Homework 3

林宇辰 2200013211

November 2023

1 Fundamental matrix estimation with ground truth matches

首先来阐述本次作业中反复用到的求解超定方程最小二乘解的数学过程,也就是求解 $min\|Ax\|$。由于我们处理的方程都是在缩放的意义下是成立的，我们不妨让待求向量的长度为 1，这可以保证数值计算的稳定性，也为使用奇异值分解求解提供了方便。因此有第一条假设

$$\|x\| = 1 \tag{1}$$

然后用奇异值分解。

$$A = UDV^T \tag{2}$$

$$U^TU = I \tag{3}$$

$$VV^T = I \tag{4}$$

$$\|Ax\| = \|UDV^Tx\| = \|DV^Tx\| = \|Dy\| \quad \text{这是因为正交矩阵 } U \text{ 是保距的} \tag{5}$$

接着取 $y = (0, ..., 1)$，这可以保证 $\|Ax\| = \sigma_n$，也就是最小的特征值。所以，在之后的计算中，我们要求解的 x 也就是 A^TA 特征值最小的特征向量。

1.1 Fit Fundamental

对于两个匹配的二维点，它们对映同一个三维点。通过这个关系建立方程。

$$x' \cong Rx + t \tag{1}$$

$$x'^TFx = 0 \tag{2}$$

$$(x'x, x'y, x', y'x, y'y, y', x, y, 1)\begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{pmatrix} = 0 \quad \text{针对一对匹配点的情况} \tag{3}$$

然后用最小二乘法来求出 Fundamental 矩阵。

1.2 Normalize

为了数值稳定性，做 Normalize。首先对所有点计算出平均点，然后进行平移，将这个平均点当作原点。接着计算所有点到原点的平均距离，将其“归一化”到 2 像素。
本任务的可视化效果如图 1和图 2所示。可以发现做过归一化的匹配效果更精确。

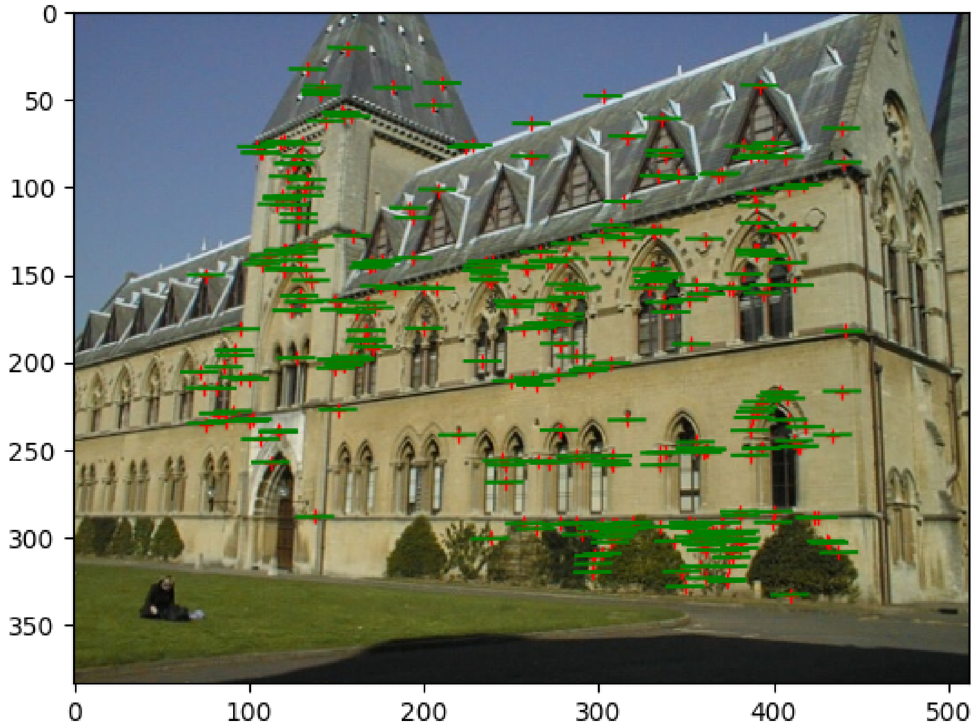

图 1: Normalized

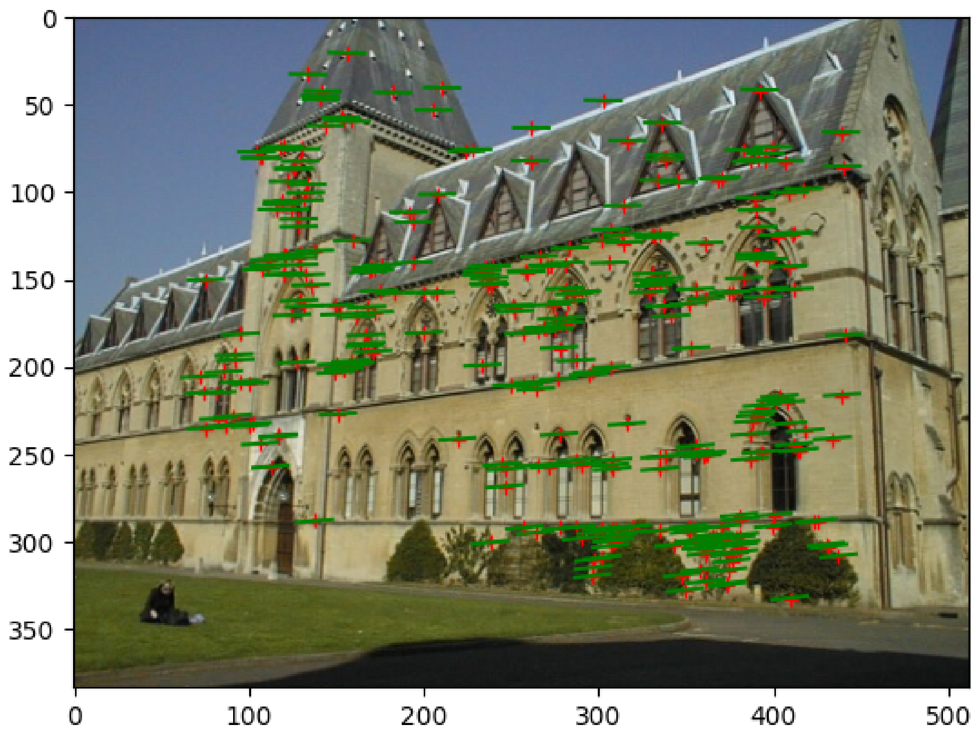

图 2: Unnormalized

2 Camera calibration

通过相机参数矩阵，世界坐标转换为相机坐标，通过这个关系建立方程。

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \cong \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{1}$$

$$\begin{bmatrix} X & Y & Z & 1 & 0 & 0 & 0 & 0 & -xX & -xY & -xZ & -x \\ 0 & 0 & 0 & 0 & X & Y & Z & 1 & -yX & -yY & -yZ & -y \end{bmatrix} \begin{pmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ p_{21} \\ p_{22} \\ p_{23} \\ p_{24} \\ p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{pmatrix} = 0 \tag{2}$$

然后运用最小二乘法就可以求出相机参数。

3 Calculate the camera matrices

对相机参数的前三列进行 RQ 分解，再将 R 变成主对角线全是 1 的上三角矩阵得到 K，同时调整 Q 得到 R。接着带入 $P = K[R|T]$ 求出 T。计算结果如下。

$$K_1 = \begin{bmatrix} 780.5675025807848 & 1.9984611557189735 & 545.6704454288864 \\ 0.0 & 779.9910940480137 & 384.1595910883893 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

$$\mathrm{R}_1 = \begin{bmatrix} 2.7981469835444743e-06 & -1.7321366886044586e-06 & -8.819862222033054e-08 \\ -4.334851081347211e-07 & -5.36355798721175e-07 & -3.219024101291231e-06 \\ 1.6793353345882447e-06 & 2.747676836103324e-06 & -6.839648271913074e-07 \end{bmatrix}$$

$$\mathrm{T}_1 = \begin{bmatrix} -0.00032618921609340663 & 0.0003927456461877954 & -0.0013288292767817086 \end{bmatrix}$$

$$\mathrm{K}_2 = \begin{bmatrix} 767.0152229883663 & 8.374810233649852 & 536.2061957476232 \\ 0.0 & 772.0225338536667 & 390.7126755637734 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

$$\mathrm{R}_2 = \begin{bmatrix} 3.7375657674989757e-06 & -7.824244178063841e-06 & -3.067701429231474e-07 \\ -1.8463520020206822e-06 & -5.502904900592349e-07 & -8.45993306156259e-06 \\ 7.60946050408235e-06 & 3.709539885689283e-06 & -1.9020324440519715e-06 \end{bmatrix}$$

$$\mathrm{T}_2 = \begin{bmatrix} 0.00127996094446251 & 0.000986035233528928 & -0.003388077115992867 \end{bmatrix}$$

$$\mathrm{K}_3 = \begin{bmatrix} -579.7909746178564 & 1.1178215085297194e-06 & 256.9915521897448 \\ 0.0 & -539.7111472878375 & 204.3175584777602 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

$$\mathrm{R}_3 = \begin{bmatrix} -0.009661932737103463 & -0.4413551418623896 & -0.8972804977569302 \\ -0.9801797201052763 & -0.17338540883523443 & 0.09583956461134308 \\ -0.19787462999999997 & 0.8804221399999999 & -0.43093212000000003 \end{bmatrix}$$

$$\mathrm{T}_3 = \begin{bmatrix} 6.485240103433924 & 1.7129909294202956 & 28.032556 \end{bmatrix}$$

$$\mathrm{K}_4 = \begin{bmatrix} -547.4691055493174 & 7.308425760695429e-06 & 258.4300935764808 \\ 0.0 & -512.9335851907326 & 204.9855420618667 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

$$\mathrm{R}_4 = \begin{bmatrix} 0.018613836563949968 & -0.6762407660646936 & -0.7364454877805992 \\ -0.9813810053615476 & -0.15319027935757834 & 0.11586227038040892 \\ -0.19116708 & 0.72057697 & -0.6665012999999999 \end{bmatrix}$$

$$\mathrm{T}_4 = \begin{bmatrix} 6.70661750829618 & 1.6948260755287177 & 28.015392 \end{bmatrix}$$

4 Triangulation

由于噪声的干扰，两条投影不一定有交点，这个任务就要求最合适的三维坐标。

$$x_1 \cong P_1 X \tag{1}$$

$$x_2 \cong P_2 X \tag{2}$$

$$\begin{pmatrix} [x_{1_\times}]P_1 \\ [x_{2_\times}]P_2 \end{pmatrix} X = 0 \tag{3}$$

$$[x_{1_\times}] = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \tag{4}$$

运用最小二乘法解出 X 即可。

5 Fundamental matrix estimation without ground-truth matches

首先用 SIFT 提取特征，然后进行匹配。最后用 RANSAC 进行随机采样。对于代价 $x'^T Fx$ 小于阈值的点，把他们视作内点。在我的超参数（见代码）选择下，可以匹配到 140 个左右的内点，平均误差在 0.06 左右。可视化效果如图 3所示。比较激进点的做法可以匹配到 500 多个内点，平均误差在 1.1 左右，如图 4所示。

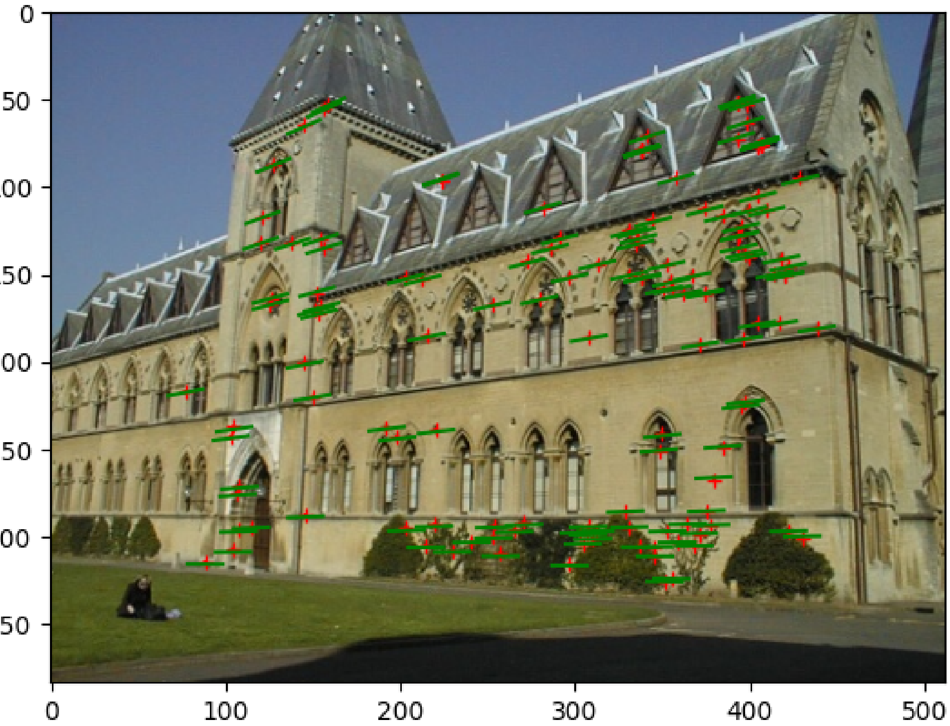

图 3: Inliers=140

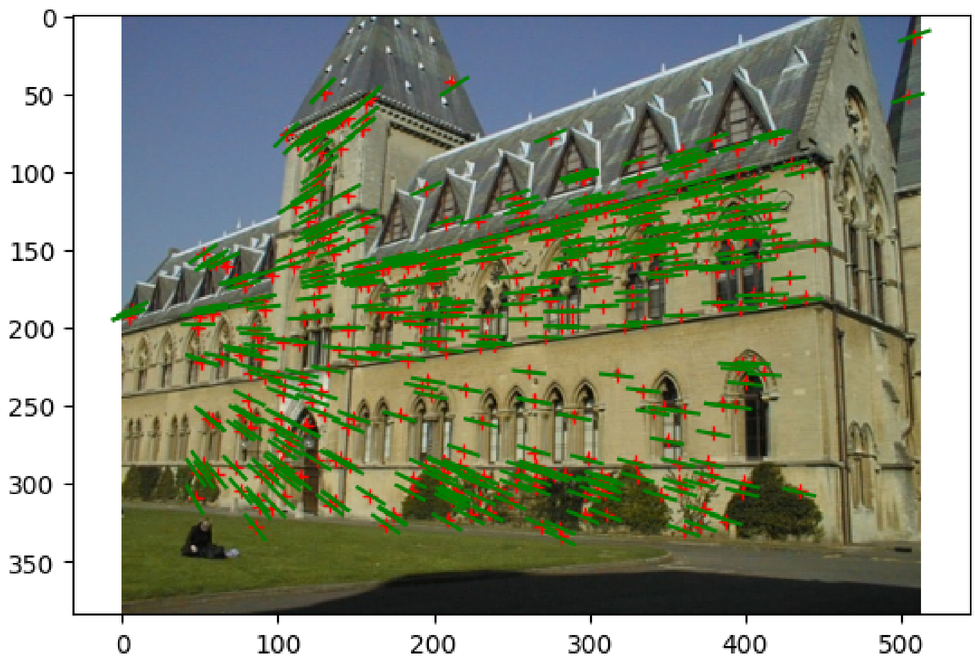

图 4: Inliers=500