

新型显示技术及应用集成系列丛书

张兆杨 安 平 张之江 沈礼权 等 编著

二维和三维视频处理及立体显示技术



 科学出版社
www.sciencep.com

内 容 简 介

由于三维视频技术是二维视频技术的延拓,本书扼要介绍了以 H. 264/AVC 和 TFT-LCD 显示器为主的二维视频技术;主要阐述了三维视频与立体图像显示技术,主要内容包括:三维视频与立体图像显示系统的组成及其发展概况,构成 3DTV 的基于双目/多视摄像阵列的三维场景获取、双视/多视视频编码、解码端的合成视绘制和基于液晶的立体显示等四个主要组成部分的基本概念与实现方法。对其中的关键技术包括多视摄像机的标定方法与预处理、基于 H. 264/AVC 时域分层 B 结构的多视视频预测结构/编码方法及发展的新方法、基于视内插和深度的视合成绘制方法以及自动立体液晶显示器技术及其优化方法进行了较深入的分析和阐述。

本书可作为通信、计算机、光电子、传媒类等研究机构和企事业单位的科研、设计人员从事研究和新技术开发的参考书,也可作为上述相关专业的本科生、硕士生、博士生的学习用书,或高校教师的教学、科研参考书。

图书在版编目(CIP)数据

二维和三维视频处理及立体显示技术/张兆杨等编著. —北京:科学出版社,2010

(新型显示技术及应用集成系列丛书)

ISBN 978-7-03-027239-3

I. ①二… II. ①张… III. ①视频信号-信号处理②液晶显示器
IV. ①TN941.1②TN141.9

中国版本图书馆 CIP 数据核字(2010)第 067528 号

责任编辑:刘宝莉 陈 婕 / 责任校对:桂伟利
责任印制:赵 博 / 封面设计:嘉华永盛

科学出版社 出版

北京东黄城根北街 16 号
邮政编码:100717

<http://www.sciencep.com>

新 蕾 印 刷 厂 印刷

科学出版社发行 各地新华书店经销

*

2010 年 4 月第 一 版 开本:B5(720×1000)

2010 年 4 月第一次印刷 印张:15

印数:1—3 000 字数:300 000

定价:42.00 元

(如有印装质量问题,我社负责调换)

前 言

视频技术正在完成从模拟视频到数字视频的过渡,高清晰度电视已逐渐进入家庭,这标志着二维视频技术的研究和发展即将告一段落,势必向三维视频技术发展。三维视频是指在原有二维视频的基础上增加了作为第三维的“深度”,且从平面图像的显示发展至立体图像的显示。2010年初,美国立体电影《阿凡达》火爆到一票难求的程度,在美国拉斯维加斯“国际消费电子展”上三维立体电视类的展品成为主角,这些均表明人们对三维影像的祈盼。

虽然三维视频的研究可上溯到约一百年之前,但由于科学和技术的限制,在20世纪末才开始形成国际上的研究热点。尤其是在2001年ISO/IEC JTC1/SC29/WG11(活动图像专家组, Moving Pictures Experts Group, MPEG)成立了3DAV Ad hoc工作组,并于2003年发布了“3DAV的应用和要求”和“3DAV探索报告”两个文件后,欧洲、美国、日本、韩国和我国的高校、研究所、企业逐步掀起了研究热潮,并开始进入各个领域的应用。目前三维视频正处于方兴未艾,从实验室研究跨向产品化、市场化的关键时期。

本书是在国家自然科学基金重点项目“自由视点多视视频以及3D立体显示基础理论与关键技术(60832003)”及另外三项国家自然科学基金项目“多视点自动立体显示关键技术的研究(60772124)”、“3DAV中的多视点视频编码及解码端视点绘制研究(60672052)”和“多视点/立体视觉的分析与合成研究(60202015)”的资助下所累积的研究成果的基础上完成的,并参考和融入了国内外现有的基本理论和研究成果。

三维视频的类型甚多,常随应用目标的不同而异,本书内容主要集中在以应用于3DTV为目标的系统组成、各组成部分的基本理论、结构和有代表性的实现方法以及与之相关的二维视频基础知识。全书共6章,其中第1章叙述了3DTV的发展概况及其组成的系统结构和关键技术,指出了不同组成系统3DTV的特点、主要应用场合及其走入家庭的前景。第3、4、6章是本书的核心部分;第3章介绍了3DTV场景获取系统,包括双目摄像机和多视点摄像阵列组成方式及其标定方法与所获取视频信号的预处理,其中包含了有关3DTV的概念与相关理论;第4章针对获取系统输出的双视或多视视频信号阐述了对其进行高效编码以压缩存储容量和传输带宽的预测结构、工具、典型算法及优化方法,以及解码端将解码的多视信号绘制合成为立体对图像信号供立体显示器显示的基本原理和代表性方法;第6章给出了立体视觉的感知原理、各种三维立体显示方式,对其中目前已可应用

的基于视差的自动立体显示器的几何模型、实现原理和最有代表性的狭缝光栅及透镜光栅立体显示器作了较深入的分析 and 阐述,并在第 5 章的基础上讨论了基于液晶显示屏的立体显示器主要参数与立体显示性能的关系。尤其对液晶屏立体显示器可能发生的失真及用户观看时可能会感觉到的视觉疲劳与不舒适感,在分析其发生原因的基础上提出了改进方法,并通过建立立体显示系统的数学模型提出一种可进行优化的设计方法。由于第 3 章和第 4 章中多视视频及其编码方法与第 6 章中基于液晶屏的三维立体显示的阐述分别是在第 2 章的二维视频处理与编码和第 5 章的二维显示技术的基础上进行介绍的,因此,有的读者在阅读第 3、4 章和第 6 章之前,需要先分别阅读第 2 章和第 5 章。

本书由张兆杨拟定全书的大纲和撰写第 1、2 章,并对全书进行统稿、修改和定稿;由安平撰写第 3、4 章;由张之江撰写第 5、6 章。其中,沈礼权和张倩分别参与了第 4.1、4.2 和 4.3 节初稿的撰写,董志华参与了第 6 章内容的整理。此外,本书在编著过程中得到了上海市经济和信息化委员会、上海世博会事务协调局、上海市科学技术委员会、上海大学通信学院和教育部新型显示技术与应用集成重点实验室的大力支持,在此表示衷心的感谢。

尽管 3D 视频与立体显示目前已有了局部应用,但总体而言,它是尚处于实验室阶段、有待逐步完善的新技术。本书是国内该领域的第一本著作,限于作者水平,书中难免存在不妥之处,谨请读者指正。

作 者

目 录

前言

第 1 章 三维视频与立体显示系统概述	1
1.1 三维立体视频与显示的基本概念	1
1.1.1 三维立体视频与显示系统的组成	1
1.1.2 不同类型的三维系统	2
1.2 3DTV 的关键技术及其发展和应用概况	4
1.2.1 3DTV 的关键技术	4
1.2.2 发展趋势和应用概况	5
1.3 本书的结构	7
参考文献	8
第 2 章 二维视频处理与编码	9
2.1 二维视频信号的数字化处理	9
2.2 二维视频信号的压缩编码原理	11
2.2.1 基本原理	11
2.2.2 熵编码	12
2.2.3 预测编码	14
2.2.4 变换编码	15
2.2.5 运动估计和补偿	18
2.3 数字视频压缩编码标准	23
2.3.1 MPEG-2	24
2.3.2 MPEG-4	32
2.3.3 H. 264	39
2.3.4 AVS 国家标准	47
2.4 H. 264 视频编码的优化和新的扩展	50
2.4.1 帧内预测模式的快速选择	50
2.4.2 帧间预测模式的快速选择	54
2.4.3 自适应亚像素运动矢量快速搜索算法	58
2.4.4 对多参考帧选择的算法优化	64
2.4.5 H. 264 视频编码扩展至多视点视频编码	69
参考文献	69

第 3 章 三维立体视频获取和预处理	71
3.1 三维立体视频的获取	71
3.1.1 双目立体视频的获取方法	71
3.1.2 多视点立体视频的获取方法	74
3.1.3 多视采集系统摄像机数量的优选	75
3.2 视频获取系统的参数标定	80
3.2.1 基于平面模板的单摄像机标定	80
3.2.2 双视立体标定法及其在多视摄像机中的应用	86
3.2.3 基于多视定位法的多视摄像机联合标定	88
3.3 多视摄像机信号的预处理	92
3.3.1 亮度与色度补偿	93
3.3.2 空间几何补偿和锐度补偿	94
参考文献	97
第 4 章 双视和多视视频编解码与绘制	99
4.1 3DTV 系统与相关概念	99
4.1.1 3DTV 系统的基本组成	99
4.1.2 双视立体和多视视频编码的要求	101
4.2 双视与多视视频编码的预测结构和工具或算法	104
4.2.1 双视立体视频编码	104
4.2.2 多视编码几种基本的预测结构	107
4.2.3 基于 H. 264 的多视视频预测结构	109
4.2.4 MVC 的若干预测工具/算法	115
4.2.5 MVC 可分级编码(SMVC)	128
4.2.6 “视频+深度”的编码方法	131
4.2.7 降低多视点编码复杂度的优化模型与方法	137
4.3 解码端的视合成绘制	143
4.3.1 自由视点 3DTV 解码框架及视绘制的概念	144
4.3.2 基于视内插的多视图像绘制方法	147
4.3.3 基于深度的图像绘制方法	149
4.3.4 完全聚焦型绘制	153
参考文献	156
第 5 章 基于二维视频的平面图像显示	158
5.1 二维视频液晶显示器的工作原理	158
5.1.1 液晶显示器的基本特性	158
5.1.2 液晶显示器的基本组成	159

5.1.3 TFT-LCD 二维视频显示器的主要技术指标	167
5.2 液晶显示器的关键部件——背光源	168
5.2.1 背光源的一般结构和基本参数	169
5.2.2 CCFL 背光源与 LED 背光源及其性能比较	171
5.3 提高 LCD 显示性能的关键技术	175
5.3.1 扩展视角的方法和技术	175
5.3.2 提高响应速度的方法和技术	182
5.4 液晶显示的新发展	186
5.4.1 反射式 LCD	186
5.4.2 低温多晶硅技术(LTPS)	189
参考文献	190
第 6 章 基于三维视频的立体图像显示	191
6.1 立体视觉的感知机理和立体显示	191
6.1.1 单眼感知与双眼感知立体信息的机理	191
6.1.2 不同方式的三维立体显示	193
6.2 三维视频自动立体液晶显示的原理	200
6.2.1 狭缝式立体显示	200
6.2.2 透镜式立体显示	202
6.3 立体视频液晶显示器特有的技术性能和问题	208
6.3.1 立体视频显示特有的技术性能	209
6.3.2 观看立体视频的视觉疲劳与不舒适感	213
6.4 液晶立体图像显示的新发展	215
6.4.1 立体显示系统的数学模型	215
6.4.2 基于模型的显示质量优化	219
6.5 结合头部跟踪的立体显示技术	224
6.5.1 基于单用户头部跟踪的立体显示	225
6.5.2 基于多用户头部跟踪的立体显示	226
6.5.3 几种头部位置跟踪方法	228
参考文献	230

第 1 章 三维视频与立体显示系统概述

视频处理、传输与显示技术经历了从黑白到彩色,从模拟电视到数字电视,从标准清晰度电视到高清晰度电视的多次变革。每次变革都提高了人们的视觉享受。人们不禁要问,下一代变革的目标是什么? 2010 年初发生了两件引起社会轰动的事件,一是立体电影《阿凡达》引起观影热潮,呈现出一票难求的现状;二是在美国拉斯维加斯“国际消费电子展”上,立体视频显示系统与相关设备成为最受人们关注的产品。这些恰恰回答了下一代变革的目标就是人们渴望已久的三维(3D)立体视频系统。目前,千家万户都在观看的二维(2D)视频平面显示系统由于丢失自然界中物体的作为第三维的深度(景深),使人的视觉缺乏立体(深度)感而有与实际的自然场景不一致(失真)的感觉。本书在介绍二维视频处理与平面显示的基础上,着重讨论三维立体视频与显示的基础理论、系统的组成、各组成部分的工作原理、构成方法及其应用。

1.1 三维立体视频与显示的基本概念

1.1.1 三维立体视频与显示系统的组成

图 1.1 是三维立体视频与显示系统的组成框图。三维立体视频与显示系统由

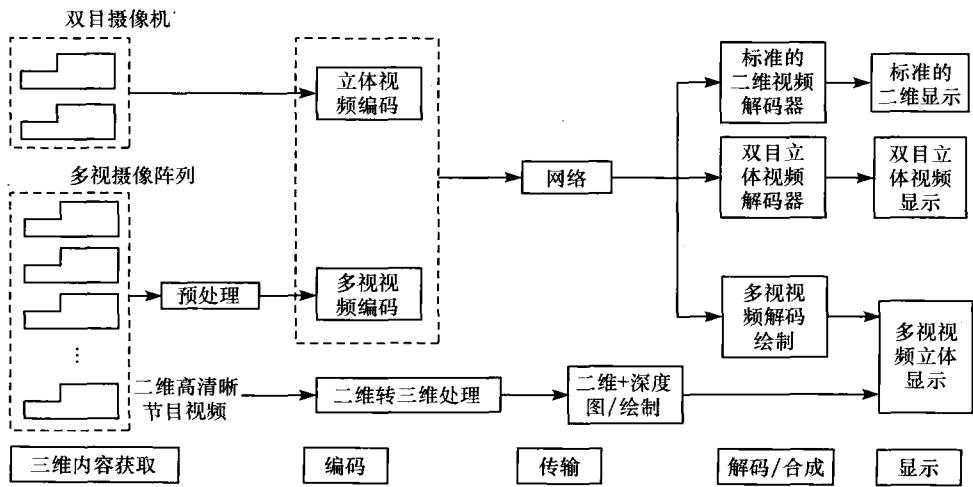


图 1.1 三维立体视频与显示系统^[1]

三维内容获取、编码、传输、解码/合成和显示 5 部分组成。其中,由摄像子系统摄取三维场景,由编码子系统将获取的视频信号去除冗余信息后编码成便于传输的视频流,经网络传输后送入解码/合成器,按照显示端的要求重建视频信号后分别送至不同类型的显示器以显示双目立体图像、多视立体图像或二维平面图像。

1.1.2 不同类型的三维系统

图 1.1 包含了不同类型的三维系统,下面按照系统的不同以及三维视频获取方法、编码、解码/合成和显示方式的不同作分类说明。

1. 基于双目的三维电视(3DTV)系统

图 1.1 的最上端是传统的双目立体视频系统,其三维场景的获取是通过光轴中心相距 6.5cm 的相互平行的相同型号的摄像机得到。左、右两路视频经立体视频编码压缩成视频流,视频流经传输后解码复原成两路视频,再在显示器中构成与人的双眼视觉对应的稍有不同的两幅图,最后由大脑合成有立体感的图像。

2. 基于多视的 3DTV 系统

图 1.1 中间部分的多视 3DTV 系统由 N 个($N>3$)摄像机阵列获取三维场景。尽管 N 个摄像机型号相同,但其内外参数难以完全一致,且 N 个摄像机在空间的位置不同,各自的光照也略有差异,因此需进行摄像机几何参数校正和亮度/色度补偿等预处理,然后再将多路视频信号经多视视频编码后压缩成视频流。由于解码后重构的 N 个视(每一路视频简称为一个视)不一定适合某个观众在显示屏前所在位置的观看要求,且为提高显示质量,故解码后需通过选择合适的一些视经绘制合成符合需要的两个视。多视 3DTV 系统的优点是,在显示屏前人能看到立体效果的视角(简称立体视角)远比双目 3DTV 系统的大,且便于应用人眼跟踪技术使人在屏前移动时所观看到的立体图像也随人的移动而变化,提高了真实感和临场感。

3. 立体视频编码与多视视频编码

目前正在使用的二维视频系统,传输的仅是单路视频。但由于数字化使数字视频的带宽远大于模拟视频,对于目前观看的标准清晰度电视,数字化后的比特率已达到 216Mb/s,为使其能在 6MHz 或 8MHz 带宽的信道内传输,压缩编码方法从 MPEG-1、MPEG-2 到 MPEG-4 或 H. 264/AVC(后文简称 H. 264)历经了近二十年的研究。对双目立体视频编码而言,将单路视频增加为双路视频,其中的一路采用目前压缩效率最高的 H. 264 编码,另一路利用两路视频之间的视间相关性或运动与视间的联合预测编码,可使两路视频编码后的总比特率约是单路视频的 1.25 倍。目前,双目立体视频编码已达到实时应用阶段。对于多视视频编码,随着视数 N 的增加(目前用的较多的是 $N=8\sim 16$),对高效压缩的要求远比双目立

体视频编码的要高,离实际的应用,尤其是实时应用尚有较大距离。

4. 与二维视频系统的兼容性

目前所有的家庭几乎都已有了播放二维视频图像的电视机,若开播双目或多视立体视频后,应使现有的遍及每个家庭的电视机也能接收到三维立体节目(尽管看到的仍是二维视频图像)。为实现此种后向兼容,在图 1.1 中,无论是在立体视频还是多视视频的编码中,基本视(双目视频或多视视频中作为参考视的 1 个视)应沿用二维视频的编码标准(如 H. 264),这样就可以由图 1.1 中标准的二维视频解码器重建双目或多视视频中的基本视,提供给标准的二维视频显示器。

5. 立体图像显示方式和立体显示器类型

1) 立体图像显示方式

此种显示可分为需佩戴“立体眼镜”的立体显示和用裸眼直接观看的自动立体显示。

(1) 佩戴立体眼镜的显示方式。

2010 年 1 月风行全球的立体影片《阿凡达》和在美国拉斯维加斯“国际消费电子展”上展出的大多数立体电视的观看都是采用佩戴立体眼镜的方式。立体眼镜可分为偏振式、快门式和分色式三种。偏振式立体眼镜是左、右眼分别使用极化方向相互垂直的偏振镜片,其中一只眼用垂直偏振,另一只眼用水平偏振,显示器投射相应的偏振光,从而使双眼分别观看左视和右视图像,并由大脑融合成立体图像。此种方式的缺点是人的头部倾斜时偏振镜片难于滤掉与之正交的偏振光,使一个视的图像漏到另一个视中,使人产生不舒适感。

快门式立体眼镜佩戴的是液晶制成的快门眼镜,当加上一定电压时改变液晶分子的排列而控制开关状态,使得一个镜片阻挡光线时另一镜片光线可通过,从而使双眼分别观看左、右视图像。此种方式的缺点是若显示器刷新频率不够高,将会产生闪烁感。

分色式立体眼镜的左、右眼镜片分别是红光或蓝光滤色片,它使得戴红光滤色片的眼睛只能看到红色图像,戴蓝色滤色片的眼睛只能看到蓝色图像,由此,使双眼分别观看左视和右视图像。它的缺点是彩色信息损失大,色调单一。

(2) 自动立体显示方式。

不戴立体眼镜由人眼直接观看的自动立体显示方式常用的有视差栅栏式和柱透镜光栅式。前者的原理是利用视差挡板分光,在液晶显示器的液晶层之前或之后安装视差栅栏,将其分为挡光和透光两部分,使得显示屏上各像素交替左、右眼对应的图像,令一幅经过匹配处理的视差图像分别投射到左、右眼,产生立体视觉。视差栅栏式的优点是工艺结构较简单,但由于被遮挡了约 50% 的光,因此显示屏

的亮度也损失了一半。

柱透镜光栅式是在显示器内部安装一排垂直排列的柱面透镜,利用每个柱面透镜对人射光的折射作用,把两幅不同的视差图像分别透射到对应于双眼的视域,使左图像聚焦于观看者左眼、右图像聚焦于观看者右眼,从而产生立体视觉。柱透镜光栅式的优点是大幅度减少了光的损失,使显示屏的亮度几乎是视差栅栏式的一倍。但是柱透镜光栅式的工艺要求高,每个透镜的截面达到了微米级,较适合用于大屏幕显示。

2) 立体显示器类型

立体图像显示器可分为单用户立体显示器和多用户立体显示器。单用户立体显示器对应于双目立体视频系统。双目立体视频由于只有两个视,能显示立体效果的区域范围小,即立体视角小,一般只适用于单个观众观看。这是因为若多人同时观看时,有的观众会因其双眼无法同时接收视差图像源而感受不到立体效果。

多用户立体显示器对应于多视视频系统。多视视频是由 N 个 ($N > 3$) 摄像机从不同视角同时拍摄相同场景获得的,可以使多个观众在不同位置接收到不同的视差图像源,因此,多视视频的立体视角宽,可供多个观众同时观看。但是,正因为屏上要容纳多对视差图像源,立体显示分辨率也就相应降低,为此应要求显示器自身的分辨率很高,且使柱透镜阵列适度倾斜,以便由垂直分辨率补偿水平分辨率,相对提高所观看立体图像的清晰度。

1.2 3DTV 的关键技术及其发展和应用概况

1.2.1 3DTV 的关键技术

在图 1.1 给出的 3 种方式 3DTV 系统(双目立体、多视立体、二维转为三维视频)中,基于多视的 3DTV 尚处在研究和逐步完善阶段,其关键技术可归纳如下。

1. 多视摄像系统的同步标定技术

目前双目立体摄像系统的技术日趋成熟,今后的研究重点将是多视摄像系统技术。前面曾提及为保持 N 个视的一致性,需将 N 个摄像机参数严格校正到一致(后文中称为摄像机标定)。目前的标定方法是基于固定焦距、光圈和固定位置的,但在实际拍摄场景时由于需变动光圈、焦距和不断移动位置,要使 N 个摄像机在变焦、变光圈、变位置时同步一致,在技术上尚有困难。

2. 多视视频编码的实时实现

尽管自 2001 年 12 月 MPEG 成立三维音视频(3DAV)工作组以来,对多视视频编码(MVC)的研究已历经 8 年,并由联合视频组(JVT)主持的对 MVC 标准化

研究在 2009 年 10 月已趋结束,很快将推出 JMVC 的最终版草案,但因其运算复杂度远比 H.264 高,离实用化尚有很大距离,尤其是对于有实时性要求的应用,因此仍需对最终版的 MVC 进行深入的算法优化研究和硬件实现研究。

3. 快速高质量的多视绘制

如何提高多视视频系统解码端绘制的质量和速度是保证立体显示器能实时优质地显示立体图像的关键技术之一。其中,基于深度的图像绘制(DIBR)是一种适用于 3DTV 中的图像绘制的方法,该方法需要准确快速地提取图像的深度,然而,现有的方法在快速和准确之间相互制约,因此,需进一步深入研究。

4. 多用户立体显示器眼跟踪技术

目前栅栏式或柱透镜光栅式立体图像显示器技术已日趋成熟。但是用于 3DTV 的立体显示器是供多人(这里称多用户)观看的,当多用户同时观看时,各用户因在屏前所处的位置和视向的不同所看到的立体图像而有所不同,当某用户在屏前从甲处移动到乙处时其看到的立体图像也有差异,这就需要用头部位置跟踪技术控制视点自适应变换(切换)。目前这一技术尚未成熟,离实用化还有很大距离,需作为关键技术进一步研究。

1.2.2 发展趋势和应用概况

1. 双目立体视频的应用及需解决的问题

1) 目前的应用概况

由双目视频组成的立体视频显示系统由于其立体视角范围小,仅适合单人观看,一般用于科学研究、航天航空、医疗、监控、教学、军事、艺术、娱乐、网上商场及网上旅游和文物展示等场合。

2) 双目立体视频进入家庭尚需解决的问题

(1) 戴立体眼镜式的家用电视。

此种方式因电视机改动少,如以 DVD 为节目源或者将一对视差图像共用一条信道交替顺序传送,其传输制式便于与现有的二维电视兼容,故较快就可进入家庭。

(2) 自动立体显示式的家用电视。

不戴立体眼镜的自动立体显示方式由于需要在现有的电视机内安装柱透镜光栅式或视差栅栏(此技术相对已较成熟),以及通过光学视角倍增器来扩大立体视角(尚在实验室阶段),预期约 3~5 年也可进入家庭。

2. 多视视频立体显示的应用及需解决的问题

1) 限制多视视频进入家庭尚存在的主要问题

多视视频及其显示系统目前尚未解决的主要问题有两点:一是其三维内容的

获取是由 N 个完全一致的摄像机同步同时完成的, N 个摄像机同步变焦、调焦、变光圈(各参数必须相同地同时改变)在技术上并不容易实现,有待发展一种完全自动的和精确的基于多视的深度/视差获取系统^[2];二是多视视频巨大的数据量尽管经 MVC 已有较大的压缩,但离可应用到 3DTV 广播上尚有很大距离,进一步优化可对运算复杂度作大幅度降低,但对传输比特率的压缩却是有限的,这就需要发展一种可根据人的头部位置选择少数视而以相当低的带宽传输视频流的结构与方法^[3]。此外,3DTV 广播不仅节目制作费用比二维的 HDTV 的更加昂贵,且演播室设备相当大的部分需要更新。因此,基于多视视频的三维立体电视机预期尚需 10 年才可进入家庭。

2) 多视视频的应用

前述的双目立体视频各种的应用,原则上也可推广应用到多视视频,尤其是在大型博览会及各种展示会上。如图 1.1 下端所示的那样,将二维高清晰度视频节目通过二维转三维处理和绘制及后期制作形成的多视视频传送到多视视频立体显示器来播放富有很强纵深立体感的三维立体视频节目。此种多视视频立体显示系统在国家自然科学基金重点项目研究的基础上,在上海世博会事务协调局、上海市科学技术委员会的资助下已于 2009 年在上海大学主持下和清华大学联合制成,且已先后在中国国际工业博览会和深圳高技术科技展示交易会上展示,并将于 2010 年上海举办的世界博览会中国国家馆中展出。图 1.2 是其中的 103 寸立体视频图像显示屏。二维转三维处理是指将已拍成的二维高清晰度电视节目通过锐度、对



图 1.2 103 英寸(右下角是 19 英寸二维显示器图像)立体视频图像显示屏

比度预处理,结合运动结构算法和深度提取技术得到“二维+深度图”,再转为多视三维节目,通过播放器在多视(图 1.2 中采用 8 个视)立体显示器上显示。

1.3 本书的结构

本书共有 6 章,内容结构安排如下:第 1 章扼要介绍三维立体视频的概念、三维系统的组成和分类、各种系统的特点以及“立体与多视视频编码”和“立体图像显示技术”的基本概念和要求。随后在此基础上对 3DTV 系统四个方面的关键技术作了说明,指出各类不同系统在设计时的要求和尚需解决的问题,并展望了预期可进入家庭的大致年份。

由于其中的立体视频和多视视频编码方法的主流是建立在单个视的已有编码标准基础上的,第 2 章对单视视频编码的原理及其主要编码标准 MPEG-2、MPEG-4、H. 264 和 AVS 作了介绍,重点是讨论 MPEG/JVT 规定 MVC 所后向兼容的 H. 264,并介绍了 MVC 与 H. 264 的关系。

第 3 章讨论了作为三维立体视频系统最前端的双视(双目)和多视视频获取的原理、方法,重点对多视视频获取系统究竟以多少个摄像机为宜作了较深入的理论分析和实验验证。由于获取多视视频的 N 个摄像机其内外参数必须一致,需要对每个摄像机进行标定(校正),该章在介绍张正友的单个摄像机标定法的基础上给出了多个摄像机的“两两标定法”,并进一步提出了基于多视定位的联合标定法。在多视视频编码时为提高其编码效率,应使每个视的亮度和色度一致,因此在 3.3 节介绍了以亮度补偿为主的对多视视频的预处理。

第 4 章讨论双目立体视频及多视视频的编码方法和视合成绘制。首先比较了双目视频、多视视频与自由视点视频的异同点,简要介绍了双视立体视频典型的预测结构及算法;其次重点介绍和分析了德国 HHI 研究院提出的、为 JVT 采纳的基于 H. 264 时域分层 B 结构的 MVC 的预测结构,并在此基础上对 MVC 的若干预测工具和编码算法作了说明;然后,提出了基于联合矢量的预测算法和对 MVC 复杂度的优化方法及 Shimizu 给出的可分级多视视频编码(SMVC)^[4],并进行了较深入的讨论;最后,介绍了最有代表性的基于视内插的视绘制、基于深度图的图像绘制(DIBR)和完全聚焦型的绘制等方法。

由于目前的三维立体视频大多是基于液晶显示器上显示的,最后两章分别给出了基于二维和三维视频的液晶显示技术。其中,第 5 章的二维视频液晶显示器已很成熟,且已出现在很多著作中,在此仅作简要的基础性介绍,便于读者在阅读时有知识的连续性。第 6 章较详细给出了立体视觉的感知机理和立体显示的各种方法,并对立体视频液晶显示器的主要参数和失真成因作了较深入的分析,在此基础上提出了改善立体显示质量的优化设计方法。最后,介绍了尚在研究和发展中

的基于人眼(头部)跟踪的立体显示技术。

参 考 文 献

- [1] Oliver S, Peter K, Thomas S. 3D Video Communication. London: John Wiley& Sons Ltd. , 2005:23.
- [2] Ozaktas H M, Levent O. Three-Dimensional Television. Berlin: Springer, 2008:317.
- [3] Tekalp A M, Kurutepe E, Civanlar. 3DTV over IP: End-to-end streaming of multiview video. IEEE signal Processing Magazing, 2007, 11: 77—87.
- [4] Shimuzu S, Kitanara M, Kimata H, et al. View scalable multiview video coding using 3D warping with depth map. IEEE Transaction on Circuits System for Video Technology, 2007, 17(11): 1485—1495.

第 2 章 二维视频处理与编码

数字视频系统是通过模拟视频信号进行数字化处理得到的,一个标准清晰度的模拟视频信号经数字化处理后其比特率增至 216Mb/s,超过了通常的广播频道或网络的传输能力,需将其中的冗余信息比特大量去除,以尽可能少的比特数对其描述,在传输后再予复原,这称为信源的压缩编码和解码。为便于相互间能通用,按照不同的应用要求,国际的相关机构制定了相应的编解码标准。先后制定的国际标准有 H. 261、MPEG-1、H. 263、MPEG-2、MPEG-4、H. 264 等,我国近年也制定了称为 AVS(先进的音视频编码标准)的国家标准。

本章在阐述二维视频信号数字化处理及其压缩编码原理后,对目前正在应用的主要标准 MPEG-2、MPEG-4、H. 264 和 AVS 作了扼要介绍,并对 H. 264 的编、解码方法及其算法优化进行了深入的分析研究,这也是第 3 章和第 4 章描述三维立体视频处理和编码的基础。

2.1 二维视频信号的数字化处理

将模拟视频信号数字化处理为数字视频信号,既可通过全信号也可通过分量信号(Y、R-Y、B-Y 或 R、G、B)转换。由于后者在从摄像机到传输的整个链路中可防止反复解码引起的质量损伤,因此目前普遍被采用,且已由 ITU-R 进行标准化,称为 BT. 601^[1,2]。

BT. 601 建议除采用对亮度信号和两个色差信号分别编码的分量编码方式外,对不同制式的信号采用相同的取样频率,对亮度信号 Y 的取样频率为 13.5MHz。由于色度信号的带宽远比亮度信号的带宽窄,所以对色度信号 U 和 V 的取样频率为 6.75MHz。每个数字有效行分别有 720 个亮度取样点和 360×2 个色差信号取样点,对每个分量的取样点都是均匀量化,对每个取样进行 8 比特精度的 PCM 编码。Y 信号黑、白电平分别对应于 16 级和 235 级,U 和 V 信号的最大正电平对应 240 级,零电平对应 128 级,最小负电平对应 16 级。这几个参数对 525 行、60 场/秒和 625 行、50 场/秒(我国电视标准)的制式都是相同的。有效取样点是指只有行、场扫描正程的样点有效,逆程的样点不在有效范围内。

对于 625 行/50 场标准,奇数场有效行为 288 行,场消隐为 24 行,一共 312 行;偶数场有效行为 288 行,场消隐为 25 行,一共 313 行;一帧 625 行与模拟电视一样。

BT. 601 建议的主取样格式为 4 : 2 : 2,其色度信号的取样率是亮度信号取样率的一半,可以理解为每一行里的 Y、U、V 的样点数之比为 4 : 2 : 2。图 2. 1(a)是 4 : 2 : 2 格式样点位置示意图,其他常用的格式还有 4 : 4 : 4 和 4 : 1 : 1,分别如图 2. 1(b)和图 2. 1(c)所示。此外,4 : 2 : 0 格式虽然不在 ITU-R BT. 601 标准中,但应用相当广泛,如图 2. 1(d)所示。

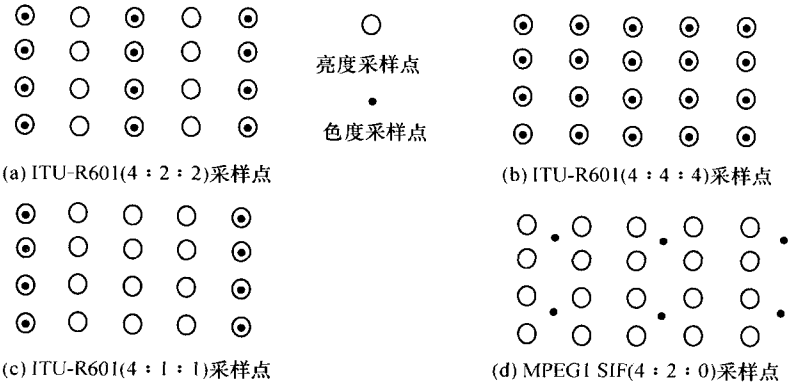


图 2. 1 不同格式的样点位置示意图

其中,在 4 : 2 : 2 取样格式时,ITU-R BT. 601 的参数如表 2. 1 所示。由表中可见,数字化之后,总比特率高达 216Mb/s,这对于常用的广播电视频道和无线传输以及 Internet 等网络是难以传输的,必须进行压缩编码。

表 2. 1 ITU-R BT. 601 参数(4 : 2 : 2 取样格式)

参 数	帧率 25Hz	帧率 30Hz	参 数	帧率 25Hz	帧率 30Hz
场率	50	60	每帧的有效行数	576	480
全帧行数	625	525	每行的有效取样数(Y)	720	720
每个取样比特数	8	8	每行的有效取样数(Cr,Cb)	360	360
总比特率	216Mb/s	216Mb/s			

在压缩编码之前,为适应各种视频编码的应用,常被转换成相应的中间格式,这些通用的中间格式对应的帧分辨率如表 2. 2 所示。

表 2. 2 中间格式

中间格式名称	亮度分量分辨率(水平×垂直)	中间格式名称	亮度分量分辨率(水平×垂直)
Sub-QCIF	128×96	4CIF	704×576
QuarterCIF(QCIF)	176×144	16CIF	1408×1152
CIF	352×288		