

信息学院人工智能专业方向

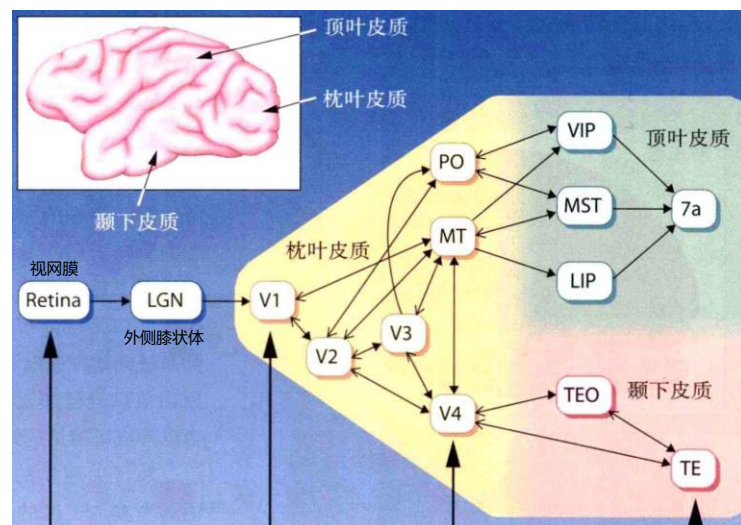
《脑与认知科学》

认知神经科学基础：物体识别

- 注：本课程部分内容整理自课程教材、参考书籍或公共资源，特此致谢！

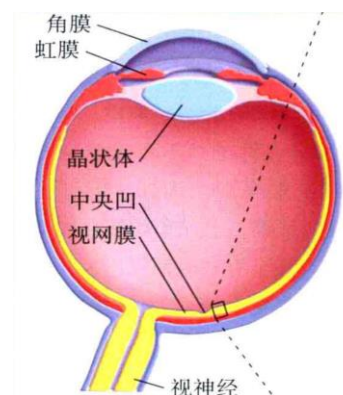
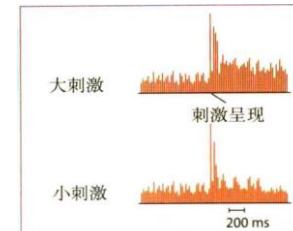
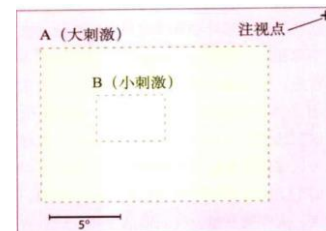
视觉信息处理的皮质通路

- 背侧通路/空间通路(where pathway)
 - 功能：处理运动和深度信息
 - 输入：视神经M通路
 - 途径：V1→V2→V3→MT
 - 输出：顶叶皮质
- 腹侧通路/内容通路(what pathway)
 - 功能：处理形态和颜色信息
 - 输入：视神经M和P通路
 - 途径：V1→V2→V4
 - 输出：颞下皮质



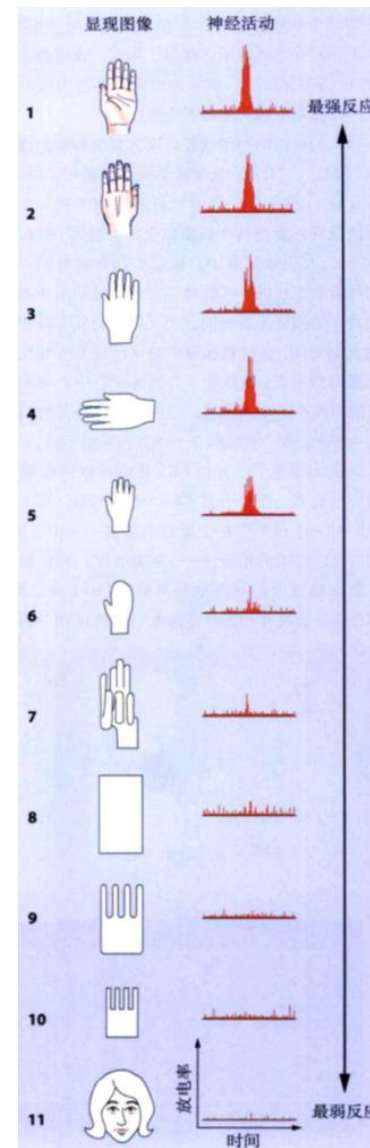
顶叶皮质神经元的特点

- 背侧通路流入的顶叶皮质神经元具有以下特点：
 - 可被物体的空间变化（出现、移动）激活，但对物体大小、形状均不敏感
 - 40%的神经元感受野位于视野中心（中央凹），其他神经元的感受野在中心视野两侧
- 以上特点有利于顶叶皮质神经元检测物体的空间信息



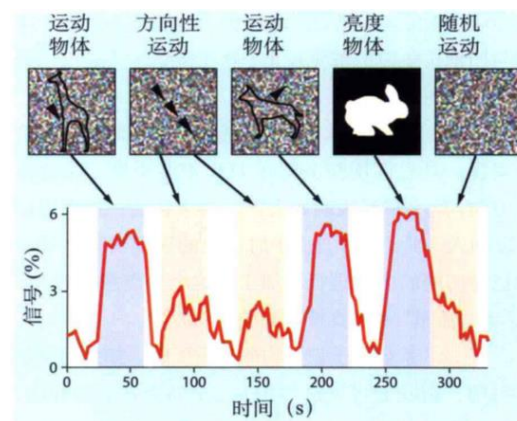
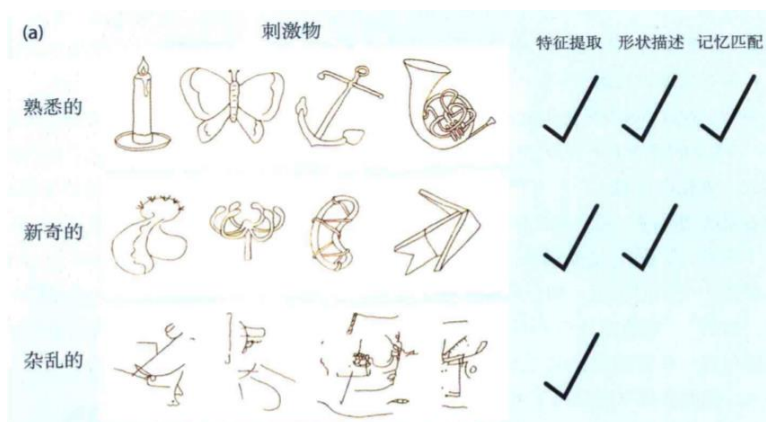
颞下皮质神经元的特点

- 腹侧通路流入的颞下皮质神经元具有以下特点：
 - 神经元的感受野全都位于视野中心（中央凹）
 - 一部分神经元可被不同形状、大小的物体所激活，但其他神经元仅能被特定复杂物体所激活
- 以上特点有利于颞下皮质神经元对物体进行识别



腹侧通路 & 物体识别

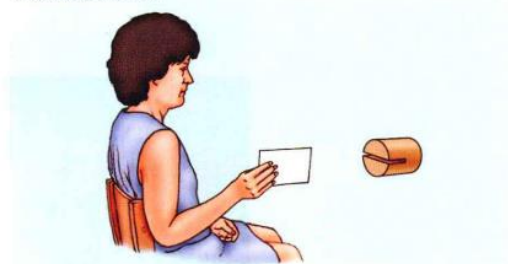
- 对正常人的脑功能研究也显示，与腹侧通路相关的皮质区域在物体识别过程中被激活



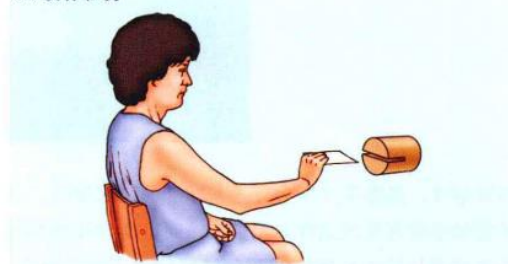
皮质通路损伤与物体识别异常

- 腹侧通路损伤→视觉失认症
 - 通过视觉识别物体能力严重下降，同时对物体的三维方向知觉存在障碍
 - 但基于视觉的运动控制和记忆表现正常
- 背侧通路损伤→视觉性共济失调
 - 能够正确识别物体，如报告凹槽的方向
 - 但不能通过视觉信息引导动作
- 上述现象也支持了腹侧通路与物体识别密切相关的结论

(a) 外显匹配任务



(b) 动作任务



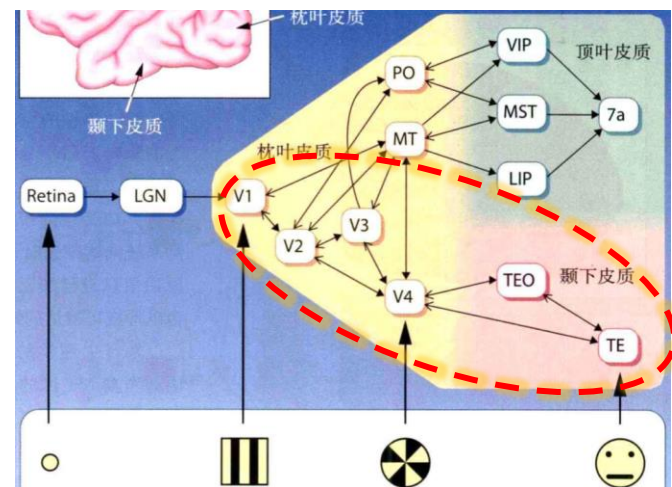
(c) 记忆任务



意识知觉和动作知觉的分离。(a) 在外显匹配任务中（意识条件），D.F. 被要求匹配卡片和凹槽的方向，她显示出严重障碍。(b) 在动作任务中，要求 D.F. 把卡片插入凹槽，她毫不犹豫地做出了正确的动作。(c) D.F. 在记忆任务中也没有障碍，证实了她对方向的知识是完好的。

皮质通路的功能联系

- 腹侧通路和背侧通路是视觉信息处理的两条关键通路，其功能分别是识别物体和检测其空间位置/运动信息
- 背侧和腹侧通路在功能上并非完全独立，而是存在密切联系
 - 顶叶对腹侧通路的视觉注意控制起关键作用
 - 如深度等空间信息对复杂场景中的物体识别也有帮助



物体识别的挑战性

- 基于视觉的物体识别是一个极其困难的问题，受到视角、距离、光照、遮挡等很多因素的影响
- 大脑的物体识别（物体知觉）能力，对上述因素以及单目、旋转等影响均非常鲁棒



我们眼中的世界的样子取决于我们的观看位置。这两张照片取自同一个景色，但是从两个不同位置在两种不同条件下拍摄的。每个观看位置展现了该景色的新视角，包括从其他观看位置看被遮挡的物体。而且，颜色随着时间和天气发生变化。尽管景色具有这样的可变性，但我们仍然能很容易地认出两张照片都是旧金山和金门大桥的照片。

视觉的物体恒常性

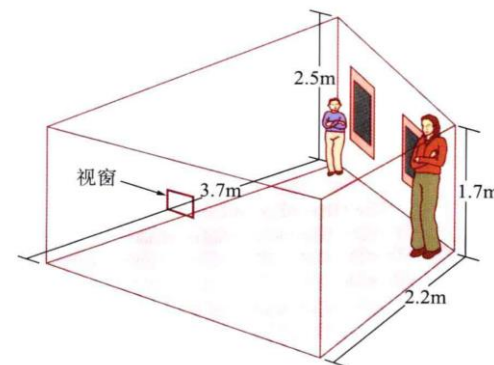
- 物体恒常性(object constancy): 在视角、距离、光照等因素发生变化的情况下, 对物体的视知觉仍旧保持不变
- 物体识别中的恒常性现象不会影响对相似物体间细微差别的感知



物体恒常性。同一辆车的 4 幅图在视网膜上的成像十分不同。尽管有这种感觉变异, 但我们能迅速识别这些图是同一辆车。

视觉的物体恒常性

- 影响物体恒常性的因素
 - 观察者的位置因素，如距离、视角等
 - 视觉系统可区分视角或物体本身变化所导致的形状改变
 - 亮度变化因素，如可视部分、阴影、反光
 - 视觉系统对物体的亮度变化不敏感
 - 遮挡因素，如复杂场景中的物体遮挡
 - 视觉系统可分离复杂场景中的不同物体



尽管感觉输入有许多变异源，包括阴影 (a) 和部分遮挡 (b)，但物体恒常性仍然能够保持。

物体识别的信息处理问题

- 大脑用于物体识别的视觉信息：

- 物体的形状
- 物体表面的颜色、纹理
- 物体的运动

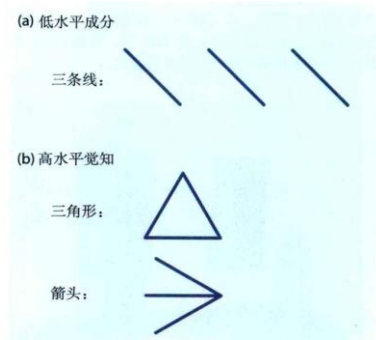
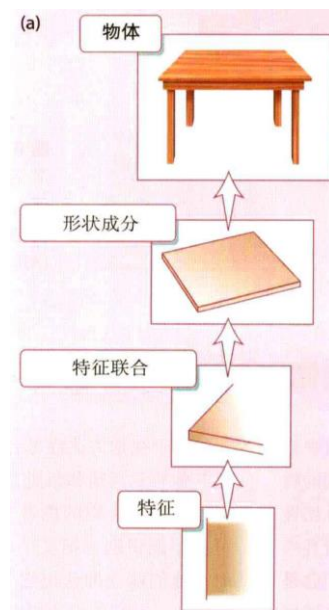


分析二维形状和三维形状。不论这些物体被画得如何不同寻常，我们识别它们都毫无问题。我们可能从未见过粉红色的大象或格子状的苹果，但我们的物体识别系统能够区分我们用以识别大象和苹果的关键特征。

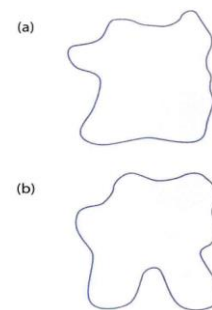
- 物体形状识别的两个关键问题：
 - 物体形状如何被表征？
 - 对物体形状的表征如何不被多种因素影响？

物体形状表征的神经机制

- 根据Hubel和Wiesel的前馈模型假说，在物体识别过程中，腹侧通路中每个皮质区域都对前一级皮质区域传来的信息进行处理，从而形成更为抽象复杂的表征，并最终与记忆中储存的物体形状表征进行匹配
 - 点 ← 外侧膝状体神经元
 - 边 ← V1中的简单细胞
 - 平行边、拐角等 ← V1中的复杂细胞
 - 复杂形状 ← 腹侧通路中的神经元
 - 显著 (salient) 的形状信息被表征
 - 物体局部 ← 腹侧通路中的神经元
 - 物体 ← 颞下皮质神经元



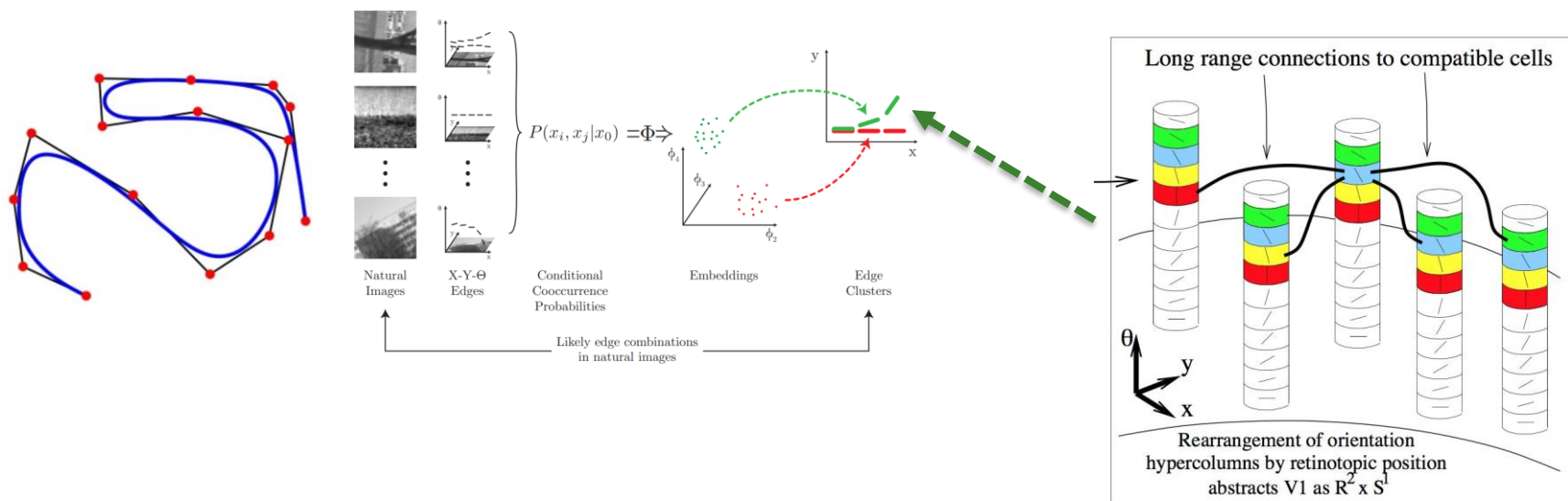
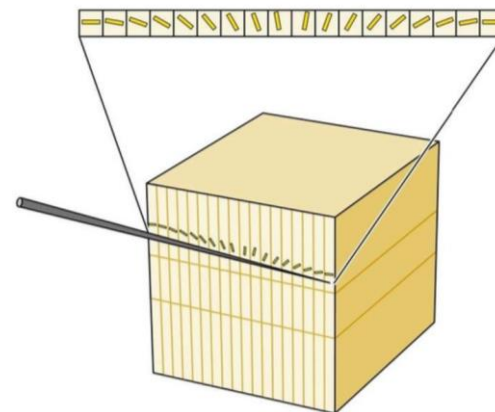
相同基本成分（三条线）根据排列可能形成不同物体（如三角形或箭头）。尽管低水平成分（a）是相同的，高水平知觉（b）却是不同的。



学习这两幅画 10 秒，然后试着根据记忆把它们画出来。你的画很可能反映了最大弯曲点是最突出的信息源这一事实。在（a）中，左边的突出物要比右边的波状边缘更有区别性。在（b）中，你画的物体肯定包括了底部把两个附件分离开的豁口。

物体形状表征的神经机制

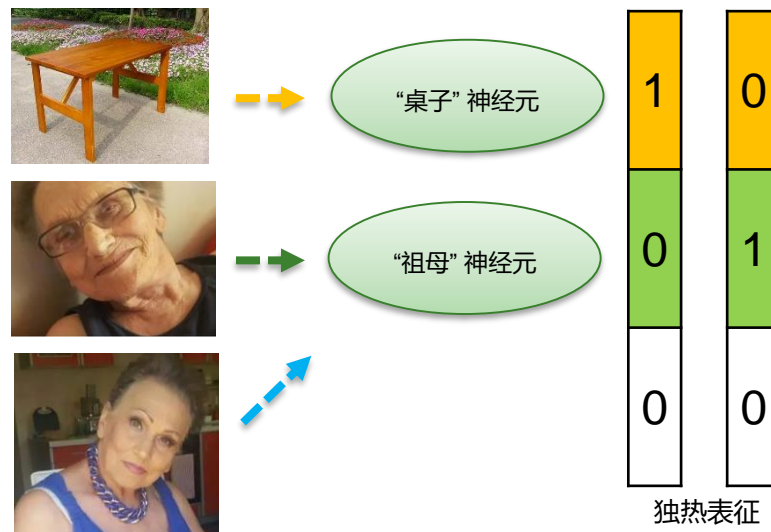
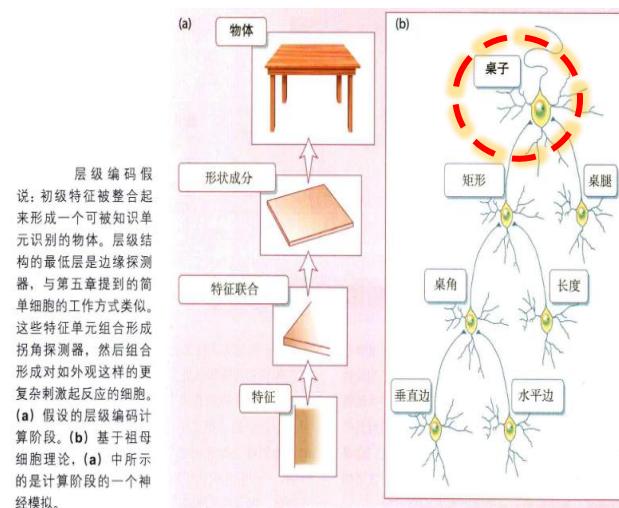
- 近期研究显示，V1区皮质柱的长程水平连接(Long-range horizontal connection)可能与复杂形状表征有关



<http://www.cs.yale.edu/homes/vision/zucker/papers/LawlorZucker2013.pdf>

物体识别的神经机制

- 物体识别的单一神经元编码假说：
通过一个可识别复杂物体的神经元（“知识单元”）激活最终产生对特定物体的知觉
 - “桌子”、“祖母”神经元
- 单一神经元编码假说本质上是假设视觉系统在物体识别时采用了独热（one-hot）表征的神经机制
 - 噪声大，很容易导致识别错误
 - 太脆弱，一个神经元死亡即导致无法识别物体
 - 泛化能力差，物体发生变化时难以适应

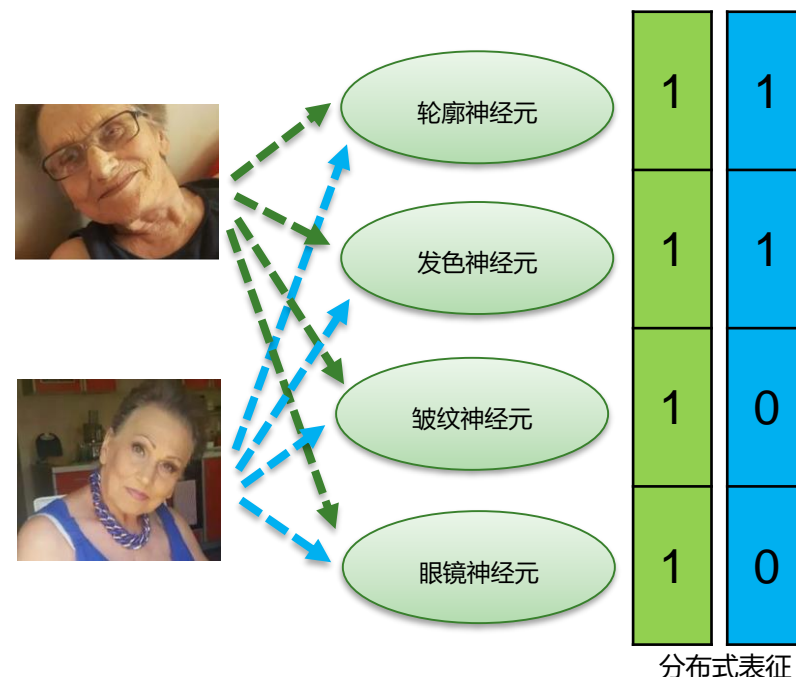


物体识别的神经机制

- 物体识别的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物体的知觉
- 集群神经元编码假说本质上是假设视觉系统在物体识别时采用了分布式(distributed)表征的神经机制
 - 符合视觉信息的分布式处理假说，可解释对视觉上相似物体的混淆现象
 - 噪声小、反脆弱，有利于物体识别的鲁棒性
 - 泛化能力强，可以适应物体的不断变化

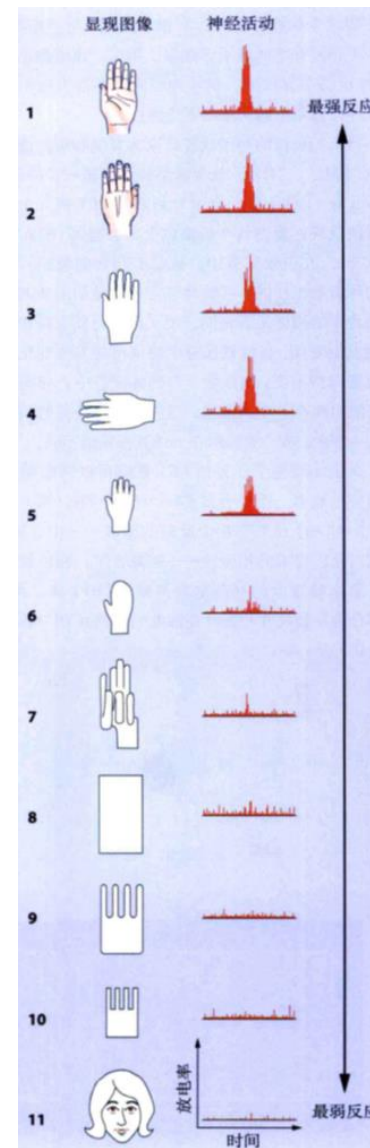


集群编码假说：由同时激活一群定义性特征来定义物体。这里“祖母”因她的眼镜、面部轮廓、头发颜色等特征共同出现而得到识别。



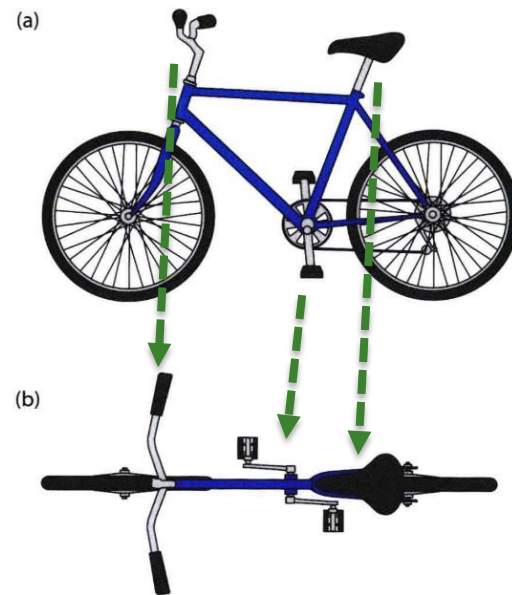
物体识别的神经机制

- 对颞下皮质神经元的单细胞研究结果也支持物体识别的集群神经元编码假说，虽然一些颞下皮质神经元可被复杂物体特异性激活，但这种特异性并不是绝对的
 - 如前面介绍的颞下皮质神经元，无论手或手套出现时都能激活这个神经元
- 另外在上述研究中，也未发现单个神经元能够被某一只手特异性激活，但人脑的知觉能力却可以识别出不同人的手



物体识别的视角相关性

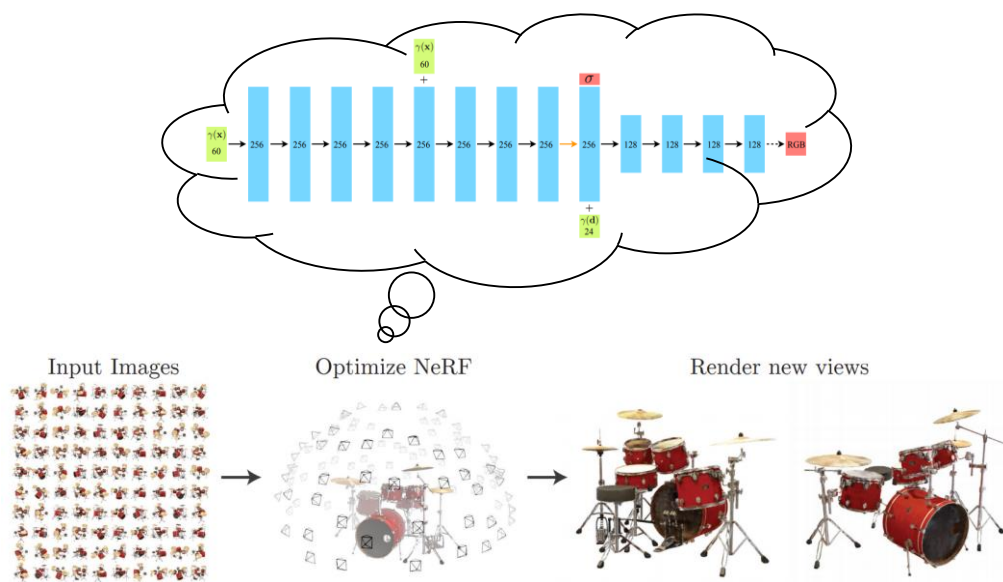
- 物体识别的视角相关(view-dependent)假说:
 - 物体识别需要依赖特定视角下的物体表征, 人类记忆中保存了大量的视角相关物体表征, 因此物体识别是将当前视角下的视觉信息与记忆存储的视角相关表征进行(快速)匹配的过程
- 问题: 当视角非常罕见甚至是全新时, 如何进行物体识别?
 - 大脑的视觉系统对新视角下的视觉信息与记忆中视角相关表征的关系进行推理
 - 对新奇物体的识别研究显示, 随着视角差异增加, 对物体是否相同的判断时间会显著增加



物体识别的视角依赖理论认为识别过程依赖于观察位置。识别一辆自行车的两幅画——一个是从侧面看而另一个从高处看——要求把不同的感觉输入与视角依赖的表征匹配。从侧面看 (a), 我们可能是把该画和已贮存的从该视角看到的自行车的表征在座位、链条和轮辐这些特征上进行匹配。从高处看 (b), 表征需要包括轴、座位、手把和脚蹬。

物体识别的视角相关性

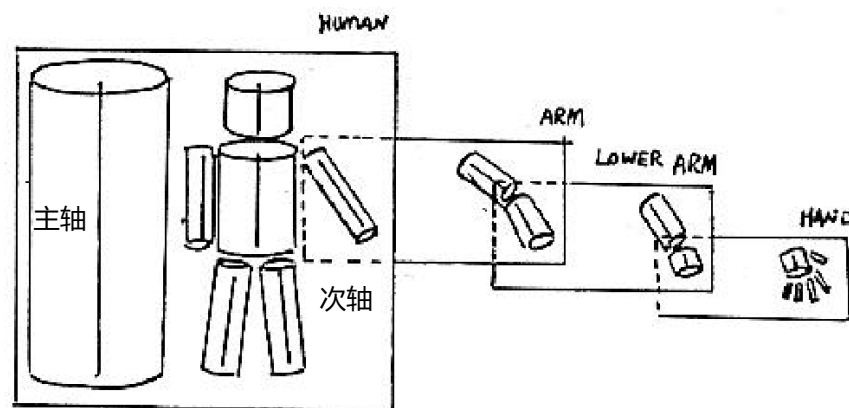
- 利用神经网络进行全新视角的图像合成研究
 - 通过静态图片进行训练，用深度神经网络获得场景的5D(空间位置 (x, y, z) 和观察方向 (θ, ϕ)) 连续表征，再通过渲染获得全新视角下的图像



<http://www.matthewtancik.com/nerf>

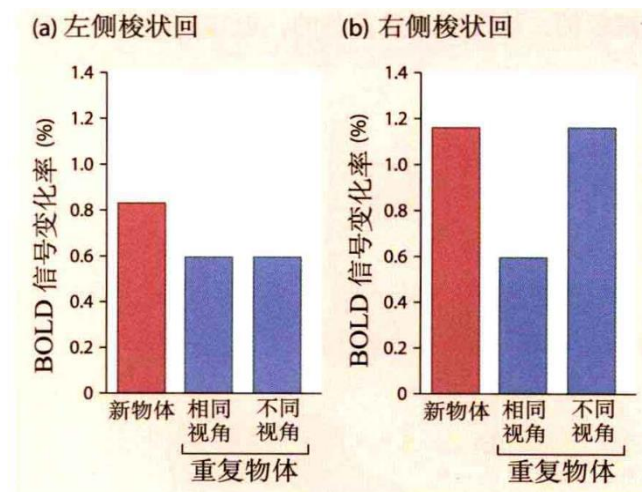
物体识别的视角相关性

- 物体识别的视角不变 (view-invariant) 假说：
 - 大脑的视觉系统通过建立了不受视角影响的物体基本表征进行物体识别
 - Marr在视觉计算理论中提出，物体识别时视角不变的基本表征是物体的主轴和次轴



物体识别的视角相关性

- 脑功能成像的实验显示，大脑中可能同时存在视角相关和视角不变的物体识别机制
- 重复抑制效应：如果某刺激最近被激活过，那么对该刺激的神经反应会更快更有效
 - 左侧大脑半球中的腹侧通路对重复出现物体的表现与视角无关→视角不变的物体识别
 - 右侧大脑半球中的腹侧通路对重复出现物体的表现与视角有关→视角相关的物体识别



重复抑制效应。(a) 不管物体是相同视角还是不同视角，在左侧腹侧枕皮质都观察到了重复抑制效应，与视角不变表征一致。(b) 相反，右侧腹侧枕皮质的活动仅当第二次是相同视角时降低，与视角依赖表征一致。

与物体识别相关的认知异常

- 视觉上的物体失认症：
 - 患者对物体的形状、颜色和动作知觉均正常，但不能根据视觉信息正确识别物体
 - 物体失认症是一种视觉认知障碍，并不是记忆缺陷
- 物体失认症可以分为两种类型：
 - 统觉性失认症：无法形成正常的物体视觉表征，与腹侧通路损伤有关
 - 联络性失认症：不能用产生的视觉表征来识别物体



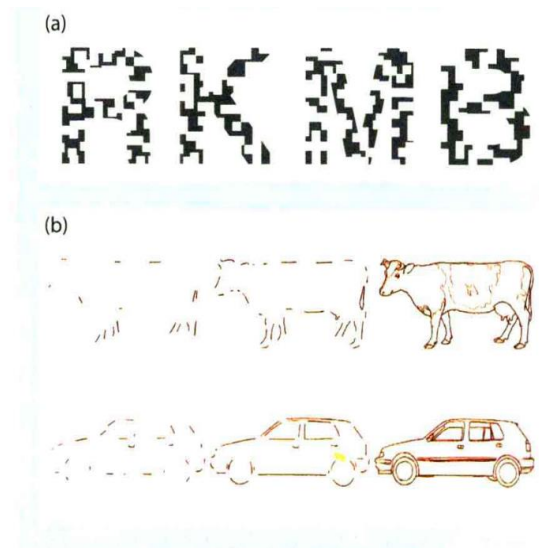
视觉失认症的测试图示。让 G. S. 命名呈现的物体，他不能给出正确的答案“密码锁”，而坚称是“电话”。密码锁和传统的旋转电话具有相似的成分——数字和刻度——意味着 G. S. 对一些信息进行了加工。另外，患者的手在模仿开密码锁的动作。



失认症与记忆缺失。要诊断失认症，我们有必要将一般性记忆障碍予以排除。(a) 视觉失认症患者不能单独通过视觉识别钥匙，但把它们拿起来后就能马上识别。(b) 记忆缺陷患者不能识别钥匙，甚至把它们拿起来后也不能。

与物体识别相关的认知异常

- 统觉性物体失认症：
 - 与右侧大脑半球的视觉皮质损伤密切相关
 - 患者对常见物体的形状、颜色知觉无明显异常，但无法正常识别物体，特别是在形状信息不显著时识别能力严重下降

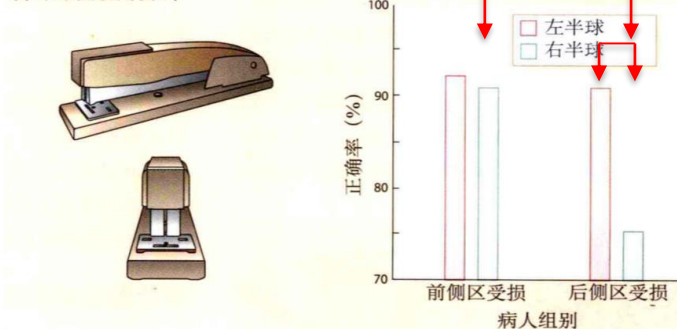


用于评估对劣化图形识别的测验。尽管左半球损伤患者有更严重的语言问题，但右半球损伤的失认症患者在完成不完整字母任务 **(a)** 时有更多困难。在 Gollin 不完整图形任务 **(b)** 中，实验者给被试呈现一系列物体图画，每个比前一个更完整。右半球损伤患者只能识别出更完整图画中的物体。

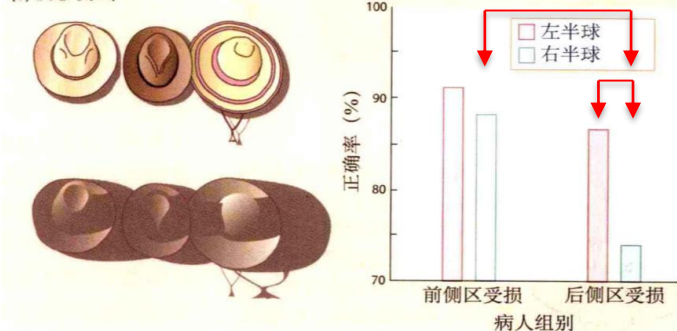
与物体识别相关的认知异常

- 实验研究显示，统觉性物体失认症患者的物体恒常性丧失
 - 统觉性物体失认症患者可以识别常见视角、光照下的物体，但难以识别非常见视角、光照下的物体
 - 当物体的视觉信息与记忆存储的物体表征可快速匹配时，患者的识别能力不受影响；但当视觉信息不常见需进一步推理时，物体识别能力显著下降
 - 右侧大脑半球在基于视觉的物体恒常性方面起主导作用

(a) 不常见视角测试

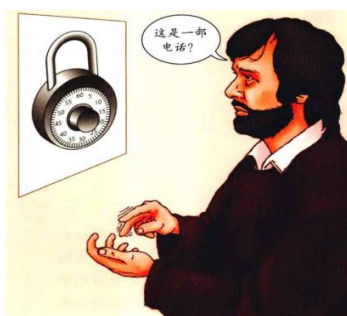


(b) 阴影测试



与物体识别相关的认知异常

- 复杂场景中的物体恒常性丧失现象：
 - 失认症患者虽可识别不同视角、光照下呈现的单独物体，但对重叠形状的识别能力很差
 - 不能快速形成对物体的整体表征，在物体识别时必须依赖于物体局部的视觉信息
 - 缺乏物体的整体表征导致无法在复杂场景中分离不同的物体



视觉失认症的测试图示。让 G. S. 命名呈现的物体，他不能给出正确的答案“密码锁”，而坚称是“电话”。密码锁和传统的旋转电话具有相似的成分——数字和刻度——意味着 G. S. 对一些信息进行了加工。另外，患者的手在模仿开密码锁的动作。

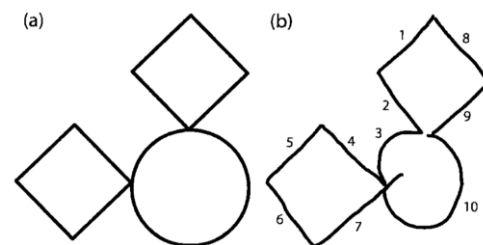


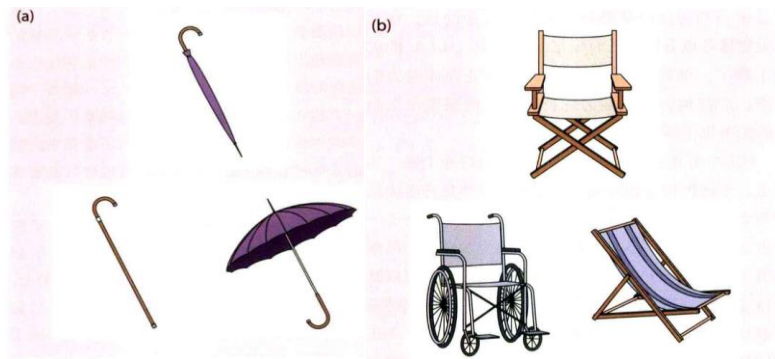
图 6.29 患者 C.K. 临摹 (a) 中的图形。他的整体成绩 (b) 很不错，我们能很容易地识别出两个菱形和圆形。数字表示他临摹这些线条的顺序。然而，不像健康被试，C. K. 并不是完整地临摹完每个物体，而是沿着外边界临摹（这样做意味着在物体间来回转换）。

与物体识别相关的认知异常

- 联络性失认症：
 - 与左侧大脑半球的皮质损伤密切相关
 - 患者的物体恒常性正常，但无法通过视觉信息对物体进行命名和描述
- 联络性失认症患者和统觉性失认症患者虽然都难以完成物体的功能匹配任务，但其机制却不相同
 - 统觉性失认症：在视觉信息不常见时无法形成物体表征
 - 联络性失认症：能形成物体的视觉表征，但无法获得名称等信息



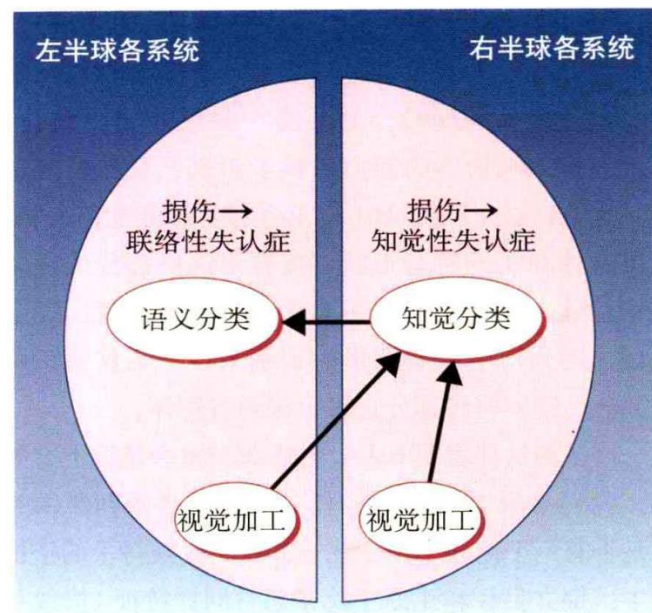
图 6.25 患者 F.R.A. 的画。尽管 F.R.A. 不能命名视觉呈现的物体，但他能够十分成功地给这些复杂图画中的各成分涂色。显然，他分解图画非常成功，但仍不能识别这些物体。



功能匹配实验

物体识别的认知模型

- 物体识别的认知模型包括知觉分类和语义分类两个阶段：
 - 知觉分类阶段：两侧大脑半球的腹侧通路协同形成物体的视觉表征，然后在右侧大脑半球中与记忆中储存的物体表征进行匹配并产生物体知觉
 - 语义分类阶段：在左侧大脑半球中将物体知觉与记忆存储的物体名称、功能等知识进行联系



物体识别的范畴特异化

- 联络性失认症患者对不同物体范畴（生物、工具等）的识别能力可存在明显差别
 - 一个患者对剪刀、椅子等物体图像的识别率约为90%，而对不同种类生物的识别能力仅有6%
- 大脑皮质中的物体知识组织方式产生了物体识别的范畴特异化现象
 - 脑损伤并不会彻底破坏记忆中全部的物体知识，即使最严重的联络性失认症患者也能保留对一部分物体的识别能力

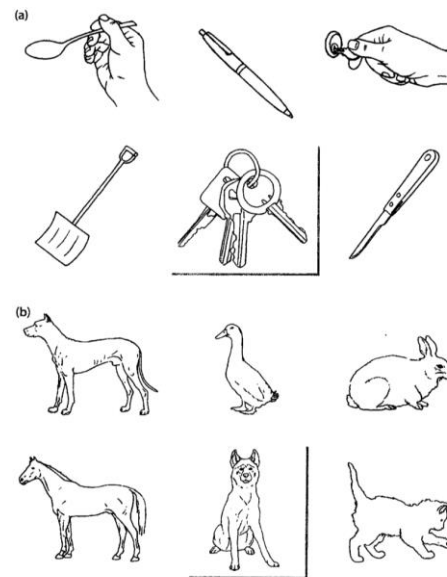
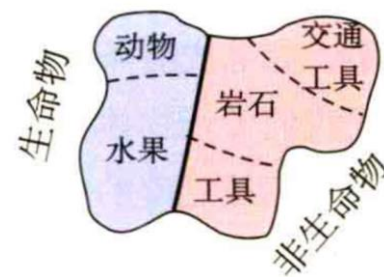


图 6.31 范畴特异性失认症测验。在每个条件下，让被试从 5 个图画中选出与方框中的探测项最匹配的一个。一个对非生命物有选择性障碍的患者可能在 (a) 例中表现差，但可能在 (b) 例中表现正常。预期对生命物有范畴特异性损伤的患者会出现相反结果。

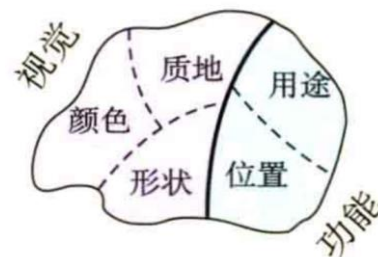
物体识别的范畴特异化

- 基于范畴的物体知识组织方式假说：物体知识依据不同的范畴关系进行层级化组织
 - 动物：鸟/猫/虎，工具：剪刀/刀/锤子
- 基于表征的物体知识组织方式假说：物体知识依据不同的表征类型进行层级化组织
 - 基于视觉表征的组织方式：依据不同物体在形状、颜色、纹理等视觉表征方面的关系
 - 皮毛条纹：猫/老虎
 - 基于功能表征的组织方式：依据不同物体在用途、位置等功能表征方面的关系
 - 切割：剪刀/刀

基于范畴的组织

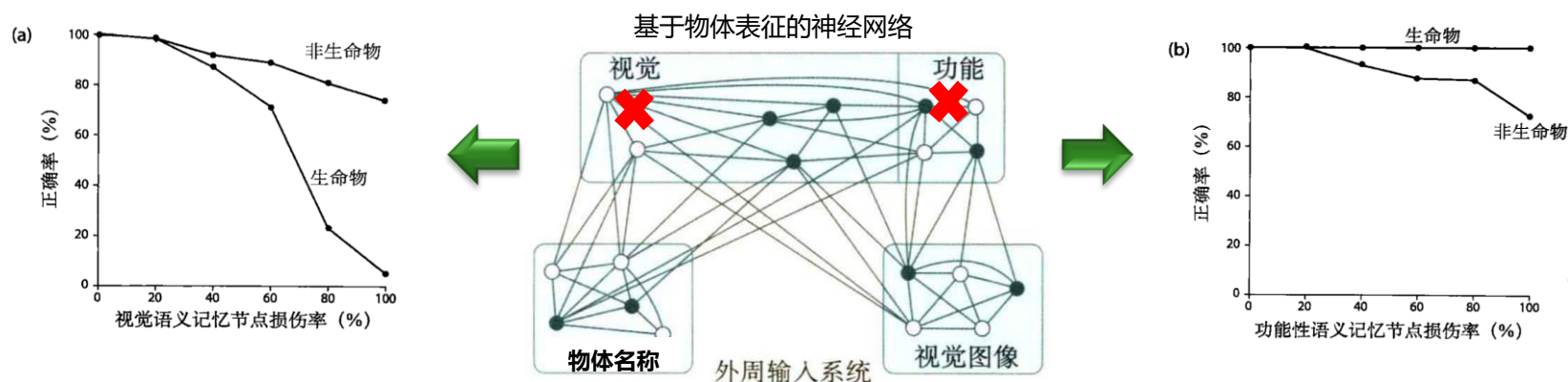


基于特征的组织



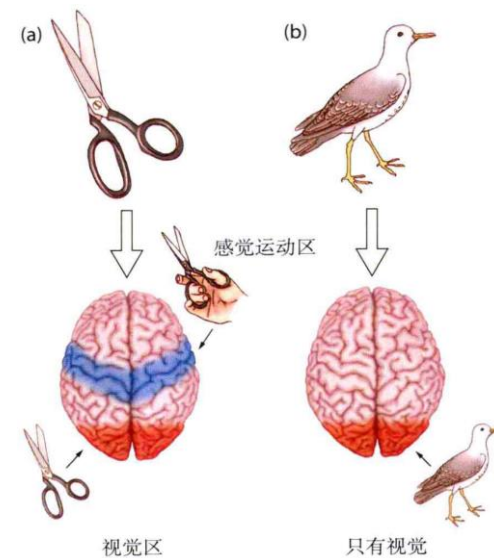
物体识别的范畴特异化

- 通过建立包含物体视觉、功能表征的人工神经网络模型，对基于表征的物体知识组织方式假说进行验证
 - 用于视觉表征的神经元无法激活 → 对生物物的识别能力出现了严重下降，而对无生命物体的识别能力降低相对有限
 - 用于功能表征的神经元无法激活 → 对生物物的识别能力不受影响，而对无生命物体的识别能力降低
- 物体知识的组织方式并非要按照对应的范畴关系，而是可以通过物体视觉、功能的分布式表征关系进行组织



物体识别的范畴特异化

- 范畴特异化的物体识别异常的本质是识别生物和无生命物体所需的信息来源不同
 - 对生物的知识高度依赖于视觉信息，例如对老虎来说，皮毛条纹是重要的视觉表征
 - 对于无生命的物体识别，则需要依赖视觉和功能信息
- 失认症患者通常对无生命物体的识别能力相对较好
 - 使用物体时的其他感觉表征补充了视觉识别能力的不足



视觉失认症的测试图示。让 G. S. 命名呈现的物体，他不能给出正确的答案“密码锁”，而坚称是“电话”。密码锁和传统的旋转电话具有相似的成分——数字和刻度——意味着 G. S. 对一些信息进行了加工。另外，患者的手在模仿开密码锁的动作。

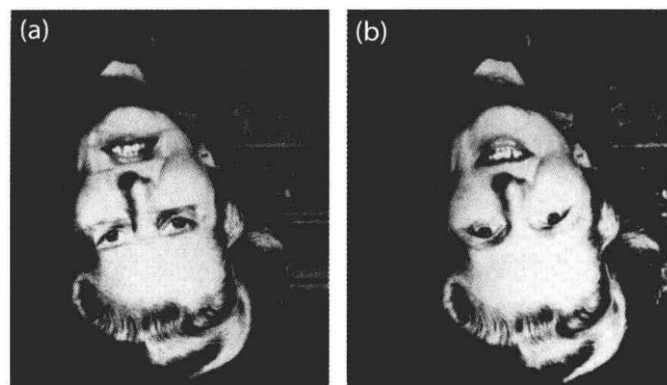
与面孔识别相关的认知异常

- 面孔失认症面孔：俗称“脸盲症”，指对陌生、熟悉甚至本人的视觉识别能力缺陷
 - 患者对面孔的识别能力明显差于对其他物体的识别能力
 - 面孔失认症和其他的物体失认症并不相同
- 人脑中是否有独立于其他物体识别的面孔识别系统？



与面孔识别相关的认知异常

- 人脑中存在独立于其他物体识别的面孔识别系统
 - 证据1：物体失认症患者无法识别物体，可以正常识别面孔
 - 证据2：面孔失认症患者可以进行精细(细粒度)物体识别，如一群绵羊中的不同个体等
 - 证据3：面孔失认症患者对倒立面孔的识别能力反而好于正立面孔



面孔识别的皮质通路

- 动物实验显示，颞下皮质中存在可被面孔特异性激活的神经元
- 通过脑成像技术，从猕猴大脑皮质中确定出多个参与面孔识别的功能区域，其中的神经元均对面孔敏感

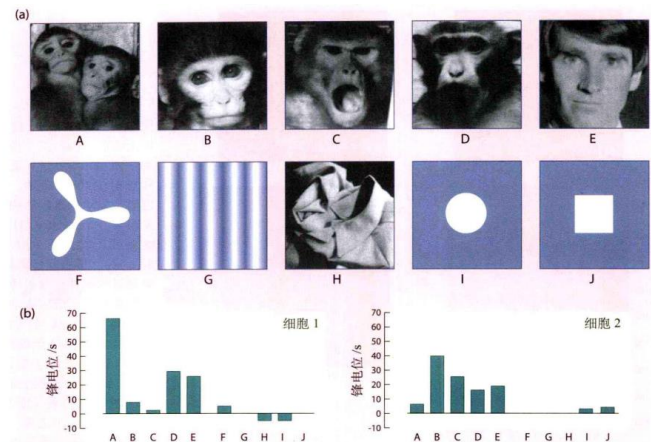
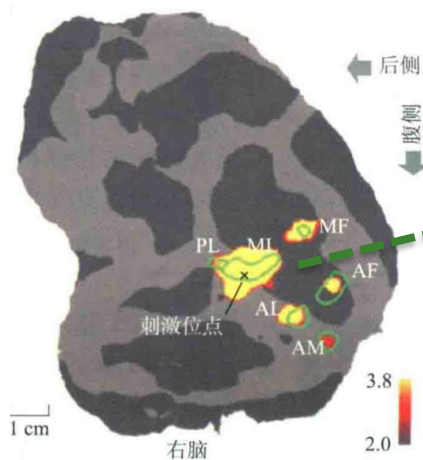
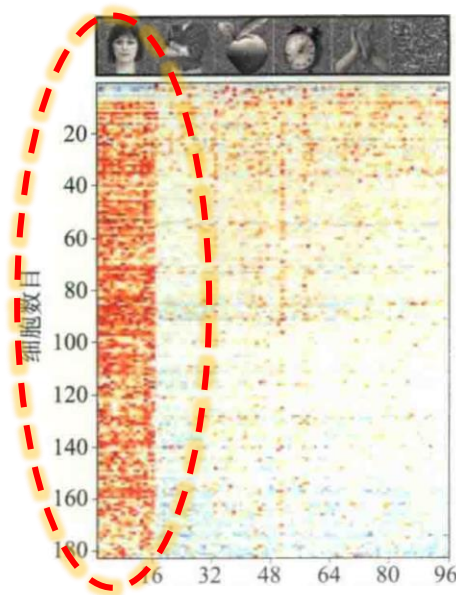
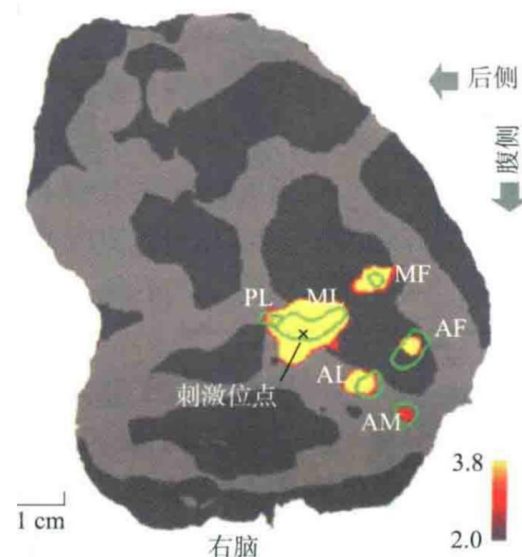


图 6.35 恒河猴颞上沟中识别面孔的细胞。图 (b) 显示两个细胞对 (a) 中 10 个刺激 (A-J) 的反应。两个细胞都对许多面孔刺激反应活跃。当动物看物体时，这些细胞活动没有变化，或者在某些情况下，这些细胞相对于基线来说实际上是被抑制了。图 (b) 表示细胞的放电变化（与没有接受刺激的基线条件相比）。



面孔识别的皮质通路

- 猕猴大脑中参与面孔识别的功能区域之间存在相互联系，一个功能区域的激活会引起其他功能区域的激活
- 面孔视觉信息的大致流动方向：ML/MF区
→AL区→AM区
 - ML/MF区中的神经元被特定视角的面孔激活
 - AL区中的神经元不仅能被特定视角的面孔激活，而且还能被这些面孔的镜像激活
 - AM区中的神经元能够被不同视角的面孔激活



面孔识别的皮质通路

- 脑成像研究显示，颞叶的梭状回面孔区（fusiform face area，也被称为FFA）在人脑的面孔识别中起重要作用
- 面孔失认症与梭状回面孔区的异常有密切联系

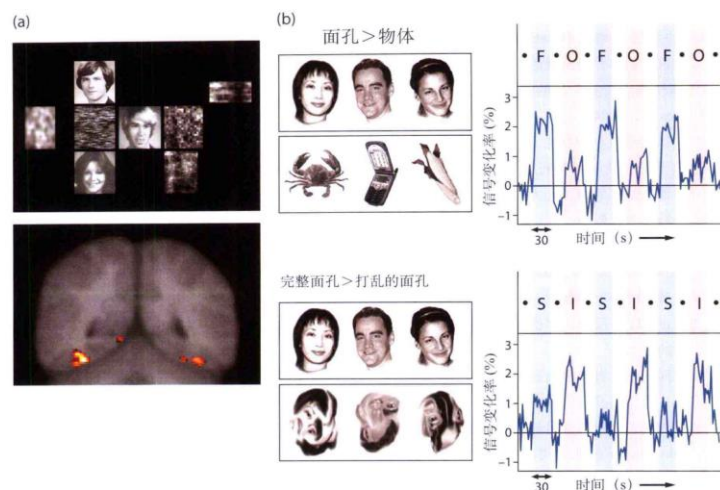
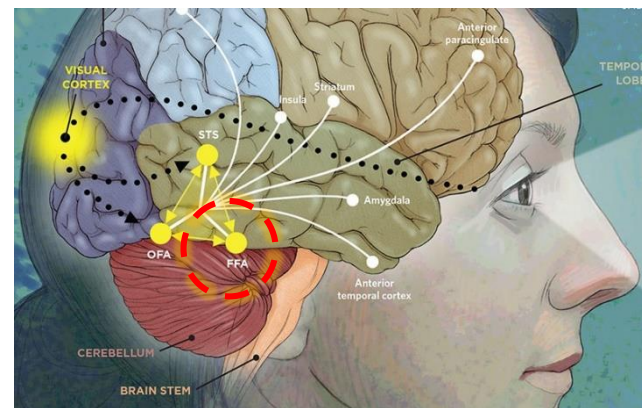


图 6.37 分离出面孔知觉时的神经区域。(a) 与仅看随机图案相比，当被试看面孔和随机图案时，fMRI 显示梭状回双侧激活。注：根据神经放射学传统，右半球在左边。(b) 在另一个 fMRI 研究中，被试看各组刺激交替变化。在一个扫描序列中，刺激在面孔和物体间变化；在另一个扫描序列中，它们在完整和打乱的面孔间变化。右列显示了扫描各种刺激序列中梭状面孔区的 BOLD 信号。在每一区组中，刺激来自不同组——面孔 (F)、物体 (O)、打乱的面孔 (S) 或完整面孔 (I)——这些区组由短间隔的注视点分隔开。结果显示，面孔呈现区组中 BOLD 信号更大。



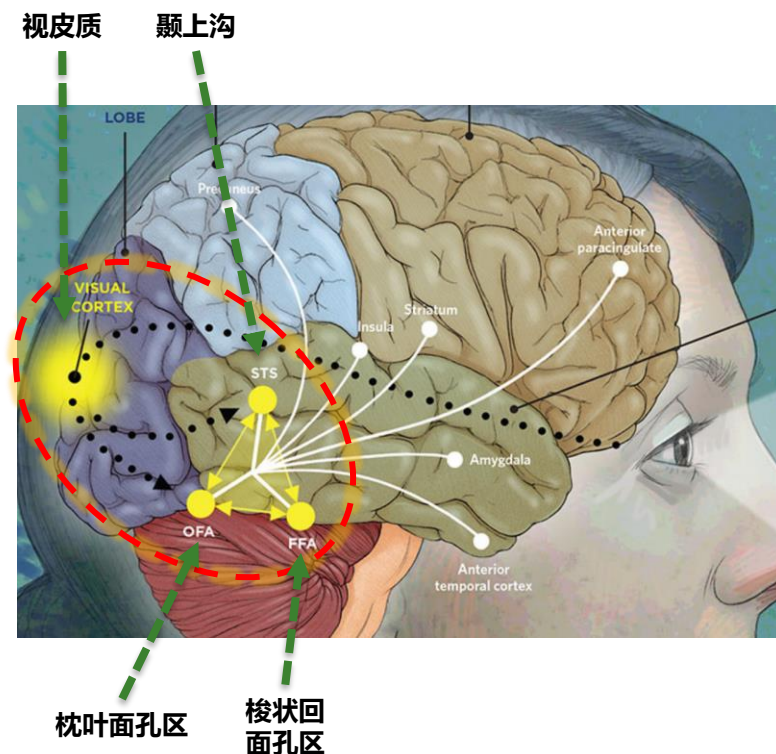
面孔识别的皮质通路

- 梭状回面孔区是否仅在面孔识别时起作用，还是在所有物体的精细识别中都发挥作用？
 - 在对其他物体进行精细识别时，除面孔区之外的梭状回区域发生激活
- 梭状回作为整体参与物体的精细识别，其中面孔区FFA负责面孔识别，而其余的梭状回区域负责其他物体的精细识别



面孔识别的皮质通路

- 面孔识别的核心皮质通路：
 - 梭状回面孔区（FFA）：面孔整体的视觉信息处理和面孔识别
 - 枕叶面孔区（occipital face area, OFA）：面孔局部的视觉信息处理
 - 颞上沟（superior temporal sulcus, STS）：面孔表情信息处理
- 视觉信息的流动方向：视皮质→枕叶面孔区/颞上沟→梭状回面孔区
 - 枕叶面孔区、颞上沟和梭状回面孔区之间存在相互联系



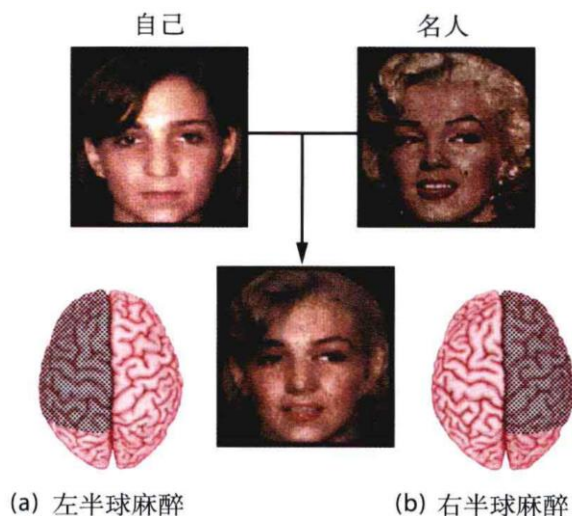
<https://www.the-scientist.com/cover-story/a-face-to-remember-36508>

面孔识别的特异化

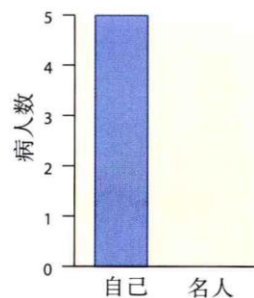
帅成这样
这让别人怎么活啊



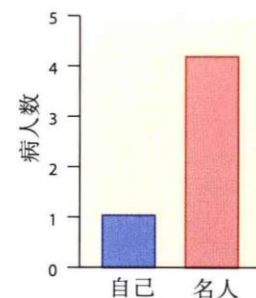
- 右侧大脑半球在自己面孔的识别（自我知觉）中起主导作用
 - 对正常人的脑功能成像研究显示，右侧大脑半球在看到自己照片时有更为明显的激活
 - 对癫痫患者的认知实验显示，右侧大脑半球可以从合成照片中识别出自己的面孔



(a) 左半球麻醉

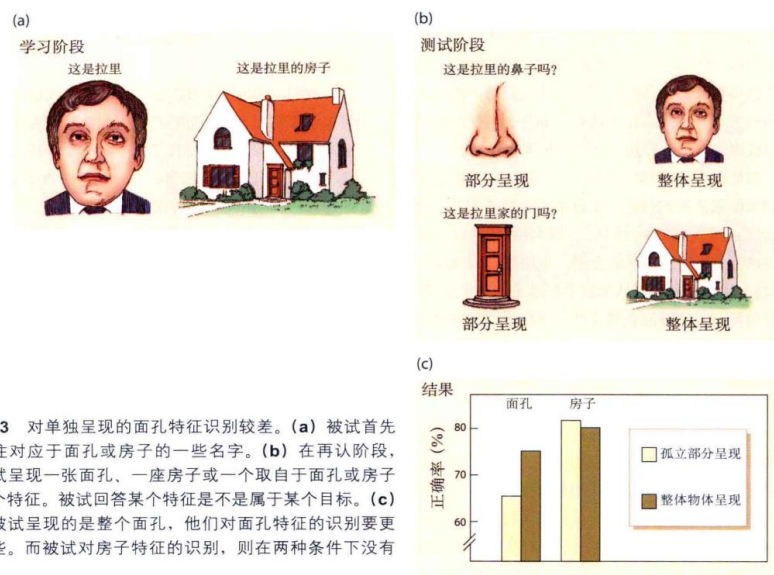


(b) 右半球麻醉



面孔识别的特异化

- 人脑中的面孔识别系统主要依赖整体而非局部的视觉信息，而其他物体的识别系统则需要依赖局部和整体的视觉信息
- 面孔失认症患者通常无法通过面孔的整体视觉信息进行识别



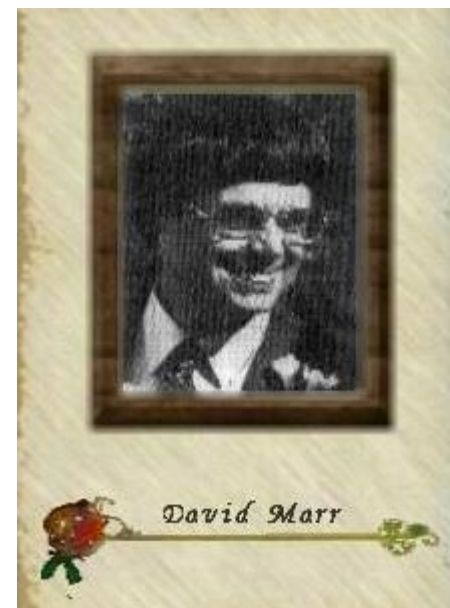
物体识别的双系统假说

- 大脑中存在整体视觉信息处理系统和局部视觉信息处理系统
- 两个信息处理系统在识别过程中的贡献与识别对象有关：
 - 单词：局部视觉信息处理系统
 - 面孔：整体视觉信息处理系统
 - 其他物体：整体和局部视觉信息处理双系统



马尔与计算机视觉

1977年，马尔第一次系统描述了计算机视觉这个领域，其中包含了计算机视觉（**Computer Vision**）和计算神经学（**Computational Neuroscience**）两个分支学科，他的工作对认知科学也产生了深远影响



http://www.sohu.com/a/155985171_337363

马尔关于视觉的计算理论

- 视觉是一种信息处理过程, 经过这个过程从图象中发现外部世界中有什么东西和它们在什么地方
- 信息处理有三个研究层次(three layers)
 - 第一个层次是信息处理的计算理论(theory)
 - 研究的是对什么信息进行计算和为什么要进行这些计算
 - 第二个层次是算法(algorithm)
 - 如何进行所要求的计算, 也就是要设计特定的算法;
 - 第三个层次是实现算法的机制(implementation)
 - 研究完成某一特定算法的计算机制



从图象推得形状信息的过程分成三个表征阶段

- **初始简图：**由图象中的变化和结构获得适当的表征
 - 包括检测亮度的变化，表征并分析局部的几何结构，以及检测光源、强光部分和透明度等照明效应
- **2.5维简图：**通过对初始简图进行一系列处理运算，推导出能反映可见表面几何特征的代表征
 - 包括表面朝向，观察者的距离、朝向和距离的不连续性，表面的反射情况以及对主要照明情况的粗略描述
- **三维模型：**实现对观察对象的三维结构在以物体为中心的坐标系中的表征
 - 在该坐标系下对物体表面性质进行描述

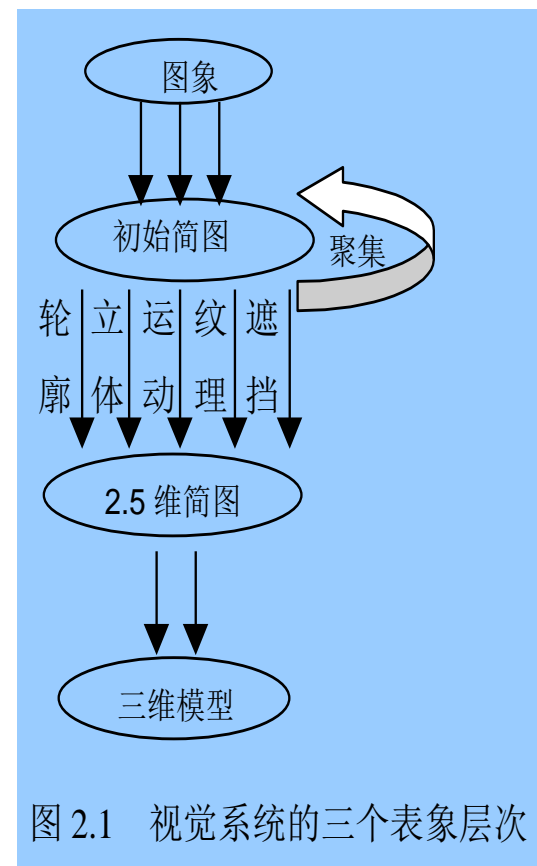
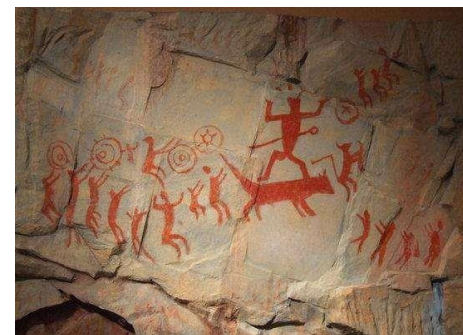
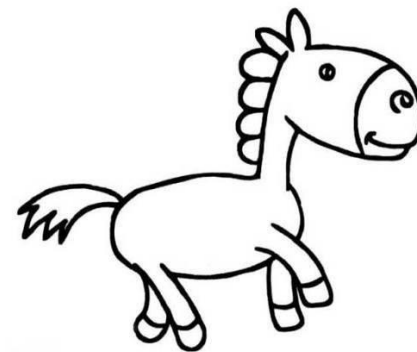


图 2.1 视觉系统的三个表象层次

初始简图

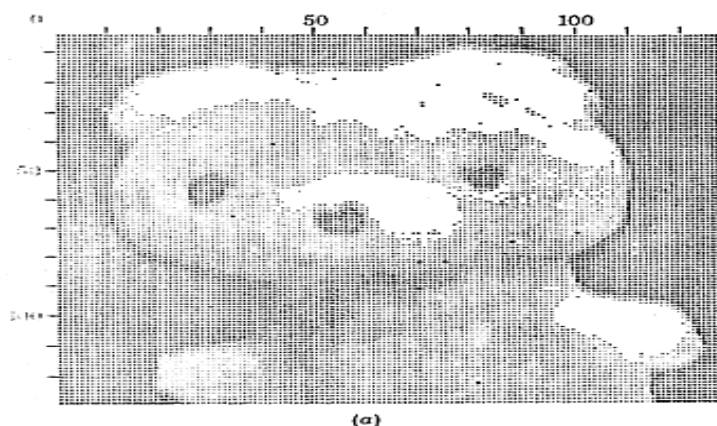
- 画家作速写时用边、线、点这样的符号就可以勾画出景物的素描。这样的素描图与实际景物在视网膜上产生的图象亮度阵列是很不相同的，但人们可以毫无困难地识别它们
- 说明视觉对图象所作的第一个运算是把它转换成一些初始符号构成的描述。这些描述所反映的不是亮度的绝对值大小，而是图象中的亮度变化和局部的几何特征



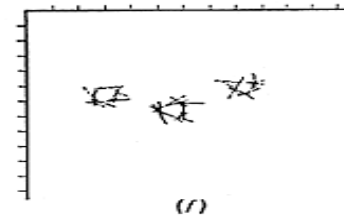
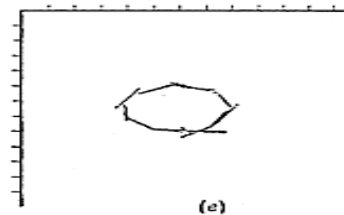
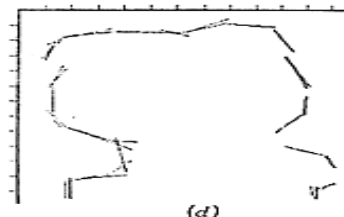
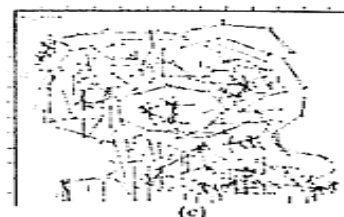
初始简图的组成

- 第一部分，用边缘段（edge segments）、线（bar）、斑点（blob）和端点（terminations）这些基元构成对图象中亮度变化的描述，这被称为未处理的初始简图
- 第二部分，利用虚拟线来完全而明确地表示几何关系，并且通过用各种方式对原始的基元进行选择。聚集和概括等过程来构成的更大、更为抽象的标记（tokens）。最后得到的描述层次可以覆盖某一尺度范围，这被称为完全的初始简图

初始简图的组成



Computer Vision 321

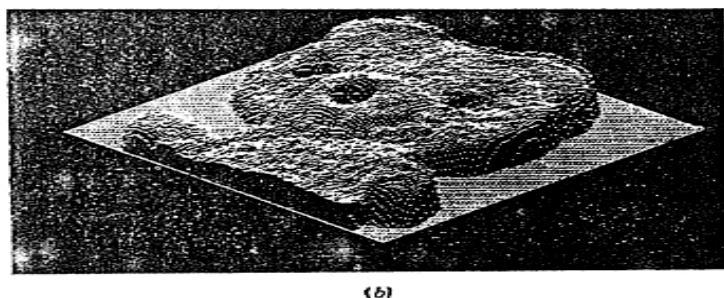


(a) 亮度图象

(b) 图象的亮度

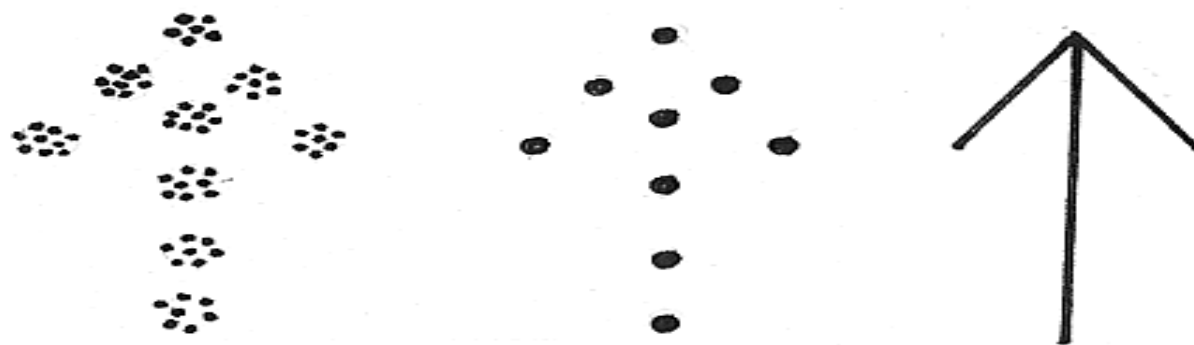
(c) 未处理的初始简图的分割，这时未进行许多聚集处理

(d) (e) (f) 是抽取完全的初始简图过程中找到的聚集



完全的初始简图

- 聚集运算：未处理的初始简图能明确表达的只是局部信息。因此还需要根据这些局部信息构成更大尺度的轮廓和区域
- Marr把聚集运算和纹理区别过程的最终结果称为完全的初始简图。这种表征完全和明确地表示区域的边界、物体的轮廓、以及基本的形状或结构



聚集过程示意图

2.5 维简图

- 在人类的视觉系统中存在双目立体视觉、运动、影调、纹理等相互独立的低层视觉处理功能模块，这些模块能提供关于表面形状的多方面信息

信息源

立体视觉

方向选择性

从运动恢复结构

光源

遮挡轮廓

其它遮挡线索

表面方向轮廓

表面纹理

表面轮廓

影调

信息类型

视差，因而可得到 $\delta\gamma$, $\Delta\gamma$ 和 S

$\Delta\gamma$

γ , $\delta\gamma$, $\Delta\gamma$ 和 S

γ 和 S

$\Delta\gamma$

$\Delta\gamma$

Δs

可能有 γ

$\Delta\gamma$ 和 γ

δs 和 Δs

在表中：

γ — 相对深度（按垂直投影），
就是观察者到表面点的距离

$\delta\gamma$ — γ 的连续或小的变化

$\Delta\gamma$ — γ 的不连续点

S — 局部表面方向

δS — S 的连续或小的变化

Δs — S 的不连续点



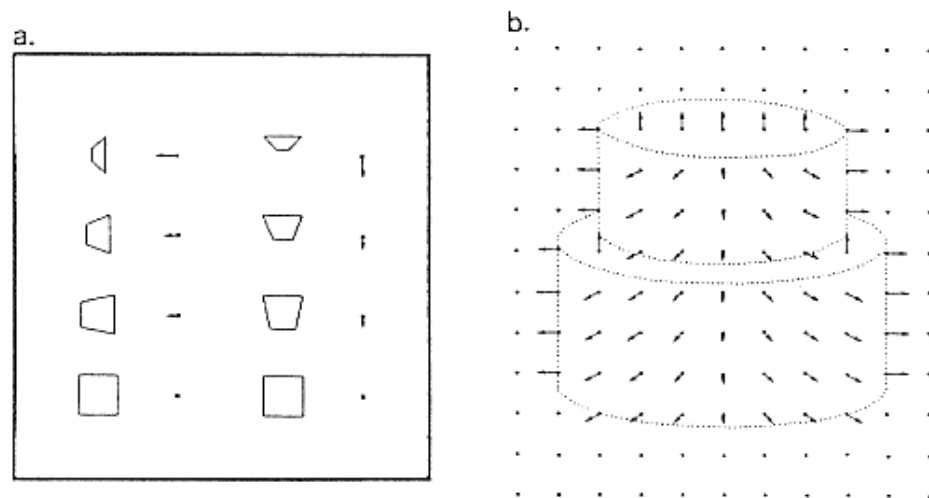
2.5维简图

1. 景物中表面上的各点至少有一个深度的或表面方向的中间表征，或同时有这两者的表征

2. 因为对不熟悉的形状也能感知它的表面方向。所以，得到表面表征是在把景物分解为物体以前

3. 因为所检测到的表面单元的方向与是单眼或双眼观察有关，所以表面表征可能完全由感知过程所支配，而受关于表面实际是什么的专门知识的影响很小

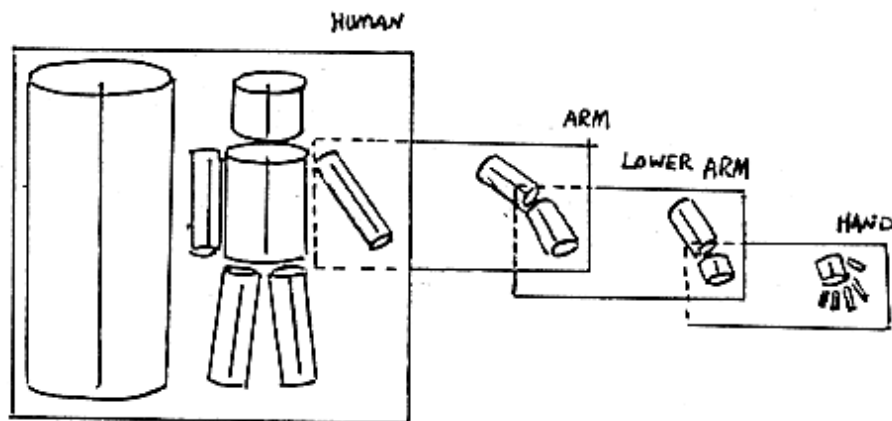
4. 此外，很可能不同的信息源可影响同一个表面表征



2.5维简图

三维模型

- 2.5维简图采用以观察者为中心的坐标系
- 适合于识别的形状表征应该建立在以物体为中心的坐标系中，这样就不随观察的方位而改变
- 有两种以物体为中心的坐标系
 - 一种是用单一坐标系来确定物体上所有点的位置
 - 第二种是分布的坐标系。对物体上的每一个部分或单独的形状特性都有一个局部的坐标系



由三维模型分级组合构成的人体表征系统

Marr计算视觉理论的困难

- 图象中某点的测量值与相应表面点性质之间的关系是不确定和多义性的，若要根据图象数据确定表面点的特性值必须增加附加的约束条件
- 是以逐点计算方式进行处理带来的难以承受的巨大计算量
- 由于缺乏目的性、缺乏高层知识反馈，从而导致三维重建框架不可行，重建算法不鲁棒
- 缺乏利用机器学习的手段来对图像物体进行识别的方法

作业

- 题目1：请观看参考资料中的视频并结合本章介绍的内容，从物体识别的认知原理出发，尝试分析解释下列问题：
 - 1) 挑战过程中哪些因素对识人造成了影响？
 - 2) 挑战者通过何种方式识别不同的人？
 - 3) 这种识人方式与日常生活中的常见识人方式有何差别？

- 参考资料：《挑战不可能》
<http://tv.cntv.cn/video/VSET10/db4e439058f1482a849a794ba3e04405>



作业

- 题目2：分布式表征是物体识别、记忆等认知过程的基本方式，目前也被广泛用于人工智能的理论研究和技术开发。如果需要设计一个可对不同汽车图像（如右图所示）进行精细识别（如识别车型、制造商、车系等）的人工智能系统，请思考如何从每一张图像中提取出与识别任务相关的分布式表征，通过这种方式得到的分布式表征可能存在哪些优点和缺点？

