Trajectory Extractor Write-up

Aubrey Clausse

August 29, 2019

1 Camera Model

1.1 Pin-Hole camera model

The pin-hole camera model is a simple model that describe the relationship between a 3D point in world coordinate frame is denoted by $\mathbf{M} = [X,Y,Z]^T$, and its projection in pixels denoted $\mathbf{m} = [u,v]^T$. We use $\tilde{\mathbf{x}}$ to denote the augmented vector by adding 1 as the last element. The relationship between a 3D point and its image projection is given by the pin-hole camera model [1]:

$$s\,\tilde{\mathbf{m}} = \mathbf{C}_{cam}\begin{bmatrix}\mathbf{R}_{cam} & \mathbf{T}_{cam}\end{bmatrix}\tilde{\mathbf{M}} \quad \text{with } \mathbf{C}_{cam} = \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

Where $\mathbf{R}_{cam}, \mathbf{T}_{cam}$ are the rotation and translation matrix called the extrinsic parameters, which define the position and rotation of the camera in the world coordinate frame. $\mathbf{C}_{cam}$ denotes the camera matrix with f the focal length and c_x, c_y the principal points set at the image center.

1.2 Camera calibration

In case the camera information is not accessible, e.g videos from online streams, we describe here a calibration method used to compute the intrinsic (focal length) and extrinsic (rotation and position) camera parameters assuming no lens distortion. Using a satellite view of the area covered by the traffic camera, we identify four or more pairs of matching points between the traffic camera view and the satellite view $(\mathbf{m}_i;\mathbf{M}_i)$, with $(\mathbf{m}_i$ pixel coordinate and $(\mathbf{M}_i$ world coordinate of the i^{th} pair.

Camera parameters such as rotation, translation and focal length of the traffic camera are found by minimizing the reprojection error:

$$\min_{\mathbf{R}_{cam},\mathbf{T}_{cam},f} \quad \sum_i \|g(\mathbf{M}_i) - \mathbf{m}_i\| \tag{2}$$

With $g : \mathbb{R}^3 \to \mathbb{R}^2$ the function that projects 3D world points onto the image plane according to the pin-hole camera model defined in Eq-1.

Note: This method can easily be extended to add distortion coefficients by adding a distortion model [1] in the projection equation Eq-1, and adding these coefficients in the optimization formulation Eq-2.

Note: When $\mathbf{R}$ is close to the identity matrix (such as in bird view), the optimization problem 2 is ill-posed. To overcome this issue in this case, we arbitrary fix the focal-length of the camera.

2 Trajectory Processing - Filtering / Smoothing

This step aims at taking the raw tracks measurements (pixel coodinates) and create smooth world frame trajectories with position, velocity and heading of the vehicles. To this end, a *Rauch-Tung-Striebel Smoother* is used because it allows to take into account the measurement uncertainty directly in pixels and the underlying kinematic constraints of the vehicles.

2.1 Process model

Let's define the *Bicycle Model* from [2]:

$$\begin{aligned}
\dot{x} &= v\,cos(\psi+\beta) \\
\dot{y} &= v\,sin(\psi+\beta) \\
\dot{v} &= u_1 \\
\dot{\psi} &= \frac{v}{l_r}sin(\beta) \\
\dot{\beta} &= u_2
\end{aligned} \tag{3}$$

With x, y position of the center of gravity, l_r, l_f distance between the center of gravity and the rear / front wheel, ψ orientation of the vehicle and β the slip angle at the center of gravity, u_1, u_2 the longitudinal acceleration control and lateral control.

Let's define the state $\mathbf{x} = [x, y, v, \psi, \beta]^T$ for the *Bicycle Model*. The process model is then defined by:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) + \mathbf{w} \tag{4}$$

For the purpose of deriving the *Rauch-Tung-Striebel* smoothing equation, we defined the partial derivative A matrix as $A = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}$:

$$A = \begin{bmatrix}
0 & 0 & cos(\psi+\beta) & -v\,sin(\psi+\beta) & -v\,sin(\psi+\beta) \\
0 & 0 & sin(\psi+\beta) & v\,cos(\psi+\beta) & v\,cos(\psi+\beta) \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{sin(\beta)}{l_r} & 0 & -\frac{v}{l_r}cos(\beta) \\
0 & 0 & 0 & 0 & 0
\end{bmatrix} \tag{5}$$

It is assumed the vehicle moves at constant velocity and sideslip angle perturbed by a zero-mean gaussian noise such that:

$$\mathbf{w} \sim \mathcal{N}(0, \mathbf{Q}_d), \ with\ \mathbf{Q_d} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \tau_v & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \tau_\beta \end{bmatrix} \tag{6}$$

2.2 Measurement model

We use the pixel position of the geometric center of the vehicle on the ground as measurement. Defining the measurement vector as $\mathbf{z} = [u, v]^T$ with u, v the pixel positions, the measurement model $\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{v}$ is derived from:

$$\mathbf{z} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tilde{\mathbf{z}} \tag{7}$$

With according to the pin-hole camera model Eq.1:

$$s\,\tilde{\mathbf{z}} = \mathbf{C}_{cam}\,\mathbf{R}_{cam}\,\mathbf{D}\,\mathbf{x} + \mathbf{C}_{cam}\,\mathbf{T}_{cam} \quad \text{with}\ \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{8}$$

For the purpose of deriving the *Rauch-Tung-Striebel* smoothing equation, we defined the partial derivative H matrix as $H = \frac{\partial \mathbf{h}}{\partial \mathbf{x}}$:

$$\begin{aligned} H &= \frac{\partial \mathbf{z}}{\partial \mathbf{x}} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{x}} \end{aligned} \tag{9}$$

And from the chain rule, we have:

$$\begin{aligned} \frac{\partial s\,\tilde{\mathbf{z}}}{\partial \mathbf{x}} &= \frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{x}} (s \otimes \mathbf{I}) + \tilde{\mathbf{z}} \frac{\partial s}{\partial \mathbf{x}} \\ &= s\,\frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{x}} + \tilde{\mathbf{z}} \frac{\partial s}{\partial \mathbf{x}} \end{aligned} \tag{10}$$

Which gives:

$$\frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{x}} = \frac{1}{s} \left(\frac{\partial s\,\tilde{\mathbf{z}}}{\partial \mathbf{x}} - \tilde{\mathbf{z}} \frac{\partial s}{\partial \mathbf{x}} \right) \tag{11}$$

With from Eq.8:

$$\frac{\partial s\,\tilde{\mathbf{z}}}{\partial \mathbf{x}} = \mathbf{C}_{cam}\,\mathbf{R}_{cam}\,\mathbf{D} \tag{12}$$

Since $s\tilde{\mathbf{z}} = [su, sv, s]^T$, we have $s = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}\ s\,\tilde{\mathbf{z}}$ which gives:

$$\frac{\partial s}{\partial \mathbf{x}} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \mathbf{C}_{cam}\,\mathbf{R}_{cam}\,\mathbf{D} \tag{13}$$

Therefore, we have:

$$\frac{\partial \tilde{\mathbf{z}}}{\partial \mathbf{x}} = \frac{1}{s}\left(\mathbf{C}_{cam}\,\mathbf{R}_{cam}\,\mathbf{D} - \tilde{\mathbf{z}}\,\begin{bmatrix}0 & 0 & 1\end{bmatrix}\mathbf{C}_{cam}\,\mathbf{R}_{cam}\,\mathbf{D}\right) \tag{14}$$

With s and $\tilde{\mathbf{z}}$ computed according to the measurement model directly at each measurement step. Then, H can easily be computed from 9 and 14.

It is assumed the measurements are perturbed by a zero-mean gaussian noise such that:

$$\mathbf{v} \sim \mathcal{N}(0, \mathbf{R}),\ with\ \mathbf{R} = \begin{bmatrix}\tau_u & 0\\ 0 & \tau_v\end{bmatrix} \tag{15}$$

2.3 Rauch-Tung-Striebel Smoother

To implement the derived RTS smoothing equations, we first convert the continous-time model defined in Sec.2.1 2.2 into a discrete model using a first order approximation according to [3] (section 8.1):

$$\begin{aligned}\mathbf{F} &= \mathbf{A} + \Delta t\mathbf{I}\\ \mathbf{Q_d} &= \mathbf{Q_c} + \Delta t\mathbf{I}\end{aligned} \tag{16}$$

Then we can use the regular RTS smoothing equations as defined in [3].

References

[1] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1330–1334, Nov 2000.

[2] P. Polack, F. Altch, B. d'Andra-Novel, and A. de La Fortelle, "The kinematic bicycle model: A consistent model for planning feasible trajectories for autonomous vehicles?," in *2017 IEEE Intelligent Vehicles Symposium (IV)*, pp. 812–818, June 2017.

[3] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley-Interscience, 2006.