**SLAM教程-直接法**

**卫翀**

**2021.05.15**

1. **SLAM的本质**

假设采用RGB-D相机采集数据，在各个时刻通过传感器获取该时刻的RGB图像以及点云（景深）信息。相机在每个时刻（例如间隔0.1s）拍摄一次RGB-D图像，各个时刻有一个对应的相机坐标系，相机坐标系的原点一般是该时刻相机中心所在位置，所获得的图像以及点云都服从这个相机坐标系。SLAM基于这些连续拍摄的RGB-D数据可估计出相机在世界坐标系中的位置同时建立世界坐标系下相机所扫过的周边环境的点云图。在这里，世界坐标系可以理解为一个事先定义好的不随相机运动发生变化的坐标系，而相机坐标系会随着相机的运动发生变化。如果假设相机的起始位姿所对应的相机坐标系为世界坐标系的化，那么SLAM本质上是要解决如下问题：估计出相邻时刻相机坐标系之间的转换关系。如果能够估计出两个时刻的相机坐标之间的转换关系就可以得到相机在任意时刻在定义的世界坐标系中的位置同时也能够将相机所获取的点云的坐标都转化到这个世界坐标系下。

1. **非线形最小二乘法**

两个时刻的相机坐标之间的转换关系是一个未知数，我们可以为这个转换关系未知数根据两个时刻所采集的图像以及点云构建一个误差函数，该函数的值随转换关系未知数的赋值变化。当未知数被赋予正确的转换关系时误差函数值趋近于0。一般采用最小二乘法来估计相机坐标系之间的转换关系。最小二乘法是通过最小化误差函数来对转换关系进行估计。如果用来表示误差函数，最小法的数学表达式子如下：

如果是非线性的，需要采用采用数值计算方法来求解。针对SLAM问题，应该是相邻两张图像所对应的相机坐标系之间的转换关系，误差函数应该是通过进行坐标转换后所观察到的误差，例如上文介绍的灰度误差就是一种观察到的误差。通过最小乘法可以得到最优的使得的值达到最小。

SLAM问题的误差函数一般是非线性的，需要采用数值计算的方法来求解最小二乘问题。我们先来介绍用Gauss–Newton法求解最小二乘法。主要思想是通过对实施一阶泰勒展开对进行近似：

之中表示的一个增量，是雅可比矩阵。Gauss–Newton法主要通过计算雅可比以及增量来实施迭代，其中算法的流程如下所示：

Step 0：给定初值；

Step 1：对第次迭代计算当前的雅可比矩阵以及误差函数值；

Step 2：求解增量方程：，其中，，

Step 3：如果足够小则停止迭代，否则令，返回第2步。

Gauss–Newton法避免了海塞矩阵的计算，但是任然面临的逆矩阵的计算问题。是半正定的，不一定存在逆矩阵。为了克服上述问题，还可以采用LM等方法来求解。

需要注意的是如果直接运用Gauss–Newton法等数值计算方法来优化会存在一个问题。数值计算方法是通过更新公式来实现迭代计算的。这个更新过程默认是对加法封闭的。通俗来说，加上一个增量之后虽然数值发生了变化但其所表达的物理含义应该没有变化。 但是在SLAM问题中，需要优化的变量即转换矩阵对加法并不封闭。转换矩阵包含了和，其中如果乘以一个增量任然是一个旋转矩阵，但是如果加上一个增量以后就不再是一个旋转矩阵了。为了能应用数值计算方法来优化转换矩阵，需要将转换矩阵变换为另外一种形式，使得对加法是封闭的。

1. **引入李群表达**

为了解决上述问题，需要引入李群表达。旋转矩阵是一种特殊的矩阵，具有如下的性质：

如果对时间求导的话可以得到：

以及

由此可以看出是一个反对称矩阵。对于一个反对称矩阵可以找到一个向量，使得：

其中是一个操作符，3X1向量加上这个操作符之后就转换成了一个3X3矩阵。通过进一步的推导可以得到：

如果将一个空间中一个点的三维坐标进行坐标变换的话，除了需要知道旋转矩阵以外还需要知道空间位移。变换坐标时将用到由和组成的4X4的变换矩阵：

我们进一步引入一个向量(引入李代数)，这个向量大小为6X1，具体的形式如下：

定义的操作如下：

可以得到

由此可见通过引入李代数可以用来表示变换矩阵。在上式中是一个雅可比矩阵：

其中：是三维向量的模长，是表示方向的单位向量，即。此外，可以通过以下公式计算：

1. **雅可比的计算方法**

设是当前图像上的一个像素点，这个像素点是三维空间中的一个点在图像上的一个投影。这张图像有一个对应的相机位姿，这个位姿又对应了一个以相机为原点的相机坐标系。在这张图像对应的相机坐标系下，三维空间中的点的坐标为。在当前图像的下一帧图像中这个三维点也被投影到了图像中的一个像素上。下一帧图像也对应了一个相机坐标系，由于相机的运动这个相机坐标系和当前图像对应的相机坐标系并不相同，如果我们能够知道两个相机坐标系之间的转换关系就可以把三维空间点的坐标转换为下一张图像对应的相机坐标系下的坐标。可以进一步通过相机模型将坐标转化为下一张图像上的像素点坐标。如果使用的转换关系是准确的话，可以得到：

其中和分别指在所给像素坐标出图像的灰度值。我们将采用一系列算法得到两个坐标系之间的转换关系是和。进一步地，我们已经介绍了可以用向量来代替和，因此可以将问题转化为求解。可以看出是的一个函数，随着变化也将变化。变化之后首先改变了的值，进一步改变了的值，最后改变了的值。在求解的过程就是寻找一个最优的，使得和之间的差最小

两个灰度之间的差也是的一个函数。

一般采用最小二乘法为了对进行优化。最小二乘法的迭代过程中一个重要步骤是计算上述误差函数的雅可比:

将雅可比根据链式法则展开后可以得到：

分别记：

首先来看，这一项其实就是根据当前的将三维空间点投影到下一帧图像中的像素处后，这个像素点和相邻像素点之间的灰阶的差，由于有横向和纵向两个方向，因此是一个1X2的矩阵，分别对应了像素处和相邻横向像素的灰阶差以及像素处和相邻纵向像素的灰阶差。

第二项可通过映射像素位置和三维空间点在下一帧图像的相机坐标系中的坐标之间的关系求得。和之间的关系如下：

其中，，，和为相机参数，可以通过相机标定得到。通过对求导可以得到如下的2X3矩阵：

第三项可以采用右扰动的方法来推导，通过在的右侧增加一个扰动，可以将转换为：

其中最终可以得到：

1. **直接法概要**

采用来表征坐标系之间的转换关系，SLAM问题由求解转换矩阵变为了求解，得到之后转换矩阵可通过公式获得。针对像素直接法考虑了两种误差，误差函数分别是：

表示空间三维坐标的轴即景深。设是像素的误差向量，优化的最小二乘的形式如下：

其中：

上述公式建立在假设服从t分布的基础之上。和都是t分布的参数。

如使用Gauss–Newton法对进行优化，那么关键一步在于建立的增量方程：

通过推导可以得到：

其中是用像素的误差向量对求导得到的雅可比矩阵，由于误差向量包含了两项误差，有六个元素，因此是一个2X6的矩阵。

其中的第一行是针对的导数。根据上面的推导可以得到的第一行等于：

其中和分别表示在相邻的下一帧图像中点的投影处和横纵方向上相邻像素点之间的灰阶的差。

的第二行是针对的导数：

将上式右侧展开得到：

根据前面介绍的推导结果可以得到：

其中和分别表示在相邻的下一帧图像中点的投影处和横纵方向上相邻像素点之间的景深的差。

1. **直接法使用过程**