计算机视觉

图像的全景拼接

图像全景拼接方案流程

1. 特征点检测
2. 创建特征点描述子
3. 特征点匹配
4. 几何变换估计
5. 坐标变换

线性几何变换

线性几何变换的定义：$T(x) = Hx$，其中 x 是 n 维坐标，H 是一个可逆矩阵

旋转变换

旋转变换表示平面上一点绕原点逆时针旋转：$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

在 $a = Ra'$ 中，"R 表达一个保持方向的旋转变换"等价于"R 为正交阵且行列式为 1"。

若 R 是一个行列式为 -1 的正交矩阵，则 R 表达的是一个平面旋转变换复合一个反射变换。

平面旋转变换的自由度为 1。

齐次坐标

齐次坐标用一个三维向量 $(x_1, x_2, x_3)^T$ 来表达二维平面上的点。

如果 $x_3 = 0$，说明这个点为一个无穷远点；如果 $x_3 \neq 0$，说明该点为一个正常点。

一个点的齐次坐标的形式不具有唯一性，$(x_1, x_2, x_3)^T$ 和 $k(x_1, x_2, x_3)^T$ 表示的是同一个点。但规范化齐次坐标 $(\frac{x_1}{x_3}, \frac{x_2}{x_3}, 1)^T$ 是唯一的。

正常坐标与齐次坐标的转换：$(x_1, x_2)^T \Leftrightarrow k(x_1, x_2, 1)^T$。

规范齐次坐标的旋转变换：$H = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$。

欧氏变换

欧氏变换由旋转和平移复合而成（不考虑反射变换）：$H = \begin{bmatrix} \cos\theta & -\sin\theta & t_x \\ \sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix}$。

平面欧氏变换有 3 个自由度。

相似变换

相似变换由欧氏变换复合一个缩放得来：$H = \begin{bmatrix} s\cos\theta & -s\sin\theta & t_x \\ s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix}$。

平面相似变换有 4 个自由度。

仿射变换

在相似变换的基础上对矩阵左上 2×2 的子矩阵 R 放宽限制为非奇异矩阵，得到仿射变换：$H = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$。

矩阵 A 的 4 个元素相互独立，带来了 4 个自由度：对 A 奇异值分解 $A = UDV^T$，U 和 D 为正交阵，相当于旋转变换，且角度相反；而 D 相当于缩放，两个方向上的缩放系数分别为 A 的两个奇异值。

矩阵 H 左上角的 2×2 子矩阵 R 的行列式 $\det(R)$ 决定了面积的变化以及是否保持方向。当 $\det(R) > 0$ 时，仿射变换保持方向；当 $\det(R) < 0$ 时，仿射变换改变方向。

平面仿射变换有 6 个自由度。

射影变换

仿射变换矩阵的第三行一定为 $(0,0,1)$。而若仅要求 H 为非奇异矩阵，就得到射影变换：$H=\begin{bmatrix}h_{11}&h_{12}&h_{13}\\h_{21}&h_{22}&h_{23}\\h_{31}&h_{32}&h_{33}\end{bmatrix}$。

射影变换可以定义在无穷远点上，而其他变换都只考虑正常点。

由于齐次坐标的不唯一性，射影变换有 8 个自由度。

特征点检测与匹配

哈里斯角点检测

对于图像上某点，以它为中心取一个窗口 W，窗口移动小量 $(\Delta x,\Delta y)$ 之后，新旧窗口的像素值差异为

$$s_W(\Delta x,\Delta y)=\sum_{(x_i,y_i)\in W}(f(x_i,y_i)-f(x_i+\Delta x,y_i+\Delta y))^2$$

进行一阶泰勒近似后得到

$$s_W(\Delta x,\Delta y)=(\Delta x,\Delta y)M(\Delta x,\Delta y)^T$$

其中

$$M=\begin{pmatrix}\sum_{(x_i,y_i)\in W}\left(\frac{\partial f}{\partial x}|_{(x_i,y_i)}\right)^2 & \sum_{(x_i,y_i)\in W}\left(\frac{\partial f}{\partial x}|_{(x_i,y_i)}\right)\left(\frac{\partial f}{\partial y}|_{(x_i,y_i)}\right)\\ \sum_{(x_i,y_i)\in W}\left(\frac{\partial f}{\partial x}|_{(x_i,y_i)}\right)\left(\frac{\partial f}{\partial y}|_{(x_i,y_i)}\right) & \sum_{(x_i,y_i)\in W}\left(\frac{\partial f}{\partial y}|_{(x_i,y_i)}\right)^2\end{pmatrix}$$

可以证明 M 为半正定矩阵，且大部分情况下为正定矩阵。

当 M 为正定矩阵时，方程 $s_W(\Delta x,\Delta y)=1$ 中 $(\Delta x,\Delta y)$ 的轨迹为一个椭圆，椭圆的半长轴长和半短轴长分别为 $\lambda_1^{-0.5}$ 和 $\lambda_2^{-0.5}$。因此越小越圆的椭圆越可能对应一个角点。

在实现时，特征值分解的计算代价较高，通常使用角点程度经验公式：$r(x)=det(M(x))-k(trace(M(x)))^2$。其中 k 为超参数，通常为 0.04 到 0.06。

选择角点不仅需要看 $r(x)$ 是否大于某个阈值，还需要进行非极大值抑制，以保证选出的角点是稀疏的。

哈里斯角点检测算法具有旋转不变性和整体光照不变性，但不具有尺度不变性，原因在于窗口大小是预先设定的，无法自动自适应。

特征描述子

对于图像上一点 x，以其为中心取一个 $s\times s$（提前设定好的值）的窗口，拉成一个列向量并进行单位化，得到的向量 d 称为块描述子，是最简单的特征描述子。

两个描述子之间的常用距离计算方式有 SSD、SAD、NCC

- 平方差之和：$SSD_{dist}(d_1,d_2)=\|d_1-d_2\|_2^2$
- 绝对差之和：$SAD_{dist}(d_1,d_2)=\|d_1-d_2\|_1^2$
- 规范化互相关距离与皮尔逊互相关系数有关：$NCC_{dist}(d_1,d_2)=1-\frac{1}{n}\frac{(d_1-\mu(d_1))\cdot(d_2-\mu(d_2))}{std(d_1)std(d_2)}\in[0,2]$
- 另一种规范化互相关距离定义：$NCC_{dist}(d_1,d_2)=\arccos\left(\frac{1}{n}\frac{(d_1-\mu(d_1))\cdot(d_2-\mu(d_2))}{std(d_1)std(d_2)}\right)\in[0,\pi]$

SIFT 特征点检测

流程：构建高斯尺度空间 > 构建 DoG 尺度空间 > 检测 DoG 极值点 > 粗略极值点位置精化与 DoG 值估计（过滤非特征点） > 计算海森矩阵（过滤边缘点） > 得到特征点的空间位置和特征尺度

SIFT 中的特征点（斑点）不仅具有位置特征，还具有大小特征。

检测斑点使用尺度归一化 LoG 算子：

$$\sigma^2\nabla^2 g=\frac{x^2+y^2-2\sigma^2}{2\pi\sigma^4}e^{-\frac{x^2+y^2}{2\sigma^2}}$$

对尺度大小参数 σ 进行取值可以得到一系列不同尺度大小的尺度归一化 LoG 算子，用这些算子与图像进行卷积并堆叠可以得到尺度空间。图像上每一点 x 在尺度空间中都有一组响应值 $\{r_i\}$，使得 r 取得极值的唯一 σ 称为斑点的特征尺度。

在实现时，为了简洁高效，可以用 DoG 算子浸提替代尺度归一化 LoG 算子：

$$DoG(\sigma)=g(x,y;k\sigma)-g(x,y;\sigma)$$

这样由于卷积运算性质，卷积输出 $DoG(\sigma)*I$ 就是 $g(x,y;k\sigma)*I-g(x,y;\sigma)*I$。

当 $k\to 1$ 时，$DoG(\sigma)\approx(k-1)\sigma^2\nabla^2 g(x,y;\sigma)$。但即时在 k 较大时，使用 DoG 算子的性能也不会受到明显影响。

在尺度空间中，如果一点的响应值比 26 个邻近值都大或都小，则得到一个候选特征点。特征点取得的极值的是极大值还是极小值与斑点的亮暗结构有关。

该算法的斑点特征尺度会随着图像的缩放而等比变化，具有尺度协变性。同时，该算法具有旋转不变性、光照不变性。

SIFT 特征描述子

SIFT 特征描述子的构建在高斯尺度空间层 g_l 上进行，在 g_l 上与特征尺度 σ 对应的高斯标准差记为 σ_1。

为了使描述子具有旋转不变性，需要确定图像邻域的主方向，并根据主方向进行方向归一化：在 DoG 空间尺度层上划定一个 $9\sigma_1\times 9\sigma_1$ 的区域 P。对于 P 上的每一点计算出其梯度，构建梯度直方图，并在直方图中选择主方向，然后进行旋转。

特征描述子的构建：将邻域划分为 4×4 个子区域，在每个子区域中构建一个 8 维的梯度方向直方图，最后将所有直方图连接起来形成一个 128 维的向量。实际实现时，每个点的贡献需要进行基于到中心距离的高斯加权。

一张图像产生的所有信息：特征点坐标、特征尺度、斑点主方向、特征描述子

射影矩阵估计

矩阵微分结论

1. $\frac{d\mathbf{a}^T\mathbf{x}}{d\mathbf{x}}=\frac{d\mathbf{x}^T\mathbf{a}}{d\mathbf{x}}=\mathbf{a}$
2. $\frac{d\mathbf{x}^T\mathbf{x} }{d\mathbf{x} }=2\mathbf{x}$$
3. $\frac{d\mathbf{y}^T(\mathbf{x})}{d\mathbf{x}}=\left(\frac{d\mathbf{y}(\mathbf{x})}{d\mathbf{x}^T}\right)^T$
4. $\frac{d\mathbf{A}\mathbf{x}}{d\mathbf{x}^T}=\mathbf{A}$
5. $\frac{d\mathbf{x}^T\mathbf{A}^T}{d\mathbf{x}}=\mathbf{A}^T$

6. $\frac{d\mathbf{x}^T\mathbf{A}\mathbf{x}}{d\mathbf{x}} = (\mathbf{A}+\mathbf{A}^T)\mathbf{x}$
7. $\frac{d\mathbf{a}^T\mathbf{X}\mathbf{b}}{d\mathbf{X}} = \mathbf{a}\mathbf{b}^T$
8. $\frac{d\mathbf{a}^T\mathbf{X}^T\mathbf{b}}{d\mathbf{X}} = \mathbf{b}\mathbf{a}^T$
9. $\frac{d(tr\mathbf{X}\mathbf{B})}{d\mathbf{X}} = \mathbf{B}^T$
10. $\frac{d|\mathbf{X}|}{d\mathbf{X}} = |\mathbf{X}|(\mathbf{X}^{-1})^T$

拉格朗日乘子

用于寻找 $f(x)$ 在等式约束 $g_i(x)=0$ 下的所有极值点。

定义 $F(x;\lambda_1,\cdots\lambda_m)=f(x)+\sum_i \lambda_i g_i(x)$，则 $f(x)$ 的所有极值点一定都是 F 的驻点。

只需要求解 $n+m$ 个方程并从解中寻找答案。

线性最小二乘

非齐次线性方程组：求 $\arg\min_x \|Ax-b\|_2^2$。由于 $\|Ax-b\|_2^2$ 是 x 的凸函数，驻点 $x_s=(A^TA)^{-1}A^Tb$ 即为全局最小值点。

齐次线性方程组：求 $\arg\min_x \|Ax\|_2^2$，其中 $\|x\|_2=1$。令 $L(x,\lambda)=\|Ax\|_2^2+\lambda(1-\|x\|_2^2)$，驻点满足 $A^TAx_0=\lambda_0 x_0, x_0^Tx_0=1$，即驻点为 A^TA 的所有特征值。其中显然最小值点为 $\min\{\lambda_i\}$。

RANSAC 单应估计

一个射影变换的单应矩阵只需要 4 个方程就能确定，而实际的特征点对数远大于 4。

由于存在异常点（离群值），直接使用最小二乘会受到较大影响。

RANSAC：每一次迭代过程中随机选择所有样本点的一个子集，并进行模型拟合，统计该模型下误差值小于某个阈值的样本点，称为当前模型的”一致集”。重复这一过程，直到次数到达上限或某次一致集的大小大于某个阈值。

单目测量

射影几何

向量操作

向量点乘：$a \cdot b = |a||b| \cos\theta = \sum a_i b_i$

向量叉乘：$a \times b = \left(\begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}, \begin{vmatrix} z_1 & x_1 \\ z_2 & x_2 \end{vmatrix}, \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \right)$

$|a \times b| = |a||b| \sin\theta$

向量混合积：$(a, b, c) = (a \times b) \cdot c$，平行六面体体积

射影几何基础

齐次坐标的几何意义：欧氏空间中过原点的直线与平面 $z = z_0$ 的交点

针对无穷远点的定义扩充

- 平面 xOy 上过 O 的任一直线都与平面 $z = 1$ 交于一个无穷远点。
- 所有无穷远点构成一条无穷远直线，即平面 xOy 与平面 $z = 1$ 的交线。
- $z = 0$ 和无穷远直线构成射影平面。
- 在三维射影空间中所有无穷远点构成一个平面，可以表示三维下的所有方向。

射影平面上的点和面

- 射影平面上的点和面具有对偶关系
- 两点确定一条直线：$l = x \times x'$
- 两线确定一个交点：$x = l \times l'$
- 无穷远直线的坐标：$(0, 0, 1)^T$，由两个不同的无穷远点确定
- 判断点在直线上：$x \cdot l = 0$

针孔相机成像模型

几种坐标系

世界坐标系（3D）：真实世界的坐标系

相机坐标系（3D）：以相机光心为原点，相机朝向为轴的坐标系

成像平面坐标系（2D）：成像平面是一个到相机光心距离为 f 且垂直于 z 轴的平面

归一化成像平面坐标系（2D）：$f=1$ 的成像平面坐标系

像素坐标系（2D）：图像坐标系

世界坐标系到相机坐标系

世界坐标系的坐标可以通过旋转和平移转换到相机坐标系。设世界坐标系下一点 $(x_w, y_w, z_w)^T$，则

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = R \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + t$$

即

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = [R, t] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

相机坐标系到成像平面坐标系

根据针孔相机模型，将相机坐标系的三维坐标投影在成像平面上：$x = f\frac{x_c}{z_c}, y = f\frac{y_c}{z_c}$

用齐次坐标来表达：

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{z_c} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}$$

转换到归一化成像平面坐标系，只需要令 $f=1$：

$$\begin{bmatrix} x_n \\ y_n \\ 1 \end{bmatrix} = \frac{1}{z_c} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}$$

成像平面坐标系到像素坐标系

成像平面坐标系与像素坐标系的不同在于**原点位置**和**度量单位**，同时由于感光器件制造工艺不完美，像素坐标系的两个轴可能不垂直。假设 v 轴与 y 轴之间的夹角为 α：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & \frac{\tan\alpha}{dx} & c_x \\ 0 & \frac{1}{dy} & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

世界坐标系到像素坐标系

综合上述三个过程，联立所有等式，得到：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z_c} \begin{bmatrix} \frac{f}{dx} & \frac{f\tan\alpha}{dx} & c_x \\ 0 & \frac{f}{dy} & c_y \\ 0 & 0 & 1 \end{bmatrix} [R, t] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

其中 $[R, t]$ 与相机相对于世界坐标系的位姿有关，称为外参数；$K = \begin{bmatrix} \frac{f}{dx} & \frac{f\tan\alpha}{dx} & c_x \\ 0 & \frac{f}{dy} & c_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$ 只与相机的物理属性有关，是相机的内参数，K 称为相机的内参矩阵

本课程不考虑参数 s，因此 $K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$

需要注意整个过程中所有变量的单位

镜头畸变

镜头畸变属于相机内参，但不在内参矩阵中

畸变的类型

- 径向畸变：镜头越小畸变越大
  - 正径向偏差：桶形畸变
  - 负径向偏差：枕形畸变
- 切向畸变：镜头和成像器件平面不平行导致的

两种类型的畸变都建模在归一化成像平面坐标系，其中径向畸变由参数 k_1, k_2, k_3 决定，切向畸变由参数 ρ_1, ρ_2 确定。这个过程**不能**用矩阵乘法表示，记为 $\mathcal{D}$。

> 鱼眼相机的畸变需要另一种模型，由参数 k_1, k_2, k_3, k_4 确定。

考虑畸变后的坐标变换等式为

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f}{dx} & 0 & c_x \\ 0 & \frac{f}{dy} & c_y \\ 0 & 0 & 1 \end{bmatrix} \mathcal{D} \left\{ \frac{1}{z_c} [R, t] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \right\}$$

每个从世界坐标到像素坐标的对应关系可以导出 2 个等式。

相机内参标定流程

旋转矩阵的处理

三维空间中的旋转可以用一个轴和一个角度 $d = n\theta$ 表示。轴角表示法和旋转矩阵可以通过 Rodrigues 公式对应起来，其中 n 为旋转矩阵对应特征值 1 的特征向量。记对应关系为 $\mathcal{R}(d_i) = R_i$。

问题框架

假设有 M 张标定板图像，每张图像有 N 个交叉点，则内参标定问题即如下优化问题：

$$\Theta^* = \arg\min_{\Theta} \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{1}{2} \left\| K \cdot \mathcal{D}\left\{ \frac{1}{z_{c_{ij}}} \left[\mathcal{R}(d_i), t_i \right] P_j \right\} - u_{ij} \right\|_2^2$$

其中 $\Theta = \{f_x, f_y, c_x, c_y, k_1, k_2, k_3, \rho_1, \rho_2, \{R_i\}_{i=1}^M, \{t_i\}_{i=1}^M\}$ 表示要优化的参数（包括外参，共 $9+6M$ 个），P_j 表示交叉点的世界坐标，u_{ij} 表示交叉点的像素坐标。

参数的初始估计

镜头畸变参数 $k_1, k_2, k_3, \rho_1, \rho_2$：可以全部初始化为 0，相当于不存在畸变

主点坐标 c_x, c_y：其大致处于图像的中心，可以初始化为 $c_x = \frac{width}{2}, c_y = \frac{height}{2}$

内参 f_x, f_y：先估计标定板到成像平面的射影变换矩阵，再考虑标定板上四条特殊直线 $y=0, x=0, y=x, y=-x$ 的消失点变换到像素坐标系的 v_1, v_2, v_3, v_4，从而得到非齐次线性方程组，最后用最小二乘法解出 f_x, f_y

1. 消失点：物理空间中一条直线的无穷远点在成像平面上的对应位置
2. 假设光心为 O, v 为成像平面上一点，则 Ov 平行于以 v 为消失点的直线
3. 假设光心为 O, u 为成像平面像素坐标系下的齐次坐标，则在相机坐标系下，射线 Ou 的方向可以表示为 $d = K^{-1}u$。
4. 假设光心为 O，x_1、x_2 为成像平面像素坐标系下的齐次坐标，则在相机坐标系下，射线 Ox_1, Ox_2 的夹角为 $\theta = \arccos \frac{x_1^T(K^{-T}K^{-1})x_2}{\sqrt{x_1^T(K^{-T}K^{-1})x_1}\sqrt{x_2^T(K^{-T}K^{-1})x_2}}$
5. 假设光心为 O，两条直线 l_1, l_2 的消失点分别为 v_1, v_2，Ov_1, Ov_2 的夹角为 θ，则 l_1, l_2 的两个夹角为 $\theta, \pi-\theta$
6. 物理空间相互垂直的两条直线的消失点 v_1, v_2 满足 $v_1^T(K^{-T}K^{-1})v_2 = 0$
7. 简化计算 $K = \begin{bmatrix} 1 & 0 & c_x \\ 0 & 1 & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} = PQ = P^{-T}(Q^{-T}Q^{-1})P^{-1}$

每张图像的外参 d_i, t_i：由成像模型，在不考虑畸变情况下从标定板坐标系到成像平面坐标系的射影变换和 $[R, t]$ 是等价的；同时标定板上点的第三维坐标一定为 0，可以忽略旋转矩阵的第三列。故 $H = \lambda[r_1, r_2, t]$。由于 $\|r_1\| = \|r_2\| = 1$，$\lambda = \|h_1\| = \|h_2\|$。因此已知 H 就可以求得 r_1, r_2, t，再求出 $r_3 = r_1 \times r_2$（叉乘顺序不可交换）。

非线性优化方法

非线性优化问题没有定势解，只能通过迭代方法求局部最优解。

连续可导函数的泰勒展开：$F(x+h) = F(x) + h^T F'(x) + \frac{1}{2} h^T F''(x) h + O(\|h\|^2)$，其中 $F'(x)$ 为梯度，$F''(x)$ 为海森矩阵。

两个定理（充分不必要条件）：

- x^* 是局部极小值点 $\Rightarrow$ x^* 是驻点（一阶导为 0）
- x_s 是驻点且 $F''(x_s)$ 正定 $\Rightarrow$ x_s 是局部极小值点

下降方法

- 二阶段法：先求方向再求距离
  - 求方向：梯度下降法、牛顿法、混合法（initial SD, final N）
  - 求距离：线搜索法
- 一阶段法：方向和距离同时确定：置信域法、阻尼法

牛顿法

- 由于目标点 x^* 满足 $F'(x^*) = 0$，我们需要尽量找到一个满足 $F'(x+h) = 0$ 的 h
- $F'(x+h) \approx F'(x) + F''(x)h \Rightarrow F''(x)h_n = -F'(x)$
- $x = x + h_n \Rightarrow$ 一阶段 / $x = x + \alpha h_n \Rightarrow$ 二阶段

置信域法

- 在置信的邻域内极小化近似的二次模型，得到近似极小点
- 根据拟合程度调整置信域半径 Δ
- 可以使用增益率 $\rho = \frac{F(x) - F(x+h)}{L(0) - L(h)}$：增益率较小时减小 Δ，增益率较大时增大 Δ

阻尼法

- 类似于置信域法，但不引入置信域的概念
- 由于 h 越大估计越有可能不准确，因此设置一个阻尼项，防止自变量走远
- $h = \arg\min_h \{L(h) + \frac{1}{2}\mu h^T h\}$
- 同样根据增益率调整阻尼系数 μ：增益率较小时增大 μ，增益率较大时减小 μ
- 求解过程（要考）：
  - $\phi_\mu(h) = L(h) + \frac{1}{2}\mu h^T h$
  - $\phi'_\mu(h) = \frac{d\left(L(h) + \frac{1}{2}\mu h^T h\right)}{dh} = \frac{d\left(F(x) + h^T c + \frac{1}{2} h^T B h\right)}{dh} = c + \frac{1}{2}(B + B^T)h + \mu h = c + Bh + \mu h = 0$

- $\Rightarrow h_{dm} = -(B+\mu I)^{-1}c$，其中 B 为对称阵
- 当 c, B 为 F 的梯度和海森矩阵时称为阻尼牛顿法。μ 很大时近似于梯度下降，μ 很小时近似于牛顿法

非线性最小二乘问题

非线性最小二乘问题的形式

$$x^* = \arg\min_x \{F(x)\} = \arg\min_x \{\frac{1}{2}\|f(x)\|^2\} = \arg\min_x \{\frac{1}{2}f(x)^T f(x)\}$$

- 泰勒展开 $f(x+h) = f(x) + J(x)h + O(\|h\|)$
- $F'(x) = (J(x))^T f(x), F''(x) = \sum_{i=1}^m f_i(x) f_i''(x) + (J(x))^T J(x)$

高斯牛顿法

- $f(x+h) \approx f(x) + J(x)h$，假设 J 列满秩
- $F(x+h) \approx L(h) \equiv \frac{1}{2}\|f(x+h)\| = \frac{1}{2}f^T f + h^T J^T f + \frac{1}{2}h^T J^T J h$
- $\frac{dL(h)}{dh} = J^T f + \frac{1}{2}(J^T J + J^T J)h = 0 \Rightarrow h_{gn} = -(J^T J)^{-1} J^T f$
- 相当于无阻尼的阻尼法或置信域无限的置信域法
- 类似牛顿法，可以是一阶也可以是二阶
- 迭代中的每一步都要求 J 列满秩，这可以保证 $J^T J$ 正定

LM 方法

- 相当于阻尼法+高斯牛顿法
- $\frac{d\left(L(h) + \frac{1}{2}\mu h^T h\right)}{dh} = 0 \Rightarrow h_{lm} = -(J^T J + \mu I)^{-1} J^T f$，其中 $(J^T J + \mu I)$ 正定
- μ 的更新与阻尼法相同
- 停止条件：一阶导极小、h 极小，循环轮数到达上限

求解框架问题

LM 法求解原问题

- 所有现代实现中都是用 LM 法解该问题
- 需要 $J(x)$，因此核心问题是计算 $\frac{dp_{ij}}{d\Theta^T}$

- 可以利用链式法则对各个参数求导（求导过程）

鸟瞰视图

- 鸟瞰坐标系和世界坐标系间是相似变换关系（假设世界坐标系原点在鸟瞰图的中心）：

$$\begin{pmatrix} x_W \\ y_W \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{H}{W} & 0 & -\frac{HN}{2M} \\ 0 & -\frac{H}{M} & \frac{H}{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_B \\ y_B \\ 1 \end{pmatrix}$$

- 求无畸变图像：对于无畸变图像中的每一个像素，将坐标经过包括畸变的变换映射到实际图像上，利用插值法在实际图像上求得对应的颜色

机器学习和 CNN 基础