

# 前言

本书介绍数字图像和视频帧图像序列中计算几何学、拓扑学和物理学的内容。Edelsbrunner [1] 的网格生成几何和 Ziegler [2] 的多胞形几何为本书所研究的三角视觉场景的几何结构提供了**计算几何学**方法的坚实基础。**平面多胞形**是由覆盖多边形内部的闭合半平面相交所定义的填充多边形。此外，可以在 Peters [3] 的计算机视觉几何基础中找到本书中对计算几何的介绍。它是由亚历山德罗夫 [4, 5] (由 Cooke 和 Finney [6] 巧妙地扩展和阐述)、Borsuk [7, 8, 9] 所引入的单元复合形拓扑结构，最近由 Edelsbrunner 和 Harer 提出的这种拓扑结构 [10] 以及由 Munch [11] 在持久性同源性方面的工作为视觉场景的**计算拓扑学**的介绍性研究提供了坚实的基础。这种拓扑形式探索了视觉场景中常见的构造、形状和结构。结合视觉场景的固有几何和拓扑结构，则需要考虑对视频中记录的结构和事件产生的计算物理学以及随之而来的关于光的精细结构的敏感性。Nye [12] 介绍了我们所关注的光和光焦散的精细结构。对光焦散的考虑导致了突变理论和光焦散折叠和尖端的出现，从而在三角化数字图像中引入了光学涡旋神经。在这种情况下，**计算物理学**与在视频帧图像中光结构的编排研究是同义词。这种光结构的编排表现为一系列从表面形状反射和折射的光的快照，为研究视觉场景中出现的结构和形状提供了坚实的基础。

研究图像目标形状在视频帧图像序列以及记录视觉场景中表面形状变化的照片序列中的持久性非常重要。表面形状会出现、经历各种光线和表面条件的变化、并最终消失。人们通常倾向于在视觉场景中寻找不寻常的物体 (自然的和人工的)，这是一种对视觉场景中连续变化和观察到的成分的瞬间持久性的默认认识。换句话说，重要的是需考虑**视觉场景形状的时空特性**。这既指对视觉场景的理解不仅包括对视觉场景的几何和拓扑的研究，也包括对光物理学、光子在视觉场景中与曲面碰撞的特性和能量的考虑。

如果考虑对表面形状的描述和从照片中 (特别是在视频帧图像中) 记录的表面形状反射的光，**物理学**就会进入图像。计算机工程也通过对光子学和反射光捕获设备的研究而进入图像。就数字图像的物理学而言，能量的变形特性很重要。有关这种能量观点的更多信息，请参见 Susskind [13, §7, p.126]。

计算几何学有助于捕获嵌入在图像目标形状中的细粒度结构。计算拓扑学能够捕捉和分析嵌入在三角化视觉场景几何结构中的**单元复合形** (顶点、线段、填充三角形、循环、涡旋、神经的集簇) 里发现的近邻 (参见，例如，Peters [14, 15])。单元复合形的同源性 (亚历山德罗夫拓扑学方法的后代 [4]) 在这里是一个重要的组成部分。**同源性**是一个数学框架，它关注空间是如何连通的，并利用代数结构，如群和映射，将空间中具有拓扑意义的子集相互联系起来 [10, §IV.1, p.79]。

**群**  $G$  是一个非空集，它配备了一个结合的**二元运算**  $\circ$ ，其中有一个恒等元素  $e$  并且  $G$  中的每个成员  $a$  都有一个逆  $b$ ，即  $a \circ b = e$ 。**循环群**  $H$  是这样一个群，其中  $G$  的每个成员都可以写成称为生成元的单个元素的正整数幂。一个循环群是**阿贝尔群**，只要对于  $G$  中的每对元素，都有  $a \circ b = b \circ a$ 。**自由阿贝尔群**是一个有多个生成元的阿贝尔群，即该群中的每个元素都可以写成  $\sum_i g_i$ ，这里生成元  $g_i$  都在  $G$  中。从同源性观点对循环群的介绍可参见 Giblin [16, A.1, p.216]。

实际上，同源性是洞察视觉场景中的各个部分如何相互连通的来源。循环群有助于以简洁的方式表示视觉场景中的各个部分是如何相互连通并接合在一起的。具有多个生成元的循环群也是指定感兴趣表面形状的重要特征的一个来源，该特征即**贝蒂数** (Betti number，自由阿贝尔群中生成元数量的计数)。H. Poincaré 根据文章 [17] 为纪念 Enrico Betti 命名了这个数。数字图像的计算几何学、拓扑学和物理学的重点是有限空间。

贝蒂观察到，有限空间具有与其维度大小和元素形状无关的属性。这些属性仅指其**各部分的连通方式**…… [17, §3, p.143]。有限有界空间区域的特性倾向于通过覆盖空间的单元复合形中的通道所连通的顶点来揭示。(例如，参见覆盖由图 1 中的表面形状占据的有界区域的通道的连通涡旋)。

有关这方面的更多信息，请参见 Tucker 和 Bailey [18]、Salepci 和 Welshinger [19]以及 Pranav、Edelsbrunner、van de Weygaert 和 Vegter [20]。



图 1 覆盖形状的嵌套、非重叠涡旋

Kaczynski、Mischaikov 和 Mrozek [21]对同源性的计算方法进行了深入研究。这里的重点是辨别和跟踪、分析和表达以及近似移动表面形状的接近性。为了应对一个视觉场景中（从一个视频帧图像到另一个视频帧图像）的连续形状变化，描述性邻近空间上的特征矢量为我们提供了一种表示形状变化的方法，这些变化要么彼此靠近要么有时相距很远。有关这方面的更多信息，请参阅 Di Concilio、Guadagni、Peters 和 Ramanna [22]。

### 什么是平面上的表面形状？

 平面上的表面形状是欧氏平面上由简单封闭曲线（形状轮廓）限定的具有非空内部（形状内容）的有限区域。

在欧氏平面中，几何结构包括顶点、线段和实心三角形（三边多面体）。多胞形是封闭半平面的交集[2]。单个多胞形是一个空间区域，内部充满，四周都有边界。在拓扑设置中，重点是将视觉场景区域分解为非常简单的多胞形，例如易于测量和分析的填充三角形。这种拓扑的基本成分是单纯复合形、形状理论和持久同源性。

这项工作背后的秘密是将封闭的数字图像区域分解成一组形状复合形，这些形状复合形为形状分析提供了基础。形状复合形用嵌套的，通常不重叠的涡旋集族（例如，参见图 1）对形状进行覆盖。图 2(b)中显示了对图 2(a)<sup>1</sup>中那不勒斯花卉的部分三角剖分样本的分解。最终结果是用例如花瓣与填充三角形（单纯复形）的集族对图像场景形状的覆盖。在每个复合形中构成神经结构的三角形集族具有共同的顶点。例如，图 2(b)中覆盖白花的复合形为我们提供了一种测量、比较、描述和分类花瓣所占据场景片段的方法。一般来说，形状分析的目标是分类、比较、量化异同，并测量形状之间的距离[23]。在本书中，重点是计算拓扑在视觉场景形状分析中的应用。



(a) 视觉场景 (b) 形状复杂性

图 2 覆盖场景形状的单纯复形样本

{脚注 1: 非常感谢 Arturo Tozzi 的这张照片。}

对于三角化视觉场景中的单元复合形，场景形状被称为神经复合形（参见图 3）的填充三角形簇所覆盖。设  $K$  是点集的有限集族。集族  $K$  的神经（用  $Nrv K$  表示）由  $K$  的所有非空子集族组成，这些子集族具有非空交集[10, §III.2, p.9]。每条神经都有其独特的形状。例如，图 2(b)中覆盖花部分的神经的形状来自围绕单个顶点的填充三角形。在三角化表面上，亚历山德罗夫神经复合形  $A$ （用  $Nrv A$  表示）是具有公共顶点的三角形集族[4, §33, p.39]（例如，参见图 4）。



图 3 样本重叠的神经复合形



图 4 亚历山德罗夫神经复合形

本书不仅介绍了数字图像中的计算几何学、拓扑学和物理学的基础知识，而且还给出了许多实际应用。应用包括：

- 应用 1：单元分裂轨迹：3.13 节，应用 3.13。
- 应用 2：最大重心星状神经：4.3 节，应用 4.3。
- 应用 3：跟踪视频帧图像形状的变化：4.13 节。
- 应用 4：法医学形态理论中的光学涡流神经：4.14 节。
- 应用 5：时空涡旋循环：重叠电磁涡旋：5.7 节，应用 5.7。
- 应用 6：嵌套非同心涡旋特征矢量集族的比较：5.11 节，应用 5.11。
- 应用 7：基于强描述性连通性的零镜头识别：5.13 节，应用 5.13。
- 应用 8：物理目标形状分类中的描述近似性：6.5 节，应用 6.5。
- 应用 9：视频帧图像中形状的近似描述性接近度：7.8 节，应用 1。
- 应用 10：视频中尖端神经系统形状分类的近似描述接近度：8.13 节，应用 2。

本书中的章节源于我在过去几年中为本科生和研究生讲授的计算机视觉课程的笔记。本书中的许多主题都得益于我与一些研究人员、研究生和博士后的讨论和交流，特别是 Sheela Ramanna、Somashekhar Amrith Naimpally (1931–2014) [24、25]，Anna Di Concilio [26、27]，Clara Guadagni [28、29]，Luigi Guadagni、Fabio Marino、Giuseppe Di Maio、Giuseppe Gerla [30]，Gerald (Jerry) Beer、Arturo Tozzi [31]，Romi Tozzi（斐波那契数  $8$  和  $\infty$ ）和鼓舞人心的 Vittorio Tozzi、Andrew Worsley [32、33、34]，Alexander Yurkin [35]，Ebubekir İnan [36、37、38、39、40、41]，Mehmet Ali Öztürk [42、43、44]，Mustafa Uçkun [43]，Özlem Tekin [44]，Orgest Zaka [45]，Brent Clark（阿基米德世界移动支点的回声），Zdzisław Pawlak [46、47、48]，Andrzej Skowron [49、50]，Jarosław Stepaniuk、Jan G.

Bazan、Marcin Wolski [51], Piotr Wasilewski、Ewa Orłowska、W. Pedrycz、William (Bill) Hankley (时间逻辑), David A. Schmidt (集合理论), Joe Campbell、Rich McBride、Iraklii Dochviri [52、53], Hemen Dutta [54], Maciej Borkowski、Surabhi Tiwari、Sankar K. Pal、Cenker Sengoz、Doungrat Chitcharoen、Chris Henry [55、56、57], Dan Lockery as well as M. Zubair Ahmad、Arjuna P.H. Don、Maxim Saltyrnakov、Enoch A-iyeh、Randima Hettiarachchi、Dat Pham、Braden Cross、Homa Fashandi、Diba Vafabakhsh、Amir H. Meghdadi、Enze Cui、Liting Han、Fateme Gorgannejad、Maryam Karimi、and Susmita Saha。

我还要感谢 M. Zubair Ahmad、Sheela Ramanna 和 Fateme Gorgannejad, 他们对本书的部分内容提供了非常有帮助的见解、建议和更正。

本书的目标读者是工程、数学和物理科学的三年级和四年级本科生、一年级研究生, 以及计算几何学、拓扑学、物理学、数字图像处理和计算机视觉领域的研究人员。对每个主要章节主题, 给出了相关的算法。以下符号用于表示各章习题的难度级别:

🚲: 快速、易于解决;

🏔️: 深入、全面。

温尼伯, 加拿大

詹姆斯 F.彼得斯

## 参考文献

(xv ~ xvii)