

第 1 章 绪论	1	第 1 单元 采集表达	
1.1 图像工程的发展	1	第 2 章 摄像机成像和标定	33
1.1.1 基本概念和定义		2.1 视觉过程	33
概括	1	2.2 亮度成像模型	35
1.1.2 图像工程发展情况		2.2.1 光度学和光源	35
回顾	3	2.2.2 从亮度到照度	36
1.2 图像理解及相关学科	5	2.3 空间成像模型	39
1.2.1 图像理解	5	2.3.1 基本摄像机模型	40
1.2.2 计算机视觉	6	2.3.2 近似投影模式	44
1.2.3 其他相关学科	8	2.3.3 一般摄像机模型	46
1.2.4 图像理解的应用领域	9	2.3.4 透镜畸变	48
1.3 图像理解理论框架	9	2.3.5 通用成像模型	49
1.3.1 马尔视觉计算理论	10	2.4 摄像机标定	51
1.3.2 对马尔理论框架的		2.4.1 标定方法分类	51
改进	14	2.4.2 标定程序和参数	53
1.3.3 关于马尔重建理论		2.4.3 两级标定法	55
的讨论	15	2.4.4 精度提升	57
1.3.4 新理论框架的研究	17	2.5 在线摄像机外参数标定	
1.3.5 从心理认知出发的		方法	58
讨论	19	2.5.1 车道线检测与数据	
1.4 深度学习简介	20	筛选	59
1.4.1 图像理解中的深度		2.5.2 优化重投影误差	60
学习	21	2.6 自标定方法	60
1.4.2 卷积神经网络的基本		2.7 结构光主动视觉系统的	
概念	21	标定	62
1.4.3 深度学习核心技术	24	2.7.1 投影模型和标定	63
1.4.4 深度学习的应用	25	2.7.2 图案分离	63
1.5 内容框架和特点	26	2.7.3 计算单应性矩阵	64
总结和复习*	29	2.7.4 计算标定参数	65
随堂测试*	29		

总结和复习*	66	4.3.4 深度和亮度图像同时采集	106
随堂测试*	66	4.4 显微镜 3-D 分层成像	106
第 3 章 压缩感知与成像	67	4.4.1 景深和焦距	106
3.1 压缩感知概述	67	4.4.2 显微镜 3-D 成像	108
3.2 稀疏表达	69	4.4.3 共聚焦显微镜 3-D 成像	110
3.3 测量矩阵及特性	70	4.5 等基线多摄像机组	111
3.3.1 采样/测量模型	70	4.5.1 图像采集	111
3.3.2 测量矩阵特性	71	4.5.2 图像合并方法	112
3.4 解码重构	73	4.6 单摄像机多镜反射折射系统	113
3.4.1 重构原理	73	4.6.1 总体系统结构	114
3.4.2 测量矩阵的校准	74	4.6.2 成像和标定模型	114
3.4.3 典型重构算法	75	总结和复习*	116
3.4.4 基于深度学习的重构算法	81	随堂测试*	116
3.5 稀疏编码与字典学习	82	第 5 章 3-D 景物表达	117
3.5.1 字典学习与矩阵分解	82	5.1 曲线和曲面的局部特征	117
3.5.2 非负矩阵分解	84	5.1.1 曲线局部特征	117
3.5.3 端元提取	85	5.1.2 曲面局部特征	120
3.5.4 稀疏编码	86	5.2 3-D 表面表达	122
3.6 压缩感知的成像应用	88	5.2.1 参数表达	122
3.6.1 单像素相机	88	5.2.2 表面朝向表达	124
3.6.2 压缩感知磁共振成像	89	5.3 等值面的构造和表达	126
总结和复习*	90	5.3.1 行进立方体算法	126
随堂测试*	90	5.3.2 覆盖算法	128
第 4 章 深度信息采集	91	5.3.3 两种算法比较	129
4.1 高维图像和成像方式	91	5.4 从并行轮廓插值 3-D 表面	131
4.1.1 高维图像种类	91	5.5 3-D 实体表达	135
4.1.2 本征图像和非本征图像	92	5.5.1 基本表达方案	135
4.1.3 深度成像方式	93	5.5.2 广义圆柱体表达	136
4.2 双目成像模式	94	总结和复习*	138
4.2.1 双目横向模式	94	随堂测试*	138
4.2.2 双目会聚横向模式	97	第 2 单元 景物重建	
4.2.3 双目轴向模式	99	第 6 章 双目立体视觉	141
4.3 深度图像直接采集	100	6.1 立体视觉模块	141
4.3.1 飞行时间法	100	6.2 基于区域的双目立体匹配	143
4.3.2 结构光法	102	6.2.1 模板匹配	143
4.3.3 莫尔等高条纹法	103	6.2.2 立体匹配	145

6.3 基于特征的双目立体匹配	150	方程	190
6.3.1 基本步骤	150	8.2.5 光度立体法求解	192
6.3.2 尺度不变特征变换	152	8.3 光度立体法进展	194
6.3.3 加速鲁棒性特征	154	8.3.1 光源标定	194
6.3.4 动态规划匹配	158	8.3.2 非朗伯表面反射模型	195
6.4 基于深度学习的立体匹配	159	8.3.3 彩色光度立体法	195
6.4.1 方法分类	159	8.3.4 3-D重建方法	195
6.4.2 立体匹配网络	160	8.4 基于GAN的光度立体法标定	196
6.4.3 基于特征级联CNN的匹配	161	8.4.1 网络结构	196
6.5 视差图误差检测与校正	162	8.4.2 损失函数	197
总结和复习*	164	8.5 从运动求取结构	198
随堂测试*	164	8.5.1 光流和运动场	198
第7章 多目立体视觉	165	8.5.2 光流方程求解	200
7.1 水平多目立体匹配	165	8.5.3 光流与表面取向	204
7.1.1 水平多目图像	166	8.5.4 光流与相对深度	206
7.1.2 倒距离	166	8.6 从分割剪影恢复形状	207
7.2 正交三目立体匹配	169	总结和复习*	208
7.2.1 基本原理	169	随堂测试*	208
7.2.2 基于梯度分类的正交匹配	172	第9章 单目单图像景物恢复	209
7.3 多目立体匹配	175	9.1 单幅图像深度估计	209
7.3.1 任意排列三目立体匹配	175	9.1.1 有监督学习方法	209
7.3.2 正交多目立体匹配	178	9.1.2 无监督学习方法	211
7.4 亚像素级视差	179	9.1.3 半监督学习方法	211
7.4.1 统计分布模型	179	9.2 从影调恢复形状	212
7.4.2 亚像素级视差计算	181	9.2.1 影调与形状	212
总结和复习*	183	9.2.2 亮度方程求解	214
随堂测试*	183	9.3 混合表面透视投影下的SFS	218
第8章 单目多图像景物恢复	184	9.3.1 改进的Ward反射模型	219
8.1 单目景物恢复	184	9.3.2 透视投影下的图像亮度约束方程	219
8.2 光度立体法	185	9.3.3 图像亮度约束方程求解	220
8.2.1 景物亮度和图像亮度	185	9.3.4 基于Blinn-Phong反射模型	221
8.2.2 表面反射特性和亮度	187	9.3.5 新图像亮度约束方程求解	222
8.2.3 景物表面朝向	189	9.4 纹理与表面朝向	224
8.2.4 反射图和亮度约束			

9.4.1 单目成像和畸变	224	11.1.1 匹配策略和类别	262
9.4.2 由纹理变化恢复朝向	225	11.1.2 匹配和配准	263
9.4.3 检测线段纹理消失点	230	11.1.3 匹配评价	264
9.4.4 确定图像外消失点	231	11.2 目标匹配	264
9.5 由焦距确定深度	233	11.2.1 对应点匹配	264
9.6 根据三点透视估计位姿	234	11.2.2 字符串匹配	265
总结和复习*	236	11.2.3 惯量等效椭圆匹配	266
随堂测试*	236	11.2.4 形状矩阵匹配	267
		11.2.5 结构匹配和量度	268
第3单元 场景解释		11.3 动态模式匹配	268
第10章 知识表达和推理	239	11.4 关系匹配	270
10.1 知识概述	239	11.5 图同构匹配	273
10.2 场景知识	240	11.5.1 图论简介	273
10.2.1 模型	240	11.5.2 图同构和匹配	275
10.2.2 属性超图	241	11.6 线条图标记和解释	277
10.2.3 基于知识的建模	243	11.6.1 轮廓标记	277
10.3 过程知识	244	11.6.2 结构推理	278
10.4 知识表达	246	11.6.3 回溯标记	279
10.4.1 知识表达要求	246	11.7 借助匹配实现配准	280
10.4.2 知识表达类型	246	11.7.1 基于特征匹配的异构遥感图像配准	280
10.4.3 基本知识表达方案	247	11.7.2 基于空间关系推理的图像匹配	281
10.4.4 人工智能中的知识表达	249	11.8 多模态图像匹配	282
10.4.5 图像理解系统中的知识模块	249	11.8.1 基于区域的技术	282
10.5 逻辑系统	250	11.8.2 基于特征的技术	283
10.5.1 谓词演算规则	250	总结和复习*	284
10.5.2 利用定理证明推理	253	随堂测试*	284
10.5.3 推理方法分类	255	第12章 场景分析和语义解释	285
10.6 语义网	256	12.1 场景理解概述	285
10.7 产生式系统	258	12.2 模糊推理	287
总结和复习*	260	12.2.1 模糊集合和模糊运算	287
随堂测试*	260	12.2.2 模糊推理方法	288
第11章 广义匹配	261	12.3 遗传算法图像解释	290
11.1 匹配概述	261	12.3.1 遗传算法原理	290

12.3.2	语义分割和解释	292	13.4.1	群体机器人的特性	329
12.4	场景目标标记	294	13.4.2	群体 SLAM 要解决的问题	330
12.5	场景分类	297	13.5	SLAM 的新动向	330
12.5.1	词袋/特征包模型	297	13.5.1	SLAM 与深度学习的结合	331
12.5.2	pLSA 模型	299	13.5.2	SLAM 与多智能体的结合	331
12.5.3	LDA 模型	302	总结和复习 *	332
12.6	遥感图像判读	304	随堂测试 *	332
12.6.1	遥感图像判读方法分类	304	第 14 章 多传感器图像信息融合		333
12.6.2	遥感图像判读知识图谱	305	14.1	信息融合概述	333
12.7	混合增强视觉认知	306	14.2	图像融合	335
12.7.1	从计算机视觉感知到计算机视觉认知	306	14.2.1	图像融合的主要步骤	335
12.7.2	混合增强视觉认知相关技术	307	14.2.2	图像融合的三个层次	336
总结和复习 *	308	14.2.3	图像融合效果评价	338
随堂测试 *	308	14.3	像素级融合方法	340
第 4 单元 研究示例			14.3.1	基本融合方法	340
第 13 章 同时定位和制图		311	14.3.2	融合方法的结合	342
13.1	SLAM 概况	311	14.3.3	小波融合时的最佳分解层数	344
13.1.1	激光 SLAM	311	14.3.4	压缩感知图像融合	345
13.1.2	视觉 SLAM	313	14.3.5	像素级融合示例	346
13.1.3	对比和结合	314	14.4	双能透射和康普顿背散射融合	348
13.2	激光 SLAM 算法	315	14.4.1	成像技术的互补性分析	348
13.2.1	Gmapping 算法	315	14.4.2	互补融合	349
13.2.2	Cartographer 算法	317	14.5	高光谱图像空间光谱特征提取	351
13.2.3	LOAM 算法	319	14.5.1	传统高光谱特征提取方法	351
13.3	视觉 SLAM 算法	320	14.5.2	基于深度学习的空间光谱特征	
13.3.1	ORB-SLAM 系列算法	321			
13.3.2	LSD-SLAM 算法	324			
13.3.3	SVO 算法	328			
13.4	群体机器人和群体 SLAM	329			

	提取方法	352		构化	377
14.6	特征级和决策级融合		15.5.2	体育比赛视频	
	方法	353		排序	380
14.6.1	贝叶斯法	353	15.5.3	家庭录像视频	
14.6.2	证据推理法	354		组织	383
14.6.3	粗糙集理论法	356	15.6	语义分类检索	387
14.7	多源遥感图像融合	359	15.6.1	基于视觉关键词的	
14.7.1	9种多源遥感数			图像分类	387
	据源	359	15.6.2	高层语义与	
14.7.2	多源遥感图像融合			气氛	389
	文献	360	15.7	基于深度学习的跨模态	
14.7.3	遥感图像的空间-光谱			检索	390
	融合	360	15.7.1	跨模态检索技术	
14.7.4	基于深度循环残差			分类	391
	网络的融合	361	15.7.2	图像标题自动	
	总结和复习*	362		生成	392
	随堂测试*	362	15.8	图像检索中的哈希	393
第15章 基于内容的图像和视频			15.8.1	有监督哈希	394
	检索	363	15.8.2	非对称监督深度离散	
15.1	图像和视频检索原理	363		哈希	394
15.2	视觉特征的匹配和检索	366	15.8.3	跨模态图像检索中的	
15.2.1	颜色特征匹配	367		哈希	395
15.2.2	纹理特征计算	368		总结和复习*	396
15.2.3	多尺度形状			随堂测试*	396
	特征	369	第16章 时空行为理解		
15.2.4	综合特征检索	370	16.1	时空技术	397
15.3	基于运动特征的视频		16.2	时空兴趣点	399
	检索	371	16.3	动态轨迹学习和分析	401
15.3.1	全局运动特征	371	16.3.1	自动场景建模	401
15.3.2	局部运动特征	372	16.3.2	学习路径	402
15.4	基于分层匹配追踪的		16.3.3	自动活动分析	404
	检索	373	16.4	动作分类和识别	406
15.4.1	检索框图	373	16.4.1	动作分类	406
15.4.2	单层图像特征		16.4.2	动作识别	407
	提取	374	16.5	结合姿态和上下文的动作	
15.4.3	多层特征提取和			分类	410
	图像检索	375	16.5.1	基于姿态模型的动作	
15.4.4	结合颜色直			分类器	411
	方图	376	16.5.2	基于上下文的动作分	
15.5	视频节目分析和索引	377		类器	411
15.5.1	新闻视频结		16.6	活动和行为建模	413

16.6.1 动作建模	413	附录A 视觉和视知觉	430
16.6.2 活动建模和 识别	416	A.1 视知觉概述	430
16.7 主体与动作联合建模	420	A.2 视觉特性	432
16.7.1 单标签主体-动作 识别	420	A.2.1 视觉的空间特性 ...	432
16.7.2 多标签主体-动作 识别	420	A.2.2 视觉的时间特性 ...	433
16.7.3 主体-动作语义 分割	421	A.2.3 视觉的亮度特性 ...	435
16.8 基于关节点的行为识别 ...	424	A.3 形状知觉	436
16.8.1 使用 CNN 作为 主干	424	A.3.1 图形和背景	436
16.8.2 使用 RNN 作为 主干	424	A.3.2 轮廓和主观轮廓 ...	440
16.8.3 使用 GCN 作为 主干	425	A.3.3 几何图形错觉	441
16.8.4 使用混合网络作为 主干	425	A.4 空间知觉	444
16.9 异常事件检测	426	A.4.1 两种空间知觉观 ...	444
16.9.1 异常事件检测方法 分类	426	A.4.2 非视觉性深度 线索	445
16.9.2 基于卷积自编码器块 学习的检测	428	A.4.3 双目深度线索	445
16.9.3 基于单类神经网络的 检测	429	A.4.4 单目深度线索	446
总结和复习*	429	A.5 运动知觉	448
随堂测试*	429	A.6 生物视觉与立体视觉	452
		A.6.1 生物视觉和双目 视觉	452
		A.6.2 从单目到双目 立体	453
		主题索引	455
		部分思考题和练习题解答*	468
		参考文献*	468