Multiple View Geometry: Exercise 6

Dr. Haoang Li, Daniil Sinitsyn, Sergei Solonets, Viktoria Ehm
Computer Vision Group, TU Munich
`Wednesdays 16:00-18:15 at Hörsaal 2, "Interims I" (5620.01.102), and on RBG Live`

Exercise: July 5, 2023

Download the ICRA 2013 paper *Robust Odometry Estimation for RGB-D Cameras* by Kerl, Sturm and Cremers from the *Publications* sections on our webpage.[1] Read the paper and focus in particular on *III. Direct Motion Estimation*.

1. **Image Warping**

   (a) Look at the warping function $\tau(\xi, \mathbf{x})$ in Eq. (9). What do $\tau(\xi, \mathbf{x})$ and $r_i(\xi)$ look like at $\xi = \mathbf{0}$?

   (b) Prove that the derivative of $r_i(\xi)$ w.r.t. ξ at $\xi = \mathbf{0}$ is

   $$\left.\frac{\partial r_i(\xi)}{\partial \xi}\right|_{\xi=\mathbf{0}} = \frac{1}{z}\begin{pmatrix} I_x f_x & I_y f_y \end{pmatrix} \left.\begin{pmatrix} 1 & 0 & -\frac{x}{z} & -\frac{xy}{z} & z + \frac{x^2}{z} & -y \\ 0 & 1 & -\frac{y}{z} & -z - \frac{y^2}{z} & \frac{xy}{z} & x \end{pmatrix}\right|_{(x,y,z)^\top = \pi^{-1}(\mathbf{x}_i, Z_1(\mathbf{x}_i))}$$

   To this end, apply the chain rule multiple times and use the following identity:

   $$\left.\frac{\partial T\left(g(\xi), \mathbf{p}\right)}{\partial \xi}\right|_{\xi=\mathbf{0}} = \begin{pmatrix} \mathrm{Id}_3 & -\hat{\mathbf{p}} \end{pmatrix} \in \mathbb{R}^{3\times 6}\,.$$

   Note: The notation $\partial f(x)/\partial x$ denotes the Jacobian matrix including all first-order partial derivatives, where the number of rows is the number of dimensions of $f(x)$, and the number of columns is the number of dimensions of x.

   (c) Following the derivation in (b), determine the derivative for arbitrary ξ

   $$\left.\frac{\partial r_i(\Delta\xi \circ \xi)}{\partial \Delta\xi}\right|_{\Delta\xi=\mathbf{0}}$$

   where $\circ$ is defined by

   $$\xi_1 \circ \xi_2 := \log\left(\exp(\widehat{\xi_1}) \cdot \exp(\widehat{\xi_2})\right)^\vee\,.$$

   $\vee\colon \mathfrak{se}(3) \to \mathbb{R}^6$ is the inverse of the hat transform.
   Hint: Rewrite the problem such that you can make use of part b).

---

[1] `http://vision.in.tum.de/publications`

1. **Image Warping**

(a) Look at the warping function $\tau(\xi, \mathbf{x})$ in Eq. (9). What do $\tau(\xi, \mathbf{x})$ and $r_i(\xi)$ look like at $\xi = \mathbf{0}$?

(b) Prove that the derivative of $r_i(\xi)$ w.r.t. ξ at $\xi = \mathbf{0}$ is

$$\frac{\partial r_i(\xi)}{\partial \xi}\bigg|_{\xi=\mathbf{0}} = \frac{1}{z}\left(I_x f_x \quad I_y f_y\right)\begin{pmatrix} 1 & 0 & -\frac{x}{z} & -\frac{xy}{z} & z+\frac{x^2}{z} & -y \\ 0 & 1 & -\frac{y}{z} & -z-\frac{y^2}{z} & \frac{xy}{z} & x \end{pmatrix}\bigg|_{(x,y,z)^\top=\pi^{-1}(\mathbf{x}_i, Z_1(\mathbf{x}_i))}$$

To this end, apply the chain rule multiple times and use the following identity:

$$\frac{\partial T\left(g(\xi), \mathbf{p}\right)}{\partial \xi}\bigg|_{\xi=\mathbf{0}} = \left(\mathrm{Id}_3 \quad -\hat{\mathbf{p}}\right) \in \mathbb{R}^{3\times 6}.$$

Note: The notation $\partial f(x)/\partial x$ denotes the Jacobian matrix including all first-order partial derivatives, where the number of rows is the number of dimensions of $f(x)$, and the number of columns is the number of dimensions of x.

(c) Following the derivation in (b), determine the derivative for arbitrary ξ

$$\frac{\partial r_i(\Delta\xi \circ \xi)}{\partial \Delta\xi}\bigg|_{\Delta\xi=\mathbf{0}}$$

where $\circ$ is defined by

$$\xi_1 \circ \xi_2 := \log\left(\exp(\widehat{\xi_1}) \cdot \exp(\widehat{\xi_2})\right)^\vee.$$

$\vee\colon \mathfrak{se}(3) \to \mathbb{R}^6$ is the inverse of the hat transform.

Hint: Rewrite the problem such that you can make use of part b).

$g \in SE(3) \qquad g(\xi) = \exp(\hat{\xi})$

if $\xi = 0 \Rightarrow g(\xi) = I_4$

$\hookrightarrow R = I_3 \quad t = 0$

$\therefore T(g(\xi), p) = Ip + 0 = p$

$\tau(0, x) = \pi\left(T\left[g(\xi), \pi^{-1}(x, Z_1(x))\right]\right)$

$= \pi\left[\pi^{-1}(x, Z_1(x))\right]$

$= x$

$r_i(\xi) = I_2(\tau(\xi, x_i)) - I_1(x_i)$

$= I_2(x_i) - I_1(x_i)$

b) $\frac{\partial r_i(\xi)}{\partial \xi} = \frac{\partial I_2(\tau(\xi, x_i))}{\partial \xi} - \frac{\partial I_1(x_i)}{\partial \xi}$

$= \frac{\partial I_2(\cdot)}{\partial \tau} \cdot \frac{\partial \tau(\xi, x_i)}{\partial \xi} = \boxed{\frac{\partial I(\tau)}{\tau(\xi, x_i)}}\ \boxed{\frac{\partial \pi(T(g, x))}{\partial T}} \cdot \boxed{\underbrace{\frac{\partial T(g, x)}{\partial \xi}}_{(I_3 \ -\hat{p})}}$

$\pi(T) = \left(\frac{f_x x}{z} - c_x, \frac{f_y y}{z} - c_y\right)^T$

$\frac{\partial \pi(T)}{\partial T} = \begin{pmatrix} \frac{f_x}{z} & 0 & -\frac{f_x x}{z^2} \\ 0 & \frac{f_y}{z} & -\frac{f_y \cdot y}{z^2} \end{pmatrix} = \frac{1}{z}\begin{pmatrix} f_x & 0 \\ 0 & f_y \end{pmatrix}\begin{pmatrix} 1 & 0 & -\frac{x}{z} \\ 0 & 1 & -\frac{y}{z} \end{pmatrix}$

$\tau(0, x) = x$

$I(\tau(0, x)) = I(x)$

$\frac{\partial I(x)}{\partial x} = (I_x, I_y)$?

c)

$$\frac{\partial r_i(\Delta\xi \circ \xi)}{\partial \Delta\xi}\bigg|_{\Delta\xi=0}$$

$$\xi_1 \circ \xi_2 := \log\left(\exp(\widehat{\xi_1}) \cdot \exp(\widehat{\xi_2})\right)^\vee.$$

$SO(3) \to so(3)$

$\phi = \ln(R)$

$\Delta\xi \circ \xi = \log(\exp(\Delta\hat{\xi}) \cdot \exp(\hat{\xi}))^\vee$

$\frac{\partial I_2(\tau)}{\partial \tau} \cdot \frac{\partial \pi(T)}{\partial T} \cdot \frac{\partial T(g, x)}{\partial \Delta\xi} = \partial T(g(\Delta\hat{\xi})g(\hat{\xi}), p)$

$= \partial T(g(\hat{\xi}), p')$

same

2. **Image Pyramids**

   In order to handle large translational and rotational motions, a coarse-to-fine scheme is applied in the paper. To go from one level l to $l+1$, the images $I^{(l)}$ (intensity) and $D^{(l)}$ (depth) are downscaled by averaging over intensities or valid depth values, respectively:

$$
\begin{aligned}
I^{(l+1)}(n,m) :&= \frac{1}{4} \cdot \sum_{(n',m') \in O(n,m)} I^{(l)}(n',m') \\
O(n,m) &= \{(2n,2m),(2n+1,2m),(2n,2m+1),(2n+1,2m+1)\} \\
D^{(l+1)}(n,m) :&= \frac{1}{|O_d(n,m)|} \cdot \sum_{(n',m') \in O_d(n,m)} D^{(l)}(n',m') \\
O_d(n,m) &= \{(n',m') \in O(n,m) : D(n',m') \neq 0\}
\end{aligned}
$$

   How does the camera matrix K change from level l to $l+1$? Write down $f_x^{(l+1)}$, $f_y^{(l+1)}$, $c_x^{(l+1)}$ and $c_y^{(l+1)}$ in terms of $f_x^{(l)}$, $f_y^{(l)}$, $c_x^{(l)}$ and $c_y^{(l)}$.

3. **Optimization for Normally Distributed** $p(r_i)$

   (a) Confirm that a normally distributed $p(r_i)$ with a uniform prior on the camera motion leads to normal least squares minimization. To this end, use

$$p(r_i|\xi) = p(r_i) = A \exp\left(-\frac{r_i^2}{\sigma^2}\right)$$

   to show that with a constant prior $p(\xi)$, the maximum a posteriori estimate is given by

$$\xi_{\text{MAP}} = \arg\min_{\xi} \sum_i r_i(\xi)^2 \, .$$

   (b) Explicitly show that the weights

$$w(r_i) = \frac{1}{r_i} \frac{\partial \log p(r_i)}{\partial r_i}$$

   are constant for normally distributed $p(r_i)$.

   (c) Show that in the case of normally distributed $p(r_i)$ the update step $\Delta\xi$ can be computed as

$$\Delta\xi = -\left(J^\top J\right)^{-1} J^\top \mathbf{r}(\mathbf{0}) \, .$$

2. **Image Pyramids**

In order to handle large translational and rotational motions, a coarse-to-fine scheme is applied in the paper. To go from one level l to $l+1$, the images $I^{(l)}$ (intensity) and $D^{(l)}$ (depth) are downscaled by averaging over intensities or valid depth values, respectively:

$$I^{(l+1)}(n,m) := \frac{1}{4} \cdot \sum_{(n',m') \in O(n,m)} I^{(l)}(n',m')$$

$$O(n,m) = \{(2n,2m),(2n+1,2m),(2n,2m+1),(2n+1,2m+1)\}$$

$$D^{(l+1)}(n,m) := \frac{1}{|O_d(n,m)|} \cdot \sum_{(n',m') \in O_d(n,m)} D^{(l)}(n',m')$$

$$O_d(n,m) = \{(n',m') \in O(n,m) : D(n',m') \neq 0\}$$

How does the camera matrix K change from level l to $l+1$? Write down $f_x^{(l+1)}$, $f_y^{(l+1)}$, $c_x^{(l+1)}$ and $c_y^{(l+1)}$ in terms of $f_x^{(l)}$, $f_y^{(l)}$, $c_x^{(l)}$ and $c_y^{(l)}$.

$x^{L+1} = \frac{x^L - 0.5}{2}$ ← 坐标调整

↖ 上采样

$x^{L+1} = \frac{1}{2} K^{L+1} X \qquad x^L = \frac{1}{2} K^L X$

$f_x^{L+1} = \frac{1}{2} f_x \qquad f_y^{L+1} = \frac{1}{2} f_y$

$c_x^{L+1} = \frac{1}{2} c_x - \frac{1}{4}$ - - -

3. **Optimization for Normally Distributed $p(r_i)$**

(a) Confirm that a normally distributed $p(r_i)$ with a uniform prior on the camera motion leads to normal least squares minimization. To this end, use

$$p(r_i|\xi) = p(r_i) = A \exp\left(-\frac{r_i^2}{\sigma^2}\right)$$

to show that with a constant prior $p(\xi)$, the maximum a posteriori estimate is given by

$$\xi_{\text{MAP}} = \arg\min_{\xi} \sum_i r_i(\xi)^2 .$$

$-\log p(r_i|\xi) = -\left[\left(-\frac{r_i^2}{\sigma^2}\right) + \log A\right]$

$\xi_{MAP} = \arg\min_{\xi} -\sum \log p(r_i|\xi) - \log p(\xi)$

$= \arg\min -\sum\left(\log A - \frac{r_i^2}{\sigma^2}\right) - \log p(\xi)$

$= \arg\min -N\log A + \frac{1}{\sigma^2}\sum r_i^2 - \log p(\xi)$

(b) Explicitly show that the weights

$$w(r_i) = \frac{1}{r_i} \frac{\partial \log p(r_i)}{\partial r_i}$$

are constant for normally distributed $p(r_i)$.

b) $\frac{\partial \log p(r_i)}{\partial r_i} = \frac{\partial \log A + \left(-\frac{r_i^2}{\sigma^2}\right)}{\partial r_i}$

$= -\frac{2}{\sigma^2} r_i$

$w(r_i) = \frac{1}{\cancel{r_i}} \cdot \left(-\frac{2}{\sigma^2}\cancel{r_i}\right) = -\frac{2}{\sigma^2}$

(c) Show that in the case of normally distributed $p(r_i)$ the update step $\Delta\xi$ can be computed as

$$\Delta\xi = -\left(J^\top J\right)^{-1} J^\top \mathbf{r}(\mathbf{0}) .$$

c) $J^T W J \Delta\xi = -J^T W r(0)$

$\cancel{W} J \Delta\xi = -J^{-T} J^T \cancel{W} r(0)$ ← diagonal matrix.

$\Delta\xi = -(J J^T)^{-1} J^T r(0)$ orthogonal matrix