

ISO/IEC 13813

钟玉琢 乔秉新 祁 卫 译

运动图象 及其伴音 通用编码国际标准 — MPEG-2



清华大学出版社

ISO/IEC 13813

运动图象及其伴音 通用编码国际标准——MPEG-2

钟玉琢 乔秉新 祁 卫 译

清华大学出版社

(京)新登字 158 号

内 容 简 介

运动图象及其伴音通用压缩编码标准是多媒体计算机、多媒体数据库、多媒体通信、常规电视数字化、高清晰度电视以及交互式电视系统中的关键技术。

本标准 MPEG-2 是在 MPEG-1 基础上的改进和扩充,目前在家用卫星广播业务、地面数字电视广播、电子影院、电子新闻采集系统、个人通信、多媒体邮件、网络数据库、家庭电视剧场、遥控监视以及点播电视系统等方面的应用方兴未艾。

本书分为系统、视频和音频三部分,全面详尽介绍了 MPEG-2 数据流的数据结构,语法和语义,编码和解码算法,还给出各部分系统框图。本书可作为从事多媒体计算机、多媒体通信、HDTV 和交互式电视技术的研制、开发及应用的工程技术人员以及大专院校有关专业的师生的参考书。

© 清华大学出版社 1997

版权所有,翻印必究。本书任何部分,未经出版者书面许可,不得以任何方式节录和翻印。

本书封面贴有清华大学出版社激光防伪标签,无标签者不得销售。

此版本只限在中国大陆地区发行销售

Original English edition published by ISO/IEC

© ISO/IEC 1994

ISO/IEC Copyright Office·Case Postale 56·CH 1211 Genève 20 Switzerland

All rights reserved. No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the publisher.

北京市版权局著作合同登记号: 01—95—572

图书在版编目(CIP)数据

ISO/IEC 13813: 运动图象及其伴音通用编码国际标准 MPEG-2/钟玉琢等译. —北京: 清华大学出版社, 1996

ISBN 7-302-02407-3

I. I… II. 钟… III. 图象识别-语音数据处理-国际标准 IV. TP391.4-65

中国版本图书馆 CIP 数据核字(96)第 25175 号

出版者: 清华大学出版社(北京清华大学校内, 邮编 100084)

印刷者: 北京市清华园胶印厂

发行者: 新华书店总店北京科技发行所

开 本: 787×1092 1/16 印张: 28.5 字数: 673 千字

版 次: 1997 年 6 月第 1 版 1997 年 6 月第 1 次印刷

书 号: ISBN 7-302-02407-3/TP·1209

印 数: 0001~3000

定 价: 39.00 元

译者序

运动图象及其伴音通用的压缩编码技术是多媒体计算机、多媒体数据库、多媒体通信、常规电视数字化、高清晰度电视(HDTV)以及交互式电视(Interactive TV)系统中的关键技术,目前正在开展的应用项目有:家庭用卫星广播业务(BSS——Broadcasting satellite Service to The Home)、地面数字电视广播(DTTB——Digital Terrestrial Television Broadcast)、电子影院(EC——Electronic Cinema)、电子新闻采集系统(ENG——Electronic News Gathering)、个人通信(IPC——Inter Personal Communication)、多媒体邮件(MMM——Multimedia Mailing)、网络数据库(NDB——Networked Database Service)、家庭电视剧场(HTT——Home Television Theatre)、遥控监视(RVS——Remote Video Surveillance)以及电视点播系统(VOD——Video On Demand)等。

要较好地解决上述应用项目中的传输和存储问题,一定要采用压缩编码技术。从1948年 Oliver 提出 PCM 编码理论开始,迄今已有 40 多年的历史,很多科学家、工程师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。它们的目的是为了从原始数据中去除信息的时间冗余、空间冗余、信息熵冗余、结构冗余、知识冗余以及视觉冗余等,保留有用信息,以便提高信息处理、传输和存储效率。

前几年世界上很多大的国际性公司,为了设计开发多媒体计算机系统以及通信系统,专门研制了自己的视频音频信息压缩解压缩算法,如 Intel 公司 DVI 系统中的 PLV 算法,为了提高速度,投资了大量研究费用,设计制造了专用芯片(Intel 公司 82750 芯片),当时在国际市场上取得了良好的经济效益和社会效益。但是随着最近几年国际标准化组织 ISO(The International Organization for Standardization)和 IEC(International Electrotechnical Commission)先后制定了:ISO/IEC 10918[连续色调静态图象的数字压缩和编码(JPEG 标准)],ISO/IEC 11172(具有 1.5Mbit/s 数据传输率数字存储媒体运动图象及其伴音的编码 MPEG-I)以及 CCITT 制定的 H. 261($P \times 64$ Kbps 视声业务视象编码方式)等国际标准,使得上述不符合国际标准的产品的进一步发展受到限制。

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相关标准化组织的研究成果,如 JPEG 和 H. 261 标准。运动图象可以看成是静止图象的一个序列,所以运动图象的帧内编码技术就采用了 JPEG 推荐的 DCT 技术。此外 MPEG 又加进了帧间压缩编码技术(运动补偿技术)。因此,可以认为 MPEG 的工作是 JPEG 工作的延续。同时,MPEG 推荐的标准尽量与 H. 261 标准兼容。到 1992 年 MPEG 正式推出了 MPEG-I 标准草案,1993 年正式通过。最近几年,由于多媒体技术、数字电视技术、多媒体通信以及交互电视技术的发展,明显感到 MPEG-I 在视频音

频分辨率和传输率方面已不能满足要求,所以 ISO/IEC 在 1994 年又推出了 ISO/IEC 13818 MPEG-2 运动图象及其伴音通用压缩编码标准。本书将全面详细地介绍该标准。全书分为三大部分:系统、视频和音频。每部分都有缩写和符号的定义,描述数据流的数据结构、语法和语义,编码和解码的算法,层次关系以及各部分和系统的框图。此外,各部分还有详细的附录,它给出了需要参考的基础理论知识、算法及专利等。该标准可供多媒体计算机、多媒体通信、HDTV 以及交互式电视技术等有关专业的科技人员和工程技术人员参考。

全书由钟玉琢、乔秉新、祁卫翻译并审阅,祁卫、刘欣、谭晖等参加了全书初稿的翻译工作。由于运动图象及其伴音的压缩编码研究正处在不断发展阶段,译者水平所限再加上时间仓促,书中内容难免存在不足或错误之处,恳请读者给予批评指正。

译 者

1995 年 5 月于北京

目 录

译者序	XI
-----------	----

第 I 部分 系 统

I-0 系统引言	3
I-0.1 传送流	4
I-0.2 程序流	6
I-0.3 传送流和程序流之间的转换	7
I-0.4 分组的原始数据流	7
I-0.5 时间模式	8
I-0.6 条件访问	8
I-0.7 多路复用宽操作	8
I-0.8 单独的流操作(PES 分组层)	9
I-0.8.1 多路分配	9
I-0.8.2 同步	9
I-0.8.3 与压缩层的关系	9
I-0.9 系统参考解码器	10
I-0.10 应用	10
I-1 系统概述	11
I-1.1 范围	11
I-1.2 规格化参考	11
I-1.3 推荐相似的国际标准	11
I-1.4 其它的参考资料	11
I-2 技术要素	13
I-2.1 定义	13
I-2.2 符号与缩写	15
I-2.2.1 算术操作符	15
I-2.2.2 逻辑操作符	16
I-2.2.3 关系操作符	16
I-2.2.4 位操作符	16
I-2.2.5 赋值	16
I-2.2.6 助记符	17
I-2.2.7 常数	17
I-2.3 描述比特流语法的方法	17
I-2.4 传送流比特流要求	19
I-2.4.1 传送流的编码结构和参数	19
I-2.4.2 传送流系统目标解码器(T-STD)	20

I-2.4.3	传送流语法和语义的说明	25
I-2.4.4	程序特殊信息(PSI)	47
I-2.5	程序流比特流要求	56
I-2.5.1	程序流的编码结构和参数	56
I-2.5.2	程序流系统目标解码器	56
I-2.5.3	程序流语法和语义的说明	60
I-2.5.4	程序流映射	64
I-2.5.5	程序流目录	66
I-2.6	程序流和原始流描述子(descriptor)	70
I-2.6.1	流描述子中各字段的语义定义	70
I-2.6.2	视频流描述子	70
I-2.6.3	视频流描述子中各字段的语义定义	71
I-2.6.4	音频流描述子	72
I-2.6.5	音频流描述子中各字段的语义定义	72
I-2.6.6	体系描述子	72
I-2.6.7	体系描述子中各字段的语义定义	73
I-2.6.8	注册描述子	74
I-2.6.9	注册描述子中各字段的语义定义	74
I-2.6.10	数据流对准描述子	74
I-2.6.11	数据流对准描述子中各字段的语义定义	74
I-2.6.12	目标背景栅格描述子	75
I-2.6.13	目标背景栅格描述子中各字段的语义定义	76
I-2.6.14	视频窗口描述子	76
I-2.6.15	视频窗口描述子中各字段的语义定义	76
I-2.6.16	条件访问描述子	77
I-2.6.17	条件访问描述子中各字段的语义定义	77
I-2.6.18	ISO 639 语言描述子	77
I-2.6.19	ISO 639 语言描述子中各字段的语义定义	78
I-2.6.20	系统时钟描述子	78
I-2.6.21	系统时钟描述子中各字段的语义定义	79
I-2.6.22	多路复用缓冲区使用描述子	79
I-2.6.23	多路复用缓冲区使用描述子中各字段的语义定义	80
I-2.6.24	版权描述子	80
I-2.6.25	版权描述子中各字段的语义定义	81
I-2.6.26	最大比特率描述子	81
I-2.6.27	最大比特率描述子中各字段的语义定义	81
I-2.6.28	私用数据指示器描述子	81
I-2.6.29	私用数据指示器描述子中各字段的语义定义	81
I-2.7	多路复用流语义的限制	82
I-2.7.1	缓冲区管理	82
I-2.7.2	系统参考时钟的编码频率	82
I-2.7.3	程序参考时钟的编码频率	82

I-2.7.4	原始流系统参考时钟的编码频率	83
I-2.7.5	显示时间标签的编码频率	83
I-2.7.6	时间标签的条件编码	83
I-2.7.7	PES 分组首部中 P-STD_buffer_size 的编码频率	84
I-2.7.8	程序流中系统首部的编码	84
I-2.7.9	系统参数受限的程序流	84
I-2.7.10	传送流	85
I-2.8	与 ISO/IEC 11172 的兼容性	86
附录 I-A	数字存储媒体的命令和控制[DSM CC]	87
I-A.0	简介	87
I-A.0.1	目的	87
I-A.0.2	未来的应用	87
I-A.0.3	优点	88
I-A.0.4	基本功能	88
I-A.1	基本元素	89
I-A.1.1	范围	89
I-A.1.2	DSM CC 应用的概况	89
I-A.1.3	DSM CC 命令和应答的传送	89
I-A.2	技术元素	90
I-A.2.1	定义	90
I-A.2.2	DSM CC 语法规范	91
I-A.2.3	DSM CC 语法规范中各字段的语义	92
I-A.2.4	控制层	92
I-A.2.5	控制层中各字段的语义	94
I-A.2.6	应答层	95
I-A.2.7	应答层中各字段的语义	95
I-A.2.8	时间码(time code)	96
I-A.2.9	时间码中各字段的语义	97
附录 I-B	CRC 解码器模型	98
I-B.0	CRC 解码器模型	98
附录 I-C	程序特殊信息	99
I-C.0	传送流中程序特殊信息的说明	99
I-C.1	简介	99
I-C.2	功能机制	99
I-C.3	从分段映射到传送流分组	100
I-C.4	重复率和随机访问	101
I-C.5	程序是什么?	101
I-C.6	Program-number 的分配	101
I-C.7	在一典型系统中 PSI 的使用	102
I-C.8	PSI 结构的关系	102
I-C.8.1	程序关联表	103
I-C.8.2	程序映射表	104

I -C. 8.3	条件访问表	104
I -C. 8.4	网络信息表	104
I -C. 8.5	Private-section()	104
I -C. 8.6	描述子	105
I -C. 9	带宽效用和信号获取时间	105
附录 I -D	ITU-T H. 222. 0 ISO/IEC 13818-1 系统时序模型和应用的内在要求	108
I -D. 0	简介	108
I -D. 0.1	时序模型	108
I -D. 0.2	音频和视频显示同步	109
I -D. 0.3	解码器中系统时间时钟的恢复	111
I -D. 0.4	SCR 和 PCR 抖动	113
I -D. 0.5	出现网络抖动时的时钟恢复	115
I -D. 0.6	用于产生彩色 sub_carrier 的系统时钟	115
I -D. 0.7	视频和音频分量的重构	117
I -D. 0.8	帧的掠过	117
I -D. 0.9	网络抖动的平滑	117
附录 I -E	数据传输应用	119
I -E. 0	几点考虑	119
I -E. 1	建议	119
附录 I -F	ITU-T H. 222. 0 ISO/IEC 13818-1 的语法图	120
I -F. 0	简介	120
I -F. 0.1	传送流语法	120
I -F. 0.2	PES 分组	121
I -F. 0.3	程序关联分段	121
I -F. 0.4	CA 分段	121
I -F. 0.5	TS 程序映射分段	122
I -F. 0.6	私用分段	122
I -F. 0.7	程序流	123
I -F. 0.8	程序流映射	123
附录 I -G	普通信息	124
I -G. 0	普通信息	124
I -G. 0.1	同步字节竞争	124
I -G. 0.2	忽略的图象状态和解码处理	124
I -G. 0.3	PID 值的选择	124
I -G. 0.4	PES 开始码字竞争	124
附录 I -H	私用数据	125
I -H. 0	私用数据	125
附录 I -I	对 ITU-T Rec. H. 222. 0 ISO/IEC 13818 已提出专利申请的公司名单	127
I -I. 0	对 ITU-T Rec. H. 222. 0 ISO/IEC 13818-1 已提出专利申请的公司	127
附录 I -J	系统一致和实时接口	129
I -J. 0	系统一致和实时接口	129
附录 I -K	“抖动引入”网络到 MPEG-2 解码器的接口	130

I-K.1	简介	130
I-K.2	网络一致模型	131
I-K.3	抖动平滑的网络说明	131
I-K.4	解码器实现示例	132
I-K.4.1	跟有 MPEG-2 解码器的网络转接器	132
I-K.4.2	集成解码器	132

第 II 部分 视 频

II-0	视频引言	137
II-0.1	目的	137
II-0.2	应用	137
II-0.3	框架和级别	137
II-0.4	分级和不可分级的语法	138
II-0.4.1	不分级语法简介	138
II-0.4.2	分级的扩展	140
II-1	范围	142
II-2	标准参照	143
II-3	定义	144
II-4	缩写和符号	149
II-4.1	算术运算符	149
II-4.2	逻辑算符	149
II-4.3	关系算符	150
II-4.4	位算符	150
II-4.5	赋值	150
II-4.6	助记符	150
II-4.7	常量	150
II-5	约定	151
II-5.1	比特流语法的描述方法	151
II-5.2	函数的定义	152
II-5.2.1	bytealigned()函数的定义	152
II-5.2.2	nextbits()函数的定义	152
II-5.2.3	next_start_code()函数的定义	152
II-5.3	保留、禁止和 marker-bit	153
II-5.4	算术精确性	153
II-6	视频比特流的语法和语义	154
II-6.1	编码视频数据的结构	154
II-6.1.1	视频序列	154
II-6.1.2	组块	159
II-6.1.3	宏块	160
II-6.1.4	块	162
II-6.2	视频比特流语法	162

II -6.2.1	起始码	162
II -6.2.2	视频序列	163
II -6.2.3	图象头	168
II -6.2.4	组块	171
II -6.2.5	宏块	172
II -6.2.6	块	175
II -6.3	视频比特流语义	175
II -6.3.1	更高层语法结构的语义规则	175
II -6.3.2	视频序列	177
II -6.3.3	序列头	177
II -6.3.4	扩展和用户数据	180
II -6.3.5	序列扩展	180
II -6.3.6	序列显示扩展	181
II -6.3.7	序列分级扩展	185
II -6.3.8	图组头	186
II -6.3.9	图象头	187
II -6.3.10	图象编码扩展	189
II -6.3.11	量化矩阵扩展	192
II -6.3.12	图象显示扩展	193
II -6.3.13	图象时域分级扩展	194
II -6.3.14	图象空域分级扩展	195
II -6.3.15	组块	195
II -6.3.16	宏块	196
II -6.3.17	块	200
II -7	视频解码过程	201
II -7.1	高层语法结构	201
II -7.2	变长码解码	202
II -7.2.1	内部块中的 DC 系数	202
II -7.2.2	其它系数	203
II -7.3	反扫描	205
II -7.3.1	加载矩阵的反扫描	206
II -7.4	反量化	206
II -7.4.1	内部 DC 系数	207
II -7.4.2	其它系数	207
II -7.4.3	饱和化	209
II -7.4.4	解谐控制	210
II -7.4.5	小结	210
II -7.5	反 DCT	211
II -7.5.1	非编码的块和跳过的宏块	211
II -7.6	运动补偿	211
II -7.6.1	预测方式	212
II -7.6.2	预测场和帧的选择	213

II-7.6.3	运动向量	214
II-7.6.4	形成预测	221
II-7.6.5	运动向量选择	222
II-7.6.6	跳过的宏块	224
II-7.6.7	组合预测	225
II-7.6.8	预测和系数数据求和	225
II-7.7	空域分级	226
II-7.7.1	更高层语法结构	227
II-7.7.2	增强层中的预测	227
II-7.7.3	空域预测的构成	227
II-7.7.4	空域和时域预测的选择和组合	232
II-7.7.5	修改运动向量预测器和运动向量选择	233
II-7.7.6	跳过的宏块	238
II-7.7.7	低层中的 VBV 缓冲区下溢	238
II-7.8	SNR 分级	238
II-7.8.1	更高的语法结构	239
II-7.8.2	宏块	241
II-7.8.3	块	242
II-7.9	时域分级	243
II-7.9.1	更高的语法结构	245
II-7.9.2	对于时域预测的限定	246
II-7.10	数据分割	246
II-7.11	混合分级	248
II-7.12	解码过程输出	249
II-8	框架和级别	252
II-8.1	ISO/IEC 11172-2 的兼容性	253
II-8.2	限定的框架间的关系	254
II-8.3	限定的级别间的关系	255
II-8.4	分级的层	256
II-8.4.1	允许的层组合	257
II-8.5	限定的框架、级别和层的参数值	258
附录 II-A	离散余弦变换	263
附录 II-B	可变长度码表	264
II-B.1	宏块寻址	264
II-B.2	宏块类型	265
II-B.3	宏块模式	269
II-B.4	运动向量	270
II-B.5	DCT 系数	270
附录 II-C	视频缓冲检验器	277
附录 II-D	算法所支持的性能	282
II-D.1	概述	282
II-D.2	视频格式	282

I -D. 2. 1	采样格式和颜色	282
I -D. 2. 2	影片定时	283
I -D. 2. 3	显示格式控制	283
I -D. 2. 4	全电视信号的透明编码	283
I -D. 3	图象质量	283
I -D. 4	数据率控制	284
I -D. 5	低延迟模式	284
I -D. 6	随机访问/信道跳换	284
I -D. 7	分级	285
I -D. 7. 1	SNR 分级在单一空域分辨率中的应用	285
I -D. 7. 2	使用 SNR 分级的多分辨率分级比特流	286
I -D. 7. 3	在数据分割中的比特率分配	287
I -D. 7. 4	时域分级	287
I -D. 7. 5	空域、SNR 和时域分级扩展的混合	289
I -D. 8	兼容性	291
I -D. 8. 1	较高和较低的分辨率格式间的兼容	291
I -D. 8. 2	和 ISO/IEC 11172-2(以及 ITU-T Rec. H. 261)的兼容	291
I -D. 9	本规范与 ISO/IEC 11172-2 的差异	292
I -D. 9. 1	IDCT 解谐	292
I -D. 9. 2	宏块填充	292
I -D. 9. 3	Run-level 换码语法	292
I -D. 9. 4	色差样本水平位置	293
I -D. 9. 5	组块	293
I -D. 9. 6	D-图	293
I -D. 9. 7	Full-pel 运动向量	293
I -D. 9. 8	高宽比信息	293
I -D. 9. 9	forward_f_code 和 backward_f_code	293
I -D. 9. 10	constrained-parameter-flag 和最大的 horizontal-size	294
I -D. 9. 11	MPEG-2 语法转换为 MPEG-1 语法	294
I -D. 10	复杂性	294
I -D. 11	编辑编码的比特流	295
I -D. 12	特技方式	295
I -D. 12. 1	解码器	295
I -D. 12. 2	编码器	296
I -D. 13	错误复原	296
I -D. 13. 1	隐藏的可能性	298
I -D. 13. 2	空间定位	301
I -D. 13. 3	时间定位	303
I -D. 13. 4	小结	303
附录 II -E	框架和级别的限定	305
I -E. 1	框架中的语法元素限定	305
I -E. 2	允许的层组合(见 I -8. 4. 1)	316

第Ⅲ部分 音频

Ⅲ-0 引言	323
Ⅲ-0.1 ISO/IEC 11172-3 降低音频编码采样率的扩充	323
Ⅲ-0.2 多通道音频的低比特率编码	323
Ⅲ-0.2.1 通用多通道音频系统	323
Ⅲ-0.2.2 多通道音频的表示	324
Ⅲ-0.2.3 多通道音频编码系统的基本参数	325
Ⅲ-1 音频概述	329
Ⅲ-1.1 范围	329
Ⅲ-1.2 标准参考	329
Ⅲ-1.2.1 相同的推荐 国际标准	329
Ⅲ-1.2.2 两个推荐 国际标准在技术内容上的等价性	329
Ⅲ-1.2.3 附加参考文献	329
Ⅲ-2 技术单元	330
Ⅲ-2.1 定义	330
Ⅲ-2.2 符号和缩写	336
Ⅲ-2.2.1 算术运算符	337
Ⅲ-2.2.2 逻辑运算符	337
Ⅲ-2.2.3 关系运算符	337
Ⅲ-2.2.4 位操作符	338
Ⅲ-2.2.5 赋值	338
Ⅲ-2.2.6 助记符	338
Ⅲ-2.2.7 常量	339
Ⅲ-2.3 描述比特流语法的方法	339
Ⅲ-2.4 ISO/IEC 11172-3 降低采样频率音频编码扩展的要求	341
Ⅲ-2.4.1 编码音频比特流语法的说明	341
Ⅲ-2.4.2 音频比特流语法的语义	343
Ⅲ-2.4.3 音频解码过程	345
Ⅲ-2.5 多通道低比特率音频编码的要求	348
Ⅲ-2.5.1 编码音频比特流语法说明	348
Ⅲ-2.5.2 音频比特流语法的语义	361
Ⅲ-2.5.3 音频解码过程	379
附录Ⅲ-A 图	393
附录Ⅲ-B 表	394
附录Ⅲ-C 编码过程	400
Ⅲ-C.1 低采样频率扩展	400
Ⅲ-C.1.1 低采样频率,层Ⅰ	400
Ⅲ-C.1.2 低采样频率,层Ⅱ	400
Ⅲ-C.1.3 低采样频率,层Ⅲ	401
Ⅲ-C.2 多通道扩展	401

■ -C. 2. 1 多通道扩展,层 I, II	401
■ -C. 2. 2 多通道扩展,层 III	409
附录 III-D 听觉心理模型	413
■ -D. 1 用于低采样频率的听觉心理模型 1	413
■ -D. 2 临界波段边界的表	425
■ -D. 3 针对较低采样频率的听觉心理模型 2	428
■ -D. 4 将國值计算分区转化为比例因子波段的表	435
附录 III-E 专利所有者清单	439

第 I 部分 系 统

