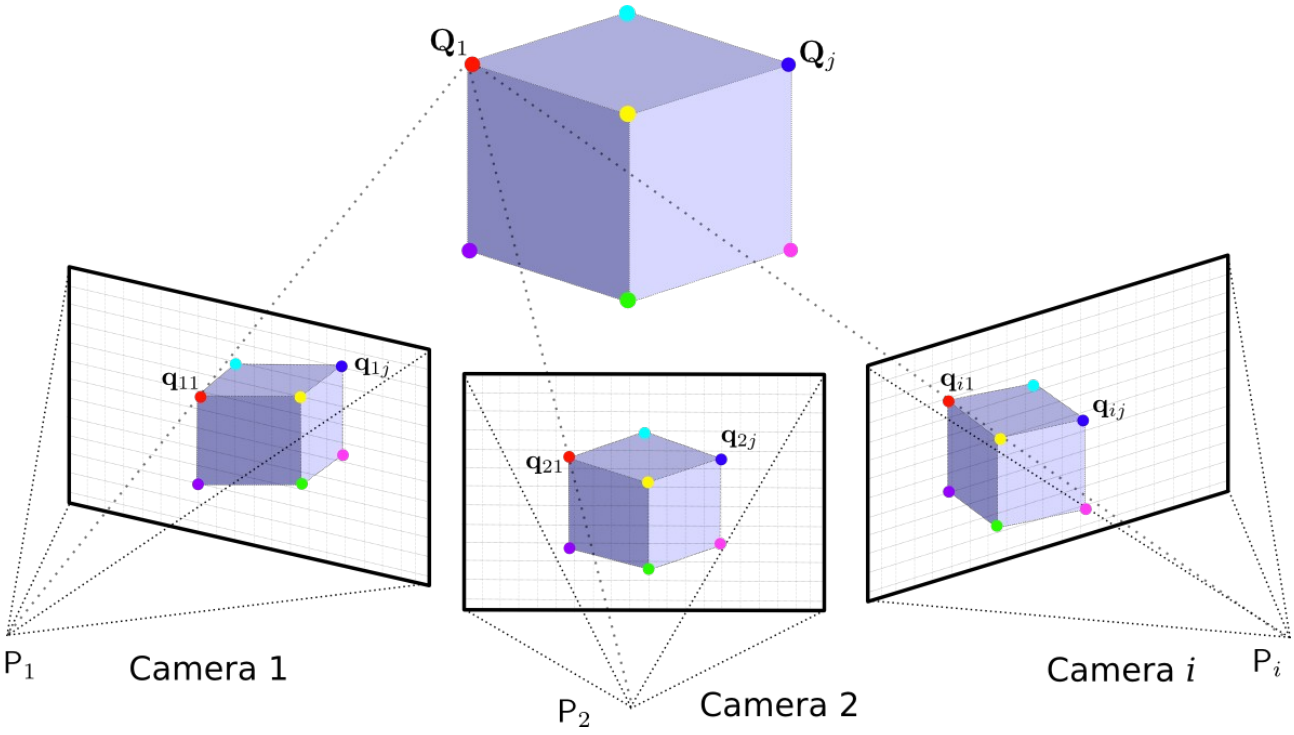
Homework 4
CSCI 677

Dixith Reddy Gomari
USC-ID:3098766483
gomari@usc.edu

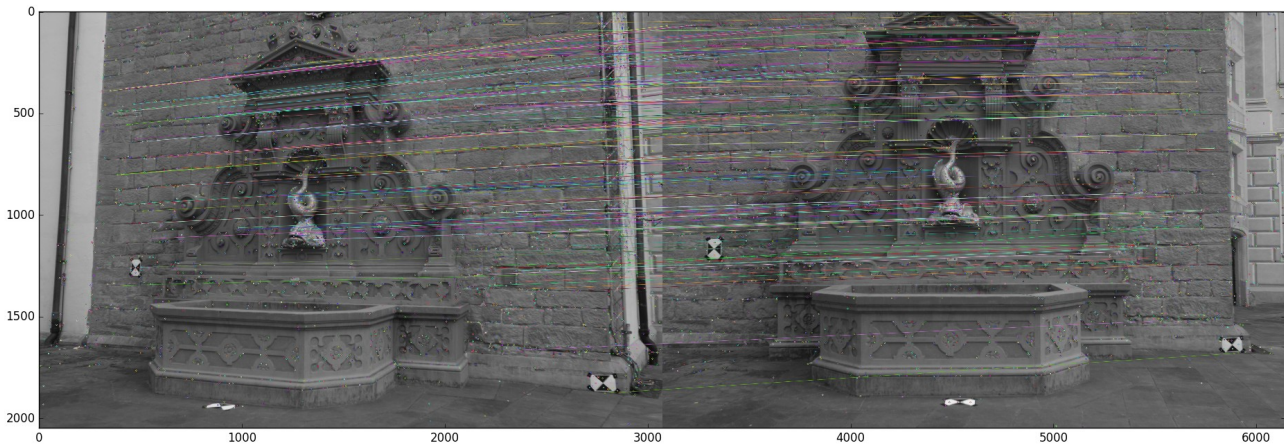
3D Reconstruction(SFM):

Structure from motion technique is used here for estimating a three dimensional structure (Water fountain in this case), from two 2-dimensional image sequence.



Results:

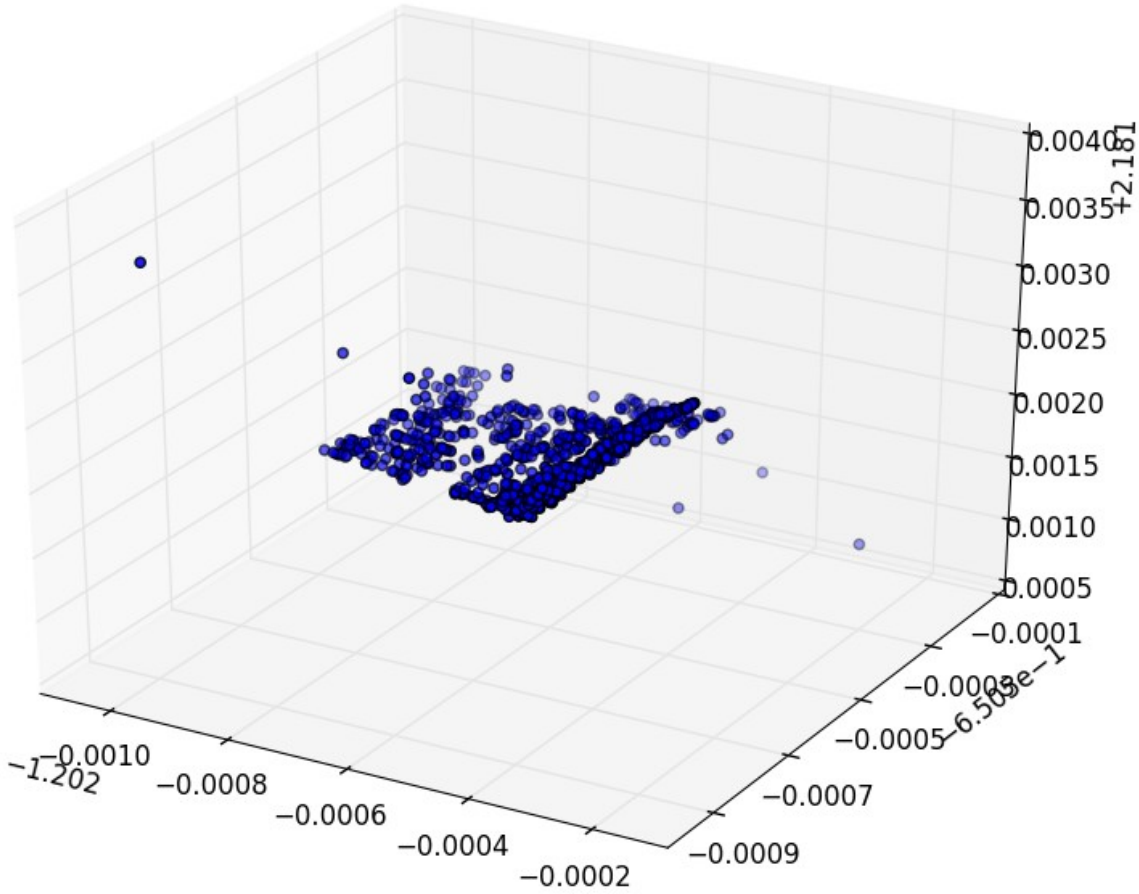
Original Images from 2 different cameras



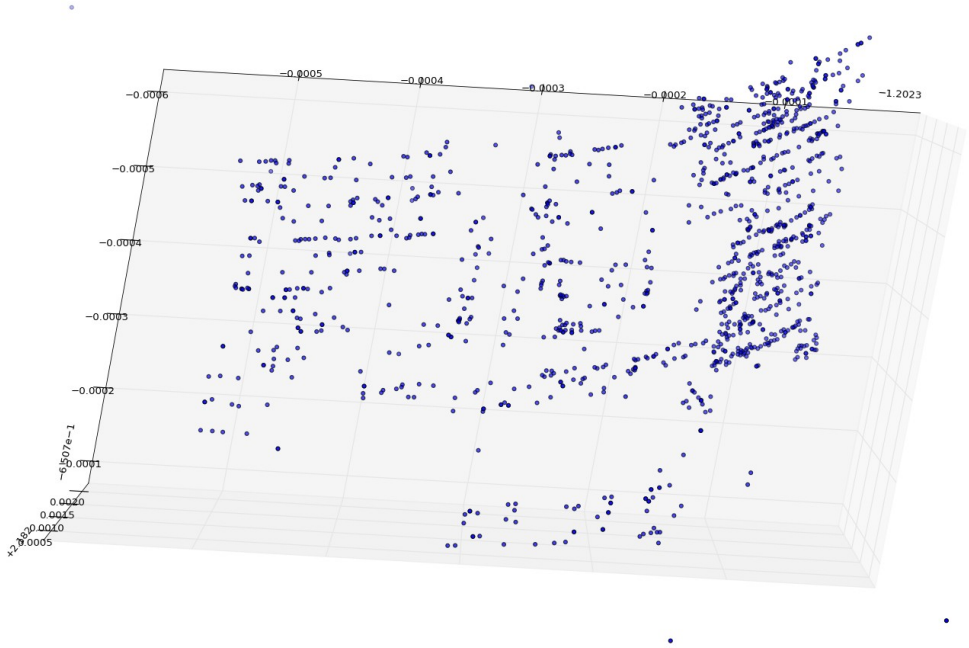
SIFT feature matching

Intermediate Results:

```
Rotation Matrix
[[ 0.93227329 -0.02683345 -0.36075821]
 [ 0.00960081  0.99872918 -0.04947573]
 [ 0.36162735  0.04266133  0.93134616]]
Translation Matrix
[[ 0.99681534]
 [ 0.02340763]
 [ 0.07623165]]
Camera Matrix for the first camera
[[  2.76000000e+03   0.00000000e+00   1.52000000e+03   0.00000000e+00]
 [  0.00000000e+00   2.76000000e+03   1.00600000e+03   0.00000000e+00]
 [  0.00000000e+00   0.00000000e+00   1.00000000e+00   0.00000000e+00]]
Camera Matrix for the Second camera
[[  3.12274785e+03  -9.21511174e+00   4.19953516e+02   2.86708244e+03]
 [  3.90295351e+02   2.79940984e+03   8.00381222e+02   1.41294085e+02]
 [  3.61627352e-01   4.26613305e-02   9.31346160e-01   7.62316468e-02]]
```



Point Cloud



Visualization after appropriate rotation of the point cloud

Discussion:

- To find structure from motion, the correspondence between images needs to be found.
- To find the correspondence between images, I have used SIFT and filtered out bad matches.
- I have used RANSAC to filter out the outlier correspondences.
- After generating the essential matrix, it can be used to get the rotation and trasnlation between the two cameras.
- To validate this result, we can verify by finding the rank of the essential matrix, which should be 2.
- Since the intrinsic parameters for both the cameras are the same, we need to generate the extrinsic matrix of the second camera with respect to the first camera.
- We can generate the camera matrix with these matrices.
- With these parameters, we can trinagulate points for visualization.
- We can generate a point cloud with a scatter plot, by appropriate rotation of this point cloud, we can clearly visualize the fountain.

Conclusion:

- This method works really well where the disparity between cameras is acceptable, that is if the images have sufficient number of matching points, we can easily reconstruct 3D images.
- 3D reconstruction is not possible with images, having insufficient number of matches.
- This method also fails, when there is huge amount of barrel or pin cushion distortion, which hinders the matching, resulting in failing of 3D reconstruction.
- This can be improved by using multple cameras by improving the probability of reconstruction.