Image Segmentation

- Region-based, edge-based and boundary-based or threshold-based

Detection of discontinuities
- For 3x3 mask, the response of the mask at any point in the image below:
  - Point detection of isolated points is determined if $|R_i| > T$ then a point has been detected
  - Def. of isolated point: a point whose gray level is significantly different from its background

Line detection
- Move a specific mask around an image
- Preferred direction of each mask is weighted with a larger coefficient than other possible directions
- Use the mask associated with one direction and thresholds its output -- $|R_i| > |R_j|
  - Horizontal, vertical, and diagonal line

Point detection mask
(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)
Example

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)

Edge detection
• the most common approach for detecting meaningful discontinuities in gray level

Basic formulation
- The difference between an edge (local) and boundary (global)
  - Edge: ability to measure gray-level transitions

- an ideal edge: a set of connected pixels
- Models of “step edge”, “ramp edge”, “roof edge”
  - Ramp edge: Edge blurring problem caused by: optics, sampling, and other image acquisition — “ramp-like” profile
  - The thickness of the ramp edge — determined by the length of the ramp

Ideal digital & ramp edge
(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

Example

(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

Random Gaussian noise:
Mean = 0
(a) σ=0
(b) σ=0.1
(c) σ=1.0
(d) σ=10.0

Gradient operators
- First derivative operator based on approximation of 2-D gradient operator
- The gradient of an image \( f(x,y) \) at location \( (x,y) \) is defined as:
  \[
  \left[ \begin{array}{c} G_x \vspace{10pt} \\ G_y \end{array} \right] = \nabla f(x,y) = \left[ \begin{array}{c} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{array} \right]
  \]
- The magnitude of this vector is denoted as \( \sqrt{G_x^2 + G_y^2} \)
- The direction of this is perpendicular to the direction of an edge direction
  \[
  \alpha(x,y) = \tan^{-1}\left( \frac{G_y}{G_x} \right)
  \]
- Robert cross-gradient operator
  - First order derivative
- Prewitt operator
- Sobel operator
  - Use a weight of 2 in the center coefficient to achieve smoothing by giving more importance to the center point
  - Have superior noise suppression to Prewitt operator
  - Harder to implement than Prewitt operator $\nabla f$

- Approximate the gradient operator by absolute values
  - The approximation will not be isotropic (invariant to rotation)
- The Laplacian operator $\nabla^2 f$
  - Is a second-order derivative defined as
  - For a 3×3 region $\nabla^2 f = f_{ij} - f_{x} - f_{y} + f_{xx} + f_{yy}$
  - A digital approximation including the diagonal neighbor is given by $\nabla^2 f = f_{ij} - f_{x} - f_{y} + f_{xx} + f_{yy}$

Edge detector masks

Prewitt and Sobel masks for diagonal edges

Gradient operators

With smoothing (5×5) averaging

Diagonal edge detection
- Is not used in its original form for edge detection
- Is unacceptably sensitive to noise
- Produces double edges, complicate segmentation
- Is unable to detect edge direction

Consists of:
- Use its zero crossing for edge location
- Use it for the complementary purpose of establishing whether a pixel is on the dark or light region of an edge

The Laplacian operator is combined with smoothing as a precursor to finding edge via zero crossing

Consider the function (Laplacian Gaussian)
- This approximation is not unique
- Capture the essential shape of \( \nabla^2 (\sigma^2 r^2 e^{-\sigma^2 r^2}) \)
- Smooth the image, and provide an image with zero crossing
- Sometime it is called the Mexican hat functions

\[ \nabla^2 (\sigma^2 r^2 e^{-\sigma^2 r^2}) \]

Zero crossing image are thinner than the gradient edges

**Drawbacks:**
- Form numerous closed loops (Spaghetti’s effect in Fig. 10.15.g)
- The computation

**Advantages:**
- Noise reduction and strong/robust performance
- Edge-finding based on gradient still are used

**Example**

(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

**EXAMPLE**

(a) Input image
(b) \( G_x \) component of the gradient
(c) \( G_y \) component of the gradient
(d) Result of edge linking

Edge linking and boundary detection
- Characterize an edge completely because of noise, break in the edge from non-uniform illumination, and other effects that introduce spurious intensity discontinuities
- Edge detection operators are followed by linking procedures to assemble pixels into meaningful edges or contours

**Local processing**
- Analyze the characteristics of pixels in a small neighborhood about every point \((x,y)\) in a small neighborhood
- All points that are similar according to a set of criteria: (1) the strength of the response of the gradient operator; (2) the direction of the gradient operator
- Similar in magnitude to the pixel at \((x,y)\) if \( \nabla^2 (\sigma^2 r^2 e^{-\sigma^2 r^2}) \)
- The predefined neighbor of \((x,y)\) is similar in edge direction to the pixel at \((x,y)\) if \( \nabla^2 (\sigma^2 r^2 e^{-\sigma^2 r^2}) \)
- A point in the neighborhood of \((x,y)\) is linked to the pixel \((x,y)\) if both magnitude and direction are satisfied
Global processing via the Hough Transform

- Points are linked by determining first if they lie on a curve of specified shape
- Find subsets if points lie on straight line
  - find all lines determined by every pair of points and find all subsets of points that are close to particular lines
- Accumulator cell
  - the number of subdivision determine the accuracy of co-linearity
  - using the line $y = ax + b$ to represent a line is that the slope approaches infinity as line approaches the vertical
  - Complexity equals to $nK$ (K. increments, n pixels)
- Problem caused by coordinates
  - a line that the slope approaches infinity--vertical line

Parameter space

(a) $xy$-plane
(b) Parameter space

Subdivision for Hough Transform

- Subdivision of the parameter plane for use in the Hough transform

- Solution: use normal representation of a line
  - $x \cos \theta + y \sin \theta = \rho$
  - the loci are sinusoidal curves that intersect at $(\rho_i, \theta_j)$
  - the range of angle $\theta$ is $\pm 90^\circ$
  - the horizontal line has $\theta = 0^\circ$

- Hough transform is applicable to $g(y;c)=0$ For example, the points lying on the circle
  - The complexity is proportional to the number of coordinates and coefficients in a given functional representation

- Edge linking problem based on the Hough transform
  - compute the gradient and threshold it to obtain a binary image
  - specify the subdivision (cells)
  - examine the counts of the accumulator cells
  - examine the relationship (continuity) between pixels in a chosen cell
Example

Global processing via Graphic-theoretic technique
- Global approach for edge detection and linking based on representing edge segments in the form of a graph and search the graph for low cost path
- Provides a rugged approach that performs well in the presence of noise
- Directed graph \( G = (N(U, V)) \)
  - a successor of the parent node
  - the cost of the entire path is
  - edge element: the boundary between pixels \( p \) and \( q \)
  - cost of edge element by pixel \( p \) and \( q \) defined as \( c(p, q) = H - (f(p) - f(q)) \) (\( H \) is the highest gray-level)
  - Sacrifice optimality for the sake of speed

Heuristic algorithm
- Does not guarantee a minimum-cost path; its advantage is speed
- Yield an optimal path only if \( h(n) \) is a lower bound on the cost of the minimum-cost path; the path is constrained to node \( n \) and \( r(n) = H(n) + g(n) \)
- If no heuristic information is available - uniform-cost algorithm of Dijkstra

Heuristic algorithm
- Does not guarantee a minimum-cost path; its advantage is speed
- Yield an optimal path only if \( h(n) \) is a lower bound on the cost of the minimum-cost path; the path is constrained to node \( n \) and \( r(n) = H(n) + g(n) \)
- If no heuristic information is available - uniform-cost algorithm of Dijkstra
Thresholding

**Foundation**
- Thresholding of Bimodal type histogram: single threshold
  - Single thresholding and Multilevel thresholding
    - Thresholding may be viewed as an operation that involves tests against a function $T$ of the form $T=f(x,y)$.
    - Global threshold depends on $f(x,y)$
    - Local threshold $g(x,y)$ and $f(x,y)$
    - Dynamic (or adaptive) threshold depends on the spatial coordinates $x$ and $y$

**The role of illumination**
- The reflection function $r(x,y)$
  - The product of illumination and reflection cause the valley of histogram eliminated, and make segmentation by a single impossible
  - Poor illumination problem: seldom work with reflection function
  - The image resulting from poor illumination could be quite difficult to segment

1. Object and background occupy comparable areas, then $T$ is the average gray value of the image
2. Object are small comparable to background, then the average value is not a good estimate, $T$ is a midway between the maximum and minimum gray levels.

- the distorted histogram $f(x,y)=i(x,y)\cdot r(x,y)$
- then $z(x,y)=\ln f(x,y) = \ln i(x,y) + \ln r(x,y)$
  - the histogram of $z(x,y)$ is given by the convolution of $i'(x,y)$ and $r'(x,y)$
  - if $i'(x,y)$ were constant, $i'(x,y)$ would be constant also, its histogram would be a simple spike (impulse)
  - if $i'(x,y)$ had a broader histogram (resulting from non-uniform illumination), the convolution process would smear the histogram of $r'(x,y)$

**Basic global thresholding**
- The simplest of thresholding—partition the image histogram by using a global threshold
- Segmentation is accomplished by
  - Scanning the image pixel
  - Labeling each pixel depending on whether the gray level of that pixel is greater or less than $T$
threshold by using a heuristic approach, based on the visual inspection of the histogram:
- Select an initial estimate for $T$
- Segment the image using $T(G_1, G_2)$
- Compute the average gray level values
- Compute a new threshold: $T = 1/2(\mu_1 + \mu_2)$
- Repeat steps 2 and 4 until the difference in $T$ is smaller than a predefined $T_0$ (convergence)
- In general, a good initial value for $T$ is the average gray level of the image
- When objects are small compared to the area occupied by the background, average value is not a good initial value.

Basic adaptive thresholding:
- uneven illumination causes a histogram that cannot be partitioned by a global threshold
  - Solution: (1) divide the original image into sub-images; (2) utilize a different threshold to segment each sub-image
  - How to divide an image? and
  - How to estimate the threshold?
  - Threshold depends on the location of the pixel in terms of sub-images

Example
- do not contain a boundary between object and background: variances < 75
- 75 < variance < 100 — composite image
- contains a boundary: variances > 100
  - using the threshold discussed in the previous section
  - the initial $T$ was selected at the point midway between the minimum and the maximum
Optimal global and adaptive thresholding

- A method for estimating thresholds that produce the minimum average segmentation error
  - Histogram may be considered as probability density function (PDF), $p(z)$
  - The overall density function is the sum or mixture of two densities
  - If the form of the densities is known, it is possible to determine an optimal threshold
  - The mixture probability density function is $p(z)=P_1p_1(z)+P_2p_2(z)$
  - An image is segmented by classifying as background all pixels with gray level is greater than a threshold $T$
  - Objective: select the value of $T$ that minimizes the average error

- Using measures based on gradient and Laplacian to deep the valley between histogram peak
  - The optimal threshold may be accomplished for other densities if the Rayleigh and log-normal densities

- Mean square error may be used to estimate a composite gray-level PDF of an image from the image histogram
  - For Example:
    - The mean square error $E$ between $p(z)$ and the image histogram $h(z) = \frac{1}{n} \sum_{i=1}^{n} p(z_i) - h(z_i)$
    - Determine analytically that minimize this mean square error is not a simple matter

GL PDF (2 regions)

- For black object on a light background
- Estimating these densities is not feasible
  - Sol: employ densities whose parameters are reasonably simple to obtain
  - One of the principal densities is Gaussian density
  - Two threshold values may be required to obtain the optimal solution
  - If the variances are equal, a single threshold is sufficient (10-3.14)
  - If $P_1=P_2$, the optimal threshold is the average of the means

Use of boundary characteristics for histogram improvement and local thresholding

- "Good" threshold are enhanced considerably if the histogram peaks are tall, narrow, and separated by deep valleys
- Improve the shape of histogram by considering only those pixels that lie on or near the edges
  - If histogram of the pixel on or near the edge between objects and background were used: improve the symmetry of the histogram peak
  - The histograms have peaks of approximately the same height
- Equal probability: improve the symmetry of the histogram peak
  $f(x,y) = \begin{cases} 0 & \text{if } \nabla f < T \\ + & \text{if } \nabla f \geq T \text{ and } \nabla f \geq 0 \\ - & \text{if } \nabla f \geq T \text{ and } \nabla f \leq 0 \end{cases}$
Threshold based on several variables

- Multi-spectral thresholding
  - 3-D histogram
  - Find clusters of points in 3-D space
  - Shortcoming
    - Cluster seeking becomes an increasingly complex

Region-based segmentation

Basic formulation
- segmentation must be complete; every pixel must be in a region
- points in a region must be connected
- the region must be disjointed
- the pixels in a segmented region \( P(R_i) = \text{TRUE} \) if all pixel in \( R_i \) have the same gray level
- \( P(R_i \cap R_j) = \text{FALSE} \) -- \( R_i \) an \( R_j \) is different

Region growing
- a procedure that groups pixel or sub-region into larger regions based on predefined start with a set of “seed” criteria-similar to the seed (specific gray-level or color)
  - points and from these grow regions by appending each seed
  - priori information is not available—compute the same set of properties
  - Criteria—gray-level, texture, color
  - the selection of similarity
  - For Ex: land-use satellite imagery depends on color

Example

- when the images are monochrome, carried out with set of descriptor based on gray levels and spatial properties (such as moments or textures)
  - The stopping rule
    - when on more pixels satisfy the criteria
    - addional criteria
      - utilize size, likeness between a candidate pixel and the pixels grown so far, and the shape of the region being grown (history of the growth)
  - seed points and seed regions (Fig 10.40)
  - If connectivity or adjacency information is not used, descriptor alone can yield misleading result

Region splitting and merging
- splitting
  - subdivide the entire region successively into smaller and smaller quadrant regions if \( P(R_i) = \text{FALSE} \)
  - quad tree — the root of the tree corresponds to the entire image and each node corresponds to a subdivision
only splitting causes the final partition would contain adjacent regions with identical properties can be solved by merging

Merging
- Is limited to groups of four blocks
- the splitting and merging steps
  - split into four disjoint quadrants
  - merge adjacent adjacent regions
  - stop when no further merging or splitting

Variations of the scheme
- Split the image into a set of blocks
- The advantage: use the same quad-tree for splitting and merging
- Texture segmentation

No further splitting
- The procedure is terminated by one final merging of regions satisfying step2
- The merged regions may be of different size
- Texture segmentation is based on the measures of texture

Segmentation by morphological watersheds
- Advantage of global thresholding speed
- Disadvantages of traditional segmentation: (1) the need to post-processing; (2) need edge linking for discontinuity
- The watershed segmentation is often produces stable segmentation results (continuous segmentation boundaries)

Basic concepts
- Based on visualizing an image in three-dimension: two spatial coordinates versus gray levels
- “Topographic” interpretation (three types of points):
  - Points belonging to a regional minimum
  - Points at which a drop of water (catchment basin or watershed)
  - Points at which water would be equally likely to fall to more than one such minimum (crest line on the topographic surface; divide line or watershed lines)
- The principal objective is to find the watershed lines
- Suppose that a hole is punched in each regional minimum and the entire topography is flooded
  - Water rise through the holes
  - Rising water in distinct catchment basin is about to merge
  - A dam is built to prevent the merging
- Reach when the tops of the dam are visible
**Dam construction**
- Based on binary images
- The simplest way
  - Construct dams separating sets of binary points based on morphological operations
- Use dilation
- Def of catchment basin and the sets of points in two regional minimum at stage $n$: $C_0(M_1)$ and $C_0(M_2)$
  - The two individual components extracted from $q$ by performing the simple AND operation $q \cap C_0(M_2)$ becomes one connected component
  - Each of the connected components is dilated by the structuring element, subject to two conditions:
    - The dilation has to be constrained to $q$
    - The dilation cannot be performed on points that would cause the sets being dilated to merge

**Flooding**
- Original image
- Topographic view
- Two slopes of flooding
- Result of further flooding
- Beginning of merging of water from two catchment basins (a short dam was built between them)
- Longer dams
- Final watershed (segmentation)

**Flooding (cont’d…)**
- Two partially flooded catchment basins at stage $n — 1$ of flooding
- Flooding at stage $n$, showing that water has