1.1

From the lectures, we know that,

$p_i = M_{3x4} * P_i$

Here M is P_1 and P_i is x_π

So, $x_1 = \lambda_1 * P_1 * x_\pi$ (Writing equal to instead of equivalent as the equation is written for a scale factor of λ_1)

$P_1^{-1} * x_1/\lambda_1 = x_\pi$ ...1

Similarly, $x_2 = \lambda_2 * P_2 * x_\pi$ ...2

From eq. 1 and 2

$$x_2 = \frac{(\lambda_2 * P_2 * P_1^{-1}) * x_1}{\lambda_1}$$

$$x_1 = \frac{\lambda_1}{\lambda_2} * P_1 * P_2^{-1} * x_2$$

Let H = $P_1 * P_2^{-1}$ and since λ_1 and λ_2 are scalars, the term $\frac{\lambda_1}{\lambda_2}$ is just a scale factor

Therefore, $x_1 \equiv H * x_2$

Hence Proved.

1.2

1. The column vector h (reshaped from H), contains 9 elements, but it has 8 degrees of freedom, as 1 element corresponds to the scale, which is not considered while working in homogeneous coordinates.

2. Since there are 8 unknowns, at least 4 point pairs are required to solve h. Because 4 point pairs will give 8 linear equations.

3.

From the Homography equation, we have

$$\begin{pmatrix} x_1^i \\ y_1^i \\ 1 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} * \begin{pmatrix} x_2^i \\ y_2^i \\ 1 \end{pmatrix}$$

$$x_1^i = \frac{h_{11} * x_2^i + h_{12} * y_2^i + h_{13}}{h_{31} * x_2^i + h_{32} * y_2^i + h_{33}} \quad \ldots . \ 1$$

$$y_1^i = \frac{h_{21} * x_2^i + h_{22} * y_2^i + h_{23}}{h_{31} * x_2^i + h_{32} * y_2^i + h_{33}} \quad \ldots . \ 2$$

$$h_{11} * x_2^i + h_{12} * y_2^i + h_{13} - h_{31} * x_2^i * x_1^i - h_{32} * y_2^i * x_1^i - h_{33} * x_1^i = 0 \quad \ldots . \ 3$$

$$h_{21} * x_2^i + h_{22} * y_2^i + h_{23} - h_{31} * x_2^i * y_1^i - h_{32} * y_2^i * y_1^i - h_{33} * y_1^i = 0 \quad \ldots . \ 4$$

Writing 3 and 4 in a Matrix form

$$\begin{pmatrix} x_2^i & y_2^i & 1 & 0 & 0 & 0 & -x_2^i * x_1^i & -y_2^i * x_1^i & -x_1^i \\ 0 & 0 & 0 & x_2^i & y_2^i & 1 & -x_2^i * y_1^i & -y_2^i * y_1^i & -y_1^i \end{pmatrix} * \begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{pmatrix} = 0_{8x1}$$

Since we need at least 4 pair points to solve for 8 unknowns in h vector, i varies from 1 to 4

The equation is of the form, $A_i * h = 0$

$$\text{Therefore, } A_i = \begin{pmatrix} x_2^i & y_2^i & 1 & 0 & 0 & 0 & -x_2^i * x_1^i & -y_2^i * x_1^i & -x_1^i \\ 0 & 0 & 0 & x_2^i & y_2^i & 1 & -x_2^i * y_1^i & -y_2^i * y_1^i & -y_1^i \end{pmatrix}$$

4. The trivial solution of h would be that h is a null vector.

Since no 2 rows or columns can be represented as a linear combination of one another, A should be a full-rank matrix.

1.3

$$x_1 = K_1 * [I \quad 0] * X$$

$$\Rightarrow x_1 = (K_1 \quad 0) * X$$

$$\Rightarrow X = (K_1 \quad 0)^{-1} * x_1 \qquad \ldots . 1$$

$$x_2 = (K_2 \quad 0) * \begin{pmatrix} R & 0 \\ 0 & 1 \end{pmatrix} * (K_1 \quad 0)^{-1} * x_1$$

$$\Rightarrow x_1 = K_1 * R^{-1} * K_2^{-1} * x_2$$

$$\text{So, } \ x_1 = H * x_2 \ \text{ where } H = K_1 * R^{-1} * K_2^{-1}$$

1.4

Since, the camera is not changing,

$$H = K * R^{-1} * K^{-1}$$

$$\Rightarrow H^2 = K * R^{-1} * R^{-1} * K$$

Note that K is a constant matrix, so H depends on $(R^{-1})^2$

If R indicates rotation by theta, then R^{-1} indicates rotation by theta in the opposite direction

Let $R^{-1} = r$

Multiplying r by r indicates 1 rotation by theta, then 1 more rotation by theta, hence effective rotation of 2Θ

Therefore, H^2 corresponds to a rotation of 2Θ

1.5

In planar homography, the scene is assumed to be planar. We assume that all the points in the real world lie on a plane, which is not really the case. Therefore, it’s not sufficient to map any arbitrary scene image to another viewpoint.

1.6

Consider 3 colinear points in 3D: X, Y, Z.

To constrain them on a single line, let's define a linear relationship among these 3 points:

$$Z = X - \Omega(X + Y)$$

The corresponding projections:

$$x = PX$$

$$y = PY$$

$$z = PZ$$

Now consider the projection of Z using the linear relation we defined above

$$z = PZ = PX - \Omega(PX + PY)$$

$$\Rightarrow z = x - \Omega(x + y)$$

Therefore, it is proved that the projection P preserves lines

2.1.1

Both the FAST detector and Harris corner detector are used to identify corners in an image. In the FAST (Features from Accelerated Segment Test) algorithm, a 16-pixel circle (a set of 16 pixels forming a circle) is selected around a pixel point p with intensity, Ip and a threshold t is ascertained. The detection happens by comparing the intensities of a subset of the selected 16 pixels to (Ip ± t). Whereas, the Harris corner detector uses a sliding window to move over a region. Since FAST employs a static technique, it's much more computationally efficient than the latter.

2.1.2

Filter banks just provide a global representation of an image. This means doesn't say much about spatial information.

The BRIEF descriptor provides an efficient way to identify binary features by comparing hamming distance by doing XOR bitwise operations.

The GIST filter bank can be used as a global descriptor for holistic representation of the scene.

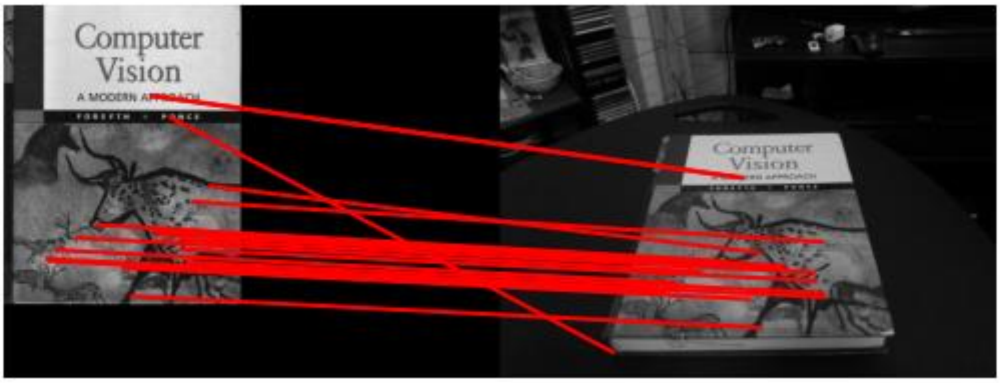
2.1.3

In BRIEF Descriptor, binary strings are used to match features. Hamming distance is used to find the match, which is nothing but a XOR operation performed on the 2 binary strings.

The nearest neighbors are defined as the points with minimum Euclidean distance from the given descriptor point. To assess the likelihood of a match being accurate, one can calculate it by comparing the distance from the closest neighbor to that of the second closest neighbor.

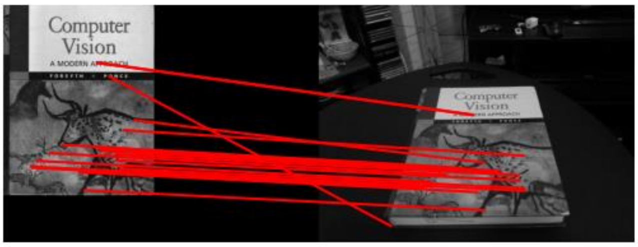
As compared to Euclidean distance, hamming distance is much faster to compute as bitwise operations on binary strings are faster than arithmetic calculations.

2.1.4

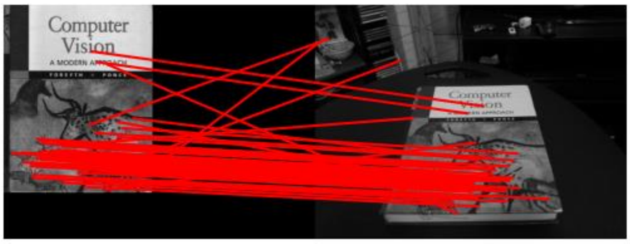


2.1.5

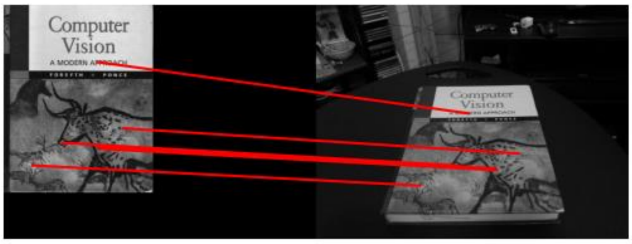
Sigma = 0.15, Ratio = 0.7



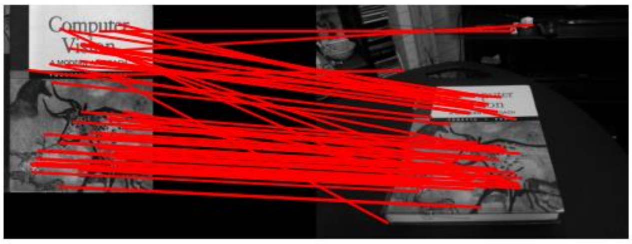
Sigma = 0.1, Ratio = 0.7



Sigma = 0.15, Ratio = 0.6



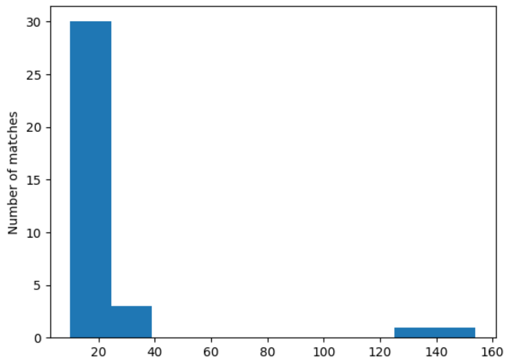
Sigma = 0.15, Ratio = 0.8



Sigma = 0.19, Ratio = 0.8

From this ablation study, it's clear that increase in ratio alone increases the no. of matches between the 2 images. With a bit lower ratio (in the range of 0.6, 0.7), lower sigma produces greater no. of matches.

2.1.6



From the histogram, it’s clear that the BRIEF descriptor is **not rotation invariant.** As the image rotates, it’s not able to match the same points.

2.2.1

See Code

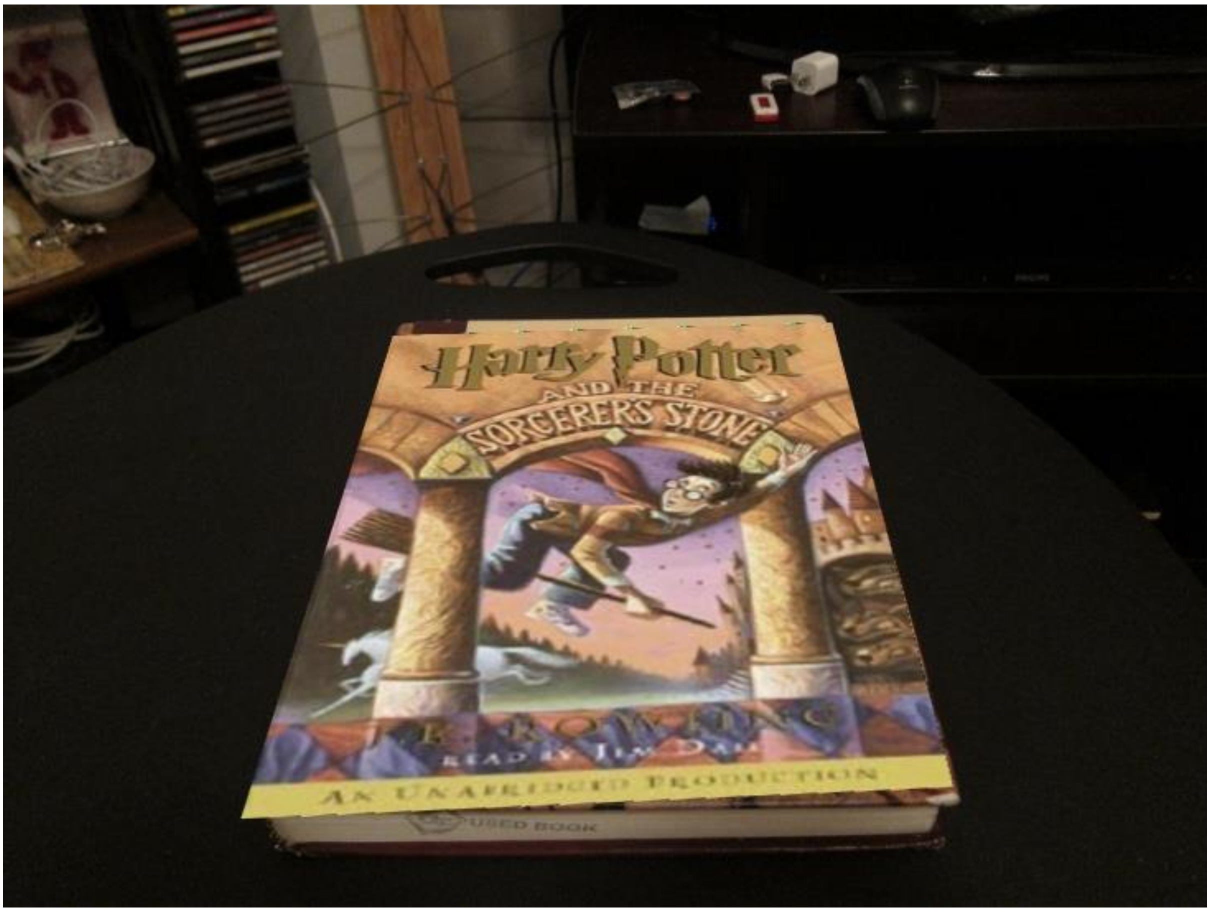
2.2.2

See Code

2.2.3

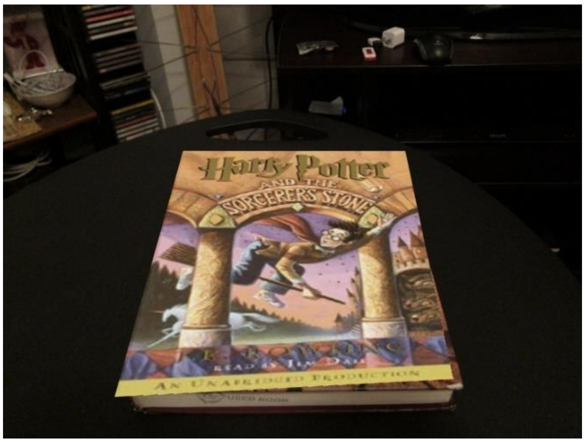
See code

2.2.4

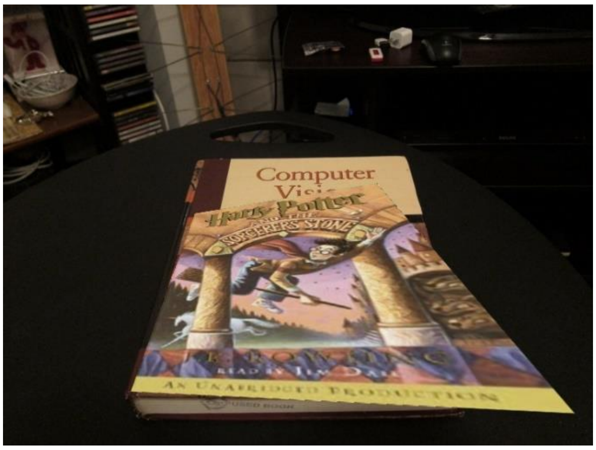


2.2.5

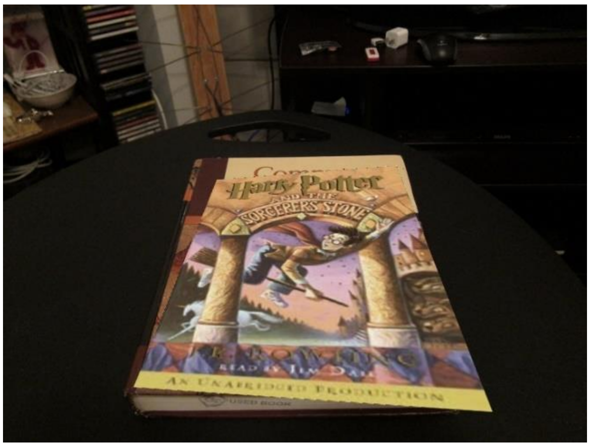
Iterations = 500, Inlier_tol = 3, Inliers = 7



Iterations = 50, Inlier_tol = 0.6, Inliers = 3



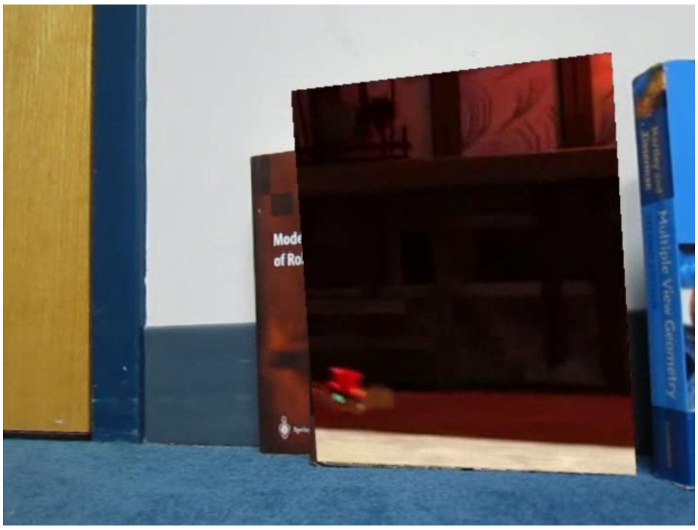
Iterations = 5000, Inlier_tol = 0.8, Inliers = 7



Tighter tolerance means, a lesser number of points are available to calculate the homography, hence even a greater number of iterations may not help to get the desired result. Loose tolerance or higher value for inlier_tol means a sufficient number of points are available to work with, hence a fewer iterations do the job.

3.1

Right



Center

Left