



第3章 环境感知



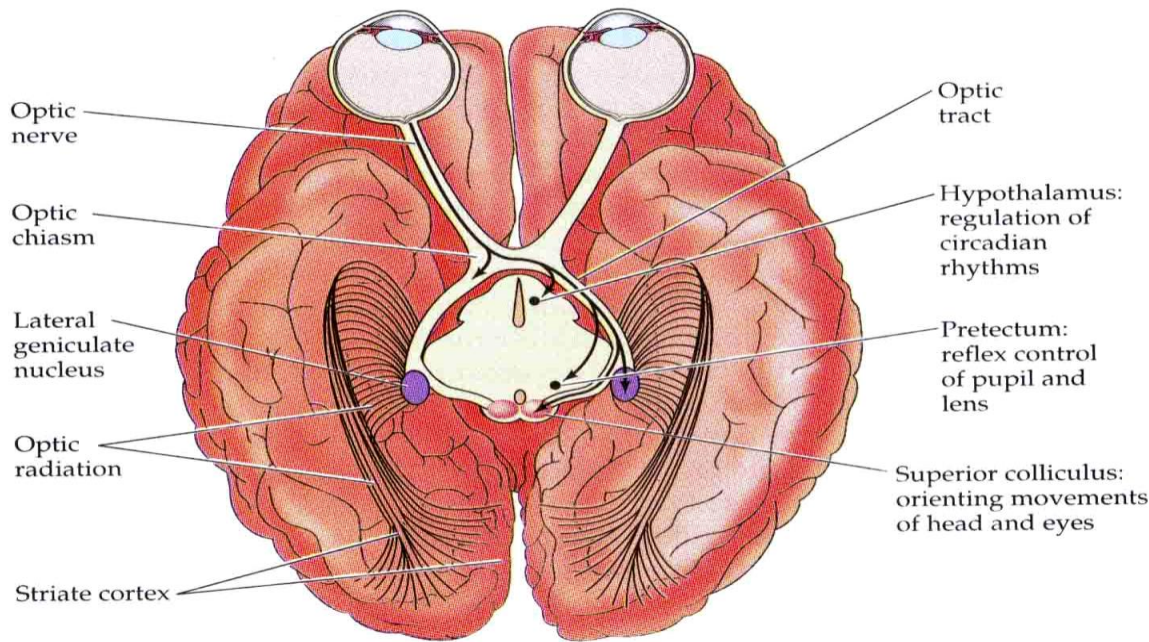
第07讲 视觉原理

在人类心智的机器模拟研究中，恐怕没有哪个方面像在视觉上那样取得如此巨大的实际应用性成功了。不仅如此，视觉感知能力是人类获取环境信息的主要渠道，在人类感知外部世界中起着举足轻重的作用。因此，作为对人类心智模拟的首要方面，也作为机器感知的典型代表，我们就从机器视觉方面来开始我们智能科学研究的讲述。



1、视觉神经通路

人类的视觉系统

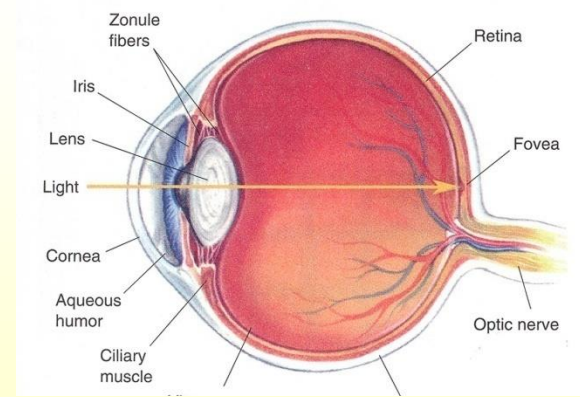
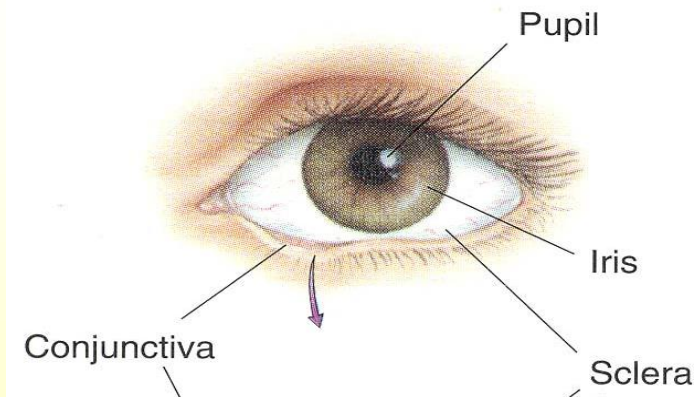


现代认知神经科学研究已经告诉我们，我们这种复杂的**整体知觉能力**完全是大脑视觉皮层中许多离散分布的特异化神经组织功能活动的产物。那么人类的视觉系统又是如何如此复杂的视觉功能的呢？让我们走进人类的视觉神经系统之中，一探究竟。



1、视觉神经通路

在人类视觉系统中，第一个起到收集视觉刺激信息的器官是**眼睛**。如图所示，眼睛是一种精巧万分的球状器官，光线由外射入可以在眼睛的底部内侧形成一个倒立的像，这一点很象我们生活中使用的照相机；所不同的是，照相机的底片是由被动感光材料构成的胶片，而眼睛的“底片”则是大脑神经外围组织的**神经细胞群构成的视网膜**，并具有**主动跟踪获取**和**能动解释信息**的能力。





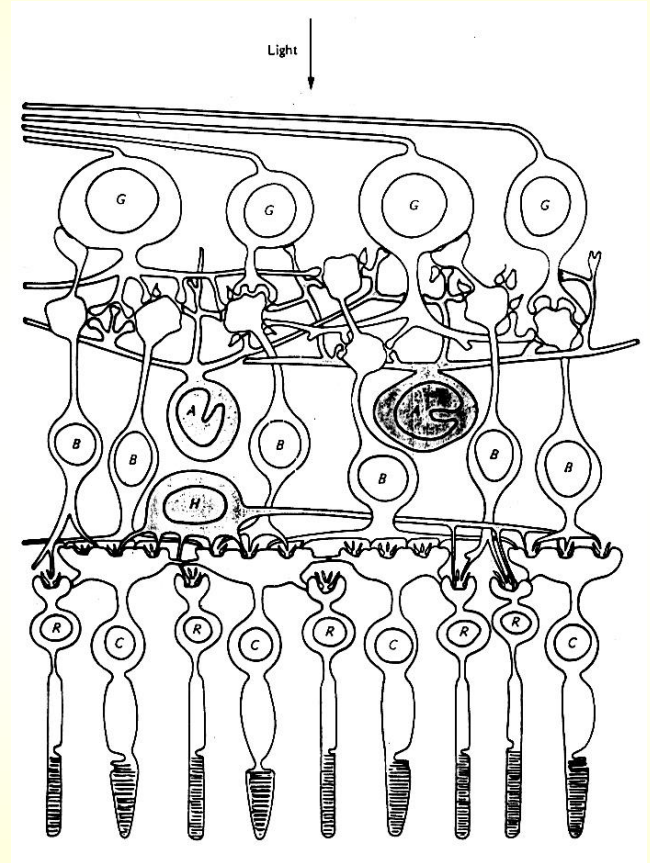
1、视觉神经通路

眼睛中的晶状体是一种可调节焦距的透明物质，而通过头 / 眼运动系统，则可以完成主动跟踪被视物体；另外视网膜中神经细胞的自主活动与相互作用又为能动解释视觉信息提供了可能。视网膜，复盖在眼睛底部的网状薄膜，是由多层神经细胞相互交错连接而成的，如图所示。最底层的那层细胞称为感光细胞，其上附着一定的色素，不匀均地分布在中央凹的周围。



1、视觉神经通路

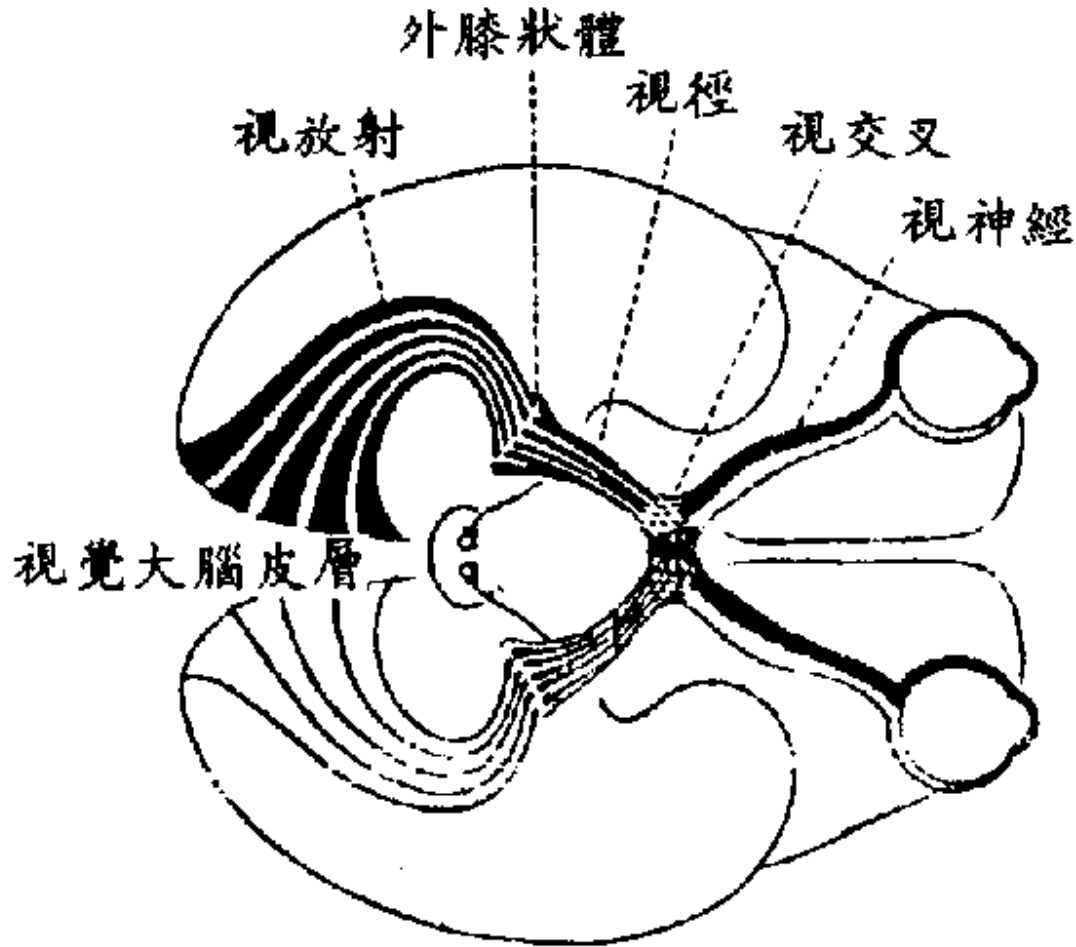
整个视网膜上，约有1亿个感光细胞，它们大体可以分为杆体细胞和锥体细胞二类。杆体细胞主要对光的明暗性敏感，因此在夜间起着主要作用。锥体细胞有三种，分别对红、绿和蓝三种波长的光起反应，而三种锥体细胞的混合反应便可对白天的色彩知觉起作用。



视网膜结构



1、视觉神经通路

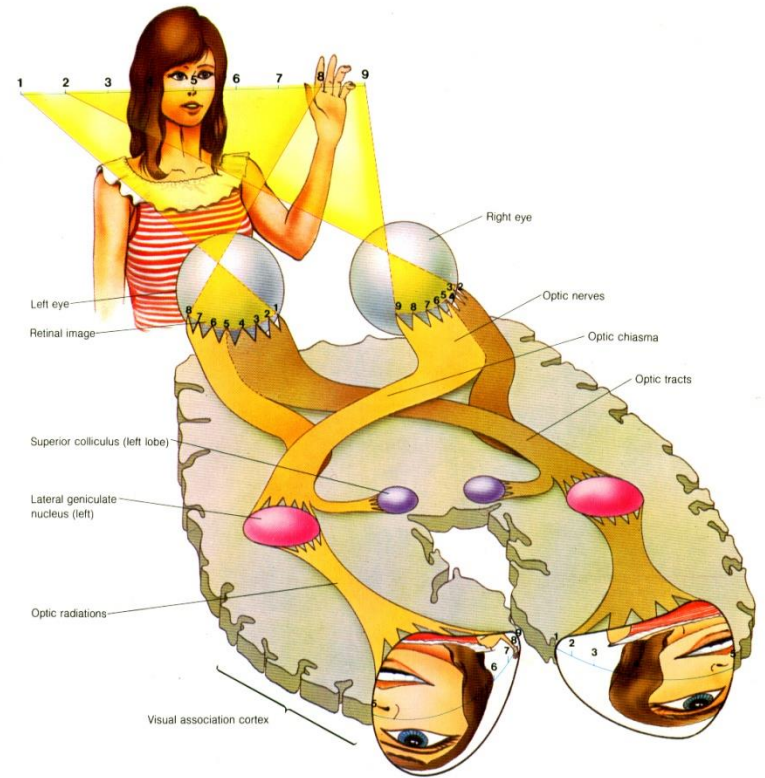


人类视觉通路示意图



1、视觉神经通路

在人类的观察活动中，眼睛的作用实际上就是将外来刺激的光照信号，经过视网膜中神经细胞的加工后，变为送入大脑的某种电脉冲编码组合模式，其代表的就是从外界观察到的景物。然后，又经过了十分复杂行进路线，到达视觉皮层。因此，不管外界的刺激是源自二维的图画还是三维的景物，从视网膜到大脑都有一个相同的将二维形状解释为三维形体的过程。



视觉观察活动过程



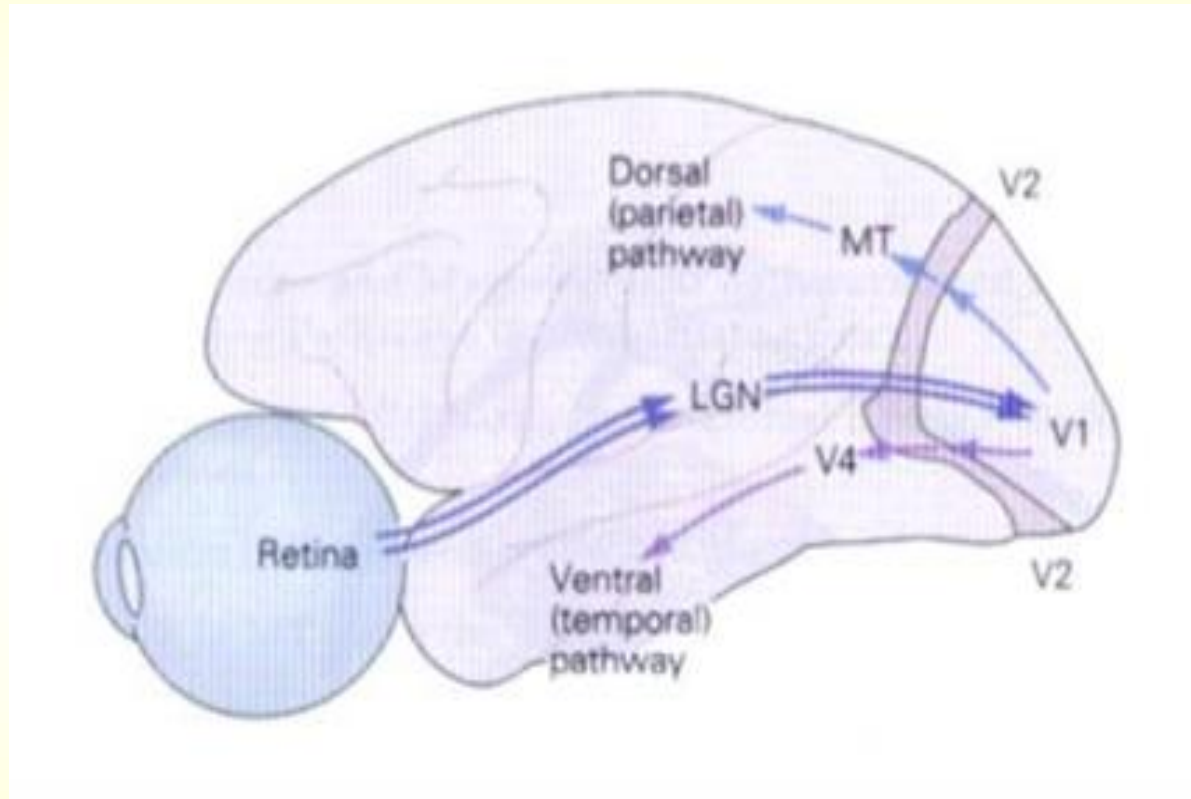
1、视觉神经通路

实际上，情况比这还要复杂，上述视觉加工过程给出的只是视觉通路的主要上行部分；而另外一小部分视束则走向内方，经上丘骨到达上丘和顶盖前区，然后再投射到丘脑枕，换元后投射到视皮层。其主要功能虽非直接与知觉感受有关，但对调节瞳孔、控制眼动等方面起着重要作用，为主动视觉的实现提供了不可替代的手段。另外，所有的视觉通路的投射活动也并非只是单向上行传递的，在各个不同通路阶段，其实普遍存在着同时的下行制约投射和并行制约投射。



1、视觉神经通路

总之，
在视觉通路中，各层次神经细胞普遍是以**相互作用**的方式进行通信的，以给出什么与哪里的视觉任务。



视觉加工的双通路结构



1、视觉神经通路

总之，整个视觉通路中的神经联接和排列方式，决定了具有引起神经细胞活动方式的视觉刺激或辩认。而且，进一步，这种联接和排列方式又是最大限度地体现了视觉通路中不同阶段视野区域的拓扑对应性、反映了对相对强度信息的最佳敏感性以及突出了所有层次神经细胞（群）对视觉信息特征把握的特异性。



2、知觉组织规律

了解了人类视觉的神经机制之后，就我们日常视觉生活而言，通过我们复杂的视觉神经系统又是如何来把握外部视觉环境信息的呢？归纳起来，人类视觉机制涉及**整体知觉**、**主观意念**，**多重选择**和**意图行为**等问题。就感知层面而言，首先遇到的就是其中的整体知觉机制。那么，人类视觉又是如何实现这种整体知觉机制的呢？下面，我们就围绕着视觉整体知觉能力，来开始我们的分析。



2、知觉组织规律

请观看右图，我想人们总是将其看作为一个猫头鹰而不是有字母、线条、圆点部分构成的拼图，这就是整体知觉效果的结果。也就是说，人们在把握视觉刺激时，并不是以自下而上逐个分析来获得生动的形象的，而是同时将整幅图案进行整合感知为有意义的形象。我们具有**整体知觉**能力。



猫头鹰简笔画



2、知觉组织规律

的确，就我们的视觉系统而言，对视觉对象的识认并不是通过部分分析，而是通过全局整体性特征的把握得以完成。对这一认识，著名的瑞士心理学家皮亚杰看得更为深刻，他在《结构主义》一书中指出：“在任何既定情境里，一种因素的本质就其本身而言是没有意义的，它的意义事实上由它和既定情境中的其他因素之间的关系所决定。总之，任何实体或经验的完整意义除非它被结合到结构中去，否则便不能被人们感觉到。”



2、知觉组织规律

其实，早在20世纪初，发轫于德国而成熟于美国的格式塔心理学派就对这种整体知觉及其规律有过全面的论述。

格式塔理论认为，形式知觉产生于该形式部分之间的关系中，而部分特性就人们所能确定的内容来说，依赖于它们所处整体的全部关系，**部分只有在整体中起作用**，离开了整体的部分是没有意义的。基于这样的认识，格式塔心理学派将反映这种**整体知觉的规律**归纳为一些**普遍性**规则，称为**知觉组织律**。



2、知觉组织规律

首先，格式塔理论认为，在整体知觉中，人们的视觉所作出的最基本的区分乃是图形与背景之间的区分，图形在背景中显现出来，反映的正是整体性知觉机制。很明显，图形倾向于轮廓更加鲜明、更好定位、更加紧密和完整，一句话更具有整体统一性；反之，背景就显得不那么整齐规范，没有什么结构可言。

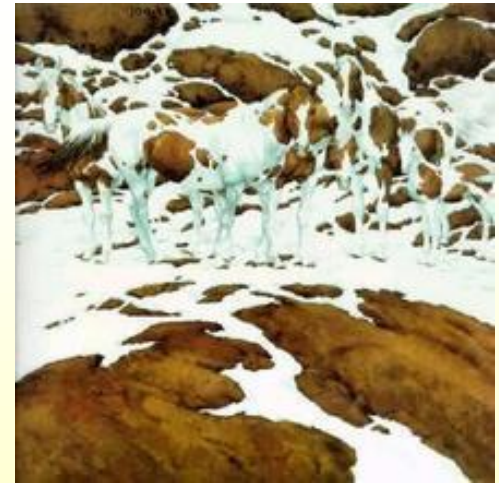


形基关系



2、知觉组织规律

不过，有时图形与背景会相互交替，你可以将图形当作背景，而将背景当作图形，这就是在可以互为形基关系图例中取得的效果。甚至有时你很难正确区分出图形与背景，但这并不影响总是存在图形与背景关系这一总的原则。

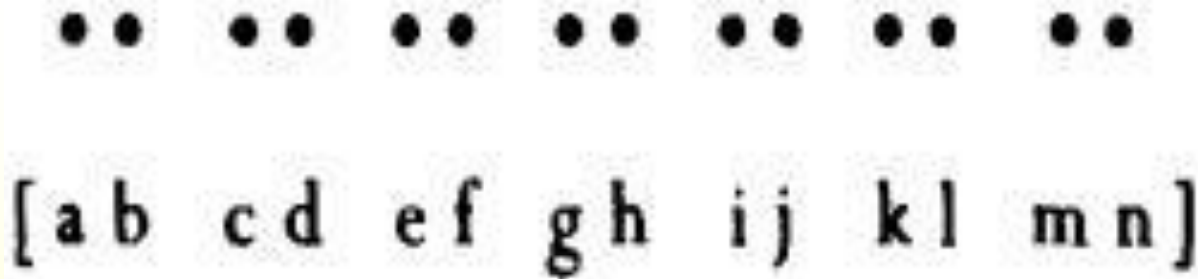




2、知觉组织规律

除了形基规律外，整体知觉的组织规律还其他一些规律，分别是：

a . **接近律**：也就是说，我们的视觉易于根据**部分**彼此之间的**邻近**或**接近**关系而组合起来并因此而得出**整体**形象。



接近律

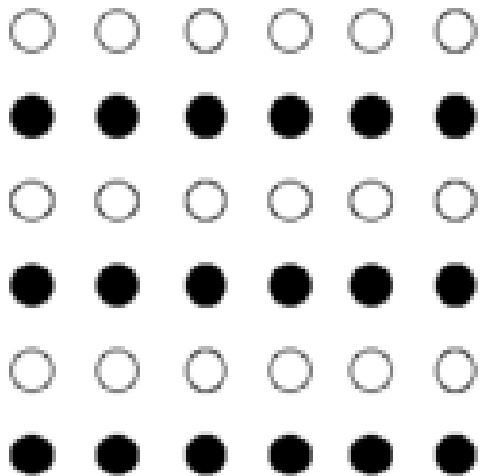


2、知觉组织规律

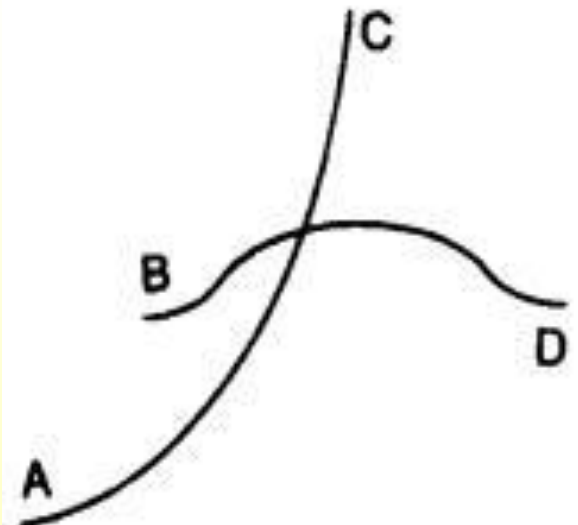
b. 相似律：指具有某种特征（形状、颜色、朝向、动向等）相似的项目，只要不被接近因素掩盖，则倾向于联合在一起。

c. 连续律：如果一套点子中有些点子显得连续或完成一个有规律的系列、或者扩展成一条简单的曲线，这套点子往往易于组织起来。

相似律



连续律

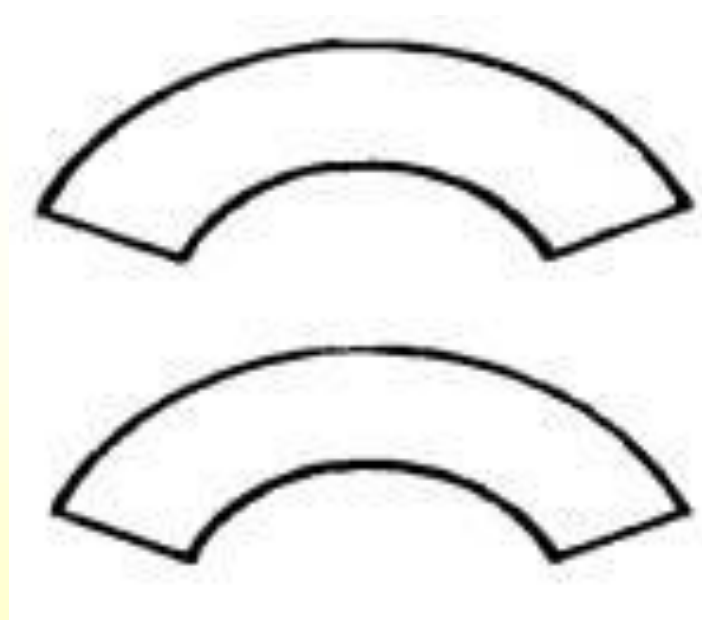




2、知觉组织规律

d. 封闭律：有时，各种视觉组织律会在知觉过程中产生竞争甚至冲突现象，此时最终知觉的结果形象往往取决于哪种因素更为重要了，其中封闭的图案往往有较大的优势。

封闭律





2、知觉组织规律

e. **简单律**：最后，整体知觉的一般规律还告诉我们，在**其它因素相同**时，人们将把**视觉对象**看成有组织的**简单规则图形**，就是说，看作是倾向于**对称、整齐、匀称、惯常**的图形，这就是知觉的简单律。

简单律





2、知觉组织规律

实际上，视觉过程中，**整体知觉**就是把**部分整合**为一个与以往**经验相关联**的完整形象，而**简单律**就意味着这一过程遵循着**经济原则**，得出一个**最简单可能的形象**去与**刺激模式相匹配**。

从这个意义上讲，**整体**知觉一方面体现了**部分只有在整体中才有意义**这一原则；另一方面也体现了**经验形成的完形在知觉中的作用**是以**经济原则**为基础的。



3、视觉感知经验

在漫长的进化历程中，我们发展了精妙的视觉系统，使得我们的视觉能力适应在复杂环境中获取有益于更好生存的视觉信息。因此，对于视觉而言，捕捉环境变化中的意义是第一位的，而没有意义的视觉信息是可以被忽略的，事实上也确实被忽略了。英国心理学家Gregory在《eyes and Brains》一书中指出：“感官接受的是生动的图案，但是我们很少看图案，我们看的是事物。相对来说，图案是无意义的标记的组合，而事物除了具有感官特征以外，还有许多其它特征。”



3、视觉感知经验

那么我们这种习以为常的视觉观看就像摄像机一样，都是现实世界的忠实反映吗？与人类视觉比较，基于摄像机的机器视觉还处于十分初步的水平，无论是形状知觉还是立体感知，机器都会遇到难以克服的困难。那么到底是什么导致了人类与机器之间存在这种不可逾越的**鸿沟**呢？

如果深谙人类的视觉运作规律，就不难发现，这种困难，从深层的机制上讲，一个集中体现的方面，便是人类所具备的**视觉恒常性**和**容错性**，难以为**机械精确**、**永远无错**的算法方式所**描述**。



3、视觉感知经验

有科学家强调，大小恒常性的原理是建立在“感知大小=感知距离×视角”这一公式之上的。比如地平线上的月亮显得比升空的月亮更远，由于月亮的视角基本固定不变，我们就感到它较大。这个就是所谓的埃默特定律。

但距离知觉又是怎么获得的呢？由于距离的知觉在很大程度上是建立在大小恒常性的视觉经验之上的，因此将大小恒常性归结为距离知觉未免又会陷入“先有鸡还是先有蛋”的二难境地。



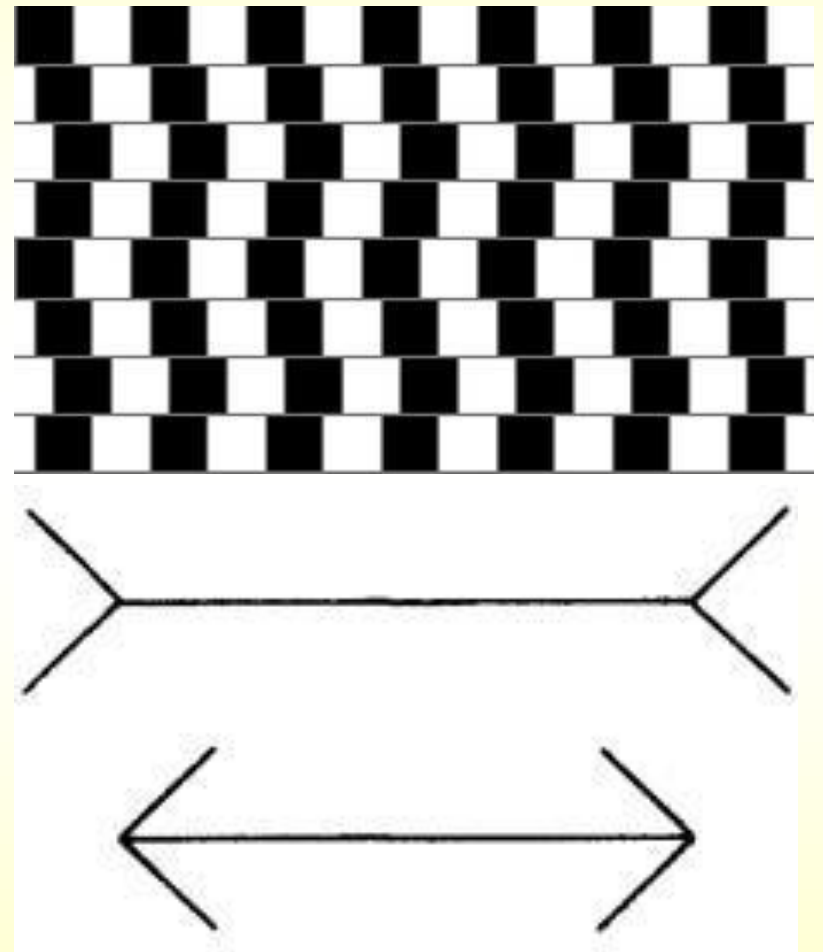
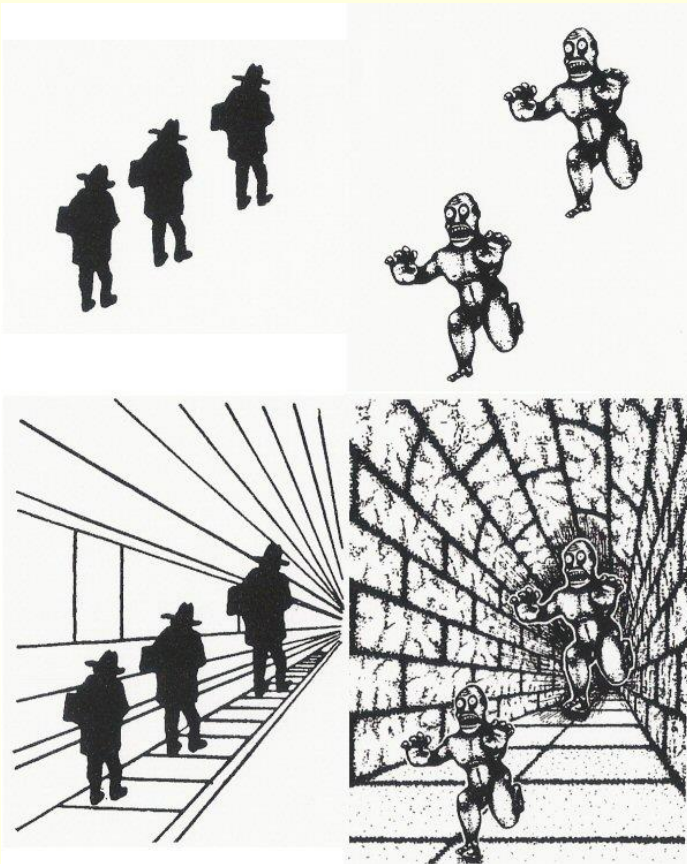
3、视觉感知经验

其实，恒常性主要来自于经验和比较。犹太大学的唐·海勒用老鼠进行了大小恒常性的实验就证实了经验对大小恒常性还是必要的。

物体与所处背景的相对比较则是恒常性的另一重要来源，而恒常性则可以看作是当前物体与已有经验比较的结果。正因为这样，有时我们的视觉会得出错误的结论，产生幻觉或错觉。经验的成见导致错觉，但人类的视觉不会因有错觉而否定经验作用，因为经验是人类赖以生存的基础。这就是所谓的视觉容错性。



3、视觉感知经验

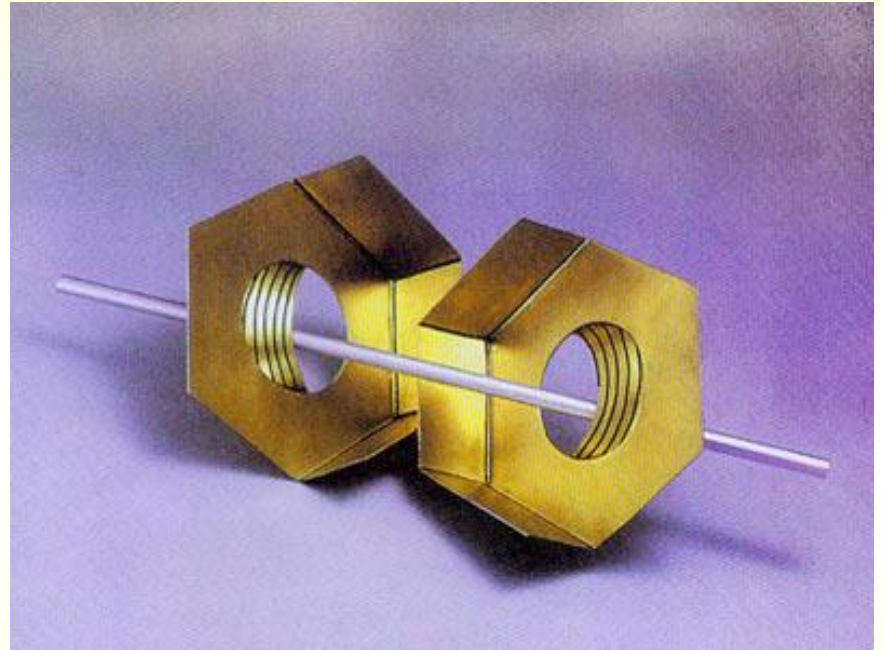


一些恒常性错觉图案



3、视觉感知经验

错觉的**极致**，便是我们可以人为地感知不可能图形。也就是说，在**实际世界中根本不可能存在的形体**，在我们的**知觉世界里竟然是合理的**。图给出了一个常见的不可能图形，你在现实中根本就造不出这样的实体来的。



不可能图形



3、视觉感知经验

实际上，错觉是我们拥有无比精湛的视觉能力的必然代价，如果有一天人类失去了错觉能力，那么也就意味着我们必然地同时也失去了精湛的视觉能力。的确，人总是免不了会犯错误，而这正是人类值得骄傲的心智特性。

机器不会出错，它总是精确地执行人类为它编制好的程序或程序的程序；因此如何使得机器也能拥有视觉恒常性并产生错觉，就成为未来视觉计算研究的一个重要课题。



3、视觉感知经验

对于人类视觉系统这种关注事物意义而忽略无意义的视觉信息的功能，几乎发展到了无以复加的地步。美国心理学家卡洛琳·布鲁墨在《视觉原理》一书的开头就强调指出：“人的头脑从外来的刺激中“毫无节制”地产生着含意。这是一个事实，你所无法逃脱的过程，不管你愿意或不愿意都在发生着的活动。你的头脑不断赋予外界事物以含意——以致有时候这些含意本不存在，完全是你的幻想创造出来的。”



3、视觉感知经验



坐在长椅上的男人



一条觅食的鬣狗



3、视觉感知经验

说明这种“**无中生意**”的极好例子是“坐在长椅上的男人”幻像。如图所示，除了一些在黑背景中的白斑块外，原本什么也没有画，但您却能从这幅图中看出意义来：一位坐在长椅上的男人形象。其实这也不是什么新鲜事，读者或许在生活中早有这样的经验：从天上的飘云中看出各种你所熟悉的造形；在斑驳的墙上看到了人物肖像、各种动作各异的动物形体；甚至还会为自然界形成的奇石怪峰，牵强附会地来冠以石猴、石龟、二郎探母等名号，等等。



3、视觉感知经验



无中生有的视觉示例



3、视觉感知经验

有时，视觉图案的意义往往存在多重选择性，对其含义的确切把握不仅依赖于主观心理定势，而且也依赖于图案本身所提供视觉刺激的微小变化。例如对于图右，你可以看作是一位标致的年轻女子微微转过脸去的侧身像，又可以看作是一位老妇人的侧身像。



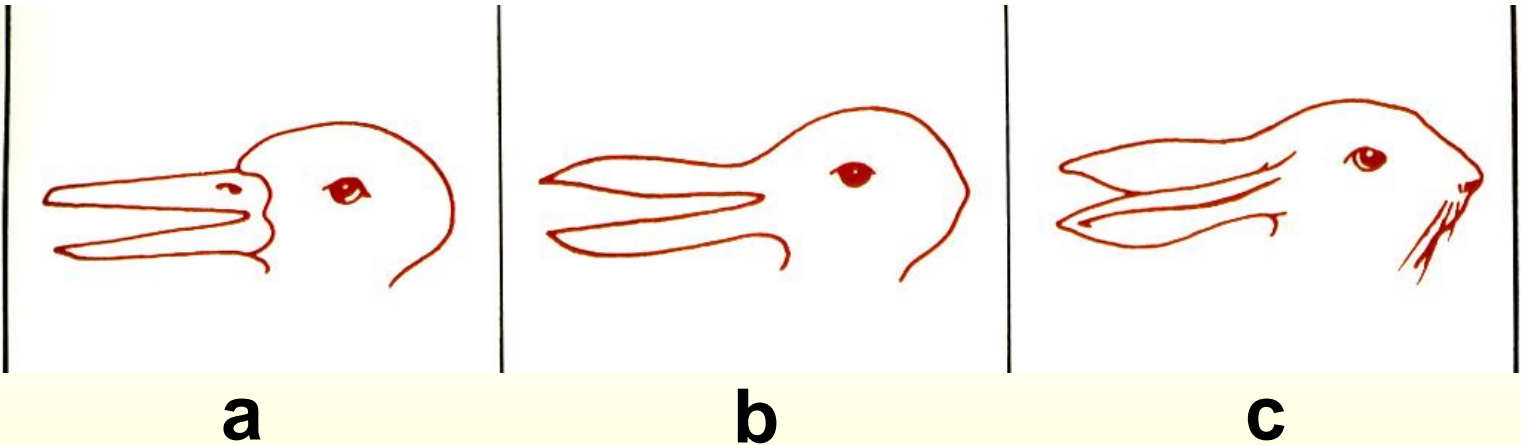
少女与老妇



3、视觉感知经验

但如果给图案**稍稍改变**；那么你就不再会有这种二难选择的困难了，参考下面的兔鸭图（b）。

此时图（a）只能被理解为鸭子而图（c）也只能被理解为兔子。



兔鸭图



3、视觉感知经验

图案多重理解的困难还不仅于此，有时对于歧义图案的多重选择也存在歧义。也就是说不是所有的歧义图案都具有确定的几种含义可供选择，有时连有几种意义本身也是含糊不清的。下一页给出了三种不同类型的歧义图案，

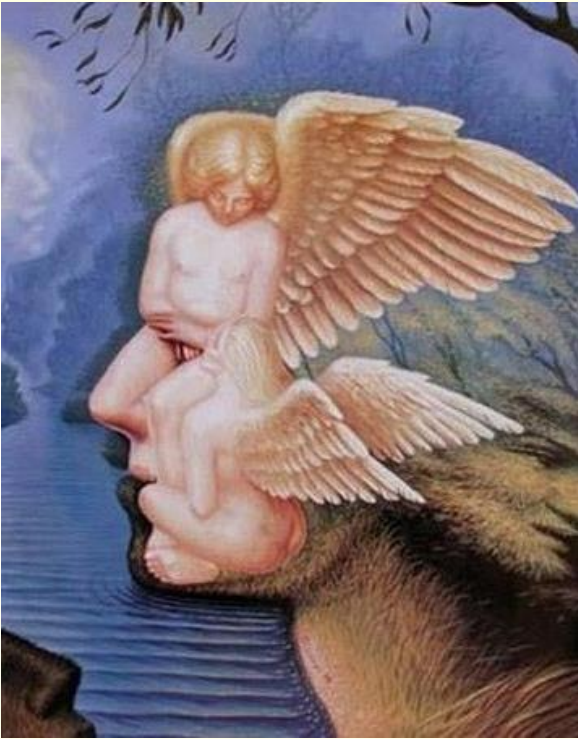
(a) 图属于“含糊不清”型歧义图，你无法知道应该将此图案理解为几个人脸；

(b) 图则属于“多重选择”型歧义图，其具有二个相互独立的并且是排他性的含义，主观理解了其一种含义就不会产生第二种含义；

(c) 图属于“一图双关”型歧义图，就象一语双关的句子一样，其同时给出了二个以上含义。于此可见，歧义图案的理解问题远远要比我们想象的还要复杂，而这一切又不过只是我们全部视觉思维机制的一部分。



3、视觉感知经验



a. 有多少张人脸？



b. 大号演奏家还是美女



c. 既是箭头又是小人



3、视觉感知经验

显然，从上面视觉现象与规律的说明中不难看出，**人类视觉系统是非常复杂的**。我们在得到环境信息刺激后，总能够对环境呈现的含义做出有效的理解，这其中所拥有的视觉机制是十分复杂的。那么，机器呢？那些总是精确地执行人类为它编制好程序的机器，是否也能拥有人类精湛奇妙的视觉能力吗？