Solution to the 4th Homework

Yimeng Zhu

November 30, 2019

Contents

1 图像去畸变 (3 分，约 1 小时)

现实生活中的图像总存在畸变。原则上来说，针孔透视相机应该将三维世界中的直线投影成直线，但是当我们使用广角和鱼眼镜头时，由于畸变的原因，直线在图像里看起来是扭曲的。本次作业，你将尝试如何对一张图像去畸变，得到畸变前的图像。

图 1: 测试图像

图1 是本次习题的测试图像 (code/test.png)，来自 EuRoC 数据集 [1]。可以明显看到实际的柱子、箱子的直线边缘在图像中被扭曲成了曲线。这就是由相机畸变造成的。根据我们在课上的介绍，畸变前后的坐标变换为:

$$\begin{cases} x_{distorted} = x(1 + k_1 r^2 + k_2 r^4) + 2p_1 xy + p_2(r^2 + 2x^2) \\ y_{distorted} = y(1 + k_1 r^2 + k_2 r^4) + p_1(r^2 + 2y^2) + 2p_2 xy \end{cases} \tag{1}$$

其中 x, y 为去畸变后的坐标，xdistorted, ydistroted 为去畸变前的坐标。现给定参数:

$$k_1 = -0.28340811, k_2 = 0.07395907, p_1 = 0.00019359, p_2 = 1.76187114e{-}05.$$

以及相机内参

$$f_x = 458.654, f_y = 457.296, c_x = 367.215, c_y = 248.375.$$

请根据 undistort_image.cpp 文件中内容，完成对该图像的去畸变操作。

注: 本题不要使用 OpenCV 自带的去畸变函数，你需要自己理解去畸变的过程。我给你准备的程序中已经有了基本的提示。作为验证，去畸变后图像如图2所示。如你所见，直线应该是直的。

图 2: 验证图像

Solution: see folder 1

2 双目视差的使用 (2 分，约 1 小时)

双目相机的一大好处是可以通过左右目的视差来恢复深度。课程中我们介绍了由视差计算深度的过程。本题，你需要根据视差计算深度，进而生成点云数据。本题的数据来自 Kitti 数据集 [2]。

Kitti 中的相机部分使用了一个双目模型。双目采集到左图和右图，然后我们可以通过左右视图恢复出深度。经典双目恢复深度的算法有 BM(Block Matching), SGBM(Semi-Global Block Matching)[[3], [4]] 等，但本题不探讨立体视觉内容 (那是一个大问题)。我们假设双目计算的视差已经给定，请你根据双目模型，画出图像对应的点云，并显示到 Pangolin 中。

本题给定的左右图见 code/left.png 和 code/right.png，视差图亦给定，见 code/right.png。双目的参数如下:

$$f_x = 718.856, f_y = 718.856, c_x = 607.1928, c_y = 185.2157$$

且双目左右间距 (即基线) 为:

$$d = 0.573m.$$

请根据以上参数，计算相机数据对应的点云，并显示到 Pangolin 中。程序请参考 code/disparity.cpp 文件。

图 3: 双目图像的左图、右图与视差

作为验证，生成点云应如图4所示。

图 4: 双目生成点云结果

Solution: see folder 2

3 矩阵运算微分 (2 分，约 1.5 小时)

在优化中经常会遇到矩阵微分的问题。例如，当自变量为向量 $\mathbf{x}$，求标量函数 $\mu(\mathbf{x})$ 对 $\mathbf{x}$ 的导数时，即为矩阵微分。通常线性代数教材不会深入探讨此事，这往往是矩阵论的内容。我在 ppt/目录下为你准备了一份清华研究生课的矩阵论课件 (仅矩阵微分部分)。阅读此 ppt，回答下列问题:

设变量为 $\mathbf{x} \in \mathbb{R}^N$，那么:

1. 矩阵 $\mathbf{A} \in \mathbb{R}^{N\times N}$，那么 $d(\mathbf{A}\mathbf{x})/d\mathbf{x}$ 是什么?
2. 矩阵 $\mathbf{A} \in \mathbb{R}^{N\times N}$，那么 $d(\mathbf{x}^T\mathbf{A}\mathbf{x})/d\mathbf{x}$ 是什么?
3. 证明:
$$\mathbf{x}^T\mathbf{A}\mathbf{x} = tr(\mathbf{A}\mathbf{x}\mathbf{x}^T). \tag{2}$$

Solution:

1. let $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, then $\mathbf{A}\mathbf{x}$ can be interpreted as:

$$\mathbf{A}\mathbf{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{bmatrix}$$

According to Definition of $d(\mathbf{A}\mathbf{x})/d\mathbf{x}$ in slides, the previous equation one can derive:

$$\begin{aligned} d(\mathbf{A}\mathbf{x})/d\mathbf{x} &= \begin{bmatrix} d(a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n)/dx_1 & \dots & d(a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n)/dx_1 \\ d(a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n)/dx_2 & \dots & d(a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n)/dx_2 \\ \vdots & \ddots & \vdots \\ d(a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n)/dx_n & \dots & d(a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n)/dx_n \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix} \\ &= \mathbf{A}^T \end{aligned}$$

Thus, $d(\mathbf{A}\mathbf{x})/d\mathbf{x} = \mathbf{A}^T$

2. let $\mathbf{A}$ and $\mathbf{x}$ be defined as previous, one can multiply out:

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$$

Further, one can derive:

$$d(\mathbf{x}^T\mathbf{A}\mathbf{x})/dx_k = d(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j)/dx_k = \sum_{i=1}^{n} a_{ik}x_i + \sum_{j=1}^{n} a_{kj}x_j = \sum_{i=1}^{n}(a_{ik} + a_{ki})x_i$$

Thus:

$$d(\mathbf{x}^T\mathbf{A}\mathbf{x})/dx = \begin{bmatrix} d(\sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j)/dx_1 \\ d(\sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j)/dx_2 \\ \vdots \\ d(\sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j)/dx_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n (a_{i1}+a_{1i})x_i \\ \sum_{i=1}^n (a_{i2}+a_{2i})x_i \\ \vdots \\ \sum_{i=1}^n (a_{in}+a_{ni})x_i \end{bmatrix} = (\mathbf{A}+\mathbf{A}^T)\mathbf{x}$$

3. From previous proof we know

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j$$

Further we have

$$\begin{aligned}\mathbf{A}\mathbf{x}\mathbf{x}^T = (\mathbf{A}\mathbf{x})\mathbf{x}^T &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{bmatrix} [x_1, x_2, \ldots, x_n] \\ &= \begin{bmatrix} (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)x_1 & \cdots \\ \vdots & \vdots \\ \cdots & (a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n)x_n \end{bmatrix}\end{aligned}$$

Therefor,

$$\begin{aligned} tr(\mathbf{A}\mathbf{x}\mathbf{x}^T) &= (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)x_1 + \cdots + (a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n)x_n \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j \\ &= \mathbf{x}^T\mathbf{A}\mathbf{x} \end{aligned}$$

Thus,

$$tr(\mathbf{A}\mathbf{x}\mathbf{x}^T) = \mathbf{x}^T\mathbf{A}\mathbf{x}$$

4　高斯牛顿法的曲线拟合实验 (3 分，约 2 小时)

我们在课上演示了用 Ceres 和 g2o 进行曲线拟合的实验，可以看到优化框架给我们带来了诸多便利。本题中你需要自己实现一遍高斯牛顿的迭代过程，求解曲线的参数。我们将原题复述如下。设有曲线满足以下方程:

$$y = \exp(ax^2 + bx + c) + w \tag{3}$$

其中 a, b, c 为曲线参数，w 为噪声。现有 N 个数据点 (x, y)，希望通过此 N 个点来拟合 a, b, c。实验中取 $N = 100$。

那么，定义误差为 $e_i = y_i - \exp(ax_i^2 + bx_i + c)$，于是 (a, b, c) 的最优解可通过解以下最小二乘获得:

$$\min_{a,b,c} \frac{1}{2} \sum_{i=1}^{N} ||y_i - \exp(ax_i^2 + bx_i + c)||^2. \tag{4}$$

现在请你书写 Gauss-Newton 的程序以解决此问题。程序框架见 code/gaussnewton.cpp，请填写程序内容以完成作业。作为验证，按照此程序的设定，估计得到的 a, b, c 应为:

$$a = 0.890912, b = 2.1719, c = 0.943629.$$

这和书中的结果是吻合的。

Solution: see folder 4

5 * 批量最大似然估计 (2 分，约 2 小时)

考虑离散时间系统:

$$\begin{aligned} x_k &= x_{k-1} + v_k + w_k, \quad & w \sim \mathcal{N}(0, Q) \\ y_k &= x_k + n_k, \quad & n_k \sim \mathcal{N}(0, R) \end{aligned}$$

这可以表达一辆沿 x 轴前进或后退的汽车。第一个公式为运动方程，v_k 为输入，w_k 为噪声; 第二个公式为观测方程，y_k 为路标点。取时间 $k = 1, ..., 3$，现希望根据已有的 v, y 进行状态估计。设初始状态 x_0 已知。

请根据本题题设，推导批量 (batch) 最大似然估计。首先，令批量状态变量为 $\mathbf{x} = [x_0, x_1, x_2, x_3]^T$，令批量观测为 $\mathbf{z} = [v_1, v_2, v_3, y_1, y_2, y_3]^T$，那么:

1. 可以定义矩阵 $\mathbf{H}$，使得批量误差为 $\mathbf{e} = \mathbf{z} - \mathbf{H}\mathbf{x}$。请给出此处 $\mathbf{H}$ 的具体形式。
2. 据上问，最大似然估计可转换为最小二乘问题:

$$\mathbf{x}^\star = \arg\min \frac{1}{2}(\mathbf{z} - \mathbf{H}\mathbf{x})^T \mathbf{W}^{-1} (\mathbf{z} - \mathbf{H}\mathbf{x}), \tag{5}$$

其中 $\mathbf{W}$ 为此问题的信息矩阵，可以从最大似然的概率定义给出。请给出此问题下 $\mathbf{W}$ 的具体取值。

3. 假设所有噪声相互无关，该问题存在唯一的解吗? 若有，唯一解是什么? 若没有，说明理由。

Solution:

1.

$$\mathbf{H} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

2.

$$\mathbf{W} = \begin{bmatrix} Q & 0 & 0 & 0 & 0 & 0 \\ 0 & Q & 0 & 0 & 0 & 0 \\ 0 & 0 & Q & 0 & 0 & 0 \\ 0 & 0 & 0 & R & 0 & 0 \\ 0 & 0 & 0 & 0 & R & 0 \\ 0 & 0 & 0 & 0 & 0 & R \end{bmatrix}$$

3. For $\mathbf{x}^\star = \arg\min \frac{1}{2}(\mathbf{z} - \mathbf{H}\mathbf{x})^T \mathbf{W}^{-1} (\mathbf{z} - \mathbf{H}\mathbf{x})$, it satisfies:

$$\frac{\partial \frac{1}{2}(\mathbf{z} - \mathbf{H}\mathbf{x})^T \mathbf{W}^{-1} (\mathbf{z} - \mathbf{H}\mathbf{x})}{\partial \mathbf{x}}\Big|_{\mathbf{x}^\star} = -\mathbf{H}^T \mathbf{W}^{-1} (\mathbf{z} - \mathbf{H}\mathbf{x}^\star) = \mathbf{0}$$

then it can be rewritten as:

$$(\mathbf{H}^T \mathbf{W}^{-1} \mathbf{H}) \mathbf{x}^\star = \mathbf{H}^T \mathbf{W}^{-1} \mathbf{z}$$

If there is a unique solution for the above equation, the necessary and sufficient condition is

$$rank(\mathbf{H}^T \mathbf{W}^{-1} \mathbf{H}) = dim(\mathbf{x}^\star) = 4$$

According to definition of $\mathbf{W}$, W^{-1} can be simply written as:

$$\mathbf{W} = \begin{bmatrix} \frac{1}{Q} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{Q} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{Q} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{R} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{R} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{R} \end{bmatrix}$$

Therefore, $\mathbf{W}^{-1}$ is symmetric positive definite. It follows:

$$rank(\mathbf{H}^T\mathbf{W}^{-1}\mathbf{H}) = rank(\mathbf{H}^T\mathbf{H}) = rank(\mathbf{H}) = 4$$

Thus, the equation has a unique solution, and the the solution can be computed as:

$$\mathbf{x}^{\star} = (\mathbf{H}^T\mathbf{W}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}^{-1}\mathbf{z}$$

References

[1] M. Burri, J. Nikolic, P. Gohl, T. Schneider, J. Rehder, S. Omari, M. W. Achtelik, and R. Siegwart, "The euroc micro aerial vehicle datasets," The International Journal of Robotics Research, 2016.

[2] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," 2012 IEEE Conference On Computer Vision And Pattern Recognition (cvpr), pp. 3354–3361, 2012.

[3] H. Hirschmuller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 2, pp. 807–814, IEEE, 2005.

[4] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," International journal of computer vision, vol. 47, no. 1-3, pp. 7–42, 2002.