arXiv:2301.04542v1 [q-bio.NC] 11 Jan 2023

三维中的良好延拓：立体视觉的神经几何学

M. V. BOLELLI[1],G. CITTI[1], A. SARTI[2], 和 S. W. ZUCKER [3]

摘要。图像曲线的经典良好延续以位置和方向为基础。皮层柱状组织、心理物理实验以及丰富的（微分）几何模型都为其提供了支持。本文将良好延续扩展到立体视觉。我们提出一种神经几何模型，其参数化同时涉及空间视差和方向视差。该模型有助于理解神经生物学，并提示神经相互作用存在一种隐式组织方式，以及一个定义良好的关联场。我们的模型阐明了对应问题背后的计算，并展示了现实世界中的良好延续如何推广平面中的良好延续。

关键词： 立体视觉、次黎曼几何、位置—方向空间、关联场、神经几何学。

1. 引言

双眼视觉是指视觉系统从二维视网膜图像出发，提供有关三维环境信息的能力。视差是深度知觉和立体视觉的主要线索之一；然而，为了提取视差，大脑需要确定右眼中的哪些特征与左眼中的哪些特征相对应，以及哪些不相对应。这会产生一个耦合问题，通常称为立体对应问题。本文的目标是提出一种关于立体视觉的知觉组织方法，将平面曲线的良好延续扩展到 $3D$ 空间曲线。

平面中的方向良好延续（视网膜拓扑坐标）是格式塔知觉组织的基础原则之一。这一原则有着悠久的历史 [88]。心理物理学研究（例如 [33, 30, 28, 31, 55]）为其提供了支持，这些研究揭示了它与轮廓统计之间的联络；生理学研究（方向选择性）也为其提供了支持，揭示了长程水平连接的作用 [12]；计算建模（[9, 78]）同样为其提供了支持，揭示了几何的重要作用。从历史上看，尽管深度中的良好延续与平面中的良好延续具有相近的起源，但前者的发展远不如后者充分。引用 Koffka[53,第 161-162] 页：

> .....一幅透视图即使以单眼观看，也不会像通过具有双眼视差的立体镜观看同一幅图那样，产生同样鲜明的深度印象.....因为在立体镜中，视差的三维力量与组织的其他三维力量相互协作；立体视觉引入的不是力量之间的冲突，而是相互增强。

本文的具体目标，是发展一种关于深度中的良好延续的理论，使之类似于二维中的轮廓组织模型。心理物理学研究表明，这应当是可行的（[50, 23, 24, 87]）；我们的重点将更多地放在数学方面。

具体而言，用 Koffka 的话来说，我们旨在建立一个关于“相互强化”的计算模型。尽管深度轮廓只比平面内的轮廓高出一个维数，但它们会带来一些微妙的新问题；

[1]数学系，意大利博洛尼亚大学。[2]CAMS，法国国家科学研究中心 - 法国社会科学高等研究院，巴黎，法国。[3]计算机科学系与生物医学工程系，耶鲁大学，美国康涅狄格州纽黑文。

这正是计算模型能够发挥启发作用的原因。首要问题是坐标的选择；当然，指定这些坐标需要一个数学框架。在平面内，位置与方向是自然的选择；曲线的平滑性则可通过曲率，或轮廓上相邻方向之间的关系来描述。对于立体视觉，左眼和右眼各自具有单眼结构。空间视差是联系二者的一个标准变量，而且众所周知，灵长类视觉系统会直接表征这一变量[71]。空间视差显然可以作为一种坐标。不过，其他生理学方面的情况则不那么明确。柱状结构在平面内的轮廓组织中作用强大，但它是单眼的；对于深度轮廓，情况很可能有所不同，至少在初级视觉皮层中存在眼优势带（见下一节）。尽管如此，方向选择性细胞为立体视觉提供输入，因此至少应涉及位置视差和方向——右眼对应一个方向，左眼可能对应另一个方向。尽管传统观点认为只有“相似”的方向才会匹配 [40, 60, 13, 16, 66], ，我们对方向视差的敏感性对此提出了质疑，因此方向视差也可能是一个变量。我们将说明，方向在立体视觉中确实发挥着重要作用，但将其表征为视差未必有效。此外，这也解决了立体视觉心理物理学中关于方向的一个经典争论：由于方向的生理实现可能与视差梯度相混淆 [63, 15],，方向可能是冗余的。但事实并非如此，因为真正起作用的是“梯度”的方向。因此，我们可以这样概述本文的主要技术目标：提供一种空间视差与方向的几何表征，以支持对良好延续的运用；这种表征既纳入生物学上的“既定事实”，又为立体对应问题奠定严谨的基础。正如曲线组织研究中那样，我们还认为，即便柱状组织尚不明确，我们的建模仍将揭示立体视觉底层的连接组学，至少能揭示其早期阶段的情况。如果像小鼠的方向柱那样 [74], ，支持立体视觉的神经结构仅隐含地体现在联络之中（而非明确排列成柱），那么这一点就尤为重要。

Hubel 和 Wiesel 引入了视差调谐神经元 [41]。他们观察到，单个神经元可以由双眼驱动，并且可以分别绘制每只眼睛的感受野（RF）。我们强调，这些单眼感受野对方向进行了调谐；有关生理学的更多内容，参见 [20] 。近期，已有研究探讨了视觉皮层的结构及其神经联络，这些结构和联络支撑着双眼视觉中立体对应关系的建立，相关研究见[68], ；而关于神经模型的综述可参见 [72]。

用于表达左右眼感受野组合的经典模型是双眼能量模型（*BEM*），最早由 [5]提出。该模型通过简单细胞的感受野轮廓编码视差，从而引出了位置视差和相位视差的可能性 [43]。然而，Read 和 Cumming [73], 在 [4], 的基础上指出，相位视差神经元往往会被错误对应对强烈激活。因此，人们普遍认为，感受野中最相关的视差仅仅是位置视差。然而，这种观点忽略了双眼之间的方向差异 [65], ，也就忽略了方向视差。尽管有人尝试扩展能量模型，以纳入感受野方向的双眼差异 [14], ，但这些尝试仍然十分有限。我们将要介绍的几何模型则直接纳入了方向差异。

基于神经模型的立体视觉数学模型还有很多。许多研究者（例如 [60]）指出，双眼的方向应当匹配，但允许存在细微差异。当然，这假设结构是正面平行的。随后，Jones 和 Malik [44] 使用了一组针对不同方向（和尺度）调谐的线性滤波器，但他们的算法并非建立在神经生理学基础之上。之后，Zucker 等人 [3, 57, 89] 构建了一个更受生物学启发的模型，探讨神经元之间的联络。他们的微分几何模型采用了位置、方向和曲率等参数，并在 $2D$视网膜平面中，以切向量表示方向，对双眼神经元进行建模。我们在此给出的结果与之相关，不过所涉及的几何更为深入。（我们将在下文中对此进行阐述。）一项较新的研究基于微分几何，具体而言是黎曼几何，

见 [64]。不过，在详细说明这些结果之前，我们先介绍将要使用的特定几何类型。它直接源于捕食动物和灵长类动物中常见的柱状组织结构。

1.1. 柱状结构与次黎曼几何。 我们提出一种次黎曼模型，用于描述立体视觉所依赖的皮层启发几何，该模型基于对来自双眼的信息中位置视差和方向差异的编码。我们以神经数学模型为基础，从 Hoffmann [37] 和 Koenderink-van Doorn [52], 的工作出发，特别关注单眼简单细胞的神经几何学（[18, 69, 70, 76, 77, 78]）。

为了阐明我们的数学方法，借助视觉皮层的抽象模型来展开讨论颇有启发。我们从单眼信息入手，这些信息在第4层被分隔到眼优势带 [56]中；这些神经元的突起延伸至浅层。我们在图 1.1中对此作了示意：图中，方向超柱阵列按照视网膜拓扑位置排列，并以主要眼输入进行着色。双眼驱动细胞往往更靠近眼优势边界，而单眼细胞则位于各带的中央。放大图突出显示了每个位置附近若干柱中的方向分布；水平连接（图中未示出）实现了这些单元之间的相互作用。由此引出本文的基本问题：*代表相邻位置处不同方向、并接受来自左眼和右眼输入的细胞群之间，究竟存在何种相互作用？* 生理学研究表明（见图1.1右图），答案在于单眼细胞和双眼细胞之间的相互作用；我们的模型从单眼细胞出发，对这种相互作用加以描述。

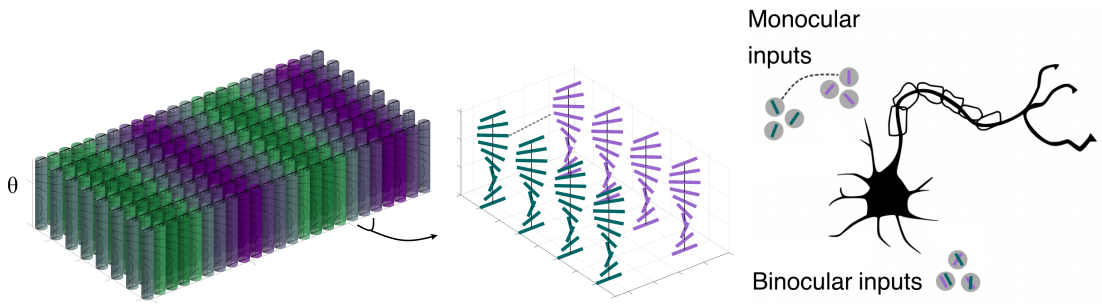


图 1.1. 视觉皮层初级视觉皮层（V1）浅层示意图。(A) 宏观组织：若干（抽象化的）方向超柱，按左眼（绿色）/右眼（紫色）的优势输入着色。颜色渐变强调了在眼优势带中心，细胞具有很强的单眼特性，而在边界处则由双眼输入驱动。(B) 放大显示若干方向柱，可见眼优势带边界处的左眼和右眼单眼细胞。这些邻近柱中的细胞将为我们的模型提供解剖学基础。(C) 较新的研究表明，单眼和双眼输入对这些细胞都很重要（根据雪貂数据改绘自[83], ）。这种更复杂的布线方式为我们模型中的连接结构提供了启示。

1.2. 非正式设定与概述。 由于本文的大部分内容都较为技术性，我们在此非正式地说明模型的主要组成部分和研究结果。我们先列出若干关键点，再对其进行直观说明。

- 立体几何具有一种数学结构，是平面曲线几何的形式扩展。在平面上，属于某条曲线的点由某个位置上的方向来描述

这些点自然表示为柱的元素（方向、位置）。在我们的模型中，它们成为抽象纤维。沿不同位置分布的纤维集合构成一个纤维丛。单目纤维的元素可以看作神经元。

- 对于立体视觉，我们需要由左眼和右眼的单眼柱构成的“乘积”纤维。立体纤维丛上的自然坐标分别是位置、位置视差以及来自左眼和右眼的方向，它们描述了每个位置上方的纤维。
- 立体系统的柱状组织，除图 1.1 所示内容外，完全不清楚。尽管 MT视觉区暗示存在表示运动方向的柱（[61, 22]），V4 也可能存在表示倾斜的柱（[36]），但据我们所知，初级视觉皮层中没有空间视差或方向视差柱的直接证据。这正是模型能够带来启发的原因。
- 曲率提供了一种“黏合剂”，使我们能够从纤维上的点过渡到附近纤维上的邻近点。这些过渡在立体纤维丛中确定了“积分曲线”。
- 积分曲线的观点提供了信息流的方向（信息在曲线束中扩散），从而揭示了潜在的神经回路。
- 积分曲线使关联场模型得以形式化。其参数描述了一簇曲线的延展情况，这与 $3D$ 在 [32, 34, 51]心理物理实验中研究的曲线十分吻合。
- 我们的形式理论解决了文献 [48, 49, 58]中的若干猜想。
- 我们的形式理论为描述对应问题提供了一个新框架，展示了三维世界中的良好延续如何推广二维平面中的良好延续。这正是双眼间一致的交互作用最为重要之处。
- 我们的形式理论对理解扭转眼球运动有直接的启示。具体而言，它指出，旋转分量并非仅仅是发育的结果，还能帮助抵消眼球运动所引起的不恰当方向视差变化。这为Listing定律赋予了新的作用，并将在一篇配套论文中讨论（准备中）。

现在我们来说明这些想法（图 1.2）。将三维刺激视为一条空间曲线 $\gamma : \mathbb{R} \longrightarrow \mathbb{R}^3$，其在注视点处具有单位切向量。由于切向量是曲线的导数，双眼细胞自然会编码指向空间 $3D$ 刺激 γ 的单位切向方向 $\dot{\gamma}$。这个空间切向量投影到左眼[1], 中形成一个切向方向，并可能投影到右眼中形成相同或不同的方向。邻近的空间切向量则投影为另一对单眼切线，如相邻柱中的活动所示。请注意，双眼神经元之间的联络如何支持沿空间曲线的一致性。我们通过立体关联场模型捕捉的，正是这种一致性关系。

由于空间曲线存在于三维空间中，因此需要两个角度来确定其在某一点处的空间切向量。换言之，单眼切向角在平面中构成一个圆；空间切向角则在三维空间中构成一个二维球面。就向左眼和右眼的投影而言，空间切向量可以用参数 $n = (\theta, \varphi)$ $\mathbb{S}^2$ 来描述（图 1.3）。因此，我们可以恰当地将立体细胞的空间——世界中任意位置处的全部空间切向量——描述为由位置和方向 $\mathbb{R}^3 \rtimes \mathbb{S}^2$构成的流形。将空间中的一个位置移至另一个位置，并将切向方向改变为新位置处的方向，这相当于在适当的流形上施加所谓的 群作用 。我们将在下一小节中非正式地介绍这些概念；附录 A 则对这些概念作更详尽的介绍。

1.2.1. 次黎曼几何。 我们生活在一个 $3D$ 距离十分常见的世界中；也就是说，这是一个由点构成的空间，任意一对点之间都定义了欧几里得距离函数。除了

[1]这里的表述并不严谨。我们所说的左眼中的切向量方向，指的是初级视觉皮层中左眼支配柱的方向

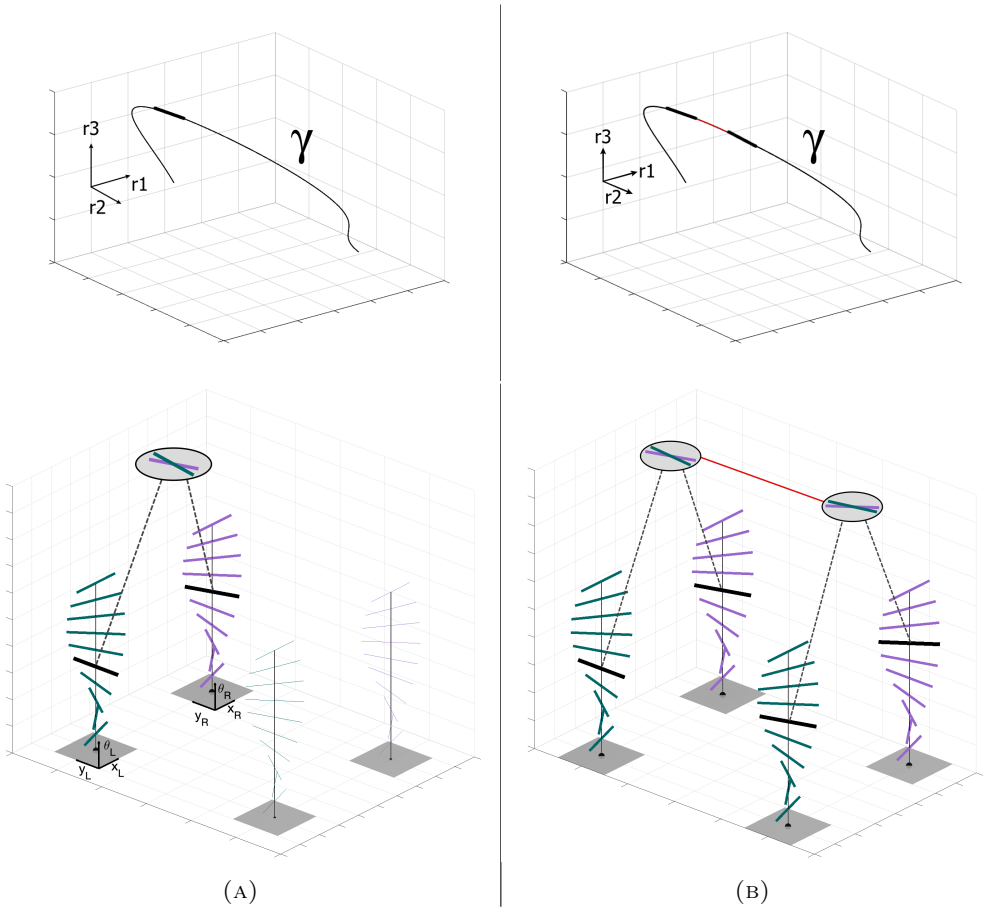


图1.2。（A）将突出显示的切向量立体投影到由左眼和右眼支配的单眼方向柱中的刺激物 $\gamma \in \mathbb{R}^3$ 上。（每条短线表示一个神经元，其方向由定向偏好表示。）双眼之间的联合活动表示空间切向量，由双眼神经元（圆圈）示意。注意两个相似但不同的单眼方向。从受到主动刺激的单眼神经元到双眼神经元的联络以虚线表示。（B）将一对连续的切向量立体投影到左、右视网膜柱中的刺激物 $\gamma \in \mathbb{R}^3$ 上。由于空间视差，每个空间切向量都投影到一对不同的单眼柱上。这四个柱的响应一致性对应于沿 γ附近位置附着的空间切向量之间的一致性。这种一致性通过双眼神经连接（实线）实现。

实际考量之外，我们可以向任何想去的方向移动。然而，汽车的移动能力受到更多限制。它们可以向前或向后移动，但不能横向移动。要向不同方向移动，汽车必须转动车轮。这里有一个基本的类比：在皮层空间中，信息可以沿切向方向移动到新的视网膜拓扑位置，也可以沿柱（定向纤维）向上或向下移动以改变方向。以这种方式移动——从某一位置上的一个方向移动到附近位置上的另一个方向——显然比在欧几里得空间中的任意移动更加受限。上述欧几里得几何是黎曼几何的一个例子；在皮层柱状空间中移动所涉及的限制

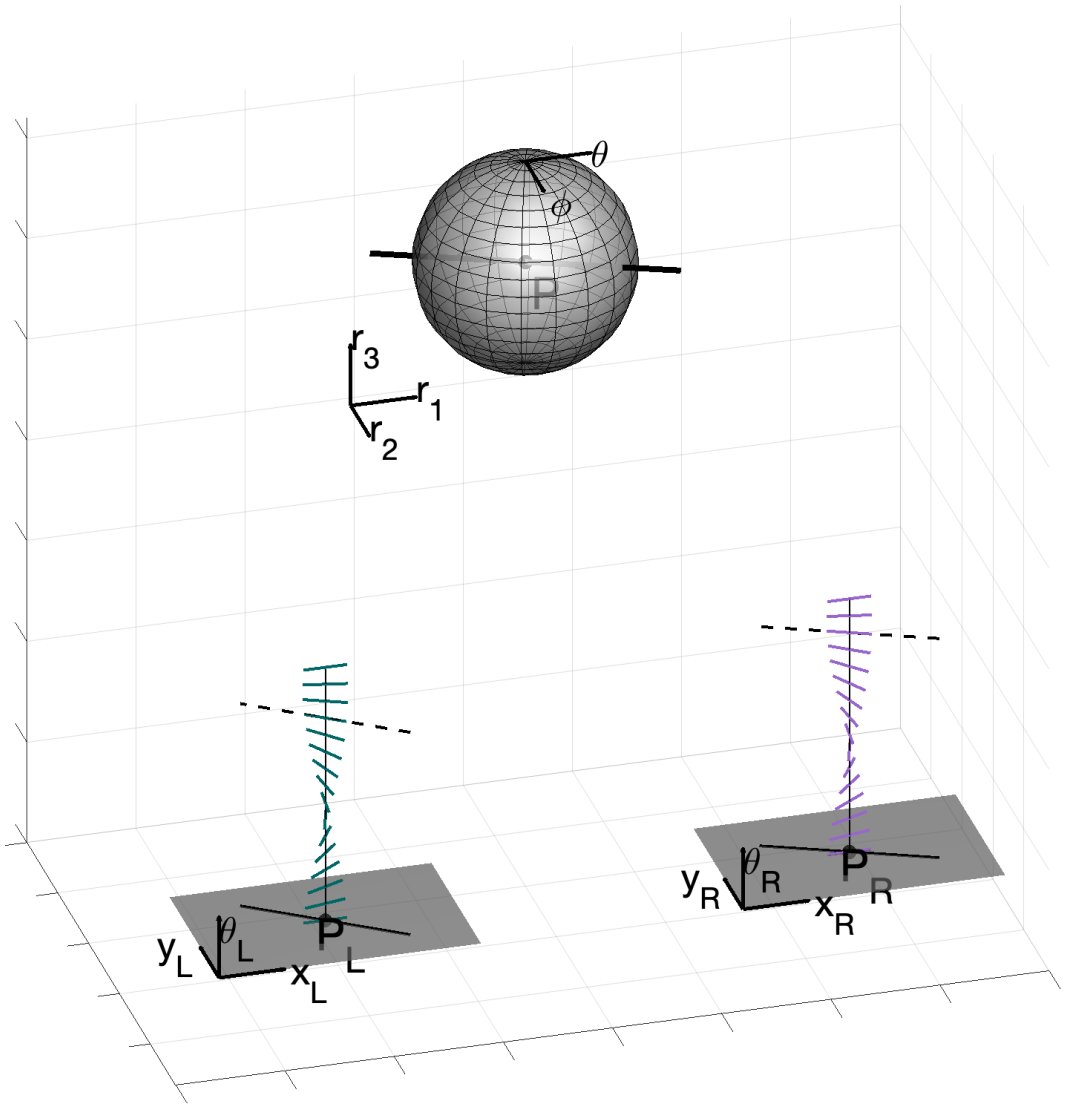


图 1.3。立体视觉的完整几何结构。请注意，立体对应问题如何帮助确立 $3D$ 切点（P，θ，ϕ）与投影 p_L 和 p_R、视差以及方向 θ_L和 θ_R之间的关系。

从而确定一种次黎曼几何。正如汽车可以沿着大多平坦的道路行驶，兴奋性神经元主要与方向相似的兴奋性神经元相连。这条神经元链表明了一条穿过次黎曼空间的路径；这类路径构成的扇形即为皮层连接，可视为关联场的神经相关物。同样，更多信息请参阅附录 A。

现在转向外部世界，我们必须能够在所有点之间移动。用更专业的方式重复上述比喻，我们为 $\mathbb{R}^3 \rtimes \mathbb{S}^2$ 赋予三维欧几里得刚体运动群 $SE(3)$的群作用。值得注意的是，该群现在作用于位置与方向的乘积空间。不过，仅凭这些关于点和方向的论述，还不足以解决立体视觉的几何问题。

正如我们在图 1.2 中所展示的，我们需要考虑邻近切线之间的关系；用几何学的语言来说，这涉及一种恰当类型的联络。因此，自然要考察次黎曼结构的积分曲线，它们的系数编码了 $3D$ 曲率和挠率这一基本概念。图 1.4 展示了一个例子。请注意， $3D$ 关联场如何包络一条空间曲线，正如 $2D$ 关联场包络一条平面曲线一样。该图以一种基本的方式阐释了本文的核心结果。

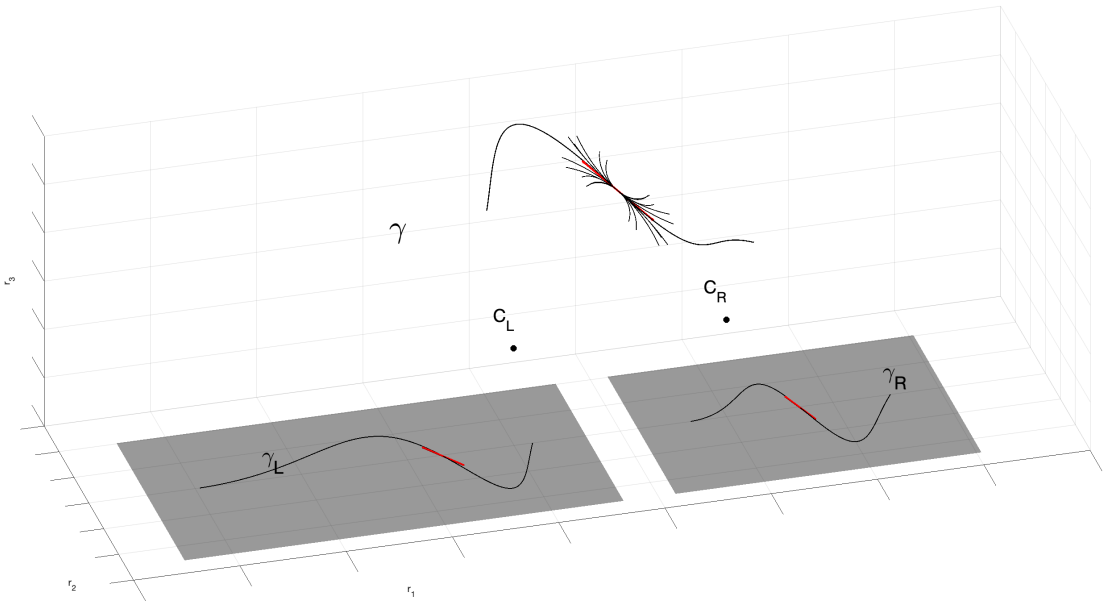


图 1.4。本文的主要结果。三维空间曲线 γ 被以一点为中心的关联场 $3D$ 所包络。形式上，这个关联场是在次黎曼几何中的积分曲线族，完全在柱状结构内计算得出。（具体而言，它由方程 (36) 描述，其中 c_1 和 c_2 在 $\mathbb{R}$中变化，但要推导出来还需要一些工作。）

1.3. **论文概述**。本文结构如下：第 2 节介绍立体视觉问题背后的几何学和神经数学背景。具体而言，我们回顾标准的立体三角测量技术，该技术用于关联一只视网膜与另一只视网膜的坐标系，并将二者结合起来重建三维空间。随后，我们简要回顾经典的单眼简单细胞神经几何，这些细胞对方向具有选择性，并介绍其背后的联络。我们还将介绍共圆性在立体视觉中的推广。第 3 节从双眼感受野轮廓出发，介绍双眼细胞的神经数学模型。首先，我们展示双眼细胞的皮层纤维丛。接下来，我们根据简单细胞的神经几何，对双眼轮廓作微分解释，并说明这与立体三角测量的结果十分吻合。然后，我们给出具有子黎曼结构的流形 $\mathbb{R}^3 \rtimes \mathbb{S}^2$ 的数学定义。最后，我们研究积分曲线以及适当的变量变换，以便将分析从皮层空间转换到外部空间。第 4 节将检验我们的几何模型是否符合心理物理实验。我们结合关于 $3D$ 知觉的心理物理学信息和形式猜想，并在此提出一个类似于 $2D$ 关联场的 $3D$ 关联场。最后，我们展示一个.....的示例

刺激的提升，并说明我们的积分曲线如何恰当地连接对应点。这展示了如何将我们的模型作为解决对应问题的基础。

2. 立体视觉与神经数学背景

2.1. 立体几何。 本小节简要回顾立体视觉背后的几何构型，以定义本文其余部分所用的变量，主要参考[29, 第 6] 章。关于完整的历史背景，请参阅 [38, 39]。

2.1.1. 立体视觉变量。我们考虑 $\mathbb{R}^3$中的全局参考系（O、i、j、k），其中 O =(0, 0, 0)，坐标为（r_1、r_2、r_3）。我们引入光学中心 C_L =（$-c$, 0, 0）和C_R =（c, 0, 0），其中 c为实数正值，并定义两个参考系：（C_L、i_L、j_L）、（C_R、i_R、j_R），分别为视网膜平面 $\mathcal{R}_L$ 和 $\mathcal{R}_R$ 的参考系，其坐标分别为（x_L、y）、（x_R、y）。在全局参考系中，我们假设视网膜平面彼此平行，其方程为 $r_3 = f$，其中 f 表示焦距。该几何设置如图 2.1 所示。

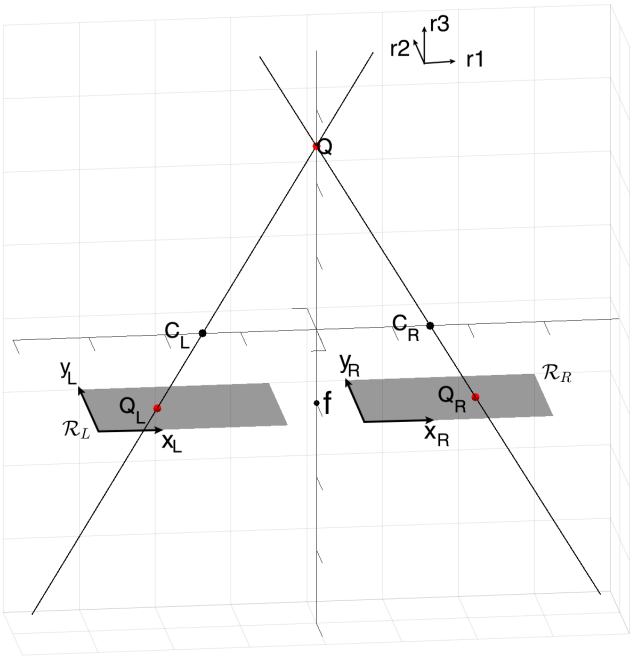


图 2.1. 通过视网膜平面 $\mathcal{R}_L$ 和 Q_R 中的点 Q_L 重建 $3D$ 空间点 Q， $\mathcal{R}_R$。

注 2.1. 如果我们知道点 $Q = (r_1, r_2, r_3)^T$ 在 $\mathbb{R}^3$ 中的坐标，那么利用透视投影很容易将其投影到两个平面上，其中已知光学中心的坐标 c 和焦距 f。这一计算分别为左、右视网膜平面定义了两个射影映射 Π_L 和 Π_R:

$$\Pi_L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2 \qquad \Pi_R : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$$

$$\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \mapsto \begin{pmatrix} \frac{f(r_1+c)}{r_3} \\ \frac{fr_2}{r_3} \end{pmatrix}, \qquad \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \mapsto \begin{pmatrix} \frac{f(r_1-c)}{r_3} \\ \frac{fr_2}{r_3} \end{pmatrix}. \tag{1}$$

证明。 左侧视网膜平面上局部坐标为 $(x_L, y)^T$ 的点，其全局坐标为$Q_L = (-c + x_L, y, f)^T$，并对应于欧几里得 $\mathbb{R}^3$中的一个点 $Q = (r_1, r_2, r_3)^T$，使得 C_L、 Q_L和 Q 共线。这意味着向量 $Q_L - C_L = (x_L, y, f)^T$与$Q - C_L = (r_1 + c, r_2, r_3)^T$平行，从而得到以下关系：

$$x_L = f\frac{r_1 + c}{r_3}, \quad y = f\frac{r_2}{r_3}. \tag{2}$$

类似地，考虑 Q_R 和 C_R，可得：

$$x_R = f\frac{r_1 - c}{r_3}, \quad y = f\frac{r_2}{r_3}. \tag{3}$$

□

按照标准定义，水平视差 是视网膜坐标之间的差值

$$d := \frac{x_L - x_R}{2}, \tag{4}$$

相差一个标量因子。此外，还可以将坐标 x 定义为两个视网膜坐标 $x := \frac{x_L + x_R}{2}$的平均值，从而得到以下变量变换：

$$\begin{cases} x = \frac{fr_1}{r_3} \\ y = \frac{fr_2}{r_3} \\ d = \frac{fc}{r_3} \end{cases} \quad \longleftrightarrow \quad \begin{cases} r_1 = \frac{xc}{d} \\ r_2 = \frac{yc}{d} \\ r_3 = \frac{fc}{d} \end{cases}, \tag{5}$$

其中，坐标集合 $(x,\ y,\ d)$ 称为独眼坐标 [46]。

2.1.2. 切向量估计。 视网膜平面上的对应点可以投影回$\mathbb{R}^3$。对于切向结构，也可以进行类似推理：如果我们知道视网膜平面上对应曲线的切向量，就可以通过反向投影恢复出 $3D$ 切向量的估计值。这里回顾一下这一结果；详细说明见[29]。

注 2.2。 设 γ_L 和 γ_R 分别为对应的左、右视网膜曲线；也就是说，它们是曲线 $\gamma \in \mathbb{R}^3$ 经过光学中心 C_L 和 C_R 、焦距为 f时的透视投影。已知左右视网膜上的切向结构后，就可以恢复切向量 $\dot{\gamma}$的方向。

证明。 从曲线 $\gamma \in \mathbb{R}^3$ 出发，将其投影到两个视网膜平面上，根据式 (1) 得到 $\gamma_L = \Pi_L(\gamma)$ 和 $\gamma_R = \Pi_R(\gamma)$。视网膜切向量通过左、右视网膜投影 $\dot{\gamma}_{L,R}(t) = (J_{\Pi_{L,R}})_{\gamma(t)}\dot{\gamma}(t)$ 的雅可比矩阵 [2] 得到：

$$\dot{\gamma}_R(t) = \begin{pmatrix} \frac{f(\gamma_3\dot{\gamma}_1 + (c-\gamma_1)\dot{\gamma}_3)}{\gamma_3(t)^2} \\ \frac{f(\gamma_3\dot{\gamma}_2 - \gamma_2\dot{\gamma}_3)}{\gamma_3^2} \end{pmatrix}, \dot{\gamma}_L(t) = \begin{pmatrix} \frac{f(\gamma_3\dot{\gamma}_1 - (c+\gamma_1)\dot{\gamma}_3)}{\gamma_3(t)^2} \\ \frac{f(\gamma_3\dot{\gamma}_2 - \gamma_2\dot{\gamma}_3)}{\gamma_3^2} \end{pmatrix}. \tag{6}$$

[2]在点 p 处取值的雅可比矩阵 $(J_\Pi)_p$ 表示如何投影位移向量（就导数、速度或方向而言）。具体来说，若 $\dot{\gamma}(t)$ 是 $\mathbb{R}^3$中的位移向量，那么矩阵乘积 $(J_\Pi)_{\gamma(t)}\dot{\gamma}(t)$ 也是一个位移向量，但位于 $\mathbb{R}^2$中。换言之，在 Π 可微的每一点上，雅可比矩阵都是Π 的微分；常见记法包括 J_Π 或 $D\Pi$。

将切向量和点延伸至 $\mathbb{R}^3$ 中，得到 $\tilde{t}_L = (\dot{\gamma}_{L1}, \dot{\gamma}_{L2}, 0)^T$，$\tilde{m}_L = (\gamma_{L1} - c, \gamma_{L2}, f)^T$，以及 $U_{t_L} = (P_L)^{-1}\tilde{m}_L \times (P_L^{-1})\tilde{t}_L$，其中投影矩阵为 $P_L = \begin{pmatrix} 1 & 0 & -c/f \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$。对于右侧结构，同理可得，其投影矩阵为 $P_R = \begin{pmatrix} 1 & 0 & c/f \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$。

那么，$U_{t_R} \times U_{t_L}$ 是一个平行于切向量 $\dot{\gamma}$ 的向量：

$$
\begin{aligned}
U_{t_R} \times U_{t_L} &= \left(\underbrace{\frac{f^4 2c(\dot{\gamma}_2\gamma_3 - \dot{\gamma}_3\gamma_2)}{\gamma_3^4}}_{\lambda(t)} \dot{\gamma}_1, \frac{f^4 2c(\dot{\gamma}_2\gamma_3 - \dot{\gamma}_3\gamma_2)}{\gamma_3^4}\dot{\gamma}_2, \frac{f^4 2c(\dot{\gamma}_2\gamma_3 - \dot{\gamma}_3\gamma_2)}{\gamma_3^4}\dot{\gamma}_3 \right)^T \\
&= \lambda(t)\left(\dot{\gamma}_1(t), \dot{\gamma}_2(t), \dot{\gamma}_3(t)\right)^T \\
&= \lambda(t)\dot{\gamma}(t).
\end{aligned} \tag{7}
$$

□

2.2. 神经数学要素。现在，我们介绍单目系统的几何建模以及平面上的良好延续的背景知识。我们的目标是阐明次黎曼几何在单目系统中的作用，这将作为从 Citti 和 Sarti 的神经数学模型出发，将其推广到立体系统的基础[18]。

2.2.1. 简单细胞的经典神经几何学。我们用皮层神经元感受野（RF）的感受剖面（RP）来建模其激活图 φ。一个经典示例是初级视觉皮层中简单细胞的感受野轮廓：这些细胞以位置 $(x,\ y)$ 和方向 θ 为中心，并被建模为一组 Gabor 滤波器 $\varphi_{\{x,y,\theta\}}$（例如，参见 [7, 21, 45]），用于处理视觉刺激。

形式上，可以将初级视觉皮层抽象为 $\mathbb{R}^2 \times \mathbb{S}^1$，即位置—方向空间，从而自然地编码 Hubel/Wiesel 超柱结构 [40]。图2.4 中的图像（A）展示了这种结构的一个示例，摘自 [9]。

根据 Citti 和 Sarti [18], 的神经数学模型，简单细胞的集合 RP 可以通过将向量 $(x,\ y)^T$ 平移，并从唯一的“母”轮廓 $\varphi_0(\xi,\ \eta)$ 旋转角度 θ 而得到

$$
\varphi_0(\xi,\eta) = \exp\left(\frac{2\pi i\xi}{\lambda}\right)\exp\left(-\frac{\xi^2+\eta^2}{2\sigma^2}\right), \tag{8}
$$

一个具有实部（偶函数）和虚部（奇函数）的 Gabor 函数（图2.2）。平移和旋转可以表示为：

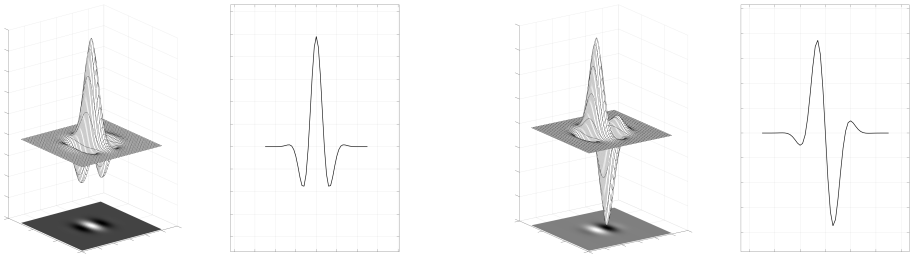
$$
T_{(x,y,\theta)}(\xi,\eta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \xi \\ \eta \end{pmatrix} + \begin{pmatrix} x \\ y \end{pmatrix}, \tag{9}
$$

其中，$T_{(x,y,\theta)}$ 表示旋转和平移群 $SE(2)$在 $\mathbb{R}^2$上的作用。该群运算将每个点（$\xi,\ \eta$）映射为一个新点（$\tilde{x},\ \tilde{y}$），其映射规律为（$\tilde{x},\ \tilde{y}$）$=T_{(x,y,\theta)}$（$\xi,\ \eta$）。因此，一般的 RP 可表示为

$$
\varphi_{(x,y,\theta)}(\xi,\eta) = \varphi_0(T^{-1}_{(x,y,\theta)}(\xi,\eta)), \tag{10}
$$

这表示群 $SE(2)$ 在感受野轮廓集合上的作用。

视网膜平面 $\mathcal{R}$ 被认定为 $\mathbb{R}^2$ 平面，其坐标为 (x,y)。当视觉刺激 $I\colon \mathcal{R} \longrightarrow \mathbb{R}^+$ 强度为 $I(x,y)$ 时激活视网膜层，以此为中心的神经元



(A) (B)

图2.2. Gabor 函数的偶部（A）和奇部（B）：二维滤波器的曲面、它们共同的二维表示以及一个一维截面。

在每个 点（x, y）产生输出 $O(x, y, \theta)$，可将其建模为 th 的积分 e
信号 I 与这组 Gabor 滤波器:

$$O(x, y, \theta) = \int_{\mathcal{R}} \varphi_{\{x,y,\theta\}}(\xi, \eta) I(\xi, \eta) d\xi d\eta, \tag{11}$$

其中，函数 I 表示视网膜图像。

¯对于固定的（x, y），我们将 θ 记为响应最大的点:

$$\max_{\theta} |O(x, y, \theta)| = |O(x, y, \bar{\theta})|. \tag{12}$$

¯于是，我们称点（x, y）被提升到点（x, y, θ）。从概念上说，理解我们的几何结构时，这一点极为重要：它说明，针对一个简单细胞 RP 对图像点进行评估后，通过显式引入方向，该图像点便被提升为一个“皮层”点。如果所有图像点都以相同方式提升，那么 $2D$ 图像 I 的等值线就会被提升为¯皮层空间 $3D$ 中的新曲线（x, y, θ）。

¯现在，我们将以向量场的形式，引入一组用于在皮层空间（x, y, θ）中移动的方向。这一点很重要，因为我们需要在这个空间中移动，跨越不同的位置和方向。从生物学角度看，这种移动对应于信息从一个柱中的细胞流向邻近柱中的另一个细胞。

首先，在方程（11）的右侧，表示了信号与 Gabor滤波器实部和虚部的积分。这两类细胞的形状不同，因此它们检测（或参与检测）不同的特征。由于奇对称细胞暗示其与边界检测有关，我们将重点讨论这类细胞，不过这主要是为了方便计算。简单细胞的输出随后可在局部近似为 $O(x, y, \theta) = -X_{3,p}(I_\sigma)(x, y)$，其中 $p = (x, y, \theta) \in SE(2)$， I_σ 是对 I进行平滑后的版本，通过与高斯核卷积得到，并且

$$X_{3,p} = -\sin\theta \partial_x + \cos\theta \partial_y, \tag{13}$$

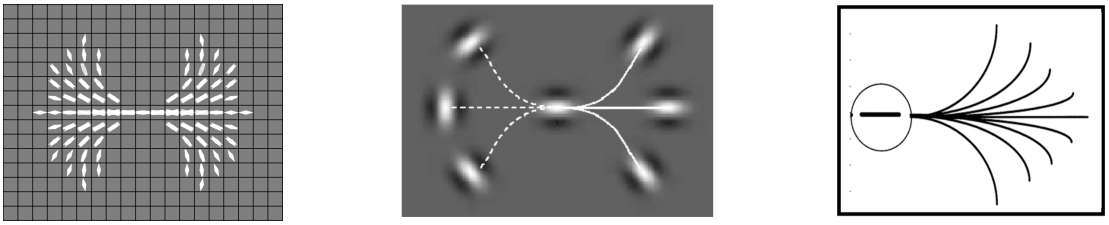
是沿方向 $\vec{X}_{3,p} = (-\sin\theta, \cos\theta, 0)^T$的方向导数。从现在起，我们将（稍微滥用一下记号）记作 $\omega^\star := \vec{X}_{3,p}$ 以提醒熟悉 1-形式的读者，这些量之间的对应关系，以及它们与霍奇星算子的关系。[3]

[3]引入这种记法的目的，也是为了说明数学模型 [18] 的一个推论；解释请参见附录 B.2.1。

现在，将向量场视为在皮层空间的每个点定义一个坐标系。于是，除上述内容外，与 $X_{3,\ p}$ 正交的向量场为：

$$X_{1,p} = \cos\theta\partial_x + \sin\theta\partial_y, \quad X_{2,p} = \partial_\theta \tag{14}$$

并且它们定义了一个到 $\mathbb{R}^2 \times \mathbb{S}^1$ 的二维可容许切丛[4] 。通过规定 $X_{1,p}$ 和 $X_{2,p}$ 正交归一，可以在这个空间上定义一个标量积：这就在 $\mathbb{R}^2 \times \mathbb{S}^1$ 上确定了一种次黎曼结构。



(A) (B) (C)

图 2.3。（A）由平面共圆性导出的、图像中心点周围兼容性的示例。亮度表示兼容性值。图片改编自 [9]。（B）从中心初始有向点出发，实线表示关联存在时图像块之间的构型，虚线表示关联不存在时的构型。图片改编自 [30]。（C）Field、Hayes 和 Hess 的关联场。图片改编自 [30]。

视觉信号沿着皮层连接以各向异性的方式传播，并且更紧密地连接方向相似的细胞。这种传播建立了刚才阐述的几何与二维轮廓整合之间的联系。这是对格式塔良好连续律的形式化表述 [53, 54]。它最初以一种更简单的形式出现，即平面上的共圆性 [67], ，用于描述相邻有向点的一致性和兼容性，并与特定的曲率值相符。这些兼容性的示例见图 2.3（A）。心理物理实验也为其提供了补充，例如 [42, 84, 86]。具体而言，Field 等人在 [30] 中描述了二维轮廓整合的关联规则，并引入了关联场这一概念。这些连接的示意图见图 2.3（B）和（C）。注意，这等价于 [67] 中按曲率取并集。神经生理学研究 [10, 12, 35, 59, 82] 表明，关联场的皮层相关物是相似（但不一定完全相同）定向偏好的细胞之间的长程水平连接。

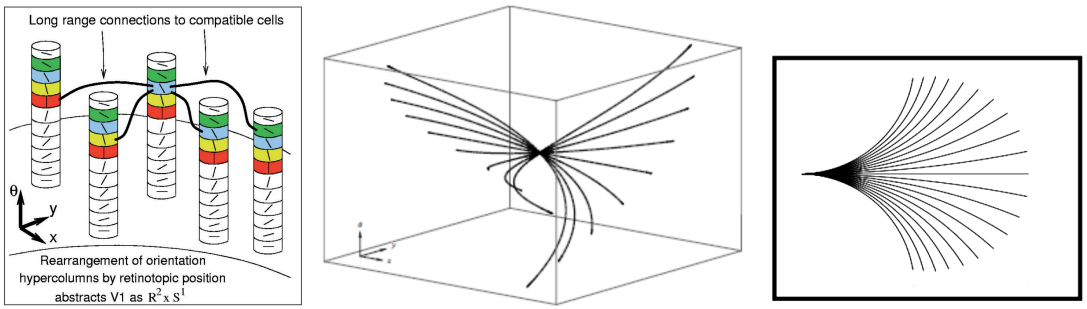
基于这些发现，Citti 和 Sarti 在 [18]中将皮层传播建模为沿向量场 X_1 和 X_2的积分曲线传播，即曲线 $\gamma\colon\ [0,T] \subset \mathbb{R} \longrightarrow \mathbb{R}^2 \times \mathbb{S}^1$由以下微分方程描述：

$$\dot{\gamma}(t) = \vec{X}_{1,\gamma(t)} + k\vec{X}_{2,\gamma(t)}, \quad t \in [0,T], \tag{15}$$

通过改变参数 $k \in \mathbb{R}$得到。（k 的作用类似于曲率。）图 2.4(B) 展示了这些曲线的一个示例。它们的 $2D$ 投影与关联场非常接近（图 2.4(B)）。

Duits 等人提出了一个相关模型 [26]。他们研究了子黎曼结构的测地线，以纳入关联场的所有适当终端条件。

[4]如附录 A.3 中所定义



(A) (B) (C)

图 2.4。（A）$(x,\ y,\ \theta)$ 坐标中的细胞方向柱。细胞之间长程的水平连接，将位置 $(x,\ y,\ \theta)$ 处的方向信号与 $(x',\ y',\ \theta')$ 处的另一方向联系起来。图片改编自 [9]。（B）由次黎曼模型[18]生成的 $\mathbb{R}^2 \times \mathbb{S}^1$ 中的水平积分曲线。（C）积分曲线族在 $(x,\ y)$ 平面上的投影。图片改编自 [18]。

2.2.2. 推广立体视觉中的共圆性。 通过观察二维曲线 γ 可在 0 点附近利用密切圆进行局部近似， $\mathbb{R}^2$ 中的共圆性概念得以发展。Zucker 等人于 [3, 57, 58] 中运用三维曲线的弗雷内微分几何推广了这一概念。

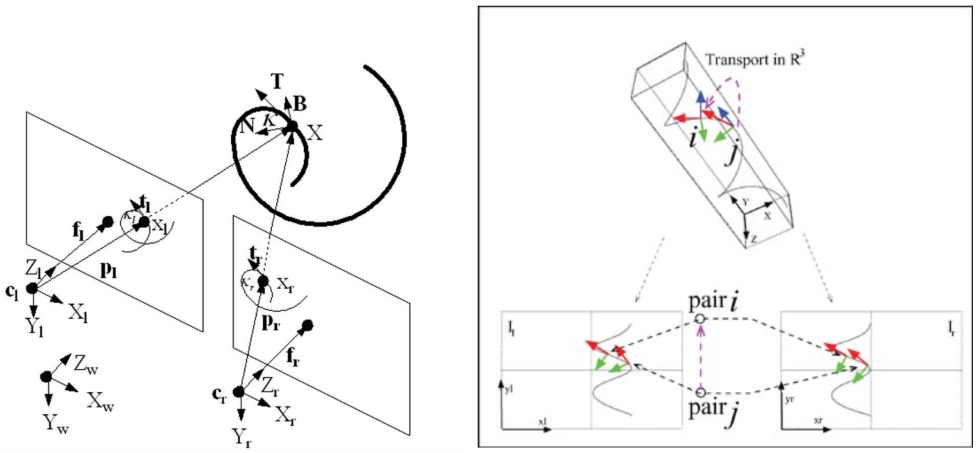
在二维情形中，使用弗雷内基底 $2D$ 近似曲线，会使曲率出现在泰勒级数展开的系数中（$1st$ 阶）；而在三维情形中，系数则同时包含曲率和挠率。因此，在 [3] 中，作者启发式地提出，用密切螺旋线替代密切圆来推广空间曲线的情形，并倾向于采用 r_3-螺旋线，以提高相机标定的稳定性。这样，方向视差便编码在螺旋线于 $3D$ 空间中的行为里：如果螺旋线被限制在正面平行平面内（此时螺旋线会变成圆），视网膜平面上的方向就没有差异；否则，沿着 $3D$ 曲线移动时，视网膜投影会呈现不同的方向。

在 [57, 58] 中，他们指出，通过将曲率变量作为特征引入两个单眼结构，并假设存在对应关系，可以从二维 Frenet 几何出发，重建曲线的 $3D$ Frenet 几何，但挠率参数除外。具体而言，他们证明：

命题 2.1。 给定一条 $3D$ 空间曲线的两个透视视图，且已完成完全标定，则可根据该曲线在空间点处的法向量 N 和曲率 k，唯一确定其投影在两幅图像中的位置、切向量和曲率。因此，可以唯一确定该空间点处的弗雷内标架 $\{T, N, B\}$ 和曲率 k。

因此，利用对弗雷内基的了解，并结合曲率变量这一基本补充，Zucker 等人提出了传输的概念。这使得$3D$ 弗雷内标架能够以一致的方式移动，并与左右视网膜平面上对应的 $2D$ 弗雷内结构相协调，从而在成对的（左、右）切向量对之间建立立体对应。见图 2.5（B）。

注 2.3. 本文提出的模型与刚才所述模型相关，但也有所不同。具体而言，为了与先前的神经几何模型保持直接兼容，我们只采用位置与方向这两个单目变量。我们不直接使用曲率，而是假设这些变量编码在联络之中；



(A) (B)

图2.5。（A）命题 2.1 的几何设置。螺旋曲线 $3D$连同弗雷内结构一起投影到左、右视网膜平面上。（B）切向量对（左—右）之间的立体对应关系。两幅图均取自 [58]。

在数学上，它们表现为参数。我们模型的一项理论成果是，现在可以严格证明关于 r_3-初级视觉皮层螺旋线的启发式假设。

我们还要提及论文 [1], ，其中曲率被视为自变量，并在 $2D$ 空间中得到了螺旋线。

3. 立体视觉的神经数学模型

3.1. 双眼轮廓。 双眼神经元接收来自左眼和右眼的输入。为便于计算，我们假设这些输入首先在初级视觉皮层的简单细胞中进行整合；这是一种广为研究的方法（[5, 20, 47, 62]）。这提供了一种初步近似，即双眼感受野轮廓可描述为单眼感受野的乘积；见图3.1，图像（A）。当然，这是一种简化——例如参见 [80], ——但它与现有的神经学发现相符。

这一双眼模型使我们能够定义视差和正面平行坐标，

$$\begin{cases} d = \frac{x_L - x_R}{2} \\ x = \frac{x_R + x_L}{2}, \end{cases} \tag{16}$$

这与式（4）中引入独眼坐标的方式完全一致。这样一来，通过变量变换（5），$(x,\ y,\ d)$便对应于$(r_1,\ r_2,\ r_3)$的神经相关物。

3.2. 双眼细胞的皮层纤维丛。 Petitot 和 Tondut 等人 [70], 在诸多研究中，将单眼简单细胞（方向选择性）的超柱结构描述为喷射纤维丛。我们着重研究纤维丛 $\mathbb{R}^2 \times \mathbb{S}^1$，其纤维为 $\mathbb{S}^1$；例如参见 [9] 等诸多研究。

在我们的设定中，双眼结构建立在单眼结构之上；请回想引言中的示意图。具体而言，对于左眼中的每个细胞，右眼中都有一整条细胞纤维；反过来，右眼中的每个细胞在左眼中也都有一整条细胞纤维。这意味着双眼空间具有一种涉及左右结构的对称性，使我们能够使用式（16）中定义的独眼坐标$(x,\ y,\ d)$。

因此，我们定义独眼视网膜 $\mathcal{R}$，并将其与 $\mathbb{R}^2$ 等同，赋予其坐标 (x, y)。纤维的结构为 $\mathcal{F} = \mathbb{R} \times \mathbb{S}^1 \times \mathbb{S}^1$，其坐标为 $(d, \theta_L, \theta_R) \in \mathcal{F}$。全空间为

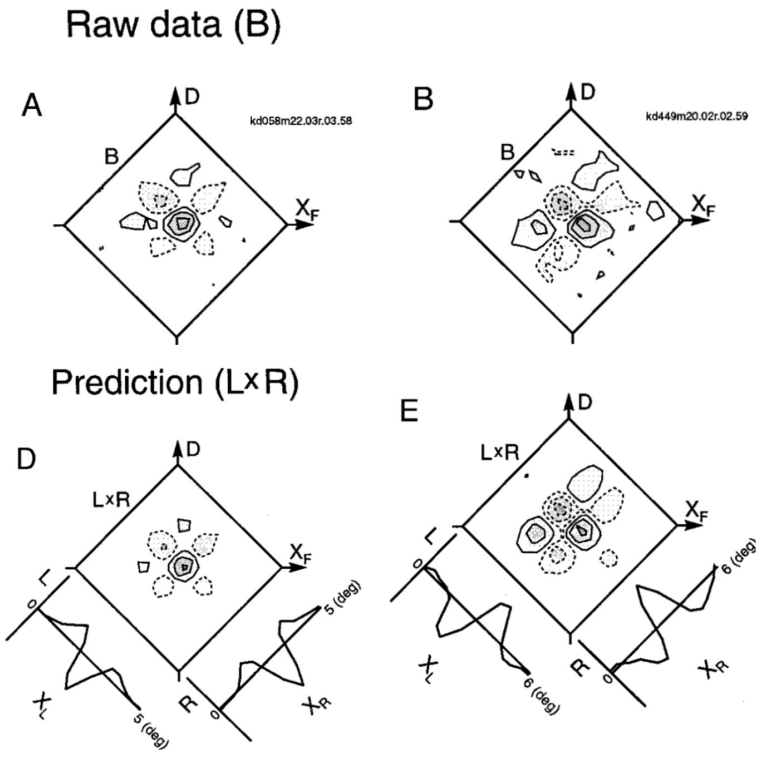


图3.1。双眼交互反应场与左右眼反应场乘积的比较，其中左右眼反应场见图2.2。左侧显示某个细胞的双眼交互反应场（原始数据）。右侧显示左右眼反应场乘积（L×R）的等高线图，以及左眼（L）和右眼（R）反应场的一维剖面。图改编自 [5]。

以平凡方式定义， $\mathcal{E} = \mathcal{R} \times \mathcal{F} = \mathbb{R}^2 \times \mathbb{R} \times \mathbb{S}^1 \times \mathbb{S}^1$，且投影 $\pi\colon \mathcal{E} \longrightarrow \mathcal{R}$ 是平凡投影 $\pi(x, y, d, \theta_L, \theta_R) = (x, y)$。对于每个 $(x, y) \in \mathcal{R}$，投影 $\mathcal{E}_{(x,y)} := \pi^{-1}(\{(x, y)\})$ 的原像均与纤维 $\mathcal{F}$ 同构，并且局部平凡化性质自然成立。

示意图见图3.2。考虑到独眼视网膜 $\mathcal{R}$ 在坐标 x上的限制 $\mathcal{R}_{|x}$ ，底座被描绘为一维。左图仅显示纤维 $\mathcal{F}$的视差分量，编码左右视网膜坐标之间的关系。右图显示了左右两侧的一维方向纤维。

3.3. 双眼能量模型。 为简化计算，正如引言中所述，我们采用经典的双眼能量模型 [5] 来描述双眼感受野轮廓。其基本思想是：双眼神经元接收来自每只眼睛的输入；如果来自左眼和右眼的输入之和 $O_L + O_R$ 为正，则双眼神经元的放电率与该输入之和的平方成正比；如果输入之和为负，则放电率为零：

$$O_B = (Pos(O_L + O_R))^2, \tag{17}$$

当 $Pos(x) = \max\{x, 0\}$ 时， O_B 双眼输出。

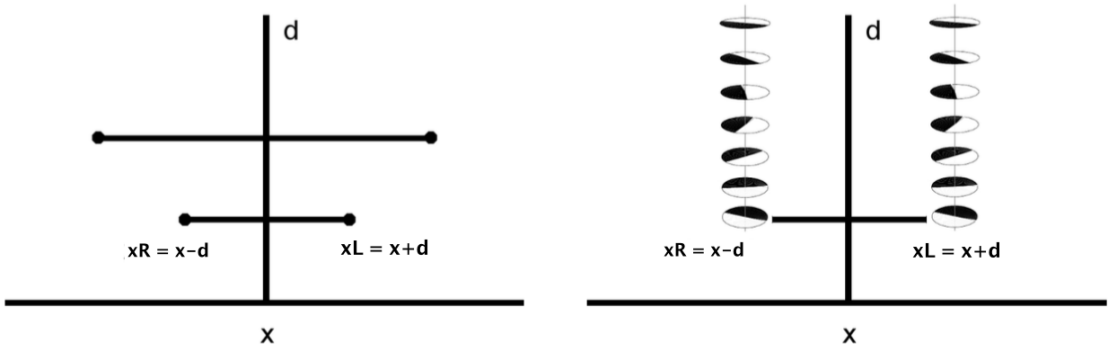


图3.2。左：二维纤维丛的示意图，以及左右视网膜坐标之间的关系。右：对左右简单细胞的一整条纤维的选择示意，其中每个x 以及每个 d均如此。

若 $O_L + O_R > 0$，则双眼简单细胞的输出可明确写为$O_B = O_L^2 + O_R^2 + 2O_L O_R$。前两项表示由单眼刺激引起的响应，而第三项 $2O_L O_R$ 可解释为双眼交互项。

随后根据细胞的输出测量其活动；对于彼此匹配概率较高的点，该活动最强。该量在 d 上的最大值就是所提取的视差。

值得注意的是，图3.1所示的双眼轮廓的神经生理学计算，假定单眼感受剖面具有一维性，忽略了有关单眼简单细胞方向的信息。然而，编码不同类型的方向视差需要这些信息。

注3.1（方向很重要）。*2001*年， [14] 的作者通过扩展 *Anzai*、*Ohzawa* 和 *Freeman* 的能量模型，将双眼感受野方向的差异纳入其中，开展了关于双眼神经元对方向视差响应的研究。最近，眼睛之间感受野方向的差异已得到证实[81]。

双眼能量模型是一种最简模型。它作为一个起点，使单眼输入得以结合。但它不足以解决立体匹配问题。

注 3.2 (联络**)**。 [68, 75] 中指出，除了将左右单眼结构相关联的特征（如信号、刺激或特定特征）的神经机制之外，还必须存在一个双眼细胞之间的联络系统，用以表征立体视觉的处理机制；另请参阅 [75] 中 *Samonds* 等人的相关论述。

3.4. **双眼感受野轮廓的微分解释**。可以用左、右感受剖面来表示式（17）中的相互作用项 $O_L O_R$：

$$
\begin{aligned}
O_L O_R &= \int \varphi_{\theta_L, x_L, y}(\tilde{x}_L, \tilde{y}_L) I_L(\tilde{x}_L, \tilde{y}_L) d\tilde{x}_L d\tilde{y}_L \int \varphi_{\theta_R, x_R, y}(\tilde{x}_R, \tilde{y}_R) I_R(\tilde{x}_R, \tilde{y}_R) d\tilde{x}_R d\tilde{y}_R \\
&= \int\int \varphi_{\theta_L, x_L, y}(\tilde{x}_L, \tilde{y}_L) \varphi_{\theta_R, x_R, y}(\tilde{x}_R, \tilde{y}_R) I_L(\tilde{x}_L, \tilde{y}_L) I_R(\tilde{x}_R, \tilde{y}_R) d\tilde{x}_R d\tilde{y}_R d\tilde{x}_L d\tilde{y}_L .
\end{aligned} \tag{18}
$$

如果固定（$\tilde{x}_R$，$\tilde{y}_R$，$\tilde{x}_L$，$\tilde{y}_L$），我们便可将双眼轮廓 $\varphi_{L,\ R} = \varphi_{\theta_R,\ x_R,\ y}\varphi_{\theta_L,\ x_L,\ y}$表示为单眼左、右轮廓的乘积。这与图3.1所示的测量轮廓一致。

命题 3.1。 双眼交互项可与由（13）定义的左右方向的叉积相关联，即单眼简单细胞的 $\omega_L^\star$ 和 $\omega_R^\star$：

$$O_L O_R = \omega_L^\star \times \omega_R^\star. \tag{19}$$

证明。 其思路是，双眼输出是单眼细胞左右作用的综合结果，因此确定了独眼坐标空间中的一个方向。该命题的详细证明见附录B。 □

为了更好地理解命题 3.1 背后的几何思想，我们回顾一下，视网膜坐标可用独眼坐标（4）表示为 $x_R = x - d$和 $x_L = x + d$，因此我们可以在坐标的 $3D$空间（$x,\ y,\ d$）中写出 $\omega_L^\star$和 $\omega_R^\star$：

$$\omega_R^\star = (-\sin\theta_R, \cos\theta_R, \sin\theta_R)^T \qquad \omega_L^\star = \ (-\sin\theta_L, \cos\theta_L, -\sin\theta_L)^T. \tag{20}$$

We 将 $\omega_{bin} := \omega_L^\star \times \omega_R^\star$ 定义为表征双眼结构的自然方向 ：

$$\omega_{bin} = \begin{pmatrix} \sin(\theta_R + \theta_L) \\ 2\sin\theta_R \sin\theta_L \\ \sin(\theta_R - \theta_L) \end{pmatrix}. \tag{21}$$

注记3.3。 方程 (21) 中的向量 ω_{bin} 可以解释为：在独眼坐标(x, y, d)中表示时，相对于 $\omega_R^\star$ 和 $\omega_L^\star$ 所定义的正交空间的交集。更准确地说，如果

$$(\omega_L^\star)^\perp = \operatorname{span}\left\{ \begin{pmatrix} \cos\theta_L \\ \sin\theta_L \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}, \qquad (\omega_R^\star)^\perp = \operatorname{span}\left\{ \begin{pmatrix} \cos\theta_R \\ \sin\theta_R \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\} \tag{22}$$

then

$$\omega_{bin} = (\omega_L^\star)^\perp \cap (\omega_R^\star)^\perp. \tag{23}$$

这些单眼结构的交集确定了一个方向，如图 *3.3* 所示。

我们先前指出，单眼奇简单细胞的作用结果是选择信息传播的方向。现在，我们将双眼的情况结合起来，说明在三维情形下，双眼神经机制也会导出一个方向。我们将在接下来的章节中看到，只要点彼此对应，这个方向就是立体刺激的切向量方向。

3.5. 与立体几何的兼容性。 我们考察在 (21) 中定义的表征双眼结构 ω_{bin} 的方向，并证明它可以与曲线 $3D$ 的切向量 $3D$ 联系起来。其思路是，该切向量既垂直于 $\omega_R^\star$，也垂直于 $\omega_L^\star$，因此其方向与向量积 $\omega_L^\star \times \omega_R^\star$的方向一致。

具体而言，我们考虑视网膜平面上的归一化切向量 t_L 和 t_R

$$t_R = (\cos\theta_R, \sin\theta_R)^T \quad t_L = (\cos\theta_L, \sin\theta_L)^T, \tag{24}$$

分别对应点（$x_R,\ y$）和（$x_L,\ y$）。考虑到 f 是 $\mathbb{R}^3$ 中视网膜平面的焦坐标，我们将这些点与其在 $\mathbb{R}^3$ 中的对应点关联起来，即

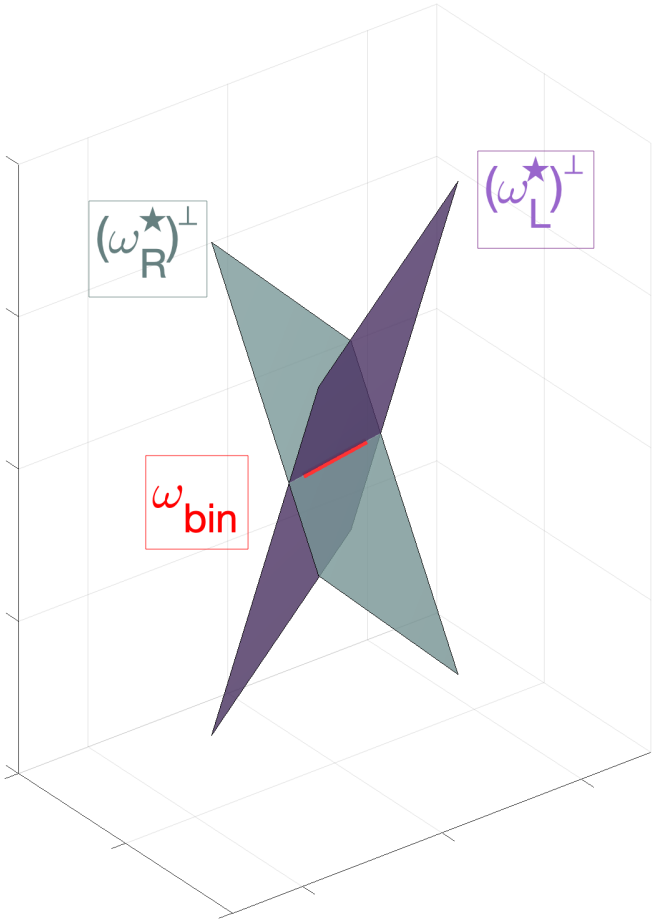


图 3.3. 初级视觉皮层 ω_{bin} 通过左右平面（$\omega_R^\star$）$^\perp$ 和（$\omega_L^\star$）$^\perp$的交集检测到的方向。红色向量对应相关的 2-形式 ω_{bin}。

$\tilde{m}_L = (x_L - c,\ y,\ f)^T$，$\tilde{m}_R = (x_R + c,\ y,\ f)^T$。应用方程（7），可以推导出三维轮廓的切向量：

$$
\begin{aligned}
U_{t_L} = P_L^{-1}\tilde{m}_L \times P_L^{-1}\tilde{t}_L &= \begin{pmatrix} x_L \\ y_L \\ f \end{pmatrix} \times \begin{pmatrix} \cos\theta_L \\ \sin\theta_L \\ 0 \end{pmatrix} = \begin{pmatrix} -f\sin\theta_L \\ f\cos\theta_L \\ x_L\sin\theta_L - y_L\cos\theta_L \end{pmatrix}, \\
U_{t_R} = P_R^{-1}\tilde{m}_R \times P_R^{-1}\tilde{t}_R &= \begin{pmatrix} x_R \\ y_R \\ f \end{pmatrix} \times \begin{pmatrix} \cos\theta_R \\ \sin\theta_R \\ 0 \end{pmatrix} = \begin{pmatrix} -f\sin\theta_R \\ f\cos\theta_R \\ x_R\sin\theta_R - y_R\cos\theta_R \end{pmatrix},
\end{aligned} \tag{25}
$$

而切向方向可通过以下方式恢复：

$$
U_{t_L} \times U_{t_R} = f \begin{pmatrix} \frac{x_L+x_R}{2}\sin(\theta_R - \theta_L) - \frac{x_R-x_L}{2}\sin(\theta_L + \theta_R) \\ y\sin(\theta_R - \theta_L) - (x_R - x_L)(\cos(\theta_R - \theta_L) - \cos(\theta_L + \theta_R)) \\ f\sin(\theta_R - \theta_L) \end{pmatrix} \tag{26}
$$

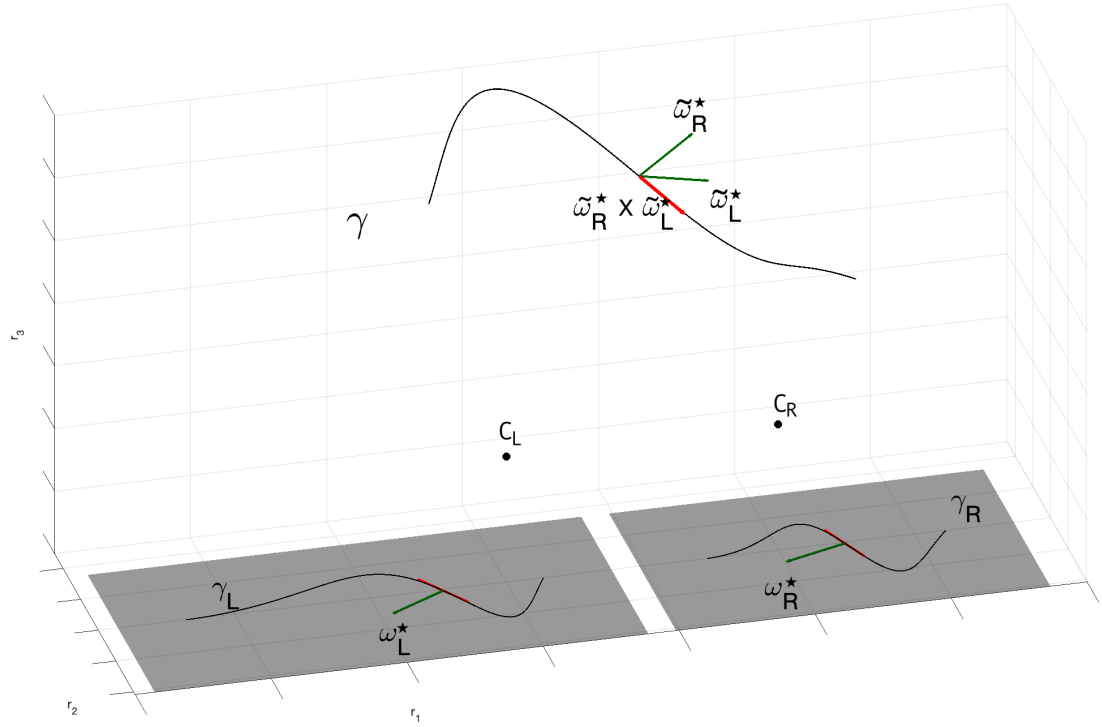


图 3.4。从视网膜平面重建三维空间。1-形式 $\omega^\star_L$ 和 $\omega^\star_R$ 分别与曲线 γ_L 和 γ_R 的法向量相对应。它们的三维对应物 $\tilde{\omega}^\star_L$ 和 $\tilde{\omega}^\star_R$ 通过叉积 $\tilde{\omega}^\star_L \times \tilde{\omega}^\star_R$ 确定曲线 $\gamma\colon\ \mathbb{R} \to \mathbb{R}^3$ 的切向量。

如果我们定义

$$\tilde{\omega}^\star_L := \frac{d}{fc} U_{t_L}, \qquad \tilde{\omega}^\star_R := \frac{d}{fc} U_{t_R} \tag{27}$$

以及相应的 2-形式 $\omega_{\mathbb{R}^3} := \tilde{\omega}^\star_L \times \tilde{\omega}^\star_R$，利用变量变换（16），我们可以观察到：

$$\tilde{\omega}^\star_L = \omega^\star_L, \quad \tilde{\omega}^\star_R = \omega^\star_R, \quad \omega_{\mathbb{R}^3} = \omega_{bin}, \tag{28}$$

相差一个标量因子。显式计算请参见附录 C。

这样，视差双眼细胞便以自然的方式将位置与方向联系起来：位置对应于 $\mathbb{R}^3$ 中的点，方向对应于 $\mathbb{S}^2$ 中的三维单位切向量。正如注 3.2 中所指出的，仅凭这些点和方向方面的论证，尚不足以解决立体视觉的几何问题，还需要考虑适当类型的联络。在 [3, 57, 58], 中，Zucker 等人提出了一种模型，将单眼结构的曲率视为一个附加变量。相反，我们主张考虑对方向具有选择性的简单单眼细胞，并将曲率的概念直接纳入联络的定义。因此，自然的做法是通过流形 $\mathbb{R}^3 \rtimes \mathbb{S}^2$ 引入知觉空间，并寻找适当的曲线。

3.6. 三维位置—方向空间中的知觉模型。我们现在推导图 1.3 中的对象。我们已经明确（第 3.5 节末尾），双眼细胞由 $\mathbb{R}^3$ 中的点和 $\mathbb{S}^2$ 中的方向参数化。空间 $\mathbb{R}^3 \rtimes \mathbb{S}^2$ 的一个元素 ξ 由 $\mathbb{R}^3$ 中的一个点 $p =$ （$p1$、$p2$、$p3$）和一个单位向量 $n \in \mathbb{S}^2$ 定义。由于该几何对象的拓扑维数

为 2，我们引入经典的球坐标（θ、φ），使得 $n = （n_1、n_2、n_3）^T \in \mathbb{S}^2$ 可以参数化为：

$$\begin{aligned} n_1 &= \cos\theta \sin\varphi \\ n_2 &= \sin\theta \sin\varphi \\ n_3 &= \cos\varphi \end{aligned} \tag{29}$$

与 $\theta \in [0, 2\pi]$ 和 $\varphi \in (0, \pi)$。使用局部坐标图所产生的歧义，通过引入覆盖奇异点的第二个坐标图得以克服。

平移和旋转通过三维特殊欧几里得群 $SE(3)$ 的群运算律来表示，从而定义群作用

$$\sigma : \mathbb{R}^3 \rtimes \mathbb{S}^2 \times SE(3) \longrightarrow \mathbb{R}^3 \rtimes \mathbb{S}^2 \text{ s.t. } \sigma((p,n),(q,R)) = (Rp+q, Rn), \tag{30}$$

其中 $(p, n) \in \mathbb{R}^3 \rtimes \mathbb{S}^2$，$(q, R) \in SE(3)$，即 $R \in SO(3)$ 三维旋转，以及 $q \in \mathbb{R}^3$。

3.6.1. 立体次黎曼几何。 在 $\mathbb{R}^3$ 中出现优先方向（与刺激的切向量相关），这就是我们赋予 $\mathbb{R}^3 \rtimes \mathbb{S}^2$ 次黎曼结构的原因，该结构偏向于由 ω_{bin} 在 $3D$ 中确定的方向。

形式上，我们考虑由向量场描述的 $\mathbb{R}^3 \rtimes \mathbb{S}^2$ 中的可容许运动：

$$\begin{aligned} Y_{\mathbb{R}^3,\xi} &= \sin\varphi\cos\theta\partial_1 + \sin\varphi\sin\theta\partial_2 + \cos\varphi\partial_3 \\ Y_{\theta,\xi} &= -\frac{1}{\sin\varphi}\partial_\theta \\ Y_{\varphi,\xi} &= \partial_\varphi \end{aligned} \tag{31}$$

其中 $\xi \in \mathbb{R}^3 \rtimes \mathbb{S}^2$ 对应 $\varphi \neq 0$，$\varphi \neq \pi$。点 ξ 处的可容许切空间[5]

$$\mathcal{A}_\xi := \text{span}\{Y_{\mathbb{R}^3,\xi}, Y_{\theta,\xi}, Y_{\varphi,\xi}\} \tag{32}$$

正如 Duits 和 Franken 在 [27] 中指出的，它编码了位置与方向之间的耦合。具体而言，向量场 $Y_{\mathbb{R}^3}$ 确定了 $\mathbb{R}^3$ 中的优先方向，而 Y_θ 和 Y_φ 则允许改变这一方向，且仅涉及 $\mathbb{S}^2$ 的方向变量。向量场 $\{Y_{\mathbb{R}^3}, Y_\theta, Y_\varphi\}$ 及其交换子在点 $\mathbb{R}^3 \rtimes \mathbb{S}^2$ 处生成切空间，从而能够沿优先方向连接流形上的任意两点（*Hörmander* 条件）。此外，还可以通过在可容许切丛 $\mathcal{A}$ 上选取标量积来定义次黎曼结构：最简单的选择是令向量场 $\{Y_{\mathbb{R}^3}, Y_\theta, Y_\varphi\}$ 构成正交归一组，并在 $\mathbb{S}^2$ 上采用由其浸入 $\mathbb{R}^3$（具有欧几里得度量）所诱导的距离。

3.6.2. 变量变换。 我们已经在方程 (5) 中将变量 (x, y, d) 变换为 (r_1, r_2, r_3)。然而，皮层坐标还包含角变量 θ_R 和 θ_L，这需要引入球坐标 θ, φ。

为了确定这些变量之间的变量变换，我们首先引入函数 $(r_1, r_2, r_3, \theta, \varphi) \xrightarrow{F} (x, y, d, \theta_L, \theta_R)$：

$$\begin{aligned} F: \quad \mathbb{R}^3 \rtimes \mathbb{S}^2 &\longrightarrow \mathbb{R}^3 \rtimes \mathbb{S}^2 \\ \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ \theta \\ \varphi \end{pmatrix} &\mapsto \begin{pmatrix} \frac{f r_1}{r_3} \\ \frac{f r_2}{r_3} \\ \frac{cf}{r_3} \\ \tan^{-1}\left(\frac{r_3 \sin\theta\cos\varphi - r_2\cos\varphi}{r_3\cos\theta\sin\varphi - (c+r_1)\cos\varphi}\right) \\ tan^{-1}\left(\frac{r_3\sin\theta\cos\varphi - r_2\cos\varphi}{r_3\cos\theta\sin\varphi - (c-r_1)\cos\varphi}\right) \end{pmatrix}, \end{aligned} \tag{33}$$

其中，视网膜右角 $\theta_R = \tan^{-1}($ r_3 θ $\frac{\varphi - r_2}{\varphi - c + r_1}$ φ 正弦 θ 余弦 余弦 余弦 正弦 $)$ 余弦 $\varphi)$ 和左视网膜角 $\theta_L = \tan^{-1}($ $\frac{r_3}{r_3}$ θ $\frac{\varphi - r_2}{\varphi - c - r_1}$ φ 正弦 θ 余弦 余弦 余弦 正弦 $)$ 余弦 $\varphi)$ 是根据方程 (6) 得出的。

[5]允许切空间的定义见附录 A。

类似地，可以定义变量替换 $(x,\ y,\ d,\ \theta_L,\ \theta_R) \xrightarrow{G} (r_1,\ r_2,\ r_3,\ \theta,\ \varphi)$:

$$
\begin{aligned}
G: \quad \mathbb{R}^3 \rtimes \mathbb{S}^2 \quad &\longrightarrow \quad \mathbb{R}^3 \rtimes \mathbb{S}^2 \\
\begin{pmatrix} x \\ y \\ d \\ \theta_R \\ \theta_L \end{pmatrix} \quad &\mapsto \begin{pmatrix} \frac{cx}{d} \\ \frac{cy}{d} \\ \frac{cf}{d} \\ \tan^{-1}\left(\frac{2\sin\theta_R \sin\theta_L}{\sin(\theta_R+\theta_L)}\right) \\ \tan^{-1}\left(\frac{\sqrt{\sin^2(\theta_R+\theta_L)+4\sin^2\theta_R\sin^2\theta_L}}{\sin(\theta_R-\theta_L)}\right) \end{pmatrix},
\end{aligned} \tag{34}
$$

其中，角度 $\theta = \tan^{-1}(2\sin\theta_R \sin\theta_L \sin(\theta_R+\theta_L))$ 和 $\varphi = \tan^{-1}(\sqrt{\sin^2(\theta_R+\theta_L)+4\sin^2\theta_R}\sin^2\theta_L \sin(\theta_R-\theta_L))$ 是通过考虑 $\tan\theta = \frac{(\vec{Y}_{\mathbb{R}^3})_2}{(\vec{Y}_{\mathbb{R}^3})_1}$ 以及 $\tan\varphi = \frac{\sqrt{(\vec{Y}_{\mathbb{R}^3})_1^2+(\vec{Y}_{\mathbb{R}^3})_2^2}}{(\vec{Y}_{\mathbb{R}^3})_3}$ 得到的。

3.6.3. 积分曲线。空间的连通性由张成 $\mathcal{A}$ 的向量场的可容许曲线描述。特别地，曲线 $\Gamma: [0,T] \longrightarrow \mathbb{R}^3 \rtimes \mathbb{S}^2$ 被称为可容许的[6]，当且仅当:

$$
\dot{\Gamma}(t) \in \mathcal{A}_{\Gamma(t)}, \leftrightarrow \dot{\Gamma}(t) = a(t)\vec{Y}_{\mathbb{R}^3,\Gamma(t)} + b(t)\vec{Y}_{\theta,\Gamma(t)} + c(t)\vec{Y}_{\varphi,\Gamma(t)}, \tag{35}
$$

其中，$a,\ b,\ c$ 是在 $[0,T]$上足够光滑的函数。我们将考虑这类可容许曲线中的一个特例，即满足 $a(t) = 1$的常系数积分曲线，因为向量场$Y_{\mathbb{R}^3}$ 表示 $3D$ 刺激的切向方向（因此永不为零）：

$$
\dot{\Gamma}(t) = \vec{Y}_{\mathbb{R}^3,\Gamma(t)} + c_1\vec{Y}_{\theta,\Gamma(t)} + c_2\vec{Y}_{\varphi,\Gamma(t)}, \tag{36}
$$

其中 c_1 和 c_2 在 $\mathbb{R}$中变化。

这些曲线可以看作 $\mathbb{R}^3$ 中的轨迹，描述沿 $\vec{Y}_{\mathbb{R}^3}$ 方向的运动；运动方向最终可能随 Y_θ 和 Y_φ而改变。积分曲线族的一个示例已在引言的图1.4中展示。

值得注意的是，当系数 c_1 和 c_2 均为零时，$3D$轨迹将是 $\mathbb{R}^3$中的直线；通过改变 $\mathbb{R}$中的系数 c_1 和 c_2，我们允许积分曲线沿弯曲的轨迹运动，在所有空间方向上扭转和弯折。

形式上，空间中的“扭转与弯曲”程度可通过引入曲率和挠率的概念来度量。接下来，我们将探究这些度量如何编码在积分曲线族的参数中，以及要获得不同类型的曲线，必须施加哪些约束。

注 3.4. 积分曲线(36) 的 $3D$ 投影记作 γ，并满足 $\dot{\gamma}(t) = (\cos\theta(t)\sin\varphi(t),\ \sin\theta(t)\sin\varphi(t),\ \cos\varphi(t))^T$。利用微分几何中的经典工具，我们可以计算曲线 $\gamma(t)$的曲率和挠率:

$$
\begin{aligned}
k &= \sqrt{(\dot{\varphi})^2 + \sin^2\theta(\dot{\theta})^2}, \\
\tau &= \frac{1}{k^2}(-\cos\varphi\sin^2\varphi(\dot{\theta})^3 - \sin\varphi\dot{\varphi}\ddot{\theta} + \dot{\theta}(-2\cos\varphi(\dot{\varphi})^2 + \sin\varphi\ddot{\varphi})).
\end{aligned} \tag{37}
$$

利用方程(36)中向量场 Y_θ 和 Y_φ 的显式表达式，可得

$$
\dot{\theta} = -\frac{c_1}{\sin\varphi}, \qquad \dot{\varphi} = c_2, \tag{38}
$$

由此可得:

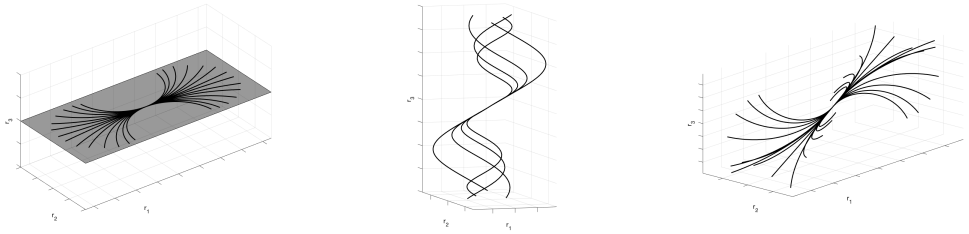
$$
\begin{aligned}
k &= \sqrt{c_1^2 + c_2^2} \\
\tau &= \frac{c_1^2 - c_2^2}{k^2} c_1 \operatorname{cotan}\varphi.
\end{aligned} \tag{39}
$$

[6]有时更倾向于使用术语水平的。

命题 3.2. 通过改变(39) 中的参数 c_1 和 c_2，其中我们显式求得了(36)的解，我们有：

1. 如果 $\varphi = \frac{\psi}{2}$，那么 $k = \sqrt{c_1^2}$、$\tau = 0$，因此曲线族 (36) 是在正面平行平面 $r_3 =$*cost* 上半径为 $1/c_1^2$ 的圆。2. 如果 $\varphi = \varphi_0$，且 $\varphi_0 \neq \pi/2$，那么 $k = \sqrt{c_1^2}$ 和 $\tau = c_1 \operatorname{cotan} \varphi_0$，因此曲线族$(36)$ 是 r_3-螺旋线。3. 如果 $\theta = \theta_0$，那么 $k = \sqrt{c_2^2}$、$\tau = 0$，因此曲线族 (36) 是在密切平面内半径为 $1/c_2^2$ 的圆。4. 如果 $c_1 = \pm c_2$，那么 $\tau = 0$，因此曲线族 (36) 是在密切平面内半径为 $1/c_2^2$ 的圆。

证明。 根据式 **(39)** 计算出的曲率和挠率以及微分几何中的经典结果，计算可立即得出。□



(A) (B) (C)

图 3.5. 改变参数 c_1 和c_2得到的积分曲线示例。**(A)** 当 $\varphi = \pi/2$ 时的圆弧。**(B)** 当 $\varphi = \pi/3$ 时的 r_3-螺旋线。**(C)** 曲率 k 恒定、挠率参数可变的曲线族。

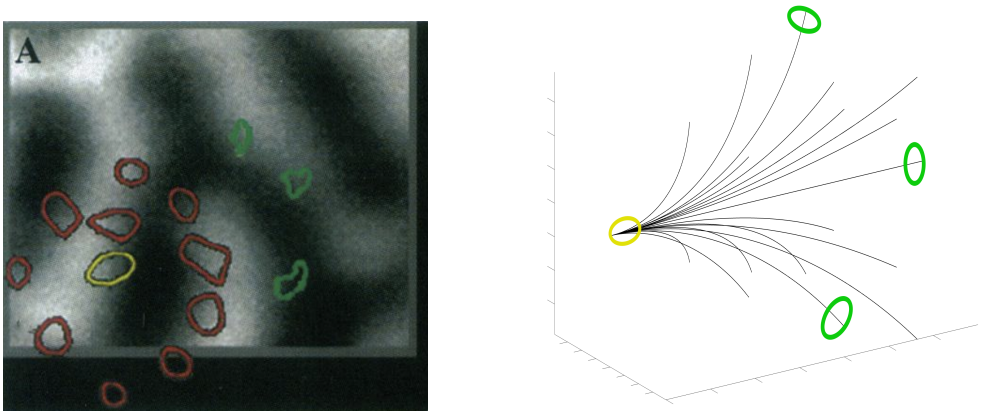
注 3.5. 如果已知曲率 k 的值，并且积分曲线 (36) 的定义中有一个自由参数 c_2，那么我们就处于命题 *2.1* 的情形。事实上，系数 c_1 是通过施加 $c_1 = \pm\sqrt{k^2 - c_2^2}$ 得到的，特别是，剩下待确定的分量就是挠率。

根据命题 3.2 和注 3.5，式 **(36)** 中积分曲线的特例示于图 3.5。

4. 与实验数据的比较

本节将介绍所提出的次黎曼模型与文献中已有的生物学现象和心理物理学现象之间的兼容性结果。

4.1. 生物学联系。 构建立体视觉次黎曼模型的基础，是一个用于曲线延续的次黎曼模型。该模型的灵感来自每个位置上的方向柱，以及相邻柱中细胞之间的联络。反过来，这些联络直接模拟了视觉皮层中的长程水平连接，而关于这些连接已有丰富而出色的生物学数据（例如 [12]）。我们还在《引言》中阐述了支持双眼处理的皮层结构特征。尽管来自每只眼睛的输入按眼优势带组织，但目前没有直接证据表明存在类似于单眼定向柱的“立体视觉柱”。不过，正如我们接下来将要展示的，有证据表明双眼细胞之间存在长程连接，而我们的模型具体说明了这些长程连接应传递哪些信息。因此，立体视觉组织的存在有其依据，但它隐含于皮层结构之中。尽管如此，仍有证据支持这一观点。



(A) (B)

图4.1。（A）将生物胞素注射结果叠加在眼优势柱分布图上，图像取自 [59] 的研究。双眼区位于单眼区（以黑白编码）的中间。从注射部位（双眼区中央的黄色圆圈）开始，斑块的扩散（红色表示密集标记，绿色表示稀疏标记）倾向于避开高度单眼化的部位，绕过眼优势柱的中心，并分布在双眼区。（B） $3D$ 对生理图像（A）的解释。

正如信息传播能够促成单眼曲线延续一样，双眼信号也会传播，从而形成连贯的双眼表征。Grinwald 团队通过立体视觉实验确立了这一点 [59] （另见图4.1(A)）：他们使用生物胞素注射，示踪物会沿神经元突起直接传播，并沉积在兴奋性突触处。因此，这项技术证实了双眼细胞之间存在长程连接。最近，Fitzpatrick 团队 [83], 利用活体钙成像进一步完善了这些结果。如图 1.1（右）所示，作者展示了立体视觉的单眼和双眼输入，并且（未展示）还揭示了其对方向的依赖性。

更准确地说， [59]显示，双眼区域之间存在选择性的各向异性 连接：生物胞素示踪剂并非均匀扩散，而是随着与注射点距离的增加呈现出很强的方向性。（单眼生物胞素注射的情况也是如此。）结合 [83], ，我们将这种各向异性解释为与（双眼）方向（[83]）有关，这恰好是我们向量场积分曲线的行为。我们的 $3D$ 关联场具有很强的方向性，信息优先沿着曲线（起点的）方向传播。图4.1的图像（B）展示了一个例子：其中叠加了彩色斑块，并呈现了积分曲线族（36），遵循 [59]提出的实验方案。

4.2. **心理物理学与关联场**。 在本节中，我们将说明，模型中由积分曲线描述的联络，可以与针对 $\mathbb{R}^3$ 中定向元素知觉组织所开展的心理物理实验得出的几何关系联系起来。我们的目标是确立：这些联络是对 $3D$ 中关联场概念的推广。

4.2.1. 迈向三维轮廓的关联场概念。两个位置—方向元素之间连续性的知觉在 $\mathbb{R}^3$ 中已得到实验研究。首先，Kellman、Garrigan 和 Shipley（[48, 49]）提出了三维可关联性，作为将 Field、Heyes 和 Hess（[30]）在 $2D$中的实验扩展至 $3D$ 的一种方式。

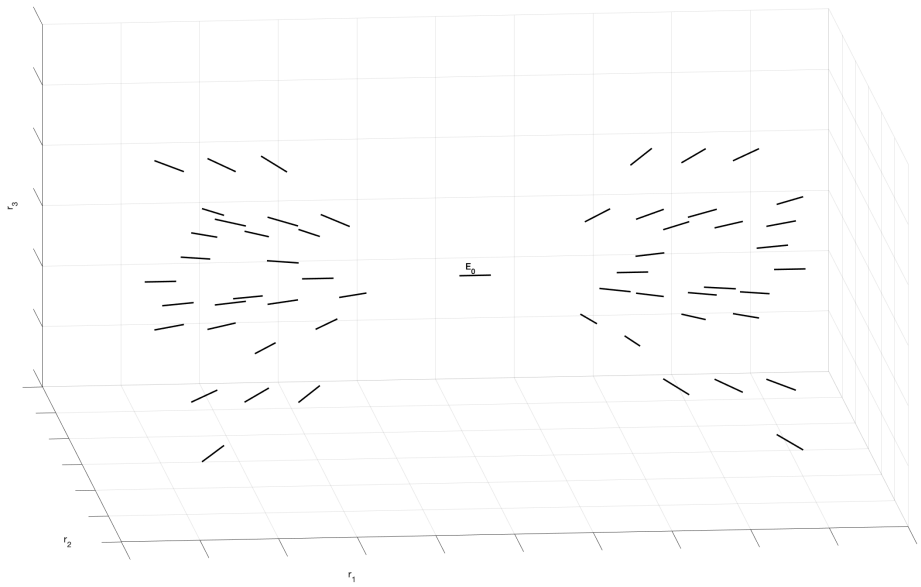


图 4.2. 以 E_0 为初始点的 $3D$ 可关联边扇形示例

特别是在 $3D$ 笛卡尔坐标系中，可以在作用点（r_1，r_2，r_3）^{T}处引入有向边 E，其方向由角度 θ 和 φ确定。在我们的情形中，该方向可通过（cos θsin φ，sin θsin φ，cos φ）^{T}所表示的方向读取。对于一条初始边 E_0，其作用点位于坐标系原点（0，0，0）^{T}，且方向位于 r_1 轴上，由 $\theta = 0$，$\varphi = \pi/2$描述；对于具有 E_0 的 $3D$—可关联边，其可能方向（θ，φ）[7] 范围为：

$$\tan^{-1}\left(\frac{r_2}{r_1}\right) \leq \theta \leq \frac{\pi}{2} \quad \text{and} \quad \frac{\pi}{2} \leq \frac{3\pi}{2} - \varphi \leq \tan^{-1}\left(\frac{r_3}{r_1}\right). \tag{40}$$

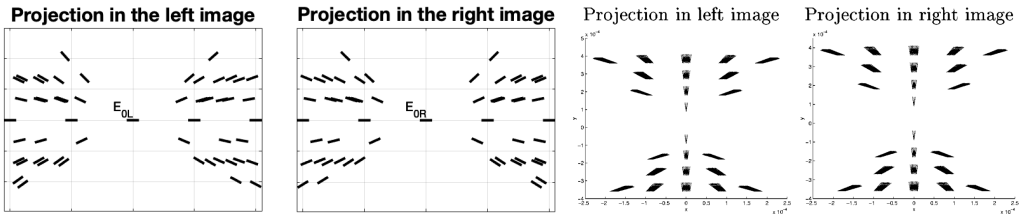
这些方程的界限以量 $\frac{\pi}{2}$ 表示，其中包含三维中的 90度约束；而由反正切定义的界限，则表示参考边 E_0 与位于任意有向点$E_{(r_1,r_2,r_3)}$ 处的一条边之间的绝对方向差异，使得该边的线性延伸与 E_0 相交；更多详情请参见 [48, 49] 。

数值模拟使我们能够直观地呈现满足 $3D$ 可关联性标准的 $3D$ 位置和方向示例。从一条初始边 E_0 开始，其端点位于 $(p_{01}, p_{02}, p_{03})^T$，方向沿 e_1 轴；对于任意点 $(p_1, p_2, p_3)^T$，我们表示可关联方向的极限 (θ, φ)。结果见图 4.2。

注 4.1. 将一族可关联点 $3D$ 投影到视网膜平面上，可以发现这些投影与 *in*[3] 的兼容场概念相符。见图 *4.3*。

心理物理学研究（见 [25, 32, 34], ）考察了适合连接这些可关联点的曲线特性。这些曲线可以很好地描述为平滑且单调。具体而言，Hess 等人（ [34] ）使用非定向轮廓元素构成轮廓，指出轮廓元素主要可以依据其在深度上的良好

[7]这里的角度 φ 已经过调整，以便与我们的坐标集相兼容。文献 [48, 49] 中角度 $\tilde{\varphi}$ 的关系可表示为：$\tilde{\varphi} = \text{acos}(\sin\varphi) + \pi$。



(A) (B)

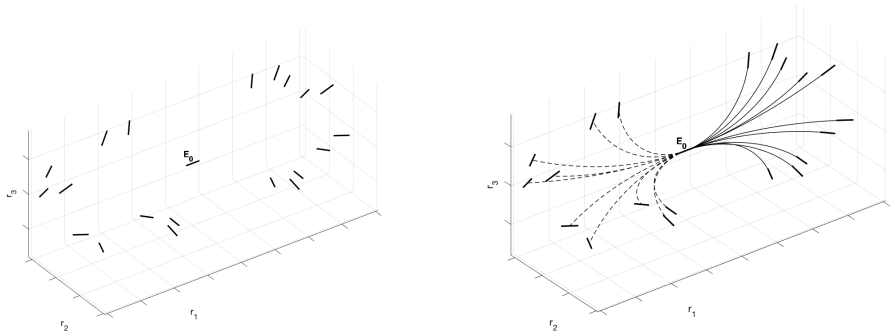
图 4.3。（A）在左右两个视网膜平面中，由 $3D$ 可关联性的几何生成的 $3D$ 关联场示例。（B） [3]的 $3D$兼容性场示例。

延续性有效地分组。Deas 和 Wilcox（[25]）较新的研究证实了这一观点；他们还发现，检测由规则深度连续性定义的轮廓，比检测不连续轮廓更快。所有这些结果都支持深度分组操作的存在，并主张将格式塔连续性和平滑性原理扩展到三维空间。最后，关于三维曲线与二维关联场之间的关系，参见 [49, 51]。这些作者认为， E_0 共面中的可关联边强度必须符合 [30]二维关联场的关系。

4.2.2. 与次黎曼模型的兼容性。为了对上一段所述知觉组织背后的关联进行建模，我们再次考虑（36）中研究的常系数族积分曲线：

$$\dot{\Gamma}(t) = \vec{Y}_{\mathbb{R}^3,\Gamma(t)} + c_1\vec{Y}_{\theta,\Gamma(t)} + c_2\vec{Y}_{\varphi,\Gamma(t)}, \text{ with } c_1, c_2 \in \mathbb{R}. \tag{41}$$

重要的是，这些曲线在局部连接了由 $3D$可关联性的几何生成的关联扇。特别是，图 4.4 的图像（B）显示了连接初始点 E_0 与 $3D$ 可关联边的水平曲线族。这些曲线使用 Matlab 求解器函数 ode45 计算得到。类比于 Field、Hayes 和 Hess 在 [30], 中的实验，我们



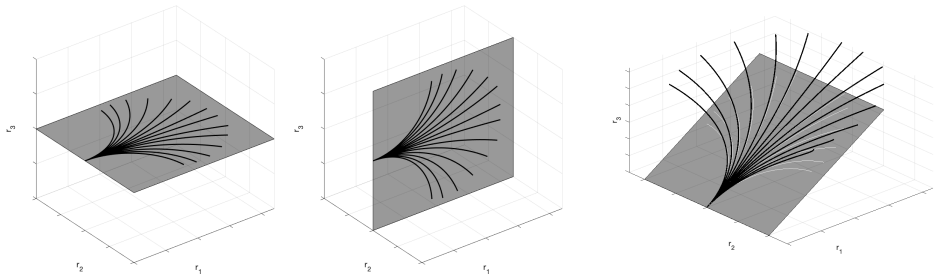
(A) (B)

图 4.4。（A）$3D$ 可关联边显示在初始边 E_0的右侧。不可关联的 $3D$ 边显示在左侧。（B）实线表示的水平积分曲线连接初始点为 E_0的 $3D$ 可关联边。点线表示的水平积分曲线不连接 $3D$ 不可关联边。

选择将不可关联边表示在起始点 E_0的左侧，而右侧

则是可关联边。因此，积分曲线的实线表示中央水平元素 E_0 与其右侧元素之间的相关性，而虚线则将起始点 E_0 与和它不相关的元素连接起来，如图像左侧所示。

将曲线限制在具有任意边缘 E 的共面邻域内，可得到不同的情形。首先，在 r_1-r_2 平面（正面平行）和 r_1-r_3 平面上，曲线为圆弧，这一点已由命题 3.2 证明。此外，对于包含边缘 E 的 $\mathbb{R}^3$ 中任意平面，我们观察到，以固定角度 φ 生成的曲线是螺旋线，并且在局部满足该平面内的二维约束。示例见图 4.5。特别是，图 4.5 的图像 (A) 和 (B) 中所示的曲线与 Citti-Sarti 模型的曲线（见图 2.3）非常吻合。



(A) (B) (C)

图 4.5。(A) 将积分曲线族限制在 e_1-e_2 平面上。(B) 将积分曲线族限制在 e_1-e_3 平面上。(C) 将曲线族限制在 $\varphi = \varphi_0$ 处。这些曲线（黑线）不是平面曲线，而是螺旋线。不过，它们在与初始边缘共面的平面上的投影（白线）满足二维约束。

4.3. 轮廓整合与立体对应问题。 虽然本文的目标并非解决立体对应问题，但我们可以说明，我们的几何方法如何有助于理解如何匹配左、右点及特征。这些想法将在 [11]中得到更充分的阐述。

受 Hess 和 Field 在 [32], 中的实验启发，我们考虑一种路径刺激 γ，将其视为嵌入随机定向元素背景中的轮廓：左、右视网膜视觉刺激如图 4.6 所示。我们首先将视网膜图像简化地提升到 $\mathbb{R}^3 \rtimes \mathbb{S}^2$的一个子集 Ω。该集合包含所有可能的对应点，这些点通过配对具有相同 y 视网膜坐标的左、右点得到，见图 4.7 的图像 (A)。集合 Ω 中包含错误匹配，即不属于原始刺激的点。对应过程的任务就是消除这些错误匹配。

我们计算每个提升点的方程 (17) 双眼输出 O_B 。这一输出可视为一种概率测度，用于提供一对左、右点之间对应关系的信息。我们只需评估哪些点最有可能相互对应，并沿视差纤维对非最大配对进行抑制。这样就能去除噪声点（图 4.7，图像 (B)）。

现在，我们直接利用深度上的良好延续。剩余的噪声元素与我们想要重建的曲线各元素的方向正交。通过数值计算连接所有剩余点对的积分曲线 (36) 的系数 c_1 和 c_2，我们可以利用 (39) 求出每一对点的曲率和挠率。图 4.8 可从矩阵的角度来解读：矩阵 M 表示每对点 ξ_i, ξ_j 在元素 M_{ij}中的曲率或挠率值。特别是，我们观察到随机点的曲率非常高，而且一般而言，挠率也偏离了最小值。

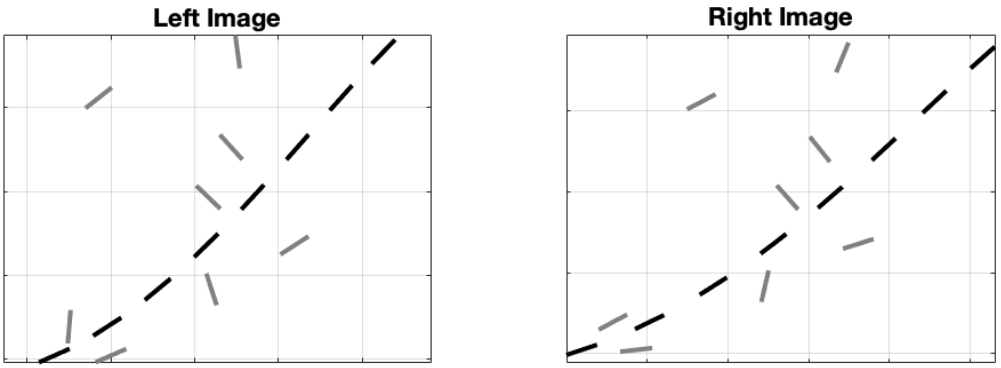
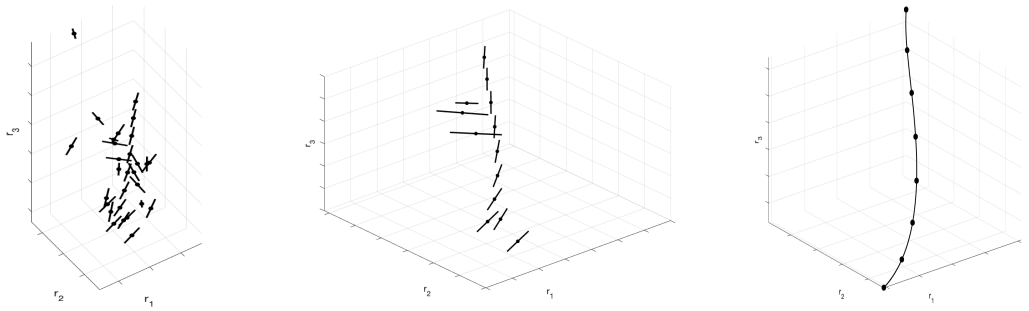


图 4.6。集合 Ω 的左、右视网膜图像。黑点是曲线 γ 上点的投影，灰点则是背景随机噪声。



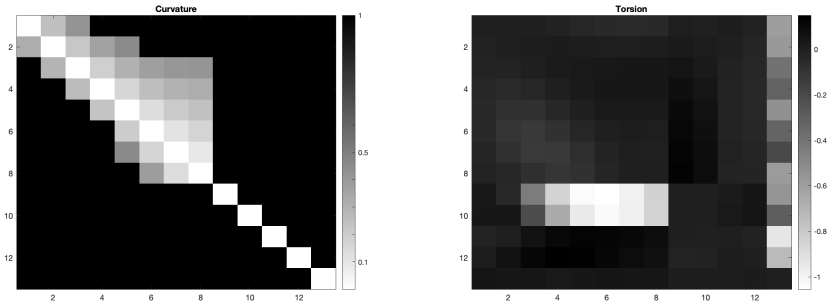
(A)　　(B)　　(C)

图 4.7。（A）将图 4.6 的左右视网膜图像提升到位置与方向 $\mathbb{R}^3 \times \mathbb{S}^2$ 空间中。（B）根据双眼输出筛选提升后的点。（C）刺激 γ 的点由积分曲线 (36) 连接。

因此，舍弃这些高值后，我们只选取曲线 γ上的三维点；如图 4.7 的图像 (C) 所示，这些点通过积分曲线良好连接。这与 [3, 57, 58], 中提出的观点一致：曲率和挠率为三维重建提供约束。

总结与结论

理解深度上的良好延续，与理解平面轮廓的良好延续一样，可以借助基本生理约束、心理物理学表现测量以及数学建模。具体而言，平面上的良好延续受到方向选择性和皮层结构（方向柱）、关联场分组表现以及几何建模的支持。我们认为，深度上的良好延续也应如此。然而，尽管心理物理学数据可能相当，生理学数据却较为薄弱，延续几何也尚未得到充分理解。在本文中，我们提出了立体视觉的神经几何，以填补这一空白。它主要受到将 $2D$ 几何类比拓展至 $3D$ 的启发，同时遵循心理物理学的约束。最终，这使我们能够明确指出，对于理解立体视觉而言，相关的几何类型



(A) (B)

图 4.8. 矩阵 M 其元素 M_{ij} 表示每一对点的曲率/挠率值 ξ_i, ξ_j。前八个点对应曲线上的点 γ，其余点为随机噪声。（A）曲率矩阵。（B）挠率矩阵。

摘要l 极为重要，且具体而言，对生理学研究很有启发。尽管一种“立体 o
柱a 这种架构从解剖结构上并不明显，但在计算上是完善的。

双眼细胞的神经几何通过双眼感受野轮廓来描述，这些轮廓是左、右单眼感受野的乘积。从双眼感受野轮廓出发，只需利用在视网膜平面中恢复的视觉刺激的位置与方向（假设已知对应点），即可重建三维空间。

具体而言，我们提出了一个定义在位置与方向空间上的次黎曼模型$\mathbb{R}^3 \rtimes \mathbb{S}^2$，用于描述相关神经细胞的知觉空间。这种几何结构有利于沿着 $3D$ 曲线刺激的切向方向。次黎曼结构的积分曲线通过其系数编码曲率和挠率的概念，并用于描述各元素之间的联络。该模型可视为二维关联场在三维场景中的扩展。具体来说，位置—方向 $3D$ 空间的次黎曼结构的积分曲线，恰好是在局部与心理物理关联场相对应的曲线。

虽然本文的目标并非解决立体对应问题，但我们已经看到，我们提出的几何为理解如何匹配左、右点及特征提供了良好的起点。模型未来的发展方向之一，是定义两个元素共同出现的概率，从而在 $3D$ 空间中识别知觉。过去已有研究提出通过对次黎曼结构进行调和分析来识别知觉，研究对象既包括 $2D$ 空间刺激 [77]，也包括 $2D+$ 时空刺激 [6]。若能开展类似的分析并将其拓展至立体视觉，将会很有意义。

致谢

MVB、GC 和 AS 得到了欧盟项目 GHAIA（具有跨学科应用的几何与调和分析，H2020-MSCA-RISE-2017）的资助。SWZ 部分得到了美国国立卫生研究院 EY031059 和美国国家科学基金会 CRCNS 1822598 的资助。

附录

A. 次黎曼几何浅介

本文采用了微分几何，尤其是次黎曼几何中的技术。本附录将以较为非正式的讨论，引领读者了解这些思想。对于希望正式了解微分几何基本工具（A.1 和 A.2 节所述内容）的读者，请参阅 [85]。若要全面、正式地学习次黎曼几何的数学知识，请参阅 [2], ；若想从更具信息性的角度了解，请参阅 [79, 第 4 章2] 和 [17]。

A.1. 切丛。 首先，设想你正站在世界某处一片光滑曲面上的一个点，离任何边界都很远。现在，你可以朝任意（$2D$）罗盘方向从这个点“走开”；例如，你可以向北走、向南走，或朝两者之间的任意方向走。如果你的步子非常非常短，那么这个（平面的）罗盘实际上刻画了你可能迈出的步子的 $2D$ 空间。这些相同的思想在微分几何中有更形式化的表述，如下所示。可以将一个切空间附加到可微流形 M （一种广义曲面）的每个点 p 上 T_pM （即罗盘加上一些描述向量运算的代数结构）。也就是说，切空间是一个实向量空间，其中包含可以沿切向穿过 $p \in M$的可能方向。如果流形是连通的，那么每个点处的切空间维数都与流形的维数相同。因此，如果流形是一个 $2D$ 曲面，那么某点处的切空间就是一个平面。一般而言，这个切平面只能在局部“近似”该曲面。

这些元素 $\vec{X}_p$ 在 p 处的切空间 T_pM 中的元素称为切向量（位于 p）。如上所述，这些切向量附着在曲面上的某一点，定义了人们可以从该点出发步行的方向。不过，现代微分几何提供了另一种解释：可以用方向导数来理解切空间中的元素。具体来说，对于每个光滑函数 f， $Xf(p) = \vec{X}_p \cdot \nabla f(p)$表示 f 沿向量 $\vec{X}_p$方向的方向导数，其中 ∇ 表示梯度向量（用适当的坐标系表示），而 $\cdot$ 表示这两个向量的标量积。我们还将记作$X_p = X_p \cdot \nabla_p$，省略函数 f。

现在我们来考虑一对方向导数 X 和 Y。如果 X 和 Y是偏导数，那么对于每个正则函数 f，都有 $XYf = YXf.$。如果 X和 Y是方向导数，一般来说， $XYf \neq YXf.$。通过显式计算可知，在每一点 p

$$[X,Y]f(p) = (XY - YX)f(p) = (J_{\vec{Y}_p}\vec{X}_p - J_{\vec{X}_p}\vec{Y}_p) \cdot \nabla f(p), \tag{42}$$

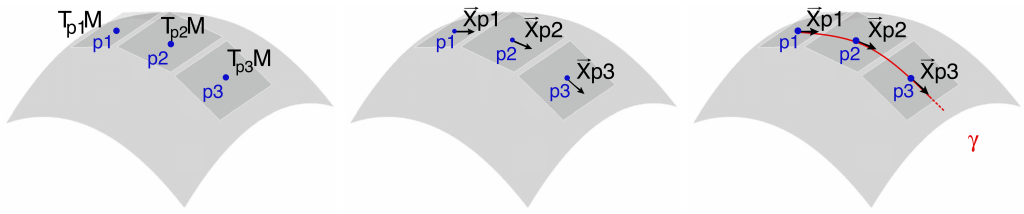
其中 $J_{\vec{X}_p}$和 $J_{\vec{Y}_p}$是 $\vec{X}_p$和 $\vec{Y}_p$的雅可比矩阵。量 $[X,Y]f$称为交换子，因为它表达了两个导数不可交换这一事实。这个概念也可以用增量来表示：可以把从点 p 出发的增量想象成在点 p处作用的向量 X_p的首端，其中 $\vec{}$是该向量。然后，表达式 $XY - YX$可以按以下几何方式得到：先从某一点出发沿 X 移动，再从其首端沿另一个量 Y移动，接着沿第一个量反向移动，最后沿第二个量反向移动。问题在于这个四边形是否闭合。形式上，这由点 p处两个元素 X 和 Y 的交换子 来刻画。

为了计算二阶导数 XYf，我们需要知道 p附近每一点处的 Yf。这引出了向量场这一更一般的概念，它抽象描述了流形上运动的点的速度场。向量场 X 以平滑的方式，为流形 M 上的每个点 p 关联一个来自该点切空间的向量 X_p，即 $\vec{}$。点与点之间不会出现突变。

既然我们已经将每个切向量与一个导子联系起来，现在就可以进一步探讨；见图 A.1。每个向量场都可以与一个常微分方程相关联，其解称为该向量场的积分曲线：它们是参数曲线，表示由该向量场描述的常微分方程的特定解。可以这样理解：想象你从某一点出发，沿着

该点处的切向量方向迈出一个无穷小的步子；此时你就来到了一个邻近点。然后，你可以从这个邻近点出发，沿着（可能不同的）切向方向再迈一步。这样持续一段时间，你便能在几何上描绘出积分曲线 $\gamma\colon\ [t_1,t_2]\subseteq\mathbb{R}\longrightarrow M$。重要的是，给定的向量场 X 在点 $\gamma(t)$ 处就是曲线在该点的切向量。更重要的是，这一点在曲线上的每一点都成立，因此积分曲线满足一个微分方程：

$$\dot{\gamma}(t)=\vec{X}_{\gamma(t)}. \tag{43}$$



(A) (B) (C)

图 A.1。（A）流形 M 中各点 $p_i, i=1,2,3$ 处的切平面 $T_{p_i}M$（较深色的平面）。（B）定义在 M 上的向量场 X：对于流形 M 中的每个点 $p_i, i=1,2,3$，我们都有一个属于该点切空间的向量 $\vec{X}_{p_i}$。（C）从 $p_1\in M$ 出发、与向量场 X 相关联的积分曲线 γ。

流形的所有切空间都可以“粘合”起来，形成一个新的可微流形，其维数是原流形的两倍，称为该流形的切丛。作为一个集合，它由 M 的各个切空间的不交并构成，即：

$$TM=\bigsqcup_{p\in M}T_pM=\{(p,X_p)\mid p\in M,\,X_p\in T_pM\}. \tag{44}$$

特别地，TM 中的一个元素可以看作一对 (p,X_p)，其中 p 是 M 中的一个点，而 X_p 是 M 在 p 处的切向量。存在一个自然投影 $\pi\colon TM\to M$，定义为 $\pi(p,X_p)=p$。它将切空间 T_pM 的每个元素映射到单个点 p。

A.2. 流形上的群作用。 实数的加法具有一种重要的代数结构，称为群。例如，它要求任意两个数的和仍是一个数；存在逆运算“-”；还存在单位运算“0”，也就是说，将任意数与 0 相加，结果仍是该数。

当群 G 作用于一个流形（例如上文中的实数）时，这意味着群中的每个元素都会以与流形自身相容的方式，对流形上的所有元素执行某种运算。更准确地说，这种作用由映射 $\sigma\colon G\times M\to M,\ (g,\ x)\mapsto g\cdot x$ 来描述；如果映射 σ 可微，那么它就是群 G 在光滑流形 M 上的（左）群作用。

例如，我们可以取二维旋转平移群 $SE(2)=\mathbb{R}^2\times\mathbb{S}^1$，并按照群运算律定义它在光滑流形 $M\subseteq\mathbb{R}^2$ 上的作用：先对流形本身进行旋转，再进行平移。这可以通过映射 $\sigma\colon SE(2)\times M\longrightarrow M,\ \sigma(g,\ p)=(Rp+q)$ 形式化表示，其中 $g=(q,\ R)\in SE(2)$，即点 $q\in\mathbb{R}^2$ 以及角度为 $\theta\in\mathbb{S}^1$的二维旋转。图 A.2 展示了一个图形示例。

现在，我们可以将这些熟悉的概念推广到皮层空间及其特殊的位置-方向结构$\times$，或推广到立体空间。

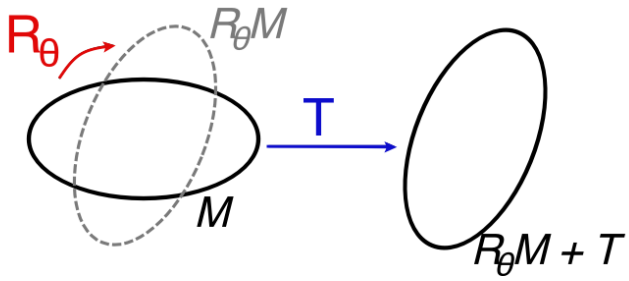
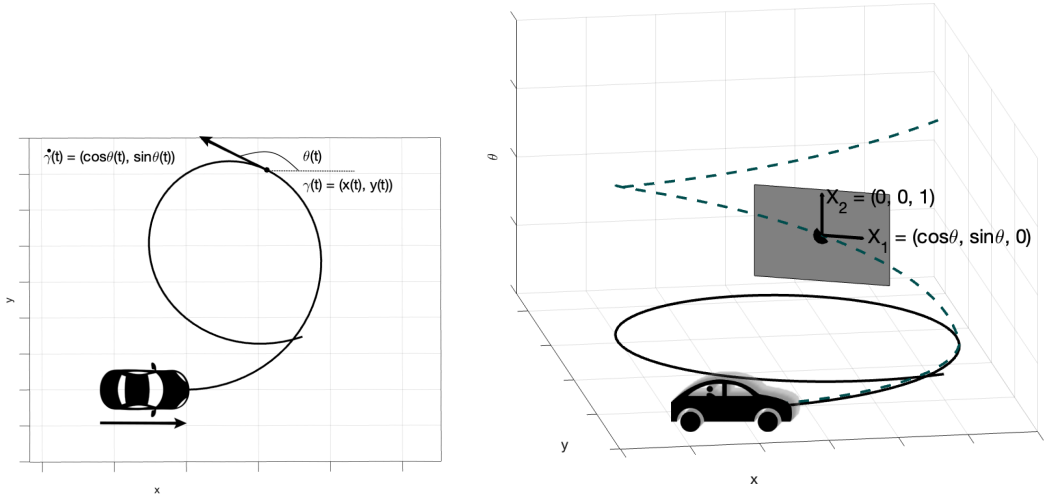


图 A.2. 旋转平移群 $SE(2)$ 在流形M （黑色椭圆）上的群作用：首先，流形旋转角度 θ，得到 $R_\theta M$；然后施加平移，使旋转后的流形在空间中移动，从而实现 $R_\theta M + T$。

A.3. **次黎曼几何**。 如上所示，在流形上移动时受到点约束，这意味着只能沿着与流形相切的方向移动，因为沿法线方向移动会离开流形。因此，对于每个点 p，允许的位移方向集合都与切平面 T_pM重合。若存在更多约束，则某些切向方向可能会被禁止。因此，我们在每个点p引入允许切空间 $\mathcal{A}_p$，它是 T_pM 中允许运动方向所构成的子空间。如果切空间 T_pM 的维数为 n，那么允许切空间 $\mathcal{A}_p$ 的维数将为 $m \leq n$。对流形上的每个点重复同样的构造，我们将各点处允许切空间的并集称为可容许切丛：$\mathcal{A} = \bigsqcup_{p\in M} \mathcal{A}_p$。如果在 $\mathcal{A}_p$上引入标量积，就能定义向量的范数，以测量这些向量的长度以及点与点之间的距离。具有这些性质的流形通常称为次黎曼流形，而允许沿任意方向运动的流形则称为黎曼流形。



(A) (B)

图 A.3. (A) 汽车在平面上移动的几何设置。(B) 在 $SE(2)$中的次黎曼形式化。路径的切向量被限制在灰色平面内，该平面由 $X_{1,\ p}$ 和 $X_{2,\ p}$张成，它们是允许的运动方向。

需要明确指出，黎曼几何是在存在物理约束时产生的，而次黎曼几何则是在存在微分约束时产生的，例如对车辆运动的描述。汽车在二维平面上行驶，但只能沿当前方向移动，或者通过转动方向盘改变当前方向。这些就是容许方向。此外，汽车无法“横向”移动（禁行方向），因此无法在保持初始位置不变的情况下直接到达其他方向，只能将两种容许运动同时组合起来，从而限制了允许的运动方式。因此，车辆所描述的轨迹是一条曲线，其切向量受限于沿着这两个容许方向。这个次黎曼问题的形式化方法是在 $SE(2)$ 中进行的：对于每个 $p \in SE(2)$，将由当前方向 $\vec{X}_{1,\,p} = (\cos\theta,\ \sin\theta,\ 0)^T$ 和旋转方向 $\vec{X}_{2,\,p} = (0,\ 0,\ 1)^T$ 所张成的子空间视为允许切空间 $A_p SE(2)$。见图 A.3。

类似地，我们可以从一个视网膜拓扑（$x,\ y$）位置移动到另一个视网膜拓扑位置（$x',\ y'$），沿着方向柱从 θ 到 θ'“向上”或“向下”移动；但我们无法在保持相同初始位置（沿着同一个方向柱移动）的情况下，从 θ 到达θ'：若要到达“禁行方向”，我们必须同时穿过不同位置和方向。这种移动限制正是欧几里得几何（或黎曼几何）与次黎曼几何的区别所在。

B. 命题 3.1 的证明

在本附录中，我们将展示如何运用微分几何的工具证明命题 3.1，尤其是微分 k–形式的概念。d

B.1. 微分形式。 在一个 n 维光滑流形 M 上，微分 k-形式是任意多重线性函数 ω：$TM^k \longrightarrow \mathbb{R}$，它以 k 个光滑向量场为输入，并输出一个标量，同时满足反对称性：

$$\omega(X_1, \ldots, X_i, \ldots, X_j, \ldots, X_k) = -\omega(X_1, \ldots, X_j, \ldots, X_i, \ldots, X_k),$$

其中 $k \leq n$ 且 $k,\ n \in \mathbb{N}$。

在特殊情况下，当 ω 是 1-形式时，值得注意的是，它是 TM的对偶空间（余切空间）中的一个元素：$\omega \in TM^* \iff \omega\colon TM \longrightarrow \mathbb{R}$。如果我们在 M上有坐标 $(x_1,\ldots,x_n)$，就可以利用 TM^*的对偶基 $\{\mathrm{d}\,x_1,\ldots,\mathrm{d}\,x_n\}$表示这些 1-形式：

$$\omega_p = f_1(\bar{x}_1, \ldots, \bar{x}_n)\,\mathrm{d}\,x_1 + \ldots + f_n(\bar{x}_1, \ldots, \bar{x}_n)\,\mathrm{d}\,x_n, \text{ with } p = (\bar{x}_1, \ldots, \bar{x}_n),$$

其中 f_i 是标量光滑函数。

此外，还可以通过楔积 $\wedge$ 将微分 k-形式 ω与微分 l-形式 η相乘，得到微分 $k+l$-形式 $\omega \wedge \eta$。更准确地说，我们关注 1-形式 ω 和 η的楔积，其计算方式为：$\omega \wedge \eta(X,\ Y) = \omega(X)\eta(Y) - \omega(Y)\eta(X)$，其中 X和 Y是 M上的向量场。

B.2. **证明的推导。**

命题 B.1. 双眼交互项 $O_L O_R$ 可以与通过 (13) 定义的左右方向的叉积相关联，即单眼简单细胞的 $\omega^\star_{p_L}$ 和 $\omega^\star_{p_R}$：

$$O_L O_R = \omega^\star_{p_L} \times \omega^\star_{p_R}. \tag{45}$$

证明。如小节 2.2.1 所述，$SE(2)$ 中简单细胞（11）的输出随后可在局部近似为 $O(x,\ y,\ \theta) = -X_{3,\,p}(I_\sigma)(x,\ y)$，其中 I_σ 是 I 的平滑版本，通过与高斯核进行卷积得到，向量场

$$X_{3,p} = -\sin\theta\partial_x + \cos\theta\partial_y, \tag{46}$$

其中 $p = (x,\ y,\ \theta) \in SE(2)$。转到对偶空间后，简单细胞的作用会在每个细胞上分别诱导出一个 1-形式：

$$\omega_p = -\sin\theta\,\mathrm{d}\,x + \cos\theta\,\mathrm{d}\,y. \tag{47}$$

因此，可以将双眼交互项改写为：

$$O_L O_R = X_{3,p_R}(I_{\sigma R})(x_R, y) X_{3,p_L}(I_{\sigma L})(x_L, y). \tag{48}$$

接下来我们将看到，这种双眼作用可以用一个 2-形式来描述，该 2-形式由两个单眼简单细胞的 1-形式定义。

我们用下标 R 表示与右侧单眼结构相对应的量，并用下标 L 表示左侧单眼结构相对应的量。因此，我们定义 $v_R := (J_{I_{\sigma_R}}\vec{X}_{3,\ p_R})X_{3,\ p_R}$，使用图像 I平滑版本的雅可比矩阵（微分），使得我们有 $\omega_{p_R}(v_R) = X_{3,\ p_R}(I_{\sigma_R}) = (J_{I_{\sigma_R}}\vec{X}_{3,\ p_R})$，因为 $\omega_{p_R}(X_{3,\ p_R}) = 1$且 $J_{I_{\sigma_R}}\vec{X}_{3,\ p_R} \in \mathbb{R}$；同样的推理也适用于左侧结构。于是可以在视网膜坐标中将（48）改写为：

$$\begin{aligned} O_L O_R =& \omega_{p_L}(v_L)\omega_{p_R}(v_R) \\ =& \omega_{p_L} \wedge \omega_{p_R}(v_L, v_R) + \underbrace{\omega_{p_R}(v_L)\omega_{p_L}(v_R)}_{=0}, \\ =& \omega_{p_L} \wedge \omega_{p_R}(v_L, v_R), \end{aligned} \tag{49}$$

利用楔积和左右视网膜坐标的性质。Th

e r视网膜坐标可用独眼坐标 (4) 表示为 $x_R = x -$ d 以及 $x_L = x + d$；那么，扩展后的左右 1-形式可写为：

$$\begin{aligned} \omega_{p_R} &= -\sin\theta_R\,\mathrm{d}\,x + \cos\theta_R\,\mathrm{d}\,y + \sin\theta_R\,\mathrm{d}\,d \\ \omega_{p_L} &= -\sin\theta_L\,\mathrm{d}\,x + \cos\theta_L\,\mathrm{d}\,y - \sin\theta_L\,\mathrm{d}\,d. \end{aligned} \tag{50}$$

利用霍奇星算子在 $\mathbb{R}^3$ 中提供的向量与 2-形式之间的同构，我们用记号 (20) [8] 将外积与叉积联系起来，如下所示：

$$\star\,(\omega_{p_L} \wedge \omega_{p_R}) = \omega^{\star}_{p_L} \times \omega^{\star}_{p_R}, \tag{51}$$

由此即可得出本文的结论。 □

在全文中，为了简化记法，我们将 $\omega_L = \omega_{p_L}$ 和 $\omega_R = \omega_p$ $_R$.

B.2.1. 数学对象的含义。 最后，我们对这一框架中引入并使用的数学工具进行探讨，以了解所提出的数学模型如何从 [18], 开始，将这些不同的数学对象分别对应到物理细胞、其作用以及作用的结果上。该模型由 Citti 和 Sarti 提出。

注释 B.1。 众所周知，奇数简单细胞（对方向具有选择性）会因刺激的出现而被激活，以选择其方向（即感知曲线的切向量）。在这一框架中， [18] 提出的模型背后的数学直觉，是将每个细胞识别为一个1-微分形式，它是余切空间中的一个元素。粗略地说，这个微分形式能够捕捉一个与刺激方向相对应的向量：这就是细胞作用的结果。形式上，该向量是切空间中的一个元素，更确切地说，它属于1-形式的核。随后，这个向量空间与细胞的作用相对应。

在一系列论文中（[78, 6, 1, 8]），同样的推理也应用于不同的细胞族，尽管这些细胞族在不同流形中具有各自不同的次黎曼结构。感兴趣的读者可参阅 [19] 以了解综述。同样，我们在双眼细胞族中也发现了相同的几何组织结构。

[8]使用记号 $\omega^{\star}$，我们将向量的各分量与 1-形式 ω 相对于对偶基的系数对应起来

注释 B.2. 本文讨论了双眼细胞，它们由单眼简单细胞组合而成。对于这些耦合简单细胞（分别对应左眼和右眼），我们形式上关联一个2-微分形式，即左、右单眼1-形式的楔积。这个2-形式同样能够捕捉一个位于该数学对象核中的向量，从而确定三维刺激方向。因此，注释 *B.1* 中的同样推理也适用于双眼细胞族。

将注释 B.1 的结果转换到具有不同维数的不同空间后，便可以使用相同的数学对象来解释不同细胞族群的行为，从几何上确定构成所研究细胞族群功能基础的数学对象。

C. 变量变换

让我们恢复 1-形式 $\tilde{\omega}_L := U_{t_L}$ 和 $\tilde{\omega}_R := U_{t_R}$ 的表达式。这里回顾变量变换（5）：

$$\begin{cases} r_1 = \frac{xc}{d} \\ r_2 = \frac{yc}{d} \\ r_3 = \frac{fc}{d} \end{cases}, \tag{52}$$

及其微分：

$$\begin{cases} \mathrm{d}\, r_1 = \frac{c}{d}\,\mathrm{d}\, x - \frac{cx}{d^2}\,\mathrm{d}\, d \\ \mathrm{d}\, r_2 = \frac{c}{d}\,\mathrm{d}\, y - \frac{cy}{d^2}\,\mathrm{d}\, d \\ \mathrm{d}\, r_3 = -\frac{fc}{d^2}\,\mathrm{d}\, d \end{cases}. \tag{53}$$

将（25）中定义的量 U_{t_L} 写成变量（r_1，r_2，r_3）中的 1-形式，我们得到：

$$\tilde{\omega}_L = -f\sin\theta_L\,\mathrm{d}\, r_1 + f\cos\theta_L\,\mathrm{d}\, r_2 + (x_L\sin\theta_L - y\cos\theta_L)\,\mathrm{d}\, r_3. \tag{54}$$

变换坐标：

$$\begin{aligned} \tilde{\omega}_L =& -f\sin\theta_L\left(\frac{c}{d}\,\mathrm{d}\, x - \frac{cx}{d^2}\,\mathrm{d}\, d\right) + f\cos\theta_L\left(\frac{c}{d}\,\mathrm{d}\, y - \frac{cy}{d^2}\,\mathrm{d}\, d\right) \\ &+ (x_L\sin\theta_L - y\cos\theta_L)\left(-\frac{fc}{d^2}\,\mathrm{d}\, d\right) \\ =& \frac{fc}{d}\left(-\sin\theta_L\,\mathrm{d}\, x + \cos\theta_L\,\mathrm{d}\, y - \sin\theta_L\,\mathrm{d}\, d\right) \\ =& \frac{fc}{d}\omega_L. \end{aligned} \tag{55}$$

因此，除一个标量因子外，我们得到：在变量（x，y，d）中，$\tilde{\omega}_L = \omega_L$。同样的推理也适用于右侧结构。

参考文献

[1] S. Abbasi-Sureshjani、M. Favali、G. Citti、A. Sarti 和 B. M. ter Haar Romeny。用于从视网膜图像中提取血管连接的五维核曲率整合。*IEEE* 图像处理汇刊，27(2):606–621，2017年。[2] A. Agrachev、D. Barilari 和 U. Boscain。次黎曼几何综合导论，第181卷。剑桥大学出版社，2019年。[3] S. Alibhai 和 S. W. Zucker。立体视觉中的基于轮廓的对应。在计算机视觉：*ECCV 2000*中，第314–330页。施普林格柏林海德堡，2000年。[4] A. Anzai、I. Ohzawa 和 R.D. Freeman。双眼视差编码的神经机制：感受野位置与相位。《神经生理学杂志》，82(2):874–890，1999年8月。[5] A. Anzai、I. Ohzawa 和 R.D. Freeman。双眼信息处理的神经机制 I：简单细胞。《神经生理学杂志》，82(2):891–908，1999年8月。[6] D. Barbieri、G. Citti、G. Cocci 和 A. Sarti。用于轮廓感知和运动整合的皮层启发几何。《数学成像与视觉杂志》，49(3):511–529，2014年。

[7] D. Barbieri、G. Citti 和 A. Sarti。 不确定性如何限定简单细胞的形状指数。数学神经科学杂志，4(1):5，2014年。[8] E. Baspinar、A. Sarti 和 G. Citti。包含频率和相位的视觉皮层次黎曼模型。数学神经科学杂志，10(1):1–31，2020年。[9] O. Ben-Shahar 和 S. W. Zucker。几何计算解释了视觉皮层中长程水平连接的投射模式。神经计算，16(3):445–476，2004年。[10] G.G. Blasdel。猴纹状皮层中的方向选择性、偏好与连续性。J. Neurosci.;《神经科学杂志》，12(8):3139–3161，1992年。[11] M. V. Bolelli 等。立体视觉的神经几何。博士论文撰写中，博洛尼亚大学和索邦大学，2023年。[12] W. H. Bosking、Y. Zhang、B. Schofield 和 D. Fitzpatrick。树鼩纹状皮层中的方向选择性及水平连接的排列。J. Neurosci.;《神经科学杂志》，17 2127，1997 6 2112( ): – 。[13] H. Bridge 和 B. G. Cumming。猕猴初级视觉皮层神经元对双眼方向差异的反应。《神经科学杂志》，21(18):7293–7302，2001年。[14] H. Bridge、B. G. Cumming 和 A. J. Parker。对初级视觉皮层神经元方向视差反应的建模。《视觉神经科学》，剑桥大学出版社，18:879–891，2001年。[15] R. Cagenello 和 B. J. Rogers。立体表面知觉中的各向异性：方向视差的作用。视觉研究，33(16):2189–2201，1993年。[16] J. T. Chang、D. Whitney 和 D. Fitzpatrick。经验依赖性重组推动双眼统一的方向皮层表征的发育。《神经元》，107(2):338–350，2020年。[17] G. Citti、L. Grafakos、C. Pérez、A. Sarti 和 X. Zhong。调和分析与几何分析。Springer，2015年。[18] G. Citti 和 A. Sarti。旋转平移空间中知觉补全的皮层模型。《数学成像与视觉杂志》，24(3):307–326，2006年2月。[19] G. Citti 和 A. Sarti。视觉神经数学，第32卷。Springer，2014年。[20] B. G. Cumming 和 G. C. DeAngelis。立体视觉的生理学。《神经科学年度评论》，24(1):203–238，2001年3月。[21] J. G. Daugman。二维视觉皮层滤波器对空间、空间频率和方向分辨率进行优化时的不确定性关系。《美国光学学会杂志A辑》，2(7):1160，1985年7月。[22] G. C. DeAngelis、B. G. Cumming 和 W. T. Newsome。MT皮层区与立体深度知觉。《自然》，394(6694):677–680，1998年。[23] L. M. Deas 和 L. M. Wilcox。通过闭合实现的格式塔分组会削弱阈上深度知觉。《视觉杂志》，14(9):14–14，2014年。[24] L. M. Deas 和 L. M. Wilcox。通过双眼视差实现知觉分组：立体良好延续的影响。《视觉杂志》，15(11):11–11，2015年。[25] L. M. Deas 和 L. M. Wilcox。通过双眼视差实现知觉分组：立体良好延续的影响。《视觉杂志》，15(11):11，2015年8月。[26] R. Duits、U. Boscain、F. Rossi 和 Y. Sachkov。通过 SE(2) 中无尖点次黎曼测地线构建关联场。《数学成像与视觉杂志》，49(2):384–417，2013年12月。[27] R. Duits 和 E. Franken。位置与方向空间上的左不变扩散及其在HARDI图像保交叉平滑中的应用。《国际计算机视觉杂志》，2011年。[28] J. H. Elder 和 R. M. Goldberg。用于轮廓知觉组织的格式塔法则的生态统计。《视觉杂志》，2(4):5–5，2002年。[29] O. Faugeras。三维计算机视觉：几何视角。麻省理工学院出版社，1993年。[30] D. J. Field、A. Hayes 和 R. F. Hess。人类视觉系统的轮廓整合：局部“关联场”的证据。视觉研究，33(2):173–193，1993年。[31] W. S. Geisler、J. S. Perry、B. J. Super 和 D. P. Gallogly。自然图像中的边缘共现可预测轮廓分组表现。视觉研究，41(6):711–724，2001年。[32] R. F. Hess 和 D. J. Field。跨深度轮廓整合。视觉研究，35(12):1699–1711，1995年。[33] R. F. Hess、A. Hayes 和 D. J. Field。轮廓整合与皮层处理。《巴黎生理学杂志》，97(2-3):105–119，2003年。[34] R. F. Hess、A. Hayes 和 F. A. A. Kingdom。深度内及跨深度的轮廓整合。视觉研究，37(6):691–696，1997年。[35] R. F. Hess、K. A. May 和 S. O. Dumoulin。轮廓整合：心理物理学、神经生理学与计算视角。《牛津知觉组织手册》，2014年。[36] D. A. Hinkle 和 C. E. Connor。猕猴V4区的三维方向调谐。《自然神经科学》，5(7):665–670，2002年。[37] W. C. Hoffman。视觉皮层是一个接触丛。《应用数学与计算》，32 167，1989 2 137( ): – 。[38] I. P. Howard。深度知觉，第1卷：基本机制。牛津大学出版社，2012年。[39] I. P. Howard 和 B. J. Rogers。双眼视觉与立体视觉。牛津大学出版社（美国），1995年。

[40] D. H. Hubel 和 T. N. Wiesel。猫视觉皮层中的感受野、双眼交互作用与功能架构。《生理学杂志》，160(1):106–154，1962年1月。[41] D. H. Hubel 和 T. N. Wiesel。猕猴的立体视觉：猕猴皮层18区中对双眼深度敏感的细胞。《自然》，225(5227):41–42，1970年。[42] R. Ivry、J. Beck 和 A. Rosenfeld。线条分离。《空间视觉》，4(2-3):75–101，1989年。[43] D Jaeger 和 J. Ranu。《计算神经科学百科全书》。Springer New York，2015年。[44] D. G. Jones 和 J. Malik。利用对应线元素，根据方向和空间频率视差确定三维形状。技术报告 UCB/CSD-91-656，电气工程与计算机科学系，加州大学伯克利分校，1991年10月。[45] J. P. Jones 和 L. A. Palmer。对猫纹状皮层简单感受野的二维 Gabor 滤波器模型的评估。《神经生理学杂志》，58(6):1233–1258，1987年12月。[46] B. Julesz。《独眼巨人式知觉的基础》。芝加哥：芝加哥大学出版社，1971年。[47] D. Kato、M. Baba、K. S. Sasaki 和 I. Ohzawa。广义池化对早期视觉皮层神经元双眼视差选择性的影响。《皇家学会哲学汇刊B辑：生物科学》，371(1697):20150266，2016年6月。[48] P. J. Kellman、P. Garrigan 和 T. F. Shipley。三维物体插值。《心理学评论》，112(3):586–609，2005年。[49] P. J. Kellman、P. Garrigan、T. F. Shipley、C. Yin 和 L. Machado。物体知觉中的三维插值：来自客观表现范式的证据。《实验心理学杂志：人类知觉与表现》，31(3):558–583，2005年。[50] S. K. Khuu、V. Honson 和 J. Kim。三维轮廓的知觉，以及亮度极性和颜色变化对其检测的影响。《视觉杂志》，16(3):31–31，2016年。[51] S. K. Khuu、V. Honson 和 J. Kim。三维轮廓的知觉，以及亮度极性和颜色变化对其检测的影响。《视觉杂志》，16(3):31，2016年2月。[52] J. J. Koenderink 和 A. J. van Doorn。视觉系统中局部几何的表征。《生物控制论》，55(6):367–375，1987年3月。[53] K. Koffka。《格式塔心理学原理》。美国纽约州纽约市：Harbinger丛书，1963年。[54] W. Kohler。格式塔心理学。《心理学研究》，31(1):XVIII–XXX，1967年。[55] M. Lawlor 和 S. W. Zucker。三阶边缘统计：轮廓延续、曲率与皮层连接。《神经信息处理系统进展》，26，2013年。[56] S. LeVay、D. H. Hubel 和 T. N. Wiesel。通过还原银染色揭示的猕猴视觉皮层眼优势柱模式。《比较神经学杂志》，159(4):559–575，1975年。[57] G. Li 和 S. W. Zucker。基于轮廓的立体对应的微分几何模型。载于《IEEE计算机视觉变分、几何与水平集方法研讨会论文集》，法国尼斯，2003年。[58] G. Li 和 S. W. Zucker。基于轮廓的立体对应中的上下文推断。《国际计算机视觉杂志》，69(1):59–75，2006年。[59] R. Malach、Y. Amir、M. Harel 和 A. Grinvald。通过光学成像和灵长类纹状皮层活体靶向生物胞素注射揭示内在连接与功能架构之间的关系。《美国国家科学院院刊》，90(22):10469–10473，1993年。[60] D. Marr 和 T. Poggio。人类立体视觉的计算理论。《伦敦皇家学会会刊B辑：生物科学》，204(1156):301–328，1979年5月。[61] J. H. Maunsell 和 D. C. Van Essen。猕猴中颞视觉区神经元的功能特性。II. 双眼交互作用与双眼视差敏感性。《神经生理学杂志》，49(5):1148–1167，1983年。[62] M. D. Menz 和 R. D. Freeman。视觉皮层中视差调谐神经元的功能连接。《神经生理学杂志》，91(4):1794–1807，2004年4月。[63] G. J. Mitchison 和 S. P. McKee。立体倾斜知觉各向异性背后的机制。《视觉研究》，30(11):1781–1791，1990年。[64] P. Neilson、M. Neilson 和 R. Bye。三维双眼视觉知觉的黎曼几何理论。《视觉》，2(4):43，2018年12月。[65] J. I. Nelson、H. Kato 和 P. O. Bishop。猫纹状皮层中双眼激活神经元对方向视差和位置视差的辨别。《神经生理学杂志》，40(2):260–283，1977年3月。[66] J.I. Nelson、H. Kato 和 P. O. Bishop。猫纹状皮层中双眼激活神经元对方向视差和位置视差的辨别。《神经生理学杂志》，40(2):260–283，1977年。[67] P. Parent 和 S. W. Zucker。轨迹推断、曲率一致性与曲线检测。《IEEE模式分析与机器智能汇刊》，11(8):823–839，1989年。[68] A. J. Parker、J. E. T. Smith 和 K. Krug。立体视觉的神经架构。《皇家学会哲学汇刊B辑：生物科学》，371(1697):20150261，2016年6月。[69] J. Petitot。《视觉的神经几何学：功能架构的数学与物理模型》。巴黎综合理工学院出版社，2008年。[70] J. Petitot 和 Y. Tondut。《走向神经几何学》。皮层纤维化、接触结构与模态主观轮廓。《数学与人文科学》，145:5–101，1999年。

[71] G. F. Poggio。猴视觉皮层中的立体视觉机制。《大脑皮层》，5(3):193–204，1995年。[72] J. C. Read。立体视觉：模型，第2873–2881页。Springer New York，美国纽约州纽约市，2015年。[73] J. C. Read 和 B. G. Cumming。用于不可能刺激的传感器或可解决立体对应问题。《自然—神经科学》，2007年10月。[74] D. L. Ringach、P. J. Mineault、E. Tring、N. D. Olivas、P. Garcia-Junco-Clemente 和 J. T. Trachtenberg。小鼠初级视觉皮层中调谐的空间聚类。《自然·通讯》，7(1):1–9，2016年。[75] J. M. Samonds、B. R. Potetz、C. W. Tyler 和 T. S. Lee。递归连接可以解释初级视觉皮层中视差处理的动态过程。《神经科学杂志》，33(7):2934–2946，2013年。[76] G. Sanguinetti、G. Citti 和 A. Sarti。旋转平移群中自然图像边缘共现的模型。《视觉杂志》，2010/2011年。[77] A. Sarti 和 G. Citti。初级视觉皮层功能架构中视觉知觉单元的构成。《计算神经科学杂志》，38(2):285–300，2015年。[78] A. Sarti、G. Citti 和 J. Petitot。初级视觉皮层的辛结构。《生物控制论》，98(1):33–48，2007年11月。[79] A. Sarti、G. Citti 和 D. Piotrowski。微分异质生成与符号功能的出现。《符号学》，2019年(230):1–34，2019年。[80] K. S. Sasaki、Y. Tabuchi 和 I. Ohzawa。猫纹状皮层中的复杂细胞在三维双眼感受野中具有多个视差探测器。《神经科学杂志》，30 13837，2010 41 13826( ): – 。[81] K. S. Sasaki、Y. Tabuchi 和 I. Ohzawa。猫纹状皮层中的复杂细胞在三维双眼感受野中具有多个视差探测器。*J. Neurosci.;*《神经科学杂志》，30(41):13826–13837，2010年。[82] K. E. Schmidt、R. Goebel、S. Löwel 和 W. Singer。共线性这一知觉分组标准反映了初级视觉皮层连接的各向异性。《欧洲神经科学杂志》，9(5):1083–1089，1997年。[83] B. Scholl、C. Tepohl、M. A. Ryan、C. I. Thomas、N. Kamasawa 和 D. Fitzpatrick。双眼突触网络支持视觉皮层神经元的眼间反应对齐。《神经元》，110(9):1573–1584，2022年。[84] J. T. S. Smits 和 P. G. Vos。点刺激中连续曲线的知觉。《知觉》，16 131，1987 1 121( ): – 。[85] L. W. Tu。《流形导论》。Springer，2011年。[86] W. R. Uttal。《三维空间中的视觉形状检测》。新泽西州希尔斯代尔：L. Erlbaum Associates，1983年。[87] W. R. Uttal。《三维空间中的视觉形状检测》。Psychology Press，2013年。[88] J. Wagemans、J. H. Elder、M. Kubovy、S. E. Palmer、M. A. Peterson、M. Singh 和 R. von der Heydt。视觉知觉中的百年格式塔心理学：I. 知觉分组与图形—背景组织。《心理学公报》，138(6):1172，2012年。[89] S. W. Zucker。立体视觉、阴影与表面：曲率约束耦合神经计算。《*IEEE*会刊》，102(5):812–829，2014年。