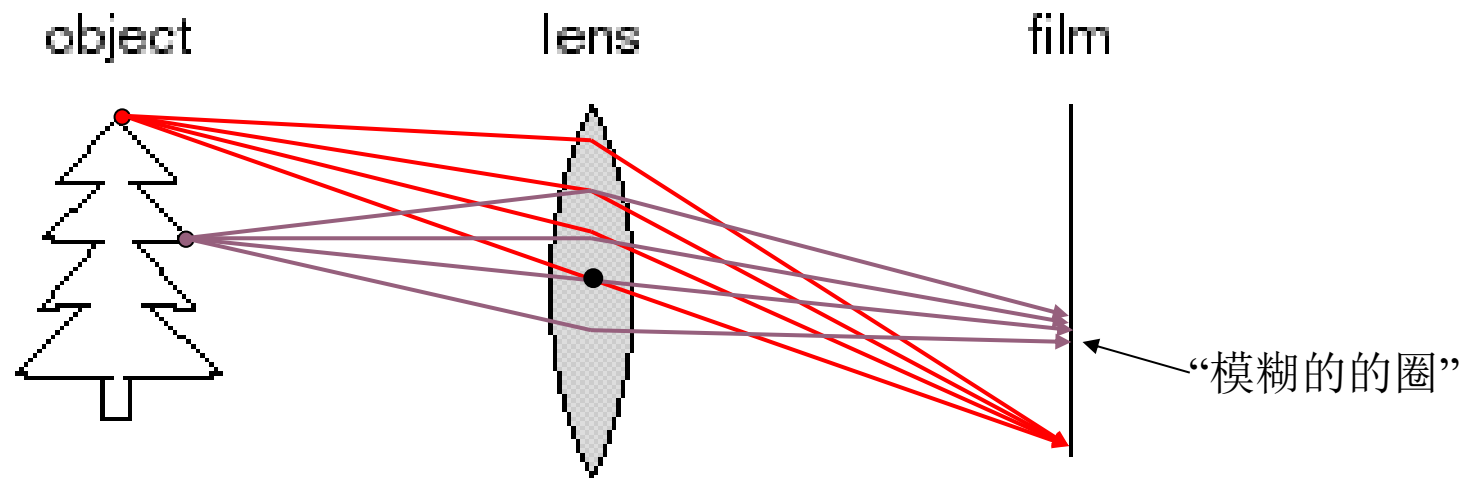


回顾—— 视觉系统与视觉误差

镜头



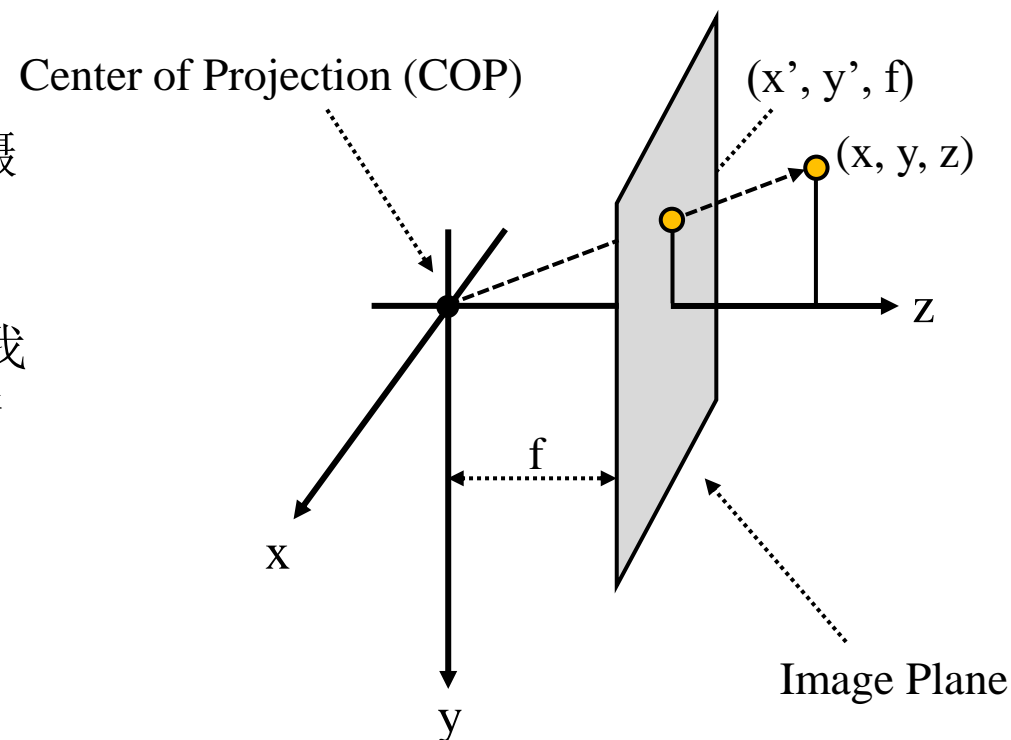
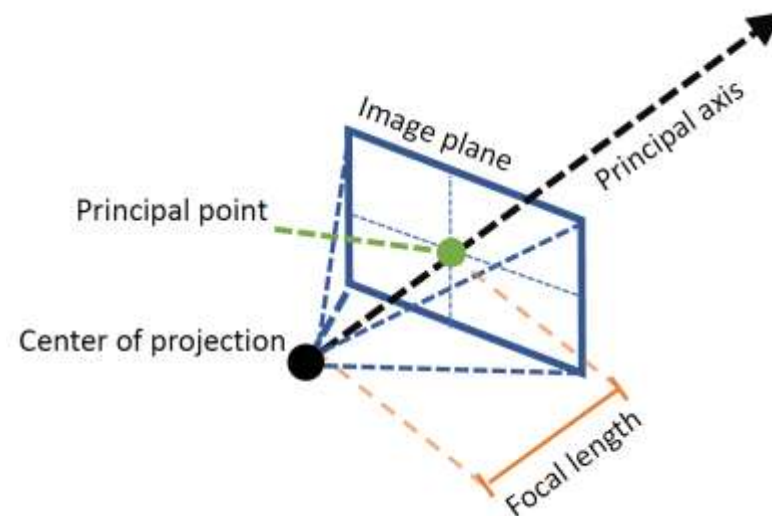
镜头通过折射光线来实现光的聚焦

- 会产生“焦点”和“焦距”
 - 其他的点会产生模糊的“圈”
- 改变镜头可以改变焦距

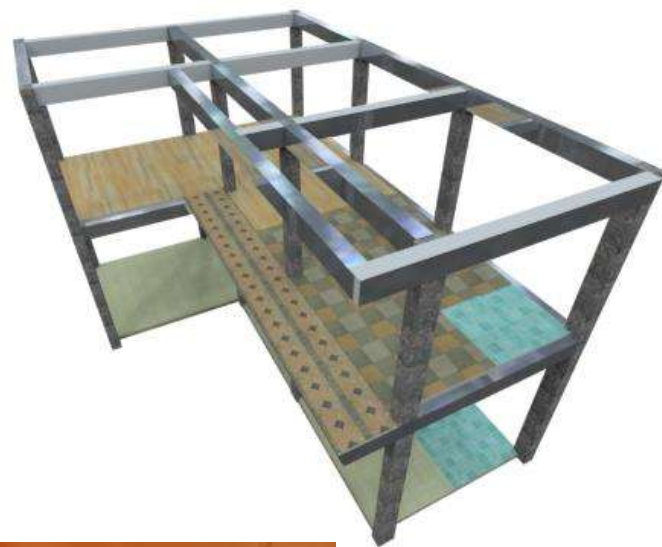
Modeling projection: 投影模型

• 坐标系统

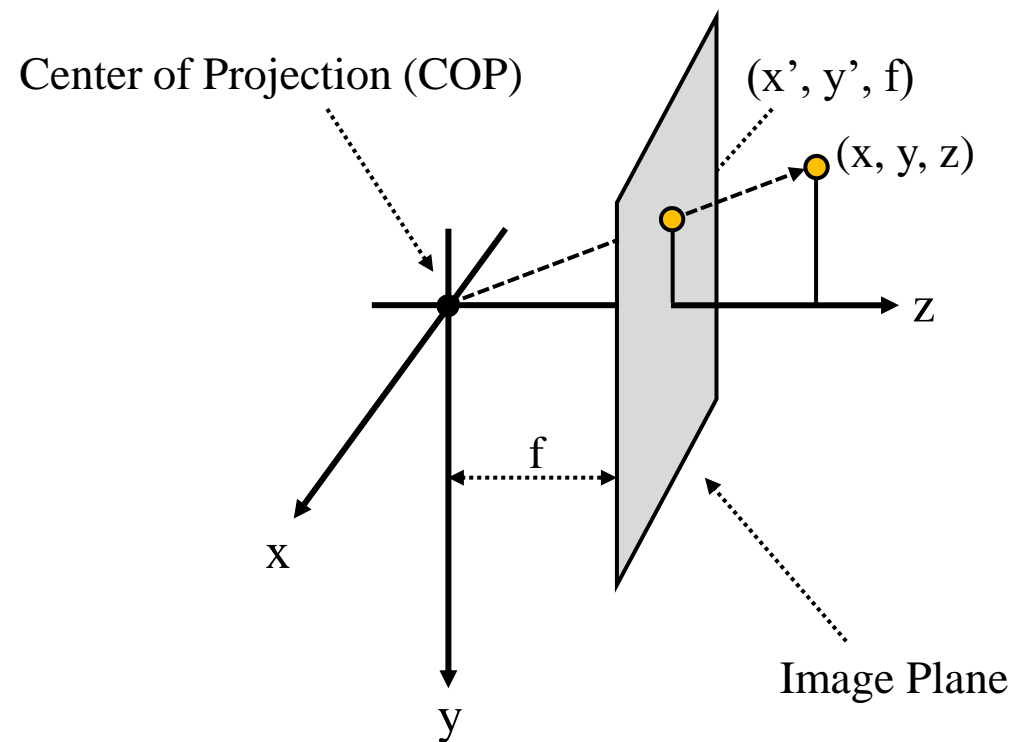
- 使用针孔相机做近似，将三维世界的信息投影到感光材料上，从而形成二维图像
- 用小孔当作坐标原点 (Center of Projection, or COP)
- 图像平面 (或 Projection Plane 投影平面) 针孔摄像机的感光材料或图像平面
- 为了方便计算图像平面与z轴垂直
- 针孔相机应朝向z轴的反方向，为了方便表示我们使用投影（透视投影）而不是相机模型来举例



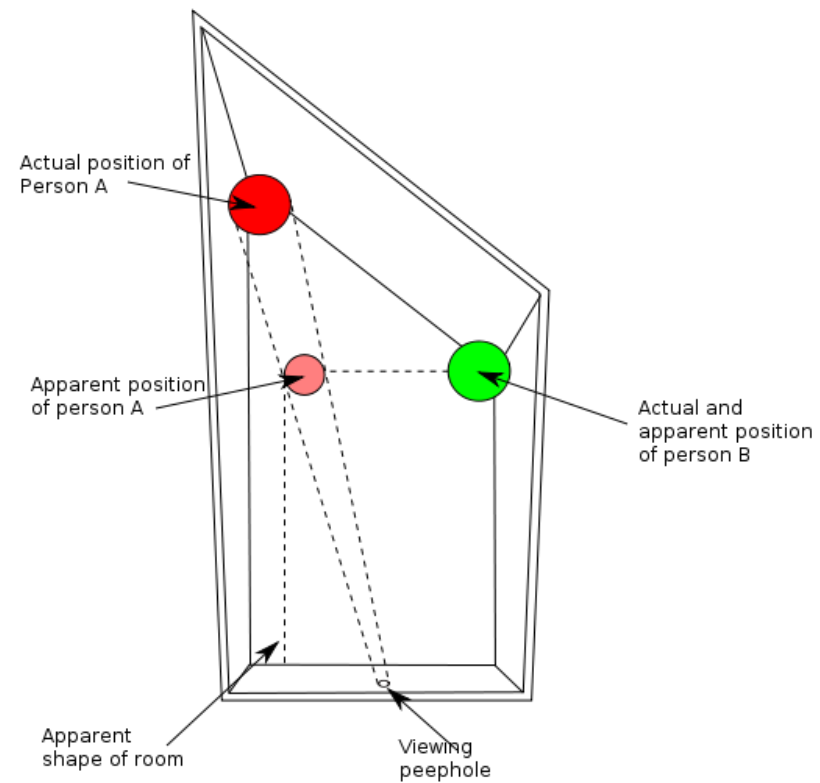
透视投影



错觉原理



Ames Room: 猜猜怎么做到的？



不同焦距拍摄效果不同



广角（短焦）



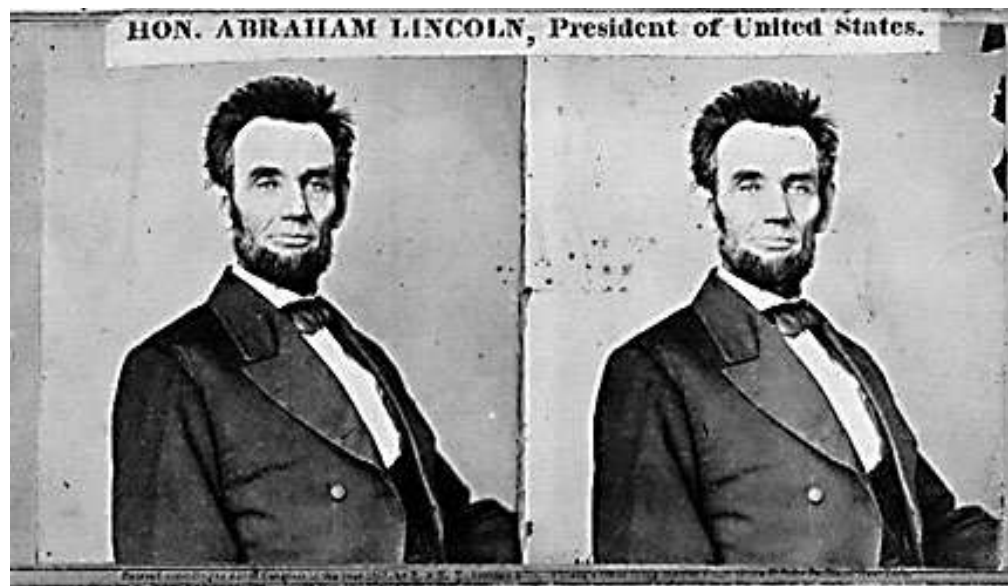
标准镜头



长焦镜头

知道为什么
专业摄影师
都用大长焦
了吧！

双目视图



- 给出两张不同视角拍摄的同一个人物
 - 我们怎么计算每个点的景深呢？
 - 基于同一个点在两张图中移动了多远！

Virtual Reality Video



Anderson, et al. *Jump: Virtual Reality Video*. SIGGRAPH Asia 2016.



Broxton, et al. *Immersive Light Field Video with a Layered Mesh Representation*. SIGGRAPH 2020.



“感知”

从 几何模型：基于几何原理和相对位置关系的模型
到 统计模型：基于数据分布和统计推断的模型

图像图形学的最终目标

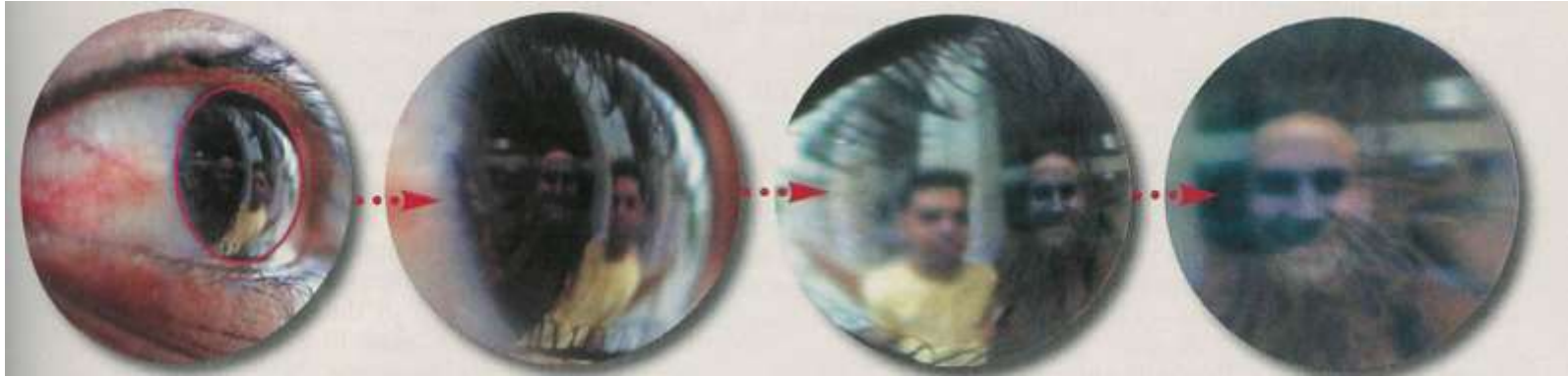
- Forensics: 取证



Source: Nayar and Nishino, "Eyes for Relighting"



Source: Nayar and Nishino, "Eyes for Relighting"



Source: Nayar and Nishino, "Eyes for Relighting"

图像图形学的最终目标



<https://www.bilibili.com/bangumi/play/ep28950?t=717>

《攻壳机动队》

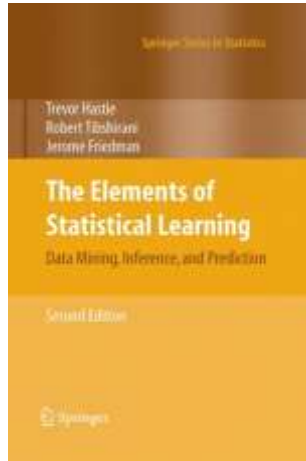
感知与理解



<https://www.bilibili.com/video/BV1EV411y7g4>

《神探夏洛克》

Pointers



Useful book (Free too!):
The Elements of Statistical Learning
Hastie, Tibshirani, Friedman

<https://web.stanford.edu/~hastie/ElemStatLearn/>



Useful set of data:
UCI ML Repository

<https://archive.ics.uci.edu/ml/datasets.html>

“感知”

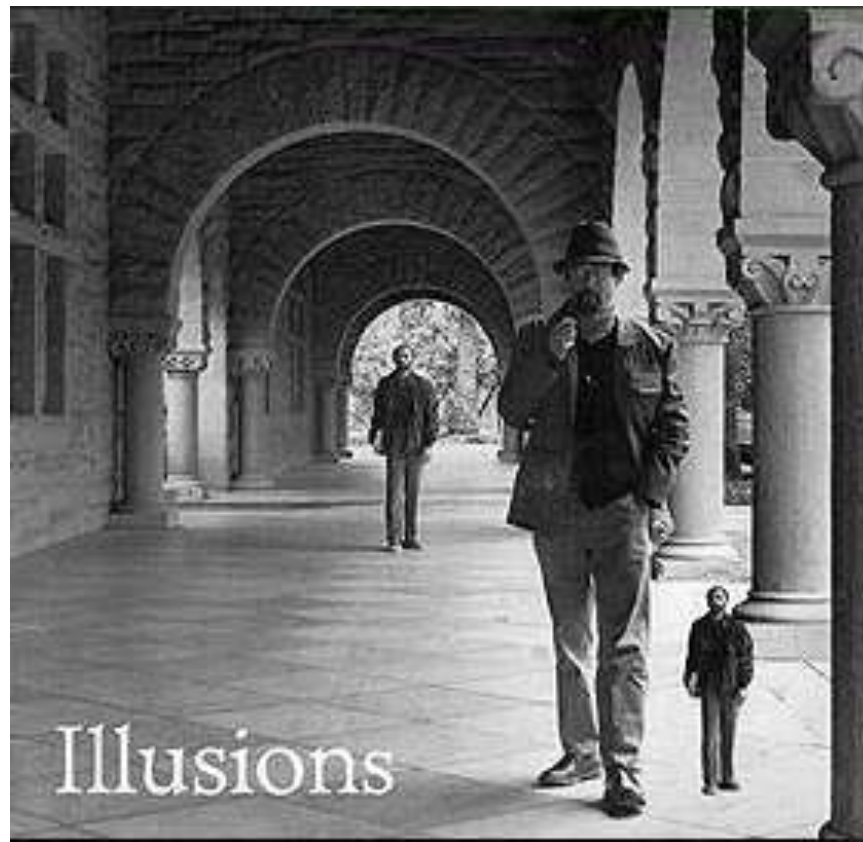
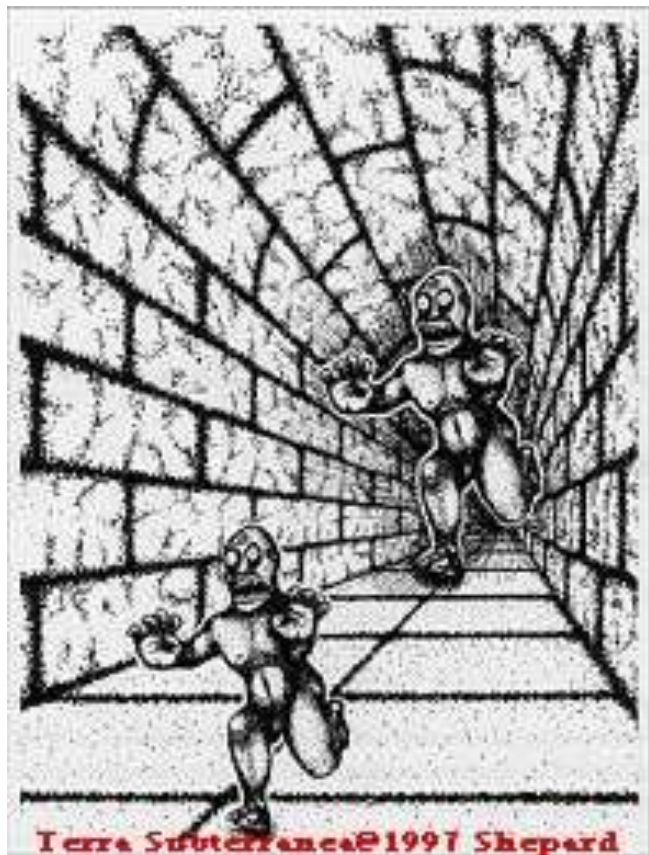
为什么 需要 统计模型？

让我们从自己的感知系统开始！

视觉系统存在错觉——透视几何系统



视觉系统存在错觉——透视线索



蓝黑还是白金？



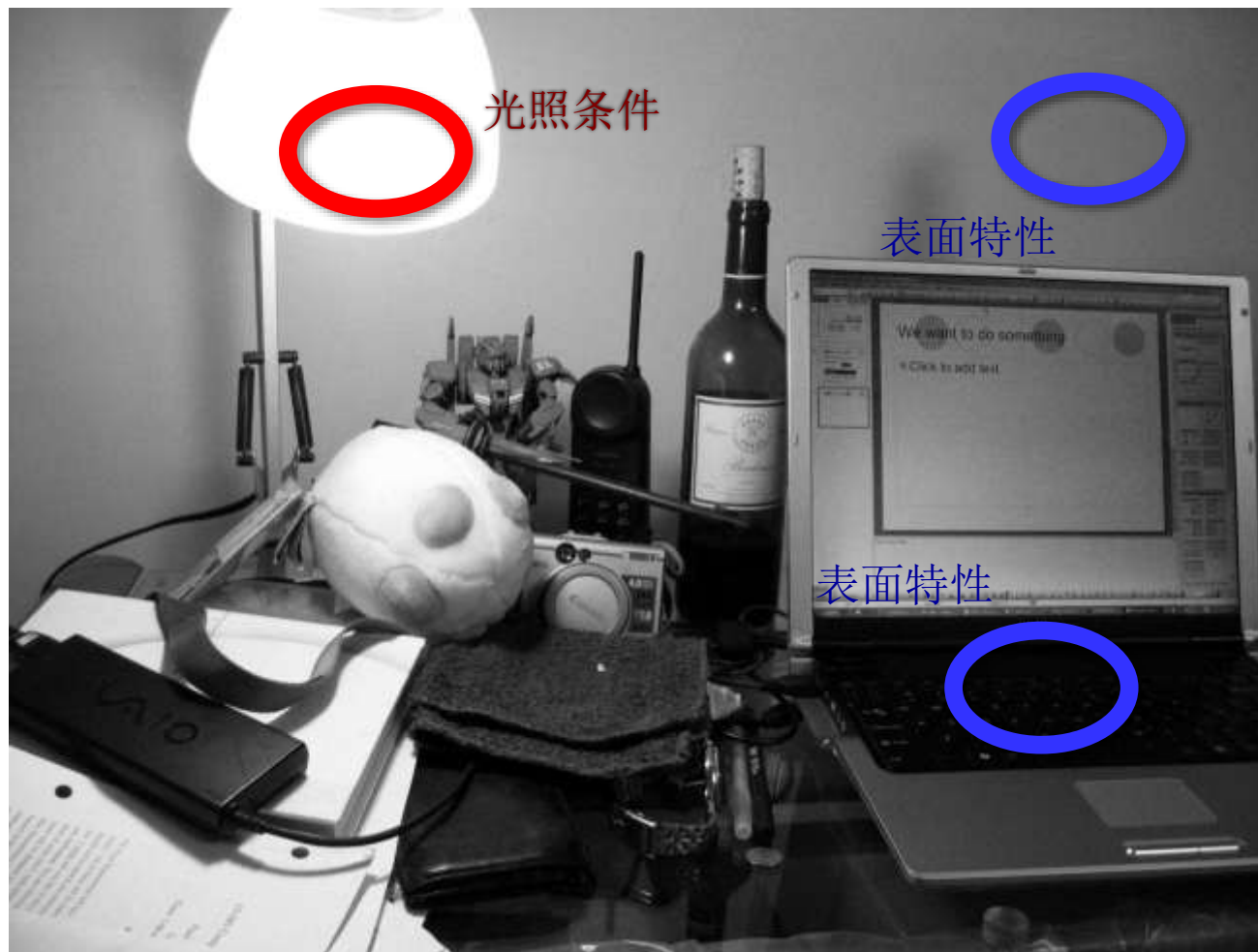
消失的圆点



With left eye shut, look at the cross on the left. At the right distance, the circle on the right should disappear (Glassner, 1.8).

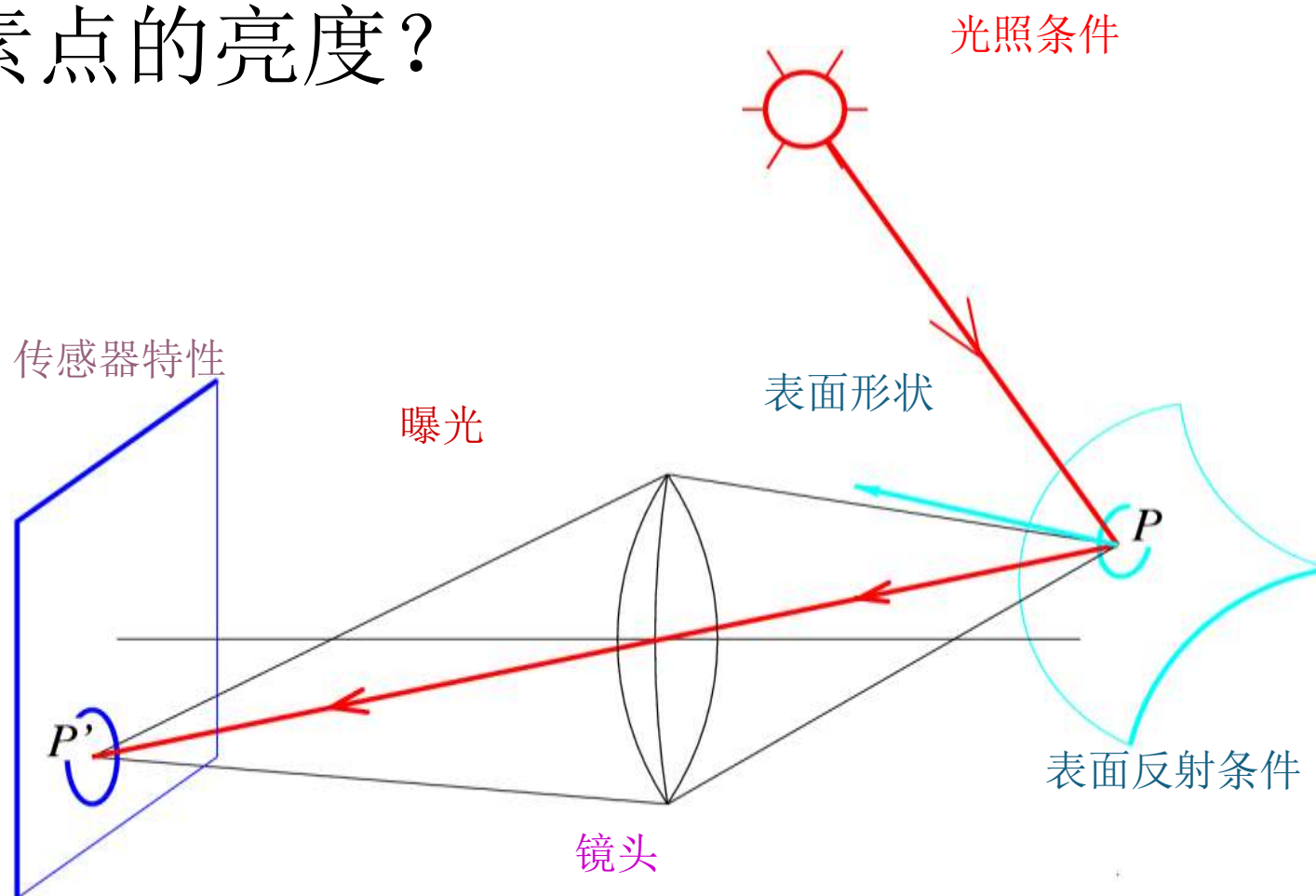
Radiometry: 辐射度量

- 什么决定了像素点的亮度?
- 相机参数
- 光照条件: 光源的强度和方向会直接影响像素点的亮度。
- 物体表面特性: 物体表面的反射率和漫反射性质会决定光线被多少吸收和多少反射到观察者或相机。



辐射度量

- 什么决定了像素点的亮度？



什么是光？

光的感知

- 怎样把辐射转换为“色彩”？
- 我们能看到哪些颜色？

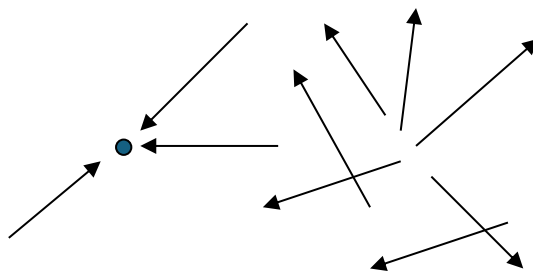
光是一种电磁辐射，它属于电磁波谱的一部分

- $R(\lambda)$ 代表其能量 (单位为watts)
 - λ 是波长



光场

- 光场是指在空间中传播的电磁辐射波动。它描述了光波的传播方向、波长和振幅。

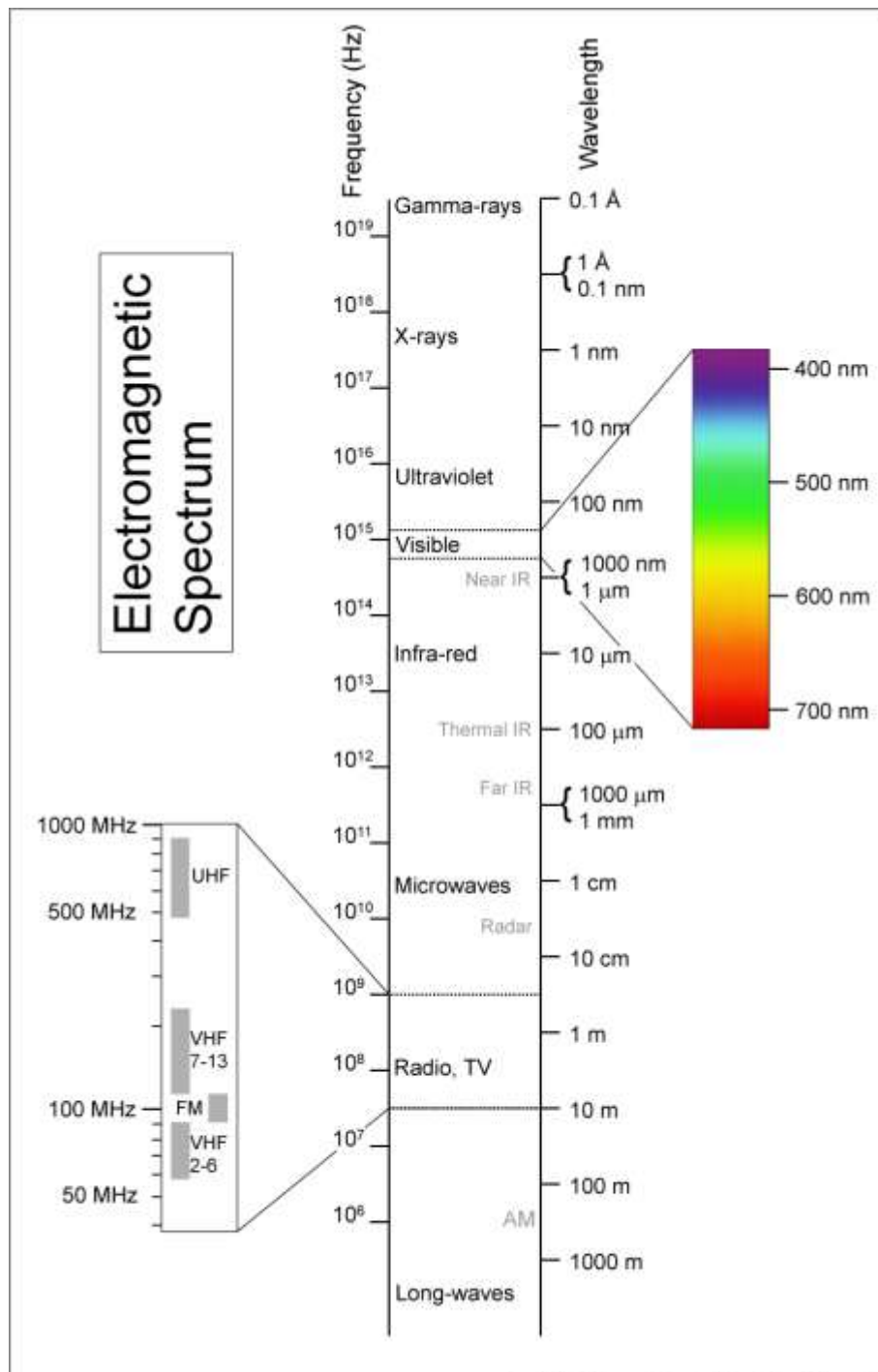


plenoptic function 定义描述光的函数:

$$R(X, Y, Z, \theta, \phi, \lambda, t)$$

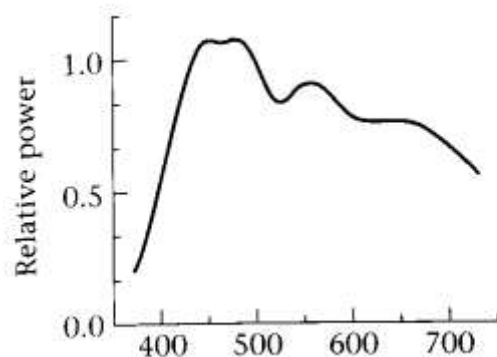
可见光

我们只能看到一部分波段的光辐射

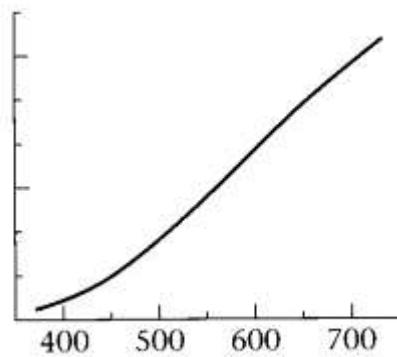


光谱

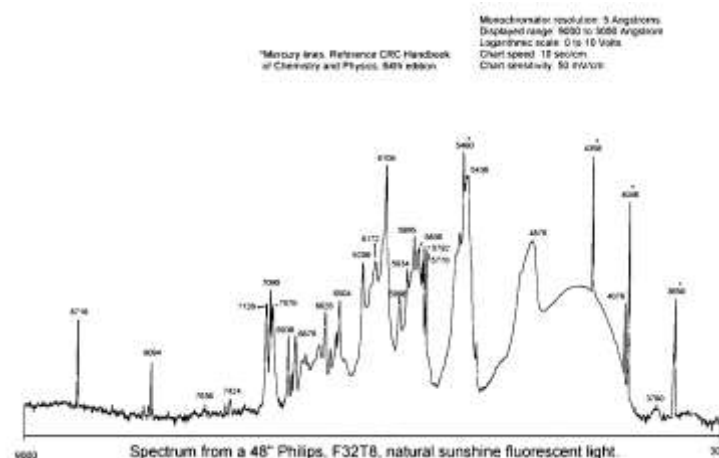
- 我们看到的外表是可见光对应的光谱
 - 光谱对应的是每个波长的强度



日光



钨丝灯

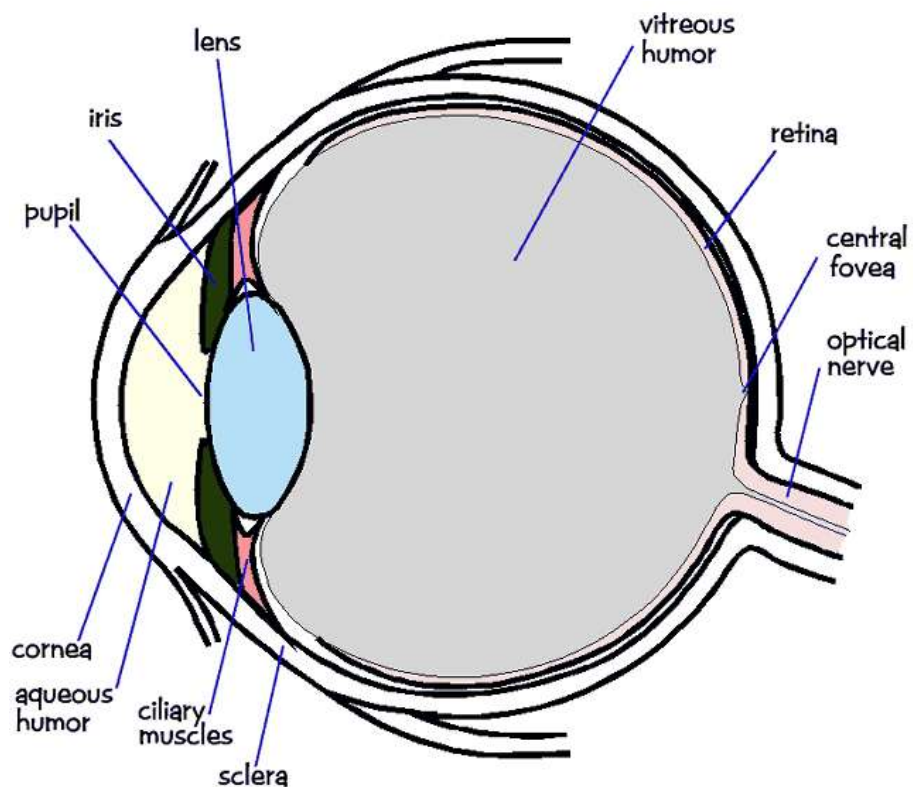


荧光灯

颜色就是从这种光谱通过复杂的形式转换过来的

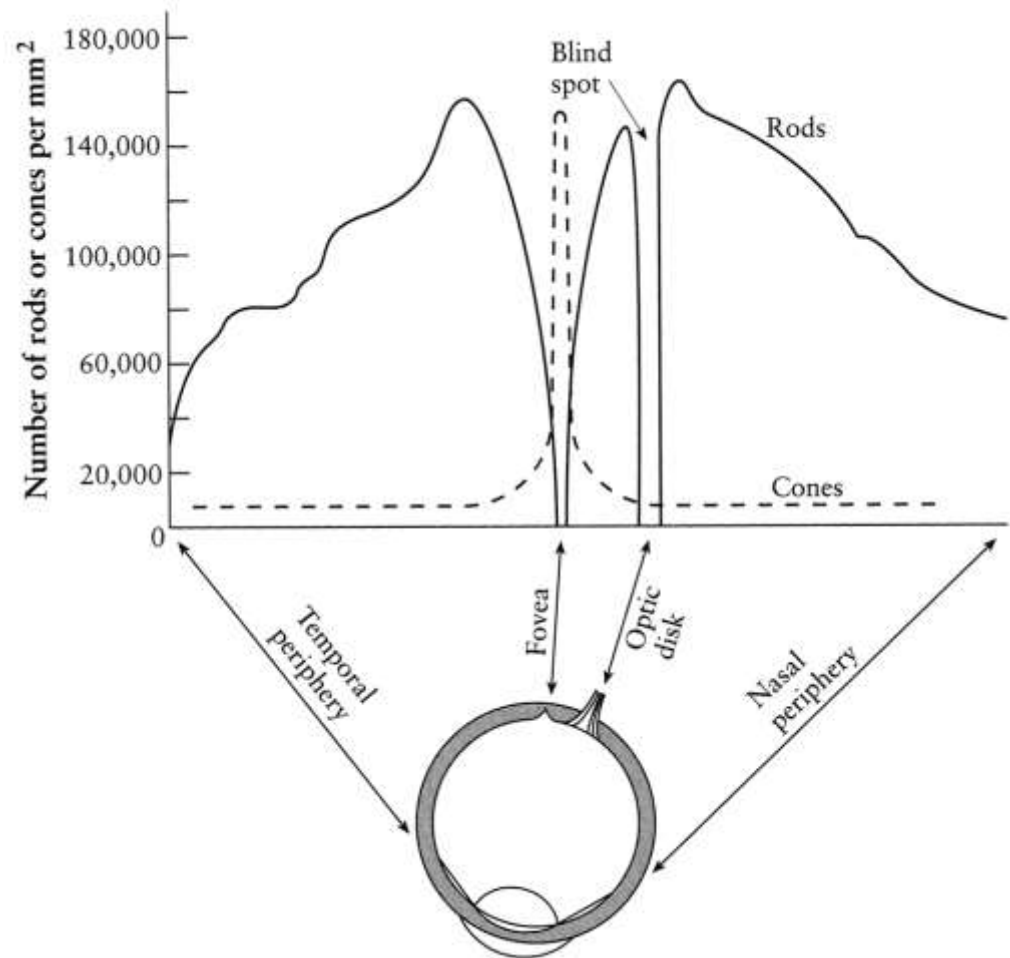
人类的视觉系统

- 颜色感知
 - 光打到视网膜上，由感光细胞捕捉到
 - Rods 视杆 和 cones 视锥
 - 这些感光细胞把光谱转换为颜色



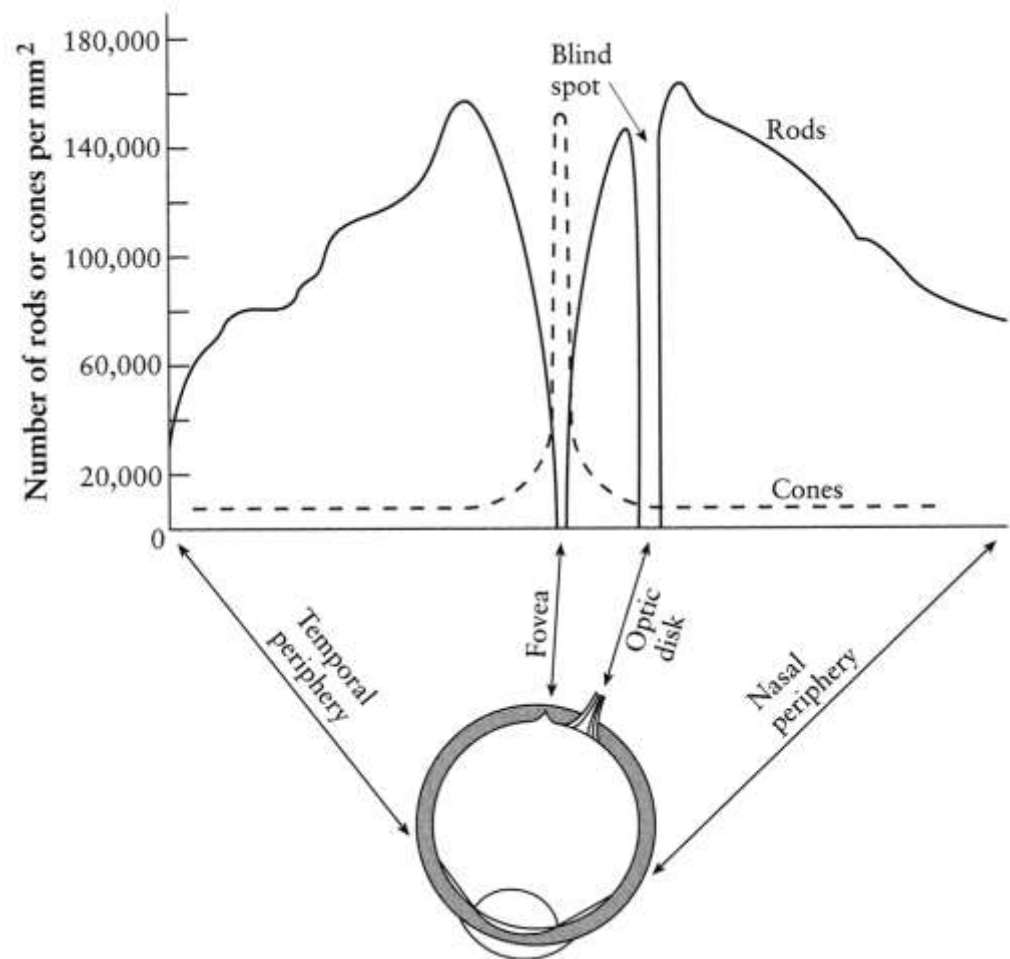
视杆与视锥

- 视杆与视锥是不均匀地分布在视网膜上的
 - 视杆对光的强度更敏感，视锥对颜色更敏感。明亮光照下，视锥细胞主导，而在暗光下，视杆细胞起主导作用
 - **Fovea 中间凹**: 视网膜上的一个小区域(1 or 2°)，包含大量视锥细胞（**没有视杆**），特别适用于高分辨率的颜色感知。它在颜色识别和高分辨率视觉中发挥关键作用。
 - 不在中央凹区域的部分，对于感知低光强度或运动检测非常有用，但视力锐度较低。



盲点

- 神经盘是视网膜上的一个区域，位于视网膜背部，是视觉信息传输的起点。它是视神经纤维束聚集的地方，这些纤维将视觉信号传送到大脑的视觉中枢。
- 视神经盘上没有感光细胞，因此在视觉场景中，与视神经盘对应的区域被视为视野中的盲点。这是我们视觉系统中的一个盲区，无法感知光或图像。



Recall: 消失的圆点



With left eye shut, look at the cross on the left. At the right distance, the circle on the right should disappear (Glassner, 1.8).

亮度、对比度

- 亮度与周围区域有光
 - 对比度: 图像或场景中不同区域之间的明暗差异。通常通过比较相邻像素的亮度差异。高对比度的区域中的像素之间有明显的亮暗差异。



- 光照恒定: 在视觉感知中, 恒定性是指在不同光照条件下, 物体的颜色和亮度看起来保持相对恒定。

人类的光感系统是动态的

- 我们的视觉系统有一个很大的**动态**的区间
 - 我们可以同时感受亮的和暗的物体
 - 我们能感受**指数级**的不同光强
 - 我们的视觉系统可以**自适应**调节

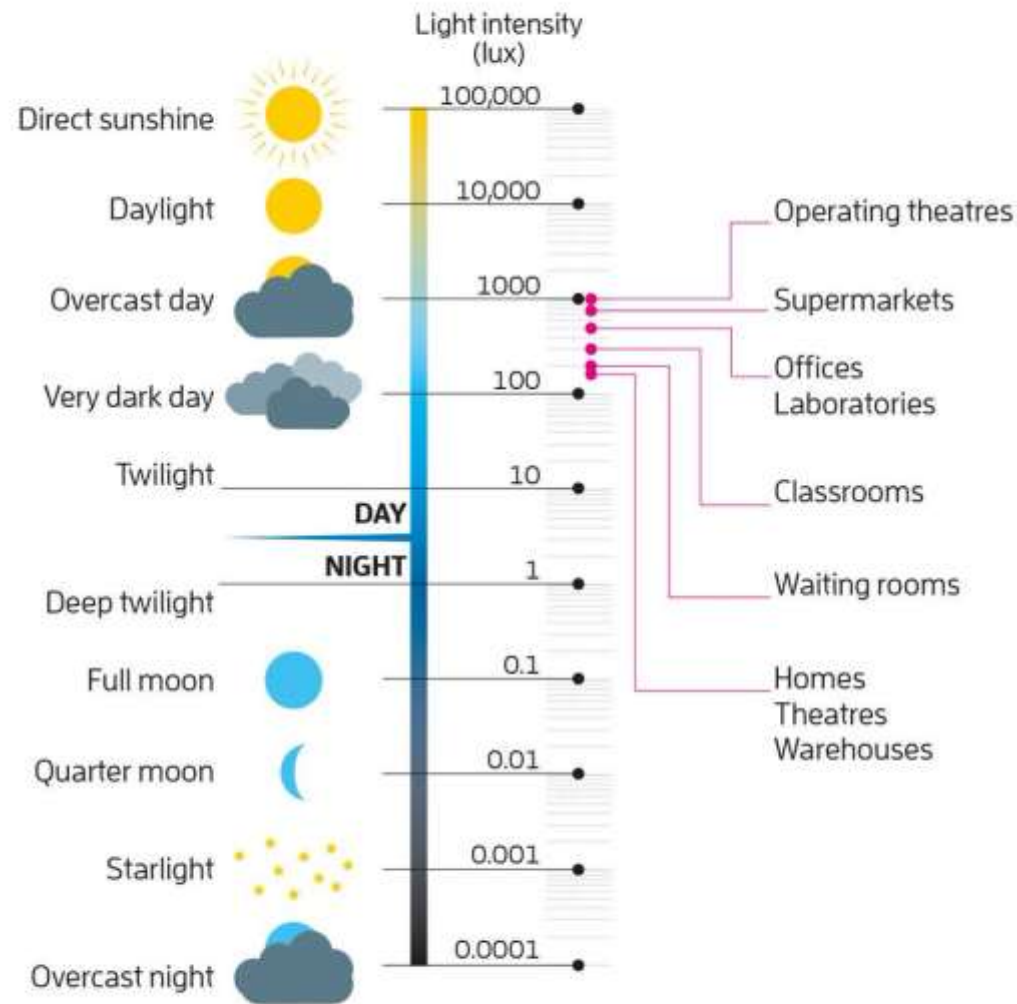
动态视觉系统

一张白纸，在明亮的室外和在没有月亮的晚上，亮度相差1,000,000,000倍

但是在一个场景内，我们可以轻松感受10倍左右的亮度差别

The light in our lives

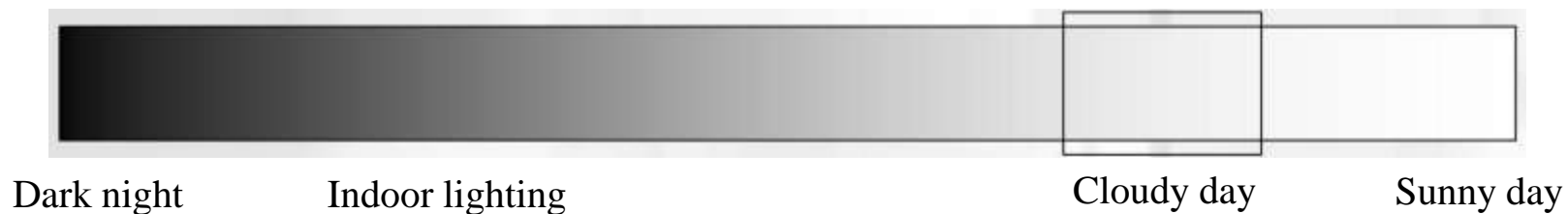
Even the brightest indoor spaces are dim compared with the outdoors in daylight



SOURCE: NATIONAL OPTICAL ASTRONOMY OBSERVATORY

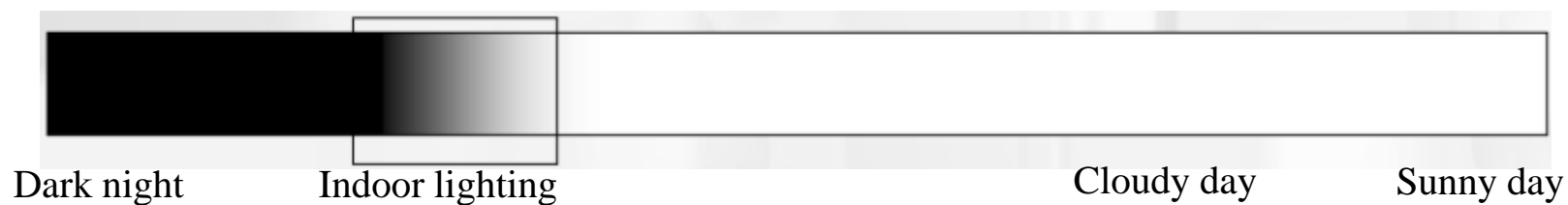
<https://threader.app/thread/1134003178515701762>

动态视觉系统



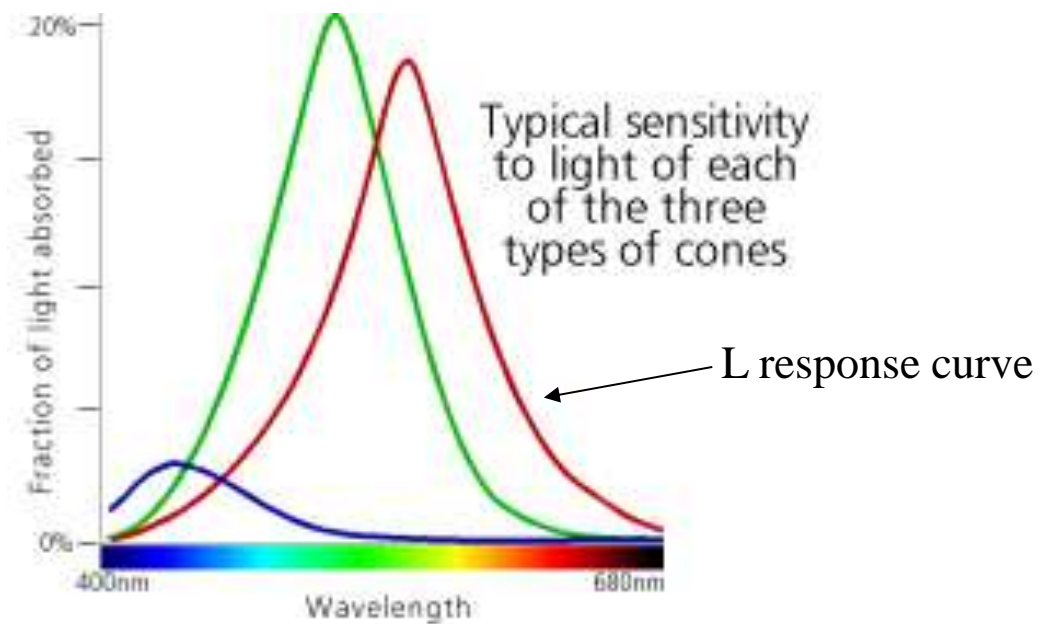
如果我们一直对整个范围都非常敏感，那么我们将无法在场景中区分亮度级别。

视觉系统通过限制其响应的“动态范围”以匹配当前的整体或“环境”光水平来解决这个问题。

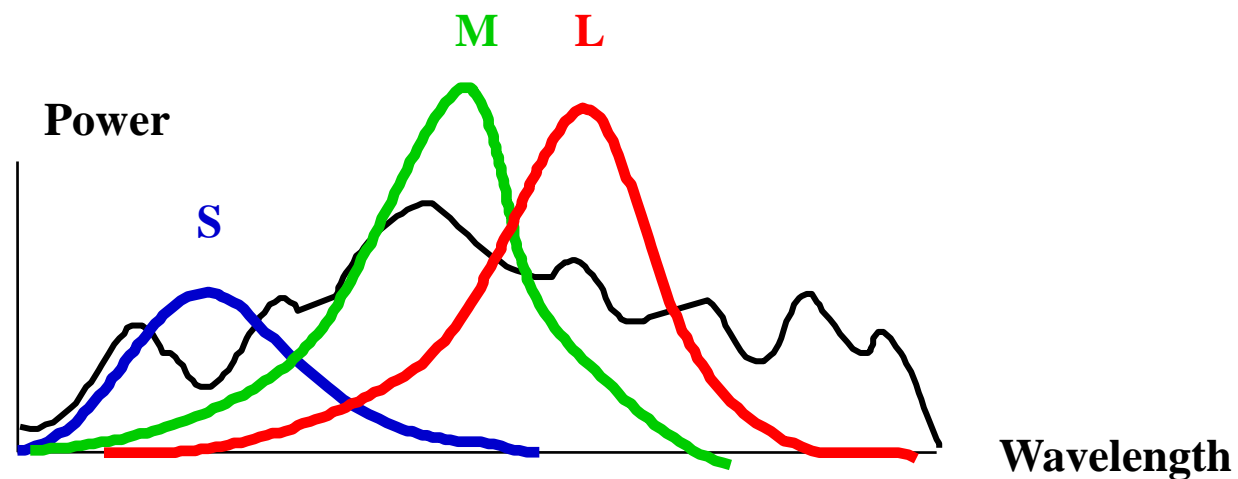


颜色感知

- 三种色锥感受颜色
 - 每种色锥对光谱中的区域更敏感
 - 这些区域相交
 - Short (S) 对应 蓝色
 - Medium (M) 对应 绿色
 - Long (L) 对应 红色
 - 总体而言，我们对红色和绿色更敏感
 - 但每个人的敏感度不同，也会随年龄变化
 - 色盲的原因是至少一种色锥出现了退化

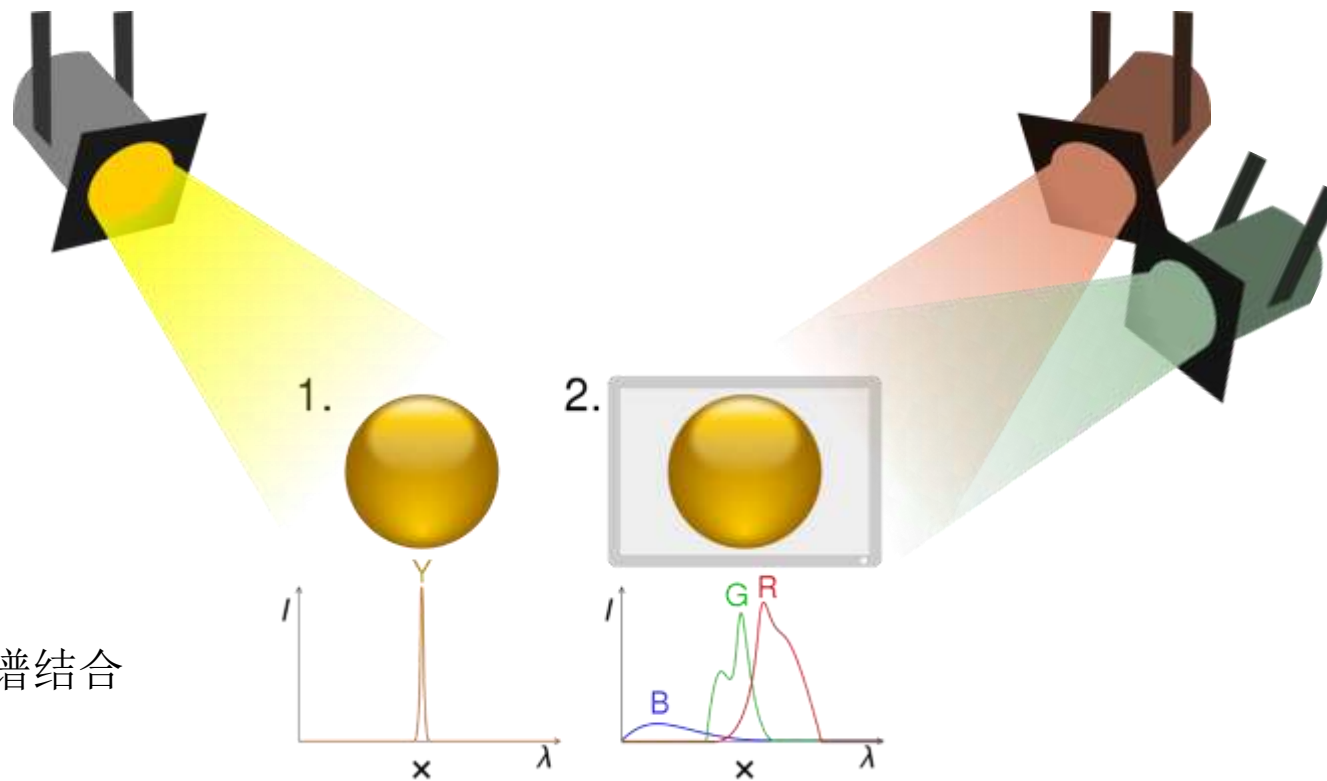


颜色感知



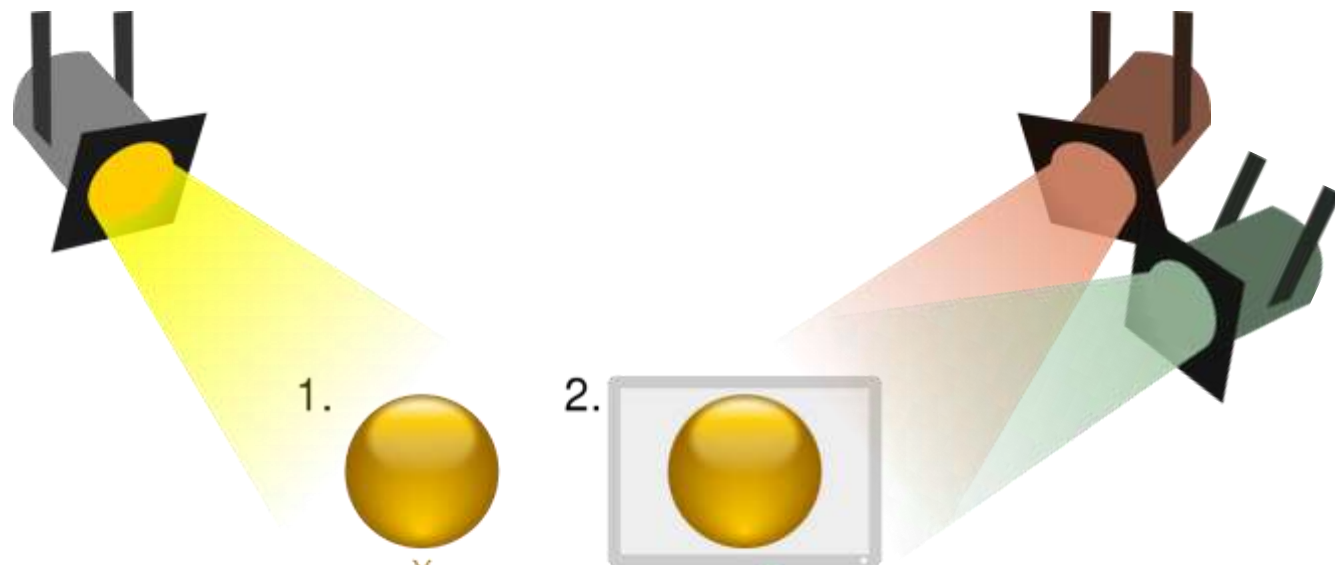
- 视杆与视锥根据光谱感受颜色
 - 在波长上做强度的积分
 - 每个视锥得到一个数字
 - Q: 我们可以直接拿视锥的感受强度来表示光谱吗?
 - A: 不行, 会丢失信息
 - 不同的光谱在某些情况下可能完全无法被分辨
 - 这种光谱被称作 **metamers** 条件等色

条件等色

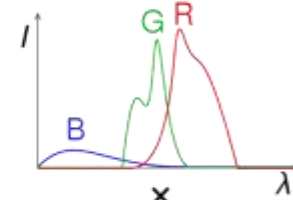
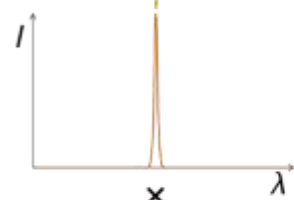


1. 与光源的光谱结合

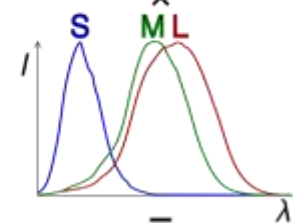
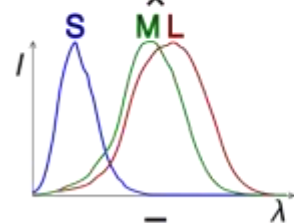
条件等色



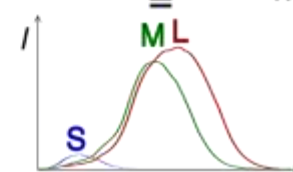
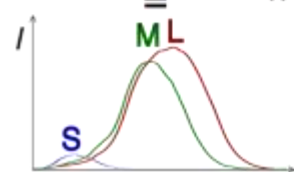
1. 与光源的光谱结合



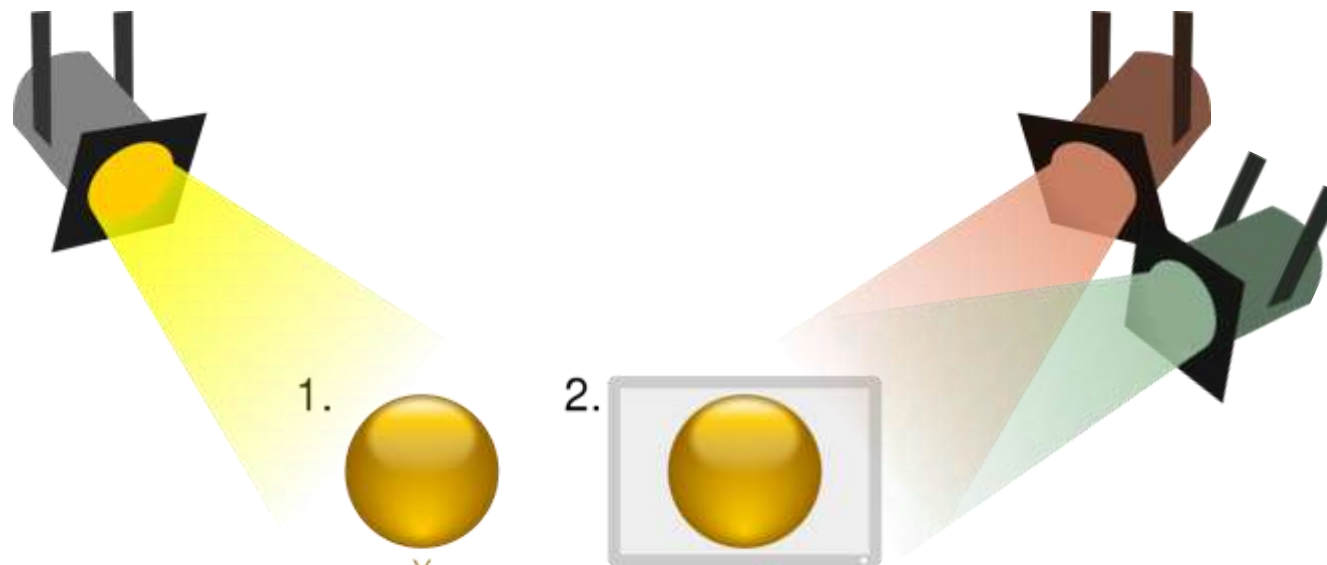
2. 光锥敏感度(S, M, L)



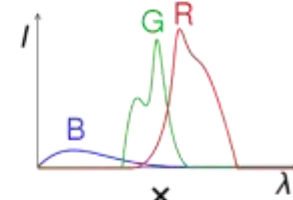
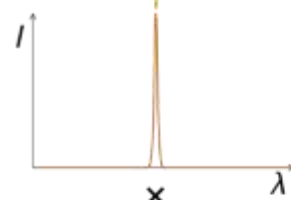
3. 1与2结合



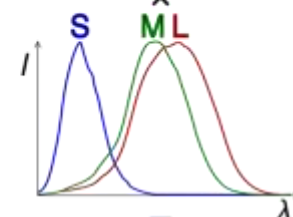
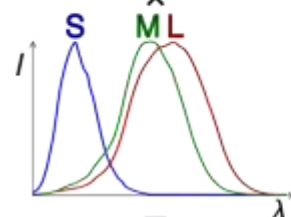
条件等色



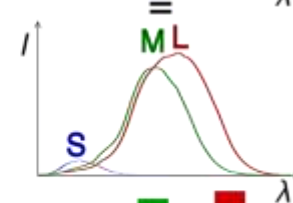
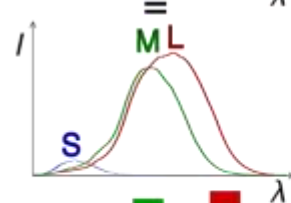
1. 与光源的光谱结合



2. 光锥敏感度(S, M, L)



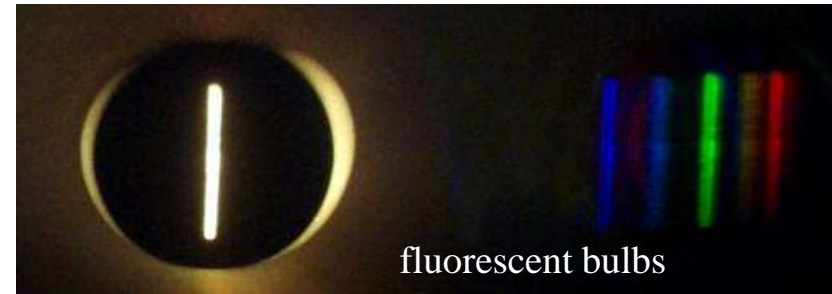
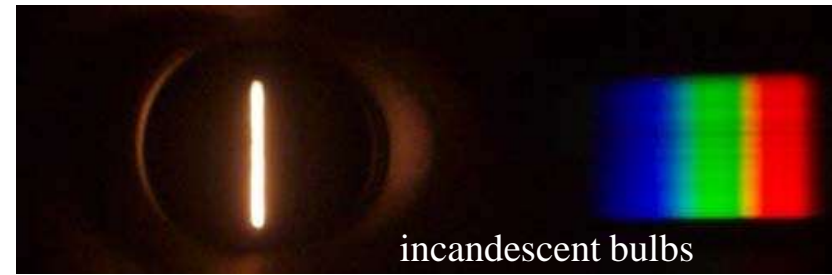
3. 1与2结合



4. 观测颜色：黄色

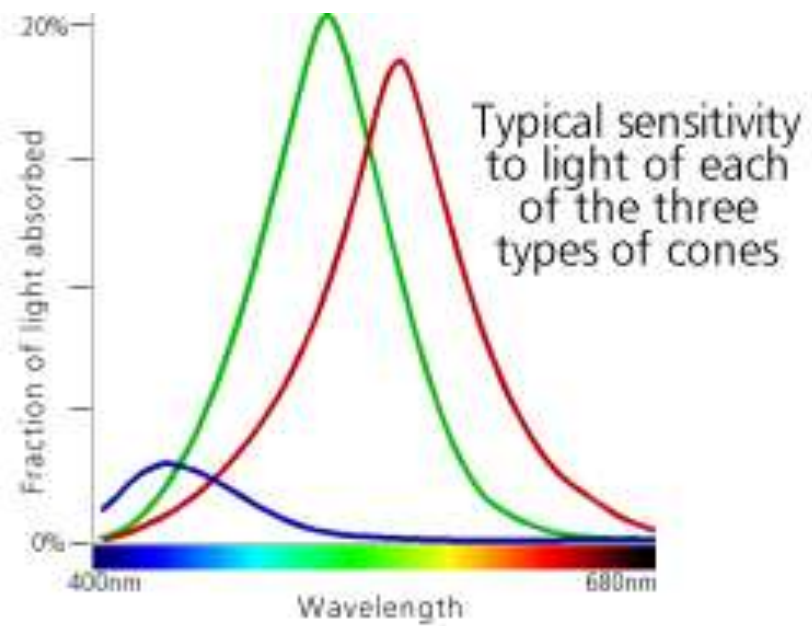
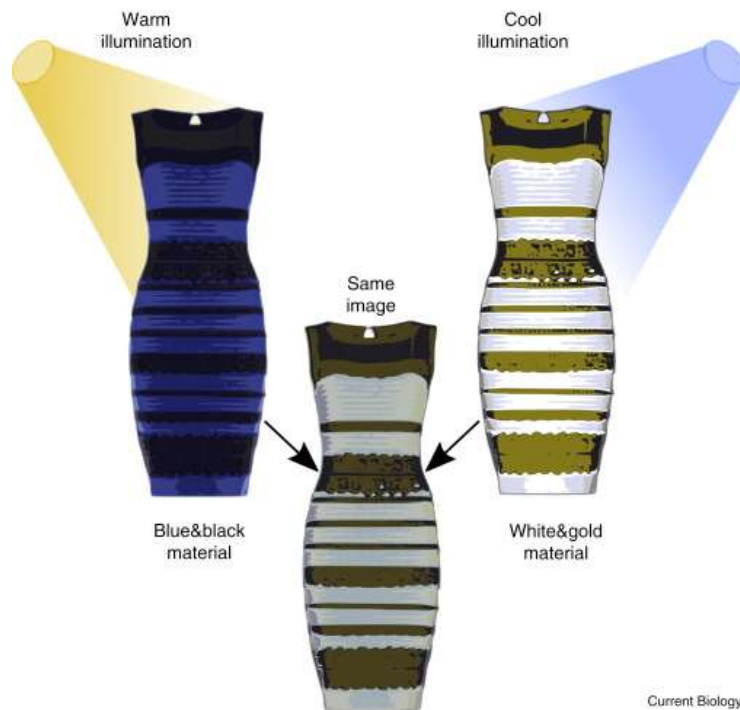


用分光镜 (Spectroscope) 区分不同的灯泡



<http://www.chemistryland.com/CHM107Lab/Exp7/Spectroscope/Spectroscope.html>

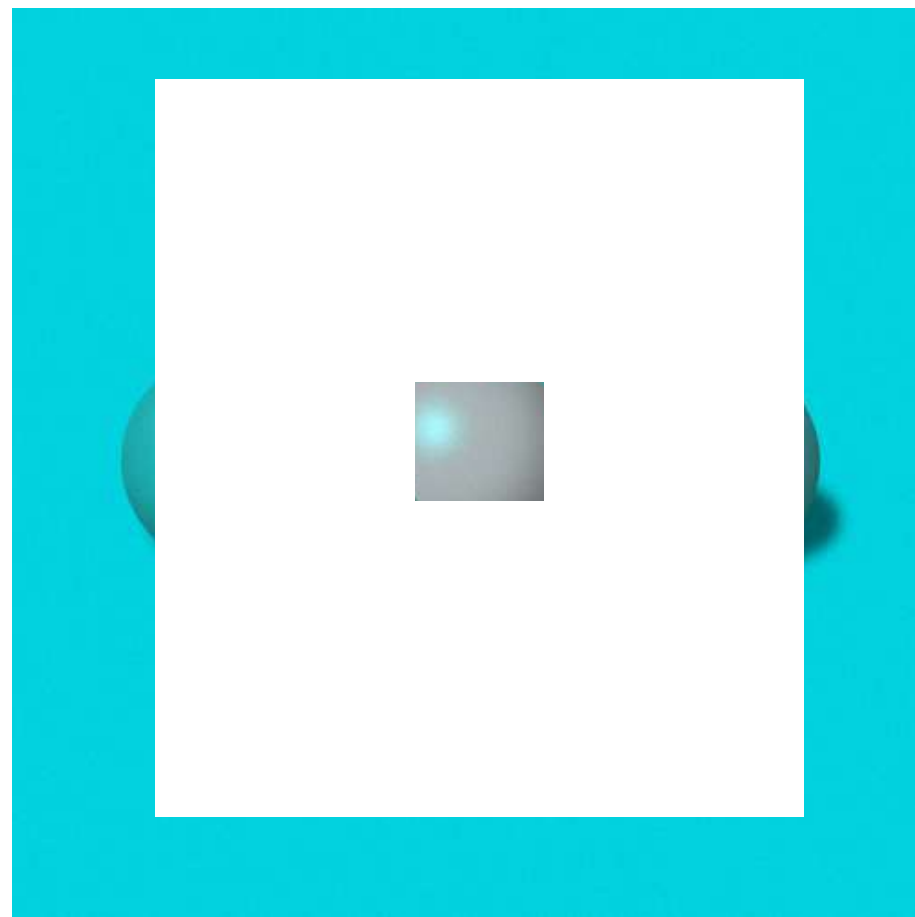
蓝黑还是白金？



中间的球是什么颜色？

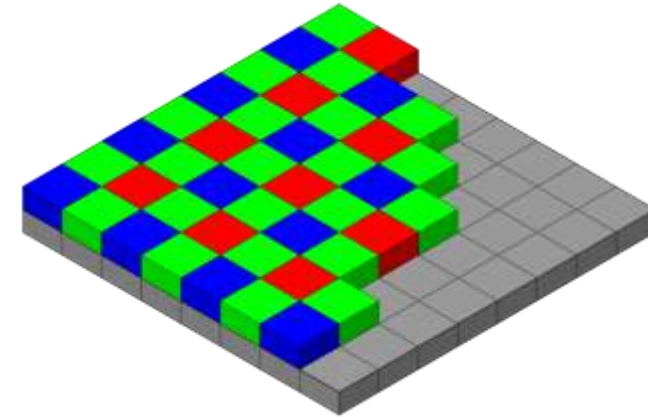


中间的球是什么颜色？

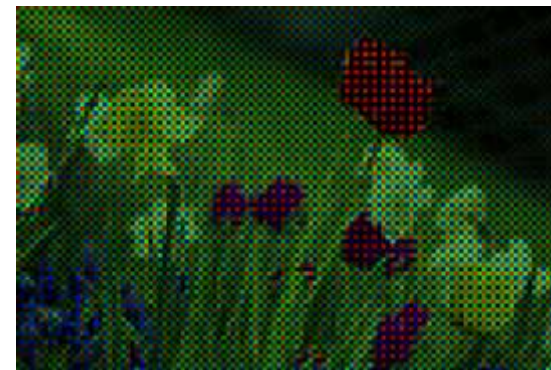


相机模型的色彩

- *Bayer filter* 贝尔滤镜分隔入射的光线，将光分为红、绿和蓝三个通道。每个像素只能透过其中一种颜色的滤镜，因此，传感器上的每个像素只能感知一种颜色的信息。
- “*demosaicked*” color image: 估计每个像素的完整彩色信息，以重建彩色图像



Bayer filter pattern in front of sensor



What the camera sees
("raw" image)



Demosaicked image

早期的图像上色

- 在发明彩色胶卷以前，Sergey Prokudin-Gorsky 根据不同光照下的黑白图像曝光来计算彩色图像。

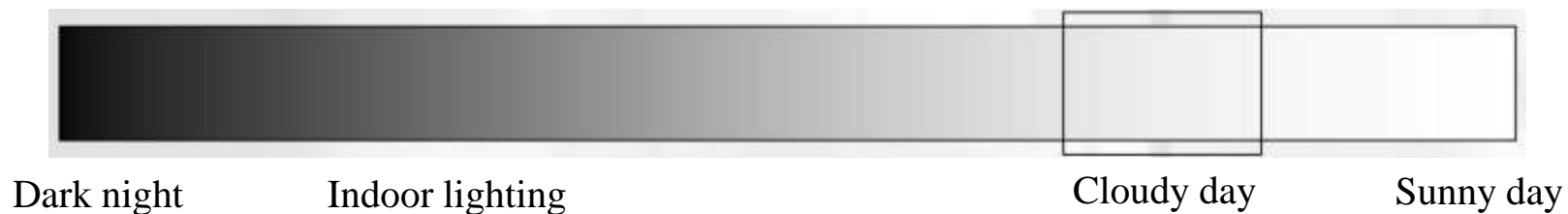


Blue, Green, Red
exposures



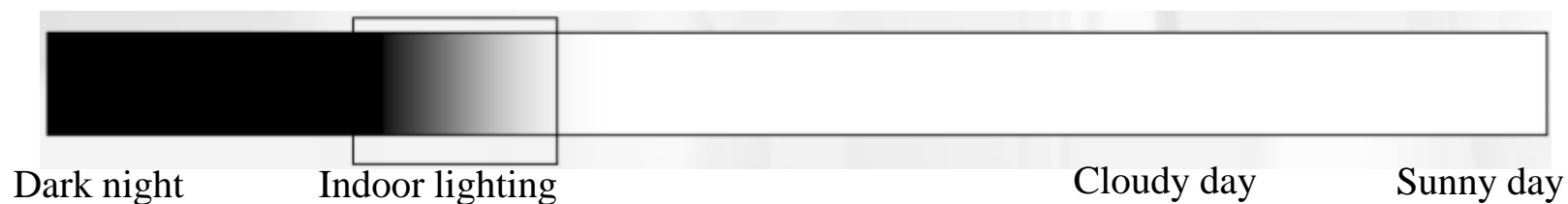
Combined color image (1911)

Recall: 动态视觉系统“尺度”

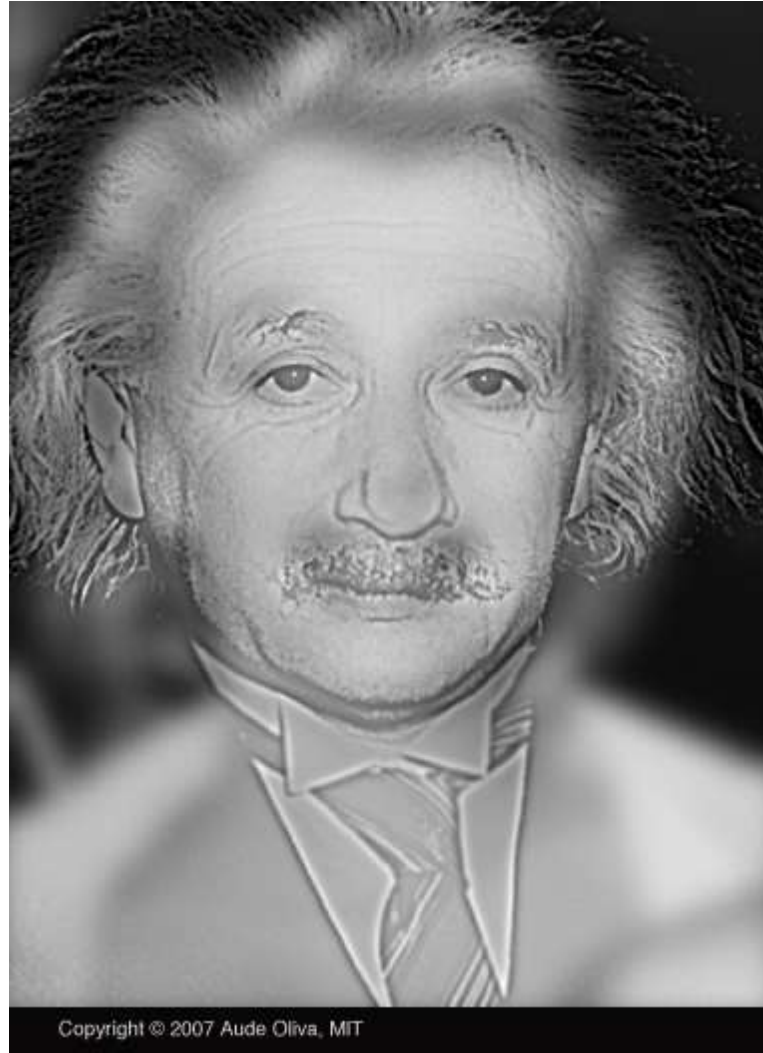


如果我们一直对整个范围都非常敏感，那么我们将无法在场景中区分亮度级别。

视觉系统通过限制其响应的“动态范围”以匹配当前的整体或“环境”光水平来解决这个问题。



另一个错觉：他是谁？

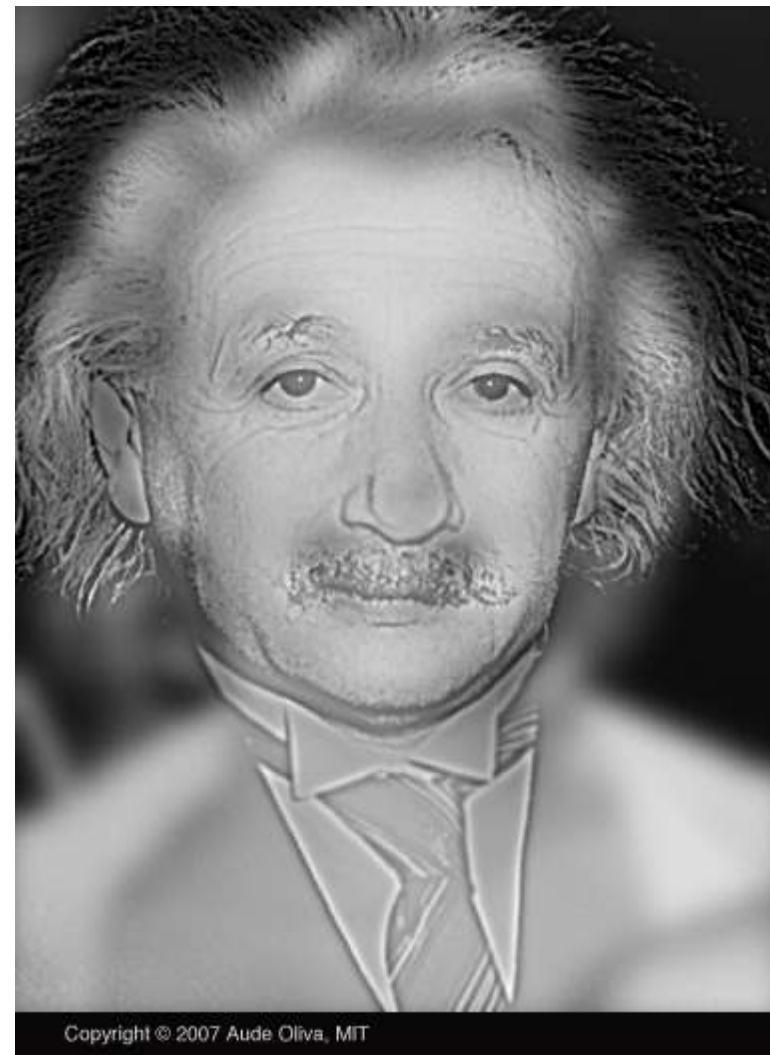
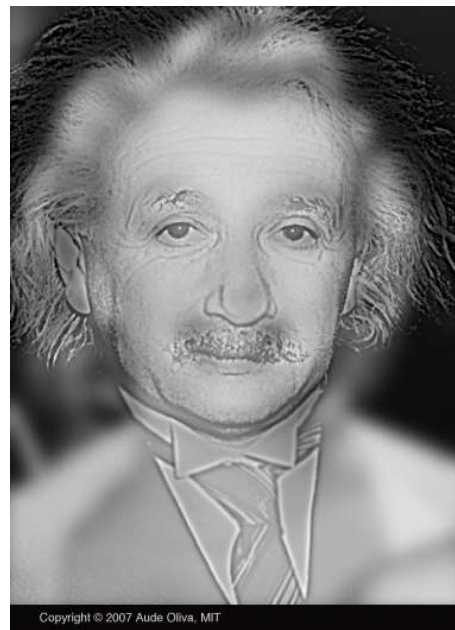


他又是谁？



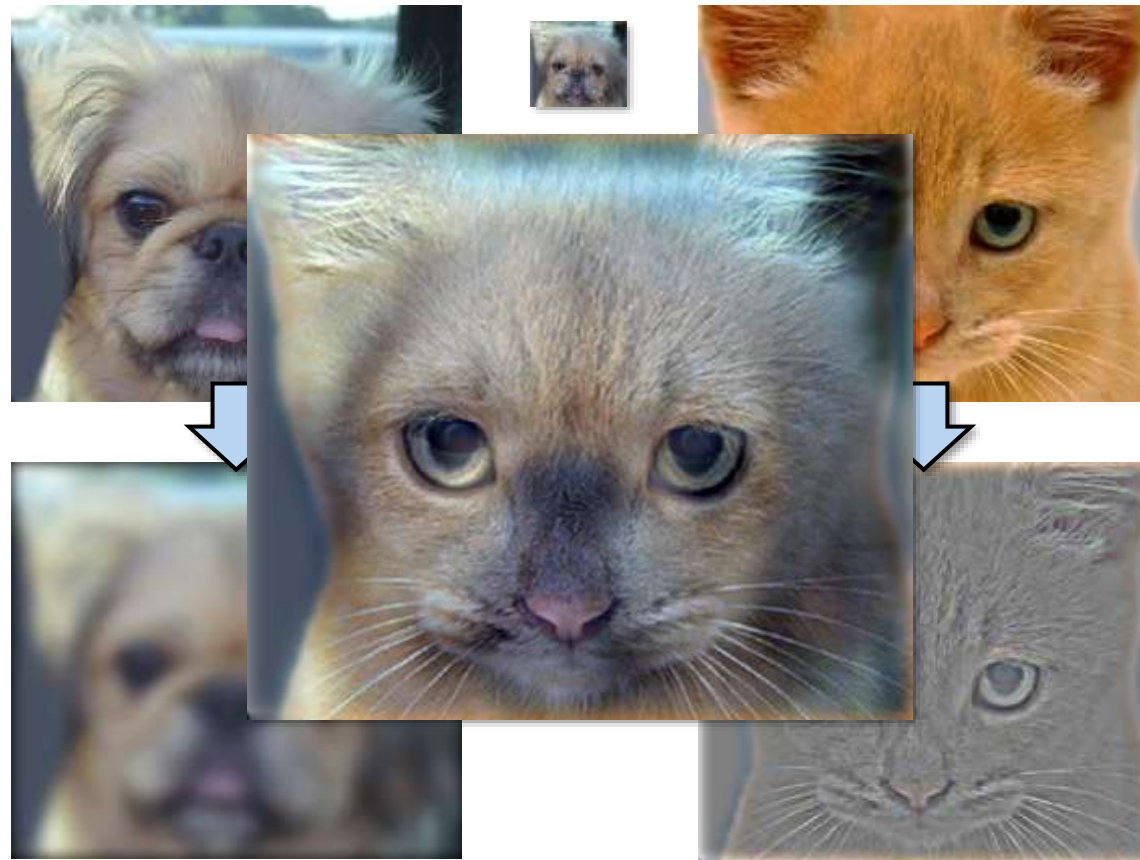
其实是同一张图！

他是谁？



Homework1

形成原理



在不同尺度下
有不同的关注点

模糊

边缘

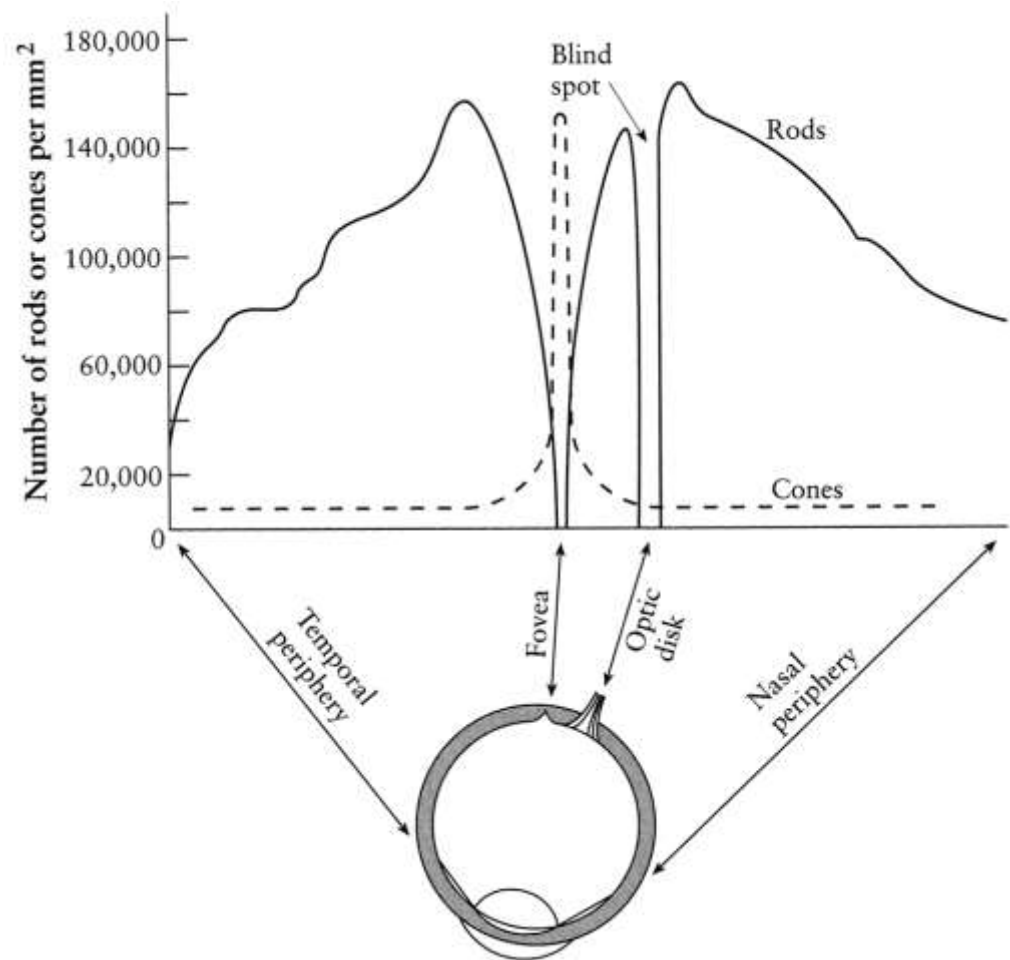
Homework 1

- 自选**2**组图做前几页slides中提到的合成图测试
- 熟悉Python基础图像处理操作、感受多尺度视觉感知的变化
- **Deadline: 2026/03/29**

回顾——光照感知

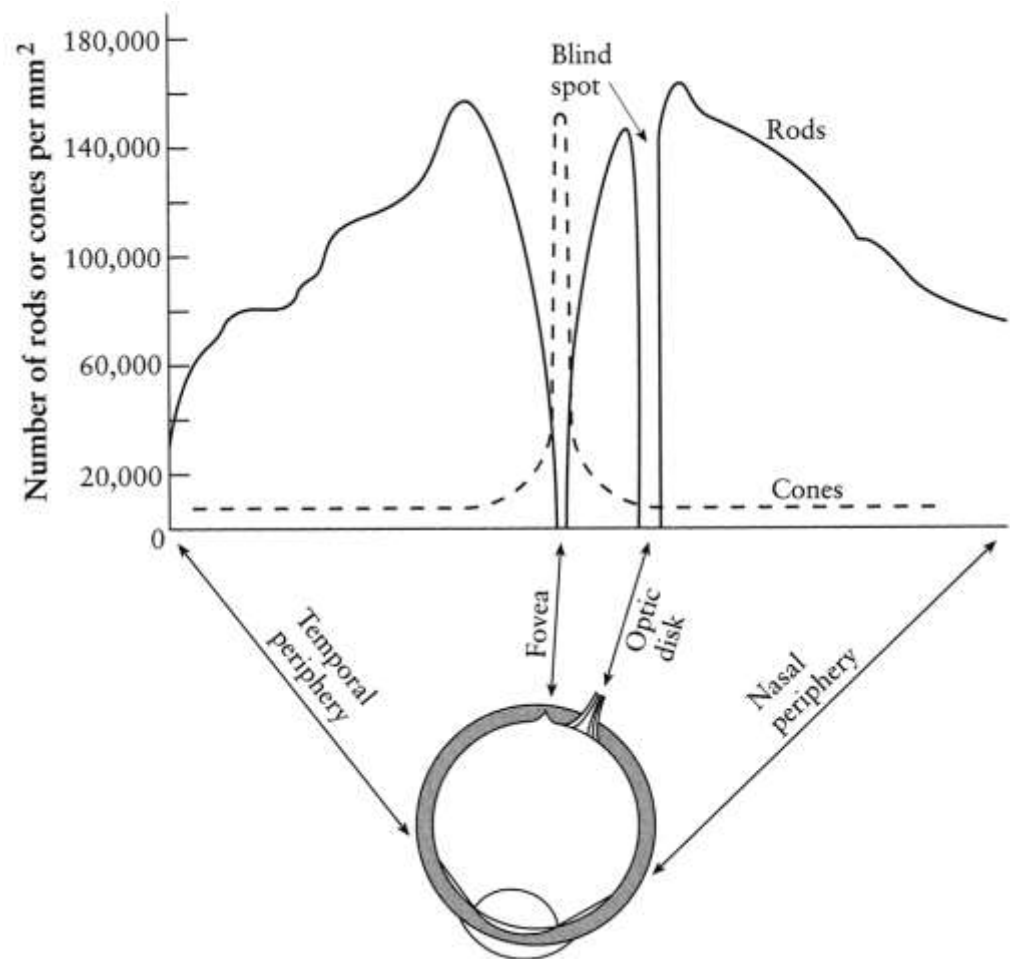
视杆与视锥

- 视杆与视锥是不均匀地分布在视网膜上的
 - 视杆对光的强度更敏感，视锥对颜色更敏感。明亮光照下，视锥细胞主导，而在暗光下，视杆细胞起主导作用
 - **Fovea 中间凹**: 视网膜上的一个小区域(1 or 2°)，包含大量视锥细胞（**没有视杆**），特别适用于高分辨率的颜色感知。它在颜色识别和高分辨率视觉中发挥关键作用。
 - 不在中央凹区域的部分，对于感知低光强度或运动检测非常有用，但视力锐度较低。



盲点

- 神经盘是视网膜上的一个区域，位于视网膜背部，是视觉信息传输的起点。它是视神经纤维束聚集的地方，这些纤维将视觉信号传送到大脑的视觉中枢。
- 视神经盘上没有感光细胞，因此在视觉场景中，与视神经盘对应的区域被视为视野中的盲点。这是我们视觉系统中的一个盲区，无法感知光或图像。



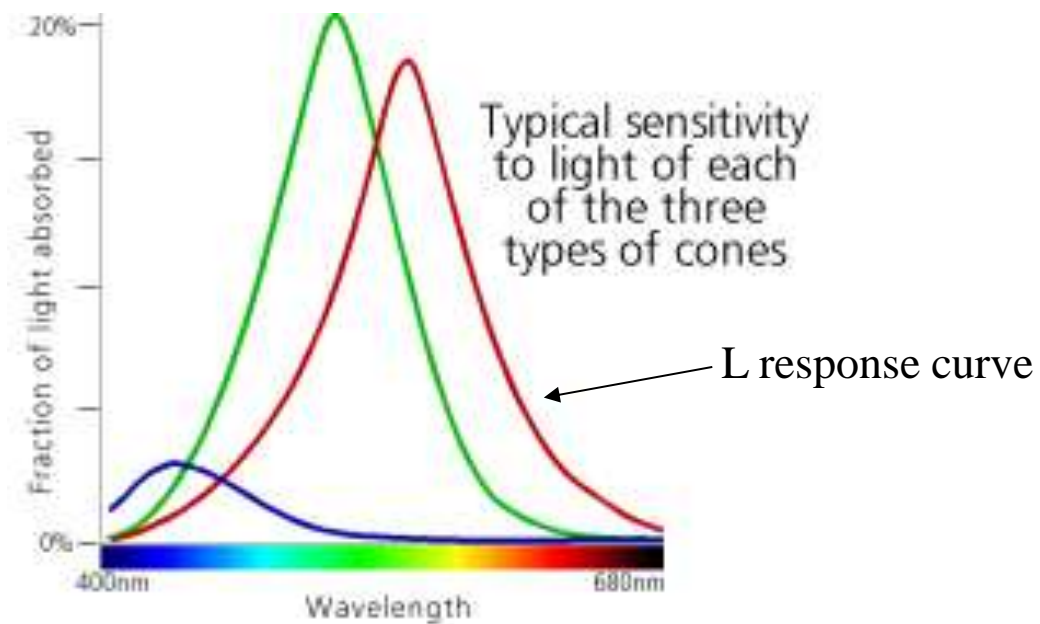
Recall: 消失的圆点



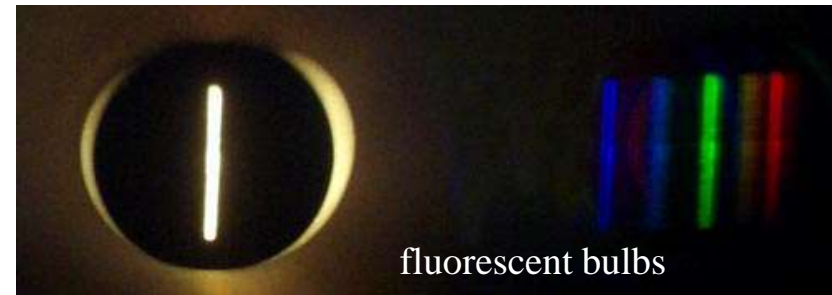
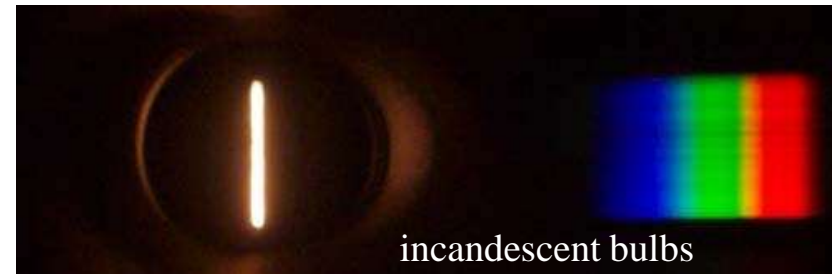
With left eye shut, look at the cross on the left. At the right distance, the circle on the right should disappear (Glassner, 1.8).

颜色感知

- 三种色锥感受颜色
 - 每种色锥对光谱中的区域更敏感
 - 这些区域相交
 - Short (S) 对应 蓝色
 - Medium (M) 对应 绿色
 - Long (L) 对应 红色
 - 总体而言，我们对红色和绿色更敏感
 - 但每个人的敏感度不同，也会随年龄变化
 - 色盲的原因是至少一种色锥出现了退化

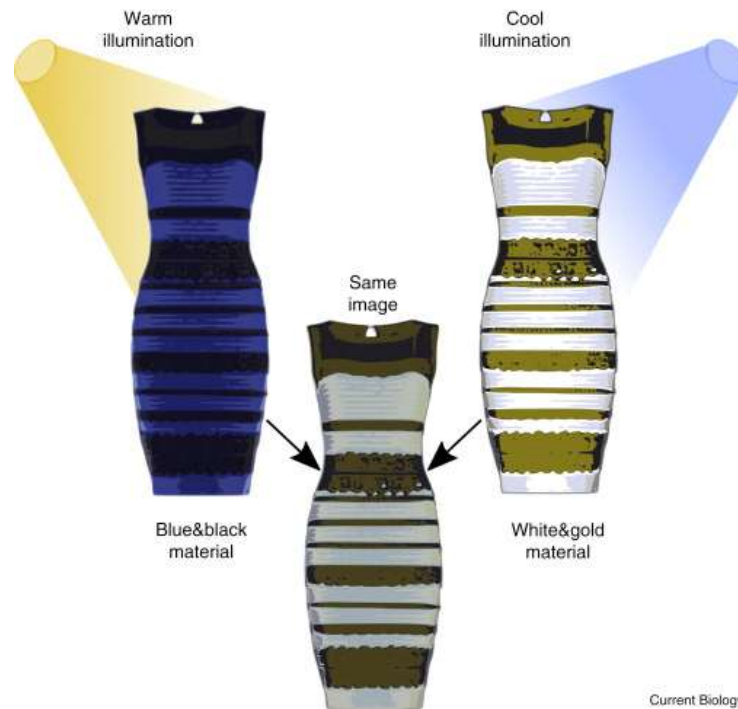


用分光镜 (Spectroscope) 区分不同的灯泡

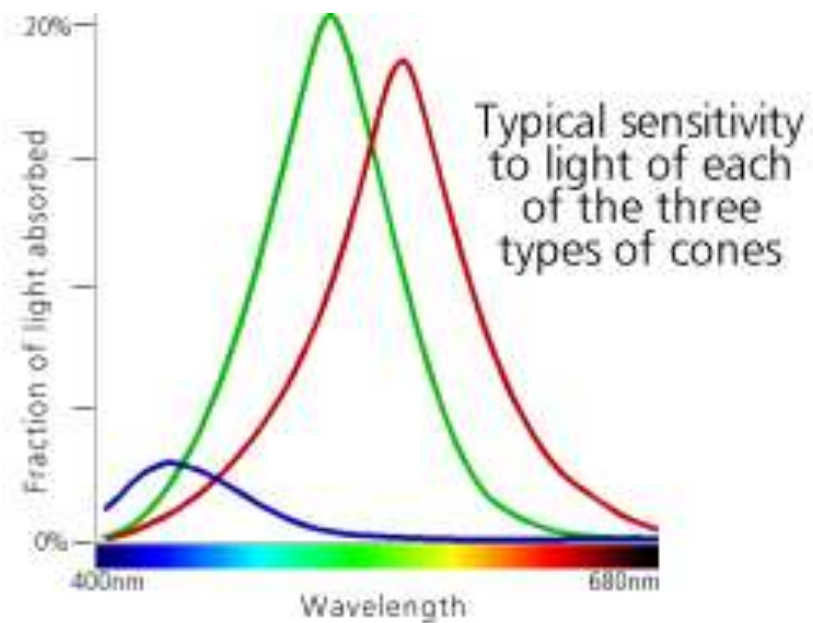


<http://www.chemistryland.com/CHM107Lab/Exp7/Spectroscope/Spectroscope.html>

蓝黑还是白金？



Current Biology



“感知”

计算机系统对数字图像和视频进行分析和理解的能力，以获得有关物体、场景、特征和动作的信息。感知的目标是模拟和模仿人类视觉系统，使计算机能够理解和解释图像中的内容。

“感知”

特征信息提取、对象识别、场景理解、运动分析、三维深度感知、
上下文感知

“感知”

从 几何模型：基于几何原理和相对位置关系的模型
到 统计模型：基于数据分布和统计推断的模型

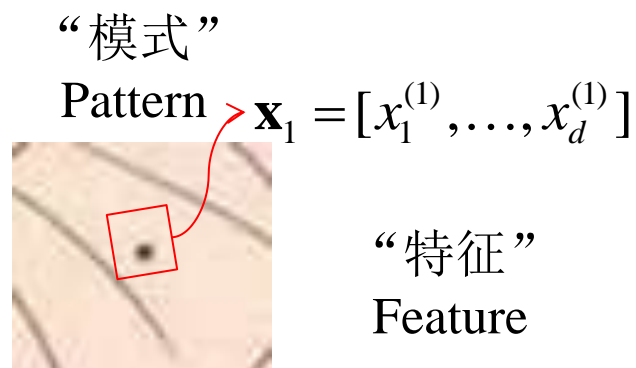
计算机如何感知？

- 让我们以人为例...
 - “音容笑貌”
 - “你掌心的痣，我总记得在哪里”



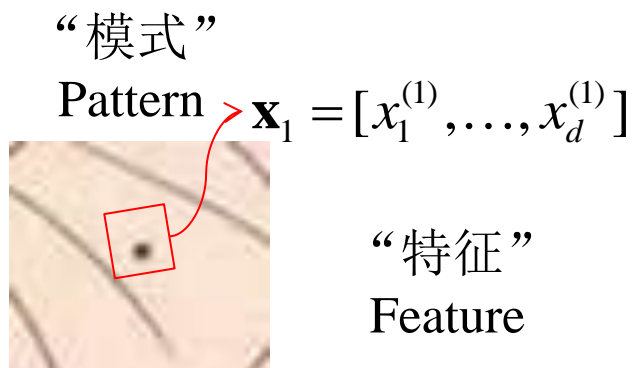
计算机如何感知？

- 让我们以人为例...
 - “音容笑貌”
 - “你掌心的痣，我总记得在哪里”



计算机如何感知？

- 让我们以人为例...
 - “音容笑貌”
 - “你掌心的痣，我总记得在哪里”



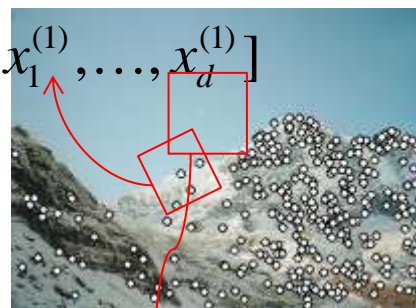
计算机的视觉认知：基于特征匹配的认识

1) 检测 Detection: 找到图中的关键点



2) 解释 Description: 在关键点周围提取特征

$$\mathbf{x}_1 = [x_1^{(1)}, \dots, x_d^{(1)}]$$



$$\mathbf{x}_2 = [x_1^{(2)}, \dots, x_d^{(2)}]$$

3) 匹配 Matching: 根据两个视角下的特征进行匹配



挑战是什么？

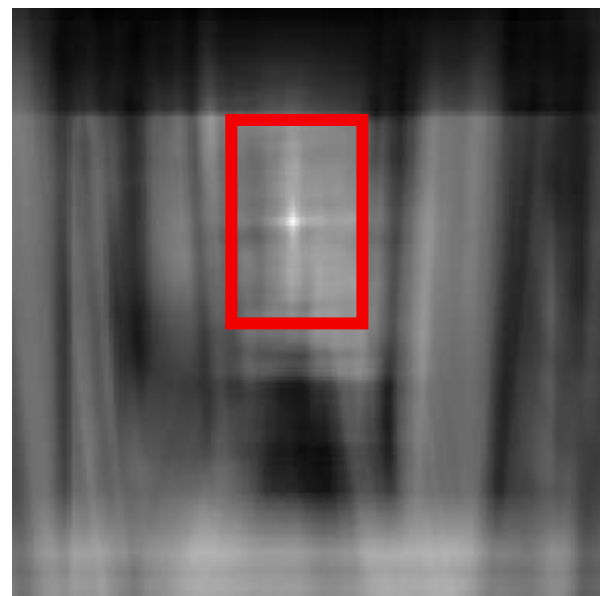
一把椅子



找出图像中的椅子



特征匹配度热图



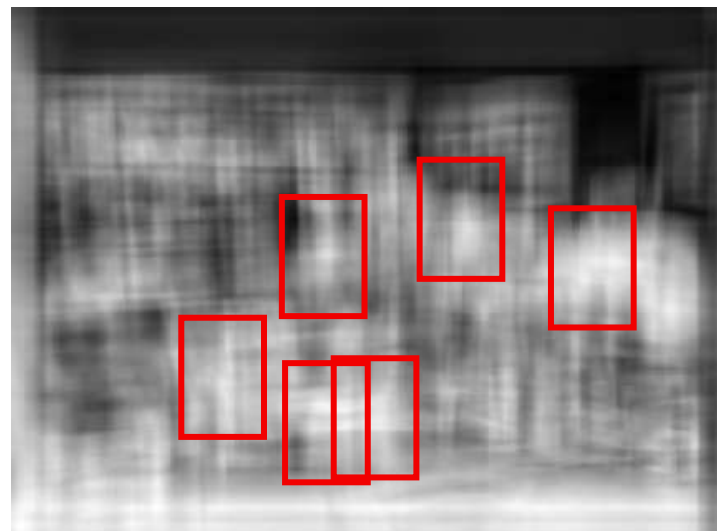
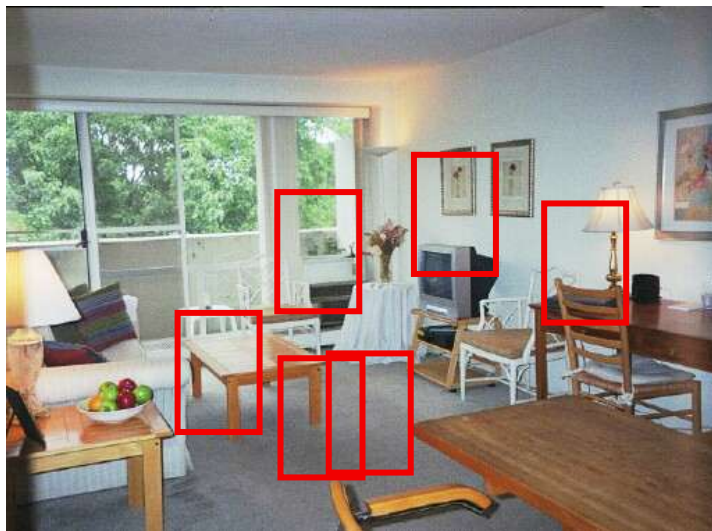
这样做有什么问题？

- 视觉信息质量可能较低
 - 模糊、遮挡、噪音

- 实时性、通用性挑战
 - 外观多变、视觉错觉

挑战是什么？

找到图中的椅子



绝大多数目标框都没有目标物体
目标物体在新场景下不会一成不变

通用性挑战：不同视角



Michelangelo 1475-1564

通用性挑战：不同光照条件

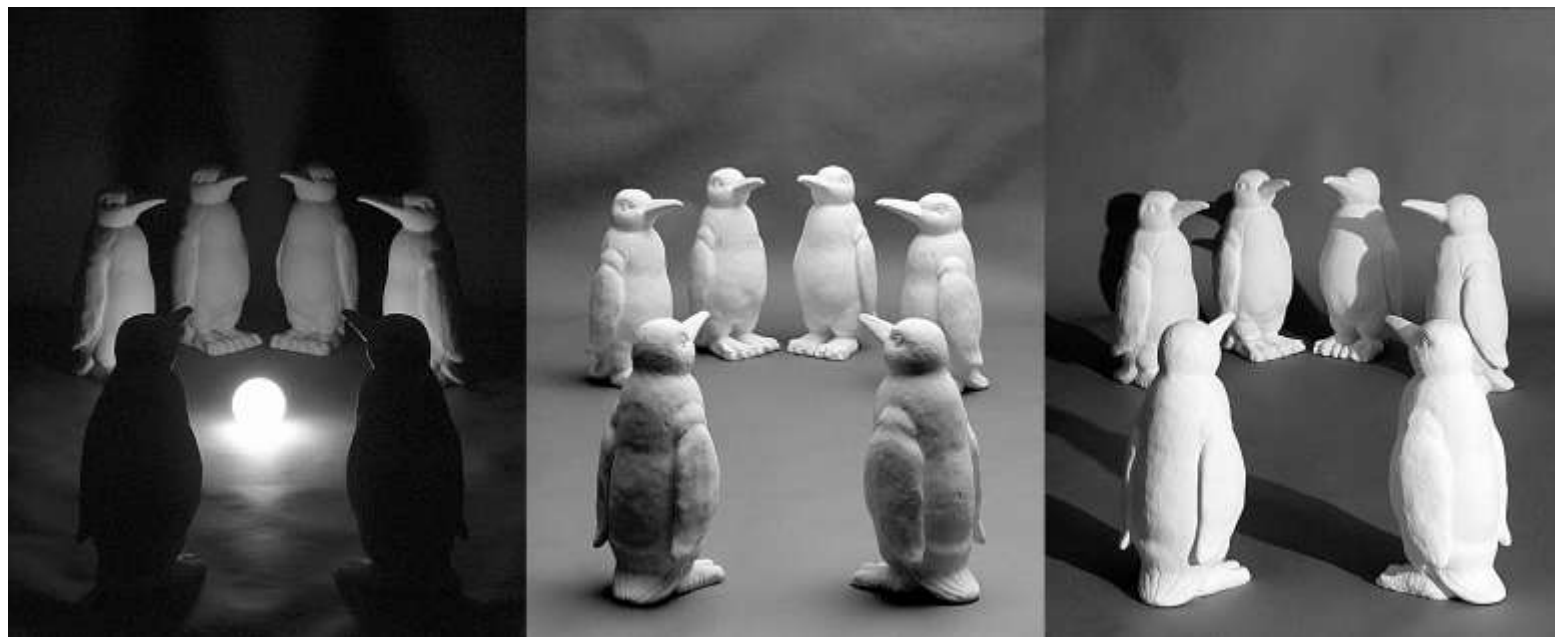


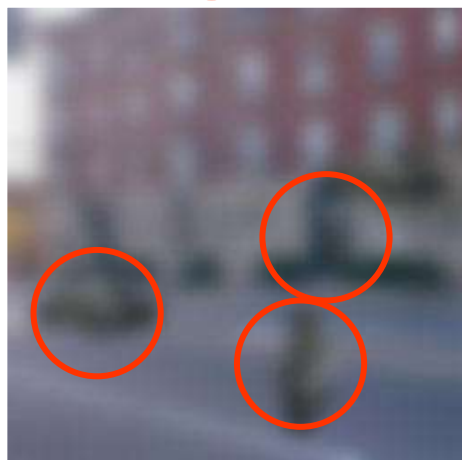
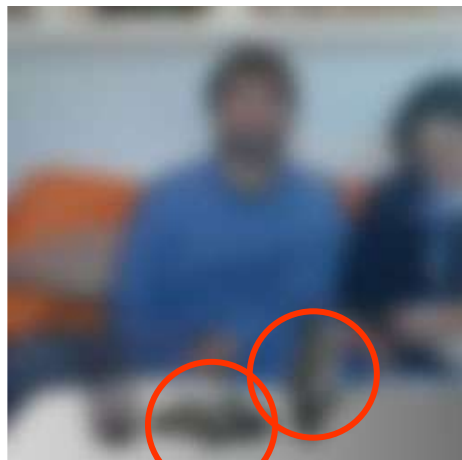
image credit: J. Koenderink

通用性难点：类内差异



同类目标外观差异很大

通用性挑战：模糊与局部歧义



单独局部无法推断内容

slide credit: Fei-Fei, Fergus & Torralba

通用性难点：极端任务



Brady, M. J., & Kersten, D. (2003). Bootstrapped learning of novel objects. *J Vis*, 3(6), 413-422



<https://www.dcard.tw/f/funny/p/233833012>

视觉错觉：感知对于人们来说同样有挑战

- 感知本身就是一个定义模糊的问题
 - 三维世界与二维世界的感知区别
 - 感知需要大量知识储备



视觉错觉下的可乐瓶
Artist Julian Beaver

Image source: F. Durand

视觉错觉

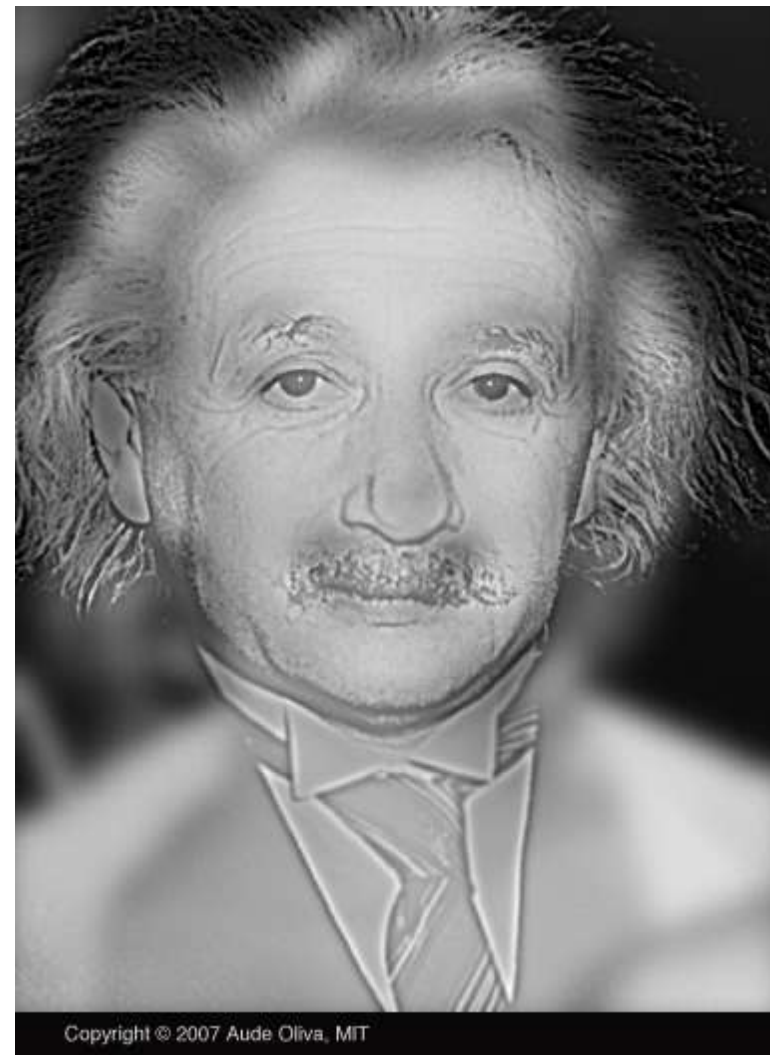
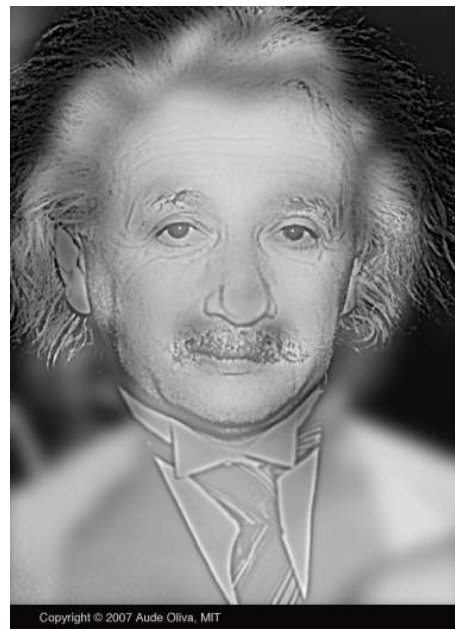


视觉错觉



With left eye shut, look at the cross on the left. At the right distance, the circle on the right should disappear (Glassner, 1.8).

尺度问题



Copyright © 2007 Aude Oliva, MIT

尺度问题: Wagon-wheel 效应



https://en.wikipedia.org/wiki/Wagon-wheel_effect

人类如何解决这些困难？

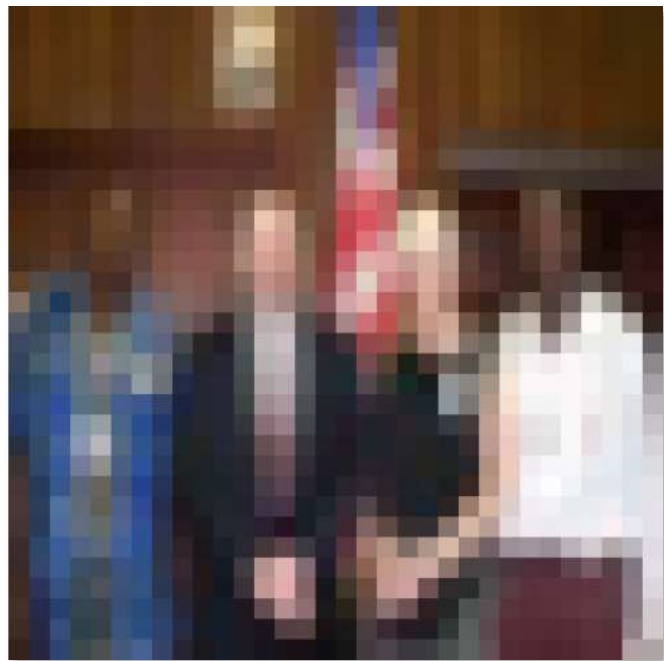
- 多次认证
 - 灵活调整自身状态
 - 总有一个最好的尺度、没有错觉的视角

- 记忆、推理
 - 经验、智能

人类的优势：调整视角



人类的优势：经验推理



Source: "80 million tiny images" by Torralba, et al.



计算机的优势：图像处理

- 增强图像的 (“Computational Photography”)



超分辨率 **放大，放大，再放大！**
(source: 2d3)



亮度调整 **照亮你的美！**
(credit: [Hasinoff et al., SIGGRAPH ASIA 2016](#))



背景虚化 **柔光双摄！** (source: [Google Research Blog](#))



图像补全 **一键修图！**
(image credit: Hays and Efros)



计算机的优势：感知可以利用大量的冗余

一个房子也许认不出来...



一百个房子认出来了！

Source: S. Lazebnik

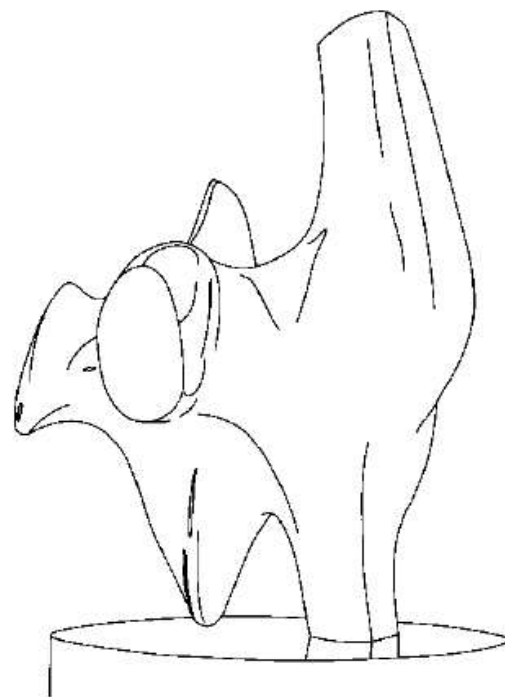
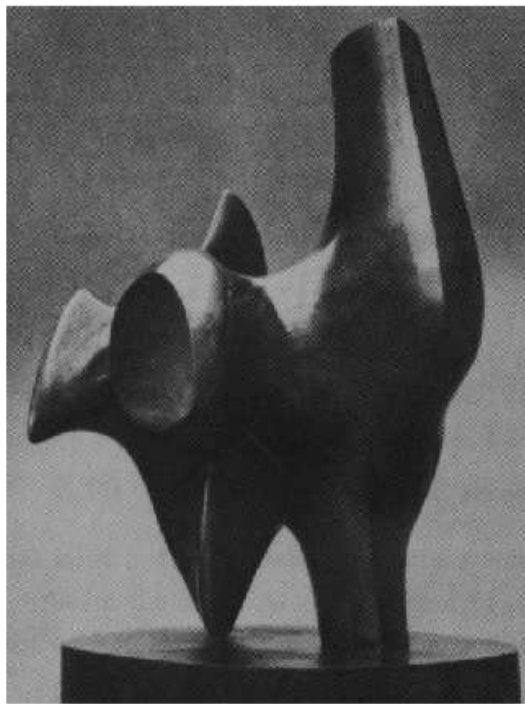
“感知”

从 几何模型：基于几何原理和相对位置关系的模型
到 统计模型：基于数据分布和统计推断的模型

我们需要统计模型快速合理地进行感知推断

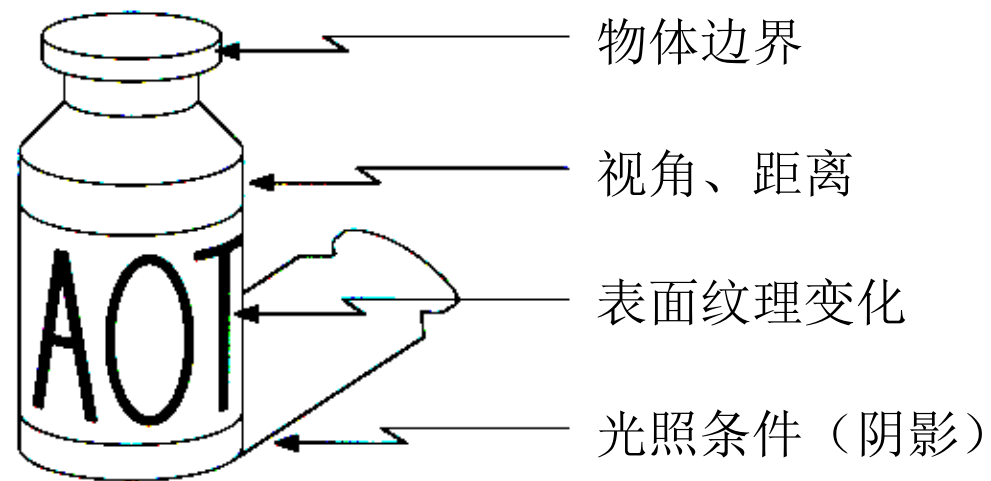
边缘检测

Edge detection 边缘检测



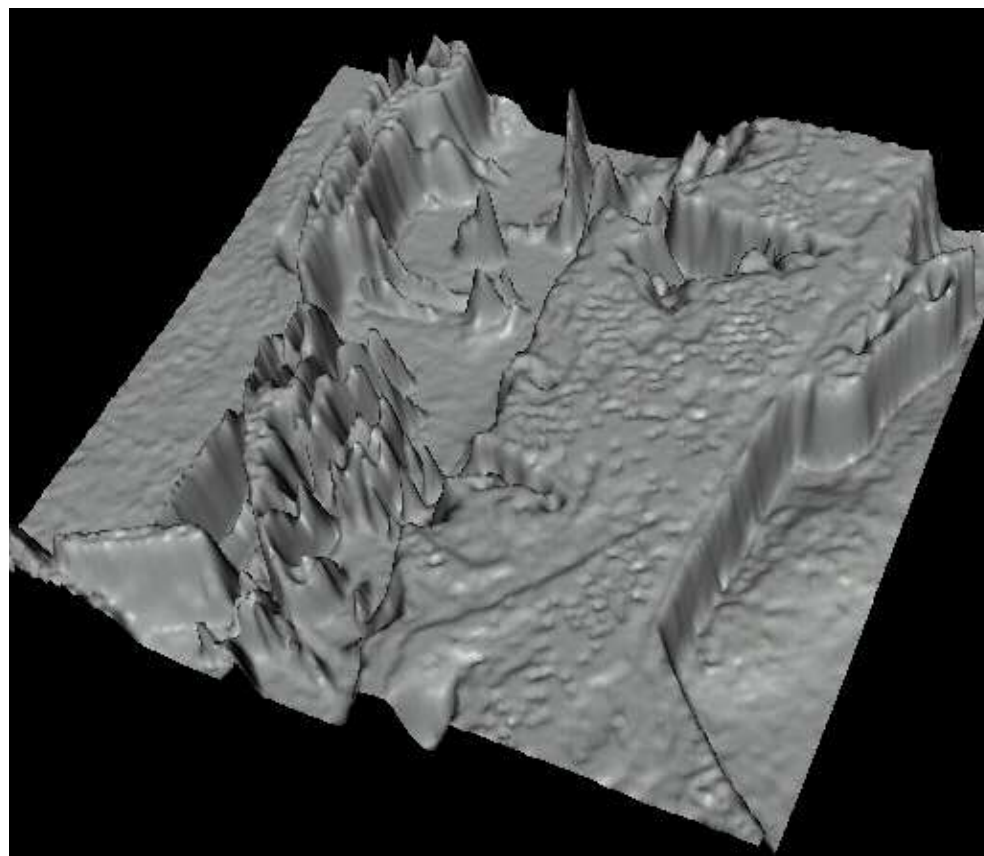
- 把2维图像转成曲线的集合
 - 可以提取场景中的显著特征
 - 比像素的信息量更大（简图）

什么是边缘？



- 很多因素可以导致边缘的出现

回顾：图像可以认为是一个像素位置到像素值的映射函数



- 边缘可以认为是那些“悬崖”

定义边缘

- 边缘时图像强度发生突变的区域

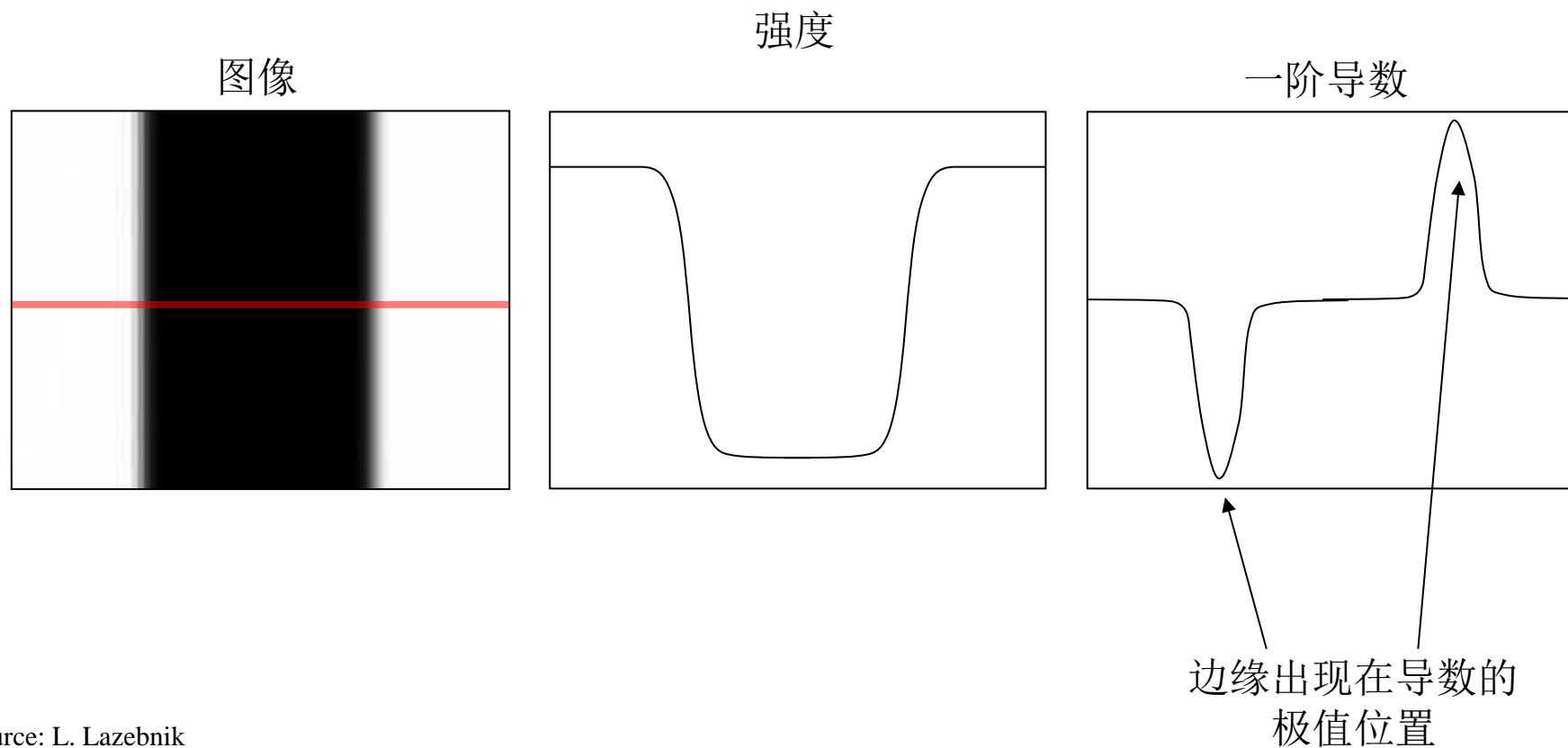


Image derivatives: 图像斜率

- 如何计算图像 $F[x,y]$ 的“斜率”？
 - Option 1: 把 F 转换为连续函数 f
 - Option 2: 做离散梯度discrete derivative (finite difference)

$$\frac{\partial f}{\partial x}[x, y] \approx F[x + 1, y] - F[x, y]$$

可以用卷积实现

$$\frac{\partial f}{\partial x} \cdot \begin{array}{|c|c|c|} \hline & & \\ \hline & & \\ \hline & & \\ \hline \end{array}$$

H_x

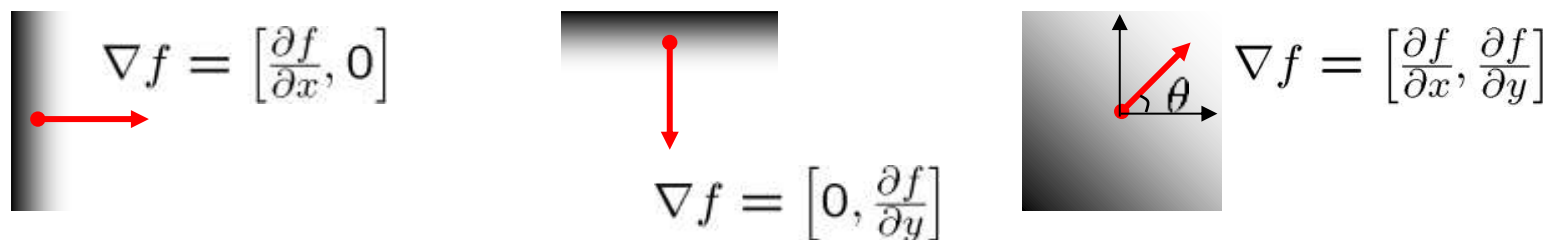
$$\frac{\partial f}{\partial y} \cdot \begin{array}{|c|c|c|} \hline & & \\ \hline & & \\ \hline & & \\ \hline \end{array}$$

H_y

Image gradient: 图像梯度

- 图像梯度定义为: $\nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right]$

这些梯度会指出哪些方向图像会发生剧烈的变化



梯度的强度定义了边缘的强度:

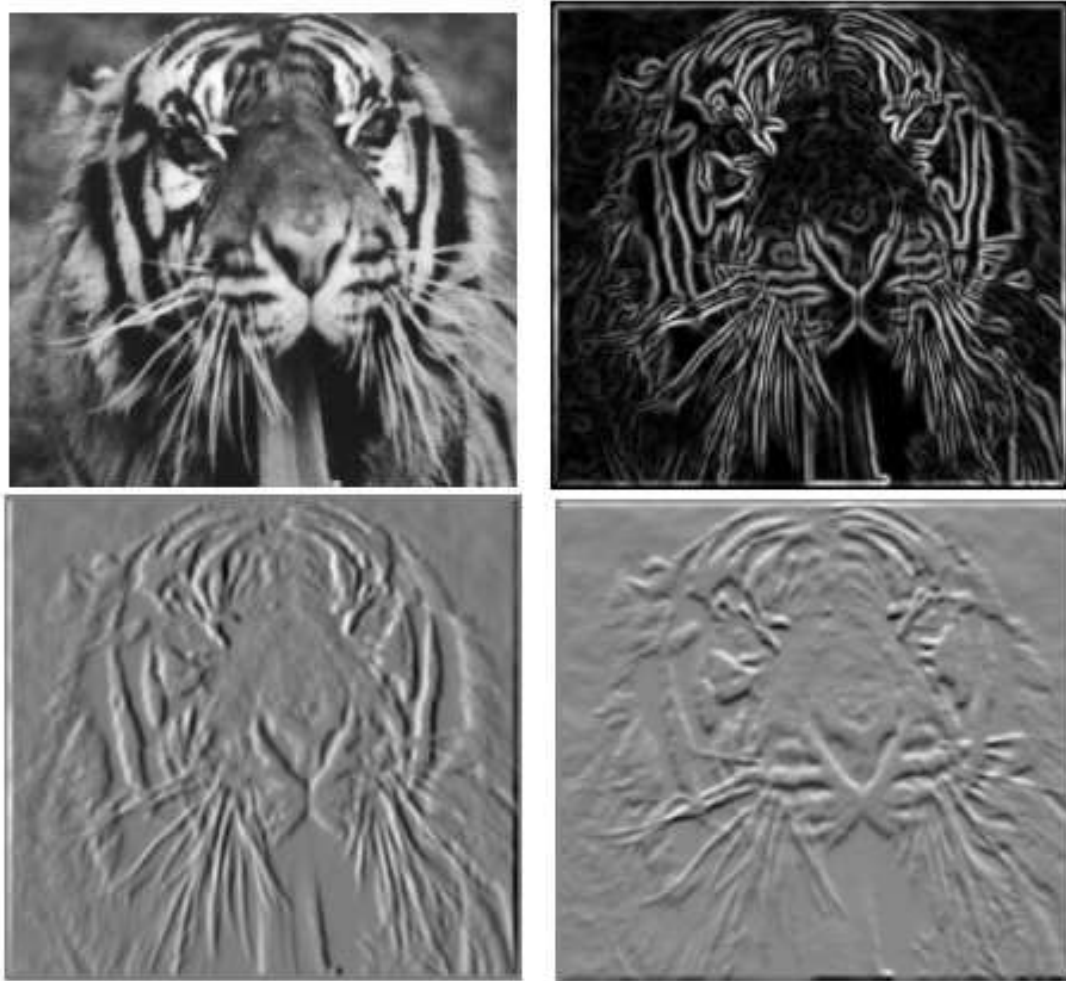
$$\|\nabla f\| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$$

我们可以计算梯度的方向:

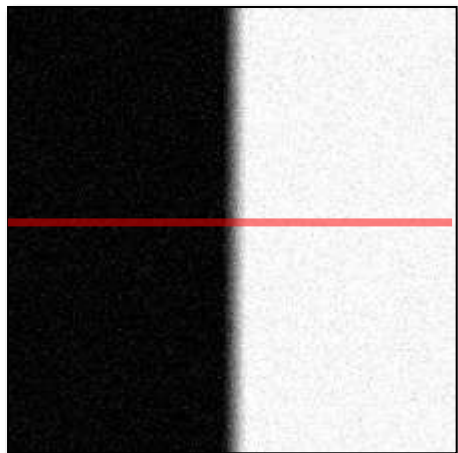
$$\theta = \tan^{-1} \left(\frac{\partial f / \partial y}{\partial f / \partial x} \right)$$

- 梯度的方向可以反应边缘的方向，他们是什么关系呢？

图像梯度

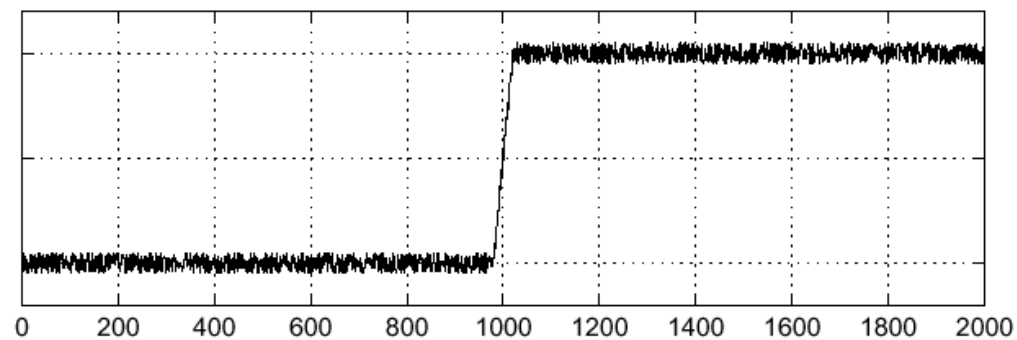


噪声的影响

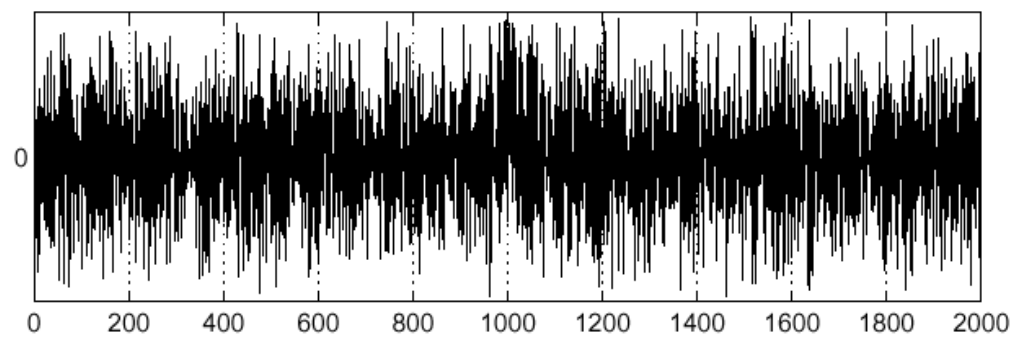


有噪声的图像

$$f(x)$$

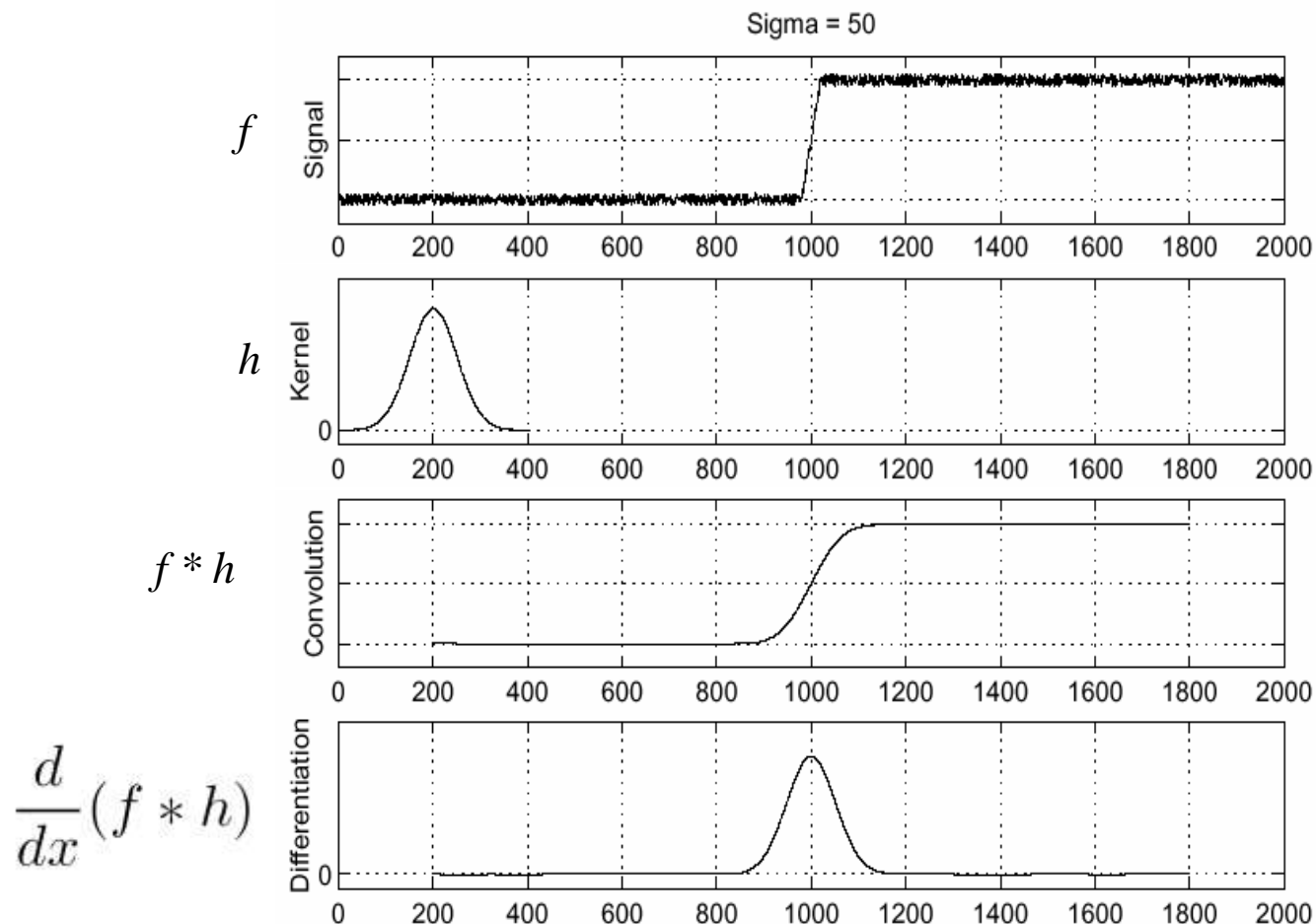


$$\frac{d}{dx}f(x)$$



哪里是边缘?

解决方案：高斯模糊

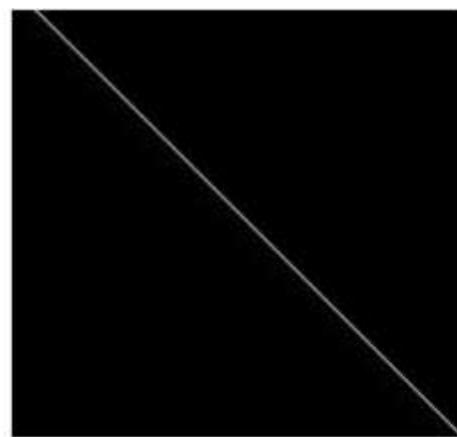


如果要找边缘，找峰值

$$\frac{d}{dx}(f * h)$$



Image with Edge



Edge Location

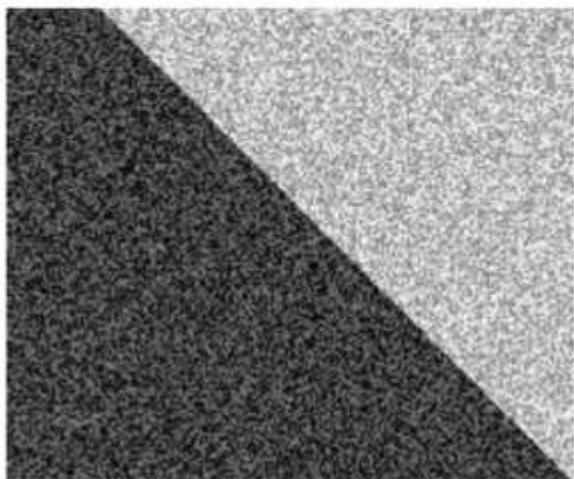
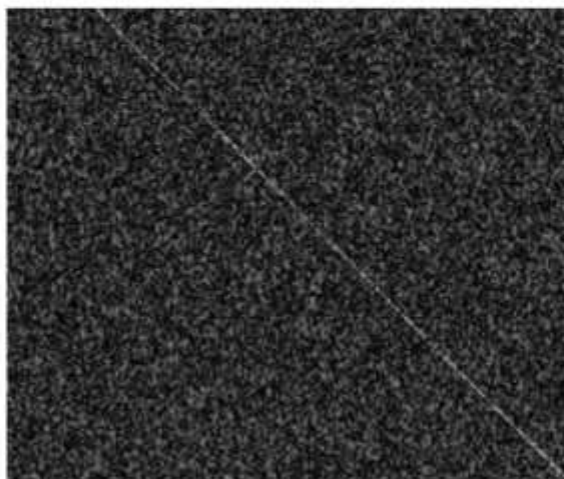
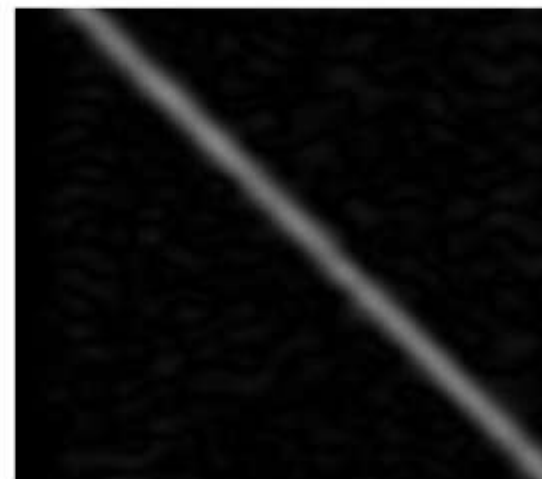


Image + Noise



Derivatives detect
edge *and* noise

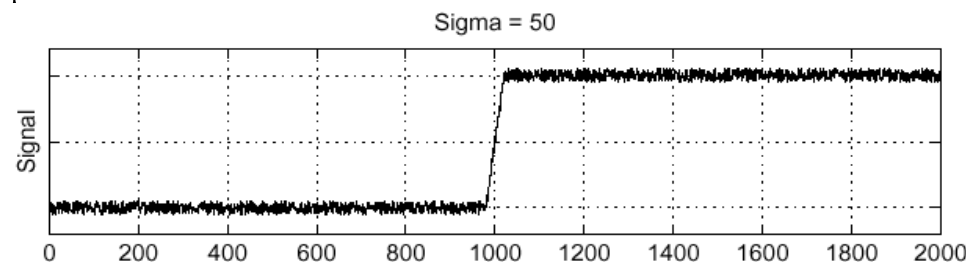


Smoothed derivative removes
noise, but blurs edge

卷积是符合交换律的

- 计算梯度可以用卷积实现，而卷积有交换律：
$$\frac{d}{dx}(f * h) = f * \frac{d}{dx}h$$
- 我们可以把计算梯度和模糊结合在一起。
 - Derivative of Gaussian (DoG)

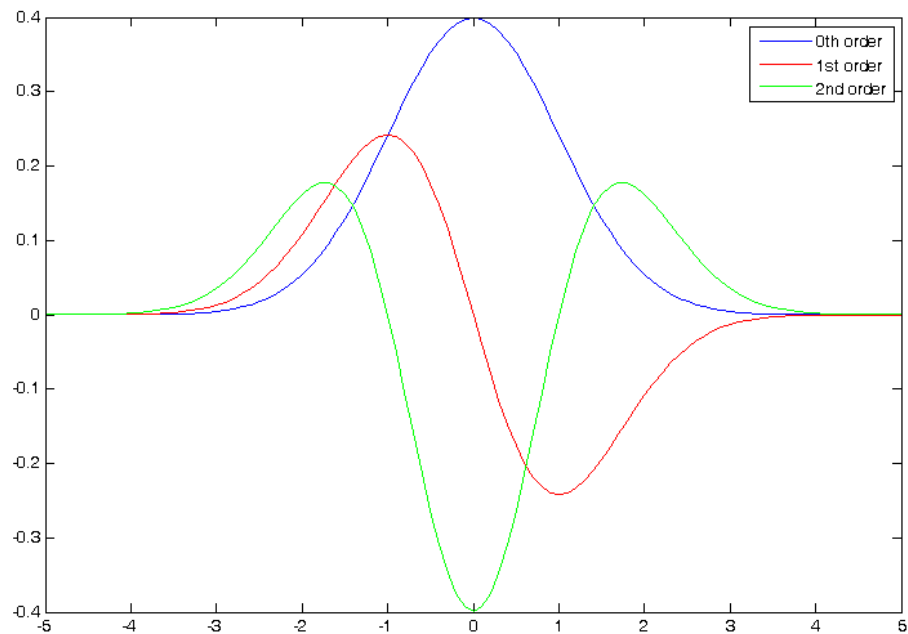
f



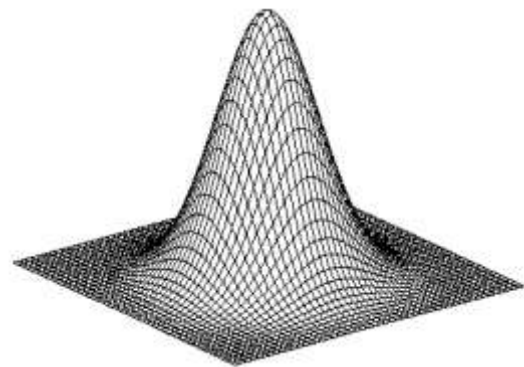
一维高斯函数和DoG

$$G_{\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

$$G'_{\sigma}(x) = \frac{d}{dx}G_{\sigma}(x) = -\frac{1}{\sigma} \left(\frac{x}{\sigma}\right) G_{\sigma}(x)$$

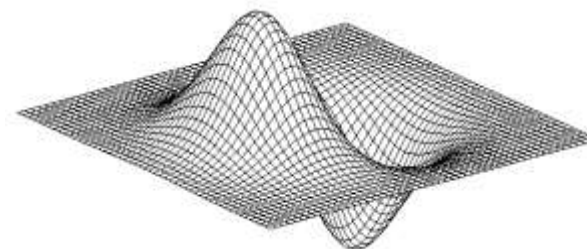


二维高斯函数和DoG



Gaussian

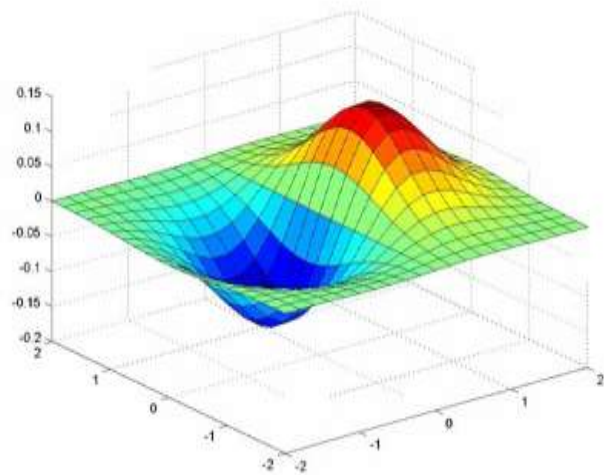
$$h_{\sigma}(u, v) = \frac{1}{2\pi\sigma^2} e^{-\frac{u^2+v^2}{2\sigma^2}}$$



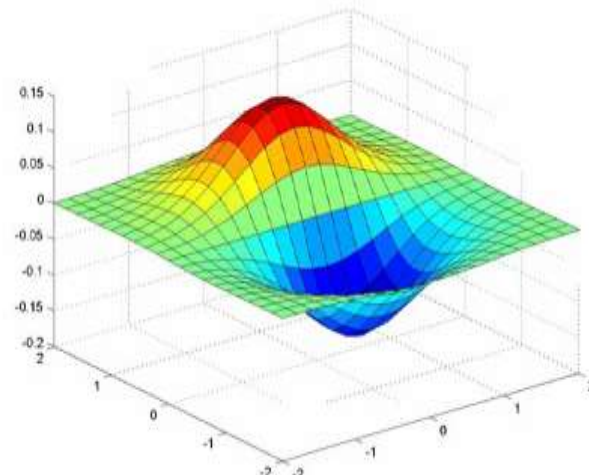
derivative of Gaussian (x)

$$\frac{\partial}{\partial x} h_{\sigma}(u, v)$$

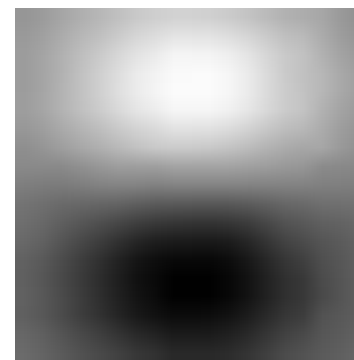
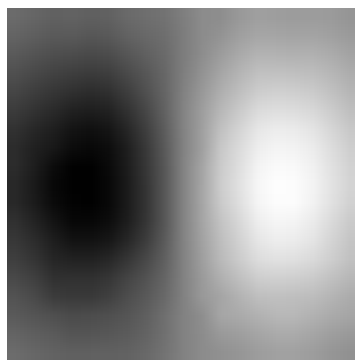
不同方向的DoG



x-direction



y-direction



Sobel 算子

- 经常用来替换DoG

$$\frac{1}{8} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

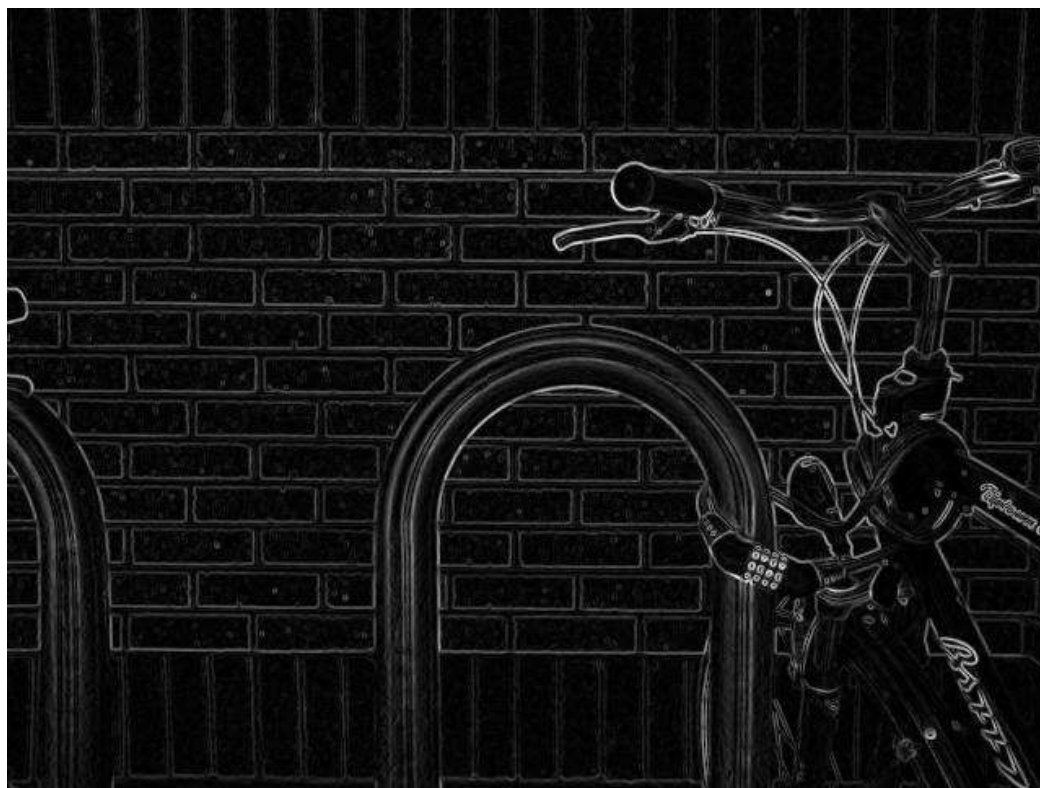
s_x

$$\frac{1}{8} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

s_y

- Sobel算子通常比DoG更加计算效率，因为它使用的是简单的3x3卷积核，而DoG通常需要更大的卷积核和更多的计算。
 - 计算简单，速度快。对于基础的边缘检测任务通常足够有效。
 - 对噪声比较敏感。可能无法检测到更复杂或微妙的边缘。

Sobel算子



Source: Wikipedia

效果



原图

Demo: <http://bigwww.epfl.ch/demo/ip/demos/edgeDetector/>

Image credit: Joseph Redmon

找出边缘



平滑梯度幅度

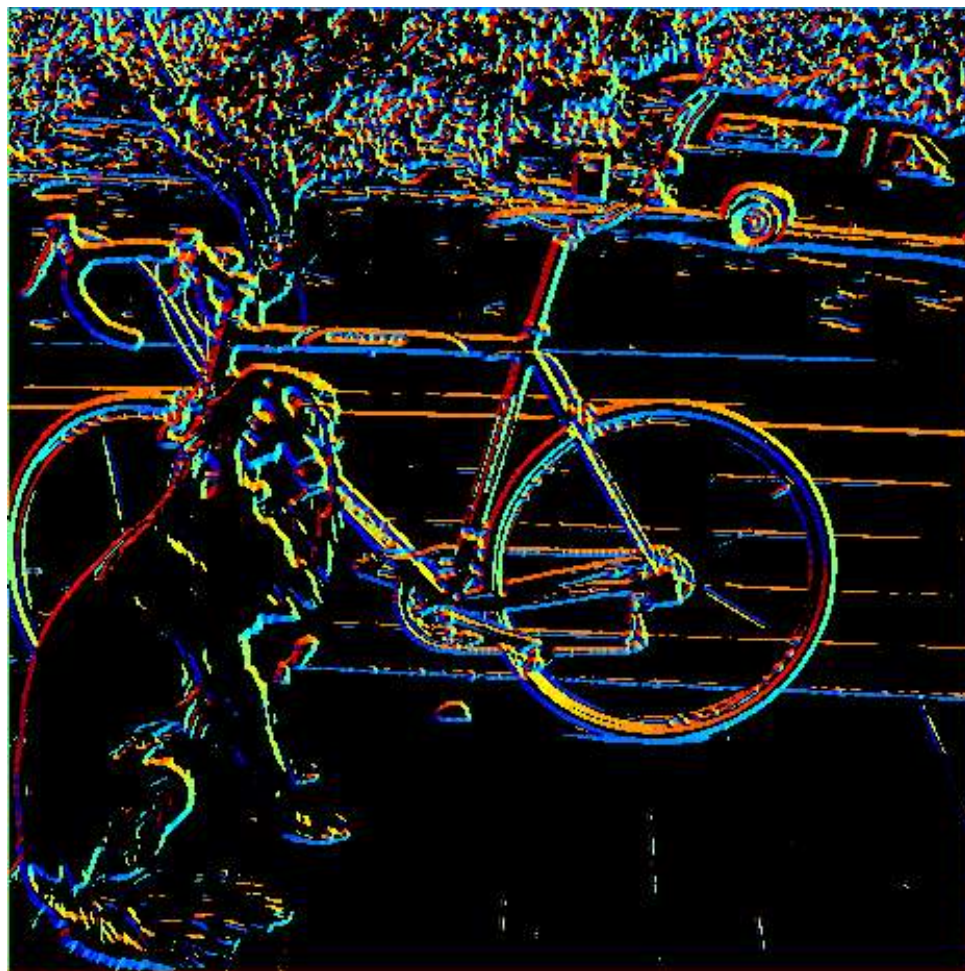
找到梯度



梯度在哪?

利用阈值

找到每个像素的方向



$$\theta = \text{atan2}(g_y, g_x)$$

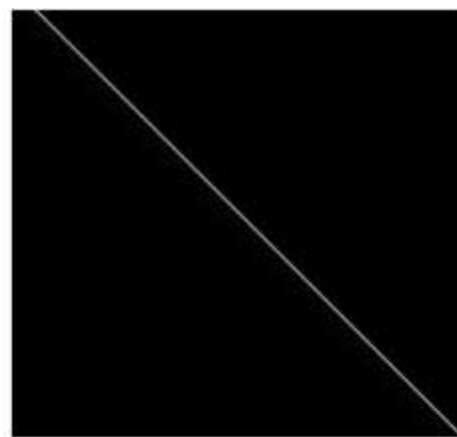
360

Gradient orientation angle

0



Image with Edge



Edge Location

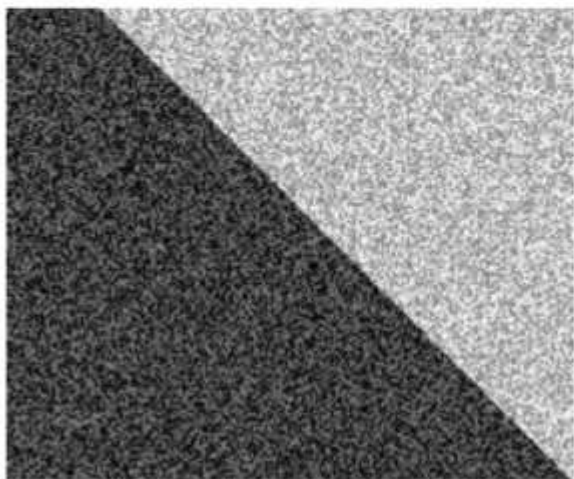
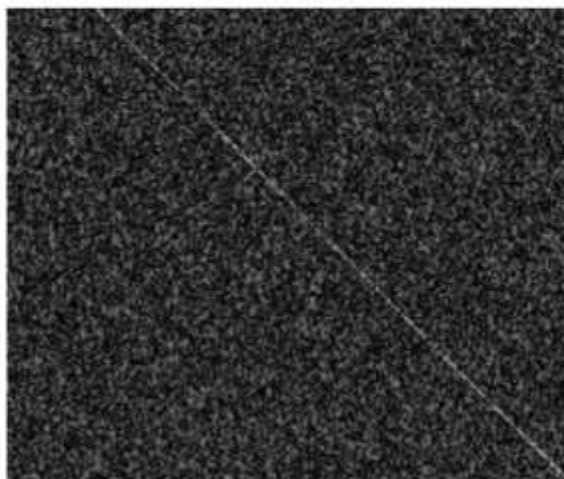
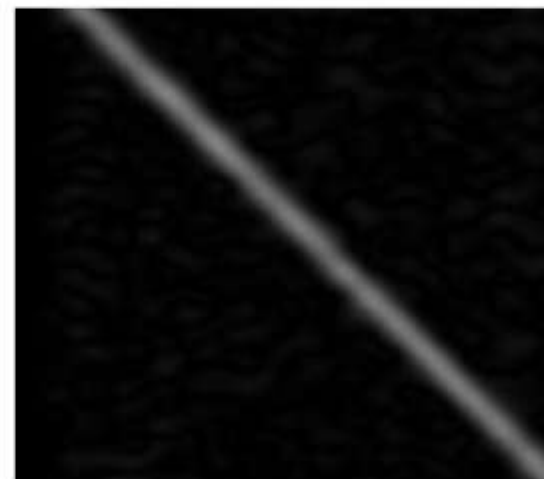


Image + Noise



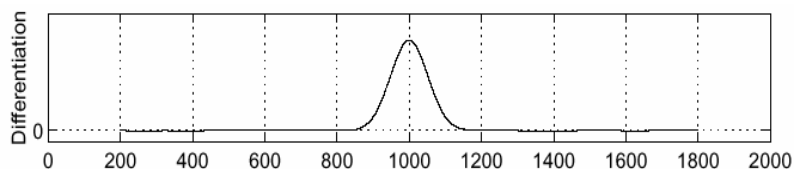
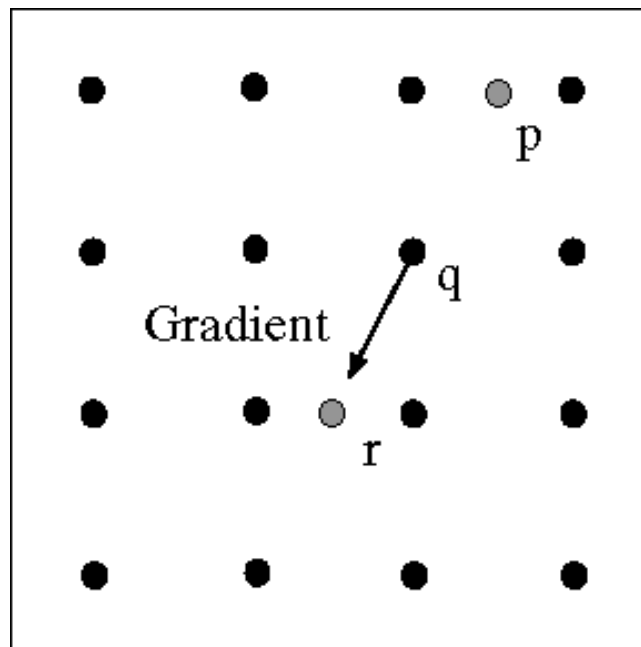
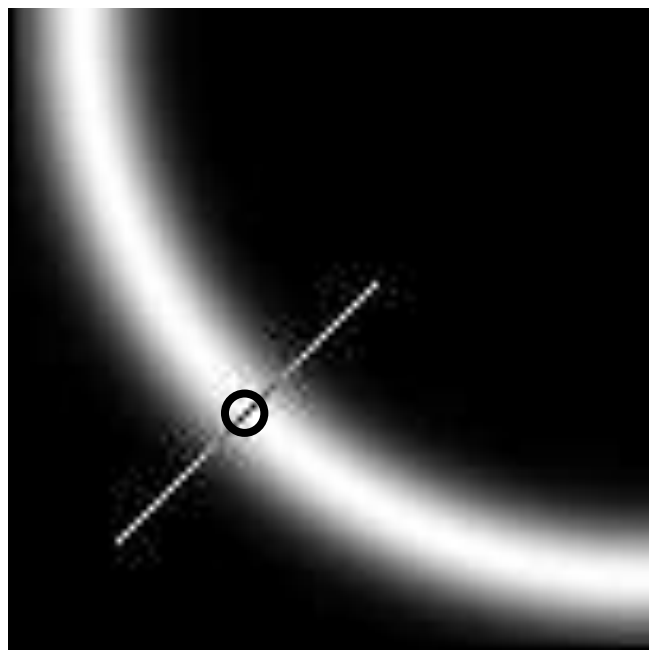
Derivatives detect
edge *and* noise



Smoothed derivative removes
noise, but blurs edge

Non-maximum suppression: 非极大抑制

将“厚边缘”细化为“细边缘”



- 对于图像中的每一个像素点，确定其梯度的方向，计算该方向子区域的梯度幅度，在该方向上，如果该像素点的梯度幅度不是局部最大值，则将其设置为0。

Before Non-max Suppression



After Non-max Suppression



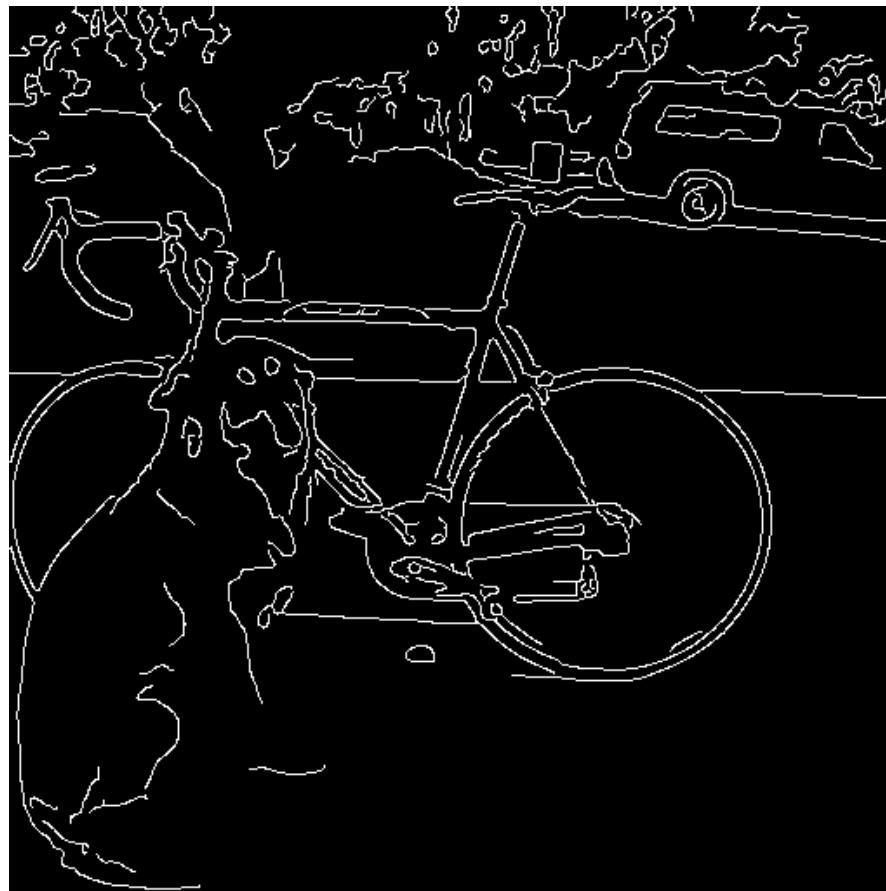
通过阈值得到边缘

- 噪声依然存在
- 我们想要一些明显边缘
- 2 个阈值, 3 种情况
 - $R > T$: 强边缘
 - $R < T$ but $R > t$: 弱边缘
 - $R < t$: 非边缘
- 为什么要两个阈值?



连接边缘

- 强边缘一定是边缘
- 如果弱边缘与强边缘联通则也是边缘
- 通过周围八个像素点进行搜索



J. Canny, *A Computational Approach To Edge Detection*, IEEE Trans. Pattern Analysis and Machine Intelligence, 8:679-714, 1986.



MATLAB: `edge(image,'canny')`

Canny边缘检测



1. 使用DoG对图像预处理

2. 找到梯度的幅度和方向



3. Non-maximum suppression

4. 连接边缘:

- 定义高低两个阈值，得到弱边缘和强边缘
- 用强边缘作为初始边缘，用弱边缘连接强边缘



Canny边缘检测器



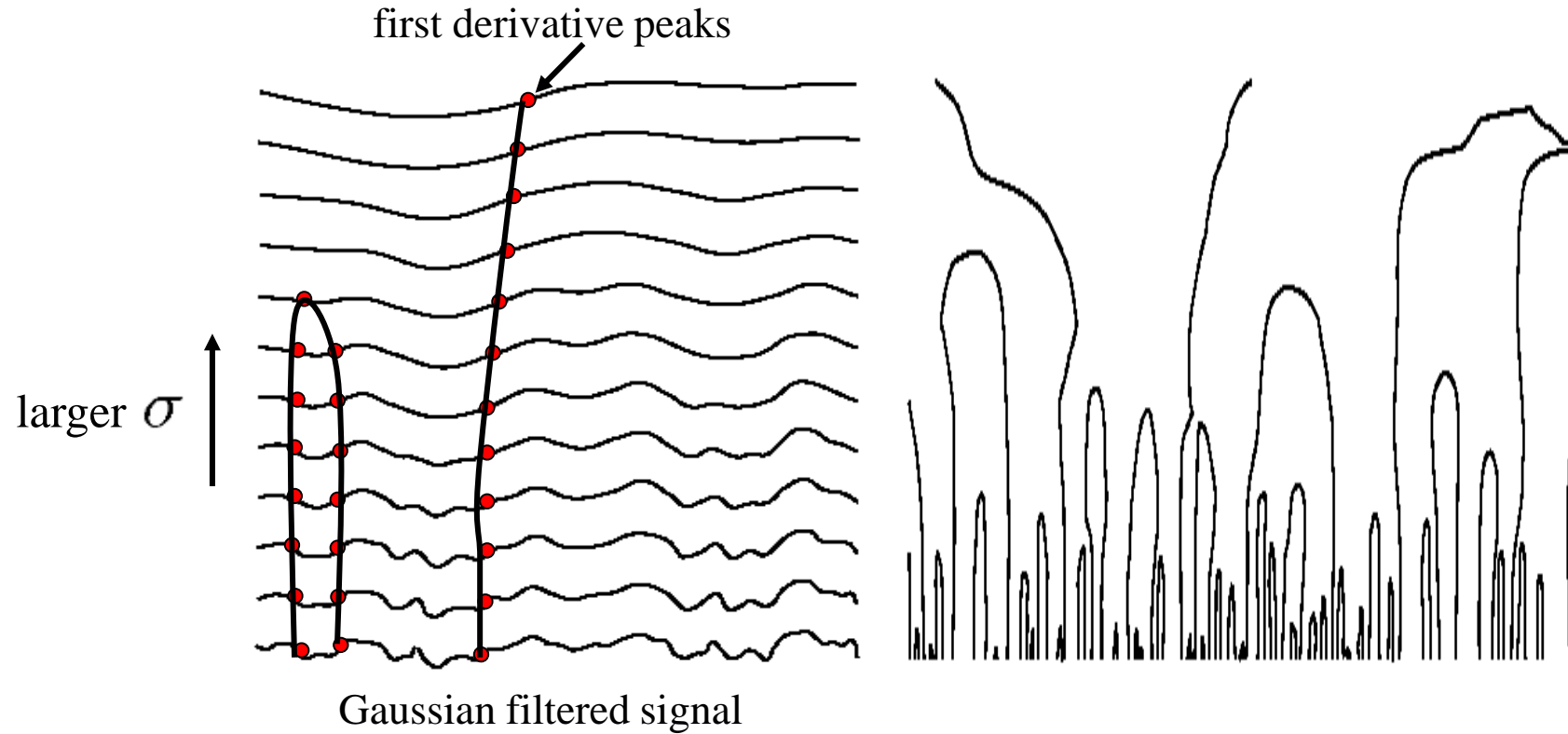
original

Canny with $\sigma = 1$

Canny with $\sigma = 2$

- 高斯滤波带宽 σ 影响最终结果
 - 高带宽边缘更“大尺度”
 - 低带宽边缘更“细粒度”

尺度空间[Witkin 83]



- 更高的尺度下（更大的带宽）：
 - 边缘的位置可能改变
 - 边缘可能合并
 - 但边缘不会再分开