

CubeSLAM 笔记

- 论文: [CubeSLAM_Monocular_3-D_Object_SLAM.pdf](#)

先前已有关于CubeSLAM的文档[Dynamic SLAM总结.pdf](#)，包含了论文的主要内容。本篇文档主要对其中未提到的细节进行一些补充。

1. 关于VP点的一些理解

VP点的坐标是二维的，公式(1)是齐次意义下相等，在归一化平面下，相等相差一个常数倍。

VP_i = K R_{C₀1} u_i R为相机系下, Cube的旋转

灭点为Cube系中的无穷远点在图像平面上的成像。考虑齐次坐标 (a, 0, 0, 1)

Cube系 $\begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} R \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix} + t \\ 1 \end{bmatrix}$

相机系下坐标 $R \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix} + t \rightarrow a R_{C_01}(1) + t$

$(a R_{31} + t) \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K (a R_{C_01}(1) + t)$

(u, v) 为成像坐标。令 $a \rightarrow \infty$

则 $VP_i = \lim_{a \rightarrow \infty} \frac{1}{a R_{31} + t} K (a R_{C_01}(1) + t)$

$= \frac{1}{R_{31}} K R_{C_01}(1) \quad (\text{归一化平面下})$

公式(1)的推导

为了理解VP点以及确定顶点坐标的过程，考虑这样的情况：两个立方体在相机系下旋转相同，只相差平移。那么两个物体的各条边都对应地互相平行，二维成像下必相交于相同坐标的三个灭点VP。固定了VP以后，各选取两个立方体上的一个顶点，得到p₁, q₁。p₁, q₁在图像平面下的坐标不相同(也有相同的可能，但他们对应的不是同一灭点)，从而利用bbox确定出来两对不同的顶点(共8个)，这8对顶点与给出的两个bbox都分别良好对应的。

2. 关于yaw角与顶点的计算

立方体被认为在地面上，故旋转矩阵可由yaw角确定。深度学习给出二维图像目标检测的bbox，CubeSLAM在bbox的边上进行等距采样，同时对yaw角进行采样。两者的一个组合可以对物体生成一个三维立方体边框，对每个立方体执行我们的打分(score)，选择得分最高的作为最终的结果。

- 然后以初始化的 yaw 角为中心的 - 45° 到 + 45° 的 90° 范围内，每隔 6° 采样一个值，采样得到 15 个偏航角。

Code

```
1 std::vector<double> obj_yaw_samples;  
2 // BRIEF linespace()函数从 a 到 b 以步长 c 产生采样的 d.  
3 linespace<double>(yaw_init - 45.0/180.0*M_PI, yaw_init + 45.0/180.0*M_PI, 6.0/180.0*M_PI, obj_yaw_samples);
```

- 为确定物体的“顶点”需要从原始边界框的最左边 `left_x_raw+5` 到最右边 `right_x_raw-5` 每隔 `top_sample_resolution` (20像素) 的距离采样一个点 `top_x_samples[i]`.

Code

```
1 int top_sample_resolution = round(min(20,obj_width_raw/10 )); // 20 pixels  
2 std::vector<int> top_x_samples;  
3 linespace<int>(left_x_raw+5, right_x_raw-5, top_sample_resolution, top_x_samples);
```

- 为确定物体的顶部，提供至少 10 个采样点（后期通过最小误差得到最合适的点），对于越小的物体需要越精细的采样。上边缘采样点如下图所示：

3. 关于立方体提案打分

考虑距离误差时提到除以对角线长，主要是减少尺度上带来的影响。

本文介绍时提到我们不对物体预设任何先验，但在打分的形状损失上，我们对长宽比的不一致进行了惩罚，实际上要求对物体的形状有一定的先验认识。

4. 关于(静态)数据关联

通过描述子匹配与对极几何检查对(静态)特征点进行匹配(动态物体上的特征点不加入数据关联)。

进行点-物体的关联，通过2-D的bbox与3-D的距离量将特征点关联到物体上，这个关联信息需要在BA过程中物体-点测量中被使用到。

进行物体-物体的关联，不同帧下的物体若共享一定数量的特征点，则对两个物体进行关联。



Fig. 5. Object association in the dynamic and occluded scenarios of KITTI 07. Green points are the nonobject points, and points in other colors are associated with objects of the same color. The front cyan moving car is not added as a SLAM landmark as no feature point is associated with it. Points in object overlapping areas are not associated with any object due to ambiguity.

5. 关于动态SLAM部分

相比于静态SLAM的BA增加了两个新的动态因子，包括物体运动因子与点-相机-物体因子。

