

## 一. 公式推导

仿照 Eq.(12)-(14)，推导当 vins 中对特征采用逆深度参数时，基于特征匀速模型的重投影误差计算形式。设世界坐标系中的第  $l$  个地图点为  $\mathbf{f}_l^w = [x, y, \lambda]^\top$  其中  $\lambda = \frac{1}{z}$ ，先将其变换到相机坐标系下：

$$\mathbf{f}_l^c = \mathbf{R}_{cb}\mathbf{R}_{wb_i}(\mathbf{f}_l^w - \mathbf{p}_{wb_i}) + \mathbf{p}_{cb} \quad (1)$$

再将其投影到图像平面，并计算重投影误差：

$$\mathbf{r}_c = [\lambda_l^i x_l^i - u_l^i \quad \lambda_l^i y_l^i - v_l^i]^\top \quad (2)$$

考虑时间戳延迟时，对特征坐标进行补偿

$$\mathbf{Z}_l^i(t_d) = [u_l^i \quad v_l^i]^\top + t_d \mathbf{V}_l^i \quad (3)$$

推导结束，不知道是不是这个意思，哈哈。

## 二. 阅读论文

总结基于 B 样条的时间戳估计算法流程，梳理论文 [1] 公式。

见下一页：

## Unified Temporal and Spatial Calibration for Multi-sensor System

### 0. ABSTRACT

多传感器融合有助于增加机器人应用中状态估计的精度和鲁棒性。为了使多传感器融合的效果得到最佳体现，各个传感器需要知道彼此的时间和空间状态 (个人理解: 时间戳对齐 外参标定)。为了达到这个目的，许多研究中采用的是两步走的策略——先进行时间戳对齐，然后在进行传感器之间的外参标定。

本文中作者提出看一种同步进行时间戳对齐和外参标定的方法。作者通过连续时间的批量估计实现同步标定的效果，并通过最大后验概率推导的方式给出了严格证明，拓展了此前的许多研究。

对于 camera 和 IMU 标定的实验证明，这种理论框架能够精确估计不同传感器的时间戳偏差 (比最小的采样周期还要小)。

### 1. INTRODUCTION

### 2. RELATED WORK

### 3. THEORY

将问题定义在 camera 和 IMU 两个传感器的 time offset 标定上，可以拓展到多个传感器上去。

#### A. Estimating Time Offsets using Basis Functions

根据 Furgale[2] 等提出的用于连续时间的批量状态估计的 Basis Function 展开这一节的推导。时变的状态量由一组已知的 analytical basis function 的加权和进行表示。For example, 一个  $D$  维的状态量  $\mathbf{x}(t)$  可以被表示为

$$\Phi(t) = [\Phi_1(t) \quad \dots \quad \Phi_B(t)], \quad \mathbf{x}(t) = \Phi(t)\mathbf{c} \quad (1)$$

其中  $\phi_b(t)$  是已知的  $D \times 1$  的 analytical function(解析函数),  $\Phi(t)$  是  $D \times B$  的矩阵, 我们要求状态量  $\mathbf{x}(t)$ , 则只用求得这个权值系数  $\mathbf{c}$  就可以了。我们将在求解 time offset 过程中遇到的误差函数定义如下:

$$\mathbf{e}_j = \mathbf{y}_j - \mathbf{h}(\mathbf{x}(t_j + d)) \quad (2)$$

其中  $\mathbf{y}_j$  是  $t_j$  时刻的测量数据,  $\mathbf{h}(\cdot)$  是预测模型,  $d$  是 time offset, 总结就是 residual =

测量值 - 预测值，这里后面要用来将公式 (2) 带入公式 (1) 得到公式 (3):

$$\mathbf{e}_j = \mathbf{y}_j - \mathbf{h}(\Phi(t_j + d)\mathbf{c}) \quad (3)$$

然后对公式 (3) 进行一阶展开 (在一个 nominal(小的、微的) value 上, 展开肯定要在 nominal value 的基础上展开啊) 得到公式 (4)

$$\mathbf{e}_j = \mathbf{y}_j - \mathbf{h}(\Phi(t_j + \bar{d})\mathbf{c}) - \mathbf{H}(\dot{\Phi}(t_j + \bar{d})\mathbf{c}\Delta t) \quad (4)$$

其中符号上面的点表示求导

$$\mathbf{H} = \frac{\partial \mathbf{h}}{\partial \mathbf{x}} \quad (5)$$

这个公式没打完，后面还有一点小的东西，太难打了。在公式 (4) 中， $\Phi(t)$  是已知的函数，因此我们认为  $\Phi(t)$  的导数  $\dot{\Phi}(t)$  也是已知的。

## B. An Example: Camera/IMU Calibration

本文中专注与 IMU 和 camera 的时间戳对齐和外参标定研究，而没有进行 general case(许多传感器) 的分析，因为这里的理论证明这里的理论分析可以应用到其他地方。

实验中采集数据方法是拿着传感器在标定板前面晃动，采集数据的持续时间在 1-2 分钟左右。实验装置设置如图 1 所示，。世界坐标系、IMU 坐标系和 camera 坐标系分别是  $F_w, F_i(t), F_c(t)$

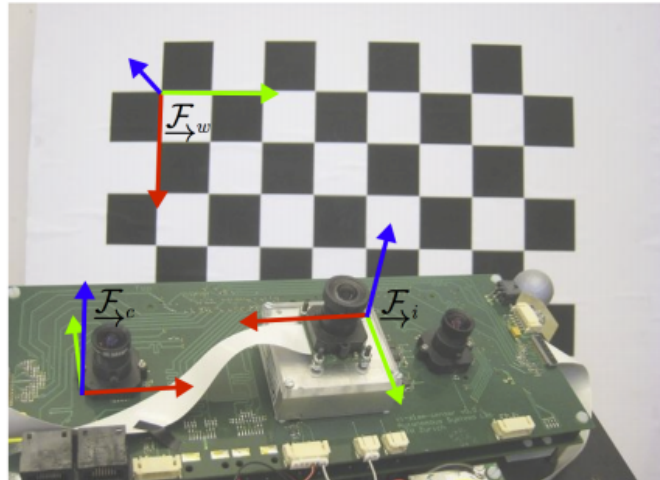


Fig. 1. This paper derives a unified framework for temporal and spatial calibration of multi-sensor systems such as this visual-inertial sensor used in the experiments.  $\mathcal{F}_w$  marks the inertial frame attached to a static calibration pattern, while  $\mathcal{F}_c$  and  $\mathcal{F}_i$  show the camera and the IMU frame respectively. By waving the sensor in front of the calibration pattern, our framework is capable of estimating the 6-degrees-of-freedom transformation between  $\mathcal{F}_c$  and  $\mathcal{F}_i$  as well as time offset between the two devices.

图 1: 实验装置

1) **Quantities Estimated** 估计了的参数 (量纲): 主要分析了五个状态变量:

- 在世界坐标系中的重力方向:  $g_w$ ,
- IMU 和相机之间的 transformation:  $T_{c,i}$
- IMU 和 camera 的 time offset:  $d$
- IMU 的 pose:  $T_{w,i}(t)$
- IMU 加速度计的 bias:  $b_a(t)$  和陀螺仪的 bias:  $b_w(t)$

2) **Parameterization of Time-Varying States** 用 B 样条拟合时变状态量, B 样条可以很好的拟合许多离散的量, 在飞机、船舶的设计以及计算机图形学中都有广泛的应用。

IMU 的位姿由一个  $6 \times 1$  的样条表示, 一个三维向量表示旋转自由度, 另外三维表示 translation。在任意时刻  $t$ , 将 IMU 坐标系下的点的坐标变换到世界坐标系下的变换矩阵如公式 (6) 所示。

$$T_{w,i}(t) = \begin{bmatrix} \mathbf{C}(\Phi(t)) & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \quad (6)$$

公式 (6) 大概的意思就是把 6 维的旋转平移向量转换为矩阵的过程, 然后速度和加速度在世界坐标系中的表示如公式 (7) 所示。

$$\mathbf{v}(t) = \dot{\mathbf{t}}(t) = \dot{\Phi}(t)\mathbf{c}_t, \quad \mathbf{a}(t) = \ddot{\mathbf{t}}(t) = \ddot{\Phi}(t)\mathbf{c}_t \quad (7)$$

给定旋转参数的情况下角速度可以通过公式 (8) 求出:

$$\omega(t) = \mathbf{S}(\varphi(t))\dot{\varphi}(t) = \mathbf{S}(\Phi(t)\mathbf{c}_\varphi)\dot{\Phi}(t)\mathbf{c}_\varphi \quad (8)$$

3) **Measurement and Process Models** IMU 的加速度和角速度数据是在  $t_k$  时刻采集得到的, 而相机的图像则是在  $t_j + d$  时刻采集得到的, 其中  $d$  表示未知的 time offset。标准的离散时间的 IMU 和相机的测量公式如公式 (9) 所示 (我到这里实在是敲不动了, 后面 3 个公式截图了 哈哈), 其中  $\alpha_k, \tilde{\omega}_k$  分别是  $t_k$  时刻的 IMU 测得的加速度和角速度。  $y_{m,j}$  是  $t_j + d$  时刻获得的路标点  $\mathbf{p}_w^m$  的像素坐标,  $t_j$  是时间戳, 而  $d$  是不确定的 time offset。

$$\alpha_k := \mathbf{C}(\varphi(t_k))^T (\mathbf{a}(t_k) - \mathbf{g}_w) + \mathbf{b}_a(t_k) + \mathbf{n}_{a_k}, \quad (9a)$$

$$\tilde{\omega}_k := \mathbf{C}(\varphi(t_k))^T \omega(t_k) + \mathbf{b}_\omega(t_k) + \mathbf{n}_{\omega_k}, \quad (9b)$$

$$\mathbf{y}_{m,j} := \mathbf{h}(\mathbf{T}_{c,i}\mathbf{T}_{w,i}(t_j + d)^{-1}\mathbf{p}_w^m) + \mathbf{n}_{y_{m,j}}, \quad (9c)$$

公式 (10) 主要是关于 IMU bias 的，文中假设 bias 是受到零均值的高斯影响 (相当于是说随机游走?)

$$\dot{\mathbf{b}}_a(t) = \mathbf{w}_a(t) \quad \mathbf{w}_a(t) \sim \mathcal{GP}(\mathbf{0}, \mathbf{Q}_a \delta(t - t')) \quad (10a)$$

$$\dot{\mathbf{b}}_\omega(t) = \mathbf{w}_\omega(t) \quad \mathbf{w}_\omega(t) \sim \mathcal{GP}(\mathbf{0}, \mathbf{Q}_\omega \delta(t - t')) \quad (10b)$$

4) **The Estimators** 本文中作者主要评估了 Section 3 B. 1) 中的五个变量的残差和雅克比。残差 = 测量值 - 预测值其中各个变量的预测方程应该是用的 analytical function(B spline)。这里应该就是本文的精华所在了，内容如公式 (11) 所示，还要仔细研究一下样条曲线。

$$\mathbf{e}_{y_{mj}} := \mathbf{y}_{mj} - \mathbf{h}(\mathbf{T}_{c,i} \mathbf{T}_{w,i}(t_j + d)^{-1} \mathbf{p}_w^m) \quad (11a)$$

$$J_y := \frac{1}{2} \sum_{j=1}^J \sum_{m=1}^M \mathbf{e}_{y_{mj}}^T \mathbf{R}_{y_{mj}}^{-1} \mathbf{e}_{y_{mj}} \quad (11b)$$

$$\mathbf{e}_{\alpha_k} := \alpha_k - \mathbf{C}(\varphi(t_k))^T (\mathbf{a}(t_k) - \mathbf{g}_w) + \mathbf{b}_a(t_k) \quad (11c)$$

$$J_\alpha := \frac{1}{2} \sum_{k=1}^K \mathbf{e}_{\alpha_k}^T \mathbf{R}_{\alpha_k}^{-1} \mathbf{e}_{\alpha_k} \quad (11d)$$

$$\mathbf{e}_{\omega_k} := \varpi_k - \mathbf{C}(\varphi(t_k))^T \omega(t_k) + \mathbf{b}_\omega(t_k) \quad (11e)$$

$$J_\omega := \frac{1}{2} \sum_{k=1}^K \mathbf{e}_{\omega_k}^T \mathbf{R}_{\omega_k}^{-1} \mathbf{e}_{\omega_k} \quad (11f)$$

$$\mathbf{e}_{b_a}(t) := \dot{\mathbf{b}}_a(t) \quad (11g)$$

$$J_{b_a} := \frac{1}{2} \int_{t_1}^{t_K} \mathbf{e}_{b_a}(\tau)^T \mathbf{Q}_a^{-1} \mathbf{e}_{b_a}(\tau) d\tau \quad (11h)$$

$$\mathbf{e}_{b_\omega}(t) := \dot{\mathbf{b}}_\omega(t) \quad (11i)$$

$$J_{b_\omega} := \frac{1}{2} \int_{t_1}^{t_K} \mathbf{e}_{b_\omega}(\tau)^T \mathbf{Q}_\omega^{-1} \mathbf{e}_{b_\omega}(\tau) d\tau \quad (11j)$$

#### 4. IMPLEMENTATION

简单介绍一下图 2 和图 3，图 2 中显示了一个稀疏的信息矩阵，是通过 L-M 方法构建的，从中可以看出各个变量之间的关系，有关联的变量之间的信息矩阵块就有值，否则就为 0。矩阵的大小受到样条的阶数和采样时间影响 (采样时间越长，样本数量越多，矩阵也就越大) 图中只显示了 0.1 秒采样时间的结果。

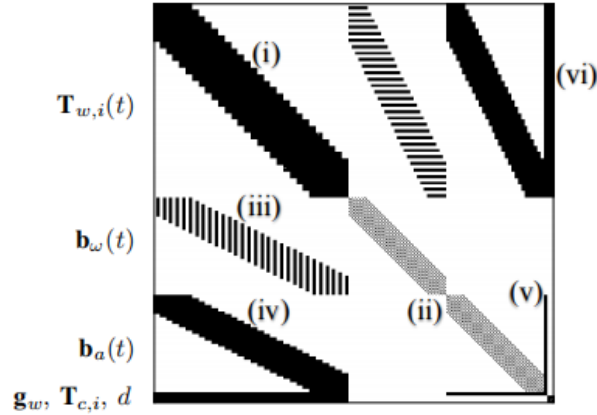


Fig. 2. A visualization of the sparsity pattern of the symmetric information matrix built during each iteration of Levenberg-Marquardt. To enhance the visibility of the matrix structure, only 0.1 seconds of data was used to generate this plot. The time-varying pose spline parameters produce a block diagonal (i) corresponding to  $\mathbf{T}_{w,i}(t)$ . The width of this diagonal depends on the spline order as well as the temporal padding. The diagonal blocks associated with the bias splines (ii) contain information from the bias motion model, (10). The parameters of the gyro bias spline,  $\mathbf{b}_{\omega}(t)$ , show striped correlation with the pose spline (iii) because they are only coupled with the IMU orientation through (11e). The parameters of the accelerometer bias,  $\mathbf{b}_a(t)$  are correlated with orientation (iv), position (iv), and gravity (v) through (11c). Finally,  $\mathbf{g}_w$ ,  $\mathbf{T}_{c,i}$  and the time offset,  $d$ , are all correlated with the  $\mathbf{T}_{w,i}(t)$  (vi) through (11a).

图 2: 信息矩阵

图 3 是 cost function 随 time offset 变化的影响，可以看到这是一个存在极小值的函数（凸的）。

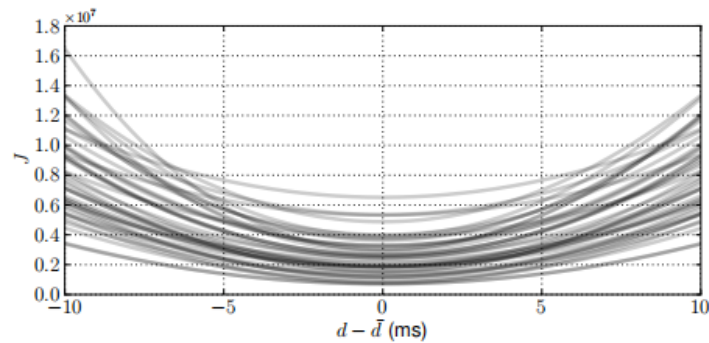


Fig. 3. The cost function evaluated for different values of the time offset,  $d$ , in the neighborhood of the minimum,  $\bar{d}$ , on 40 real datasets from Section V-B. In this neighborhood, the cost function is convex.

图 3: cost function 和 time offset

后面的实验和总结先不写了，来不起了。不过多传感器融合这一块，我接触得比较少，读了这个论文感觉打开了新世界的大门啊，先码一下，后面再看。

## 7. 自己的一点思考

在这些状态估计问题中 (主要是 SLAM 的几种方法), 我粗略地根据优化方法分为一下三种:

- filter based 的目标是求最大后验概率
- 纯视觉的 optimization based 方法是求最小重投影误差
- VIO 中也是求最大后验概率

上述三种方法其实都是在求最大后验概率, 重投影误差的计算过程中也应该引入像素的协方差矩阵。

信息矩阵是协方差矩阵的 inverse

最小二乘求解的问题中, 也是先把求最大后验概率的问题转化为求解  $\max \text{ a posterior} = \min -\log(\text{MAP})$

然后再求  $-\log(\text{MAP})$  的导数, 那么这个 MAP 的形式是怎么样的呢, 好像没有具体的推导 VIO 中也是视觉重投影误差构建残差 IMU 的测量值和预测值共同残差, 然后求这两个加上前面的先验信息共同构成的 residual block 的导数

进一步求得状态量 (camera pose + IMU states) 的最大后验估计。

covariance 在其中起到的作用:

- 视觉重投影误差的  $\text{residual} = \text{covariance} * \text{residual}$
- IMU states 的  $\text{residual} = \text{covariance} * \text{residual}$

进一步, residual 对 states 求导得到 Jacobian 的时候, 也要带上这个 covariance。同时 covariance 的均值和方差也会受到随机变量的影响 (ba, bg) 因此有了 ch3 第三节的  $F = (A\Delta t + I)$  和  $G = (B\Delta t)$

这个信息矩阵和协方差矩阵还有最小化 residual、求最大后验概率等, 我感觉属实是有点复杂了, 到现在都还没完全理清楚, 慢慢打磨吧, 最后还是想问一下梁哥, 你觉得我上面的理解, 有存在偏差的地方吗

## 参考文献

- [1] Furgale, P., Rehder, J., Siegwart, R. (2013, November). Unified temporal and spatial calibration for multi-sensor systems. In 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (pp. 1280-1286). IEEE.
- [2] Furgale, P., Barfoot, T. D., Sibley, G. (2012, May). Continuous-time batch estimation using temporal basis functions. In 2012 IEEE International Conference on Robotics and Automation (pp. 2088-2095). IEEE.