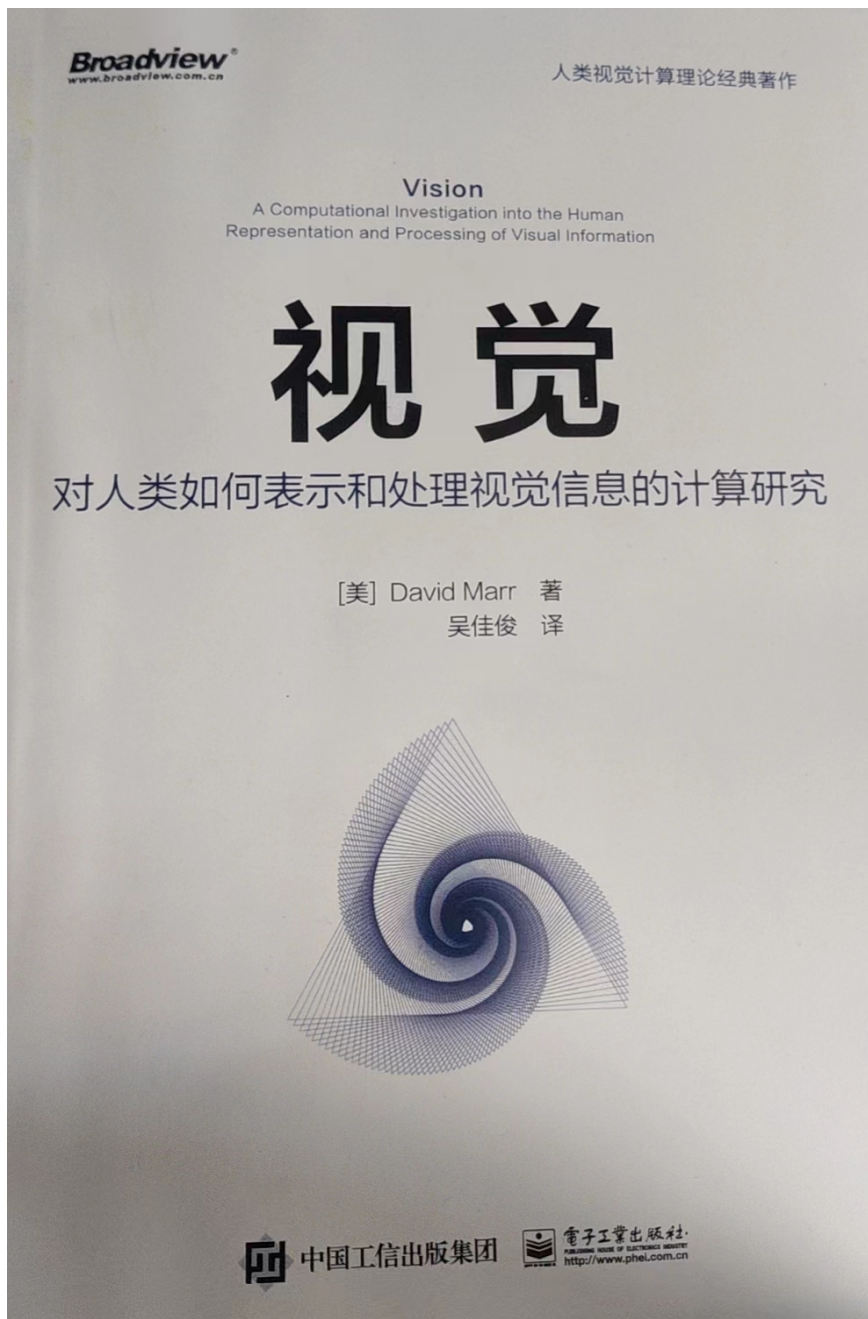




Vision-Ch01

☰ Author	David Marr
🔗 Link	https://zhuanlan.zhihu.com/p/272384722
📊 Score	★★★★★
📅 Completed	@November 3, 2023 → November 4, 2023
⚙️ Status	Done
📄 Type	Book



Notes

综述

Key takeaways

Quotes

Summary

这就是我从pdd上16块买的那本盗版视角哈哈哈哈哈

Notes

- 背景：牛顿（1704）奠定了现代色彩视觉工作的基石 → 亥姆霍次（1910）的生理光学论文 → Wertheimer（1912, 1923）注意到视频中连续 → 格式塔学派注重整体性后失败 → 实验心理学盛行主要分析具体性质和表现

▼ 综述

知觉心理学(1, 2, 3)、心理物理学 (4) 起源于适应性和阈值检测研究的结合

1. Brindley, 1970 : **三色理论 (trichromatism)** [未全面理解知觉, 此研究主要重点在性质]
 2. 迈尔斯 (Miles, 1931) 和瓦拉赫与奥康奈尔 (Wallach & O'Connell, 1953) 的实验, 它们证实了在适宜的条件下, 一个陌生的三维形状可以只从变化中的单眼投影 (注: 单只眼睛看到的二维图像) 中被正确地感知。
 3. 1960年贝拉·朱尔兹 (Bela Julesz) 设计了计算机生成的**随机点立体图 (random-dot stereograms)** [数字电子计算机的发展使双目视觉的表现分析研究成为可能]
 4. 坎贝尔和罗布森 (Campbell & Robson, 1968) 关于独立的**空间频率调谐通道 (spatial frequency-tuned channels)** 即我们早期阶段的感知器官, 对特定尺度或空间间隔内**图像的亮度变化敏感的通道是存在的** (引发大量文章, 视觉感知特征定量描述)
 5. 1971年, 罗杰·N·谢巴德 (Roger N. Shepard) 和杰奎琳·梅茨勒 (Jacqueline Metzler) : **心理旋转 (mental rotation)** , 重点在提出的问题类型, **表示 (representation)**
 6. **放大器的发展** : 阿德里安 (Adrian, 1928) 和他的同事能够记录伴随着神经信号传输的微小电压变化。→ 放大技术的改进, 这使记录单个神经元成为可能 → 细胞“**感受野 (receptive field)**”的概念 (Hartline, 1940) and 哈佛学派著名的一系列关于神经元在视觉通路更深层次行为的研究 (Kuffler, 1953 ; Hubel & Wiesel, 1962, 1968)
 7. 心理学感兴趣的问题可以通过神经生理学实验来启发: 巴洛 (Barlow, 1953) 对青蛙视网膜神经节细胞的研究。[对单个神经细胞的活动描述, 是一个对于理解神经系统功能的、足够完整的描述。] **还原论**
 8. 当时, 还原论的方法似乎很可能最终取得成功。1950~1970发展比较迅速。
 9. 但在深处, 有些东西出了问题。没有一项新的研究成功地阐明了视觉皮质的连接。
 10. 最终, 在图宾根的托马索·波吉奥 (Tomaso Poggio) 和我自己 (Marr & Poggio, 1977; Marr, 1977 b) 独立达成但共同表述的: 必须有一个额外的认识层次, 在这个层次上, 我们才能**分析和理解感知过程中所执行的信息处理任务的特征**, 将问题作为信息处理任务的**分析**。没有它, 就不可能真正了解所有这些神经元的功能。
- “过程”和“表示”

- 表示 (representation) ，例如数字的表示形式可以有二进制、阿拉伯数字、罗马数字
- 信息的表示方式会极大地影响用它做不同事情的难易程度。
- 过程 (process) → 加法性质 → 通用定理

Key takeaways

- 计算什么、为什么计算、约束条件唯一地定义
- 映射、映射的抽象属性被定义、适当性 (appropriateness) 和充分性 (adequacy)
- 信息处理的三个层次：计算理论、表示与算法、硬件实现

Quotes

层次具体含义：

计算理论	表示和算法	硬件实现
计算的目的是什么？为什么计算是合适的？执行计算的策略逻辑是什么？	这个计算理论怎么能被简化呢？具体来说，输入和输出的表示是什么，以及转换的算法是什么？	如何在物理上实现表示和算法？

当人们在20世纪70年代初对这类问题进行反思时，人们逐渐意识到，缺失某些重要的东西，它们在神经生理学或心理物理学两门学科中都没有出现。关键的观察是，神经生理学和心理物理学的任务是描述细胞或被试的行为，而不是解释这种行为。

但在20世纪60年代，几乎没有人意识到机器视觉是困难的。造成这种误解的原因是我们人类本身就很擅长视觉。

通过将解释划分为不同的层次，就可以对计算的内容和原因作出明确的说明，并构建理论，说明计算的内容在某种意义上是最优的，或保证正确运行。去掉了特别的因素，用坚实的基础取代了启发式的计算机程序，在此基础上可以建立一个真正的学科。

过程：我们研究的一个中心主题是，分离出那些足以定义一个过程、且普遍适用于整个世界的约束条件。

1. 处于层次的一个极端——顶级层次的是设备的抽象计算理论。在**计算理论**中，设备的表现由从一种信息到另一种信息的**映射**来刻画，这种**映射的抽象属性**要被准确定义，并且它对手头任务的**适当性（appropriateness）**和**充分性（adequacy）**也要被证明。
2. 处于中间层次的是**输入和输出的表示形式**的选择，以及用于将前者**转换**为后者的**算法**。
3. 处于另一个极端的是如何在物理上**实现算法和表示**的细节——可以说是详细的计算机体系结构。

通过理解正在解决的问题的本质，而不是通过研究它被实现的具体机制（和硬件），可能更容易理解算法。

Summary

- 综述人话概括：过去人们不会试图去从整体理解知觉感觉，也不会去构建整个系统的运行规律，只是单调的从性质和表现进行分析，没有试着去理解它们。三色理论（trichromatism）**主要重点在性质；数字电子计算机的发展使双目视觉的表现分析研究成为可能**；1971年，心理旋转（mental rotation）研究中的问题类型值得人们思考：**表示（representation）**；上半叶**放大器技术**的发展与改进可以让人们**单个记录神经元，这是一种还原论的思路**，当时五六十年代的一些研究比较有进展，人们对此都有期待，希望可以从神经元探索出知觉感觉的奥秘，然后到七十年代这个思路忽然进行不下去了，”没有一项新的研究成功地**阐明了**视觉皮质的连接“。
- 不同的表示方法针对不同的问题，各自都有各自的妙用。
- 信息处理的三个层次：计算理论、表示与算法、硬件实现