COMPUTER VISION
Assignment 2B Report
M22MA003

Question 1 : What is a Homography Matrix? Write down steps to compute the Homography Matrix in detail with clear illustrative figures.

First, usefulness of Homography Matrix:
The planar homography is the relation between the two planes via the homography matrix.
A homography matrix is a transformation matrix that describes the relationship between two perspective views of a planar object in 3D space. It is used to project points in one image to their corresponding positions in another image to accommodate different outlooks and perspectives. Homography arrays are widely used in computer vision, especially in areas such as image stitching, object recognition, and augmented reality.

The steps for calculating the homography matrix are as follows:-

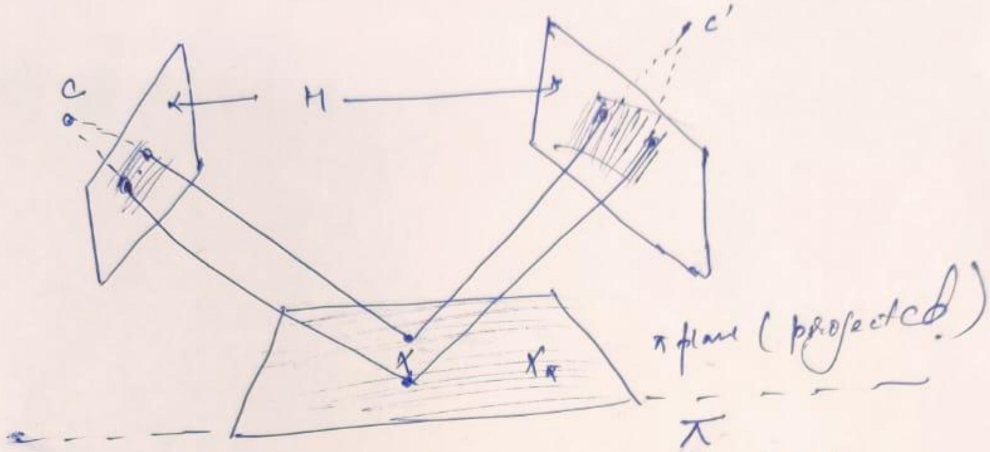
1.Select corresponding points in both images.
2.Represent the points as homogeneous coordinates.
3.Construct a system of linear equations using the homogeneous coordinates.
The equation can be written:

[x1' y1' 1 0 0 0 -x1x1' -y1x1'] [h11] =[0]
[0 0 0 x1' y1' 1 -x1y1' -y1y1'] [h12] = [0]
[x2' y2' 1 0 0 0 -x2x2' -y2x2'] [h13] =[0]
[0 0 0 x2' y2' 1 -x2y2' -y2y2'] [h21] =[0]
[x3' y3' 1 0 0 0 -x3x3' -y3x3'] [h22] =[0]
[0 0 0 x3' y3' 1 -x3y3' -y3y3'] [h23] =[0]
[x4 'y4' 1 0 0 0 -x4x4' -y4x4'] [h31] =[0]
[0 0 0 x4' y4' 1 -x4y4' -y4y4'] [h32] =[1]
where x1', y1', x2', y2', x3', y3', x4', y4' are the coordinates of the corresponding points in the second image, x1, y1, x2, y2, x3, y3, x4, y4 are the coordinates in the first point corresponding to the image.

Solving Systems of Equations: Next we have to solve the homogeneous systems of equations using **Least Square** method and **SVD**. This will result in a vector h containing the nine elements of the homography matrix H.

Normalization of the homography matrix: Normalize the homography matrix by dividing all the elements of the homography matrix by h33 to ensure that the last element of the matrix is equal to 1. This step is necessary to avoid scale and numerical errors.
Reshape the normalized homography matrix into a 3x3 matrix, and the matrix can be used.

The Mathematical illustrations are shown below:



Here we have assumed x as a ~~vector~~ matrix $= \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$

and $x' = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$ in homogeneous coordinate.

we can write both as,

$$* \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

* s.t., h $\downarrow$ Homography Matrix

$$Ah = 0 \quad \text{s.t } \|h\|^2 = 1$$

Now, $c * \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} & h & \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$

we need to remove c and again by formulating we get, $Ah = 0$

where, A can be written as —

$$A = \begin{bmatrix} -x & -y & -1 & 0 & 0 & 0 & ux & uy & u \\ 0 & 0 & 0 & -x & -y & -1 & vx & vy & v \end{bmatrix}$$

and $h = [h_1\ h_2\ h_3\ h_4\ h_5\ h_6\ h_7\ h_8\ h_9]^T$.

Here, there are more than 4 points.

hence, we need to use least square soln. to find h.

i.e, $\min_h \|Ah\|_2$; rank(A) = 3

Hence $H = (T')^{-1} H' T$

So we need to minimise

$\min_h \|Ah\|_2$ s.t $\|h\|_2 = 1$

i.e, $\min_h \|Ah\|_2$ s.t $(1 - h^T h) = 0$

$A \in R^{m \times 9}$ $m \geq 8$

$h \in R^{9 \times 1}$

Using Lagrange, we write the above constrained optimis as:-

$\min_h \|Ah\|_2^2 + \lambda(1 - h^T h)$

Taking partial derivative with respect to h & equating to zero,

we get,
$$(A^T A - \lambda I) h = 0$$
$$\det(A^T A - \lambda I) = 0$$

λ is one of the eigen values the $A^T A$ with eigen vector h.

on using SVD $A = U \Sigma V^T$, h can be any of the columns of V.

we will take that h corresponding to min eigenvalue $A^T A$. i.e, last eigenvector/col. of V.

Hence, let h be the vector obtain in the above step.

Next we need to normalise the vector h

last:- After it perform Reshaping of vector to turn h to form H (homography matrix) 3x3.

Question 2 : What is stereo matching? Write down three applications of stereo matching. You are allowed to use web/books, but write the answer in your own words.

- ➔ Stereo matching is a computer vision technique used to find the corresponding points between two or more images taken from different viewpoints. The goal of stereo matching is to reconstruct the 3D structure of the scene captured by the images.
- ➔ Stereo matching algorithms typically consist of three main steps: feature extraction, correspondence matching, and depth estimation. In the feature extraction step, distinctive features are extracted from the left and right images, such as corners or edges. In the correspondence matching step, the features from the left and right images are compared to find corresponding features in both images. This is often done by computing a distance metric between the features. In the depth estimation step, the positions of the corresponding features are triangulated to estimate the depth of the scene.
- ➔ One of the most common approaches for stereo matching is to use a disparity map, which is a matrix that contains the disparity value for each pixel in the stereo images. The disparity map can be computed using various techniques, including local matching, global matching, and semi-global matching. These techniques use different algorithms to find the best matching pixels or features between the stereo images.
- ➔ In local matching, the disparity is computed for each pixel by searching for the best matching pixel within a small window around it. This approach is simple and fast but can lead to errors in regions with textureless or occluded pixels.
- ➔ In global matching, the disparity is computed by considering the entire image and finding the best matching pixels or features globally. This approach is more accurate but requires more computation and can be sensitive to lighting and contrast changes.
- ➔ In semi-global matching, the disparity is computed by combining the advantages of both local and global matching techniques. It uses a cost aggregation method to compare the matching pixels or features locally and globally to produce a more accurate disparity map.

Stereo matching can be challenging due to factors such as occlusions, textureless regions, and lighting changes, which can affect the accuracy of the correspondences and the resulting depth maps.

APPLICATION OF THE STEREO MATCHING ARE:

- ➢ Stereo matching has many applications, such as 3D reconstruction.
- ➢ Object tracking, autonomous navigation, and augmented reality.
- ➢ It is widely used in robotics, automotive, and entertainment industries, among others.

Question 3: Write down steps to stitch images to create the panorama. Use the three Taj Mahal Images provided with this assignment to create one panorama. Show panorama into one.

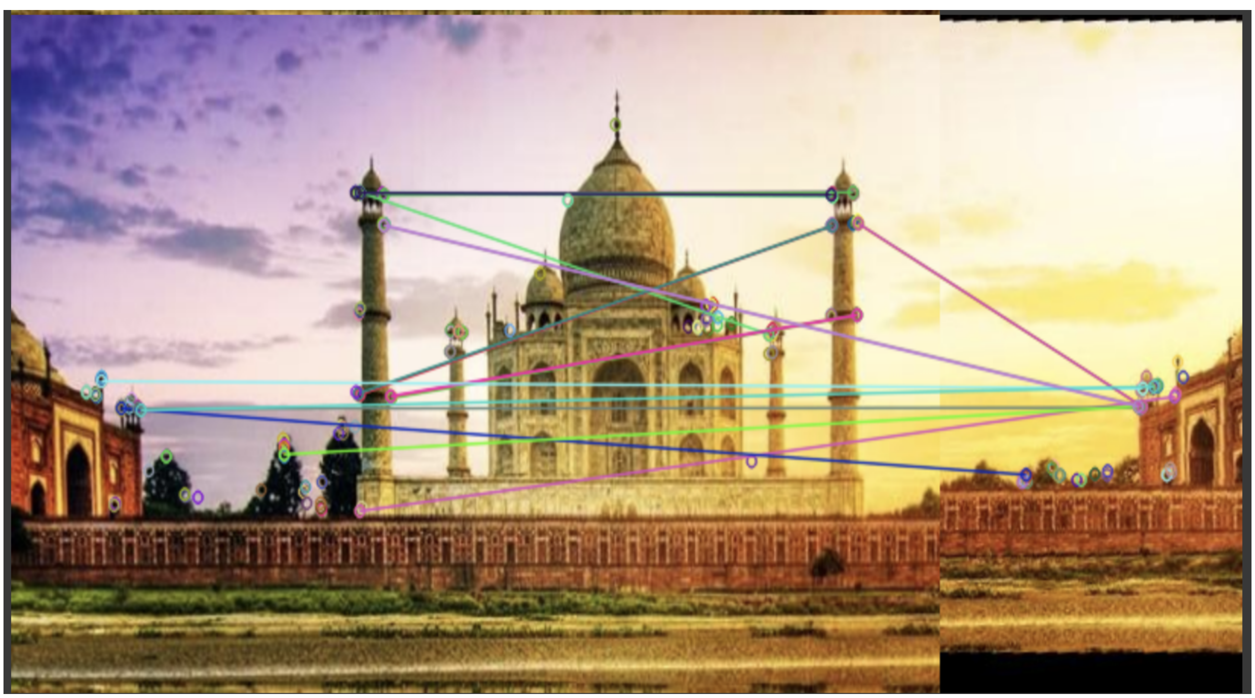
Steps Followed :-

1. Find the keypoints and descriptors with ORB.
2. Brute-Force matcher is used which takes the descriptor of one feature in the first set and it is matched with all other features in the second set using distance calculation.
3. Distance is the similarity score between two keypoints.
4. Creates a list of all matches, just like keypoints and sorts them in the order of their distance.
5. Overlay one image over the image and obtain a resultant image
6. Follow the same steps for the third image to create the panorama.

Given Images :-

Results with Image 1 and Image 2:-

Final Result with above resultant image and Image 3:-



Colab Link : [Solution3_2B]