

## 第二章 人类视觉

### 2.1 人类视觉简介

了解人类视觉的构成、信息处理过程，对机器视觉研究人员来讲是非常具有启发性和吸引力的。本章从人眼的基本构造出发，介绍视觉通路、视觉感受野及视觉信息的基本处理过程。本章的大部分内容取自文献[寿 1997]。

人眼所能看到的光谱范围，只是电磁辐射波范围的很小一部分，其波长范围从 380 纳米到 780 纳米，这段波长称为可见光谱，如图 2.1 所示。在可见光范围内，不同波长的光给人以不同的色彩感觉，不同强度的光及不同强度分布的光刺激人眼，在人脑中产生不同的光强，颜色，形状等视觉信息。

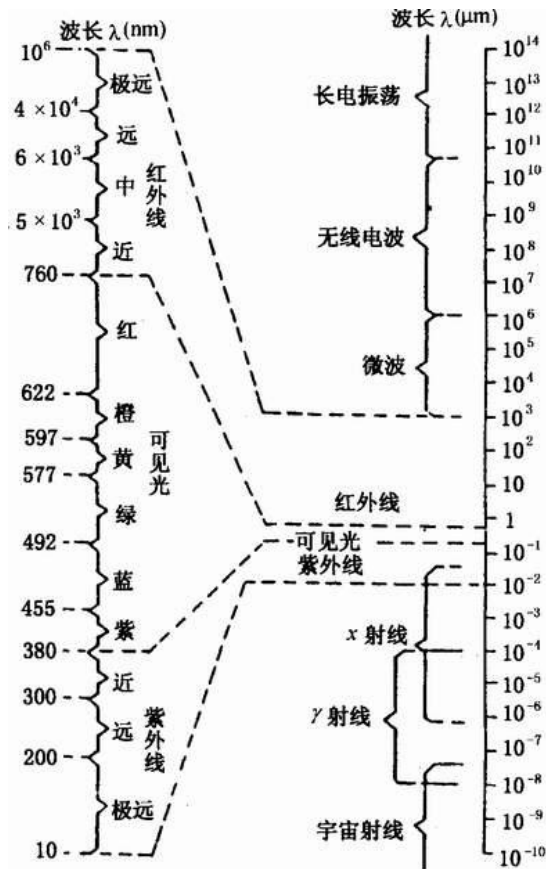


图 2.1 电磁辐射波谱和可见波谱

#### 2.1.1 人眼的构造

人的眼睛是一个前后直径大约 23 毫米的近似球状体，由眼球壁和眼球构成，如图 2.2 所示。眼球壁的正前方，占整个眼球壁面积的  $1/6$  是一层弹性的透明组织，叫做角膜。角膜具有屈光功能，光线经角膜发生屈折进入眼内。眼球壁外层的其余  $5/6$  是白色的不透明膜，叫做巩膜，它主要起巩固、保护眼球的作用。眼球壁的中层包括虹膜、睫状体和脉络膜。脉络膜含有丰富的色素细胞，呈黑色，起着吸收外来散光的作用，消除光线在眼球内部的乱反射。虹膜在角膜的后面，位于晶体的前面，虹膜中央的圆孔叫做瞳孔。虹

膜可以使瞳孔扩大和缩小,睫状体位于虹膜后面,其内部有睫状肌,起调节晶体的作用.眼球壁内层是视网膜(retina)和视神经乳头.视网膜是一种透明薄膜,是眼球的感光部分.视网膜由锥体细胞(cone cell)和杆体细胞(rod cell)两种感光细胞组成,其中,锥体细胞约有650万个,杆体细胞约有1亿个.

视网膜可分为以视轴为中心直径约6毫米的中央区和周边区.中央区有一直径约2毫米(折合6度视角)呈黄色区域,称为黄斑.黄斑中央有一小凹,叫做中央凹(fovea),面积约1平方毫米.人类视觉的中央凹没有杆体细胞,只有锥体细胞,其密度高达每平方毫米150,000.离开中央凹,锥体细胞急剧减少,而杆体细胞急剧增多,在离开中央凹20度的地方,杆体细胞最多.中央凹的锥体细胞密度很高,是产生最清晰视觉的地方.杆体细胞主要是在黑暗的条件下起作用,同时还负责察觉物体的运动.因此,常将锥体细胞称为明视觉细胞,将杆体细胞称为暗视觉细胞.

眼球内包括晶体、房水及玻璃体,它们都是屈光介质.晶体为一扁球形的弹性透明体,位于玻璃体与虹膜之间,睫状体的收缩可改变晶体的屈光力,使外界的对象能在视网膜上形成清楚的影像.

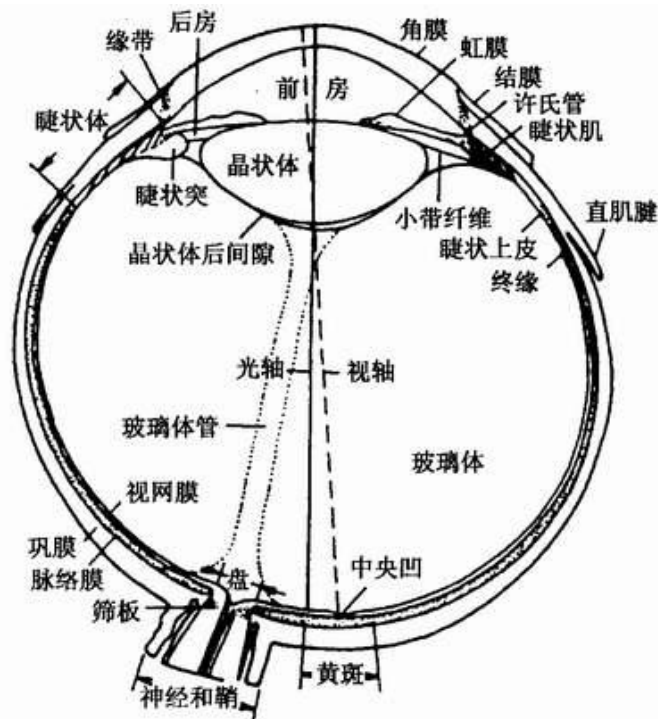


图 2. 2 眼睛 (右眼) 解剖图[Wall 1942]

## 2. 1. 2 视觉通路

视觉通路见图 2. 3. 到达视网膜的光线经杆体和锥体细胞转换为神经信号,并经视网膜中的神经节细胞(ganglion cell, GC)加工,传出视网膜.经神经节细胞加工的神经信号,经过视交叉部分地交换神经纤维后,再形成视束,传到中枢的许多部位,其中包括丘脑的外膝体或外膝核(lateral geniculate nucleus, LGN)、四叠体上丘、顶盖前区和皮层等.上丘与眼动等视反射有关,顶盖前区与调节反射、瞳孔反射有关,外膝体和视皮层都直接与视觉知觉有关.神经节细胞轴突在外膝体换神经元后,由外膝体神经元直接经放射到视皮层,这时视束的大部分纤维去向,称为视觉的第一视通路.视束的一小部分

纤维走向内方，经上丘臂，到达上丘和顶盖前区。上丘浅层神经元再透射到丘脑枕换元后，再透射到视皮层，上丘还有纤维直接透射到视皮层。由于这条通路不经过外膝体，故称为第二视觉通路。

人们普遍认为，视神经信号主要是通过视觉第一通路到达视皮层，因此，第一视觉通路在视觉过程中起着主要的作用，而第二视觉通路的作用及其微小，可以忽略不计，如图 2.4 所示。

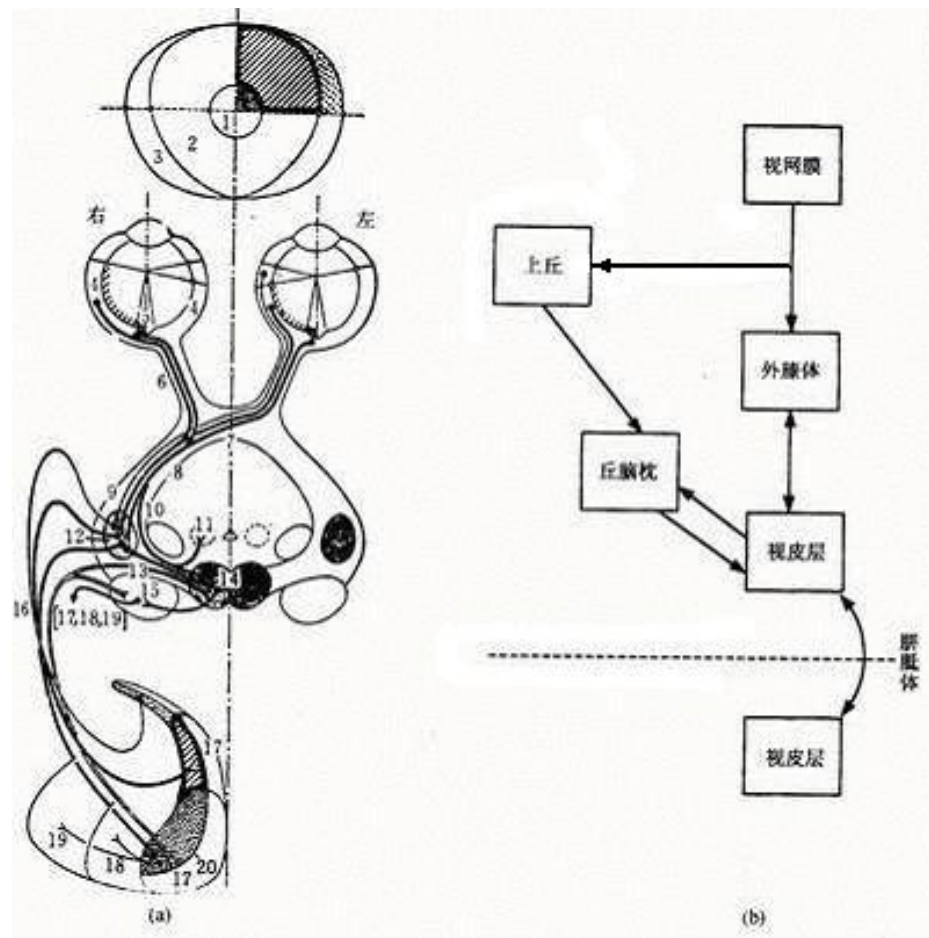


图 2.3 人类视觉通路[Nieuwenhuys1979]

## 2. 1. 3 视觉信息处理过程

基本视觉信息包括：亮度，形状，运动，颜色，深度知觉等，其中亮度是最基本的视觉信息，因为没有亮度就谈不上什么形状、运动、颜色等视知觉。亮度是一种外界辐射的物理量在我们视觉中反映出来的心理物理量。物体的形状主要是由物体在视觉空间上的亮度分布，颜色分布或运动状态不同而显示出来的。雪地上的一张白纸，荒野中潜伏的狼，都比较难以觉察。相反，雪地上的一张彩纸，绿草地上等一朵红花，荒野中奔跑的狼，都比较容易发现，这是由于物体形状因亮度、颜色、运动等因素而突显出来，易于被人眼分辨出来。视觉系统对运动目标十分敏感，特别是低等动物更是如此，如蛙类的视网膜对运动特别敏锐，而对静止的目标却视而不见。颜色知觉是一种主观感觉，目前只在感受器细胞水平和心理物理的宏观水平上得到了较深入的研究。研究表明

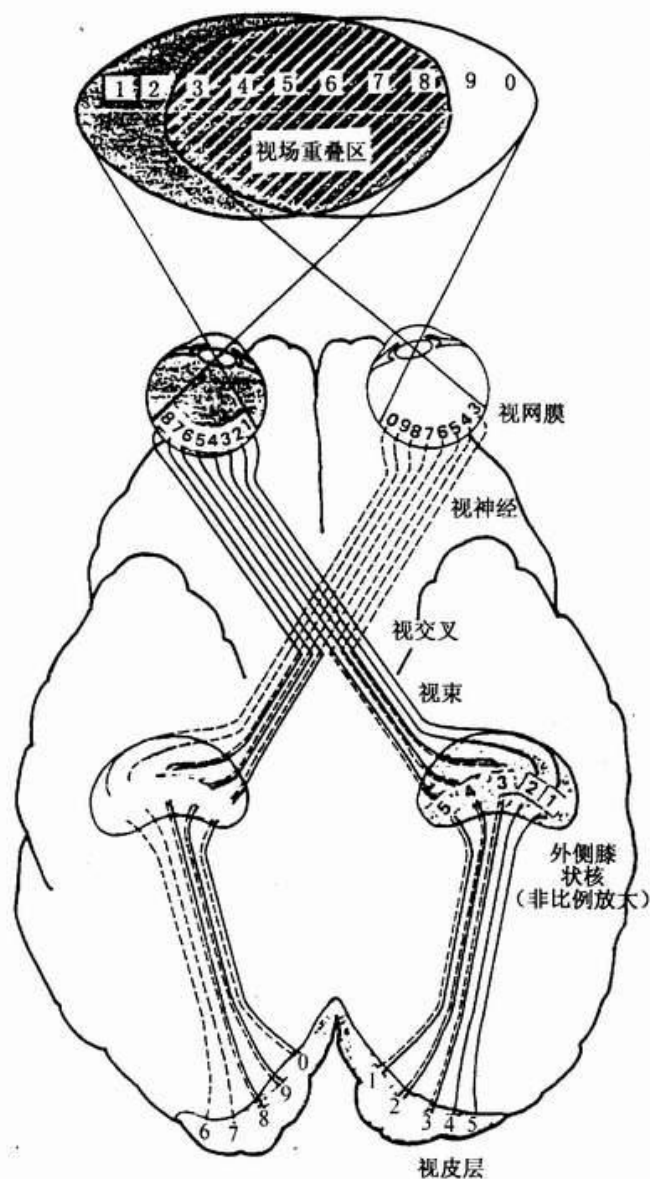


图 2.4 视觉通路简化模型(Lindsay&Norman1972)

[Wald1964]，猴和人的视网膜中，含有三种不同的锥体细胞，每一种锥体细胞对不同的光谱，其敏感性也是不同的。三种锥体细胞对光谱的敏感峰值分别在 430 纳米，540 纳米和 570 纳米。这三个峰值段正对应着光谱中的红、绿、蓝区域。实验进一步表明，当三种颜色按一定比例同时刺激人眼时，会产生各种颜色感觉。其中有一种比例会使得颜色感觉完全消失，只有亮度感觉。这就是所谓的色觉三变量性(tri-variance of color vision)，说明颜色只取决于三个基本的输入量，这也是色觉三基色原理的基础。

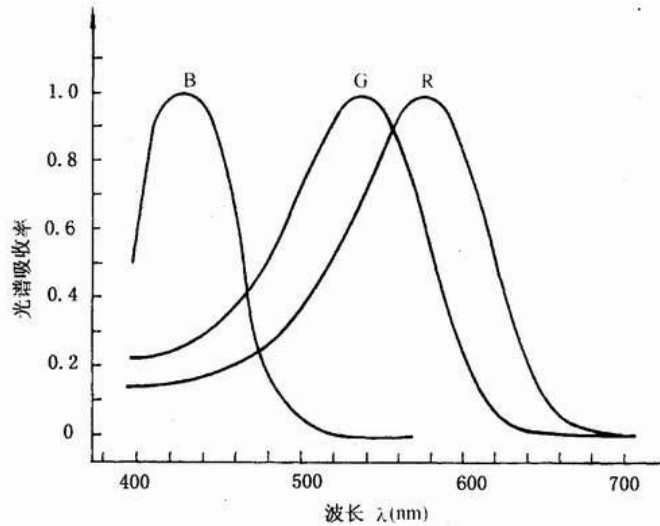


图 2.5 三种感受器的光谱敏感示意图[Wald 1964]

人类的视觉不仅要识别物体的形状和颜色，而且要随时地作用于物体，例如，伸手拿一本书，躲开汽车或障碍物，把足球踢入球门等，这一切活动都需要判断我们与被作用物体的距离。立体知觉就是指这种判断物体距离或深度的感觉。正常的双眼视觉都可以提供高度的立体感。外界目标在视网膜上的象是二维的，而且同一物体在左右眼的视网膜上的成像有着微小的差异，比如，用你的手轮流遮挡你的左、右眼，会发现同一个物体在左右眼中的位置是不同的。实际上，这种不同为立体视觉提供了最基本的信息——视差(disparity)。[Julesz 1960]首次使用随机点立体图对作为刺激研究了立体视知觉，如图 2.4 所示，这种刺激图形排除了所有单眼视觉引入的第二视觉线索，而只保留了视差信息。受试者在观察随机立体图对时，用左右眼分别观看其中一个图几秒钟，当两个图融合后，便会感受到有一个方形平面从背景中突显出来。除了双眼视差提供深度信息外，还有许多单眼的信息可以产生深度信息的估计，如物体的重叠、透视(近大远小，近清晰远模糊，近亮远暗)、明暗、纹理及运动。当然这种深度估计在质量和感觉上仍无法与视差立体知觉的深度相比。

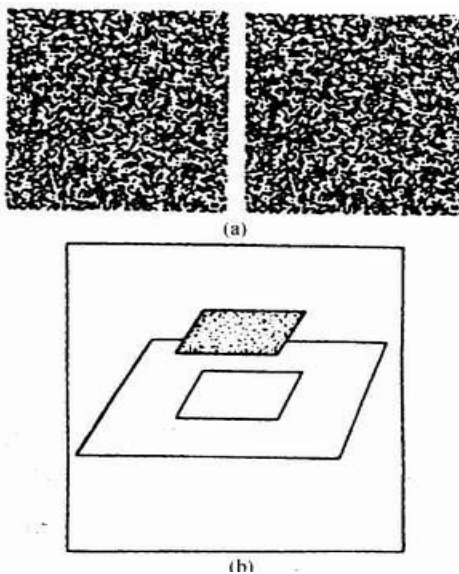


图 2.6 立体视觉测试图[Julesz 1960]

## 2. 2 感受野

在视觉通路上，视网膜上的光感受器(杆体细胞和锥体细胞)通过接受光并将它转换为输出神经信号而来影响许多神经节细胞、外膝状体细胞以及视觉皮层中的神经细胞。反过来，任何一种神经细胞(除起支持和营养作用的神经胶质细胞外)的输出都依赖于视网膜上的许多光感受器。我们称直接或间接影响某一特定神经细胞的光感受器细胞的全体为该特定神经细胞的感受野(receptive field)。

1953 年 Kuffler 首次阐明猫的视网膜神经节细胞(GC)的感受野在反应敏感性的空间分布是一个同心圆[Kuffler 1953]，即感受野一般是由中心的兴奋区域和周边的抑制区域构成的同心圆结构，称为 On-型感受野，还有一类感受野是由中心抑制和周边兴奋区域的同心圆构成，称为 Off-型感受野。图 2. 5 所示的是猫视网膜 GC 的感受野及其反应形式，其中(a)为 On-型感受野，(f)为 Off-型感受野，(b)为小光点单独刺激 On-型感受野中心时，细胞发放频率增加的示意图。(c)表明当用面积正好覆盖 On-型感受野中心的光斑刺激感受野中心时，可以得到 GC 的最大兴奋型反应。(d)表示当用面积正好覆盖 On-型感受野周边的光斑刺激周边时，得到 GC 的抑制型反应。(e)表示当用大面的弥散光照射 On-型 GC 时，它们倾向于彼此抵消，得到较弱的兴奋型反应。Rodieck 于 1965 年提出了同心圆拮抗式(homocentric opponent)感受野的数学模型，如图 2. 6 所示，它由一个兴奋作用强的中心机制和一个作用较弱但面积更大的抑制性周边机制构成[Rodieck 1965]。这两个具有相互拮抗作用的机制，都具有高斯分布的性质，但中心机制具有更高的峰敏感度，而且彼此方向相反，故称相减关系，又称高斯差模型(Difference of Gaussians, DOG)。

神经节细胞对落入其感受野内的对比度有选择性的敏感性要比落入整个感受野上的总光强的信息更加敏感。图 2. 7 为一个对比边刺激物位于 On-型感受野不同位置时，其 GC 反应放电频率变化。图 2.7(b)表明，当刺激物从左移到感受野的中心区域边界且没有覆盖中心区域时，GC 对这一位置的对比边缘反应要比同样光强的弥散光覆盖整个感受野(图 2.7(a))所引起的反应更强。GC 感受野的这种对明暗对比边缘特别敏感的性质，可以解释心理物理学中著名的马赫带效应(Mach Band)。马赫是 19 世纪著名的奥地利物理学家，他在观察一个亮度渐变的边缘时，发现主观感觉在亮度的一端呈现一个特别亮的亮带，在暗的一端呈现一个特别暗的暗带，如图 2. 8 所示，这就是马赫带。图 2. 8 的上方给出了 On-型感受野的位置示意图，位置 3 的 GC 反应最为强烈，它对应于图 2.7(b)的位置，位置 4 相当于图 2. 6(a)的位置，因此，强度仍然较强，但不是最强烈的反应，位置 2 对应于图 2. 7(c)的位置，反应最小，且是抑制性的，即低于黑暗中无刺激时的 GC 自发放电水平。位置 1 整个 GC 感受野处于黑暗中，因此只有 GC 本身的自发放电水平，但比位置 2 的放电水平要强一些。

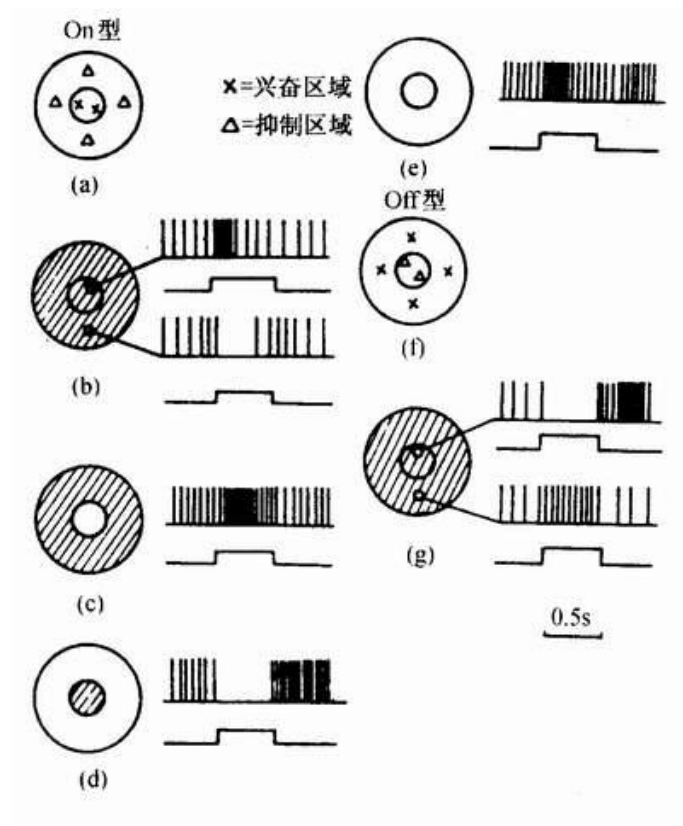


图 2.7 猫视网膜神经节细胞的感受野及其反应形式

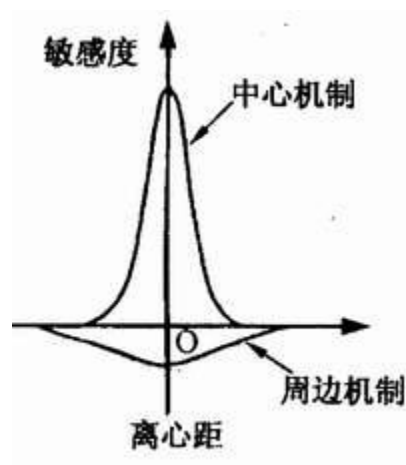


图 2.8 Rodieck 的视网膜感受野神经节细胞数学模型

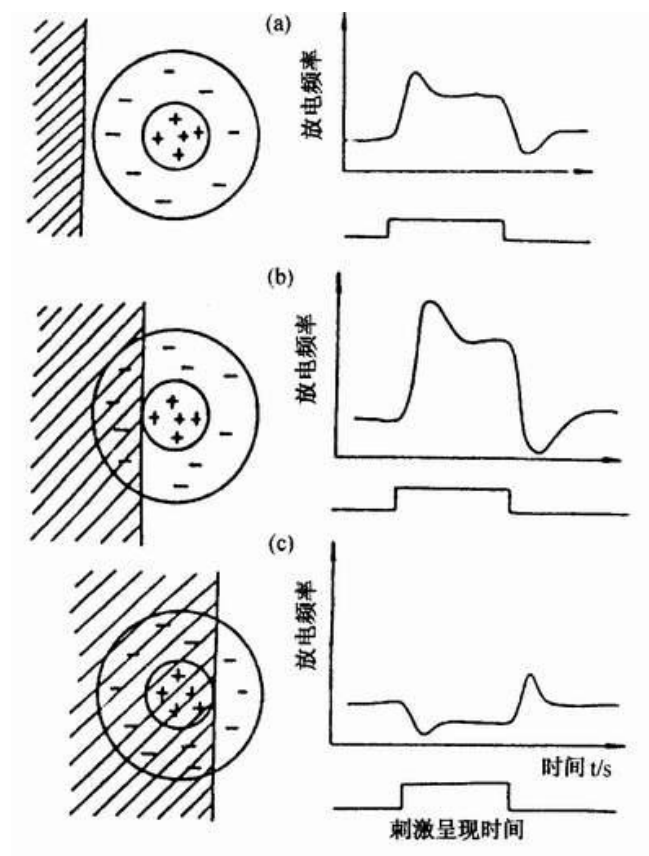


图 2.9 亮暗刺激物位于 On-型感受野不同位置时，其 GC 反应放电频率图  
[Levine 1981]

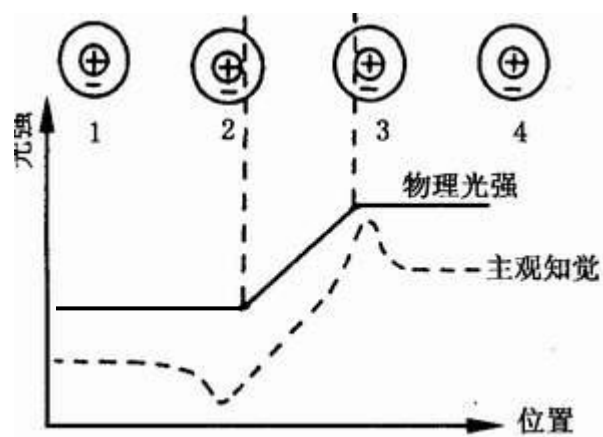


图 2.10 马赫带现象及其生理基础

在视觉皮层中，除了同心圆状的感受野外，还有更复杂的感觉野响应特性，它们对其感受野中的特定方向的线段敏感。所有视觉通道上的神经细胞，按其感受野在一个视



网膜或两个视网膜上，可分为单眼神经细胞和双眼神经细胞。所有神经节细胞、外膝体细胞和简单细胞都是单眼的，复杂细胞约有半数单眼，半数双眼。双眼细胞又可进一步分为右眼主导、左眼主导和双眼均衡的三种。

## 2. 3 视觉信息的多层次并行处理

许多神经科学家对视觉信息的并行分块处理进行了深入的研究。所谓的并行分块处理是指不同视觉性质的视觉信息成分按不同的神经通道预处理并输入视皮层，由不同性质的皮层细胞分别进行分析处理。以英国科学家 Zeki 为代表的神经科学家认为，人类视觉系统使用更加精巧的策略或办法来统一不同性质的信息，即在几个不同水平上相互作用来多级地处理复杂的视觉信息，达到感知周围多彩生动的视觉世界[Zeki 1993]。

### 2. 3. 1 视觉信息的多层串行处理机制

Hubel 和 Wiesel 首先提出视觉信息是多级串行处理的[Hubel 1959]。由视觉通路可以看到，视网膜、外膝体、视皮层构成对视觉信息处理的多级串行处理，特别是视皮层，表现出更为复杂的多级分层。

视网膜由三层细胞组成。从最外到最内为感受器细胞层(receptor cell, RC)，双极细胞层(bipolar cell, BC)和神经节细胞层(GC)，GC 的轴突形成视神经。这三层的每一层均包含有不只一类细胞，各层之间以及一层之内的细胞形成广泛的联系。这里需要指出的是光线传递方向与视网膜信息传递方向相反，即光线首先到达神经节细胞层，最后到达感受器细胞，称这种视网膜为倒转视网膜(inverse retina)

视网膜神经节细胞轴突形成视神经，经视交叉和视束到达外膝体。外膝体属丘脑，是眼睛到视皮层通路的中继站。猴的外膝体细胞在组织上分为六层，各层之间几乎没有任何细胞。而来自两个视网膜的纤维分别按照一定的规律投射到外膝体各层。外膝体对信息的处理是并行的。

### 2. 3. 2 视觉信息的并行处理

#### (1) X、Y 和 W 通道

空间拮抗(spatial opponent)是感受野的一种基本作用性质，但感受野还有一些其它性质，其中最重要的是它的反应时间特性和线性特性。比如，猫的神经节细胞可按其反应的空间—时间总和性质而划分为两类，一类细胞反应的空间总和性质大体符合 Rodieck 模型，即它们的感受野的兴奋和抑制作用可以线性相加，称为 X 细胞，另一类神经节细胞的空间和性质是非线性的，称为 Y 细胞。外膝体的神经元也可按其空间总和性质划分为 X 和 Y 细胞，并且在传递信息过程中，X 型神经节细胞总是与 X 型外膝体神经元发生联系，Y 型神经节细胞总是与 Y 型外膝体神经元发生联系，而 X 型外膝体神经元多数传至视皮层简单细胞和超级复杂细胞，而多数 Y 型神经元只传给视皮层复杂细胞。由此可见，视觉系统内存在一个 X 和 Y 通道，它们在功能上表现出 X 细胞的感受野可能与空间信息的检测与传递有关，而 Y 细胞的感受野可能与时间信息的检测与传递有关。除了 X 和 Y 细胞外，人们在猴、猫视网膜上还发现一种称为 W 型的神经节细胞，其感受野与 Y 细胞大小相仿，但轴突直径特别细，因而动作电位在其轴突上传导速度也最慢。W 细胞的轴突主要传至中脑上丘部，是控制眼球运动的。

#### (2) On—型和 Off—型通道

在视网膜上，On—型和 Off—型细胞是一种均匀镶嵌式的排列，其总数基本相等，

而在 LGN，它们开始呈现一定程度的分离，实验充分证明，On—型通道和 Off—型通道在 LGN 到视皮层是充分地平行分离的。比如，在猴视网膜水平细胞、双极细胞处用药物选择性地阻断 On—型通路，可以取消神经节细胞，LGN 和视皮层的 On—型反应，但对 Off—型细胞的反应和视皮层细胞方位、方向选择毫无影响。

### (3) 左眼、右眼通道

在视网膜、外膝体及视皮层构成的视觉通道上，外膝体内的每一个细胞均只接受单眼输入，双眼信息并不相混。视皮层中的细胞多为双眼输入细胞，但大多数细胞总是呈现对某眼输入刺激的反应占优势。视皮层的左、右眼优势柱，与其对应的细胞产生连接。双眼视差信息是立体视觉产生的基础。实验证明，猫的不同视差信息处理是经 X，Y，W 通道分别处理的。

### (4) 空间频率通道

视网膜、外膝体的 X 和 Y 型细胞都分别对高、低空间频率反应有显著差异。光栅适应性心理学实验有力地支持视觉信息是按空间频率强弱不同的多通道进行分析处理的。现已经有充分的实验证据表明猫和猴的皮层 17 区均存在着空间频率功能柱，在同一柱内的细胞最优空间频率相同。具有高空间频率的功能柱集中在 17 区中央皮层，具有较低空间频率的功能柱向 17 区中央的周边区域扩散分布。

### (5) 颜色信息处理通道

在视网膜感受器细胞水平，颜色信息即被三种光谱敏感性不同的(红、绿、蓝)锥体细胞所分别处理，侏儒型和平底型双极细胞处理颜色信息，而杆体双极细胞无色觉。在神经节细胞水平，猴 B 型节细胞处理信息，故接受其平行投射的外膝体小细胞层司色觉信息处理，而 A 型细胞无色觉，故平行地接受其输入的 LGN 大细胞层亦无色觉。经外膝体小细胞层细胞所处理的颜色信息，被  $V_1$  区(17 区)的细胞色素氧化酶染斑点内皮层细胞进一步地加工处理。

### (6) 空间方位信息通道

自 Hubel 和 Wiesel 20 世纪 50 年代末开始的视皮层细胞研究工作以来，人们一直认为视觉方位敏感性是视皮层细胞的独有的功能。20 世纪 80 年代初，Levick 等证明神经节细胞也具有方位敏感性，不过其敏感性很弱[Levick19??]。寿天德和 Leventhal 证明约 80% 的外膝体中继细胞具有弱但确实存在的方向敏感性，这些外膝体神经元具有与视网膜神经节细胞相似的、向心的最优方位分布规律，并且最优方位的细胞在外膝体内部已经聚集在一起[寿 19]。因此，视网膜、外膝体和视皮层形成一个方位信息通道。

### (7) 运动方向信息通道

视网膜节细胞中的少数 W 型兴奋—抑制中心细胞对运动刺激有方向敏感性。寿天德等人认为猫视网膜 22% 的 X 型和 34% 的 Y 型节细胞具有显著的方向敏感性[寿 1995]。Thompson 等人证明猫外膝体约有 1/3 的 X 和 Y 细胞具有方向敏感性[Thompson 1994]。这些方向敏感性 X 和 Y 细胞可能为强的皮层方向选择性形成作出贡献。

## 2.4 视觉信息的集成和反馈

现在知道猴大脑皮层确定与视觉有关的区域约有 35 个以上，它们既平行又分级处理着各种不同的视觉信息。解剖学证据说明 35 个视觉皮层区之间存在广泛的交互投射。英国科学家 Zeki 提出关于视觉皮层信息传递集成的“多级同步集成”假说[Zeki 1993]，用于解释视皮层各特殊区域之间的相互作用机理，以及由此实现思想和行为的集成。Zeki 指出，视觉皮层信息的集成不是以部位上的会聚为主，而是一种多级集成，在几个不同水平上的相互作用来实现。支持这

种“多级同步集成”假说的证据很多。例如，色觉区  $V_4$  与运动区  $V_5$  均有投射到顶皮层的内顶沟；顶皮层区和颞皮层均有输出到额叶，但它们的输入在第三个区域内空间上很少重叠，而且各自有其自己的领地； $V_4$  和  $V_5$  也都投射到颞叶皮层，但根本就不存在直接的重叠。

脑可能使用更加精巧的策略或办法来集成不同性质的信息，即在几个不同水平上相互作用来多级地处理复杂的视觉信息，达到感知周围多彩生动的视觉世界。Zeki 认为，更高级视觉皮层向  $V_1$  和  $V_2$  区“再进入”反馈输入信息，对于某些概念的形成具有十分重要的作用。正如前述，视觉皮层之间几乎毫无例外地存在着交互投射，甚至丘脑外膝体也接受大量的视皮层下行投射。以  $V_2$  区为例， $V_5$  区  $\rightarrow V_2$  区是弥散于整个  $V_2$  区（虽然在宽带部分最为密集），所以不但宽带而且窄带、亮度之间都有来自  $V_5$  的投射纤维，因此  $V_5$  区可以影响  $V_2$  区窄带内细胞的颜色信息处理；颜色信息处理区  $V_4$  对  $V_2$  区内所有的带区均有弥散性投射，从而对  $V_2$  区各带区内细胞投射到  $V_5$  和  $V_3$  区的信息处理产生影响。这样，返回性的“再进入”信息通路不仅返回到原有视区的输入神经元所在的亚区，而且分布到整个前级视区，因此是非模块、不易定位的和多级弥散性的，从而将分工明确的高级视觉皮层内的形状、颜色和运动信息联系起来。总之，脑内整个视知觉是由几个视皮层和通路同时活动的产物。

以上是人类近 40 年来对高级生物视觉的研究成果，这些研究成果给我们研究和设计机器视觉系统提供了很好的模型或生物支持。这里需要指出，对生物视觉信息处理过程的研究难度十分巨大，目前的研究成果仅仅是生物视觉信息处理机理的及其微小的一部分。