

计算机视觉基础

章毓晋

清华计算机图书·译丛

Foundations of Computer Vision

计算机视觉基础

本书介绍计算机视觉的基础内容，比较侧重计算几何和目标检测方面。本书对图像网格的构建和叠加、德劳内三角剖分和沃罗诺伊镶嵌、多边形拼贴、图像拓扑等都有比较全面深入的介绍，并对图像结构给出了直观可视的描述。书中提供了大量相应的Matlab程序，可结合原理学习进行实验，以进一步加深理解并解决实际问题。

本书可作为信号与信息处理、通信与信息系统、电子与通信工程、模式识别与智能系统、计算机科学等学科大学高年级本科生或研究生专业课教材和教学参考书，还可供涉及图像和机器视觉技术的应用行业（如生物医学、电视广播、工业自动化、文档识别、机器人、电子医疗设备、遥感测绘、增强现实、智能交通和军事侦察等）的科技工作者和从业者参考和自学。

课件下载·样书申请



书圈

清华社官方微信信号



扫 我 有 惊 喜



定价：98.00元

清华计算机图书·译丛

计算机视觉基础

清华大学出版社



Springer

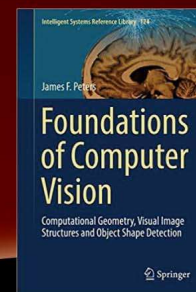
清华计算机图书·译丛

Foundations of Computer Vision

计算机视觉基础

[加] 詹姆斯·彼得斯 (James F. Peters) 著

章毓晋 译



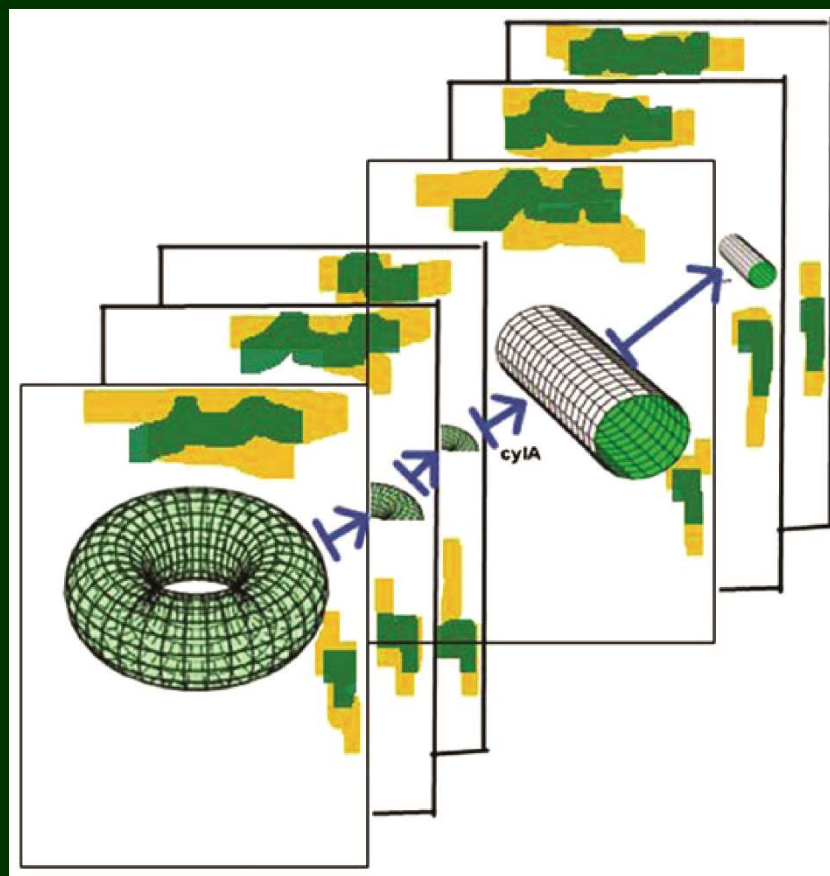
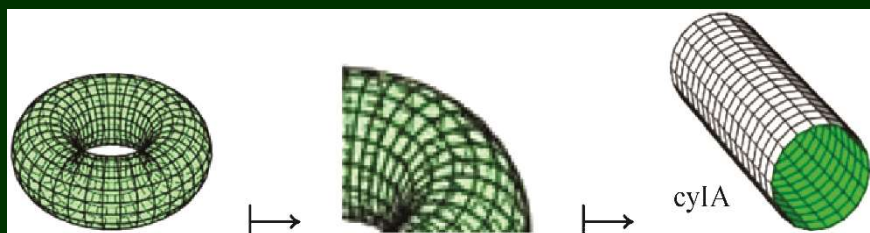
清华大学出版社



第9章 后记：形状适合计算机视觉环境的地方

在视频帧图像的系列中，存在一个变形系列
(由 $I \rightarrow$ 表示)

这种形状变化的现象在检测和比较图像目标形状时很重要

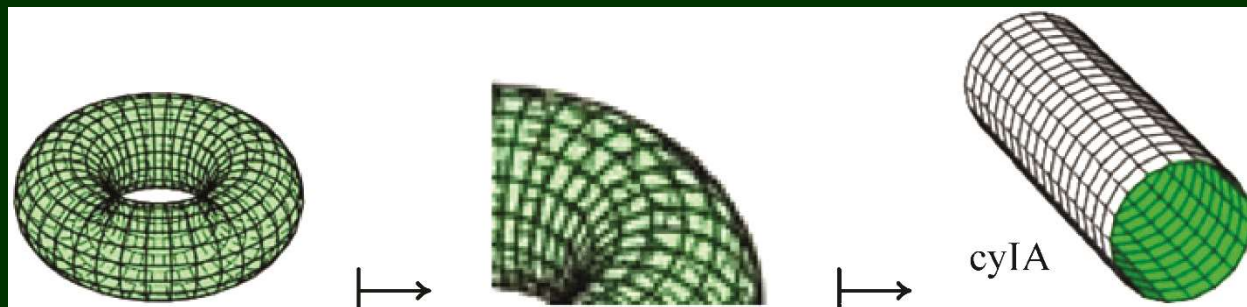




9.1 自然场景中的目标形状

形状是难以捉摸的生物，它们漂浮进出人们有时会感知的自然场景，存储在记忆中并被相机所记录

形状变化的现象在检测和比较图像目标形状时会很重要，目标形状在视频中某一帧里以一种形式出现并且在同一视频中的另一帧里以或多或少改变了的形式再现





9.1 自然场景中的目标形状

使用了三个工具从数字图像中提取对图像目标及其形状的检测和分类有用的信息

几何、拓扑和算法

图像几何以各种形式出现

图像中的原型几何结构是各种图像区域，沃罗诺伊镶嵌中的多边形（沃罗诺伊区域）和德劳内三角剖分中的三角形（德劳内三角形区域）



9.1 自然场景中的目标形状

图像中几何结构体系中的上一层是检测包含核和辐条的最大核聚类（MNC），这些核和辐条定义了网状神经结构

在图像几何结构的体系中进一步向上，可以发现围绕每个MNC核的精细边缘集和粗糙边缘集

将这个图像几何体更进一步（沿精细轮廓边界向外移动），将可以识别围绕每个精细轮廓的路线轮廓

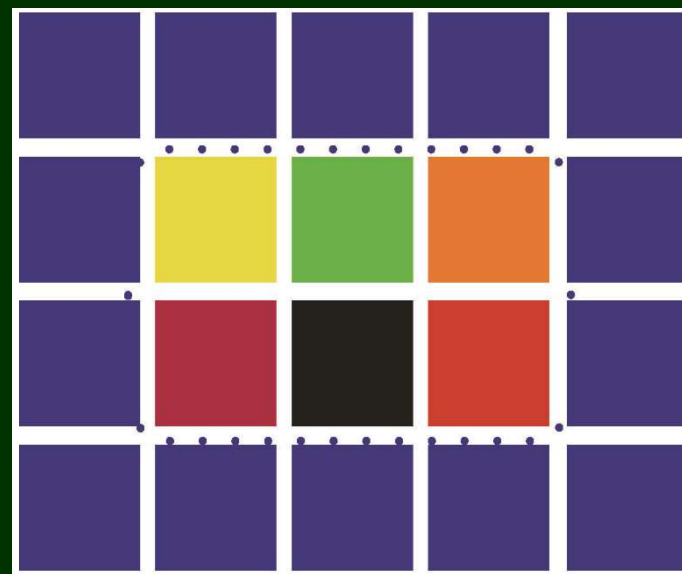


9.1 自然场景中的目标形状

图像拓扑提供了在对图像区域的分析和分类中有用的结构

图像拓扑中的主要结构是开集。开集基本上是一组元素，不包括其边界上的元素

开集 $X = \{\text{yellow}, \text{green}, \text{orange}, \text{red}, \text{white}, \text{black}\}$





9.1 自然场景中的目标形状

图像拓扑是图像开集 X 上的开集 τ 的集合，具有以下属性：

- (1) 空集 \emptyset 是开集且 \emptyset 在 τ 中
- (2) 集合 X 是开集且 X 在 τ 中
- (3) 如果 A 是 τ 中的开集的子集合，那么

$$\bigcup_{B \in A} B \text{ 是 } \tau \text{ 中的开集}$$

- (4) 如果 A 是 τ 中的开集的子集合，那么

$$\bigcap_{B \in A} B \text{ 是 } \tau \text{ 中的开集}$$



9.1 自然场景中的目标形状

在其上具有拓扑 τ 的开集 X 被称为拓扑空间

拓扑空间：二元组 (X, τ)

一个开放的沃罗诺伊区域是一个镶嵌数字图像上的沃罗诺伊区域，其内部包含所有像素，但不包括其边缘

镶嵌数字图像上的数字沃罗诺伊拓扑是开放的沃罗诺伊区域的集合，它们满足拓扑的属性



9.1 自然场景中的目标形状

图像目标形状检测

图像目标形状检测中的技巧是将视频帧序列上的形状变化视为可能在初始视频帧图像中检测到的原始形状的近似

通过在视频中检测特定形状（称为目标形状），然后观察与目标形状大致相同的类似形状的出现，可实现目标形状跟踪



9.2 形状估计

基本思想是双重的。首先，需要一些测量图像目标形状的方法。其次，需要确定何时一个形状与另一个形状大致相同

图像区域的形状周边是连接的边缘集

如果在边缘 e 和 e' 之间存在通路，则边 e 和 e' 是连接的。由连接的直边缘组成的周边称为边缘集。只要边缘集中的每对直边都连接在一起，则边缘集是连接的边缘集



9.2 形状估计

一对无人机交通视频帧



目标无人机视频帧区域










无人机视频帧中的样本区域



9.2 形状估计

要从每个视频帧中获取形状周长，执行以下操作：

- (1) 选取视频帧、
- (2) 选取网格生成点集合 S
- (3) 选取视频帧图像 $\in \{\text{img1}, \text{img2}\}$
- (4) 在上叠加沃罗诺伊图 $V(S)$ ，即镶嵌的，用沃罗诺伊区域 $V(s)$ 覆盖，使用每个生成点（网点，种子点） $s \in S$
- (5) 在图像图 $V(S)$ 中识别MNC（称为MNC(s)）
- (6) 在中识别粗糙边缘轮廓MNCedgelet（视频帧中的目标MNC形状周长）



9.2 形状估计

(7) 在获得img中的目标MNC形状周长（称为MNCedgelet T ）之后重复步骤（3），以获得样本视频帧图像MNC粗糙边缘轮廓MNCedgelet R



(a) 视频帧目标形状周边

(b) 区域形状周边示例



9.2 形状估计

(8) 接下来从嵌入式MNC周边提取一对纯平面形状周长

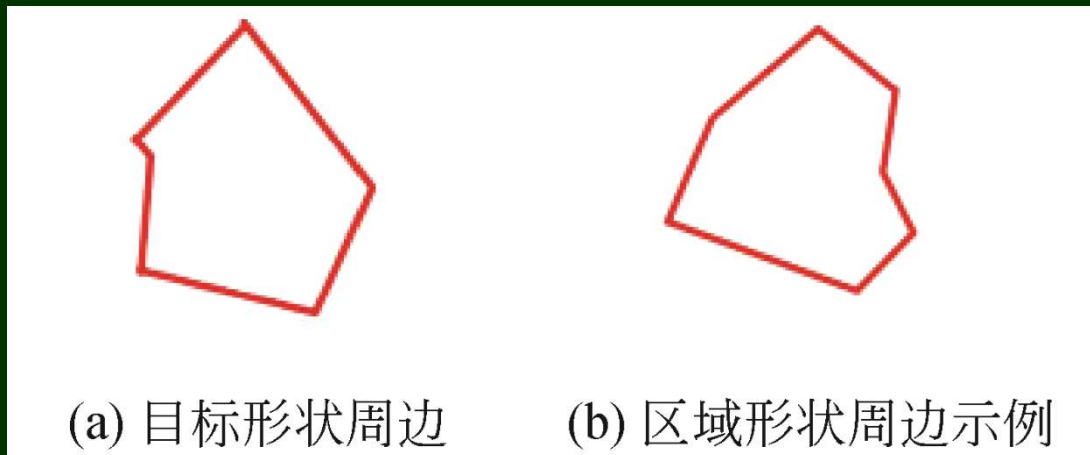
(9) 选取形状周边 $\text{edgelet} \in \{\text{MNCedgelet}T, \text{MNCedgelet}R\}$

(10) 从 edgelet 中提取形状周边 shape （此步骤的结果是纯平面形状周边，没有考虑图像MNC）

(11) 在获得第一个MNC形状周边（称之为 $\text{edgelet}T$ ）之后重复步骤（9），以获得样本MNC形状周边 $\text{edgelet}R$ （视频帧中的样本MNC形状周边）



9.2 形状估计



- (12) 在算法10中使用 $\text{edgelet}T$ 和 $\text{edgelet}R$ 作为输入（计算目标和样本MNC形状周长之间的相似性）
- (13) 计算 $\text{shapeSimilarity}(\text{edgelet}T, \text{edgelet}R)$



译者（章毓晋）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/
(下载更新的讲稿和教材修改表)