

**Assignment #4**  
**10/30/2020**

**Prepared  
for  
Prof. Morato**

**Prepared  
by  
Abedin Sherifi**

## **Introduction:**

The main goal of this coding assignment was our exposure to Mono SLAM based on Extended Kalman Filter (EKF). I personally have been having a hard time understanding EKF especially the theory behind it. Nevertheless, applying EKF to monocular SLAM. However, Mono SLAM alone seemed doable to some degree. My efforts therefore are focused primarily on the Mono SLAM application to a video I took with my wife in Boston.

I will present in detail some of the algorithms I used in this coding assignment. One of the first algorithms used was the ORB detector. One of the tasks within SLAM is the localization of landmarks in our environment. The ORB detector is a very good detector in localizing such landmarks. The ORB detector is a very fast and accurate detector which is based out of the FAST algorithm for keypoint detection and on the BRIEF algorithm for calculation of descriptors. Once keypoints are located on a frame, descriptors are matched from the current frame to the previous frame using the Brute Force matcher.

One of the tasks of this homework assignment was to present the camera positions and the landmarks in a point cloud. What is a point cloud? A point cloud is a set of data points in 3D space. The 3D point cloud will consist of landmark data points as well as camera position points. Multiple frames are run in order to get depth information in 2D which assists in displaying the point cloud in 3D. One of the main tasks in Mono SLAM is to figure out the camera position. The camera angle could change frame to frame. One of the ways to figure this out is through the fundamental matrix. The fundamental matrix is used for uncalibrated cameras. The essential matrix is essentially the same thing but is used for calibrated cameras. The fundamental matrix is calculated through the use of the 8-point algorithm. The 8-point algorithm as the name suggests takes in 8 correspondence points and outputs a homogeneous linear equation. The fundamental matrix was computed through cv2.findFundamentalMat function of opencv.

Camera has two characteristics. It has intrinsic and extrinsic parameters. One of the main matrices is the intrinsic matrix. The intrinsic matrix is made of the image size and the focal length of the camera. My camera focal length is equal to 339. The extrinsic matrix consists of the rotation matrix and the translation vector. The camera matrix is the multiplication of the intrinsic matrix and the extrinsic matrix. The rotation and translation matrices are calculated through the use of the recoverPose function in opencv.

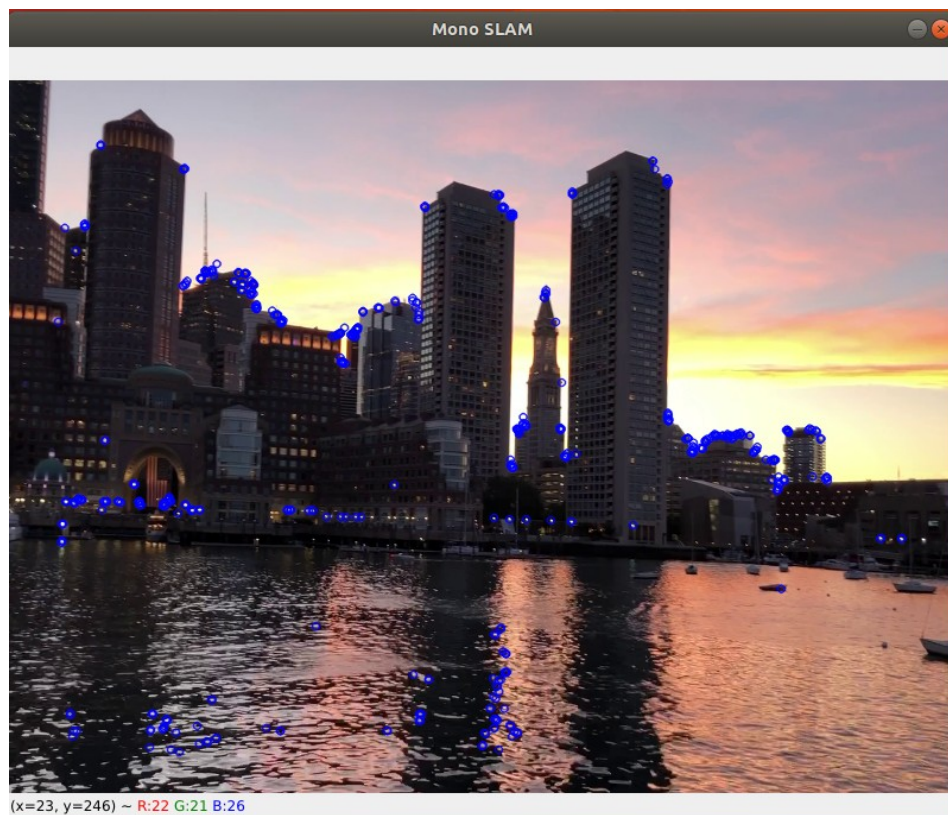
Initially we are going to get 2D points. In order to go from 2D point to 3D point, we need to multiple the camera matrix with the 2D point at time t.

## **Results:**

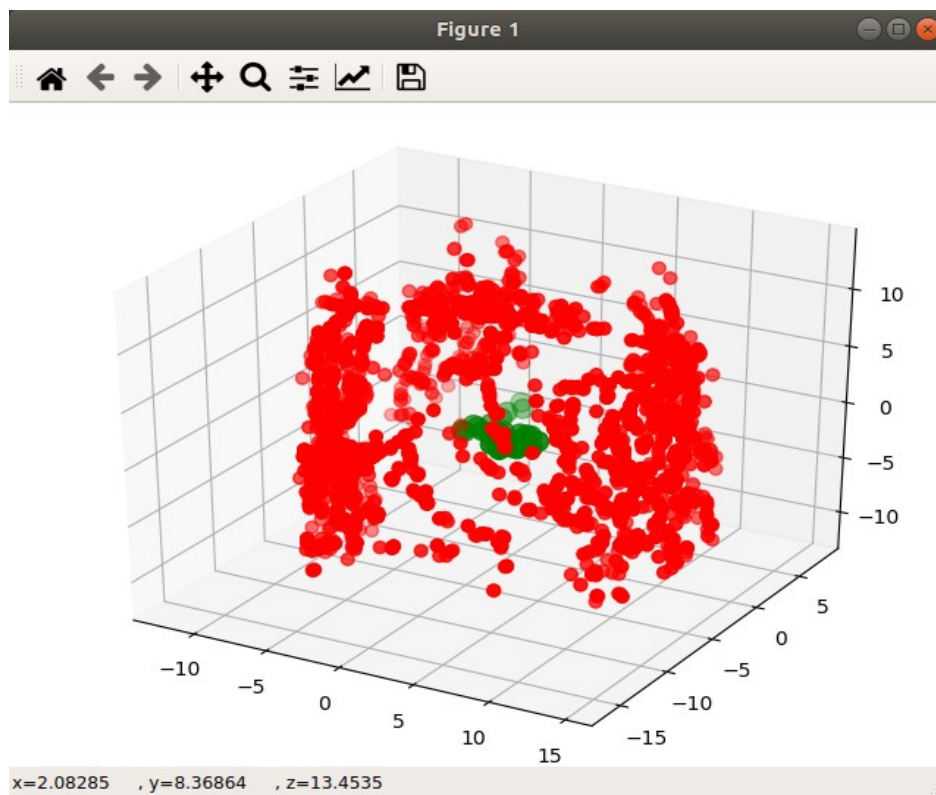
A snapshot of the 3D point cloud of landmarks and camera positions as well as ORB detector keypoints on the video capture and camera position points are given below:

```
Current Camera Location Per Frame: [0.00596461], [-0.61640987], [0.00088086]
Current Camera Location Per Frame: [0.26282509], [-1.58285607], [0.00298517]
Current Camera Location Per Frame: [0.1675244], [-2.57830456], [0.00268497]
Current Camera Location Per Frame: [-0.61082744], [-3.20613161], [0.0014178]
Current Camera Location Per Frame: [-0.74537988], [-4.19703805], [0.00114109]
Current Camera Location Per Frame: [-0.25165435], [-3.32742028], [0.00133573]
Current Camera Location Per Frame: [-0.50845977], [-4.2938832], [0.00071012]
Current Camera Location Per Frame: [0.30316399], [-3.70970345], [0.00165271]
Current Camera Location Per Frame: [1.28063291], [-3.49863055], [0.00330447]
Current Camera Location Per Frame: [2.11494322], [-4.04992428], [0.00454363]
Current Camera Location Per Frame: [2.49460823], [-3.1248004], [0.00401358]
Current Camera Location Per Frame: [2.75552651], [-4.09015708], [0.00686506]
Current Camera Location Per Frame: [2.93525589], [-3.106441], [0.00699459]
Current Camera Location Per Frame: [2.99558312], [-4.10461948], [0.0075871]
Current Camera Location Per Frame: [3.20794966], [-5.08180938], [0.00817108]
Current Camera Location Per Frame: [3.10987916], [-4.08663013], [0.00883518]
Current Camera Location Per Frame: [3.21876234], [-3.09257558], [0.0088325]
Current Camera Location Per Frame: [4.06041884], [-3.63258534], [0.0107811]
Current Camera Location Per Frame: [4.23408415], [-4.61738966], [0.01168007]
Current Camera Location Per Frame: [3.34549032], [-4.15871027], [0.00791]
Current Camera Location Per Frame: [3.57286348], [-5.13251775], [0.00857158]
Current Camera Location Per Frame: [2.74672687], [-5.69598651], [0.00745424]
Current Camera Location Per Frame: [1.87542141], [-5.20524551], [0.00769674]
Current Camera Location Per Frame: [1.50456799], [-4.2765541], [0.00779948]
```

**Figure 1. Current Camera Position Per Frame**



**Figure 2. ORB Detector Keypoints on a Video Frame**



**Figure 3. The 3D Point Cloud of Landmarks in Red and Camera Positions in Green**