

## Chapter1 绪论

1. 认知心理学家所说的“表征”是什么意思？神经学家说的“表征”又是什么意思？

答：表征指的是将现实世界的物理属性和抽象形成知识。

在认知心理学方面，“表征”即指心理表征。

心理表征：（1）认知心理学的核心概念之一，指外部世界对象在心里的定义（包含抽象或者具体的定义），代表着心理活动中信息或知识的表达和记录方式。（2）物理属性与心理表征具有其一致性。

在神经科学方面，“表征”即指神经表征。

（1）外部世界对象的属性在神经信号或生物系统中的显示。

（2）例如边缘、颜色、运动

2. 认知神经科学中“相关性”和“因果性”研究的目的是什么？他们分别采用了哪些研究手段或研究技术？（p21 页 ppt、课本 p139）

答：相关性研究主要为了探寻哪些大脑区域或者神经反应对于感知或认知的特定方面是有关的。

因果性研究的目的是探寻哪些大脑区域或者神经反应对于感知或认知的特定方面是必要的。（P73，76）

3. 提供一些你认为的好和不好的视觉认知工程的例子。

答：

好的：工厂产品质量的检测。不仅大大减少了该方面工作所需要的人力，还提高了检测精度，利于产品的质量检测。

不好的：景区人群计数。这个问题有更加简便的方式来做，完全无需使用视觉工程的知识，会使得问题复杂化。比如可以借助 GPS 定位来确定一个区域里面的人数。

## Chapter2 大脑功能

1. 脑叶分为哪几个部分？各自的功能是什么？

大脑的四个脑叶分别是额叶、顶叶、颞叶、枕叶，每种脑叶分别主管不同的功能。1、额叶，位于大脑半球的前部，主管运动、语言、智能、情感等，与推理、计划、问题解决有关。2、顶叶，位于大脑半球的后顶部，主管感觉、阅读等，处理触觉、压力、温度以及疼痛等知觉，并能调节注意力或分配空间注意力。3、颞叶，位于大脑半球的外侧，人们感受声音、确认自己的空间位置的最高中枢就在这里，由许多处理听觉、语言和某方面记忆的亚区组成。4、枕叶，位于大脑半球的后部，主管视觉、阅读等，被称为视觉皮层。

## Chapter4 视觉机制

### 1. 人类视觉为什么会产生马赫带、赫曼方格这样的视觉错觉现象？

**答：**马赫带效应（Mach band effect）指的是一种主观的边缘对比效应。当观察两块亮度不同的区域时，边界处亮度对比加强，使轮廓表现得特别明显。马赫带效应的出现是人类的视觉系统造成的。生理学对马赫带效应的解释是：人类的视觉系统有增强边缘对比度的机制。

视网膜神经网络由三级神经元构成。第一级神经元是光感受器细胞（视锥和视杆细胞），它们首先将光能转换为神经元的细胞膜内外电位变化，进而影响第二级神经元双极细胞和水平细胞的活动。双极细胞再将处理后的信息传递给第三级神经元，即神经节细胞和无长突细胞；只有第三级神经元的神经节细胞才能产生神经脉冲（动作电位），将整个视网膜神经网络信息处理结果传入脑中。

介于光感受器和双极细胞之间的水平细胞体积较大，能够从大量的光感受器接受兴奋性输入，也能对光感受器和双极细胞发出负反馈信号，产生侧抑制作用。正是水平细胞的这种侧抑制作用，造成双极细胞和神经节细胞的中心-周边同心圆颞颞式的视觉感受野（在视觉系统中任何一级单个神经元都在视网膜上有一块属于自己的小“视野”——感受野，只有适宜的光刺激到感受野时，该神经元才产生反应），从而使视网膜具有检测空间对比度的功能。正是由于神经节细胞有这种同心圆颞颞式感受野，使得其对亮暗边界处于其两侧周边分界线上时，反应最大或最小。

**赫曼方格：**侧抑制假说，认为激活较弱的光感受细胞信号被旁边的强激活细胞的信号所覆盖，因此不集中注意力，一些白色光源信号会被灰色的所覆盖。

人类视网膜的侧抑制现象。为了聚焦，视网膜里的神经节细胞会增强焦点处接收到的刺激，减弱这一点之外的刺激。但是看着方格时，会导致细胞误会所收到的刺激。一些科学家认为这种错觉会让那些视力没有聚焦到的白点隐身。

### 2 大脑处理视觉信息的过程是什么？

- (1) **视网膜：**视网膜中的神经元，根据不同的形态和功能主要分为五大类：光感受器负责光电转换，双极神经元负责接收光感受器输出的信号并传递给下游的视网膜神经节细胞，而信息从光感受器到双极细胞以及从双极细胞到神经节细胞的传递过程分别受到水平细胞和无长突细胞的调节，视网膜神经节细胞对信息进行加工整合后将电信号向下一级脑区外膝体的中继细胞进行传递。
- (2) **外膝体：**外膝体其中的神经元也是规则地分层排列，不同类型的神经元分布在不同分层，接收不同亚型的视网膜神经节细胞的输入，再将这些信息传输给初级视皮层。
- (3) **视皮层：**视皮层主要包括初级视皮层（又称作纹状皮层或视觉第一区域，即 V1 区）和纹外皮层（如视觉第二、第三、第四、第五区域等，即 V2、V3、V4、MT）。初级视皮层接收来自外侧膝状体的信息，然后通过 V2 和 V3 传递 V4、MT 区以及更高的脑区。整个的皮层信息处理过程由两条并行

的通路完成：V1、V2、V4 等组成的腹侧通路主要处理物体形状、颜色等信息；V1、V2、MT 等组成的背侧通路主要负责对运动等信息的感知。

### 3 结合前面即将的内容，分析计算机视觉中的卷积神经网络可能存在哪些真实神经系统中的处理过程？

答：

- (1) 视网膜需要从外界的光信号中提取出有效的特征信息，如物体颜色、大小、形状、深度、运动信息等，而卷积网络本质上也是在做特征提取。
- (2) **信息压缩**：人的视觉系统有约一亿个感光细胞，但是有且仅有 100 万视神经将外界的光信号传到了大脑。在卷积网络里面下采样操作也是一种信息筛选、压缩的过程。
- (3) **反向传播**：光线是先通过和光无关的西部，到达眼球后方的光感细胞之后再反向传入神经，卷积网络的参数更新也是反向传播的过程。
- (4) **横向整合**：在视觉系统中，横向整合通常调整各层之间的横向连接，全连接层也起到了信息整合的作用。
- (5) **局部感受野**：视觉神经元仅仅只对有限空间区域的光信号产生反应，而卷积核正好具有局部连接的特性。
- (6) 层与层之间的信号传导，每个层都有自己的功能。

## Chapter5 物体识别机制

### 1. 物体识别的神经编码假说有哪些？各自的特点是什么？

答：

#### (1) 单一神经编码假说：

是一种“独热表征”的神经机制  
噪音大，很容易导致识别错误  
十分脆弱，一个神经元的死亡会导致无法识别对象  
泛化能力差，物体发生变化时候难以适应

(2) 层级神经元编码假说：腹侧通路每个皮质区都对前一级区域来的信息来进行处理，从而形成更加抽象复杂的表征，并最终与记忆中存储的物体形状表征进行匹配。

视网膜和 LGN 神经元关注点  
V1 中的简单细胞关注边  
V1 中的复杂细胞关注平行边、拐角  
腹侧通路神经云关注复杂形状、物体局部  
颞下皮质神经元分辨物体

(3) 集群神经元编码假说：通过多个相关神经元的集体激活最终产生对特定对象的知觉。

分布式表征的神经机制，符合视觉信息的分布式处理假说  
可解释对视觉相似的对象混淆现象  
噪音小、反脆弱、有利于识别对象的鲁棒性  
泛化能力强、可以适应对象的不断变化

2. 谈一下你对物体识别过程中“自顶向下”、“自底向上”、“由粗到精”等策略的理解。

答：

“自顶向下”指的是在目标识别的过程中，有约 10%的信息通过上丘和枕核，进一步输入额叶皮质进行处理，额叶皮质的输入可以影响沿着腹侧通路的处理。

额叶利用早期场景分析和对当前上下文的知识生成有关场景的预测，将这些自上而下的预测与沿着颞叶皮层腹侧通路进行的自下而上的分析进行比较，通过限制可能性领域来更快的识别物体。

“自底向上”指的是视皮质不同层次细胞功能专一化，每一层次提取不同尺度上的信息进而进行信息整合。视觉信号特征在视皮质之间以既并行又串行的方式传递。

其中，v1 区负责编码边缘，线条等简单特征，只对特定方向有反应；v2 对错觉轮廓有反应，是色调敏感区；v3 区负责视觉信息的传递；v4 区是色彩感知的主要区域，参与曲率计算，运动方向选择和背景分离；MT 负责运动知觉。随着逐层之间的传递，视觉感受野也增大，视觉特征由局部迈向整体。

“由粗到精”主要是由于目标识别有快慢两种通路，其中快速通路对对象整体进行识别，慢速通路对对象的局部进行识别。因此，对目标的识别是先识别轮廓整体信息，后识别局部信息，也就是由粗到精的过程。

3. 以神经系统信息流动顺序和信息处理加工过程为依据设计无人车行驶过程中遇到障碍时可能的处理流程。

- (1) 获取位于当前道路上目标车辆的预定区域内的障碍物信息。
- (2) 根据获取到的图像信息进行特征提取，获得障碍物的特征信息。
- (3) 将目前障碍物的特征信息与知识库内存储的信息进行比对（查找是否有该情况的处理方法），如果有则按照该方法，目标车辆采取行动。或是在内部仿真目前的路况，模拟目标车辆采取不同行驶策略规避障碍物的多个模拟结果。
- (4) 根据特征比对的结果/模拟结果选取最优的目标车辆行驶策略。
- (5) 按照（4）中的策略采取行动。

## 第七章 运动控制

1. 那些类型的运动可以在没有反馈的条件下仍然存在？反馈信息是如何被运动系统用作实时控制和学习的？

答：运动指令起自于脊髓靠下的部分的运动可以在没有反馈的条件下仍然存在，如膝跳反射，当膝盖被外界敲击时，四头肌伸展，引起肌肉纺锤体的感受器放电。感觉信号传至脊髓背根并直接激活  $\alpha$  运动神经元，使得四头肌收缩。

知觉和动作是紧密关联的，一个区域既可以参与知觉也可以参与运动控制。运动信息会造成周围环境的改变，外部来源的信息会通过外部环路（小脑、顶叶皮质和外侧运动前区）来指导视觉相关的运动。

在学习的早期阶段，外侧运动区由于有大量投向顶叶皮质的纤维联系，最适于在运动计划之下综合外部信息，如果我们学习到了产生运动和其效果之间的对

应关系，就可以预期动作的结构，从而更多的利用内侧运动结构的功能。

**2. 帕金森病患者为什么会有运动障碍？给出一种基于基底神经节生理性能的解释。为什么外界输入多巴胺可以缓解症状？**

**答：**基底神经节帮助解决备选动作作为最终支配运动器官而竞争的问题。强大的抑制性基线活动可以使其保持对运动系统的抑制，使有可能发生动作的皮质表征激活而不引起肌肉活动。随着特定的运动计划得到激活，对特定神经元的抑制信号削弱。而多巴胺是奖赏通路的关键递质，偏向于产生特定的反应而不是其他的反应，人们脑中的多巴胺神经元会随着年龄的增长而萎缩，当这些萎缩的比例过大时，会导致帕金森。

外界输入多巴胺可以在一定程度上环境多巴胺神经元萎缩而导致的多巴胺缺乏的问题，可以缓解行为紊乱的现象。

## 第十章 语言

**1. 心理词典在大脑中是如何组织的？在单词的语义识别过程中，针对单词的高水平表征和低水平表征的区别是什么？**

**答：**心理词典是关于语义、句法和词形信息的心理存储器。

心理词典被认为是以特异性信息网络的形式组织起来的，特异性信息网络在所谓的词素水平上以单词的形式存在，在词元上以单词的句法特征的形式存在，在词元水平上，单词的语义特征也被表征出来了，这种词义信息定义了概念水平，单词的词义知识是通过概念水平来表征的，在心理词典中，这些表征组织涉及到了单词之间的关系，也就是说在这个网络中倾向于联系紧密的单词。

还有一种观点认为心理词典中的表征是根据单词间的意义关系组织起来的。支持这种观点的证据是利用词汇判断任务的语义启动研究。

在单词的语义识别过程中，低水平表征指的是从感觉输入建构的表征，高水平表征指的是单词被加工之前的语境建构的表征，其所表征的对象不同。

**2. 请简述 ChatGPT 的整体训练流程。结合人类语言和 Transformer 的理论特点，尝试分析为什么无监督学习能够在 ChatGPT 的训练中发挥重要的作用。**

**答：**训练流程主要分为两个步骤：预训练和微调。

**预训练：**在预训练阶段，模型通过学习大量无标签文本数据来掌握语言的基本结构和语义规律。这些数据主要来源于互联网，包括书籍、新闻文章、博客、论坛等。训练过程中，模型使用一种名为“掩码语言模型”（Masked Language Model, MLM）的方法。这意味着在训练样本中，一些词汇会被随机掩盖，模型需要根据上下文信息预测这些被掩盖的词汇。通过这种方式，ChatGPT 学会了捕捉文本中的语义和语法关系。

**微调：**在微调阶段，模型使用特定任务的标签数据进行训练，以便更好地适应不同的应用场景。这些标签数据通常包括人类生成的高质量对话，以及与特定任务相关的问答对。在微调过程中，模型学习如何根据输入生成更准确、更相关的回复。

语言本就具有低噪音、高度抽象、信息密度大的特性，因此适合无监督/自监督学习。从人类的文明发展中也不难看出，语言是人类对世界认知的抽象、总

结。

Transformer 使用自注意力（self-attention）机制，允许模型在处理序列数据时，关注与当前词汇相关的其他词汇，从而捕捉文本中的长距离依赖关系，而语言正是具有强上下文依赖的特性，和 transformer 的特性十分契合，这也是无监督可以在 ChatGPT 训练中发挥较好效果的部分原因。