

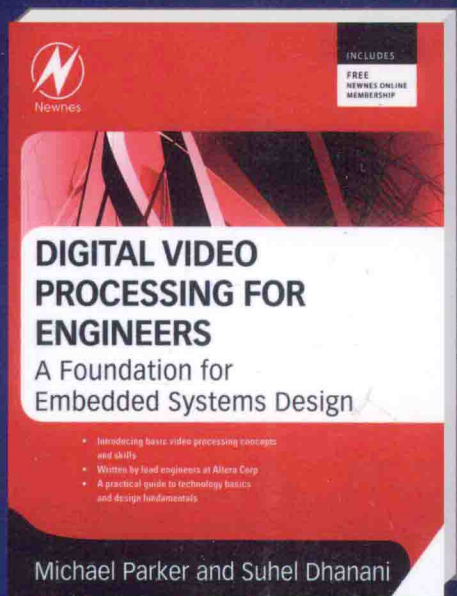
本书是从事数字视频处理相关工作的工程师必读的入门书。

FPGA工程师通过本书能够深入理解数字视频处理的概念、标准以及具体实现。

本书的两位作者均在Altera公司负责相关产品线，具有非常丰富的经验。



电子与嵌入式系统
设计译丛



Digital Video Processing for Engineers
A Foundation for Embedded Systems Design

嵌入式系统数字视频处理 权威指南

[美] Michael Parker Suhel Dhanani 著

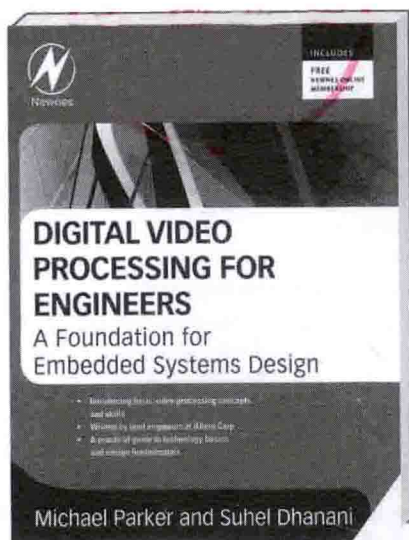
邓天平 译



机械工业出版社
China Machine Press



电子与嵌入式系统
设计译丛



Digital Video Processing for Engineers
A Foundation for Embedded Systems Design

嵌入式系统数字视频处理 权威指南

[美] Michael Parker Suhel Dhanani 著

邓天平 译



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

嵌入式系统数字视频处理权威指南/ (美) 帕克 (Parker, M.), (美) 迪哈那尼 (Dhanani, S.) 著; 邓天平译. —北京: 机械工业出版社, 2014.8

(电子与嵌入式系统设计译丛)

书名原文: Digital Video Processing for Engineers: A Foundation for Embedded Systems Design

ISBN 978-7-111-47104-2

I. 嵌… II. ①帕… ②迪… ③邓… III. 数字视频系统-数字信号处理-指南 IV. TN941.3-62

中国版本图书馆CIP数据核字 (2014) 第133204号

本书版权登记号: 图字: 01-2013-8811

Digital Video Processing for Engineers: A Foundation for Embedded Systems Design

Michael Parker and Suhel Dhanani

ISBN: 978-0-12-415760-6

Copyright © 2013 by Elsevier Inc. All rights reserved.

Authorized Simplified Chinese translation edition published by the Proprietor.

Copyright © 2014 by Elsevier (Singapore) Pte Ltd. All rights reserved.

Printed in China by China Machine Press under special arrangement with Elsevier (Singapore) Pte Ltd. This edition is authorized for sale in China only, excluding Hong Kong SAR, Macau SAR and Taiwan. Unauthorized export of this edition in violation of the Copyright Act. Violation of this Law is subject to Civil and Criminal Penalties.

本书简体中文版由 Elsevier (Singapore) Pte Ltd. 授权机械工业出版社在中国大陆境内独家出版和发行。本版仅限在中国境内 (不包括香港特别行政区、澳门特别行政区及台湾地区) 出版及标价销售。未经许可之出口, 视为违反著作权法, 将受法律之制裁。

本书封底贴有 Elsevier 防伪标签, 无标签者不得销售。

本书引入了数字视频处理领域最新的技术, 而且将数字视频处理和 FPGA 技术相结合, 以直观的方式、用最少的数学公式推导阐述了视频处理的核心概念和标准, 并辅以参考设计、循序渐进的 FPGA 开发实例、核心标准和系统结构映射, 对使用 FPGA 的工程师如何在项目中应用数字视频处理技术提供实际的指导。

嵌入式系统数字视频处理权威指南

[美] Michael Parker 等著

出版发行: 机械工业出版社 (北京市西城区百万庄大街22号 邮政编码: 100037)

责任编辑: 朱 瑛

责任校对: 董纪丽

印 刷: 北京市荣盛彩色印刷有限公司

版 次: 2014年8月第1版第1次印刷

开 本: 186mm×240mm 1/16

印 张: 10

书 号: ISBN 978-7-111-47104-2

定 价: 59.00元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88378991 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzsj@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光/邹晓东

译者序

数字视频涵盖了众多应用领域，随着计算机网络技术的不断发展和数字视频应用需求的多样化，对于数字视频的研究和应用也越来越广泛。在许多情况下，没有采用视频处理的传统行业现在也迫切需要将视频技术集成到其产品中。这种趋势对系统设计者和应用工程师提出了掌握视频技术的新要求。

当前，对于视频编码技术的研究不再仅仅局限于压缩特性，而渐渐开始向网络适应性、用户交互性等方面转移。因此，近几年来，视频编码技术一方面继续以混合编码为框架研究如何进一步提高压缩特性，另一方面不断地向可伸缩编码、多视点编码、智能编码等多分支方向发展。

本书不但引入了数字视频处理领域的最新技术，而且将数字视频处理和 FPGA 技术相结合，这是目前该领域的热点，因此本书处于国内先进水平。

本书图文并茂，以直观的方式、用最少的数学公式推导阐述了视频处理的核心概念和标准，并辅以参考设计、循序渐进的 FPGA 开发实例、核心标准和系统结构映射，对使用 FPGA 的工程师如何在项目中应用数字视频处理技术提供实际的指导。

本书共 21 章，第 1 章介绍了视频技术在现实世界中的应用；第 2 章介绍了数字视频处理的基本概念，如像素、帧、分辨率和颜色空间等；第 3 章介绍了采样和混叠的基本概念，如奈奎斯特采样准则；第 4 章介绍了数字滤波器基本知识，如中值滤波器、FIR 滤波器等；第 5 章介绍了视频缩放的基本概念；第 6 章阐述了如何进行视频的去隔行；第 7 章介绍了 alpha 混合的基本原理及其硬件实现；第 8 章描述了图像传感器信号处理的硬件实现及其应用；第 9 章介绍了通用的视频接口，如 SDI、HDMI、DVI 和 VGA 等；第 10 章介绍了视频旋转的基本概念；第 11 章阐述了熵的基本概念以及预测编码和量化的基本原理；第 12 章介绍了视频的频域表示方法；第 13 章阐述了图像压缩的基础知识及 JPEG 实现；第 14 章介绍了视频压缩的基础知识，如块尺寸、运动估计等；第 15 章介绍了 MPEG-2 和 H.264 视频压缩标准；第 16 章描述了视频噪声及压缩失真的基本问题；第 17 章阐述了视频调制和传输；第 18 章介绍视频信号如何在 IP 网络上传输；第 19 章描述了视频分割和聚焦；第 20 章阐述了构建视频处理系统时存储器的设计；第 21 章阐述了如何在 FPGA 上构建视频处理系统。

纵观全书，我认为有如下特点。

1. 重视基础

本书专门针对那些没有视频处理经验，但是需要在他们的设计中采用视频系统的工程师，详细介绍了基本的视频处理概念，如数字视频的基本概念：像素、帧、分辨率和颜色空间、采样和混叠等；介绍了基本的视频接口，如 SDI、HDMI、DVI 和 VGA 等；阐述了熵的基本

概念、预测编码和量化的基本原理、图像压缩的基础知识以及 JPEG 实现等。这些概念为读者扩展他们的视频知识提供了坚实的理论基础。

2. 强调实际应用

目前, 数字视频处理涉及的领域越来越广, 本书基于这些应用, 强调了实际处理过程中的滤波、去隔行、alpha 混合、视频旋转、调制和传输等。这些技术可以广泛应用于数字影院、医学成像、视频监控和军事等领域。

3. 内容丰富、新颖

本书所讨论的内容涉及有关数字视频处理的几乎所有问题, 如去隔行、alpha 混合、视频旋转、视频压缩及其标准等。

本书作者 Michael Parker 和 Subel Dhanani 是全球领先的 FPGA 视频 IP 和设备供应商 Altera 公司首席工程师。在 FPGA 和数字视频处理方面积累了丰富的经验。本书是他们多年研发工作的结晶。

本书专为那些需要理解数字视频基本概念和应用的人员编写, 使得他们能够构建自己的视频系统或者集成第三方视频技术到他们的产品中。本书也可以作为需要运用到视频技术的行业工程师的参考资料。

感谢机械工业出版社为出版这本书所做的努力。特别感谢机械工业出版社的张国强编辑, 没有他的努力, 本书今天能够付梓出版是难以想象的。

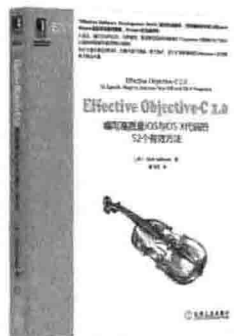
还要感谢华中科技大学电子与信息工程系的领导和同事在本书翻译和出版过程中给予的支持。感谢我的家人, 感谢他们在背后默默的支持和鼓励。

由于译者学术水平和英语水平有限, 译文肯定存在未发现的错误或不当之处, 敬请读者给予批评指正。意见请寄至华中科技大学电子与信息工程系 (430074), 也可以发送电子邮件至 dengtp@hust.edu.cn。

邓天平

2014 年 3 月于华中科技大学

推荐阅读



目 录

译者序

第 1 章 现实世界中的视频 / 1

第 2 章 视频处理介绍 / 3

- 2.1 数字视频: 像素和分辨率 / 3
- 2.2 数字视频: 像素和比特 / 4
- 2.3 数字视频: 颜色空间 / 5
- 2.4 视频处理性能 / 6

第 3 章 采样和混叠 / 7

- 3.1 奈奎斯特采样准则 / 10

第 4 章 数字滤波器介绍 / 12

- 4.1 中值滤波 / 12
- 4.2 FIR 滤波器 / 13
- 4.3 FIR 滤波器结构 / 13
- 4.4 频率响应计算 / 16

第 5 章 视频缩放 / 20

- 5.1 理解视频缩放 / 21
- 5.2 实现视频缩放 / 23
- 5.3 不同长宽比的视频缩放 / 25
- 5.4 结论 / 27

第 6 章 视频去隔行 / 28

- 6.1 基本去隔行技术 / 29
- 6.2 运动自适应去隔行: 基础 / 31
- 6.3 逻辑要求 / 33
- 6.4 节奏检测 / 34
- 6.5 结论 / 35

第 7 章 alpha 混合 / 36

- 7.1 引言 / 36

- 7.2 alpha 混合的概念和数学表达 / 36

- 7.3 alpha 混合的硬件实现 / 37

- 7.4 创建不同的背景 / 38

- 7.5 结论 / 38

第 8 章 图像传感器信号处理 / 39

- 8.1 CMOS 传感器基础 / 39

- 8.2 Bayer 去马赛克的简单硬件实现 / 41

- 8.3 军用电光红外系统传感器处理 / 41

- 8.4 结论 / 43

第 9 章 视频接口 / 44

- 9.1 SDI / 44

- 9.2 显示端口 / 45

- 9.3 HDMI / 45

- 9.4 DVI / 45

- 9.5 VGA / 46

- 9.6 CVBS / 46

- 9.7 S-视频 / 46

- 9.8 分量视频 / 47

第 10 章 视频旋转 / 48

- 10.1 插值 / 49

第 11 章 熵、预测编码和量化 / 50

- 11.1 熵 / 50

- 11.2 哈夫曼编码 / 51

- 11.3 马尔可夫源 / 52

- 11.4 预测编码 / 53

- 11.5 差分编码 / 53

- 11.6 无损压缩 / 54

11.7 量化 / 55

11.8 分贝 / 57

第 12 章 频域表示 / 59

12.1 DFT 和 IDFT 方程 / 60

12.1.1 第一个 DFT 例子 / 61

12.1.2 第二个 DFT 例子 / 62

12.1.3 第三个 DFT 例子 / 63

12.1.4 第四个 DFT 例子 / 63

12.2 快速傅里叶变换 / 65

12.3 离散余弦变换 / 68

第 13 章 图像压缩基本原理 / 72

13.1 基准 JPEG / 72

13.2 直流缩放 / 72

13.3 量化表 / 73

13.4 熵编码 / 74

13.5 JPEG 扩展 / 75

第 14 章 视频压缩基本原理 / 77

14.1 块尺寸 / 77

14.2 运动估计 / 79

14.3 帧处理次序 / 80

14.4 压缩 I 帧 / 81

14.5 压缩 P 帧 / 81

14.6 压缩 B 帧 / 82

14.7 速率控制和缓冲 / 82

14.8 量化比例因子 / 83

第 15 章 从 MPEG 到 H.264 视频压缩 / 87

15.1 MPEG-2 / 87

15.2 H.264 视频压缩标准 / 91

15.2.1 H.264 的级和类 / 91

15.2.2 支持 4×4 整数 DCT 的
H.264 / 94

15.2.3 H.264 对数量化 / 95

15.2.4 H.264 频率相关的
量化 / 95

15.2.5 H.264 像素插值 / 95

15.2.6 H.264 可变块尺寸运动
估计 / 95

15.2.7 H.264 去块滤波器 / 96

15.2.8 H.264 时域预测 / 96

15.2.9 H.264 运动自适应帧和场
编码 / 96

15.2.10 H.264 空间预测 / 97

15.2.11 H.264 熵编码 / 97

15.2.12 H.264 冗余片 / 97

15.2.13 H.264 任意片排序 / 97

15.2.14 H.264 灵活的宏块
排序 / 97

15.2.15 H.264 其他的特性 / 97

15.3 数字影院应用 / 98

第 16 章 视频噪声和压缩失真 / 99

16.1 椒盐噪声 / 99

16.2 蚊式噪声 / 100

16.3 块效应 / 101

第 17 章 视频调制和传输 / 103

17.1 复调制和解调 / 103

17.2 调制信号带宽 / 105

17.3 脉冲成形滤波器 / 107

17.4 升余弦滤波器 / 109

17.5 信号上转换 / 114

17.6 数字上转换 / 115

第 18 章 IP 视频 / 119

18.1 IP 基础 / 119

18.2 封装 / 120

18.3 流媒体视频 / 120

18.4 传输协议 / 121

18.5 IP 传输 / 122

18.6 互联网视频问题 / 123

18.7 流式视频 / 124

18.8 多播视频 / 125

- 18.9 视频会议 / 125
- 第 19 章 分割和聚焦 / 127**
 - 19.1 聚焦测量 / 127
 - 19.1.1 梯度技术 / 128
 - 19.1.2 方差技术 / 129
 - 19.2 分割 / 130
 - 19.2.1 模板匹配 / 130
 - 19.2.2 FPGA 实现 / 133
- 第 20 章 构建视频处理系统的存储器考虑 / 135**
 - 20.1 帧缓冲器 / 135
 - 20.2 计算外置存储器的带宽需求 / 137
 - 20.3 计算片上存储器 / 140
 - 20.4 结论 / 142
- 第 21 章 基于 FPGA 的视频系统的调试 / 143**
 - 21.1 时序分析 / 143
 - 21.1.1 检查设计是否满足时序 / 144
 - 21.1.2 如果不满足时序, 修正你的设计 / 144
 - 21.2 系统控制台调试器 / 145
 - 21.3 检查时钟和复位是否正常工作 / 145
 - 21.4 时钟和流量控制的视频流 / 146
 - 21.5 调试工具 / 146
 - 21.6 从时钟到流量控制视频流的转换 / 148
 - 21.7 从流量控制到时钟视频流的转换 / 149
 - 21.8 自由运行的流视频接口 / 149
 - 21.9 不足的内存带宽 / 150
 - 21.10 检查流内数据 / 150
 - 21.11 小结 / 151

现实世界中的视频

视频最初作为纯粹的模拟技术出现。用相机将连续图像捕获在电影流中，通过闪光来播放电影，以显示运动电影中的每一帧，当帧率足够高时就显示连续的动作。然而，容易看见闪烁！

早期的广播电视中，改进的系统采用亮度（即光强度）信息表示模拟信号。为了传输图像，使用连续的水平扫描发送亮度信息。足够的水平扫描构成二维图像。电视和监视器采用阴极射线枪发射电子流，以激发用磷粉涂敷的屏幕。缓慢消退的荧光粉有助于消除闪烁。阴极射线枪扫描连续的行，通过磁路控制，每一行刚好在前一行的下面。扫描速度如此迅速，因此图像以每秒钟 25~30 帧的速率重现。亮度信号用于控制电子流的强度。

水平同步信号用于分离水平扫描周期。水平同步时间是在每个行扫描线结束后的短脉冲。其时间长度必须确保电子枪能返回到屏幕的左边，以准备下次扫描。同样，垂直同步信号出现在最后即最底端扫描结束后，用于分离每个视频帧。垂直同步间隔时间比水平同步时间要长很多，其使得电子枪能从屏幕的右下角返回到左上角，以开始新的一帧扫描，如图 1-1 所示。

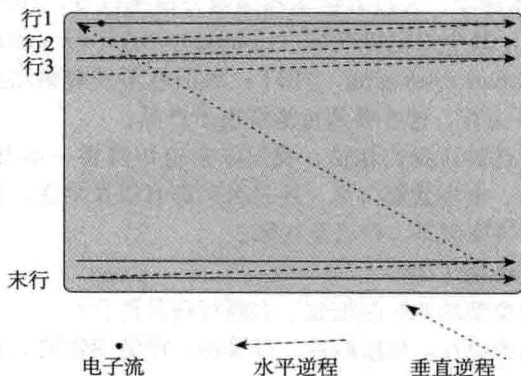


图 1-1 视频光栅扫描

然后，颜色信息以红和蓝色调的形式加入，称之为色度信息。其叠加于亮度信号之上，因此，彩色电视系统后向兼容于黑白电视系统。

现代电视信号采用数字方式表示、存储和传输。数字表示开创了视频的各种新应用。数字视频处理在各种各样的市场中日益重要，如视频监控、视频会议、医学成像、包含无人机 (Unmanned Aerial Vehicle, UAV) 的军用成像、武器瞄准和夜视、广播、数码相机、工业显示和消

费电子。所有的这些都经历了从标清到高清以及超高清视频处理的历程，这些大规模的升级周期长达十年。在某些场合，采用数字技术的系统已经代替了旧式的模拟视频处理设备。

在许多情况下，传统的没有采用视频处理的行业现在也必须集成视频技术到他们的产品。例如汽车中的后置摄像头、娱乐中心、“车道偏离”和“平视显示器”；网络服务器和路由器中的视频数据处理；军事系统中提供场景意识的视频共享及合并技术；军用和商业机载系统中的监控和导引；机器人系统；安全监控系统和很多其他应用中的面部和其他特征识别（如车牌）。这种趋势对系统设计者和实现工程师提出了理解视频技术的新要求。本书是专为那些需要理解基本概念和应用的人员编写，使得他们能够构建自己的视频系统，或者集成第三方的视频技术到他们的产品。

另外的目标读者来自许多需要视频技术的行业，他们从事技术性的市场营销和管理，因此，需要理解基本概念和应用，而不会被细节和实现的复杂性吓倒。

这些新的视频应用规模正迅速增长。例如，下面是一些已公开的预测：

- ABI^①研究表明，视频监控市场有望爆炸性增长，公司预测其收入将从2006年的大约135亿美元显著增长到2013年的460亿美元。这些数字包含相机、计算机和存储器、专业服务和硬件基础结构：所有都实现了端到端的安全系统。
- 依据温豪斯（Wainhouse）的研究，视频会议所有的终端市场将从2007年的13亿美元增长到2013年的49亿美元。视频会议基础结构的产品收入，包含微控制器（Micro Control Unit, MCU）、网关和网守，预计将同时增加到7.25亿美元。
- 高清普及率：在美国和欧洲，高清信号的研发、存储、编辑和传输还有许多工作要做。

在接下来的5年，数字影院将会加速增长，在2010~2011年，10 000个美国电影院将要升级。数字影院极大推动了高清和4K（4096×2160）分辨率视频处理方面的设计。

在2010年，基于技术进步、人口老龄化和保健方法的改变，160亿美元的美国医疗图像产品行业将每年增加6%，其中以电脑断层（Computerized Tomography, CT）扫描仪、核磁共振成像（Magnetic Resonance Imaging, MRT）和正电子发射断层显像（Positron Emission Tomography, PET）设备为首，这些将超过消费电子产品。

所有的这些和更多的趋势让我们相信，我们非常迫切需要一本书，揭开视频处理的神秘面纱。无论是专业工程师、市场营销人员，还是公司高官以及学生，都需要理解：

- 就颜色、比特和分辨率而言，什么是视频？
- 视频传输的典型方法是什么？
- 什么是视频处理的典型功能？是缩放、去隔行还是混合？
- 涉及构建视频处理的设计，如帧缓冲、行缓冲、存储器带宽、嵌入式控制等，有哪些典型的挑战？
- 什么是视频压缩？
- 视频如何调制、编码和传输？

这些概念为读者扩展他们的视频知识提供了坚实的理论基础。这本书将是为此这个学科的工程师和学生撰写的第一本教材。

① 美国著名市场研究公司 ABI research。——译者注

视频处理介绍

对视频进行大小调整、净化或者压缩操作的处理，正在以数字处理方式增加，并且迅速在商业和家庭设备中变得普及。

本书以数字方式研究视频，为此我们将讨论像素、颜色空间等。我们假设视频是由像素构成，并且一排像素构成一行，行的集合构成视频帧。在有些章节，我们将简单讨论旧的模拟格式，但本书中主要是以数字方式呈现。

本书是一本引导性教材，且是阐述数字视频概念的第一本书，因此数字视频主要是通过图片以及少量的数学表达来阐述。

2.1 数字视频：像素和分辨率

数字视频由像素构成，你可以将像素看作电视机屏幕上的一个点。在一个视频帧中有许多像素，并且1秒钟内有许多帧，一般是每秒60帧。

当你观看电视时，有不同的分辨率，如标准清晰度（Standard Definition, SD）、720P 逐行扫描（Progressive Scan）的高清（High Definition, HD）或1080P的高清。分辨率决定你的电视机显示的像素数。图2-1显示了这些不同分辨率的像素数。对你看到的同样的视频帧来说，标清视频有30万像素，而1080P电视有超过两百万的像素。无疑高清看起来更好。

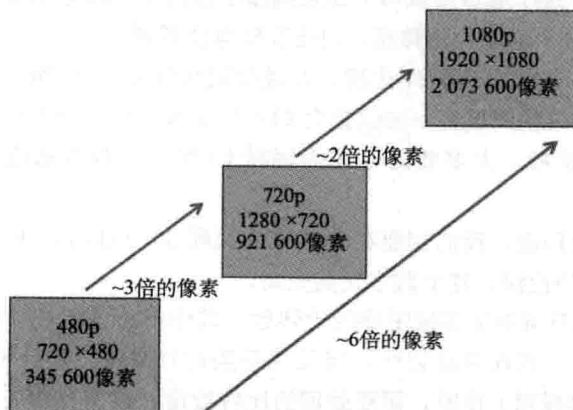


图 2-1 每帧视频中增加的像素数

令人关注的是，旧式的阴极射线管（Cathode Ray Tube, CRT）电视只有标清（SD）一半的分辨率，因此，对每个视频帧，从 CRT 电视到新的 1080P 电视，你的眼睛看到的有超过 12 倍的像素。

像素的数量有着巨大的差别。

举另外一个例子，当苹果创造了新的应用于 iPhone4 的视网膜显示后，其在消费电子中应用特别普遍。相比于旧的 iPhone3 320×480 的分辨率，新的 iPhone4 分辨率为 960×640 像素。因此苹果公司找到了在同样尺寸的屏幕上增加 4 倍像素数的方法。

像素的数量也决定了用于处理这些像素的硬件复杂度。因为所有的操作都是以比特为单位，下面我们来看看像素如何转换成比特。

2.2 数字视频：像素和比特

每个像素都有唯一的颜色，其是原色红、绿、蓝的组合。关键取决于红色、绿色和蓝色的多少，并且其数量的多少由像素值精确描述。像素的值用比特表示，且用到的比特越多，表示就越准确。然而，必须记住，这些比特的存储、处理和和设备之间的传输是十分昂贵的。因而必须找到一个完美的平衡。

每个像素有红（R）、绿（G）和蓝（B）三个分量。虽然还有别的方法表示这个，但我们首先研究红、绿和蓝。我们均采用 8 比特存储红色、绿色和蓝色的值，对红、绿和蓝的每个 8 比特，具有 2^8 即 256 种不同的值。在这种情况下，我们称之为颜色深度为 8 或 8 比特色深。

有些 HD 视频编码采用 10 比特甚至 12 比特色深，也就是每个 RGB 分量用 10 或 12 比特来编码。

色深的比特数越多越好，但这些比特必须相加。比如 8 比特色深，每个像素需要 $8 \times 3 = 24$ 比特来表示它的值。

现在我们来看看你家里的平板电视。你可能记得这个电视是 1080P，但售货员可能说 1920×1080 的分辨率。这个意思是说每个视频帧显示在这个平板电视上有 1080 行，并且每行有 1920 个像素。因此你早就在讨论像素，只是你没有注意到。

综上所述，因为每个像素需要 24 比特，并且视频每行有 1920 像素，每帧有 1080 行，这意味着你的平板电视在每帧图像显示的信息为 $24 \times 1920 \times 1080 = 49\,766\,400$ 比特。近似为 5 千万比特，也就是 50M 比特，大多数的电视为每秒 60 帧，一些新的电视甚至可以达到每秒 120 帧。

为了给你带来观赏乐趣，我们需要在每秒钟内处理 30 亿比特，也就是 3G 比特。并且这只是每秒 60 帧和 8 比特色深，这个数字还能更高。

表 2-1 显示了不同分辨率下每帧图像的比特数。其中每个像素用 30 比特表示，也显示了隔行扫描视频的影响——现在只需记住，当视频是隔行扫描时，其分辨率减半。这个表是为了让你了解当采用数字视频工作时，需要处理的比特数量。数字视频处理是计算密集型工作，特别是对高清分辨率，其主要原因是数字视频的像素数量是如此之多。

表 2-1

图像尺寸	帧尺寸 (总像素数)	帧大小 (假定每个像素 30bit)
1920×1080p	1920×1080=2M 像素	60Mbit
1920×1080i	1920×1080×0.5=1M 像素	30Mbit
1280×720p	1280×720=900K 像素	27Mbit
SD 720×480p	720×480×0.5=173K 像素	5.19Mbit

2.3 数字视频：颜色空间

颜色空间就是我们指定、产生和可视化颜色的方法。每个像素具有一定的颜色，简单地说，其为红色、绿色和蓝色的组合。我们用 8 比特表示每个颜色的值。如果像素是纯红的，像素的红色分量将会是 11111111，且另外的两个分量（蓝和绿）将是 00000000。

当这些值加到一起的时候，我们看到红色。如果另外两个颜色的值不为 0，那么得到的颜色是红色和一些绿色及一些蓝色的混合。这个颜色空间是相加的——得到的像素颜色是每个颜色的强度的和，如图 2-2 所示。

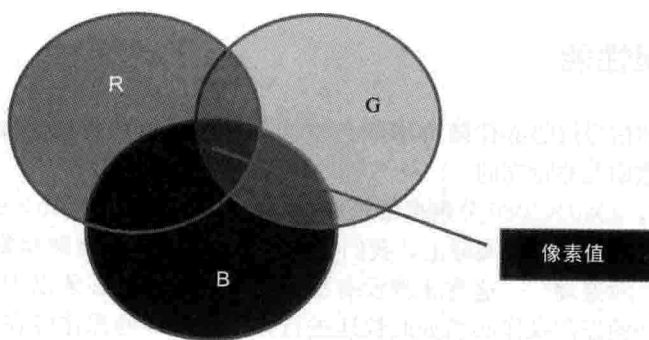


图 2-2

RGB 颜色模型用于旧式的 CRT 电视以及现在的液晶电视上显示颜色。每个值激发 CRT 面板上的红色、绿色和蓝色的荧光粉。然而对数字电视，得到的像素值存储在硬件中，然后被转换成电压，再点亮屏幕上的像素。还有包括伽马校正等其他问题，我们将在后面研究。

印刷机通过青色、紫色、黄色和黑色墨水在纸上的反射率和吸收率来描述颜色刺激。因而，它们工作在不同的颜色空间。

颜色空间有许多，其中更感兴趣的是 YCrCb 颜色空间，这种颜色空间通过明度（亮度）和 Cr, Cb 来表示编码的像素值，其中 Y 和 Cr, Cb 是 RGB 的组合，因为人眼对明度或亮度信号比较敏感，而对彩色不敏感，所以这种表示颜色的方法很有用。当像素值被分解成亮度和彩色时，我们能较少的比特（低分辨率）来编码彩色信息，而且人眼不会觉得有差别。

YCrCb 是编码 RGB 彩色的另一种方法，并且在处理中能用较少的比特，但是在视频显示前，我们必须将其重新转换成 RGB。

从一个颜色空间 (RGB) 到另一个颜色空间 (YCrCb) 转换像素值的方法是将 RGB 空间

中的每个分量乘以一个固定的常数——见图 2-3。

$$Y = R \times 0.299 + G \times 0.587 + B \times 0.114$$

$$Cr = R \times (-0.169) + G \times (-0.332) + B \times 0.500 + 128$$

$$Cb = R \times 0.500 + G \times (-0.419) + B \times (-0.0813) + 128$$

图 2-3

在硬件方面，需要乘法器和加法器实现这些操作。任何性能较好的处理器都能实现，具有内置数字信号处理（Digital Signal Processing, DSP）能力的可编程门阵列（Field-Programmable Gate Array, FPGA）当然也能很好和快速地实现这一操作。

当你开始从一个颜色空间到另一个颜色空间转换像素值时，每个步骤都有多个转换。

例如：

转换 RGB 到 YCrCb → 传输

→ 转换回 RGB → 处理视频

→ 转换回 YCrCb → 传输

→ 转换回 RGB

→ 显示

2.4 视频处理性能

任何视频处理的信号链势必伴随许多颜色空间的转换。这些转换必须以像素速率进行，而高清视频的像素速率是非常高的。

比如每秒 60 帧，1920×1080 分辨率的视频，每秒钟有 1920×1080×60 个像素，这意味着每秒钟有 124.4 百万的像素，实际上，我们现在还忽略了和每帧视频相关的时序信息。

如果视频必须实时处理——这意味着没有缓冲——那么像素必须以 124.4MHz 的频率来处理。每个像素比特的任何操作必须如此快地进行，以致同样的操作能在 1 秒钟内对颜色空间中的 124.4 百万像素进行。换句话说，频率为 124.4MHz。实际上，因为我们必须考虑每帧视频的时序信息，频率约为 148MHz。

这个频率是非常重要的，因为无论你选择什么处理平台，都必须以这个频率工作。也必须记住每个像素都是由颜色平面组成，并且每个颜色平面是用一定量的比特来表示的。例如，和每个颜色 R、G、B 有关的比特对应于一个颜色平面，我们假定每个颜色平面有 8 比特，并且采用简单的 RGB 颜色空间。回到处理速度，每个像素的 24 比特不得不以 148MHz 的频率操作。对于 FPGA，因为能采用 24 比特的硬件处理电路，这是相对容易的，但如果你使用 8 比特的 DSP，因为只能操作 8 比特，那么你必须使这个 DSP 以 3×148MHz 频率工作以跟上进入的像素。实际上，高清视频的操作通常在 32bit 的 DSP 或处理器上进行。

采样和混叠

无论什么用数字表示时，都必须被采样。这意味着连续的、不间断的目标被数字化时，必须用时间或空间上紧密相连的许多样本表示。如果样本足够紧密，那么目标看起来就是连续的。视频在时间上的采样，意味着景物快照的获取足够快，那么当回放时，就会给人连续运动的视频印象。在视频系统中，较为普遍的是使用每秒 25~30 幅图像，并且广泛被许多观众所接受（在隔行扫描视频中，这个数将加倍，这将在后面才涉及）。

空间采样应用于图像，在垂直和水平方向上，采用足够多的像素表示图像，以致图像看起来既逼真又清晰。当缩放图像时，采样率可以改变，因为这是从一个分辨率到另一个分辨率的转换过程。

为了获取模拟信号并且将其数字化，我们需要对信号采样。对一维信号，我们将要用到模-数转换器（Analog-to-Digital Converter, ADC）。照相机必须同时在两个方向对信号采样，并且在每一个瞬间获得许多样本。不像传统的模拟照相机，利用光的化学反应曝光胶卷，数码照相机使用半导体器件转换光信号到电信号，而且电信号可以同时用许多模-数转换器采样。普遍采用的是电荷耦合器件（Charge-Coupled Device, CCD）。新的选择是采用互补金属氧化物半导体（Complementary Metal Oxide Semiconductor, CMOS）技术。由于这些器件能同时获取几百万的样本（这对应于数码相机的百万像素级），并且将它们转换为表示单个像素的电信号。再用 ADC 将它们转换为数字形式。

颜色采样可以通过各种各样的方法实现，但常见的方法是使用 Bayer 滤波器，它是颜色滤波矩阵。它能在每个像素的位置将进来的光过滤为不同的颜色，如红色、绿色和蓝色，并且分别记录每个颜色像素的光强度。更新的方法是将色敏传感器植入 CMOS 技术中。有许多别的因素影响这个过程，如光圈、快门速度、缩放和聚焦，在这里我们不予讨论。

为了简单地说明采样和混叠的有关概念，我们将采用一维信号。ADC 将在快速的时间间隔内测量信号，并且得到样本（就是图像和视频中的像素）。ADC 将输出数字信号，其与相应瞬间的模拟信号的幅度成比例。这可以用只有闪光灯照明物体的情况来类比。你只有在闪光灯闪烁时才能看见物体。如果物体是静止的，那么所看到的和使用正常的连续光源时很相似。但当在闪光灯下观察运动物体时，事情变得很有趣。如果物体运动很迅速，那么运动现象看起来和正常光照情况下有较大的不同。当物体的运动相当慢，或者我们足够降低闪光灯的闪烁频率，我们还是能看到奇怪的结果。我们能直观地知道，相对于被照亮的物体的运动速率，闪光灯的闪烁频率更重要。只要闪光灯的闪光速率远快于物体的运动速率，运动看起来非常连续、

正常。但当闪光灯的闪光速率低于物体的运动速率时，运动看起来就很奇怪，经常像慢动作。尽管我们能看到物体是运动的，但我们失去了运动的连续感和流畅感。

我们再举一个例子，如画在索引卡片上的一个字符的简单动画影片。为了说明这个字符跳跃或降落的运动，我们可以画出 20 或 40 张卡片，每张卡片表示同一字符在这个连续运动的一个阶段，从一张卡片到另一张卡片只是很小的运动变化。那么，拿着一张索引卡片的边沿，并且快速翻阅到另一张的边沿，卡片上的字符看起来就是连续地跳跃或降落。这也可以应用于视频处理，屏幕大约以每秒 30 次的速度更新，对我们来说，足够快的速度使我们感觉不到是单独的帧，因此运动看起来就是连续的。

因此，如果我们采样信号的速度快于信号的变化，我们将得到非常准确的信号采样值，但如果采样速度过低，我们将看到失真的信号。

下面的图形表示两个不同的正弦信号被采样。在图 3-1 中运动缓慢的信号（低频）能用指定的采样率准确地表示。但是，图 3-2 所示的快速运动的信号（高频）却不能用我们的采样率准确表示。实际上，采样值看起来好像是一个运动缓慢的信号（低频），如图 3-2 中的虚线所示。这表明对给定频率的信号，足够快的采样率是非常重要的。

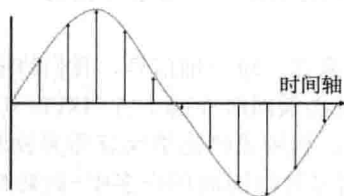


图 3-1 低频信号的采样（箭头表示采样间隔）

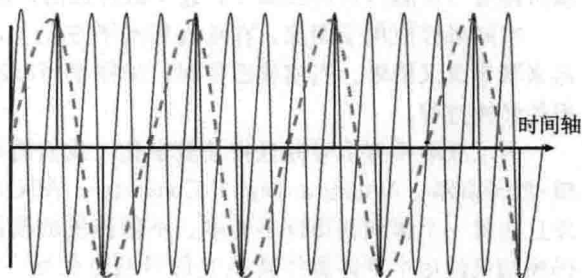


图 3-2 高频信号的采样（和图 3-1 同样的采样间隔）

如果我们将采样点相连并且平滑信号，那么虚线表示有多少采样信号出现。注意，由于实际信号（实线）在采样间隔之间变化如此迅速，所以间隔之间的运动在采样信号中是看不见的。采样信号表现为低频信号，而不是实际信号。这个影响称之为混叠。

我们需要一种方法来确定我们必须以多快的速度采样，才能准确表示给定的信号。我们也必须更好地理解当混叠出现时将会发生什么。这看起来可能很奇怪，但在有些情况下，混叠是很有用的。

再次回到闪光灯的类比，并且尝试另一个实验。设想一个旋转的轮子，在它的边沿上有一个小圆点。设定闪光灯每 $1/8$ 秒闪烁一次，也就是每秒钟闪烁 8 次。图 3-3~图 3-10 表示 6 次闪光后，我们将看到什么，这和轮子的旋转速度有关系。在所有的图片中，时间从左到右依次增加。

事实上一旦开始采样信号（也就是闪烁闪光灯），我们不能确定在两次闪光之间发生了什么。我们本能地假定信号（例子中的小圆点）采取最短路径从一次闪光出现的位置到紧接着的闪光位置。但是，正如我们在上面的例子中所看到的，这可能是错误的。小圆点可能以相反的方向沿圆周运动（采取最长路径）到达我们看到的下一次闪光位置。或者设想圆周以假定的方向旋转，但是每次闪光它旋转了一整圈再加上“额外的一点点”（9Hz 示意图）。我们