基于优化的IMU与视觉信息融合

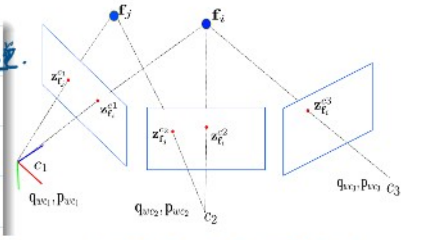
2019年11月23日 14:50

① 基于 Bundle Adjustment 的 VIO 融合.

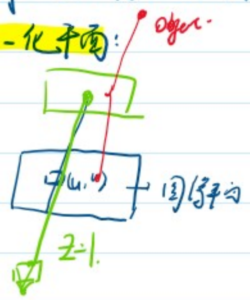
已知:

状态量初始值: 特征点的三维坐标. 相机的位姿.

系统测量值: 特征点在不同图像上的图像坐标.



Tipps: 归一化平面:



① 将特征点从三维空间中, 取与相机前 z=1 的距离形成一个投影平面.

② 空间中的一个三维点首先投到归一化平面上, 再通过一步单位的映射关系得到像素坐标映射成像素坐标

$$\begin{bmatrix} u \\ v \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}_{(m)}$$

符号定义: q: 旋转四元数. p: 平移向量. f: 特征点3D坐标.

c_i: 第 i 个相机系. $\pi(\cdot)$: 投影函数. $z_{f_j}^{c_i}$: c_i 对 f_j 的观测.

Σ_{ij}: 协方差 (权重)

解决方法: 构建误差函数, 利用最小二乘得到状态量的最优估计.

$$\underset{q,p,f}{\arg\min} \sum\sum \| \pi(q_{wc_i}, p_{wc_i}, f_j) - z_{f_j}^{c_i} \|_{\Sigma_{ij}}$$

Input: A vector function $f : \mathcal{R}^m \to \mathcal{R}^n$ with $n \geq m$, a measurement vector $\mathbf{x} \in \mathcal{R}^n$ and an initial parameters estimate $\mathbf{p}_0 \in \mathcal{R}^m$.
Output: A vector $\mathbf{p}^+ \in \mathcal{R}^m$ minimizing $||\mathbf{x} - f(\mathbf{p})||^2$.
Algorithm:

```
k := 0; ν := 2; p := p0;
A := J^T J; ε_p := x − f(p); g := J^T ε_p;
stop:=(||g||_∞ ≤ ε1); μ := τ ∗ max_{i=1,...,m}(A_ii);
while (not stop) and (k < k_max)
    k := k + 1;
    repeat
        Solve (A + μI)δ_p = g;
        if (||δ_p|| ≤ ε2(||p|| + ε2))
            stop:=true;
        else
            p_new := p + δ_p;
            ρ := (||ε_p||^2 − ||x − f(p_new)||^2)/(δ_p^T(μδ_p + g));
            if ρ > 0
                stop:=(||ε_p|| − ||x − f(p_new)|| < ε4||ε_p||);
                p = p_new;
                A := J^T J; ε_p := x − f(p); g := J^T ε_p;
                stop:=(stop) or (||g||_∞ ≤ ε1);
                μ := μ ∗ max(1/3, 1 − (2ρ − 1)^3); ν := 2;
            else
                μ := μ ∗ ν; ν := 2 ∗ ν;
            endif
        endif
    until (ρ > 0) or (stop)
    stop:=(||ε_p|| ≤ ε3);
endwhile
p+ := p;
```

Section 2. 最小二乘问题的求解:

假设损失函数 $F(x)$ 是可导并且平滑的，因此，二阶泰勒展开：

$$F(x+\Delta x) = F(x) + J\Delta x + \frac{1}{2}\Delta x^T H \Delta x + O(\|\Delta x\|^3)$$

其中 J 和 H 为损失函数 F 对变量 x 的一阶导和二阶导矩阵.

当正定 x 处导数为0，则称这个点为驻点.

此时当 H 为正定矩阵 → 局部最小值

负定矩阵 → 局部最大值

不定矩阵 → 鞍点

linear search.

① 最速下降法：适用于迭代的开始阶段：

搜索的方向为最速下降方向

缺点：最优值附近震荡收敛慢

② 牛顿法：适用于最优值附近

局部

在最优点 x 附近，如果 $x+\Delta x$ 是最优解，则损失函数对 Δx 的导数等于 0

$$\frac{\partial}{\partial \Delta x}\left(F(x) + J\Delta x + \frac{1}{2}\Delta x^T H \Delta x\right) = J^T + H\Delta x = 0$$

得 $\Delta x = -H^{-1}J^T$

缺点：二阶导矩阵计算复杂.

③ 阻尼法：Damp Method.

$$F(x+\Delta x) \approx L(\Delta x) \equiv F(x) + J\Delta x + \frac{1}{2}\Delta x^T H \Delta x$$

求以下函数的最小化：

$$\Delta x \equiv \arg\min_{\Delta x}\left\{L(\Delta x) + \frac{1}{2}\mu \Delta x^T \Delta x\right\}$$

↳ 防止步长过大

其中，$\mu \geq 0$ 为阻尼因子，$\frac{1}{2}\mu\Delta x^T\Delta x = \frac{1}{2}\|\Delta x\|_2$ 是惩罚项.

对新的损失函数求一阶导，并令其等于 0. 得

$$L'(\Delta x) + \mu\Delta x = 0$$

$$\Rightarrow J^T + (H + \mu I)\Delta x = 0$$

非线性最小二乘：

残差函数 $f(x)$ 为非线性函数，对其一阶泰勒近似有

$$f(x+\Delta x) \approx l(\Delta x) = f(x) + J\Delta x.$$ 残差函数f的雅克比矩阵

代入损失函数：

$$F(x+\Delta x) \approx L(\Delta x) = \frac{1}{2} l(\Delta x)^T l(\Delta x)$$
$$= \frac{1}{2} f^T f + \Delta x^T J^T f + \frac{1}{2}\Delta x^T J^T J \Delta x$$
$$= F(x) + \Delta x^T J^T f + \frac{1}{2}\Delta x^T J^T J \Delta x$$

$Jx=0 \Leftrightarrow x=0$

这样损失函数就近似成了一个二次函数，并且如果雅克比是满秩的，则 J^TJ 是正定的，损失函数有最小值。

若：$x \neq 0 \Rightarrow Jx \neq 0$
$\Rightarrow (Jx)^T(Jx) > 0$
$\Rightarrow J^TJ$ 是正定矩阵。

另外，易得：$F'(x) = (J^T f)^T = -b^T$

以及：$F''(x) = J^T J$.

$F'(x+\Delta x) = J^T f + J^T J \cdot \Delta x = 0$
$-b + H\Delta x = 0$
$\Delta x \quad H \cdot \Delta x_{gn} = b$

令一阶导等于0，则有 (Gaussian - Newton)

$$\underbrace{J^T f}_{-b} + \underbrace{J^T J}_{H} \cdot \Delta x = 0$$

$$\Rightarrow H \Delta x_{gn} = b. \quad \Rightarrow \text{normal equation}$$

The Levenberg - Marquardt Method.

$$(J^T J + \mu I)\Delta x_{lm} = -J^T f \quad \text{with } \mu \geq 0$$

阻尼因子的作用：

$\mu > 0$，保证 $(J^T J + \mu I)$ 正定，迭代朝着下降方向进行。

μ 非常大时，则 $\Delta x_{lm} = -\frac{1}{\mu} J^T f = -\frac{1}{\mu} F'(x)^T$ 接近最速下降法

μ 比较小时，则 $\Delta x_{lm} \approx \Delta x_{gn}$ 接近高斯牛顿法

阻尼因子初始值的选取：

阻尼因子 μ 大小是相对于 J^TJ 的元素而言。半正定的信息矩阵 J^TJ 特征值 $\{\lambda_j\}$ 和对应的特征向量为 $\{v_j\}$。对 J^TJ 做特征值分解后有：$J^TJ = V\Lambda V^T$，可得：

$$\Delta x_{lm} = -\sum_{j=1}^{n} \frac{v_j^T F'^T}{\lambda_j + \mu} v_j$$

已知最大特征值在一个数量级，但特征值的计算代价很高，只能利用下面的技巧。最大特征值的数量级，近似和对角线上最大元素的数量级一致。

所以一个简单的 μ_0 初始值的策略就是：

$$\mu_0 = \tau \cdot \max\{(J^T J)_{ii}\}, \quad \tau \in [10^{-8}, 1]$$

对角线上元素

初始值的 τ 为 10^{-8}

较近

利用上述表达，所有对角线上的元素为对角线上的元素

·

阻尼因子μ的更新策略

①定性分析：

① 如果 $\Delta x \to F(x)\uparrow$，则 $\mu\uparrow \to \Delta x\downarrow$，增大阻尼减小步长，拒绝本次迭代

② 如果 $\Delta x \to F(x)\downarrow$，则 $\mu\downarrow \to \Delta x\uparrow$，减小阻尼增大步长，加快收敛，减少迭代次数。

②定量分析：

$$\rho = \frac{F(x) - F(x+\Delta x_{lm})}{\underbrace{L(0) - L(\Delta x_{lm})}_{>0}} = \frac{\text{实际下降}}{\text{近似下降}}$$

$$L(0) - L(\Delta x_{lm}) = -\Delta x_{lm}^T J^T f - \frac{1}{2}\Delta x_{lm}^T J^T J \Delta x_{lm}.$$

$$\overset{\text{令}-J^Tf}{=} -\frac{1}{2}\Delta x_{lm}^T(-2b + (J^TJ + \mu I - \mu I)\Delta x_{lm})$$

$$= \frac{1}{2}\Delta x_{lm}^T(\mu\Delta x_{lm} + b)$$

若 $\rho < 0$，则 $F(x)\uparrow$，应该 $\mu\uparrow \to \Delta x\downarrow$，增大阻尼，减小步长。

$\rho > 0$ 且比较大，减小 μ，让LM接近GN更快收敛。

$\rho > 0$ 但比较小，增大 μ，缩小迭代步长。

一种策略：

$\rho < 0.25$ $\quad \mu = 2 * \mu$

$\rho > 0.75$ $\quad \mu = \mu/3$

Nielsen策略：（g2o, ceres） $\Rightarrow$ 减小了μ的振荡。

if $\rho > 0$

$\mu := \mu * \max\{\frac{1}{3}, 1-(2\rho-1)^3\}$; $\quad v := 2$

else:

$\mu := \mu * v$ $\quad v := 2 * v$

鲁棒核函数 ⇔ M-Estimation / Adjust-estimation.

引言：最小二乘中遇到outlier怎么处理？

⇒ Triggs Correction：

鲁棒核函数直接作用在残差 $f_k(x)$ 上，最小二乘函数变成如下形式：

$$\min_x \frac{1}{2}\sum_k \rho(\|f_k(x)\|^2).$$

将误差的平方记作：$s_k = \|f_k(x)\|^2$，则鲁棒核误差函数进行二次泰勒展开：

$$\frac{1}{2}\rho(s) = \frac{1}{2}(\text{const.} + \rho'\cdot\Delta s + \frac{1}{2}\rho''\Delta^2 s)$$

$$\Delta s_k = \|f_k(x+\Delta x)\|^2 - \|f_k(x)\|^2$$

$$\approx \|f_k + J_k\Delta x\|^2 - \|f_k(x)\|^2.$$

$$= 2f_k^T J_k\Delta x + \Delta x^T J_k^T J_k \Delta x.$$

回代有: $\frac{1}{2}\rho(s) = \frac{1}{2}(const) + \frac{1}{2}\rho'[2f_k^T J_k \cdot \Delta x + \Delta x^T J_k^T J_k \cdot \Delta x]$
$+\frac{1}{2}\rho''[2f_k^T J_k \Delta x + (\Delta x)^T J_k^T J_k \cdot \Delta x]^2 -$

忽略 Δx 3次以上高阶项

$\approx \rho' f_k J_k \Delta x + \frac{1}{2}\rho' \Delta x^T J_k^T J_k \Delta x + \rho''(\Delta x)^T J_k^T f_k f_k^T J_k \Delta x + const$

$= \rho' f_k J_k \Delta x + \frac{1}{2}(\Delta x)^T J_k^T(\rho' I + 2\rho'' f_k f_k^T) J_k \Delta x + const$

对于Δx求导，令其等于0,得到:

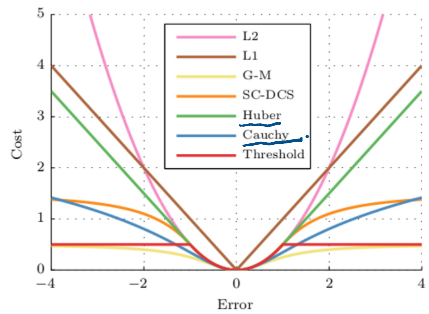
$$\sum_k J_k^T(\rho' I + 2\rho'' f_k f_k^T) J_k \Delta x = -\sum_k \rho' f_k^T J_k.$$

$$\sum_k J_k^T \boxed{W} J_k \Delta x = -\sum_k \boxed{\rho'} f_k^T J_k.$$

example. Cauchy cost function:

$\rho(s) = c^2 \log(1 + \frac{s}{c^2})$，其中 c 为控制参数.

$\rho' = \frac{1}{1+\frac{s}{c^2}} \qquad \rho'' = -\frac{1}{c^2}(\rho'(s))^2$



控制参数的选定:
95% efficiency rule:
如果残差为正态分布:
Huber $c = 1.345$
Cauchy $c = 2.3849$ ⟺ IRLS
如果残差非正态分布,就估计残差方差,然后对残差归一化
median absolute residual 方法.
$\sigma = 1.482\, med(|med(r) - r_i|)$