Computer Vision HW 2

Harsh Dhruva

andrewID – hdhruva

Q1.1 Homography

Say we are projecting point P through plane Π onto the camera planes of cameras 1 and 2 corresponding to $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$

We have,

$$x_1 \equiv P_1 P$$

$$x_2 \equiv P_2 P$$

$$P_2^{-1} x_2 \equiv P$$

$$\therefore \quad x_1 \equiv P_1 (P_2^{-1} x_2)$$

$$x_1 \equiv P_1 P_2^{-1} x_2$$

$$\boxed{x_1 \equiv H x_2}$$

Q1.2 Correspondences

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

$$\mathbf{x}_1^i \equiv H\mathbf{x}_2^i$$

$$\begin{bmatrix} x_1^i \\ y_1^i \\ 1 \end{bmatrix} \equiv \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_2^i \\ y_2^i \\ 1 \end{bmatrix} \longrightarrow \quad x_1^i = \frac{h_{11}x_2^i + h_{12}y_2^i + h_{13}}{h_{31}x_2^i + h_{32}y_2^i + h_{33}}, \quad y_1^i = \frac{h_{21}x_2^i + h_{22}y_2^i + h_{23}}{h_{31}x_2^i + h_{32}y_2^i + h_{33}}$$

$$h_{11}x_2^i + h_{12}y_2^i + h_{13} - h_{31}x_2^i x_1^i - h_{32}y_2^i x_1^i - h_{33}x_1^i = 0$$

$$h_{21}x_2^i + h_{22}y_2^i + h_{23} - h_{31}x_2^i y_1^i - h_{32}y_2^i y_1^i - h_{33}y_1^i = 0$$

$$\begin{pmatrix} x_2^i & y_2^i & 1 & 0 & 0 & 0 & -x_2^i x_1^i & -y_2^i x_1^i & -x_1^i \\ 0 & 0 & 0 & x_2^i & y_2^i & 1 & -x_2^i y_1^i & -y_2^i y_1^i & -y_1^i \\ & & \vdots & & & & & & \end{pmatrix} \begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ \vdots \\ h_{33} \end{pmatrix} = 0 \Rightarrow A_i h = 0$$

$i = 1 \ldots N$

8X9 for homography

$$\therefore \boxed{A_i = \begin{pmatrix} x_2^i & y_2^i & 1 & 0 & 0 & 0 & -x_2^i x_1^i & -y_2^i x_1^i & -x_1^i \\ 0 & 0 & 0 & x_2^i & y_2^i & 1 & -x_2^i y_1^i & -y_2^i y_1^i & -y_1^i \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}}$$

4. Trivial solution for **h** is h = 0.
   Matrix **A** is full rank because for a non-trivial solution of **h** to exist, rank(A) = 8 since the matrix A is of dimension 8x9 to compute a homography by solving for the values in **h.**
   The rank of the matrix A corresponds to the number of non-zero eigenvalues of the matrix, which correspond to the number of linearly independent eigenvectors. Since DOF of **h** is 8, we need 8 linearly independent eigenvectors to solve for the values in **h**, to obtain non trivial solutions.

Q1.3 Homography under rotation

$$x_1 = K_1[I\ 0]X = [K_1 I\ 0]X = F_1 X \text{ , where } \overset{3\times4}{F_1 = [K_1 I\ 0]}$$

$$x_2 = K_2[R\ 0]X = [K_2 R\ 0]X = F_2 X \text{ , where } F_2 = [K_2 R\ 0]$$

$$\therefore \quad x_1 = \overset{3\times4}{F_1} X$$

$$x_2 = F_2 X$$

$$F_2^{-1} x_2 = X$$

$$x_1 = \overset{3\times4}{F_1}(\overset{4\times3}{F_2^{-1}} x_2)$$

$$x_1 = F_1 F_2^{-1} x_2$$

$$\boxed{x_1 = H x_2}$$

Q1.4 Understanding homographies under rotation

Let us say the first second and third orientations of the camera are **X1, X2,** and **X3** respectively all separated by the angle of rotation θ

$$X_1 = H X_2$$

and, $$X_2 = H X_3$$

rotate

$$X_1 \xrightarrow{\theta} X_2$$

$$X_2 \xrightarrow{\theta} X_3$$

$$\therefore X_1 \xrightarrow{2\theta} X_3$$

$$\therefore X_1 = H(H X_3)$$

$$\boxed{X_1 = H^2 X_3}$$

Q1.5 Limitations of the planar homography

Planar homography can be applied only under the assumption that the world is represented as a 2D plane. With an arbitrary image with different viewpoints, the sets of corresponding points necessary to match may not lie on a single plane, thus making it difficult to warp and match the viewpoints.

Q1.6 Behavior of lines under perspective projections

$$x \equiv PX$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \equiv P \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \equiv A \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + b$$

$$\therefore \quad x = \frac{1}{\lambda}\left([X \; Y \; Z]\, a_1 + b_1\right) = K_1$$

$$y = \frac{1}{\lambda}\left([X \; Y \; Z]\, a_2 + b_2\right) = K_2$$

$\therefore$ There exists a line $x + y - K_1 - K_2 = 0$

for the projection of P in $\boldsymbol{x}$

Q2.1.1 FAST Detector

FAST detector uses a neighborhood of 16 pixels around a particular pixel p and compares the intensity of these pixels to be outside a certain threshold of intensity of pixel p, whereas the Harris corner detector deals with computing gradients of the pixel p and then moving the small window around a pixel to measure the amount of change occurring in the pixel values.

The computational performance of the FAST detector is better since first, it only compares the intensity to the four pixels in the "pixel circle" to determine if it is an interest point. Then it compares the intensity in the rest of the pixels in the pixel circle around it if it considers it as an interest point to further confirm. Harris detector is computationally more intensive since you have to calculate several gradient values, the harris value, and then find all the pixels that exceed a certain threshold and are the local maxima within a certain window.

Q2.1.2 BRIEF Descriptor

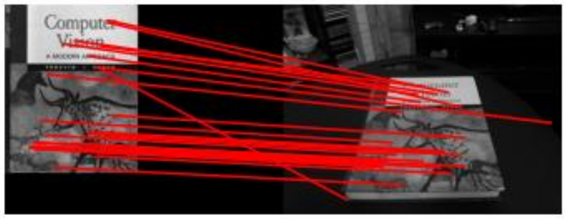
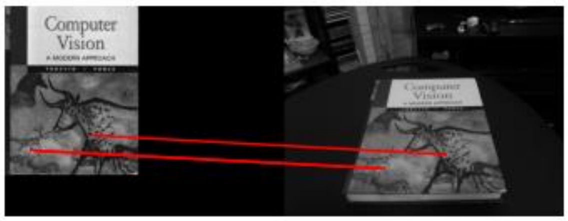
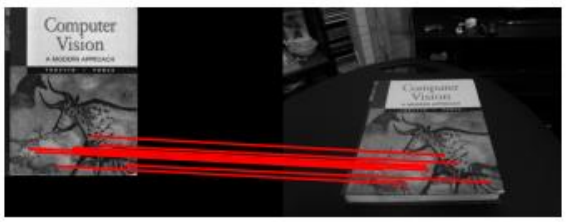
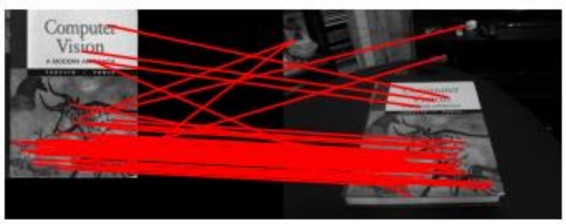
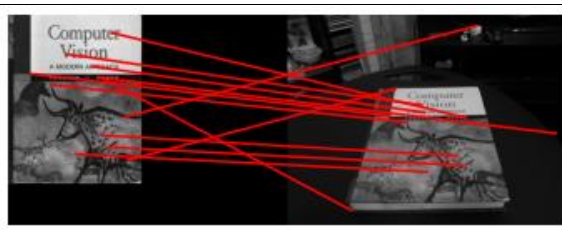
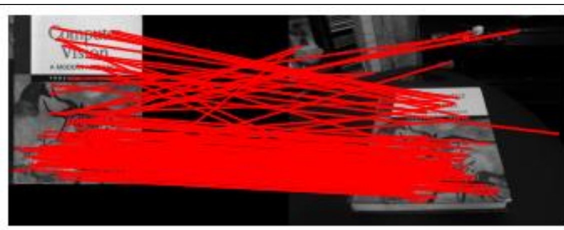
The BRIEF descriptor represents an image patch as a binary string, therefore representing the important features in the image as matrices of binary values, making it easier to compare to other images, making matching or detecting easier. We can use any of the filter banks (with additional parts to the algorithm) to describe features in the image. GIST descriptor uses Gabor filter bank over different scales and blocks in the image, to capture the structure of the images. We used filter banks for scene to scene matching in hw1, by extracting the features in the images.

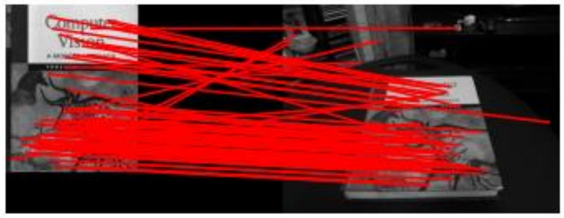
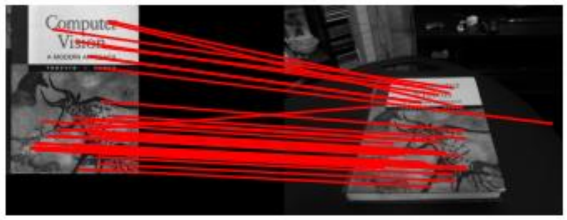
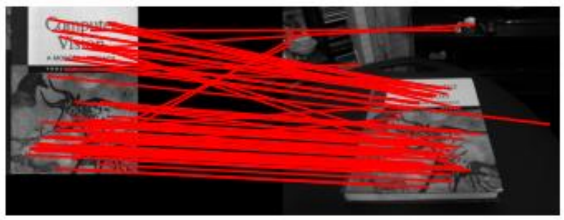
Q2.1.3 Matching Methods

Hamming distance can be used to measure the distance between two descriptors or bit vectors, and then we can look at the relative frequency of hamming distances between the descriptors and use the Nearest Neighbor search in the hamming space to find the descriptor closest to the one we are looking for such that the hamming distance in minimized.

The hamming distance is used for comparing two binary data strings, whereas the Euclidean distance is particularly used to compute the distance between two real valued vectors. In our setting, we are representing the image region as a binary string, therefore using Euclidean distance would not be as informative as the hamming distance.

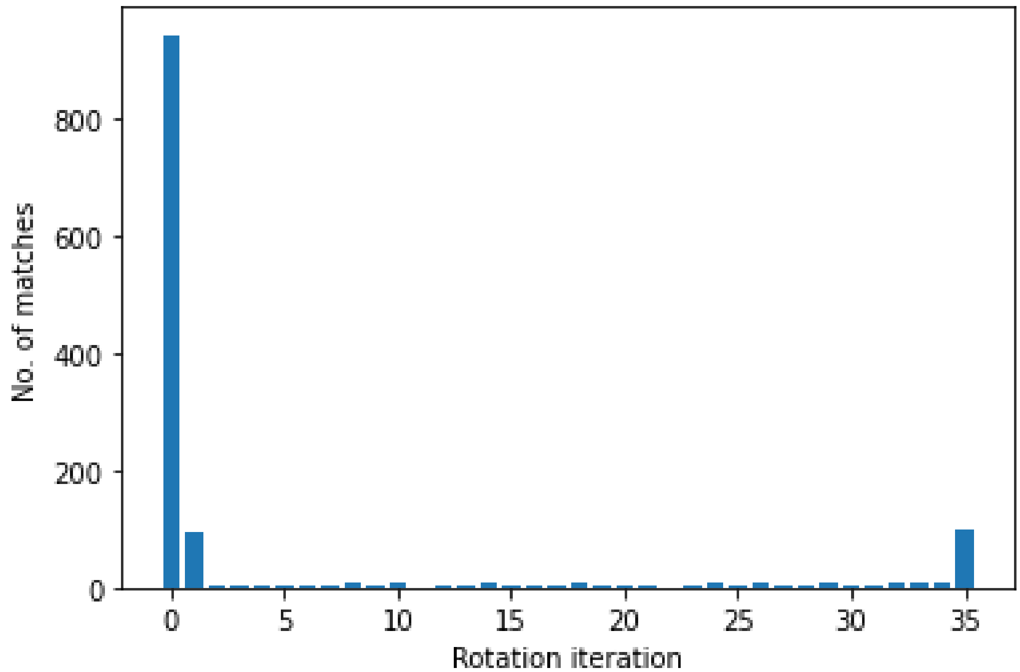
Q2.1.5 Feature Matching Parameter Tuning

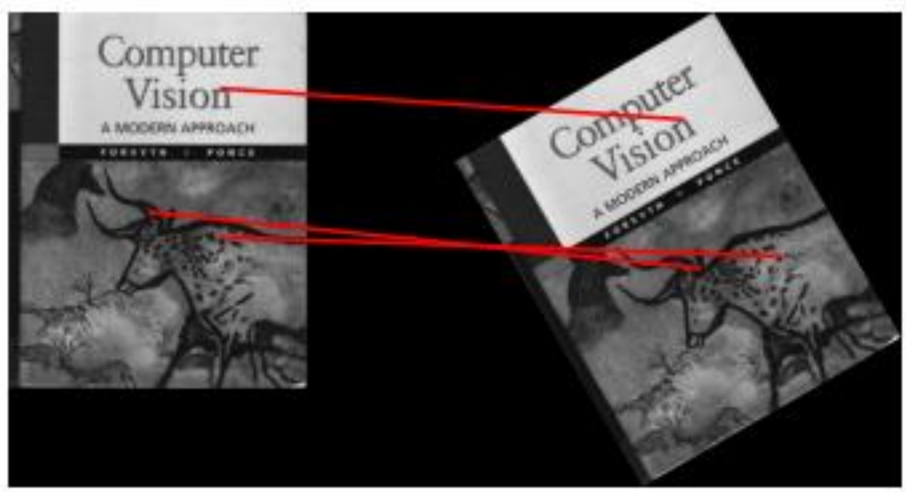
Sigma	Ratio		Comments
0.15 (default)	0.7 (default)		Matches using default sigma and ratio values
0.15	0.5		As ratio is decresed, we see only two matches, since the ratio constricts the nearest neighbors
0.10	0.5		However as the sigma value is reduced too, the lower threshold allows more points to be within nearest neighbors
0.10	0.7		Here the sigma is kept the same and ratio is increased. We see a lot more matches, however some points are matching with background due to lower threshold
0.20	0.7		As threshold increases fewer matches are seen. We also see the black background match to the black in the textbook image
0.10	0.8		Here there are a lot more matches due to relaxed threshold and ration, allowing for more nearest neighbors

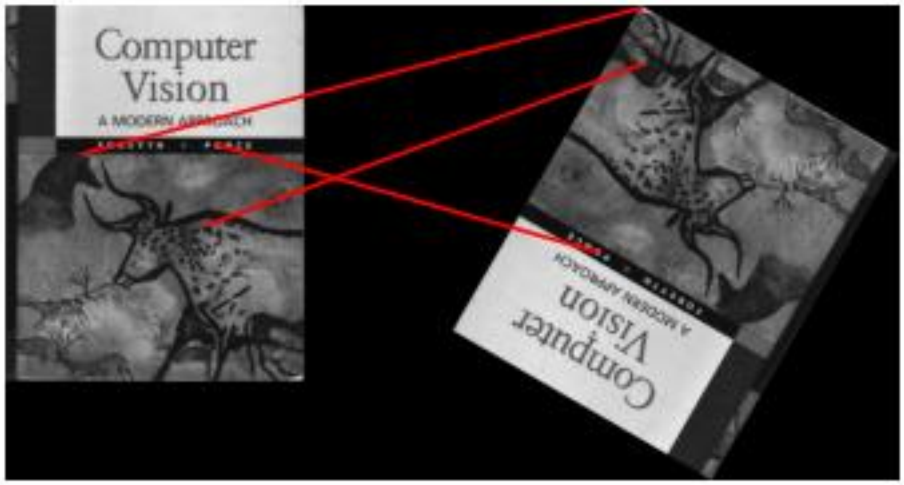
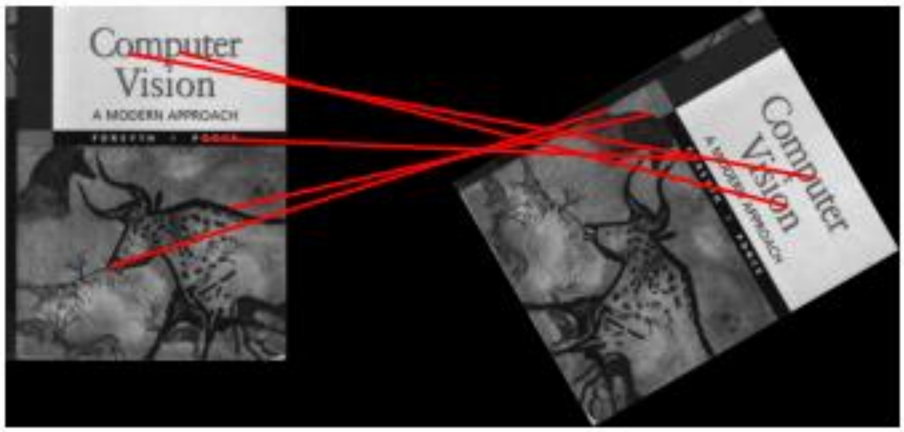
0.12	0.8		Slightly lower matches than previous picture as threshold is tightened a bit by increasing
0.13	0.7		Here threshold and ratio are both tightened, seeing less matches, but better accuracy.
0.13	0.8		Loosening ratio increases the matches, although reducing accuracy.
0.20	0.10		Too big of a sigma and very relaxed ratio yields no particular matches

Q2.1.6 BRIEF and Rotations

Histogram: Each rotation iteration corresponds to 10 degree rotation in the image

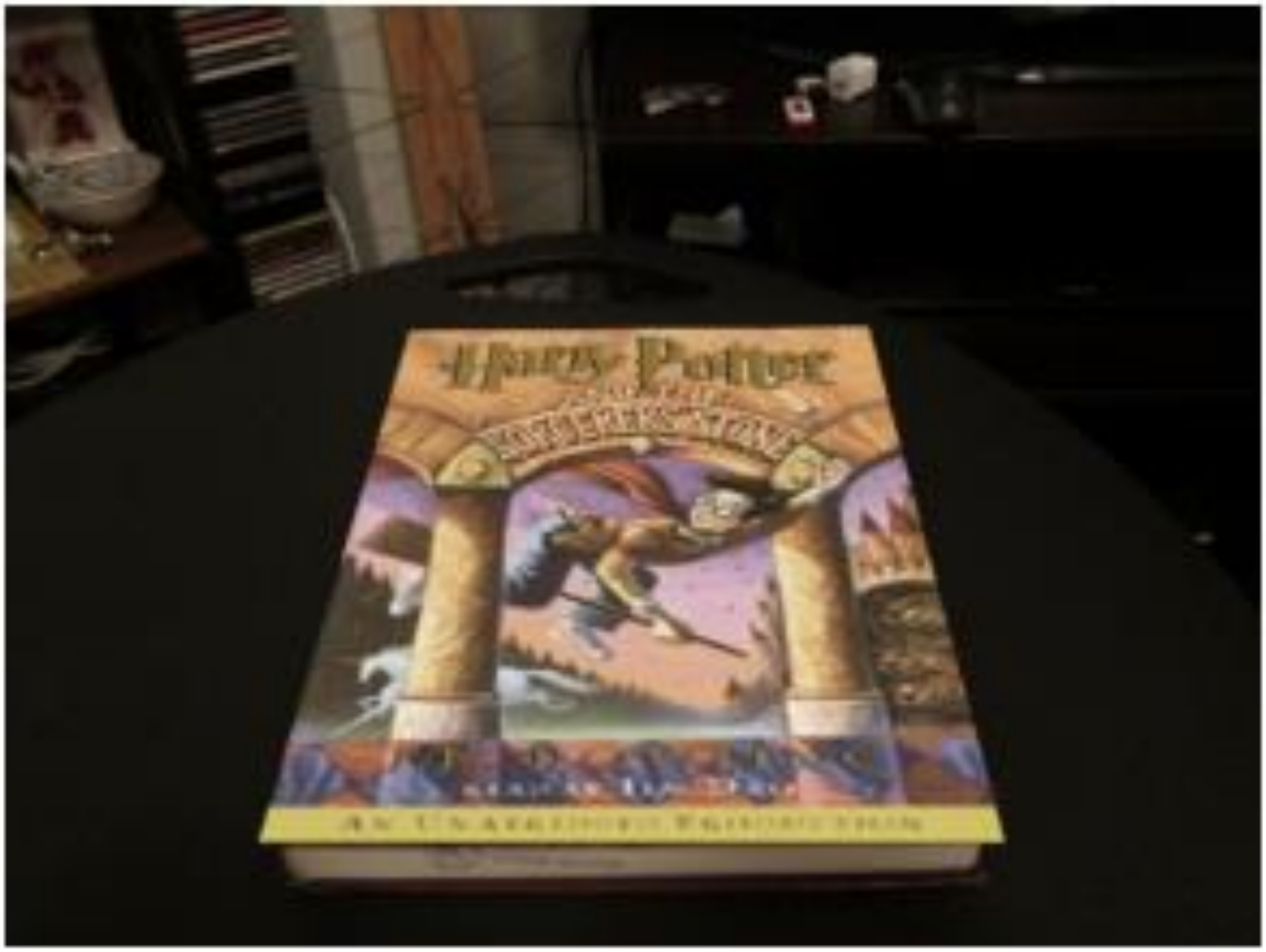


30°	

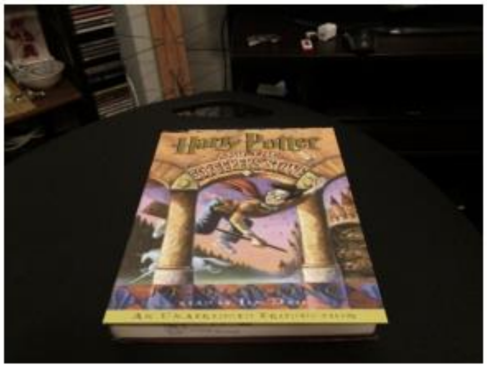
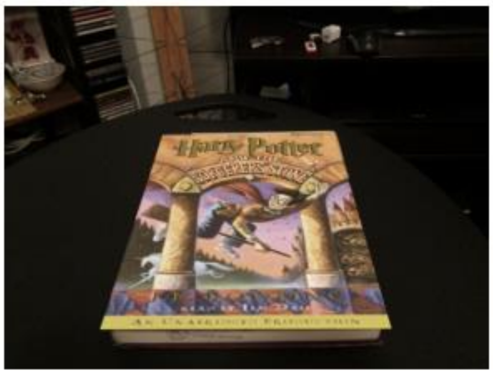
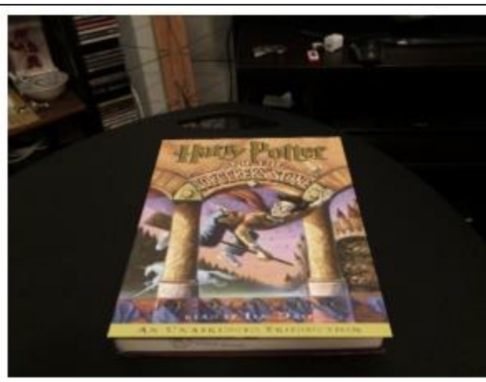
150°	
300°	

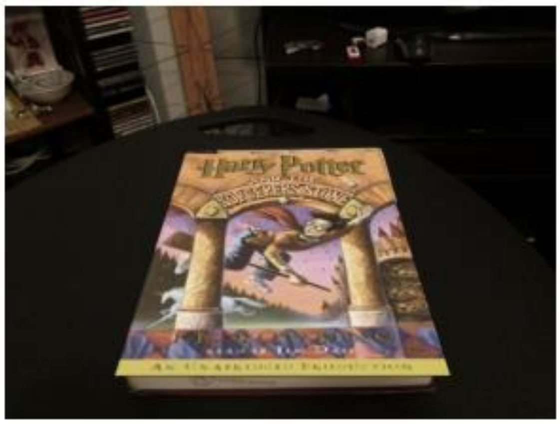
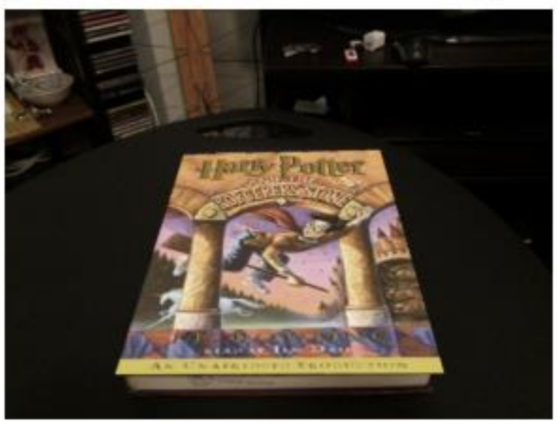
The BRIEF detector algorithm does not account for rotation invariance. As the image is rotated, the patches that the algorithm considers to describe them as bit vectors are not considered to be rotated, therefore describing each patch with a different set of bits, thus making the hamming distance quite, large between the corresponding patches in the original and the rotated image. This makes the number of points that are matched to be significantly lower.

Q2.2.4 Putting it together



Q2.2.5 RANSAC Parameter Tuning

max iters	inlier tol		Comments
500 (default)	2 (default)		Fit well with default parameters
500	1		Not much of difference was observed when tol was reduced by 1.
1000	1		Here increasing the max iters while keeping the tol 1 will not yield a significant difference as the homography computed matches well.

1000	0.01		Reducing tol too much with only 1000 iterations yielded a bad homography. Reason being that not enough iterations provided to yield the best set of points to compute the homography. Some background points were probably considered to be used.
10000	0.01		Increasing the iterations by 10 times than the previous study yielded in a decently sufficient matching.
10000	2		Increasing the iterations results in best possibility for a good match with an appropriate tolerance. Decreasing the tolerance too much will not yield in a great match, however increasing it will yield in an average match since the four points for homography are chosen at random, therefore it could pass the tolerance test in short iterations.

Q3.1: Incorporating video

Video is saved in result

Q4.2x: Create a Simple Panorama

Two Individual Images