CHAPTER 10

IMAGE SEGMENTATION

CHAPTER 10: IMAGE SEGMENTATION

- Segmentation subdivides an image into its constituent regions or objects.
- Image segmentation algorithms are generally based on one of 2 basic properties of intensity values:
  - Discontinuity
  - Similarity
- <u>Algorithms based on discontinuity</u>: the approach is to partition an image based on abrupt changes in intensity (e.g., edges).
- <u>Algorithms based on similarity:</u> the approach is to partition an image into regions that are similar according to a predefined set of criteria.
- Segmentation bridges the gap between <u>low-level</u> image processing and <u>high-level</u> image processing.

DETECTION OF DISCONTINUITIES

- 3 basic types of gray-level discontinuities:
  - Points
  - Lines
  - Edges
- The most common way to look for discontinuities is to use a mask.

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

3x3 mask

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

3x3 region of the image

- Response of the mask at any point in the image: $R = \Sigma w_i z_i$

Gray level of the pixel associated with w_i

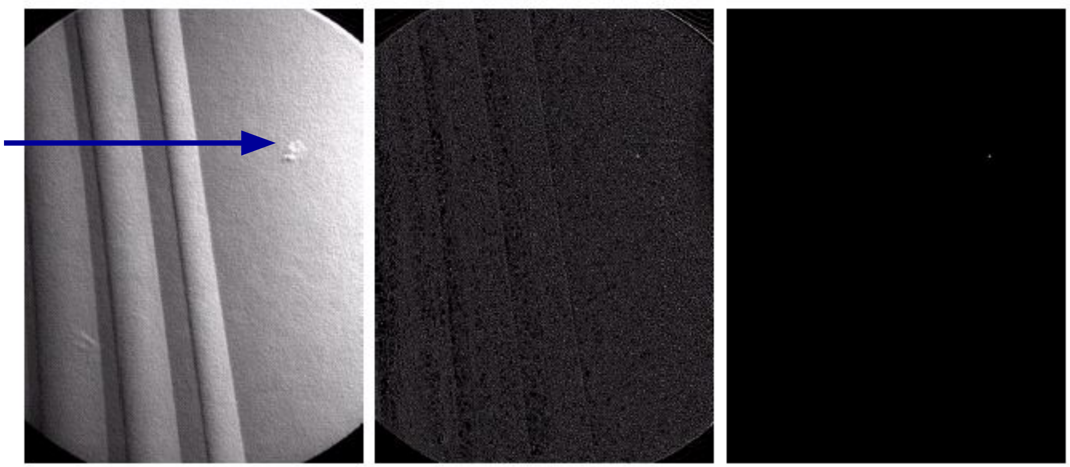
POINT DETECTION

A point is detected at the location on which the mask is centered if $|R| \geq T$, $T \geq 0$.

The sum of the coefficients is 0.

−1	−1	−1
−1	8	−1
−1	−1	−1

A single black pixel embedded within the porosity.



Application of the mask

T = 90% of the highest absolute pixel value in (c)

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.)

LINE DETECTION

Mask responses

R_1 **R_2** **R_3** **R_4**

FIGURE 10.3 Line masks.

Horizontal

−1	−1	−1
2	2	2
−1	−1	−1

+45°

−1	−1	2
−1	2	−1
2	−1	−1

Vertical

−1	2	−1
−1	2	−1
−1	2	−1

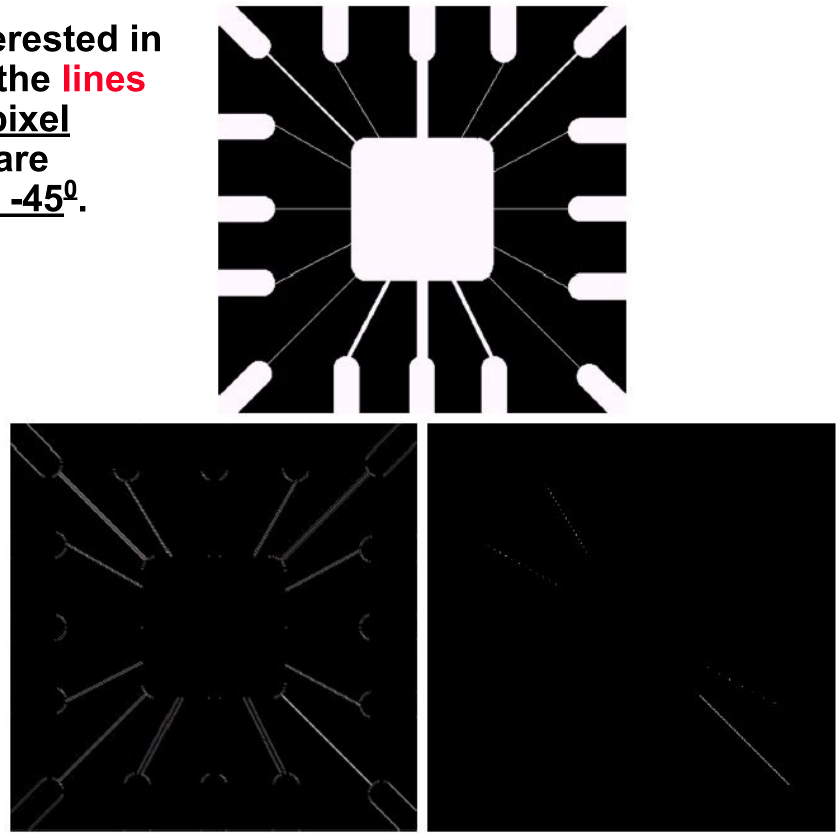
−45°

2	−1	−1
−1	2	−1
−1	−1	2

At a certain point, $|R_i| > |R_j|$, for all $j \neq i$: the point is more likely associated with a line in the direction of mask *i*.

AN EXAMPLE OF LINE DETECTION

We are interested in finding all the **lines** **that are** **<u>1 pixel thick</u>**, **and are** **<u>oriented at -45^{0}</u>**.



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).

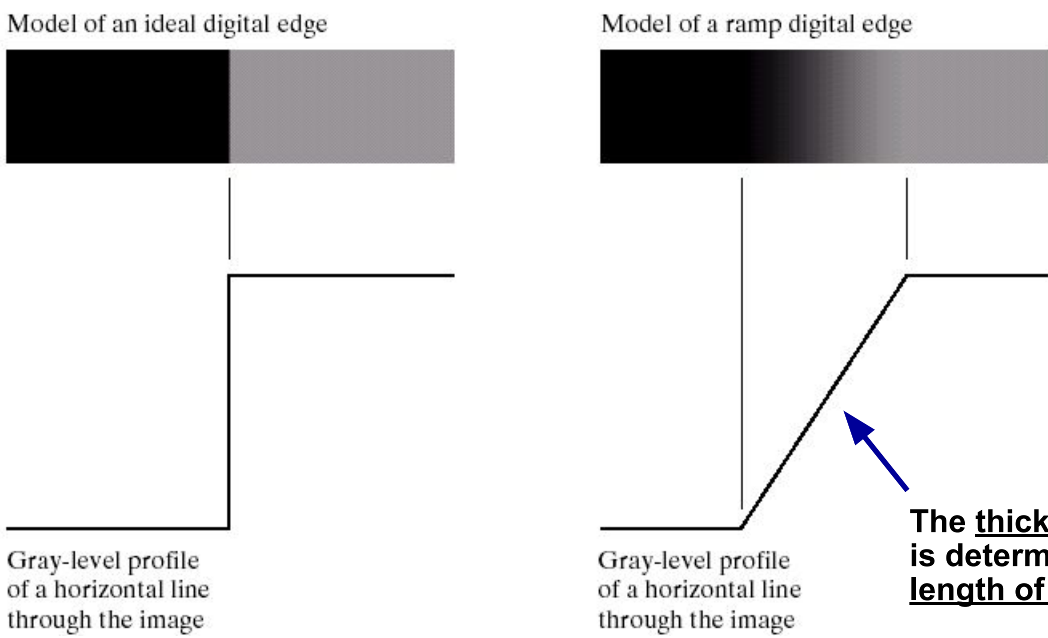
Absolute value of the result

T **= the max value in the image**

EDGE DETECTION

- Edge detection is the most common approach for detecting meaningful discontinuities in gray level.
- Intuitively, an edge is a set of connected pixels that lie on the boundary of 2 regions.
- We will use first and second digital derivatives for edge detection in images.
- Difference between an edge and a boundary:
  - Edge is local.
  - Boundary is global.
- In practice, edges are blurred.
  - The degree of blurring is determined by a number of factors.
    - Quality of the acquisition system
    - Sampling rate
    - Illuminations conditions
- Hence, edges are modeled as having a ramplike profile.

MODELING OF A EDGE



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.

The <u>thickness of the edge</u> is determined by the <u>length of the ramp</u>.

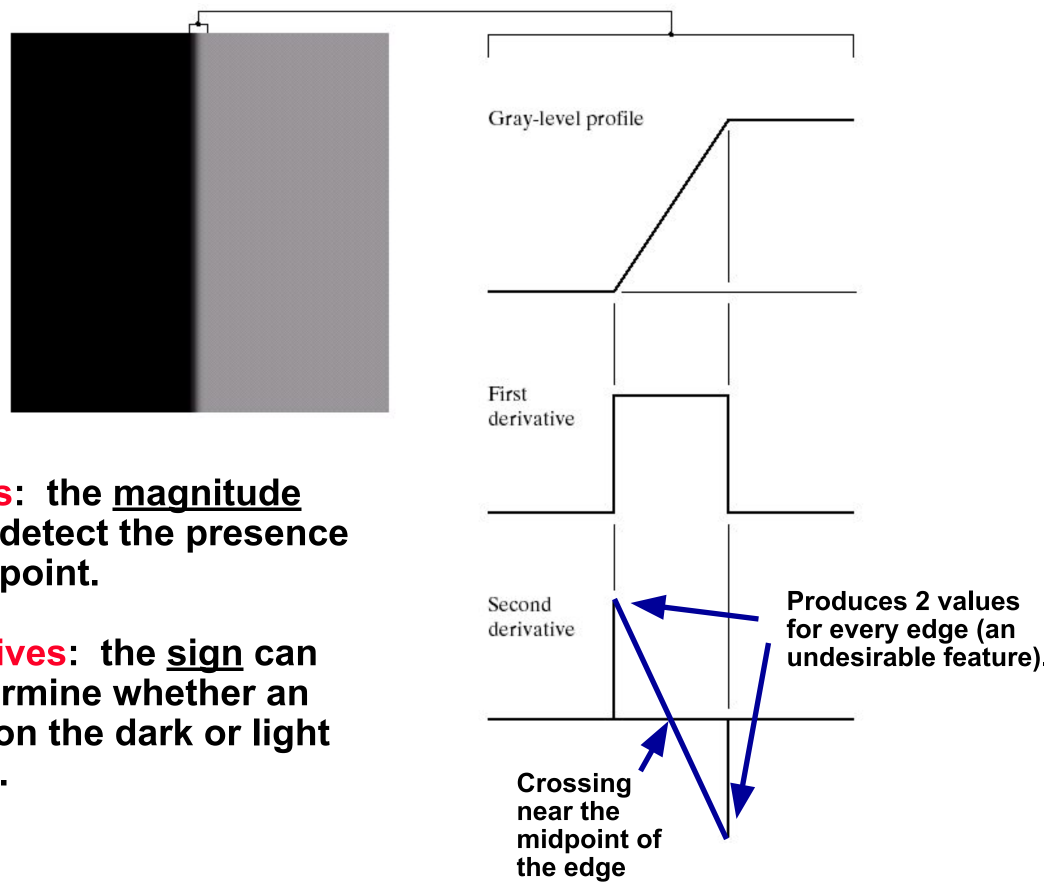
Sharp edges tend to be thin.

Blurred edges tend to be thick.

DETAIL NEAR AN EDGE

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.



First derivatives: the magnitude can be used to detect the presence of an edge at a point.

Second derivatives: the sign can be used to determine whether an edge pixel lies on the dark or light side of an edge.

SENSITIVITY OF DERIVATIVES TO NOISE

Although the noise is almost invisible in the images, both derivatives are sensitive to noise.

Image segment | 1st D | 2nd D

Noise free

Gaussian noise: mean =0, σ=0.1

Gaussian noise: mean =0, σ=1.0

Gaussian noise: mean =0, σ=10.0

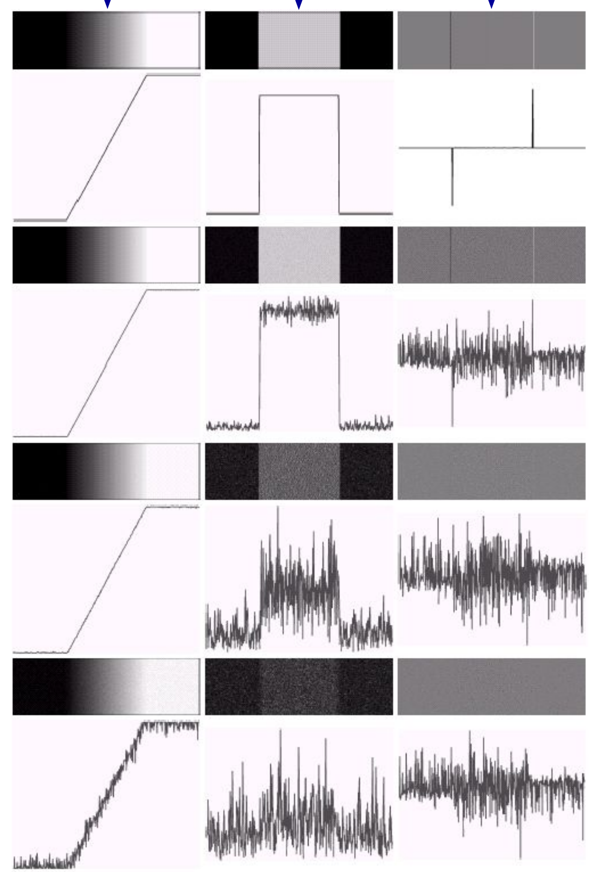


FIGURE 10.7 First column: images and gray-level profiles of a ramp edge corrupted by random Gaussian noise of mean 0 and σ = 0.0, 0.1, 1.0, and 10.0, respectively. Second column: first-derivative images and gray-level profiles. Third column: second-derivative images and gray-level profiles.

a
b
c
d

DEFINITION OF AN EDGE POINT

- To be classified as a meaningful edge point, the transition in gray level associated with that point has to be significantly stronger than the background at that point.
- Two definitions of edge point
  - Its 1^{st} order derivative > T.
  - Zero crossing of its 2^{nd} order derivative.
- Edge: a set of edge points that are connected according to a predefined criterion of connectedness.
- Edge segment: the edge is short relative to the dimensions of the image.
- The edge definitions do NOT guarantee success in finding edges in an image. They simply give us a formalism to look for them.

THE GRADIENT

a	
b	c
d	e
f	g

FIGURE 10.8 A 3 × 3 region of an image (the z's are gray-level values) and various masks used to compute the gradient at point labeled z_5.

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

2x2 filters

$G_x = (z_9 - z_5)$

−1	0
0	1

0	−1
1	0

$G_y = (z_8 - z_6)$

Roberts

$G_x = (z_7+z_8+z_9)-(z_1+z_2+z_3)$

−1	−1	−1
0	0	0
1	1	1

−1	0	1
−1	0	1
−1	0	1

$G_y = (z_3+z_6+z_9)-(z_1+z_4+z_7)$

Prewitt

3x3 filters

$G_x = (z_7+2z_8+z_9)-(z_1+2z_2+z_3)$

−1	−2	−1
0	0	0
1	2	1

−1	0	1
−2	0	2
−1	0	1

$G_y = (z_3+2z_6+z_9)-(z_1+2z_4+z_7)$

Sobel

$\nabla f = [G_x^2+G_y^2]^{1/2}$

(not preferred)

$\nabla f \cong |G_x|+|G_y|$

(more attractive)

$\alpha(x,y) = \tan^{-1}(G_y/G_x)$

DIAGONAL EDGE DETECTION

0	1	1
−1	0	1
−1	−1	0

−1	−1	0
−1	0	1
0	1	1

Prewitt

0	1	2
−1	0	1
−2	−1	0

−2	−1	0
−1	0	1
0	1	2

Sobel

a b
c d

FIGURE 10.9 Prewitt and Sobel masks for detecting diagonal edges.

AN EXAMPLE OF HORIZONTAL & VERTICAL EDGE DETECTION

The roof tile, horizontal brick joints, and horizontal segments of the windows.

1200x1600 pixels

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, $|G_x| + |G_y|$.

The corner near the wall, the lamp post, vertical brick joints, and vertical segments of the windows.

The <u>directionality</u> of the 2 components is evident.

PRINCIPLE EDGE DETECTION

The contribution made to image detail by the wall bricks is significant.
This level of detail is often not desirable. One way to reduce it to smooth the image.

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.

The response of each mask now shows almost no contribution due to the bricks.

The principle edges dominate the result.

Note that averaging caused the response of all edges to be weaker.

EMPHASIS ON DIAGONAL EDGE DETECTION

0	1	2
−1	0	1
−2	−1	0

If it is important to emphasize the diagonal edges, a diagonal mask can be used.

−2	−1	0
−1	0	1
0	1	2

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).

The <u>stronger</u> diagonal response is evident.

The response in H & V directions is <u>weaker</u>.

THE LAPLACIAN

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

3x3 region of the image

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

$$\nabla^2 f = 4z_5 - (z_2 + z_4 + z_6 + z_8)$$

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

$$\nabla^2 f = 8z_5 - (z_1 + z_2 + z_3 + z_4 + z_6 + z_7 + z_8 + z_9)$$

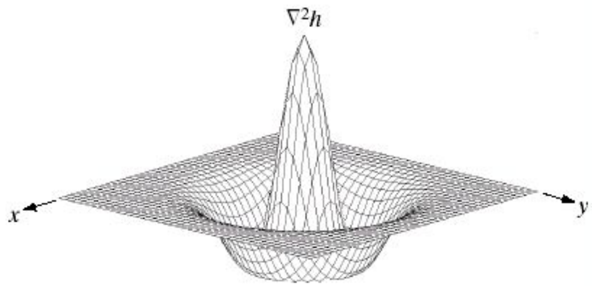
(diagonal neighbors included)

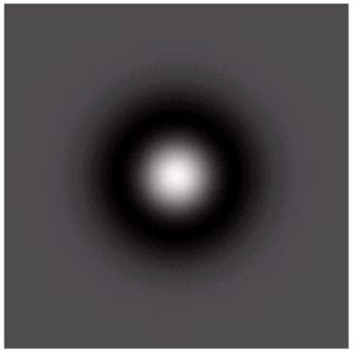
ROLE OF THE LAPLACIAN IN SEGMENTATION

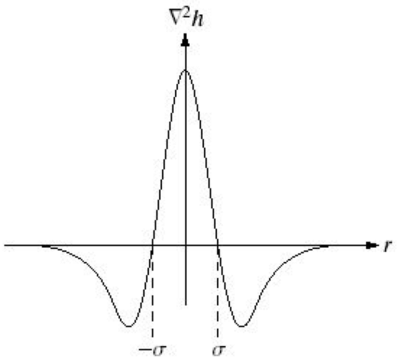
- In general, the Laplacian is <u>not used in its original form</u> for edge detection for several reasons:
  - It is unacceptably sensitive to noise.
  - Its magnitude produces double edges (an undesirable effect).
  - It is unable to detect edge direction.
- Hence, the role of the Laplacian in segmentation consists of:
  - Using its zero-crossing property for edge location
  - Using it to check whether a pixel is on the dark or light side of an edge
- Consider the smoothing function $h(r) = -e^{-\frac{r^2}{2\sigma^2}}, r^2 = x^2 + y^2, \sigma:\text{SD}$
- The Laplacian of *h*: $\nabla^2 h(r) = -\left[\frac{r^2 - \sigma^2}{\sigma^4}\right] e^{-\frac{r^2}{2\sigma^2}}$ (Laplacian of a Gaussian)

THE LAPLACIAN OF A GAUSSIAN (LoG)

<u>The second derivative is a linear operation</u>: convolving an image with $\nabla^2 h$ is the same as convolving the image with *h* first and then computing the Laplacian of the result.







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

A mask that <u>approximates</u> $\nabla^2 h$.

The coefficients sum to zero.

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).

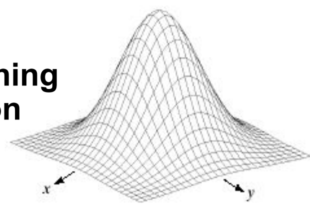
COMPARISON OF 2 APPROACHES FOR EDGE FINDING

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.)

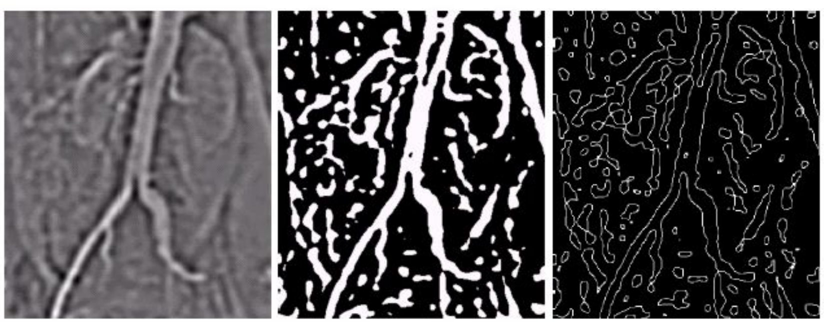
Sobel gradient

Spatial smoothing function



−1	−1	−1
−1	8	−1
−1	−1	−1

Laplacian mask



LoG **Thresholded LoG** **Zero-crossings**

Comparison between (b) and (g):

1. **Edges in the zero-crossing image are <u>thinner</u>.**
2. **The edges determined by <u>zero-crossings form</u> closed loops (a serious drawback of the method).**
3. **The <u>computation</u> of zero-crossings presents a challenge in general. More sophisticated techniques are often required to obtain acceptable results.**
4. **Zero-crossing methods are of interest because of their <u>noise reduction capabilities</u> and potential for rugged performance.**
5. **Because of the noted limitations, the gradient-based edge finding techniques are still used more frequently.**

EDGE LINKING & BOUNDARY DETECTION

- The edge detection methods normally yield pixels lying only on edges.
- In practice, this set of pixels seldom characterizes an edge completely for several reasons:
  - Noise
  - Breaks in the edge from non-uniform illumination
  - Other reasons that introduce spurious intensity discontinuities.
- Edge detection algorithms are followed by linking procedures to assemble edge pixels into meaningful edges.
- We'll look at several basic approaches.
  - Local processing
  - Global processing via the Hough Transform
  - Global processing via graph-theoretic techniques

LOCAL PROCESSING

- A simple approach for linking edge points is to analyze the characteristics of pixels in a small neighborhood (3x3, 5x5) about every point (x,y) labeled an edge point.
- All points that are similar according to a predefined set of criteria are linked.
- This forms an edge of pixels that share the specified criteria.
- 2 principal properties for establishing similarity:
  - The strength of the response of the gradient operator
  - The direction of the gradient operator
- An edge pixel at (x_0,y_0) in the predefined neighborhood of (x,y) is similar in magnitude to the pixel at (x,y) if

$$|\nabla f(x,y) - \nabla f(x_0,y_0)| \leq E,\ E > 0.$$

- An edge pixel at (x_0,y_0) in the predefined neighborhood of (x,y) has an angle similar to the pixel at (x,y) if

$$|\alpha(x,y) - \alpha(x_0,y_0)| \leq A,\ A > 0.$$

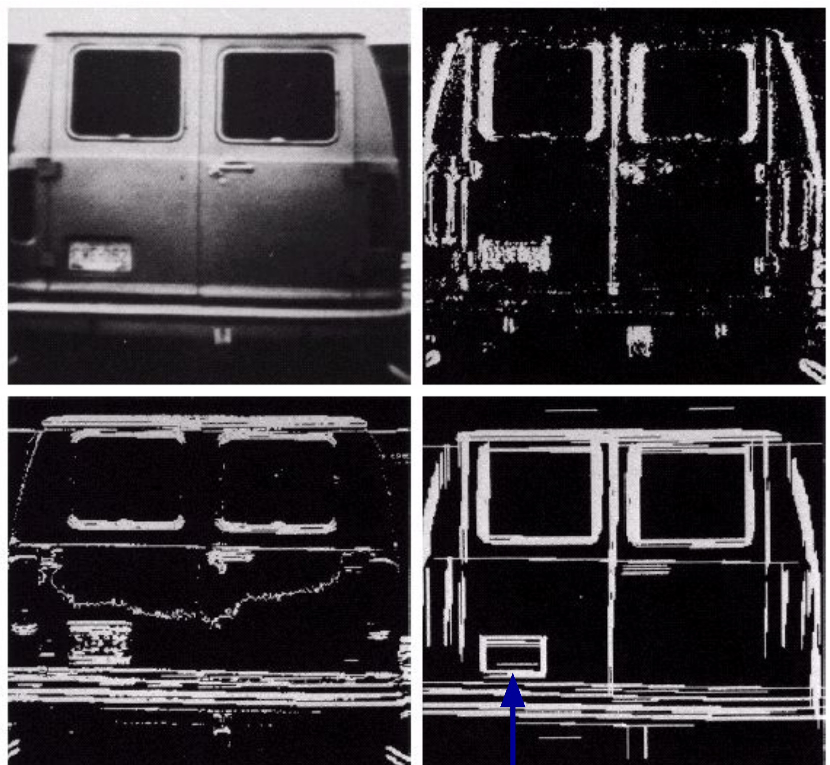
- The point at (x_0,y_0) is linked to the pixel at (x,y) if both criteria are satisfied.

AN EXAMPLE OF LOCAL PROCESSING

The objective is to find rectangles whose sizes make them candidates for <u>license plates</u>.

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.)



G_y

G_x

$E = 25$
$A = 15$

One of the few rectangles detected.
In the US, the width-to-height ratio is 2:1.

THRESHOLDING

Image thresholding has intuitive properties and simple implementation: It enjoys a central position in applications of image segmentation.

We now introduce thresholding in a more formal way.

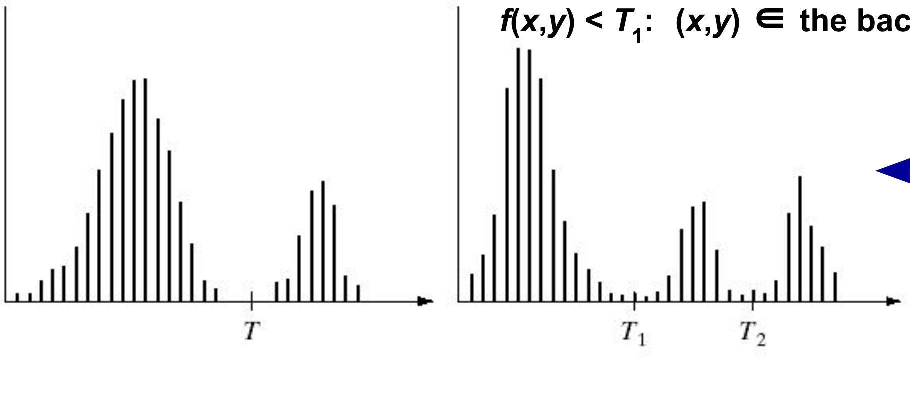
$f(x,y) > T$: an object point
$f(x,y) < T$: a background point

$T_1 < f(x,y) < T_2$: $(x,y) \in$ object class 1

$f(x,y) > T_2$: $(x,y) \in$ object class 2

$f(x,y) < T_1$: $(x,y) \in$ the background

An image composed of light objects on a dark background.



An image composed of 2 types of light objects on a dark background.

FIGURE 10.26 (a) Gray-level histograms that can be partitioned by (a) a single threshold, and (b) multiple thresholds.

In general, segmentation problems requiring multiple thresholds are best solved using region growing methods.

THRESHOLDING FUNCTION

- Thresholding may be viewed as an operation that involves tests against a function T of the form

$$T = T[x, y, p(x,y), f(x,y)]$$

Denotes some local property of (x,y)
e.g., average gray level of a
neighborhood centered at (x,y)

- A thresholded image is defined as

$$g(x,y) = \begin{cases} 1 & \text{if } f(x,y) > T \\ 0 & \text{if } f(x,y) \le T \end{cases}$$

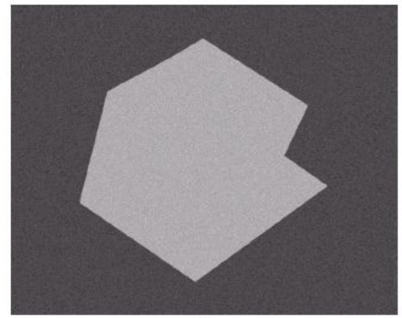
- T depends only on f: the threshold is global.
- T depends on *both* f and p: the threshold is local.
- T depends on f, p, and (x,y): the threshold is dynamic or adaptive.

THE ROLE OF ILLUMINATION

Simple image formation model:

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

$0 < i(x,y) < \infty$
$0 < r(x,y) < 1$



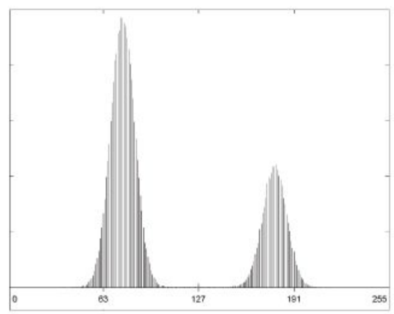
$r(x,y)$

Easy to segment

a
b c
d e

FIGURE 10.27
(a) Computer generated reflectance function.
(b) Histogram of reflectance function.
(c) Computer generated illumination function.
(d) Product of (a) and (c).
(e) Histogram of product image.

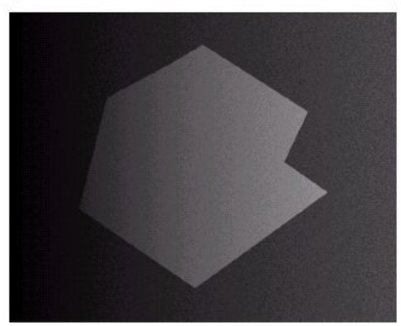
Histogram of $r(x,y)$

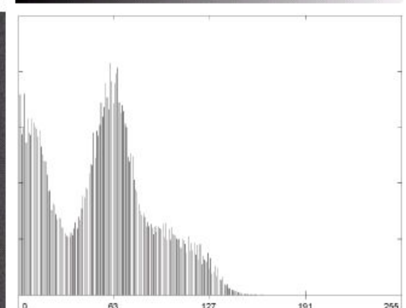


$i(x,y)$

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

Difficult to segment



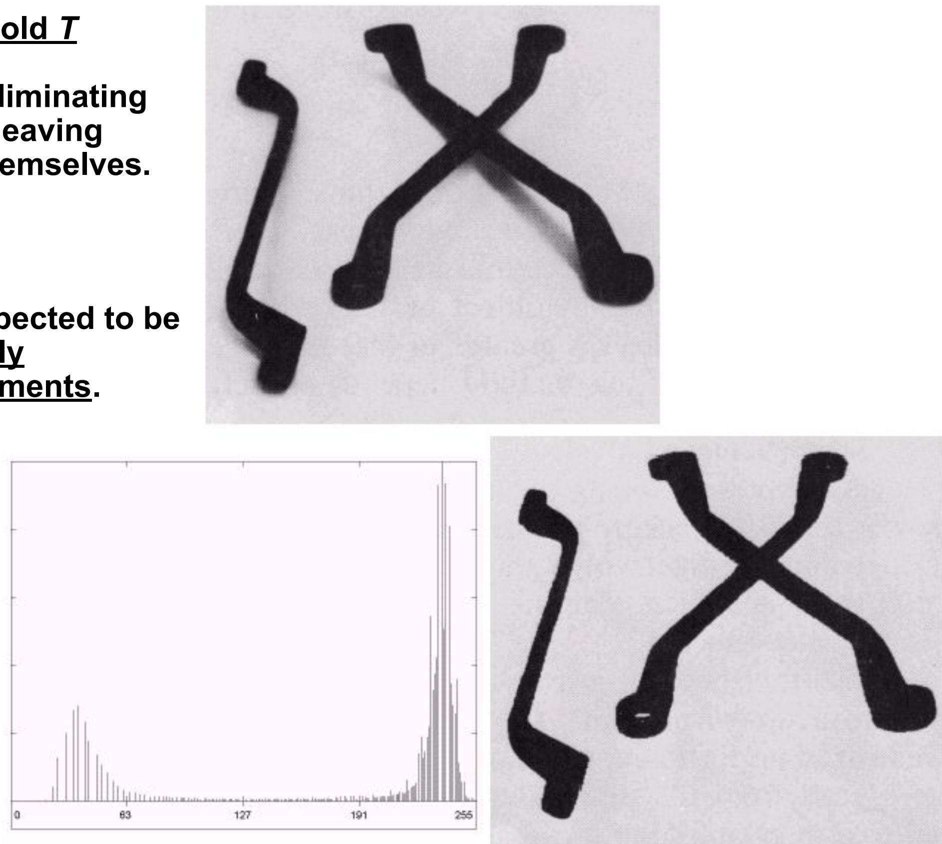


BASIC GLOBAL THRESHOLDING

The <u>chosen threshold *T*</u> achieved a clean segmentation by eliminating the shadows, and leaving only the objects themselves.

This type of global thresholding is expected to be successful in <u>highly controlled environments</u>.

For example, in industrial applications, <u>illumination can be controlled</u>.



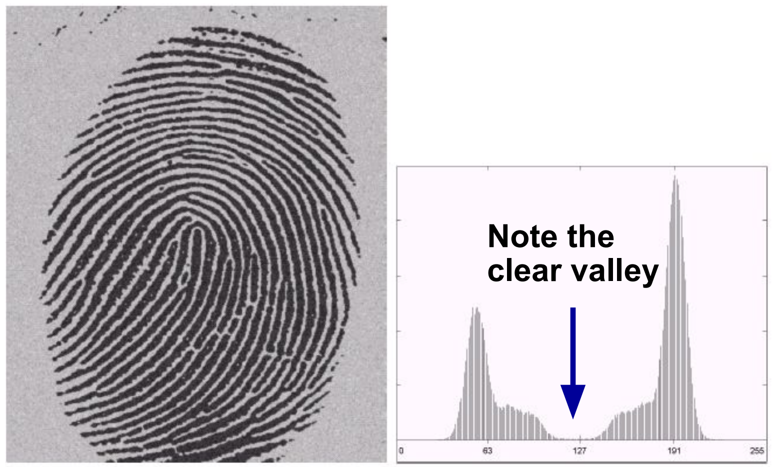
a
b c

FIGURE 10.28 (a) Original image. (b) Image histogram. (c) Result of global thresholding with T midway between the maximum and minimum gray levels.

AUTOMATIC DETERMINATION OF *T*

- In the previous example, the value of T was chosen by a <u>visual inspection</u> of the histogram.
- **Algorithm to determine T automatically**
  1. Select an initial estimate for T.
  2. Segment the image using T.
     - G_1: all pixel values $> T$
     - G_2: all pixel values $\leq T$
  3. Compute μ_1 and μ_2 in regions G_1 and G_2.
  4. Compute a new threshold value: $T = 0.5(\mu_1+\mu_2)$
  5. Repeat steps (2)-(4) until $\Delta T \leq T_0$.
- **Estimation of T_0**
  - <u>Background and object occupy comparable areas</u>: T_0 is the average gray level of the image.
  - <u>Objects are small compared to the area occupied by the background (or vice versa)</u>: T_0 is a value midway between the max and min gray levels.

AN EXAMPLE OF AUTOMATIC DETERMINATION OF *T*



a b
c

FIGURE 10.29 (a) Original image. (b) Image histogram. (c) Result of segmentation with the threshold estimated by iteration. (Original courtesy of the National Institute of Standards and Technology.)

$T = 125$

$T_0 = 0$

After 3 iterations: $T = 125.4$

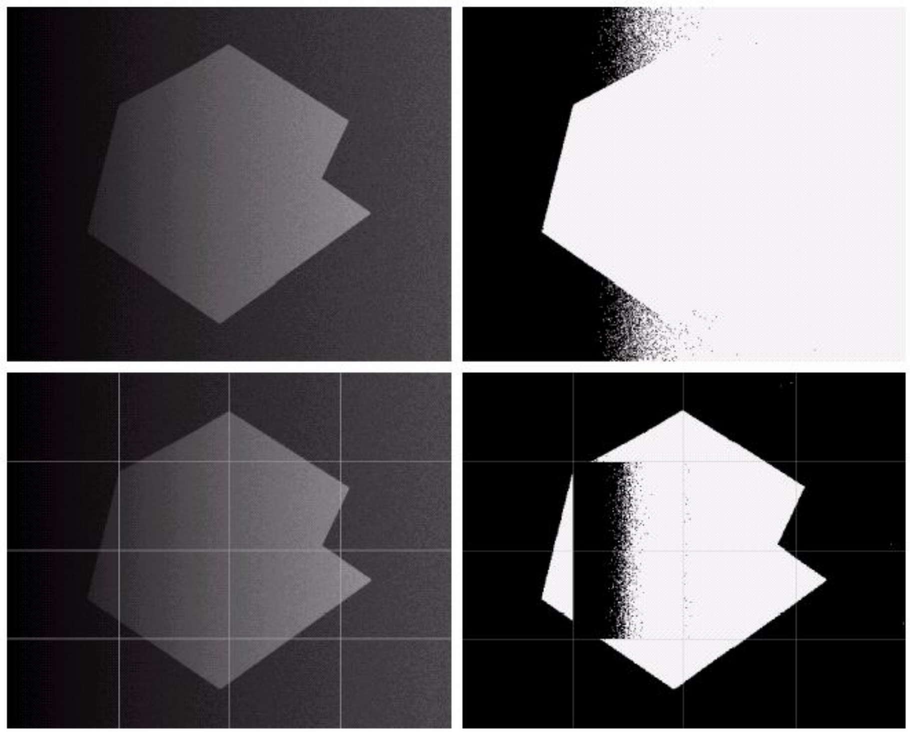
BASIC ADAPTIVE THRESHOLDING

- Imaging factors (uneven illumination, etc.) can transform a perfectly segmentable histogram into a histogram that cannot be partitioned effectively by a single global threshold.
- How do we handle such <u>situations</u>?
- One approach is to divide the image into subimages, and utilize a <u>different threshold</u> for each subimage.
- Key issues
  - How do we <u>subdivide</u> the image?
  - How do we <u>estimate the threshold</u> for each subimage?
- The threshold used for each pixel depends on the <u>location</u> of the pixel.
- Hence, this type of thresholding is adaptive.

AN EXAMPLE OF BASIC ADAPTIVE THRESHOLDING

a b
c d

FIGURE 10.30 (a) Original image. (b) Result of global thresholding. (c) Image subdivided into individual subimages. (d) Result of adaptive thresholding.



The global threshold is manually placed in the valley of the histogram.

Subimages not containing a boundary: $\sigma^2 < 75$

Subimages containing a boundary: $\sigma^2 > 100$

Each subimage with $\sigma^2 > 100$:

T automatically determined.

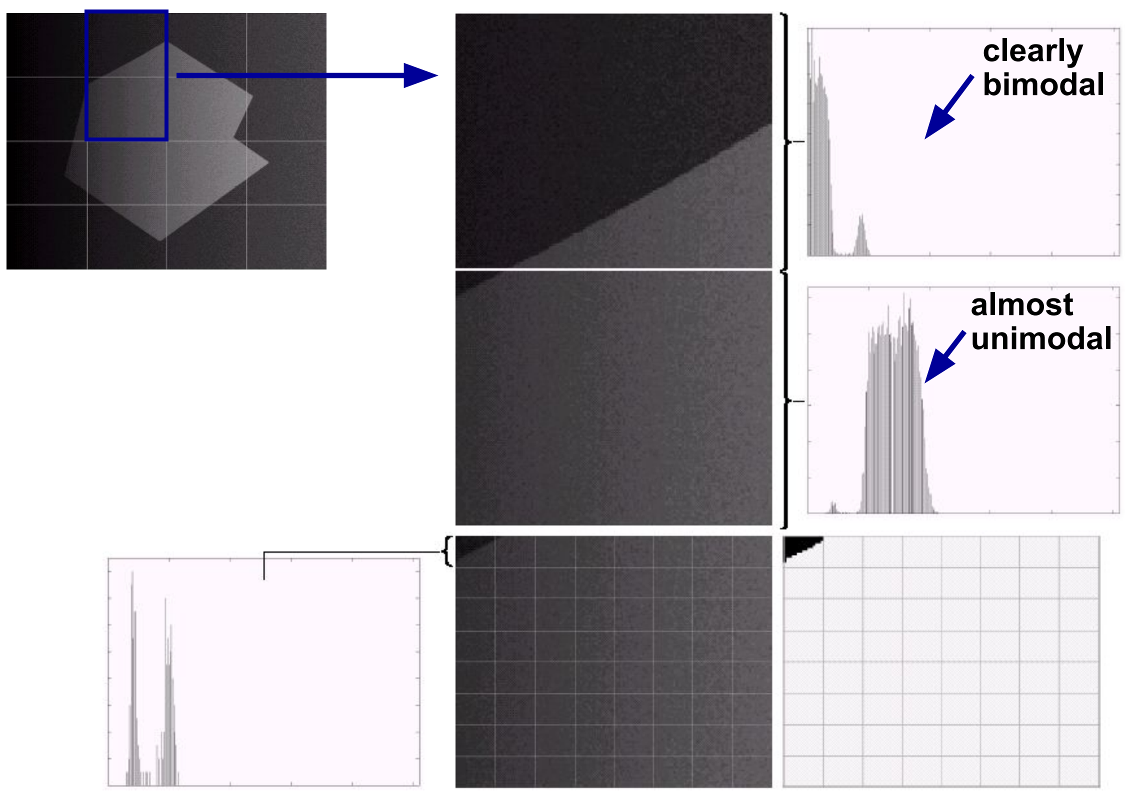
T_0: midway between the min and max gray levels.

16 subimages

All subimages with $\sigma^2 < 100$:

Treated as one composite image, and a single T estimated using the same algorithm.

SUBDIVISION OF A PROBLEM SUBIMAGE



THRESHOLDS BASED ON SEVERAL VARIABLES

- **Multispectral thresholding**: generalization of gray level thresholding
- In color images, each pixel is characterized by 3 RGB values.
- **3D histogram**:
  - An image with 3 RGB components, each having 16 possible levels.
  - 16x16x16 cube
  - Each cell of the cube contains the # of pixels whose RGB components have values corresponding to the coordinates defining the location of the cell.
  - Each entry is divided by the total # of pixels in the image to form a normalized histogram.
- Color segmentation can be based on <u>any color model</u>.
  - HSI model: segmentation using the H and S components is attractive because it involves 2D data clusters.

AN EXAMPLE OF COLOR SEGMENTATION

The color image has 3 16-level RGB components.

Thresholding about one of the histogram clusters corresponding to facial tones.

Thresholding about a cluster close to the red axis.

a b c

FIGURE 10.39 (a) Original color image shown as a monochrome picture. (b) Segmentation of pixels with colors close to facial tones. (c) Segmentation of red components.

T: a distance of <u>one cell</u>
Pixels with components outside the cell: <u>black</u>
Pixels with components inside the cell: <u>white</u>

REGION-BASED SEGMENTATION

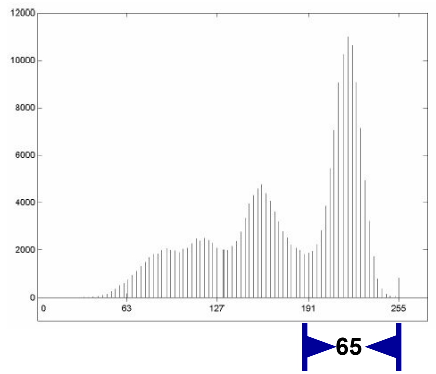
- The objective of segmentation is to partition an image into regions.
- Previous approaches
  - Finding boundaries between regions based on discontinuities in gray levels
  - Thresholding based on the distribution of pixel properties
- We will now discuss techniques that are based on finding the regions directly.
- Basic formulation
  - R: the entire image region; n subregions: $R_1, R_2, \ldots, R_n$
  - $\cup R_i = R$
  - R_i is a connected region, $i = 1,2,\ldots,n$
  - $R_i \cap R_j = \varnothing$ for all i and j, $i \neq j$
  - $P(R_i)$ = TRUE, $i = 1,2,\ldots,n$, where $P(R_i)$ is a logical predicate defined over the points in R_i
  - $P(R_i \cup R_j)$ = FALSE for $i \neq j$

REGION GROWING

- Region growing: a procedure that groups pixels or subregions into larger regions based on predefined criteria.
- Basic approach
  - Start with a set of "seed" points
  - Grow regions by appending to each seed those neighboring pixels with similar properties
- Selection of the seed points
  - Often depends on the nature of the problem
  - No information is available:
    - Compute at every pixel the same set of properties that will be used to assign pixels to regions during the growing process.
    - If the result of the computations shows cluster of values, the pixels whose properties place them near the centroid of these clusters can be used as seeds.
- Selection of similarity criteria depends on (1) problem under consideration, and (2) type of image data available.
- Connectivity or adjacency information is often needed.
- Formulation of a stopping rule

AN EXAMPLE OF REGION GROWING

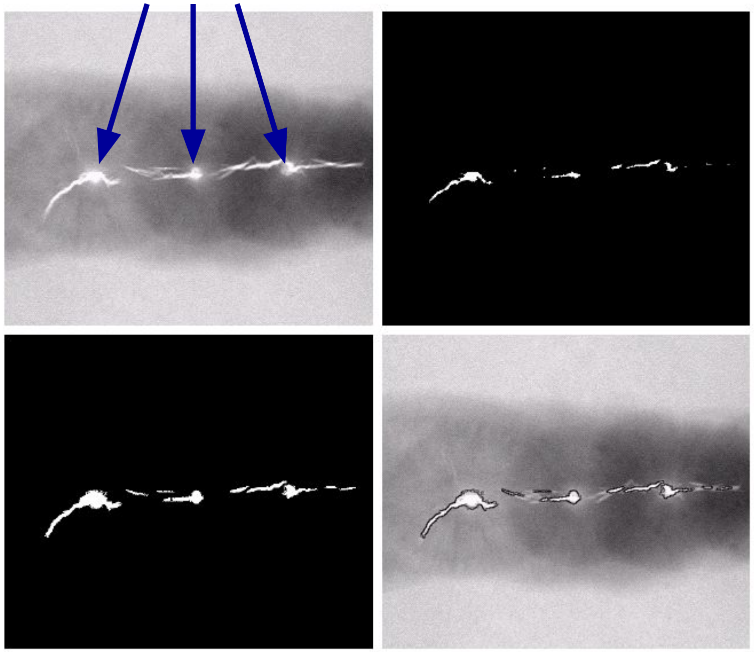
Histogram of (a)



a b
c d

FIGURE 10.40 (a) Image showing defective welds. (b) Seed points. (c) Result of region growing. (d) Boundaries of segmented defective welds (in black). (Original image courtesy of X-TEK Systems, Ltd.).

Several cracks and porosities



Selection of initial seeds:

All pixels with a value of 255

Selection of criteria for region growing:

1. **|any pixel – seed| < 65**
2. **To be included in a region, the pixel has to be 8-connected to at least one pixel in that region.**

No stopping rules were needed: the criteria for region growing were sufficient to isolate the features of interest.

REGION SPLITTING AND MERGING

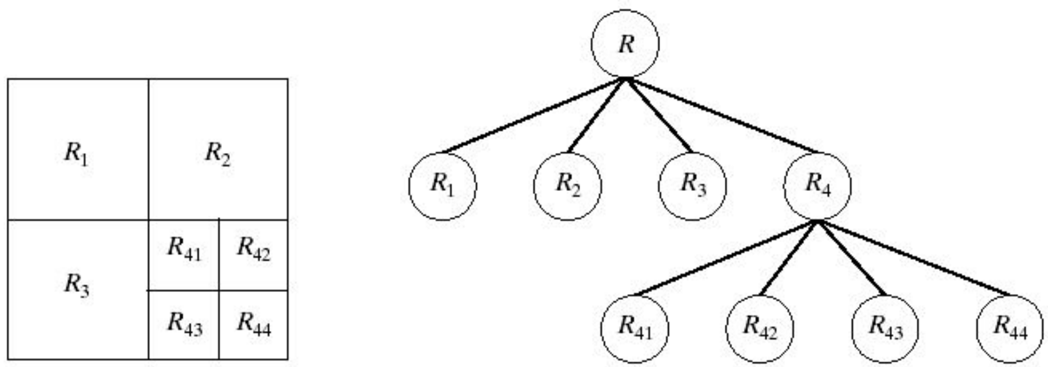
- The previous method grows regions from a set of seed points.
- An alternative is to <u>subdivide an image</u> initially into a set of arbitrary, disjointed regions, and then <u>merge and/or split the regions</u> in an attempt to satisfy the conditions stated earlier.
- Split/merge algorithm
  - R: the entire region, P: the selected predicate
  - If $P(R)$ = FALSE, divide the image into 4 quadrants.
  - At any step
    - Split into 4 disjoint quadrants any region R_i for which $P(R_i)$ = FALSE.
    - Merge any adjacent regions R_j and R_k for which $P(R_j \cup R_k)$ = TRUE.
    - Stop when no further merging or splitting is possible.
- Several variations of the algorithm are possible.

QUADTREE REPRESENTATION

The **root** correspond to the entire image

a b

FIGURE 10.42
(a) Partitioned image.
(b) Corresponding quadtree.



A **quadtree**: a tree in which nodes have exactly 4 descendants

AN EXAMPLE OF REGION SPLITTING AND MERGING

$P(R_i)$ = TRUE if at least 80% of the pixels in R_i have the property $|z_j - m_i| \leq 2\sigma_i$

Gray level of jth pixel in R_i

Mean gray level in R_i

If $P(R_i)$ = TRUE, the values of all pixels in R_i are set to m_i.

The shading and the stem were erroneously eliminated.

a b c

FIGURE 10.43 (a) Original image. (b) Result of split and merge procedure. (c) Result of thresholding (a).

segmentation is perfect.

T is placed midway between the two principal peaks of the histogram.

SEGMENTATION BY MORPHOLGICAL WATERSHEDS

- We have discussed segmentation based on 3 principal concepts:
  - Detection of discontinuities
  - Thresholding
  - Region processing
- Each approach have advantages and disadvantages
  - Speed is an advantage of global thresholding
  - The need for postprocessing is a disadvantage in methods based on detecting discontinuities in gray levels
- Segmentation by watersheds embodies many of the concepts of the other 3 approaches, and often produces more stable segmentation results.
- Topographic and hydrology concepts have proven useful in the development of region segmentation methods.
- A monochrome image is considered to be an altitude surface.
  - High-amplitude pixels correspond to ridge points.
  - Low-amplitude pixels correspond to valley points.
- If a drop of water were to fall on any point of the altitude surface, it would move to a lower altitude until it reached a local minimum.
- The accumulation of water in the vicinity of a local minimum is called a catchment basin.
- All points that drain into a common catchment basin are part of the same watershed.