

# Machine Vision: Stereo Vision

Wu Wei

[ww@hdu.edu.cn](mailto:ww@hdu.edu.cn)



# 立体视觉

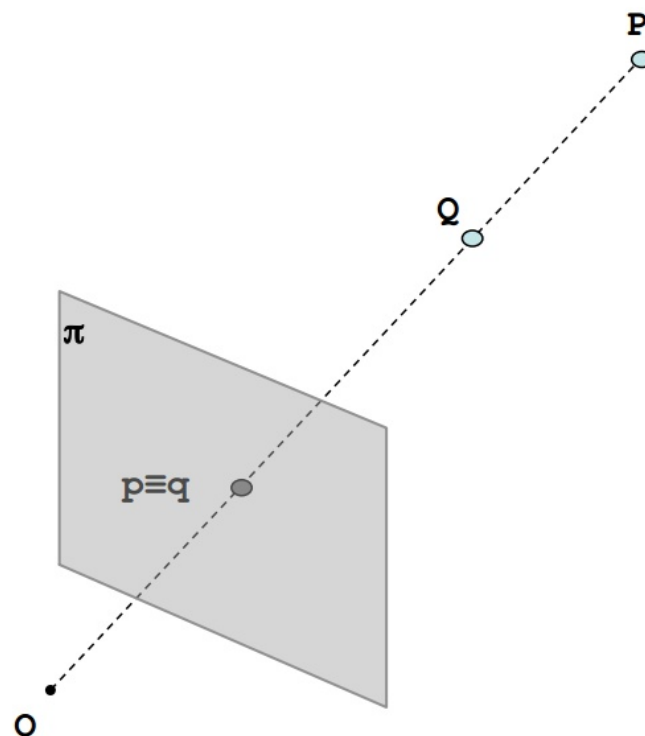
---

- 能够从双目或者多目相机中提取深度图像的技术
- 与双目立体视觉系统、稠密立体算法、立体视觉应用有关
- 能够实时或者硬件实现



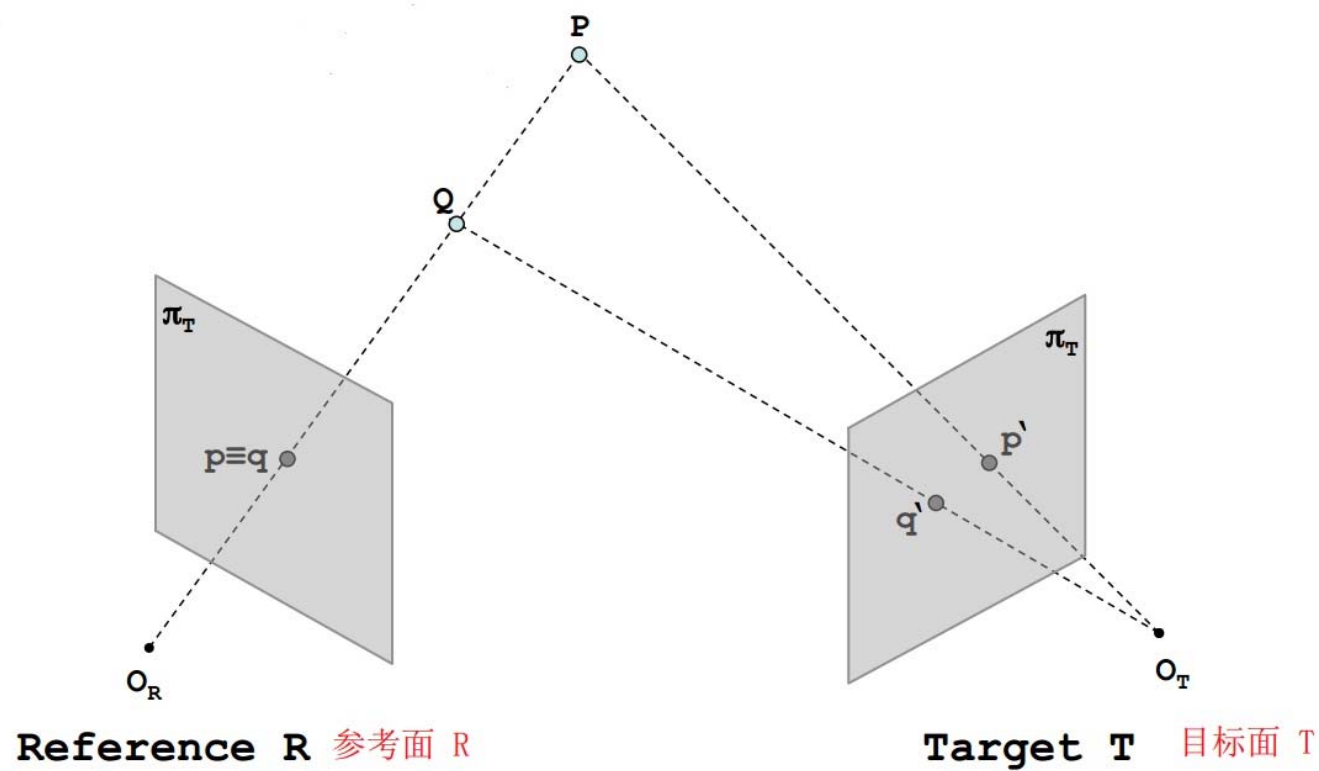
# 单目相机？

---



# 单目相机？

---



# 双目立体

---

## ■ 双目立体视觉深度相机简化流程

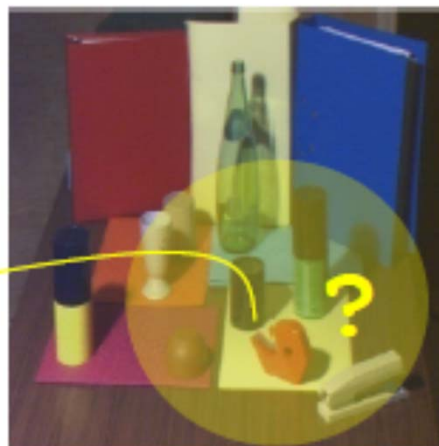
- 首先需要对双目相机进行标定，得到两个相机的内外参数、单应矩阵。
- 根据标定结果对原始图像校正，校正后的两张图像位于同一平面且互相平行。
- 对校正后的两张图像进行像素点匹配。
- 根据匹配结果计算每个像素的深度，从而获得深度图。



# 双目相机



Reference (R)

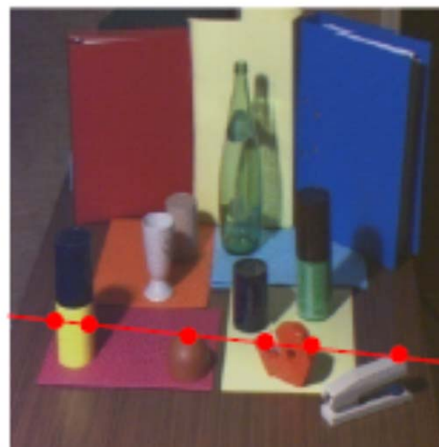


Target (T)

2D search domain ?



Reference (R)

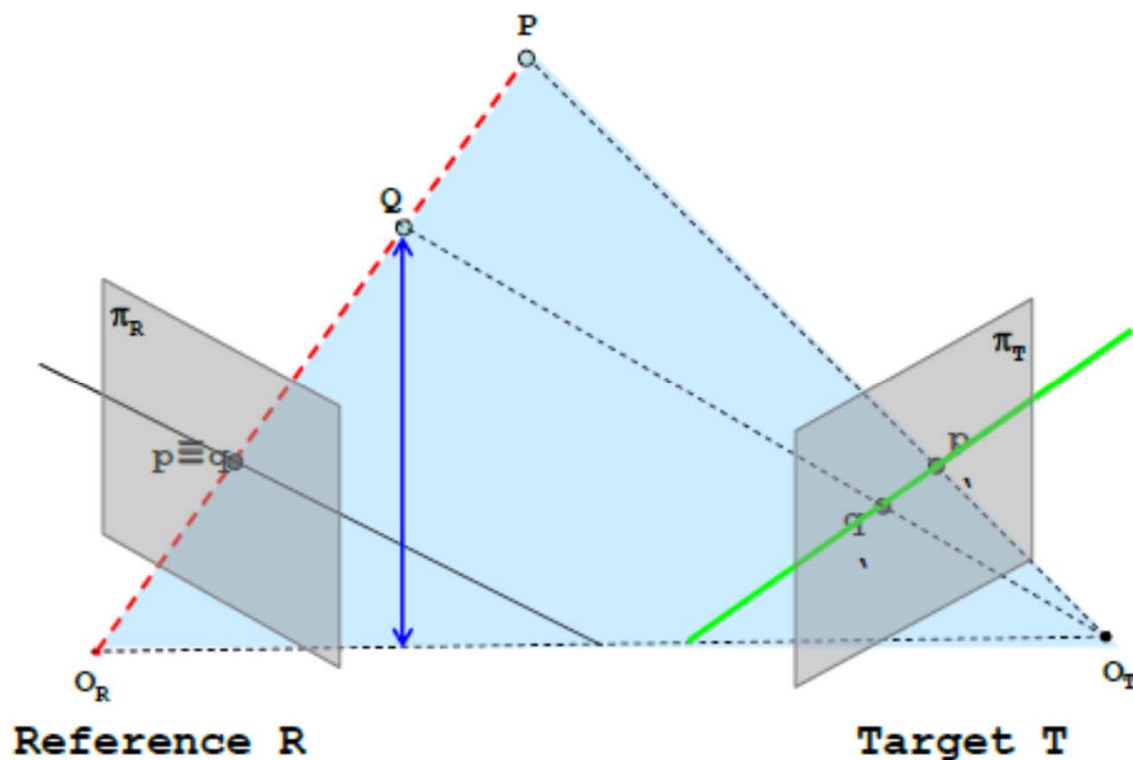


Target (T)

No!! Thanks to the  
epipolar constraint



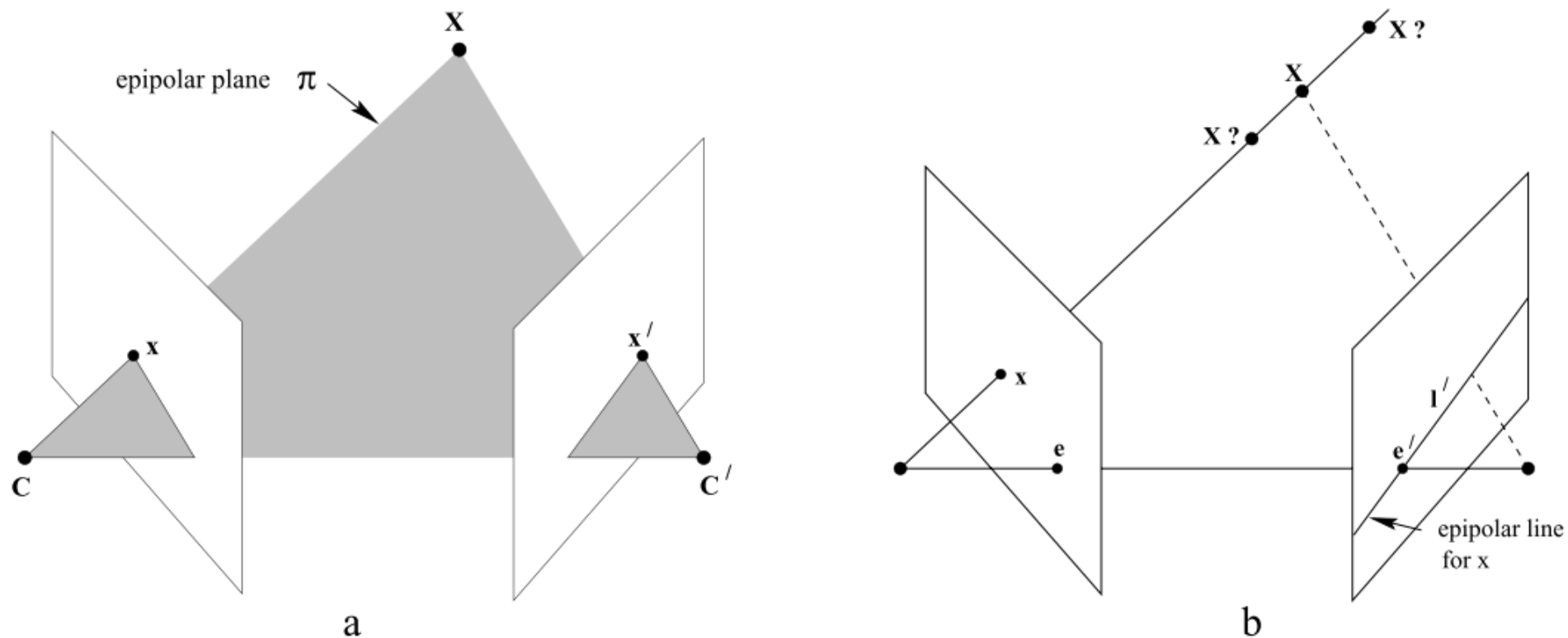
# 极线约束



- 对于参考图像R而言，现实场景中的P与Q点在其像平面 $\pi_R$ 上被投影成为一个点 $p=q$ 。
- 极线约束规定，属于（红色）视线的点对应位于目标图像T的图像平面 $\pi_T$ 上的绿线上。



# 对极几何

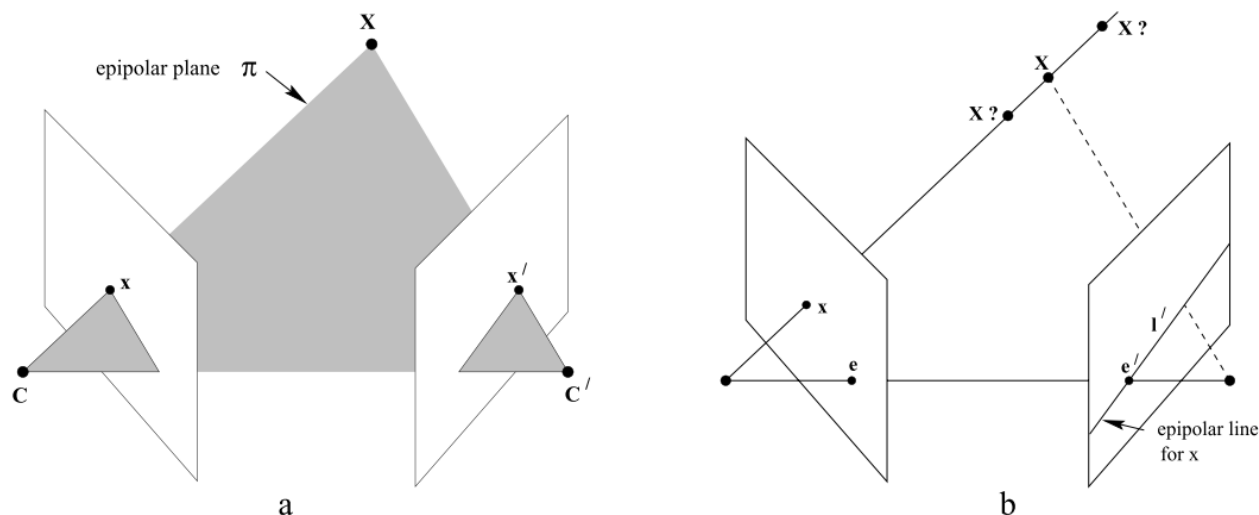


上图a中两个摄像机的中心分别是 $C$ 与 $C'$ ，而 $X$ 为一个三维空间点，它在两个摄像机的成像平面上的投影点分别是 $x$ 与 $x'$ 。





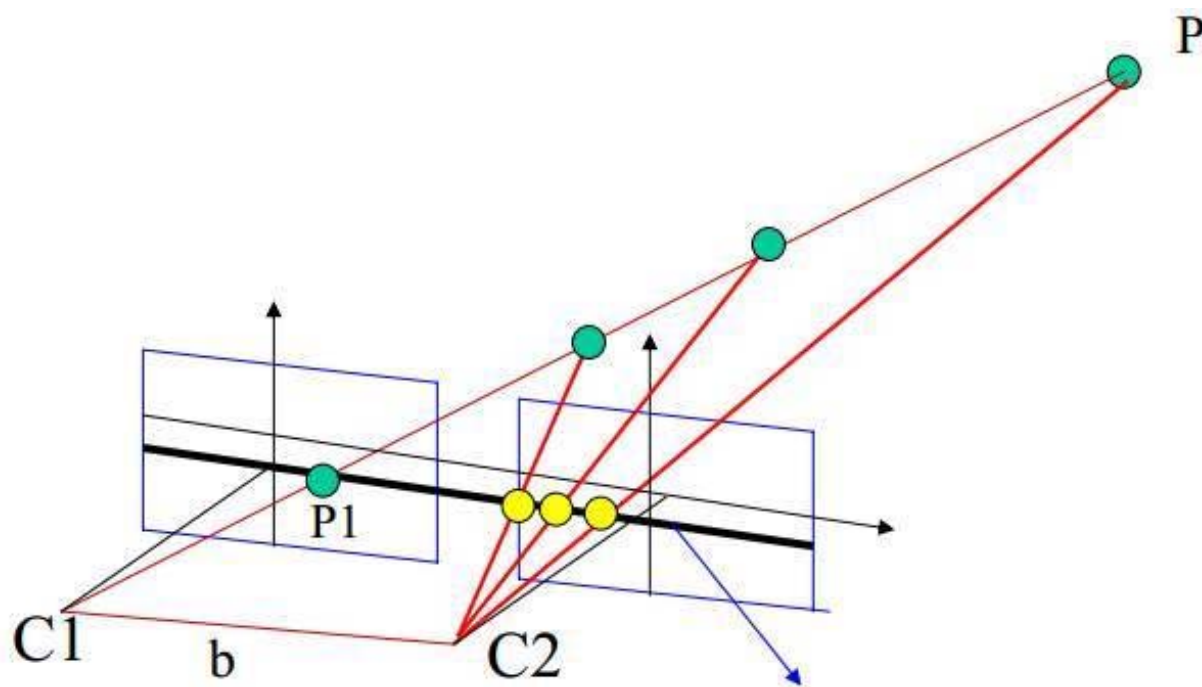
# 对极几何



- 基线：两个摄像机光心的连线 $CC'$ 。
- 对极点：上图b中的 $e$ 与 $e'$ ，它们分别是一幅视图中另一个摄像机中心的像。二维表示为基线 $CC'$ 与两个成像平面的交点。
- 对极平面：是一张包含基线的平面，存在着对极平面的一个集合（以基线为轴转动），上图中的一个例子就是 $CXC'$ 。
- 对极线：对极平面与图像平面的交线。上图中的例子是 $xe$ 与 $x'e'$ ，一个成像平面上的所有的对极线相交于对极点。



# 极线约束



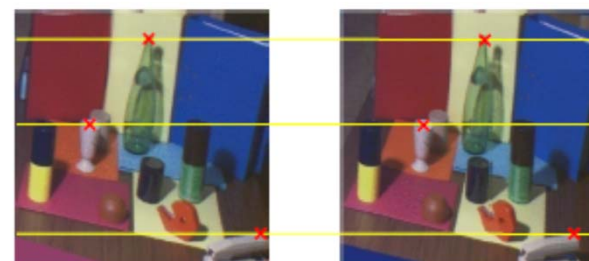
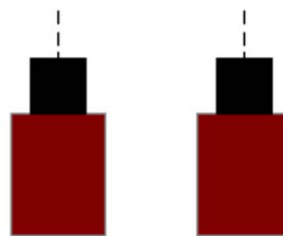
对于一幅图像上的每一点 $P_1$ ，在第二幅图像上，任何与该点 $P_1$ 匹配的点 $P'$ 必然在对极线 $l'$ 上，该对极线是过点 $P_1$ 与第一个摄像机中心 $C_1$ 的射线在第二幅图像上的投影。



# 极线约束

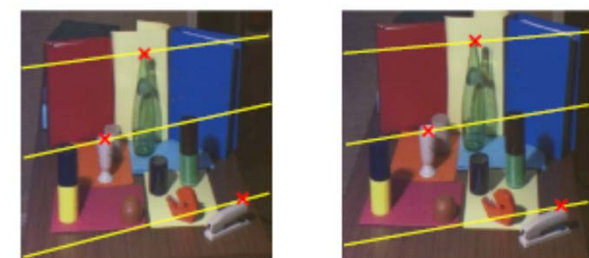
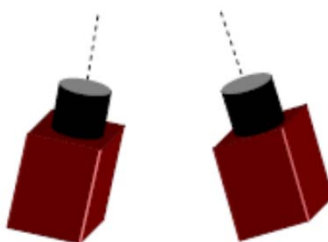
- 对于不同的立体摄像机模型，视差矢量的结构也有所不同

- **平行式立体视觉模型：** 由于两摄像机的光轴互相平行，导致左右图像间的对应极线相互平行，且位于相同的图像水平扫描线上。因而视差矢量平行于图像的水平扫描线，此时视差矢量实际上退化为标量。



Stereo pair in standard form

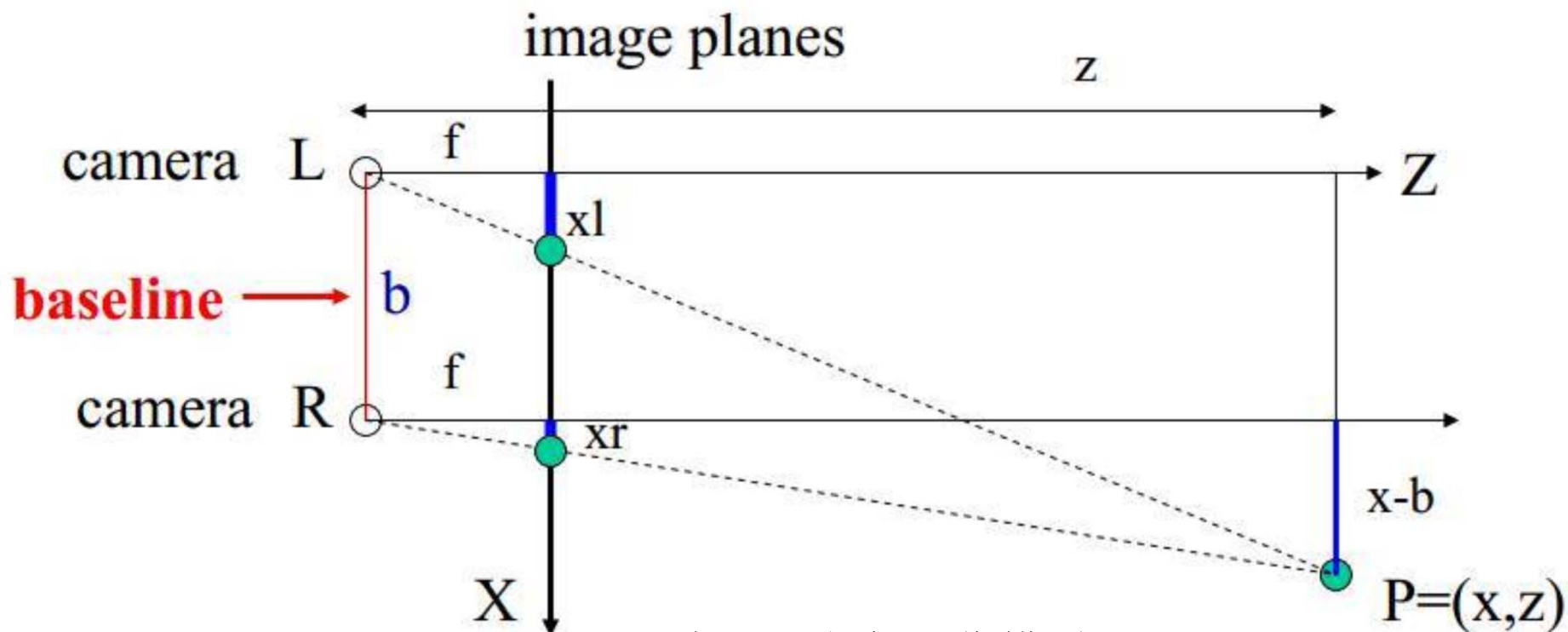
- **汇聚式视觉模型：** 左右图像间的对应极线不再平行，因而视差矢量也就不平行于图像水平扫描线，视差矢量的两分量都不为零，如下图所示。



Original stereo pair



# 深度与视差



理想双目相机成像模型

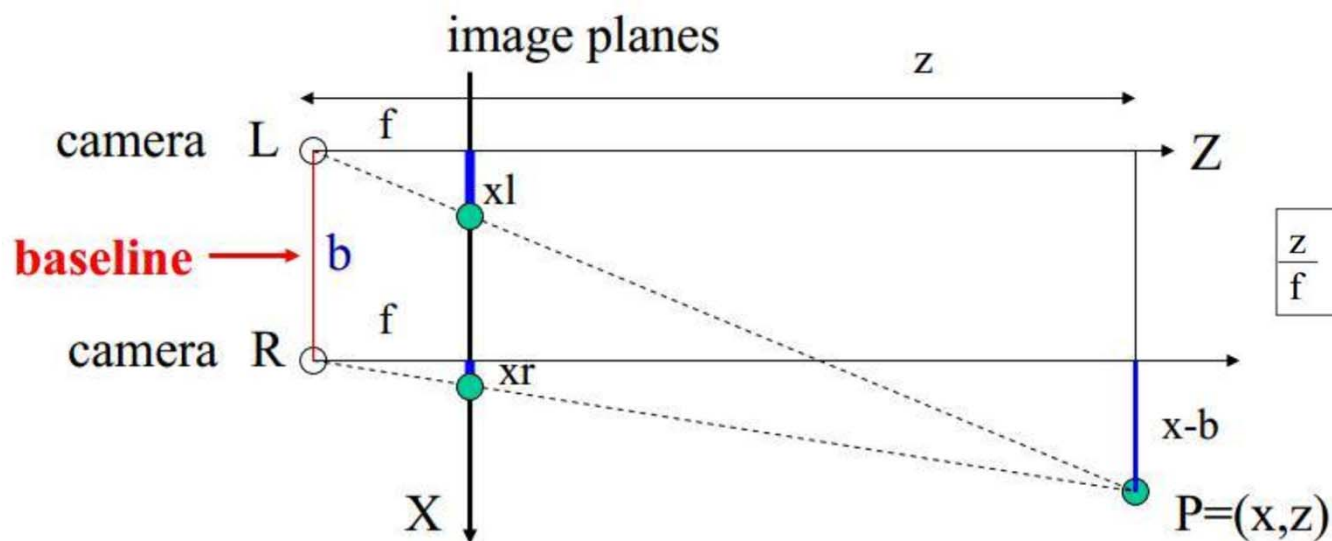
$$\frac{z}{f} = \frac{x}{x_l}$$

$$\frac{z}{f} = \frac{x-b}{x_r}$$

$$\frac{z}{f} = \frac{y}{y_l} = \frac{y}{y_r}$$



# 深度与视差



$$\frac{z}{f} = \frac{x}{x_l}$$

$$\frac{z}{f} = \frac{x-b}{x_r}$$

$$\frac{z}{f} = \frac{y}{y_l} = \frac{y}{y_r}$$

$$\text{Depth } z = f \cdot b / (x_l - x_r) = f \cdot b / d$$

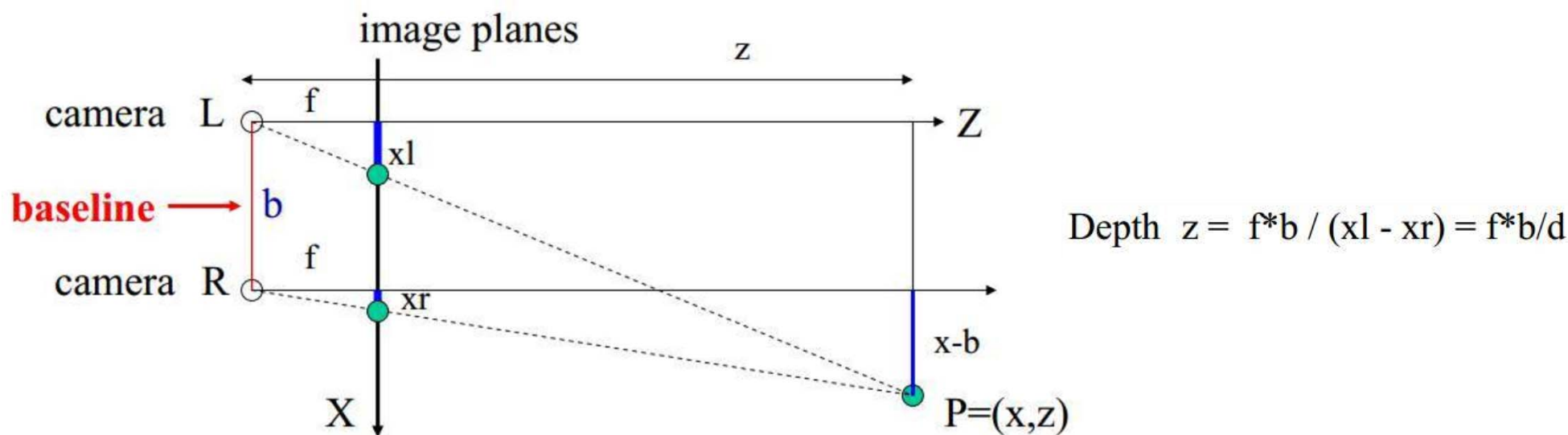
$$x = x_l \cdot z / f \quad \text{or} \quad b + x_r \cdot z / f$$

$$y = y_l \cdot z / f \quad \text{or} \quad y_r \cdot z / f$$

Y-axis is  
perpendicular  
to the page



# 深度与视差



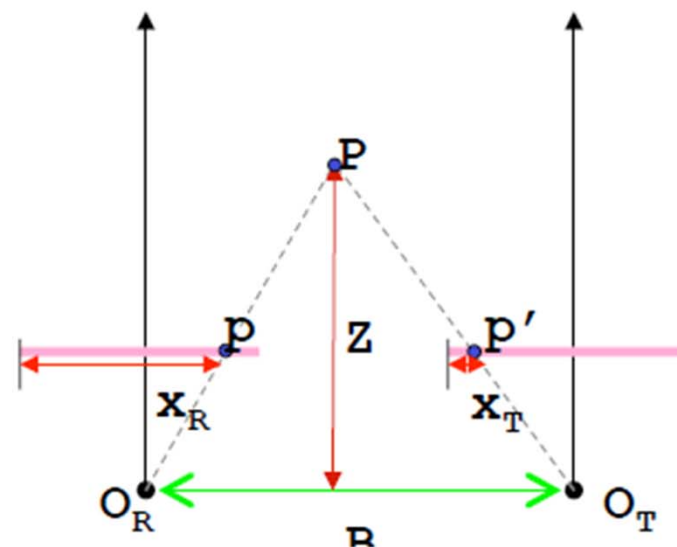
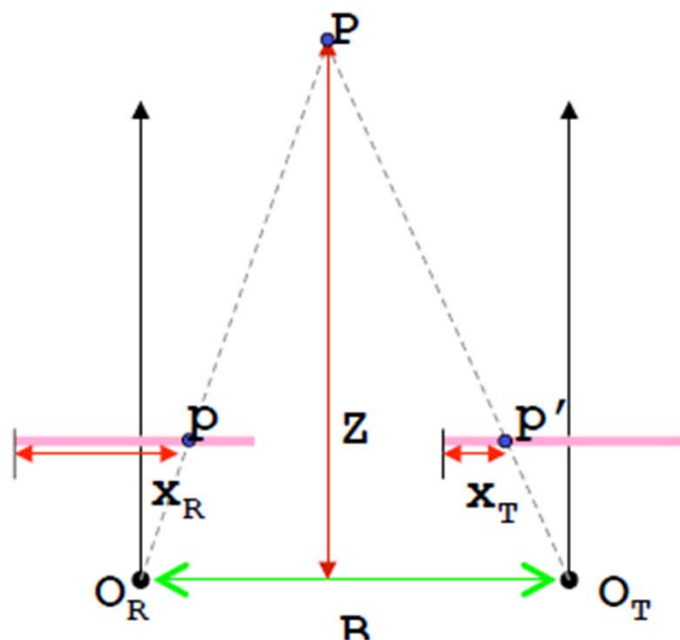
理想情况下双目立体视觉相机深度值计算原理

- 相机焦距 $f$ ，左右相机基线 $b$ 。
- 视差 $d$ 。需要知道左相机的每个像素点 $(x_l, y_l)$ 和右相机中对应点 $(x_r, y_r)$ 的对应关系。



# 深度与视差

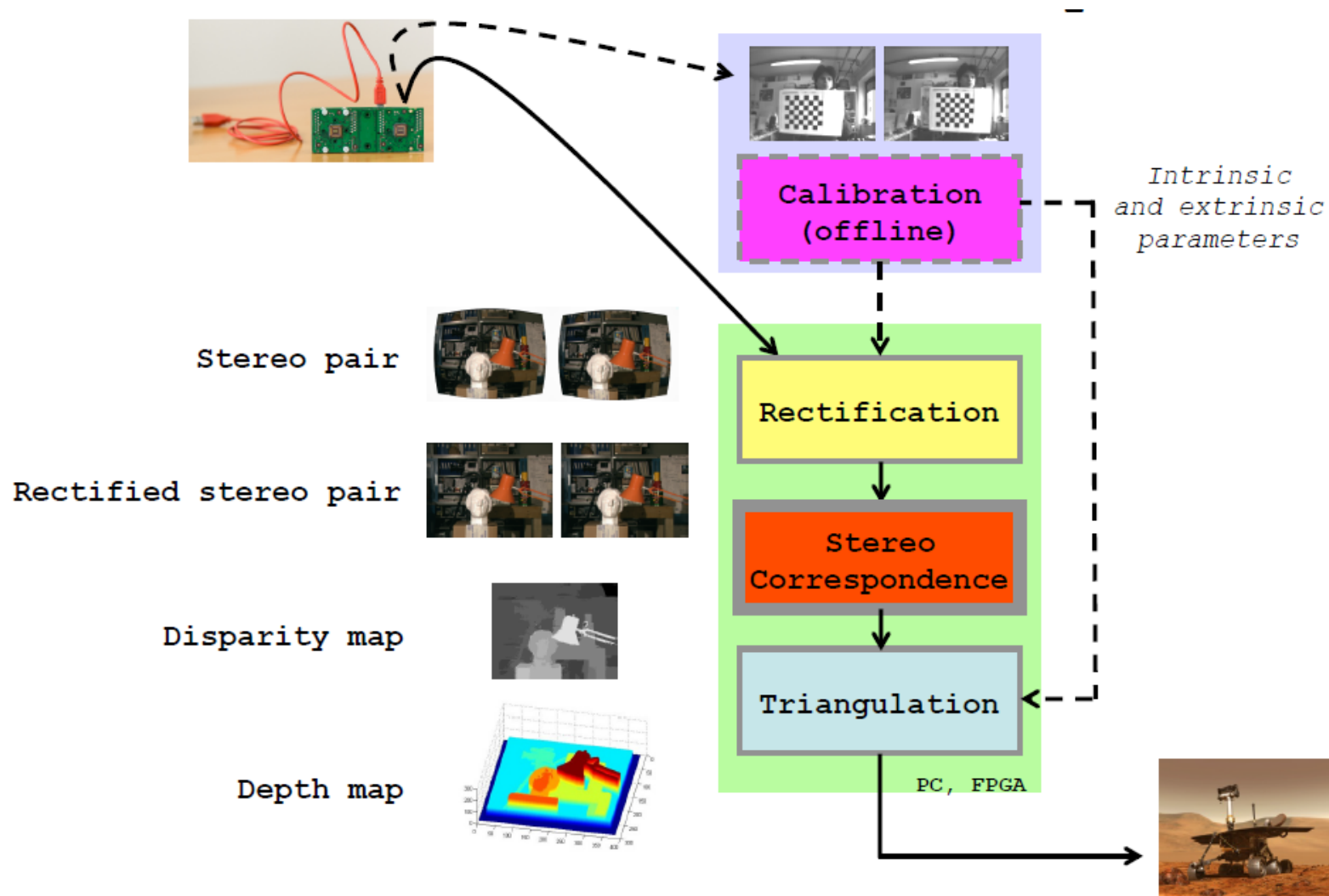
---



物体距离相机越近，视差越大



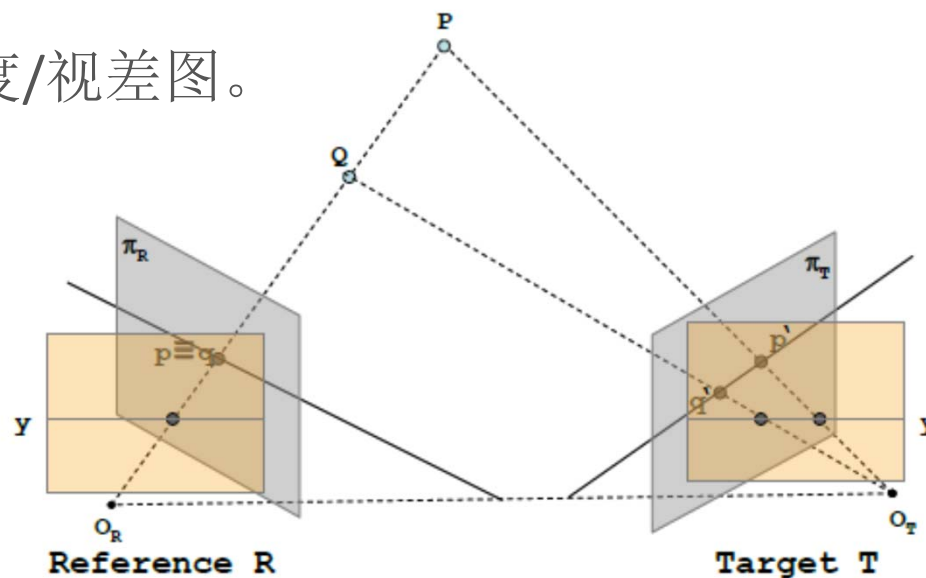
# 视差估计的基本流程





# 相机标定

- 通过双摄设备采集图像，此时图像是存在镜头畸变的，在进行扫描线对齐之前要进行离线标定以消除镜头畸变。扫描线对齐的过程叫做镜头矫正（rectification），经过这步之后就可以进行1D的匹配点搜索（stereo correspondence）了。随后通过三角形相似原理得到相应的深度/视差图。



# 双目校正

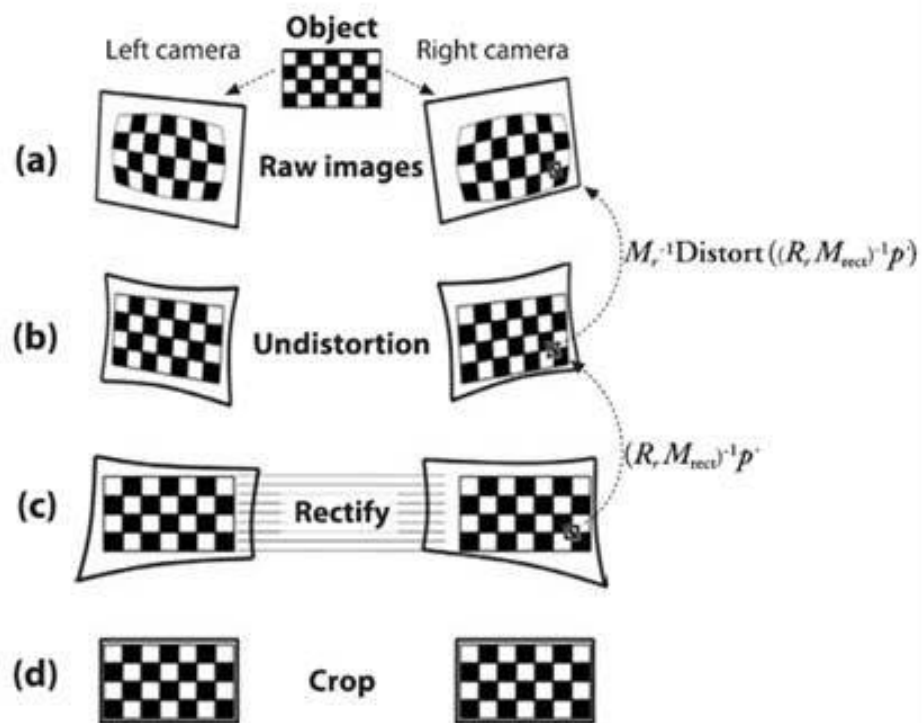
## 双目校正

原始视图

消除畸变

双目校正

图像剪裁



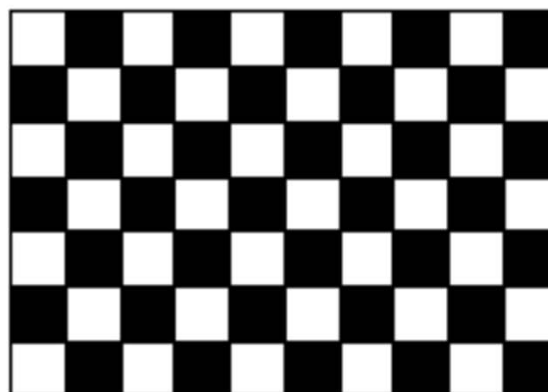
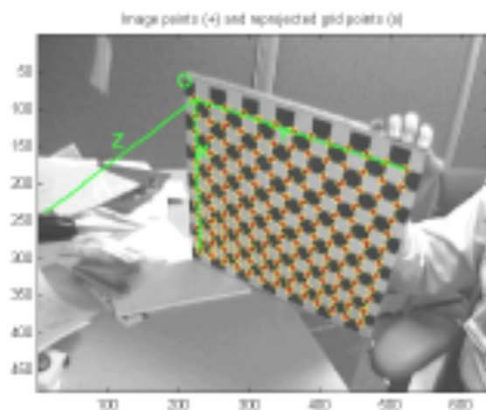
# 离线标定

---

## ■ 标定的目标:

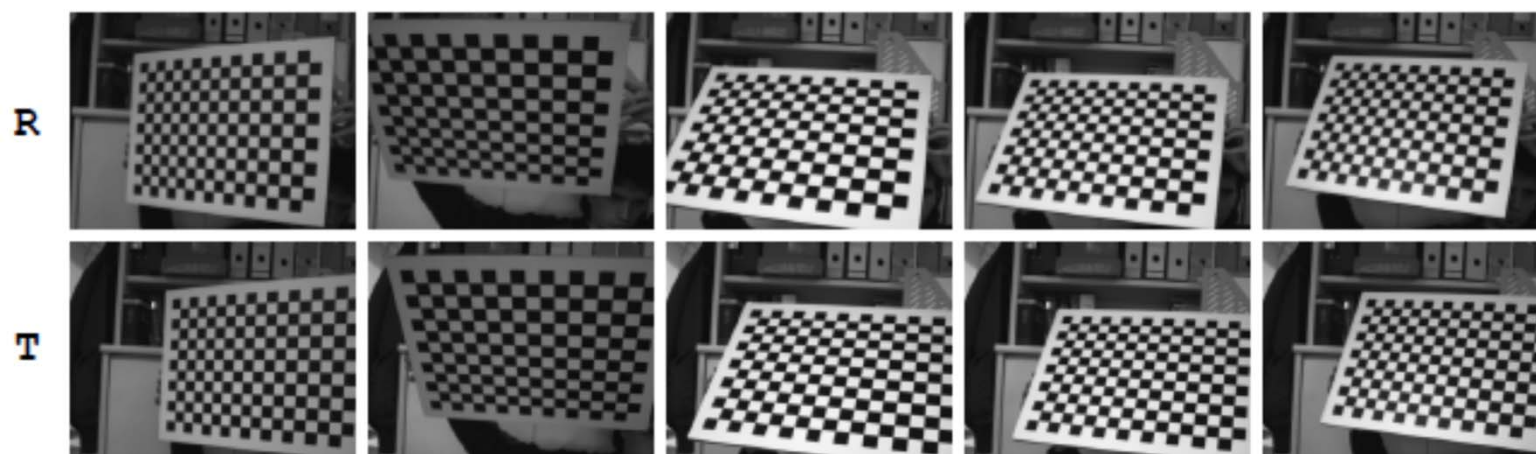
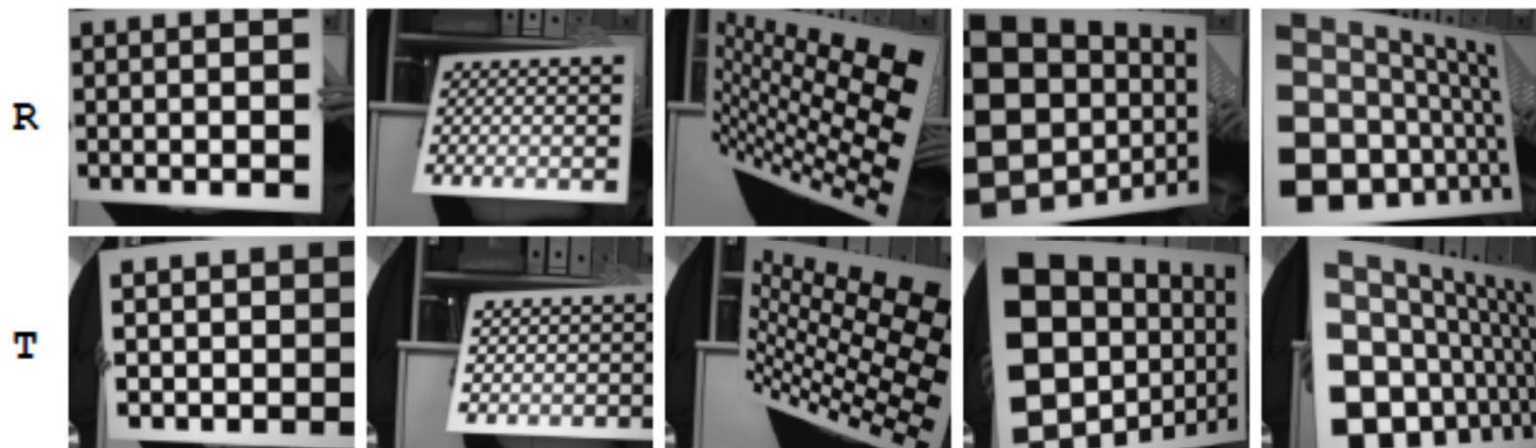
- 相机内参：焦距、图像中心、镜头畸变参数
- 相机外参：排列相机使其对齐的参数

（注意的是，相机标定的话一般需要10对以上的图像）

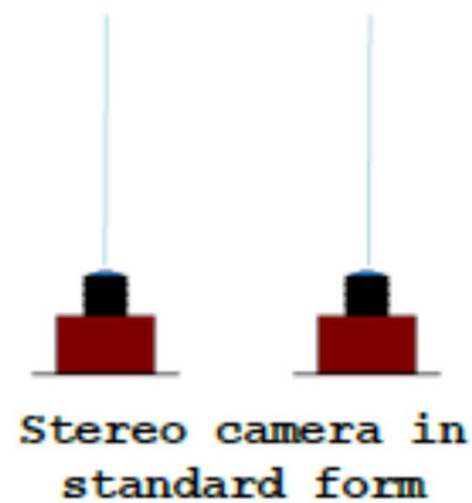
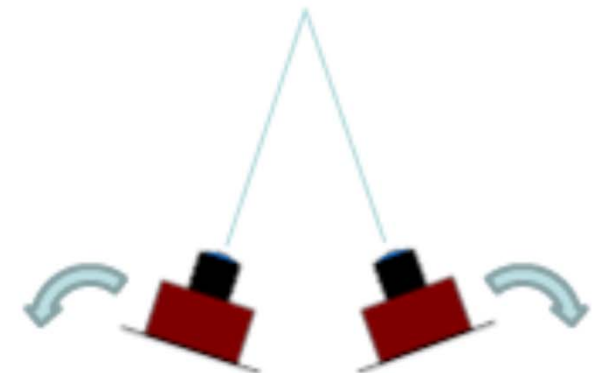
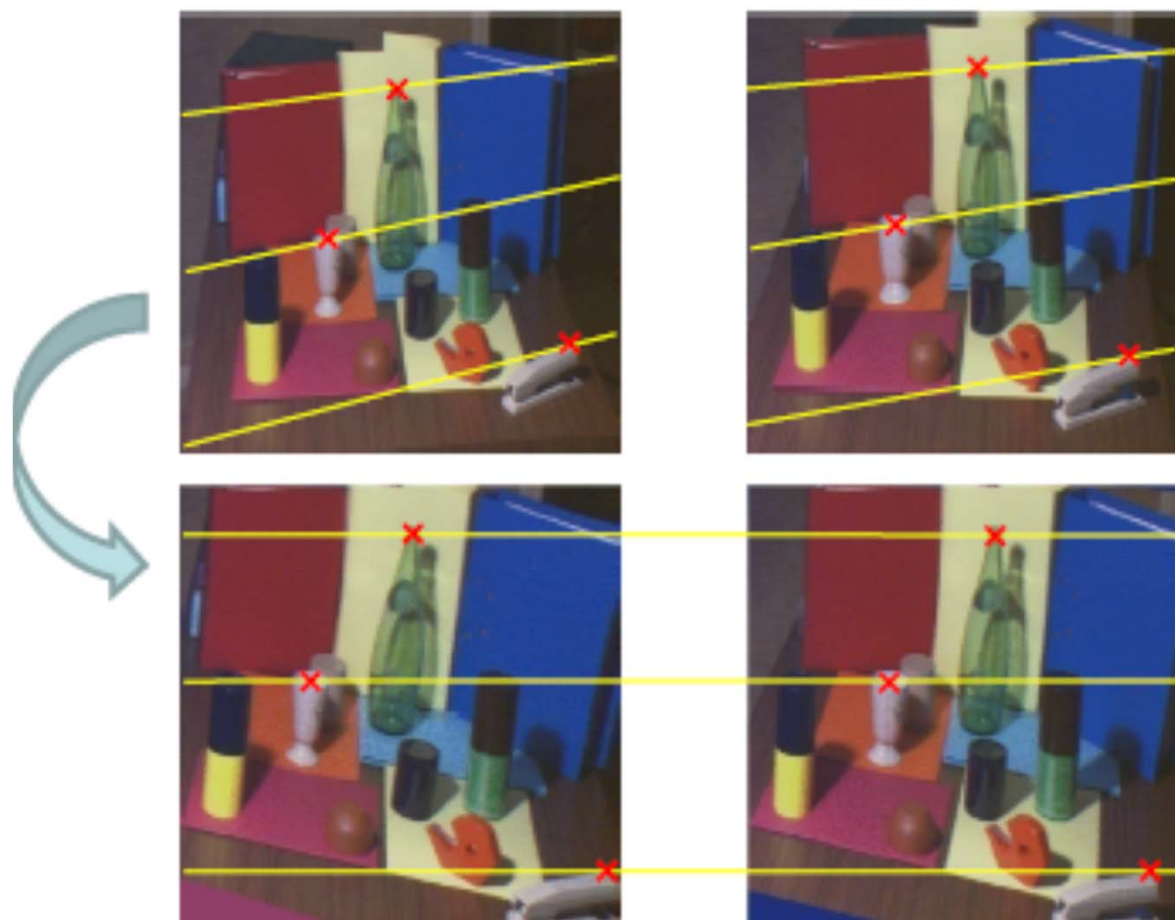


# 离线标定

---



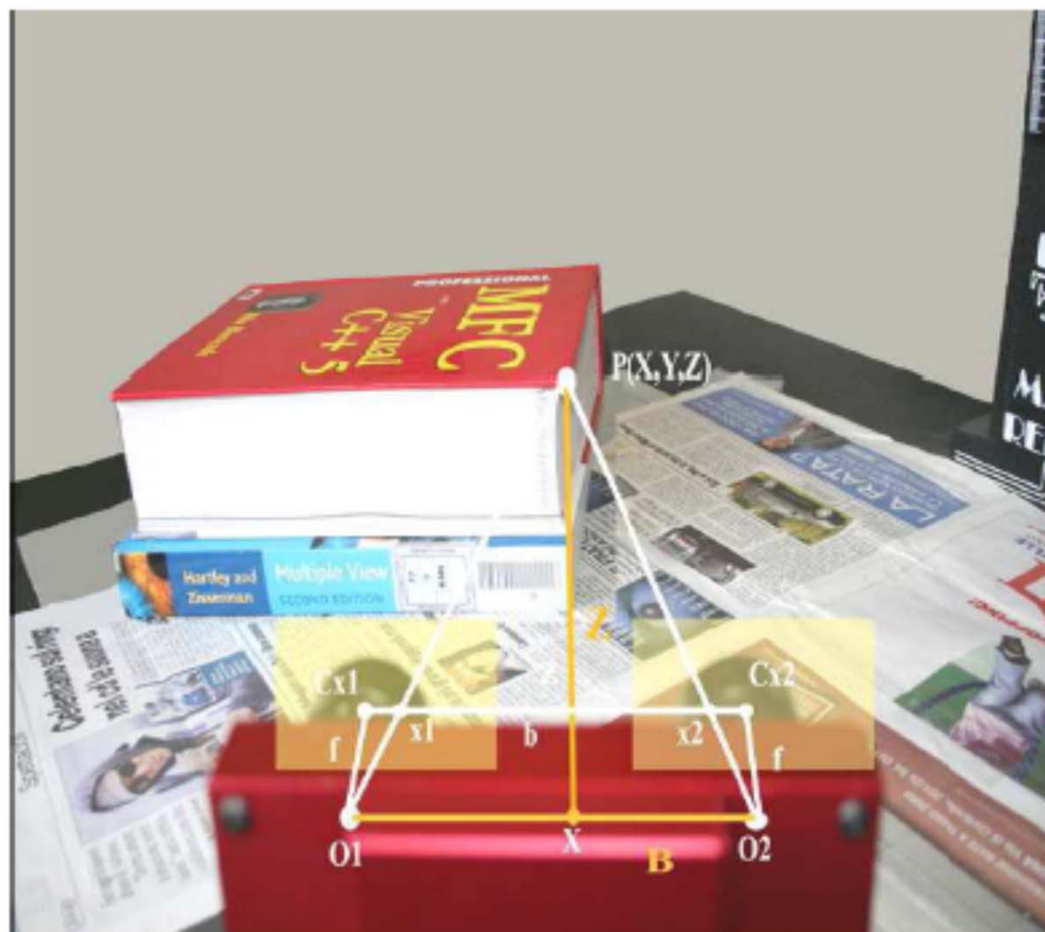
# 匹配校正





# 深度与视差

---



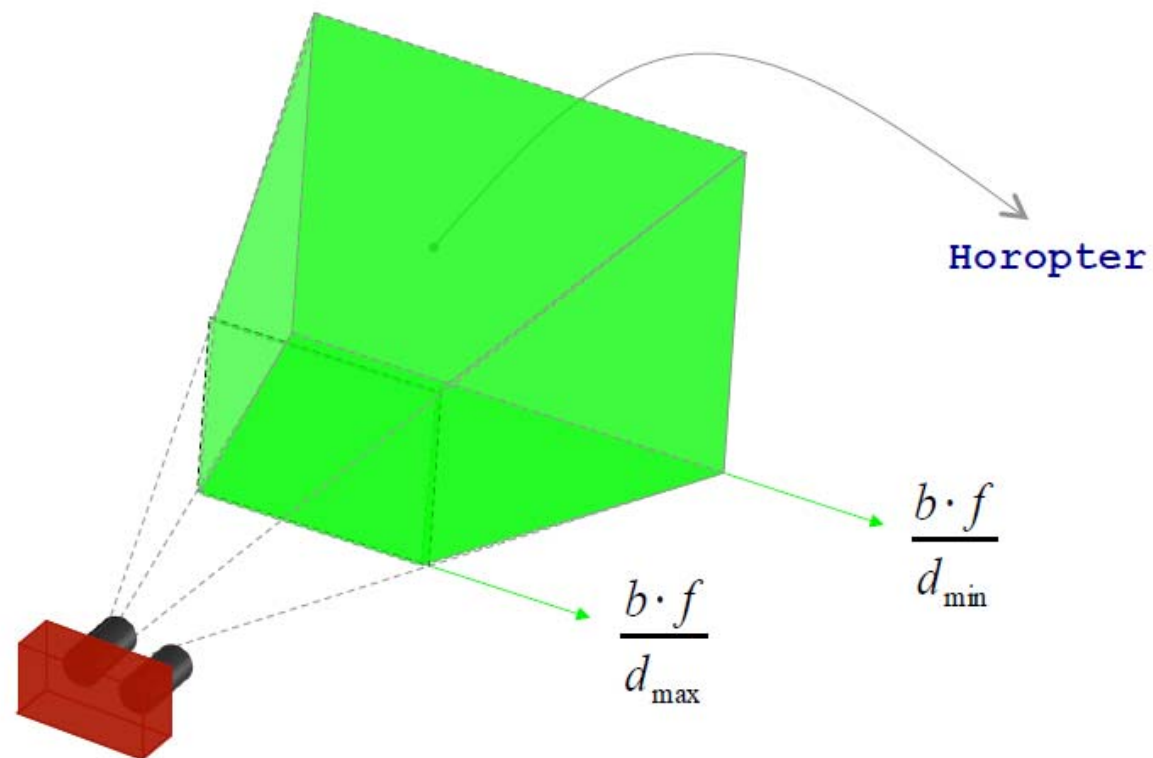
Disparity and depth

---



# 视界

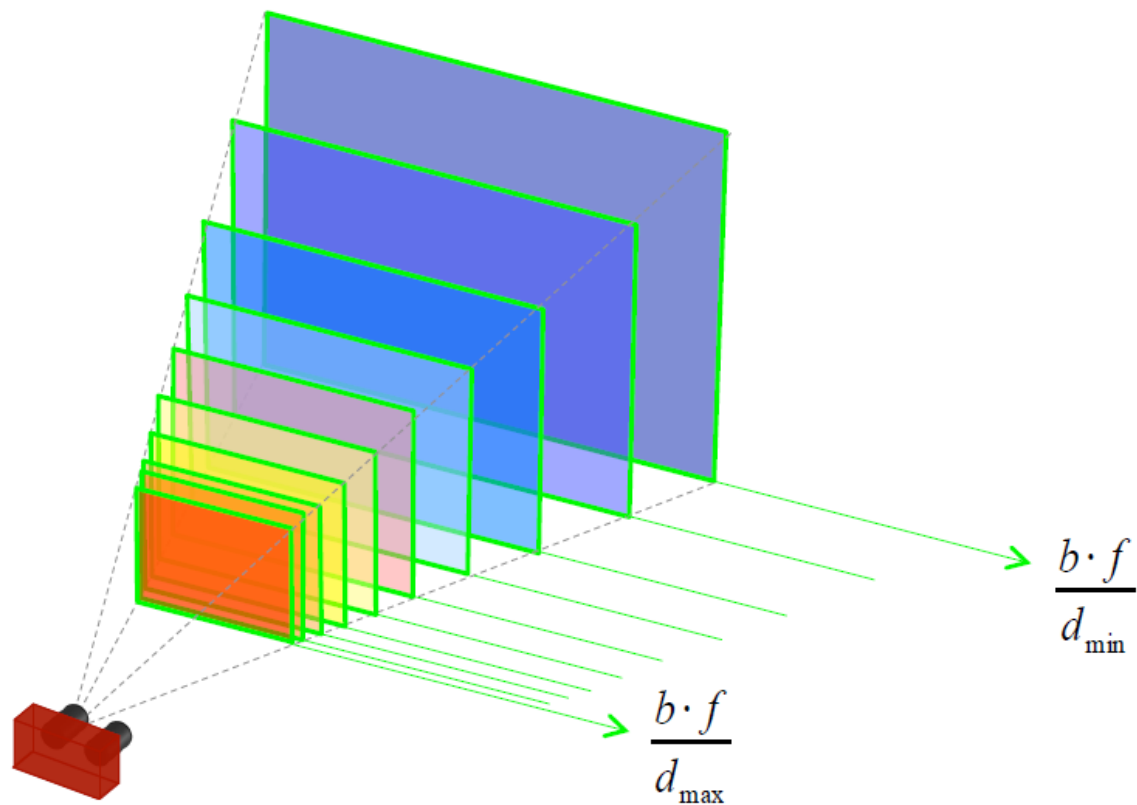
---



双摄装置，基线为 $b$ ，焦距为 $f$ ，那么双摄的视界被视差 $d$ 范围所限定 $\{d_{\min}, d_{\max}\}$ ，如图中绿色包裹的区域



# 视界

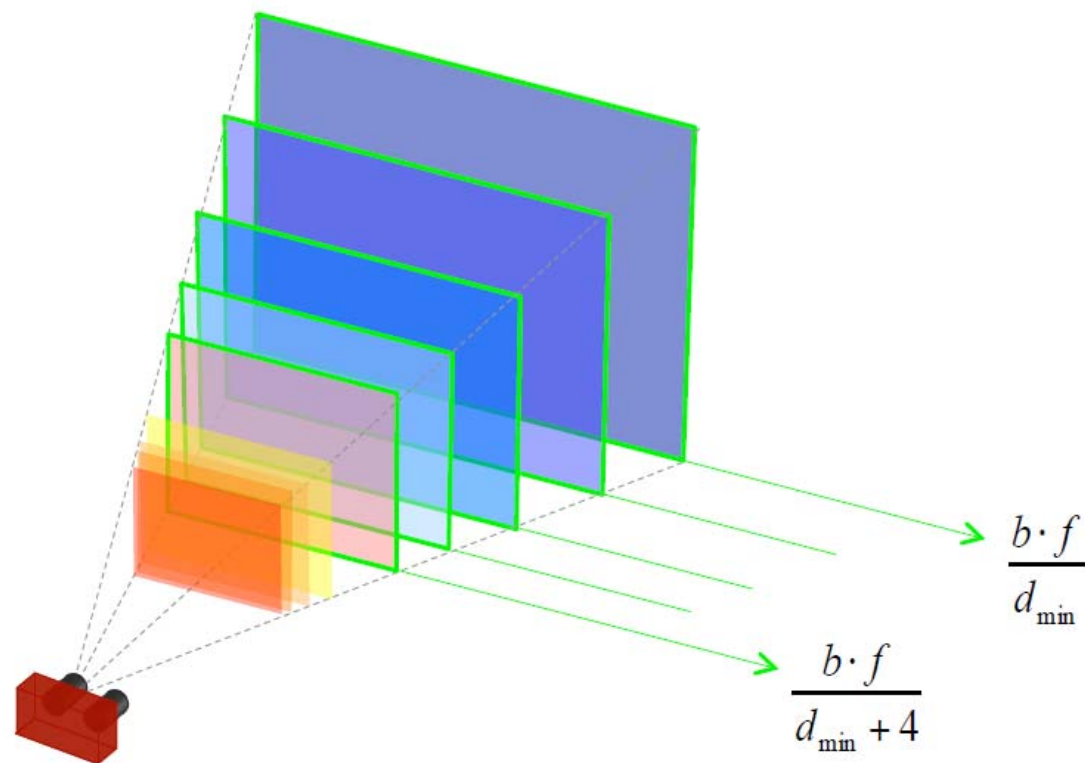


- 深度是通过利用立体匹配系统将视差离散成一系列平行的平面来测量的；每一层平面对应着一个视差。
- 可以通过超像素的方法得到效果更好的深度图。





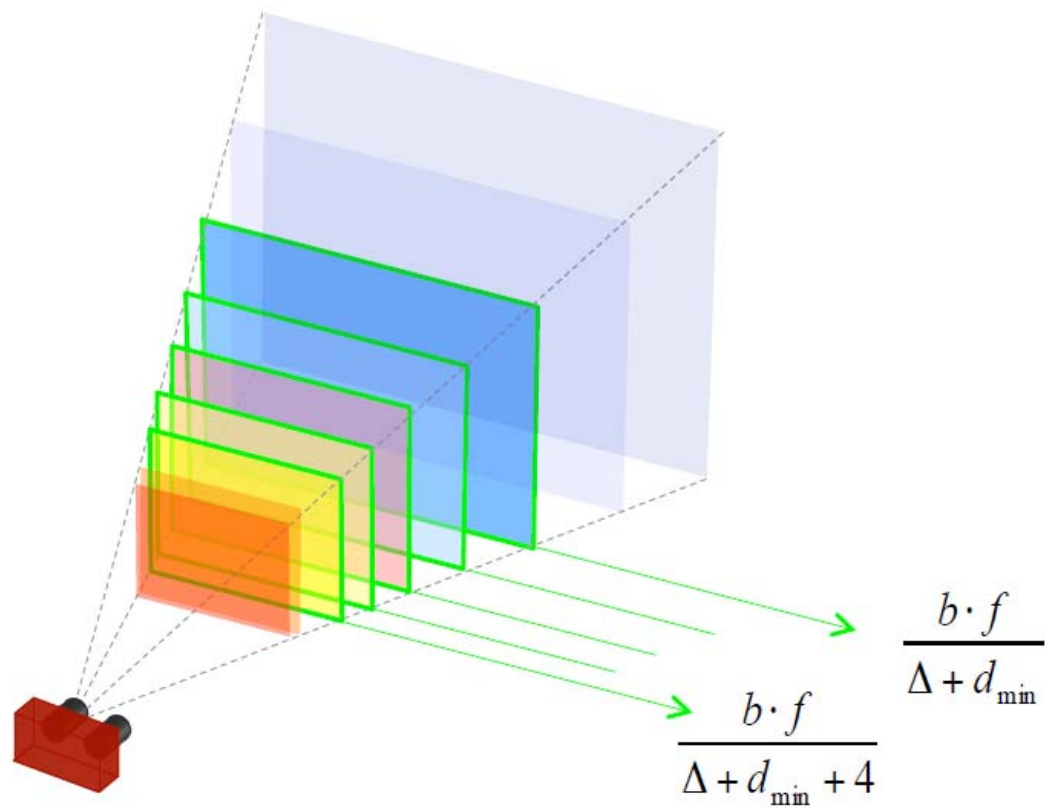
# 视界



图为5个视差 $\{d_{\min}, d_{\min}+4\}$ 组成的视场



# 视界

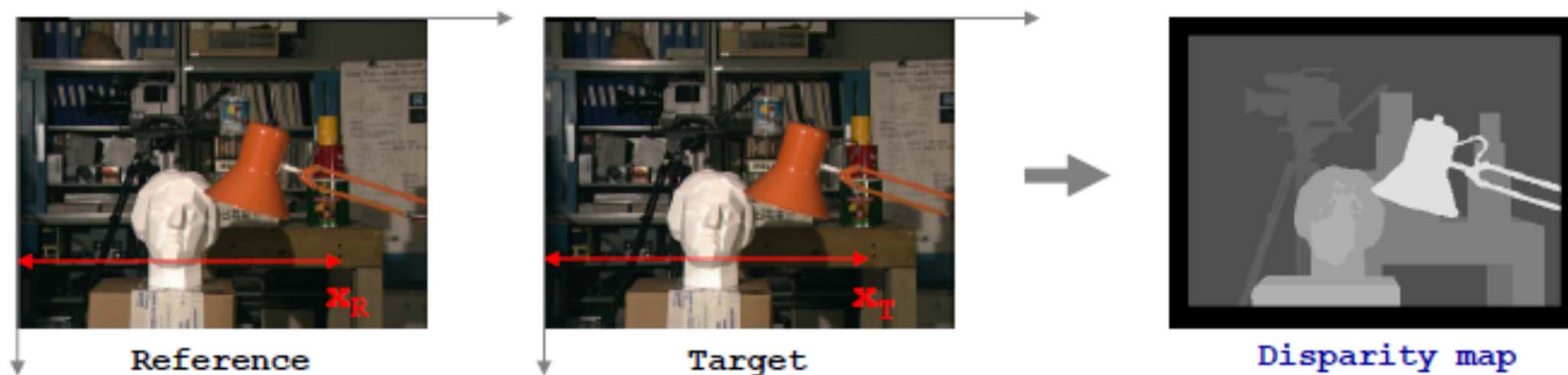


- 图为5个视差 $\{\Delta + d_{\min}, \Delta + d_{\min} + 4\}$ 组成的视场
- $\Delta > 0$ 时，视场收缩并向相机靠近



# 深度与视差

---



所谓视差就是匹配对中对对应点之间 $x$ 方向上的差异,可以将这种差异转换成为灰度图（越近越白）

# 深度估计

---



Traditional  
algorithm

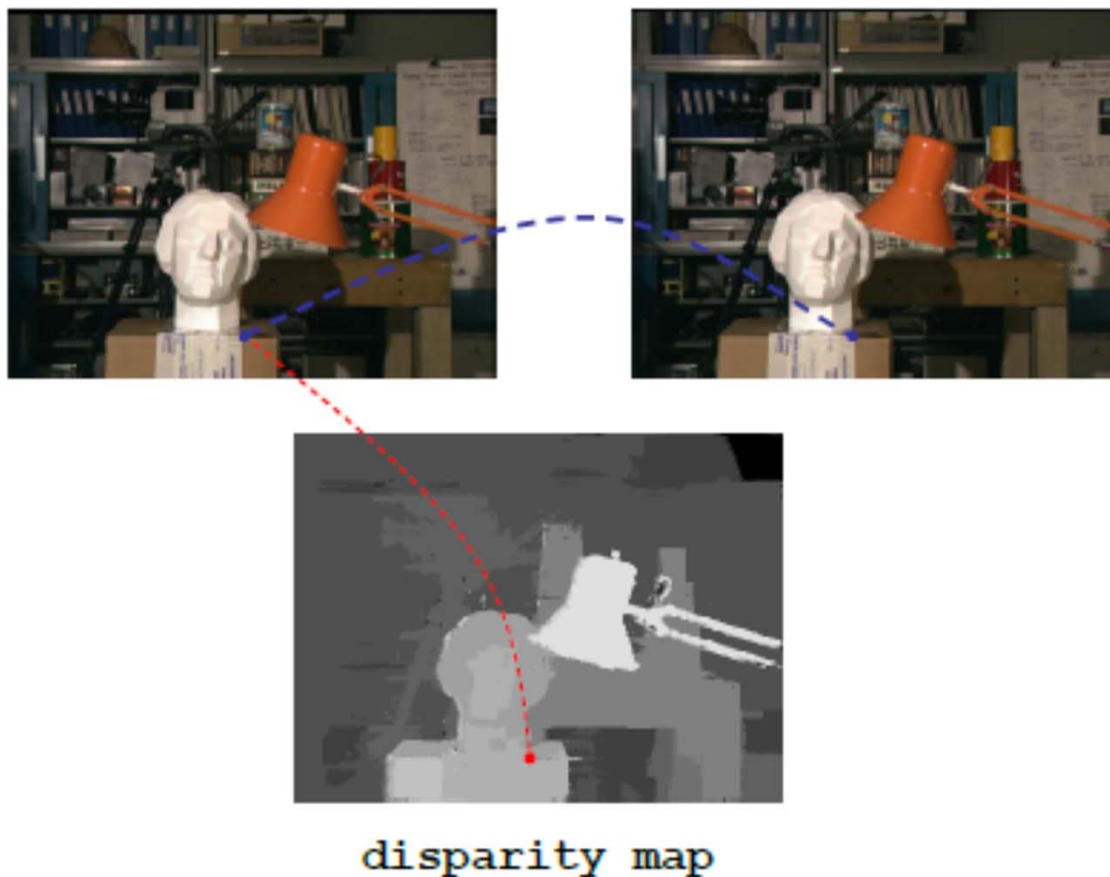


State of the art  
(e.g. ICCV 2011)



# 立体匹配

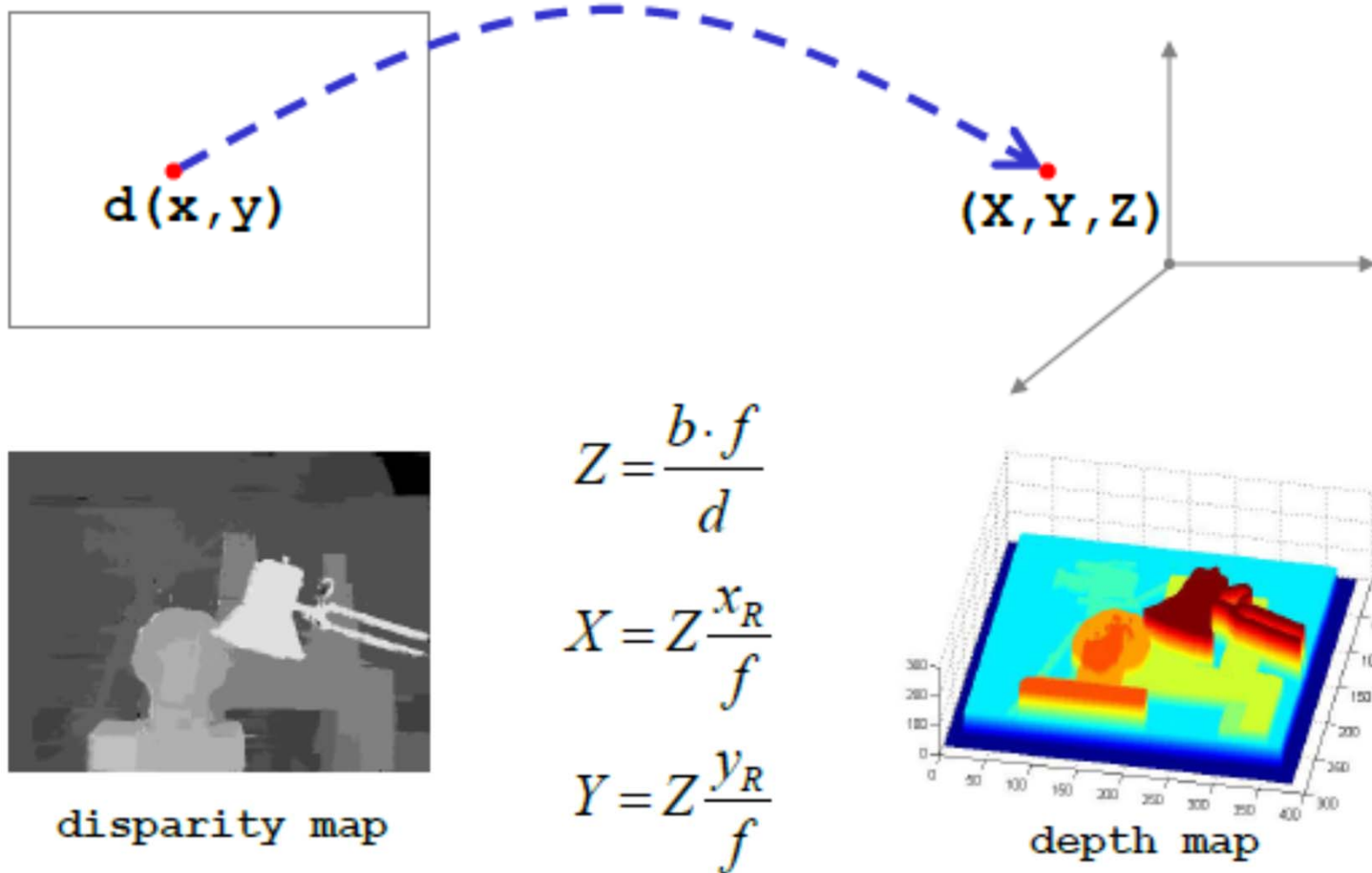
---



利用标定步骤得到的相机的内参对相机镜头畸变进行校正，同时对其扫描线。



# 三角测量



给定视差图像，基线长度以及焦距可以通过三角计算得到当前位置对应的3D位置。

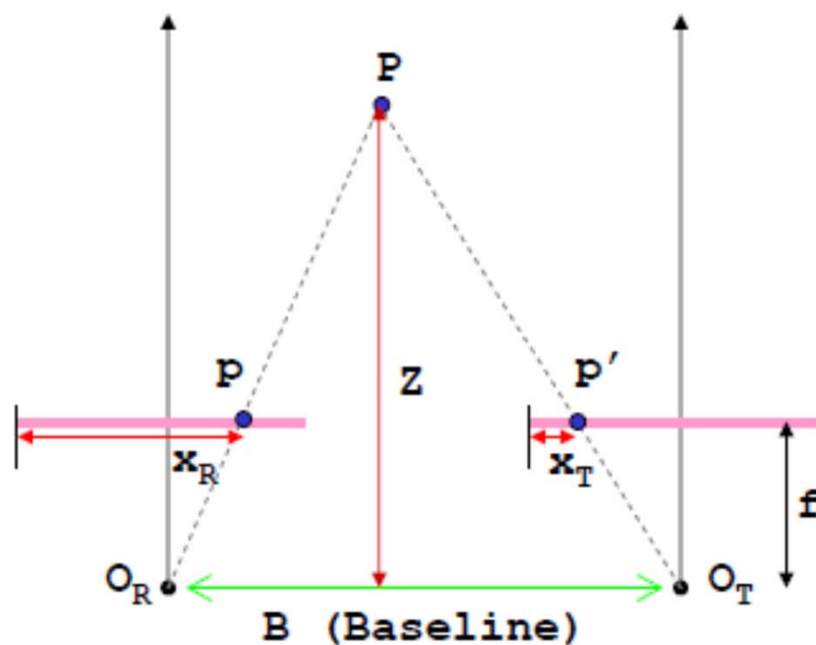


---

# Thanks



# 深度与视差



可以发现：  $PO_R O_T$  与  $PP'' p''$  是相似三角形，由于相似三角形原理

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

其中，  $x_R - x_T$  就是视差，  $Z$  表示深度，  $B$  为基线，  $f$  是焦距。





# 离线标定

---

