联合立体声编码〔音频〕**: 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〕**: 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〕**: ISO/IEC 11172-3 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〕**: ISO/IEC 11172-2 以及 ISO/IEC 11172 的本部分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〕**: 表示单色信号的矩阵、块或单个象素, 符合 CCIR Rec 601 定义的与色差信号相联系。表示亮度的符号是 Y。
- 2.1.86 宏块〔视频〕**: 从图象的亮度分量的一个 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。宏块有时指象素数据, 有时指象素值的编码表示, 以及其它的本部分 ISO/IEC 11172 所定义的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〕**: 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〕**: 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〕**: 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〕**: 利用运动矢量来改善象素值预测的效率。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象素值。
- 2.1.92 运动估值〔视频〕**: 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〕**: 用于运动补偿的二维矢量, 它提供从当前图象的坐标位置到参考图象坐标的偏移。
- 2.1.94 MS 立体声〔音频〕**: 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〕**: 宏块或图象的编码不仅用到其自身的信息而且还用到其它