

Digital Image Processing 2017fall

Project Report

Name: Wang Yiqing
Student Number: 515 030 910 456
Email: WangYiqing_2015@sjtu.edu.cn

2017-12-10

Contents

1	Project 1 - Histogram Equalization	3
1.1	Project Proposal	3
1.2	Preliminaries	3
1.3	Histogram equalization	3
1.4	Discussion	6
1.5	Implementation	6
2	Project 2 - Spatial Enhancement Methods	7
2.1	Project Proposal	7
2.2	Preliminaries	7
2.2.1	Basic Spatial Filtering	7
2.3	Experiment - Combining Spatial Enhancement	8
2.4	Discussion	9
2.5	Implementation	9
3	Project 3 - Filtering in Frequency Domain	12
3.1	Project Proposal	12
3.2	Preliminaries	12
3.2.1	Summary of Fourier transform properties	12
3.2.2	Steps for filtering in the frequency domain	12
3.3	Image Smoothing Using Frequency Domain Filters	13
3.3.1	Ideal lowpass filters	13
4	Project 4 - Image Restoration	14
4.1	Project Proposal	14
4.2	Preliminaries	14
4.2.1	Noise Models	14
5	Project 6 - Geometric Transformation	15
5.1	Project Proposal	15
5.2	Preliminary	15
5.2.1	Spatial Transform	15
5.2.2	Interpolation	15
5.3	Experiment	16
5.4	Implementation	17
5.5	Discussion	19
6	Project 8 - Morphological Processing	20
6.1	Project Proposal	20
6.2	Preliminaries	20
6.2.1	Basic morphological operations	20
6.2.2	Morphological restoration	20
6.3	Task-1 Opening by reconstruction	21
6.4	Task-2 Hole filling	22
6.5	Task-3 Border clearing	22
6.6	Implementation	23
6.7	Discussion	24
7	Project 9 - Image Segmentation	25
7.1	Project Proposal	25
7.2	Preliminaries	25
7.2.1	Edge detection	25
7.2.2	Basic edge detection	25
7.2.3	More advanced techniques for edge detection	25

1 Project 1 - Histogram Equalization

1.1 Project Proposal

In this project we implement histogram equalization. First, plot the histogram of an image, then implement histogram equalization and display the equalized image and histogram. Test images are *Fig1.jpg*, *Fig2.jpg*.

1.2 Preliminaries

Let r denote the intensity of a pixel in the original image and $r \in [0, L - 1]$. We consider a transformation

$$s = T(r) \quad 0 \leq r \leq L - 1 \quad (1.1)$$

that produce an output intensity level s for every pixel in the input image having intensity r . We assume that:

(a) $T(r)$ is a monotonically increasing function in the interval $0 \leq r \leq L - 1$; and

(b) $0 \leq T(r) \leq L - 1$ for $r \in [0, L - 1]$.

Condition (a) and (b) guarantee the existing of inverse function $r = T^{-1}(s)$ which is monotonically increasing. The central idea is that *intensity levels of an image can be viewed as random variables in the interval $[0, L - 1]$ and can be described by probability density function (PDF)*. Let $p_r(r)$ and $p_s(s)$ to be the PDF of r and s respectively. Thus we can apply the basic result of probability theory that *if $T(r)$ is differentiable over the range of interest, we have this below relationship between $p_r(r)$ and $p_s(s)$:*

$$p_s(s) = p_r(r) \left| \frac{dr}{ds} \right| \quad (1.2)$$

Then consider the cumulative distribution function(CDF) which is exactly used here as the transformation function $T(r)$

$$s = T(r) = (L - 1) \int_0^r p_r(w) dw \quad (1.3)$$

Now we are ready to prove the transformation really does the histogram equalization. Firstly, compute the derivatives. The last = is from the assumption (a) stating the monotonically increasing.

$$\frac{ds}{dr} = \frac{dT(r)}{dr} = (L - 1)p_r(r) = \left| \frac{ds}{dr} \right| \quad (1.4)$$

Thus, substituting the result for Eq.1.2, we get the desired result

$$p_s(s) = p_r(r) \frac{1}{(L - 1)p_r(r)} = \frac{1}{L - 1} \quad (1.5)$$

which show $p_s(s)$ follows uniform distribution.

1.3 Histogram equalization

Histogram equalization on an image of size $M \times N$ is like a discrete version of the process in the preliminaries section.

$$p_r(r_k) = \frac{n_k}{MN} \quad k = 0, 1, 2, \dots, L - 1 \quad (1.6)$$

where n_k is the number of pixels that have intensity r_k . The discrete form of Eq.1.3 is

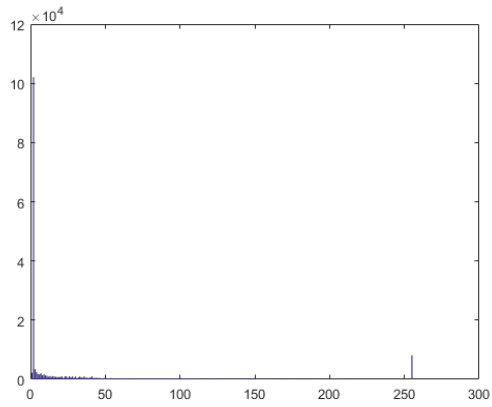
$$s_k = T(r_k) = (L - 1) \sum_{j=0}^k p_r(r_j) = \frac{L - 1}{MN} \sum_{j=0}^k n_k \quad k = 0, 1, 2, \dots, L - 1 \quad (1.7)$$

Based on these discrete form equation, I implement the matlab function and test them on *Fig1.jpg*, *Fig2.jpg* and get the results listed in 1.1 and 1.2 respectively.

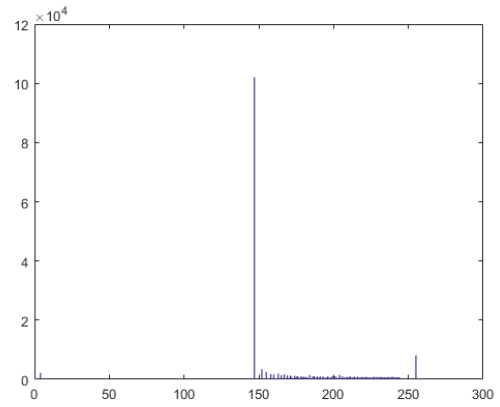


(a)

(b)

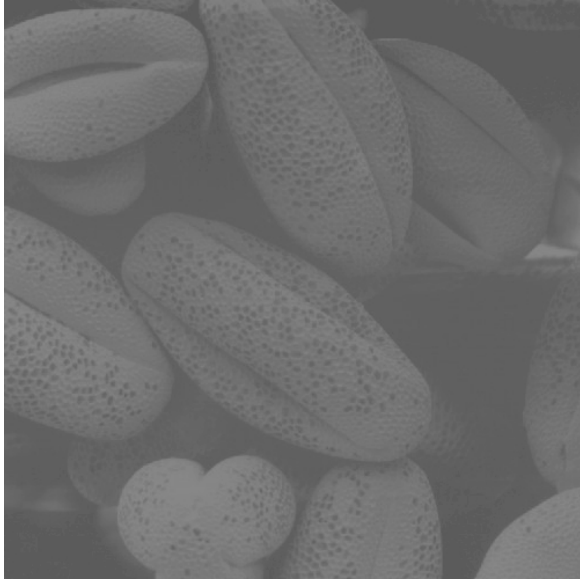


(c)

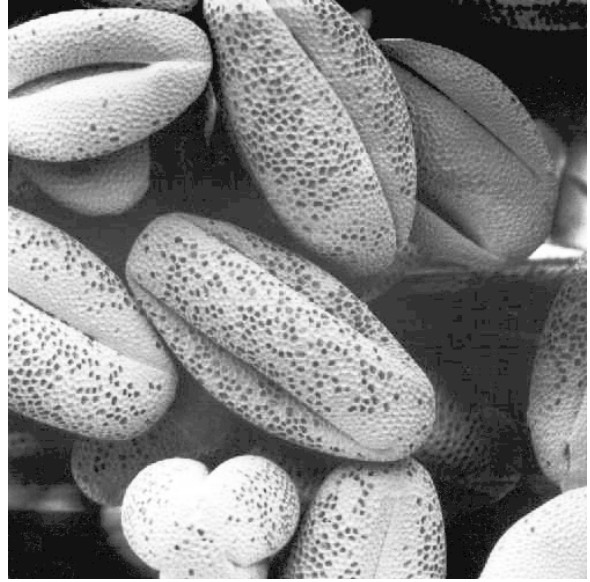


(d)

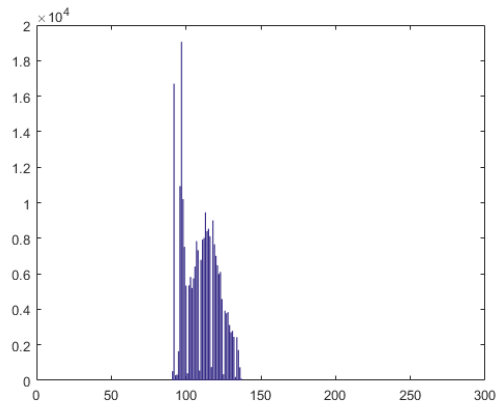
Figure 1.1: Results of Fig1.jpg. (a)Original image. (b)Processed image after applying histogram equalization. (c)Histogram of original image. (d)Histogram of processed image.



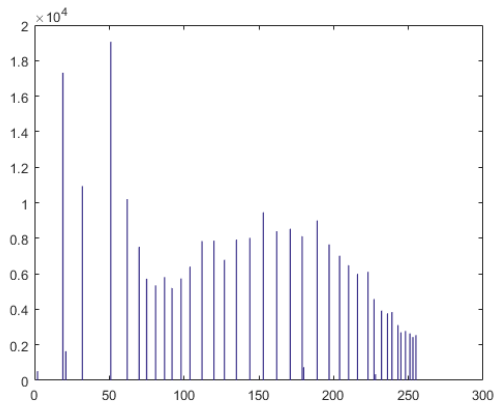
(a)



(b)



(c)



(d)

Figure 1.2: Results of Fig2.jpg. (a)Original image. (b)Processed image after applying histogram equalization. (c)Histogram of original image. (d)Histogram of processed image.

1.4 Discussion

We can see that histograms are quite different from continuous functions because there are gaps between each horizontal value. So the histogram of the equalized image is like a sparse version of the original image. The result of Fig2(low contrast ratio) shows an increasing on contrast ratio and it's more useful than the original result. However, Fig1(high contrast ratio) is not a good example of histogram equalization. The affect is just like a intensity transition. We can simply draw a not so serious conclusion that histogram equalization is useful for images with low contrast ratio but not for images with high contrast ratio.

1.5 Implementation

There is some key part of my implementation.

```
1 function [x, y] = histShow( imgf )
2 %HISTSHOW display the histogram graph of the imgf
3 % x - the horizontal axis of histogram ,
4 % y - the vertical axis of histogram
5 g = imgf(:) + 1;
6 n = length(g);
7 x = (1 : 256);
8 y = zeros(1, 256);
9 for i = (1 : n)
10     y(g(i)) = y(g(i)) + 1;
11 end
12
13 end
14
15 function [ imgg ] = histEqual( imgf )
16 %HISTEQUAL
17 %
18 [x, y] = histShow(imgf);
19 T = zeros(1, 256);
20 a = 0;
21 g = imgf(:);
22 n = length(g);
23 for i = (1 : 256)
24     T(i) = a + y(i);
25     a = T(i);
26 end
27 T = round(255 * T / n);
28 for i = (1 : n)
29     g(i) = T(g(i)+1);
30 end
31 imgg = reshape(g, size(imgf));
32
33 end
```

2 Project 2 - Spatial Enhancement Methods

2.1 Project Proposal

Implement the image spatial enhancement task showed in text book Figure 3.43. The steps involve Laplacian filter, Sobel filter, average filter and power law transformation.

2.2 Preliminaries

2.2.1 Basic Spatial Filtering

In general, linear spatial filtering of an image of size $M \times N$ with a filter of size $m \times n$ is

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t) \quad (2.1)$$

where $a = (m - 1)/2$ and $b = (n - 1)/2$. There are two closely related concepts related to spatial filtering. One is *correlation* and the other is *convolution*. Correlation of filter $w(x, y)$ is defined as

$$\sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t) \quad (2.2)$$

while convolution is defined as

$$\sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x - s, y - t) \quad (2.3)$$

The confusion point is that we usually call spatial filter as convolution filter, convolution mask or convolution kernel but the terms are not necessarily refer to the true convolution operation defined in Eq.2.3.

Smoothing spatial filters

Smoothing filters are used for blurring and for noise reduction. The simplest smoothing filter is averaging filters, which is also called lowpass filter. $M \times N$ averaging filter can be represented as

$$g(x, y) = \frac{\sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t)}{\sum_{s=-a}^a \sum_{t=-b}^b w(s, t)} \quad (2.4)$$

The simple idea of filter size choosing is that choose the approximately the same size of the noisy you want to reduce. Average filter is an example of linear spatial filters. There are also nonlinear order-static spatial filters using the ranking information of each pixel. We will talk about order-static filters in *Project 8*.

Sharpening spatial filters

The principal objective of sharpening is to highlight transitions in intensity. Because averaging is analogous to integration, it is logical to conclude that sharpening can be accomplished by spatial differentiation. Thus we consider the sharpening filters based on first- and second-order derivatives.

Second-order derivatives - the Laplacian

Laplacian is defined as

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (2.5)$$

In x-direction and y-direction we have

$$\frac{\partial^2 f}{\partial x^2} = f(x + 1, y) - f(x, y) - (f(x, y) - f(x - 1, y)) = f(x + 1, y) + f(x - 1, y) - 2f(x, y) \quad (2.6)$$

$$\frac{\partial^2 f}{\partial y^2} = f(x, y + 1) - f(x, y) - (f(x, y) - f(x, y - 1)) = f(x, y + 1) + f(x, y - 1) - 2f(x, y) \quad (2.7)$$

Thus we get the discrete Laplacian of two variables:

$$\nabla^2 f(x, y) = f(x + 1, y) + f(x - 1, y) + f(x, y + 1) + f(x, y - 1) - 4f(x, y) \quad (2.8)$$

Based on Eq.2.8 we have 4 masks displayed in Fig.2.1. The basic way we use Laplacian for image sharpening is

$$g(x, y) = f(x, y) + c [\nabla^2 f(x, y)] \quad (2.9)$$

0	1	0	1	1	1	0	-1	0	-1	-1	-1
1	-4	1	1	-8	1	-1	4	-1	-1	8	-1
0	1	0	1	1	1	0	-1	0	-1	-1	-1
(a)			(b)			(c)			(d)		

Figure 2.1: 4 Laplacian masks. **(a)**Direct implementation of Eq.2.8. **(b)**Extent (a) by adding two diagonal terms. **(c)(d)** sign inverse of (a)(b) relatively. More common in practice.

where $c = 1$ if we use Fig.2.1c or Fig.2.1d.

First-order derivatives - the gradient

First-order derivatives in image processing using the magnitude of the gradient. The gradient vector is

$$\nabla f \equiv \text{grad}(f) \equiv \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (2.10)$$

The magnitude of vector ∇f defined as

$$M(x, y) = \text{mag}(\nabla f) = \sqrt{g_x^2 + g_y^2} \approx |g_x| + |g_y| \quad (2.11)$$

is the value of the rate of change in the direction of the gradient vector. The second = is a frequently used approximate to avoid square roots. A widely used approximate filter masks implementing gradient is Sobel defined as

$$g_x = \frac{\partial f}{\partial x} = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3) \quad (2.12)$$

and

$$g_y = \frac{\partial f}{\partial y} = (z_3 + 2z_6 + z_9) - (z_1 + 4z_2 + z_7) \quad (2.13)$$

3×3 Sobel filter is displayed in Fig.2.2

-1	-2	-1	-1	0	1
0	0	0	-2	0	2
1	2	1	-1	0	1

Figure 2.2: 3×3 Sobel mask. Left is on x direction, right is on y direction.

2.3 Experiment - Combining Spatial Enhancement

We conduct the experiment on the Fig.2.3a image of whole body bone scan. Our goal is to enhance this image by sharpening it and by bringing out more of the skeletal detail. The narrow dynamic ranging of the intensity levels and high noise content make this problem difficult. The process and result images are displayed in Fig.2.3 and Fig.2.4. Here, I describe the detail process of the enhancement.

We first use Laplacian with the mask in Fig.2.1d to extract the sharp transitions in intensity. Note that Fig.2.3b is a scaled result because the numerical matrix result contains negative values which can not be visualized as image by matlab default function. After obtaining Laplacian, we add Laplacian(not the scaled one by the real laplacian) to the original image to get the sharpened Fig.2.3c. This step is just follow Eq.2.8. In Fig.2.3d we compute the Sobel gradient of original image using masks in Fig.2.2. Edges are more dominant than ones in Laplacian. In Fig.2.4e we use 5×5 averaging filter to smooth the Sobel gradient for noisy reduction purpose. Its not proper to use median filter because medical image processing requires high level of convince. In Fig.2.4f, we multiply the Laplacian with Sobel. Multiply seems like a strange operation but we can consider it as use

Sobel to mask Laplacian in order to strengthen edges and reduce noisy. In Fig.2.4g we sum up the product and the original image. Finally in Fig.2.4h we increase the dynamic range of the sharpened image by powerlaw transformation which is defined as

$$s = cr^\gamma \quad (2.14)$$

where we use $\gamma = 0.5$ and $c = 1$. We use $\gamma < 1$ to spread the intensity level and note that histogram equalization would not give use satisfying result as discussed in Project 1.

2.4 Discussion

We finally get a satisfy result in the Fig.2.4h which shows significant new visible features. There are two tricks we should pay attention to. First, the operations may produce intensity levels out of the range $[0, 255]$. In this kind of cases, a direct matlab command *imshow()* will cause problem. So we have to use scale function. While coding, I made a mistake that I didnt take the numerical type into consideration. I gain an experience that before computation, transform the matrix type to double at first.

2.5 Implementation

Here I paste the key parts of my matlab implementation.

```

1 % key part of function erosion
2 for i = (1:M-m+1)
3     for j = (1:N-n+1)
4         x = imgf(i:i+m-1, j:j+n-1);
5         if sum(sum(x.*B)) == m*n
6             g(i+downshift, j+rightshift) = 1;
7         end
8     end
9 end
10
11 % key part of dilation_reconstruction
12 k_times = 0;
13 while( ~isequal(f0, f1) )
14     k_times = k_times + 1;
15     f0 = f1;
16     f1 = geodesic_dilation(f1, G, B);
17 end
18
19 % key part of geodesic_dilation
20 imgg = dilation(imgf, B) & G;
21
22 % key part of opening_reconstruction
23 for i=(1:n_size)
24     f_erosion = erosion(f_erosion, B);
25 end
26 [imgg, k_times] = dilation_reconstruction(f_erosion, imgf, ones(3,3)); % here
    can not use ones(51, 1)

```

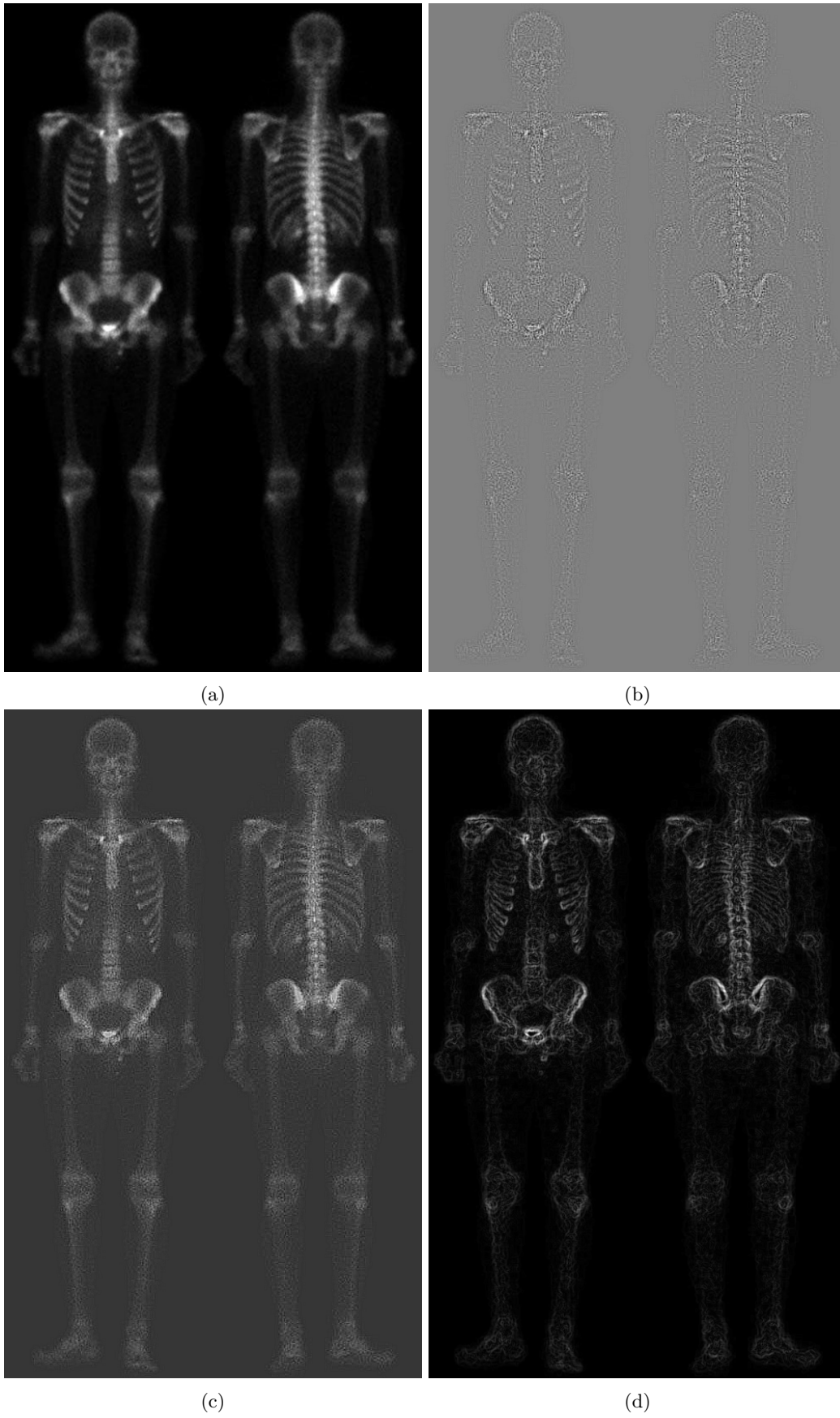


Figure 2.3: **(a)**Original image of whole body bone scan. **(b)**Laplacian of (a)(Rescaled to $[0,255]$). **(c)**Sharpened image obtained by add (a) and (c). **(d)**Sobel gradient of (a).

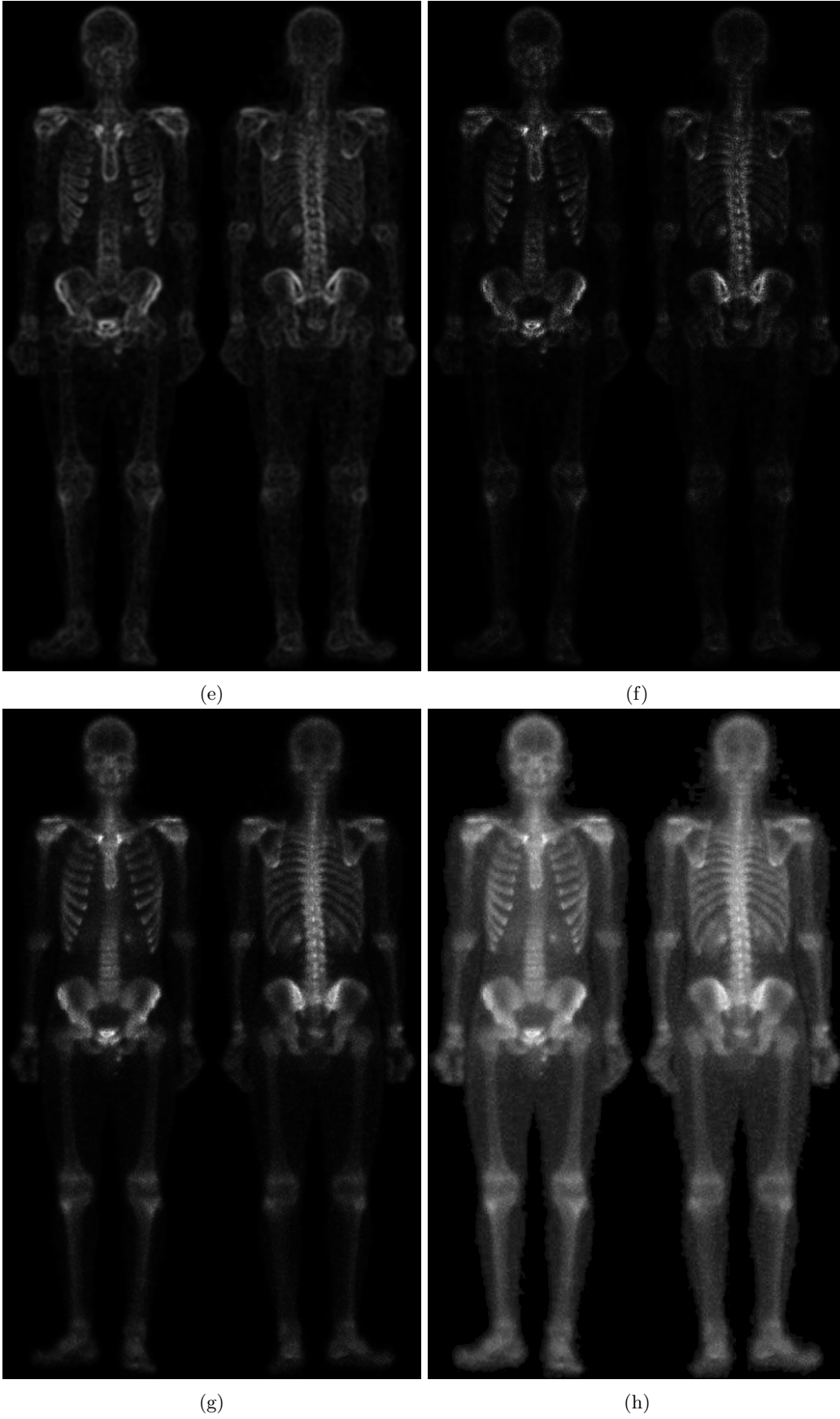


Figure 2.4: (e)Sobel image smoothed with 5×5 averaging filter. (f)Mask image formed by the product of (c) and (e). (g)Sharpened image obtained by applying a power-law transformation to (g). (h)Compare (g) and (h).

3 Project 3 - Filtering in Frequency Domain

3.1 Project Proposal

Implement the ideal, Butterworth and Gaussian lowpass and highpass filters and the results under different parameters using the image `character_test_pattern.tif`

3.2 Preliminaries

3.2.1 Summary of Fourier transform properties

Table 3.1: Summary of useful formulas.

Name	Expression(s)
1D FT	$F(u) = \int_{-\infty}^{\infty} f(t)e^{-j2\pi ut} dt \quad (3.1)$
1D IFT	$f(t) = \int_{-\infty}^{\infty} F(u)e^{j2\pi ut} du \quad (3.2)$
1D DFT	$F(u) = \sum_{t=0}^{M-1} f(t)e^{-j2\pi ut/M} \quad (3.3)$
1D IDFT	$f(t) = \frac{1}{M} \sum_{u=0}^{M-1} F(u)e^{j2\pi ut/M} \quad (3.4)$
2D FT	$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)e^{-j2\pi(ux+vy)} dx dy \quad (3.5)$
2D DFT	$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)e^{-j2\pi(ux/M+vy/N)} \quad (3.6)$
2D IDFT	$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v)e^{j2\pi(ux/M+vy/N)} \quad (3.7)$
Power spectrum	$P(u, v) = F(u, v) ^2 \quad (3.8)$

3.2.2 Steps for filtering in the frequency domain

1. Given an input image $f(x, y)$ of size $M \times N$, obtain the padding parameters $P = 2M$ and $Q = 2N$. Form a padded image, $f_p(x, y)$, of size $P \times Q$ by appending zeros.
2. Multiply $f_p(x, y)$ by $(-1)^{x+y}$ to center its transform.
3. Compute the DFT, $F(u, v)$ of the image from centered padded image.
4. Generate a real, symmetric filter function $H(u, v)$ of size $P \times Q$ with center at coordinates $(P/2, Q/2)$. Form the product $G(u, v) = H(u, v)F(u, v)$.
5. Obtain the preprocessed image: $g_p(x, y) = \{\text{real}[\mathcal{F}^{-1}[G(u, v)]]\}(-1)^{x+y}$ where the real part is selected in order to ignore parasitic complex components resulting from computational inaccuracies.
6. Obtain the final processed result, $g(x, y)$, by extracting the top left $M \times N$ quadrant of $g_p(x, y)$

3.3 Image Smoothing Using Frequency Domain Filters

Edges and other sharp intensity transitions such as noise in an image contribute significantly to the high-frequency content of its Fourier transform. Hence, smoothing is achieved in the frequency domain by high-frequency attenuation.

3.3.1 Ideal lowpass filters

Ideal lowpass filters (ILPF) is very sharp as it is specified by the function

$$H(u, v) = \begin{cases} 1 & \text{if } D(u, v) \leq D_0 \\ 0 & \text{otherwise} \end{cases} \quad (3.9)$$

where D_0 is a positive constant and $D(u, v)$ is the distance between (u, v) in frequency domain and the center of the frequency rectangle; that is

$$D(u, v) = [(u - P/2)^2 + (v - Q/2)^2]^{1/2} \quad (3.10)$$

One way to establish a set of standard cutoff frequency loci is to compute circles that enclose specified amounts of total image power P_T .

$$P_T = \sum_{u=0}^{P-1} \sum_{v=0}^{Q-1} P(u, v) \quad (3.11)$$

The percentage of power enclosed by the circle of radius D_0 with origin at the center of the frequency rectangle is

$$\alpha = 100 \left[\sum_u \sum_v P(u, v) / P_T \right] \quad (u, v) \text{ is inside the circle} \quad (3.12)$$

4 Project 4 - Image Restoration

4.1 Project Proposal

First implement the random noise generation functions. The random noise functions include uniform, Gaussian, Rayleigh, Gamma, exponential and impulse. Requirement is that only uniform noise is allowed to use built-in function while any other noise should be generated based on uniform noise with self-defined functions. The test image is *Fig0503.tif*. Second follow the process shown in textbook *FIGURE 5.7*, *FIGURE 5.8*, *FIGURE 5.10*, *FIGURE 5.11*, *FIGURE 5.12*. Each figure displays a process of add some specified type of noise and then use restoration method to reduce noise.

4.2 Preliminaries

4.2.1 Noise Models

5 Project 6 - Geometric Transformation

5.1 Project Proposal

Implement 4 geometric transform functions - rotate by a specified angle, translate by a given vector, scale by a given rate using nearest neighbor and scale by a given rate using bilinear interpolation.

5.2 Preliminary

5.2.1 Spatial Transform

A spatial transformation of an image is a geometric transformation of the image coordinate system. First we define the coordinate system in an image: *axis-X* is from left to right while *axis-Y* is from top to bottom. In a spatial transformation each point (x, y) of origin image A is mapped to point $(u, v) = (f_1(x, y), f_2(x, y))$ in a new coordinate system.

Affine transformation is any transformation that preserve collinearity and ratios of distances. Rotations, translations, magnifications, shears and their combinations are all affine transformations. Any single affine operation can be expressed as

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (5.1)$$

We extend the 2D point to 3D in order to use a homogeneous matrix multiplication notation. Then we can represent translation, rotation, scaling under this notation. We denote the 3×3 transformation matrix as T

$$T = \begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix} \text{ Translation by } (x_0, y_0) \quad (5.2)$$

$$T = \begin{bmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ Scale by } s_1 \text{ and } s_2 \quad (5.3)$$

$$T = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ Rotate by } \theta \text{ (clockwise)} \quad (5.4)$$

$$T = \begin{bmatrix} 1 & \alpha & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ Horizontal shearing by } \alpha \quad (5.5)$$

$$T = \begin{bmatrix} 1 & 0 & 0 \\ \beta & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ Vertical shearing by } \beta \quad (5.6)$$

Another advantage of matrix multiplication notation for affine transformation is that we can easily get composite affine transformation

$$T = \prod_i T_i \quad (5.7)$$

as well as the inverse transformation T^{-1}

5.2.2 Interpolation

After conduct geometric spatial transformation, the new grid points may not locate exactly on grid points in the new coordinate system. Interpolation is needed in this kind situation to find out the value of the grid points in the new coordinate system. Here, we talk about two methods of interpolation - nearest neighbor interpolation and bilinear interpolation.

Nearest neighbor interpolation is the simplest and fastest interpolation method. In practice, we just use the $(x, y, 1)^T = \text{round}(T^{-1}(u, v, 1)^T)$ as nearest neighbor interpolation.

In bilinear interpolation, we first calculate (x', y') in O_{xy} which is the original phase of a grid point (u, v) in O_{uv} . Then find out the smallest rectangle that surround (x', y') which is formed by 4 grid points $\{(x_0, y_0), (x_1, y_1), (x_2, y_2), (x_3, y_3)\}$ in O_{xy} . Thus we calculate the value at grid point (x', y') as well as (u, v) in O_{uv} in the following way including 2 level of linear interpolation:

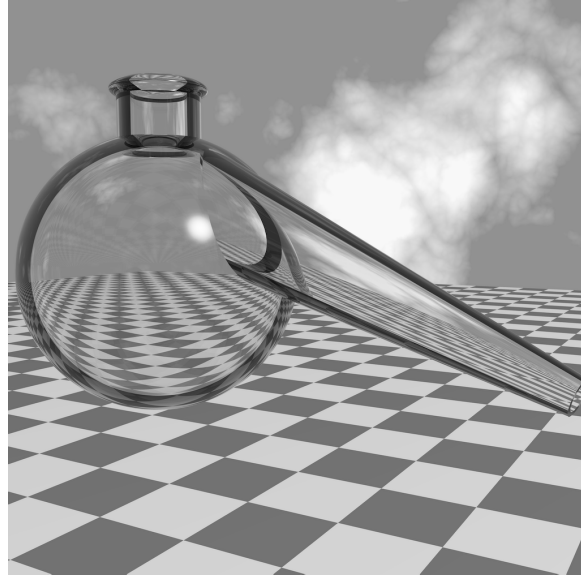
$$g_{0,1}(x', y') = \frac{x_1 - x'}{x_1 - x_0} f(x_0, y_0) + \frac{x' - x_0}{x_1 - x_0} f(x_1, y_1) \quad (5.8)$$

$$g_{3,2}(x', y') = \frac{x_2 - x'}{x_2 - x_3} f(x_3, y_3) + \frac{x' - x_3}{x_2 - x_3} f(x_2, y_2) \quad (5.9)$$

$$g(x', y') = \frac{y_2 - y'}{y_2 - y_1} g_{0,1}(x', y') + \frac{y' - y_1}{y_2 - y_1} g_{3,2}(x', y') \quad (5.10)$$

5.3 Experiment

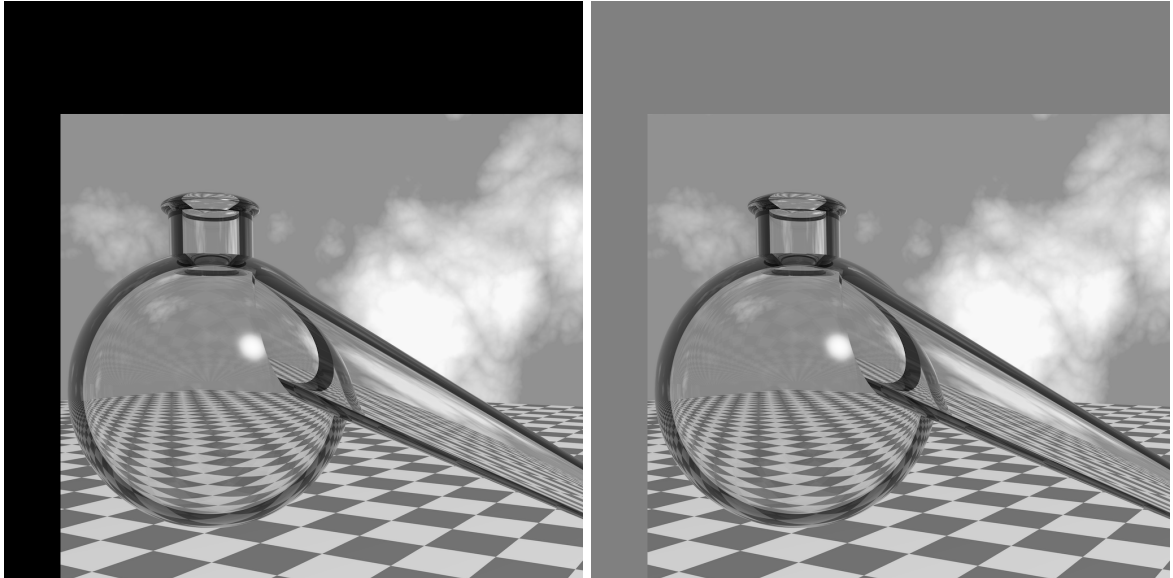
In this experiment we use Fig.5.1 for testing.



(a)

Figure 5.1: Origin test image

Translate The image is translated by 100 pixel at right and by 200 pixel down. The result is display in Fig.5.2.



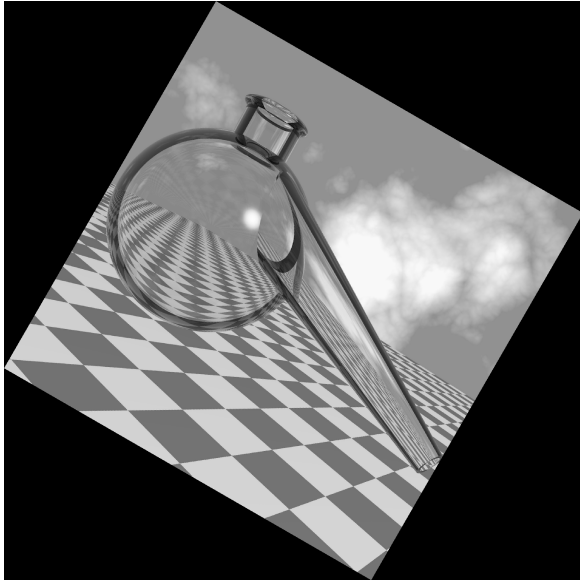
(a) Use nearest neighbor interpolation.

(b) Use bilinear interpolation.

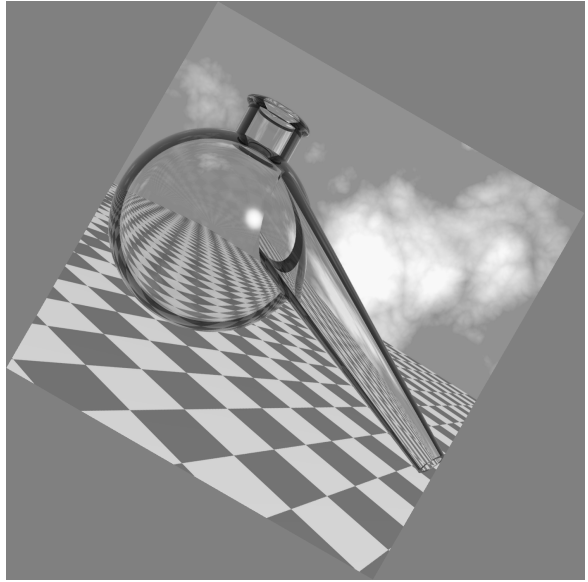
Figure 5.2: Results of translate 100 from left to right and 200 down.

Rotation The image is rotated by $\pi/6$ clockwise. The result is display in Fig.5.3.

Scaling The image is scaled by 1.5 at axis-x and by 2 at axis-y. The result is display in Fig.5.3.



(a) Use nearest neighbor interpolation.



(b) Use bilinear interpolation.

Figure 5.3: Results of rotation. $\pi/6$ clockwise.



(a) Use nearest neighbor interpolation.



(b) Use bilinear interpolation.

Figure 5.4: Results of scaling. Times 1.5 at axis-X and times 2 at axis-Y.

5.4 Implementation

The main part of this implementation is on the function *mytransform*. The input parameters are three necessary parameters - input image, 3×3 affine transform matrix, interpolation method and 3 optional parameters - horizontal and vertical display boundaries and filling colors for empty pixels.

The idea is simple: First, find out the output image's size based only on the transform matrix and input image. The approach is to calculate the coordinates for four rectangle vertices. After consider the user defined boundary, we get the size of final output rectangle range. Then use the each grid coordinate (u, v) to calculate the corresponding (x, y) in input. After get (x, y) , get the pixel value $g(u, v)$ according to the interpolation methods. To get the little square where (x, y) is in, another function *bilinear_rect* is written.

The code of *mytransform* and *bilinear_rect* is listed below:

```

1 function [ imgg ] = mytransform( imgf, matT, interpolation, varargin )
2 %MYTRANSFORM

```

```

3 %
4 ip = inputParser;
5 ip.addParameter('fillValue', 0);
6 ip.addParameter('boundX', [0, 0]);
7 ip.addParameter('boundY', [0, 0]);
8 ip.parse(varargin{:});
9 para=ip.Results;
10
11 boundX = para.boundX;
12 boundY = para.boundY;
13 fillValue = para.fillValue;
14 setBound = 1;
15 if (boundX(2)==boundX(1) || boundY(1)==boundY(2))
16     setBound=0;
17 end
18
19 [M, N] = size(imgf);
20 X = [1 N N 1];
21 Y = [1 1 M M];
22 tmp = matT * [X; Y; ones(1,4)];
23
24 minX = floor(min(tmp(1,:)));
25 minY = floor(min(tmp(2,:)));
26 maxX = ceil(max(tmp(1,:)));
27 maxY = ceil(max(tmp(2,:)));
28
29 if (setBound==1)
30     minX = boundX(1);
31     maxX = boundX(2);
32     minY = boundY(1);
33     maxY = boundY(2);
34 end
35 imgg = zeros(maxY-minY, maxX-minX);
36
37 for i=(minY:maxY)
38     for j=(minX:maxX)
39         tmp = matT \ [j;i;1];
40         u = tmp(1);
41         v = tmp(2);
42         if (isequal('bilinear', interpolation))
43             [X, Y, out] = bilinear_rect(tmp, M, N);
44             if (out==1)
45                 imgg(i-minY+1, j-minX+1) = fillValue;
46                 continue;
47             end
48             g_12 = (X(2)-u)/(X(2)-X(1))*imgf(Y(1),X(1)) + (u-X(1))/(X(2)-X(1))*
                     imgf(Y(1),X(1));
49             g_34 = (X(3)-u)/(X(3)-X(4))*imgf(Y(4),X(4)) + (u-X(4))/(X(3)-X(4))*
                     imgf(Y(3),X(3));
50             imgg(i-minY+1, j-minX+1) = (Y(3)-v)/(Y(3)-Y(2))*g_12 + (v-Y(2))/(Y
                     (3)-Y(2))*g_34;
51         end
52         if (isequal('nearest', interpolation))
53             if (round(u)>=1 && round(u)<=N && round(v)>=1 && round(v)<=M)
54                 imgg(i-minY+1, j-minX+1) = imgf(round(v), round(u));
55             end
56         end
57     end
58 end
59 end
60

```

```

61 function [ X, Y, out ] = bilinear_rect(orig_p , M, N)
62 u = orig_p(1);
63 v = orig_p(2);
64 X = zeros(1,4);
65 Y = zeros(1,4);
66 X(1) = floor(u);
67 X(3) = ceil(u);
68 if(X(1)==X(3))
69     X(3) = X(1)+1;
70 end
71 Y(1) = floor(v);
72 Y(3) = ceil(v);
73 if(Y(1)==Y(3))
74     Y(3) = Y(1)+1;
75 end
76 out = 0;
77 if(X(1)<1 || X(3)>N || Y(1)<1 || Y(3)>M)
78     out = 1;
79 end
80 X(4) = X(1);
81 X(2) = X(3);
82 Y(4) = Y(3);
83 Y(2) = Y(1);
84 end

```

5.5 Discussion

When we are implementing the coordinate transform, we should pay attention to that the coordinate of an image is not the same as the one in mathematics. For example, for efficiency reason, we usually loop 'y' outer and loop 'x' inner. Hence, be careful to use the loop variables correctly while indexing the image matrix. Another difference is that in mathematics, we use counter-clock direction as positive direction for rotation while here we use clockwise direction as positive direction. In this experiment, I didn't see large difference between bilinear interpolation and nearest neighbor interpolation. However, in practice, we should consider the efficiency and effects when we are choosing the proper interpolation algorithm.

6 Project 8 - Morphological Processing

6.1 Project Proposal

Implement the "Opening by reconstruction", "Filling holes" and "Border clearing" operations on textbook chapter 9.5. The task is to reproduce the results in Figure 9.29, 9.31 and 9.32.

6.2 Preliminaries

6.2.1 Basic morphological operations

Morphology offers a unified and powerful approach to numerous image processing problem. These operations defined based on set theory. In this project we consider that morphological operations are conducted on binary images (preprocessing is required for gray-level images). The below table describes the basic widely used morphological processing.

Table 6.1: Summary of basic morphological operations.

Operation	Equation	Comments
Translation	$(B)_z = \{w w = b + z, b \in B\}$	Translation the origin of B to point z .
Reflection	$\hat{B} = \{w w = -b, b \in B\}$	Reflects all elements of B about the origin of this set.
Complement	$A^c = \{w w \notin A\}$	Set of points not in A .
Difference	$A - B = \{w w \in A, w \notin B\}$	Set of points that belong to A but not to B .
Dilation	$A \oplus B = \{z (\hat{B}_z) \cap A \neq \emptyset\}$	Expands the boundary of A .
Erosion	$A \ominus B = \{z (B)_z \subseteq A\}$	Contracts the boundary of A .
Opening	$A \circ B = (A \ominus B) \oplus B$	Smooths contours, breaks narrow isthmuses, and eliminates small islands and sharp peaks.
Closing	$A \bullet B = (A \oplus B) \ominus B$	Smooths contours, fuses narrow breaks and long thin gulfs and eliminates small holes.
Hit-or-miss transform	$A \otimes B = (A \ominus B_1) \cap (A^c \ominus B_2)$	The set of points at which, simultaneously B_1 found a hit in A and B_2 found a match in A^c .

6.2.2 Morphological restoration

With the basic operations, we can discuss a powerful morphological transformation *morphological restoration* that involves two images and a structuring element. One image, the *marker*, contains the starting points for the transformation. The other image, the *mask*, constrains the transformation.

Geodesic dilation and erosion

Geodesic dilation and erosion are the central concepts to morphologic reconstruction. Let F denote the maker and G denote the mask and $F \subseteq G$. The *geodesic dilation* of size 1 of the marker with respect to the mask, denoted by $D_G^{(1)}(F)$, is defined as

$$D_G^{(1)}(F) = (F \oplus B) \cap G \quad (6.1)$$

The geodesic dilation of size n of F with respect to G is defined as

$$D_G^{(n)}(F) = D_G^{(1)} \left[D_G^{(n-1)}(F) \right] \quad (6.2)$$

with $D_G^{(0)}(F) = F$. Similarly, the *geodesic erosion* of size 1 of marker F with respect to mask G is defined as

$$E_G^{(1)}(F) = (F \ominus B) \cup G \quad (6.3)$$

The geodesic erosion of size n of F with respect to G is defined as

$$E_G^{(n)}(F) = E_G^{(1)} \left[E_G^{(n-1)}(F) \right] \quad (6.4)$$

with $E_G^{(0)}(F) = F$. Geodesic dilation and erosion are duals with respect to set complementation.

Morphological reconstruction by dilation and by erosion

Morphological reconstruction by dilation of a mask G from a marker F , denoted $R_G^D F$ is defined

$$R_G^D(F) = D_G^{(k)}(F) \quad (6.5)$$

with k such that

$$D_G^{(k)}(F) = D_G^{(k+1)}(F) \quad (6.6)$$

Morphological reconstruction by erosion of mask G from a marker F , denoted $R_G^E(F)$ is defined

$$R_G^E(F) = E_G^{(k)}(F) \quad (6.7)$$

with k such that

$$E_G^{(k)}(F) = E_G^{(k+1)}(F) \quad (6.8)$$

6.3 Task-1 Opening by reconstruction

The opening by reconstruction of size n of an image F is defined as the reconstruction by dilation of F from the erosion of size n of F ; that is

$$O_R^{(n)}(F) = R_F^D [(F \ominus nB)] \quad (6.9)$$

where $(F \ominus nB)$ indicates n erosions of F by B .

Figure 6.1a is the original image for this task. We are interested in extracting the characters that contain long, vertical strokes. The origin image is of size 918×2018 pixels. The approximate average height of the tall characters is 50 pixels. Thus we use a structuring element of size 51×1 pixels to erode the original image. The erosion image is shown in Figure 6.1b. For comparison, Figure 6.1c shows the opening of Figure 6.1a with the same structuring element. Using erosion image (b) as the marker and the original image (a) as mask, we restored the characters containing long vertical strokes accurately via opening by reconstruction. This result is displayed in Figure 6.1d.

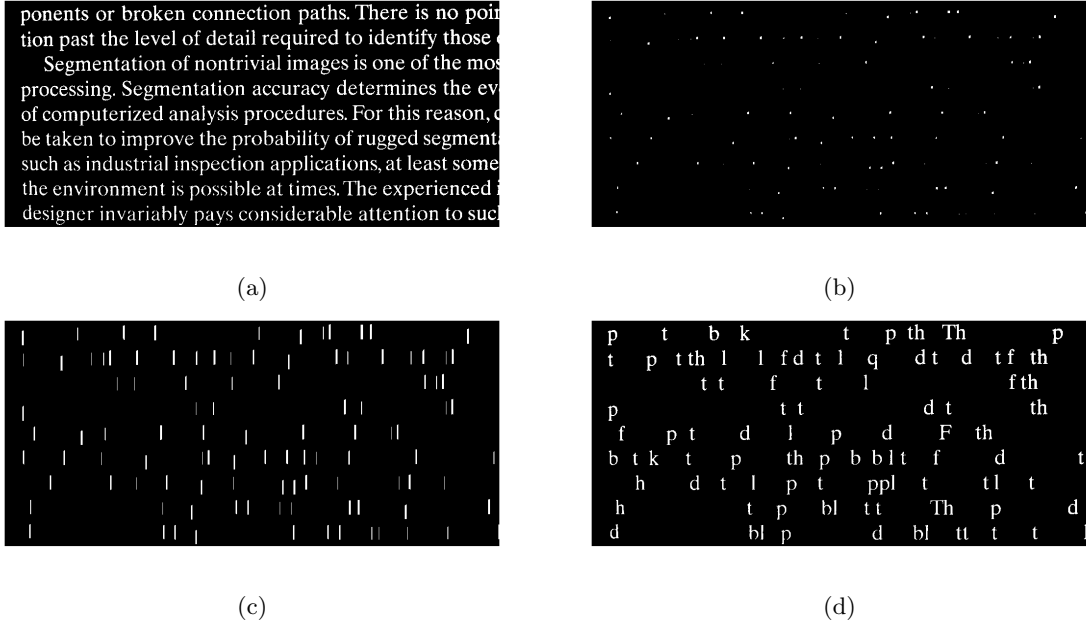
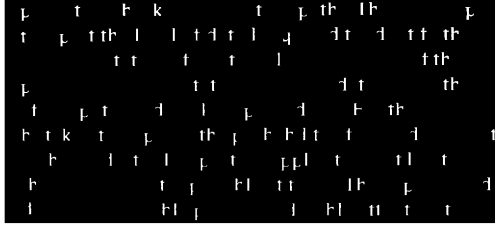
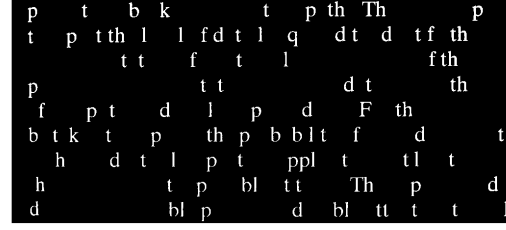


Figure 6.1: Task1-opening by reconstruction. (a)Original binary image 918×2018 . (b)Erosion of (a) with a structuring element of size 51×1 pixels. (c)Opening of(a) with the structuring element 51×1 , shown for comparison. (d)Result of opening by reconstruction.

The size of geodesic dilation reconstruction is 76. This process cost about 7 minutes. I output some of the intermediate results in Figure 6.2 that help us understand the reconstruction better.



(a)



(b)

Figure 6.2: See task-1's process in details. (a)Dilation of size 20. (b)Dilation of size 76, the final results.

6.4 Task-2 Hole filling

Here we develop a fully automated procedure based on morphological reconstruction. Let $I(x, y)$ denote a binary image and we form a marker F that is 0 everywhere, except at the image border; that is,

$$F(x, y) = \begin{cases} 1 - I(x, y) & \text{if } (x, y) \text{ is on the border of } I \\ 0 & \text{otherwise} \end{cases} \quad (6.10)$$

Then

$$H = [R_{I^c}^D(F)]^c \quad (6.11)$$

is a binary image equal to I with all holes filled.

I use 3 structuring element. The hole filling images are shown in Figure 6.3. A detailed process of dilation reconstruction with intermediate results are shown in Figure 6.4. The reconstruction takes 479 steps of dilation, which costs a very long period of about 40 minutes.

ponents or broken connection paths. There is no point past the level of detail required to identify those
Segmentation of nontrivial images is one of the most processing. Segmentation accuracy determines the evolution of computerized analysis procedures. For this reason, care be taken to improve the probability of rugged segmentation, such as industrial inspection applications, at least some the environment is possible at times. The experienced designer invariably pays considerable attention to such

(a)

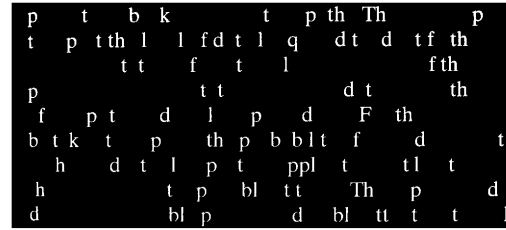


(c)

ponents or broken connection paths. There is no point past the level of detail required to identify those

Segmentation of nontrivial images is one of the most processing. Segmentation accuracy determines the evolution of computerized analysis procedures. For this reason, care be taken to improve the probability of rugged segmentation, such as industrial inspection applications, at least some the environment is possible at times. The experienced designer invariably pays considerable attention to such

(b)



(d)

Figure 6.3: Task2-hole filling. (a)Original binary image 918×2018 . (b)Complement of (a). Used as a mask. (c)Marker image. Seems like a whole black one, but in fact with some white on border. (d)Result of opening by reconstruction.

6.5 Task-3 Border clearing

Removing objects that touch the border is a useful work because it can screen images so that only complete objects remain for further processing. The marker $F(x, y)$ is defined as:

$$F(x, y) = \begin{cases} I(x, y) & \text{if } (x, y) \text{ is on the border of } I \\ 0 & \text{otherwise} \end{cases} \quad (6.12)$$

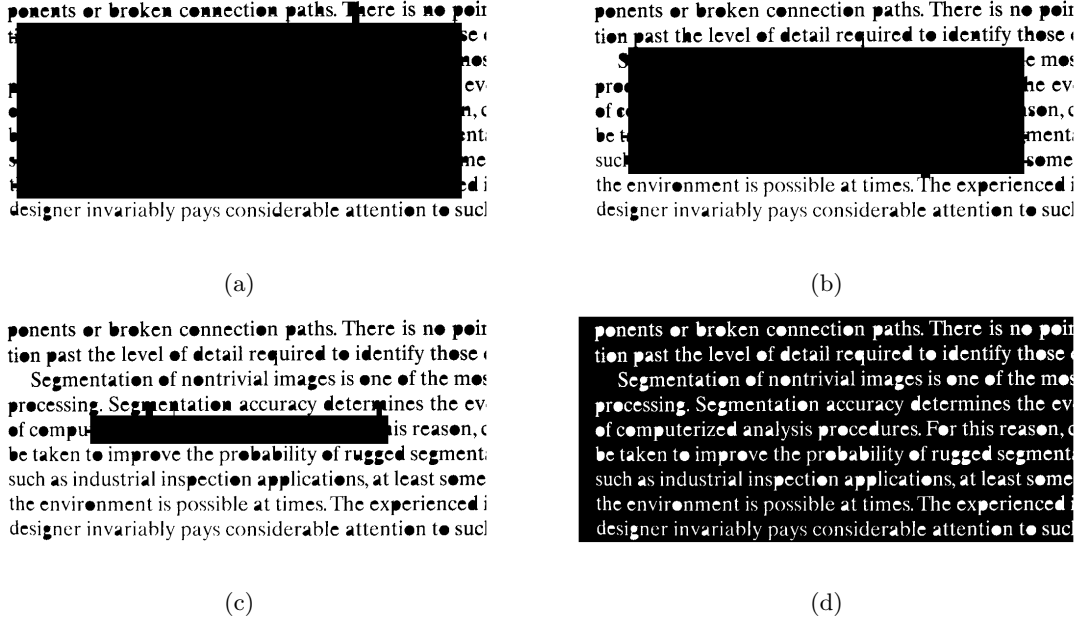


Figure 6.4: The process of hole filling in details. (a)After 100 steps of dilation. (b)After 200 steps of dilation. (c)After 400 steps of dilation. Much closer to success! (d)The final result with completion.

First we computes morphological reconstruction $R_I^D(F)$ and then computes the desired image X

$$X = I - R_I^D(F) \quad (6.13)$$

I use structuring element of size 3×3 . This task is much easier than task 2 and takes just 21 steps of dilation in about 2 minutes. Results are shown in Figure 6.5.

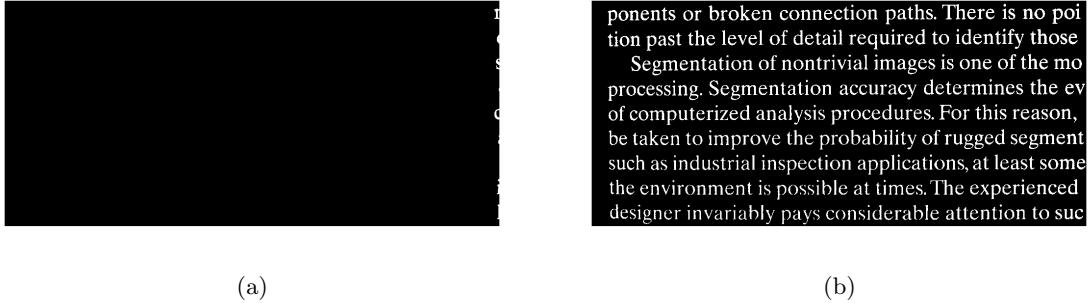


Figure 6.5: Border clearing. (a)After use border marker for 21-dilation reconstruction. The border letter. (b)The final result after subtraction.

6.6 Implementation

Matlab functions *erosion*, *dilation*, *geodesic_dilation*, *dilation_reconstruction* and *opening_reconstruction* are implemented for this project. I made a **mistake** on *opening_reconstruction* at first that I still use 51×1 as the maker while calling *dilation_reconstruction*. The wrong output image is the same as (c) as all the shorter horizontal adjacent relationship in our wanted characters was damaged. So for correction, we must use 3×3 as structuring element for *dilation_reconstruction* called in *opening_reconstruction*.

In the below code frame, I list the key part of these functions. Other process of calling these functions in main script are trivial so omitted here.

```

1 % key part of function erosion
2 for i = (1:M-m+1)
3     for j = (1:N-n+1)
4         x = imgf(i:i+m-1, j:j+n-1);
5         if sum(sum(x.*B)) == m*n

```

```

6         g(i+downshift, j+rightshift) = 1;
7     end
8 end
9 end
10
11 % key part of dilation_reconstruction
12 k_times = 0;
13 while( ~isequal(f0, f1) )
14     k_times = k_times + 1;
15     f0 = f1;
16     f1 = geodesic_dilation(f1, G, B);
17 end
18
19 % key part of geodesic_dilation
20 imgg = dilation(imgf, B) & G;
21
22 % key part of opening_reconstruction
23 for i=(1:n_size)
24     f_erosion = erosion(f_erosion, B);
25 end
26 [imgg, k_times] = dilation_reconstruction(f_erosion, imgf, ones(3,3)); % here
    can not use ones(51, 1)

```

6.7 Discussion

These three tasks show the basic idea of morphological reconstruction used in feature extraction. We start from a marker which can be obtained easily. Then we use the mask as the constrain to conduct set operations. More advanced topics are morphological operations on gray intensity images and image segmentation. This is a very interesting subproject. The main obstacle here is **time**! I tried to do much vectorization but still have a big problem that 20-step-dilation cost about 2 minutes on the 918×2018 image but the complexity is only about $918 \times 2018 \times 20 \times 9 \approx 4 \times 10^8$. I think the computation of this complexity need no more than 2 sec on language like C++ or python. However, I tried the matlab toolbox *IPT* and it's just as fast as the expectation (within 2 sec). This is a strange but interesting problem.

7 Project 9 - Image Segmentation

7.1 Project Proposal

There are two parts of project 9. One task is for edge detection: implement the Roberts, Prewitt, Sobel, the Marr-Hildreth and the Canny edge detectors. The test image is *building.tif*. The other task is to implement the Otsus method of thresholding segmentation, and compare the results with the global thresholding method using test image *polymersomes.tif*.

7.2 Preliminaries

7.2.1 Edge detection

The central idea of edge detection is that local changes in intensity can be detected using derivatives. We have the following conclusions which show that first- and second-order derivatives are particularly well suited for this purpose: (1) First-order derivatives generally produce thicker edges in an image. (2) Second-order derivatives have a stronger response to fine details such as thin lines, isolated points, and noise. (3) Second-order derivatives produce a double-edge response at ramp and step transitions in intensity. (4) The sign of the second derivative can be used to determine whether a transition into an edge is from light to dark or dark to light.

7.2.2 Basic edge detection

The tool of choice for finding edge strength and direction at location (x, y) of an image f is the gradient.

$$\nabla f \equiv \text{grad}(f) \equiv \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (7.1)$$

The magnitude of vector ∇f defined as

$$M(x, y) = \text{mag}(\nabla f) = \sqrt{g_x^2 + g_y^2} \approx |g_x| + |g_y| \quad (7.2)$$

is the value of the rate of change in the direction of the gradient vector. The second \approx is a frequently used approximate to avoid square roots. The direction of gradient vector is given by the angle

$$\alpha(x, y) = \tan^{-1} \left[\frac{g_y}{g_x} \right] \quad (7.3)$$

The direction of an edge at an arbitrary point (x, y) is orthogonal to the direction $\alpha(x, y)$. Here we mention three masks (Figure 7.1) Roberts, Prewitt and Sobel which can be used to compute the gradient of the center point with convolution operation.

-1	0	0	0	0	-1
0	1	1	1	0	0
Roberts					
-1	-1	-1	-1	0	1
0	0	0	-1	0	1
1	1	1	-1	0	1
Prewitt					
-1	-2	-1	-1	0	1
0	0	0	-2	0	2
1	2	1	-1	0	1
Sobel					

Figure 7.1: 3 masks for gradient computation: Roberts, Prewitt, Sobel

7.2.3 More advanced techniques for edge detection