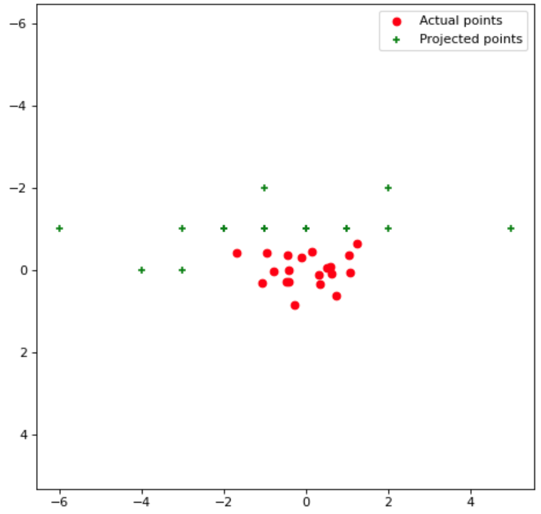
CS 5330 Programming Assignment 3

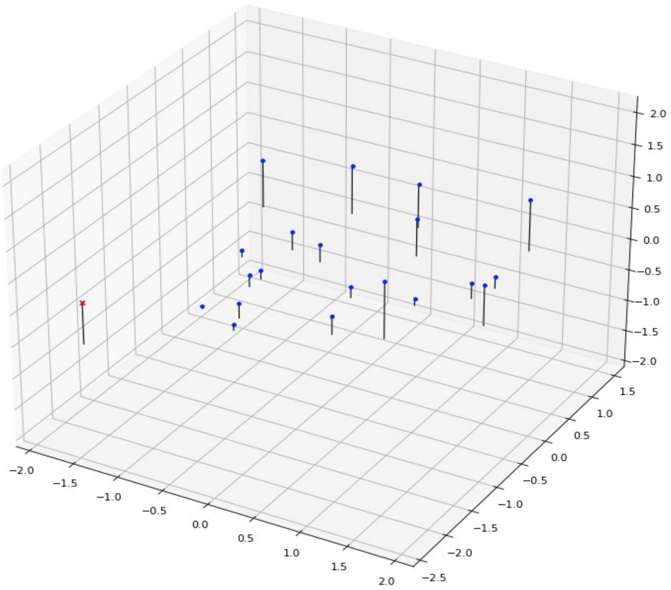
[Aditya Sri Prasad Appana]
[appana.a@northeastern.edu]
[appana.a]
[001006587]

Part 1: Projection matrix

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

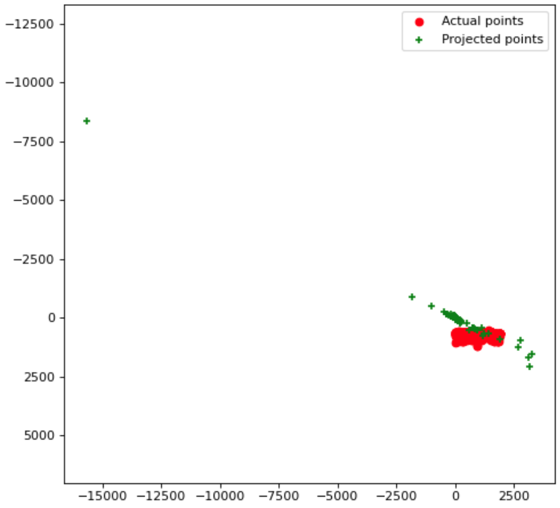


[insert visualization of camera center for the CCB image here]

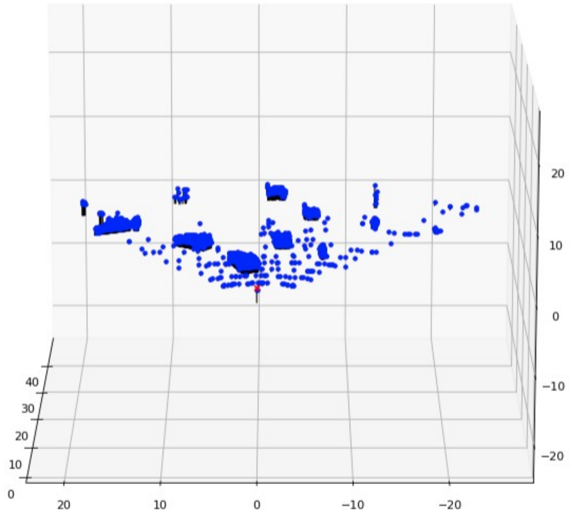


Part 1: Projection matrix

[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 projection matrix maps 3d world coordinates to 2d image coordinates.]

[What quantities can the camera matrix be decomposed into?

The projection camera matrix can be decomposed into intrinsic matrix K and an extrinsic matrix [R|t]]

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

1)Focal length of camera.

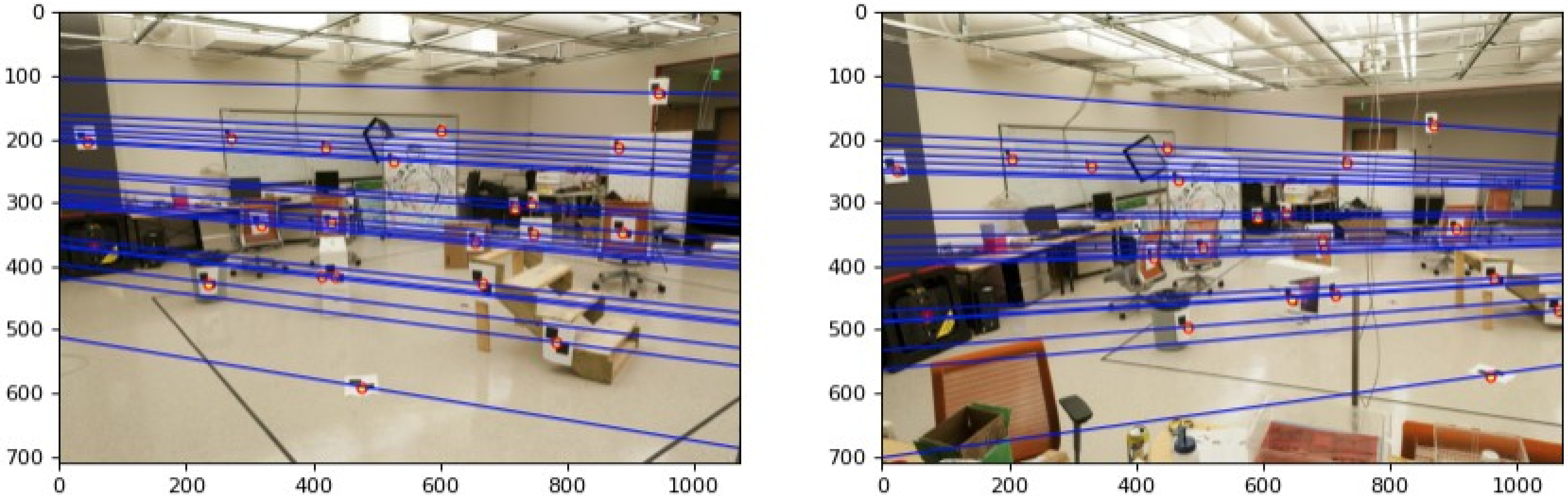
2)optical center of camera.

3)Rotation (pitch,roll yaw).

4)camera translation]

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 epipolar constraint says, for any given point in one image, multiplying it with an essential matrix gives us an epipolar line on another image.(Incase of calibrated cameras). When using uncalibrated cameras, we use Fundamental matrix along with camera intrinsics to ensure the epipolar constraint, which results in points on one image being projected to epipolar lines on other image.]

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

Epipoles and epipolar lines will become a point. There will not be any epipolar plane so we can consider that epipoles and epipolar lines converge to a point.

]

Part 2: Fundamental matrix

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

It means Image planes of cameras are parallel to each other and to the baseline]

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

A fundamental matrix for corresponding points in a stereo image pair satisfies the epipolar constraint and is defined as x'Fx =0 (where x,x' are corresponding points). Now there could be a lot of F matrices which satisfy this equation. For all such matrices, we can factor out a scalar to get another F. Hence we say Fundamental matrix is defined up to a scale.]

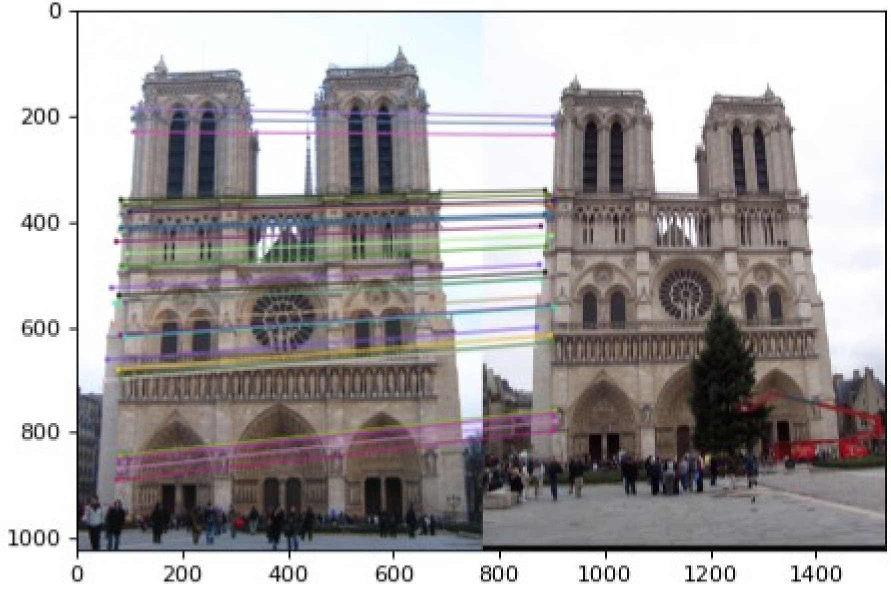
[Why is the fundamental matrix rank 2?

The fundamental matrix F represents a mapping from 2 dimensional projective plane of the first stereo image to epipolar lines through e. So it represents a mapping from 2d space to 1 d space. So rank is 2.

]

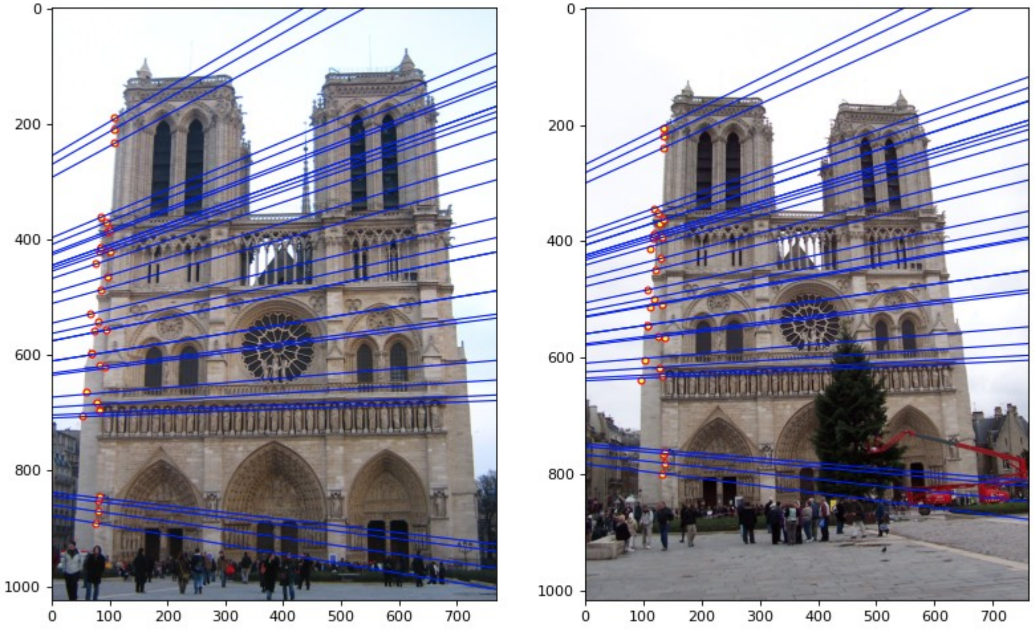
Part 3: RANSAC

[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?
With 8 point correspondences,
Notre Dame: iterations: 12
Mt. Rushmore : iterations : 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 for Notre Dame : 42

for MtRushmore :42

We can use more than 9 point correspondences to solve for Fundamental matrix, but the number of Ransac iterations tripled from using just 8 points. So it's not necessary to use more than 9 point correspondences. Infact just 8 point correspondences are sufficient. The downside of using more points is we need to run more number of iterations to get the similar probability of success.

]

[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?

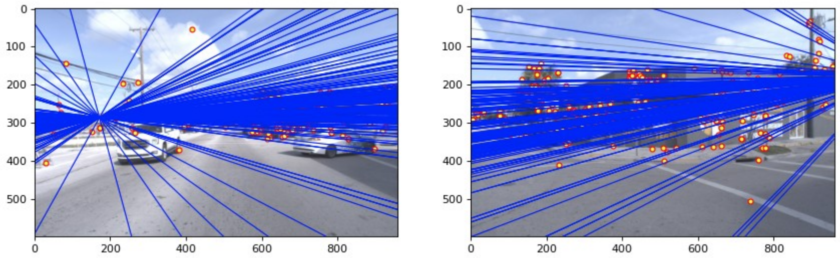
With 8 point correspondences,

Number of iterations = 116

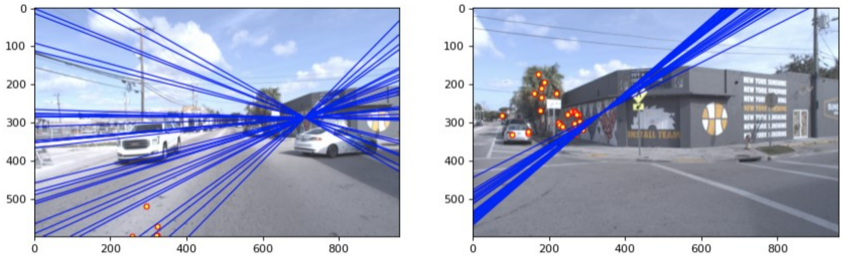
]

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.
The method without ransac, directly estimates the Fundamental matrix using the hartleys normalized 8 point algorithm. It results in an inaccurate Fundamental Matrix due to error prone matches.
The later method uses RANSAC to estimate the best Fundamental matrix. Uses inliers to best estimate a Fundamental Matrix.
The RANSAC based method seems to perform well.
]

[Why do these differences appear?

The Sift point correspondences are not always reliable, they are prone to errors. So when we are directly estimating Fundamental matrix with 8 such error prone correspondences, we get a faulty Fundamental matrix. To fix this issue, we can use RANSAC to select the inlier point correspondences(best 8 samples) from all the point pairs, and estimate Fundamental matrix.]

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

Ransac method is more robust as it is not sensitive to outliers. We fit a model to the inlier point correspondences and estimare the best Fundamental matrix for these inlier point correspondences. This gives a more robust Fundamental matrix estimate for same 8 points. Instead of directly estimating Fundamental matrix with random 8 SIFT point pairs.
]

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 the 2 consecutive images of the ego vehicles camera, we can calculate the SIFT matches, and Use RANSAC estimation for Fundamental matrix(which internally uses Part 2 normalized Fundamental matrix estimation) to get an optimal Fundamental matrix, and the best matches(inliers). We can then use these to compute Essential Matrix. As the Camera intrinsics (K matrix) for both images are same, we can easily compute Essential matrix. From this essential matrix and the best point matches, we can decompose the Essential matrix using svd to get the Rotation and translation of the camera. W can then track the rotation and translation of camera to get an overall ego motion of the camera.]

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

We need the camera intinsics(K matrix) to get an essential matrix using Fundamental matrix.

]

Part 5: Visual odometry

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