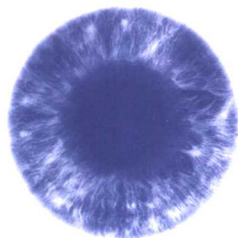


[英] Iain E.G. Richardson 著

欧阳合 韩军 译

H.264 和 MPEG-4



视频压缩

新一代多媒体的视频编码技术

国防科技大学出版社

H.264 和 MPEG-4 视频压缩

——新一代多媒体的视频编码技术

[英] Iain E. G. Richardson 著

欧阳合 韩 军 译

国防科技大学出版社

湖南·长沙

编著书名: H.264 and MPEG-4 Video Compression — Video Coding for Next-generation Multimedia

原著简体中文版由 John Wiley & Sons Ltd 授权国防科技大学出版社翻译并出版。

申明: 未经许可, 本出版品内容的任何部分不得翻版, 不得存于检索系统或以电子、机械、影印、录音、扫描等任何形式进行传播。

版权所有: John Wiley & Sons Ltd, England

图书在版编目(CIP)数据

H.264 和 MPEG-4 视频压缩: 新一代多媒体的视频编码技术/[英]理查森(Richardson, I. E. G.) 著; 欧阳合, 韩军译. —长沙: 国防科技大学出版社, 2004.11

ISBN 7-81099-140-X

I. H… II. ①理… ②欧… ③韩… III. 视频信号-数据压缩
IV. TN919.81

中国版本图书馆 CIP 数据核字(2004)第 107515 号

著作权合同登记号 图字: 军-2004-031 号

国防科技大学出版社出版发行
电话: (0731)4572640 邮政编码: 410073
E-mail: gfkdcbs@public.cs.hn.cn
责任编辑: 唐卫藏 责任校对: 耿 筠
新华书店总店北京发行所经销
国防科技大学印刷厂印装

*

开本: 787×1092 1/16 印张: 20.25 字数: 386 千
2004 年 11 月第 1 版第 1 次印刷 印数: 1-3 000 册

ISBN 7-81099-140-X/TN·13

定价: 39.00 元

译者的话

当前数字视频技术发展迅速。近年来，国际上相继推出了 MPEG-4 及 H.264 视频压缩标准，它对未来的数字电视、无线视频等产业的发展起到了巨大的推动作用。为了国内的广大科研工作者及国内大专院校师生对当前的数字视频技术有一个比较全面、系统的了解，上海杰得微电子有限公司总裁欧阳合博士组织公司的科技人员对《H.264 and MPEG-4 VIDEO COMPRESSION》一书进行了翻译。

由欧阳合博士和韩军博士对全书进行总的翻译与校对，具体翻译者：第 1、6 章史翠竹；第 2 章戴一村；第 3 章杨眉；第 4 章罗霖；第 5 章唐谦、罗霖；第 6 章韩军；第 7 章肖晟；第 8 章唐谦。校对者：第 1、2 章林晓芸；第 3、4 章伍宽；第 5、6 章万凯；第 6 章史翠竹；第 7、8 章王璐。

公司副总裁邹艳为本书的翻译在人员的安排、与编辑部联系等方面做了大量的事务性工作。

对本书翻译的不足之处，望广大读者及时指正。

关于作者

Iain Richardson 是苏格兰阿伯丁的 Robert Gordon 大学的讲师和研究者。他于 1990 年和 1999 年分别在 Heiort-Watt 大学和 Robert Gordon 大学获得了硕士和博士学位。自 1993 年以来,他积极从事视频压缩系统的研究工作,共发表了 40 多篇期刊、会议文章和两本专著。在他领导下的 Robert Gordon 大学的图像通信技术研究小组,是多家公司的视频压缩技术的指导专家。

前言

过去三年，视频编码标准组一直致力于“先进视频编码”标准的制定。“先进视频编码”标准也称为 ITU-T H. 264 和 ISO/IEC 14496 (MPEG-4) 的第 10 部分。这项工作使我们深入地涉及到视频压缩的设计。现在我们要看看在这方面已经完成了哪些工作。

Richardson 博士虽然不是直接的参与者，但他对新标准给出了高水准的、最新的描述和分析。本书是非常适时的，因为标准才刚刚完成。

新的 H. 264/AVC 标准试图在相对广泛领域中进行应用，并提供了合适的解决方案，包括如下方面：

- 通过电缆，卫星，电缆调制解调器，DSL 和陆地的广播；
- 光介质和磁介质存储设备上的交互或顺序存储，如 DVD；

- 通过 ISDN, Ethernet, LAN, DSL, 无线和移动网络和调制解调器的会话服务;

- 视频点播或电缆调制解调器, DSL, ISDN, LAN 和无线网络上的多媒体流服务;

- DSL, ISDN 上的多媒体信息服务。

H. 264/AVC 支持的比特率范围和图像大小也非常广泛, 适于从手机和电话拨号的极低码率、低帧率、像邮票大小的图像分辨率的应用, 直到具有娱乐功能的标准清晰电视和高度清晰电视服务。标准为编码视频提供了灵活的系统接口, 使得视频内容的使用几乎适合于所有的网络和信道环境。同时, 标准在技术方面也着重于两个主要目的: 一个是高编码效率; 另一个是网络环境下对于传统的矩形图像和摄像机镜头视频内容的编码健壮性要求。其他一些有趣的特性(当前不是主流的)被省去了(至少在标准的第一版本中不涉及), 比如支持任意形状视频对象的编码、比特率可分级编码、4:2:2 和 4:4:4 的色度格式以及每个色度分量超过 8 个比特的色彩采样精确度等。

在新的 H. 264/AVC 标准的工作中, 采用了许多新的技术。为了提高编码效率, 包括改进的运动预测设计在内的技术有:

- 用较小的块做可变大小的运动补偿;
- 四分之一像素精度的运动补偿;
- 图像边界外的运动矢量估算;
- 多参考帧的运动补偿;
- 显示顺序和参考顺序的去耦合;
- 图像表示方法不依赖所使用一个参考图像;
- 加权预测;
- 改进的“跳转”和“直接”运动参考;
- 基于方向性空间预测的帧内编码;
- 环内去块效应滤波器。

除了改进的预测方法外, 在其他方面的设计也改善了编码效率, 包括:

- 小尺寸块的变换;
- 块的分级变换;
- 短字长变换;
- 精确匹配的变换;

- 算术熵编码;
- 基于上下文的熵编码。

为了增强健壮性和不同网络环境下操作的灵活性, 标准中包含一些关键的设计:

- 参数集结构;
- NAL (网络提取层) 单元的语法结构;
- 灵活的条带大小;
- 灵活的宏块顺序;
- 任意的条带顺序;
- 编码图像的冗余处理;
- 数据分流;
- 基于图像转换的 SP/SI 同步。

在 H. 264/AVC 标准之前, 国际标准化组织公布了 MPEG-4 Part 2 (视频) 编码标准。这个标准有了很大程度的创新, 以创造性和灵活性提供了表示数字视频内容的能力, 特别是对视频“对象”的编码、可分级特性、扩展的 N-bit 采样精度和 4:4:4 的色度格式, 而且它支持人工合成视频场景的编码。它为广泛的应用提供了很多不同需求的设计 (称为档次 “profile”, 目前总共有 19 个)。H. 264/AVC 标准 (只有 3 个档次) 回归到范围相对较窄的传统应用上, 重点在于由普通相机所拍摄的矩形视频图像的高效压缩且能在网络上可靠传输的视频流。H. 264/AVC 并不试图涵盖 MPEG-4 视频标准的所有领域。MPEG-4 视频标准虽然不是最近发布的标准, 但仍然是近年来视频编码技术发展的里程碑, 它的编码性能仍然有待开发。

许多人知道, 一个标准首先是由媒体发布的, 而不是我们通过阅读、研究标准而了解的。我个人的观点是, 如果你想要了解一个标准, 你就应该得到一份复印件, 去阅读它, 去参照它, 在内容、界限和性能上把它作为最高的权威。没有任何指南或综述能够提供比通过仔细分析标准本身而获得更多的知识。

同时, 标准的规范文档 (至少对视频编码来说) 也不能完全取代在这一领域的一本好的技术书籍。标准规范的编写主要保证精确、一致、完整和正确, 它的可读性不强。标准倾向于丢弃那些对遵循标准不是绝对必要的信息, 例如, 很多人可能会吃惊地发现, 视频编码标准几乎不提编码器如何工作, 或者

编码器如何设计。其实，标准允许用任何方法来设计编码器，只要它能产生被正确解码的比特流就可以了，而不管解码出来的图像质量如何。然而，如果从编码器的角度考虑，人们通常只能了解编码的原理，而几乎所有的书（包括这本书）都是从编码器的角度来了解视频编码的。一本好书，比如这本书，将会告诉你为什么要这样设计，怎样应用这样的设计。然而，一个好标准只会告诉你它只能产生这样的码流，并不做任何解释。

就 H. 264/AVC 标准或 MPEG-4 视频标准而言，我强烈地推荐初学者阅读一些介绍性的概述文章，比如本书，甚至去阅读诸如 H. 261 或 MPEG-1 这些比较旧、比较简单的标准文档，并试着理解它们。数字视频编解码器的设计并不太复杂，而且这些年也没有太大的改变——但是这些基本的原理已经被包裹在一层又一层的技术特性里了，核心部分的简单原理反而变得模糊了。整个 H. 261 规范只有 25 页，其中 17 页的内容描述了以后的各种视频编码标准的核心技术。相比之下，H. 264/AVC 和 MPEG-4 视频标准分别超过了 250 和 500 页，含有大量的技术细节（虽然完全去掉了一些主要内容，例如，如何根据格式编码视频等）。它们包含的内容对于专家来说也是难以完全理解的。

Richardson 博士的这本书并没有面面俱到，然而，他的方法却是包含了丰富的信息，对于初学者理解关键的概念是大有裨益的，他的方法是从概念上优于其他视频编码文献（在某些方面也比较客观），这样的观念和出版的实时性使本书成为了此领域有重要贡献的技术性文献。

Gray J. Sullivan

Gray J. Sullivan 博士简介

Gray J. Sullivan 博士是联合视频小组 (JVT) 的主席, JVT 开发了 H. 264/AVC 这个最新的视频编码的国际标准。ITU-T 视频编码专家组 (VCEG) 和 ISO/IEC 运动图像专家组 (MPEG) 共同负责的这个项目, 并得到最终完成。

Gray J. Sullivan 博士也是 ITU-T 的先进视频编码大会报告起草人, 七年来他一直领导着 VCEG (ITU-T Q. 6/SG16); 他也是 ITU-T 和 MPEG 视频的联络代表, 从 2001 年 3 月到 2002 年 5 月出任 MPEG (ISO/IEC JTC1/SC29/WG11) 视频的主席。

目前 Gray J. Sullivan 博士是 Microsoft 公司的 eHome A/V 平台组的视频标准和技术的项目经理。在微软, 他设计了 Microsoft Windows 操作系统的 Direct X 视频加速的 API/DDI 及其扩展接口。

序

随着数字电视、网络视频流和 DVD 技术的广泛应用，视频压缩已经成为广播和娱乐媒体的最基本的组成部分。数字电视和 DVD 的成功是建立在十多年前的 MPEG-2 标准上的，这个标准曾经被证明是很实用的，但是现在看来，已经有点过时了。用更高效、更实用的、并能很好地利用更强的处理能力的技术取代 MPEG-2 的趋势已经很明显了。最近，对于哪种技术将取代 MPEG-2 的争论还在不断持续，其中，比较看好的是国际标准 MPEG-4 和 H. 264。

本书的目的在于对这两种技术提供一个清晰和客观的不含偏见的介绍，使开发者、工程师、学者和学生能更好地理解和有效地运用这些技术。视频和图像压缩是一个复杂而广泛的课题，本书讨论的内容将仅限于对标准本身（对于 MPEG-4 视频，只针对其支持的“自然实物”视频进行编码的工具）

以及直接体现标准特点的视频编码概念。这本书的内容是基于实际的应用展开的，并特别注重对实际应用有用的那些工具与特点，希望能对开发者或这些标准的使用者提供一些有益的帮助。

感谢很多朋友对构思本书内容框架所提供的帮助。我收到了很多朋友对我写的《视频编解码器设计 (Video Codec Design)》提出的很有价值的意见和建议。尤其感谢 Gary Sullivan 为本书提供的翔实而有益的注解、修正和建议，并给本书作序。特别感谢 Harvay Hanna (Impact Labs Inc), Yafan Zhao (The Robert Gordon University) 和 Aitor Garay 阅读了本书的部分章节，并给予建议；感谢联合视频组 (Joint Video Team) 提供了很多 H.264 标准的细节；感谢 John Wiley & Sons 公司的编辑人员 (尤其是永远给我耐心的支持的 Kathryn Sharples)；感谢 Phyllis 长期以来的支持；最后感谢 Freya 和 Hugh 对许诺已久的 Storybook Glen 之旅的耐心等待。

我衷心希望您觉得这本书有趣易读，尤其是有用。在我的主页 (<http://www.vcodex.com/>) 上可以找到更多的资料和链接。在此，感谢读者的反馈、建议和意见。我的联系方式也可以从我的主页上获得。

Iain Richardson

术 语 表

4:2:0(采样)	采样方法是色度分量的水平分辨率和垂直分辨率,是亮度分量的一半
4:2:2(采样)	采样方法是色度分量的水平分辨率,是亮度分量的一半
4:4:4(采样)	采样方法是色度分量的水平分辨率与垂直分辨率与亮度分量相同
算术编码	降低冗余度的编码方法
人工处理的图像	图像的视觉失真
ASO	任意条带顺序,条带的编码可以不是光栅扫描顺序
BAB	二值 Alpha 块,用于标示区域的边界(MPEG-4 视频)
BAP	身体动画参数
块	作变换的宏块区域(8×8 或 4×4)
块匹配	矩形图像区域上的运动估计
块效应	图像的方块或矩形失真区域
B 帧(条带)	双向运动补偿预测的编码图像(条带)
CABAC	基于上下文的自适应二值算术编码
CAE	基于上下文的算术编码
CAVLC	基于上下文的自适应变长编码
色度	色彩差异分量
CIF	通用中间格式,一种彩色图像的格式
CODEC	编解码器
色彩空间	表示彩色图像的方法
DCT	离散余弦变换
直接预测	一种编码模式,此模式中不传输运动矢量
DPCM	差分脉冲编码调制
DSCQS	双激励连续质量等级,主观质量衡量的等级和方法
DWT	离散小波变换
熵编码	降低冗余的编码方法

错误隐藏	编码图像的后处理,以消除或减小可见的错误效应
Exp-Golomb	指数哥伦布变长码
FAP	脸部动画参数
FBA	脸部和身体动画
FGS	精细颗粒可分级
场	隔行视频序列的偶数行或奇数行
流程图	变换算法的图示表示(或者算法本身)
FMO	灵活的宏块顺序,宏块的编码顺序可以不是光栅扫描顺序
完全搜索	一种运动估计的算法
GMC	全局运动补偿,用于整个编码对象的运动补偿(MPEG-4 视频)
GOP	图像组,编码视频图像的集合
H.261	一种视频编码标准
H.263	一种视频编码标准
H.264	一种视频编码标准
HDTV	高清晰度电视
哈夫曼编码	降低冗余的编码方法
HVS	人体视觉系统,用来观察和解释视频图像的系统
混合(CODEC)	以运动补偿和变换为特征的编解码器
IEC	国际电工委员会,一个标准体
帧间(编码)	使用时间预测或补偿的视频帧编码
隔行扫描(视频)	表示为一系列场的视频数据
帧内(编码)	不使用时间预测的视频帧编码
I 帧(条带)	编码不需要参考其他帧的图像(或条带)
ISO	国际标准化组织,一个标准体
ITU	国际电信联盟,一个标准体
JPEG	联合图像专家组,ISO 的委员会(也是一个图像编码标准)
JPEG2000	一种图像编码标准
延迟	通过通信系统的延迟
级别	符合一定特性的参数集合(适用于档次)
环内滤波器	编码或解码反馈环路里的空间滤波器
宏块	帧编码的单元(通常是 16×16 个像素的区域)
宏块分块	有自己的运动矢量的宏块子区域
宏块子分块	有自己的运动矢量的宏块子区域
媒体处理器	具有媒体编码和处理特性的处理器

术语表

运动补偿	视频帧预测的运动模型
运动估计	两个或多个视频帧之间的相对运动估计
运动矢量	用于运动补偿的指示位移块或区域的矢量
MPEG	运动图像专家组,ISO/IEC 的委员会
MPEG-1	一种多媒体编码标准
MPEG-2	一种多媒体编码标准
MPEG-4	一种多媒体编码标准
NAL	网络提取层
主观质量	衡量算法的视频图像视觉质量
OBMC	重叠块运动补偿
图像(编码)	编码的(压缩的)视频帧
P 帧(条带)	使用参考帧作运动补偿预测的编码图像(条带)
档次	视频编解码器的工具集
逐行扫描(视频)	由完整帧的序列来表示视频数据
PSNR	峰值信噪比,一种客观质量的量度方法
QCIF	四分之一的通用中间格式
量化	降低标量或矢量的精度
码率控制	编码视频信号的比特率的控制
码率失真	编解码器性能的量度(一定范围的编码比特率的失真)
RBSP	原始字节序列载荷
RGB	红/绿/蓝色彩空间
环状效应	解码图像中尖锐边缘周围的“波纹”状的痕迹
RTP	实时传输协议
RVLC	可逆变长码
分级编码	把信号编码成若干层
SI 条带	用于在编码的比特流之间切换的帧内条带
条带	编码图像的一个区域
SNHC	合成的图像,自然的图像混合编码
SP 条带	用于在编码的比特流之间切换的帧间条带
Sprite	可以包含在一系列解码帧中的纹理区域(MPEG-4 视频)
统计冗余	由数据的统计分布引起的冗余
演播室质量	无损或接近无损的视频质量
主观质量	由人眼感觉到的视频质量
主观冗余	由主观上不重要的数据带来的冗余

半像素(运动补偿)	整数像素位置插值形成的参考区域的运动补偿预测
测试模型	描述视频编码标准的参考实现的软件模型和文档
纹理	图像或残差数据
树结构运动补偿	以分块尺寸的灵活可变为特征的运动补偿(H.264)
TSS	三步搜索法,一种运动估计算法
VCEG	视频编码专家组,ITU 的委员会
VCL	视频编码层
视频包	合适打包的编码单元
VLC	变长码
VLD	变长解码器
VLE	变长编码器
VLSI	超大规模集成电路
VO	视频对象
VOP	视频对象面
VQEG	视频质量专家组
加权预测	来自两个参考区域的预测样点做尺度变换的运动补偿
YCbCr	亮度,蓝色度,红色度的色彩空间
YUV	一个色彩空间(见 YCbCr)

目 录

关于作者	(I)
前言	(III)
Gray J. Sullivan 博士简介	(VII)
序	(IX)
术语表	(XI)

第 1 章 概 述

1.1 场 景	(1)
1.2 视频压缩	(3)
1.3 MPEG-4 与 H.264	(6)
1.4 本书结构	(7)
1.5 参考文献	(8)

第 2 章 视频格式和图像质量

2.1 概 要	(9)
2.2 自然视频场景	(9)
2.3 视频捕获	(10)
2.3.1 空间采样	(10)
2.3.2 时序采样	(13)
2.3.3 帧和场	(13)
2.4 色彩空间	(14)
2.4.1 RGB 色度空间	(14)
2.4.2 YCbCr 色度空间	(15)
2.4.3 YCbCr 的采样格式	(17)
2.5 视频格式	(19)
2.6 图像质量	(21)
2.6.1 主观视频质量评估	(22)