法律声明

- 本课件包括：演示文稿，示例，代码，题库，视频和声音等，小象学院拥有完全知识产权的权利；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或机构不得盗版、复制、仿造其中的创意，我们将保留一切通过法律手段追究违反者的权利。

- 课程详情请咨询
  - 微信公众号：小象学院
  - 新浪微博：小象AI学院

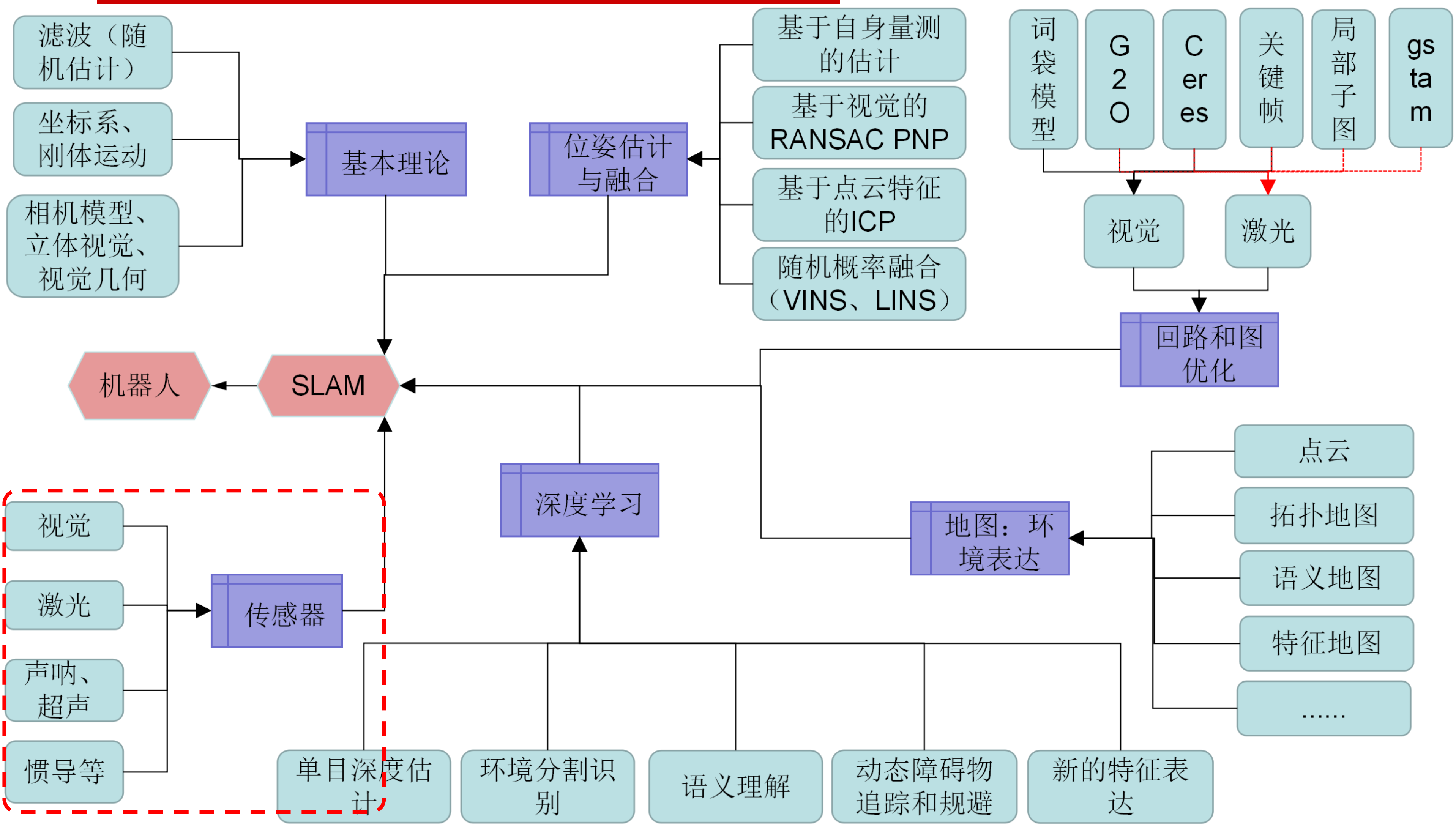
第四讲：

SLAM的传感器

主讲：杨亮、殷鹏

GitHub链接：https://github.com/EricLYang/courseRepo

总结


滤波（随机估计）
坐标系、刚体运动
相机模型、立体视觉、视觉几何
基本理论
位姿估计与融合
基于自身量测的估计
基于视觉的RANSAC PNP
基于点云特征的ICP
随机概率融合（VINS、LINS）
词袋模型
G2O
Ceres
关键帧
局部子图
gstam
视觉
激光
回路和图优化
机器人
SLAM
深度学习
地图：环境表达
点云
拓扑地图
语义地图
特征地图
......
视觉
激光
声呐、超声
惯导等
传感器
单目深度估计
环境分割识别
语义理解
动态障碍物追踪和规避
新的特征表达

提纲

- SLAM传感器综述
- 视觉类传感器（单目、双目和RGBD相机）
- 主动类传感器-- 激光
- 视觉和激光实例

SLAM传感器综述

Simultaneous **L**ocalization **A**nd **M**apping:　同步定位与构图

输入

输出

可见光视觉

IMU

激光雷达

声呐

SLAM

我在当前环境的什么位置？

当前所在的环境是个什么样子？

SLAM传感器综述

- SLAM 传感器
  - 可见光类
    - 相机：单目、双目...
  - 不可见光类
    - 激光：2D/3D
    - 红外
    - TOF
  - 混合
    - RGBD
  - 声传感器
    - 超声波：2D/3D
    - 雷达

[1]Chong, T. J., Tang, X. J., Leng, C. H., Yogeswaran, M., Ng, O. E., & Chong, Y. Z. (2015). Sensor technologies and simultaneous localization and mapping (SLAM). Procedia Computer Science, 76, 174-179.

SLAM传感器综述

LASER SCANNERS

2D LIDAR
(SICK, HOKUYO, PEPPERL FUCHS, OMRON...)

2D MULTI-LAYER LASER SCANNER
(IBEO)

3D LIDAR
(VELODYNE, NEPTEC OPAL)

3D CAMERAS

DIRECT TIME OF FLIGHT IMAGERS
(ASC)

PHASE SHIFT TIME OF FLIGHT IMAGERS
(MESA IMAGING, IFM, FOTONIC, SOFTKINETIC)

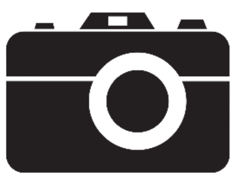
STEREO CAMERAS
(STEREOLABS, VISLAB, MULTISENSE)

SONAR/RADAR

2D SONARS
(BLUEVIEW)

3D SONARS
(BLUEVIEW)

Mono Camera



Infrared Camera

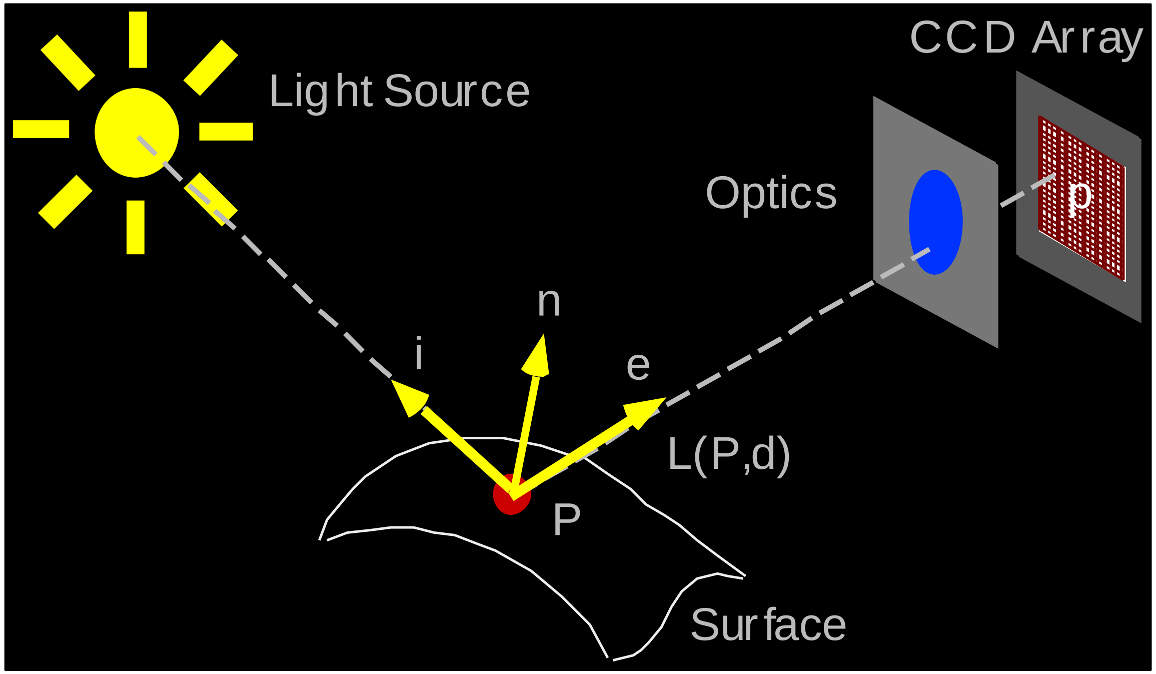
提纲

- SLAM传感器综述
- 视觉类传感器（单目、双目和RGBD相机）
- 主动类传感器-- 激光
- 视觉和激光实例

视觉类传感器

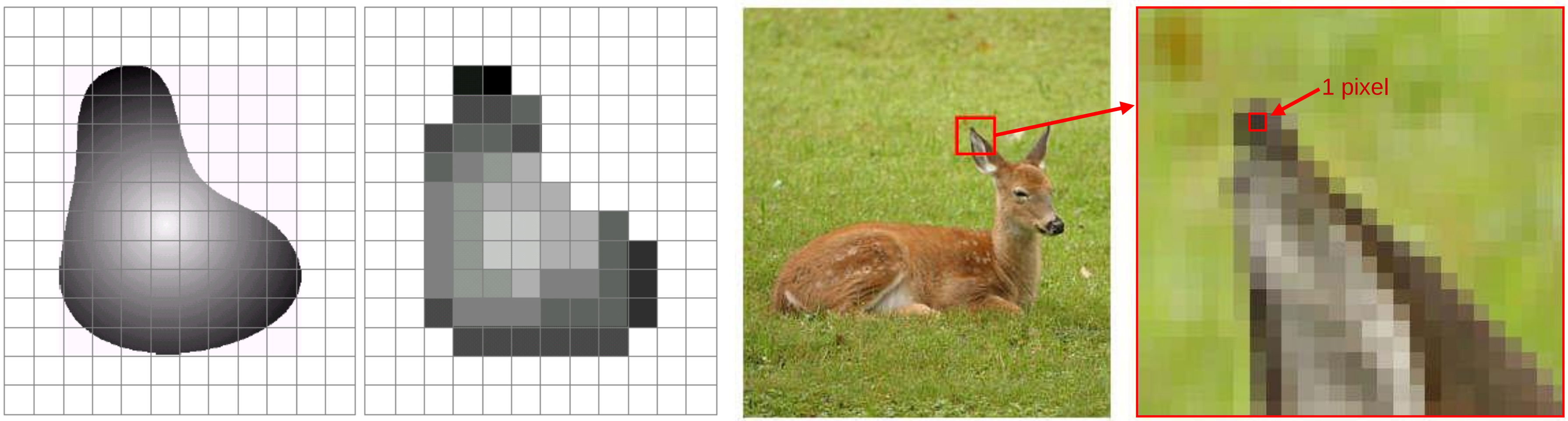
相机的作用：一种将可见光或者电磁谱捕捉记录为一张图像（通常方便人或者设备读取或者处理）

对于数字图像：表示为 通过具有有限数值表示的二维矩阵（单一通道），也叫做像素



视觉类传感器

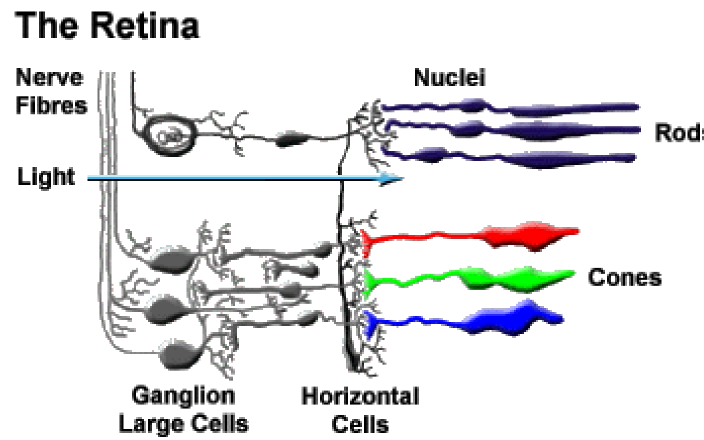
- 像素通常表示为一个灰度值、颜色等信息。并且数字化图像的灰度值是对真实场景的近似



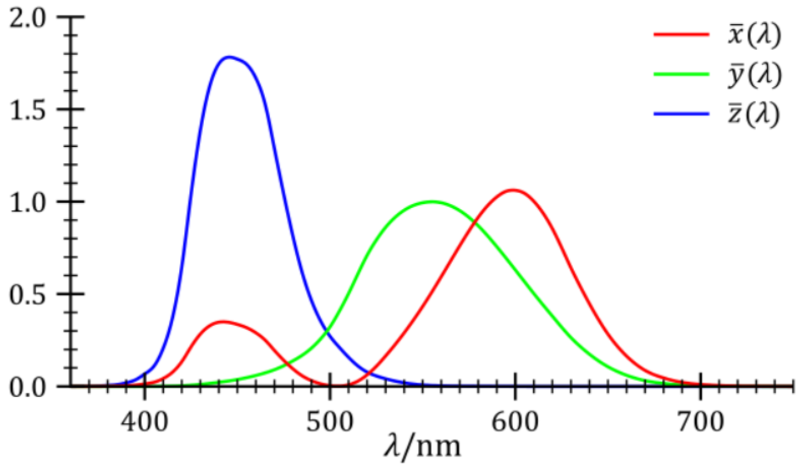
视觉类传感器

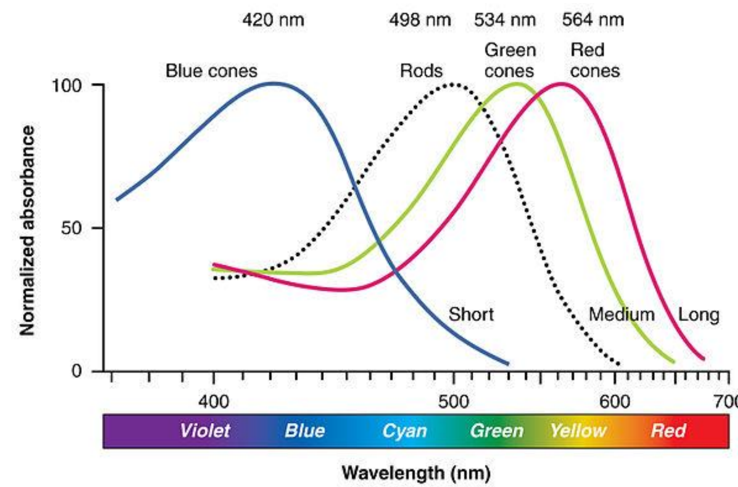
- Cones : sensitive to red, green, and blue light are 63, 31, and 6 percent, respectively

- Rods : are exquisitely sensitive and operate only in low-light conditions



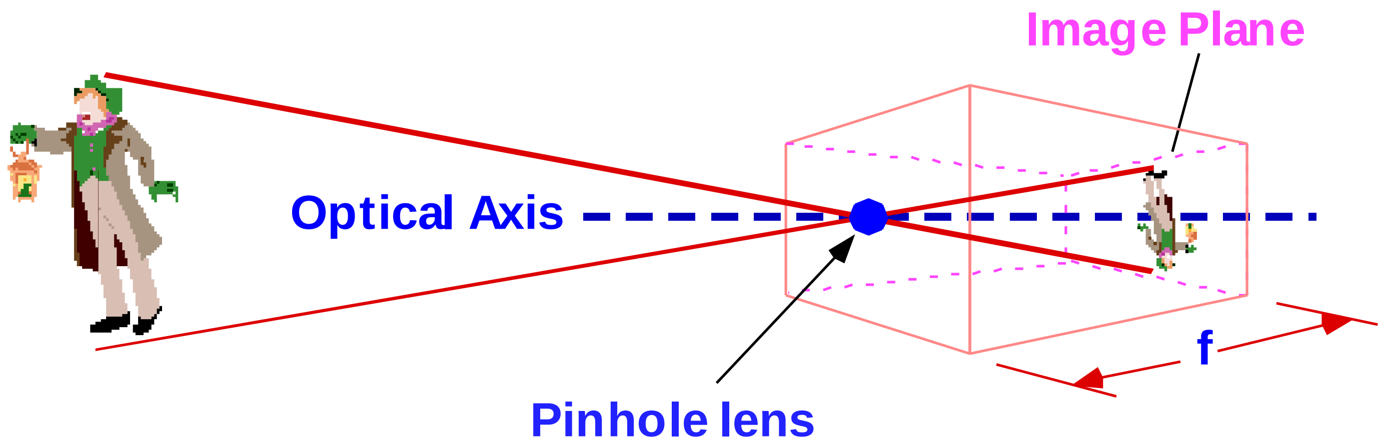
CIE 1931 color space





$$Y_{\text{linear}} = 0.2126R_{\text{linear}} + 0.7152G_{\text{linear}} + 0.0722B_{\text{linear}}$$

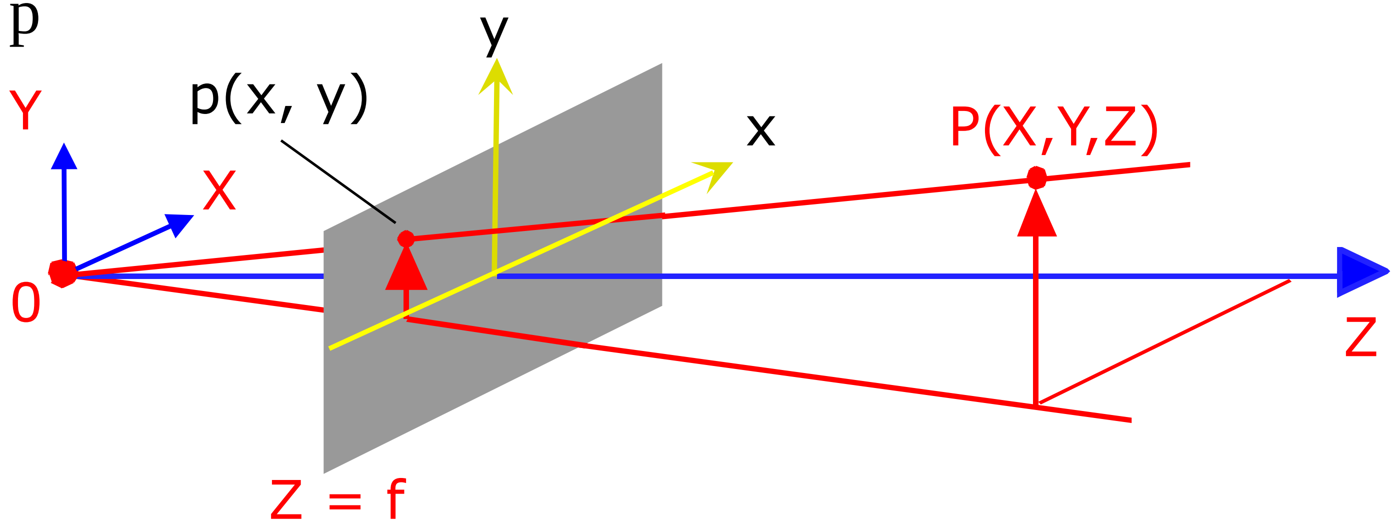
视觉类传感器



- 相机成像（小孔成像模型）
  - Image inverted
  - Size reduced
  - Image is dim
  - No direct depth information
- f 叫做镜头的焦距
- 这就是： perspective projection

视觉类传感器

- 假设给定了世界坐标系下的点P，在相机平面对应的投影p



相机的中心即投影中心

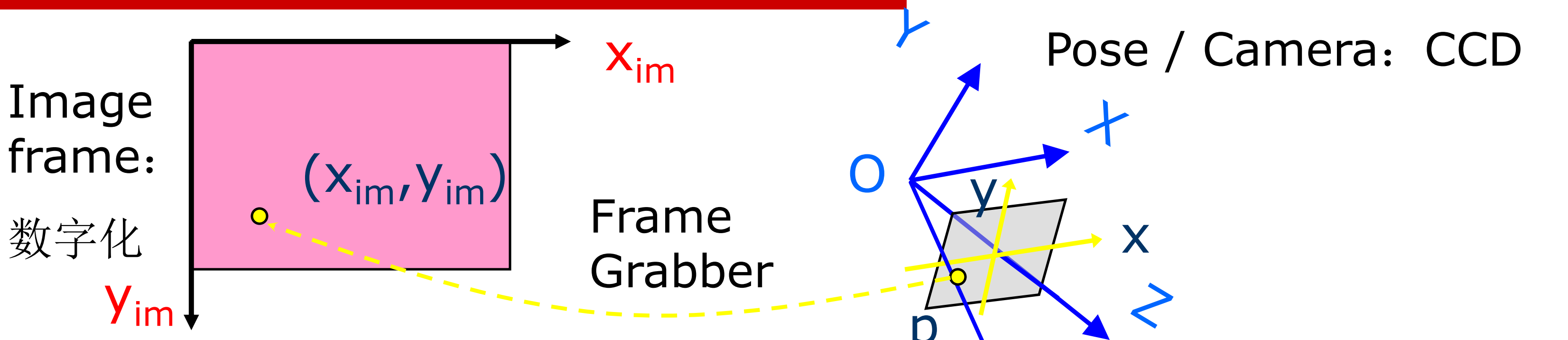
Z轴沿着光学轴

图像平面的深度 Z = f；对应轴和世界坐标轴平行x // X 和 y//Y

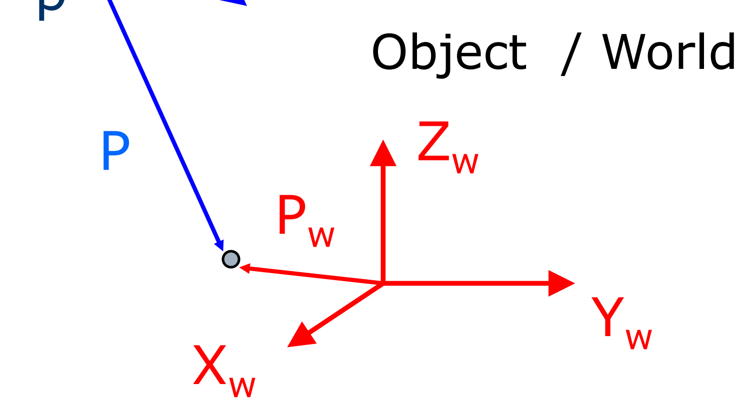
$$x = f\,\frac{X}{Z}$$

$$y = f\,\frac{Y}{Z}$$

视觉类传感器

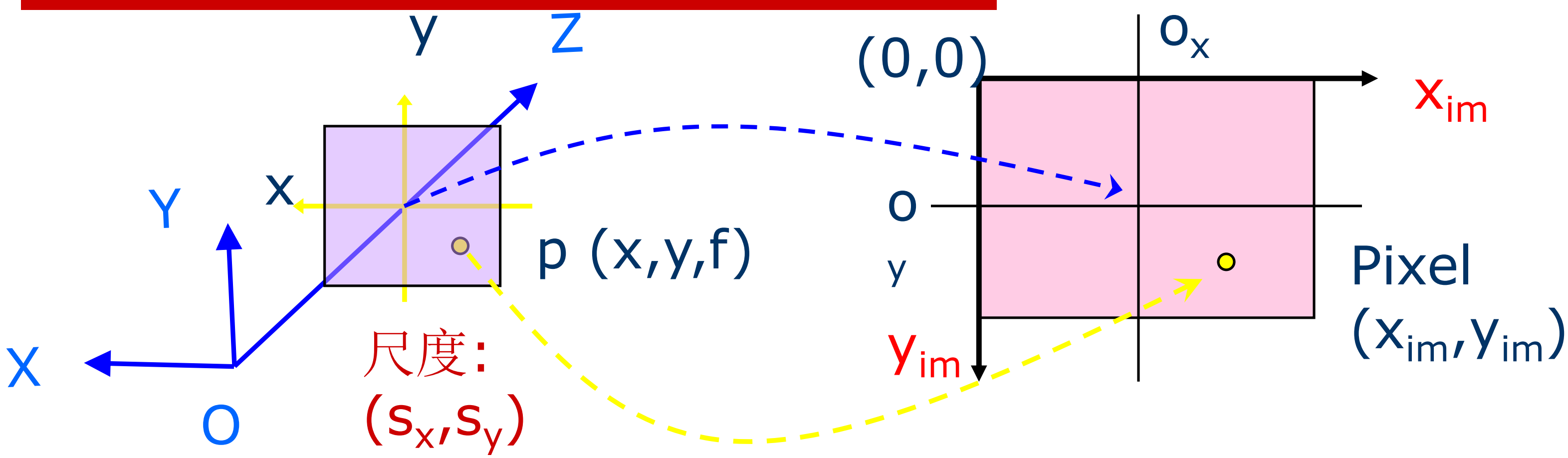


- 坐标系
  - Frame coordinates (x_{im}, y_{im}) pixels
  - Image coordinates (x,y) in mm
  - Camera coordinates (X,Y,Z)
  - World coordinates (X_w,Y_w,Z_w)



- 相机的参数
  - 内参 Intrinsic Parameters (of the camera and the frame grabber): link the **frame coordinates** of an image point with its corresponding **camera coordinates**
  - 外参 Extrinsic parameters: define the location and orientation of the **camera coordinate system** with respect to the **world coordinate system**

视觉类传感器



- 从成像坐标系到帧坐标系
  - Image center
  - Directions of axes
  - Pixel size

$$x = -(x_{im} - o_x)s_x$$

$$y = -(y_{im} - o_y)s_y$$

- 从3D到2D
  - Perspective projection

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

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

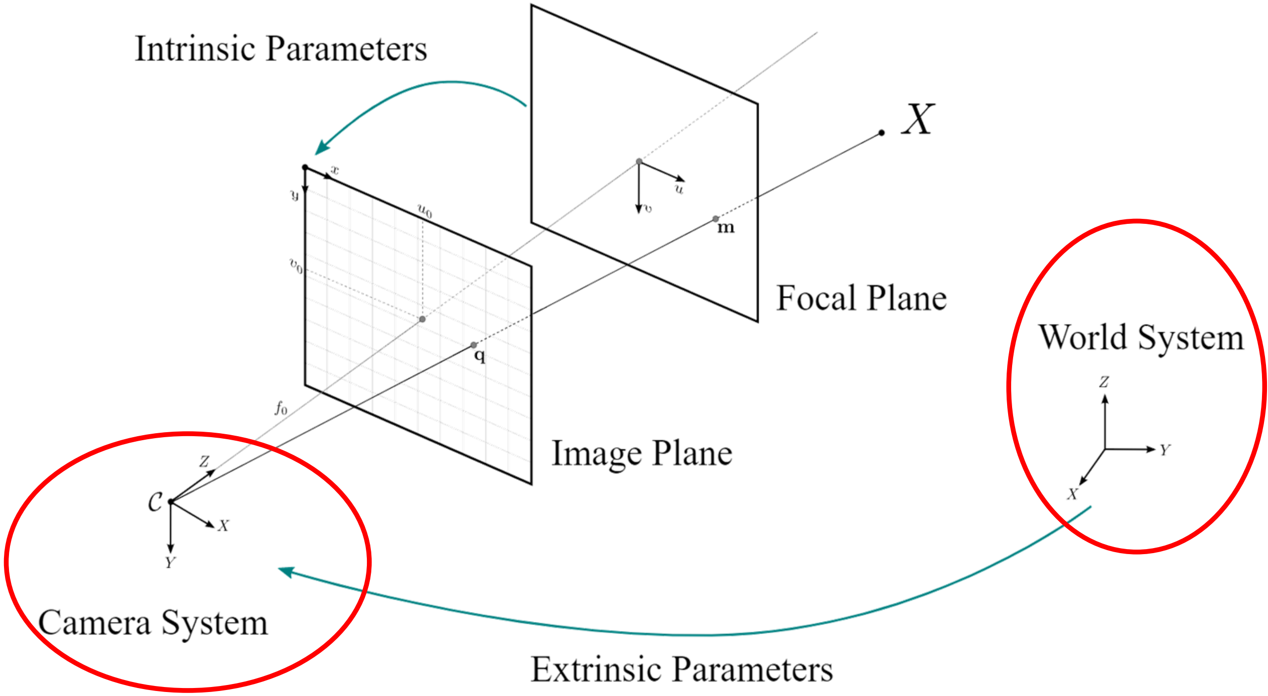
- 内参
  - (o_x, o_y) : image center (in pixels)
  - (s_x, s_y) : effective size of the pixel (in mm)
  - f: focal length

视觉类传感器

- 从世界坐标系到相机坐标系

相机看到的，只是在他自己的世界里，不是那个不名的世界

$$\mathbf{P} = \mathbf{R}\,\mathbf{P_w} + \mathbf{T}$$



- 外参
  - 平移：T
  - 旋转：R

视觉类传感器

$$\mathbf{P} = \mathbf{R}\,\mathbf{P_w} + \mathbf{T}$$

旋转：

$$\mathbf{R} = \left(r_{ij}\right)_{3\times 3} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_1^T \\ \mathbf{R}_2^T \\ \mathbf{R}_3^T \end{bmatrix}$$

平移：

$$\mathbf{T} = (T_x, T_y, T_z)^T$$

也即：

$$\mathbf{P} = \mathbf{R}\mathbf{P_w} + \mathbf{T} = \begin{pmatrix} r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x \\ r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y \\ r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z \end{pmatrix} = \begin{bmatrix} \mathbf{R}_1^T\mathbf{P_w} + T_x \\ \mathbf{R}_2^T\mathbf{P_w} + T_y \\ \mathbf{R}_3^T\mathbf{P_w} + T_z \end{bmatrix}$$

视觉类传感器

帧坐标系到成像平面：

$$x = -(x_{im} - o_x)s_x$$
$$y = -(y_{im} - o_y)s_y$$

世界坐标系到相机坐标系（要在自己的坐标系下看）：

$$\mathbf{P} = \mathbf{R}\mathbf{P}_{\mathbf{w}} + \mathbf{T} = \begin{pmatrix} r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x \\ r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y \\ r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z \end{pmatrix} = \begin{bmatrix} \mathbf{R}_1^T\mathbf{P}_{\mathbf{w}} + T_x \\ \mathbf{R}_2^T\mathbf{P}_{\mathbf{w}} + T_y \\ \mathbf{R}_3^T\mathbf{P}_{\mathbf{w}} + T_z \end{bmatrix}$$

成像坐标系到帧坐标系

$$(x, y) = (f\frac{X}{Z},\ f\frac{Y}{Z})$$

投影到成像坐标系

完整的投影：

$$x_{im} - o_x = -f_x \frac{r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

$$y_{im} - o_y = -f_y \frac{r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

视觉类传感器

完整的投影：

$$x_{im} - o_x = -f_x \frac{r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

$$y_{im} - o_y = -f_y \frac{r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{M}_{\mathbf{int}} \mathbf{M}_{\mathbf{ext}} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

$$\mathbf{M}_{\text{int}} = \begin{bmatrix} -f_x & 0 & o_x \\ 0 & -f_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{M}_{ext} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix}$$

这就是Ransac
PNP的求解的目标
3D -> 2D

标定

著名的八点法

$$x' = x_{im} - o_x = -f_x \frac{r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

$$y' = y_{im} - o_y = -f_y \frac{r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

$$x' f_y (r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y) = y' f_x (r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x)$$

α = fx/fy

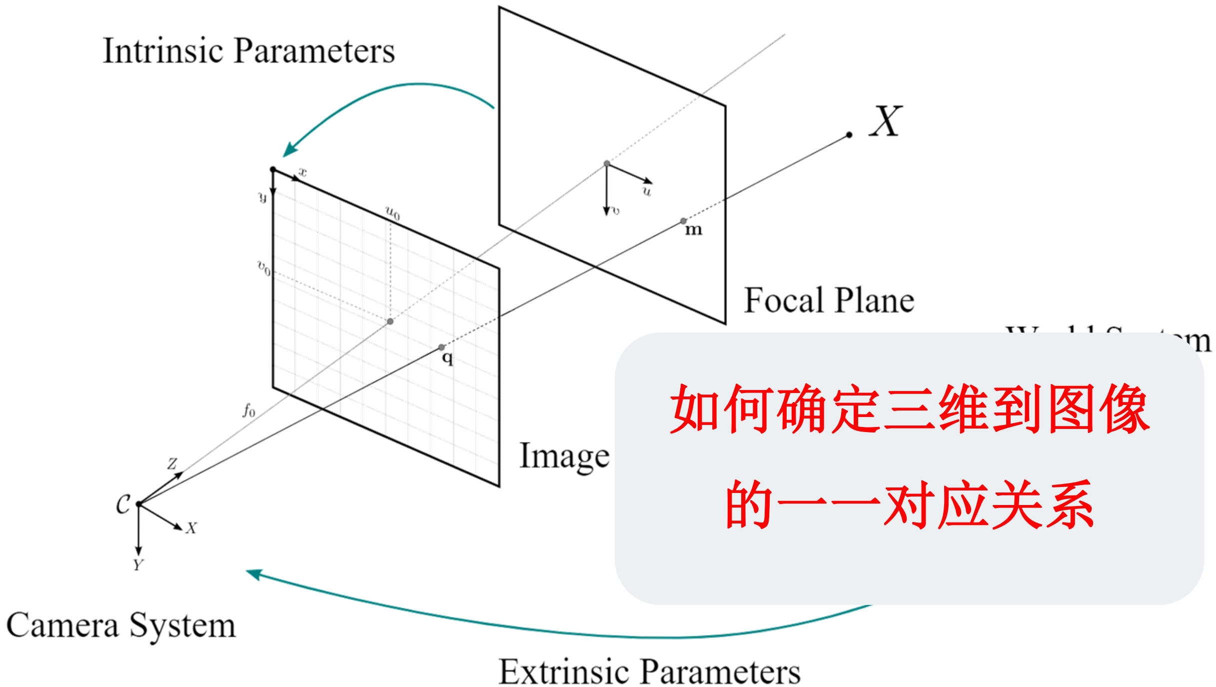
$$x'(r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y) = y' \alpha (r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x)$$

$$x_i X_i r_{21} + x_i Y_i r_{22} + x_i Z_i r_{23} + x_i T_y - y_i X_i (\alpha r_{11}) - y_i Y_i (\alpha r_{12}) - y_i Z_i (\alpha r_{13}) - y_i (\alpha T_x) = 0$$

$$x_i X_i v_1 + x_i Y_i v_2 + x_i Z_i v_3 + x_i v_4 - y_i X_i v_5 - y_i Y_i v_6 - y_i Z_i v_7 - y_i v_8 = 0$$

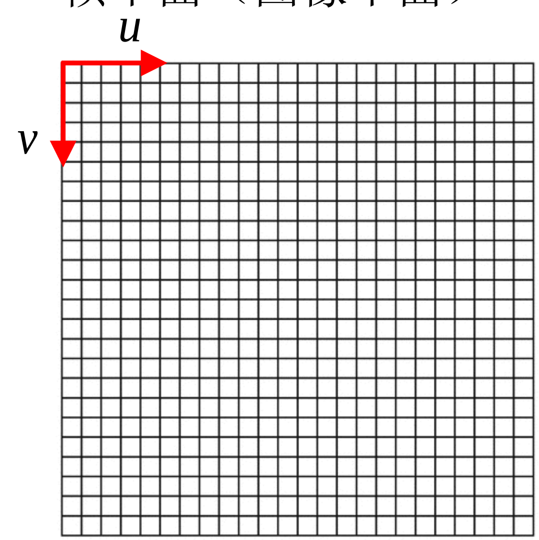
八个未知数，需要对应的八个点

视觉类传感器



如何确定三维到图像的一一对应关系

帧平面（图像平面）



世界到相机 → 3D到成像平面 → 成像到图像

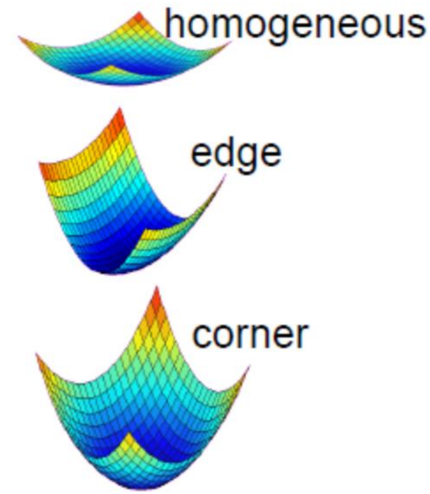
外参：也就是我们所说的单步旋转平移的解算

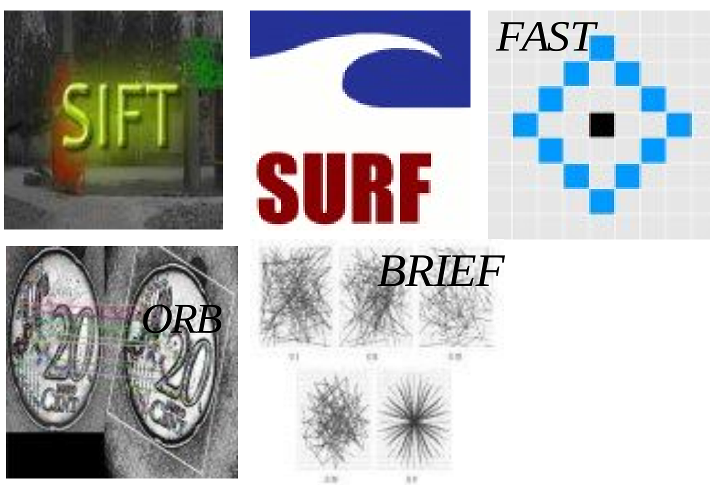
特征

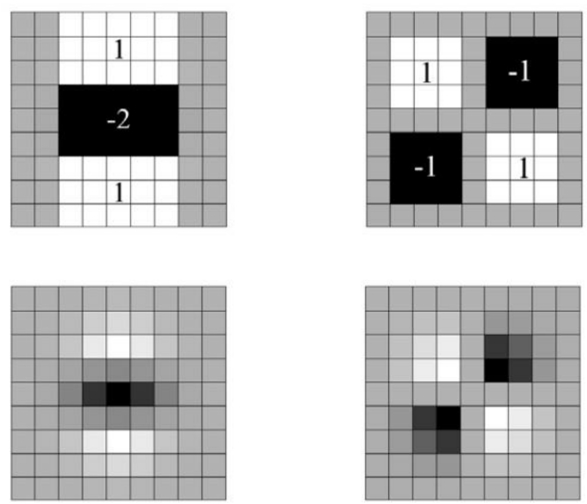
特征：

并没有什么统一、明确的定义，但是，想想它的用处：压缩了、将底层信息给予了更高层的表达。

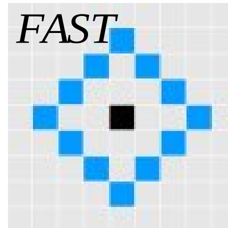
- 高对比度区域：一般在物体的边缘（视察）
- 避免Edge

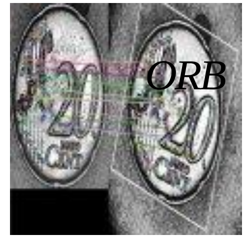


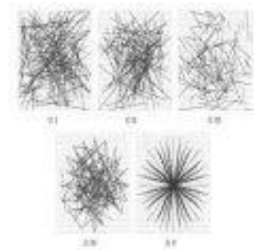




特征





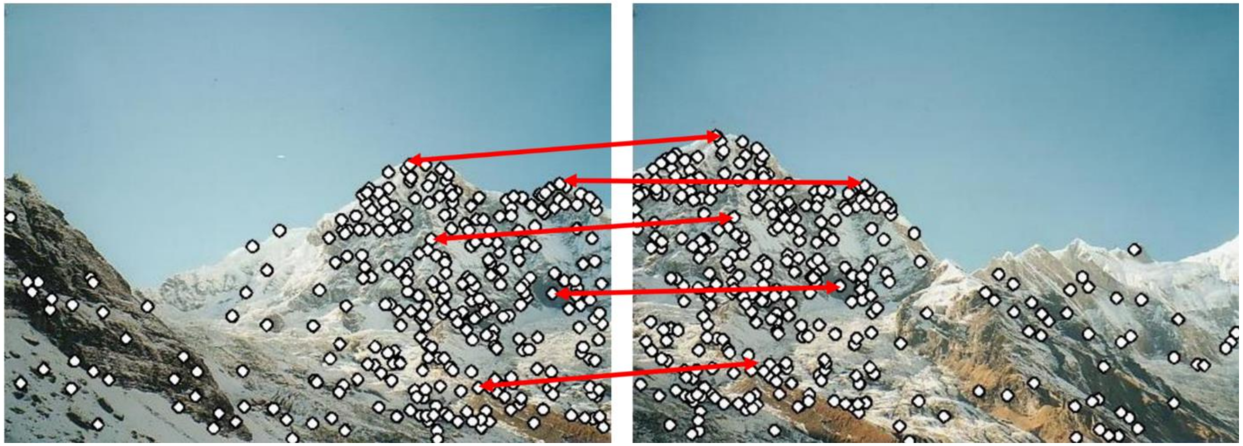


	Corner Detector	Blob Detector	Rotation Invariant	Scale Invariant	Affine Invariant	Repeatability	Localization Accuracy	Robustness	Efficiency
Haris	x		x			+++	+++	++	++
Shi-Tomasi	x		x			+++	+++	++	++
FAST	x		x	x		++	++	++	++++
SIFT		x	x	x	x	+++	++	+++	+
SURF		x	x	x	x	+++	++	++	++
CENSURE		x	x	x	x	+++	++	+++	+++

https://docs.opencv.org/3.3.0/db/d27/tutorial_py_table_of_contents_feature2d.html

特征

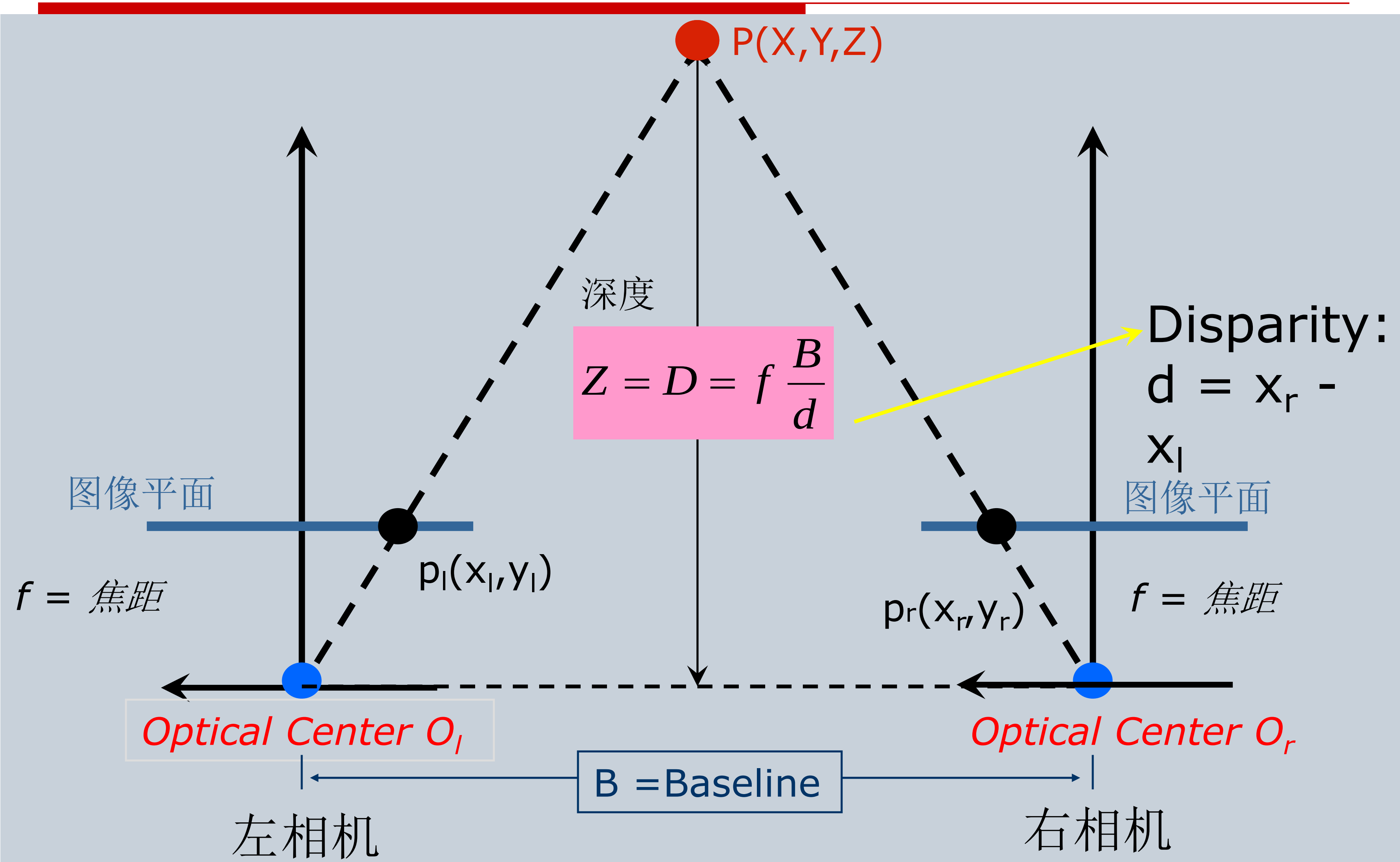
那么，特征提取和匹配怎么走？



特征检测 → 特征描述器 → 特征匹配

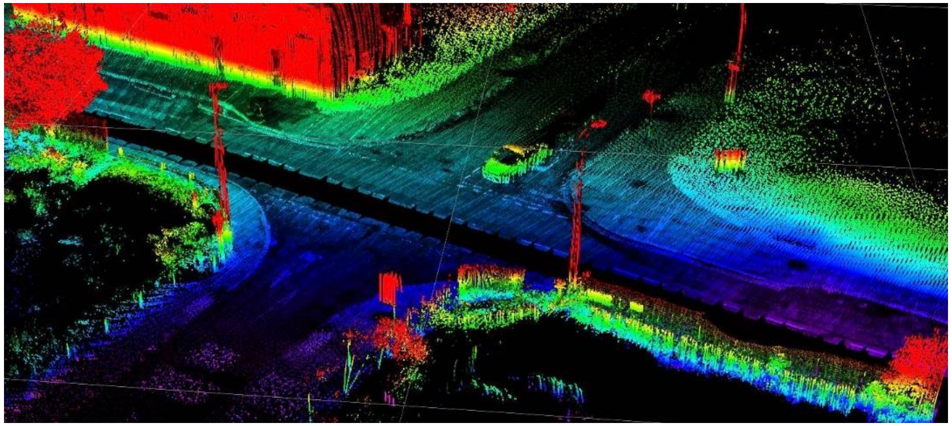
http://opencv-python-tutroals.readthedocs.io/en/latest/index.html

双目立体视觉


P(X,Y,Z)
深度
$Z = D = f\,\frac{B}{d}$
Disparity:
d = x_r - x_l
图像平面
图像平面
$p_l(x_l,y_l)$
$p_r(x_r,y_r)$
f = 焦距
f = 焦距
Optical Center O_l
Optical Center O_r
B =Baseline
左相机
右相机

提纲

- 激光原理与优势
- 激光的应用
- 激光特征和点云匹配

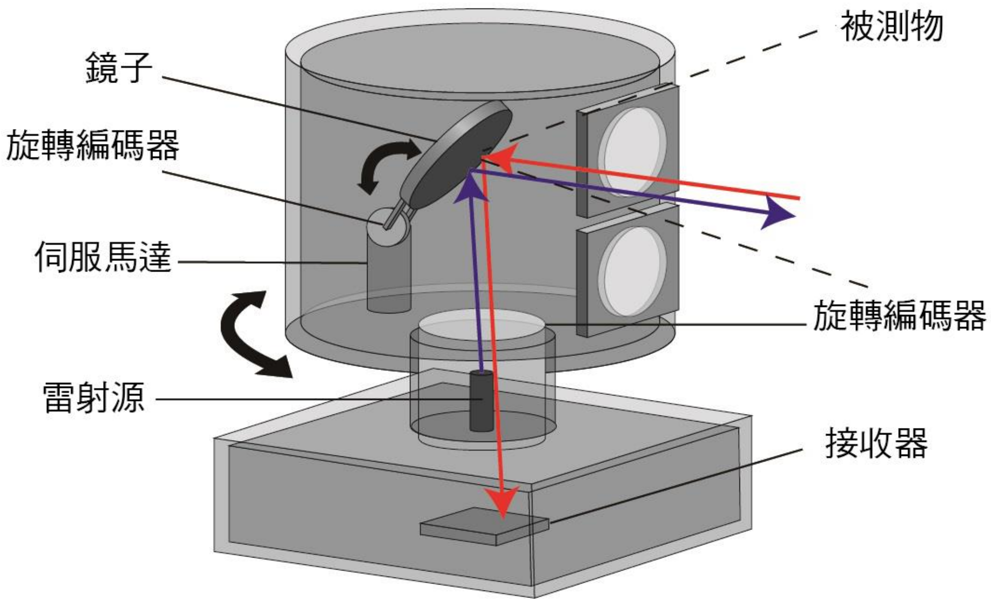


激光传感器

- 2D激光传感器

- 3D激光传感器

激光传感器成像原理



Laser ID	Vertical Angle
0	-15°
1	1°
2	-13°
3	-3°
4	-11°
5	5°
6	-9°
7	7°
8	-7°
9	9°
10	-5°
11	11°
12	-3°
13	13°
14	-1°
15	15°

世界到激光传感器 → 3D点云到车体坐标系

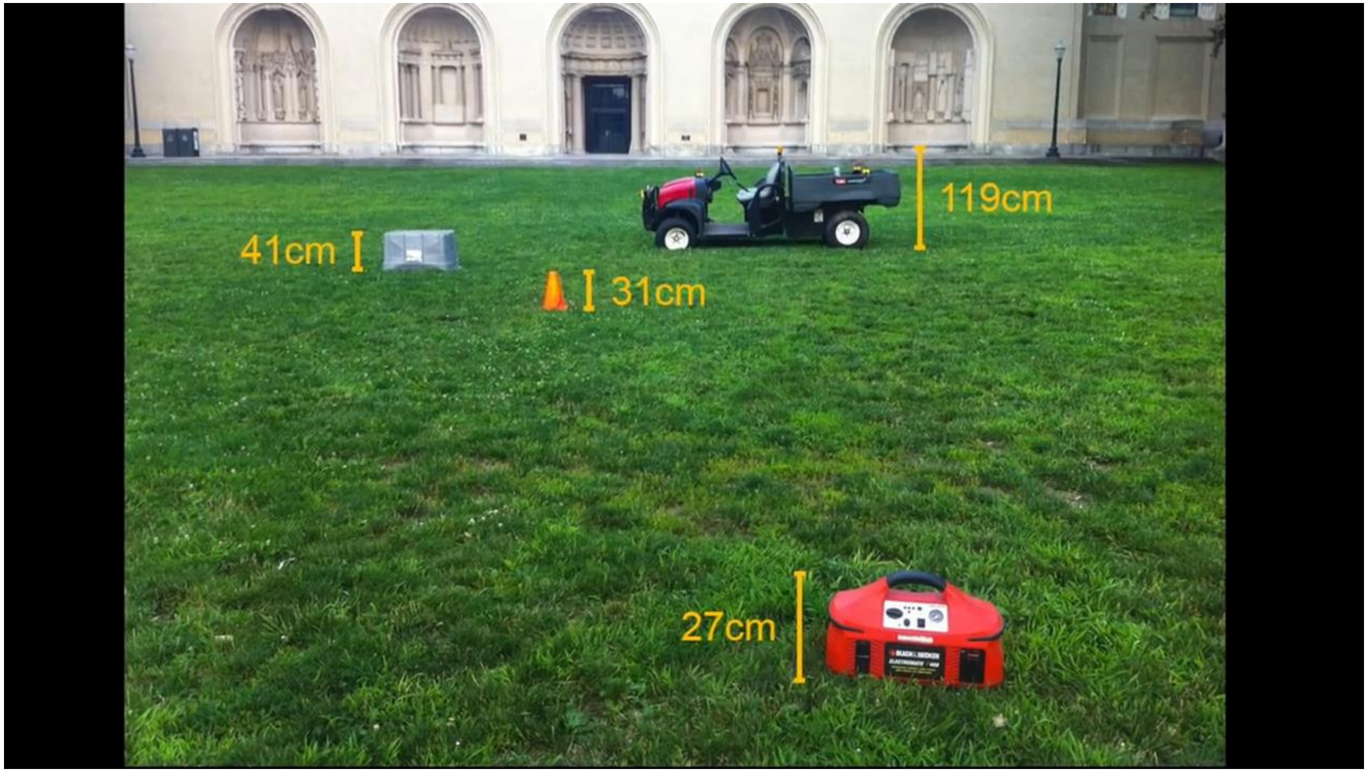
外参：激光传感器与车体几何中心的转化关系

激光信息特性

- 激光优势
  - 不受光照、季节条件影响，适合全日时运行；
  - 检测精度高，有效检测范围可达到70~100m
  - 可以根据对象材质的不同波长的反射回率
- 激光劣势
  - 价格高昂
  - 收特殊天气（雨，雪，雾，沙尘）等天气影响
  - 无法返回丰富纹理信息，且分辨率低，视野受限。

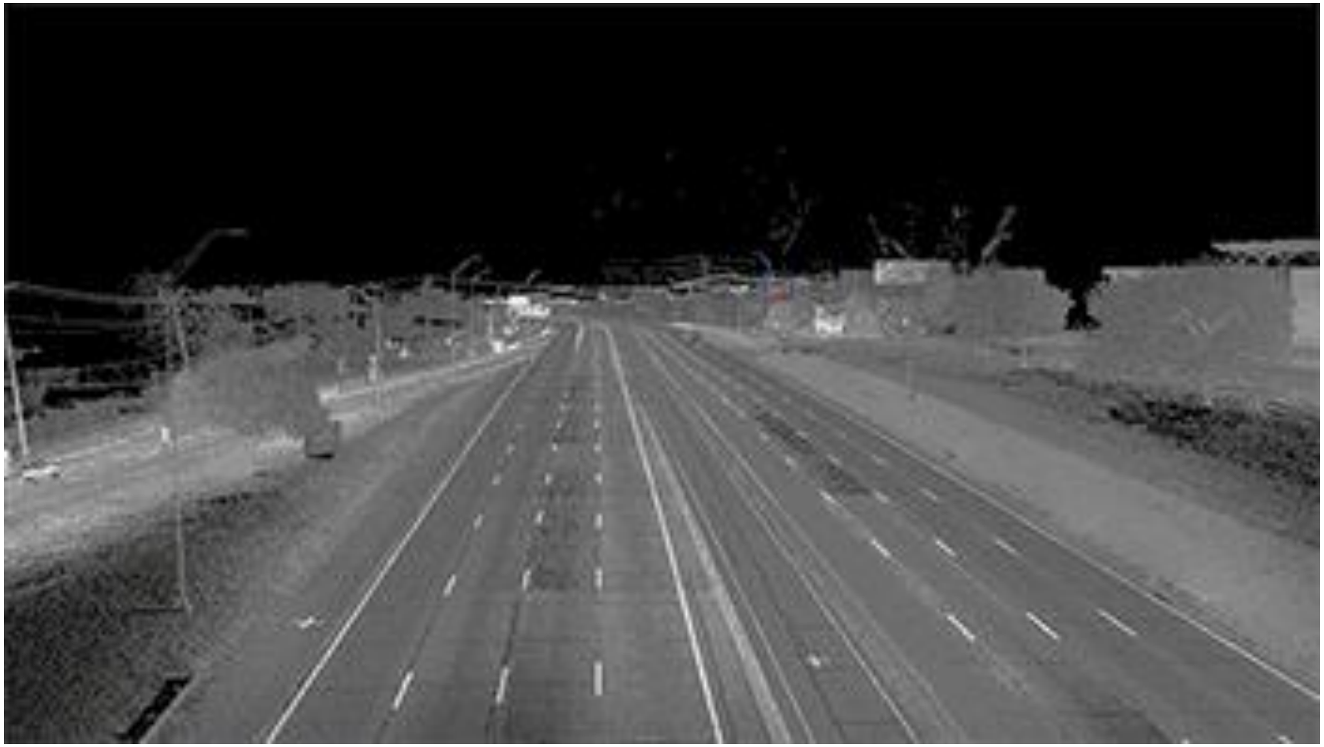
激光的应用价值

- 激光里程计



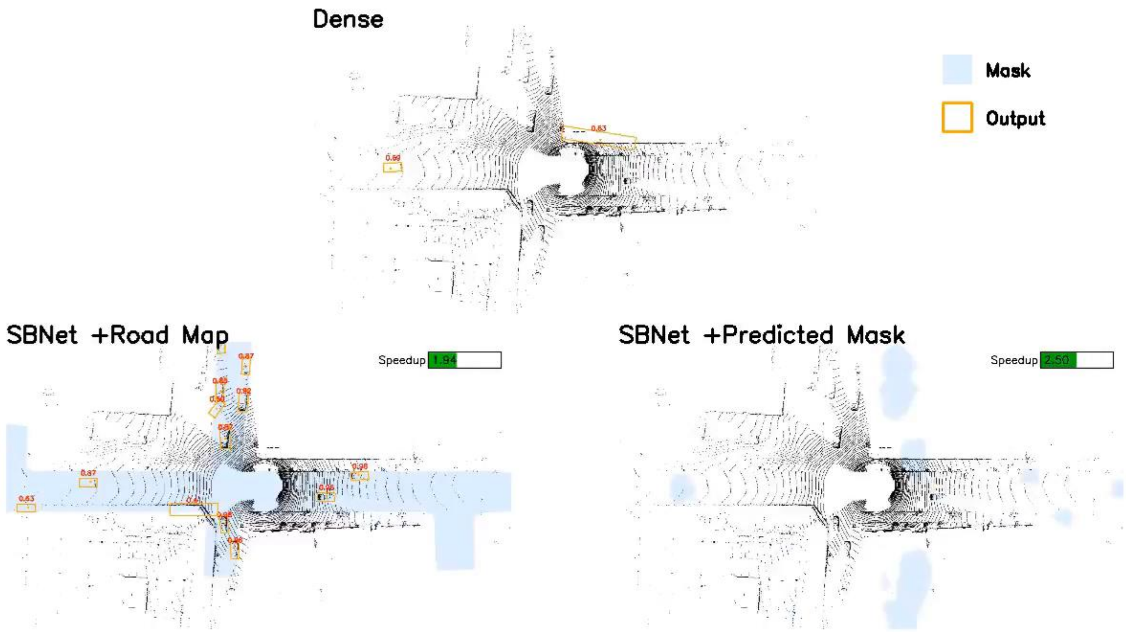
激光的应用价值

- 高精度离线地图



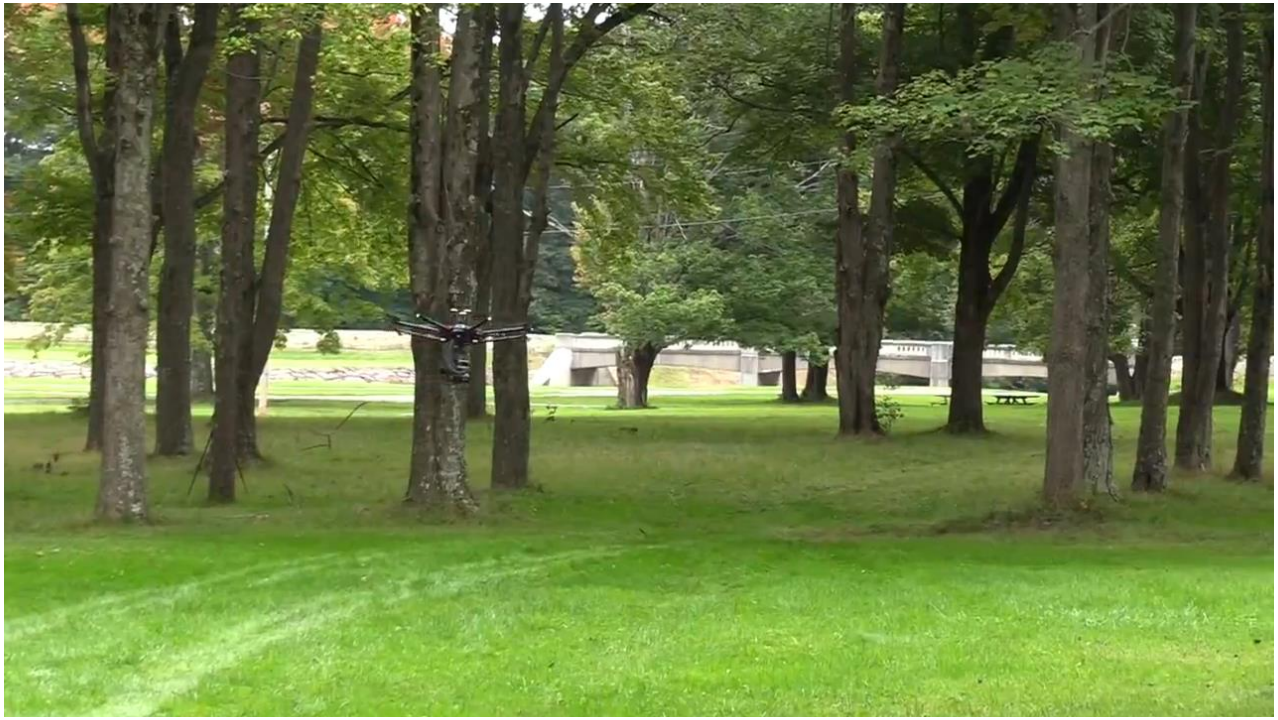
激光的应用价值

多目标识别



激光的应用价值

- 静态障碍物设备与实时运动规划



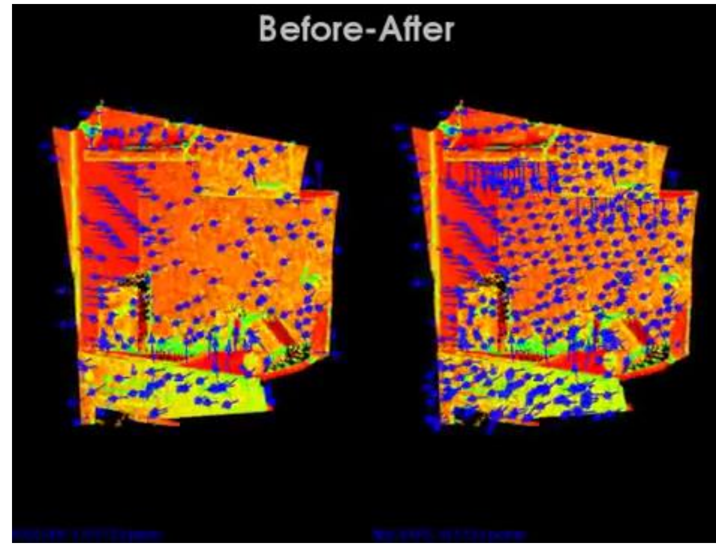
激光特征与匹配

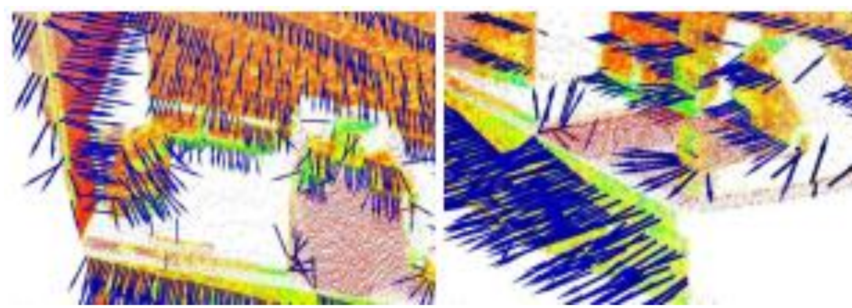
- 表面法向量 （Surface Normal ）
- 点特征直方图 （Point Feature Histograms）
- 旋转图片特征 （Spin-Image）
- 迭代最近邻点（Iterative Closest Points）

激光特征：法向量特征

法向量计算：

1. 曲面重建技术，从获取的点云数据集中得到采样点所对应的曲面，然后从曲面模型中计算表面法线；
2. 根据邻近点云数据集近似推断表面法向量（最小二乘法或者PCA方法）

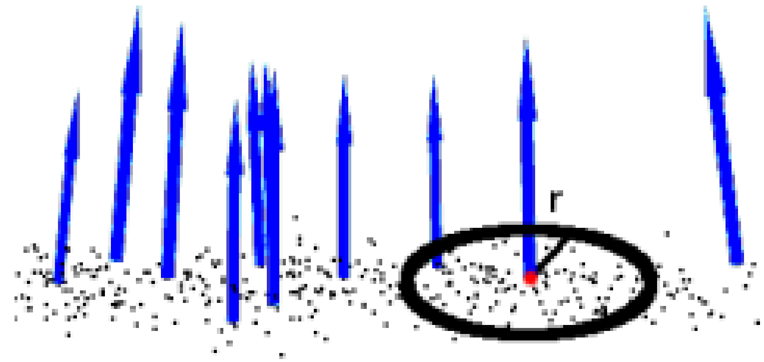




激光特征：法向量特征

计算最近点均值信息

$$x = \bar{p} = \frac{1}{k} \cdot \sum_{i=1}^{k} p_i$$



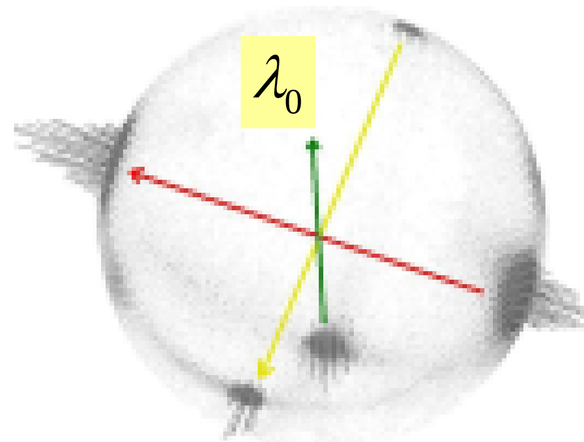
估计最小二乘城协方差矩阵

$$C = \frac{1}{n} \sum_{i=1}^{n} \zeta_i \left(P_i - \bar{P}\right)\left(P_i - \bar{P}\right)^T ,$$

$$C \cdot \vec{v}_j = \lambda_j \cdot \vec{v}_j , j \in \{0,1,2\}$$

$$0 \le \lambda_0 \le \lambda_1 \le \lambda_2$$

法向向量

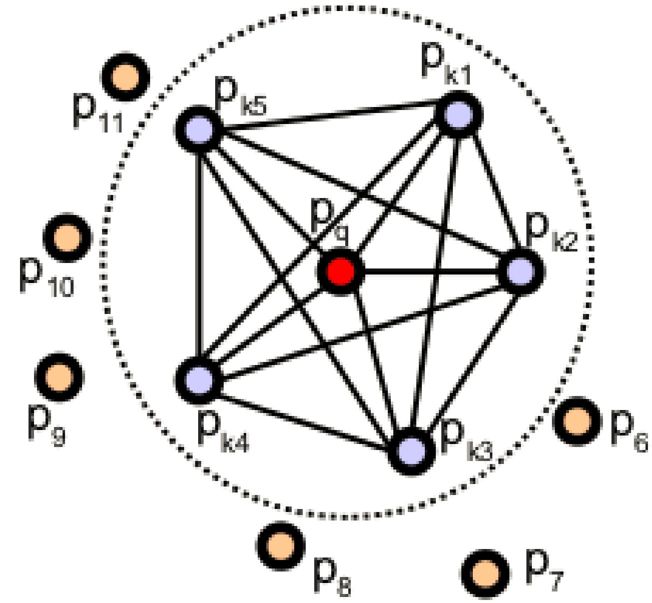


$$\sigma_p \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2}$$

曲率信息

激光特征：点特征直方图

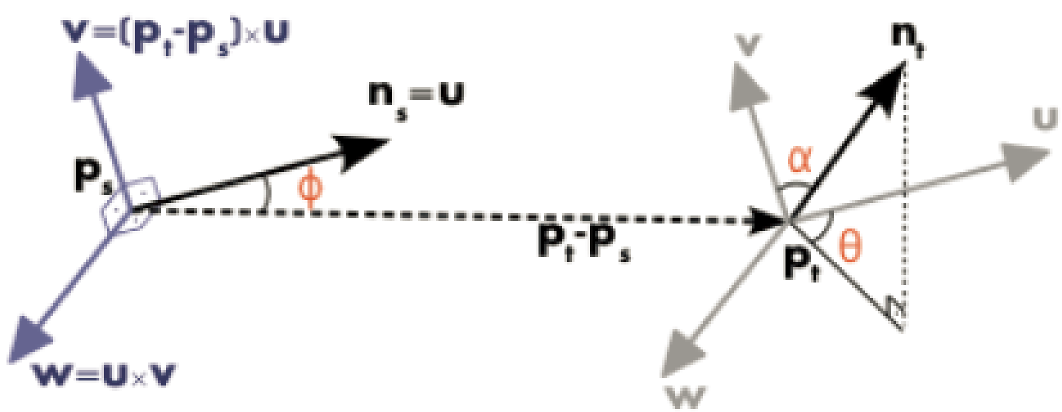
虽然表面法向量可以用于快速计算几何对象的法向量和曲率信息，但是无法提取更丰富的特征信息。如果想衡量两个姿态点之间的关联度，我们需要12量的描述空间。



$$u = n_s,$$

$$v = u \times \frac{(p_t - p_s)}{\|p_t - p_s\|_2},$$

$$w = u \times v$$



激光特征：点特征直方图

$$d = \left\| p_t - p_s \right\|_2 ,$$

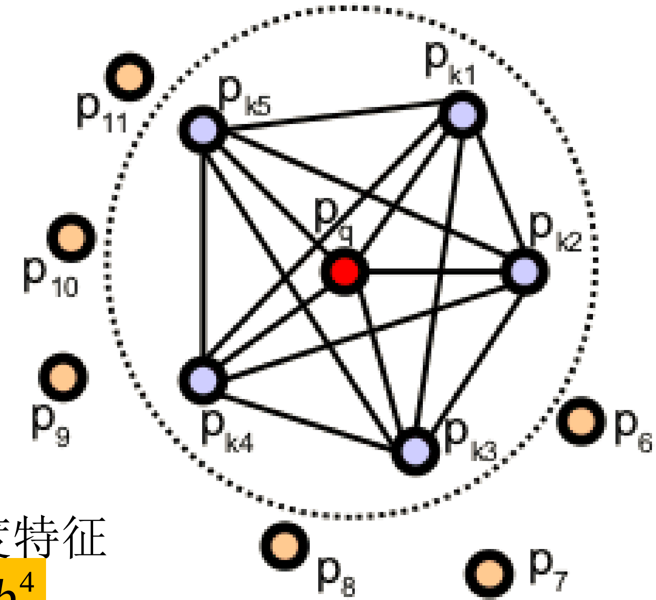
$$\alpha = v \cdot n_t ,$$

$$\phi = u \times \frac{(p_t - p_s)}{d} ,$$

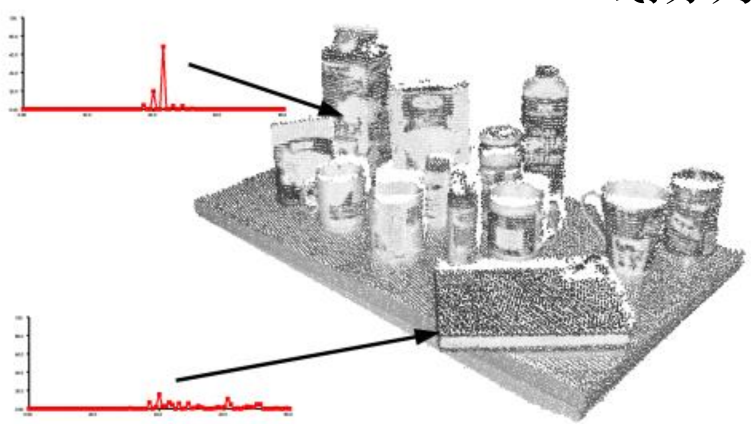
$$\theta = \arctan\left(w \cdot n_t , u \cdot n_t \right)$$

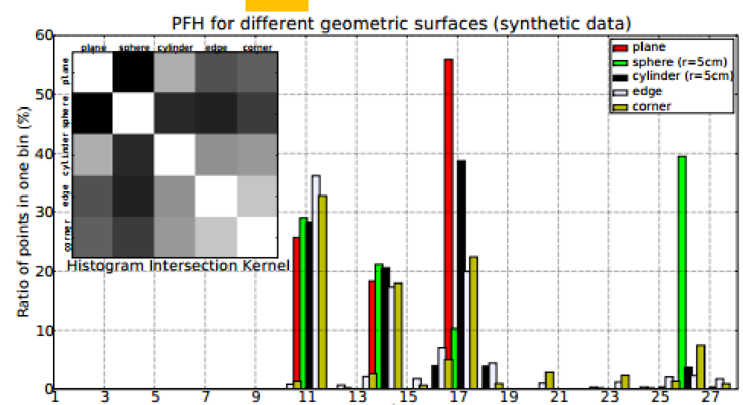
只需要4个变量

$$\langle d, \alpha, \phi, \theta \rangle$$



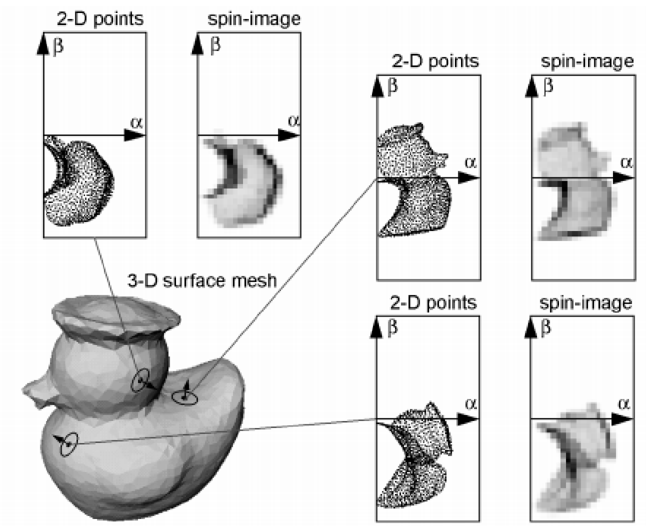
建立直方图，每一维度特征划分为b，bin的大小为 b^4

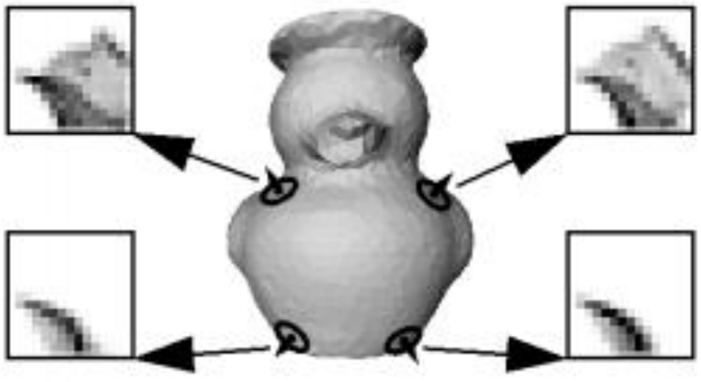




激光特征：旋转图片

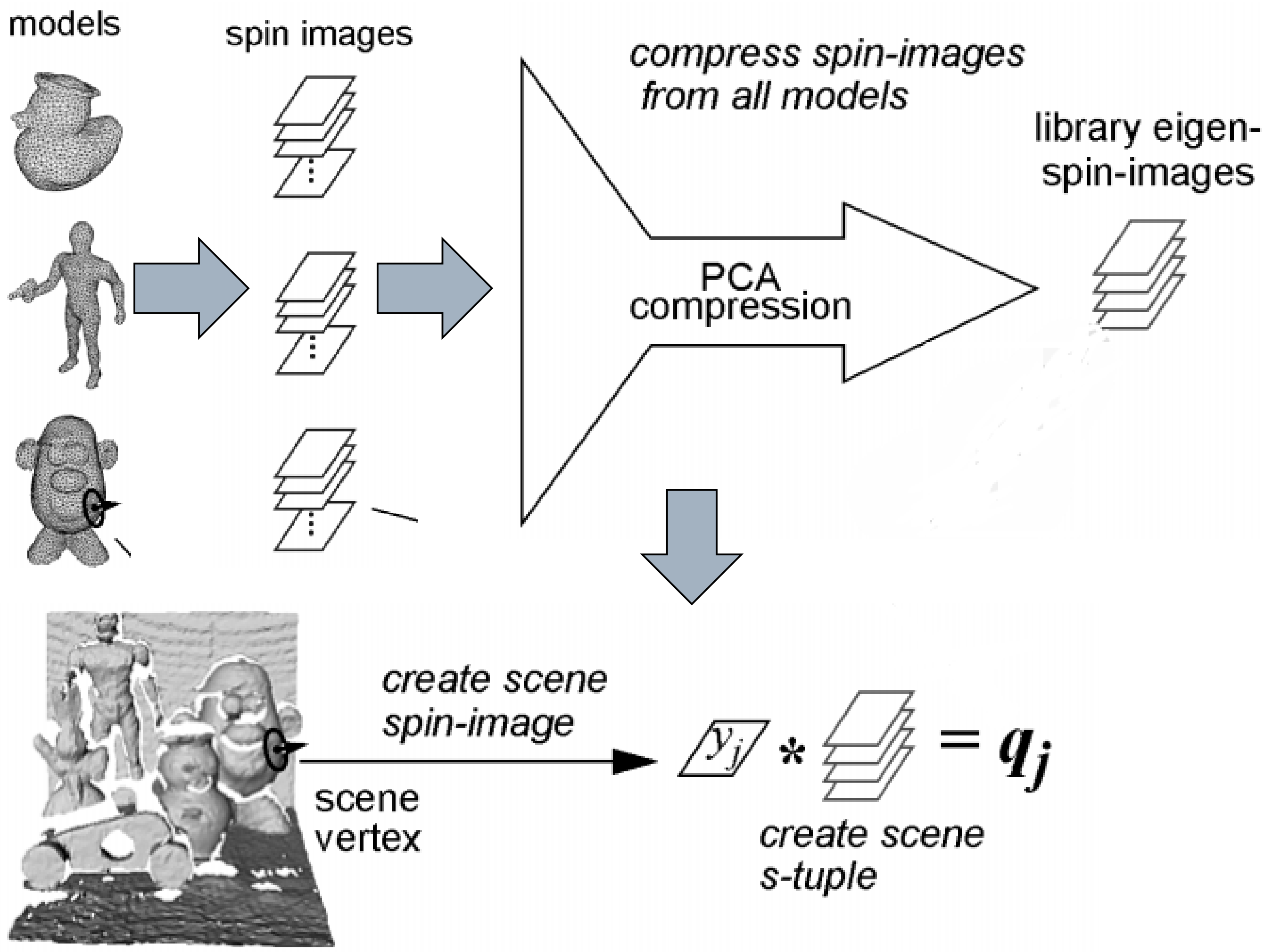
到目前为止，所有的特征只能计算局部的法向量统计特性，无法提取更丰富的几何特性，如视觉信息可以提取纹理信息，那么是否也有类似的方法提取点云的‘纹理’信息呢？





(b)

激光特征：旋转图片



models
spin images
compress spin-images
from all models
library eigen-
spin-images
PCA
compression
create scene
spin-image
scene
vertex
y_j * $= q_j$
create scene
s-tuple

激光点云匹配

迭代最近邻匹配（ICP）：

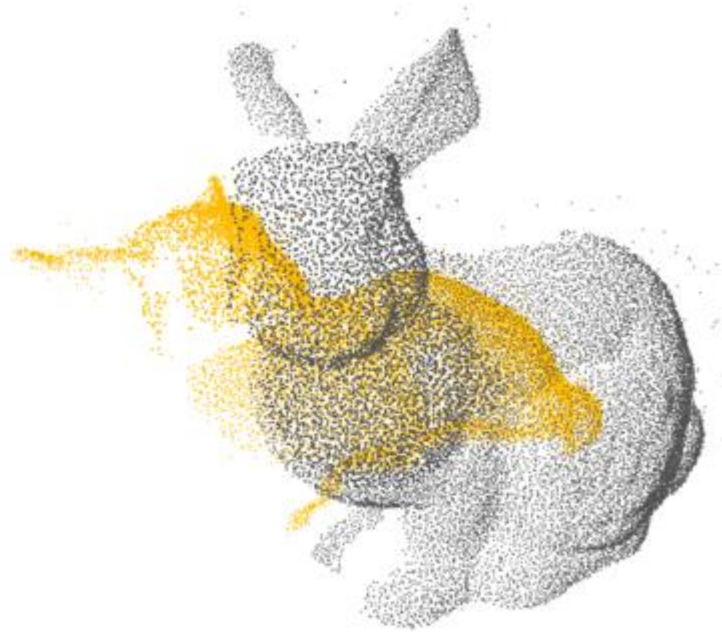
给定两组待匹配点云

$$X = \{x_1, ..., x_n\}$$

$$P = \{p_1, ..., p_n\}$$

最小化一下方程，求得最优的旋转平移矩阵。

$$E(R, \mathrm{t}) = \frac{1}{N_p} \sum_{i=1}^{N_p} \left\| x_i - Rp_i - t \right\|^2$$



激光点云匹配

$$E(R,\mathrm{t}) = \frac{1}{N_p}\sum_{i=1}^{N_p}\left\| x_i - Rp_i - t \right\|^2$$

$$\mu_x = \frac{1}{N_x}\sum_{i=1}^{N_x} x_i, \mu_p = \frac{1}{N_p}\sum_{i=1}^{N_p} p_i$$

SVD求解矩阵W

$$W = \sum_{i=1}^{N_p}(x_i - \mu_x)(p_i - \mu_p)$$

$$= \mathrm{U}\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} V^T$$

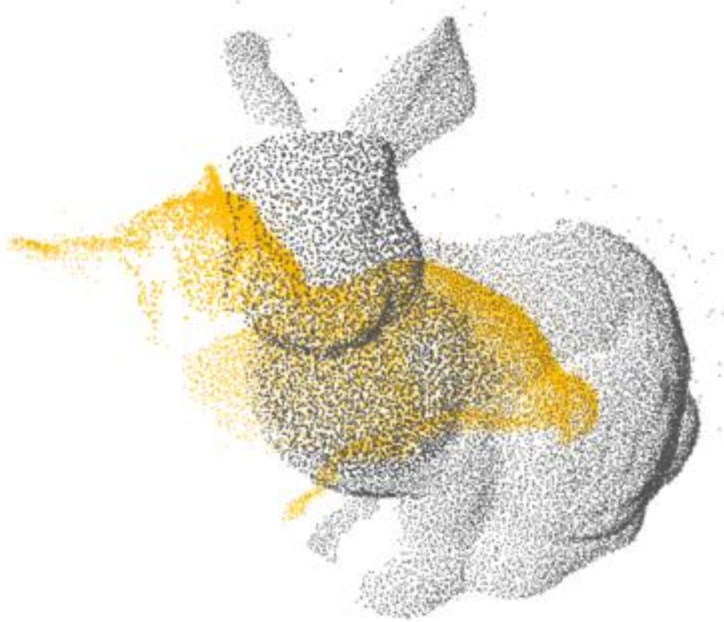
$$U, V \in R^{3x3}$$

$$\sigma_1 \ge \sigma_2 \ge \sigma_3$$

$$R = UV^T$$

$$t = \mu_x - R\mu_p$$

$$E(R,t) = \sum_{i=1}^{N_P}\left(\left\| x_i - \mu_x \right\|^2 + \left\| y_i - \mu_y \right\|^2\right) - 2(\sigma_1 + \sigma_2 + \sigma_3)$$



Reading and Homework

Reading:

- Rusu, Radu Bogdan. "Semantic 3d object maps for everyday manipulation in human living environments." *KI-Künstliche Intelligenz* 24.4 (2010): 345-348.
- Johnson, Andrew E., and Martial Hebert. "Using spin images for efficient object recognition in cluttered 3D scenes." *IEEE Transactions on pattern analysis and machine intelligence*21.5 (1999): 433-449.
- Besl, Paul J., and Neil D. McKay. "A method for registration of 3-D shapes." *IEEE Transactions on pattern analysis and machine intelligence* 14.2 (1992): 239-256.
- Rusu, Radu Bogdan, Nico Blodow, and Michael Beetz. "Fast point feature histograms (FPFH) for 3D registration." *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009.

Homework, try and understand the basic ICP method.

- https://github.com/ClayFlannigan/icp.git

Q&A