COL780: Computer Vision

Assignment-2 Report

Bogam Sai Prabhath
2023AIB2079
Indian Institute of Technology Delhi
March 7, 2024

1.Problem Statement:

The assignment is about stitching the images of a scean captured by a camera.the stitching of images have be done from scratch with out any libraries.the output image is generated at high resolution by combining them into a single panorama image also other part is about capturing the frames of the input video at particular intervals of time and generating the panorama image using the captured frames.

TASK 1:

1.Image preprocessing:

As the images are at high resolution in RGB for the finding of the keypoints and further processing i have generated the gray scale image which contain the original height and width of only one layer.there is no compressinon of the original image for finding keypoints.

2.SIFT:

1. **Scale-Space Extrema Identification:**
    1.1. **Normalization:** The grayscale values of the image are adjusted to fall within a 0 to 1 range.
    1.2. **Upsampling:** To increase the resolution, the original image is upscaled to twice its size, introducing more sample points than the original image provided.
    1.3. **Octaves Formation:** Five sets (octaves) are constructed, each containing three scale levels (thus totaling six images per octave due to adding three extra images for the scales). These images undergo Gaussian blurring at varying scales, starting with a standard deviation of 1.6 and increasing incrementally according to the formula kn⋅σkn⋅σ, where k=2−sk=2−s. The Gaussian-blurred image is represented as L(x,y,σ)=G(x,y,σ)⋅I(x,y)L(x,y,σ)=G(x,y,σ)⋅I(x,y).
    1.4. **Difference of Gaussians (DoG):** Four DoG images are generated by subtracting Gaussian-blurred images at adjacent scales for each octave. The DoG is calculated as D(x,y,σ)=L(x,y,k⋅σ)−L(x,y,σ)D(x,y,σ)=L(x,y,k⋅σ)−L(x,y,σ).
    1.5. **Keypoint Selection:** Pixels that stand out as local extrema (either maxima or minima) within a 3×3×3 neighborhood are marked as potential key points.
2. **Refinement of key points:**
    2.1. **Low-Contrast Extrema Removal:** Local extrema with a DoG value below 0.0025 are considered unstable due to low contrast and are thus discarded.
    2.2. **Edge Response Elimination:** Keypoints positioned along edges are filtered out by examining the principal curvatures derived from the Hessian matrix HH, applying a threshold defined by Tr(H)2Det(H)<(r+1)2rDet(H)Tr(H)2<r(r+1)2, where r=10.
3. **Orientation Attribution:**
    3.1. **Histogram Formation:** A histogram of orientations is constructed from the gradients of sample points around each keypoint. This histogram comprises 36 bins, representing a full 360-degree range.
    3.2. **Weighted Contributions:** Contributions to the histogram are weighted by both the magnitude of the gradient and a Gaussian

window, with the window's σσ set at 1.5 times the scale of the keypoint.

3.3. **Peak Identification:** The highest peak in the histogram signifies the keypoint's primary orientation. Other significant peaks, reaching at least 80% of the highest peak's value, create additional keypoints with corresponding orientations. Peaks are fine-tuned by fitting a parabola to the three closest histogram values.

4. **Descriptor Construction:**

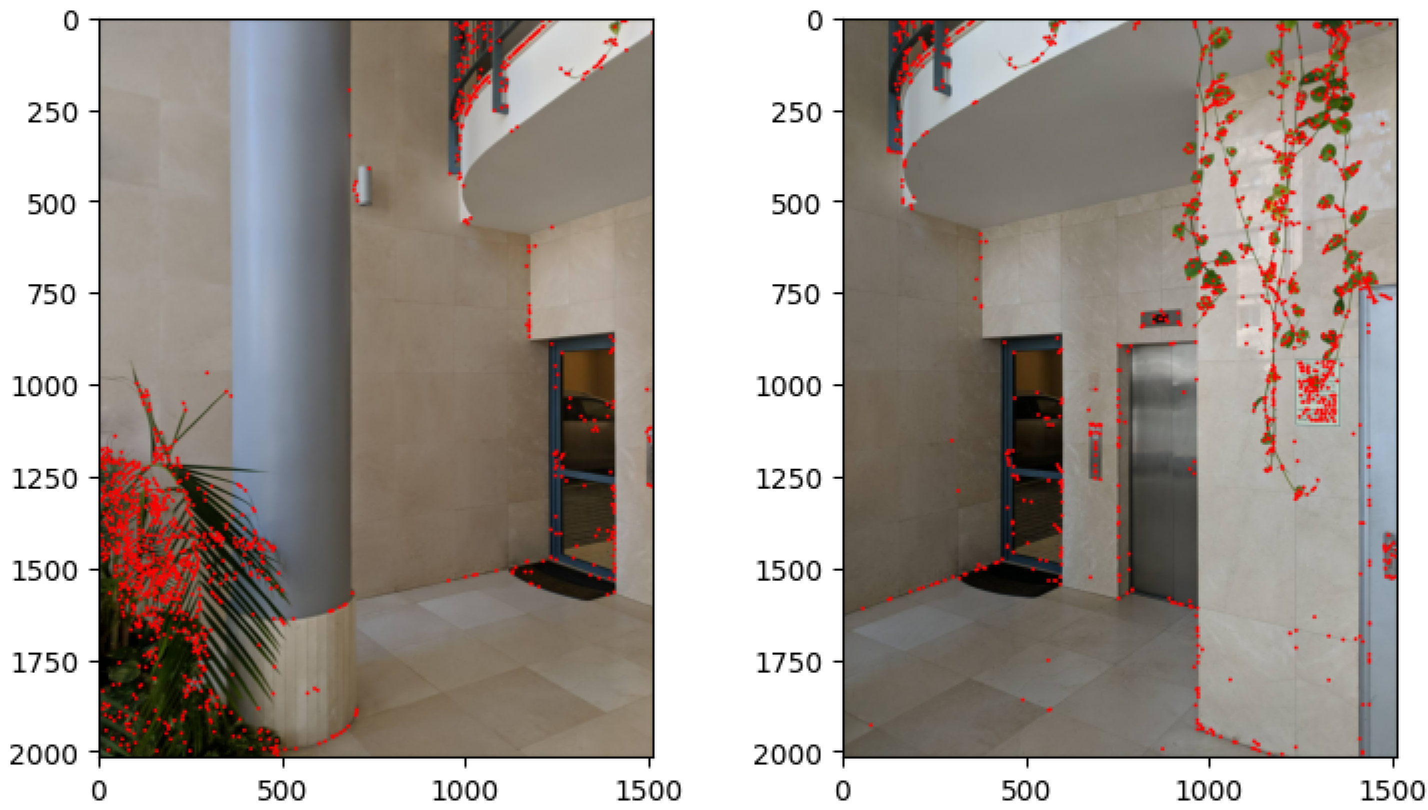
4.1. **Sampling Gradients:** The magnitudes and orientations of image gradients around each keypoint are sampled, with the degree of Gaussian blur determined by the keypoint's scale.

4.2. **Orientation Normalization:** To ensure orientation invariance, gradient orientations are adjusted relative to the keypoint's main orientation.

4.3. **Weighting Function:** A Gaussian function, where σσ is half the descriptor window's width, is used to weight each sample point's magnitude.

4.4. **Histogram Aggregation:** Orientation histograms are compiled for 4x4 sample regions, each with eight orientation bins, forming a comprehensive descriptor.

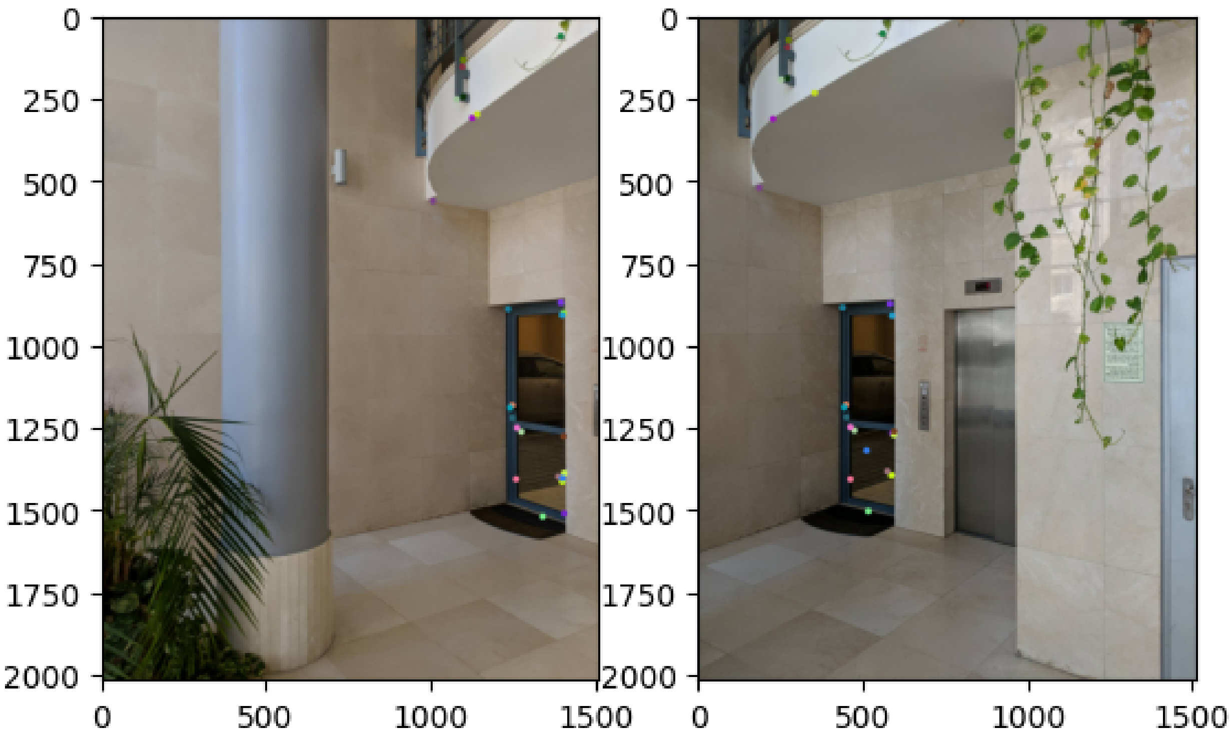
4.5. **Descriptor Vector:** The descriptor consists of a 128-element feature vector, aggregating all entries from the orientation histograms associated with each keypoint.

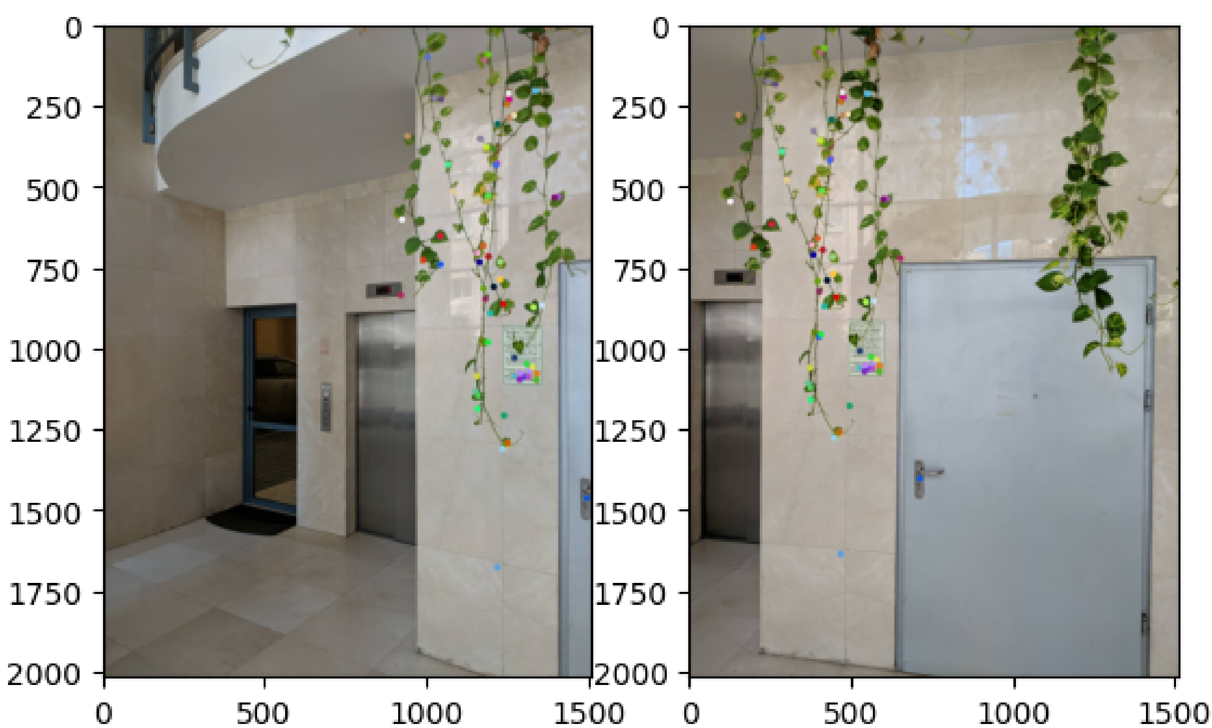



0 250 500 750 1000 1250 1500 1750 2000
0 500 1000 1500
0 250 500 750 1000 1250 1500 1750 2000
0 500 1000 1500
0 250 500 750 1000 1250 1500 1750 2000
0 500 1000 1500
0 250 500 750 1000 1250 1500 1750 2000
0 500 1000 1500

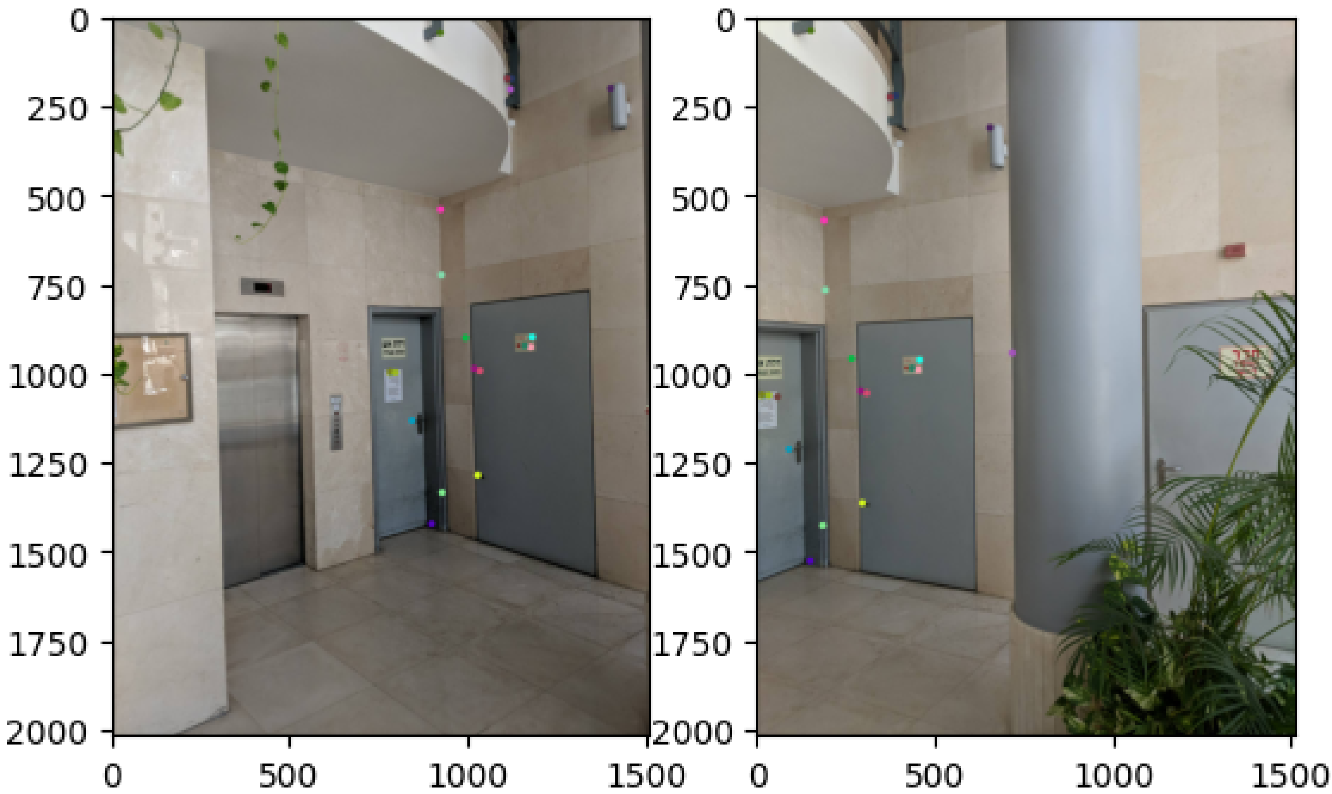
3.matching key points:

After getting all the keypoints and descriptors find make the adjacent image pairs and find the minimum euclidian distance between all posible pairs and consider those points whose second best distance is less than the threshold=0.6 as best match.





0 250 500 750 1000 1250 1500 1750 2000
0 500 1000 1500
0 250 500 750 1000 1250 1500 1750 2000
0 500 1000 1500
0 250 500 750 1000 1250 1500 1750 2000
0 500 1000 1500
0 250 500 750 1000 1250 1500 1750 2000
0 500 1000 1500



4. Image Registration and Stitching

Creating a panoramic image from multiple photographs involves a three-step process: Left Stitch, Right Stitch, and Final Stitch.

1.Left Stitch:

This initial phase focuses on sequentially stitching images from the left end of the set towards the center. For any given pair of adjacent images, the process unfolds as follows:

1.1 Homography Matrix Calculation: Utilizing the RANSAC algorithm, the homography matrix for the left image relative to the coordinate system of the right image is determined. This matrix is pivotal for mapping the geometric transformation between the two images.

1.2 Back-Warping**:** The left image is geometrically adjusted (back-warped) to align with the coordinate system of the right image, facilitating seamless integration.

1.3 Alignment and Feathering: Following alignment, a feathering technique is applied at the junction of the two images to blend them smoothly, eradicating any visible seams.

2.Image Stitching Process:

Post keypoint matching, the ensuing steps aim to unify these images into a panoramic view:

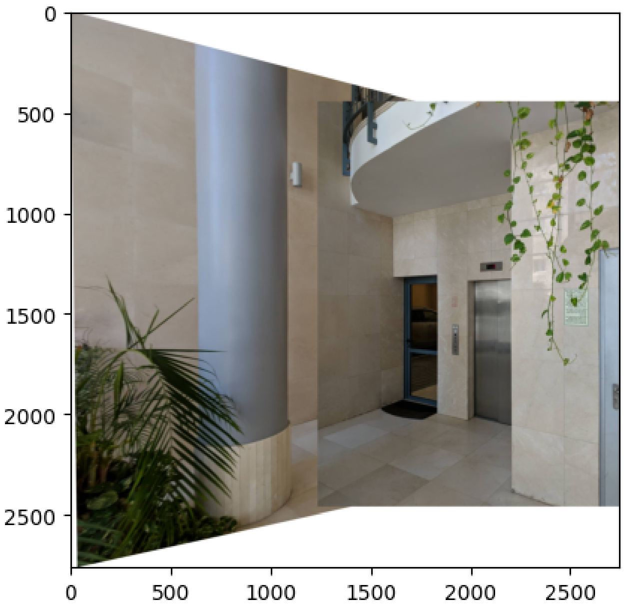
2.1 Homography Determination: This step involves calculating the homography matrix essential for point transformation from one image to another. This matrix aids in aligning corner points from one image with their corresponding locations in the adjacent image.

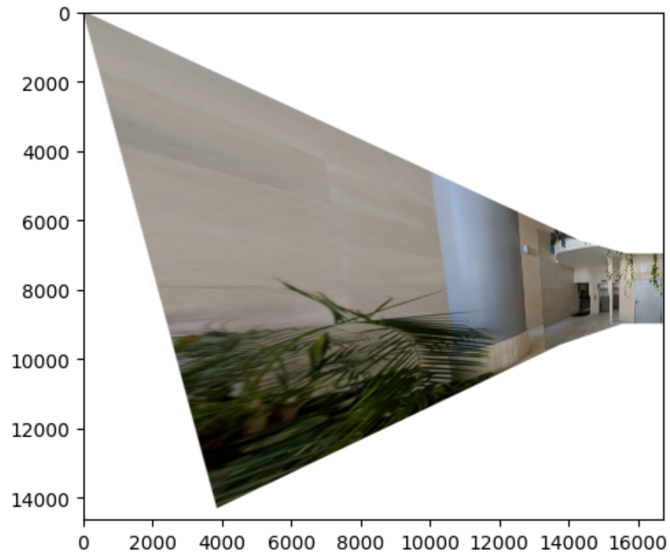
2.3 Warp Perspective and Integration: The next phase is the perspective warp of the transformed image onto a new canvas, which will also accommodate the additional image. Due to potential discontinuities arising from a direct overlay, a reverse mapping technique is employed. This technique utilizes the inverse of the homography matrix to precisely position pixels from the source image onto the destination canvas, ensuring continuity and filling any gaps.

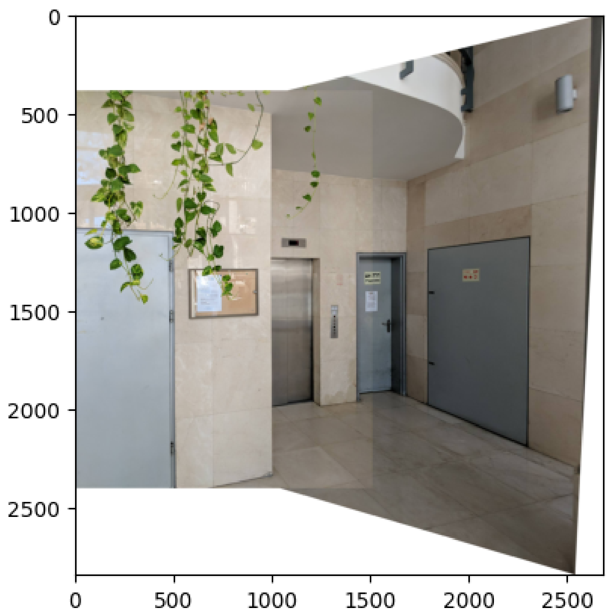
2.4 Sequential Stitching: The newly formed composite image now acts as the 'left' image for the subsequent stitching operation with the next 'right' image in the series. This iterative process continues until all images on the left half are cohesively stitched together.

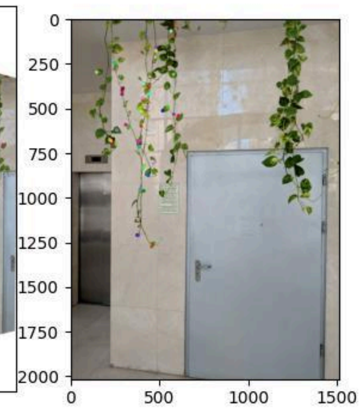
This structured approach to image stitching combines multiple photographs, accounting for geometric disparities and ensuring a smooth gradient at image junctions, creating a seamless panoramic image.

The stitched images are as below:

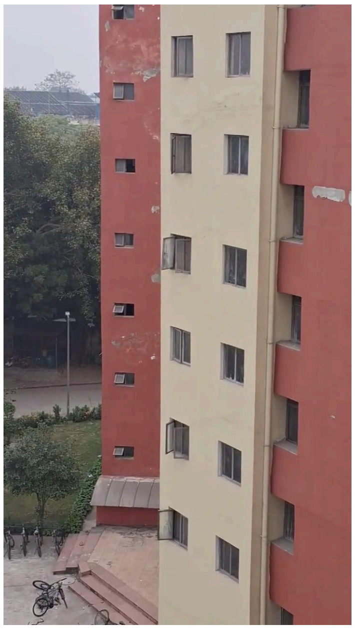


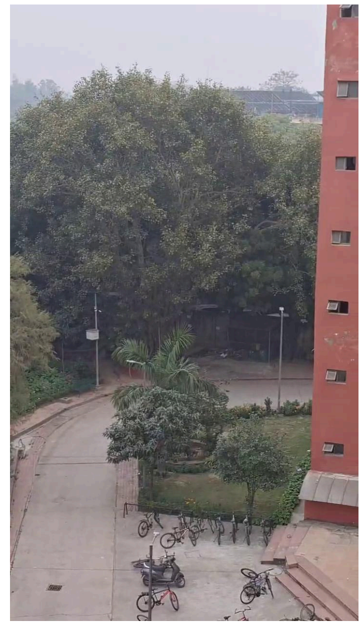


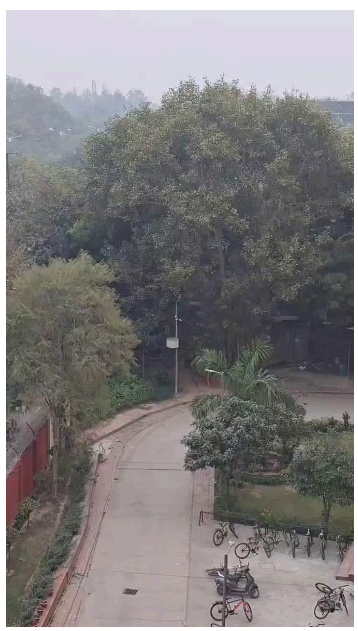




Task2 :
For the 2nd task the images are extracted at a particular interval of time i have restricted to get 5 images frames from the video the images samples are as below:







After the extraction of frames the proces is same for as Task 1.

The stitched images for given videos are as below:

Libraries used:

1.numpy: used for matrix calculations and equation solving
2.cv2: used for reading, saving, gray scale conversion.
3.os: os used to read files in path
4.random: used to generate 4 random numbers in homography.
5.sys: used for reading arguments.

References:

1.https://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf