



课程介绍

- 教材和主要参考书
- **第一章**
 - 图论工程、图论分析、分析基础(定量)
 - 课程内容概括，授课进度和安排，先修知识
- 相关情况和信息
 - 练习题/作业
 - 考核方式
 - 清华网络学堂

教材和主要参考书



☀ 图象工程（第4版）



<http://oa.ee.tsinghua.edu.cn/~zhangyujin/C3.htm>



课程介绍

- 教材和主要参考书
- **第一章**
 - 图象工程，图象理解，理论框架
 - 课程内容概括，授课进度和安排，先修知识
- 相关情况和信息
 - 练习题/作业
 - 考核方式
 - 清华网络学堂



教材和主要参考资料



<http://oa.ee.tsinghua.edu.cn/~zhangyujin/>

HOMEPAGE

<http://oa.ee.tsinghua.edu.cn/~zhangyujin/>

Contents

讲义

章毓晋		Yu-Jin ZHANG	
工作信息		Information	
教学课程		Teaching	
编著书籍		Books	
研究文章		Papers	
个人简历		Biography	
上次修改: 2017年6月30日		Last modified: June 30th, 2017	
图像工程研究室成员		Members of Image Engineering Laboratory	

您是2011年10月1日起第  位访问本主页的学者



第1章 绪论

- 1.1 图象工程的发展
- 1.2 图象理解及相关学科
- 1.3 图象理解理论框架
- 1.4 内容框架和特点



1.1 图象工程的发展

图象：

用各种观测系统以不同形式和手段观测客观世界而获得的，可以直接或间接作用于人眼并进而产生视知觉的实体

图象（广义/抽象） \supset 图像（狭义/具体）

图象技术：

图象技术在广义上是各种与图象有关的技术的总称



1.1 图象工程的发展

图象工程

图象工程学科是将数学、光学等基础科学的原理，结合在图象应用中积累的各种技术经验而发展起来的

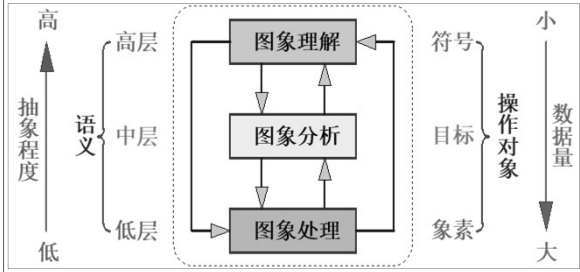
一个对整个图象领域进行研究应用的新学科

一个对各种图象技术进行综合集成的研究和应用的整体框架



1.1 图象工程的发展

图象工程三层次框图





1.1 图象工程的发展

图象工程整体框架

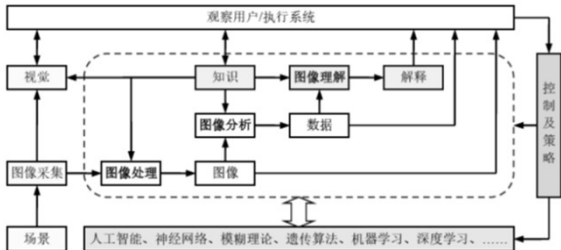


图 1.1.2 图像工程整体框架



1.1 图象工程的发展

图象工程文献综述:

- 文献综述系列

每年《中国图象图形学报》5月那一期
从**1996**年开始, 已历时**23**年, 涉及**15**
种刊物, **14348** (59323) 篇论文

- 主要目的

- (1) 概括我国图象工程发展现状
- (2) 帮助查阅有关文献
- (3) 对期刊编者和论文作者提供参考

大类	名称	小类	名称和主要内容
A	图像处理	A1	图像获取（包括各种成像方式方法、图像采集及存储、摄像机校准等）
		A2	图像重建（从投影等重建图像，间接成像等）
		A3	图像增强和恢复等（包括变换、滤波、复原、修补、校正等）
		A4	图像（视频）压缩编码（包括算法研究、国际标准实现等）
		A5	图像数字水印和图像信息隐藏
		A6	图像多分辨率处理（超分辨率重建、图像分解和插值、分辨率转换等）
B	图像分析	B1	图像分割，边缘和角点（感兴趣点/控制点）检测
		B2	目标表达、描述、测量（包括二值图处理分析等）
		B3	目标特性（颜色、纹理、形状、空间、结构、运动等）的提取分析
		B4	目标检测和识别（目标 2-D 定位、追踪、提取、鉴别和分类等）
		B5	人体生物特征提取和验证（包括人体、人脸和器官的检测、定位与识别）
C	图像理解	C1	图像匹配和融合等（包括序列、立体图的配准、镶嵌等）
		C2	场景恢复（3-D 表达、建模、重构或重建等）
		C3	图像感知和解释（包括语义描述、信息模型、专家系统、机器学习、推理等）
		C4	基于内容的图像和视频检索
		C5	时空技术（3-D 运动分析，姿态检测，对象跟踪，行为判断和理解）
D	技术应用	D1	硬件，系统和快速（/并行）算法
		D2	通信、视频传输（包括电视、网络、广播等）
		D3	文档、文本（包括文字，数字，符号等）
		D4	生物、医学
		D5	遥感、雷达、测绘
		D6	其他（不在以上各类的应用）
E	综述评论	E1	综述（概括图像处理/分析/理解，或综合新技术）



1.2 图象理解及相关学科

1.2.1 图象理解

1.2.2 计算机视觉

1.2.3 其他相关学科

1.2.4 图象理解的应用领域



1.2.1 图象理解

- 图象工程高层
- 在图象分析基础上，结合人工智能和认知理论，进一步研究图象中各目标性质及之间的相互联系，并理解图象内容的含义以及解释原来的客观场景，从而指导和规划行动
- 目前研究成果和应用实例相对比较少
- 限于目前计算机系统的能力和图象理解技术的水平，“系统”完成较低层的工作，而人需要接着系统完成剩下的较高层的工作



1.2.2 计算机视觉

- 人类视觉过程：一个复杂的从感觉（感受到对3-D世界之2-D投影——图象）到知觉（由2-D图象认知3-D世界内容和含义）的过程
- 计算机视觉是指用计算机实现人类的视觉功能
 - 仿生学（模仿）的方法：参照人类视觉系统的结构原理建立相应的操作模块完成类似的功能和工作
 - 工程的（模拟）方法：从分析人类视觉过程的功能着手，并不去刻意模拟人类视觉系统内部结构，而仅考虑系统的输入和输出，并采用任何现有的可行的手段实现系统功能



1.2.2 计算机视觉

- 图象是表达视觉信息的一种物理形式，图象理解必须借助计算机，并基于对视觉信息的处理和分析来进行
- 计算机视觉作为一门学科，与许多以图象作为主要研究对象的学科，特别是图象理解有着非常密切的联系和不同程度的交叉
- 在许多场合和情况下，将图象理解和计算机视觉看作是专业和背景不同的人习惯使用的不同术语更为恰当



1.2.3 其他相关学科

- 模式识别
图象模式识别、分类 (~图象分析)
- 人工智能
指由人类用计算机模拟、执行或再生某些与人类智能有关功能的能力和技术
视觉功能是人类智能的一种体现
- 机器视觉
更关注图象的获取, 系统的构造和算法的实现
- 计算机图形学



1.2.4 图象理解的应用领域

- (1) 工业视觉（生产自动化）
- (2) 人机交互（应用领域和实例）
- (3) 视觉导航（交通、军事）
- (4) 虚拟现实（模拟训练）
- (5) 图象自动解释、判读
- (6) 对人类视觉系统和机理、人脑心理和生理的研究等



1.3 图象理解理论框架

1.3.1 马尔视觉计算理论

1.3.2 对马尔理论框架的改进

1.3.3 关于马尔重建理论的讨论

1.3.4 新理论框架的研究



1.3.1 马尔视觉计算理论

- 马尔1982年出版了《视觉》一书
- Marr D. Vision — A Computational Investigation into the Human Representation and Processing of Visual Information. W.H. Freeman, 1982
- 一个理解视觉信息处理的框架
- 要先理解视觉目的再去理解其中细节



1.3.1 马尔视觉计算理论

1. 视觉是一个复杂的信息加工过程

- 视觉是一个远比人所想象更为复杂的信息加工任务和过程，而且其难度常不为人们所正视
- ✓ 为理解视觉这个复杂的过程，要解决两个问题
 - 视觉信息的表达问题：某些信息是突出的和明确的，另一些信息则是隐藏的和模糊的
 - 视觉信息的加工问题：对信息处理、分析、理解，将不同表达形式转换，逐步抽象



1.3.1 马尔视觉计算理论

2. 视觉信息加工的三个要素

① 计算理论；② 算法实现；③ 硬件实现

① 计算理论

可计算性问题：一个任务要用计算机完成，它应该是可以被计算的

一般对于某个特定的问题，如果存在一个程序，对于给定的输入，这个程序都能在有限步内给出输出，这个问题就是可计算的



1.3.1 马尔视觉计算理论

2. 视觉信息加工的三个要素

① 计算理论

- 视觉问题是否可用现代计算机计算？
- 视觉可计算性：对计算机给定输入，能否得到人类视觉可获得的类似结果？
- 两方面的研究内容：① 计算的是什么以及为什么要计算它们；② 提出一定的约束条件，它们可唯一地确定最终得到的运算结果



1.3.1 马尔视觉计算理论

2. 视觉信息加工的三个要素

② 算法实现

需要给加工所操作的实体选择一种合适的表达

选择加工的输入和输出表达

确定完成表达转换的算法

① 一般情况下可以有許多可选的表达

② 算法的确定常取决于所选的表达

③ 给定一种表达，可有多种完成任务的算法



1.3.1 马尔视觉计算理论

2. 视觉信息加工的三个要素

③ 硬件实现

有了表达和算法，在物理上如何实现算法也是必不可少的

实时性要求 \Rightarrow 专用硬件实现

- (1) 算法的确定常依赖于物理上实现算法硬件的特点
- (2) 同一个算法也可由不同的技术途径实现



1.3.1 马尔视觉计算理论

2. 视觉信息加工的三个要素

视觉信息加工三要素的含义

要素	名称	含义和所解决的问题
1	计算理论	什么是计算目的，为什么要这样计算
2	表达和算法	怎样实现计算理论，什么是输入输出表达，用什么算法实现表达间的转换
3	硬件实现	怎样在物理上实现表达和算法，什么是计算结构的具体细节



1.3.1 马尔视觉计算理论

2. 视觉信息加工的三个要素

- 它们之间有一定的逻辑因果联系，但并无绝对的依赖关系
- 实际上看成两个层次更恰当
- 一旦有了计算理论，表达和算法与硬件实现是互相影响的





1.3.1 马尔视觉计算理论

3. 视觉信息的三级内部表达

(1) 基素表达 (primal sketch)

- 一种2-D表达，它是图象特征的集合，描述了物体上属性发生变化的轮廓部分
- 只用基素表达不能保证得到对场景的唯一解释



(a)



(b)



(c)



(d)



(e)

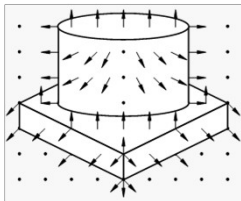


1.3.1 马尔视觉计算理论

3. 视觉信息的三级内部表达

(2) 2.5-D表达 (2.5-D sketch)

- ① 将物体可见表面正交投影分解成单元表面
- ② 用法线（箭头/针）代表单元表面的取向
- ③ 将各法线画出，叠加于物体轮廓内可见面上





1.3.1 马尔视觉计算理论

3. 视觉信息的三级内部表达

(2) 2.5-D表达 (2.5-D sketch)

- 本征图，表示了物体表面面元的朝向
- 既表达了一部分物体轮廓的信息（这与基素表达类似），又表达了以观察者为中心、可观察到的物体表面的取向信息
- 与人所理解的3-D物体一致（可见物体轮廓以内目标的3-D信息，如边界、深度，反射特性等）



1.3.1 马尔视觉计算理论

3. 视觉信息的三级内部表达

(2) 2.5-D表达 (2.5-D sketch)

- 可将2.5-D图转化成 (相对) 深度图
- 给定 $z(x, y)$ 对 x 和 y 的偏导 p 和 q , 理论上讲可通过在平面上沿任意曲线的积分来恢复 $z(x, y)$

$$z(x, y) = z(x_0, y_0) + \int_{(x_0, y_0)}^{(x, y)} (pdx + qdy)$$

- 为最小化误差可选择 $z(x, y)$ 满足

$$\iint_I [(z_x - p)^2 + (z_y - q)^2] dx dy$$

{P.15}



1.3.1 马尔视觉计算理论

3. 视觉信息的三级内部表达

(3) 3-D表达 (3-D representation)

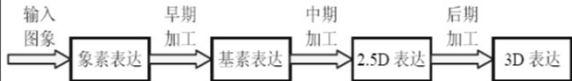
- 以物体为中心 (即也包括了物体不可见部分) 的表达形式
- 在以物体为中心的坐标系中描述3-D物体的形状及其空间组织
- ① 空间占有数组, ② 单元分解, ③ 几何模型
- 广义圆柱体表达方法 {第5章}



1.3.1 马尔视觉计算理论

3. 视觉信息的三级内部表达

- 从计算机或信息加工的角度来说，视觉可计算性问题可分成几个步骤，步骤之间是某种表达形式，而每个步骤都是把前后两种表达形式联系起来的计算/加工方法





1.3.1 马尔视觉计算理论

3. 视觉信息的三级内部表达

表 1.3.2 视觉可计算性问题的表达框架

名称	目的	基元
图像	表达场景的亮度或物体的照度	像素（值）
基素图	表达图像中亮度变化位置、物体轮廓的几何分布和组织结构	零交叉点、端点、角点、拐点、边缘段、边界等
2.5-D 图	在以观察者为中心的坐标系中表达物体可见表面的取向、深度、轮廓等性质	局部表面朝向（“针”基元）、表面朝向不连续点、深度、深度上不连续点
3-D 图	在以物体为中心的坐标系中，用体元或面元集合描述形状和形状的空间组织形式	3-D 模型，以轴线为骨架，将体元或面元附在轴线上



1.3.1 马尔视觉计算理论

4. 视觉信息处理按照功能模块的形式组织

- 把视觉信息的系统看成由一组功能相对独立的模块所组成的思想，不仅有计算方面进化论和认识论的论据支持，而且某些功能模块已经能用实验的方法分离出来
- 人通过使用多种线索或从它们的结合来获得各种本征视觉信息。这启示计算机视觉系统应该包含许多模块，每个模块获取某一特定的视觉线索，结合不同的模块来最终完成视觉任务



1.3.1 马尔视觉计算理论

5. 计算理论形式化表示必须考虑约束条件

- 图象采集过程中，原始场景中信息会发生变化
 - (1) 当3-D的场景被投影为2-D图象时，丢失了物体深度和不可见部分的信息
 - (2) 图象是从特定视角获取的，同一物体的不同视角图象会不同，另外由于物体遮挡也会丢失信息
 - (3) 成象投影使得照明、物体几何形状和表面反射特性、摄象机特性、光源与物体和摄象机之间的空间关系等都被综合成单一的图象灰度值，很难区分
 - (4) 在成象过程中不可避免地会引入噪声和畸变



1.3.1 马尔视觉计算理论

5. 计算理论形式化表示必须考虑约束条件

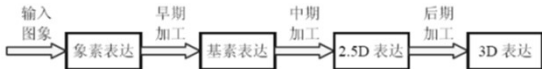
- 适定的问题：如果它的解是：① 存在的；
② 唯一的；③ 连续地依赖于初始数据
- 由于前述原始场景中信息发生变化的原因，使得将视觉处理问题作为光学成像过程逆问题的求解方法不适定，成了病态问题
- 利用约束条件可改变病态问题
通过给计算问题加上约束条件可使它含义明确从而能够获解



1.3.2 对马尔理论框架的改进

- 四个有关整体框架的问题:

- (1) 框架中输入是被动的
- (2) 框架中加工目的不变，总是恢复场景中物体的位置和形状等（3-D表达）
- (3) 框架未足够重视高层知识的指导作用
- (4) 整个框架中信息加工过程基本自下而上，单向流动，没有反馈

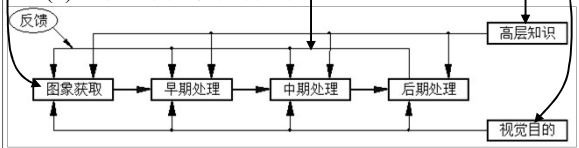




1.3.2 对马尔理论框架的改进

- 四个方面的相应改进:

- (1) 增加了图象（主动）获取模块
- (2) 根据视觉目的进行决策
- (3) 利用高层知识可解决低层信息不足的问题
- (4) 增加了反馈控制流向





1.4 内容框架和特点

课程内容结构

- 15章正文和1个附录
 - 主要内容分四个单元
 - 总结与复习
 - ✓ 参考文献（介绍）
 - ✓ 思考题和练习题
- (1) 采集表达

(2) 景物重建

(3) 场景解释

(4) 研究示例



1.4 内容框架和特点

图象技术

数学、计算机的应用领域

1-D信号的推广，特殊的模式识别技术

学习图象技术所需的基础知识

数学：线性代数，随机过程

计算机科学：软件编程

电子学：信号处理，电路原理



联系信息



- ✎ 通信地址：北京清华大学电子工程系
- ✎ 邮政编码：100084
- ✎ 办公地址：清华大学，罗姆楼，6层305室
- ✎ 办公电话：(010) 62798540
- ✎ 传真号码：(010) 62770317
- ✎ 电子邮件：zhang-yj@tsinghua.edu.cn
- ✎ 个人主页：oa.ee.tsinghua.edu.cn/~zhangyujin/