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`

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

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