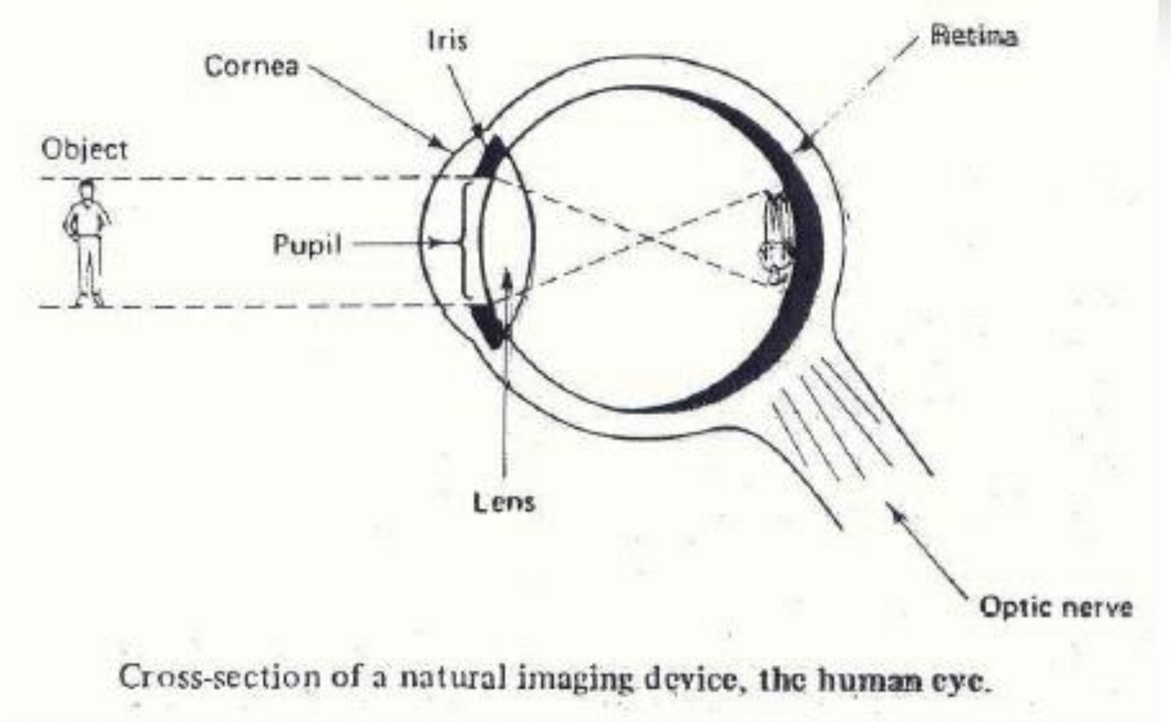
Camera models & Two-view geometry & Stereo Camera

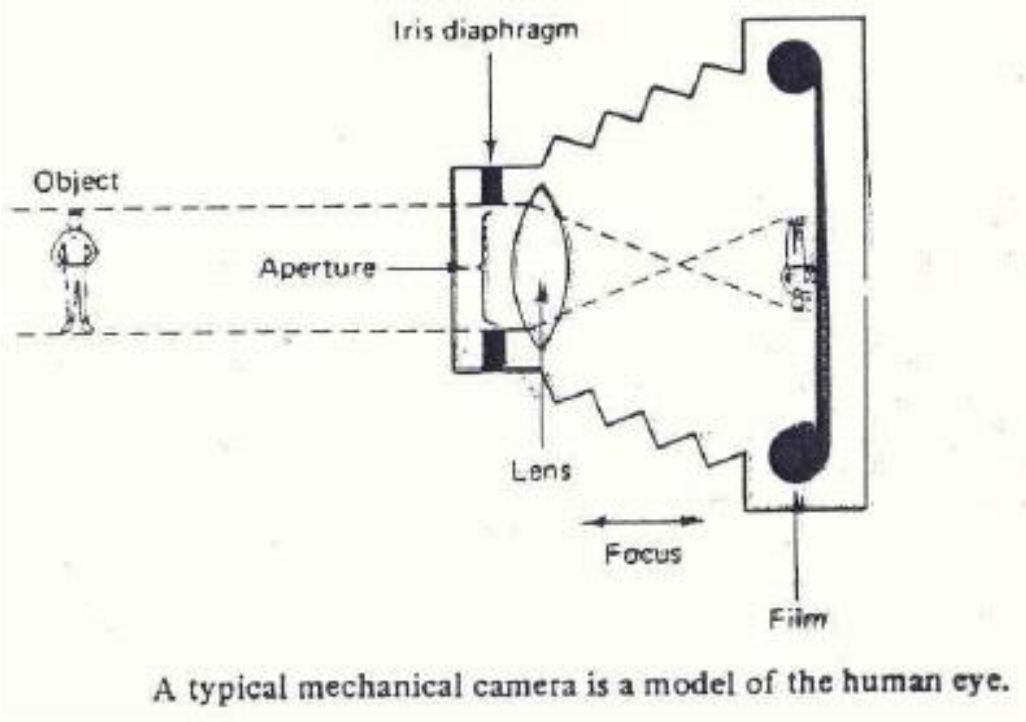
taoguo
20180402

Contents

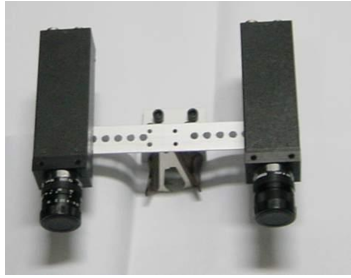
The eyes



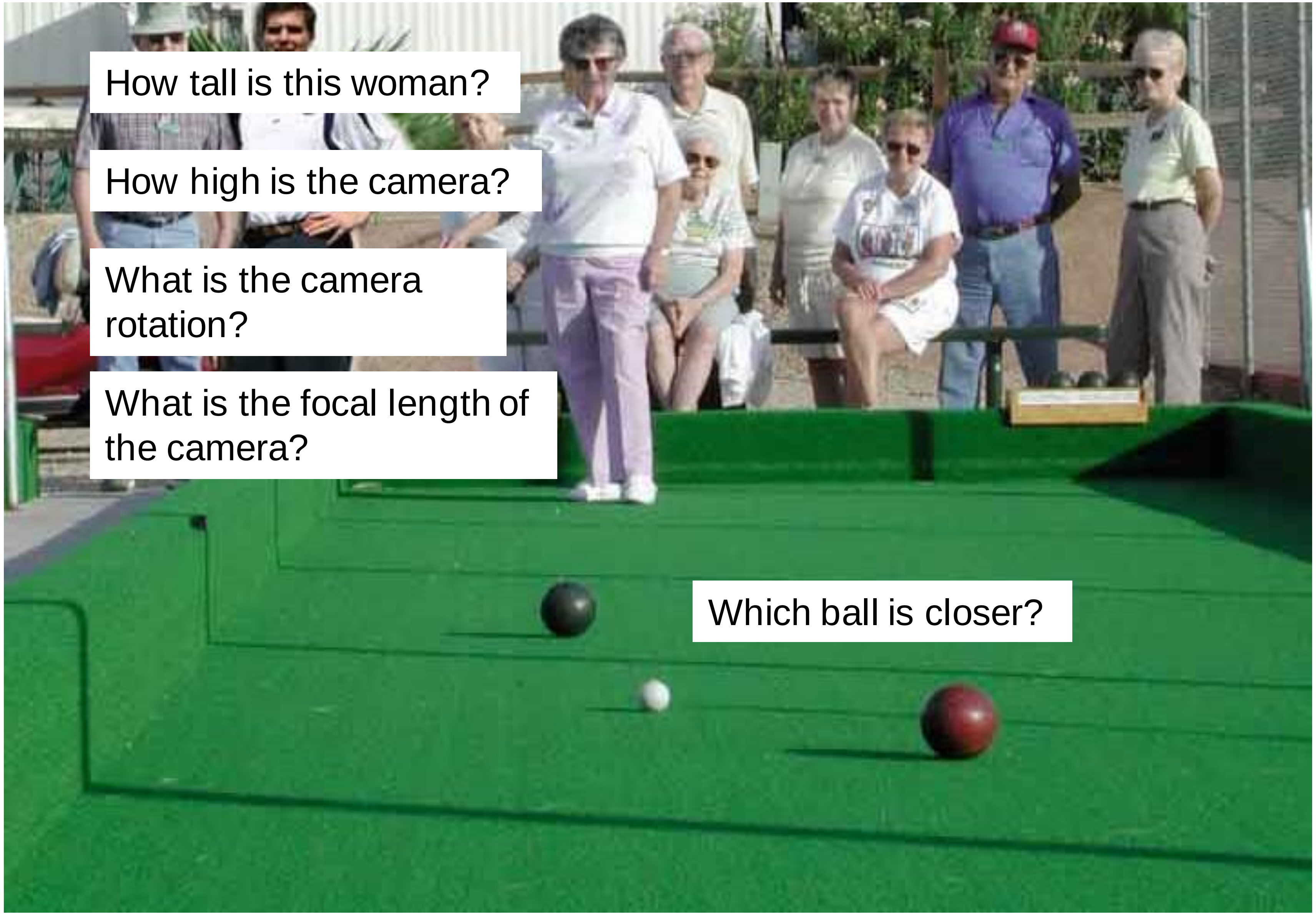
Cross-section of a natural imaging device, the human eye.



A typical mechanical camera is a model of the human eye.

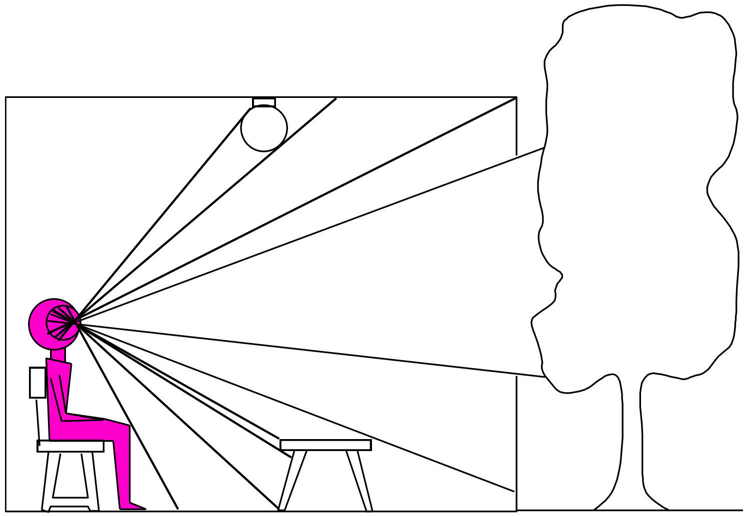


Camera and World Geometry



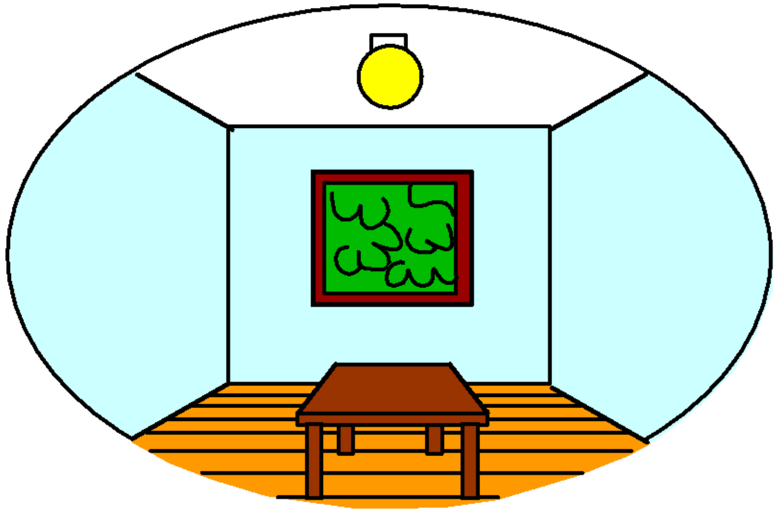
Dimensionality Reduction Machine (3D to 2D)

3D world



Point of observation

2D image

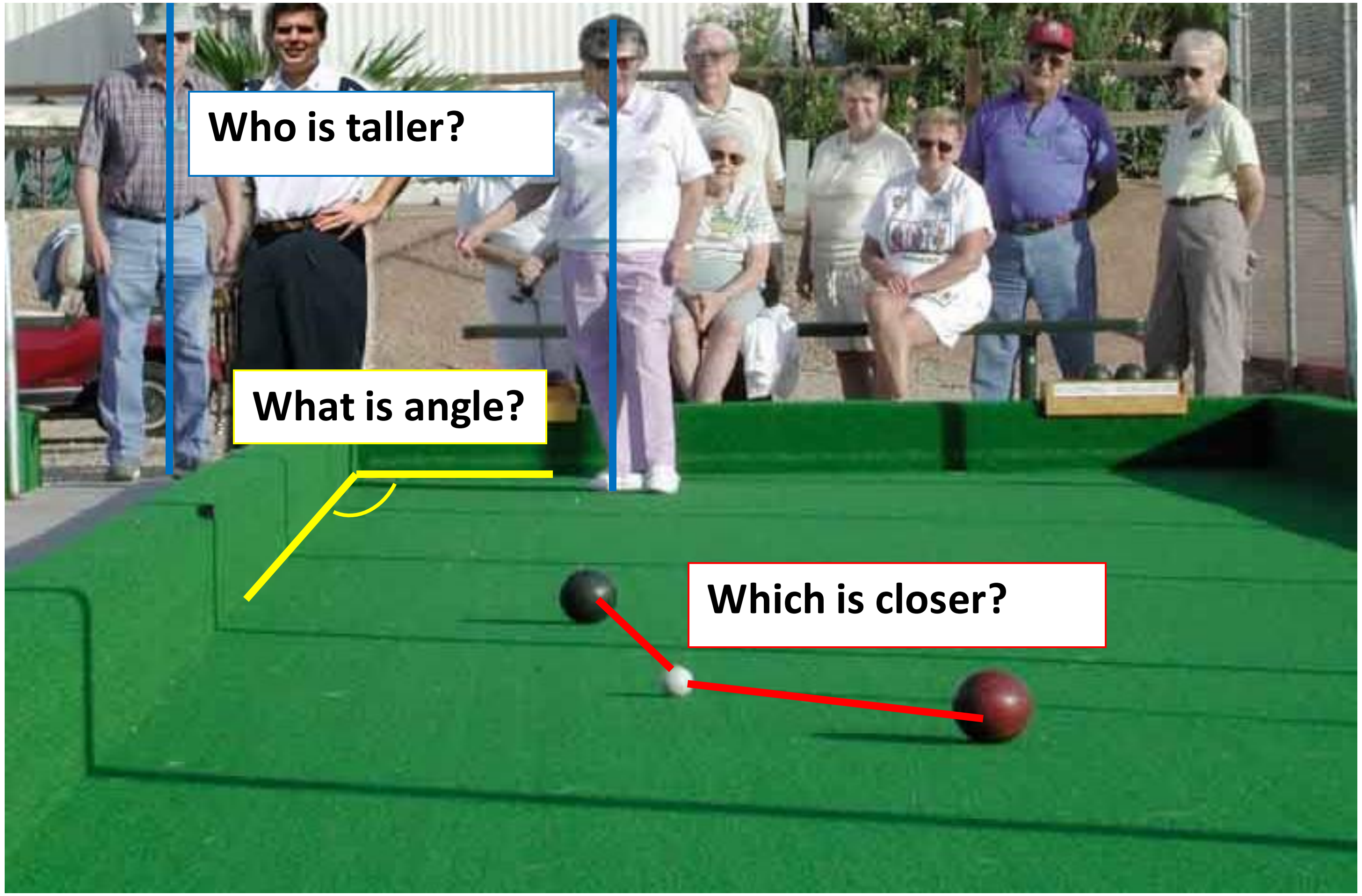


Projection can be tricky…

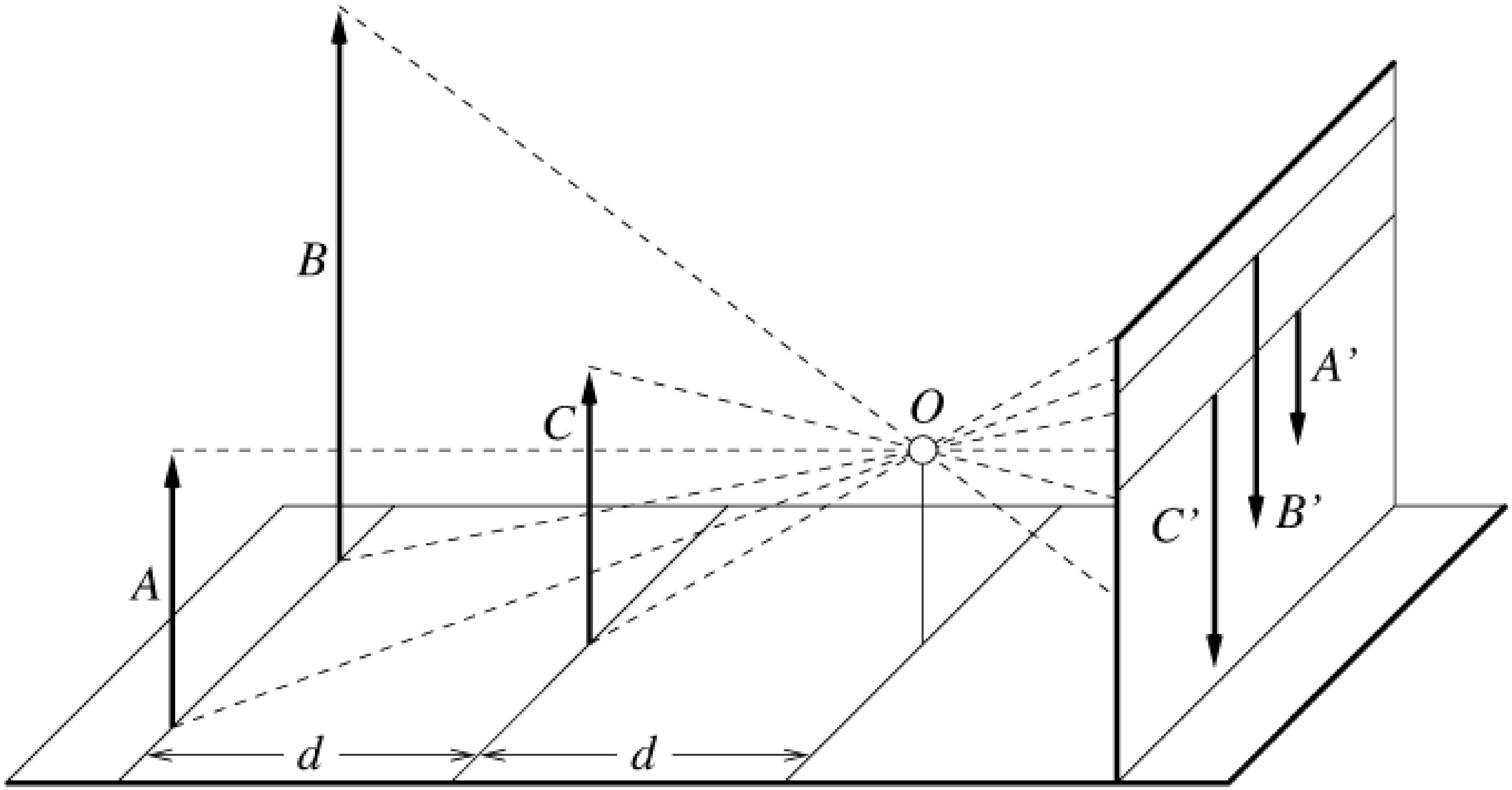
Projective Geometry

What is lost?

- Length
- Angle
- Parallelity



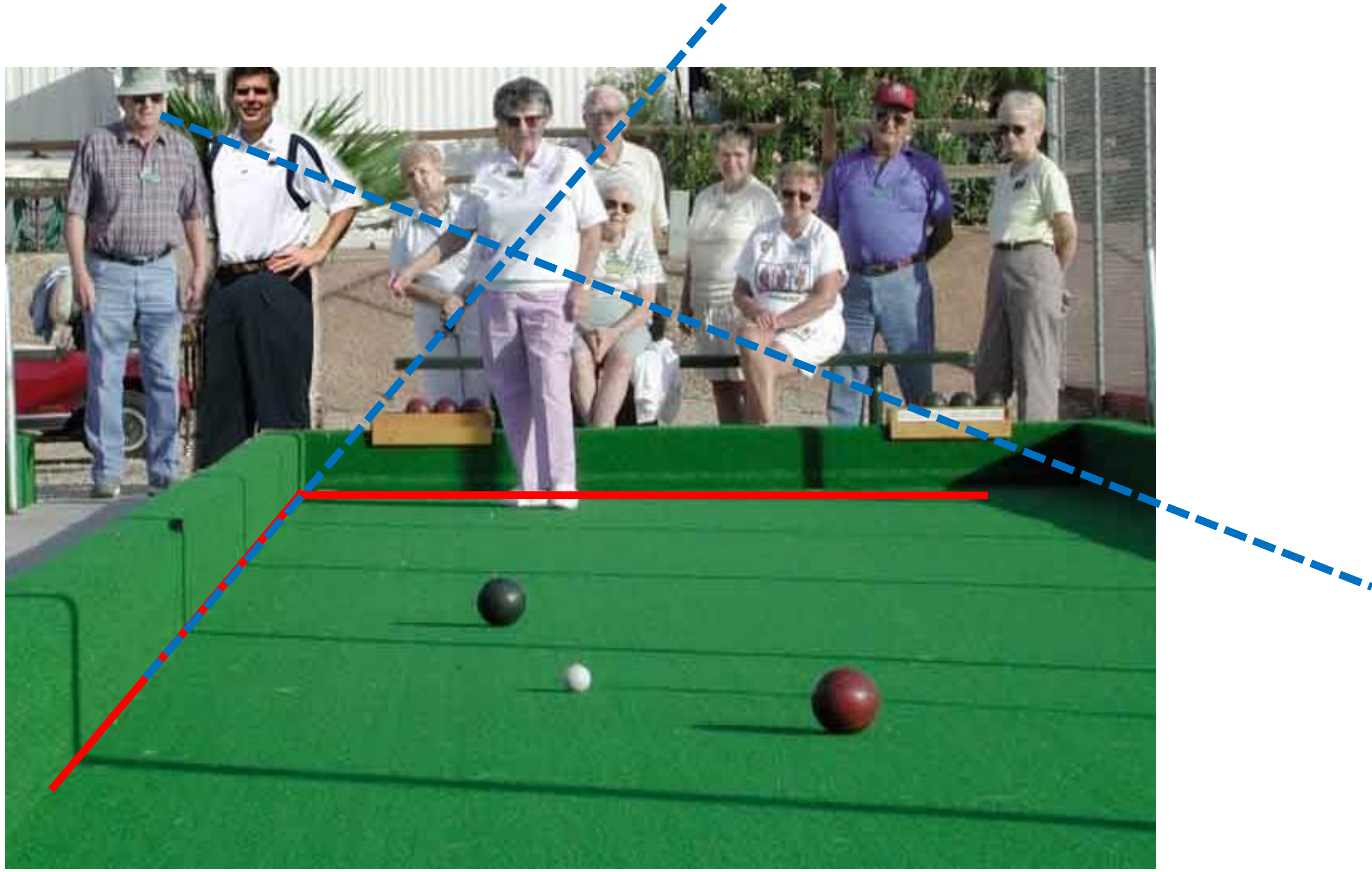
Length is not preserved



Projective Geometry

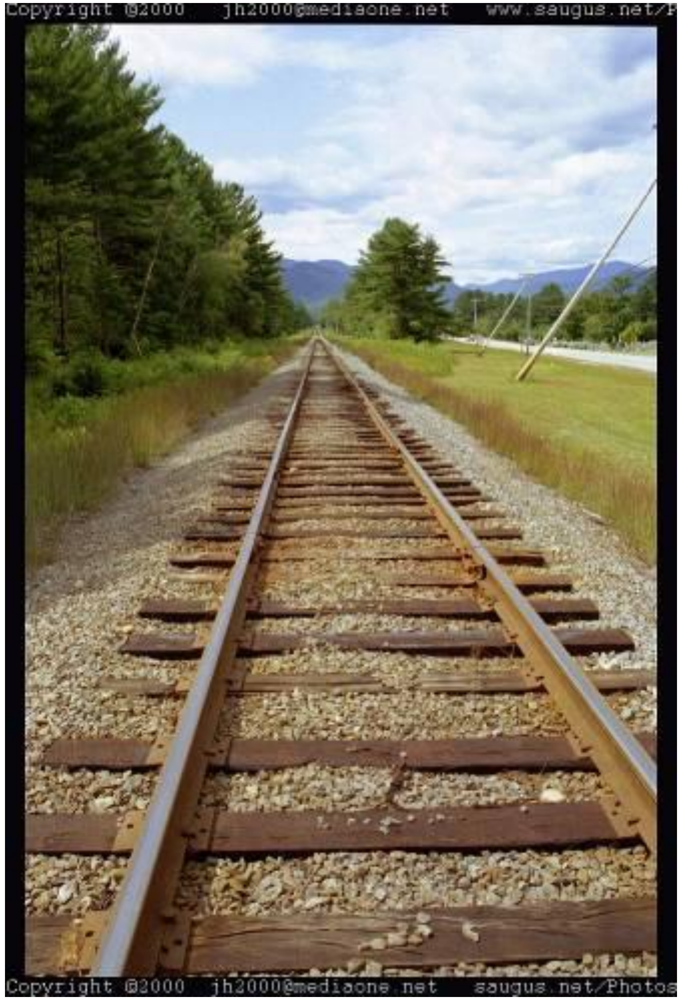
What is preserved?

- Straight lines are still straight

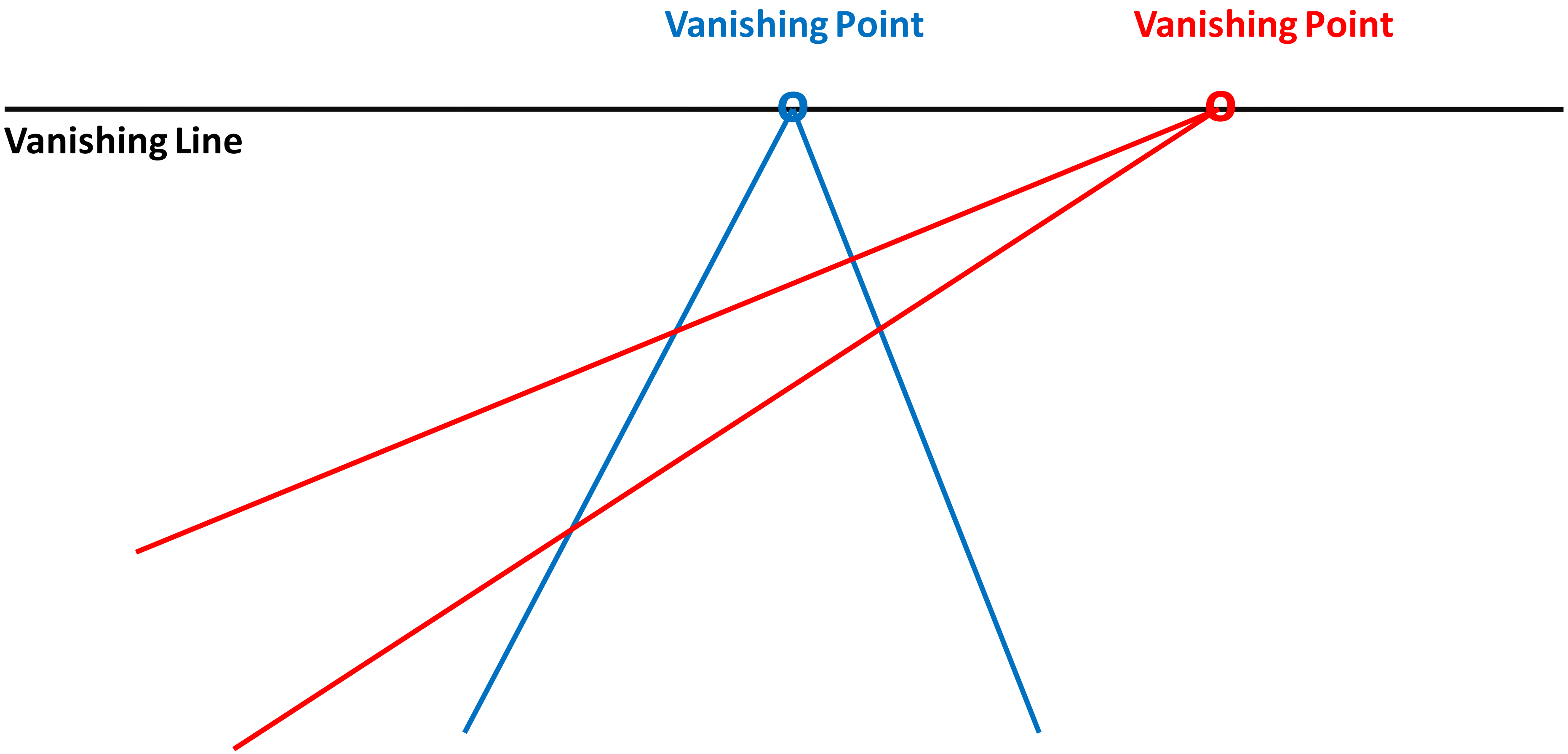


Vanishing points and lines

Parallel lines in the world intersect in the image at a "vanishing point"

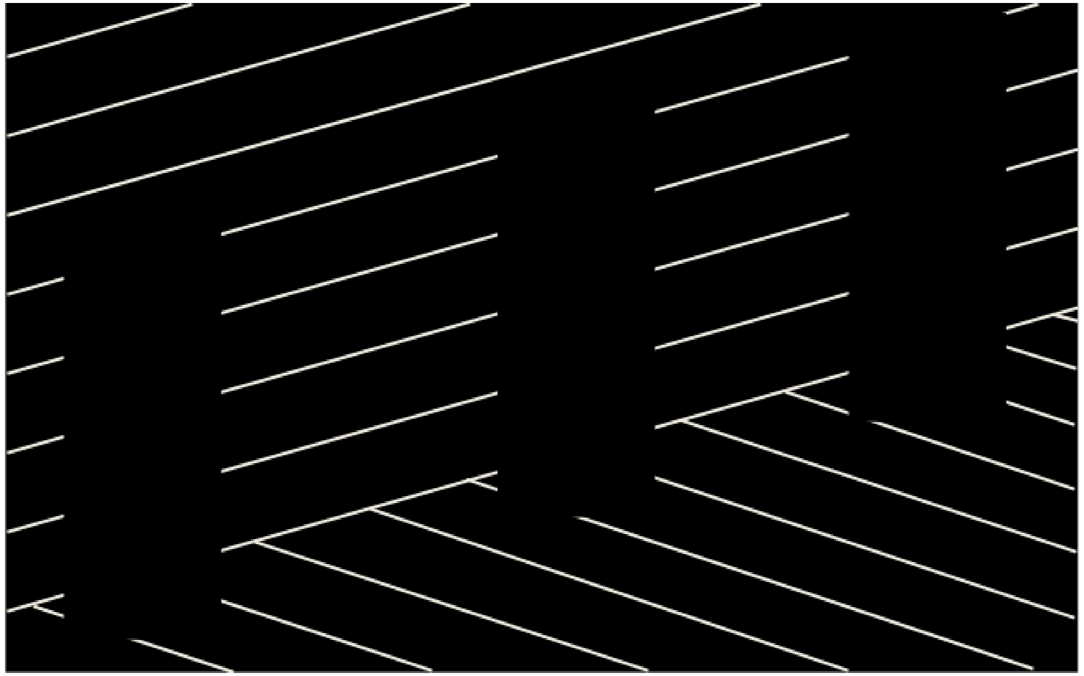


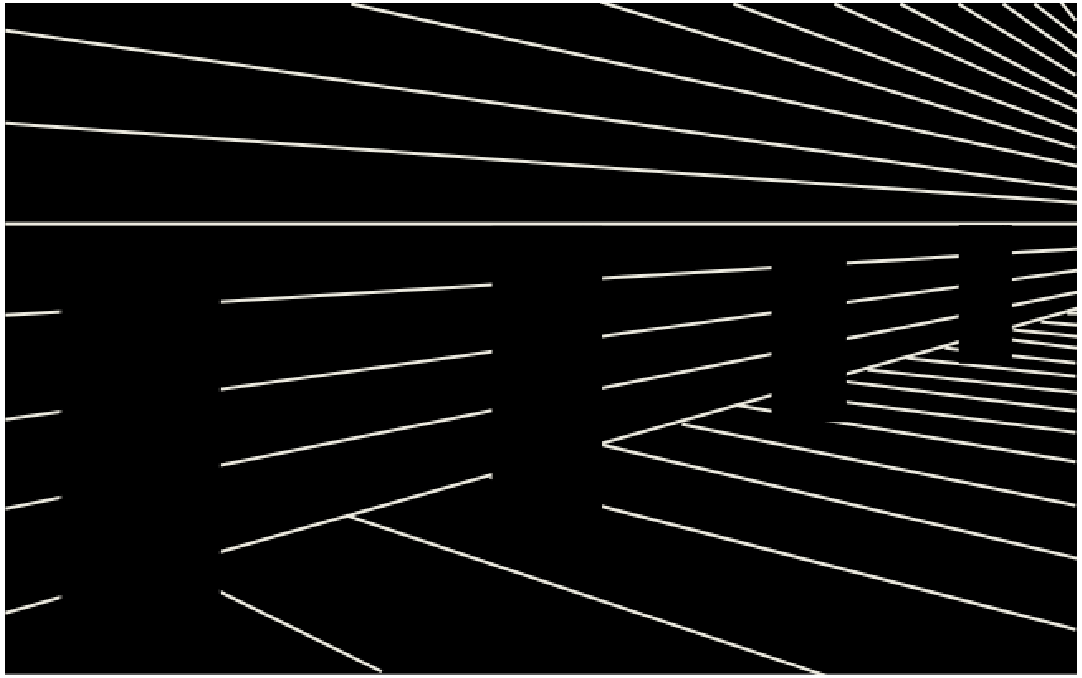
Vanishing points and lines



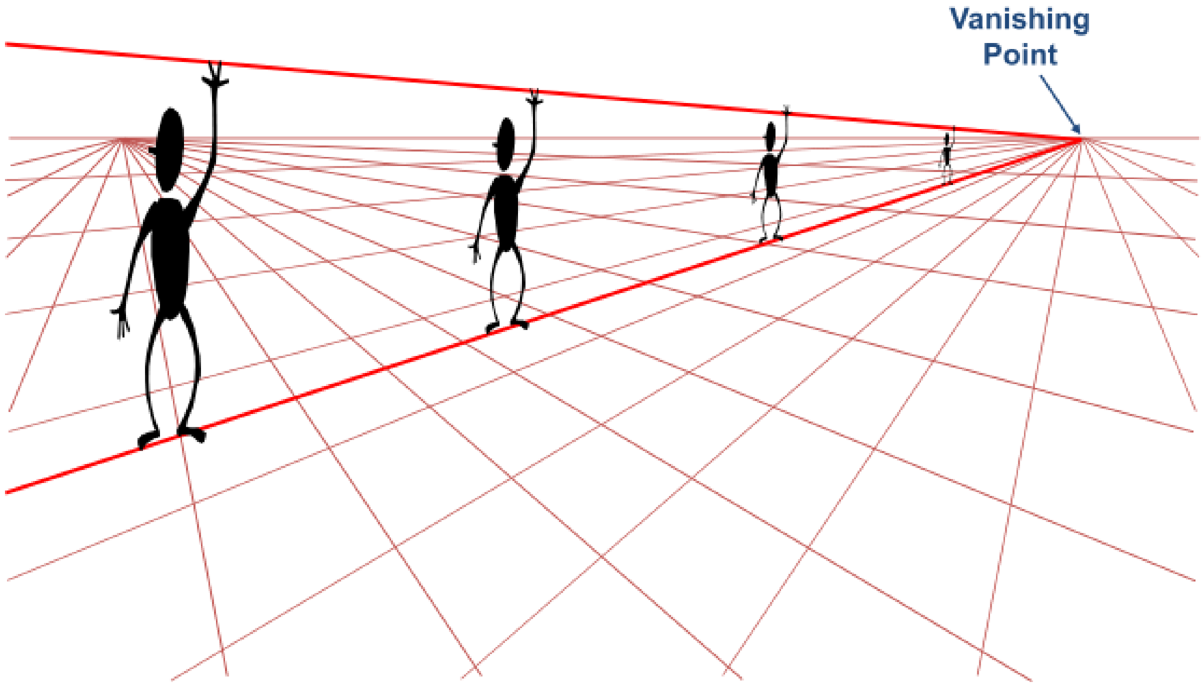
Vanishing points and lines

透视暗示

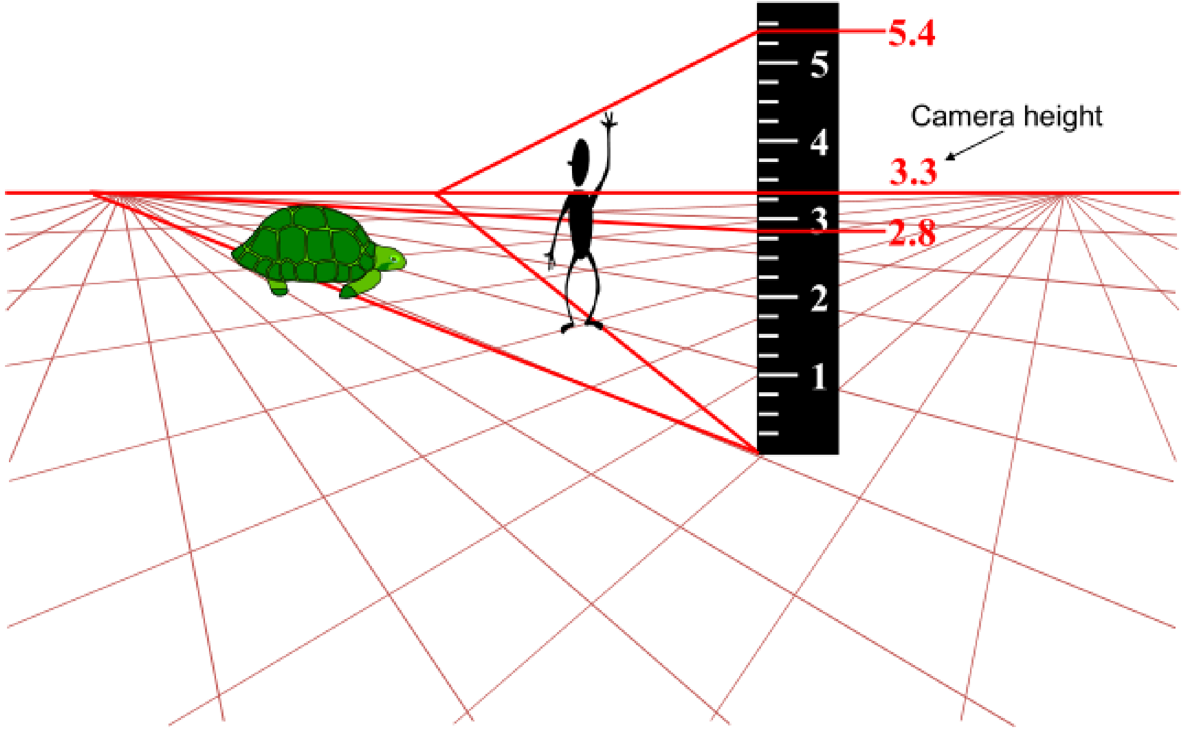




高度比较



高度测量

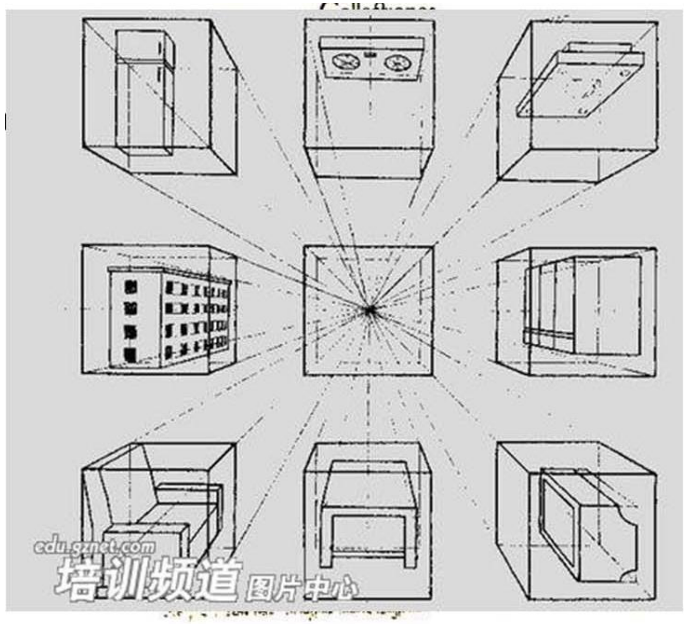


Contents

- **几何：**研究某个空间里的图形在变换之后保持不变性质的学科；
- **欧式几何：**研究在欧式变换下，保持不变的性质（欧式性质）的几何；
  - 旋转和平移都属于欧式变换
  - 欧式性质包括：长度、角度、平行性等

射影几何

- Pappus(290-350) 帕普斯，希腊数学家，提出交比、对合等概念；
- 在文艺复兴时期，人们在绘画和建筑艺术方面非常注意和大力研究如何在平面上表现实物的图形，产生透视法；
- Desargues(1591-1661) 笛沙格，法国工程师及建筑师，引入无穷远元素，透视定理，极点，极线等概念，创立**射影几何**

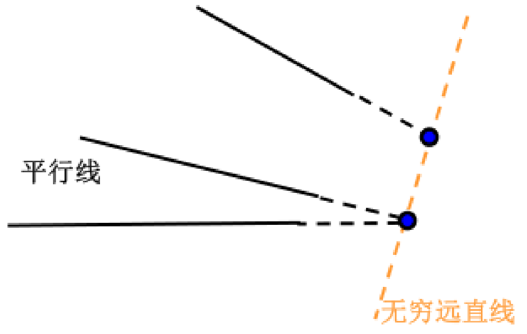


无穷远元素

- 平行线相交于一个无穷远点
- 平行平面相交于一条无穷远直线

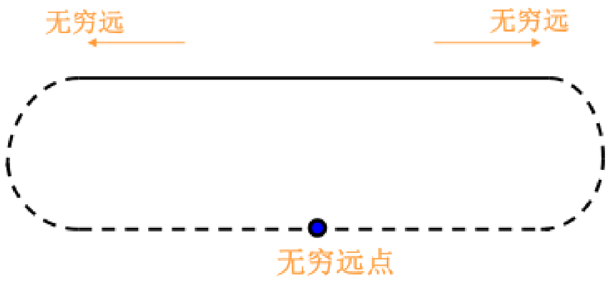
无穷远直线(lines at infinity)

- 在一个平面上，所有的无穷远点组成一条直线，称为该平面的无穷远直线



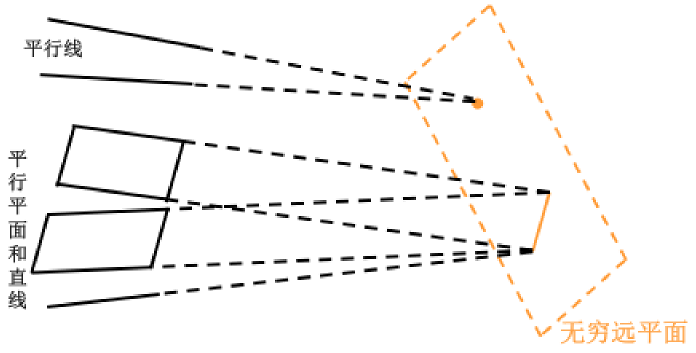
无穷远点(points at infinity)

- 在一条直线上只有唯一一个无穷远点
- 所有的一组平行线共有一个无穷远点



无穷远平面(planes at infinity)

- 三维空间中的所有无穷远点组成的一个平面，称为这个空间的无穷远平面



射影空间(Projective space)

- 对$\boldsymbol{n}$维欧氏空间加入无穷远元素，并对有限元素和无穷远元素不加区分，则它们共同构成$\boldsymbol{n}$维射影空间，记作$\mathbb{P}^n$
  - 一维射影空间是一条射影直线，由欧氏直线和它的无穷远点构成；
  - 二维射影空间是一个射影平面，由欧氏平面和它的无穷远直线构成；
  - 三维射影空间由我们所在空间与无穷远平面构成

齐次坐标

- 在欧氏空间中建立坐标系后，点与坐标是一一对应的
- 引入无穷远点后，无穷远点无欧氏坐标
- 为了刻化无穷远点的坐标，引入齐次坐标

齐次坐标与非齐次坐标

- 在n维空间中，建立欧氏坐标系后，每一个有限点的坐标为$(m_1, m_2, \dots, m_n)$，对任意$n+1$个数$x_1, x_2, \dots, x_n, x_0$，如果满足

$$x_0 \neq 0, \frac{x_1}{x_0} = m_1, \frac{x_2}{x_0} = m_2, \dots, \frac{x_n}{x_0} = m_n.$$

则$(x_1, x_2, \dots, x_n, x_0)$称为该点的齐次坐标。

- $(m_1, m_2, \dots, m_n)$称为非齐次坐标

齐次坐标与非齐次坐标

非齐次坐标到齐次坐标转换

$$(x, y) \Rightarrow \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \qquad (x, y, z) \Rightarrow \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

齐次2D(图像)坐标　　齐次3D(场景)坐标

齐次坐标到非齐次坐标转换

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} \Rightarrow \left(\frac{x}{w}, \frac{y}{w}\right) \qquad \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \Rightarrow \left(\frac{x}{w}, \frac{y}{w}, \frac{z}{w}\right)$$

无穷远点齐次坐标

- 不全为0的数$x_1, x_2, \dots, x_n$组成的坐标

$$(x_1, x_2, \dots, x_n, 0)$$

称为无穷远点的齐次坐标

交比(Cross ratio)

- 对于共线的4个点P_1, P_2, P_3, P_4,比例

$$\frac{(\theta_1-\theta_3)(\theta_2-\theta_4)}{(\theta_2-\theta_3)(\theta_1-\theta_4)}$$

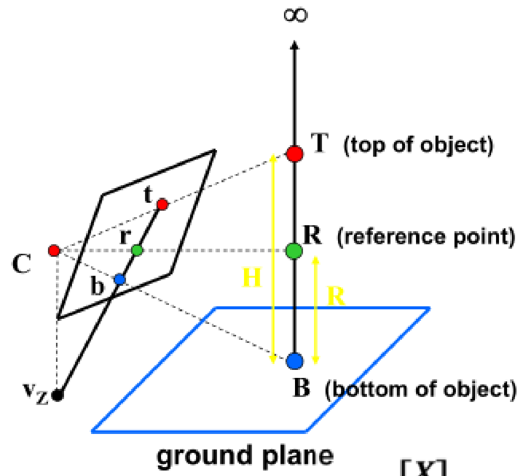
称为(P_3, P_4)关于(P_1, P_2)的交比，记作

$$(P_1, P_2; P_3, P_4)$$

其中θ_i为射影参数

- 设四个不同的共线点中的三点及其交比值为已知，则第四点必唯一确定

高度测量



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

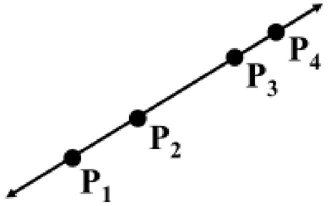
scene cross ratio

$$\frac{\|t-b\|\|v_Z-r\|}{\|r-b\|\|v_Z-t\|}=\frac{H}{R}$$

image cross ratio

场景中的点记为 $P=\begin{bmatrix}X\\Y\\Z\\1\end{bmatrix}$ 图像点记为 $p=\begin{bmatrix}x\\y\\1\end{bmatrix}$

交比



- 射影不变量
  - 在射影变换下保持不变的量
- 共线4点的交比

$$\frac{\|P_3-P_1\|\|P_4-P_2\|}{\|P_3-P_2\|\|P_4-P_1\|} \qquad P_i=\begin{bmatrix}X_i\\Y_i\\Z_i\\1\end{bmatrix}$$

- 可以交换上述点的顺序

$$\frac{\|P_1-P_3\|\|P_4-P_2\|}{\|P_1-P_2\|\|P_4-P_3\|}$$

  - 4！ = 24不同的顺序（只有6个不同的值）
- 交比是射影几何中的基本不变量

射影变换

- 记$\mathbb{P}^n, \mathbb{P}^{n\prime}$是两个由点构成的射影空间，$\boldsymbol{T}$是由$\mathbb{P}^n$到$\mathbb{P}^{n\prime}$的映射。如果$\boldsymbol{T}$保持：

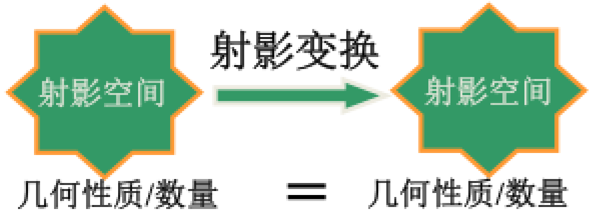
1）点和线的结合关系。如：点在直线上、直线通过点等；

2）共线的4个点的交比

则$\boldsymbol{T}$称为$\boldsymbol{n}$维射影变换

射影几何

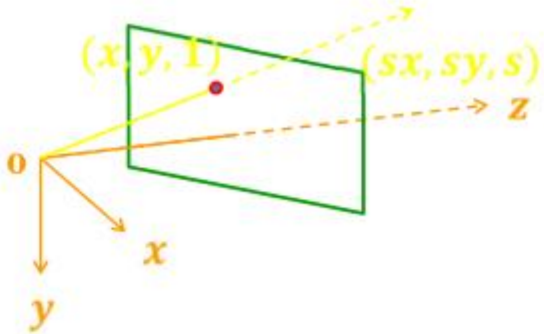
- 射影几何：研究射影空间中，在射影变换下保持不变的性质的几何学



点

- 射影平面$\mathbb{P}^2$中的点$(\boldsymbol{x},\boldsymbol{y})$在欧氏空间$\mathbb{R}^3$中可以看作一条从原点出发的射线$(\boldsymbol{sx},\boldsymbol{sy},\boldsymbol{s})$
  – 射线上的所有点是等价的：

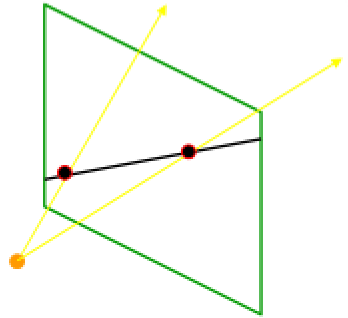
$$(\boldsymbol{x},\boldsymbol{y},\mathbf{1}) \cong (\boldsymbol{sx},\boldsymbol{sy},\boldsymbol{s})$$



- 射影几何中的2D点如何表示？ $\hat{\mathbf{x}} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$

直线

- 射影平面$\mathbb{P}^2$中的直线在欧氏空间$\mathbb{R}^3$中看作是由许多从原点出发的射线所构成的平面
  – 每一条射线$(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})$满足：$\boldsymbol{ax}+\boldsymbol{by}+\boldsymbol{cz}=\mathbf{0}$

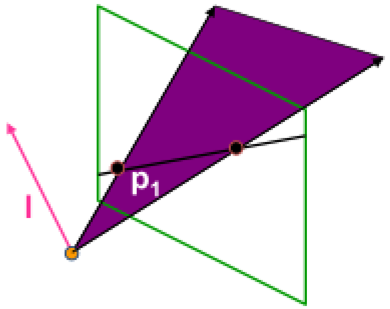


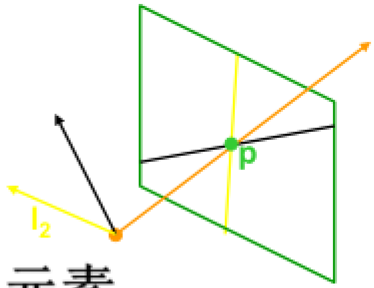
- 射影几何中的线如何表示？ $\mathbf{l} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ ，这其中隐含的表示是一个直线方程

$$ax+by+c=0$$

点线对偶

- 直线$\boldsymbol{l}$用三维齐次坐标表示
- $\boldsymbol{l}^T\boldsymbol{p} = 0 \Rightarrow \boldsymbol{l} \perp \boldsymbol{p}$
- 由$\boldsymbol{p}_1, \boldsymbol{p}_2$构成的直线$\boldsymbol{l}$是？
  - $\boldsymbol{l} \perp \boldsymbol{p}_1, \boldsymbol{l} \perp \boldsymbol{p}_2 \Rightarrow \boldsymbol{l} = \boldsymbol{p}_1 \times \boldsymbol{p}_2$
  - $\boldsymbol{l}$是平面的法向
- 直线$\boldsymbol{l}_1, \boldsymbol{l}_2$的交点$\boldsymbol{p}$是？
  - $\boldsymbol{p} \perp \boldsymbol{l}_1, \boldsymbol{p} \perp \boldsymbol{l}_2 \Rightarrow \boldsymbol{p} = \boldsymbol{l}_1 \times \boldsymbol{l}_2$
- 点与直线称为射影平面的对偶元素





无穷远点

- 平行线的交点
  - 两平行直线$\boldsymbol{l} = (\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})^T, \boldsymbol{l}' = (\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}')^T$，两直线的交点为$\boldsymbol{l} \times \boldsymbol{l}' = (\boldsymbol{c}' - \boldsymbol{c})(\boldsymbol{b}, -\boldsymbol{a}, \boldsymbol{0})^T$，忽略尺度因子$(\boldsymbol{c}' - \boldsymbol{c})$，得到点$(\boldsymbol{b}, -\boldsymbol{a}, \boldsymbol{0})^T$
- 齐次坐标为$\boldsymbol{x}_\propto = (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0})^T$的点称为无穷远点
- 无穷远点没有欧式坐标

无穷远直线

- 平面上所有无穷远点构成的集合
- 由于所有无穷远点$\boldsymbol{x}_\propto$满足方程：

$$\boldsymbol{0} \cdot \boldsymbol{x} + \boldsymbol{0} \cdot \boldsymbol{y} + \boldsymbol{1} \cdot \boldsymbol{0} = \boldsymbol{0}$$

故无穷远直线$\boldsymbol{l}_\propto = (\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{1})^T$

二维射影变换

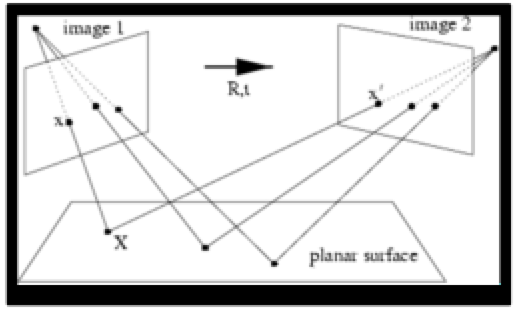
- 射影变换是射影平面上的可逆齐次线性变换，这个变换可用一个非奇异$\mathbf{3\times3}$矩阵$\boldsymbol{H}$表示为：

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

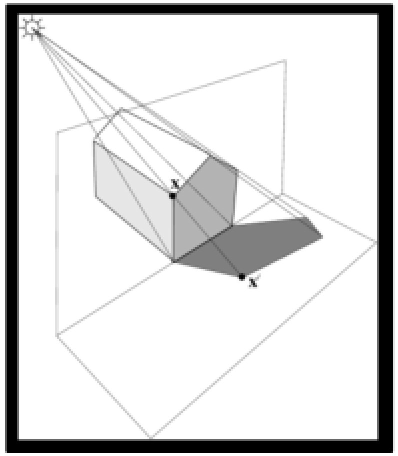
或记为$\boldsymbol{x}' = \boldsymbol{H}\boldsymbol{x}$

- 射影变换又称为单应(**Homography**)，矩阵$\boldsymbol{H}$称为射影变换矩阵或单应矩阵
- 射影变换有**8**个自由度

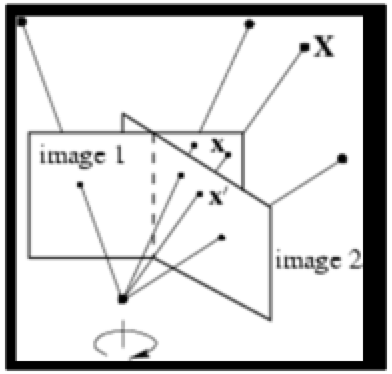
二维射影变换



两射影变换的复合



图像和它的阴影

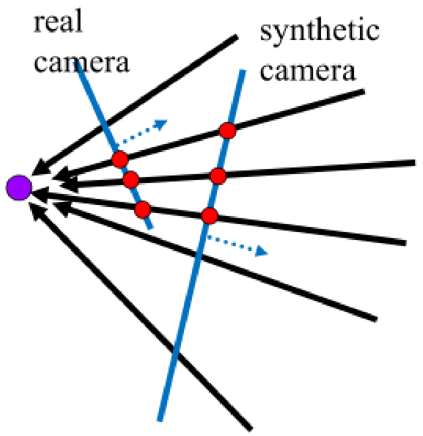


射影中心相同的两幅图像

求解单应矩阵

- 由于**H**相差一个常数因子，故只有**8**个自由度
  - 一对对应点可以提供**2**个约束方程，**4**对对应点可以确定**H**
- 如果$n > 4$，求解最小二乘问题：
$$minimize\|Ah - 0\|^2$$
  - $A^T A$的最小特征值对应的特征向量即为解

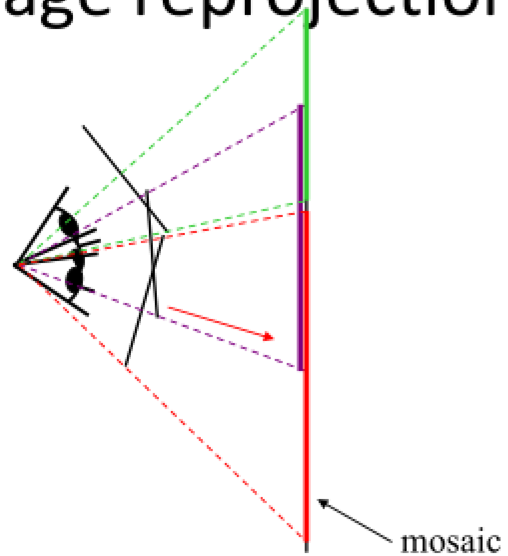
Mosaics: 生成合成视角（synthetic views）



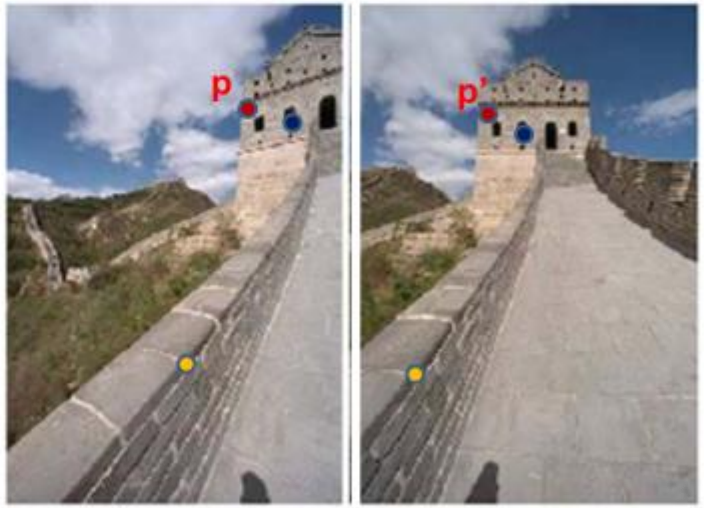
射影变换

只要是同一个光心的投影，就可以合成任意视角图像

Image reprojection

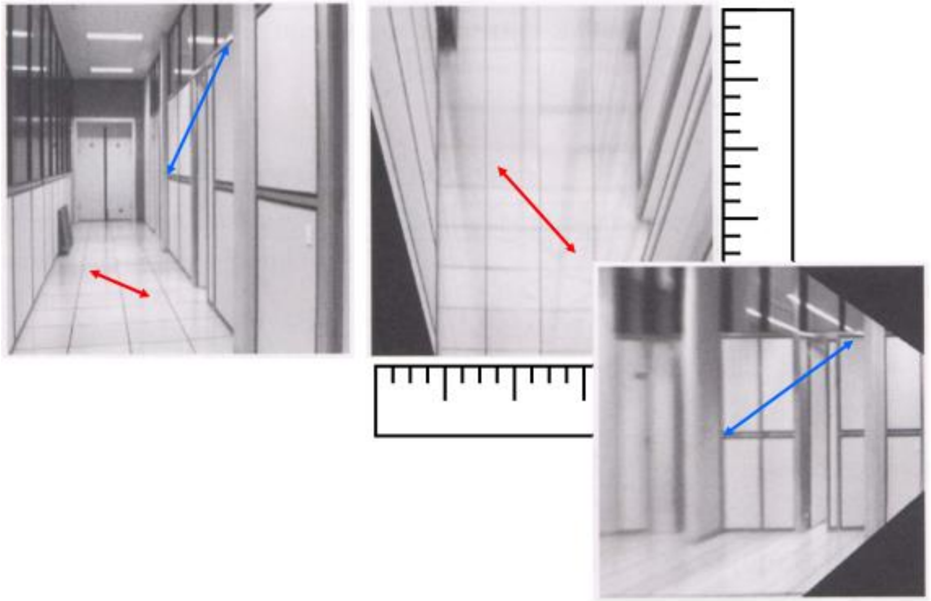


Homography

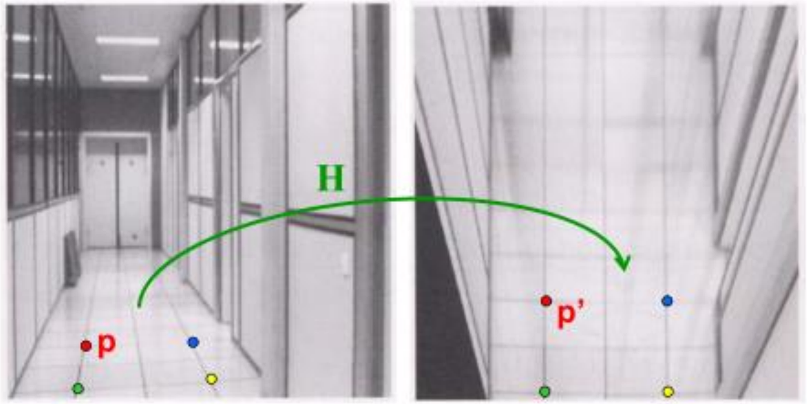


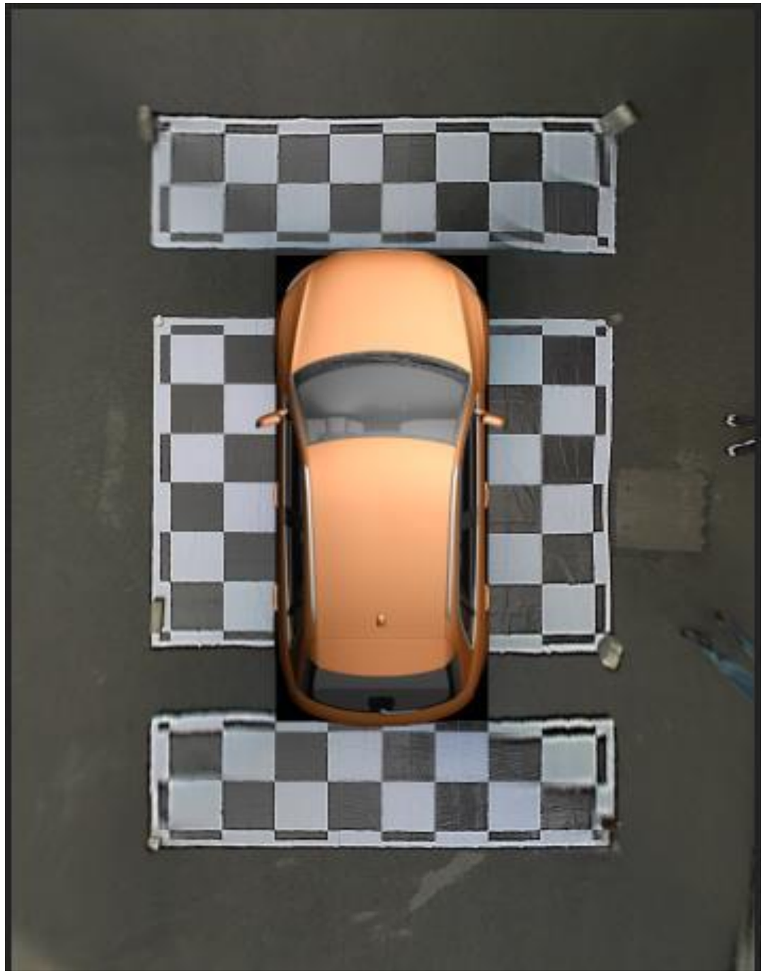
Mosaics

平面测量



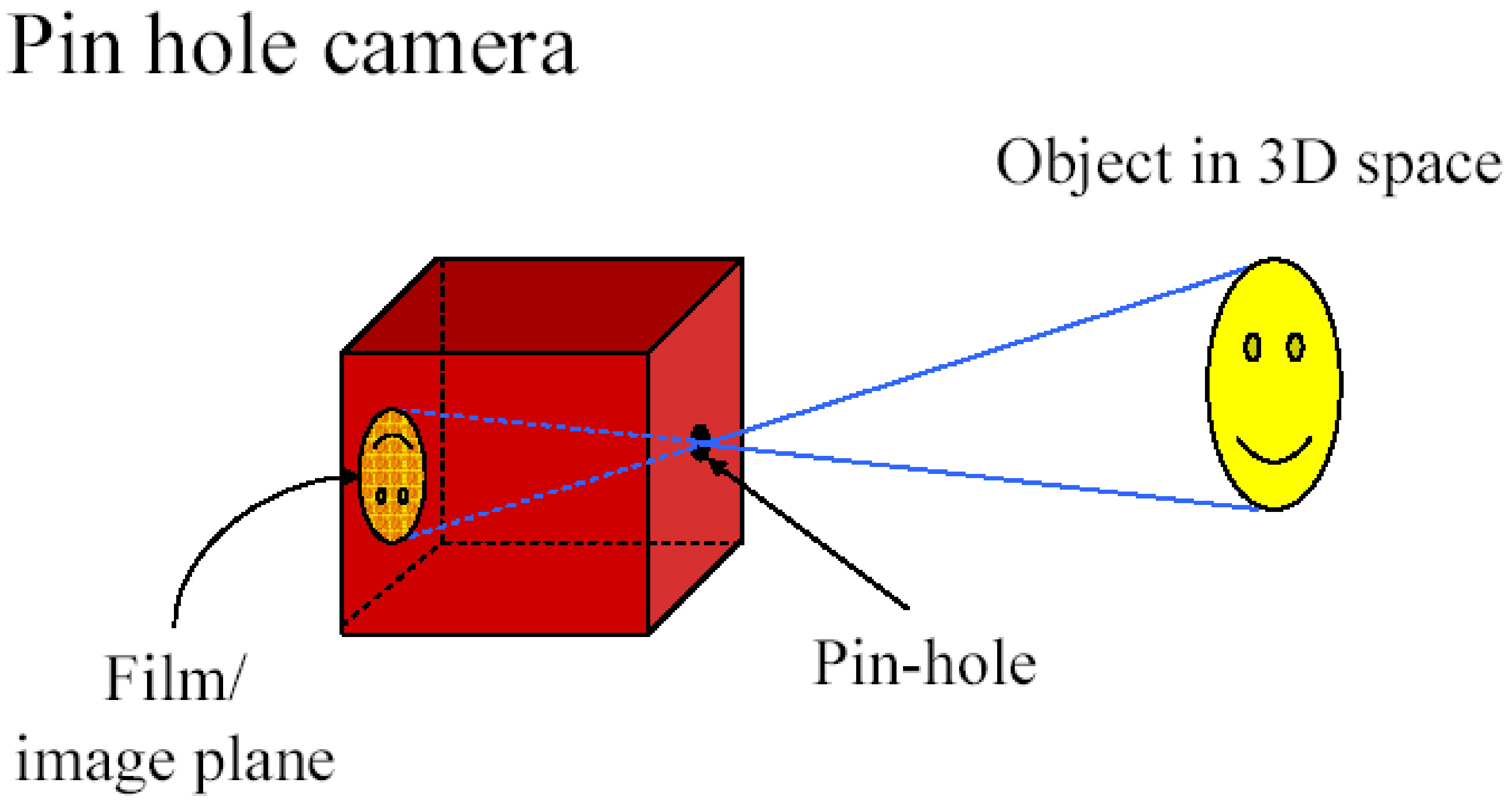
图像校正



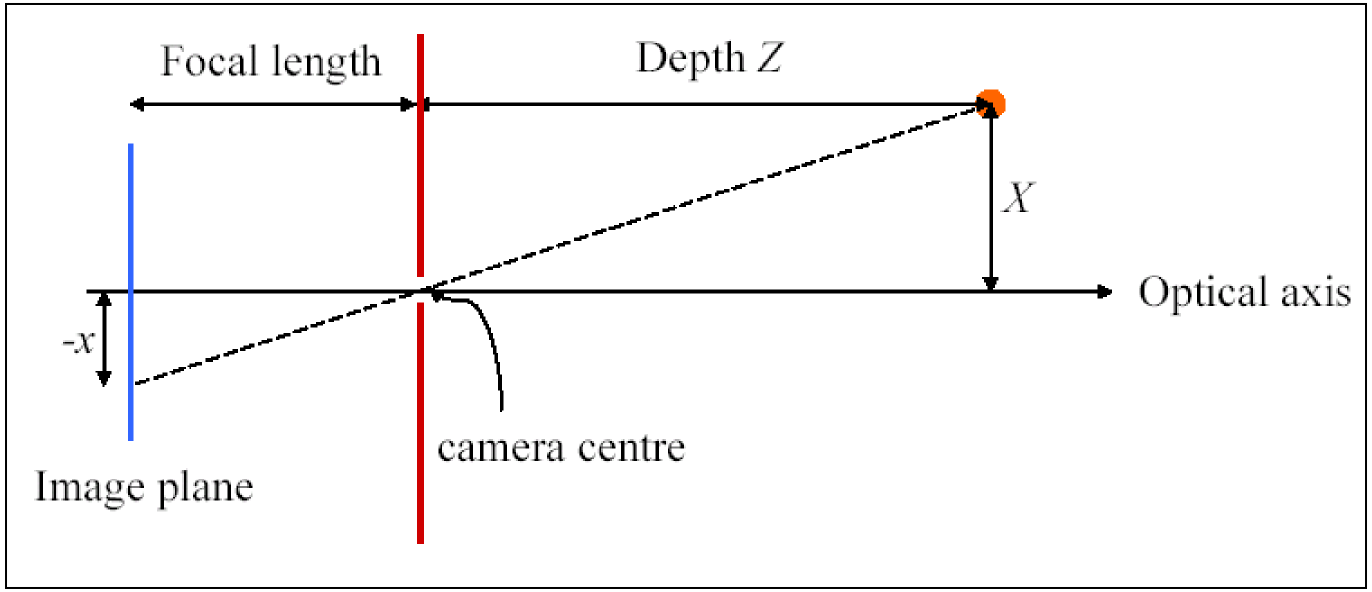


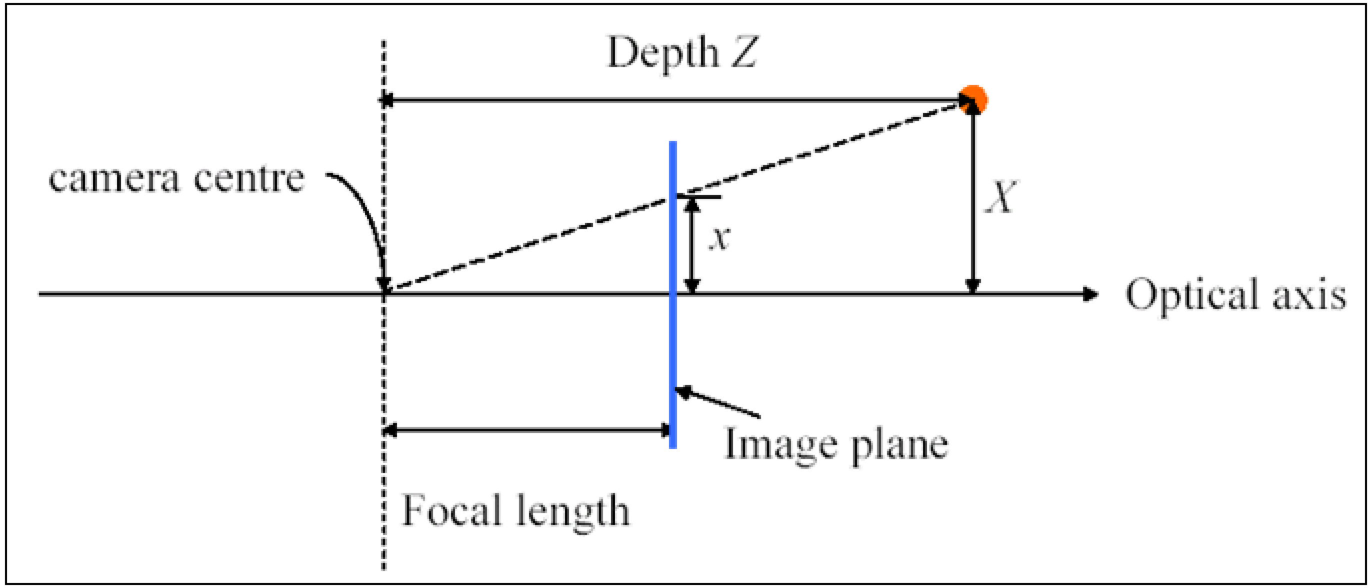
Contents

Camera models



Camera models

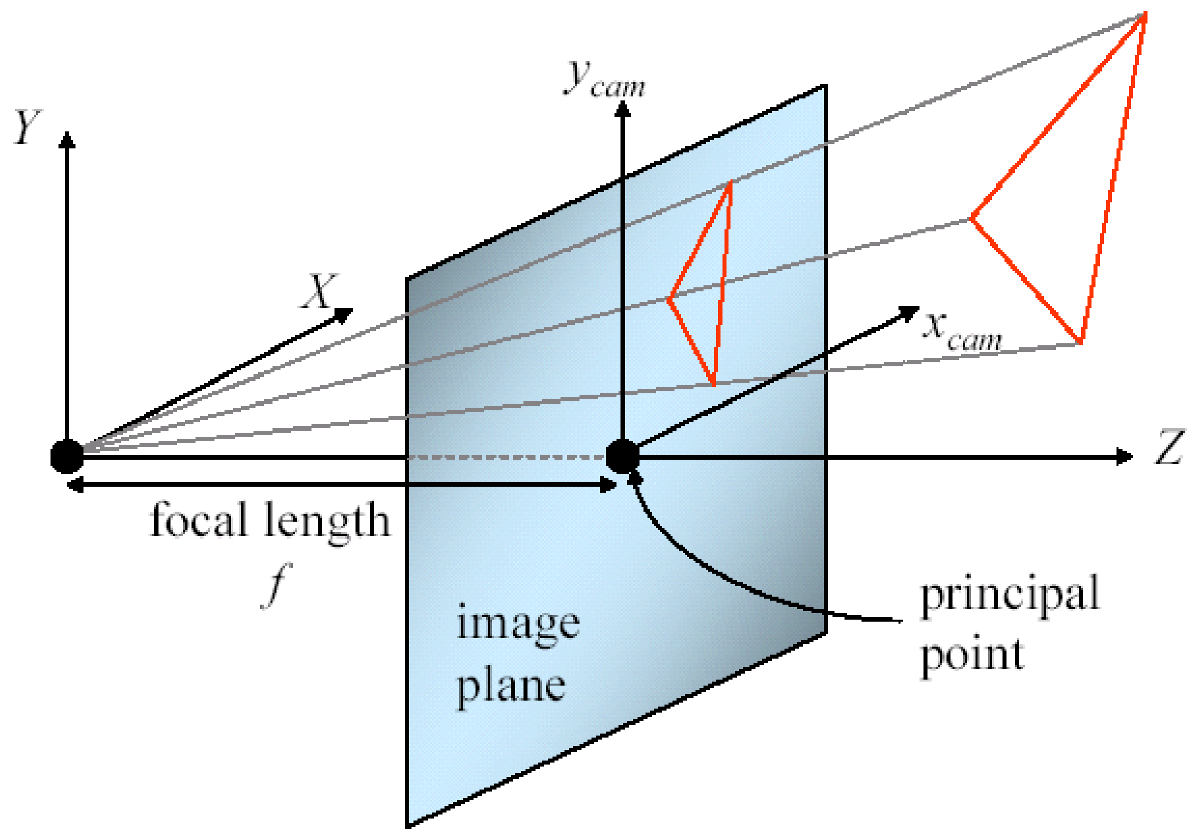




$$x = \frac{f}{Z} X$$

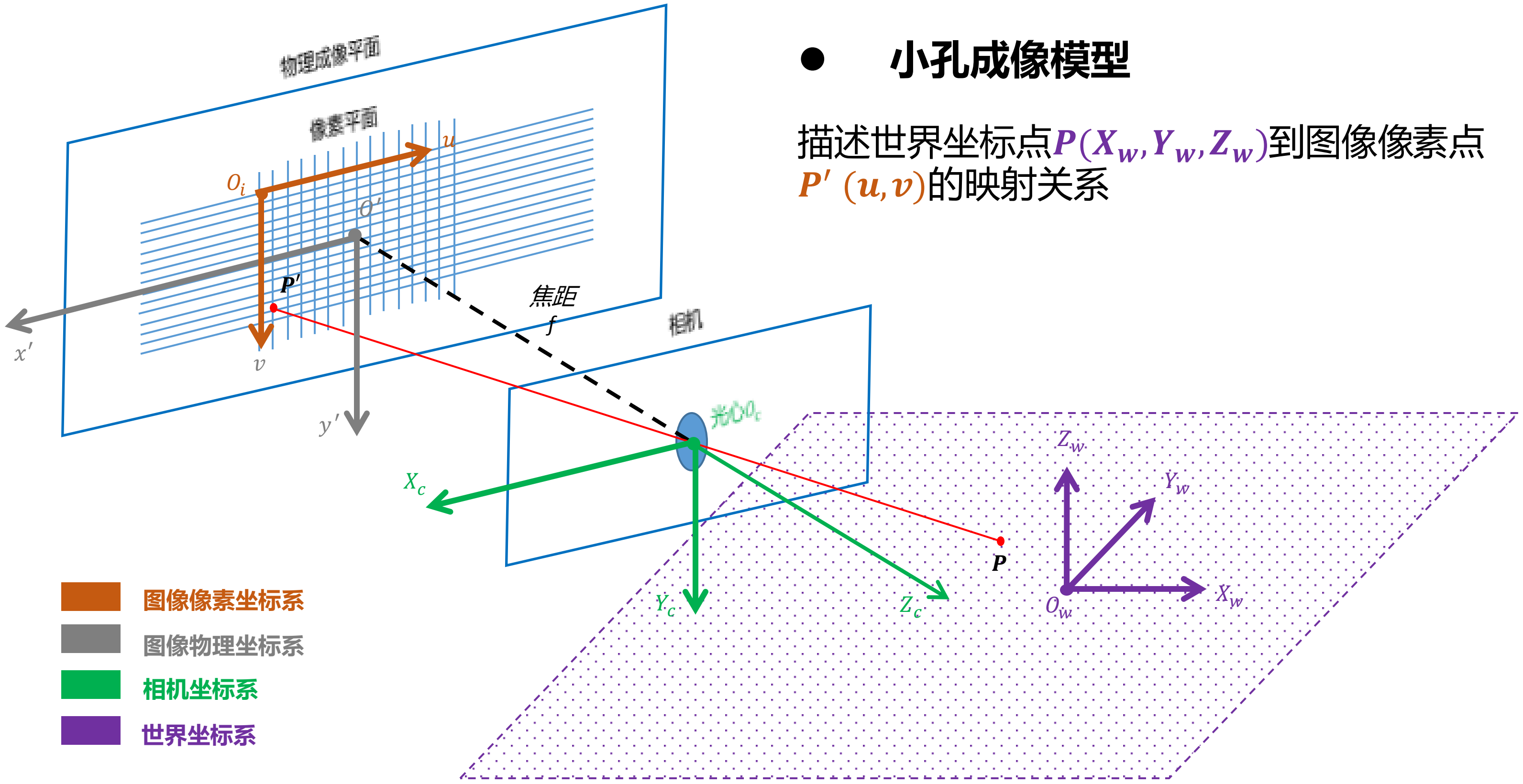
Camera models

We'll look at this in 3D



$$x_{cam} = \frac{f}{Z} X \qquad y_{cam} = \frac{f}{Z} Y$$

Camera models



- **小孔成像模型**

描述世界坐标点$P(X_w,Y_w,Z_w)$到图像像素点$P'\ (u,v)$的映射关系

世界坐标系 (X_w,Y_w,Z_w) → 刚体变换 → **相机坐标系** (X_c,Y_c,Z_c) → 相似变换 → **图像物理坐标系** (x',y') → 缩放变换 → **图像像素坐标系** (u,v)

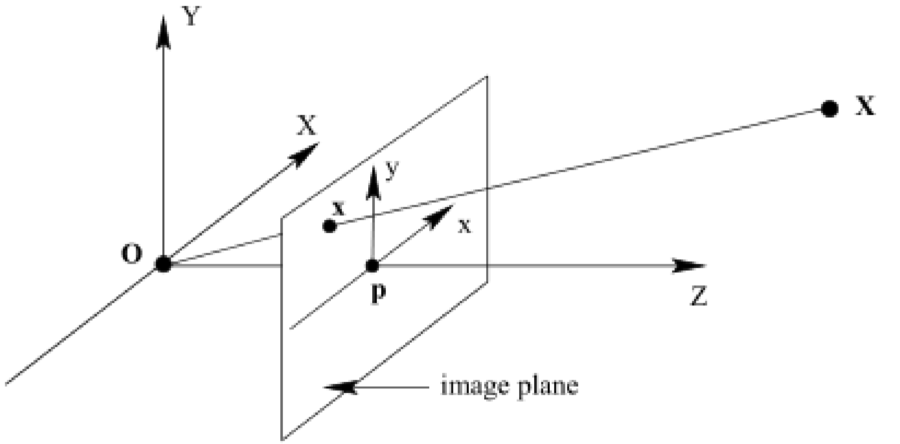
$$Z_C\begin{bmatrix}u\\v\\1\end{bmatrix}=\begin{bmatrix}\alpha_x & 0 & u_0 & 0\\ 0 & \alpha_y & v_0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}R & t\\ 0 & 1\end{bmatrix}\begin{bmatrix}X_w\\Y_w\\Z_w\\1\end{bmatrix}=P\begin{bmatrix}X_w\\Y_w\\Z_w\\1\end{bmatrix}$$

Imaging Geometry

Perspective projection

$$\lambda \begin{pmatrix} x \\ y \\ f \end{pmatrix} = \begin{pmatrix} \mathrm{X} \\ \mathrm{Y} \\ \mathrm{Z} \end{pmatrix}$$

where $\lambda = \mathrm{Z}/f$.



This can be written as a linear mapping between homogeneous coordinates (the equation is only up to a scale factor):

$$\begin{pmatrix} x \\ y \\ f \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} \mathrm{X} \\ \mathrm{Y} \\ \mathrm{Z} \\ 1 \end{pmatrix}$$

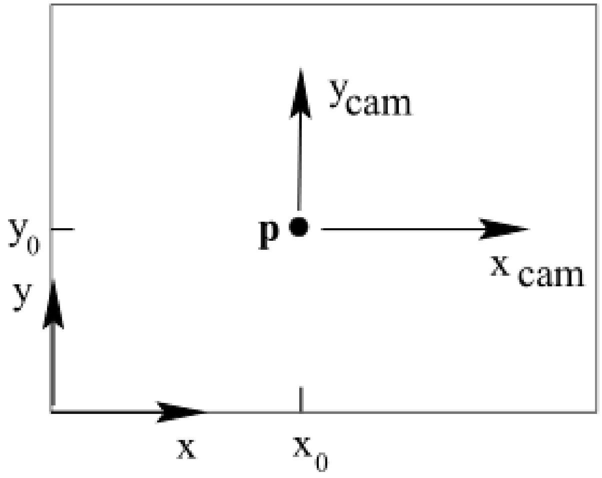
where a 3×4 projection matrix represents a map from 3D to 2D.

Image Coordinate System

Internal camera parameters

$$k_x\, x_{\text{cam}} = x - x_0$$
$$k_y\, y_{\text{cam}} = y - y_0$$

where the units of k_x, k_y are $[\text{pixels}/\text{length}]$.

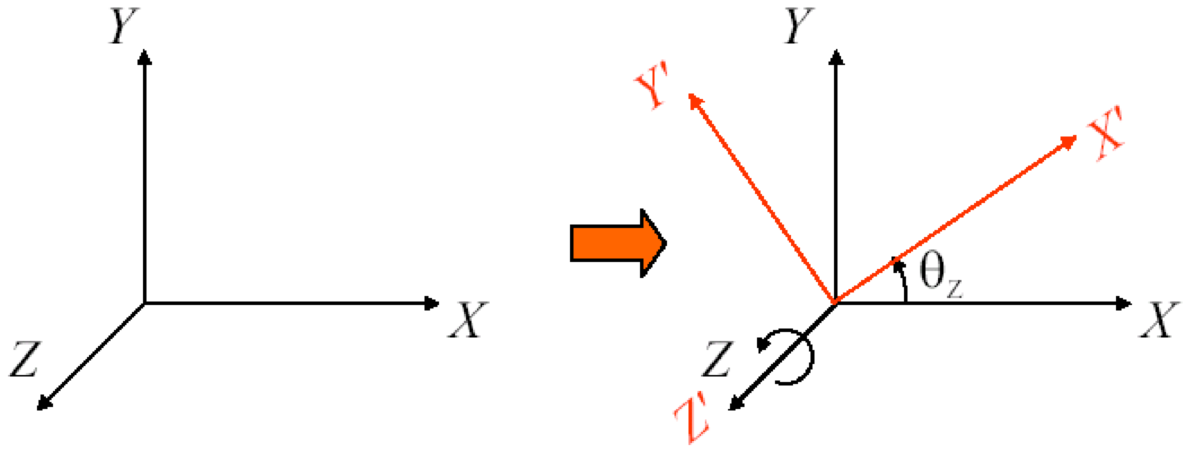


$$\mathbf{x} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \frac{1}{f}\begin{bmatrix} \alpha_x & & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}\begin{pmatrix} x_{\text{cam}} \\ y_{\text{cam}} \\ f \end{pmatrix} = \mathrm{K}\begin{pmatrix} x_{\text{cam}} \\ y_{\text{cam}} \\ f \end{pmatrix}$$

where $\alpha_x = f k_x,\ \alpha_y = f k_y$.

World Coordinate System

Rotating about the Z-axis in 3D ...



$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \cos\theta_Z & -\sin\theta_Z & 0 \\ \sin\theta_Z & \cos\theta_Z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix}$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \cos\theta_Z & -\sin\theta_Z & 0 \\ \sin\theta_Z & \cos\theta_Z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix}$$

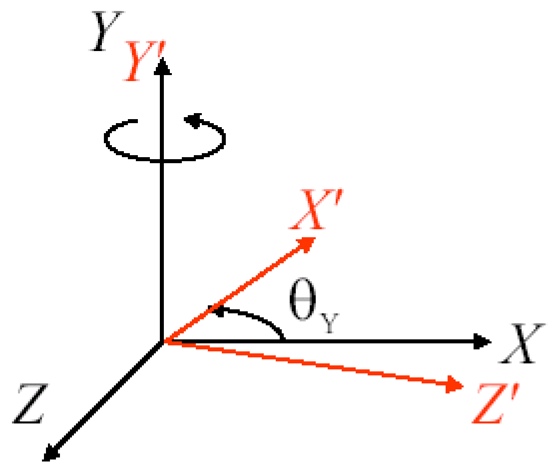
$$\mathbf{X} = \mathbf{R}_Z(\theta_Z)\,\mathbf{X}'$$

World Coordinate System

Can also rotate around the Y axis ...

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \cos\theta_Y & 0 & \sin\theta_Y \\ 0 & 1 & 0 \\ -\sin\theta_Y & 0 & \cos\theta_Y \end{bmatrix} \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix}$$

$$\mathbf{X} = \mathbf{R}_Y(\theta_Y)\,\mathbf{X}'$$

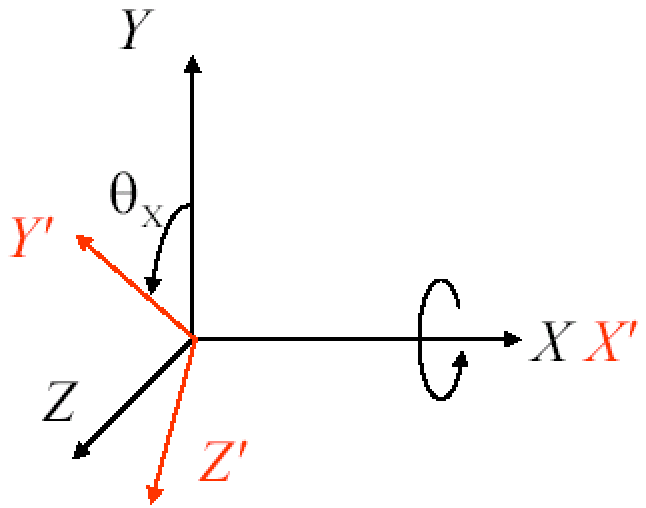


World Coordinate System

or the X axis ...

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_X & -\sin\theta_X \\ 0 & \sin\theta_X & \cos\theta_X \end{bmatrix} \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix}$$

$$\mathbf{X} = \mathbf{R}_X(\theta_X)\,\mathbf{X}'$$



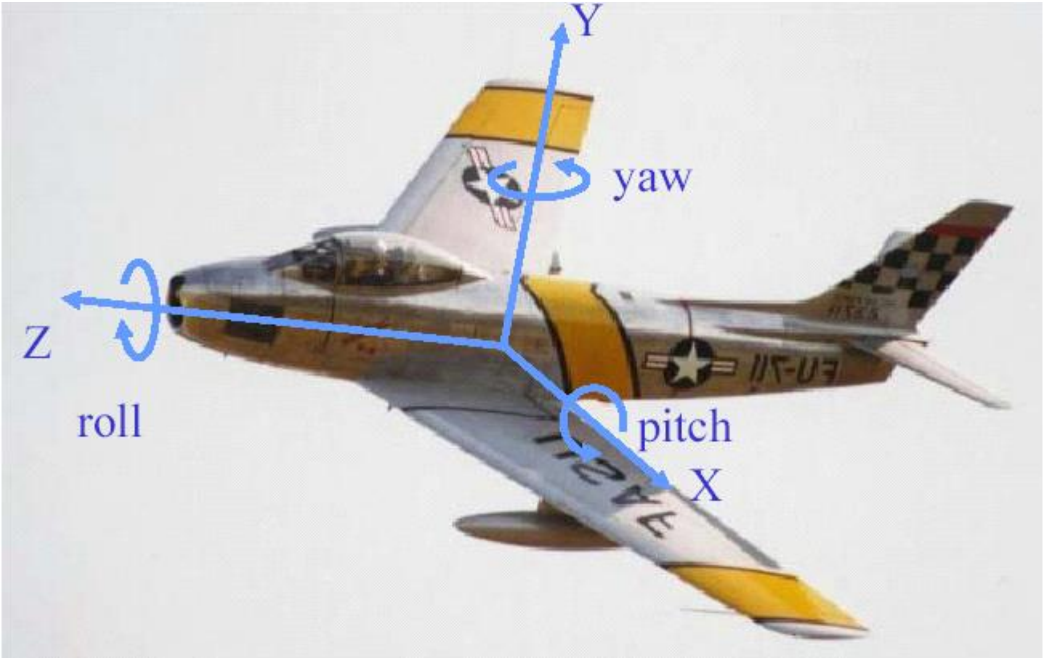
World Coordinate System

Arbitrary rotation between the camera and WCS …

- Elementary rotations are rotations about the current axes.
- The 3 elementary rotations in 3D are
  - $\mathbf{R}_Z(\theta_Z)$ rotation about the Z-axis
  - $\mathbf{R}_Y(\theta_Y)$ rotation about the Y-axis
  - $\mathbf{R}_X(\theta_X)$ rotation about the X-axis

World Coordinate System

Arbitrary rotation between the camera and WCS …



- Roll, Pitch, Yaw (RPY) angles describe rotation with respect to a reference frame *fixed to an object*.

- $\mathbf{R}(\theta_X, \theta_Y, \theta_Z) = \mathbf{R}_Z(\theta_Z)\ \mathbf{R}_Y(\theta_Y)\ \mathbf{R}_X(\theta_X)$
  - **1)** $\mathbf{R}_Z(\theta_Z)$ is a rotation about the Z-axis
  - **2)** $\mathbf{R}_Y(\theta_Y)$ is a rotation about the Y-axis
  - **3)** $\mathbf{R}_X(\theta_X)$ is a rotation about the X-axis

World Coordinate System

Relative Rotation ...

$$\mathbf{X} = \mathbf{R_x}(\theta_x)\mathbf{R}_y(\theta_y)\mathbf{R}_z(\theta_z)\mathbf{X}^{'}$$

$$\mathbf{X} = \mathbf{R}\mathbf{X}^{'}$$

$$\mathbf{R}_X(\theta_X) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_X & -\sin\theta_X \\ 0 & \sin\theta_X & \cos\theta_X \end{bmatrix}$$

$$\mathbf{R}_Y(\theta_Y) = \begin{bmatrix} \cos\theta_Y & 0 & \sin\theta_Y \\ 0 & 1 & 0 \\ -\sin\theta_Y & 0 & \cos\theta_Y \end{bmatrix}$$

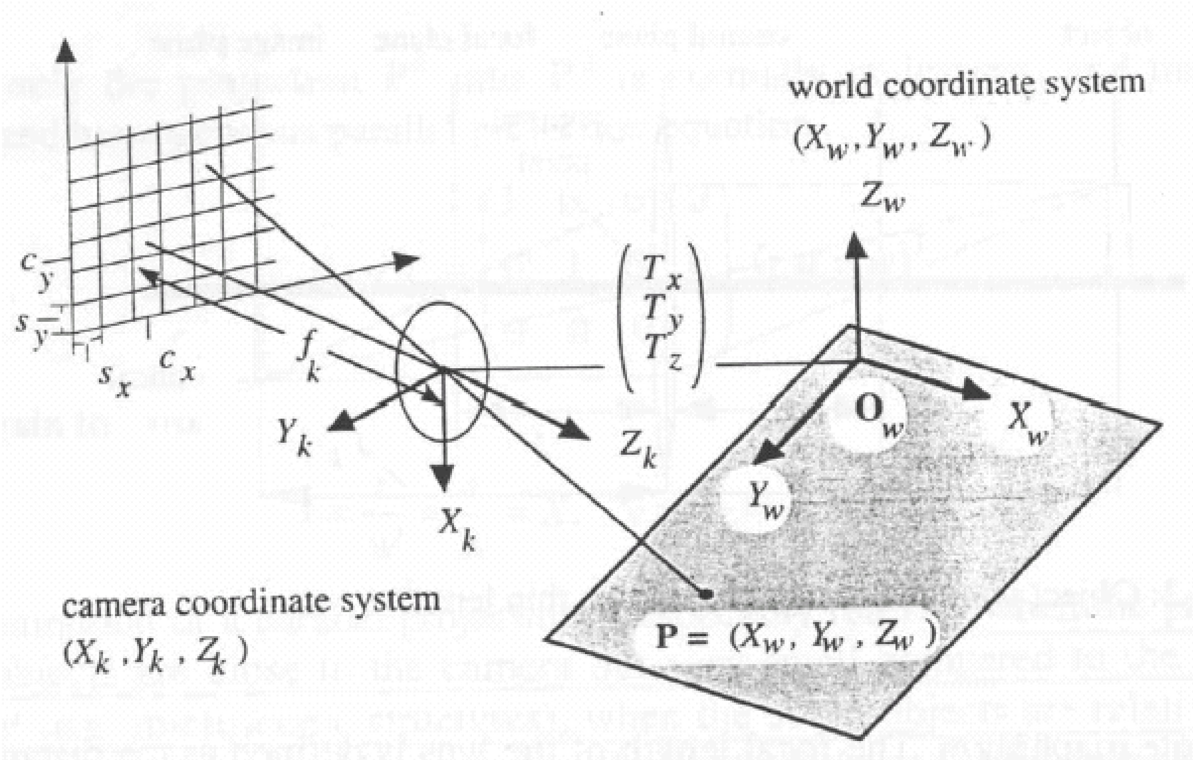
$$\mathbf{R}_Z(\theta_Z) = \begin{bmatrix} \cos\theta_Z & -\sin\theta_Z & 0 \\ \sin\theta_Z & \cos\theta_Z & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

World Coordinate System

Relative Translation ...

$$\mathbf{X} = \mathbf{X}' + \mathbf{T}$$

World Coordinate System



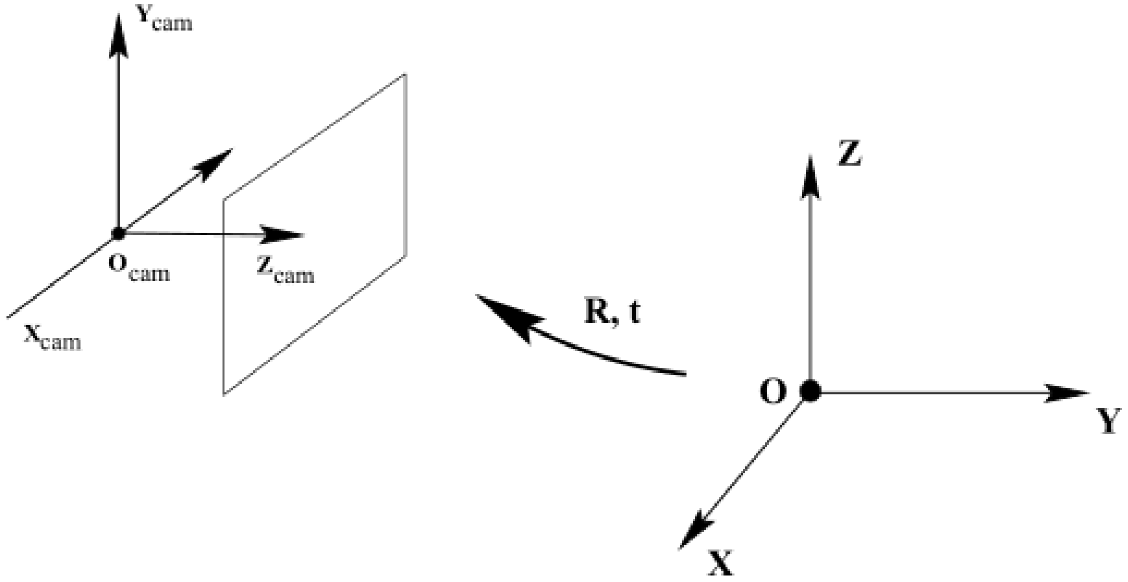
$$\begin{bmatrix} X_k \\ Y_k \\ Z_k \end{bmatrix} = \mathbf{R}\left(\begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \mathbf{T}\right)$$

World coordinate, camera coordinate and image coordinate systems

World Coordinate System

External camera parameters

$$\begin{pmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \\ 1 \end{pmatrix} = \begin{bmatrix} \mathtt{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$



Euclidean transformation between world and camera coordinates

- $\mathtt{R}$ is a 3×3 rotation matrix
- $\mathbf{t}$ is a 3×1 translation vector

Camera models

Concatenating the three matrices,

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \mathtt{K} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathtt{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{pmatrix} \mathrm{X} \\ \mathrm{Y} \\ \mathrm{Z} \\ 1 \end{pmatrix} = \mathtt{K}\,[\mathtt{R}|\ \mathbf{t}]\,\mathbf{X}$$

which defines the 3×4 projection matrix from Euclidean 3-space to an image as

$$\mathbf{x} = \mathtt{P}\mathbf{X} \qquad \mathtt{P} = \mathtt{K}\,[\mathtt{R}|\ \mathbf{t}] = \mathtt{K}\mathtt{R}[\mathtt{I}|\ \mathtt{R}^\top\mathbf{t}]$$

Note, the camera centre is at $(\mathrm{X}, \mathrm{Y}, \mathrm{Z})^\top = -\mathtt{R}^\top\mathbf{t}$.

In the following it is often only the 3×4 form of $\mathtt{P}$ that is important, rather than its decomposition.

Camera models

The camera model for perspective projection is a linear map between homogeneous point coordinates

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \left[\quad \mathtt{P} \ (3 \times 4) \quad \right] \begin{pmatrix} \mathrm{X} \\ \mathrm{Y} \\ \mathrm{Z} \\ 1 \end{pmatrix}$$

Image Point   Scene Point

$$\mathbf{x} \quad = \quad \mathtt{P} \quad \mathbf{X}$$

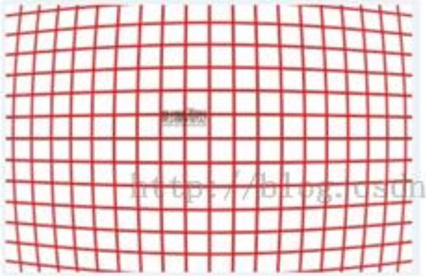
- The camera centre is the null-vector of $\mathtt{P}$
e.g. if $\mathtt{P} = [\mathtt{I}|0]$ then the centre is $\mathbf{X} = (0, 0, 0, 1)^\top$.
- $\mathtt{P}$ has 11 degrees of freedom (essential parameters).
- $\mathtt{P}$ has rank 3.

Camera models

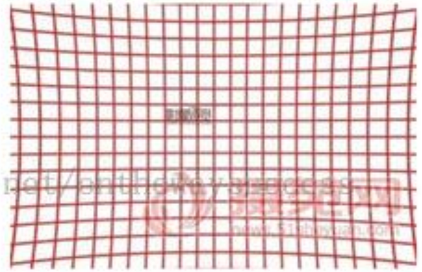
真实环境中，摄像机镜头并不是严格的小孔成像模型，会产生非线性畸变，一般只关心径向畸变

- **径向畸变**

产生原因是光线在远离透镜中心的地方比靠近中心的地方更加弯曲径向畸变主要包含桶形畸变和枕形畸变两种。下面两幅图是这两种畸变的示意：



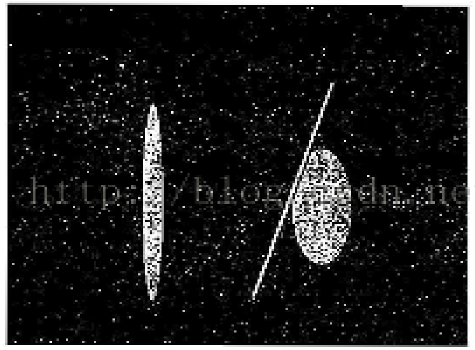
桶形畸变

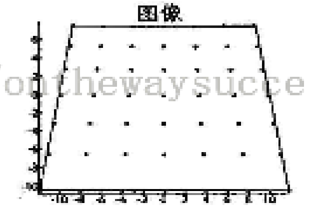


枕形畸变

- **切向畸变**

产生的原因透镜不完全平行于图像平面，这种现象发生于成像仪被粘贴在摄像机的时候；





Camera models

摄像机畸变模型（OpenCV）的数学表达式为：

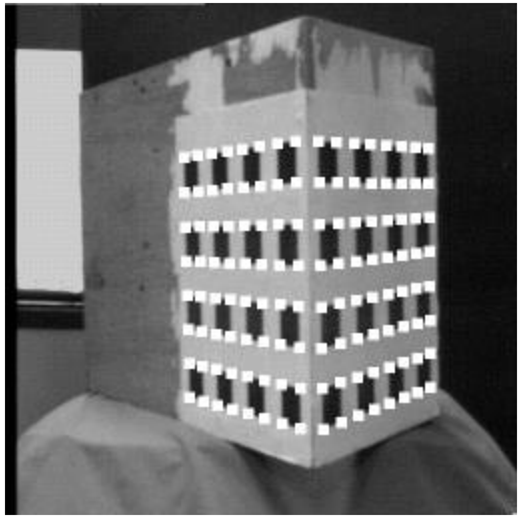
$$\begin{cases} \breve{u} = u + (u - u_0)[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \\ \breve{v} = v + (v - v_0)[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \end{cases}$$

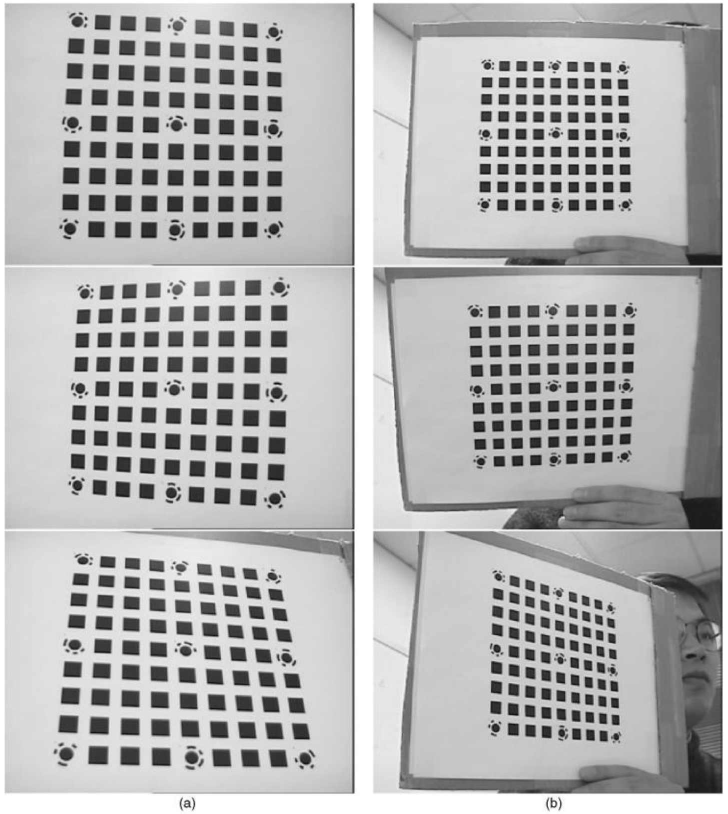
$$\begin{bmatrix} (u - u_0)(x^2 + y^2) & (u - u_0)(x^2 + y^2)^2 \\ (v - v_0)(x^2 + y^2) & (v - v_0)(x^2 + y^2)^2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} \breve{u} - u \\ \breve{v} - v \end{bmatrix}$$

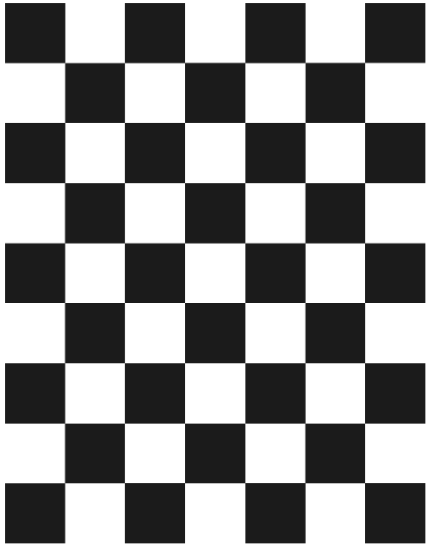
其中$(\breve{u}, \breve{v})$是有畸变图像坐标，(u, v)是通过小孔成像模型计算出的图像坐标，(x, y)是相应的图像物理坐标；

Camera calibration

- 在摄像机位置相对固定情况下，对于同一平面内的任意一点，单应性矩阵都是不变的；
- 我们可以找到世界坐标系中的多个点(X_{Wi}, Y_{Wi}, Z_{Wi})和其对应的图像坐标系中的点(u_i, v_i)的坐标，建立方程组，求解出该标定平面的单应性矩阵M；
- 根据求得的单应性矩阵M，再去求解相机内参、外参及畸变系数；







张正友标定法

张正友标定法因其模板容易制作，操作简单，成为最常用的相机标定方法

假设选定的所有点世界坐标Z=0：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = s\begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_x & v_0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} r_1 & r_2 & t \end{bmatrix}\begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix}$$

定义单应性矩阵H：

$$H = s\begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_x & v_0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} r_1 & r_2 & t \end{bmatrix} = sM[r_1, r_2, t]$$

H写成列向量形式：

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix}$$

根据旋转向量两两正交且长度不变性质：

$$r_1^T r_2 = 0, \|r_1\| = \|r_2\| = 1$$

$$h_1^T (M^{-1})^T M^{-1} h_2 = 0, h_1^T (M^{-1})^T M^{-1} h_1 = h_2^T (M^{-1})^T M^{-1} h_2$$

假设矩阵B如下：

$$B = (M^{-1})^T M^{-1} = \begin{bmatrix} 1/f_x^2 & 0 & -c/f_x^2 \\ 0 & 1/f_y^2 & -c/f_y^2 \\ -c/f_x^2 & -c/f_y^2 & c_x^2/f_x^2 + c_y^2/f_y^2 + 1 \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix}$$

可见B是一个对称矩阵，所以B的有效元素只剩下6个，另这6个元素构成向量b

$$b = [B_{11} \quad B_{12} \quad B_{22} \quad B_{13} \quad B_{23} \quad B_{33}]^T$$

张正友标定法

在此基础上，做进一步纯数学化简

$$h_i^T B h_j = v_{ij}^T b$$

可以计算得：

$$v_{ij} = [h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i2}h_{j2}, h_{i3}h_{j1} + h_{i1}h_{j3}, h_{i2}h_{j3}, h_{i3}h_{j3}]^T$$

利用上面的约束，可以得到方程组

$$\begin{bmatrix} V_{12}^T \\ (v_{11} - v_{12})^T \end{bmatrix} b = 0$$

- 上式包括4个未知量，如果需要求解这4个未知量，则需要2个单应性矩阵，2个单应性矩阵在两个约束下得到4个方程，也就可以得到全部的四个内参了。
- 为获得两个不同的单应性矩阵，需要改变摄像机与标定板的相对位置拍摄两张不同的照片。
- 通过至少含一个棋盘格的两幅图像，应用上述公式就可以估算出B ，得到B后，利用cholesky分解求取摄像机的内参矩阵。

张正友标定法

对于外参矩阵，可以通过下面的公式求解

$$\begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix} = sM \begin{bmatrix} r_1 & r_2 & t \end{bmatrix}$$

对上式进行化简，可以得到

$$s = 1 \big/ \left\| M^{-1} h_1 \right\| = 1 \big/ \left\| M^{-1} h_2 \right\|, r_1 = sM^{-1} h_1, r_2 = sM^{-1} h_2, r_3 = r_1 \times r_2, t = sM^{-1} h_3$$

至此，在不考虑畸变的条件下，内外参求解完毕

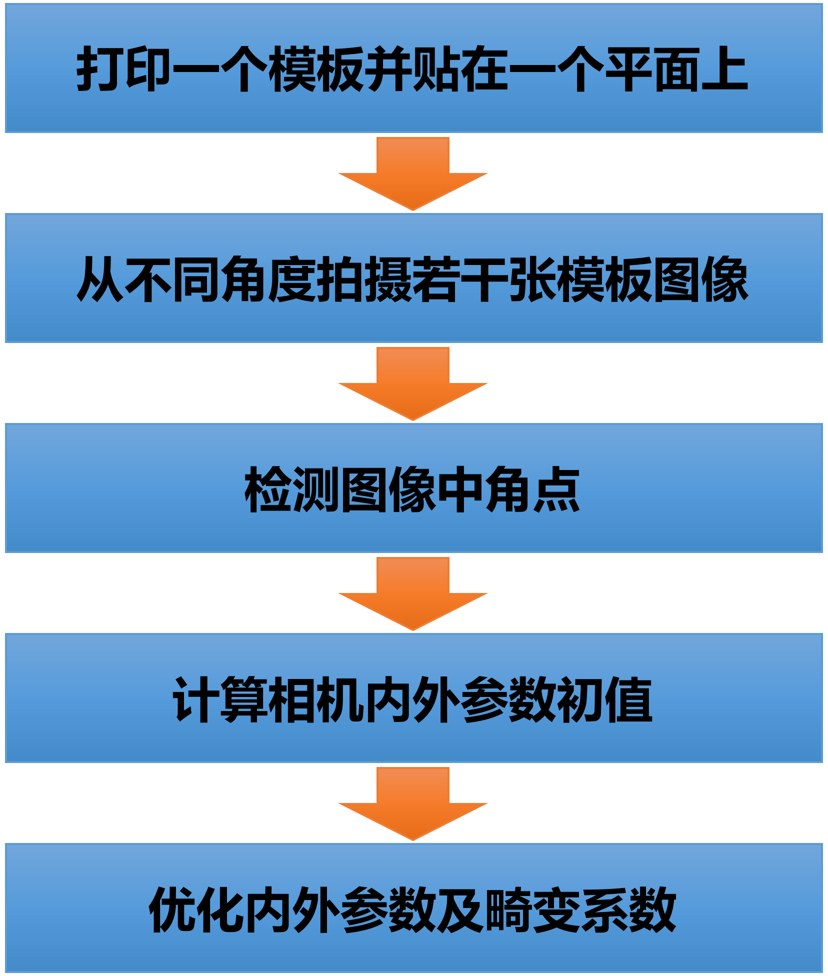
若进一步考虑透镜畸变情况，根据畸变模型，畸变前后的坐标关系为

$$\begin{bmatrix} x_{dis} \\ y_{dis} \end{bmatrix} = (1 + k_1 r^2 + k_2 r^4) \begin{bmatrix} x_p \\ y_p \end{bmatrix}$$

根据上式，如果有多张图像（多组对应点），可以使用最小二乘法对内外参数以及畸变系数进行优化

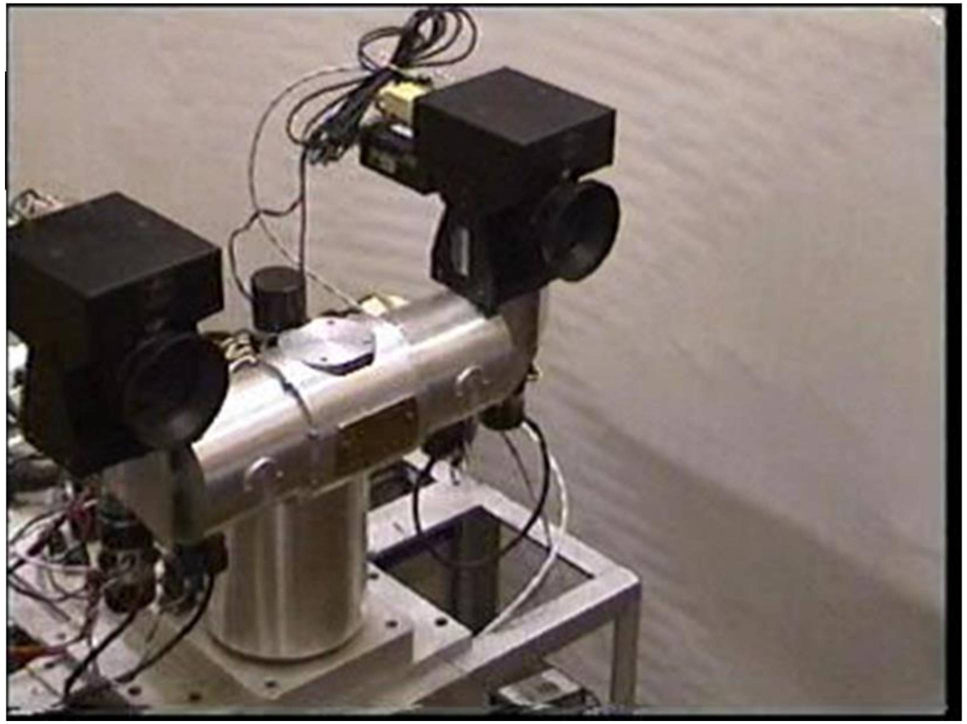
$$\sum_{i=1}^{n} \sum_{j=1}^{m} \| \mathbf{m}_{ij} - \hat{\mathbf{m}}(\mathbf{A}, \mathbf{R}_i, \mathbf{t}_i, \mathbf{M}_j) \|^2,$$

张正友标定法


打印一个模板并贴在一个平面上
从不同角度拍摄若干张模板图像
检测图像中角点
计算相机内外参数初值
优化内外参数及畸变系数

Contents

Two-view geometry



- **具有两个相机的立体设备**
  - 同时获取两幅图像
  - 相机相对位置固定
  - 一般用于测距（Stereo）
- **单个移动相机**
  - 顺序获取两张图像
  - 一般用于移动位姿估计（SLAM）

Two-view geometry

- Cameras $\mathtt{P}$ and $\mathtt{P}'$ such that

$$\mathbf{x} = \mathtt{P}\mathbf{X} \qquad \mathbf{x}' = \mathtt{P}'\mathbf{X}$$

- Baseline between the cameras is non-zero.

Given an image point in the first view, where is the corresponding point in the second view?

What is the relative position of the cameras?

What is the 3D geometry of the scene?

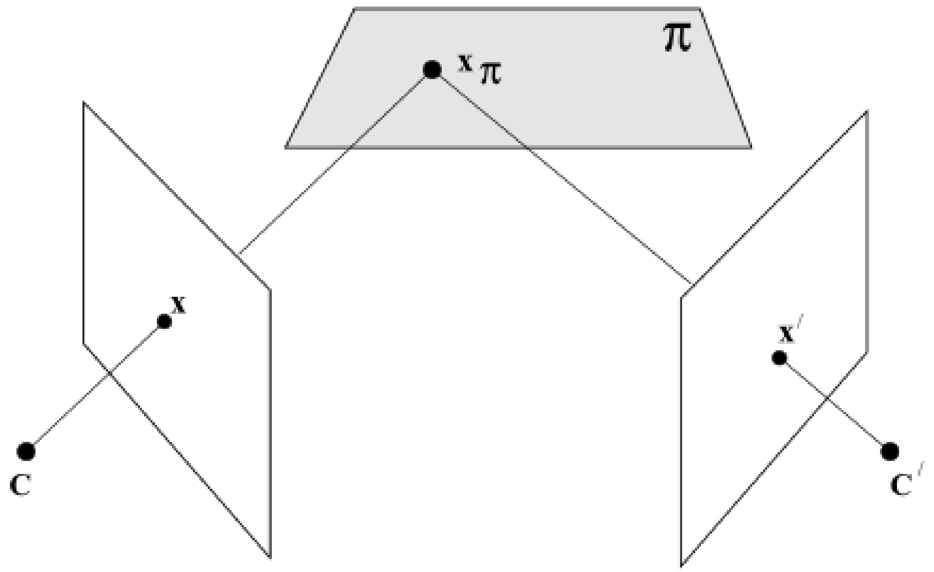
Images of Planes

Projective transformation between images induced by a plane

$$\mathbf{x} = \mathtt{H}_{1\pi}\mathbf{x}_{\pi} \qquad \mathbf{x}' = \mathtt{H}_{2\pi}\mathbf{x}_{\pi}$$

$$\mathbf{x}' = \mathtt{H}_{2\pi}\mathbf{x}_{\pi}$$

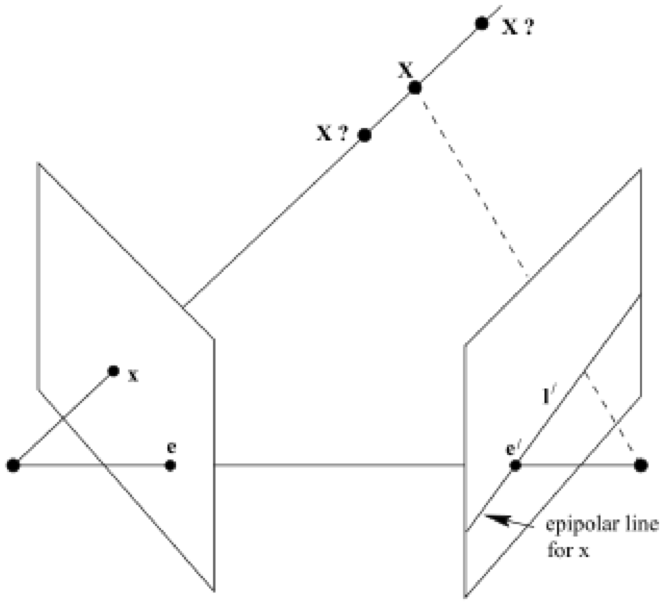
$$= \mathtt{H}_{2\pi}\mathtt{H}_{1\pi}^{-1}\mathbf{x} = \mathtt{H}\mathbf{x}$$



- $\mathtt{H}$ can be computed from the correspondence of four points on the plane.

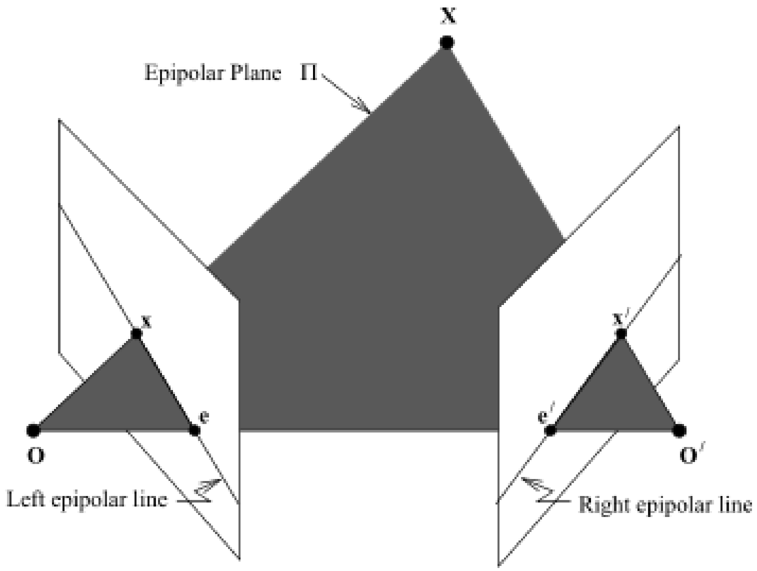
Correspondence Geometry

Given the image of a point in one view, what can we say about its position in another?



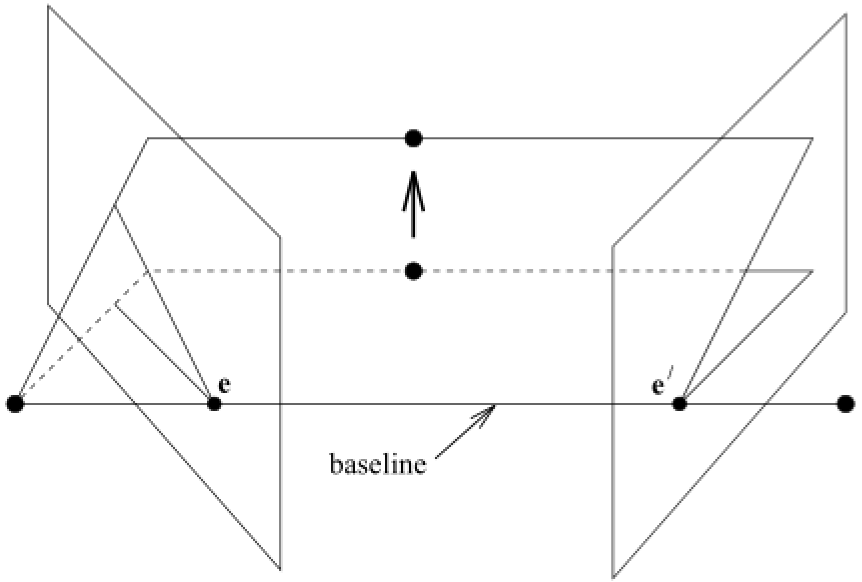
- A point in one image "generates" a line in the other image.
- This line is known as an epipolar line, and the geometry which gives rise to it is known as epipolar geometry.

Epipolar Geometry



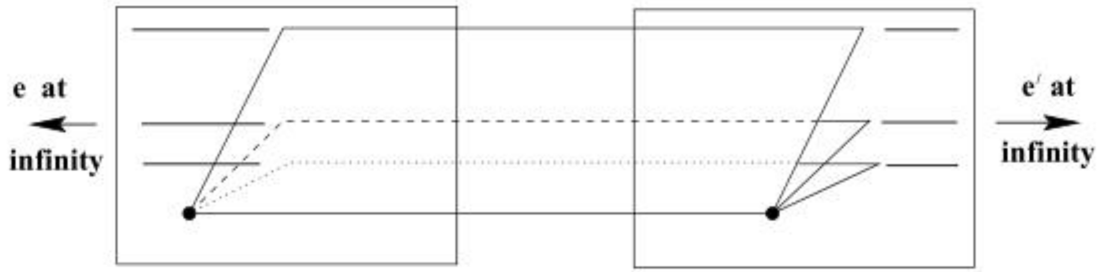
- The epipolar line $\mathbf{l}'$ is the image of the ray through $\mathbf{x}$.
- The epipole $\mathbf{e}'$ is the point of intersection of the line joining the camera centres—the baseline—with the image plane.
- The epipole is also the image in one camera of the centre of the other camera.
- All epipolar lines intersect in the epipole.

Epipolar pencil



As the position of the 3D point $\mathbf{X}$ varies, the epipolar planes "rotate" about the baseline. This family of planes is known as an epipolar pencil. All epipolar lines intersect at the epipole.

Epipolar geometry example



Epipolar geometry depends only on the relative pose (position and orientation) and internal parameters of the two cameras, i.e. the position of the camera centres and image planes. It does not depend on structure (3D points external to the camera).

Homogeneous Notation Interlude

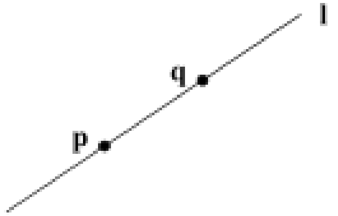
- A line $\mathbf{l}$ is represented by the homogeneous 3-vector

$$\mathbf{l} = \begin{pmatrix} l_1 \\ l_2 \\ l_3 \end{pmatrix}$$

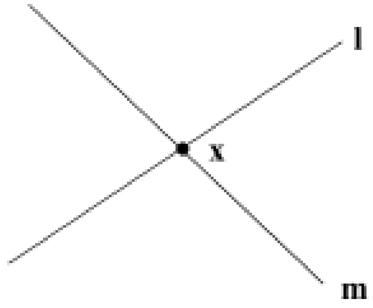
for the line $l_1 x + l_2 y + l_3 = 0$. Only the ratio of the homogeneous line coordinates is significant.

- point on line: $\mathbf{l}.\mathbf{x} = 0$ or $\mathbf{l}^\top \mathbf{x} = 0$ or $\mathbf{x}^\top \mathbf{l} = 0$

- two points define a line: $\mathbf{l} = \mathbf{p} \times \mathbf{q}$



- two lines define a point: $\mathbf{x} = \mathbf{l} \times \mathbf{m}$



Matrix notation for vector product

The vector product $\mathbf{v} \times \mathbf{x}$ can be represented as a matrix multiplication

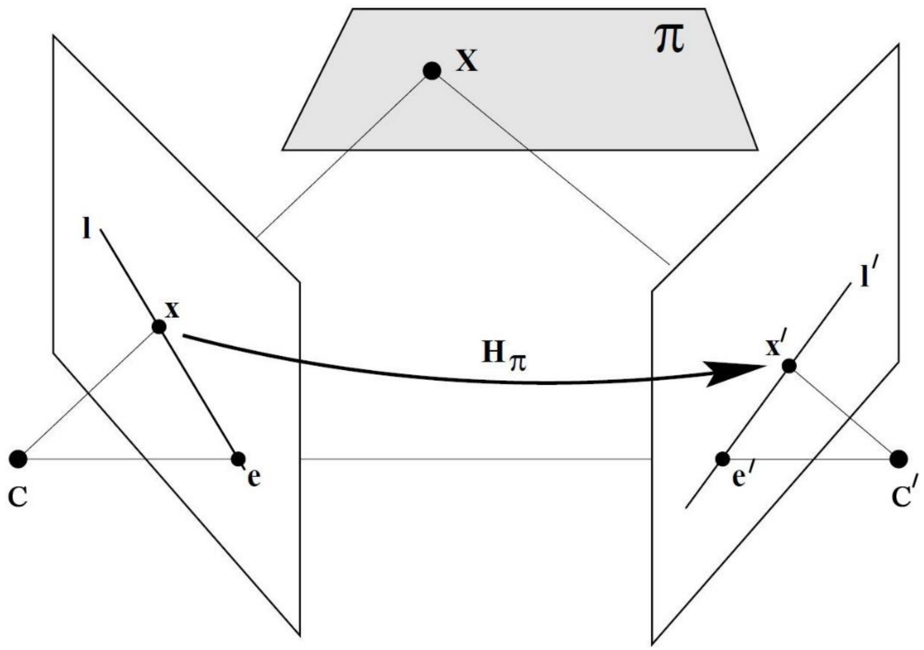
$$\mathbf{v} \times \mathbf{x} = [\mathbf{v}]_\times \mathbf{x}$$

where

$$[\mathbf{v}]_\times = \begin{bmatrix} 0 & -v_z & v_y \\ v_z & 0 & -v_x \\ -v_y & v_x & 0 \end{bmatrix}$$

- $[\mathbf{v}]_\times$ is a 3×3 skew-symmetric matrix of rank 2.
- $\mathbf{v}$ is the null-vector of $[\mathbf{v}]_\times$, since $\mathbf{v} \times \mathbf{v} = [\mathbf{v}]_\times \mathbf{v} = \mathbf{0}$.

Fundamental Matrices



基本矩阵

基本矩阵是对极几何的代数表示。给定一对图像，对于图像上的每点 **x**，在另一幅图像中存在一条对应的对极线 **l**’。在第二幅图像上，任何与该点 **x** 匹配的点 **x**’必然在对极线 **l**’上。该对极线是过点 **x** 与第一个摄像机中心 **C** 的射线在第二幅图像上的投影。因此，存在一个从一幅图像上的点到另一幅图像上与之对应的对极线的映射

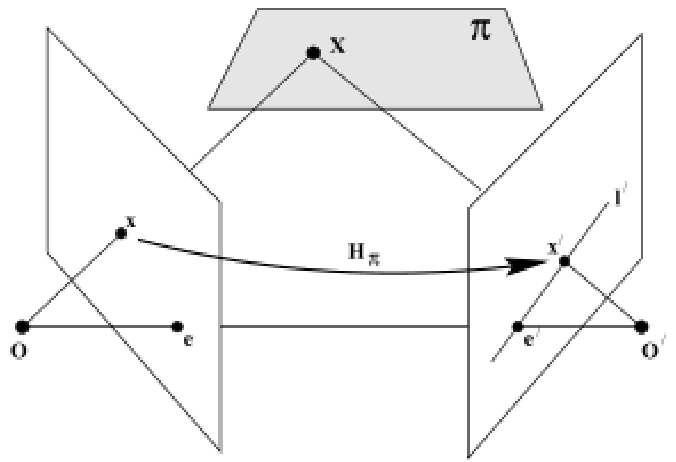
$$\mathbf{x} \rightarrow \mathbf{l}'$$

这个映射是一个（奇异）对射，即是由**基本矩阵** F 表示的从点到直线的射影映射。

Fundamental Matrices

$$\mathbf{x}'^{\top} \, \mathtt{F} \, \mathbf{x} = 0 \qquad \mathbf{l}' = \mathtt{F} \, \mathbf{x}$$

- $\mathtt{F}$ is a 3×3 rank 2 homogeneous matrix
- $\mathtt{F}^{\top}\mathbf{e}' = \mathbf{0}$
- It has 7 degrees of freedom
- Counting: $2 \times 11 - 15 = 7$.
- Compute from 7 image point correspondences



Step 1: Point transfer via a plane $\mathbf{x}' = \mathtt{H}_\pi \mathbf{x}$

Step 2 : Construct the epipolar line $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}' = [\mathbf{e}']_\times \mathbf{x}'$

$\mathbf{l}' = [\mathbf{e}']_\times \mathtt{H}_\pi \mathbf{x} = \mathtt{F}\mathbf{x}$

$\mathtt{F} = [\mathbf{e}']_\times \mathtt{H}_\pi$

This shows that $\mathtt{F}$ is a 3×3 rank 2 matrix.

Fundamental Matrices

- **基本矩阵的性质**

  假设两幅图像由中心不重合的摄像机获得，则基本矩阵 F 为对所有的对应点 $\mathbf{x}\leftrightarrow\mathbf{x}'$，都满足

$$\mathbf{x}'^{T}F\mathbf{x} = 0$$

- F 是秩 2、自由度 7 的齐次矩阵
- **点对应：** 如果 $\mathbf{x}$ 和 $\mathbf{x}'$是对应的图像点，那么$\mathbf{x}'^{T}F\mathbf{x} = \mathbf{0}$
- **对极线：**
  - ✓ $\mathbf{l}'=F\mathbf{x}$ 是对应于 $\mathbf{x}$ 的对极线
  - ✓ $\mathbf{l}=F^{T}\mathbf{x}'$是对应于 $\mathbf{x}'$的对极线
- **对极点：**
  - ✓ $F\mathbf{e} = \mathbf{0}$
  - ✓ $F^{T}\mathbf{e}' = \mathbf{0}$
- **由摄像机矩阵 P、P'进行计算：**
  - ✓ 一般摄像机

    $F=[\mathbf{e}']_{\times}P'P^{+}$，其中 P^{+}是 P 的伪逆，$\mathbf{e}'=P'C$ 且 $PC=\mathbf{0}$.
  - ✓ 规范摄像机，$P=[I|\mathbf{0}]$，$P'=[M|\mathbf{m}]$，

    $F=[\mathbf{e}']_{\times}M=M^{-T}[\mathbf{e}]_{\times}$，其中 $\mathbf{e}'=\mathbf{m}$ 且 $\mathbf{e}=M^{-1}\mathbf{m}$
  - ✓ 非无穷远摄像机 $P=K[I|\mathbf{0}]$，$P'=K'[R|\mathbf{t}]$

    $F=K'^{-T}[\mathbf{t}]_{\times}RK^{-1}=[K'\mathbf{t}]_{\times}K'RK^{-1} = K'^{-T}RK^{T}[KR^{T}\mathbf{t}]_{\times}$

Essential Matrices

本质矩阵

本质矩阵是归一化图像坐标下的基本矩阵的特殊形式。考虑分解为 $P=K[R|\mathbf{t}]$的摄像机矩阵，并令 $\mathbf{x}=P\mathbf{X}$ 为图像上一点，如果知道标定矩阵 K，用它的逆矩阵作用于 $\mathbf{x}$，得到点$\hat{\mathbf{x}} = K^{-1}\mathbf{x}$，则$\hat{\mathbf{x}} = [R|\mathbf{t}]\mathbf{X}$，其中$\hat{\mathbf{x}}$是图像上的点在**归一化坐标**下的表示。它可以被视为空间点 $\mathbf{X}$ 在摄像机$[R|\mathbf{t}]$的标定矩阵等于单位矩阵 I 情形下的像。摄像机矩阵$K^{-1}P = [R|\mathbf{t}]$称为**归一化摄像机矩阵**。

现在，考虑一对归一化的摄像机矩阵 $P=[I|\mathbf{0}]$和 $P'=[R|\mathbf{t}]$。与归一化摄像机矩阵对应的基本矩阵按惯例称为**本质矩阵**，它具有如下形式

$$E = [\mathbf{t}]_{\times}R = R[R^T\mathbf{t}]_{\times}$$

用归一化图像坐标表示对应点 $\mathbf{x}\leftrightarrow\mathbf{x}'$时，本质矩阵的定义方程是

$$\hat{\mathbf{x}}'^T E\hat{\mathbf{x}} = 0$$

又有$\hat{\mathbf{x}} = K^{-1}\mathbf{x}$且$\hat{\mathbf{x}}' = K'^{-1}\mathbf{x}'$，则 $\mathbf{x}'^T K'^{-T} E K^{-1}\mathbf{x} = 0$，把它与基本矩阵关系式$\mathbf{x}'^T F\mathbf{x} = 0$比较便可得到基本矩阵和本质矩阵之间的关系是

$$E = K'^T F K$$

➢ **本质矩阵的性质**

本质矩阵 $E=[\mathbf{t}]_{\times}R$ 只有 5 个自由度：旋转矩阵 R 和平移矢量 $\mathbf{t}$ 各有 3 个自由度，但是有一个整体尺度因子的多义性——与基本矩阵一样，本质矩阵也是一个齐次量。

基础矩阵（F）：

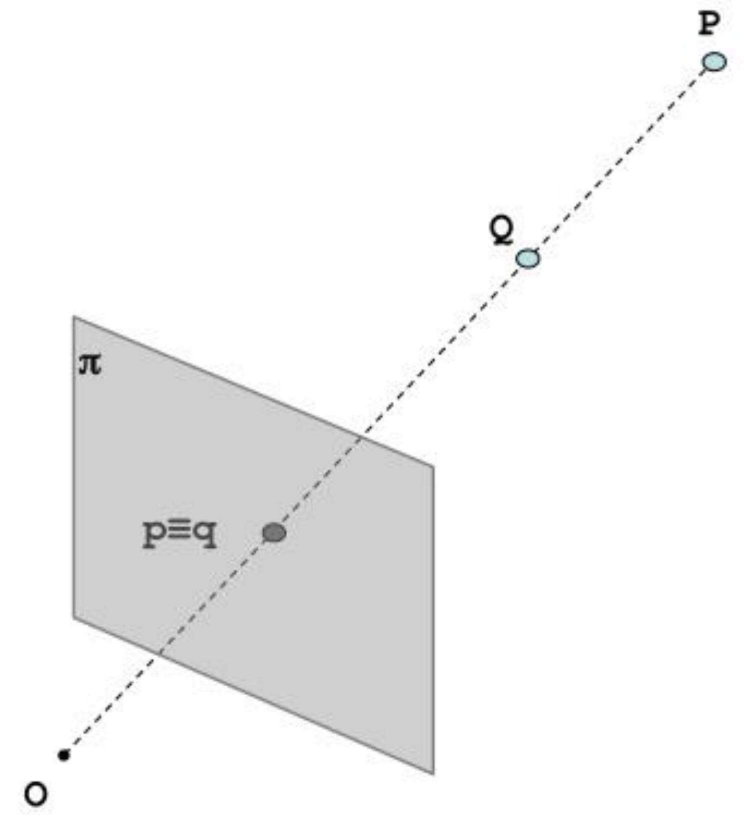
空间中同一点在2个图像坐标系上不同(二维)坐标之间的关系

本质矩阵（E）：

空间中同一点在2个相机坐标系中不同(三维)坐标之间的关系

Contents

Single Camera

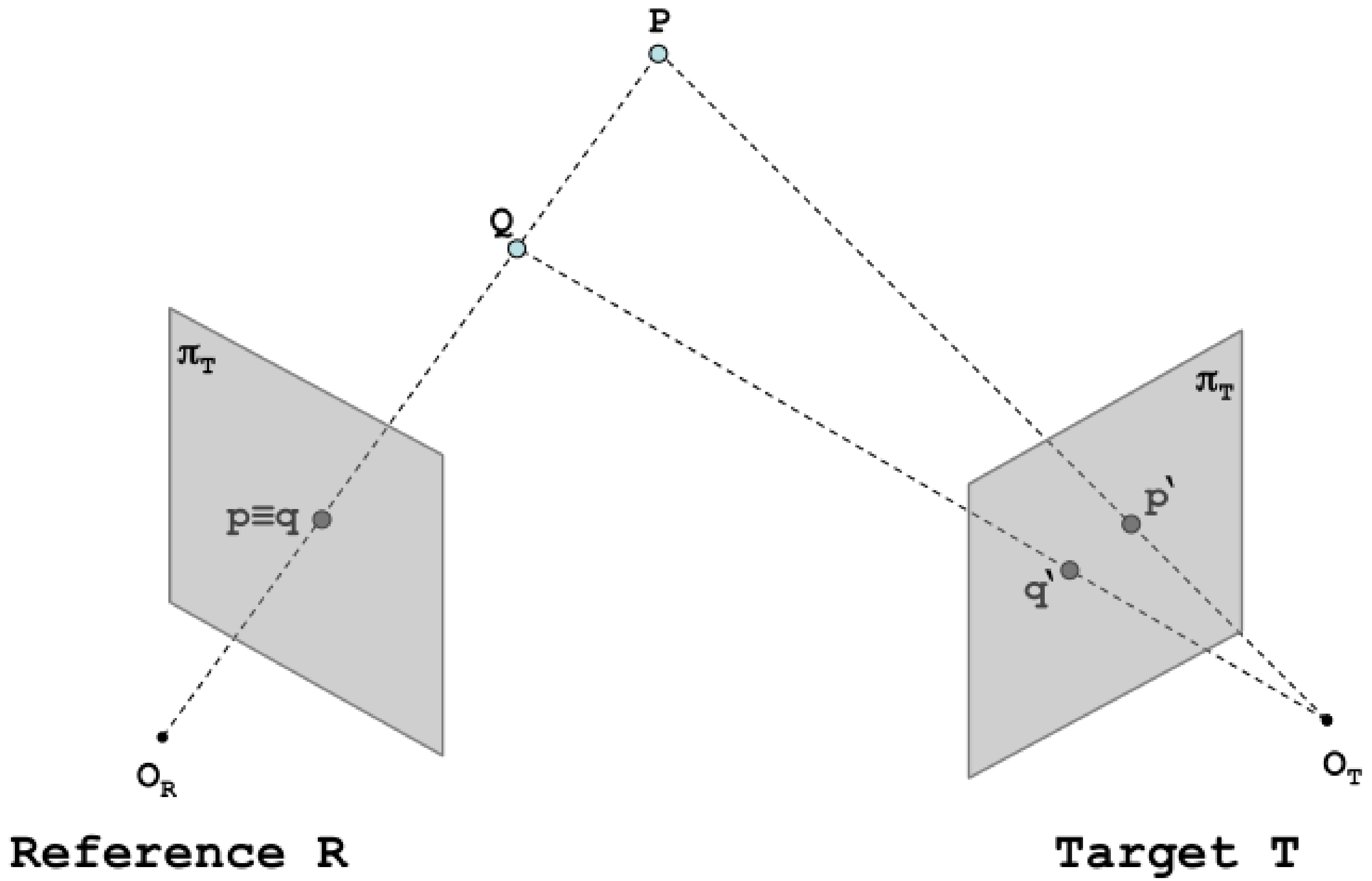


π: image plane

O: optical center

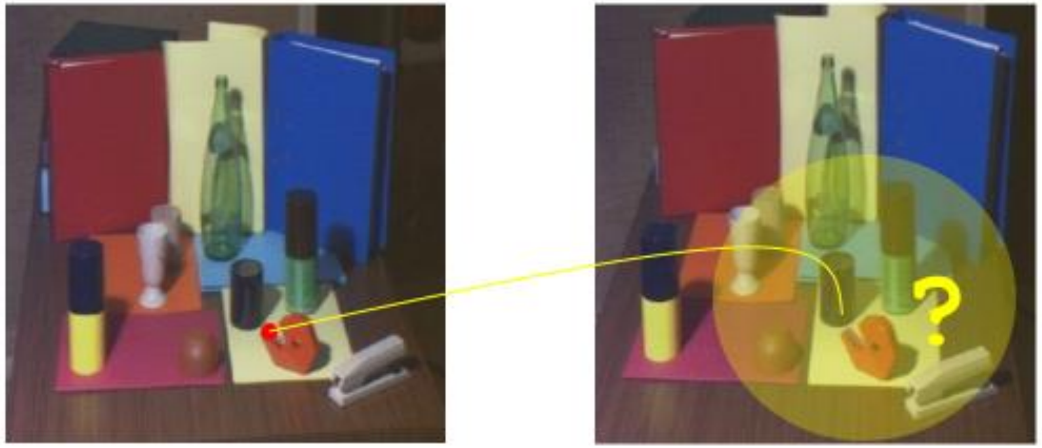
- Both (real) points (P and Q) project into the same image point (p ≡ q)
- This occurs for each point along the same line of sight
- Useful for optical illusions…

Stereo Camera



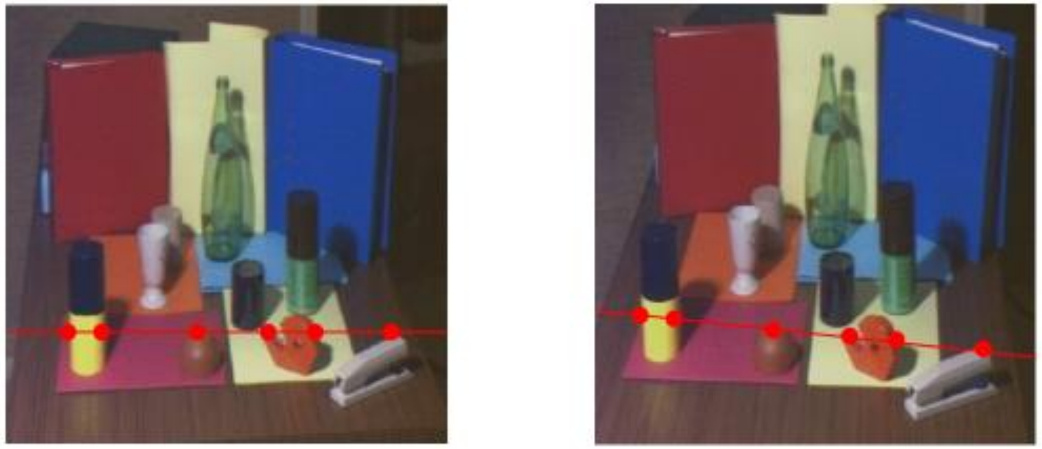
With two (or more) cameras we can infer depth, by means of triangulation, if we are able to find corresponding (homologous) points in the two images

How to solve the correspondence problem?



Reference (R) Target (T)

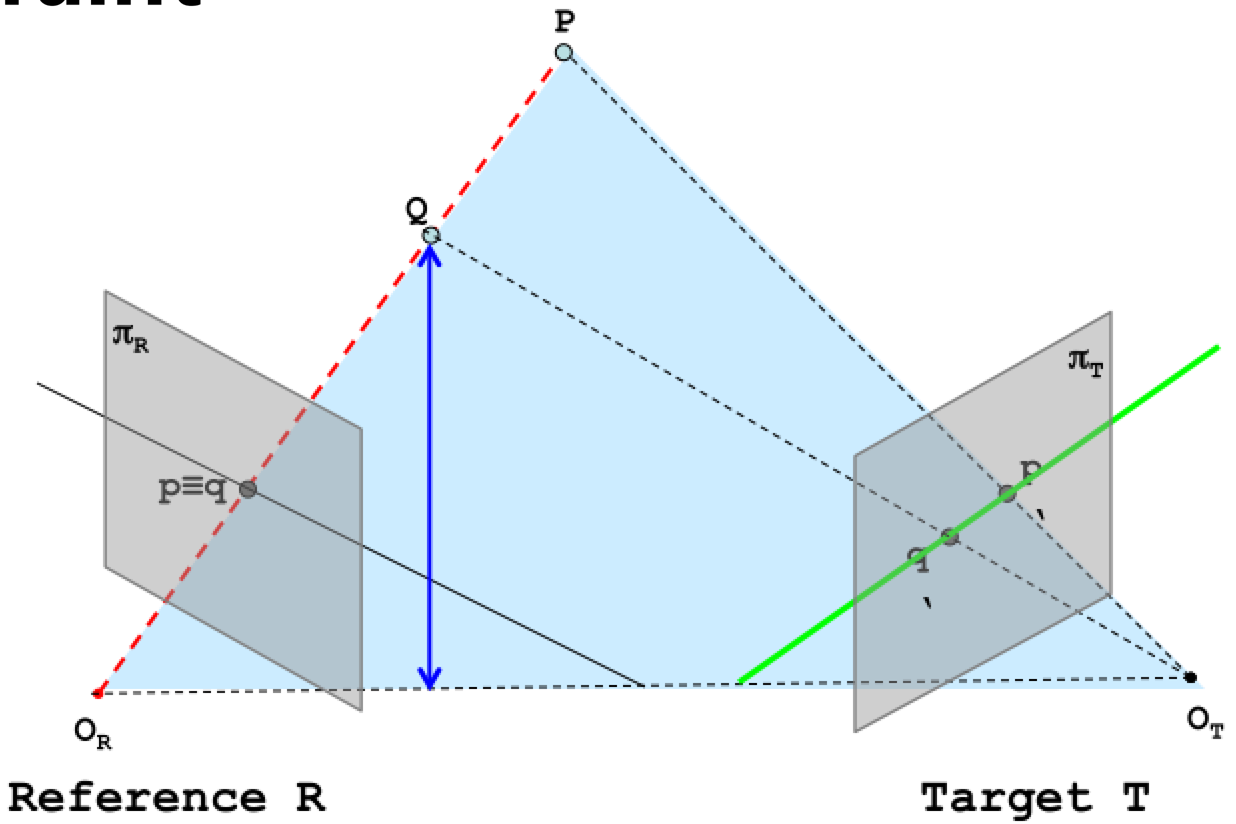
2D search domain ?



Reference (R) Target (T)

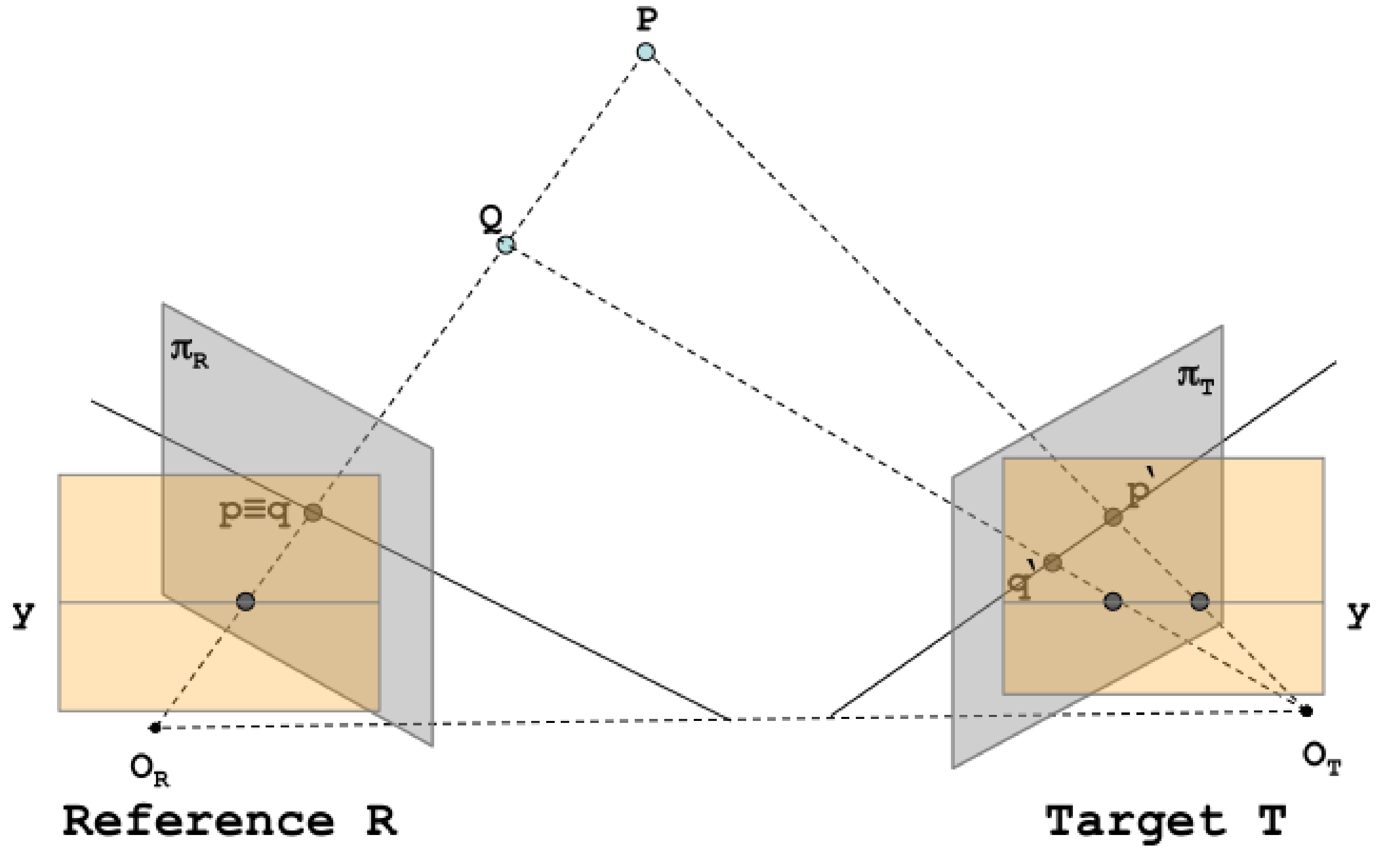
No!! Thanks to the epipolar constraint

Epipolar constraint



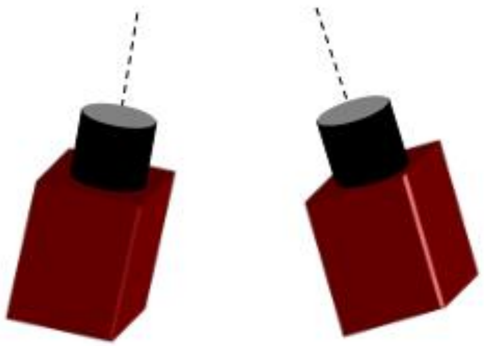
- Consider two points P and Q on the same line of sight of the reference image R (both points project into the same image point p≡q on image plane π_R of the reference image)
- The epipolar constraint states that the correspondence for a point belonging to the (red) line of sight lies on the green line on image plane π_T of target image

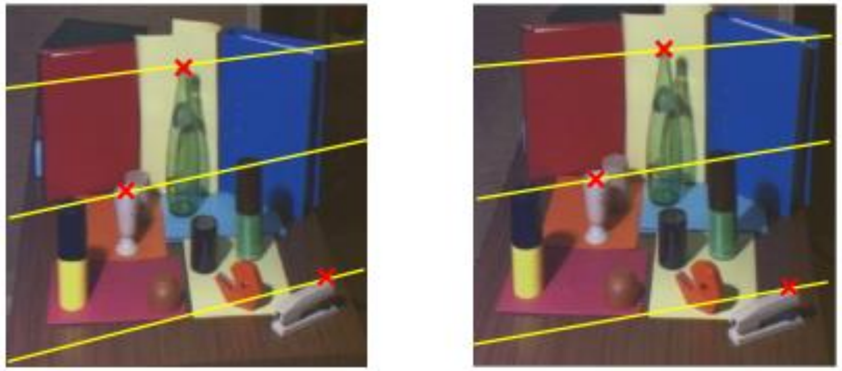
Stereo camera in standard form



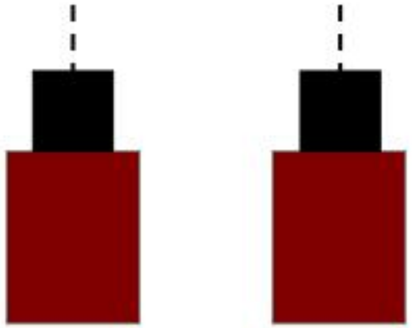
Once we know that the search space for corresponding points can be narrowed from 2D to 1D, we can put (virtually) the stereo rig in a more convenient configuration (standard form) - corresponding points are constrained on the same image scanline

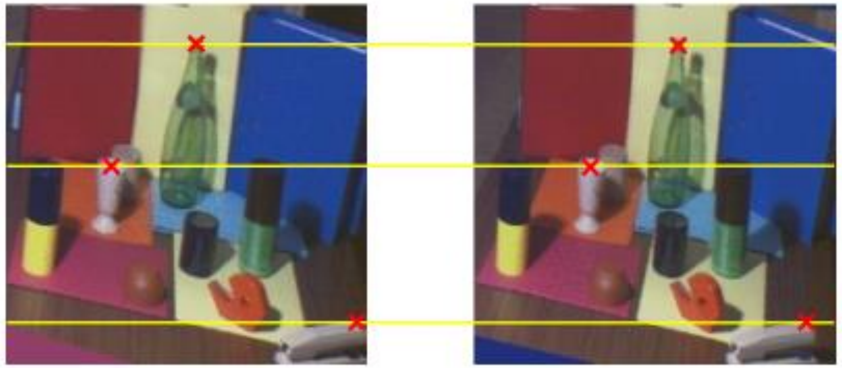
Stereo camera in standard form





Original stereo pair



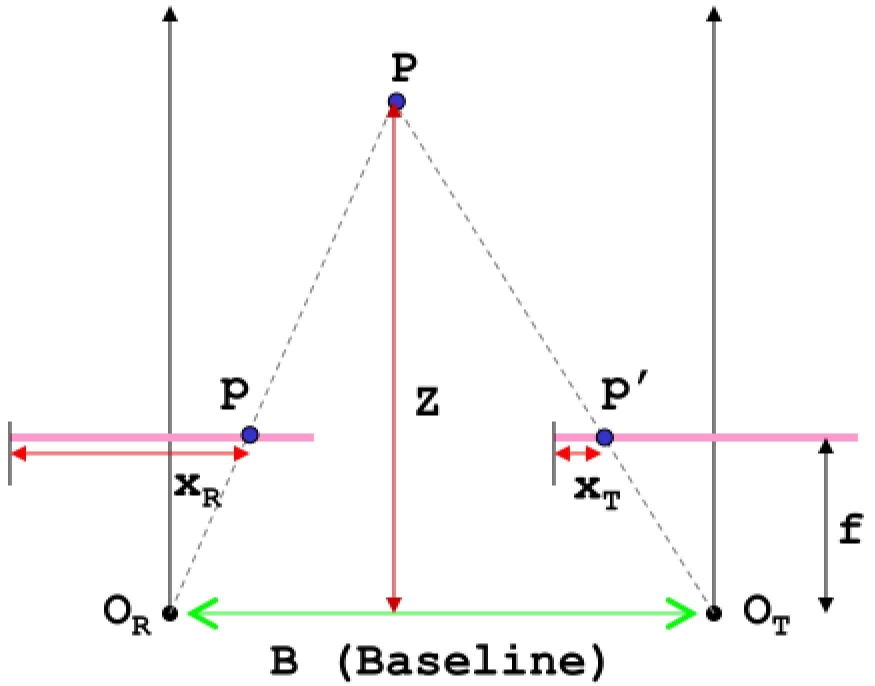


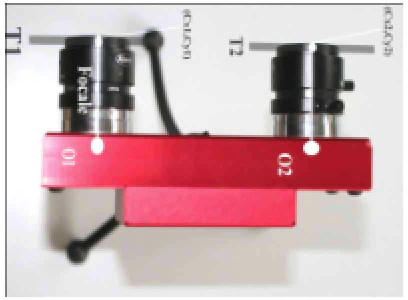
Stereo pair in standard form

Bouguet极线校正方法

Cameras are "*perfectly*" aligned
and with the same *focal length*

Disparity and depth



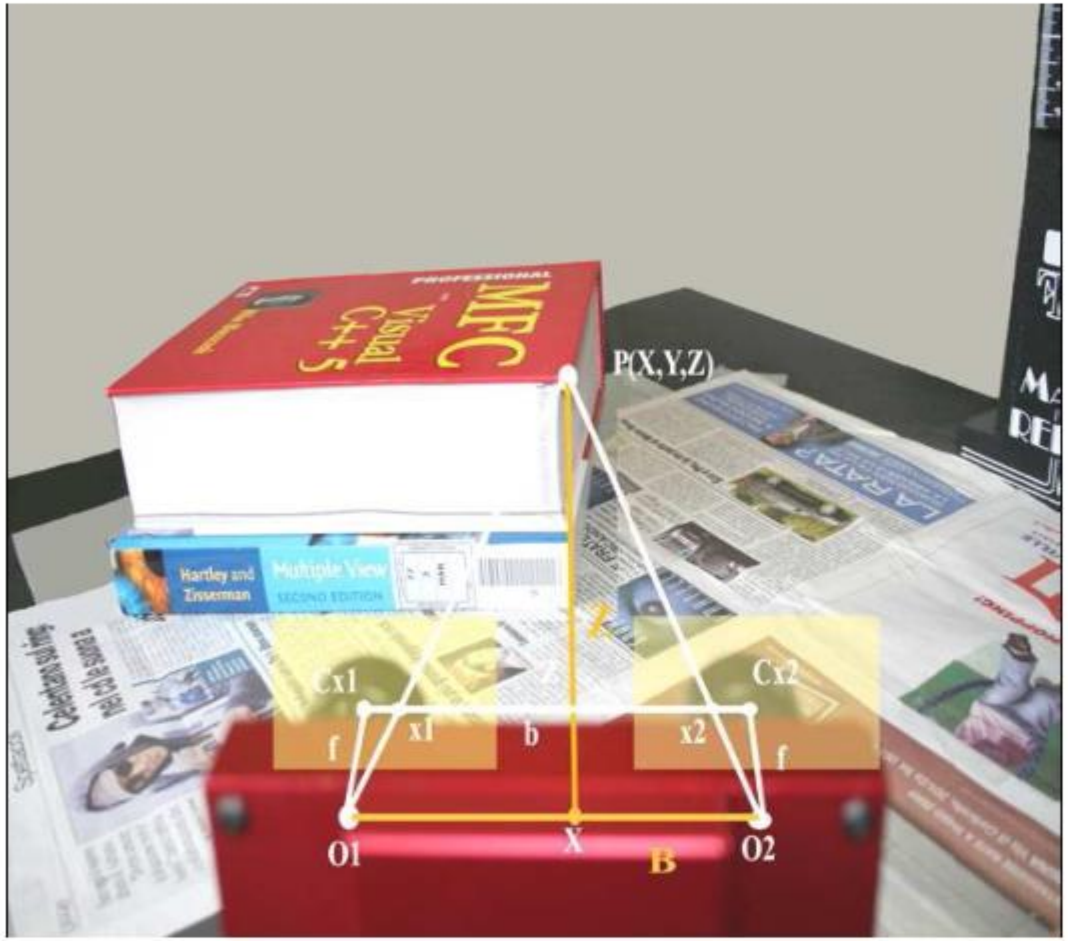


With the stereo rig in standard form and by considering similar triangles (PO_RO_T and Ppp'):

$$\frac{b}{Z} = \frac{(b+x_T)-x_R}{Z-f} \quad \Rightarrow \quad Z = \frac{b \cdot f}{x_R - x_T} = \frac{b \cdot f}{d}$$

$x_R - x_T$ is the **disparity**

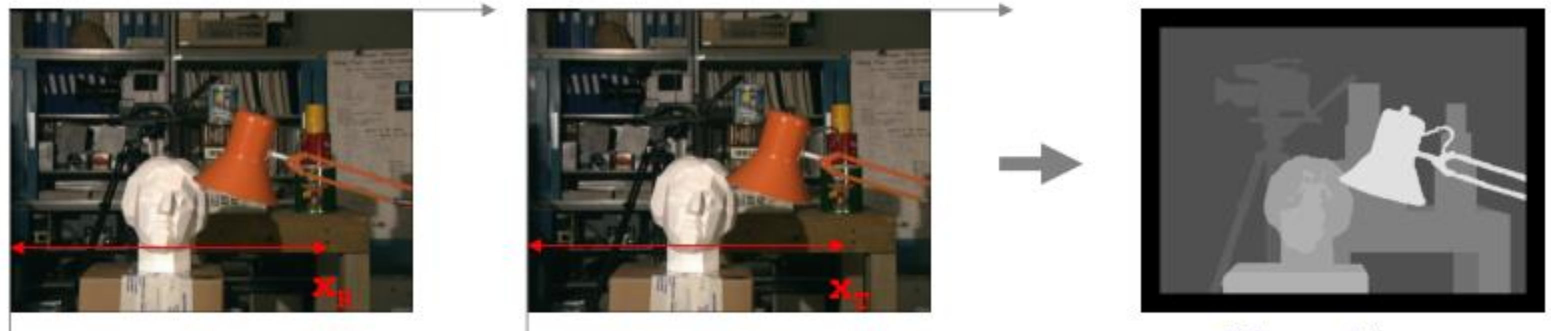
Disparity and depth



Disparity and depth

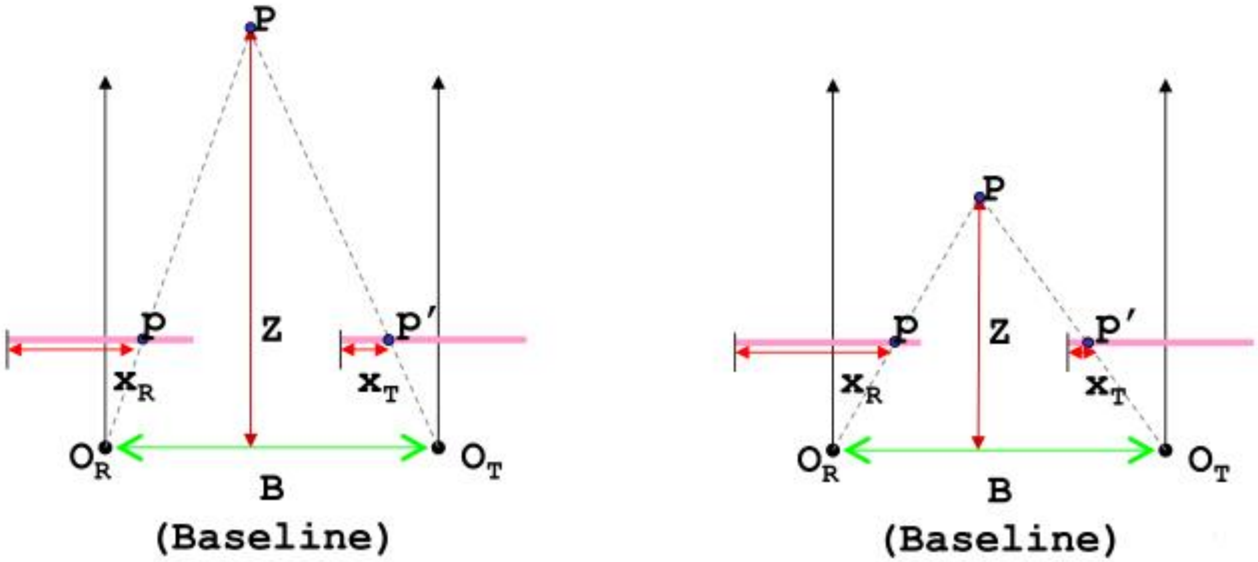
Disparity and depth

The disparity is the difference between the x coordinate of two corresponding points; it is typically encoded with greyscale image (closer points are brighter).



Reference Target Disparity map

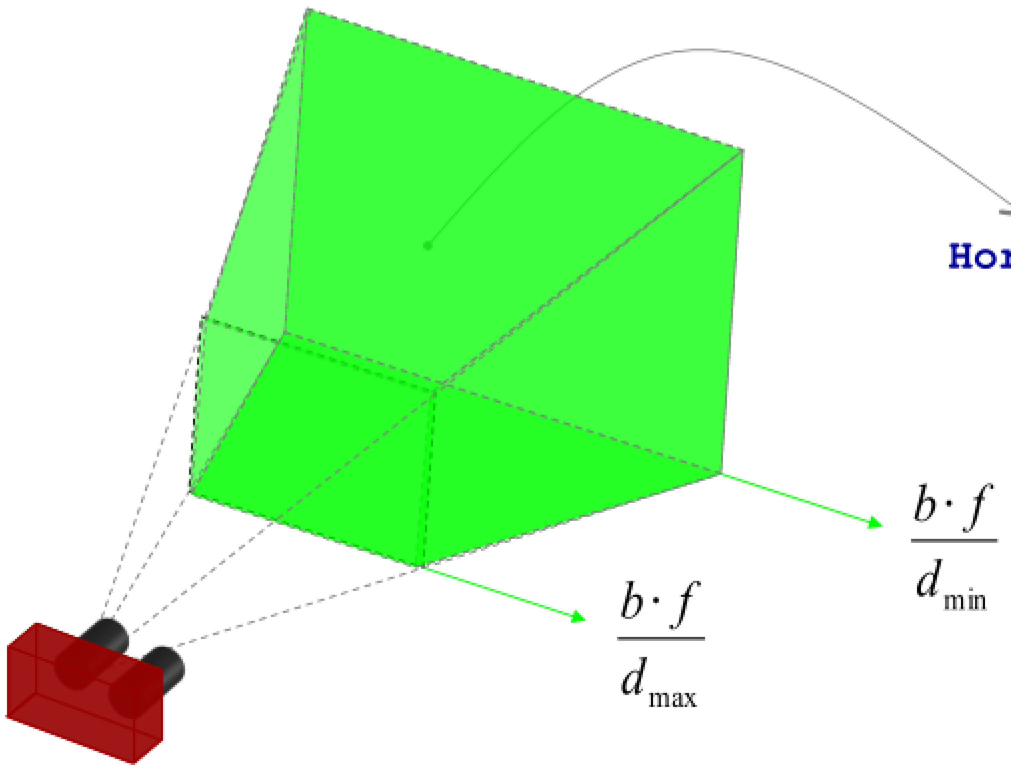
Disparity is higher for points closer to the camera



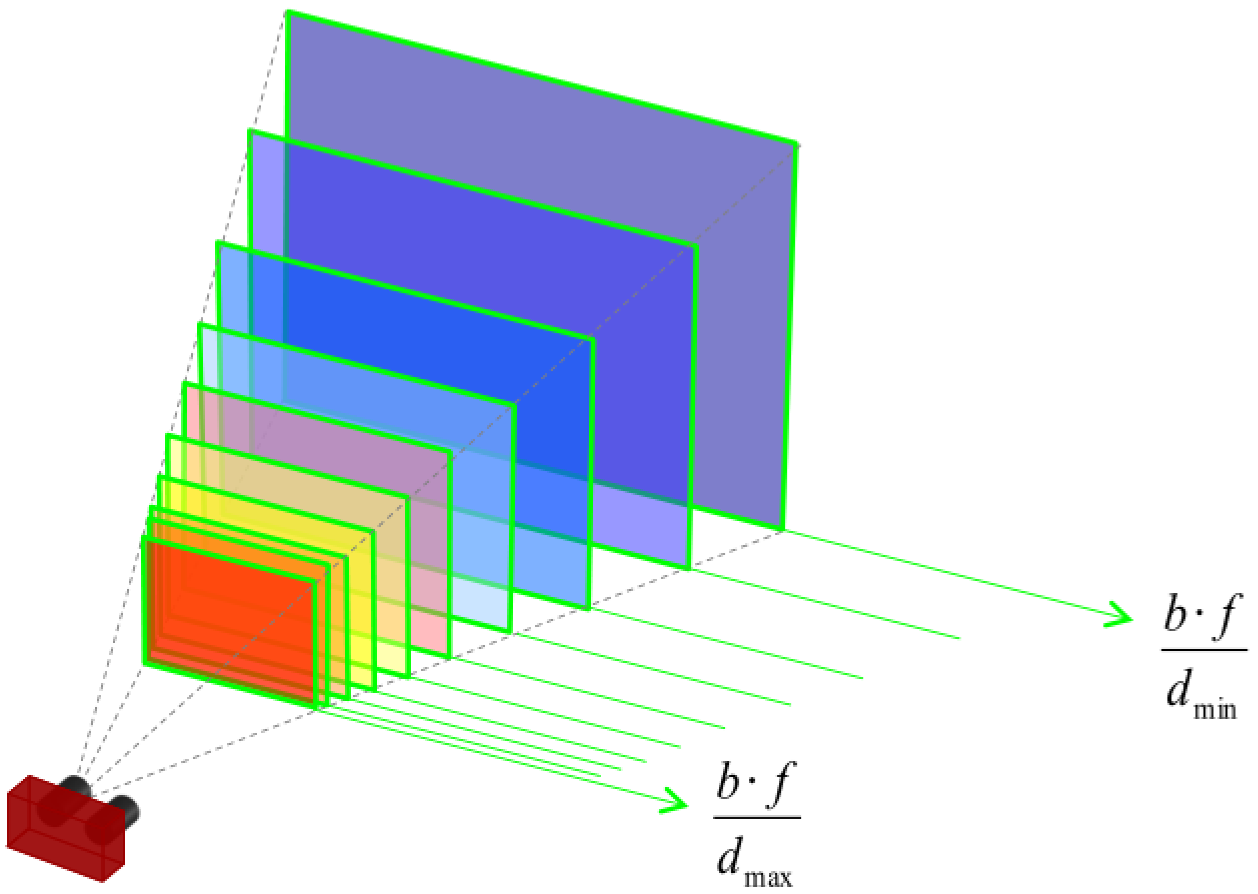
Disparity and depth

Range field (Horopter)

Given a stereo rig with baseline b and focal length f, the range field of the system is constrained by the disparity range [d_{min}, d_{max}].

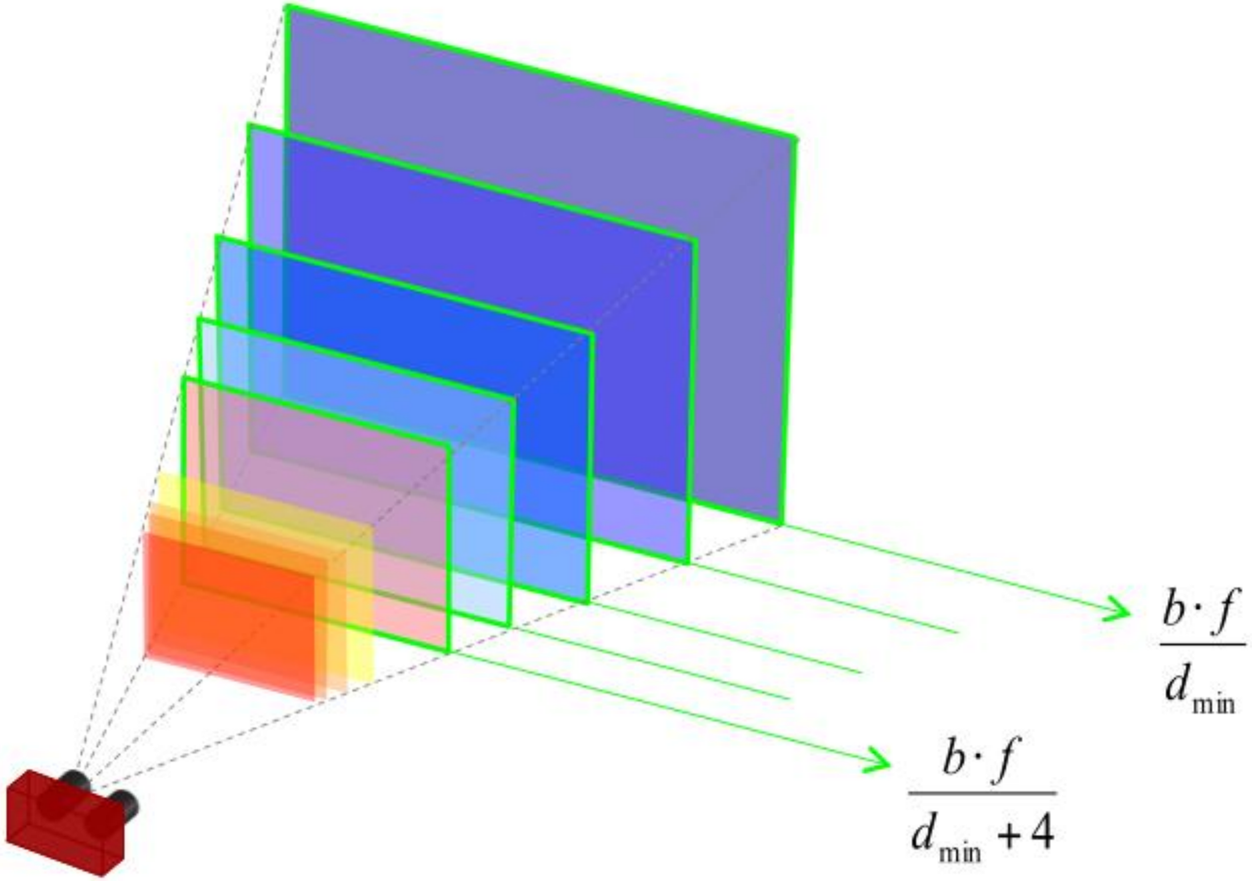


Disparity and depth



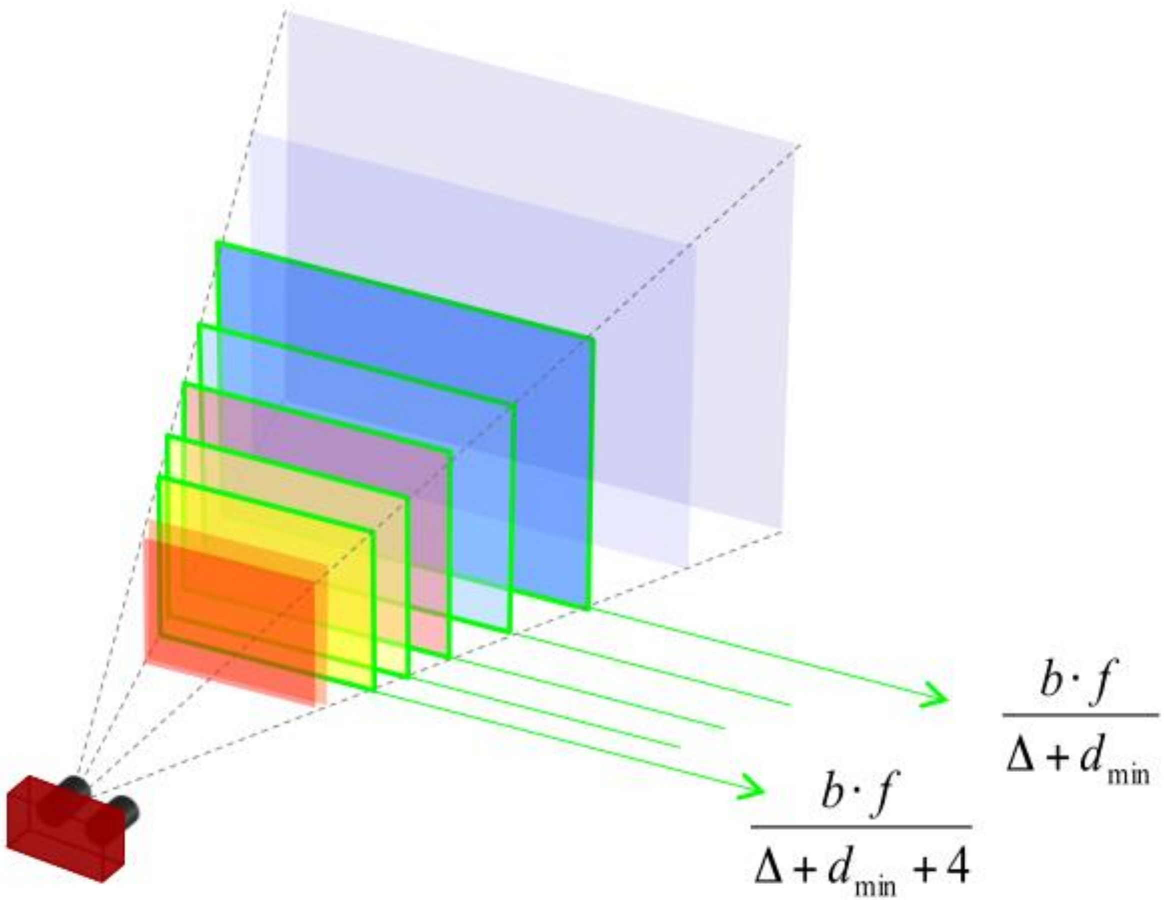
- Depth measured by a stereo vision system is discretized into parallel planes (one for each disparity value)
- A better (virtual) discretization can be achieved with subpixel techniques (see Disparity Refinements)

Disparity and depth



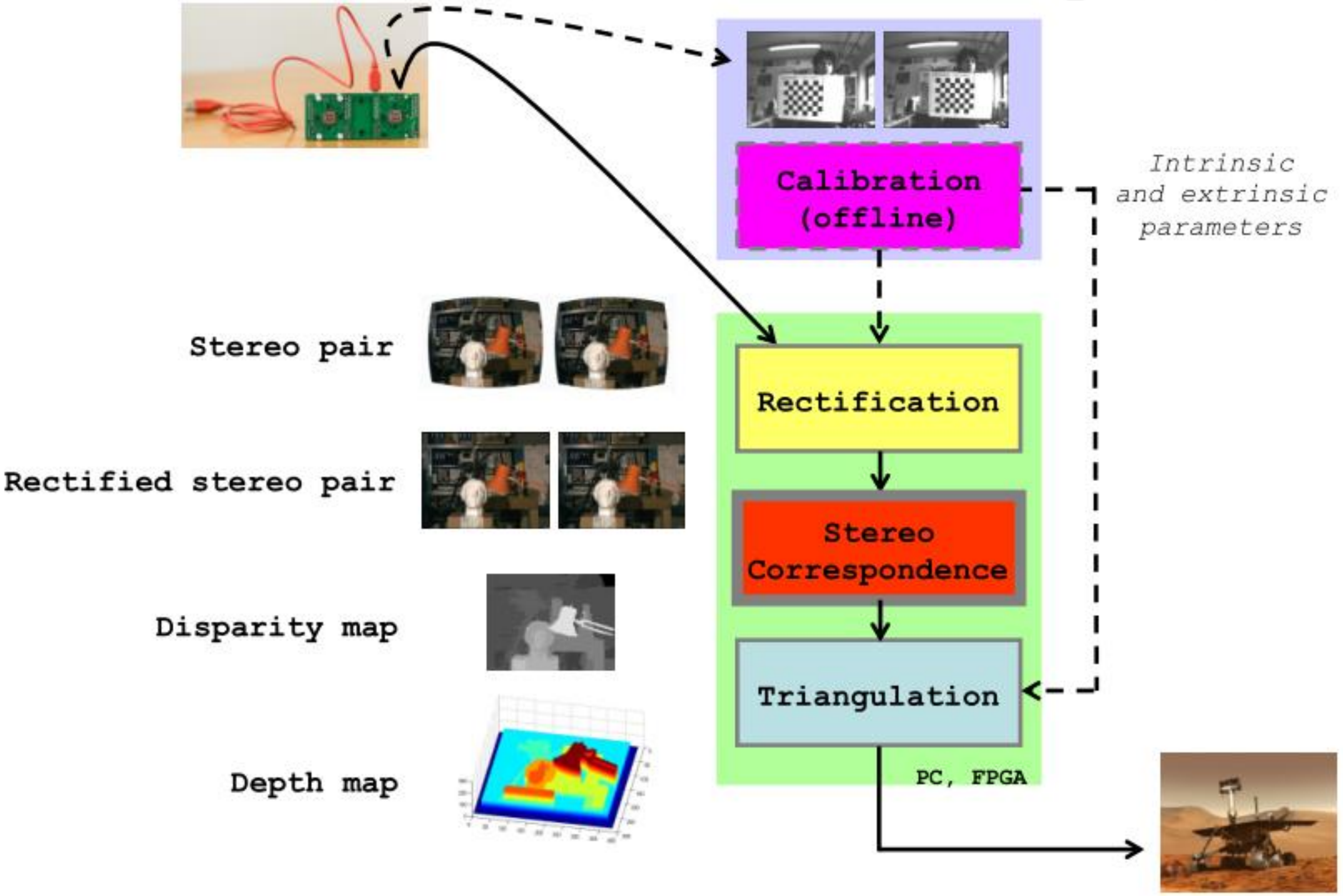
- The range field (horopter) using 5 disparity values $[d_{min}, d_{min}+4]$

Disparity and depth



- Using 5 disparity values $[\Delta+d_{min}, \Delta+d_{min}+4]$
- With $\Delta>0$, *horopter* gets closer and shrinks (depth and obviously area/volume)

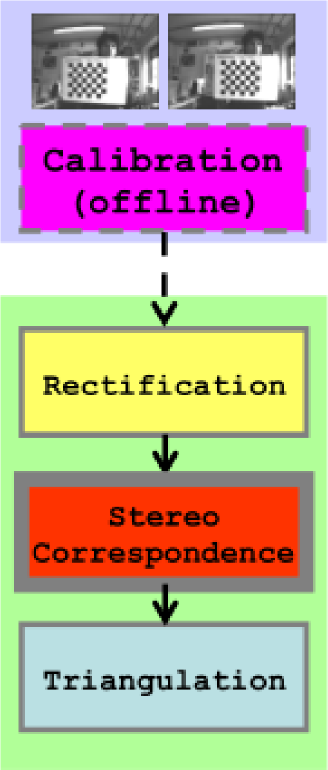
Overview of a stereo vision system



Stefano Mattoccia

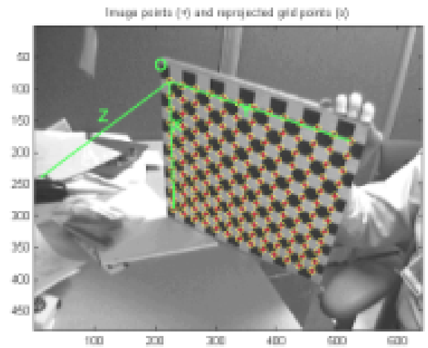
Calibration (offline)

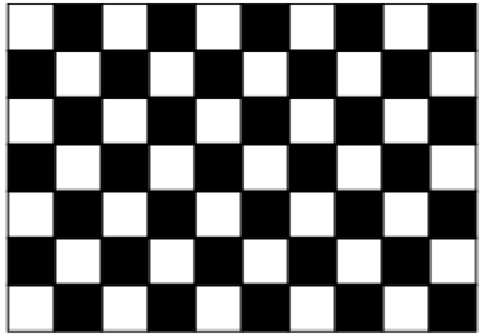
Offline procedure aimed at finding:



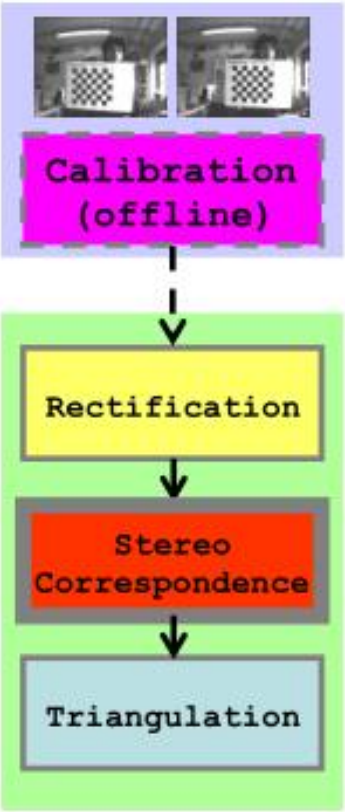
- Intrinsic parameters of the two cameras (focal length, image center, parameters of lenses distortion, etc)

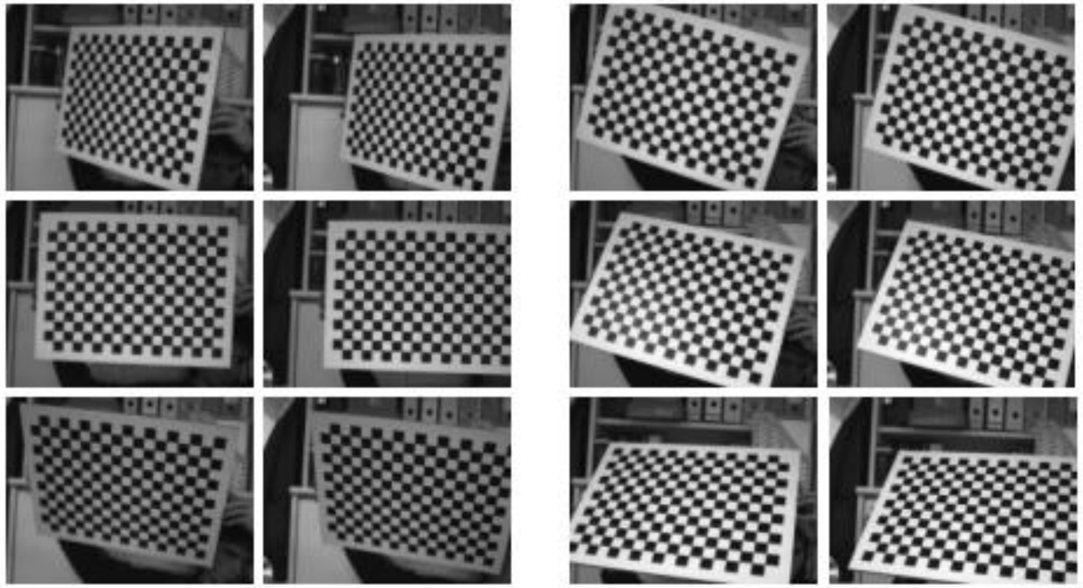
- Extrinsic parameters (R and T that aligns the two cameras)





Calibration is carried out acquiring and processing 10+ stereo pairs of a known pattern (typically a checkerboard)





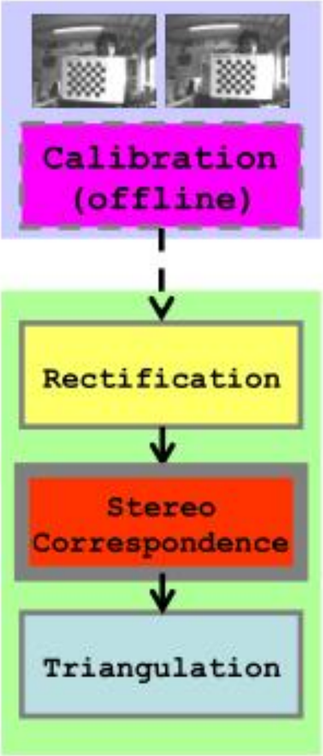
- Calibration is available in OpenCV [39] and Matlab [40]
- A detailed description of calibration can be found in [20,21,22]
- Next slides show 20 stereo pairs used for calibrating a stereo camera

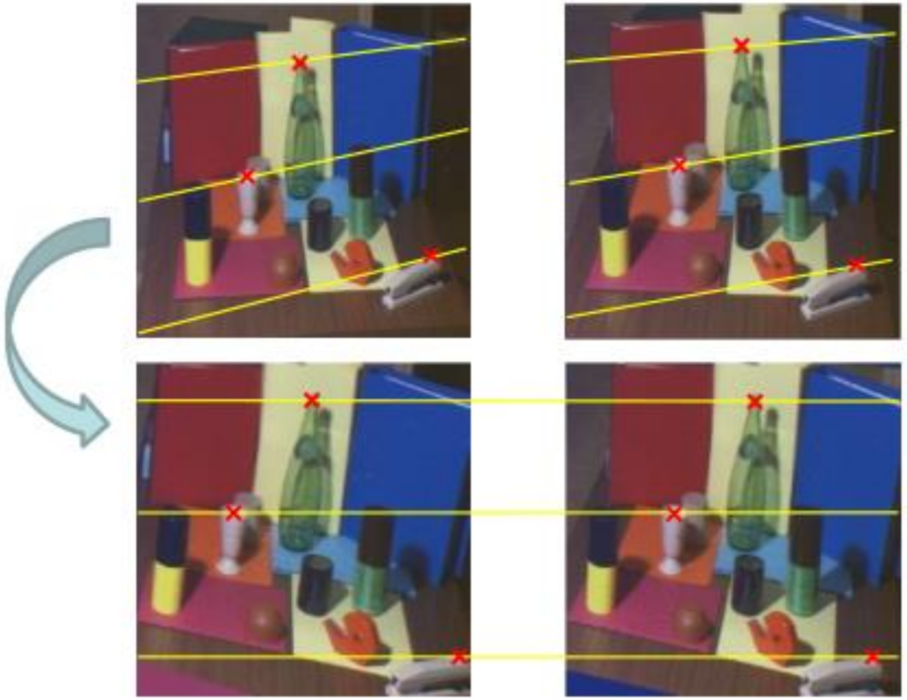
Rectification

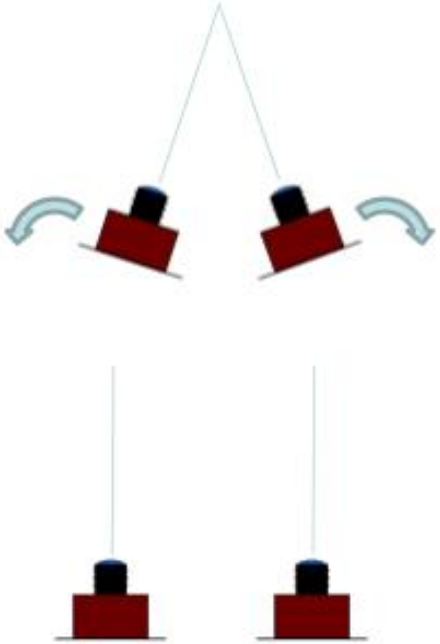
Using the information from the calibration step:

a) removes lens distortions

b) turns the stereo pair in standard form



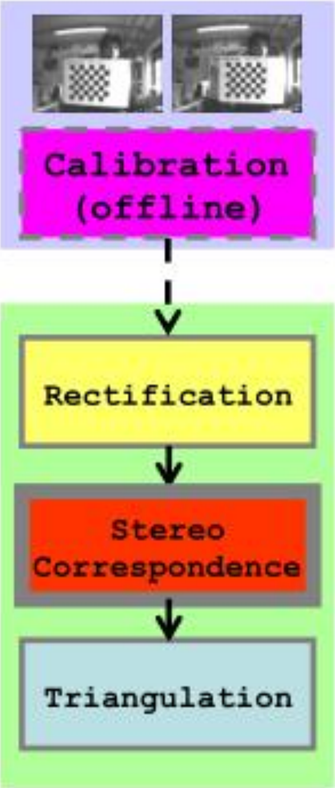


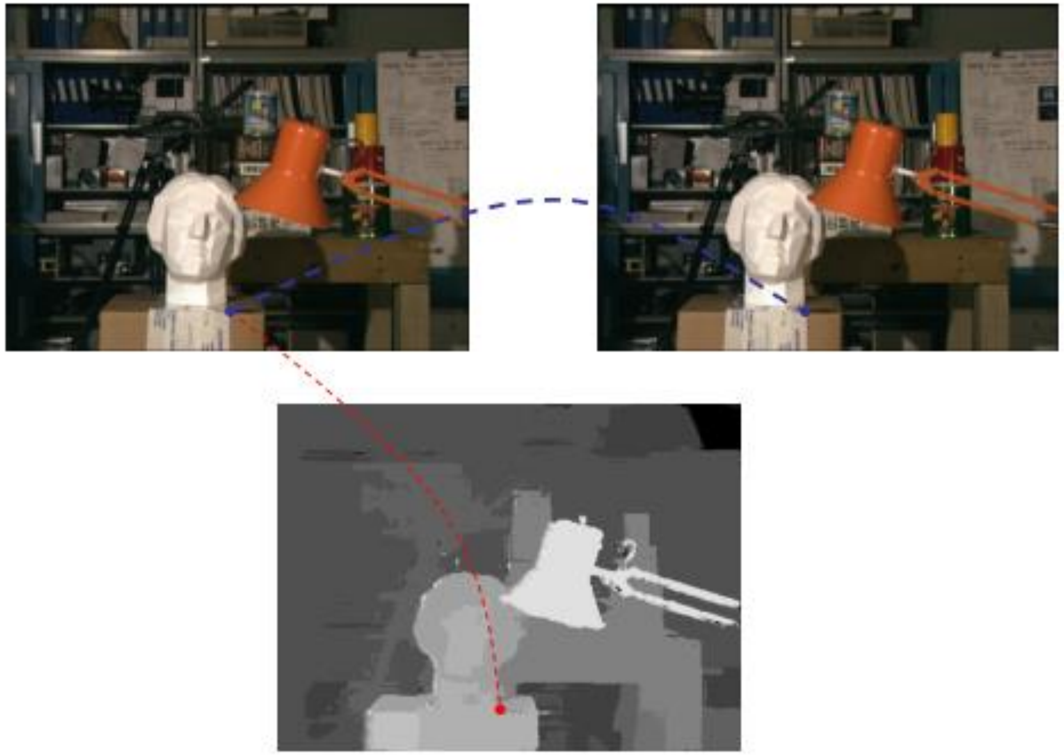


Stereo camera in standard form

Stereo correspondence

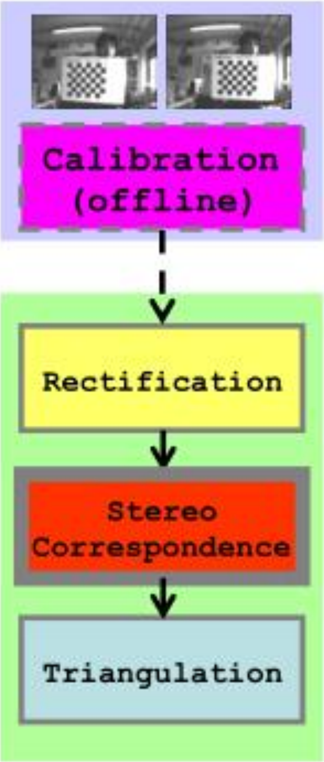
Aims at finding homologous points in the stereo pair.



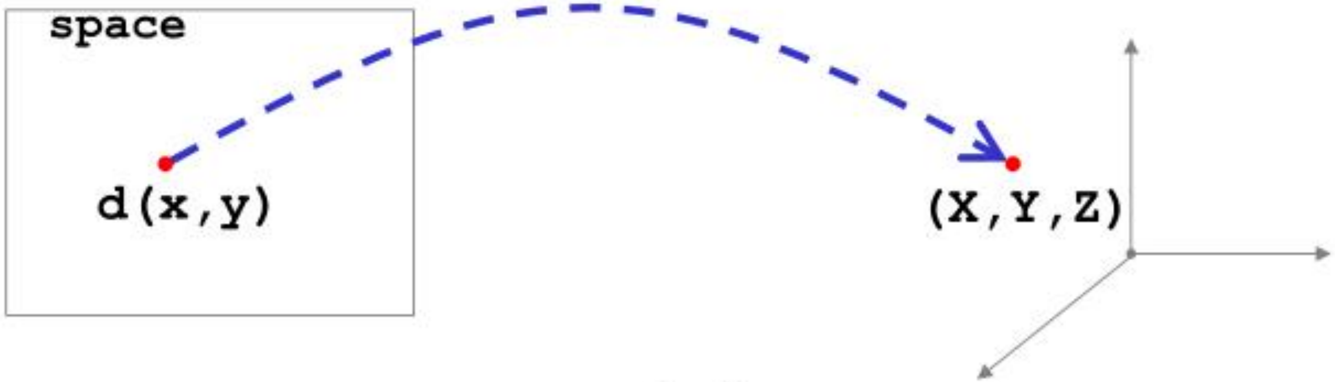


disparity map

Triangulation



Given the disparity map, the baseline and the Focal length (calibration): triangulation computes the position of the correspondence in the 3D space

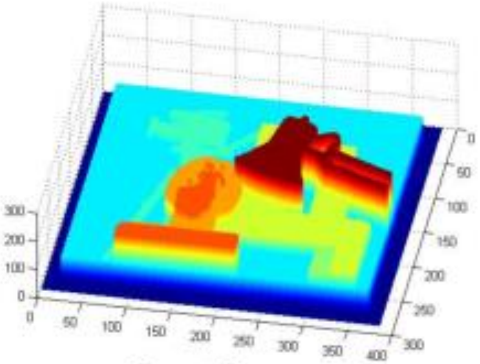


disparity map

$$Z = \frac{b \cdot f}{d}$$

$$X = Z\frac{x_R}{f}$$

$$Y = Z\frac{y_R}{f}$$



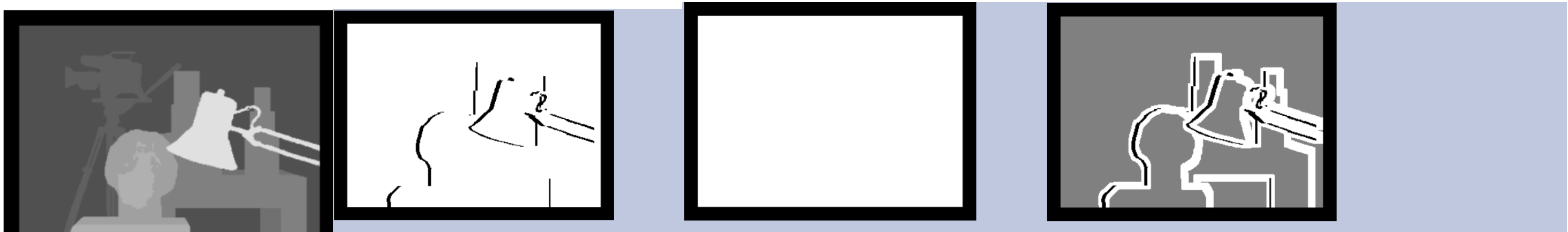
depth map

评价准则：

在三个区域*，不同指定视差阈值下计算错误像素点的平均个数作为排名依据：

***三个区域**

1、non-occluded 无遮挡区域；
2、全部区域；
3、不连续区域（depth discontinuities）



Non-occluded regions (white) and occluded and border regions (black).
Errors are only evaluated in the white regions.

All (including half-occluded) regions (white) and border regions (black).
Errors are only evaluated in the white regions.

Regions near depth discontinuities[1] (white), occluded and border regions (black), and other regions (gray).
Errors are only evaluated in the white regions.

***多个指标：**

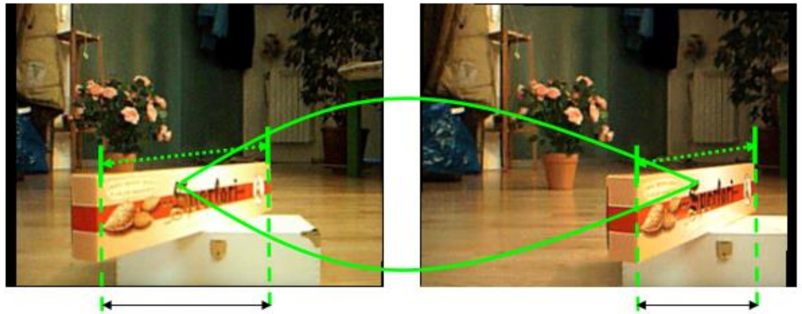
Out-Noc: Percentage of erroneous pixels in non-occluded areas
Out-All: Percentage of erroneous pixels in total
Avg-Noc: Average disparity / end-point error in non-occluded areas
Avg-All: Average disparity / end-point error in total
Density: Percentage of pixels for which ground truth has been provided by the method

Why is stereo correspondence so challenging?

Photometric distortions and noise

Specular surfaces

Foreshortening

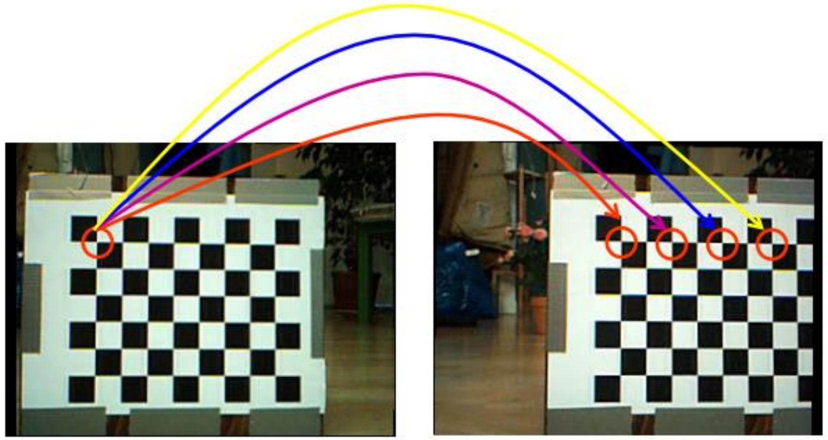


Perspective distortions

Uniform/ambiguous regions

Why is stereo correspondence so challenging?

Repetitive/ambiguous patterns

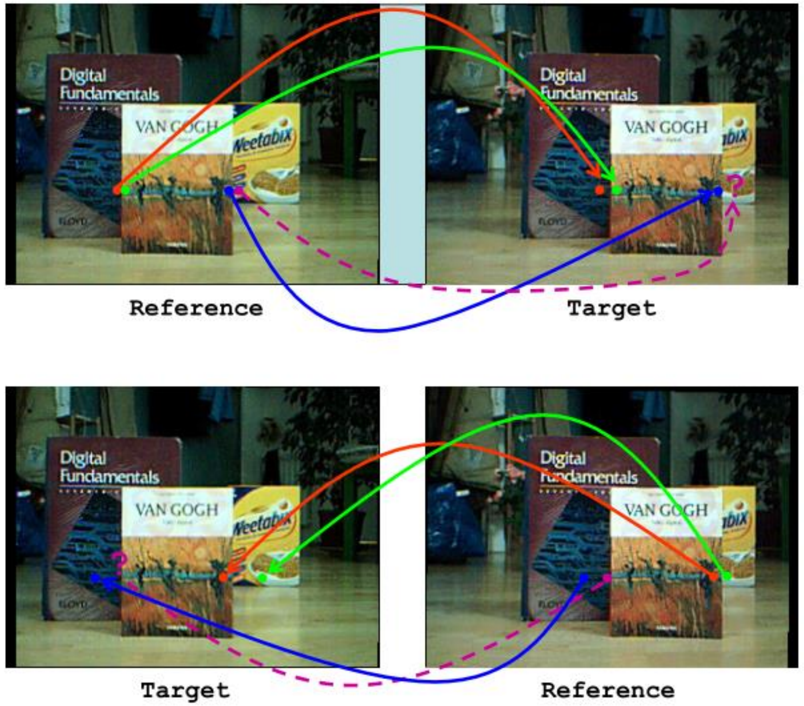


How to reduce ambiguity... ?

Transparent objects

Occlusions and discontinuities 1/2

Occlusions and discontinuities 2/2

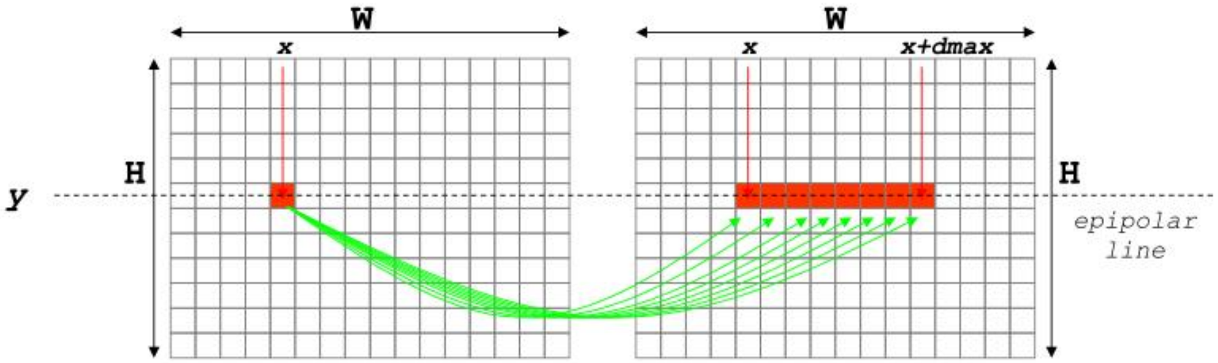


Stereo correspondence

The simplest (naive and unused) local approach:

Reference (R)

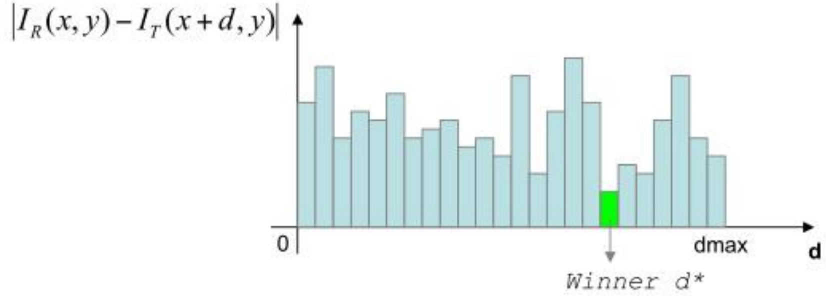
Target (T)



Reference (R)

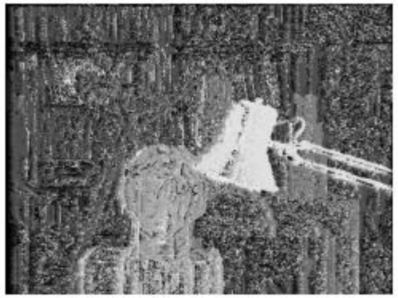
Target (T)

- matching cost (1): pixel-based absolute difference between pixel intensities
- disparity computation (3): Winner Takes All (WTA)



Reference

Groundtruth



Result (disappointing)

Stereo correspondence

How to improve the results of the naive approach ?

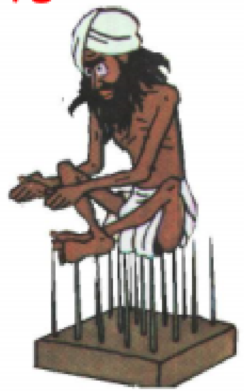
Basically exist two different (not mutually exclusive) strategies:

- Local algorithms use the simple WTA disparity selection strategy but reduce ambiguity (increasing the signal to noise ratio (SNR)) by aggregating matching costs over a support window (aka kernel or correlation window). Sometime a smoothness term is adopted. Steps 1+2 (+ WTA)

- Global (and semi-global*) algorithms search for disparity assignments that minimize an energy function over the whole stereo pair using a pixel-based matching cost (sometime the matching cost is aggregated over a support). Steps 1+3

Both approaches assume that the scene is piecewise smooth. Sometime this assumption is violated...

This hypothesis is implicitly assumed by local approaches while it is explicitly modelled by global approaches



reference

1. 《计算机视觉中的数学方法》 吴福朝 著
2. 《计算机视觉-摄像机标定》 高伟 中国科学院自动化研究所
3. “Multiple View Geometry”, Richard Hartley and Andrew Zisserman, CVPR June 1999
4. “Stereo Vision:Algorithms and Applications”, Stefano Mattoccia, Department of Computer Science (DISI) University of Bologna
5. Z. Zhang, “A flexible new technique for camera calibration”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11), 1330-1334.