Image Segmentation

Tien-Ying Kuo

NTUT

Segmentation

- Segmentation:
  - Input: Image ,Output: Attribute
  - Subdivides an image into constituent regions or objects of interests.
- Segmentation is based on:
  - Discontinuity:
    - Edge/Abrupt changes in intensity
      ➔ Edge detection+ edge connection (assemble edge segments into longer edge)
  - Similarity:
    - Thresholding ➔ Fast Speed, popular
    - Region growing, splitting and merging ➔ morphological/watershed seg
- Segmentation Accuracy:
  - When to stop: No past the level of detail required to identify those objects.
  - Good sensor: Diminishing the contribution of irrelevant image detail.

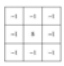
Detection of Discontinuities

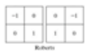
Detection of Discontinuities

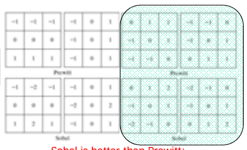
- Detect by masking (sum to 0) $R=\sum_{i=1}^{9} w_i z_i$
  - Point Detection
    - Isolated point: $|R| \geq T$ (T controls the degree)
  - Line Detection

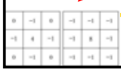
    Case 1: Ex: $|R_1|>|R_j|$, j=2,3,4 ➔ Point is most likely associated with horizontal line

    Case 2: we are interested in detect a line in a specified direction
  - Edge Detection
    - Gradient
    - Laplacian
      - Double Edge: undesired for segmentation
      - Unable to detect edge direction







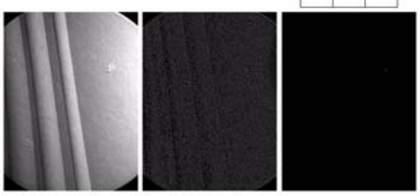


Sobel is better than Prewitt:
Slightly superior in noise-suppression

Point Detection

T: set the interests, 90% of the highest absolute pixel value

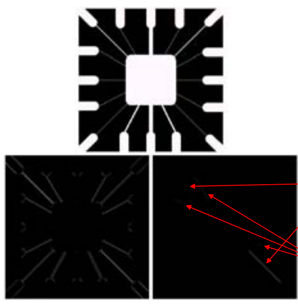
−1	−1	−1
−1	8	−1
−1	−1	−1



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.)

Industry inspection: a porosity in turbine blade

Line Detection



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).

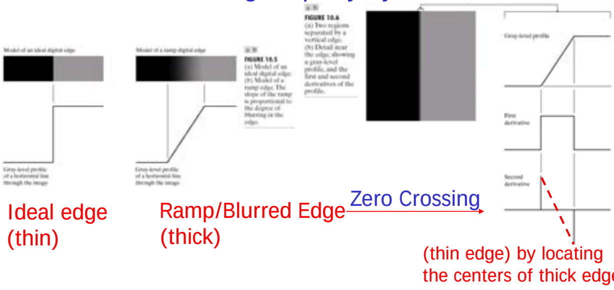
Because TH=max value
Only detect one-pixel thick

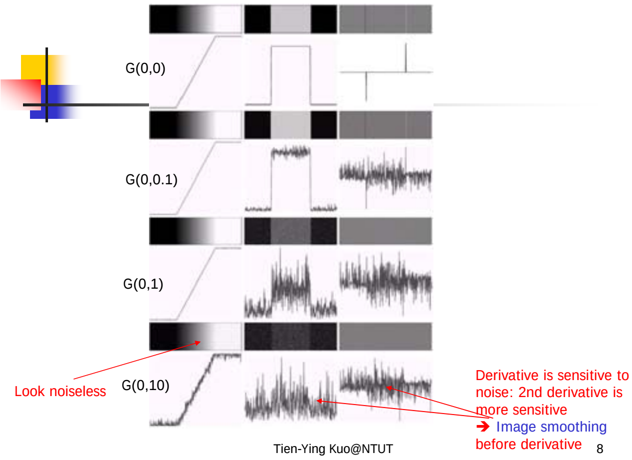
Point detection can remove isolated points

Industry inspection: find -45° Line with one pixel thick

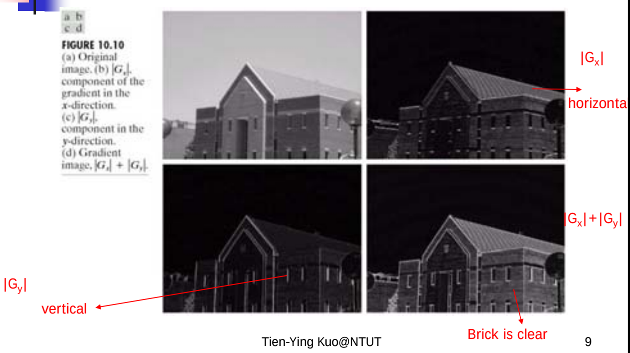
Edge Detection

- Zero Crossing Property by 2nd derivative

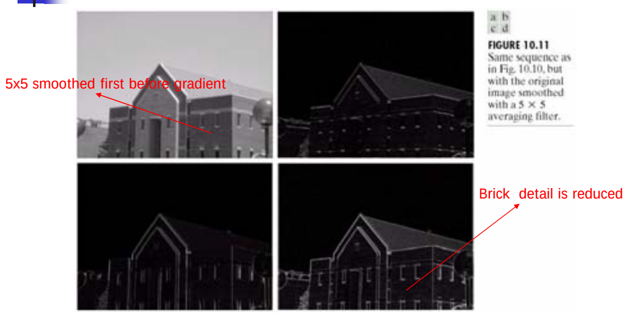




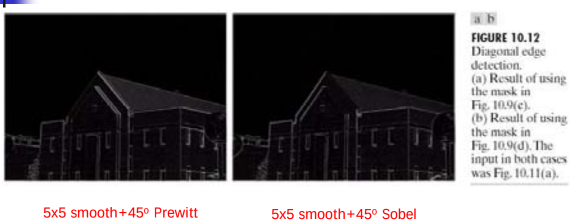
Edge detection by Gradient



Smooth+Gradient



Diagonal edge detection



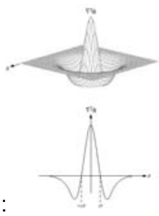
Gaussian LPF + Laplacian

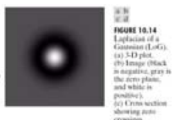
- To reduce the double edge, and the effect of noise
- Gaussian LPF (σ controls the degree of blurring):

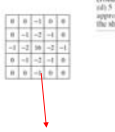
$$h(r) = -e^{-\frac{r^2}{2\sigma^2}} \quad , r^2 = x^2 + y^2$$

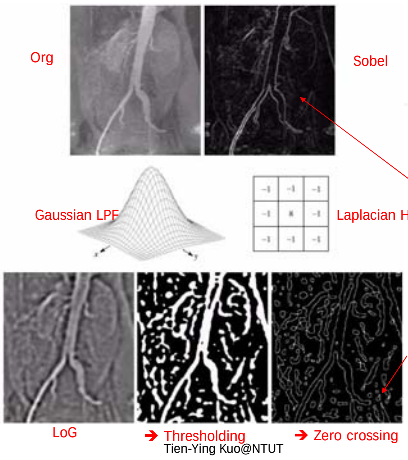
- Laplacian of Gaussian (LoG)(Mexican hat fun):

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









Zero crossing vs Sobel

Advantage:
1. Thinner Edge
2. Noise reduction

Drawback:
1. Many closed loop (Spaghetti effect)
2. Computation

Edge Linking and Boundary Detection

Edge Linking

- Assemble edge pixels into meaningful edges
- Why needed:
  - Non-uniform illumination ➔ break the edge
  - Noise
- Methods:
  - Local Processing
  - Global Processing
    - Hough Transform
    - Graph-Theoretic Techniques

Local Processing

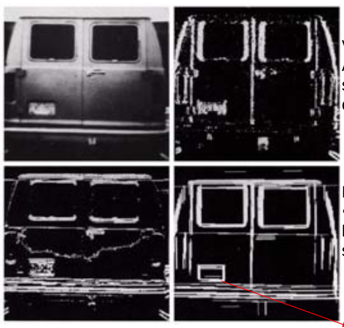
- A point in the predefined neighborhood of (x,y) is linked to the pixel at (x_0,y_0) if
  - Magnitude is similar
    $|\nabla f(x,y) - \nabla f(x_0,y_0)| \le E$, where E>0
  - **and** Angle is similar
    ($\alpha(x,y)$ is **perpendicular** to the direction of the gradient vector at the point)
    $|\alpha(x,y) - \alpha(x_0,y_0)| \le A$, where A>0 $\quad \alpha(x,y) = \tan^{-1}\left(\frac{G_y}{G_x}\right)$
- Bookkeeping processing:
  - A record must be kept of linked points as the center of the neighborhood is moved from pixel to pixel
  - Assiagn **a different gray level** to each set of linked edge pixels

Local Processing

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.)

E=25, A=15º



Vertical (G_y): Applying the similarity to every column

Horizontal (G_x): Applying the similarity to every row

Horizontal+Vertical ➔removing small breaks and isolated short segments

Use 2:1 ratio to locate the license plate

Global Processing – Hough Transform

- Global Processing:
  - Consider global relationships between pixels
  - Points are linked by determining first if they lie on a curve of specified shape
- Problem: Given n points, find subsets of them that lie on straight lines
- Straightforward Approach
  - Step 1: Find all lines determined by every pair of points ($n(n-1)/2 \sim n^2$)
  - Step 2: Find all subsets of points that close to particular lines ($(n)(n(n-1))/2 \sim n^3$)
  - This approach is computationally prohibitive

Global Processing – Hough Transform

- Hough Transform [1962]
  - Convert xy-plane image space (x,y) to ab-plane parameter space (a,b).

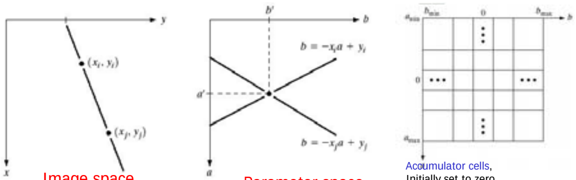


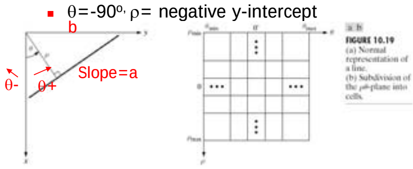
Image space
$y_i = ax_i + b$

Parameter space
$b = -x_i a + y_i$

Accumulator cells,
Initially set to zero,
Increased by 1
Complexity=nK, if K increments in cells
Usually smaller than n^2

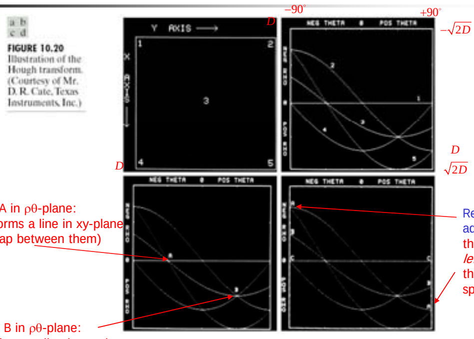
Global Processing – Hough Transform

- Difficulty in ab-plane: Vertical line → infinity slope (a)
- Change (a,b) ab-plane to (ρ, θ) ρθ-plane.
  (x ·cosθ + y ·sinθ = ρ)
  - θ∈[-90°,90°]
  - θ=0°, ρ= positive x-intercept
  - θ=90°, ρ= positive y-intercept
  - θ=-90°, ρ= negative y-intercept



Global Processing – Hough Transform

FIGURE 10.20 Illustration of the Hough transform. (Courtesy of Mr. D. R. Cate, Texas Instruments, Inc.)



Intersection A in ρθ-plane:
Point 1,3,5 forms a line in xy-plane
(Possibly a gap between them)

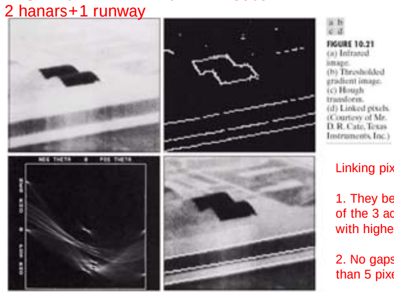
Reflective adjacency at the *right* and *left* edges of the parameter space

Intersection B in ρθ-plane:
Point 2,3,4 forms a line in xy-plane
(a gap between them)

Global Processing – Hough Transform

Original (aerial infrared)
2 hanars+1 runway

Sobel



Hough of Sobel

Linked pixels

Linking pixels by

1. They belonged to one of the 3 accumulators with highest count
2. No gaps were longer than 5 pixels

Global Processing – Hough Transform

- Summary:
  - Step 1: Compute the gradient of an image and threshold it to obtain a binary image
  - Step 2: Specify subdivisions in the ρθ-plane
  - Step 3: Examine the counts of the accumulators cells for high pixel concentrations
  - Step 4: Examine the relationship (principally for continuity, i.e. no gap) between pixels (in a chosen cell by finding the pixel distance in a given accumulator cells during traverse)

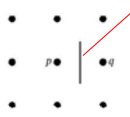
Global Processing – Graphic-Theoretic Technique

- Definition:
  - Graph G=(N,U)
    - N: nodes,
    - U: unordered pairs $\{(n_i,n_j)\}$ of distinct elements of N
    - (n_i,n_j): arc
  - Directed Graph: a graph in which the arcs (n_i,n_j) are directed
    - n_i: parent node, n_j: successor, n_i to n_j: expansion
    - Level 0: start/root node, Last level : goal node
    - Cost $c(n_i,n_j)$ is assocaited with every arc(n_i,n_j)
    - A path $\{n_1,n_2,\ldots,n_k\}$ where n_i is a sucessor of n_{i-1}
    - Cost of path: $c = \sum_{i=2}^{k} c(n_{i-1}, n_i)$

Global Processing – Graphic-Theoretic Technique

- Edge element $(x_p,y_p)(x_q,y_q)$: the boundary between two pixels p and q, such that **p and q are 4-neighbors**.
- Edge: A sequence of connceted edge elements

Edge element



Assume the edge start in the top row and terminate in the last row

Gray Level

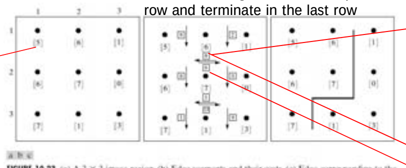


FIGURE 10.23 (a) A 3 × 3 image region. (b) Edge segments and their costs. (c) Edge corresponding to the lowest-cost path in the graph shown in Fig. 10.24.

Cost c(p,q)=H-[f(p)-f(q)]
- H is the highest gray-level value (7)
- p: right-hand side of travel direction
- q: left-hand side of travel direction

c(p,q)=8=7-(6-7)

c(p,q)=6=7-(7-6)

Global Processing – Graphic-Theoretic Technique

Edge corresponding to the lowest cost path

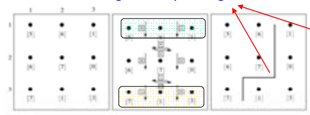
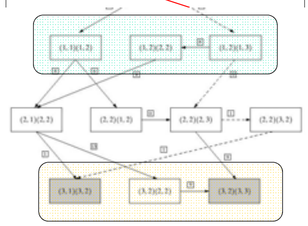


FIGURE 10.23 (a) A 3 × 3 image region. (b) Edge segments and their costs. (c) Edge corresponding to the lowest-cost path in the graph shown in Fig. 10.24.

Assume the edge start in the top row and terminate in the last row

The min cost path is shown dashed



- Finding a min-cost path is not trivial in terms of computation
- Some approach is to sacrifice optimality for the sake of speed (see textbook)
- Graphic Theoretic Approach is a rugged approach in the presence of noise, but is more complicated than the methods so far.

Global Processing – Graphic-Theoretic Technique

- Sub-optimum: use the opimum path for five levels of heuristic down from the point

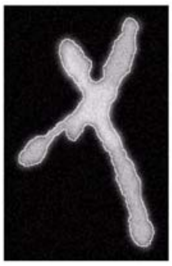


FIGURE 10.25 Image of noisy chromosome silhouette and edge boundary (in white) determined by graph search.

Noisy chromosome

Thresholding

Thresholding

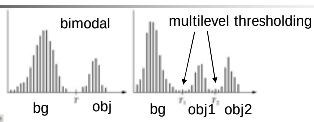


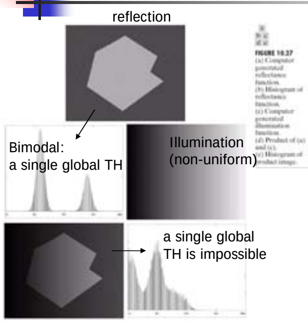
FIGURE 10.26 (a) Gray-level histograms that can be partitioned by (a) a single threshold, and (b) multiple thresholds.

- A thresholded image g(x,y) is defined as

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

- In general: $T=T[x,y,p(x,y),f(x,y)]$
  - Where f(x,y): gray level at point (x,y)
  - P(x,y) some local property of this point (ex: ag gray of neighborhood of (x,y))
- Global thrshold: T depends only on f(x,y) (gray-level)
- Local threshould: T depends on both f(x,y) and p(x,y)
- Dynamic/adaptive threshold: T depends on (x,y)

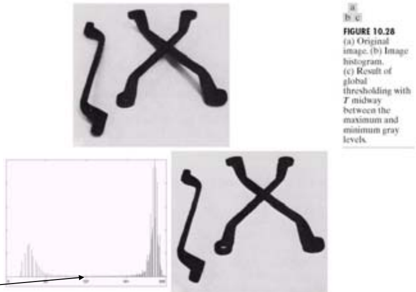
The role of Illumination



f(x,y)=i(x,y)·r(x,y)

- To thresholding an image f(x,y)=i(x,y)·r(x,y)
  - When access to the illumination source is available
    - Project the illumination pattern onto a **constat, white** reflective surface → g(x,y)=k·i(x,y)
  - A normalized function of f(x,y) h(x,y)=f(x,y)/g(x,y)=r(x,y)/k
  - If r(x,y) can be segmented by a single TH, then h(x,y) can be segmented by a single threshold TH/k

Global Thresholding (Heuristic)

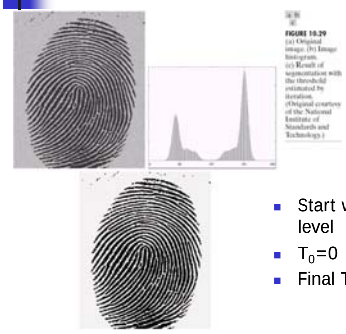


TH= midway between the max and min gray levels
(heuristic approach – based on visual inspection of the histogram)

Global Thresholding (Auto)

- **Step 1**: Select an initial estimate for T
- **Step 2**: Segment the image using T.
  Produce tow groups of pixels: G_1 (>T) and G_2 (≤T)
- **Step 3**: Compute average gray level u_1 and u_2 for the pixels in regions G_1 and G_2
- **Step 4**: Compute a new threshold: $T= (u_1+u_2)/2$
- **Step 5**: Repeat **Steps 2 through 4** until the difference in T in successive iterations is smaller than a predefined parameter T_0

Global Thresholding (Auto)



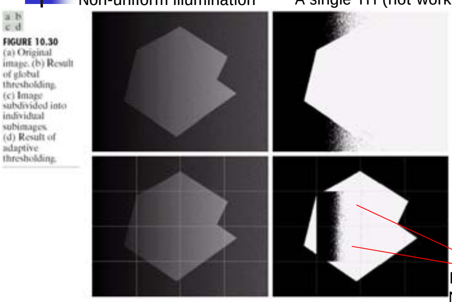
- Start with average gray level
- T_0=0
- Final TH T=125.4→125

Adaptive Thresholding

- Problem: Uneven illumination
  → cannot be partitioned by a single global threshold
- Solution: divide the original image into sub-images and then utilize a different threshold to segment each sub-image
- Key issue:
  - how to subdivide the image
  - How to estimate the threshold for each sub-image

Adaptive Thresholding

Non-uniform illumination A single TH (not work)

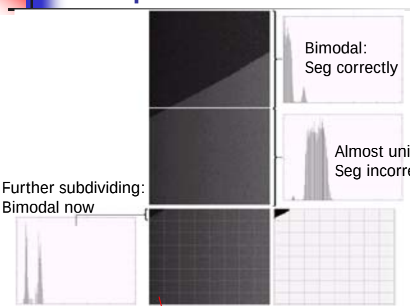


For each sub-image
- Var < 100:
  treated as one composite image
- Var ≥ 100:
  use midway between the min and max in sub-image as TH

Improper,
Need further subdividing

sub-image (1→4 twice) Sub-image

Adaptive Thresholding



Bimodal:
Seg correctly

Almost uni-modal:
Seg incorrectly

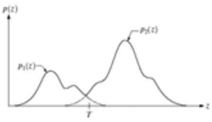
Further subdividing:
Bimodal now

Still uni-modal, and avg gray level gets closer to the object

Optimal Global and Adaptive Thresholding

- Optimal: Estimate a threshold to produce the min average seg error

FIGURE 10.32 Gray-level probability density functions of two regions in an image.



- Mixture probability: $p(z) = P_1 p_1(z) + P_2 p_2(z)$ where $P_1 + P_2 = 1$
- Minimize $E(T) = P_2 E_1(T) + P_1 E_2(T)$
  - Where $E_1(T) = \int_{-\infty}^{T} p_2(z)dz$ error pdf: should be bg, but classified as obj

    $E_2(T) = \int_{-\infty}^{T} p_2(z)dz$ error pdf: should be obj, but classified as bg
- By dE(T)/dT, solution: $\boxed{P_1 p_1(T) = P_2 p_2(T)}$
  - If P_1=P_2, the optimum TH is the intersect of p1(z) and p2(z) curves

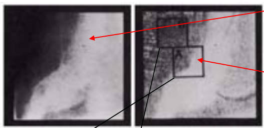
Optimal Global and Adaptive Thresholding

- Usually we model p(z) pdf as Gaussian to get analytical expression for T
  - Gaussian Model: $p(z) = \frac{P_1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(z-\mu_1)^2}{2\sigma_1^2}} + \frac{P_2}{\sqrt{2\pi}\sigma_2} e^{-\frac{(z-\mu_2)^2}{2\sigma_2^2}}$
  - Analytic solution: $AT^2 + BT + C = 0$
    where
    $A = \sigma_1^2 - \sigma_2^2$
    $B = 2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)$
    $C = \sigma_1^2\mu_2^2 - \sigma_2^2\mu_1^2 + 2\sigma_1^2\sigma_2^2 \ln(\sigma_2 P_1 / \sigma_1 P_2)$
  - If $\sigma^2 = \sigma_1^2 = \sigma_2^2$ $\boxed{T = \frac{\mu_1 + \mu_2}{2} + \frac{\sigma^2}{\mu_1 - \mu_2}\ln\left(\frac{P_2}{P_1}\right)}$
    - If P_1=P_2 or σ=0, $\boxed{T = \frac{\mu_1 + \mu_2}{2}}$ $\boxed{e_{ms} = \frac{1}{n}\sum_{i=1}^{n}(p(z_i) - h(z_i))^2}$
  - For real n-pt histogram, find a better fit of Guassian p(z) by
- Similar to Raleigh, log-normal…

Optimal Global and Adaptive Thresholding

FIGURE 10.35 Cardioangiogram showing superimposed ... (Chow and Kaneko.)

FIGURE 10.33 A cardioangiogram before and after preprocessing. (Chow and Kaneko.)



Histogram (solid)

Two Gaussian fit (x, o)

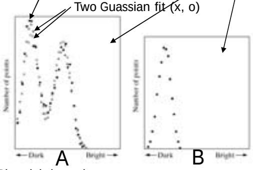


FIGURE 10.34 Histograms (black dots) of (a) region A, and (b) region B in Fig. 10.33(b). (Chow and Kaneko.)

A: Bimodal: boundary presented

B: unimodal: no boundary

- Preprocessing:
  - Log transform: contrast enh
  - Image subtraction: remove spinal column
  - Several images average: reudce random noise
- 256x256 to 7x7 grid with 50% overlap (each grid: 64x64)
- Curve fit (histogram → Gaussian by minimizing e_{ms})
  - Only regions with bimodal are assigned TH
  - TH of remaining regions are interpolated.
  - Second interpolation was carried out point by point by using neighboring TH
- Bindary decision:

$$f(x,y) = \begin{cases} 1 & \text{,if } f(x,y) \ge T_{xy}, \quad \text{(object)} \\ 0 & \text{,if } f(x,y) < T, \quad \text{(background)} \end{cases}$$

Use of Boundary Characteristics for Histogram Improvement and Local Thresholding

- Observation:
  - Good Thresholding if the histogram peaks are tall, narrow, symmetric and separated by deep valleys.
  - Improve the shape of histogram by considering the pixel on or near the edge. i.e., work on gradient or Laplacian iamge
  - Laplacian can yield information regarding whether a given pixel lies on the dark or light side of an image. (give more inform than gradient)

Use of Boundary Characteristics for Histogram Improvement and Local Thresholding

FIGURE 10.36 Image of a handwritten stroke coded by using Eq. (10.3-16). (Courtesy of IBM Corporation.)

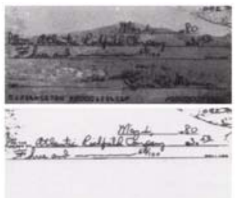
- Three-level image is formed using gradient and Laplacian

$$s(x,y) = \begin{cases} 0 & \text{if } \nabla f < T \quad \text{(not on an edge)} \\ + & \text{if } \nabla f \ge T \text{ and } \nabla^2 f \ge 0 \text{ (dark side)} \\ - & \text{if } \nabla f \ge T \text{ and } \nabla^2 f < 0 \text{ (light side)} \end{cases}$$

- A horizontal/vertical scan line containing a section of an object has the transition: (…)(-,+)(0 or +)(+,-)(…)

Use of Boundary Characteristics for Histogram Improvement and Local Thresholding

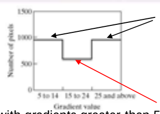
FIGURE 10.37 (a) Original image. (b) Image segmented by local thresholding. (Courtesy of IBM Corporation.)



org

seg

FIGURE 10.38 Histogram of pixels with gradients greater than 5. (Courtesy of IBM Corporation.)



Two dominant modes

TH (midpoint of valley)

Histo of pixels with gradients greater than 5

Thresholds Based on Several Variables

- Some sensor can allow multispectral thresholding (more than one variable to characterize)
  - Ex: Color image: 3D histogram of 16x16x16 RGB cube (each has 16 possible levels)
    - The problem becomes to finding clusters of points in 3-D space (Cluster-seeking methods)
  - Ex: HIS: H and S are important ➔ 2-D data clusters (easier)
- Example:

a b c Color image Seg by facial tones Seg by red components (scarf & flower)

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.

Region-Based Segmentation

- Region-Based: Finding the regions directly by splitting or merging
- Formulation: segmentation is a process to partition R (entire image) into n subregions, $R_1, R_2, \ldots, R_n$, such that
  - Every pixel must be in a region: $\bigcup_{i=1}^{n} R_i = R$
  - R_i is a coonected regions, i=1,2,…,n (a region must be connected)
  - Must be disjoin, $R_i \cap R_j = \phi$ for all i and j, $i \neq j$
  - $P(R_i)$=TRUE for i=1,2,…,n, where $P(R_i)$ is a logical predicate defined over the points in set R_i
  - $P(R_i \cup R_j)$=FALSE for $i \neq j$

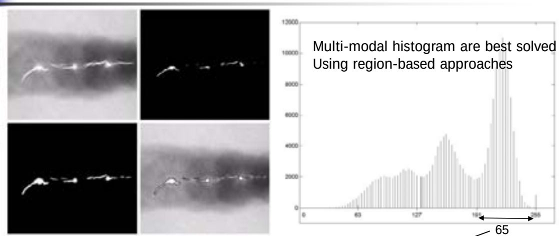
Region-Based Growing

- Region Growing is a procedure that groups pixels or sub-regions into larger regions based on predefined criteria
- Start with a set of "seed" points and grow by appending to each seed if neighboring pixels that have similar properties
- Predefined criteria/similar properties:
  - Depending on the type of image data
    - Land-use stellite: color
    - Monochrome: gray levels and spatial properties(moments, texture)
  - Connectivity or adjacency must be considered. (or misleading)
  - Stopping rule: stop when no more pixels satisfy criteria
    - Not use: gray level, texture, color (local in nature, not into history of growth)
    - Used: Size, likeness between a candidate pixel and the pixels grown so far
      Ex: A comparison of the gray level of a candidate and the average gray level of the grown region

Region-Based Growing

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.)



Multi-modal histogram are best solved Using region-based approaches

65

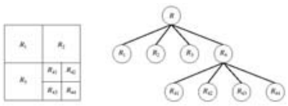
- Initial seed point: all pixels having vlaues of 255, and clustered into seed regions.
- Merge:
  - Abs gray diff between any pixel and the seed < 65
  - 8-connected (The use of connectivity is fundamental in solving this problem)

Region Splitting and Merging

- Quadtree splitting:
  - R: entire image
  - If P(R)=FASE, we divide the image into four disjoin qudrants R_i
  - Merge any adjacent regions R_j and R_k for which $P(R_j \cup R_k)$=TRUE
  - Repeat the qudarnt splitting if $P(R_i)$=FASE (step 2)
  - Step when no further mergin or splitting is possible
- Merging is required, because if only splitting is used, the final partition likely would contain adjacent regions with identical properties

a b

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



Region Splitting and Merging

a b c

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

org | By splitting and merging (segmented perfectly) | By thresholding (midpoint of two peaks) (shading of stem and leave were erroneously eliminated)

- Splitting and Merging:
  - Define $P(R_i)$=TRUE if at least 80% of the pixels in R_i has the property $|z_j - m_i| \leq 2\sigma_i$, ($m_i$: mean, σ_i: std dev in R_i)
  - If $P(R_i)$=TRUE under this condition, the values of all the pixels in R_i were set to m_i
- Texture segmentation is beased on using measures of texture for the predicates $P(R_i)$, which property is based on m_i and σ_i in a region.

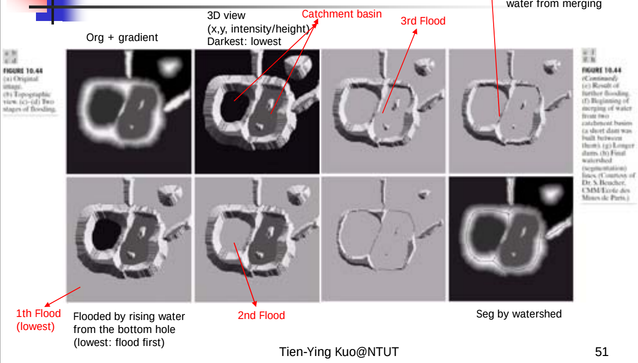
Segmentation by Morphological Watersheds

Watersheds

- Watersheds:
  - Embodies the concepts of
    - Detection of discontinuities
    - Thresholding
    - Region processing
  - Provide a simple framework for incorporating knowledge-based constraints
  - Based on Gradient+Morphological

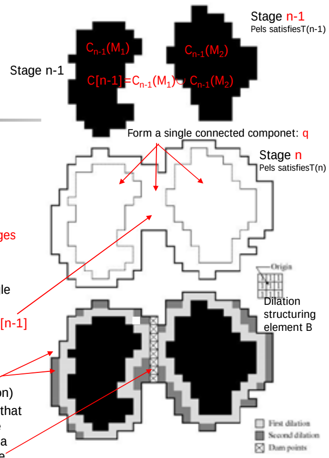
Watersheds

Short dam (single pixel width) is built to prevent water from merging



Watershed

- Work on Binary image:
  - Gradient on original image first
  - Define $T(n)=\{(s,t)|g(s,t)<n\}$
  - $T(n)$ to set regional minimal M_i, and its catchment basin $C_n(M_i)$. Let $C[n]=\cup C_n(M_i)$
  - Using $T(n-1)$, $T(n)$ to obtain two binary images at stages n and n-1, C[n-1] and C[n]
- Compare C[n-1] and C[n]
  - If two connected components form one single connected component q
  - A dam has to be built in some pixels of $q\cap C[n-1]$ using dilation on C[n-1]
- Dilation has to satifiy two conditions
  - The dilation has to be constrained to q (the center of B can be located in q during dilation)
  - The dilation cannot be performed on points that would cause the sets being dilated to merge (become a single connected component) or a dam (single pixel width) has to be built there.

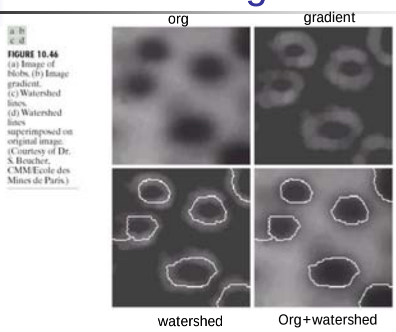


Watershed Segmentation

- $q\cap C[n-1]$ has three possibilities
  - Empty: new minimum (new area with min=n) is encounter
  - Contains one connected component of C[n-1]: q is incorporated into C[n-1] to form C[n] (q lies within catchment basin of some regional min)
  - Contains more than one connected component of C[n-1]: a one-pixel-thick dam can be built by dilating $q\cap C[n-1]$ within a 3x3 structuring element of 1's, and constraining the dilation to q

---

Watershed Segmentation



Watershed with marker

- Problem of Watershed Segmentaiton:
  - Over-segmentation due to noise and other local irregularities of the gradient

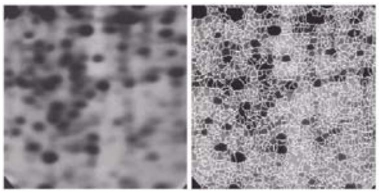


FIGURE 10.47 (a) Electrophoresis image. (b) Result of applying the watershed segmentation algorithm to the gradient image. Oversegmentation is evident. (Courtesy of Dr. S. Beucher, CMM/Ecole des Mines de Paris.)

Org   Over-segmentation

Watershed with marker

- Solution:
  - Preprocessing with a smoothing filter before gradient
  - Definition of a set of criteria that markers must satisfy
    - A marker is a connected component belonging to an image
    - Internal markers: Objects of interest
    - External markers: Background

Watershed with marker

Internal marker

External marker

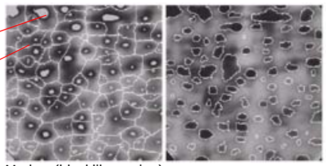


FIGURE 10.48 (a) Image showing internal markers (light gray regions) and external markers (watershed lines). (b) Result of segmentation. Note the improvement over Fig. 10.47(b). (Courtesy of Dr. S. Beucher, CMM/Ecole des Mines de Paris.)

- Internal Marker (blocklike region):
  - A region that is surrounded by points of higher "altitude"
  - Such that the points in the region form a connected component
  - In which all the points in the connected component have the same gray-level value
- External Marker
  - Perform watershed under the restriction that these internal markers be the only allowed regional minima
- Once Internal and External Markers (Region) are set
  - Perform any segmentation technique by partition each of these regions into two: a single object and its background
  - Ex: Operate on a single watershed for each region

The Use of Motion in Segmentation

Spatial Techniques - Difference

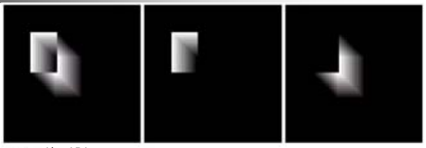
- Image frame f(x,y,t)
  - x,y,t: Integer for digitalized video
- Difference image $d_{ij}(x,y)$ (same size as f)
  - $$d_{ij}(x,y)=\begin{cases}1 & \text{if } |f(x,y,i)-f(x,y,j)|>T \\ 0 & \text{otherwise}\end{cases}$$
  - T: control the the bound of slight illumination variation
  - Isolated points
    - Caused by noise or slow movement of objects
    - Removed by forming 4- or 8-connected regions of 1 in d_{ij}, and ignoring that has less than a predetermined number of entries (size).
    - Removed by filtering or ADI (next slide)

Spatial Techniques - ADI

- Accumulative difference image (ADI)
  - Consider the changes at a pixel location over several frames
  - Compare the reference image R(x,y)=f(x,y,1) with every subsequent image in the sequence f(x,y2,),f(x,y,3),…,f(x,y,n)
- Three types of ADI (all counter start out with zero value)
  - Absolute ADI:
    - $$A_k(x,y)=\begin{cases}A_{k-1}(x,y)+1 & \text{if } |R(x,y)-f(x,y,k)|>T \\ A_{k-1}(x,y) & \text{otherwise}\end{cases}$$
  - Positive ADI:
    - $$P_k(x,y)=\begin{cases}P_{k-1}(x,y)+1 & \text{if } [R(x,y)-f(x,y,k)]>T \\ P_{k-1}(x,y) & \text{otherwise}\end{cases}$$
  - Negative ADI:
    - $$N_k(x,y)=\begin{cases}N_{k-1}(x,y)+1 & \text{if } [R(x,y)-f(x,y,k)]<-T \\ N_{k-1}(x,y) & \text{otherwise}\end{cases}$$

Spatial Techniques - ADI

Object is moving in a southeasterly direction at a speed of 5*sqrt(2) pixels per frame



a b c   Abs ADI   Pos ADI   Neg ADI

FIGURE 10.49 ADIs of a rectangular object moving in a southeasterly direction. (a) Absolute ADI. (b) Positive ADI. (c) Negative ADI.

- Abs ADI: Union of Pos and Neg
- Pos ADI: Equal to
  - The size of moving objects
  - The location of moving objects in reference frame
- Neg ADI: Determine
  - The direction of the moving objects
  - The speed of the moving objects

Spatial Techniques – Stationary Reference Frame

- Difference tends to cancel stationary components, leaving only noise or moving objects
- Build a reference from a set of images conaining one or more moving objects becomes necessary
- Procedure:
  - Monitor the change in postive ADI
    ➔ Determine the initial position of a moving object
    ➔ Object can be removed by subtraction
  - When the positive ADI stops changing, copy from this image the area previously occupied by the moving object in the initial frame