Multiple View Geometry: Solution Sheet 8

Prof. Dr. Daniel Cremers, Simon Weber, Tarun Yenamandra
Computer Vision Group, TU Munich
`Wednesdays 16:00-18:15 at Hörsaal 2, "Interims I" (5620.01.102), and on RBG Live`

Exercise: July 13th, 2022

Part I: Theory

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 $\xi = \mathbf{0}$?

For $\xi = \mathbf{0}$, we have

$$T(g(\mathbf{0}), \mathbf{p}) = T(\mathrm{Id}_4, \mathbf{p}) = \mathrm{Id}_3\mathbf{p} + \mathbf{0} = \mathbf{p}\,.$$

Thus, $\tau(\mathbf{0}, \mathbf{x})$ becomes

$$\tau(\mathbf{0}, \mathbf{x}) = \pi\left(\pi^{-1}(\mathbf{x}, Z_1(\mathbf{x}))\right) = \mathbf{x}\,,$$

where the last equality follows from inserting the formulas for π and π^{-1}. Finally,

$$r_i(\mathbf{0}) = I_2(\tau(\mathbf{0}, \mathbf{x}_i)) - I_1(\mathbf{x}_i) = I_2(\mathbf{x}_i) - I_1(\mathbf{x}_i)\,.$$

(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:

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

Since $I_1(\mathbf{x}_i)$ does not depend on ξ, we only need to look at the first term in $r_i(\xi)$. It is a composition of the functions I_2, π and T. Applying the chain rule gives

$$\left.\frac{\partial r_i(\xi)}{\partial \xi}\right|_{\xi=\mathbf{0}} = \left.\frac{\partial I_2(\mathbf{y})}{\partial \mathbf{y}}\right|_{\mathbf{y}=\pi(T(g(\mathbf{0}), \pi^{-1}(\mathbf{x}_i, Z_1(\mathbf{x}_i))))} \cdot \left.\frac{\partial \pi(\mathbf{p})}{\partial \mathbf{p}}\right|_{\mathbf{p}=T(g(\mathbf{0}), \pi^{-1}(\mathbf{x}_i, Z_1(\mathbf{x}_i)))} \cdot \left.\frac{\partial T(g(\xi), \pi^{-1}(\mathbf{x}_i, Z_1(\mathbf{x}_i)))}{\partial \xi}\right|_{\xi=\mathbf{0}}\,.$$

Now, we know that $T(g(\mathbf{0}), \mathbf{p}) = \mathbf{p}$, so we can write

$$\left.\frac{\partial r_i(\xi)}{\partial \xi}\right|_{\xi=\mathbf{0}} = \left.\left[\left(\nabla I_2\left(\pi(\mathbf{p})\right)\right)^\top \cdot \frac{\partial \pi(\mathbf{p})}{\partial \mathbf{p}} \cdot \begin{pmatrix} \mathrm{Id}_3 & -\hat{\mathbf{p}} \end{pmatrix}\right]\right|_{\mathbf{p}=\pi^{-1}(\mathbf{x}_i, Z_1(\mathbf{x}_i))}$$

The second and third term are

$$\frac{\partial \pi(\mathbf{p})}{\partial \mathbf{p}} = \begin{pmatrix} \frac{\partial}{\partial x}\left(\frac{f_x x}{z}\right) & \frac{\partial}{\partial y}\left(\frac{f_x x}{z}\right) & \frac{\partial}{\partial z}\left(\frac{f_x x}{z}\right) \\ \frac{\partial}{\partial x}\left(\frac{f_y y}{z}\right) & \frac{\partial}{\partial y}\left(\frac{f_y y}{z}\right) & \frac{\partial}{\partial z}\left(\frac{f_y y}{z}\right) \end{pmatrix} = \frac{1}{z}\begin{pmatrix} f_x & 0 \\ 0 & f_y \end{pmatrix}\begin{pmatrix} 1 & 0 & -\frac{x}{z} \\ 0 & 1 & -\frac{y}{z} \end{pmatrix}$$

$$\begin{pmatrix}\mathrm{Id}_3 & -\hat{\mathbf{p}}\end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & z & -y \\ 0 & 1 & 0 & -z & 0 & x \\ 0 & 0 & 1 & y & -x & 0 \end{pmatrix}$$

Performing the matrix multiplication and using

$$\pi(\mathbf{p}) = \pi\left(\pi^{-1}(\mathbf{x}_i, Z_1(\mathbf{x}_i))\right) = \mathbf{x}_i$$

(see (a)) as well as

$$(\nabla I_2(\mathbf{x}_i))^\top \begin{pmatrix} f_x & 0 \\ 0 & f_y \end{pmatrix} = \begin{pmatrix} I_x(\mathbf{x}_i) f_x & I_y(\mathbf{x}_i) f_y \end{pmatrix}$$

leads to the desired result.

(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

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

We observe that due to associativity of matrix multiplication

$$T\left(g(\Delta\xi \circ \xi), \mathbf{p}\right) = T\left(g(\Delta\xi) g(\xi), \mathbf{p}\right) = T\left(g(\Delta\xi), T\left(g(\xi), \mathbf{p}\right)\right)$$

and thus

$$\left.\frac{\partial T\left(g(\Delta\xi \circ \xi), \mathbf{p}\right)}{\partial \Delta\xi}\right|_{\Delta\xi=\mathbf{0}} = \left.\frac{\partial T\left(g(\Delta\xi), T\left(g(\xi), \mathbf{p}\right)\right)}{\partial \Delta\xi}\right|_{\Delta\xi=\mathbf{0}} = \begin{pmatrix}\mathrm{Id}_3 & -\widehat{\mathbf{p}'}\end{pmatrix} \in \mathbb{R}^{3\times 6}$$

with

$$\mathbf{p}' = T\left(g(\xi), \mathbf{p}\right) .$$

Thus analogously to (b) we can derive

$$\left.\frac{\partial r_i(\Delta\xi \circ \xi)}{\partial \Delta\xi}\right|_{\Delta\xi=\mathbf{0}} = \left.\left[\left(\nabla I_2\left(\pi(\mathbf{p}')\right)\right)^\top \cdot \frac{\partial \pi(\mathbf{p}')}{\partial \mathbf{p}'} \cdot \begin{pmatrix}\mathrm{Id}_3 & -\widehat{\mathbf{p}'}\end{pmatrix}\right]\right|_{\mathbf{p}'=T(g(\xi),\pi^{-1}(\mathbf{x}_i, Z_1(\mathbf{x}_i)))} .$$

Notice that the unlike in (b), the image gradient is now evaluated at the warped point $\tau(\xi, \mathbf{x}_i)$. With this we get

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

$$\mathbf{p}' = T(g(\xi), \pi^{-1}(\mathbf{x}_i, Z_1(\mathbf{x}_i))) ,$$

which is very similar to the result in (b), just that everything (including the image gradient) is evaluated at the transformed point $\mathbf{p}'$. What follows are some additional remarks.

Note: In this exercise we use what is called "Forward Additive" direct image alignment, whereas the 2013 ICRA paper describes "Forward Compositional" approach. The difference in the latter is that you always compute an intermediate "warped" image in every iteration and you compute image gradients for that. For further explanation of the two approaches see Christian Kerl's Master thesis[1] (Sections 4.2.1. and 4.2.2.).

Note: It is in principle also possible to linearize as

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

however, the derivative becomes more complicated and slower to compute, since the derivative of the exponential map for $\mathrm{SE}(3)$ for $\xi \neq \mathbf{0}$ is a bit more involved. Moreover, since with that we only use a linear approximation of the exponential map of $\mathrm{SE}(3)$, convergence properties of the resulting optimization algorithm can be slower, as suggested by some author's experiments. Lastly, one needs to then also perform the update step as $\xi_{\text{new}} = \Delta\xi + \xi_{\text{old}}$ and take special care at rotations close to $180°$ (e.g. wrapping around to $-180°$).

2. **Image Pyramids**

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 $c_y^{(l)}$.

Looking at how each pixel coordinate transforms from one image level l to the next, $l+1$, we have

$$2\mathbf{x}^{(l+1)} + \tfrac{1}{2} = \mathbf{x}^{(l)} \quad \Rightarrow \quad \mathbf{x}^{(l+1)} = \tfrac{1}{2}\mathbf{x}^{(l)} - \tfrac{1}{4} .$$

Plugging into the latter the relations $\mathbf{x}^{(l+1)} = \frac{1}{Z}K^{(l+1)}\mathbf{X}$ and $\mathbf{x}^{(l)} = \frac{1}{Z}K^{(l)}\mathbf{X}$ results in

$$f_x^{(l+1)} = \tfrac{1}{2}f_x^{(l)} , \quad f_y^{(l+1)} = \tfrac{1}{2}f_y^{(l)} , \quad c_x^{(l+1)} = \tfrac{1}{2}c_x^{(l)} - \tfrac{1}{4} , \quad c_y^{(l+1)} = \tfrac{1}{2}c_y^{(l)} - \tfrac{1}{4} .$$

Note: Remember that we define the $(0,0)$ continuous image coordinate to lie in the *center* of the top-left pixel.

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

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

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

[1] https://vision.in.tum.de/_media/spezial/bib/kerl2012msc.pdf

Inserting into Eq. (15) gives

$$\xi_{\text{MAP}} = \arg\min_{\xi} \left(-N \log A + \frac{1}{\sigma^2} \sum_i r_i(\xi)^2 - \log p(\xi) \right) = \arg\min_{\xi} \sum_i r_i(\xi)^2 \,,$$

since $-N \log A$ and $-\log p(\xi)$ are just constant shifts and $\frac{1}{\sigma^2}$ is only a scaling, and none of them changes the argmin.

(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 $p(r_i)$.

$$w(r_i) = \frac{1}{r_i} \frac{\partial \log p(r_i)}{\partial r_i} = \frac{1}{r_i} \frac{\partial \left(\log A - \frac{r_i(\xi)^2}{\sigma^2} \right)}{\partial r_i} = \frac{1}{r_i} \left(0 - \frac{2 r_i}{\sigma^2} \right) = -\frac{2}{\sigma^2} = \text{const}(r_i)$$

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

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

Eq. (21) reads

$$J^\top W J \Delta\xi = -J^\top W \mathbf{r}(\mathbf{0}) \,,$$

with W a diagonal matrix with constant diagonal entries $W_{ii} = w(r_i) = -\frac{2}{\sigma^2}$.

$$\Rightarrow \quad W = -\frac{2}{\sigma^2}\text{Id} \quad \Rightarrow \quad -\frac{2}{\sigma^2} J^\top J \Delta\xi = \frac{2}{\sigma^2} J^\top \mathbf{r}(\mathbf{0}) \quad \Rightarrow \quad \text{claim} \,.$$