

SLAM理论与系统

同济大学计算机科学与技术学院
—朱亚萍



引言

1. 机器人如何观测外部世界？
2. 以相机为主的视觉SLAM中，观测主要是指相机成像过程。
3. 相机成像过程：
 - ◆ 一张照片由很多个像素组成，每个像素记录了色彩或亮度信息。
 - ◆ 物体反射/发出的光线，投影在相机的成像平面上。
 - ◆ 相机的感光器件接收到光线，产生测量值，得到像素，形成照片。
4. 上述过程能否用数学原理来描述？

第四讲 相机模型

1. 针孔相机模型与图像
2. 实践：OpenCV/RGB-D图像拼接

1. 针孔相机模型

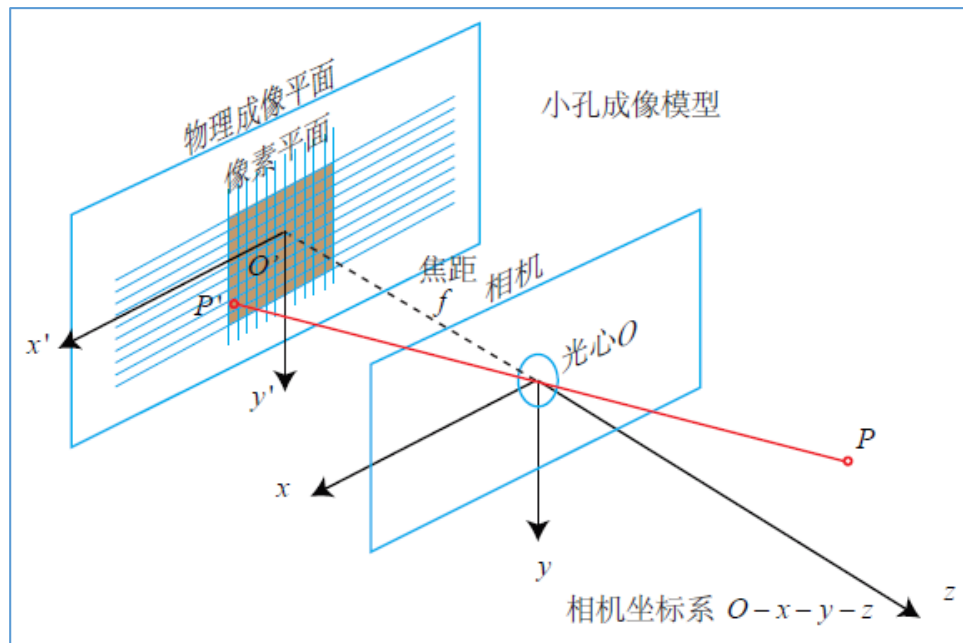
1. 针孔相机模型

- 现实生活中存在大量的照片
 - 照片记录了真实世界在成像平面上的投影
 - 将三维世界的坐标点（单位为米）映射到二维图像平面（单位为像素）
 - 这个过程可以用一个几何模型进行描述
 - 普通相机可以用针孔模型很好地近似



1. 针孔相机模型

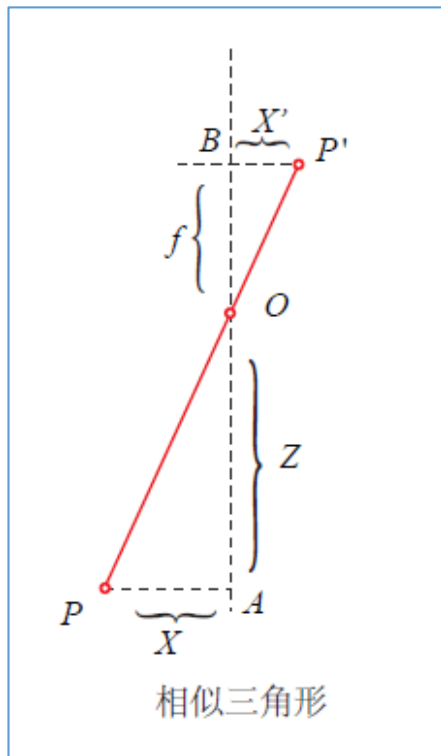
- 蜡烛投影实验：
 - 一支点燃的蜡烛放在暗箱的前方，蜡烛的光透过暗箱上的小孔投影到暗箱后方的平面上形成一个倒立的蜡烛图像
- 几何建模：
- 相机坐标系 $O-x-y-z$ ：
 - ◆ z 轴指向相机前方
 - ◆ x 轴向右
 - ◆ y 轴向下（从左侧看向右侧）
 - ◆ O 为摄像机的光心，即针孔模型的针孔



1. 针孔相机模型

- 空间点P经过小孔O投影，落在物理成像平面 $O' - x' - y'$ 上，成像点为 P' ：
- P点坐标： $[X, Y, Z]^T$
- P' 点坐标： $[X', Y', Z']^T$
- 物理成像平面到小孔的距离（焦距）： f

三角形相似： $\frac{Z}{f} = -\frac{X}{X'} = -\frac{Y}{Y'}$.



1. 针孔相机模型

- 空间点P经过小孔O投影，落在物理成像平面 $O' - x' - y'$ 上，成像点为 P' ：

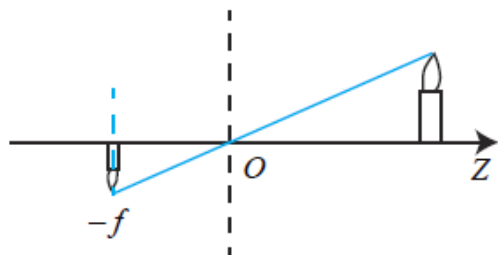
三角形相似：
$$\frac{Z}{f} = -\frac{X}{X'} = -\frac{Y}{Y'}$$

翻转在前面：
$$\frac{Z}{f} = \frac{X}{X'} = \frac{Y}{Y'}$$

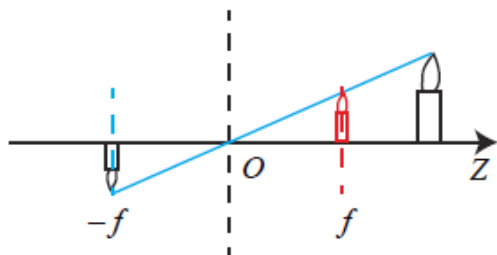
整理之：

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

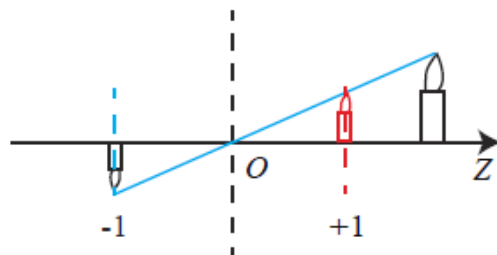
$$Y' = f \frac{Y}{Z}$$



真实成像平面



对称的成像平面



归一化成像平面

1. 针孔相机模型

$$\begin{aligned} X' &= f \frac{X}{Z} \\ Y' &= f \frac{Y}{Z} \end{aligned}$$

- 相机：像素，需要在成像平面上对像进行**采样**和**量化**
 - 设在物理成像平面上固定着一个像素平面 $o - u - v$ ，像素平面上 P' 的**像素坐标**： $[u, v]^T$ 。
 - **像素坐标系**：原点 o' 位于图像的左上角， u 轴向右与 x 轴平行， v 轴向下与 y 轴平行
 - 像素坐标系与成像平面之间相差了一个**缩放**和一个**原点的平移**

1. 针孔相机模型

- 假设像素坐标在u轴上缩放了 α 倍，在v轴上缩放了 β 倍；同时，原点平移了 $[c_x, c_y]^T$ ：
- P' 的坐标与像素坐标 $[u, v]^T$ 的关系（成像平面到像素坐标）：

$$\begin{cases} u = \alpha X' + c_x \\ v = \beta Y' + c_y \end{cases}$$

将 $\begin{cases} X' = f \frac{X}{Z} \\ Y' = f \frac{Y}{Z} \end{cases}$ 代入，把 αf 合并成 f_x ，把 βf 合并成 f_y ：

得：

$$\begin{cases} u = f_x \frac{X}{Z} + c_x \\ v = f_y \frac{Y}{Z} + c_y \end{cases}$$

f 的单位为米， α 、 β 的单位为像素/米， f_x 、 f_y 和 c_x 、 c_y 的单位为像素。

1. 针孔相机模型

转化成矩阵形式（左侧使用齐次坐标，右侧不需要）：

$$\begin{cases} u = f_x \frac{X}{Z} + c_x \\ v = f_y \frac{Y}{Z} + c_y \end{cases}$$

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{Z} \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \triangleq \frac{1}{Z} K P.$$

进一步改写：

$$Z \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \triangleq K P.$$

中间的量组成的矩阵 K 称为相机的内参数

内参通常在相机生产之后就已固定。

1. 针孔相机模型

$$Z \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \triangleq KP$$

- P_{uv} : 相机坐标系下的坐标
- 相机不停运动: P 的相机坐标与它的世界坐标 (P_w) 相差一个变换:

$$Z P_{uv} = Z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K (R P_w + t) = K T P_w.$$

根据相机的当前位姿变换到
相机坐标系下

隐含了一次齐次坐标到非齐次坐标的变换

- 相机的位姿 R, t 或 T 称为相机的外参
 - 外参会随着相机运动发生改变
 - 是SLAM中待估计的目标, 代表着机器人的轨迹

1. 针孔相机模型

$$Z P_{uv} = Z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K (R P_w + t) = K T P_w.$$

- 另一个角度看投影过程：

- 把一个世界坐标点转换到相机坐标系，再除掉它最后一维的数值

该点距离相机成像平面的深度

- 相当于把最后一维进行归一化处理，得到P点在相机归一化平面上的投影：

$$(R P_w + t) = \underbrace{[X, Y, Z]^T}_{\text{相机坐标}} \rightarrow \underbrace{[X / Z, Y / Z, 1]^T}_{\text{归一化坐标}}$$

- 归一化坐标：相机前方 $z=1$ 处的平面上的一个点
- $z = 1$ 平面：归一化平面
- 归一化坐标左乘内参就得到像素坐标，因此可以将像素坐标看成对归一化平面上的点进行量化测量的结果
- 如果对相机坐标同时乘以非零常数，归一化坐标都一样，说明点的深度信息在投影过程中丢失了

单目视觉无法得到像素点的深度值

1. 针孔相机模型

- 畸变
 - 针孔前的镜头会引入畸变



广角镜头畸变



鱼眼镜头畸变

1. 针孔相机模型

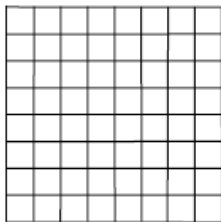
- 相机前方加入的透镜对成像过程中光线的传播产生影响：
 - 透镜自身的形状对光线传播的影响 ➡ 径向畸变
 - 在机械组装过程中，透镜和成像平面不完全平行，使得光线穿过透镜投影到成像平面时位置发生变化 ➡ 切向畸变

1. 针孔相机模型

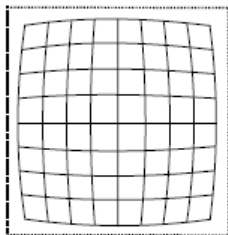
- 径向畸变：
 - 在实际拍摄的照片中，真实的直线通过透镜在图片中变成了曲线
 - 越靠近图像的**边缘**，这种现象**越明显**
 - 实际加工制作的透镜一般是中心对称的，使得畸变通常**径向对称**：

- ◆ 桶形畸变：图像放大率随着与光轴之间的距离**增加而减小**
- ◆ 枕形畸变：图像放大率随着与光轴之间的距离**增加而增大**

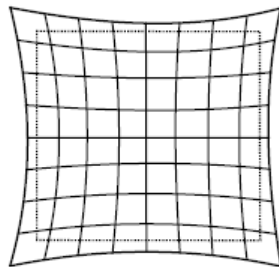
穿过图像中心与光轴有交点的直线保持形状不变



正常图像



桶形失真



枕形失真

1. 针孔相机模型

- 切向畸变：
 - 相机的组装过程中，透镜和成像面无法严格平行

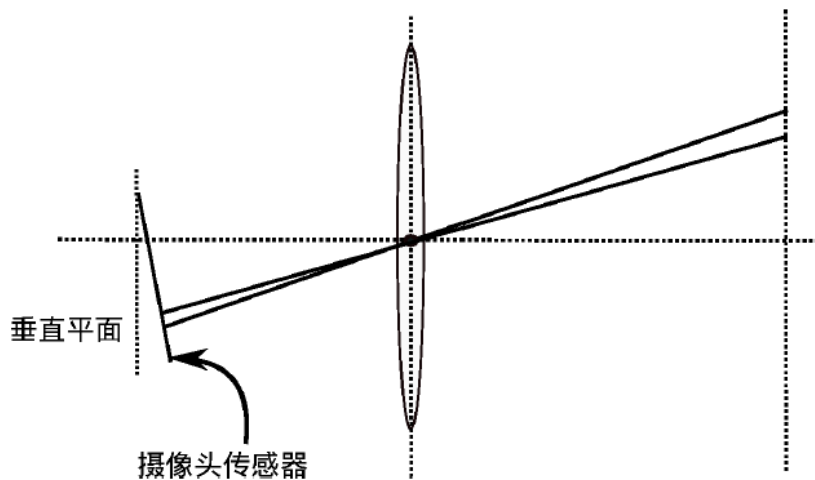


图 5-4 切向畸变来源示意图。

1. 针孔相机模型

- 平面上任意一点：
 - 笛卡尔坐标: $[x, y]^T$
 - 极坐标: $[r, \theta]^T$ (r 表示离原点的距离; θ 表示与水平轴的夹角。)
- 径向畸变：
 - 坐标点沿着长度方向发生了变化 δr , 即距离原点的长度发生了变化
 - 无论桶形或枕形畸变, 都是随着离中心距离的增加而增加
 - 畸变可以用与距中心距离有关的二次及高次多项式函数进行纠正

1. 针孔相机模型

- 用与距中心距离有关的二次及高次多项式函数**纠正畸变**:

$$\begin{aligned}x_{distorted} &= x \left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6 \right) \\y_{distorted} &= y \left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6 \right)\end{aligned}$$

- ◆ 对于畸变较小的图像**中心区域**，畸变纠正主要是 k_1 起作用；
- ◆ 对于畸变较大的**边缘区域**主要是 k_2 起作用；
- ◆ 对**畸变很大**的摄像头（如鱼眼），可以加入 k_3 畸变项对畸变进行纠正

普通摄像头用这两个系数
就能很好地纠正径向畸变

1. 针孔相机模型

- 切向畸变：

- 坐标点沿着切线方向发生了变化，即水平夹角发生了变化 $\delta\theta$

- 纠正函数：

$$x_{distorted} = x + 2p_1xy + p_2(r^2 + 2x^2)$$

$$y_{distorted} = y + p_1(r^2 + 2y^2) + 2p_2xy$$

1. 针孔相机模型

- 联合上述两种畸变的纠正函数，对于相机坐标系中的任一点P，可以通过5个畸变系数得到该点在像素平面上的正确位置：

- 将三维空间点投影到归一化图像平面，设归一化坐标为 $[x, y]^T$ ；

- 对归一化平面上的点进行径向畸变和切向畸变纠正：

$$\begin{cases} x_{distorted} = x(1 + k_1r^2 + k_2r^4 + k_3r^6) + 2p_1xy + p_2(r^2 + 2x^2) \\ y_{distorted} = y(1 + k_1r^2 + k_2r^4 + k_3r^6) + p_1(r^2 + 2y^2) + 2p_2xy \end{cases}$$

- 将纠正后的点通过内参数矩阵投影到像素平面，得到该点在图像上的正确位置：

$$\begin{cases} u = f_x x_{corrected} + c_x \\ v = f_y y_{corrected} + c_y \end{cases}$$

- 上述5个畸变纠正项，在实际应用中可以灵活选择不同的组合。

1. 针孔相机模型

- 两种去畸变处理方法：
 - 先对整张图像进行去畸变，得到去畸变后的图像，然后讨论此图像上的点的空间位置（在SLAM中更加常见）
 - 先考虑图像中的某个点，然后按照去畸变方程，讨论它去畸变后的空间位置
- 当图像去畸变之后，可以直接用针孔模型建立投影关系，无需再考虑畸变。

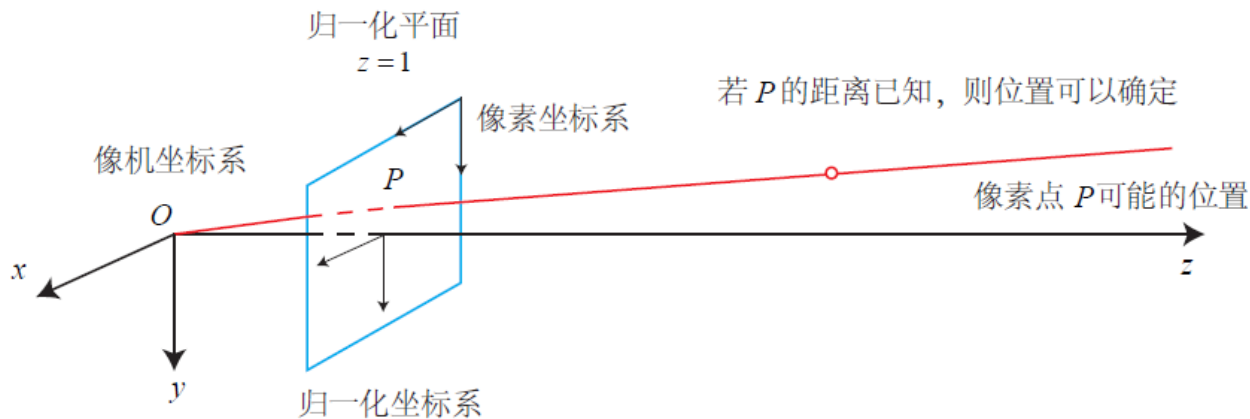
1. 针孔相机模型

- 单目相机的成像过程：

1. 首先，世界坐标系下有一个固定的点 P ，世界坐标为 P_w ；
2. 由于相机在运动，它的运动由 R, t 或变换矩阵 $T \in SE(3)$ 描述。 P 的相机坐标为：
 $\tilde{P}_c = RP_w + t$ 。
3. 这时的 \tilde{P}_c 仍有 X, Y, Z 三个量，把它们投影到归一化平面 $Z = 1$ 上，得到 P 的归一化相机坐标： $P_c = [X/Z, Y/Z, 1]^T$ ^①。
4. 最后， P 的归一化坐标经过内参后，对应到它的像素坐标： $P_{uv} = KP_c$ 。

1. 针孔相机模型

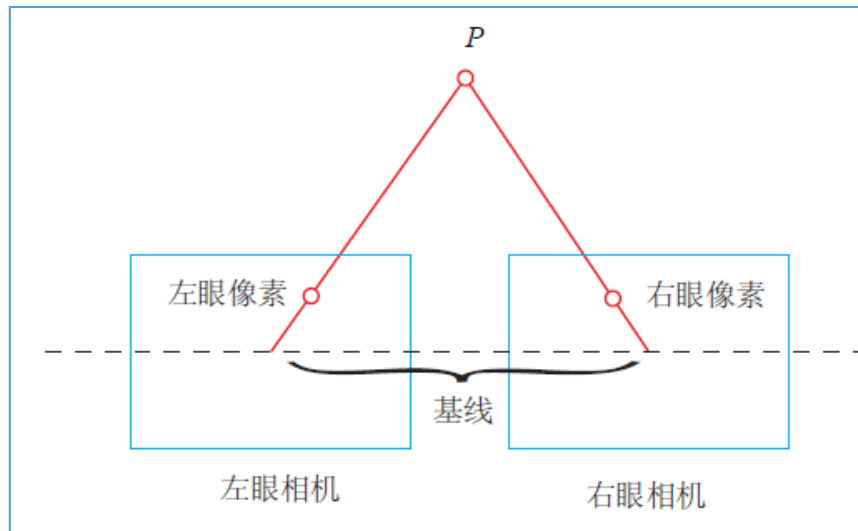
- 针孔相机模型 -- 单个相机的成像模型
- 仅根据一个像素，无法确定空间点的具体位置（从相机光心到归一化平面连线上的所有点，都可以投影到该像素上）
- 需要确定P的深度 → 双目/RGB-D相机



1. 针孔相机模型

- 双目模型

- 人眼：左右眼的景物**差异**判断物体距离
- 双目相机：同步采集左右相机的图像，计算图像间**视差**，估计像素深度
- 左右双目相机水平放置，两个相机的光圈中心都位于x轴上
 - 左右相机中心距离称为**基线**（记作 b ）



1. 针孔相机模型

- 双目模型

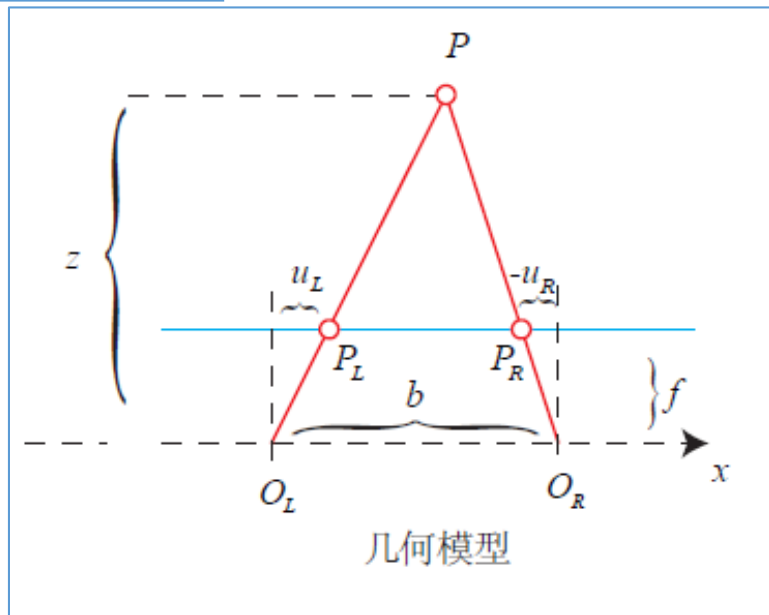
- 左右像素的几何关系：

$$\frac{z - f}{z} = \frac{b - u_L + u_R}{b}.$$

- 整理得

$$z = \frac{fb}{d}, \quad d = u_L - u_R.$$

- 视差d与距离成反比：视差越大，距离越近
 - 基线b越长，双目能测到的最大距离越远；反之，小型双目器件只能测量很近的距离。



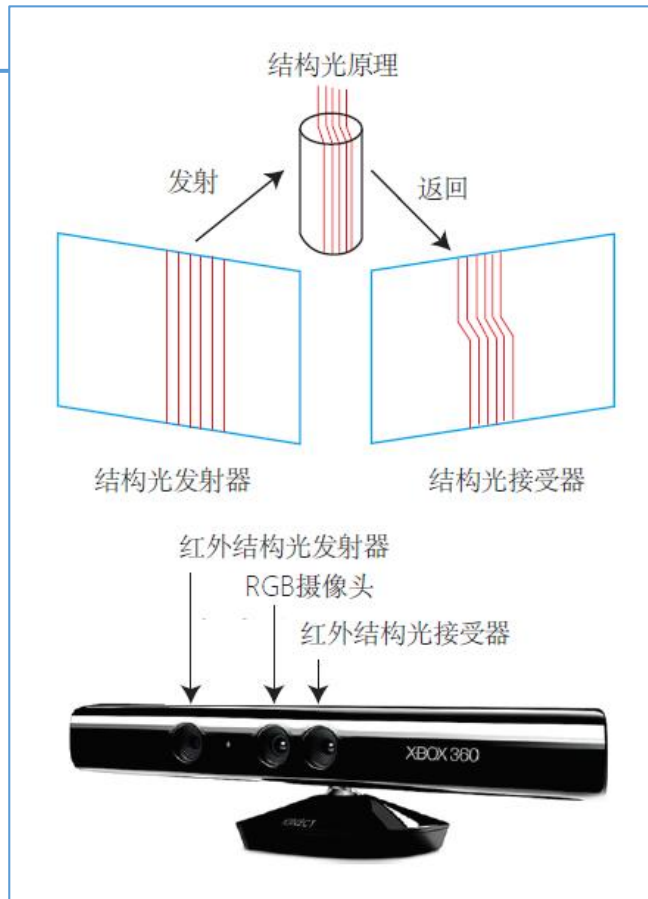
d称为**视差**（disparity），可以用来估计一个像素与相机之间的距离；视差最小为一个像素，因此双目能测量的深度有理论上的最大值：fb；虽然距离公式简单，但d不容易计算

1. 针孔相机模型

- RGB-D相机：物理手段测量深度
 - 能够主动测量每个像素的深度：向探测目标发射一束光线
 - ◆ 红外结构光
 - ◆ 飞行时间（ToF）

1. 针孔相机模型

- RGB-D相机：
 - 红外结构光：根据返回的**结构光图案**，计算物体与自身的距离，例如，Kinect 1代、Project Tango 1代、Intel RealSense等。



1. 针孔相机模型

- RGB-D相机：
 - 飞行时间（ToF）：相机向目标发射脉冲光，然后根据发送到返回之间的光束**飞行时间**，确定物体与自身的距离，例如，Kinect 2代、ToF传感器等。



1. 针孔相机模型

- RGB-D相机：

- 测量深度之后，按照生产时的相机摆放位置，自己完成深度与彩色图像素之间的配对，输出一一对应的彩色图和深度图
- 可以在同一个图像位置，读取到色彩信息和距离信息，计算像素的3D相机坐标，生成点云（Point Cloud）
- 既可以在图像层面对RGB-D数据进行处理，也可在点云层面处理

1. 针孔相机模型

- RGB-D相机:

- 优势:

- ◆ 能够实时测量每个像素点的距离

- 劣势:

- ◆ 由于使用发射-接收的测量方式, 使用范围受限
 - ◆ 用红外光进行深度测量时, 容易受日光或其他传感器发射的红外光干扰, 无法在室外使用
 - ◆ 如果不加调制的话, 同时使用多个RGB-D相机可能产生互相干扰
 - ◆ 无法测量透射材质的物体的位置
 - ◆ 成本、功耗上处于劣势

1. 针孔相机模型

- 相机成像后，生成了图像
- 图像在计算机中以矩阵形式存储（二维数组）

- 灰度图

$$I(x, y) : \mathbb{R}^2 \mapsto \mathbb{R}$$

- 需要对感光度量化成数值，例如0~255之间的整数
 - 二维数组表达

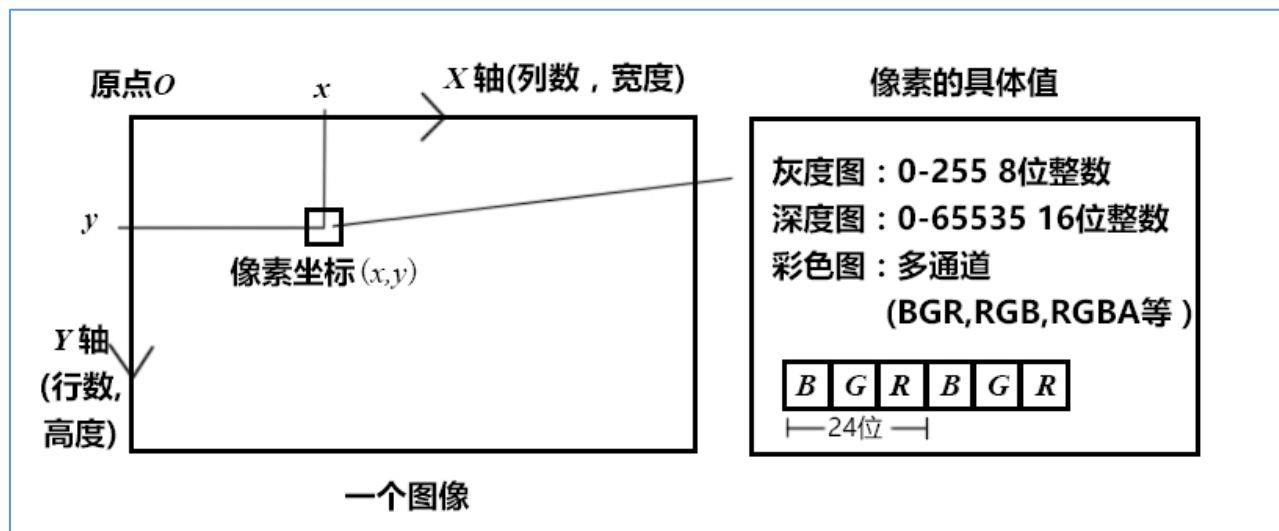
```
unsigned char image[480][640];
```


1. 针孔相机模型

- 访问图像像素

```
unsigned char pixel = image[y][x];
```

- 图像坐标示意



2. 实践：OpenCV/图像拼接

计算机中的图像

- 安装OpenCV，实现图像读取、显示、像素遍历、复制、赋值
- 利用OpenCV的去畸变函数对图像进行处理（见文件夹畸变图像）
- 构建点云