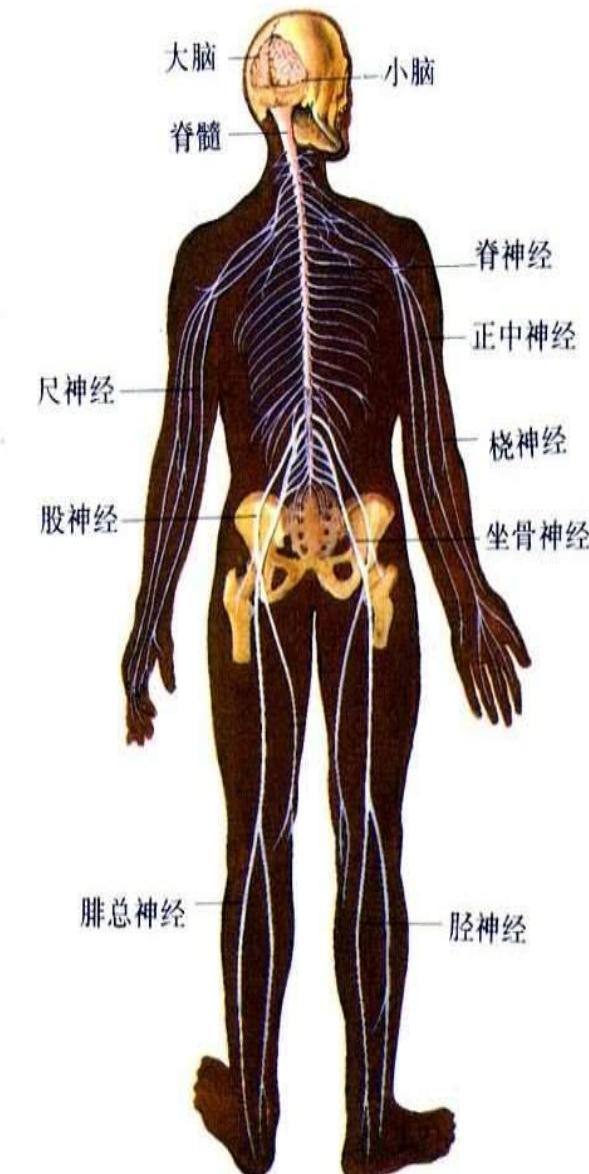
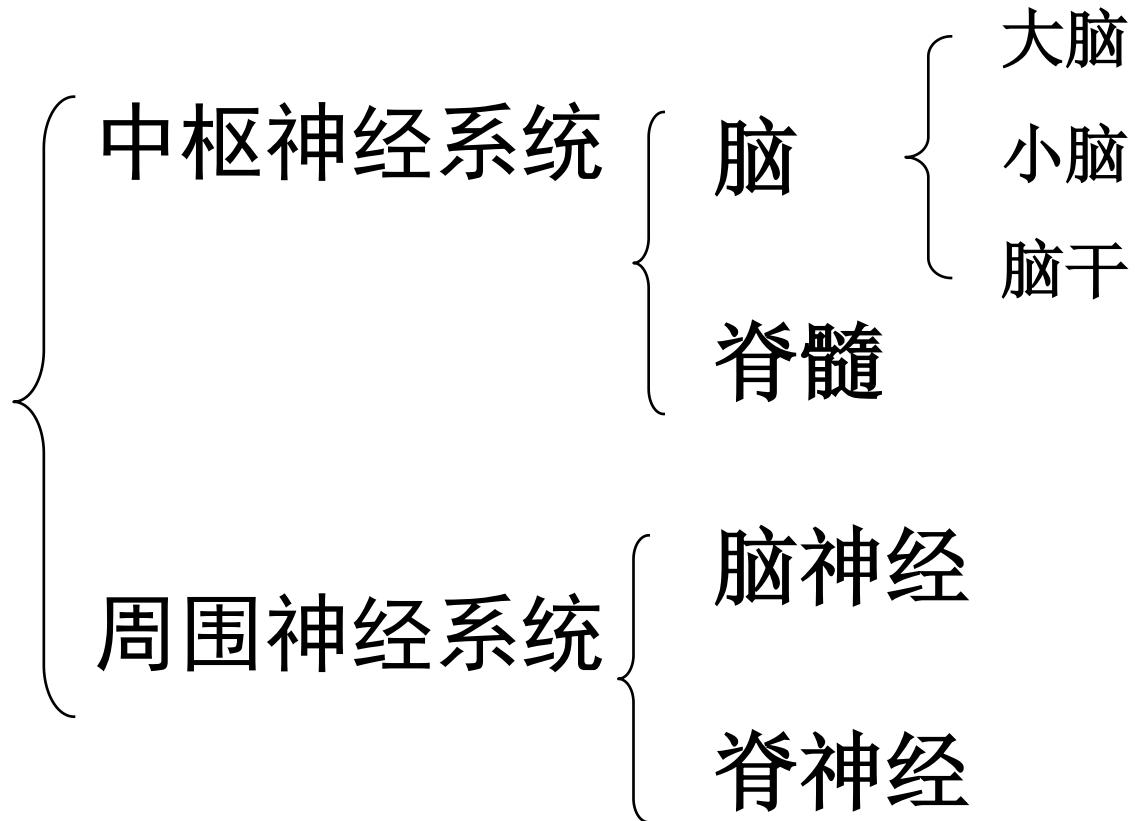


Visual Perception

人体神经系统组成



人的神经系统

人体对信息的感知

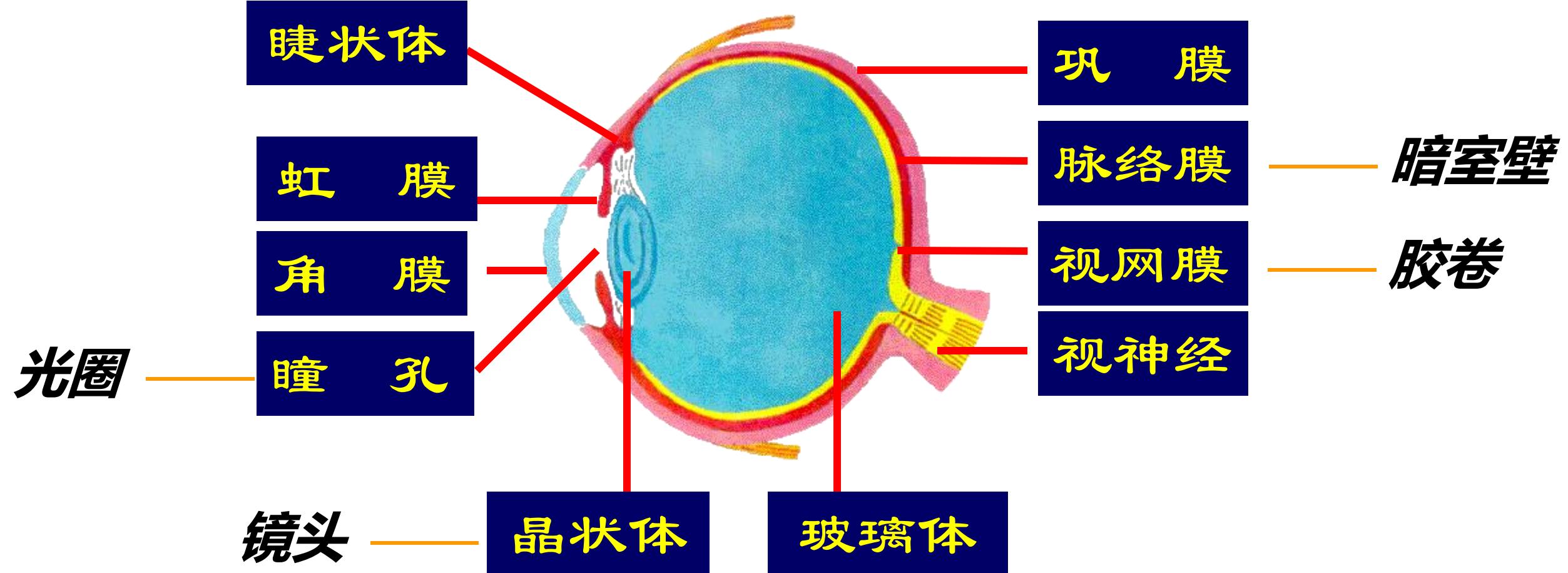
人体对信息的感知是通过人体的感觉器官

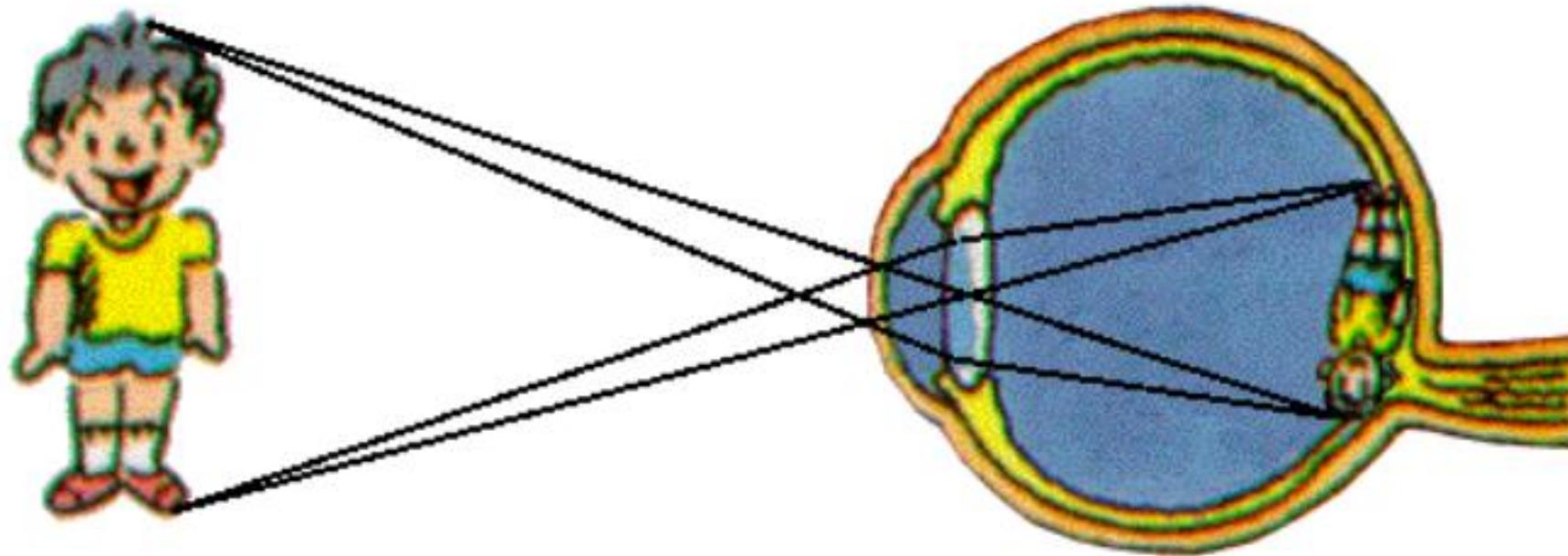
皮肤中的感受器（触觉）：耳、眼、鼻、舌等

眼睛 是心灵的窗户

是人体获得外界信息最多的器官

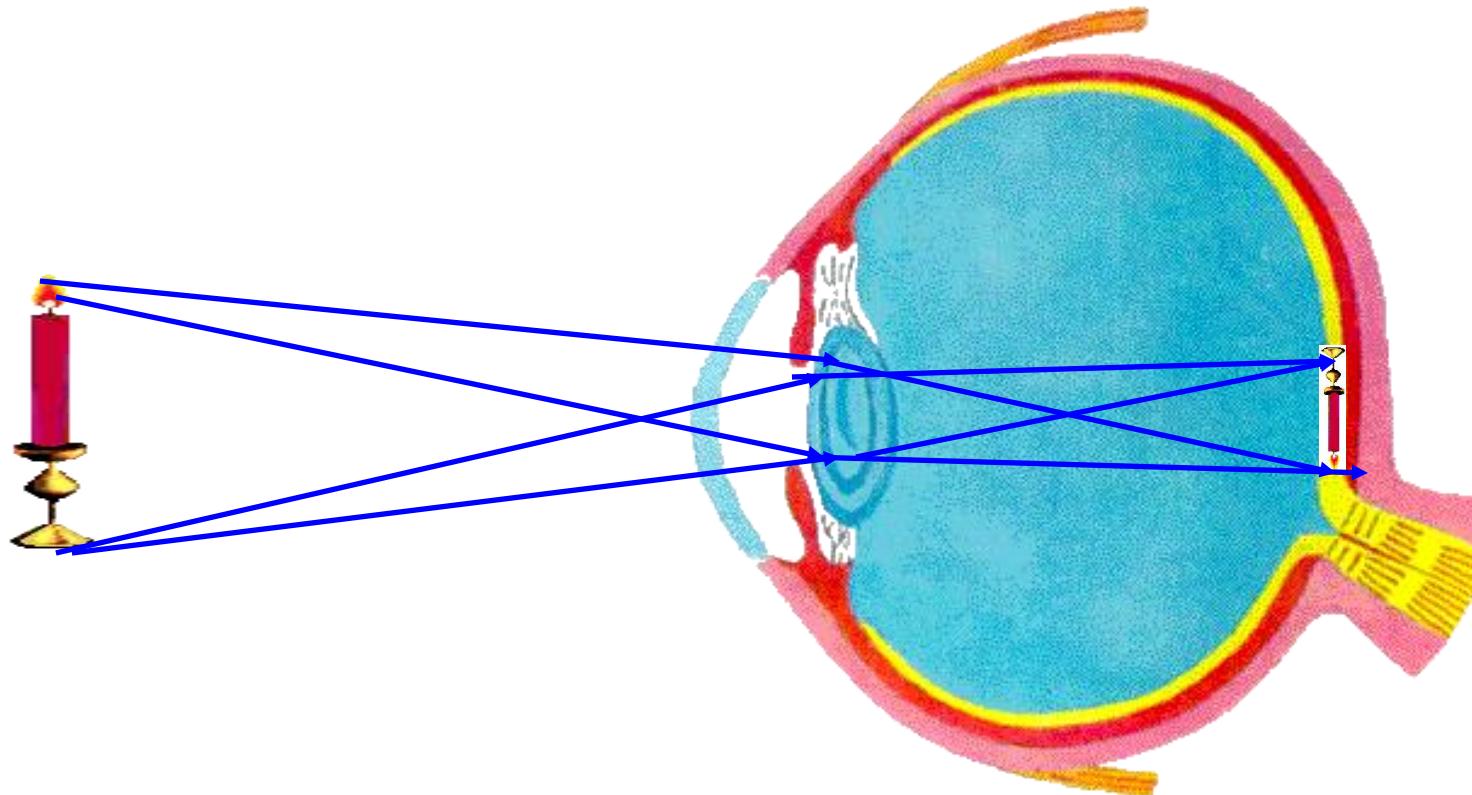
眼球 vs 照相机

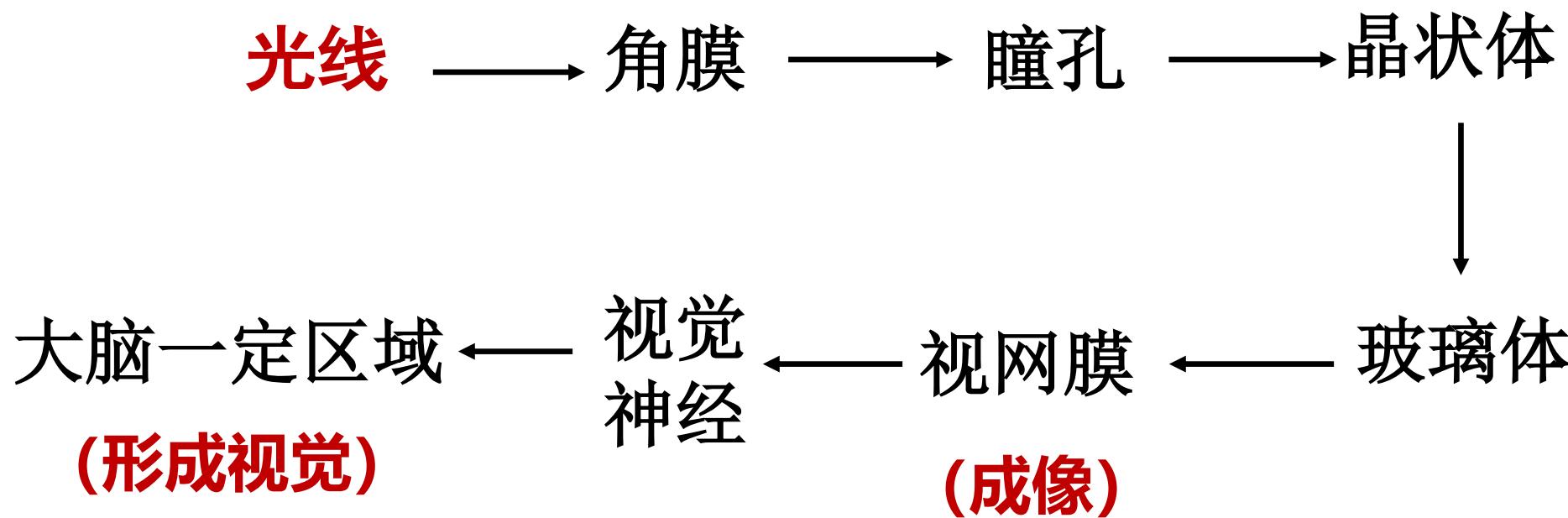
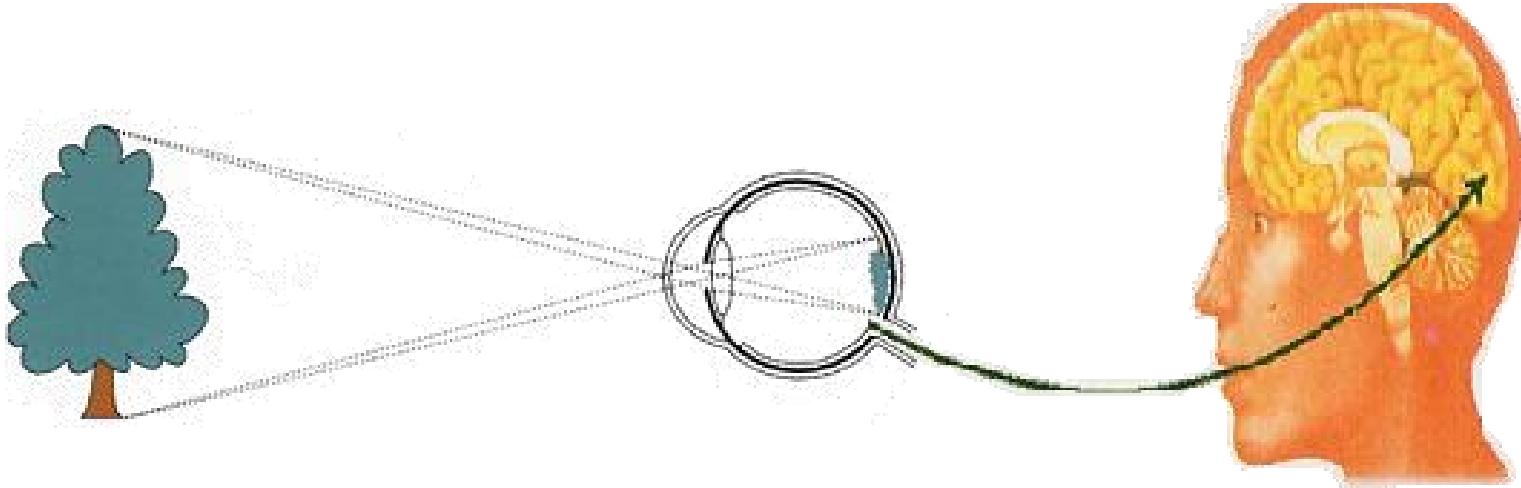




物体在视网膜上成像

视觉的形成





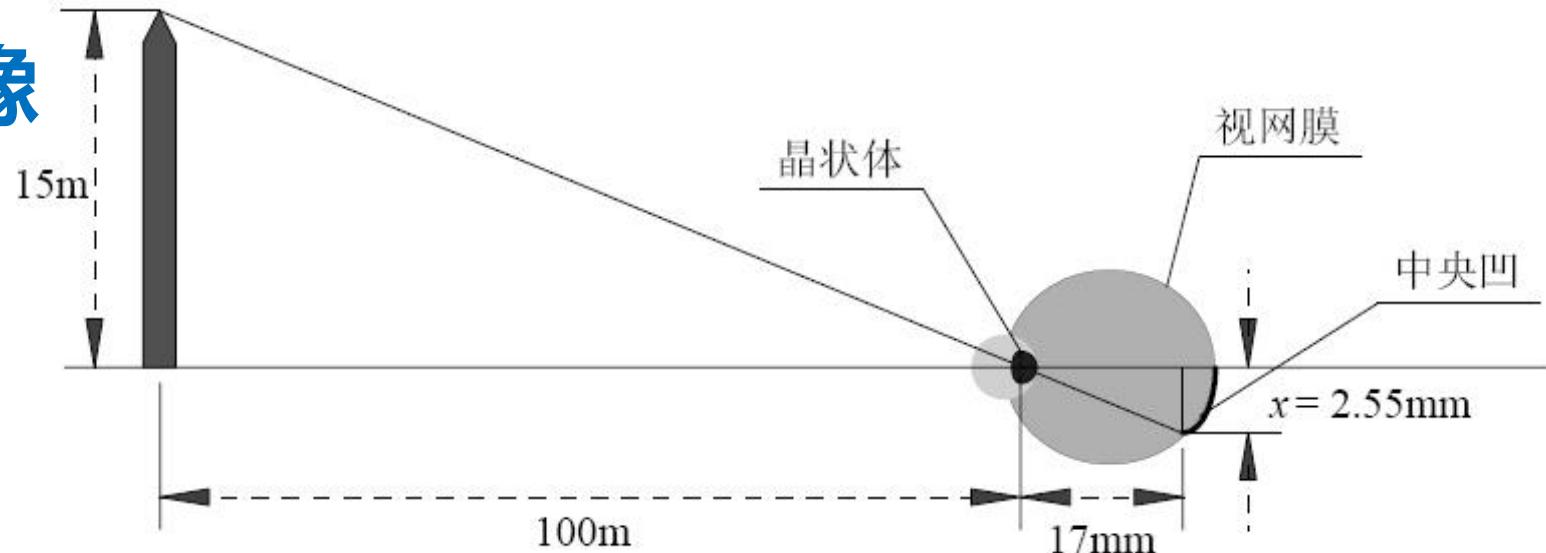
视觉原理

- 视觉过程

- 包括：光学过程、化学过程、神经处理过程

1. 光学过程

- 物体在视网膜上成像



视觉原理

2. 化学过程

- 接受光的能量并形成视觉图案（确定成像的亮度或颜色）

- 锥细胞和柱细胞（视网膜表面）

(1) 锥细胞：数量少，对颜色很敏感

锥细胞视觉：明视觉或亮光视觉

(2) 柱细胞：数量多，分辨率比较低；不感受颜色并对低照度较敏感

柱细胞视觉：暗视觉或微光视觉

视杆体与视锥体

- 视网膜上有杆状体和锥状体两类**视觉接收器**；
- **视锥体(Cons)**: 结构上短而粗，数量少, 约6.5 million, 光灵敏度较低，它们提供**明视(Photopic Vision)**，其响应光亮度范围比视杆体要高5 ~ 6个数量级。在中间亮度范围是两种视觉细胞同时起作用。视锥体集中分布在视网膜中心。
- **视杆体(Rods)**: 细长而薄，数量上约100 million, 它们提供**暗视(Scotopic Vision)**, 即在低几个数量级亮度时的视觉响应，其光灵敏度高。

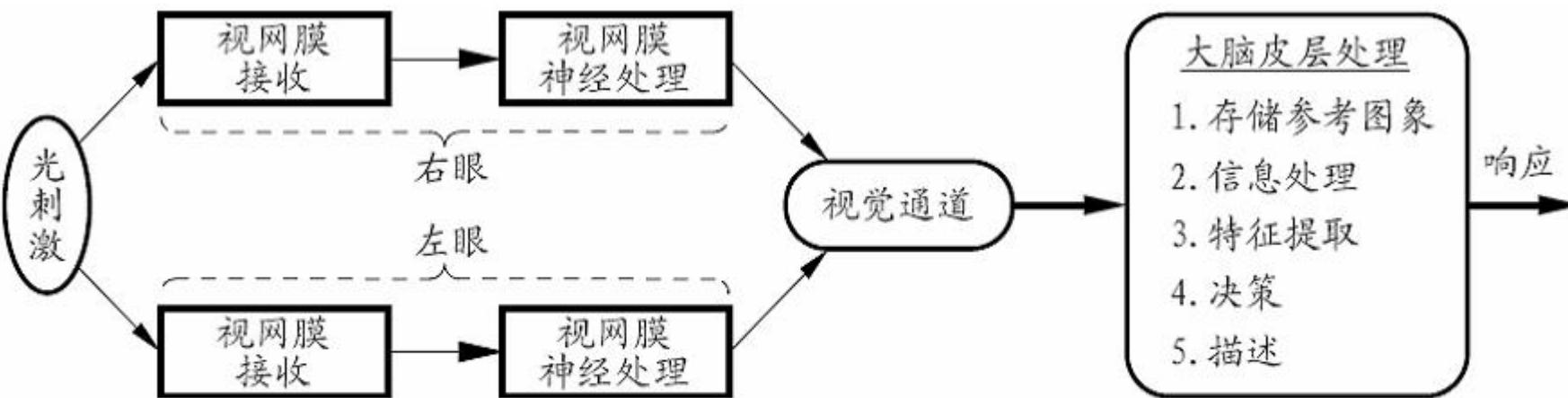
视觉原理

3. 神经处理过程

- 对光的感觉转换为对景物的知觉
- 每个视网膜接收单元都与一个神经元细胞借助突触 (synapse) 相连
- 每个神经元细胞借助其它的突触与其它细胞连接，从而构成光神经 (optical nerve) 网络
- 光神经进一步与大脑中的侧区域(side region of the brain)连接，并到达大脑中的纹状皮层(striated cortex)
- 对光刺激产生的响应经过一系列处理最终形成关于场景的表象，从而将对光的感觉转化为对景物的知觉

视觉原理

• 整体视觉过程



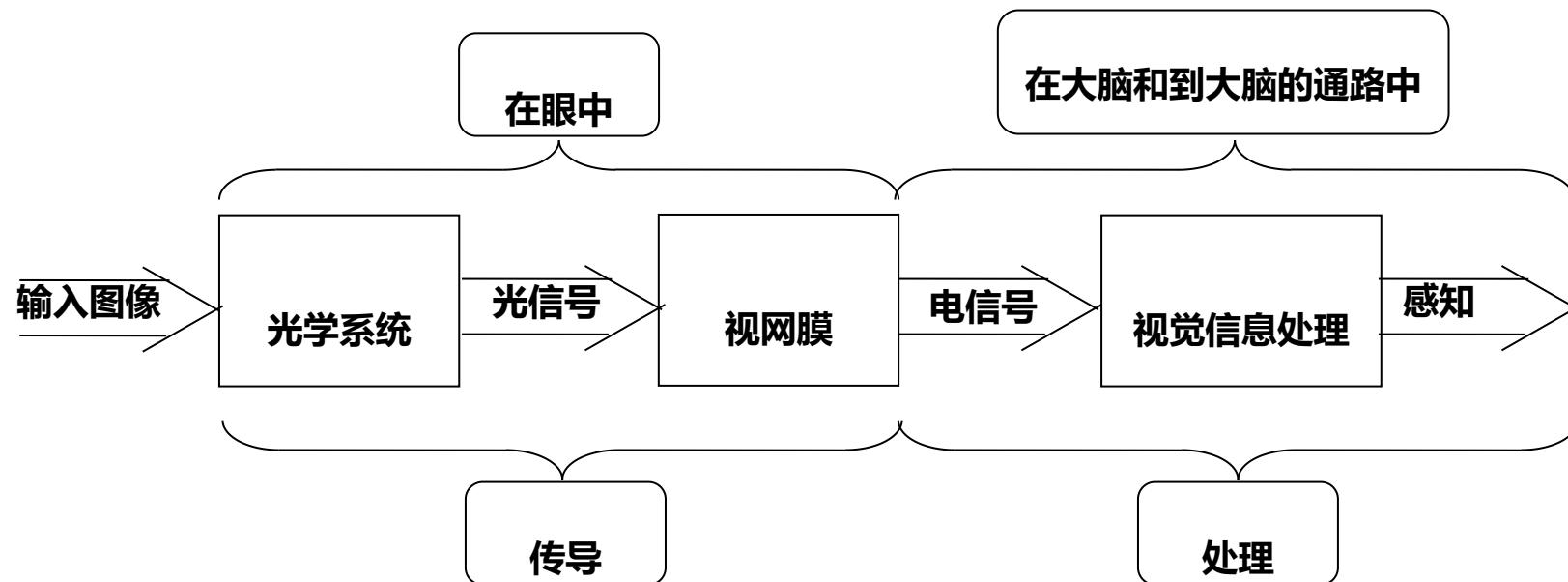
视觉过程流图

视觉 = “视” + “觉”

人类视觉系统的组成

人类视觉的**眼—脑系统**看成一个有生命的**光学变换器和信息处理系统**:

1. **光学系统**: 由于有关的神经活动最少, 因此最为简单
2. **视网膜**: 把光信号转变成电信号, 并进行某些细胞一级的处理
3. **视觉信息处理**: 从视网膜到大脑皮层的视觉通路上所完成的复杂处理



视觉感知

- 视觉感知是视觉的内在表象。
- 视觉感知包括两个不同的感知层次：
 - 视觉的**低级感知层次**
 - 视觉的**高级感知层次**

视觉低级感知层次

• 视觉系统从外界获取图像。就是在眼睛视网膜上获得周围世界的光学成像，然后由视觉接收器（杆状体和锥状体在视网膜上作为视觉接收器），将光图像信息转化为视网膜的神经活动电信号，最后通过视神经纤维，把这些图像信息传入大脑，由大脑获得图像感知。

视觉低级感知层次

- 光图像激活视杆体或视锥体时，发生**光电化学反应**，同时**产生视神经脉冲**，视觉系统散布视神经中有80万神经纤维，**视觉系统传播视神经脉冲**。许许多多的视杆体和视锥体相互连接到神经纤维上。
- 视觉系统的可视**波长范围**：380nm ~ 780nm；
- 视觉系统的可响应**亮度范围**：1 ~ 10个量级的幅度范围。

视觉高级感知层次

- 大脑对视神经纤维传送来的图像信息进行分析和理解，通过图像获得对周围世界感知的**信息和知识**。
- 人们对大脑的高级感知层次至今知之甚少，仍是生理学、神经科学、生物物理学、生物化学等研究的重要课题。

视觉和视觉感知

- “**视觉是一个信息处理过程。** 它能从外部世界的图像中得到一个既对观察者有用又不受无关信息干扰的描述。”
- “**视觉感知**又是与过去留存于**记忆**中的同类活动有关，视觉储积起大量的视觉意象。记忆形象可用于对知觉对象的辨认、解释和补充。”
- **使计算机具有人类视觉能力**，研究人类的视觉感知，模仿人类的视觉感知，是研究工作的重要途径。

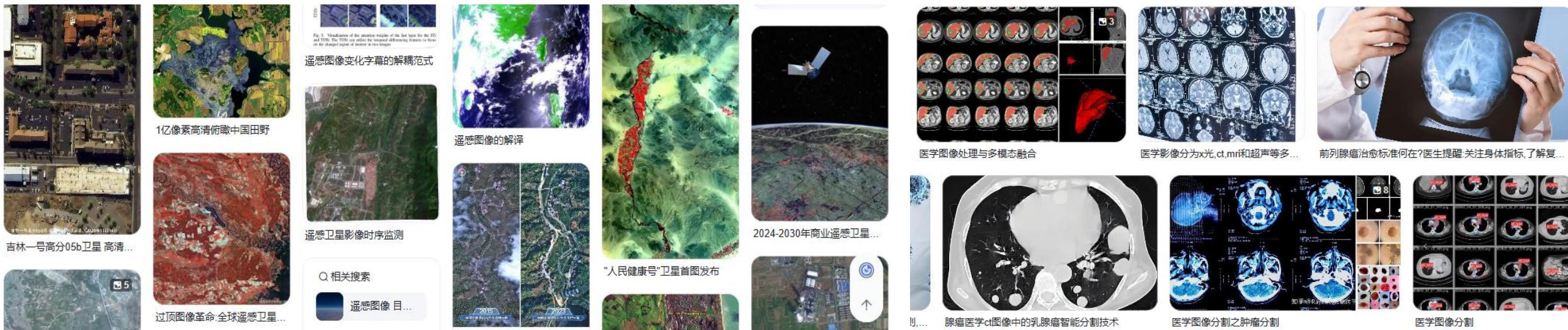
视觉特性

- 视觉特性是视觉的外在表现；
- 图像是周围世界的一种映射，对于运动图像，空间坐标 x, y, z 都是时间 t 的函数，若在连续的不同时间获取图像，可以获得序列图像：

$$\{I_1, I_2, \dots, I_n\}$$

视觉特性

- 对于按**不同波段**获取图像, 可获得彩色图像或不同波段的图像信号(如遥感图像、医学图像等); 对于按**不同视角**, 即不同的x,y,z间相互关系, 可以得到不同视角的不同图像。

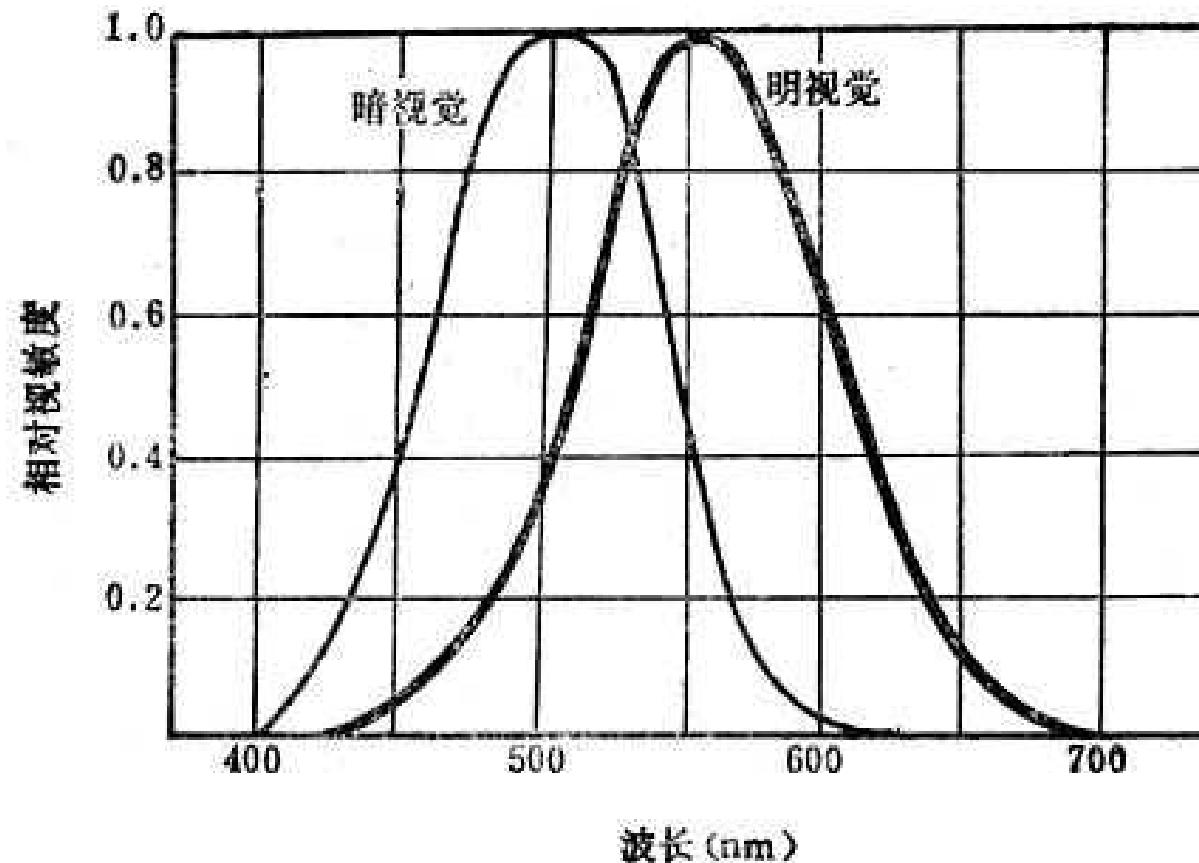


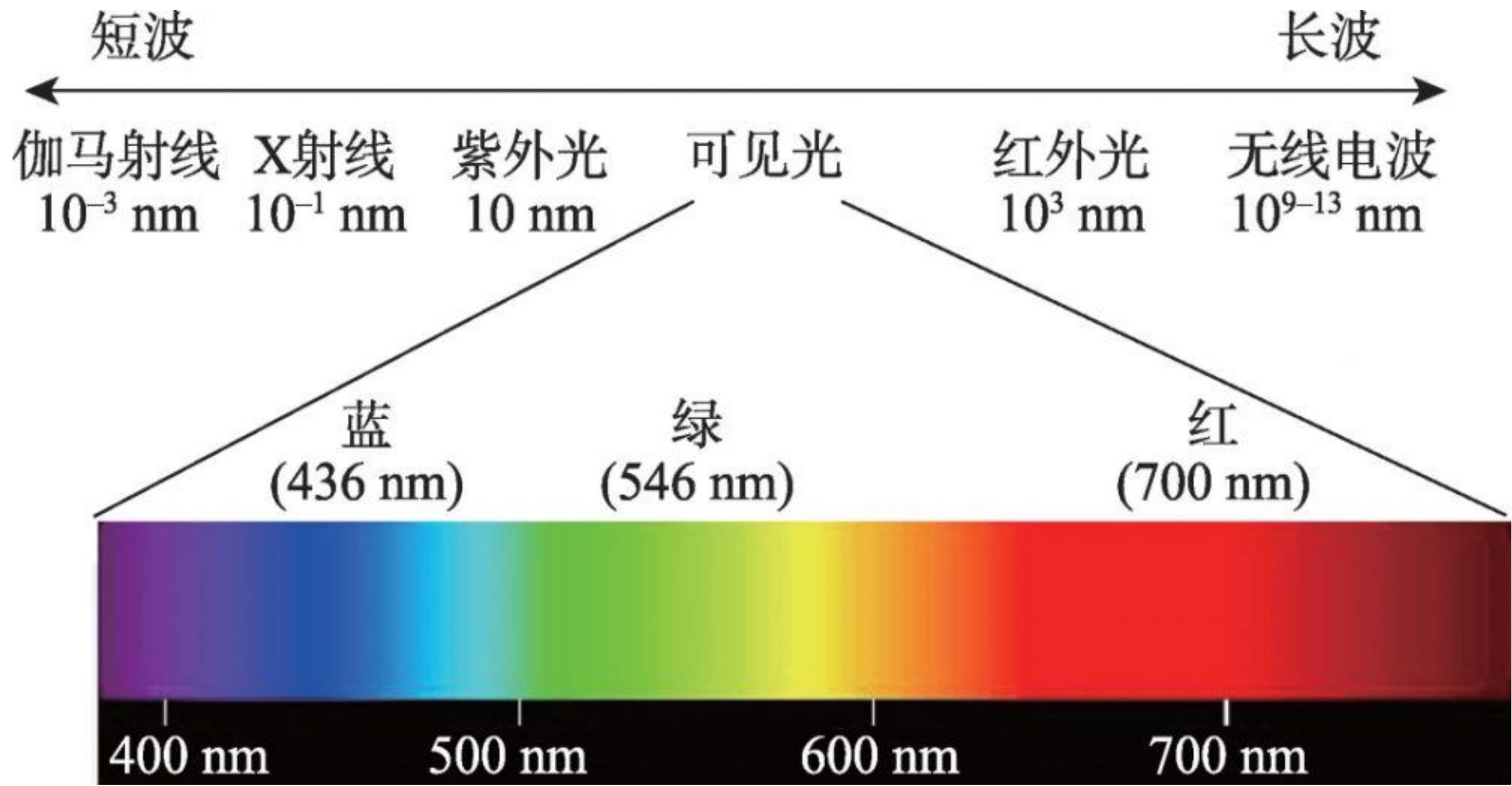
因此, 视觉现象包括视觉对**光强**, 对各种**波长**、**彩色光谱效应**, 对**物体边缘**等空间频率变化的响应, 以及视觉对**时间瞬时变化运动**的响应。

相对视敏函数

- 人眼对不同波长的光有不同的敏感度，不同波长而幅射功率相同的光不仅给人以不同的色彩感觉，而且亮度感觉也不同。
- **视敏函数：**描述人眼视敏特性的物理量为视敏函数和相对视敏函数。在相同亮度感觉的条件下，不同波长光辐射功率 $V(\lambda)$ 的倒数用来衡量人眼对各波长光明亮感觉的敏感程度。
- 对于人眼， $V(\lambda)$ 是钟形曲线。

- 视杆体和视锥体的相对视敏曲线有所不同：
 - 对视锥体亮视情况，在 $V(\lambda) = 555\text{nm}$ 时绿光亮度最敏感
 - 对视杆体暗视情况，则 $V(\lambda) = 505\text{nm}$ 时绿光最敏感





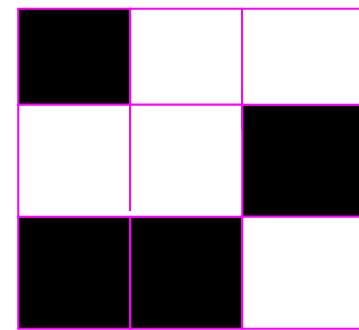
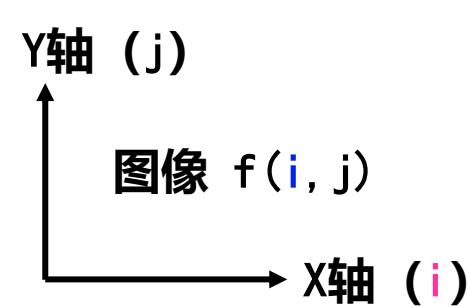
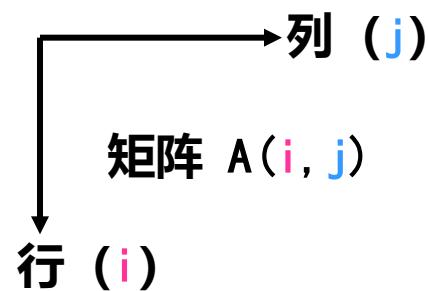
人类视觉感知能力

- 人类视觉系统在对**物体识别**上有特殊强大的功能；但在对灰度、距离和面积的绝对估计上却有某些欠缺；
- 以**传感器单元的数目**比较：视网膜包含接近130 millions 光接收器，这极大地**大于**CCD (Charge Couple Device, 电荷耦合器件) 上的传感器单元数；
- 以**每次执行运算的数目**比较：和计算机的时钟频率相比，神经处理单元的开关时间将比之大约慢 10^4 倍；

即使如此，人类的视觉系统是比计算机视觉系统要强大得多。它能**实时分析复杂景物以使我们能即时的反应。**

数字图像的概念与描述

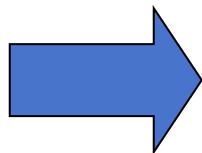
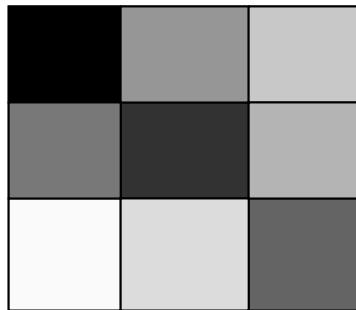
- 所谓的数字图像的描述是指**如何用一个数值方式来表示一个图像。**
- **数字图像**是图像的数字表示，**像素**是其最小的单位。
- 可以用矩阵来描述数字图像。
- 描述数字图像的矩阵目前采用的是**整数阵**，即每个像素的亮暗，用一个整数来表示。



$$I = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

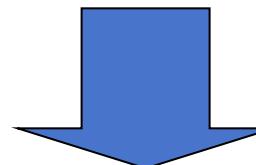
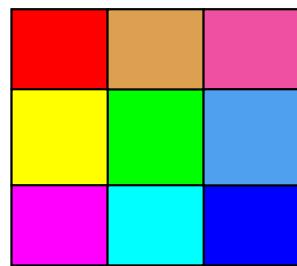
黑白图像

数字图像的概念与描述



$$I = \begin{bmatrix} 0 & 150 & 200 \\ 120 & 50 & 180 \\ 250 & 220 & 100 \end{bmatrix}$$

灰度图像：灰度级



彩色图像：RGB

$$R = \begin{bmatrix} 255 & 240 & 240 \\ 255 & 0 & 80 \\ 255 & 0 & 0 \end{bmatrix}$$

$$G = \begin{bmatrix} 0 & 160 & 80 \\ 255 & 255 & 160 \\ 0 & 255 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 80 & 160 \\ 0 & 0 & 240 \\ 255 & 255 & 255 \end{bmatrix}$$

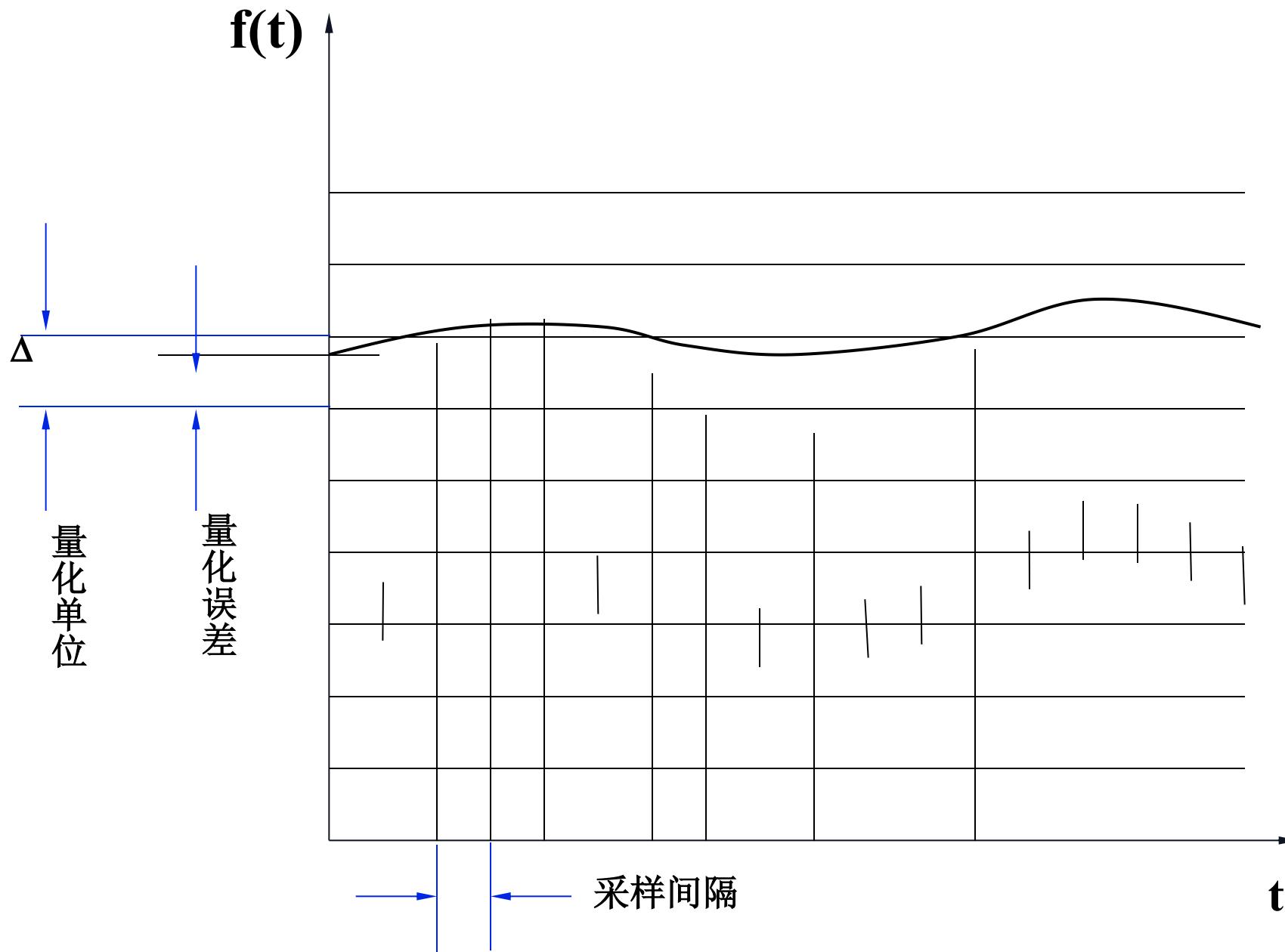
图像的数字化

- 我们日常生活中见到的图像一般是**连续形式的模拟图像**，可由一个二维连续函数 $f(x, y)$ 来描述。

其中： (x, y) 是图像平面上任意一个二维坐标点， $f(x, y)$ 则是该点颜色的深浅。
- 图像处理的方法有**模拟式**和**数字式**两种。
- 数字图像处理的一个先决条件就是**将连续图像经采样、量化（离散）化**，转换为数字图像。

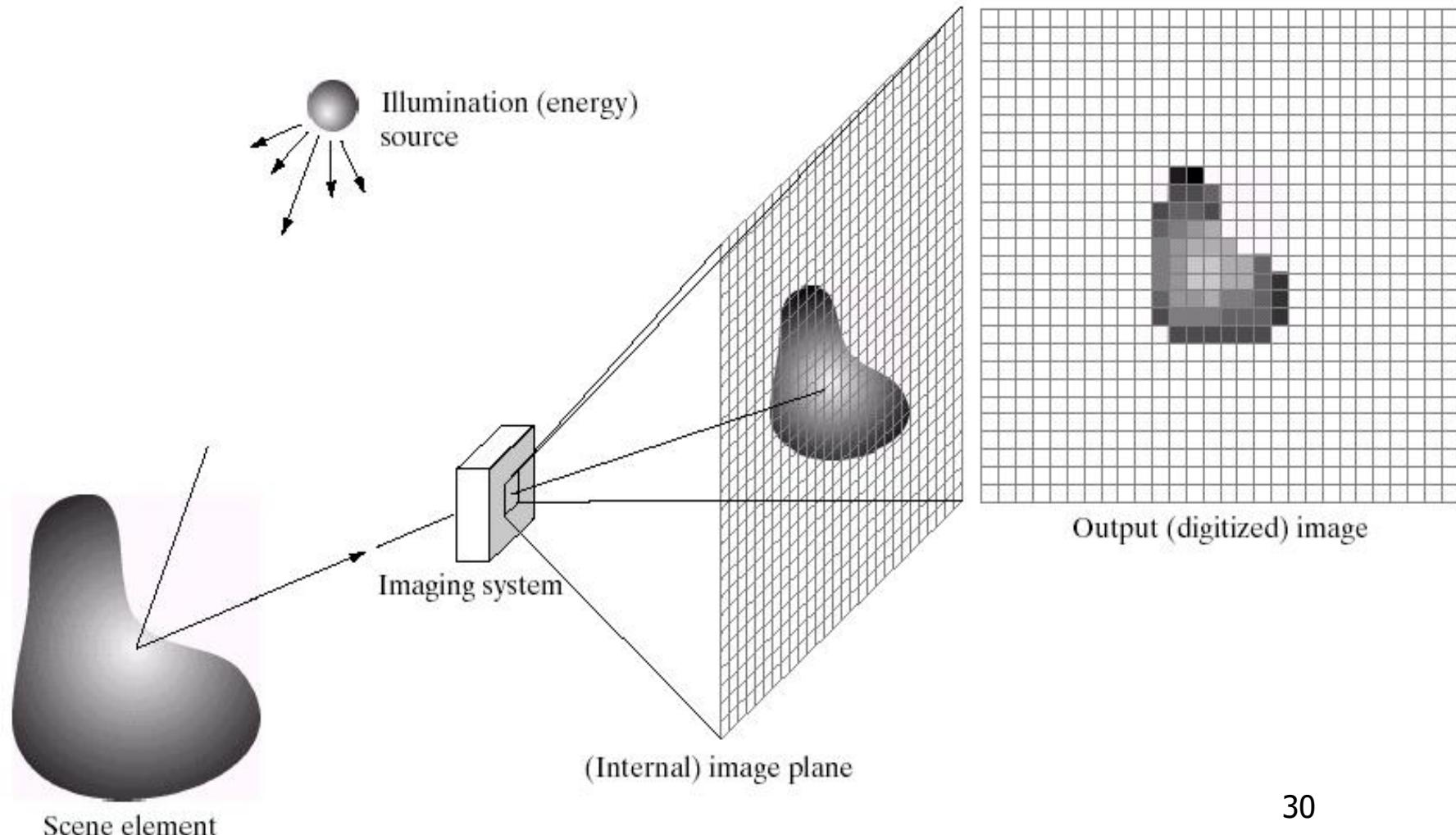
图像的数字化

- 数字化的过程（A/D转换）：是将光电传感器产生的模拟量转换为数字量，以便计算机处理；
 - 转换过程：**采样、量化、编码**；
 - 转换精度：ADC位数（模数转换器 Analog to Digital Converter）；
 - 转换速度：采样速率；
 - 量化误差。



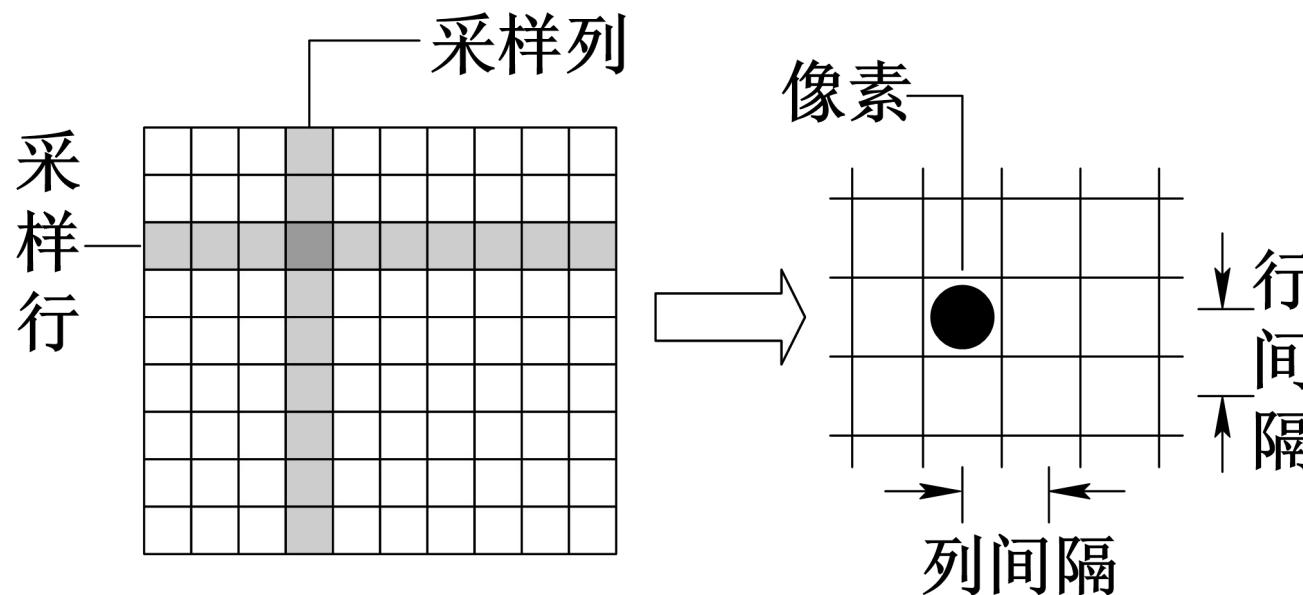
图像的数字化

- 经数字化的图像方可用计算机来处理。
- 图像 $f(x, y)$ 必须在空间上和在颜色深浅的幅度上都进行数字化：
 - **空间坐标 (x, y) 的数字化：图像采样**
 - **颜色深浅幅度的数字化：灰度级量化**



图像的数字化——采样间隔

- 采样时的注意点：**采样间隔**的选取。
- 采样间隔太小，则增大数据量；
- 采样间隔太大，则会发生信息的混叠，导致细节无法辨认。



图像的数字化——分辨率

- **分辨率是指映射到图像平面上的单个像素的景物元素的尺寸。**

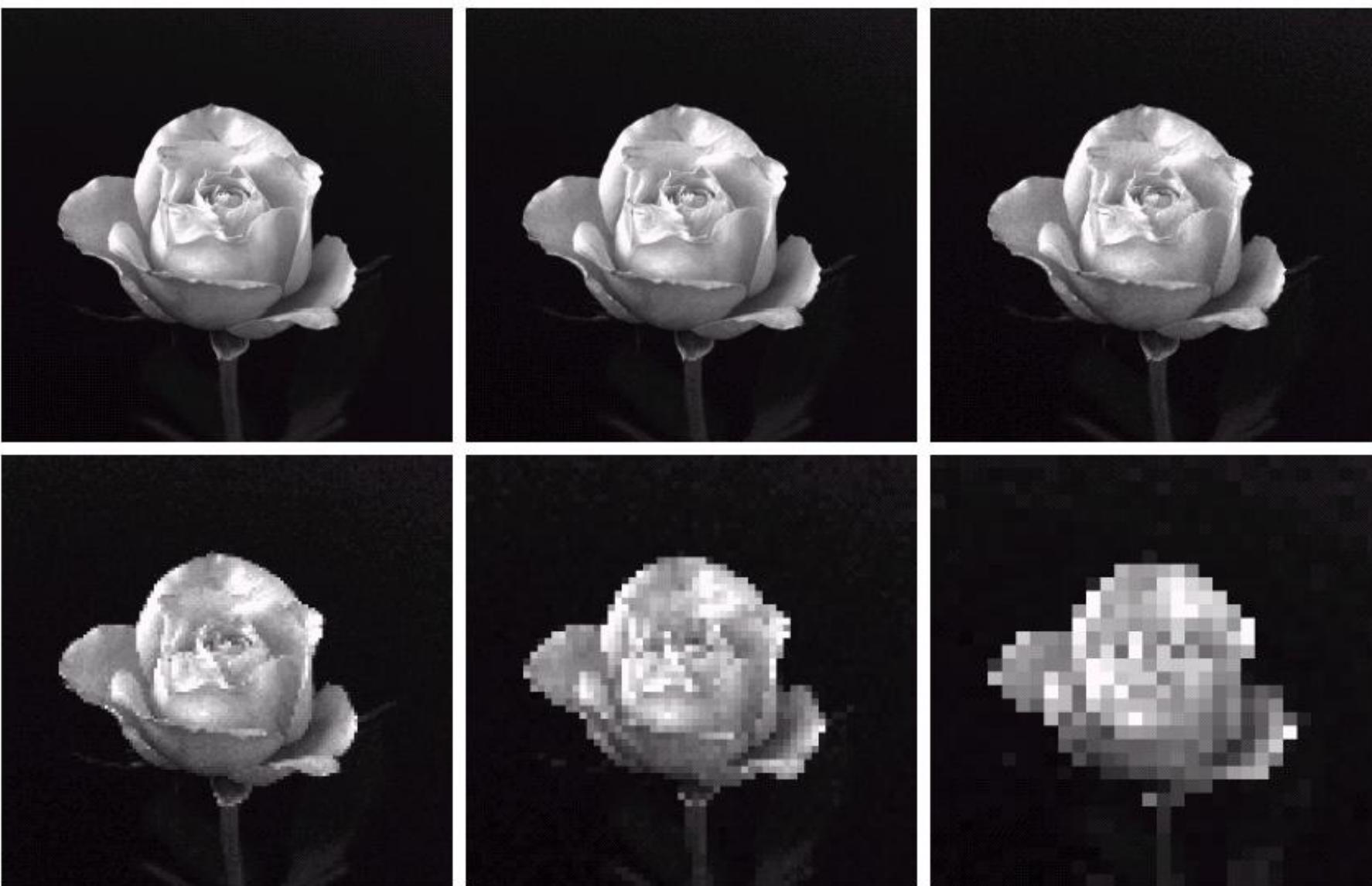
- 单位：像素/英寸，像素/厘米

(如：扫描仪的指标 300dpi)

- **分辨率指精确测量和再现一定尺寸的图像所必需的像素个数。**

- 单位：像素×像素

(如：数码相机指标1210万像素 (4000×3000))



a	b	c
d	e	f

FIGURE 2.20 (a) 1024×1024 , 8-bit image. (b) 512×512 image resampled into 1024×1024 pixels by row and column duplication. (c) through (f) 256×256 , 128×128 , 64×64 , and 32×32 images resampled into 1024×1024 pixels.

图像的数字化——量化

- **量化**是将各个像素所含的明暗信息离散化后，用数字来表示。一般的量化值为整数。
- 充分考虑到人眼的识别能力之后，目前非特殊用途的图像均为8bit量化，即采用0 ~ 255的整数来描述“从黑到白”。
- 在3bit以下的量化，会出现**伪轮廓现象**。



图像的数字化——量化

量化可分为均匀量化和非均匀量化。

1. 均匀量化: 简单地在灰度范围内等间隔量化。
2. 非均匀量化: 是对像素出现频度少的部分量化间隔取大，而对频度大的量化间隔取小。

一般情况下，对灰度变化比较平缓的部分用比较多的量化级，
在灰度变化比较剧烈的地方用比较高的采样密度。

模型量化(Model Quantization): 将深度学习模型中原本以高精度浮点数(如 FP32、FP16)存储和计算的权重、激活值等参数，转换为低比特宽度的数据类型(如INT8、INT4)，以减少模型大小、降低计算开销，并提升推理速度。

图像质量

- 数字化时，关键是要决定：
 - 采样点数 $M \times N$ (行和列)
 - 量化级别 G (灰度级数)
- 为了便于处理，采样点数 N 与量化级别 G 都为 2 的幂次方，即 $M=2^m, N=2^n, G=2^k$ (m, n, k 均为正整数)。记录一幅图像所需的字节 (byte) B 为：

$$B = M \times N \times k \div 8$$

- 一般地说，图像质量随 M 、 N 和 k 的增加而增高。

采样传感器

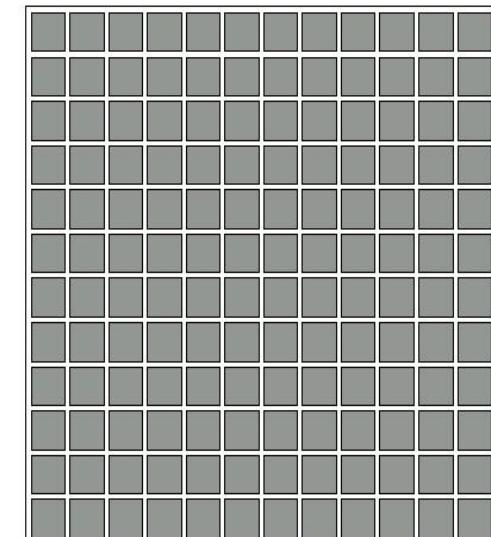
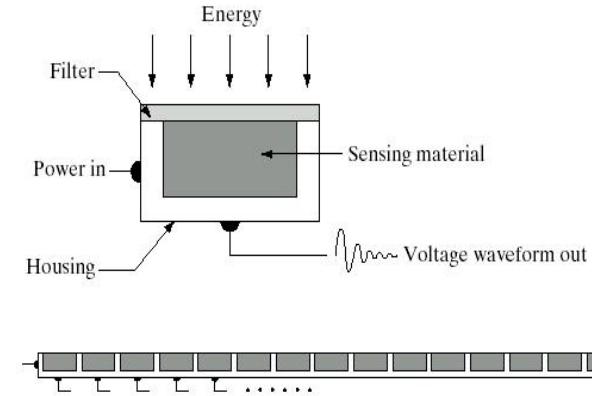
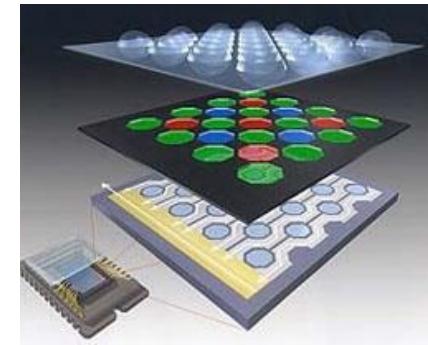
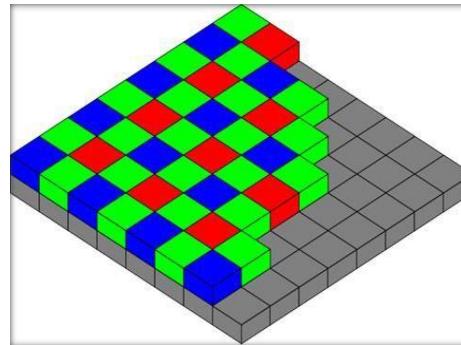
- **CCD**(Charge Couple Device): 电荷耦合器件

- 利用感光二极管(photodiode)进行光电转换，将图像转换为数字数据
- 线阵CCD、面阵CCD

- **CMOS**(Complementary Metal Oxide Semiconductor):

互补性金属氧化物半导体。

- CMOS和CCD一样都是可用来感受光线变化的半导体。通过带负电和带正电的晶体管实现基本功能。产生的电流即可被处理芯片纪录和解读成影像。
- 价格低廉、制造工艺较简单且省电，耗电量约是普通CCD的 $1/3$ 。
- 主要问题是在处理快速变化的影像时，由于电流变化过于频繁而过热。

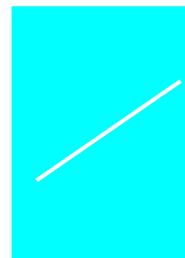


数字图像的存储格式

- 数字图像是由排成矩形点阵的像素组成的。
- 图像有不同的**编码方式**: **无压缩、无损压缩、有损压缩**。
- 图像文件通常与操作系统有关: Windows、Linux、Mac；
- 把一幅图像记录进文件时, 必须同时记录下各像素在点阵中的位置及像素的灰度值。
- 图像数据文件: **文件头+数据流**
 - 文件头是有关图像整体的信息数据块, 除记录图像的尺寸外, 还记录诸如像素的位长、图像的颜色表等有关信息。文件头之后才是图像的数据流。

数字图像的存储格式

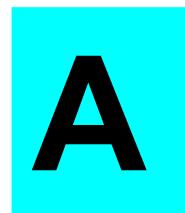
- **静态图像文件** 格式： BMP/DIB、 JPEG、 Tiff、 Gif、 WMF。
- **动态图像文件** 格式： MPEG4、 SWF (Shockwave Format) 、 AVI、 n AVI、 rm、 rmvb、 MOV、 MKV.....
- **点阵图 (Bit)** : 位图
- **矢量图 (Vector)** : 向量图，是用称之为矢量的直线、曲线、多边形和填充的色块来描绘图形的。



类型: line

起点坐标 xx.xxx,yy.yyy 终点坐标 xx.xxx,yy.yyy

颜色、线型、线宽



类型: 文字

起点坐标 xx.xxx,yy.yyy

颜色、字体、字型、字号

点阵图



放大后的点阵图



矢量图



放大后的矢量图

