



高等院校
通信与信息专业规划教材

数字视频信息 处理与传输教程

刘富强 主编



机械工业出版社
CHINA MACHINE PRESS



高等院校通信与信息专业规划教材

数字视频信息处理与传输教程

刘富强 主编

王新红 钱建生 刘 云 王传旭 编著



机械工业出版社

本书介绍了数字视频的基础、数字视频编码、数字视频信息处理、数字视频信息检索、数字视频信息传输等内容,较全面地介绍了数字视频技术及应用。

本书内容详尽,重点突出,注重理论和实践的结合,每章后还配有习题供读者练习。本书可作为信息工程、计算机、通信、自动化等专业高年级本科生和研究生教材或教学参考书,也可作为工程技术人员的参考资料或培训教材。

图书在版编目(CIP)数据

数字视频信息处理与传输教程/刘富强主编. —北京:机械工业出版社, 2004.7

(高等院校通信与信息专业规划教材)

ISBN 7-111-14787-1

I. 数... II. 刘... III. 视频信号—数字技术—高等学校—教材
IV. TN941.3

中国版本图书馆CIP数据核字(2004)第061446号

机械工业出版社(北京市百万庄大街22号 邮政编码 100037)

策划:胡毓坚

责任编辑:蔡岩

责任印制:石冉

三河市宏达印刷有限公司印刷·新华书店北京发行所发行

2004年8月第1版·第1次印刷

787mm×1092mm $\frac{1}{8}$ ·18.25印张·451千字

0001—5000册

定价:26.00元

凡购本图书,如有缺页、倒页、脱页,由本社发行部调换

本社购书热线电话:(010) 68993821、88379646

封面无防伪标均为盗版

高等院校通信与信息专业规划教材

编委会名单

(按姓氏笔画排序)

编委员会主任	乐光新			
编委会副主任	张文军	张思东	杨海平	徐澄圻
编委会委员	王金龙	冯正和	刘增基	李少洪
	邹家禄	吴镇扬	赵尔沅	南利平
	徐惠民	彭启琮	解月珍	
秘书长	胡毓坚			
副秘书长	许晔峰			

出版说明

为了培养 21 世纪国家和社会急需的通信与信息领域的高级科技人才,为了配合高等院校通信与信息专业的教学改革和教材建设,机械工业出版社会同全国在通信与信息领域具有雄厚师资和技术力量的高等院校,组成阵容强大的编委会,组织长期从事教学的骨干教师编写了这套面向普通高等院校的通信与信息专业规划教材,并且将陆续出版。

这套教材将力求做到:专业基础课教材概念清晰、理论准确、深度合理,并注意与专业课教学的衔接;专业课教材覆盖面广、深度适中,不仅体现相关领域的最新进展,而且注重理论联系实际。

这套教材的选题是开放式的。随着现代通信与信息技术日新月异地发展,我们将不断更新和补充选题,使这套教材及时反映通信与信息领域的新发展和新技术。我们也欢迎在教学第一线有丰富教学经验的教师及通信与信息领域的科技人员积极参与这项工作。

由于通信与信息技术发展迅速而且涉及领域非常宽,这套教材的选题和编审难免有缺点和不足之处,诚恳希望各位老师和同学提出宝贵意见,以利于今后不断改进。

机械工业出版社
高等院校通信与信息专业规划教材编委会

前 言

人类获得信息的 75% 来自人的视觉,在计算机网络上传输的视频信息只能是数字化的视频,即数字视频。本书介绍了数字视频基础、数字视频编码、数字视频信息处理、数字视频信息检索、数字视频信息传输等内容,较全面地介绍了数字视频技术及应用。

数字视频基础包括第 1~3 章,分别介绍全书的概述、数字视频的描述、数字视频的采集。

数字视频编码包括第 4~5 章,分别介绍了数字视频图像压缩的基本方法和图像压缩标准等。

数字视频信息处理包括第 6~9 章,分别介绍了二维运动估计、三维运动估计、数字视频滤波、数字视频水印等。

数字视频信息检索包括第 10~12 章,分别介绍了视频信息检索、视频信息检索标准、基于内容检索的应用系统。

数字视频信息传输包括第 13~16 章,分别介绍了数字视频传输网络、通信协议、流媒体技术、数字视频信息传输的应用。

本书内容详尽,重点突出,注重理论和实践的结合,每章后有习题供读者练习。本书可作为信息工程、计算机、通信、自动化等专业高年级本科生和研究生教材或教学参考书,也可作为工程技术人员的参考资料或培训教材。

同济大学王新红博士撰写了第 3、9、10、11、12 章,中国矿业大学钱建生教授/博士撰写了第 13、14、15、16 章,青岛科技大学刘云教授/博士撰写了第 4、5 章,王传旭博士撰写了第 6、7、8 章,同济大学刘富强教授撰写了第 1、2 章,全书由刘富强教授组织编写并统稿。

感谢张祥光博士、李世银博士、程德强博士,硕士研究生朱春生、王佑贞、李庆、曾智慧、张欢等也参与了本书的辅助性工作。

感谢上海市自然科学基金(02ZD14105)等的资助。感谢同济大学教材、学术著作出版基金委员会资助。

由于作者水平有限,疏漏和错误之处在所难免,敬请各位同行专家、读者批评指正。

刘富强

目 录

出版说明

前言

第1章 绪言 1

1.1 数字视频概述 1

1.2 本书章节安排 2

第2章 数字视频描述 4

2.1 时变图像构成模型 4

2.1.1 三维运动模型 4

2.1.2 几何投影 5

2.1.3 成像的光学模型 6

2.2 模拟视频 6

2.2.1 电视的扫描 6

2.2.2 彩色广播电视信号 8

2.2.3 视频信号的形式 9

2.3 数字视频 10

2.3.1 数字视频基础 10

2.3.2 数字视频的结构 12

2.4 数字视频格式 16

2.4.1 MPEG 数字视频格式 16

2.4.2 AVI 数字视频格式 17

2.4.3 MOV 数字视频格式 18

2.4.4 ASF 数字视频格式 19

2.4.5 RA/RM 数字视频格式 20

2.5 习题 21

第3章 数字视频采集 23

3.1 图像的视觉基础 23

3.1.1 图像的概念 23

3.1.2 分辨率 23

3.1.3 眼的视觉特性 24

3.2 时空采样 24

3.2.1 模拟和数字视频的采样 25

3.2.2 二维结构采样 26

3.2.3 二维周期采样 27

3.2.4 三维结构上的采样 28

3.3 采集图像的量化 30

3.3.1 均匀量化 31

3.3.2 非均匀量化 32

3.4 采样结构的转换 33

3.4.1 一维信号的插值与抽取 33

3.4.2 采样点阵转换 38

3.5 CCD 图像传感器 39

3.5.1 工作原理与结构 39

3.5.2 主要技术参数 41

3.5.3 应用 42

3.6 视频采集卡 45

3.6.1 视频采集卡简介 45

3.6.2 视频采集卡原理概述 45

3.6.3 视频采集卡的分类 47

3.6.4 视频采集卡的技术指标 47

3.7 习题 48

第4章 数字视频图像压缩 51

4.1 图像数据压缩编码概述 51

4.1.1 数字视频图像压缩的必要性和可能性 51

4.1.2 图像压缩系统的组成 53

4.1.3 视频图像压缩编码的主要方法及其分类 53

4.2 视频压缩的评价指标和图像质量的评判标准 55

4.2.1 视频压缩的评价指标 55

4.2.2 图像质量的评判标准 56

4.3 变换编码 57

4.3.1 变换编码基本原理 57

4.3.2 K-L 变换 57

4.3.3 离散余弦变换(DCT)编码 59

4.3.4 子带编码 59

4.4 预测编码与帧间压缩方法 60

4.4.1 预测编码 60

4.4.2 运动估值与运动补偿	62	6.2.2 光流估计法中克服孔径问 题的几种方法	110
4.5 小波变换编码	65	6.3 基于像素的运动估计	112
4.5.1 傅里叶(Fourier)分析和小 波分析	65	6.3.1 用运动平滑约束正规化	112
4.5.2 小波变换在图像压缩编码 中的应用	68	6.3.2 使用多点邻域	112
4.6 模型基编码	71	6.3.3 像素递归法	113
4.7 分形编码	72	6.4 基于块的运动估计	114
4.7.1 分形编码概述	72	6.4.1 块运动的两种模型	114
4.7.2 数字图像的分块式分形编码	77	6.4.2 相位相关法	116
4.7.3 运动图像的分形压缩方法	80	6.4.3 块匹配法	117
4.7.4 关于分形图像编码研究的新 进展	82	6.4.4 通用块运动估计	119
4.8 基于对象的视频编码	82	6.5 全局运动估计	121
4.8.1 基于对象的视频编码框架	83	6.5.1 健壮估计器	121
4.8.2 MPEG-4 视频编码	83	6.5.2 直接估计	121
4.8.3 MPEG-4 的应用领域	92	6.5.3 间接估计	122
4.9 习题	93	6.6 基于区域的运动估计	122
第5章 图像压缩标准	95	6.6.1 基于运动的区域分割	123
5.1 动态图像压缩标准 H.261	95	6.6.2 联合区域分割和运动估计	123
5.2 低码率视频压缩标准 H.263	96	6.7 习题	124
5.2.1 H.263 的图像数据结构	97	第7章 三维运动估计	125
5.2.2 H.263 的 PB 帧模式	98	7.1 三维运动估算中的基本概念	125
5.2.3 H.263 的运动补偿和高 级预测方式	99	7.1.1 正交位移量模型	125
5.3 数字视频压缩标准 MPEG-1	100	7.1.2 透视位移场模型	126
5.3.1 MPEG-1 视频的编码	100	7.2 基于特征的三维运动估计	126
5.3.2 MPEG-1 视频的解码	102	7.2.1 正交投影下已知形状的物体	126
5.3.3 MPEG-1 系统	102	7.2.2 透视投影下已知形状的物体	127
5.4 数字视频压缩标准 MPEG-2	104	7.2.3 平面物体	128
5.5 多媒体应用标准 MPEG-4	105	7.2.4 使用外极线的未知形 状的物体	128
5.6 习题	106	7.3 基于点对应法的运动估计	132
第6章 二维运动估计	107	7.3.1 基于正交模型的三维 运动估计方法	132
6.1 二维运动估算中的基本概念	107	7.3.2 基于透视模型的三维运 动方法	134
6.1.1 二维运动与视在运动	107	7.4 直接运动估计	137
6.1.2 二维运动场模型与二维 运动估算	108	7.4.1 图像信号模型和运动	137
6.2 基于光流方程的运动估计	109	7.4.2 已知形状的物体	138
6.2.1 光流方程的推导	109	7.4.3 对平面物体的三维估计	139
		7.4.4 健壮估计	140

7.5 习题	142	10.5 习题	180
第 8 章 数字视频滤波	143	第 11 章 视频信息检索标准	181
8.1 按运动轨迹的补偿滤波	143	11.1 概述	181
8.1.1 任意运动轨迹	143	11.2 MPEG-7	182
8.1.2 恒速整体运动	143	11.2.1 MPEG-7 的术语	182
8.1.3 加速运动	144	11.2.2 MPEG-7 的组成	183
8.2 时空滤波去噪方法	145	11.2.3 MPEG-7 多媒体描述方案	184
8.2.1 帧间滤波及其实现方法	145	11.2.4 MPEG-7 视频	185
8.2.2 自适应运动补偿滤波	146	11.2.5 MPEG-7 与基于内容检索 的关系	187
8.3 视频图像的多帧维纳去卷积	148	11.2.6 MPEG-7 的应用	188
8.4 习题	150	11.3 MPEG-21	188
第 9 章 数字视频水印技术	151	11.3.1 MPEG-21 的基本概念	189
9.1 概述	151	11.3.2 MPEG-21 的目标和用户	189
9.1.1 数字水印技术的产生背景	151	11.3.3 MPEG-21 的组成要素	190
9.1.2 数字水印技术的基本思想	151	11.3.4 数字项声明 DID	191
9.1.3 数字水印技术的特性	152	11.3.5 数字项标识 DII	192
9.2 数字水印技术的基本原理	153	11.3.6 知识产权管理和保护 IPMP	193
9.2.1 数字水印的嵌入算法	153	11.3.7 MPEG-21 的关键问题	194
9.2.2 数字水印的提取和检测算法	155	11.4 习题	195
9.3 数字水印技术的分类	156	第 12 章 视频信息检索的应用系统	196
9.4 数字水印技术的主要应用	159	12.1 概述	196
9.5 习题	160	12.2 QBIC 系统	197
第 10 章 视频信息检索	161	12.3 Visualseek 系统	198
10.1 概述	161	12.4 Photobook 系统	199
10.1.1 视频信息检索的必要性	161	12.5 VideoQ 系统	200
10.1.2 传统信息检索的局限性	162	12.6 其他系统	202
10.1.3 视频信息检索的发展过程	162	12.7 习题	202
10.2 基于内容的视频信息检索	163	第 13 章 数字视频传输网络技术	203
10.2.1 系统结构	164	13.1 多媒体视频对通信网的 要求	203
10.2.2 特点	165	13.1.1 视频对网络的要求	203
10.2.3 应用领域	166	13.1.2 音频对网络的要求	204
10.3 基于内容的静止图像检索	166	13.1.3 图像传输对网络的要求	204
10.3.1 颜色特征	166	13.1.4 RSVP 与多媒体通信	205
10.3.2 纹理特征	169	13.2 宽带视频传输网络技术	208
10.3.3 形状特征	173	13.2.1 变速率数字视频技术	208
10.4 基于内容的运动视频检索	176	13.2.2 宽带交换技术——ATM	209
10.4.1 镜头检测	176	13.2.3 SDH 传输技术	216
10.4.2 关键帧提取	178		
10.4.3 运动信息的提取和应用	179		

13.3 宽带接入技术	219	14.5.1 IPv6 产生的背景	245
13.3.1 100Mbit/s 快速以太网	219	14.5.2 IPv6 的设计目标	245
13.3.2 高速局域网——FDDI	220	14.5.3 IPv6 增强特性	245
13.3.3 高速数字用户线 HDSL	221	14.5.4 IPv6 数据报文格式	246
13.4 宽带无线接入(LMDS)技术	222	14.5.5 IPv6 的地址空间	247
13.4.1 LMDS 概述	222	14.5.6 IPv6 的扩展首部	250
13.4.2 LMDS 业务和特点	223	14.6 习题	253
13.4.3 LMDS 网络组成	224	第 15 章 流媒体技术	254
13.5 视频传输宽带 IP 技术	225	15.1 基本概念	254
13.5.1 IP 技术简介	225	15.1.1 流式传输	254
13.5.2 ATM 和 IP 技术	226	15.1.2 流媒体技术原理	255
13.5.3 IP over SDH	228	15.1.3 流媒体文件格式与体系结构	256
13.5.4 IP over WDM	229	15.2 流媒体播放方式	258
13.6 习题	229	15.2.1 基于 IP 的基本播放方式	258
第 14 章 通信协议	230	15.2.2 IP 组播技术	259
14.1 RTP/RTCP 协议	230	15.2.3 智能流技术(SureStream)	260
14.1.1 RTP 实时传输协议	230	15.3 流媒体技术的开发与应用	261
14.1.2 RTCP 协议	231	15.3.1 流媒体技术开发工具介绍	261
14.1.3 RTP/RTCP 传输方式的实施	233	15.3.2 流媒体技术的应用	265
14.2 实时流协议 RTSP	234	15.4 习题	265
14.2.1 协议支持的操作	234	第 16 章 数字视频信息传输的应用	266
14.2.2 方法定义	235	16.1 数字视频服务器	266
14.2.3 基于 RTP 的视频传输应用协议族结构模型	235	16.1.1 PC 式视频服务器	266
14.3 RSVP 协议	236	16.1.2 嵌入式视频服务器	267
14.3.1 RSVP 概述	236	16.2 多媒体数字视频监控系統概述	268
14.3.2 RSVP 协议模型	236	16.2.1 多媒体数字视频监控系统的发展	268
14.3.3 RSVP 的预留方式	238	16.2.2 多媒体数字视频监控系统的特点	269
14.3.4 基于 RSVP 的 QOS 参数控制报文设计	239	16.3 数字视频传输系统	269
14.3.5 RSVP 的不足	241	16.3.1 工业环境数字视频监控	269
14.4 前向差错控制(FEC)方法	241	16.3.2 无线远程图像监控	271
14.4.1 概述	241	16.3.3 基于 H.264 的视频传输矩阵	272
14.4.2 前向误差校正法 FEC 的性能参数	241	16.3.4 智能建筑数字视频监控	273
14.4.3 前向误差校正法 FEC 的通用方法	242	16.3.5 变电站/所现场监控系统	274
14.5 宽带 IP 网协议-IPv6	245	16.3.6 智能交通数字视频监控	275
		16.3.7 银行数字视频监控系统	276

16.3.8 超市防盗数字视频监控系统 ...	277	报警系统.....	278
16.3.9 医疗行业数字视频监控系统 ...	278	16.4 习题	279
16.3.10 消防远程图像监控自动		参考文献	280

第1章 绪 言

1.1 数字视频概述

视觉是人类最重要的感觉,也是人类获取信息的主要来源。据统计,人类从外界获取的信息中,75%来自视觉。视频信息与其他的信息形式相比,具有直观、具体、生动等诸多显著优点,并且视频所包含的信息量很大。“百闻不如一见”、“一图值千言”等成语都说明了这一特点。因此,各种视频技术的研究和应用一直吸引着国内外广大科技人员的关注。随着多媒体技术及 Internet 网络的迅速发展,各种信息采集和生产手段在不断投入使用,例如,数码照相机、摄像机、扫描仪、打印机等的使用,使视频信息的来源不断扩大。同时,人们对视频的应用也越来越广泛,视频点播、可视电话、视频会议等应用形式不断涌现。视频信息正在对人们的生活方式和社会发展起着越来越重要的作用。

视频是一组在时间轴上有序排列的图像,是二维图像在一维时间轴上构成的图像序列,又称为动态图像、活动图像或者运动图像。它不仅包含静止图像所包含的内容,还包含场景中目标运动的信息和客观世界随时间变化的信息。电影、电视等都属于视频的范畴。早期的视频主要指模拟的视频信号,随着各种电子技术的发展以及全球数字化进程的推进,视频的采集设备和采集方式有了很大的进展,直接采集数字视频的设备得到了广泛的开发和应用。

针对模拟视频,数字视频有很多的优越性:

(1) 存储便利。模拟视频信号大多存储在磁带或胶片上,磁带对存储环境的温度、湿度要求较严格,因此,长期存放和使用磨损都会对磁带造成物理损伤,如掉磁粉、磁带变形等,这将严重地影响重放的图像质量。而数字存储就没有这些问题。特别是随着数字存储技术的发展,以及存储设备价格的降低,存储的成本也在不断下降,同时,各种压缩编码技术层出不穷,也大大降低了数字视频所需的存储空间。

(2) 图像质量好。模拟视频信号抗干扰能力差,在传输和存储过程中会受到各种干扰和引入各种噪声,并且经多次传输或复制后,会不断积累噪声,导致图像质量下降。而数字视频的复制不会引入任何错误,图像质量不会下降。另外,数字视频中还可采用纠错编码技术,提高抗干扰的能力。

(3) 便于提供新业务。数字视频的数字表示形式和传输手段可以使它很容易地与现有的数字通信和计算机技术相结合,能为人们提供全新的、更灵活、更方便的服务,如视频点播等。

(4) 便于编辑。对模拟视频的编辑非常不方便,需要对原有电影复制或重新复制才可以进行视频剪辑,通常会破坏媒体的载体。而数字视频的编辑不必破坏原有载体,可进行非线性剪辑和视频创作,并且可以多人共享同一视频源。

特别是近年来宽带网络技术的发展,为数字视频技术提供了广阔的机遇。网络的带宽正在逐渐地由十兆,过渡到百兆、千兆,甚至更宽,网络作为视频应用的瓶颈正在被打破。而且各种压缩算法、多媒体通信协议也在不断的完善和发展,这些都为发展更多、更广泛的视频应用

奠定了良好的基础,相信不久的将来,视频应用将遍及我们工作、生活的方方面面。

1.2 本书章节安排

本书主要介绍数字视频信息处理和传输的相关技术和主要研究成果,各个章节的安排和主要内容如下:

第1章对全书进行了概括性介绍。主要介绍了数字视频的概念,以及数字视频相对于模拟视频的优越性,并对全文的章节安排进行了简单介绍。

第2章描述数字视频的基本知识。首先介绍构成时变图像的数学模型,包括三维运动模型、三维场景的几何投影和成像的光学模型。然后概述了模拟视频和数字视频的基本组成原理和概念。最后介绍了几种常用的数字视频格式。

第3章介绍数字视频采集技术。主要介绍视频采集的基本理论和概念。介绍了图像的视觉基础后,详细描述了表示模拟和数字视频的二维和三维采样结构、二维矩阵栅格上的采样、三维点阵上对时空信号数字的采样以及采样结构变换的有关问题,并对采样后信号的量化进行了讨论,最后介绍了CCD图像传感器原理和视频采集卡的有关知识。

第4章介绍数字视频图像压缩方法。先概括叙述了图像数据压缩的必要性和可能性,以及视频压缩的评价指标和图像质量的评判标准,然后对一些主要的压缩编码方法进行了详细介绍,包括变换编码、预测编码与帧间压缩编码、小波变换编码、模型基编码、分形编码以及基于对象的视频编码。

第5章介绍图像压缩的标准。主要介绍了几个国际上流行的图像压缩标准,它们分别是动态图像压缩标准H.261、低码率视频压缩标准H.263以及MPEG工作组提出的数字视频压缩标准MPEG-1、MPEG-2和MPEG-4。

第6章介绍二维运动估计采用的主要算法。首先给出了二维运动估算中的一些基本概念,然后分别介绍二维运动估计中常用的算法,包括基于光流方程的运动估计、基于像素的运动估计、基于块的运动估计、基于全局的运动估计和基于区域的运动估计。

第7章介绍三维运动估计方法。在对三维运动估计的基本概念做出描述后,详细介绍了基于特征的估计方法、基于点对应法的运动估计和直接法运动估计。

第8章介绍数字视频滤波技术。主要介绍帧间滤波方法,包括基于运动补偿的滤波方法,时空滤波去噪方法以及多帧维纳去卷积。

第9章介绍数字视频水印技术。首先对数字水印技术进行了概括性描述,然后介绍了数字水印技术的嵌入和提取的基本原理,最后介绍数字水印技术的分类和主要应用。

第10章介绍视频信息检索技术。首先对视频信息检索的发展情况进行简单回顾后,介绍了基于内容的检索系统的结构,并根据其处理对象的不同,详细介绍了静止图像检索和运动视频检索。

第11章介绍视频信息检索的标准。分别对多媒体内容描述接口MPEG-7和多媒体框架MPEG-21进行了详细的介绍。

第12章介绍视频信息检索的应用系统。首先对当前出现的视频信息检索系统进行了简单阐述,然后对一些著名的系统,包括QBIC系统、Visualeek系统、Photobook系统、VideoQ系统,分别进行了描述。

第 13 章介绍数字视频传输网络技术。简要介绍了多媒体视频对通信网的要求,又对 ATM、SDH 等宽带视频传输网络技术进行了描述,然后分别介绍了千兆快速以太网、FDDI、HDSL 等宽带接入技术以及宽带无线接入技术 LMDS,最后介绍了视频传输的宽带 IP 技术。

第 14 章介绍支持数字视频传输的通信协议。包括传输协议 RTP/RTCP、实时流协议 RTSP、资源预留协议 RSVP、前向差错控制方法 FEC 和宽带 IP 网络协议 IPv6。

第 15 章介绍流媒体技术。介绍了流媒体的基本概念、流媒体技术原理、流式视频文件格式,分析比较了基于 IP 的几种基本通信方式的特点,并对其在流媒体播放中的应用进行了讨论,最后介绍了当今最流行的几种流媒体技术开发工具及其主要应用。

第 16 章介绍数字视频传输系统及其应用。首先介绍了数字视频服务器,分析了 PC 式视频服务器和嵌入式视频服务器的结构和特点,然后对多媒体数字视频监控系统进行了概述,最后结合视频监控的实际应用介绍了具有视频传输和远程控制特点的多媒体数字视频信息传输系统。

第2章 数字视频描述

人们获取的信息的 75% 来自视觉系统,实际就是图像信息。在图像的基础上加上时间因素,就形成视频,因此视频有时又称时基媒体。从最早出现的模拟视频发展到数字视频是一大飞跃,数字视频具有易储存、易编辑等优良特性,正获得越来越广泛的应用。本章将首先在 2.1 节中对时变图像的构成模型进行简单介绍,然后在 2.2 节和 2.3 节中分别介绍模拟视频和数字视频的基础知识,最后 2.4 节介绍目前常见的几种数字视频格式的结构。

2.1 时变图像构成模型

图像是由三维空间场景投影到二维平面形成的,并且随着时间的变化,其场景中的实体也将发生运动,这种二维图像随时间的变化就构成了时变图像,通常用函数 $s_c(x_1, x_2, t)$ 来表示。数字视频是对时变图像进行时空采样而形成的。

场景中实体的运动和摄像机的运动都将影响到时变图像的构成,所以我们首先介绍实体及摄像机的三维运动模型,然后针对如何将三维场景映射到二维图像平面,分别讨论三维到二维的几何投影和成像的光学模型问题。

2.1.1 三维运动模型

在经典运动学中,三维运动可分为刚体运动和非刚体运动。在刚体运动中,实体的三维点之间的相对距离不随时间而改变,也就是说,运动实体的三维结构模型可表示为一个不变形的表面。而在非刚体运动的情况下,运动实体的三维结构模型可表示为一种可变形的表面。

在笛卡尔坐标系中,一个刚体的三维位移表示为:

$$\mathbf{X}' = \mathbf{R}\mathbf{X} + \mathbf{T} \quad (2-1)$$

其中 $\mathbf{X}$ 和 $\mathbf{X}'$ 分别表示实体点在 t 和 t' 时刻相对于旋转体中心的坐标:

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}, \mathbf{X}' = \begin{bmatrix} X'_1 \\ X'_2 \\ X'_3 \end{bmatrix}$$

$\mathbf{T}$ 是一个三维变换矢量:

$$\mathbf{T} = \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix}$$

$\mathbf{R}$ 是一个 3×3 的旋转矩阵,如果用相对于三个坐标轴 X_1, X_2, X_3 旋转的欧拉角 θ, ψ 和 ϕ 来表示,则:

$$\mathbf{R} = \begin{bmatrix} 1 & -\Delta\phi & \Delta\psi \\ \Delta\phi & 1 & -\Delta\theta \\ -\Delta\psi & \Delta\theta & 1 \end{bmatrix}$$

当 t 和 t' 的瞬时间隔 Δt 趋于零,对三维模型式(2-1)取极限,可以得到三维空间中的点 (X_1, X_2, X_3) 的瞬时速度的表达式:

$$\begin{bmatrix} \dot{X}_1 \\ \dot{X}_2 \\ \dot{X}_3 \end{bmatrix} = \begin{bmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} + \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix}$$

其中, Ω_i, V_i 表示各个方向上的角度和线性矢量, $i=1,2,3$ 。用紧缩的形式表示为:

$$\dot{\mathbf{X}} = \boldsymbol{\Omega} \times \mathbf{X} + \mathbf{V}$$

对非刚体的三维运动建模引入了变形体的模型,可以将模型(2-1)扩充为:

$$\mathbf{X}' = (\mathbf{D} + \mathbf{R})\mathbf{X} + \mathbf{T}$$

式中, $\mathbf{D}$ 是一个任意变形矩阵。

2.1.2 几何投影

时变图像实际上是一个随时间变化的三维场景的二维投影,可以表示为一个从四维空间到三维空间的映象:

$$f: R^4 \rightarrow R^3$$

$$(X_1, X_2, X_3, t) \rightarrow (x_1, x_2, t)$$

其中 (X_1, X_2, X_3) 为三维全局坐标, (x_1, x_2) 为二维图像平面坐标, t 为连续时间变量。通常有两类投影:透视投影和正交投影。

透视投影相当于在实体和图像平面之间放置了一个透镜,所有从实体出发的光线都通过该透镜的中心,透镜的中心又称为投影中心。从实体点 (X_1, X_2, X_3) 和图像点 $(x_1, x_2, 0)$ 向轴 X_3 作垂线,根据形成的相似三角形,推导可得:

$$x_1 = \frac{fX_1}{f - X_3}$$

$$x_2 = \frac{fX_2}{f - X_3}$$

正交投影是假设所有从三维实体出发到图像平面的光线互相平行,当图像平面和全局坐标系统的 $X_1 - X_2$ 平面相平行时,正交投影如图 2-1 所示,则:

$$x_1 = X_1 \quad x_2 = X_2$$

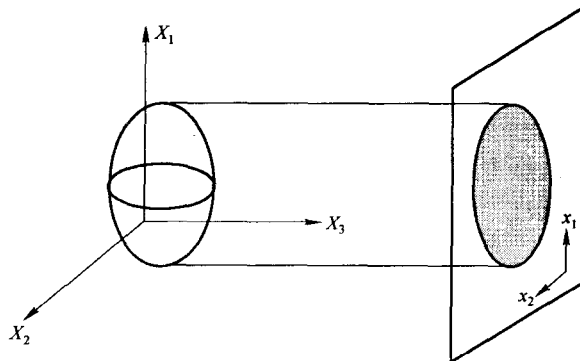


图 2-1 正交投影模型

2.1.3 成像的光学模型

图像强度可以认为与场景中的实体反射的光线数量成比例。场景的反射包含两部分:朗伯反射和镜面反射。如果镜面反射部分可以被忽略,那么这样的表面就称为朗伯表面。

如果用一个均匀强度的点光源照射朗伯表面,则得到的图像强度为:

$$s_c(x_1, x_2, t) = \rho N(t) \cdot L$$

其中, ρ 是表面的反射率; $L = (L_1, L_2, L_3)$ 是平均发光方向上的单位矢量; $N(t)$ 是在空间位置 $(X_1, X_2, X_3(X_1, X_2))$ 和时间 t 的场景的单位表面法线, 可以表示为:

$$N(t) = (-p, -q, 1) / (p^2 + q^2 + 1)^{1/2}$$

其中 $p = \partial X_3 / \partial x_1, q = \partial X_3 / \partial x_2$ 。

假设平均照明方向 L 保持不变, 则可以把由于运动的光学效应而引起的图像强度变化表示如下:

$$\frac{ds_c(x_1, x_2, t)}{dt} = \rho L \cdot \frac{dN(t)}{dt}$$

在点 (X_1, X_2, X_3) 上的法线矢量 N 变化率可近似的表示为:

$$\frac{dN}{dt} \approx \frac{\Delta N}{\Delta t}$$

其中, ΔN 表示在间隔 Δt 内, 由于点 (X_1, X_2, X_3) 到点 (X'_1, X'_2, X'_3) 的三维运动引起的法线矢量的方向变化。它可以表示为:

$$\begin{aligned} \Delta N &= N(X'_1, X'_2, X'_3) - N(X_1, X_2, X_3) \\ &= \frac{(-p', -q', 1)}{(p'^2 + q'^2 + 1)^{1/2}} - \frac{(-p, -q, 1)}{(p^2 + q^2 + 1)^{1/2}} \end{aligned}$$

其中, p' 和 q' 为 $N(X'_1, X'_2, X'_3)$ 的分量, 由下式给定

$$\begin{aligned} p' &= \frac{\partial X'_3}{\partial x'_1} = \frac{\partial X'_3}{\partial x_1} \frac{\partial x_1}{\partial x'_1} = \frac{-\Delta\phi + p}{1 + \Delta\phi p} \\ q' &= \frac{\partial X'_3}{\partial x'_2} = \frac{\Delta\theta + q}{1 - \Delta\theta q} \end{aligned}$$

2.2 模拟视频

模拟视频指的是视频的记录、存储和传输以模拟的形式进行。视频信号最初是用于广播电视的, 也就是说是要经过传输, 尤其是无线传输送到接收机上, 由于视频图像信息量巨大, 因此必须对视频信号作一定的处理, 以节省带宽。为此, 对全电视信号在清晰度、闪烁性、叠加彩色后与黑白图像的兼容性、所占用的带宽等各方面都作了精心的权衡与安排, 研究设计出目前的黑白/彩色全电视信号标准。

2.2.1 电视的扫描

构成一幅图像的基本单元是像素。在电视系统中, 将按空间位置分布的图像按一定顺序分解成与像素对应的时间信号。而电视接收机或电视监视器则按同样的顺序将时间信号在相对应的位置上转变为光信号, 从而将电视图像在显示器屏幕上显示出来。