生物视觉机理简介

胡占义
中国科学院自动化研究所
模式识别国家重点实验室

奇妙无穷的生物视觉系统

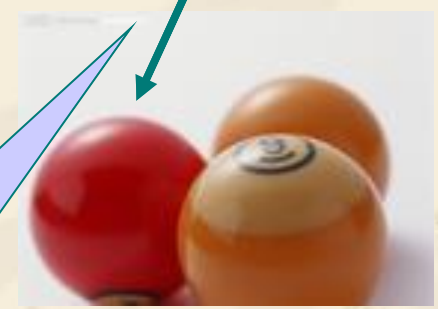
Magic Biological Vision system

人类奇妙无穷的物体识别能力

- 个体辨识（identification)
  - 要求发现细微差别：识别双胞胎
- 种类识别 （categorization)
  - 不相似的物体分为一类，如苹果和葡萄：水果类；相似的物体分为不同类：苹果和台球

个体辨识

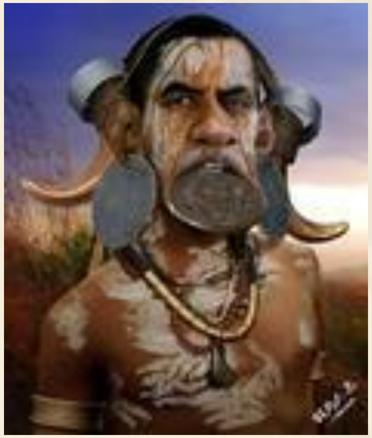
水果



看上去更相似

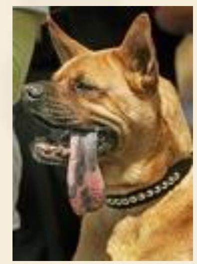
种类识别

对漫画识别的困惑



大脑是如何表达物体的？





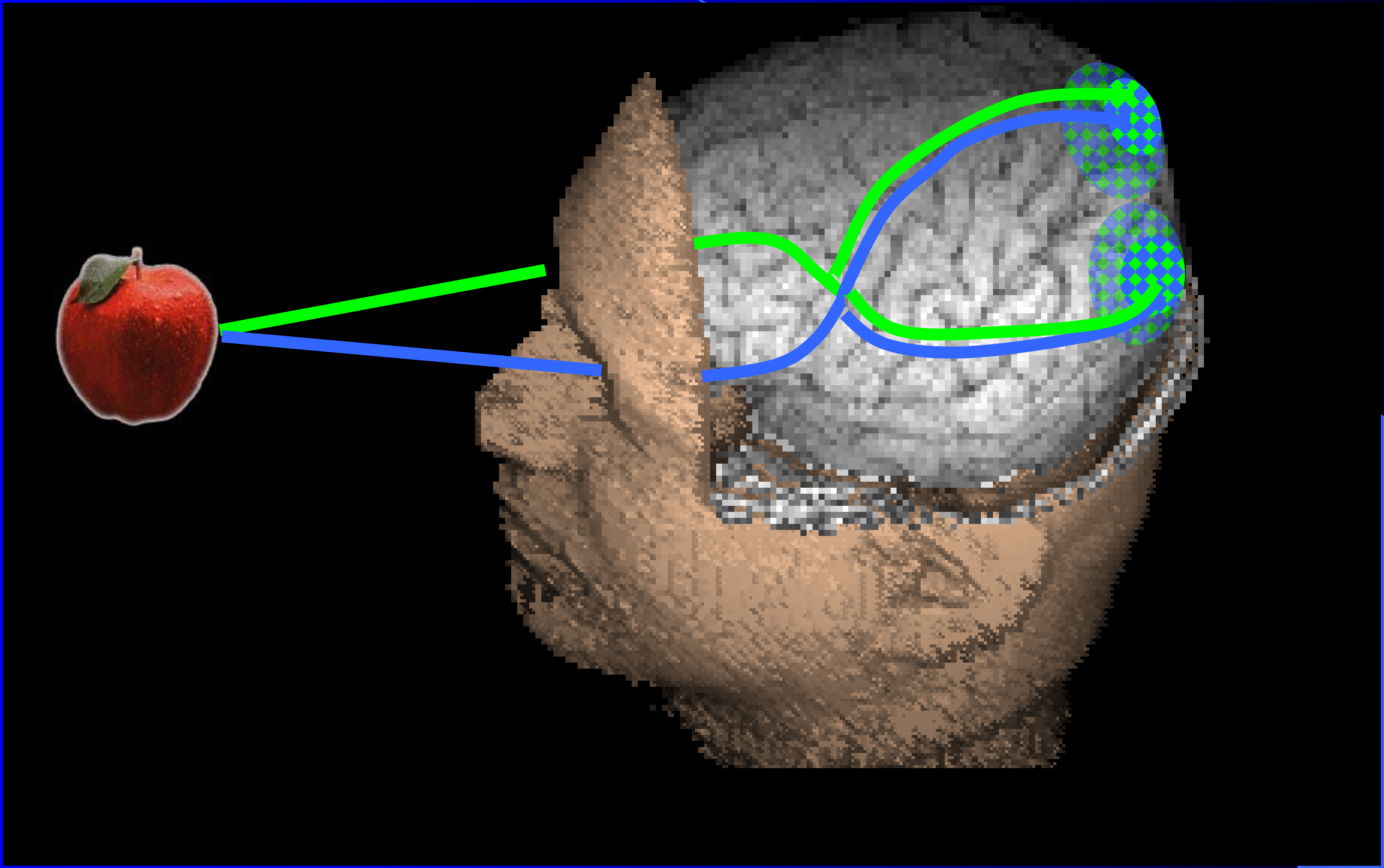


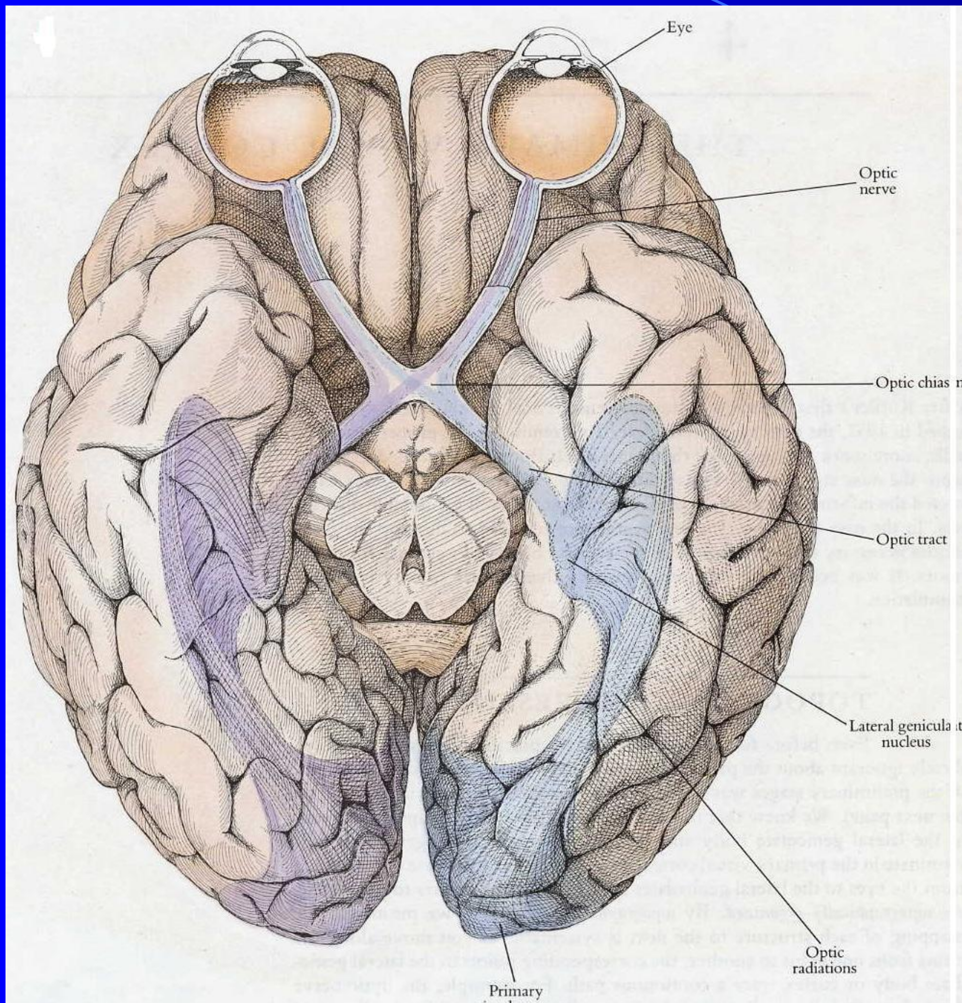
视觉系统是如何估计运动和协调动作的？

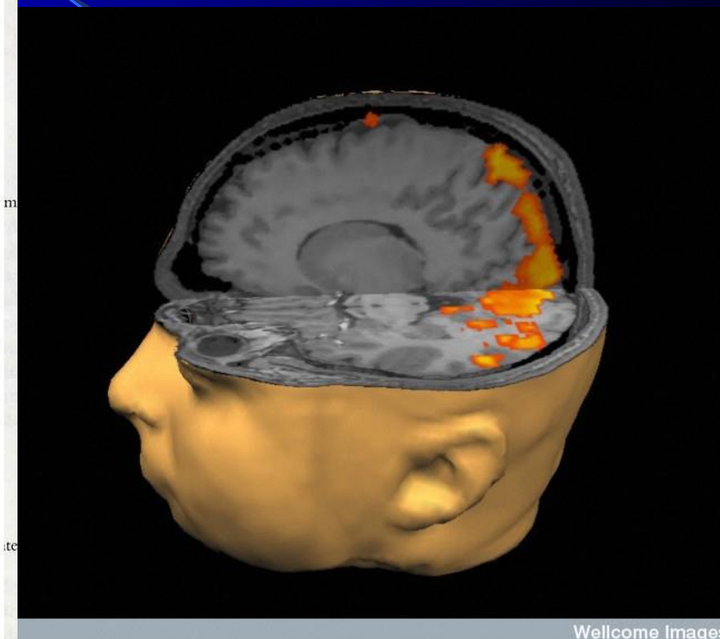
在所有的残疾人中，盲人是最不幸的

我们无时无刻不在使用视觉系统，但我们对之了解甚少

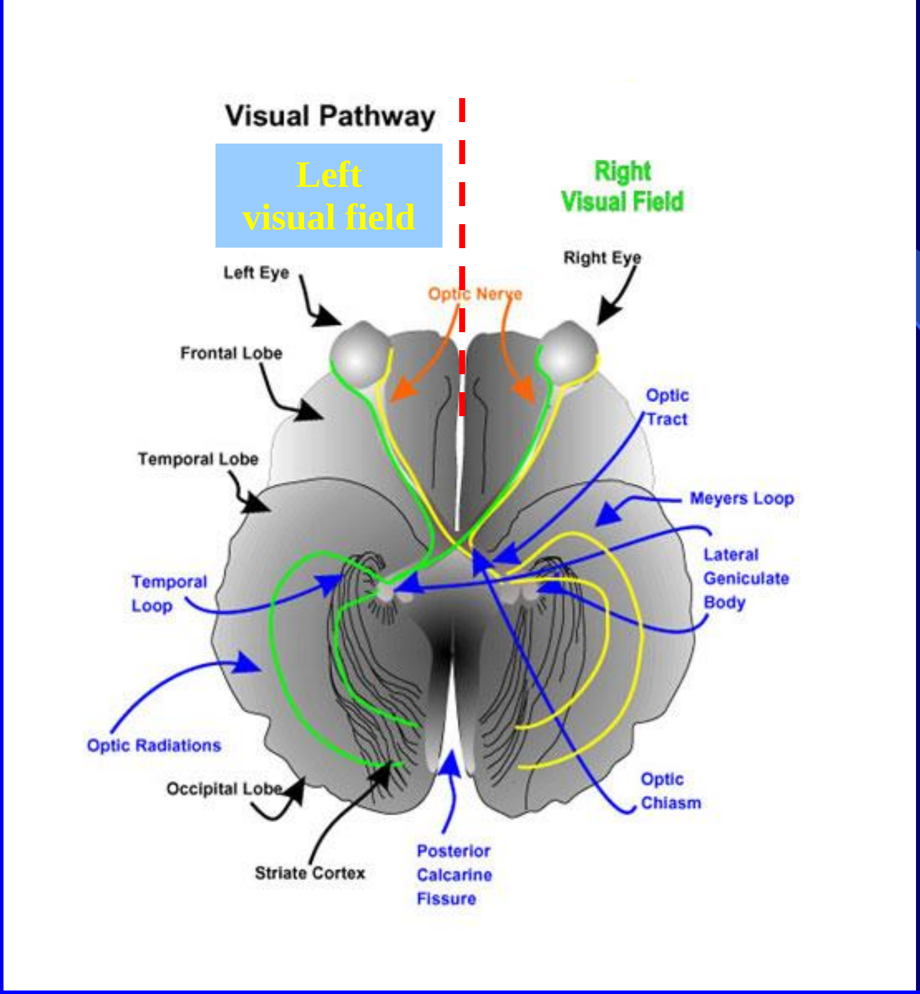
→ 生物视觉简介



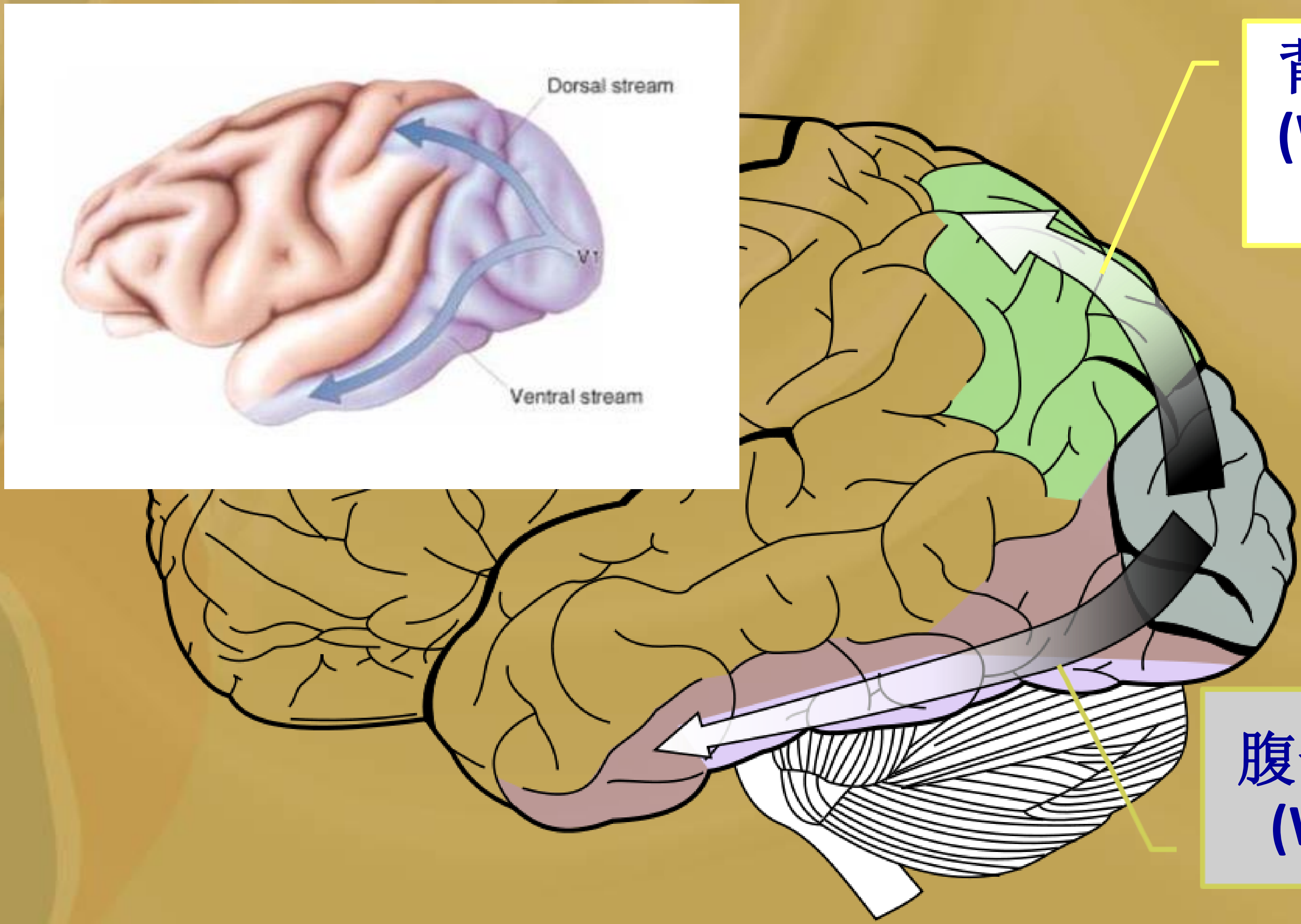

Eye
Optic nerve
Optic chiasm
Optic tract
Lateral geniculate nucleus
Optic radiations
Primary visual cortex

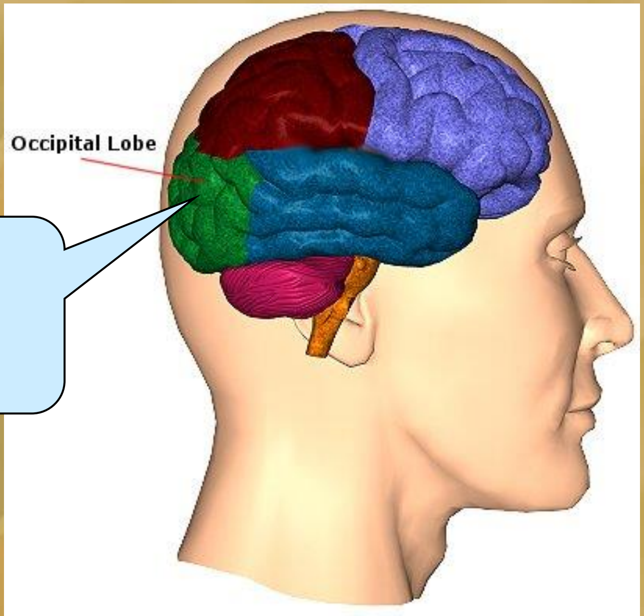

Wellcome Images

视觉通道：visual pathway

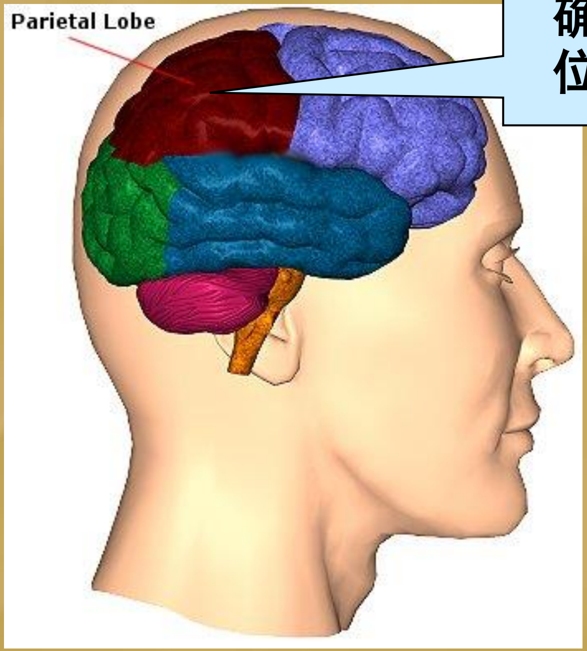


视觉通路概述

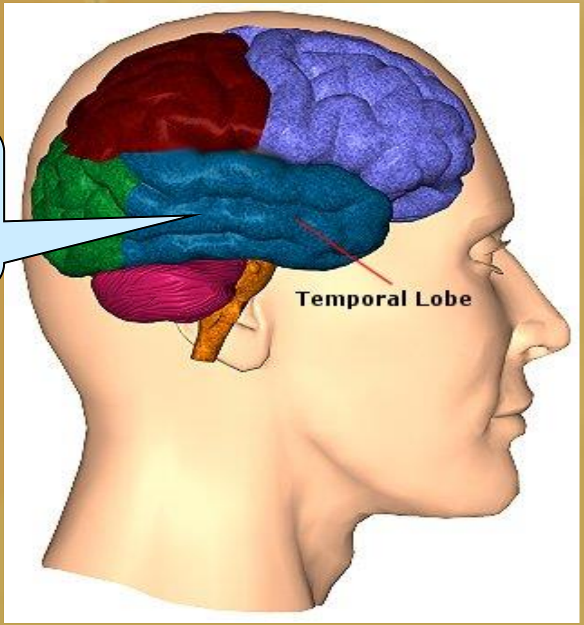




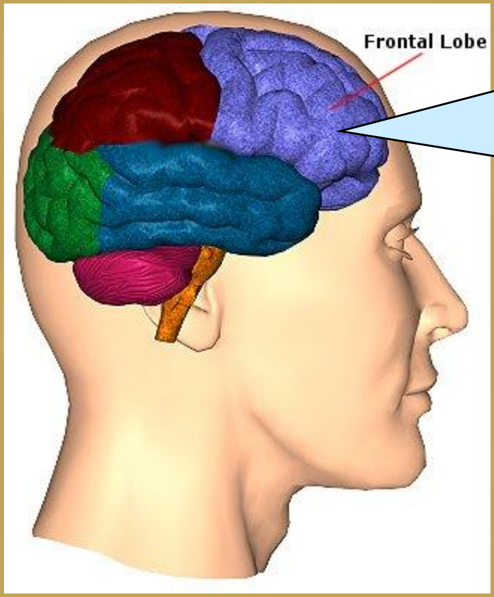
底层共性处理



确定物体的位置

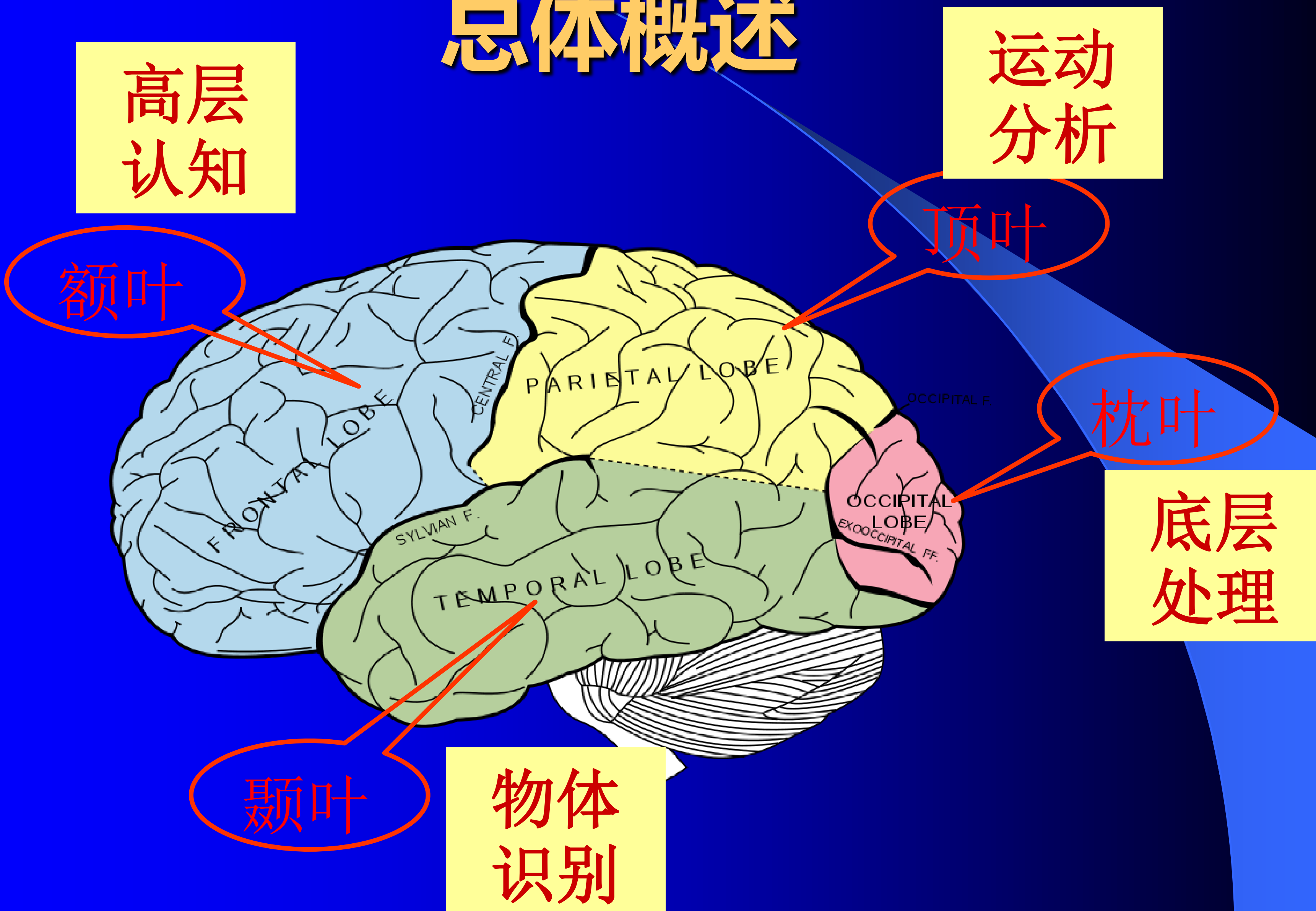


识别物体



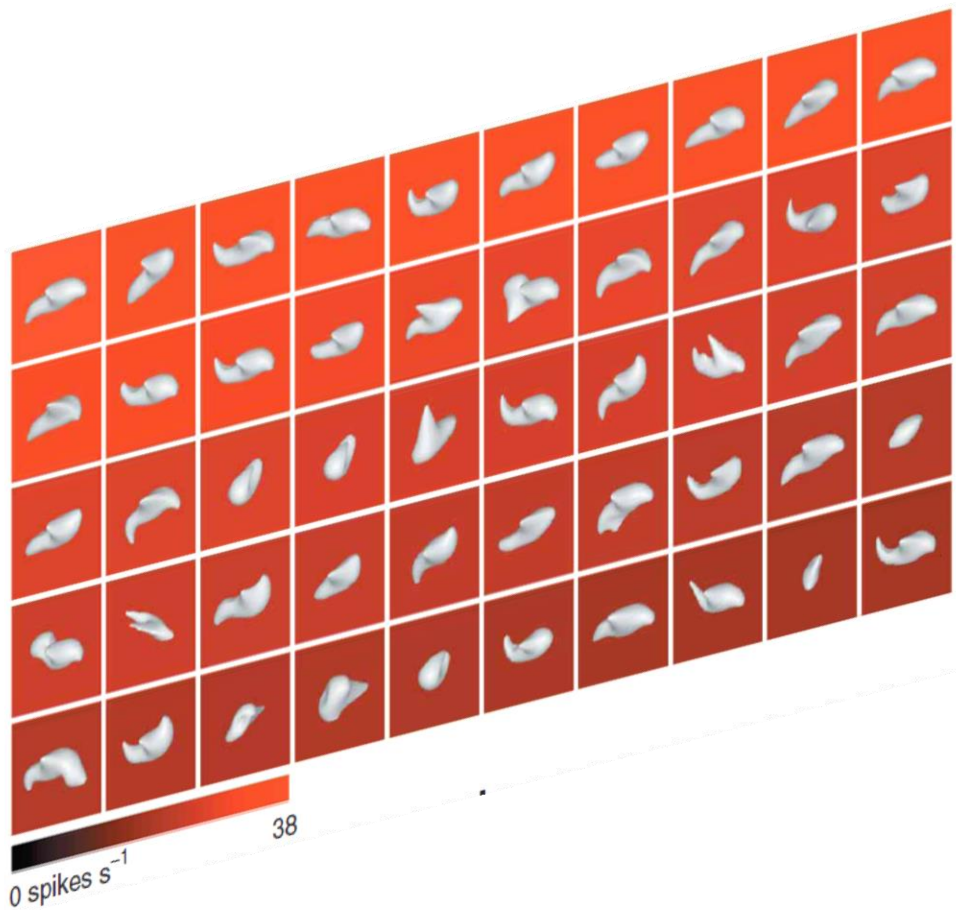
记忆。视觉、听觉、触觉融合

总体概述

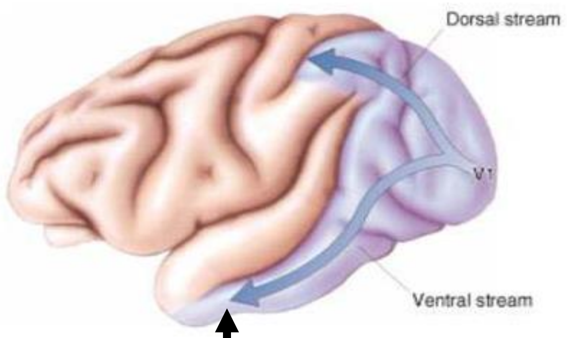


高层认知
额叶
运动分析
顶叶
枕叶
底层处理
颞叶
物体识别
FRONTAL LOBE
CENTRAL F.
PARIETAL LOBE
OCCIPITAL F.
OCCIPITAL LOBE
EXOOCCIPITAL FF.
SYLVIAN F.
TEMPORAL LOBE

IT 是视觉物体识别的最高区域
一个IT神经元所倾向的三维形状

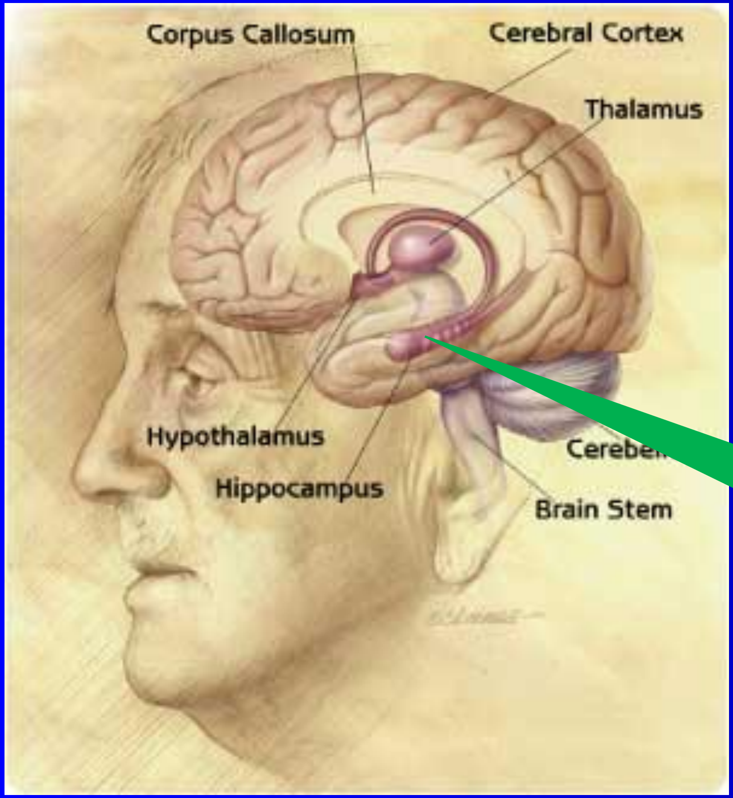


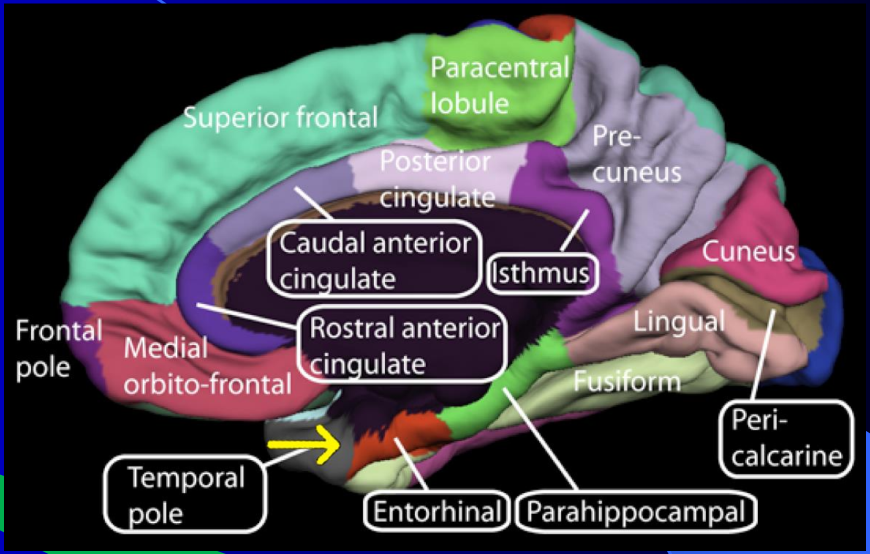
对应问题在IT区已解决



IT:
Inferotemporal cortex

多模态集成





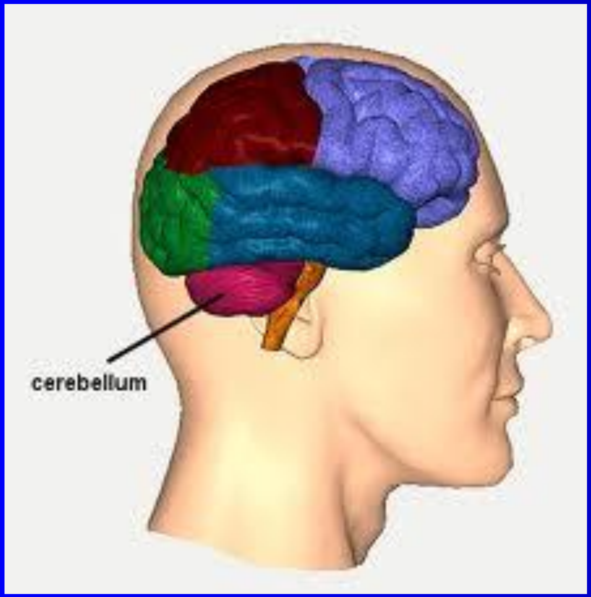
当看到萨达姆侯赛因**照片**，看到“萨达姆侯赛因”**这几个字**，或听到“萨达姆侯赛因”这**几个字的声音**，同一个神经元都放电，说明多模态信息在这个神经元上已得到了集成和统一表达

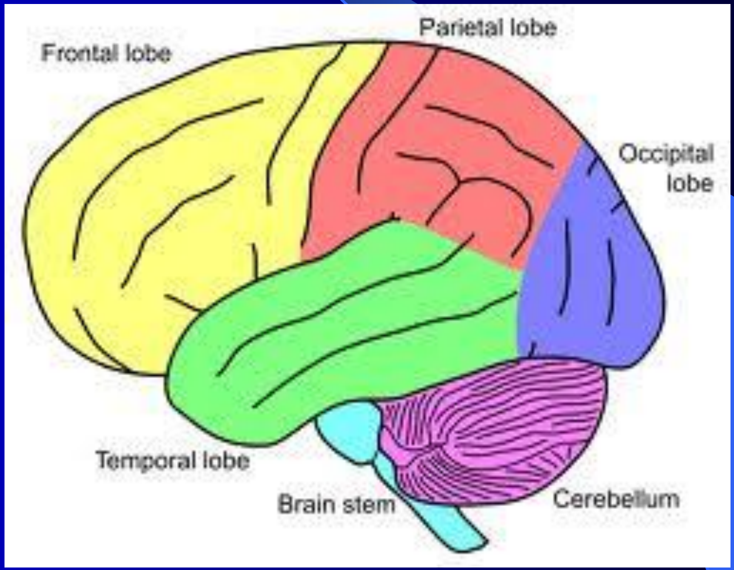
Quian Quiroga R, Kraskov A, et al. (2009). Explicit encoding of multimodal percepts by single neurons in the human brain. *Current Biology* **19**(15): 1308-1313.

人类与恒河猴的皮层分配比例

	人类	恒河猴
总皮层	950 平方厘米	12.6 平方厘米
视觉皮层	总皮层 27%	总皮层 52%
V1区	总皮层 2.2%	总皮层 13%
听觉皮层	总皮层 8.0%	总皮层 3.0%
运动皮层	总皮层 7%	总皮层 8 %
其它感觉	7%	10%
认知等	51%	25%

Cerebellum （小脑：运动控制）





智能与神经元多少联系不大

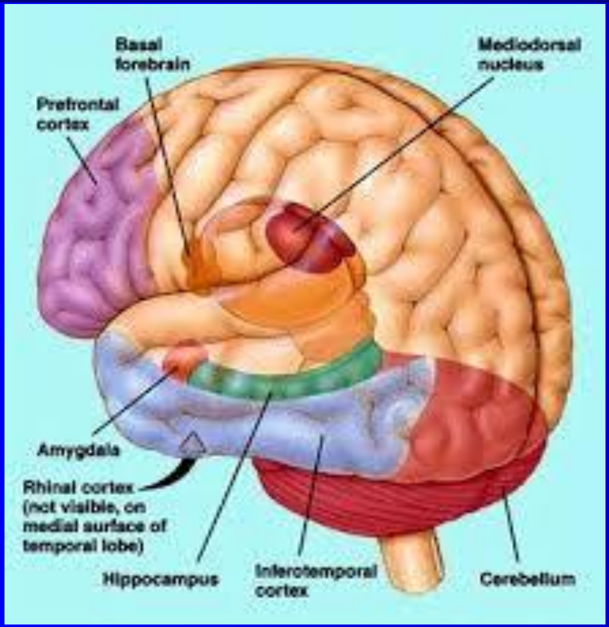
Brain	Cerebellum	entire cortex	Rest of the brain
86	69	17	< 1 billions
Mass	10%	82%	
Neuron	72%	19%	

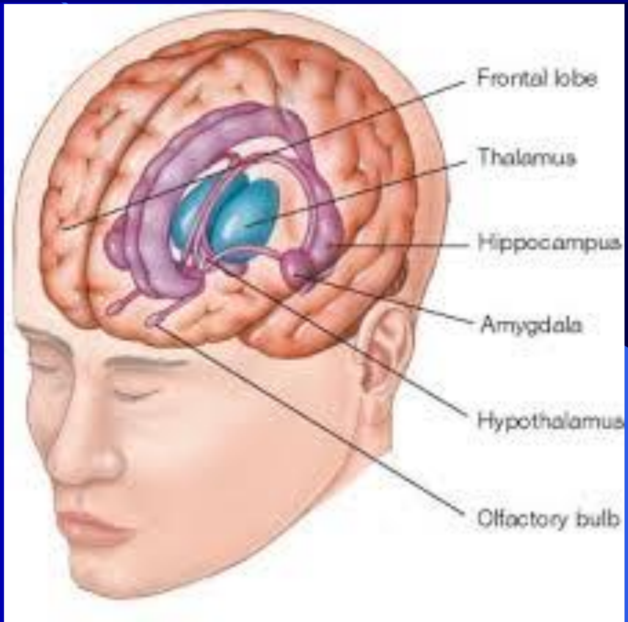
Human cerebral cortex volume is 2.75 times larger than in chimpanzees, but has only 1.25 times more neurons. Shariff G. A (1953), Cell counts in the primate cerebral cortex, Journal of Comparative Neurology, 98(3):381-400

Brain has 85 billions nonneuronal cells

Azevedo F. A. C et al. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primal brain, Journal of Comparative Neurology, 513(5):532-541, 2009

Amygdala（杏仁核：恐惧）





Amygdala: unconciousness function: Reflective acts

If a pattern associated with danger in the past is recognized by the amygdala, it sends an impulse along a direct connection to the brain stem, which then activates the flight-or-fight response and rings the alarm.

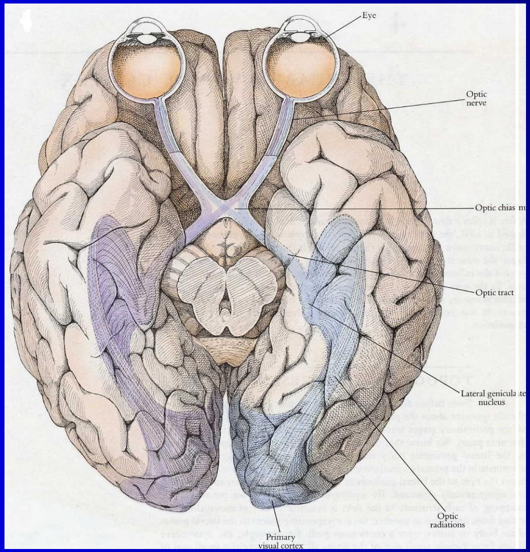
Jump after seeing a snake: I had seen it, but I did not know I had seen it. (Unconsciousness action)

左右脑功能有显著差异

1: Why did the apple fall down ? (right Hemisphere)

2: Noting pushed it, why does not is go up ? (left hemisphere)

误解： 视网膜是一台照相机



视网膜（retina)比照相机复杂的多

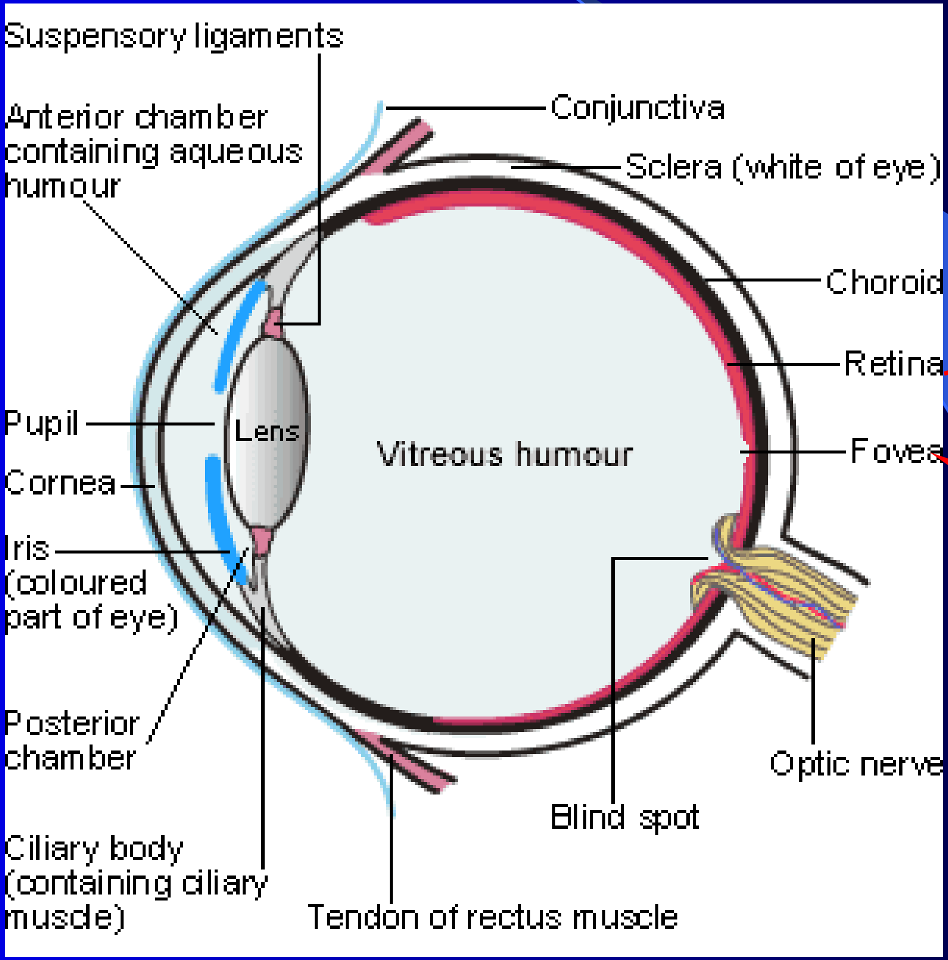
视觉通道

视网膜

外膝体

视皮层

高级区域



视网膜

中央凹

视觉通道

视网膜

↓

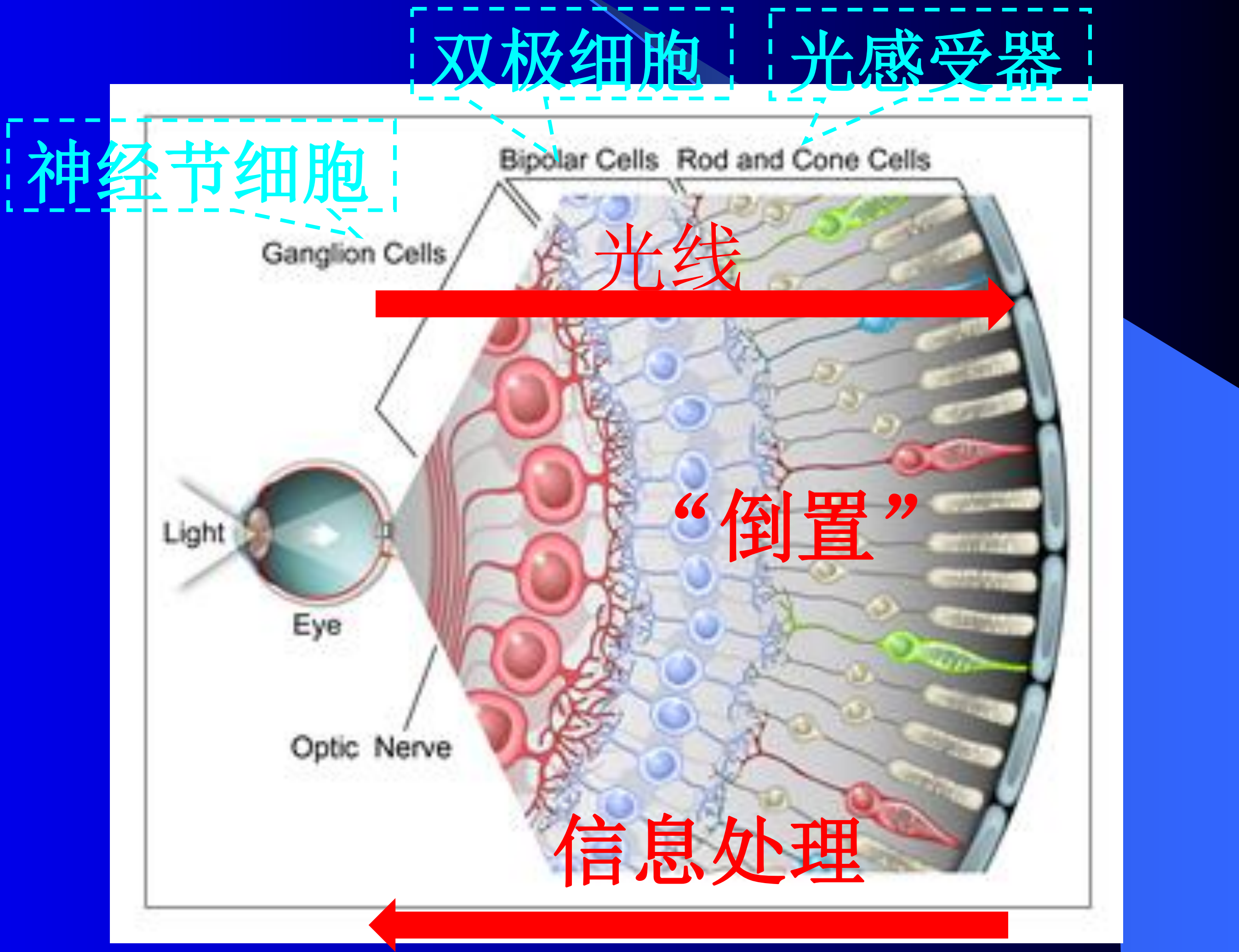
外膝体

↓

视皮层

↓

高级区域



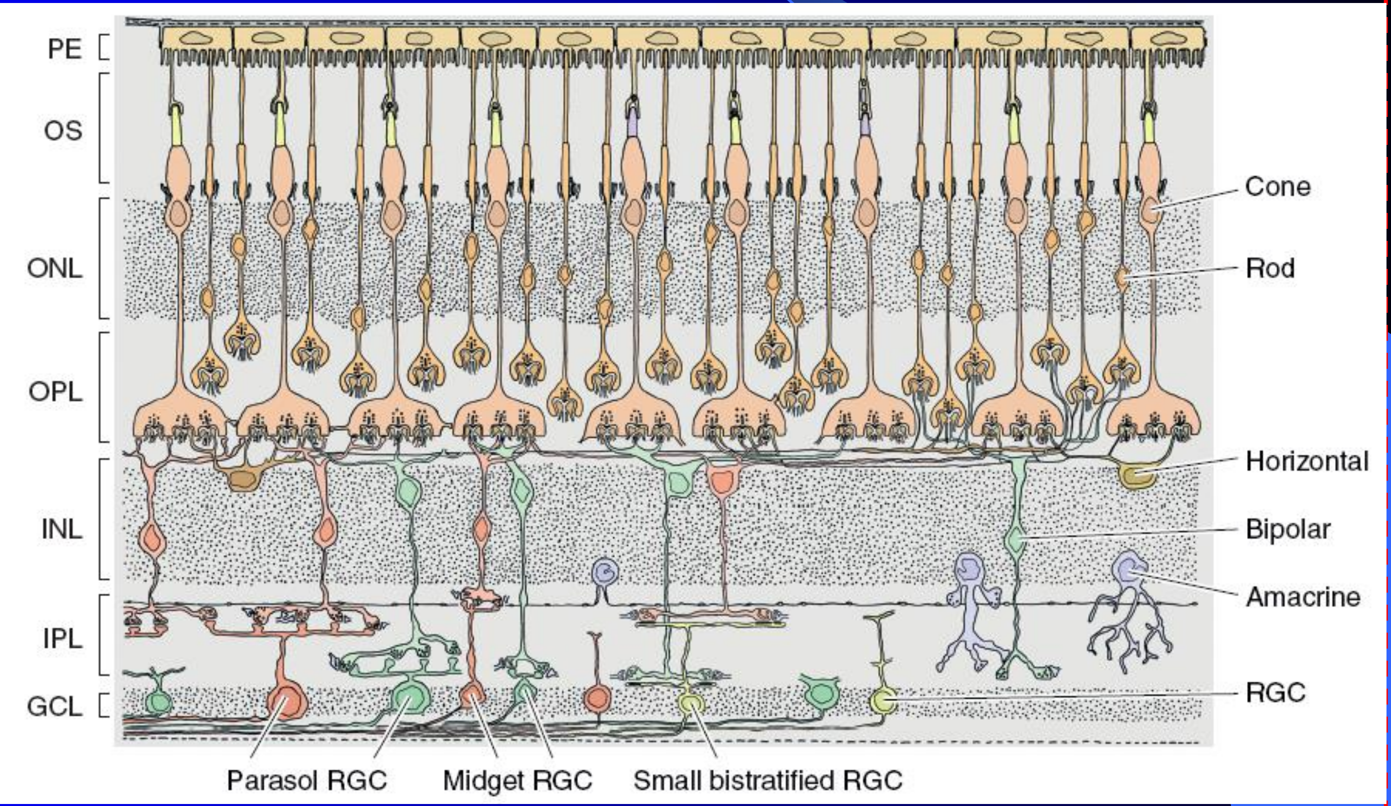
视觉通道

视网膜

外膝体

视皮层

高级区域



Field G, Chichilnisky E (2007). Information processing in the primate retina: circuitry and coding. *Annu. Rev. Neurosci*. **30**: 1-30.

视觉通道

视网膜

外膝体

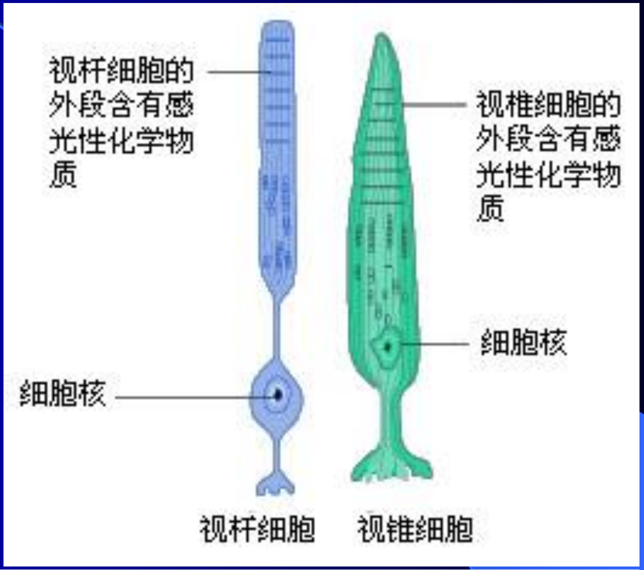
视皮层

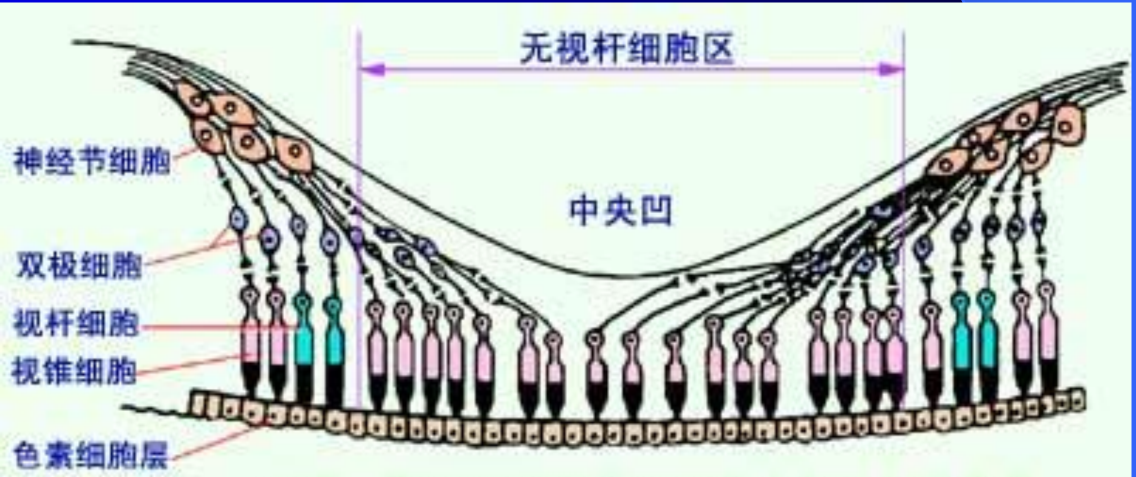
高级区域

光感受器

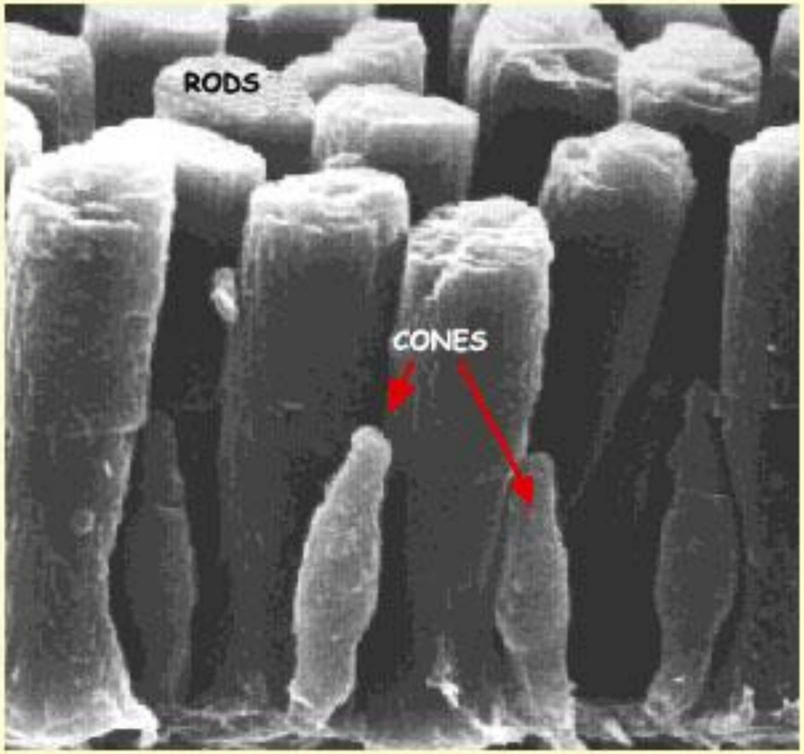
双极细胞

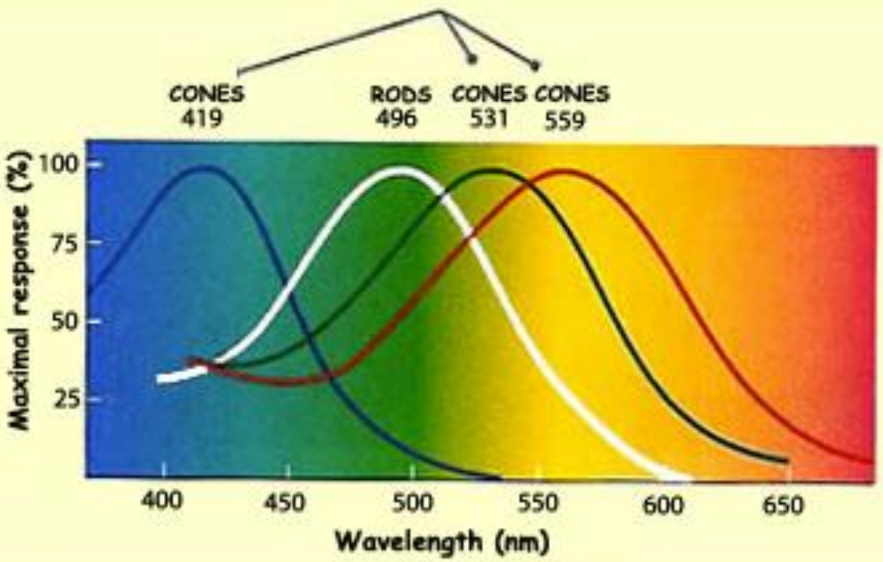
神经节细胞





视杆、视锥细胞在中央凹附近的分布示意图





视杆：120 millions

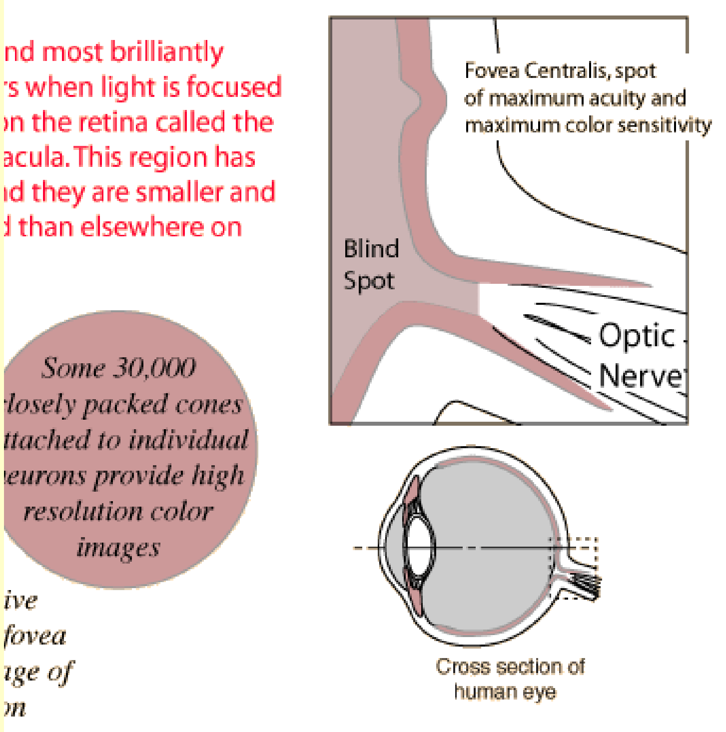
视锥：6-7 millions

红色锥：64%

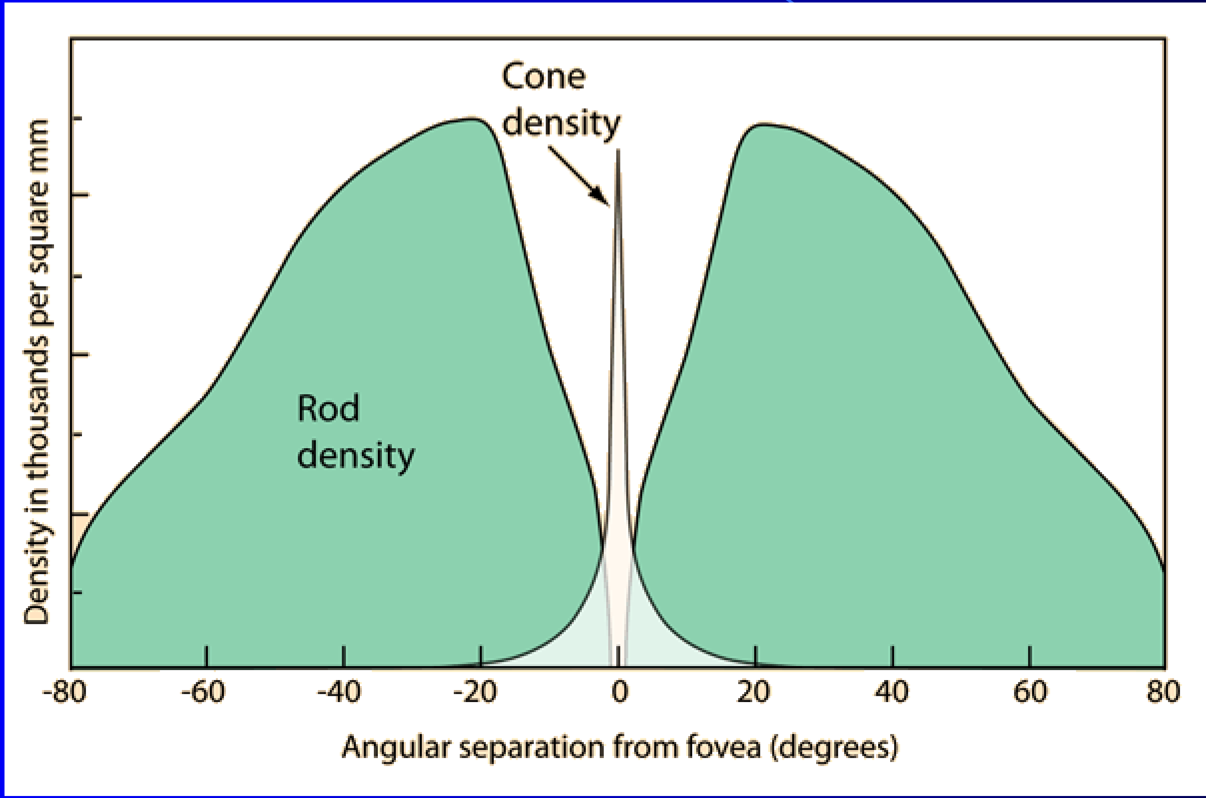
绿色锥： 32%,

蓝色锥： 2%

视锥主要分布在中央凹直径0.3 毫米，中央凹无视杆



视锥和视杆分布：视锥主要分布在中央凹



$$\frac{No-of-P-cells}{No-of-M-cells}\approx 10 \quad \frac{No-of-K-cells}{No-of-M-cells}\approx 1$$

高度压缩

Midget GC: 80% GC

Parasol GC: 20% GC

Cones

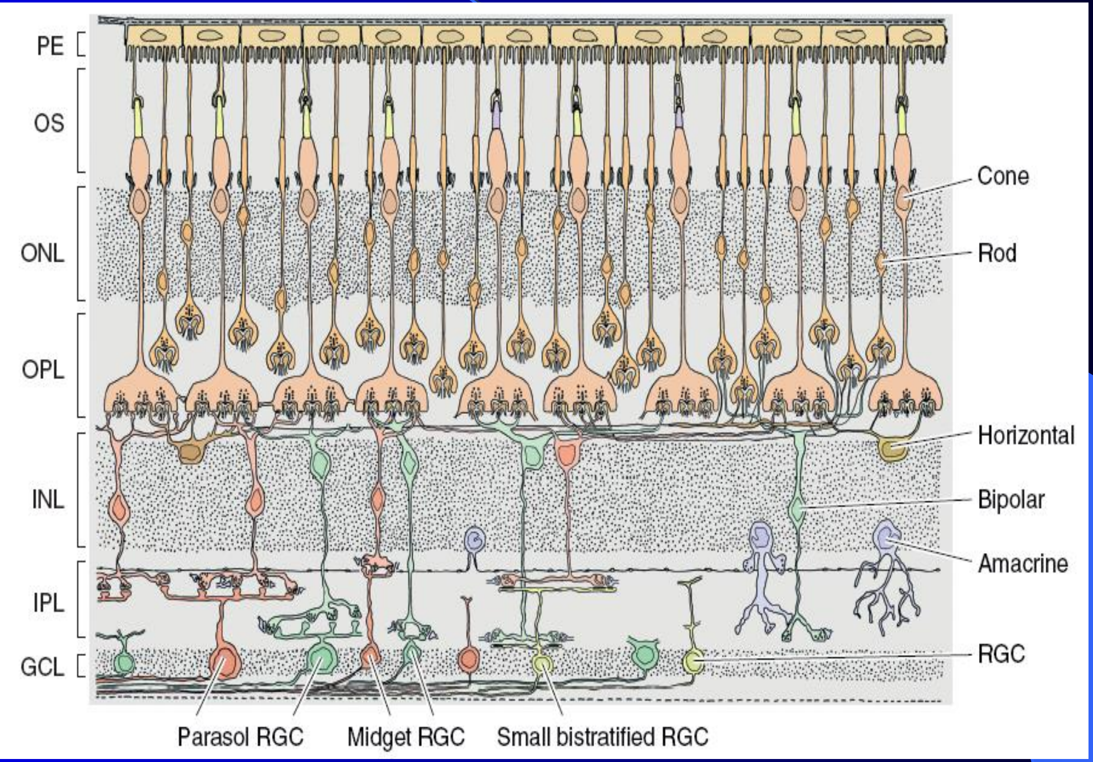
5-6 million

Rods

120 million

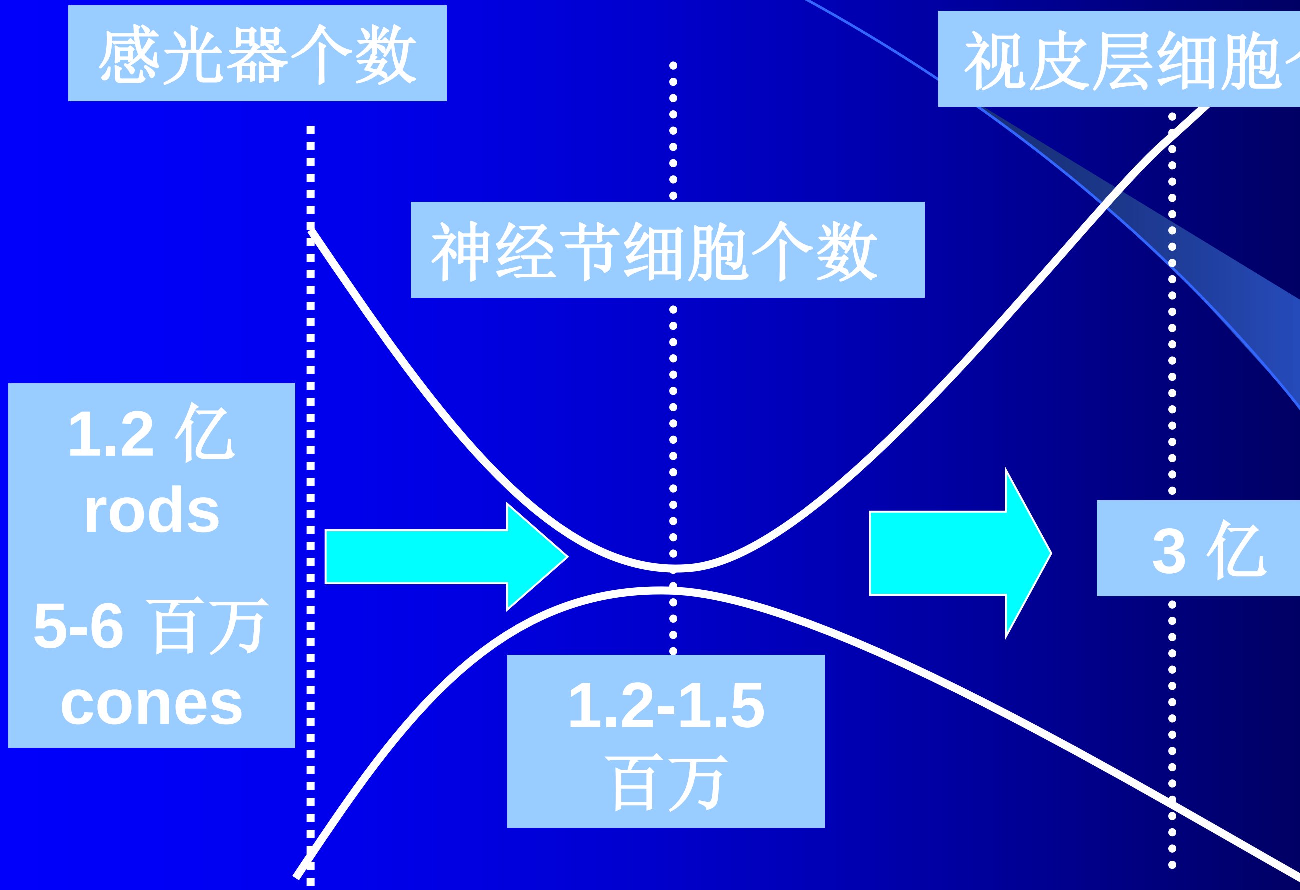
GCs

1.2-1.5 million



现在任何一台照相机的分辨率都大于二百万像素

从高度压缩到高度复杂的过程



越到后面的层，细胞个数越来越少。所以，原则上说，DNN后面层的节点要少于前面层。越到后面的层，细胞越来越大，连接越来越多。说明越来越接受多模态信息，处理的对象越来越复杂

Why such a compression ?

神经元对某个刺激形状敏感不等于意识中的感知。只有神经计算达到某个高级层次时，才可以上升到意识中感知的范畴

如从神经计算的观点看，视差（disparity）是双眼视觉形成立体感的基础。人类双眼首次在视皮层V1区融合，然后在V2区，V4区， IT 区逐步计算。但V1区并没有形成我们日常生活中深度的概念。大脑到底在那个区域形成了我们意识中真正的深度的概念，目前仍是一个有待进一步研究的问题

Stupid Vision ?

Magritte’s rider on a horse

Famous but controversial statement about intelligence

" What magical trick makes us intelligent? The trick is that there is no trick. The power of intelligence stems from our vast diversity, not from single, perfect principle".

Marvin Minsky

Minsky and Papert in 1969
networks that lack a hidden layer and
thus cannot learn more difficult
mappings such as XOR

Isaac Asimov

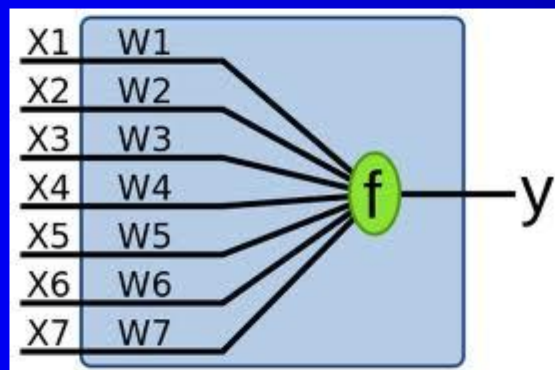
Writer of Science fiction
" two people he would admit were more intelligent than he was"

Carl Edward Sagan ,
Astronomer

AI winter

Symbolism

"throwing out the baby with the bathwater"



Perceptrons: an introduction to computational geometry is a book written by Marvin Minsky and Seymour Papert

Minsky and Papert in 1969
networks that lack a hidden layer
and
thus cannot learn more difficult
mappings such as XOR

谢谢

欢迎批评指正

Henry Gustav Molaison Feb. 26, 1926 – Dec. 2, 2008

H.M. had his hippocampus removed to prevent otherwise intractable epilepsy, in 1957. He then developed the inability to learn new episodic information (anterograde amnesia), as well as some degree of forgetting of previously learned knowledge (retrograde amnesia). But he remembered how to talk, the meanings of different words and objects, how to ride a bike, and could learn all manner of new motor skills. This was a clear indication that the hippocampus is critical for learning only some kinds of new knowledge.

he could also learn new semantic information, but that this occurred relatively slowly, and the learned knowledge was more brittle in the way it could be accessed, compared to neurologically intact people. This further clarifies that the hippocampus is critical for episodic, but not semantic learning. However, for most people semantic information can be learned initially via the hippocampus, and then more slowly acquired by the neocortex over time. One indication that this process occurs is that HM lost his most recent memories prior to the surgery, more than older memories (i.e., a temporally-graded retrograde gradient, also know as a Ribot gradient). Thus, the older memories had somehow become consolidated outside of the hippocampus, suggesting that this gradual process of the neocortex learning information that is initially encoded in the hippocampus, is actually taking place. We discuss this process in the next section.