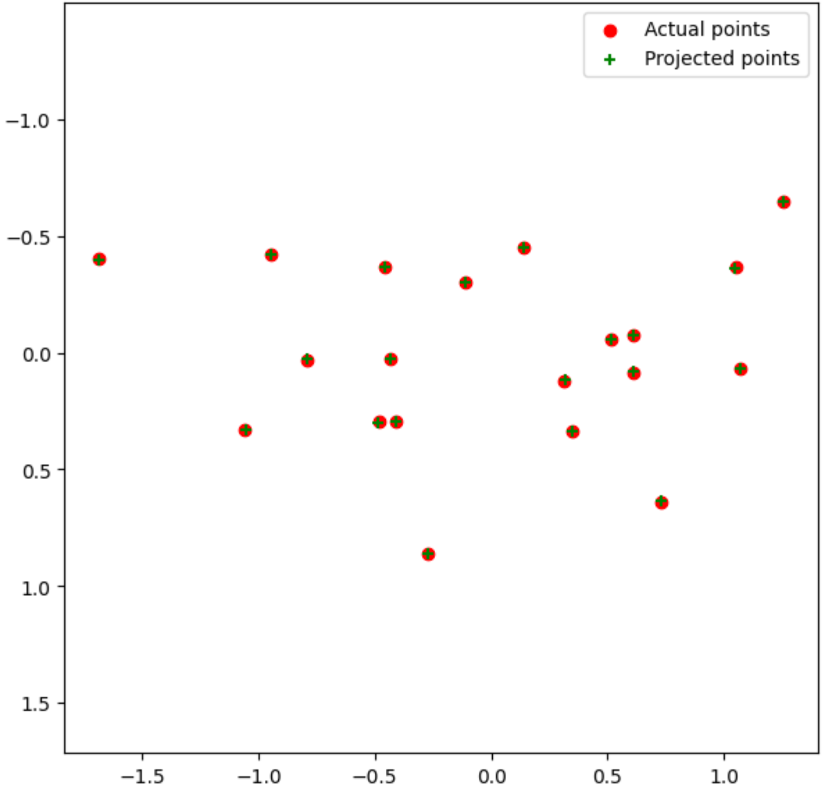
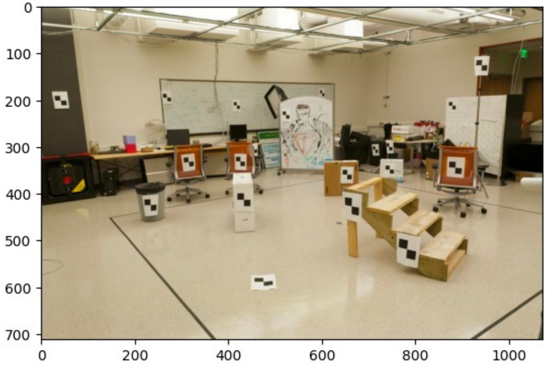
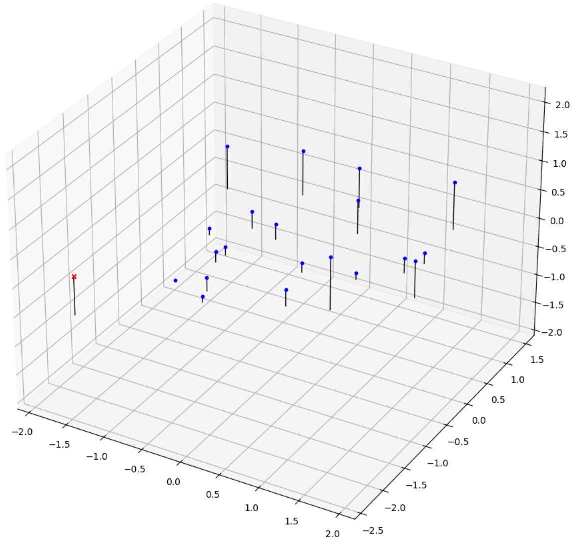
CS 4476 Project 3

Vidit Pokharna
vidit@gatech.edu
vpokharna3
903772087

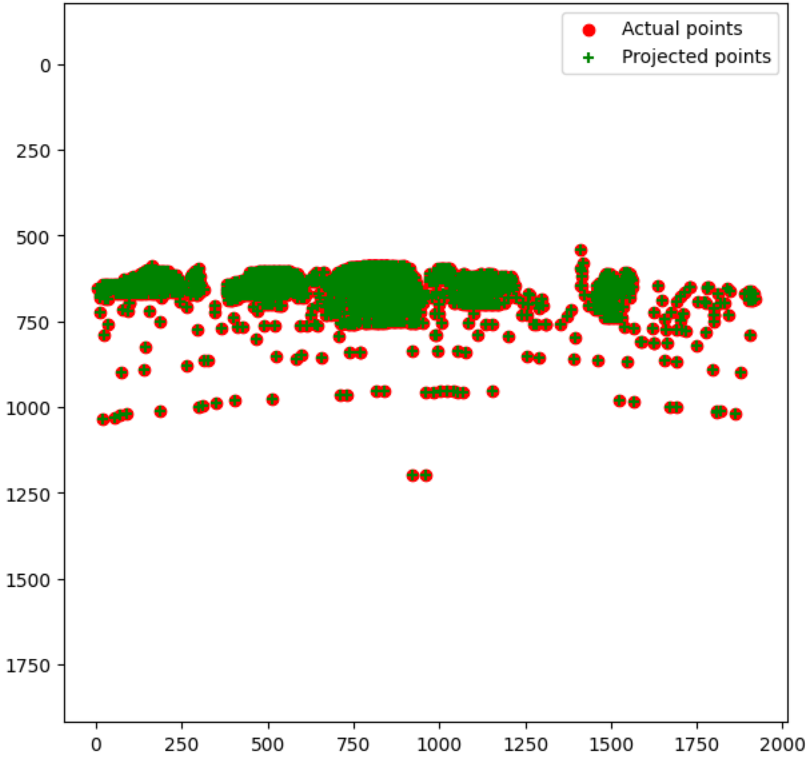
Part 1: Projection matrix

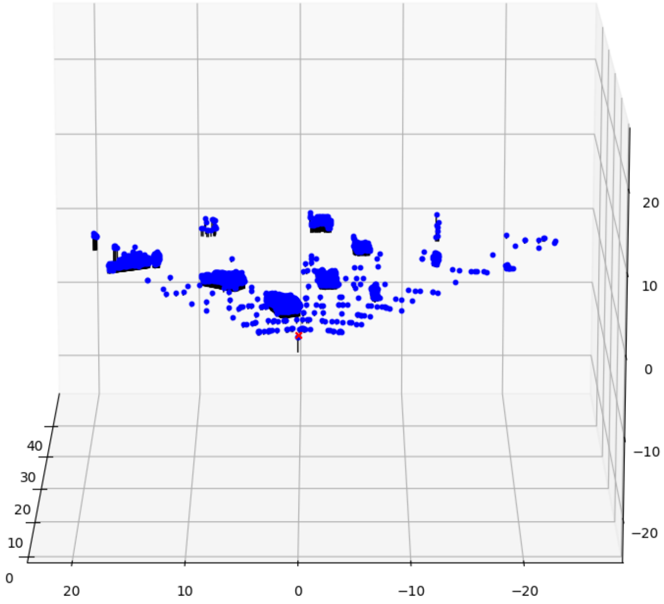






Part 1: Projection matrix





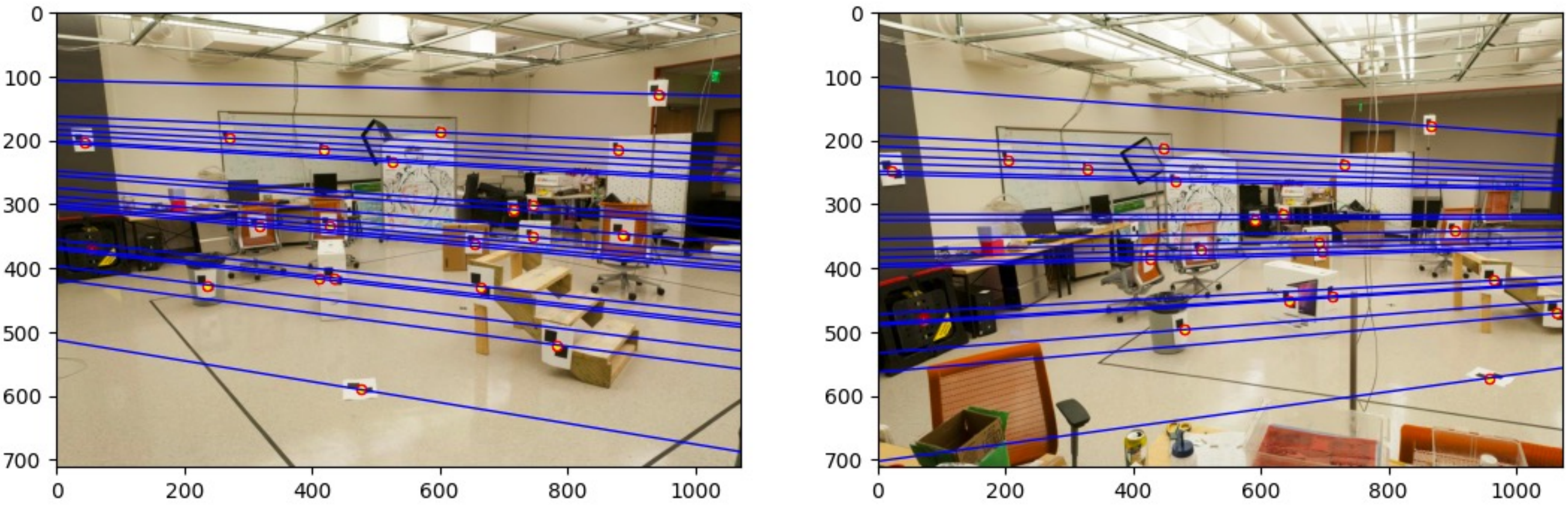
Part 1: Projection matrix

The camera matrix relates 3D world coordinates to 2D image coordinates.

The camera matrix can be decomposed into intrinsic and extrinsic parameters.

Factors that affect the camera projection matrix include lens distortion, the focal length of the camera, and the relative position and orientation of the camera to the scene being imaged.

Part 2: Fundamental matrix



Part 2: Fundamental matrix

Points in one image are projected onto epipolar lines in the other image by the fundamental matrix due to the constraint that the corresponding points in stereo images must lie along the epipolar lines. This is a result of the geometry of stereo vision where the epipolar lines are the lines of intersection between the image planes and the epipolar plane (the plane containing the baseline between the two camera centers and the observed point in space).

When the camera centers (the projection centers) are within the images, the epipoles will be located within the image boundaries. The epipolar lines will radiate out from the epipoles. This occurs because the epipoles are the projection of one camera center onto the image plane of the other camera. When the camera centers are within the images, it means that each camera is able to capture the other camera within its field of view, hence the epipoles and epipolar lines will be within the captured images.

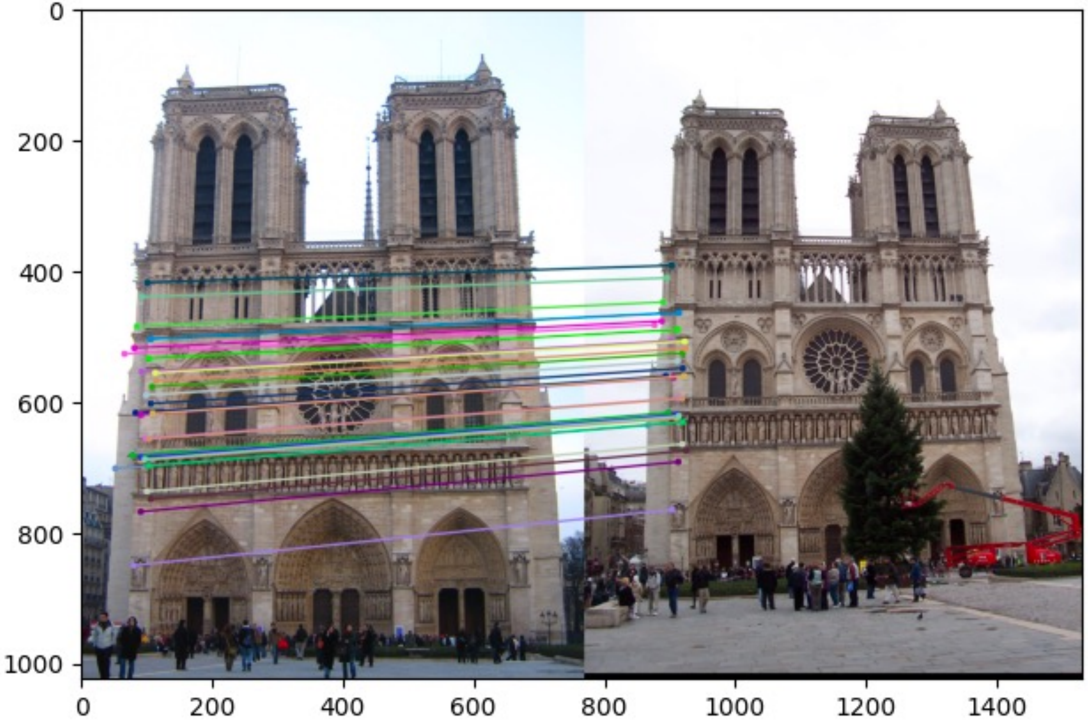
Part 2: Fundamental matrix

If the epipolar lines are all horizontal across two images, it usually means that the images have been rectified. Image rectification is the process of transforming images to ensure that the corresponding epipolar lines in both images become collinear and parallel to the horizontal axis of the image plane. This simplifies the problem of finding corresponding points in stereo vision because it reduces the search space to one dimension along the horizontal lines.

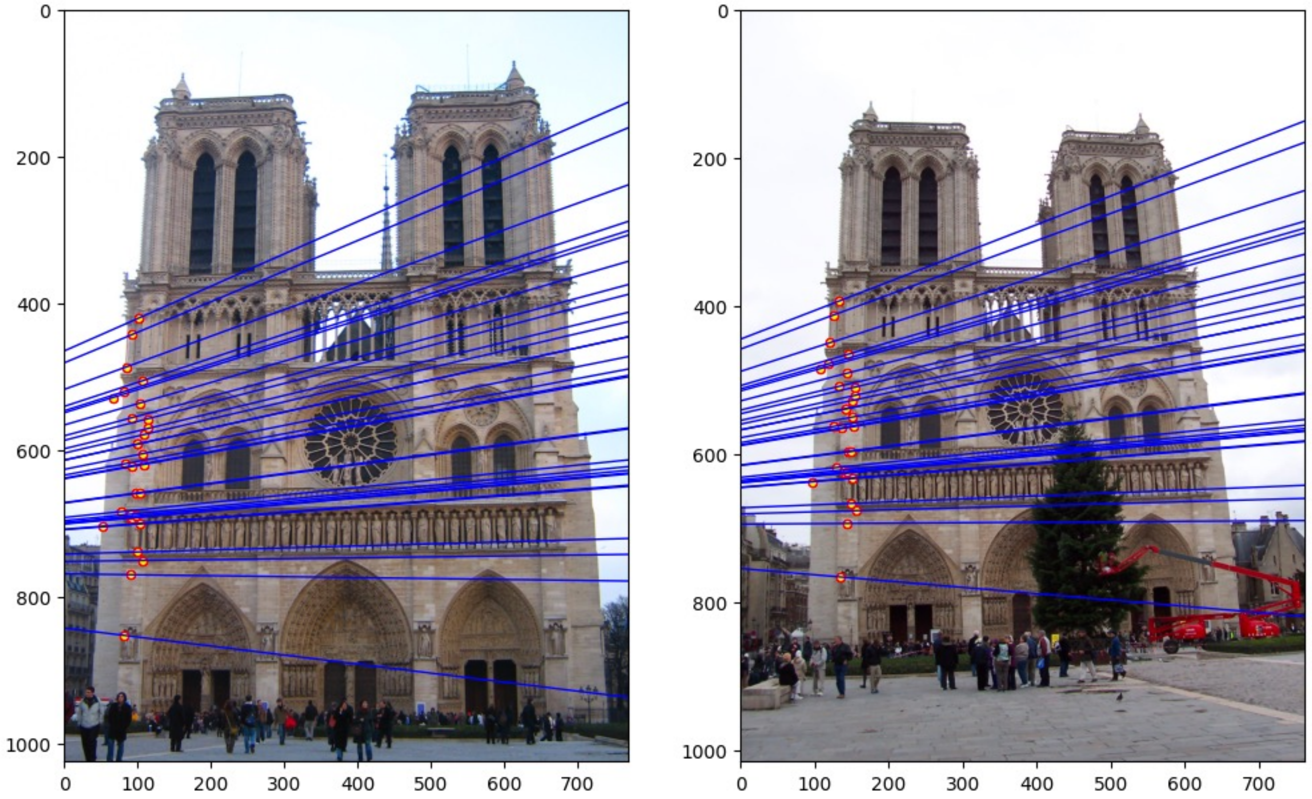
The fundamental matrix is defined up to a scale because it represents a projective transformation, and its computation involves homogeneous coordinates. In homogeneous coordinates, multiplying the coordinates by a non-zero scalar does not change the point's location in projective space. Therefore, the fundamental matrix, which relates corresponding points in two images using homogeneous coordinates, is only determined up to a scale factor.

The fundamental matrix is of rank 2 because it encapsulates the epipolar constraint that the corresponding points in stereo images must lie on their respective epipolar lines. Mathematically, the epipolar constraint is expressed as a bilinear form, which results in a matrix equation that inherently has one degree of freedom less than the dimension of the space. This is reflected in the singular value decomposition of the fundamental matrix, where one singular value is zero, reducing the rank to 2. This also corresponds to the fact that the epipoles are not in the image plane, resulting in a matrix that cannot be of full rank.

Part 3: RANSAC



Part 3: RANSAC



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 if there are 9 points?

$$\frac{log(1-0.999)}{log\left(1-(1-0.1)^{9}\right)} = 14.09525$$

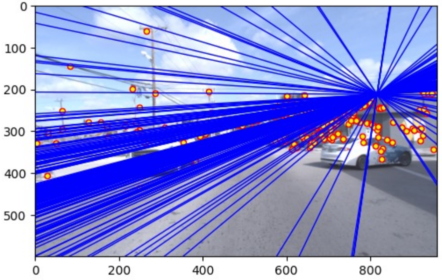
One might imagine that if we had more than 9 point correspondences, it would be better to use more of them to solve for the fundamental matrix. Investigate this by finding the # of RANSAC iterations you would need to run with 18 points.

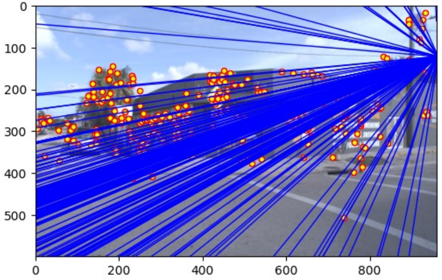
$$\frac{log(1-0.999)}{log\left(1-(1-0.3)^{9}\right)} = 167.70302$$

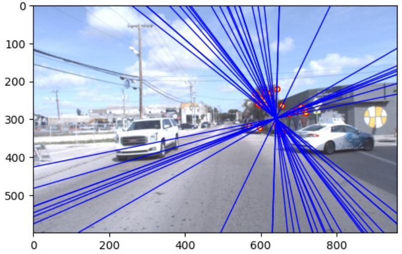
If our dataset had a lower point correspondence accuracy, say 70%, what is the minimum # of iterations needed to find the fundamental matrix with 99.9% certainty?

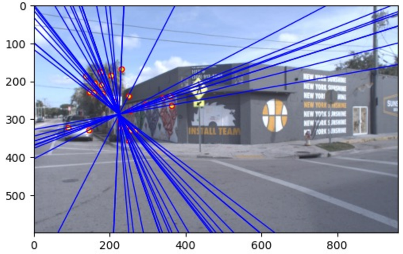
$$\frac{log(1-0.999)}{log\left(1-(1-0.1)^{18}\right)} = 42.47521$$

Part 4: Performance comparison









Part 4: Performance comparison

The linear method calculates the fundamental matrix by solving a set of linear equations derived from point correspondences. It assumes that all correspondences are correct and can be sensitive to noise and outliers. The RANSAC method, on the other hand, is an iterative method that randomly selects a subset of point correspondences to estimate the fundamental matrix and identifies inliers that fit the model well. This method is more robust to outliers but can be computationally more intensive.

The differences in performance arise because the linear method does not account for the presence of outliers which can significantly affect the estimation of the fundamental matrix. RANSAC continuously refines the model by focusing on inliers and discarding outliers, which often leads to a more accurate estimation in real-world scenarios where outliers are common.

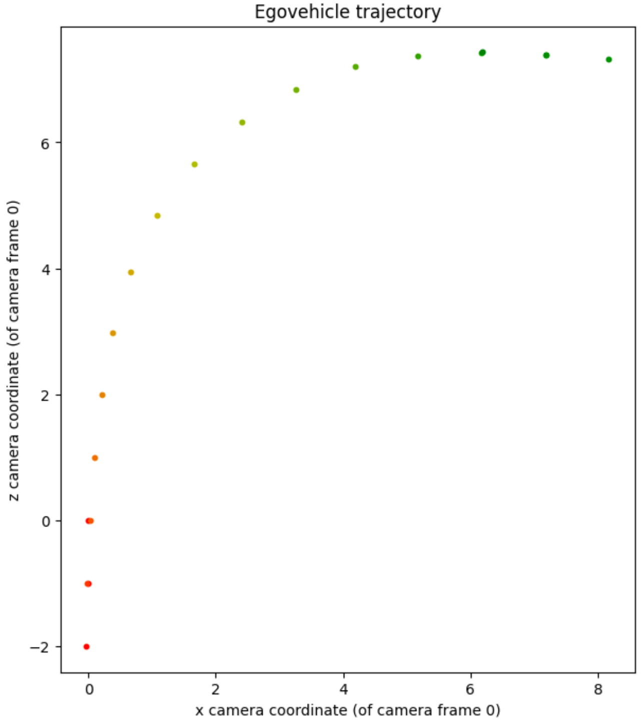
RANSAC should generally be more robust in real applications because real-world data often contains outliers due to incorrect matches, sensor noise, or other errors. RANSAC's iterative process of selecting random subsets of the data for model estimation allows it to handle such outliers effectively

Part 5: Visual odometry

To determine the ego-motion, which is the 3D motion of the robot between two frames, we can use the fundamental matrix to find correspondences between points in consecutive frames and then calculate the essential matrix if you have the camera's intrinsic parameters. From the essential matrix, we can extract the rotation and translation components that represent the motion of the camera (and hence the robot) between the two frames.

To recover the ego-motion from the fundamental matrix, we need camera parameters, which include the focal length and the optical center of the camera. These parameters allow us to compute the essential matrix from the fundamental matrix. The intrinsic parameters are necessary because the fundamental matrix alone only encodes the epipolar geometry without considering the camera's internal configuration, which is essential for recovering the actual motion in 3D space.

Part 5: Visual odometry



Part 6: Panorama Stitching

Introduction
This script stitches two images together to create a single panoramic image.

Dependencies
- Python 3.8 or higher
- OpenCV 4.5.3 or higher
- NumPy 1.19.5 or higher

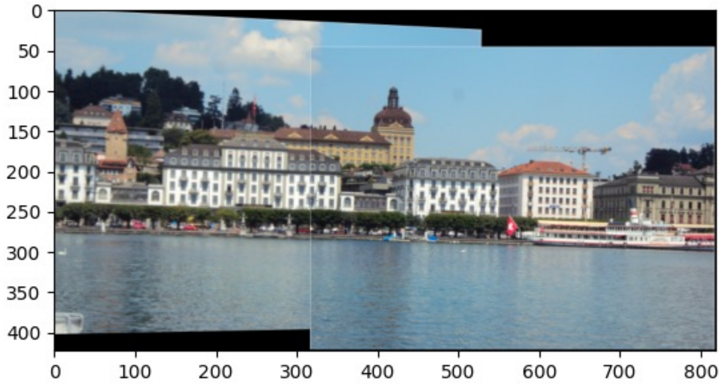
Setup
```bash
pip install opencv-python==4.5.3 numpy==1.19.5

Usage
To stitch two images into a panorama:
1. Save your images in a known directory.
2. Ensure the script panorama_stitch.py is in the same directory as your images.
3. Run the script using the following command, replacing imageA.jpg and imageB.jpg with your image filenames:

python panorama_stitch.py imageA.jpg imageB.jpg

The script will output a single image file panorama.jpg, which is the stitched panoramic image of the two input images.
```

Part 6: Panorama Stitching



Conclusion

Accurate estimation of the fundamental matrix via RANSAC is crucial in computer vision for applications such as 3D reconstruction, which is essential in virtual reality and augmented reality to align virtual objects with the real world. It aids in robotics and autonomous vehicles for navigation and obstacle avoidance by understanding the spatial layout. It helps to correct and stitch images for accurate map creation, and in surveillance, it assists in object tracking. Video stabilization is another area that benefits from accurate motion estimation. Furthermore, these applications become more accessible as they can be applied with low-cost and less sophisticated cameras, broadening the impact of computer vision technologies.

Extra Credits

The warrior image was not given to us, I could not proceed further with this. The TA’s were also unresponsive on Ed Discussion.