

DSP 应用大观

4

张旭东 主编

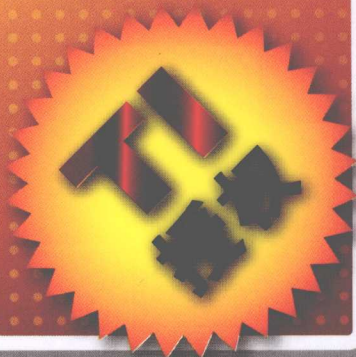
TI DSP

在视频传输和处理中的应用

- 围绕应用，重点突出实用
- 全面覆盖，资料及时、权威
- 专家筛选，内容深入、经典



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



DSP 应用大观

TI DSP在视频传输和处理中的应用

张旭东 主编

電子工業出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书主要介绍 TI DSP 在视频传输和处理中的一些应用实例,分四部分内容。第一部分介绍 DSP 和视频技术的一些基本概念,包括 DSP 系统开发的基本框架和视频编码基础;第二部分讨论 DSP 在 JPEG 和 MPEG 实现中的应用;第三部分是 H.264 在 DSP 上的实现,分别给出在 TMS320C6416 和 DM642 平台上的实现,包括 H.260 在 DSP 上实现的软件结构优化和算法效率优化;第四部分讲解 DSP 在图像增强方面的应用,这部分还介绍了从 MATLAB 算法到 DSP 实现的开发过程。本书可供视频传输和处理的技术人员阅读,也可作为相关专业大学高年级本科生和硕士研究生的参考资料。

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

TI DSP 在视频传输和处理中的应用 / 张旭东主编.—北京:电子工业出版社,2009.11

(DSP 应用大观)

ISBN 978-7-121-09782-9

I. T … II. 张 … III. ① 数字信号—信息处理系统—应用—图像通信—数据传输 ② 数字信号—信息处理系统—应用—视频信号—信号处理 IV. TN919.8

中国版本图书馆 CIP 数据核字(2009)第 197080 号

责任编辑:万子芬(wzf@phei.com.cn)

印 刷:北京市李史山胶印厂

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:16.25 字数:416 千字

印 次:2009 年 11 月第 1 次印刷

印 数:3 500 册 定价:36.00 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 zlts@phei.com.cn,盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010) 88258888。

丛书序言

DSP 技术的发展与应用,正在我国教育界、科技界和工程界蓬勃地展开。数以百计的大学建设了 DSP 技术实验室,开设了相关的课程和实验;大量的相关教材、技术手册和应用书籍得到编写、编译和出版。更为重要的是,基于 DSP 技术的研究和开发,无论是涉及的范围,还是达到的深度,都令人叹为观止。以两年一度的 TI DSP 大赛为例,每次都有数十所大学的上百支代表队参赛,参赛者所表现出来的选题的广度、算法研究的深度,算法实现和系统设计及实现的娴熟程度,无不令人振奋。

随着教学、科研的发展和深入,教师、学生、以及科研和工程技术人员已经不再满足于对 DSP 的粗浅了解;市场的发育,对 DSP 技术的发展和运用也提出了越来越高的要求。在这样的形势下,编写和出版一套 DSP 应用汇编丛书,就成为一种强烈的需求,并迅速在出版社、TI 公司以及编写者之间达成了共识。

我们也注意到,在全球范围内,随着 DSP 技术应用范围的扩大和应用程度的深入,通用 DSP 器件的增幅在逐步减缓,而基于 DSP 核的各种 SoC、ASSP 以及嵌入式系统,正在以更快的速度发展。对于 DSP 工程师来说,开发算法并将算法在 DSP 芯片或 DSP 核上实现,还将仍然是长期的重要任务。本丛书的编写和出版,正是基于这样的认识和理解。

这套丛书是这样设计的:

按应用领域来分类,先在几个重要的领域,例如,通信信号处理、图像/视频信号处理、音频/语音信号处理、工业控制、通用信号处理算法、DSP 接口与软件工具等,各出一个选题。每个选题以 TI 网站上公开的 Application notes 为基本内容,为了便于读者理解和使用,各书的编译者对所介绍的内容,都不同程度增加了补充性的介绍。

这套丛书是开放的,这里所指的开放,包含以下两重意思:一方面,随着各领域的技术进步,新的算法和新的器件层出不穷,本丛书对新的算法及其实现的介绍也会继续下去;另一方面,欢迎广大的读者对丛书的选题和内容提出意见和建议,更欢迎有志者加入编写者的行列。

本丛书第一批选题的作者,是各高校多年从事 DSP 技术研究和实践的教师,以及他们的一些研究生,他们在各自的领域具有长期的知识积累和丰富的实践经验,为本丛书的选题、编写和出版付出了辛勤的劳动。

TI 公司对本丛书所使用的文档予以了授权,TI(中国)大学计划对丛书的编写和出版给予了一贯的支持和鼓励。电子工业出版社的编辑们,首先提出了本丛书的创意,积极参与了选题策划和论证,认真地完成了编辑和出版工作。在此,对所有为本丛书的选题、编写、出版做出贡献的单位和人士,致以深切的谢意和敬意。

希望这套丛书的出版,能对推动我国 DSP 技术的教育和应用起到微薄的作用,衷心希望得到广大读者的支持、意见和建议。



电子科技大学教授
2008 年 4 月

前 言

本书主要介绍 TI DSP 在视频传输和处理中的一些应用实例,可供工程技术人员和大学高年级学生参考。按照丛书主编彭启琮教授的设想,丛书根据不同的应用领域,以整理和编译 TI 的“Application Notes”为主,这个设想对于基本算法、语音处理等分册是非常适合的。的确, TI 的“Application Notes”中有大量的应用算法和系统的技术报告,但对于视频传输和应用领域,情况有些不同。在准备本书材料的时候,有关 H.264 这样最新的视频编码标准的技术报告很少,而且除了标准的概要介绍外,几乎没有实质性的内容,这是可以理解的,因为 H.264 在 TI DSP 上实现的编解码器有很高的商业价值,没有人会把这样的技术公开在网络资源中。鉴于这种情况,本书做了一些调整,对于 JPEG、MPEG 这些传统方法,我们整理了“Application Notes”中的相关技术报告,而对于 H.264 的相关技术,增加了我们实验室的相关学位论文和技术报告作为主要内容。

本书分为四部分。第一部分包括前两章,第 1 章概要,叙述 DSP 和视频技术的一些基本概念,并介绍 DSP 系统开发的基本框架;第 2 章给出视频编码技术的介绍,由于目前一线工程师大多缺乏视频编码的系统知识,本章概要性的介绍是有用的。

第二部分包括第 3 章、第 4 章,分别介绍 TI DSP 在 JPEG 和 MPEG 实现中的应用,这两章的体系是一样的,首先介绍标准的原理,然后编译了相关标准实现的技术报告。

第三部分由第 5 章、第 6 章构成,介绍 H.264 在 TI DSP 的实现,分别给出了在 TMS320C6416 和 DM642 平台上的实现。第 5 章是一篇完整的技术报告,包括 H.264 的详细介绍, H.264 在 DSP 上实现的软件结构优化和算法效率优化两个方面的内容,实际上,一个复杂算法在 DSP 上的实现都会存在软件优化和算法优化两个部分,缺一不可很难达到好的性能,尽管本章介绍的是一个阶段性的研究结果,但其包含的内容对进入该领域的人很有参考价值;第 6 章则介绍了 H.264 在 TMS 320DM642 上的实现。

第四部分由最后两章构成,介绍 TI DSP 在图像增强方面的应用,在低照度的视频监控环境下,图像增强是很实用的工具,这部分还介绍了从 MATLAB 算法到 DSP 实现的开发过程。

本书是一本应用汇编性的图书,并不是一本完整的教材,本书实际是实验室集体工作的结果,本书主编只是将多人的工作汇编成册,并进行了整理和统稿,第 3 章、第 4 章的标准实现源自 TI 的“Application Notes”,第 5 章的主要内容来自魏振宇的硕士学位论文,第 6 章的主要内容来自范嘉略的技术报告,第 7 章、第 8 章的主要内容来自张众的技术报告,在此,谨向他们表示感谢,同时,感谢彭启琮教授的帮助和指导,感谢 TI 大学计划部沈洁经理长期对实验室的支持。

目 录

第 1 章 技术基础概要	(1)
1.1 数字视频编码标准的演进	(1)
1.2 数字视频编码国际标准概述	(3)
1.2.1 国际电信联盟 (ITU-T) 视频标准 H 系列	(3)
1.2.2 MPEG 系列视频标准	(4)
1.3 DSP 系统开发的基本流程	(6)
1.3.1 DSP 的发展及特点	(6)
1.3.2 DSP 系统的设计与开发	(8)
1.4 视频处理算法开发平台	(11)
1.4.1 DSP 程序开发的基本流程	(11)
1.4.2 DM642 开发平台	(15)
1.4.3 XDS560 JTAG 仿真器	(17)
1.4.4 DSP/BIOS 实时内核	(21)
1.4.5 CCS (Code Composer Studio)	(25)
1.4.6 软/硬件接口	(26)
1.4.7 一个示例程序	(27)
参考文献	(31)
第 2 章 视频图像压缩编码基础	(32)
2.1 数字图像编码概述	(32)
2.2 图像的表达和编码质量的评价	(33)
2.2.1 静止图像格式	(33)
2.2.2 视频序列的常用格式	(35)
2.2.3 编码质量的评价	(37)
2.3 信息理论基础和熵编码	(38)
2.3.1 离散信源的熵表示	(38)
2.3.2 信源编码定理	(41)
2.3.3 Huffman 编码	(43)
2.3.4 算术编码	(45)
2.3.5 行程编码	(47)
2.3.6 有记忆信源的编码问题	(48)

2.4	量化	(49)
2.4.1	率失真函数	(49)
2.4.2	标量量化	(51)
2.5	预测编码	(55)
2.6	变换编码	(59)
2.6.1	一般图像变换	(59)
2.6.2	DCT 变换	(64)
2.6.3	变换编码	(66)
2.6.4	基于 HVS 的量化与码率分配	(68)
2.6.5	量化系数的扫描和表示方法	(70)
2.6.6	一个编码实例	(72)
2.7	块匹配运动估计与补偿	(73)
2.7.1	运动矢量的快速搜索算法	(75)
2.7.2	变块大小的分层运动估计	(79)
2.7.3	分数像素运动估计	(84)
2.7.4	重叠运动补偿预测 (OMCP)	(87)
2.7.5	双向预测	(88)
2.8	序列图像编码算法	(89)
2.9	各种图像压缩标准的应用目标和主要技术	(91)
	参考文献	(94)
第 3 章	TMS320C6000 实现 JPEG 编解码器	(95)
3.1	JPEG 编码标准	(95)
3.1.1	JPEG 标准的工作模式	(95)
3.1.2	基本工作模式	(96)
3.1.3	其他工作模式	(101)
3.2	JPEG 在 C6000 上的实现	(104)
3.2.1	JPEG 编码器	(105)
3.2.2	JPEG 解码器	(111)
	参考文献	(115)
第 4 章	MPEG 编码标准及其在 DSP 上的实现	(116)
4.1	MPEG-1 视频压缩标准	(116)
4.1.1	SIF 格式	(117)
4.1.2	MPEG-1 视频编码	(118)
4.1.3	MPEG-1 视频解码	(124)
4.1.4	MPEG-1 的其他问题	(125)
4.2	MPEG-2	(125)
4.2.1	MPEG-2 的运动估计	(126)

4.2.2	MPEG-2 的变换和扫描	(127)
4.2.3	MPEG-2 的可分级编码模式	(128)
4.2.4	MPEG-2 分档和分层	(129)
4.3	MPEG-4	(130)
4.3.1	MPEG-4 的组成	(130)
4.3.2	MPEG-4 视频编码原理	(133)
4.3.3	MPEG-4 中视频编码器的实现	(134)
4.3.4	MPEG-4 中的差错控制方法	(138)
4.3.5	MPEG-4 中的解码技术	(140)
4.4	基于 MS320C62x 的 MPEG-2 视频解码器实现	(141)
4.4.1	软件实现概述	(141)
4.4.2	算法描述	(142)
4.4.3	解码器的实现	(143)
4.4.4	与解码器的连接	(144)
4.4.5	程序的运行	(149)
	参考文献	(150)
第 5 章	TMS320C6416 实现 H.264	(151)
5.1	H.264 概述	(151)
5.2	H.264 视频编解码器	(152)
5.3	H.264 的结构框架	(154)
5.3.1	H.264 的档和层	(154)
5.3.2	H.264 支持的视频格式	(156)
5.3.3	H.264 的码流格式	(156)
5.3.4	H.264 的帧结构	(157)
5.4	H.264 具体技术概述	(158)
5.4.1	帧内预测编码	(159)
5.4.2	运动估计	(160)
5.4.3	整数 DCT 变换	(164)
5.4.4	熵编码	(165)
5.5	实现 H.264 编解码的 TMS320C6416 平台	(167)
5.5.1	TMS320C6416 简介	(167)
5.5.2	CPU 的技术特点	(169)
5.5.3	NVDK (Network Video Development Kit) 简介	(171)
5.6	H.264 在 NVDK 上的实现与优化	(175)
5.6.1	算法选择	(176)
5.6.2	编码器代码移植	(179)
5.6.3	代码优化	(180)
5.6.4	程序优化结果	(185)

5.7 算法优化	(185)
5.7.1 快速整像素运动估计算法——ARPS-4	(186)
5.7.2 基于早停止技术的亚像素运动估计快速算法	(188)
5.7.3 快速运动估计算法实验结果与分析	(190)
5.7.4 快速模式选择算法	(191)
参考文献	(198)
第 6 章 H.264 编码器在 TMS320DM642 上的实现和优化	(199)
6.1 TMS320DM642 EVM 介绍	(199)
6.1.1 DM642 的缓存结构	(200)
6.1.2 TMS320DM642 的视频接口	(201)
6.2 DSP 平台的程序开发问题	(202)
6.3 编码器实现	(203)
6.3.1 算法基本流程	(203)
6.3.2 代码移植	(204)
6.4 代码优化	(207)
6.4.1 项目级优化	(207)
6.4.2 指令级优化	(208)
6.4.3 缓存优化	(211)
6.4.4 优化结果	(214)
6.5 程序示例	(214)
参考文献	(221)
第 7 章 使用 CCS 开发视频图像增强算法	(222)
7.1 直方图均衡化的基本原理	(222)
7.2 实现代码	(224)
7.3 调试	(225)
7.3.1 DSP/BIOS 错误调试	(225)
7.3.2 使用 LOG 模块输出信息	(227)
7.4 算法性能优化	(228)
7.4.1 如何评估一个 DSP 算法的性能	(228)
7.4.2 程序优化	(231)
参考文献	(233)
第 8 章 使用 MATLAB 开发 DSP 的图像处理算法	(234)
8.1 MATLAB LINK FOR CODE COMPOSER STUDIO	(234)
8.1.1 背景介绍	(234)
8.1.2 安装配置	(235)

8.2 示例程序	(235)
8.3 使用 EMBEDDED MATLAB 构造 SIMULINK 模块	(240)
8.3.1 Embedded MATLAB 简介	(242)
8.3.2 如何使用 Embedded MATLAB 开发 Simulink Blocks	(243)
8.4 使用 MATLAB 开发的视频图像增强算法	(244)
参考文献	(246)

第1章 技术基础概要

随着通信技术和信号处理技术的发展，人们对多媒体信号的需求越来越多，要求的质量也越来越高，如何在现有的技术水平和硬件条件下，实现合理、优化、实时的多媒体通信终端设备和多媒体信息存储设备一直是近年来信号处理领域和相关产业界关注的话题。

我们知道，多媒体通信终端平台的实现主要有两点。第一，需要有快速、稳定的处理器作为多媒体处理的硬件平台；第二，需要有适合网络或者无线通信的视频通信协议，两者的结合才能产生高效的多媒体通信设备。多媒体存储设备则要求高质量的压缩编码和有效的检索、浏览功能。

多媒体信号处理主要针对音频、视频信号，本系列丛书中已有专门讨论音频处理的分册，所以本书重点介绍视频图像处理的实现技术。

目前，随着数字信号处理器（DSP）的高速发展，实现高效的视频图像处理有了可能，尤其是 TI（Texas Instruments）公司的 TMS320C64 系列产品及后续的 DM64X 系列，Davinci 系列和 OMAP 系列产品等，具有高主频、多流水线、高并行度及专用的视频信号处理指令和接口等优点，使其成为视频处理领域性能优异的 DSP 芯片类型之一。

针对当前网络带宽还不够，无线通信信道误码率较高的情况，稳定的低码率视频信号的需求一直都是必要的，而视频存储设备同样要求较高压缩率和方便的浏览功能。因此，各种视频压缩编码标准被制定并侧重于不同的应用领域，其中，H.264/AVC 是 ITU-T 视频编码专家组和 ISO/IEC 运动图像专家组联合提出的最新一代的视频编码标准。目前，在工程实际中，针对不同需要，有多种视频编码标准同时存在，可以根据需要选择合适的标准。

本书讨论用 DSP 实现数字视频的编码、传输和处理。本章首先概要介绍关于视频图像编码标准的发展，以及利用 DSP 实现视频处理的一些基本的流程，概要介绍 TI DSP 开发中的基本工具 CCS，并以 DM642 为参考平台，给出了一个视频图像采集和显示的例子。关于视频编码的关键技术在第 2 章给出更加详细的讨论，后续章节则介绍在几种不同 DSP 平台上实现典型视频处理算法的一些关键技术。关于各种 DSP 的 CPU 和指令集的详细叙述，请参考相关文档。

1.1 数字视频编码标准的演进

随着数字视频处理技术的快速发展，各种数字视频应用已经广泛深入到我们的日常生活中，如数字电视、视频电话、视频会议和视频监控，而第三代移动通信（3G）更是把无线视频传输作为其一大特点，这些快速发展的视频应用无疑推动了数字视频编码标准的制定。事实上，从 20 世纪 80 年代开始，ISO/IEC 和 ITU-T 这两大国际组织就已经不断推出一系列针对不同应用领域的数字视频编码标准，这其中包括 ISO/IEC 的 MPEG 系列和 ITU-T 的 H.26X

系列, 图 1.1 是数字视频编码标准的发展过程。

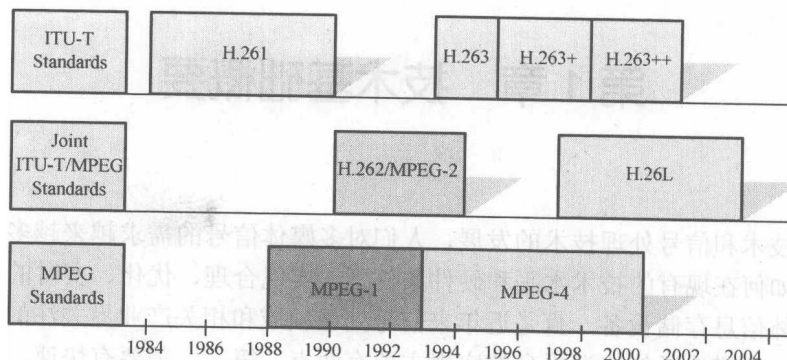


图 1.1 数字视频编码标准的发展过程

从图中可以看出, ITU-T 和 ISO/IEC 基本上是在相互独立的情况下开发各自的系列标准, 其应用领域是不同的。MPEG 系列主要应用于数字娱乐 (如 DVD、数字电视) 或视频监控, 而 H.26X 系列主要应用于实时的视频通信, 包括视频电话、视频会议, 等等。

虽然这些视频编码标准所采用的具体技术细节不同, 但是它们都遵循运动估计、运动补偿、变换编码、熵编码这样的编码框架, 之所以采用这样的模式, 是为了最大限度地去除视频序列中的冗余信息, 保持适度的复杂性。

通常, 视频序列中主要包括三种冗余信息: 空间冗余、时间冗余和统计冗余。

空间冗余是指视频信号中同一幅图像内相邻或相近像素之间具有的相关性。变换编码的目的就是去除这种冗余度。一般来说, 所选择的变换应具有较好的能量集中特性, 使得变换之后图像的能量集中在少数几个系数中。KLT 变换是目前能量集中性最好的一种变换, 但是由于 KLT 变换的矩阵是不固定的, 计算复杂度较大, 因此, 目前大多数视频编码器中采用的是 DCT 变换。与 KLT 变换相比, DCT 变换的能量集中性能非常接近, 但是计算简便, 有快速算法, 因此得到了广泛的应用。空间冗余压缩技术也是静态图像编码 (如 JPEG) 采用的主要技术。

时间冗余是指视频信号中不同帧图像的像素之间具有的相关性, 特别是在运动幅度比较小的序列中, 这种相关性更加明显。运动补偿用来消除这种时间冗余度, 它基于差分编码的思想, 只对当前信号与参考信号之间的差值进行编码。即使一帧图像中原始像素的能量很高, 经过运动补偿后得到的残差信号的能量也可以显著降低, 有利于降低码率。

针对时间冗余的消除, 目前最常采用的是运动估计和补偿技术 (具体技术细节参考第 2 章), 运动估计的目的就是为当前信号寻找最合适的参考信号。运动估计可以采用多种方式, 包括基于块匹配的方法, 基于光流的方法, 等等。基于块匹配方法具有较好的准确性, 并且运算量相对其他方法较低, 因此在现代视频编码器中通常都采用这种方法。为了进行块匹配, 一幅图像通常被分为若干个大小相同的单元, 称为宏块 (宏块的大小通常为 16×16), 运动估计过程对每个宏块单独进行。

视频序列中的第三种冗余——统计冗余, 是指码流中各符号间的相关性。为了消除这种冗余, 一般采用熵编码的方法, 最有名的熵编码方法是 Huffman 编码和算术编码, 曾经被广泛地应用在各种图像和视频的编码算法中。后来人们对熵编码的算法进行了不断的完善, 但

基本的思想并没有很大的变化。

除了上述三个主要部分,一般实用的编码器中通常还包含编码控制模块及码率控制模块,限于篇幅这里就不再赘述了。

1.2 数字视频编码国际标准概述

1.2.1 国际电信联盟 (ITU-T) 视频标准 H 系列

1. H.261

H.261 是最早出现的视频编码标准,在 1988—1993 年之间完成了协议的制定和修订。它首次采用了运动补偿预测编码加 DCT 变换的方式,奠定了以后视频编码技术的基础。其输出码率是 64 kb/s 的整数倍 (1~31)。H.261 主要是针对 ISDN 的会议电视和可视电话等应用制定的,通过缓冲器控制产生恒定的输出码率。H.261 通过均衡图像质量和运动来优化带宽,所以,当图像快速运动时,质量会下降。

H.261 支持 CIF (352×288) 和 QCIF (176×144) 两种视频图像格式,采用简单的渐进扫描模式,帧率为 30 Hz。在 H.261 协议中,只允许使用 I 帧 (Intra Frame) 和 P 帧 (Predict Frame) 两种帧模式,运动估计采用整像素运动矢量,所以,在 H.261 标准中,压缩码率不是特别高。

2. H.263

H.263 是为了支持低速率通信而制定的标准,主要应用于 PSTN, ISDN 和无线网络,但也能够适应较大的动态范围,不限于低码率。对 CIF 图像,在 128kb/s~1Mb/s 码率范围内, H.263 可以获得比 MPEG-1 更好的压缩效果。

虽然 H.263 是从 H.261 基础上发展起来的,但是,它在低码率时,能够在不增加太多复杂度的情况下,提供更高的图像质量。此外, H.263 还包括 4 个可选模式,能够进一步提高压缩质量。

自从 1996 年 3 月 H.263 问世以来,就不断根据需求发展,先后出现了 H.263+, H.263++, H.26L 等后续标准。

1998 年, ITU-T 推出了 H.263 的第 2 版,叫做 H.263+。H.263+ 基于 H.263 标准,提出了一些新的可选特性以扩展 H.263 的应用领域并增强其编码效率。这些特性包括灵活视频格式、分级编码、补充增强信息 (Supplemental Information)、增强 PB 帧编码及高级帧内编码方法等。其中分级编码包括空域分级、时域分级、SNR 分级模式,这几种模式的选择,增强了压缩信号的抗干扰能力,使得 H.263 码流适合于网络的传输。

H.263++ 是 H.263 的第 3 版,它在 H.263+ 的基础上,又增加了很多优化的选项,这些优化选项被整合后,作为必需的选项,构成了 H.264 的前身——H.26L。

H.263 系列标准具有里程碑的意义,在其中首次提出了很多概念,如变块大小的运动估计、初始运动矢量预测、无限制运动估计、多参考帧运动估计等都被其后很多标准所沿用。

3. H.264

H.264 是 ITU-T 和 ISO/IEC 联合制定的最新编码标准,它最先由 ITU-T 的 VCEG 于 1997 年提出,目标是提出一种更高性能(相对于当时的 H.263)的视频编码标准,其前身是由 H.263 发展起来的 H.26L,这里 L 代表 Long term,以区别 H.263 版本 2 和版本 3。由于其相对于 MPEG-4 的优良表现,2001 年年底,ISO/IEC 的 MPEG 加入到标准的开发中,与 VCEG 组成 JVT。到 2002 年年底,H.264 完成了所有技术工作,2003 年年底正式成为官方标准。该标准在 ITU-T 中被称为 H.264,而在 ISO/IEC 中成为 MPEG-4 的 Part 10 (Advanced Video Coding Profile)。

目前,H.264 标准已经公布并被广泛采用,从各种实验来看,它达到了当初提出的目标,目前,很多 DSP 芯片具有标准分辨率下实现 H.264 编码器的能力和高清晰度下的实时解码能力。

1.2.2 MPEG 系列视频标准

1. MPEG-1

MPEG-1 是于 1991 年制定完成的。该标准主要是为了视频存储媒体(如 VCD)而制定的,目标应用码率为 1~1.5 Mb/s,提供 25 帧 CIF (352×288) VHS 质量的图像。MPEG-1 提供了视频序列的随机读取,快进和快退,视频序列的反向播放和压缩码流的可编辑性等功能。

MPEG-1 在典型的运动补偿预测编码(MCPC)框架基础上,应用了双向预测技术和半像素搜索,可以提供更好的编码质量和更高的压缩比。

双向预测技术允许编码帧参考前向和后向的参考帧,这样,MPEG-1 首次引入了 B 帧,B 帧不被用于其他 B 帧或 P 帧的运动补偿预测,因此,可以容忍更大的失真,并能提供相对于 I 帧和 P 帧更大的压缩比,大大提高了视频序列的压缩效率。

1/2 精度的运动估计,通过对整像素位置的亮度和色度进行插值,得到亚像素位置的信息,再进行运动估计。由于亚像素位置信息是对整像素的平滑,由此进行运动估计的结果会使得残差能量更低,压缩质量更好,但由此带来的运算复杂度也更高。

此外,MPEG-1 中还包含完整的音频编码、系统控制及一致性测试的规范。VCD 的广泛流行说明了 MPEG-1 的成功。MPEG-1 的图片组结构,使得其便于浏览,因此也被广泛应用于视频监控设备中。

2. MPEG-2

MPEG-2 是在 MPEG-1 基础上进一步发展起来的音视频编码标准,主要应用于广播级高质量音视频领域。目标码率为 3~35 Mb/s。该标准自 1990 年开始制定,于 1994 年完成标准化工作,目前被广泛应用于 SDTV, HDTV, DVD, DVB 等领域,在商业上取得了很大成功。在视频编码方面,与 MPEG-1 相比,MPEG-2 加入了两个主要的新技术:对隔行扫描的支持和可伸缩性编码。

对隔行扫描, 在视频的每帧图像中, 相邻行属于不同的场, 一帧图像被分成两场: 顶场和底场, 也称奇场和偶场。在景物存在快速垂直运动时, 相邻行的相关性会降低, 影响编码效率。MPEG-2 通过支持场编码来解决这个问题。场编码也是电视信号采用的形式, MPEG-2 支持该项技术, 也使之和电视视频兼容。

可伸缩性编码, 顾名思义, 就是根据信道质量或终端解码性能, 对传输的码流进行实时调整, 只传输或解码其中部分码流, 仍可以得到完整的解码图像, 只是相对于解码全部码流后的图像, 或降低了分辨率, 或降低了质量, 或减少了帧率。MPEG-2 支持 4 种可分级编码模式: 数据划分、SNR 分级、空间分级和时域分级, 大大增强了该标准的实用性。

3. MPEG-4

MPEG-4 标准的制定是于 1994 年启动的, 到 1999 年, 第一个 MPEG-4 官方版本正式发表。与 H.263 系列相对应, MPEG-4 也能支持低于 64kb/s 的视频应用, 同时它还能支持广播级的应用。它的制定目标是为了支持多种媒体的应用, 特别是多媒体信息基于内容的检索和访问。

为了实现基于内容的视频交互及检索, MPEG-4 的一大特色是提出了视频对象的概念, 并引入了基于对象的编码技术。MPEG-4 以 VO (Video Object) 的概念来实现基于内容的表示。VO 的构成依赖于具体应用和系统实际所处的环境, 在要求超低比特率的情况下, VO 可以是一个矩形帧, 与传统标准兼容。VO 也可以是场景中某一物体或某一层面, 为画面中被分割出来的不同物体。每个 VO 由三类信息来描述: 运动信息, 形状信息和纹理信息。

此外, MPEG-4 还应用了很多技术, 例如, 多边形匹配、亚像素搜索、动态的块选择、重叠的运动估计、形状编码技术、sprite 技术及可扩展性编码, 等等。

但由于实际存在对复杂视频中对象自动分割的困难, 目前实时实现的 MPEG-4 编解码器并没有实现基于对象的编码技术。当前 MPEG-4 最流行也最成功的是其简单档 (Simple Profile) 和增强的简单档 (Advanced Simple Profile)。前者基本和 H.263 类似, 后者在 H.263 的基础上引入了 1/4 精度运动估计和全局运动估计技术, 基于对象的概念和技术并未包括在这两档中。

4. MPEG-7 和 MPEG-21

MPEG 专家组还制定了两个多媒体技术标准, 分别是 MPEG-7 和 MPEG-21, 它们并不是视频压缩技术的标准, 下面对它们进行简要介绍。

在 MPEG-4 之后, MPEG 专家组致力于 MPEG-7 “多媒体内容描述接口” (Multimedia Content Description Interface) 的制定。MPEG-7 可以形容为“基于语义的表示”。它指定一个用于描述各种多媒体信息的描述符标准集, 还将定义其他描述符的方法标准化, 也包括标准化描述符及其相互关系的结构。MPEG-7 标准化了一种描述语言, 如 Description Definition Language (DDL)。视听材料 (如静止图像, 图形, 3D 模型, 音频, 视频) 和关于这些材料在一个多媒体表达中是如何结合等信息, 通过 MPEG-7 的描述就能被索引和搜索。

MPEG-21 是为了支持电子内容传输和电子商务而出现的国际标准,随着多媒体技术的迅猛发展,运营商为各自用户提供了丰富的信息和媒体业务,用户几乎可以随时随地享受这些服务。但是,对于不同网络之间用户的互通问题,至今仍没有成熟的解决方案。为了解决以上问题,MPEG 从 2000 年 6 月开始着手定义 21 世纪多媒体应用的标准化技术——MPEG-21 “Multimedia Framework”,MPEG-21 是一个可互操作和高度自动化的框架,而且这个框架还考虑到了数字版权管理 DRM (Digital Rights Management) 的要求、对象化的多媒体接入及使用不同网络和终端进行传输等问题。

关于视频图像编码的基本方法和关键技术,在第 2 章有更详细的叙述。

1.3 DSP 系统开发的基本流程

DSP 是一种非常适合进行实时数字信号处理的微处理器,它具有可编程性好、可靠性高,灵活性大,易于大规模集成等优点,从诞生的第一天起就为业界所推崇。随着人们对实时信号处理要求的不断提高和大规模集成电路技术的迅速发展,DSP 芯片技术也不断推陈出新,发生着日新月异的变化。例如,传统的 DSP 是功能相对单纯的为信号处理结构优化设计的专用处理器,但近年来,DSP 和其他处理器技术的结合,出现了针对各种类型应用的真正的片上解决方案,例如,TI 的 DM642 是很好的视频解决方案,Davinci 和 OMAP 都是多核系统,集成了 ARM 和 DSP 多核处理器,可用于复杂的视频处理设备和移动设备。DSP 芯片在视频处理中的应用也越来越广泛。

1.3.1 DSP 的发展及特点

世界上第一款公认的 DSP 是于 1978 年诞生在 AMI 公司的 S2811,1979 年 Intel 公司推出的商用可编程器件 2920 是 DSP 芯片发展史上的重要里程碑。这两种芯片都没有现代 DSP 所拥有的单周期乘法器,但是已经是 DSP 芯片的雏形。1980 年日本 NEC 公司推出的 μ PD7720 是第一款具有乘法器的商用 DSP 芯片。

在 DSP 领域最成功的企业应该是美国的德州仪器公司 (Texas Instruments, TI),自 1982 年 TI 公司推出 TMS320 系列 DSP 中的第一款定点 DSP (TMS32010) 以来,TI 的 DSP 已经发展了若干代,有 C1x, C2x, C2xx, C5x, C54x, C55x, C62x, C64x 等定点 DSP; 有 C3x, C4x, C67x 等浮点型 DSP; 此外还有 C8x 多处理器的 DSP, 近年推出的针对多媒体处理的 DM64x, Davinci, 以及针对移动终端的 OMAP 等。从最初的 16 位 DSP 发展到如今的 64 位 DSP, 目前, 处理能力最高的 C64x DSP 的最高主频已经达到 1.1GHz, 处理能力为 8 800 MIPS。

除了 TI 公司以外,还有一些厂家的 DSP 产品在市场上占有一定份额。美国 AD (Analog Device) 公司较著名的 DSP 芯片有 ADSP2101/2103/2105, ADSP2111/2115, ADSP2161/2162/2163/2164/2165/2166, ADSP2171/2173/2181 等定点 DSP, 此外还有 ADSP/21000/21020, ADSP21060/21062 浮点 DSP。还有一些公司,如 AT&T, Motorola, NEC 公司的 DSP 芯片也各有特点。

DSP 主要的特点可以概括为以下几点。

1) 采用哈佛结构

总线结构可以分为两种,一种是冯·诺伊曼结构,一种是哈佛结构。早期微处理器内部大都采用前者,其特点是程序和数据共用一个存储空间,程序数据总线共享,采用时分复用的方式共享总线。缺点是执行指令时只能串行执行,执行速度慢,吞吐量低。当高速运算时,不但不能同时取指令和操作数,而且会造成传输通道上的瓶颈。DSP 内部采用哈佛总线结构,数据存储空间和程序存储空间独立,数据总线和程序总线分开,这样就能够同时取数据(数据内存)和指令(程序内存),减少了冲突,大大提高了内存的访问速度。

TI 公司的 DSP 采用改进的哈佛结构,修正的地方有两点:一是数据空间和程序空间能够交换数据,有交叉数据通道;二是具有高速缓冲器,能够让 CPU 高速访问,大大节省读取指令和数据的时间,提高了运行速度。

2) 流水线技术

所谓流水线操作,就是一个时钟周期同时进行多条指令的操作,取指令和执行指令同时进行,从而减少指令执行时间。DSP 执行一条指令,需要经过取指令、译码、取操作数和执行等几个阶段,每一阶段成为一级流水,当执行本条指令的时候,就可以同时进入下面指令的取指令、译码、取操作数的阶段。同时执行的指令个数称为流水深度,不同产品的流水深度不同,AD 公司的 ADSP 深度为 2 级, TI 公司的 C64x 和 Motorola 公司的 568xx 深度为 5 级。

3) 多总线结构

许多 DSP 芯片内部都是多总线结构,这样就能够 1 个时钟周期内多次访问程序空间和数据空间。例如, TMS320C54x 内部有 4 条总线,可以在 1 个时钟周期内从程序存储器取 1 条指令,从数据存储器取 2 个操作数和向数据存储器写 1 个操作数,大大提高了 DSP 的运算能力。

4) 特殊的指令系统

DSP 往往有一些功能强大的指令,这些指令都是专为数字信号处理领域的应用而设计的。例如, TMS320C64x 系列 DSP 的 LDDW 与 LDNDW 指令能够一次取 8 字节的数据。ADD4 一次能计算 4 组 8 位加法,此外, DSP 还有一些为专用领域设计的指令,如视频处理指令。C64xDSP 中的 SUBABS4 指令能够一次执行 4 组 8 位数据求差绝对值运算, DOTPU4 指令能够一次执行 4 组 8 位数之间的点乘运算,这些指令能够大大提高视频处理中运动估算法的运算速度。

5) 多处理单元

DSP 内部一般都有多个处理单元,它们在一个时钟周期内能够同时进行运算,这样大大提高了 DSP 的并行处理能力。如 TI 公司的 C64x 系列 DSP 具有 8 个功能单元,为两组 S (Shift), M (Multiply), D (Data address), L (Logic) 单元。