ECE661HW4

yang1243

September 2018

1 Logic

1.1 Harris Corner Detector

1.1.1 compute derivatives of image

We first need to compute the derivatives of every pixels of the image using Haar Filter. For a window of length 6 for example, the Haar filter is defined as:

$$\frac{\partial}{\partial x} = \begin{bmatrix} -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 \end{bmatrix} \tag{1}$$

$$\frac{\partial}{\partial y} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix} \tag{2}$$

The length of window should be around $4 * \sigma$ and be a even number.

1.1.2 find corner

For a pixel in the image, define a window to determine whether there is a corner in this pixel. This window's length is better an odd number. In this window, below matrix is computed for all pixels in this window.

$$C = \begin{bmatrix} \sum d_x^2 & \sum d_x d_y \\ \sum d_x d_y & \sum d_y^2 \end{bmatrix} \tag{3}$$

For above matrix, direct threshold $\frac{Det(C)}{Tr(C)^2}$ to find if this pixel is a corner.

1.1.3 use SSD and NCC to match points

SSD is below equation

$$SSD = \sum_i \sum_j |f_1(i,j) f_2(i,j)|^2 \tag{4}$$

NCC is below equation

$$NCC = \frac{\sum\sum(f_1(i,j) - m_1)(f_2(i,j) - m_2)}{\sqrt{\sum\sum(f_1(i,j) - m_1)^2 (f_2(i,j) - m_2)^2}} \tag{5}$$

in which m_1 and m_2 is the mean of $f_1(i,j)$ and $f_2(i,j)$ respectively.

To use SSD or NCC, define a window. Apply this window to a interest point to obtain the information around this window. Apply SSD or NCC to two such windows, the higher the NCC, or the lower the SSD, the more possible for a pair of interest points to match.

1.2 SIFT

1.2.1 find local maxima

In this first step, we need to first build the DoG pyramid. Then find the local maxima respect to x-axis, y-axis and scale. So one pixel has to be compared with 26 volumetric neighborhood to be maxima.

1.2.2 find true maxima with Taylor expansion

For a maximum found in step 1, we use Taylor series to find the true extremum. The true locatio of extremum is given by

$$\vec{x} = -H^{-1}(\vec{x}_0) \ldots J(\vec{x}_0) \tag{6}$$

1.2.3 week out weak maxima

Weekd out weak extremum by thresholding $D(\vec{x}_0) > 0.03$ and some extrema whose support is beyond the edge of img.

1.2.4 build a descriptor for interest point

For the surviving interest point, we build a descriptor of this point regarding the dominant loacl orientation, which include a gradient magnitude and a gradient orientation.

1.2.5 normalize the descriptor

After normalizing the descriptor, this descriptor can be used to find correspond with another image.

1.3 SURF

SURF is designed to identify the interest points that exhibit significant invariance to changes in scale, viewpoint and illumination.

SURF places the interest points at locations where the determinant of the Hessian is maximized. The Hessian matrix is given as below

$$H(x,y,\sigma) = \begin{bmatrix} \frac{\partial^2}{\partial x^2} ff(x,y,\sigma) & \frac{\partial^2}{\partial x \partial y} ff(x,y,\sigma) \\ \frac{\partial^2}{\partial x \partial y} ff(x,y,\sigma) & \frac{\partial^2}{\partial y^2} ff(x,y,\sigma) \end{bmatrix} \tag{7}$$

The Hessian matrix in SURF is computed from Integral Image, where every pixel in the image is represented as

$$I(x,y) = \sum_{i=0}^{i \le x} \sum_{j=0}^{j \le y} f(i,j) \tag{8}$$

Using this representation, the derivatives of each pixel can be computed by six element of $I(x,y)$.

Then we can find the maxima respect to x-axis, y-axis and scale space. Unlike the SIFT where we use a pyramid, here in SURF we do not change the image size with different scale. Thus, interpolation or starting from an upsampled version of the image is required to avoid too large a jump.

After finding the interest point using SURF, we compute the descriptor of each interest point to match between two images as SIFT.

2 Conclusion

From this experience, the NCC is always better than SSd. However, since NCC is too good. Sometimes there is no matches detected for two similar picture. Compared to NCC and SSD, the SIFT and SURF is always good to use. And sometimes it is necessary to tune the parameters, such as thresholds to get best result.

3 Experiment

3.1 Origin image

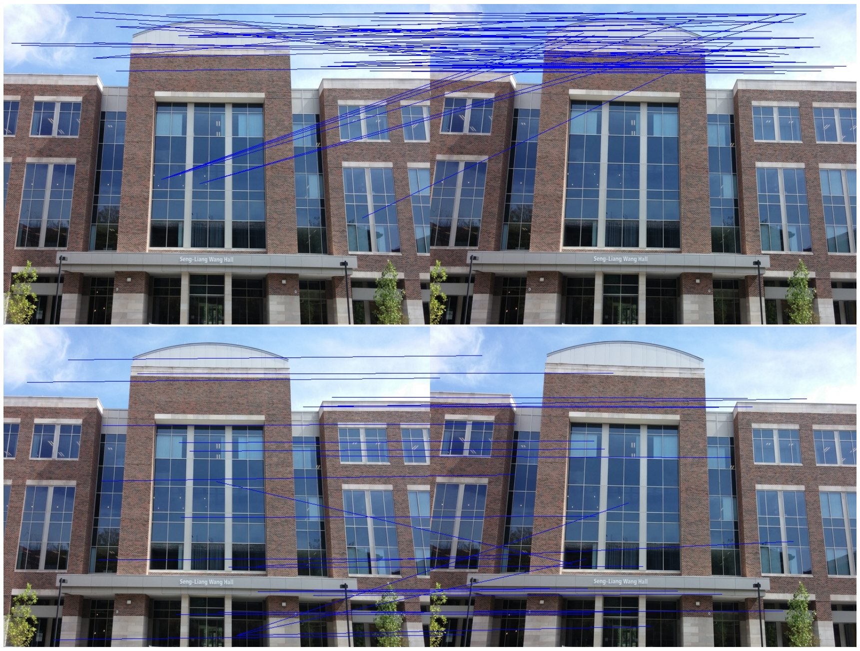

U-HAUL
ONE-WAY & IN-TOWN MOVES


U-HAUL
ONE-WAY & IN-TOWN MOVES
2405

3.2 Harris Corner

3.2.1 pic. 1

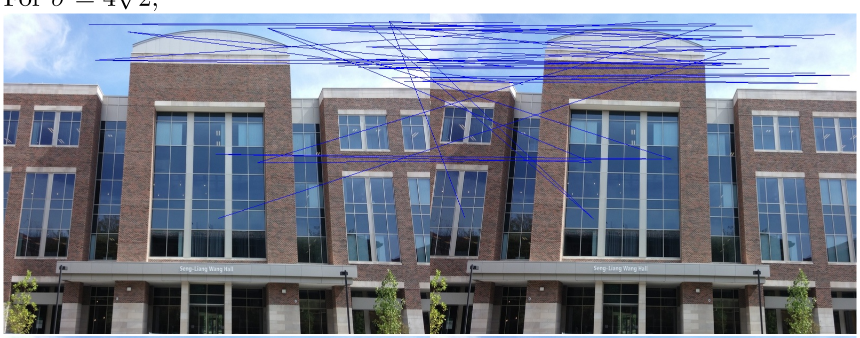
For $\sigma = \sqrt{2}$,



For $\sigma = 2\sqrt{2}$,

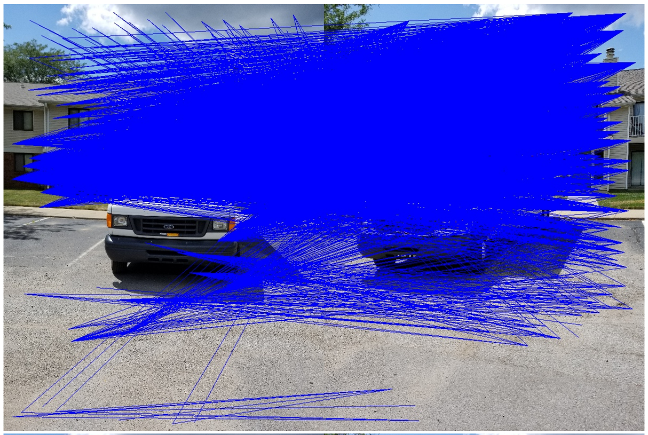
For $\sigma = 3\sqrt{2}$,

For $\sigma = 4\sqrt{2}$,

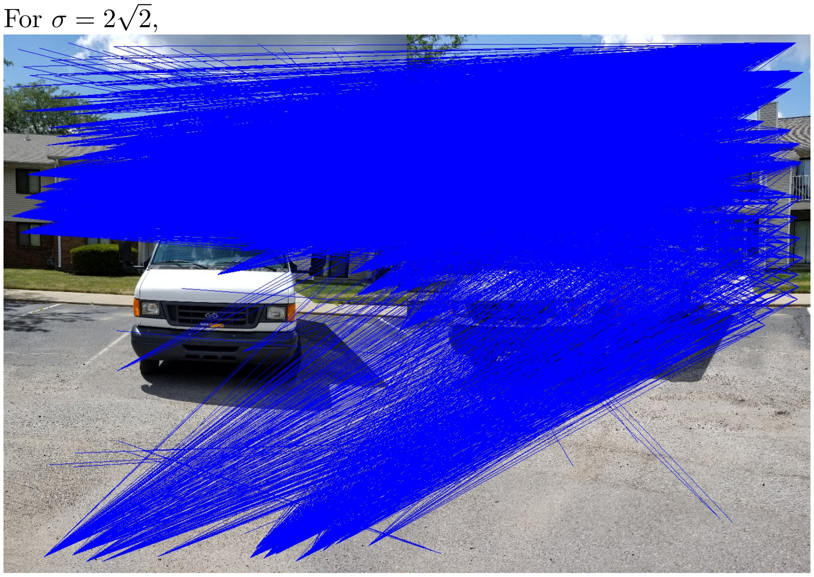


3.2.2 pic. 2

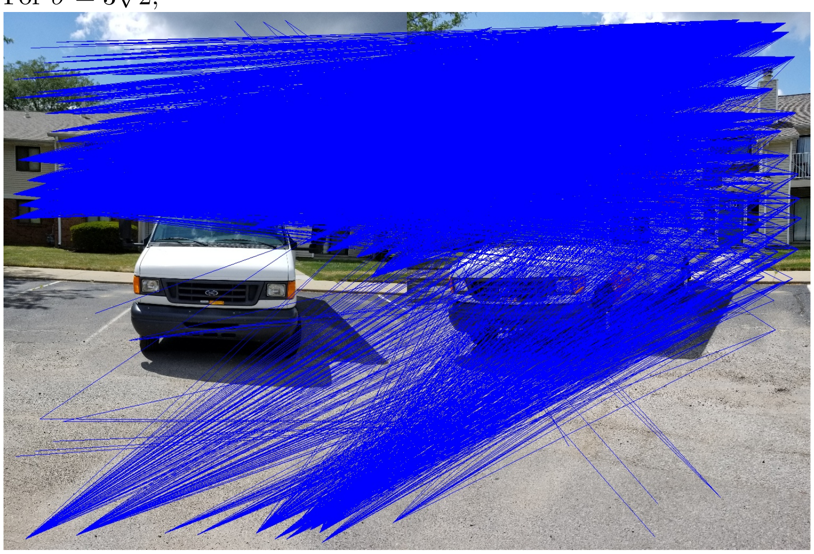
For $\sigma = \sqrt{2}$,



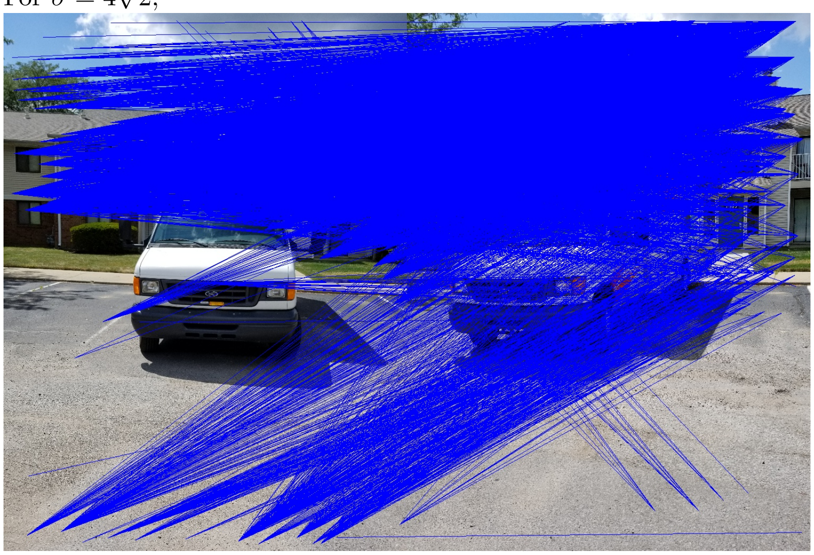
For $\sigma = 2\sqrt{2}$,



For $\sigma = 3\sqrt{2}$,

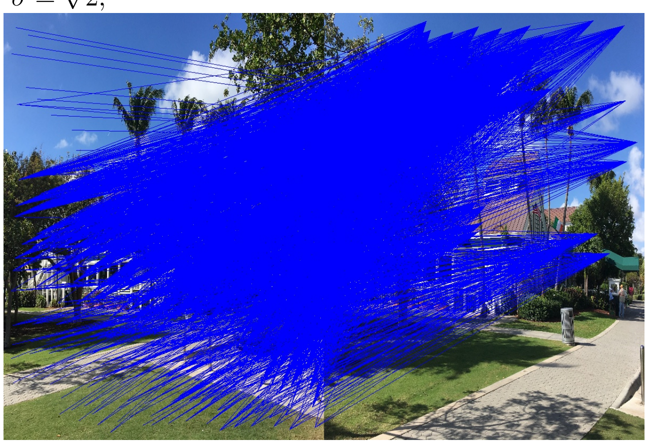


For $\sigma = 4\sqrt{2}$,



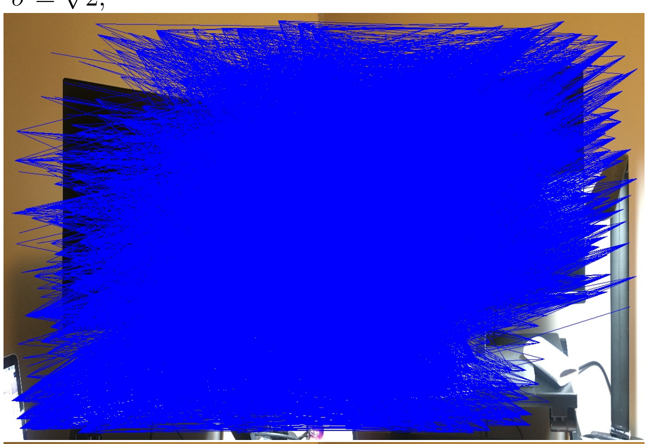
3.2.3 pic. 3

For $\sigma = \sqrt{2}$,

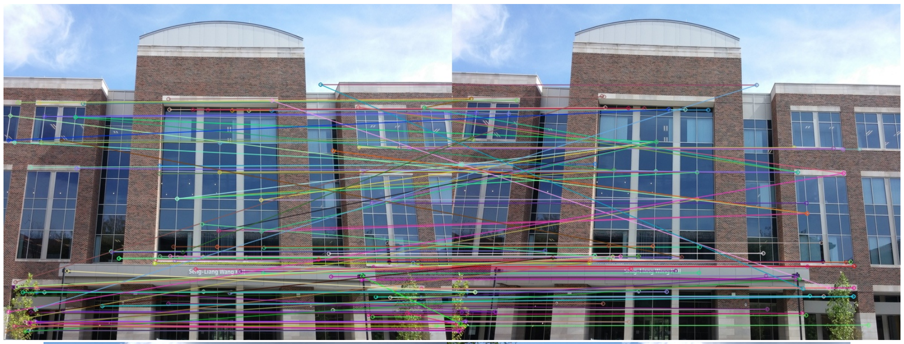


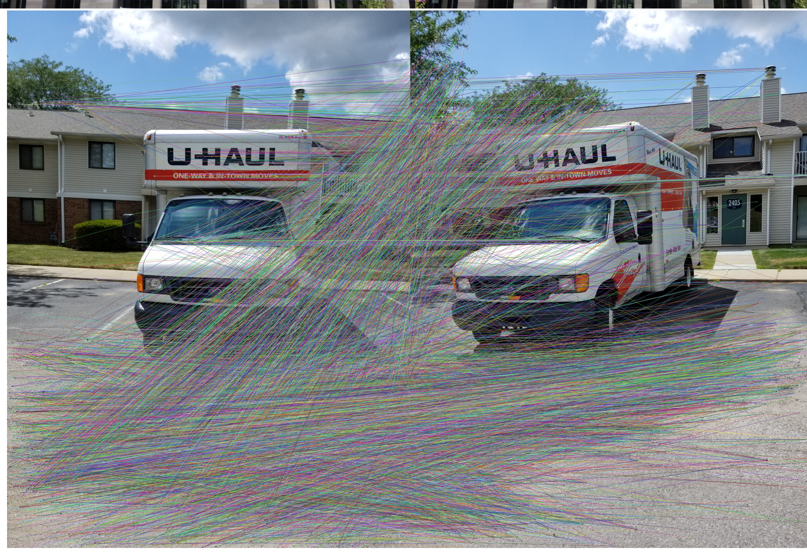
3.2.4 pic. 3

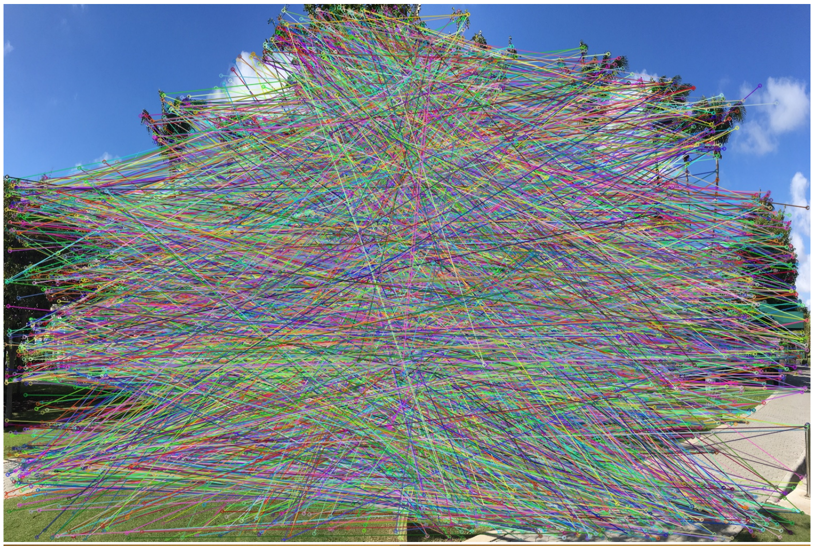
For $\sigma = \sqrt{2}$,



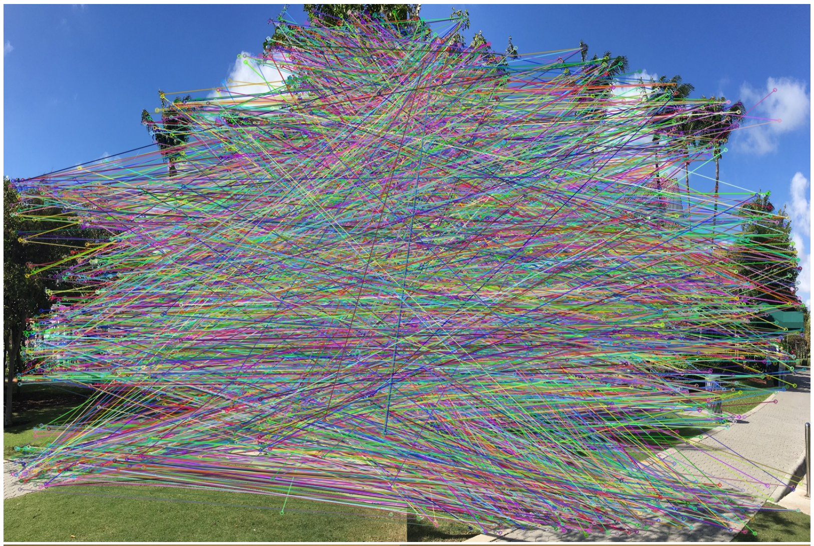
3.3 SIFT

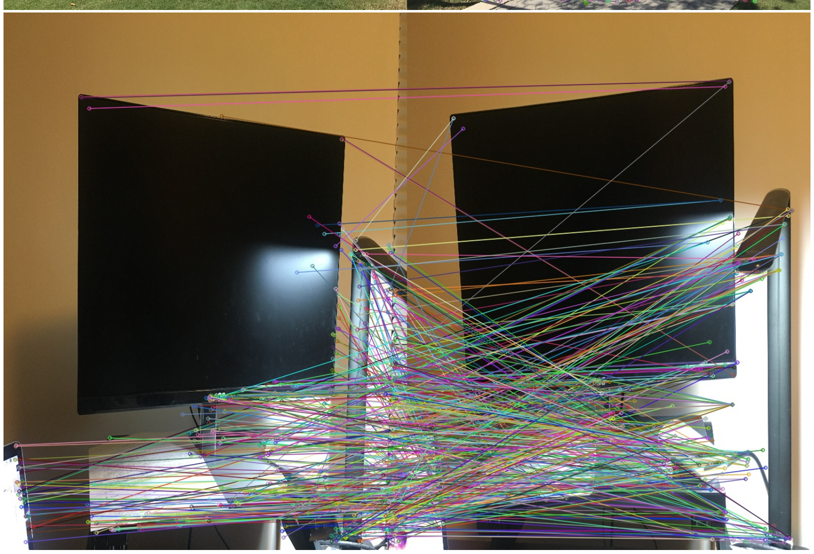






3.4 SURF





4 Code

Code is submitted with this report.