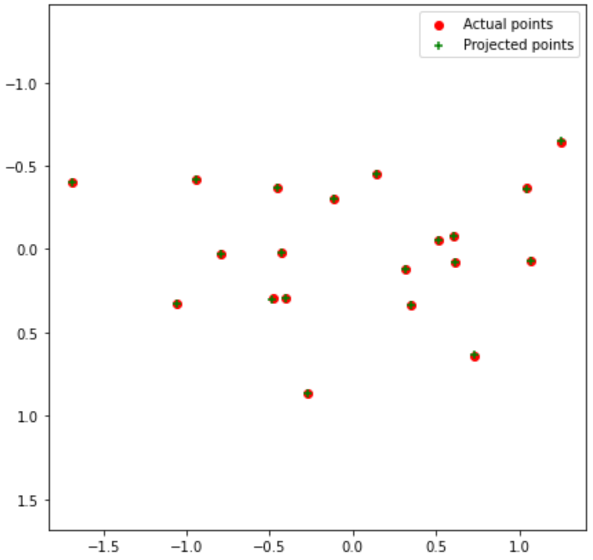
CS 5330 Programming Assignment 3

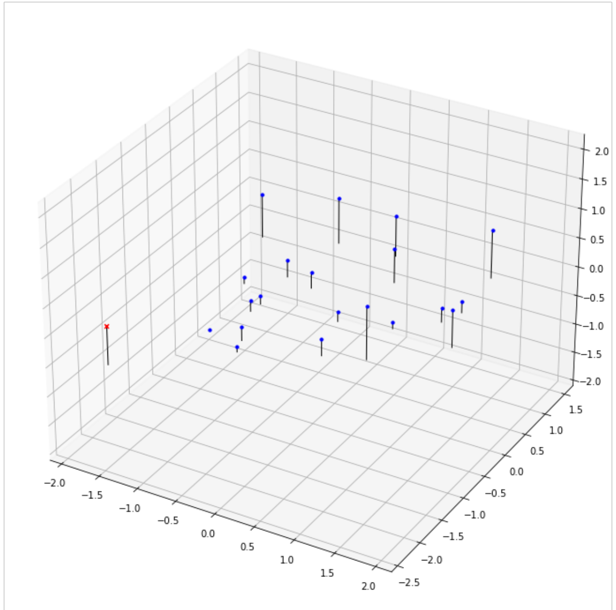
[Sai Nikhil]
[thirandas.s@northeastern.edu]
[thirandas.s]
[001564864]

Part 1: Projection matrix

[insert visualization of projected 3D points and actual 2D points for the CCB image we provided here]

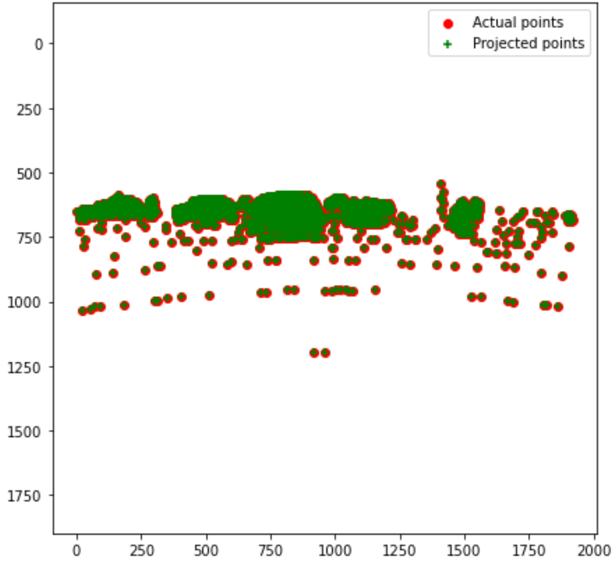


[insert visualization of camera center for the CCB image here]

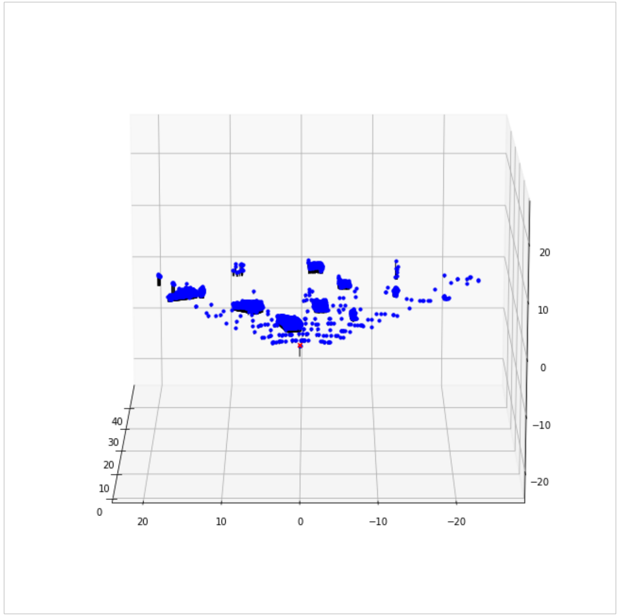


Part 1: Projection matrix

[insert visualization of projected 3D points and actual 2D points for the Argoverse image we provided here]



[insert visualization of camera center for the Argoverse image here]



Part 1: Projection matrix

[What two quantities does the camera matrix relate?]

Camera Matrix relates $\boldsymbol{p}_i$ and $\boldsymbol{X}_i$. $\boldsymbol{p}_i \equiv \boldsymbol{M}\boldsymbol{X}_i$.

$\boldsymbol{X}_i$ is the coordinate of the point in 3D world and $\boldsymbol{p}_i$ is the projection of $\boldsymbol{X}_i$ on 2D image plane.

[What quantities can the camera matrix be decomposed into?]

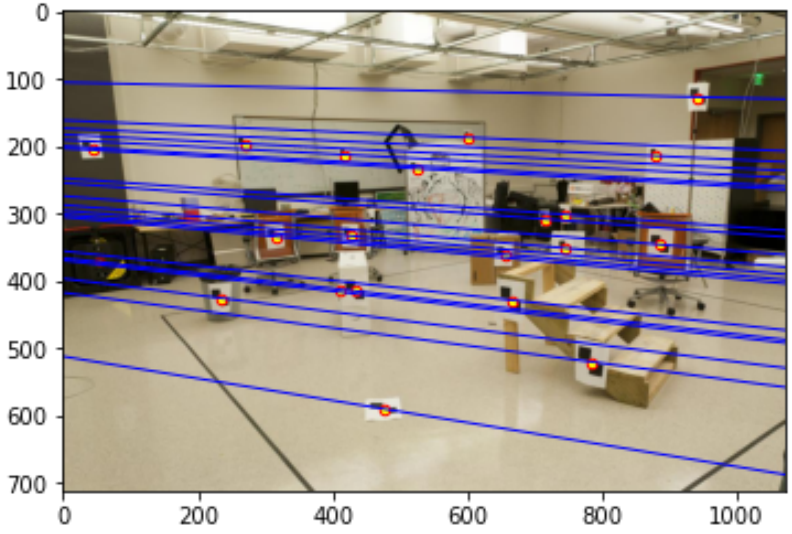
Camera matrix can be decomposed into intrinsic (focal length, offset, skewness) and extrinsic quantities (rotation and translation).

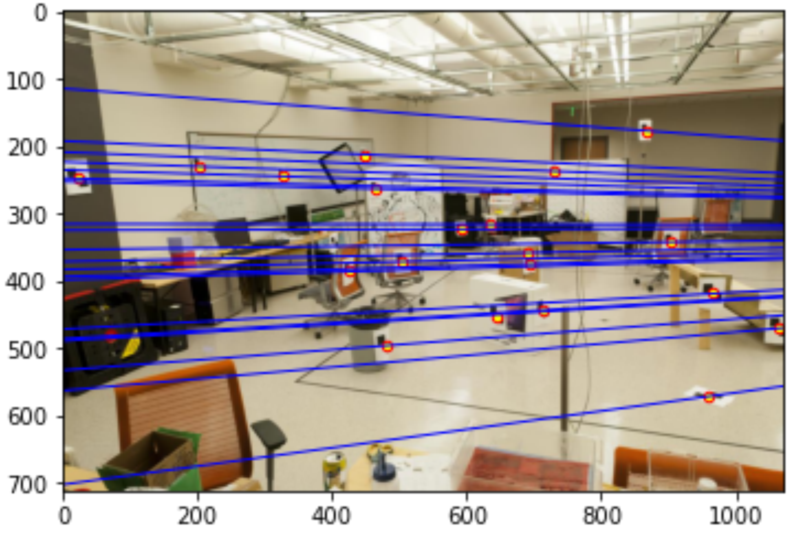
[List any 3 factors that affect the camera projection matrix.]

1) Focal length
2) Offset
3) Skewness
4) Rotation in 3D world
5) Translation in 3D world

Part 2: Fundamental matrix

[insert visualization of epipolar lines on the CCB image pair]





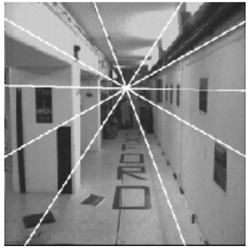
Part 2: Fundamental matrix

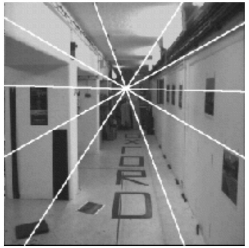
[Why is it that points in one image are projected by the fundamental matrix onto epipolar lines in the other image?]

We know that $\boldsymbol{p'^T F p} = 0$. Similarly,$\boldsymbol{p^T F^T p'} = 0$. This is of the form, $\boldsymbol{ax} = 0$. Which is the equation of a straight line. Also, the equation says that points of second image, $\boldsymbol{p'^T}$are related to $\boldsymbol{Fp}$, which is the equation of epipolar line of the first image image. Likewise, points of first image, $\boldsymbol{p^T}$are related to $\boldsymbol{F^T p'}$, which is the epipolar line of second image.

[What happens to the epipoles and epipolar lines when you take two images where the camera centers are within the images? Why?]

This case is similar to forward motion shown below.





In this case, epipole has same coordinates in both images.

Points move along lines radiating from the "Focus of expansion". Also, the epipolar lines are the same and concurrent at same point, the epipole.

Part 2: Fundamental matrix

[What does it mean when your epipolar lines are all horizontal across the two images?]

This case happens when the image planes of two cameras are parallel to each other. Epipolar lines fall along the horizontal scan lines of the images.

[Why is the fundamental matrix defined up to a scale?

The fundamental matrix is defined by the equation $\boldsymbol{p'^T F p} = 0$. But once we have an $\boldsymbol{F}$ that solves this equation for given set of pixels $(\boldsymbol{p'^T}, \boldsymbol{p})$, we can always multiply $\boldsymbol{F}$ by any scalar $\boldsymbol{a}$ and still solve the equation $\boldsymbol{p'^T aFp} = 0$, since we can factor $\boldsymbol{a}$ and eliminate it with the 0 on the other side. Thus, $\boldsymbol{F'} = \boldsymbol{aF}$ is also a valid fundamental matrix. This means fundamental matrix is defined up to a scale.

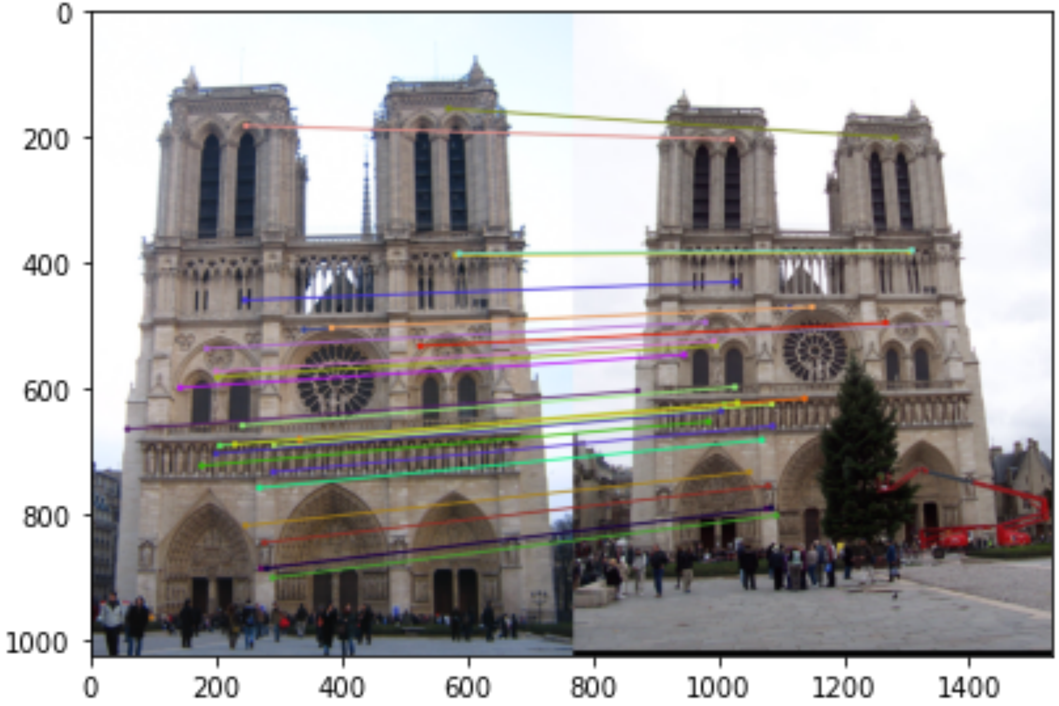
[Why is the fundamental matrix rank 2?]

We know that,

$\boldsymbol{E} = [\boldsymbol{t_x}]\boldsymbol{R}$, where $\boldsymbol{t_x} = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix}$. The third row can be obtained as a linear combination of the first 2 rows. Hence, rank of $\boldsymbol{E}$ is 2. We also know that $\boldsymbol{F} = \boldsymbol{K'^{-T} E K^{-1}}$. Since, $\boldsymbol{K'^{-T}}$ and $\boldsymbol{K^{-1}}$ are invertible matrices, they do not alter the rank of $\boldsymbol{F}$. Hence, $\boldsymbol{F}$ also has rank of 2.

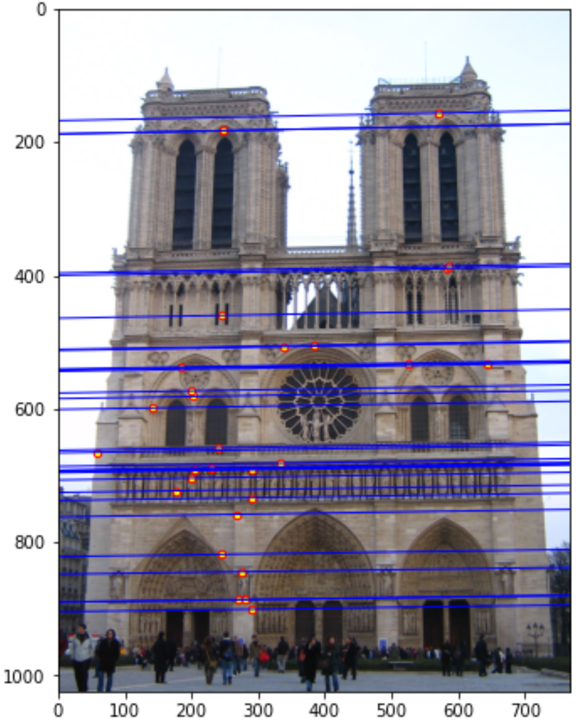
Part 3: RANSAC

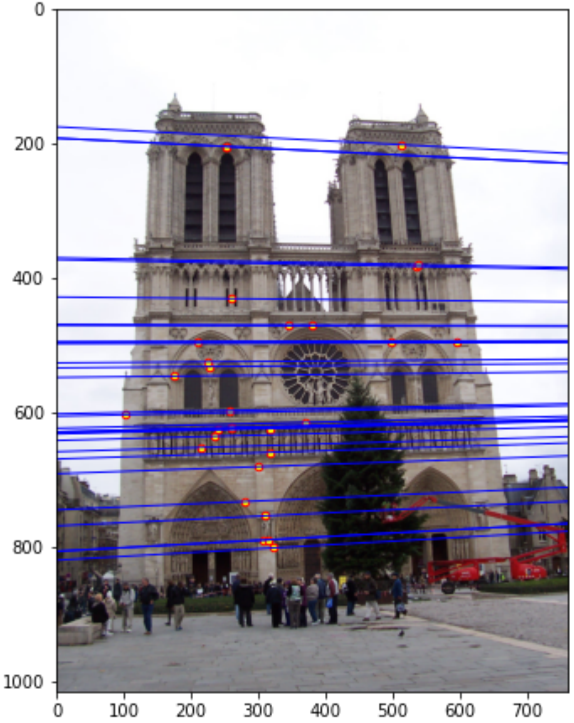
[insert visualization of correspondences on Notre Dame after RANSAC]



Part 3: RANSAC

[insert visualization of epipolar lines on the Notre Dame image pair]





Part 3: RANSAC

[How many RANSAC iterations would we need to find the fundamental matrix with 99.9% certainty from your Mt. Rushmore and Notre Dame SIFT results assuming that they had a 90% point correspondence accuracy?]

$$N = \left\lceil \log \frac{1-0.999}{1-0.9^8} \right\rceil = 13$$

[One might imagine that if we had more than 9 point correspondences, it would be better to use more of them to solve for the fundamental matrix. Investigate this by finding the # of RANSAC iterations you would need to run with 18 points.]

$$N = \left\lceil \log \frac{1-0.999}{1-0.9^{18}} \right\rceil = 43$$

We need to run RANSAC for more iterations because we are incorporating more noise into the data.

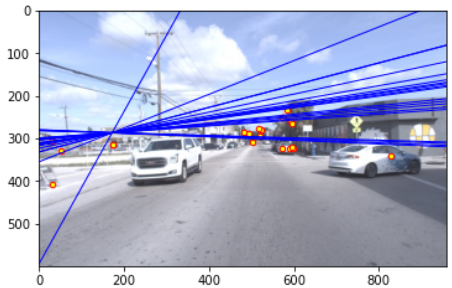
[If our dataset had a lower point correspondence accuracy, say 70%, what is the minimum # of iterations needed to find the fundamental matrix with 99.9% certainty?]

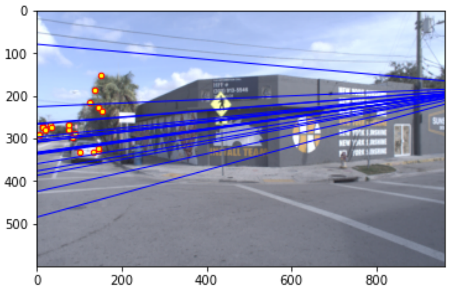
$$N = \left\lceil \log \frac{1-0.999}{1-0.7^8} \right\rceil = 117$$

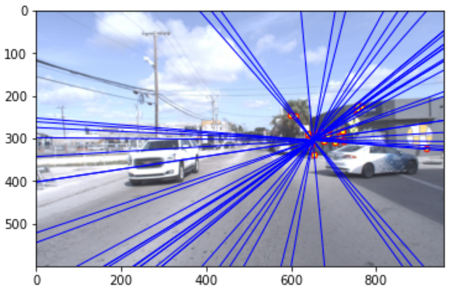
Part 4: Performance comparison

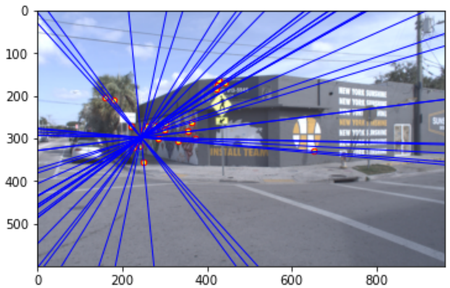
[insert visualization of epipolar lines on the Argoverse image pair using the linear method]

[insert visualization of epipolar lines on the Argoverse image pair using RANSAC]









Part 4: Performance comparison

[Describe the different performance of the two methods.]
The epipolar lines of second/first image are not converging at epipole of first/second in linear method whereas they are converging when RANSAC is used.

[Why do these differences appear?]

The presence of outliers has influenced linear method so much that the points in one image are not passing through epipolar lines of the other image. The epipoles are also not converging at the same point. However, RANSAC doesn’t have this problem as it is robust to outliers.

[Which one should be more robust in real applications? Why?]

In general RANSAC is more robust as it is not heavily influenced by outliers. This is because, the hyperparameters in RANSAC are threshold and number of iterations, both of are them are independent of outliers.

Part 5: Visual odometry

[How can we use our code from part 2 and part 3 to determine the “ego-motion” of a camera attached to a robot (i.e., motion of the robot)?]

In part 2, we estimate fundamental matrix using 8-point algorithm and in part 3 we estimate the fundamental matrix using RANSAC inside which we use part 2 after randomly sampling 8 points. We first calculate camera intrinsic parameters. After this, we go through all image frames from ego-motion and take two successive image frames, i and i+1. We then calculate essential matrix from the fundamental matrix and Camera Matrix. Then we recover pose from essential matrix and the top 30 inlier points in both images obtained from RANSAC. This will recover relative camera rotation and translation using cheirality check. The cheirality check ensures that the scene points should be in front of the camera. It means that the triangulated 3D points should have a positive depth. We then use previous world frame pose, to place this camera in world frame assuming 1 meter translation for unknown ambiguity.

[In addition to the fundamental matrix, what additional camera information is required to recover the ego-motion?]

From above algorithm, we can see that we need essential matrix to recover ego-motion. Since, essential matrix is related to fundamental matrix using, $\boldsymbol{F} = \boldsymbol{K}'^{-\boldsymbol{T}}\boldsymbol{E}\boldsymbol{K}^{-\boldsymbol{1}}$, we need the camera matrix $\boldsymbol{K}$ and $\boldsymbol{K}'$. In case of Visual Odometry, both $\boldsymbol{K}$ and $\boldsymbol{K}'$ are the same as it is the same camera.

Part 5: Visual odometry

[Attach a plot of the camera's trajectory through time]