# SLAM知识点汇总

孔大庆 2020-2-5

## 1.相机模型

### 1.1理想透射模型——透镜成像与小孔成像

在研究像素点与实际物理环境映射关系之前，需要了解成像原理。实际数码相机拍摄的过程是一个光学成像的过程，具体来说是一个透镜成像的非线性过程，透镜成像原理如下：

v

f

u

P

图1-1 透镜成像示意图

相机的镜头是一组透镜用图中的椭圆表示，感光元件构成的成像平面P用图中平行四边形表示（应该是矩形，视角调整为平行四边形）。横轴虚线表示光轴，平行于光轴的光线汇聚到竖轴虚线上为透镜的焦点，焦点到透镜中心的距离叫焦距用f表示，实际物体到透镜中心的距离叫物距用u表示，而透射所成的像到透镜中心的距离叫相距离用v表示，三者之间满足：

(1-1)

实际成像平面P就在焦点附近，也即是焦距f可以近似为相距v，将透镜中心当作小孔，可将上述透镜模型当作小孔成像模型，这也是实际相机应用最多的模型。

### 1.2坐标系定义

在研究图片像素点位置到真实物理点位置的映射关系时，需要在不同坐标系对其进行表达，这里涉及到四个坐标系：世界坐标系、相机坐标系、图像坐标系和像素坐标系。下图是物理世界一点经过小孔成像原理到像素平面的坐标表示：

Zc(Z)

像素坐标系

图像坐标系

相机坐标系

m

x

Y

xw

Yw

M

Yc

xc

Zw

P

世界坐标系

图1-2 成像过程中的坐标系

上述图中点M为三维空间点，m为M在图像平面的投影。图中展示的坐标系有：

世界坐标系：是客观三维世界的绝对坐标系，也称客观坐标系。因为数码相机安放在三维空间中，我们需要世界坐标系这个基准坐标系来描述数码相机的位置，并且用它来描述安放在此三维环境中的其它任何物体的位置，用（Xw, Yw, Zw）表示其坐标值。

相机坐标系（光心坐标系）：以相机的光心为坐标原点，X 轴和Y 轴分别平行于图像坐标系的 X 轴和Y 轴，相机的光轴为Z 轴，用（Xc, Yc, Zc）表示其坐标值。

图像坐标系：以CCD 图像平面的中心为坐标原点，X轴和Y 轴分别平行于图像平面的两条垂直边，用( x , y )表示其坐标值。图像坐标系是用物理单位（例如毫米）表示像素在图像中的位置。

像素坐标系：以 CCD 图像平面的左上角顶点为原点，X 轴和Y 轴分别平行于图像坐标系的 X 轴和Y 轴，用(u , v )表示其坐标值。数码相机采集的图像首先是形成标准电信号的形式，然后再通过模数转换变换为数字图像。每幅图像的存储形式是M × N的数组，M 行 N 列的图像中的每一个元素的数值代表的是图像点的灰度。这样的每个元素叫像素，像素坐标系就是以像素为单位的图像坐标系。

### 1.3坐标变换

#### 1.3.1 图像坐标系和像素坐标系

像素坐标系和图像坐标系共处一个平面，上述像素坐标系的定义可用下图展示二者关系：

v

u

O’

x

y

O

图1-3 图像坐标系与像素坐标系关系图

那么对于图像坐标系上的一点m(x,y)，其对应的像素坐标可表示为：

(1-2)

采用齐次坐标方式表达如下：

(1-3)

其中是图像坐标原点在像素坐标系下的坐标，dx和dy是单个像素对应在图像坐标系x轴和y轴下的尺寸。根据公式(1-3)可得由已知像素坐标转化为图像坐标的公式为：

(1-4)

#### 1.3.2相机坐标系和图像坐标系以及像素坐标系

由图1-2点M在相机坐标系坐标定义为(,)，而其在图像坐标系对应的相点m(x,y,1)。很容易可以利用相似三角形得到二者之间的关系如下：

(1-5)

则有

(1-6)

根据公式(1-3)可得相机坐标转化为像素坐标的公式为：

(1-7)

令K=，可以发现K仅与相机本上有关，我们称其为内参矩阵，相机内参标定即是标定此参数。同时，我们实际使用中会忽略图像坐标，之间研究相机坐标系和像素坐标系关系。通常我们已知像素坐标，通过下式可以转化为归一化相机坐标：

(1-8)

因为像素坐标是2D坐标不包含3D信息，所以单个像素坐标无法推导除实际点的深度，但是可以通过上式得到归一化相机坐标系的坐标值。

#### **1.3.3世界坐标系和相机坐标系**

世界坐标系可以是客观坐标系，相对不变的，相机运动需要在世界坐标系中表达，用表示相机坐标系原点在世界坐标系下坐标，用表示从相机坐标系到世界坐标系旋转矩阵。如果点M在世界坐标系坐标为(,)，而在相机坐标系下坐标为(,)，那么二者转化关系如下：

+ (1-9)

用其次坐标系表示：

(1-10)

同样可得世界坐标系坐标转化到相机坐标系公式：

(1-11)

其中和描述相机相对于实际坐标系的位姿称为相机外参数，是slam中需要求解的量。

### 1.4相机模型

## 投影几何

### 2.1 ICP算法原理

ICP（Iterative Closest Point）常用于3D-3D点云配准，这里介绍如何通过ICP算法恢复两组3D点云之间的运动信息R，t。

将求运动R，t问题转换为如下最小二乘问题：

(2-1)

这里和为匹配点（实际上他们有可能不是正确的匹配点），这样一个最小二乘问题可以进一步分析。取两组点云的质心（就是点云坐标均值）、，令，。（2-1）可变形为：

因为

所以上式可化简为：

（2-2）

为了取得最小值，可令，得，所以只要求出R，t直接可得。

对（2-2）公式中关于R项进一步展开可得：

=

=

=)

=

=

因为均是正交矩阵，所以 也是正交矩阵，且满足=1，上述公式可转化为：

=

因为奇异值、、，、、1，所以当、、均为1时，上式取得最大值。最终可得 ，因此有：

(2-3)

### 2.2对极约束

投影几何中很重要的一个定理就是对极约束，道理很简单，就是对于同一个3D点它在不同相机下所成的像存在一个约束——匹配点x在令一个图像映射为一条极线，而其对应的匹配点必在这条极线上。如下图所示：

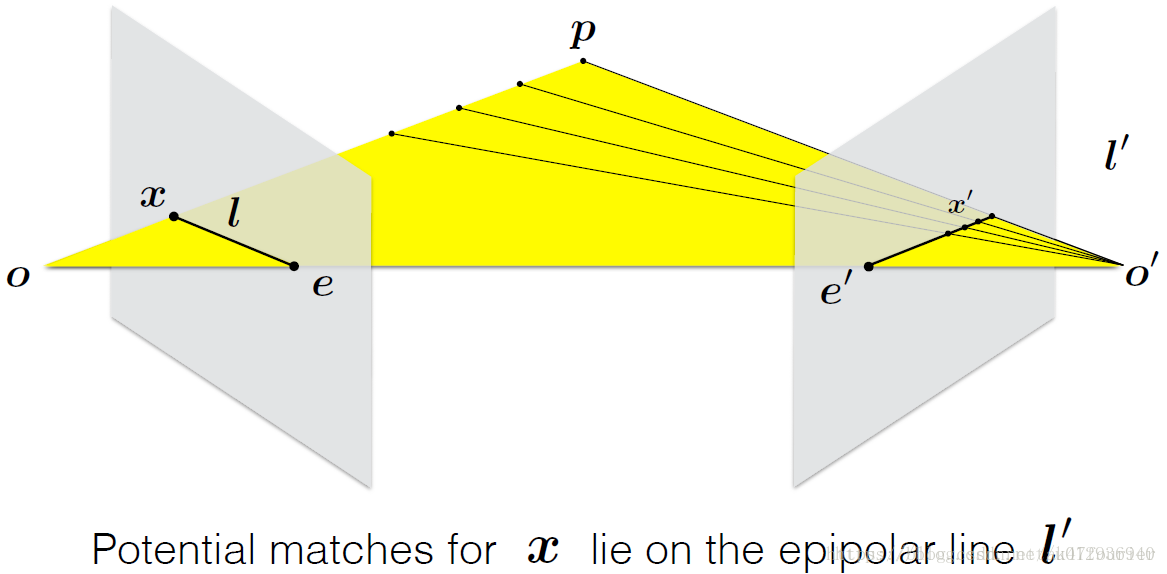


图2-1 对极约束

可以用基础矩阵来表示极线，即

(2-4)

因为在该极线上，所以有如下公式：

(2-5)

#### 2.2.1 基础矩阵和本质矩阵

推导式（2-5）中矩阵E：

图2-1中点和是同一点p在两帧图像归一化坐标系下的投影，u和是其对应的像素点。根据公式（1-8）有：

(2-6)

设s和分别对应匹配点的深度信息，则s和分别是点p在两帧图像坐标系下的坐标。假设，所在帧相对所在帧的位姿关系为R、t，那么有

(2-7)

等式两边同时与t叉乘，得

(2-8)

然后两边同时左乘，得

(2-9)

在将(2-6)带入上式，可得：

（2-10）

上式中有被称为本质矩阵，被称为基础矩阵，很显然有：

(2-11)

#### 2.2.2 八点法求基础矩阵

根据公式（2-10），且*F*矩阵是3x3矩阵，也即是有9个元素，因此理论上构造出9个方程即可求得*F*，实际上*F*具有尺度等价性——F乘以一个非零的常数，对极约束仍然成立，因此实际需要8个方程即可，事实上常规处理是令。

对于一对匹配点和，将其带入公式，得：

所以只需要8对匹配点即可求得*F。*

#### 2.2.3基础矩阵实际求解过程

实际使用8个点对求得基础矩阵往往是不准确的，因为实际情况匹配点往往会有噪声、数值舍入误差和误匹配等情况，仅仅用上面的方程求得结果往往是不稳定的，因此在8点法基础上做必要的改进。这部分可以直接参照orb\_slam2中求解基础矩阵的代码，具体求解基础矩阵步骤如下：

**1. 图像坐标归一化**

在求解Af = 0过程中，系数矩阵可能会因为像素坐标的各个分量数量级相差太大而造成不稳定，图像坐标归一化的目的是将所有的匹配点坐标都进行标准化

1. 对所有点进行平移，使其原点在这些点的形心上；
2. 对所有点进行伸缩，使得所有点到中心的平均距离为。

可以用归一化矩阵来描述上述归一化过程：

(2-12)

其中

经过上述归一化处理，可以得到两个归一化矩阵和，利用8点法可以计算出归一化坐标系下的基础矩阵，然后将其转换到正常坐标系得

(2-13)

**2. 最小二乘求解**

归一化坐标可以一定程度上降低噪音和误匹配得影响，但是还是不够。实际可以匹配的点数远远多余8对情况下，可以用最小二乘方法求得最优解。实际求解过程中只要对系数矩阵A进行SVD分解，找最小特征值对应的特征向量即可（参见附录ii）。

1. **RANSAC迭代求最优**

RANSAC随机采样一致性算法是常用的随机优化模型的方法，网上有很多关于该算法原理介绍，这里就不再赘述。采用RANSAC迭代寻找满足大多数匹配点对极约束的基础矩阵，这样能提高基础矩阵求解的准确度。

综上采用这些方法，基本上可以获得比较好的基础矩阵。这里可以参照orb-slam2初始化中计算基础矩阵的部分。需要注意的是求解出的t是标准化的尺度，无法得到真实的尺度信息，因为如果同时扩大t的尺度和3D点的深度，对极约束依然成立，所以单目的缺点就在于无法知道实际环境的尺度，后续恢复的位置信息和路标点都是建立在第一帧的尺度基础上的。

另外，在使用8点法求基础矩阵时会遇到蜕化情况，比如：所有特征点比较集中，特征点几乎在一个平面上，以及两帧之间只有纯旋转。如果是特征点是同一平面或近似同一平面的点，可以通过求解单应矩阵来恢复R，t。

#### 2.2.4本质矩阵恢复运行信息R、t

根据对极约束使用八点法求基础矩阵，根据(2-11)求本质矩阵或者使用五点法直接求本质矩阵，之后就是从本质矩阵中恢复运动信息R、t。首先，本质矩阵自由度为5，包括三个旋转自由度和三个平移自由度，但是因为存在整体尺度因子多义性（此处为一直没有理解），所以为5。另外研究E的性质可以发现，对于一个3x3矩阵为本质矩阵的充要条件是其奇异值有两个相等一个为零，即

(2-14)

原因如下：

，S为反对称矩阵，对于反对称矩阵可以写称如下形式：

(2-15)

在相差一个符号的意义下，diag(1,1,0)W，其中

(2-16)

可将E进行变形为：

(2-17)

最终得：

(2-18)

而对于不同符号的*Z*，旋转矩阵也可以为，所以*R*有两种情况。而对于平移t，因为*S*是t的反对称矩阵，则St=0，则t对应其奇异值分解最后一个特征值向量，即

(2-19)

#### 2.2.5单应矩阵

### 2.3 三角测量恢复3D点

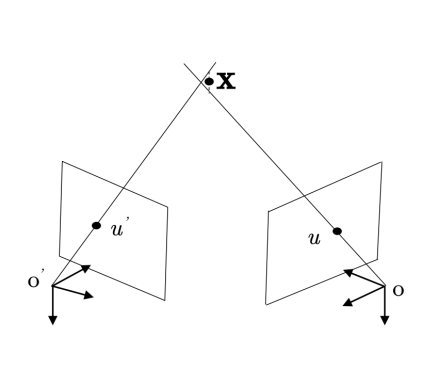
基于特征点法的视觉SLAM算法流程多是用：

1. 基础矩阵或单应矩阵初始化前两帧位姿，同时用三角化方法恢复出3D路标点；
2. 图像输入提取特征点与3D路标点进行匹配；
3. 匹配成功用Pnp3D-2D的算法求解位姿；
4. 再用三角化方法恢复一部分路标点，如此反复。

这其中无法避免的是三角化路标点，三角化算法比较简单，单目和双目有所区别，下面进行介绍。

#### 2.3.1 单目三角化

如下图所示，两相机不同位置看到3D点x，且已知摄像机位姿、。求3D点x的位置。



设和u为匹配的像素点，根据像素坐标系和世界坐标系关系，可得：

( 2-15)

取第一个式子，按照DLT算法，可得：

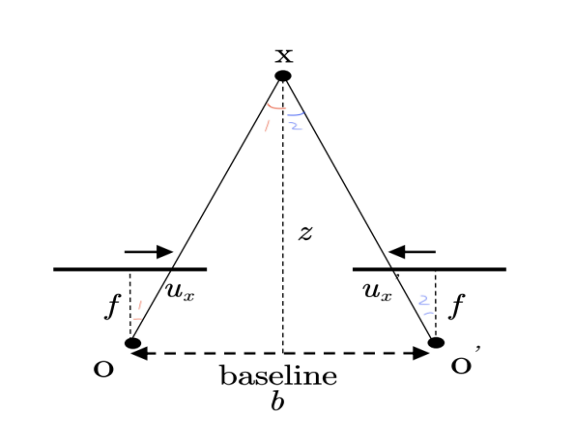
(2-16)

观察上式，第三式 = 第一式子\*（-）+ 第二式子\*（-），因此第三式子没有意义。这样将公式（2-15）同样进行展开可得类似（2-16）的结果，并将二者连立，得：

(2-17)

上述式子三个未知数，四个方程，可采用最小二乘算法，直接对其进行SVD分解，则最小特征值对应的特征向量即是x的解。

#### 2.3.2 双目三角化

对于双目匹配好的特征点，且图像进行了极线矫正，三角化就相当简单了。具体来说此时3D点坐标仅仅和它在双目中的投影视差以及双目的基线有关。

经过极线矫正后，匹配的特征点在同一行，设它们在像素坐标系的横坐标分别为和，双目基线用b表示，f为焦距。很显然有下式成立：

(2-18)

将式中第二和第三行带入第一行，得

(2-19)

其中d称为视差，知道z值后，通过公式（1-8），得

(2-20)

### 2.4 3D-2D Pnp算法

## 三维空间运动

### 3.1坐标系和向量

对于三维空间中的刚体，其运动有六个自由度：三自由度平移运动和三自由度旋转运动，统称为物体的位姿（pose）。设定参考坐标系，则从该坐标系原点指向运动物体的向量可以用来表示该运动物体的位置信息，而运动物体本身固连的坐标系相对于参考系的旋转关系则表达了其姿态信息。

#### 3.1.1向量

设向量在坐标系，即有

(3-1)

通过上式，可得：

(3-2)

#### 3.1.2向量内积

设向量和在坐标系下坐标分别为和，其内积公式如下：

(3-3)

#### 3.1.3向量外积

上述向量和的外积公式如下：

(3-4)

### 3.2旋转运动参数化方式

对于共原点的两个参考坐标系，和，向量其在参考系下坐标为，在参考系下坐标为。那么有：

(3-5)

**1）向量和向量坐标**

该式似乎不好理解，和是不同坐标系下的坐标，这个坐标乘各自的基向量怎么就相等了呢，而对于笛卡尔坐标系基向量基本上是x:、y:、z:，那么上述等式就更不可能相等了。误解在于没有搞清楚和的含义，二者都是坐标系的基向量（注意是向量），表示坐标系的方位，显然不同坐标系的基向量指向是不完全一致的，所以就不能参数化为上述的x，y，z。

该式的意思是不论这个向量在哪个坐标系，它其实是同一个向量，是不会因为定义的坐标系不同而出现这个向量的方向和大小发生变化。问题关键就在于如何理解将这个向量在不同坐标系下参数化，还存在（3-5）的等式呢。其实很简单，搞懂向量坐标的含义就明白了。

向量坐标表面上看表示的是向量在坐标系下的位置，实际上是各个基向量的线性组合，因为不同坐标系的基向量指向不完全相同，所以坐标也不会相同（比如对于同一个向量，在不同坐标系的坐标分别为，你可能就会猜到这两个坐标系只是z轴方向相反而已）。所以你回看公式（3-1）中，这里面的坐标是标量，表征基向量的大小而已，向量还是那个向量，它不再是具体的数字和坐标。因此，上述公式可以改为：

（3-6）

所以一定要把向量和向量坐标区分清楚。只有理解了这一点才能理解清楚坐标系旋转。

上面是从向量的层面理解的，实际上也可以从坐标的层面理解。坐标是向量在具体坐标系下的参数化表达，我们也可以将坐标系的基向量参数化，这样可以得到数值上的等式。比如向量在坐标系下坐标为，在坐标系坐标系下坐标是，显然坐标系是由坐标系绕z轴旋转180度得到的，设为参考坐标系，其基向量为

x2:、y2:、z2:，则基向量在参考坐标系下表示为

x1:、y1:、z1:，带入公式（3-5），则得。

1. **坐标系旋转**

在1）中我们从向量和坐标两个角度解释了公式（3-5），从坐标表示的方法上看，我们可以直接建立同一个向量在不同坐标系下存在的坐标关系等式，前提是其中一个坐标系基向量要在另一个坐标系中进行参数化表达。那么也就是说只要我们掌握了这个参数化表达方法，我们可以轻易地将这个向量坐标从一个坐标系转换到另一个坐标系。

实际上将一个坐标系的基向量放到参考坐标系中参数化，其实相当于知道了这个坐标系到参考系的转换关系。坐标系间的旋转关系可以用3x3的旋转矩阵来表示称方向余弦矩阵，还可以仅用一个旋转轴和旋转度数表示称为角轴，也可以使用欧拉角来表示旋转，还可以使用四元数表示。下面将一一进行介绍。

#### 3.2.1方向余弦矩阵

对于公式(3-5)，左右同时左乘，有：

(3-6)

因为坐标系坐标轴都是单位长度的向量，所以上式各个元素向量内积为基向量间的余弦值，因此称为方向余弦矩阵。而坐标系变化如下：

这说明坐标系右乘方向余弦矩阵即可得坐标系，所以坐标系间转换是通过右乘旋转矩阵将一个坐标系变成另一个坐标系，而坐标变换恰好是左乘旋转矩阵。所以导航解算的时候，更新旋转矩阵时，都时右乘旋转变化量得到当前的旋转位姿，而当我们需要知道惯性系下的速度和加速度时，又需要将这些机体系下的向量坐标左乘这个旋转矩阵得到惯性系下的向量坐标。也即是：

#### 3.2.2角轴

旋转运动可以描述为绕一个旋转轴旋转一定角度，可以用一个三维单位矢量来表示旋转轴，用表示旋转的角度，称这个向量为角轴。上述方向余弦矩阵描述的是坐标系相对于坐标系的旋转关系，同样可以用角轴来表示。

角轴转换为方向余弦矩阵，可以使用罗德里格斯公式：

(3-7)

方向余弦矩阵R和角轴是李群和李代数的关系，后续介绍李群和李代数时候再证明罗德里格斯公式。

#### 3.2.3四元数

#### 3.2.4欧拉角

## 附录

### i. DLT（直接线性变换）原理

凡是满足如下关系的数学模型，即可通过DLT算法求解矩阵M：

其中和是空间上向量，是尺度值，为待求的变换矩阵。由上式可以变换为下面公式：

另，最终可得：

0

固为A零空间向量，当A矩阵行数大于M的维度可以使用SVD分解，求最小二乘解为A最小特征值对应的右奇异矩阵。

### ii. 齐次线性方程**=0**的近似解为矩阵奇异值分解最小特征值对应的特征向量

设定是超定方程，且方程无解。为此将其改为如下最小二乘问题：

对A做SVD分解有：

带入上式，得：

很显然，当=时（最小特征值对应特征向量），上式可取最小值。