从零开始手写 VIO

第一讲 概述与课程介绍

贺一家、高翔、崔华坤

2019 年 10 月 10 日

目录

Section 1

课程内容与提要

课程与内容提要

- 本课程专注于视觉 +IMU 融合定位的基础理论和实现，作为视觉 SLAM 的进阶课程
- 在本课程中，你将学习到的重点内容有：
  - IMU 的工作原理和噪声方程
  - 视觉与 IMU **紧耦合**的基础理论
  - 从零开始实现 VIO 紧耦合优化器（**仅基于 Eigen**）

相比视觉 SLAM 基础课程，本次课程更加注重 **IMU 融合的理论推导与实现**。

课程与内容提要

讲师介绍：

贺一家，中科院自动化所博士，“白巧克力亦唯心”博主，研究方向包括视觉 SLAM、多传感器信息融合。CSDN 博客专家，发表 SCI 和 EI 论文数篇。

课程与内容提要

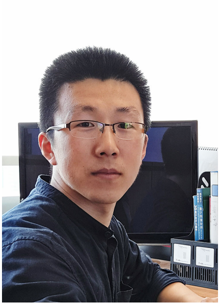
讲师介绍：

高翔，清华大学博士，慕尼黑工业大学博士后，《视觉 SLAM 十四讲》、《机器人学中的状态估计》作者、译者。曾在 RAS、IROS 等期刊会议发表论文。

课程与内容提要

讲师介绍：

崔华坤，北京工业大学物理学专业毕业，北京臻迪科技 SLAM 算法工程师，研究方向为基于多传感器融合的无人机定位导航。



课程与内容提要

- 课程时间安排：公式推导为主，代码为辅
  - 课程提要，约 1 小时
  - IMU 原理与模型，约 2.5 小时
  - 后端优化与 VIO BA, 约 2.5 小时
  - 滑动窗口算法与 FEJ, 约 2.5 小时
  - 后端求解系统与代码，约 2.5 小时
  - 前端介绍，约 1.5 小时
  - VIO 初始化，约 1.5 小时
  - VIO 系统回顾与展望，约 1 小时
- 课后作业时间安排：以编程为主，公式推导为辅，每周约 8 小时。

课程与内容提要

预备知识

- 数学：线性代数、微积分、概率论
- 编程：Linux、C++、OpenCV
- 英语：文献阅读
- 专业知识：2D/3D 计算机视觉、图像处理

Section 2

VIO 概述

VIO 概述

深蓝学院 shenlanxueyuan.com

VIO：（Visual-Inertial Odometry）

以视觉与 IMU 融合实现里程计

IMU（Inertial Measurement Unit），惯性测量单元

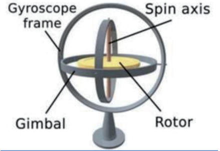
- 典型 6 轴 IMU 以较高频率（$\geq 100\text{Hz}$）返回被测量物体的角速度与加速度
- 受自身温度、零偏、振动等因素干扰，积分得到的平移和旋转容易漂移

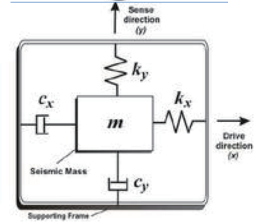
视觉 Visual Odometry

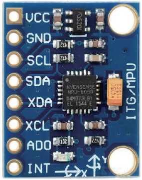
- 以图像形式记录数据，频率较低（$15 - 60\text{Hz}$ 居多）
- 通过图像特征点或像素推断相机运动

VIO 概述

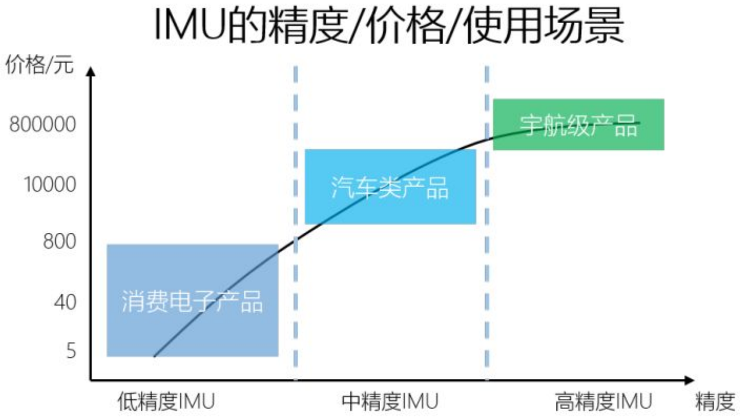
六自由度 IMU 本身由一个陀螺仪和一个加速度计组成，分别测量自身的角速度和加速度。







VIO 概述



手机等电子产品多使用价格低廉的 MEMS IMU（如 MPU 6050），自动驾驶类则多使用几万元的 IMU（如 Apollo 中使用的 Novatel SPAN-IGM-A1）[1]。

[1]图片来自https://zhuanlan.zhihu.com/p/32693377

VIO 概述

IMU 与视觉定位方案优势与劣势对比：

方案	IMU	视觉
优势	快速响应 不受成像质量影响 角速度普遍比较准确 可估计绝对尺度	不产生漂移 直接测量旋转与平移
劣势	存在零偏 低精度 IMU 积分位姿发散 高精度价格昂贵	受图像遮挡、运动物体干扰 单目视觉无法测量尺度 单目纯旋转运动无法估计 快速运动时易丢失

VIO 概述

整体上，视觉和 IMU 定位方案存在一定互补性质：

- IMU 适合计算短时间、快速的运动；
- 视觉适合计算长时间、慢速的运动。

同时，可利用视觉定位信息来估计 IMU 的零偏，减少 IMU 由零偏导致的发散和累积误差；反之，IMU 可以为视觉提供快速运动时的定位。

VIO 概述

深蓝学院 shenlanxueyuan.com

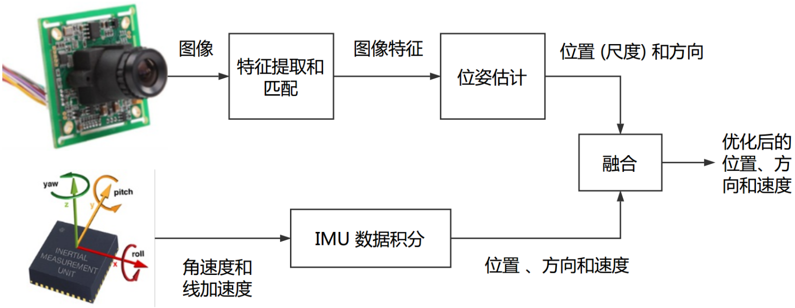
IMU 数据可与多种定位方案融合

- 自动驾驶中通常用 IMU+GPS/差分 GPS/RTK 的融合定位方案，形成 GNSS-INS 组合导航系统，达到厘米组定位精度；
- 头戴式 AR/VR 头盔则多使用视觉 +IMU 的 VIO 定位系统，形成高帧率定位方案。

融合方案

松耦合

将 IMU 定位与视觉/GNSS 的位姿直接进行融合，融合过程对二者本身不产生影响，作为后处理方式输出。典型方案为卡尔曼滤波器。

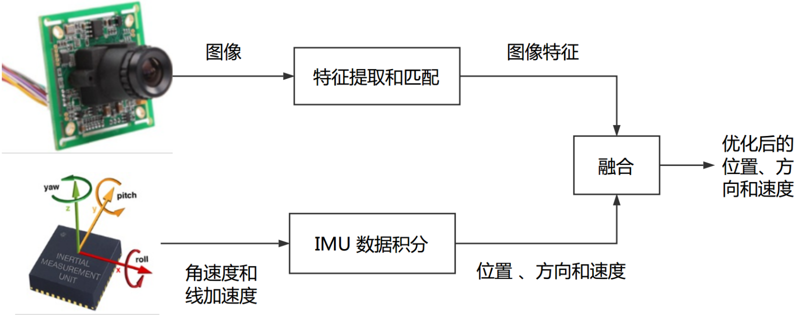


融合方案

深蓝学院 shenlanxueyuan.com

紧耦合

融合过程本身会影响视觉和 IMU 中的参数（如 IMU 的零偏和视觉的尺度）。典型方案为 MSCKF 和非线性优化。



VIO 概述

为什么要使用紧耦合

- 单纯凭（单目）视觉或 IMU 都不具备估计 Pose 的能力：视觉存在尺度不确定性、IMU 存在零偏导致漂移；
- 松耦合中，视觉内部 BA 没有 IMU 的信息，在整体层面来看不是最优的。
- 紧耦合可以一次性建模所有的运动和测量信息，更容易达到最优。

VIO 概述

本课程要探讨的问题

- IMU 的测量数据表达了系统的什么状态，受哪些噪声影响？
- 如何建立一个带有 IMU 测量信息和视觉特征点信息的非线性优化问题并进行求解？
- 该问题随着时间将发生怎样的演变？

Section 3

预备知识回顾

预备知识回顾

数学符号约定:

- 普通变量: a, b, c
- 矩阵和向量: $\mathbf{A}, \mathbf{B}, \mathbf{v}$
- 集合: $\mathbb{R}, \mathbb{Z}$
- 特殊集合: $\mathcal{F}, \mathcal{G}$
- 希腊字母和向量: α, $\boldsymbol{\alpha}$
- 李代数: $\mathfrak{so}(3)$,$\mathfrak{se}(3)$

预备知识回顾

三维刚体运动:

- 我们通常在机器人/车辆上定义各种坐标系，如:
  - 世界坐标系 W;
  - IMU 坐标系 I;
  - 相机坐标系 C;
- 坐标系之间变换关系由一个 SE(3) 给出。如 I 到 W 系的变换矩阵为: $\mathbf{T}_{WI}$:

$$\mathbf{T}_{WI} = \begin{bmatrix} \mathbf{R}_{WI} & \mathbf{t}_{WI} \\ \mathbf{0}^\top & 1 \end{bmatrix} \in \mathbb{R}^{4\times 4}. \tag{1}$$

- $\mathbf{R}_{WI}$ 为 3×3 的旋转矩阵，$\mathbf{t}_{WI}$ 为平移向量。
- $\mathbf{T}_{WI}$ 右乘一个 I 系下（齐次）坐标，将得到该点 **W** 系下坐标。

预备知识回顾

约定：

- 当某个量表达坐标系的转换关系时，写在右下脚标，例如 $\mathbf{T}_{WB}$。
- 当表达矢量在某坐标系中取的坐标时，写在右上角标，如 $\mathbf{v}^w$ 表达速度矢量在 World 系坐标。
- $\mathbf{I}$ 系也称为 Body 系。
- 定义明确时，有时会省略该脚标，我们会直接谈论 $\mathbf{R}, \mathbf{t}$ 这样的量。
- 不刻意区分齐次和非齐次坐标，因为在程序中可以自动完成转换，且无歧义。
- 默认以 $\mathbf{T}_{WI}$ 表达并存储 IMU 的定位信息，而不是 $\mathbf{T}_{IW}$。二者实际互为逆，存储哪一类区别不大，视习惯而定。
- 同理，$\mathbf{T}_{WI}$ 的平移部分可直接视作 IMU 在世界中的坐标，从而进行绘图或可视化操作。

预备知识回顾

四元数：

- 旋转矩阵 $\mathbf{R}$ 亦可用四元数 $\mathbf{q}$ 描述。
- 四元数 $\mathbf{q}$ 有一个实部和三个虚部。我们把实部写在前：

$$\mathbf{q} = [q_0, q_1, q_2, q_3]^\top \text{ 或 } \mathbf{q} = [w, x, y, z]^\top \tag{2}$$

其中 q_0 为实部，$[q_1, q_2, q_3]^\top$ 为虚部。因为实部为标量，虚部为矢量，所以也可记为：

$$\mathbf{q} = [s, \mathbf{v}]^\top. \tag{3}$$

其中 s 为标量，$\mathbf{v}$ 为虚部的矢量。

预备知识回顾

四元数的乘法记了好多次都没记下来。
这次用大白话试试：
两个四元数1和2相乘，得到一个新的四元数
四元数的实部是12的实部相乘，减去12的虚部点乘
四元数的虚部是一个三维向量，它由三个三维向量相加得到：
1的实部乘以2的虚部
2的实部乘以1的虚部
1的虚部叉乘2的虚部

四元数之间可以进行乘法

$$\begin{aligned}\mathbf{q}_a \otimes \mathbf{q}_b = &\, w_a w_b - x_a x_b - y_a y_b - z_a z_b \\ &+ \left(w_a x_b + x_a w_b + y_a z_b - z_a y_b\right) i \\ &+ \left(w_a y_b - x_a z_b + y_a w_b + z_a x_b\right) j \\ &+ \left(w_a z_b + x_a y_b - y_a x_b + z_a w_b\right) k.\end{aligned} \tag{4}$$

或：

$$\mathbf{q}_a \otimes \mathbf{q}_b = \left[s_a s_b - \mathbf{v}_a^\top \mathbf{v}_b, s_a \mathbf{v}_b + s_b \mathbf{v}_a + \mathbf{v}_a \times \mathbf{v}_b \right]^\top. \tag{5}$$

此外，四元数可类似复数，定义加减、模长、逆、共轭等运算，不一一展开。

预备知识回顾

深蓝学院 shenlanxueyuan.com

单位四元数可表达任意三维旋转，且无奇异性。

四元数和角轴的转换关系：
假设某个旋转运动的旋转轴为单位向量 $\mathbf{u}$，绕该轴旋转的角度为 θ，那么它对应的单位四元数为：

角轴到四元数：
实部：角度的一半的余弦
虚部：角度的一半的正弦，乘以单位向量

$$\mathbf{q} = \begin{bmatrix} \cos\frac{\theta}{2} \\ \mathbf{u}\sin\frac{\theta}{2} \end{bmatrix} \tag{6}$$

当旋转一段微小时间，即角度趋于 0 时，容易有：

小量旋转的四元数近似：
实部为1
虚部为角度除以2

$$\Delta\mathbf{q} = \begin{bmatrix} \cos\frac{\delta\theta}{2} \\ \mathbf{u}\sin\frac{\delta\theta}{2} \end{bmatrix} \approx \begin{bmatrix} 1 \\ \mathbf{u}\frac{\delta\theta}{2} \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{1}{2}\boldsymbol{\delta\theta} \end{bmatrix} \tag{7}$$

其中 $\delta\theta$ 的方向表示旋转轴，模长表示旋转角度。

预备知识回顾：对时间求导

深蓝学院 shenlanxueyuan.com

角速度:

$$\boldsymbol{\omega} = \lim_{\Delta t \to 0} \frac{\delta \boldsymbol{\theta}}{\Delta t} \tag{8}$$

四元数时间导数:

总结：
1.旋转矩阵的时间导数 = 角速度叉乘旋转矩阵
C· = -w^C
2.角轴（李代数）的时间导数 ：
w = J φ·
角速度等于左雅克比J乘以李代数的导数
3.四元数的时间导数 = 四元数右乘一个虚四元数：实部为0，虚部为角速度的一半；

$$\begin{aligned}
\dot{\mathbf{q}} &\triangleq \lim_{\Delta t \to 0} \frac{\mathbf{q}(t+\Delta t) - \mathbf{q}(t)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{\mathbf{q} \otimes \Delta \mathbf{q} - \mathbf{q}}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{\mathbf{q} \otimes \left(\begin{bmatrix} 1 \\ \frac{1}{2}\delta \boldsymbol{\theta} \end{bmatrix} - \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix} \right)}{\Delta t} \\
&= \mathbf{q} \otimes \begin{bmatrix} 0 \\ \frac{1}{2}\boldsymbol{\omega} \end{bmatrix}
\end{aligned} \tag{9}$$

预备知识回顾

除了利用四元数求导，亦可利用李代数进行旋转求导。
使用旋转矩阵 $\mathbf{R}$ 时，角速度为 $\boldsymbol{\omega}$，那么 $\mathbf{R}$ 相对于时间的导数可写作：

该式子在左乘习惯和右乘习惯下写法不同，可能对应的w坐标也不一样？

$$\dot{\mathbf{R}} = \mathbf{R}\boldsymbol{\omega}^{\wedge} \tag{10}$$

该式被称为**泊松公式**（Possion's equation），其中 $^{\wedge}$ 为反对称矩阵算子：

$$\boldsymbol{\omega}^{\wedge} = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \tag{11}$$

在本课程中亦记作：$\boldsymbol{\omega}_{\times}$ 或者 $[\boldsymbol{\omega}]_{\times}$，表达含义相同。

预备知识回顾

$\mathfrak{so}(3)$ 导数：
在优化带有旋转的函数时，通常计算一个增量 $\phi \in \mathfrak{so}(3)$，然后用它更新当前估计值：

$$\boxed{\mathbf{R} \leftarrow \mathbf{R} \exp(\phi^\wedge).} \tag{12}$$

其中 exp 为 $\mathfrak{so}(3)$ 至 $\mathrm{SO}(3)$ 上的指数映射。
本课程习惯为右乘，实际当中左右乘等价，仅为习惯上的差别。
注：(i) 不同的 $\mathbf{R}$ 函数，具体的导数形式也不同。(ii) 在程序中，不必区分 $\mathbf{R}$ 是以矩阵存储或是以四元数存储，只需按照该式更新即可。

预备知识回顾

常见的一些雅可比（以自变量为 $\mathbf{R}$ 举例）:
旋转点的左扰动雅可比:

$$
\begin{aligned}
\frac{\partial\left(\mathbf{Rp}\right)}{\partial\boldsymbol{\varphi}} &= \lim_{\boldsymbol{\varphi}\to\mathbf{0}} \frac{\exp\left(\boldsymbol{\varphi}^{\wedge}\right)\exp\left(\boldsymbol{\phi}^{\wedge}\right)\mathbf{p}-\exp\left(\boldsymbol{\phi}^{\wedge}\right)\mathbf{p}}{\boldsymbol{\varphi}} \\
&= \lim_{\boldsymbol{\varphi}\to\mathbf{0}} \frac{\left(\mathbf{I}+\boldsymbol{\varphi}^{\wedge}\right)\exp\left(\boldsymbol{\phi}^{\wedge}\right)\mathbf{p}-\exp\left(\boldsymbol{\phi}^{\wedge}\right)\mathbf{p}}{\boldsymbol{\varphi}} \\
&= \lim_{\boldsymbol{\varphi}\to\mathbf{0}} \frac{\boldsymbol{\varphi}^{\wedge}\mathbf{Rp}}{\boldsymbol{\varphi}} = \lim_{\boldsymbol{\varphi}\to\mathbf{0}} \frac{-(\mathbf{Rp})^{\wedge}\boldsymbol{\varphi}}{\boldsymbol{\varphi}} = \boxed{-(\mathbf{Rp})^{\wedge}.}
\end{aligned}
\tag{13}
$$

预备知识回顾

小结：
雅克比的定义是旋转后的点Rp，对旋转本身来求导：
左雅克比的结果是旋转后的点取反对称，即-(Rp)^
右雅克比的结果是旋转前的点取反对称，再左乘R，即-Rp^

旋转点的右扰动雅可比：

$$
\begin{aligned}
\frac{\partial(\mathbf{Rp})}{\partial\varphi} &= \lim_{\varphi\to\mathbf{0}} \frac{\mathbf{R}\exp(\varphi^\wedge)\mathbf{p}-\mathbf{Rp}}{\varphi} \\
&= \lim_{\varphi\to\mathbf{0}} \frac{\mathbf{R}(\mathbf{I}+\varphi^\wedge)\mathbf{p}-\mathbf{Rp}}{\varphi} \\
&= \lim_{\varphi\to\mathbf{0}} \frac{\mathbf{R}\varphi^\wedge\mathbf{p}}{\varphi} = \lim_{\varphi\to\mathbf{0}} \frac{-\mathbf{Rp}^\wedge\varphi}{\varphi} = \boxed{-\mathbf{Rp}^\wedge.}
\end{aligned} \tag{14}
$$

预备知识回顾

Q1:是什么？
两个旋转矩阵相乘，得到新的旋转矩阵，该矩阵对矩阵1和2的偏导数
Q2:怎么表达和推导？
在李代数空间进行求导。分子和分母都是

旋转连乘的雅可比

$$\begin{aligned}\frac{\mathrm{d}\ln\left(\mathbf{R}_1\mathbf{R}_2\right)^\vee}{\mathrm{d}\mathbf{R}_2} &= \lim_{\boldsymbol{\phi}\to\mathbf{0}}\frac{\ln\left(\mathbf{R}_1\mathbf{R}_2\exp\left(\boldsymbol{\phi}^\wedge\right)\right)^\vee-\ln\left(\mathbf{R}_1\mathbf{R}_2\right)^\vee}{\boldsymbol{\phi}} \\ &= \lim_{\boldsymbol{\phi}\to\mathbf{0}}\frac{\ln\left(\mathbf{R}_1\mathbf{R}_2\right)^\vee+\mathbf{J}_r^{-1}\boldsymbol{\phi}-\ln\left(\mathbf{R}_1\mathbf{R}_2\right)^\vee}{\boldsymbol{\phi}} \\ &= \mathbf{J}_r^{-1}(\ln(\mathbf{R}_1\mathbf{R}_2)^\vee)\end{aligned} \tag{15}$$

其中用到了对 $\forall\mathbf{R}$:

这是哪个常见的公式，要这么翻译：SO3上右乘一个小量，相当于so3上加上该小量对应的李代数，乘以右雅克比的逆（注意在哪个点）

$$\ln\left(\mathbf{R}\exp\left(\boldsymbol{\phi}^\wedge\right)\right)^\vee=\ln\left(\mathbf{R}\right)^\vee+\mathbf{J}_r^{-1}\boldsymbol{\phi} \tag{16}$$

和 $\mathbf{J}_r^{-1}$ 为 $\mathrm{SO}(3)$ 上的右雅可比：

$$\mathbf{J}_r^{-1}(\theta\boldsymbol{\omega})=\frac{\theta}{2}\cot\frac{\theta}{2}\mathbf{I}+\left(1-\frac{\theta}{2}\cot\frac{\theta}{2}\right)\boldsymbol{\omega}\boldsymbol{\omega}^{\mathrm{T}}+\frac{\theta}{2}\boldsymbol{\omega}^\wedge. \tag{17}$$

预备知识回顾

小结：旋转连乘的雅克比
关键是计算该旋转结果处的右雅克比的逆，记为M
对R2求导，结果就是M
对R1求导，结果是M右乘R2的转置

旋转连乘的雅可比：

$$\begin{aligned}\frac{\mathrm{d}\ln\left(\mathbf{R}_1\mathbf{R}_2\right)^{\vee}}{\mathrm{d}\mathbf{R}_1} &= \lim_{\phi\to 0}\frac{\ln\left(\mathbf{R}_1\exp\left(\phi^{\wedge}\right)\mathbf{R}_2\right)^{\vee}-\ln\left(\mathbf{R}_1\mathbf{R}_2\right)^{\vee}}{\phi} \\ &= \lim_{\phi\to 0}\frac{\ln\left(\mathbf{R}_1\mathbf{R}_2\exp\left(\left(\mathbf{R}_2{}^{\top}\phi\right)^{\wedge}\right)\right)^{\vee}-\ln\left(\mathbf{R}_1\mathbf{R}_2\right)^{\vee}}{\phi} \\ &= \mathbf{J}_r^{-1}(\ln(\mathbf{R}_1\mathbf{R}_2)^{\vee})\mathbf{R}_2^{\top}\end{aligned} \tag{18}$$

SO3的伴随性质是什么？一个旋转
李群空间上左乘RT，右乘R，相当于：
李代数空间上左乘RT。
作用其实就是交换了顺序

这里用到了 SO(3) 的伴随性质：

$$\boxed{\mathbf{R}^{\top}\exp\left(\phi^{\wedge}\right)\mathbf{R}=\exp\left(\left(\mathbf{R}^{\top}\phi\right)^{\wedge}\right)} \tag{19}$$

预备知识回顾

有关 $\mathrm{SE}(3)$：由于 $\mathrm{SE}(3)$ 李代数性质复杂，在 VIO 中，我们通常使用 $\mathrm{SO}(3)+\mathbf{t}$ 的形式表达旋转和平移。对平移部分使用矢量更新而非 $\mathrm{SE}(3)$ 上的更新。

习题

1. VIO 文献阅读

阅读 VIO 相关综述文献如[a]，回答以下问题：

- 视觉与 IMU 进行融合之后有何优势？
- 有哪些常见的视觉 +IMU 融合方案？有没有工业界应用的例子？
- 在学术界，VIO 研究有哪些新进展？有没有将学习方法用到 VIO 中的例子？

你也可以对自己感兴趣的方向进行文献调研，阐述你的观点。

[a] Jianjun Gui et al. “A review of visual inertial odometry from filtering and optimisation perspectives”. In: *Advanced Robotics* 29.20 (2015), 1289–1301. ISSN: 0169-1864. DOI: {10.1080/01691864.2015.1057616}.

习题

两种更新方式：
R右乘小量w对应的SO3
或者q右乘一个四元数（实部为1，虚部为w/2）

2. 四元数和李代数更新

课件提到了可以使用四元数或旋转矩阵存储旋转变量。当我们用计算出来的 ω 对某旋转更新时，有两种不同方式：

$$\begin{aligned} & \mathbf{R} \leftarrow \mathbf{R}\exp\left(\boldsymbol{\omega}^{\wedge}\right) \\ \text{或} \quad & \mathbf{q} \leftarrow \mathbf{q} \otimes \left[1, \tfrac{1}{2}\boldsymbol{\omega}\right]^{\top} \end{aligned} \tag{20}$$

请编程验证对于小量 $\boldsymbol{\omega} = [0.01, 0.02, 0.03]^{\mathrm{T}}$，两种方法得到的结果非常接近，实践当中可视为等同。因此，在后文提到旋转时，我们并不刻意区分旋转本身是 q 还是 $\mathbf{R}$，也不区分其更新方式为上式的哪一种。

习题

深蓝学院 shenlanxueyuan.com

3. 其他导数

使用右乘 $\mathfrak{so}(3)$，推导以下导数：

$$\frac{\mathrm{d}\left(\mathbf{R}^{-1}\mathbf{p}\right)}{\mathrm{d}\mathbf{R}} \tag{21}$$

$$\frac{\mathrm{d}\ln\left(\mathbf{R}_1\mathbf{R}_2^{-1}\right)^{\vee}}{\mathrm{d}\mathbf{R}_2} \tag{22}$$