



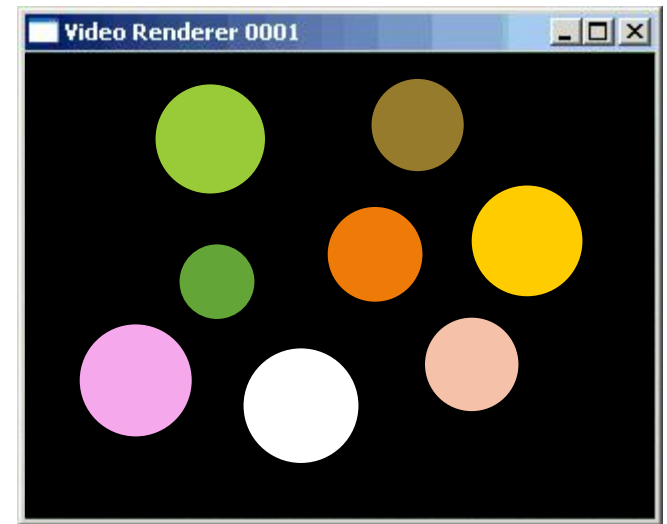
Phân Đoạn Ảnh (1)

SEGMENTATION

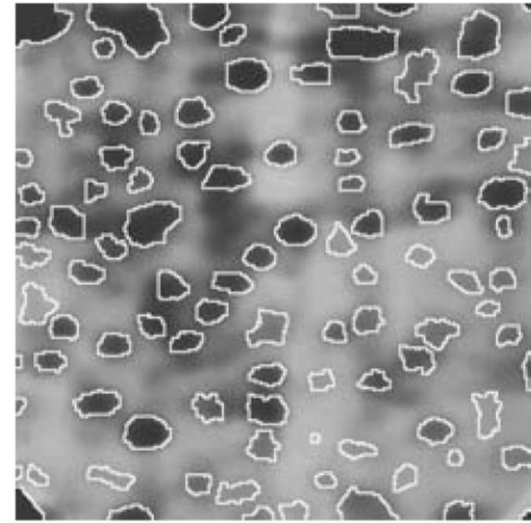
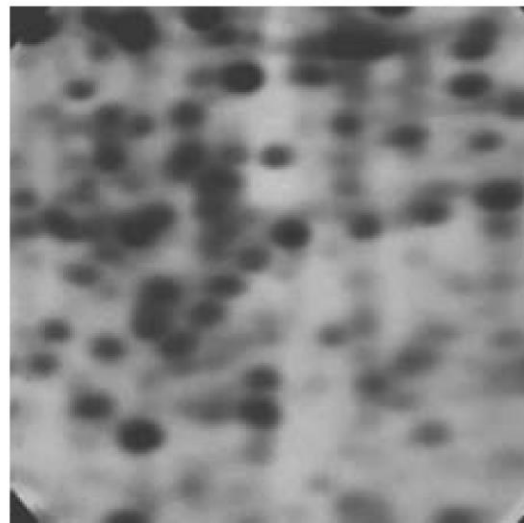
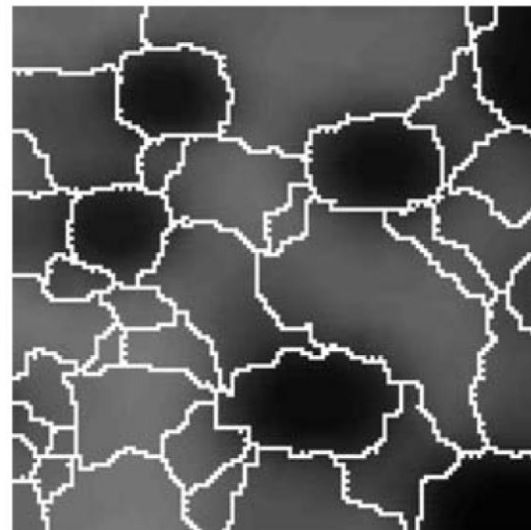
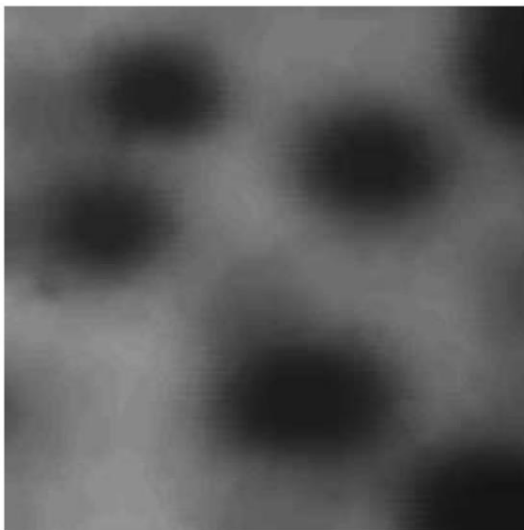
Phân đoạn ảnh là gì ?

- ❑ Phân đoạn ảnh là quá trình chia ảnh thành các thành phần hay các đối tượng khác nhau (Gonzales).
- ❑ Phân đoạn ảnh là quá trình nhóm các pixel có cùng thuộc tính (Efford)
- ❑ Mức độ phân chia sẽ phụ thuộc vào từng bài toán.

Phân đoạn ảnh là gì ?

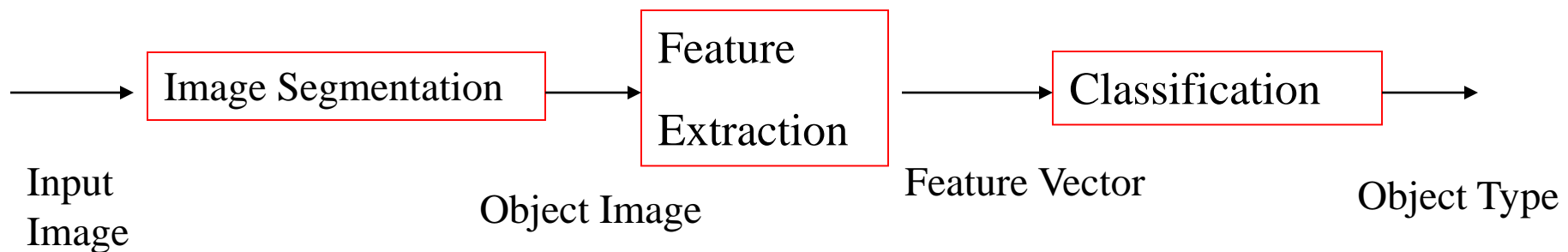


Phân đoạn ảnh là gì ?



Phân đoạn ảnh là gì ?

- ❑ Phân đoạn ảnh là một trong những bước quan trọng nhất của quá trình phân tích ảnh.
- ❑ Thông thường, phân đoạn ảnh thường được xét như công đoạn đầu tiên trong bài toán nhận dạng mẫu (pattern recognition).



Hai loại phân đoạn ảnh

- ❑ **Phân đoạn hoàn toàn:** phân chia ảnh thành một tập các vùng độc lập tương ứng với các đối tượng trong ảnh.
- ❑ **Phân đoạn cục bộ:** các vùng được phân chia không tương ứng với các đối tượng. Ảnh được chia thành các vùng thuần nhất theo một thuộc tính được chọn: độ sáng, màu, cấu trúc,...

Hai loại phân đoạn ảnh



Tổng quan về phân đoạn ảnh

1. Phân đoạn dựa trên cạnh Edge

- Tìm đường biên giữa các vùng kề nhau

2. Phân đoạn dựa trên ngưỡng (Threshold)

- Tìm các vùng bằng cách nhóm các pixel có giá trị mức xám giống nhau

3. Phân đoạn dựa trên vùng (Region)

- Tìm các vùng một cách trực tiếp bằng phương pháp tăng trưởng (growing) hoặc phân chia (splitting)

4. Phân đoạn dựa trên sự chuyển động (Motion)

- Tìm các vùng bằng cách so sánh các frame liên tiếp trong một đoạn video để xác định các vùng tương ứng với sự di chuyển của đối tượng

1. Edge-Based Segmentation

- ❑ Sự không liên tục (dis-continuity) trong một vùng được xem như đường biên của vùng.
- ❑ Sự không liên tục trong ảnh có thể là:
 - Points
 - Lines
 - Edges
- ❑ Các kĩ thuật để phân đoạn:
 - Point detection
 - Edge (pixel) detection
 - Edge formation from edge pixels
 - ❖ Edge linking
 - ❖ Hough transformation

1.1. Dò tìm điểm (Point)

Sử dụng mặt nạ (mask/template)

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

trong đó w_i : hệ số mặt nạ

Đáp ứng (response) của mặt nạ tại một điểm bất kỳ:

$$R = \sum_{i=1}^9 w_i z_i$$

z_i là giá trị mức xám của pixel.

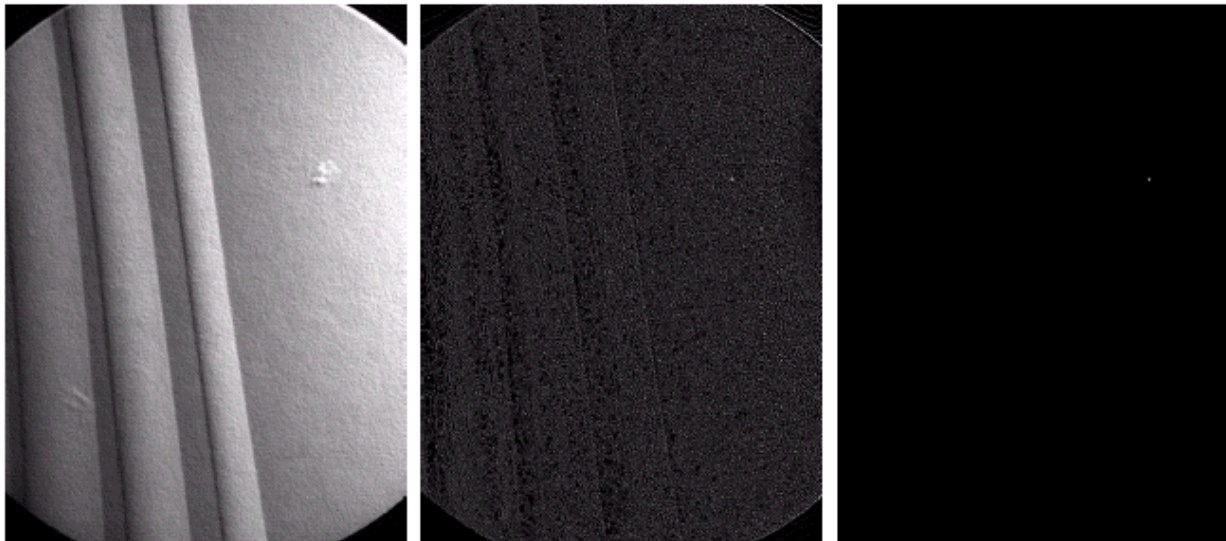
Chú ý về các vấn đề tràn số tại các pixel tại biên.

1.1. Dò tìm điểm (Point)

Dò tìm các điểm kích thước 1 pixel.

Điểm được dò thấy nếu $|R| > T$, T là giá trị không âm và phụ thuộc vào từng bài toán.

-1	-1	-1
-1	8	-1
-1	-1	-1

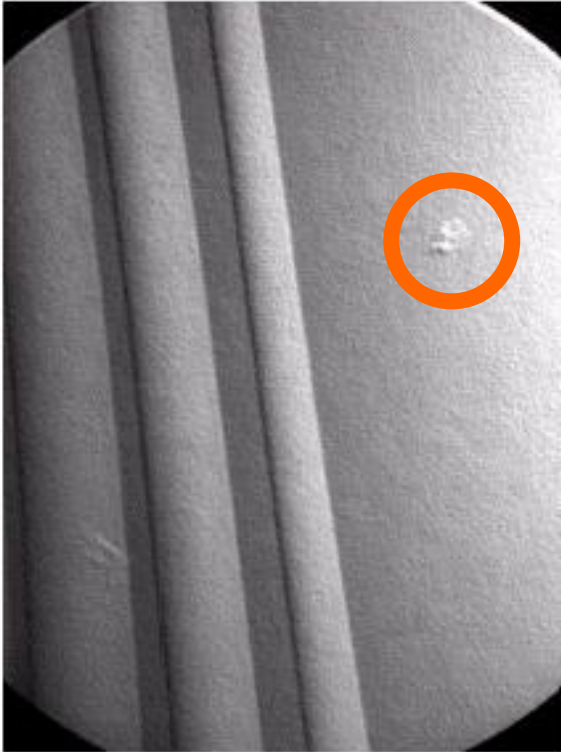


a
b c d

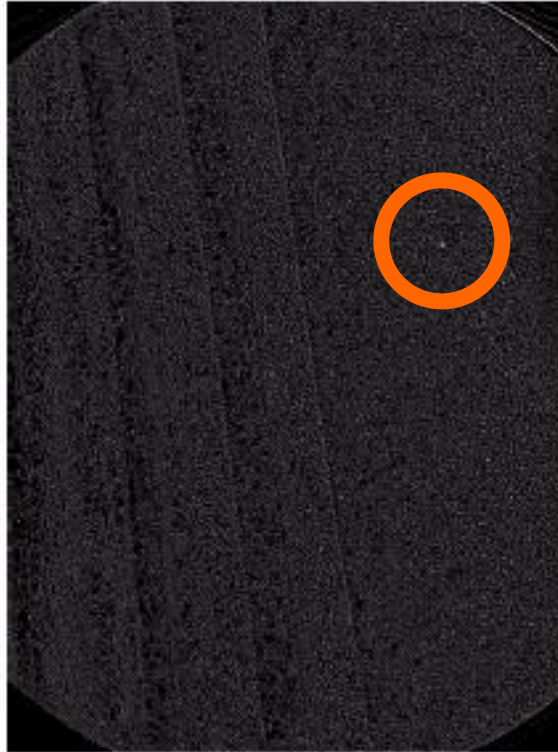
FIGURE 10.2

(a) Point detection mask.
(b) X-ray image of a turbine blade with a porosity.
(c) Result of point detection.
(d) Result of using Eq. (10.1-2).
(Original image courtesy of X-TEK Systems Ltd.)

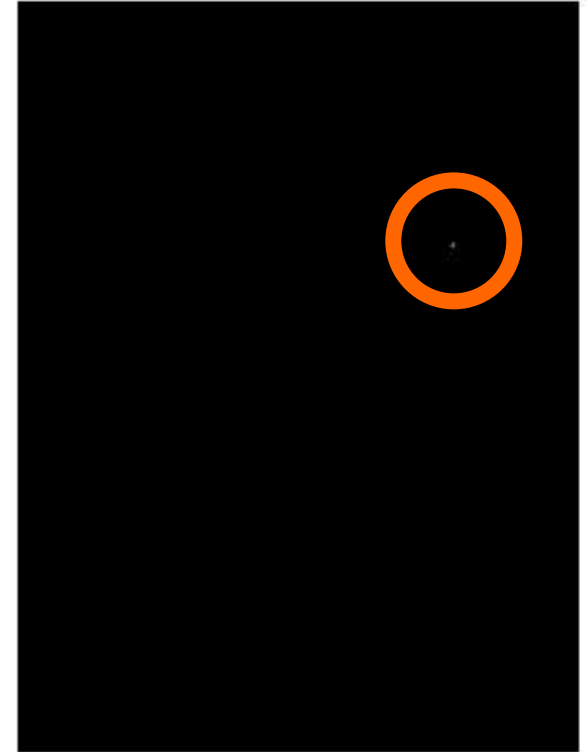
1.1. Dò tìm điểm (Point)



X-ray image of
a turbine blade



Result of point
detection



Result of
thresholding

1.1. Dò tìm điểm (Point)

Ví dụ: Trường hợp ảnh nhị phân

Để dò tìm được các pixel có dạng sau thì ta chọn $T=?$

0	0	0
0	1	0
0	0	0

$$|R|=8$$

0	0	0
0	1	0
0	0	1

$$|R|=7$$

Chọn giá trị $T=7$

1.2. Dò tìm đoạn thẳng (độ rộng 1 pixel)

FIGURE 10.3 Line masks.

-1	-1	-1	-1	-1	2	-1	2	-1	2	-1	-1
2	2	2	-1	2	-1	-1	2	-1	-1	2	-1
-1	-1	-1	2	-1	-1	-1	2	-1	-1	-1	2
Horizontal			+45°			Vertical			-45°		

Nhận xét:

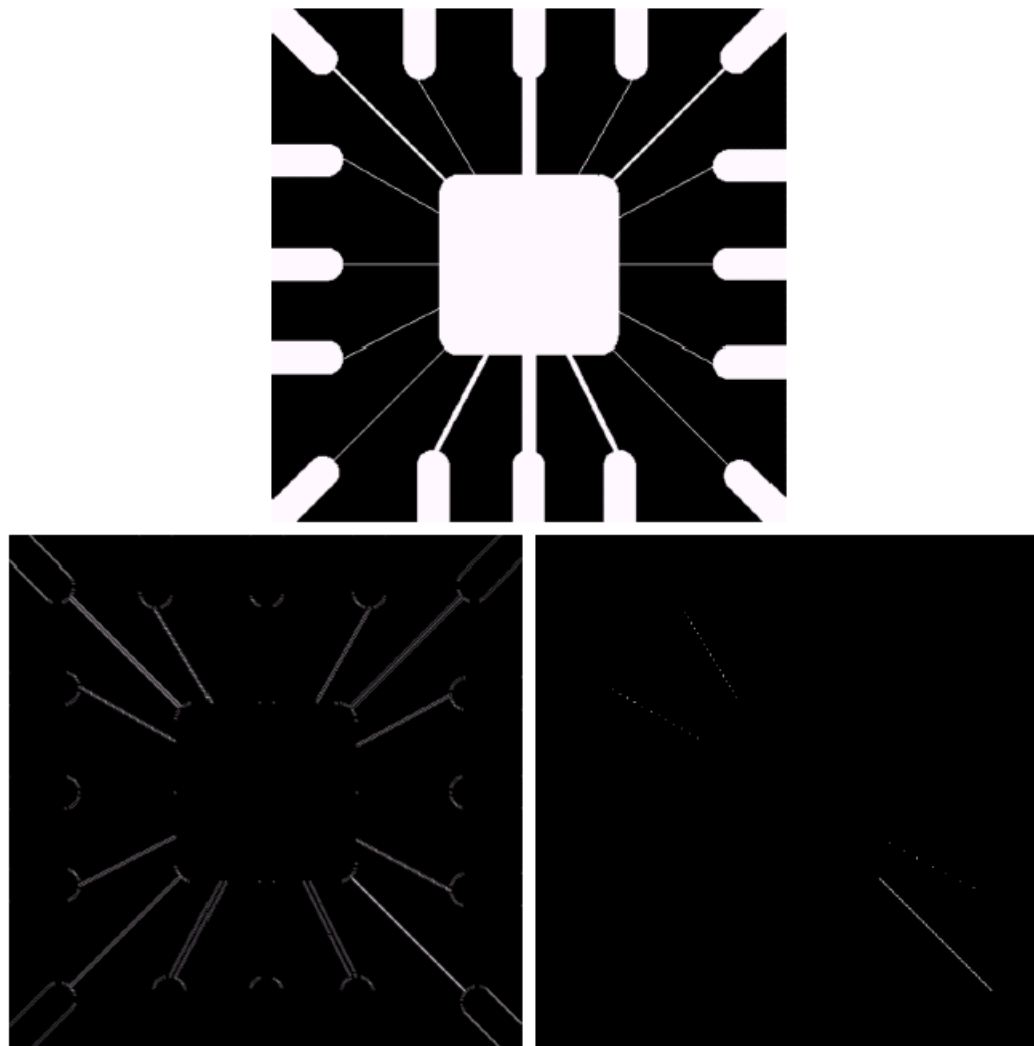
- Tổng các hệ số trong mặt nạ là zero
- Đối với vùng thuần nhất thì nếu thiết kế mặt nạ sao cho $\sum w_i = 0$ thì hiển nhiên ta không tìm thấy đối tượng bất chấp giá trị T ($T \geq 0$)

Việc dò tìm này có thể dựa trên một tập giả thuyết gồm nhiều mặt nạ cho trước W , rồi tính tập đáp ứng R .

- Nếu $\exists R_i : |R_i| > |R_j| \quad \forall i \neq j$

Thì ta nói pixel tại vị trí đang xét có tính chất W_i . (Chọn đáp ứng lớn nhất tại vị trí đang xét).

1.2. Dò tìm đoạn thẳng (độ rộng 1 pixel)



a
b c

FIGURE 10.4

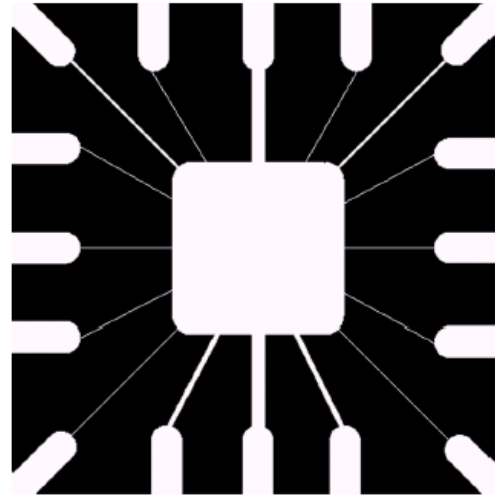
Illustration of line detection.

(a) Binary wire-bond mask.

(b) Absolute value of result after processing with -45° line detector.

(c) Result of thresholding image (b).

1.2. Dò tìm đoạn thẳng (độ rộng 1 pixel)



Binary image of a wire bond mask

After
processing
with -45° line
detector



Result of
thresholding
filtering result

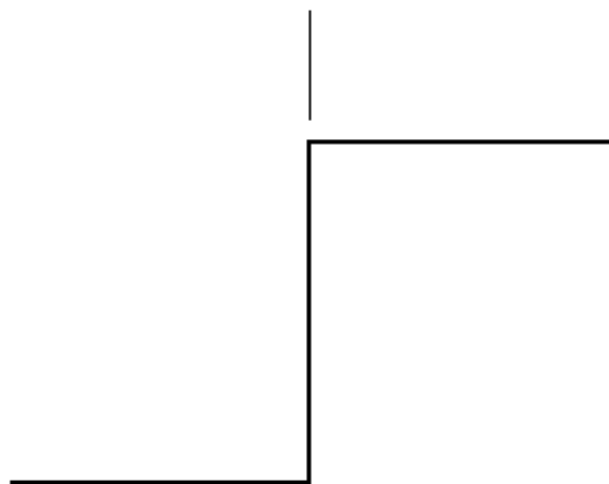


1.3. Dò tìm cạnh (Edge Detection)

- ❑ Dò điểm và đường thẳng là trường hợp đặc biệt của dò tìm cạnh.
- ❑ Độ thay đổi giữa đối tượng và nền thông qua đường biên.
 - Đây là phương pháp thông dụng nhất để dò tìm tính bất liên tục vì phương pháp sử dụng mặt nạ là quá lý tưởng.
 - Định nghĩa cạnh: cạnh là biên giữa hai vùng có giá trị mức xám khác biệt nhau.
 - Vì ta đang xét tốc độ thay đổi của mức xám → dùng đạo hàm.

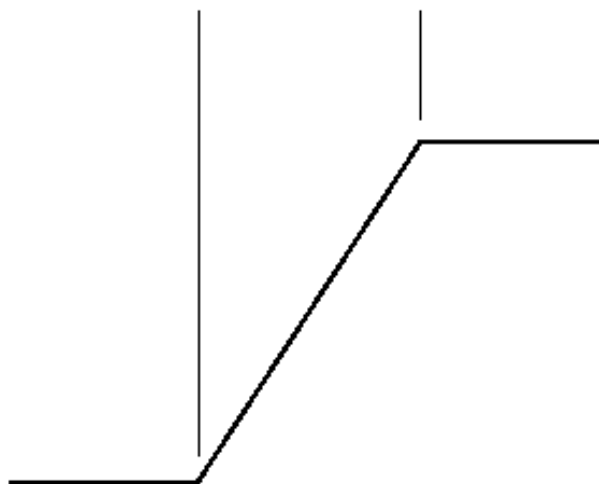
1.3. Dò tìm cạnh (Edge Detection)

Model of an ideal digital edge



Gray-level profile
of a horizontal line
through the image

Model of a ramp digital edge



Gray-level profile
of a horizontal line
through the image

a b

FIGURE 10.5

(a) Model of an ideal digital edge.
(b) Model of a ramp edge. The slope of the ramp is proportional to the degree of blurring in the edge.

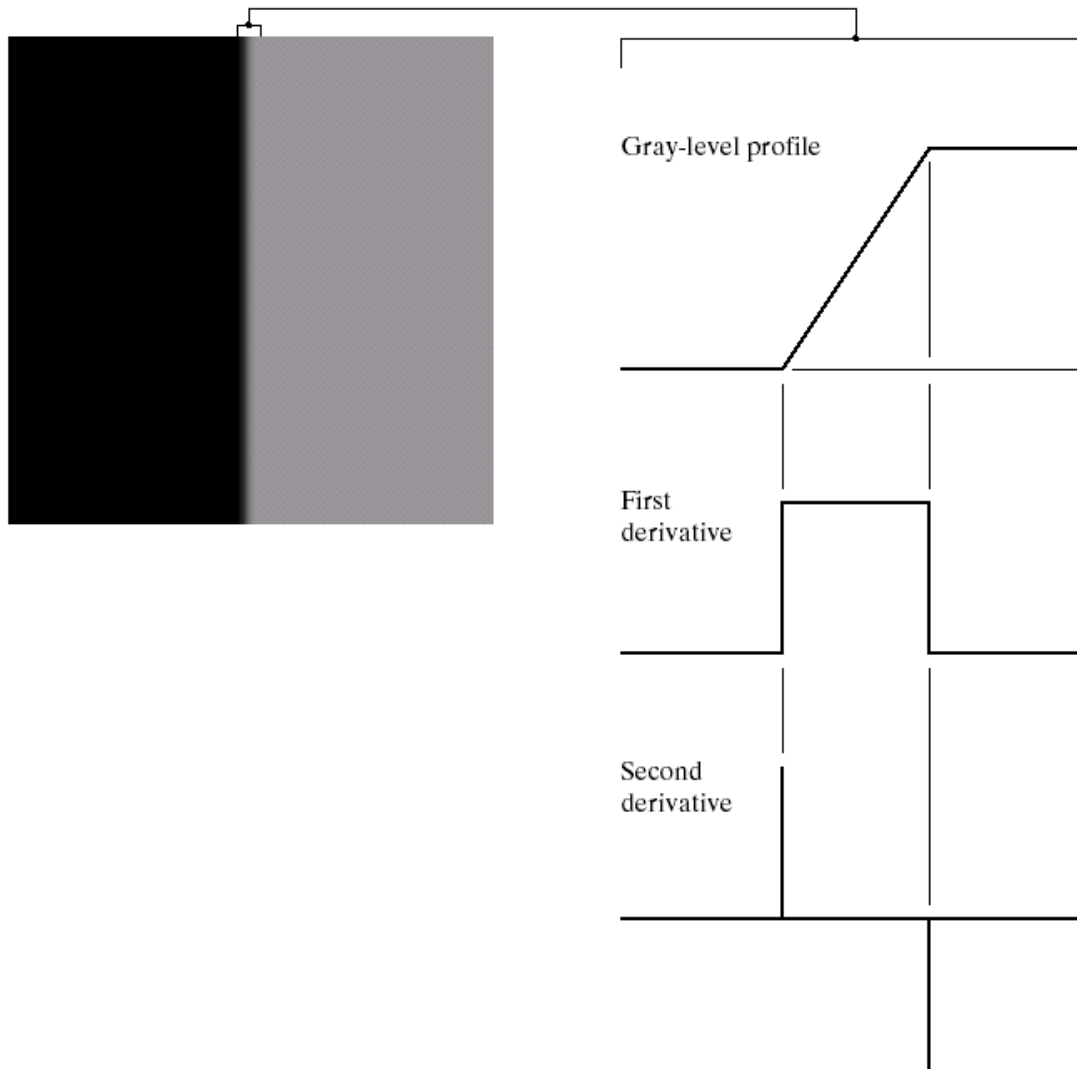
1.3. Dò tìm cạnh (Edge Detection)

a b

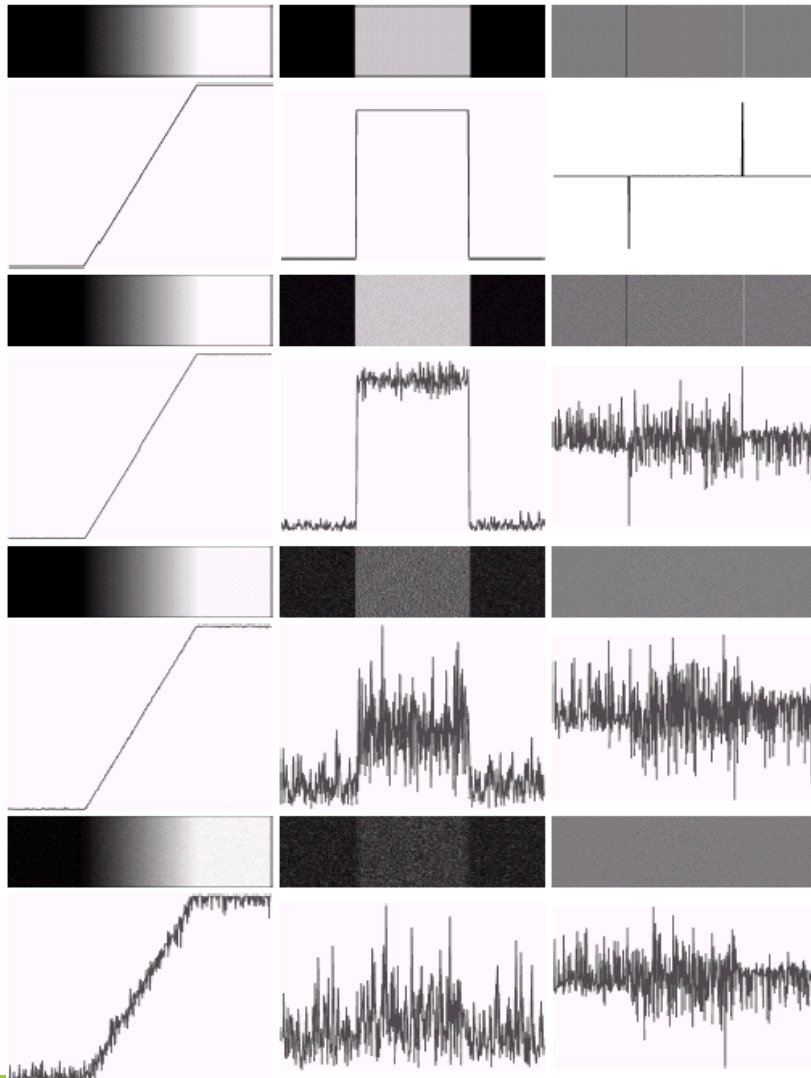
FIGURE 10.6

(a) Two regions separated by a vertical edge.

(b) Detail near the edge, showing a gray-level profile, and the first and second derivatives of the profile.



1.3. Dò tìm cạnh (Edge Detection)



Nhận xét:

- ❖ Đạo hàm cấp 1 dương khi chuyển từ tối sang sáng.
- ❖ Đạo hàm cấp 1 âm khi chuyển từ sáng sang tối.
- ❖ Đạo hàm cấp 2 đổi dấu tại trung điểm của sự chuyển dịch dương ở phần tối và âm ở phần sáng.
- ❖ => Độ lớn của đạo hàm cấp 1 có thể được sử dụng để dò tìm cạnh ảnh.

FIGURE 10.7 First column: images and gray-level profiles of a ramp edge corrupted by random Gaussian noise of mean 0 and $\sigma = 0.0, 0.1, 1.0$, and 10.0 , respectively. Second column: first-derivative images and gray-level profiles. Third column: second-derivative images and gray-level profiles.

a
b
c
d

1.3. Dò tìm cạnh (Edge Detection)

- ❑ Đối với ảnh 2 chiều thì gradient (đạo hàm theo hai thành phần x và y) được sử dụng thay cho đạo hàm cấp 1 và toán tử Laplace thay cho đạo hàm cấp 2.
- ❑ Gradient của hàm liên tục $f(x,y)$ tại (x,y) là vector: ∇f được gọi là “hướng của tốc độ thay đổi cực đại của f tại (x,y) ”.
- ❑ Độ lớn của gradient, trong xử lý ảnh thường được gọi tắt là “gradient”.

1.3. Dò tìm cạnh (Edge Detection)

$$|\nabla f| = \nabla f = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad (1)$$

Là “độ tăng cực đại của $f(x,y)$ trên từng đơn vị khoảng cách theo hướng ∇f ”

Đôi khi ∇f được tính xấp xỉ là:

$$\nabla f = \left| \frac{\partial f}{\partial x} \right| + \left| \frac{\partial f}{\partial y} \right| \quad (2)$$

Hướng của vector gradient:

$$\alpha(x, y) = a \tan\left(\frac{\partial f}{\partial y} / \frac{\partial f}{\partial x}\right) \quad (3)$$

Vì ảnh là rời rạc nên đạo hàm tại 1 pixel bất kỳ phải được xấp xỉ.

1.3.0. Toán tử Robert

Tương đương hai cửa sổ sau:

-1	0
0	1

0	-1
1	0

1.3.1. Toán tử ∇_1 :

Đây là một xấp xỉ đơn giản nhất:

$$\nabla_{1x} = I(x, y) - I(x-1, y)$$

$$\nabla_{1y} = I(x, y) - I(x, y-1)$$

Tương đương hai cửa sổ sau:

0	0	0
-1	1	0
0	0	0

0	-1	0
0	1	0
0	0	0

1.3.1. Toán tử ∇_1 :

Về cài đặt:

- Có thể dùng thuật toán tích chập nhanh 3x3 nhưng có thể thao tác trực tiếp trên ma trận ảnh nguồn, bắt đầu từ pixel cuối để đảm bảo tính toàn vẹn dữ liệu.
- Chỉ mất biên trên và biên trái (do chỉ đi từ dưới lên và từ phải sang trái).

1.3.1. Toán tử ∇_1 :

Nhược điểm của ∇_1 :

- Không tính gradient tại (x,y) mà tại
- Nhạy cảm đối với nhiễu.

$$\left(x - \frac{1}{2}, y - \frac{1}{2}\right)$$

Xác định cạnh:

- $I(x,y)$ được xem là một pixel thuộc cạnh nếu $\nabla f > T$, trong đó T có thể là giá trị giữa của vùng mức xám, ∇f có thể dùng công thức (1) hay (2).

1.3.1. Toán tử ∇_1 :

Thuật toán:

- Tính ∇_1 cho từng pixel và lưu ngược lại ma trận ảnh nguồn dùng công thức (1) hoặc (2).
- Tính ngưỡng $T=(\text{pixel}_{\max}-\text{pixel}_{\min})/2$.
- Gán biên trên và biên trái về 0.
- Đối với các pixel khác:
 - $I(x,y)=\text{pixel}_{\max}$ nếu $\nabla f > T$, ngược lại:
 - $I(x,y)=\text{pixel}_{\min}$

1.3.2. Toán tử ∇_2 :

Được xấp xỉ :

- $\nabla_{2x} = I(x+1, y) - I(x-1, y)$
- $\nabla_{2y} = I(x, y+1) - I(x, y-1)$

Tương đương hai cửa sổ sau:

0	0	0
-1	0	1
0	0	0

0	-1	0
0	0	0
0	1	0

→ Về mặt liên tục, do lấy 2 pixel trừ nhau dẫn đến không chính xác và có thể khác hoàn toàn.

1.3.3. Toán tử Prewitt:

Tương đương hai cửa sổ sau:

-1	0	1
-1	0	1
-1	0	1

∇_x

-1	-1	-1
0	0	0
1	1	1

∇_y

1.3.4. Toán tử Sobel:

Tương đương hai cửa sổ sau:

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

∇_x

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

∇_y

1.3.4. Toán tử Sobel:

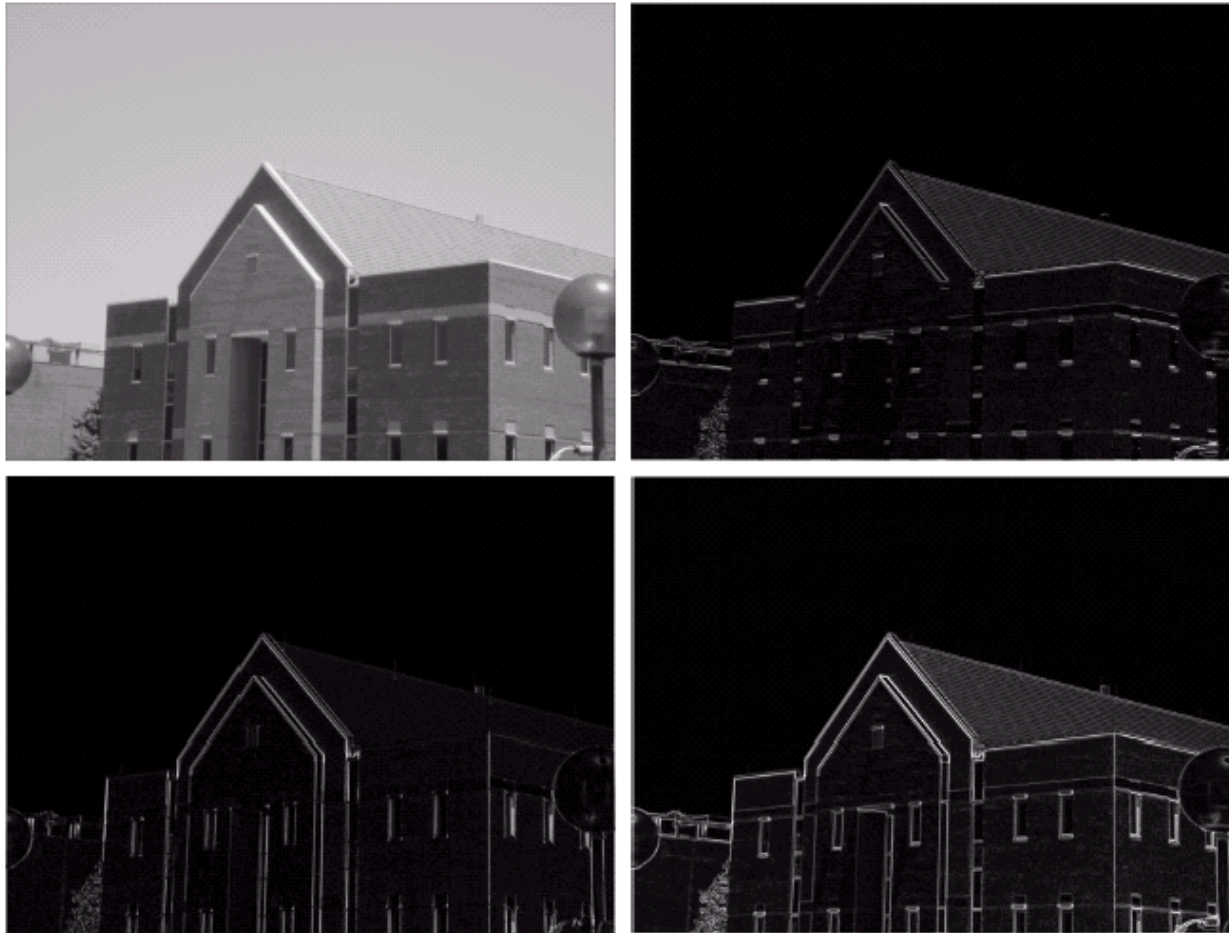
a b
c d

FIGURE 10.10

(a) Original image. (b) $|G_x|$, component of the gradient in the x -direction. (c) $|G_y|$, component in the y -direction. (d) Gradient image, $|G_x| + |G_y|$.



1.3.4. Toán tử Sobel:

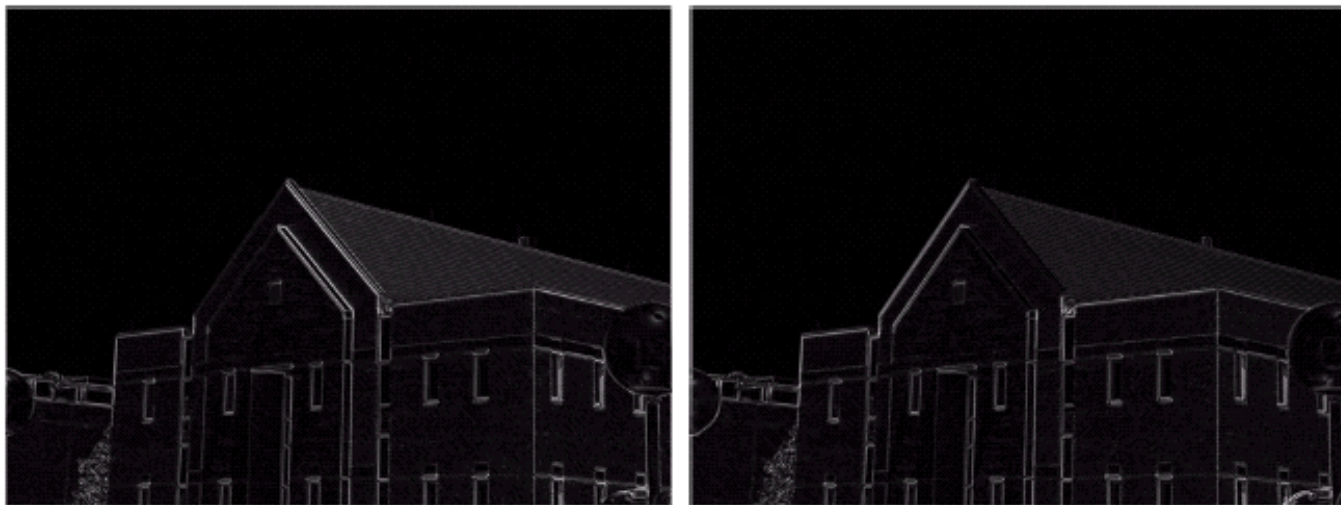


a	b
c	d

FIGURE 10.11

Same sequence as in Fig. 10.10, but with the original image smoothed with a 5×5 averaging filter.

1.3.4. Toán tử Sobel:



a b

FIGURE 10.12

Diagonal edge detection.

(a) Result of using the mask in Fig. 10.9(c).

(b) Result of using the mask in Fig. 10.9(d). The input in both cases was Fig. 10.11(a).

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

Use diagonal Sobel operator shown in figure 10.9(d)

1.3.5. Toán tử Kirsh:

Tương đương với các cửa sổ sau:

-3	-3	5
-3	0	5
-3	-3	5

-3	5	5
-3	0	5
-3	-3	-3

5	5	5
-3	0	-3
-3	-3	-3

5	5	-3
5	0	-3
-3	-3	-3

5	-3	-3
5	0	-3
5	-3	-3

-3	-3	-3
5	0	-3
5	5	-3

-3	-3	-3
-3	0	-3
5	5	5

-3	-3	-3
-3	0	5
-3	5	5

1.3.5. Toán tử Kirsh:

Nhận xét:

- Tổng $\sum w_i = 0$
- Đây không phải là xấp xỉ gradient
- Mô hình kiểu thay đổi mức xám gần cạnh ở nhiều hướng khác nhau.
- Ảnh I được tích chập với tất cả các mặt nạ tại từng vị trí pixel. Đáp ứng của toán tử Kirsh là giá trị lớn nhất trong 8 mặt nạ.

1.4. Mô hình Marr-Hildreth

Vào những năm 70s, David Marr đã nghiên cứu những hiện thị của loài thú có vú và đã tổng kết thành 5 điểm chính:

Các đặc trưng của ảnh tự nhiên ở nhiều thang khác nhau. Chẳng có toán tử nào có thể hoạt động tốt trên mọi thang, vì thế cần có sự kết hợp kết quả của nhiều toán tử trên từng thang.

Một quang cảnh tự nhiên không bao gồm các mẫu nhiễu xạ hoặc các hiệu ứng tựa sóng. Vì thế phải sử dụng một phương pháp lấy trung bình địa phương (smoothing) nào đó.

Bộ lọc làm trơn tối ưu phù hợp với các hệ thống nhìn sinh học là bộ lọc Gaussian.

Khi có sự thay đổi về cường độ (có 1 cạnh) thì đạo hàm cấp 1 có cực trị. Cực trị này tương ứng với giao điểm zero của đạo hàm cấp 2.

Toán tử vi phân độc lập hướng cấp thấp nhất là toán tử Laplacian.

1.4. Mô hình Marr-Hildreth

Từ 5 tổng kết này, Marr đã đề ra thuật toán dò tìm cạnh như sau:

Tích chập ảnh với hàm Gaussian 2 chiều.

Áp dụng toán tử Laplacian lên ảnh xoắn, ảnh kết quả gọi là L .

Các pixel cạnh là các pixel có giao điểm zero trong L .

1.4. Mô hình Marr-Hildreth

Tích chập ảnh I với các hàm Gaussian có độ lệch chuẩn khác nhau. Độ lệch chuẩn σ đóng vai trò các thang khác nhau trong trường hợp này.

- Hàm Gaussian 2 chiều:

trong đó σ : độ lệch chuẩn
$$G_{\sigma}^{(x,y)} = e^{\frac{-x^2+y^2}{2\sigma}}$$

- Để thực hiện thao tác tích chập, hàm Gaussian phải được lấy mẫu để tạo ra một ảnh (cửa sổ) 2 chiều nhỏ.

Toán tử Laplacian:

- Trong không gian liên tục:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

- Trong không gian rời rạc:

0	-1	0
-1	4	-1
0	-1	0

-1	-1	-1
-1	8	-1
-1	-1	-1

1.4. Mô hình Marr-Hildreth

Nhận xét:

Thay vì thực hiện tích chập ảnh rồi áp dụng toán tử Laplacian, ta có thể tính Laplacian của hàm Gaussian trước, rồi lấy mẫu hàm kết quả.

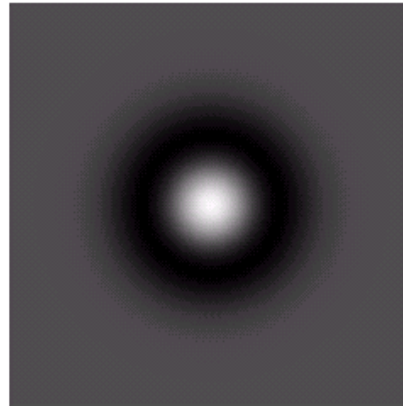
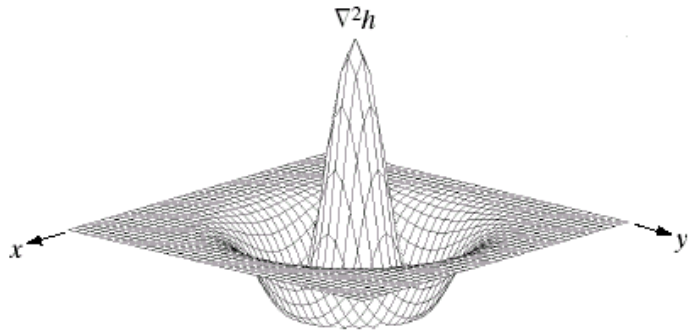
$$\nabla^2 G_\sigma = \frac{r^2 - 2\sigma^2}{\sigma^4} e^{-\frac{r^2}{2\sigma^2}}, r^2 = x^2 + y^2$$

Vấn đề:

Phải chọn nhiều giá trị σ khác nhau như thế nào ? \rightarrow Đưa ra thông số thực nghiệm, chẳng hạn $\sigma_1=1.2$, $\sigma_2=2.8$.

Kích thước cửa sổ ? $\rightarrow w = 3.35\sigma + 0.33$

Laplacian and Mexican Hat



a b
c d

FIGURE 10.14

Laplacian of a Gaussian (LoG).

(a) 3-D plot.

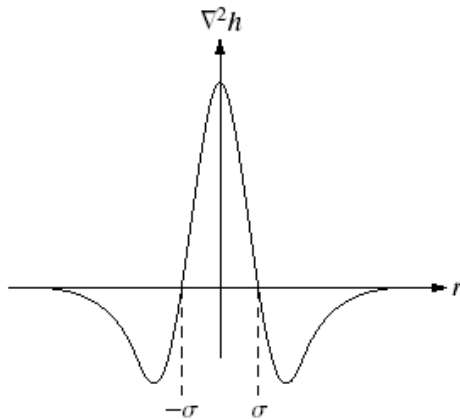
(b) Image (black is negative, gray is the zero plane, and white is positive).

(c) Cross section showing zero crossings.

(d) 5×5 mask approximation to the shape of (a).

0	-1	0
-1	4	-1
0	-1	0

-1	-1	-1
-1	8	-1
-1	-1	-1



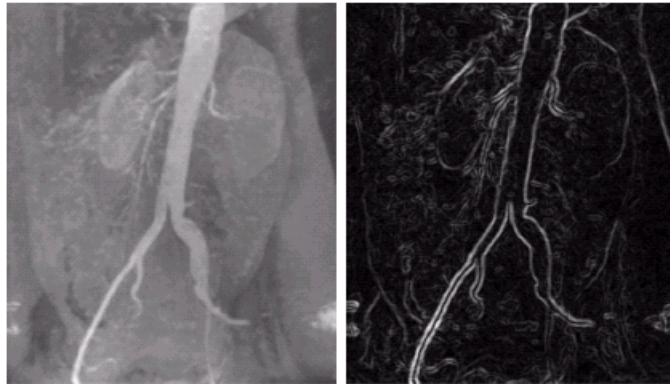
0	0	-1	0	0
0	-1	-2	-1	0
-1	-2	16	-2	-1
0	-1	-2	-1	0
0	0	-1	0	0

LoG operator

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

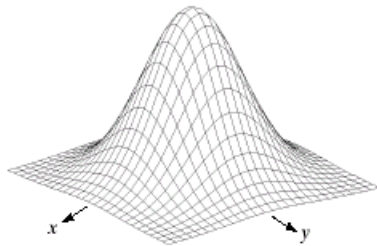
$$\nabla^2 h(r) = \nabla^2 \left(-e^{-r^2/(2\sigma^2)} \right) = \frac{\sigma^2 - r^2}{\sigma^4} e^{-r^2/(2\sigma^2)}$$

1.4. Mô hình Marr-Hildreth



a b
c d
e f g

FIGURE 10.15 (a) Original image. (b) Sobel gradient (shown for comparison). (c) Spatial Gaussian smoothing function. (d) Laplacian mask. (e) LoG. (f) Thresholded LoG. (g) Zero crossings. (Original image courtesy of Dr. David R. Pickens, Department of Radiology and Radiological Sciences, Vanderbilt University Medical Center.)



-1	-1	-1
-1	8	-1
-1	-1	-1



Boundary Extraction

Edge detection classifies individual pixels to be on an edge or not.

Isolated edge pixels is more likely to be noise rather than a true edge.

Adjacent or connected edge pixels should be linked together to form boundary of regions that segment the image.

Edge linking methods:

- Local processing
- Hough transform
- Graphic theoretic method
- Dynamic programming

Local Processing Edge Linking

An edge pixel will be linked to another edge pixel within its own neighborhood if they meet two criteria:

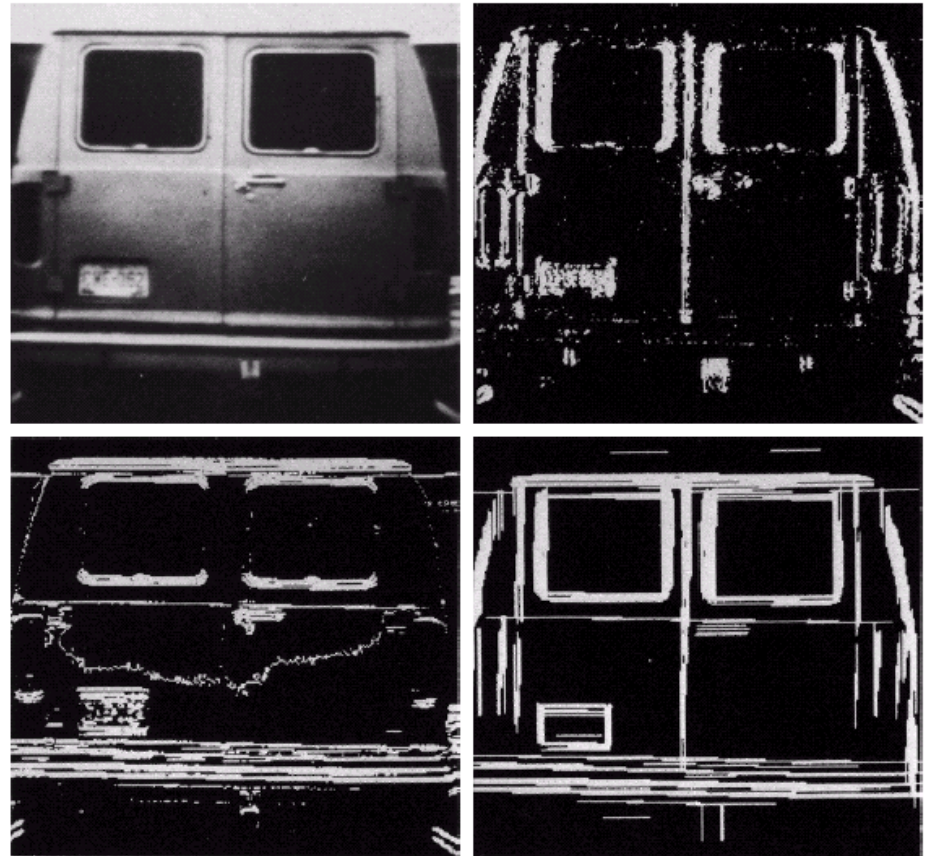
$$|\nabla f(x, y) - \nabla f(x_o, y_o)| \leq E$$

$$|\alpha(x, y) - \alpha(x_o, y_o)| \leq A$$

a	b
c	d

FIGURE 10.16

(a) Input image.
(b) G_y component of the gradient.
(c) G_x component of the gradient.
(d) Result of edge linking. (Courtesy of Perceptics Corporation.)



Global Processing Edge Linking: Hough Transform

Find a subset of n points on an image that lie on the same straight line.

Write each line formed by a pair of these points as

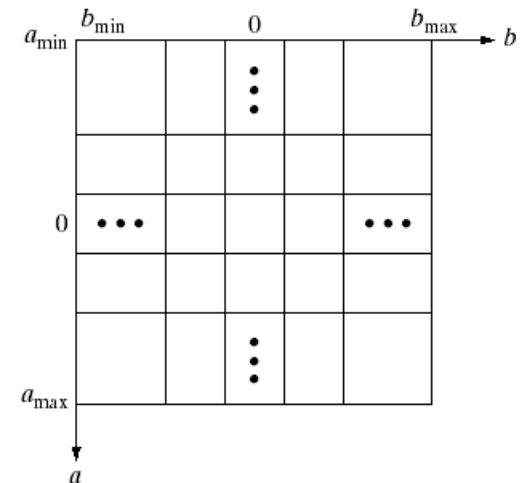
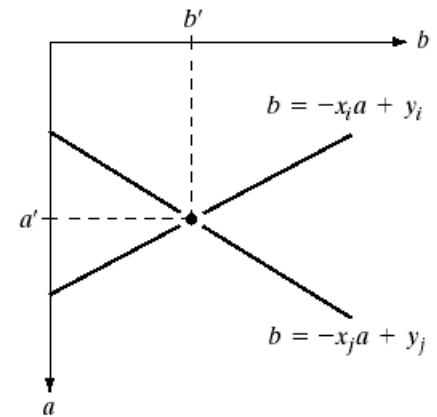
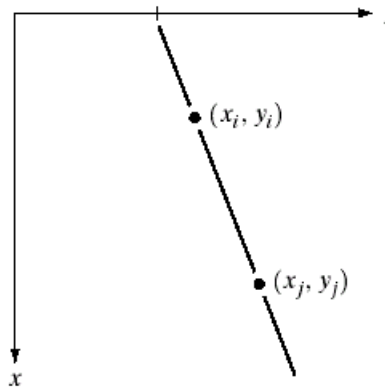
$$y_i = ax_i + b$$

Then plot them on the parameter space (a , b):

$$b = -x_i a + y_i$$

All points (x_i, y_i) on the same line will pass the same parameter space point (a, b) .

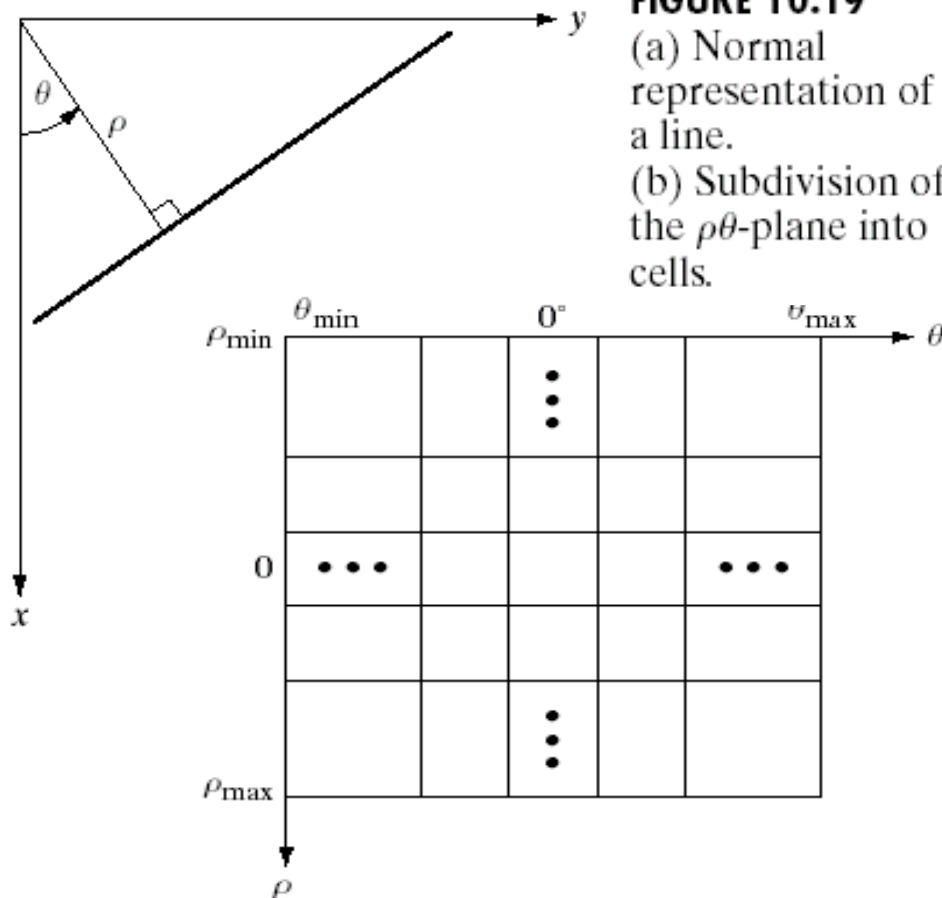
Quantize the parameter space and tally # of times each point falls into the same accumulator cell. The cell count = # of points in the same line.



Hough Transform in (ρ, θ) plane

FIGURE 10.19

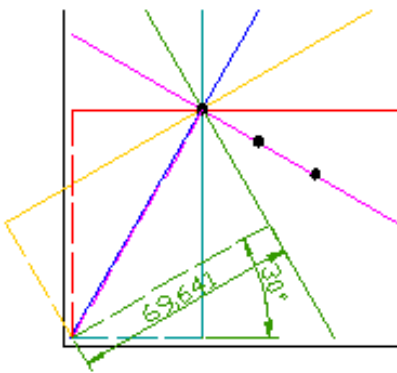
(a) Normal representation of a line.
(b) Subdivision of the $\rho\theta$ -plane into cells.



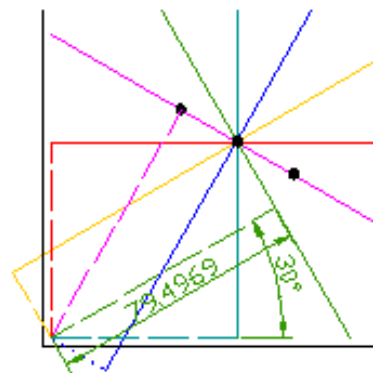
To avoid infinity slope, use polar coordinate to represent a line.

$x \cos \theta + y \sin \theta = \rho$
Q points on the same straight line gives Q sinusoidal curves in (ρ, θ) plane intersecting at the same (ρ_i, θ_i) cell.

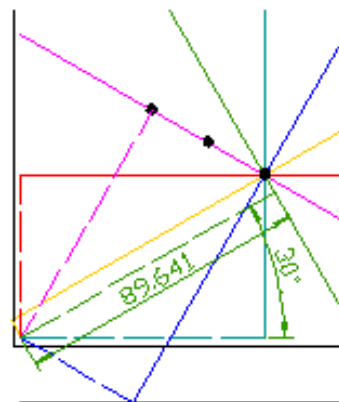
Hough Transform



Angle	Dist.
0	40
30	69.6
60	81.2
90	70
120	40.6
150	0.4



Angle	Dist.
0	57.1
30	79.5
60	80.5
90	60
120	23.4
150	-19.5



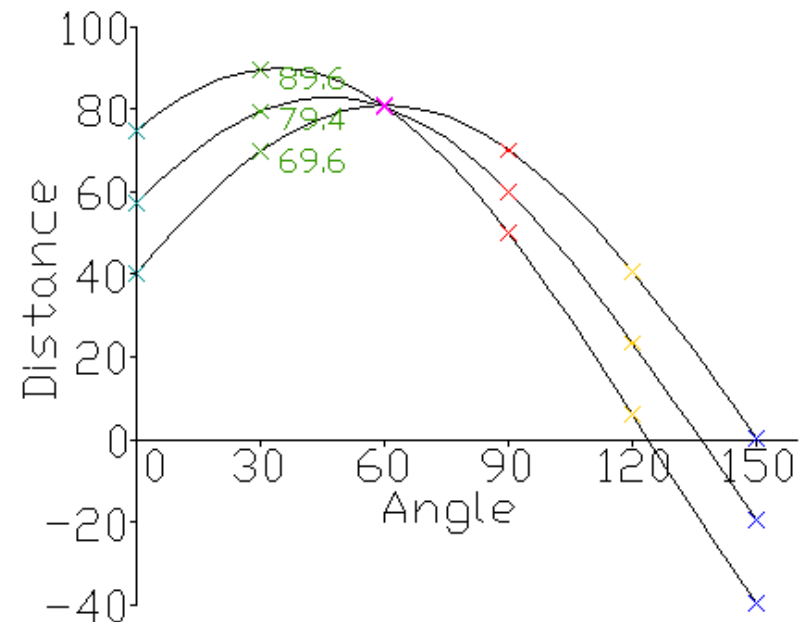
Angle	Dist.
0	74.6
30	89.6
60	80.6
90	50
120	6.0
150	-39.6

For each data point, a number of lines are plotted going through it, all at different angles. These are shown here as solid lines.

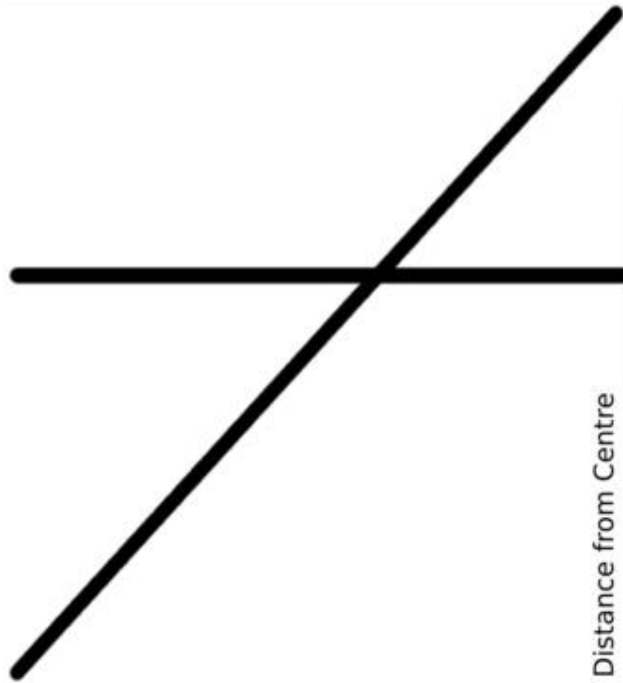
For each solid line a line is plotted which is perpendicular to it and which intersects the origin. These are shown as dashed lines.

The length and angle of each dashed line is measured. In the diagram above, the results are shown in tables.

This is repeated for each data point and graphed as length against angle, called Hough space.



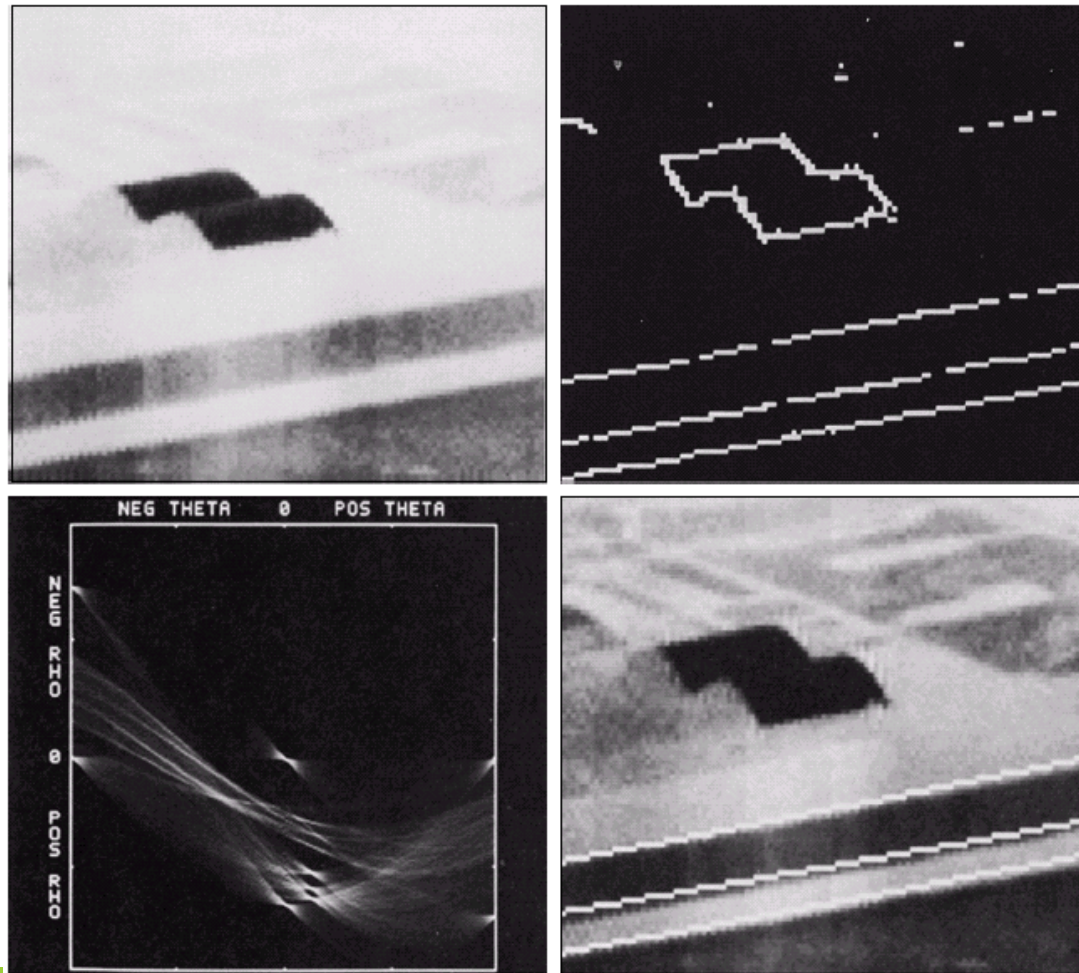
Input Image



Rendering of Transform Results



Example



a	b
c	d

FIGURE 10.21

(a) Infrared image.

(b) Thresholded gradient image.

(c) Hough transform.

(d) Linked pixels.
(Courtesy of Mr. D. R. Cate, Texas Instruments, Inc.)