

ADI 处理器实用丛书

数字视频处理原理 及DSP实现

◎ 邢延超 皇甫伟 编著



本书特色：

- 内容系统，涵盖视频开发方方面面
- 叙述由浅入深，强调实用性
- 突出软件设计原则及工具包使用



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

ADI 处理器实用丛书

数字视频处理原理及 DSP 实现

邢延超 皇甫伟 编著

電子工業出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书在介绍数字视频信号处理基本概念和常用算法的基础上,从实用性的角度出发,借助 ADI 公司的 Blackfin 系列 DSP 处理器平台,讨论了嵌入式视频处理的相关问题。主要内容包括:绪论、数字视频基础、数字信号处理与嵌入式开发、基于 Blackfin 处理器的最小视频系统、视频应用设计原则及基础应用简介、图像与视频处理软件开发包、视频运动分析及应用、视频编解码理论及实现、视频时空滤波及实现。后面几部分中包含了运动跟踪、H.264 编解码和视频去交错应用等具体应用。

本书可作为对视频处理和 DSP 开发感兴趣的科研技术人员的参考,也可作为信息与信号处理方向研究生和高年级本科生的参考书。

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

数字视频处理原理及 DSP 实现 / 邢延超, 皇甫伟编著. —北京: 电子工业出版社, 2011 12

(ADI 处理器实用丛书)

ISBN 978-7-121-15417-1

I. ①数… II. ①邢… ②皇… III. ①数字视频系统—数字信号处理 IV. ①TN941.1

中国版本图书馆 CIP 数据核字(2011)第 253409 号

策划编辑: 竺南直 徐蔷薇

责任编辑: 谭丽莎 文字编辑: 王凌燕

印 刷: 北京京科印刷有限公司

装 订:

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: 18 字数: 461 千字

印 次: 2011 年 12 月第 1 次印刷

印 数: 3 000 册 定价: 39.00 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。

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。

服务热线: (010) 88258888。

序 言

这些年，在与电子技术领域的工程师、学者以及大学师生交流的时候，他们的聪明才智和创新能力给我留下了深刻的印象。而他们所做的设计和项目，无一不让我感觉到中国工程师队伍成长之快，和中国电子行业巨大的发展潜力。但另一方面，他们的经历和成功，也带给了我很多思考。

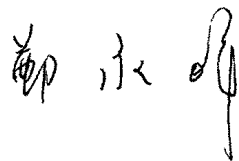
ADI 在模拟和数字信号领域中已经发展了 40 多年。在这几十年间，我们不断推动技术的创新和进步，不断提高相关领域的各类产品性能以满足客户的广泛需求，包括消费类、通信、医疗、运输和工业等方面。令人欣慰的是，至 2009 年，ADI 已经拥有遍布世界各地的 60,000 余家客户。而通过大学计划、培训、研讨会等活动所积累起来的资源更是不计其数。如何让我们的客户，让 ADI 技术产品的使用者和爱好者，真正准确、有效、快捷地掌握相关知识与设计技巧，是我们需要考虑的，也是我们为所有用户提供的非常重要的服务之一。

经过多年的运行和完善，ADI 已经拥有了一整套对中国工程师以及在校工科类学生的培养计划，如每年一届的中国大学创新设计竞赛，在高校建立的联合实验室，各类线上线下研讨会，还有在多个城市开展的高水平培训课程等等。这些计划架起了 ADI 与用户之间最直接、最有效的沟通桥梁。同时，为了使更多的电子技术领域从业者和爱好者了解数字信号处理和电子产品设计理念，我们还邀请了业内具有较深影响力的专家、学者、教授共同编写并出版一套基于 ADI 模拟和数字产品的应用技术丛书。

该丛书详细介绍了 ADI 产品在医疗电子，通信，工业仪器仪表，汽车电子等行业的应用，以理论与实际案例相结合的方式为读者们讲解了世界先进处理器的设计与使用。

丛书的出版凝聚了来自清华大学、西安电子科技大学、青岛理工大学、解放军理工大学、厦门大学、天津大学、黑龙江大学、中国科学技术大学、辽宁工业大学等多所院校老师丰富的经验和智慧。在此，感谢他们对 ADI 出版计划的大力支持。同时，也感谢电子工业出版社的竺南直博士对本丛书的出版所作出的贡献！

衷心希望能得到读者朋友的意见反馈，在你们提出的问题和建议下，我们将不断完善 ADI 丛书，不断完善 ADI 的产品和技术，与客户们一起共同开拓中国市场。



ADI 公司亚太区副总裁

前 言

近年来,数字视频处理在智能监控、消费类电子、机器人视觉、工业检测与控制及网络视频会议等领域的应用发展迅猛。同时,许多视频应用产品都在朝着小型化、便携化、低功耗和易升级等方向发展。随着高速数字处理器技术的发展,嵌入式数字视频处理已经成为重要的发展方向,具有客观的市场前景。

DSP 不仅具有高性能的数字信号处理能力,同时还具有开发速度快、升级方便且成本低等明显优势,在该领域发挥着越来越重要的作用。Blackfin 系列 DSP 是 ADI 公司开发的基于微信号结构的高性能数字信号处理器,非常适合于各种音频、视频、通信及复杂控制等领域,市场占有率也名列前茅。

本书从实用性的角度出发,首先介绍了数字视频处理的基本概念和典型算法,以及 Blackfin DSP 的基础知识和基本使用方法。然后就视频处理系统的硬件搭建和软件开发进行详细介绍,实现了一个包括视频输入、输出和传输等基本功能的最小视频系统,作为进一步开发复杂系统的基础。接下来讨论了视频开发过程中应遵循的设计原则,提供了典型的参考开发模板。考虑到复杂图像处理系统的开发已经进入模块化、集成化的开发阶段,充分利用已有的、经过优化的开发包则是快速、高效开发高性能视频处理产品的必然趋势,本书最后部分结合 ADI 公司提供的 Blackfin 图像和视频开发包(类似于 OpenCV),介绍了运动分析、视频编解码、视频滤波的基本原理,以及在 Blackfin 处理器上的实现方法。

本书共 9 章。第 1 章是绪论,介绍数字视频处理的概念、历史、应用和主要研究内容。第 2 章介绍数字视频处理的基本理论和处理模型。第 3 章介绍 Blackfin DSP 的结构、特性和工作原理,还讨论了 DSP 应用开发的一般过程和集成开发环境 VisualDSP++ 的基本使用。第 4 章是基于 Blackfin 处理器的最小视频系统,实现了视频的输入、输出和数据传输,为更复杂的处理系统的开发打好了基础。第 5 章讨论了视频系统的设计原则,提供了典型的视频开发模板,并用具体例子进行了说明。第 6 章介绍 Blackfin 图像与视频处理开发工具包,它提供了丰富的软件模块,能够显著加速学习和开发过程,文中还介绍了一些典型模块的使用。第 7 章视频运动分析及应用,介绍运动估算、运动分割和运动跟踪的理论和算法,以及基于图像视频开发包的实现方法。第 8 章介绍视频编解码的基本理论和算法,并针对 H.264 基线类介绍了基于 Blackfin 的代码移植和优化策略,并详细介绍了 Blackfin 提供的 H.264 编解码器的使用方法。第 9 章介绍视频时域和空域滤波的基本理论,并介绍了基于运动检测的自适应去交错效应的实现算法。

本书的编写得到了 ADI 公司的大力支持,为本书提供了丰富的文档和网络资源。邢延超主要负责第 1 章至第 3 章及第 7 章至第 9 章的撰写;皇甫伟主要负责第 4 章至第 6 章的撰写。在材料整理、录入和校对方面,王治中、庞秀娟、王清贵等做了大量的工作。另外,本书的完成更离不开家庭作为坚强的后盾。在此一并向他们表示衷心感谢!

由于本人水平有限,书中错误之处在所难免,恳请广大读者批评指正。

编著者

2011 年 4 月

目 录

第 1 章 绪论	1
1.1 数字视频处理概述	1
1.2 数字视频的发展历史	2
1.3 数字视频处理的应用	3
1.4 数字视频处理的研究内容	5
1.5 数字视频处理系统概述	8
1.5.1 视频信号采集 ^[15,22]	8
1.5.2 数字视频编解码 ^[20,21]	10
1.5.3 数字视频存储	10
1.5.4 数字视频显示 ^[15]	11
1.5.5 数字视频处理	13
1.6 嵌入式数字视频处理系统 ^[34,35,36]	13
1.7 研究现状与发展前景	16
第 2 章 数字视频基础	17
2.1 人类视觉机理	17
2.1.1 人眼视觉特性 ^[14,17]	17
2.1.2 人类视觉系统模型	20
2.2 颜色感知与表示模型 ^[51]	22
2.2.1 颜色感知机理	22
2.2.2 颜色模型	23
2.3 视频获取与显示	25
2.3.1 彩色视频成像原理 ^[13]	25
2.3.2 视频摄像机	26
2.3.3 视频显示	26
2.3.4 复合视频与分量视频 ^[51]	27
2.3.5 伽马校正	28
2.4 模拟视频技术 ^[13]	28
2.4.1 模拟视频信号	28
2.4.2 视频光栅扫描	29
2.4.3 模拟电视系统	30
2.5 数字视频技术 ^[15,16,21]	31
2.5.1 模拟视频信号数字化表示	31
2.5.2 数字视频的特点及应用	32
2.5.3 ITU-T BT.601 数字视频标准	33
2.6 视频模型 ^[13,14]	33

2.6.1	照明模型	34
2.6.2	摄像机模型	34
2.6.3	物体模型	37
第 3 章	数字信号处理与嵌入式开发	39
3.1	数字信号处理基础及 DSP 系统应用 ^[4~9,28,37]	39
3.2	Blackfin 处理器简介 ^[28,37~39]	42
3.3	Blackfin 处理器架构	46
3.3.1	Blackfin 处理器架构概述	46
3.3.2	Blackfin 处理器内核基础知识	49
3.3.3	数据运算指令简介	51
3.3.4	地址运算指令简介	54
3.3.5	Blackfin 内存结构	55
3.3.6	事件处理	57
3.3.7	DMA 控制器	59
3.3.8	系统接口	63
3.4	ADSP 开发过程	67
3.5	集成开发套件 VisualDSP++ 简介 ^[31,32]	74
3.5.1	开发工具及其特点	74
3.5.2	利用 IDDE 进行 DSP 程序开发	76
3.5.3	调试工具	80
第 4 章	基于 Blackfin 处理器的最小视频系统	87
4.1	数字视频处理系统构成	87
4.2	Blackfin 处理器与评估板简介 ^[38,39]	90
4.2.1	ADSP-BF533: 高性能的通用 Blackfin 处理器	91
4.2.2	ADSP-BF561: 用于消费者多媒体的 Blackfin 对称多核处理器	92
4.2.3	EZ-KIT Lite for ADSP-BF533	93
4.2.4	EZ-KIT Lite for ADSP-BF561	94
4.3	Blackfin 处理器与视频外设之间的连接	94
4.3.1	Blackfin 处理器上的视频接口——PPI	95
4.3.2	将 Blackfin 处理器连接至视频源 ^[50]	96
4.3.3	连接至显示设备 ^[49]	97
4.3.4	连接视频源和显示设备的原则和技巧	98
4.4	数字视频信号标准简介 ^[19~21]	99
4.5	基于 ADSP-BF561 的视频采集	100
4.5.1	Blackfin 系统服务 ^[39]	101
4.5.2	Blackfin 设备驱动模型	103
4.5.3	视频采集硬件组成	105
4.5.4	视频输入数据流	106
4.5.5	视频输入实现过程	107

4.6	基于 Blackfin 处理器的视频输出	112
4.6.1	视频输出数据流	112
4.6.2	视频显示实现过程	112
4.6.3	基于 Blackfin 处理器的视频传输	115
4.7	基于 ADSP-BF533 的视频采集与显示	116
4.7.1	硬件平台初始化部分	116
4.7.2	初始化中断服务	117
4.7.3	初始化 DMA	117
4.7.4	初始化 PPI	118
4.8	视频采集回放及编码系统的实现	118
4.9	视频 Sobel 边缘提取系统	124
第 5 章	视频应用设计原则及基础应用简介	133
5.1	视频应用开发模板	133
5.1.1	视频开发模板综述	133
5.1.2	视频开发模板类型	134
5.1.3	针对 Blackfin 处理器的优化	136
5.1.4	使用视频开发模板	136
5.1.5	视频开发模板应用举例	137
5.1.6	视频开发模板组合使用	138
5.2	Blackfin 处理器视频处理框架	138
5.2.1	内存使用原则	139
5.2.2	PPI 采集和显示的 DMA 模式	141
5.3	视频基础应用举例	145
5.3.1	解交错	145
5.3.2	解交错扫描速率转换	145
5.3.3	像素处理	146
5.3.4	色度再采样和颜色转换	147
5.3.5	缩放和裁切	148
5.3.6	显示处理	149
第 6 章	图像与视频处理软件开发包	151
6.1	Blackfin 软件开发包介绍	151
6.1.1	SDK 的安装与使用	152
6.1.2	SDK 中的应用简介	152
6.1.3	受限的软件	153
6.2	图形和视频处理软件开发包介绍	154
6.2.1	图像处理开发包	154
6.2.2	视频处理开发包	155
6.3	Hough 变换及其实现	155
6.3.1	Hough 变换基本原理 ^[10,12]	155

6.3.2	图像处理开发包中的 Hough 变换函数	157
6.3.3	基于图像处理开发包的实现	159
6.4	腐蚀与膨胀运算的实现	160
6.4.1	形态学基本知识 ^[10,12]	160
6.4.2	腐蚀与膨胀的开发包实现	165
6.5	人脸检测	165
6.5.1	基于 Adaboost 学习的人脸检测 ^[55]	166
6.5.2	基于图像处理开发包的人脸检测实现	170
6.5.3	人脸跟踪算法的设计	174
6.6	图像处理软件包的内存使用	175
6.6.1	内存移动流程	176
6.6.2	一维内存移动 API	177
6.6.3	二维内存移动 API	179
6.6.4	使用乒乓缓冲区进行内存移动	181
第 7 章	视频运动分析及应用	184
7.1	运动估算	184
7.1.1	基于帧差的运动分析	185
7.1.2	基于块的二维运动分析	186
7.1.3	基于光流场的二维运动分析	189
7.1.4	基于像素递归的二维运动分析	191
7.2	运动分割	192
7.2.1	基于背景差分的方法	193
7.2.2	背景图像更新	194
7.2.3	帧间差分方法	198
7.2.4	目标检测	199
7.2.5	基于光流的方法	199
7.3	运动目标跟踪	201
7.3.1	基于特征的跟踪方法	203
7.3.2	基于变形模型的跟踪方法	204
7.3.3	基于区域的跟踪方法	205
7.3.4	卡尔曼 (Kalman) 滤波器	206
7.3.5	粒子滤波器	208
7.4	光流计算的实现	209
7.4.1	Lucas-Kanade 算法	210
7.4.2	块匹配算法	211
7.4.3	金字塔型光流	212
7.5	前景目标检测的实现	213
7.5.1	初始化对象检测库	213
7.5.2	基于视频开发包的前景对象检测的实现	216

7.5.3	前景对象检测中的基础算法	217
7.6	Kalman 滤波器的实现	219
7.6.1	开发包中的 Kalman 滤波器 API	219
7.6.2	基于 API 的 Kalman 滤波器实现过程	222
7.7	视频交通流检测系统设计	222
7.7.1	硬件平台	223
7.7.2	软件设计和实现	223
第 8 章	视频编解码理论及实现	228
8.1	视频编码基本理论与技术	228
8.1.1	信源编码的信息论基础 ^[1~3]	229
8.1.2	无损压缩	229
8.1.3	变换编码	231
8.1.4	预测编码	233
8.2	视频编码国际标准	234
8.2.1	H.261 视频编码标准	235
8.2.2	H.263 视频编码标准	237
8.2.3	H.264 视频编码标准	240
8.2.4	其他视频编码标准	242
8.3	基于 Blackfin 的 H.264 视频编解码系统设计	243
8.4	ADI 提供的 H.264 视频编码实现	248
8.4.1	H.264 基线编码器概述	248
8.4.2	H.264 基线编码器库的使用	250
8.4.3	H.264 基线编码器 API 介绍	255
第 9 章	视频时空滤波及实现	260
9.1	视频时空滤波技术	261
9.1.1	运动轨迹模型	261
9.1.2	运动补偿滤波	263
9.1.3	运动自适应滤波	266
9.1.4	运动补偿上行变换	266
9.2	基于运动检测的自适应去交错	268
9.3	视频滤波中的二维卷积运算	273
参考文献		276

第1章 绪 论

1.1 数字视频处理概述

众所周知,视觉是人类感知信息最重要的途径,人类从外部世界获取的信息 70%~80% 是通过视觉获取的。随着电子信息技术的发展,人们已经能够利用各种电子设备完成视频信息的采集、编解码、存储、传输和处理等操作,为高效地分析和处理客观世界提供了丰富的手段。其中模拟视频信号已经应用了几十年,至今仍在使用,最原始也是最常见的通用视频标准包括 NTSC、PAL 和 SECAM 等。其他的现代消费模拟视频传输系统包括 S-Video、分量视频等系统。现代视频系统通常在前端和后端分别采用模数转换和数模转换技术,在系统内部采用数字信号处理技术来对视频内容进行处理,大大提高了系统的灵活性,并具有相应的低成本及集成特性,因而数字化产品更吸引消费者。随着电子、通信、计算机等技术的飞速发展,数字视频产品性能不断提高,同时价格也能够为大众所接受,因而具有广大的潜在市场。其中,基于数字信号处理器(DSP)的嵌入式多媒体系统由于具有体积小、功耗小、成本低等优点,使得该类产品越来越实用。从事数字视频处理系统和产品的研发人员不仅需要掌握数字视频处理的基本理论、方法和技术,同时还要了解嵌入式系统软件开发平台的相关知识。本书正是从这两个方面出发,本着实用化的思想来介绍基本开发过程和技术,希望能够为读者尽快掌握基于 DSP 平台开发视频处理系统和产品提供帮助。

视频信息分为图像和视频两大类,前者是静态的,后者是动态的。与静态的图像相比,视频信息无论是在存储空间需求、传输带宽需求,还是在处理的实时性上,都对整个处理系统提出了更高的要求。在数字信号处理应用的早期,由于存储、带宽及处理器处理能力的限制,人们在谈到数字信号处理时,更多关注的是语音和图像信号的处理。随着电子、通信、计算机等相关技术的飞速发展,上述三方面的限制在很大程度上得到了解决或改善,又由于视频信号处理对社会和经济发展具有重大的推动作用,人们开始积极地展开了对视频信号处理的研究与应用。

视频信号本身是模拟的,传统的广播电视信号就是一种典型的模拟视频信号,它由摄像机通过电子扫描将随时间和空间变化的景物进行光电转换后,得到一维的时间函数的电信号,其电平的高低反映了景物的色彩值。模拟视频信号在传输、存储、处理和交互操作等方面具有很大的局限性。为此,可以将视频信号数字化,得到数字视频信号。数字视频信号便于传输,便于多媒体通信,便于存储、处理和加密,无噪声累积,便于设备的小型化。随着数字电路和微电子技术的进步,特别是超大规模集成电路的快速发展,使得数字视频信号的优点越来越突出,应用越来越广泛。例如,高清晰度电视(HDTV)、多媒体、

视频会议、移动视频、监控系统、医疗设备、航空航天、教育、电影等。

数字视频信号处理系统主要包括视频信号的采集、数字化、视频编解码、存储、处理、传输、回放等主要模块。每项功能都由相应的设备来完成,如摄像机负责视频信号的采集,模数转换器负责信号的数字化,编解码完成视频编解码等。广义地讲,上述内容都属于数字视频处理的范畴。与之相对应,狭义的数字视频信号处理主要指,对已数字化的视频信号进行某种特殊功能的分析和加工,如帧间滤波、视频压缩、图像增强、运动目标跟踪及动态场景分析等。但是与语音信号和图像信号的数字化处理相比,数字视频信号数据量巨大,而且对处理的实时性要求高。通用微处理器在实现数字信号处理上效率不高,一般难以满足数字视频处理的实时性需求,因此产生了专门针对数字信号处理的微处理器 DSP (Digital Signal Processor, 数字信号处理器)。

本书将介绍数字视频处理系统的各个环节的理论和系统实现,重点介绍基于 DSP 的各种算法的实现。

1.2 数字视频的发展历史

数字视频的发展历史与计算机的发展史、与计算机能够处理的信息类型密切相关。自 20 世纪 40 年代计算机诞生以来,计算机大约经历了以下几个发展阶段:数值计算阶段和多媒体阶段。

数值计算阶段是计算机发展的初级阶段,这个时期它只能处理数值数据,主要用于解决科学计算与工程计算中的数值计算问题。例如,世界上第一台电子计算机 ENIAC 就是为美国国防部解决弹道计算问题、编制射击表而产生的。

接下来是多媒体阶段。随着电子技术的发展,各种支持图形、图像和语音的设备的问世,计算机逐渐进入多媒体时代,处理的信息载体扩展到文本、图形、图像、声音等多种类型,这不仅扩大了计算机的应用领域,同时也极大地改善了人机接口的友好性。由于视觉信息(如图形、图像)能够直观明了、生动形象地传达有关对象的信息,因而在多媒体计算机中占有重要的地位。

进入多媒体阶段后,基于计算机系统的视频信息的获取、处理、存储、显示等得到了广泛认可和应用。由于普遍使用的是数字计算机,其内核只能处理数字信号,所以我们研究的是数字视频处理技术。数字视频的发展大致可以分为初级阶段、主流阶段和高级阶段。

初级阶段主要是在微型计算机上增加简单的视频功能,利用计算机来处理活动画面。但此时由于设备尚未普及,一般都是面向视频制作领域的专业人员,普通 PC 用户还无法奢望在自己的计算机上实现视频功能。

在初级阶段,数字视频的发展比较缓慢,其主要原因是数字视频的处理、存储和传输能力都不足,无法胜任所需要的处理、存储和传输能力。例如,一分钟的真彩色数字视频需要 1.5GB 的存储空间,而早期一般台式机配备的硬盘容量大约是几百兆,显然无法胜任如此大的数据量。

尽管存在这些限制,人们还是采用一些折中方法来对数字视频进行处理。比如用计算机捕获单幅视频图像,将其以一定的文件格式存储起来,利用图像处理软件进行处理,将结果保存下来用于需要的各种场合。后来,又实现了在计算机上显示活动的视频,虽然画面时断时续而且持续时间很短,但毕竟使图像动了起来,给人们带来了崭新的体验。

后来,随着计算机软硬件性能的不断提高,以及视频采集设备、大容量存储设备、视频显示设备等不断升级,最终使得视频捕获、存储、播放在个人台式机上成为可能。由此进入了数字视频处理的主流阶段,即模拟视频不再是视频处理的主流。这其中非常关键的是压缩解压缩(Codec)技术的成熟,压缩可以极大地降低数据量(达数十上百倍)。压缩的实现又分为纯软件压缩和硬件辅助压缩两种,前者方便易行,成本低,但是速度较慢,难以满足实时性的要求;后者速度快,但是成本高。如果尺寸较小、帧速较慢,通过视频压缩后1分钟的视频数据只需20MB而不再是1.5GB,压缩比高达1:75。这样就能够将视频文件保存在硬盘上,也可以进行动态的显示,虽然分辨率仅为 160×120 ,帧速率只有15帧/秒,色彩也只有256色,但画面毕竟活动起来了。这样数字视频处理就不再是专业人员的特权,而成为个人计算机的必备功能。

高级阶段又称成熟的多媒体计算机阶段,各种计算机外设产品日益齐备,数字影像设备争奇斗艳,视音频处理硬件与软件技术高度发达,这些都为数字视频的流行起到了推动的作用。在这个阶段,数字视频被进一步地标准化。

在数字视频的编解码方面。首先,为了PAL、NTSC和SECAM电视制式之间确定共同的数字化参数,国际无线电咨询委员会(CCIR)制定了广播级质量的数字电视标准,称为CCIR 601标准。在该标准中,对采样速率进展迅速,包括图像尺寸、帧率、采样结构、颜色空间选择等都做了严格的规定。这种未压缩的数字视频数据量对目前的计算机和网络来说是无法实现的,因此需要对其进行数字压缩。这就要求对压缩算法和存储格式进行标准化,现在普遍采用的压缩标准包括MPEG系列的MPEG-1、MPEG-2和MPEG-4,以及H系列的H.261、H.263和H.264等。这些标准的确定使不同系统之间可以交换和共享数字视频内容。

视频内容数字化以后,就可以采用(二维)数字信号处理技术进行灵活地处理,如进行视频滤波、图像增强、图像缩放、运动估计、目标检测与跟踪等。数字视频处理的应用不断丰富,如智能视频监控、视频增强、视频滤波等。当前数字视频处理的理论和技术成为研究的前沿和热点,同时必将不断拓宽视频处理应用的范围。

1.3 数字视频处理的应用

数字视频在监控、无线视频网关和数码相机等方面有广泛的应用,这些应用的共同点是它们都要求视频和图像的处理。这些处理可能是视频图像质量的优化,如锐化和白平衡,也可能是视频和图像的压缩/解压缩,如MPEG4或H.264等。主要的视频应用包括以下几类。

1. 数字电视机顶盒

机顶盒负责在用户端同步解码、处理和播放视频、音频和数据流,有些机顶盒还支持对 DVB 或 DVD 其他 MPEG 数据流的编码功能。这些机顶盒一般功能比较简单,主要是负责接收数据流然后解码或编码并进行显示。但是由于是进行流处理,并且多半是 MPEG2 质量的视频流,所以需要处理能力较强。

2. 数字视频监控系统

数字视频监控系统相对传统视频监控系统来说是一个显著的飞跃,它增加了更多的智能特性,如可以进行移动报警,通过视频处理方法在监控区域发现移动物体就报警;数字化的视频档案更加便于存储、管理和远程传输,如可以通过自动视频分析方法来找到目的片段,数字化信息通过网络更便于传输。数字视频监控系统最主要的要求是可编程性,数字处理能力也是实现这些特点的保证。另外,视频输入和网络接口也不可或缺。

3. IP 视频电话/视频会议

IP 视频电话是将实时的多通道视频集成到现有的 IP 语音电话中,视频会议则是在多点间实现视频电话会议。目前,这两种应用多采用 H.263 或 MPEG4,芯片要完成这两种制式的编解码,同时完成图像采集处理的任务,单芯片处理能力要求较强。由于要进行视频采集,因此接口要求支持视频输入,要实现 TCP/IP 协议则要求有以太网接口的输出支持。

4. 个人数字视频播放器/点播机

其主要任务是完成视频解码,前者更强调便携性,而后者则强调丰富的节目源。在该应用中,处理性能并不是第一位要考虑的因素,而整个系统的低能耗十分关键。另外,要求芯片的接口比较丰富以支持 LCD 显示屏、CFC 或其他存储卡或硬盘接口。

5. 无线视频网关

无线视频网关位于 MSC 内,用于将任何制式的视频内容转成手机支持的格式,从而将这些内容在无线网络上发送。由于无线网络对环境的依赖比较强,不同连接的信道质量会有所不同,因此要求网关针对信道质量,将视频内容转换到该信道所支持的码率上。由于网关要与其他网络接口,要求芯片的外围接口支持如 ATM 或以太网等网络总线。

除此以外,还有各种丰富的数字视频处理系统,如数字摄录机、数字照相机等。这些应用总体上分为两大类:一类是高性能,要求多通道处理或多编码制式,或追求高图像质量,同时由于技术的不断发展,它又需要较强的可编程性支持未来的技术发展;另一类是便携应用,强调小型化、低功耗,并能支持多种移动存储接口及 USB、1394 或其他高速总线。

1.4 数字视频处理的研究内容

尽管数字视频系统多种多样,处理目的各不相同,但视频处理的基本内容和相关技术可以归纳为以下几个方面^[13~17,18~21,23~25]。

1. 视频信号表示与标准化

模拟电视系统通常用光栅扫描方式,即在一定的时间间隔内电子束以从左到右、从上到下的方式扫描采集荧光屏表面。扫描方式通常有逐行扫描和隔行扫描两种。经行场扫描得到的一维连续时间信号包括图像信号、行同步信号、行消隐信号、场同步信号、场消隐信号和前后均衡脉冲等。视频信号的基本参数主要有清晰度、分解力、宽高比、行频、场频和帧频等。模拟电视系统标准主要有 NTSC 制、PAL 制和 SECAM 制,它们对上述各种参数作出了不同的标准化,方便了视频设备的兼容和视频内容的交互。

模拟视频信号经过数字化处理后,就变成由一帧帧数字图像组成的图像序列,即数字视频信号。每帧图像由 N 行、每行 M 个像素组成,每个像素用 N_b 比特表示。为了便于国际节目交换及 625 行 PAL 制系统与 NTSC 制系统之间的兼容,1982 年 CCIR (国际无线电通信咨询委员会) 制定了 CCIR 601 数字视频标准,1993 年变更为国际电信联盟无线通信部门 ITU-T BT.601 建议,定义了对应于 525 行和 625 行电视演播室的数字编码参数。三个分量信号 Y 、 $R-Y$ 、 $B-Y$ 的抽样频率分别为 13.5MHz、6.75MHz 和 6.75MHz,三个分量在一行中抽样点数的比例为 4:2:2,且每帧的行数相同,故称为 4:2:2 标准,此外还有图像质量更高的 4:4:4 和较低图像质量的 4:1:1 标准或 4:2:0 标准。

2. 二维运动估计

建立在数字图像处理技术的基础上,数字视频处理主要关注对视觉信息动态部分的分析、处理和利用。运动分析广泛应用于计算机视觉、目标跟踪、视频监视、视频压缩等应用场合。不同应用场合对运动估计的要求不同,有时强调估计结果与实际运动尽可能一致,称为真实运动分析;有时则强调获取更高压缩率,不要求估计结果与实际运动的对应关系。

运动估计处理的对象是视频图像序列,是三维场景在图像平面上的投影。在不同时刻拍摄到的投影图像上,运动物体的投影坐标是不同的。运动分析与估计的研究内容就是测量计算投影坐标在图像平面上的变化,并用来分析运动物体的结构,估计物体的运动参数。由于摄像机的投影过程是不可逆的,投影过程不可避免地要丢失一些信息,因此估计物体的真实运动和结构是非常困难的。

二维运动估计的基本问题是估计运动前后相邻时刻两幅图像上对应点的坐标,即二维运动矢量。三维运动估计的基本问题是利用一组对应点坐标估计和确定运动物体的结构和运动参数。通常运动估计算法均假设物体点的亮度(或颜色)在其运动轨迹上保持不变。目前,运动分析方法主要有两种:一是,根据时间相邻的两幅或多幅图像求解物体的运动

参数和三维结构信息；二是，图像序列的光流分析方法。

二维运动估计方法主要有以下几类：

(1) 基于光流的运动估计，利用光流场和各种约束条件求解各个位置上的运动矢量。

(2) 基于像素的运动估计，包括位移帧差、多点邻域约束、像素递归法、基于贝叶斯准则的方法等。

(3) 基于块的运动估计，它假定每个块只做二维平移运动，已成功应用于各种视频压缩编码标准。

(4) 基于网格的运动估计，利用网格的形变体现物体的形状变化。

(5) 基于区域的运动估计，将视频图像分割为多个区域，每个区域对应一个特定的运动，然后可以为每个区域估计运动参数。

(6) 多分辨率运动估计，利用视频图像的多分辨率原理，首先在最小分辨率层进行运动估计，然后逐层进行，最后得到最高分辨率的运动场。

3. 三维运动估计

三维运动估计广泛应用于机器视觉、视频监控、交通控制等场合。三维运动估计的基本任务是分析和估计三维场景中物体运动的情况。由于物体的结构与投影结果密切相关，所以同时需要估计三维物体的结构参数。由于很难精确估计出物体的运动结构，因此需要作出一些合理的假设来简化运动模型，如刚体运动限制等。三维运动和结构的估计方法可以分为两类：一是，间接估计法，根据已经给出的二维运动矢量（如特征点对、光流场）来估计二维运动和结构参数；二是，直接估计法，根据视频图像的空时亮度信息来估计三维运动和结构。

三维运动估计主要有以下几种方法：

(1) 基于特征对应的运动估计，属于间接运动估计方法，如最小二乘估计方法。

(2) 基于光路的运动估计，也属于间接运动估计方法，又分为正交投影下的运动估计和透视投影下的运动估计，首先估计二维图像的光流场，然后根据光流和光流参数模型估计三维运动参数和结构参数。

(3) 直接运动估计法，如用图像梯度代替光流矢量，然后基于平面运动模型直接估计运动参数的方法等。

4. 运动目标分割

三维场景中通常包含多个运动物体，运动分割就是将视频序列中属于各个不同运动的像素标记出来。运动分割与运动估计密切相关，精确的运动分割可以提高运动估计的准确性，反之亦然。在基于内容的视频描述中，为了得到描述视频的每个对象，需要通过运动分割将运动物体与背景分割开来。运动分割的主要方法包括基于空时图像的直接分割方法、基于光流场参数模型的分割方法及运动分割和运动估计同时进行的方法。

直接分割方法属于分割优先的方法，首先基于图像亮度信息进行区域分割，然后估计每个区域的运动参数。基于光流的分割法属于运动优先的方法，首先估计光流场，然后对

光流场建模并进行运动分割。最常见的光流场参数模型有 6 参数仿射流模型和 8 参数二次流模型。运动分割可以通过将具有相同模型的参数矢量归为一类得到。运动分割和运动估计同时进行的方法, 由于运动分割和运动估计是相互依赖和相互促进的, 因此同时进行运动估计和运动分割可以得到更好的结果。

5. 运动目标跟踪

运动跟踪的基本任务是通过将摄像机拍摄到的图像序列进行分析, 计算出目标物体在各图像帧上的位置, 并给出物体运动的估计。运动跟踪通常基于一组特征对应或光流矢量来描述运动的发展过程, 其中物体特征或光流矢量称为运动跟踪的观察对象。运动跟踪主要包括三个方面的内容: 运动模型、观察模型和跟踪方法。

运动模型可以是简单的匀速运动模型, 也可以是更加复杂的存在旋转、形变等的模型。运动模型又可以分为二维运动模型和三维运动模型, 一般主要研究刚体运动模型。观察模型可以是一组特征对应关系或光流矢量。跟踪方法利用运动模型和观察模型估计最佳的运动模型参数, 分为两类: 一类是分批法, 如非线性最小二乘法, 基本思想是在获取一定的观察数据后对现有全部数据进行处理; 另一类是递归法, 如 Kalman 滤波器等, 只要获取的数据可以更新运动参数, 递归算法就能很好地工作, 通过逐步迭代得到理想的跟踪结果。

6. 视频滤波

在视频采样与数字化过程中, 由于设备噪声、采样方式转换、隔行扫描模式、场景运动等因素往往造成视频质量的下降, 有时则是希望利用现有视频信号重构超分辨率的高清晰度视频信号。这都属于视频滤波研究的内容, 它具体又包括运动补偿滤波、噪声滤波、图像复原、上下行变换、超分辨率等内容。各个研究内容有各种丰富的处理技术, 如噪声滤波就有帧内滤波、运动自适应滤波和运动补偿滤波等不同种类的技术。

7. 视频压缩编码

视频原始数据量巨大, 难以直接进行存储和传输, 因此数据压缩具有特别重要的意义。视频压缩主要采用有损压缩方式, 与静态图像压缩相比, 视频信号压缩可以充分利用帧间相关性去除时间冗余度。因此视频压缩编码又分为帧内压缩和帧间压缩两部分, 帧内压缩与静态图像压缩类似, 主要利用变换编码(如 DCT)来实现。帧间压缩则通过计算图像块的运动矢量, 利用参考帧信息来重构预测帧的信息, 因而能够极大地提高压缩效率。现在通用的 MPEG-2、H.264 等编码标准普遍采用这一思想。

近些年来则提出了基于模型的编码技术, 它通过将场景分割为相互独立的不同目标来进行单独编码, 在显示端重构出整个场景, 不仅能够取得更高的编码效率, 同时也为交互式视频提供了基础。

数字视频处理的内容还有很多, 如三维立体视频、数字视频水印技术等, 这里就不一一介绍了。