16-720 Computer Vision: Homework 2
Feature Descriptor & Homographies & RANSAC

Abhishek Bhatia (abhatia1@andrew.cmu.edu)

2/18/2016

1) Keypoint Detector

1.1) Gaussian Pyramid

Figure 1: Gaussian Pyramid

1.2) Computing Visual Words

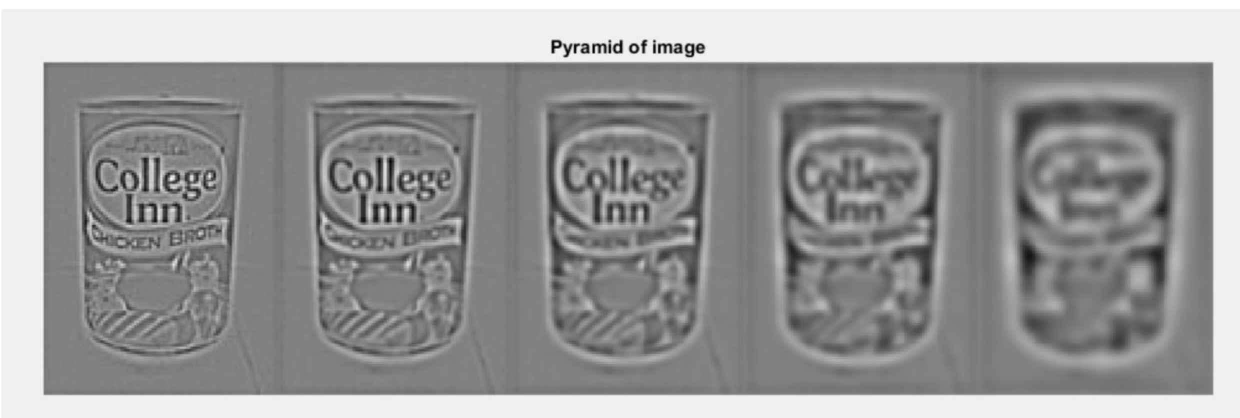

Figure 2: DoG Pyramid

1.5) Edge Suppression and Detecting Extrema

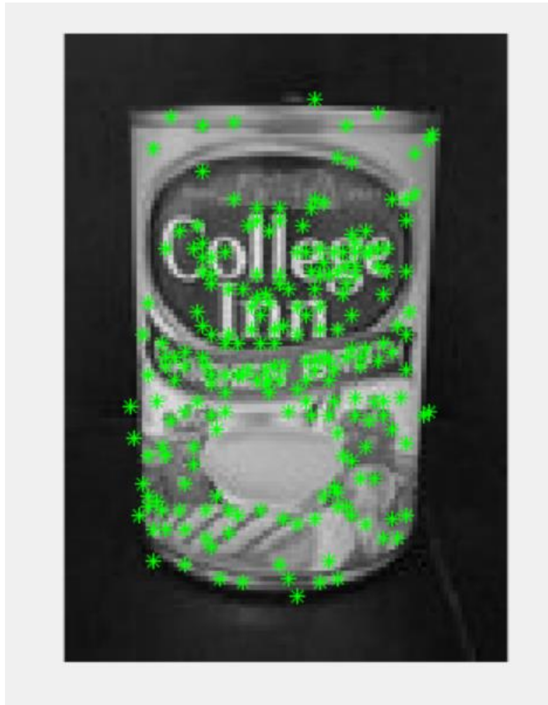

Figure 3: Interest Point (keypoint) Detection with edge suppression for Model_chickenbroth.jpg

2) BRIEF Descriptor

2.4) testMatch

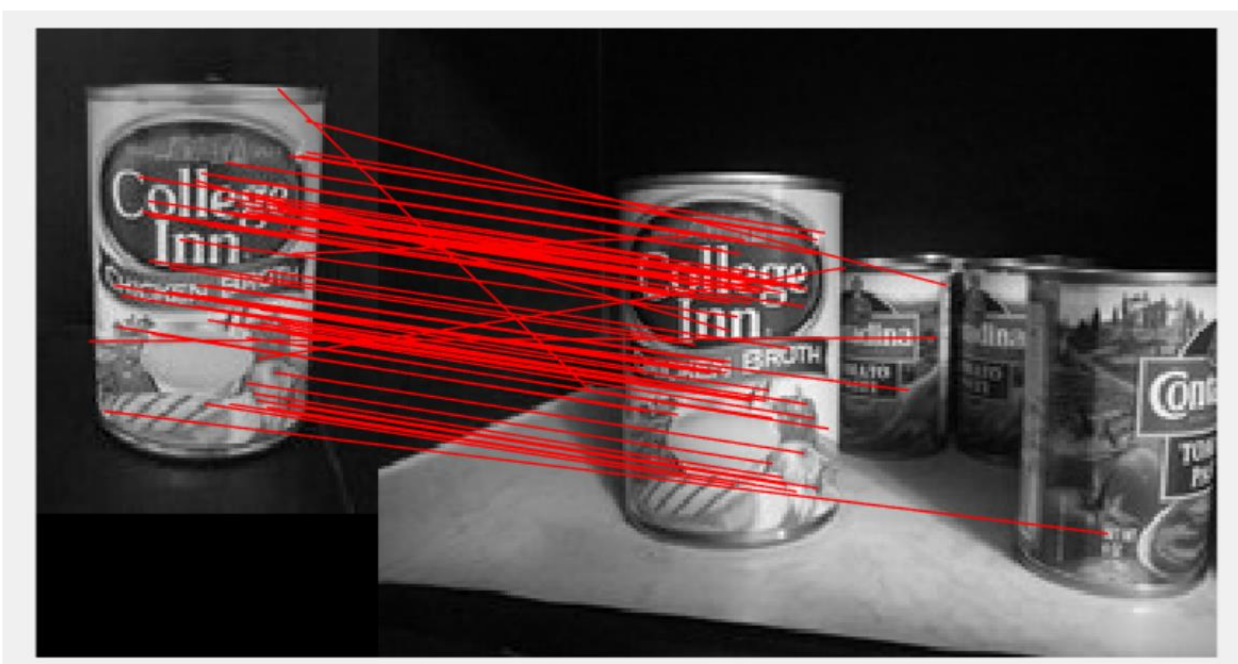

Figure 4: BRIEF matches for model_chickenbroth.jpg and chickenbroth_01.jpg

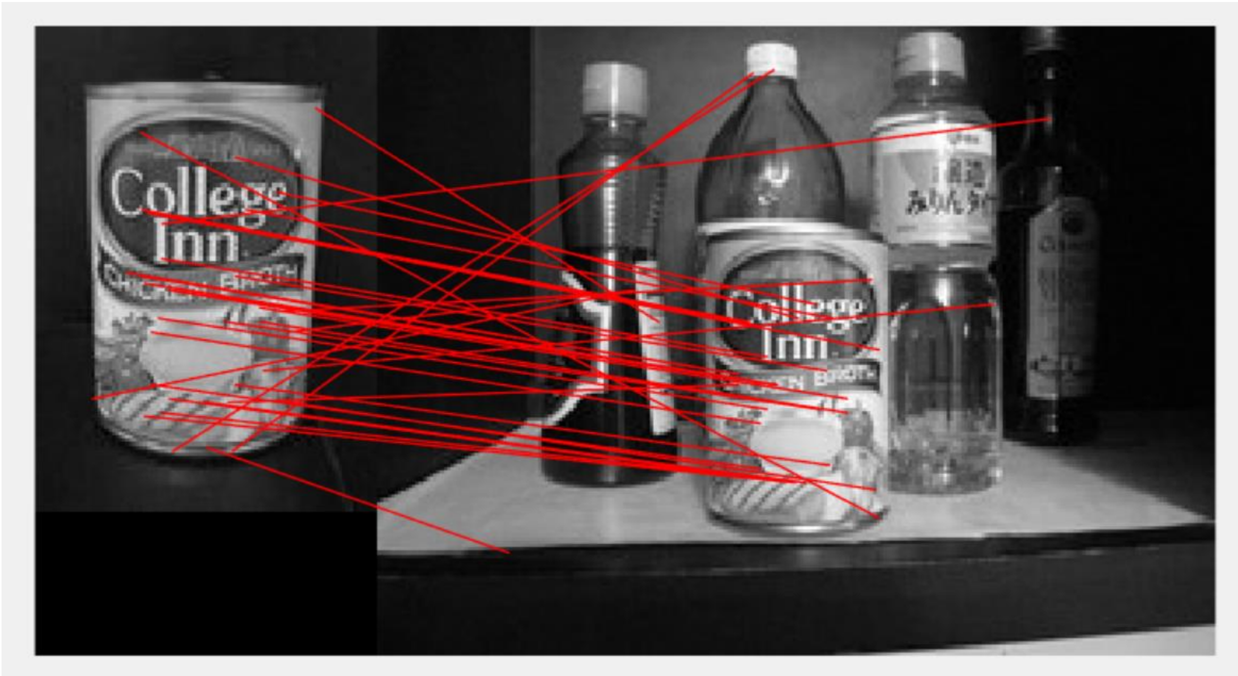

Figure 5: BRIEF matches for model_chickenbroth.jpg and chickenbroth_02.jpg


Figure 6: BRIEF matches for incline_L.png and Incline_R.png

Figure 7: BRIEF matches for pf_pile.jpg and pf_scan_scaled.jpg

Figure 8: BRIEF matches for pf_desk.jpg and pf_scan_scaled.jpg

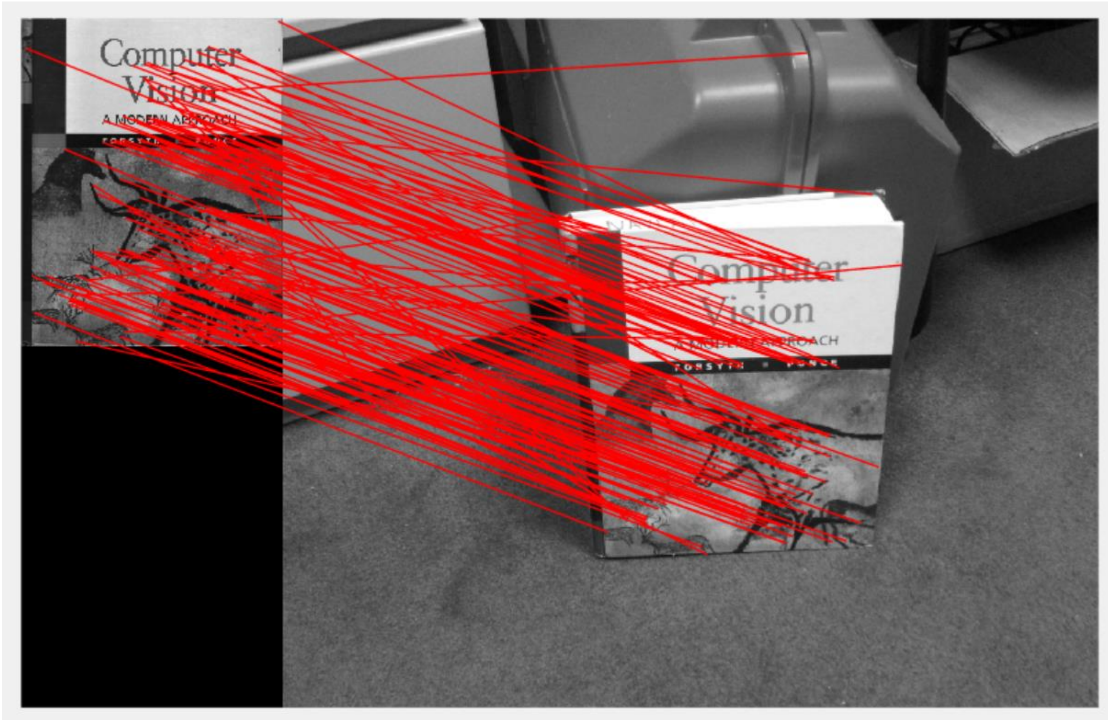

Figure 9: BRIEF matches for pf_stand.jpg and pf_scan_scaled.jpg

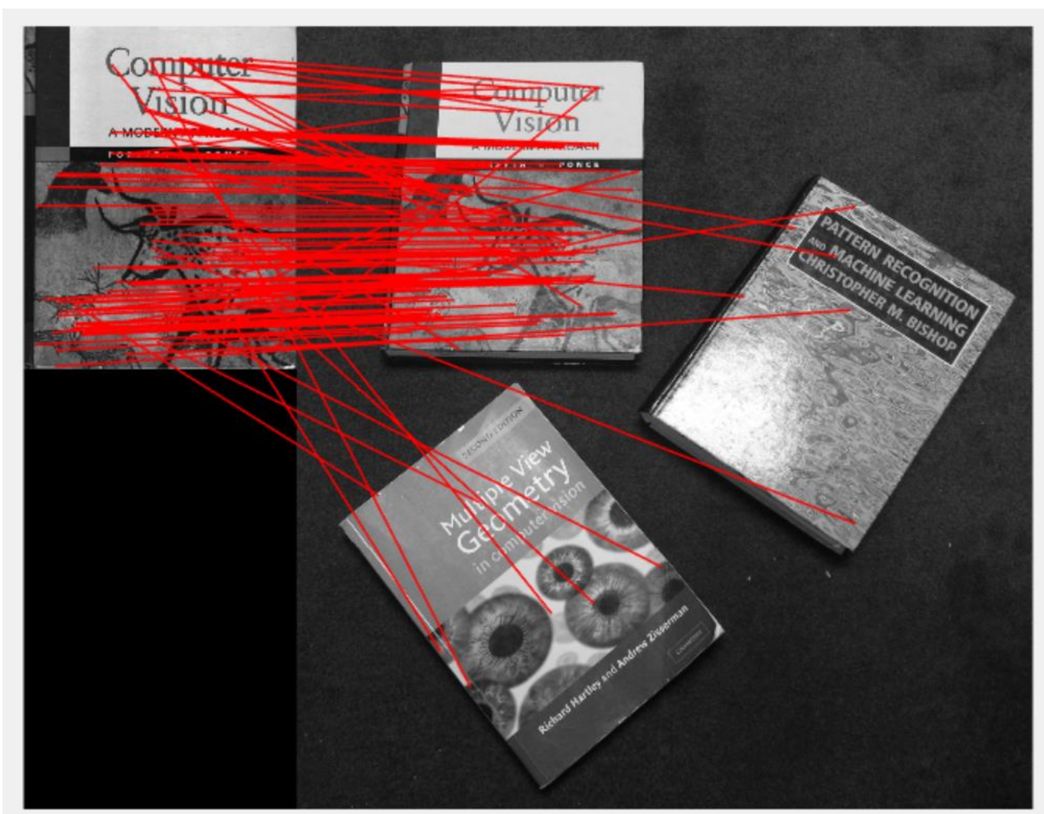

Figure 10: BRIEF matches for pf_floor.jpg and pf_scan_scaled.jpg

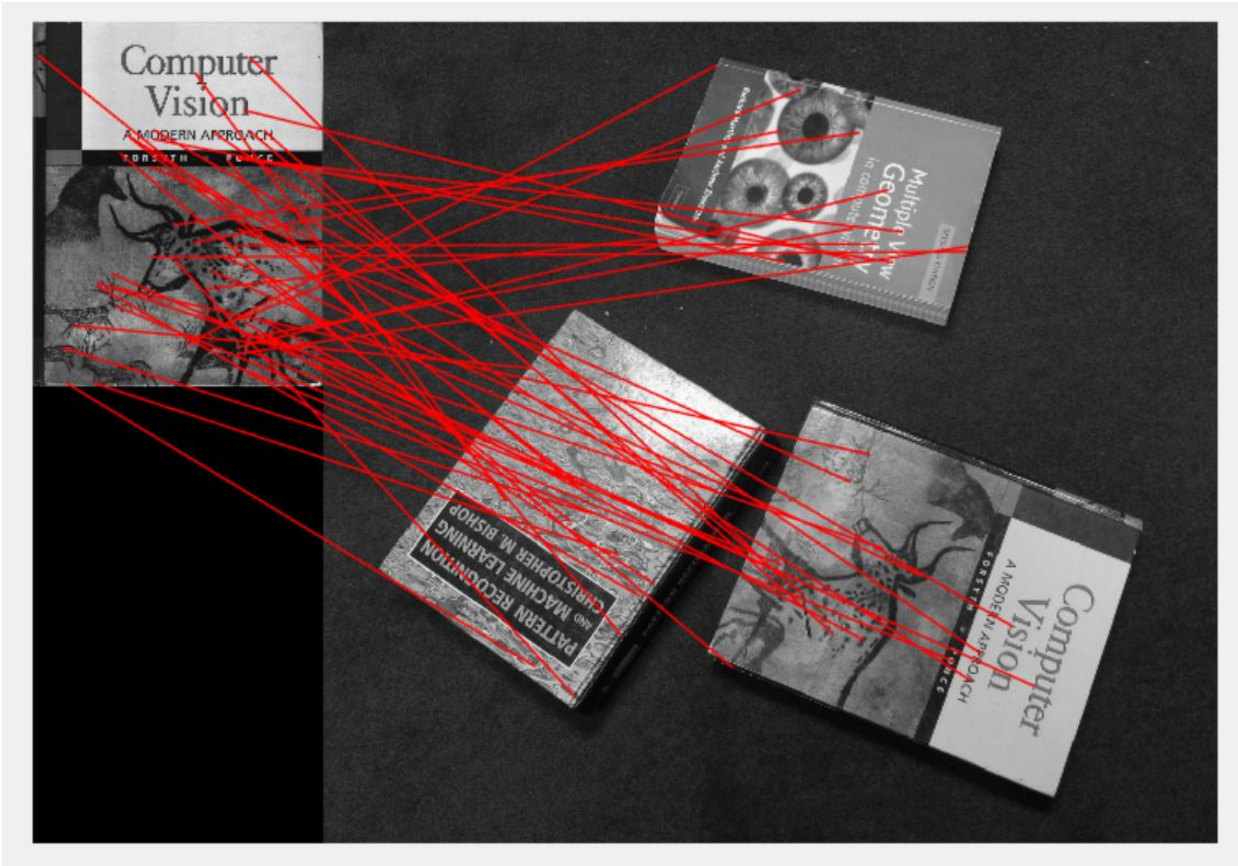

Figure 11: BRIEF matches for pf_floor_rot.jpg and pf_scan_scaled.jpg

Cases that perform worse or better:

1) ChickenBroth: Of all the chickenbroth images, the comparison of model_chickenbroth and chickenbroth_02 is the worst probably because of presence of other bottles with text labels. These text labels provide similar interest points that the descriptor tries to match between the two images and gives false results. Besides, other chickenbroth combinations performed relatively better.
2) Incline: The incline images have lots of matches mostly because of the scale of images. The two images cover great area with many buildings and interest points. But, the presence of similar buildings and features is also the reason for many outliers. In general, the descriptor performed good on the incline images as it found many true matches.
3) PF: The pf image set performs poorly of the three sets because of multiple reasons that include rotation and multiple books with similar features and interest points present in the same image. BRIEF descriptor performs poorly in cases of rotation and scale because it lacks rotation and scale invariance.

2.5) BRIEF and rotations

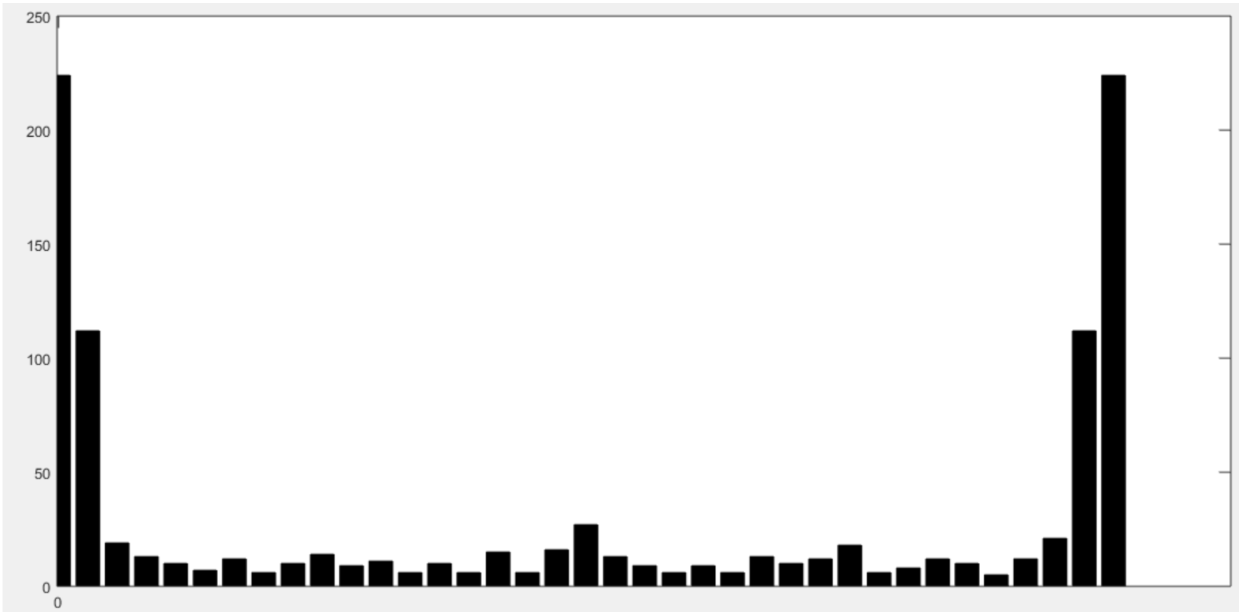

Figure 12: Rotation test on model_chickenbroth.jpg. Y_axis represents the number of matches and X_axis represents the rotation angles from 0 to 360 degree at increments of 10 degrees.

From the Figure 12, it can be seen how the number of matches computed between 2 images where second image in continuously rotated at an increments of 10 degrees vary. The number of matches decrease considerably as soon as the second image start to rotate, but this decrease is not continuous, the number of matches decrease upto 50 degrees and then increase from 50-90 degree. The similar pattern seems to be repeating. It is also observed that the number of matches at 90 degrees and 180 degrees are considerably high than other values. Now, we know that BRIEF works by comparing pixel intensities between pixels selected randomly from square patches (patchWidth x patchWidth) around the interest points between different images. With rotation, the ideal interest region around the interest point also rotates but using this algorithm the descriptor searches for matches within the same square patch, neglecting the rotation. This causes the matches to decrease with rotation. Certain cases like 90 degree rotation and 180 degree rotation perform considerably better because the patch still contains the interest region due to the symmetry around the interest point.

3) Planar Homographies

3) Planar Homographies:

3-1) $p = \{p^1, p^2, \ldots p^N\}$

$q = \{q^1, q^2 \ldots q^N\}$

$$p^i \equiv H q^i \quad \{H \rightarrow \text{homography}\} \quad — (1)$$

$$p^i = \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}, \quad q^i = \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix}$$

a) using (1)

$$p^i \equiv H q^i \Rightarrow \lambda p^i = H q^i$$

$$\Rightarrow \lambda \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix}$$

$$\Rightarrow \lambda = H_{31} u_i + H_{32} v_i + H_{33} \quad — (2)$$

$$x_i (H_{31} u_i + H_{32} v_i + H_{33}) = H_{11} u_i + H_{12} v_i + H_{13}$$

$$\Rightarrow H_{31} x_i u_i - H_{11} u_i + H_{32} x_i v_i - H_{12} v_i + H_{33} x_i - H_{13} = 0 \quad — (3)$$

Similarly, $H_{31} y_i u_i - H_{21} u_i + H_{32} y_i v_i - H_{22} v_i + H_{33} y_i - H_{23} = 0 \quad — (4)$

From (3) and (4)

$$\begin{bmatrix} -u_i & -v_i & -1 & 0 & 0 & 0 & x_i u_i & x_i v_i & x_i \\ 0 & 0 & 0 & -u_i & -v_i & -1 & y_i u_i & y_i v_i & y_i \end{bmatrix} \begin{bmatrix} H_{11} \\ H_{12} \\ H_{13} \\ H_{21} \\ H_{22} \\ H_{23} \\ H_{31} \\ H_{32} \\ H_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Similarly, for each N points, we can derive 2N independent linear equations of the above form and express them together as:

$$\boxed{Ah = 0}$$

where h is a vector of the elements of H and A is a matrix composed of elements derived from point coordinates.

b) ~~#~~ h has 9 elements corresponding to each element of Homography matrix.

c) Homography matrix (H) has 9 elements, but the degree of freedom of this Homography matrix is 8 because of scale ambiguity. In other words, the homography matrix H is used to determine the homogenous coordinates of a 2D point on one camera into another camera:

$$p \equiv Hq$$

Then any scalar non-zero multiple of the ~~homogenous~~ homography matrix generates the same point:

$$(cH)q \equiv cp$$

$$\text{and} \quad cp = p$$

because of scale ambiguity of homogenous coordinates.

This reduce one degree of freedom of homography.

Now, to solve for 8 variables, we need 8 independent linear equations which can be generated by 4 correspondance points pairs, as from 3 and 4 we see that each correspondance point pair generates 2 linear equations.

d) To solve $Ax = 0$

or minimize this homogeneous linear least squeres system, we will use Singular Value decomposition (SVD)

We need to find a solution that minimizes $||Ah||^2$ with the constraint $||h|| = 1$.

when we compute the singular value decomposition of A, we get

$$A = USV^T$$

Now, $\min ||Ah||^2$ happens when we take the right singular vector, i.e column from V that corresponds to smallest singular value from S.

This column gives us the vector h that minimizes $||Ah||^2$ and we reshape this vector into the Homography matrix.

5) Panorama

5.1) With Clipping

Figure 13: Warped Image output of incline_R.png

5.2) Without Clipping

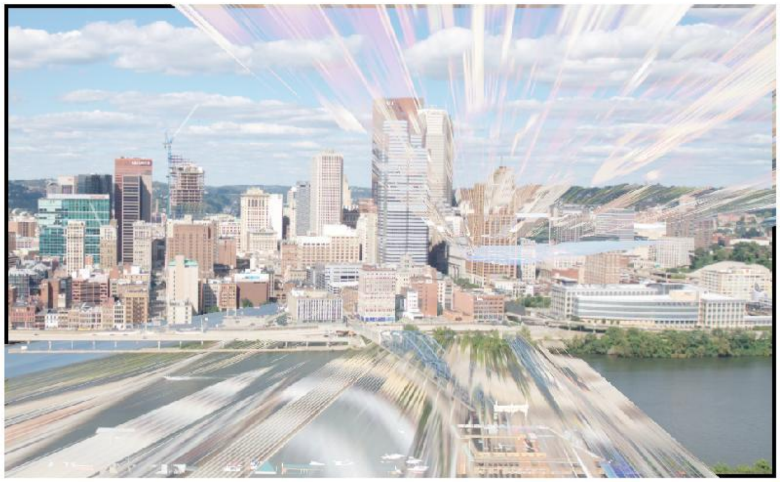

Figure 14: Blended Image output of incline_L.png and incline_R.png, without clipping

6.2) Ransac (Image Panorama)

Figure 15: Image Panorama (With Clipping). With homography estimated using RANSAC.

Figure 16: Image Panorama (Without Clipping). With homography estimated using RANSAC.

Refrences:

1) Homography Estimation: https://cseweb.ucsd.edu/classes/wi07/cse252a/homography_estimation/homography_estimation.pdf