



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

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	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

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。

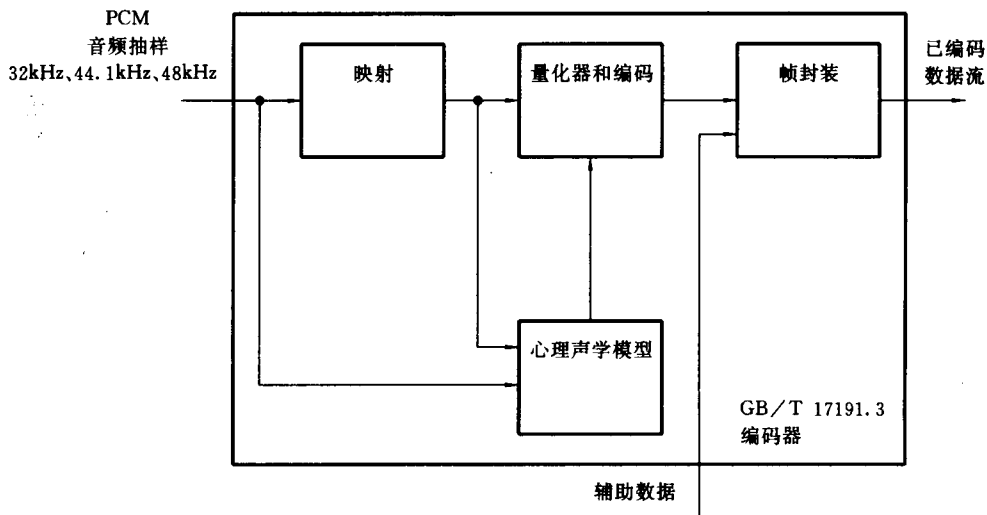


图1 音频编码器基本结构框图

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。

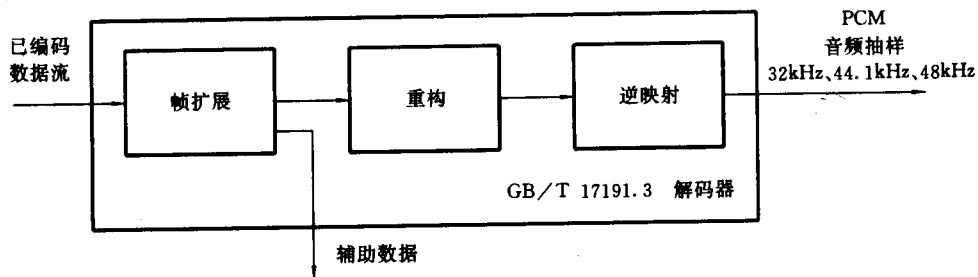


图 2 音频解码器结构框图

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 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 I 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〕