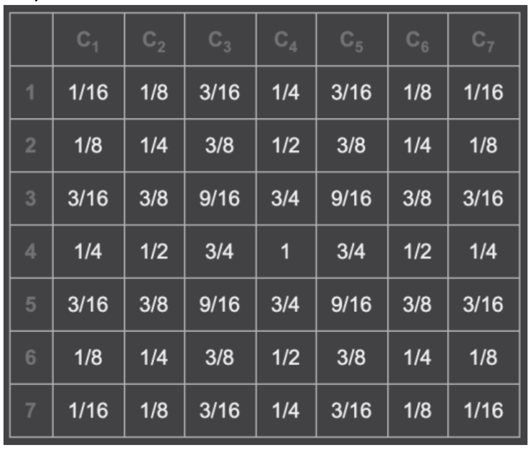
Fan Vincent Ka Chun 1002563380

Question 1

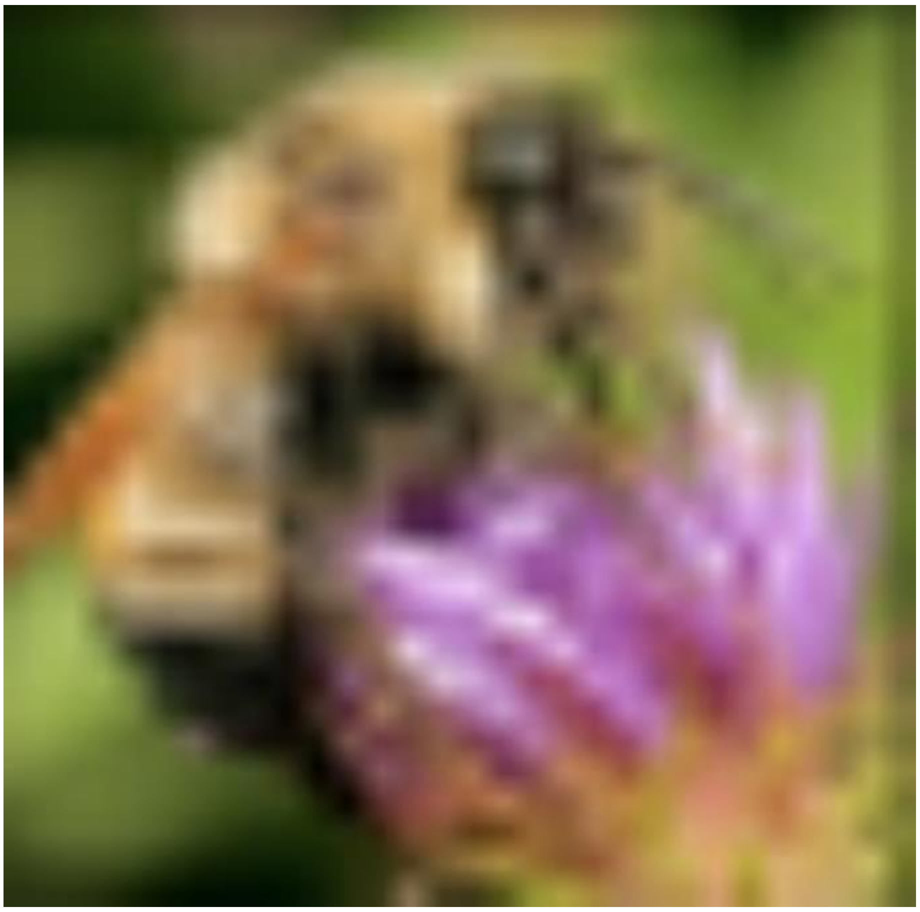
(a) We can certainly perform the same operation with a 2D filter. In our case of enlarging it 4 times, the 2D filter would be as follow:



	C1	C2	C3	C4	C5	C6	C7
1	1/16	1/8	3/16	1/4	3/16	1/8	1/16
2	1/8	1/4	3/8	1/2	3/8	1/4	1/8
3	3/16	3/8	9/16	3/4	9/16	3/8	3/16
4	1/4	1/2	3/4	1	3/4	1/2	1/4
5	3/16	3/8	9/16	3/4	9/16	3/8	3/16
6	1/8	1/4	3/8	1/2	3/8	1/4	1/8
7	1/16	1/8	3/16	1/4	3/16	1/8	1/16

We can observe that the above is the product with a column matrix and row matrix of the filter from the lecture which was extended for quadrupling the image size. Note that I did not include the 2 rows and 2 columns of zeroes since it does not contribute to the final result.

The result after perform convolution with the 1D filter (0, 1/4, 2/4, 3/4, 1, 3/4, 2/4, 1/4, 0) on both x and y direction is as follow:



We can see that the result is very successful.

(b) Following similar pattern as above, we can generalize the formula of the filter to be the product of column matrix and row matrix with the format of:

$$(0, 1/n, 2/n, \ldots, n/n, \ldots, 2/n, 1/n, 0)$$

This works because we are constructing a separable filter which we already know that each separated filter can indeed up-sample our image in one direction. Thus, the 2D separable can just perform what we did for (a) in one go.

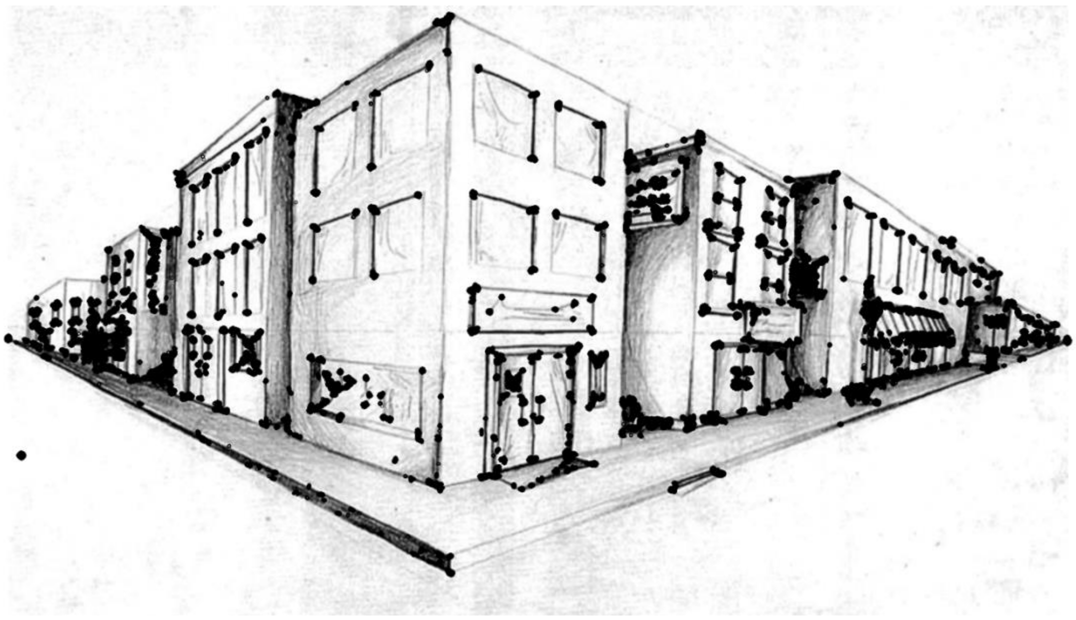
The results for using this 2D filter is shown below:

We can observe that the final result is the same as above which proved that our generalized 2D filter works the same as perform 1D up sampling twice.

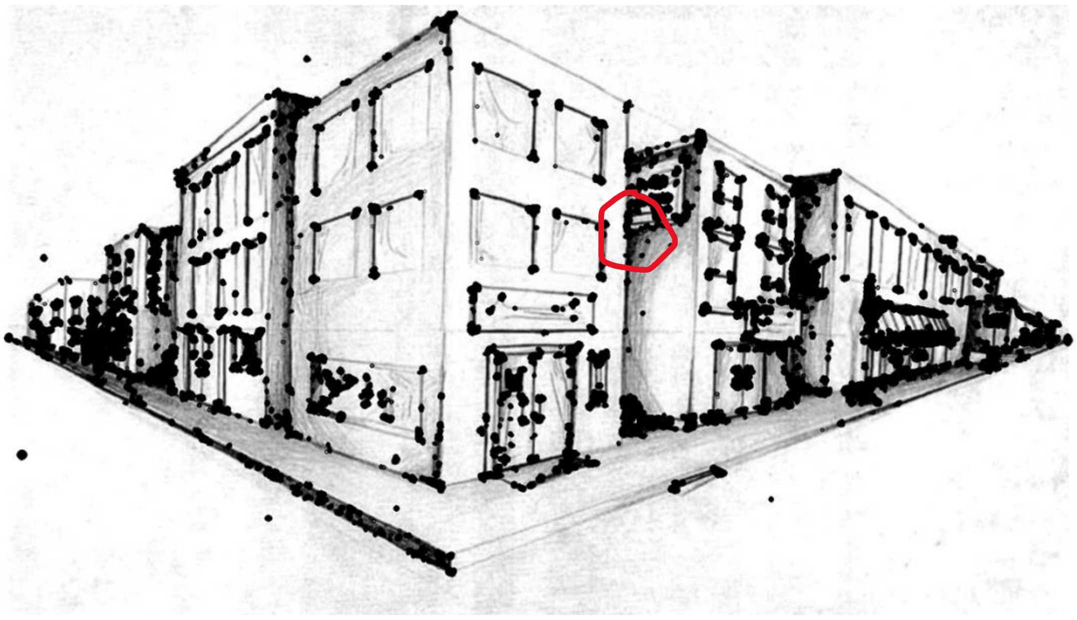
Fan Vincent Ka Chun 1002563380

Question 2

(a) Harris Detector result:



Brown Detector results:

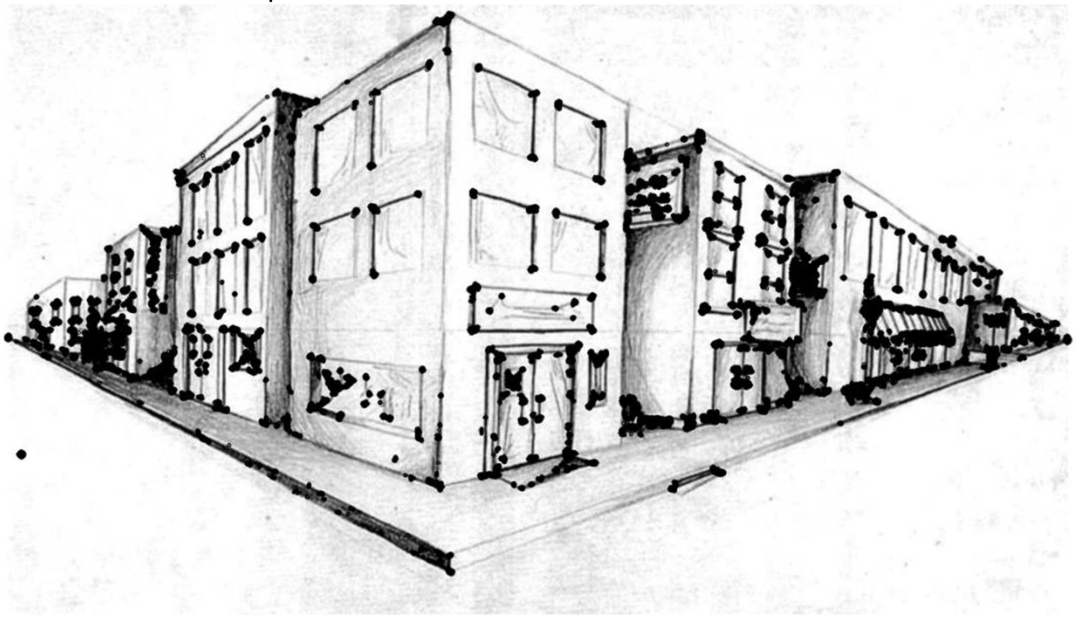


We can observe that there are more points detected with Brown metrics because Brown's method takes the harmonic mean which has a smoother change than computing difference of determinant and trace, thus it would allow more points to be considered as corners. It is especially smooth in region where λ_1 is close to λ_2. For example, region circled in red, we can see a transition in that corner.

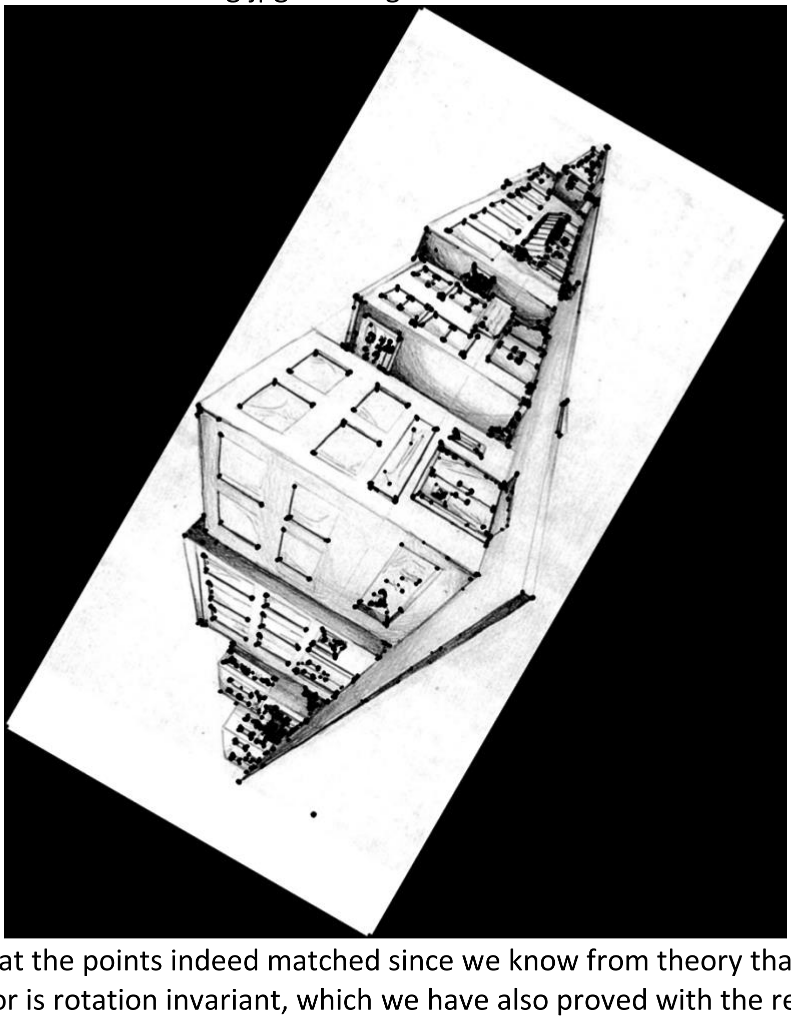
Note that in my implementation of both methods, I only took the points where it is close to the max of the values. The range was determined through trials and the quality of the output, which I then drew small black circles to show where the corners are at.

Fan Vincent Ka Chun 1002563380

(b) Harris Detector result from part a:

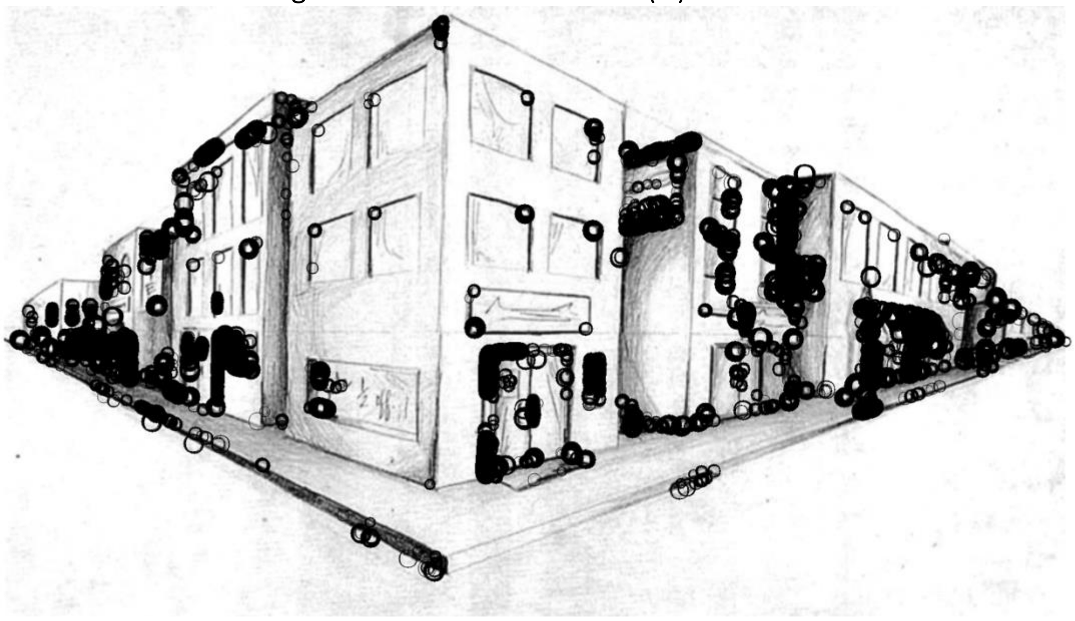


Harris Detector result of building.jpg rotating 60^o:



We can see that the points indeed matched since we know from theory that Harris corner detector is rotation invariant, which we have also proved with the results.

(c) Here is the result of calculating LOG with different scales (σ):



I drew circles of different size with corresponding scale.
In my implementation, I choose σ between 3 and 12 to guarantee to include as many interest points as possible. I also choose to pick up the points with LoG values lower than 60% of the minimum of the entire array. In the image above, we can see circles with different sizes, which indicates different scale level. Note that most interest points are cluttered near the sides of the images in most scales.

(d) I picked GLOH which is a modified SIFT, here are the steps of the algorithm:

1. Our scale invariant interest point detector gives scale p for each keypoint
2. For each keypoint, we take the Gaussian blurred image at corresponding scale p
3. Compute the gradient magnitude and orientation in neighborhood of each keypoint proportional to the detected scale
4. Compute dominant orientation of each keypoint via a histogram of gradient orientations, each bin covers 10^o. Then take the peak of histogram.
5. Compute a 272-dimensional descriptor computed from 17 locations, each location has a histogram of 16 orientation bins relative to the dominant orientation.
6. Perform PCA on the 272-dimensional descriptor to reduce the vector size to 128 which would be the same size of a SIFT descriptor.

Fan Vincent Ka Chun 1002563380

Question 3

(a) We know the gaussian filter is given to us as:

$$g(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

So, after taking the second partial derivatives, we have:

$$\frac{\partial^2}{\partial x^2} g(x, y, \sigma) = \frac{1}{2\pi\sigma^2}\left(\frac{x^2 - \sigma^2}{\sigma^4}\right) e^{-\frac{x^2+y^2}{2\sigma^2}}$$

$$\frac{\partial^2}{\partial y^2} g(x, y, \sigma) = \frac{1}{2\pi\sigma^2}\left(\frac{y^2 - \sigma^2}{\sigma^4}\right) e^{-\frac{x^2+y^2}{2\sigma^2}}$$

So, the Laplacian would be:

$$\nabla^2 g(x, y, \sigma) = \frac{1}{2\pi\sigma^2}\left(\frac{x^2 - \sigma^2}{\sigma^4} + \frac{y^2 - \sigma^2}{\sigma^4}\right) e^{-\frac{x^2+y^2}{2\sigma^2}}$$

$$= \frac{1}{2\pi\sigma^2}\left(\frac{x^2 + y^2}{\sigma^4} - \frac{2\sigma^2}{\sigma^4}\right) e^{-\frac{x^2+y^2}{2\sigma^2}}$$

$$= -\frac{1}{2\pi\sigma^2}\left(\frac{2}{\sigma^2} - \frac{x^2 + y^2}{\sigma^4}\right) e^{-\frac{x^2+y^2}{2\sigma^2}}$$

$$= -\frac{1}{2\pi\sigma^2}\frac{2}{\sigma^2}\left(1 - \frac{x^2 + y^2}{\sigma^4}\frac{\sigma^2}{2}\right) e^{-\frac{x^2+y^2}{2\sigma^2}}$$

$$= -\frac{1}{\pi\sigma^4}\left(1 - \frac{x^2 + y^2}{2\sigma^2}\right) e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Which is exactly what we expected from the lecture slides!

We can also observe that LoG is indeed a separable filter since it is just a sum of second partial derivatives of the Gaussian filter. Since we know that a 2D Gaussian filter is separable, the second derivative of 2D Gaussian filter must also be separable. Therefore, we can conclude that the LoG is separable.

(b) I will now show how LoG can be approximated with DoG:

$$DoG = g(x,y,\sigma_1) - g(x,y,\sigma_2)$$
$$= \frac{1}{2\pi\sigma_1^2}e^{-\frac{x^2+y^2}{2\sigma_1^2}} - \frac{1}{2\pi\sigma_2^2}e^{-\frac{x^2+y^2}{2\sigma_2^2}}$$
$$= \frac{1}{2\pi\sigma_1^2\sigma_2^2}\left(\sigma_2^2 e^{-\frac{x^2+y^2}{2\sigma_1^2}} - \sigma_1^2 e^{-\frac{x^2+y^2}{2\sigma_2^2}}\right)$$
$$= \frac{1}{2\pi\sigma_1^2\sigma_2^2}\left(\sigma_2^2 - \sigma_1^2 e^{-\frac{x^2+y^2}{2\sigma_2^2}} e^{\frac{x^2+y^2}{2\sigma_1^2}}\right)e^{-\frac{x^2+y^2}{2\sigma_1^2}}$$
$$= \frac{1}{2\pi\sigma_1^2\sigma_2^2}\left(\sigma_2^2 - \sigma_1^2 e^{\frac{(x^2+y^2)(\sigma_2^2-\sigma_1^2)}{2\sigma_1^2\sigma_2^2}}\right)e^{-\frac{x^2+y^2}{2\sigma_1^2}}$$

Using the Taylor expansion of exponential function and taking the first 2 terms:

$$\cong \frac{1}{2\pi\sigma_1^2\sigma_2^2}\left(\sigma_2^2 - \sigma_1^2\left(1 + \frac{(x^2+y^2)(\sigma_2^2-\sigma_1^2)}{2\sigma_1^2\sigma_2^2}\right)\right)e^{-\frac{x^2+y^2}{2\sigma_1^2}}$$
$$= \frac{1}{2\pi\sigma_1^2\sigma_2^2}(\sigma_2^2 - \sigma_1^2)\left(1 - \frac{(x^2+y^2)}{2\sigma_2^2}\right)e^{-\frac{x^2+y^2}{2\sigma_1^2}}$$

Assume $\sigma_1^2 \approx 2\sigma^2, \sigma_2^2 \approx \frac{\sigma^2}{2}$

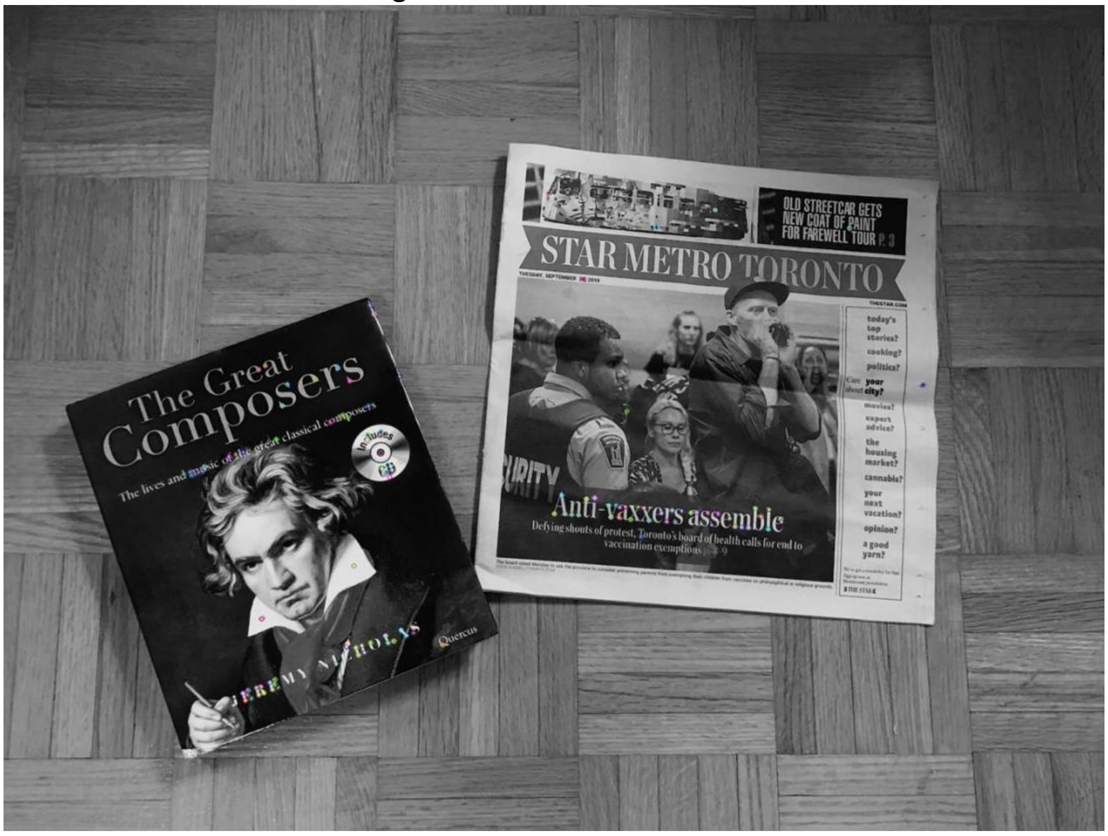
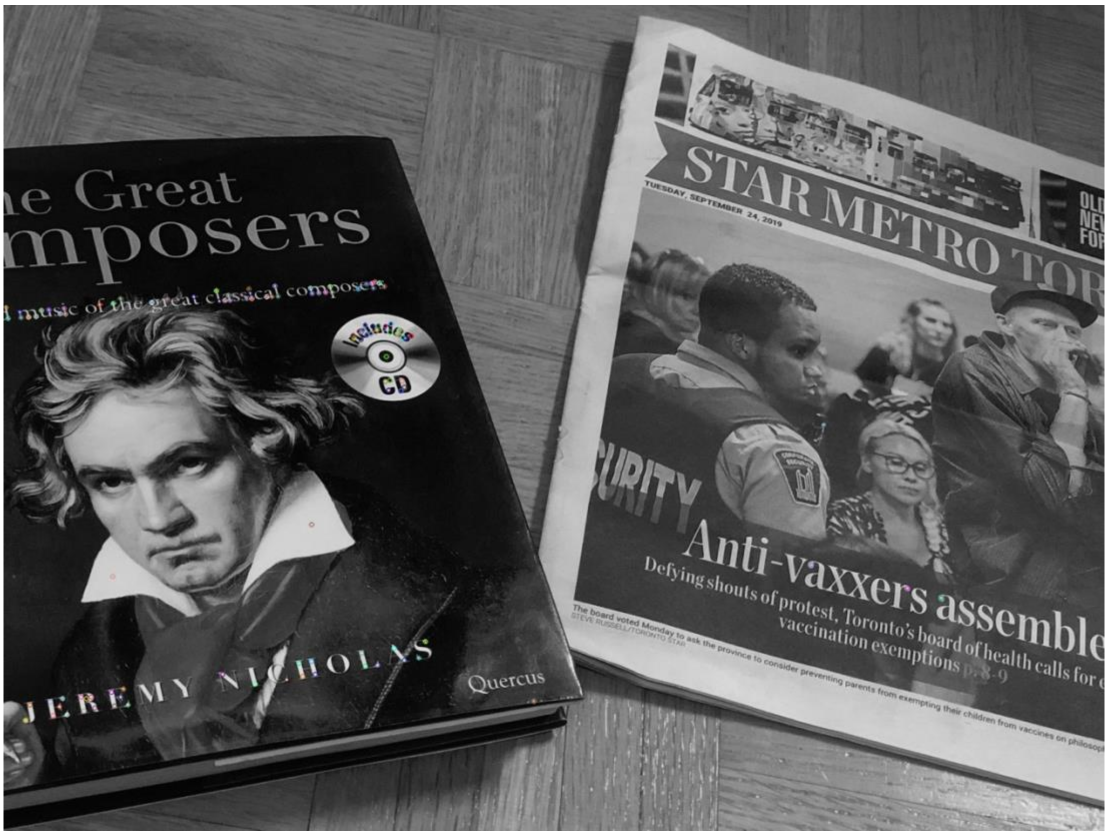
$$= \frac{1}{2\pi\sigma^4}\left(\frac{\sigma^2}{2} - 2\sigma^2\right)\left(1 - \frac{(x^2+y^2)}{2\sigma^2}\right)e^{-\frac{x^2+y^2}{4\sigma^2}}$$
$$= -\frac{1}{\pi\sigma^4}\left(\sigma^2 - \frac{\sigma^2}{4}\right)\left(1 - \frac{(x^2+y^2)}{2\sigma^2}\right)e^{-\frac{x^2+y^2}{4\sigma^2}}$$
$$\cong -\frac{1}{\pi\sigma^4}\left(1 - \frac{x^2+y^2}{2\sigma^2}\right)e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Which is approximately the Laplacian of Gaussian given σ is relatively small. Note that if the choices of $\sigma_1^2$and σ_2^2 changes from above, it can lead to a less accurate approximation of the LoG and result in inaccurate result since the difference of gaussian cannot create the Mexican hat shape. For example, if σ_1^2 is larger and σ_2^2 is smaller, than the “dip” would be less profound which is not close to a Laplacian of Gaussian graph.

Fan Vincent Ka Chun 1002563380

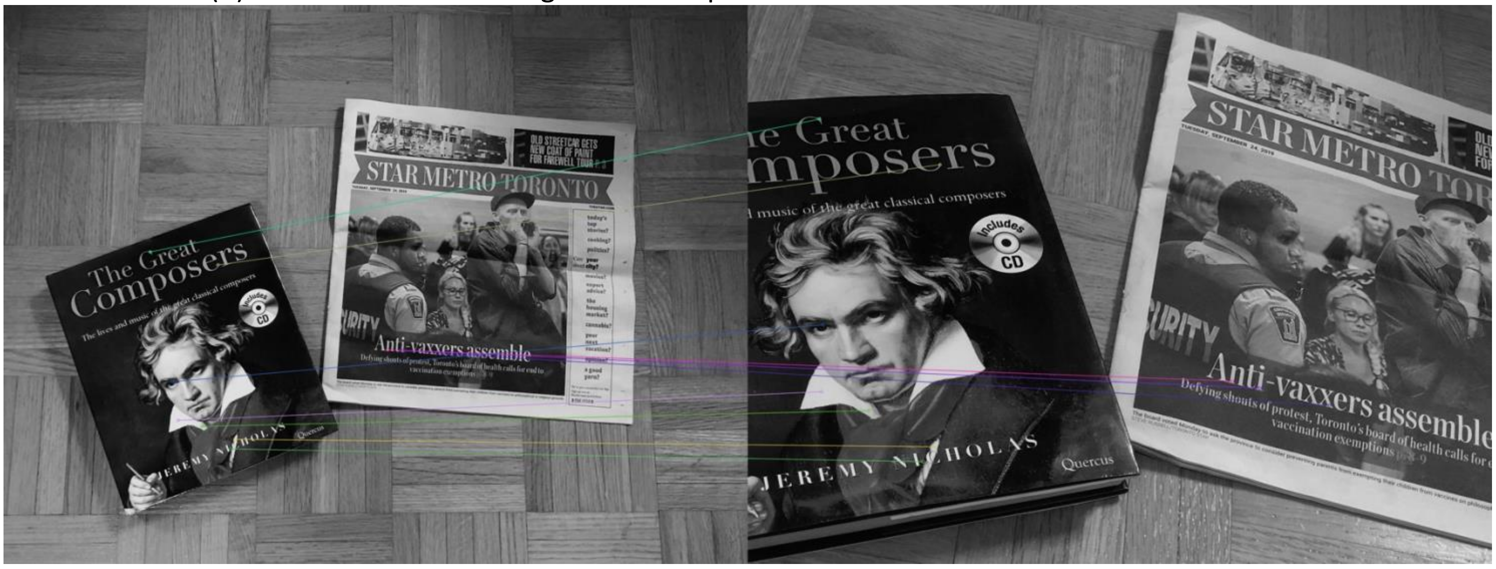
Question 4

(a) Here are the results after using cv2.sift:


The Great Composers
The lives and music of the great classical composers
Includes CD
JEREMY NICHOLAS
Quercus
OLD STREETCAR GETS NEW COAT OF PAINT FOR FAREWELL TOUR P. 3
STAR METRO TORONTO
Anti-vaxxers assemble
Defying shouts of protest, Toronto's board of health calls for end to vaccination exemptions

The Great Composers
Includes CD
JEREMY NICHOLAS
Quercus
STAR METRO TORONTO
TUESDAY, SEPTEMBER 24, 2019
Anti-vaxxers assemble
Defying shouts of protest, Toronto's board of health calls for end to vaccination exemptions

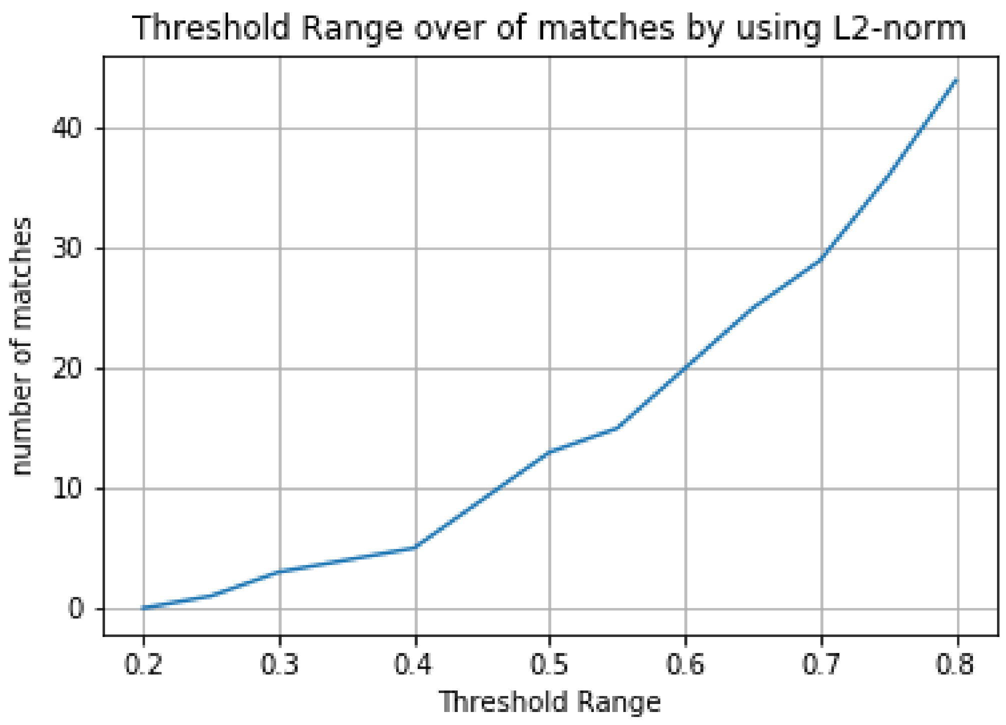
It might be hard to see the feature points here, but I only took the top 400 feature points by their scores (measured in SIFT algorithm as the local contrast) to shorten the computation time.

(b) Here is the result using the best experimental value:



The best value that I used was 0.6 as my ratio. Note that when I sort the matches according to the distance, I also removed matches that are close to each other. Otherwise it will be showing less than 10 since some of them were within couple pixels of each other.

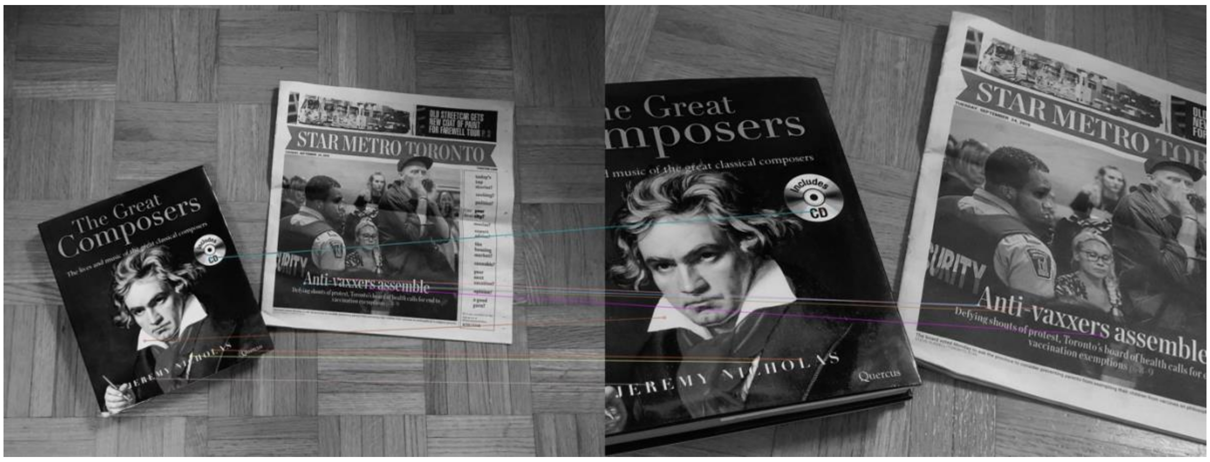
Here is the plot of number of matches versus threshold:



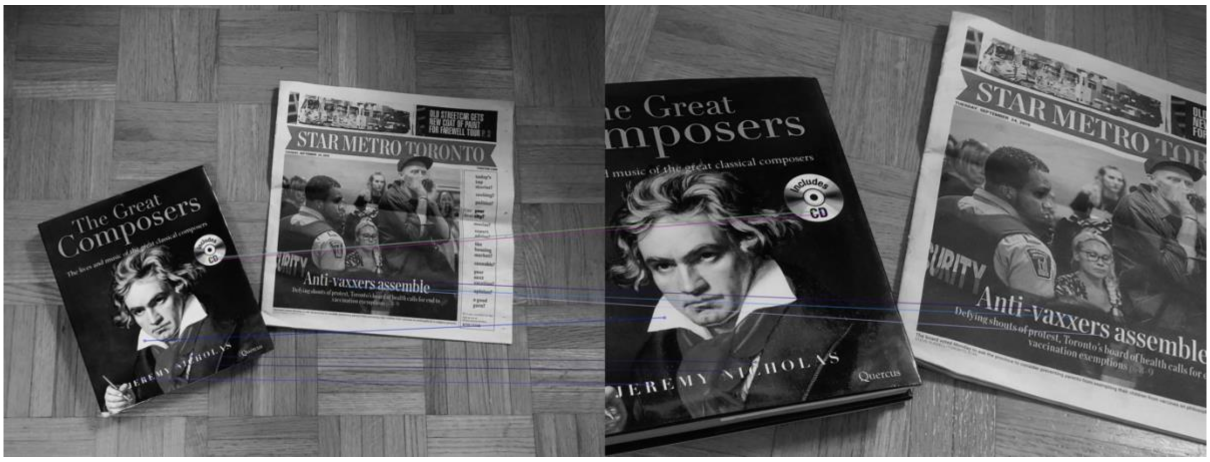
We can observe that it is closer to an exponential graph than a linear graph. We can also see that at 0.6 threshold, I am getting 20 matches which is more than enough to find the top 10 matches with including any false negatives.

Fan Vincent Ka Chun 1002563380

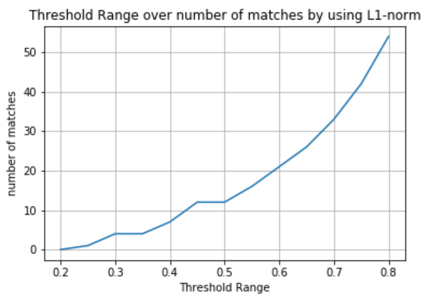
(c) L1 Norm results:

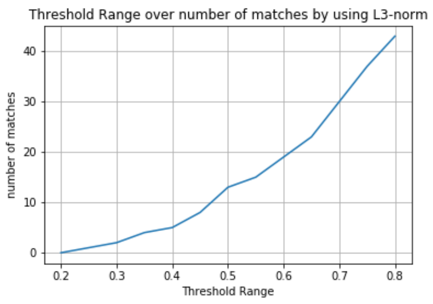


L3 Norm results:



Graphs:

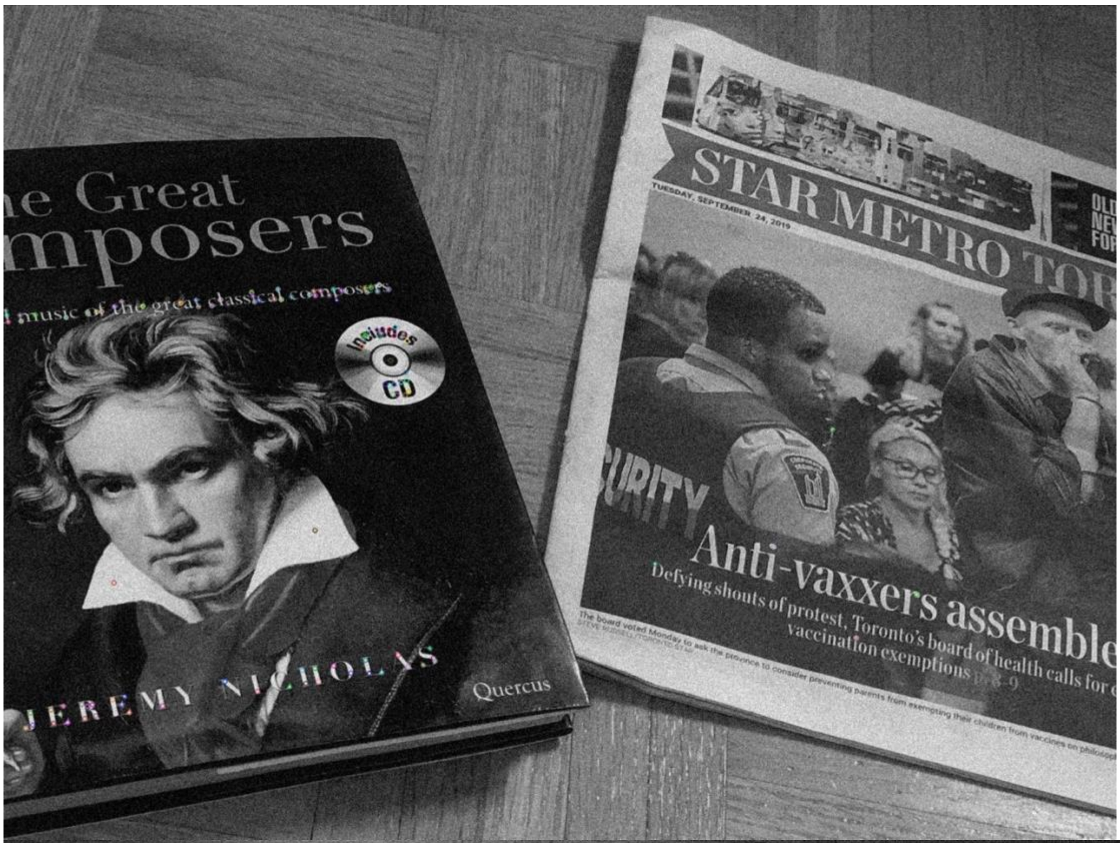


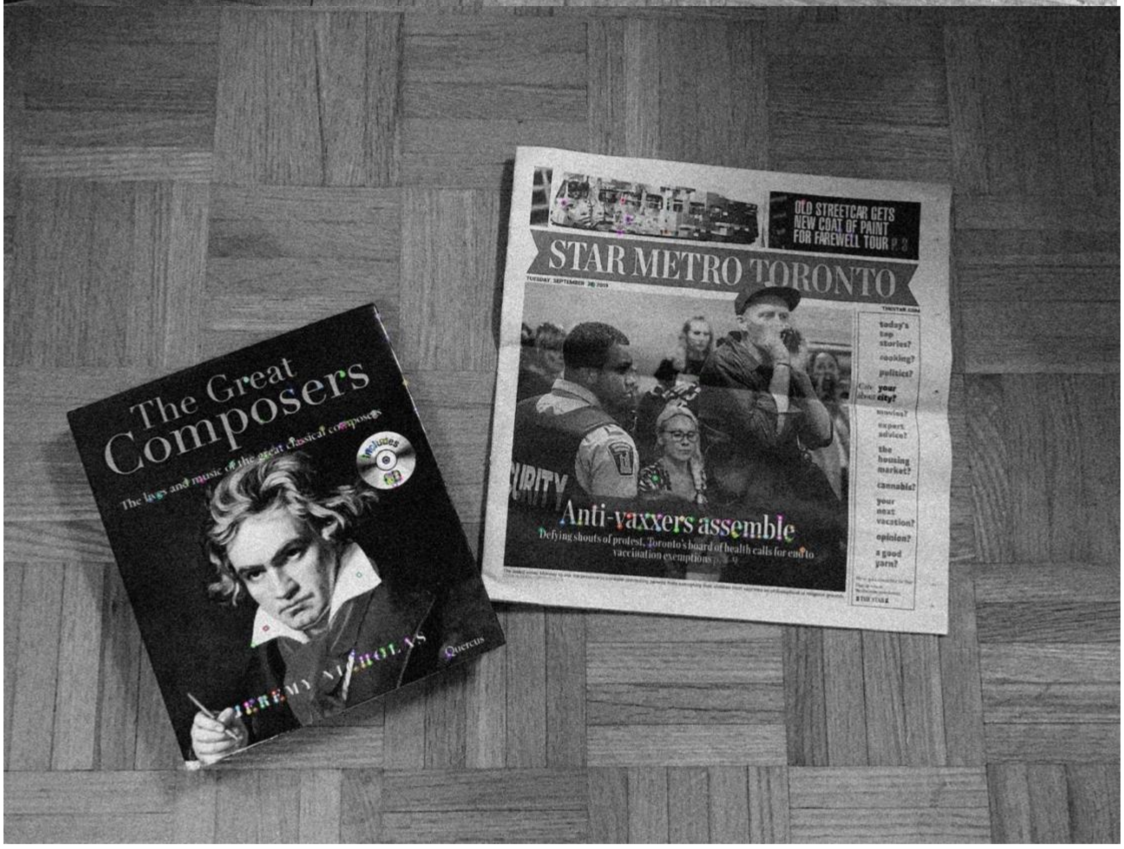


We can observe from the graphs that the number of matches with L3 is lower than L1 when the threshold is above 0.6, then when it's under 0.6 the number of matches is approximately the same.
The L3 performs better because L3 would "enlarge" the difference when taking the cube of the difference, thus would allow less false positives to be included in the matches since the cubed difference would be much larger.
In other words, L3 penalizes larger error but increasing the magnitude of the difference, however it would be insensitive to small error, which in our case would not matter as much.

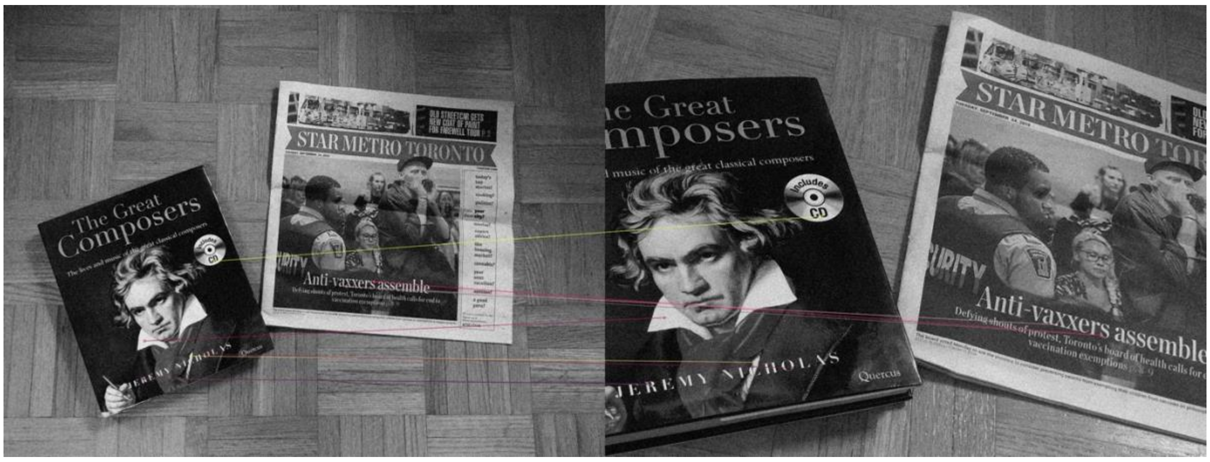
Fan Vincent Ka Chun 1002563380

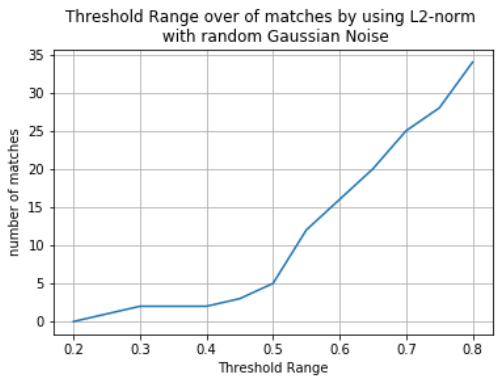
(d) Here are the results:


e Great
mposers
l music of the great classical composers
Includes CD
JEREMY NICHOLAS
Quercus
STAR METRO TOR
TUESDAY, SEPTEMBER 24, 2019
CURITY
Anti-vaxxers assemble
Defying shouts of protest, Toronto's board of health calls for e
vaccination exemptions


The Great
Composers
The lives and music of the great classical composers
Includes CD
JEREMY NICHOLAS
Quercus
OLD STREETCAR GETS NEW COAT OF PAINT FOR FAREWELL TOUR
STAR METRO TORONTO
CURITY
Anti-vaxxers assemble
Defying shouts of protest, Toronto's board of health calls for end to vaccination exemptions

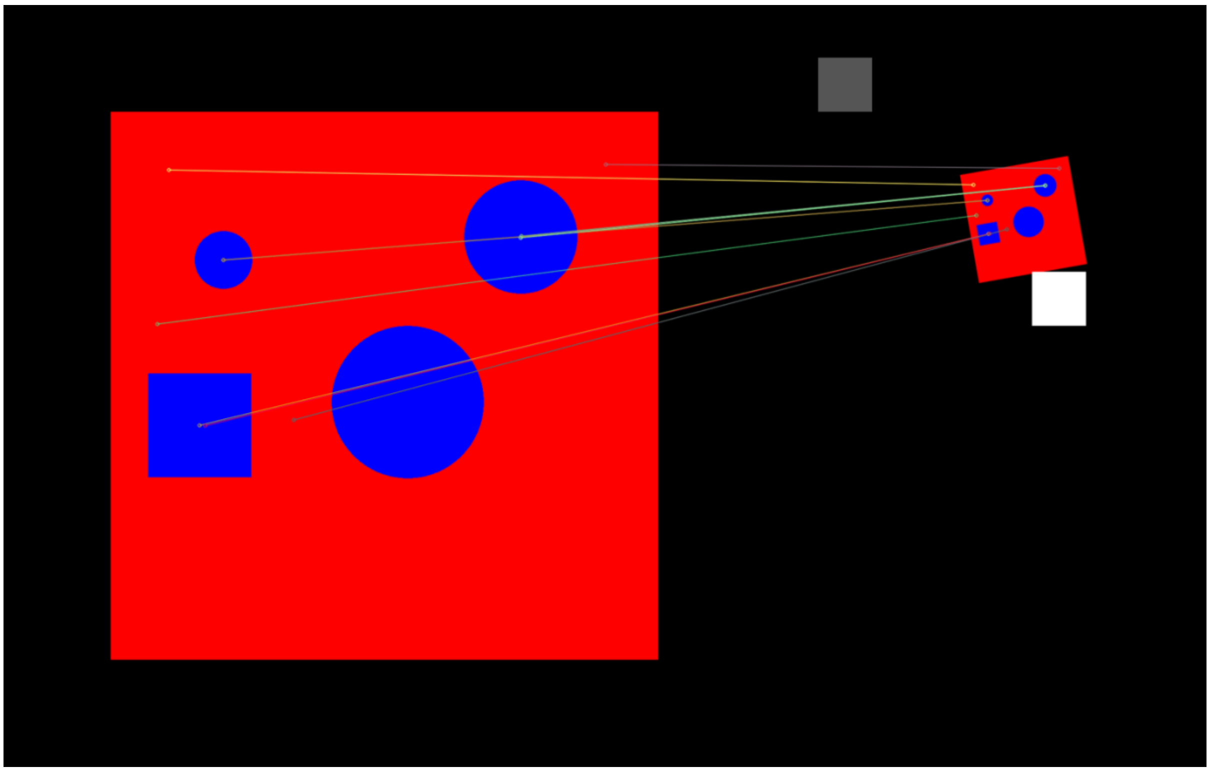
Fan Vincent Ka Chun 1002563380

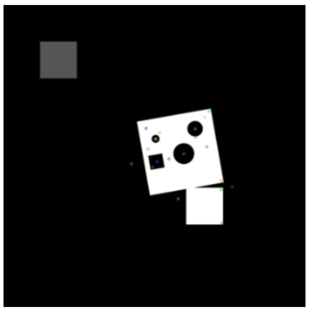



Threshold Range over of matches by using L2-norm
with random Gaussian Noise
number of matches
Threshold Range
0
5
10
15
20
25
30
35
0.2
0.3
0.4
0.5
0.6
0.7
0.8

I used the same procedures as above and I found that the number of mismatches is a lot higher, we can observe that from the graph and see that the number of matches is lower than using L-2 norm. The noise that I added would alter the pixel's intensity, which would accumulate and alter the magnitude and orientation of the descriptors. Therefore, ultimately affecting the matching process due to the slight change of descriptors and result in less good matches.

(e) My function first split up the 3 channels, then find the interest points with SIFT in each channel. Then I would go through each of the matches and take the match's coordinates and check with the original picture's RGB value. If the RGB value of the both pictures at the match's coordinates (or within a threshold), then we know the interest points also matches in color.
Here are the results:





As we Can see, the matches are very accurate, and none of the matches between the corners of the squares matched, which would happen in a single channel. For example, in red channel, the red square would match with one of the corners of white square. However, in our result, that has been eliminated since their RGB values does not match.

Note that after inspecting the green channel, cv2.sift does not return any keypoints since it is just a pitch black for colourtemplate.jpg so I choose to ignore it for my calculation, but one can easy add the interest point and my function would still be able to handle the extra keypoints.