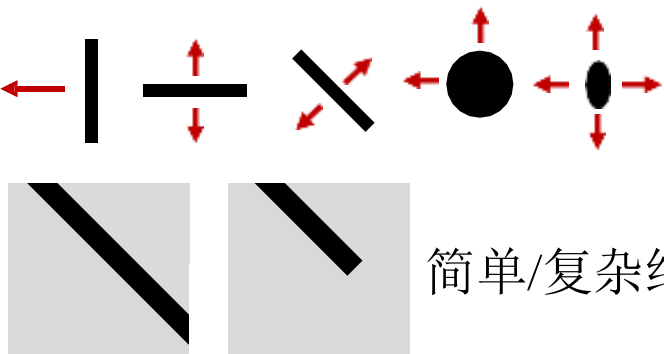
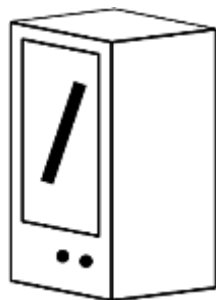


回顾——导论

计算机视觉： 建立系统对视觉数据进行处理、感知、推理

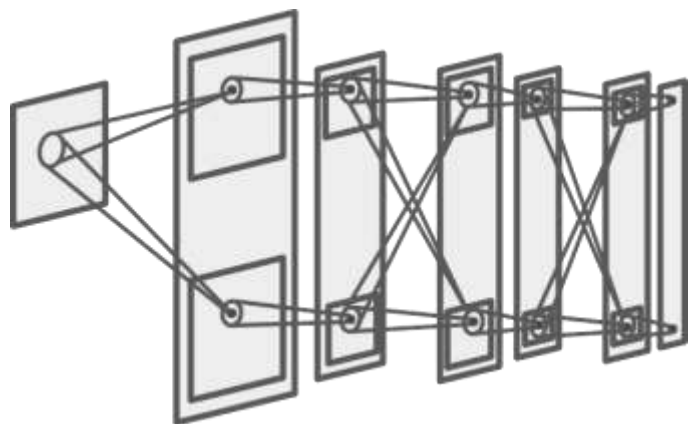
- 图像图形学
 - 我们怎么获得视觉信息
 - 无监督学习，更关注数学建模
- 视觉感知与理解
 - 我们能从视觉信息中获得什么
 - 监督学习，更关注数据驱动
- 视觉内容生成
 - 我们怎么创造新的视觉信息
 - 自监督学习，模型与数据兼备

从“类脑”到“认知机”，再到深度学习

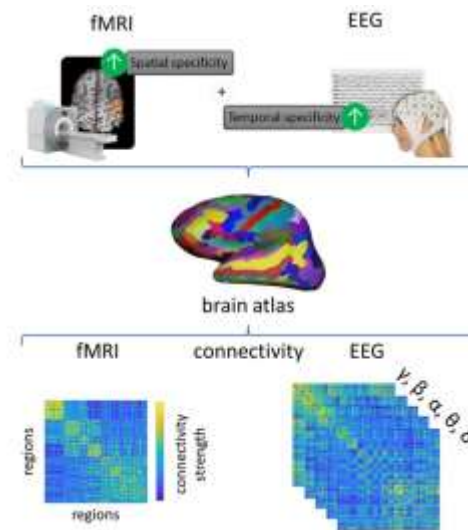


简单/复杂细胞

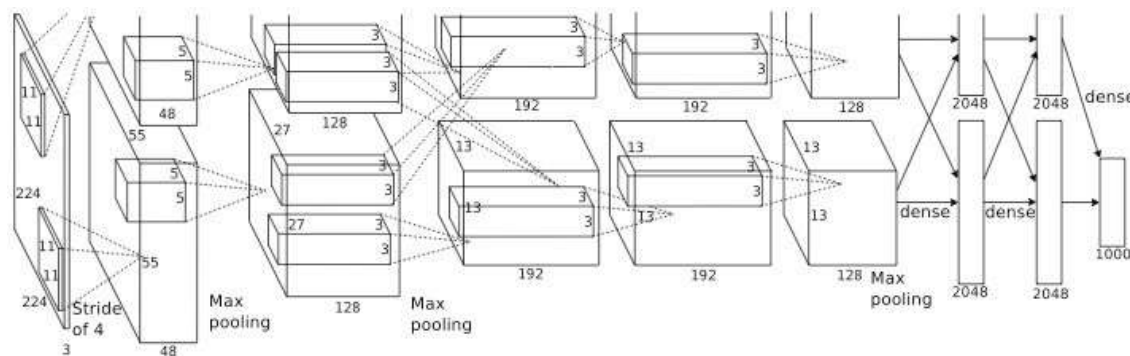
No response



认知机

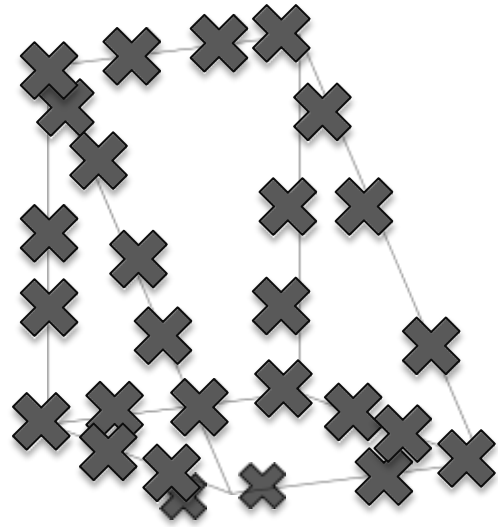


fMRI/EEG



AlexNet

对视觉信息的建模



Input image



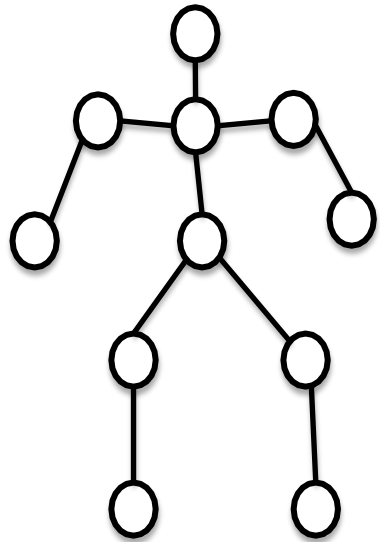
Edge image



2 1/2-D sketch

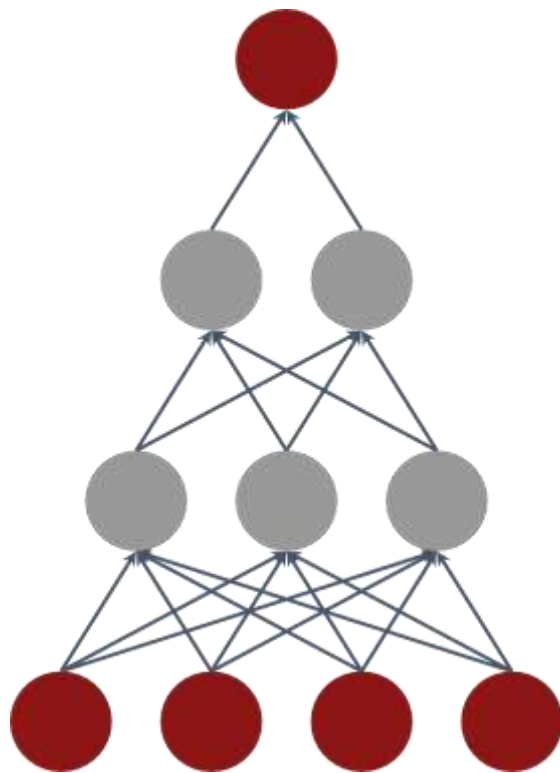


3-D model





计算资源



算法



数据

图像图形学

相机与投影几何

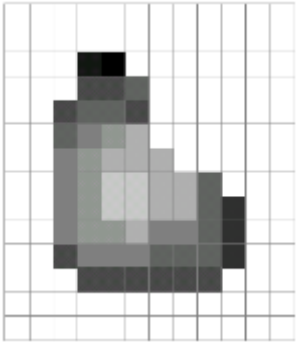
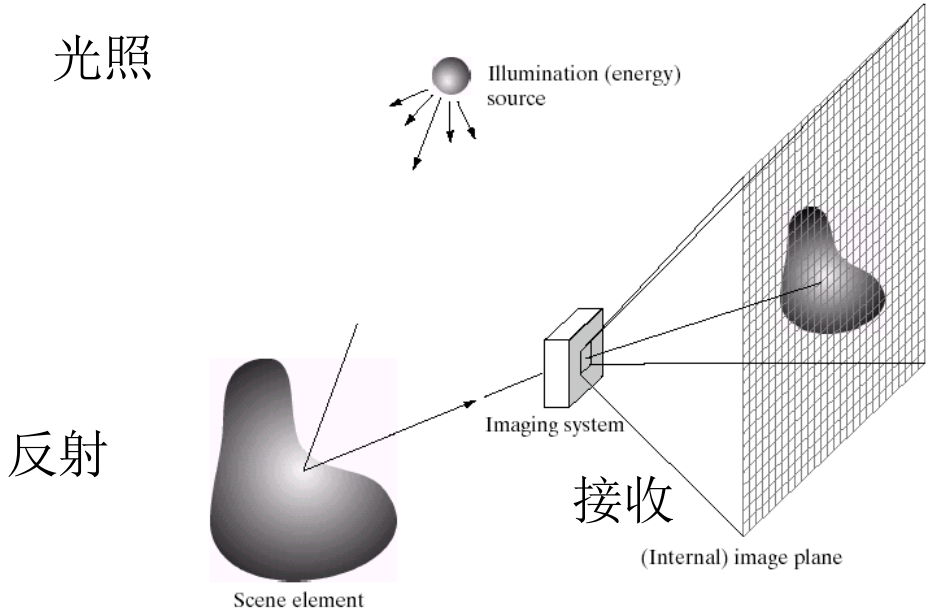
计算机视觉

- 图像图形学（Graphics）
 - 我们怎么获得视觉信息
 - 无监督学习，更关注数学建模
- 感知、识别、推理
 - 我们能从视觉信息中获得什么
 - 监督学习，更关注数据驱动
- 生成
 - 我们怎么创造新的视觉信息
 - 自监督学习，模型与数据兼备

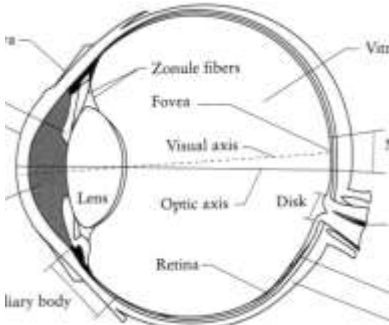
图像图形学

- 我们怎么获得视觉信息
- 无监督学习，更关注数学建模
- 照相机模型与成像原理
- 图像处理基础
- 二维几何
- 三维几何
- 次要目的：提升学生的拍照、修图水平
- 首要目的：为感知理解提供基础

什么是图像？



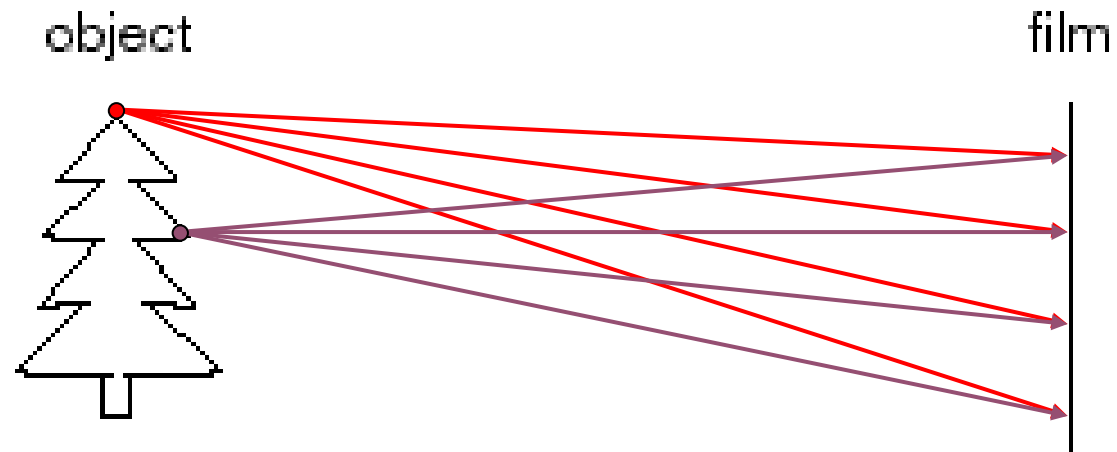
数字照相机



眼睛

Source: A. Efros

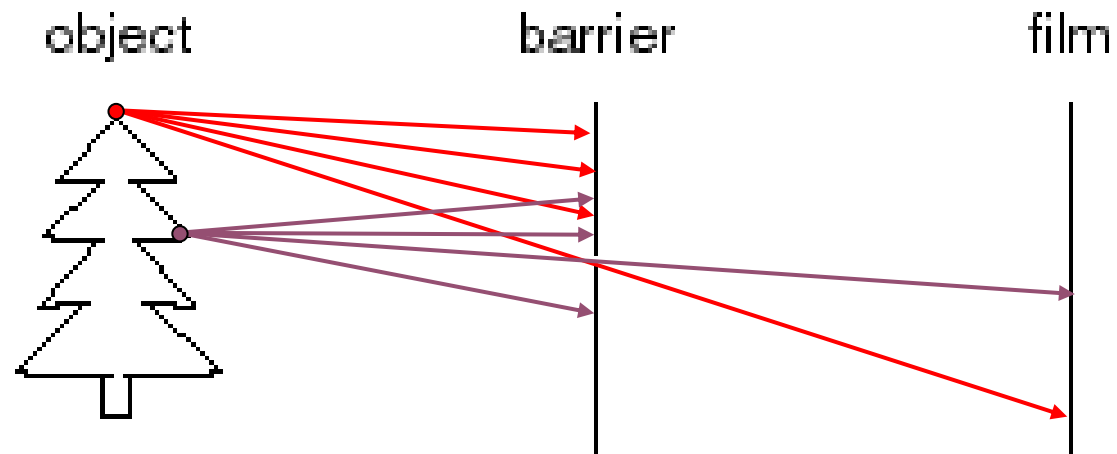
图像的形成



假如你来设计几个照相机...

- 把胶卷放直接在物体前面
- 这是一个好的设计吗？
- 我们会得到什么？

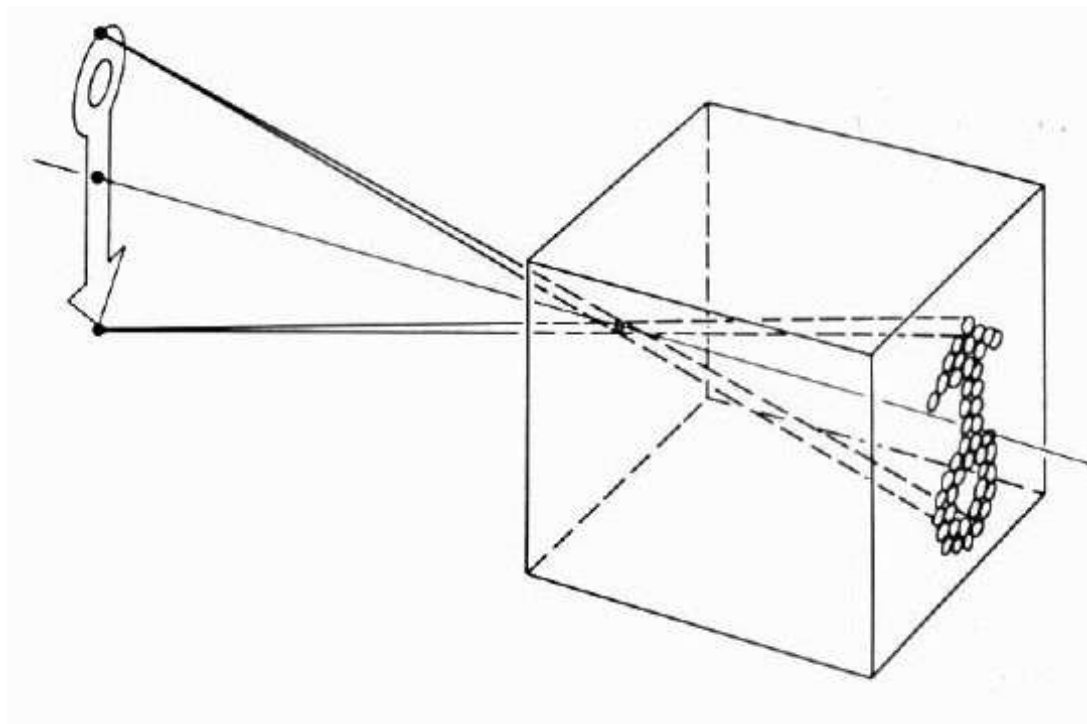
针孔照相机



加一块挡板挡住绝大多数光线

- 没有模糊了！
- 图像在胶卷上会怎么呈现呢？

Camera Obscura 暗箱



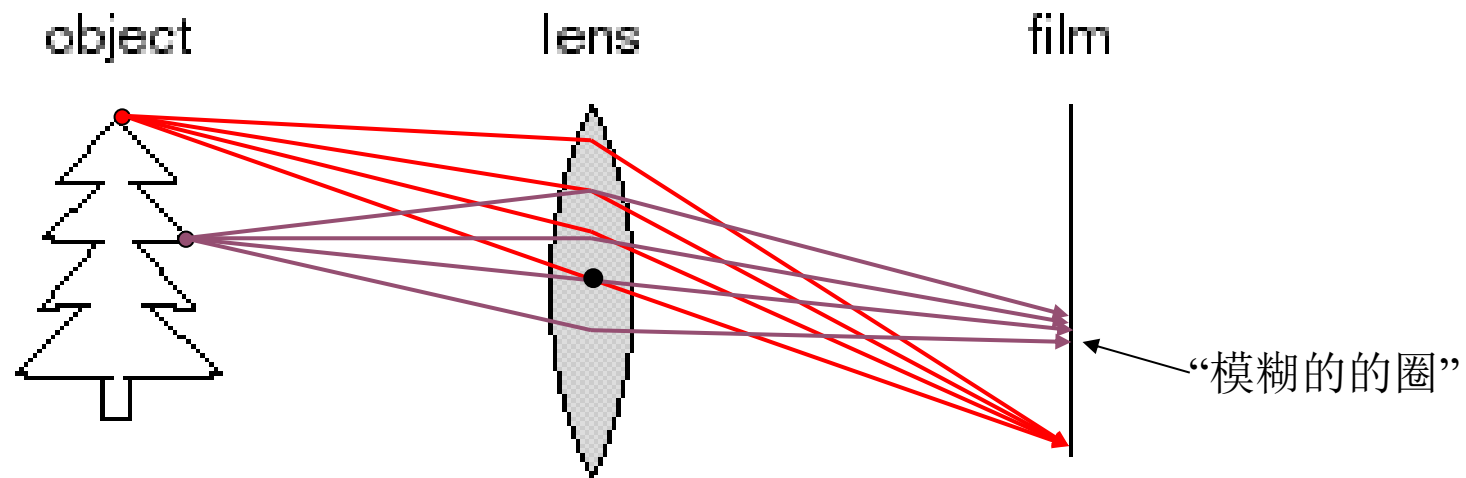
为什么相片还是会有模糊？



我们最常调节照相机的什么设定？



镜头

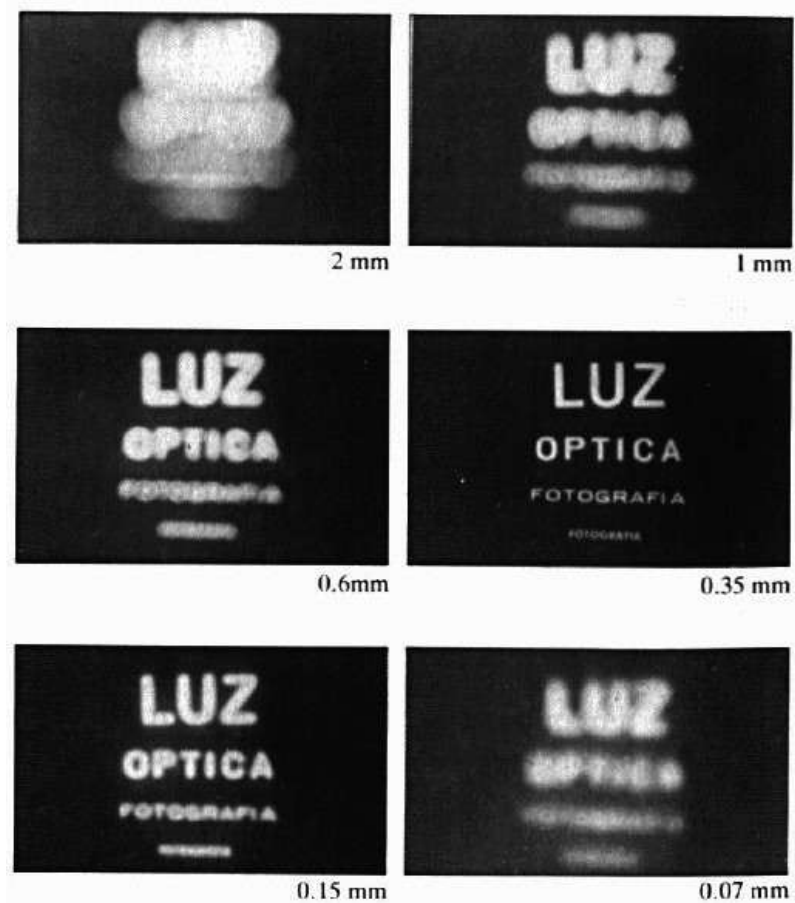


镜头通过折射光线来实现光的聚焦

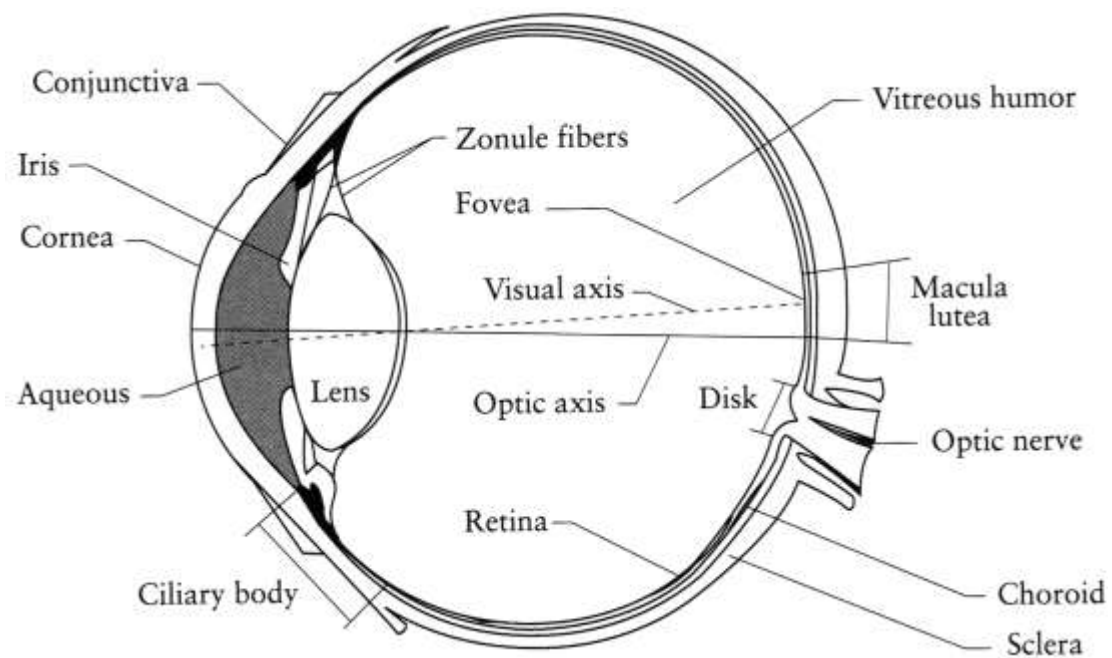
- 会产生“焦点”和“焦距”
 - 其他的点会产生模糊的“圈”
- 改变镜头可以改变焦距

焦距

不同焦距可以调节
清晰度



眼睛



人眼可以近似认为是一个照相机

- **Iris** 虹膜 – 控制瞳孔大小
- **Pupil** 瞳孔 – 调节“光圈”
- 胶卷在哪里?
 - **retina** 视网膜

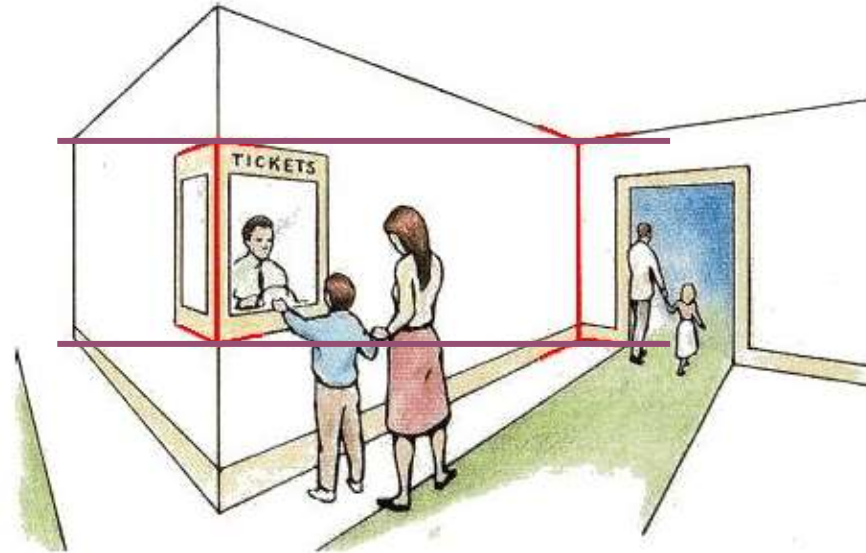
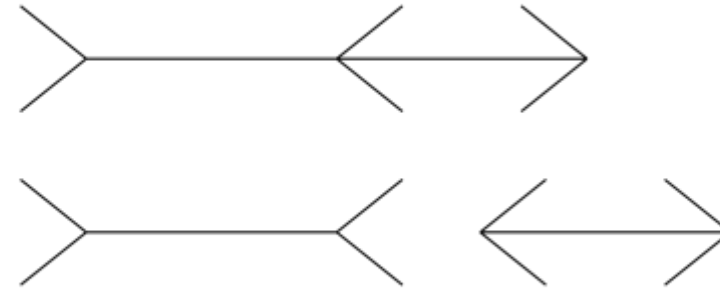
Projection 投影: What You See (with Eyes) is What You Get?



Projection 投影：换个角度看



Müller-Lyer Illusion

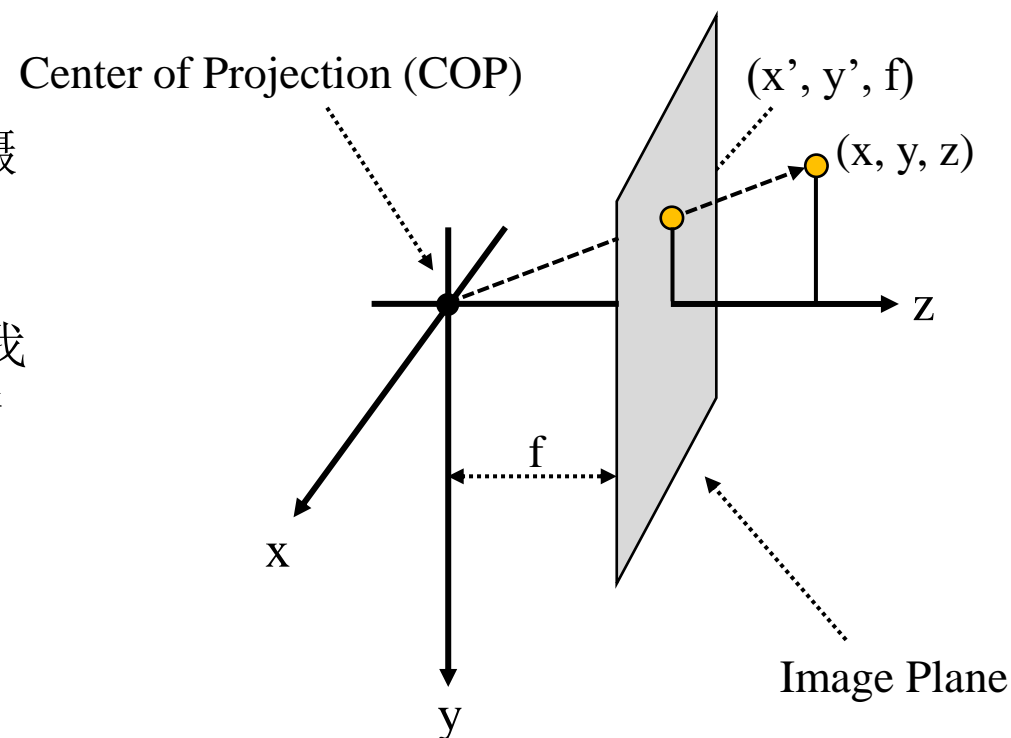
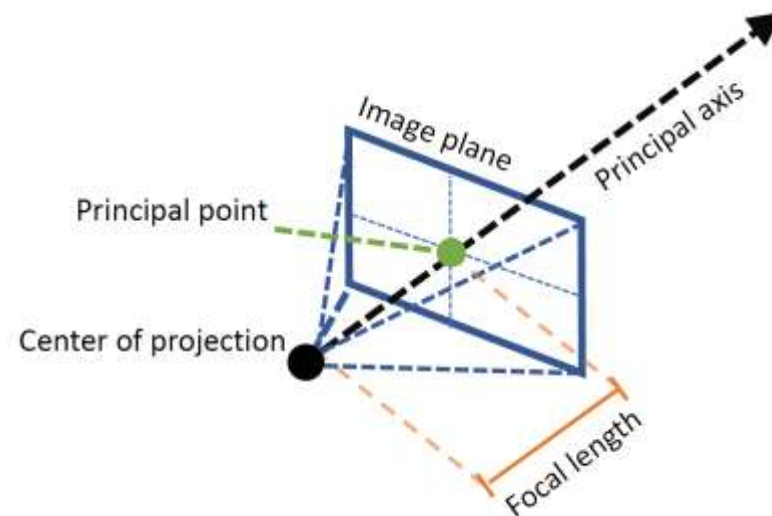


https://en.wikipedia.org/wiki/Müller-Lyer_illusion

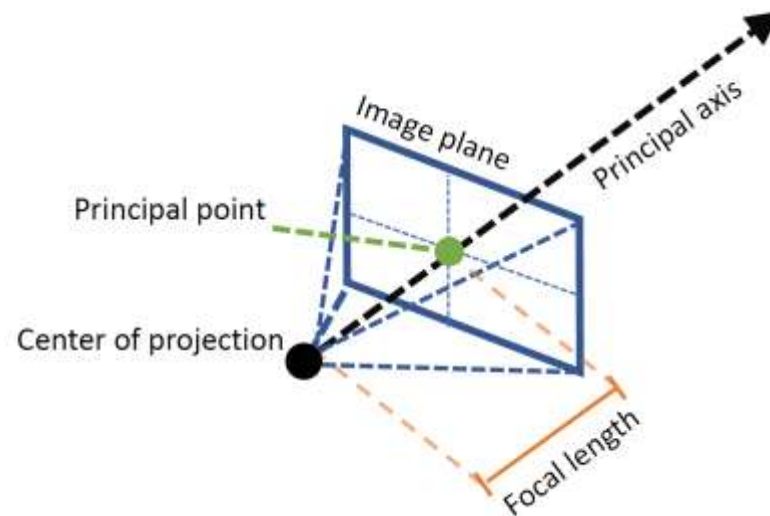
Modeling projection: 投影模型

- 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



Modeling projection: 投影模型



- 投影公式

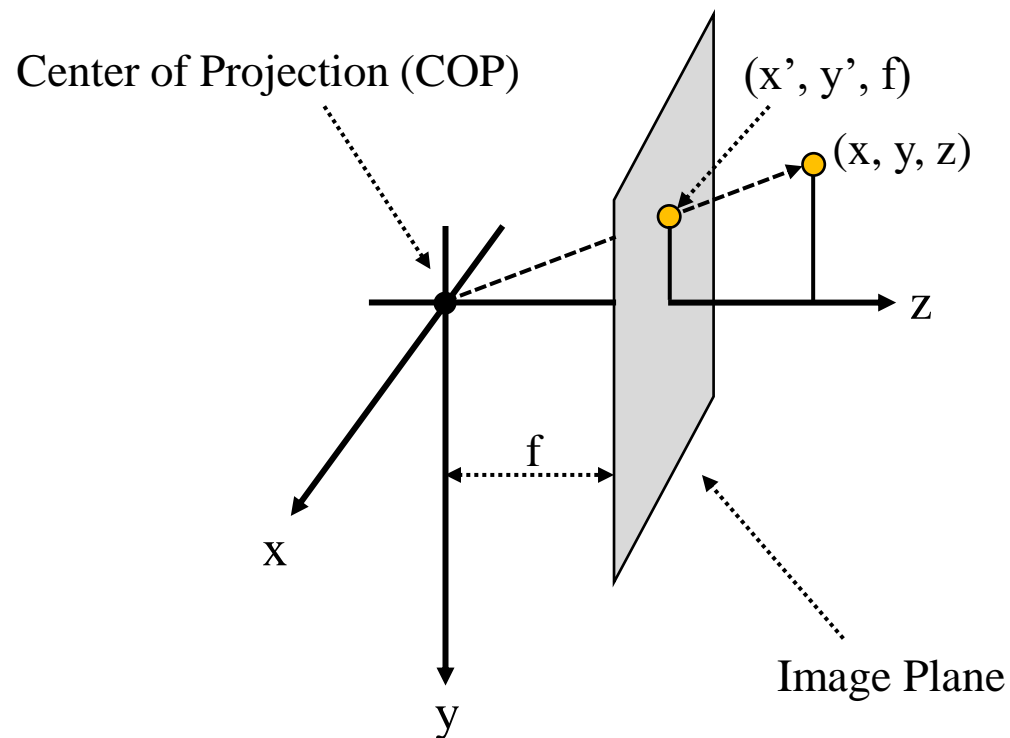
- 三维点投影到二维图像平面上: from (x,y,z) to COP

- 相似三角形

$$(x, y, z) \rightarrow (f \frac{x}{z}, f \frac{y}{z}, f)$$

- 在图像平面:

$$(x, y, z) \rightarrow (f \frac{x}{z}, f \frac{y}{z})$$



Homogeneous coordinates : 齐次坐标

- 投影模型并非线性变换 ($/z$), 不方便用矩阵形式表示
 - 使用齐次坐标

$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad \text{场景齐次坐标}$$

从齐次坐标转换到三维坐标

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

Perspective Projection: 透视投影

把投影在齐次坐标用矩阵相乘表示:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z/f \end{bmatrix} \Rightarrow \left(f \frac{x}{z}, f \frac{y}{z} \right)$$

继而转化为平面坐标

这就是 **透视投影**

- 左边的矩阵是 **投影矩阵**

透视投影

如果最后一列全部是0，放缩矩阵会发生什么？

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z/f \end{bmatrix} \Rightarrow (f \frac{x}{z}, f \frac{y}{z})$$

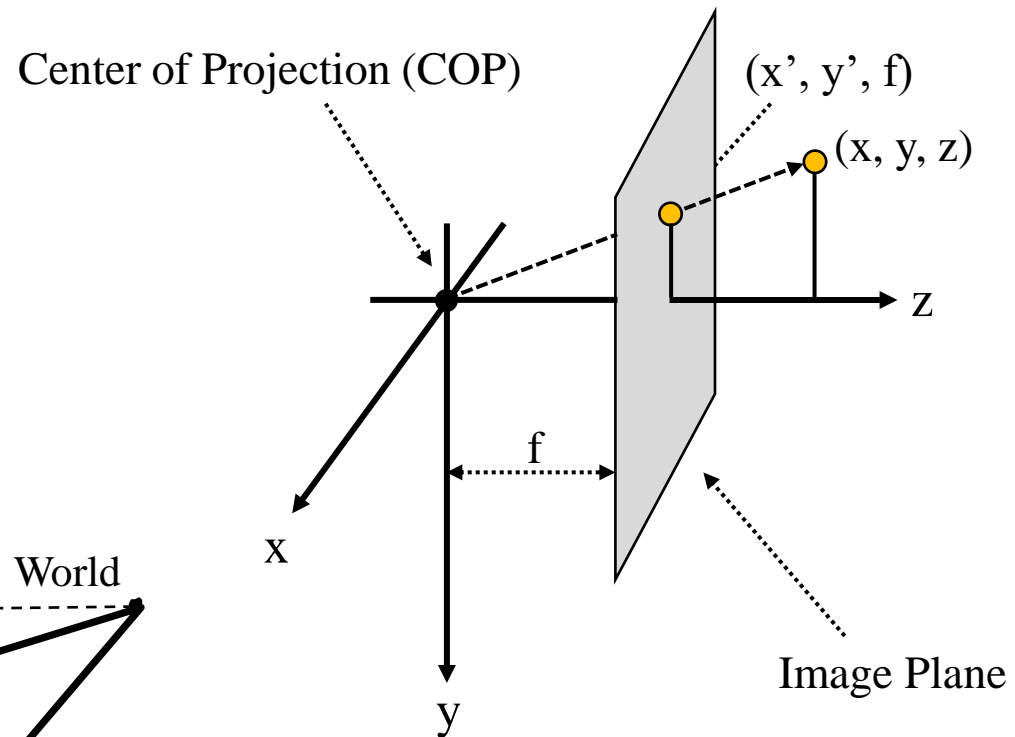
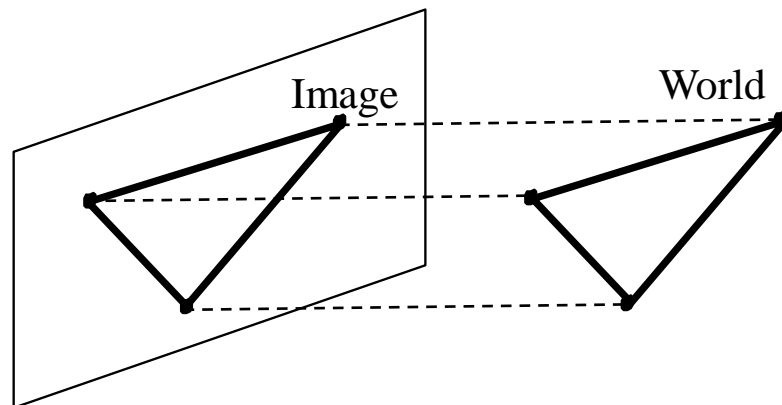
f 放缩:

$$\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} fx \\ fy \\ z \end{bmatrix} \Rightarrow (f \frac{x}{z}, f \frac{y}{z})$$

放缩矩阵的值对结果没有影响！

Orthographic projection: 正交投影

- 一种特殊的透视投影
- 图像平面和原点距离无限远



$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \Rightarrow (x, y)$$

变种

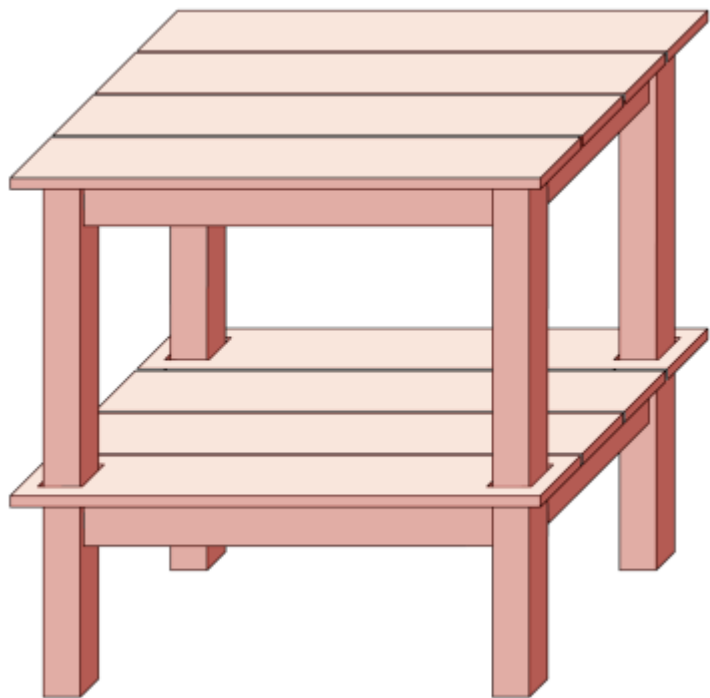
- **Scaled orthographic:** 缩放正交投影，允许将场景中的物体在某个方向上进行缩放，而不会影响其在其他方向上的大小
 - 也叫“Weak Perspective 弱透视投影”

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1/d \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1/d \end{bmatrix} \Rightarrow (dx, dy)$$

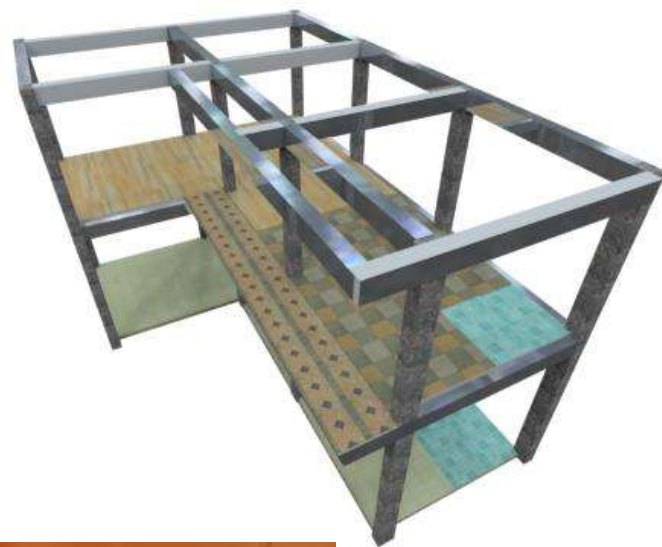
- **Affine projection:** 仿射投影
 - 也叫“Paraperspective 平行透视”

$$\begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

正交投影

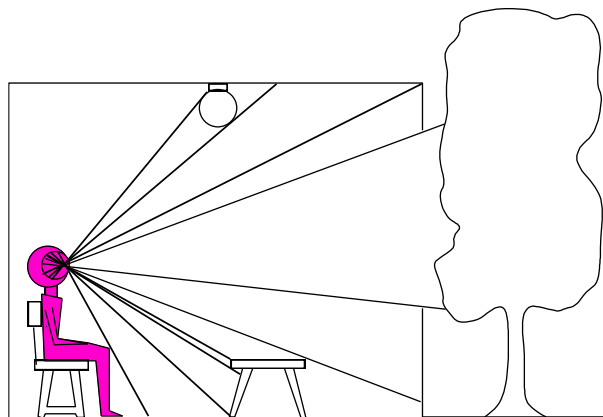


透视投影



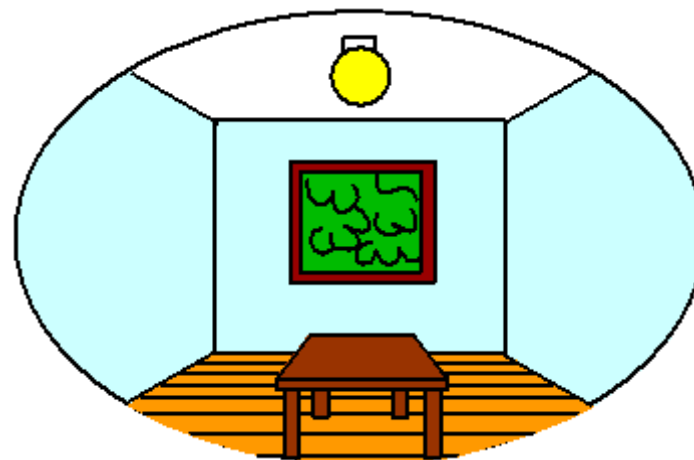
投影与维度下降 (3D to 2D)

3D world



Point of observation

2D image

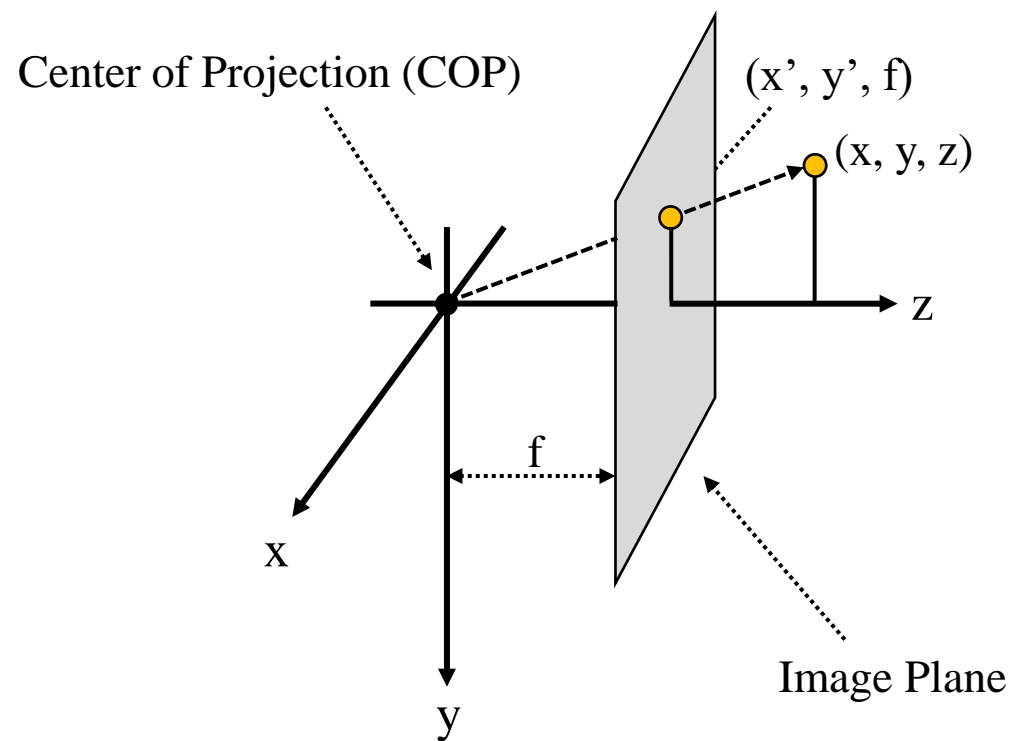


我们损失了什么？

- 视角（相机坐标）
- 距离（长度）

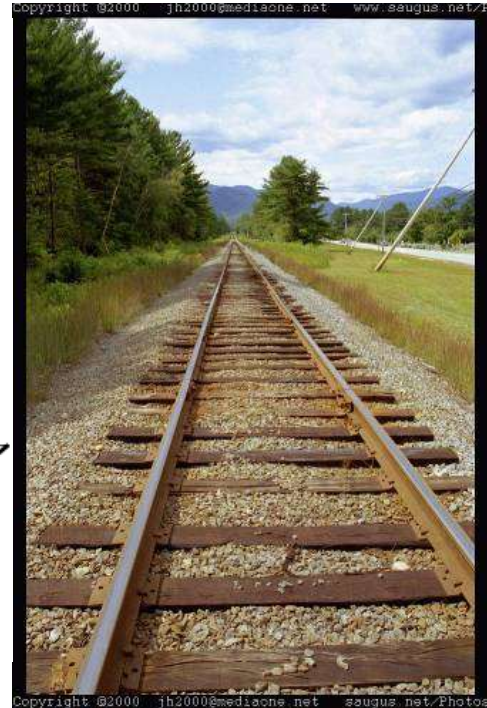
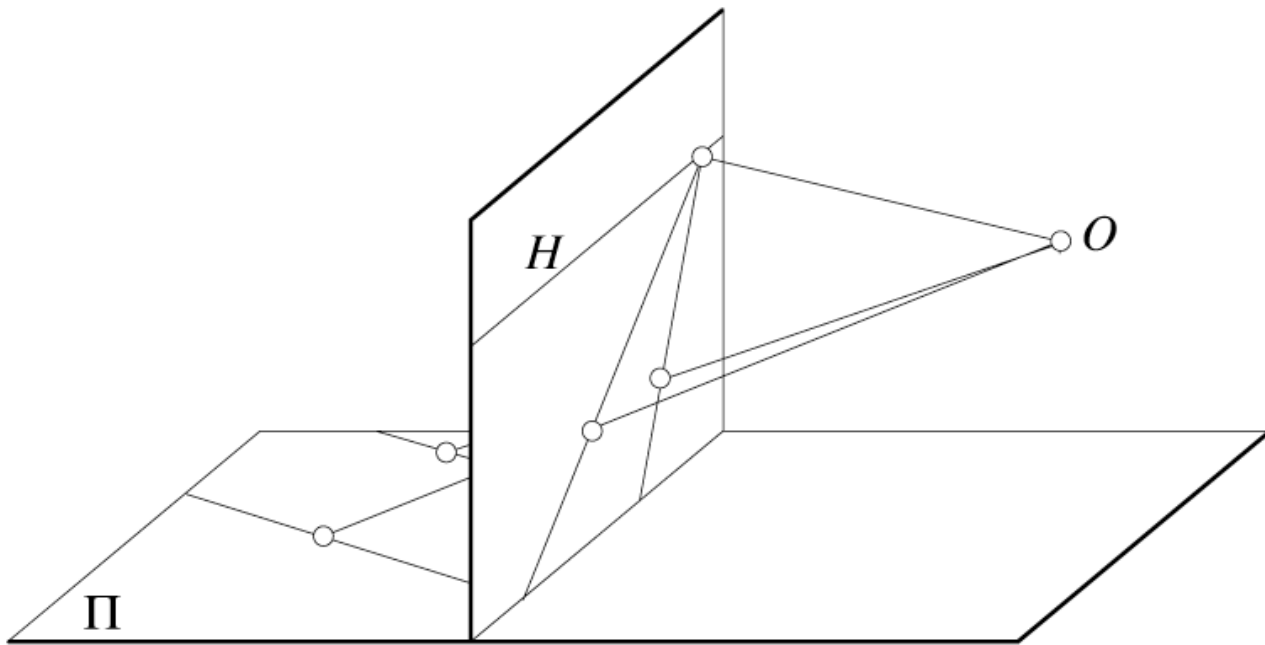
投影的性质

- Many-to-one: 原点出发同一条射线的点会投影到图像的同一点
- 点 \rightarrow 投影之后还是点
- 线 \rightarrow 线
 - 过原点（或延长线）的线会变成点
- 平面 \rightarrow 平面
 - 过原点（或延长面）的面会变成线



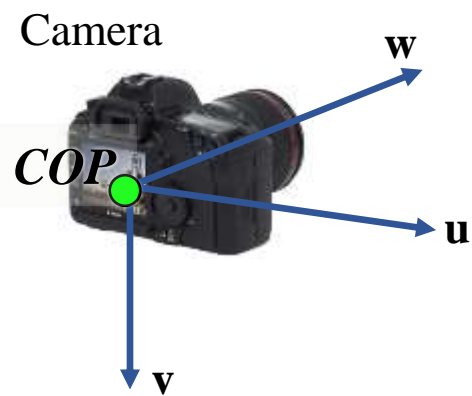
投影的性质

- 两条平行线也有相交的那一天：消失点 vanishing point
 - 图像里有多少个消失点？
 - 图像里是否还是两条平行线没有相交的那一天呢？



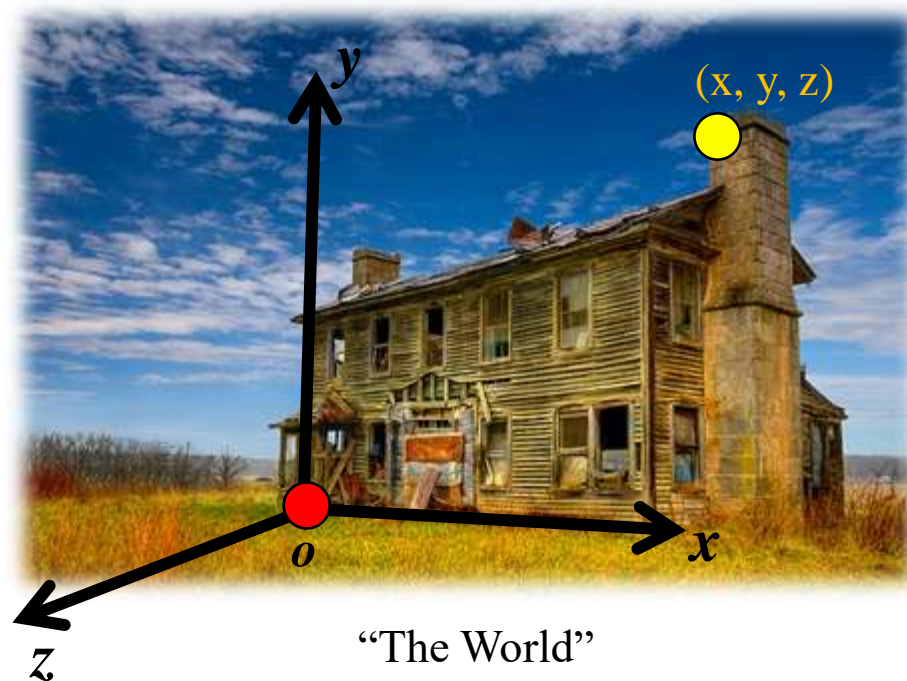
相机参数

- 我们怎么做相机的几何模型？怎么把现实世界的点 (x, y, z) 投影到图像像素点上？

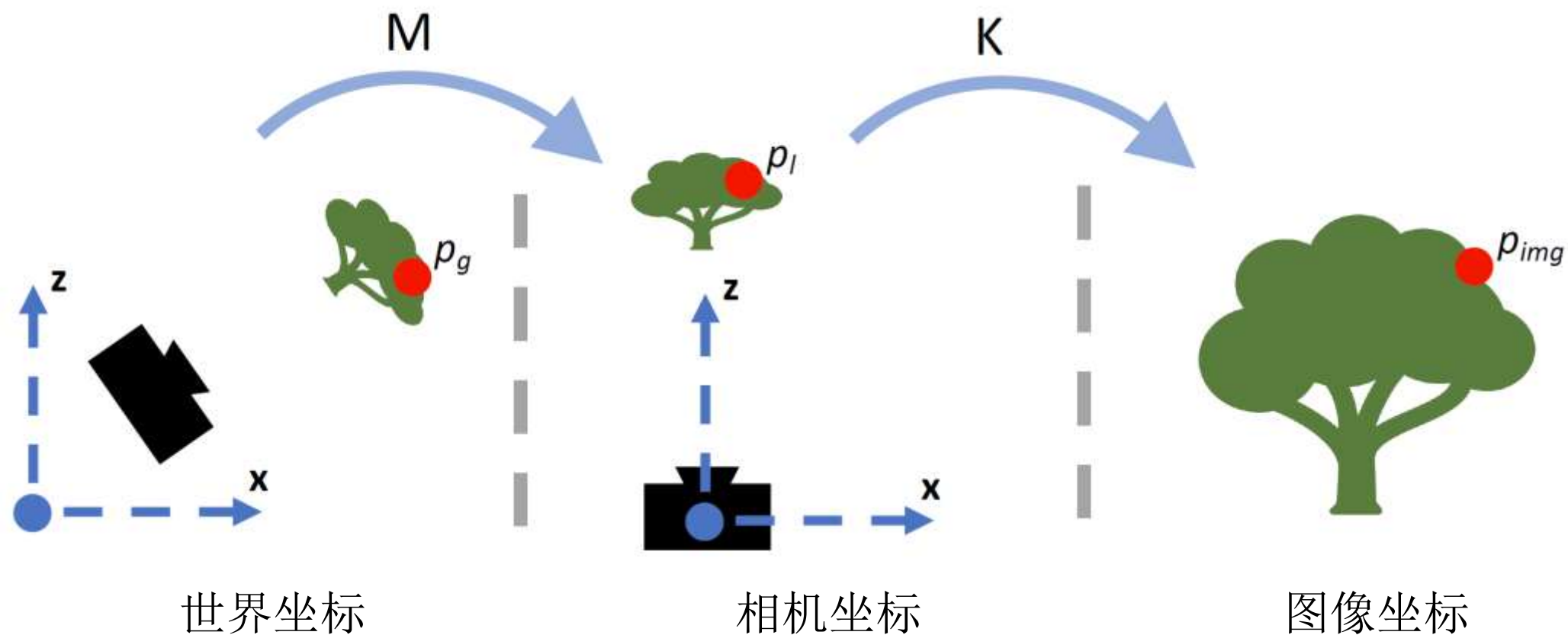


三个重要的坐标系统:

1. 世界坐标
2. 相机坐标
3. 图像坐标



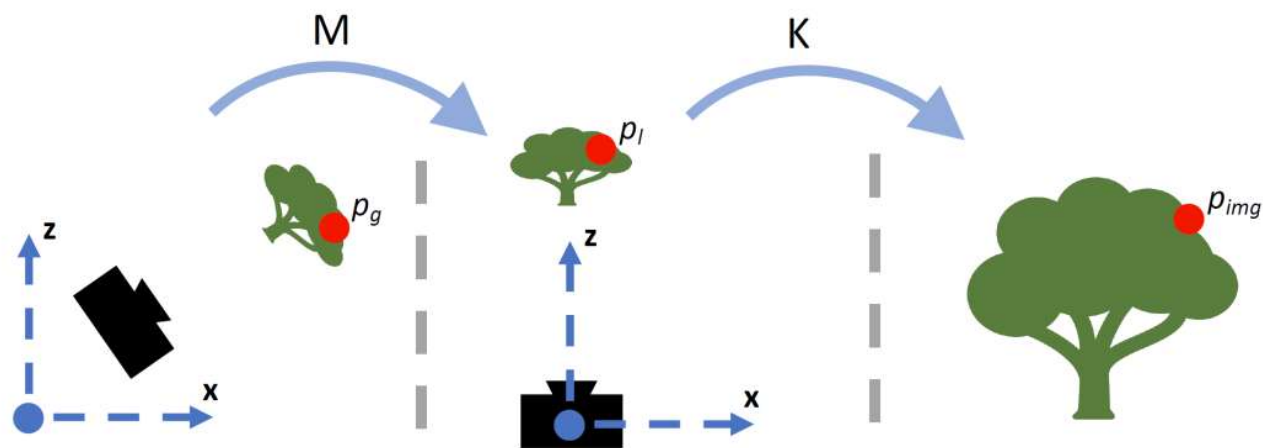
坐标系叠加



相机参数

把世界坐标系下的 (x, y, z) 投影到图像上

- 把 (x, y, z) 转换到相机坐标系下
- 前提是需要什么？
 - 相机位置 (世界坐标系下)
 - 相机朝向 (世界坐标系下)
- 然后把点投影到图像坐标系上
 - 需要知道相机内参



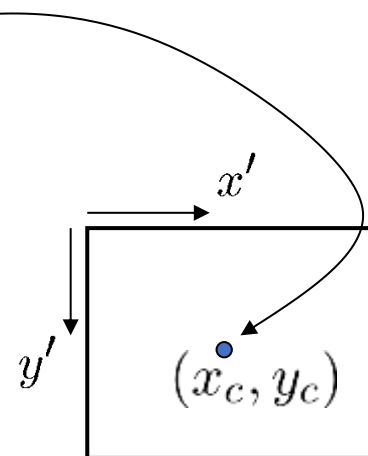
相机参数

相机有很多参数

- 相机光学中心与世界坐标系的变换 \mathbf{T}
- 图像平面旋转角度 \mathbf{R}
- 焦距 f , 图像中心点 (c_x, c_y) , 图像分辨率 α
- 蓝色为“外参,” 红色为“内参”

投影公式

$$\mathbf{x} = \begin{bmatrix} sx \\ sy \\ s \end{bmatrix} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathbf{\Pi X}$$



- 投影将是很多变换的叠加
- 以上参数可以把这些变换拆解

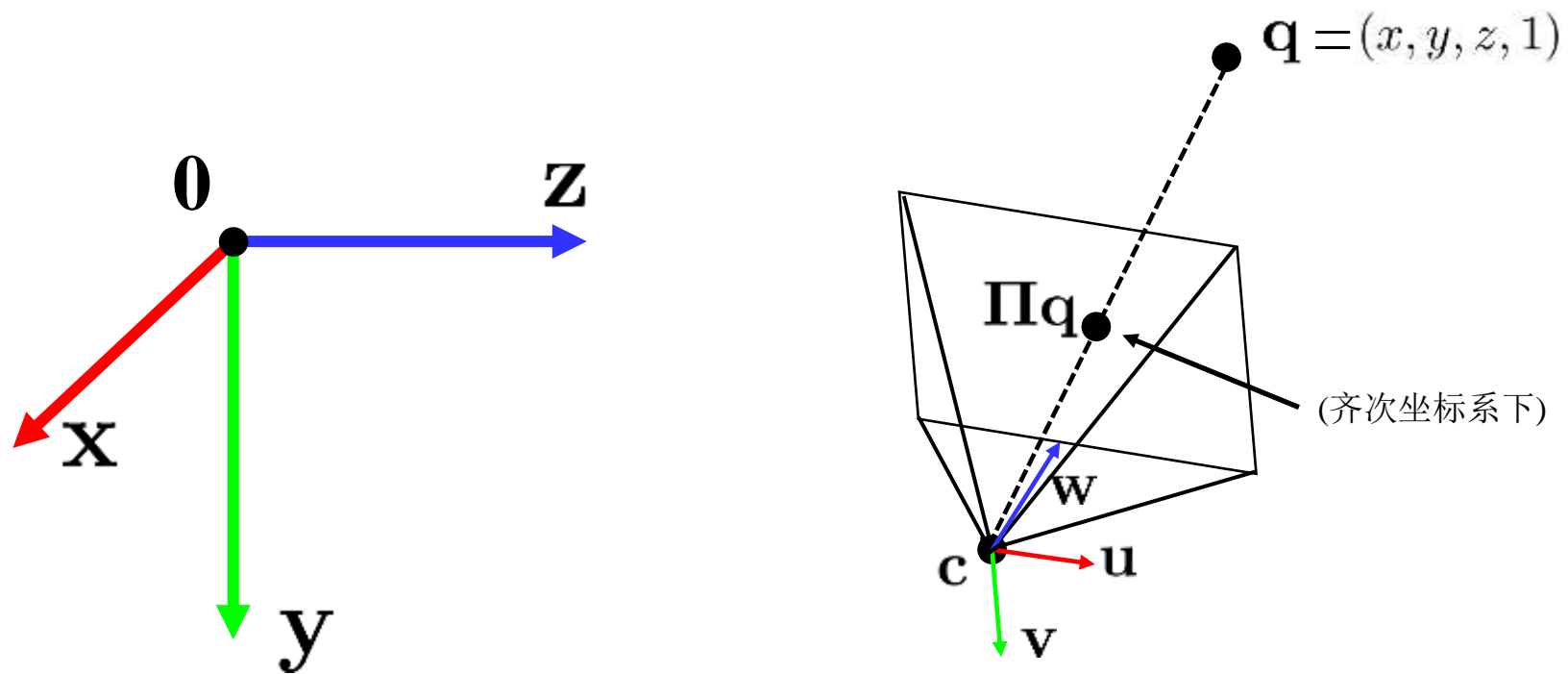
$$\mathbf{\Pi} = \begin{bmatrix} f & s & c_x \\ 0 & \alpha f & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R}_{3 \times 3} & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{T}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 0 \end{bmatrix}$$

内参 投影 旋转 坐标变换

identity matrix

- 请注意：这里只是简单模型，真实相机很复杂！

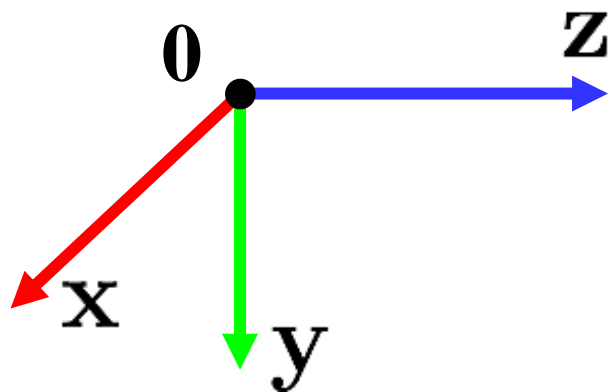
投影矩阵



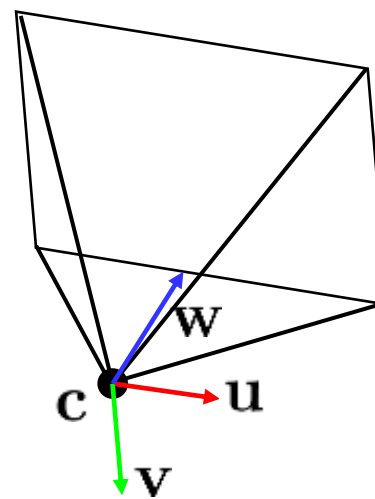
Π : 如何将世界坐标中的点 (x, y, z)
投影到图像坐标 (u, v) 上

外参

- 投影中心为原点, x 轴向右, y 轴向下, z 轴向前

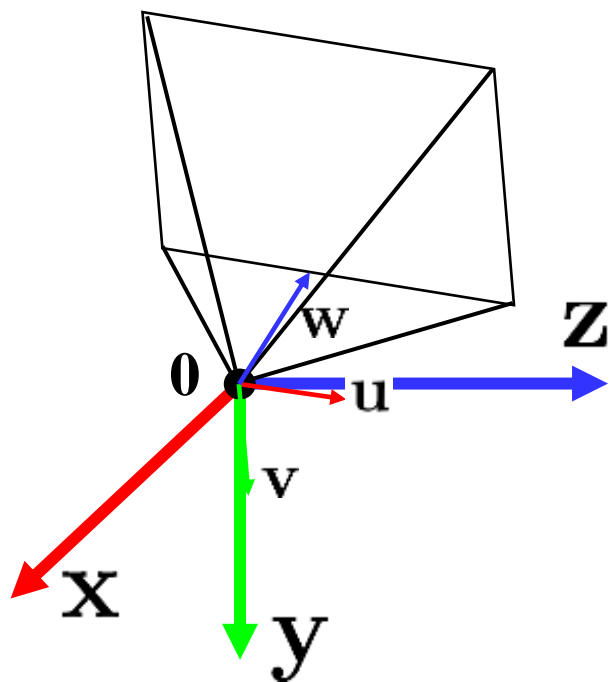


Step 1:世界坐标系的变换 $-c$



外参

- 投影中心为原点, x轴向右, y轴向下, z轴向前



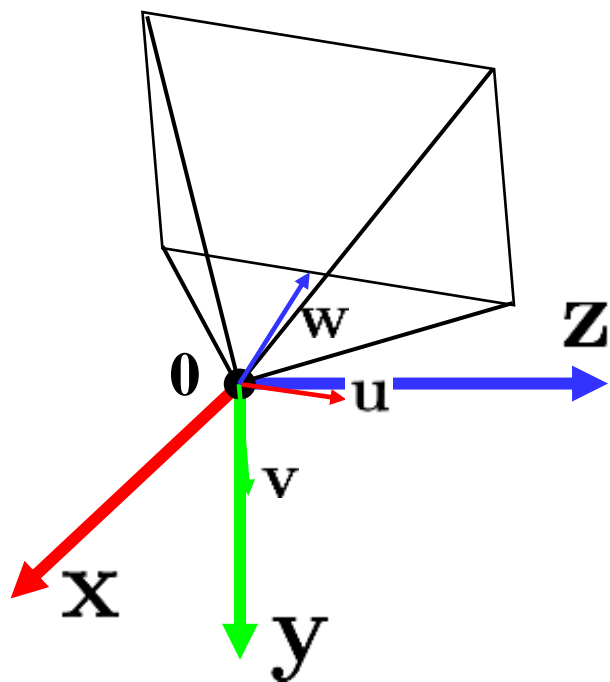
Step 1:世界坐标系的变换 $-\mathbf{c}$

矩阵形式:

$$\mathbf{T} = \begin{bmatrix} \mathbf{I}_{3 \times 3} & -\mathbf{c} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

外参

- 投影中心为原点, x轴向右, y轴向下, z轴向前



Step 1:世界坐标系的变换 $-c$

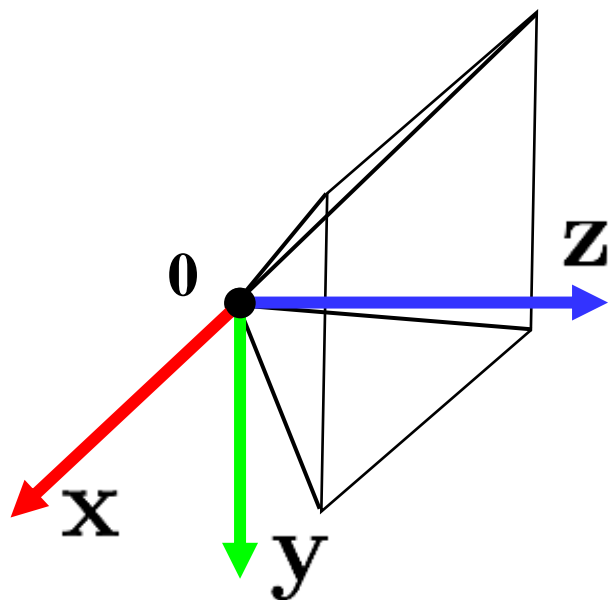
Step 2: 旋转 \mathbf{R}

$\mathbf{R} = \begin{bmatrix} \mathbf{u}^T \\ \mathbf{v}^T \\ \mathbf{w}^T \end{bmatrix}$

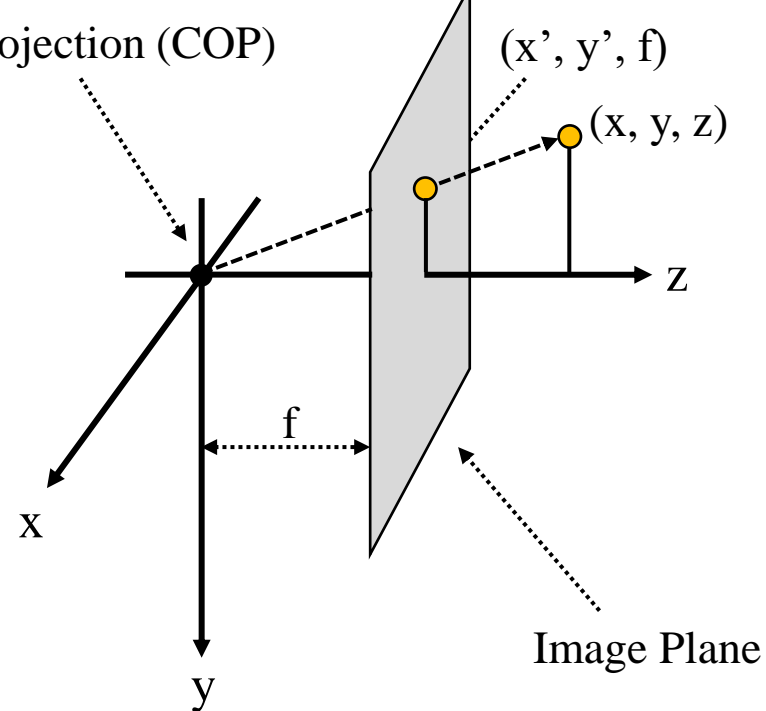
3x3 rotation matrix

外参

- 投影中心为原点, x轴向右, y轴向下, z轴向前



Center of Projection (COP)



Step 1:世界坐标系的变换 $-c$

Step 2: 旋转 \mathbf{R}

$$\mathbf{R} = \begin{bmatrix} \mathbf{u}^T \\ \mathbf{v}^T \\ \mathbf{w}^T \end{bmatrix}$$

(with extra row/column of [0 0 0 1])

透视投影

$$\underbrace{\begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{K}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

K
(内参)

(将三维射线上的点投影到二维平面)

一般来说, $\mathbf{K} = \begin{bmatrix} f & s & c_x \\ 0 & \alpha f & c_y \\ 0 & 0 & 1 \end{bmatrix}$ (上三角矩阵)

α : 纵横比 (一般是1除非不是方形图像)

s : 倾斜 (一般是0 除非图像平面与相机光轴之间的倾斜)

(c_x, c_y) : 主点 (一般是(w/2,h/2) 除非出现偏移)

一般相机的内参模型

$$\mathbf{K} = \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

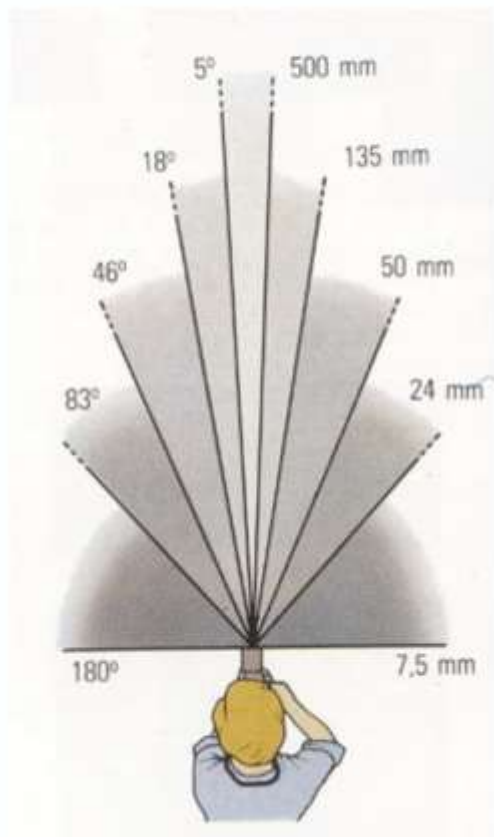
- **2D 仿射变换**：放缩 f (焦距) + 变换 (c_x, c_y) (主点)

焦距

Focal length in practice



不同焦距有不同的
视野



24mm



50mm



135mm



Fredo Durand

焦距

- 可以当作放大



24mm



50mm



200mm



800mm

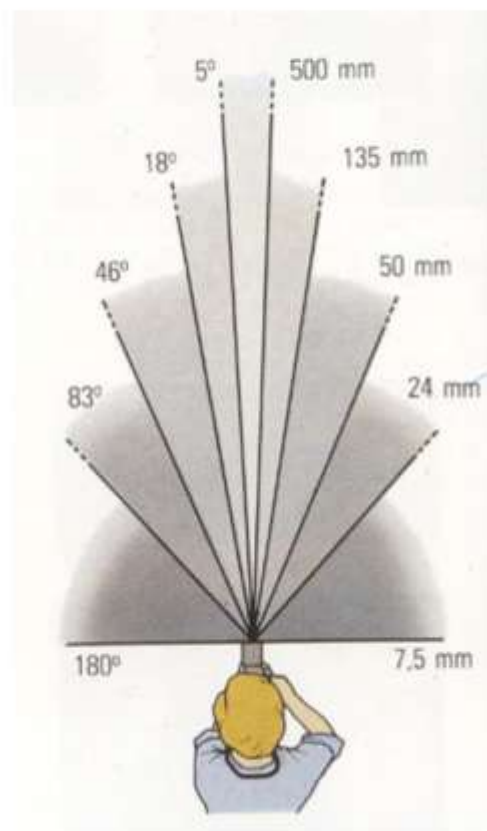


- 也可以认为是 *视野的变换*

焦距 Focal length = cropping



不同焦距可以看清
不同距离的景色



24mm



50mm

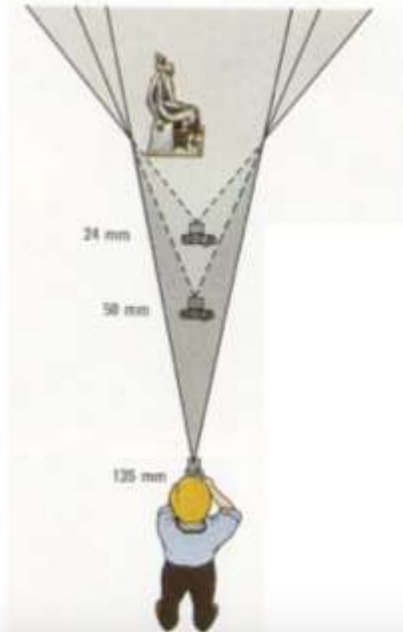


135mm



Focal length vs. viewpoint

- **Telephoto makes it easier to select background (a small change in viewpoint is a big change in background).**



Grand-angle 24 mm



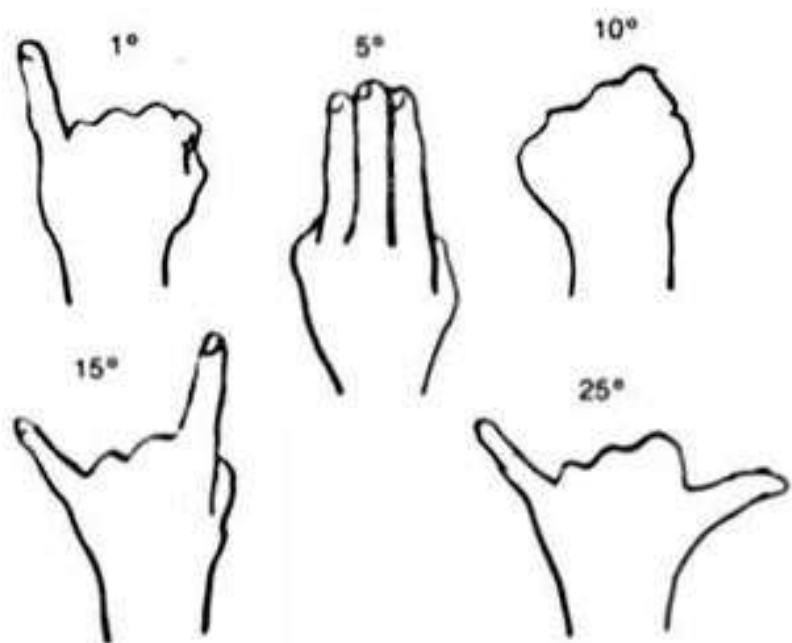
Normal 50 mm









Longue focale 135 mm

视场角与剪裁因子 CheatSheet

APS-C Crop Body Measurement Table



Lens	After 1.62 Multiplier	APS-C Sensor (1.62 lens multiplier) Canon 60D, 7D, 70D, T3i, T4i	Hand Positions
18mm	29.16mm	Three hands wide at full arms length.	
28mm	45.36mm	Slightly less than two hands wide at full arms length.	
35mm	56.7mm	One hand + width of one fist at full arms length.	
50mm	81mm	One hand wide + width of thumb at full arms length.	
55mm	89.1mm	Slightly less than one hand wide at full arms length.	
85mm	137.7mm	Inside edge of thumb to tip of forefinger wide with hand in "L" shape, thumb up.	

固定前景下不同焦距拍摄效果



不同焦距拍摄效果不同



广角（短焦）

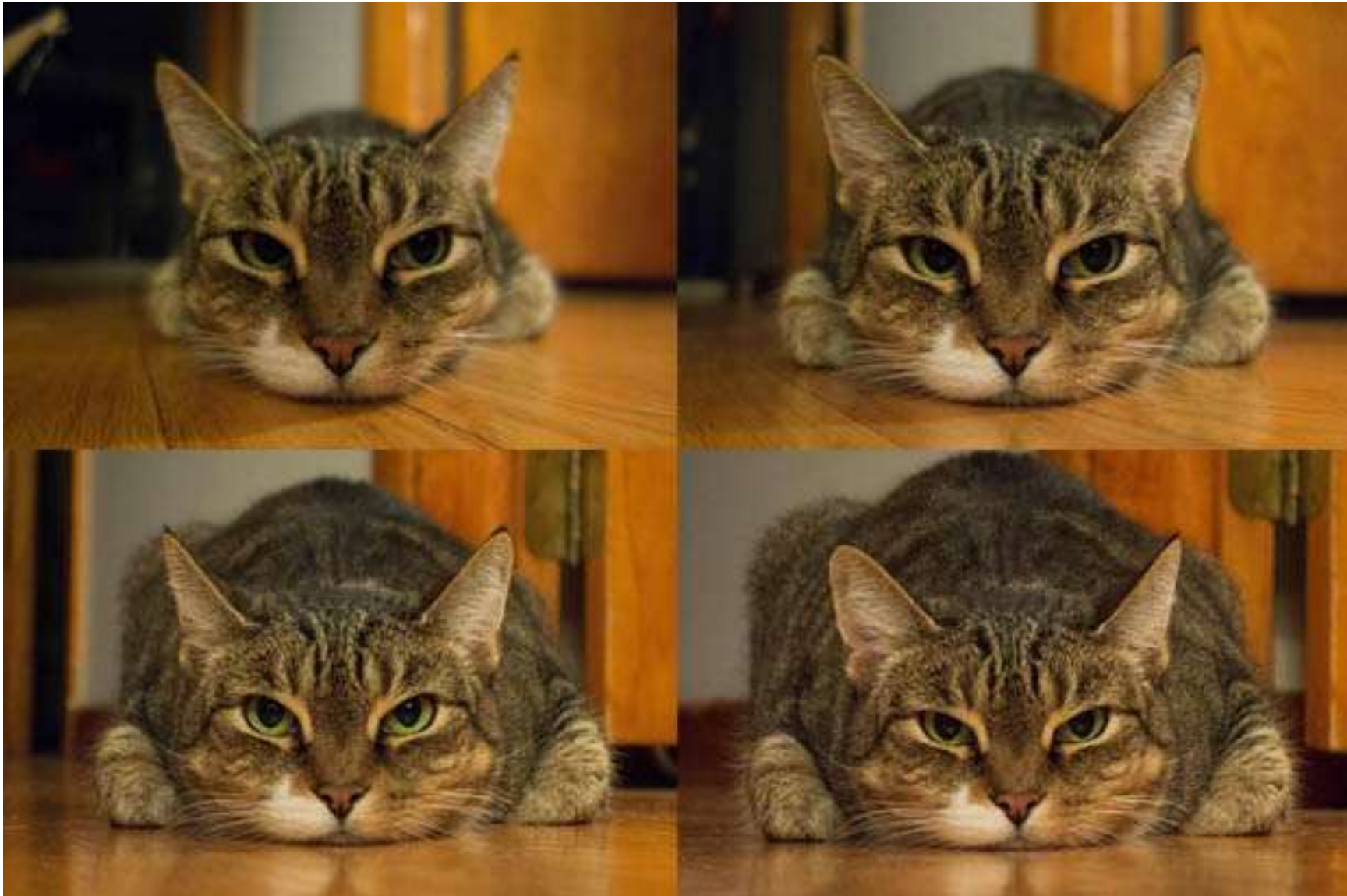


标准镜头



长焦镜头

知道为什么
专业摄影师
都用大长焦
了吧！



<http://petapixel.com/2013/01/11/how-focal-length-affects-your-subjects-apparent-weight-as-seen-with-a-cat/>

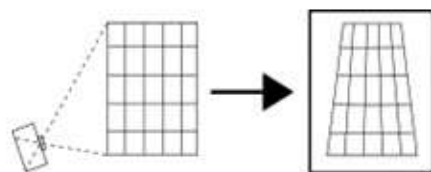
透视失真

- 拍摄建筑物时经常出现这种情况：

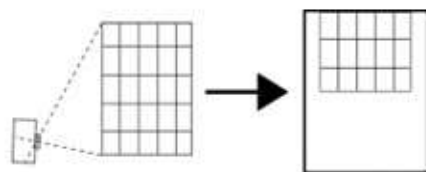


透视失真

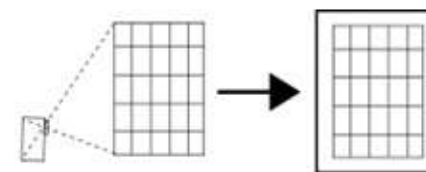
- 怎么办呢？



向上倾斜摄像头
建筑物上下比例失衡

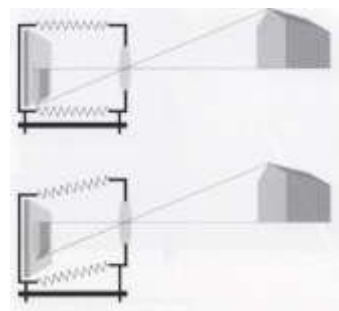


保持照相机水平，则建筑
物上面部分拍不到



把透镜向上移就可以
解决这个问题

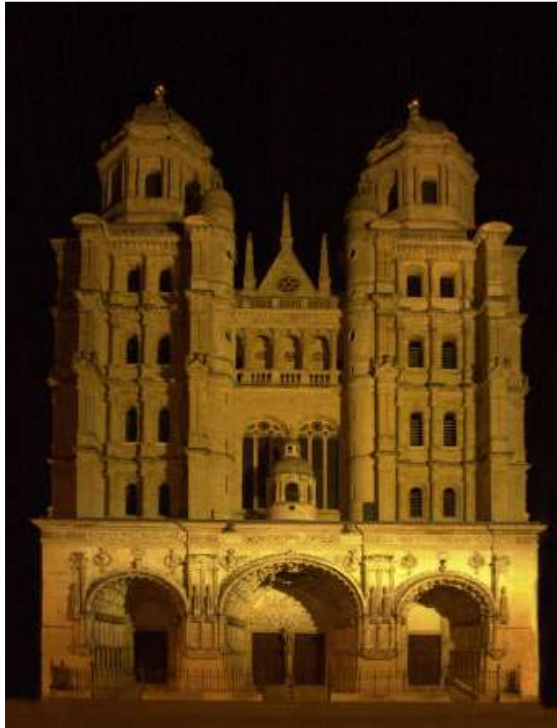
- 解决方案: view camera (lens shifted w.r.t. film)



http://en.wikipedia.org/wiki/Perspective_correction_lens

透视失真

- 效果:



透视失真

- 水平失真如何解决？

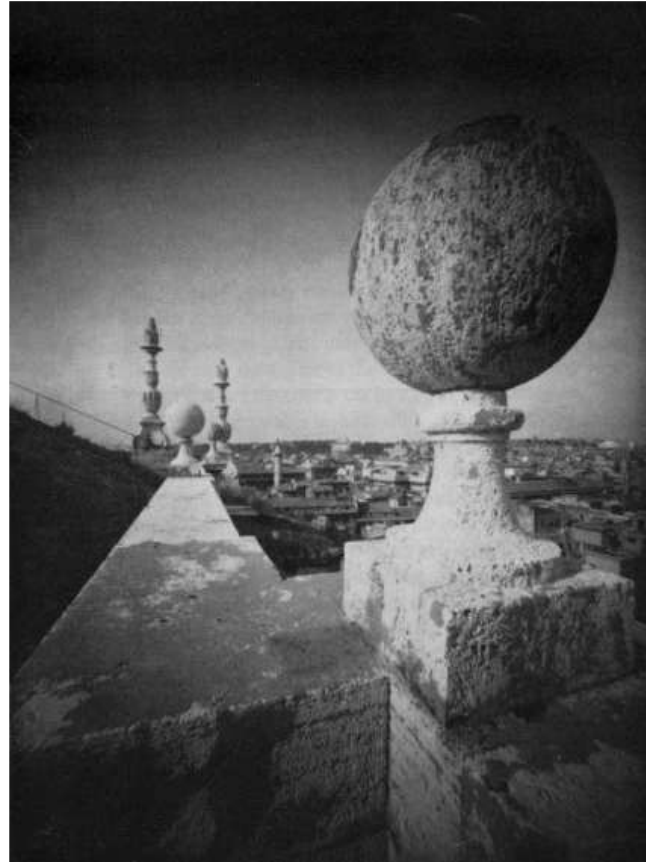


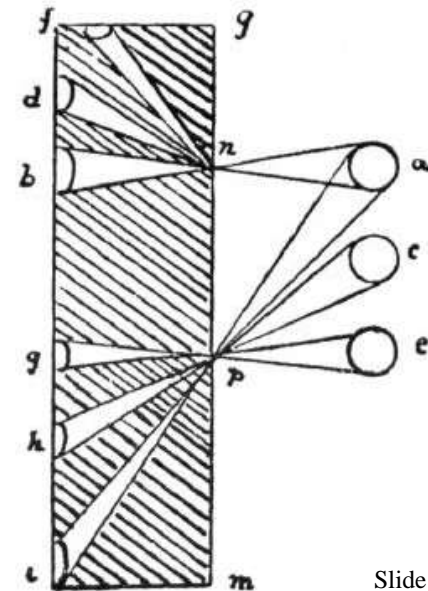
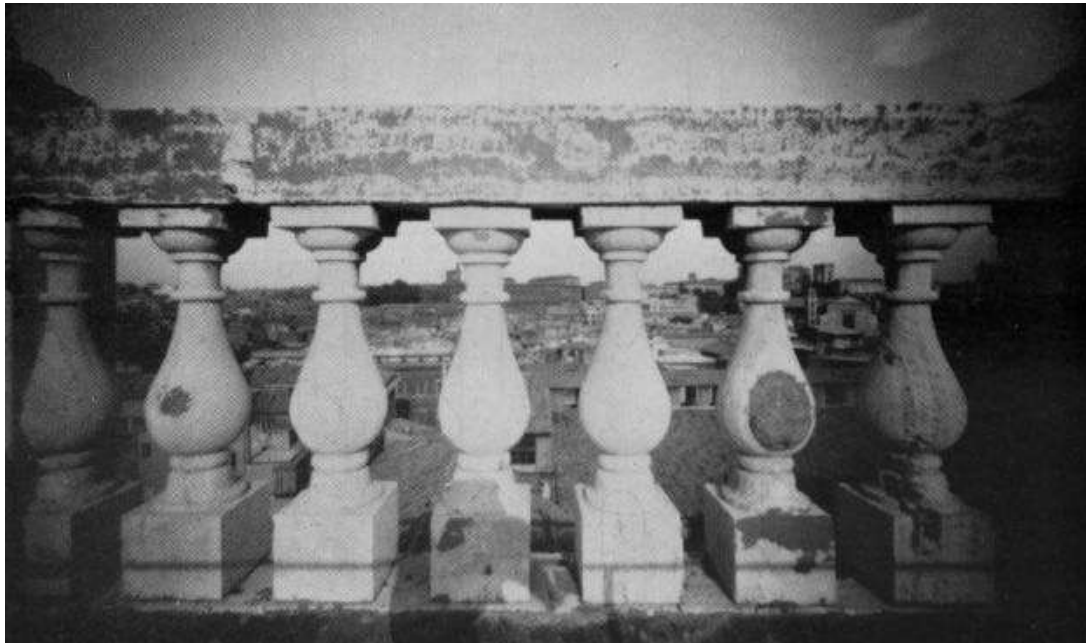
Image source: F. Durand

透视失真: 人像



透视失真

- 外面的柱子看起来更大
- Problem pointed out by Da Vinci



Slide by F. Durand

Distortion-Free Wide-Angle Portraits on Camera Phones



(a) A wide-angle photo with distortions on subjects' faces.



(b) Distortion-free photo by our method.

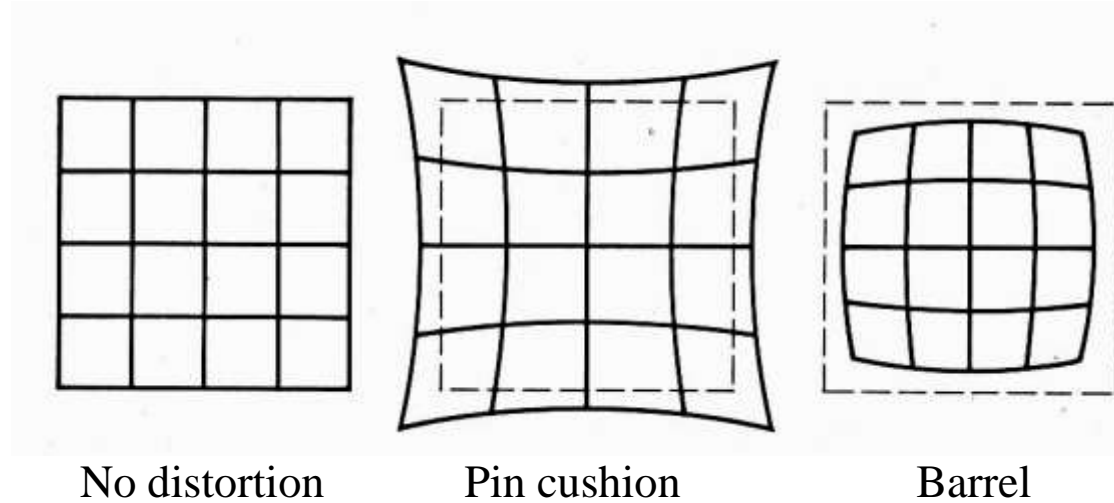
YiChang Shih, Wei-Sheng Lai, and Chia-Kai Liang, Distortion-Free Wide-Angle Portraits on Camera Phones, SIGGRAPH 2019

https://people.csail.mit.edu/yichangshih/wide_angle_portrait/

怎么利用透视失真？



畸变 Distortion

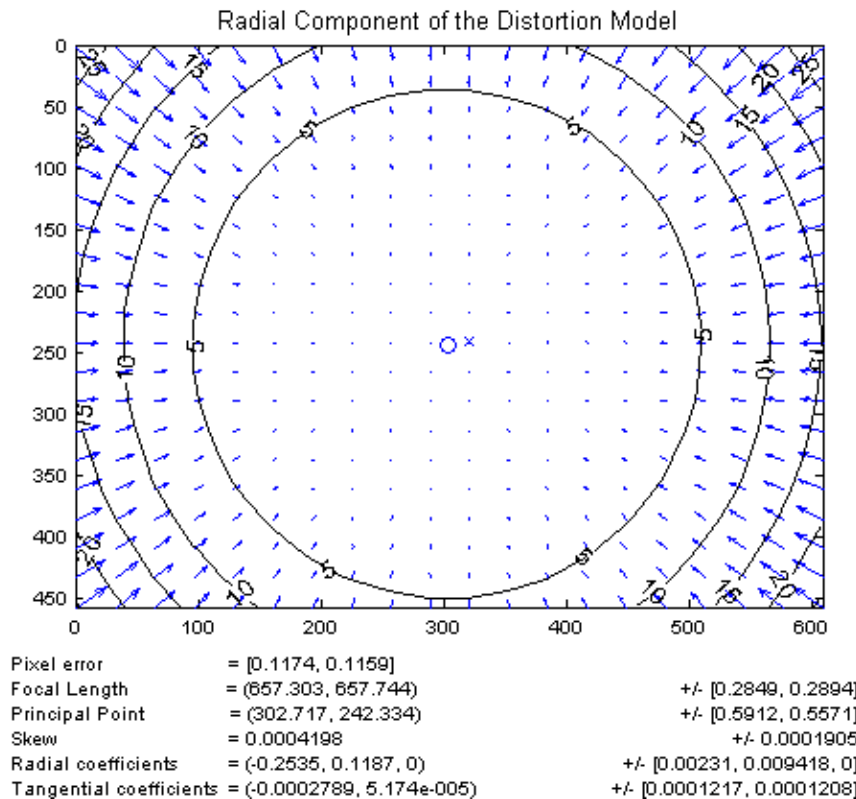


- Radial distortion

- 一般由于镜头瑕疵导致（广角镜头尤为明显）
- 一般在边缘区域出现



Radial distortion



- Arrows show motion of projected points relative to an ideal (distortion-free lens)

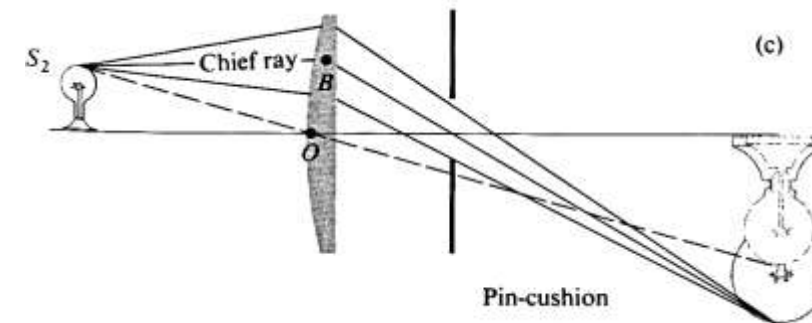
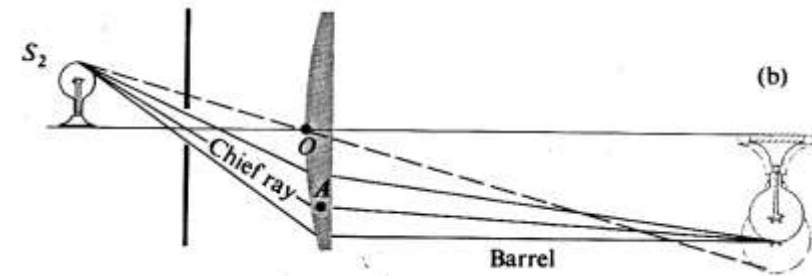
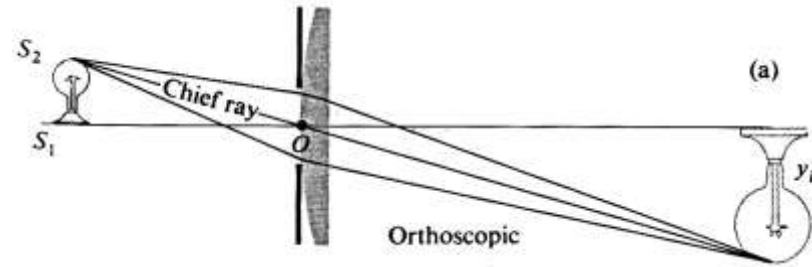
[Image credit: J. Bouguet http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/example.html]

Correcting radial distortion



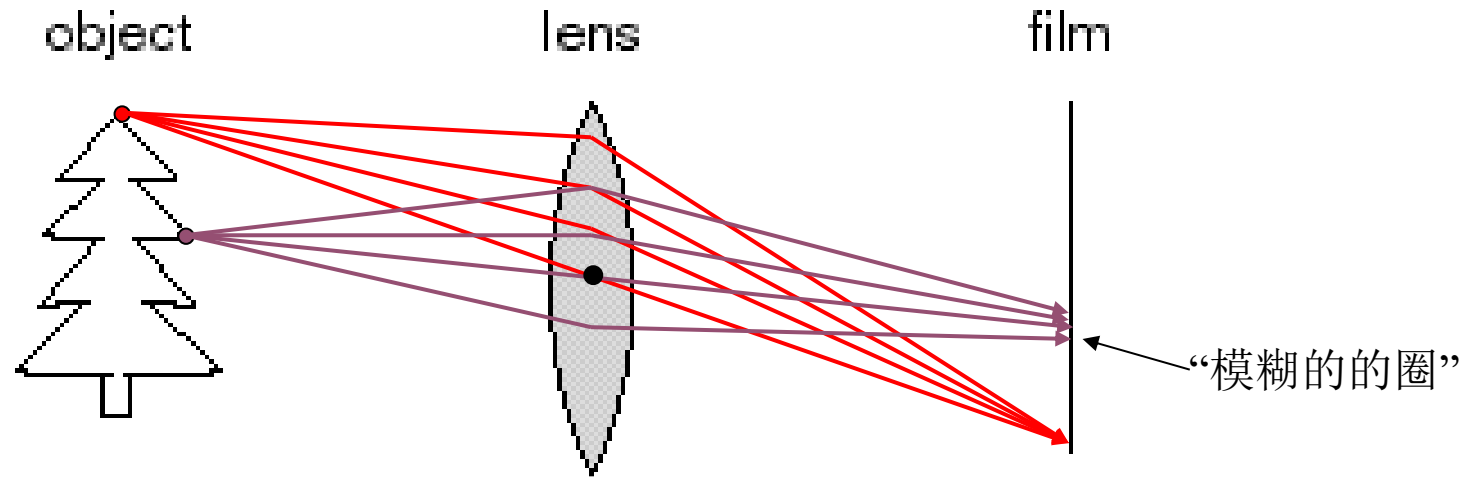
from [Helmut Dersch](#)

Distortion



回顾——相机与投影几何

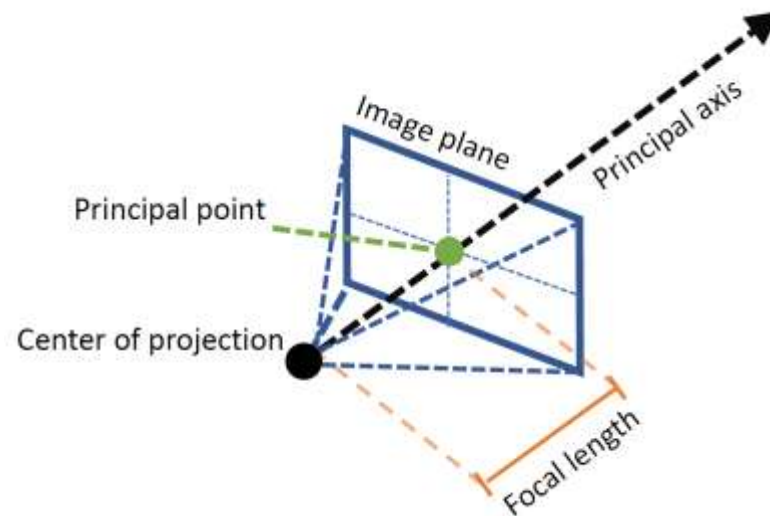
单孔相机与镜头



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

投影模型：三维到二维的线性变换



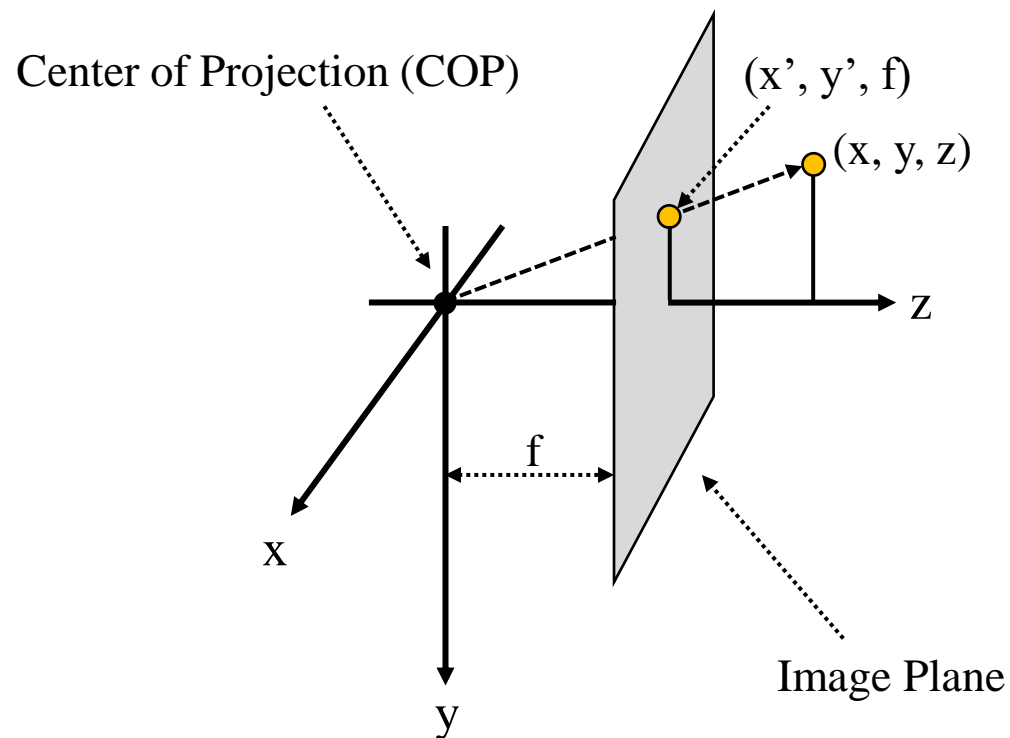
- 投影公式

- 三维点投影到二维图像平面上: from (x,y,z) to COP
- 相似三角形

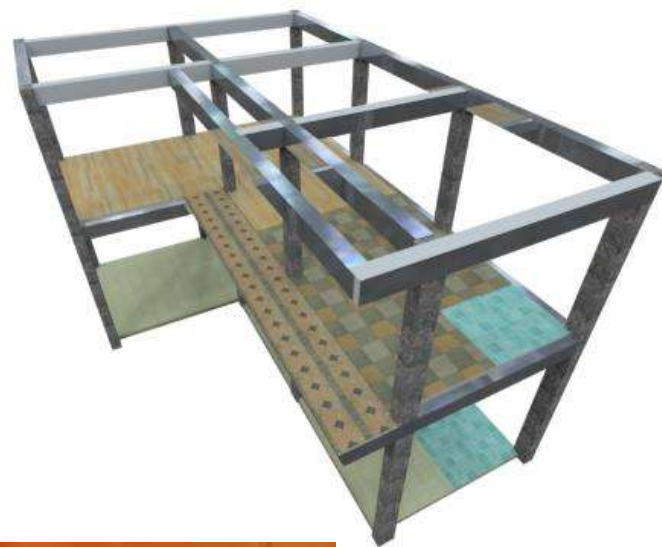
$$(x, y, z) \rightarrow (f \frac{x}{z}, f \frac{y}{z}, f)$$

- 在图像平面:

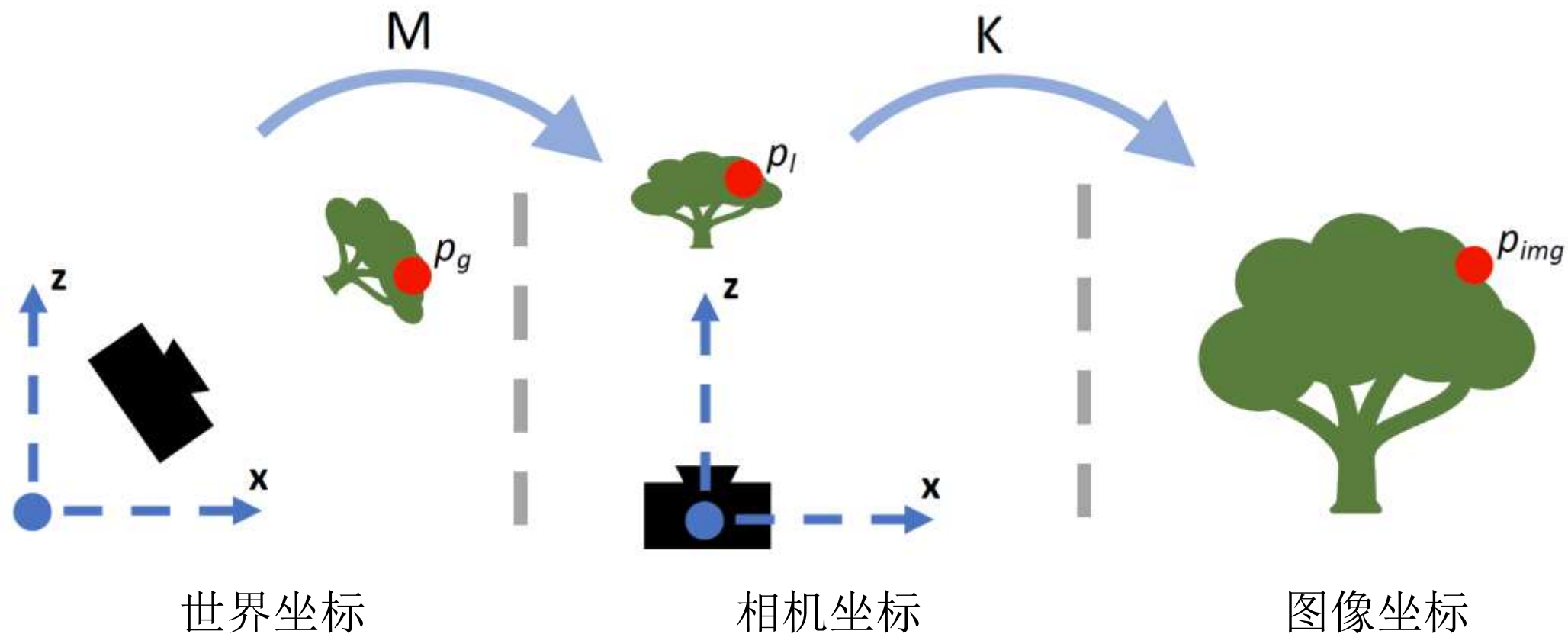
$$(x, y, z) \rightarrow (f \frac{x}{z}, f \frac{y}{z})$$



透视投影：近大远小



相机投影：坐标系叠加



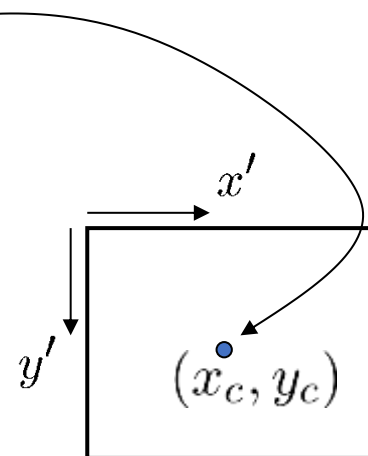
相机参数与投影矩阵的关系

相机有很多参数

- 相机光学中心与世界坐标系的变换 \mathbf{T}
- 图像平面旋转角度 \mathbf{R}
- 焦距 f , 图像中心点 (c_x, c_y) , 图像分辨率 α
- 蓝色为“外参,” 红色为“内参”

投影公式

$$\mathbf{x} = \begin{bmatrix} sx \\ sy \\ s \end{bmatrix} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathbf{\Pi X}$$



- 投影将是很多变换的叠加
- 以上参数可以把这些变换拆解

$$\mathbf{\Pi} = \begin{bmatrix} f & s & c_x \\ 0 & \alpha f & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R}_{3 \times 3} & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{T}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 0 \end{bmatrix}$$

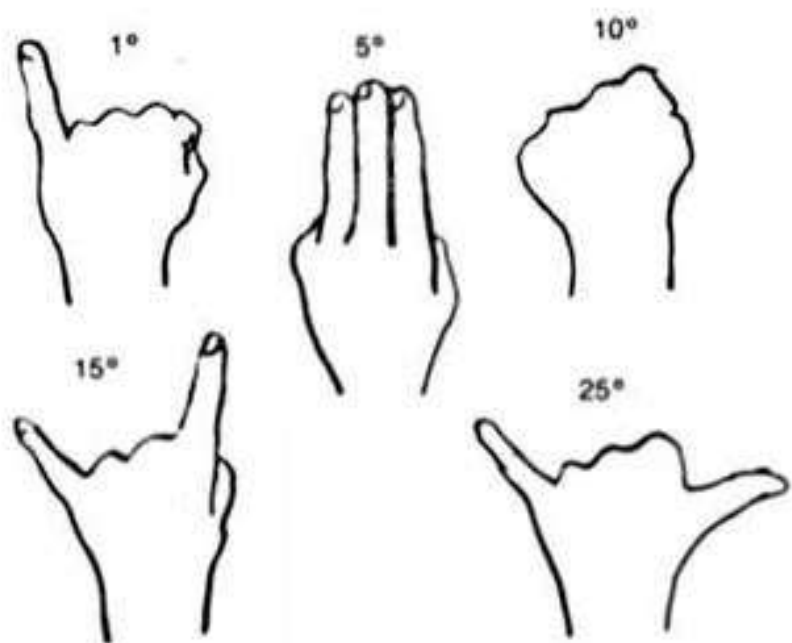
内参 投影 旋转 坐标变换







identity matrix

- 请注意：这里只是简单模型，真实相机很复杂！

视场角与剪裁因子 CheatSheet

APS-C Crop Body Measurement Table



Lens	After 1.62 Multiplier	APS-C Sensor (1.62 lens multiplier) Canon 60D, 7D, 70D, T3i, T4i	Hand Positions
18mm	29.16mm	Three hands wide at full arms length.	
28mm	45.36mm	Slightly less than two hands wide at full arms length.	
35mm	56.7mm	One hand + width of one fist at full arms length.	
50mm	81mm	One hand wide + width of thumb at full arms length.	
55mm	89.1mm	Slightly less than one hand wide at full arms length.	
85mm	137.7mm	Inside edge of thumb to tip of forefinger wide with hand in "L" shape, thumb up.	

不同焦距拍摄效果不同



广角（短焦）



标准镜头

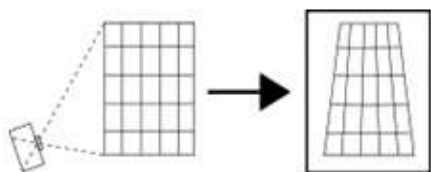


长焦镜头

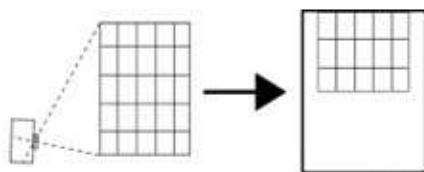
知道为什么
专业摄影师
都用大长焦
了吧！

透视失真

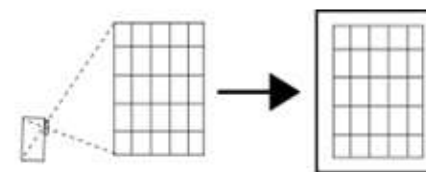
- 怎么办呢？



向上倾斜摄像头
建筑物上下比例失衡

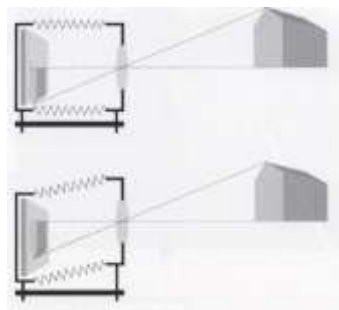


保持照相机水平，则建筑
物上面部分拍不到



把透镜向上移就可以
解决这个问题

- 解决方案: view camera (lens shifted w.r.t. film)



http://en.wikipedia.org/wiki/Perspective_correction_lens

怎么利用透视失真？



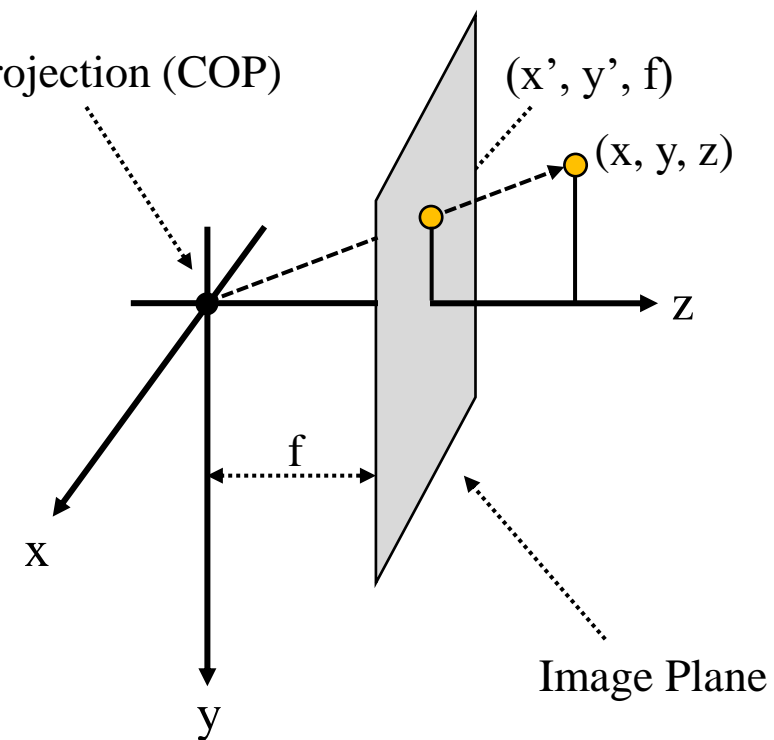
视觉错觉



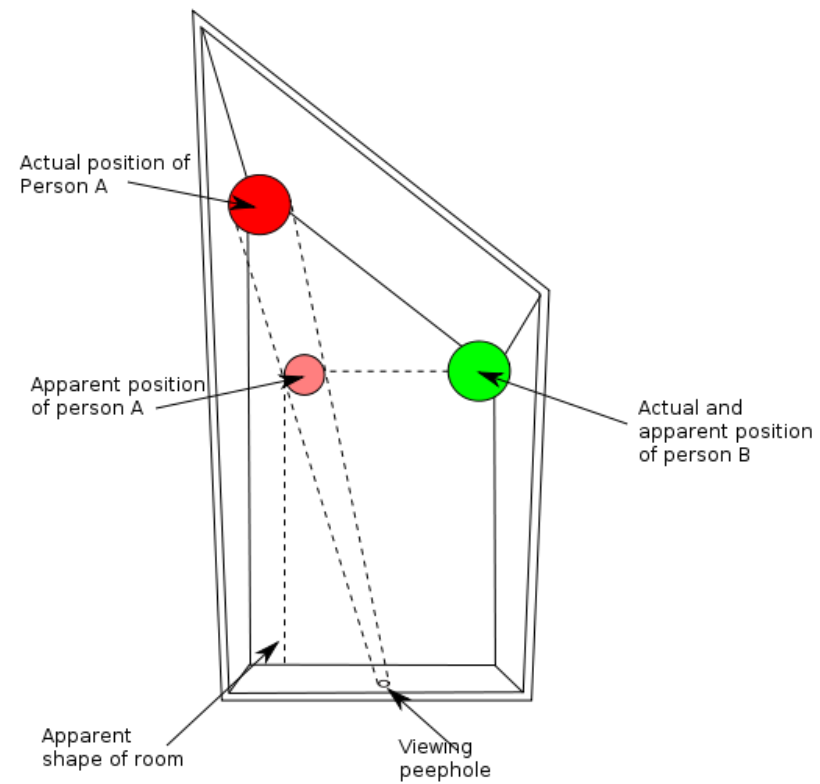
错觉原理



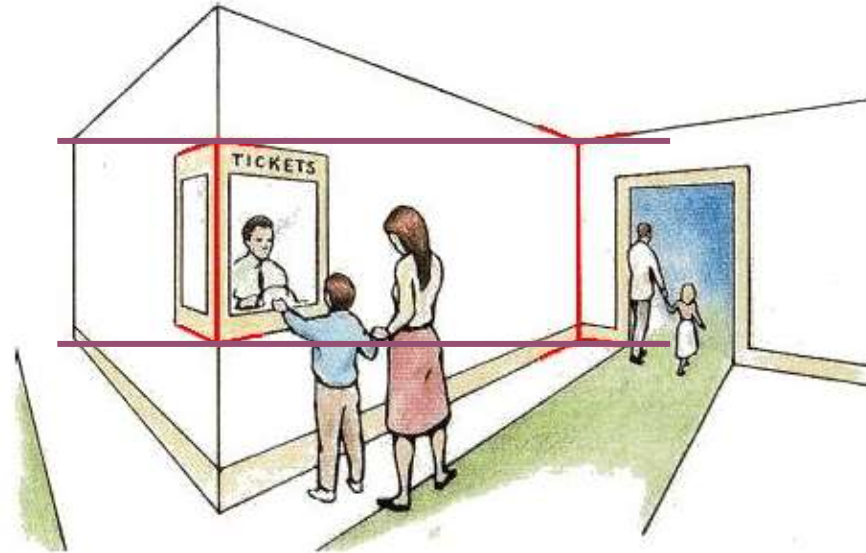
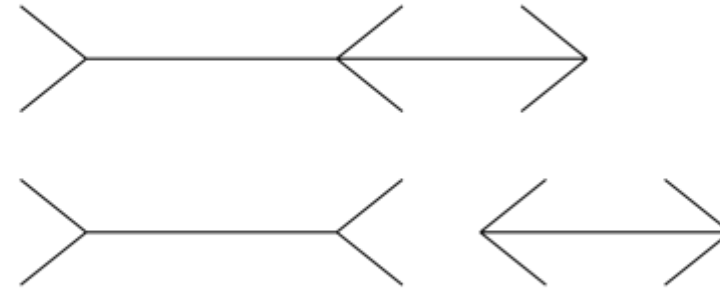
Center of Projection (COP)



Ames Room: 猜猜怎么做到的？



Müller-Lyer Illusion



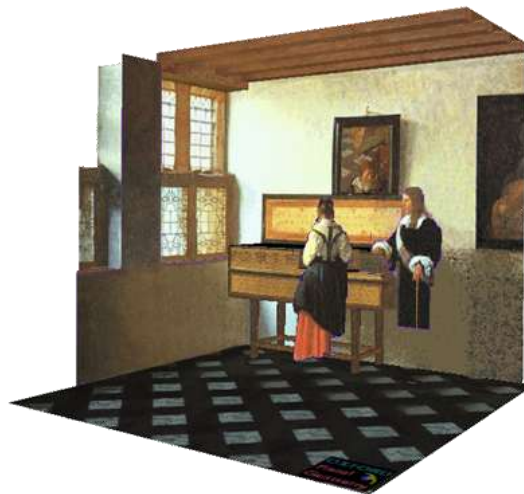
https://en.wikipedia.org/wiki/Müller-Lyer_illusion

单视图与立体视图

透视几何的应用



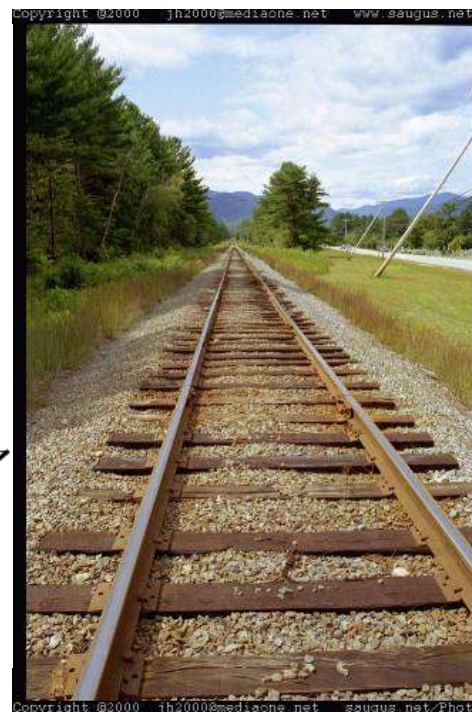
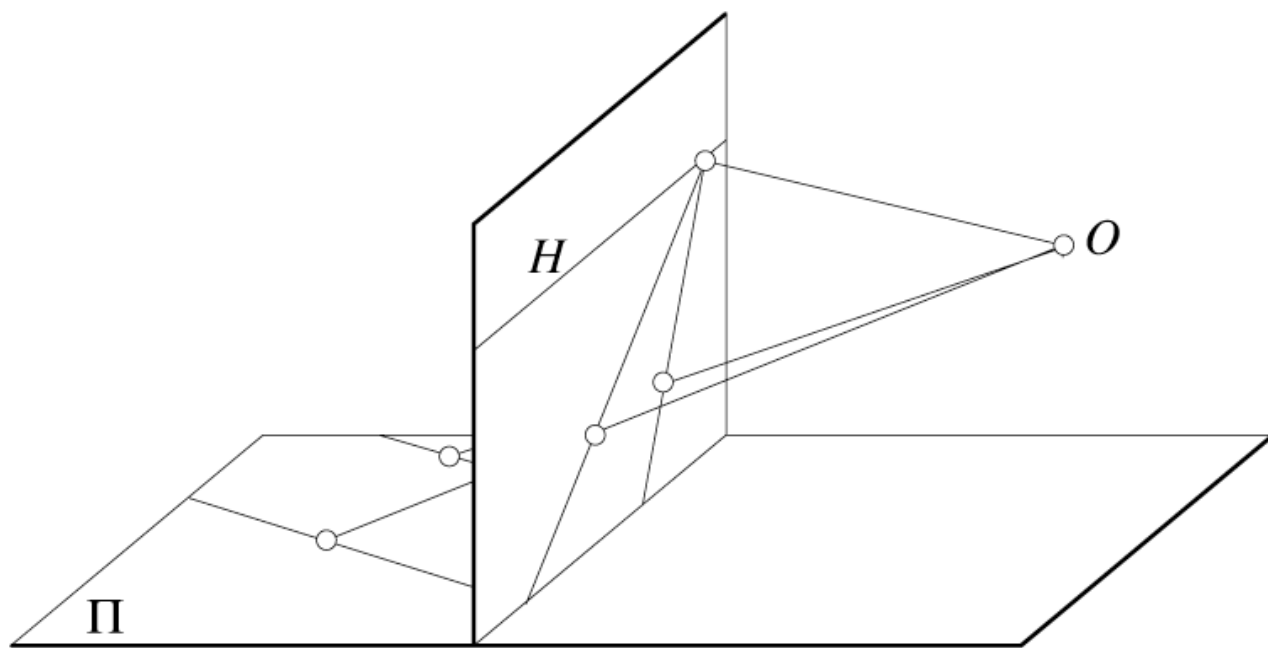
Vermeer's *Music Lesson*



Reconstructions by Criminisi et al.

投影的性质

- 两条平行线也有相交的那一天：消失点 vanishing point



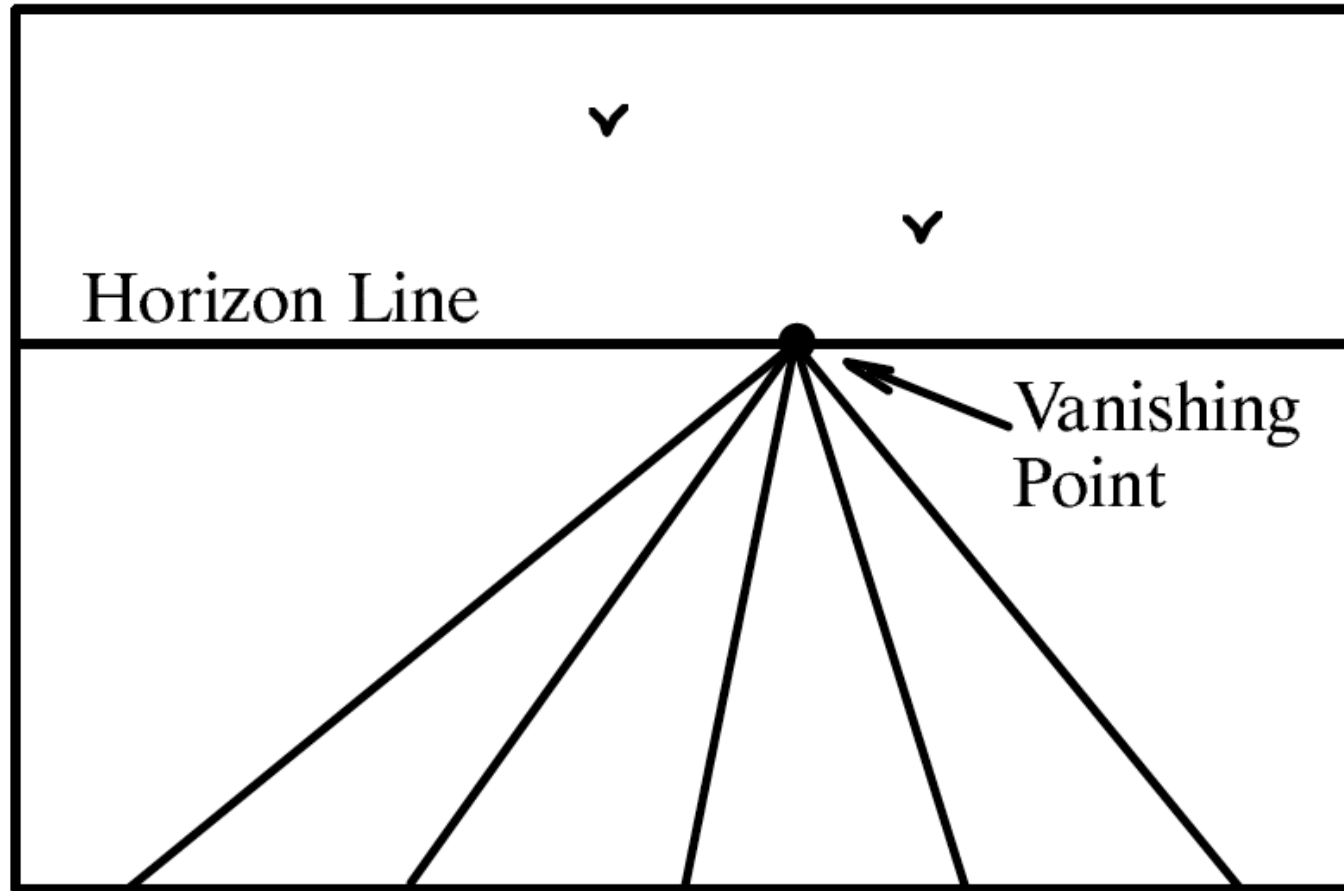
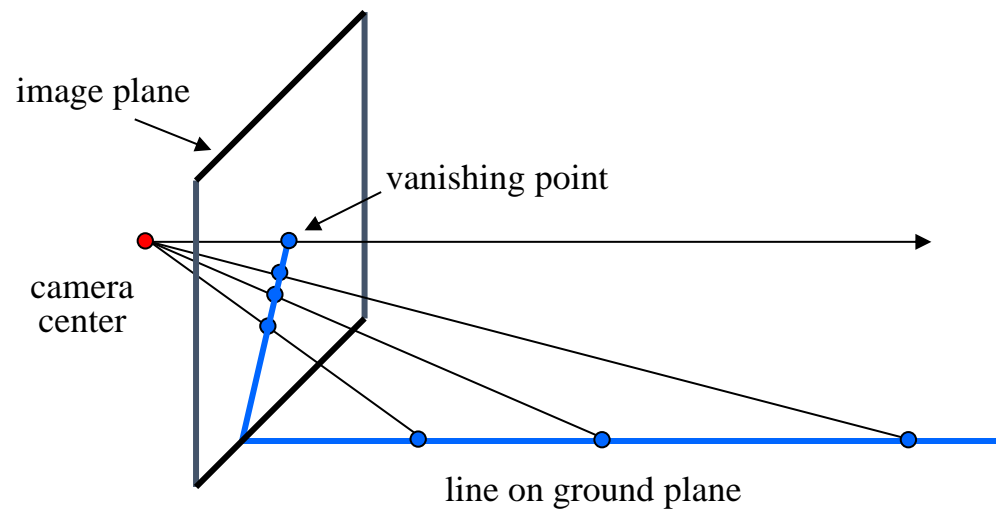


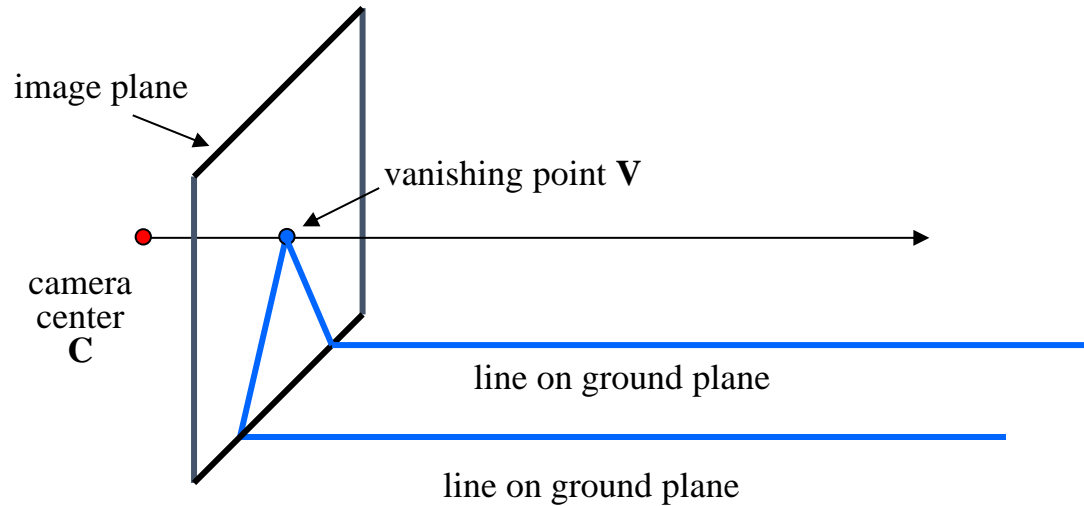
Figure 23.4

A perspective view of a set of parallel lines in the plane. All of the lines converge to a single vanishing point.

消失点: 无穷远的投影



消失点: 地平面与平行线

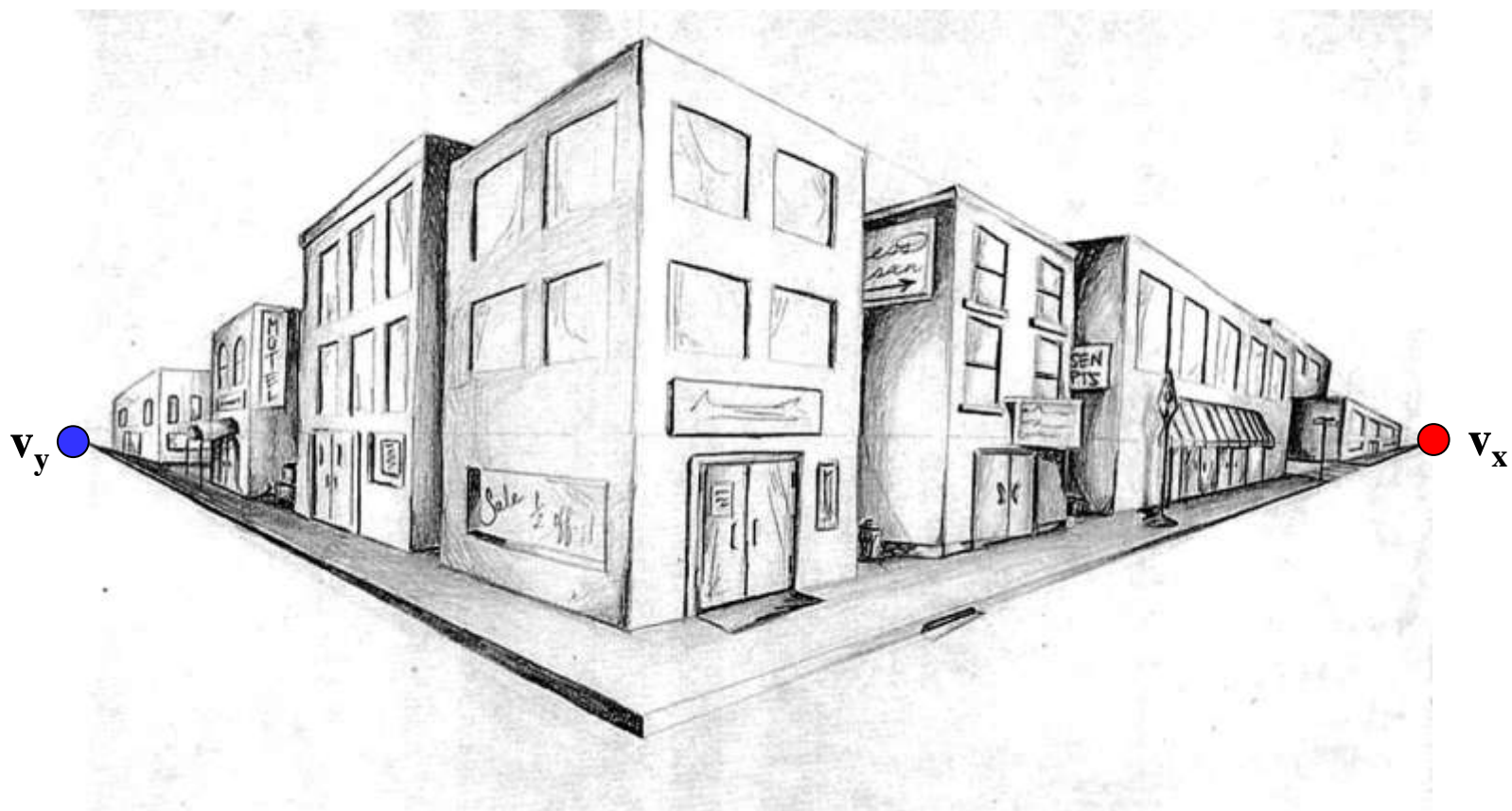


- 性质
 - 地平面中的一组平行线有同一个消失点
 - C 到 v 的射线也与这些线平行
 - 一张图像可能有多于一个消失点
 - 图像上的所有点都有可能是消失点

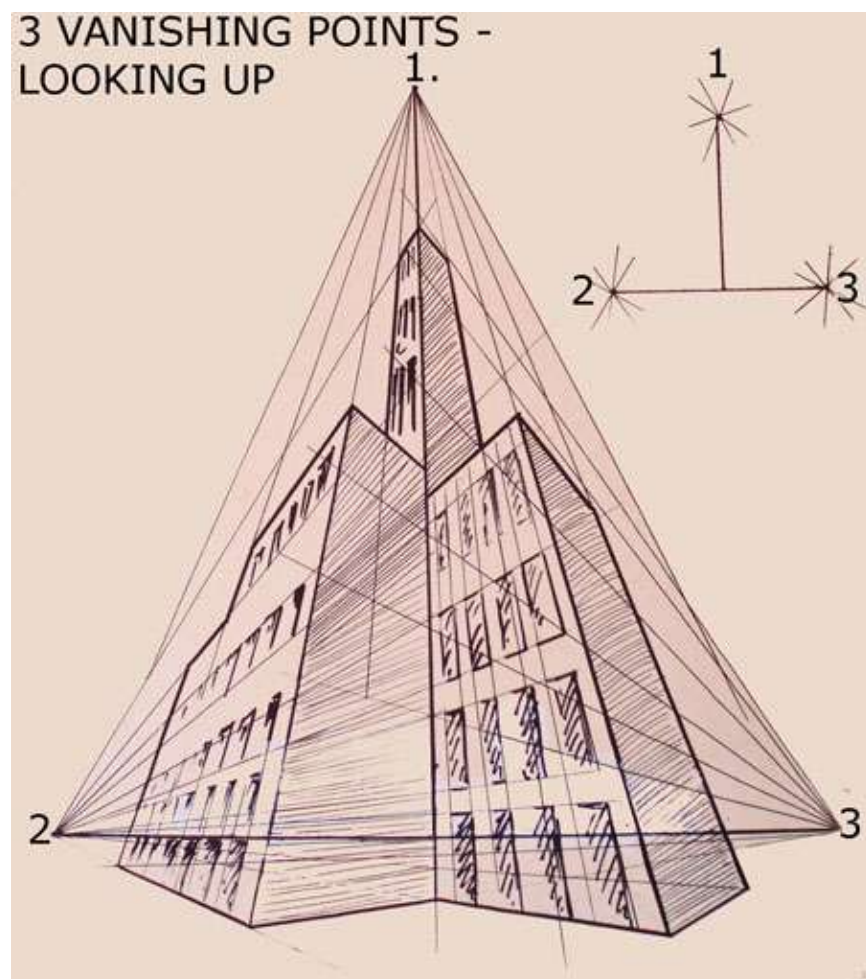
一个消失点



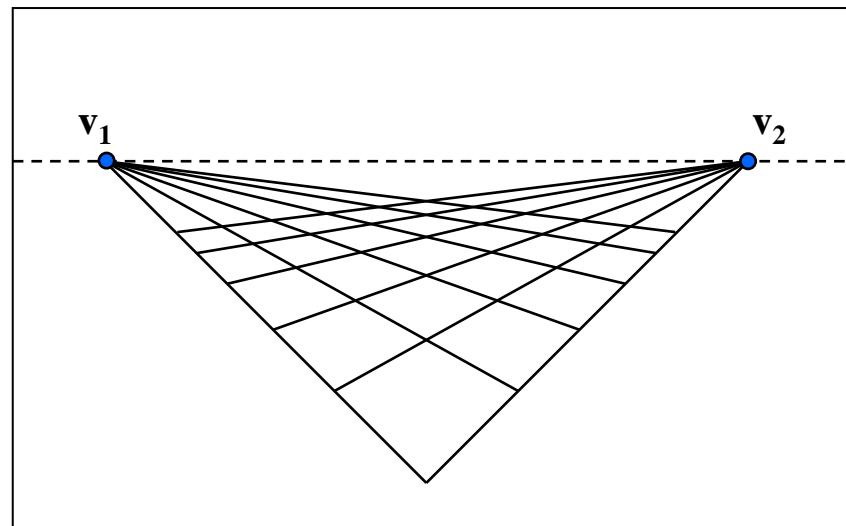
几个消失点？



三个消失点

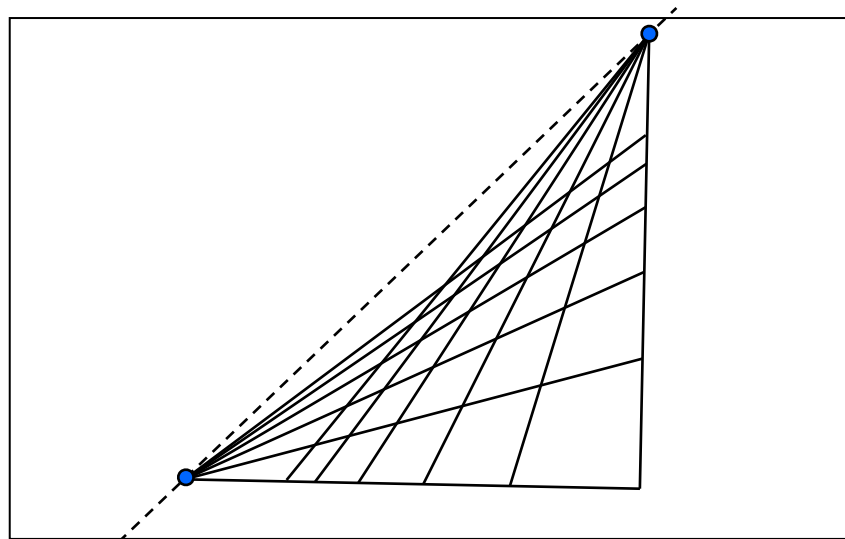


消失线



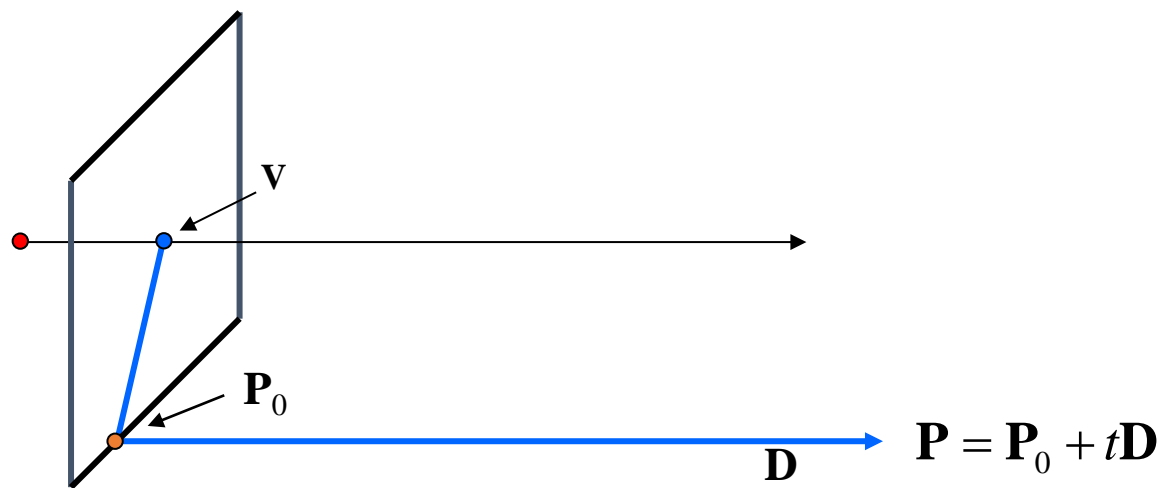
- 消失点的集合
 - 平行线的集合定义一个消失点
 - 所有这些消失点构成一个水平线
 - 或者叫消失线
 - 不同的平面会对应不同的消失线

消失线

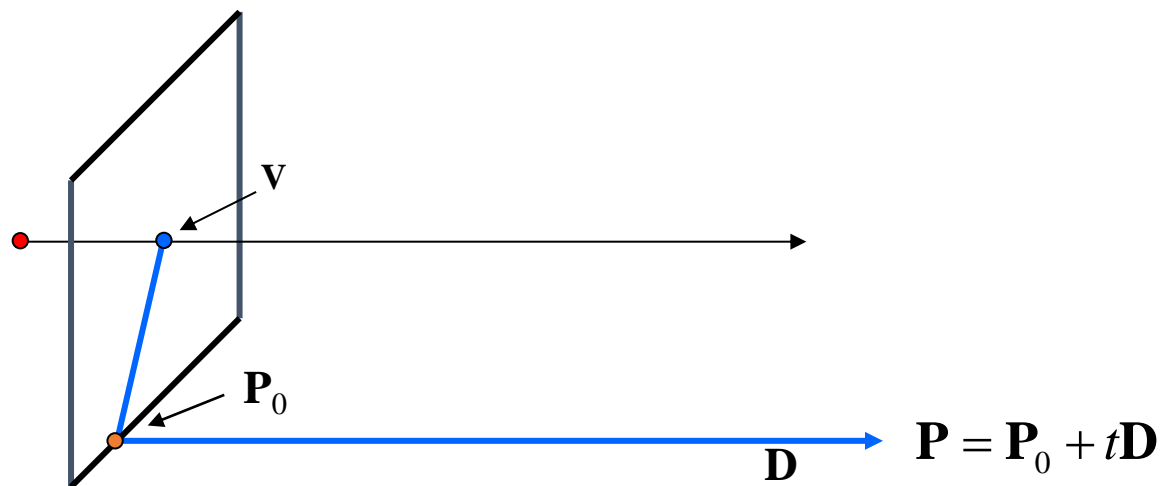


- 消失点的集合
 - 平行线的集合定义一个消失点
 - 所有这些消失点构成一个水平线
 - 或者叫消失线
 - 不同的平面会对应不同的消失线

计算消失点



计算消失点



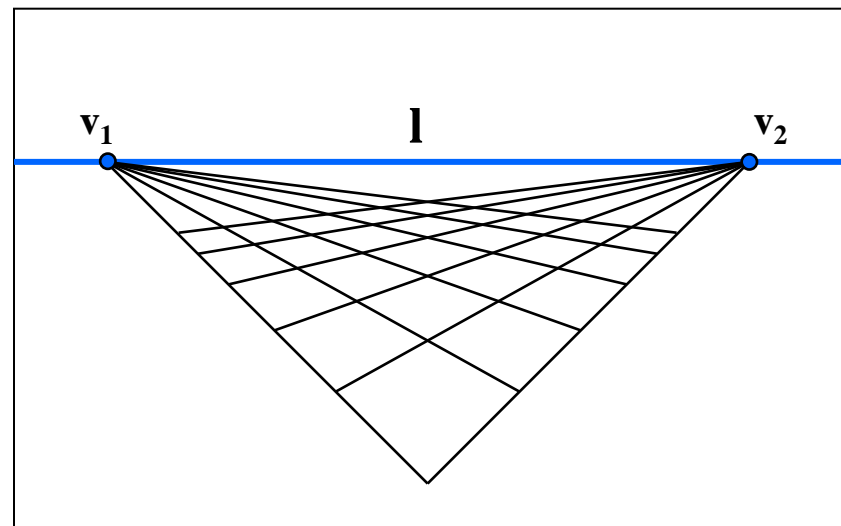
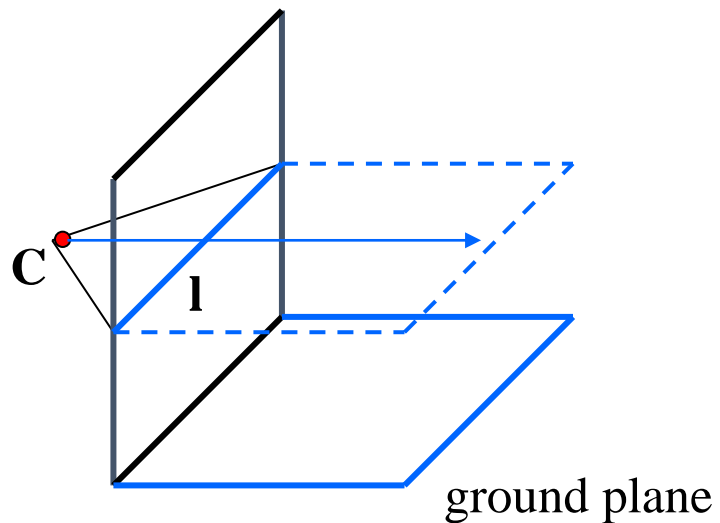
$$\mathbf{P}_t = \begin{bmatrix} P_x + tD_x \\ P_y + tD_y \\ P_z + tD_z \\ 1 \end{bmatrix} \cong \begin{bmatrix} P_x / t + D_x \\ P_y / t + D_y \\ P_z / t + D_z \\ 1/t \end{bmatrix}$$

• 性质

$$\mathbf{v} = \mathbf{I}\mathbf{P}_\infty$$

- \mathbf{P}_∞ 是在无穷远的点, \mathbf{v} 是它的投影
- 仅仅与射线方向相关
- 平行线 $\mathbf{P}_0 + t\mathbf{D}$, $\mathbf{P}_1 + t\mathbf{D}$ 投影相交在 \mathbf{P}_∞

计算消失线



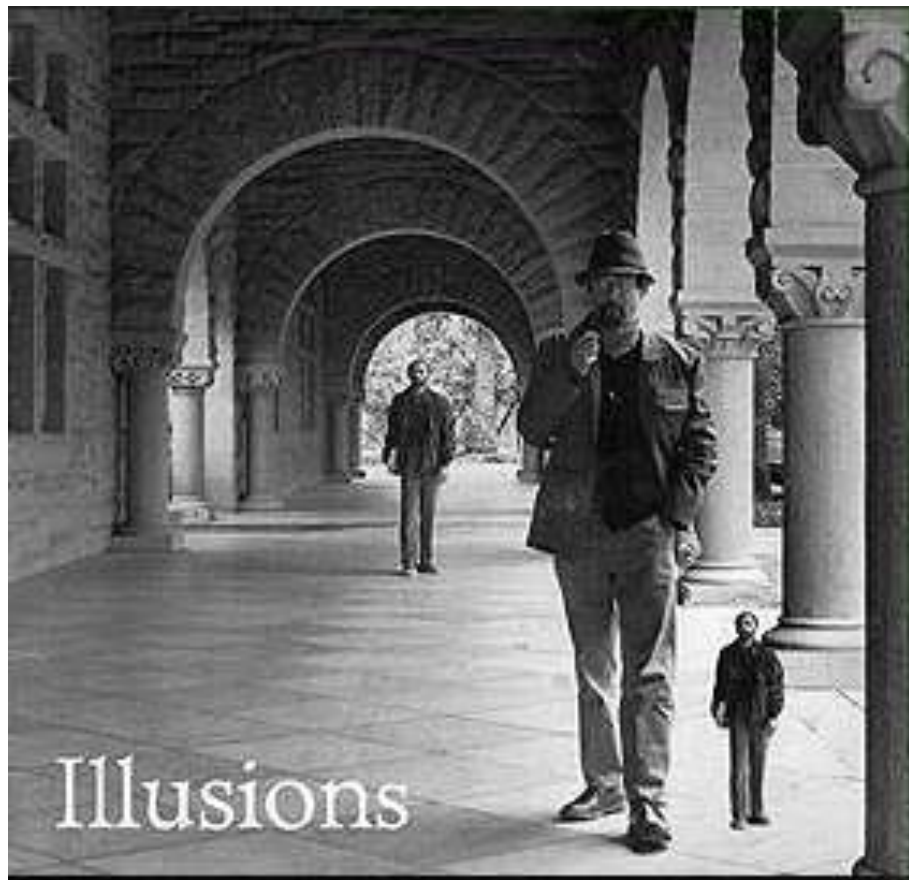
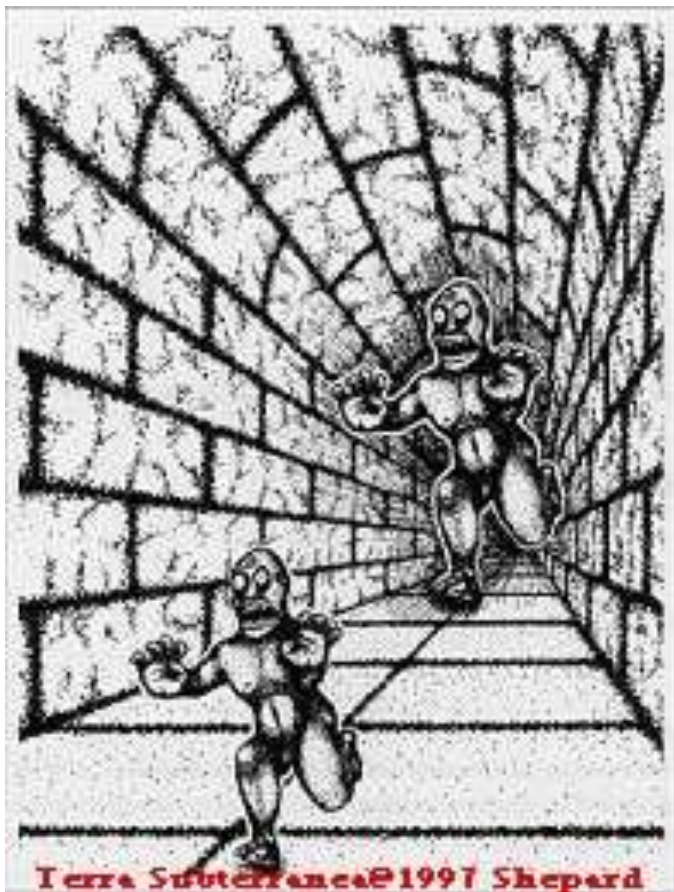
• 性质

- **l** 是图像平面与过C的水平面的相交线
- 可以通过地平面的两组平行线计算**l**
- 所有与**C** 等高的点都会投影到**l**
 - 高于**C** 的点会投影在**l**的上方
- 为计算物体高度提供了方法!

跳伞者在相机之上还是之下？



错觉又来了



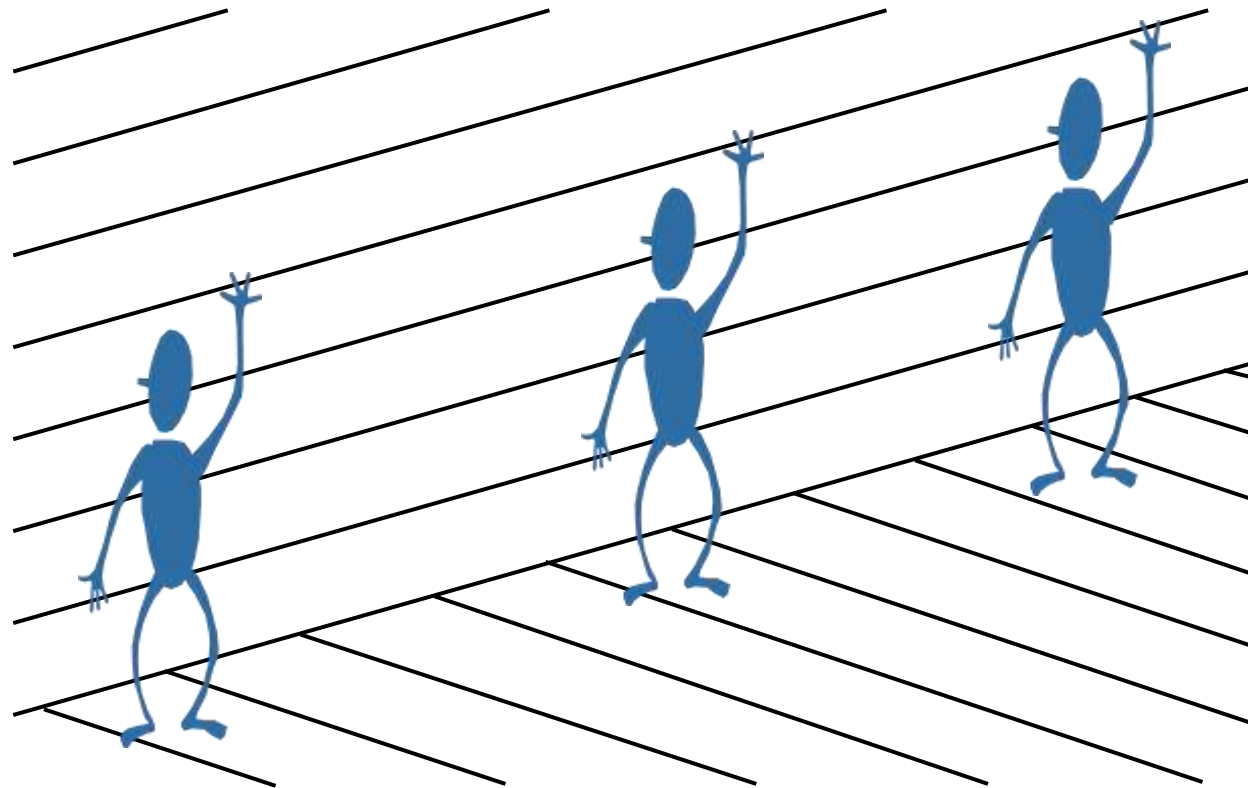
错觉又来了



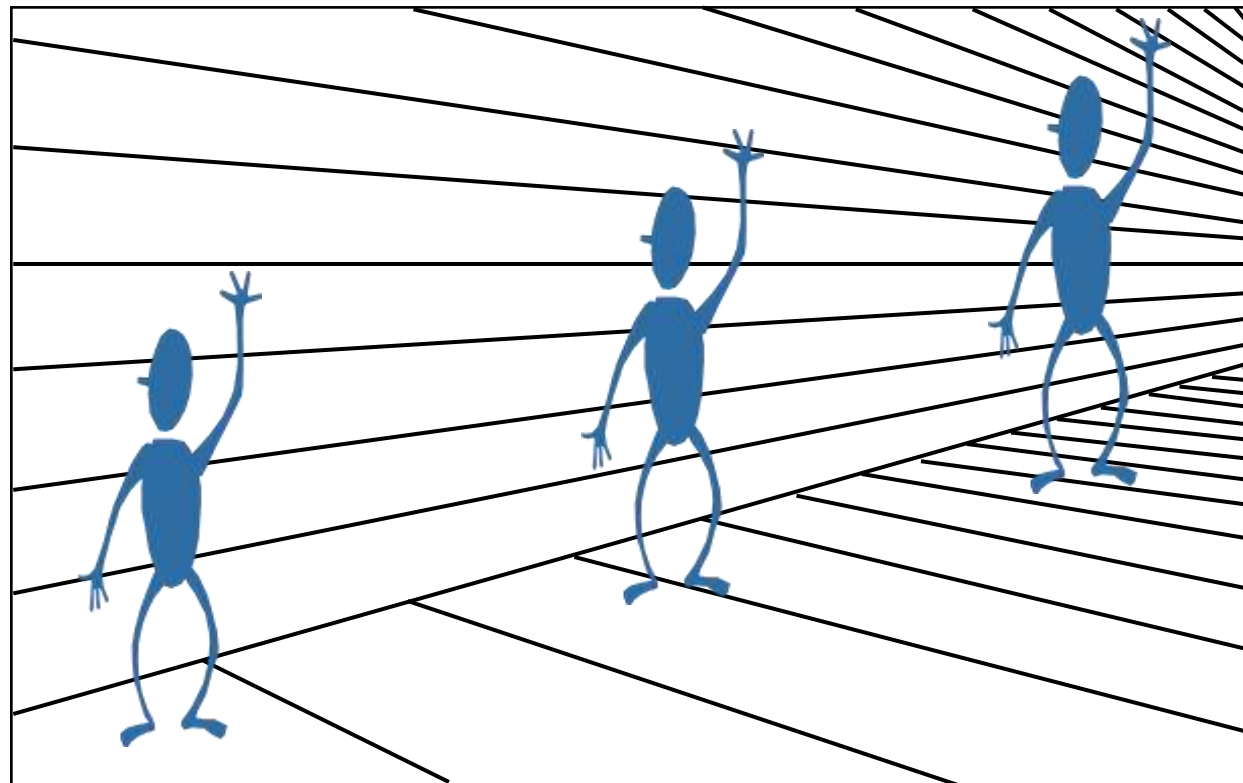
错觉又来了



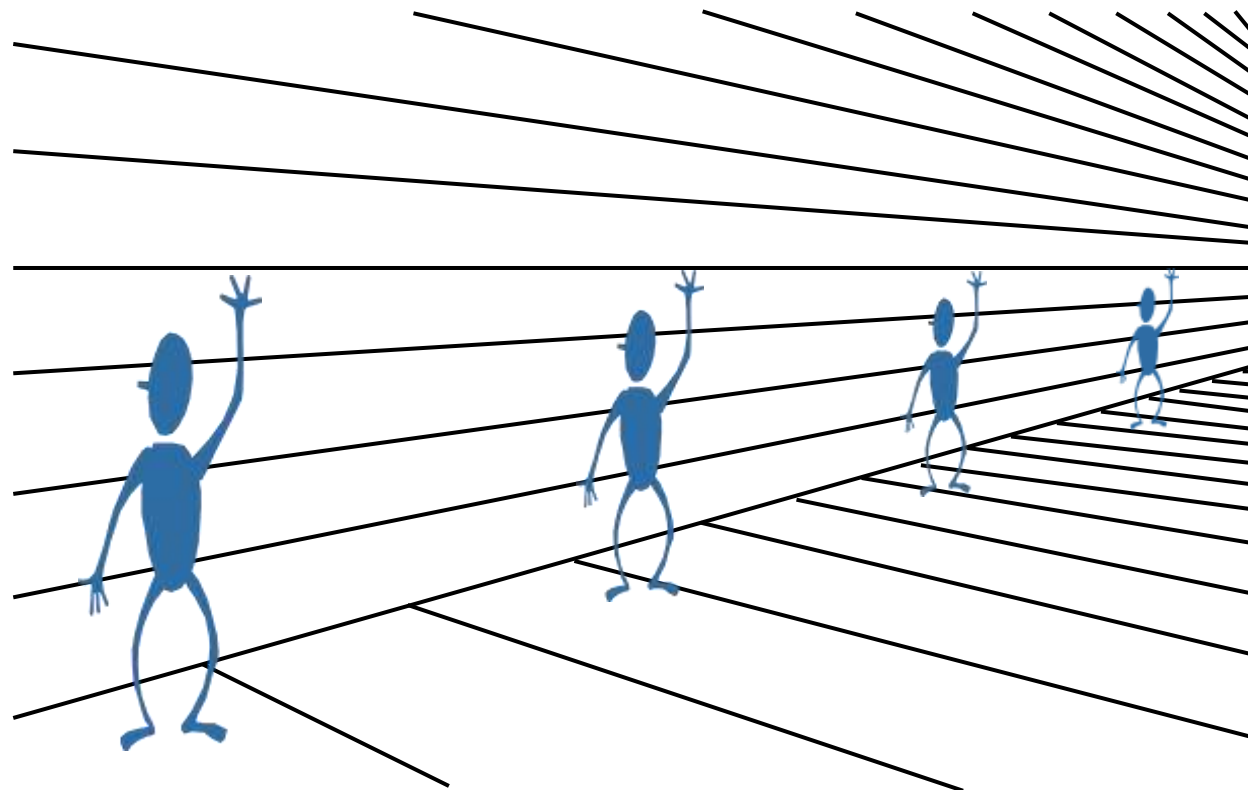
Perspective Cues: 透视线索



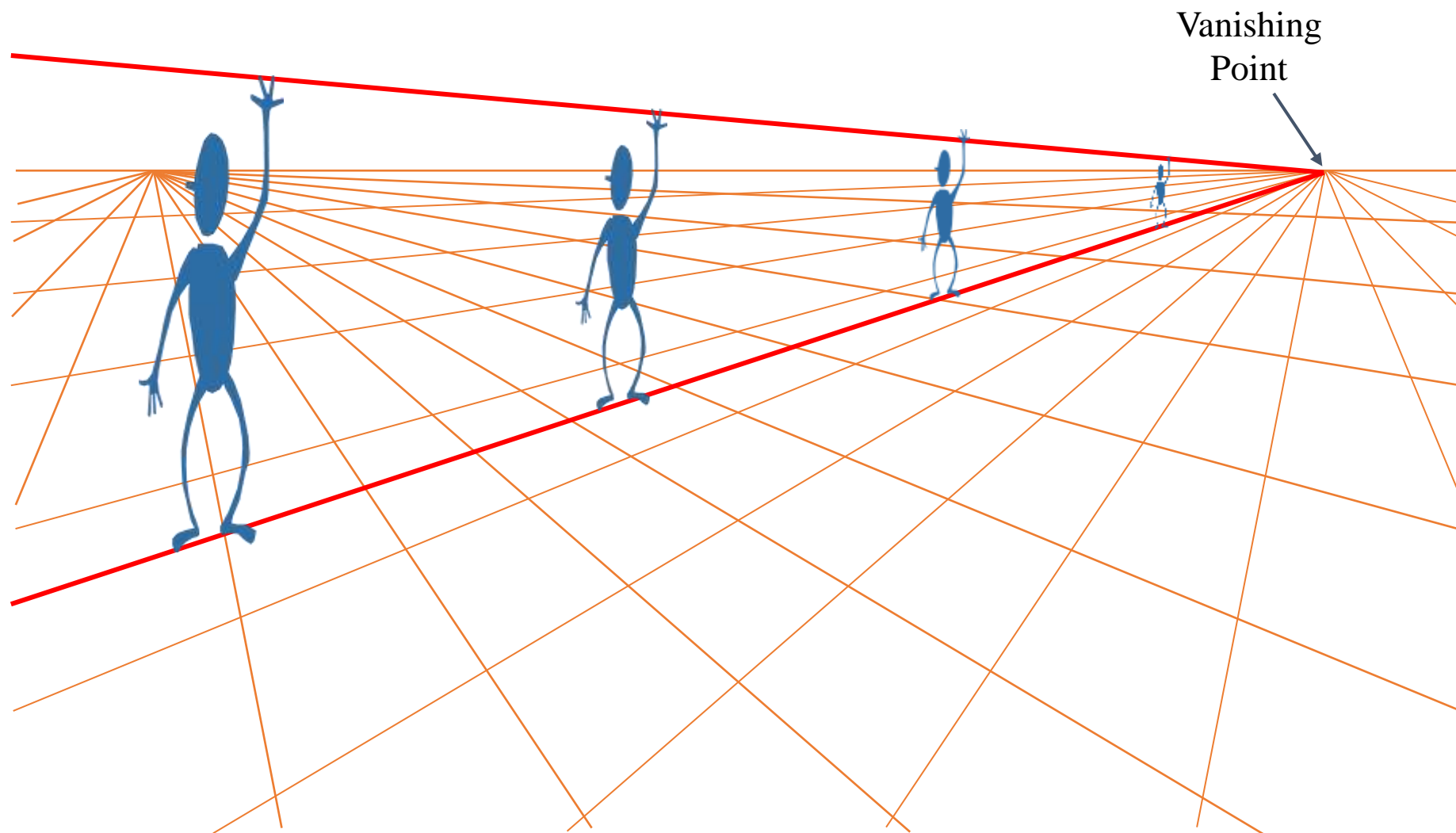
Perspective Cues: 透视线索



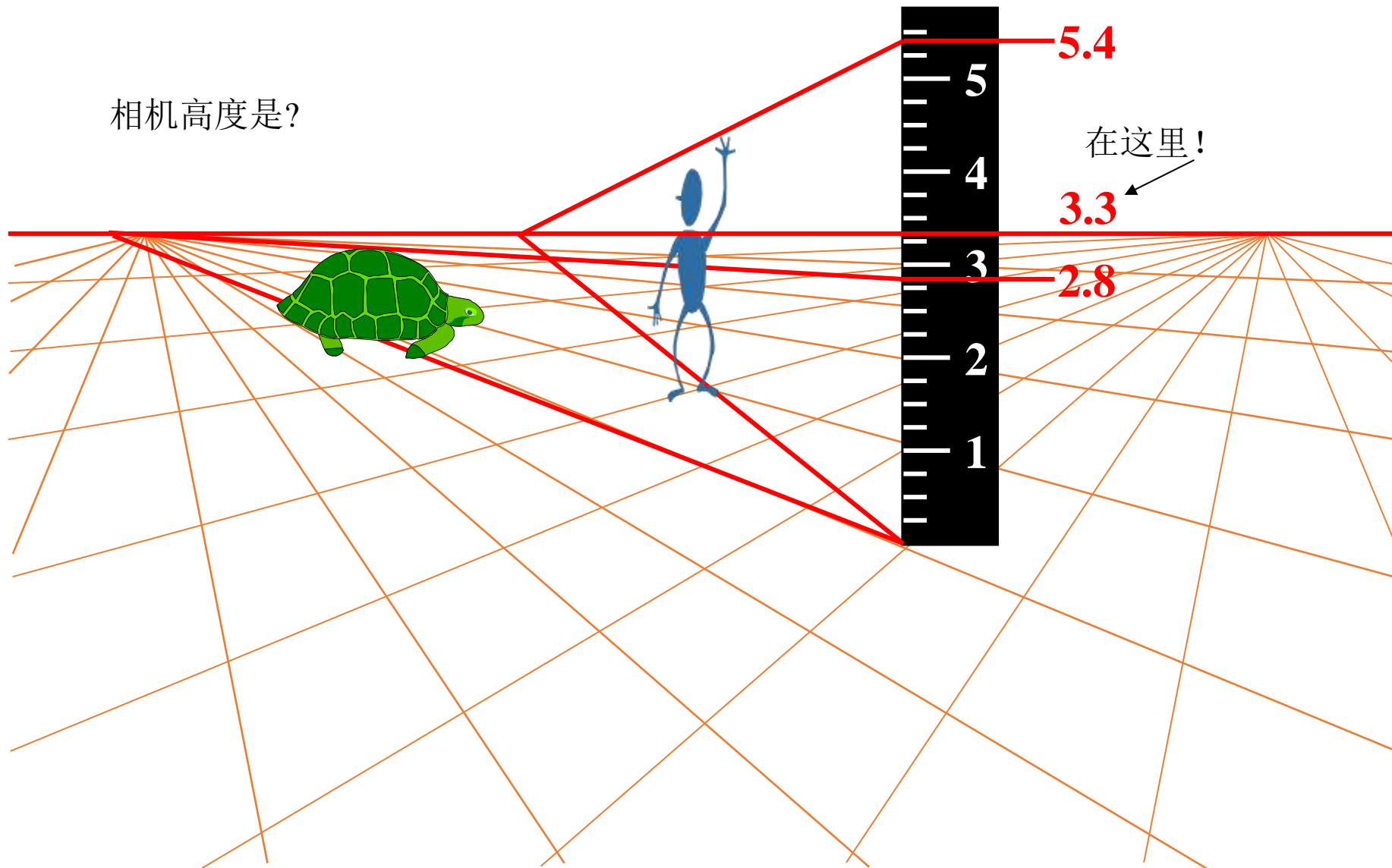
Perspective Cues: 透视线索



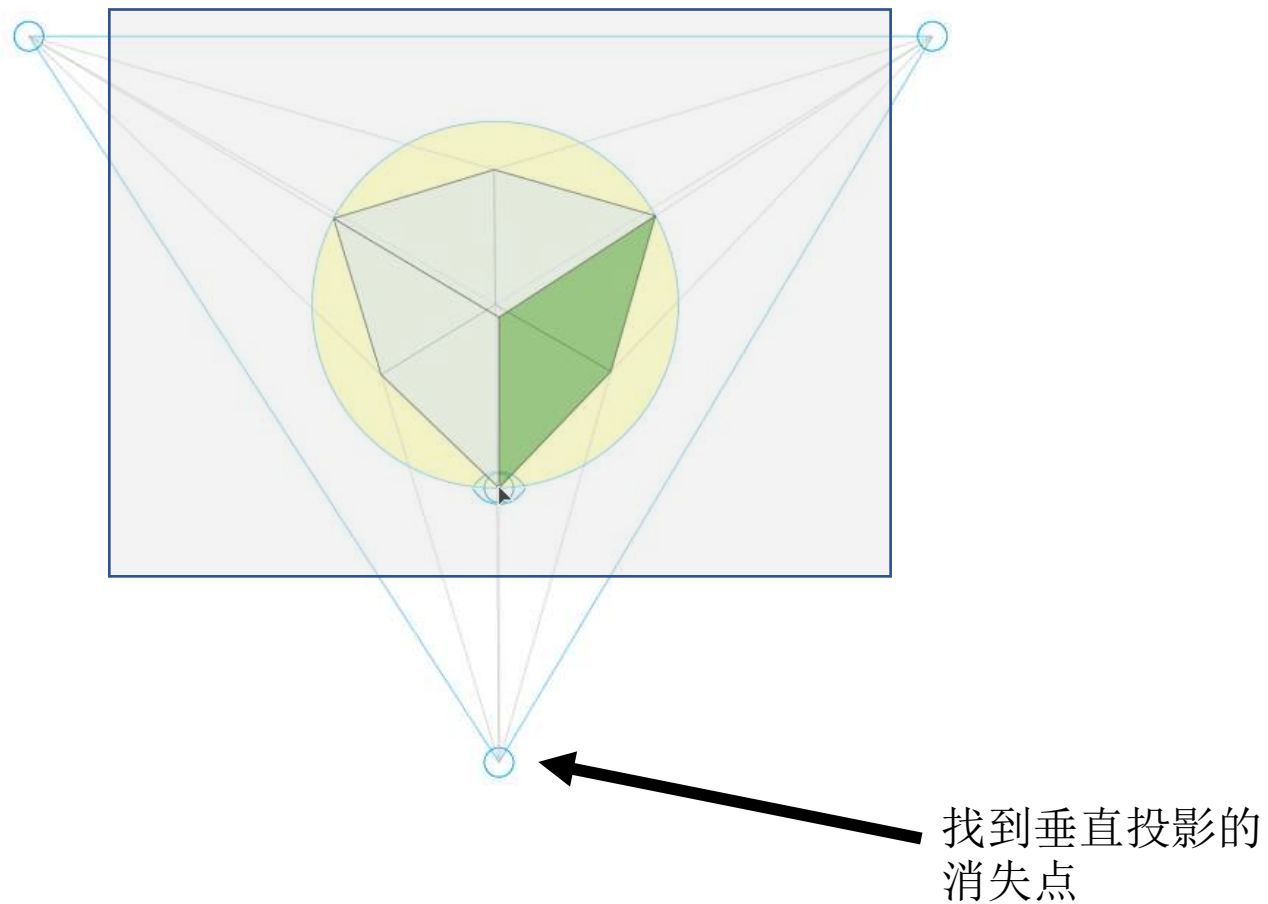
对比高度



计算高度

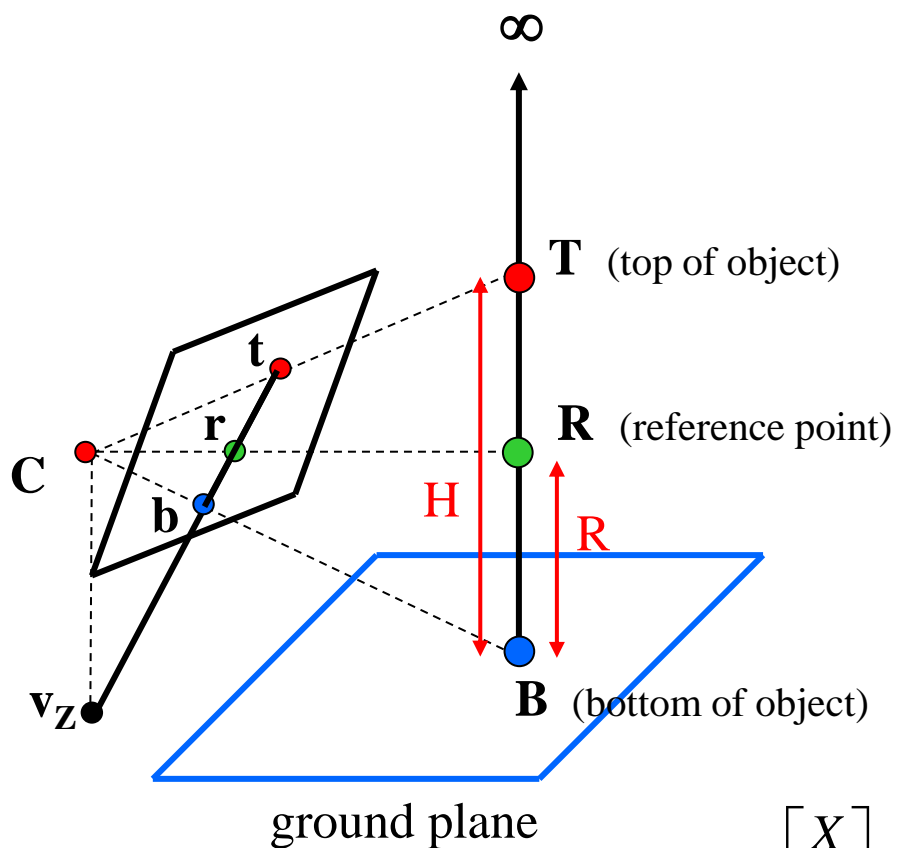


垂直消失点



计算高度

- 线段长度比例投影以后不变



$$\frac{\|\mathbf{T} - \mathbf{B}\| \|\infty - \mathbf{R}\|}{\|\mathbf{R} - \mathbf{B}\| \|\infty - \mathbf{T}\|} = \frac{H}{R}$$

scene cross ratio

$$\frac{\|\mathbf{t} - \mathbf{b}\| \|\mathbf{v}_Z - \mathbf{r}\|}{\|\mathbf{r} - \mathbf{b}\| \|\mathbf{v}_Z - \mathbf{t}\|} = \frac{H}{R}$$

image cross ratio

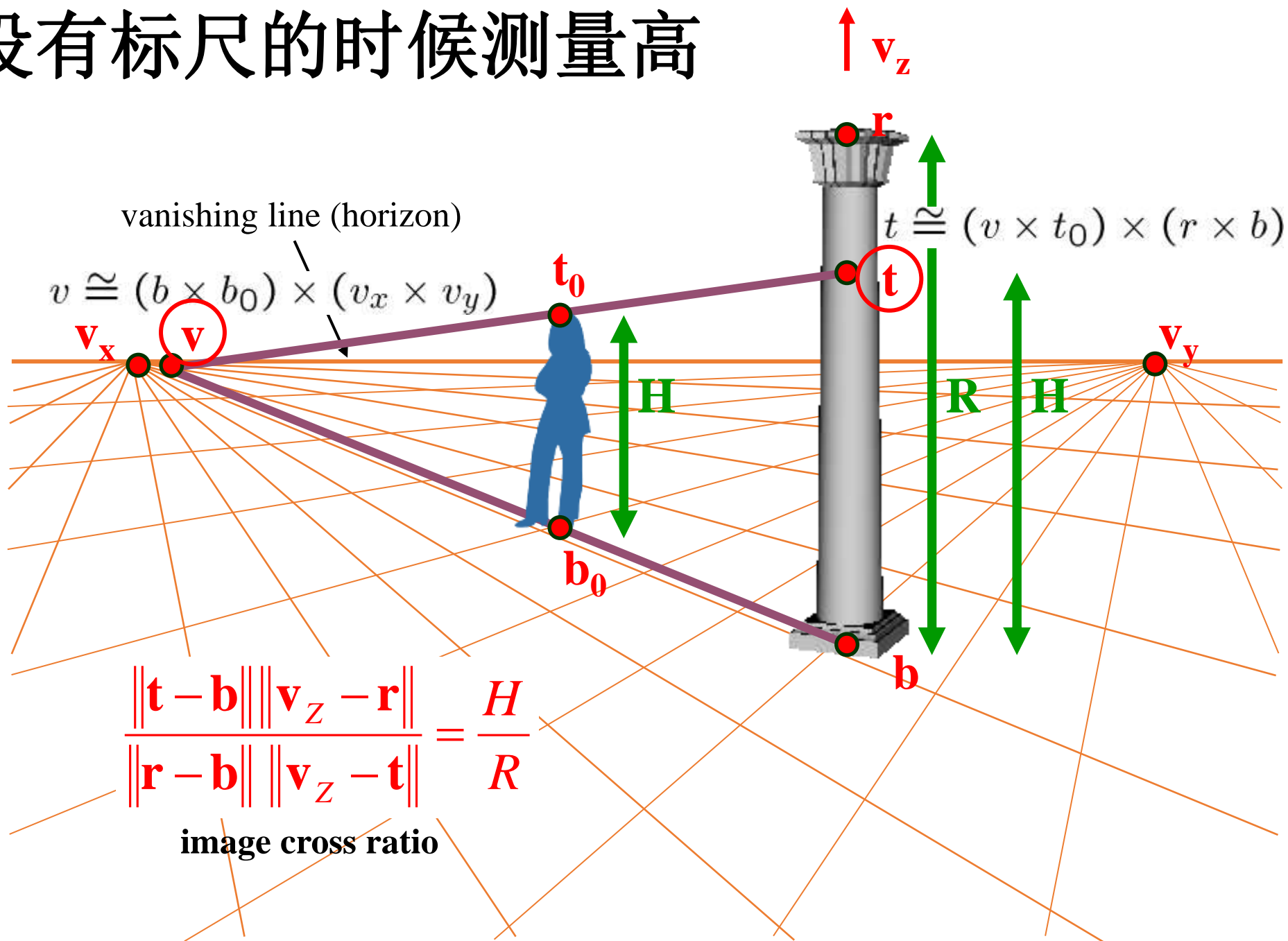
scene points represented as

$$\mathbf{P} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

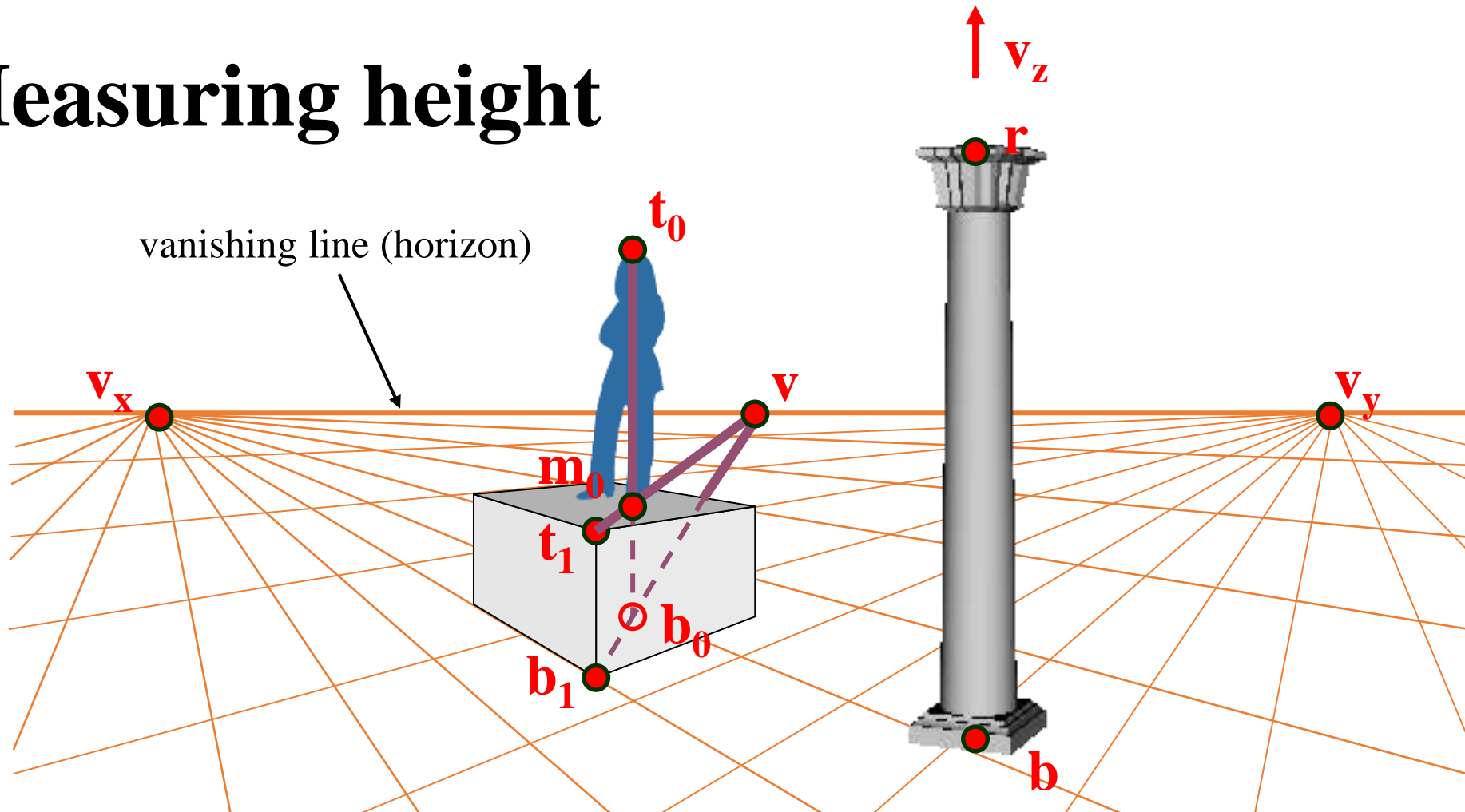
image points as

$$\mathbf{p} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

在没有标尺的时候测量高度



Measuring height



What if the point on the ground plane b_0 is not known?

- Here the person is standing on the box, height of box is known
- Use one side of the box to help find b_0 as shown above

单张图像的三维重建



St. Jerome in his Study, H. Steenwick

Bringing Pictorial Space to Life: Computer Techniques for the Analysis of Paintings.
Antonio Criminisi, Martin Kemp, Andrew Zisserman. 2002.

单张图像的三维重建

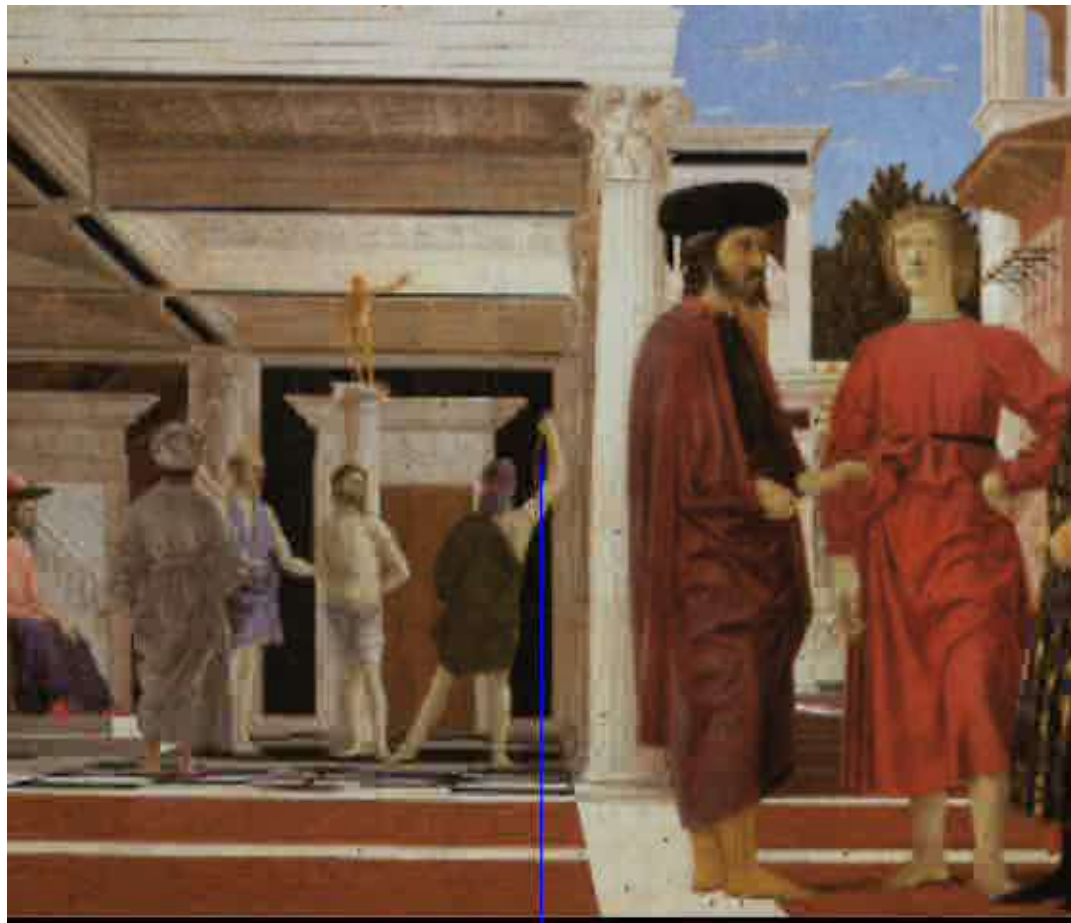


单张图像的三维重建



Flagellation, Piero della Francesca

单张图像的三维重建



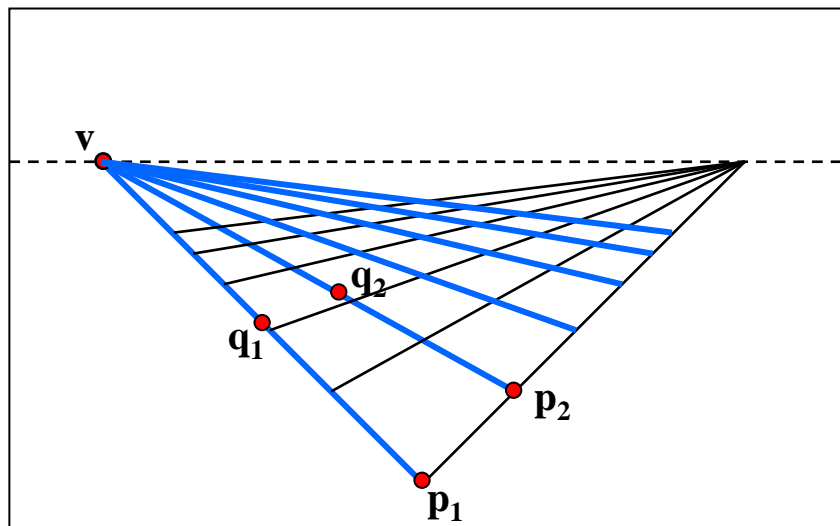
video by Antonio Criminisi

单张图像的三维重建



Flagellation. Piero della Francesca. c1453.

从线段计算消失点



Intersect p_1q_1 with p_2q_2

$$v = (p_1 \times q_1) \times (p_2 \times q_2)$$

怎么找两条平行的线段？怎么知道线段的长度？

相机校准

- 目标：估计相机参数
 - Version 1: 解一下方程

$$\mathbf{X} = \begin{bmatrix} wx \\ wy \\ w \end{bmatrix} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathbf{M}\mathbf{X}$$

- Version 2: 分别计算相机参数
 - 内参 (焦距, 主点, 分辨率)
 - 外参 (转角, 位置)

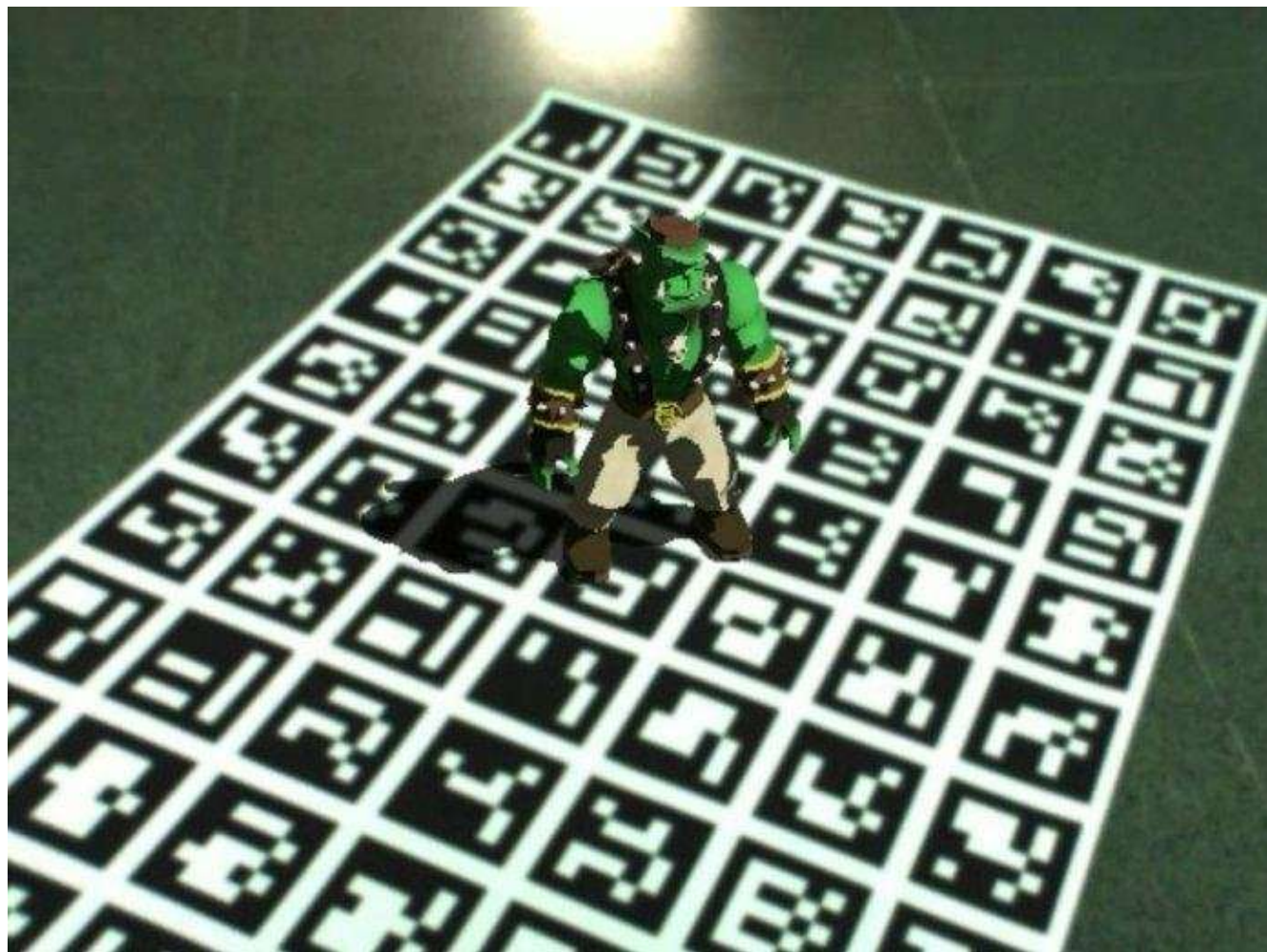
相机校准

$$\mathbf{\Pi} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} = [\boldsymbol{\pi}_1 \quad \boldsymbol{\pi}_2 \quad \boldsymbol{\pi}_3 \quad \boldsymbol{\pi}_4]$$

- $\boldsymbol{\pi}_1 = \mathbf{\Pi}[1 \ 0 \ 0 \ 0]^T = \mathbf{v}_x$ (X vanishing point)
- similarly, $\boldsymbol{\pi}_2 = \mathbf{v}_y$, $\boldsymbol{\pi}_3 = \mathbf{v}_z$
- $\boldsymbol{\pi}_4 = \mathbf{\Pi}[0 \ 0 \ 0 \ 1]^T =$ projection of world origin

$$\mathbf{\Pi} = [\mathbf{v}_x \quad \mathbf{v}_y \quad \mathbf{v}_z \quad \mathbf{o}]$$

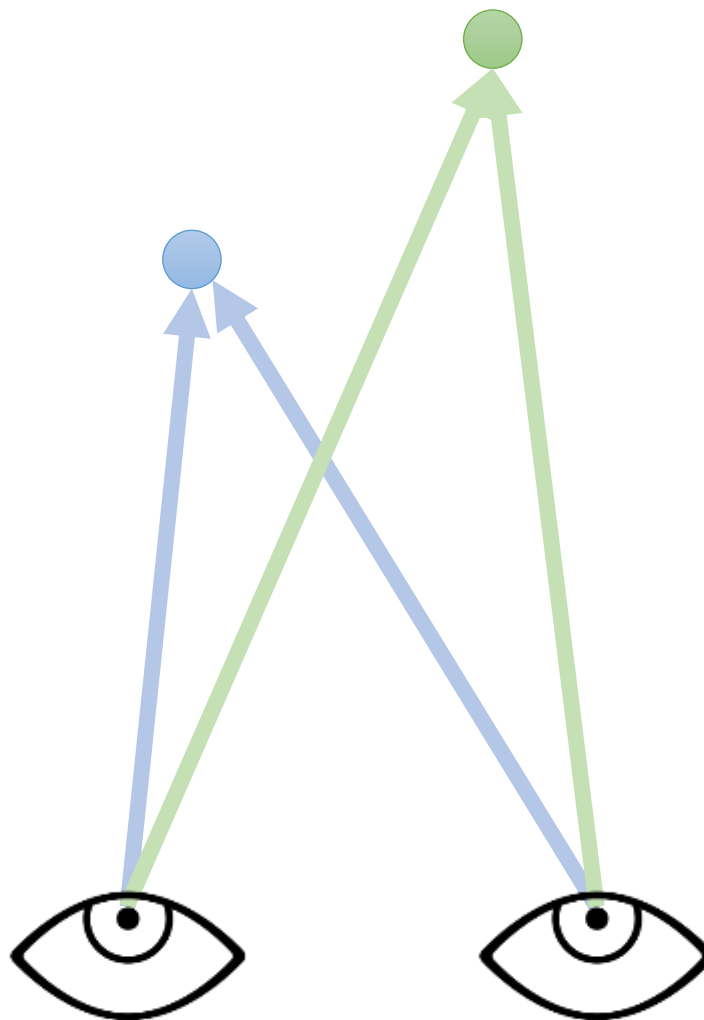
AR 增强现实



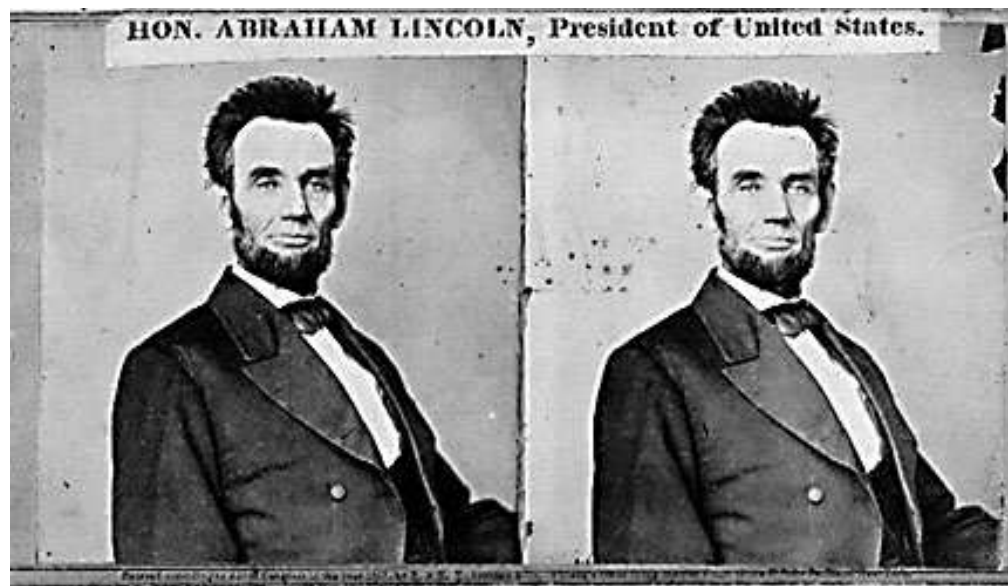
ArUco

双目视图：测距

- 在知道相机参数的情况下，若不借助额外工具，怎样测距？
- 使用双目照相机！

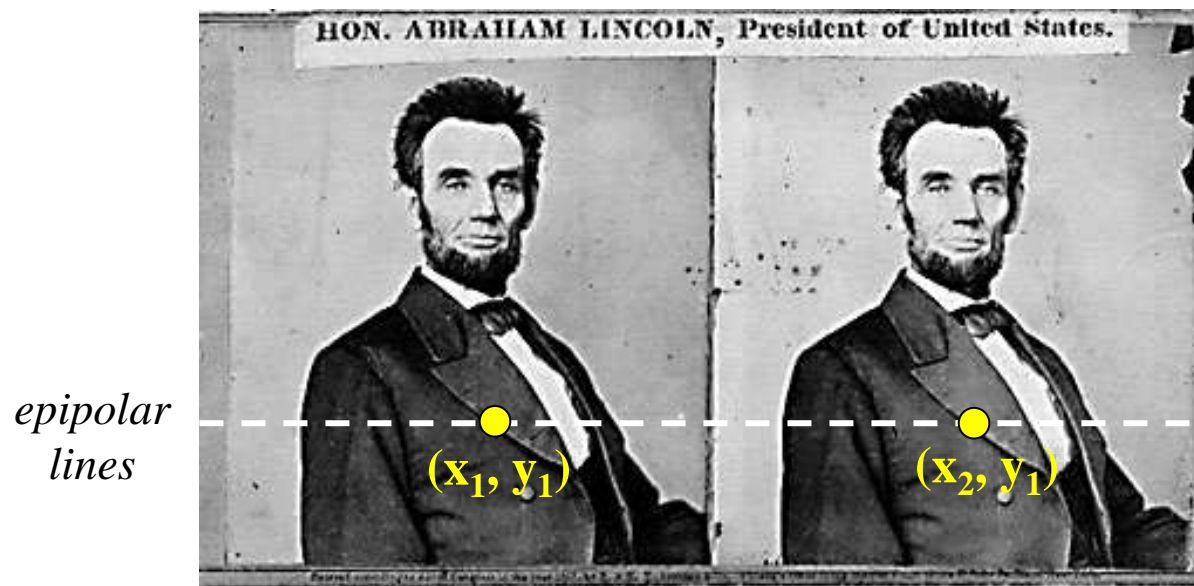


双目视图



- 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！

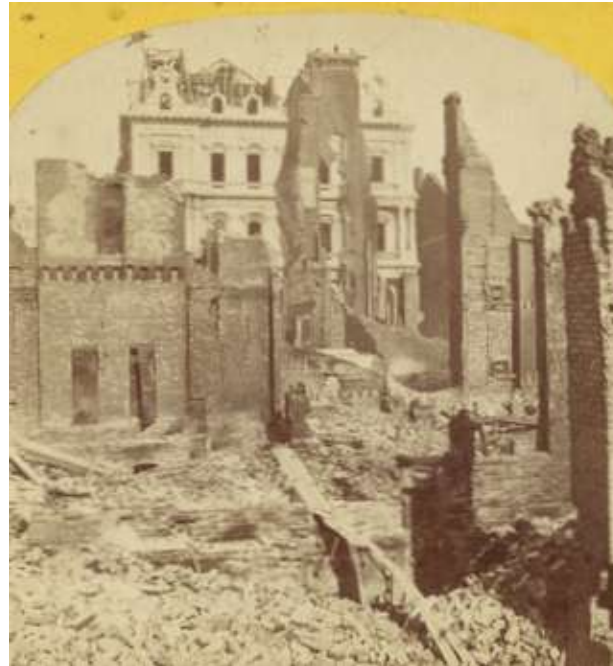
Epipolar Geometry: 极线几何



极线：在两个相机图像之间的几何关系。
每个像素在一个图像上都有对应的一条极线在另一个图像上。

$$\mathbf{x}_2 - \mathbf{x}_1 = \text{the } \textit{disparity} \text{ of pixel } (\mathbf{x}_1, y_1)$$

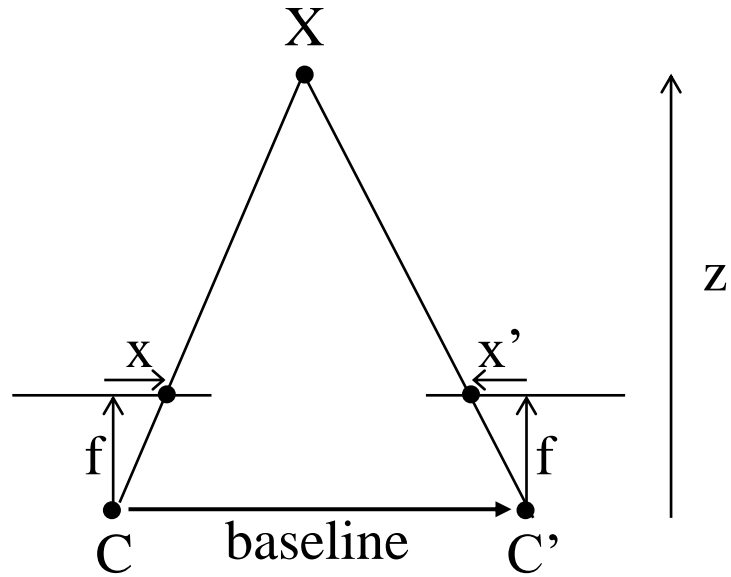
Disparity = inverse depth



<http://stereo.nypl.org/view/41729>

(Or, hold a finger in front of your face and wink each eye in succession.)

Depth from disparity: 通过差异计算景深



$$disparity = x - x' = \frac{baseline * f}{z}$$

实时双目设备



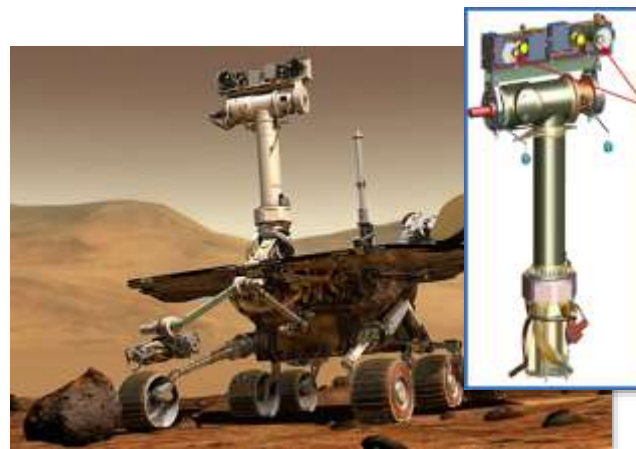
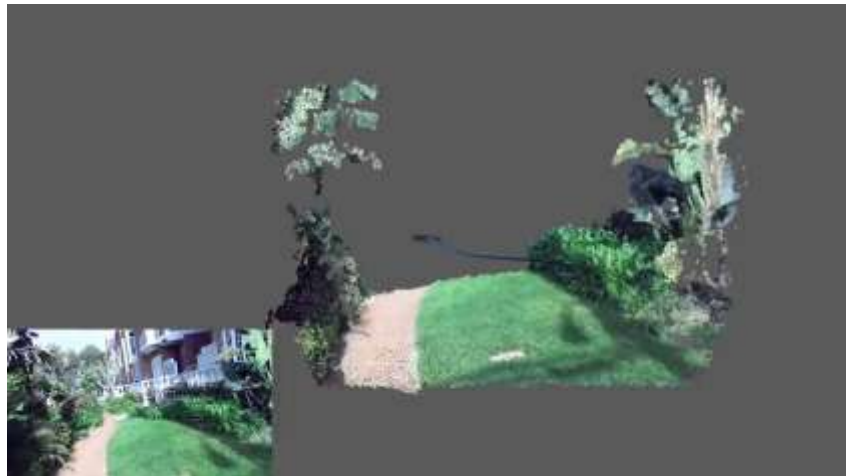
[Nomad robot](#) searches for meteorites in Antarctica

- 机器人巡航

目标：对世界进行三维建模



ZED 2i
Camera





原始镜头效果



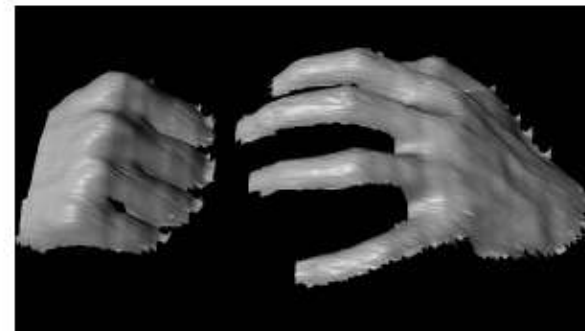
校正
后

结构光

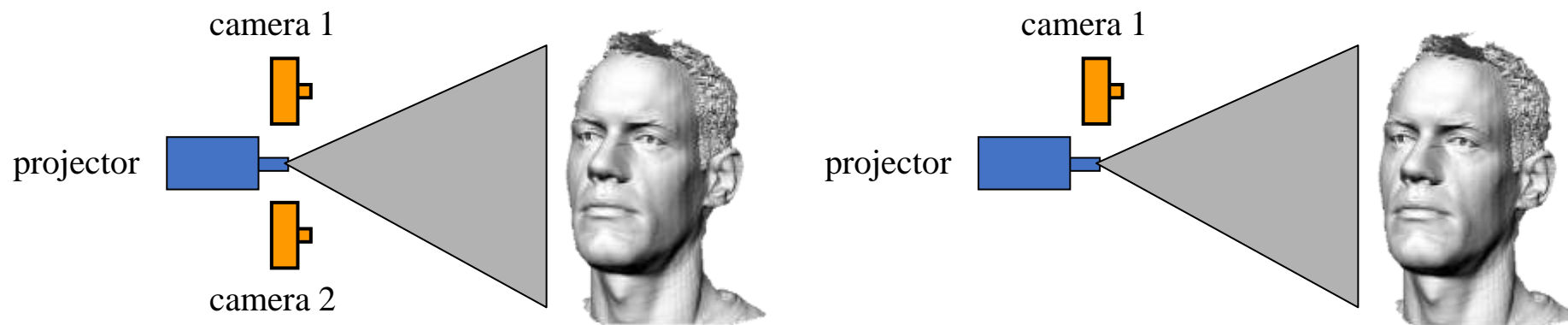
1.物体反应: 说明物体如何反应投射的光图案。光图案在物体表面上会发生形变, 这个形变取决于物体表面的形状和距离。

2.相机捕捉: 提到一个或多个相机用于捕捉投射到物体上的光图案, 记录光图案的形变和变化。

3.深度计算: 解释如何根据捕捉到的图像与投射图案的变化来计算物体表面上不同点的深度信息。这通常需要计算图案的位移或形变。



Li Zhang's one-shot stereo



- 投射已知的、结构化的光图案到物体表面
 - 通过观察物体对光图案的扭曲或变形来推断物体的三维结构
 - Kinect, iPhone X (using IR)等设备测深度的手段

结构光



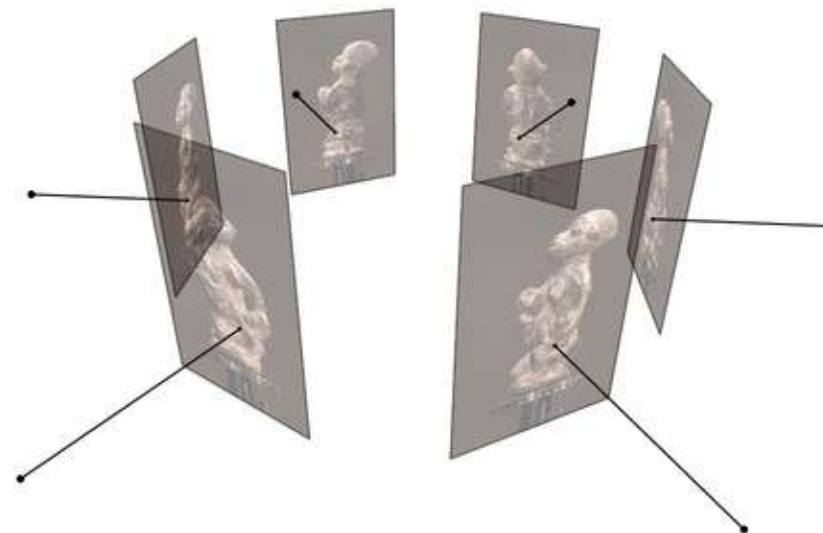
<https://ios.gadgetsacks.com/news/watch-iphone-xs-30k-ir-dots-scan-your-face-0180944/>

多视角立体视图

Input: (已知相机参数下的) 多视角图像

Output: 3D模型

多视角相对双目更全面、更精准



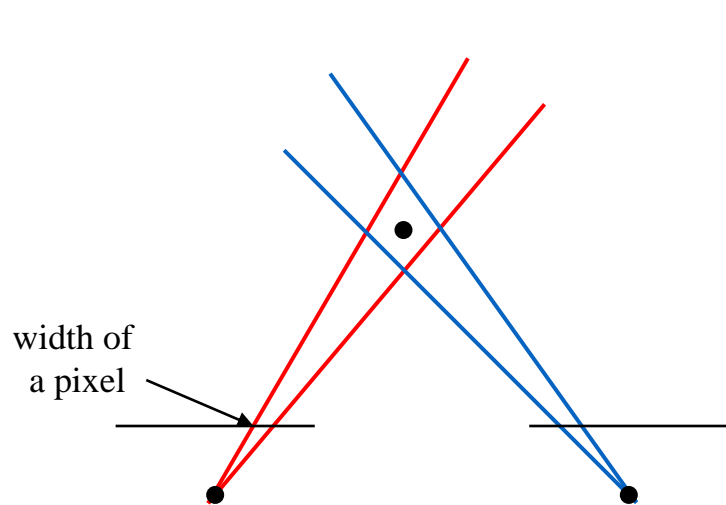
Binocular Stereo



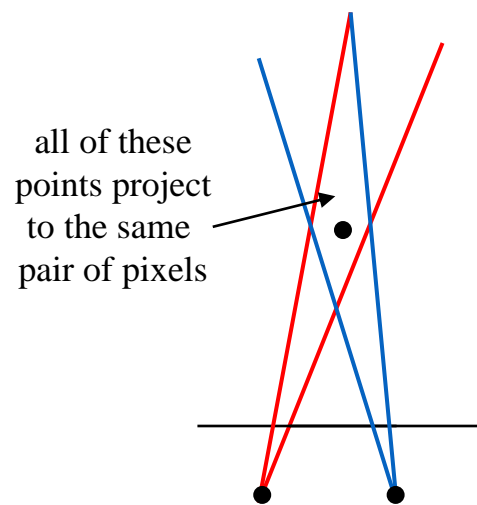
Multi-view stereo



双目视图 vs 多视角视图



Large Baseline

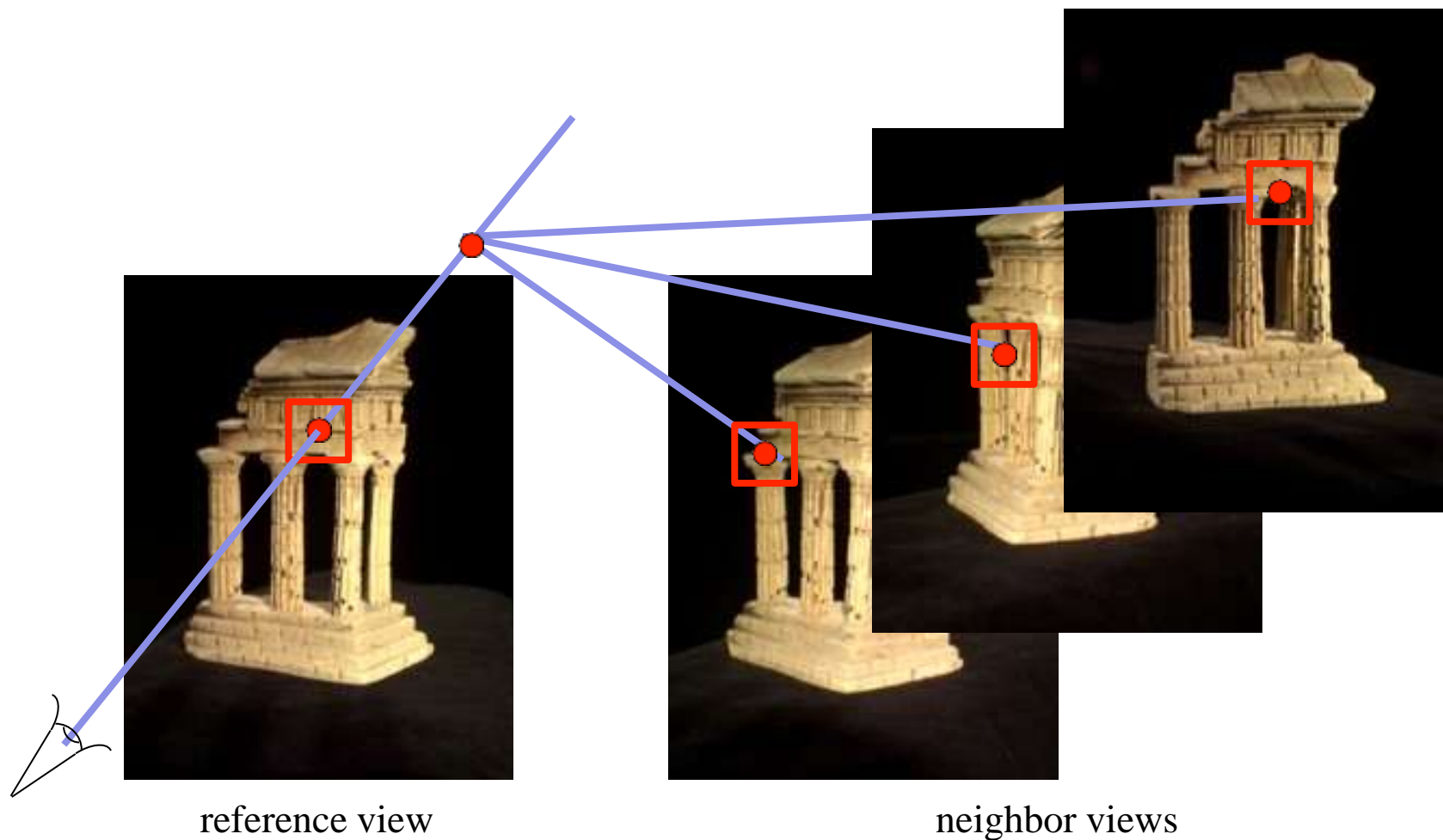


Small Baseline

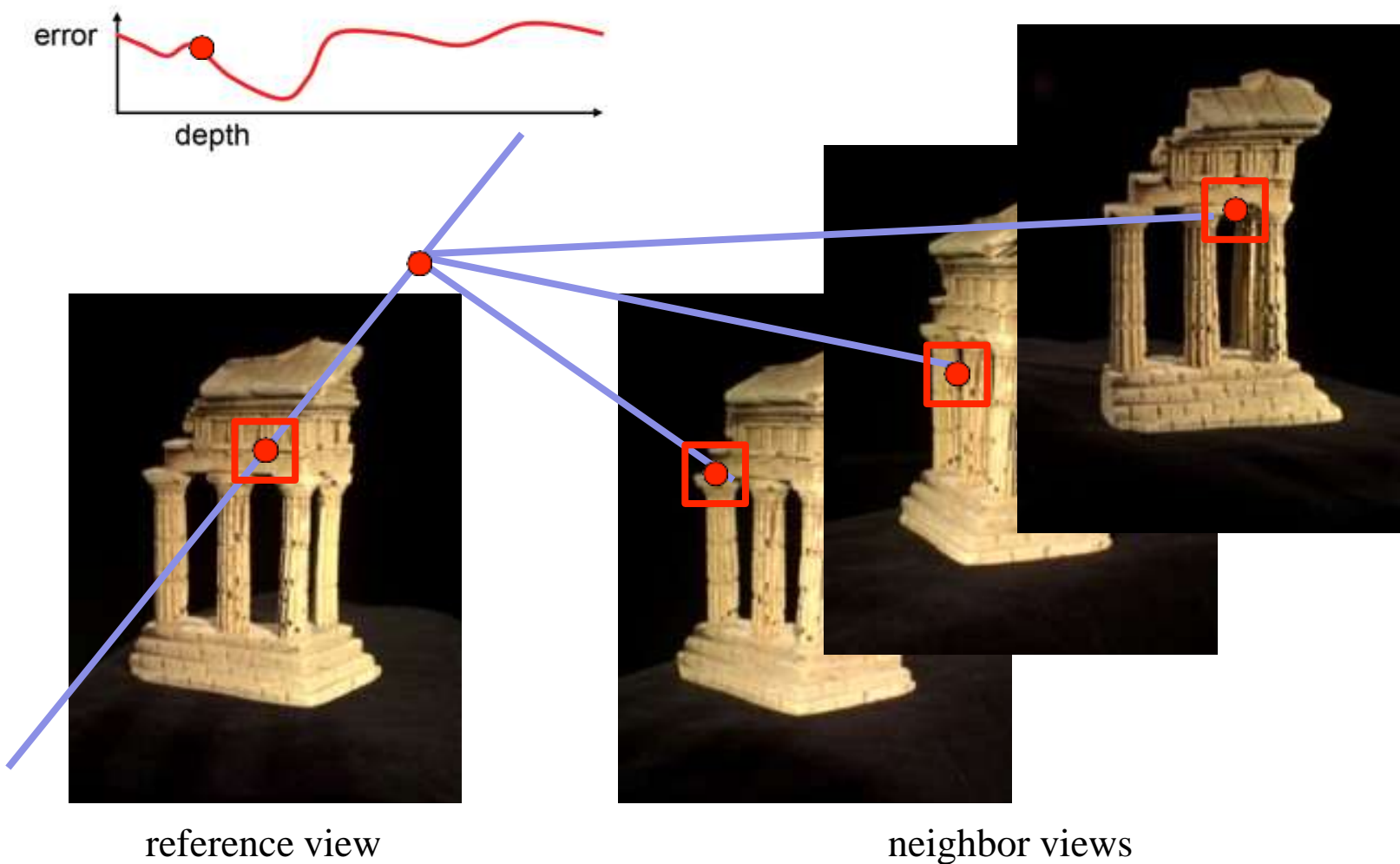
哪种baseline比较好?

- Too small: 相差小, 精度较低
- Too large: 难以找到匹配的点

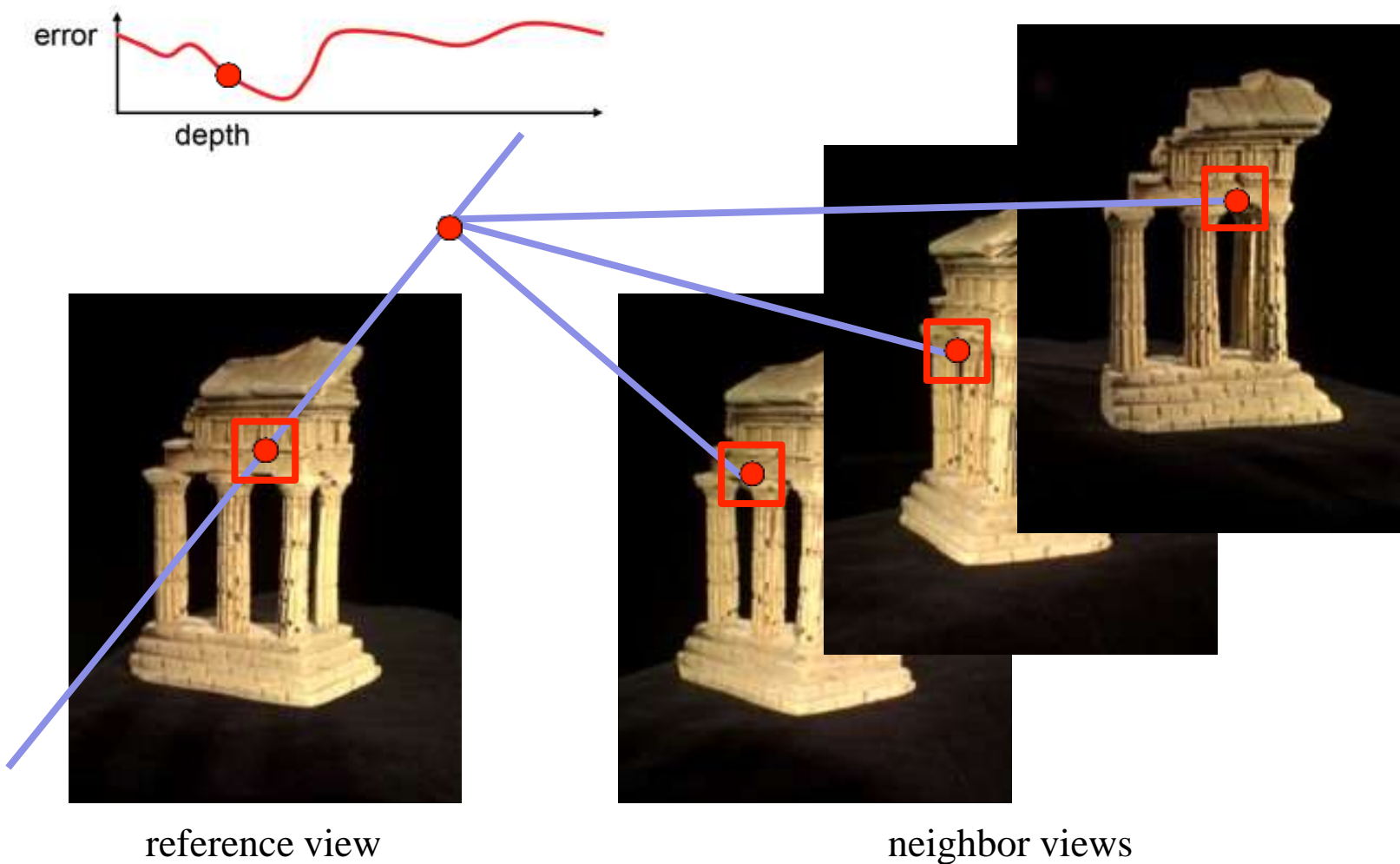
多视角视图



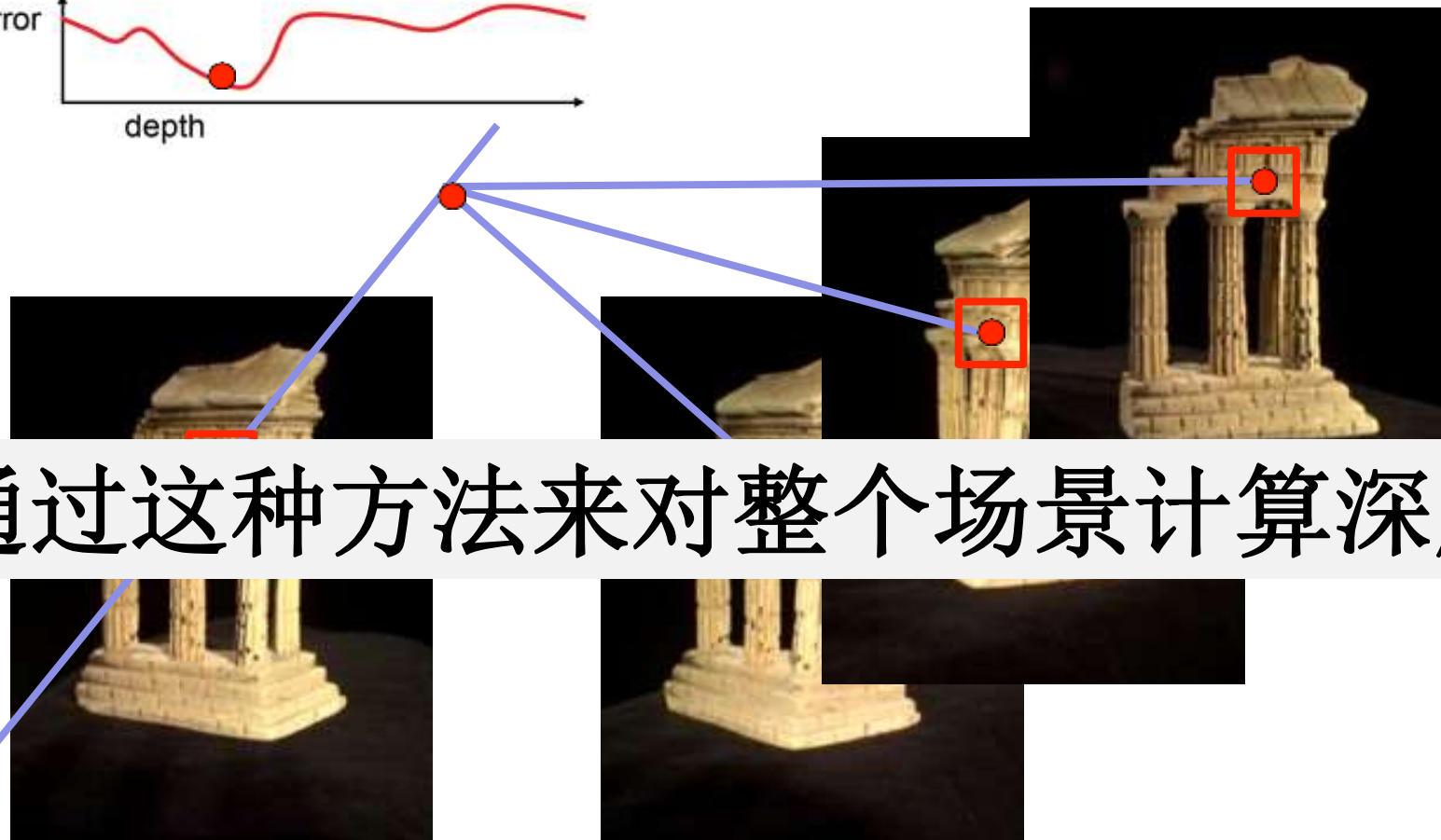
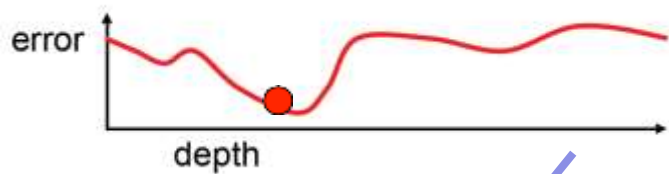
多视角视图



多视角视图



多视角视图



通过这种方法来对整个场景计算深度

reference view

neighbor views



Data SIO, NOAA, U.S. Navy, NGA, GEBCO

Google earth

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.



拓展：形状



双目视图通过相机视差来计算深度

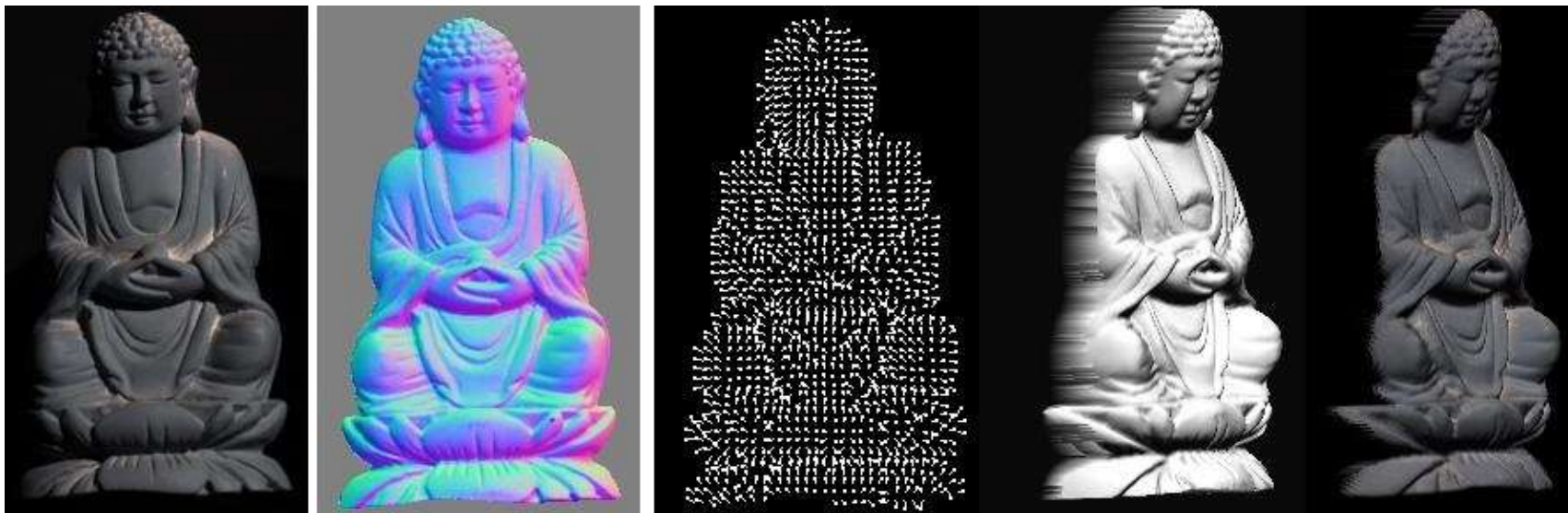
拓展：形状



利用多个光源和亮度来计算形状

Photometric Stereo: 光度立体

形状表示方法



Input
(1 of 12)

Normals (RGB
colormap)

Normals (vectors)

Shaded 3D
rendering

Textured 3D
rendering

大作业——17周周日截止

- 不会提供公共计算资源
- 分组，最多每组5人
 - 分组在**第四周**确定，之后不得更改
- 课题：物体检测/图像生成
 - 每组内部用统一模型，提交统一报告
 - 每个组员使用自己图片，组员之间不可以重复，每个同学提供自己的实验结果
- 可以自行定义与课程相关的题目
 - 如自行选题，需要在**第四周**前联系确定是否合适

“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

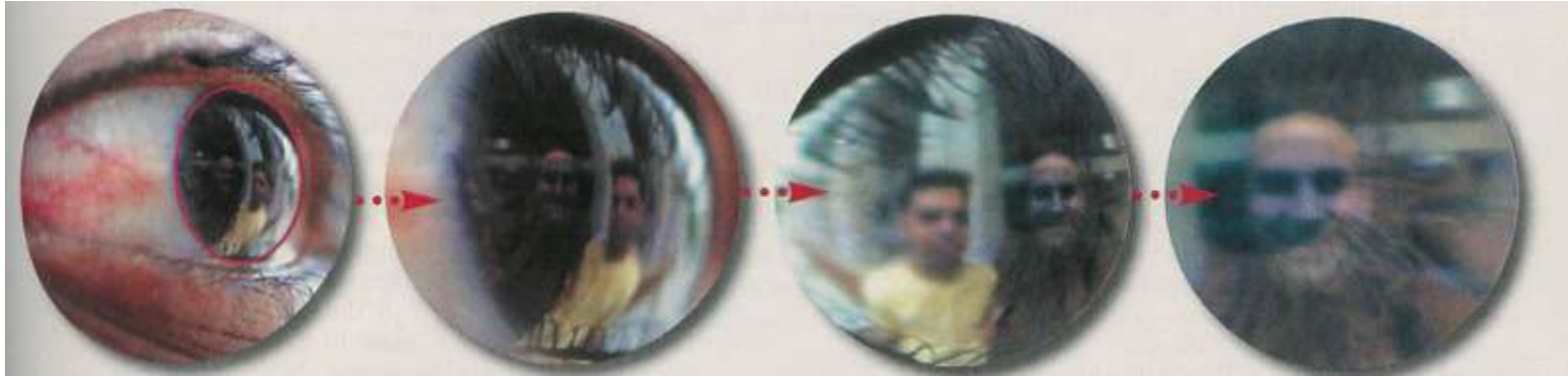
- Forensics: 取证



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》

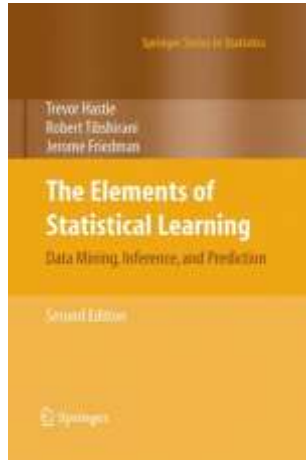
感知与理解



<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”

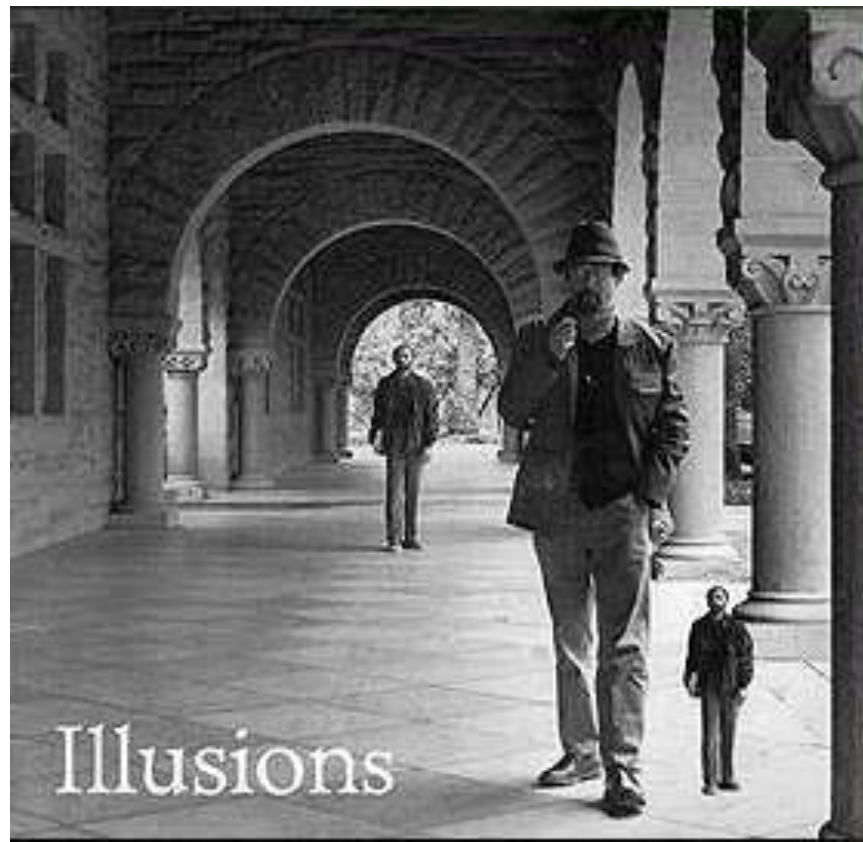
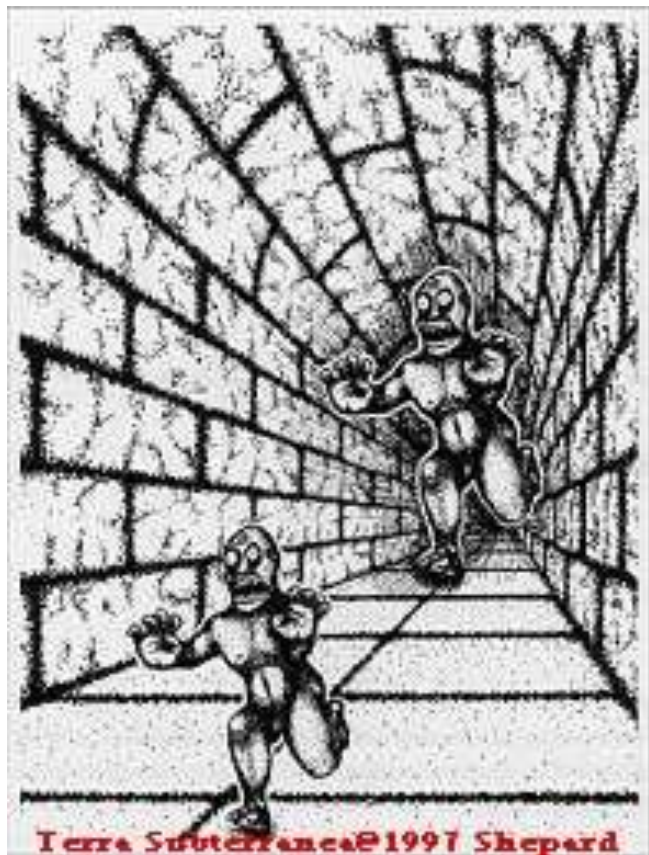
为什么 需要 统计模型？

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？



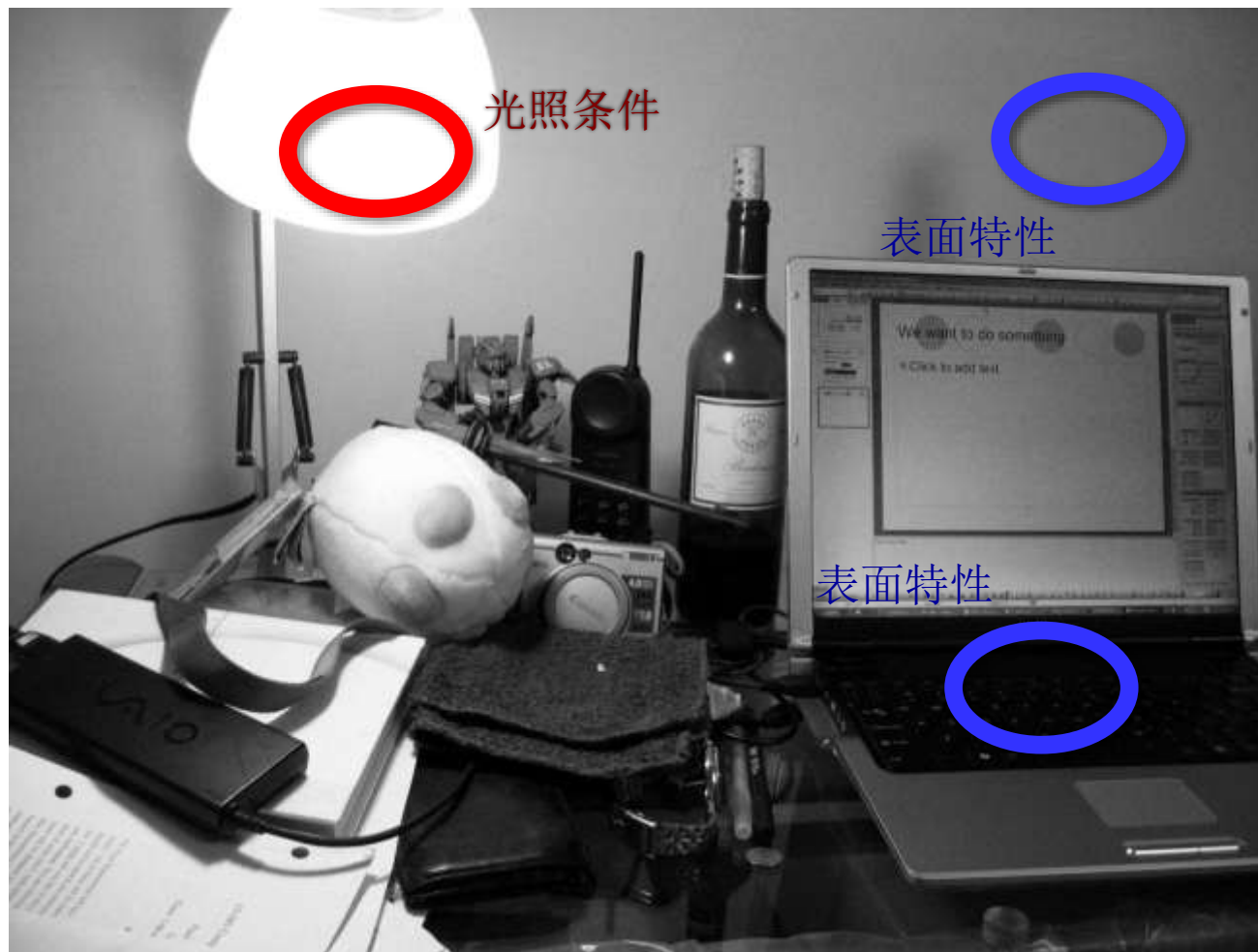
消失的圆点



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).

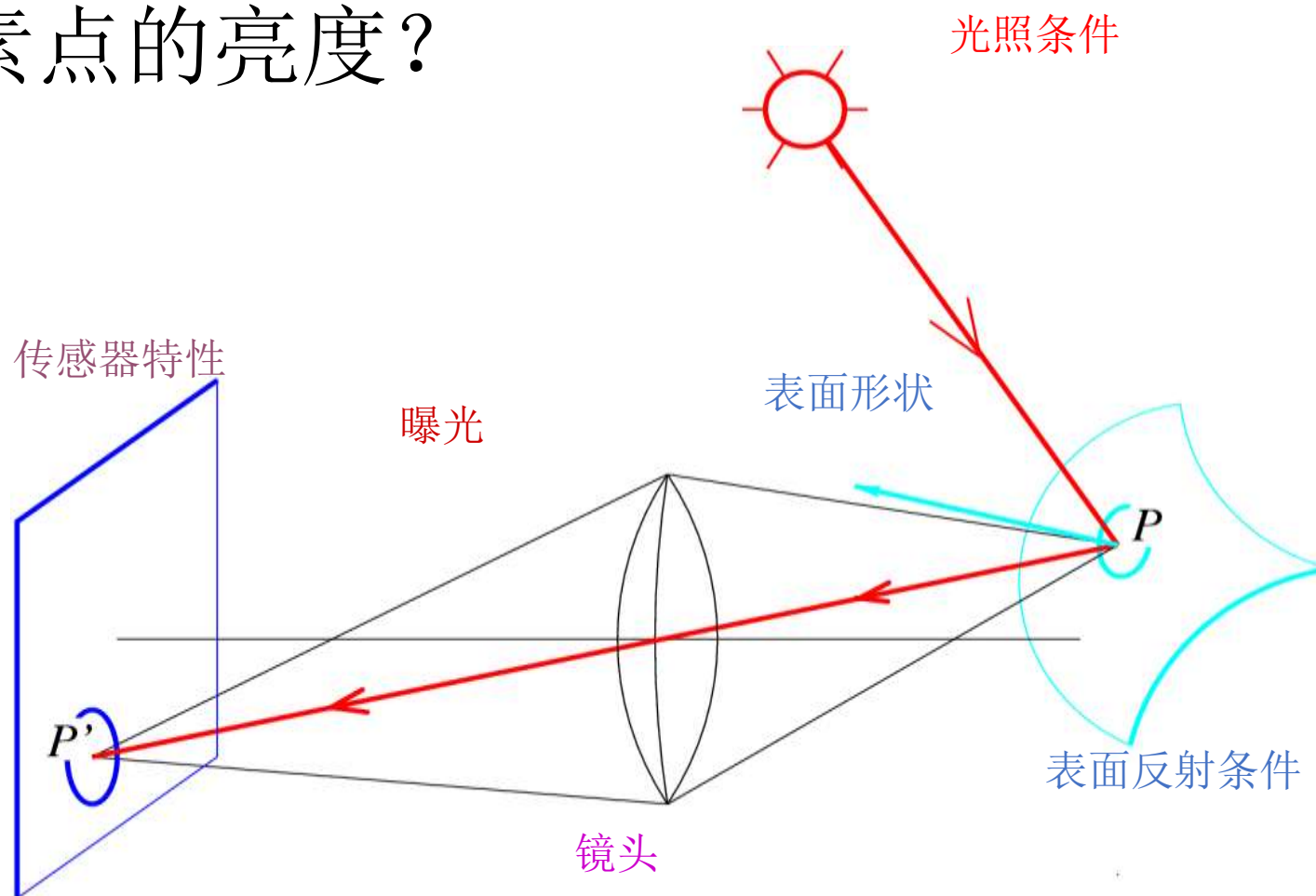
Radiometry: 辐射度量

- 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？



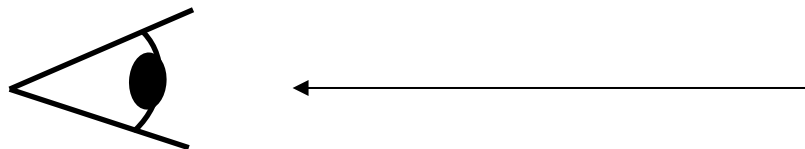
什么是光？

光的感知

- 怎样把辐射转换为“色彩”？
- 我们能看到哪些颜色？

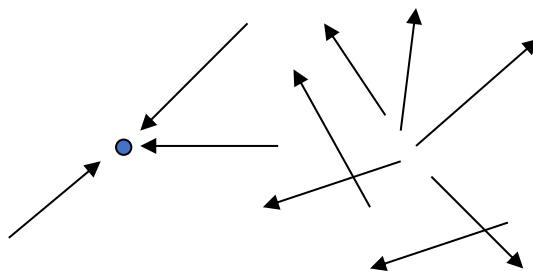
光是一种电磁辐射，它属于电磁波谱的一部分

- $R(\lambda)$ 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。

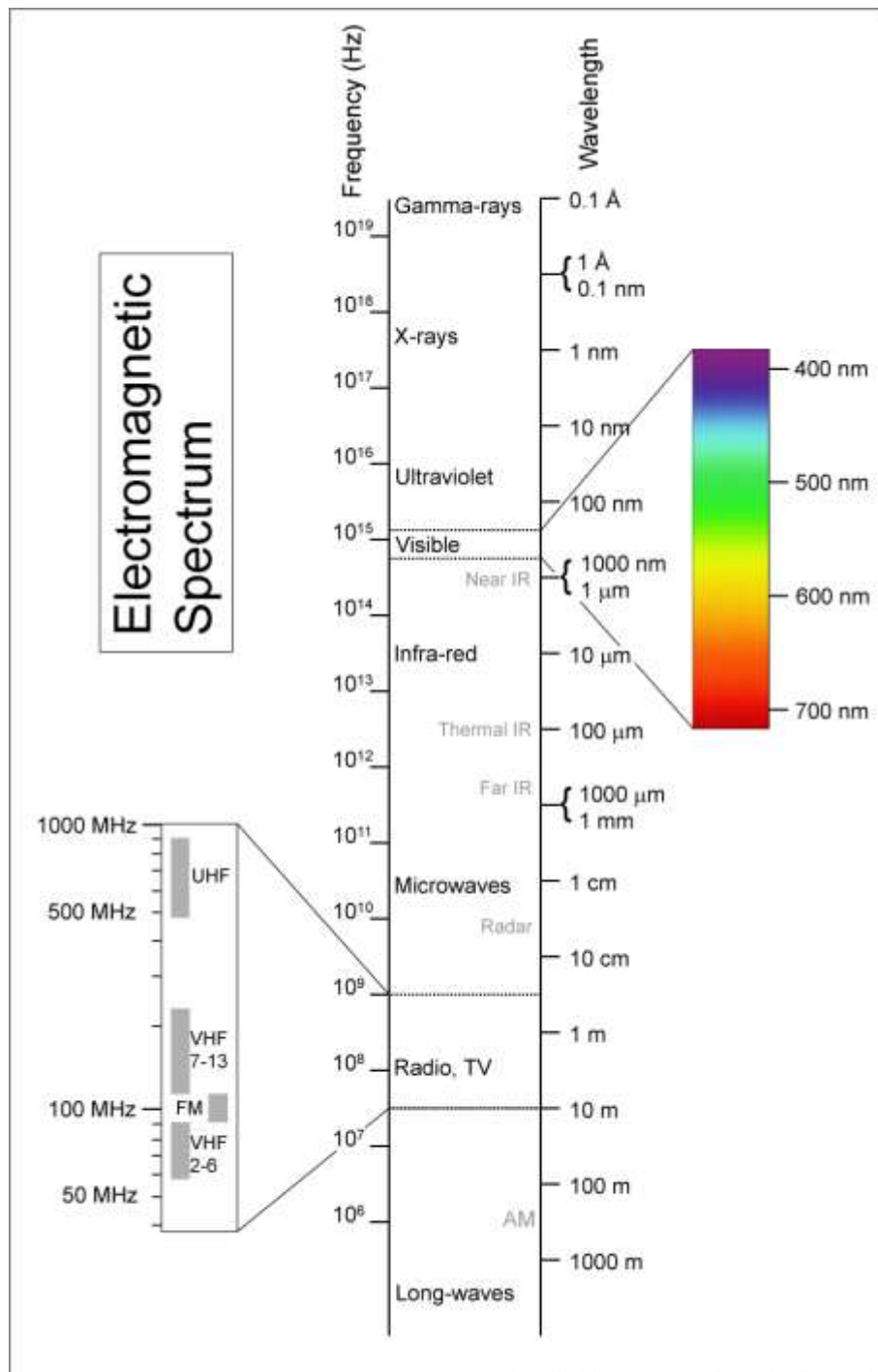


plenoptic function 定义描述光的函数:

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

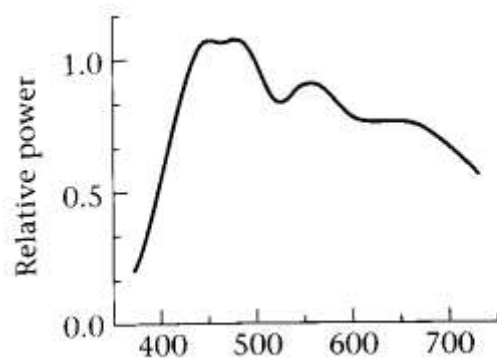
可见光

我们只能看到一部分波段的光辐射

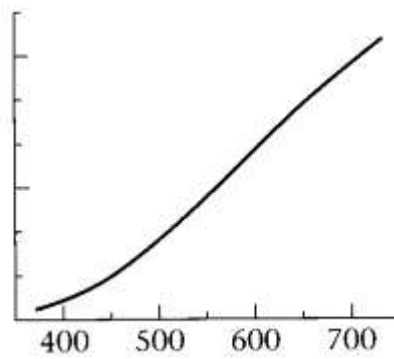


光谱

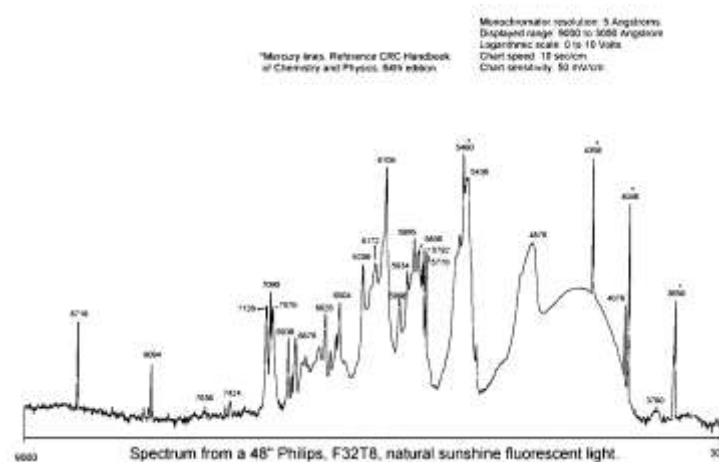
- 我们看到的外表是可见光对应的光谱
 - 光谱对应的是每个波长的强度



日光



钨丝灯

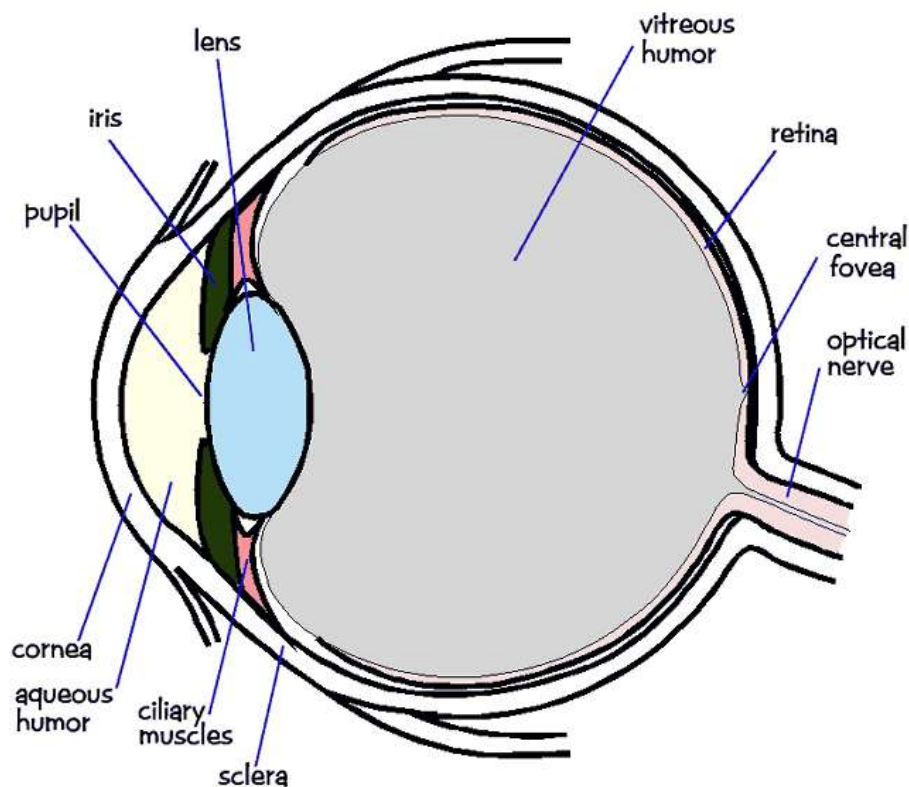


荧光灯

颜色就是从这种光谱通过复杂的形式转换过来的

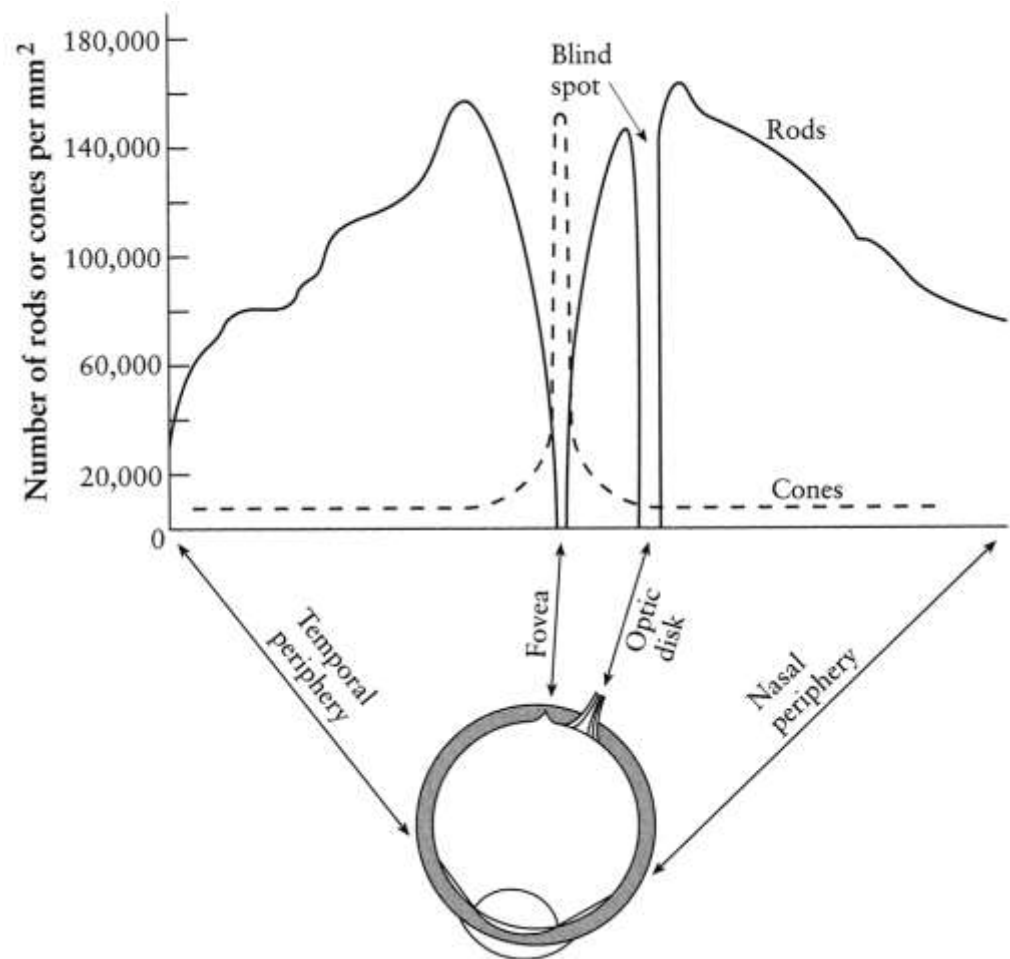
人类的视觉系统

- 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



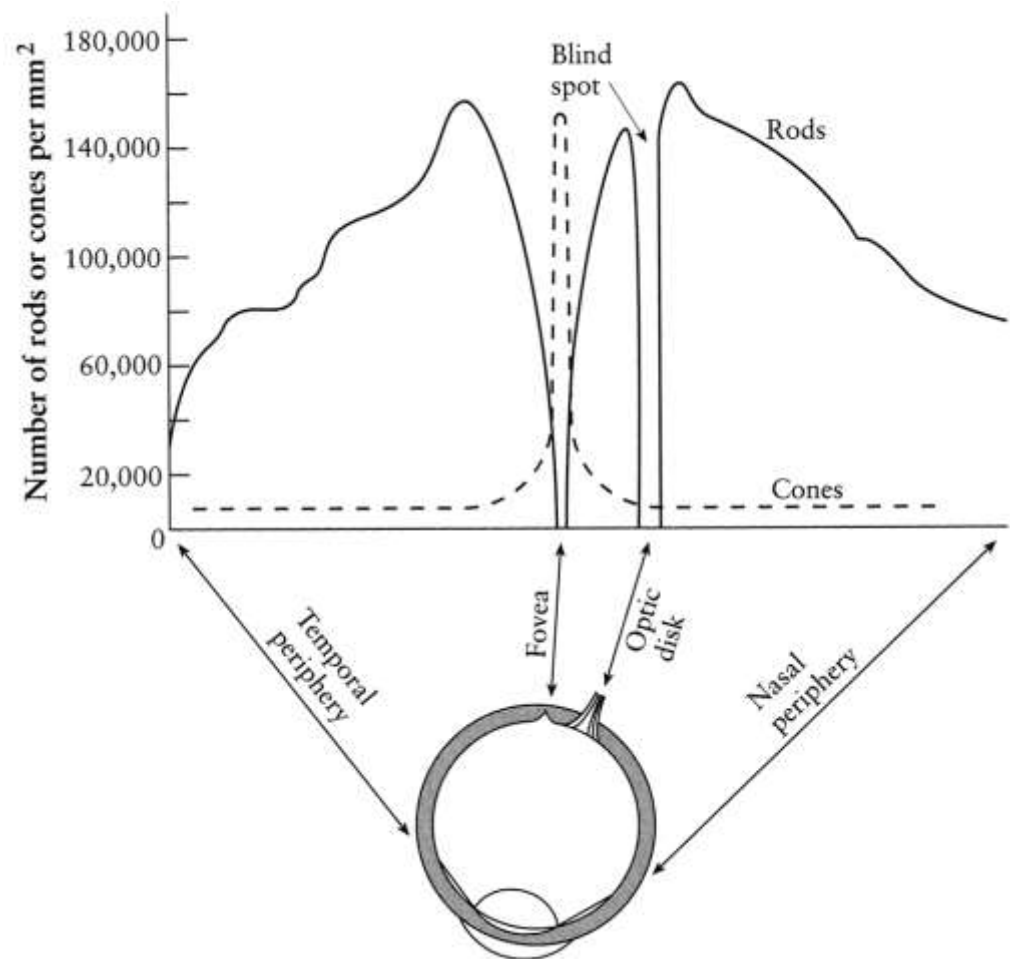
视杆与视锥

- 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

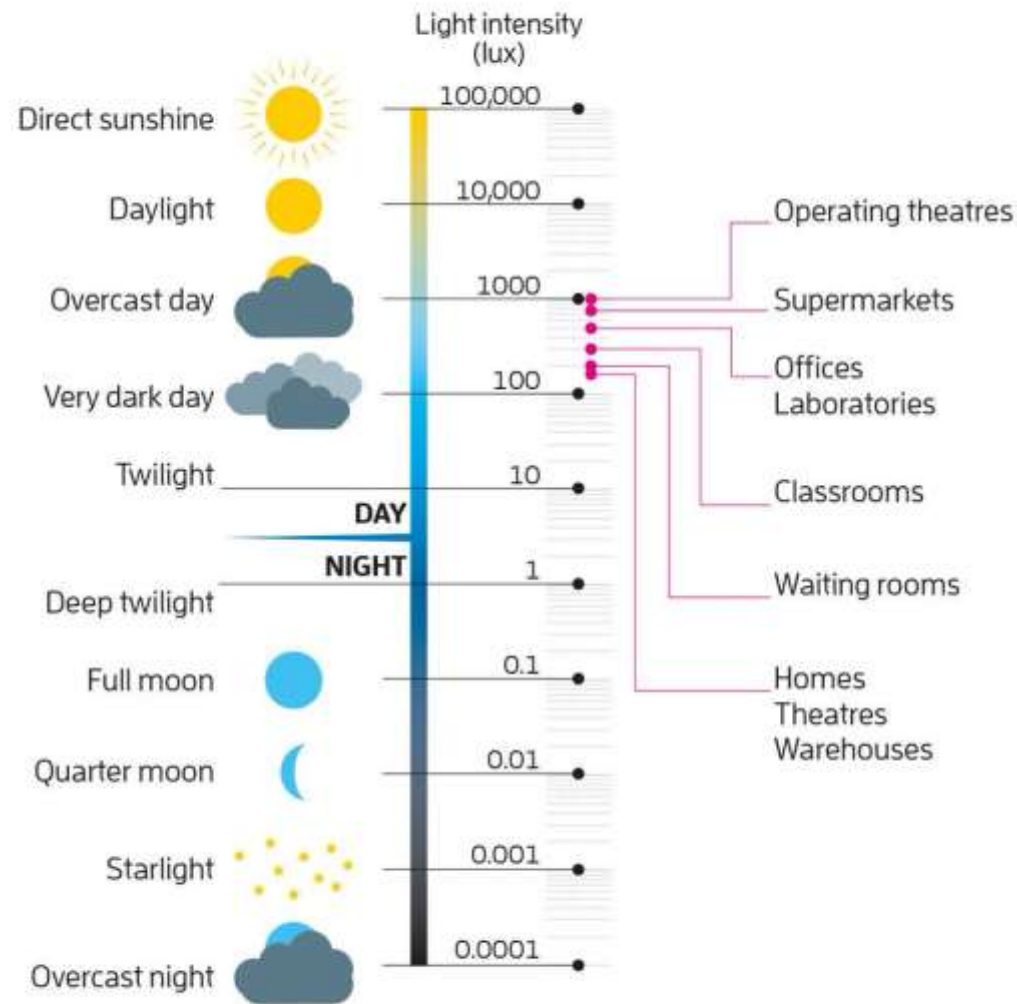
动态视觉系统

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

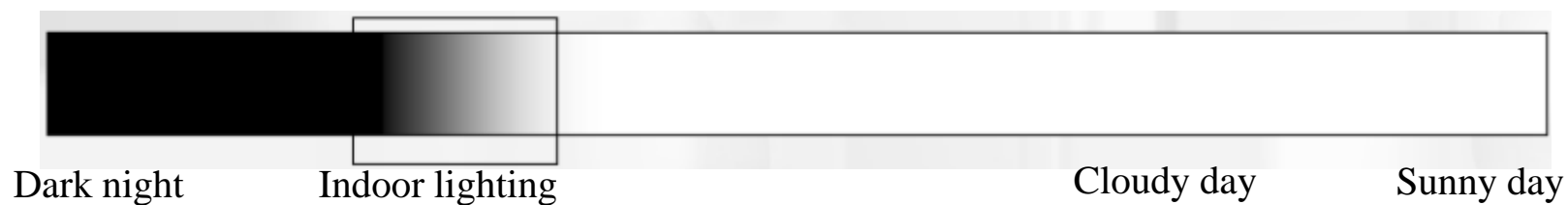
<https://threader.app/thread/1134003178515701762>

动态视觉系统



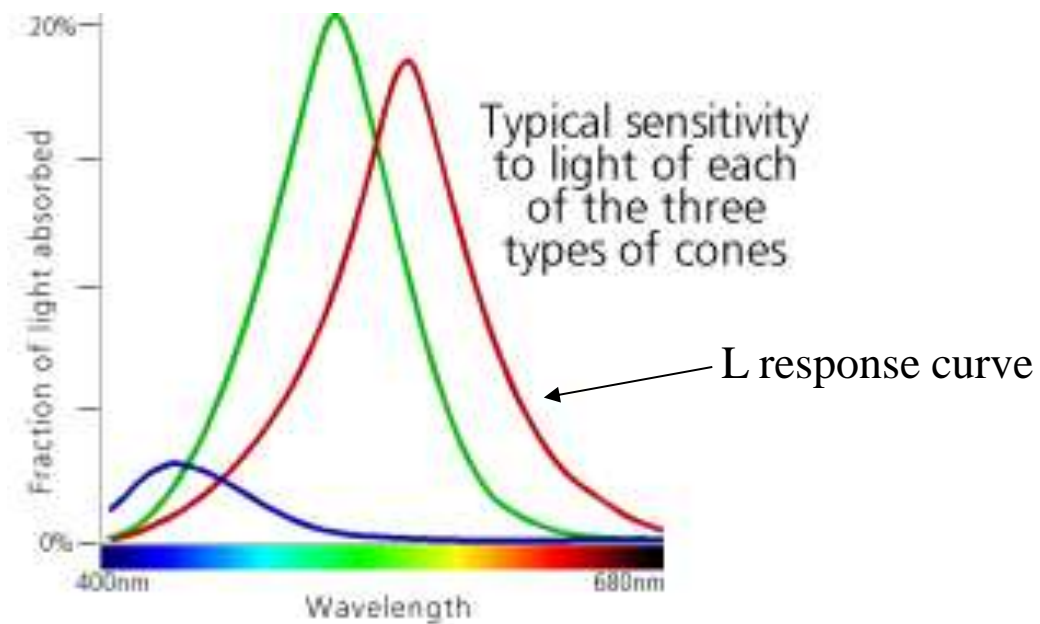
如果我们一直对整个范围都非常敏感，那么我们将无法在场景中区分亮度级别。

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。

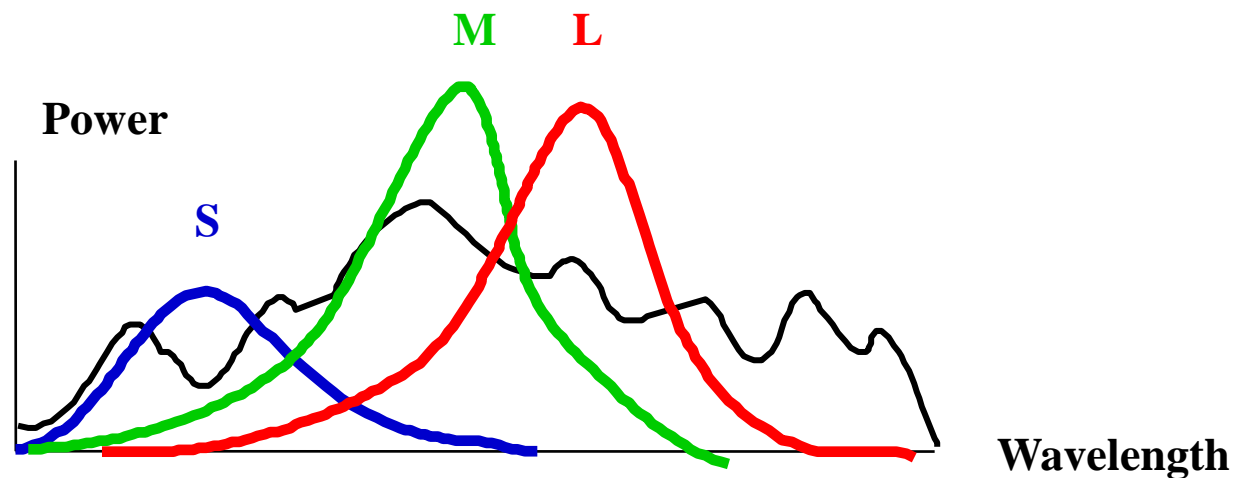


颜色感知

- 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

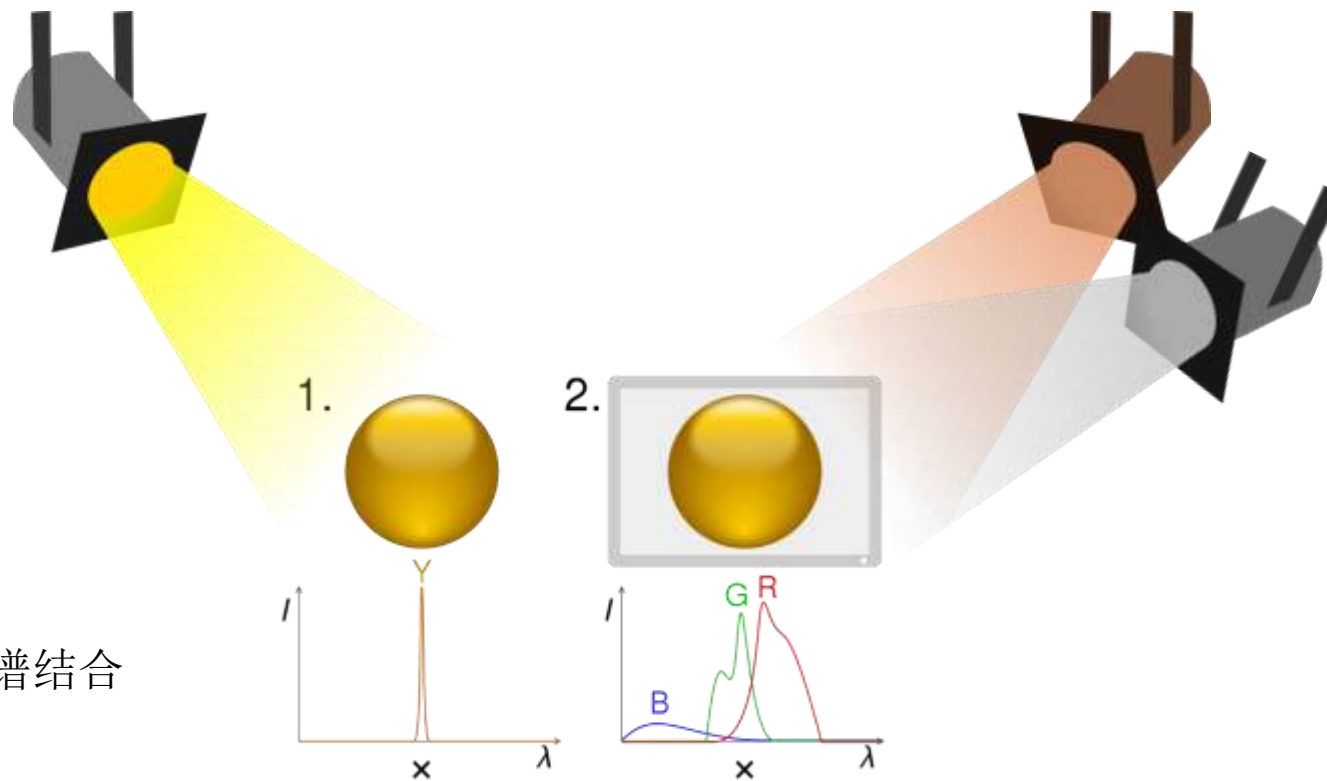


颜色感知



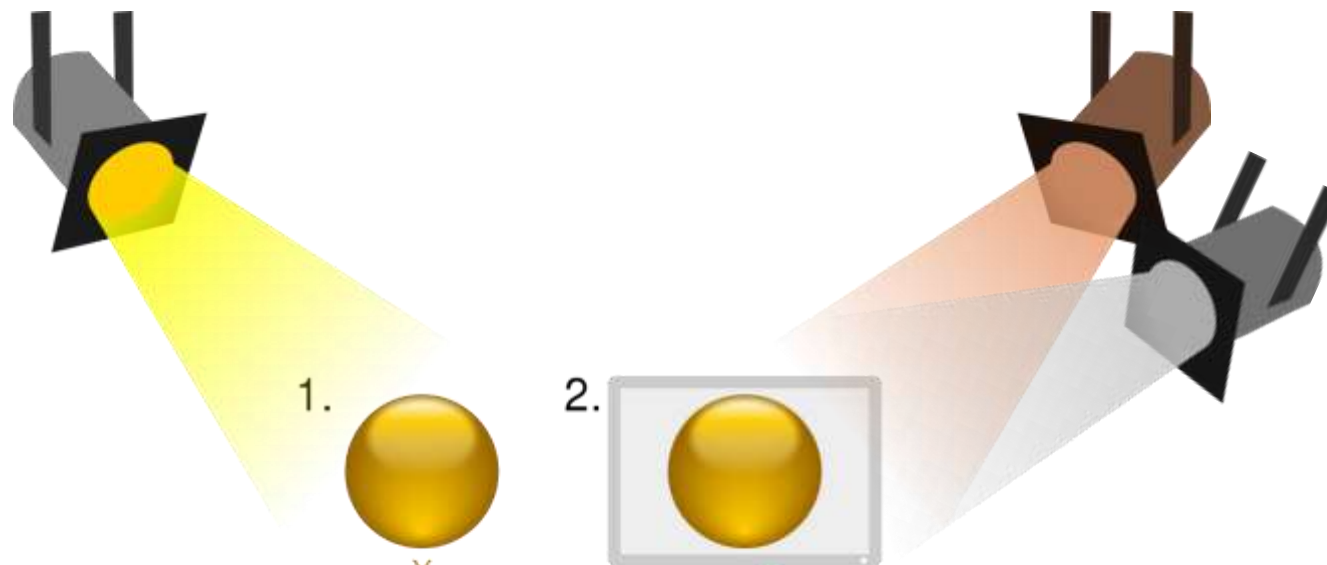
- 视杆与视锥根据光谱感受颜色
 - 在波长上做强度的积分
 - 每个视锥得到一个数字
 - Q: 我们可以直接拿视锥的感受强度来表示光谱吗?
 - A: 不行, 会丢失信息
 - 不同的光谱在某些情况下可能完全无法被分辨
 - 这种光谱被称作 **metamers** 条件等色

条件等色

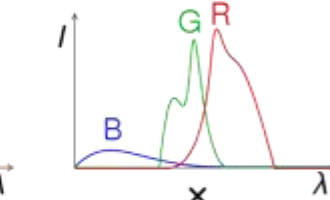
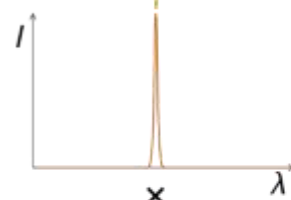


1. 与光源的光谱结合

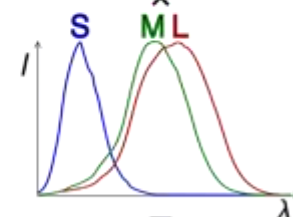
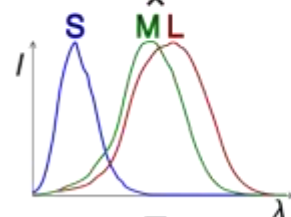
条件等色



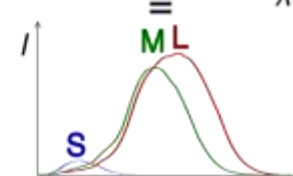
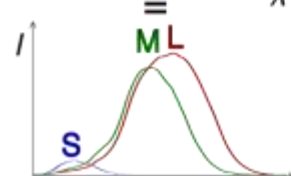
1. 与光源的光谱结合



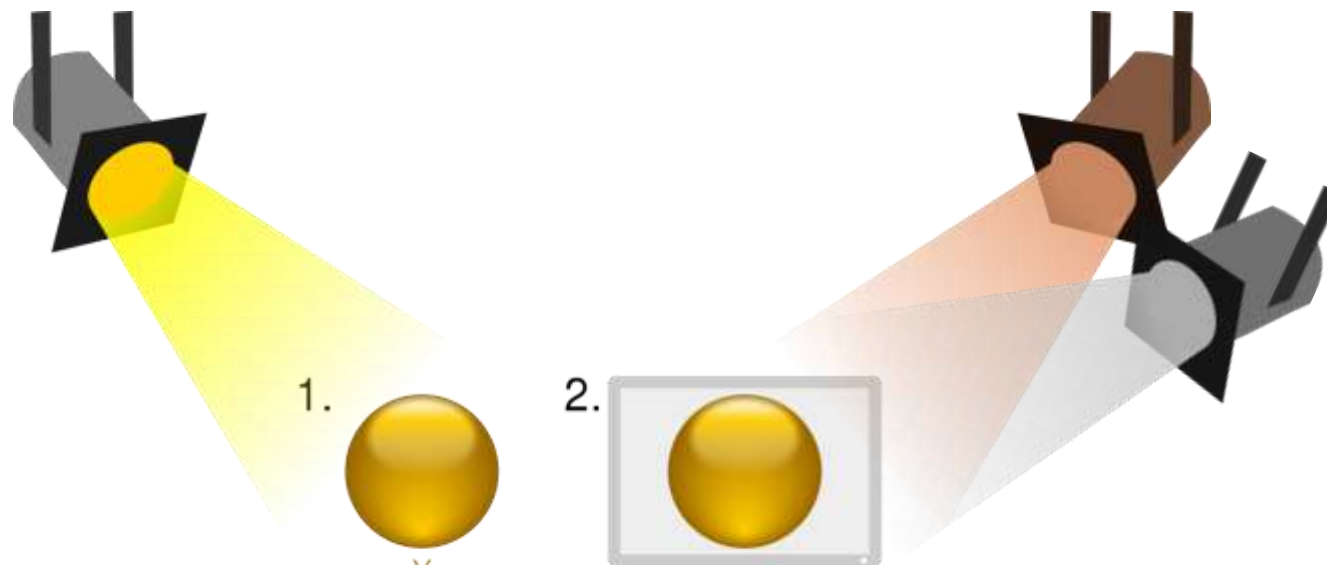
2. 光锥敏感度(S, M, L)



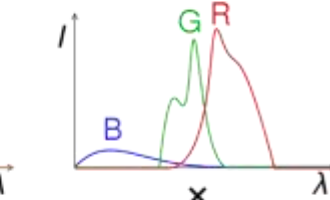
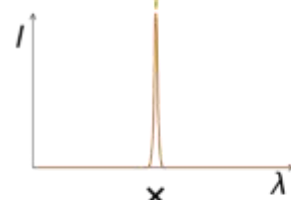
3. 1与2结合



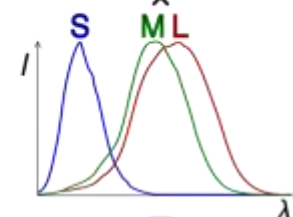
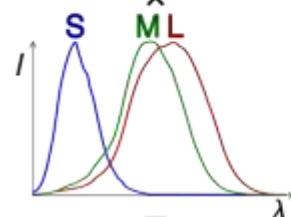
条件等色



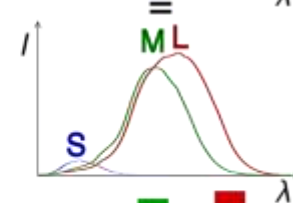
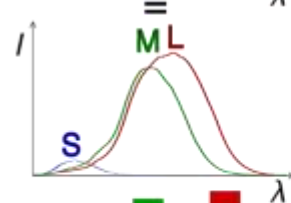
1. 与光源的光谱结合



2. 光锥敏感度(S, M, L)



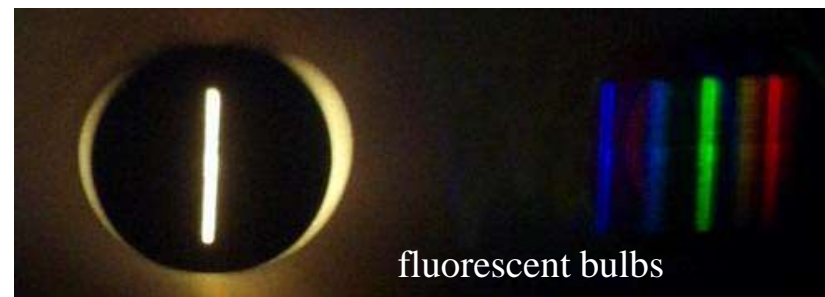
3. 1与2结合



4. 观测颜色: 黄色

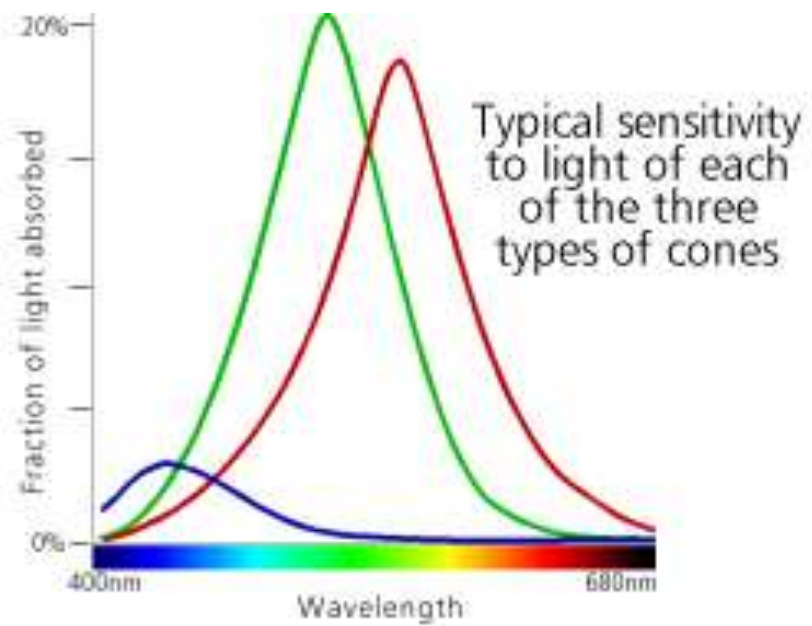
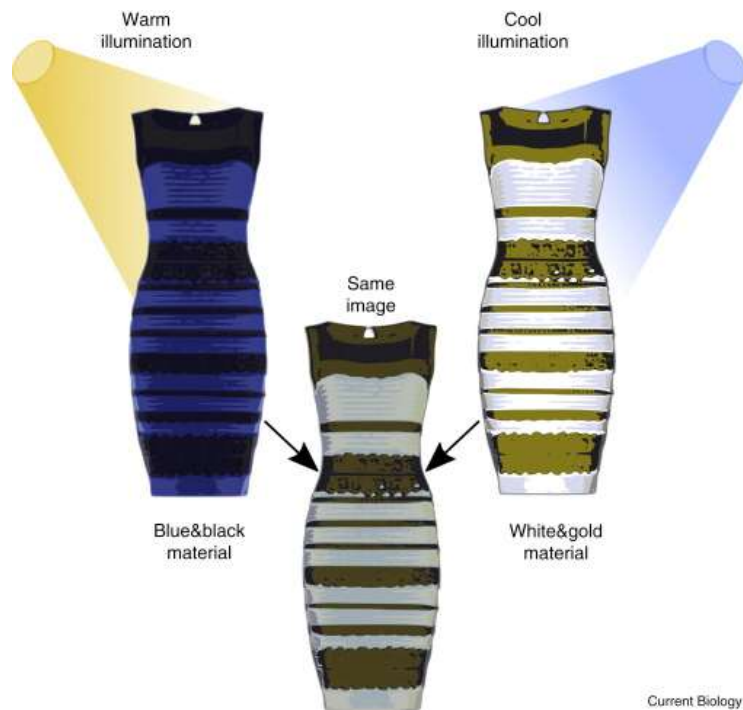


用分光镜 (Spectroscope) 区分不同的灯泡



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

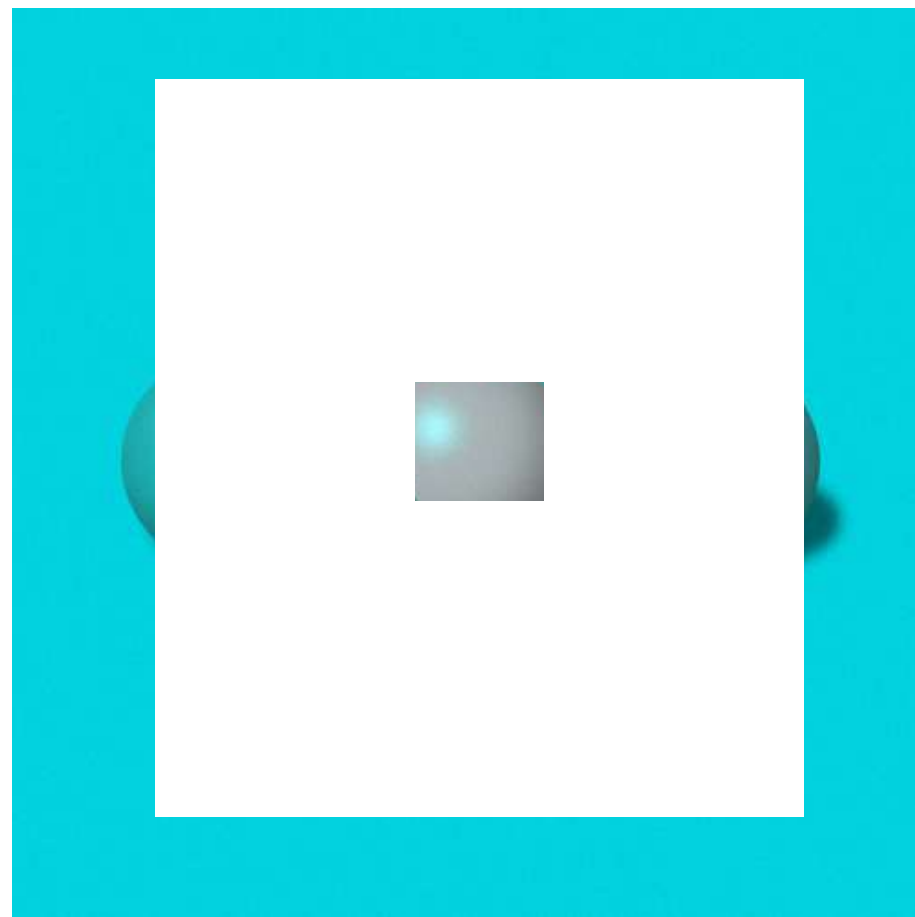
蓝黑还是白金？



中间的球是什么颜色？

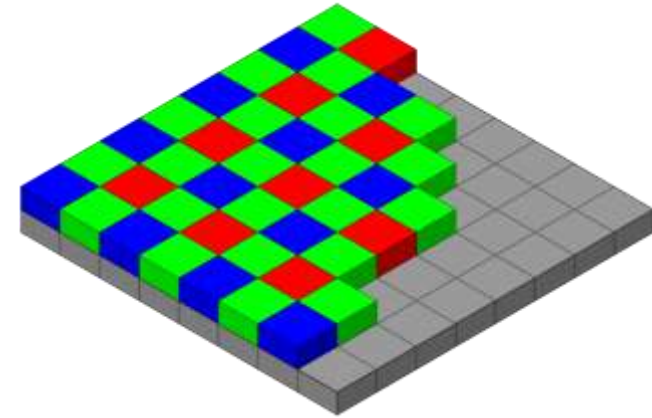


中间的球是什么颜色？

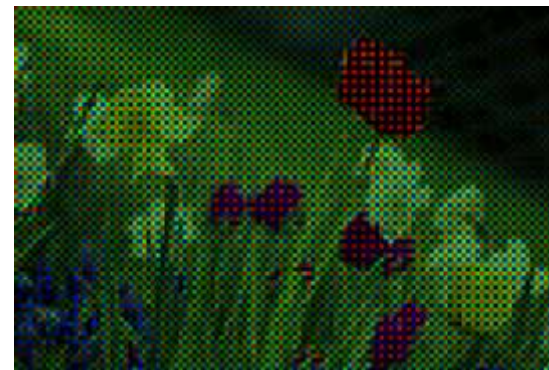


相机模型的色彩

- *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



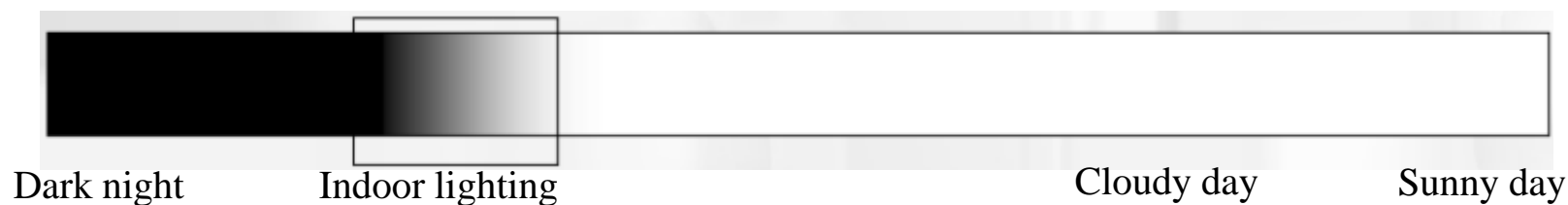
Combined color image (1911)

Recall: 动态视觉系统“尺度”

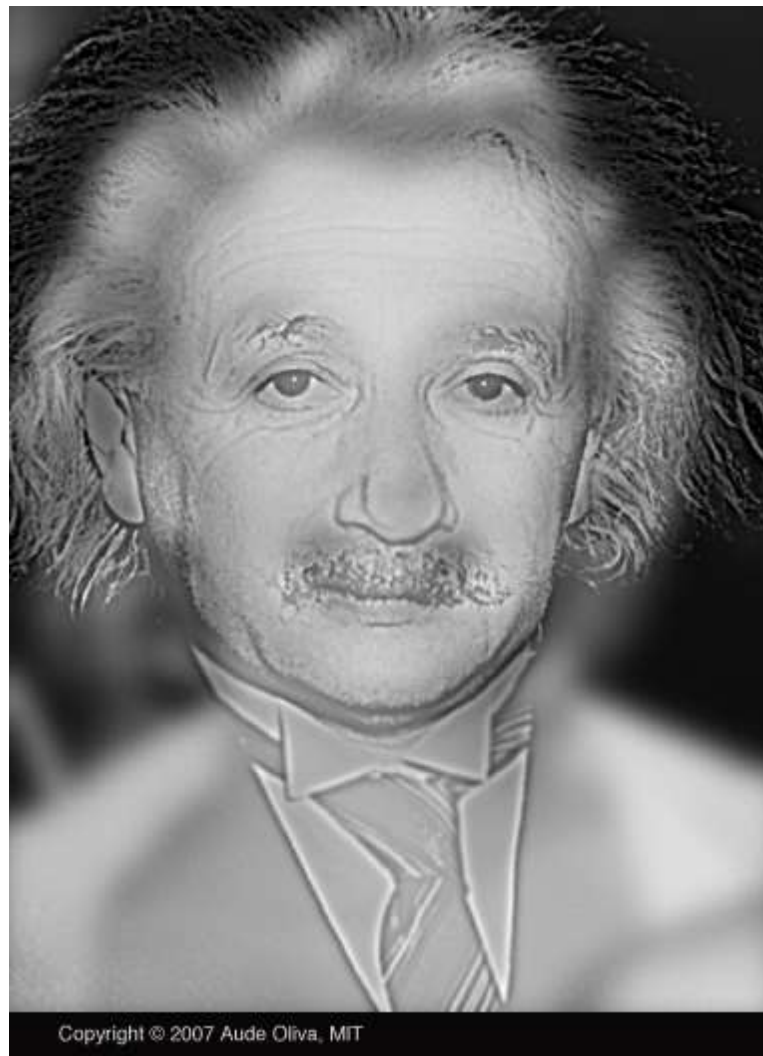


如果我们一直对整个范围都非常敏感，那么我们将无法在场景中区分亮度级别。

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？

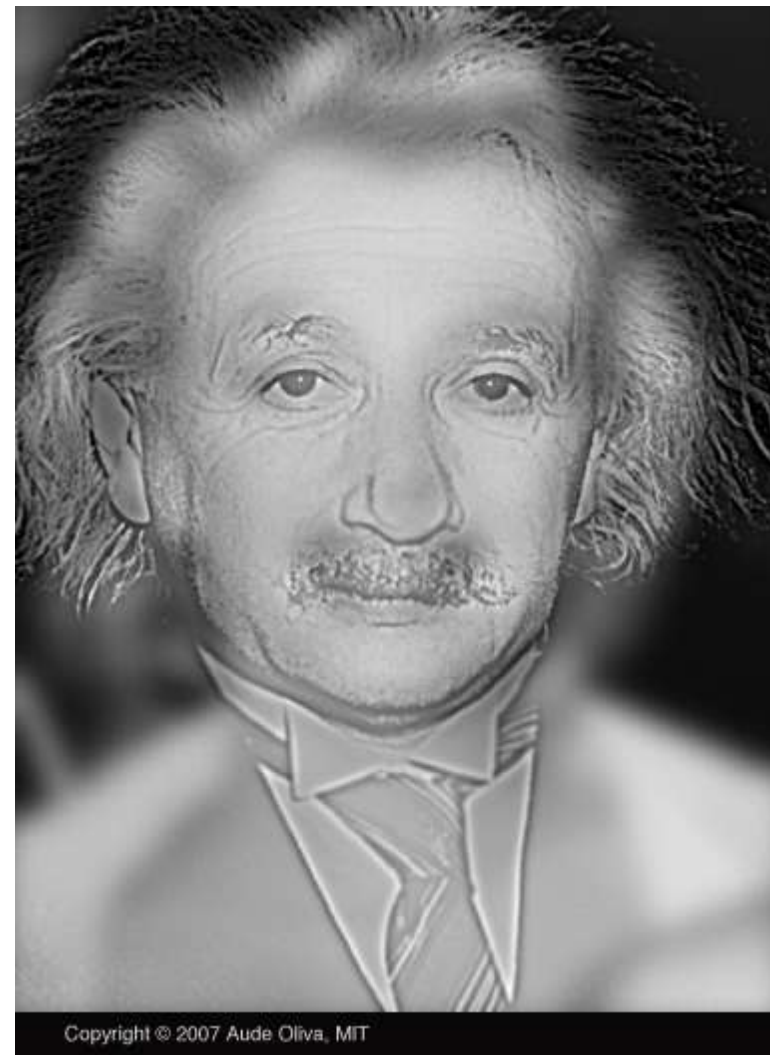
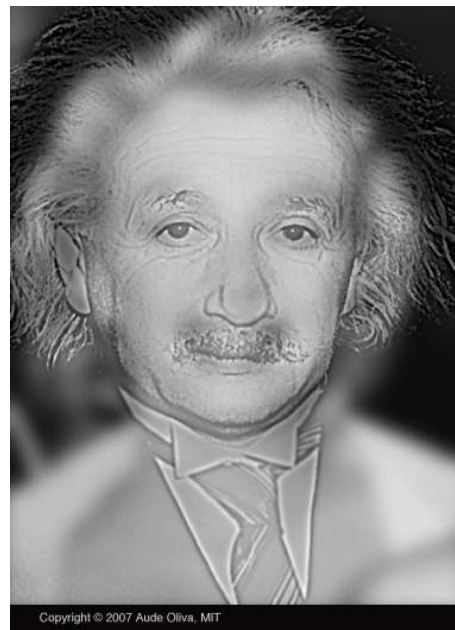


他又是谁？



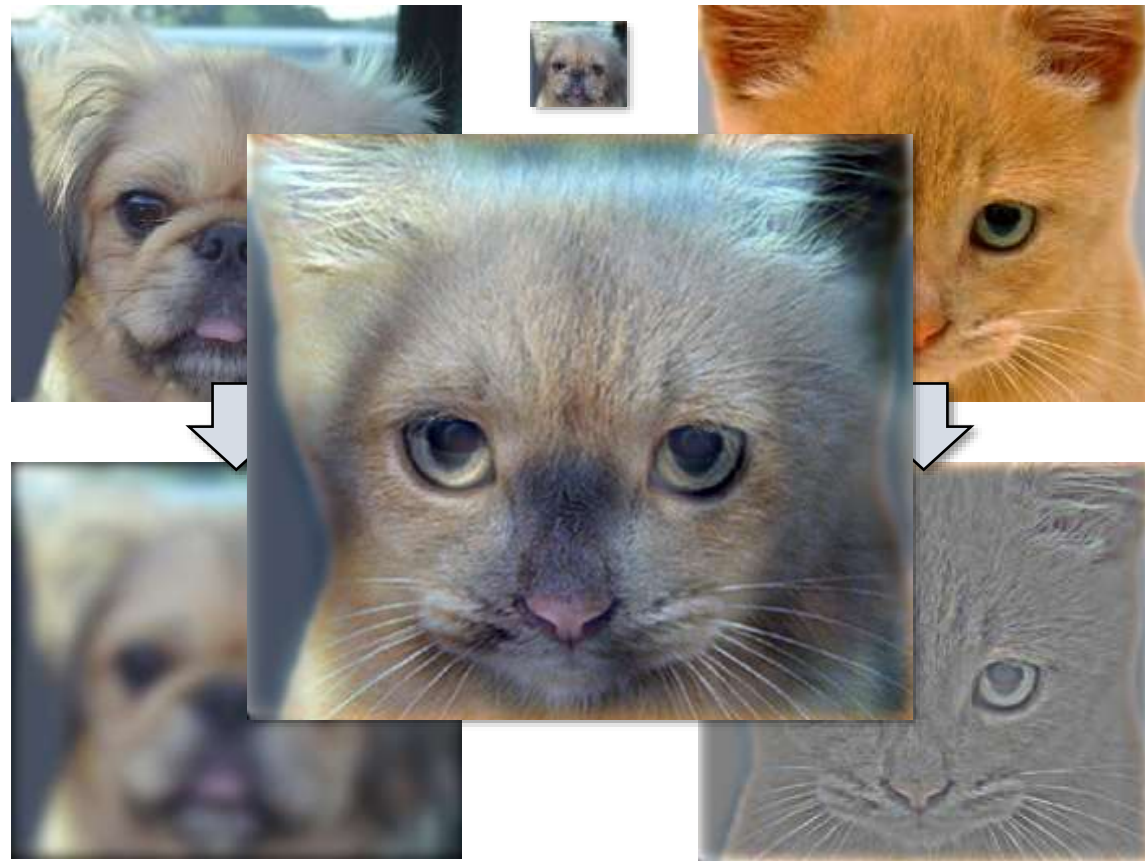
其实是同一张图！

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**