

Wei-Yuan Wen, Yen-Ting Chen, Shun-Jen Lee
A53202335, A53214417, A53209671
University of California, San Diego

November 2, 2016

HOMEWORK 2

Problem 1. Epipolar Geometry Theory [6 pt]

Description:

Suppose a camera calibration gives a transformation (R, T) such that a point in the world maps to the camera by ${}^C P = R^W P + T$.

1. Given calibrations of two cameras (a stereo pair) to a common external coordinate system, represented by R_1, T_1, R_2, T_2 , provide an expression that will map points expressed in the coordinate system of camera 1 to that of camera 2.
2. What is the length of the baseline of the stereo pair.
3. Give an expression for the Essential Matrix in terms of R_1, T_1, R_2, T_2

Solution:

Problem 2. Epipolar Geometry [4 pt]

Description:

Consider two cameras whose image planes are the $z = 1$ plane, and whose focal points are at $(12, 0, 0)$ and $(12, 0, 0)$. Well call a point in the first camera (x, y) , and a point in the second camera (u, v) . Points in each camera are relative to the camera center. So, for example if $(x, y) = (0, 0)$, this is really the point $(12, 0, 1)$ in world coordinates, while if $(u, v) = (0, 0)$ this is the point $(12, 0, 1)$.

1. Suppose the points $(x, y) = (8, 7)$ is matched with disparity of 6 to the point $(u, v) = (2, 7)$. What is the 3D location of this point?
2. Consider points that lie on the 3-D line $X + Z = 2, Y = 0$. Use the same stereo set up as before. Write an analytic expression giving the disparity of a point on this line after it projects onto the two images, as a function of its position in the right image. So your expression should only involve the variables u and d (for disparity). Your expression only needs to be valid for points on the line that are in front of the cameras, i.e. with $Z > 1$.

Solution:

Problem 3. Reconstruction Accuracy [3pt]

Description:

Characterize the accuracy of the 2D stereo system below. Assume the only source of noise is the localization of corresponding points in the two images (in other words, the only source of error is the disparity). Discuss the dependence of the error in depth estimation ($\Delta Z'$) as a function of baseline width (b), focal length (f), and depth (Z').

Solution:

As mentioned in lecture 8, we assume that the localization positions of corresponding points in the two images are X_L and X_R . By using similar triangles, we have:

$$X' = \frac{bX_L}{(X_L - X_R)}, Z' = \frac{bf}{(X_L - X_R)}, \text{ where } (X_L - X_R) \text{ is the disparity.}$$

To observe the error in depth estimation with respect to the noise of disparity, we differentiate Z' to the disparity. We get:

$$\frac{dZ'}{d(X_L - X_R)} = \frac{-bf}{(X_L - X_R)^2} \rightarrow \Delta Z' = \frac{-bf}{(X_L - X_R)^2} \Delta(X_L - X_R)$$

By substituting Z' into the equation, the result becomes:

$$\Delta Z' = \frac{-Z'^2}{bf} \Delta(X_L - X_R)$$

Hence, we can observe that ($\Delta Z'$) is proportional to the square of estimated depth (Z'^2) and inversely proportional to the product of baseline width (b) and focal length (f).

Problem 4. Filters as Templates [16pt]

Description:

In this problem we will play with convolution filters. Filters, when convolved with an image, will fire strongest on locations of an image that look like the filter. This allows us to use filters as object templates in order to identify specific objects within an image. In the case of this assignment, we will be finding cars within an image by convolving a car template onto that image. Although this is not a very good way to do object detection, this problem will show you some of the steps necessary to create a good object detector. The goal of this problem will be to teach some pre-processing steps to make vision algorithms be successful and some strengths and weaknesses of filters. Each problem will ask you to analyze and explain your results. If you do not provide an explanation of why or why not something happened, then you will not get full credit. Provide your code in the appendix.

4.1 Warmup [3pt]

First you will convolve a filter to a synthetic image. The filter or template is `filter.jpg` and the synthetic image is `toy.png`. These files are available on the course webpage. You may want to modify the filter image and original slightly. I suggest `filter_img = filter_img - mean(filter_img(:))`. To convolve the filter image with the toy example, in Matlab you will want to use `conv2`. The output of the convolution will create an intensity image. Provide this image in the report. In the original image (not the image with its mean subtracted), draw a bounding box of the same size as the filter image around the top 3 intensity value locations in the convolved image. Describe how well you think this technique will work on more realistic images? Do you foresee any problems for this algorithm on more realistic images?

4.2 Detection Error [3pt]

We have now created an algorithm that produces a bounding box around a detected object. However we have no way to know if the bounding box is good or bad. In the example images shown above, the bounding boxes look reasonable, but not perfect. Given a ground truth bounding box (g) and a predicted bounding box (p), a commonly used measurement for bounding box quality is $\frac{p \cap g}{p \cup g}$. More intuitively, this is the number of overlapping pixels between the bounding boxes divided by the total number of unique pixels of the two bounding boxes combined. Assuming that all bounding boxes will be axis-aligned rectangles, implement this error function and try it on the toy example in the previous section. Choose 3 different ground truth bounding box sizes around one of the Mickey silhouettes. In general, if the overlap is 50% or more, you may consider that the detection did a good job.

4.3 More Realistic Images [8pt]

Now that you have created an algorithm for matching templates and a function to determine the quality of the match, it is time to try some more realistic images. The file, `cartemplate.jpg`, will be the filter to convolve on each of the 5 other car images (`car1.jpg`, `car2.jpg`, `car3.jpg`, `car4.jpg`, `car5.jpg`). Each image will have an associated text files that contains 2 x, y coordinates (one pair per line). These coordinates will be the ground truth bounding box for each image. For each car image, provide the following:

- A heat map image.
- A bounding box drawn on the original image.

- The bounding box overlap percent.
- A description of what pre-processing steps you needed to do to achieve this overlap.
- An explanation of why you felt these steps made sense.

4.4 Invariance [2pt]

In computer vision there is often a desire for features or algorithms to be invariant to X. One example briefly described in class was illumination invariance. The detection algorithm that was implemented for this problem may have seemed a bit brittle. Can you describe a few things that this algorithm was not invariant to? For example, this algorithm was not scale-invariant, meaning the size of the filter with respect to the size of the object being detected mattered. One filter size should not have worked on everything.

Solution:

4.1 Warmup:

First, I pre-processed the filter by using `filter_img = filter_img - mean(filter_img(:))`. Then I used `conv2` to convolve the filter to the toy image and got an intensity image. The intensity image is shown as Fig. 1(a).

To draw bounding boxes around the top 3 intensity value, I first found the top 3 intensity value and got the responding coordinates. Then I used `rectangle` to draw bounding boxes of the same size as the filter image around these coordinates in the original image. The result is shown as Fig. 1(b).

I think this technique won't work well on more realistic images since the background of the input image is pretty simple. Also, the filter is pretty as the same size of the object we want to find, so we won't need to do much pre-processing steps on the filter image. If we apply this algorithm on more realistic images, the maximum of the intensity image may not be located at the object we want to find, since the background will be complicated and distracting. This will mislead us to draw the bounding box on different location in the image. On the other hand, we may not use only a simple filter image to find various kind of object in the same category, or we need to do some more pre-processing steps to get an accurate bounding box.

The codes are attached as Listing 1 in Appendix.

4.2 Detection Error:

I wrote a function named `detectError`, which takes ground truth box and predicted bounding box as two arguments. Since I have the length and width and at least 2 coordinates of these rectangles, I can easily compute the number of overlapping pixels of the two bounding boxes between the bounding boxes and then divide it by the total number of unique pixels of the two bounding boxes. I then constructed 3 ground truth bounding boxes around one of the Mickey silhouettes and found the overlap is more than 50%.

The codes are attached as Listing 3 in Appendix.

4.3 More Realistic Images:

For `car1` image:

- Bounding box overlap percent: 94.72%.
- Pre-processing steps:
 - a) Subtracted both filter image and `car1` image by their means.
 - b) Crop the filter.
 - c) Rescale the filter image to the size of ground truth bounding box.
 - d) Flip the filter image in x and y dimensions before convolution.
- Explanation:
 - a) I want the object in `car1` that looks like the filter image has maximums of intensity values, so I subtracted both filter image and `car1` image by their means.
 - b) I first cropped the filter image to get rid of the blank area, because I don't want the blank to affect my result of convolution.
 - c) Then I rescaled the filter image to the size of ground truth bounding box, this way the intensity image will be more accurate and make the maximum locate at the position of `car1`.
 - d) To do the convolution, we need to make sure the filter image is upside down and reversed from left to right, so I flipped it in x and y dimensions right before convolution.
- The result is shown as Fig. 2(a) and Fig. 2(b).

For `car2` image:

- Bounding box overlap percent: 89.59%.
- Pre-processing steps:
 - a) Subtracted both filter image and `car2` image by their means.
 - b) Crop the filter.
 - c) Rescale the filter image to the size of ground truth bounding box.
 - d) Reverse the filter image from left to right.
 - e) Flip the filter image in x and y dimensions before convolution.
- Explanation:
 - a) I want the object in `car2` that looks like the filter image has maximums of intensity values, so I subtracted both filter image and `car2` image by their means.
 - b) I cropped the filter to get rid of the blank area, because I don't want the blank to affect my result of convolution.
 - c) Then I rescaled the filter image to the size of ground truth bounding box, this way the intensity image will be more accurate and make the maximum locate at the position of `car2`.
 - d) I reversed the filter image from left to right, since the head of the car in the `car2` image points toward right.
 - e) To do the convolution, we need to make sure the filter image is upside down and reversed from left to right, so I flipped it in x and y dimensions right before convolution.
- The result is shown as Fig. 3(a) and Fig. 3(b).

For `car3` image:

- Bounding box overlap percent: 50.80%.

- Pre-processing steps:
 - a) Subtracted both filter image and `car3` image by their means.
 - b) Crop the filter.
 - c) Rescale the filter image to the size of the rear part of the car in `car3` image.
 - d) Flip the filter image in x and y dimensions before convolution.
- Explanation:
 - a) I want the object in `car3` that looks like the filter image has maximums of intensity values, so I subtracted both filter image and `car3` image by their means.
 - b) I cropped the filter to get rid of the blank area, because I don't want the blank to affect my result of convolution.
 - c) Then I rescaled the filter image to the size of the rear part of the car in `car3` image, because I thought somehow that rear part looks more alike with the filter image than the side part. This way the intensity image will be more accurate and make the maximum locate at the position of `car3`.
 - d) To do the convolution, we need to make sure the filter image is upside down and reversed from left to right, so I flipped it in x and y dimensions right before convolution.
- The result is shown as Fig. 4(a) and Fig. 4(b).

For `car4` image:

- Bounding box overlap percent: 65.35%.
- Pre-processing steps:
 - a) Subtracted both filter image and `car2` image by their means.
 - b) Crop the filter.
 - c) Add 0.1 on each pixel in the filter image.
 - d) Rescale the filter image to the size of ground truth bounding box.
 - e) Reverse the filter image from left to right.
 - f) Flip the filter image in x and y dimensions before convolution.
- Explanation:
 - a) I want the object in `car4` that looks like the filter image has maximums of intensity values, so I subtracted both filter image and `car4` image by their means.
 - b) I cropped the filter to get rid of the blank area, because I don't want the blank to affect my result of convolution.
 - c) I found the top-left area in `car4` image has similar brightness with the filter image, which means the intensity is likely to locate at that area. Thus I add 0.1 to every pixels in the filter image to make it brighter.
 - d) Then I rescaled the filter image to the size of ground truth bounding box, this way the intensity image will be more accurate and make the maximum locate at the position of `car4`.
 - e) I reversed the filter image from left to right, since the head of the car in the `car2` image points toward right.
 - f) To do the convolution, we need to make sure the filter image is upside down and reversed from left to right, so I flipped it in x and y dimensions right before convolution.

- The result is shown as Fig. 5(a) and Fig. 5(b).

For `car5` image:

- Bounding box overlap percent: 74.40%.
- Pre-processing steps:
 - Subtracted both filter image and `car5` image by their means.
 - Crop the filter.
 - Add 0.1 on each pixel in the filter image.
 - Rescale the filter image to the size of ground truth bounding box.
 - Reverse the filter image from left to right.
 - Flip the filter image in x and y dimensions before convolution.
- Explanation:
 - I want the object in `car5` that looks like the filter image has maximums of intensity values, so I subtracted both filter image and `car5` image by their means.
 - I cropped the filter to get rid of the blank area, because I don't want the blank to affect my result of convolution.
 - I found that rhino in `car5` image and the right part of the forest has similar brightness with the filter image, which means the intensity is likely to locate at that area. Thus I add 0.1 to every pixels in the filter image to make it brighter.
 - Then I rescaled the filter image to the size of ground truth bounding box, this way the intensity image will be more accurate and make the maximum locate at the position of `car5`.
 - I reversed the filter image from left to right, because I thought the trunk of truck in the `car5` image somehow seems like the head of the filter image.
 - To do the convolution, we need to make sure the filter image is upside down and reversed from left to right, so I flipped it in x and y dimensions right before convolution.
- The result is shown as Fig. 6(a) and Fig. 6(b).

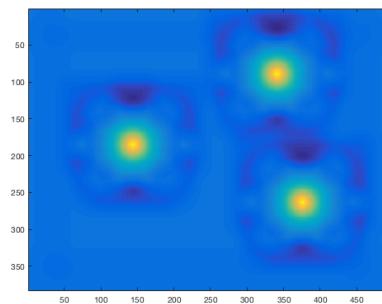
The codes are attached as Listing 1 to Listing 5 in Appendix.

4.4 Invariance:

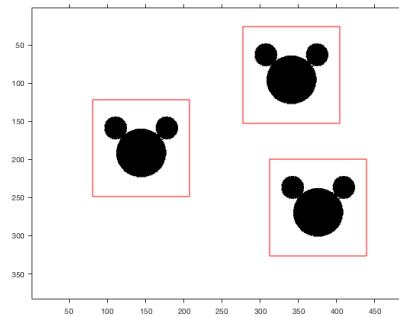
- Not color-invariant, the color/brightness of the filter with respect to the color/brightness of the object being detected mattered. We may need to adjust the brightness of the filter to find the correct object in the image. For example, in the `car4` and `car5` images, the darker area like smoke, forest and rhinos will mislead us to detect wrong objects.
- Not shape-invariant, the silhouette of the filter with respect to the silhouette of the object being detected mattered. For example, in the `car3` image, we can see that the silhouette of a car may not appear at the correct location we expect.

Result:

4.1 Warmup:



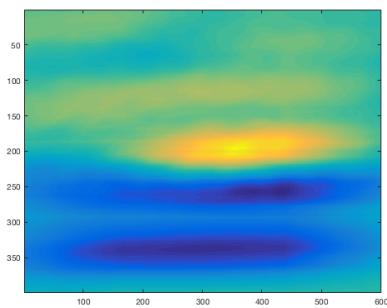
(a)



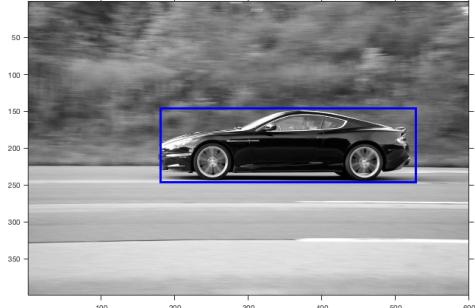
(b)

Figure 1: (a) Intensity image, (b) Bounding boxes.

4.3 More Realistic Images:



(a)



(b)

Figure 2: Outputs for `car1`. (a) Heat map, (b) Bounding boxes.

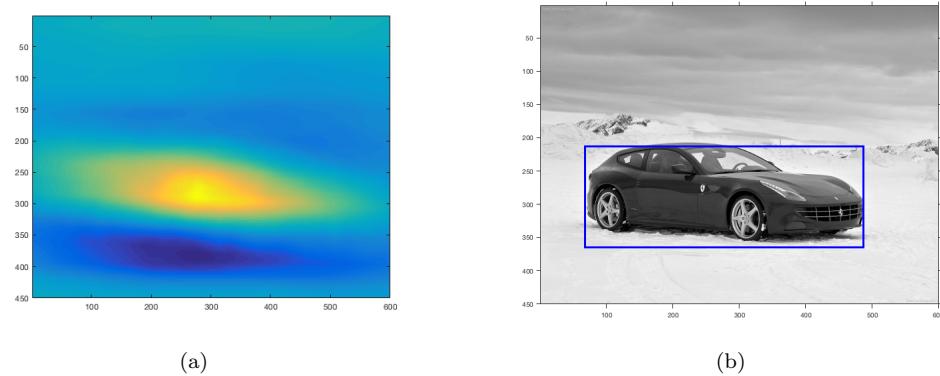


Figure 3: Outputs for `car2`. (a) Heat map, (b) Bounding boxes.

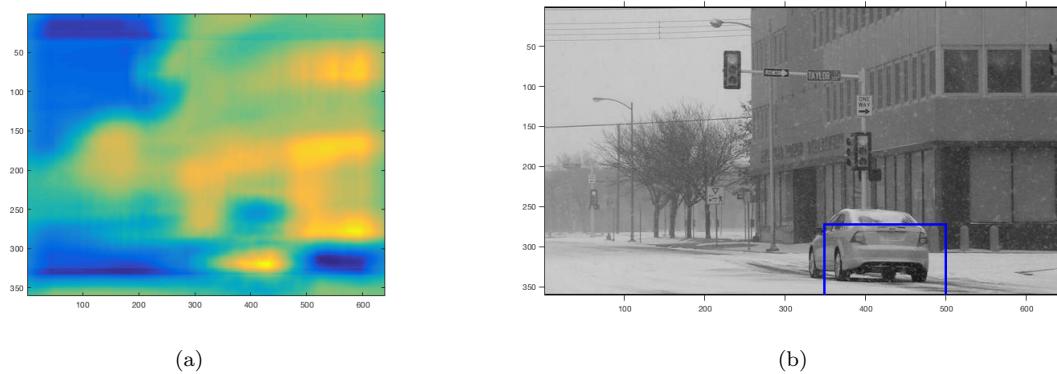


Figure 4: Outputs for `car3`. (a) Heat map, (b) Bounding boxes.

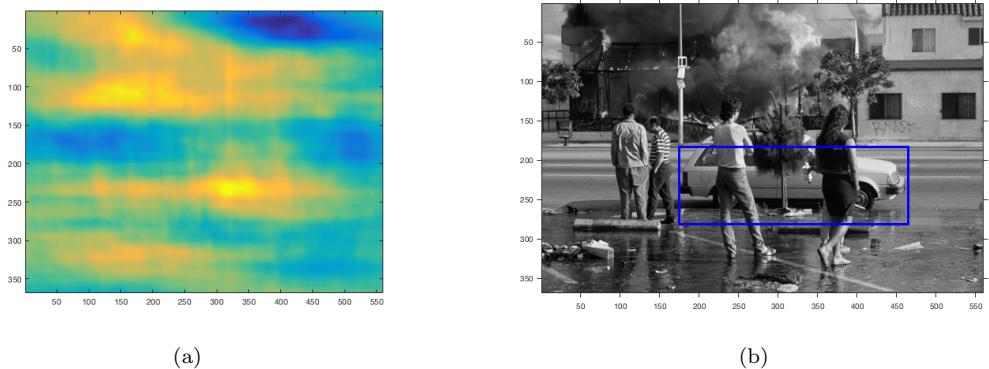


Figure 5: Outputs for `car4`. (a) Heat map, (b) Bounding boxes.

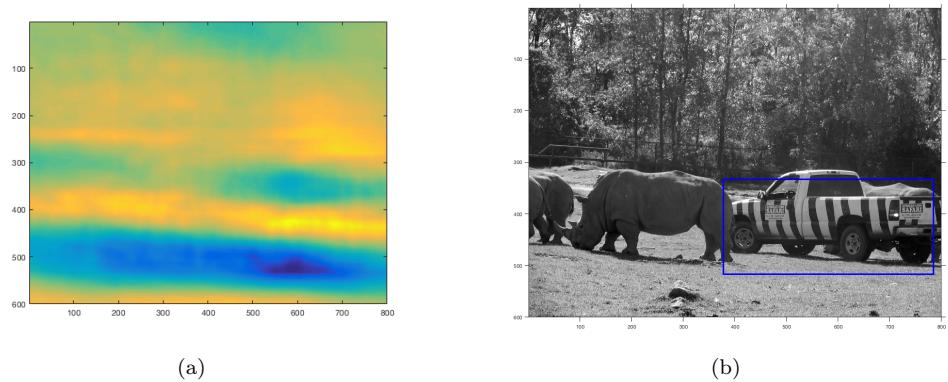


Figure 6: Outputs for `car5`. (a) Heat map, (b) Bounding boxes.

Problem 5. Canny Edge Detection [9pt]

Description:

In this problem, you have to write a function to do CANNY EDGE DETECTION. The following steps need to be implemented.

- **Smoothing [1pt]:** First, we need to smooth the images to remove noise from being considered as edges. For this assignment, use a 5×5 Gaussian kernel filter 1 with $\sigma = 1.4$ to smooth the images.

$$K = \frac{1}{159} \begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 12 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix} \quad (1)$$

- **Gradient Computation [2pt]:** After you have finished smoothing, find the image gradient in the horizontal and vertical directions. You can use Sobel operators 2 as your filter kernel to calculate G_x and G_y . G_x and G_y are the gradients along the x and y axis respectively. s_x and s_y are the corresponding kernels. Compute the gradient magnitude image as $|G| = \sqrt{G_x^2 + G_y^2}$. The edge direction as each pixel is given as $G_\theta = \tan^{-1}\left(\frac{G_x}{G_y}\right)$.

$$s_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, s_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}. \quad (2)$$

- **Non Maximum Suppression [3pt]:** Our desired edges need to be sharp, not thick like the ones in gradient image. Use non maximum suppression to preserve all local maxima and discard the rest. You can use the following method to do so:

For each pixel do:

- Round the gradient direction θ to the nearest multiple of 45° in a 8-connected neighbourhood.
- Compare the edge strength at the current pixel to the pixels along the $+ve$ and $-ve$ gradient direction in the 8-connected neighbourhood.
- Preserve the values of only those pixels which have maximum gradient magnitudes in the neighbourhood along the $+ve$ and $-ve$ gradient direction.

- **Hysteresis Thresholding [3pt]:** Choose optimum values of thresholds and use the thresholding approach given in lecture 7. This will remove the edges caused due to noise and colour variations.

Compute the images after each step and select suitable thresholds that retains most of the true edges. For this question use the image `geisel.jpeg` in the data folder.

Solution:

For smoothing, I used `conv2.m` and applied the Gaussian filter with $\sigma = 1.4$ to smooth the image. The result is shown as Fig. 8(a).

For gradient computation, I use s_x and s_y to calculate G_x and G_y respectively. Then I computed the gradient magnitude image using $|G| = \sqrt{G_x^2 + G_y^2}$ and get the edge direction at each pixel by computing $G_\theta = \tan^{-1}(\frac{G_x}{G_y})$. The result is shown as Fig. 8(b).

For non maximum suppression, I first rounded every gradient direction at each pixel to the nearest multiple of 45° . And for each pixel, I checked if the gradient magnitude of two pixels along the *+ve* and *-ve* gradient direction in the 8-connected neighbourhood is smaller than that of current pixel. If both magnitudes are smaller, I added the current pixel to the edge candidate. The result is shown as Fig. 8(c).

For hysteresis thresholding, I first added all the pixels, which have values higher than τ_{high} , from edge candidate into an dynamic array. For each pixel in the array, I added it into a result map and added pixels within 8-connected neighbourhood of it into the array if those pixels' gradient magnitude are higher than τ_{low} and are not already in the result map. The result is shown as Fig. 8(d).

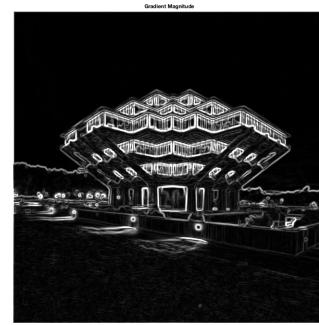
The codes are attached as Listing 6 to Listing 7 in Appendix.

Result:

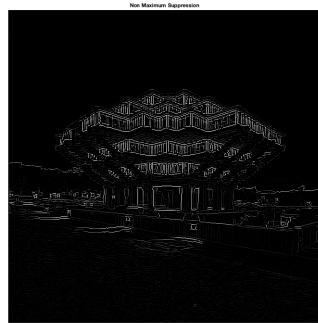
Figure 7: Original Image



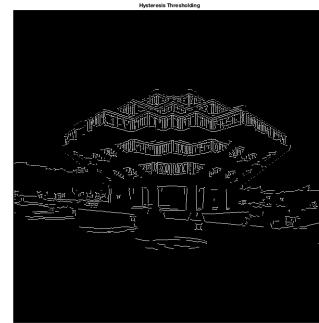
(a)



(b)



(c)



(d)

Figure 8: (a) Smoothed Image using Gaussian filter with $\sigma = 1.4$, (b) Gradient Magnitude, (c) Non Maximum Suppression edges, (d) Final Edge after Hysteresis Thresholding.

Problem 6. Sparse Stereo Matching [27pt]**Description:**

In this problem we will play around with sparse stereo matching methods. You will work on two image pairs, a warrior figure and a figure from the Matrix movies (`warrior2.mat` and `matrix2.mat`). These files both contain two images, two camera matrices, and sets of corresponding points (extracted by manually clicking the images). For illustration, I have run my code on a third image pair (`dino2.mat`). This data is also provided on the webpage for you to debug your code, but you should only report results on warrior and matrix. In other words, where I include one (or a pair) of images in the assignment below, you will provide the same thing but for BOTH matrix and warrior. Note that the matrix image pair is harder, in the sense that the matching algorithms we are implementing will not work quite as well. You should expect good results, however, on warrior. To make the TAs extra happy, make the line width and marker sizes bigger than the default sizes.

6.1 Corner Detection [8pt]

The first thing we need to do is to build a corner detector. This should be done according to <http://cseweb.ucsd.edu/classes/fa16/cse252A-a/lec7.pdf>. Your file should be named `CornerDetect.m`, and take as input

```
corners = CornerDetect(Image, nCorners, smoothSTD, windowHeight)
```

where `smoothSTD` is the standard deviation of the smoothing kernel, and `windowSize` the size of the smoothing window. In the lecture the corner detector was implemented using a hard threshold. Do not do that but instead return the `nCorners` strongest corners after non-maximum suppression. This way you can control exactly how many corners are returned. Run your code on all four images (with `nCorners = 20`) and show outputs.

Solution:

To detect the corners, we follow the method in lecture 7.

1. Filter image with a Gaussian.
2. Compute the gradient in both x and y directions for every pixel.
3. Move window over image, and for each window location:
 - (1) Construct the matrix C over the window.

$$C(x, y) = \begin{pmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{pmatrix}$$

- (2) Find the eigenvalue for C using `eig()` in matlab.
- (3) Store the smaller eigenvalue for every pixel in a matrix.
4. Perform non-maximum suppression by comparing the stored eigenvalue in every pixel with its neighbors(a window). If the eigenvalue is the maximum in the window, we take this pixel as a candidate of a corner.
5. We sort all the candidates by their eigenvalues, and choose the largest nCorners pixels as our result.

Result:

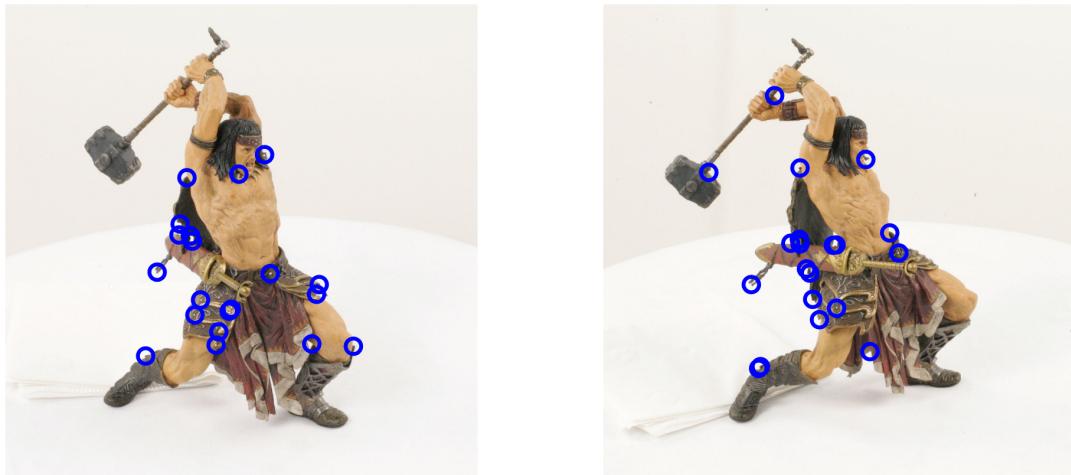


Figure 9: Corner Detection: Warrior

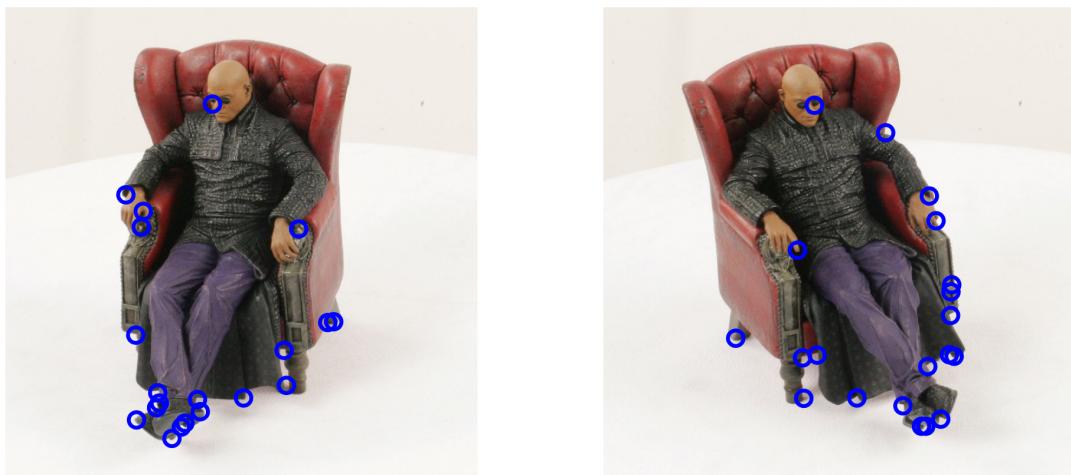


Figure 10: Corner Detection: Matrix

6.2 SSD Matching [2pt]

Write a function `SSDmatch.m` that implements the SSD matching algorithm for two input windows. Include this code in your report (in appendix as usual).

Solution:

To implement the SSD matching algorithm, we simply vectorizes the two input windows and then take the 2-norm of their difference.

6.3 Naive Matching [4pt]

Equipped with the corner detector and the SSD matching code, we are ready to start finding correspondances. One naive strategy is to try and find the best match between the two sets of corner points. Write a script that does this, namely, for each corner in `image1`, find the best match from the detected corners in `image2` (or, if the SSD match score is too low, then return no match for that point). You will have to figure out a good threshold (`SSDth`) value by experimentation. Write a function `naiveCorrespondanceMatching.m` and call it as below. Examine your results for 10, 20, and 30 detected corners in each image. In your report, only include your results for 10 corners, so that the figure will not be cluttered. `naiveCorrespondanceMatching.m` will call your SSD matching code. The parameter `R` below, is the radius of the patch used for matching.

```
ncorners = 10;
corners1 = CornerDetect(I1, ncorners, smoothSTD, windowSize));
corners2 = CornerDetect(I2, ncorners, smoothSTD, windowSize));
[I, corsSSD] = naiveCorrespondanceMatching(I1, I2, corners1, corners2, R, SSDth);
```

Solution:

We simply follow the instruction above to search the best match in `image2` for every corner in `image1`. We adjust the SSD threshold so that it can have the highest percentage of correct matches. In the end, we got 4 correct matches among 7 matches in `Warrior` images and 3 correct matches among 9 matches in `Matrix` images.

Result:

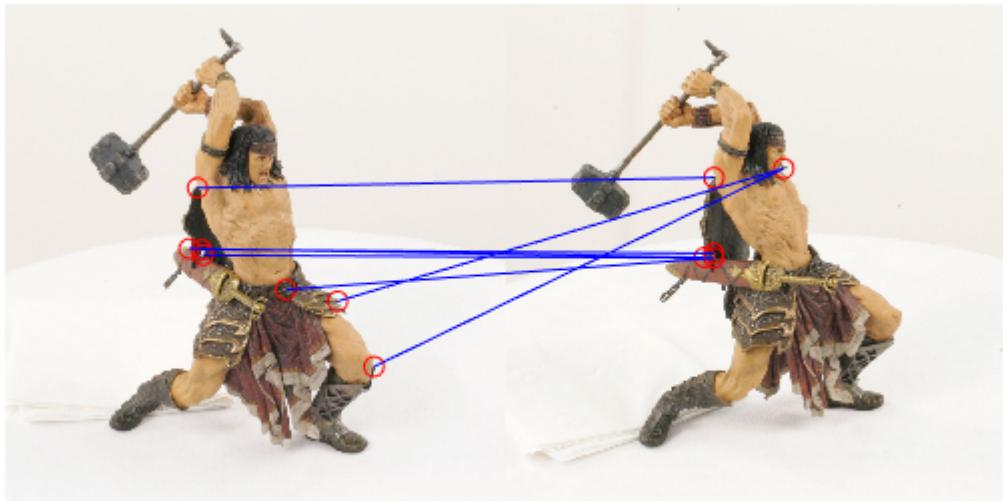


Figure 11: Naive Matching: Warrior

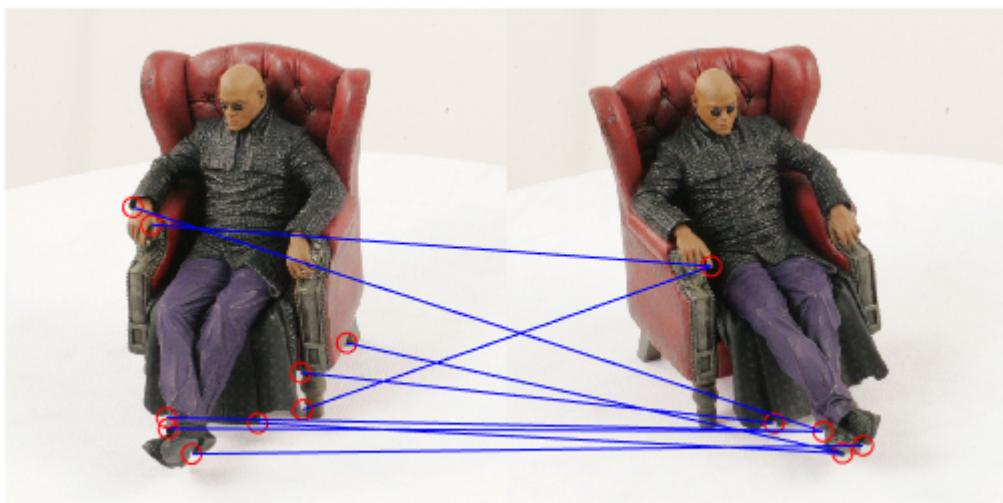


Figure 12: Naive Matching: Matrix

6.4 Epipolar Geometry [3pt]

Using the provided `fund.m` together with the provided points `cor1`, and `cor2`, calculate the fundamental matrix, and then plot the epipolar lines in both images pairs as shown below. Plot the points and make sure the epipolar lines go through them. You might find the supplied `linePts.m` function useful when you are working with epipolar lines.

Solution:

To compute the epipolar line in image1 for the corner in image2, we use $l_1 = \mathbf{F}^T p_2$

To compute the epipolar line in image2 for the corner in image1, we use $l_2 = \mathbf{F}p_1$

The \mathbf{F} above is the fundamental matrix.

Result:

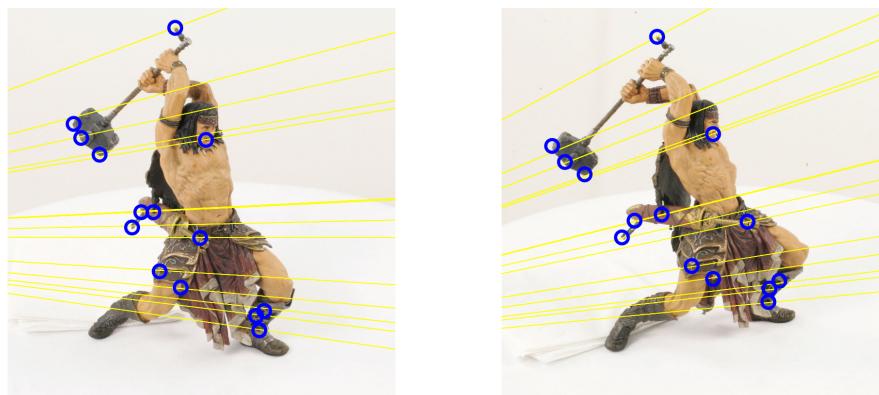


Figure 13: Epipolar Lines: Warrior

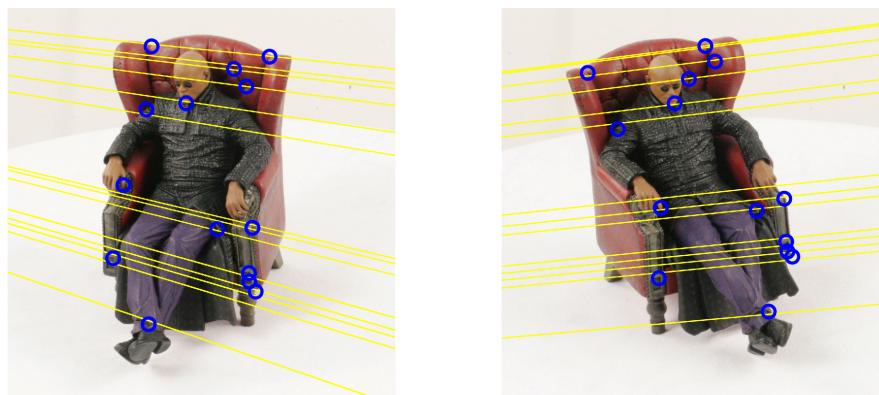


Figure 14: Epipolar Lines: Matrix

6.5 Epipolar Geometry Based Matching [5pt]

We will now use the epipolar geometry to build a better matching algorithm. First, detect 10 corners in Image1. Then, for each corner, do a linesearch along the corresponding epipolar line in Image2. Evaluate the SSD score for each point along this line and return the best match (or no match if all scores are below the SSDth). R is the radius (size) of the SSD patch in the code below. You do not have to run this in both directions, but only as indicated in the code below.

```
ncorners = 10;
F = fund(cor1, cor2);
corners1 = CornerDetect(I1, ncorners, smoothSTD, windowSize));
corsSSD = correspondanceMatchingLine( I1, I2, corners1, F, R, SSDth);
```

Solution:

To do the linesearch, we have to interpolate the pixels from the line. We interpolate the pixels by first using the supplied `linePts.m` to find the two end points of the line. Then for every x and y in between these end points, we compute a corresponding y and x by the slope of the line.

The other processes are as same as the processes in 6.3.

Result:

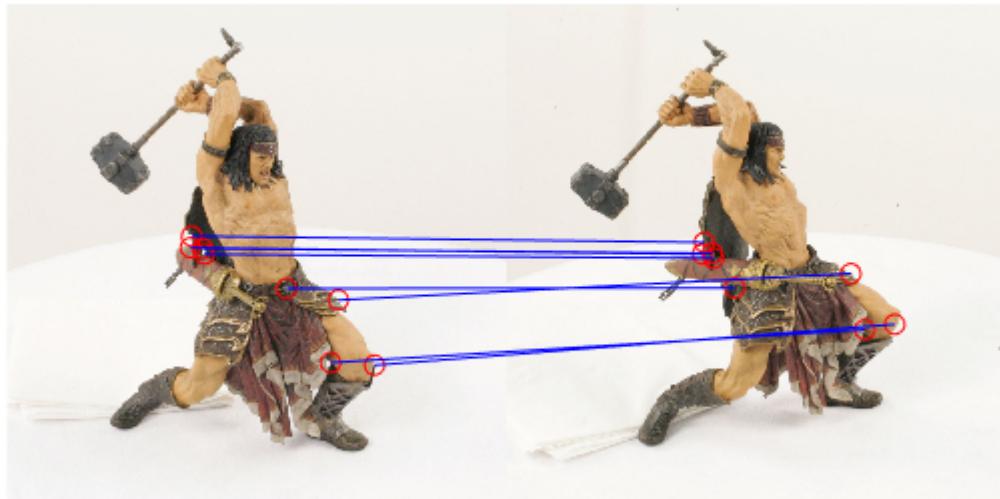


Figure 15: Epipolar Geometry Based Matching: Warrior

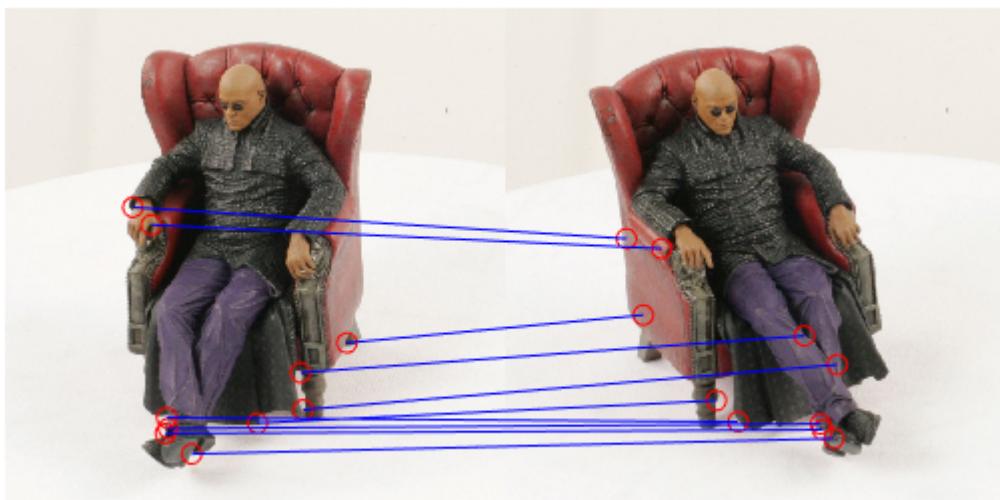


Figure 16: Epipolar Geometry Based Matching: Matrix

6.6 Triangulation [5pt]

Now that you have found correspondences between the pairs of images we can triangulate the corresponding 3D points. Since we do not enforce the ordering constraint the correspondences you have found are likely to be noisy and to contain a fair amount of outliers. Using the provided camera matrices you will triangulate a 3D point for each corresponding pair of points. Then by reprojecting the 3D points you will be able to find most of the outliers. You should implement the linear triangulation method described in lecture. P_1 and P_2 below are the camera matrices. Also write a function, `findOutliers.m`, that reprojects the world points to Image2, and then determines which points are outliers and inliers respectively. For this purpose, we will call a point an outlier if the distance between the true location, and the reprojected point location is more than 20 pixels.

```
outlierTH = 20;
F = fund(cor1, cor2);
ncorners = 50;
corners1 = CornerDetect(I1, ncorners, smoothSTD, windowSize));
corsSSD = correspondanceMatchingLine( I1, I2, corners1, F, R, SSDth);
points3D = triangulate(corsSSD, P1, P2);
[ inlier, outlier ] = findOutliers(points3D, P2, outlierTH, corsSSD);
```

Display your results by showing, for Image2, the original points in black circles, the inliers as blue plus signs, and the outliers as red plus signs, as shown in Figure 7. Compare this outlierplot with Figure 6. Does the detected outliers correspond to false matches?

Solution: To implement the triangulation, we use the linear triangulation method described in lecture. First, for a pair of corresponding corners in image1 and image2, we each construct two lines that passes the corner. We choose the lines simply by

$$\begin{aligned}x + y - (x_1 + y_1) &= 0 \\x - y - (x_1 - y_1) &= 0 \\x + y - (x_2 + y_2) &= 0 \\x - y - (x_2 - y_2) &= 0\end{aligned}$$

, where $(x_1, y_1), (x_2, y_2)$ are the corners in image1 and image2.

Follow the method mentioned in the lecture, we construct the matrix \mathbf{A} , where

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & -(x_1 + y_1) \\ 1 & -1 & -(x_1 - y_1) \\ 1 & 1 & -(x_2 + y_2) \\ 1 & -1 & -(x_2 - y_2) \end{pmatrix}$$

We solve the null space for $\mathbf{AX} = 0$ by using singular value decomposition and we take the last column of matrix \mathbf{V} as our answer.

To detect outliers, we project the 3D points obtained by triangulation back to image2 and see if the distance between the projection location and the original point is greater than `outlierTH`. We project the 3D point to image2 by $\mathbf{x}_2(reproject) = \mathbf{P}_2\mathbf{X}$, where \mathbf{P}_2 is the camera matrix for image2.

Result:

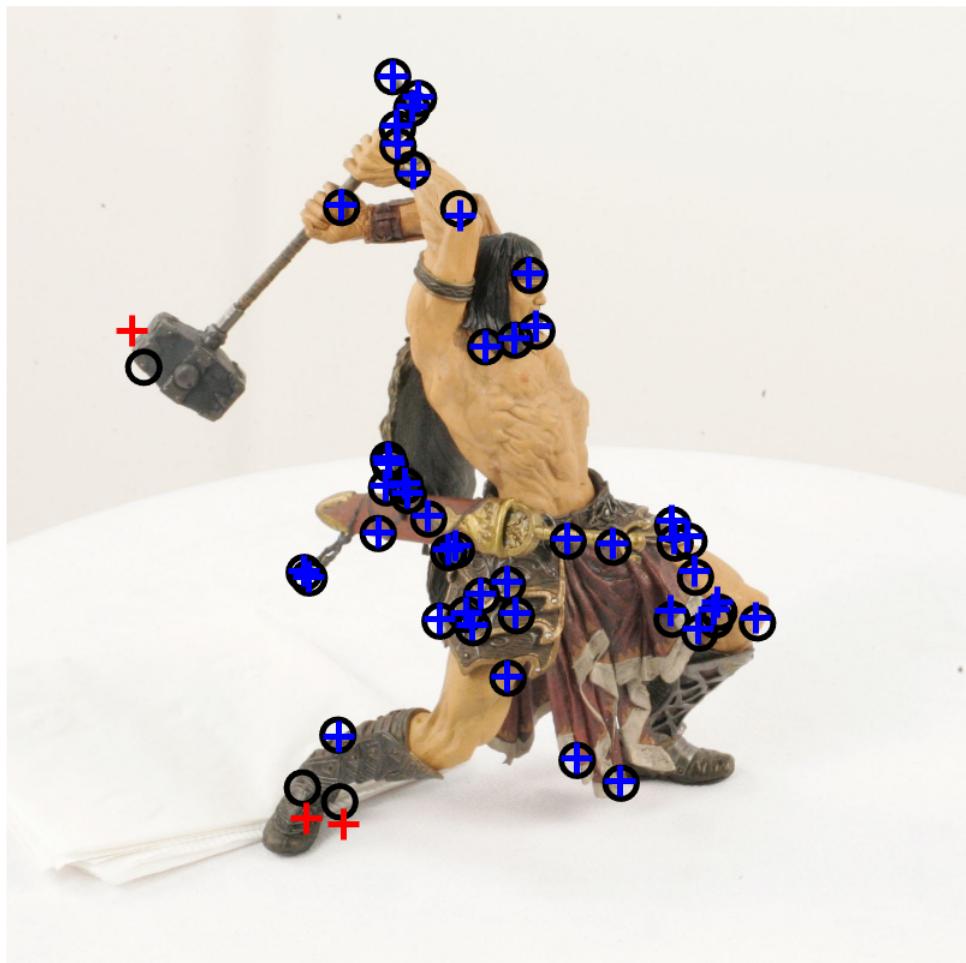


Figure 17: Result of Reprojection: Warrior

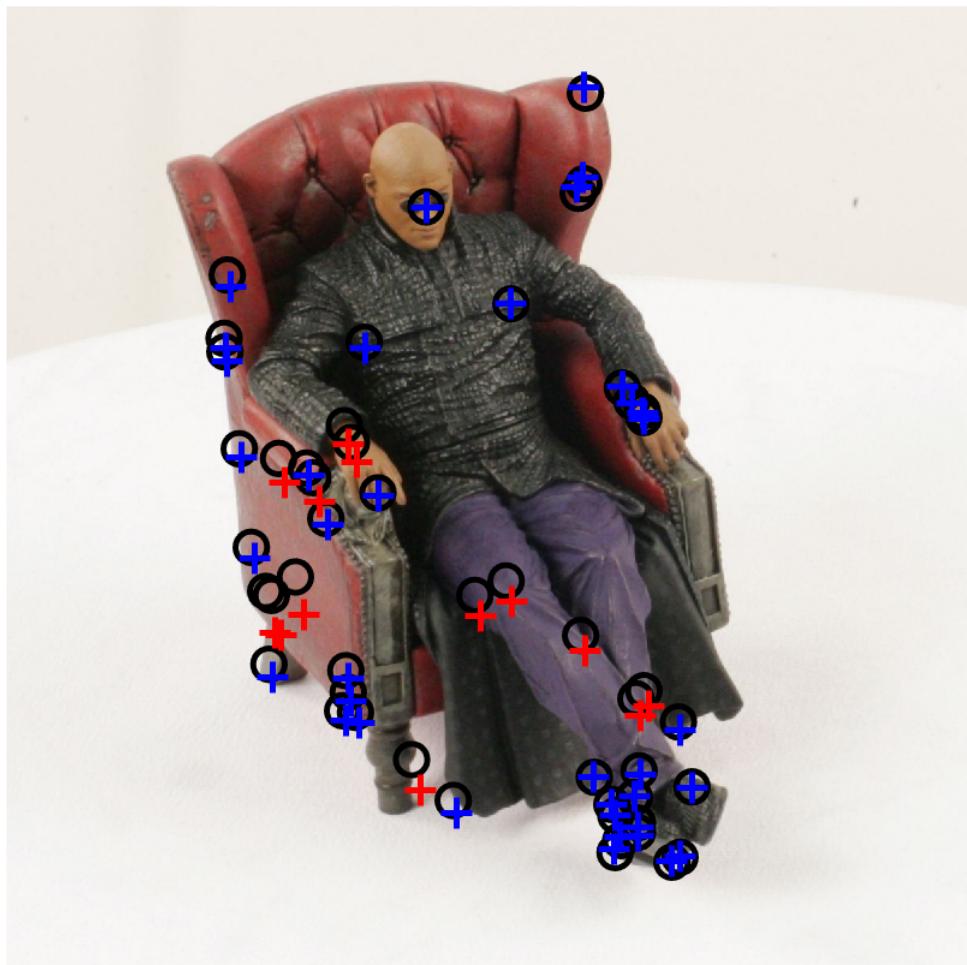


Figure 18: Result of Reprojection: Matrix

Appendix

Listing 1: Script of Problem 4

```
1 clear
2 clc
3 %% Problem 4.1 Warmup
4 filter_img = loadAndPreprocess('filter.jpg');
5
6 toy_ori_img = imread('toy.png');
7 toy_img = loadAndPreprocess('toy.png');
8 toy_conv = conv2(toy_img, filter_img, 'same');
9 figure;
10 imagesc(toy_conv);
11
12 % find maximum intensity
13 max_intensity = max(toy_conv(:));
14
15 max_index = zeros(1,2,3);
16 count = 1;
17 for i = 1:size(toy_ori_img, 1)
18     for j = 1:size(toy_ori_img, 2)
19         if toy_conv(i,j) == max_intensity
20             max_index(1,1,count) = i;
21             max_index(1,2,count) = j;
22             count = count + 1;
23         end
24     end
25 end
26 len = size(filter_img, 1);
27 width = size(filter_img, 2);
28 figure;
29 imshow(toy_ori_img);
30 hold on;
31 for i = 1:3
32     y = max_index(1,1,i)-len/2;
33     x = max_index(1,2,i)-width/2;
34     rectangle('Position',[x y width len], 'EdgeColor', 'r');
35 end
36 %% Problem 4.2 Detection Error
37 figure;
38 imshow(toy_ori_img);
39 hold on;
40 ground_truth_box_1 = [322,213,105,105];
41 ground_truth_box_2 = [325,215,110,110];
42 ground_truth_box_3 = [320,220,120,120];
43 rect_1 = [x y width len];
```

```

44 rectangle( 'Position' , rect_1 , 'EdgeColor' , 'r' );
45 rectangle( 'Position' , ground_truth_box_3 , 'EdgeColor' , 'r' );
46 overlap = detectError( rect_1 , ground_truth_box_3 );
47 %% Problem 4.3 More Realistic Images
48 car_template_img = loadAndPreprocess( 'cartemplate.jpg' );
49
50 % car1
51 car1_ori_img = imread( 'car1.jpg' );
52 car1_img = loadAndPreprocess( 'car1.jpg' );
53 car1_in = fopen( 'car1.txt' );
54 car1_rect = textscan( car1_in , '%d.%d' );
55 box_len = car1_rect{1}(2) - car1_rect{1}(1);
56 box_width = car1_rect{2}(2) - car1_rect{2}(1);
57 car1_box = [ car1_rect{1}(1) car1_rect{2}(1) box_len box_width ];
58 % car1 filter preprocessing
59 car1_template = imcrop( car_template_img , [90 280 1630 620] );
60 car1_template = imresize( car1_template , [box_width box_len] );
61 % car1 overlap finding process
62 heat_map1 = constructHeatMap( car1_img , car1_template );
63 rect_car1 = constructBoundingBox( car1_ori_img , heat_map1 , car1_box );
64 % rectangle( 'Position' , car1_box , 'EdgeColor' , 'b' );
65 overlap1 = detectError( rect_car1 , car1_box );
66 % car2
67 car2_ori_img = imread( 'car2.jpg' );
68 car2_img = loadAndPreprocess( 'car2.jpg' );
69 car2_in = fopen( 'car2.txt' );
70 car2_rect = textscan( car2_in , '%d.%d' );
71 box_len = car2_rect{1}(2) - car2_rect{1}(1);
72 box_width = car2_rect{2}(2) - car2_rect{2}(1);
73 car2_box = [ car2_rect{1}(1) car2_rect{2}(1) box_len box_width ];
74 % car2 filter preprocessing
75 car2_template = imcrop( car_template_img , [90 280 1630 620] );
76 car2_template = imresize( car2_template , [box_width box_len] );
77 car2_template = flip( car2_template , 2 );
78 % car1 overlap finding process
79 heat_map2 = constructHeatMap( car2_img , car2_template );
80 rect_car2 = constructBoundingBox( car2_ori_img , heat_map2 , car2_box );
81 % rectangle( 'Position' , car2_box , 'EdgeColor' , 'b' );
82 overlap2 = detectError( rect_car2 , car2_box );
83 % car3
84 car3_ori_img = imread( 'car3.jpg' );
85 car3_img = loadAndPreprocess( 'car3.jpg' );
86 car3_in = fopen( 'car3.txt' );
87 car3_rect = textscan( car3_in , '%d.%d' );
88 box_len = car3_rect{1}(2) - car3_rect{1}(1);
89 box_width = car3_rect{2}(2) - car3_rect{2}(1);

```

```

90 car3_box = [car3_rect{1}(1) car3_rect{2}(1) box_len box_width];
91 % car3 filter preprocess
92 car3_template = imcrop(car_template_img, [90 260 1630 640]);
93 car3_template = imresize(car3_template, [68 95]);
94 % car3 overlap finding process
95 heat_map3 = constructHeatMap(car3_img, car3_template);
96 rect_car3 = constructBoundingBox(car3_ori_img, heat_map3, car3_box);
97 % rectangle('Position', car3_box, 'EdgeColor', 'b');
98 overlap3 = detectError(rect_car3, car3_box);
99 %% car4
100 car4_ori_img = imread('car4.jpg');
101 car4_img = loadAndPreprocess('car4.jpg');
102 car4_in = fopen('car4.txt');
103 car4_rect = textscan(car4_in, '%d.%d');
104 box_len = car4_rect{1}(2) - car4_rect{1}(1);
105 box_width = car4_rect{2}(2) - car4_rect{2}(1);
106 car4_box = [car4_rect{1}(1) car4_rect{2}(1) box_len box_width];
107 % car4 filter preprocess
108 car4_template = imcrop(car_template_img, [90 280 1630 620]);
109 car4_template = car4_template + 0.1;
110 car4_template = imresize(car4_template, [box_width box_len]);
111 car4_template = flip(car4_template, 2);
112 % car4 overlap finding process
113 heat_map4 = constructHeatMap(car4_img, car4_template);
114 rect_car4 = constructBoundingBox(car4_ori_img, heat_map4, car4_box);
115 % rectangle('Position', car4_box, 'EdgeColor', 'b');
116 overlap4 = detectError(rect_car4, car4_box);
117 %% car5
118 car5_ori_img = imread('car5.jpg');
119 car5_img = loadAndPreprocess('car5.jpg');
120 car5_in = fopen('car5.txt');
121 car5_rect = textscan(car5_in, '%d.%d');
122 box_len = car5_rect{1}(2) - car5_rect{1}(1);
123 box_width = car5_rect{2}(2) - car5_rect{2}(1);
124 car5_box = [car5_rect{1}(1) car5_rect{2}(1) box_len box_width];
125 % car5 filter preprocess
126 car5_template = imcrop(car_template_img, [90 280 1630 620]);
127 car5_template = imresize(car5_template, [box_width box_len]);
128 car5_template = car5_template + 0.1;
129 car5_template = flip(car5_template, 2);
130 % car5 overlap finding process
131 heat_map5 = constructHeatMap(car5_img, car5_template);
132 rect_car5 = constructBoundingBox(car5_ori_img, heat_map5, car5_box);
133 % rectangle('Position', car5_box, 'EdgeColor', 'b');
134 overlap5 = detectError(rect_car5, car5_box);

```

Listing 2: loadAndPreprocess.m used in Script of Problem 4

```
1 function [ output_img ] = loadAndPreprocess( input_img )
2     output_img = imread(input_img);
3     output_img = im2double(output_img);
4     output_img = output_img - mean(output_img (:));
5 end
```

Listing 3: detectError.m used in Script of Problem 4

```
1 function [ overlap ] = detectError( predicted_box , ground_truth_box )
2     intersection_area = rectint(predicted_box , ground_truth_box );
3     union_area = predicted_box(3)*predicted_box(4) + ...
4         ground_truth_box(3)*ground_truth_box(4) - intersection_area ;
5     overlap = double(intersection_area)/double(union_area);
6 end
```

Listing 4: constructHeatMap.m used in Script of Problem 4

```
1 function [ heat_map ] = constructHeatMap( input_img , filter_img )
2     filter_img = flip(filter_img , 1);
3     filter_img = flip(filter_img , 2);
4     heat_map = conv2(input_img , filter_img , 'same');
5     figure;
6     imagedc(heat_map);
7 end
```

Listing 5: constructBoundingBox.m used in Script of Problem 4

```
1 function [ rect ] = constructBoundingBox( ori_img , heat_map , ...
2     ground_truth_box )
3     % find maximum intensity
4     max_intensity = 0;
5     max_index = zeros(1,2);
6     for i = 1:size(ori_img , 1)
7         for j = 1:size(ori_img , 2)
8             max_intensity = max(heat_map(i,j) , max_intensity);
9         end
10    end
11    for i = 1:size(ori_img , 1)
12        for j = 1:size(ori_img , 2)
13            if heat_map(i,j) == max_intensity
14                max_index(1,1) = i;
15                max_index(1,2) = j;
16            end
17        end
18    end
19    len = ground_truth_box(4);
```

```

20     width = ground_truth_box(3);
21 %     len = size(filter_img , 1);
22 %     width = size(filter_img , 2);
23 figure;
24 imshow(ori_img);
25 hold on;
26 y = max_index(1,1)-len/2;
27 x = max_index(1,2)-width/2;
28 rectangle('Position',[x y width len], 'EdgeColor','b', 'LineWidth', 3);
29 rect = [x y width len];
30 end

```

Listing 6: Script of Problem 5

```

1 clc;
2 %% Problem 5.1 Smoothing
3 geisel = imread('geisel.jpeg');
4 geisel = im2double(geisel);
5 figure;
6 imshow(geisel);
7 title('Input_Image');
8 K = 1/159*[2,4,5,4,2;4,9,12,9,4;5,12,15,12,5;4,9,12,9,4;2,4,5,4,2];
9 gaussian = conv2(geisel, K, 'same');
10 figure;
11 imshow(gaussian);
12 title('Gaussian_Filter_with_sigma=1.4');
13 %% Problem 5.2 Gradient Computation
14 sx = [-1,0,1;-2,0,2;-1,0,1];
15 sy = [-1,-2,-1;0,0,0;1,2,1];
16 Gx = conv2(gaussian, sx, 'same');
17 Gy = conv2(gaussian, sy, 'same');
18 for i = 1:length(Gx)
19     for j = 1:length(Gy)
20         G(i,j) = sqrt(Gx(i,j)^2 + Gy(i,j)^2);
21         edge_direction(i,j) = atan2(Gy(i,j), Gx(i,j));
22         edge_direction(i,j) = rad2deg(edge_direction(i,j));
23         if (edge_direction(i,j) < 0)
24             edge_direction(i,j) = edge_direction(i,j) + 180;
25         end
26     end
27 end
28 figure;
29 imshow(G);
30 title('Gradient_Magnitude');
31 %% Problem 5.3 Non Maximum Suppression
32 edge_candidate = zeros(size(G));
33 for i = 2:length(G)-1

```

```

34 for j = 2:length(G)-1
35     round_direction(i,j) = round2nearest45(edge_direction(i,j));
36     switch(round_direction(i,j))
37         case 0 % east and west directions
38             if G(i,j) > G(i,j-1) && G(i,j) > G(i,j+1)
39                 edge_candidate(i,j) = G(i,j);
40             end
41         case 45 % north east and south west directions
42             if G(i,j) > G(i-1,j+1) && G(i,j) > G(i+1,j-1)
43                 edge_candidate(i,j) = G(i,j);
44             end
45         case 90 % north and south directions
46             if G(i,j) > G(i-1,j) && G(i,j) > G(i+1,j)
47                 edge_candidate(i,j) = G(i,j);
48             end
49         case 135 % north west and south east directions
50             if G(i,j) > G(i-1,j-1) && G(i,j) > G(i+1,j+1)
51                 edge_candidate(i,j) = G(i,j);
52             end
53     end
54 end
55 end
56 figure;
57 imshow(edge_candidate);
58 title('Non-Maximum-Suppression');
59 %% Problem 5.4 Hysteresis Thresholding
60 high = 0.4;
61 low = 0.1;
62 A = [];
63 for i = 2:length(edge_candidate)-1
64     for j = 2:length(edge_candidate)-1
65         if edge_candidate(i,j) > high
66             A = [A; i j];
67         end
68     end
69 end
70 result = zeros(size(edge_candidate));
71 while ~isempty(A)
72     cur = A(size(A,1),:);
73     A(size(A,1),:) = [];
74     result(cur(1),cur(2)) = 1;
75     if cur(1)>1 && cur(1)<length(edge_candidate)-1 && cur(2)>1 ...
76         && cur(2)<length(edge_candidate)-1
77         if edge_candidate(cur(1)-1,cur(2)-1) > low ...
78             && result(cur(1)-1,cur(2)-1) == 0
79         A = [A; cur(1)-1 cur(2)-1];

```

```

80      elseif edge_candidate(cur(1)-1,cur(2)) > low ...
81          && result(cur(1)-1,cur(2)) == 0
82          A = [A; cur(1)-1 cur(2)];
83      elseif edge_candidate(cur(1)-1,cur(2)+1) > low ...
84          && result(cur(1)-1,cur(2)+1) == 0
85          A = [A; cur(1)-1 cur(2)+1];
86      elseif edge_candidate(cur(1),cur(2)-1) > low ...
87          && result(cur(1),cur(2)-1) == 0
88          A = [A; cur(1) cur(2)-1];
89      elseif edge_candidate(cur(1),cur(2)+1) > low ...
90          && result(cur(1),cur(2)+1) == 0
91          A = [A; cur(1) cur(2)+1];
92      elseif edge_candidate(cur(1)+1,cur(2)-1) > low ...
93          && result(cur(1)+1,cur(2)-1) == 0
94          A = [A; cur(1)+1 cur(2)-1];
95      elseif edge_candidate(cur(1)+1,cur(2)) > low ...
96          && result(cur(1)+1,cur(2)) == 0
97          A = [A; cur(1)+1 cur(2)];
98      elseif edge_candidate(cur(1)+1,cur(2)+1) > low ...
99          && result(cur(1)+1,cur(2)+1) == 0
100         A = [A; cur(1)+1 cur(2)+1];
101     end
102 end
103 figure;
104 imshow(result);
105 title('Hysteresis_Thresholding');

```

Listing 7: round2nearest45.m used in Script of Problem 5

```

1 function [ degree_out ] = round2nearest45( degree_in )
2     if 0 < degree_in && degree_in < 22.5
3         degree_out = 0;
4     elseif 22.5 <= degree_in && degree_in < 67.5
5         degree_out = 45;
6     elseif 67.5 <= degree_in && degree_in < 112.5
7         degree_out = 90;
8     elseif 112.5 <= degree_in && degree_in < 157.5
9         degree_out = 135;
10    else
11        degree_out = 0;
12    end
13 end

```

Submitted by Wei-Yuan Wen, Yen-Ting Chen, Shun-Jen Lee on November 2, 2016.