EE5175: Image Signal Processing - Lab 3 report

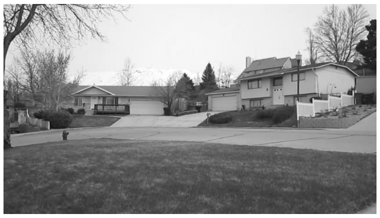
Ruban Vishnu Pandian V (EE19B138)

February 21, 2023

Contents

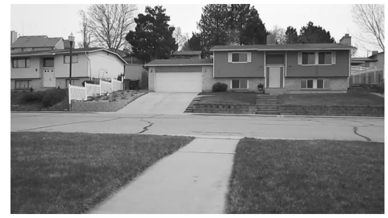
1 Image mosaicing

Aim: Given three images of the same scene obtained by panning the camera from left to right, the homographies relating the images must be found and using them, a single image must be formed by stitching the three images (The scene is assumed very far away from the camera so that it can be assumed to be planar and hence, homography can be applied).



(a) Left image

(b) Middle image



(c) Third image

2 Calculating homography from point correspondences:

Homography is a matrix which relates two images of a planar scene. Given point correspondences (x, y) and (x', y'), we can write the following equation:

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

Normalizing w.r.t z' and writing the equations with the homography elements as the unknowns, we have:

$$\begin{bmatrix} -x & -y & -1 & 0 & 0 & 0 & xx' & yx' & x' \\ 0 & 0 & 0 & -x & -y & -1 & xy' & yy' & y' \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Based on the matrix equation above, we can form a 8x9 matrix using four point correspondences. Finding the nullspace of that matrix would give us the homography matrix elements. A general homography matrix will have eight unknowns in it. This is because all scaled versions of a homography matrix are similar and hence, we can set an element value as '1' in the matrix by dividing every other element by it. Hence, only 8 unknowns are there.

3 Homography computation using SIFT and RANSAC:

Scale invariant feature transform (SIFT) is an algorithm to detect local features in an image. In the given assignment, SIFT algorithm is used to obtain the point-to-point correspondences between the images img1 and img2 and also between img2 and img3. Using these point-to-point correspondences, homographies relating the images are found.

To find the homography matrices, the **Random sample consensus (RANSAC)** algorithm is used. In this algorithm, the following steps are followed:

1. Randomly four different point correspondences are picked to compute the homography matrix as discussed in the previous section.

2. Once the homography is obtained, it is applied on the remaining point correspondences as shown:

$$\begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix} = H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$x'' \leftarrow \frac{x''}{z''}$$
$$y'' \leftarrow \frac{y''}{z''}$$

3. Once the estimates are obtained, euclidean error is calculated as follows:

$$\sqrt{(x - x'')^2 + (y - y'')^2}$$

   If it is less than a threshold ϵ, that point correspondence is added to the consensus set.

4. After applying this procedure for all point correspondences, the relative size of the consensus set is found. If it is greater than a threshold fraction, the algorithm is terminated and the homography is assumed to be found for the given thresholds.

5. Else, the algorithm is restarted from the first step.

Once we have the homography matrices relating the images, we can stitch the three images using target-to-source mapping.

4 Image mosaicing using the homography matrices:

To stitch the images together, an empty canvas of suitable dimensions is created. Let the homography relating img2 and img1 be H_{21} for img2 and img3, let it be H_{23}. It is assumed that the final mosaiced image is taken from the frame of reference of the second image. For target co-ordinates are (i, j), the following target-to-source mappings are done:

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = H_{21} \begin{bmatrix} i \\ j \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = I \begin{bmatrix} i \\ j \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} = H_{23} \begin{bmatrix} i \\ j \\ 1 \end{bmatrix}$$

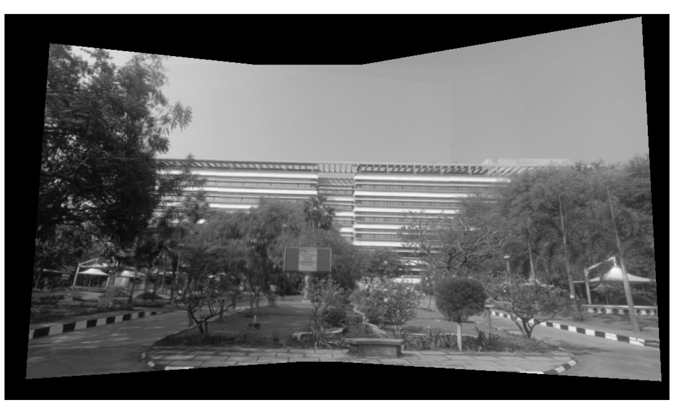
Once the source co-ordinates on each image are obtained, the respective source intensities are obtained using Bilinear interpolation. A final target intensity is obtained by averaging the source intensities. The final mosaiced image is given below:

Mosaiced image

Note: The homographies are calculated using point correspondences under the assumption that the image center is the origin and not the top left corner of the image. Also, the mosaiced image dimensions are taken to be $Nx = 1.6Nx\_original$ and $Ny = 2.5Ny\_original$. To fit the image within the canvas, offsets are also used: $Offset\_x = 10$ and $Offset\_y = 90$. These values are hard-coded based on the given images.

The New academic complex (NAC) building was captured from a location far away from it by panning the camera from left to right. The image dimensions are 480 X 640. Hard-coded target image parameters: $Nx = 1.3Nx\_original$, $Ny = 1.68Ny\_original$, $Offset\_x = -10$ and $Offset\_y = 20$. These images were stitched together using the same code. The captured images and the mosaiced result are given below:

(a) Left image (b) Middle image (c) Third image



Mosaiced own image

5 Observations and Conclusions

1. The mosaic image looks convincing to some extent. Rotations, which are not fronto parallel, are still captured by the homography matrices.

2. At points where two images are stitched, the transition is not so smooth. This is because points at those locations have different intensities in different images and hence, when averaged and looked as one whole image, the transition is clearly visible.

3. The underlying assumption that the scene is very far away from the camera is essential since it implies that the scene is somewhat planar. The whole homography concept is applicable only if the scene is planar.