Auto-Calibration with Stop Signs

Yuhan Liu, Yunhai Han

Jul 31, 2020

Autonomous Vehicle Laboratory

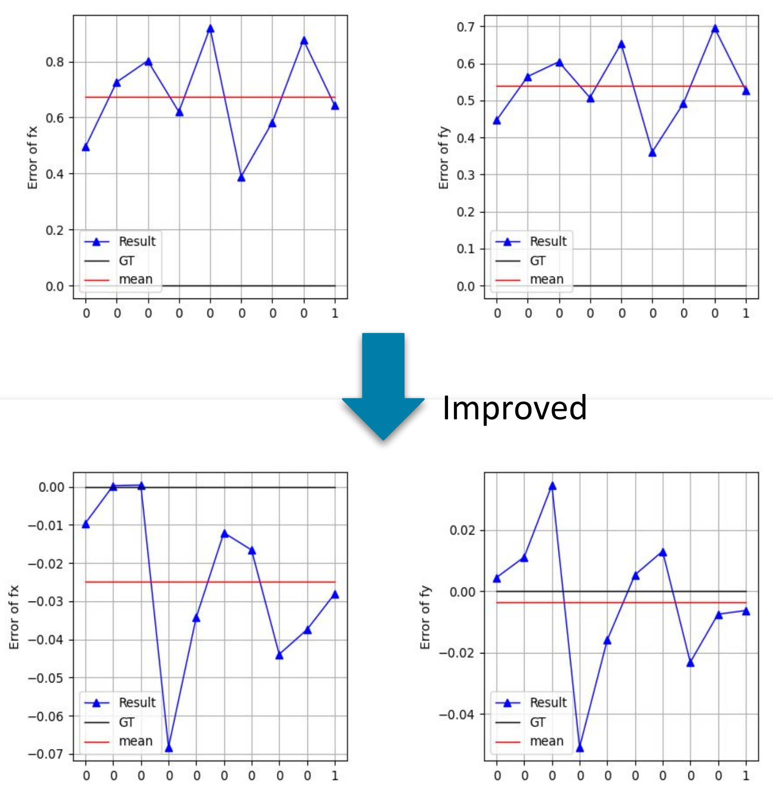
Cognitive Robotics

Overview
calibrate intrinsics using estimated stop sign corners

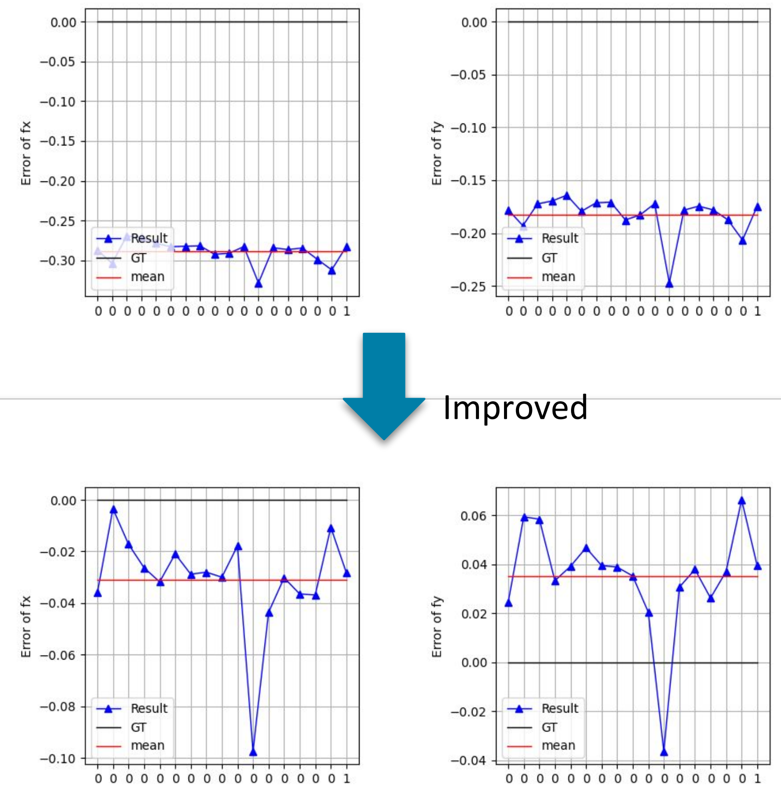
- Calibration Results on AVL Dataset
- System Pipeline (Revisit)
- Line Improvement Method
- Problems Unsolved and Future Works

Calibration Results on AVL Dataset

1. Extract the images in rosbag’s, and then calibrate both camera 1 and camera 6 (intrinsic matrices);
2. Compare the calibration results with the groundtruth obtained using chessboard calibration;
3. Compare the relative errors before and after improvement.



Camera 1

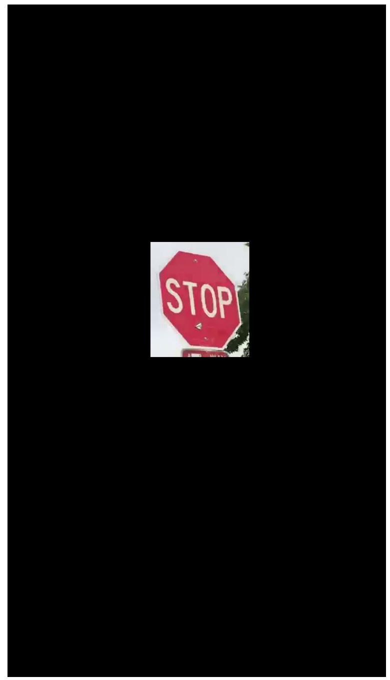


Camera 6

System Pipeline

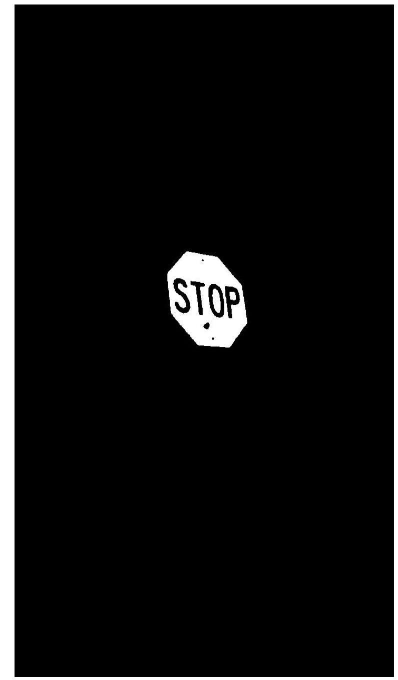
Video Frame

Doing
- YOLO
- RCNN



Bounding Box

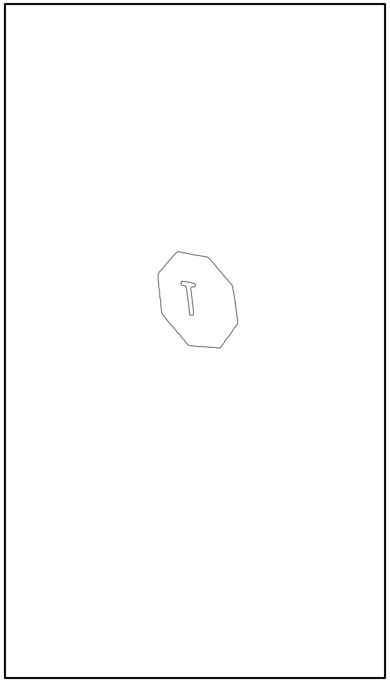
Current
HSV



Pixel-wise Mask

System Pipeline

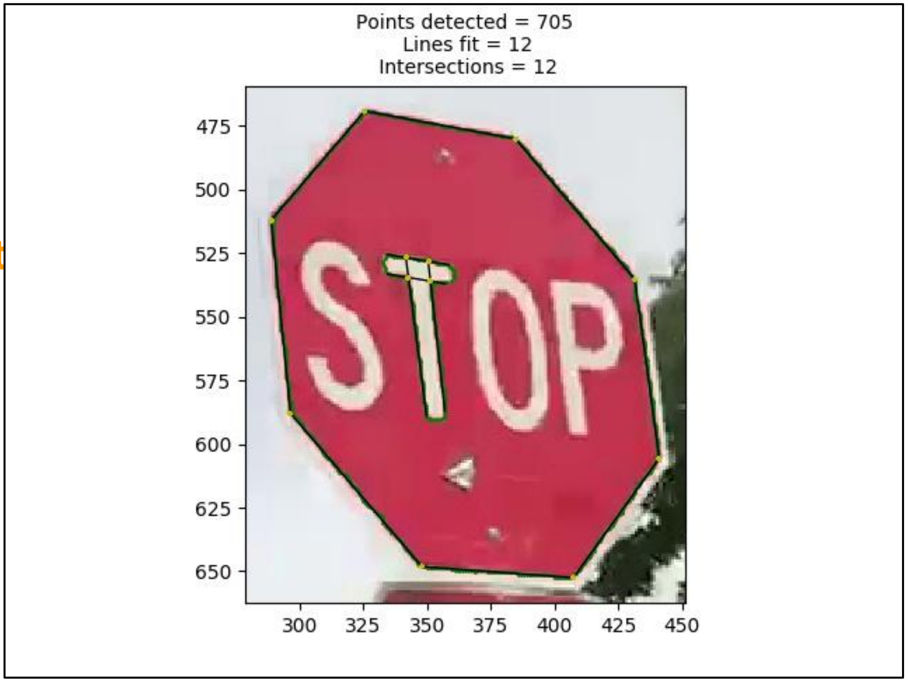
Current
Canny Edge
+Devernay
Sub-Pixel
Correction*



Edge Points

Current
SVD
+RANSAC
+Line
Improvement

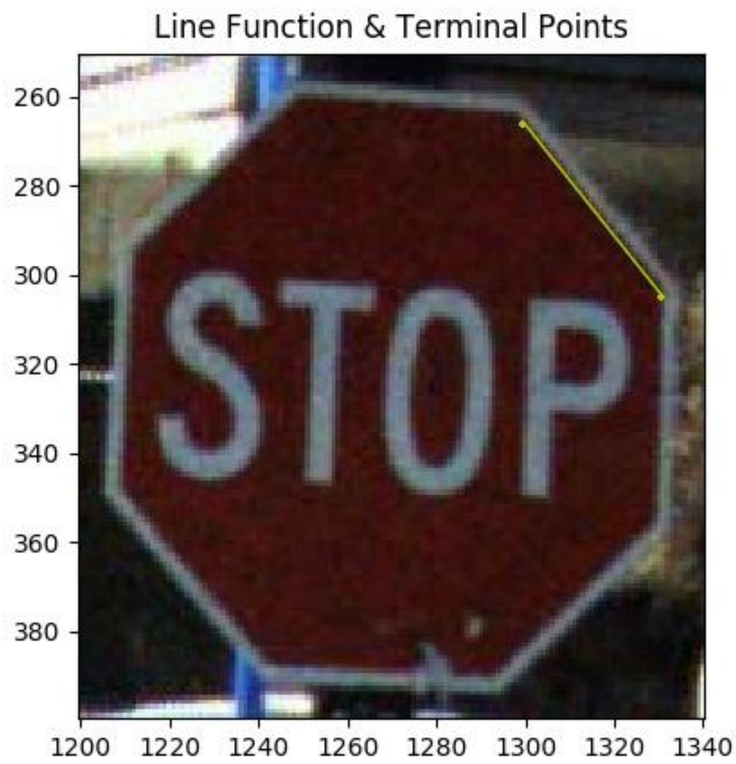
Future
+Data
Association

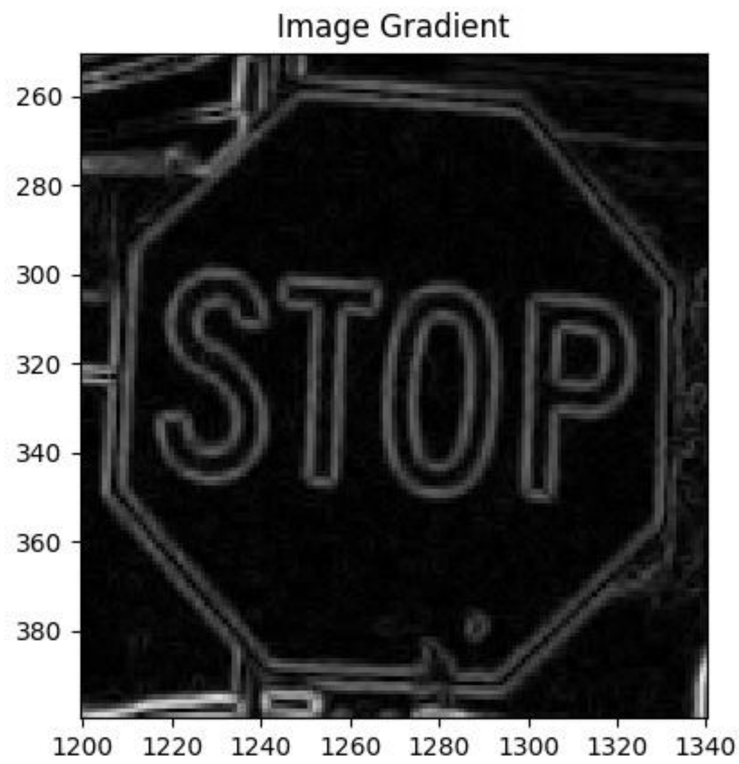


Line Estimation + Intersections

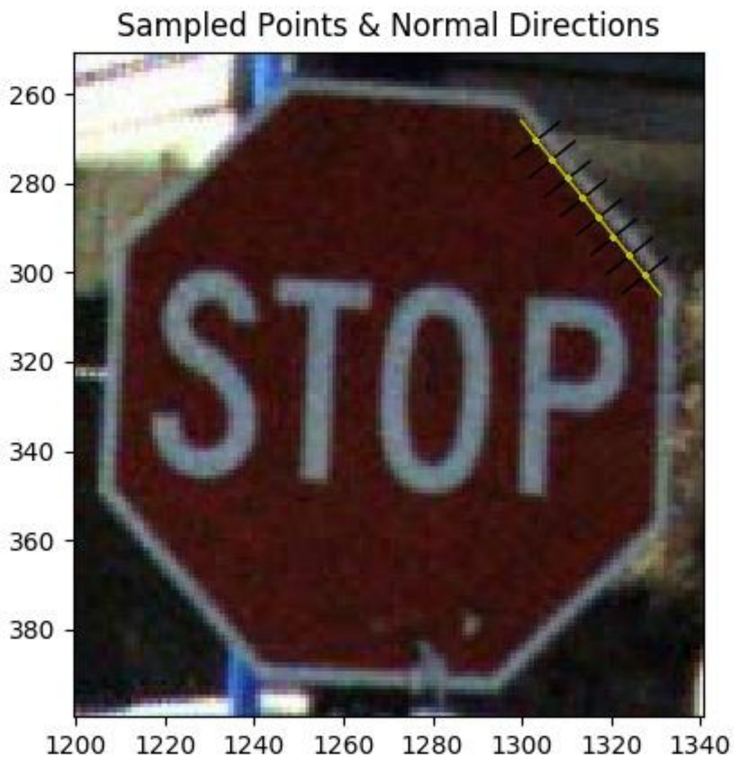
*: R.G. Gioi & G. Randall(2017), "A Sub-Pixel Edge Detector: an Implementation of the Canny/Devernay Algorithm"

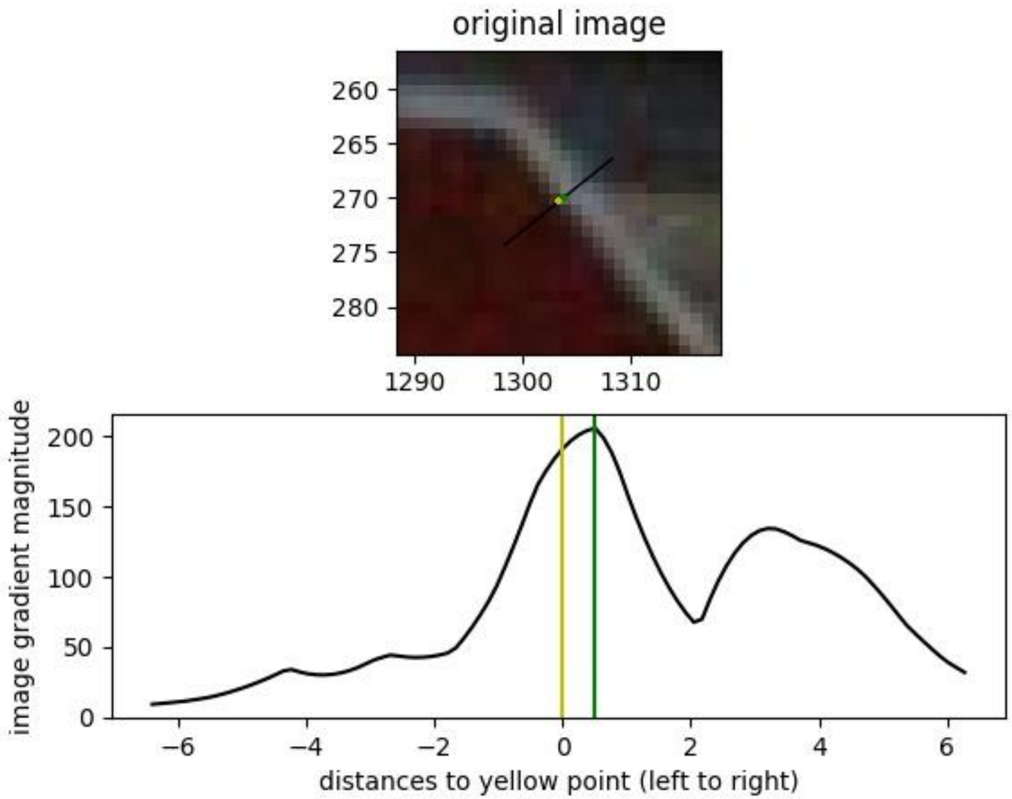
Line Improvement Method - Inputs



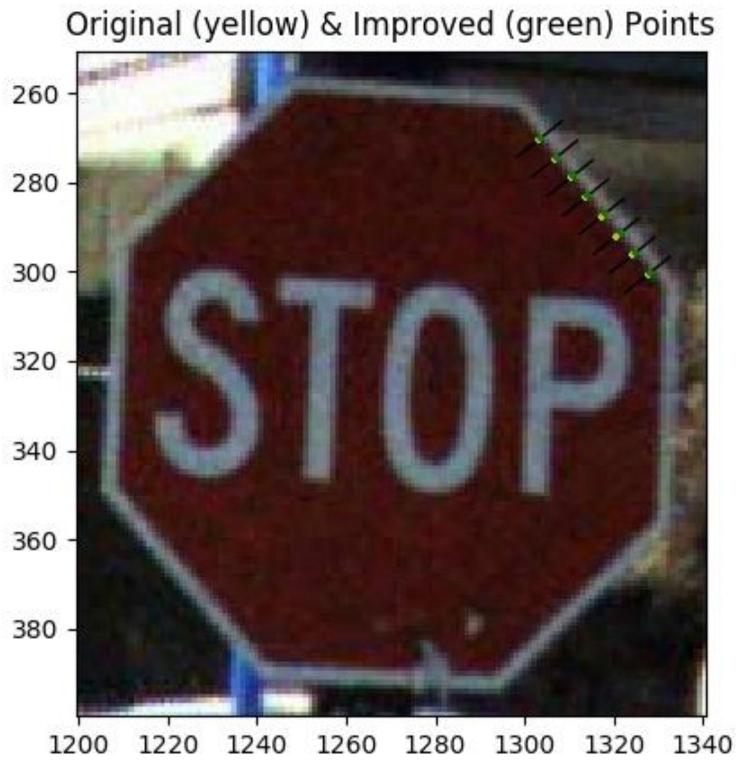


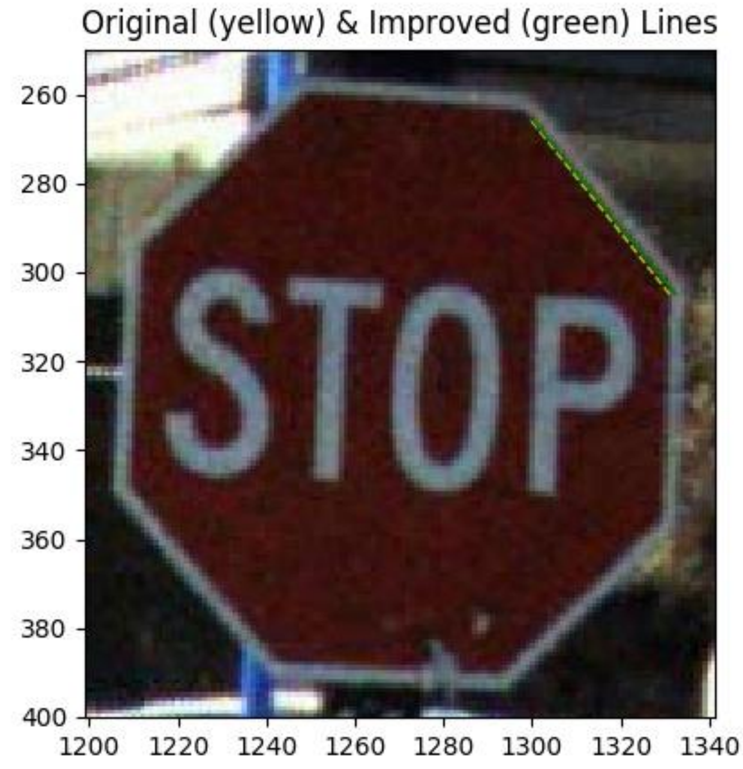
Line Improvement Method - Steps



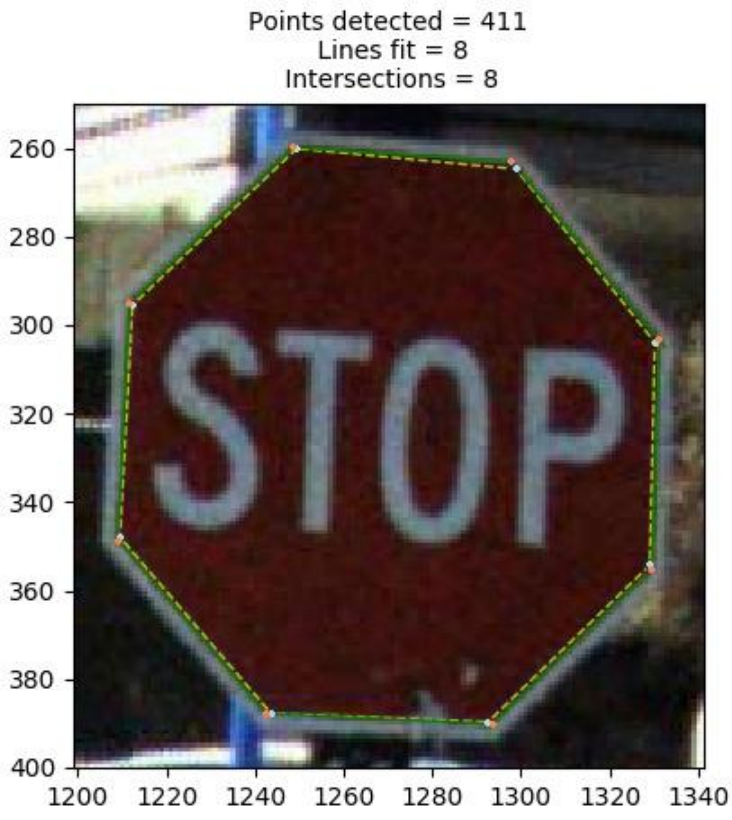


Line Improvement Method - Steps

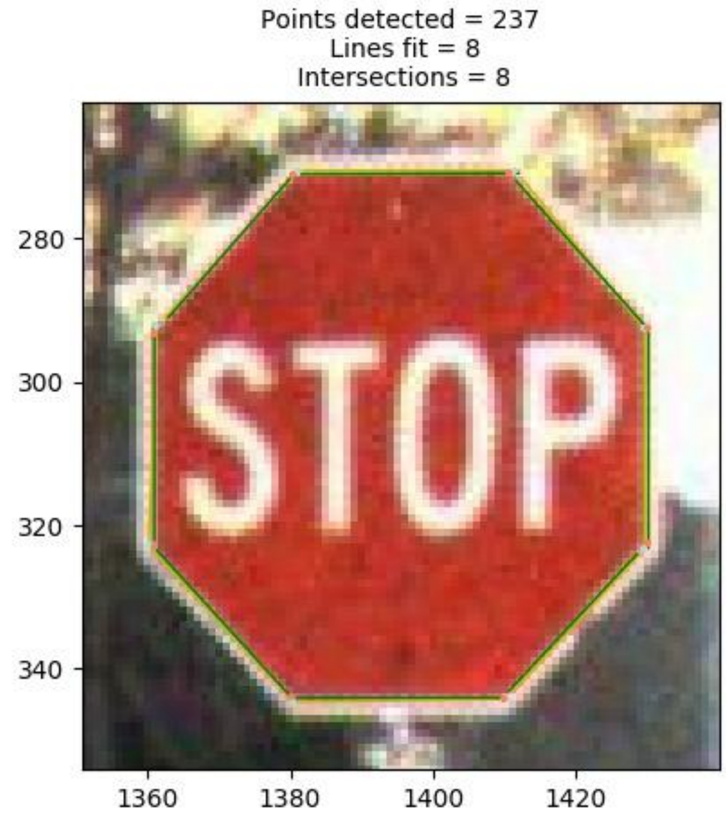




Line Improvement Method - Results

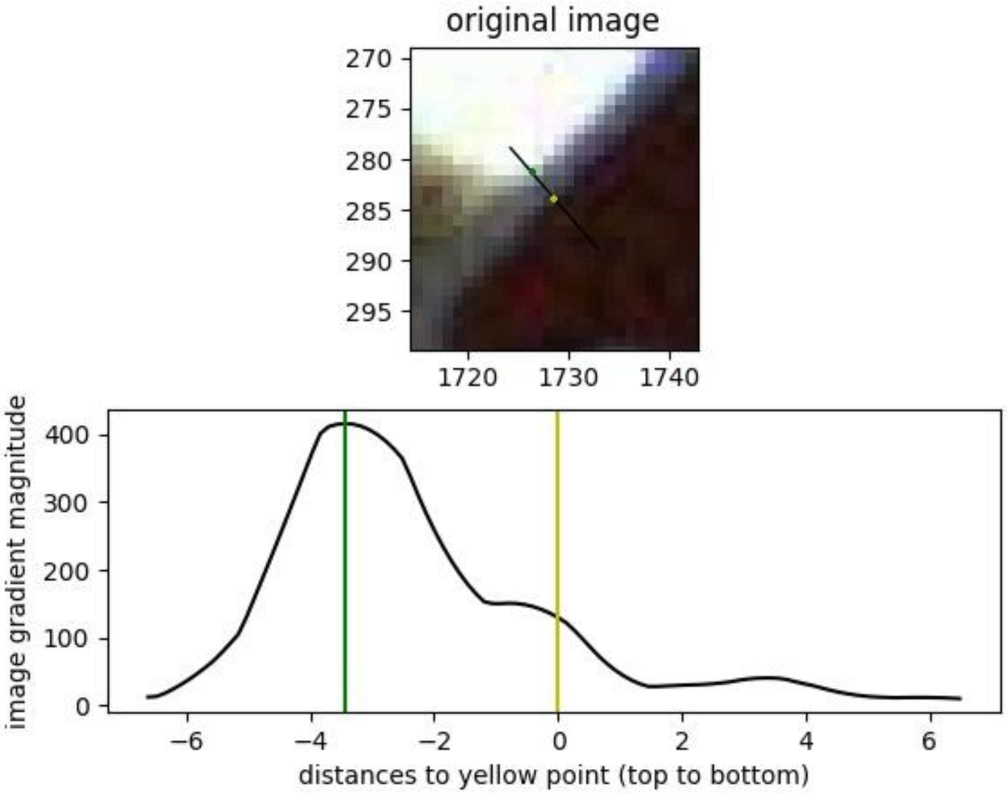


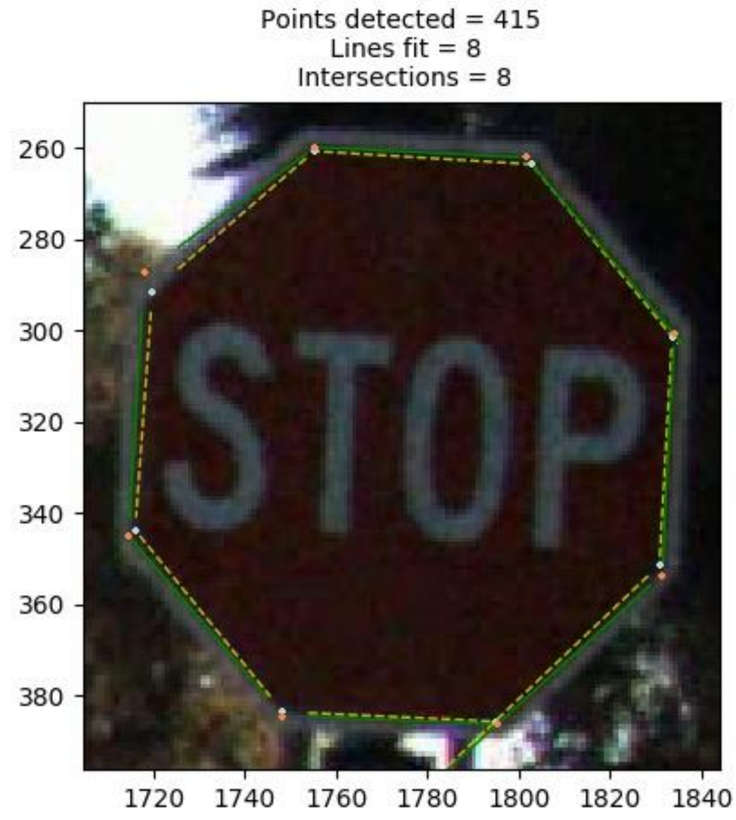
Example 1



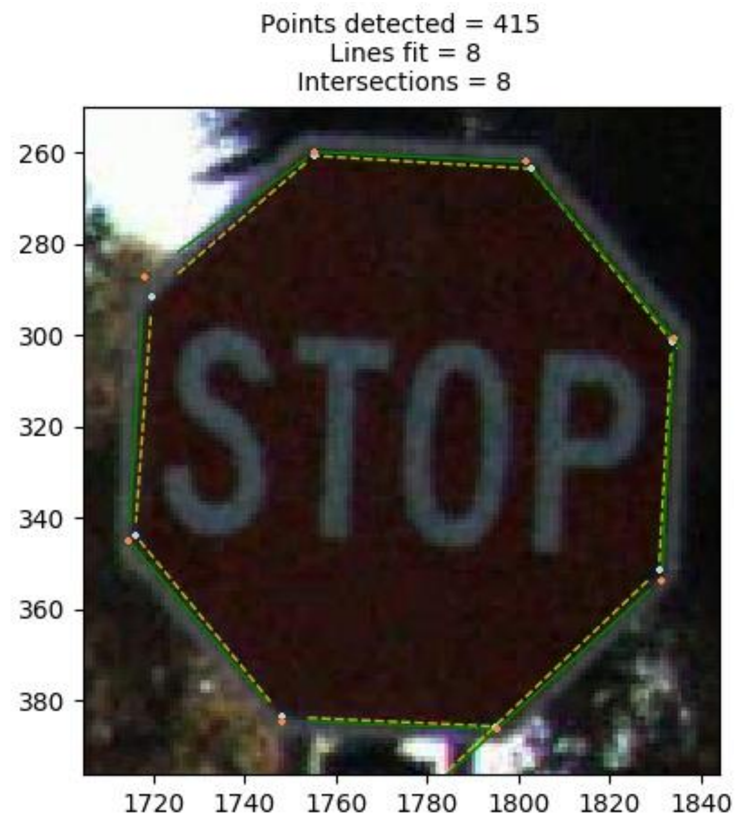
Example 2

Line Improvement Method - Problem

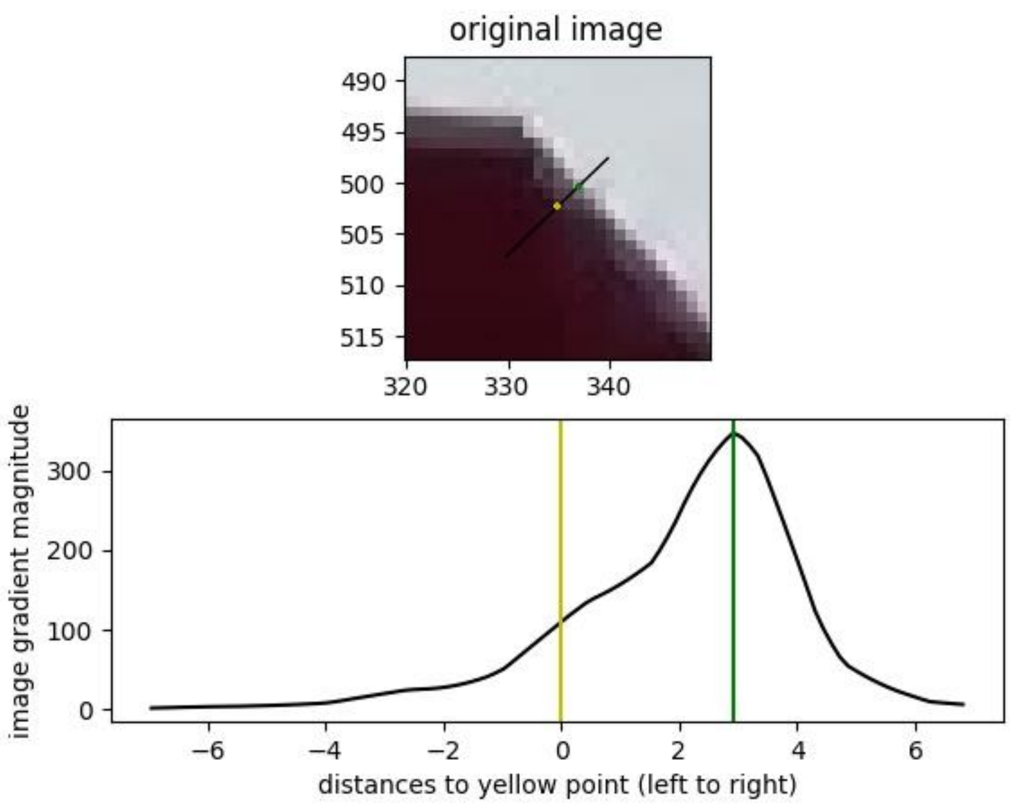


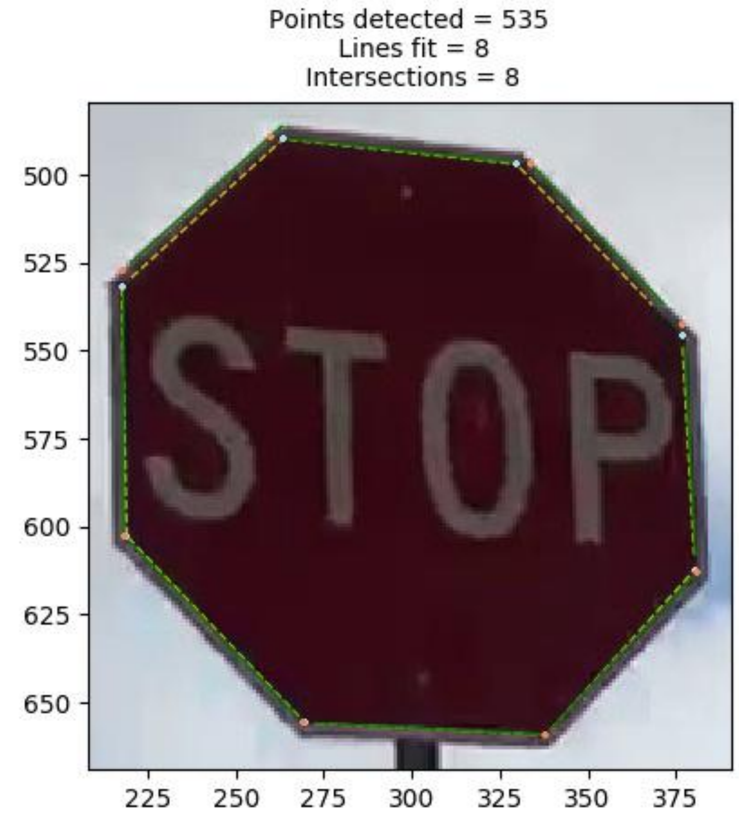


Line Improvement Method - Problem



Line Improvement Method - Problem





Problems Unsolved & Future Works

1. Experiment on More Datasets
   - AVL Dataset
   - Waymo Open Dataset
2. Stop Sign Detector
3. Improvement Method 2
   combine sequential info from video frames
4. Manual Bad Result Removal -> Automatic Algorithm

Improvement 2 - Bundle Adjustment

Here we list some basic transform that are used in the derivatives:

$$ {}^{c_2}_{w}T = {}^{c_2}_{c_1}T \, {}^{c_1}_{w}T \tag{1} $$

Where, ${}^{c_2}_{w}T = \begin{bmatrix} {}^{c_2}_{w}R & {}^{c_2}P_w \\ 0 & 1 \end{bmatrix}$, ${}^{c_2}_{c_1}T = \begin{bmatrix} {}^{c_2}_{c_1}R & {}^{c_2}P_{c_1} \\ 0 & 1 \end{bmatrix}$, ${}^{c_1}_{w}T = \begin{bmatrix} {}^{c_1}_{w}R & {}^{c_1}P_w \\ 0 & 1 \end{bmatrix}$.

In most cases, the sensors can record the relative pose from the second frame c_2 to the first frame c_1. Hence, instead of using ${}^{c_2}_{c_1}T$, ${}^{c_1}_{c_2}T$ is the more practical information we can directly obtain from sensors. The relation between the two transform is simple:

$$ {}^{c_2}_{c_1}T = {}^{c_1}_{c_2}T^{-1} = \begin{bmatrix} {}^{c_1}_{c_2}R^T & -{}^{c_1}_{c_2}R^T \, {}^{c_1}P_{c_2} \\ 0 & 1 \end{bmatrix} \tag{2} $$

Hence, the equation (1) can be denoted as:

$$ \underbrace{\begin{bmatrix} {}^{c_2}_{w}R & {}^{c_2}P_w \\ 0 & 1 \end{bmatrix}}_{\text{unknown}} = \underbrace{\begin{bmatrix} {}^{c_1}_{c_2}R^T & -{}^{c_1}_{c_2}R^T \, {}^{c_1}P_{c_2} \\ 0 & 1 \end{bmatrix}}_{\text{known}} \underbrace{\begin{bmatrix} {}^{c_1}_{w}R & {}^{c_1}P_w \\ 0 & 1 \end{bmatrix}}_{\text{unknown}} \tag{3} $$

Finally, we can obtain the following equations that are useful.

$$ \begin{cases} {}^{c_2}_{w}R = & {}^{c_1}_{c_2}R^T \, {}^{c_1}_{w}R \\ {}^{c_2}P_w = & {}^{c_1}_{c_2}R^T \, {}^{c_1}P_w - {}^{c_1}_{c_2}R^T \, {}^{c_1}P_{c_2} \end{cases} \tag{4} $$

Improvement 2 - Bundle Adjustment

In the original paper, only the pose information with respect to each frame has bee used for building constraints. And then, we can use DLT method to solve the equations of system.

Before we study how to build new constraints, we first introduce some basic equations:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = sK \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = sK \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \tag{7}$$

By abuse of notation, we still use $\mathbf{Q}$ to denote a point on the model plane, but $\mathbf{Q} = [X, Y]^T$ since Z is always equal to zero. In turn, $\overline{\mathbf{M}} = [X, Y, 1]^T$. Therefore, a model point $\mathbf{M}$ and its image $\mathbf{m}$ is related by a homography $\mathbf{H}$:

$$\tilde{\mathbf{m}} = \mathbf{H}\overline{\mathbf{M}} \quad \text{with} \quad \mathbf{H} = sK [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}] \tag{8}$$

Given an image of the model plane, an homography can be estimated by pairs of matched points. Let's denote it by $\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 \end{bmatrix}$. From (8), we have

$$\begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 \end{bmatrix} = sK \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix}, \tag{9}$$

$$\begin{aligned} h_1 &= sKr_1 \quad \text{or} \quad r_1 = \lambda K^{-1} h_1 \\ h_2 &= sKr_2 \quad \text{or} \quad r_2 = \lambda K^{-1} h_2 \\ h_3 &= sKt \quad \text{or} \quad t = \lambda K^{-1} h_3 \end{aligned} \tag{10}$$

Improvement 2 - Bundle Adjustment

where s, λ is an arbitrary scalar. Using the knowledge that r_1 and r_2 are orthonormal($r_1^{\mathrm{T}} r_2 = 0$ and $\|r_1\| = \|r_2\|$, that is $r_1^{\mathrm{T}} r_1 = r_2^{\mathrm{T}} r_2$), we have

$$\begin{aligned} \mathrm{h}_1^T K^{-T} K^{-1} \mathrm{h}_2 = 0 \\ \mathrm{h}_1^T K^{-T} K^{-1} \mathrm{h}_1 = \mathrm{h}_2^T K^{-T} K^{-1} \mathrm{h}_2 \end{aligned} \tag{11}$$

The traditional method can build two constraints for each frame. It dose not require that the relative poses between two sequential frames are known. In this special application, we study how to add new constraints by considering relative poses.

4.3.1 New constraints

$$\begin{aligned} {}^{c1}_{w}R &= \begin{bmatrix} \mathbf{r}_{11} & \mathbf{r}_{12} & \mathbf{r}_{13} \end{bmatrix} \\ {}^{c2}_{w}R &= {}^{c1}_{c2}R^T\, {}^{c1}_{w}R = {}^{c1}_{c2}R^T \begin{bmatrix} \mathbf{r}_{11} & \mathbf{r}_{12} & \mathbf{r}_{13} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_{21} & \mathbf{r}_{22} & \mathbf{r}_{23} \end{bmatrix} \end{aligned} \tag{12}$$

Hence, for the first frame:

$$\begin{bmatrix} \mathbf{h}_{11} & \mathbf{h}_{12} & \mathbf{h}_{13} \end{bmatrix} = s_1 K \begin{bmatrix} \mathbf{r}_{11} & \mathbf{r}_{12} & {}^{c_1}P_w \end{bmatrix} \tag{13}$$

and for second frame:

$$\begin{aligned} \begin{bmatrix} \mathbf{h}_{21} & \mathbf{h}_{22} & \mathbf{h}_{23} \end{bmatrix} &= s_2 K \begin{bmatrix} \mathbf{r}_{21} & \mathbf{r}_{22} & {}^{c_2}P_w \end{bmatrix} \\ &= s_2 K \begin{bmatrix} {}^{c_1}_{c_2}R^T \mathbf{r}_{11} & {}^{c_1}_{c_2}R^T \mathbf{r}_{12} & {}^{c_1}_{c_2}R^T\, {}^{c_1}P_w - {}^{c_1}_{c_2}R^T\, {}^{c_1}P_{c_2} \end{bmatrix} \end{aligned} \tag{14}$$

Improvement 2 - Bundle Adjustment

Right now, we have:

$$\begin{aligned} \mathrm{r}_{11} &= \lambda_1 K^{-1} h_{11} \\ \mathrm{r}_{12} &= \lambda_1 K^{-1} h_{12} \\ \mathrm{r}_{11}^{\mathrm{new}} &= \lambda_2 {}^{c_1}_{c_2}R K^{-1} h_{21} \\ \mathrm{r}_{12}^{\mathrm{new}} &= \lambda_2 {}^{c_1}_{c_2}R K^{-1} h_{22} \end{aligned} \tag{18}$$

The possible new constraints are:${\mathrm{r}_{11}^{\mathrm{new}}}^T \mathrm{r}_{12} = 0$ or ${\mathrm{r}_{11}^{\mathrm{new}}}^T \mathrm{r}_{11}^{\mathrm{new}} = {\mathrm{r}_{11}}^T \mathrm{r}_{11}$(the same for $\mathrm{r}_{12}^{\mathrm{new}}$ and r_{12}).

The question is:**How to use them for computation?**

1. ${\mathrm{r}_{11}^{\mathrm{new}}}^T \mathrm{r}_{12} = 0$

   The equation can be written as:

$$h_{21}^T K^{-T} {}^{c_1}_{c_2}R^T K^{-1} h_{12} = 0 \tag{19}$$

${}^{c_1}_{c_2}R^T$ is in the middle of $K^{-T}K^{-1}$ and that would make it difficult to solve this linear equations(it is not a symmetric matrix any more).

The symmetric matrix $K^{-T}K^{-1}$ actually describes the image of the absolute conic and since it is symmetric, after a few steps of derivatives, we can obtain:

Improvement 2 - Bundle Adjustment

$$\mathbf{B} = \mathbf{K}^{-T}\mathbf{K}^{-1} \equiv \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix}$$
$$= \begin{bmatrix} \frac{1}{f_x^2} & 0 & \frac{-c_x}{f_x^2} \\ 0 & \frac{1}{f_y^2} & -\frac{c_y}{f_y^2} \\ \frac{-c_x}{f_x^2} & -\frac{c_y}{f_y^2} & \frac{c_x^2}{f_x^2} + \frac{c_y^2}{f_y^2} + 1 \end{bmatrix} \tag{20}$$

Note that B is symmetric, defined by a 6D vector

$$\mathbf{b} = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T \tag{21}$$

Let the i th column vector of $\mathbf{H}$ be $\mathbf{h}_i = [h_{i1}, h_{i2}, h_{i3}]^T$. Then, we have

$$\mathbf{h}_i^T \mathbf{B} \mathbf{h}_j = \mathbf{v}_{ij}^T \mathbf{b} \tag{22}$$

with

$$\mathbf{v}_{ij} = [h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i2}h_{j2}, h_{i3}h_{j1} + h_{i1}h_{j3}, h_{i3}h_{j2} + h_{i2}h_{j3}, h_{i3}h_{j3}]^T \tag{23}$$

Therefore, the two equations in (11) can be rewritten as two homogeneous equations in b:

$$\begin{bmatrix} \mathbf{v}_{12}^T \\ (\mathbf{v}_{11} - \mathbf{v}_{22})^T \end{bmatrix} \mathbf{b} = \mathbf{0} \tag{24}$$

Improvement 2 - Bundle Adjustment

2. $\mathbf{r}_{11}^{\text{new}\,T}\mathbf{r}_{11}^{\text{new}} = \mathbf{r}_{11}{}^{T}\mathbf{r}_{11}$

The equation can be written as:

$$\begin{aligned} \lambda_2^2 h_{21}^T K^{-T}\, {}^{c_1}_{c_2}R^T\, {}^{c_1}_{c_2}R K^{-1} h_{21} &= \lambda_1^2 h_{11}^T K^{-T} K^{-1} h_{11} \\ \lambda_2^2 h_{21}^T K^{-T} K^{-1} h_{21} &= \lambda_1^2 h_{11}^T K^{-T} K^{-1} h_{11} \end{aligned} \tag{25}$$

Similarly($\mathbf{r}_{12}^{\text{new}\,T}\mathbf{r}_{12}^{\text{new}} = \mathbf{r}_{12}{}^{T}\mathbf{r}_{12}$),

$$\lambda_2^2 h_{22}^T K^{-T} K^{-1} h_{22} = \lambda_1^2 h_{12}^T K^{-T} K^{-1} h_{12} \tag{26}$$

If so, the quotient of λ_1 by λ_2 can be considered as one unknown variable, which results in the equations system that only involves one extra unknown variable.

Besides, the most interesting result is that it does not need to know the relative pose between the two camera frames.

To do one step further, we can obtain:

$$\begin{aligned} \frac{\lambda_2^2}{\lambda_1^2} h_{22}^T K^{-T} K^{-1} h_{22} - h_{12}^T K^{-T} K^{-1} h_{12} = 0 \\ \frac{\lambda_2^2}{\lambda_1^2} h_{21}^T K^{-T} K^{-1} h_{21} - h_{11}^T K^{-T} K^{-1} h_{11} = 0 \end{aligned} \tag{27}$$

Improvement 2 - Bundle Adjustment

The only different part between 11 and 27 is that the quotient is unknown(for equations in 11, since it only involves single image, the scalars on both sides of the equations are the same, thus can be safely cancelled out). Here comes the tricky part:both scalars(λ_1 and λ_2) and intrinsic matrix K are unknown.

Our idea to deal with this problem is that:

1. Compute intrinsic matrix using previous algorithm
2. Plug the estimated intrinsic matrix into equation 27. Ideally, the quotient of λ_2^2 by λ_1^2 should be the same each equation. However, due to the noise, it can not be this case.
3. Hence, it would be interesting if we could modify the value of f_x and f_y to make the error as small as possible($^{upper}\frac{\lambda_2^2}{\lambda_1^2} - {}^{bottom}\frac{\lambda_2^2}{\lambda_1^2}$).
4. We could simply extend to the scenario that there are more than two sequential frames and the total error could be minimized.

We will first do the experiments to compare the error($\sum(^{upper}\frac{\lambda_2^2}{\lambda_1^2} - {}^{bottom}\frac{\lambda_2^2}{\lambda_1^2})$) when we use groundtruch and the estimated results.

If such property holds true, We also have to figure out how to modify f_x and f_y to reduce the error.

Improvement 2 - Filter

4.3.2 Filter method

From the other perspective, we can build the model as follow:

$$\begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = s_1 K \begin{bmatrix} \mathbf{r}_{11} & \mathbf{r}_{12} & {}^{c_1}P_w \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \tag{28}$$

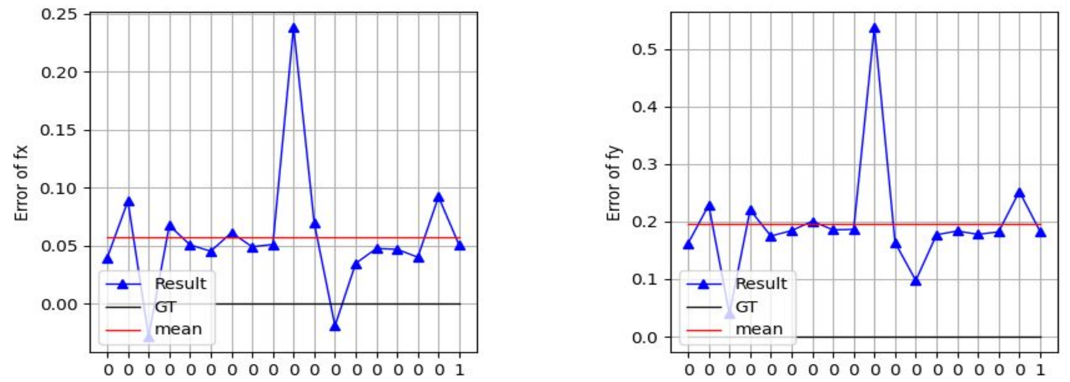
$$\begin{aligned}
\begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} &= s_2 K \begin{bmatrix} \mathbf{r}_{21} & \mathbf{r}_{22} & {}^{c_2}P_w \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \\
&= s_2 K {}^{c_1}_{c_2}R^T \begin{bmatrix} \mathbf{r}_{11} & \mathbf{r}_{12} & {}^{c_1}P_w - {}^{c_1}P_{c_2} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \\
&= s_2 K {}^{c_1}_{c_2}R^T \left(\begin{bmatrix} \mathbf{r}_{11} & \mathbf{r}_{12} & {}^{c_1}P_w \end{bmatrix} - \begin{bmatrix} 0 & 0 & {}^{c_1}P_{c_2} \end{bmatrix} \right) \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \\
&= s_2 K {}^{c_1}_{c_2}R^T \begin{bmatrix} \mathbf{r}_{11} & \mathbf{r}_{12} & {}^{c_1}P_w \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} - s_2 K {}^{c_1}_{c_2}R^T \begin{bmatrix} 0 & 0 & {}^{c_1}P_{c_2} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \\
&= \frac{s_2}{s_1} K {}^{c_1}_{c_2}R^T K^{-1} \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} - s_2 \begin{bmatrix} 0 & 0 & K {}^{c_1}_{c_2}R^T {}^{c_1}P_{c_2} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \\
&= \frac{s_2}{s_1} K {}^{c_1}_{c_2}R^T K^{-1} \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} - s_2 K {}^{c_1}_{c_2}R^T {}^{c_1}P_{c_2}
\end{aligned} \tag{29}$$

For the final expression, ${}^{c_1}_{c_2}R^T$, ${}^{c_1}P_{c_2}$ are known; $\begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix}$ can be estimated by previous algorithms; K is the camera intrinsic matrix and s_1, s_2 are two different unknown scalars.

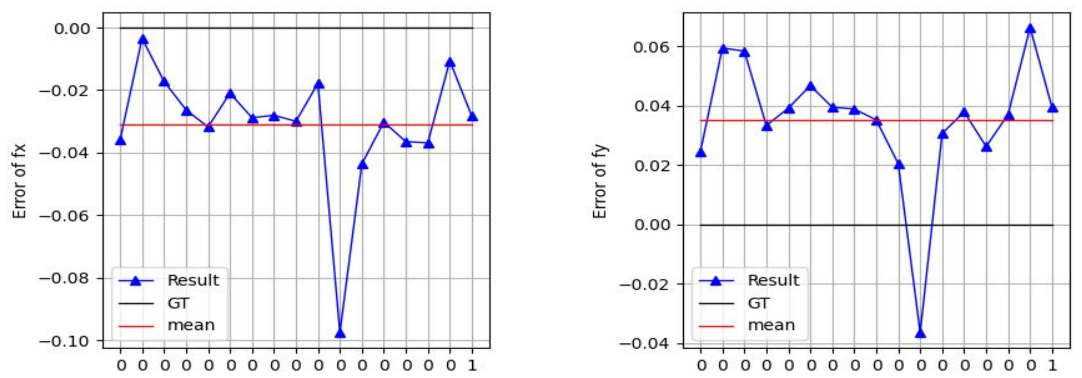
Problems:

1. Accuracy of relative pose estimation
2. Unsynchronized sensor data record

Manual Work



Bad frame/stop sign 3 removed (manually)



Thank you!