Bundle adjustment

1 Numerical Exercises

In the following, we will derive the first steps for the "Closed-form solution of visual-inertial structure from motion" Martinelli et al., IJCV14. To stay consistent with the paper, we introduce the following notation for the problem. A platform with a camera and an IMU moves in a 3D environment relative to a global frame. Vectors in this global frame are written in lower-case letters, e.g. $\boldsymbol{w}(\tau)$. Furthermore, assume that the platform frame coincides with the camera frame, which we call local frame. We will adopt upper-case letters to denote vectors in this frame, e.g. $\boldsymbol{W}_t(\tau)$ expressed in the local frame at t. The rotation occurred during the time interval (t_1, t_2) is $C_{t_2}^{t_1}$. Finally, C^t will denote the rotation matrix between the global frame and the local frame at time t. Thus, we can express a vector in the world frame $\boldsymbol{w}(\tau)$ using the rotation matrix C^t as $\boldsymbol{w}(\tau) = C^t \boldsymbol{W}_t(\tau)$.

1. In a first step, state the position of the platform $\boldsymbol{r}(t)$ expressed in the world frame at any time $t \in [T_{in}, T_{fin}]$ using a double integral. Assume that you have access to the correct platform acceleration $\boldsymbol{a}(\tau)$, the initial position $\boldsymbol{r}(T_{in})$ and the initial velocity $\boldsymbol{v}(T_{in})$, which are all expressed in the global frame.
   Solution
   $$\boldsymbol{r}(t) = \boldsymbol{r}(T_{in}) + \boldsymbol{v}(T_{in})\Delta t + \int_{T_{in}}^{t} \int_{T_{in}}^{\tau} \boldsymbol{a}(\xi) d\xi d\tau$$
   Where $\Delta t = t - T_{in}$

2. Rewrite the position of the platform $\boldsymbol{r}(t)$ by simplifying the double integral using integration by parts.
   Solution
   The integration by parts formula states:
   $$\int_a^b u(x)v'(x)dx = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x)dx$$
   Thus, we can rewrite the double integral $\int_{T_{in}}^{t} \int_{T_{in}}^{\tau} \boldsymbol{a}(\xi) d\xi d\tau$ using integration by parts
   $$\int_{T_{in}}^{t} \int_{T_{in}}^{\tau} \boldsymbol{a}(\xi) d\xi d\tau = \int_{T_{in}}^{t} u(\tau)v'(\tau)d\tau$$
   $$u(\tau) = \textstyle\int_{T_{in}}^{\tau} a(\xi)d\xi \qquad u(\tau)' = a(\tau)$$
   $$v'(\tau) = 1 \qquad v(\tau) = \tau$$

$$\int_{T_{in}}^{t} \int_{T_{in}}^{\tau} \boldsymbol{a}(\xi) d\xi d\tau = [u(\tau)v(\tau)]_{T_{in}}^{t} - \int_{T_{in}}^{t} u'(\tau)v(\tau)d\tau = [\int_{T_{in}}^{\tau} \boldsymbol{a}(\xi)d\xi\tau]_{T_{in}}^{t} - \int_{T_{in}}^{t} \boldsymbol{a}(\tau)\tau d\tau$$

$$\begin{aligned} \int_{T_{in}}^{t} \int_{T_{in}}^{\tau} \boldsymbol{a}(\xi) d\xi d\tau &= [u(\tau)v(\tau)]_{T_{in}}^{t} - \int_{T_{in}}^{t} u'(\tau)v(\tau)d\tau = [\int_{T_{in}}^{\tau} \boldsymbol{a}(\xi)d\xi\tau]_{T_{in}}^{t} - \int_{T_{in}}^{t} \boldsymbol{a}(\tau)\tau d\tau \\ &= \int_{T_{in}}^{t} \boldsymbol{a}(\xi)td\xi - \int_{T_{in}}^{t} \boldsymbol{a}(\tau)\tau d\tau = \int_{T_{in}}^{t} (t-\tau)\boldsymbol{a}(\tau)d\tau \end{aligned}$$

   In the last step, we just replaced the variable ξ with τ.

3. In practice, the accelerometer does not provide the acceleration $\boldsymbol{a}(\tau)$ in the global frame. Instead, it provides a sensor measurment $\boldsymbol{A}_\tau(\tau)$ expressed in the local frame at time τ, which includes the gravitation vector $\boldsymbol{G}_\tau$ and a constant bias $\boldsymbol{B}$. For this exercise, we omit the IMU noise.
   Based on this influences, state the sensor measurement of the IMU $\boldsymbol{A}_\tau(\tau)$ as a function of $\boldsymbol{G}_\tau$, $\boldsymbol{B}$ and $\boldsymbol{A}_\tau^{inertial}(\tau)$, which is the inertial acceleration of the platform expressed in the local frame at time τ.
   Solution
$$\boldsymbol{A}_\tau(\tau) = \boldsymbol{A}_\tau^{inertial}(\tau) - \boldsymbol{G}_\tau + \boldsymbol{B}$$
   Note that the gravity comes with a minus since, when the platform does not accelerate (i.e. $\boldsymbol{A}_\tau^{inertial}(\tau)$ is zero), the accelerometer perceives an acceleration which is the same of an object accelerated upward in absence of gravity.

4. In a next step, state the position vector $\boldsymbol{r}(t)$ using $\boldsymbol{A}_\tau(\tau)$ as well as the platform rotation $C_{T_{in}}^{\tau}$ between timestep T_{in} and τ. Simplify the resulting formula.
   Solution
$$\boldsymbol{r}(t) = \boldsymbol{r}(T_{in}) + \boldsymbol{v}(T_{in})\Delta t + \int_{T_{in}}^{t} (t-\tau)\boldsymbol{a}(\tau)d\tau = \boldsymbol{r}(T_{in}) + \boldsymbol{v}(T_{in})\Delta t + \int_{T_{in}}^{t} (t-\tau)C_{T_{in}}^{\tau}(\boldsymbol{A}_\tau^{inertial}(\tau) - \boldsymbol{G}_\tau + \boldsymbol{B})d\tau$$
   We can simplify the above equation by rewriting the gravity vector in the world frame $\boldsymbol{g}$ and using $\Delta t = t - T_{in}$
$$\boldsymbol{r}(t) = \boldsymbol{r}(T_{in}) + \boldsymbol{v}(T_{in})\Delta t + \boldsymbol{g}\frac{\Delta t^2}{2} + C^{T_{in}}[\boldsymbol{S}_{T_{in}}(t) - \Gamma(t)\boldsymbol{B}]$$
   where
$$\boldsymbol{S}_{T_{in}}(t) = \int_{T_{in}}^{t} (t-\tau)C_{T_{in}}^{\tau}\boldsymbol{A}_\tau(\tau)d\tau$$
$$\Gamma(t) = \int_{T_{in}}^{t} (t-\tau)C_{T_{in}}^{\tau}d\tau$$

5. Is it possible to measure with an IMU sensor all variables needed for computing $\boldsymbol{r}(t)$?
   Solution
   Yes. The matrix $C_{T_{in}}^{\tau}$ can be obtained from the angular speed during the interval $[T_{in}, \tau]$ provided by the gyroscopes. Hence, also the matrix $\Gamma(t)$ can be obtained by directly integrating the gyroscope data during the interval $[T_{in}, t]$. Finally, the vector $\boldsymbol{S}_{T_{in}}(t)$ can be obtained by integrating the data provided by the gyroscopes and the accelerometers delivered during the interval $[T_{in}, t]$.

6. Since we are interested in combining a RGB camera with an IMU for VIO, let us now suppose that N point-features are observed in the camera. Let us denote their position in the physical world with $\boldsymbol{p}^i,\ i = 1, ..., N$. According to our notation,$\boldsymbol{P}_t^i(t)$ will denote their position at time t in the local frame at time t.
   Write down the formula to convert the feature coordinates in the local $\boldsymbol{P}_t^i(t)$ to the global frame $\boldsymbol{p}^i$.
   Solution
$$\boldsymbol{p}^i = \boldsymbol{r}(t) + C^{T_{in}}C_{T_{in}}^{t}\boldsymbol{P}_t^i(t)$$

7. Finally, using the above equation, we can relate the location of one feature $\boldsymbol{P}_{T_{in}}^i(T_{in})$ at timestep T_{in} to the relative location $\boldsymbol{P}_t^i(t)$ at timestep t. Both vectors are expressed in the corresponding local frame. State the resulting formula.
   Solution
   We can rewrite the equation found in subquestion 6 for timestep T_{in} as follows:
$$\boldsymbol{p}^i - \boldsymbol{r}(T_{in}) = C^{T_{in}}\boldsymbol{P}_{T_{in}}^i(T_{in})$$

By plugging the above equation again into the equation found in subquestion 6, we obtain

$$C^{T_{in}} C^{t}_{T_{in}} \boldsymbol{P}^{i}_{t}(t) = \boldsymbol{r}(t) - \boldsymbol{r}(T_{in}) + C^{T_{in}} \boldsymbol{P}^{i}_{T_{in}}(T_{in})$$

We can then pre multiply the above equation with the rotation matrix $(C^{T_{in}})^{-1}$

$$C^{t}_{T_{in}} \boldsymbol{P}^{i}_{t}(t) = (C^{T_{in}})^{-1}(\boldsymbol{r}(T_{in}) - \boldsymbol{r}(t)) + \boldsymbol{P}^{i}_{T_{in}}(T_{in})$$

Replacing $\boldsymbol{r}(T_{in}) - \boldsymbol{r}(t)$ with the formula found in 4 leads to

$$\begin{aligned}
C^{t}_{T_{in}} \boldsymbol{P}^{i}_{t}(t) =& (C^{T_{in}})^{-1}(-\boldsymbol{v}(T_{in})\Delta t - \boldsymbol{g}\frac{\Delta t^2}{2} - C^{T_{in}}[\boldsymbol{S}_{T_{in}}(t) - \Gamma(t)\boldsymbol{B}]) + \boldsymbol{P}^{i}_{T_{in}}(T_{in}) \\
=& \boldsymbol{P}^{i}_{T_{in}}(T_{in}) - (C^{T_{in}})^{-1}(\boldsymbol{v}(T_{in})\Delta t + \boldsymbol{g}\frac{\Delta t^2}{2}) + \Gamma(t)\boldsymbol{B} - \boldsymbol{S}_{T_{in}}(t) \\
=& \boldsymbol{P}^{i}_{T_{in}}(T_{in}) - (\boldsymbol{V}_{T_{in}}(T_{in})\Delta t + \boldsymbol{G}_{T_{in}}\frac{\Delta t^2}{2}) + \Gamma(t)\boldsymbol{B} - \boldsymbol{S}_{T_{in}}(t)
\end{aligned}$$

with

$$\boldsymbol{V}_{T_{in}} = (C^{T_{in}})^{-1}\boldsymbol{v}(T_{in}) \qquad \boldsymbol{G}_{T_{in}} = (C^{T_{in}})^{-1}\boldsymbol{g}$$

The final equation includes only variables expressed in the local frame at timestep T_{in}. Thus, we can use it to find the camera transformation from timestep T_{in} to timestep t.

By further reformulating the equation in subquestion 7, we obtain a linear system of equations, which can be solved for the camera poses and the scale of the features.