深蓝学院
shenlanxueyuan.com

视觉SLAM理论与实践

选题 5：基于检测框的重建

Outline

深蓝学院
shenlanxueyuan.com

- Part1: Define the question
- Part2: Method

Part1: Define the question

- Why object?

  1. SLAM maps need immediate semantic information.

  2. Depend on how we introduce the semantic information:

  Detection : CubeSLAM, QuadricSLAM, ClusterVO, etc.

  Segmentation : AVP-SLAM, MID-Fusion,etc.

Part1: Define the question

- Why Quadric?

  1. Compact Representation(9 degrees of freedom)

  2. Capture size, position and orientation of object

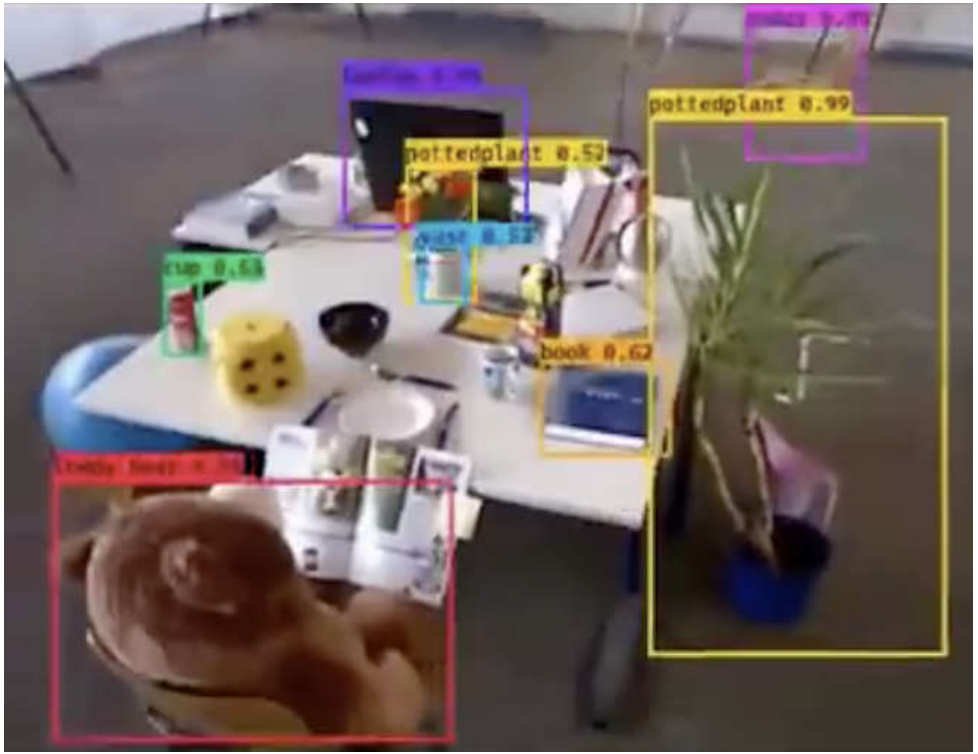
  3. Convenient

- Integration

  Solving for camera pose, landmark pose and shape parameters simultaneously.

Method

- Pre-required
- Problem setup
- Initialization



Method : Pre-required

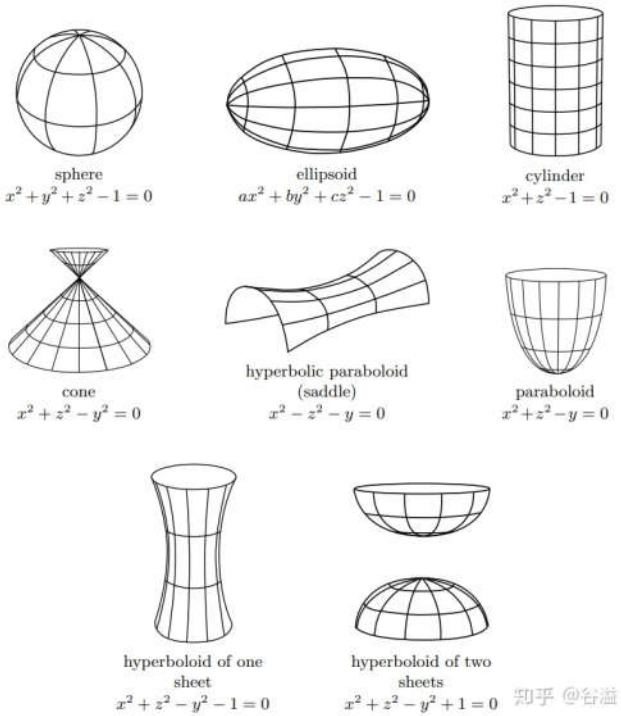
- Quadrics(projective geometry)

General expression：

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + J = 0$$

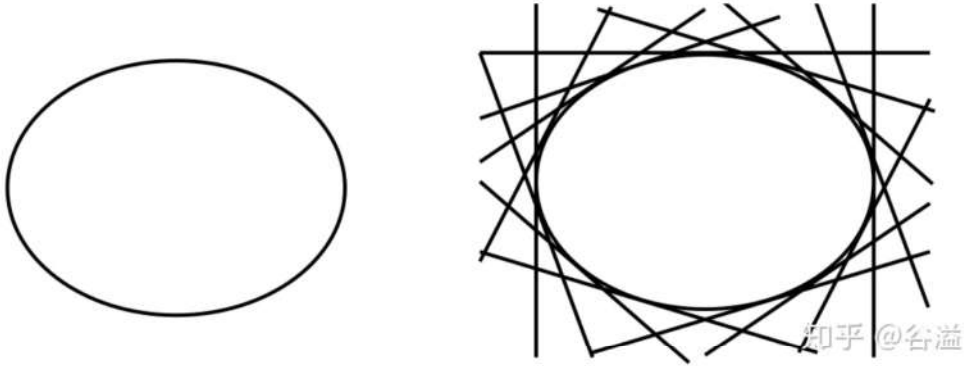
More elegant：$\mathbf{x}^t Q\mathbf{x} = 0$

$$\text{where } \mathbf{x} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}, \mathbf{x}^t = \begin{bmatrix} x & y & z & w \end{bmatrix}, \text{and } Q = \begin{bmatrix} a & b & c & d \\ b & e & f & g \\ c & f & h & i \\ d & g & i & j \end{bmatrix}$$



Method : Pre-required

- Dual conic



左边是二次曲线，右边是对偶二次曲线

二次曲线的表达式：

$$\mathbf{p}^T C \mathbf{p} = 0$$

对偶二次曲线的表达式：

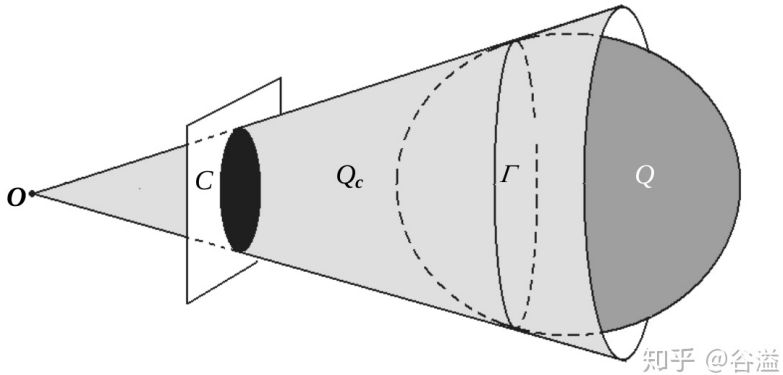
$$\mathbf{l}^T C^* \mathbf{l} = 0$$

Method : Pre-required

- Dual quadrics

In dual form, a quadric is defined by a set of tangential planes such that the planes form an envelope around the quadric.

$$\pi^T Q^* \pi = 0$$



Method : Pre-required

设相机投影矩阵为P： $P = K[R|t]$ ，包含了相机的内外参，现在我给你一个结论：

$$C^* = PQ^*P^T$$

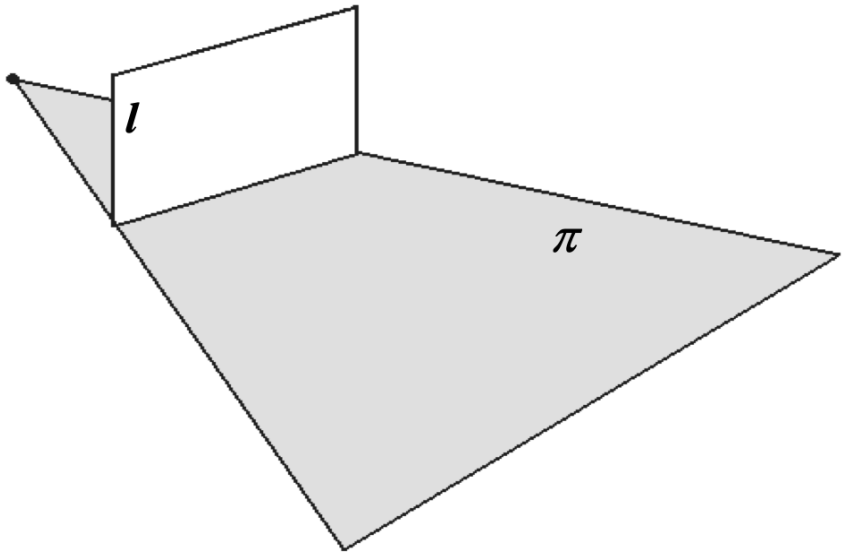
然后我要证明它：

C^* 中的任意一条直线，有 $l^T C^* l = 0$ ，而 l 与光心O构成一个平面（记作 π ，也叫做直线l的反投影平面，这里我们一直把它作为列向量，其实就是平面参数ABCD），这个平面与二次曲面相切，

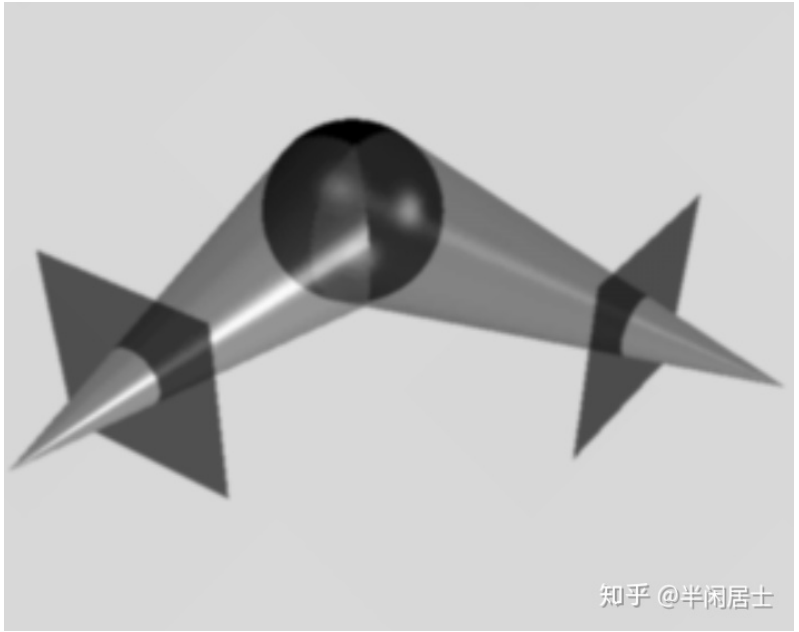
对于空间点X，假设它在平面 π 上面，那么投影到图像上，也满足直线方程： $l^T(PX) = 0$ ；

也就是 $X^T(P^T l) = 0$ ， 所以 $\pi = P^T l$ ，而 $\pi^T Q^* \pi = 0$ ，所以 $l^T PQ^* P^T l = 0$ ， 至此也就得到了 $C^* = PQ^*P^T$ ，也就是有了C，因为转换关系很简单。

Method : Pre-required



图像直线 l 的反投影平面是 $\pi = P^T l$



通过多帧观测恢复曲面

Method : Optimization

深蓝学院 shenlanxueyuan.com

为什么引入 Q^* 而不是直接使用 Q ？

Q^* 是Q的对偶形式，二者很容易转换，使用 Q^* 是为了和检测框一起使用，如果你的椭球算的很准，框检测的也很准，那么检测框应该正好框住这个椭球的投影，也就是相切关系，你应该联系到了，直线方程有了，假设相机内外参你也有了，学计算机的人都知道要干嘛了：

$$\hat{Q}^* = \arg\min_a ||l_{i,k}^T P_i Q^* P_i^T l_{i,k}||^2$$

Method : Optimization

深蓝学院 shenlanxueyuan.com

Constrained Dual Quadric Parametrization

Similar to [14], we parametrize dual quadrics as:

$$\mathbf{Q}^* = \mathbf{Z}\,\breve{\mathbf{Q}}^*\,\mathbf{Z}^\mathsf{T} \tag{1}$$

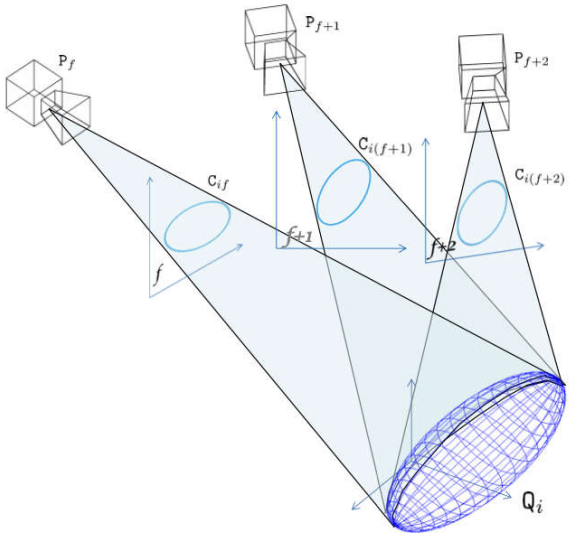
where $\breve{\mathbf{Q}}^*$ is an ellipsoid centred at the origin, and $\mathbf{Z}$ is a homogeneous transformation that accounts for an arbitrary rotation and translation. Specifically,

$$\mathbf{Z} = \begin{pmatrix} \mathbf{R}(\boldsymbol{\theta}) & \mathbf{t} \\ \mathbf{0}_3^\mathsf{T} & 1 \end{pmatrix} \text{ and } \breve{\mathbf{Q}}^* = \begin{pmatrix} s_1^2 & 0 & 0 & 0 \\ 0 & s_2^2 & 0 & 0 \\ 0 & 0 & s_3^2 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{2}$$

椭球面的加约束形式

$$\mathbf{q} = (\theta_1, \theta_2, \theta_3,\ t_1, t_2, t_3,\ s_1, s_2, s_3)^\mathsf{T}$$

优化变量



Method : Optimization

深蓝学院 shenlanxueyuan.com

Constrained Dual Quadric Parametrization

有了Q的估计之后，根据之前的二次曲面投影公式，我们能在图像上得到一个投影，也就是说，在图像上，我能预测一个box，用公式概括就是：

$$\beta(x_i, q_i) = \hat{b}_{i,j}$$

有了预测和检测结果，就可以构造geometric error做优化。公式都给了，按着推一下就行了，计算bounding box有点说法，首先你得知道正常case下，怎么根据椭圆预测box，然后再想想特例，比方与图像边界交叉的case，最后按着作者给的算法过一遍，很快就明白了：

Method : Optimization

深蓝学院 shenlanxueyuan.com

接下来就是根据概率模型，做MAP估计，比较常规，照搬了：

$$
\begin{aligned}
X^*, Q^* &= \underset{X,Q}{\operatorname{argmin}} -\log P(X, Q|U, B) \\
&= \underset{X,Q}{\operatorname{argmin}} \underbrace{\sum_i \|f(\mathbf{x}_i, \mathbf{u}_i) \ominus \mathbf{x}_{i+1}\|^2_{\Sigma_i}}_{\text{Odometry Factors}} \\
&\quad + \underbrace{\sum_{ij} \|\mathbf{b}_{ij} - \boldsymbol{\beta}_{(\mathbf{x}_i, \mathbf{q}_j)}\|^2_{\Lambda_{ij}}}_{\text{Quadric Landmark Factors}}
\end{aligned}
$$

LM也好，DOG-LEG也好，把它做出来就行了，Odometry Factors你随意发挥，哪怕直接拿gt来做我认为都可以，那就不叫slam了，但是可以先验证一下后面的项的求解是否正确。

Initialization

深蓝学院 shenlanxueyuan.com

类似于常规SLAM的操作，初始化的时候一般会做一个SFM，把点固定下来，这里是固定一些物体的二次曲面

$$\boldsymbol{\pi}_{ijk}^{\mathsf{T}} \mathbf{Q}_{(\hat{\mathbf{q}}_j)}^{*} \boldsymbol{\pi}_{ijk} = 0$$

Reshape:

$$(\pi_1^2,\ 2\pi_1\pi_2,\ 2\pi_1\pi_3,\ 2\pi_1\pi_4,\ \pi_2^2,\ 2\pi_2\pi_3,\ \ldots,\ 2\pi_2\pi_4,\ \pi_3^2,\ 2\pi_3,\ \pi_4^2) \cdot (\hat{q}_1, \hat{q}_2, \ldots, \hat{q}_{10})^{\mathsf{T}} = 0$$

$$\begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} = \left| \sqrt{ -\frac{\det \mathbf{Q}}{\det \mathbf{Q}_{33}} \begin{pmatrix} \lambda_1^{-1} \\ \lambda_2^{-1} \\ \lambda_3^{-1} \end{pmatrix} } \right|$$

这部分水平有限，推不出来

Remember to constrain it to be a ellipsoid, refer to [2]

Reference

QuadricSLAM: Dual Quadrics from Object Detections as Landmarks in Object-oriented SLAM

3D Object Localisation from Multi-view Image Detections

计算机视觉中的数学方法

https://www.zhihu.com/question/394814665/answer/1250621518

在线问答

深蓝学院
shenlanxueyuan.com

Q&A

深蓝学院
shenlanxueyuan.com

感谢各位聆听！

Thanks for Listening