



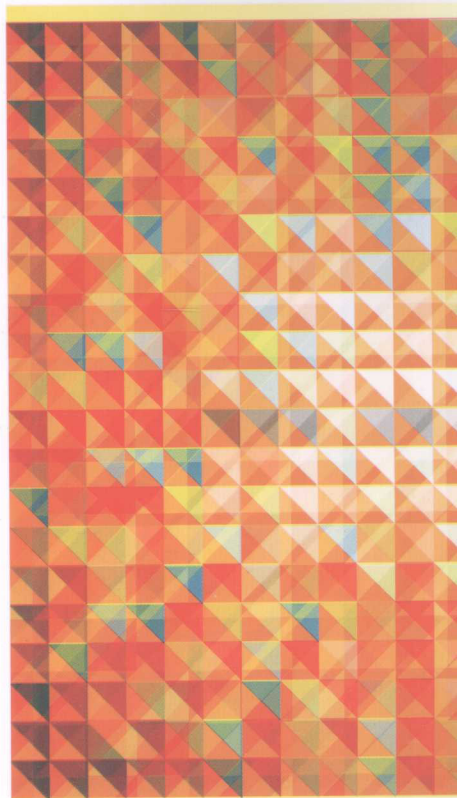
高端图像与视频新技术丛书

H.265/HEVC

视频编码

新标准及其扩展

▶ 朱秀昌 刘峰 胡栋 编著



在给出最新视频编码国际标准H.265/HEVC及其扩展部分全貌的基础上，着重介绍相关的视频编码原理、高效编码工具、高层语法语义和主要技术规范。



中国工信出版集团



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

高端图像与视频新技术丛书

H.265/HEVC——视频编码 新标准及其扩展

朱秀昌 刘 峰 胡 栋 编著

電子工業出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书从视频编码国际标准,特别是 HEVC 标准所涵盖的技术内容出发,主要介绍了新一代视频编码标准所采用的技术规范、基本工作原理、核心编码技术和新的编码工具。同时还介绍了 HEVC 标准新近扩展部分的主要内容,包括扩展的技术、工具和指标。

全书共 13 章,第 1、2 章分别简要介绍了视频编码的基础知识和历年来的国际标准;第 3~10 章,分 8 个部分按顺序介绍了 HEVC 的核心技术内容,包括编码结构、帧内预测、帧间预测、变换和量化、熵编码、环路编码、并行处理和高层语法;第 11、12 章介绍了 HEVC 标准的最新扩展内容,即多层编码、可分级编码、多视点编码和 3D 编码;第 13 章从应用的角度简要介绍了 HEVC 的实现复杂度和软硬件实现的考虑。

本书可作为从事信号处理、通信工程、计算机应用、广播电视、机器人视觉及自动控制等领域的工程技术人员、大专院校相关专业的教师、高年级学生或研究生学习、理解和掌握视频编码最新国际标准和关键编码技术的参考用书。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有,侵权必究。

图书在版编目(CIP)数据

H.265/HEVC: 视频编码新标准及其扩展 / 朱秀昌, 刘峰, 胡栋编著. —北京: 电子工业出版社, 2016.7
ISBN 978-7-121-29038-1

I. ①H… II. ①朱… ②刘… ③胡… III. ①视频编码—研究 IV. ①TN762

中国版本图书馆 CIP 数据核字(2016)第 128704 号

责任编辑: 赵 娜

印 刷: 三河市华成印务有限公司

装 订: 三河市华成印务有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本: 787×1 092 1/16 印张: 21.75 字数: 550 千字

版 次: 2016 年 7 月第 1 版

印 次: 2016 年 7 月第 1 次印刷

定 价: 59.00 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010) 88254888, 88258888。

质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式:(010) 88254694。

前言



伴随着大数据的洪流和无处不在的移动通信发展，以及互联网、物联网的全面覆盖，一个以大数据、大流通为特征的信息时代已经开启。由于视频信息具有真实、具体、生动、全面、信息量大等特点，人们对各种视频信息的需求与日俱增，使得大数据中海量视频信息的占比已过半，因此，保证海量视频信息的高效、优质、标准化的压缩日显重要。

目前视频压缩编码技术面临着三方面的压力：一是由于视频分辨率迅速提高而引起的数据量激增，如第一代视频编码标准面对的基本上是 352×288 的 CIF 视频，第二代视频编码标准面对的主要是 768×576 的 SDTV 视频，如今第三代视频编码标准面对最低的也是 1920×1080 的 HDTV 视频，需压缩视频的数据量大约是以每代 4 倍的速度向上翻，而视频编码效率，每代大致翻一番，这里还未考虑超高清视频、3D 视频、多视点视频等数据量更大的情况。二是由于视频应用的迅速普及而引起的视频信息需求激增，现在各行各业都需要视频信息，几乎每个人都需要视频信息，移动视频还带来每时每刻和无处不在的视频需求，这些需求使得视频采集、传输、存储的数量不断增长。三是视频压缩标准化带来的压力，各方面的视频信息需要相互交流，需要有标准化的格式来保证各种类型的视频信息的畅通，这就要求制定统一的国际标准，并随着技术的发展而不断修订、扩展和换代。

尽管半导体芯片的密度不断提高，网络带宽也不断增加，可以减轻一部分对视频编码的压力，但是远远消解不了上述三方面给视频编码带来的压力。虽然目前已有一系列的国际视频编码标准，但其效率仍然远不能满足当前视频信息的应用需求。诚然，这些都是标准制定者面临的巨大压力，同时，这也理所当然地成为从事视频编码的工程技术人员不断研究新技术、制定新标准的强大驱动力。

为达到对视频数据的高效、标准化压缩的目的,从 20 世纪 80 年代末以来,国际电联 (ITU) 和国际标准化组织 (ISO) 先是分别、后是联手矢志不渝地进行视频编码国际标准的制定工作。迄今为止,已历经三代。第一代以 ITU-T 的 H.261、H.263 建议和 ISO/IEC 的 MPEG-1、MPEG-2 标准为代表;第二代则是以两大组织联合制定的 H.264/ MPEG-4 AVC 标准为代表;第三代也是以两大组织联合制定的并正在扩展中的 HEVC 标准为代表。

HEVC 即“高效视频编码”是该标准的俗称或统称,在 ITU-T 称其为 H.265 建议,在 ISO/IEC 称其为 23008-2 标准。虽然名称各异,但其中的内容是一样的,现已更新为包括扩展部分的第 3 版。本书主要介绍新一代视频编码国际标准 HEVC 及其扩展中的主要压缩编码原理、高效编码工具、高层语法语义和主要技术规范。本书的要点和特点如下。

扼要介绍了和国际标准相关的视频编码原理,梳理了近 30 年来一系列视频编码的国际标准,简述了这些标准体系及发展历程演进。着重分析了 HEVC 第 1 版的关键编码工具,包括编码结构、帧内预测、帧间预测、变换和量化、熵编码、环路滤波、高层语法、并行处理等。在 HEVC 实现方面也给出了简单的说明和几个实例。针对 HEVC 第 2 版的扩展,介绍了在统一的多层编码概念、可分级视频编码、多视点和 3D 视频编码方面的核心编码技术的要点。HEVC 在总体上继承了 H.264/AVC,但在各个方面都有新发展,使得总编码效率提高了一倍,在适当的章节与 H.264/AVC 标准进行了简单的对比,有助于加深对 HEVC 的理解。

本书以从事计算机技术、通信技术以及电视技术的工程技术人员、大专院校有关专业的高年级学生、研究生或教师为主要阅读对象。全书的内容可以划分为四个部分。

第一部分(第 1、2 章)。第 1 章介绍和视频编码有关的基本概念和相关技术,如视频的数字化,数字视频信号的统计特性,混合编码的框架和主要方法,图像质量的评价等。在此基础上,第 2 章介绍了 ISO 和 ITU-T 两大国际标准化组织的 H.261/263 和 MPEG-1/2 第一代标准、H.264/MPEG-4 AVC 第二代标准,以及当前 HEVC 第三代视频编码国际标准的总体概况和标准化进程。

第二部分(第 3~10 章)。主要介绍了 HEVC 第 1 版(2013 年 3 月)中重要编码工具的技术原理和技术要求。包括:第 3 章的视频编码结构,重点是四叉树图像的划分技术;第 4 章的帧内预测编码,重点是高达 35 种方向的帧内预测技术;第 5 章的帧间预测编码,重点是运动信息的高效 Merge 模式的生成技术;第 6 章的变换和量化,重点是多尺寸 DCT、小尺寸 DST 和改进的量化、伸缩方式;第 7 章的熵编码,重点是高效的基于上下文的二进算术编码,尤其是它的常规模式、旁路模式和上下文模型更新技术;第 8 章的环路滤波,重点是简洁高效的去方块滤波和新增的样点自适应补偿技术;第 9 章的并行处理,重点是新增的基于片和基于波前并行处理技术;第 10 章的高层语法,重点是改进的 NAL 单元表示、新增的图像类型和参数集技术。

第三部分(第 11、12 章)。这两章是有关 HEVC 第 2 版(2014 年 10 月)和第 3 版(2015 年 4 月)中新编码扩展部分内容。其中第 11 章主要介绍了有关统一的多层编码概念和可分级视频编码的关键技术。第 12 章主要介绍了有关多视点视频编码和 3D 视频编码的关键技术。

第四部分(第 13 章)。简单介绍 HEVC 的参考软件(HM),编码复杂度比较,软硬件实现的设计考虑,给出了几个 HEVC 编解码实现的简例。

在本书的编写过程中，得到我们研究团队的多位教师和研究生的帮助和参与，引用了他人的研究报告、著作和论文，具体出处在书后的参考文献中一一列出。在此，对有助于本书编写的师生和参考文献的作者一并表示深切的谢意。对在本书选题、确定内容和编辑整理过程中付出辛勤劳动的董亚峰编辑表示由衷的感谢。

由于当今视频编码技术的飞速发展，同时受限于作者的学术水平，加上编写时间仓促，书中的错误和不足之处在所难免，真诚地欢迎广大读者予以批评指正。

作 者

2015 年 10 月于南京

目 录



第 1 章 视频编码基础	1
1.1 数字视频信号	1
1.1.1 视频信号的采集	2
1.1.2 视频信号的数字化	3
1.1.3 视频信号的显示	6
1.1.4 数字视频的格式	7
1.1.5 高清和超高清视频	10
1.2 视频信号的统计特性	10
1.2.1 图像的自相关函数	11
1.2.2 像素差值的自相关函数	12
1.3 混合编码	13
1.3.1 预测编码	13
1.3.2 变换编码	14
1.3.3 运动估计和运动补偿	15
1.3.4 混合编码框架	18
1.4 量化和熵编码	19
1.4.1 量化	19
1.4.2 Zig-zag 扫描	19
1.4.3 熵编码	20
1.5 率失真优化	20
1.5.1 图像的信源熵	20
1.5.2 率失真定理	22

1.5.3	失真率函数	25
1.5.4	有记忆信源的处理	26
1.5.5	率失真优化编码	26
1.6	图像质量的评价	29
1.6.1	主观质量评价方法	29
1.6.2	客观质量评价方法	30
1.6.3	SSIM 质量评价方法	31
	本章参考文献	32
第 2 章	视频编码的国际标准	34
2.1	H.26x 标准	35
2.1.1	H.261 标准	35
2.1.2	H.263 标准	38
2.2	MPEG-x 标准	39
2.2.1	MPEG-1 标准	40
2.2.2	MPEG-2 标准	40
2.2.3	MPEG-4 标准	42
2.3	H.264/AVC 标准	43
2.3.1	多方向帧内预测	44
2.3.2	多模式运动估计	44
2.3.3	整数变换和熵编码	45
2.3.4	差错控制	45
2.4	AVS 标准	46
2.5	VC-1 标准	46
2.6	HEVC 标准	47
2.6.1	HEVC 标准的进程	48
2.6.2	HEVC 技术概要	52
	本章参考文献	55
第 3 章	HEVC 的编码结构	57
3.1	H.264/AVC 的编码结构	57
3.1.1	宏块灵活划分	58
3.1.2	图像的条划分	58
3.1.3	档次和水平	59
3.2	HEVC 的网络适配和编码方式	61
3.2.1	视频编码层和网络提取层	61
3.2.2	三种编码方式	63
3.3	HEVC 的四叉树划分	63

3.3.1 图像的取样格式	64
3.3.2 编码树单元和编码单元划分	65
3.3.3 预测单元划分	67
3.3.4 变换单元划分	68
3.3.5 CTU 划分实例	69
3.4 HEVC 的条和片划分	70
3.4.1 条划分	71
3.4.2 片划分	72
3.4.3 条/片划分实例	74
3.5 HEVC 的档次、水平和等级	75
3.5.1 档次	76
3.5.2 水平	77
3.5.3 等级	77
本章参考文献	79
第 4 章 HEVC 的帧内预测	80
4.1 帧内编码	80
4.1.1 空域预测编码	80
4.1.2 最佳线性预测	82
4.2 H.264/AVC 的帧内预测	82
4.2.1 亮度 4×4 块的预测模式	83
4.2.2 亮度 16×16 块的预测模式	85
4.2.3 色度 8×8 块的预测模式	85
4.3 HEVC 的帧内预测模式	86
4.3.1 帧内预测 PU 的划分	86
4.3.2 亮度 PU 的帧内预测模式	87
4.3.3 色度 PU 的帧内预测模式	89
4.4 HEVC 的帧内预测过程	90
4.4.1 参考像素的准备	90
4.4.2 参考像素的平滑滤波	91
4.4.3 计算预测值	94
4.4.4 边界值的平滑	99
4.4.5 模式信息的编码	100
本章参考文献	101
第 5 章 HEVC 的帧间预测	103
5.1 帧间预测编码	103
5.1.1 帧间预测方式	103

5.1.2	基于块的运动估计	105
5.1.3	运动矢量的预测	106
5.2	H.264/AVC 的帧间预测	108
5.2.1	多模式宏块划分	108
5.2.2	高精度运动估计	108
5.2.3	双向预测条	110
5.3	HEVC 的帧间预测	111
5.3.1	帧间预测 PU 的划分	111
5.3.2	子像素插值	113
5.4	HEVC 的运动参数编码	116
5.4.1	运动参数的编码传送	116
5.4.2	Merge 模式	117
5.4.3	Skip 模式	122
5.4.4	Inter 模式	122
5.4.5	帧间预测模式的选择	123
	本章参考文献	124
第 6 章	HEVC 的变换和量化	126
6.1	变换与量化	126
6.1.1	离散余弦变换和正弦变换	126
6.1.2	量化和量化失真	128
6.2	H.264/AVC 的变换与量化	131
6.2.1	4×4 整数 DCT 变换	131
6.2.2	变换系数的量化	132
6.3	HEVC 残差的整数变换	135
6.3.1	残差二叉树 (RQT)	135
6.3.2	整数 DCT 变换	136
6.3.3	4×4 整数 DST 变换	138
6.4	HEVC 变换系数的量化	139
6.4.1	量化参数和量化步长	140
6.4.2	量化和反量化计算	140
6.4.3	加权量化矩阵	142
6.5	HEVC 变换块的编码表示	144
6.5.1	量化后系数的扫描	144
6.5.2	变换系数的表示	146
6.5.3	变换跳过	149
	本章参考文献	149

第7章 HEVC 的熵编码	151
7.1 熵编码	151
7.1.1 熵编码的要求	151
7.1.2 定长编码	153
7.1.3 变长编码	153
7.2 算术编码	155
7.2.1 一般算术编码	156
7.2.2 自适应算术编码	157
7.2.3 二进制算术编码	158
7.2.4 自适应二进制算术编码	160
7.3 HEVC 的算术编码	160
7.3.1 CABAC 框架	161
7.3.2 二进制化	162
7.3.3 上下文模型	165
7.3.4 常规编码模式	165
7.3.5 旁路编码模式	170
7.4 上下文建模和更新	171
7.4.1 上下文关系	171
7.4.2 上下文模型的初始化	171
7.4.3 上下文模型的更新	175
本章参考文献	177
第8章 HEVC 的环路滤波	178
8.1 环路滤波	178
8.1.1 方块效应的产生	178
8.1.2 环内滤波和环外滤波	180
8.2 H.264/AVC 的去方块滤波	181
8.2.1 自适应去方块滤波	181
8.2.2 边界强度测定	182
8.2.3 去方块滤波过程	184
8.3 HVEC 的环路滤波	185
8.3.1 自适应去方块滤波	186
8.3.2 样点自适应补偿	186
8.4 HEVC 的去方块滤波	187
8.4.1 去方块滤波单元	187
8.4.2 边界强度的判定	188
8.4.3 滤波强度的判定	189

8.4.4	去方块滤波过程	192
8.5	HEVC 的样值自适应补偿	194
8.5.1	信号失真及补偿	194
8.5.2	SAO 的两种模式	195
8.5.3	带补偿 (BO) 模式	196
8.5.4	边缘补偿 (EO) 模式	197
8.5.5	SAO 的模式选择和参数共享	200
	本章参考文献	202
第 9 章	HEVC 的并行处理	204
9.1	视频编码的并行处理	204
9.1.1	并行处理的主要方式	205
9.1.2	功能并行	206
9.1.3	数据并行	207
9.1.4	流水线并行	207
9.2	HEVC 的并行处理工具	209
9.2.1	片并行处理	210
9.2.2	波前并行处理	211
9.3	HEVC 的各级并行处理	213
9.3.1	GOP 级并行处理	213
9.3.2	图像级并行处理	213
9.3.3	条、片级并行处理	215
9.3.4	块级并行处理	215
9.3.5	指令级并行处理	216
9.4	去方块滤波的并行处理	217
	本章参考文献	219
第 10 章	HEVC 的高层语法	221
10.1	HEVC 语法特点	221
10.1.1	新增语法结构和元素	222
10.1.2	基本语法表示	224
10.2	H.264/AVC 语法提要	226
10.2.1	码流的分层结构	226
10.2.2	NAL 单元语法	226
10.2.3	Slice 语法	228
10.2.4	参数集	229
10.3	HEVC 的 NAL 单元	230
10.3.1	字节流格式	231

10.3.2 一般 NAL 单元语法	232
10.3.3 NAL 单元头语法	233
10.4 HEVC 的接入图像	236
10.4.1 帧内随机接入图像	236
10.4.2 前置图像	238
10.4.3 后置图像	239
10.5 HEVC 的参数集	241
10.5.1 三类参数集	241
10.5.2 视频参数集 (VPS)	243
10.5.3 序列参数集 (SPS)	245
10.5.4 图像参数集 (PPS)	245
10.6 HEVC 的参考图像集	246
10.6.1 参考图像集	246
10.6.2 参考图像列表	249
10.7 HEVC 的 SEI 和 VUI	250
10.7.1 补充增强信息 (SEI)	250
10.7.2 视频可用信息 (VUI)	251
本章参考文献	253
第 11 章 HEVC 的多层和可分级编码扩展	255
11.1 HEVC 编码扩展的进程	256
11.2 HEVC 统一的多层编码	257
11.2.1 多层编码的结构	258
11.2.2 多层编码的工具	259
11.3 HEVC 的多层扩展	260
11.3.1 层和子层	260
11.3.2 接入单元	260
11.3.3 视频参数集扩展	260
11.4 可分级视频编码	261
11.4.1 常用可分级编码方法	262
11.4.2 H.264/AVC 的可分级编码	265
11.5 HEVC 的可分级扩展	268
11.5.1 SHVC 的编码框架和性能	268
11.5.2 上采样滤波器	270
11.5.3 层间纹理预测	271
11.5.4 层间运动预测	272
11.5.5 SHVC 编码一例	273
本章参考文献	273

第 12 章 HEVC 的多视点和 3D 编码扩展	276
12.1 立体视频编码	276
12.1.1 立体视频和多视点视频	277
12.1.2 立体视频的采集和显示	278
12.1.3 多视点视频编码	280
12.1.4 H.264/AVC 的多视点编码	282
12.2 HEVC 的多视点扩展	284
12.2.1 MV-HEVC 编码系统	285
12.2.2 多视点编码工具	287
12.2.3 虚拟视点的合成	290
12.3 HEVC 的 3D 扩展	292
12.3.1 立体图像的深度图	293
12.3.2 深度图的编码	294
12.3.3 深度图的编码工具	295
本章参考文献	298
第 13 章 HEVC 的实现	300
13.1 HEVC 的参考软件 HM	300
13.2 HEVC 的复杂度	304
13.2.1 功能单元的复杂度	304
13.2.2 HM 的编码复杂度	309
13.2.3 HM 的解码复杂度	311
13.2.4 和 H.264/AVC 比较	313
13.3 HEVC 编码器的实现考虑	313
13.3.1 软件实现考虑	313
13.3.2 硬件实现考虑	314
13.4 HEVC 的解码实验	315
13.4.1 HEVC 的测试序列	315
13.4.2 基于 ARM 的解码	317
13.4.3 基于 X86 的解码	318
13.4.4 解码性能分析	319
13.5 HEVC 的编解码器简例	319
13.5.1 基于 DSP 的 HEVC 解码器	320
13.5.2 HEVC 解码器芯片	321
13.5.3 HEVC 编码器芯片	322
13.5.4 HEVC 编码系统	323
本章参考文献	324
缩略语 (Abbreviations)	326

第 1 章

视频编码基础

在信息化时代，视频技术和应用的发展，特别是高清（HD）、超高清（UHD）、3D 和多视点（Multi-View）视频技术的兴起，海量的视频信息已如潮涌般深入到我们工作和生活的方方面面。据统计，视频流数据在 2015 年占到了整个互联网流量的 90% 左右，说夸张一点，整个互联网络差不多就是视频网了。尽管近年来网络带宽和传输能力增加迅速，但仍远不能满足海量视频数据的传输和存储要求。因此视频信息量的高效压缩过去是、现在是、在可预见的未来也必然是解决这一矛盾的重要技术措施之一。

本书通过对最新的视频编码国际标准 High Efficiency Video Coding, HEVC, 及其扩展的介绍，帮助读者了解和国际标准有关的视频编码的基本原理、主要规范和关键技术。视频编码（Video Coding）又称视频压缩（Video Compression）技术，属于数字视频信号处理的范畴，要想掌握视频编码的原理和技术，必须对数字视频处理方面的相关基本内容有一定的了解。本章将提供这方面的一些基本内容，包括视频信号的采集、显示和数字化，视频信号统计特性，图像质量评价方法，以及最基本的视频编码技术。

1.1 数字视频信号

随着视频技术的发展，传统的模拟摄像机已经逐步被淘汰，可以直接获取数字图像的数字摄像机等已占据主流地位，大大方便了人们对视频信号的采集、处理、存储和传输。但数字视频设备的普及，并不表示视频的数字化不再重要，这是因为视频信号的“源”（来源）和“宿”（归宿）的形式都是连续的、模拟的，只是在中间的处理过程中采用了数字化的方法。所以，了解视频的数字化过程有助于对数字视频的采集和显示有深入的了解，从而有助于对数字视频信号的编码处理有深入的理解。

1.1.1 视频信号的采集

摄像机是视频信号的主要来源,随着数字化的进程,数字摄像机已经成为主流的视频信号来源,它输出的是数字视频信号或压缩的数字视频信号,同时也可以输出模拟视频信号。数字摄像机和模拟摄像机的光学镜头和感光器件部分基本是一样的,两者的区别在光电转换后续的处理电路不同,下面进行简单介绍。

1. 模拟摄像机

早期的光导管模拟摄像机已遭淘汰,现在常用的是电荷耦合器件(Charge Coupled Device, CCD)和互补氧化金属半导体(Complementary Metal-Oxide Semiconductor, CMOS)摄像机。

CCD 摄像机内的核心部件是一种固态半导体面阵集成电路,即 CCD 感光芯片,它由若干行、若干列的离散感光单元排列而成传感阵列。摄像机所对准的场景的光线通过镜头聚焦投射到阵列上,每个感光单元由于光照的作用而产生与输入光强成正比的输出电荷。这些电荷通过适当的逻辑电路,按照逐单元、逐行的顺序,在一帧的时间内将整个阵列的所有单元的电压送出,经 AD 变换、信号处理和 DA 变换后形成标准的模拟视频信号输出。

CMOS 摄像机的传感器的工作过程和 CCD 类似,但集成度比较高,处理过程比较简单,因此采集速度可以做到每秒数百帧,功耗也比较小,耗电量大约为同类 CCD 的 1/3。CMOS 的不足之处在于传感噪声比 CCD 高,但这一差距正在逐步缩小,所以今天的便携式摄像设备几乎都使用 CMOS 传感器。

实际上,无论 CCD 还是 CMOS 本身都没有参与图像感应(光电转换),它们使用的是同一类“光电二极管”传感器(半导体 PN 结),能够转换光线的强弱为电压的高低,光线进入图像半导体越多,电子产生的电子也越多,从传感器输出的电压也越高。而 CCD 和 CMOS 主要作用是将图像半导体中出来的电信号快速地储存、传输到信号处理部分。

感光平面阵列中每一行基元数(像素数)的多少和行数的多少,决定了所摄图像清晰度的高低。常用的 CCD、CMOS 摄像机的像素数在 40~800 万,如 1080p 视频(1920×1080)的像素数约为 200 万。

2. 数字摄像机

数字视频信号有两种获得的途径,一种是直接的方式,另一种是间接的方式。所谓间接方式是指将模拟视频信号数字化以后产生数字视频。近年来随着电子领域数字化的进程,越来越多的直接输出数字视频信号的数字摄像机已成主流。这样的摄像机可以直接和计算机、数字图像设备相连接,而不需要经过 A/D 转换。

数字摄像机和模拟摄像机的构成中感光部分基本相同,不同的是数字摄像机增加了数字化信号处理部分。数字摄像机对经过 CCD(或 CMOS)光电转换得到的视频信号进行数字化,获得数字视频信号输出,或再经过数字信号处理、数据压缩,最终输出已压缩的数字视频信号。

在这个意义上可以将数字摄像机称为“数字处理摄像机”。

现在大多数的数字摄像机都具备符合 IEEE1394 接口和 HDMI 接口规范的输出。IEEE1394 俗称“火线”(Fire Wire)接口, HDMI (High Definition Multimedia Interface) 是高清多媒体接口, 包括了数字视频和音频数据。它们已普遍用于和 PC 机或其他设备相连, 高速传送视音频信号。当然, 这类摄像通常还带有普通模拟复合视频输出及 S-Video 分量视频输出。

1.1.2 视频信号的数字化

通常意义下的图像是光强度的二维分布, 是平面坐标 (x, y) 的函数 $f(x, y)$ 。如果是一幅彩色图像, 函数值还应反映出色彩变化, 可用 $f(x, y, \lambda)$ 表示, 其中 λ 为波长; 如果是活动彩色图像, 即视频图像, 还应是时间 t 的函数, 可表示为 $f(x, y, \lambda, t)$; 当然, 如果是立体视频图像, 还应是空间深度 z 的函数, 可表示为 $f(x, y, z, \lambda, t)$ 。对常规图像来说, $f()$ 是一个连续、非负、有界函数, 即 $0 \leq f(x, y, z, \lambda, t) < \infty$ 。

视频图像的连续性包含了三个方面的含义, 即空间位置延续的连续性, 每一个位置上光强度变化的连续性和时间方向变化的连续性。连续的视频图像无法用计算机等数字化设备进行处理、传输或存储, 所以必须将连续(模拟)的视频信号转变为离散(数字)的视频信号, 这个转变的过程称为视频信号的数字化, 它一般包含三项操作: 图像的采样、样值的量化和量化值的编码。

1. 图像的采样

图像在空间上的离散化过程称为采样或取样 (Sampling), 被选取的点称为采样点或样点, 也称为像素 (Pixel)。在采样点上的函数值称为采样值或样值。采样实际上就是用空间有限采样点的值来代替连续坐标上的函数值。一幅图像应取多少样点才能够完全由这些样点来重建原图像呢? 如果样点取得过多, 增加了用于表示这些样点的信息量; 如果样点取得过少, 则有可能会丢失原图像所包含的信息。所以最少的样点数应该满足一定的约束条件: 由这些样点, 采用某种方法能够完全重建原图像。实际上, 这就是二维采样定理的内容。采样定理和信号的频谱有关, 下面从图像信号的频谱说起。

(1) 图像信号的频谱

定义二维图像信号 $f(x, y)$ 的傅里叶 (Fourier) 变换和反变换分别为

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy \quad (1.1)$$

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{j2\pi(ux+vy)} du dv \quad (1.2)$$

其中, u, v 是频率域坐标, $F(u, v)$ 称作 $f(x, y)$ 的傅里叶变换或频谱, 它表明了空间频率成分与二维图像信号之间的相互关系, 其物理意义是 $f(x, y)$ 可由频率域上的各谐波分量迭加得到。

实际的二维图像, 其傅里叶变换一般是在频率域上是有界的, 即信号频谱的有效成分总是落在一定的频率域范围之内。主要依据在于: 图像中景物的复杂性具有一定的限度, 其中大部