CSC420: Assignment 2

Xiangyu Kong, kongxi16

October 9, 2019

Problem 1

See *question_1.py* for implementation.

1. The result of applying linear interpolation to up-sample the image by 4 times are shown in Figure 1.

   Figure 1a shows the first round of convolution along the first axis.

   Figure 1b shows the second round of convolution along the second axis on top of the result of the first round.

   Comparing to the original image, the up-sampled image maintains most of the information because the images look alike.

(a) First round (b) Second Round (c) Original Image

Figure 1

To achieve the same operation (up-sampling by a factor of 4), we can multiply the 1-dimensional filter by itself to get a 2-dimensional filter. This filter will be separable, and will produce the same result as convolving with the 1-dimensional filter twice on different directions.

In this specific case of up-sampling by a factor of 4, we calculate that

$$h = \begin{bmatrix} 0.25 & 0.5 & 0.75 & 1 & 0.75 & 0.5 & 0.25 \end{bmatrix}$$

$$h \times h^T = \begin{bmatrix} 0.0625 & 0.125 & 0.1875 & 0.25 & 0.1875 & 0.125 & 0.0625 \\ 0.125 & 0.25 & 0.375 & 0.5 & 0.375 & 0.25 & 0.125 \\ 0.1875 & 0.375 & 0.5625 & 0.75 & 0.5625 & 0.375 & 0.1875 \\ 0.25 & 0.5 & 0.75 & 1 & 0.75 & 0.5 & 0.25 \\ 0.1875 & 0.375 & 0.5625 & 0.75 & 0.5625 & 0.375 & 0.1875 \\ 0.125 & 0.25 & 0.375 & 0.5 & 0.375 & 0.25 & 0.125 \\ 0.0625 & 0.125 & 0.1875 & 0.25 & 0.1875 & 0.125 & 0.0625 \end{bmatrix}$$

2. The generalized two dimensional linear interpolation reconstruction filter is

$$h = \begin{bmatrix} \frac{1}{d} & \dots & \frac{d-1}{d} & 1 & \frac{d-1}{d} & \dots & \frac{1}{d} \end{bmatrix}$$

$$h \times h^T = \begin{bmatrix} \frac{1}{d}^2 & \dots & \frac{d-1}{d} \times \frac{1}{d} & 1 \times \frac{1}{d} & \frac{d-1}{d} \times \frac{1}{d} & \dots & \frac{1}{d}^2 \\ \frac{d-1}{d} \times \frac{1}{d} & \ddots & \dots & \dots & \dots & \dots & \vdots \\ \vdots & \dots & \dots & 1 & \dots & \dots & \vdots \\ \frac{d-1}{d} \times \frac{1}{d} & \dots & \dots & \dots & \dots & \dots & \frac{d-1}{d}^2 \end{bmatrix}$$

Since this is a separable filter (composed of $h \times h^T$), it is equivalent of applying the 1D linear interpolation filter twice in both direction.

The results of the 2D linear interpolation filter convolution is shown in Figure 2a. Comparing to the result of applying 1D linear interpolation filter twice in two directions (Figure 2b), the results do not have any difference.

(a) 2D Filter (b) 1D Filter (c) Original Image

Figure 2

Problem 2

See *question_2.py* for implementations.

1. The outputs of the corner detection functions are listed in Figure **??**.

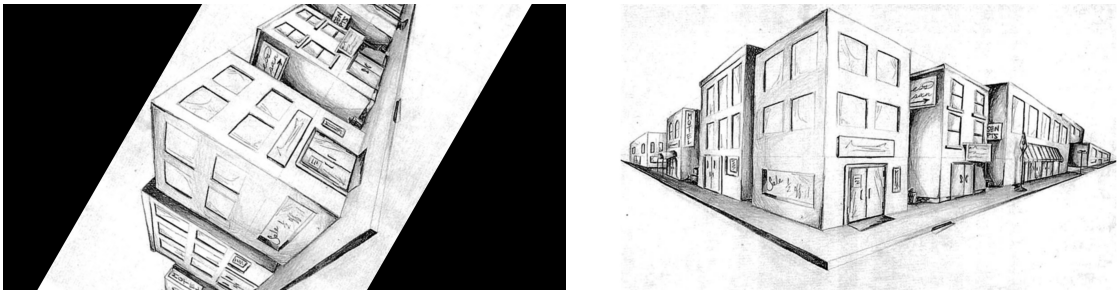
   After tuning with different alpha values, $\alpha = 0.06$ was picked because it produced the least amount of noise.

Figure 3: Harris corner detection

Figure 4: Brown corner detection

The only way for us to calculate the R scores without using determinant or trace is to calculate the eigenvalues for each window's error term. However, this is extremely computationally heavy and does yield a better result than computing the trace and determinant.

2. First we rotate the original image to produce Figure 5a



(a) Rotated image (b) Original Image

Figure 5

By applying Harris corner detection to the rotated image, we get Figure 6. We can see that the detected corners are also rotated.

Figure 6: Rotated Haris corner detection

By rotating the detected corners, we get the following Figure 7.

To prove that the detected points are also rotated by the same angle, we count the number of detected corners in the rotated image ($num\_total\_rotated$). Then we also count the number of corners in the original image corresponding to the corners in the rotated image (num_match). Then we take the fraction of those two $\frac{num\_match}{num\_total\_rotated}$. The result was 99% which is sufficient to prove that the detected corners are the same.

Note that we didn't count the total number of detected corner in the original image ($num\_total\_original$) and compare it with the number of corners detected in the rotated image because rotating the image will caused some information loss, and this measurement won't be accurate.

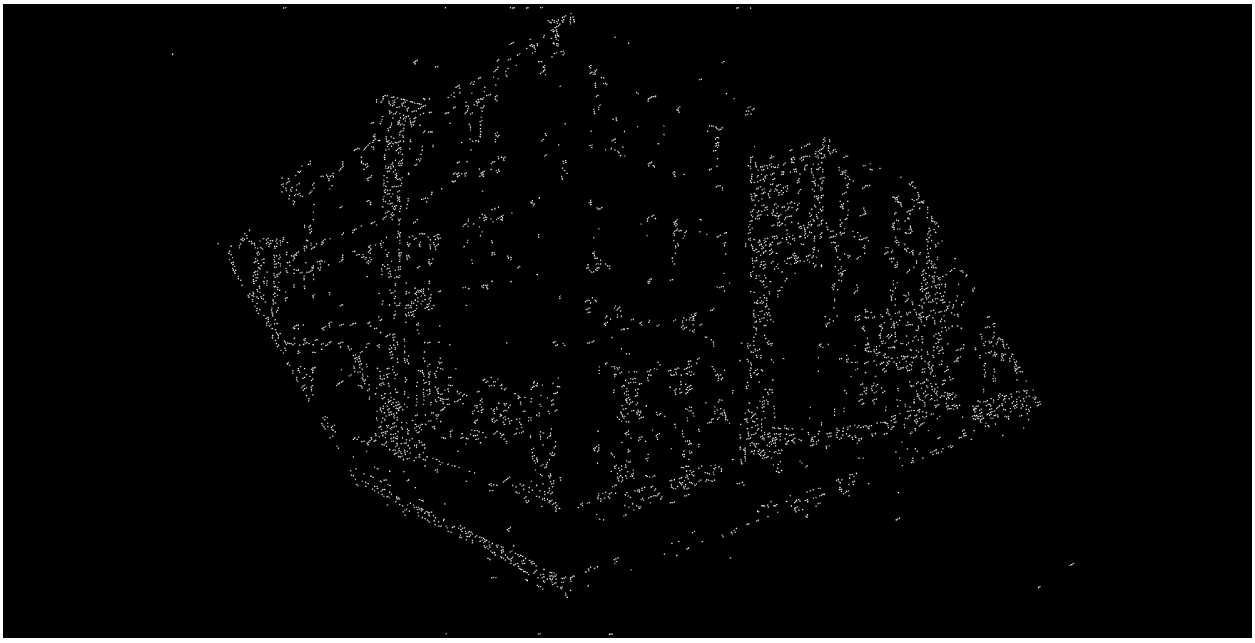


Figure 7: Un-rotate detected corners

3. The generated results are shown in Figure 8.

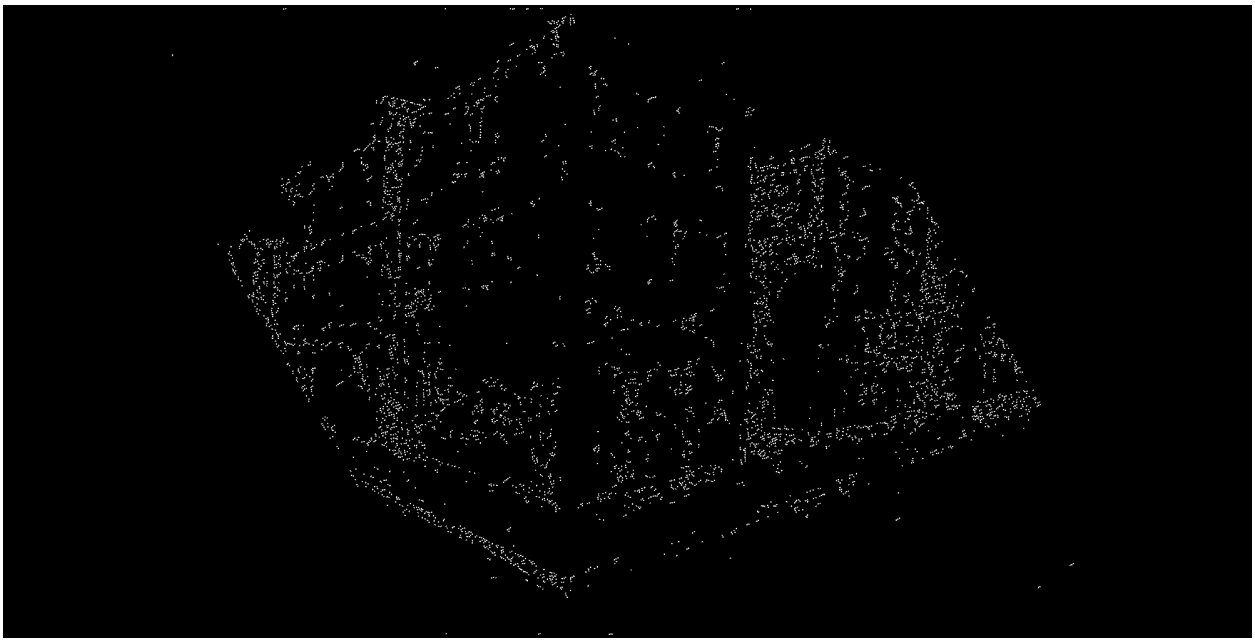


Figure 8: Interest points

By circling out the interest points on the original image, we get Figure **??**.

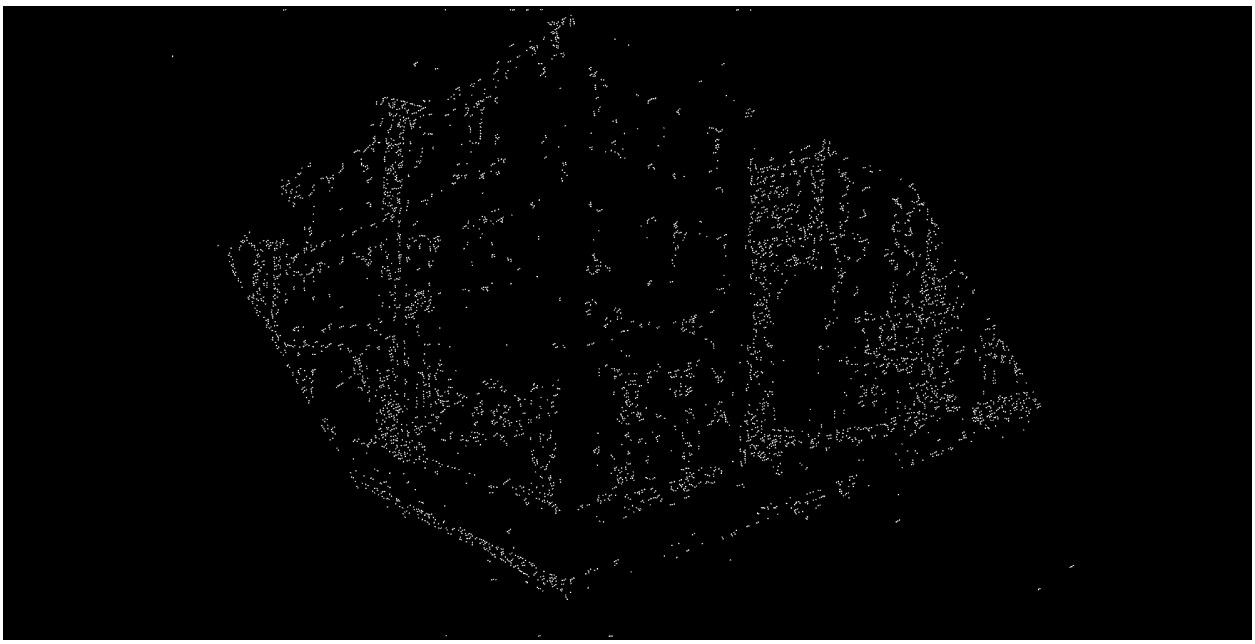


Figure 9: Image with interest points

4. The selected descriptor is SURF (Speed Up Robust Feature).

Key point Detection

In SURF, key points are detected using Box Filters that approximates LoGs. These Box Filters are easier to calculate using the Integral images. Integral images are sum of all the intensities of the image $I_\Sigma(X) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(i,j)$.

The blob detector used in SURF is Hessian matrix

$$H(p,\sigma) = \begin{bmatrix} L_{xx}(p,\sigma) & L_{xy}(p,\sigma) \\ L_{yx}(p,\sigma) & L_{yy}(p,\sigma) \end{bmatrix} \tag{1}$$

where $L_{xx}(p,\sigma)$ etc. is the convolution of the second-order derivative of gaussian with the image $I(x,y)$ at the point p.

The determinant of the Hessian matrix is used as a measure of local change around the point and points are chosen where this determinant is maximal

Key point Descriptor

In order to invariant to image rotation, we identify a reproducible orientation for the interest points. Because of that, first calculate the Haar wavelet responses in x and y direction within a circular neighbourhood of radius 6s around the interest point, with scale s (sampling step is depend on s) at which the interest point was detected. The size of wavelet which is scale depended and its side length is 4s.

Once the wavelet responses are calculated and weighted with a Gaussian =2s centered at the interest point. The responses are represented as points in a space with the horizontal response strength along the abscissa and the vertical response strength along the ordinate. Find maximum the sum of all responses which is wavelet response in every sliding window (/3 window orientation) (see figure 4). The horizontal and vertical responses within the window are summed. From these two horizontal and vertical, summed responses then yield a local orientation vector. The orientation of the interest point can be defined by finding the longest such vector over all windows.

To extract the descriptor, square region which size is 20s are constructed on interested points. Examples of such square regions are illustrated in figure 5

The wavelet responses dx and dy are summed up over each sub-region and form a first set of entries in the feature vector. In order to bring in information about the polarity of the intensity changes, extract the sum of the absolute values of the responses, —dx— and —dy—, each sub-region has a four-dimensional descriptor vector V for its underlying intensity structure V = (dx , dy , —dx — , —dx —). Concatenating this for all, 4 x 4 sub-regions, and this results in a descriptor vector of length is 64. The wavelet responses are invariant to a bias in illumination (offset) and Invariance to contrast (a scale factor) is achieved by turning the descriptor into a unit vector.

Problem 3

1. Gaussian:

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

Gaussian First derivative w.r.t x:

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

Gaussian Second derivative w.r.t x:

$$\begin{aligned}\frac{\partial^2 G(x,y,\sigma)}{\partial x^2} &= -\frac{1}{2\pi\sigma^2}(\frac{x^2}{\sigma^4} e^{-\frac{x^2+y^2}{2\sigma^2}} - \frac{1}{\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}) \\ &= \frac{1}{2\pi\sigma^2} \frac{x^2-\sigma^2}{\sigma^4} e^{-\frac{x^2+y^2}{2\sigma^2}}\end{aligned}$$

Similarly, we can derive the Gaussian Second derivative w.r.t y:

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

Laplacian of Gaussian:

$$\begin{aligned}\nabla^2 G(x,y,\sigma) &= \frac{\partial^2 G(x,y,\sigma)}{\partial x^2} + \frac{\partial^2 G(x,y,\sigma)}{\partial y^2} \\ &= \frac{1}{2\pi\sigma^2} \frac{1}{\sigma^4} e^{-\frac{x^2+y^2}{2\sigma^2}} (x^2+y^2-2\sigma^2) \\ &= -\frac{x^2+y^2-2\sigma^2}{2\pi\sigma^6} e^{-\frac{x^2+y^2}{2\sigma^2}} \\ &= -\frac{1}{\pi\sigma^4}(1-\frac{x^2+y^2}{\sigma^2}) e^{-\frac{x^2+y^2}{2\sigma^2}}\end{aligned}$$

The Laplacian of Gaussian is NOT separable because it can not be separated into the product of two vectors.

2. Difference of Gaussian

$$G(x,y,\sigma_1) - G(x,y,\sigma_2) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{x^2+y^2}{2\sigma_1^2}} - \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{x^2+y^2}{2\sigma_2^2}}$$

Problem 4

1.