



多视点视频编码 的关键技术

THE KEY TECHNOLOGY OF
MULTI-VIEW VIDEO CODING

■ 孟丽丽 谭艳艳 张化祥 著



中国工信出版集团



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

多视点视频编码的关键技术

孟丽丽 谭艳艳 张化祥 著

電子工業出版社

Publishing House of Electronics Industry

北京 · BEIJING

内 容 简 介

本书系统地介绍了三种多视点视频压缩技术,即多边信息的分布式多视点视频编码、兼容标准的高效立体视频编码和鲁棒的多描述多视点视频编码。本书的研究内容对3D视频的压缩和鲁棒传输具有重要理论指导价值。

本书可以作为高等院校电子信息、通信工程、计算机等相关专业的教师和学生的参考用书,也可供有关工程技术人员参考阅读。

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

多视点视频编码的关键技术/孟丽丽,谭艳艳,张化祥著. —北京:
电子工业出版社, 2018.9

ISBN 978-7-121-35076-4

I. ①多… II. ①孟… ②谭… ③张… III. ①视频编码—研究 IV. ①TN762

中国版本图书馆CIP数据核字(2018)第218173号

策划编辑:徐蔷薇

责任编辑:刘小琳 特约编辑:孙悦

印刷:北京虎彩文化传播有限公司

装订:北京虎彩文化传播有限公司

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:720×1000 1/16 印张:9.5 字数:115千字

版次:2018年9月第1版

印次:2018年9月第1次印刷

定价:58.00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)88254888,88258888。

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。

本书咨询联系方式:(010)88254438。

前 言

近年来,视频显示技术已经从 2D 视频发展到 3D 立体视频。多视点视频是 3D 视频一种重要的表述格式。多视点视频是由同一时刻不同角度的摄像机对同一场景采集的视频序列,它能更生动、准确地呈现场景。然而,多视点视频的数据量很大,其存储和传输都非常困难,如何高效和鲁棒地压缩多视点视频数据已成为当前视频编码领域的研究热点。

本书围绕多视点视频编码,对多边信息的分布式多视点视频编码、兼容标准的高效立体视频编码和鲁棒的多描述多视点视频编码几种视频压缩技术进行了深入研究,主要创新性研究成果如下。

(1) 研究多边信息的分布式多视点视频编码方法。该编码方法能降低编码端的复杂度、避免摄像机之间的通信,还可以更好地利用边信息的相关性;同时,基于贝叶斯准则研究多边信息的联合条件概率密度函数,并将其应用在视频编码方法中。实验结果证明,该方法能有效提高编码效率。

(2) 研究兼容标准的高效立体视频编码方法。现有的绝大多数编码器都还是针对 2D 视频的,无法直接处理多视点视频。因此,本书在现有视频编码标准的基础上,研究具有灵活预测模式和自适应预测结构的兼容标准的高效立体视频编码方法,该方法有效地提高了率失真性能。

(3) 研究鲁棒的多描述多视点视频编码方法。传统多视点视频编码码流一旦遇到网络丢包或比特传输错误,解码质量就会严重下降。具有鲁棒性的多描述多视点视频编码方法能有效地解决这个问题。首先,针对多视点帧内编码,研究基于随机偏移量化器和均一偏移量化器的多描述多视点帧内编码;然后,针对多视点帧间编码,研究基于内插补偿预处理的多描述多视点视频编码,该方法能有效地重建丢失块或有误差块的信息,从而进一步提高块的重建质量;最后,通过实验及与相关算法的比较,证明该方法的鲁棒性和有效性。

感谢北京交通大学赵耀教授、加拿大西蒙弗雷泽大学 Jie Liang 教授、福建工程学院潘正祥教授及山东师范大学信息科学与工程学院对笔者研究工作的支持;同时,感谢国家自然科学基金青年项目(项目编号:61402268)和国家自然科学基金面上项目(项目编号:61572298)对本书研究工作的资助。

由于笔者学识水平所限,本书难免存在不足之处,恳请各位专家和广大读者给予批评指导。

著 者

2018 年 6 月

目 录

1

绪论	1
1.1 研究背景与研究意义	3
1.2 多视点视频编码的研究现状	9
1.2.1 基于传统 2D 视频编码的多视点视频编码研究	10
1.2.2 基于运动估计和视差估计的多视点视频编码研究 ..	11
1.2.3 基于合成视点预测的多视点视频编码研究	14
1.2.4 分布式多视点视频编码研究	17
1.2.5 基于多描述编码的多视点视频编码研究	19
1.3 本书内容安排与组织结构	20

2

多边信息的分布式多视点视频编码	25
2.1 引言	27
2.2 多边信息分布式视频编码	30
2.2.1 系统描述	30
2.2.2 最优重建	33

2.2.3 基于贝叶斯准则的联合条件概率密度函数和 相关噪声模型	35
2.3 基于贝叶斯准则的多边信息分布式多视点视频编码	38
2.4 实验结果	40
2.4.1 不同二进制码的实验结果	40
2.4.2 多边信息分布式视频编码的实验结果	43
2.4.3 多边信息分布式多视点视频编码的实验结果	48
2.5 本章小结	49

3

兼容标准的高效立体视频编码	51
3.1 引言	53
3.2 基于 H.264 的传统立体视频编码	55
3.3 基于灵活预测模式兼容标准的立体视频编码	57
3.3.1 灵活预测模式	58
3.3.2 基于灵活预测模式的立体视频编码系统	60
3.4 基于自适应预测结构兼容标准的立体视频编码	62
3.4.1 自适应预测结构	62
3.4.2 基于自适应预测结构的立体视频编码系统	65
3.5 实验结果	66
3.6 本章小结	70

4

鲁棒的多描述多视点视频编码	73
4.1 引言	75
4.2 基于随机偏移量化器的多描述多视点帧内 编码	81
4.2.1 MDROQ 系统描述	81
4.2.2 MDROQ 期望失真的一般表达式	83
4.3 基于均一偏移量化器的多描述多视点帧内 编码	87
4.3.1 低码率联合重建方法的比较	88
4.3.2 MDUOQ 系统描述	90
4.4 多描述多视点帧内编码的优化	92
4.4.1 系统框图及 DCT 域的维纳滤波	92
4.4.2 期望失真的建模	96
4.4.3 迭代优化算法	98
4.4.4 均一偏移量化器死区间隔的优化	99
4.5 鲁棒的多描述立体视频编码	100
4.5.1 多描述立体视频编码的系统设计	100
4.5.2 内插补偿预处理算法	103
4.6 理论分析	105
4.6.1 MDROQ 与 MDLTPC 的理论比较	105

4.6.2 MDROQ 的 1D 数据理论界限和实验结果 107

4.7 实验结果 109

4.7.1 多描述多视点帧内编码的实验结果 109

4.7.2 鲁棒的多描述立体视频编码实验结果 118

4.8 本章小结 121

5

总结与展望 123

5.1 总结 125

5.2 展望 125

参考文献 127

1

绪 论

1.1 研究背景与研究意义

随着实时场景采集、传输、演示的迫切需求和信号处理技术的发展, 视频显示技术已经从 2D (2 Dimensional) 发展到 3D (3 Dimensional) [1,2], 3D 电影、3D 电视等开始进入人们的生活[3,4]。

除了 3D 视频处理技术的进步, 3D 视频的内容也越来越丰富。3D 电影的数量每年都在增加, 其票房在影院总票房中已占很高比例。几家著名的电影制片厂声称, 未来他们只制作和发售 3D 电影; 一些主要的投资者也已经决定升级 3D 电影设备。3D 电影的票房和投资力度说明消费者对 3D 视频的接受和喜爱。然而, 当前 3D 视频普及到日常生活中还是比较困难的。首先, 当前观看 3D 视频需要佩戴特殊的眼镜, 这是非常不方便的; 其次, 观看者长时间观看 3D 视频, 其眼睛会不舒服。

3D 视频通常有两种数据表示格式: 一种是包含 n ($n \geq 2$) 个视点的多视点视频 (Multiview Video, MVV); 另一种是结合深度信息的单路或多路视点视频 (Multiview Video plus Depth, MVD)。本书研究的是多视点视频 (不包含深度信息)。多视点视频是由同一时刻不同角度的摄像机对同一场景采集的视频序列, 它能更生动准确地呈现场景[5]。图 1.1 所示为多视点视频实例——Ballroom, 其中 $X_{i,j}$ 表示第 i 个视点在第 j 个时刻的视频帧, 该视频序列是由图 1.2 所示的采集系统 (平行

排列的摄像机)获得的。这个采集系统是由三菱电子研究实验室搭建的^[6]。目前,其他一些机构也构建了实际的多视点视频系统。例如,日本名古屋大学搭建了包含100个摄像机的采集系统,该采集系统中的摄像机有1D线形、1D弧形和2D矩阵排列三种排列方式(见图1.3);斯坦福大学构建了52个摄像机矩阵系统[见图1.4(a)]^[7];微软亚洲研究院也已经发展了32个摄像机的实时多视点视频系统[见图1.4(b)]^[8],其中一些技术已经应用在MPEG 3DAV中^[9];笔者所在研究团队搭建了10个摄像机的采集和控制平台(见图1.5),摄像机的型号为Stingray F046C,分辨率为 780×582 像素,帧率为60fps。图1.5(a)为4个水平排列的摄像机;图1.5(b)为6个环形排列的摄像机,其中的环形升降支架,可以有效调节摄像机的位置。此外,该平台配备了高性能工作站、同步触发模块、磁盘阵列及相机控制器等,可以实现同步多视点视频采集。

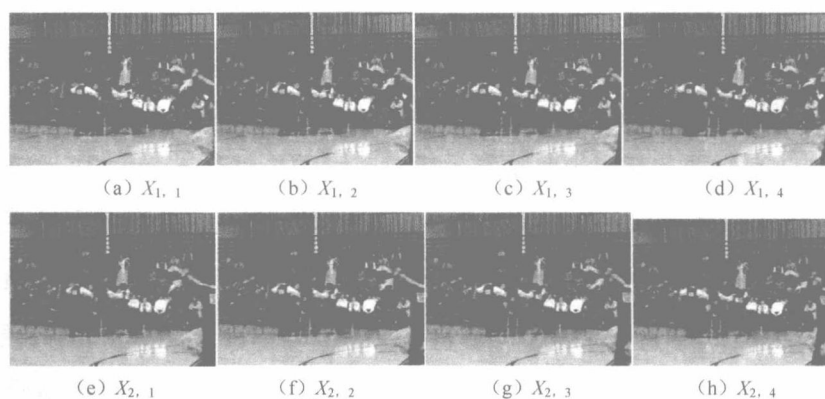


图 1.1 多视点视频实例——Ballroom

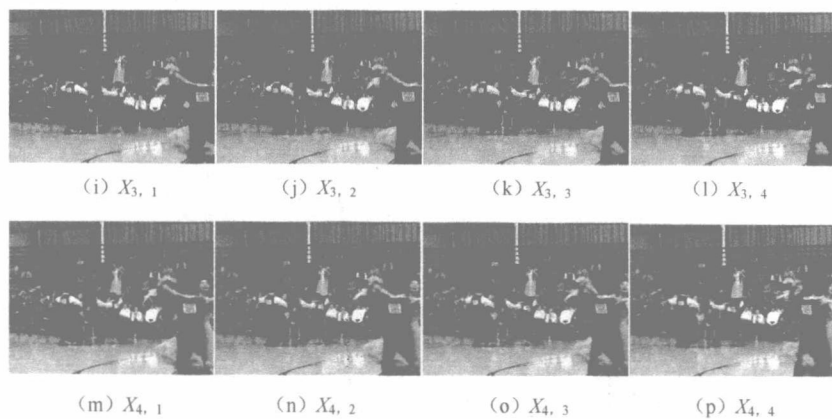


图 1.1 多视点视频的实例——Ballroom (续)

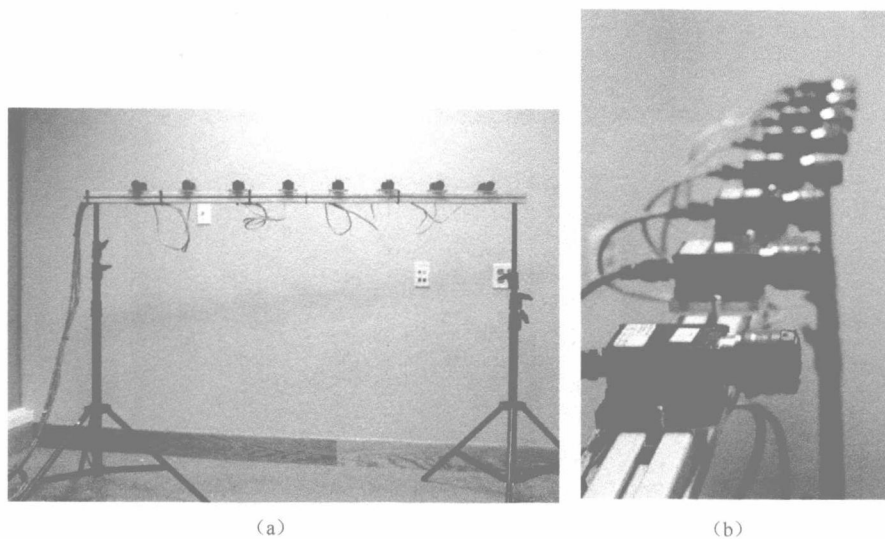
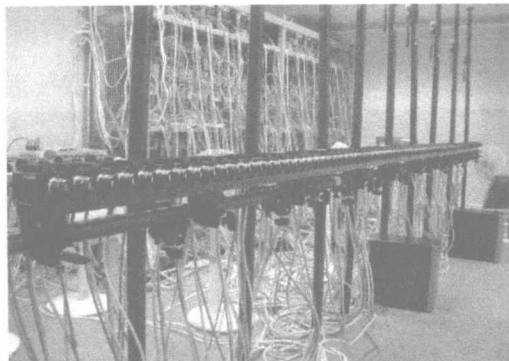


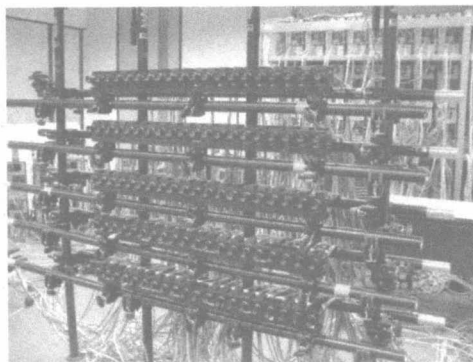
图 1.2 平行排列的摄像机



(a) 1D线形

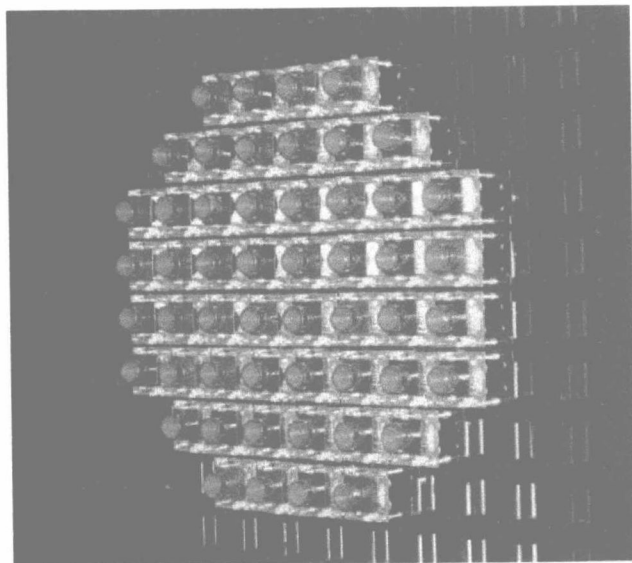


(b) 1D弧形

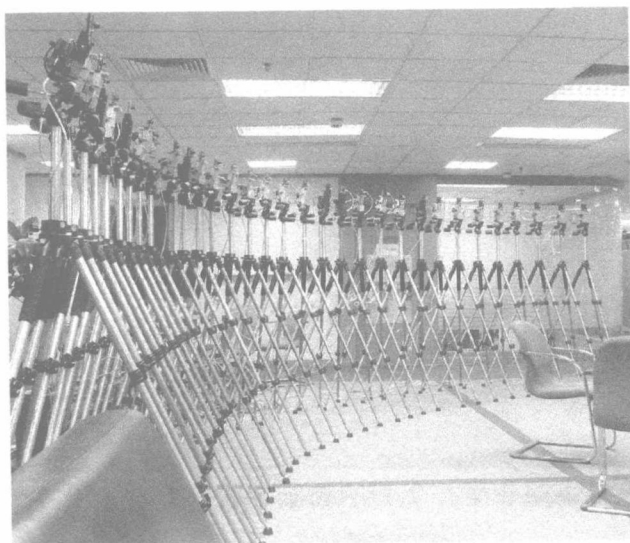


(c) 2D矩阵排列

图 1.3 100 个摄像机的三种排列形式

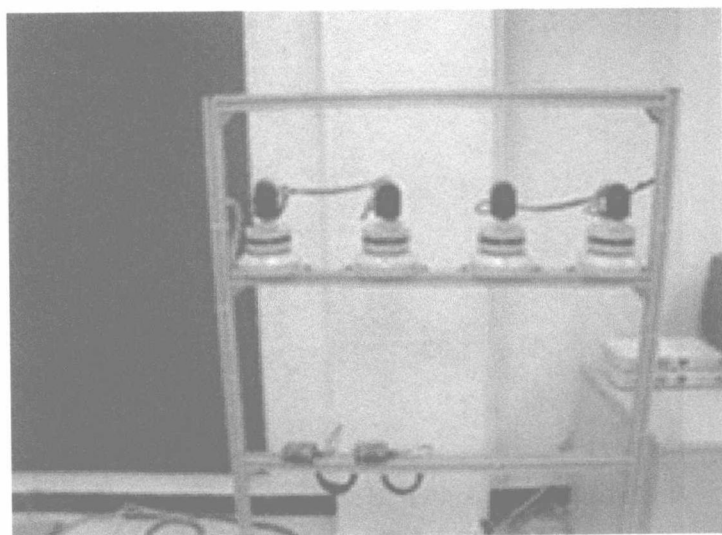


(a) 52个摄像机的矩阵排列

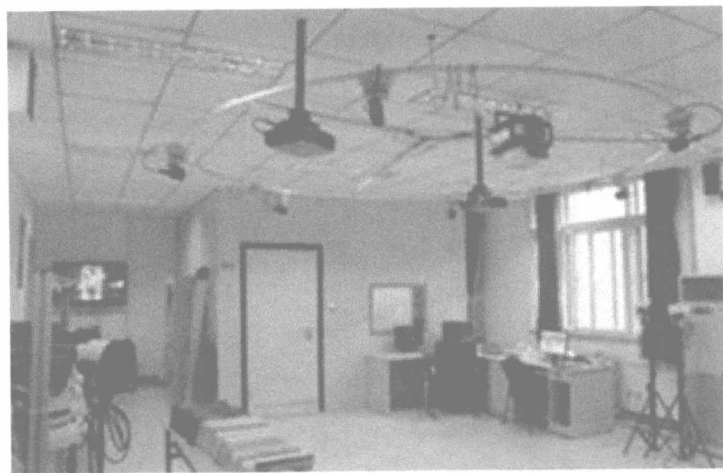


(b) 32个摄像机的弧度排列

图 1.4 摄像机的排列



(a) 4个水平排列的摄像机



(b) 6个环形排列的摄像机

图 1.5 笔者所在研究团队搭建的采集设备