

图 象 理 解

（第4版）

章毓晋

清华大学电子工程系 100084 北京

第3单元 场景解释

- 第10章 知识表达和推理
- 第11章 广义匹配
- 第12章 场景分析和语义解释

通过学习、推理、与模型的匹配等来解释场景的内容、特性、变化、态势或趋向

知识指导对客观世界认识和理解

将从图象中获得的信息与已有的解释场景的模型进行匹配

对场景的高层次解释和语义描述

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第2页

第10章 知识表达和推理

- 10.1 知识基础
- 10.2 场景知识
- 10.3 过程知识
- 10.4 知识表达
- 10.5 逻辑系统
- 10.6 语义网
- 10.7 产生式系统



第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第3页

10.1 知识基础

1、知识的作用：

- 图象理解是一个非常复杂的过程，困难主要来源于两点：一是有大量的数据要处理，另一是缺乏从已知的像素矩阵到所需结果（对图象内容的细节把握）的基本工具
- 为克服困难，一方面需要对问题的一般性加以限制。这比较直接，如限制未知条件或限制对期望结果的范围或精度
- 另一方面需要将人类的知识引入理解过程

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第4页

10.1 知识基础

1、知识的作用：

- 人类在理解世界时用到许多经验和知识，因此在图象理解与计算机视觉的研究中人们也试图有效地借助对知识的利用以提高分析理解的可靠性和效率
- 知识在整个视觉过程的各个阶段都可以起到重要的作用（例如在图象分割中……）
- 把握场景的含义常需要进行推理，知识也是推理的基础

第10讲

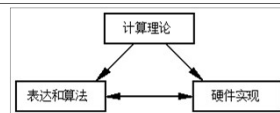
章毓晋 (TH-EE-IE) ZHANG YU JIN

第5页

10.1 知识基础

2、知识的分类：


- 可将（图象理解）知识分成三类：
 - ① 程序知识：与诸如选择算法、设置算法参数等操作有关
 - ② 视觉知识：与图象形成模型有关（较低）
 - ③ 世界知识：关于问题领域的整体知识
- 也可将（图象理解）知识分成两类
场景知识和过程知识 ⇒ 10.2节，10.3节




第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第6页



10.1 知识基础



3、知识的表达:

- 用什么样的形式来表示知识

用原始图象表达目标不可靠，因为对目标的成像结果与对它们的照明和观察方向等都有关


低层工作的结果常是某种表达，如双眼视觉系统接收同一个场景的两幅不同的图象并产生一个明确表示深度信息的表达

高层次工作中对表达的选择很不明确，至今为止的研究都表明并不能从一个目标的表现来恢复其几何性质并用作其表达


第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第7页



10.2 场景知识



10.2.1 模型


10.2.2 属性超图

10.2.3 基于知识的建模


第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第8页



10.2.1 模型




- 有关场景的知识主要包括有关客观世界中目标的事实特性，一般局限于某些确定的场景中（场景的先验知识）
- 这类知识常用模型表示，也直接称为模型
- 模型这个词反映了任何自然现象都只能在一定程度（精确度或准确度）上进行描述这样一个事实（对问题的一般性加了限制）
- 过限定逆问题**：描述参数少，验证数据多
- 欠限定问题**：数据太少，需要增加限定条件


第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第9页



10.2.1 模型



- 2-D模型**表示图象的特性

优点：可直接用于图象或图象特性的匹配

缺点：不易全面表达3-D空间中物体的几何特点及物体之间的联系
- 3-D模型**包含了有关场景中3-D目标位置和形状的特性以及它们之间的联系


优点：灵活、通用，可适合多种不同的场合

缺点：建立模型和场景描述间的匹配联系比较困难，产生模型所需计算量也很大


第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第10页



10.2.1 模型



- 常用的2-D模型可分为两类
- 图象模型**：全局的，整体的


把整个图象的描述去和场景的模型匹配
- 目标模型**：局部的，部分的

仅将局部图象的描述去和模型匹配
- 场景知识对图象理解很重要。它在很多情况下可帮助给出对场景的唯一解释，因为根据模型常可限定问题的条件和变化的种类


第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第11页



10.2.2 属性超图



- 一种对3-D物体属性（场景知识）的表达方法
- 物体通过属性对（集合）的形式来表达
- 一个属性对是一个有序对，可记为 (A_i, a_i) ，其中 A_i 是属性名， a_i 是属性值
- 一个属性集可表达为 $\{(A_1, a_1), (A_2, a_2), \dots, (A_n, a_n)\}$
- 一个属性图表示为 $G = [(V) (A)]$ ，其中 V 为结点集， A 为弧集合。每个元素 \leftrightarrow 属性集

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第12页

10.2.3 基于知识的建模

- 从知识出发可以对三维场景和目标建立模型
- 建模可根据对目标所获图象数据来逐步进行
- 建模本质上是一个学习的过程，类似认知
- 认知：人在多次看到某一物体后，会将该物体的各种特征抽象出来，从而得到对物体的描述，并存储于大脑中，以备日后使用的
- 学习：学习者有目的，学习者需思考
- 训练：无目的的学习，老师有目的

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第13页

10.2.3 基于知识的建模

- (通用) 建模是一个复杂的问题
 - (1) 模型应包含目标的全部信息，很困难所以有时不(需/能)建立完整模型
 - (2) 模型的复杂度也较难确定
- 在建模中，与应用领域有关的模型知识或场景知识的使用是非常重要的

例如，计算机辅助设计

5.5.1节的几何模型

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第14页

10.2.3 基于知识的建模

- 根据几何模型法建立目标模型时，应考虑
 - (1) 客观世界由物体组成，每个物体又可分解为处在不同层次的几何元素，模型的数据结构应能将这些反映出来
 - (2) 对任何几何元素的表达都要使用一定的坐标系，为了表达方便，各层次所采用的坐标系可以不同，这样一来，模型各层次之间应有坐标转换所需的信息

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第15页

10.2.3 基于知识的建模

- 根据几何模型法建立目标模型时，应考虑
 - (3) 在同一个层次最好使用相同的数据结构
 - (4) 模型对特征的表达可分为显式的和隐式的

例如，对曲面的显式表达直接给出曲面上各点应满足的条件方程，由于根据曲面的方程可以计算出它们的交线，所以可认为对曲面的显式表达也是对曲面交线的隐式表达

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第16页

10.3 过程知识

- 运用知识(进行加工)的知识 (cybernetics)
- 控制(论)：借助运用知识的知识进行决策
- 根据控制顺序和控制程度的不同来分类

控制

分层控制

- (1) 由底向上控制
- (2) 由顶向下控制

无分层控制

- (3) 反馈控制
 - 有序反馈
 - 无序(动态)反馈
- (4) 异层控制
 - 反馈到底层
 - 不反馈到底层

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第17页

10.3 过程知识

(1) 由底向上控制

整个过程基本上完全依赖于输入图象数据(所以也叫数据驱动)，而有关目标的知识(模型)只在匹配场景的描述时使用

知识

解释

场景描述

图象特征

图象

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

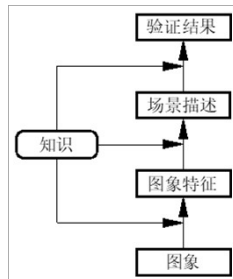
第18页

3

10.3 过程知识

(2) 由顶向下控制

整个过程都由模型（基于知识）控制（所以也叫模型驱动控制），多用于检验结果，即对目标预测并验证假设



第10讲

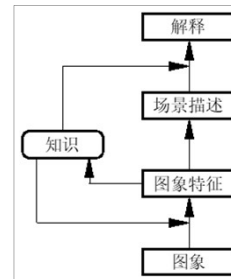
章毓晋 (TH-EE-IE) ZHANG YU JIN

第19页

10.3 过程知识

(3) 反馈控制

将部分加工分析结果从较高层反馈到较低层的处理阶段以增进低层的性能，这里先验知识是必要的



第10讲

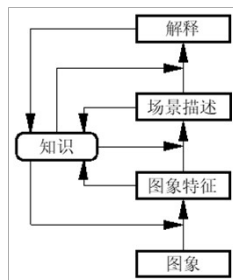
章毓晋 (TH-EE-IE) ZHANG YU JIN

第20页

10.3 过程知识

(4) 异层控制

基本思想是在任何时刻都选择最有效的处理手段来进行处理，每个阶段的部分处理结果都将反映在其后的处理过程中



第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第21页

10.4 知识表达

10.4.1 知识表达要求

10.4.2 知识表达类型

10.4.3 图象理解系统中的知识模块

10.4.4 基本知识表达方案

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第22页

10.4.1 知识表达要求

- (1) 充分表达：通用，即有能力表达有关领域中各种各样所需的知识
- (2) 充分推理：有能力管理知识表达结构，并且借助从旧知识中推理出来的新知识可以导出新的表达结构
- (3) 有效推理：利用所有可获得的信息进行有用的推理，有能力把附加的信息结合到表达结构中去，而这些附加的信息可用于把推理机能的重点放到最有希望的方向上去
- (4) 有效地获取知识：可帮助有能力很方便地获取新的知识

第10讲

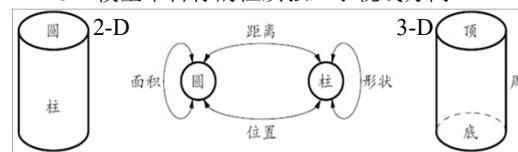
章毓晋 (TH-EE-IE) ZHANG YU JIN

第23页

10.4.2 知识表达类型

2-D和3-D线图模型

- 基本元素是区域，结点反映了区域的特性，连线反映了区域间的关系
- 2-D模型中目标的性质与观察方向有关
- 3-D模型中目标的性质独立于视线方向



第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第24页



10.4.3 图象理解系统中的知识模块



{P.240}

1. 软件系统解决问题所需的知识模块
2. 一般图象处理知识模块
3. 用户界面知识模块
4. 自动学习模块

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第25页



10.4.4 基本知识表达方案



{P.241}

1. 守护程序 (demons)
2. 黑板系统 (blackboards)
3. 图象函数库
4. 图象处理语言
5. 图象理解代理和基于代理的表达

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第26页



10.5 逻辑系统



谓词逻辑 (predicate logic): 一种知识类型
可表达命题和借助事实知识库推出新事实

10.5.1 谓词演算规则

10.5.2 利用定理证明来推理

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第27页



10.5.1 谓词演算规则



- 逻辑系统基于一阶谓词演算 (符号形式语言)
- 4种基本元素: ① 谓词符号, ② 函数符号, ③ 变量符号, ④ 常量符号

表 10.5.1 谓词示例

语句	谓词
图像 I 是数字图像	$DIGITAL(I)$
图像 J 是扫描图像	$SCAN(J)$
组合数字图像 I 和扫描图像 J	$COMBINE [DIGITAL(I), SCAN(J)]$
图像 I 中有个像素 p	$INSIDE(p, I)$
目标 x 位于目标 y 后面	$BEHIND(x, y)$

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第28页



10.5.1 谓词演算规则



- 谓词也称为原子, 将原子用逻辑连词结合起来得到子句
- 常用逻辑连词有: “ \wedge ”(AND), “ \vee ”(OR), “ \sim ”(NOT), “ \Rightarrow ”(IMPLIES)

表 10.5.2 子句示例

语句	子句
图像 I 是数字图像和扫描图像	$DIGITAL(I) \wedge SCAN(I)$
图像 I 是数字图像或模拟图像	$DIGITAL(I) \vee ANALOGUE(I)$
图像 I 不是数字图像	$\sim DIGITAL(I)$
如果图像 I 是扫描图像, 则图像 I 是数字图像	$SCAN(I) \Rightarrow DIGITAL(I)$
一幅图像或者是数字图像或者是模拟图像	$(\forall x) [DIGITAL(x) \vee ANALOGUE(x)]$
有个目标在图像中	$(\exists x) INSIDE(x, I)$

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第29页



10.5.1 谓词演算规则



子句形式 \Leftrightarrow 非子句形式

- 子句形式句法 (条件) \Rightarrow (结论)
逻辑表达式形式: $(\forall x_1 x_2 \dots x_k) [A_1 \wedge A_2 \wedge \dots \wedge A_n \Rightarrow B_1 \vee B_2 \vee \dots \vee B_m]$

$(\forall x)[IMAGE(x) \wedge DIGITAL(x) \Rightarrow GRAY(x) \vee COLOR(x)]$

- 非子句形式句法
逻辑表达式包括: 原子、逻辑连词、存在量词 (\exists) 和全称量词 (\forall)

$(\forall x)[IMAGE(x) \wedge DIGITAL(x) \vee GRAY(x) \vee COLOR(x)]$

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第30页

10.5.2 利用定理证明来推理

- 推理规则：**（WFF = 合适/合理公式）
将它应用于WFF的集合可以产生新的WFF
被推导出来的 WFF 称为定理
顺序地应用推理规则就构成了对定理的证明

表 10.5.5 推理规则示例

推理规则	定义
取式 (Modus Ponens)	从 $W_1 \wedge (W_1 \Rightarrow W_2)$ 推论出 W_2
拒式 (Modus Tollens)	从 $\sim W_2 \wedge (\sim W_1 \Rightarrow W_2)$ 推论出 W_1
投影	从 $W_1 \wedge W_2$ 推论出 W_1
全称规定化	从 $(\forall x) W(x)$ 推论出 $W(c)$

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第31页

10.5.2 利用定理证明来推理

证明逻辑表达式的正确性

- 在谓词演算中，可使用两种基本方法：
 - 去匹配表达式中子句形式的项
思路是先证明子句的非与事实不符，这样就可表明所需证明的子句是成立的
 - 用一个与证明数学表达式类似的过程直接对非子句形式进行操作
利用规则和定理进行推演直到结论

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第32页

10.5.2 利用定理证明来推理

证明逻辑表达式的正确性

用子句表达来证明：

- $\Rightarrow \text{INSIDE}(p, I)$
 - $\Rightarrow \text{DIGITAL}(I)$
 - $(\forall x, y)[\text{DIGITAL}(y) \Rightarrow \text{DISCRETE}(x)]$
 - $\text{DISCRETE}(p) \Rightarrow \{\text{结论的非}\}$
- 预解：(5) $\Rightarrow \text{DISCRETE}(p) \quad \{(3)\text{左} \Leftrightarrow (2)\text{右}\}$
- (4)和(5)的解是空子句，这个结果是矛盾的
(4)不能成立，证明了DISCRETE(p)的正确性

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第33页

10.5.2 利用定理证明来推理

证明逻辑表达式的正确性

- 用非子句表达来证明：③ $\{\sim A \Rightarrow B \Leftrightarrow A \vee B\}$
- 替换 (1) $(\forall x, y)[\text{INSIDE}(x, y) \wedge \text{DIGITAL}(y)] \wedge [\sim \text{INSIDE}(x, y) \wedge \sim \text{DIGITAL}(y) \vee \text{DISCRETE}(x)]$
- 投影 (2) $[\text{INSIDE}(p, I) \wedge \text{DIGITAL}(I)] \wedge [\sim \text{INSIDE}(p, I) \wedge \sim \text{DIGITAL}(I) \vee \text{DISCRETE}(p)]$
- 分配 (3) $\text{INSIDE}(p, I) \wedge [\sim \text{INSIDE}(p, I) \vee \text{DISCRETE}(p)]$
- 简化 (4) $\text{INSIDE}(p, I) \wedge \text{DISCRETE}(p)$
- 投影 (5) $\text{DISCRETE}(p)$

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第34页

10.5.2 利用定理证明来推理

- 谓词演算的一个基本结论就是所有的定理都可在有限时间内得到证明
- 析解推理规则的基本步骤：
 - 先将问题的基本元素表达成子句形式
 - 然后求可以匹配的隐含表达式的前提和结果
 - 再通过替换变量以使原子相等来进行匹配
 - 匹配后所得到的（称为预解的）子句包括不匹配的左右两边

{例10.5.5}定理证明现在转化成要解出子句以产生空子句

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第35页

10.6 语义网

- 语义网络是编了号的有向图
给出图象中各元素间联系的直观表示

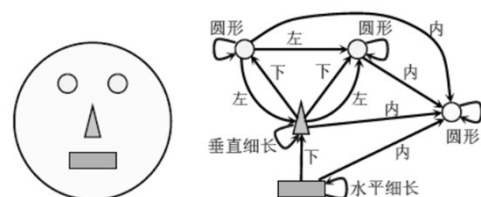


图 10.6.1 人脸模型和它的语义网示例

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第36页

10.6 语义网

表达和处理以子句形式描述知识的语义网络

- ① 数字图象 I 中有一个象素 p
- ② 图象 J 是扫描图象
- ③ 如果象素在数字图象中, 那么它是离散的

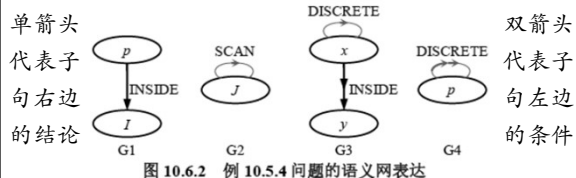


图 10.6.2 例 10.5.4 问题的语义网表达

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第37页

10.6 语义网

通过合并子图进行推理

{例 10.6.1}

- (1) $\Rightarrow \text{INSIDE}(p, I)$
- (2) $\Rightarrow \text{SCAN}(J)$
- (3) $(\forall x, y)[\text{INSIDE}(x, y) \Rightarrow \text{DISCRETE}(x)]$
- (4) $\text{DISCRETE}(p) \Rightarrow$

合并: 用双箭头的弧消去单箭头的弧

将 G13 与 G4 合并得到空图

图 10.6.3 通过合并子图进行推理

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第38页

10.7 产生式系统

- 基于规则的系统
- 模块化的知识表达系统
- 产生式系统把知识表达成称为产生式规则的守护程序 (关键情况发生就启动) 的集合, 每个产生式规则都具有条件-动作 (condition-action) 对的形式
- 产生式系统的典型应用是专家系统, 专家系统是具有解决特定问题能力 (专家知识) 的人-机系统

第10讲

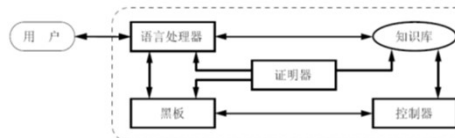
章毓晋 (TH-EE-IE) ZHANG YU JIN

第39页

10.7 产生式系统

专家系统的基本构成单元 (示例)

- 语言处理器用作用户和系统间的通信接口
- 证明器将系统的动作解释给用户
- 黑板记录当前数据、工作假设、中间决策等
- 知识库包括过程、视觉和世界知识
- 控制器提供解决问题的知识以及冲突-解决策略



第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第40页

10.7 产生式系统

- 专家系统中的规则一般具有如下典型形式:
if (conditions) then (actions)
- 上述规则可读为: 如果条件1和条件2和……和条件 m 为真, 那么就执行动作1和动作2和……和动作 n
- 专家系统规则和子句表达式之间有相似性
- 专家系统规则一般较易确定和读取
- 例10.5.4中的物理定理可以表达成规则形式
如果 象素 p 在图象 I 中, 图象 I 是数字图象
那么 象素 p 是离散的

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第41页

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/

第10讲

章毓晋 (TH-EE-IE) ZHANG YU JIN

第42页