Digital Image Processing
School of Information Science and Technology
ShanghaiTech University

Assignment 1

Due time: 23:59, March 15th, 2023

---

1 Notes

This homework has **100 points** in total.

Please submit your homework to blackboard with a zip file named as **DIP2023_ID_Name_hw1.zip**. The zip file should contain three things: **a folder named 'codes' storing your codes**, **a folder named 'images' storing the original images**, and **your report named as report_ID_Name_hw1.pdf**. The names of your codes should look like **'p1a.m'** (for (a) part of Problem 1), so that we can easily match your answer to the question. Make sure all paths in your codes are relative path and we can get the result directly after running the code. Please answer in English.

Please complete all the coding assignments using MATLAB. All core codes are required to be implemented by yourself (without using relevant built-in functions). Make sure your results in the report are the same with the results of your codes. Please explain with notes at least at the key steps of your code.

2 Policy on plagiarism

This is an individual homework. You can discuss the ideas and algorithms, but you can neither read, modify, and submit the codes of other students, nor allow other students to read, modify, and submit your codes. Do not directly copy ready-made or automatically generated codes, or your score will be seriously affected. We will check plagiarism using automated tools and any violations will result in a zero score for this assignment.

3 Problem sets

Problem 1 (30 pts)

$$\begin{bmatrix} 6 & 1 & 2 & 1 & (2) \\ 2 & 3 & 5 & 3 & 8 \\ 1 & 0 & 1 & 2 & 3 \\ 3 & 2 & 4 & 5 & 2 \\ (1) & 5 & 3 & 4 & 0 \end{bmatrix}$$

(a) Calculate the D_4, D_8, D_m distance between the pair of points(framed in parentheses) in the matrix above, mark the shortest 4-,8-,m-path respectively. Show the results in your report. (V=1,2,3) (10 pts)

(b) An affine transformation of coordinates is given by

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = A \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Please perform the following transformation on **jetplane.tif**, write out the affine transformation matrix, then show the results:

(1) Scaling and shearing. Set x-scaling factor to 2 and y-scaling factor to 3, then shear the image horizontally 4 units. (5 pts)

(2) Translation and Rotation. Move 2 units to the left and move down 5 units, then rotate 45 degrees clockwise (about the origin). (5 pts)

(Hint: The main focus of this problem is to solve the affine transformation matrix, and using built-in functions like imwarp() to obtain the transformed image is allowed.)

(c) Perform bit-plane slicing on **lena_gray_256.tif** to get bit plane 1-8 and show the results in your report. Which kind of bit-plane usually contains more effective information, low bit or high bit? Why? (10 pts)

Solution:

(a) **4-path:** The path is shown in the following matrix with red marks:

$$\begin{bmatrix} 6 & 1 & 2 & 1 & 2 \\ 2 & 3 & 5 & 3 & 8 \\ 1 & 0 & 1 & 2 & 3 \\ 3 & 2 & 4 & 5 & 2 \\ 1 & 5 & 3 & 4 & 0 \end{bmatrix}$$

So the D_4 distance is 8.

8-path: The path is shown in the following matrix with red marks:

$$\begin{bmatrix} 6 & 1 & 2 & 1 & 1 \\ 2 & 3 & 5 & 3 & 8 \\ 1 & 0 & 1 & 2 & 3 \\ 3 & 2 & 4 & 5 & 2 \\ 1 & 5 & 3 & 4 & 0 \end{bmatrix}$$

So the D_8 distance is 4.

m-path: The path is shown in the following matrix with red marks:

$$\begin{bmatrix} 6 & 1 & 2 & 1 & 1 \\ 2 & 3 & 5 & 3 & 8 \\ 1 & 0 & 1 & 2 & 3 \\ 3 & 2 & 4 & 5 & 2 \\ 1 & 5 & 3 & 4 & 0 \end{bmatrix}$$

So the D_m distance is 8.

(b) (1) The coordinate of point (x, y) after scaling with x-scaling factor 2 and y-scaling factor 3 is $(2x, 3y)$, then the scaling matrix is $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Moreover, the coordinate of point (x, y) after shearing in y-axis direction with 4 units is $(x, 4x + y)$, then the shearing matrix is $\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Therefore, the whole affine transform matrix is $\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 8 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.
The processed result is shown in Figure 1.

(2) The coordinate of point (x, y) after translation is $(x+5, y-2)$, so the translation matrix is $\begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$. As for rotation, the coordinate of point (x, y) after rotation 45° clockwise is $(x \cos 45° + y \sin 45°, y \cos 45° - x \sin 45°)$, so the rotation matrix is $\begin{bmatrix} \cos 45° & \sin 45° & 0 \\ -\sin 45° & \cos 45° & 0 \\ 0 & 0 & 1 \end{bmatrix}$ Therefore, the whole affine transform matrix is $\begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 5 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & -2 \\ 0 & 0 & 1 \end{bmatrix}$. The processed result is shown in Figure 2.

(c) The result is shown in Figure 3. From the result, it is obvious that the high bits have more effective information since those high bits slices contain more semantic and structural information while those low bits slices contain almost noise.

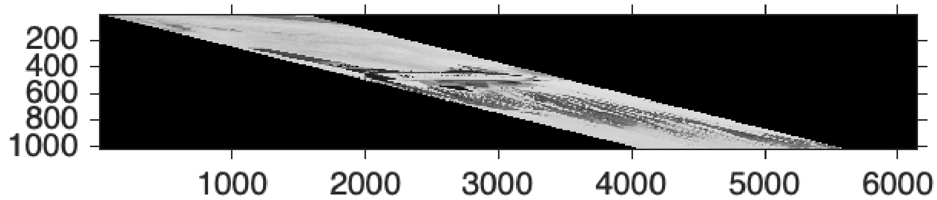


Figure 1: Image after scaling and shearing

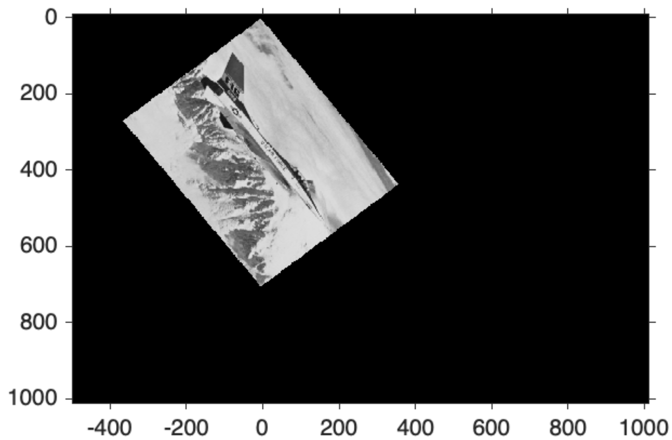


Figure 2: Image after translation and rotation

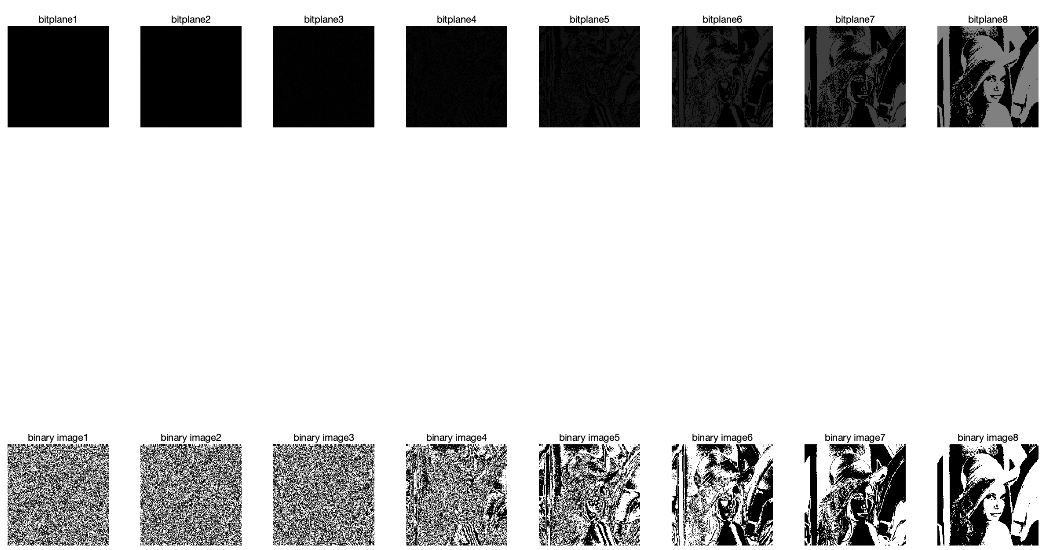


Figure 3: Bitplanes slicing results. The top line shows the bitplanes with original values 0 or 2^i and the bottom line shows the binary map with rescaled values 0 or 255.

Problem 2 (30 pts)

(a) Compute the histogram of **einstein_low_contrast.tif**. Perform histogram equalization (HE) on it, show the result. Show the histogram of the processed image, too. Why can HE enhance the contrast? (10 pts)

(b) Match the histgram of **lena_color.tif** to **peppers_color.tif**. Show the result in your report. (8 pts)

(c) Perform contrast limited adptive histogram equalization(CLAHE) on **man_in_house.png**. Contrast the result to HE method. Show the results in your report. Explain why CLAHE has a better effect.(Reference: http://cas.xav.free.fr/Graphics%20Gems%204%20-%20Paul%20S.%20Heckbert.pdf, page474-485) (12 pts)

(Hint: If you can effectively eliminate the checkerboard effect and balance the efficiency of the algorithm, you will get a bonus. Built-in functions like hist(), histogram(), histeq() and adapthisteq() are not allowed.)

Solution:

(a) The histogram of input image is shown in Figure 4(a), the result after HE is shown in Figure 4(c) and the histogram of result is shown in Figure 4(b).

The reason why HE can enhance the contrast is that HE transforms the original image distribution into a nearly uniform distribution, which enlarges the intensity difference of area where its original contrast is low.

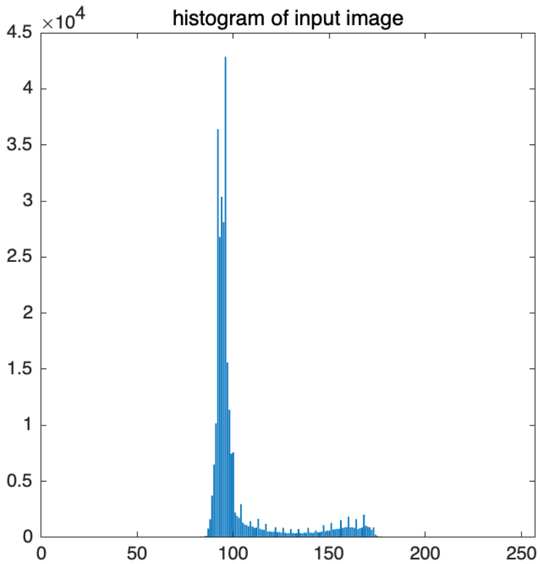
(b) The result after histogram matching is shown in Figure 5.

(c) The result of HE and CLAHE are shown in Figure 6(a) and 6(b). The detailed parameter settings of CLAHE are as follows:

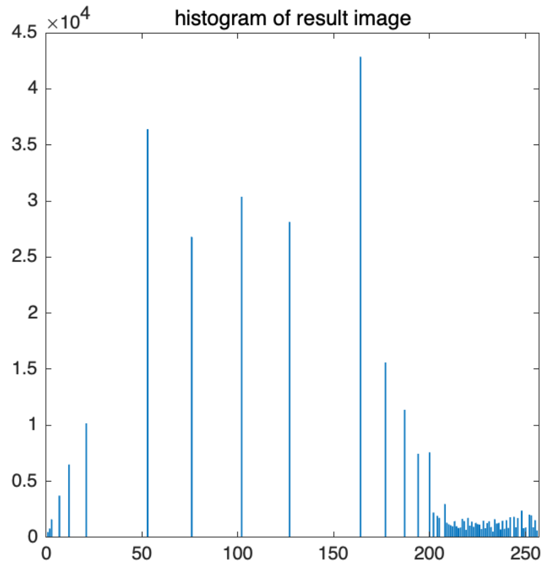
- padding (mirroring) size = 30 pixels
- updated patch length = 2 pixels
- clipping limit = 0.3 (for distribution, not histogram)

The reasons why CLAHE has a better effect lie in two aspects:

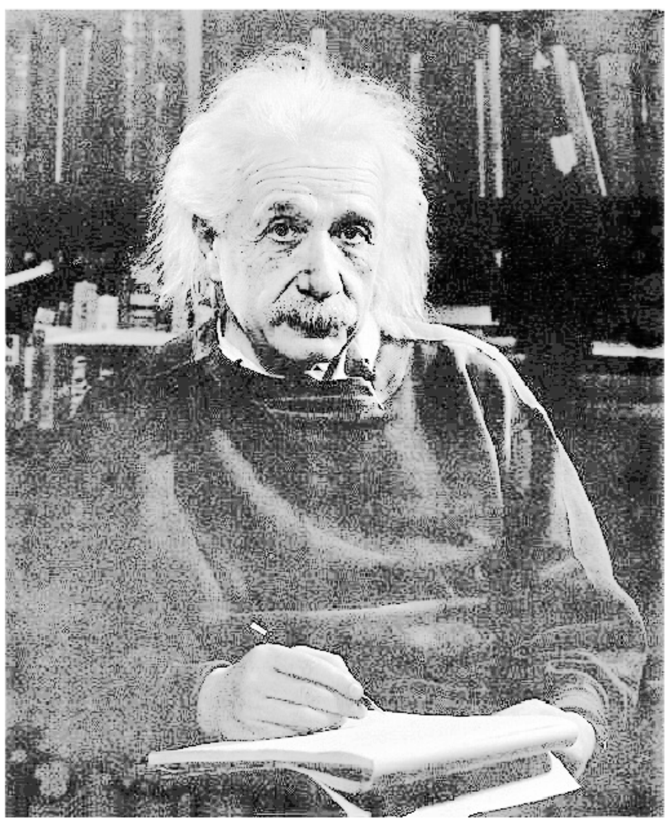
- CLAHE only cares about the intensity distribution under the moving window, those local distributions prevent pixels to be updated from being affected by other dense-intensity areas which are far away from them, so the areas where the contrast is low can get higher contrast.
- CLAHE limits the slop of histogram, this property prohibits too dense distribution after HE, which reduces artifacts like shining.



(a) Histogram of input imag



(b) Histogram of result image



(c) Result after HE

Figure 4

Figure 5: Histogram matching result

(a) HE result

(b) CLAHE result

Figure 6

Problem 3 (40 pts)

(a) Implement full convolution and cropped convolution of the given matrices, show the results in your report. (8 pts)

$$\begin{bmatrix} 6 & 4 & -1 & 0 & 1 \\ 1 & -3 & -4 & 3 & 2 \\ 0 & 3 & 5 & -2 & 1 \\ 9 & -1 & -3 & 4 & 5 \\ -2 & -5 & 2 & 3 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 4 \\ -3 & -1 & 1 \\ 5 & 2 & -1 \end{bmatrix}$$

(Hint: Pay attention to the difference between convolution and correlation.)

(b) Describe the noise type on **circuitboard-a.tif**, **circuitboard-b.tif**, **circuitboard-c.tif** and **circuitboard-d.tif**, then choose the best filter you think for each of them to reduce the noise. Show the results in your report. (12 pts)

(Hint: Choose the best kernel size you think.)

(c) Filter the image **house.tif** in both x direction and y direction with 3*3 Sobel mask and Laplacian mask. Show the results in your report. (8 pts)

(d) Perform image sharpening on **house.tif** using LoG filter(with the $\sigma^2 = 1$ and another two values you choose yourself) and unsharpen mask method(choose the best k you think). Show the results in your report. (12 pts)

Solution:

(a) The full convolution result is

$$\begin{bmatrix} 6 & 16 & 31 & 14 & -3 & 2 & 4 \\ -17 & -19 & -1 & -12 & -12 & 15 & 9 \\ 27 & 43 & 24 & 6 & 10 & -3 & 5 \\ 14 & -5 & -14 & 8 & 25 & 24 & 19 \\ -29 & 10 & 54 & 14 & -12 & 15 & 4 \\ 51 & 0 & -39 & 9 & 35 & 9 & -5 \\ -10 & 21 & 22 & 14 & 4 & -3 & 0 \end{bmatrix}$$

and the cropped convolution result is

$$\begin{bmatrix} -19 & -1 & -12 & -12 & 15 \\ 43 & 24 & 6 & 10 & -3 \\ -5 & -14 & 8 & 25 & 24 \\ 10 & 54 & 14 & -12 & 15 \\ 0 & -39 & 9 & 35 & 9 \end{bmatrix}$$

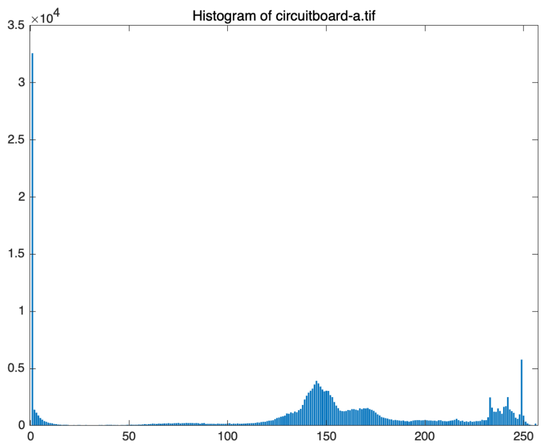
The detailed implementation is shown in the file: codes/p3a.mlx.

(b)
- circuitboard-a.tif: Peper noise (histogram shown in 7(a)) → max filter, result shown in 8(a).
- circuitboard-b.tif: Salt noise (histogram shown in 7(b)) → min filter, result shown in 8(b).
- circuitboard-c.tif: Salt-and-pepper noise (histogram shown in 7(c))→ median filter, result shown in 8(c).
- circuitboard-d.tif: Gaussian noise (histogram shown in 7(d))→ smooth filter, result shown in 8(d).

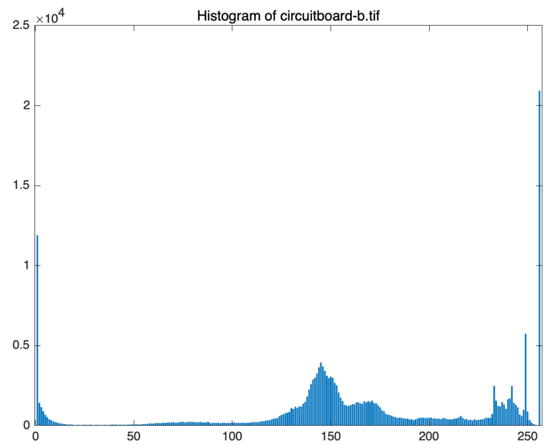
Remark: All the kernels have size 3×3.

(c) Results after applying Sobel mask and Laplacian mask are shown in 9(a), 9(b), 9(c) and 9(d) respectively.

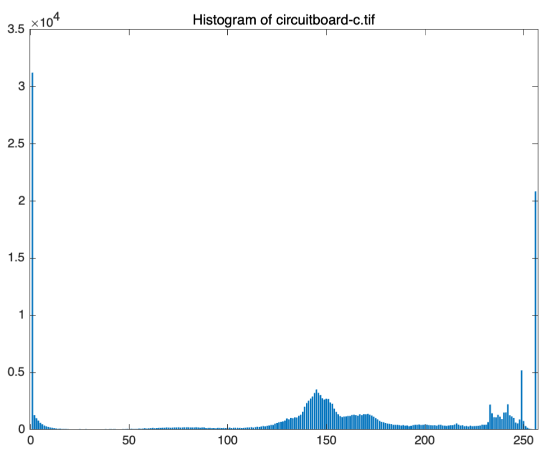
(d) Results after applying LoG filter and unsharpen mask are shown in 10(a) 10(b) 10(c) and 10(d) respectively.



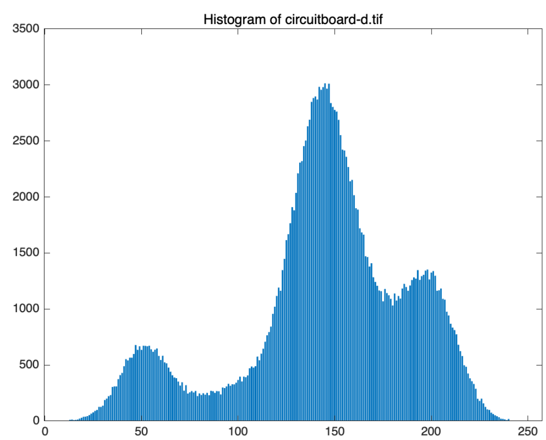
(a) Histogram of circuitboard-a.tif



(b) Histogram of circuitboard-b.tif

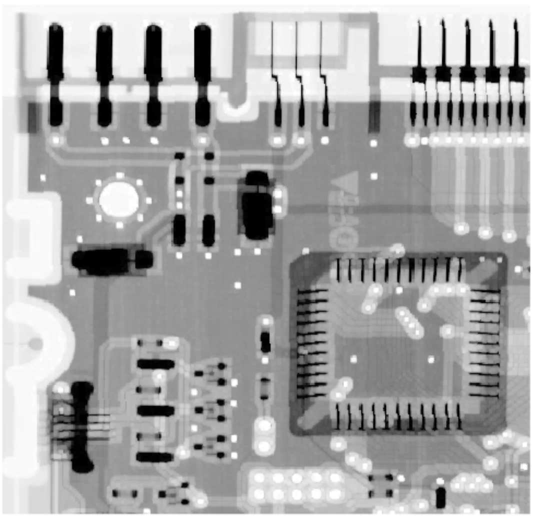


(c) Histogram of circuitboard-c.tif

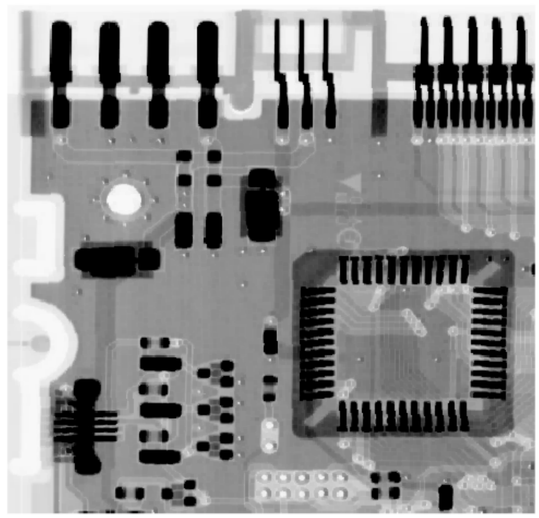


(d) Histogram of circuitboard-d.tif

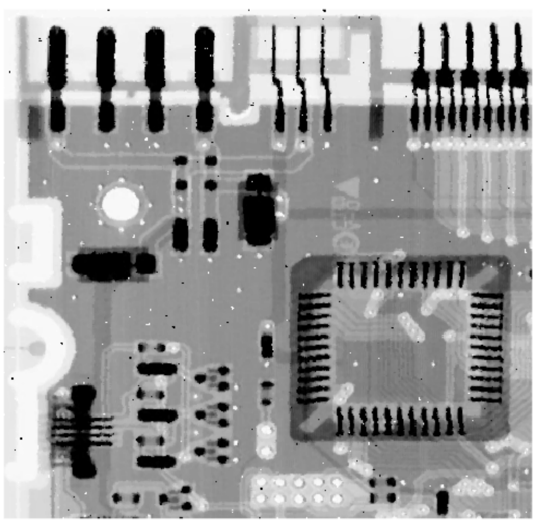
Figure 7



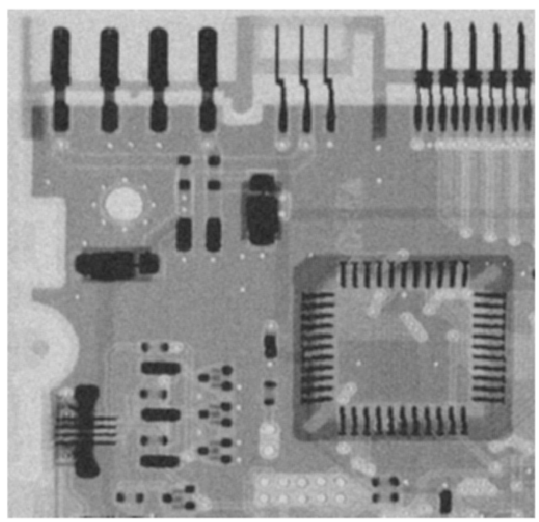
(a) denoised result after max filter



(b) denoised result after min filter

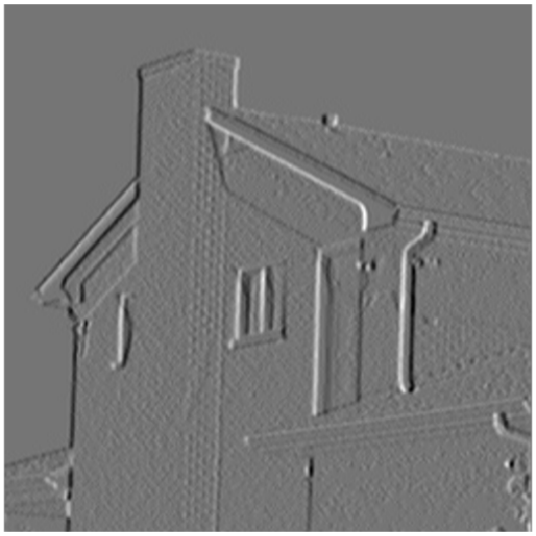


(c) denoised result after median filter

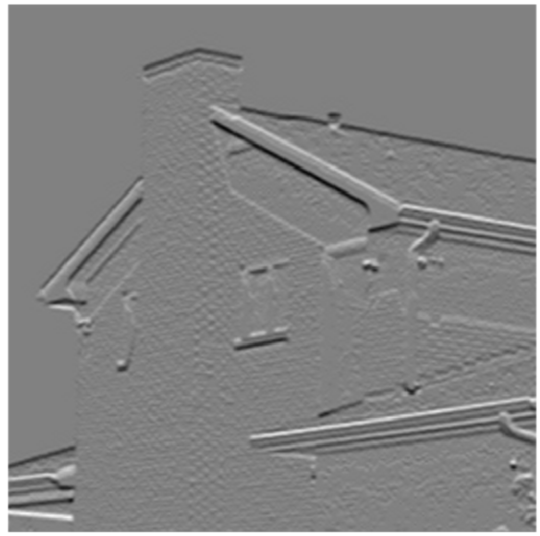


(d) denoised result after smooth filter

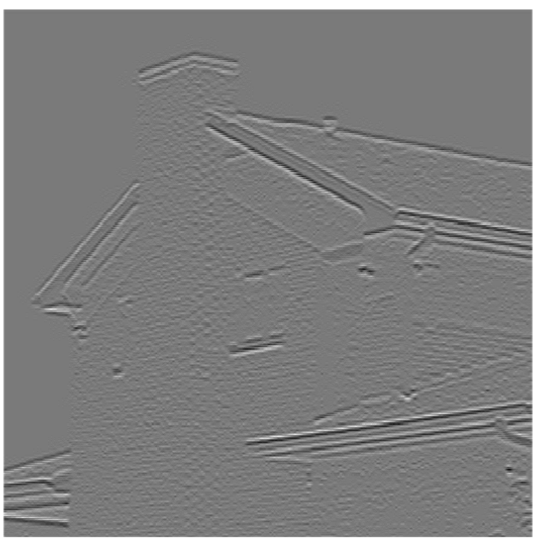
Figure 8



(a) Result after applying Sobel mask in x direction



(b) Result after applying Sobel mask in y direction



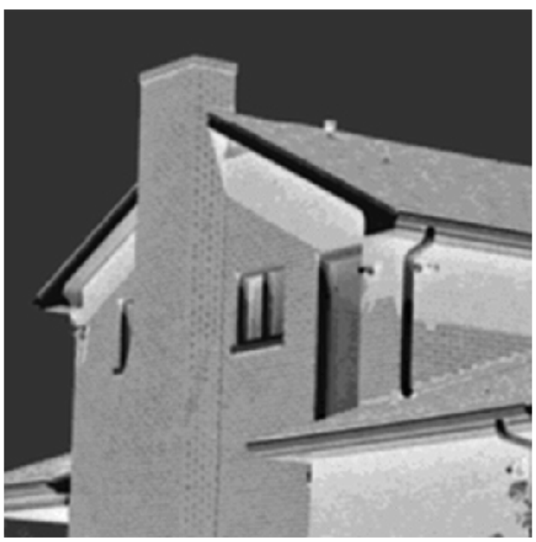
(c) Result after applying Laplacian mask in x direction

(d) Result after applying Sobel mask in y direction

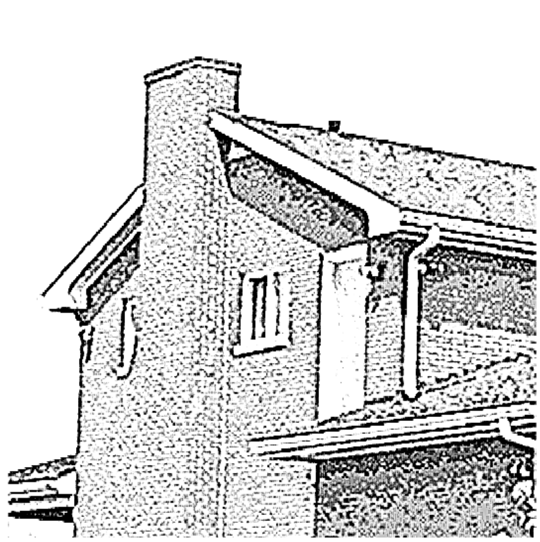
Figure 9: Those images are processed to map their original value interval to $[0, 255]$ for visualization purposes.

(a) Result after applying LoG filter with $\sigma = 1$

(b) Result after applying LoG filter with $\sigma = 0.5$



(c) Result after applying LoG filter with $\sigma = 0.25$



(d) Result after applying unsharpen mask

Figure 10: The Result of LoG filter is processed to fit intensity interval $[0, 255]$ for visualization purposes. The unsharpen mask uses Gaussian kernel to get the smooth image and sets $k = 100$