Knarula, Kartik Narula
16720- Computer Vision
HW2

Q1.1 Homography (5 points)
Let the point be X in the space
Then we have,
$X_1=P_1X$ and $X_2=P_2X$
Or $X_1=[A_1\ B_1]\ X$ and $X_2=[A_2\ B_2]\ X$
Where, A_1, A_2 and B_1, B_2 are 3✕3 and 3✕1 matrices respectively

We can write,
$X_1=A_1X^P+B_1$ and $X_2=A_2X^P+B_2$, where X^P is the object point in non-homogeneous coordinates
Hence, $X^P=A_1^{-1}(X_1-B_1)$
Therefore,
$X_2=A_2A_1^{-1}(X_1-B_1)+B_2$

$=A_2A_1^{-1}X_1-A_2A_1^{-1}B_1+B2$

$=A_2A_1^{-1}X_1+C$, where $C= -A_2A_1^{-1}B_1+B2$, a 3✕1 matrix

X_2 can be represented as some linear combination of X_1. Hence, the homography exists for invertible matrices A_1 and A_2,

Q1.2
1. H has 8 degrees of freedom
2. 4 point pairs are required to solve H
3.

Knarula, Kartik Narula
16720- Computer Vision
HW2

$$x_2 = H x_1$$

for matching points (u, v) in x_2 and (x, y) in x_1 we have

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

or $u = \dfrac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}$

and $v = \dfrac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}}$

or $h_{11}x + h_{12}y + h_{13} - h_{31}xu - h_{32}yu - h_{33}u =$

and $h_{21}x + h_{22}y + h_{23} - h_{31}xv - h_{32}yv - h_{33}v$

We can write above two equations and similar equations for 4 set of points a

$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1u_1 & -y_1u_1 & u_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1v_1 & -y_1v_1 & v_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2u_2 & -y_2u_2 & -u_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2v_2 & -y_2v_2 & -v_2 \\ \vdots & & & \vdots & & & \vdots & & \\ \vdots & & & \vdots & & & \vdots & & \end{pmatrix} \begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ \vdots \\ \vdots \\ h_{33} \end{pmatrix} = 0$$

where $(x_1, y_1)(u_1, v_1)$ $(x_2, y_2)(u_2, v_2)$ $(x_3, y_3)(u_2, v_2)$ $(x_4, y_4)(u_4, v_4)$

Knarula, Kartik Narula
16720- Computer Vision
HW2

The above equation is of form
$Ah = 0$

Hence $A =$

$$
\begin{bmatrix}
x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1u_1 & -y_1u_1 & -u_1 \\
0 & 0 & 0 & x_2 & y_2 & 1 & -x_2v_1 & -y_2v_2 & -v_2 \\
x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2u_2 & -y_2u_2 & -u_2 \\
0 & 0 & 0 & x_2 & y_2 & 1 & -x_2v_2 & -y_2v_2 & -v_2 \\
x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3u_3 & -y_2u_3 & -u_3 \\
0 & 0 & 0 & x_3 & y_3 & 1 & -x_3v_3 & -y_3v_3 & -v_3 \\
x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4u_4 & -y_2u_4 & -u_4 \\
0 & 0 & 0 & x_4 & y_4 & 1 & -x_4v_4 & -y_4v_4 & -v_4
\end{bmatrix}
$$

4. The trivial solution for the equation would be when H is the null vector and would hold true when the points x_1 and x_2 are in the same plane that includes the camera centers. In other words, the camera centers form the basis of the set of points in x_1 and x_2.

   The matrix will be full rank(9). The eigenvalues of the matrix will be positive. All the vectors will be transformed into a 9*1 space and we will get 9 independent eigen vectors.

Knarula, Kartik Narula
16720- Computer Vision
HW2

Q1.3

$x_1=K_1[I\ 0]X$, $X=[X^P\ 1]^T$, where X^P is the image point in non-homogeneous
$x_1=[K_1X^P]+0$, where K_1X^P is a 3*1 matrix and 0 represents a zero vector of size 3*1
Similarly, $x_2=[K_2RX^P]+0$

Hence, we have,
$X^P=K_1^{-1}x_1$ and $X^P=K_2^{-1}Rx_2$
Therefore, $K_1^{-1}x_1=K_2^{-1}Rx_2$

Or, $x_1=K_1K_2^{-1}Rx_2$
Or, $x_1=Hx_2$, where $H=K_1K_2^{-1}R$
Hence, homography exists for two cameras with pure rotation.

Knarula, Kartik Narula
16720- Computer Vision
HW2
1.4

A rotation matrix for angle θ is given by:

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Rotation matrix for angle 2θ

$$R(2\theta) = \begin{bmatrix} \cos 2\theta & -\sin 2\theta & 0 \\ \sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2\theta - \sin^2\theta & -2\sin\theta\cos\theta & 0 \\ 2\sin\theta\cos\theta & \cos^2\theta - \sin^2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= [R(\theta)]^2$$

Homography matrix for rotation 2θ

$$H_{2\theta} = K\,R(2\theta)$$

$$= [K\,R(\theta)]^2 = [H(\theta)]^2$$

$$\left\{ \begin{array}{l} K^2 = K \\ = \text{constant} \\ \text{for a camera} \end{array} \right\}$$

Q1.5
Planar Homography cannot map scenes that are in different planes, except for the scenes that are separated by pure rotation.

Knarula, Kartik Narula
16720- Computer Vision
HW2

1.6

Let a line be

$$ax + by = c$$

and (x, y) be points on this line.

$x_1 = Px$ (where x_1 and y_1 are projected points)

$y_1 = Py$

or $a\frac{x_1}{P} + b\frac{y_1}{P} = c$ (from original eq)

$\Rightarrow$

$$ax_1 + by_1 = cP$$

Hence, the relationship between points is preserved.

Perspective projection preserves lines.

Knarula, Kartik Narula
16720- Computer Vision
HW2

2.1.1
The following points may be noted for FAST detector and Harris Corner

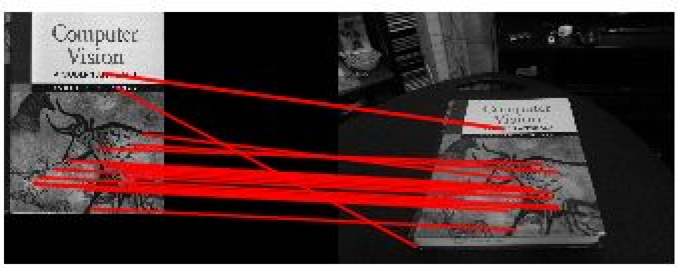
a) Both FAST and Harris scan for regions across the interest point and qualify it as a feature if some calculated quantity exceeds some threshold value.
b) FAST is based on the brightness values while Harris calculates the gradients.
c) Harris uses a gaussian weighting scheme to weight the points around the interest point but FAST does not.
d) FAST detector is found to be faster than Harris detector.

2.1.2

The BRIEF descriptor is binary while the histogram representation of pixels created using the filter-banks is discrete. We can use filter banks as descriptors but I don't think it will be as accurate and robust as BRIEF. The success of filter banks depends on more number of hyperparameters(K, alpha etc) while BRIEF is based on random sampling.

2.1.3
Hamming distance is more computationally friendly. It is just the number of 1's in the XOR operation between the two vectors.
Moreover, for binary vectors, euclidean distance is just the square root of the hamming distance. Counting is a faster operation than multiplying(a lot of terms are multiplied by 0 in this case, for calculating Euclidean distance)

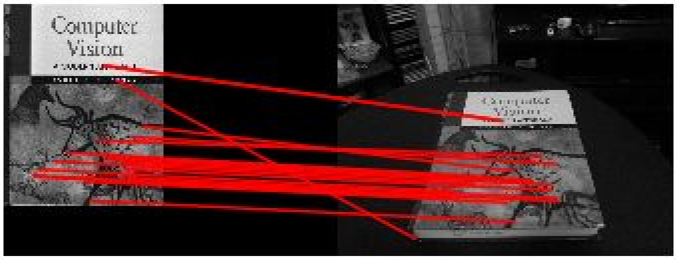
2.1.4



Number of matched points=19

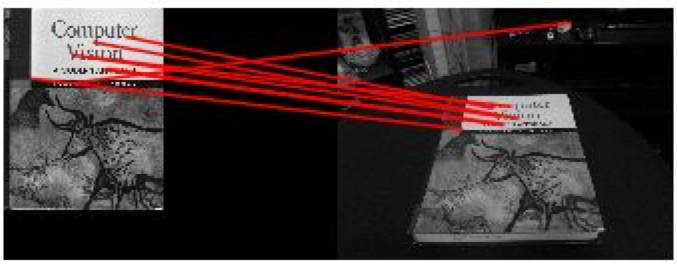
Knarula, Kartik Narula
16720- Computer Vision
HW2

2.1.5
sigma=0.15, ratio-0.7



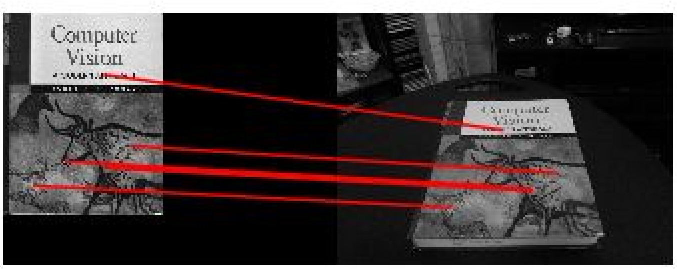
Number of matched points=19
Number of features detected in image1=945
Number of features detected in image2=477

sigma=0.3, ratio=0.7



Number of matched points=8
Number of features detected in image1=320
Number of features detected in image2=132
Less number of features are detected in locs 1 and locs 2. Sigma is the brightness threshold for fast. We should decrease sigma if we want to detect more corners(features)

sigma=0.15, ratio=0.6

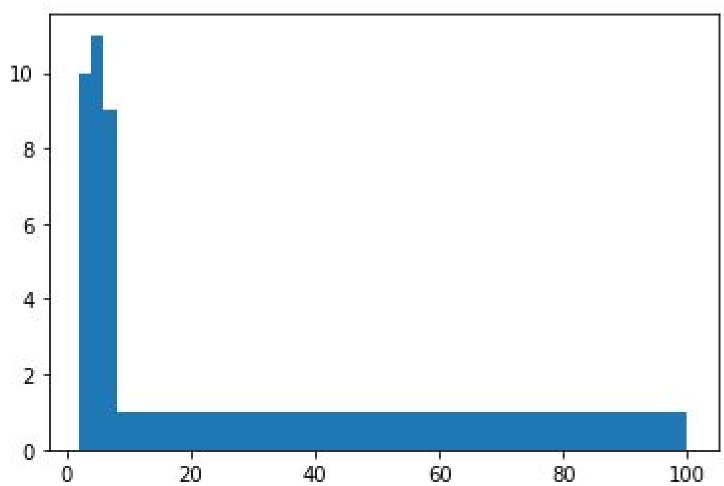


Number of matched points=6
Number of features detected in image1=945
Number of features detected in image2=477
The number of features detected for both the images was same as the default case, but the number of matches dropped from 19 to 6.

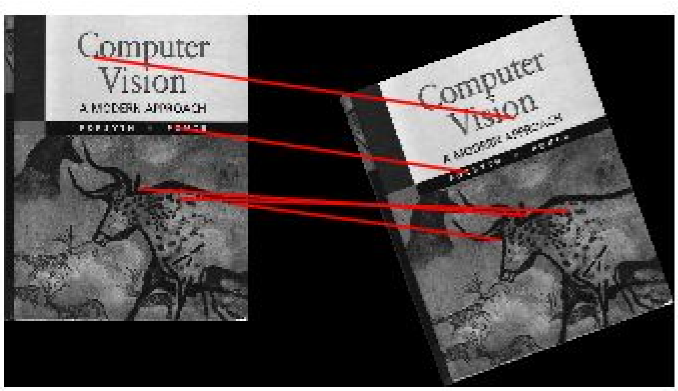
Knarula, Kartik Narula
16720- Computer Vision
HW2

The parameter sigma is the part of the corner detection function and acts as the brightness threshold for the fast detector. We should decrease this value if we want more corners to be detected.
The parameter ratio is the part of the descriptor and is the maximum ratio of distances between the first and second closest descriptors in the second set of descriptors. Decreasing the value reduces the number of matches identified.

2.1.6

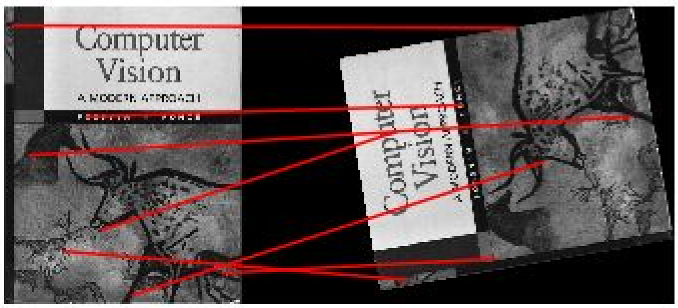


The above figure is the histogram for the number of matches. The one value at extreme right corresponds to the rotation of 360 degrees where the BRIEF descriptor finds the maximum number of matches. The other values are less than 10.

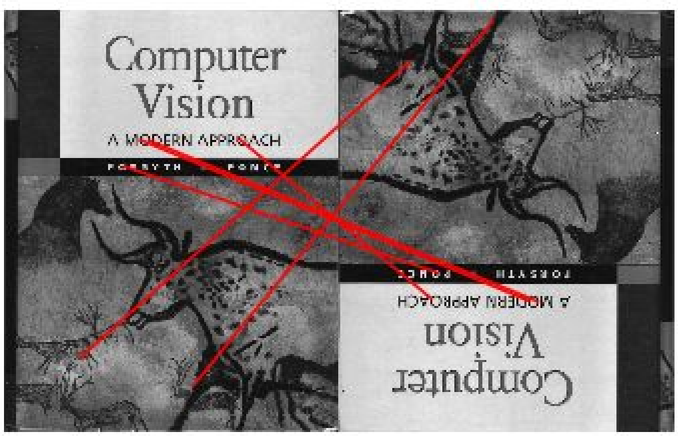


Matches with the image rotated 20 degrees

Knarula, Kartik Narula
16720- Computer Vision
HW2



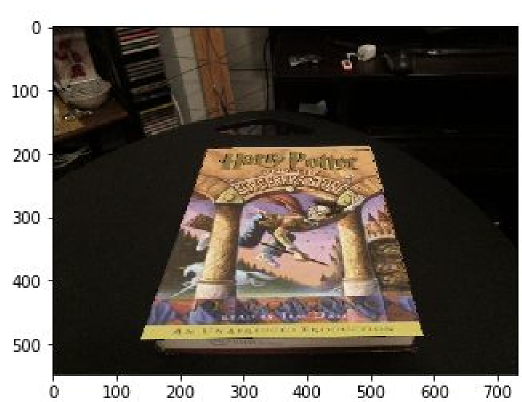
Matches with image rotated 100 degrees



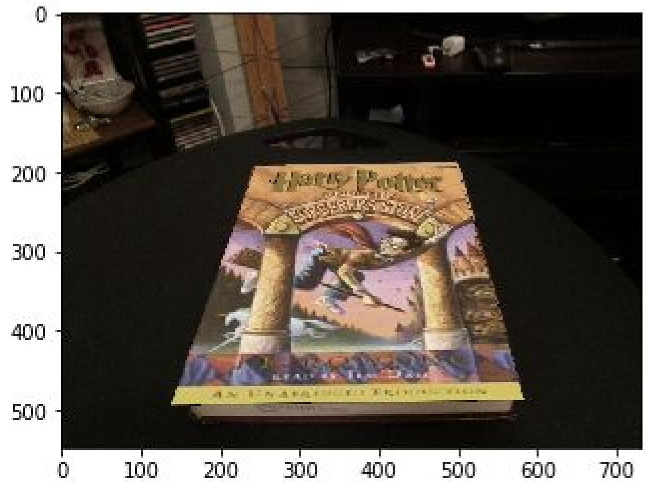
Matches with image rotated 180 degrees

BRIEF works with a set rule of matching pattern and fails to detect matches when the image is oriented in such a way that random sampling method is not able to detect matches.

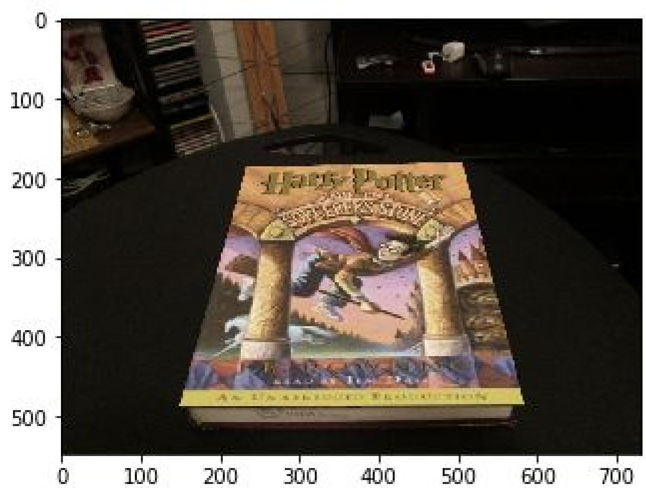
2.2.4



Knarula, Kartik Narula
16720- Computer Vision
HW2
2.2.5
max_iter=500, inlier_tol=2

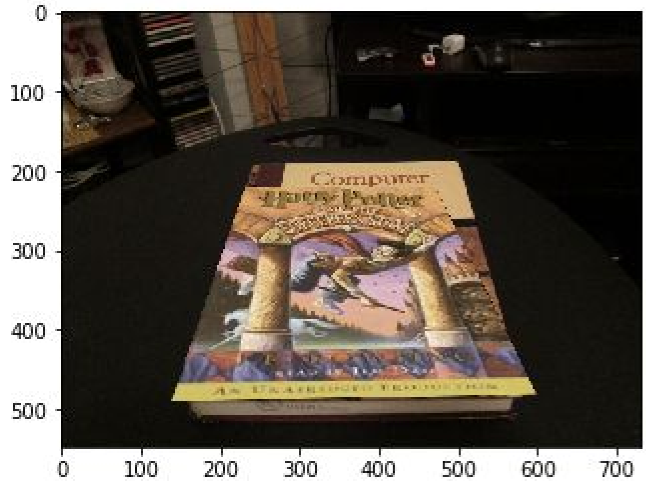


max_iter=500, inlier_tol=0.5



max_iter=10, inlier_tol=2

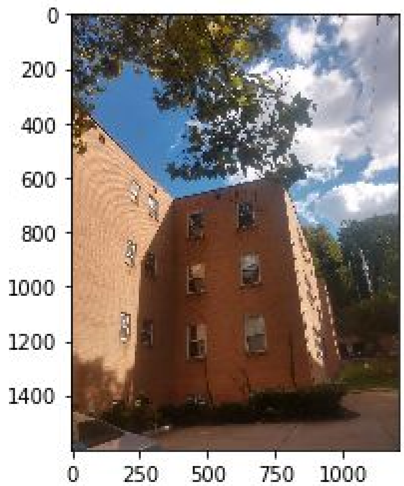
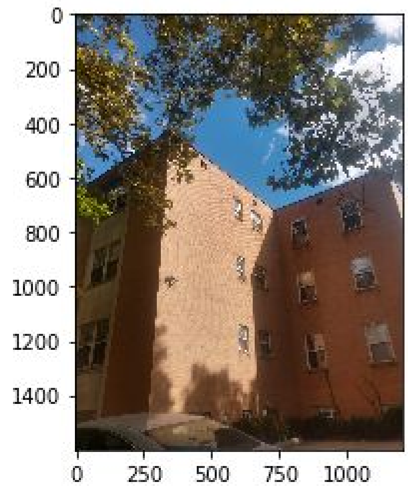
Knarula, Kartik Narula
16720- Computer Vision
HW2

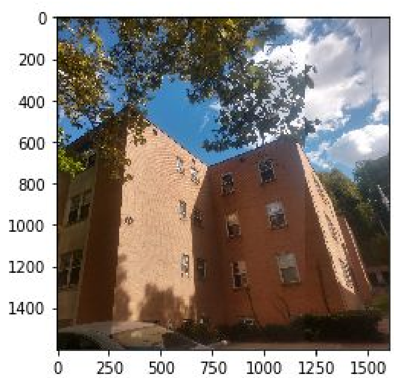


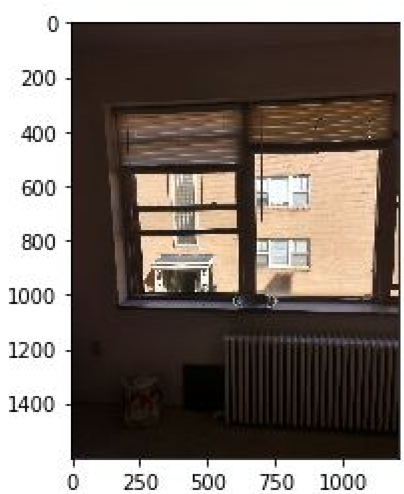
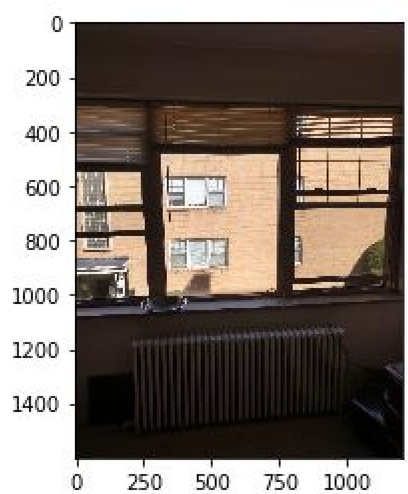
We get a better fit by increasing the number of iterations and reducing the inlier tolerance. Increasing the number of iterations iterates the RANSAC algorithm for larger number of random samples and reports the best homography from those samples. Reducing the inlier tolerance includes only the closed fit points in the homography. So, reducing the inlier tolerance will reduce the chances of including outliers while calculating homography.

Knarula, Kartik Narula
16720- Computer Vision
HW2

4.2x Panorama



Knarula, Kartik Narula
16720- Computer Vision
HW2