

# Advanced Image Processing

Final Exam.

Student No.: \_\_\_\_\_ Name: \_\_\_\_\_ 109/12

## 一、是非題(T/F test) (30%)

- ( O ) 1. If degradation is systematic (有條理的), it can be suppressed (抑制) by brightness correction. The brightness correction method can be used only if the image degradation process is stable.

[PPT 5-1 p.4]

## Pixel brightness transformations

### • Position-dependent brightness correction

- The sensitivity of image acquisition and digitization devices may depend on position in the image.
- Uneven object illumination (不均匀的光源) is also a source of degradation (退化).
- If degradation is systematic (有條理的), it can be suppressed (抑制) by brightness correction.
- A multiplicative error coefficient  $e(i, j)$  describes the change from the ideal.
- Assume  $g(i, j)$  is the original undegraded image and  $f(i, j)$  is degraded version.

$$f(i, j) = e(i, j)g(i, j)$$

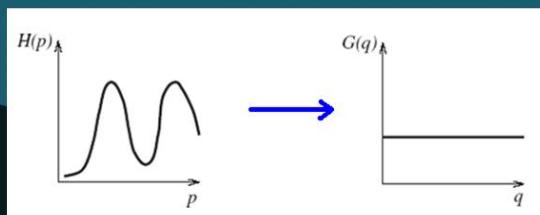
- ( O ) 2. The aim of histogram equalization is to create an image with equally distributed brightness levels over the whole brightness scale.

[PPT 5-1 p.9]

## Gray-scale transformation

### • Histogram equalization

- The aim is to create an image with equally distributed brightness levels over the whole brightness scale.
- Histogram equalization enhances contrast for brightness values close to histogram maxima, and decreases contrast near minima.



- ( O ) 3. If the area of the image is invariant under a geometric transformation, then its Jacobia determinant  $J = 1$ .
- [PPT 5-1 p.22] If the area of the image is invariant under the transformation, then  $J = 1$ .

## Pixel co-ordinate transformation

- A geometric transform applied to the whole image may change the co-ordinate system, and a **Jacobia determinant** / provides information about how the co-ordinate system changes.

$$J = \left| \frac{\partial(x', y')}{\partial(x, y)} \right| = \begin{vmatrix} \frac{\partial x'}{\partial x} & \frac{\partial x'}{\partial y} \\ \frac{\partial y'}{\partial x} & \frac{\partial y'}{\partial y} \end{vmatrix}$$

- If the transformation is **singular** (has no inverse), then  $J = 0$ .
- If the **area** of the image is **invariant** under the transformation, then  $J = 1$ .

© 2015 Cengage Learning Engineering. All Rights Reserved.

22

- 
- 

### [補充]

- Two basic steps of a geometric transform
  - First:** pixel co-ordinate transformation
  - Second:** brightness interpolation

© 2015 Cengage Learning Engineering. All Rights Reserved.

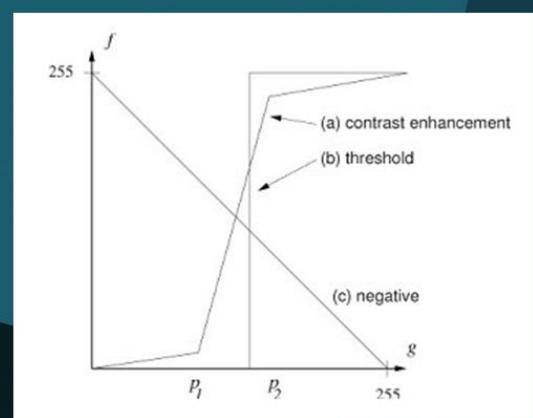
14

- ( X ) 4. Gray-scale transformations (**do not**) depend on the position of the pixel in the image.
- [PPT 5-1 p.7] Gray-scale transformations do not depend on the position of the pixel in the image.

## Gray-scale transformation

- Gray-scale transformations **do not depend on the position of the pixel in the image**.
- A transformation  $\mathcal{T}$  of original brightness  $p$  from scale  $[p_0, p_k]$  into brightness  $q$  from a new scale  $[q_0, q_k]$  is given by  

$$q = \mathcal{T}(p)$$
- The figure shows the most common gray-scale transformations.



© 2015 Cengage Learning Engineering. All Rights Reserved.

7

-

- ( X ) 5. The bilinear transformation includes typical geometric transformations such as rotation and translation, but **not** scaling.
- [PPT 5-1 p.19]

## Pixel co-ordinate transformation

- **Bilinear transform**

$$x' = a_0 + a_1x + a_2y + a_3xy$$

$$y' = b_0 + b_1x + b_2y + b_3xy$$

- The geometric transform can be approximated by a linear transform.

- **Affine transform**

$$x' = a_0 + a_1x + a_2y$$

$$y' = b_0 + b_1x + b_2y$$

- The affine transformation includes typical geometric transformations such as rotation, translation, scaling, and skewing.

© 2015 Cengage Learning Engineering. All Rights Reserved.

19

- ( X ) 6. The position error of nearest-neighborhood brightness interpolation is at **least** (most) half a pixel.
- [PPT 5-1 p.25] interpolation(插值)

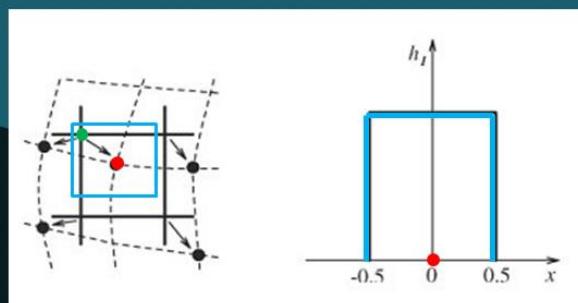
## Brightness interpolation

- Three most common brightness interpolation methods are **nearest neighbor**, **linear** and **bi-cubic**.

- **Nearest-neighborhood interpolation**

$$f_1(x, y) = g_s(\text{round}(x), \text{round}(y))$$

- The position error of nearest-neighborhood interpolation is at most half a pixel.



© 2015 Cengage Learning Engineering. All Rights Reserved.

27

- ( O ) 7. Image smoothing uses redundancy in image data to suppress noise, usually by some form of averaging of brightness values in some neighborhood.

### • Image smoothing

- Image smoothing uses redundancy (多餘;重複) in image data to suppress noise, usually by some form of averaging of brightness values in some neighborhood  $\mathcal{O}$ .
- Goal: suppressing noise
- Problem: blurring sharp edges
- Solution: edge preserving
  - Usually using non-linear methods
  - For example: computing the average only from points in the neighborhood which have similar properties to the point being processed.

© 2015 Cengage Learning Engineering. All Rights Reserved.

32

- ?( X ) 8. “Averaging according to inverse gradient” is a (non-)linear method that avoids edge blurring.

[PPT 5-1 p.32]

## Image smoothing

### • Averaging according to inverse gradient

- Within a convolution mask of odd size, the inverse gradient  $\delta$  of a point  $(i, j)$  with respect to the central pixel  $(m, n)$  is defined as

$$\delta(i, j) = \frac{1}{|g(m, n) - g(i, j)|}$$

- If  $g(m, n) = g(i, j)$ , then we define  $\delta(i, j) = 2$ .
- So  $\delta(i, j) \in (0, 2]$  and is smaller at the edge than in the interior of a homogeneous region.
- The kernel function  $h$

$$h(i, j) = 0.5 \frac{\delta(i, j)}{\sum_{(m,n) \in \mathcal{O}} \delta(m, n)}$$

Moreover, the mask coefficient corresponding to the central pixel is defined as  $h(i, j) = 0.5$ .

© 2015 Cengage Learning Engineering. All Rights Reserved.

39

- ( O ) 9. Median filter is a nonlinear smoothing method.

• [PPT 5-1 p.43] Median filtering:

## Image smoothing

### • Median filtering

- In probability theory, the **median** divides the higher half of a probability distribution from the lower half.
- Median filter is a **non-linear** smoothing method.

#### Algorithm 5.3 Efficient median filtering

1. Set  $t = \frac{mn}{2}$

( $m, n$  are the numbers of rows and columns of the median window and both odd, round  $t$ )

2. Position the window at the beginning of a new row, and sort its contents.

Construct a histogram  $H$  of the window pixels, determine the median  $m$ , and record  $n_m$ , the number of pixels with intensity less than or equal to  $m$ .

3. For each pixel  $p$  in the leftmost column of intensity  $p_g$ , perform

$$H[p_g] = H[p_g] - 1$$

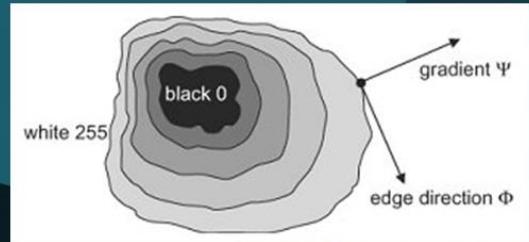
Further, if  $p_g < m$ , set  $n_m = n_m - 1$ .

- ( O ) 10. The edge direction is rotated with respect to the gradient direction by  $-90^\circ$ .

[PPT 5-2 p.5] The edge direction  $\phi$  is rotated with respect to the gradient direction  $\psi$  by  $-90^\circ$

## Edge detectors

- An edge is a property attached to **an individual pixel** and is calculated from the image function behavior in a neighborhood of that pixel.
- It is a **vector variable** with two components.
  - **Magnitude**
    - The edge magnitude is the magnitude of the gradient.
  - **Direction**
    - The **edge direction**  $\phi$  is rotated with respect to the **gradient direction**  $\psi$  by  $-90^\circ$ .
    - The gradient direction gives the direction of maximum growth of the function.



- ( O ) 11. The Prewitt operator approximates the first derivative. The gradient is estimated in eight (for a  $3 \times 3$  convolution mask) possible directions, and the convolution result of largest magnitude indicates the gradient direction.

### • Prewitt operator

- The Prewitt operator approximates the **first derivative**.
  - The gradient is estimated in **eight** (for a  $3 \times 3$  convolution mask) possible directions, and the convolution result of **greatest magnitude** indicates the **gradient direction**.
  - Some examples of  $3 \times 3$  masks
- $$h_1 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \quad h_2 = \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix} \quad h_3 = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \dots$$
- Larger masks are possible, for example,  $5 \times 5$  or  $7 \times 7$ .

PS. The direction of the gradient is given by the mask giving **maximal response**. This is also the case for all the following operator approximating the first derivative.

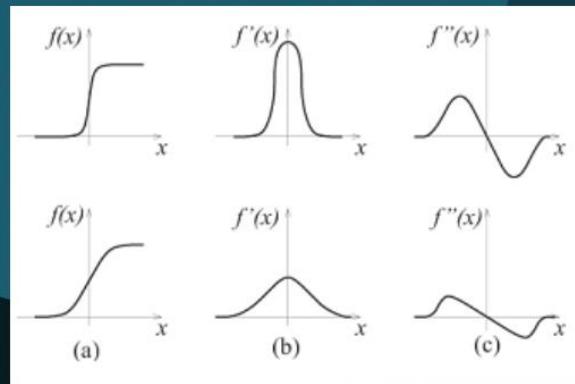
© 2015 Cengage Learning Engineering. All Rights Reserved.

12

- ( X ) 12. Zero-crossings is an edge detection technique using the **first**-(**second**) derivative.

### • Zero-crossings

- An edge detection technique using the **second derivative**.
- The **first** derivative of the image function should have an extremum at the position corresponding to the image.
- The **second** derivative should be **zero** at the same position.
- However, it is **much easier** and **more precise** to find a zero-crossing position than an extremum.



© 2015 Cengage Learning Engineering. All Rights Reserved.

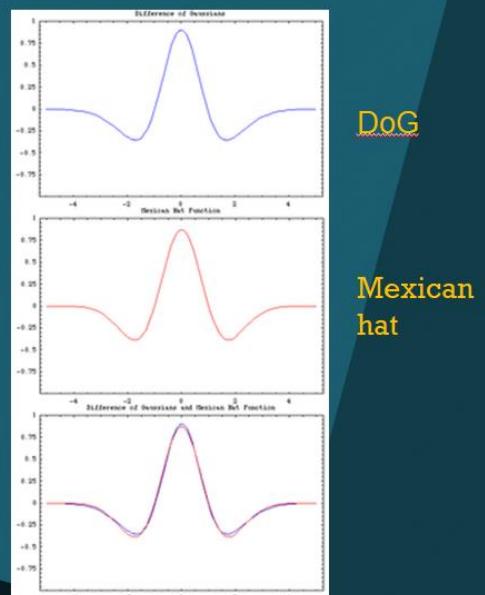
16

- ( X ) 13. The definition of **Laplacian of Gaussian (LoG)** (**the difference of Gaussians (DoG)**) operator is the difference of two Gaussian averaging masks with substantially different variance.

## Zero-crossings of the second derivative

- LoG v.s. DoG

- The  $\nabla^2 G$  operator can be very effectively approximated by convolution with a mask that is the difference of two Gaussian averaging masks with substantially different  $\sigma$ .
- The method is called the difference of Gaussians (DoG)



DoG

Mexican hat

Comparison of difference of Gaussian with Mexican hat wavelet  
[https://en.wikipedia.org/wiki/Difference\\_of\\_Gaussians](https://en.wikipedia.org/wiki/Difference_of_Gaussians)

© 2015 Cengage Learning Engineering. All Rights Reserved.

22

- ( X ) 14. The one-response-(detection) criterion of edge detector indicates that the important edges should not be missed.

- Canny edge detection

- It is optimal for step edges corrupted by white noise.
- The optimality of the detector is related to three criteria.
  - The detection criterion
    - The important edges should not be missed.
    - There should be no spurious (假的) edges.
  - The localization criterion
    - The distance between the actual and located position of the edge should be minimal.
  - The one response criterion
    - Minimizing multiple responses to a single edge

© 2015 Cengage Learning Engineering. All Rights Reserved.

28

- ( O ) 15. The Laplacian operator for edge detection has the same properties in all directions and is therefore invariant to rotation.

# Edge detectors

- The **gradient magnitude**  $|\text{grad } g(x, y)|$  and **gradient direction**  $\psi$  are continuous image functions calculated as

$$|\text{grad } g(x, y)| = \sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2}$$

$$\psi = \arg\left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right)$$

where  $\arg(x, y)$  is the angle (in radians; 強度) from the  $x$  axis to  $(x, y)$ .

- Laplacian:** (obtained only edge magnitudes **without** orientations)

$$\nabla^2 g(x, y) = \frac{\partial^2 g(x, y)}{\partial x^2} + \frac{\partial^2 g(x, y)}{\partial y^2}$$

- The Laplacian has the same properties in all directions and is therefore **invariant to rotation**.

© 2015 Cengage Learning Engineering. All Rights Reserved.

7

- ( O ) 16. Non-maximal suppression is an edge thinning technique.

## Canny edge detection

- Substituting in equation (5.56) for  $G_n$  from equation (5.54)

$$\frac{\partial}{\partial \mathbf{n}} (G_n * f) = \frac{\partial}{\partial \mathbf{n}} \left( \frac{\partial G}{\partial \mathbf{n}} * f \right) = \frac{\partial^2}{\partial \mathbf{n}^2} (G * f) = 0 \quad (5.57)$$

- This equation illustrates how to find local maxima in the direction perpendicular to the edge (**non-maximal suppression**—Algorithm 6.4)

- Non-maximal suppression** is an edge thinning technique.

- The **magnitude of the edge** is measured as

$$|G_n * f| = |\nabla(G * f)|$$

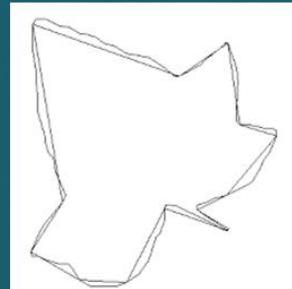
© 2015 Cengage Learning Engineering. All Rights Reserved.

30

- ( O ) 17. Edge detectors based on parametric models describe edges more precisely than convolution-based edge detectors. However, their computational requirements are much higher.

## Parametric edge models

- Once the facet model parameters are available for each image pixel, edges can be detected as extrema of the first directional derivative and/or zero-crossings of the second directional derivative of the local continuous facet model functions.
- Edge detectors based on parametric models describe edges more precisely than convolution-based edge detectors.
  - They carry the potential for subpixel edge localization.
  - However, their computational requirements are much higher.



[http://cs.joensuu.fi/~koles/approximation/Ch3\\_5.html](http://cs.joensuu.fi/~koles/approximation/Ch3_5.html)

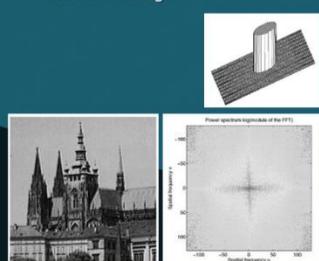
© 2015 Cengage Learning Engineering. All Rights Reserved.

38

- ( O ) 18. Edge can be enhanced by high-pass frequency-domain filtering and noise can be suppressed by low-pass frequency-domain filtering.

### Local pre-processing in the frequency domain

- Low-pass frequency-domain filtering
  - Noise can be suppressed.
  - A blurred image

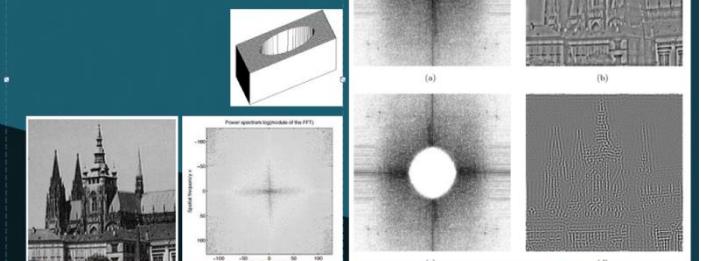


© 2015 Cengage Learning Engineering. All Rights Reserved.

6

### Local pre-processing in the frequency domain

- High-pass frequency-domain filtering
  - Edge can be enhanced.



© 2015 Cengage Learning Engineering. All Rights Reserved.

7

- ( O ) 19. Otsu's algorithm is one kind of optimal thresholding.

## Optimal thresholding

- Optimal thresholding

- To model the histogram of an image using a weighted sum of two or more probability densities with normal distribution.
- The threshold is set as the gray-level closest to the minimum probability between the maxima of the normal distributions, which results in minimum error segmentation.
- Otsu's algorithm
  - To test each possible threshold, and to compute the gray-level variances of both foreground and background each one implies.
  - The threshold is at the place where a weighted sum of these variances is minimal.

© 2015 Cengage Learning Engineering. All Rights Reserved.

12

- ( X ) 20. Let  $D_0$  be the cut-off frequency coincides with the dispersion  $\sigma$ ,  $D(u, v)$  be an isotropic (等向的) filter (for example,  $D(u, v) = \sqrt{u^2 + v^2}$ ), then the Fourier spectrum of the filter  $G(u, v) = \exp\left(-\frac{1}{2}\left(\frac{D(u, v)}{D_0}\right)^2\right)$  is a high(low)-pass Gaussian filter.

## Local pre-processing in the frequency domain

- The Gaussian low-pass filter

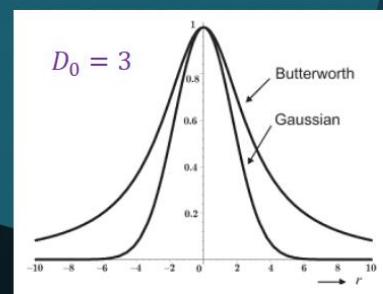
- The cut-off frequency  $D_0$  coincides (一致) with the dispersion  $\sigma$ .
- The Fourier spectrum of a low-pass Gaussian filter  $G_{low}$  is

$$G_{low}(u, v) = \exp\left(-\frac{1}{2}\left(\frac{D(u, v)}{D_0}\right)^2\right)$$

- The Butterworth low-pass filter

$$B_{low}(u, v) = \frac{1}{1 + \left(\frac{D(u, v)}{D_0}\right)^n}$$

The usually Butterworth filter degree is  $n = 2$ .



© 2015 Cengage Learning Engineering. All Rights Reserved.

11

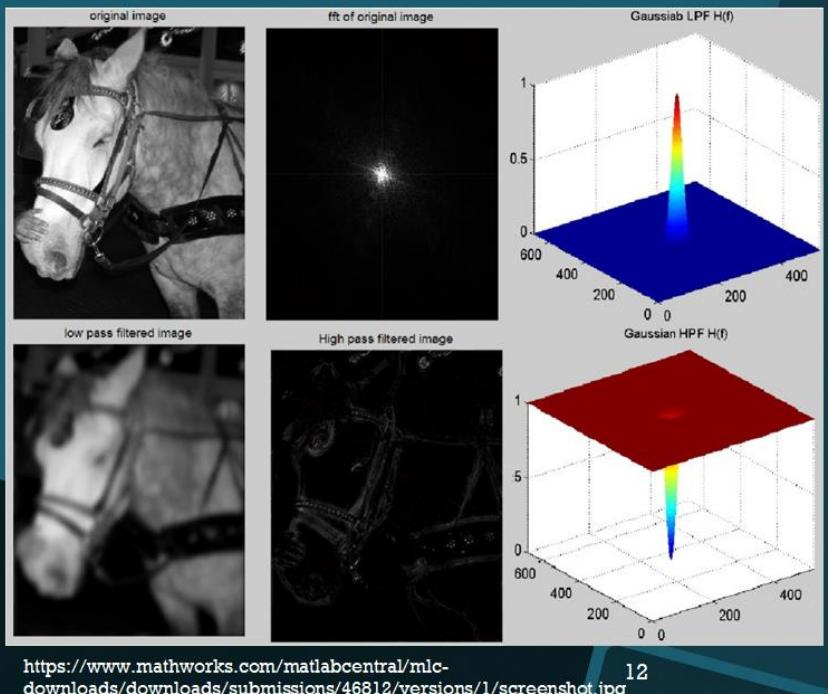
- ( X ) 21. The high-pass filter is created easily from the low-pass filter. If the Fourier frequency spectrum of a low-pass filter is  $H_{low}$ , the high-pass filter can be created by just flipping it vertically  $\textcolor{red}{H_{high} = 1/H_{low}}$ . ( $H_{high} = 1 - H_{low}$ )

## Local pre-processing in the frequency domain

### • High-pass filter

- The high-pass filter is created easily from the low-pass filter.
- If the Fourier frequency spectrum of a low-pass filter is  $H_{low}$ , the high-pass filter can be created by just flipping it vertically

$$H_{high} = 1 - H_{low}$$



<https://www.mathworks.com/matlabcentral/mlc-downloads/downloads/submissions/46612/versions/1/screenshot.jpg>

12

- ( O ) 22. A corner in an image can be defined as a pixel in whose immediate neighborhood there are two dominant, different edge directions.

### • Corners

- A **corner** in an image can be defined as a pixel in whose immediate neighborhood there are **two dominant**(明顯的), **different edge directions**.
- Corner detectors are **not** usually very robust. One can detect corners using two or more images.
- Corner detector
  - Moravce detector**
  - Zuniga-Haralick operator**
  - Harris corner detector**

- ( O ) 23. Homomorphic filtering is used to remove multiplicative noise.

## Local pre-processing in the frequency domain

### • Homomorphic (同態的) filtering

- Another useful pre-processing technique operating in the frequency domain
- Homomorphic filtering is used to remove multiplicative noise.  
PS: In signal processing, the term **multiplicative noise** (乘性雜訊) refers to an unwanted random signal that gets multiplied into some relevant signal during capture, transmission, or other processing. An important example is **the speckle noise** (斑點雜訊) commonly observed in radar imagery.
- The assumption is the image function  $f(x, y)$  can be factorized as a product of two **independent** multiplicative components in each pixel.

$$f(x, y) = i(x, y)r(x, y)$$

where  $i(x, y)$  is the **illumination** and  $r(x, y)$  is the **reflectance**

© 2015 Cengage Learning Engineering. All Rights Reserved.

13

[補充]:Homomorphic filtering 的作法

P.11 在這 filter 內的值全部 (不再全為 0. 或 1)  
 P.12 high pass = 1 - low pass

¶ 三節  
 P.25 Homomorphic : 去掉乘性雜訊 (multiplicative noise)

P.14. Homomorphic Filtering  $\Rightarrow$  方法取 log

$$f(x, y) = i(x, y) \cdot r(x, y)$$

↓ log

$$z(x, y) = \log f(x, y) = \log i(x, y) + \log r(x, y)$$

↓ FFT

$$Z(u, v) = I(u, v) + R(u, v)$$

| A HRF filter the filtering

$$S = H_r * Z = H_r * I + H_r * R$$

↓ 做 inverse FFT

$$s(x, y) = \mathcal{F}^{-1}(S(u, v))$$

↓ IR e

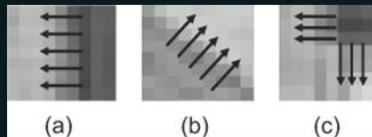
$$g(x, y) = \exp(s(x, y))$$

- ( X ) 24. In Harris corner detector, let  $\lambda_1$  and  $\lambda_2$  be the eigenvalues of the Harris matrix A, a corner has been found if both eigenvalues are **small.(large)**

## Detection of corners

- **Harris corner detector**

- Let  $\lambda_1, \lambda_2$  be the eigenvalues of A, three distinct cases can occur:
  - Both eigenvalues are small.
    - Image f is flat at the examined pixel.
    - There is no edges or corners in this location.
  - One eigenvalue is small and the second large.
    - The local neighborhood is ridge(山脊)-shaped.
    - Significant change of f occurs if a small movement is made perpendicularly to the ridge.
  - Both eigenvalues are rather large.
    - A small shift in any direction causes significant change of f.
    - A corner has been found.



**Figure 5.33:** The Harris corner detector according to eigenvalues of the local structure matrix. (a), (b) Ridge detected, no corner at this position. (c) Corner detected.  
© Cengage Learning 2015.

- **The Harris matrix A is symmetric and positive semi-definite.**

- ( O ) 25. Invariance to monotonic transformations is one property of maximally stable extremal Regions (MSERs).

## Detection of maximally stable extremal regions

- Maximally Stable Extremal Regions (MSERs) have the following properties.
  - **Invariance to monotonic transformations.**  
If  $I(p) < I(q)$  then  $M(I(p)) = I'(p) < I'(q) = M(I(q))$   
where  $M$  is a transformation.
  - **Invariance to adjacency preserving (continuous) transformations.**
  - **Stability**, since only extremal regions whose support is virtually unchanged over a range of thresholds are selected.
  - **Multi-scale detection**. Since no smoothing is involved, both very fine and very large structure is detected.
  - The set of all extremal regions can be enumerated in  $\mathcal{O}(n \log \log n)$ , almost in linear time for 8 bit images.

- ( O ) 26. A complete segmentation of an image  $R$  is a finite set of regions  $R_1, R_2, \dots, R_s$ ,  
 where  $R = \bigcup_{i=1}^s R_i$ , and  $R_i \cap R_j = \emptyset, i \neq j$ .

## Segmentation

- The main goal of image segmentation is to **divide an image into parts** that have a strong correlation with objects or areas of the real world contained in the image.
- A complete **segmentation** of an image  $R$  is a finite set of regions  $R_1, \dots, R_s$ ,

$$R = \bigcup_{i=1}^s R_i,$$

$$R_i \cap R_j = \emptyset, i \neq j.$$



<https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html>

© 2015 Cengage Learning Engineering. All Rights Reserved. 3

- ( O ) 27.  $p$ -tile(瓦)thresholding method uses a prior information to choose a threshold  $T$  such that  $1/p$  of the image area has gray values less than  $T$ .

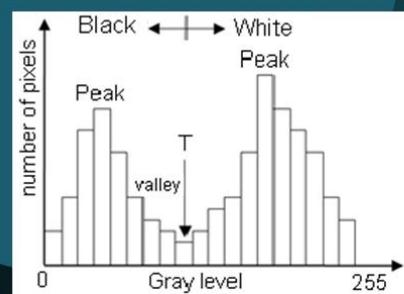
## Threshold detection methods

### • **$p$ -tile thresholding**

- Using a **prior information**, to choose a threshold  $T$  such that  $1/p$  of the image area has gray values less than  $T$ .
- For example, on a page of printed text the characters  $1/p$  cover of the sheet area.

### • **Bimodel histogram**

- Pixels of objects form one of its peaks, while pixels of the background form the second peak.



[https://www.researchgate.net/figure/233814424\\_fig1\\_Figure-1-Histogram-of-a-sample-gray-level-bimodal-image-T-is-the-threshold-value](https://www.researchgate.net/figure/233814424_fig1_Figure-1-Histogram-of-a-sample-gray-level-bimodal-image-T-is-the-threshold-value)

© 2015 Cengage Learning Engineering. All Rights Reserved.

10

- ( X ) 28. The Hough transform was designed to detect straight lines and curves. However, this approach is (not) very sensitive to imperfect data or noise.

## Hough transforms

- The Hough transform was designed to detect **straight lines and curves** [Hough, 1962].
  - A **big advantage** of this approach is **robustness of segmentation result**. Segmentation is not to sensitive to imperfect data or noise.
- Suppose we seek instances of curves satisfying the equation  $f(\mathbf{x}, \mathbf{a}) = 0$ , where  $\mathbf{a}$  is an  $n$ -dimensional vector of curve parameters.
  - Apply an edge detection method to detect edges.
  - Then we can apply Hough transform method.

© 2015 Cengage Learning Engineering. All Rights Reserved.

3

- ( X ) 29. Homogeneity (同質性) is an important property of regions. Let  $S$  be the total number of regions in an image, and  $H(R_i)$  be a binary homogeneity evaluation of the region  $R_i$ , the criteria for homogeneity can be defined as  $H(R_i) = \text{TRUE}, i = 1, 2, \dots, S$ , and  $H(R_i \cup R_j) = \text{FALSE}, i \neq j$ . ( $R_i$  adjacent to  $R_j$ )

## Region-based segmentation

- **Region growing techniques** are generally **better** in noisy images, where borders are extremely difficult to detect.
- **Homogeneity (同質性)** is an important property of regions.
- The criteria for homogeneity can be based on gray-level, color, texture, shape, model, etc.
- Region must satisfy the conditions:

$$H(R_i) = \text{TRUE}, \quad i = 1, 2, \dots, S \quad (6.22)$$

$$H(R_i \cup R_j) = \text{FALSE}, \quad i \neq j, R_i \text{ adjacent to } R_j$$

where  $S$  is the total number of regions in an image

$H(R_i)$  is a binary homogeneity evaluation of the region  $R_i$

Resulting regions of the segmented image must be both **homogeneous** and **maximal**.

© 2015 Cengage Learning Engineering. All Rights Reserved.

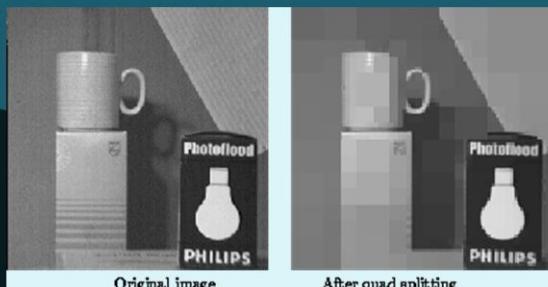
16

- ( X ) 30. Region splitting is the opposite of region merging. Region splitting begins with each pixel (whole image) represented as a single region.

## Region splitting

- Region splitting is the opposite of region merging.
- It begins with the whole image represented as a single region.
- Region splitting methods generally use similar homogeneity criteria as region merging methods.
- Compare with region merging, region splitting does not result in the same segmentation even if the same homogeneity criteria are used.

Why?



[http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL\\_COPIES/MARBLE/medium/segment/split.htm](http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MARBLE/medium/segment/split.htm)

© 2015 Cengage Learning Engineering. All Rights Reserved.

19

## 二、簡答證明題(assay question)

1. (2%) Here shows an edge detection mask, what is the value of  $a$ ?

$$h = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & a & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

ANS: -12

2. (2%) Here shows a smoothing mask, what is the value of  $b$ ?

$$h = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & b & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

ANS: 4

3. (2%) What kind of objects can be detected by the following masks?

$$h = \begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix}$$

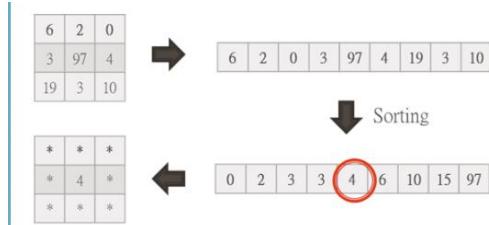
ANS: 水平線



4. (6%) (a) How to remove the salt-and-pepper noises?  
 (b) How to remove the periodic (周期性的) noises?  
 (c) How to remove Gaussian noises?

ANS:

(a) 用 Median Filter 非常適合去除影像中較為極端的雜訊



[補充]

### Image smoothing

- Median filtering ~~會出什麼問題~~
- The main disadvantage of median filtering in a rectangular neighborhood is its damaging of the ~~細線和角會被破壞~~ thin lines and sharp corners.
- This can be avoided if another shape of neighborhood is used.
- For example, if horizontal/vertical lines need preserving, a neighborhood such as that in the figure can be used.

Horizontal/vertical line preserving neighborhood for median filtering

© 2015 Cengage Learning Engineering. All Rights Reserved. 46

(b) 先轉換到 frequency domain 做 filtering 後 再轉換回 spatial domain

(c) 可以用(b)的方法 也可以用 averaging 的 filter(ex:  $\frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$  這樣的filter)

5. (3%) Compare with region merging, region splitting does not result in the same segmentation even if the same homogeneity criteria are used. Why?

ANS: region merging 跟 region splitting 可能會因為執行順序不一樣而有不一樣的結果。比如：一張圖片 [20 15 10 5] 用 15 為分界。在 merging 時會分成 [15 15 15] [5]，這時要再分裂就會還是 [15 15 15] [5]，回不去了。

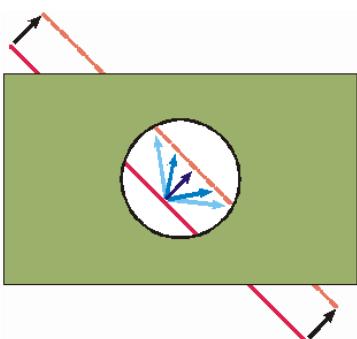
6. (2%) Hough transform is a very powerful technique for detection. What is the main problem of Hough transform?

ANS: Hough 變換算法的特點導致其時間複雜度和空間複雜度都很高，並且在檢測過程中只能確定直線方向，丟失了線段的長度信息。

7. (4%) Corners is one kind of interest points for the correspondence problem.
- (a) (2%) What is a corner? (please give the definition of a corner)
- (b) (2%) Corners serves better than edges (lines) when the correspondence problem is to be solved because of the aperture problem (孔徑問題). What is the aperture problem?

ANS:

- (a) A corner in an image can be defined as a pixel in whose immediate neighborhood there are two dominant(明顯的), different edge directions.
- (b) 當只能小範圍的觀察直線時，會無法判斷直線的移動方向。如下圖你會不知道是往哪個方向移動。



8. (2%) The following image is one example of corrupted (篡改;塗改) image. How to remove this kind of noise?



apply image averaging only to pixels in that interval

9. (4%) What is the meanings of the difference of Gaussian? Can you give an application which can use the convolution masks of the difference of Gaussian?

ANS:

- (a) The  $\nabla^2 G$  operator can be very effectively approximated by convolution with a mask that is the difference of two Gaussian averaging masks with substantially different  $\sigma$ .

The method is called  
the difference of Gaussians (DoG)

- (b) 可以用在 Zero-crossing 時做輔助

10. (4%) Zero-crossings is an edge detection technique. What is zero-crossings? What is the main advantages of zero-crossings?

ANS:

- (a) An edge detection technique using the second derivative.  
 (b) it is much easier and more precise to find a zero-crossing position than an extremum.

11. (2%) Three most common brightness interpolation methods are nearest neighbor, linear and bi-cubic. Please explain the nearest neighbor interpolation method.

ANS:

$$f_1(x, y) = g_s(\text{round}(x), \text{round}(y))$$

12. (3%) Please explain what the hysteresis (滯後作用) approach used by Canny edge detection is.

ANS:

Usually a weak edge pixel caused from true edges will be connected to a strong edge pixel while noise responses are unconnected

# Hysteresis Thresholding

- **Double threshold:** Select high and low threshold values
  - If an edge pixel's gradient value is higher than the high threshold value, it is marked as a **strong edge pixel**.
  - If an edge pixel's gradient value is smaller than the high threshold value and larger than the low threshold value, it is marked as a **weak edge pixel**.
  - If an edge pixel's value is smaller than the low threshold value, it will be suppressed.
- **Edge tracking by hysteresis (滯後作用)**
  - Usually a **weak edge pixel** caused from true edges will be connected to a **strong edge pixel** while noise responses are unconnected.
  - To track the edge connection, blob analysis is applied by looking at a weak edge pixel and its 8-connected neighborhood pixels.
  - As long as there is one strong edge pixel that is involved in the blob, that weak edge point can be identified as one that should be preserved.

[https://en.wikipedia.org/wiki/Canny\\_edge\\_detector](https://en.wikipedia.org/wiki/Canny_edge_detector)

© 2015 Cengage Learning Engineering. All Rights Reserved.

33

13. (2%) A straight line can be represented by the equation  $x_2 = kx_1 + q$  with two the parameters  $k$  and  $q$ . Why this equation is not suitable to represent a straight line in Hough transforms?

ANS: 因為  $k$  與  $q$  值為固定，所以可以產生無限多種可能的直線

## Hough transforms

- Similarly, if a **straight line** is sought we would look for solutions of the equation

$$x_2 = kx_1 + q$$

the parameter space is given by  $(k, q)$ . (a 2D parameter space)

- However, the gradient  $k$  is **unbounded**.
- In this case we reformulate the equation as

$$s = x_1 \cos\theta + x_2 \sin\theta$$

and seek solutions  
in  $(s, \theta)$  space.

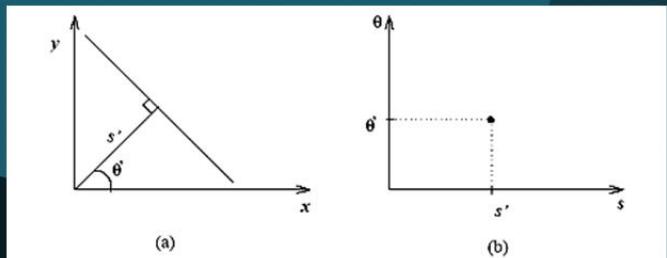
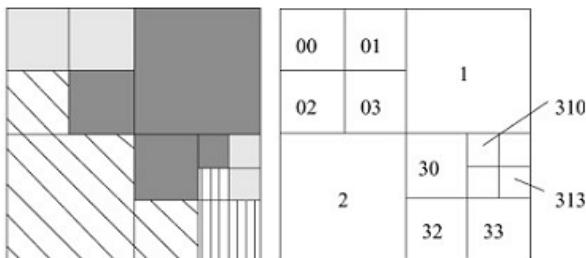


Figure 6.35: Hough transform in  $s, \theta$  space. (a) Straight line in image space. (b)  $s, \theta$  parameter space.

14. (4%) (a) (2%) Please draw the corresponding quadtree of the following image.

(b) (2%) What is the **drawback** of segmentation quadtree?



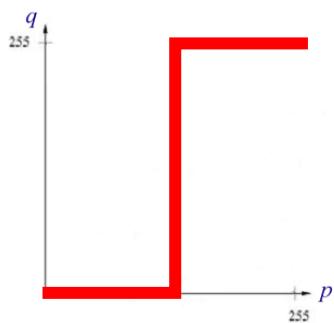
(a)

(b) the square-region shape assumption

15. (2%) A transformation  $T$  of original brightness  $P$  from scale  $[p_0, p_k]$  into brightness  $q$  from a new scale  $[q_0, q_k]$  is given by  $q = T(p)$ . Given an original image as Figure (a), can you draw the transformation  $T$  if the output image is (b)?



(a) original image



Function  $T$



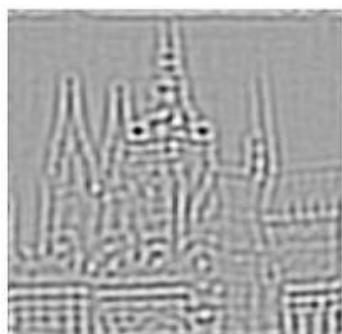
(b) output image

16. (4%) (1) (2%) Figure (a) shows the original image, and Figure (b) shows the filtered image. What kind of filters is used in the frequency domain? Low-pass filters? High-pass filters? Band-pass filters? Or homomorphic filtering? ANS: Band-pass filters

(2) (2%) If the filtered image is shown in Figure (c), then what kind of the filters is used in the frequency domain? ANS: low-pass filters



(a)



(b)



(c)

17. (4%) Here shows Otsu's threshold detection algorithm, please fill in the blanks.

### Algorithm 6.2 Otsu's threshold detection

1. For an image  $I$  on  $G$  gray levels, compute the gray-level histogram  $H(0), H(1), \dots, H(G-1)$ . Normalize the histogram by dividing through by the number of pixels in  $I$  – the histogram now represents the probability of each gray level.
2. For each possible threshold  $t = 0, 1, \dots, G-2$ , partition the histogram into background  $B$  (gray-levels less than or equal to  $t$ ), and foreground  $F$  (gray-levels more than  $t$ ).
3. Compute  $\sigma_B(t)$ ,  $\sigma_F(t)$ , the variance of the background and foreground gray-levels. Compute the probability of a pixel being background

$$w_B(t) = \sum_{j=0}^t H(j) \text{ and } w_F(t) \text{ similarly. Set}$$

$$\sigma(t) = \frac{\sigma_B(t)\sigma_F(t)}{w_B(t)\sigma_B(t) + w_F(t)\sigma_F(t)}$$

$$\text{and select as threshold } \hat{t} = \arg \min_t (\sigma(t)).$$

18. (3%) Here shows a histogram equalization algorithm, please fill in the blanks.

### Algorithm 5.1 Histogram equalization

1. For an  $N \times M$  image of  $G$  gray-levels, initialize an array  $H$  of length  $G$  to 0.
2. From the image histogram: Scan every pixel  $p$  – if it has intensity  $g_p$ , perform

$$H[g_p] = H[g_p] + 1$$

Then let  $g_{\min}$  be the minimum  $g$  for which  $H[g] > 0$ .

3. Form the cumulative image histogram  $H_c$ :

$$H_c[0] = H[0]$$

$$H_c[g] = H_c[g-1] + H[g], g = 1, 2, \dots, G-1$$

Let  $H_{\min} = H_c[g_{\min}]$ .

4. Set

$$T[g] = \text{round} \left( \frac{H_c[g] - H_{\min}}{MN - H_{\min}} (G-1) \right)$$

5. Rescan the image and write an output image with gray-levels  $g_q$ , setting

$$g_q = T[g_p]$$

## 19. (4%) Given a 2D Gaussian filter

$$G(x, y) = e^{-(x^2+y^2)/2\sigma^2}$$

where  $\sigma$  is the standard deviation. Please derive  $\frac{\partial^2 G}{\partial x^2}$ , and the Laplacian of Gaussian  $\nabla^2 G$ .

$$(\text{PS. } \nabla^2 G = \frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2})$$

$G(x, y) = e^{-\frac{(x^2+y^2)}{2\sigma^2}}$ $\frac{\partial G}{\partial x} = -\frac{x}{\sigma^2} \times e^{-\frac{(x^2+y^2)}{2\sigma^2}} = -\frac{x}{\sigma^2} \times e^{-\frac{(x^2+y^2)}{2\sigma^2}}$ $\frac{\partial^2 G}{\partial x^2} = -\frac{1}{\sigma^2} \times e^{-\frac{(x^2+y^2)}{2\sigma^2}} + \left[ + \frac{x^2}{\sigma^4} \times e^{-\frac{(x^2+y^2)}{2\sigma^2}} \right]$ $= \frac{1}{\sigma^2} \times e^{-\frac{(x^2+y^2)}{2\sigma^2}} \times \left( \frac{x^2}{\sigma^2} - 1 \right)$ $\frac{\partial G}{\partial y} = -\frac{y}{\sigma^2} \times e^{-\frac{(x^2+y^2)}{2\sigma^2}}$ $\frac{\partial^2 G}{\partial y^2} = \frac{1}{\sigma^2} \times e^{-\frac{(x^2+y^2)}{2\sigma^2}} \times \left( \frac{y^2}{\sigma^2} - 1 \right)$	$\nabla^2 G = \frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2}$ $= \left[ \frac{1}{\sigma^2} \times e^{-\frac{(x^2+y^2)}{2\sigma^2}} \times \left( \frac{x^2}{\sigma^2} - 1 \right) \right] + \left[ \frac{1}{\sigma^2} \times e^{-\frac{(x^2+y^2)}{2\sigma^2}} \times \left( \frac{y^2}{\sigma^2} - 1 \right) \right]$ $= \frac{1}{\sigma^2} \times e^{-\frac{(x^2+y^2)}{2\sigma^2}} \times \left( \frac{x^2+y^2}{\sigma^2} - 2 \right)$
---	--

20. (2%) A geometric transform is a vector function  $T$  that maps the pixel  $(x, y)$  to a new position  $(x', y')$ .  $T = (T_x, T_y)$ , where  $x' = T_x(x, y)$ ,  $y' = T_y(x, y)$ . The geometric transform can be approximated by an affine transform.

$$x' = a_0 + a_1x + a_2y; \quad y' = b_0 + b_1x + b_2y$$

A Jacobia determinant  $J$  can provides information about how the co-ordinate system changes.  $J = \left| \frac{\partial(x', y')}{\partial(x, y)} \right|$

Please show the Jacobia determinant for the transform.

The image shows handwritten mathematical work on a piece of paper. It starts with the definition of the Jacobia determinant  $J$  as a 2x2 matrix of partial derivatives:

$$J = \begin{vmatrix} \frac{\partial x'}{\partial x} & \frac{\partial x'}{\partial y} \\ \frac{\partial y'}{\partial x} & \frac{\partial y'}{\partial y} \end{vmatrix}$$

Below this, two equations are given:

$$x' = a_0 + a_1x + a_2y \Rightarrow \begin{cases} \frac{\partial x'}{\partial x} = a_1 \\ \frac{\partial x'}{\partial y} = a_2 \end{cases}$$

$$y' = b_0 + b_1x + b_2y \Rightarrow \begin{cases} \frac{\partial y'}{\partial x} = b_1 \\ \frac{\partial y'}{\partial y} = b_2 \end{cases}$$

To the right of these, the Jacobia determinant is simplified:

$$\bar{J} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \underline{a_1b_2 - a_2b_1}$$

- The **Jacobia determinant** for the **bilinear transform**

$$x' = a_0 + a_1x + a_2y + a_3xy$$

$$y' = b_0 + b_1x + b_2y + b_3xy$$

$$J = \begin{vmatrix} \frac{\partial x'}{\partial x} & \frac{\partial x'}{\partial y} \\ \frac{\partial y'}{\partial x} & \frac{\partial y'}{\partial y} \end{vmatrix} = \begin{vmatrix} a_1 + a_3y & a_2 + a_3x \\ b_1 + b_3y & b_2 + b_3x \end{vmatrix}$$

$$= (a_1 + a_3y)(b_2 + b_3x) - (b_1 + b_3y)(a_2 + a_3x)$$

$$= a_1b_2 - a_2b_1 + (a_1b_3 - a_3b_1)x + (a_3b_2 - a_2b_3)y$$

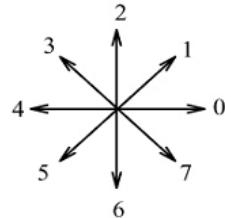
- The **Jacobia determinant** for the **affine transform**

$$J = a_1b_2 - a_2b_1$$

21. (6%) Given an inner boundary tracing algorithm as follows.

**Algorithm: Inner boundary tracing in 8-connectivity**

1. Search the image from top left until a pixel  $P_0$  of a new region is found; this has the minimum column value of all pixels of that region having the minimum row value.  $P_0$  is a starting pixel of the region border.
- Define a variable  $dir$  which stores the direction of the previous move along the border from the previous border element to the current border elements. Assign  
 $dir=7$  if the border is detected in 8-connectivity.
2. Search the  $3 \times 3$  neighborhood of the current pixel in an anti-clockwise direction, beginning the neighborhood search in the pixel positioned in the direction  
 $(dir+7) \bmod 8$  if  $dir$  is even. (8-connectivity)  
 $(dir+6) \bmod 8$  if  $dir$  is odd. (8-connectivity)
- The first pixel found with the same value as the current pixel is a new boundary element  $P_n$ . Update the  $dir$  value.
3. If the current boundary element  $P_n$  is equal to the second border element  $P_1$ , and if the previous border element  $P_{n-1}$  is equal to  $P_0$ , stop. Otherwise repeat step 2.
4. Pixels  $P_0, P_1, \dots, P_{n-2}$  are now the detected inner border.



**Questions:**

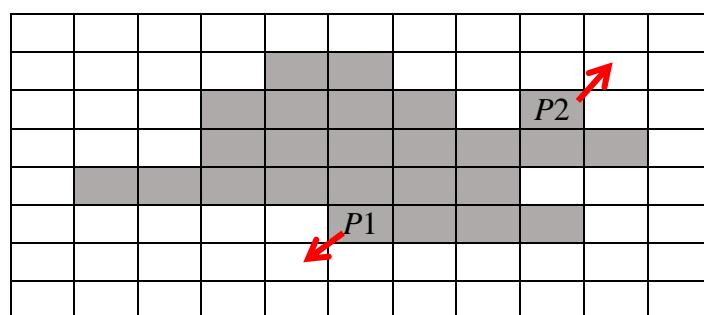
(1) (2%) What are the  $dir$  values at  $P1$  and  $P2$ ?

ANS:

P1: 的  $dir = 7$

P2: 的  $dir = 3$

(2) (4%) What is the starting search pixel after  $P1$  and  $P2$ ? Please mark them in the image.



P1: 下一個: 7 為奇數  $\Rightarrow (7+6)\%8=5$

P2: 下一個: 3 為奇數  $\Rightarrow (3+6)\%8=1$

22. (3%) Here define four kind of inner boundary pixels of a region  $R$ :

If  $Q$  denote outside the region  $R$ , then pixel  $P \in R$  is

- a LEFT pixel of  $R$  if  $P_4(P) \in Q$
- a RIGHT pixel of  $R$  if  $P_0(P) \in Q$
- an UPPER pixel of  $R$  if  $P_2(P) \in Q$
- a LOWER pixel of  $R$  if  $P_6(P) \in Q$

3	2	1
4	$P$	0
5	6	7

For example,  $P_4(P)$  denotes the pixel immediately to the left of pixel  $P$ .

Let  $LEFT(R)$ ,  $RIGHT(R)$ ,  $UPPER(R)$ ,  $LOWER(R)$  represent the corresponding subsets of  $R$ .

How to define the extended boundary EB?

$$EB = \{P : P \in LEFT(R)\} \cup \{P : P \in UPPER(R)\}$$

$$\cup \{ \underline{\text{P0}} : P \in RIGHT(R) \} \cup \{ \underline{\text{P7}} : P \in RIGHT(R) \}$$

$$\cup \{ \underline{\text{P6}} : P \in LOWER(R) \}$$