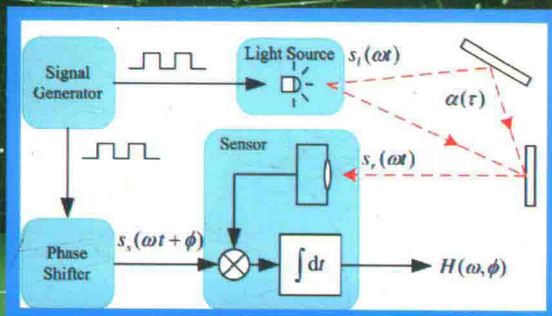


“十二五”国家重点图书出版规划项目



3D Video Processing and Communication

立体视频处理与通信

戴琼海 曹汛 著

Dai Qionghai Cao Xun



清华大学出版社

“十二五”国家重点图书出版规划项目

3D Video Processing and Communication

立体视频处理与通信

戴琼海 曹汛 著

Dai Qionghai Cao Xun

清华大学出版社

北 京

内 容 简 介

视频技术的发展经历了从黑白到彩色,从标清到高清,而今迈入了立体的时代。当前,立体视频在影视动漫、医疗诊断、军事仿真等众多领域有着广泛而深入的应用。典型的立体视频系统包括立体内容的生成、压缩、传输、重构和显示。针对当前立体视频内容匮乏和压缩、传输数据量大的困境,本书着重阐述立体视频的生成和通信问题。本书从立体拍摄、立体建模以及平面视频立体转换三个方面详细介绍立体视频产生的原理、方法,同时对于立体视觉产生的机制、主要的立体视频格式、立体失真因素等直接影响立体视频观感的重要问题进行阐述。本书旨在从立体视频产生的原理、立体视频格式、立体视频的压缩和传输等多个方面全面阐释立体视频和通信的基础理论和知识。

本书适合作为高等院校电子通信、计算机、自动化等信息学科,尤其是图像和视频处理专业的研究生教材,也适合科研院所从事立体视频方面的研究工作者参考。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

立体视频处理与通信/戴琼海,曹汛著.--北京:清华大学出版社,2016

ISBN 978-7-302-42574-8

I. ①立… II. ①戴… ②曹… III. ①立体声设备—视频设备 ②通信原理
IV. ①TN912.27 ②TN948.57 ③TN911

中国版本图书馆 CIP 数据核字(2016)第 005268 号

责任编辑:王一玲

封面设计:常雪影

责任校对:时翠兰

责任印制:李红英

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:北京鑫海金澳胶印有限公司

经 销:全国新华书店

开 本:175mm×245mm 印 张:10.25 彩 插:2 字 数:224 千字

版 次:2016 年 6 月第 1 版

印 次:2016 年 6 月第 1 次印刷

印 数:1~2000

定 价:49.00 元

产品编号:062696-01



人类尝试体验立体效果的历史其实十分久远。早在公元前 400 年,欧几里得就已经指出:“人类之所以能洞察立体空间,主要是因左右眼所看到的景物不同而产生”;到了公元 1838 年,英国的 Wheatstone 制作了立体镜用来观察立体效果;1922 年,第一部立体电影《The power of love》就在洛杉矶上映。尽管人们做了很多的尝试,直到 20 世纪 90 年代末,微电子、材料科学、计算机视觉等关键技术的突破,才使得立体视频技术在各个领域中的应用成为可能。

立体视频技术通过对场景信息采集、建模及深度计算,进行立体重建与显示,实现视觉信息的宽视场、高沉浸感、高逼真呈现,是视觉信息获取、传播与显示的新一代前沿技术。欧洲、美国、日本和韩国等纷纷启动了立体视频技术研究计划。我国有关科研计划在“十二五”中也将立体视频作为重要的研究发展方向之一,并且将在“十三五”中持续大力发展。立体视频系统包括立体内容的生成、压缩、传输、重构和显示。作为立体视频系统的重要组成部分,并针对当前立体视频内容匮乏的困境,本书着重研究立体视频的生成问题。立体视频的产生本质上是获取视频的三维信息,本书从立体拍摄、立体建模以及平面视频立体转换三个方面详细介绍立体视频产生的原理、方法,同时对于立体视觉产生的机制、主要的立体视频格式、立体失真因素等直接影响立体视频观感的重要问题进行阐述。希望通过本书激发国内更多研究者和专业技术人员从事立体视频领域的工作,为我国在下一代视频技术的国际竞争中占据有利地位做出贡献。

本书的编撰承蒙国家自然科学基金、科技部 863 计划、科技重大专项的支持,得益于清华大学自动化系宽带网数字媒体实验室多年以来师生研究的经验和成果总结,在此一并致谢。

作 者

2015 年年末 于清华园

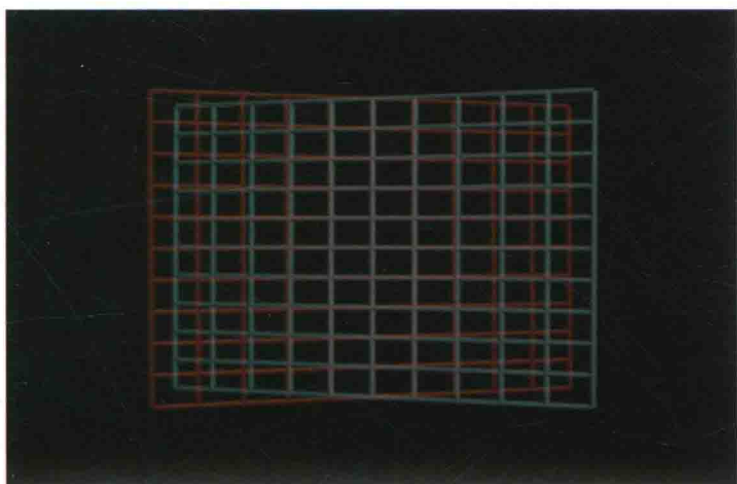


图 2.5 梯形失真



(a) 参考图像

(b) 目标图像

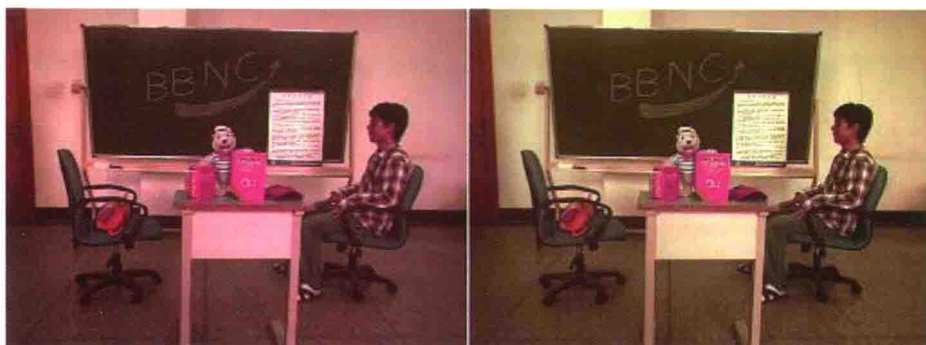
图 2.17 离线颜色校准采用的参考图像和目标图像



(a) 参考图像

(b) 校准后的目标图像

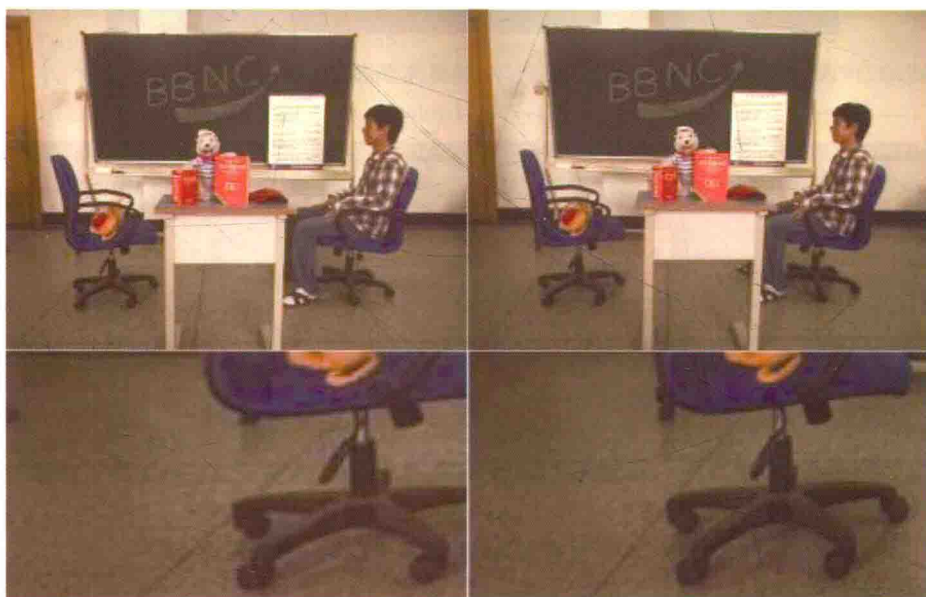
图 2.18 灰色世界法颜色校准结果



(a) 人为改变的目标图像

(b) 校准后的目标图像

图 2.19 灰色世界法颜色校准结果



(a) 参考图像

(b) 校准后的目标图像

图 2.21 直方图互相关法的颜色校准结果



图 3.12(c) 颜色分割图



(a) 参考图像

(b) 校准后的目标图像

图 2.24 特征点能量函数优化法颜色校准结果

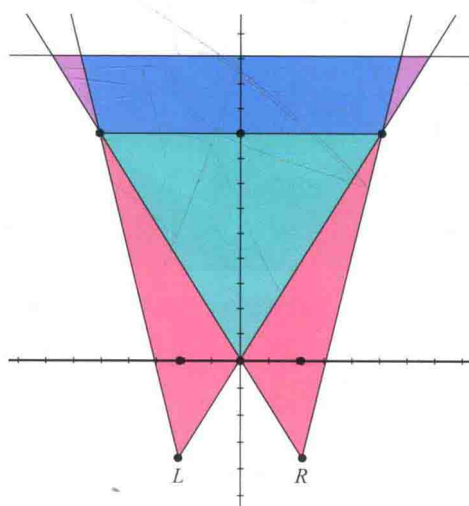


图 7.9 安全区域

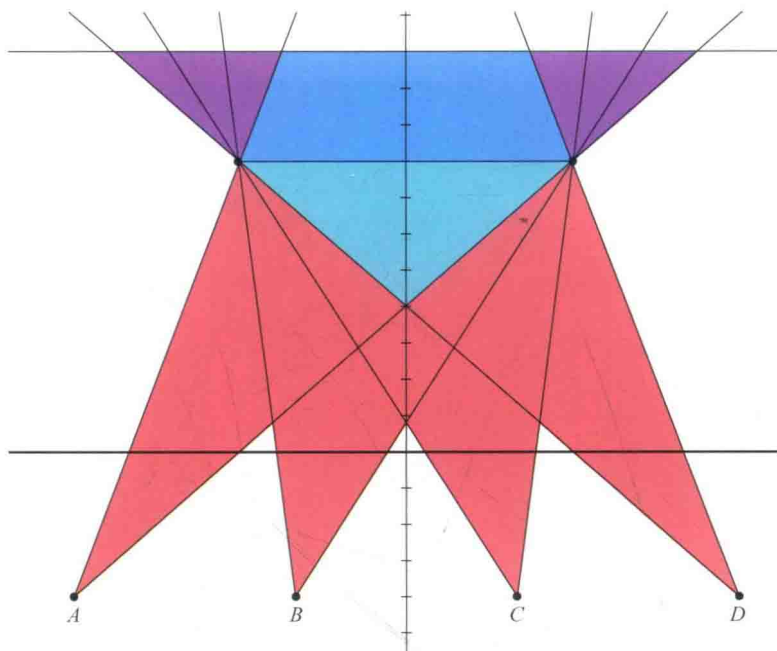


图 7.11 多目渲染安全区域



图 8.1 红青立体图像,利用不同颜色的分色眼镜分别观看左右眼视频

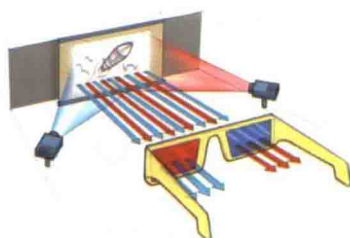


图 8.7 分色眼镜及其工作原理

第 1 章 绪论	1
1.1 立体视频技术的研究背景及意义	1
1.2 立体视频生成技术综述	2
1.3 立体视频压缩编码技术综述	2
1.4 主要研发机构和公司	4
1.4.1 主要研发机构	4
1.4.2 主要研发公司	5
第 2 章 立体视频拍摄技术	8
2.1 立体图像的基本原理	8
2.1.1 人眼立体感知生理学特征	8
2.1.2 立体失真	11
2.2 立体视频采集硬件系统	14
2.2.1 高清同步采集设备	15
2.2.2 数据处理服务器	18
2.3 立体视频采集软件系统	19
2.3.1 相机参数标定	19
2.3.2 离线颜色校准	21
2.4 本章小结	27
第 3 章 平面视频立体转换技术	28
3.1 本章概述	28
3.2 深度恢复算法	28
3.2.1 基于遮挡分析与视觉关注的深度恢复算法	28
3.2.2 基于运动跟踪和运动分析的 2D 转 3D 算法	36
3.3 2D 转 3D 的立体生成、感知和调节方法	45
3.3.1 立体视图的生成	46
3.3.2 立体深度的感知	49
3.3.3 立体效果的调节	51
3.3.4 立体视觉扭曲	53

3.3.5	小结	55
3.4	2D 转 3D 方案分析	56
3.4.1	技术方案和软件系统	56
3.4.2	视频分析与检测	59
3.4.3	深度计算和立体合成	64
3.5	CPU/GPU 加速和实时处理	68
3.6	本章小结	70
第 4 章	立体视频的压缩与编码	71
4.1	高效立体视频编码	71
4.2	具有自由视点的立体视频编码技术	75
4.2.1	多视点视频和深度图的联合编码技术	75
4.2.2	深度辅助的多视点视频编码技术	76
4.2.3	纹理和深度联合的几何宏块划分	84
第 5 章	立体视频编码率失真模型	97
5.1	视频编码率失真模型	97
5.2	区域自适应的虚拟视点绘制失真估计方法	99
5.3	立体视频编码率失真模型	105
5.3.1	编码码率与量化步长之间的关系模型	105
5.3.2	编码失真与量化步长之间的关系模型	108
第 6 章	立体视频编码码率分配	110
6.1	立体视频编码码率分配研究背景	110
6.2	多视点视频与深度图的码率分配	111
6.2.1	立体视频编码码率分配数学模型	111
6.2.2	高效立体视频编码码率分配技术	113
第 7 章	基于模型的立体图像渲染方法	117
7.1	图形学基础	117
7.2	双路立体渲染系统	119
7.2.1	双目成像模型	119
7.2.2	摄像机参数对深度感知的影响	122
7.2.3	安全区域	124
7.2.4	远景处理	125
7.3	任意视角数目立体渲染系统	126
7.3.1	虚拟摄像机组架设方法	126

7.3.2 虚拟摄像机组安全区域	126
7.4 基于现有 CAD 软件的立体渲染方案	127
7.4.1 摄像机设置	127
7.4.2 基于 3ds Max 的立体渲染	128
7.4.3 基于 Unigraphics 的立体渲染	130
7.5 摄像机自标定技术	130
7.6 最小点法的自标定技术概括	131
7.7 本章小结	132
第 8 章 立体终端显示技术	133
8.1 常见立体图像格式	133
8.2 常见立体显示终端	138
第 9 章 总结与展望	144
9.1 立体视频处理技术总结	144
9.2 立体视频处理与通信的未来展望	145
参考文献	147

1.1 立体视频技术的研究背景及意义

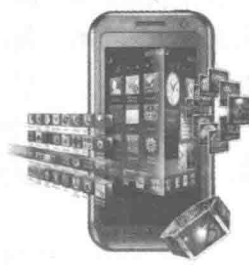
毋庸置疑,现在处于立体产业发展的一个高峰期,与立体产业相关的立体采集、立体重建、立体内容制作以及立体显示器制造工艺等理论和技术日趋成熟。立体技术和产品在动画和游戏开发、电影电视、医疗成像、远程教育和视频会议、工业产品模拟和仿真、广告和广播^[1,2]等多个领域得到了迅速而广泛的应用(如图1.1)。



(a) 立体电视



(b) 立体笔记本



(c) 立体手机

图1.1 立体显示终端

立体视频技术是一种能够提供立体感的视频技术,是继高清视频之后,未来多媒体技术的重要发展方向。近年来,立体视频技术受到学术界与产业界的广泛关注。自1922年第一部立体电影问世以来,为了满足人们对震撼和逼真的视觉享受的不断追求,研究人员一直致力于如何提高立体电影显示效果和改善观赏的舒适度。2010年电影《阿凡达》的成功向人们展示了立体影像独特的魅力,使立体视频走出了实验室,步入大众的视野。3D立体大片风靡全球,预示着立体时代的到来。与传统的平面视频相比,立体视频增加了画面场景的视差、深度信息,从而给观看者带来了三维立体的观感,以及身临其境的观影体验。立体视频的蓬勃发展带来影音播放技术的变革,高清立体视频将取代平面视频,成为多媒体技术的主要

表现形式。借此时机,立体电视、立体手机、立体投影等各类立体终端纷纷走向市场,进入普通百姓的家中。同时,在工业设计、远程医疗、航空航天等领域,立体视频技术也被广泛地应用,其产业价值及学术价值可见一斑。

尽管如此,目前立体产业领域依然存在着这样的矛盾:立体显示设备的飞速普及和立体显示片源的严重匮乏。解决目前立体显示内容短缺的途径大体上有两种,一是设计新的拍摄设备,如双目立体摄像机,利用现有的视觉方法获取立体信息,二是转换已经拍摄好的单目(monocular)的视频。一方面,可供立体显示的内容极度匮乏,另一方面,现有的平面媒体资源,诸如大量的电影和影视资源又十分充足;如何能够获取现有平面媒体资源中的立体信息成了摆在立体研究领域的科学工作者面前的一大难题。各类视频的形成如图 1.2 所示。因此,解决好从现有平面媒体获取立体信息,做到平面视频转立体视频,不仅能够解决目前立体视频产业的巨大瓶颈——立体视频内容匮乏的问题,而且由于问题本身的难度,也具有很高的学术意义。同时,在获取到立体视频内容之后,需要在终端显示之前对其进行相应的表示,由于视频数据量十分庞大,因此本书也涉及最新的立体视频编码和传输的相关内容。



图 1.2 立体视频内容生产的主要方法

1.2 立体视频生成技术综述

立体视频生成技术现阶段主要包括三种途径,一是立体视频拍摄,二是立体建模,三是平面视频立体转换(2D 转 3D)。如图 1.2 所示:立体拍摄通过架设两个摄像机来模拟人的左右眼,分别获取对同一场景、不同位置拍摄的视频信息;立体建模通过将场景和物体的三维信息进行建模,从而直接获得立体视频信息;平面视频立体转换(2D 转 3D)通过对平面视频加入深度的信息获取立体视频。

1.3 立体视频压缩编码技术综述

海量的立体视频数据与有限的传输带宽和存储空间使得立体视频压缩编码技术成为了立体视频发展的主要瓶颈之一。与传统的单目视频相比,立体视频包含的内容广泛。从显示方式来看,立体视频通常包括双目立体视频、多目立体视频、自由视点立体视频以及全息立体视频等等。从数据表示方式来看,立体视频通常又可以分为基于彩色视频加深度图的立体视频、基于表面网格的立体视频、基于纹理映射的立体视频表示方法。上述立体视频的数据格式多种多样,对存储、传输以及显示

的要求也不尽相同。因此,国际国内的研究机构针对不同的立体视频表示方法以及应用方式设计了相应的高效压缩编码方法。其中,最为主要的压缩编码研究包含以下两类。

1. 基于彩色视频加深度图的立体视频的压缩编码方法

基于彩色视频加深度图的立体视频是在传统的单目立体视频的基础上引入多目彩色视频和深度图而形成的。由于其与传统的单目视频系统具有良好的兼容性,因此其在立体视频系统中应用获得了广泛的应用。不难发现,由多目彩色视频和深度图组成的立体视频包含大量相关视频图像,在时域、空域以及视间三个维度引入了大量的冗余。为了有效去除上述立体视频的相关性,研究人员在采用传统的时域预测编码和变换编码来压缩立体视频的时域和空间冗余之外,还引入了视间预测编码来压缩不同视点之间的冗余。为了进一步保证前向兼容性,视间预测编码同样采用时域预测编码技术来实现。目前,该压缩编码方案已经在国际上获得了较大范围的关注和认同。国际标准化领域的两大巨头——国际电信联盟 ITU 和国际标准化组织 ISO,于 2012 年联合成立了相应的标准化工作组,即 Joint Collaborative Team on 3D Video Coding Extension Development 来联合制定基于彩色视频加深度图的立体视频压缩编码技术标准。目前包括 Qualcomm、Samsung、Sony、Sharp、MediaTek 和华为等国际知名公司均参与其中。该组织提出了多项具有不同兼容性的立体视频编码方案,如与传统单目视频编码标准 H. 264 相兼容的 3D-ATM 标准以及与最新的高性能视频编码标准相兼容的 3D-HTM 标准。尽管上述视频编码技术标准还在进一步的讨论中,但是可以预见在不久的将来,在上述组织和公司的推动下,该立体视频编码技术标准将获得更加广泛的应用。

2. 基于表面网格的立体视频的压缩编码方法

作为立体视觉技术发展中另一种重要的立体视频表示方法,表面网格技术在立体电影制作、三维动画、计算机仿真等领域同样获得了广泛的应用。表面网格是通过记录每一个顶点的绝对三维坐标以及与其相邻顶点间的相对位置关系来表示三维场景信息。与基于彩色视频加深度图的立体视频相比,基于表面网格的表示方法能够获得对三维场景更加精确的近似,但同样带来了更大的数据量。三维网格的压缩编码技术经历了较长时间的发展,从静态三维网格压缩到动态三维网格压缩,从单一码率的三维网格压缩到码率可伸缩的三维网格压缩。其中,对于目前应用最为广泛的静态三维网格的压缩编码技术,国际标准化组织 ISO 下所属的运动图像专家组(ISO/IEC JTC1/SC29/WG11, MPEG)正在研究并制定其相应的技术标准,并将在立体视频、三维动画以及计算机仿真等领域获得广泛的应用。尽管在上述研究的过程中取得了相当长足的进步,然而到目前为止,动态三维网格的压缩仍然是一个非常棘手的难题。压缩动态三维网格所带来的较高时间和空间复杂度使得其难以直接应用于立体视频。但是由于其对三维场景、动态三维网格具有其无可替代的

优势,因此,动态三维网格的压缩技术将在未来数十年成为三维网格压缩编码技术的热点之一。

总的看来,伴随着立体视频应用的大量涌现,立体视频的压缩编码技术将成为立体视频走向实际应用中不可或缺的一个重要环节。针对不同类型立体视频的压缩编码技术将在未来获得更加广泛的研究和应用。

1.4 主要研发机构和公司

立体视频技术尚在快速发展阶段,很多国家和地区都在结合自己的优势进行战略布局。欧洲(以飞利浦公司为代表)主要布局在立体内容和立体显示等基础技术领域;美国依靠好莱坞优势(以DDD,In-Three为代表)布局立体内容尤其是立体影视领域;日韩(以Sony、Samsung为代表)布局立体摄像、立体电视、立体手机等应用领域;台湾(以Himax,MStar为代表)则结合自己的集成电路优势布局立体处理芯片尤其是平面视频立体转换(2D转3D)芯片和立体视频处理集成芯片。

1.4.1 主要研发机构

Philips研究院长期致力于3DTV系统的构建,并且主导了2DplusDepth为核心的ATTEST项目,同时Philips公司还是全球立体显示器的主要生产商之一。在2D转3D领域,其代表性产品为基于“Wowvx”技术的半自动2D转3D转换盒,实现输入为平面视频,输出为2DplusDepth格式立体视频的转换功能。

斯坦福大学(Stanford University)的立体重建小组开发了Make3D软件,该软件利用多种深度线索来估计单幅图像的3D结构。在Make3D的方案中^[26],他们首先将单幅图像分成小块,同时对视觉深度线索和图像各部分之间的相对关系建模,然后利用马尔可夫随机场(MRF)来推导一组“平面参数”,这些参数即包含了各图像块的3D位置和方向数据。在该方案中,除了图像场景结构由各个小平面组成这一假定,没有对图像场景进行其他显性的结构假定,这使得该算法对其他未训练过的场景也能构建出相当准确的3D结构。

台湾大学电机系陈良基教授领导的感知视觉实验室与奇景光电(Himax)合作,利用人眼感知特性,研发了应用于立体电视的全自动2D转3D技术,并利用奇景光电的芯片制造和代工实力,目前已经推广到部分立体电视产品中。

清华大学BBNC实验室的2D转3D小组^[27]同时研发了半自动2D转3D软件 and 全自动2D转3D软件,其中半自动软件已经被江苏广电、景象等多家单位所采用,而全自动软件则已嵌入到美国Sonic公司的视频播放和编辑软件“Roxio Creator”中,并已经在全球销售。目前清华大学也已经开发出“清立方”全自动2D转3D芯片和机顶盒,并与海尔、富士通、长虹等多家厂商合作进行市场推广。

浙江大学、上海大学、宁波大学等多家大学也在立体内容制作、立体传输、立体显示等 3D 相关领域进行持续研究并取得了一定的科研成果。整体而言,我国在立体视频相关领域核心竞争力有待发展,但立体内容制作和立体视频处理领域尚在快速发展阶段,技术上还大有可为,产业化需求十分迫切。

1.4.2 主要研发公司

调研发现,进行立体视频内容生成技术的研发企业比较多,主要集中在视频播放和编辑软件、立体电视、立体笔记本、蓝光播放器、机顶盒等产品研发领域。主流电视厂商,比如 Samsung、LG、Sony、Fujitsu 等均通过独立研发或技术合作方式,生产具有 2D 转 3D 功能的立体电视。主流芯片、FPGA、Soc 以及专业数字媒体处理器提供商也在研发自己的 2D 转 3D 模块,并利用自己代工或市场占有优势推广,国际竞争比较激烈。

美国 HDLogix 公司开发了一种“ImageIQ3D”实时 2D 转 3D 方案,该方案^[28]宣称利用了多个深度线索。例如:通过光流法运动估计提取场景的运动深度信息;通过霍夫(Hough)或者雷登(Radon)变换提取场景建筑物的竖直线、自然远景的水平线等来寻找几何深度渐变趋势;使用点扩散函数估计器(Point Spread Function Estimator)来估计散焦模糊区域的位置。最后,ImageIQ3D 还宣称利用了一种基于超分辨率的统计方案来得到鲁棒和连续的深度结果。

日本 JVC 公司设计了一种面向立体电视信号的实时全自动 2D 转 3D 方法并已在硬件平台实现^[29],该方法建立了三种简单的立体模型来模拟近似的自然场景,并利用三种模型和图像颜色的冷/暖视觉差异理论指导每一帧视频信号的深度图求取。该方法尽量避免每一帧视图的质量损失,计算复杂度低,并且对硬件的存储空间要求较低,也基本不需要人工交互和参数调节,但是该方法也存在模型数量较少,深度图求取不准确,立体效果不明显等问题。

表 1.1 和表 1.2 展示了当前典型的立体视频技术研发企业、相关产品及应用。通过调研也可发现,目前实时全自动转换的效果不能跟专业制作的立体电影相比,因此如单独做成产品,其竞争力和生命期有限。基于市场考虑,必须跟多媒体软件、数字电视或其他终端集成为完整的解决方案。

表 1.1 硬件研发企业简介

公 司	公 司 类 型	芯片功能、参数	市 场 应 用
SANYO(三洋电机)	家电厂商	1998 年研发,LSI,从影像颜色、帧时延中提取深度信息	三洋电视
Himax(奇景光电)	芯片厂商	与台大电机系合作,近期准备发布第二版算法(基于人类视觉感知特性)	三星、索尼、夏普等立体电视
Mstar	芯片厂商	图像平移技术,SOC 芯片设计,集成到数字音视频处理芯片	长虹、海信

续表

公 司	公 司 类 型	芯片功能、参数	市 场 应 用
Altera	芯片厂商	将 DDD 公司的 TriDef Core 嵌入式 3D 图像处理器运行在 Altera Arria® GX FPGA 上	三星、现代等立体电视
Dialog	半导体公司	全球首款 2D/3D 影像转换实时处理芯片,型号: DA8223. 图像分析,分离前景和背景	带光栅立体显示的手机和平板电脑
ECT inc	芯片厂商	持续研发 ECT223 系列芯片,并开始研发 2D-Multiview	LG 立体电视
Panasonic(松下)	家电厂商	FPGA,对图像分块分析,可手动调整视差和立体感强弱	2D-3D 专用转换器(JVC)
Thoshiba(东芝)	家电厂商	与 IBM 合作研发处理器 CELL Broadband Engine,并研发 Tri-vector 2D-3D 技术	Thoshiba 3D“CELL”TV
Sony(索尼)	家电、数码厂商	芯片	3D 手机、电脑、XBR-HX900 系列 3DTV
Trident Microsystems(泰鼎)	机顶盒与电视半导体解决方案供应商	将 2D to 3D 功能集成到芯片 PNX5130 中,具有简单的内容识别功能	第三大的数字电视芯片生产商
Wistron(纬创资通)	全球最大 OEM 专业代工	与 DDD、宏基合作开发全球首款 3D NB,具有 2D to 3D 功能	2009 年 11 月推出 Aspire5738DG
InnoVision(收购 e-MDT)	显卡制造、OEM/ODM	开发 EM3100 芯片,实现 2D-3D 实时处理	未知
Sensio	3D 芯片研发	S3D220 ASIC 芯片	Vizio(台湾瑞轩)3DTV
Quartics	半导体公司	开发 QV1721 视频处理器,集成有 2D-3D 功能	宏基数码产品
Marvell	IC 代工厂	芯片集成 2D/3D 功能	数字电视芯片提供商

表 1.2 软件研发企业简介

软 件 公 司	公 司 类 别	2D-3D 功能、参数	产 品 应 用
DDD	3D 内容制作	动态深度线索提取	Tridef 系列软件
Roxio	数字媒体服务	清华大学研发	Cineplay 立体播放器
ArcSoft	数字媒体服务	高清实时转换	Sim3D 立体播放器
Cyberlink	数字媒体服务	高清实时转换	PowerDVD 立体播放器
HDLogix	数字媒体服务	CPU/GPU 加速,各种深度线索	ImageIQ3D 立体播放器
3D Mention LLC	3D 软件、芯片研发	全自动 2D-3D	手机软件、芯片
StereoD	3D 影视制作	开发 VDX 半自动软件	参与《阿凡达》、《最后的气宗》等影视制作
InThree, Sony Image Works Pictures	3D 影视制作	非全自动、非实时;采用“Roto”以及人海战术	《爱丽丝梦游仙境》立体版等