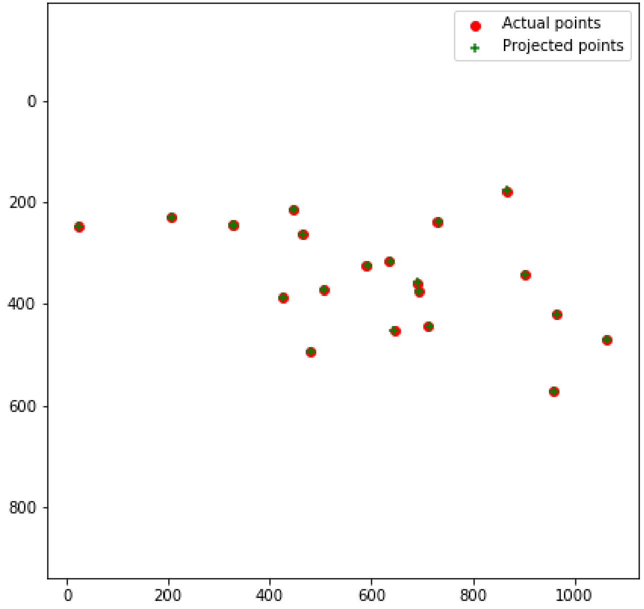
# CS 5330 Programming Assignment 3

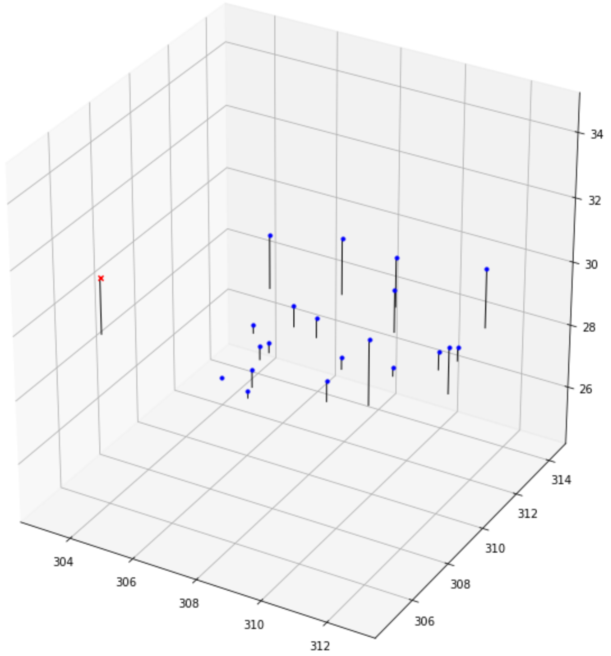
Grania Machado
machado.g@northeastern.edu
machado.g
001517781

# Part 1: Projection matrix

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

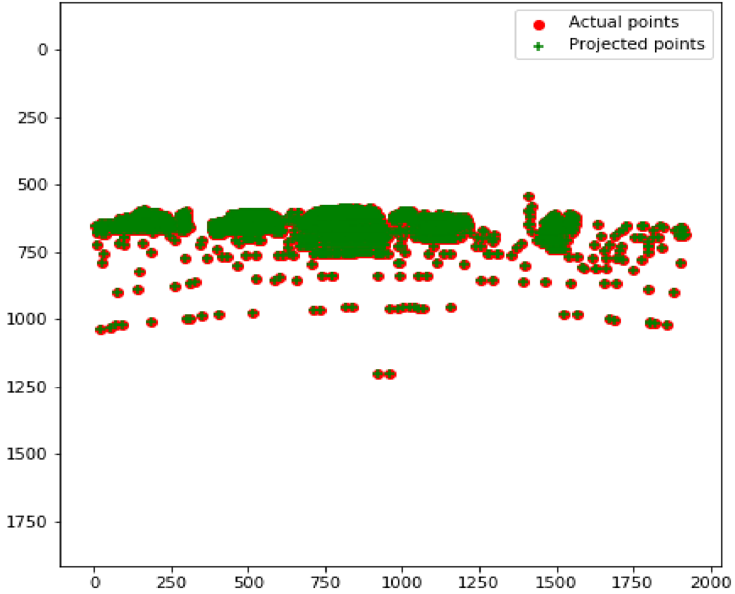



[insert visualization of camera center for the CCB image here]

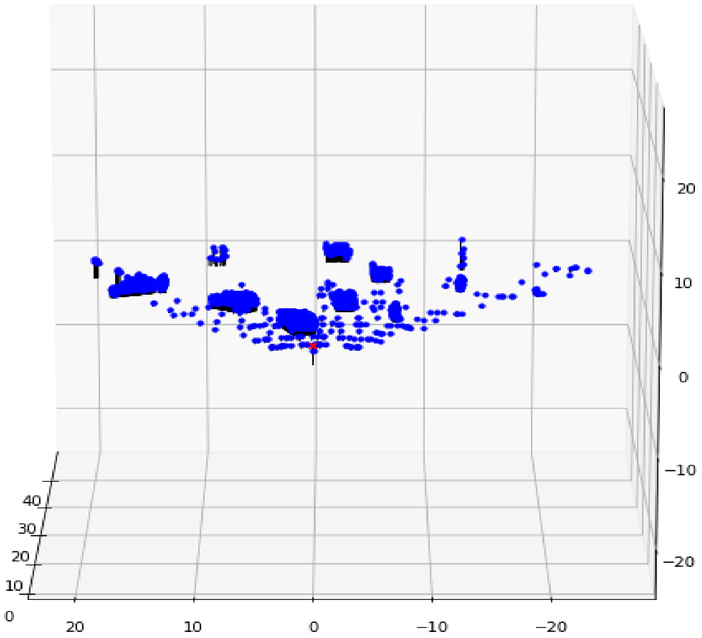


# Part 1: Projection matrix

[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?]

$$x = PX$$

Where, x = 2D image point
P = camera matrix
X = 3D world point
Homogenous 3D points to 2D image points

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

$$P = K[R|t]$$

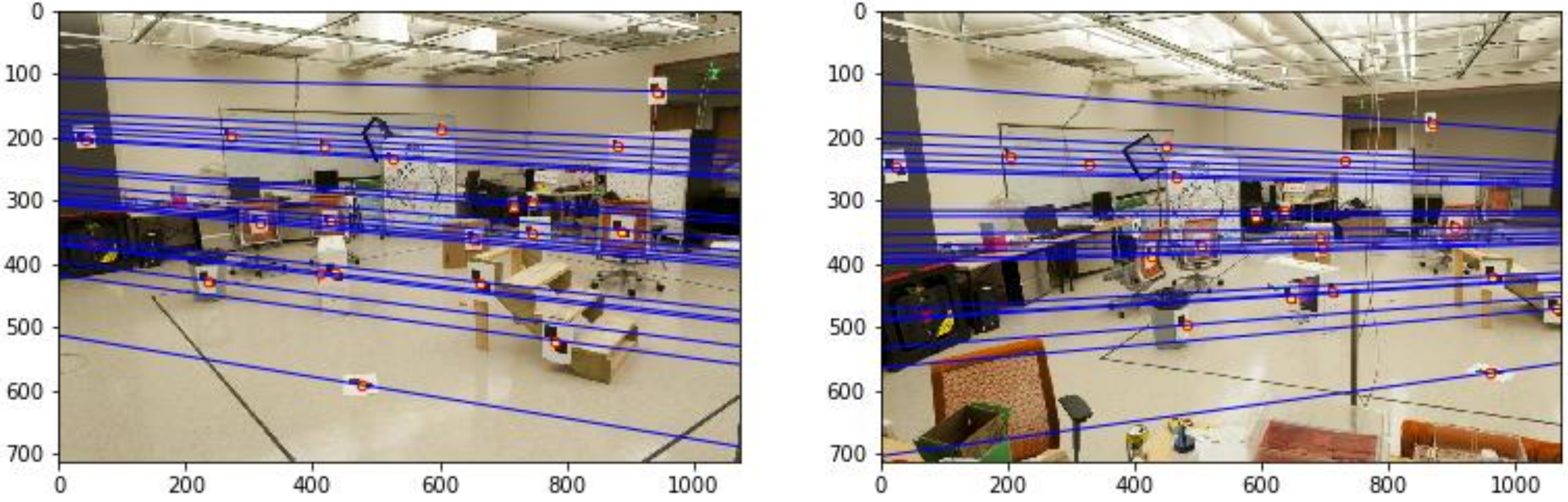
Where, K = 3x3 intrinsics
R = 3x3 3D rotation
t = 3x1 3D translation
Intrinsic and Extrinsic parameters

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

1. Camera (perspective projection)
2. Image (intrinsic camera parameters)
3. World (extrinsic camera parameters)

# 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?]

The mapping between a point and its epipolar line is given by the fundamental matrix. Consider a pair of images, here for each point $x_L$ in first image, there exists a corresponding epipolar line l in the other image. In second image, we have any point $x_R$ matching the point $x_L$, then this point $x_R$ must lie on the epipolar line l. Thus, there is a map $x_L \rightarrow l$ from a point in one image to its corresponding epipolar line in the other image which makes the projection of epipolar lines by fundamental matrix.

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

When two images have camera centers within the image then the epipoles e and e' will be located at infinity since the baseline joining the centers $O_L$, $O_R$ is parallel to the image planes, and intersects the image plane at infinity. The epipolar lines are parallel to an axis of each image plane.

# Part 2: Fundamental matrix

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

If two camera image planes coincide, then epipolar lines also coincide ($e_L$–$X_L$ = $e_R$–$X_R$) which means that the epipolar lines are parallel to the line $O_L$–$O_R$ between the centers of projection and can be aligned with the horizontal axes of the two images. Thus, it means that if you look simply along a horizontal line, you can find the matching point in the other image for each point in the first image.

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

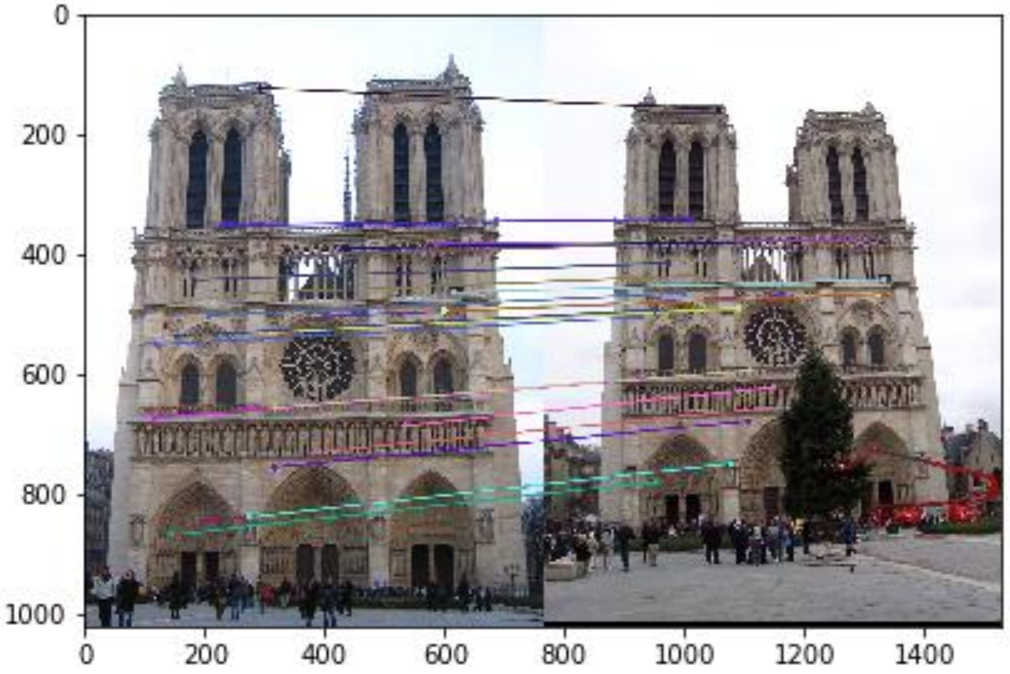
The fundamental matrix is defined by, $x_L^T *F*x_R = 0$. Once we have calculated the value for F, it solves equation for pixels $x_L$, $x_R$. Now, if we multiply F by any scalar a , we solve equation as : $x_L^T*a*F*x_R = 0$. Hence, F'=a*F is also a valid fundamental matrix. This can be simply put as there is a unique fundamental matrix, up to a scale.

[Why is the fundamental matrix rank 2?]

Fundamental matrix F represents a 3x3 mapping from the 2-dimensional projective plane of the first image to the pencil of epipolar lines through the epipole e. In this way, it speaks to a mapping from a 2-dimensional onto a 1-dimensional projective space and must subsequently have rank 2.

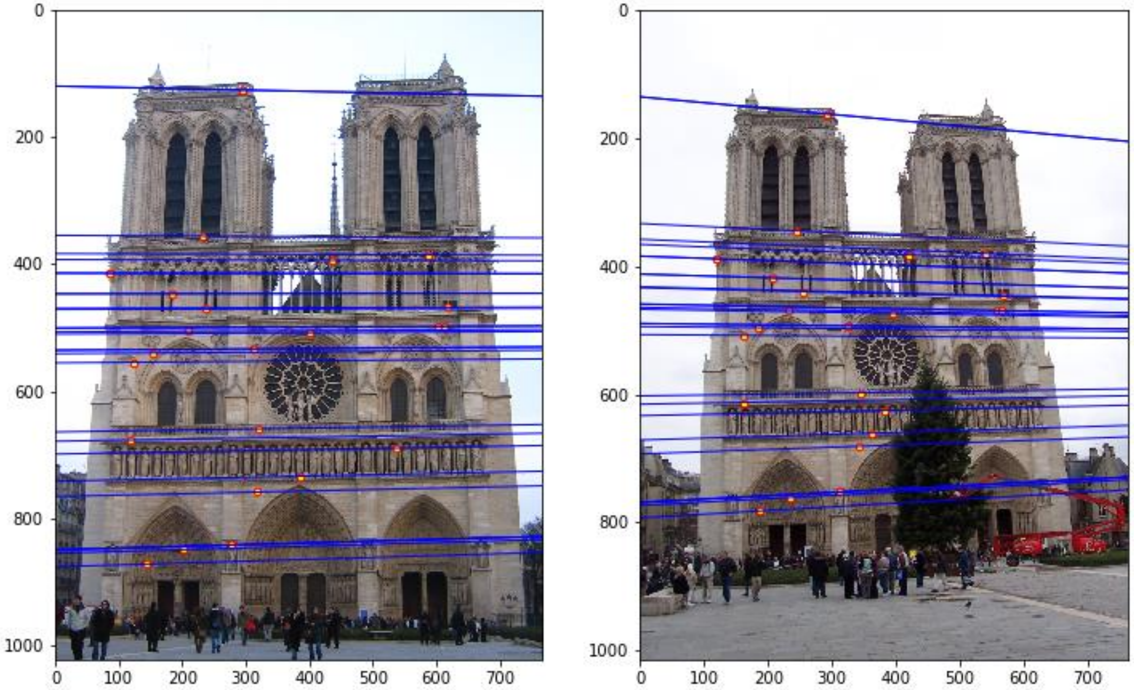
# Part 3: RANSAC

[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?]
Mt. Rushmore : 12
Notre Dame : 12

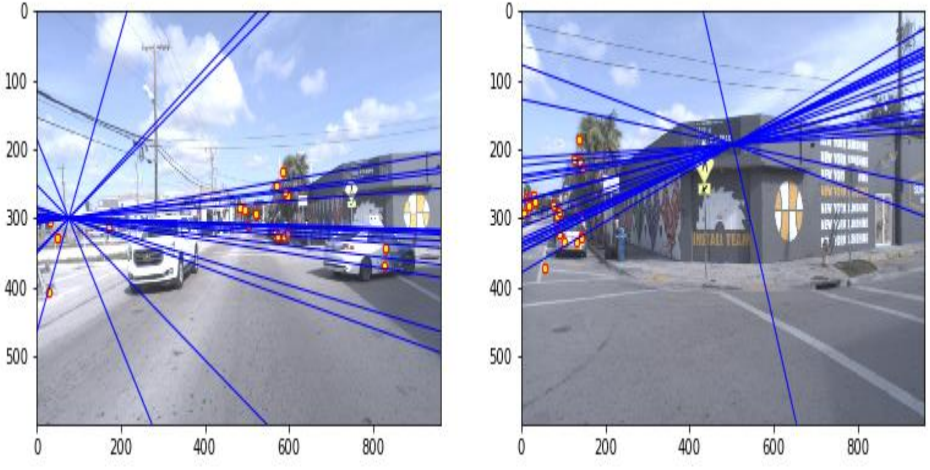
[One might imagine that if we had more than 9 point correspondences, it would be better to use more of them to solve for the fundamental matrix. Investigate this by finding the # of RANSAC iterations you would need to run with 18 points.]
# of RANSAC iterations : 42

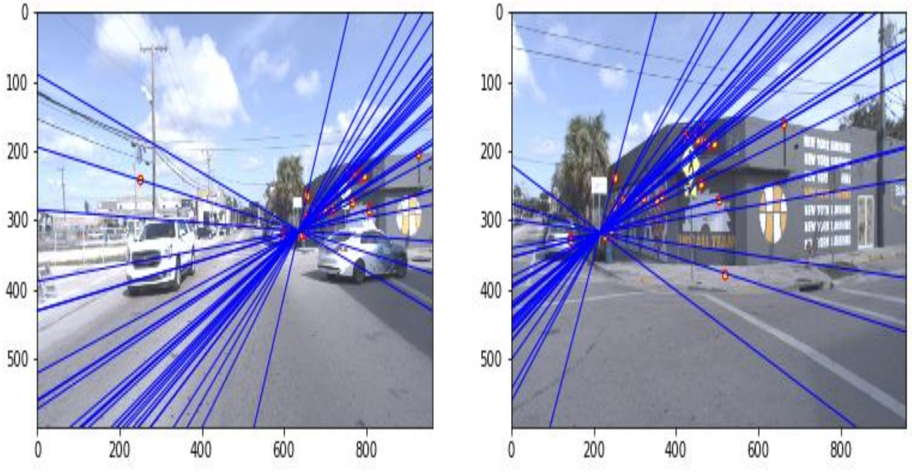
[If our dataset had a lower point correspondence accuracy, say 70%, what is the minimum # of iterations needed to find the fundamental matrix with 99.9% certainty?]
# of iterations needed : 116

# 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.]

Linear method considers all points i.e. inliers and outliers while constructing its epipolar lines and hence the intersection of the epipolar lines does not give accurate image matching results. RANSAC through repeated sub-sampling tries to draw its epipolar lines considering majorly inliers and small subset size, thus, intersection of its epipolar lines gives more accurate results for image matching.

[Why do these differences appear?]

Linear method tries to optimally fit lines of all data points, including inliers as well as outliers hence it generally will produce a bad fit. On the contrary, RANSAC tries to exclude the outliers and only use inliers to draw lines in its calculations, hence produces more accurate results.

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

Linear method should be more robust in real applications as compared to RANSAC as ransac is based on two major assumptions; sufficient inliers to agree on a good model and outliers will not repeatedly choose a single model. However, with real world data, it is hard to guarantee that these assumptions will be met as its possible to have highly skewed distributions. This causes ransac to discard large portion of data without justification in order to find a good fit for model which is harmful in real applications as we lose important data.

# 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)?]

From parts 2 & 3, we use the fundamental matrix F, instrinsic camera matrices K and K’, with the matrix multiplication formula $E = K'^T * F * K$ to obtain the Essential matrix E. We then use obtained Essential matrix E and ransac fundamental matrix for pose estimation using 3D world points from triangulation and bundle adjustment to plot the trajectory of ego-motion of the camera.

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

To recover ego-motion, we need geometric properties of camera like focal length ($f_x$, $f_y$), principal point offset ($x_0$, $y_0$), as well as axis skew to compute the intrinsic camera matrix. Additionally, we also need information about the type of camera (like single camera, stereo camera), lens distortion, film size and camera calibration.

# Part 5: Visual odometry

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