Structure From Motion (SFM)

The camera intrinsic matrix given in file IntrinsicsMCV.mat was printed (not shown in .ipynb file)

```
print(intrinsic_matrix)

[[720.6605049    0.         420.22416865]
 [  0.         720.05791688 316.64872007]
 [  0.           0.           1.        ]]
```

In the above matrix the element at 0,0 is focal length of first camera and element at 1,1 is focal length of second camera. Elements at 0,2 and 1,2 are the offsets of both the cameras respectively.

We used Bundle Adjustment followed by least square optimizer to reduce the value of cost function for bundle adjustment given by:

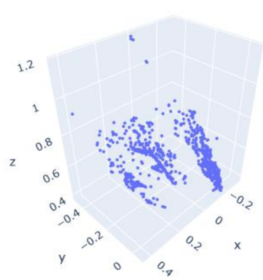
$$F = \sum_{i=1}^{N_c} \sum_{j=1}^{N_p} \|p_{ij} - \pi(R_i P_j + T_i)\|^2$$

We plotted sparse 3D point cloud for 5, 15 and 25 frames and observed them. We concluded following:

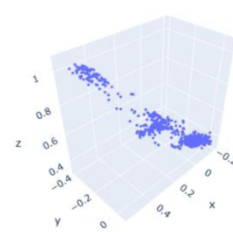
1) The depth values looks more tuned
2) The differences in depth of ‘G’ and ‘F’ figures looks more enhanced.

I am pasting the 3D point cloud of 5 frames and 25 frames for comparison purpose here:

With 5 frames



With 25 frames



PLANE SWEEP

We now use the camera rotation matrices and translation vectors obtained from the bundle adjustment along with the sequence provided in the shared assignment folder in a plane-sweeping framework.

It involves us to find homography. We know the formula for finding homography as :

$$H_{d,i} = K[R_i - (nT_i^T)/d]K^{-1}$$

As we are only concerned with the planes parallel to the reference frame, n=$[0,0,-1]^T$. K is the intrinsic matrix for the camera in above equation.

We plotted the 3d reconstruction with different number of frames and found that its quality improves as we increase the number of frames with more texture.