* File object\_detection\_zed\_live is the original file as provided in the stereolabs repository.
* File object\_detection\_zed\_modi is the modified file allowing us to specify the object class to be displayed during object detection. Currently, “person” has been specified in the program.

To be done:

* Saving the output in csv with real world coordinates(X,Y,Z), frame\_id, class, etc.
* Using this csv to run simple Kalman filter to provide tracking results (provide unique id to each person)

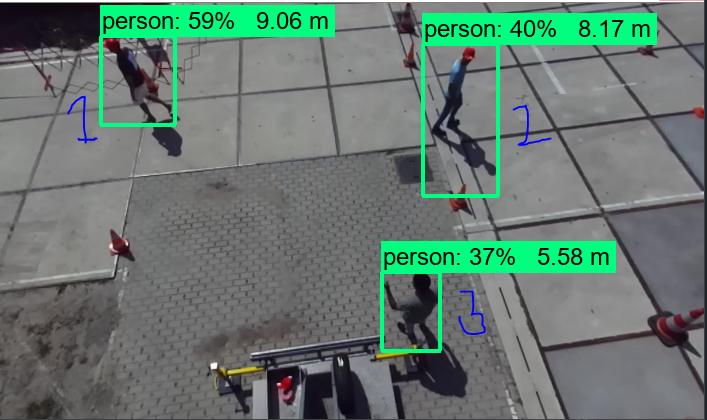
Working on object\_detection\_zed\_rishabh.py

* Difficult to save the output in a video as the loop is working over each frame and destroys the information from the previous frame. Thus, trying to use only one frame and understand the program and the values being passed during calculations.

Input file: frame\_3205.svo (only one frame, contains depth information)

08/08/19

* Running the .py file on one frame .svo file. It has three persons identified.



* Results from print(type(x)), print(x, y, z) (line159,160).

<class 'numpy.float64'>

-4.0493793 -2.4885128 7.7368784

<class 'numpy.float64'>

2.2758591175079346 -1.797508716583252 7.653509140014648

<class 'numpy.float64'>

0.8682291507720947 1.2250481843948364 5.382748126983643

* Results from print(type(x\_center)), print(x\_center, y\_ center) (line123,124)

<class 'int'>

137 77

<class 'int'>

460 116

<class 'int'>

411 308

Top left corner is the origin with x=width=704, y=height=416

* Print(type(ymin)), print(ymin, xmin, ymax, xmax)

<class 'float'>

0.08055010437965393 0.1426280438899994 0.2914063632488251 0.24879199266433716

<class 'float'>

0.0994887501001358 0.5999946594238281 0.46315687894821167 0.7085766792297363

<class 'float'>

0.6470780968666077 0.541461169719696 0.8345361351966858 0.6262728571891785

These values are raw coordinated of boxes (probably pixel coordinates?) with origin again at top left corner of the image. The global x\_max = 1.00, y\_max=1.00.

* print(type(distance)) print(distance) Provides distance of the object from the camera.

<class 'float'>

9.33555341073363

<class 'float'>

8.159998383628016

<class 'float'>

5.58368138988692

The distance is calculated as sqroot of x2+y2+z2. This is because x,y,z represent the coordinate values and the

* print (type(min\_y\_r)), print((min\_y\_r, min\_x\_r, max\_y\_r, max\_x\_r )) Line 135 at research\_distance\_box = 30

(47, 107, 107, 167)

<class 'int'>

(87, 430, 147, 490)

<class 'int'>

(278, 381, 338, 440)

<class 'int'>

At research\_distance\_box = 5

(72, 132, 82, 142)

<class 'int'>

(112, 455, 122, 465)

<class 'int'>

(303, 406, 313, 416)

<class 'int'>