Report 1.0

This report contains the following

- Understanding the module better
- Gaining an understanding on the math concerned in this particular module
- Links and references which helped me start of or better understand the module

Understanding the module

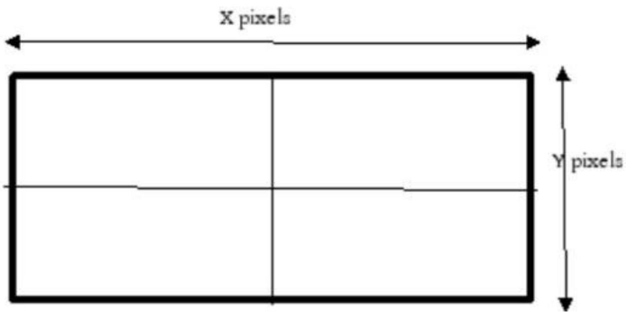
The GT Sensing vehicles which has a camera on the dashboard which faces straight. The outer world is of 3D and these 3D positions are projected onto 2D surfaces of a pinhole camera. This transformation is mathematical and can be better understood here (Prinicples of Pinhole Camera).

Before we look into the mathematics involved, one prerequisite is to understand the pinhole camera and explore the terminology associated with it. As per discussion the camera on the dashboard has 3-axis **x** (parallel to the image plane), **y** (the axis going upwards) , z (the axis in line with the optical axis). The camera has some intrinsic parameters f_x, f_y, C_x, C_y. These parameters are derived from camera calibration technique which can be done using the following tutorial (Camera Calibration using openCV). The end result of this calibration technique would be the parameters which helps to build the camera matrix (C_m).

The C_m allows to perform projections from the 3D space to 2D and vice versa. To calculate the size (height, length, width) of objects from a 2D image, the following approach may be used. Find the size of the object in the image and then scale it up using a scaling factor. This scaling factor in general is the pixels per meter in the respective axis.

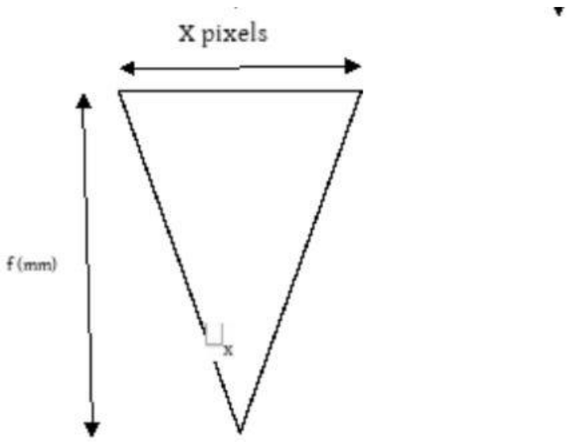
Mathematics concerned with it (Appendix A explained in detail):

Front View:



Assume a sign board whose **image size** is x-pixels * y-pixels

From top View:



$$\frac{X}{2}\text{pixels} = f\tan\left(\frac{\square}{2}\right)\text{ mm}$$

$$1 \text{ pixel in x direction} = \frac{2f\tan\left(\frac{\square_x}{2}\right)}{X}\text{ mm}$$

Distance from center in a projection f mm away.

Using trigonometry, we know the tan of an angle is opposite/adjacent. So as discussed, we would be needing to know who much each pixel would compare against the image size(in mm). Here 1 pixel in x direction is equivalent to how many mm (on the image) is found, similarly from another project we would need to find 1 pixel in y direction is how many mm.

Now to relate to the actual dimensions of the object in the 3D space. Assume we have 'l' 'w' and 'h' as the corresponding measurements in real. Y(numerical) pixels correspond to height 'h'. Using properties of similar angle triangles, we have the following.

h/l = Ypixels/f_{mm}

An important conversion here would be Y pixels = y (the numerical number of pixels) * pixels per mm in the y direction. This has been calculated above. For example 23pixels would be 23 * ppx/mm in the y direction.

This allows us to establish a relation between h and l (A). Similar procedure in another perspective would allow us to generate a relation between w and l (B). from (A) and (B) we also get a relation between w and h.

Before the capture of a second image let the car move by a distance of 'd'. let x_1 and y_1 correspond the number of pixels the sign occupies before moving distance 'd' and let x_2 and y_2 correspond to the number of pixels the sign occupies after moving distance 'd'. Trigonometry allows us to establish a relation between 'l', 'd', and other pixel values, which gives the actual dimensions.

$$\text{we have : } l = d\frac{x_1}{x_2 - x_1} \quad \text{and} \quad w = l\tan(\square_2) = d\frac{x_1}{x_2 - x_1}\,\frac{2x_2\tan\left(\frac{\square_x}{2}\right)}{X}$$

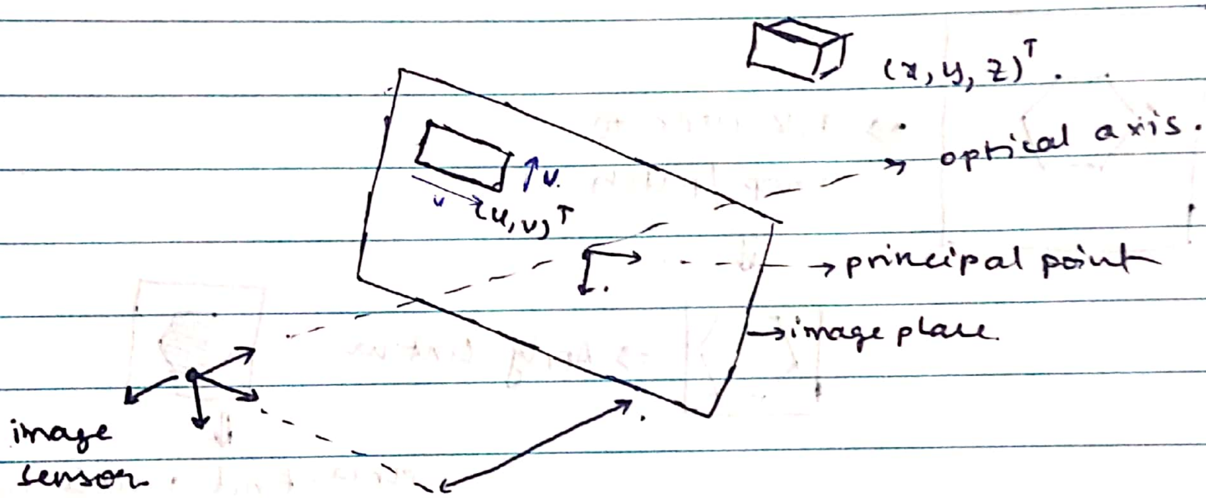
Notes with more detailed drawings

(4) PART-1. MATHEMATICS BEHIND PHOTOMETRY

- Sai Siddartha Maram.

(*) Understanding Camera [pinhole] Mathematically.

The regular pinhole camera projects a 3D plane onto a 2D surface. [inverted].



So, for a known focal length

[3D→2D] $u = \frac{Xf}{Z}$ and $v = \frac{Yf}{Z}$. [Some scaling factor may be applied].

CALIBRATION OF CAMERA:

Using openCV standard method to calibrate camera we get f_x, f_y, c_x, c_y and f

this allows us to form.

$$\text{Camera Matrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}.$$

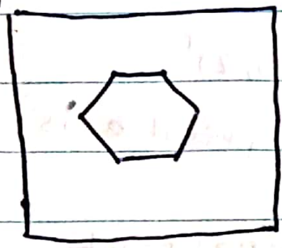
Answering each part at a time mathematically.

problem (4) x,y cordinates of a point

solution: User places cursor, extract x,y coordinates w.r.t Image. [Straightforward].

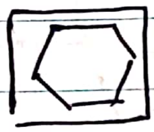
problem (3) The Area of a pothole (or) Patched area.

solution (3)

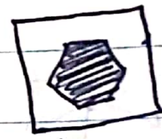


⇒ ask user to crop / patch

⇓



⇒ Apply contour



⇓

area = find area contour()

↙

area * scaling_factor.

problem (2) Height of a sign post

solution (2)

step 1: f_x, f_y, c_x, c_y and f.

step 2: Calculate pixels per millimeter (px/mm) on camera

step 3: Size of object in ~~mirror~~ camera

$$= \text{object\_size\_in\_pixels} / (px/mm)$$

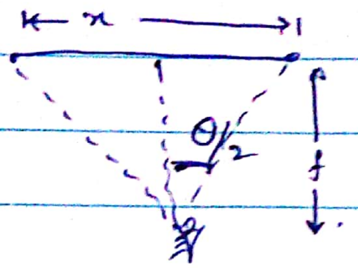
Step 4: Real size

$$= \frac{(\text{distance} * \text{size\_of\_object\_in\_image\_sensor})}{\text{focal length}}.$$

Understanding Appendix A and Problem 1.

→Assumption 1: tangential and Barrel Correction.

Calculating: the value of 1 pixel in → x axis



$\tan \theta/2 = \frac{x/2}{f}$.

$2f \tan \theta/2 = x$. ①

Similarly,

$2f \tan \theta/2 = y$.

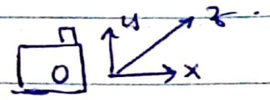
Using unitary method.

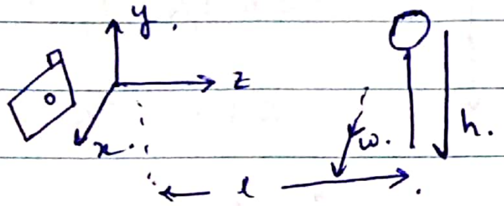
each pixel is

$1 = \frac{2f}{x} \tan \theta/2$

↳ in 'y' direction use y in the denominator

- The camera is set up on the dashboard of a vehicle.
- Looking ahead.
- We need where the sign touches the ground. w.r.t to camera
-



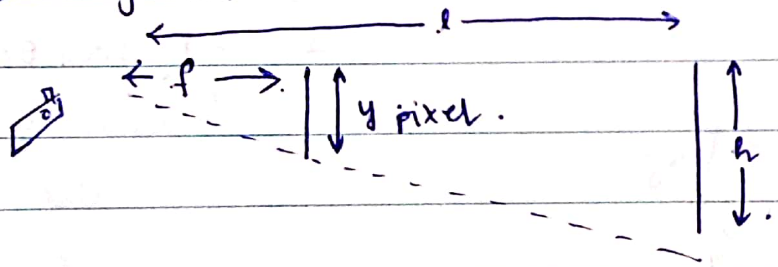


L, h, w are the values we want displacement of the sign. w.r.t to cameras.

in (+z) (-y) (+x).

l w h.

focal length 'f'

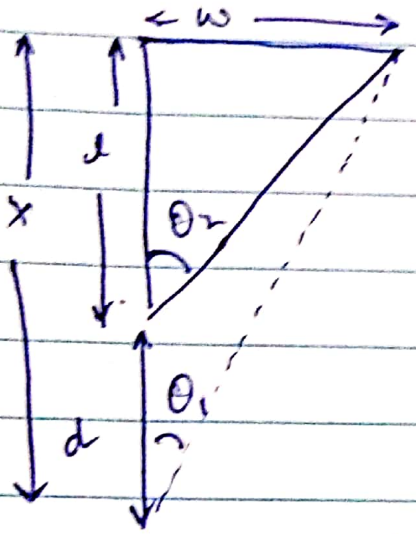


$$\frac{h}{d} = \frac{y\,\text{pixel}}{f} = \frac{y\left(\frac{2f\tan(\theta/2)}{Y}\right)}{f}$$

$$l = \frac{Yh}{2y\tan(\theta_y/2)}$$

$$W = \frac{2x\tan(\theta_x/2)}{X}$$

$$w = \frac{x\,Y \tan(\theta_x/2)\,h}{y\,X \tan(\theta_y/2)}.$$



$w/(l+d)$

$$\tan\theta_1 = \frac{2x_1 \tan(\theta_x/2)}{X}$$

$$\tan\theta_2 = \frac{2x_2 \tan(\theta_x/2)}{X}$$

$$l = d\,\frac{x_1}{x_2 - x_1}$$

$$w = l \tan\theta_2$$

$$= d\,\frac{x_1}{x_2 - x_1} \cdot \frac{2x_2 \tan\left(\frac{\theta_x}{2}\right)}{X}.$$