

Homework 2

Thomas Zhang; AndrewID: tczhang

October 2023

Q1.1

Let a point x_0 in 3d space be represented by the homologous coordinates $\begin{bmatrix} x_{0,1} \\ x_{0,2} \\ x_{0,3} \\ 1 \end{bmatrix}$. Then given two 3×4 camera projection matrices \mathbf{P}_1 and \mathbf{P}_2 .

$$x_1 = \mathbf{P}_1 x_0 = \begin{bmatrix} \mathbf{P}_{11}^T x_0 \\ \mathbf{P}_{12}^T x_0 \\ \mathbf{P}_{13}^T x_0 \end{bmatrix} \quad x_2 = \mathbf{P}_2 x_0 = \begin{bmatrix} \mathbf{P}_{21}^T x_0 \\ \mathbf{P}_{22}^T x_0 \\ \mathbf{P}_{23}^T x_0 \end{bmatrix}$$
$$\mathbf{x}_1 \equiv \mathbf{H} \mathbf{x}_2$$
$$\mathbf{P}_1 x_0 \equiv \mathbf{H} \mathbf{P}_2 x_0$$

Let λ be some scaling factor.

$$\mathbf{P}_2^{-1} \mathbf{P}_1 x_0 - \lambda \mathbf{H} x_0 = 0$$

Thus \mathbf{H} exists and can be found using:

$$\mathbf{P}_2^{-1} \mathbf{P}_1 = \mathbf{H}$$

Q1.2

1. h has 8 degrees of freedom. H has 9 variables, however one of the degrees of freedom is lost due to the scale factor.
2. 4 point pairs are required to solve h. (Each point pair can find two variables).

3.

$$\begin{bmatrix} x_1^i \\ y_1^i \\ 1 \end{bmatrix} = \begin{bmatrix} h_{1,1} & h_{1,2} & h_{1,3} \\ h_{2,1} & h_{2,2} & h_{2,3} \\ h_{3,1} & h_{3,2} & h_{3,3} \end{bmatrix} \begin{bmatrix} x_2^i \\ y_2^i \\ 1 \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} x_1^i \\ y_1^i \\ 1 \end{bmatrix} = \begin{bmatrix} x_2^i h_{1,1} + y_2^i h_{1,2} + h_{1,3} \\ x_2^i h_{2,1} + y_2^i h_{2,2} + h_{2,3} \\ x_2^i h_{3,1} + y_2^i h_{3,2} + h_{3,3} \end{bmatrix}$$

Convert the coordinates back to non-homogeneous coordinates.

$$\begin{bmatrix} x_1^i \\ y_1^i \end{bmatrix} = \begin{bmatrix} x_2^i h_{1,1} + y_2^i h_{1,2} + h_{1,3} \\ x_2^i h_{3,1} + y_2^i h_{3,2} + h_{3,3} \\ x_2^i h_{2,1} + y_2^i h_{2,2} + h_{2,3} \\ x_2^i h_{3,1} + y_2^i h_{3,2} + h_{3,3} \end{bmatrix}$$

Multiply both sides by denominator.

$$\begin{bmatrix} x_1^i x_2^i h_{3,1} + x_1^i y_2^i h_{3,2} + x_1^i h_{3,3} \\ y_1^i x_2^i h_{3,1} + y_1^i y_2^i h_{3,2} + y_1^i h_{3,3} \end{bmatrix} = \begin{bmatrix} x_2^i h_{1,1} + y_2^i h_{1,2} + h_{1,3} \\ x_2^i h_{2,1} + y_2^i h_{2,2} + h_{2,3} \end{bmatrix}$$

$$\begin{bmatrix} x_2^i h_{1,1} + y_2^i h_{1,2} + h_{1,3} - x_1^i x_2^i h_{3,1} - x_1^i y_2^i h_{3,2} - x_1^i h_{3,3} \\ x_2^i h_{2,1} + y_2^i h_{2,2} + h_{2,3} - y_1^i x_2^i h_{3,1} - y_1^i y_2^i h_{3,2} - y_1^i h_{3,3} \end{bmatrix} = 0$$

$$\begin{bmatrix} x_2^i h_{1,1} + y_2^i h_{1,2} + h_{1,3} + 0 + 0 - x_1^i x_2^i h_{3,1} - x_1^i y_2^i h_{3,2} - x_1^i h_{3,3} \\ 0 + 0 + 0 + x_2^i h_{2,1} + y_2^i h_{2,2} + h_{2,3} - y_1^i x_2^i h_{3,1} - y_1^i y_2^i h_{3,2} - y_1^i h_{3,3} \end{bmatrix} = 0$$

$$\begin{bmatrix} x_2^i & y_2^i & 1 & 0 & 0 & 0 & -x_1^i x_2^i & -x_1^i y_2^i & -x_1^i \\ 0 & 0 & 0 & x_2^i & y_2^i & 1 & -y_1^i x_2^i & -y_1^i y_2^i & -y_1^i \end{bmatrix} \begin{bmatrix} h_{1,1} \\ h_{1,2} \\ h_{1,3} \\ h_{2,1} \\ h_{2,2} \\ h_{2,3} \\ h_{3,1} \\ h_{3,2} \\ h_{3,3} \end{bmatrix} = 0$$

$$\mathbf{A}_i = \begin{bmatrix} x_2^i & y_2^i & 1 & 0 & 0 & 0 & -x_1^i x_2^i & -x_1^i y_2^i & -x_1^i \\ 0 & 0 & 0 & x_2^i & y_2^i & 1 & -y_1^i x_2^i & -y_1^i y_2^i & -y_1^i \end{bmatrix}$$

4. The trivial solution of h is the 0 vector (vector of 0's). The matrix A is full rank because all the rows and columns are linearly independent (no column/row is a linear combination of the other columns/rows). The dimensions of matrix is equal to its null space + its rank. So if A is full rank then and if 0 is not an eigenvalue of A, then the only solution to $Ah = 0$ is the trivial solution.

Q1.3

$$x_1 = \mathbf{K}_1 [\mathbf{I} \ 0] \mathbf{X}$$

$$x_2 = \mathbf{K}_2 [\mathbf{R} \ 0] \mathbf{X}$$

$$\mathbf{K}_1 [\mathbf{I} \ 0] \mathbf{X} = \mathbf{H} \mathbf{K}_2 [\mathbf{R} \ 0] \mathbf{X}$$

$$\mathbf{K}_1 [\mathbf{I} \ 0] \mathbf{X} = \mathbf{H} \mathbf{K}_2 [\mathbf{R} \ 0] \mathbf{X}$$

$$\mathbf{K}_1 \mathbf{I} \mathbf{X} = \mathbf{H} \mathbf{K}_2 \mathbf{R} \mathbf{X}$$

$$\mathbf{K}_1 \mathbf{X} = \mathbf{H} \mathbf{K}_2 \mathbf{R} \mathbf{X}$$

$$\mathbf{K}_1 = \mathbf{H} \mathbf{K}_2 \mathbf{R}$$

$$\mathbf{K}_1 \mathbf{R}^{-1} \mathbf{K}_2^{-1} = \mathbf{H}$$

This means that there exists an \mathbf{H} such that $\mathbf{x}_1 \equiv \mathbf{H}\mathbf{x}$

Q1.4

$$\mathbf{K}_1 \begin{bmatrix} \mathbf{I} & 0 \end{bmatrix} \mathbf{X} = \mathbf{H} \mathbf{K}_2 \begin{bmatrix} \mathbf{R} & 0 \end{bmatrix} \mathbf{X}$$

Dropping the subscript for K_1 and K_2 since they are equal.

$$\mathbf{K}^{-1} \mathbf{K} \begin{bmatrix} \mathbf{I} & 0 \end{bmatrix} \mathbf{X} = \mathbf{K}^{-1} \mathbf{H} \mathbf{K} \begin{bmatrix} \mathbf{R} & 0 \end{bmatrix} \mathbf{X}$$

$$\begin{bmatrix} \mathbf{I} & 0 \end{bmatrix} \mathbf{X} = \mathbf{K}^{-1} \mathbf{H} \mathbf{K} \begin{bmatrix} \mathbf{R} & 0 \end{bmatrix} \mathbf{X}$$

$$\mathbf{I} = \mathbf{K}^{-1} \mathbf{H} \mathbf{K} \mathbf{R}$$

$$\mathbf{R}^{-1} = \mathbf{K}^{-1} \mathbf{H} \mathbf{K}$$

$$\mathbf{K} \mathbf{R}^{-1} \mathbf{K}^{-1} = \mathbf{H}$$

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

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

$$\mathbf{K} \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{K}^{-1} = \mathbf{H}^2$$

$$\mathbf{K} \begin{bmatrix} \cos^2(\theta) - \sin^2(\theta) & 2\cos(\theta)\sin(\theta) & 0 \\ -2\cos(\theta)\sin(\theta) & -\sin^2(\theta) + \cos^2(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{K}^{-1} = \mathbf{H}^2$$

$$\mathbf{K} \begin{bmatrix} \cos(2\theta) & \sin(2\theta) & 0 \\ -\sin(2\theta) & \cos(2\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{K}^{-1} = \mathbf{H}^2$$

Thus \mathbf{H}^2 applies a rotation of 2θ

Q1.5

Planar homography assumes that the scene of interest is on a single plane, but most real world areas are not on a single plane, resulting in distortions when using planar homography. Additionally, planar homography is not sufficient if the change in position of the camera between the two images is very large.

Q1.6

A line in 3d can be described with a parametric equation $\mathbf{P}(t) = \vec{u}_0 + t\vec{d}$ where u_0 is some point on the line and \vec{d} is the direction vector of the line.

Let \mathbf{M} be the perspective projection matrix for projecting 3D to 2D.

The projection of the line $P(t)$ to 2-dimensions is then $\mathbf{MP}(t) = \mathbf{M}(\mu_0 + t\vec{d}) = \mathbf{M}\vec{u}_0 + t\mathbf{M}\vec{d}$. This is the parametric form of a line in 2D where $\mathbf{M}\vec{u}_0$ is the projection of \vec{u}_0 onto the 2-D plane and $\mathbf{M}\vec{d}$ is the direction vector projected to the 2-D plane.

Thus a perspective projection from 3-D to 2-D preserves lines.

Q2.1

1. The FAST detector chooses a circle of 16 pixels around a pixel of interest. Then a pixel is considered an edge if at least n contiguous pixels on the circle are greater than $\mathbf{I} + t$ or less than $\mathbf{I} - t$ where I is the intensity of the pixel of interest, t is some constant we chose and n is some threshold that is less than 16. The Harris corner requires sliding a window over each pixel to calculate derivatives in a inertia matrix: $\mathbf{M} = \Sigma \begin{bmatrix} \mathbf{I}_x^2 & \mathbf{I}_x \mathbf{I}_y \\ \mathbf{I}_x \mathbf{I}_y & \mathbf{I}_y^2 \end{bmatrix}$. Then using the determinant and trace of \mathbf{M} to calculate the score of the pixel R and if the score of the pixel exceeds some threshold then it is considered an edge. The fast detector is much less computationally expensive than the Harris detector because the Harris detector requires a gaussian window over the image to generate \mathbf{M} , the FAST detector simply looks at 16 pixels at a fixed distance from the pixel of interest.

2. The BRIEF detector assigns a feature vector to each pixel using a binary vector representation of the neighborhood of that pixel. The neighborhood of a pixel is the square around the pixel with some defined height and width and is referred to as a "patch". Before converting the patch to a binary feature vector the patches are smoothed using a gaussian filter. The advantage of the Brief detector is that it is very quick to generate a BRIEF descriptor for an image since generating the binary values requires a small number of comparisons. Additionally comparing two different images becomes comparing two binary feature vectors.

In contrast the other filter banks used in class are a collection of multiple filters that each get applied to the image via convolution and the output for each filter is used as a descriptor for the image. This is much slower than the BRIEF detector method due to the amount of convolution required. Lastly comparing two images using the filter banks is slower because the feature vectors are not binary and require more sophisticated distance calculation metrics. Any of the filter banks could be used as a descriptor for an image.

3. The Hamming distance between two vectors is simply the number of differing elements between the two vectors. (e.g. $[1 \ 0 \ 1]$ and $[0 \ 0 \ 0]$ have a hamming distance of 2 since the first and last elements of the two vectors are different.)

Nearest neighbor finds the nearest neighbor of a point via some distance metric. This is normally the euclidean distance between two points. Using the hamming distance is advantageous compared to conventional euclidean distance in calculation time. Conventional euclidean distance between two points follows the formula:

$$\sqrt{\sum_i^n ((v_i - u_i)^2)}$$

Where v and u are vectors with n elements. This is much more computationally complex than the Hamming distance which is simply counting the number of different elements between \vec{u} and \vec{v} .

4. Image generated from MatchPics function.

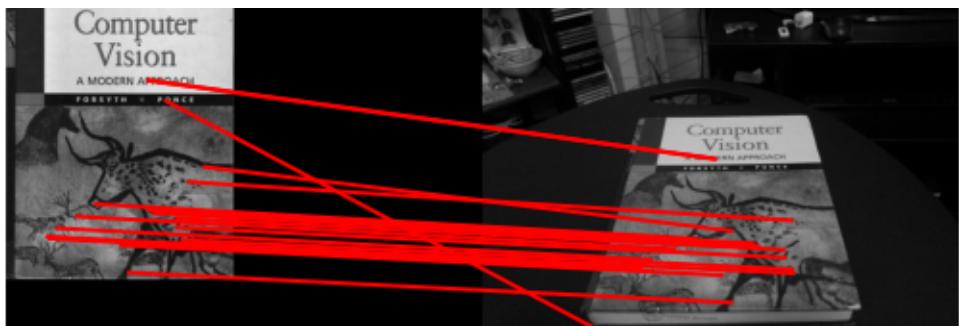
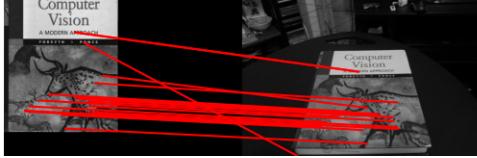
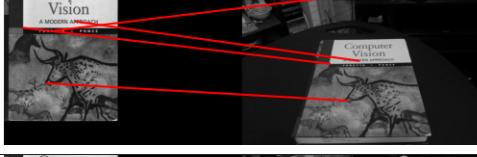
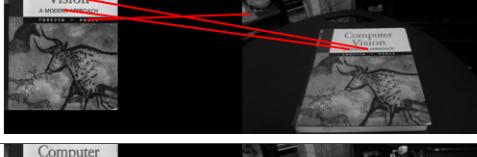
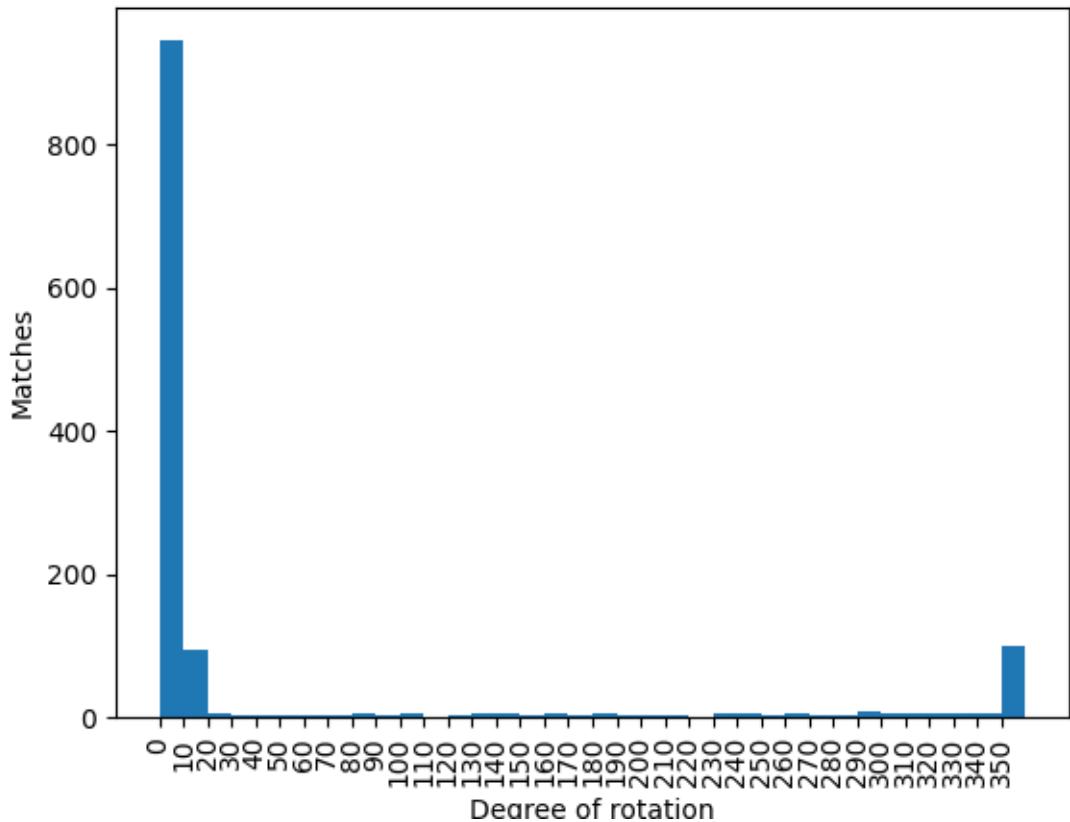


Figure 1: $\sigma = 0.15$, ratio = 0.7

5. Based on the ablation table study, number of matched points decreases with increasing sigma and increases with increasing ratio.

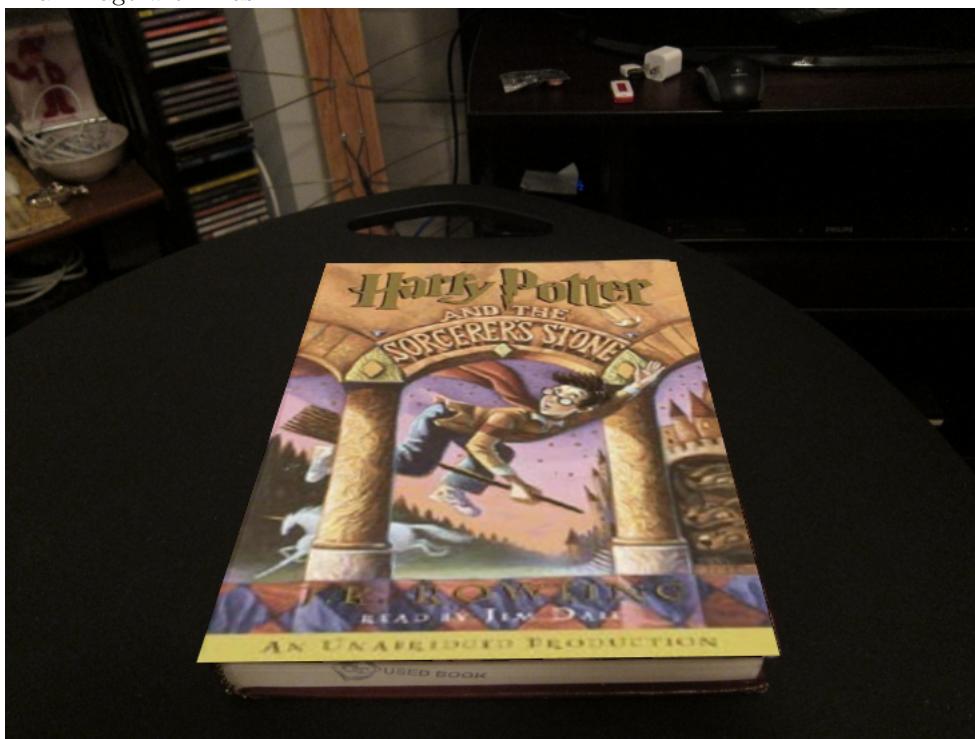
sigma	ratio	image
0.15	0.7	
0.34	0.7	
0.50	0.7	
0.15	0.5	
0.15	0.9	



The number of matches is much higher when the angle of rotation is close to 0 (or 360) degrees. This might be because image description generated by BRIEF is sensitive to angles and thus when there is a large amount of rotation, BRIEF fails to identify matches.

Q2.2.4

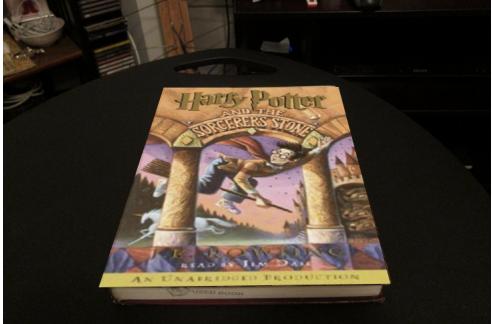
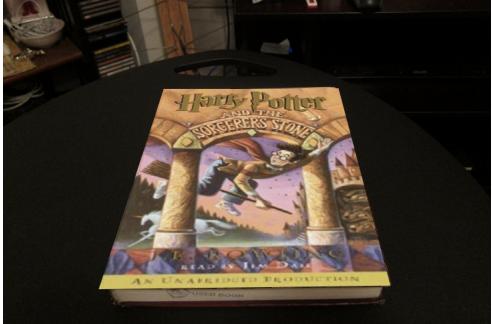
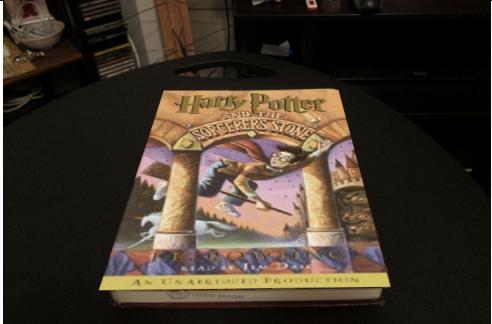
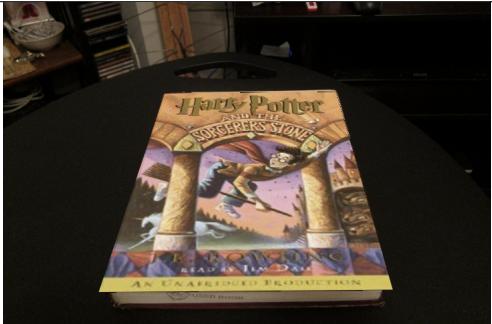
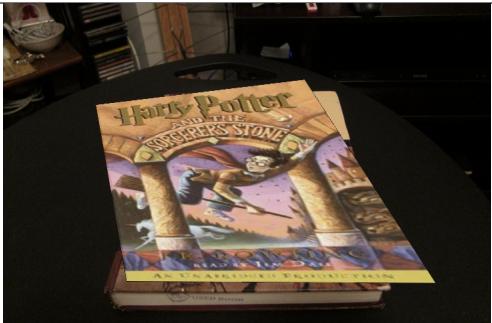
- 4) The image isn't filling up the same space as the book because the initial size of the cv_cover and hp_cover images are different. Thus the homography isn't mapping the points on the cover to the points of the book on the desk correctly. To fix this we can resize hp_cover to match the size of cv.cover.
- 6) Final Image with Mask



Q2.2.5

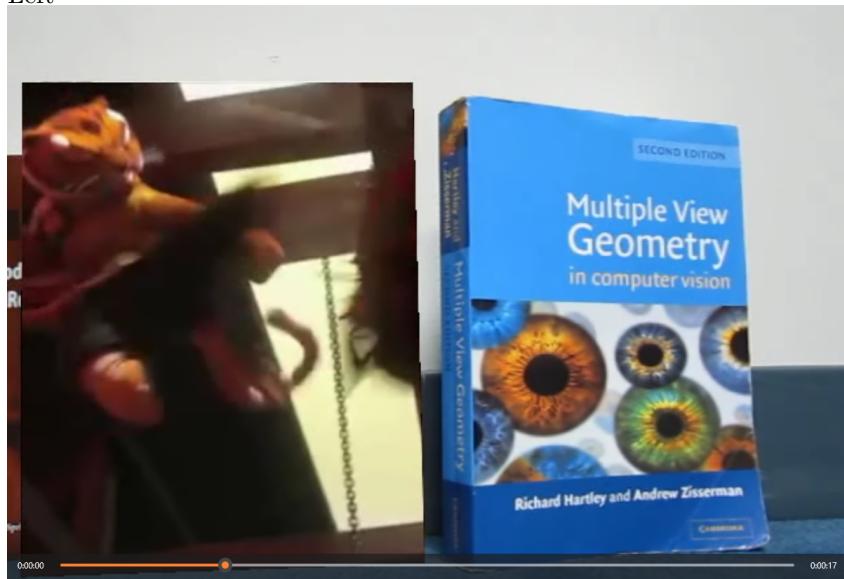
Changing the maximum iterations does not appear to have a large effect on the mapping. This is most likely because the set of 4 points that best map the matched points is likely found relatively early during the RANSAC process.

Increasing the tolerance too much causes the alignment to become bad because a set of points that performs a poor mapping becomes able to score high when counting inliers.

Maximum Iterations	Inlier Tolerance	image
50	2	
500	2	
1000	2	
500	0.5	
500	30	

Q3.1

Left



Center



Right

