

## DSO windowed optimization 代码 (1)

这里不想解释怎么 `marginalize`, 什么是 First-Estimates Jacobian (FEI), 这里只看代码, 看看Hessian矩阵是怎么构造出来的。

## 1 优化流程

整个优化过程, 也是 Levenberg–Marquardt 的优化过程, 这个优化过程在函数 `FullSystem::makeKeyFrame()` 中被调用, 也是在确定当前帧成为关键帧, 并且用当前帧激活了窗口中其他帧的 `immaturePoints` 之后, 过程在 `FullSystem::optimize()` 函数中。

优化的目标值, 是所有需要优化点的逆深度、相机的4个内参数, 窗口中8个帧的状态量。

`FullSystem::optimize()` 函数流程大致如下。首先 `FullSystem::linearizeAll(false)` 把相关的导数计算一下, 然后在所有的 `residual` 中找到 `isLinearized` 为 `false` 的 residual, 调用 `PointFrameResidual::applyRes(true)`, 设置它们的 `PointFrameResidual::ResStat` (按照

`FullSystem::optimize(immaturePoint)` 的结果), 如果是正常的点(`ResState::IN`)调用 `EFResidual::takeDataF()`把 `EFResidual::JpJdF` 设置一下。这就算完成了优化的准备工作。

随后进入循环体, 进行最高有限次的优化循环。每一次优化都有可能使得整体能量升高, 所以每次优化前调用 `FullSystem::backupState()` 保存当前所有得优化参数的 state 和 step, 然后调用 `FullSystem::step()` 进行优化(得到的优化化量 step)。最后调用 `FullSystem::loadStateFromBackup()` 将优化结果生效。

`FullSystem::linearizeAll(false)` 计算新的能量, 如果能量高了, 就使用 `FullSystem::loadStateFromBackup()` 将结果回滚。

在跳出循环体之后调用一次 `FullSystem::linearizeAll(true)`, 效果是将优化之后成为 outlier 的 residual 剔除, 剩下正常的 residual 调用一次 `EFResidual::takeDataF()` 计算新的 `EFResidual::JpJdF` (这里涉及到 FEI)。注意一下 `FullSystem::linearizeAll()` 参数为 `false` 和 `true` 的区别。

`FullSystem::solveSystem()` 设置优化化量 step 的过程在 `EnergyFunctional::resubstituteIMT()` 中, 能帮助理解 `idepth` 是如何更新的 (`SchurComplement` 将 `idepth` 剔除最终需要矩阵求逆的系统, 而这些 `idepth` 的优化化量 step 是如何计算的?)。

## 2 导数准备

需要遍历每一个 `PointFrameResidual` 将与这个 `Residual` 相关的导数计算出来, 再进行优化。而这些计算出来的相关导数被存储在 `RawResidualJacobian` 中。

<https://github.com/JakobEngel/dso/blob/master/src/OptimizationBackend/RawResidualJacobian.h#L32>

在优化过程中 `Frame`, `Pointlessian`, `PointFrameResidual` 都有与之对应的实体, 分别是 `EFFrame`, `EFPoint`, `EFResidual`, 通过保留指针, 保存层次之间的索引。

在计算 `RawResidualJacobian` 时, 是计算 `PointFrameResidual` 的 `J`, 尔后会将这个 `J` 移到 `EFResidual` 的 `J`, 并且计算 `EFResidual::JpJdF`, 这个过程在 `EFResidual::takeDataF()` 中。所以这里把 `JpJdF` 的计算过程写出来, 弄清 `JpJdF` 的意义。

<https://github.com/JakobEngel/dso/blob/5fb2c065b1638e10ccf049a6575ede4334ba673/src/OptimizationBackend/EnergyFunctionalStruct.cpp#L37>

```
struct RawResidualJacobian
{
    EIGEN_MAKE_ALIGNED_OPERATOR_NEW;
    // ===== new structure: save independently =====
    VecNRF resF; // typedef Eigen::Matrix<float,MAX_RES_PER_POINT,1> VecNRF; MAX_RES_PER_POINT == 8

    // the two rows of d[x,y]/d[x1].
    Vec6f Jpxdx1[2]; // 2x6

    // the two rows of d[x,y]/d[C].
    VecCf JpdC[2]; // 2x4

    // the two rows of d[x,y]/d[idepth].
    Vec2f JpdD; // 2x1

    // the two columns of d[r]/d[x,y].
    VecNRF Jidx[2]; // 9x2

    // = the two columns of d[r] / d [ab]
    VecNRF JabF[2]; // 9x2

    // = Jidx * T * Jidx (inner product). Only as a shorthand.
    Mat22f JIdx2; // 2x2
    // = JabT * JIdx (inner product). Only as a shorthand.
    Mat22f JabIdx; // 2x2
    // = JabT * Jab (inner product). Only as a shorthand.
    Mat22f Jab2; // 2x2
};

};

以上变量的类型中出现 [NR], 说明该变量是存储了每一个 pattern 点的信息。
```

现在将这些变量对应的导致一一列出:

- `VecNRF resF`: 对应  $\frac{\partial r_2}{\partial \mathbf{x}}$ , 1x8, 这里的  $r_2$  是对于一个点, 八个 pattern residual 组成的向量。
- `Vec6f Jpxdx1[2]`: 对应  $\frac{\partial r_2}{\partial \mathbf{C}}$ , 2x6, 注意这里的  $\mathbf{C}$  是像素坐标。 (我一般把像素坐标写成  $\mathbf{x}$ , 对应代码中的变量 `[Ku]`, 归一化坐标写成  $\mathbf{x}'$ , 对应代码中的矩阵 `[u]` )
- `VecCf JpdC[2]`: 对应  $\frac{\partial r_2}{\partial \mathbf{C}}$ , 2x4, 这里的  $\mathbf{C}$  指相机内参  $[f_x, f_y, c_x, c_y]^T$ 。
- `Vec2f JpdD`: 对应  $\frac{\partial r_2}{\partial p_1}$ , 2x1, 注意是对 host 帧的逆深度求导。
- `VecNRF Jidx[2]`: 对应  $\frac{\partial r_2}{\partial \mathbf{x}_1}$ , 8x2, 这个和 target 帧上的影像梯度相关。
- `VecNRF JabF[2]`: 对应  $\frac{\partial r_2}{\partial \mathbf{x}_1}$ ,  $\frac{\partial r_2}{\partial \mathbf{x}_2}$ , 8x1, 8x1。
- `Mat22f JIdx2`: 对应  $\frac{\partial r_2}{\partial \mathbf{x}_1}^T \frac{\partial r_2}{\partial \mathbf{x}_2}$ , 2x8 8x2, 2x2。
- `Mat22f JabJdx`: 对应  $\frac{\partial r_2}{\partial \mathbf{x}_1}^T \frac{\partial r_2}{\partial \mathbf{x}_2}$ , 2x8 8x2, 2x2, 这里的  $J_{21}$  存储  $\begin{bmatrix} a_{21} \\ b_{21} \end{bmatrix}$ 。
- `Mat22f Jab2`: 对应  $\frac{\partial r_2}{\partial \mathbf{x}_1}^T \frac{\partial r_2}{\partial \mathbf{x}_2}$ , 2x8 8x2, 2x2。
- `JpJdF`: 对应  $\frac{\partial r_2}{\partial p_1}$ , 8x1。

在 `PointFrameResidual::linearize` 中对这些变量进行了计算。

<https://github.com/JakobEngel/dso/blob/master/src/FullSystem/Residuals.cpp#L78>

在计算时使用了投影过程中的变量, 现在将这些变量与公式对应。投影过程标准公式如下:

$$\begin{aligned} x_2 &= K\rho_2(R_2\rho_1^{-1}K^{-1}x_1 + t_2) \\ &= Kx'_2 \end{aligned}$$

变量的对应关系如下:

```
K1P = K^-1x1 = x'_1
p1P = R21K^-1x1 + rho1t21 = rho1^-1rho1K^-1x2
drescale = rho2^-1
[u, v, w, 1]^T = K^-1x2 = x'_2
[Ku, Kv, Kw, 1]^T = x2
1. Vec2f JpdD = partial derivative of x2 with respect to rho2
d_d_x = drescale * (PRE_tT11_0[0]+PRE_tT11_0[2]*u+PRE_tT11_0[0,0]);
d_d_y = drescale * (PRE_tT11_0[1]+PRE_tT11_0[2]*v+PRE_tT11_0[1,1]);
d_d_z = drescale * (PRE_tT11_0[2]+PRE_tT11_0[3]*w+PRE_tT11_0[2,2]);
d_d_w = drescale * (PRE_tT11_0[3]+PRE_tT11_0[4]*u+PRE_tT11_0[3,0]);
d_d_u = drescale * (PRE_tT11_0[0]+PRE_tT11_0[1]*v+PRE_tT11_0[0,1]);
d_d_v = drescale * (PRE_tT11_0[1]+PRE_tT11_0[2]*w+PRE_tT11_0[1,1]);
d_d_w = drescale * (PRE_tT11_0[2]+PRE_tT11_0[3]*u+PRE_tT11_0[2,1]);
d_d_u = drescale * (PRE_tT11_0[3]+PRE_tT11_0[4]*v+PRE_tT11_0[3,1]);
d_d_v = drescale * (PRE_tT11_0[4]+PRE_tT11_0[5]*w+PRE_tT11_0[4,1]);
d_d_w = drescale * (PRE_tT11_0[5]+PRE_tT11_0[6]*u+PRE_tT11_0[5,1]);
d_d_u = drescale * (PRE_tT11_0[6]+PRE_tT11_0[7]*v+PRE_tT11_0[6,1]);
d_d_v = drescale * (PRE_tT11_0[7]+PRE_tT11_0[8]*w+PRE_tT11_0[7,1]);
d_d_w = drescale * (PRE_tT11_0[8]+PRE_tT11_0[9]*u+PRE_tT11_0[8,1]);
d_d_u = drescale * (PRE_tT11_0[9]+PRE_tT11_0[10]*v+PRE_tT11_0[9,1]);
d_d_v = drescale * (PRE_tT11_0[10]+PRE_tT11_0[11]*w+PRE_tT11_0[10,1]);
d_d_w = drescale * (PRE_tT11_0[11]+PRE_tT11_0[12]*u+PRE_tT11_0[11,1]);
d_d_u = drescale * (PRE_tT11_0[12]+PRE_tT11_0[13]*v+PRE_tT11_0[12,1]);
d_d_v = drescale * (PRE_tT11_0[13]+PRE_tT11_0[14]*w+PRE_tT11_0[13,1]);
d_d_w = drescale * (PRE_tT11_0[14]+PRE_tT11_0[15]*u+PRE_tT11_0[14,1]);
d_d_u = drescale * (PRE_tT11_0[15]+PRE_tT11_0[16]*v+PRE_tT11_0[15,1]);
d_d_v = drescale * (PRE_tT11_0[16]+PRE_tT11_0[17]*w+PRE_tT11_0[16,1]);
d_d_w = drescale * (PRE_tT11_0[17]+PRE_tT11_0[18]*u+PRE_tT11_0[17,1]);
d_d_u = drescale * (PRE_tT11_0[18]+PRE_tT11_0[19]*v+PRE_tT11_0[18,1]);
d_d_v = drescale * (PRE_tT11_0[19]+PRE_tT11_0[20]*w+PRE_tT11_0[19,1]);
d_d_w = drescale * (PRE_tT11_0[20]+PRE_tT11_0[21]*u+PRE_tT11_0[20,1]);
d_d_u = drescale * (PRE_tT11_0[21]+PRE_tT11_0[22]*v+PRE_tT11_0[21,1]);
d_d_v = drescale * (PRE_tT11_0[22]+PRE_tT11_0[23]*w+PRE_tT11_0[22,1]);
d_d_w = drescale * (PRE_tT11_0[23]+PRE_tT11_0[24]*u+PRE_tT11_0[23,1]);
d_d_u = drescale * (PRE_tT11_0[24]+PRE_tT11_0[25]*v+PRE_tT11_0[24,1]);
d_d_v = drescale * (PRE_tT11_0[25]+PRE_tT11_0[26]*w+PRE_tT11_0[25,1]);
d_d_w = drescale * (PRE_tT11_0[26]+PRE_tT11_0[27]*u+PRE_tT11_0[26,1]);
d_d_u = drescale * (PRE_tT11_0[27]+PRE_tT11_0[28]*v+PRE_tT11_0[27,1]);
d_d_v = drescale * (PRE_tT11_0[28]+PRE_tT11_0[29]*w+PRE_tT11_0[28,1]);
d_d_w = drescale * (PRE_tT11_0[29]+PRE_tT11_0[30]*u+PRE_tT11_0[29,1]);
d_d_u = drescale * (PRE_tT11_0[30]+PRE_tT11_0[31]*v+PRE_tT11_0[30,1]);
d_d_v = drescale * (PRE_tT11_0[31]+PRE_tT11_0[32]*w+PRE_tT11_0[31,1]);
d_d_w = drescale * (PRE_tT11_0[32]+PRE_tT11_0[33]*u+PRE_tT11_0[32,1]);
d_d_u = drescale * (PRE_tT11_0[33]+PRE_tT11_0[34]*v+PRE_tT11_0[33,1]);
d_d_v = drescale * (PRE_tT11_0[34]+PRE_tT11_0[35]*w+PRE_tT11_0[34,1]);
d_d_w = drescale * (PRE_tT11_0[35]+PRE_tT11_0[36]*u+PRE_tT11_0[35,1]);
d_d_u = drescale * (PRE_tT11_0[36]+PRE_tT11_0[37]*v+PRE_tT11_0[36,1]);
d_d_v = drescale * (PRE_tT11_0[37]+PRE_tT11_0[38]*w+PRE_tT11_0[37,1]);
d_d_w = drescale * (PRE_tT11_0[38]+PRE_tT11_0[39]*u+PRE_tT11_0[38,1]);
d_d_u = drescale * (PRE_tT11_0[39]+PRE_tT11_0[40]*v+PRE_tT11_0[39,1]);
d_d_v = drescale * (PRE_tT11_0[40]+PRE_tT11_0[41]*w+PRE_tT11_0[40,1]);
d_d_w = drescale * (PRE_tT11_0[41]+PRE_tT11_0[42]*u+PRE_tT11_0[41,1]);
d_d_u = drescale * (PRE_tT11_0[42]+PRE_tT11_0[43]*v+PRE_tT11_0[42,1]);
d_d_v = drescale * (PRE_tT11_0[43]+PRE_tT11_0[44]*w+PRE_tT11_0[43,1]);
d_d_w = drescale * (PRE_tT11_0[44]+PRE_tT11_0[45]*u+PRE_tT11_0[44,1]);
d_d_u = drescale * (PRE_tT11_0[45]+PRE_tT11_0[46]*v+PRE_tT11_0[45,1]);
d_d_v = drescale * (PRE_tT11_0[46]+PRE_tT11_0[47]*w+PRE_tT11_0[46,1]);
d_d_w = drescale * (PRE_tT11_0[47]+PRE_tT11_0[48]*u+PRE_tT11_0[47,1]);
d_d_u = drescale * (PRE_tT11_0[48]+PRE_tT11_0[49]*v+PRE_tT11_0[48,1]);
d_d_v = drescale * (PRE_tT11_0[49]+PRE_tT11_0[50]*w+PRE_tT11_0[49,1]);
d_d_w = drescale * (PRE_tT11_0[50]+PRE_tT11_0[51]*u+PRE_tT11_0[50,1]);
d_d_u = drescale * (PRE_tT11_0[51]+PRE_tT11_0[52]*v+PRE_tT11_0[51,1]);
d_d_v = drescale * (PRE_tT11_0[52]+PRE_tT11_0[53]*w+PRE_tT11_0[52,1]);
d_d_w = drescale * (PRE_tT11_0[53]+PRE_tT11_0[54]*u+PRE_tT11_0[53,1]);
d_d_u = drescale * (PRE_tT11_0[54]+PRE_tT11_0[55]*v+PRE_tT11_0[54,1]);
d_d_v = drescale * (PRE_tT11_0[55]+PRE_tT11_0[56]*w+PRE_tT11_0[55,1]);
d_d_w = drescale * (PRE_tT11_0[56]+PRE_tT11_0[57]*u+PRE_tT11_0[56,1]);
d_d_u = drescale * (PRE_tT11_0[57]+PRE_tT11_0[58]*v+PRE_tT11_0[57,1]);
d_d_v = drescale * (PRE_tT11_0[58]+PRE_tT11_0[59]*w+PRE_tT11_0[58,1]);
d_d_w = drescale * (PRE_tT11_0[59]+PRE_tT11_0[60]*u+PRE_tT11_0[59,1]);
d_d_u = drescale * (PRE_tT11_0[60]+PRE_tT11_0[61]*v+PRE_tT11_0[60,1]);
d_d_v = drescale * (PRE_tT11_0[61]+PRE_tT11_0[62]*w+PRE_tT11_0[61,1]);
d_d_w = drescale * (PRE_tT11_0[62]+PRE_tT11_0[63]*u+PRE_tT11_0[62,1]);
d_d_u = drescale * (PRE_tT11_0[63]+PRE_tT11_0[64]*v+PRE_tT11_0[63,1]);
d_d_v = drescale * (PRE_tT11_0[64]+PRE_tT11_0[65]*w+PRE_tT11_0[64,1]);
d_d_w = drescale * (PRE_tT11_0[65]+PRE_tT11_0[66]*u+PRE_tT11_0[65,1]);
d_d_u = drescale * (PRE_tT11_0[66]+PRE_tT11_0[67]*v+PRE_tT11_0[66,1]);
d_d_v = drescale * (PRE_tT11_0[67]+PRE_tT11_0[68]*w+PRE_tT11_0[67,1]);
d_d_w = drescale * (PRE_tT11_0[68]+PRE_tT11_0[69]*u+PRE_tT11_0[68,1]);
d_d_u = drescale * (PRE_tT11_0[69]+PRE_tT11_0[70]*v+PRE_tT11_0[69,1]);
d_d_v = drescale * (PRE_tT11_0[70]+PRE_tT11_0[71]*w+PRE_tT11_0[70,1]);
d_d_w = drescale * (PRE_tT11_0[71]+PRE_tT11_0[72]*u+PRE_tT11_0[71,1]);
d_d_u = drescale * (PRE_tT11_0[72]+PRE_tT11_0[73]*v+PRE_tT11_0[72,1]);
d_d_v = drescale * (PRE_tT11_0[73]+PRE_tT11_0[74]*w+PRE_tT11_0[73,1]);
d_d_w = drescale * (PRE_tT11_0[74]+PRE_tT11_0[75]*u+PRE_tT11_0[74,1]);
d_d_u = drescale * (PRE_tT11_0[75]+PRE_tT11_0[76]*v+PRE_tT11_0[75,1]);
d_d_v = drescale * (PRE_tT11_0[76]+PRE_tT11_0[77]*w+PRE_tT11_0[76,1]);
d_d_w = drescale * (PRE_tT11_0[77]+PRE_tT11_0[78]*u+PRE_tT11_0[77,1]);
d_d_u = drescale * (PRE_tT11_0[78]+PRE_tT11_0[79]*v+PRE_tT11_0[78,1]);
d_d_v = drescale * (PRE_tT11_0[79]+PRE_tT11_0[80]*w+PRE_tT11_0[79,1]);
d_d_w = drescale * (PRE_tT11_0[80]+PRE_tT11_0[81]*u+PRE_tT11_0[80,1]);
d_d_u = drescale * (PRE_tT11_0[81]+PRE_tT11_0[82]*v+PRE_tT11_0[81,1]);
d_d_v = drescale * (PRE_tT11_0[82]+PRE_tT11_0[83]*w+PRE_tT11_0[82,1]);
d_d_w = drescale * (PRE_tT11_0[83]+PRE_tT11_0[84]*u+PRE_tT11_0[83,1]);
d_d_u = drescale * (PRE_tT11_0[84]+PRE_tT11_0[85]*v+PRE_tT11_0[84,1]);
d_d_v = drescale * (PRE_tT11_0[85]+PRE_tT11_0[86]*w+PRE_tT11_0[85,1]);
d_d_w = drescale * (PRE_tT11_0[86]+PRE_tT11_0[87]*u+PRE_tT11_0[86,1]);
d_d_u = drescale * (PRE_tT11_0[87]+PRE_tT11_0[88]*v+PRE_tT11_0[87,1]);
d_d_v = drescale * (PRE_tT11_0[88]+PRE_tT11_0[89]*w+PRE_tT11_0[88,1]);
d_d_w = drescale * (PRE_tT11_0[89]+PRE_tT11
```